Humboldt-Universität zu Berlin – Geographisches Institut

SEISMIC VULNERABILITY ASSESSMENT OF BUILT ENVIRONMENTS WITH REMOTE SENSING

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

im Fach Geographie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät II

der Humboldt-Universität zu Berlin

von

Christian Geiß

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:

Prof. Dr. Elmar Kulke

Gutachter/in:

Prof. Dr. Tobia Lakes

Dr. Hannes Taubenböck

Prof. Dr. Günter Strunz

Tag der wissenschaftlichen Aussprache: 12. 12. 2014

SEISMIC VULNERABILITY ASSESSMENT OF BUILT ENVIRONMENTS WITH REMOTE SENSING

ACKNOWLEDGEMENTS

I want to express my deep gratitude to everyone who contributed to this work.

I would like to express my special gratitude to a few persons:

Hannes Taubenböck whose guidance, ideas, motivation and support were invaluable. Thank you for making this work possible. Tobia Lakes for very valuable advice and guiding me along the way, and Günter Strunz for being available as referee.

I want to thank all my colleagues at DLR for their support and letting me have a great time.

Thank you Michael Wurm for valuable discussions and important help during the years and Martin Klotz for the support especially during the final stage.

Special thanks to Patrick Aravena Pelizari and Marianne Jilge for your great motivation and work.

Marc Wieland and Massimiliano Pittore (GFZ Potsdam) for valuable discussions and Hideomi Gokon for the warm hospitality during my stays in Japan.

All my friends, especially Bastian, Jona (Munich), Jona (Berlin), Jan, Lukas, Thomas, and Michael. Thank you SoYeon for the help with proofreading the manuscript and beyond.

My parents Marion and Ottmar and my sister Vanessa for their loving support.

This work was supported by Helmholtz-EOS (Earth Observation System).

It was a lot of fun.

ABSTRACT

The impact of natural hazards such as earthquakes on mankind has increased dramatically over the last decades. Global urbanization processes and increasing spatial concentration of exposed elements such as people, buildings, infrastructure, and economic values in earthquake prone regions induce seismic risk at a uniquely high level. This situation, when left unmitigated, is expected to cause unprecedented death tolls, enormous economic and ecological losses, critical infrastructure and service failures, and poses a major threat for civil security, and a sustainable development in the future. To mitigate those perils requires detailed knowledge about seismic risks. As an important constituent element of seismic risk, the seismic vulnerability of the built environment has to be assessed. In particular, it is crucial to know about the behavior of the building inventory under seismic load. However, information about building inventories and affiliated seismic vulnerability is often outdated, unavailable, or simply not existent in many earthquake prone regions around the globe. Especially for analyses in development countries, which are often characterized by highly dynamic development and transformation processes related to urbanization, remote sensing is perceived as a promising tool for an economical up-to date, area-wide, and spatially consistent data collection in general. Numerous sensor systems with a high spatial resolution are nowadays available and enable a detailed detection, extraction, delineation, and characterization of urban morphology.

The main goal of the thesis was to develop and evaluate tailored methods and procedures that allow for a viable seismic assessment of the built environment with remote sensing data. In particular, methods from the machine leaning domain were adapted to estimate vulnerability levels of buildings and homogeneous urban structures based on features derived from remote sensing and by incorporation of *in situ* knowledge.

In this sense, we show for an earthquake prone test region in Indonesia (i.e., the city of Padang) how to derive a set of features for a comprehensive characterization of urban morphology based on multi-sensor remote sensing data. The data comprises multispectral IKONOS imagery, a normalized digital surface model derived from SAR measurements, and multi-temporal LANDSAT data. Features with significant explanatory power and the most suitable features for estimating seismic vulnerability levels of buildings are identified in combination with ancillary *in situ* information. Regression and classification models are learned that allow to estimate seismic vulnerability levels of buildings with viable accuracies. Subsequently, we introduce a hierarchical supervised classification approach to estimate seismic building structural types. Seismic building structural types characterize the main load-bearing structure of a building and, thus, reflect the behavior und seismic load. Models are learned under consideration of scarce *in situ* observations and are applied to the building inventory of Padang. To demonstrate the applicability for earthquake loss estimations, scenario-based loss estimations are presented based on empirical fragility functions for designated types.

Analyses on individual building level need to be based on remote sensing data with a very high spatial resolution. This hampers utilization capabilities for large areas due to data availability and costs as well as processing requirements. When aiming at spatially continuous and consistent assessment approaches that are applicable for large regions, on a national level, or even globally, those data represent a clear limitation. To lower this limitation, we alter the spatial scale of analysis and propose concepts and methods to assess homogeneous urban structures. Urban structures represent areas in urban environments with a homogeneous composition of its constituent elements (e.g., buildings). Thus, a coarser level of urban morphology is addressed when compared to building level. However, this allows relying on remote sensing data with a lower spatial resolution but larger spatial coverage. In particular, we use height information from the TanDEM-X mission in combination with multispectral RapidEye data. TanDEM-X represents a spaceborne radar interferometer, which delivers a global digital surface model with an unprecedented consistent spatial resolution of 0.4 arcseconds. To extract heights of elevated objects in urban environments (i.e., buildings), we propose jointly a novel region growing-based progressive morphological filter and a selective post-classification processing scheme. This is specifically done to cope with challenges that arise from the spatial resolution of the data and terrains with a high orographic energy. Based on the extracted information and multispectral RapidEye data, experiments are carried out in Istanbul (Turkey), which aim to assess the seismic vulnerability of urban structures under consideration of *in situ* information. Results suggest that a mixture of spectral and height related features provide most viable accuracies.

Overall, this thesis provides some promising results, which show that remote sensing has a high capability to contribute to a rapid screening assessment of the seismic vulnerability of buildings and urban structures. Further work can build upon these results and may challenge empirical findings in further case studies, enhance developed and applied methods, transfer concepts and approaches to other sensor systems/data sources, or apply data and methodologies within integrative and holistic risk assessment strategies.

Zusammenfassung

Die Auswirkungen von Naturgefahren wie Erdbeben haben in den vergangenen Dekaden dramatisch zugenommen. Globale Urbanisierungsprozesse und eine Zunahme der räumlichen Konzentration von exponierten Elementen wie Menschen, Gebäude, Infrastruktur und ökonomische Werte induzieren ein ungekanntes Risiko in erdbebengefährdeten Regionen. Wenn keine Abschwächung des Risikos erfolgt werden dramatische Folgen in der Zukunft erwartet. Diese umfassen eine beispiellose Anzahl an Todesopfer, enorme ökonomische und ökologische Verluste und Ausfälle bezüglich kritischer Infrastruktur und Versorgung. Um derartige Gefährdungen abzuschwächen sind detaillierte Informationen über seismisches Risiko notwendig. Die seismische Verwundbarkeit von Siedlungsarealen ist dabei als zentrale, konstituierende Komponente von seismischem Risiko zu berücksichtigen. In diesem Zusammenhang ist es von besonderem Interesse das Verhalten von Gebäudeinventaren unter einem bestimmten Erdbebeneinfluss abschätzen zu können.

Allerdings sind Informationen bezüglich Gebäudeinventaren und zugehöriger seismischer Verwundbarkeit in vielen erdbebengefährdeten Regionen häufig nicht aktuell, uneingeschränkt verfügbar oder existent. Besonders für Analysen in Entwicklungsländern, die gleichzeitig von hochdynamischen Entwicklungs- und Transformationsprozessen charakterisiert sind, wird Fernerkundung als vielversprechende Methode gesehen, die eine ökonomische, aktuelle und räumlich konsistente Datenerhebung erlaubt. Heutzutage sind zahlreiche Sensorsysteme mit einer hohen räumlichen Auflösung verfügbar die eine detaillierte Detektion, Extraktion, Abgrenzung und Charakterisierung der urbanen Morphologie erlauben.

Das Hauptziel der Arbeit war es maßgeschneiderte Methoden zu entwickeln die eine Bewertung der seismischen Vulnerabilität von Siedlungsräumen, basierend auf Fernerkundungsdaten, durchführbar machen. Im Besonderen wurden Methoden aus dem Bereich des maschinellen Lernens adaptiert, um Verwundbarkeitsstufen von Gebäuden und homogenen Siedlungsstrukturen zu bestimmen. Hierfür wurden Merkmale aus Fernerkundungsdaten abgeleitet und mit *in-situ*-Daten verknüpft.

In diesem Zusammenhang zeigen wir für eine erdbebengefährdete Testregion in Indonesien (Padang) wie eine Gruppe von Merkmalen aus komplementären Fernerkundungsdaten abgeleitet werden kann, die eine umfassende Charakterisierung der urbanen Morphologie erlaubt. Die Daten setzen sich aus Multispektralaufnahmen des Sensors IKONOS, einem normalisierten Oberflächenmodell aus Radarmessungen und multitemporalen Landsat-Aufnahmen zusammen. Merkmale mit einem signifikanten Erklärungsgehalt und die geeignetsten Merkmale in Beug auf eine Schätzung des Vulnerabilitätsniveaus der Gebäude werden unter Berücksichtigung von *in-situ*-Informationen identifiziert. Wir bilden Regressions- und Klassifikationsmodelle die eine Schätzung des seismischen Vulnerabilitätsniveaus von Gebäuden mit hohen Genauigkeiten erlauben.

Nachfolgend stellen wir einen hierarchischen, überwachten Klassifikationsansatz vor, der es erlaubt seismische Gebäudestrukturtypen zu identifizieren. Seismische Gebäudestrukturtypen charakterisieren den Konstruktionstyp eines Gebäudes und reflektieren das Verhalten unter Erdbebeneinfluss. Die Modelle werden unter Berücksichtigung von *in-situ*-Informationen gebildet und auf das Gebäudeinventar von Padang angewendet. Um die Relevanz des Ansatzes in Bezug auf erdbebenbezogene Schadensschätzungen darzustellen werden szenario-basierte Schadenschätzungen präsentiert. Dabei werden Fragilitätsfunktionen für verschiedene Typen verwendet, die empirisch auf Basis eines Erdbebenereignisses am 30. September 2009 in Padang abgeleitet wurden.

Grundsätzlich benötigen Analysen für Einzelgebäude die Verwendung von Fernerkundungsdaten mit einer sehr hohen räumlichen Auflösung. Dies erschwert gegenwärtig die Nutzungsmöglichkeiten für große Siedlungsflächen auf Grund der Datenverfügbarkeit und zugehörigem Prozessierungsaufwand. Wenn man auf räumlich kontinuierliche und konsistente Bewertungsansätze abzielt, die für große Regionen, auf einem nationalen Level oder sogar global anwendbar sind, bedingen eben genannte Daten deutliche Limitierungen.

Um diese Limitierungen abzuschwächen verändern wir die räumliche Analyseskala und schlagen Konzepte und Methoden vor um homogene urbane Strukturen zu bewerten. Urbane Strukturen repräsentieren Siedlungsflächen die von einer homogenen Zusammensetzung ihrer konstituierenden Elemente (z.B. Gebäude) gekennzeichnet sind. Entsprechend wird ein gröberer Level der urbanen Morphologie, im Vergleich zu Einzelgebäuden, adressiert. Allerdings erlaubt dies die Einbindung von Fernerkundungsdaten mit einer geringeren räumlichen Auflösung und größerer räumlicher Abdeckung. Wir verwenden Höheninformationen der TanDEM-X Mission in Verbindung mit multispektralen RapidEye-Daten. TanDEM-X ist ein satellitengestützter Radar-Interferometer, der ein globales digitales Oberflächenmodell mit ungekannter räumlicher Bodenauflösung von 0.4 Bogensekunden liefert. Um aus diesen Daten Objekte in Siedlungsräumen zu extrahieren die von der Erdoberfläche erhaben sind (Gebäude), schlagen wir ein neues Verfahren („region growing-based progressive morphological filter“) in Verbindung mit einem selektiven „Post-Klassifikationsschema“ vor. Dieses Vorgehen ist Herausforderungen geschuldet, die sich aus der räumlichen Auflösung der TanDEM-X Daten und Terrain mit einer hohen orographischen Energie ergeben.

Basierend auf dieser Information und Multispektraldaten des RapidEye-Systems werden Experimente in Istanbul (Türkei) vorgenommen, die auf die seismische Vulnerabilitätsbewertung von urbanen Strukturen abzielen. Analog zu bisherigen Vorgehensweisen werden aus Fernerkundungsdaten Merkmale extrahiert und mit *in-situ*-Informationen verknüpft. Die Ergebnisse weißen darauf hin, dass eine kombinierte Nutzung von Spektral- und Höheninformation die besten Modellschätzungen ermöglichen.

Insgesamt kann diese Arbeit einige vielversprechende Ergebnisse präsentieren die zeigen, dass Fernerkundung ein hohes Potenzial hat, um zu einer schnellen Vulnerabilitätsbewertung von Gebäuden und urbanen Strukturen beizutragen. Zukünftige Arbeiten können daran anknüpfen und beispielsweise empirische Erkenntnisse in weiteren Fallstudien anzweifeln, eine Verbesserung der Methodik vornehmen, Konzepte und Ansätze auf andere

Sensorsysteme oder Datenquellen übertragen oder Daten und Methoden im Rahmen von holistischen Risikobewertungsstrategien anwenden.

CONTENTS

CHAPTER I: INTRODUCTION

CHAPTER II: REMOTE SENSING CONTRIBUTING TO ASSESS EARTHQUAKE RISK: FROM A LITERATURE REVIEW TOWARDS A ROADMAP

CHAPTER III: ASSESSMENT OF SEISMIC BUILDING VULNERABILITY FROM SPACE

CHAPTER IV: ESTIMATION OF SEISMIC BUILDING STRUCTURAL TYPES USING MULTI-SENSOR REMOTE SENSING AND MACHINE LEARNING TECHNIQUES

LIST OF FIGURES

LIST OF TABLES

Chapter I

Introduction

1. TRANSFORMATIONS OF HUMAN HABITATS AND EARTHQUAKE RISK

Fundamental transformations of human habitats impose new challenges with regard to natural hazard risks. It is estimated that more people live in urban environments than in rural areas since the year 2008 (UN, 2011). This migration process, from rural to urban areas, is expected to continue, and current estimates foresee a share of two-thirds of the world population living in cities in 2050. Generally, the world population is expected to increase from little more than ~7 billion nowadays to ~9.3 billion in 2050. During the same time period, the absolute number of people living in urban environments is expected to increase from ~3.6 billion to ~6.3 billion. Thus, urban environments worldwide are expected to absorb completely the projected population growth and also draw in some rural population (UN, 2012).

Besides these prospective numbers, the past 40 years were subject to a witnessed transformation. In 1975 ~38% of the global population were urban residents, in 1990 already ~43%, and nowadays little more than 50% live in urban environments (Ibid.). These figures illustrate a global urbanization process with unprecedented magnitude and dynamic. The increase of urban population has occurred during a time period that is comparatively short with respect to the return time of severe earthquakes. Earthquakes that had little impact in the past, when they hit sparsely populated and spatially fragmented settlement areas, are nowadays expected to shake urban agglomerations with millions of people (Bilham, 2009). Moreover, rapid urban growth, which occurs especially in developing countries, is often accompanied by the construction of unplanned and highly vulnerable settlements. These frequently change over short time scales with respect to location and extent (Wieland et al., 2012).

These described dynamic global urbanization processes, which are accompanied by increasing spatial concentration of exposed elements such as people, buildings, infrastructure, and economic values in earthquake prone regions, induce seismic risk at a uniquely high level. When left unmitigated, this situation is expected to cause unprecedented death tolls, enormous economic and ecological losses, and poses a considerable threat for civil security, and a sustainable development in the future (Bilham, 2009; Holzer and Savage, 2013; Tucker, 2013). To mitigate those perils requires detailed knowledge about seismic risks. As an important constituent element of seismic risk, the seismic vulnerability of the built environment has to be assessed. In particular, it is crucial to have information about the building inventory and its behavior with respect to a certain level of ground shaking (Erdik et al., 2003).

Especially for earthquake loss estimation (ELE) modeling, the gathering of building inventory and vulnerability information represents normally the most time-consuming and expensive aspect (Dunbar et al., 2003). Conventional approaches to assess the seismic vulnerability of the building inventory, which incorporate an exclusive application of detailed *in situ* building-by-building analysis by structural engineers, may provide very detailed and high quality information, but are decreasingly able to cope with the high spatiotemporal dynamics of urban environments. On the contrary, information collected on a very broad spatial level such as spatially aggregated census data hampers the consideration of small-scale hazard effects in a downstream risk model (Wieland et al., 2012). Hence, building inventory data and affiliated seismic vulnerability information is often outdated, spatially aggregated and discontinuous, and in many earthquake prone regions of the world simply not existent.

2. SEISMIC VULNERABILITY ASSESSMENT AND REMOTE SENSING

The assessment and monitoring of the physical seismic vulnerability of built environments is a challenging task, especially when spatially consistent and large-area evaluations are required. In this sense, numerous studies emphasize that remote sensing can play a valuable role in supporting the extraction of relevant features for pre-event vulnerability analysis of urban environments (French and Muthukumar, 2006; Mueller et al., 2006; Sarabandi et al., 2008; Taubenböck et al., 2009a; Borfecchia et al., 2010; Sahar et al. 2010; Borzi et al. 2011; Deichmann et al., 2011; Wieland et al., 2012; Pittore and Wieland, 2013; and GEM, 2013). The intrinsic advantage of remote sensing is the ability to offer an overview of building stocks and serve as a screening method for derivation of seismic vulnerability related features, such as shape characteristics, height, roof material, year of construction, structural type and spatial context. This topic is subject to a lively research and has gained much scientific contemplation in the past few years. Thereby, the research environment is constituted primarily by two distinctive science communities: remote sensing and earthquake engineering. Nowadays, concepts and methods are reaching a level, where they are found to be relevant and being accepted in both communities. This is evident, since the following studies are published in established journals of both communities.

Taubenböck et al. (2009), and Borzi et al. (2011) characterize the built environment by means of remote sensing data and retrieve specific fragility functions for designated building types. Borfecchia et al. (2009) assess the vulnerability of buildings in a hybrid way, by combining *in situ* ground truth for selected buildings with information derived from remote

sensing data. Supervised classification techniques are subsequently applied to classify the residual building inventory. Wieland et al. (2012), and Pittore and Wieland (2013) use multitemporal Landsat data to discriminate homogeneous urban structures based on an image segmentation approach and semantically annotate them by utilizing a supervised classification scheme. Derived urban structures are subject to a more detailed analysis of the building stock with VHR optical data and a ground-based omnidirectional imaging system. The sensed information is combined with ancillary information (i.e., information from the world housing encyclopedia) for a subsequent probabilistic seismic vulnerability assessment.

The cited studies provide promising empirical evidence with respect to the viability of the approaches. However, a number of points related to conceptual, methodological, and application-oriented considerations remain unaddressed, and give rise to a further scientific contemplation. In this sense, the subsequent section outlines the particular research objectives of this thesis. Generally, please note that in-depth distinctions to previous research can be found in the introduction of chapters III-VI.

3. RESEARCH OBJECTIVES AND STRUCTURE OF THE THESIS

As stated, recent studies introduced approaches to assess built environments under consideration of remote sensing data and presented interesting empirical evidence with respect to viability. However, the affiliated scientific canon regarding deployed concepts, developed methods, and used data is not comprehensively elaborated. In addition, empirical evidence on the viable application of remote sensing for seismic vulnerability assessment of built environments is still scarce. In this sense, this thesis is governed by the overarching goal to develop, apply, and evaluate tailored methods and procedures that allow for a seismic vulnerability assessment of the built environment based on remote sensing data.

To this purpose, we first review a comprehensive body of scientific literature to give an overview on current concepts of risk, hazard, and vulnerability and contributing remote sensing related methods and applications in chapter II. From a qualitative and quantitative analysis of the literature, we identify research trends of the past years and uncover unaddressed perspectives and research gaps. Conclusions and direction for future research are drawn from a scientific, technical, multi- and transdisciplinary, and political perspective.

Chapter III reports outcomes of an empirical case study carried out in the city of Padang (Indonesia). It aims is to answer the questions: (i) Which features can be derived from satellite remote sensing data that best explain seismic building vulnerability?; (ii) How suitable are

features derived from satellite remote sensing data for estimating seismic building vulnerability levels? These questions are addressed under consideration of *in situ* information, which characterizes the seismic vulnerability levels of buildings according to two engineering based assessment methods (i.e., a scoring method and a categorization according to the European Macroseismic Scale 98). We deploy very high resolution multispectral data (i.e., IKONOS), height information from a normalized digital surface model, and spatiotemporal analyses generated from LANDSAT data, and regression analysis and machine learning algorithms to address the raised research questions.

Chapter IV complements previous analyses carried out in Padang by focusing on application-oriented perspectives with respect to earthquake loss estimation. Therefore, we propose a tailored hierarchical classification procedure to estimate seismic building structural types. We combine this information with available fragility functions for designated building types, which were derived empirically after the last severe earthquake in Padang on 30th September 2009. Scenario-based earthquake loss estimations are given to demonstrate the relevance of the approach.

Up to this point of the thesis, the assessment of individual buildings is addressed. However, this induces the deployment of very high spatial resolution remote sensing data, which hampers today's utilization capabilities for larger areas due to data costs and processing requirements. When aiming at spatially continuous and consistent assessment approaches, which are applicable for large regions, on a national level, or even globally, those data represent a clear limitation. This is why we alter the spatial level of analysis and subsequently propose concepts and methods to assess homogeneous urban structures. Urban structures represent areas in urban environments with a homogeneous composition of its constituent elements. Thus, a coarser level of urban morphology is addressed when compared to building level. However, this allows relying on remote sensing data with a lower spatial resolution but larger spatial coverage. In particular, we exploit multispectral data from the RapidEye constellation and elevation information from the TanDEM-X mission.

To be able to infer valid height information of built environments in the subsequent analyses, we first design an algorithm for a viable normalization of TanDEM-X digital surface model data in challenging, non-flat terrain (chapter V).

Chapter VI documents a procedure which comprises four main steps dedicated to: i) delineation of urban structures by means of an original and tailored unsupervised data segmentation procedure with scale optimization; ii) characterization of urban structures by a joint exploitation of multi-sensor data; iii) selection of most feasible features under

consideration of *in situ* vulnerability information; iv) estimation of seismic vulnerability levels of urban structures within a supervised learning framework. Experimental results are reported for the mega city Istanbul (Turkey), which faces an enormous seismic threat. The last chapter (VII) is dedicated to synthesizing outcomes of the five preceding chapters. Moreover, ideas and directions for future research are drawn.

Chapters II-VI were written as standalone manuscripts to be published in international journals with peer-review. The different manuscripts contain sections related to introductory perspectives, study area and data, methods, results, discussions, and conclusions. This determines recurrence of a limited amount of material through these sections and thesis. Chapters II-VI were submitted or published in the time period from November 2011 until September 2014.

Chapter II

Geiß C, Taubenböck H (2013) Remote sensing contributing to assess earthquake risk: from a literature review towards a roadmap. *Natural Hazards*, 68, 7-48.

Chapter III

Geiß C, Taubenböck H, Tyagunov S, Tisch A, Post J, Lakes T (2014) Assessment of seismic building vulnerability from space. *Earthquake Spectra*, 30(4), 1553-1584.

Chapter IV

Geiß C, Aravena Pelizari P, Marconcini M, Sengara W, Edwards M, Lakes T, Taubenböck T () Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, in press.

Chapter V

Geiß C, Wurm M, Breunig M, Felbier A, Taubenböck H () Normalization of TanDEM-X DSM data in urban environments with morphological filters. Submitted to *IEEE Transactions on Geoscience and Remote Sensing*, minor revisions.

Chapter VI

Geiß C, Jilge M, Lakes T, Taubenböck H () Seismic vulnerability assessment of urban structures with multi-sensor remote sensing. Submitted to *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*.

Authors' contributions to the individual chapters

Chapter II

Christian Geiß and Hannes Taubenböck elaborated the concept of the paper. Christian Geiß conducted the literature review and wrote vast parts of the manuscript. Especially section 5 of the manuscript was written by both authors.

Chapter III

Christian Geiß elaborated the concept of the paper, conducted vast parts of analyses, and wrote most parts of the manuscript. Hannes Taubenböck contributed to conceptual points, conducted preliminary work by compiling data *in situ* and providing derived information, and revised the manuscript. Sergey Tyagunov developed the categorization of the *in situ* data according to the European Macroseismic Scale 98 and wrote the affiliated section in the manuscript. Anita Tisch contributed to regression analyses. Joachim Post revised the manuscript. Tobia Lakes contributed on a conceptual level and revised the manuscript in great detail.

Chapter IV

Christian Geiß elaborated the concept of the paper, conducted vast parts of analyses, and wrote vast parts of the manuscript. Patrick Aravena Pelizari conducted vast parts of analyses jointly with Christian Geiß and helped to write the manuscript. Mattia Marconcini helped to design methods and procedures and revised the manuscript in great detail. Wayan Sengara and Mark Edwards collected *in situ* data after the earthquake in Padang on 30th September 2009. Tobia Lakes contributed on a conceptual level and revised the manuscript in great detail. Hannes Taubenböck contributed on a conceptual level, provided data, and revised the manuscript in great detail.

Chapter V

Christian Geiß elaborated the concept of the paper, designed and implemented algorithms, conducted vast parts of analyses, and wrote vast pasts of the manuscript. Michael Wurm helped to design experiments, depict results, and revise the manuscript. Markus Breunig contributed to the computation and provision of intermediate TanDEM-X DSM data and wrote the affiliated section in the manuscript. Andreas Felbier computed and provided the "urban footprint" data and wrote the affiliated section in the manuscript. Hannes Taubenböck contributed on a conceptual level and revised the manuscript in great detail.

Chapter VI

Christian Geiß elaborated the concept of the paper, designed and implemented methods, and wrote the manuscript. Marianne Jilge carried out a vast part of data processing. Tobia Lakes and Hannes Taubenböck contributed on a conceptual level and revised the manuscript in great detail.

Chapter II

Remote sensing contributing to assess earthquake risk: from a literature review towards a roadmap

Natural Hazards, 2013, 68, 7-48

Christian Geiß and Hannes Taubenböck

doi: 10.1007/s11069-012-0322-2
Received 3rd November 2011; accepted 27th July 2012

ABSTRACT

Remote sensing data and methods are widely deployed in order to contribute to the assessment of numerous components of earthquake risk. While for earthquake hazard related investigations, the use of remotely sensed data is an established methodological element with a long research tradition, earthquake vulnerability–centred assessments incorporating remote sensing data are increasing primarily in recent years. This goes along with a changing perspective of the scientific community which considers the assessment of vulnerability and its constituent elements as a pivotal part of a comprehensive risk analysis. Thereby, the availability of new sensors systems enables an appreciable share of remote sensing first. In this manner, a survey of the interdisciplinary conceptual literature dealing with the scientific perception of risk, hazard and vulnerability reveals the demand for a comprehensive description of earthquake hazards as well as an assessment of the present and future conditions of the elements exposed. A review of earthquake-related remote sensing literature, realized both in a qualitative and quantitative manner, shows the already existing and published manifold capabilities of remote sensing contributing to assess earthquake risk. These include earthquake hazard-related analysis such as detection and measurement of lineaments and surface deformations in pre- and post-event applications. Furthermore, pre-event seismic vulnerability–centred assessment of the built and natural environment and damage assessments for post-event applications are presented. Based on the review and the discussion of scientific trends and current research projects, first steps towards a roadmap for remote sensing are drawn, explicitly taking scientific, technical, multi- and transdisciplinary as well as political perspectives into account, which is intended to open possible future research activities.

1. INTRODUCTION

According to official estimates, the magnitude 7.0 earthquake in Haiti in 2010 killed 316,000 people, 300,000 were injured and 1.3 million people were displaced, 97,294 houses were destroyed and further 188,383 damaged (USGS, 2011). Other prominent examples of the last decennium are the earthquakes in Eastern Sichuan in China in 2008 (87,587 casualties), in Pakistan in 2005 (86,000 casualties) or in Bam in Iran in 2003 (30,000 casualties). These immense figures emphasize the devastating impact of earthquakes, especially in countries that are not prepared for such an event. In 2008, for example, the NatCatSERVICE of the Munich Re's database on natural hazards documented 750 loss events, with only 12 % from earthquakes (Munich Re Group, 2011). However, in the same year, 43 % of all fatalities and 43 % overall economic losses around the globe were caused by earthquakes.

In general, urban populations are expected to be killed by earthquakes in the foreseeable future in greater numbers than in the documented past (Bilham, 2009). Many hot spots for devastating future earthquakes are known and well documented (Dilley et al., 2005). For instance the mega city Istanbul, Turkey, with its estimated 15 million inhabitants is threatened by a 30–70 % probability of a major earthquake ($M_w > 7$), in the next 30 years (Parsons, 2004).

However, ''the international community's response to disasters has been mostly reactive, with only limited budget invested in prevention. (…) Even if there were a willingness to invest in prevention, the question would be: where?'' (Peduzzi, 2006: 171).

Prior to this question, the concept of risk itself is subject to a vibrant debate within the scientific community, which constitutes an alteration of the perception of risk. While hazard-oriented research strategies dominated the past (Lewis, 1999; Zhang et al. 2002), more integrative approaches to assess risk and its components, incorporating also human, societal and cultural factors, are in the focus of the scientific community nowadays (Turner et al., 2003; Pelling, 2003; Cardona, 2004; Adger, 2006; Birkmann, 2006b; Thywissen, 2006; Mercer et al., 2007; Fuchs, 2009). This shift is based on the realization that natural hazards do not have an intrinsic dangerous character itself (Cannon, 1994; Wisner, 2007), but become disastrous if an unfavourable combination of several parameters comes together. In this regard, natural disasters are perceived as ''un-natural'' (O'Keefe et al., 1976; WB/UN, 2010), not solely triggered by natural events but also connected to social, economic, ecological, social and political aspects (Burton et al., 1993; Blaikie et al., 1994; Alexander, 2002; Schneiderbauer and Ehrlich, 2004; Smith and Petley, 2009). Therefore, disasters can be viewed as a consequence of a complex reciprocity between potentially damaging physical

events such as earthquakes, tsunamis, droughts, floods, storms, etc. and the vulnerability of the built and natural environment, society and economy, which are preconditioned by human behaviour (Birkmann, 2006b) and often constitute and manifest development problems (UN, 1994; Yodmani, 2001; Alcantara-Ayala, 2002; Guinau et al., 2005; Schneiderbauer and Ehrlich, 2006).

To consider these framing conditions for disaster management and develop strategies to reduce disaster risk, several integrative research concepts have been designed and postulated (e.g., Cutter, 1996; Turner et al., 2003; Wisner et al., 2003; Bogardi and Birkmann, 2004). Still, there is a lack within the scientific community concerning common ontologies and definitions (Hufschmidt, 2011) especially about terms like ‘‘risk’’ and ‘‘vulnerability’’ (Timmermann, 1981; Cutter, 2003; Thywissen, 2006). This is mainly due to an inflationary usage of those terms within several research contexts and scientific disciplines (Cutter, 1996) and the need to work in the context of different social and environmental conditions (Hufschmidt, 2011).

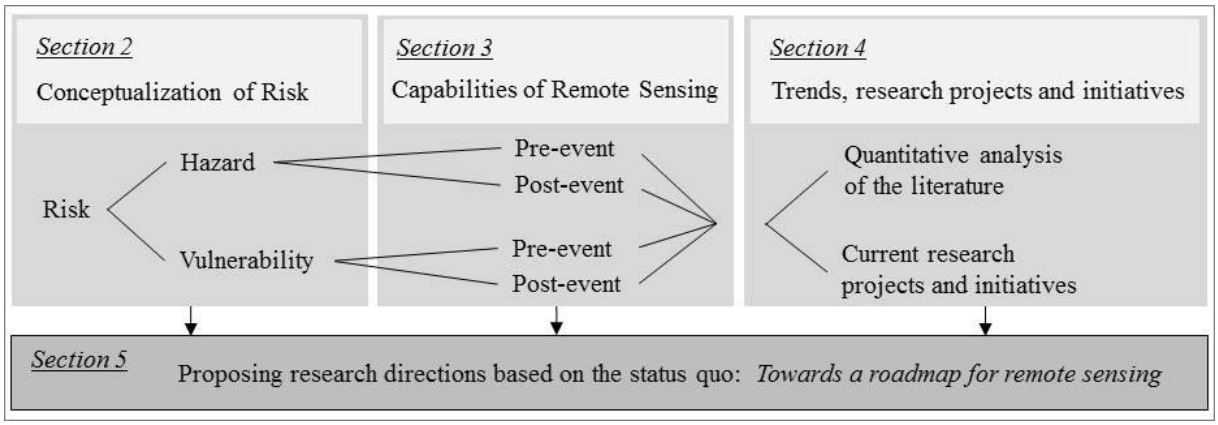


Fig. II-1. Structure and content of the paper.

In order to be able to categorize and discuss the role of remote sensing against this framing background of risk, this article gives a multidisciplinary overview of concepts and definitions and the associated theoretical assumptions and implications first. Despite this divergence about basic terminological questions, there is an increasing convergence about the need for robust and reliable indicators and methods to identify and assess risk and its components and utilize that information to reduce disaster risk (Kasperson et al., 2005; Villagran De Leon, 2006; Birkmann, 2006b; Peduzzi et al., 2009; Heltberg et al., 2009). It has been stated numerous times that one of the most difficult issues in assessing risk is the gathering of appropriate data (Birkmann 2006a; Ehrlich and Zeug, 2008; Ehrlich et al., 2010). In this manner, remote sensing is perceived as promising tool for an economical, up-to-date and area-

wide information collection in general (Dech, 1997; Paylor et al., 2005; Mueller et al., 2006; French and Muthukumar, 2006; Chiroiu et al., 2006; Esch et al., 2009; Joyce et al., 2009a; Guo, 2010). In particular, previous studies demonstrated that remote sensing is able to tackle certain aspects of earthquake risk. These reach from hazard-related analyses (e.g., Fu et al., 2004; Stramondo et al., 2005; Philip, 2010) to vulnerability-centred assessments (e.g., Taubenböck et al., 2008, 2009a; Ehrlich et al., 2010; Deichmann et al., 2011). Thus, the second research target of this article is to identify and discuss these potentials and at the same time show the limitations of remote sensing contributing to assess earthquake risk.

From a temporal perspective, remote sensing contributions to assess earthquake risk have a long research tradition. Remote sensing imagery is already used for hazard-related applications since the advent of research-oriented satellite systems and sensors four decades ago (Tronin, 2010). However, especially for vulnerability-related analysis, remote sensing is a less-established methodological element (Nassel and Voigt, 2006) and is perceived increasingly only in recent years as a valuable source of information (Deichmann et al., 2011). In order to disclose such developments and identify general trends, a quantitative analysis of the literature is performed reaching from 2011 back to 1991, which is completed by presenting actual research projects and initiatives in section 4. Finally, based on the previous chapters as well as on experiences in projects, advisory boards, at conferences and interviews with experts, a synthesis is given in section 5, which addresses explicitly scientific, technical, multi- and transdisciplinary and political perspectives in order to propose possible future research. The structure of this paper is visualized in Fig. II-1, which shows the aggregated content of the respective sections.

2. CONCEPTUALIZATION OF RISK

From a very general perspective, risk can be described as probability to suffer loss, damage and negative consequences (Burby, 1991; Brooks, 2003), referred to a present or specified future time period (Lafond and Gosselin, 1994; Coburn et al., 1994). Projected to the context of disaster and risk management, the probability of harmful consequences or expected losses is explicitly linked to the influence of a particular hazard for a given area (Downing et al., 2001) to a given element at danger or peril (Schneiderbauer and Ehrlich, 2006). However, risk cannot solely be characterized as a function of a hazard, describing the possibility of physical harm, since elements exposed to a certain hazard can react differently in order of their degree of vulnerability, or resilience, respectively, if understood as antonym of vulnerability (Adger

et al., 2005), and therefore modify the realization of risk (Cannon, 1993; Chapman 2001; Schneiderbauer and Ehrlich, 2004; Samuels et al., 2009). In this manner, the risk for a particular system (e.g. city) can be described on the basis of two distinctive factors: (1) A potentially damaging event, phenomenon or human activity, which is constituted by likelihood of occurrence, intensity, frequency and location and (2) the vulnerability, which characterizes the degree of susceptibility of the elements exposed to that particular source and therefore manifest the relationship of the degree of exposure and the degree of damage (UN/ISDR, 2004).

This conceptual superstructure of risk is very prevalent within the scientific community (Wisner et al., 2003; Birkmann, 2007), deployed in various theoretical and conceptual approaches and applications (e.g., Blaikie et al., 1994; Garatwa and Bollin, 2002; Bollin, 2003; Rashed and Weeks, 2003; Sarewitz et al., 2003; UN/ISDR, 2004; Taubenböck et al., 2008; Müller et al. 2011). Birkmann (2006b, 2007) reveals that the hazard event is primarily perceived as an external factor (see also van Dillen, 2004), while the term vulnerability describes the intrinsic characteristics of a system (Bohle, 2001). However, there are several definitions that adapt this separation of hazard and vulnerability, but divide the vulnerability part in further subcomponents (White et al., 2005; Villagran De Leon, 2006). A definition that is exhaustively used by the earthquake disaster risk community denotes elements at risk, which are understood as objects potentially adversely affected such as people, properties, infrastructure or economic activities, as autonomous component (UNDRO, 1979; Crichton 1999; Granger, 2003; Masure, 2003; Schneiderbauer and Ehrlich, 2004; Peduzzi et al., 2005; Sinadinovski et al., 2005; Dilley et al., 2005; UNDP/ERRRP, 2009; Ehrlich et al., 2010; Deichmann et al., 2011). Thereby, the calculation of potential losses within frameworks of earthquake risk models such as HAZUS (FEMA, 2010), OpenQuake (GEM, 2011) or RiskScape (RiskScape, 2012) is carried out by combining determined hazard parameters, quantified and characterized exposed elements and their assessed vulnerability. Beyond, there are definitions that put vulnerability in a less broader context and therefore explicitly address within the risk definition components like, for example, resilience (Thywissen, 2006), deficiencies in preparedness (Villagran De Leon, 2006) or coping capacities (Davidson, 1997; Hahn, 2003) separately.

In contrast, especially definitions and concepts evolved from global environmental and climate change research do not follow the separation of hazard and vulnerability (Cutter, 1996; Füssel and Klein, 2006), but take the hazard and the exposure to it into account by considering anthropogenic actions, which indeed are effecting the environment and therefore

some of the hazards (Villagran De Leon, 2006). However, we maintain the conceptual superstructure of the UN/ISDR (2004) definition, since we focus on earthquake hazards, which are beyond the influence of human kind (Bollin, 2003) and discuss different perspectives on risk components at the respective passage. For overviews of risk definitions with respect to different scientific disciplines, the reader is referred to, for example, Brooks (2003), Schneiderbauer and Ehrlich (2004) and Thywissen (2006). Carreno et al. (2007) provide an introductory chronological description of the development of epistemological terms and concepts in the context of earthquake risk. In conclusion, according to Deichmann et al. (2011), it should be noted that all components of natural disaster risk show an inherent variability over space, and therefore, risk identification and assessment must rely on data and information with spatial reference.

2.1. HAZARD

Based on a wider understanding, hazards represent potentially damaging events, phenomenon or human activities, which may have negative consequences on human aspects (loss of life or injury), elements of the built environment (property damage) and environmental components. Hazards can be single, sequential or combined, both in their origin and effects (UN/ISDR, 2004). For example, earthquakes may trigger other earthquakes (e.g., Pinar et al., 2001; Erdik et al., 2004; McCloskey et al., 2005; Marsan and Lengline, 2008) and/or secondary effects such as landslides (e.g., Bommer and Rodriguez, 2002; Huang and Li, 2009), tsunamis (e.g., Sahal et al., 2009; Taubenböck et al., 2009b; Rosenau et al., 2010; Strunz et al., 2011) or fires (e.g., Girgin, 2011), also in a cascading way (Schneiderbauer and Ehrlich, 2004), which can even be more disastrous than the initial hazard (Korup, 2010).

Brooks (2003) notes that in certain definitions, there is some ambiguity as to whether hazards represent a trigger event or the outcome of such events. Schneiderbauer and Ehrlich (2004) state that especially in the disaster-related literature there are divergent views whether hazards should only describe naturally induced events or also include events that are triggered by human activities. In this context, Garatwa and Bollin (2002) make a distinction between real natural hazards such as earthquakes, volcanic eruptions and storms and socio-natural hazards such as floods, droughts and forest fires, which are also caused or exacerbated due to human intervention in nature. Analogously, Smith and Petley (2009) show the spectrum of environmental hazards reaching from geophysical events to human activities, where hazards with a high level of human causation such as air pollution, industrial accidents or bushfires are more voluntary in terms of their acceptance and more diffuse in terms of their disaster

impact in contrast to highly involuntary hazards such as earthquakes, tsunamis, etc. As already mentioned above, Schneiderbauer and Ehrlich (2006) emphasize that hazards often have interrelated causes, and therefore, the allocation of a hazard to one class may be difficult. However, if describing certain hazard-related risk components that are at the same time most likely hazard-specific (Brooks, 2003), it is crucial to determine basic characteristics such as magnitude (only events are considered as extreme, when some common level is exceeded), duration (persistence of a hazardous event), speed of onset (time between first occurrence of an event and its peak), temporal spacing (sequencing of events ranging from random to periodic), spatial extent (space covered by a hazardous event) and spatial dispersion (pattern of distribution over the space the impact can occur; Gravely, 2001, quoted in Schneiderbauer and Ehrlich, 2004).

In this regard, earthquake hazards are typically characterized based upon the basic concepts of probabilistic seismic hazard assessment (Cornell, 1968) and described by probabilities of occurrence of earthquakes with a specified amplitude of interest for a given period of time (Robinson et al., 2006; UNDP/ERRRP, 2009; see Panza et al., 2011 for a critical discussion of the shortcomings of this method). In order to describe potential earthquake hazards, Sinadinovski et al. (2005) distinguish between regional seismic models (description of the chance of an earthquake of a given magnitude occurring in a specified time period in various parts of the region), attenuation models (general description how earthquake ground shaking or intensity decreases with distance away from the earthquake source) and site response models (description how local soil conditions will affect the ground shaking experienced during an earthquake). In detail, a pre-event earthquake parameter evaluation may include earthquake intensity, peak velocity, predominant period, potential earthquake source, liquefaction potential, etc. and also take the analyses of potential secondary threats into account (Tang and Wen, 2009) based on empirical, analytical or hybrid methods (e.g., Calvi et al., 2006; Crowley and Bommer, 2006).

2.2. VULNERABILITY

An international widespread definition of vulnerability in the context of risk research describes vulnerability as ‘‘the conditions determined by physical, social, economic and environmental factors or processes, which increase the susceptibility of a community to the impact of a hazard’’ (UN/ISDR, 2004). Although this definition provides little information about the specific factors or processes that constitute vulnerability, it is clear that vulnerability characterizes the combined conditions and situations of various elements of interest, which

make them prone being affected from the influences of an hazardous event (Bogardi et al., 2005; Birkmann et al., 2011). Vulnerability has to be understood as a predictive concept, since vulnerability is a present attribute that describes and characterizes possible future harm (Wolf, 2011). Certainly, it should also be noted that within vulnerability research different schools of thought exist with partially significant differences regarding definitions and associated concepts (Birkmann et al., 2011; see also Birkmann, 2006b; Füssel and Klein, 2006; Thywissen, 2006), since vulnerability is perceived basically as something abstract without a simple, clear notion (Hilhorst and Bankoff, 2004). However, Wolf (2011) argues that especially from a methodological point of view, which incorporates certain strategies to assess vulnerability, these differences are less conceptual but primarily terminological.

The *concept of vulnerability* emerged within the social sciences as a reaction to a solely hazard-oriented reception of disasters and risk in the early 1970s (Schneiderbauer and Ehrlich, 2004; Villagran De Leon, 2006). But already Timmermann (1981) stressed the heterogenic definitions of that term, which made it almost useless for a careful description. Bogardi and Birkmann (2004) note that vulnerability does not represent a clear scientific concept, what causes the paradox that the scientific community tries to measure vulnerability, but cannot define it precisely yet (Birkmann, 2006b). Rashed and Weeks (2003) state that assessing vulnerability and risk towards natural hazards can be regarded as an illstructured problem, without a unique, clearly identifiable and objectively optimal solution. Therefore, it is no surprise that within the actual scientific literature numerous definitions and pre-analytic frameworks have emerged, assigning a definition with respect to a specific epistemological view of several research disciplines (Wisner 2004; Brooks 2003) and enhance it with individual connotations in order to make the term operable for different research question. An overview of several definitions can be studied, for example, in Schneiderbauer and Ehrlich (2004), Green (2004), Thywissen (2006), Villagran De Leon (2006), Adger (2006), Fuchs (2009) and Cutter et al. (2009). Kasperson et al. (2005) and Füssel (2007) provide an overview of the evolution of concepts and approaches to vulnerability research, respectively.

Within the *social sciences*, vulnerability refers to the susceptibility of a given population or social system to harm from exposure to multiple stressors (Alwang et al., 2001). It is directly linked to the ability to prepare for, respond to, and recover from hazards and disasters. Social vulnerability therefore explicitly focuses on demographic and socioeconomic factors (Briuglio, 1995) that increase or attenuate the impacts of hazard events on local populations (Cutter et al., 2009). The perspective of numerous definitions and concepts are human centred (Bohle et al., 1994; Adger et al., 2001; Cannon et al., 2003), incorporating the role of human

agency explicitly by attributing the driving forces of vulnerability on the social, political and economic pressures imposed on individuals, which constrain their responses and ability to cope with disasters (Forsyth, 2004; Adger, 2006; Walton et al., 2008). In this manner, Wisner (2004) tracks the progression of vulnerability from root causes to dynamic pressures to unsafe conditions, which then interact with natural events. Bohle (2001) emphasizes the double structure of vulnerability with an external and internal side, whereas the external side includes the exposure to potentially damaging events, and the internal side relates to the capacity to cope with, resist and recover from the impact of a hazard. Viewed from these perspectives, vulnerability depends almost as much on the preparedness and coping capacity of the affected society as on the natural hazards itself (Bogardi and Birkmann, 2004).

Within the *natural sciences*, the focus is on physical damage assessment and related adaption processes, where vulnerability of a given entity (system, sector, region, etc.) may tentatively be defined as the expected damage resulting from anticipated environmental perturbations in consideration of the expected transformation and adaption processes (Corell et al., 2001). For example, the ‘‘end point’’ definition of Bogardi et al. (2005) views vulnerability as the residual segments of natural impacts that cannot be targeted by adaption processes. In this context, adaption-based approaches explicitly consider the examination of the adaptive capacity (O’Brien et al., 2004), which is necessary to improve the resilience and robustness of a certain system (Smit and Wandel, 2006). Therefore, the framework presented by Turner et al. (2003) sets vulnerability in the context of a human–environmental system with multiple perturbations and stresses, encompassing exposure, sensitivity and resilience as discrete but interlinked components of vulnerability.

Definitions and approaches evolved from *disaster management* concentrate upon potentials of the population to overcome and recover from the impact of a hazard (IFRC, 1999). Blaikie et al. (1994) define vulnerability with respect to this kind of perspective as characteristics of a person or group in terms of their capacity to anticipate, cope with, resist and recover from the impact of a natural hazard. Furthermore, it involves a combination of factors that determine the degree to which someone’s life and livelihood are put at risk by a discrete and identifiable event in nature or in society. Therefore, Schneiderbauer and Ehrlich (2006) state that the focus of disaster management research also shifted from ‘‘*hazard assessment*’’ to ‘‘*vulnerability analysis*’’.

However, approaches evolved from *engineering sciences* mainly concentrate upon short-term intensive extern influences on certain valued elements (e.g. buildings) and their related susceptibility by defining vulnerability as the severity of failure in terms of its consequences

(Coburn et al., 1994; Burton et al., 1993). The concern is not how long the failure lasts but how costly it is (Correira et al., 1987). From this point of view, vulnerability can, for example, be quantified by deriving an empirical relation between number of affected people and impact of a natural hazard (Vrijling et al., 1995). Especially within the engineering focused earthquake risk community (Smith, 2005), vulnerability is linked to the probability of collapse of buildings and critical lifelines considering specific earthquake scenarios and taking potential human casualties and economic losses into account (Menoni et al., 2002; Sinadinovski et al., 2005). Based on damage assessment approaches (Whitman, 1973), methodologies are developed that focus on physical components of seismic vulnerability. For instance, an assessment of damage for loss estimation studies combines parameters that are hazard related such as the determination of macro seismic intensity and peak ground acceleration with the analysis of the seismic vulnerability of built-up structures by methods such as displacement response spectra (Crowley et al., 2004) or capacity spectrum (Freeman, 2004), depending on conceptual assumptions whether a single building is studied in detail or idealized classes of buildings are considered for scenario-oriented analyses (Calvi, 1999). In this regard, the likely damage to structures is modelled by expressing their vulnerability by damage functions and fragility curves (Robinson et al., 2006; FEMA, 2010). Seismic loss is therefore described as a function of exposure, which is represented by the amount of human activity at a certain location (e.g. stock of infrastructure), vulnerability, which describes the susceptibility of the infrastructure stock, the hazard, which is expressed by the likelihood of occurrence of a specified ground motion at a certain location and damage loss conversion, which refers to the mean damage ratio (Crowley et al., 2006, quoted in Daniell, 2009; Daniell, 2011).

Yet, *integrative approaches*, for example, incorporating human, financial, social and administrative aspects for describing earthquake effects (Coburn and Spence, 1992) are established, reaching widespread methodologies such as HAZUS-MH (FEMA, 2010), where also multi-hazard risk assessments are considered (Carreno et al., 2007). More extensively, Cardona (1999) developed a holistic approach evaluating disaster risk, where exposed elements are assessed in dependence of several aspects and dimensions of vulnerability. Vulnerability is therefore characterized based upon exposure and physical susceptibility, which is hazard dependent, as well as social and economic fragilities and lack of resilience or ability to cope and recover, which both are considered as hazard independent (Cardona and Hurtado, 2000). By overlaying and weighting normalized risk parameters that may come from a multidisciplinary perspective, earthquake risk is spatially expressed in a commensurable

way aiming at a holistic view (Carreno et al., 2007; Carreno et al. 2009a; Carreno et al., 2009b). Adapting this holistic perspective, Birkmann (2006b) stresses that vulnerability should not be viewed as an isolated feature, but can be embedded within the context of a dynamic process, which implies that the focus has simultaneously to be on vulnerabilities, coping capacities and intervention tools in order to reduce vulnerability. Thywissen (2006) reveals that within earthquake engineering the susceptibility is often quantified by means of a damage ratio but especially intangible parameters related to, for example, environmental, institutional or human factors, which can often hardly be quantified properly, are set aside. However, all of the presented multidisciplinary definitions and concepts have the core notion ‘‘potential for disruption or harm’’ in common (Wisner, 2004) and connect a specified system (e.g. region, social group, sector) with risk (Füssel, 2005), respectively. This perspective implies that the present and future conditions of the elements exposed can be viewed as the central elements of vulnerability (UNDP, 2004; Cardona, 2004; Thywissen, 2006), whereby the impacts as well as the receptors of natural hazards are considered (Fuchs et al., 2011).

3. CAPABILITIES OF REMOTE SENSING IN ORDER TO ASSESS EARTHQUAKE RISK

Generally, remote sensing techniques are widely deployed for contributing to numerous aspects of earthquake risk. They provide valuable information for both hazard- and vulnerability-related research. Regarding the disaster management cycle, which can be split into four phases (Cartwright, 2005), including pre-event phases such as reduction (mitigation) and readiness (preparedness), as well as post-event phases such as response and recovery, remote sensing has a numerous share in each of it (Joyce et al., 2009a). As already described, secondary effects of an earthquake, such as tsunamis, landslides, fires, are a critical part of a comprehensive and integrative risk assessment (Deichmann et al., 2011). However, in this review, only risk components that are directly related to earthquakes are considered. For remote sensing literature contributing to secondary effects, the reader is referred to, for example, Roessner et al. (2005), Hong et al. (2007), Han et al. (2009) regarding landslides, Yamazaki and Matsuoka (2007), and Pesaresi et al. (2007) with respect to tsunamis or Gitas et al. (2008) contributing to fire hazards.

3.1. HAZARD PARAMETERS

Remote sensing for earthquake hazard research evolved simultaneously to the first appearance of commercially available satellite images in the 1970s, which were used to map active faults

and structures (Tronin, 2006). Nowadays, applications contributing to the understanding and documentation of location, slip rates as well as the kinematics and dynamics of active faults on interseismic temporal scales are based on a wide spectrum of air- and spaceborne remote sensing reaching from optical sensors to radar systems (Tralli et al., 2005).

3.1.1. PRE-EVENT

Several aspects of pre-event earthquake hazard analysis are tackled by means of remotely sensed data. Especially in pre-event geological observations, remote sensing addresses the need for quantitative observational parameters on landforms, land cover and tectonic features (Philip, 2010). Deichmann et al. (2011) note that remote sensing can contribute valuable information for microscale zonation by deriving information for producing geological, seismic or soil maps. Thereby, digital suface models (DSMs) and multi-spectral imagery proved to be a valuable data source for a detailed, spatially consistent and thematically suitable site characterization (Yong et al., 2008a, b; Shafique et al., 2012). In this manner, Theilen-Willige (2010), Reif et al. (2011) and Shafique et al. (2011) analyse geomorphologic/topographic features and settings of earthquake hazard–prone areas by extracting geomorphometric/seismotectonic parameters based on digital elevation model (DEM) data, whereas, for example, Sitharam et al. (2006), Demirkesen (2008) or Duarah and Phukan (2011) also incorporate multi-spectral imagery for terrain analysis and the description of lineaments. Based on DEMs derived from light detection and ranging (LIDAR) surface expression of faulting can be mapped with a high accuracy and detail (Cunningham et al., 2006; Begg and Mouslpoulou, 2007). By integrating optical data (Walker, 2006; Kaya et al., 2004; Gonzalez et al., 2010) and DEMs derived from stereoscopic imagery (Fu et al., 2004), active faults can be mapped less detailed but for a larger spatial extent.

By using multi-temporal synthetic aperture radar (SAR) data, pre-seismic land-surface deformations in the amount of centimetres can be measured based on the concepts and techniques of differential interferometric SAR (D-InSAR) (Kuzuoka and Mizuno, 2004; Stramondo et al., 2007; Bayuaji et al., 2010). Weston et al. (2012) compare earthquake source models determined by InSAR and seismic data and find InSAR to be a valuable technique for the estimation of earthquake source parameters such as location, seismic moment, and fault geometry. This technique has proven to be advantageous compared to, for example, GPS-based measurements in terms of costs, coverage and data accessibility (Fornaro et al., 2009) and is used in order to evaluate the seismic potential of a region (Yen et al., 2008), quantify aseismic accumulation of strains between events (Fielding et al., 2004) and calculate slip rates

of segments (Ding and Huang, 2011), what also possibly allows to identify precursory surface deformations (Tsai et al., 2006). In this manner, Tralli et al. (2007) present a conceptual case for the assimilation of InSAR measurements into the HAZUS-MH earthquake module.

Based on data from thermal sensors, short-term temperature increases prior to several earthquake events have been described numerous times for both surface and atmosphere (Gorny et al., 1988; Tronin, 1996; Ouzounov and Freund, 2004; Saraf and Choudhury, 2005; Ouzounov et al., 2006; Choudhury et al., 2006; Yurur, 2006; Ouzounov et al., 2007; Yang and Guo, 2010; Saradjian and Akhoondzadeh, 2011; Chen et al., 2011), what may hold information for earthquake prediction and warning (see also Saraf et al. (2009) for a discussion of reasons of correlation of these phenomena). Also, other anomalies related to earthquakes that can be detected by remote sensing instruments are discussed in the literature (Tronin, 2010) such as ionospheric (Kakinami et al., 2010) or cloud (Gup and Xie, 2007) anomalies.

3.1.2. POST-EVENT

Post-event hazard-related applications deal mainly with the quantification and measurement of earthquake-induced changes of the land surface. Takano and Maeda (2009) and Maeda and Takano (2010) present an approach to detect land-surface deformation induced by earthquakes based on spaceborne microwave radiometer data with a high timely resolution. Lineament analysis based on pre- and post-event optical data has shown significant changes in terms of number, spatial distribution and arrangement (Arellano-Baeza et al., 2006; Liu and Haselwimmer, 2006). Coseismic effects of strong earthquakes can cause gravity pertubations, which are detectable by spaceborne sensors dedicated to gravity field measurements (Tralli et al., 2005). For instance, the GRACE (Gravity Recovery and Climate Experiment) satellites enabled the measurement of gravity changes related to the 26 December 2004 Sumatra–Andaman (Indonesia) earthquake (Han et al., 2006).

Nevertheless, one of the main post-event application fields of remote sensing is surface deformation mapping, for what both optical and SAR data are utilized (Tronin, 2010). Approaches based on optical data use methods and concepts of change detection based on pre- and post-event data (Saraf, 2000) applying a sub-pixel correlation technique in order to be able to quantify horizontal movements with a high accuracy (Van Puymbroeck et al., 2000; Remi and Avouac, 2002; Dominguez et al., 2003; Avouac et al., 2006; Leprince et al., 2007; Ayoub et al., 2009).

Differential SAR interferometry (D-InSAR) is a widely established method for mapping significant surface-deformation signatures associated with faults, fractures and subsidences (Massonnet et al., 1993; Massonnet, 1995; Massonnet and Feigl, 1998) induced by earthquakes. For example, Reale et al. (2011) show the potential of a SAR system to monitor post-seismic deformations with a high timely resolution in an operational way. As already mentioned, there are also pre-event or inter-event applications of this technique, but it is regarded as the best tool for studying earthquake deformations especially at the moment after the shock (Tronin, 2010). Based on phase difference of multi-temporal radar observations acquired before and after a hazard event when deformation has occurred, ground deformation and displacements on the Earth's surface in range of centimetres and millimetres can be measured in a spatially continuous way (Tralli et al. 2005; Joyce et al. 2009a). Pre- and post-event SAR data were acquired, and D-InSAR techniques were applied, for example, for the 17 January 1995 Kobe (Japan) (Ozawa et al., 1997), 9 July 1998 Azores (Catita et al., 2005), 26 December 2003 Bam (Iran) (Stramondo et al., 2005; Erten et al., 2010), 24 February 2004 Al Hoceima (Morocco) (Tahayt et al., 2009), 26 December 2004 Sumatra (Indonesia) (Chini et al., 2008), 14 November 2007 Tocopilla (Chile) (Motagh et al., 2010), 12 May 2008 Sichuan (China) (Chini et al., 2010; Liou et al., 2010; Zhang et al., 2010) or 12 January 2010 Port-au-Prince (Haiti) (Eineder et al., 2010) earthquakes to quantify co- and post-seismic deformations.

Since InSAR provides precise measurements with a distance of a few kilometres away from the actual fault but is limited in generating complete deformation information in the near fault zone, Tronin (2010) states that a combined use of optical and SAR data–based displacement mapping techniques could be supplementary, whereas high-resolution optical data can be used for post-event deformation mapping in the epicentre areas.

3.2. VULNERABILITY PARAMETERS

Remote sensing for earthquake vulnerability assessment is a less long-established research field than for earthquake hazards itself. While some examples explicitly exploit and evaluate the capabilities of remote sensing generally (e.g. Rathje and Adams, 2008; Taubenböck et al., 2008) or explicitly address certain vulnerability components (e.g. Chen, 2002; Mueller et al., 2006), it can be stated that this part of earthquake risk is only in recent years subject to a more intense contemplation, especially from the remote sensing point of view. However, valuable efforts were made contributing to different vulnerability aspects of earthquake risk based on remote sensing concepts, data and methods, which are presented in the following section.

3.2.1. PRE-EVENT

Generally, remote sensing–based mapping of land cover/land use, and components of the built environment such as buildings, infrastructures or lifelines on a wide geographical scope (Polli and Dell'Acqua, 2011) can be substantial and beneficial for risk assessment and management (Tralli et al., 2005).

The potentials of remote sensing contributing to create (earthquake) risk building inventories and subsequent assessment of their physical vulnerability are discussed in, for example, French and Muthukumar (2006), Mueller et al. (2006), Ehrlich and Zeug (2008), Taubenböck et al. (2009a). Vulnerability-related building parameters that can be extracted from remote sensing data incorporate building footprint, height, shape characteristics, roof materials, location, period of construction and structure type. Especially very high and high spatial resolution optical imagery is found to be suitable to quantify and characterize the building stock based on manual cartographic methods, statistical enumeration of samples (Ehrlich et al., 2010) or automatic image information extraction methods (Sahar et al., 2010; Borzi et al., 2011). Especially the latest generation of optical spaceborne sensors are perceived as a breakthrough for operational applications especially where no alternative data source is available such as in third world countries and smaller and medium size remote urban areas (Deichmann et al., 2011). Combining several optical sensors and LiDAR data allows the automated evaluation of seismic building vulnerability with a high accuracy (Borfecchia et al., 2010), whereas, for example, the combined use of optical and SAR data is used to derive crucial parameters such as building footprint and floor number (Polli and Dell'Acqua, 2011).

In order to contribute to the assessment of demographic vulnerability components such as the regionalized number of population, spatial disaggregation approaches of population census data based on remotely sensed data are proposed (Dobson et al., 2000; Chen, 2002). Aubrecht et al. (2012) provide an overview on available multi-level geospatial information and modelling approaches from global to local scales that could serve as inventory for people involved in disaster-related areas. In the absence of or when only outdated information on total population is available, approaches for spatial extrapolation of punctual population data have been presented (Taubenböck et al., 2007).

Based on proxy variables derived from very high–resolution optical and LiDAR data Ebert et al. (2009) assess social vulnerability for an urban environment. Based on comparable data sets, Taubenböck (2011) uses physical proxies of the urban environment in order to build correlations of urban structure types (e.g. slums) with socio-economic parameters, such as income of the people, to spatially assess social vulnerability parameters, in this case financial

capacity to recover. Zeng et al. (2012) use medium resolution optical imagery for modelling social vulnerability. More specifically, Prasad et al. (2009) address social vulnerability by using very high optical data for the identification of socioeconomic clusters within the built environment in order to subsequently assess seismic risk.

3.2.2. POST-EVENT

Post-event applications that utilize remote sensing data and techniques deal mainly with the identification of land cover/land use change induced by an earthquake (Chang and Tang, 2010). Analogous to pre-event studies, the focus of such applications is on the identification, description and assessment of the present and future conditions of the built and natural environment, and hence the elements exposed.

Since remote sensing is the only technology capable of immediately capturing the damage situation over a large affected area after a hazard event (Vu and Ban, 2010), it has become a valuable tool during emergency response (Dell'Acqua et al., 2009). Especially for post-disaster rapid damage detection, mapping and assessment, recovery and rescue information in cases of a disaster event, remote sensing has proven its potential and can be identified as an operational tool to support decision makers with up-to-date spatial information (Joyce et al. 2009b; see also Saito and Spence, 2004; Voigt et al., 2007; van den Broek et al. 2009; Lang and Tiede, 2010).

However, especially rapid damage mapping is still primarily based on manual interpretation in order to avoid long (pre-)processing times (Trianni and Gamba, 2009) and to provide the acquired accuracy for rescue teams in events such as the 12 May 2008 Sichuan (China) (Wang et al., 2009b), 12 January 2010 Port-au-Prince (Haiti) (DLR-ZKI, 2010) or 11 March 2011 Tohoku (Japan) (ICSMD, 2011a) earthquakes. Thereby, limitations due to misinterpretations and data availability are still a challenging task (Kerle, 2010). Recognized damage is typically limited to severely damaged structures (Jaiswal et al., 2011), and patterns of damage are more likely to be accurately mapped whereas the quantification of damage intensity appears not feasible (Corbane et al., 2011).

Automated approaches are presented solely using post-event optical (Kaya et al., 2005; Vu et al., 2005) or SAR data (Balz and Liao, 2010), where the pre-event situation of buildings is, for example, simulated (ibid.; Wang et al., 2009a). However, the use of change detection–based methods, which compare pre- and post-earthquake images, is perceived to deliver more accurate and reliable results in general (Li et al. 2008; Ehrlich et al., 2009). This is due to the finding that damage assessment is principally a change-detection problem, where the mapping

classes are correlated with the level of damages experienced by the structures in the area of interest (Trianni and Gamba, 2009). Based on optical pre- and post-event space- and airborne imagery, concepts and methods are developed in order to detect change and subsequently derive earthquake damage of buildings in an automated way (Mitomi et al. 2002; Yano and Yamazaki, 2006; Adams et al., 2004; Kosugi et al., 2004; Turker and San, 2004; Turker and Cetinkaya, 2005; Kumar et al., 2006; Yamazaki and Kouch, 2006; Sertel et al., 2007; Yamazaki and Matsuoka, 2007; Teimouri et al., 2008; Turker and Sumer, 2008; Aydöner and Maktav, 2009; Aldrighi and Dell'Acqua, 2009; Li et al., 2010; Vu and Ban, 2010), also incorporating pre- and post-event LiDAR data in order to reach a high level of morphologic detail and accuracy (Vu et al., 2004; Li et al., 2008).

As data acquisition of optical imagery shortly after an earthquake may be limited due to cloud coverage and weather conditions and flight campaigns in remote areas can often not be carried out quickly, approaches are developed to assess earthquake damage to build structures by using pre- and post-event SAR data (Hoffmann, 2007; Gamba et al., 2007; Matsuoka and Yamazaki, 2010; Schmitt et al., 2010) or combine pre-event optical and post-event SAR imagery (Stramondo et al., 2006; Chini et al., 2009; Brunner et al., 2010).

Also other elements of the built environment are under investigation, such as critical infrastructures–like transportation networks (Huang et al., 2009; Wang et al., 2011) and remote sensing–based monitoring of reconstruction activities have been reported (Guo et al., 2010). Recent applications also deal with the assessment of ecological aspects of vulnerability of a system such as investigations of the impact of the 2008 Wenchuan (China) earthquake. These investigations focus on wildlife habitats using optical remote sensing data (Xu et al., 2009; Deng et al., 2010) and DEM data (Yu et al., 2011) or quantify damage to vegetation based on pre- and post-event optical data (Ge et al., 2009).

4. TRENDS, RESEARCH PROJECTS AND INITIATIVES

4.1. TRENDS: WHAT WAS IN THE CENTRE OF RESEARCH?

Numerous devastating natural disasters in the 1980s triggered the United Nations to proclaim the International Decade for Natural Disaster Reduction (IDNDR). Against this institutional background, the Tokyo Declaration of 1989 was followed by the forming of a UN Special Committee and the establishment of national secretaries and committees (Verstappen, 1995).

In order to give a comprehensive picture on general developments, technological innovations and ‘‘hot’’ research topics on remote sensing capabilities in the earthquake risk

context since the proclaimed time period, a quantitative analysis and categorization of the literature is performed. This analysis is both meant to complement the previous review and serve as a profound basis for the subsequent discussion about future research activities in the next chapter.

4.1.1. SEARCH CRITERIA AND SAMPLING METHODOLOGY

Based on the assumption that researchers tend to publish in well-established journals (Caron et al., 2008), peer-reviewed journals listed in the Thomson Reuther's ISI Web of Knowledge are screened from 1991 to July 2011. This is believed to give a significant overview of main developments and trends without aiming for completeness though.

First, title, keywords and abstract of all papers listed are screened for a mandatory combination of the buzzwords ''remote sensing'' and ''risk'', ''earthquake'', ''hazard'', 'vulnerability'' and the search terms' lingual variations such as plural, adjective, respectively. This implies that remote sensing is regarded as a central methodological element in the respective paper which is worth being already pointed out or at least mentioned in the title, keywords or abstract and not only in the full manuscript. This might be especially true for papers published in journals related to research disciplines with a focus not strictly bound to remote sensing (such as this journal). In contrast, for papers published in remote sensing journals, there might be less emphasis of the central methodological element necessary, since a relation to remote sensing is normally a prerequisite to publish in such a journal. Therefore, all 23 journals listed in the Thomson Reuther's subject category ''remote sensing'' are screened separately for the search terms without combing them with the buzzword ''remote sensing''.

The search results in a large error of commission, but what is found to be necessary in order to keep the error of omission as low as possible. By manual inspection of the respective search results, papers are identified that have a direct relation to earthquakes and its components such as the analysis and applications already presented above. For reasons of consistency, for example, secondary threats that are tackled by remote sensing are not integrated. One may think of numerous papers that meet these search criteria like investigations of earthquake triggered landslides by means of remote sensing or tsunami damage observed from remote sensing data. Also papers that are identified based on very general search combinations, for example, ''remote sensing'' and ''risk'', such as investigations of droughts or malaria risks integrating remote sensing are naturally not further considered.

4.1.2. CATEGORIZATION OF THE LITERATURE

Based on the sampling scheme and exclusion criteria, 251 papers were identified for the period from 1991 to end of July 2011. Following the presented categorization of the literature, the papers are indexed in dependency if the focus is on hazard or vulnerability components for pre- or post-event applications. Additionally, a categorization of the papers corresponding to the geometric resolution of the respective sensors deployed has been made.

There is a considerable trade-off between the geometric resolution of remote sensing data and the aerial coverage of a scene or data set (Rathje and Adams, 2008). Generally, remote sensing data with a coarse geometric resolution provide large spatial coverage (in terms of scene size), what implies that fewer data sets are needed to evaluate a large area. In contrast, a high geometric resolution implies relatively small spatial coverage (in terms of scene size) (ibid.).

In Tab. II-1, several remote sensing systems deployed in earthquake research are listed with affiliated geometric resolution, swath, and revisit capability. The categorization according to the geometric resolution of the sensors is adapted from the scheme presented by Neer (1999) (revealed in Möller, 2011) for optical sensors. Being aware that, for example, SAR data do not contain the same thematic information as optical data with the same geometric resolution, the categorization is later on adapted also for non-optical sensors for terms of consistency. Based on the quantitative sensor characteristics, a qualitative categorization of the sensors is made. Tab. II-1 shows the assumed spatial scales the respective analyses are typically applied on. These reach from ‘‘focal’’ to ‘‘local’’ to ‘‘regional’’ to ‘‘national’’. Generally, it can be assumed that data with a coarser geometric resolution and larger spatial coverage per scene are able to contribute to overall evaluation of pre- and postevents studies, whereas it is yet both difficult and expensive to obtain high geometric resolution data over the entire area threatened or affected by an earthquake (Rathje and Adams, 2008). Clearly, the geometric resolution of the sensors, affiliated scene sizes and the spatial scales analyses are typically applied on have also to be set in context of the objects to be analyzed. There may be applications where no substantial added value may arise answering the respective research questions based on data with a higher geometric resolution. For instance, the use of coarse resolution remote sensing data in order to analyse large-scale phenomena and objects such as large active faults may deliver appropriate results, whereas an earthquake-related evaluation of small-scale objects such as buildings may, by trend, represent only a rough estimation based on such data. However, the categorization also aims to identify characteristic spatial scales for respective application fields. In addition, it should

be noted that some of the sensors shown in Tab. II-1 in the category ‘‘national’’ can be considered more as sensors allowing ‘‘continental’’ or even ‘‘global’’ analysis, but are integrated in this category for terms of lucidity.

TAB. II-1. OVERVIEW OF SEVERAL REMOTE SENSING SYSTEMS DEPLOYED IN EARTHQUAKE RESEARCH

Platform/Satellite	Sensor/Mode	Geometric resolution (Nadir) [m]	Swath [km]	Revisit capability	Categorization of geometric resolution according to Neer (1999)	Spatial scales analysis are typically applied on according to swath/scene size
Airborne	LiDAR	0.5-1	daily coverage of 1-100 km²	Mobilized to order	very high and high resolution	*focal*
Worldview	Panchromatic Multispectral	0.46 1.85	16.4	1.1 days		
Quickbird	Panchromatic Multispectral	0.6 2.4	16.5	1.5-3 days		
Ikonos	Panchromatic Multispectral	1 4	11	1.5-3 days		
Cosmo-Skymed^	Spotlight	<1	10	~37 hours		
TerraSAR-X	Spotlight Stripmap	1 3	10 30	11 day repeat cycle; 2.5 day revisit capability		
Radarsat-2	Ultra-fine	3	20	Every few days		
SPOT-5	Panchromatic Multispectral	5 10	60-80 60-80	11 times every 26 days	medium resolution	*local*
Rapid Eye	Multispectral	6.5	77 x 1500	1 day		
ALOS	AVNIR PALSAR (Fine)	10 10	70 40-70	Several times per year as per JAXA acquisition plan		
Radarsat-1/-2	Fine	8	50	Every few days		
Landsat-5	TM Multispectral	30	185	Every 16 days	coarse resolution	*regional*
Lansat-7*	ETM+ Panchromatic ETM+ Multispectral	15 30				
TerraSAR-X	ScanSAR	18	100	11 day repeat cycle; 2.5 day revisit capability		
Radarsat-1/-2	Standard Wide	25 30	100 150	Every few days		
ERS-2		30	100	35-day repeat cycle		
Envisat	ASAR standard	30	100	36-day repeat cycle		
ALOS	PALSAR (ScanSAR)	100	250-350	Several times per year as per JAXA acquisition plan	very coarse resolution	*national*
Radarsat-1/-2	ScanSAR wide	100	500	Every few days		
Terra /Acqua	MODIS	250, 500, 1000	2300	At least twice daily for each satellite		
NOAA	AVHRR	1100	2399	Several times per day		
Envisat	ASAR ScanSAR	1000	405	36-day repeat cycle		

^ Figures quoted for one satellite constellation
* Problems with scan line corrector resulting in data gaps

Source: Sensor/Mode characteristics “Geometric resolution (Nadir) [m]”, “Swath [km]”, and “Revisit capability” according to Joyce et al. (2009). Characteristics for “Airborne LiDAR” are taken from Rathje and Adams (2008).

4.1.3. RESULTS

In Fig. II-2, the 251 identified peer-reviewed journal articles are visualized according to their assignment in terms of risk components, spatial scale and year of publication. As described, the papers are subdivided depending if the research is hazard or vulnerability related and

further differentiated in pre- and post-event research. The spatial scales are to be understood according to the sensor categories in Tab. II-1. For multi-sensoral approaches, the category of the sensor with the highest resolution deployed is used. The temporal differentiation corresponds to the publication year of the respective paper. Note that the quantitative axis is logarithmically scaled. Furthermore, the absolute number of papers is visualized in a cumulative manner over time.

First, the quantitative characteristics of the overall categories are noticeable. Continuous research was published with contributions to assess earthquake hazards and its components both for pre- and post-event analyses and applications (hazard-related pre-event: 97 papers; hazard-related post-event: 76 papers). Especially since the middle of the last decade, postevent damage analysis has increased (damage related post-event: 64 papers), whereas the number of papers dealing with pre-event vulnerability components is still comparatively low (vulnerability-related pre-event: 14 papers).

In detail, pre-event hazard analysis on a (supra-) national scale dealt mainly with the quantification of possible earthquake precursors such as thermal anomalies of the earth surface and atmosphere. In contrast, regional, local and focal analysis focussed primarily on an identification, description and measurement of geomorphometric features and seismotectonic parameters such as detection of faulting. Especially analyses and applications which can be associated with the post-hazard category are based on the development of change-detection techniques such as D-InSAR for SAR-based or sub-pixel correlation techniques for optical data in order to quantify earthquake-induced surface deformations. Being less long in the interests of research, analyses and applications contributing to the post-event impact-related part of earthquakes mainly contributed to assess structural damage of the built environment. The advent of high- and very high–resolution optical, SAR and LiDAR sensors enabled the quantification of earthquake damage with viable accuracies, whereas studies on a regional level offer quick overview assessments.

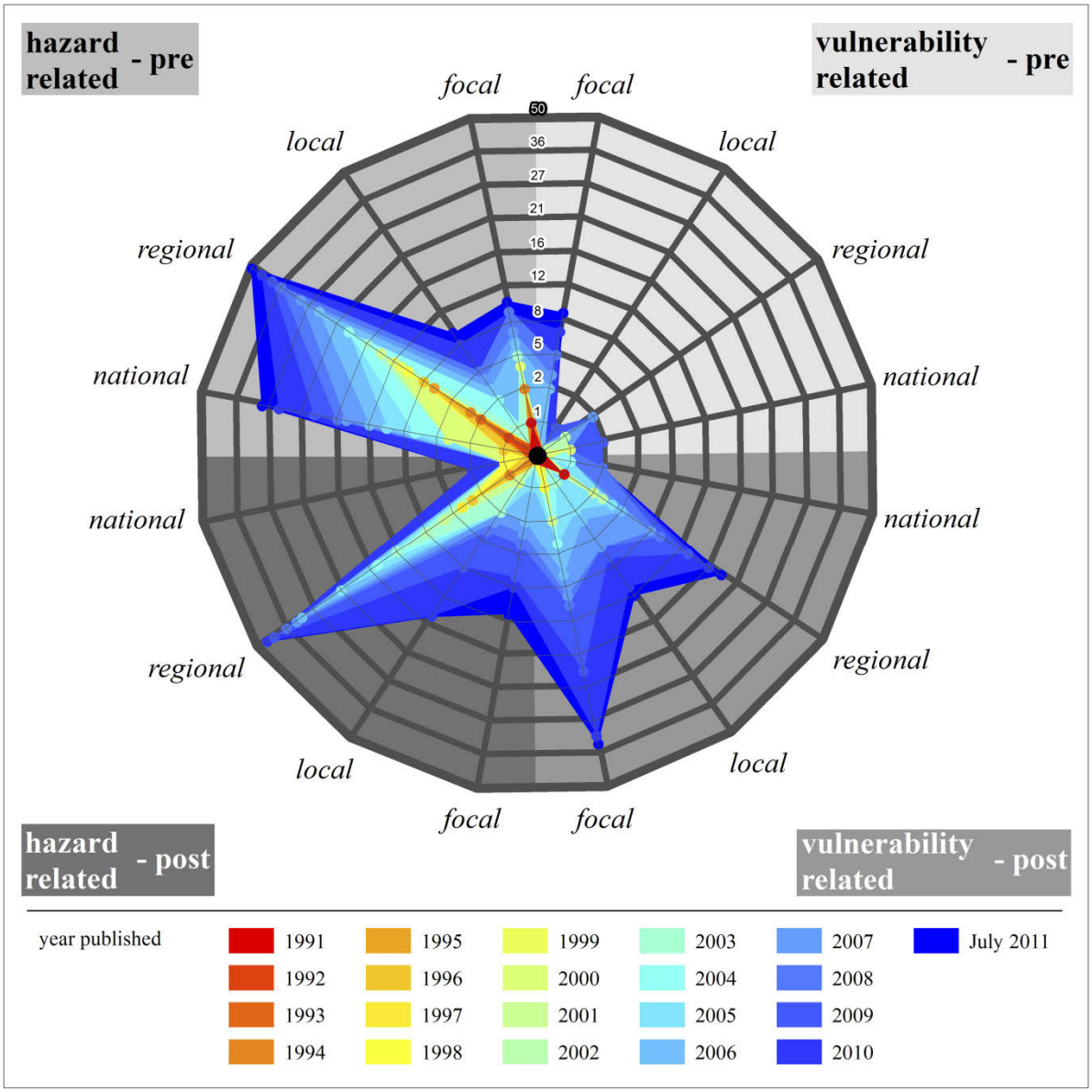


Fig. II-2. Visualization of the number of peer-reviewed journal articles differentiated according to risk components, spatial scales and year of publication based on a logarithmic scaled axis.

Less found, but increasing since the last years, are concepts, analysis and applications focusing on the pre-event vulnerability part of earthquake risk that mainly deal with the detection and assessment of elements at risk such as buildings and their physical vulnerability using latest generation air and spaceborne remote sensing sensors. In this context, remote sensing can be regarded as a widely deployed source of data and established methodological element for hazard-centred studies, which was already used since the availability of satellite images (Tronin, 2006), whereas especially for pre-event vulnerability- related research remote sensing can still be considered as relatively new source of information (Nassel and Voigt, 2006). We may conclude that this development is triggered both by the changing focus of the

scientific community that the assessment of vulnerability and its constituent elements is an important part of a comprehensive risk analysis and at the same time by the availability of sensors that first enabled an appreciable share of remote sensing capabilities to these components of earthquake risk.

Finally, it should be mentioned that, beside the discussed developments and trends concerning the specific earthquake risk categories, an overall increase in peer-reviewed literature can be found. This increase follows a general increase in peer-reviewed publications related to remote sensing, naturally connected to the impressive pace of technical innovations in this scientific field (Blaschke, 2010). Additionally, the increase is also triggered by the scientific community perceiving remote sensing as a valuable tool in earthquake risk–related investigations. This finding is, for example, quantitatively manifested in documented research such as two special issues in remote sensing journals that deal with the detailed analysis of the Wenchuan (China) earthquake on the basis of concepts, data and methods of remote sensing (Satellite observations of the Wenchuan Earthquake, 12 May 2008, International Journal of Remote Sensing Volume 31, Issue 13 (Singh, 2010); Special Section on Remote Sensing of the Wenchuan Earthquake, Journal of Applied Remote Sensing, Volume 3 (Guo, 2009)).

4.2. CURRENT RESEARCH PROJECTS AND INITIATIVES

With the designation of the 1990s as the International Decade for Natural Disaster Reduction (IDNDR), the United Nations General Assembly recognized the importance of reducing the impact of natural disasters. Since then, many different initiatives have been started to strengthen international science and stakeholders to benefit to risk-reduction capabilities:

- On the one hand, political initiatives and programmes are especially focusing on the use of earth observation data to tackle the multifaceted problem of earthquake risk management. The intergovernmental Group on Earth Observation (GEO) is coordinating international efforts to build a Global Earth Observation System of Systems (GEOSS). This initiative was launched in response to calls for action by the 2002 World Summit on Sustainable Development, aiming at exploiting the growing potential of remotely sensed data to support decision-making. With the topic ‘‘disasters’’ as one of nine ‘‘societal benefit areas’’ (SBA) within GEO, the political significance of this topic becomes obvious (GEO, 2011). The Committee on Earth Observation Satellites (CEOS), as a space component of GEOSS, coordinates civil space–borne missions to prevent unnecessary overlap. Since 2008, a CEOS Disaster SBA team coordinates all actions with respect to disaster management (CEOS, 2012).

- The Global Monitoring for Environment and Security (GMES) is a joint initiative of the European Commission and European Space Agency (ESA), which aims at achieving an autonomous and operational Earth observation capacity and is the European Union's contribution to GEOSS. Pre-operational implementation of disasterrelated mapping services has been demonstrated in multidisciplinary pre-cursor projects within the European Commission Framework Programme such as SAFER (FP7; SAFER, 2011) and PreView (FP6; PreView, 2012) aiming at real-time emergency response services or G-MOSAIC (FP7; G-MOSAIC, 2012) aiming to support early warning and crisis management operations. With GMES Initial Operation (GIO) Emergency Management Support (GIO-EMS), the goal is establishment of the operational service for post- and preparation phases (GIO, 2012). Within these frameworks, remote sensing plays a critical role as actual source of information for multiple hazards in rapid mapping and response applications.
- Furthermore, the United Nations Office for Outer Space Affairs for instance is promoting international cooperation in the peaceful uses of outer space (UNOOSA, 2011). Within, a 24-h hotline is operating as the United Nations focal point for satellite imagery requests during disasters and manages the United Nations Platform for Space- based Information for Disaster Management and Emergency Response (UN SPIDER, 2011), whereas the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA, 2012) acts as an operational coordinator in crisis situations. In 2000, the European, French and Canadian Space Agencies initiated the ''International Charter Space and Major Disasters'' (ICSMD, 2011b). The Charter represents an institutionalized cooperation between commercial remote sensing systems operators (Digital Globe, GeoEye, Spotimage, etc.), and (inter)national operators (JAXA, USGS, ESA, etc.). The goal is to provide rescue teams and local stakeholders with actual (geo-)information immediately after an hazardous event. Thereby, the Charter is responsible for the initial satellite image acquisition and processing/mapping/analyzing facilities such as UNOSAT (UNOSAT, 2011), DLR-ZKI (Voigt et al., 2007), SERVIR (SERVIR, 2012), SERTIT (SERTIT, 2012) or e-GEOS (e-GEOS, 2012) transform remotely sensed data into actual information for emergency response. The derived information such as damage maps is subsequently used by humanitarian relief organizations such as ''Medicins Sans Frontie`res'', ''German technical relief agency'' or ''Red Cross'' in order to direct supporting measures.

- Another major initiative from the private sector is the Global Earthquake Model (GEM) financed mainly by re-insurance companies—bringing together state-of-the-art science, regional, national and international organizations as well as individuals, in a global collaborative effort that aims at a lasting impact on seismic risk assessment (GEM, 2011). Remote sensing is among manifold disciplines—for example, civil engineering, seismology, architecture or social sciences—a crucial part to provide spatial, quantitative information for, among others, an intended global exposure and consequences databases. Remote sensing data and methods are developed for data inventory capturing, for interdisciplinary combination with in situ field measurements, gathering of existing earthquake databases, crowdsourcing, vulnerability estimation methods, etc. The initiative clearly reveals the community's efforts for multi-, inter and transdisciplinary progress in earthquake risk assessment. Within the context of regional programmes such as the Earthquake Model Central Asia (EMCA, 2011), or the Earthquake Model of the Middle East Region (EMME, 2011), remote sensing is also used both for the assessment of seismic hazard and vulnerability.
- On the other hand, manifold initiatives have been set up to coordinate and proclaim integrated, multi- and transdisciplinary programs for risk reduction. Among others, the International Council for Science (ICSU) proposed in 2006 that despite all the existing or already-planned activities on natural hazards, an integrated research programme on disaster risk reduction, integrated across the hazards, disciplines and geographical regions, is an imperative (ICSU, 2011). Within the United Nations, the ''International Strategy for Disaster Reduction'' (ISDR) is coordinating disaster risk reduction and ensuring synergies between the manifold players (UN/ISDR, 2011). For instance, the International Disaster and Risk Conference (IDRC) is complementing UN-ISDR's governmental policy and strategy-oriented focus by concentrating on providing a network for experts, practitioners and institutions from science, technology, business and civil society (IDRC, 2011). Further initiatives are, for example, the Global Facility for Disaster Reduction and Recovery managed by the World Bank (GFDRR, 2011), which funds, for example, GIS-based platforms for risk analysis such as CAPRA (GFDRR, 2012). Therein, remote sensing represents a critical source of data. Another example is the integrated Disaster Risk Management programme managed by the Asian Development Bank (ADB, 2011).
- The Understanding Risk Network (UR, 2012) is an example initiative connecting a global community of experts and practitioners to provide a permanent online space

where the community can share ideas and collaborate. Furthermore, conferences are held every 2 years. On a more stakeholder-related level, NGOs are found to have the potential to play a significant role in natural disaster mitigation and preparedness (Benson et al., 2001), since they can operate on grassroots level with communities, enjoy comparatively high operational flexibility and work often with the most marginalized groups in society (UN/ISDR, 2006).

5. PROPOSING RESEARCH DIRECTIONS BASED ON THE STATUS QUO: TOWARDS A ROADMAP FOR REMOTE SENSING

This literature review of remote sensing data, methods, applications and products shows a huge methodological and thematic spectrum and proves high relevance to the risk community: the capability to provide consistent, up-to-date, independent and large-area spatial data for basically any location around the globe to support and analyse hazard and vulnerability-related questions at different spatial scales is beyond controversy. However, the qualitative and quantitative review also identifies key gaps in research and demonstrates differences between theoretical capabilities of remote sensing and the status quo. Based on this review, as well as on interviews with experts, experiences in related projects, advisory boards and at conferences, we propose research directions for remote sensing in order to increase the notion on earthquake risk and the field's impact and relevance for the different groups involved.

It is obvious that the issue of earthquake risk assessment and emergency management is of multidimensional complexity with different groups such as politicians, stakeholders (insurance) industry, different disciplines within science (e.g. seismology, geophysics, civil engineering, remote sensing, social sciences) as well as people themselves having different perspectives, different unsolved questions and different open issues. In this manner, we aim to derive, discuss and suggest a path for future research directions and initiatives according to these different points of view.

5.1. RECOMMENDATIONS FOR FUTURE SCIENTIFIC RESEARCH

In general, *future scientific research* needs to incorporate the documentation and understanding of earthquake hazards and related measurable effects as well as the characterization and assessment of the vulnerability of the elements exposed within a clearly accepted conceptual risk framework.

However, the manifold conceptual frameworks reviewed regarding risk and its components lack common taxonomy and nomenclature. We argue for a demystification of terminology and acceptance of new conceptual approaches only if significant value adding is achieved. This is crucial in order to overcome conceptual barriers by setting aside often emphasized differences, which are found to be less conceptual but rather terminological (Wolf, 2011). Since a fragmented understanding of epistemological frameworks is at danger to miss out innovative ideas and bundling strengths (Hufschmidt, 2011), the development of a common, framing taxonomy and ontology in the context of risk research has to be attained—as already postulated by, for example, Brooks (2003) or Janssen and Ostrom (2006).

From a remote sensing perspective, manifold case studies have been carried out; however, examples evaluating the current maximum capabilities of remote sensing to assess earthquake risk in a systematic way are still absent. Thus, we propose integrative studies incorporating all remotely sensed data sets available to date—from LiDAR to SRTM, from SAR with different polarizations to panchromatic to multi-, super- and hyperspectral airborne and spaceborne data to ground-based remote sensing data with different geometric resolutions and repetition rates—to quantitatively and qualitatively evaluate the capability and limitations of multi-source remote sensing for pre- and post event hazard and vulnerability analysis at different spatial scales. When brought out for areas that are representative for earthquake-prone regions, such systematic evaluations would demonstrate the specific capabilities and effectiveness of individual data sources. Subsequently, the results can be the foundation in order to provide consistent and area-wide seismic risk assessment and monitoring in a standardized and operational way by developing automated thematic processors.

More specifically, as shown, remote sensing techniques allow for monitoring tectonic activities such as surface deformations from space. However, nowadays systematic and area-wide monitoring is mostly restricted by data availability. For instance, Lundgren et al. (2004) show these capabilities of surface-deformation monitoring, but for the limited area of a volcanic region. However, missions that are capable of monitoring tectonic activities and geo-tectonic threats from space with a high accuracy and high timely resolution are planned or in a conceptual phase. The proposed TanDEM-L mission (Moreira et al., 2011) aiming at a systematic, large-area interferometric monitoring will be a chance to better understand the Earth's dynamic surface processes and to improve prediction capabilities in the long term. For instance, the understanding of correlations between displacements and strain build-up or relaxation may be improved (Eineder et al., 2009; Minet et al., 2008). More promptly, Salvi et al. (2012) emphasize that the Sentinel-1 (ESA, 2011) data acquisition strategy allows

effective coverage for interferomteric data over seismically active regions at global level by acquiring data in Interferometric Wide swath mode with 250-km swath, 5 x 20 m geometric resolution (in range and azimuth respectively), a minimum revisit of 12 days with one satellite, and 6 days with both. These data are found to have the potential to substantially improve scientific knowledge and allow operational monitoring of the seismic cycle. By means of such data, a geodetic monitoring from space and thus quantification of changes regarding seismic hazards becomes viable.

Beyond this, Tronin (2006) states that future earthquake hazard–related applications may incorporate gas analysers with a high spatial resolution and sensitivity, which can be helpful for earthquake prediction and warning.

Regarding vulnerability, the above-proposed high-end studies would allow for a systematic analysis of direct or indirect correlations between the phenomenon of vulnerability (e.g. building stability or time-dependent population distribution) and remotely sensed data. To accomplish such a proposition, systematic strategies and substantial databases for validation including pre- and post-disaster information on, for example, building stocks, their seismic vulnerability and (eventually) experienced damage is needed. The latter may be the basis to develop consistent and robust damage scales for remote sensing estimates according to different data such as optical and SAR with various resolutions as proposed by Rathje and Adams (2008).

However, also for the assessment of pre-event seismic building vulnerability, common scales are needed in order to link remote sensing observations to scales that are designed to work on a comparatively high level of aggregation such as the EMS-98 (Grünthal et al., 1998) as well as locally adapted building codes. In this manner, remote sensing is not to be regarded as a panacea in general and can be further extended when, for example, combined with approaches such as automated ground-based data collection (Wieland et al., 2012), crowdsourcing (Heipke, 2010) or collective sensing (Blaschke et al., 2011).

Although many regions of the world are characterized by a lack of available data, one may think of refining, for example, the existing aggregated geospatial information related to earthquake risk such as global building inventories (Jaiswal et al., 2010) or regional census data using remote sensing data for spatial disaggregation (Setiadi et al., 2010) and enhancement also. Vulnerability is less static but rather highly varying over space and time. Many earthquake prone regions of the world show a high dynamic in terms of, for example, settlement development. Thus, up-to-date assessment, monitoring and modelling of the

expectable timely developments of the elements exposed and their conditions is an essential task.

We conclude the scientific recommendations with the observation that benchmarking of the manifold scientific contributions is largely absent. More research about research is mandatory to identify research gaps, promising solutions, dead ends or urgent needs for well-directed prioritization of future science (see also Taubenböck and Geiß 2012 for a comment).

5.2. SYNTHESIS FROM THE QUANTITATIVE ANALYSIS OF LITERATURE

Overall, the *quantitative analysis of literature* on remote sensing for earthquake risk analysis clearly reveals an academic void for vulnerability studies in the pre-event phase. The lion's share of research studies on pre-event vulnerability have to deal with the constraint of small-area coverage. From our point of view, highly detailed analysis on building level are essential to demonstrate the applicability of remote sensing generally, but the intrinsic advantage of the bird's eye view of remote sensing is large-area coverage. The availability, costs, data handling and processing requirements especially of highresolution data represent nowadays a clear limitation regarding an area-wide deployment (Rathje and Adams, 2008). Against this background, it can be stated that still strong efforts have to be undertaken working towards an area-wide, detailed and yet integrative derivation of earthquake risk parameters in an operational way.

Thus, we propose further research that explicitly addresses settlement scales on a coarser morphological level such as structure types that can be detected, characterized and assessed properly and cost-effective (see also Wyss, 2012) when using remotely sensed data with geometric resolution characteristics that enable at the same time large-area coverage.

Furthermore, future research has to be directed to vulnerability assessments beyond isolated, singular case studies. Thus, generic methodologies and frameworks are needed, which allow for applicability (implying transferability) around the globe. However, adjustments due to local and cultural idiosyncrasies cannot be bypassed. As complexity and investment for vulnerability analysis on highest scale are significant, we propose to direct research to a coarser level of abstraction, but therefore allowing area-wide coverage in a standardized way and if favoured adjustment to highest spatial and thematic detail. This also goes along with the idea of the research initiative of GEM to provide data and methods—among other disciplines—to basically allow generating systematic and standardized results all across the world. In order to prioritize tasks and areas of interest, research has systematically

to be directed to earthquake risk hot spots (Dilley et al., 2005) and subsequently hierarchically brought out on multiple spatial scales.

5.3. FUTURE NEEDS FROM A TECHNICAL PERSPECTIVE

From a *technical perspective*, the constantly increasing availability and accessibility of modern remote sensing technologies provides new opportunities for a wide range of applications. However, still today, one major constraint is data costs. These are often too high for large-area coverage, or if smaller areas are needed, data costs often are beyond the willingness or capabilities of local authorities. While satellite data are relatively low priced, some applications need high repetition rates or three-dimensional analysis, where mostly airborne and thus cost-intensive remote sensing is required. Future spaceborne missions such as the Sentinel programme of the ESA (ESA, 2011) intend to provide continuity and guarantee the availability of ERS, Envisat and SPOT-like observations (Berger and Aschbacher, 2012) to service providers and users since the technical lifetime of other missions will come to an end and in particular make data available free of charge.

Among others, launched, planned or proposed missions such as TanDEM-X (global DEM data with a geometric resolution of ~12 m and relative vertical accuracy of ±2 m; Lopez-Dekker et al., 2011), ALOS-2 (L-band SAR system with a geometric resolution of 1–10 m; JAXA, 2011), the RADARSAT constellation (planed as a medium resolution C-band mission it also includes high-resolution modes at 3 and 5 m, which were primarily designed for disaster management; CSA, 2012), DESDynI (L-band SAR with a geometric resolution of ~10 m and a multiple beam LiDAR instrument with a geometric resolution of ~25 m and 1 m vertical accuracy; DESDynI, 2011), CARTOSAT-3 (multispectral sensor with a geometric resolution of 0.25 m panchromatic; Katti et al., 2007), ALOS-3 (optical sensor with a geometric resolution of 0.8 m panchromatic and the capability to take stereo images with a swath of 50 km; Suzuki, 2012), WorldView-3 (superspectral sensor with a geometric resolution of 0.31 m panchromatic; DigitalGlobe, 2012) or EnMAP (hyperspectral sensor with a geometric resolution of 30 m; Heldens et al., 2011) have the potential to play a key role in future earthquake research. Thereby, the tasking of satellites (constellations) becomes more important in order to be able to provide the right images at the right time with the appropriate geometric resolution (ESA II 2012).

Furthermore, technical developments such as unmanned aerial vehicles may become a viable option especially for post-event rapid damage mapping (Bendea et al., 2008; Suzuki et al., 2008).

At the same time, the foreseeable increase in data volume will induce new challenges in terms of data storage, handling, mining and processing techniques. Investment in processing is still comparatively high due to mostly not fully automation of information generation procedures. During processing, adjustments are often needed due to, for example, different atmospheric conditions, land cover types, different user's requirements or the algorithms are still in experimental status. Therefore, robust methodologies are needed, as, for example, automated settlement detection approaches based on optical (Pesaresi and Gerhardinger, 2011; Pesaresi et al., 2011) or SAR (Esch et al., 2010; Gamba et al., 2011) data, which build the basis for implementing fully automated processing chains reaching from data acquisition to user-ready products. In order to address the computational requirements of especially time critical applications, the use of approaches such as high performance computing models (Plaza and Chang, 2008) appears promising.

We observe that regarding quality indicators the remote sensing literature predominantly focuses on accuracy assessment of the resulting product. However, this is not enough without a throughout discussion on performance analysis of the approach deployed with indicators such as degree of automation, level of semantic information derived, timeliness of product generation, robustness due to changes in input images or input parameters or economic viability considerations.

On the topic of risk assessment using remotely sensed data an unmanageable number of projects have been, are, or will be running. In general, reporting and thus information sharing is a critical problem that can be tackled with the possibilities to disseminate geospatial data and results through, for example, online solutions such as GeoNode (GeoNode, 2012). Also virtual globes such as Google Earth allow visualizing and sharing data, analysis and results in an easily accessible way.

Moreover, technological advancements in recent years have made it possible for Volunteer and Technical Communities (V&TCs) such as OpenStreetMap (Ramm et al., 2010), Ushahidi (Ushahidi, 2012), CrisisMappers (CrisisMappers, 2012), Virtual Disaster Viewer (VDV, 2012), Google Map Maker (Google, 2012), InSTEDD (InSTEDD, 2012) and others to provide increasing support to disaster risk management and emergency response efforts with spatial knowledge. OpenStreetMap even coordinates a Humanitarian OSM Team (HOSMT, 2012). Thereby, the principle of open data sharing is of crucial mutual benefit for participating parties and especially for the remote sensing community in support of non-duplicative classifications, target-oriented product derivation, post-classification product enhancement or the availability of validation data.

However, regarding geodata, missing data documentation, standards such as proposed by the Open Geospatial Consortium (OGC, 2012), compatibility or sharing of software is a not negligible problem. Generally, we experienced that municipalities/governments are working with different software environments and individual definitions of describing the spatial domain. Feasible ways are possibly initiatives such as, for example, INSPIRE, which in that particular case aim at establishing an infrastructure for spatial information in Europe to support environmental policies on a community level, and policies or activities which may have an impact on the environment (INSPIRE, 2012). To fully use and integrate the advantages of remote sensing products and enable comparisons, standards are required, on the data format, the land cover classes mapped, the spatial scales and units used, etc.

5.4. THE NEED FOR TRANSDISCIPLINARITY

Especially, the last paragraph includes the *need for transdisciplinarity*, since scientific results are valueless if they do not transform into practical value. In order to overcome the paradox of ‘‘knowing better and loosing even more’’ (White et al., 2001), the use of knowledge in hazard management has to be considered explicitly, whereby Fekete (2011) states that an assessment of the application of scientific results by the users should also be a critical part of research.

In this manner, projects such as LinkER (LinkER, 2012) explicitly support the implementation of operational service products in emergency response.

In general, the authors experienced a knowledge gap between the stakeholder’s and the remote sensing community. The knowledge gap includes, for example, data availability, data costs, data requirements or operational procedures, capabilities of products and thus the development of applying remote sensing to earthquake hazard and vulnerability assessment as well as post-event capabilities. Comprehensive surveying of the user and stakeholder community, including consolidation of user demand, setting of precise pass/fail criteria on selected quality-of-service parameters and the specification of user requirements is necessary (ESA II, 2012). Cartographic representation turns out to be an important aspect for the acceptance of the derived indicators. Examples are WebGIS, easily available data and applications with the capability to visualize and calculate results specific to the individual needs of particular stakeholders.

Another constraint is the difference between requirements and capabilities regarding accuracy of the respective products: the synoptic overview of remote sensing in the previous section shows area-wide and spatially highly detailed information extraction for various applications, but, for example, the accuracy of extracted pre-event vulnerability– related

information cannot compete with cadastral data sets. On the one hand, accuracies of 80–90% and sometimes even higher provide an objective basis for decisions and should increase credibility of EO-based solutions. On the other hand, these earth observation products are not established at the current legal foundation and now need to find juristic acceptance.

From a transdisciplinary perspective, the question arises to what extent the remote sensing community wants to engage for promotion of usage of its own data, results and products, or if this is to be seen as a political task. In any case, the promotion should be emphasized to, for example, major risk conferences such as the International Disaster and Risk Conference (IDRC) or major insurance industry conferences such as RIMS (RIMS, 2012).

The engagement for promotion has already been pushed by an ESA workshop (ESA II, 2012), which brought together the earth observation service industry, scientists and insurance and reinsurance companies. It was initiated in order to identify needs and common aims for the future usage of remotely sensed data. Key findings address inter alia the demand for a consistent and systematic provision of data and derived products for the extent of coverages required (most likely from national to global coverage). Beyond this, the establishment of new business models (e.g. based on a market survey) for EO data and products are demanded in order to reduce respectively share the cumulative costs for large-area coverage and provide standardized delivery mechanisms dealing with barriers related to licensing, since there is a clear message that says that industry budgets are limited for EO-based information. Preferred solutions when it comes to information distribution are platforms such as PERILS (PERILS, 2012), which provide a single access point to data and information.

5.5. THE NEED FOR MULTIDISCIPLINARITY

The need for truly multidisciplinary research regarding natural hazards is often still little addressed (Fuchs et al., 2011). However, the integration of data, information, techniques, tools, perspectives, concepts and/or theories from two or more disciplines is not only promising to lead to new knowledge, but a must to tackle such multifaceted problems of pre- and post-earthquake hazard and vulnerability analysis. As one example, the established service of earthquake loss estimates in near real time allows specifying this need for multidisciplinary integration.

Real-time estimates of damage to buildings are based on calculating the acceleration of ground motion in settlements near a reported earthquake and depend critically on accurate knowledge of hypocentre and magnitude. However, the highest uncertainties for real-time assessments of casualties depend mostly on too little knowledge on the basic structures of the

affected buildings (Wyss et al., 2006). The example of the indirect correlations between the phenomenon of building stability and remotely sensed data (see Sect. 5.1), requires systematic evidence of dependencies. This can only be achieved with sufficient databases on georeferenced buildings and their structural vulnerability.

The ways in which real-time loss estimates can be improved include the following. (1) Deepen the database on population and building stock (Wyss, 2004)—a global requirement for social sciences, civil engineering, architecture, remote sensing, geography and beyond this, also governments. (2) Refine the earthquake source model used for strong motion calculations to allow extended sources of a length appropriate for the announced moment magnitude and (3) reduce the uncertainty in hypocentral depth estimates by a catalogue of regional most likely depths for the entire globe—requirements to geophysics, soil sciences and seismology. Finally, (4) educate local disaster managers on the benefits and limitations of international loss estimates in real time (ibid.)—requirements to capacity development and transdisciplinary communication. One may add that research on transferring damage assessment into economic losses is an obvious and necessary follow-up.

As it becomes more and more understood and accepted that single disciplines are decreasingly able to progress individually, we would like to emphasize this paradigm shift and propose the willingness for an open dialogue about expectations, requirements, best practice guidelines (e.g. in form of white papers), capabilities and limitations beyond long established communities. This could serve as a ground-breaking step to transfer remote sensing products into value-added products for the manifold players within the risk community.

5.6. A POLITICAL PERSPECTIVE

Last but not least, from a political perspective, the societal loss of earthquakes is and will be enormous (see Sect. 1), for example, expressed in financial loss the Haiti earthquake in 2010 was estimated to 7.8 billions US dollars. Despite manifold scientific innovations, the problem of earthquake prediction in a deterministic sense has not been solved yet (Zschau et al., 2002).

Thus, continuous promotion and financing of integrated research and innovative technologies for prevention and early warning measures is crucial. As already mentioned above, an integrated research programme on disaster risk reduction, integrated across the hazards, disciplines and geographical regions, is an imperative (ICSU, 2011). Coordination between manifold funding agencies, streamlining of procedures and frameworks setting priorities based on political directions, scientific advice and user requirements need to be

established to ensure non-duplicative projects, well-directed progress and implementation phases. Initiatives with a voluntary nature such as, for example, GEO (GEO, 2011) may serve as valuable institutionalized platforms in order to achieve the mentioned goals, but are nowadays confined due to their restricted financial resources and limited commitment of members.

A noble, but unrealistic goal would be the implementation of a general framework for a systematic global collection of all initiatives and projects, data, research results, information products, developed tools, participating stakeholders and a documentation of lessons learnt and best practice examples. A commendable example is demonstrated by the ‘‘Atlas of Vulnerability and Resilience Research’’, which collects all research studies, theses, reports and even minor studies that deal with vulnerability and resilience in the context of disasters for Germany, Austria and Switzerland (Fekete and Hufschmidt, 2012).

Along this, an improvement in operational interfaces, a networking platform to increase user involvement, foster discussions and collaborations among the manifold stakeholders along the chain of risk management is essential. A commendable, initial example of this proposition is for instance the above-mentioned Understanding Risk Network (UR, 2012). However, an institutional responsibility for these tasks is unclear, not to mention knowledge about general responsibilities at global to local levels.

From a remote sensing perspective, the ensuring of continuity of earth observation sensors and data availability and the subsequent derivation of precise, objective and reliable information is essential to promote sustainable development. As an example, making use of remote sensing applications to track the temporal and spatial growth of urban areas as well as the changes of exposure of vulnerable elements is a crucial topic. Also, assisting institutions in rapid, precise and objective assessments of impacts of earthquakes as a way to facilitate the delivery of humanitarian assistance to those communities affected by earthquakes is a necessary task.

However, the main challenge within the field of disaster reduction is to change people’s perception so that they can recognize this notion of disasters as the outcome of a development process whereby societies have implicitly generated vulnerabilities and risks, which become evident during the disaster (Villagran de Leon, 2006). With its fascinating images from above remote sensing is pre-destinated to reach the broad public. This can be the basis for raising awareness and creating consciousness of potentially affected people; this is a promising step for a major reduction in their personal vulnerability. The transparent communication of complex scientific results in an easy accessible environment and understandable language is

of crucial importance, also when addressing stakeholders and politicians. However, it is clear that the step from scientific results to policy-making is challenging and often neglected, unfortunately from both sides.

Moreover, this review of the manifold different aspects—scientifically, technically, trans- and multidisciplinary and politically—without claiming to be complete, is meant as contribution towards an open dialogue among all involved parties for promoting remote sensing capabilities to support sustainable development.

ACKNOWLEDGEMENTS

The authors thank Michael Eineder (DLR) for sharing his thoughts about future earthquake hazard-related research and missions, Annette Fröhlich (DLR), Harald Mehl (DLR), Eike-Marie Nolte (KIT) and Juan Carlos Villagran de Leon (UNOOSA) for sharing their thoughts on the political perspective, Friedemann Wenzel (KIT) for hints regarding especially future scientific and technical issues and Tobia Lakes (HU Berlin) for her comments on the initial manuscript. Furthermore, the authors want to acknowledge Beverly Adams (Guy Carpenter) for the fruitful discussions regarding the applicability of EO data and information related to the insurance industry. The authors also wish to thank the two anonymous reviewers for their constructive comments and suggestions. The research was founded partially by Helmholtz-EOS (Earth Observation System).

Chapter III

Assessment of seismic building vulnerability from space

Earthquake Spectra, 30(4), 1553-1584

Christian Geiß, Hannes Taubenböck, Sergey Tyagunov, Anita Tisch, Joachim Post and Tobia Lakes

doi: http://dx.doi.org/10.1193/121812EQS350M
Received 26th December 2012; revised 13th February 2013; accepted 18th March 2013

ABSTRACT

This paper quantitatively evaluates the suitability of multi-sensor remote sensing to assess the seismic vulnerability of buildings for the example of the city Padang, Indonesia. Features are derived from remote sensing data to characterize the urban environment and are subsequently combined with *in situ* observations. Machine learning approaches are deployed in a sequential way to identify meaningful sets of features that are suitable to predict seismic vulnerability levels of buildings. When assessing the vulnerability level according to a scoring method, the overall mean absolute percentage error is 10.6%, if using a supervised Support Vector Regression approach. When predicting EMS-98 classes, the results show an overall accuracy of 65.4% and a *Kappa* statistic of 0.36, if using a naïve Bayes learning scheme. This study shows potential for a rapid screening assessment of large areas which should be further explored in the future.

1. INTRODUCTION

Most casualties from earthquakes are associated with collapsing buildings. Therefore the rapid urbanization observed in earthquake prone regions places more people at risk than ever before. As a consequence, the death toll in urban areas is expected to reach unprecedented levels (Bilham, 2009). Developing countries are characterized by dynamic urban growth with large shares of unplanned, spontaneous and often highly vulnerable settlements. Simultaneously, these settlements are highly variable over short time scales (Wieland et al., 2012). In this regard the continuous assessment and monitoring of the seismic vulnerability of buildings is a challenging task, especially when large-area evaluations are required. Numerous studies emphasize that remote sensing can play a valuable role in supporting the extraction of relevant features for pre-event vulnerability analysis of built-up structures (French and Muthukumar, 2006; Mueller et al., 2006; Sarabandi et al., 2008; Taubenböck et al., 2009a; Borfecchia et al., 2010; Sahar et al., 2010; Borzi et al., 2011; Deichmann et al., 2011; Wieland et al., 2012; GEM, 2013). The intrinsic advantage of remote sensing is the ability to offer an overview of building stocks and serve as a screening method for derivation of building vulnerability related features, such as shape characteristics, height, roof material, year of construction, structure type and spatial context (Geiß and Taubenböck, 2013).

Approaches evolved from engineering science to assess the seismic vulnerability of buildings such as the quantification of displacement response spectra (Crowley et al., 2004), capacity spectrum (Freeman, 2004), or damage probability matrices (e.g. EMS-98; Grünthal et al., 1998) consider single structures, studied in a detailed analytical way, or aggregated and idealized classes of buildings, that can be assessed for large-area applications (Calvi et al., 2006). Existing studies have focused on two different approaches linking remote sensing data to seismic building vulnerability: (i) by defining a direct relation, by for example, vulnerability curves based on features that can be gained from remote sensing data (Taubenböck et al. 2009a; Borzi et al., 2011) or (ii) by using remotely sensed data primarily for spatial inter- and extrapolation of *in situ* surveys, by for example, supervised classification techniques (Borfecchia et al., 2010). The former approach can only perform well if finding and defining valid vulnerability curves based on the available input features for respective built-up structures (Taubenböck et al., 2009a). Analogously, the accuracy of the latter approach is highly depending on finding significant and robust proxy variables that have a high correlation with *in situ* observations. Typically only a small number of features that influence seismic building vulnerability, such as building height or shape features are reflected directly or can be gained from remote sensing data. Naturally, both approaches are

further dependent on the accuracy of the derived features and the precision of the results may vary with respect to local idiosyncrasies and different built-up structures. To date, studies have evaluated the potential of remote sensing in a solely qualitative manner (Mueller et al., 2006) or presented results that emphasize the viability of the use of remote sensing (Borfecchia et al., 2010), but lack the identification and documentation of the necessary and meaningful features. The goal of this paper comprises the quantitative evaluation of the potential and limitations of satellite remote sensing for assessing seismic building vulnerability in Padang, Indonesia. High resolution optical imagery, height information from a normalized digital surface model (nDSM), and multi-temporal medium resolution optical data, are used to calculate features that characterize the urban environment. These are combined with *in situ* data about the buildings' vulnerability levels. The *in situ* assessed vulnerability levels are based on a scoring method and the well-established EMS-98 scheme.

By using techniques of machine learning feature selection we aim to identify the most relevant features that can be used for the estimation of seismic building vulnerability. Techniques of regression analysis are also utilized to determine the strength, direction and significance of independent variables on the dependent variable, namely *in situ* assessed vulnerability level. Therefore, our first research question is as follows:

1. Which features can be derived from satellite remote sensing data that best explain seismic building vulnerability?

Subsequently, we assess the accuracy of supervised regression and classification techniques to answer the question:

2. How suitable are features derived from satellite remote sensing data for estimating seismic building vulnerability levels?

2. EXPERIMENT SETTING AND DATA

2.1. STUDY AREA AND GENERAL EXPERIMENT SCHEME

We chose the city of Padang, Indonesia because it is situated in one of the most earthquake-prone regions in developing countries worldwide. Padang is located in West Sumatra and is the capital city of the Sumatera Barat province. It has approximately one million inhabitants and represents the third largest city on the island of Sumatra (Fig. III-1). The dynamic urban system of Padang is characterized by a high concentration of inhabitants and infrastructure. Padang has supra-regional relevance with an international airport and railway connection. The city plays an essential economical role for the coastal region and mountainous hinterland. Off

the coast of Padang the Sunda Arc marks an active convergent plate boundary, placing the city in a zone of extremely high probability of severe earthquakes and secondary effects such as tsunamis (Taubenböck et al., 2009b). On September 30th 2009, Padang was hit by an earthquake with a moment magnitude of M_w = 7.6. Despite its size, the Sunda megathrust was not ruptured and the stress on the Mentawai segment, which was accumulated over 200 years, was not significantly reduced. The megathrust strain-energy budget remains at a high level, threatening a great, and also tsunamigenic earthquake with a magnitude M_w > 8.5 on the Mentawai patch (McCloskey et al., 2010).

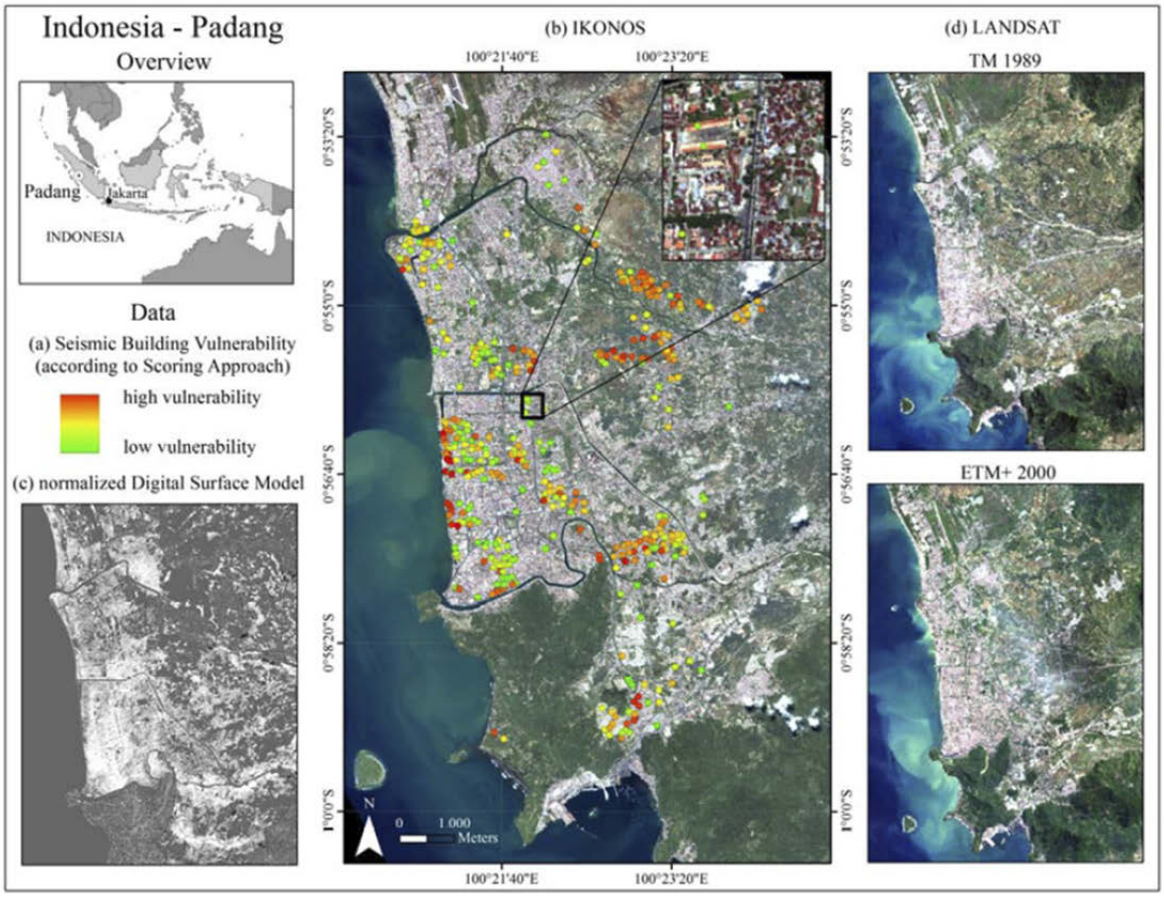


Fig. III-1. Overview on the location of the study area Padang, Indonesia, *in situ* and remote sensing data; (a) *in situ* assessed buildings represented as points and superimposed on (b) multispectral IKONOS imagery; (c) normalized digital surface model (nDSM) as basis for height estimations; (d) multitemporal LANDSAT data for spatiotemporal analyses.

To evaluate the potential of remote sensing for assessing the seismic vulnerability of Padang's buildings we followed the schematic workflow shown in Fig. III-2. Based on remote sensing data we calculated features on two different spatial levels, building and block level (see Tab. III-2 for a detailed list of features). Subsequently, the *in situ* data with affiliated vulnerability information was added. Techniques of machine learning based feature selection

and regression analyses were then used to identify features that were most suitable to assess seismic vulnerability levels and quantify their explanatory content. Under consideration of the results, we built several supervised regression and classification models and assessed the accuracies of the predictions. In-depth explanations are given throughout sections "experiment setting and data" and "methods".

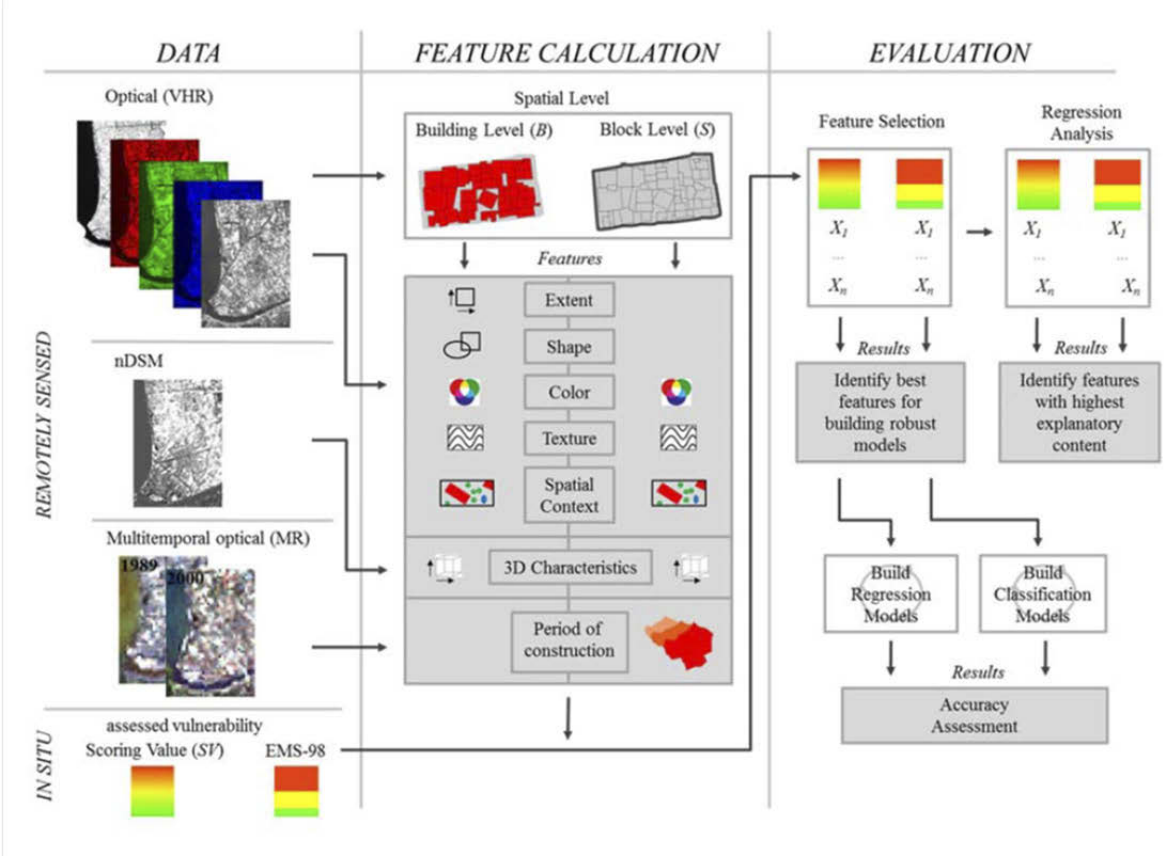


Fig. III-2. Overview of the general experiment scheme followed in this study, which is subdivided according to the main categories "data", "feature calculation" and "evaluation". Detailed explanations are given throughout sections "experiment setting and data" and "methods".

2.2. IN SITU DATA AND SEISMIC VULNERABILITY ASSESSMENT

The *in situ* data on seismic vulnerability in Padang was collected in February/March, 2008 within the "*Last-Mile*" project (Taubenböck et al., 2009b). The building inventory database compiled for Padang, based on a ground truth survey, includes information about physical characteristics of 434 buildings in the city. The sampling scheme of the buildings aimed at both the incorporation of all existing housing types of Padang (from informal small shacks to high-rise commercial buildings) and broad spatial coverage (from the urban center to periphery suburbs) (Fig. III-1). The database includes information about geometry, material of bearing structures and walls, foundations and local soil conditions, material of the roof, type of building, etc. For most of the inspected buildings the dataset is supplemented by results of

physical tests providing information about the reinforcement and quality of concrete of the main bearing structures. Additionally, for every building in the database there is an indication of the damage level due to previous earthquakes (Taubenböck et al., 2013). Based on this data, a vulnerability scoring approach and a classification according to the EMS-98 scale was carried out to assess the seismic vulnerability of the buildings as explained in the next two subsections. We incorporate both methods in this study to provide an assessment that reflects local idiosyncrasies and expert knowledge (scoring approach), and an assessment according to a more generalizable and wide-spread scheme (EMS-98). The working definition of seismic building vulnerability of this paper is related to engineering driven definitions, whereby vulnerability describes the probability of damage to a building under specified earthquake influence (Whitman, 1973).

2.2.1. DATA FROM SCORING APPROACH

We incorporate data from a vulnerability scoring approach, which was carried out after the *in situ* data collection within the "*Last-Mile*" project. It is based on an indexing method (see e.g., Calvi et al., 2006 for a general description of such methods) that incorporates local expert knowledge. The aforementioned documented building parameters are first expressed quantitatively on a normalized scale. High values express characteristics that are considered as favorable, regarding the buildings' seismic vulnerability. Subsequently, individual weights for the respective parameters are assigned based on expert observations made in the study area during previous research. For instance, an established method to assess the stability properties of reinforced concrete is the aforementioned Schmidt rebound hammer test. Completed structural survey results in the study area confirmed the hammer test outcomes to be the most important indicator and are therefore given the highest weight. Finally the weighted values of the respective parameters are summed up. For brevity the reader is referred to Mück et al. (2013) for a detailed description of conceptual and methodological details of this approach. Fig. III-3 shows the distribution of the scoring values within an interval of [10.85, 25.6], a mean value of 18.4 and a standard deviation of 2.89. Lower scoring values express higher building vulnerability, while higher values represent lower vulnerability.

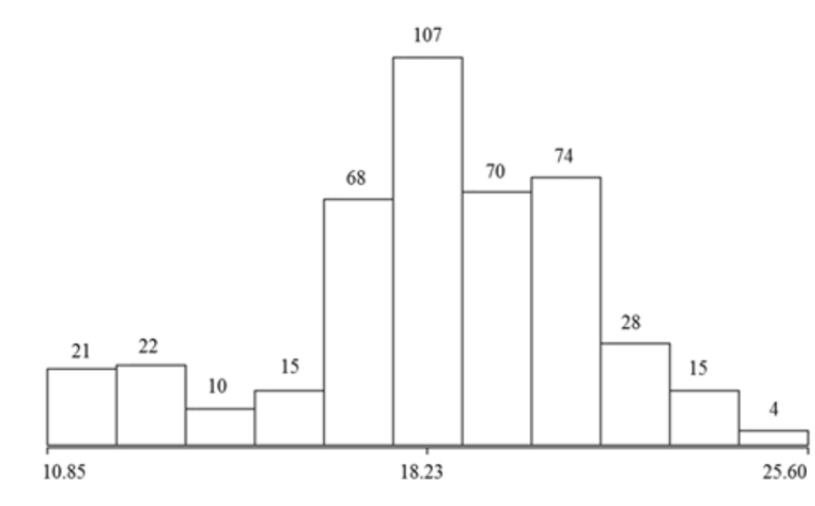


Fig. III-3. Distribution of the scoring values for 434 *in situ* surveyed buildings within an interval of [10.85, 25.6].

2.2.2. EMS-98 CLASSIFICATION DATA

The classification scheme of the European Macroseismic Scale (EMS-98) (Grünthal et al., 1998) was initially designed for Europe. However, it is accepted as a useful point of reference for other areas when collapse data are absent (Jaiswal et al., 2011). According to the EMS-98 different types of buildings are classified into six vulnerability classes, denoted alphabetically from A (highest vulnerability) to F (lowest vulnerability). The classification depends primarily on the building material and the type of structure, taking into consideration a variety of additional factors (such as constructional and architectural features, quality of materials and workmanship, age and state of preservation, etc.), which may affect the seismic performance of the buildings. As mentioned, this kind of information is contained in the *in situ* data set. Additionally, available reports of the West Sumatra Earthquake of September 30th, 2009, containing descriptions of damages for different types of buildings in the affected area, are considered (Sengara et al., 2010). It is worth noting that among the main causes of damage reported, poor quality of materials and construction was mentioned. This should be taken into account when assessing the seismic vulnerability of existing buildings in the area.

As a first step, all buildings were differentiated based on the material of the main structural elements: steel, concrete, bamboo, and wood. For the case of steel structures, according to the EMS vulnerability table, the class E can be assigned as the most probable one with the range of probable classes C-F. For wooden structures the class D can be assigned with the range of B-E. Bamboo structures are not included in the EMS-98, however we considered that they are similar to, but slightly more vulnerable than wooden structures. We therefore assumed a proper range for these structures is C-D (with the less probable range of B-E). In the case of

concrete structures, we had to keep in mind the existing uncertainty of the structural type (in the database there is no indication if the structures are made of concrete, masonry units, or reinforced concrete frame). Therefore for concrete structures we assigned the class C initially, with the less probable range of A-E. Secondly, we took into account weaknesses that are mentioned in the inventory database. If necessary, the corresponding modifiers were applied to the initial vulnerability class. Assessment of weaknesses should be building-type-specific, by examining if the principal rules of earthquake-resistant design (including quality, regularity, homogeneity, ductility, overall integrity and stability of the structure) are observed. The following essential weakness characteristics were considered: wide openings of the ground floor (which may cause soft storey effects), very thin bearing elements (columns and beams) or their absence, lack of reinforcement (or low reinforcement ratio) of the main structural elements, poor results from the hammer tests. If some of the listed weaknesses can be identified, the vulnerability class of such buildings was downgraded. Besides weaknesses there may be strengths (e.g. relatively high reinforcement ratio or enlarged dimensions of the bearing structures), however, we do not take those into account due to the reported poor quality of materials and workmanship in the area in general. Another illustrative indicator of seismic performance of structures is the damage observed from previous earthquakes. Based on this we downgraded the vulnerability class in the case of moderate to severe damage to class B or even to the range of A-B, depending on the structural type. Tab. III-1 shows the EMS-98 classes assigned and the affiliated number of buildings. For some classes there were only a sparse numbers of instances, and so some classes were aggregated. The bold letter indicates the most likely vulnerability class. This is done to have a more sufficient number of samples per class for applying supervised learning approaches.

TAB. III-1. EMS-98 CLASSES AND AFFILIATED NUMBER OF INSTANCES

Initial classification	*Number of instances*	*Aggregated classification*	*Number of instances*
A-B	7	A-**B**	242
B	235		
B-C	26	B-**C**	149
C	123		
C-D	32	C-**D**-E	43
D-E	3		
E	8		

2.3. REMOTE SENSING DATA

As mentioned, no remote sensing data set or derived products offer the complete set of features that are frequently used for analytical engineering assessments. However, remote sensing offers a huge spectrum of sensor systems that may deliver useful data for a subset of seismic vulnerability indicators. Active remote sensing, such as airborne LiDAR (Light Detection And Ranging) measurements, are frequently used to extract buildings heights (e.g., Sirmacek et al., 2012). When combined with optical data, high resolution Synthetic Aperture Radar (SAR) can be utilized to derive crucial building features such as footprint and height (e.g., Polli and Dell'Acqua, 2011). The latest generation of multispectral spaceborne sensors such as IKONOS, QuickBird-2, GeoEye-1 or WorldView-2 enables satellite based detection, characterization and assessment of objects in urban environments, on a large scale (e.g. Maktav et al., 2005; Weng and Quattrochi, 2006; or Rashed and Jürgens, 2010). Especially for developing countries and remote areas, the airborne acquisition of data is still a challenging task in terms of flight campaign preparation and keeping data costs reasonable. We therefore utilize data with resolution characteristics that can be achieved by spaceborne sensors for this study, which are comparatively low priced and easily available, although we are aware that especially airborne sensors, can achieve higher resolutions. For the evaluation of spaceborne remote sensing capabilities multispectral IKONOS imagery, height information from a nDSM and LANDSAT data were acquired within the "Last-Mile" project (Taubenböck et al., 2009b).

The optical IKONOS imagery (acquisition date: 2005-04-12) covers a spectral range of 0.445-0.853 µm, with a geometric resolution of 1 m in the panchromatic band and 4 m in the multispectral (blue, green, red, nir) bands. The satellite has a revisit capability of 1.5-3 days and the sensor's swath-width of 11 km allows covering large parts of cities. The data were pansharpened and atmospherically corrected using the ATCOR (Atmospheric and Topographic Correction) model (Richter, 1996). Height information was derived from a digital surface model (DSM) and a digital terrain model (DTM). Measurements from the models were based on the return signals received by two radar antennas mounted on an aircraft, and the application SAR interferometry techniques (see Li et al., 2004 for a detailed description of the technique applied). Both data sets have a geometric resolution of 5 m and a Root-mean-square-error *(RMSE)* regarding the vertical accuracy of 1 m. To receive relative height information of elevated objects a nDSM is calculated by subtracting the height values of the DTM from the height values of the DSM using map algebra. Although the height information for this study was acquired based on an airborne sensor, the data resolution

characteristics can also be achieved or even refined when using along-track stereo data of spaceborne sensors, such as Cartosat-1, IKONOS, or WorldView-2 (Sirmacek et al., 2012). In addition, we used data from the LANDSAT sensors Thematic Mapper (acquisition date: 1989-07-25) and Enhanced Thematic Mapper (acquisition date: 2000-07-15). Both sensors have 7 multispectral bands covering a spectral range of 0.45-2.35 µm with a maximum geometric resolution of 30 m (TM) and 15 m (ETM+) (Irish, 2008). The satellites have a revisit capability of 16 days and the data can be accessed free-of-charge within a public image archive which dates back to 1972.

3. METHODS

3.1. FEATURE CALCULATION

Before we calculated features from the remote sensing data we first manually digitized the building footprints and derived building blocks from a road network (Taubenböck et al., 2008a). In the subsequent steps, features were calculated for both. The building level was used for calculating features that refer to characteristics of the respective building (designated by the subscript *B*), whereas the block level (designated by the subscript *S*), was used to describe the spatial setting the respective buildings are embedded in. This is why we chose to use building blocks derived from a street network rather than artificial spatial units, such as quadratic objects. This allows us to reflect the urban morphology, which is constituted by distinct areas that are generally irregularly shaped. Simultaneously, the difficulty of having to determine the optimal kernel size *a priori* is avoided (Herold et al., 2003).

The calculated features relate to the building objects' two-dimensional extent and the description of their shape characteristics to directly reflect the influence of geometry on the vulnerability level. Through manual image interpretation a further distinction of buildings with "flat" and "non-flat" roofs was made. In addition, the multispectral information available was used to calculate both 1^{st} and 2^{nd} order statistical values on the building level as well as on the block level. Thereby, the spectral information on building level primarily serves as a descriptor of roof surface material and arrangement (Mueller et al., 2006); whereas the information on building block can be utilized to describe distinct urban structures (Herold et al., 2003; Wu et al., 2008). Mean and standard deviation values of the different image bands as well as band ratios, which are intended to emphasize spectral dissimilarities, were calculated. Rotation-invariant texture measures for the panchromatic and near-infrared band were calculated using both the co-occurrence matrix (GLCM, Haralick et al., 1973) and grey

level difference vector (GLDV, Weszka et al., 1976). Texture is the term used to characterize tonal or gray-level variations in an image. It has been demonstrated that it can provide supplementary information to overcome a lack of spectral resolution (Pacifici et al., 2009).

Features explicitly describing the spatial context consist of the calculated distance to neighboring buildings, the area of building blocks and the average size of the buildings located within. Furthermore, spatial metrics such as proportion measures of land cover classes "buildings", "sealed", "grass/meadow", "trees", and combined classes such as "vegetation" and "impervious surface" are calculated based on an urban land cover map derived from the IKONOS data. The urban land cover map has a high overall accuracy of 97% correctly classified pixels based on an automated object based, hierarchical classification methodology and subsequent manual enhancement (Taubenböck et al., 2009b). Additionally, a semantic classification which is built on physical features that describe the urban morphology is incorporated. The classification describes the socio-economic status of the population by distinguishing "slums", "suburbs", "low income areas", "medium income areas", and "high income areas" (Taubenböck et al., 2009c). Housing clusters that describe the socioeconomic level of occupants represent a proxy variable that has already been used in previous studies for the assessment of seismic risk (Prasad et al., 2009). The incorporation of height information allows the calculation of 3D features such as building floor number, which was calculated with an accuracy of 86.7% (Taubenböck et al., 2009c), floor space, ratio of diameter and height as well the average building height within a building block. Slope values are calculated at the block level to describe topographic location. By analyzing two Landsat images, the period of construction was derived at the individual building level based on a semi-automated post classification change detection procedure, with an overall classification accuracy of 89.1% for 1989 and 92.4% for 2000 (Taubenböck et al., 2008b).

Overall, each building object is represented by a 132-dimensional feature vector, whereby 73 features are related to individual buildings, and 59 provide block level information. The features calculated (i) aim to reflect features that went into the calculation of the *in situ* values when they can be quantified by means of remote sensing data, such as building height or geometry, (ii) represent the spectrum of features frequently utilized in previous studies on remote sensing based building vulnerability assessment, and (iii) features that were used to discriminate urban built-up structures by means of remote sensing previously.

TAB. III-2. LIST OF FEATURES DERIVED FROM REMOTE SENSING DATA.

Sensor/ Data			Feature	Description
IKONOS	SPEC. INFO 1st ORD.	EXTENT	Area $_B$	
			Perimeter $_B$	
			Width $_B$	Width of polygon's main line
			Length $_B$	Length of polygon's main line
			Length/Width $_B$	
		SHAPE	Asymmetry $_B$	Description of a building's geometry by approximate comparison of the shape with 2-dimensional geometrical forms such as square, rectangle or ellipse
			Elliptic Fit $_B$	
			Radius of Smallest Enclosing Ellipse $_B$	
			Radius of Largest Enclosing Ellipse $_B$	
			Rectangular Fit $_B$	
			Roundness $_B$	
			Density $_B$	
			Main Direction $_B$	Direction of a building's longer eigenvector
			Border Index $_B$	Description of how jagged/smooth a building's shape is
			Shape Index $_B$	
			Compactness $_B$	
			Roof Type $_B$	Discrimination of buildings with "flat" and "non-flat" roofs
		MEAN	Max. Diff. $_{B,S}$	Maximum difference of the image layers' mean intensity value
			Brightness $_{B,S}$	Mean brightness of the image layers
			Blue(1) $_{B,S}$	Mean spectral value
			Green(2) $_{B,S}$	
			Red(3) $_{B,S}$	
			Nir(4) $_{B,S}$	
		STDV.	Blue(1) $_{B,S}$	Standard deviation of the spectral values
			Green(2) $_{B,S}$	
			Red(3) $_{B,S}$	
			Nir(4) $_{B,S}$	
		BAND RATIOS	[M(4)-M(3)]/[M(4)+M(3)] $_{B,S}$	Mean spectral value (M) of respective band used for transformation according to formula
			Brightness×M(1)/M(2) $_{B,S}$	
			Brightness×M(1)/M(3) $_{B,S}$	
			Brightness×M(2)/M(3) $_{B,S}$	
			M(1)/M(2) $_{B,S}$	
			M(1)/M(3) $_{B,S}$	
			M(1)/M(4) $_{B,S}$	
			M(2)/M(3) $_{B,S}$	
			M(2)/M(4) $_{B,S}$	
			M(3)/M(4) $_{B,S}$	
			M(2)/[M(1)+M(2)+M(3)+M(4)] $_{B,S}$	
			[M(4)/M(3)]/[M(1)+M(2)+M(3)+M(4)] $_{B,S}$	
			[M(3)-M(2)]/[M(3)+M(2)] $_{B,S}$	
			[M(3)-M(1)]/[M(3)+M(1)] $_{B,S}$	
nDSM		3D FEATS	Floor Number $_B$	Building height expressed in floors
			Floor Space $_B$	Area×Floor Number
			Diameter/Height $_B$	Length/Floor Number
			Width/Height $_B$	Width/Floor Number
			Average Building Height $_S$	Average building floor number per building block
	SLOPE		Slope $_S$	Average slope value per building block; Slope value per pixel corresponds to avg. degrees in 3 x 3 window
LANDSAT			TM ETM+	Period of construction $_S$	Semi-automated classification of "urban"/"non-urban" areas for 1989 and 2000

Subscripts: B = Building Level, S = Block Level, Pan = Panchromatic, Nir = Near-Infrared, M = Mean Value;

Sensor/ Data			Feature	Description
SPEC. INFO 2nd ORD.	Texture after Haralick (all directions)	GLCM	Angular 2nd Moment $_{B\,pan\,nir,\,S\,pan\,nir}$	Also called "Energy"; Quantification of texture uniformity by counting pixel pair repetitions
			Contrast $_{B\,pan\,nir,\,S\,pan\,nir}$	Difference between the highest and the lowest values of a contiguous set of pixels
			Dissimilarity $_{B\,pan\,nir,\,S\,pan\,nir}$	Similar to "Contrast" but weighting the diagonal linearly and not exponentially
			Std. Dev. $_{B\,pan\,nir,\,S\,pan\,nir}$	Measures the dispersion of values around the mean
			Mean $_{B\,pan\,nir,\,S\,pan\,nir}$	Pixel value is weighted by the frequency of its occurrence in combination with a certain neighbor pixel value
			Homogeneity $_{B\,pan\,nir,\,S\,pan\,nir}$	Describes homogeneity by measuring the presence of near diagonal elements in a GLCM
			Entropy $_{B\,pan\,nir,\,S\,pan\,nir}$	Quantifies the disorder by approximating the textural uniformity
			Correlation $_{B\,pan\,nir,\,S\,pan\,nir}$	Measures the linear dependency of pixel values
		GLDV	Angular 2nd Moment $_{B\,pan\,nir,\,S\,pan\,nir}$	Same as respective GLCM measure but based on a grey level difference vector
			Contrast $_{B\,pan\,nir,\,S\,pan\,nir}$	
			Mean $_{B\,pan\,nir,\,S\,pan\,nir}$	
			Entropy $_{B\,pan\,nir,\,S\,pan\,nir}$	
	SPAT CONTEXT		Distance to 1st, 2nd, 3rd Neighbor Building $_B$	
			Area Building Block $_S$	
			Average Building Size $_S$	Average area of the buildings per building block
			Degree of Building Density $_S$	Percent of building block area covered by buildings
			Share of LC class "sealed" $_S$	Shares of respective land cover (LC) classes per building block
			Share of LC class "grass/meadow" $_S$	
			Share of LC class "tree(s)" $_S$	
			Share of Vegetation $_S$	Share of LC class "grass/meadow" + Share of LC class "tree(s)" $_S$
			Share of Impervious Surface $_S$	Degree of building density + Share of LC class "sealed" $_S$
			Structure Type $_S$	Classification of semantic structure types "Slum", "Suburb", "Low Income", "Medium Income", "High Income"

3.2. FEATURE SELECTION ANALYSIS

The selection of features to be used for regression and classification models is generally a difficult task, especially when dealing with a huge number of features as in this study. These often exhibit redundancy, are highly correlated, and suffer from the "Hughes phenomenon". The latter describes the effect that for a limited amount of samples the predictive power decreases as the dimensionality of the feature vector increases (Hughes, 1968). Therefore, two

machine learning based feature selection algorithms were applied on the data set. The feature selection approaches were chosen because they can handle both regression problems and evaluate discrete valued variables. Furthermore, one can discriminate algorithms which evaluate individual features and those which assess subsets of features (Hall and Holmes, 2003). As such, the Relief-F (Kononenko, 1994) approach was chosen because it enables to rank individual features. Additionally the Correlation-based Feature Selection (CFS) method was chosen, since it enables the scoring of the value of groups of features (Liu et al., 2002).

3.2.1. RELIEF-F

The focus of the Relief approach is to rank features according to how well their values enable the discrimination of cases that are near to each other. The underlying assumption is that a suitable feature should have different values for cases from different classes, and similar values for cases from the same class (Liu and Schumann, 2005). This principle is implemented by random sampling of an instance from the input data and the subsequent locating of its nearest neighbor from the same and opposite class. Values of the features of the nearest neighbors are then compared to the sampled instance and used to update relevance values for each feature. We used the enhanced approach "Relief-F" presented by Kononenko (1994), which allows handling of multi-class data (Robnik-Šikonja and Kononenko, 2003). Relief-F also smoothes the influence of noise in the data by averaging the contribution of k nearest neighbors of each sampled case (Hall and Holmes, 2003).

Primarily two parameters need to be defined when calculating relevance values: the number of cases m and the number of nearest neighbors k. According to Liu and Schumann (2005), m is also set to be the number of all training cases since a larger number of cases implies more reliable approximation, and k is set to 10, without weighting the nearest neighbors according to their distance. We tested several values for k (10, 20, 30, 40), the results of which showed little sensitivity to k values in this study, with only very moderate variations of the relevance value W(A). W(A) gives an indication in the interval [–1, 1] to what degree the respective feature is relevant. If $W(A) > 0$ then there is some degree of relevance, whereas features with a value $W(A) < 0$ are not relevant.

3.2.2. CORRELATION-BASED FEATURE SELECTION (CFS)

Under the premise that suitable groups of features contain variables highly correlated with the feature to be predicted, and are uncorrelated with each other, the CFS approach evaluates the value of subsets of features. Thereby, a matrix of feature-class and featurefeature correlations

from the training data is calculated first. The degree of intercorrelation between two features or the correlation between feature *X* and class *Y,* which is in the range of [0,1] is quantified as

$$r_{x,y} = 2.0 \times \left[\frac{gain}{H(X) + H(Y)}\right] \quad \text{(III-1)}$$

where $gain = H(X) + H(Y) - H(X,Y)$ represents the information gain between features and classes, and *H(X)* is the entropy of the feature (Liu et al., 2002). Subsequently, a score value is assigned by using a heuristic in the form of

$$Merit_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \quad \text{(III-2)}$$

where $Merit_s$ is the heuristic of a feature subset *S* with *k* features, $\overline{r_{cf}}$ represents the average feature-class correlation, and $\overline{r_{ff}}$ represents the average feature-feature intercorrelation. The numerator gives an indication of how predictive a set of features is and the denominator reveals the redundancy among them. Subsets that contain irrelevant features (with low feature-class correlation $\overline{r_{cf}}$) and features with a high redundancy (high feature-feature correlation $\overline{r_{ff}}$) are evaluated as unsuitable subsets of the feature space (Liu and Schumann, 2005). Since features are threatened independently, the CFS approach is not able to identify strongly interacting features. Nonetheless, empirical studies revealed that this method is able to identify useful features under moderate levels of interaction (Hall and Holmes, 2003). We used the greedy stepwise search, where the CFS approach starts from the empty set of features with a stopping criterion of five consecutive fully expanded non-improving subsets (preliminary experiments showed that the result of feature selection in this study is not sensitive with respect to the specification of the stopping criterion). Then the subset with the highest merit within an interval of [0, 1] revealed during the search, was selected (Liu et al., 2002).

3.3. REGRESSION ANALYSIS

To further explore the strength, direction and significance of the features derived from remote sensing data regarding the dependent variable (*in situ* assessed seismic building vulnerability), we deployed simple ordinary least squares (OLS) regression models for the scoring approach and generalized ordered logit regression models to examine the influence of the features on the respective EMS-98 classes. The basic linear regression model takes the form

$$y = a + \beta x + \varepsilon \quad \text{(III-3)}$$

with y being the depended variable, x the independent variable, α represents the intercept, β expresses the slope of the relationship between the two variables, and ε is an error term. Regarding the situation where there is more than one independent variable, the regression model is typically expressed as follows

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \varepsilon \quad \text{(III-4)}$$

where β_0 is the intercept and $\beta_1 - \beta_n$ represent the slope coefficients for the independent variables x_1 - x_n (see e.g. Kleinbaum et al., 1998 and Montgomery et al., 2001 for a thorough discussion of the simple and multi-linear case). The OLS method minimizes the sum of squared vertical deviations between the observed values in the data set and the values predicted by the linear approximation. Therefore, the regression line describes, as close as possible, the original values of the dependent variable.

When the dependent variable has a categorical character, the probability of falling into the categories 1 to i of the dependent variable with I categories is set in relation to the probability of falling in the categories $i+1$ to I:

$$Logit(Y_{1\ldots i/i+1\ldots I}|\mathrm{X}) = b_{0i} - b_1 X_1 - \ldots - b_k X_k \quad \text{(III-5)}$$

To take into account the ordinal character of the EMS-98 classification, a generalized ordered logistic regression model is used (see Williams, 2006 for a detailed description). The main advantage of generalized ordered logistic regression over ordinal logistic regression is that one does not have to make the assumption that the influence of a predictor is the same for each stage (proportional-odds or parallel-slopes assumption) (Peterson and Harrell, 1990). The applied partial proportional odds model tests the parallel assumption for each of the independent variables and calculates unique beta-coefficients for those violating the assumption. Generalized ordered logistic regression models are therefore less restrictive than ordinal regression models, but more convincing than multinomial logistic regression models. In ordered logistic regressions one single regression constant is calculated for each category of the dependent variable.

The regression analyses are solely applied with the features contained in the subsets with the highest merit as evaluated by the CFS approach. This group of features delivered the best results regarding the supervised regression and classification approaches that are described in the next section. Furthermore, we eliminated collinear variables. To track the composition of the models' performance the explanatory variables were grouped according to the feature categories in Tab. III-2 and entered into the regression equations in a stepwise, hierarchical way.

3.4. SUPERVISED REGRESSION AND CLASSIFICATION

To predict the respective scoring values and EMS-98 classes, we deploy supervised regression and classification techniques that are based on the delineation of functions from labeled training data. Each instance of the training data is constituted by a dependent variable and an *n*-dimensional vector of independent variable(s). The supervised learning scheme analyzes the training data and a generalized regression function (for continuous dependent variables) or classifier (for discrete dependent variables) is delineated to correctly estimate new examples.

3.4.1. SUPERVISED REGRESSION

For estimating the scoring values we compare the merits of multi-linear regression models and Support Vector Machine (SVM) based regression models which are able to represent non-linear boundaries between classes. As described above, multi-linear regression is based on the assumption that the dependent variable Y and its predictors $X_1, X_2, \ldots, X_n$ are directly related by a linear combination. Since linear regression models predict poorly in the presence of a nonlinear or non-additive relationship, a nonlinear Support Vector Regression (SVR) approach is additionally utilized. SVMs determine a suitable set of parameters that places a decision surface, the so called hyperplane, between the different classes of training samples according to their position in an *n*-dimensional feature space. The optimal separating hyperplane is identified as the maximized margin between the different classes and the hyperplane. In a modified form SVMs can also be applied for function estimation (see Smola and Schölkopf, 2004). Detailed theoretical background of SVMs is given in Vapnik (1995, 1998), Cortes and Vapnik (1995), and Burges (1998).

For the calculation of the regression models, the sample data were separated in a stratified manner to use one half of the samples for building the models, and the other half for validation. Regarding the linear regression approach, we excluded collinear features. As a measure of the relative goodness of fit, the Akaike information criterion was used for model selection. For the nonlinear approach we used the Sequential Minimal Optimization algorithm for regression as proposed by Smola and Schölkopf (1998) with a poly kernel and a regression optimizer introduced by Shevade et al. (2000). Regarding the comparison of the actual and estimated vulnerability values, we calculated a set of statistical accuracy measures: Mean Error (*ME*), Mean Percentage Error (*MPE*), Mean Absolute Error (*MAE*), Mean Absolute Percentage Error (*MAPE*), Standard Deviation (*StDev*), Relative Standard Deviation (*RStDev*), Pearson product-moment correlation coefficient (*R*), and *RMSE*.

3.4.2. SUPERVISED CLASSIFICATION

For estimating seismic vulnerability classes according to the EMS-98, we tested several supervised classification techniques (SVMs, radial basis function networks, backpropagation multilayer perceptrons, and random forests) and finally selected a simple naïve Bayes approach, since it outperformed the other classification techniques for a specific set of features. This has been evaluated with respect to the overall accuracy and Cohen's *kappa* statistic (Foody, 2002). The simple naïve Bayes approach represents a probabilistic classifier applying Bayes' theorem with strong underlying (naïve) independence assumptions. The basic assumption is the presence (respectively absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature, given the class variable. The aim is to assign an object I to one of a discrete set of categories C_1, C_2, …, C_m by using its observable features X_1, X_2, ..., X_n. The probability of I belonging to a respective category is calculated by applying Bayes' theorem. This is further reduced by assuming mutual conditional independence. Subsequently, I is assigned to the category with the greatest probability (Boyles et al., 2007). For a detailed description see e.g., Duda and Hart (1973) or Lewis (1998).

Analogous to the regression approach, the sample data were separated in a stratified manner, with half of the samples for building the classifier and the other half for its validation. As it can be seen in Tab. III-1, the data set used for training the classifier is imbalanced. This is when an uneven distribution of data patterns exists and the number of training instances of a majority class is much larger compared to other minority classes. Hence, the classifier is bias-prone and tends to favor the majority class (Nguyen et al., 2008). To increase the size of the minority classes and balance the class distribution, we oversampled the training data by using the Synthetic Minority Over-sampling Technique (SMOTE; see Chawla et al., 2002). Thereby, new instances are generated based on the "known" distribution to improve the generalization capacity of the learned classifier. The synthetic instances are added in the space between minority examples, emphasizing the class border in favor of the minority class. This principle is applied since the emphasis of class borders is useful in learning efficient discriminative classifiers (Cieslak et al., 2006). As mentioned, the classification outcomes are evaluated by calculating the overall accuracy and the *kappa* statistic. Additionally the user's (precision) and producer's (recall) accuracies as well as the receiver operating characteristics (ROC Area) for the individual classes are calculated.

4. RESULTS AND DISCUSSION

4.1. FEATURE SELECTION ANALYSIS

4.1.1. SCORING APPROACH

For the scoring approach the Relief-F algorithm reveals 70 features with a relevance value W(A) > 0. Due to limited space, only the first 21 most important features are presented what equals the number of features revealed by the CFS approach (Tab. III-3). One can observe similarities regarding the selected features. Most of the features that have a relevance value W(A) > 0 are included also in the CFS subset; exceptions include four spectral features of 1st order and two features of 2nd order. Furthermore the features "Roof Type $_B$", "Asymmetry $_B$" "Perimeter $_B$", Floor Number $_B$", "GLCM Homogeneity $_{S\ pan}$" and "GLCM Angular 2nd Moment $_{S\ pan}$" are merited as very influential by both algorithms, although they have completely different search heuristics.

From the 21 features ranked as most important by the Relief-F approach, 12 are related to buildings, whereas 9 are related to the block level. Thereby, the "Structure Type $_S$" is merited as most important. This gives an indication that the combined use of physical features is suitable to discriminate homogeneous urban areas that show similar vulnerability characteristics. This feature is followed by roof type and average building height per block. The building height can be considered as an important feature in general, since the floor number of the individual buildings is ranked 7th and represents a feature which is also considered relevant by the CFS approach. The subsequent features that are top-ranked primarily describe the geometry and extent of individual buildings. The merit of the best subset found from the CFS approach is 0.506. The CFS approach reveals a group of features that consists primarily of building level features (16), rather than of block level features (5). Analogous to the Relief-F results, features that characterize the geometry of the individual buildings are included in the set. The CFS subset also includes numerous features that are related to spectral information of 1st and 2nd order. In contrast to the Relief-F results, features that explicitly describe the spatial context were not included.

TAB. III-3. FEATURES REVEALED FOR THE VULNERABILITY SCORING APPROACH USING THE RELIEF-F AND CFS FEATURE SELECTION

Relief-F (Ranker)			CfsSubsetEval (Greedy Stepwise)	
Value	*Feature Name*	*Feature category*	*Feature Name*	*Feature category*
0.0435	Structure Type $_{S}$	IKONOS - Spatial context	Perimeter $_{B}$	IKONOS – Extent
0.0239	Roof Type $_{B}$	IKONOS - Shape	Asymmetry $_{B}$	IKONOS – Shape
0.0142	Avg. Building Height $_{S}$	DSM - 3D Features	Density $_{B}$	
0.0110	Width $_{B}$	IKONOS - Extent	Roof Type $_{B}$	
0.0109	Perimeter $_{B}$		M(1)/M(4) $_{B}$	IKONOS - Spectral Information (1st order)
0.0106	Dist. 3rd Neighbor Building $_{B}$	IKONOS - Spatial context	M(2)/M(3) $_{B}$	
0.0105	Floor Number $_{B}$	DSM - 3D Features	M(2)/M(4) $_{B}$	
0.0098	Length $_{B}$	IKONOS - Extent	M(1)/M(2) $_{B}$	
0.0076	Area $_{B}$		M(3)/M(4) $_{B}$	
0.0071	Dist. 2nd Neighbor Building $_{B}$	IKONOS - Spatial context	[M(4)-M(3)]/ [M(4) + M(3)] $_{B}$	
0.0069	GLCM Homogeneity $_{S\ pan}$	IKONOS - Spectral Information (2nd order)	M(2)/ [M(1)+M(2)+M(3)+ M(4)] $_{B}$	
0.0064	Degree of Building Density $_{S}$	IKONOS - Spatial context	M(2)/M(4) $_{S}$	
0.0062	Share of LC class "sealed" $_{S}$		M(2)/ [M(1)+M(2)+M(3)+ M(4)] $_{S}$	
0.0060	Dist. 1st Neighbor Building $_{B}$		[M(4)/M(3)]/ [M(1)+M(2)+Mean(3)+ M(4)] $_{S}$	
0.0059	Average Building Size $_{S}$		GLDV Angular 2nd Moment $_{B\ pan}$	IKONOS - Spectral Information (2nd order)
0.0056	Effective Area $_{B}$	DSM - 3D Features	GLDV Angular 2nd Moment $_{B\ nir}$	
0.0055	Area Building Block $_{S}$	IKONOS - Spatial context	GLCM Angular 2nd Moment $_{S\ pan}$	
0.0052	[M(3)-M(1)]/ [M(3)+M(1)] $_{S}$	IKONOS - Spectral Information (2nd order)	GLCM Homogeneity $_{B\ pan}$	
0.0045	Asymmetry $_{B}$	IKONOS - Shape	GLCM Homogeneity $_{B\ nir}$	
0.0042	GLCM Angular 2nd Moment $_{S\ pan}$	IKONOS - Spectral Information (2nd order)	GLCM Homogeneity $_{S\ pan}$	
0.0042	GLCM Standard Deviation $_{B\ pan}$		Floor Number $_{B}$	DSM - 3D Features

4.1.2. EMS-98 CLASSIFICATION

Regarding the EMS-98 data set, the Relief-F approach scored only eight features with a value W(A) < 0. The best CFS subset has a merit of 0.106 and consists of nine features (Tab. III-4), whereby these features all have a positive W(A) value. The features "Structure Type $_S$", "Floor Number $_B$", "Std. Dev. Blue(1) $_S$", "Compactness $_B$", "Share of Impervious Surface $_S$", "Std. Dev. Red(3) $_S$", which are part of CFS subset are also ranked as very influential by the Relief-F approach. The majority of features refer to the block level for both methods. Similar to the results of the scoring approach, the "Structure Type $_S$" is merited as most important by the Relief-F approach, followed by the estimated period of construction. Again, the height characteristics of the urban environment play an important role, since the individual height of buildings and the average building height per block are considered very influential. The CFS subset contains features related to geometry and spectral characteristics. Additionally, features which describe the spatial context where the buildings are embedded in, are given notable scores.

TAB. III-4. FEATURES REVEALED FOR THE EMS-98 CLASSIFICATION USING THE REFLIEF-F AND CFS APPROACH

Relief-F (Ranker)			CfsSubsetEval (Greedy Stepwise)	
Value	*Feature Name*	*Feature category*	*Feature Name*	*Feature category*
0.0589	Structure Type $_S$	IKONOS - Spatial context	Area $_B$	IKONOS – Extent
0.0212	Period of construction $_S$	LANDSAT – TM, ETM+	Compactness $_B$	IKONOS – Shape
0.015	Roof Type $_B$	IKONOS - Shape	Std. Dev. Blue(1) $_S$	IKONOS - Spectral Information (1st order)
0.0143	Avg. Building Height $_S$	DSM - 3D Features	Std. Dev. Red(3) $_S$	
0.0142	Main Direction $_B$	IKONOS - Shape	GLCM Entropy $_{B\,pan}$	IKONOS - Spectral Information (2nd order)
0.0132	GLCM Std. Dev. $_{S\,pan}$	IKONOS - Spectral Information (2nd order)	Average Building Size $_S$	IKONOS - Spatial context
0.0124	Floor Number $_B$	DSM - 3D Features	Share of Impervious Surface $_S$	
0.0113	[M(3)-M(1)]/[M(3)+M(1)] $_S$	IKONOS - Spectral Information (1st order)	Structure Type $_S$	
0.0112	GLCM Homogeneity $_{S\,pan}$	IKONOS - Spectral Information (2nd order)	Floor Number $_B$	DSM - 3D Features

4.2. REGRESSION ANALYSIS

4.2.1. SCORING APPROACH

Results from the OLS models (Tab. III-5) indicate that the building height and features that describe the geometry of individual buildings are the most important remotely sensed determinants of seismic building vulnerability, with significance influence at the 99.9 percent level. The combined use of these features results in an adjusted coefficient of determination of 0.254, which covers approximately 85% of the explained share of variance. The additional features related to 1st and 2nd order spectral information alone, can raise the model's R^2 value to 0.299. This finding is reasonable since the first mentioned features are reflected by comparable features part of the scoring value calculation itself, although for example, the geometry parameters are approximated mathematically in a completely different manner.

4.2.2. EMS-98 CLASSIFICATION

Tab. III-6 summarizes the results of the generalized ordered logistic model. It shows that having a higher value of "Area $_B$" increases the likelihood of being in class B-C or C-D-E, and therefore to be less vulnerable. Compact structures are more likely to be assigned to a class of higher vulnerability. In contrast, Low Income Structure Types (Structure Type $_S$ = Low Income) again have a higher chance of belonging to a class of lower vulnerability in comparison to slum structure types (note that no differences are observed between other structure types and the slum structure). Unique beta-coefficients are calculated for the influence of the "Floor Number $_B$" on the different EMS-98 classes. This indicates that the parallel assumption was not met for this variable. Having a higher "Floor Number $_B$" increases the chance of being in EMS-98 group A-B in comparison to the others. An even stronger negative effect of "Floor number $_B$" is found for the probability of belonging to EMS-98 class B-C. The latter finding is contrary to the effect of "Floor Number $_B$" regarding the scoring approach. However, this observation reflects oppositional assumptions associated with the different assessment approaches. Regarding the scoring approach, an increase in building height increases the value, implying lower vulnerability. Based on this it was assumed that high rise buildings represent engineered structures that are built in a more sophisticated way than informal (low rise) settlements. For the EMS-98 there were no such unidirectional assumptions made and the analysis results indicate that higher buildings are more likely to be more vulnerable. This is consistent with observations derived from damage data collected after the M_w = 7.6 event on September 30th, 2009 (Sengara et al., 2010).

TAB. III-5. RESULTS OF SIMPLE OLS MODELS REGARDING THE SCORING APPROACH (COEFFICIENTS AND STANDARD ERROR). THE GROUPED FEATURES WERE ENTERED INTO THE REGRESSION EQUATIONS IN A STEPWISE HIERARCHICAL WAY

Feature	*Model 1*		*Model 2*		*Model 3*		*Model 4*	
	Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
Floor Number $_B$	1.034508***	.1132338	.7293502***	.1144819	.6392977***	.11527	.5193538***	.1165116
Perimeter $_B$			.0090334***	.0024904	.0087243***	.0024959	.0085121***	.0025175
Asymmetry $_B$			3.31487***	.888291	3.354441***	.8640369	3.112394***	.8651691
Density $_B$			5.073578***	1.123404	5.063067***	1.081441	4.660547***	1.063173
Roof Type $_B$			.1545023	.2558068	.1839185	.2544025	.2649701	.2615001
GLCM Homogeneity $_{B\,pan}$					1.07571	1.315121	.4871516	1.336938
GLCM Homogeneity $_{B\,nir}$					1.350518*	.5738469	.7967434	.6515387
GLCM Angular 2nd Moment $_{S\,pan}$					396.783*	170.6297	190.7972	217.2341
GLCM Homogeneity $_{S\,pan}$					-11.61803	9.150705	-1.461658	10.16228
M(3)/M(4) $_B$							1.167996	1.689854
M(2)/[M(1)+M(2)+M(3)+M(4)] $_B$							12.98874*	6.32278
[M(4)/M(3)]/ [M(1)+M(2)+M(3)+M(4)] $_S$							-29.81346	105.0541
M(2)/M(4) $_S$							1.475905	1.412028
Number of observations	434		434		434		434	
adj. R²	0.165		0.254		0.273		0.299	

*P < 5%, **P < 1%, ***P < 0.1%

TAB. III-6. RESULTS OF GENERALIZED ORDERED LOGIT MODELS REGARDING THE EMS-98 CLASSIFICATION (COEFFICIENTS AND STANDARD ERROR). FIRST SECTION REFERS TO THE RELATION OF CLASS A-**B** AND C-**D**-E; SECOND SECTION REFERS TO THE RELATION OF CLASS B-**C** AND C-**D**-E (ONLY RELEVANT FOR "FLOOR NUMBER B"). THE GROUPED FEATURES WERE ENTERED INTO THE REGRESSION EQUATIONS IN A STEPWISE HIERARCHICAL WAY

Feature	*Model 1*		*Model 2*		*Model 3*		*Model 4*	
	Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
Floor Number $_B$	.4678125***	.0929906	.3403922***	.0980104	.2258272*	.1060161	.1886913+	.1086677
Area $_B$			.0006703***	.0001668	.0006324***	.0001679	.0005222*	.000228
Compactness $_B$			-1.604815***	.4134367	-1.44446***	.415815	-1.51715***	.4242106
Std. Dev. Blue(1) $_S$					.0344184**	.0128353	.0272497+	.0154555
GLCM Entropy $_{B\,pan}$					.0190522	.0723382	.0119984	.0725885
Average Building Size $_S$							.0010974	.0009604
Share of Impervious Surface $_S$							.0203762	.0211656
Structure Type $_S$= Suburb							.7969139	.5182687
Structure Type $_S$=Low Income							.9920697*	.4511381
Structure Type $_S$=Medium Income							.6634897	.4319753
Structure Type $_S$=High Income							.8390549	.703233
Floor Number $_B$	-.0793007	.1457927	-.2491776	.155034	-.3606772*	.1624929	-.3851584*	.1644611
Area $_B$			.0006703***	.0001668	.0006324***	.0001679	.0005222*	.000228
Compactness $_B$			-1.604815***	.4134367	-1.44446***	.415815	-1.51715***	.4242106
Std. Dev. Blue(1) $_S$					.0344184**	.0128353	.0272497+	.0154555
GLCM Entropy $_{B\,pan}$					.0190522	.0723382	.0119984	.0725885
Average Building Size $_S$							.0010974	.0009604
Share of Impervious Surface $_S$							.0203762	.0211656
Structure Type $_S$=Suburb							.7969139	.5182687
Structure Type $_S$=Low Income							.9920697*	.4511381
Structure Type $_S$=Medium Income							.6634897	.4319753
Structure Type $_S$=High Income							.8390549	.703233
Number of observations	434		434		434		434	

*P < 5%, **P < 1%, ***P < 0.1%

4.3. SUPERVISED REGRESSION AND CLASSIFICATION

4.3.1. SUPERVISED REGRESSION

The multi-linear and SVR models are calculated based on (i) the originate 132-dimensional feature vector, (ii) the 70 features with a positive degree of relevance as evaluated by the Relief-F algorithm, (iii) the group of 21 features with the highest merit as revealed by the CFS approach, and (iv) the further reduced set of features as used for the regression analysis (RA) (collinear features were removed from the subsets (i),(ii), and (iii) for the linear models). The evaluated model estimation results are given in Tab. III-7.

TAB. III-7. EVALUATED RESULTS OF THE MODEL PREDICTIONS OF DIFFERENT FEATURE SETS COMPARED TO REFERENCE VALUES

Multi-linear Regression								
Used features	*MAE*	*MAPE*	*ME*	*MPE*	*StDev*	*RStDev*	*R*	*RMSE*
All features	2.33	13.32	–0.56	–3.02	2.56	14.32	0.43	2.98
Relief-F W(A) > 0	2.02	11.84	–0.11	–0.58	2.07	11.29	0.48	2.62
CFS subset	1.84	10.99	–0.17	–0.91	1.52	8.33	0.56	2.41
CFS subset (RA)	1.80	10.81	–0.20	–1.10	1.45	7.96	0.58	2.37
Support Vector Regression								
Used features	*MAE*	*MAPE*	*ME*	*MPE*	*StDev*	*RStDev*	*R*	*RMSE*
All features	1.88	11.26	0.02	0.13	1.91	10.37	0.53	2.48
Relief-F W(A) > 0	1.79	11.07	0.43	2.31	1.72	9.13	0.57	2.42
CFS subset	1.72	10.61	0.23	1.25	1.54	8.28	0.59	2.36
CFS subset (RA)	1.73	10.67	0.24	1.31	1.63	8.73	0.59	2.36

When using the CFS feature set and the SVR approach the best results are achieved, with lowest *MAE* (1.72) / *MAPE* (10.61) and *RMSE* (2.36) and highest linear correlation ($R = 0.59$) of the model estimates. This demonstrates the viability of the approach. Furthermore, the nonlinear approach shows better predictions compared to the linear approach regarding the respective set of features used. The actual vulnerability values and the predicted values for the best model are visualized in a scatter plot (Fig. III-4). One can observe that the model overestimates low scoring values, buildings with a high seismic vulnerability – and underestimates high scoring values – characterizing buildings with a low seismic vulnerability. This observation is also true for all other models calculated.

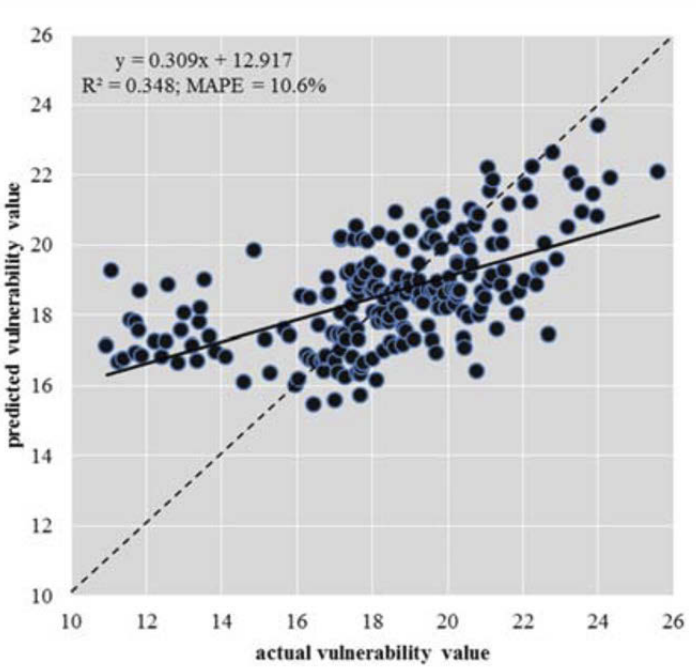


Fig. III-4. Scatter plot of the actual and predicted scoring values (SV) by using the CFS feature subset and the SVR approach.

4.3.1. SUPERVISED CLASSIFICATION

As mentioned above, the best result of all classification approaches in combination with different feature sets evaluated (analogous to the different feature sets used for building the supervised regression models), was achieved with the naïve Bayes classifier and the features revealed by the CFS approach. With this set of features only, the classifier performed considerably better than the other classification approaches evaluated. This is most likely attributable to the low pair-wise correlation of the features in this subset. The overall accuracy for the classification is 65.4% with a "fair" (Landis and Koch, 1977) *Kappa* statistic of 0.36. The classification performance for the respective EMS-98 classes is shown in Tab. III-8. One can find the classifications for the most vulnerable buildings according to the EMS-98 classification (class A-**B**) to be the most accurate with a classification accuracy of 0.696 (user's) and 0.793 (producer's), respectively. In contrast, the classification outcomes for the least vulnerable structures according to the EMS-98 classification (class C-**D**-E) are less feasible, with a user's accuracy of 0.294 and a producer's accuracy of 0.227, and the lowest ROC Area values. The strong confusion and poor predictive performance regarding class C-**D**-E is most likely caused by the small number of instances and high intra-class variability.

TAB. III-8. RESULTS FOR THE EMS-98 CLASSES OF THE SUPERVISED CLASSIFICATION APPROACH USING THE CFS SUBSET FEATURES

Class	User's Accuracy (Precision)	Producer's Accuracy (Recall)	ROC Area
A-**B**	0.696	0.793	0.666
B-**C**	0.661	0.554	0.686
C-**D**-E	0.294	0.227	0.601
Weighted Average	0.643	0.654	0.645

5. CONCLUSIONS

This study shows that an indirect correlation between physical information in the (urban) environment, drawn from remote sensing data and seismic vulnerability of buildings, exists. We demonstrated how to derive and identify meaningful features using a combination of remote sensing data and how to quantitatively evaluate their explanatory power. By means of a sequential procedure of calculating features from very high resolution multispectral data, height information and spatiotemporal analyses, and of applying machine learning approaches, the seismic vulnerability of buildings can be estimated with viable accuracies. Furthermore, we identified features that have high explanatory content and are most useful for the estimation of seismic building vulnerability, in terms of a preliminary screening. The features merited as most useful, their influence and direction vary considerably in dependency of the in situ seismic vulnerability assessment applied. However, features such as building height characteristics and features related to the geometry of the individual buildings turned out to generally explain seismic vulnerability to a significant degree.

We believe that this study generated insight into the capabilities of remote sensing for assessing seismic building vulnerability, and can be helpful for further empirical case studies in other parts of the world. Future research activities should comprise the systematic evaluation of remote sensing data according to wide-spread vulnerability assessment methods, such as the EMS-98, to make results comparable. In addition, the evaluation of the capabilities of remote sensing for estimating the seismic building structural type is a crucial task (see e.g., Sarabandi and Kiremidjian, 2007 for related work). Latest and future spaceborne missions, such as WorldView-2 and 3, CARTOSAT-3 or ALOS-3 have enhanced resolution characteristics, which will allow for the calculation of more sophisticated feature vectors (Novack et al., 2011). It is believed that this will entail more accurate predictive

performance of the models and further increase the applicability of remote sensing to assess the seismic vulnerability of buildings.

We conclude that remote sensing data and methods have a high capability to support large area assessments of building vulnerability, indicating the need for systematic application and validation of our findings. Lastly, we want to trigger an open dialogue between the remote sensing and earthquake engineering community, to better understand how remote sensing data can be linked to assessment approaches from engineering science in a robust, standardized and transferable way, to define common scales and enable systematic large-area assessments and monitoring of dynamic earthquake prone urban areas around the globe.

ACKNOWLEDGEMENTS

The research was founded partially by Helmholtz-EOS (Earth Observation System). The authors wish to thank the three anonymous reviewers for their very constructive comments and suggestions and Sarah Banks for her help with proofreading the manuscript.

Chapter IV

Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques

ISPRS Journal of Photogrammetry and Remote Sensing, forthcoming

Christian Geiß, Patrick Aravena Pelizari, Mattia Marconcini, Wayan Sengara, Mark Edwards, Tobia Lakes and Hannes Taubenböck

doi: 10.1016/j.isprsjprs.2014.07.016
Received 12th July 2013; revised 20th June 2014; accepted 30th July 2014

ABSTRACT

Detailed information about seismic building structural types (SBSTs) is crucial for accurate earthquake vulnerability and risk modeling as it reflects the main load-bearing structures of buildings and, thus, the behavior under seismic load. However, for numerous urban areas in earthquake prone regions this information is mostly outdated, unavailable, or simply not existent. To this purpose, we present an effective approach to estimate SBSTs by combining scarce *in situ* observations, multi-sensor remote sensing data and machine learning techniques. In particular, an approach is introduced, which deploys a sequential procedure comprising five main steps, namely calculation of features from remote sensing data, feature selection, outlier detection, generation of synthetic samples, and supervised classification under consideration of both Support Vector Machines and Random Forests. Experimental results obtained for a representative study area, including large parts of the city of Padang (Indonesia), assess the capabilities of the presented approach and confirm its great potential for a reliable area-wide estimation of SBSTs and an effective earthquake loss modeling based on remote sensing, which should be further explored in future research activities.

1. INTRODUCTION

Increasing spatial concentration of exposed elements such as people, buildings, infrastructure or economic values in earthquake prone regions induce seismic risk at an unprecedented high level. In particular, urban areas in developing countries are characterized by a large amount of vulnerable buildings. At the same time, a very dynamic urban growth is accompanied by the construction of unplanned, spontaneous and highly vulnerable settlements. Thus, local governments and stakeholders face the problem of continuously updating their knowledge on the current building stock and simultaneously assessing exposed buildings area-wide to efficiently establish and adjust preparedness measures (Sarabandi and Kiremidjian, 2007; Taubenböck et al., 2009a; Wieland et al., 2012). Especially for earthquake loss estimation (ELE) modeling, the gathering of building inventory and vulnerability information represents normally the most time-consuming and expensive aspect (Dunbar et al., 2003).

The exclusive application of conventional approaches such as detailed *in situ* building-by-building analysis by structural engineers is decreasingly able to cope with this situation. Instead, in the last few years remote sensing has proven its great potential to extract relevant features for pre-event vulnerability analysis of built-up structures for large areas (Geiß and Taubenböck, 2012). So far, different approaches have been presented in the literature. By means of characteristics extracted from remote sensing data, Taubenböck et al. (2009a) and Borzi et al. (2011) reconstruct and characterize the built environment and retrieve specific fragility functions for designated building types. Pittore and Wieland (2012) use remote sensing data for delineating and characterizing homogeneous built-up areas. The vulnerability of the building inventory is determined in combination with information from a ground-based omnidirectional imaging system. Similarly, Borfecchia et al. (2009) assess the vulnerability of buildings in a hybrid way, namely by combining *in situ* ground truth for selected buildings with information derived from remote sensing data. Supervised classification techniques are subsequently used to classify the residual building inventory. Geiß et al. (2013) combine detailed *in situ* seismic vulnerability information with features describing the urban morphology derived from remote sensing data. Supervised regression and classification techniques are then applied to evaluate the suitability for an area-wide assessment. The aforementioned studies deploy very heterogeneous approaches with respect to the vulnerability levels or classes to be estimated. Taubenböck et al. (2009a), Borzi et al. (2011), and Borfecchia et al. (2009) use rather specific definitions, whereas Pittore and Wieland (2012) and Geiß et al. (2013) incorporate also more transferable, yet generalized, assessment

schemes, such as the European Macroseismic Scale (EMS-98; Grünthal et al. 1998). However, none of the cited studies focus on the estimation of seismic building structural types (SBSTs). SBSTs characterize the main load-bearing structure of a building. This is the most affecting factor for earthquake damage and, accordingly, it is generally the first property considered for categorizing a building. Further frequently considered parameters that may reflect the seismic performance also comprise the number of storeys, the period of construction or the presence of structural irregularities (Coburn and Spence, 2002). A function for individual SBSTs can be determined that relates the magnitude of the seismic hazard to the damage probability of the structures (Calvi et al., 2006). This enables the prediction of the probable damage distribution of the building inventory with respect to a certain level of seismic hazard (Douglas, 2007). Additionally, SBSTs can also contribute to the assessment of the seismic vulnerability according to schemes such as the EMS-98.

In the pioneering work of Sarabandi and Kiremidjian (2007), information derived from remote sensing data is combined with ancillary (geo-)information to estimate SBSTs. In particular, they use very high resolution optical imagery to derive the building inventory and calculate features describing the height, extent, shape, and roof type characteristics of individual buildings. In addition, they use tax assessor data to compile information about occupancy and age. Subsequent to that, supervised classification techniques (Classification and Regression Trees (CART) and multinomial logistic regression) are deployed to estimate SBSTs. In this paper we propose considerable conceptual and methodological differences to estimate SBSTs. The plethora of sensors systems that provide useful and complementary information yields the possibility to substitute ancillary (geo-)information and, thus, fully rely on remote sensing to reconstruct and characterize the building inventory. Due to e.g. data availability it may be crucial to gain independence from proprietary sources of information (e.g. tax assessor data). In addition, a complementary set of remote sensing data allows to characterize the building inventory in an exhaustive manner and to encode for instance also spatial context information in the classifier. This in turn opens a good opportunity to boost predictive performance of learned models. An exhaustive characterization of the building inventory based on a comprehensive set of features simultaneously suggests relying on classification approaches that are able to cope efficiently with high-dimensional data sets. Moreover, SBSTs ground truth is very costly to obtain and at the same time is afflicted with uncertainties induced by an often challenging assignment process. This induces the general need for a more tailored approach, which is able to lower those uncertainties and can cope with the scarcity of *in situ* elaborated ground truth.

To address these considerations, the objective of this paper is to introduce an approach for estimating SBSTs area-wide based on scarce *in situ* ground truth and complementary multi-sensor remote sensing data by means of a sequential procedure of advanced machine learning techniques. More specifically, we exploit very high resolution multispectral imagery, multi-temporal medium resolution multispectral data, as well as height information from a normalized digital surface model (nDSM) to derive a comprehensive set of features characterizing the urban environment. Different feature selection techniques are then employed to reduce the dimensionality of the resulting dataset and identify the most relevant features. Outlier detection is applied to prune those objects from the data for which the available *in situ* information cannot be considered reliable. To tackle the scarcity of SBSTs ground-truth data, additional synthetic samples are generated. Finally, different SBSTs are estimated by means of advanced supervised classification techniques. In particular, both Support Vector Machines (SVM) (Vapnik, 1998; Schölkopf and Smola, 2002) and Random Forests (RF) (Breiman, 2001) are considered due to their capability of effectively handling complex remote sensing classification problems (Camps-Valls and Bruzzone, 2009; Gislason, 2006). Since spatially distributed estimation of SBSTs represents a critical input for ELE models, we illustrate the applicability of the presented approach within scenario-based loss estimations for the city of Padang, Indonesia.

The remainder of the paper is organized as follows. Section 2 characterizes the study site and data base for this study. In section 3 the methods are described and results and discussions are revealed in section 4. The paper is concluded and an outlook is given in section 5.

2. STUDY SITE AND DATA

2.1. STUDY SITE: PADANG, INDONESIA

The presented study focuses on the city of Padang (Indonesia), which is situated in one of the most earthquake-prone regions worldwide. Padang is located on the island of Sumatra (mainly on the coast and to some extent sited beneath the mean sea level) and is the capital city of the Sumatera Barat province. It represents the third largest city on the island with approximately one million inhabitants. The dynamic urban system of Padang is constituted by a high concentration of population, infrastructure and economic values. The city has supra-regional relevance with an international airport and railway connection and possesses an essential economical role for the coastal region and the mountainous hinterland.

The Sumatra subduction zone represents one of the most active plate tectonic margins in the world (Petersen et al., 2004). The Australian plate plunges beneath the Sunda block of the Eurasian plate with convergence rates between ~56 and 62 mm/yr. (Chlieh et al., 2008) (Fig. IV-1a). The associated complex plate boundary setting leads to thrust earthquakes on the subduction fault, strike-slip earthquakes on the Sumatran fault, deeper earthquakes within the subducting lithosphere, and volcanic earthquakes (McCaffrey, 2009). Accordingly, the city is located in a region characterized by extremely high probability of severe earthquakes, as well as secondary effects such as tsunamis (Chlieh et al., 2008; Taubenböck et al., 2009b). As an example, in the afternoon of 30th September 2009, Padang was hit by an earthquake with a moment magnitude of M_w = 7.6 (Taubenböck et al., 2013). Overall, the earthquake event affected an area with a population of 1.2 million and caused 1,195 fatalities. 144,000 buildings collapsed or were significantly damaged. In Padang, 383 people died and 431 were seriously injured, primarily due to collapsing buildings (EERI, 2009; BNPB, 2009). Despite the size of the event, the Sunda megathrust was not ruptured and the stress on the Mentawai segment, which was accumulated over 200 years, has not been significantly reduced. The megathrust strain-energy budget remained substantially at a high level and the threat of a great, also tsunamigenic earthquake with a magnitude M_w > 8.5 on the Mentawai patch is unabated (McCloskey et al., 2010).

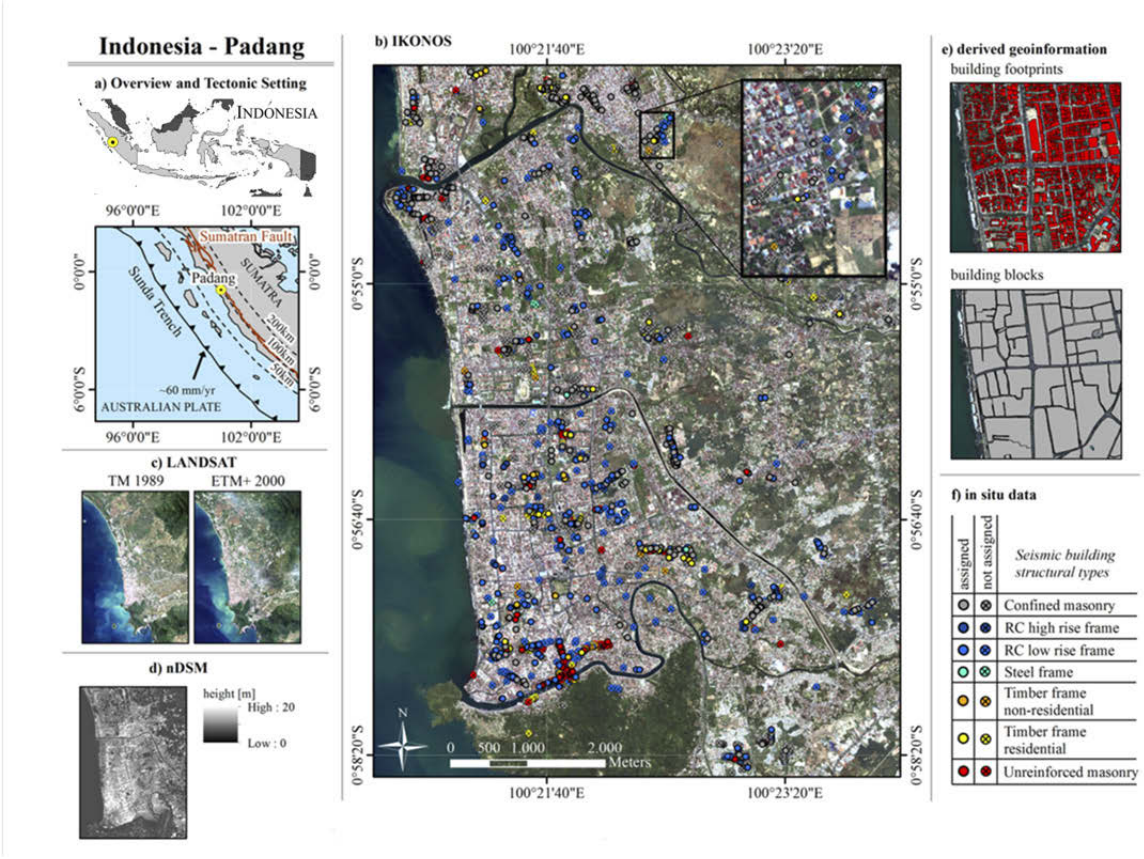


Fig. IV-1. Overview on the location of the study area and acquired data. (a) overview on the location of Padang and tectonic setting; basic active structural elements of the obliquely convergent Sumatran plate boundary are shown; dashed lines parallel to the trench are the 50, 100, and 200km depth of the megathrust (Chlieh et al., 2008); (b) very high resolution multispectral IKONOS imagery; (c) multitemporal LANDSAT data; (d) nDSM with object heights; (e) derived geoinformation consisting of building footprints and buildings blocks; (f) *in situ* data regarding SBSTs that are superimposed on (b).

2.2. REMOTE SENSING DATA

The remote sensing data used in our study have been acquired and processed within the "Last-Mile" project (Taubenböck et al., 2009b) and comprise a multispectral IKONOS image, multitemporal LANDSAT data, as well as height information from a nDSM (all of them co-registered to UTM 47S projection and WGS-84 datum). The multispectral IKONOS image was acquired on 12th April 2005 and covers a spectral range between 0.445 and 0.853 µm, with a geometric resolution of 1 m for the panchromatic band and 4 m for the 4 multispectral bands. The data were pan-sharpened and atmospheric correction was performed using the ATCOR (Atmospheric and Topographic Correction) model (Richter, 1996; Taubenböck et al., 2009b) (Fig. IV-1b). Multitemporal LANDSAT data were acquired by the Thematic Mapper sensor and the Enhanced Thematic Mapper sensor on 25th July 1989 and 15th July 2000, respectively. Both images are characterized by 7 multispectral bands covering a spectral range between 0.45 and 2.35 µm at 30 m spatial resolution (Fig. IV-1c). Height information is

derived by means of a digital surface model (DSM) and a digital terrain model (DTM), both derived from airborne radar data acquired by pair antennas and processed using SAR interferometry techniques (Li et al., 2004). The DTM was derived based on measurements of the bare ground contained in the original radar data and by manually reviewing and editing (Intermap, 2010). The data sets have a geometric resolution of 5 m and a height Root-mean-square-error (RMSE) of 1 m. To get relative height information of elevated objects a nDSM is calculated by subtracting the height values of the DTM from the height values of the DSM (Fig. IV-1d). Interested readers can refer to Taubenböck et al. (2009b) for a more detailed description of data acquisition and preprocessing.

2.3. DERIVED GEO-INFORMATION AND IN SITU DATA

Basic geo-information had already been derived from the remote sensing data and provided for this study. Within the "Last-Mile" project 87,573 building footprints were digitized from the IKONOS imagery by means of a manual photointerpretation procedure. They represent the core of Padang's building inventory. Additionally, building blocks had been derived from a closed-meshed road network (Taubenböck et al., 2008) (Fig. IV-1e). Both information layers serve as basis for the calculation of features, which is explained in section 3.1.

About four weeks after the earthquake on 30th September 2009 a field survey in the affected region took place in the framework of the Australia-Indonesia Facility for Disaster Reduction (AIFDR) jointly led by the Institut Teknologi Bandung (ITB) and Geoscience Australia. The primary objective of the survey was to undertake a population based inspection of buildings of all types and all damage levels. The results allowed inferring knowledge regarding the vulnerability of a range of building types present in the surveyed region and representative for others in Indonesia (Sengara et al., 2010). Overall, 3896 buildings were surveyed and each of them was assigned to a specific structural system, wall type, roofing type, floor type, number of storeys, usage, and the degree of damage suffered from the earthquake event. To conduct a vulnerability assessment and to derive fragility curves, the surveyed buildings were categorized according to SBSTs, which reflect a similar behavior under seismic load. In particular, the following classes were considered: "Confined masonry" (*CM*), "Reinforced concrete high" (*RC high*), "Reinforced concrete low" (*RC low*), "Steel frame" (*SF*), "Timber frame residential" (*TF res*), "Timber frame non-residential" (*TF non-res*), and "Unreinforced masonry" (*URM*) (Fig. IV-1f). From the whole amount of surveyed buildings, 2779 are located in the study area. The position of each building was recorded with a GPS device (Fig. IV-1b) and digital pictures were also taken (ibid.). Nevertheless, due to

inaccuracies in the GPS positioning, only 561 buildings could be unambiguously assigned to their corresponding building footprint extracted from remote sensing imagery (Fig. IV-1e). Unfortunately, only two samples remained for the structural type *SF*. As this class represents a relatively rare but striking SBST, the corresponding *in situ* ground truth was extended with 12 additional samples derived from another data set compiled in February/March 2008 during the "Last-Mile" project (Taubenböck et al., 2009b). The histogram depicting frequencies of different SBSTs of the final *in situ* data set is shown in Fig. IV-2. Descriptive statistical analyses were carried out to check whether the final *in situ* data set is consistent with all surveyed buildings and the results revealed a very good agreement.

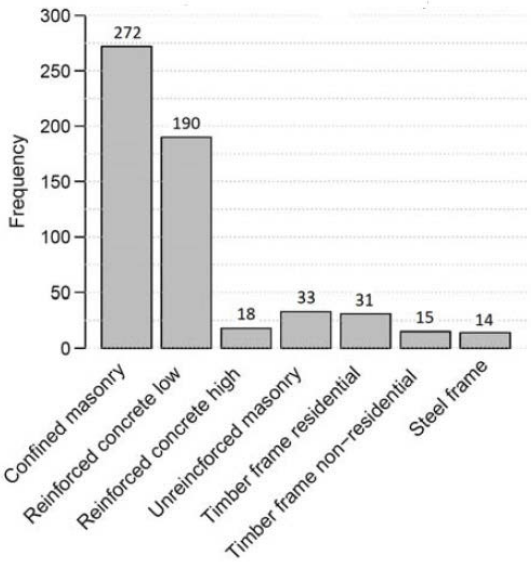


Fig. IV-2. Frequency of labeled samples (overall: 573) according to different SBSTs of the final *in situ* data set.

3. METHODS

Based on the remote sensing and *in situ* data, we carry out a sequential procedure to estimate SBSTs. Fig. IV-3 gives a schematic overview from the data sets used, the chronology of the procedure to the targeted SBSTs classification map. A set of features is derived from the remote sensing data at two different spatial levels, building and block level (section 3.1.). The hierarchical supervised classification approach is described in section 3.2. Outliers in the *in situ* data and building inventory are identified first. Therefore, a subset based feature selection technique (section 3.2.1.) is used to create a suitable group of features for building robust one-class classification models based on the *in situ* data. The models are built by means of a one-class support vector machine (OC-SVM, section 3.2.2.) approach and are applied on both *in situ* data and building inventory. Subsequent to outlier identification, multiclass classification

models are built in three consecutive steps. The remaining *in situ* samples are used to identify useful groups of features for building robust models by applying subset and ranker based feature selection techniques (section 3.2.1.). To tackle scarcity of the *in situ* data and learn efficient discriminative classifiers, synthetic training samples are generated by means of an oversampling technique (section 3.2.3.). Based on the generated feature groups and oversampled training data, multiclass classification models are learned by using SVM and RF (section 3.2.4.). Finally, the most accurate model is applied on the building inventory to estimate SBSTs spatially distributed.

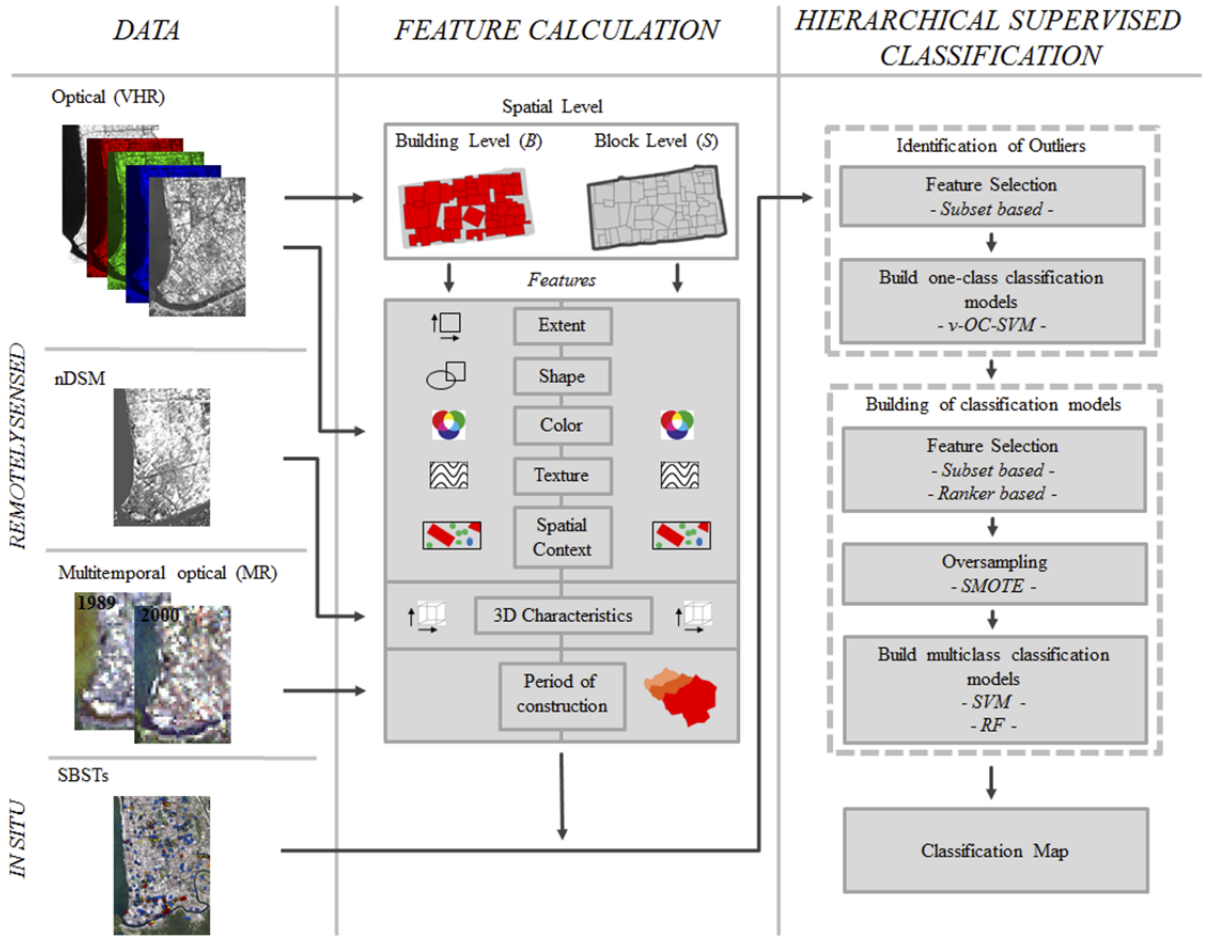


Fig. IV-3. Overview of the framework and processing steps followed in this study. Detailed explanations are given throughout the text in section 3.1 which describes the feature calculation, and section 3.2 which describes the hierarchical supervised classification approach.

3.1. CALCULATION OF FEATURES FROM REMOTE SENSING DATA

For a reliable estimation of SBSTs, numerous features have been extracted for both the aforementioned individual building footprints and building blocks (identified in the following by the subscripts *B* and *S*, respectively) (Fig. IV-1e and Fig. IV-3). Generally, the building

footprints allow characterizing individual buildings, whereas the building block layer characterizes the spatial setting which the respective buildings are embedded in. We chose to use building blocks derived from a street network rather than artificial spatial units, such as quadratic objects. This allows us to reflect the urban morphology, which is constituted by distinct areas that are generally irregularly shaped, more naturally. Simultaneously, the difficulty of having to determine the optimal kernel size *a priori* is avoided (Herold et al., 2003).

We use a set of features, that was introduced and explained in detail in Geiß et al. (2013), where it was used to evaluate the potential of remote sensing to assess the seismic vulnerability levels of buildings (see Tab. IV-1a). In particular, the features relate to the two-dimensional extent of buildings as well as the description of their shape characteristics. In addition, statistical values of 1st and 2nd order were extracted from the available IKONOS imagery at building and block level. The first serve as a descriptor of roof surface material and arrangement whereas the latter are intended to describe the composition of distinct urban structures. Mean and standard deviation values of the different image bands as well as band ratios, which are intended to emphasize spectral dissimilarities, were calculated. Additionally, rotation-invariant texture measures for the panchromatic and near-infrared band were computed using both the co-occurrence matrix (GLCM) and grey level difference vector (GLDV).

Features explicitly aiming to describe the spatial context are calculated at block level and consist of the area of building blocks and the average size of the buildings located within. Furthermore, spatial metrics such as proportion measures of land cover classes are computed. Based on a urban land cover map derived in Taubenböck et al. (2009b) (which exhibited an OA of 97%), proportions of land cover classes “buildings”, “sealed”, “grass/meadow”, “trees”, and “impervious surface”, which represents a combination of “buildings” and “sealed” were calculated per block. Additionally, a semantic classification (“Structure Type s”), which is built on physical features that describe the urban morphology, is incorporated. The classification describes the socio-economic status of the population by distinguishing “slums”, “suburbs”, “low income areas”, “medium income areas”, and “high income areas”. Beyond, the incorporation of height information allows the calculation of 3D features such as building floor number, floor space, ratio of diameter and height, ratio of width and height, as well the average building height within a building block. The mean slope for each building block was calculated to describe topographic location characteristics. By analyzing two Landsat images from 1989 and 2000, the period of construction is approximately described

based on a post classification change detection procedure, which aims to map the urban extent at the respective time step. For a more comprehensive description of all the features listed in Tab. IV-1a, the reader is referred to Geiß et al. (2013).

In addition, we introduce a number of new features in this study (Tab. IV-1b). They characterize the buildings' two-dimensional extent, shape and spatial context. They were chosen since they turned out to be beneficial in previous studies for discriminating different morphologically homogeneous urban structures based on geospatial data (Steiniger et al., 2008; Colaninno et al., 2011). Overall, each building object is represented by a 145-dimensional feature vector, whereby 79 features are calculated based on the individual building footprints, and 66 are calculated based on the building blocks. In this manner, a perceptual coherence (Steiniger et al., 2008) of physical appearance, spatial composition and context, and temporal development of the urban morphology and the main load-bearing structure of buildings is assumed.

TAB. IV-1. FEATURES DERIVED FROM REMOTE SENSING DATA TO CHARACTERIZE THE URBAN MORPHOLOGY; A) SET OF FEATURES EXTRACTED AS DESCRIBED IN GEIß ET AL. (2013); B) ADDITIONAL FEATURES THAT ARE RELATED TO THE EXTENT AND SHAPE OF INDIVIDUAL BUILDINGS AND THE SPATIAL CONTEXT THEY ARE EMBEDDED IN

FEATURES

a)

IKONOS Extent/Shape	IKONOS Spectral 1st order	IKONOS Spectral 2nd order	IKONOS Spatial context
Area $_B$	Mean Max. Diff. $_{B,S}$	GLCM Angular 2nd Moment $_{B\,pan\,nir,\ S\,pan\,nir}$	Area Building Block $_S$
Perimeter $_B$	Mean Brightness $_{B,S}$	GLCM Contrast $_{B\,pan\,nir,\ S\,pan\,nir}$	Average Building Size $_S$
Width $_B$	Mean Blue(1) $_{B,S}$	GLCM Dissimilarity $_{B\,pan\,nir,\ S\,pan\,nir}$	Degree of Building Density $_S$
Length $_B$	Mean Green(2) $_{B,S}$	GLCM Std. Dev. $_{B\,pan\,nir,\ S\,pan\,nir}$	Share of LC class "sealed" $_S$
Length/Width $_B$	Mean Red(3) $_{B,S}$	GLCM Mean $_{B\,pan\,nir,\ S\,pan\,nir}$	Share of LC class "grass/meadow" $_S$
Asymmetry $_B$	Mean Nir(4) $_{B,S}$	GLCM Homogeneity $_{B\,pan\,nir,\ S\,pan\,nir}$	Share of LC class "tree(s)" $_S$
Elliptic Fit $_B$	StDev. Blue(1) $_{B,S}$	GLCM Entropy $_{B\,pan\,nir,\ S\,pan\,nir}$	Share of Impervious Surface $_S$
Radius of Smallest Enclosing Ellipse $_B$	StDev. Green(2) $_{B,S}$	GLCM Correlation $_{B\,pan\,nir,\ S\,pan\,nir}$	Structure Type $_S$
Radius of Largest Enclosing Ellipse $_B$	StDev. Red(3) $_{B,S}$	GLDV Angular 2nd Moment $_{B\,pan\,nir,\ S\,pan\,nir}$	
Rectangular Fit $_B$	StDev. Nir(4) $_{B,S}$	GLDV Contrast $_{B\,pan\,nir,\ S\,pan\,nir}$	
Roundness $_B$	[M(4)-M(3)]/[M(4)+M(3)] $_{B,S}$	GLDV Mean $_{B\,pan\,nir,\ S\,pan\,nir}$	
Density $_B$	Brightness×M(1)/M(2) $_{B,S}$	GLDV Entropy $_{B\,pan\,nir,\ S\,pan\,nir}$	
Main Direction $_B$	Brightness×M(1)/M(3) $_{B,S}$		
Border Index $_B$	Brightness×M(2)/M(3) $_{B,S}$	**nDSM 3D**	**LANDSAT temporal**
Shape Index $_B$	M(1)/M(2) $_{B,S}$	Floor Number $_B$	Period of construction $_S$
Compactness $_B$	M(1)/M(3) $_{B,S}$	Floor Space $_B$	
	M(1)/M(4) $_{B,S}$	Diameter/Height $_B$	
	M(2)/M(3) $_{B,S}$	Width/Height $_B$	
	M(2)/M(4) $_{B,S}$	Average Building Height $_S$	
	M(3)/M(4) $_{B,S}$	Slope $_S$	
	M(2)/[M(1)+M(2)+M(3)+M(4)] $_{B,S}$		
	[M(4)/M(3)]/ [M(1)+M(2)+M(3)+M(4)] $_{B,S}$		
	[M(3)-M(2)]/[M(3)+M(2)] $_{B,S}$		
	[M(3)-M(1)]/[M(3)+M(1)] $_{B,S}$		

b)

IKONOS Extent/Shape	Description	IKONOS Spatial context	Description
Area/Perimeter $_B$	Area to perimeter ratio	Area/Perimeter $_S$	See description at "IKONOS Extend/Shape"; Calculated for Block Level
Shape Index II $_B$	Schumm's longest axis to area ratio (MacEachren 1985)	Shape Index II $_S$	
Shape Index III $_B$	Building perimeter to perimeter of a circle with the same area ratio; Gravelius's compactness measure (Gravelius 1914)	Shape Index III $_S$	
Building Elongation $_B$	Length to width ratio of the minimum bounding rectangle (Steiniger et al. 2008)	Building Elongation $_S$	
Building Orientation $_B$	Orientation of the major axis of the minimum bounding rectangle	Building Orientation $_S$	Homogeneity of alignment
		Proximity Sum $_{B,S}$	Summed/Averaged distance to 5 nearest neighbors
		Proximity Avg. $_{B,S}$	
		No. Buildings Buffer 100m $_B$	Count of all buildings intersected by a buffer with 100m and 200m distance (Steiniger et al. 2008)
		No. Buildings Buffer 200m $_B$	
		Area to Buffer Area 100m $_B$	Area to buffer area ratio for a buffer with 100m and 200m distance (Steiniger et al. 2008)
		Area to Buffer Area 200m $_B$	

Subscripts: B = Building Level, S = Block Level, Pan = Panchromatic, Nir = Near-Infrared; M = Mean Value

3.2. HIERARCHICAL SUPERVISED CLASSIFICATION APPROACH

3.2.1. FEATURE SELECTION WITH FILTERS

For taking into account the high dimensionality of the input data, several feature selection algorithms were deployed. The selection of features to be used for building robust classification models is generally a difficult task, especially when dealing with a large number of features as it is the case in this study. Thereby, high-dimensional feature vectors often exhibit redundancy, show inter-correlations and may suffer from the "Hughes phenomenon"

(which states that for a limited amount of samples the predictive power decreases as the dimensionality of the feature vector increases (Hughes, 1968)). Furthermore, multivariate classification methods can be prone to over-fitting. Filtering out the least promising features and thus reducing the dimensionality of the feature vector may attenuate the aforementioned problems (Guyon, 2003).

Feature selection techniques can be categorized into filters and wrappers. That latter evaluate features by using accuracy estimates provided by the actual classification algorithm, which is deployed subsequent to feature selection. Thus, the classifier needs to be trained and accuracy estimation needs to be performed for each iteration of the evaluation process. This can lead to large processing times what make them unpractical for extensive studies on high-dimensional data sets (Kohavi and John, 1997). Accordingly, in our study we employed filter methods for feature selection, as they operate independently with respect to the classifier. Among filter methods one can discriminate algorithms which evaluate individual features and those which assess subsets of features (Hall and Holmes, 2003). Overall, six different filter methods were applied in the context of the presented work (by means of the WEKA software environment (Hall et al., 2009)), namely Information Gain (*IG*), Gain Ratio (*GR*), Chi-Squared (χ^2), Pearson product-moment correlation coefficient (*COR*), and Relief-F (*RelF*), which all allow feature ranking, as well as the Correlation-based Feature Selection (*CFS*) which allows evaluating feature groups.

IG belongs to the group of information theory indices, which evaluates a feature by means of the Shannon entropy. The decrease in entropy of a class is evaluated when a feature is considered (Duch, 2006). A drawback of the *IG* measure is that it tends to be biased towards features with a large number of distinct values. *GR*, which is a largely-employed modification of *IG*, aims to avoid this bias by a simple normalization utilizing the class entropy (Van Hulse, 2009). The χ^2 method deploys the χ^2 statistic to evaluate the strength of the relationship between each independent feature and a class (ibid.). *COR* measures the relation of a feature and a class based on Pearson's Correlation coefficient (Hall et al., 2009) and was used as a benchmark in this study. Focusing on the expectation that useful features should differentiate between instances from different classes and have similar values for instances from the same class, *RelF* (Kononenko, 1994; Robnik-Šikonja and Kononenko, 2003) ranks features according to their ability to discriminate between neighboring instances. For multiclass problems *RelF* randomly samples an instance from the data and locates its *k* nearest neighbors from the same and different classes. The feature values of the nearest neighbors are compared to the sampled instance and used to up-date relevance scores for each feature. This procedure

is repeated for a number of instances *m* that has to be specified by the user (Hall and Holmes, 2003). The *CFS* method (Hall, 1999) uses a best first search algorithm to identify a group of possibly suitable subsets. Each considered subset is subsequently merited by means of an entropy based heuristic. Subsets with high feature-class correlation and low feature-feature inter-correlation are merited best (Hall and Holmes, 2003). *IG*, *GR*, χ^2, and *COR* are bivariate procedures. Each feature is evaluated independently of all other features in the data set. This is in contrast to the procedure of the *RelF* and *CFS* approach.

For feature selection methods that require discrete values (i.e. *IG*, *GR*, χ^2), numerical features were discretized according to the method of Fayyad and Irani (1993). Concerning the *RelF* approach, we set *m* to the number of all instances, since a larger value of *m* implies a more reliable approximation (Kononenko, 1994). We tested several values for the number of nearest neighbors to be considered. However, the results are hardly sensitive with respect to this parameter in this study, thus the number of neighbors *k* was set to 10. For the *CFS* method we used a stopping criterion for the search heuristic of 5 consecutive fully expanded non-improving subsets. The subset with the highest merit revealed during the search was selected. The results of the feature selection algorithms served for the creation of feature subsets that were used for building several multiclass classification models (section 3.2.4). 14 subsets containing the *n* "best ranked" features (*n* = 5, 10, …, 50, 60, 80, 100, 120) were created based on the results of each ranker method. In addition, one feature subset was built according to the result from the *CFS* technique (*n* = 28). Altogether this leads to a total amount of 72 data sets, one containing the original number of features and 71 represent feature reduced subsets.

3.2.2. OUTLIER DETECTION WITH OC-SVM

Outlier detection is applied to exclude objects from the data for which the available *in situ* information cannot be considered reliable. The appearance of outliers can be related to several sources. The set of *in situ* samples may not be fully representative for all buildings in the area and the SBSTs, as defined in the context of the *in situ* survey, may not cover all structures. Inaccuracies may have been occurred in the surveying of single buildings, leading to the assignment of inappropriate labels. Mislabeling also cannot be excluded in the assignment process of the *in situ* data. Beyond, feature values of building objects can be biased due to errors or noise in the underlying remote sensing data.

To lower those uncertainties we employ *v*-OC-SVM, which were introduced by Schölkopf et al. (1999) as support vector method for novelty detection. It is based on the general

principle of SVM, which determine appropriate parameters that construct a decision surface, the optimal separating hyperplane, between the classes of training samples with respect to their position in an *n*-dimensional feature space (Vapnik, 1995, 1998). With the benefits of SVM, the one class classifier is able to capture the support region (i.e., where the density is large) without the need of prior assumptions about the distribution of the data. Therefore, the target class is described by a function that maps the majority of instances to a region where the function is nonzero. To achieve this, the origin of the feature space is first treated as the only available member of the non-target class (i.e., as an outlier). Then, a hyperplane with maximum margin separation from the origin is identified. Analogous to the usual SVM framework, outliers in the training data are handled by *slack variables* (Muñoz-Marí et al., 2010).

To keep computational costs low and mitigate over-fitting, we identified a group of valuable features for the one-class models by using the *CFS* technique (section 3.2.1.). The *in situ* data with the reduced set of features were split according to the SBSTs class labels. Subsequently, for each of them we determine a specific *v*-OC-SVM model (all support vector methods in this paper were carried out with the LibSVM package by Chang and Lin (2011)). We used Gaussian radial basis function (RBF) kernels as commonly used in literature when addressing environmental applications since they showed generally good performances in plenty of studies (e.g., Camps-Valls and Bruzzone, 2009; Volpi et al., 2013). The application of the *v*-OC-SVM approach with a RBF kernel requires adjusting the parameter *v,* which represents the expected percentage of outliers in the training set and the kernel-width parameter γ. However, it is difficult to tune free parameters if only target labeled samples are available in the training data. In such situations solely the true positive rate (sensivity) can be calculated, whereas the error counterpart (specificity) cannot. To overcome this limitation, the free parameter selection was determined by evaluating $\arg\max_{\theta}\left\{\frac{\mathrm{OA}[\%]}{\#\mathrm{SV}}\right\}$, where θ is the set of free parameters (i.e., v and γ), OA is the overall accuracy and *#SV* the number of support vectors. This heuristic enforces high OA while simultaneously limiting model complexity keeping a low number of *SV* (Muñoz-Marí et al., 2010). For both v and γ we performed a grid search varying v in the range $\{0.01, \ldots, 0.1\}$ in 0.01 steps and γ in the range $\{10^{-2}, \ldots, 10^{2}\}$ in power of $\sqrt{10}$ steps, respectively. OA for each model was estimated by a 4-fold cross-validation strategy. Tab. IV-2 shows the determined hyper-parameter combinations for the respective SBSTs with number of *SV*, OA estimates and corresponding values of the evaluation heuristic.

TAB. IV-2. DETERMINED HYPER-PARAMETERS (I.E., v, γ) FOR THE RESPECTIVE SBSTS WITH AFFILIATED NUMBER OF SV *(#SV)*, OA ESTIMATES (OA [%]) AND CORRESPONDING VALUES OF THE EVALUATION HEURISTIC (OA[%]/#SV)

SBSTs	v	γ	#SV	OA[%]	OA[%]/#SV
CM	0.01	0.01	8	96.31	12.04
RC low	0.01	0.01	5	95.74	19.15
RC high	0.01	0.02	5	66.67	13.33
URM	0.01	0.02	5	84.85	16.97
TF res	0.01	0.07	6	77.42	12.90
TF non-res	0.01	0.01	3	53.33	17.78
SF	0.01	0.01	5	50.00	10.00

A v-OC-SVM model was trained for each SBST based on the identified hyper-parameter combination and applied to the building inventory. Building objects that were not identified as class members were removed from the *in situ* data. Analogous, instances from the building inventory that were not assigned to one of the 7 SBST classes were removed (Fig. IV-3b).

3.2.3. OVERSAMPLING FOR HANDLING SCARCE IMBALANCED DATA SETS

To tackle scarcity of the *in situ* data and address problems associated with class imbalance, we oversampled the training data. As can be seen in Fig. IV-2, the *in situ* data for the investigated classes exhibit an uneven distribution and for some of them only a very small number of samples is available. If a class is characterized by a small number of samples, it is difficult to uncover regularities and thus construct accurate decision boundaries. In addition, the classifiers are bias-prone and tend to favor the majority classes (Nguyen et al., 2009).

Accordingly, we employed the SMOTE approach (Synthetic Minority Over-sampling Technique; Chawla et al., 2002), which generates new samples by interpolating between existing instances rather than simply duplicating original samples to avoid over-fitting. Thereby, the k nearest neighbors of a minority sample within the affiliated minority class are first identified. Depending on the amount of over-sampling required, some of the nearest neighbors are randomly selected and synthetic samples are generated along the line between the minority example and its selected nearest neighbors in the feature space (He and Garcia, 2009). It has been shown that the error rate caused by class imbalance decreases when the size of the training data increases (Japkowicz and Stephen, 2002). For this reason, we choose to keep the influence of synthetic samples as low as possible and at the same time equal for each class. Thus, we oversampled *in situ* ground-truth data for all considered classes by 500%, by keeping the *a priori* distribution of the classes. It is ensured that every class consists of at least

60 instances and the classification models deployed later on are not biased due to an alteration of the class distribution. According to the recommendations of Chawla et al. (2002) the number of nearest neighbors to be considered was set to 5.

3.2.4. MULTICLASS CLASSIFICATION WITH SVM AND RF

For the actual estimation of SBSTs, we deploy SVM and RF since both algorithms showed excellent performances in previous studies, especially when high-dimensional data sets are considered. Contrarily to the OC-SVM approach described in Section 3.2.2., the multiclass SVM approach aims to discriminate two or more target classes from each other. To cope with class overlap or the existence of noise in the training data, soft margin SVM were introduced (referred to as *C*-SVM). This technique represents a modification of the maximum margin approach using relaxed separation constrains that allow for the possibility of instances on the incorrect side of the respective margin boundary (Cortes and Vapnik, 1995).

The designated description of SVM refers to binary classification. To extend it to classification problems involving more than two classes normally a parallel architecture made of an ensemble of binary classifiers is utilized (Hsu and Lin, 2002; Melgani and Bruzzone, 2004). Following this way, one can discriminate one-against-the-rest or one-against-one methods. We chose the latter, because this technique allows producing equal or better performances, while featuring a favorable trade-off between accuracy, computational costs and algorithm complexity (Hsu and Lin, 2002). As for the OC-SVM approach, also here we used RBF kernels. Learning the most appropriate *C*-SVM in conjunction with a RBF kernel requires the definition of the cost-parameter C (which controls the trade-off between the maximization of the margin and minimization of the classification error) and the kernel-width parameter γ. Tuning of C and γ was addressed by a grid search strategy based on 10-fold cross validation. Generalization accuracy is evaluated in terms of estimated kappa statistic κ (which allows considering both omission and commission errors; Foody, 2004) on the average of 5 independent trials. In conformity with the recommendations of Hsu et al. (2010), a coarse grid-search with values of $C = \{2^{-4}, 2^{-3}, \ldots, 2^{12}\}$ and $\gamma = \{2^{-5}, 2^{-4}, \ldots, 2^{3}\}$ was performed. Subsequently, a refined grid-search in the neighborhood of the resulting C and γ pair was conducted to determine the final parameterization. A *C*-SVM was trained for each feature subset (section 3.2.1.) as well as for the whole feature set.

RF represent a decision-tree-based ensemble learning method for classification and regression introduced by Breiman (2001). Ensemble learning methods build a prediction model by utilizing the strength of a collection of simple base models. Therefore, RF grow

multiple decision trees on random subsets of the training data. The high variance among individual trees, letting each tree vote for the class membership, and assigning the respective class according to the majority of the votes, allows the accurate and robust classification of unseen data with little need for fine-tuning, even in the presence of many noisy variables (ibid.; Stumpf and Kerle, 2011). Compared to the classification and regression tree approach (Breiman, 1984) used in Sarabandi and Kiremidjian (2007), RF represents a bagged predictor and thus always yields higher accuracies (Han and Kamber, 2006).

The parameters that need to be specified for generating a RF model consist of the number of classification trees to be grown n_{tree} and the number of features m_{try} used at each node. Both parameters are evaluated by means of the RF inherent out-of-bag (OOB) error measure on the average of 10 independent trials for this study (RF deployed in this paper were carried out with the randomForest package (Liaw and Wiener, 2002) within the statistical computation environment R (R Core Team, 2013)). An increase of n_{tree} causes an improvement of classification accuracy until the performance converges. Since adding more trees to the model does not induce over-fitting, it is possible to run past the point of conversion to obtain more confident OOB values. However, adding more trees simultaneously rises computation time. To determine a suitable tradeoff, models were trained with all features for n_{tree} = 1, 2, 5, 10, 20, …, 50, 100, 200, …, 1000, 2000, …, 5000, 10000. We observed that the conversion point for our data set is located between a n_{tree} value of 200 and 300. To provide a reliable OBB error estimate and maintaining the computation times in a reasonable range, we chose a n_{tree} value of 500. This is in a good agreement with the RF parameter study performed by Genuer et al. (2008). According to Breiman (2001) a value for $m_{try} = \sqrt{p}$, with p denoting the number of input features, yields near optimum classification results. However, studies revealed that larger values of m_{try} might perform better for high-dimensional data sets (Breiman, 2002; Genuer et al., 2008). Accordingly, the original training data with all features were used to check for a possible increase in performance. As suggested by Genuer et al. (2008), classification models were learned with nine m_{try} values (1, $\sqrt{p}/2$, $\sqrt{p}$, $2\sqrt{p}$, $4\sqrt{p}$, $p/4$, $p/2$, $3p/4$, p), with n_{tree} = 500 for each model. We observed that $m_{try} = \sqrt{p}$ shows near optimum results, coping with the dimensionality of the whole feature set. Thus, we used this parameterization since it is also very favorable from a computational point of view.

4. RESULTS AND DISCUSSION

4.1. DETECTED OUTLIERS

Seven OC-SVM models were learned from the labeled samples to identify and exclude outliers from both the *in situ* data and the building inventory. Each model represents a descriptor of the SBST class it was trained on. Overall, from the 573 labeled samples 32 samples were detected as outliers and thus were removed from the *in situ* data set (Tab. IV-3a). Subsequently, each of the seven models was applied separately to the building inventory. The amount of buildings classified as class members and outliers according to each model is shown in Tab. IV-3b.

TAB. IV-3. RESULTS OF THE OUTLIER DETECTION APPROACH FOR THE *IN SITU* DATA (A) AND BUILDING INVENTORY (B) WITH AFFILIATED NUMBERS AND SHARES OF CLASS MEMBERS AND OUTLIERS

SBSTs	a) *In situ* data				b) Building inventory			
	Class members		Outliers		Class members		Outliers	
	abs.	percent	abs.	percent	abs.	percent	abs.	percent
CM	265	97.43	7	2.57	79,543	90.83	8,030	9.17
RC low	186	97.89	4	2.11	77,963	89.03	9,610	10.97
RC high	14	77.78	4	22.22	4,543	5.19	83,030	94.81
URM	29	87.88	4	12.12	63,560	72.58	24,013	27.42
TF res	25	80.65	6	19.35	56,113	64.08	31,460	35.92
TF non-res	12	80.00	3	20.00	29,113	33.24	58,460	66.76
SF	10	71.43	4	28.57	14,188	16.20	73,385	83.80

As can be seen from Tab. IV-3b, notable overlapping between individual classes exists. Especially *CM*, *RC low*, and *URM* show a certain degree of conformity. *TF res* and *TF non-res* seem to be more distinctive from other SBSTs, but feature some interference between each other. *RC high* and *SF* appear to be the most distinctive SBSTs, what seems plausible since these SBSTs are generally characterized by a unique physical appearance. However, the task was not to discriminate one class from another, but to identify instances that feature a distinctive dissimilarity in relation to the *in situ* data. Overall, 5587 instances from the entire building inventory were not assigned to any of the seven SBSTs by the learned models and were therefore excluded.

4.2. RELEVANCE OF FEATURES

To identify the most relevant features for building robust multiclass classification models several feature selection algorithms were applied on the data. The features evaluated as most

important are shown in Fig. IV-4. It contains the 15 best ranked features according to the different feature ranking algorithms and the subset revealed from the *CFS* approach. These features contributed the largest shares to raise model performance for both SVM and RF (section 4.3.; Fig. IV-5).

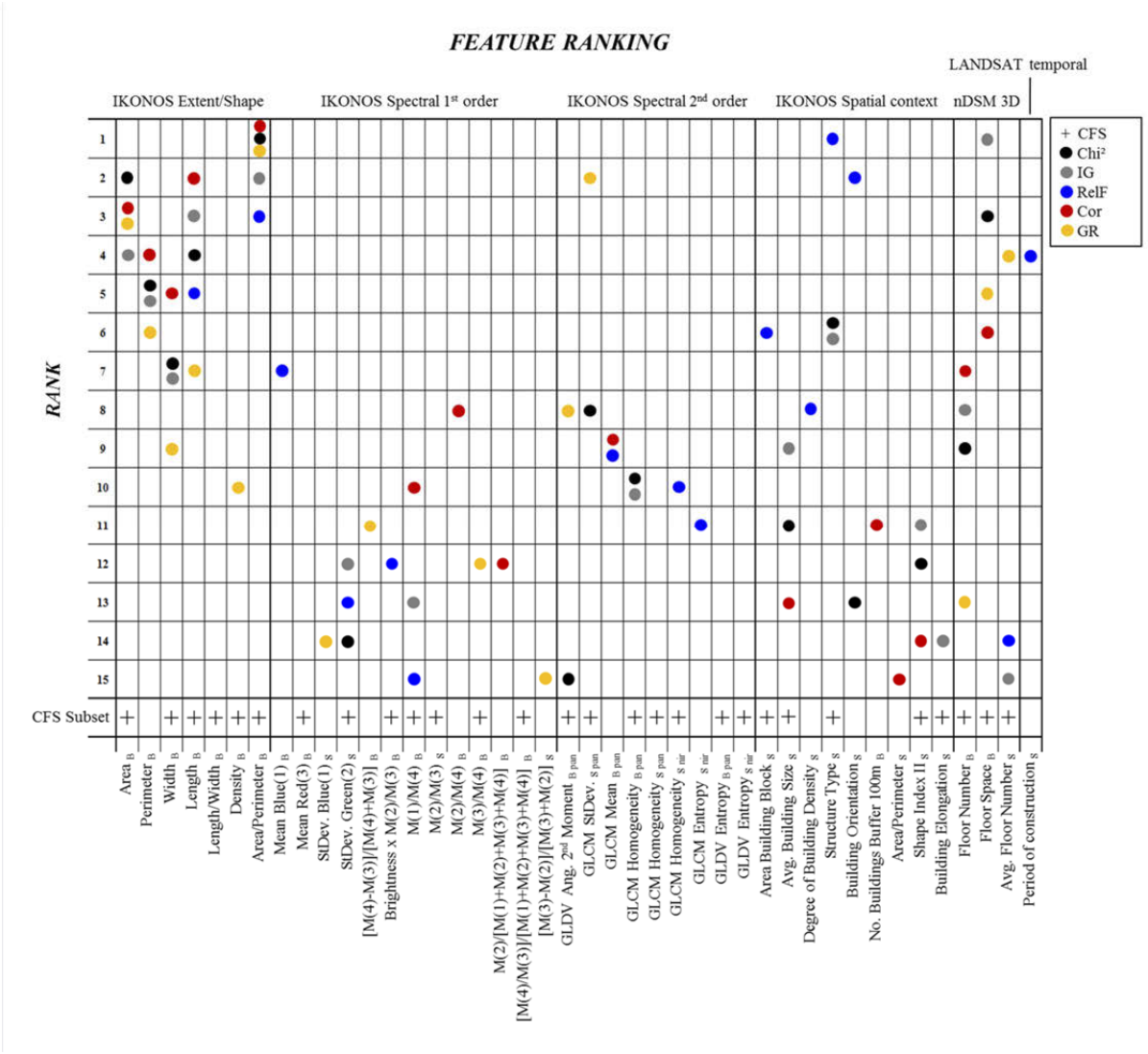


Fig. IV-4. Benchmarking of the most relevant features. The x-axis lists the features. The y-axis shows the ranks of the respective features according to the feature ranking algorithms and additionally the subset revealed from the *CFS* approach.

Overall, 42 different features of the original 145 feature vector appear in the ranking or the *CFS* subset. 23 features represent individual building characteristics and 19 represent block level information. This shows that both spatial levels are valuable to add specific information for SBSTs estimation. The IKONOS Extent/Shape features are evaluated as very valuable and mostly appear at the top of the ranking. Especially the features "Area/Perimeter $_B$", "Length $_B$", and "Area $_B$" have a very prominent position. Our findings hence go in line with earlier

work of Sarabandi and Kiremidjian (2007) who stress the importance of features on the extent of buildings to characterize seismic vulnerability. IKONOS Spatial context and nDSM 3D features show a more disperse ranking pattern and thus seem to be slightly less valuable. Interestingly, in relation to Geiß et al. (2013), the expanded feature set (Tab. IV-1b) contains some highly ranked features for both categories IKONOS Extent/Shape and IKONOS Spatial context. Features belonging to the groups IKONOS Spectral 1st and 2nd order primarily appear at the bottom positions of the ranking. However, features from these groups are noteworthy included. Overall, the LANDSAT temporal feature, which is intended to reflect approximate periods of constructions based on data from 1998 and 2000, appears to be less relevant for Padang. However, it is ranked high by the *RelF* approach, which shows distinctive performances for feature sets with a small number of features, as described in the next section.

4.3. MULTICLASS CLASSIFICATION MODELS

The classification model performances, decomposed according to the built feature sets, are reported separately for SVM (Fig. IV-5a) and RF (Fig. IV-5b). As mentioned, it can be seen that adding the first 15 best ranked features, according to the different feature selection techniques, induces a steep rise of performance. Subsequently, the increase of performance becomes more gradually. When comparing SVM and RF, it can be seen that the performance curve of the RF models start with higher κ values for the subsets containing solely five features. This trend remains until approximately 15-20 features are considered. At this point the model performances become relatively aligned. Adding further features, the estimated κ values of SVM models are higher than the estimated κ values of RF models.

It is worth mentioning that the RF models built on the features ranked by the *RelF* approach show a comparatively distinctive performance with a few number of features. Especially the ten best ranked features induce a distinctive performance for both SVM and RF, what indicates the usefulness of these features (see Fig. IV-4) for estimating SBSTs. This feature set is constituted by features from different feature categories. Analogous, the best five features as evaluated by the *GR* approach show distinctive performances and represent also features from different categories. Contrarily, the best five features as evaluated by the *COR* approach belong only to one category and perform worst. Moreover, the *CFS* subset covers a broad spectrum of features from different categories and shows a distinctive positive performance. Overall, this gives an indication that features from different categories are complementary and hence yield higher performances. Regarding the actual κ values, it has to be accentuated that accuracy estimates are very optimistic and do very unlikely reveal real

generalization capabilities on unseen data. Due to the application of the SMOTE approach, the training data contains a large fraction of synthetic samples that are very similar to the original *in situ* data. This leads to well bordered clusters in feature space, what makes it possible to learn robust models on the basis of very scarce *in situ* data. Performing a cross-validation with such a data set tends to produce very optimistic accuracy estimates. However, the accuracy estimates in terms of κ statistic served for the comparison and selection of models.

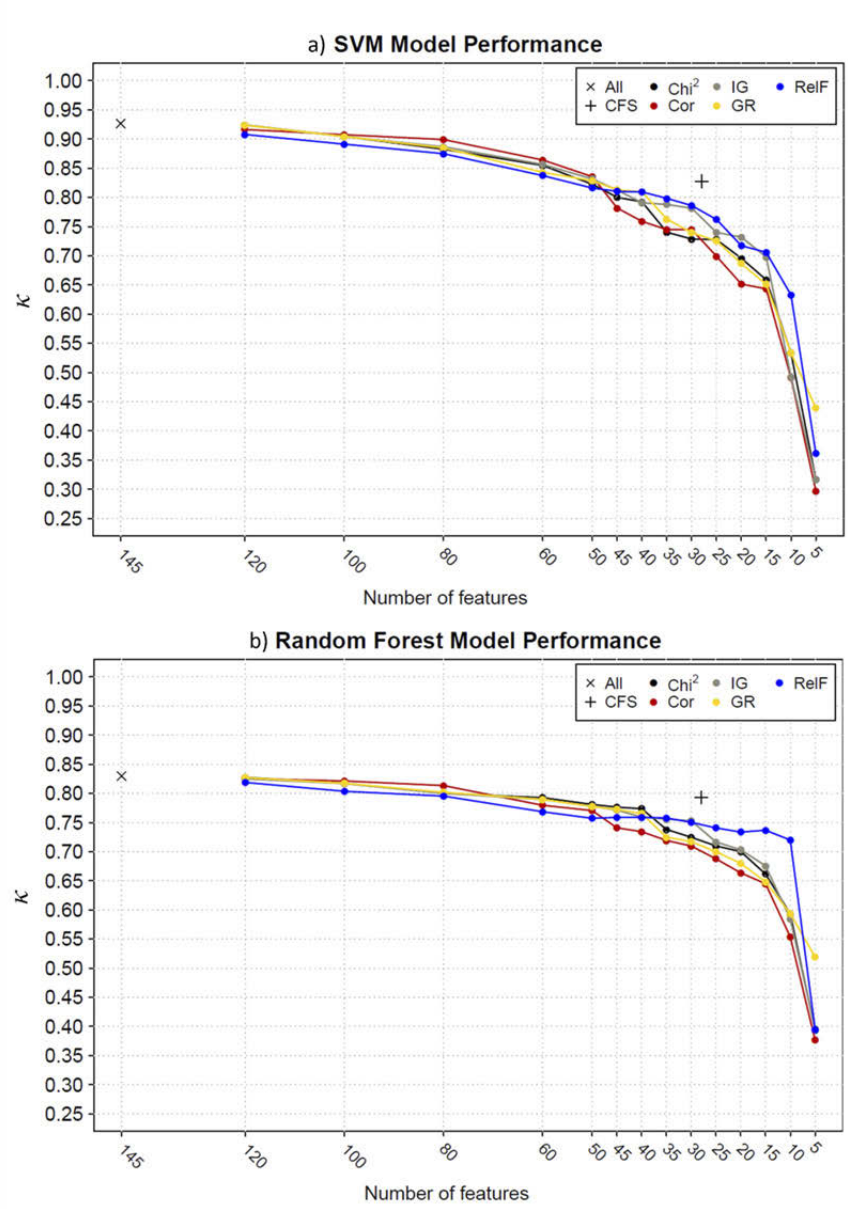


Fig. IV-5. Performance of the SVM (a) and RF (b) models as a function of the different feature sets. The feature sets were built based on the feature selection techniques described in section 3.2.1. The x-axis lists the number of features contained in the respective feature set and the y-axis reveals the corresponding κ statistic of the learned models.

4.4. CLASSIFICATION OF BUILDING INVENTORY AND PLAUSIBILIZATION OF SBSTs ESTIMATION

For the application and further plausibilization of the approach, we chose the SVM model built with the unreduced feature vector, since it yields the highest estimated κ statistic of all models. The spatially distributed estimation of SBSTs can be seen in Fig. IV-6a. Analogous to the shares of the different SBSTs of the *in situ* data, the building inventory of Padang is dominated by *CM* and *RC low* buildings. The first is more dominant in remote parts of the city and the latter shapes the central parts. Spatial concentrations of *URM* buildings appear in different sections of the city, whereas *RC high* buildings can be found primarily in the core area parallel to the coast line. *TF* buildings can be found in many parts of the city. They occur both in a diffuse and clustered manner. *SF* buildings appear most dominantly in the south-eastern parts. Outliers primarily represent very huge buildings that are located in central parts or small, informal structures that are located in very remote parts. As described in the previous section, the estimated κ statistic of the models is most likely too optimistic. Thus, we use the *in situ* samples that could not be unambiguously assigned to individual buildings (see section 2.3.) for a further plausibilization. The location of the affiliated building is not exactly known, but must be very close to the respective *in situ* sample.

A visual inspection reveals a good agreement in large areas. Five sections of Fig. IV-6b exemplify the visual plausibility checks. The results appear reasonable especially for the SBSTs *CM*, *RC low*, and *URM*. For timber frame buildings it appears that there were some confusions between *TF res* and *TF non-res*. Due to the fact that the unverified samples do not comprise *SF* buildings and a very small number of *RC high*, we do not make statements about these SBSTs. In order to support validity of estimated accuracies, we checked if there is a corresponding classified building object within a radius of 50 m from each of the 2221 unverified *in situ* buildings. This analysis resulted in an OA of 84%, what underlines the plausibility of the results.

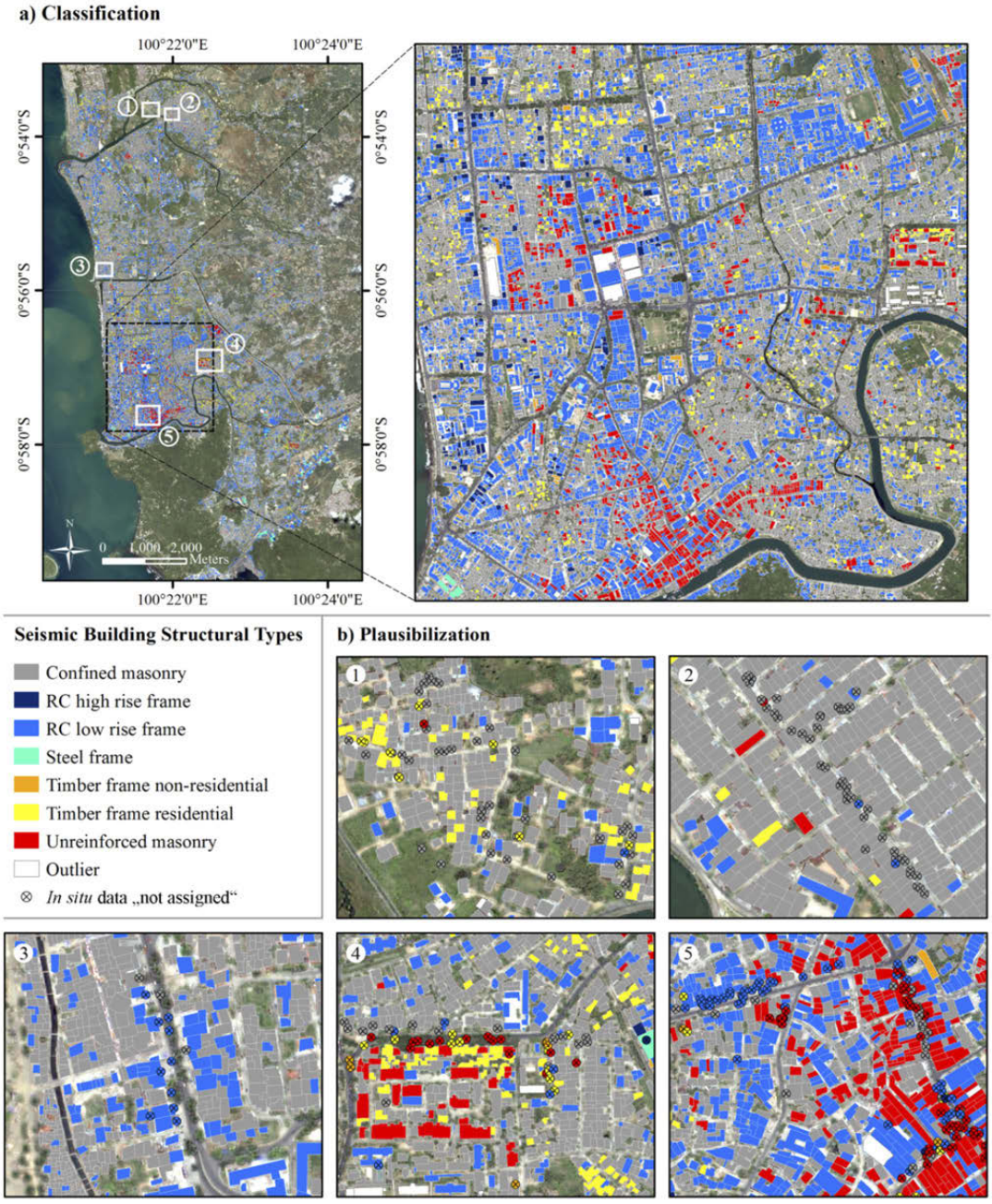


Fig. IV-6. a) Spatially distributed estimation of SBSTs by application of learned classification model to building inventory; b) Plausibilization of estimation by utilizing the spatially unverified *in situ* samples.

4.5. APPLICATION FOR EARTHQUAKE LOSS ESTIMATION

For the illustration of the applicability of the approach for ELE modeling, scenario-based loss estimations for Padang are presented. Fragility functions in the form of cumulative lognormal distributions have been derived from data collected after the 30th September 2009 event (Fig. IV-7a, Sengara et al., 2010) for different SBSTs. They relate the Modified Mercalli Intensity

(MMI) to a damage index *DI*. The latter represents an economic measure for damage and is constituted by the ratio of repair cost and total building reconstruction cost. For the presented scenarios, we assume that site effects do not play a significant role in the intensity distribution. Thus, the MMI is assumed to be spatially invariant within the city of Padang, what is consistent with observations from the aforementioned event (ibid.). Accordingly, Fig. IV-7b reveals calculated building inventory loss for several MMIs for the presented study area. In addition, spatially distributed building damage based on a *violent* MMI of 9 is shown in Fig. IV-7c. The estimation of densities regarding damaged buildings was carried out with a *KDE* approach (Bailey and Gatrell, 1995), with a Gaussian kernel and an adaptive kernel bandwidth that ensures a sample size of 50 buildings. Note that "*SF*" buildings and outliers are not included in the analyses due to missing fragility functions.

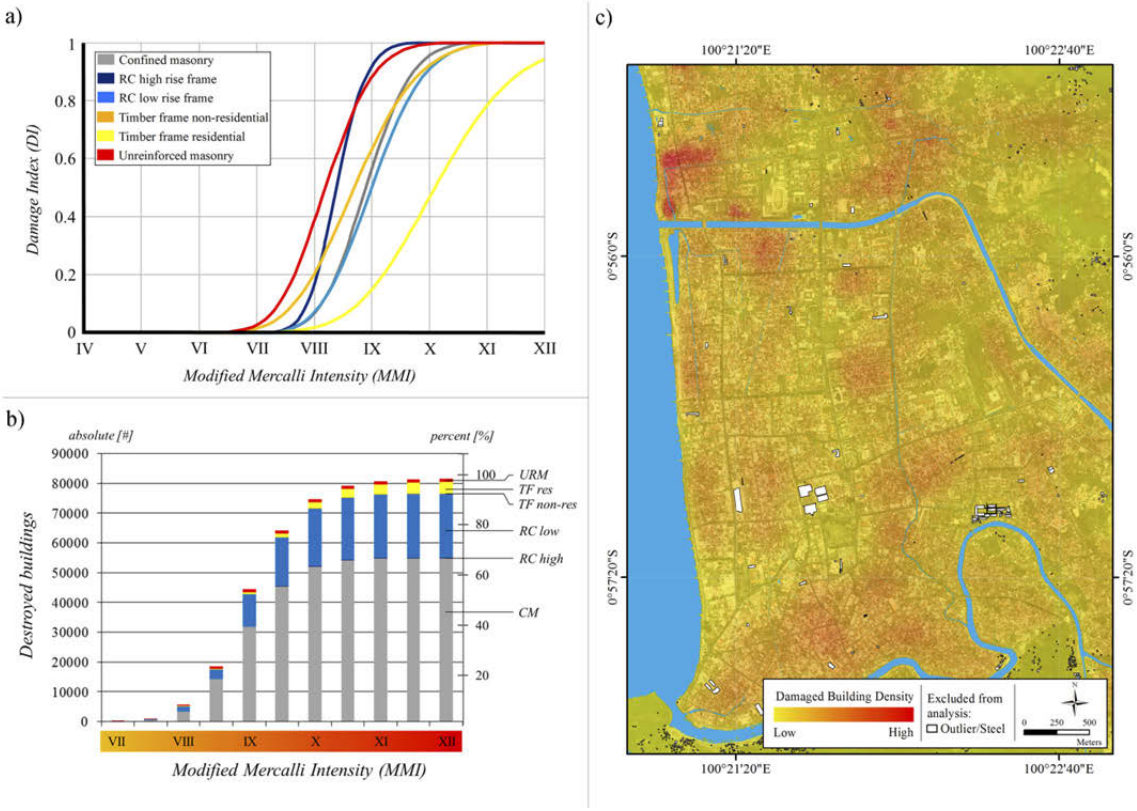


Fig. IV-7. Application for scenario-based ELE; a) Fragility functions for different SBSTs. They were derived empirically by Sengara et al. (2010) for the buildings of Padang after the 30th September 2009 earthquake; b) building inventory loss for several MMIs; c) spatially distributed building damage for a MMI of 9.

The presented approach allows to quantify building damage in a detailed way and make damage estimations spatially explicit by e.g., localizing hot spots within a city. These are key features for earthquake loss modeling and predictions.

5. CONCLUSIONS AND OUTLOOK

This study aims to demonstrate how to estimate seismic building structural types by means of a remote sensing based approach and to evaluate the suitability of the data and methods. As there is no obvious correlation between the digital information of pixels and SBSTs, we showed how to derive sets of valuable features to characterize the urban environment. For this purpose we utilized very high resolution multispectral EO-data, height information and multi-temporal medium resolution multispectral EO-data. It turned out that a combinatory use of features from different feature categories related to building shape and height, spatial context, and spectral information appears most promising. In this manner, we proposed a hierarchical supervised classification scheme that adapts techniques from the machine learning domain to estimate SBSTs. By means of a sequential procedure including *feature selection*, *outlier detection*, *generation of synthetic samples*, and *learning non-parametric SVM and RF classification models*, SBSTs could be estimated spatially distributed with plausible accuracies. Thus, we conclude that remote sensing data and methods have a high capability to support large area estimation of SBSTs.

A spatially distributed and accurate estimation of SBSTs is a critical input for ELE models as illustrated in section 4.5. and thus represents a relevant contribution to seismic risk mitigation and preparedness activities. Furthermore, related to an actual earthquake hazard, post-event earthquake loss estimations for response and recovery activities can be made, e.g. also advanced earthquake early warning systems include the capacity for the rapid assessment of damage (Picozzi et al., 2013). Such information would also be useful for a guided and thus more accurate post-event damage mapping (Dell'Acqua and Gamba, 2012). Beyond, remote sensing has the capability to quantify exposed people (Taubenböck et al., 2009b), what also allows the estimation of human casualties within the aforementioned model setting.

However, from a technical perspective, future research can utilize the enhanced resolution characteristics of latest and future spaceborne missions, such as WorldView-2 and 3, CARTOSAT-3 or ALOS-3, what allows the calculation of more sophisticated feature vectors (Novack et al., 2011). In addition, the application of computationally intensive wrapper methods for feature selection should be investigated. Regarding the actual classification task, the application of a semisupervised approach (e.g., Bruzzone et al., 2006), which also encodes some knowledge from the unlabeled data, appears promising especially when only very few labeled samples are available. To this end, active learning methods (e.g., Tuia et al., 2009) enable a guided selection of the most feasible samples. Nevertheless, the most exigent task regarding future research comprises the systematic and comprehensive collection of accurate

georeferenced *in situ* data for both SBSTs and experienced earthquake damage by structural engineers. Only this way the remote sensing community can fully demonstrate the usability of EO-data for SBSTs estimation and ELE modeling. This is why we see a need to trigger an open dialogue between the remote sensing and earthquake engineering community to share data and gather a common understanding about e.g., typologies. We believe that only a close interdisciplinary collaboration will enable systematic and valid large-area estimations of SBSTs and earthquake loss of dynamic earthquake prone urban areas around the globe.

ACKNOWLEDGMENTS

The research was founded partially by Helmholtz-EOS (Earth Observation System). The authors also would like to thank the DFG/BMBF special Programme ‘‘Geotechnologies’’—Early Warning Systems in Earth Management. Sponsorship Code: 03G0643A-E. Furthermore, we highly appreciate the encouragement of Jonathan Griffin (Geoscience Australia) and the Australia-Indonesia Facility for Disaster Reduction (AIFDR) regarding the cooperation and data sharing.

Chapter V

Normalization of TanDEM-X DSM data in urban environments with morphological filters

Submitted manuscript

Christian Geiß, Michael Wurm, Markus Breunig, Andreas Felbier and Hannes Taubenböck

ABSTRACT

TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurements) is a spaceborne radar interferometer, which delivers a global digital surface model with an unprecedented spatial resolution. This allows resolving objects above ground such as buildings. Extracting and characterizing those objects in an automated manner represents a challenging problem, but opens simultaneously a broad range of large-area applications. In this paper, we discuss and evaluate the suitability of morphological filters for the derivation of normalized digital surface models from the TanDEM-X mission in complex urban environments and introduce a novel region growing-based progressive morphological filter procedure. This approach is jointly proposed with a post-classification processing scheme to specifically allow for an accurate reconstruction of urban morphology even in challenging terrain. The introduced filter approach comprises a multistep procedure using concepts of morphological image filtering, region growing and interpolation techniques. It is based on the idea of progressive morphological filters that aim to discriminate ground and non-ground pixels in the digital surface model based on algebraic set operations. The post-classification processing scheme adapts techniques of object based image analyses (OBIA) to refine regions of classified non-ground pixels. Digital terrain models are subsequently generated by interpolating between identified ground pixels. Experimental results obtained with intermediate digital elevation model (IDEM) data for an area that covers large parts of Izmir (Turkey) reveal promising empirical evidence. The approaches perform beneficial compared to basic morphological filter based methods, especially in terrain with high orographic energy and steep areas.

1. INTRODUCTION

The TanDEM-X mission acquires data for a global digital surface model (DSM) with unprecedented resolution characteristics (Krieger et al., 2007). The spatial resolution of 0.4 arcseconds allows resolving objects in urban environments above ground such as buildings. Extraction of those objects opens a broad range of large-area applications, which are to-date unfeasible due to data availability and costs. In this sense, spatial analyzes can be extended by including threedimensional characteristics of urban environments and rely on data which is available for large areas consistently. Examples of this type of applications comprise analyzes concerning urbanization processes (Griffith et al., 2010; Taubenböck et al., 2012; Taubenböck et al., 2014), characterization of urban morphology (Baud et al., 2010; Graesser et al., 2012; Taubenböck et al., 2013), and the contribution to natural disaster related mitigation measures (Geiß and Taubenböck, 2013; Pittore and Wieland, 2013; Geiß et al., in press), among others.

When aiming at the extraction of those objects in an automated manner, it is common to derive a digital terrain model (DTM) from the data first. A DTM contains elevation measurements of the bare earth without including objects above ground. A DTM allows computing a normalized digital surface model (nDSM). The latter then naturally comprises elevation information of objects above ground. Numerous approaches have been postulated to retrieve DTM information from a DSM. Thereby, a large share of approaches is related to the derivation of DTMs from very high resolution DSMs, with ground sampling distances usually smaller than one meter. These kinds of DSMs are frequently derived from stereoscopic optical acquisitions, interferometric SAR measurements, or laser scanning (LiDAR) (Sirmacek et al., 2012; Wurm et al., 2011; Gamba and Houshmand, 2000). Especially the latter has triggered the development of ground filtering algorithms (Kraus and Pfeiffer, 1998; Axelsson, 1999; Zhang et al., 2003; Sithole and Vosselmann, 2004; Liu, 2008). However, lower spatial resolution of TanDEM-X DSM data hampers the use of many approaches.

With the benefit of a very high spatial resolution, for instance the existence of a distinct difference between the slope of terrain and that of non-ground objects such as buildings and trees can be exploited (Vosselman, 2000; Sithole, 2001). This is a prerequisite that generally cannot be expected given the acquisition characteristics and spatial resolution of TanDEM-X DSM data. Moreover, for this study we only consider approaches that operate on greyscale images and neglect waveform related approaches. This is due to user-oriented considerations, which include the fact that elevation measurements of the TanDEM-X mission are converted into regular, greyscale grid images, including several tailored correction procedures such as

masking of water areas, before dissemination. Regarding techniques that are suitable for application on greyscale grid images, approaches based on mathematical morphology (Zhang et al., 2003; Haralick et al., 1987; Weidner and Förstner, 1995; Kilian et al., 1996; Chen et al., 2007; Pingel et al., 2013), are common due to their simplicity in accordance with good performance. By means of compositions of algebraic set operations (Haralick and Shapiro, 1991), objects in the image such as buildings, can be identified since their elevation is usually higher than those of surrounding ground measurements.

However, mathematical morphology-based filters are prone to errors in steep terrains (Meng et al., 2010; Maguya et al., 2013). Generally, the biggest challenge when applying morphological methods is to keep terrain features unchanged while using large window sizes (Chen et al., 2007). Ground points are mistakenly treated as non-ground points in such settings, what results in an overestimation of objects and their height in the final nDSM (Sithole and Vosselmann, 2004; Liu, 2008; Meng et al., 2010). Besides these general challenges, the spatial resolution of TanDEM-X data poses individual challenges to a suitable approach. When aiming at the extraction of urban land cover objects such as buildings, the spatial resolution aggravates the use of algorithms that are designed for data with ground sampling distances considerably smaller than objects of interest. In particular, a large share of approaches uses the concept of progressive morphological filters that aim to separate ground/bare earth from non-ground/objects first (e.g., Zhang et al., 2003; Chen et al., 2007; Pingel et al., 2013). This allows progressively altering free parameters in dependence of the window size of the morphological filter in order to cope with different kind of land cover objects in the data (Mayer, 2000; Zhang et al., 2003). For instance, low elevation difference thresholds are used for small sizes of the filter window to eliminate objects such as bushes, small trees or cars. However, given the spatial resolution of the TanDEM-X data, even the smallest size of the filter window (3 × 3 pixels) comprises a size that already can easily exceed a building. In such a setting, i.e., applying parameters that are non-adaptive with respect to the window size, the conceptual advancement of a progressive procedure cannot be exploited the same way as it was done in previous studies.

In this paper, we discuss and evaluate the suitability of morphological filters for the derivation of normalized digital surface models from the TanDEM-X mission in complex urban environments. More significant, a novel multistep procedure using progressive morphological image filtering, region growing, and interpolation techniques is described. The approach is designed to yield beneficial results also in settings where free parameters are nonadaptive with respect to the window size. Moreover, it is intended to address and lower

errors particularly associated with steep terrain. In this manner, the strategy adopted in this work aims at a significant reduction of omission errors while taking the risk of a solely moderate increase of commission errors when separating ground and non-ground pixels. However, accuracy of classified ground and non-ground pixels have a varying influence on the quality of the final nDSM, which is dependent on the terrain characteristics of the image domain and cannot be completely determined *a priori*. To provide a suitable tradeoff between reduction of omission errors and increase of commission errors, regions of pixels that were dominantly classified as non-ground by the introduced procedure are spatially refined. Therefore, a postclassification processing scheme is proposed, which adapts techniques of object based image analysis (OBIA; Blaschke, 2010). To evaluate the suitability of introduced methods, experiments are carried out with intermediate digital elevation model (IDEM) data for an area that covers large parts of Izmir (Turkey). We focus on the extraction of elevated objects in urban environments (i.e., buildings), which are characterized by steep terrains and a high orographic energy.

The remainder of the paper is organized as follows. Section II details morphological filters in a consecutive manner: the most basic morphological filter, progressive morphological filters, and the proposed region growing-based progressive morphological filter procedure are explained in section 2.1. to 2.3., respectively. In subsection 2.4., the post-classification processing scheme is introduced. Section 3 is used for the description of data sets and experiments, whereas section 4 reports the actual results of conducted experiments. Concluding remarks are given in section 5.

2. MORPHOLOGICAL FILTERS FOR DERIVATION OF NORMALIZED DIGITAL SURFACE MODELS

To retrieve heights of objects that are elevated from the earth's surface such as buildings and vegetation from a DSM, the elevation of the terrain has to be removed. To this purpose, a DTM which represents the bare earth surface has to be derived first. With this information a nDSM can be calculated, which contains all objects above the terrain, by subtracting the DTM from the DSM:

$$nDSM = DSM - DTM. \tag{V-1}$$

This way the DSM is decomposed in terrain (DTM) and object heights (nDSM).

2.1. MORPHOLOGICAL FILTER (MF)

A very common approach to derive a DTM is to conduct a morphological opening operation (Haralick et al., 1987) on each measurement of the DSM. The approach sequentially executes a minimum and maximum filter with a structuring element *w*, e.g., a square window, on a grid of continuous surface elevation measurements *Z*. In the terminology of mathematical morphology, minimum filtering represents a special erosion operation, where the structuring element *w* has constant elevation values (Weidner and Förstner, 1995). For an individual surface elevation measurement $z_{x,y}$, the erosion of surface elevation *z* at *x* and *y* is defined as

$$\bar{z} = \min_{(x,y)\in \mathrm{w}} (z_{x,y}). \tag{V-2}$$

The eroded surface $\bar{Z}$ is subject to maximum filtering, which represents a special form of dilation and is analogously defined as

$$\bar{\bar{z}} = \max_{(x,y)\in \mathrm{w}} (\bar{z}_{x,y}). \tag{V-3}$$

To fully eliminate the objects of interest (i.e., buildings) in the opened surface $\bar{\bar{Z}}$, the window size has to be chosen so that the structuring element *w* always exceeds an object's outline. Thus, the window must have a side length of $2 \times d_{max} + 1$, with d_{max} being the largest number of pixels between an object of interest pixel and the next ground pixel, and can be determined empirically for an arbitrary image domain.

Although this classical morphological approach has proven its viability especially in flat terrains (Kraus et al., 2011), it features some conceptual drawbacks. Every single measurement that does not represent a minimum in *w* is altered and the opened surface $\bar{\bar{Z}}$ is usually lower than the original measurements (Zhang et al., 2003; Kraus et al., 2011). Additionally, the shape of the structuring element *w* (e.g., square or disk) emerges in the resulting image.

2.2. PROGRESSIVE MORPHOLOGICAL FILTER (PMF)

To lower drawbacks related to MF, approaches were presented which introduce a major methodological difference. They aim to separate bare earth (BE) pixels from object (OBJ) pixels first and interpolate a DTM from the identified BE pixels. This is frequently done within a progressive procedure, which accounts for different non-ground objects with various sizes. An opening operation is conducted on the DSM with an initial window size of the structuring element. The opened surface is subtracted from the original DSM to compute a preliminary nDSM (pnDSM; ΔZ). This allows identifying BE pixels by applying an elevation difference threshold θ. In this manner, pixels are classified as BE when they fall below a

certain value for θ. Subsequently, the opened surface is subject to a renewed opening with increased size of the structuring element to identify additional BE pixels. Thereby, θ can be varied in dependence of the size of the structuring element to address the identification of non-ground objects of differing magnitudes (e.g., also in an adaptive way in dependency of the slope of the terrain). The procedure is repeated until a maximum window size is exceeded. As explained in the previous section, the maximum window size depends on the size of the largest object of interest. Finally, the identified BE pixels can be used for interpolating a DTM (Zhang et al., 2003; Chen et al., 2007; Pingel et al., 2013).

In urban environments non-ground objects primarily consist of bushes, trees, cars, and buildings. Small non-ground objects such as cars and trees are removed within the first iterations while buildings will be removed when they are exceeded by larger sizes of the structuring element. Analogously, low elevation difference thresholds are used for small sizes of the filter window to eliminate objects such as bushes, small trees or cars. Larger thresholds are applied for window sizes which correspond to objects such as buildings. To ensure complete identification of buildings, θ must be set to the lowest building height in the study area. As stated in the introductory section, the spatial resolution of the TanDEM-X data hampers the exploitation of the conceptual advancement of a progressive procedure. Even the smallest size of the filter window (3 × 3 pixels) comprises a size (~1,300m² in our study area) that already can easily exceed a building. In such a setting, i.e., using parameters that are non-adaptive with respect to the window size, a progressive procedure cannot be exploited the same way as it can be done for data with ground sampling distances considerable smaller than the objects of interest.

Moreover, PMF perform generally well in flat terrains but at the same time they are prone to errors in steep terrains, similar to MF. There are two basic errors that arise in identifying BE pixels in the DSM. One is error of commission (False Positives, Type I error) that describes pixels which are mistakenly classified as BE. The second is error of omission that describes pixels that are mistakenly not classified as BE (False Negatives, Type II error). In steep terrain an increased number of BE pixels is mistakenly identified as OBJ pixels (omission error) because they feature a distinctive difference of the elevation with respect to surrounding pixels similar to e.g., buildings. This leads simultaneously to overestimations of objects and their height in the final nDSM. Generally, the influence of omission and commission errors on the quality of the final nDSM cannot be considered equally and is strongly dependent on the terrain characteristics. For a perfectly flat terrain, an extremely high omission error would yield an acceptable nDSM, since only a few correctly classified BE

pixels would allow an adequate interpolation and representation of the terrain. Contrarily, in steep terrain the nDSM is prone to contain large fractions of terrain information when BE pixels are not exhaustively identified.

2.3. REGION GROWING-BASED PROGRESSIVE MORPHOLOGICAL FILTER (RPMF)

To address problems particularly associated with non-flat terrain, the strategy adopted here aims at a significant reduction of omission errors while taking the risk of a solely moderate increase of commission errors compared to the PMF approach. An overview of the approach, referred to as region growing-based progressive morphological filter (RPMF), is given in Fig. V-1.

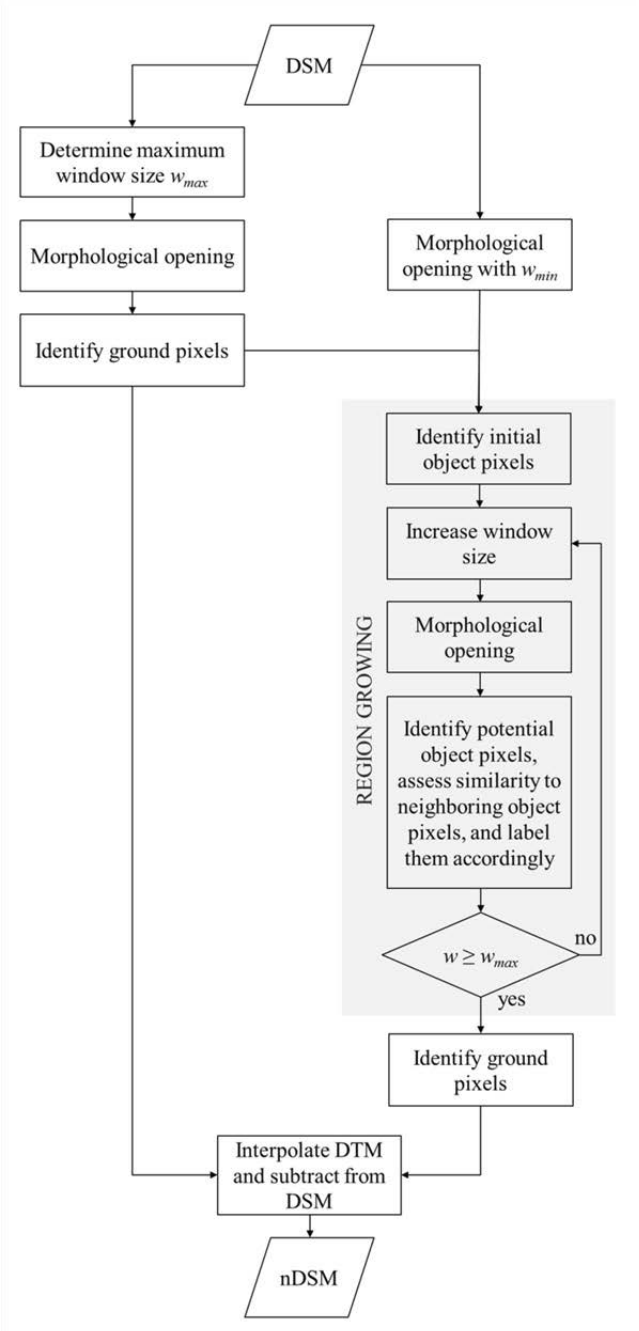


Fig. V-1. Flowchart of the region growing-based progressive morphological filter for the derivation of a nDSM from a DSM.

The approach can be subdivided according to three consecutive main steps. First, the DSM is opened with the maximum window size w_{max} and BE pixels are identified by applying an elevation difference threshold. This is done to retain BE pixels that can be considered reliable (low commission error) but not exhaustive (high omission error). Second, a procedure is conducted for the remaining pixels, which aims to identify pixels which represent elevated objects (i.e., buildings; OBJ pixels). Ideally, all non-identified pixels then represent BE pixels, what allows interpolating solely areas that are covered by elevated objects. To this purpose, a

region growing procedure is deployed. Lastly, all identified BE pixels are used for interpolating a DTM, which enables computing the final nDSM.

Regarding the region growing procedure, initial OBJ pixels are identified which are used as seed pixels. Fig. V-2 illustrates the image processing steps and criteria applied. The figure is divided according to the strategy adopted for objects that are exceeded by the minimum size of the structuring element (V-2a) and objects that are not exceeded by it (V-2b). To identify objects that are exceeded by w_{min}, a pnDSM is calculated by subtracting $\bar{\bar{Z}}_{w_{min}}$ from Z. OBJ pixels are identified by the application of a threshold θ on the pnDSM. An alternative strategy was adopted to account for objects that are not exceeded by w_{min}. These objects are fully preserved in $\bar{\bar{Z}}_{w_{min}}$ and are thus not contained in the pnDSM. Pixels at the border of an elevated object are identified by subtracting $\bar{Z}_{w_{min}}$ from $\bar{\bar{Z}}_{w_{min}}$. Subsequently, these pixels can be classified as OBJ pixels by sequentially combining an edge extraction filter and contrast segmentation.

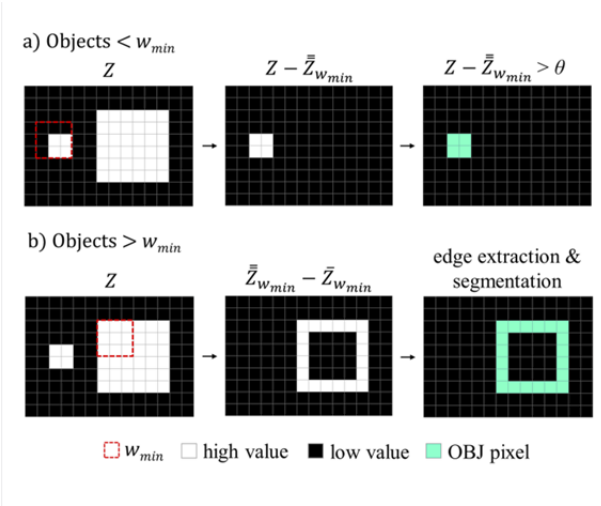


Fig. V-2. Idealized procedure for the identification of initial OBJ pixels that represent the basis for the region growing procedure. (a) Objects that are exceeded by the minimum window size (w_{min}) can be identified by the application of a threshold θ on the pnDSM ($Z - \bar{\bar{Z}}_{w_{min}}$). (b) Border pixels of objects that are not exceeded by w_{min} are identified by subtracting $\bar{Z}_{w_{min}}$ from $\bar{\bar{Z}}_{w_{min}}$. They are classified as OBJ by combining edge extraction filter and segmentation. Section a) of this figure visualizes step (5) and (6) of Algorithm 1 and section b) visualizes step (7). The identification of initial OBJ pixels with IDEM TanDEM-X data is exemplified in Appendix A (Fig. V-8(a)).

In particular, we rely on edge extraction by Lee-Sigma edge detection filtering (Lee, 1983). The Lee- Sigma filter represents an adaptive filter based upon the spatial domain. It utilizes the sigma probability of the Gaussian distribution to smooth variations in the image by averaging only those neighborhood pixels which have intensities within a fixed range of standard deviations of the center pixel. Consequently, edges and linear features are preserved. The filter is applied on the image data generated by subtracting $\bar{Z}_{w_{min}}$ from $\bar{\bar{Z}}_{w_{min}}$ to extract

bright edges, i.e., distinctive transitions of brightness, which ideally represent boarder pixels of building objects exceeded by w_{min}. To classify these pixels as OBJ in an automated manner, a contrast segmentation approach is deployed. The filtered image is segmented into dark and bright pixels. To maximize the contrast between resulting dark pixels and bright pixels, a threshold is iteratively evaluated and accordingly adjusted (Trimble, 2012).

After the identification of initial OBJ pixels that serve as seed pixels, the window of the morphological opening is increased linearly

$$w_k = 2\,x\,k + 1 \tag{V-4}$$

where k = 1, 2, …, d_{max}, with d_{max} being the largest number of pixels between an object pixel and the next ground pixel (see also section V-2.1.). In this manner, the filtered surface obtained from the previous iteration is opened with an increased window size. At each iteration additional OBJ pixels are identified if they exceeded θ in the respective newly calculated pnDSM and fulfill a similarity constraint with respect to already classified OBJ pixels. The similarity (sim) of a potential OBJ pixel (potOBJ) and adjacent OBJ pixel(s) (OBJ pixels that share a common boundary with a potOBJ pixel) is evaluated as

$$sim(potOBJ, OBJ) = \begin{cases} 1, \mid \mu\left(\Delta Z_{OBJ}\right) - \Delta Z_{potOBJ} \mid \leq \gamma \\ 0, else \end{cases} \tag{V-5}$$

where ΔZ represents the pnDSM value, μ the mean value and γ represents a threshold. This process is repeated during each iteration until a stable situation is reached, where no more potOBJ pixels are classified as OBJ pixels. Some magnifications of both the identification of initial OBJ pixels that serve as seed pixels as well as of the actual region growing procedure with IDEM TanDEM-X data are visualized in Fig. V-8(a) and V-8(b), respectively (Appendix V-A). The whole procedure to identify BE pixels is also described in the pseudocode under Algorithm 1. Lastly, all unclassified pixels are considered as BE and are used in conjunction with previously identified BE pixels to interpolate a final DTM. In this manner, an exact interpolation method is favorable to keep BE pixels unaltered.

Algorithm 1 Identification of bare earth (BE) pixels with RPMF

Inputs:
Z: A regular grid of continuous surface elevation measurements
w_{min}: Minimum window size
w_{max}: Maximum window size
θ: Elevation difference threshold
γ: Threshold for similarity constraint
Output:
S: Set of pixels that represent ground measurements

1: Open Z by applying sequentially (V-2) and (V-3) with w_{max}
2: Compute preliminary nDSM by subtracting $\bar{\bar{Z}}_{w_{max}}$ from Z
3: Identify BE pixels that are below elevation difference threshold θ and add them to S
4: Open Z by applying sequentially (V-2) and (V-3) with w_{min}
5: Compute preliminary nDSM by subtracting $\bar{\bar{Z}}_{w_{min}}$ from Z
6: Identify initial OBJ pixels that are above elevation difference threshold θ and yet unclassified
7: Subtract eroded surface from opened surface $\bar{Z}_{w_{min}}$ from opened surface $\bar{\bar{Z}}_{w_{min}}$ and identify unclassified, initial OBJ pixels by applying edge extraction and contrast segmentation
8: **repeat**
9: Increase window w by (V-4)
10: Open Z by applying sequentially (V-2) and (V-3) with w
11: Compute preliminary nDSM by subtracting $\bar{\bar{Z}}_{w}$ from Z
12: **repeat**
13: Identify potential OBJ pixels that are above elevation difference threshold θ, have adjacent OBJ pixel(s) and are yet unclassified
14: Assess similarity of potential OBJ pixels with respect to already identified OBJ pixels with similarity constraint (V-5) and label them as OBJ pixels when below γ
15: **until** the number of classified OBJ pixels equals the number of classified OBJ pixels from previous iteration
16: **until** $w \geq w_{max}$
17: Add all unclassified pixels to S

2.4. SELECTIVE POST-CLASSIFICATION PROCESSING WITH OBJECT BASED IMAGE ANALYSIS (RPMV-SOBV)

As explained in the previous section, the strategy of RPMF is to identify OBJ pixels and subsequently consider all remaining unclassified pixels as BE. This procedure is prone to deliver dissatisfactory results when dealing with building objects with irregular roof surfaces, or when considerable local variations of elevation values occur. Especially when applying restrictive γ values for the similarity constraint (V-5) during the region growing procedure, pixels may remain unclassified although they are part of a continuous building object. Since

pixels that remain unclassified after the region growing procedure will be considered as BE, such errors will result in partially eliminated building objects in the final nDSM. To lower those errors, we introduce a post-classification processing scheme that adapts concepts of object based image analyses (Blaschke, 2010), referred to as region growing-based progressive morphological filter with selective object based voting (RPMF-SOBV). The selective object based voting scheme is constituted by three consecutive steps and is illustrated in Fig. V-3. (i) The pnDSM computed with w_{max} is segmented into homogeneous image regions. (ii) Segments are selected that contain solely OBJ and unclassified pixels after the region growing procedure of RPMF. (iii) All pixels of selected segments are classified according to the maximum class probability of the respective segment.

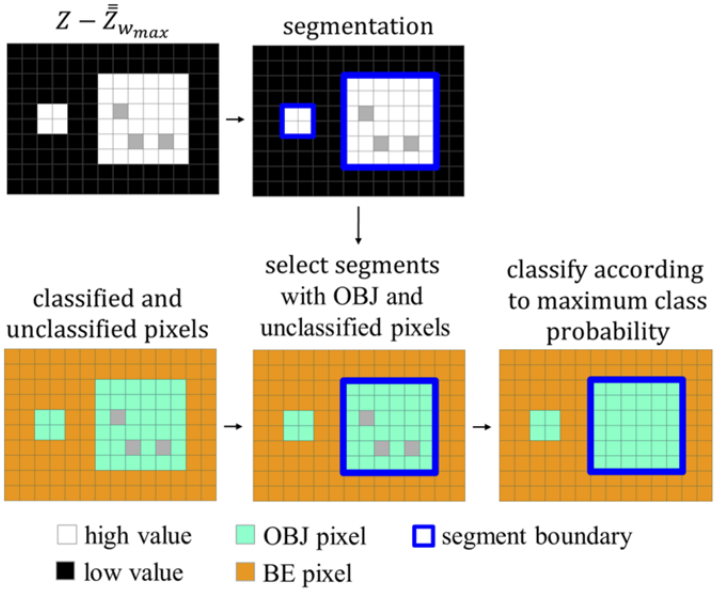


Fig. V-3. Scheme for selective post-classification processing of regions with OBJ pixels and unclassified pixels after the region growing procedure of RPMF. First, a preliminary nDSM is calculated with w_{max} to ensure that all building objects are contained. A segmentation algorithm is used to discriminate homogenous image regions. Segments are selected that contain solely OBJ pixels and unclassified pixels. Pixels of selected segments are classified according to the maximum class probability of the respective segment. The selective post-classification processing of IDEM TanDEM-X data is exemplified in Appendix B (Fig. V-9).

Regarding step (i), the pnDSM is computed with w_{max} to ensure that all building objects are contained. Hence, segmentation aims at the delineation of building footprints, which are homogeneous in terms of their response in the pnDSM. We deploy multiresolution segmentation based on fractal net evolution (*FNEA*) approach (Benz et al., 2004). This is a region-growing segmentation algorithm to segment image data without *a priori* knowledge. Some parameters need to be specified when segmenting image data with this approach. In this manner, we suggest to put more emphasize on shape heterogeneity rather than on grey-value heterogeneity. This is due to the fact that man-made features such as buildings have distinct

shape and size properties, unlike e.g., natural features. Analogously, the weights for heterogeneity of smoothness and compactness can be maintained equal. However, the main difficulty lays in the objective determination of an optimal segmentation scale. We deploy the objective function introduced by the authors of (Espindola et al., 2006) to find domain-specific optimal segmentation scales objectively. Based on the assumption that optimal segmentation maximizes intrasegment homogeneity and intersegment heterogeneity, a measure is calculated by incorporating intrasegment variance (σ^2) and Moran's I (I). The intrasegment variance (σ^2) with respect to the pnDSM is calculated as

$$\sigma^2 = \frac{\sum_{i=1}^{n} A_i {\sigma^2}_i}{\sum_{i=1}^{n} A_i} \tag{V-6}$$

where A_i and ${\sigma^2}_i$ represent the area and intrasegment variance of segment i. The intrasegment variance σ^2 is the weighted average, with the areas of segments being the weights. As mentioned, Moran's I is used as a measure of intersegment heterogeneity

$$I = \frac{N}{\sum_i \sum_j e_{ij}} \cdot \frac{\sum_i \sum_j e_{ij} (\mu_S(\Delta Z)_i - \mu(\Delta Z))(\mu_S(\Delta Z)_j - \mu(\Delta Z))}{\sum_i (\mu_S(\Delta Z)_i - \mu(\Delta Z))^2} \tag{V-7}$$

where N is the number of segments indexed by i and j, $\mu_S(\Delta Z)$ the mean pnDSM value of a segment, $\mu(\Delta Z)$ is the mean pnDSM value of all segments, and e_{ij} represents the spatial weight between segments i and j, implemented in our case analogous to a queen contiguity as follows

$$e_{ij} = \begin{cases} 1, & \text{if } i,j \text{ are adjacent neighbour segments} \\ 0, & \text{else.} \end{cases} \tag{V-8}$$

With the determined variance and autocorrelation measure the objective function is calculated by summing up normalized values of σ^2 and I

$$F(\sigma^2, I) = \frac{{\sigma^2}_{max} - \sigma^2}{{\sigma^2}_{max} - {\sigma^2}_{min}} + \frac{I_{max} - I}{I_{max} - I_{min}}. \tag{V-9}$$

The maximum value of function $F(\sigma^2, I)$ is a statistical indicator of optimal segmentation.

From the generated segments only those are selected that contain solely OBJ and unclassified pixels after the region growing procedure of RPMF. For the selected segments a decision rule is deployed (Huang et al., 2014), where a probability function is maximized

$$C(m) = \arg\max_{v \in C} (\tilde{P}_{m,v}) \tag{V-10}$$

with C being the labeling space constituted by class labels "OBJ" and "unclassified", $C(m)$ represents the final label of pixel m, and $\tilde{P}_{m,v}$ denotes the probability of pixel m belonging to class v. In accordance with a crisp voting scheme revealed in (ibid.), we compute

$$P_{s,v} = \frac{1}{N_s} \sum_{u \in s} \tau(C(u) = v) \quad \text{(V-11)}$$

where τ is an indicator function capturing the number of times that the pixels u within a selected segment s feature class label v, and N_s is the number of pixels in selected segment s. The new class label of the pixels in selected segments is determined according to the maximum probability, as indicated in (10).

The most common processing paradigm of object based image analysis comprises first a segmentation followed by an 'objectification' (i.e., classification; Blaschke et al., 2014). In contrast to that, RPMF-SOBV relies on per-pixel classification of RPMF, which is followed by a segmentation, and then classifies selected image regions based on certain class fractions of the superobjects of individual pixels. This procedure is designed to cope naturally with the resolution characteristics of the data, which hampers the use of a pure object based approach. The latter is frequently intended to model 'meaningful image objects' (ibid.). However, even the smallest spatial entity (i.e., pixel) may correspond to an object of interest (i.e., building). Moreover, in such a setting, a pure object based approach, comprising the complete aggregation of the image pixels according to certain scale(s), may hardly be applied to TanDEM-X data without violating the Shannon sampling theorem. The theorem states that modeled objects should be of the order of one tenth of the dimension of the sampling scheme – the pixel – to ensure that they will be completely independent of their random position and their orientation in relation to the sampling scheme (Blaschke, 2010). In this manner, RPMF-SOBV is intended to benefit from both the per-pixel and object based image analyses paradigm. The selective post-classification processing of IDEM TanDEM-X data is exemplified in Appendix V-B (Fig. V-9).

3. DESCRIPTION OF DATA SETS AND EXPERIMENTS

3.1. TANDEM-X INTERMEDIATE DIGITAL ELEVATION MODEL (IDEM) DATA

The TanDEM-X satellite was launched in June 2010 and is operating jointly in a unique helix tandem formation with its twin radar satellite TerraSAR-X, which is in space since June 2007. Since December 2010, however, the two satellites are operationally acquiring data to generate a seamless global multi-coverage digital elevation model using single-pass interferometry in bi-static mode (Krieger et al., 2007). Since August 2011 the data is being processed operationally. Single- and multi-baseline interferometric phase-unwrapping (Rossi et al., 2012), manual quality inspection, water detection (Wendleder et al., 2012), height comparison

with ICESat data (Huber et al., 2009), and calibration including tilt and offset correction is carried out in order to provide a homogenous dataset for a regional mosaic (Gruber et al., 2012). In our study, two tiles of the so called TanDEM-X Intermediate Digital Elevation Model (IDEM) (Wessel et al., 2013) are used (Tile N38E026 & N38E027). The IDEM consists only of the best quality single-baseline processed data of the first global coverage. Insufficient acquisitions affected by phase-unwrapping errors are excluded. The data sets cover large parts of the city of Izmir (Turkey) and its hinterland. The terrain features a high orographic energy with partially steep areas. The spatial resolution of the data of 0.4 arc seconds corresponds to ~12m in Izmir.

3.2. URBAN FOOTPRINT (UF) DATA

To focus our experimental analysis on urban environments, an approach which was introduced by (Esch et al., 2010), was deployed to discriminate "built-up" and "non-built-up" land cover. The approach was implemented as a fully automated image analysis procedure, which is currently applied to delineate Urban Footprints (UF) from single polarized Strip Map imagery of the TanDEM-X mission (Esch et al., 2013) globally. The high classification accuracy, which exceeds consistently an overall accuracy of 94% and a κ statistic of 0.75 for representative case studies (ibid.), allows us to spatially focus on man-made structures within urban environments (i.e., buildings) and at the same time neglect conceptual and methodological considerations related to other objects that are elevated from the earth's surface (i.e., vegetation). The UF of Izmir comprises a large range of different man-made structures reaching from e.g., small and very low-rise informal settlements (Gecekondular) to large industrial buildings.

3.3. DESCRIPTION OF EXPERIMENTS

Experimental analyses are based on a reference data set, which contains BE and OBJ pixels. 1000 points were randomly generated within the UF of Izmir. By careful manual inspection of the TanDEM-X data and additional utilization of VHR optical imagery, 605 BE pixels and 186 OBJ pixels in the TanDEM-X data could be identified. The 209 remaining pixels could not be unambiguously allocated and thus were excluded from the reference data set.

The experimental analyses are organized in three main parts. The first analysis aims at demonstrating the effectiveness of the proposed RPMF, and RPMF-SOBV approaches to separate BE from OBJ pixels. Therefore, a comparison with the PMF approach is provided in

terms of computed False Negative and False Positive Rates. The approaches' accuracies are studied in dependency of their free parameters. In the second analysis, the accuracies of generated DTMs are assessed. The quality of the generated DTMs are assessed based on the 605 verified BE pixels. Therefore, the deviations of the DTMs from BE pixels of the TanDEM-X data were calculated. In this manner, we do not assess absolute quality of DTMs with respect to a certain reference but their relative quality. Thus, this analysis serves as a further comparison of methods. It is carried out under the premise that low deviations of DTM values from BE pixel values of the TanDEM-X DSM (which represent terrain information) are favorable. In the third analysis we visually assess the quality of the final nDSMs according to the different DTM generation approaches and discuss favorable and less favorable characteristics of respective strategies.

4. EXPERIMENTAL RESULTS

4.1. RESULTS OF ANALYSIS I: ACCURACY ASSESSMENT OF BE AND OBJ PIXEL CLASSIFICATION

This analysis assesses the ability of PMF, RPMF, and RPMF-SOBV to distinguish BE and OBJ pixels. MF is not considered, since it does not feature the inherent concept of the other approaches to separate BE from OBJ pixels first and interpolate a DTM from the identified BE pixels. The classification accuracies of PMF, RPMF, and RPMF-SOBV are studied in dependency of their free parameters. Free parameters were kept constant during the progressive increase of the size of w. This is important since free parameters are frequently altered in dependency of the size of w when working with data with a very high spatial resolution. As explained in section 1 and 2.1., low θ values are used for small sizes of w to eliminate objects such as bushes, small trees or cars. With increasing size of w, θ values are also increased to account for objects such as buildings. However, regarding the TanDEM-X data, even the smallest size of w (3×3 pixels) comprises a size that already can easily exceed a building. To take account for large industrial buildings in the southeastern part of the study area, the maximum size of w was set to 15 pixels for all approaches. Additionally, we chose a square shaped structuring element. The height threshold was assessed in the interval $\theta \in [2, 2.2, ..., 3.6]$ and the similarity constraint was assessed in the interval $\gamma \in [0.6, 0.8, \ldots, 2]$. The basic parameters of the different approaches are also bundled in Tab. V-I.

TAB. V-1. BASIC PARAMETERS FOR PMF, RPMF, AND RPMF-SOBV USED IN THE EXPERIMENTS

Filter approach	Shape of structuring element	w_{min} [pixels]	Increment steps for windows (k in Equation (4))	w_{max} [pixels]	θ [meters]	γ [meters]
PMF	square	3	1, 2, …, 7	15	2, 2.2, ..., 3.6	—
RPMF	square	3	1, 2, …, 7	15	2, 2.2, ..., 3.6	0.6, 0.8, …, 2
RPMF-SOBV	square	3	1, 2, …, 7	15	2, 2.2, ..., 3.6	0.6, 0.8, …, 2

Regarding the edge extraction filter procedure, which is part of RPMF (and then naturally also relevant for RPMF-SOBV), we apply a σ value of 4 (preliminary experiments showed that results are hardly sensitive with respect to the specification of σ). Regarding the segmentation which is an essential processing step of RPMF-SOBV, we put more emphasize on shape heterogeneity (Shape: 0.7) rather than on grey-value heterogeneity (0.5) as suggested. To identify an optimal segmentation scale that can be used for RPMF-SOBV, we calculated the objective function for scales in the range of [5, …, 15] when segmenting the pnDSM computed with w_{max}. The affiliated objective function is revealed in Fig. V-4. It can be seen that the maximum value, which represents a statistical indicator for optimal segmentation, is reached with a scale factor of 9 for our data set. Hence, segmentation used for RPMF-SOBV was carried out with this scale factor.

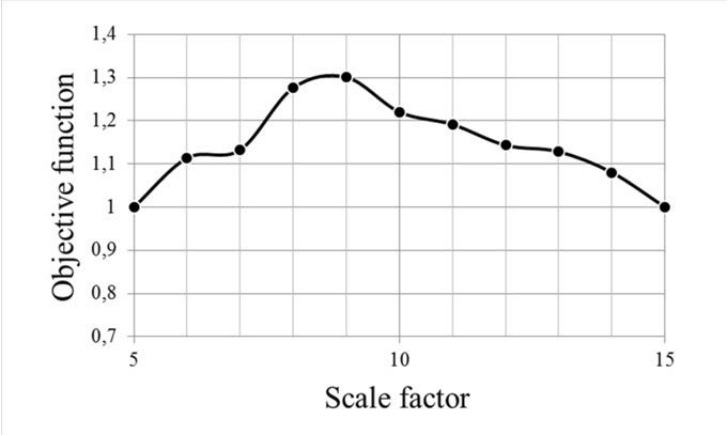


Fig..V-4. Objective function calculated for scales in the range of [5, …, 15] when segmenting the pnDSM computed with w_{max}. The maximum value of the objective function is a statistical indicator for optimal segmentation. The maximum value for our data set is reached with a scale factor of 9.

Based on the reference data, False Negative (FN) and False Positive (FP) Rates are calculated for PMF, RPMF, and RPMF-SOBV with the specified parameters. Results are revealed in Fig. V-5. In Fig. 5(a) it can be seen that FN rates decrease continuously with an increasing θ value and FP rates increase continuously with increasing θ value for PMF, RPMF, and RPMF-SOBV. This is reasonable since many building objects are lower than a high θ value in

this range what leads to an increase of FP. Analogously, terrain is predominantly classified as elevated objects with a low θ value what results in a high FN rate. Generally, it can be observed that all approaches face the problem of considerably misclassifying terrain as elevated objects, what results in a generally quite high level of FN rates.

When comparing PMF and RPMF, it is obvious that FN rates of RPMF are consistently lower and simultaneously FP rates are consistently higher. FN rates of RPMF-SOBV are slightly higher than FN rates of RPMF and FP rates are slightly above FP rates of PMF. This uncovers a remarkable performance characteristic of RPMF-SOBV, which is also exemplified in Fig. 5(b). The figure shows FN and FP rates for a certain θ value (2.6m) as a function of γ. It can be seen that RPMF-SOBV combines favorable performance characteristics of both PMF and RPMF. Regarding FN rates it is aligned to the better performance of RPMF. In contrast, regarding FP rates it is aligned to the better performance of PMF. Generally, when increasing γ, we observe that FN and FP rates associated with RPMF and RPMF-SOBV become more aligned with the error rates of PMF for the given interval. In addition, the respondence of FN and FP rates become less sensitive. Fig. 5(c) visualizes the gain of accuracy (decrease of FN rate) and loss of accuracy (increase of FP rate) of both RPMF and RPMF-SOBV compared to PMF. Generally, it can be observed that the decrease in FN exceeds the increase of FP for both RPMF and RPMF-SOBV compared to PMF considerably, while maintaining FN rates in a reasonable range. Compared to RPMF, the decrease of FN rates of RPMF-SOBV is more moderate, however, the increase of FP remains simultaneously on a very low level.

As already mentioned, the classification errors have a varying influence on the quality of the DTM that cannot be fully determined *a priori*. It is primarily dependent on the terrain characteristics (i.e., flat or steep). Hence, further analysis related to the quality assessment of the resulting DTMs is carried out in the next section.

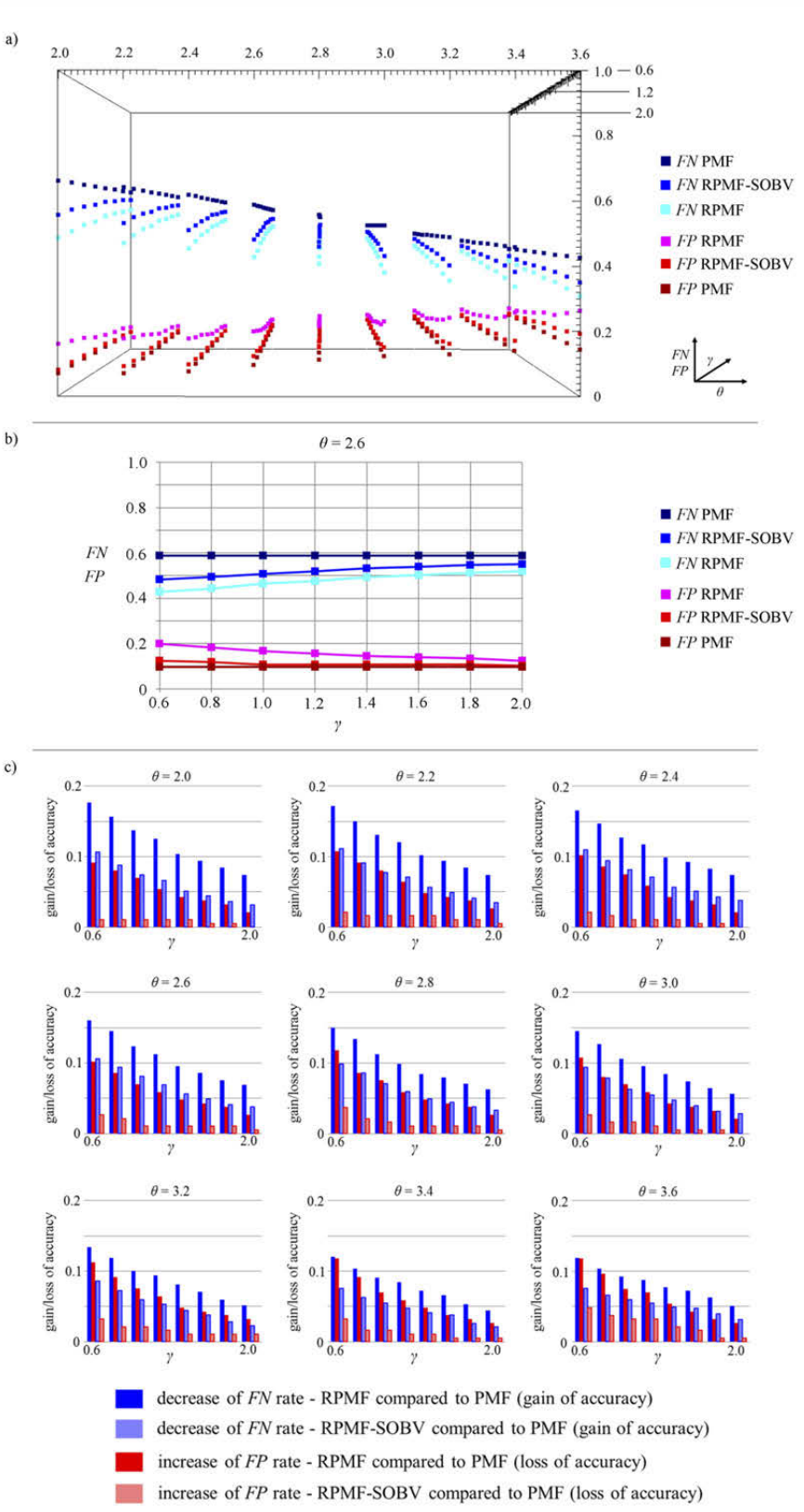


Fig. 5. (a) False Negative (FN) and False Positive (FP) rates achieved with PMF, RPMF, and RPMF-SOBV (plotted on the yaxis) in dependency of the approaches' free parameters. The free parameters are height threshold θ (plotted on the x-axis in meters) for PMF, RPMF, and RPMF-SOBV, and similarity constraint γ (plotted on the z-axis in meters; since this parameter is only relevant for RPMF and RPMF-SOBV it does not vary for PMF). (b) FN and FP rates for a certain θ value (2.6m) as a function of γ. (c) Gain of accuracy (decrease of FN rate) and loss of accuracy (increase of FP rate) of both RPMF and RPMFSOBV compared to PMF.

4.2. RESULTS OF ANALYSIS II: ACCURACY ASSESSMENT OF DTMS

DTMs were generated based on the BE pixels classified by PMF, RPMF, and RPMFSOBV. Additionally, the surface generated by applying an opening on Z with w_{max} (MF) is evaluated. Analogous to the first analysis, the maximum size of w was set to 15 pixels for all approaches. A quite restrictive θ value (i.e., 2.6m) for PMF, RPMF, and RPMF-SOBV was chosen to account for the aforementioned very small and low buildings, as the general aim is to extract every single building. Additionally, a reasonable tradeoff between FN and FP is featured this way, as can be seen in Fig. V-5. Analogously, a quite restrictive value for the similarity constraint γ (i.e., 0.8m; only relevant for RPMF and RPMF-SOBV) was chosen, to achieve a favorable tradeoff regarding the overall gain of accuracy with respect to PMF. Analogous to the authors of (Zhang et al., 2003) we used a Kriging approach (Wackernagel, 2003) to interpolate BE pixels to a DTM surface. We used ordinary Kriging on the BE pixel sets with spherical semivariogram models and adaptive search radii to ensure a minimum number of BE pixels to be included.

As explained in section 3.3, quality of the generated DTMs is assessed based on the 605 verified BE pixels and is intended to serve as a further comparison of methods. In this sense, low deviations of DTM values from BE pixel values of the TanDEM-X DSM (which represent terrain information) are favorable. Deviations and absolute deviations in meters (m) are visualized in Fig. V-6 on the basis of boxplots (depicting median, interquartile range (IQR), as well as 1.5 times IQR). It can be observed that in general all approaches have the tendency to underestimate terrain heights. This will result in overestimations of object heights in the nDSM. However, RPMF shows consistently for both error measures the smallest deviations with e.g., lowest medians and spreads of interquartile range (IQR). It is followed by RPMFSOBV, with just slightly higher deviations. The deviations of PMF are noticeably higher, and MF shows least favorable performance characteristics with e.g., highest medians and largest spreads of IQR. These results reflect also results of analysis I. The ability to identify terrain exhaustively is expressed in a low FN rate. In analysis I, RPMF showed lowest FN rates, followed by RPMF-SOBV and PMF (see Fig. V-5(a)).

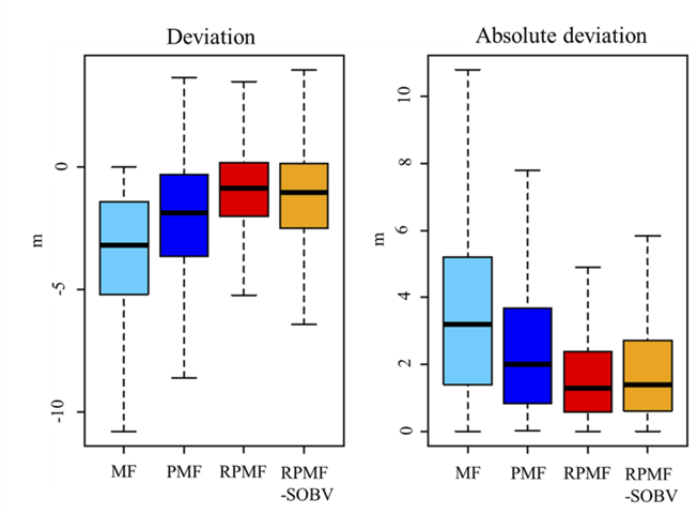


Fig. 6. Boxplots (depicting median, interquartile range (IQR), as well as 1.5 times IQR) for illustration of deviations, and absolute deviations of the DTMs with respect to BE pixels of the TanDEM-X data in meters.

Besides, we calculated mean error (ME), mean absolute error (MAE), root-mean-square error (RMSE), and linear deviations at the 90th percentile (LDP90) (results are shown in Tab. V-1). In accordance with the results depicted in Fig. 6, RPMF shows the lowest deviations with respect to MAE (1.76m), ME (-1.07m), and RMSE (2.5m). It is directly followed by RPMFSOBV, which shows slightly worse quality measures. PMF comes up with a MAE of 2.84m, a ME of -2.36m and a RMSE of 4.12m. These values are considerably above RPMF and RPMF-SOBV but more favorable compared to the values of MF, which shows highest deviations. LDP90 reveals that 90 percent of the deviations are smaller than 3.68m and 4.35m when using RPMF and RPMF-SOBV, respectively, compared to considerably higher deviations for PMF (6.21m) and MF (7.6m). Overall, the results of this analysis are consistent with results of Analysis I and suggest that RPMF features a superior ability to identify terrain with a comparatively low error of omission. Nevertheless, it is directly followed by RPMFSOBV. A more distinct decrease of performance can be observed when using PMF. MF generally shows least favorable accuracies.

Tab. V-2. MEAN ERROR (ME), MEAN ABSOLUTE ERROR (MAE), ROOT-MEAN-SQUARE ERROR (RMSE), AND LINEAR DEVIATIONS AT THE 90TH PERCENTILE (LD_{P90}) OF THE DIFFERENT APPROACHES IN METERS

approach	MAE	ME	RMSE	LD_{P90}
MF	4.12	-3.71	4.96	7.60
PMF	2.84	-2.36	4.12	6.22
RPMF	1.76	-1.07	2.50	3.68
RPMF-SOBV	2.07	-1.4	3.13	4.35

4.3. RESULTS OF ANALYSIS III: VISUAL ASSESSMENT OF NDSMS

The DTMs generated for the aforementioned analysis were subtracted from the TanDEM-X DSM to visually assess the quality of the resulting nDSMs. When values smaller than zero occurred in the resulting nDSM, these values were set to zero since they represent artifacts related to the DTM calculation method. “Non-urban” areas are faded out by integrating UF data.

Computed nDSMs are visualized in Fig. V-7(a) for focus areas in steep (1) as well as flat (2) terrain. Additionally, magnification (3) is intended to exemplify a situation, which motivates the use of RPMF-SOBV. It focuses on large buildings with irregular roof surfaces. The first column shows VHR optical imagery for comparison. Subsequent to the optical imagery, nDSMs based on MF, PMF, RPMF, and RPMF-SOBV are presented. It can be observed that the overall height levels of the objects differ significantly with MF showing the highest level and RPMF the lowest. When comparing the individual nDSMs to the optical imagery, it is obvious that many pixels that represent bare earth in reality have misleadingly a height value in the MF and PMF based nDSM. In contrast, nDSMs calculated based on RPMF and RPMF-SOBV look fewer affected and generally exhibit a crisper and less blurry appearance.

The overestimation of objects’ heights by MF and PMF can be dominantly observed in steep terrain (1), also partly for RPMF-SOBV. In the central part of the focus area coherent regions with exaggerated object heights appear. These are caused by BE pixels that were misclassified as OBJ pixels (the height difference to neighbor pixels is high due to the steepness of the terrain). In such a situation, the interpolation procedure cannot rely on sufficient BE pixels and thus underestimates the terrain heights. This causation is also revealed in the corresponding height profile (1) in part (b) of the figure (shaded relief with profile lines in yellow are shown in the last section of the upper part). Terrain heights of MF,

PMF, and RPMF-SOBV purport too low levels, whereas RPMF follows the actual terrain height very closely. In flat terrain (height profile 2 in (b)), a general characteristic of MF becomes obvious as this approach frequently produces terrain models with a too low height level. PMF follows the terrain height more closely and naturally. However, the profile reveals a fractional underestimation of the actual terrain heights. On the contrary, RPMF and RPMF-SOBV hardly underestimate terrain heights. The surface generated by RPMF partially even reflects a volatile behavior of height values of the DSM. This can be related to FP errors which cause regions in the interpolated surface that are elevated towards the DSM surface. Interestingly, the surface generated based on RPMF-SOBV appears to be less prone to this behavior. This is specifically illustrated in height profile 3 in (b). It can be observed that the terrain surface produced by RPMF is considerably elevated towards the DSM surface, what leads to underestimated height values in the nDSM as can be seen in section (a), magnification 3. Generally, this error source motivated the introduction of RPMF-SOBV. As can be seen from this example, such errors can be largely avoided with RPMF-SOBV and more valid surface estimations are achievable.

Overall, results are unambiguously in line with the results of analysis I and II, and reflect the capability of the approaches to distinguish bare earth from elevated objects correctly. In this manner, the level and proportions of FN and FP rates as evaluated explicitly in analysis I are directly reflected in the computed nDSMs. Hence, analyzes suggest that RPMF and RPMF-SOBV yield more favorable performance characteristics compared to PMF and MF. In that sense, RPMF features a distinct capability to identify bare earth pixels in an exhaustive manner, what is especially advantageous in steep terrain. Nevertheless, the method may perform less advantageous in situations where elevation values of elevated objects feature a volatile behavior. In such situations RPMF is prone to omit classifying OBJ pixels. This motivated the introduction of the post-classification processing scheme to spatially refine classification outcomes of the region growing procedure of RPMF. Analyses uncover a remarkable characteristic of RPMF-SOBV. It yields a favorable tradeoff between a significant decrease of omission errors while keeping errors of commission on a very low level. This makes the approach generally applicable in various settings.

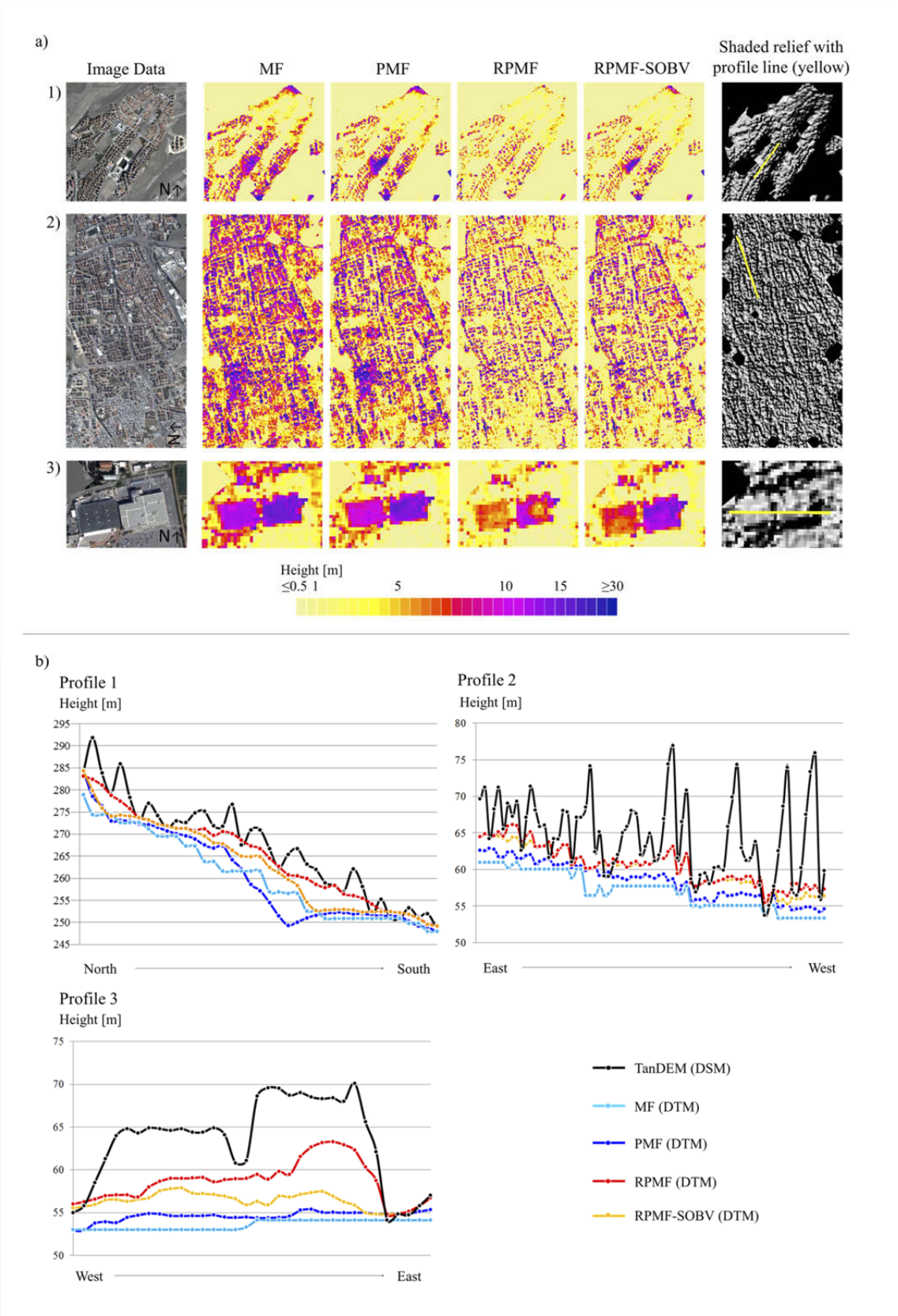


Fig. V-7. (a) Visual comparisons of nDSMs based on MF, PMF, RPMF and RPMF-SOBV with particular focus on (a) steep terrain, (b) flat terrain, and (c) large buildings. (b) Affiliated height profiles of DTM surfaces generated by the different approaches.

5. CONCLUSIONS

In this paper, we addressed the calculation of normalized Digital Surface Models from TanDEM-X data in urban environments. Numerous approaches to derive nDSM information have been postulated in scientific literature. However, distinct resolution characteristics of TanDEM-X hamper the use of many existing algorithms and pose individual challenges to a suitable approach. Hence, we jointly proposed a novel region growing-based progressive morphological filter procedure with a post-classification processing scheme to specifically allow for an accurate reconstruction of urban morphology even in challenging terrain. The former is based on a multistep procedure, which sequentially and iteratively executes progressive morphological image filtering and region growing to identify ground pixels. These are subsequently used for the interpolation of a DTM which allows normalizing the DSM. The post-classification processing scheme adapts techniques of object based image analyses to spatially refine regions of classified non-ground pixels.

Experiments were carried out with Intermediate Digital Elevation Model data for an area that covers large parts of Izmir (Turkey). Results confirm the interest of the proposed approaches and reveal beneficial performance characteristics compared to basic morphological filter based approaches. This is evident especially in terrain with a high orographic energy and steep areas. In particular, when separating ground from non-ground pixels, RPMF features a remarkable decrease of omission errors which exceeds an increase of commission errors compared to a basic progressive morphological filtering procedure. The decrease of omission errors of RPMF-SOBV is more moderate compared to RPMF, however, the increase of commission errors remains simultaneously on a very low level. This makes the approach specifically suitable for an application in various settings.

The TanDEM-X mission will deliver a globally consistent digital surface model with an unprecedented spatial resolution. Having suitable methods available for an automated extraction of objects above ground in urban environments will open a broad range of areawide applications. These are related to the analysis and monitoring of urbanization processes, characterization of urban morphology, and natural disaster mitigation, among others. Hence, valuable contributions to different fields of research can be achieved when extracting relevant information based on presented data and methods.

6. APPENDIX A

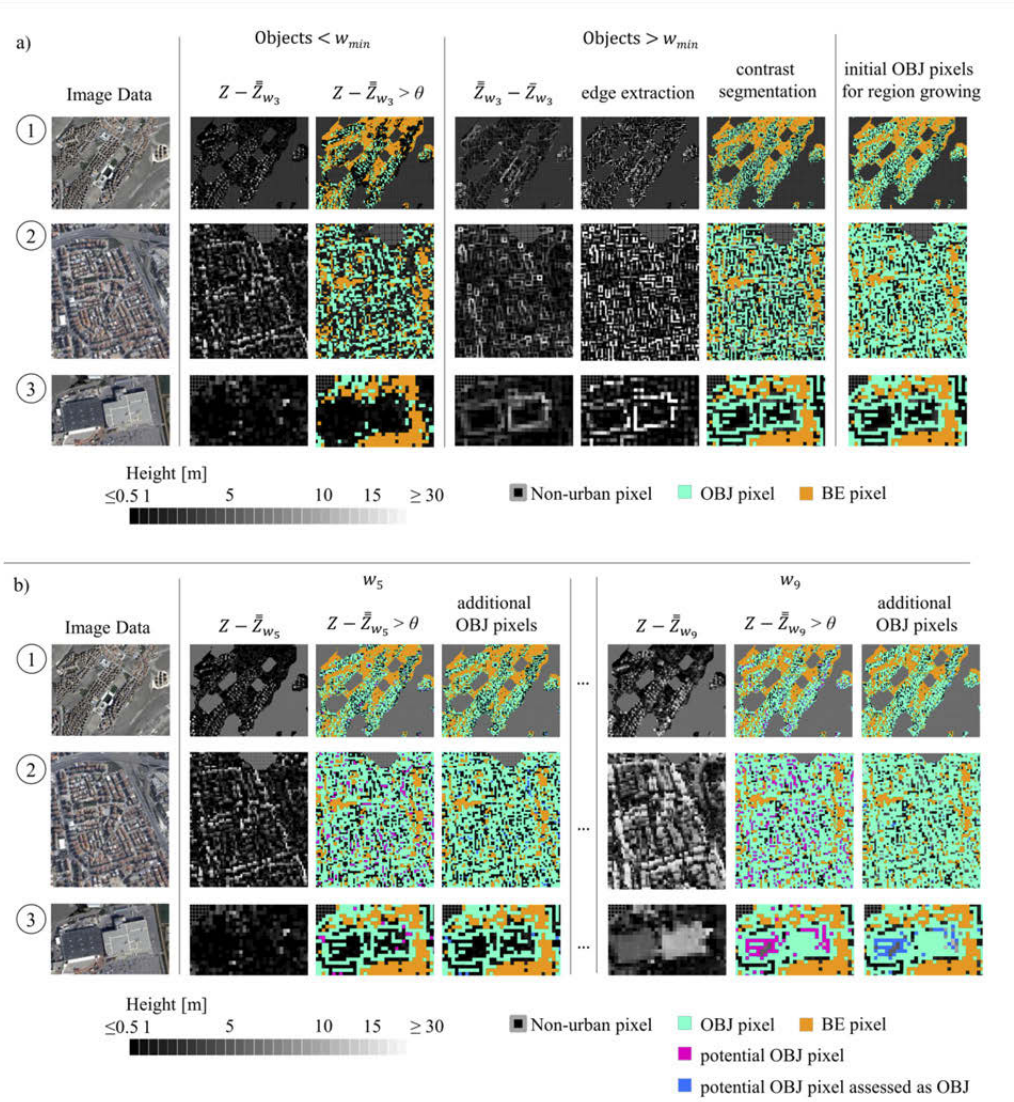


Fig. V-8. Exemplification of the procedure for the identification of initial OBJ pixels (a) that represent the basis for the region growing procedure (b) with IDEM TanDEM-X data. a) Objects that are exceeded by the minimum window size (w_{min}) can be identified by the application of a threshold θ on the pnDSM ($Z - \bar{\bar{Z}}_{w_{min}}$). Border pixels of objects that are not exceeded by w_{min} are identified by subtracting $\bar{Z}_{w_{min}}$ from $\bar{\bar{Z}}_{w_{min}}$. They are classified as OBJ by combining edge extraction filter and segmentation. These initial OBJ pixels are used for the region growing procedure shown in (b). The size of the structuring element w is increased and additional OBJ pixels are identified if they exceeded θ in the respective newly calculated pnDSM and fulfill a similarity constraint with respect to already classified OBJ pixels. In this example free parameters θ and γ were set to 2.6m and 0.8m, respectively.

7. APPENDIX B

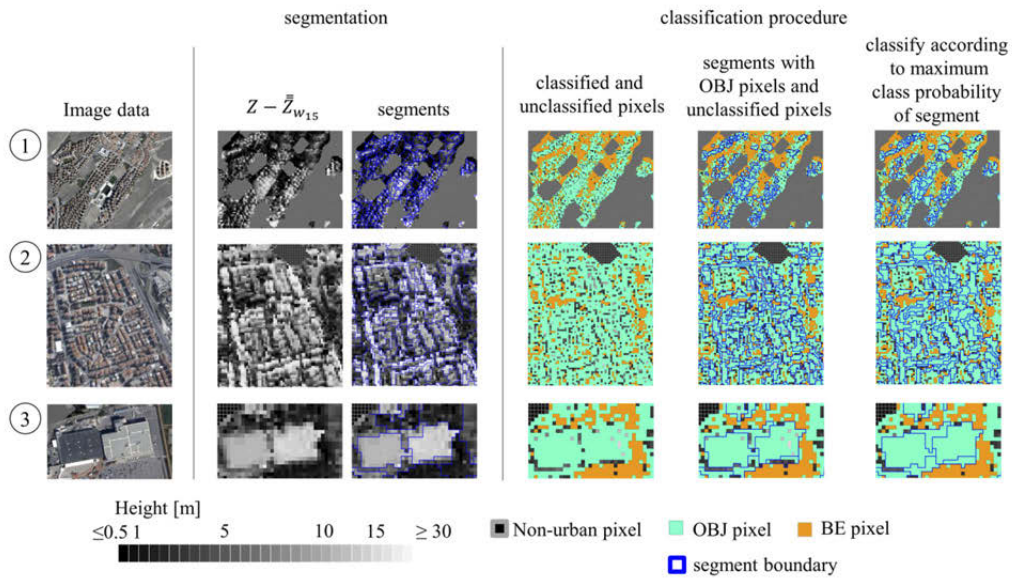


Fig. V-9. Exemplified selective post-classification processing scheme. First, a preliminary nDSM is calculated with w_{max} to ensure that all building objects are contained. A segmentation algorithm is used to discriminate homogenous image regions. Segments are selected that contain solely OBJ pixels and unclassified pixels after the region growing procedure of RPMF. Pixels of selected segments are classified according to the maximum class probability of the respective segment. In this example, w_{max} corresponds to 15 pixels, and free parameters θ and γ were set to 2.6m and 0.8m, respectively.

ACKNOWLEDGMENTS

The research was founded partially by Helmholtz-EOS (Earth Observation System). Additionally, we want to acknowledge the support by the European Commission's Seventh Framework Programme [FP7/2007-2013], under grant agreement no. 312972 "Framework to integrate Space-based and in-situ sENSing for dynamic vUlnerability and recover Monitoring". We thank Andreas Schmitt (DLR) for remarks on the initial manuscript and the anonymous reviewers for their helpful comments.

Chapter VI

Seismic vulnerability assessment of urban structures with multi-sensor remote sensing

Submitted manuscript

Christian Geiß, Marianne Jilge, Tobia Lakes and Hannes Taubenböck

ABSTRACT

The impact of natural disasters such as earthquakes on mankind has increased dramatically over the last decades. Global urbanization processes and increasing spatial concentration of exposed elements induce seismic risk at a uniquely high level. The increase of urban population has occurred during a time period that is comparatively short with respect to the return time of severe earthquakes. Earthquakes that had little impact in the past, when they hit sparsely populated and spatially fragmented settlement areas, will nowadays shake urban agglomerations with millions of people. To mitigate this situation determines detailed knowledge about seismic risk. In particular, it is crucial to have information about the building inventory and its behavior with respect to a certain level of ground shaking.

Numerous studies emphasize that remote sensing can play a valuable role in supporting the extraction of relevant features for pre-event vulnerability analysis of urban environments. However, majority of approaches operate on building level. This induces the deployment of very high spatial resolution remote sensing data what hampers nowadays utilization capabilities for larger areas due to data costs and processing requirements. In this manuscript, we alter the spatial scale of analysis and propose concepts and methods to assess homogeneous urban structures. A procedure is designed which comprises four main steps dedicated to: i) delineation of urban structures by means of a tailored unsupervised data segmentation procedure with scale optimization; ii) characterization of urban structures by a joint exploitation of multi-sensor data; iii) selection of most feasible features under consideration of *in situ* vulnerability information; iv) estimation of seismic vulnerability levels of urban structures within a supervised learning framework.

Experimental results obtained for the earthquake-prone mega city Istanbul confirm viability of procedures. When estimating damage grades for Istanbul's district Zeytinburnu with a support vector regression approach, best models are characterized by mean absolute percentage errors less than 11% and fairly strong goodness of fit ($R > 0.75$). When aiming to identify types of urban structures (i.e., urban structures determined by large industrial/commercial buildings that can be considered as highly vulnerable, and urban structures determined by tall detached residential buildings that can be considered as slightly vulnerable), results obtained with soft margin support vector machines showed distinctive increase of accuracies when compared to results obtained with ensembles of one-class support vector machines. Ensembles of one-class support vector machines were not able to exceed moderate agreements, with κ statistics slightly above 0.45. Instead, soft margin support vector

machines allowed to obtain κ statistics showing substantial and even excellent agreements ($\kappa > 0.6$ up to $\kappa > 0.8$). Analyzes provide promising empirical evidence, which confirms the potential of remote sensing to support the seismic vulnerability assessment of urban structures.

1. INTRODUCTION

The impact of natural disasters such as earthquakes on mankind has increased dramatically over the last decades. Global urbanization processes and increasing spatial concentration of exposed elements such as people, buildings, infrastructure, and economic values in earthquake prone regions induce seismic risk at a uniquely high level. Thereby, the increase of urban population has occurred during a time period that is comparatively short with respect to the return time of severe earthquakes. Earthquakes that had little impact in the past, when they hit sparsely populated and spatially fragmented settlement areas, will nowadays shake urban agglomerations with millions of people. This situation, when left unmitigated, is expected to cause unprecedented death tolls, enormous economic and ecological losses, critical infrastructure and service failures, and poses a significant threat for civil security, and a sustainable development in the future (Bilham, 2009; Holzer and Savage, 2013; Tucker, 2013). To mitigate those perils requires detailed knowledge about seismic risks. As an important constituent element of seismic risk, the seismic vulnerability of the built environment needs to be assessed. In particular, it is crucial to have information about the building inventory and its behavior with respect to a certain level of ground shaking (Erdik et al., 2003).

Generally, conventional approaches to assess the seismic vulnerability of the building inventory, which incorporate an exclusive application of detailed *in situ* building-by-building analysis by structural engineers, may provide very detailed and high quality vulnerability information, but are decreasingly able to cope with the high spatiotemporal dynamics of urban environments. On the contrary, information collected on a very broad spatial level such as spatially aggregated census data hampers the consideration of small-scale hazard effects in a downstream risk model (Wieland et al., 2012). Hence, building inventory data and affiliated seismic vulnerability information is often outdated, spatially aggregated and discontinuous, and in many earthquake prone regions of the world simply not existent.

Remote sensing has already proven its great potential to extract relevant features for pre-event vulnerability analysis of buildings. The intrinsic advantage of remote sensing is the ability to offer an overview of building stocks and serve as a screening method for derivation of building vulnerability related features, such as shape characteristics, height, roof material, period of construction, structure type and spatial context (Geiß and Taubenböck, 2013). This topic is subject to a lively research and has gained much scientific contemplation in the past few years. Thereby, the research environment is constituted primarily by two distinctive

science communities: remote sensing and earthquake engineering. Nowadays, concepts and methods have reached a level, where they are found to be relevant and being accepted in both communities. This is evident since the following studies are published in established journals of both communities. Taubenböck et al. (2009), and Borzi et al. (2011) characterize the built environment by means of remote sensing data and retrieve specific fragility functions for designated building types. Borfecchia et al. (2009) assess the vulnerability of buildings in a hybrid way, by combining *in situ* ground truth for selected buildings with information derived from remote sensing data. Supervised classification techniques are subsequently applied to classify the residual building inventory. Geiß et al. (2014 a, 2014 b) combine detailed *in situ* seismic vulnerability information with features describing the urban morphology derived from remote sensing data. Supervised regression and classification techniques are then applied to evaluate the suitability for an area-wide seismic vulnerability assessment and earthquake loss estimation. All cited studies present promising empirical evidence with respect to the viability of the approaches. However, from a conceptual point of view, all approaches operate on building level. This induces the deployment of very high spatial resolution remote sensing data what hampers utilization capabilities for larger areas due to data costs and processing requirements. When aiming at spatially continuous and consistent assessment approaches that are applicable for large regions, on a national level, or even globally, those data represent nowadays a clear limitation.

To alleviate these restrictions, analyzes can be addressed at a coarser level of the urban morphology. From a conceptual point of view, urban environments are perceived as systems that exhibit a hierarchical arrangement (Banzaf and Höfer, 2008). At the lowest level of aggregation, individual objects such as buildings can be considered, followed by their homogeneous assemblage at block level (i.e., urban structures), and end with a spatially indifferent urban area at the highest level of aggregation. In this manuscript, we alter the spatial scale of analysis compared to the aforementioned studies and propose concepts and methods to assess vulnerability for physically homogeneous urban structures. Thus, a coarser level of urban morphology is addressed when compared to building level. However, this allows relying on remote sensing data with a lower spatial resolution but larger spatial coverage.

This idea goes along with the concept of urban structure types (UST). UST represent distinct and homogeneous spatial entities in terms of physical arrangement of their constituent elements such as surface materials, affiliated environmental characteristics (e.g., micro-climate), or functional properties such as land use (Heiden et al., 2012; Pauleit and Duhme,

2000). From a remote sensing perspective, the concept of urban structure types has been exploited within different applications. Previous studies emphasize the viability of this concept in order to e.g., identify distinctive settlement types (Niebergall et al., 2008; Wurm et al., 2009; Baud et al., 2010), classify urban biotopes (Bochow et al., 2007), support urban (micro-)climate modeling (Heldens et al., 2012; Bechtel and Daneke, 2012), or monitor urban dynamics (Banzhaf and Höfer, 2008). As can be seen from the cited studies, definitions about target class(es) to be mapped and characterized are very heterogeneous, what render "urban structure types" rather a concept then an accepted typology with defined specification. Hence, we use the term urban structures to generally describe distinctive and homogeneous assemblages of land cover/land use elements.

In the context of seismic vulnerability research already Mueller et al. (2006) note that urban structures can be helpful in order to assess the seismic vulnerability of buildings, since they represent a characteristic grouping of idealtypic building types and the spatial context they are embedded in. In this manner, Wieland et al. (2012) and Pittore and Wieland (2013) use multitemporal Landsat data to discriminate homogeneous urban structures based on an image segmentation approach and semantically annotate them by utilizing a supervised classification scheme. Derived urban structures are intended to serve as strata for a more detailed analysis of the building stock with VHR optical data and a ground-based omnidirectional imaging system. The sensed information is combined with ancillary information (i.e., information from the world housing encyclopedia) for a subsequent probabilistic seismic vulnerability assessment. Hence, the approach delineates and determines urban structures without prior knowledge and assesses the seismic vulnerability *a posteriori*.

In contrast to that, here, we suggest a procedure, which learns prior seismic vulnerability knowledge for urban structures. Hence, the primal objective of this manuscript is to introduce an original procedure for seismic vulnerability assessment of urban structures with earth observation data. In particular, we combine multispectral data and elevation measurements in order to characterize the urban morphology and use techniques of object based image analysis (Pal and Pal, 1993; Blaschke, 2010) and statistical learning (Chen and Ho, 2008; Camps-Valls et al., 2014) to assess the vulnerability of urban structures under consideration of *in situ* information.

Generally, a number of remote sensing systems appear to be promising for the characterization of urban structures. In this manuscript, we exploit multispectral and elevation data, which feature a lower spatial resolution compared to systems with the highest spatial resolution available nowadays (e.g., QuickBird-2, Geoeye-1, WorldView-2, etc.), but allow at

the same time a larger spatial coverage. In particular, we use data from the RapidEye constellation (Tyc et al., 2005), Tandem-X mission (Krieger et al., 2007), and the Landsat archive (Tucker, 2004). The RapidEye constellation consists of five equally designed satellites, which operate in a single sun-synchronous orbit. The system appears to be promising for area-wide analyses of urban structures, since it offers a high spatial resolution of 6.5m (pixel size), in conjunction with a swath width of 77 km, and a maximum acquisition capacity of 1500 km per orbit (Tigges et al., 2013; also missions such as e.g., Sentinel-2 (Drusch et al., 2012) can be considered as an interesting opportunity with respect to a suitable characteristic of spatial resolution and swath size). TanDEM-X is a spaceborne radar interferometer, which acquires data for a seamless global digital surface model (DSM) with an unprecedented, globally consistent spatial resolution of ~0.4 arcseconds. This allows resolving objects that are elevated from the earth's surface in urban environments such as buildings. Lastly, data from the Landsat archive, which dates back to 1972, is considered since they enable a characterization of spatiotemporal developments of urban structures (Taubenböck et al., 2012).

Based on the remote sensing data and subsequently derived information layers, a procedure is designed which comprises four main steps dedicated to: i) delineation of urban structures by means of tailored unsupervised data segmentation procedure with scale optimization; ii) characterization of urban structures by a joint exploitation of multi-sensor data; iii) selection of most feasible features under consideration of *in situ* vulnerability information; iv) estimation of seismic vulnerability levels of urban structures within a supervised learning framework.

Three challenging experiments are designed, which are intended to demonstrate feasibility of procedures and discuss viability of results. Thereby, we relate our experimental analyses to Istanbul (Turkey), since the city faces an enormous seismic threat (Erdik et al., 2003). In addition, *in situ* observations are available for relevant parts of the city. The remainder of the paper is organized as follows. Section 2 gives an overview on study area, utilized remote sensing data and reference information. Section 3 presents developed and deployed methods. We use section 4 to give an overview on the concept and objective of designed experiments and report and discuss outcomes of experiments in section 5. Concluding remarks are given in section 6.

2. STUDY AREA AND DATA

2.1. ISTANBUL, TURKEY

Istanbul (41° 1` N, 28° 58` O) has about 14.2 million inhabitants (TurkStat, 2014), what constitutes it a mega city. Some sources estimate the number of people living in Istanbul to be even higher, since numerous people are not registered and, thus, do not appear in official statistics. Massive migration during past decades induced a complex and erratic city structure, which comprises small and low-rise informal settlements (i.e., Gecekondular), large industrial buildings, and high-rise residential and commercial buildings, among others. About one fifth of the building inventory was built within the past 15 years (Cakti, 2013). Besides, many existing buildings were expanded without permission by new stories and extensions, what alters structural properties. Moreover, authorized real estate projects were already modified during the construction phase. It is estimated that the majority of buildings in Istanbul does not correspond to any structural standards and lack regulatory supervision. Hence, it is assumed that most of the buildings are not shake resistant, although a law to enforce resistant design was issued in 1999.

Almost all earthquakes in Turkey are related to tectonic movement (Erdik et al., 1999), and Istanbul is located close to the North Anatolian Fault (NAF; Fig. VI-1a). The NAF represents an active right-lateral strike-slip fault. Previous earthquakes along the fault indicate a westbound regime (Stein et al., 1997) and give rise to a possible earthquake close to Istanbul. According to Parsons (2004), a 35-70% probability for the occurrence of an earthquake with a magnitude >7 along the NAF within the next 30 years is expected.

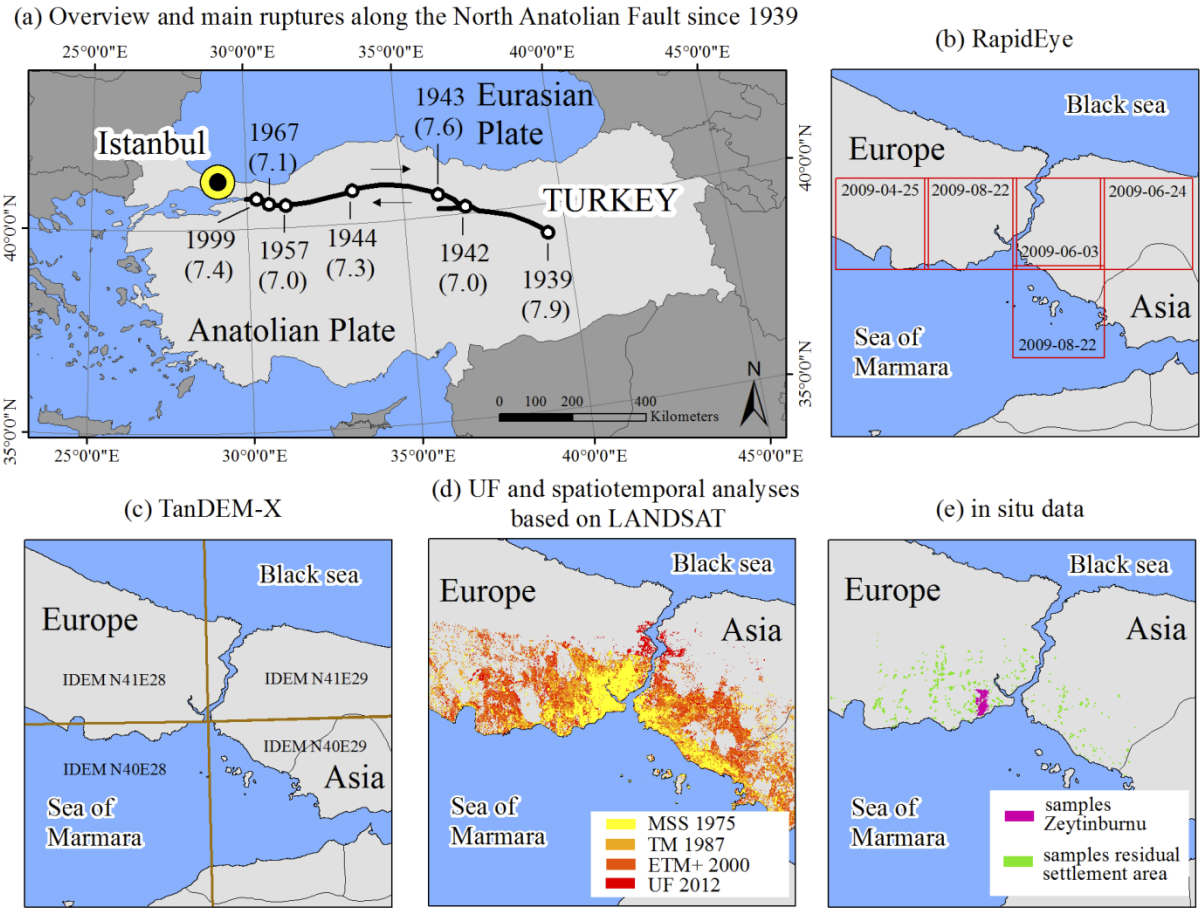


Fig. VI-1. Overview on the location of the study area, tectonic setting and acquired data; (a) location of Istanbul and main ruptures along the North Anatolian Fault since 1939, which indicate a westbound regime; (b) acquired RapidEye data; (c) Tiles of TanDEM-X Intermediate DSM data; (d) *in situ* information for the city district Zeytinburnu and for residual settlement areas in Istanbul.

2.2. REMOTE SENSING DATA: RAPIDEYE, TANDEM-X, AND LANDSAT

We acquired five RapidEye images which were collected between April and August 2009. They cover the main parts of the settlement area of Istanbul (Fig. VI-1b). Cloud coverage was minimal and imagery was delivered as level 2A. Hence, data were subject to geometric and radiometric sensor corrections and were calibrated to at-sensor radiances with a pixel size of 6.5m at nadir. In the experiments, we deploy the blue (0.440-0.510 µm), green (0.520-0.590 µm), red (0.630-0.685 µm), and nir (0.760-0.850 µm) band, and neglect RapidEye's specific red edge band (0.690-0.730 µm). We want to rely on image bands that are frequently available for multispectral scanners and thus render methods and results more transferable and ubiquitous.

The TanDEM-X satellite was launched in June 2010 and is operating jointly in a helix tandem formation with its twin radar satellite TerraSAR-X, which is in space since June 2007. Both satellites are operationally acquiring data to generate a seamless global multi-coverage

digital elevation model using single-pass interferometry in bi-static mode (Krieger et al., 2007). Since August 2011 the data is being processed operationally. Single- and multi-baseline interferometric phase-unwrapping, manual quality inspection, water detection, height comparison with ICESat data, and calibration including tilt and offset correction is carried out to provide a homogenous dataset for a regional mosaic. In our study, four tiles of the so-called TanDEM-X Intermediate Digital Elevation Model (IDEM) are used (Tile N40E028, N40E029, N41E028, N41E029; Fig. VI-1c). The IDEM consists only of the best quality single-baseline processed data of the first global coverage. Insufficient acquisitions affected by phase-unwrapping errors are excluded. Data sets cover the city and its hinterland. The spatial resolution of the data of 0.4 arc seconds corresponds to ~12 m in Istanbul.

For spatiotemporal analysis we use multitemporal LANDSAT data, which were acquired by the Multispectral Scanner (MSS), Thematic Mapper sensor (TM), and the Enhanced Thematic Mapper sensor (ETM+) in 1975, 1987, and July 2000, respectively. Images are constituted by 4 (MSS) and 7 (TM, ETM+) multispectral bands covering a spectral range of 0.500-1.100 µm (MSS) and 0.450-2.350 µm (TM, ETM+) at 79 m and 30 m spatial resolution, respectively.

2.3. IN SITU DATA: SEISMIC BUILDING VULNERABILITY REFERENCE INFORMATION

Two different kind of in situ information were available for the experiments. (i) For Istanbul's district Zeytinburnu (40° 59` N, 28° 54` O), a spatially continuous assessment of the buildings according to the capacity spectrum method was available. The approach relates the capacity of the structure (in the form of a pushover curve) with the seismological impact on the structure (in the form of response spectra). The intersection of the two functions approximates the response of the structure (Freeman, 2004). Thus, it allows quantifying expectable damage grades. The necessary *in situ* information to adapt the functions for Zeytinburnu was gathered and provided by Taubenböck et al. (2009). (ii) In addition, ground-based GPS-photos for large parts of the settlement area were available from Google Panoramio™. Generally, such information can support a rapid visual screening assessment of buildings (FEMA, 2002). Here, the information was used to identify urban structures according to distinctive types as identified by the capacity spectrum method. Please find a more detailed description of procedures in section 4.1., where we describe the experiments.

3. METHODS

Presented data are subject to a multistep procedure to assess seismic vulnerability of urban structures. A block scheme of the procedure is reported in Fig. VI-2. First, remote sensing data are pre-processed. The multispectral data are atmospherically corrected by means of the software tool ATCOR-2 (Richter, 2008), and a mosaic is generated from the different images (section 3.1.). Data from the TanDEM-X mission serves for both the derivation of a settlement mask (i.e., urban footprint; section 3.1.1.), and the computation of a normalized digital surface model (nDSM; section 3.1.2.). Landsat imagery are deployed for spatiotemporal analyzes, and are subject to a post classification change detection procedure (section 3.1.3.). Subsequently, four main steps are carried out within a supervised learning framework: i) section 3.2. is used to describe an unsupervised image segmentation procedure with scales optimization to allow for an objective delineation of urban structures; ii) section 3.3. details multi-scale calculations of features from both the multispectral data and height information from the nDSM for a comprehensive characterization of urban structures; iii) in section 3.4. features are grouped according to different underlying remote sensing data and segmentation scales to better understand the respective value within the process of supervised learning; in addition, supervised feature selection techniques are deployed to possibly identify the most suitable feature sets for model learning; iv) lastly, in section 3.5. we reveal three different formulation of support vector machines (SVM), for multi- and one-class problems, and function estimation within a supervised learning framework, which were used in the experiments.

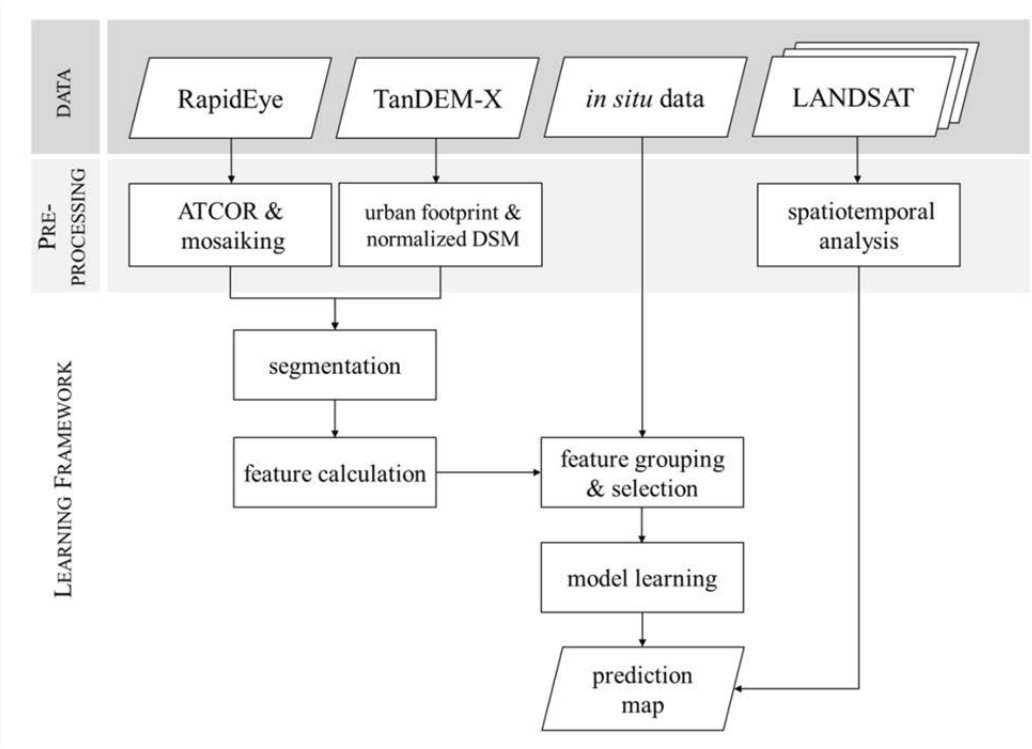


Fig. VI-2. Block scheme of the proposed procedure.

3.1. PRE-PROCESSING

Multispectral RapidEye data were subject to atmospheric corrections with ATCOR-2 (ibid.) to account for radiometric distortions. The five images were radiometrically adjusted with a histogram match and merged to a spatially continuous mosaic. Generally, data (i.e., RapidEye and TanDEM-X imagery) have been resampled (by means of a nearest neighbor interpolation) and properly coregistered to a common spatial resolution of 5 m.

3.1.1. DERIVATION OF SETTLEMENT MASK FROM TANDEM-X

To focus analyses on urban environments, we deploy a fully automated image analysis procedure to discriminate "built-up" and "non-built-up" land cover (referred to as "urban footprint – UF"; Esch et al., 2013). Based on the TanDEM-X mission, which collects two global data sets of very high resolution synthetic aperture radar images between 2011 and 2013, built-up areas can be delineated globally with an unknown spatial detail. The binary layer is generated with an unsupervised classification scheme accounting for both the original backscattering amplitude and extracted texture characteristics. A high classification accuracy, which exceeds consistently an overall accuracy of 94% and a κ statistic of 0.75 for representative case studies (ibid.), ensures a reliable delineation of built-up areas. The original

urban footprint data were subject to some generalization procedures in this study. We deployed morphological operators (i.e., closing was performed by sequentially carry out dilation and erosion; Haralick et al. 1987) to ensure a spatially non-fragmented representation of the settlement area of Istanbul. The utilized urban footprint of Istanbul covers an area of 501.89 km^2 and is visualized in Fig. VI-1d.

3.1.2. CALCULATION OF NORMALIZED DSM FROM TANDEM-X

Height characteristics of urban structures are important features with respect to seismic vulnerability. We use elevation measurements of TanDEM-X IDEM to compute a nDSM, which comprises elevation information of objects above ground. To this purpose, a digital terrain model (DTM) is derived from the data first with a region growing-based progressive morphological filter procedure (see Chapter V). This approach was recently proposed to address general challenges associated with the use of morphological filters in non-flat terrain, and overcome individual challenges related to the spatial resolution of TanDEM-X data. It comprises a multistep procedure using concepts of morphological image filtering, region growing and interpolation techniques. It is based on the idea of progressive morphological filters that aim to discriminate ground and non-ground pixels in the digital surface model based on algebraic set operations. Some free parameters need to be determined when using this approach. To ensure extraction of all buildings present in the area under investigation, the structuring element must always exceed a building's outline. The side length of the structuring element was determined empirically for eight individual subsets of the study. In addition, an elevation difference threshold and similarity constraint needs to be fixed. Both were set according to previous experimental analyses, to ensure a favorable tradeoff between decrease of omission errors and increase of commission errors when classifying ground pixels (i.e., 2.6 m for elevation difference threshold and 0.8 for similarity constraint). Identified ground pixels are interpolated to a DTM based on Inverse Distance Weighting (Shepard, 1968). Lastly, the DTM was subtracted from the DSM to receive the final nDSM.

3.1.3. SPATIOTEMPORAL ANALYSIS WITH LANDSAT

To uncover the spatiotemporal evolvement of urban structures in Istanbul, we integrate results from a post classification change detection procedure provided by Taubenböck et al. (2012). Based on LANDSAT imagery, the settlement area is classified (Fig. VI-1d). The approach deploys spectral classification in conjunction with a temporal hierarchical scheme. This means that the classified settlement area for a past time step is used as a spatial condition

when classifying urbanized areas for the more current time steps. The individual classification feature viable accuracies. The overall accuracy for 1975 is 93.7% (κ = 0.82), 92.4% (κ = 0.79) for the year 1987, and 90.8% (κ = 0.82) for 2000, respectively.

Data form the Landsat archive are used quite frequently to estimate the period of construction of built environments in seismic vulnerability studies (e.g., Taubenböck et al., 2008; Borfechia et al., 2009, Wieland et al., 2012). Here, we assign the estimated period of construction posterior to model learning (see Fig. VI-2 in this section and Fig. VI-10 in the experimental results section). This reduces the computational burden compared to an assignment prior to model learning, since this way a model must be learned based on labeled samples for every class of interest and time step. However, we simultaneously ensured a stratified collection of labeled samples across areas of all estimated periods of construction.

3.2. SEGMENTATION PROCEDURE FOR DELINEATION OF URBAN STRUCTURES

This section is dedicated to the description of an unsupervised segmentation procedure with scale optimization for delineation of urban structures. In contrast to natural environments, urban man-made structures have been identified as few examples of objects within a landscape with distinct and crisp boundaries, which feature often also an irregular shape (Herold et al., 2003). This makes the utilization of object based image analysis techniques feasible and especially the latter limits the use of uniform spatial entities such as quadratic objects. Image segmentation represents the basis for object based images analysis and aims at the delineation of meaningful real-world objects (Blaschke, 2010). In this study, segmentation focuses on the distinction of intra-urban areas, which are homogeneous in terms of their response in the multispectral and elevation data. For a joint exploitation, a brightness layer is computed first

$$B = \frac{1}{n_v} \sum_{i=1}^{n_v} z_{i(v)} \qquad \text{(VI-1)}$$

with $z_{i(v)}$ being the visible bands, NIR, and nDSM, and n_v the number of bands (i.e., 5). Segments represent basic spatial units to delineate and characterize homogeneous urban structures later on. We deploy multiresolution segmentation based on fractal net evolution (*FNEA*) approach (Benz et al., 2004). This is a bottom-up region-growing segmentation algorithm. It starts from individual pixels and merges pixels in dependence of user-defined constraints related to spectral and geometrical properties of modelled segments. Based on this technique, several automated routines are carried out in a sequential way to ensure parameterization and outcomes of optimized multiresolution image segmentation to be

objective, data-driven and generic. A schematic overview of the segmentation procedure is given in Fig. VI-3.

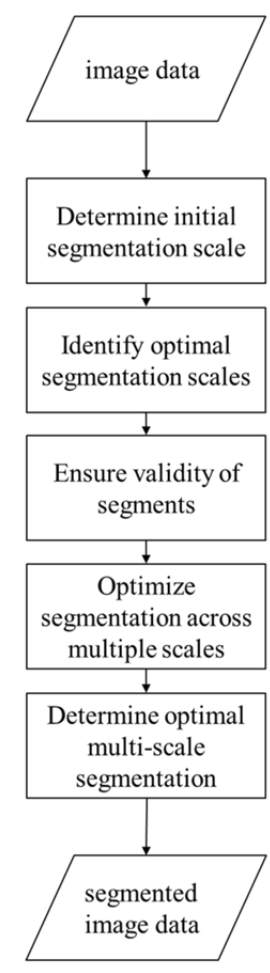


Fig. VI-3. Schematic processing steps of the segmentation procedure.

3.2.1. SEGMENTATION

First, an initial image level is built, which includes all "built-up" areas from the UF data set (Section 3.1.1.; exemplifications are provided in Appendix A, Fig. VI-13a). The subsequent segmentation is limited to areas identified as "built-up". The multiresolution segmentation approach is controlled by scale, color, and shape, with shape being composed of compactness and smoothness. The scale parameter is a function of the geometric resolution of the image data and defines the maximum allowable heterogeneity of modeled segments, with a lower scale parameter resulting in a higher number of segments. We suggest to put more emphasize on shape heterogeneity rather than on grey-value heterogeneity. This is due to the fact that man-made features such as urban structure have distinct shape and size properties, unlike e.g., natural features. Analogously, weights for heterogeneity of smoothness and compactness can

be maintained equally. However, the main difficulty lays in the determination of a scale which is suitable to represent different kinds of urban structures adequately. In this manner, the usage of a single scale is prohibited, since urban structures feature several orders of magnitude of spatial extend. Thus, we adapt a multilevel segmentation procedure proposed by Esch et al. (2008), which compares the distinctiveness of segments generated at multiple scales. Prior to this, we deploy the objective function introduced by Espindola et al. (2006) to find scene-specific optimal segmentation scales objectively. Based on the assumption that optimal segmentation maximizes intrasegment homogeneity and intersegment heterogeneity, a measure is calculated by incorporating intrasegment variance and Moran's *I*.

During the procedure, we suggest to maintain constant weights for color and shape criteria and create segments at ascending scales in the interval [h_s, 150]. Thereby, h_s represents the initial scale. In order to determine an appropriate initial scale and ensure that generated segments can represent valid delineations of urban structures, some constraints are introduced. In this sense, we define that valid image segments must correspond to the Shannon sampling theorem. This theorem states that modeled segments should be of the order of one tenth of the dimension of the sampling scheme – the pixel – to ensure that they will be completely independent of their random position and their orientation in relation to the sampling scheme (Blaschke, 2010). Thus, we define that there should be at least ten pixels to represent a valid segment in terms of size, and five pixels to represent a valid segment in terms of width

$$valid\ segment = \begin{cases} 1, \text{if } \dfrac{\sqrt{A}}{c} \cdot \dfrac{1}{10} > 1 \wedge S < 3 \\ 1, \text{if } \dfrac{d}{c} \cdot \dfrac{1}{5} > 1 \wedge S \geq 3 \\ 0, else \end{cases} \tag{VI-2}$$

where A is the area of a segment, c represents the pixel size of the image data, d is the width of a segment, and S a shape complexity index calculated as perimeter to boundary ratio

$$S = \frac{q}{2 \cdot r \cdot \pi}, \qquad r = \sqrt{\frac{A}{\pi}} \tag{VI-3}$$

where q is the perimeter of a segment, and r is the radius of circle with the same surface area. Analogous to Hengl (2006), we chose a threshold of 3 to differentiate between compact ($S < 3$) and narrow/long ($S \geq 3$) segments. Generally, a suitable h_s should include a comparatively small share of non-valid segments and represent a local maximum regarding

the decline in shares of non-valid segments. This principle and the selection of h_s for our application is illustrated in the experimental setup section 4.2. (Fig. VI-8a).

Following the aforementioned heuristic of Espindola et al. (2006), for each of the created scales in the interval [h_s, 150] the intrasegment variance (σ^2) with respect to the brightness layer was calculated

$$\sigma^2 = \frac{\sum_{i=1}^{n} A_i {\sigma^2}_i}{\sum_{i=1}^{n} A_i} \tag{VI-4}$$

where A_i and ${\sigma^2}_i$ represent the area and intrasegment variance of segment *i*. The intrasegment variance σ^2 is the weighted average, with the areas of each segment being the weights. As a measure of intersegment heterogeneity Moran's *I* is used

$$I = \frac{N}{\sum_i \sum_j e_{ij}} \cdot \frac{\sum_i \sum_j e_{ij}(\mu B_i - \mu\bar{B})(\mu B_j - \mu\bar{B})}{\sum_i (\mu B_i - \mu\bar{B})^2} \tag{VI-5}$$

where *N* is the number of segments indexed by *i* and *j*; μB is the mean brightness of a segment; $\mu\bar{B}$ is the mean brightness of all segments; and e_{ij} is the spatial weight between segments *i* and *j*, implemented in our case analogous to a queen contiguity as follows

$$e_{ij} = \begin{cases} 1, & \text{if } i,j \text{ are adjacent neighbour segments} \\ 0, & \text{else.} \end{cases} \tag{VI-6}$$

With the determined variance and autocorrelation measure the objective function is calculated by summing up normalized values of σ^2 and *I*

$$F(\sigma^2, I) = \frac{{\sigma^2}_{max} - \sigma^2}{{\sigma^2}_{max} - {\sigma^2}_{min}} + \frac{I_{max} - I}{I_{max} - I_{min}}. \tag{VI-7}$$

Generally, the objective function is calculated for ascending scales. Thereby, σ^2 values represent a monotonic function: for the sequel $({\sigma^2}_n)_{n\in\mathbb{R}}$ the condition ${\sigma^2}_{n+1} \geq {\sigma^2}_n$ is fulfilled. The maximum value of function $F(\sigma^2, I)$ is a statistical indicator of optimal segmentation (section 4.2., black arrow in Fig. VI-8b). As discussed above, a single optimal scale is not sufficient to take into account the relationship between the spatial structure of an image and the structure of urban environments. Thus, we deploy a plateau objective function $F(u)$, which was introduced by Martha et al. (2011), in order to obtain multiple optimal scales. The function is defined as

$$F(u) = F(\sigma^2, I)_{max} - \sigma \tag{VI-8}$$

where $F(\sigma^2, I)_{max}$ represents the maximum value of the objective function, and σ is the standard deviation of the function calculated for all scales. Optimal segmentation scales should be located above the plateau function, since these scales have high external and low internal heterogeneity levels, and the balance between under- and oversegmentation is still present (ibid.). As initial segmentation scale for the subsequent multilevel optimization procedure, we select the segmentation of the first scale above the plateau objective function (Section 4.2., black arrow in Fig. VI-8b; exemplified in Appendix A, Fig. VI-13b, ioMRS).

3.2.2. OPTIMIZATION PROCEDURE

Segments that do not represent valid segments according to the Shannon sampling theorem criteria defined above (exemplified in Appendix A, Fig. VI-13c-I, nvS) are merged with adjacent neighbor segments that have the smallest mean brightness difference of all adjacent neighbor segments (Fig. VI-13c-II, ioMRS-m). Subsequently, a second hierarchical scale level is created and the mean percentage difference (*mPD*) between subsegment level (*L1*) and supersegment level (*L2*) is calculated as

$$mPD = \frac{|\mu B_{L1} - \mu B_{L2}|}{\mu B_{L2}} \tag{VI-9}$$

where μB is the mean brightness of the respective super- and subsegments. Similar to the simplified and data-driven version of the multiscale optimization approach proposed by Lu et al. (2011), we regard segments as "real" subsegment if their *mPD* exceeds the mean *mPD* of all subsegments by more than two standard deviations

$$real\ subsegment = \begin{cases} 1, & mPD > 2\sigma_{mPD} \\ 0, & \text{else.} \end{cases} \tag{VI-10}$$

Identified "real" subsegments (exemplified in Appendix A, Fig. VI-13d-I, rSS) are transferred to the supersegment level (Fig. VI-13d-II). Then, the similarity of transferred adjacent subsegment is evaluated in terms of their mean brightness values, and two adjacent segments are merged (Fig. VI-13d-III) if they fulfill the following condition

$$similarity_{segment1,segment2} = \begin{cases} 1, & |\mu B_{segment1} - \mu B_{segment2}| \leq \gamma \\ 0, & \text{else} \end{cases} \tag{VI-11}$$

with γ being a threshold. The procedure is repeated for the remaining scales above the plateau objective function, whereby the result of the previous cycle becomes the subsegment level in each step, and a number of segments are merged to create a supersegment level above, according to the respective scale factor (ibid.). The complete procedure is intended to provide an optimized segmentation that comprises distinctive and valid segments independent of their

particular scale. However, as the multilevel optimization procedure and the merging of non-valid segments alters image segments and alignment, we calculated the objective function again for the processed levels in order to identify the most sufficient final segmentation (Section 4.2., black arrow in Fig. VI-8c; exemplified in Appendix A, Fig. VI-13e, fS).

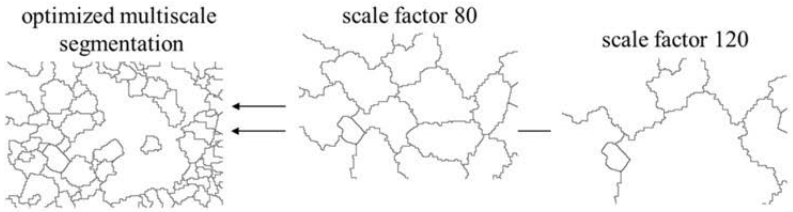
3.3. FEATURE CALCULATION FOR CHARACTERIZATION OF URBAN STRUCTURES

Features based on the RapidEye and TanDEM-X nDSM data were computed to allow for a comprehensive characterization of urban structures. Thereby, a perceptual coherence (Steiniger et al., 2008) of seismic vulnerability properties and physical appearance as well as composition of constituent elements of urban structures, as measured with remote sensing, is assumed.

We extracted statistical measures of 1st and 2nd order from both the multispectral RapidEye data and TanDEM-X nDSM (Fig. VI-4). Measures of central tendency and measures of spread of the different image bands are deployed. The same measures are computed for the nDSM to take into account height characteristics of discriminated areas. Additionally, rotation-invariant texture measures for the optical and nDSM data were computed based on the co-occurrence matrix (GLCM; Haralick et al., 1973). It could be shown that texture can provide valuable information when aiming at discrimination of urban structures (Herold et al., 2003), and generally allow to overcome a lack of spectral information (Pacifici et al., 2009). The last group of features consists of spatial metrics. By using the NDVI (Rouse et al., 1974), we quantify the share of vegetation per segment. Thereby, pixels which exceed a certain NDVI value (i.e., 0.3) are considered as vegetation. To specifically describe height characteristics of elevated objects that do not represent vegetation within a segment (i.e., buildings), we first identify pixels that exceed a certain nDSM value (i.e., 2.6m). Besides the share of elevated objects per segment, measures of central tendency and spread of elevated objects are computed.

FEATURES

computation of features for three segmentation levels



		RapidEye	TanDEM-X
Measures of 1st order	Measures of central tendency	Mean (ρR, ρG, ρB, ρNIR, ρBrightness, ρMax. Diff.) Median (ρR, ρG, ρB, ρNIR)	Mean (nDSM) Median (nDSM)
	Measures of spread	Standard deviation (ρR, ρG, ρB, ρNIR) Variance (ρR, ρG, ρB, ρNIR) Minimum (ρR, ρG, ρB, ρNIR) Maximum (ρR, ρG, ρB, ρNIR) Range (ρR, ρG, ρB, ρNIR) Interquartile range (ρR, ρG, ρB, ρNIR)	Standard deviation (nDSM) Variance (nDSM) Minimum (nDSM) Maximum (nDSM) Range (nDSM) Interquartile range (nDSM)
Measures of 2nd order	Texture based on GLCM	Homogeneity(ρBrightness) Contrast (ρBrightness) Dissimilarity (ρBrightness) Entropy (ρBrightness) Angular 2nd moment (ρBrightness) Mean (ρBrightness) Standard deviation (ρBrightness) Correlation (ρBrightness)	Homogeneity(nDSM) Contrast (nDSM) Dissimilarity (nDSM) Entropy (nDSM) Angular 2nd moment (nDSM) Mean (nDSM) Standard deviation (nDSM) Correlation (nDSM)
Spatial metrics		Share of vegetation	Share of elevated objects Mean height of elevated objects Standard deviation of height of elevated objects Variance of height of elevated objects Minimum height of elevated objects Maximum height of elevated objects Range of height of elevated objects

ρ = reflectance

Fig. VI-4. Features derived from remote sensing data and for different segmentation scales to characterize the urban morphology.

We calculate the aforementioned features not only for the optimized multiscale segmentation, which generally provides the spatial entities for subsequent procedures, but also for two segmentations with a larger scale factor (i.e., scale factor 80 and 120), which were not subject to the optimization procedure. This was done to ensure a sufficient consideration of spatial context information of urban structures. In fact, it is very challenging to fully exclude oversegmentation by means of an unsupervised segmentation procedure. When urban structures are subject to oversegmentation, a characteristic appearance with respect to

spectral-spatial composition may not be reflected. However, including information from coarser segmentation scales allows to model spatial context relations in feature space adequately. Generally, this idea is consistent with approaches that deploy supersegment information to enhance classification accuracy (e.g., Bruzzone & Carlin, 2006; Johnson & Xie, 2013).

3.4. FEATURE GROUPING AND SELECTION

Two different strategies were followed to bundle features into feature sets to be used in the experiments. First, a thematic grouping was carried out to quantify the usability of the different remote sensing data sets and evaluate the value of features from multiple segmentation levels. Beside the unreduced feature vector containing all computed features, we grouped all optical features, all height related features, all optical features from the optimized segmentation level, all height related features from the optimized segmentation level, and both all optical and height related features from the optimized segmentation level.

The second strategy to bundle features is motivated by the circumstance that feature vectors with a high dimensionality often exhibit redundancy, are highly correlated, and may suffer from the "Hughes phenomenon" (which states that for a limited amount of samples the predictive accuracy decreases as the dimensionality of the feature vector increases (Hughes, 1968)). Accordingly, we applied two machine learning-based feature selection algorithms on the data. The two approaches represent filter methods, which work independently with respect to the classifier, what is in contrast to the concept of wrapper methods. They were chosen since they can handle both regression problems and evaluate discrete valued variables. We used the Relief-F approach (Kononenko, 1994), because it enables to rank individual features, and deployed the Correlation-based Feature Selection (CFS) method (Hall, 1999), since it enables the scoring of the value of groups of features. The applied filter methods run supervised and aim to identify the best features for building robust regression and classification models.

The Relief-F approach ranks features according to the assumption that useful features should be able to differentiate between instances from different classes and have similar values for instances from the same class (Kononenko, 1994, Robnik-Šikonja and Kononenko, 2003). Therefore, an instance from the data is randomly sampled and k nearest neighbors from the same and opposite classes are located. Feature values of the k nearest neighbors are compared to the sampled instance and are used to up-date relevance scores for each feature. This procedure is repeated until a number of instances m was considered. Thereby, m is a free

parameter, but frequently set to the maximum number of labeled instances available in order to achieve a reliable approximation (Kononenko, 1994).

CFS deploys a best first search algorithm to identify a group of subsets, which are possibly suitable. The identified subsets are evaluated by means of an entropy based heuristic. The heuristic favors subsets with a high feature-class correlation and low feature-feature inter-correlation. Relief-F and CFS are both multivariate procedures, which evaluate features in dependence of other features in the data set.

3.5. SUPPORT VECTOR MACHINES

Support Vector Machines emanate from the field of machine leaning and represent a family of non-parametric approaches for supervised classification and regression (Vapnik, 1995). They are based on the structural risk minimization principle, which suggest a tradeoff between the accuracy of an approximation and the complexity of the affiliated approximation function. SVMs determine a suitable set of parameters to fit a decision surface, the so-called hyperplane, between different classes of labeled samples. To deal with nonlinearly problems, the labeled samples are mapped through a nonlinear transformation $\phi(\cdot)$ from the input space $\mathcal{X}$ into a space of higher dimensionality $\mathcal{H}$. In that space, the optimal separating hyperplane maximizes the margin between the patterns of the different classes and the hyperplane. The maximized margin can be described by two additional, marginal hyperplanes that border the samples closest to the separating surface, the so-called support vectors (Burges, 1998; Leinenkugel et al., 2011). Only those samples are needed to define the model, what allows to build robust models with a high generalization capability based on a comparatively small number of labeled training samples. A linear separation in $\mathcal{H}$ corresponds to a nonlinear separation in the original input space $\mathcal{X}$ (Volpi et al., 2013). This general principle of SVMs allows to learn nonlinear relations and is visualized in Fig. VI-5a-d.

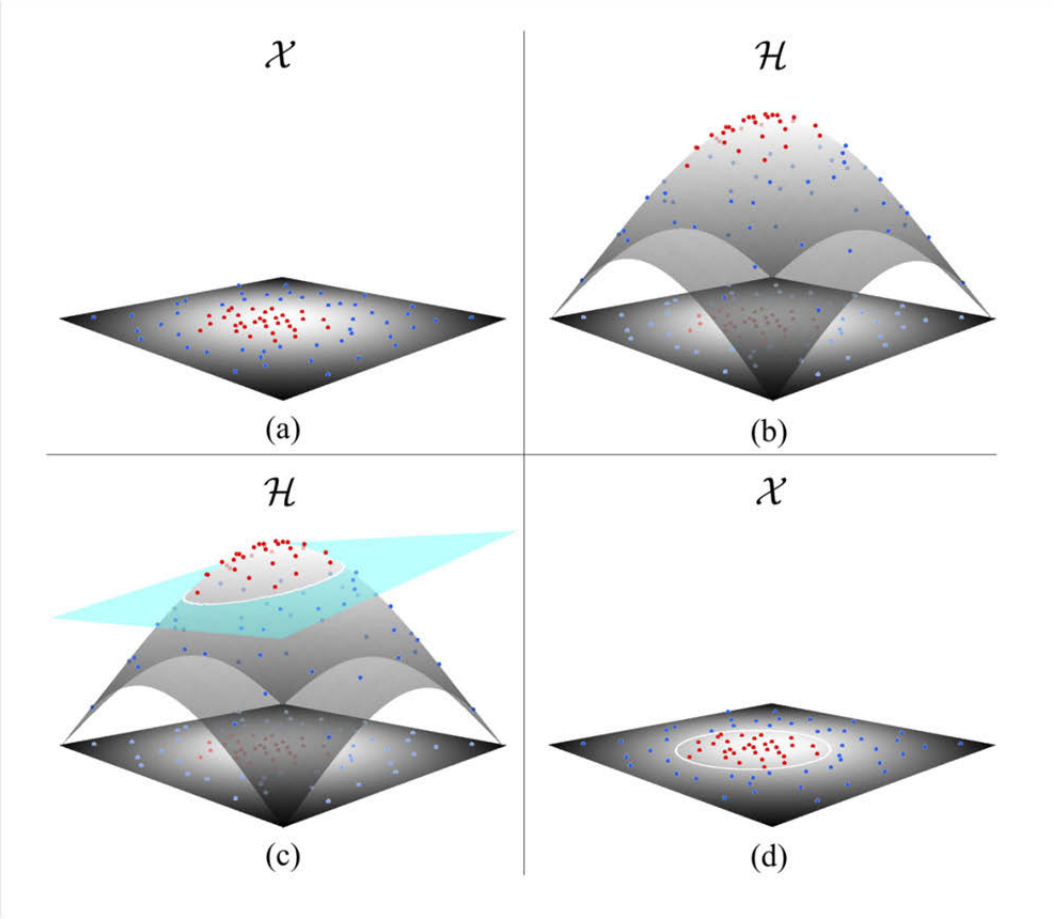


Fig. VI-5. Idealized procedure for generation of a nonlinear decision function by SVM; (a) Dataset with two classes (red and blue dots) that are non-linearly separable in $\mathcal{X}$ are mapped through a nonlinear transformation $\phi(\cdot)$ into a space of higher dimensionality $\mathcal{H}$ (b); a linear separation becomes possible in that space and a hyperplane (cyan) with maximum margin is fitted (c), what corresponds to a nonlinear decision function in $\mathcal{X}$ (d).

Different formulations of SVM allow to render a prediction problem in three ways: (i) if two or more classes are to be discriminated based on labeled samples of all classes present in the data under investigation, the framework of *C*-SVMs can be utilized; (ii) if solely labeled samples are available for the class(es) of interest and not for all classes present in the data, one-class *v*-SVM can be deployed; (iii) if the statistical level of measurement is higher and corresponds to an interval or ratio scale, a Support Vector Regression approach allows to estimate a function from the training samples. We depict these three formulations with affiliated minimization objective and decision function in the subsequent three subsections to clarify methodological divergences, which are relevant for differing real-world situations. Those may evolve when aiming to assess urban structures according to different engineering-related methods and a varying amount of prior knowledge available. For a more detailed theoretical and application-oriented background of SVM, the reader can refer to Vapnik

(1995, 1998), Burges (1998), Chang and Lin (2001), and Schölkopf and Smola (2002, 2004). In the context of remote sensing, e.g., Melgani and Bruzzone (2004), Camps-Valls and Bruzzone (2005, 2009), Mountrakis et al. (2011), or Salcedo-Sanz et al. (2014) provide comprehensive literature.

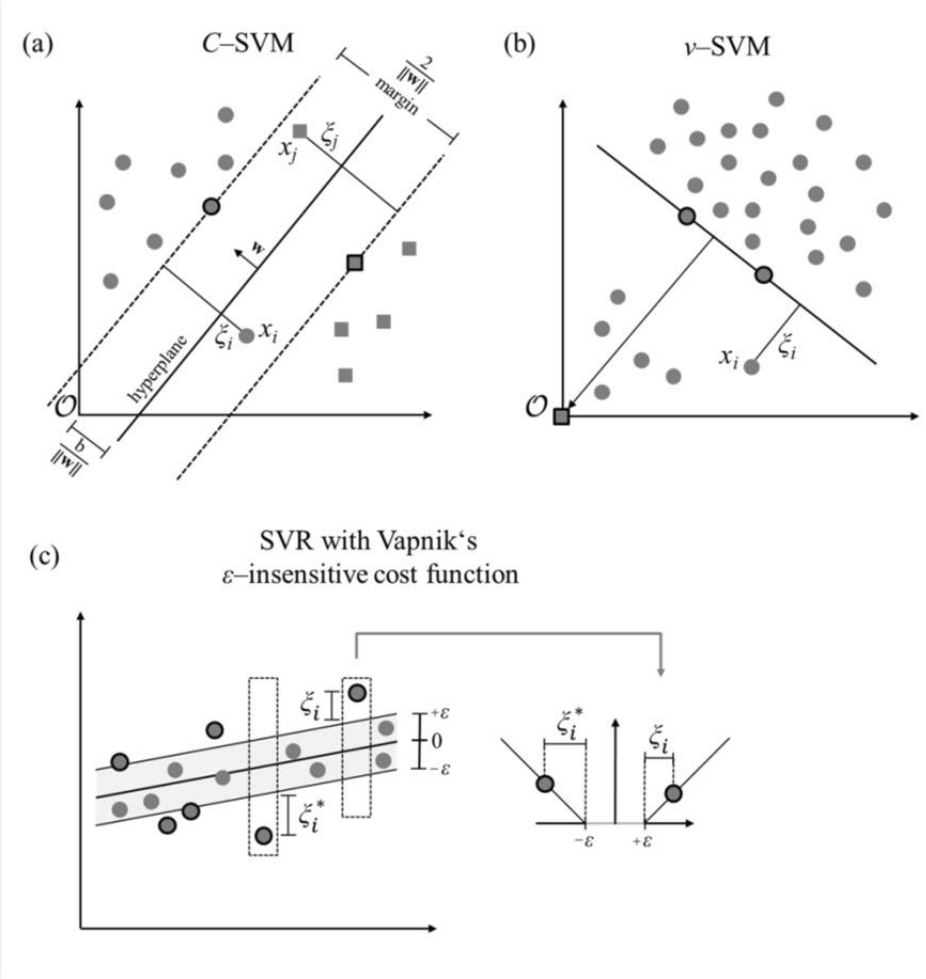


Fig. VI-6. Different formulations of SVM; (a) *C*-SVM provide a linear decision function with relaxed separation constraints, which allow for labeled samples lying on the incorrect side of the hyperplane (modified adaption from Melgani and Bruzzone, 2004); (b) *ν*-OC-SVM treat the origin of the feature space ($\mathcal{O}$) as the only available sample of the non-target class first, and fit a hyperplane with maximum margin from the origin (modified adaption from Muñoz-Marí et al., 2010); (c) in SVR all samples outside a fixed tube with size ε (i.e., support vectors) are penalized by applying a cost function. Here, Vapnik's ε-insensitive cost function is deployed, which accounts for a linear penalization (modified adaption from Schölkopf and Smola, 2002, and Salcedo-Sanz et al., 2014).

3.5.1. MULTICLASS CLASSIFICATION WITH *C*-SVM

The *C*-SVM method was introduced by Cortes and Vapnik (1995) to cope with class overlap or the existence of noise in the training data in multiclass problems. This technique represents a modification of the maximum margin approach using relaxed separation constraints that allow for the possibility of instances on the incorrect side of the respective margin boundary.

Consider a data set with labeled instances $\{\mathbf{x}_i, y_i\}_{i=1}^n$, with $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \{-1, +1\}$. The data is mapped through a nonlinear transformation $\phi(\cdot)$ to a space with a higher dimension. An appropriate determination of $\phi(\cdot)$ ensures that the transformed samples are more likely to be linearly separable in the higher dimensional space (Camps-Valls and Bruzzone, 2005). Then, the minimization objective of *C*-SVM is formulated as

$$\min_{\mathbf{w},\xi_i,b} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{n} \xi_i \right\} \tag{VI-12}$$

subject to

$$y_i(\langle \phi(\mathbf{x}_i), \mathbf{w} \rangle + b) \geq 1 - \xi_i \qquad \forall i = 1, \ldots, n \tag{VI-13}$$

$$\xi_i \geq 0 \qquad \forall i = 1, \ldots, n \tag{VI-14}$$

where **w** represents the normal perpendicular to the optimal separating hyperplane and *b* is the nearest distance to the origin ($\mathcal{O}$) of the coordinate system. These parameters constitute a linear classifier in $\mathcal{H}$, which separates the labeled samples of different classes with maximum margin. To enhance generalization capabilities and reduce over-fitting, positive slack variables ξ_i are introduced to account for labeled samples lying on the incorrect side of the respective margin boundary (Fig. 6a). The constant *C* determines the trade-off between maximizing the margin and the number of incorrectly classified samples (training errors). An optimal parameterization of *C* can be determined empirically. It can be noted that equation (12) is constituted by two distinctive terms that are clearly interpretable: the objective is to minimize simultaneously both the norm of the model weights, $\|\mathbf{w}\|^2$, which is equivalent to the maximization of the margin, and the committed errors $\sum_{i=1}^n \xi_i$ (Salcedo-Sanz et al., 2014).

The minimization objective of equation (12) is reformulated from its primal form to its dual form by introducing Lagrange multipliers, so that it can be solved with quadratic programming techniques efficiently (Schölkopf and Smola, 2002). Finally, a decision function is given that allows to assign a class label to an instance of unknown class membership $\mathbf{x}_*$

$$f(\mathbf{x}_*) = \operatorname{sgn}\left(\sum_{i=1}^{n} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_*) + b \right) \tag{VI-15}$$

with α_i being the Lagrange multipliers and *K* being a kernel function. The Lagrange multipliers are determined by optimization and feature nonzero values for instances lying on the margin – the support vectors (Cortes and Vapnik, 1995; Camps-Valls and Bruzzone, 2005; 2009). The kernel function *K* is expressed as the dot product of mapped instances $K(\mathbf{x}_i, \mathbf{x}_j) =$

$\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$. Hence, the outcome of the decision function (15) only relies on the dot product of the vectors in the input space $\mathcal{X}$. This allows to avoid an explicit projection to a space of higher dimensionality, since this formulation yields the same solution (also known as the kernel trick). This property of the SVM algorithm enables a very efficient computation of decision functions for data with a very high dimensionality. A number of different kernels with varying characteristics exist. However, in environmental applications it is common to use Gaussian RBF kernels, that take the form $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$, due to their interpretability (they express local similarity) in accordance with favorable performance properties (Volpi et al., 2013). Hence, we rely on this kernel for all support vector methods in the experiments.

3.5.2. ONE-CLASS CLASSIFICATION WITH ν-SVM

One-class *ν*-SVM were introduced by Schölkopf et al. (1999) as support vector method for novelty detection. They can be deployed in situations where the objective is to identify only one or few classes of interest from all classes present in the data - simultaneously having only labeled samples of the classes of interest available.

Generally, the strategy of *ν*-OC-SVM is to capture the support region (i.e., where the density is large) without the need of prior assumptions about the distribution of the data. Therefore, the target class is described by a function that maps the majority of instances to a region where the function is nonzero. To achieve this, the origin of the feature space is first treated as the only available member of the non-target class (i.e., as an outlier). Then, a hyperplane with maximum margin separation from the origin is identified (Fig. 6b). To separate the data from the origin, the minimization objective is as follows

$$\min_{\mathbf{w},\xi_i,p} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu l}\sum_i \xi_i - p \right\} \qquad \text{(VI-16)}$$

subject to

$$\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle \geq p - \xi_i \qquad \forall i = 1, \ldots, l \qquad \text{(VI-17)}$$

$$\xi_i \geq 0 \qquad \forall i = 1, \ldots, l \qquad \text{(VI-18)}$$

where $\mathbf{w}$ represents a vector, which is perpendicular to the hyperplane, and p is the distance to the origin. Parameter $\nu \in (0,1]$ controls the tradeoff between an upper bound of fraction of margin errors (as can be seen from (16), outliers in the training data are handled by slack variables analogous to the *C*-SVM framework) and a lower bound of fraction of support

vectors (i.e., model complexity) (Schölkopf et al., 2001, Muñoz-Marí et al., 2010). Again, by utilizing Lagrange multipliers and a kernel function, the final decisions function to assign a class label to an instance of unknown class membership $\mathbf{x}_*$ is obtained by

$$f(\mathbf{x}_*) = \operatorname{sgn}\left(\sum_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_*) - p\right). \tag{VI-19}$$

In practice, *v*-OC-SVM can be used in problems when more than a single specific class of interest is considered at a time by employing an ensemble of one-class classifiers. Thereby, each classifier is trained on a specific class of interest. When manifold classifiers assign a class label to an instance, a heuristic can be applied to the outputs (e.g., a winner-take-all rule based on prior or posterior probabilities) to determine the final class membership (Marconcini et al., 2014).

3.5.3. FUNCTION ESTIMATION WITH SVR

SVR allows approximating a function from training data, when the statistical level of measurement of the target variable corresponds to an interval or ratio scale (i.e., $y \in \mathbb{R}$). In accordance to the presented SVM framework, SVR defines a linear model over samples that are mapped to a higher dimensional space via a nonlinear function. A common SVR formulation deploys Vapnik's ε-insensitive cost function, in which errors up to ε are not penalized, and further errors are subject to a linear penalization (Fig. VI-6c). Thereby, SVR determines weights $\mathbf{w}$ by minimizing a regularized functional

$$\min_{\mathbf{w},\xi_i,\xi_i^*,b}\left\{\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i(\xi_i + \xi_i^*)\right\} \tag{VI-20}$$

subject to

$$y_i - (\langle\phi(\mathbf{x}_i), \mathbf{w}\rangle + b) \leq \varepsilon + \xi_i \qquad \forall i = 1, \ldots, n \tag{VI-21}$$

$$(\langle\phi(\mathbf{x}_i), \mathbf{w}\rangle + b) - y_i \leq \varepsilon + \xi_i^* \qquad \forall i = 1, \ldots, n \tag{VI-22}$$

$$\xi_i, \xi_i^* \geq 0 \qquad \forall i = 1, \ldots, n \tag{VI23}$$

where ξ_i and ξ_i^* are positive slack variables, which quantify the distances of the labeled training samples that lie outside of the ε-insensitive tube to the boarder of the tube (Fig. VI-6c). The regularization parameter C determines the tradeoff between the flatness of the function and the tolerance to observed errors. Similar to the previous formulations, the minimization objective can be solved by introducing linear restrictions (21)-(23) into (20) using Lagrange multipliers α_i, calculate the Karush-Kuhn-Tucker conditions, and solve the

dual problem with quadratic programming techniques (Drucker et al., 1997; Schölkopf and Smola, 2002; Verrelst et al., 2012; Salcedo-Sanz et al., 2014). Subsequently, the final estimation function is given by

$$f(\mathbf{x}_*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_*) + b. \quad \text{(VI-24)}$$

4. DESCRIPTION OF EXPERIMENTS AND EXPERIMENTAL SETUP

4.1. DESCRIPTION OF EXPERIMENTS

Experiments are designed with three different formulations of the supervised prediction problem. Generally, suggested procedures can be considered relevant when prior information about seismic vulnerability is available for parts of an area of interest but not for an area of interest in a spatially continuous and complete way. Here, we render the prediction problem in three ways to address operational requirements that may evolve in many real-life cases.

(i) For the district Zeytinburnu we first aggregate the damage grade, as determined with the capacity spectrum method, from building level to structure level. Hence, the aim is to estimate the damage grade for spatial entities, where no damage grade information is available based on SVR approach (Fig. VI-7a).

(ii) Then, we use the ground-based GPS-photos to identify urban structures that are either determined by large industrial/commercial buildings or tall detached residential buildings. Generally, the predominant building type can be seen as proxy for the vulnerability of an area (Wieland et al., 2012). In this study, those two structure types revealed distinctive properties with respect to the capacity spectrum assessment. Thereby, urban structures determined by large industrial/commercial buildings can be considered as highly vulnerable, whereas urban structures determined by tall detached residential buildings can be considered as slightly vulnerable. In the second experiment, we aim to identify those two urban structures within the settlement area of Istanbul, disregarding all other potentially present classes in the area. Thereby, only labeled samples are available for the classes of interest and not for all classes present in the data (Fig. VI-7b). This is a very challenging but realistic task. The availability of information comprising all urban structures present in the area of interest is infrequently satisfied in real-life situations. We deploy an ensemble of *v*-OC-SVM to address this problem.

(iii) However, in experiment three, we investigate feasibility of procedures when also labeled samples of other classes are available (Fig. VI-7c). This renders the classification problem fully supervised and we approach it with *C*-SVM.

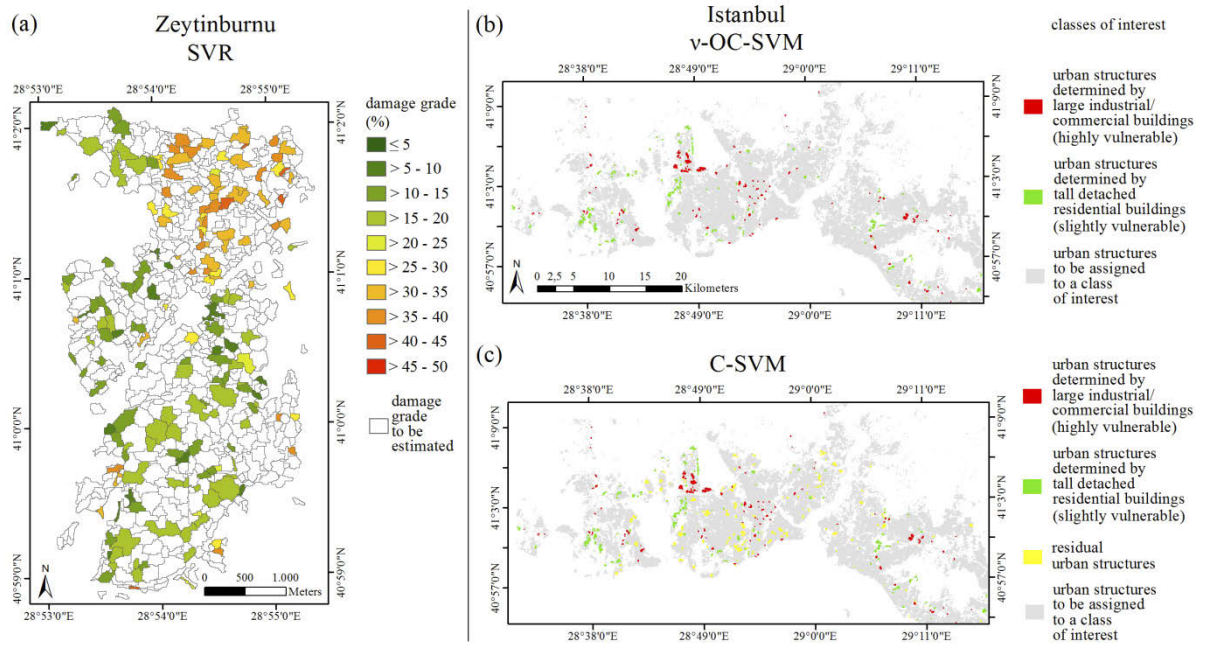


Fig. VI-7. Concept and data for experiments; (a) damage grade for Zeytinburnu which is to be estimated with SVR; (b) available labeled samples of classes of interest deployed for learning ensembles of ν-OC-SVM; (c) setting for *C*-SVM: additional labeled samples for classes of residual urban structures are available what renders the classification problem fully supervised.

4.2. EXPERIMENTAL SETUP

Regarding free parameters of multiresolution segmentation, we put more emphasize on shape heterogeneity (Shape: 0.7) rather than on grey-value heterogeneity (0.5) as suggested in section 3.2. First, an initial scale h_s was determined. Therefore, the ratio of valid and non-valid segments, as evaluated with (2), was computed. Fig. VI-8a shows the corresponding function. We chose scale 24, since it features a small share of non-valid segments and the decline in shares of non-valid segments with respect to neighboring scales is comparatively explicit. Hence, the objective function is calculated for the interval [24, 150], and the outcome is plotted in Fig. VI-8b. It can be seen that optimal segmentation scales above the plateau objective function ($F(u)$ = 1.432) are located between 24 and 38. Hence, those scales were subject to the optimization procedure described in section 3.2.2 (γ in similarity constraint (11) was set to 5). Outcomes are reassessed with the objective function in order to determine the final segmentation (black arrow in Fig. VI-8c).

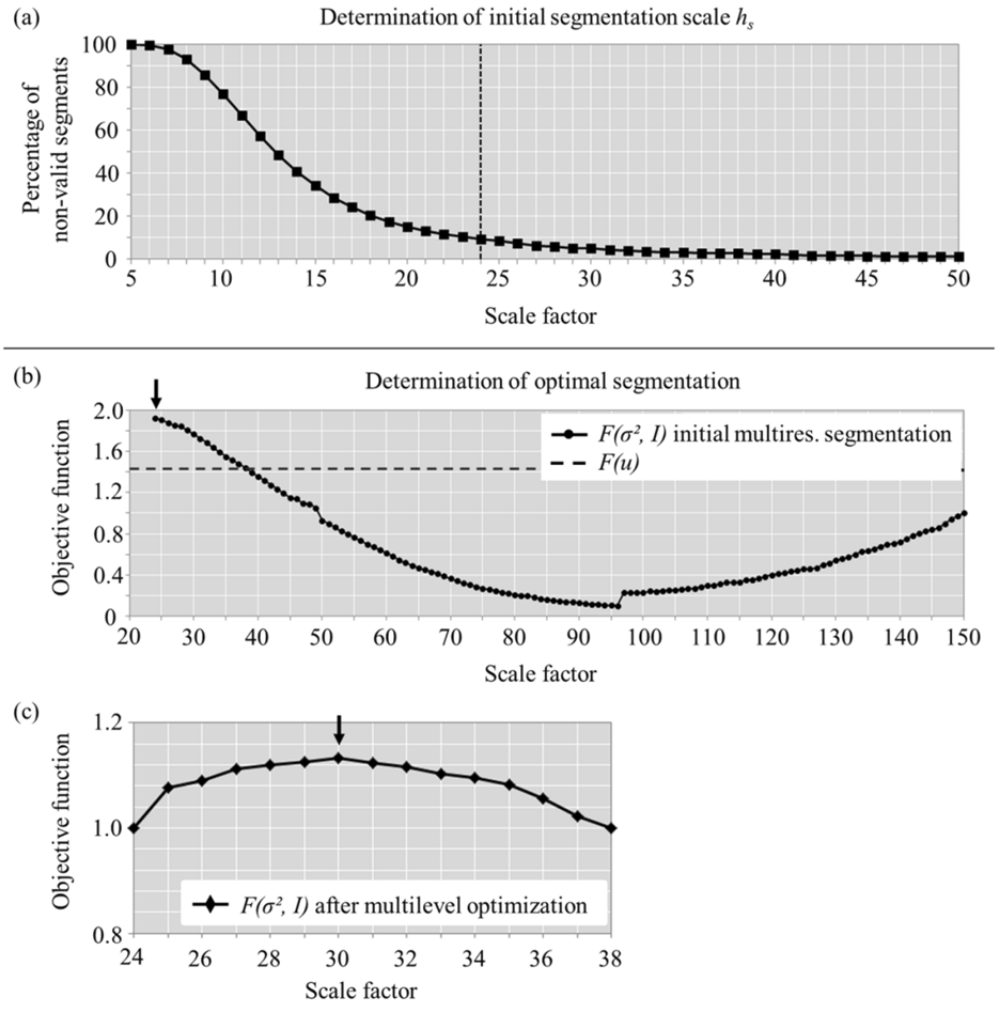


Fig. VI-8. Outcomes of multilevel image segmentation optimization procedure.

Regarding the feature selection algorithms, we tested varying numbers for neighbors to be considered regarding the Relief-F approach. Finally, *k* was consistently set to 10, since it was found that results are hardly sensitive for this parameter in this study. Based on Relief-F, we compiled four feature sets, containing the 10, 20, and 50 best ranked features, and a set with all features that have a positive degree of relevance (w > 0). For the CFS approach we deployed a stopping criterion of 5 consecutive fully expanded, non-improving subsets. The subset with the highest merit according to the evaluation heuristic during the search was selected. Together with the thematically grouped features (sets include features from different remote sensing data sources and segmentation scales) computed features were subdivided in eleven different feature sets.

As mentioned, we use RBF kernels for all support vector methods (support vector methods in this paper were carried out with the LibSVM package by Chang and Lin (2001)). Features were normalized in feature space, yielding normalized kernel functions (Graf et al., 2003).

Regarding the SVR approach, the number of samples has been varied to test sensitivity with respect to accuracy. We learn models with randomly drawn 25%, 50%, 75%, and 100% of the labeled samples available and estimate generalization capabilities based on 5-fold cross-validation. To avoid skewed results, five different training sets sharing the same number of labeled instances were created. It was made sure that samples contained in one set are also contained in the affiliated set with a larger number of samples (i.e., the samples randomly drawn in one of the five sets with 25% of available labeled samples are also contained in the corresponding set with 50% of available labeled samples and so on), to avoid a biased quantification of the effect of training set size on prediction accuracy. Generalization capability is evaluated in terms of mean absolute percentage error (MAPE) and reported as average of three independent trials (some Pearson's correlation coefficients R are also reported in the results section to account for the models' goodness of fit). Model parameters were optimized with respect to MAPE in the ranges $\sigma = \{10^{-1}, \ldots, 10\}$, $C = \{1, \ldots, 100\}$, and $\varepsilon = \{10^{-6}, 10^{-3}\}$.

Application of the ν-OC-SVM approach with a RBF kernel requires adjusting the parameter ν, and the kernel-width parameter γ. Generally, it is difficult to tune free parameters if only target labeled samples are available in the training data. In such situations solely the true positive rate (sensivity) can be calculated, whereas the error counterpart (specificity) cannot (i.e., a suitable model cannot be distinguished from a fatally underfitted model since also the latter would yield a high accuracy). To overcome this limitation, the free parameter selection was determined by evaluating $\arg\max_{\theta} \left\{\frac{\text{OA}[\%]}{\#\text{SV}}\right\}$, where θ is the set of free parameters (i.e., ν and γ), OA is the overall accuracy and $\#SV$ the number of support vectors. This heuristic enforces high OA while simultaneously limiting model complexity keeping a low number of SV (Muñoz-Marí et al., 2010). For both ν and γ we performed a grid search varying ν in the range $\{0.01, \ldots, 0.1\}$ in 0.01 steps and γ in the range $\{10^{-2}, \ldots, 10^{1}\}$ in power of $\sqrt{10}$ steps, respectively. OA for each model was estimated by a 5-fold cross-validation strategy. The parameter combination yielding the highest value of the evaluation heuristic was chosen. In case a segment is voted as belonging to the class of interest by multiple ν-OC-SVM models, reasonably, we assign the segment to the residual class since results can be considered non-distinctive, and we generally observe considerable errors of commission in preliminary model runs. In the experiments, we used 50% of available samples for training

the models and 50% for validation. Results are reported based on five independent trials. Evaluation is based on estimated κ statistic (which allows considering both omission and commission errors and is thus not biased by class distribution; Foody, 2004).

For the *C*-SVM, we adopt a one-against-one scheme (Hsu and Lin, 2002), since we deal with more than two classes to be discriminated. Learning the most appropriate *C*-SVM in conjunction with a RBF kernel requires the definition of the cost-parameter *C* and the kernel-width parameter γ. Tuning of *C* and γ was addressed by a grid search strategy based on 5-fold cross-validation. Generalization accuracy is evaluated in terms of estimated κ on the average of three independent trials. In conformity with the recommendations of Hsu et al. (2010), a grid-search with values of $C = \{2^{-4}, 2^{-3}, \ldots, 2^{12}\}$ and $\gamma = \{2^{-5}, 2^{-4}, \ldots, 2^{3}\}$ was performed. Analogous to the SVR approach, we reveal generalization capabilities as a function of feature sets with different shares of available labeled samples.

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1. ESTIMATION OF DAMAGE GRADES FOR THE DISTRICT ZEYTINBURNU WITH SVR

For Zeytinburnu, mean absolute percentage errors with respect to estimated damage grades are reported in Fig. VI-9a-c as functions of training set size. The different figures and functions correspond to MAPE obtained with different feature sets. Generally, it can be noted that graphs reveal distinctive differences with respect to accuracy. When using solely features calculated from the elevation data (i.e., nDSM) MAPE are considerably higher than corresponding MAPE obtained with features computed from the optical data (Fig. VI-9a). In particular, MAPE decrease from 20.76% (± 0.94; R = 0.163), achieved with elevation features, to 16.23% (± 0.36; R = 0.544), achieved with the optical features, when all samples were deployed within the cross validation procedure. Analogously, when examining accuracies obtained with features computed not only from the optimized segmentation level but also from supersegments (Fig. VI-9b), we also observe a considerable decrease in MAPE from 14.54% (± 0.41; R = 0.596) to 11.64% (± 0.30; R = 0.753). Besides, an explicit gain of accuracy can be observed when incorporating supersegment information compared to a single segmentation (e.g., decrease of MAPE with both optical and elevation features from 16.34% to 11.85%; simultaneously, goodness of model fit raises from R = 0.521 to R = 0.741). Fig. VI-9a and b also reveal that a joint use of optical and elevation information does not increase accuracy remarkably compared to accuracies obtained with optical features alone (difference in MAPE is 0.36% and 0.30%, respectively). Lastly, Fig. VI-9c shows that the application of

filter methods allow to identify beneficial feature sets for the estimation of damage grades. Thereby, lowest MAPE of all feature sets (i.e., 10.74%; $R = 0.778$) could be obtained with the 50 highest ranked features from the Relief-F algorithm.

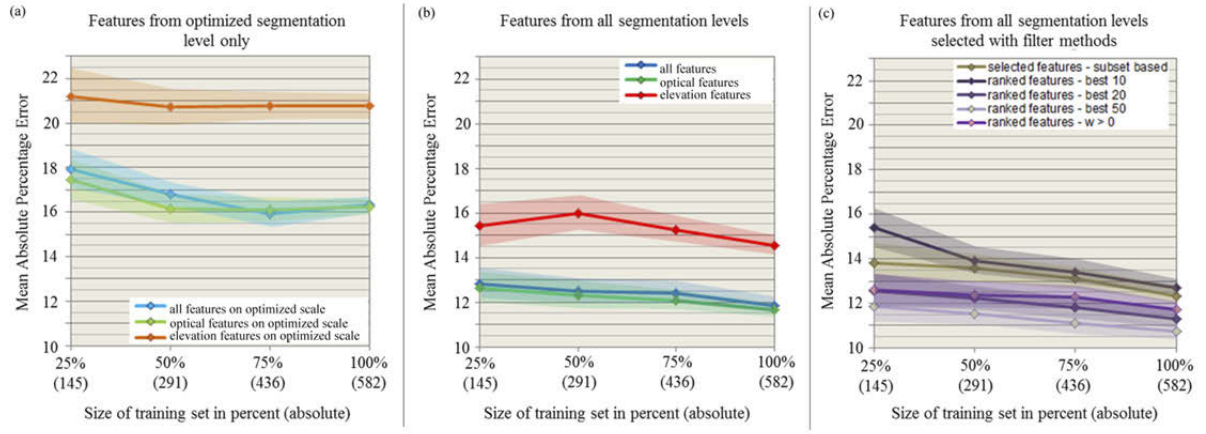


Fig. VI-9. Mean absolute percentage errors (reported as mean and standard deviation of different model runs) as a function of training set size; (a) MAPE obtained with features from the optimized segmentation level only; (b) MAPE obtained with features from all three segmentation levels; (c) MAPE obtained with features selected by filter methods.

These numerical results suggest that features from optical data allow estimating damage grades with viable accuracies, whereas elevation features, as derived from the TanDEM-X data, do not allow estimating damage grades in a viable way. Here, also a joint exploitation of both data sets does not feature a clear improvement in accuracy. In contrast, the idea to exploit supersegment information proved to enhance accuracy of estimates considerably. This is in line with previous experimental analyses that confirm an increase of prediction accuracy, when modeling spatial context information in feature space with supersegment information (e.g., Bruzzone & Carlin, 2006; Johnson & Xie, 2013). However, this observation simultaneously suggests that homogeneous urban structures could not consistently be discriminated properly based on the optimized segmentation, and that oversegmentation has occurred. However, in contrast to undersegmentation, oversegmentation leaves the possibility to gain accurate mapping results after classification (i.e., a possibly oversegmented urban structure, which may be represented by two segments instead of one, can be mapped properly if both segments are assigned to the same class). In this sense, the incorporation of supersegment information allowed to gain viable accuracies in the experiments. In addition, deployed feature selection algorithms proved useful to alleviate problems associated with high-dimensional feature vectors in conjunction with a comparatively small number of labeled

samples, since best accuracies were achieved with subsets of the features. Thereby, best models are characterized by MAPE less than 11% and fairly strong goodness of fit ($R > 0.75$).

As a further means, an application of a model to estimate damage grades for Zeytinburnu is shown in Fig. VI-10a. It was learned with 50% of available labeled samples and 50% were held out for validation. A MAPE of 13.00% and R of 0.754 was obtained based on 50 features ranked highest with Relief-F. Generally, a good spatial agreement of reference and model estimates can be observed. Areas characterized by a high vulnerability and high damage grades in the northern part of Zeytinburnu feature also highest model estimates. Instead, a vast majority of the central and southern parts correspond to moderate or low damage grades in both maps. Simultaneously, as can be seen from the scatter plot, the model overestimates damage grades for instances with low damage grades, and underestimates high damage grades.

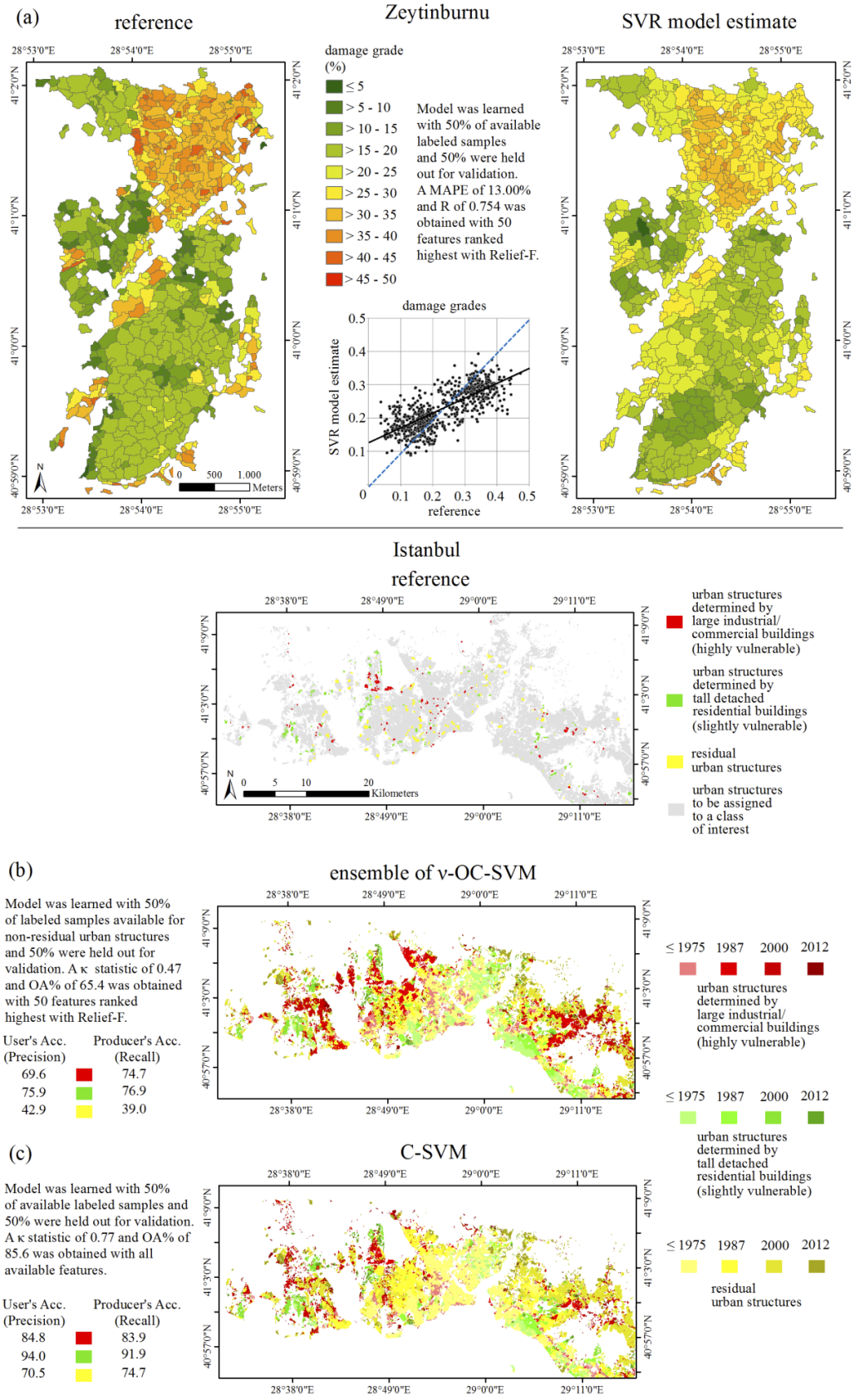


Fig. VI-10. Application of learned models; (a) estimated damage grade with SVR for Zeytinburnu; (b) classified urban structures of Istanbul according to an ensemble of *ν*-OC-SVM; (c) classified urban structures of Istanbul according to a fully supervised *C*-SVM

5.2. ASSIGNMENT TO CLASSES OF INTEREST WITH AN ENSEMBLE OF ν-OC-SVM

We aimed to identify two distinctive urban structures (i.e., urban structures determined by large industrial/commercial buildings that can be considered as highly vulnerable, and urban structures determined by tall detached residential buildings that can be considered as slightly vulnerable) within the settlement area of Istanbul, disregarding all other potentially present classes in the area. Only labeled samples were available for the classes of interest in this experiment and not for all classes present in the data, what rendered the prediction problem extremely challenging.

Prediction accuracies, as evaluated with estimated κ statistics for different feature sets and model runs, are shown in Fig. VI-11. Analogous to results of SVR, we observe considerable differences in accuracy with respect to deployed feature sets. It can be noticed that κ statistics achieved with features from the optimized segmentation level indicate agreements just slightly better than chance (i.e., κ statistics vary from –0.03 to 0.24). Again, elevation features used alone perform worst. Interestingly, when incorporating knowledge from supersegments, a consistent increase of inter-rater agreement can be observed for all three feature groups. Corresponding κ statistics show a mean increase in terms of absolute values by 0.09, 0.05, and 0.12, respectively. Thereby, the unreduced feature vector allows to exhibit fair agreements with κ statistics between 0.23 and 0.34. Nevertheless, applied strategies to reduce dimensionality of the feature vector proved useful since features sets generated with filter methods generally provide highest κ statistics. Best models feature a moderate agreement, with κ statistics larger than 0.45.

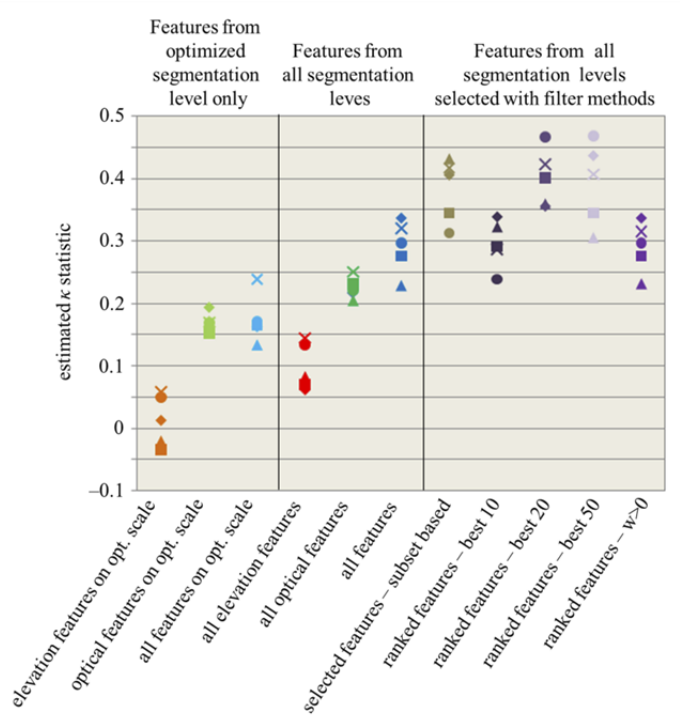


Fig. VI-11. κ statistic obtained with an ensemble of ν-OC-SVM for different feature sets. Samples were split in a stratified manner and models were learned with 50% of available samples and 50% were used for validation; five different model runs with a differing composition of samples were carried out (marked with different symbols, i.e., x, ○, Δ, ◊, □).

Interestingly, in accordance with the findings regarding the SVR approach, also here we observe that elevation features used alone yield least favorable results. However, in contrast to SVR, a joint exploitation with optical features yields a substantial improvement in accuracy when simultaneously considering supersegment information (mean κ statistics increase from 0.17 to 0.29). Results are comparable to those obtained with models learned from filter-based features sets. In this sense, feature selection algorithms proved useful again, since best results were obtained with 20 and 50 features ranked highest with Relief-F. However, κ statistics cannot exhibit a moderate agreement (i.e., best models feature κ statistics between 0.4 and 0.47) what questions viability in real-life situations. Although benchmark accuracies that need to be met in real life situations are largely absent within this rather emerging application context, results correspond to the challenging nature of the prediction problem.

An ensemble of models was applied to assign urban structures of Istanbul to the classes of interest (Fig. VI-10b). It was learned with 50% of labeled samples available for non-residual urban structures and 50% were held out for validation. A κ statistic of 0.47 and OA% of 65.4 was obtained based on 50 features ranked highest with Relief-F. From the affiliated user's and producer's accuracies it can be seen that urban structures determined by tall detached

residential buildings could be identified most accurately, closely followed by urban structures determined by large industrial/commercial buildings. The residual class shows lowest accuracies. In particular, large parts of the settlement area were assigned to either non-residual class. The validity of this vast assignment can be questioned, in particular in conjunction with results from the subsequent experiment (discussion follows at the end of the subsequent subsection). In turn, this suggests a large error of commission, evolving from broad, non-sufficient decision functions of the models describing the distributions of the two classes of interest.

5.3. ASSIGNMENT TO CLASSES OF INTEREST WITH *C*-SVM

To complement results obtained with an ensemble of *v*-OC-SVM, we also rendered the prediction problem fully supervised and approached it with *C*-SVM. Results are reported in terms of estimated κ statistics in Fig VI-12a-c. As for the two previous experiments, also here, we observe considerable differences with respect to accuracy in dependence of deployed feature sets. When using solely elevation features computed from the optimized segmentation level, κ statistics feature a fair to moderate agreement (i.e., mean κ statistics vary between 0.42 ± 0.09 and 0.46 ± 0.05). In contrast, optical features allow to achieve substantial κ statistics larger than 0.6 most of the time (Fig. VI-12a). Again, usability of supersegment information is expressed in a further increase of κ statistics, which definitely met a substantial agreement (i.e., $\kappa > 0.6$) when using optical features or jointly exploit optical and elevation features (Fig. VI-12b). Thereby, a joint use of optical and elevation features yields an increase of κ between 0.022 and 0.044 compared to the use of optical features alone. In this sense, best models show excellent κ statistics larger 0.8. Regarding feature sets composed with filter methods, we observe that only a sufficient number of features ensures viable κ statistics with a substantial or excellent agreement. Thereby, features that show a positive degree of relevance as evaluated with Relief-F ($w > 0$; $n = 152$) consistently performed best (Fig. VI-12c).

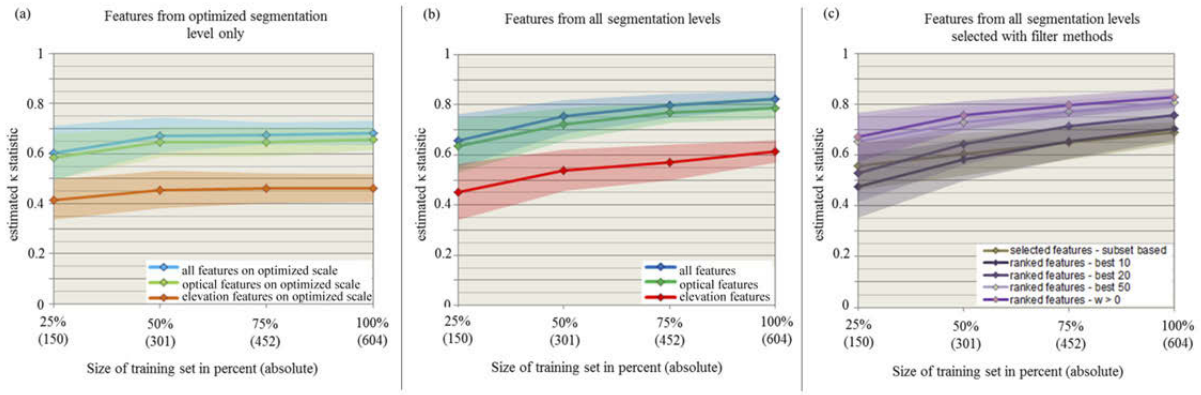


Fig. VI-12. κ statistics (reported as mean and standard deviation of different model runs) as a function of training set size; (a) κ obtained with features from the optimized segmentation level only; (b) κ obtained with features from all three segmentation levels; (c) κ obtained with features selected by filter methods.

Results obtained with *C*-SVM show distinctive increase of accuracies when compared to results obtained with ensembles of *v*-OC-SVM. Ensembles of *v*-OC-SVM could not exceed moderate agreements, with κ statistics slightly above 0.45. Instead, *C*-SVM allowed to obtain κ statistics showing substantial and even excellent agreement ($\kappa > 0.6$ up to $\kappa > 0.8$). Generally, this is reasonable and meets expectations, since *C*-SVM can rely on more prior knowledge. In fact, it is much more challenging to estimate the support of multiple multi- or high-dimensional distributions, than to discriminate known target classes from each other. Nevertheless, the use of labeled samples from classes others than the classes of interest proved very useful to achieve accuracies that may be needed in real-life cases. Hence, results confirm viability of proposed procedures to support an assessment of urban structures with respect to their seismic vulnerability.

The application of a *C*-SVM model to assign urban structures of Istanbul to the classes of interest is shown in Fig. VI-10c. It was learned with 50% of available labeled samples and 50% were held out for validation. A κ statistic of 0.77 and OA% of 85.6 was obtained with all available features. User's and producer's accuracies reveal that urban structures determined by tall detached residential buildings could be identified very accurately (accuracies > 90.0), and also urban structures determined by large industrial/commercial buildings feature very good agreements (accuracies ~ 84.0). A moderate decrease in accuracy can be observed for the class of residual urban structures (user's acc. of 70.5 and producer's acc. of 74.7). However, in distinction to the results obtained with an ensemble of *v*-OC-SVM, the majority of the settlement area of Istanbul is assigned to residual urban structures and urban structures of particular interest emerge now in a spatially very explicit way. As just one example, consider urban structures determined by large industrial and commercial buildings in the

southern part of the Asian side. Those clearly emerge in a linear manner - in reality - along an arterial road.

6. SUMMARY, CONCLUSIONS, AND FUTURE PERSPECTIVES

Global urbanization processes and increasing spatial concentration of exposed elements induce seismic risk at a uniquely high level. To mitigate affiliated perils determines detailed knowledge about building inventories and their behavior under a certain level of ground shaking. In this manuscript, we suggested procedures which learn prior seismic vulnerability information to identify seismic vulnerability of urban structures based on remote sensing. In contrast to numerous studies that operate on building level, we therefore addressed a coarser level of the urban morphology. This allowed relying on remote sensing data with a lower spatial resolution but larger spatial coverage. In this sense, we exploited data from the RapidEye constellation, TanDEM-X mission, and Landsat archive. A sequential procedure comprising the (i) delineation of urban structures by means of a tailored unsupervised data segmentation procedure with scale optimization, (ii) characterization of urban structures by a joint exploitation of multi-sensor data, (iii) selection of most feasible features under consideration of *in situ* vulnerability information, and (iv) estimation of seismic vulnerability levels of urban structures within a supervised learning framework, was proposed.

Numerical results obtained for the city of Istanbul confirm viability of procedures. When estimating damage grades for Istanbul's district Zeytinburnu with SVR, best models are characterized by mean absolute percentage errors less than 11% and fairly strong goodness of fit ($R > 0.75$). When aiming to identify types of urban structures (i.e., urban structures determined by large industrial/commercial buildings that can be considered as highly vulnerable, and urban structures determined by tall detached residential buildings that can be considered as slightly vulnerable), results obtained with *C*-SVM showed distinctive increase of accuracies when compared to results obtained with ensembles of *v*-OC-SVM. This is reasonable since *C*-SVM could rely on more prior knowledge. Ensembles of *v*-OC-SVM were not able to exceed moderate agreements, with κ statistics slightly above 0.45. Instead, *C*-SVM allowed to obtain κ statistics showing substantial and even excellent agreements ($\kappa > 0.6$ up to $\kappa > 0.8$).

Besides, during experiments we observed that elevation information, as provided by the nDSM, used alone does not allow to learn viable models. However, when jointly exploited with optical data, learned models exhibit viable accuracies. It is worth noting that the idea to

model spatial context relations in feature space with supersegment information has also shown very favorable results and in most cases enabled to retrieve viable accuracies at all. In this regard, feature selection algorithms proved generally useful to alleviate problems associated with high-dimensional feature vectors in conjunction with a comparatively small number of labeled samples. All in all, analyzes provide promising empirical evidence, which confirms the potential of remote sensing to support the seismic vulnerability assessment of urban structures.

However, from a methodological perspective, future work may exploit semi-supervised approaches (e.g. Bruzzone et al., 2006), which encode some knowledge from the unlabeled data also. They may yield favorable accuracies especially in situations where only very few labeled samples are available. Ascertainment of labeled samples can generally be considered to be very costly within this application context. Therefore, the idea of a targeted estimate, i.e., identifying the classes of interest by having only labeled samples of the classes of interest available and not for all classes present in the data, should be further followed. In this manuscript, we deployed ensembles of *v*-OC-SVM for this task. Thereby, identification of most suitable hyperparameters is generally very difficult and we used a heuristic which evaluates overall accuracy and model complexity, expressed by number of support vectors. However, more sophisticated approaches may identify more suitable combinations of free parameters (e.g., approaches that aim to characterize and assess the shape of the decision function of learned models). In addition, unsupervised pre-classification of the feature space has shown promising results recently (Krawczyk et al., 2014). Besides, other machine learning one-class approaches (e.g., ensembles of one class random forests; Desir et al., 2013) can be considered as an interesting opportunity. Regarding alternative techniques for function estimation, Gaussian process regression (Rasmussen & Williams, 2006) has shown distinctive accuracies recently, when compared to other machine learning regression techniques (e.g., Verrelst et al., 2012).

From a conceptual point of view, an intensive exchange with the earthquake engineering community appears exigent. Definition of typologies that need to be estimated in a standardized way and commitment of benchmark accuracies that need to be met would allow a rigorous evaluation of approaches. This may contribute to a further determination of the role of remote sensing within this emerging research field.

7. APPENDIX A

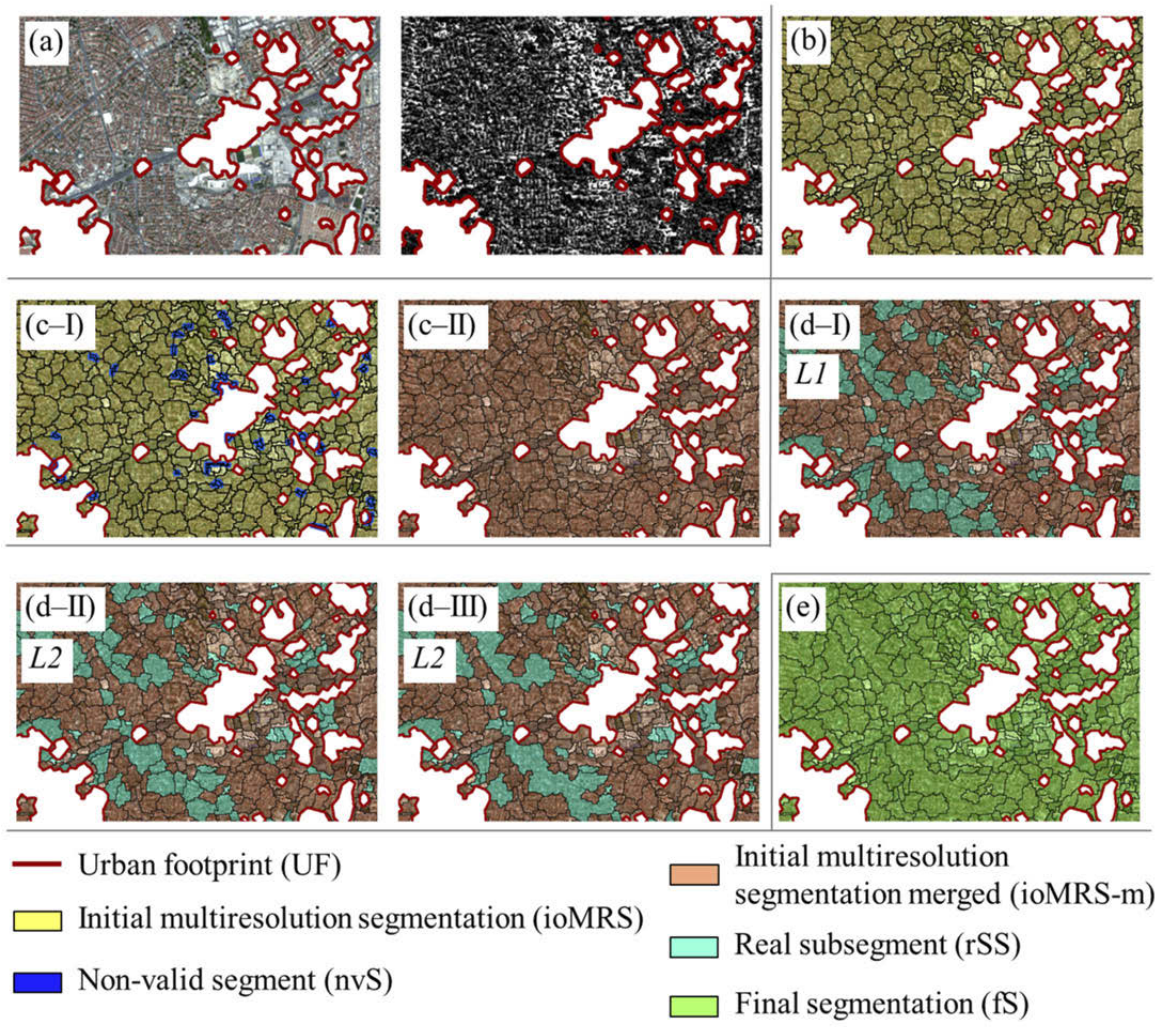


Fig. VI-13. Illustration of some basic processing steps of the image segmentation procedure. It starts with the incorporation of the UF as basic segmentation level (a). The initial optimal multiresolution segmentation (b) (ioMRS) is determined with a plateau objective function. Then, segments are identified that do not correspond to the Shannon sampling theorem (c-I) and are considered as non-valid segments (nvS). Subsequently, non-valid segments are merged with most similar adjacent segments (c-II) (ioMRS-m). Subsequently, distinctive subsegments (d-I) (rSS) are determined and transferred to the supersegment level (d-II). They are merged if a similarity constraint is fulfilled (d-III). The final segmentation (fS) is identified by reassessment with the objective function after multilevel segmentation procedure (e).

ACKNOWLEDGEMENTS

We want to acknowledge the support by the European Commission's Seventh Framework Programme [FP7/2007-2013], under grant agreement no. 312972 "Framework to integrate Space-based and in-situ sENSing for dynamic vUlnerability and recover Monitoring". In addition, research was founded partially by Helmholtz-EOS (Earth Observation System). We

thank Andreas Felbier and Thomas Esch (DLR) for computing and providing the urban footprint data, and the TanDEM-X science team for providing TanDEM-X intermediate DSM data for the test site.

Chapter VII

Synthesis

1. SUMMARY AND MAIN CONCLUSIONS

From a general perspective, the overarching goal of this thesis was to develop, apply, and evaluate tailored methods and procedures that allow for a seismic vulnerability assessment of the built environment based on remote sensing data. This thesis provides very promising results, which show that remote sensing has a high capability to contribute to a rapid screening assessment of the seismic vulnerability of buildings and urban structures. Experimental results obtained for the city of Padang (Indonesia) and Istanbul (Turkey), confirm the viability of combining features derived from remote sensing with *in situ* observations to estimate seismic vulnerability levels of buildings and urban structures. To this purpose, we used and designed supervised learning procedures. However, this also determines both the availability of training data and finding robust features from remote sensing that allow for a discriminative characterization of *in situ* observations. In this sense, adaption to other cities (in other regions/countries) must be done with care, and local idiosyncrasies cannot be bypassed. Nevertheless, when being able to infer models with viable accuracies, as shown in this thesis, an assessed built environment allows to quantify expectable earthquake loss and make damage estimations spatially explicit in a consistent manner. These are key features of significant earthquake loss modeling and predictions and naturally enable mitigation measures.

Moreover, a number of *more detailed conclusions* can be drawn. In this sense, a *review of the scientific literature* uncovered a number of points that motivated and directed research objectives throughout this thesis. Generally, it is found that earthquake vulnerability–centered assessments incorporating remote sensing data are increasing primarily in recent years. This goes along with a changing perspective of the scientific community which considers the assessment of vulnerability and its constituent elements as a pivotal part of a comprehensive risk analysis. Thereby, the availability of new sensors systems enables an appreciable share of remote sensing first. Vulnerability-related building parameters that were extracted from remote sensing data in previous studies incorporate building footprint, height, shape characteristics, roof materials, location, period of construction and structure type (Mueller et al., 2006; Sarabandi et al., 2008; Taubenböck et al., 2009a; Borfecchia et al., 2010; Sahar et al., 2010; Borzi et al., 2011). Very high and high spatial resolution optical imagery is found to be suitable to quantify and characterize the building stock based on manual cartographic methods, statistical enumeration of samples (Ehrlich et al., 2010), or automatic image

information extraction methods (Sahar et al., 2010; Borzi et al., 2011). Especially the latest generation of optical spaceborne sensors are perceived as a breakthrough for operational applications particularly where no alternative data source is available such as in third world countries and smaller and medium size remote urban areas (Deichmann et al., 2011). Combining optical sensors and LiDAR data allows for the automated evaluation of seismic building vulnerability with a viable accuracy (Borfecchia et al., 2010), whereas, for example, the joint use of optical and SAR data allows to derive crucial parameters such as building footprint and floor number (Polli and Dell'Acqua, 2011).

However, the review showed also that previous studies have evaluated the potential of remote sensing for seismic vulnerability assessment in a solely qualitative manner (Mueller et al., 2006) or presented results that emphasize the viability of the use of remote sensing (Borfecchia et al., 2010), but lack the identification and documentation of the necessary and meaningful features. That is why we raised the questions: (i) Which features can be derived from satellite remote sensing data that best explain seismic building vulnerability?; (ii) How suitable are features derived from satellite remote sensing data for estimating seismic building vulnerability levels? To answer these questions, we derived a set of features from high resolution optical imagery, height information from a normalized digital surface model, and multi-temporal medium resolution optical data to characterize the urban environment for the city of Padang (Indonesia). Features are computed on individual building level, and in addition on building block level to characterize the spatial context the buildings are embedded in. *In situ* information was integrated to describe the vulnerability level of buildings according to a scoring method and the EMS-98 scheme. To identify features that best explain seismic building vulnerability, we deployed ordinary least squares regression models for the scoring method and generalized ordered logit regression models to examine the influence of the features on the respective EMS-98 class. The influence and direction of the features vary considerably in dependency of the applied *in situ* seismic vulnerability assessment. However, features describing building height characteristics and features related to the geometry of the individual buildings turned out to generally explain seismic vulnerability to a significant degree. When assessing the vulnerability level according to the scoring method, the overall mean absolute percentage error is 10.6%, if using a supervised Support Vector Regression approach and a certain subset of the features, as identified with a feature selection algorithm. When predicting EMS-98 classes, the results show an overall accuracy of 65.4% and a κ statistic of 0.36, if using a naïve Bayes learning scheme. The individual urban environment of Padang with respect to existing buildings, their physical appearance and affiliated seismic

vulnerability hampers a non-adapted transfer of empirical findings to other earthquake prone regions. Although local idiosyncrasies cannot be bypassed, results provide useful insights into drivers of seismic building vulnerability as measurable with remote sensing, and give details about expectable viability of procedures when aiming to estimate seismic vulnerability levels of buildings.

From an *application-oriented point of view*, the assessment of the seismic vulnerability of the building inventory is driven by the affiliated possibility to estimate earthquake loss prior to a potentially damaging event. In this sense, we aimed to complete the previous analyses for Padang by estimating seismic building structural types. They reflect the main load-bearing structure of buildings and, thus, the behavior under seismic load. In addition, information about building seismic structural types can support an assessment according to e.g., the EMS-98. To this purpose, we designed a hierarchical supervised classification scheme. It is based on five consecutive main steps dedicated to: (i) calculation of features from remote sensing data, (ii) feature selection, (iii) outlier detection, (iv) generation of synthetic samples, (v) and supervised classification under consideration of both support vector machines and random forests. Experimental results obtained for Padang confirm the potential for a viable area-wide and spatially consistent estimation of seismic building structural types. Thereby, it turned out that a combinatory use of features from different feature categories related to building shape and height, spatial context, and spectral information appeared most promising. To demonstrate the applicability of the approach for earthquake loss estimation, we combined this information with available fragility functions for designated building types, which were derived empirically after the last severe earthquake in Padang on 30th September 2009. This way, the approach allowed to quantify building damage in a detailed way and make damage estimations spatially explicit by e.g., localizing hot spots within a city. These are key features for significant earthquake loss modeling and predictions.

Up to this point of the thesis, analyses were carried out on *building level*. As discussed, from a conceptual point of view, this induces the deployment of very high spatial resolution remote sensing data. Nowadays, data costs and processing requirements hamper utilization capabilities for larger areas. When aiming at spatially continuous and consistent assessment approaches that are applicable for large regions, on a national level, or even globally, those data represent a clear limitation. In accordance with the synthesis from a quantitative analysis of the scientific literature reviewed in chapter III, which suggested to "addresses settlement scales on a coarser morphological level such as structure types that can be detected, characterized and assessed properly and cost-effective (see also Wyss 2012) when using

remotely sensed data" (Geiß and Taubenböck, 2013, p. 29), we alter the spatial scale of analysis. As mentioned, from a conceptual point of view, urban environments are perceived as systems that exhibit a hierarchical arrangement. At the lowest level of aggregation, individual objects such as buildings can be considered, followed by homogeneous urban structures, and end with a spatially indifferent settlement area at the highest level of aggregation. Hence, a coarser level of urban morphology was addressed when compared to building level in the subsequent analysis. However, this simultaneously allowed relying on remote sensing data with a lower spatial resolution but larger spatial coverage. In particular, we exploited multispectral data from the RapidEye constellation, elevation information from the TanDEM-X mission, and data from the Landsat archive.

Before, we aimed to extract elevated objects in urban environments (i.e., buildings) from TanDEM-X data to be able to infer valid height information in the subsequent analyses. To this purpose, it is common to derive a digital terrain model from the data first. A digital terrain model contains elevation measurements of the bare earth without including objects above ground. This information allows computing a normalized digital surface model. The latter then naturally comprises elevation information of objects above ground. To retrieve this information it is common to deploy approaches based on mathematical morphology due to their simplicity in accordance with good performance. However, we introduce a novel region growing-based progressive morphological filter for this task to address challenges associated with the use of morphological filters in non-flat terrain, and overcome individual challenges related to the spatial resolution of TanDEM-X data. The approach comprises a multistep procedure using concepts of morphological image filtering, region growing and interpolation techniques. It is based on the idea of progressive morphological filters that aim to discriminate ground and non-ground pixels in the digital surface model based on algebraic set operations. Experimental results obtained for large parts of the settlement are of Izmir (Turkey) reveal promising empirical evidence. Errors of omission could be considerably reduced while enforcing a solely moderate increase of commission errors when discriminating ground and non-ground pixels in the digital surface model. Thus, a beneficial performance compared to basic morphological filter-based methods in terrain with high orographic energy and steep areas could be shown.

Chapter VI documents an approach which is intended to alleviate restrictions that are nowadays inherent to the use of very high resolution remote sensing data. As discussed, a coarser level of urban morphology was addressed when compared to building level. We suggested procedures which learn prior seismic vulnerability information to assess seismic

vulnerability of urban structures in a spatially continuous and consistent manner based on remote sensing. Four main steps are carried out, which are dedicated to: i) delineation of urban structures by means of an original and tailored unsupervised data segmentation procedure with scale optimization; ii) characterization of urban structures by a joint exploitation of multi-sensor data; iii) selection of most feasible features under consideration of *in situ* vulnerability information; iv) estimation of seismic vulnerability levels of urban structures within a supervised learning framework. We render the prediction problem in three ways to address possible real-world, operational requirements. Experimental results are reported for the mega city Istanbul (Turkey), which faces an enormous seismic threat. Numerical results confirm viability of procedures. When estimating damage grades for Istanbul's district Zeytinburnu with a Support Vector Regression approach, best models are characterized by mean absolute percentage errors less than 11% and fairly strong goodness of fit ($R > 0.75$). When aiming to identify types of urban structures (i.e., urban structures determined by large industrial/commercial buildings that can be considered as highly vulnerable, and urban structures determined by tall detached residential buildings that can be considered as slightly vulnerable), results obtained with *C*-SVM showed distinctive increase of accuracies when compared to results obtained with ensembles of *v*-OC-SVM. This is reasonable since *C*-SVM could rely on more prior knowledge. Ensembles of *v*-OC-SVM were not able to exceed moderate agreements, with κ statistics slightly above 0.45. Instead, *C*-SVM allowed to obtain κ statistics showing substantial and even excellent agreements ($\kappa > 0.6$ up to $\kappa > 0.8$). Besides, during experiments we observed that elevation information, as provided by the nDSM, used alone does not allow to lean viable models. However, when jointly exploited with optical data, learned models exhibit viable accuracies. It is worth noting that the idea to model spatial context relations in feature space with supersegment information (i.e., we computed features for two additional supersegment levels) has also shown very favorable results and in most cases enabled to retrieve viable accuracies at all. In this regard, feature selection algorithms proved generally useful to alleviate problems associated with high-dimensional feature vectors in conjunction with a comparatively small number of labeled samples. All in all, analyses provide promising empirical evidence, which confirms the potential of remote sensing to support the seismic vulnerability assessment of urban structures.

2. FUTURE RESEARCH

Overall, this thesis provides very promising results, which show that remote sensing has a high capability to contribute to a rapid screening assessment of the seismic vulnerability of buildings and urban structures. Further work can build upon these results and may challenge empirical findings in further case studies, enhance developed and applied methods, transfer concepts and approaches to other sensor systems/data sources, or apply data and methodologies within integrative and holistic risk assessment strategies.

From a *technical point of view*, it would be interesting if follow-up research would systematically evaluate the current high-end capabilities of remote sensing to characterize urban environments and subsequently assess their vulnerability. For instance, evaluation of usability of detailed 3-D building models (e.g., Bulatov et al., 2014; Xiong et al., 2014) is largely absent, and current studies (including this thesis) deployed very basic models (i.e., building footprints with elevation information). For characterization and discrimination of urban structures, spectral super-resolution information as collected by e.g., Sentinel-2, or full-polarized synthetic aperture radar data, may yield interesting opportunities.

From a *methodological point of view*, future work may enhance deployed methods and also expand procedures by relevant components. For the enhanced characterization of urban environments, interesting features may be deployed by further analysis of spatial, structural, and contextual relations (e.g., Histogram of Oriented Gradients, Line Support Regions, lacunarity etc.; Graesser et al., 2012). In this sense, we deployed GLCM measures for the characterization of image textures. However, texture measures based on local binary patterns (Ojala et al., 2002; Musci et al., 2013) can be considered as an interesting opportunity to enhance predictive performance of models. Regarding the actual classification task, future work may exploit semi-supervised approaches (e.g. Bruzzone et al., 2006), which encode some knowledge from the unlabeled data also. This may yield favorable accuracies especially in situations where only very few labeled samples are available. Ascertainment of labeled samples can generally be considered to be very costly within this application context. Therefore, active learning strategies (Tuia et al., 2009; Tuia et al., 2011) may be followed to solve estimation of vulnerability levels efficiently, based on remote sensing data in a supervised fashion. Active learning deploys predefined heuristics to rank unlabeled instances in the domain under analysis that are considered the most valuable for improvement of estimation accuracy. Once instances are selected they are labeled and the process is iterated. Thus, active learning starts with a minimal, non-optimal set of labeled samples and the model

builds the optimal set of labeled samples itself to minimize prediction error. Recent approaches also include the spatial domain for this task (Stumpf et al., 2014; Pasolli et al., 2014), and consider labeling costs emerging from ground surveys (i.e., labeling costs of samples are modeled by a cost function that exploits road information and digital elevation model data; Demir et al., 2014). Such approaches may guide civil engineers through the building inventory to collect most valuable *in situ* information efficiently.

From a *conceptual point of view*, this thesis followed a very narrow definition of seismic vulnerability, i.e., vulnerability was regarded as the expected degree of damage (loss) to a given element at risk resulting from a given level of seismic hazard (i.e., ground shaking). However, also with respect to a more holistic understanding of vulnerability, which may explicitly consider e.g., social, economic, or political factors, information describing the expectable behavior of the building inventory may be crucial, and can be combined with other data for a holistic risk assessment by e.g., index overlays (Taubenböck et al., 2008). Nevertheless, a spatially detailed and continuously assessed building inventory opens the possibility to estimate potential earthquake loss under consideration of small scale hazard effects. In this sense, validation of model estimates based on experienced damage is an exigent task to gain insights about viability of remote sensing-based procedures. Lastly, an intensive exchange with the earthquake engineering community appears exigent. Definition of typologies that need to be estimated in a standardized way and commitment of benchmark accuracies that need to be met would allow a rigorous evaluation of approaches. This may contribute to a further determination of the role of remote sensing within this emerging research field.

REFERENCES

A

Adams JB, Huyck KC, Mansouri B, Eguchi TR, Shinozuka M (2004) Application of high-resolution optical satellite imagery for post-earthquake damage assessment: The 2003 Boumerdes (Algeria) and Bam (Iran) earthquakes. Research progress and accomplishments: 2003-2004, Multidisciplinary Center for Earthquake Engineering Research (MCEER), pp 173-186

ADB (2011) Asian Development Bank. Disaster risk management. Available online at http://www.adb.org/disaster/ Accessed 17 Oct 2011

Adger WN, Hughes TP, Folke C, Carpenter SR, Rockstroem J (2005) Social-ecological resilience to coastal disasters. Science 309:1036-1039

Adger WN, Kelly PM, Ninh HN (2001) Living with environmental change. Social vulnerability, adaption and resilience in Vietnam. Research Global Environmental Change. Routhledge, London

Adger WN (2006) Vulnerability. Global Environmental Change 16:268-281

Alcantara-Ayala I (2002) Geomorphology, natural hazards, vulnerability and prevention of natural disasters in developing countries. Geomorphology 47:107-124

Aldrighi M, Dell'Acqua F (2009) Mode-based method for matching of pre- and postevent remotely sensed images. IEEE Geosci Remote Sens Lett 6(2):317-321

Alexander DE (2002) Principles of emergency planning and management. Terra Publishing, Harpenden

Alwang J, Siegel PB, Jorgensen SL (2001) Vulnerability: A view from different disciplines. Social protection discussion paper series No. 0115. Social Protection Unit, Human Development Network, The World Bank, Washington D.C.

Arellano-Baeza AA, Zverev AT, Malinnikov VA (2006) Study of changes in the lineament structure, caused by earthquakes in South America by applying the lineament analysis to the Aster (Terra) satellite data. In: Natural Hazards and Oceanographic Processes from Satellite Data, vol 37. Advances in Space Research. 37(4), pp. 690-697

Aubrecht C, Özceylan D, Steinnocher K, Freire S (2012) Multi-level geospatial modeling of human exposure patterns and vulnerability indicators. Taubenböck H, Post J, Strunz G (eds.) Remote Sensing contributing to mapping earthquake vulnerability and effects. Special Issue in Natural Hazards. To be published

Avouac JP, Ayoub F, Leprince S, Konca O, Helmberger DV (2006) The 2005, Mw 7.6 Kashmir earthquake: Sub-pixel correlation of ASTER images and seismic waveforms analysis. Earth and Planetary Science Letters 249:514-528

Axelsson, P (1999) Processing of laser scanner data – algorithms and applications," ISPRS J. Photogramm. 54:138-147

Aydöner C, Maktav D (2009) The role of the integration of remote sensing and GIS in land use/land cover analysis after an earthquake. Int J Remote Sens 30(7):1697-1717

Ayoub F, Leprince S, Avouac JP (2009) Co-registration and correlation of aerial photographs for ground deformation measurements. ISPRS Journal of Photogrammetry and Remote Sensing 64:551-560

B

Bailey TC, Gatrell AC (1995) Interactive Spatial Data Analysis, John Wiley and Sons, New York

Balz T, Liao M (2010) Building-damage detection using post-seismic high-resolution SAR satellite data. Int J Remote Sens 31(13):3369-3391

Banzhaf E, Höfer R (2008) Monitoring Urban Structure Types as Spatial Indicators With CIR Aerial Photographs for a More Effective Urban Environmental Management. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1(2):129–138

Baud I, Kuffer M, Pfeffer K, Sliuzas R, Karuppannan S (2010) Understanding heterogeneity in metropolitan India: The added value of remote sensing data for analyzing sub-standard residential areas. International Journal of Applied Earth Observation and Geoinformation 12:359-374

Bayuaji L, Sumantyo JTS, Kuze H (2010) ALOS PALSAR D-InSAR for land subsidence mapping in Jakarta, Indonesia. Canadian Journal of Remote Sensing 36:1-8

Bechtel B, Daneke C (2012) Classification of Local Climate Zones Based on Multiple Earth Observation Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(4):1191-1202

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58(3/4):239-258

Begg JG, Mouslpoulou V (2007) Rangitaiki Plains: another veil is lifted: New lidar data from Bay of Plenty, New Zealand. Geological Society of New Zealand Miscellaneous Publications 123A, New Zealand

Bendea H, Boccardo P, Dequal S, Giulio Tonolo F, Marenchino D, Piras M (2008) Low cost UAV for post-disaster assessment. Proceedings of The XXI Congress of the International Society for Photogrammetry and Remote Sensing, Beijing (China), 3-11 July 2008

Benson C, Twigg J, Myers M (2001) NGO Initiatives in Risk Reduction: An Overview. Disasters 25(3):199-215

Berger M, Aschbacher J (2012) Preface: The sentinel missions-new opportunities for science. Remote Sensing of Environment 120:1-2

Bilham R (2009) The seismic future of cities. Bulletin of earthquake engineering 7:839-887

Birkmann J (2006a) Indicators and criteria for measuring vulnerability: Theoretical bases and requirements. In: Birkmann J (ed) Measuring vulnerability to natural hazards. United Nations University Press, New York, pp 55-77

Birkmann J (2006b) Measuring vulnerability to promote disaster-resilient societies: Conceptual frameworks and definitions. In: Birkmann J (ed) Measuring vulnerability to natural hazards. United Nations University Press, New York, pp 9-54

Birkmann J (2007) Risk and vulnerability indicators at different scales: Applicability, usefulness and policy implications. Environmental Hazards 7(1):20-31

Birkmann J, Krings S, Vollmer M, Wolfertz J, Welle T, Kühling W, Meisel K, Wurm M, Taubenböck H, Gähler M, Zwenzner H, Roth A, Voigt S, Dech S (2011) Indikatoren zur Abschätzung von Vulnerabilität und Bewältigungspotentialen am Beispiel von wasserbezogenen Naturgefahren in urbanen Räumen. Forschung im Bevölkerungsschutz, 13, Bonn, Germany

Blaikie P, Cannon T, Davis I, Wisner B (1994) At Risk: Natural hazards, peoples vulnerability and disasters. Routledge, London

Blaschke T (2010) Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65:2-16

Blaschke T, Hay GJ, Kelly M, Lang S, Hofmann P, Addink E, Queiroz Feitosa R, van der Meer F, van der Werff H, van Coillie F, Tiede D (2014) Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87:180-191

Blaschke T, Hay GJ, Weng Q, Resch B (2011) Collective Sensing: Integrating geospatial Technologies to understand urban systems - An overview. Remote Sensing 3(8):1743-1776

BNPB – Badan Nasional Penanggulangan Bencana (National Disaster Management Agency, Indonesia), 2009. Laporan Harian Pusdalops BNPB Minggu, http://www.bnpb.go.id (Accessed 18 May, 2013)

Bochow M, Segl K, Kaufmann H (2007) Automating the built-up process of feature-based fuzzy logic models for the identification of urban biotopes from hyperspectral remote sensing data. Urban Remote Sensing Joint Event, Paris, France

Bogardi J, Birkmann J (2004) Vulnerability assessment: The first step towards sustainable risk reduction. In: Malzahn D, Plapp T (eds) Disaster and society - from hazard assessment to risk reduction. Logos Berlin, pp 72-82

Bogardi JJ, Villagrán De Léon JC, Birkmann J, Renaud F, Sakulski D, Chen X, Affeltranger B, Mensa A, Kaplan M (2005) Vulnerability in the context of climate change. Human Security and Climate Change, An International Workshop Holmen Fjord Hotel, Asker, near Oslo, 21–23 June 2005, p. 14

Bohle H-G (2001) Vulnerability and criticality perspectives from social geography. IHDP Update 2/2001, Newsletter of the International Human Dimensions Programme on Global Environmental Change 2:1-7

Bohle H-G, Downing TE, Watts MJ (1994) Climate change and social vulnerability. Towards a sociology and geography of food insecurity. Global Environmental Change 4 (1):37-48

Bollin C (2003) Community-based disaster risk approach. Experience gained in Central America. GTZ - Deutsche Gesellschaft für technische Zusammenarbeit, Eschborn

Bommer JJ, Rodriguez CE (2002) Earthquake-induced landslides in Central America. Engineering Geology 63:189-220

Borfecchia F, Pollino M, De Cecco L, Lugari A, Martini S, La Porta L, Ristoratore E, Pascale C (2010) Active and passive remote sensing for supporting the evaluation of the urban seismic vulnerability. Italian Journal of Remote Sensing 42(3):129-141

Borzi B, Dell'Acqua F, Faravelli M, Gamba P, Lisini G, Onida M, Polli D (2011) Vulnerability study on a large industrial area using satellite remotely sensed images. Bulletin of Earthquake Engineering 9:675-690

Boyles S, Fajardo D, Waller ST (2007) A Naïve Bayesian Classifier for Incident Duration Prediction. In Proceedings of the 86th Annual Meeting of the Transportation Research Board. Washington D.C., USA

Breiman L (2001) Random Forests. Machine Learning 45:5–32

Breiman L (2002) Manual on setting up, using, and understanding random forests V3.1. Department of Statistics, University of California, Berkeley, http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf (Accessed 8 Feb, 2013)

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and Regression Trees. Wadsworth

Briuglio L (1995) Methodological and practical considerations for constructing socio-economic indicators to evaluate disaster risk. In: Programme on Information and Indicators for Risk Management. Institute of Environmental Studies, University of Columbia. IADB-ECLAC-IDEA, Manizalez, Colombia

Brooks N (2003) Vulnerability, risk and adaption: A conceptual framework. Tyndall Centre for Climate Change Research, working paper 38(16)

Brunner D, Lemoine G, Bruzzone L (2010) Earthquake damage assessment of buildings using VHR optical and SAR imagery. IEEE Transactions on Geoscience and Remote Sensing 48:2403-2420

Bruzzone L, Carlin L (2006) A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images. IEEE Transactions on Geoscience and Remote Sensing 44(9):2587-2600

Bruzzone L, Chi M, Marconcini M (2006) A Novel Transductive SVM for Semisupervised Classification of Remote-Sensing Images. IEEE Transactions on Geoscience and Remote Sensing 44(11):3363–3373

Bulatov D, Häufel G, Meidow J, Pohl J, Solbrig P, Wernerus P (2014) Context-based automatic reconstruction and texturing of 3D urban terrain for quick-response tasks. ISPRS Journal of Photogrammetry and Remote Sensing 93:157-170

Burby RJ (1991) Sharing environmental risks: How to control governments' losses in natural disasters. Westview Press, Boulder

Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge discovery 2:121–147

Burton I, Kates RW, White GF (1993) The environment as hazard. 2nd edn. Guildford Press, New York

C

Cakti E (2013) Issues with the earthquake vulnerability of Istanbul. Natural Hazards 68(1):227-228

Calvi GM (1999) A displacement-based approach for vulnerability evaluation of classes of buildings. Journal of Earthquake Engineering 3(3):411-438

Calvi GM, Pinho R, Magenes G, Bommer JJ, Restrepo-Velez LF, Crowley H (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. ISET Journal of Earthquake Technology 43(3):75-104

Camps-Valls G, Bruzzone L (2005) Kernel-based methods for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing 43:1351-1362

Camps-Valls G, Bruzzone L (2009) Kernel Methods for Remote Sensing Data Analysis, John Wiley & Sons, New York.

Camps-Valls G, Tuia D, Bruzzone L, Benediktsson J A (2014) Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. IEEE Signal Processing Magazine 31(10):45-54

Cannon T (1993) A hazard need not a disaster make: vulnerability and the causes of ‘natural’ disasters. In: Merriman PA, Browitt CWA (eds) Natural disasters: Protecting vulnerable communities. Thomas Telford, London, pp 92-105

Cannon T (1994) Vulnerability analysis and the explanation of "natural disasters". In: Varley A (ed) Disasters, development and environment. John Wiley and Sons, Chichester, New York, Brisbane, Toronto, Singapore, pp 13-29

Cannon T, Twigg J, Rowell J (2003) Social vulnerability. Sustainable livelihoods and disasters. Report to DFID Conflict and Humanitarian Assistance Department (CHAD) and Sustainable Livelihood Support

Cardona OD (1999) Environmental management and disaster prevention: Two related topics - A holistic risk assessment and management approach. In: Ingleton J (ed) Natural disaster management. Tudor Rose, IDNDR, London, UK

Cardona OD, Hurtado JE (2000) Holistic seismic risk estimation of a metropolitan center. In: Proceedings of 12th World Conference of Earthquake Engineering, Auckland, New Zealand

Cardona OD (2004) The need for rethinking the concepts of vulnerability and risk from a holistic perspective: A necessary review and criticism for effective risk management. In: Bankoff G, Frerks G, Hilhorst D (eds) Mapping Vulnerability. Disasters, Development & People. Earthscan, London, pp 37-51

Caron C, Roche S, Goyer D, Jaton A (2008) GIScience journals ranking and evaluation : An international delphi study. Transactions in GIS 12(3):293-321

Carreno M-L, Cardona OD, Barbat AH (2007) Urban Seismic Risk Evaluation: A Holistic Approach. Natural Hazards 40:137-172

Carreno M-L, Cardona OD, Marulanda M-C, Barbat AH (2009) Holistic urban seismic risk evaluation of megacities: Application and robustness. In: Mendez-Victor LA, Oliveira CS, Azevedo J, Ribeiro A (eds) The 1755 Lisbon Earthquake: revisited. Springer, pp 167-183

Carreno M-L, Cardona OD, Suarez DC, Barbat AH (2009) Holistic Evaluation of risk in the framework of the urban sustainability. In: II Congreso Internacional de Medida y Modelización de la Sostenibilidad (ICSMM 09), Barcelona, Spain, 2009

Cartwright S (2005) National civil defence emergency management plan order 2005. Published under the authority of the New Zealand Government, Wellington, New Zealand

Catita C, Feigl KL, Catalao J, Miranda JM, Victor LM (2005) InSAR time series analysis of the 9 July 1998 Azores earthquake. Int J Remote Sens 26(13):2715-272

CEOS 2012 Committee on Earth Observation Satellites. Available online at http://www.ceos.org/images/PDFs/CEOS08Brochure.pdf Accessed 12 Jul 2012

Chang CC, Lin CJ (2001) LIBSVM: A library for support vector machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm (Accessed 12 Jan, 2013)

Chang L, Tang Z (2010) Using remote sensing technology to assess land - use changes after the Northridge Earthquake. Disaster Advances 3:5-10

Chapman D (2001) Natural hazards, vol 2. University Press, Oxford

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 16:321-357

Chen CH, Ho, P-GP (2008) Statistical pattern recognition in remote sensing. Pattern Recognition 41(9):2731-2741

Chen K (2002) An approach to linking remotely sensed data and areal census data. International Journal of Remote Sensing 23:37-48

Chen Q, Gong P, Baldocchi D, Xie G (2007) Filtering Airborne Laser Scanning Data with Morphological Methods. Photogramm. Eng. Rem. S. 73(2):175-185

Chen SY, Liu PX, Liu LQ, Ma J, Chen GQ, Hu XY (2011) Comparative analysis between land surface temperatures obtained by field measurement and satellite remote sensing and its implication in earthquake research. Chinese Journal of Geophysics-Chinese Edition 54:747-755

Chini M, Atzori S, Trasatti E, Bignami C, Kyriakopoulos C, Tolomei C, Stramondo S (2010) The May 12, 2008, (Mw 7.9) Sichuan Earthquake (China): Multiframe ALOS-PALSAR DInSAR analysis of coseismic deformation. IEEE Geosci Remote Sens Lett 7(2):266-270

Chini M, Bignami C, Stramondo S, Pierdicca N (2008) Uplift and subsidence due to the 26 December 2004 Indonesian earthquake detected by SAR data. Int J Remote Sens 29:3891-3910

Chini M, Pierdicca N, Emery WJ (2009) Exploiting SAR and VHR Optical Images to Quantify Damage Caused by the 2003 Bam Earthquake. IEEE Transactions on Geoscience and Remote Sensing 47(1):145-152

Chiroiu L, Adams B, Saito K (2006) Advanced techniques in modeling, response and recovery. In: Oliveira CS, Roca A, Goula X (eds) Assessing and managing earthquake risk. Springer, Dordrecht, The Netherlands, pp 427-448

Chlieh M, Avouac JP, Sieh K, Natawidjaja DH, Galetzka J (2008) Heterogeneous coupling of the sumatran megathrust constrained by geodetic and paleogeodetic measurements. Journal of Geophysical Research 113(B5), B05305

Choudhury S, Dasgupta S, Saraf AK, Panda S (2006) Remote sensing observations of pre-earthquake thermal anomalies in Iran. Int J Remote Sens 27:4381-4396

Cieslak DA, Chawla NV, Striegel A (2006). Combating imbalance in network intrusion datasets. In Proceedings of the IEEE International Conference on Granular Computing, pp. 732–737

Coburn AW, Spence RJS (1992) Earthquake protection. John Wiley & Sons Ltd, Chichester, UK

Coburn AW, Spence RJS (2002) Earthquake protection. John Wiley & Sons Ltd, New York, USA

Coburn AW, Spence RJS, Pomonis A (1994) Training Manual: Vulnerability and Risk assessment. 2 edn. UNDP Disaster Management Training Programme, Cambridge Architectural Research Limited, The Oast House, Malting Lane, Cambridge, UK

Colaninno N, Roca J, Pfeffer K (2011) An automatic classification of urban texture: form and compactness of morphological homogeneous structures in Barcelona. 51st European Congress of the Regional Science Association International, pp. 1–20

Corbane C, Carrion D, Lemoine G, Broglia M (2011) Comparison of Damage Assessment Maps Derived from Very High Spatial Resolution Satellite and Aerial Imagery Produced for the Haiti 2010 Earthquake. Earthquake Spectra 27:199-218

Cornell AC (1968) Engineering Seismic Risk Analysis. Bulletin of the Seismological Society of America 58(5):1583-1606

Correira F, Santos M, Rodrigues R (1987) Engineering risk in regional drought studies. In: Duckstein L, Plate EJ (eds) Engineering reliability and risk in water resources. Nijhoff M, Dordrecht/Boston, pp 61 - 86

Corell R, Cramer W, Schellnhuber HJ (2001) Methods and models of vulnerability research, analysis and assessment. In: Symposium at Potsdam Sustainability Days, Potsdam, 2001

Cortes C, Vapnik V (1995) Support Vector Networks. Machine Learning 20(3):273-297

Crichton D (1999) The risk triangle. In: Ingleton J (ed) Natural disaster management. Tudor Rose, London, pp 102-103

CrisisMappers (2012) The humanitarian technology network. Available online at http://crisismappers.net/ Accessed 12 Jul 2012

Crowley H, Bommer J (2006) Modelling seismic hazard in earthquake loss models with spatially distributed exposure. Bulletin of Earthquake Engineering 4:249-273

Crowley H, Pinho R, Bommer JJ (2004) A probabilistic displacement-based vulnerability assessment procedure for earthquake loss estimation. Bulletin of Earthquake Engineering 2:173-219

Crowley H, Pinho R, Bommer JJ, Bird JF (2006) Development of a displacement-based method for earthquake loss assessment. Report 2006/01. European School for Advanced Studies in Reduction of Seismic Risk (ROSE School), Pavia, Italy

CSA (2012) Canadian Space Agency - RADARSAT Mission description. Available online at http://www.asc-csa.gc.ca/eng/satellites/radarsat/description.asp Accessed 8 May 2012

Cunningham D, Grebby S, Tansey K, Gosar A, Kastelic V (2006) Application of airborne LiDAR to mapping seismogenic faults in forested mountainous terrain, southeast Alps, Slovenia. Geophysical Research Letters 33:L20308, doi:20310.21029/22006GL027014

Cutter S (2003) Social Vulnerability to Environmental Hazards. Social Science Quarterly 84(2):242-261

Cutter S, Emrich C, Webb J, Morath D (2009) Social vulnerability to climate variability hazards: A review of the literature. Final Report to Oxfam America

Cutter SL (1996) Vulnerability to environmental hazards. Progress in Human Geography 20(4):529-539

D

Daniell J (2009) Open source Procedure for assessment of loss using global earthquake modelling – OPAL-GEM Project. CEDIM Research Report 09/01, Karlsruhe, Germany

Daniell JE (2011) Open source procedure for assessment of loss using global earthquake modelling software (OPAL). Natural Hazards and Earth System Sciences 11:1885-1900

Davidson R (1997) An urban earthquake disaster risk index. John A. Blume Earthquake Engineering Center, Department of Civil Engineering. Stanford University, Stanford

Dech S (1997) Anwendung der Satellitenfernerkundung. Von der geowissenschaftlichen Forschung zum operationellem Einsatz. DLR-Forschungsbericht 97, 52

Deichmann U, Ehrlich D, Small C, Zeug G (2011) Using high resolution satellite data for the identification of urban natural disaster risk. Global Facility for Disaster Reduction and Recovery: Washington, DC, USA, 80p.

Dell'Acqua F, Lisini G, Gamba P (2009) Experiences in optical and SAR imagery analysis for damage assessment in the Wuhan, May 2008 Earthquake. IEEE International Geoscience and Remote Sensing Symposium 2009, vols 1-5:2417-2420

Dell'Acqua F, Gamba P (2012) Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. Proceedings of the IEEE 100(10):2876-2890

Demir B, Minello L, Bruzzone, L (2014) Definition of Effective Training Sets for Supervised Classification of Remote Sensing Images by a Novel Cost-Sensitive Active Learning Method. IEEE Transactions on Geoscience and Remote Sensing 52(2): 1272-1284

Demirkesen AC (2008) Digital terrain analysis using Landsat-7 ETM+ imagery and SRTM DEM: a case study of Nevsehir province (Cappadocia), Turkey. International Journal of Remote Sensing 29:4173-418

Deng X, Jiang Q, Ge Q, Yang L (2010) Impacts of the Wenchuan earthquake on the Giant Panda nature reserves in China. Journal of Mountain Science 7:197-206

DESDynI (2011) Deformation, Ecosystem Structure and Dynamics of Ice - Mission concept. Available online at http://desdyni.jpl.nasa.gov/mission/ Accessed 7 Oct 2011

Desir C, Bernard S, Petitjean C, Heutte L (2013) One class random forests. Pattern Recognition, 46(12):3490-3506

DigitalGlobe (2012) WorldView-3. Available online at http://www.digitalglobe.com/downloads/WorldView3-DS-WV3-Web.pdf Accessed 24 May 212

Dilley M, Chen RS, Deichmann U, Lerner-Lam AL, Arnold M, Agwe J, Buys P, Kjekstad O, Lyon B, Yetman G (eds) (2005) Natural disaster hotspots. A global risk analysis, vol 5. Disaster risk management series. The World Bank Hazard Management Unit, Washington D.C.

Ding X, Huang W (2011) D-InSAR monitoring of crustal deformation in the eastern segment of the Altyn Tagh Fault. International Journal of Remote Sensing 32:1797-1806

DLR-ZKI (2010) Earthquakes in Haiti on January 12, 2010. Available online at http://www.zki.dlr.de/article/1262 Accessed 24 Jun 2011

Drusch M, Del Bello U, Carlier S, Colin O, Fernandez V, Gascon F, Hoersch B, Isola C, Laberinti P., Martimort, P, Meygret A, Spoto F, Sy O, Marchese F, Bargellini P (2012) Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment 120:25-36

Dobson JE, Bright EA, Coleman PR, Durfee RC, Worley BA (2000) LandScan: A global population database for estimating populations at risk. Photogrammetric Engineering and Remote Sensing 66:849-857

Dominguez S, Avouac JP, Remi M (2003) Horizontal coseismic deformation of the 1999 Chi-Chi earthquake measured from SPOT satellite images: Implications for the seismic cycle along the western foothills of central Taiwan. Journal of geophysical research 108 (0.1029/2001JB000951)

Douglas J (2007) Physical vulnerability modelling in natural hazard risk assessment. Natural Hazards and Earth System Science 7(2):283-288

Downing TE, Butterfield R, Cohen S, Huq S, Moss R, Rahman A, Sokona Y, Stephen L (2001) Vulnerability indicies: Climate change impacts and adaption. Policy Series 3. United Nations Development Programme, Nairobi, Kenya

Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V (1997) Support vector regression machines. Advances in Neural Information Processing Systems, 9:155-161

Duarah BP, Phukan S (2011) Understanding the tectonic behaviour of the Shillong plateau, India using remote sensing data. Journal of the Geological Society of India 77:105-112

Duch W (2006) Filter methods, Feature Extraction: Foundations and Applications (Studies in Fuzziness and Soft Computing), Guyon I, Gunn S, Nikravesh M, Zadeh LA, Eds., chapter 3, pp. 90–117, Springer, New York

Duda R, Hart P (1973) Pattern Classification and Scene Analysis, John Wiley and Sons, New York

Dunbar PK, Bilham, RG, Laituri, MJ (2003) Earthquake Loss Estimation for India Based on Macroeconomic Indicators, Risk Science and Sustainability. NATO Science 112:163-180

E

Ebert A, Kerle N, Stein A (2009) Urban social vulnerability assessment with physical proxies and spatial metrics derived from air- and spaceborne imagery and GIS data. Natural Hazards 48:275-294

EERI – Earthquake Engineering Research Institute (2009) Learning from Earthquakes – The Mw 7.6 Western Sumatra Earthquake of September 30, 2009, EERI Special Earthquake Report, Dec. 2009, USA

e-GEOS (2012) e-GEOS: Emergency Response Mapping. Available online at http://www.eurimage.com/applications/emergency.html Accessed 11 Jul 2012

Ehrlich D, Guo H, Molch K, Ma JW, Pesaresi M (2009) Identifying damage caused by the 2008 Wenchuan earthquake from VHR remote sensing data. International Journal of Digital Earth 2(4):309-326

Ehrlich D, Zeug G (2008) Assessing disaster risk of building stock. Methodology based on earth observation and geographical information systems. JRC Scientific and Technical Reports. European Comission Joint Research Centre - Institute for the Protection and Security of the Citizen.

Ehrlich D, Zeug G, Gallego J, Gerhardinger A, Caravaggi I, Pesaresi M (2010) Quantifying the building stock from optical high-resolution satellite imagery for assessing disaster risk. Geocarto International 25 (4):281-293

EMCA (2011) Earthquake Model Central Asia. Available online at http://www.emca-gem.org/ Accessed 21 Sep 2011

EMME (2011) Earthquake Model of the Middle East Region. Available online at http://www.emme-gem.org/ Accessed 21 Sep 2011

Eineder M, Friedrich A, Minet C, Bamler R, Flerit F, Hajnsek I (2009) Scientific requirements and feasibility on an L-band mission dedicated to measure surface deformation. IEEE International Geoscience and Remote Sensing Symposium, 12.-17. July 2009, Cape Town, South Afrika

Eineder M, Yague-Martinez N, Minet C, Fritz T, Parizzi A (2010) Haiti earthquake surface shift estimation using TerraSAR-X data. Available online at http://supersites.earthobservations.org/TerraSAR-X-DLR-01-27.pdf Accessed 5 Jul 2011

Erdik M., Aydinoglu N, Fahjan Y, Sesetyan K, Demircioglu M, Siyahi B, Durukal E, Ozbey C, Biro Y, Akman H, Yuzugullu O (2003) Earthquake risk assessment for Istanbul metropolitan area. Earthquake Engineering and Vibration 2(1):1-23

Erdik M, Biro YA, Onur T, Sesetyan K, Birgoren G (1999) Assessment of earthquake hazard in Turkey and neighboring regions. Annali di Geofisica, 42(6):1125-1138

Erdik M, Demircioglu M, Sesetyan K, Durukal E, Siyahi B (2004) Earthquake hazard in Marmara Region, Turkey. Soil Dynamics and Earthquake Engineering 24:605-631

Erten E, Reigber A, Hellwich O (2010) Generation of three-dimensional deformation maps from InSAR data using spectral diversity techniques. ISPRS Journal of Photogrammetry and Remote Sensing 65:388-394

ESA (2011) GMES Sentinels - Sentinels Overview. Available online at http://www.esa.int/esaLP/SEM097EH1TF_LPgmes_0.html Accessed 30 Sep 2011

ESA II (2012) ESA - Insurance workshop on earth observation (Natural Catastrophes). Workshop Summary Report, Esrin, Italy, 23-24 Feb 2012. Available online at http://earth.eo.esa.int/workshops/Insurance/files/Insurance_Industry_EO_wshop_report_FINAL.pdf Accessed 13 Jul 2012

Esch T, Dech S, Roth A, Schmidt M, Taubenböck H, Heldens W, Thiel M, Wurm M, Klein D (2009) Monitoring and assessment of urban environments using space-borne earth observation data. In: Kreck A, Rumor M, Zlatanova S, Fendel E (eds) Urban and Regional Data Management. Taylor & Francis Group, London, pp 385-398

Esch T, Marconcini M, Felbier A, Roth A, Heldens W, Huber M, Schwinger M, Taubenböck H, Müller A, Dech S (2013) Urban Footprint Processor – Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. IEEE Geoscience and Remote Sensing Letters, 10(6):1617-1621

Esch T, Thiel M, Bock M, Roth A, Dech S (2008) Improvement of Image Segmentation Accuracy Based on Multiscale Optimization Procedure. IEEE Geoscience and Remote Sensing Letters 5(3):463-467

Esch T, Thiel M, Schenk A, Roth A, Müller A, Dech S (2010) Delineation of urban footprints from TerraSAR-X data by analyzing speckle characteristics and intensity information. IEEE Transactions on Geoscience and Remote Sensing 48(2):905-916

Espindola GM, Camara G, Reis IA, Bins LS, Monteiro AM (2006) Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. International Journal of Remote Sensing, 27(14):3035-3040

F

Fayyad UM, Irani KB (1993) Multi-interval discretization of continuous-valued attributes for classification learning, in IJCAI, pp. 1022–1029

Fekete, A. (2011). Spatial disaster vulnerability and risk assessments: challenges in their quality and acceptance. Natural Hazards doi:10.1007/s11069-011-9973-7

Fekete A, Hufschmidt G (2012) Atlas of Vulnerability and Resilience Research: Case studies for Germany, Austria and Schwitzerland, to be published.

FEMA (2002). Rapid Visual Screening of Buildings for Potential Seismic Hazards. A Handbook, FEMA 154, 2nd edition

FEMA (2010) HAZUS - MH MR5. Multi-hazard loss estimation methodology - earthquake model. Technical manual. Department of Homeland Security, Emergency Prepardness and Response Directorate. Washington, D.C.

Fielding EJ, Wright TJ, Muller J, Parsons BE, Walker R (2004) Aseismic deformation of a fold-and-thrust belt imaged by synthetic aperture radar interferometry near Shahdad, southeast Iran. Geology 32(7):577-580

Foody GM (2002) Status of land cover classification accuracy assessment. Remote Sensing of Environment 80: 185-201

Foody GM (2004) Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogrammetric Engineering and Remote Sensing 70: 627–633

Fornaro G, Pauciullo A, Serafino F (2009) Deformation monitoring over large areas with multipass differential SAR interferometry: a new approach based on the use of spatial differences. International Journal of Remote Sensing 30:1455-1478

Forsyth T (2004) Critical political ecology: The politics of environmental science. Routledge, London

Freeman AS (2004) Review of the development of the capacity spectrum method. ISET Journal of Earthquake Technology 41(1):1-13

French SP, Muthukumar S (2006) Advanced technologies for earthquake risk inventories. Journal of Earthquake Engineering 10(2):207-236

Fu B, Ninomiya Y, Lei X, Toda S, Awata Y (2004) Mapping active fault associated with the 2003 Mw 6.6 Bam (SE Iran) earthquake with ASTER 3D images. Remote Sensing of Environment 92:153-157

Fuchs S (2009) Susceptability versus resilience to mountain hazards in Austria - Paradigms of vulnerability revisited. Natural Hazards and Earth System Sciences 9:337-352

Fuchs S, Kuhlicke C, Meyer V (2011) Editorial for the special issue: vulnerability to natural hazards - the challenge of integration. Natural Hazards 58:609-619

Füssel HM (2007) Vulnerability: A generally applicable conceptual framework for climate change research. Global Environmental Change-Human and Policy Dimensions 17:155-167

Füssel HM, Klein RJT (2006) Climate change vulnerability assessments: An evolution of conceptual thinking. Climatic Change 75:301-329

Füssel HM (2005) Vulnerability in climate change research: A comprehensive conceptual framework. University of California, International and Area Studies. Breslauer Symposium No. 6, Paper 6

G

Gamba P, Aldrighi M, Stasolla M (2011) Robust Extraction of Urban Area Extents in HR and VHR SAR Images. IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing 4(1):27-34

Gamba P, Dell'Acqua F, Trianni G (2007) Rapid damage detection in the Bam area using multitemporal SAR and exploiting ancillary data. IEEE Transactions on Geoscience and Remote Sensing 45(6):1582-1589

Gamba P, Houshmand, B (2000) Digital Surface Models and Building Extraction: A Comparison of IFSAR and LIDAR Data. IEEE Trans. Geosci. Remote Sens. 38(4):1959-1968

Garatwa W, Bollin C (2002) Disaster Risk Management: Working Concept. GTZ - Gesellschaft für technische Zusammenarbeit, Eschborn, Germany.

Ge Y, Xu J, Liu Q, Yao Y, Wang R (2009) Image interpretation and statistical analysis of vegetation damage caused by the Wenchuan earthquake and related secondary disasters. Journal of Applied Remote Sensing 3, 031660 (26 May 2009). doi: 10.1117/1.3141726

Geiß C, Aravena Pelizari P, Marconcini M, Sengara W, Edwards M, Lakes T, Taubenböck H (2014b) Estimation of seismic buildings structural types using multi-sensor remote sensing and machine learning techniques. ISPRS Journal of Photogrammetry and Remote Sensing, minor revisions

Geiß C, Taubenböck H (2013) Remote sensing contributing to assess earthquake risk: from a literature review towards a roadmap. Natural Hazards 68:7-48

Geiß C, Taubenböck H, Tyagunov S, Tisch A, Post J, Lakes T (2013) Assessment of seismic building vulnerability from space. Earthquake Spectra, in press, doi: 10.1193/121812EQS350M.

Geiß C, Wurm M, Breunig M, Felbier A, Taubenböck H () Normalization of TanDEM-X DSM data in urban environments with morphological filters. Submitted to IEEE Transactions on Geoscience and Remote Sensing

GEM (2011) Global Earthquake Model. Available online at http://www.globalquakemodel.org/ Accessed 21 Sep 2011

Genuer R, Poggi JM, Tuleau C (2008) Random forests: some methodological insights, Rapport de recherche RR-6729, INRIA.

GEO (2011) Group on earth observations. Disasters: Strategic Target. Available online at http://www.earthobservations.org/geoss_di_tar.shtml Accessed 17 Oct 2011

GeoNode (2012) GeoNode Online Platform. Available online at http://geonode.org/ Accessed 18 Jul 2012

GFDRR (2011) Global facility for disaster reduction and recovery. Available online at http://www.gfdrr.org/gfdrr/ Accessed 17 Oct 2011

GFDRR (2012) GFDRR Case Study: Central American Probabilistic Risk Assessment (CAPRA). Available online at http://www.gfdrr.org/docs/Snapshot_CAPRA.pdf Accessed 18 Jul 2012

GIO (2012) GMES Initial Operations (GIO). Available online at http://portal.ems-gmes.eu/web/guest;jsessionid=a5adb6c432f96a27da343aee7a81 Accessed 21 Sep 2011

Girgin S (2011) The natech events during the 17 August 1999 Kocaeli earthquake: aftermath and lessons learned. Natural Hazards and Earth System Sciences 11(4):1129-1140

Gislason PO, Benediktsson JA, Sveinsson JR (2006) Random Forests for land cover classification. Pattern Recognition Letters 27:294-300

Gitas IZ, Polychronaki A, Katagis T, Mallinis G (2008) Contribution of remote sensing to disaster management activities: A case study of the large fires in the Peloponnese, Greece. International Journal of Remote Sensing 29 (6):1847-1853

G-MOSAIC (2012) G-MOSAIC: GMES Pilot Services for Security. Available online at http://www.gmes-gmosaic.eu/ Accessed 12 Jul 2012

Gonzalez PJ, Chini M, Stramondo S, Fernandez J (2010) Coseismic horizontal offsets and fault-trace mapping using phase correlation of IRS satellite images: The 1999 Izmit (Turkey) earthquake. IEEE Transactions on Geoscience and Remote Sensing 48:2242-2250

Google (2012) Google Map Maker. Available online at http://www.google.com/mapmaker Accessed 12 Jul 2012

Gorny VI, Salman AG, Tronin AA, Shilin BV (1988) The earth's outgoing IR radiation as an indicator of seismic activity. Proc Acad Sci USSR 301:67-69

Graf ABA, Smola AJ, Borer S (2003) Classification in a Normalized Feature Space Using Support Vector Machines. IEEE Transactions on Neural Networks, 14:597-605

Granger K (2003) Quantifying storm tide risk in cairns. Natural Hazards 30:165-185

Gravelius H (1914) Grundriß der gesamten Gewässerkunde. Band I: Flußkunde (Compendium of Hydrology, vol. I. Rivers, in German), Goschen, Berlin, Germany

Graesser J, Cheriyadat A, Vatsavai RR, Chandola V, Long, J, Bright, E (2012) Image Based Characterization of Formal and Informal Neighborhoods in an Urban Landscape. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(4):1164-1176

Gravely D (2001) Risk, hazard and disaster. University of Canterbury, New Zealand.

Gruber A, Wessel B, Huber M, Roth A (2012) Operational TanDEM-X DEM calibration and first validation results," ISPRS J. Photogramm. Remote Sen. 73:39-49

Green C (2004) The evaluation of vulnerability to flooding. Disaster Prevention and Management 13(4):323-329

Griffith P, Hostert P, Gruebner O, van der Linden S (2010) Mapping megacity growth with multi-sensor data Remote Sensing of Environment 114:426-439

Grünthal G, Musson R, Schwarz J, Stucchi M (1998) European Macroseismic Scale. Cahiers de Centre Européen de Géodynamique et de Seismologie, vol 15, Luxembourg

Guinau M, Pallas R, Vilaplana JM (2005) A feasible methodology for landslide susceptibility assessment in developing countries: A case-study of NW Nicaragua after Hurricane Mitch. Engineering Geology 80(3-4):316-327

Guo H (2009) Guest editorial: Remote sensing of the Wenchuan earthquake. Journal of Applied Remote Sensing 3:031699 (26 May 2009). doi: 10.1117/1.3153899

Guo H (2010) Understanding global natural disasters and the role of earth observation. International Journal of Digital Earth 3(3):221-230

Guo HD, Ma JW, Zhang B, Li Z, Huang J, Zhu LW (2010) Damage consequence chain mapping after the Wenchuan earthquake using remotely sensed data. Int J Remote Sens 31(13):3427-3433

Gup G, Xie G (2007) Earthquake cloud over Japan detected by satellite. Int J Remote Sens 28(23):5375-5376

Guyon I (2003) An introduction to variable and feature selection. Journal of Machine Learning Research 3:1157–1182

H

Hahn H (2003) Indicators and other instruments for local risk management for communities and local governments. Document prepared as part of the documents related to the Project Local Risk Management for Communities and Local Governments. The German Technical Cooperation Agency (GTZ) for IADB.

Hall MA (1999) Correlation-based Feature Selection for Machine Learning, Ph.D. dissertation, Department of Computer Science, The University of Waikato, Hamilton, New Zealand

Hall M, Eibe F, Holmes G, Pfahringer B, Reutemann P, Witten I (2009) The WEKA data mining software: An update. SIGKDD Explorations 11(1):10–18

Hall M, Holmes G (2003) Benchmarking attribute selection techniques for discrete class data mining, IEEE Transactions on Knowledge and Data Engineering 15:1–16

Han J, Kamber M (2006) Data mining: concepts and techniques. Elsevier, Morgan Kaufmann, 2nd edition

Han S-C, Shum CK, Bevis M, Ji C, Kuo C-Y (2006) Crustal dilatation observed by GRACE after the 2004 Sumatra-Andaman earthquake. Science 313:658-662

Han Y, Liu H, Cui P, Su F, Du D (2009) Hazard assessment on secondary mountain-hazards triggered by the Wenchuan earthquake. Journal of Applied Remote Sensing 3: 031645 (26 May 2009). doi:031645-031645-031616

Haralick R, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics SMC-3:610–621

Haralick RM, Shapiro LG (1991) Computer and Robot Vision. Reading, MA: Addison-Wesley Longman Publishing Co.

Haralick R M, Sternberg S R, Zhuang X (1987) Image Analysis Using Mathematical Morphology. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(4):532-550

He H, Garcia E (2009) Learning from imbalanced data. IEEE T. Knowl. Data En. 21(9):1263–1284

Heiden U, Heldens W, Roessner S, Segl K, Esch T, Mueller A (2012) Urban structure type characterization using hyperspectral remote sensing and height information. Landscape and Urban Planning, 105:361-375

Heipke C (2010) Crowdsourcing geospatial data. ISPRS Journal of Photogrammetry and Remote Sensing 65:550-557

Heldens W, Esch T, Heiden U (2012) Supporting urban micro climate modelling with airborne hyperspectral data. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, 23-27 July, Munich, Germany, pp. 1598-1601

Heldens W, Heiden U, Esch T, Stein E, Müller A (2011) Can the Future EnMAP Mission Contribute to Urban Applications? A Literature Survey. Remote Sens 3(9):1817-1846

Heltberg R, Siegel PB, Jorgensen SL (2009) Addressing human vulnerability to climate change: Toward a 'no-regrets' approach. Global Environmental Change-Human and Policy Dimensions 19(1):89-99

Hengl T (2006) Finding the right pixel size. Computers & Geosciences 32:1283-1298

Herold M, Liu X, Clarke K (2003) Spatial metrics and image texture for mapping urban land use, Photogrammetric Engineering and Remote Sensing 69:991-1001

Hilhorst D, Bankoff G (2004) Mapping vulnerability. In: Bankoff G, Frerks G, Hilhorst D (eds) Mapping vulnerability. Disasters, Development & Peolpe. Eartscan, London, pp 1-24

Hoffmann J (2007) Mapping damage during the Bam (Iran) earthquake using interferometric coherence. Int J Remote Sens 28(6):1199-1216

Holzer TL, Savage JC (2013) Global Earthquake Fatalities and Population. Earthquake Spectra 29(1):155-175

Hong Y, Adler R, Huffmann G (2007) Use of satellite remote sensing data in the mapping of global landslide susceptibility. Natural Hazards 43:245-256

HOSMT (2012) Humanitarian Open Street Map Team. Available online at http://hot.openstreetmap.org/ Accessed 12 Jul 2012

Hsu CW, Chang CC, Lin CJ (2010) A practical guide to support vector classification, Technical report, Department of Computer Science, National Taiwan University.

Hsu CW, Lin CJ (2002) A comparison of methods for multi-class support vector machines. IEEE T. Neural Networ. 13(2):415–425

Huang RQ, Li WL (2009) Analysis of the geo-hazards triggered by the 12 May 2008 Wenchuan Earthquake, China Bulletin of Engineering Geology and the Environment 68(3):363-371

Huang X, Lu Q, Zhang L, Plaza A (2014) New Postprocessing Methods for Remote Sensing Image Classification: A Systematic Study. IEEE Trans. Geosci. Remote Sens. in press

Huang X, Wei C, Li H (2009) Remote sensing analysis of the distribution and genetic mechanisms of transportation network damage caused by the Wenchuan earthquake. Journal of Applied Remote Sensing 3: 031650 (26 May 2009). doi: 10.1117/1.3154387

Huber M, Wessel B, Kosmann D, Felbier A, Schwieger V, Habermeyer M, Wendleder A, Roth A (2009) Ensuring globally the TanDEM-X height accuracy: Analysis of the reference data sets ICESat, SRTM and KGPS-tracks. Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 769-772, Cape Town, South Africa, 2009

Hufschmidt G (2011) A comparative analysis of several vulnerability concepts. Natural Hazards 58:621-643

Hughes GF (1968) On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory 14(1):55–63

I

ICSMD (2011a) Earthquake in Japan. Available online at http://www.disasterscharter.org/web/charter/activation_details?p_r_p_1415474252_assetId=ACT-359 Accessed 25 Sep 2011

ICSMD (2011b) International Charter on Space and Major Disasters. Charter On Cooperation To Achieve The Coordinated Use Of Space Facilities In The Event Of Natural Or Technological Disasters Rev.3 (25/4/2000).2. Available online at http://www.disasterscharter.org/web/charter/charter Accessed 25 Sep 2011

ICSU (2011) International Council for Science. Disaster risk - integrated research on disaster risk. Available online at http://www.icsu.org/what-we-do/interdisciplinary-bodies/irdr Accessed 17 Oct 2011

IDRC (2011) International disaster risk conference. Available online http://www.idrc.info/ Accessed 17 Oct 2011

IFRC (1999) Vulnerability and capacity assessment: An international federation guide. Geneva, Switzerland

INSPIRE (2012) Infrastructure for Spatial Information in the European Community. Available online at http://inspire.jrc.ec.europa.eu/ Accessed 13 Jul 2012

Intermap (2010) Product handbook and quick start guide. Standard edition, v4.4, http://www.intermap.com/images/handbook/producthandbook.pdf (Accessed 7 Jul, 2012)

InSTEDD (2012) Innovative Support to Emergencies Diseases and Disasters. Available online at http://instedd.org/ Accessed 12 Jul 2012

Irish R (2008) Landsat 7 Science data user's handbook, Landsat Project Science Office, Issue 21, 186 pp

J

Jaiswal K, Wald D, D'Ayala D (2011) Developing Empirical Collapse Fragility Functions for Global Building Types. Earthquake Spectra 27(3):775-795

Jaiswal K, Wald D, Porter K (2010) A global building inventory for earthquake loss estimation and risk management. Earthquake Spectra 26:731-748

Janssen AM, Ostrom E (2006) Resilience, vulnerability, and adaption: A cross-cutting theme of the International Human Dimensions Programme on Global Environmental Change (editorial). Global Environmental Change 16:237-239

Japkowicz N, Stephen S (2002) The class imbalance problem: A systematic study. Intell. Data Anal. 6(5):429-449

JAXA (2011) Advanced Land Observing Satellite-2. Available online at http://www.jaxa.jp/projects/sat/alos2/index_e.html Accessed 7 Oct 2011

Johnson B, Xie Z (2013) Classifying a high resolution image of an urban area using super-object information. ISPRS Journal of Photogrammetry and Remote Sensing 83:40-49

Joyce K, Belliss S, Samsonov S, McNeill S, Glassey P (2009a) A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. Progress in Physical Geography 33(2):183-207

Joyce K, Wright K, Samsonov S (2009b) Remote sensing and the disaster management cycle. In: Jedlovec G (ed) Advances in Geoscience and Remote Sensing vol. 48, 7, INTECH, pp 317-346

K

Kakinami Y, Liu JY, Tsai LC, Oyama KI (2010) Ionospheric electron content anomalies detected by a FORMOSAT-3/COSMIC empirical model before and after the Wenchuan Earthquake. International Journal of Remote Sensing 31:3571-3578

Kasperson J, Kasperson R, Turner BL, Hsieh W, Schiller A (2005) Vulnerability to Global Environmental Change. In: Kasperson J, Kasperson R (eds) The social contours of risk. Volume II: Risk Analysis, Corporations & the Globalization of Risk. Earthscan, London, pp 245-285

Katti RV, Thyagarajan K, Shankara NK, Kiran Kumar SA (2007) Spacecraft technology, Special section: Indian space programme. Current Science 93(12):1715-1736

Kaya S, Curran PJ, Llewellyn G (2005) Post-earthquake building collapse: a comparison of government statistics and estimates derived from SPOT HRVIR data. Int J Remote Sens 26(13):2731-2740

Kaya S, Muftuoglu O, Tuysuz O (2004) Tracing the geometry of an active fault using remote sensing and digital elevation model: Ganos segment, North Anatolian Fault zone, Turkey. Int J Remote Sens 25(19):3843-3855

Kerle N (2010) Satellite-based damage mapping following the 2006 Indonesia earthquake--How accurate was it? International Journal of Applied Earth Observation and Geoinformation 12(6):466-476

Kilian J, Haala N, Englich M (1996) Capture and elevation of airborne laser scanner data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 31(B3):383-388

Kleinbaum DG, Kupper LL, Mueller KE, Nizam A (1998) Applied regression analysis and other multivariable methods, 3rd ed., Duxbury Press, Pacific Grove, CA, 736pp

Kohavi R, John GH (1997) Wrappers for feature subset selection. Artif. Intell. 97(1): 273–324

Kononenko I (1994) Estimating attributes: Analysis and extension of RELIEF, in Proceedings of the European Conference on Machine Learning, Springer, Catania, Italy, pp. 171–182

Korup O (2010) COGEAR - Coupled seismogenic geohazards in alpine regions. Module 1: Earthquake-triggered landslides - spatial patterns and impacts. Swiss Federal Research Institutes WSL/SLF, Davos, Switzerland.

Kosugi Y, Sakamoto M, Fukunishi M, Lu W, Doihara T, Kakumoto S (2004) Urban Change Detection Related to Earthquakes Using an Adaptive Nonlinear Mapping of High-Resolution Images. IEEE Geosci Remote Sens Lett 1(3):152-156

Krauss T, Arefi H, Reinartz P (2011) Evaluation of selected methods for extracting digital terrain models from satellite borne digital surface models in urban areas. International Conference on Sensors and Models in Photogrammetry and Remote Sensing, SMPR, 2011

Kraus K, Pfeiffer N (1998) Determination of terrain models in wooded areas with airborne laser scanner data," ISPRS J. Photogramm. 53:193-203

Krawczyk B, Wozniak M, Cyganek B (2014) Clustering-based ensembles for one-class classification. Information Sciences, 264:182-195

Krieger G, Moreira A, Fiedler H, Hajnsek I, Werner M, Younis M, Zink M (2007) TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. IEEE Transactions on Geoscience and Remote Sensing, 45(11):3317-3341

Kumar KV, Martha TR, Roy PS (2006) Mapping damage in the Jammu and Kashmir caused by 8 October 2005 Mw 7.3 earthquake from the Cartosat-1 and Resourcesat-1 imagery. Int J Remote Sens 27(20):4449-4459

Kuzuoka S, Mizuno T (2004) Land deformation monitoring using PSInSAR technique. In: Proceedings of International Symposium on Monitoring, Prediction and Mitigation of Disasters by Satellite Remote Sensing, Awaji, Hyogo, Japan, January 2004, pp 176-181.

L

Lafond G, Gosselin A (1994) A survey on perceived risks. In: Martin LRG (ed) Risk assessment and management: emergency planning perspectives. pp 44-59

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159-174

Lang S, Tiede D (2010) Rapid geospatial reporting. Erfahrungen mit automatisierter Schadensanalyse nach den Erdbeben in Haiti und Chile. GIS Business 2:42-43

Lee JS (1983) Digital image smoothing and the sigma filter. Comput. Vision Graph. 24(2):255-269

Leinenkugel P, Esch T, Künzer C (2011) Settlement detection and impervious surface estimation in the Mekong Delta using optical and SAR remote sensing data. Remote Sensing of Environment, 115:3007–3019

Leprince S, Ayoub F, Klinger Y, Avouac JP (2007) Co-registration of optically sensed images and correlation (COSI-Corr): an operational methodology for ground deformation measurements. IEEE International Geoscience and Remote Sensing Symposium, July 2007, Barcelona, Spain, pp 1943-1946

Lewis DD (1998) Naïve Bayes at forty: The independence assumption in information retrieval, Proceedings of ECML

Lewis J (1999) Development in disaster-prone places: Studies of vulnerability. Intermediate Technology Publications. Intermediate Technology Publications, London

Li M, Cheng L, Gong J, Liu Y, Chen Z, Li F, Chen G, Chen D, Song X (2008) Post-earthquake assessment of building damage degree using LiDAR data and imagery. Science in China Series E: Technological Sciences 51:133-143

Li P, Xu H, Guo J (2010) Urban building damage detection from very high resolution imagery using OCSVM and spatial features. International Journal of Remote Sensing 31:3393-3409

Li X, Tennant K, Lawrence G (2004) Three-dimensional mapping with airborne IFSAR based STAR technology – Intermap's experiences, International Archives Photogrammetry and Remote Sensing Vol. XXXV-B3, pp. 261–266

Liu X (2008) Airborne LiDAR for DEM generation: some critical issues. Prog. Phys. Geog. 32:31-49

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News: The Newsletter of the R Project 2(3):18–22

LinkER (2012) linkER -Supporting the implementation of operational GMES services in Emergency Response. Available online at http://www.zki.dlr.de/project/1394 Accessed 8 May 2012

Liou YA, Kar SK, Chang L (2010) Use of high-resolution FORMOSAT-2 satellite images for post-earthquake disaster assessment: a study following the 12 May 2008 Wenchuan Earthquake. International Journal of Remote Sensing 31:3355-3368

Liu JG, Haselwimmer, CE (2006) Co-seismic ruptures found up to 60 km south of the Kunlun fault after 14 November 2001, Ms 8.1, Kokoxili earthquake using Landsat-7 ETM+ imagery. International Journal of Remote Sensing 27:4461-4470

Liu H, Li J, Wong L (2002) A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns, Genome Informatics 13:51-60

Liu Y, Schumann M (2005) Data mining feature selection for credit scoring models. Journal of the Operational Research Society:1-10

López-Dekker P, Prats P, De Zan F, Schulze D, Krieger G, Moreira A (2011) TanDEM-X first DEM Acquisition: A crossing orbit experiment. IEEE Geoscience and Remote Sensing Letters 8(5):943-947

Lu P, Stumpf A, Kerle N, Casagli N (2011) Object-oriented Change Detection for Landslide Rapid Mapping. IEEE Geoscience and Remote Sensing Letters, 8(4):701-705

Lundgren P, Casu F, Manzo M, Pepe A, Berardino P, Sansosti E, Lanari R (2004) Gravity and magma induced spreading of Mount Etna volcano revealed by satellite radar interferometry. Geophysical Research Letters 31, L04602. doi: 10.1029/2003GL018736

M

MacEachren AM (1985) Compactness of geographic shape: comparison and evaluation of measures. Geogr. Ann. B. 67(1):53-67

Maeda T, Takano T (2010) Detection Algorithm of Earthquake-Related Rock Failures From Satellite-Borne Microwave Radiometer Data. IEEE Transactions on Geoscience and Remote Sensing 48(4):1768-1776

Maguya AS, Junttila V, Kauranne T (2013) Adaptive algorithm for large scale dtm interpolation from lidar data for forestry applications in steep forested terrain. ISPRS J. Photogramm. 85:74-83

Maktav D, Erbek FS, Jürgens C (2005) Remote sensing of urban areas. International Journal of Remote Sensing 26: 655–659

Marconcini M, Fernandez-Prieto D, Buchholz T (2014) Targeted Land-Cover Classification. IEEE Transactions on Geoscience and Remote Sensing 52(7):4173–4193

Martha TR, Kerle N, van Westen CJ, Jetten V, Kumar KV (2011) Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis. IEEE Transactions on Geoscience and Remote Sensing, 49(12):4928-4943

Mayer S (2000) Extraction of tree groups from high-resolution digital surface models. International Conference on Image Processing 3: 712-715, Vancouver, BC, Canada, 10-13 Sep. 2000

McCaffrey R (2009) The Tectonic Framework of the Sumatran Subduction Zone. Annu. Rev. Earth Pl. Sc. 37: 345–366

McCloskey J, Lange D, Tilmann F, Nalbant SS, Bell AF, Natawidjaja DH, Rietbrock A (2010) The September 2009 Padang earthquake. Nature Geoscience 3:70-71

Marsan D, Lengline O (2008) Extending earthquakes' reach through cascading. Science 319 (5866):1076-1079

Massonnet D (1995) Application of remote sensing data in earthquake monitoring. Advances in Space Research 15(11):37-44

Massonnet D, Feigl K (1998) Radar interferometry and its application to changes on the earth surface. Reviews of Geophysics 36:441-500

Massonnet D, Rossi M, Carmona C, Adragna F, Peltzer G, Feigl K, Rabaute T (1993) The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364:138-142

Masure P (2003) Variables and indicators of vulnerability and disaster risk for land-use and urban or territorial planning. Information and indicators program for disaster risk management, IADB - ECLAC - IDEA.

Matsuoka M, Yamazaki F (2010)Comparative analysis for detecting areas with building damage from several destructive earthquakes using satellite synthetic aperture radar images. Journal of Applied Remote Sensing 4:041867 (18 Nov 2010), doi: 10.1117/1.3525581

McCloskey J, Nalbant SS, Steacy, S (2005) Indonesian earthquake: Earthquake risk from co-seismic stress. Nature 434:291-291

Melgani F, Bruzzone L (2004) Classification of hyperspectral remote sensing images with support vector machines. IEEE Trans. Geosci. Remote Sens. 42(8):1778–1790

Meng X, Currit N, Zhao K (2010) Ground filtering algorithms for airborne LiDAR data: a review of critical issues. Remote Sens. 2(3):833-860

Menoni S, Pergalani F, Boni MP, Petrini V (2002) Lifelines earthquake vulnerability assessment: a systemic approach. Soil Dynamics and Earthquake Engineering 22(9-12):1199-1208

Mercer J, Dominey-Howes D, Kelman I, Lloyd K (2007) The potential for combining indigenous and western knowledge in reducing vulnerability to environmental hazards in small island developing states. Environmental Hazards 7(4):245-256

Minet C, Eineder M, Bamler R, Hajnsek I, Friedrich A (2008) Requirements for an L-band SAR-Mission for global monitoring of tectonic activities. USEReST '08, 11-14 November 2008, Naples, Italy

Mitomi H, Matsuoka M, Yamazaki F (2002) Application of automated damage detection of buildings due to earthquakes by panchromatic television Images. Proceedings of the 7th U.S. National Conference on Earthquake Engineering, Boston.

Montgomery DC, Peck EA, Vining GG (2001) Introduction to Linear Regression Analysis, 3rd edition, Wiley, New York, NY, 672 pp

Möller M (2011) Systemvergleich hochauflösender optischer Satellitenfernerkundungssensoren. In: Strobl J, Blaschke T, Griesebner G (eds) Angewandte Geoinformatik 2011, Beiträge zum 23. AGIT-Symposium Salzburg, 2011. Wichmann Verlag, Heidelberg.

Moreira A, Krieger G, Younis M, Hajnsek I, Papathanassiou K, Eineder M, De Zan F (2011) Tandem-L: A mission proposal for monitoring dynamic earth processes. IEEE International Geoscience and Remote Sensing Symposium, 25-29 July 2011, Vancouver, Canada

Motagh M, Schurr B, Anderssohn J, Cailleau B, Walter T R, Wang R, Villotte J P, (2010) Subduction earthquake deformation associated with 14 November 2007, Mw 7.8 Tocopilla earthquake in Chile: Results from InSAR and aftershocks. Tectonophysics 490(1-2):66-68

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66:247–259

Mück M, Taubenböck H, Post J, Wegscheider S, Strunz G, Sumaryono S, Ismail FA (2013) Assessing building vulnerability to earthquake and tsunami hazard using remotely sensed data. In press

Mueller M, Segl K, Heiden U, Kaufmann H (2006) Potential of high-resolution satellite data in the context of vulnerability of buildings. Natural Hazards 38:247-258

Müller A, Reiter J, Weiland U (2011) Assessment of urban vulnerability towards floods using an indicator-based approach – a case study for Santiago de Chile. Natural Hazards and Earth System Sciences 11:2107–2123

Munich Re Group (2011). Topics geo, natural catastrophes 2008 – Analyses, assessments, positions. Available online at http://www.amre.com/pdf/topicsgeo_2008.pdf Accessed 19 Sep 2011

Muñoz-Marí J, Bovolo F, Gomez-Chova L, Bruzzone L, Camp-Valls G (2010) Semisupervised one-class support vector machines for classification of remote sensing data. IEEE Trans. Geosci. Remote Sens. 48(8):3188–3197

Musci M, Queiroz Feitosa R, Costa GAOP, Fernandes Velloso ML (2013) Assessment of Binary Coding Techniques for Texture Characterization in Remote Sensing Imagery. IEEE Geoscience and Remote Sensing Letters, 10(6):1607-1611

N

Nassel M, Voigt S (2006) Vulnerability assessment of the built environment. In: Birkmann J, Fernando N, Hettige S, Amarasinghe S, Jayasingham T, Paranagama D, Nandana MDA, Nassel M, Voigt S, Grote U, Engel S, Schraven B, Wolferts J (eds) Rapid and multidimensional vulnerability assessment in Sri Lanka at the local level, UNU-EHS. Colombo, Bonn, pp 10-22.

Neer JT (1999) High resolution imaging from space – A commercial perspective on a changing landscape. International Archives of Photogrammetry and Remote Sensing XXXII (7C2):132-143

Nguyen GH, Bouzerdoum A, Phung A (2008) A supervised learning approach for imbalanced data sets. Pattern Recognition:1-4

Niebergall S, Loew A, Mauser W (2008) Integrative Assessment of Informal Settlements Using VHR Remote Sensing Data – The Delhi Case Study. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3(1):193-205

Novack T, Esch T, Kux H, Stilla U (2011) Machine learning comparison between WorldView-2 and QuickBird-2-Simulated imagery regarding object-based urban land cover classification. Remote Sensing 3:2263–2282

O

O'Brien K, Leichenko R, Kelkar U, Venema H, Aandahl G, Tompkins H, Javed A, Bhadwal S, Barg S, Nygaard L, West J (2004) Mapping vulnerability to multiple stressors: climate change and globalization in India. Global Environmental Change-Human and Policy Dimensions 14(4):303-313

OGC (2012) Open Geospatial Consortium. Available online at http://www.opengeospatial.org/ Accessed 8 May 2012

Ojala T, Pietikainen M, Topi Mäenpää T (2002) Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7):971-987

O'Keefe P, Westgate K, Wisner B (1976) Taking the naturalness out of natural disasters. Nature 260:566-567

Ouzounov D, Bryant N, Logan T, Pulinets S, Taylor S (2006) Satellite thermal IR phenomena associated with some of the major earthquakes in 1999-2003. Physics and Chemistry of the Earth, Parts A/B/C 31(4-9):154-163

Ouzounov D, Freund F (2004) Mid-infrared emission prior to strong earthquakes analyzed by remote sensing data. Advances in Space Research 33(3):268-273

Ouzounov D, Liu D, Kang C, Cervone G, Kafatos M, Taylor P (2007) Outgoing long wave radiation variability from IR satellite data prior to major earthquakes. Tectonophysics 431:211-220

Ozawa S, Murakami M, Fujiwara S, Tobita M (1997) Synthetic aperture radar interferogram of the 1995 Kobe earthquake and its geodetic inversion. Geophysical Research Letters 24(18):2327-2330

P

Pacifici F, Chini M, Emery W (2009) A neural network approach using multi-scale textural metrics from very high resolution panchromatic imagery for urban land-use classification. Remote Sensing of Environment 113(6):1276-1292

Pal R, Pal K (1993) A review on image segmentation techniques. Pattern Recognition, 26 (9):1277-1294

Panza GF, Irikura K, Kouteva M, Peresan A, Wang Z, Saragoni R (2011) Advanced Seismic Hazard Assessment. Pure and Applied Geophysics 168:1-9

Parsons T (2004) Recalculated probability of M ≥7 earthquakes beneath the Sea of Marmara, Turkey. Journal of Geophysical Research 109, B05304

Pasolli E, Melgani F, Tuia D, Pacifici F, Emery WJ (2014) SVM Active Learning Approach for Image Classification Using Spatial Information. IEEE Transactions on Geoscience and Remote Sensing, 52(4): 2217-2233

Pauleit S, Duhme F (2000) Assessing the environmental performance of land cover types for urban planning. Landscape and Urban Planning, 52:1-20

Paylor II ED, Evans DL, Tralli DM (2005) Remote sensing and geospatial information for natural hazards characterization. ISPRS Journal of Photogrammetry & Remote Sensing 59:181-184

Peduzzi P (2006) The disaster risk index: Overview over a quantitative approach In: Birkmann J (ed) Measuring vulnerability to natural hazards - Towards disaster resilient societies. United Nations University, New York, pp 171-181

Peduzzi P, Dao H, Herold C (2005) Mapping disastrous natural hazards using global datasets. Natural Hazards 35:265-289

Peduzzi P, Dao H, Herold C, Mouton F (2009) Assessing global exposure and vulnerability towards natural hazards: the Disaster risk index. Natural Hazards and Earth System Sciences 9(4):1149-1159

Pelling M (2003) The vulnerability of cities - Natural disasters and social resilience. Earthscan Publications Ltd., London

PERILS (2012) PERILS AG. Available online at http://www.perils.org/ Accessed 17 Jul 2012

Pesaresi M, Ehrlich D, Caravaggi I, Kauffmann M, Louvrier C (2011) Toward Global Autmatic Built-Up Area Recognition Using Optical VHR Imagery. IEEE Journal of Selected Topics in Applied Earth Observatins and Remote Sensing 4(4):923-934

Pesaresi M, Gerhardinger A (2011) Improved textural built-up presence for automatic recognition of human settlements in arid regions with scattered vegetation. IEEE Journal of selected topics in applied earth observation and remote sensing 4(1):16-26

Pesaresi M, Gerhardinger A, Haag F (2007) Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. Int J Remote Sens 28(13-14):3013-3036

Petersen MD, Dewey J, Hartzell S, Mueller C, Harmsen S, Frankel AD, Rukstales K (2004) Probabilistic seismic hazard analysis for Sumatra, Indonesia and across the Southern Malaysian Peninsula. Tectonophysics 390(1-4):141-158

Peterson BL, Harrell FEJ (1990) Partial proportional odds models for ordinal response variables. Applied Statistics 39(2):205-217

Philip G (2010) Remote sensing data analysis for mapping active faults in the northwestern part of Kangara Valley, NW Himalaya, India. Int J Remote Sens 28(21):4745-4761

Picozzi M, Bindi D, Pittore M, Kieling K, Parolai S (2013) Real-time risk assessment in seismic early warning and rapid response: a feasibility study in Bishkek (Kyrgyzstan). J Seismol 17:485-505

Pinar A, Honkura, Y., Kuge K (2001) Seismic activity triggered by the 1999 Izmit earthquake and its implications for the assessment of future seismic risk. Geophysical Journal International 146(1):F1-F7

Pingel TJ, Clarke KC, McBride WA (2013) An improved simple morphological filter for the terrain classification of airborne LIDAR data. ISPRS J. Photogramm. 77:21-30

Pittore M, Wieland M (2012) Towards a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. Nat. Hazards, doi: 10.1007/s11069-012-0475-z

Plaza AJ, Chang C-I (2008) High Performance Computing in Remote Sensing. Chapman & Hall/CRC: Boca Raton, Fl., 466 pp.

Polli D, Dell'Acqua F (2011) Fusion of optical and SAR data for seismic vulnerability mapping of buildings In: Prasad S, Bruce LM, Chanussot J (eds) Optical remote sensing. Advances in signal processing and exploitation techniques. Springer, Heidelberg, pp 329-341

Prasad JSR, Singh Y, Kaynia AM, Lindholm C (2009) Socioeconomic clustering in seismic risk assessment of urban housing stock. Earthquake Spectra 25:619-641

PreView 2012 PreView: Geo-information services for risk management on a European level. Available online at http://www.preview-risk.com/site/FO/scripts/myFO_accueil.php?flash=1&lang=EN Accessed 12 Jul 2012

R

R Core Team (2013) R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria

Ramm F, Topf J, Chilton S (2010) OpenStreetMap: Using and Enhancing the Free Map of the World. Uit Cambridge Ltd. 352 pp.

Rashed T, Jürgens C (2010) Remote sensing of urban and suburban areas. Remote Sensing and Digital Image Processing, Vol. 10, 1st Edition, XII, 352 pp

Rashed T, Weeks J (2003) Assessing vulnerability to earthquake hazards through spatial multicriteria analysis of urban areas. Int J Geogr Inf Sci 17(6):547-576

Rasmussen CE, Williams CKI (2006) Gaussian Processes for Machine Learning. New York: The MIT Press

Rathje EM, Adams BJ (2008) The role of remote sensing in earthquake science and engineering: Opportunities and challenges. Earthquake Spectra 24(2):471-492

Reale D, Nitti DO, Peduto D, Nutricato R, Bovenga F, Fornaro G (2011) Postseismic deformation monitoring with the COSMO/SKYMED constellation. IEEE Geoscience and Remote Sensing Letters 8:696-700

Reif D, Grasemann B, Faber RH (2011) Quantitative structural analysis using remote sensing data: Kurdistan, northeast Iraq. Aapg Bulletin 95:941-956

Remi M, Avouac JP (2002) Deformation due to the 17 August 1999 Izmit, Turkey, earthquake measured from SPOT images. Journal of geophysical research 107. doi:10.1029/2000JB000102

Richter R (1996) A spatially adaptive fast atmospheric correction algorithm, International Journal of Remote Sensing 17:1201-1214

RIMS (2012) The Risk Management Society. Available online at http://www.rims.org/Pages/Default.aspx Accessed 20 Jul 2012

RiskScape (2012) Easy-to-use multi-hazard impact and risk assessment tool. Available online at http://www.globalquakemodel.org/ Accessed 21 Sep 2011

Robinson D, Dhu T, Schneider J (2006) Practical probabilistic seismic risk analysis: A demonstration of capability. Seismological Research Letters 77(4):453-459

Robnik-Šikonja M, Kononenko I (2003) Theoretical and empirical Analysis of ReliefF and RReliefF. Machine Learning Journal 53:23-69

Roessner S, Wetzel H-U, Kaufmann H, Sarnagoev A (2005) Potential of satellite remote sensing and GIS for landslide hazard assessment in Southern Kyrgyzstan (Central Asia). Natural Hazards 35:395-416

Rosenau M, Nerlich R, Brune S, Oncken O (2010) Experimental insights into the scaling and variability of local tsunamis triggered by giant subduction megathrust earthquakes. Journal of Geophysical Research-Solid Earth 115. doi: B09314 10.1029/2009jb007100

Rossi C, Gonzalez FR, Fritz T, Yague-Martinez N, Eineder M (2012) TanDEM-X calibrated Raw DEM generation. ISPRS J. Photogramm. 73:12-20

Rouse JW, Haas RH, Schell JA, Deering DW (1974) Monitoring vegetation Systems in the Great Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, Greenbelt: NASA SP-351, 3010-317

S

SAFER (2011) Services and Applications For Emergency Response. Available online at http://safer.emergencyresponse.eu/site/FO/scripts/myFO_accueil.php?lang=EN Accessed 21 Sep 2011

Sahal A, Roger J, Allgeyer S, Lemaire B, Hebert H, Schindele F, Lavigne F (2009) The tsunami triggered by the 21 May 2003 Boumerdes-Zemmouri (Algeria) earthquake: field investigations on the French Mediterranean coast and tsunami modelling. Natural Hazards and Earth System Sciences 9(6):1823-1834

Sahar L, Muthukumar S, French P (2010) Using aerial imagery and GIS in automated building footprint extraction and shape recognition for earthquake risk assessment of urban inventories. IEEE Transactions on Geoscience and Remote Sensing 48(9):3511-3520

Saito K, Spence R (2004) Rapid damage mapping using post-earthquake satellite images. In: IEEE International Geoscience and Remote Sensing Symposium, August 2004, Anchorage, AK, USA, pp 2272 – 2275

Salcedo-Sanz S, Rojo-Alvarez JL, Martinez-Ramon M, Camps-Valls G (2014) Support vector machines in engineering: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(3): 234-267

Salvi S, Stramondo S, Funning GJ, Ferretti A, Sarti F, Mouratidis A (2012) The Sentinel-1 mission for the improvement of the scientific understanding and the operational monitoring of the seismic cycle. Remote Sensing of Environment 120:164-174

Samuels P, Gouldby B, Klijn F, Messner F, van Os A, Sayers P, Schanze J, Udale-Clarke H (2009) Language of risk—project definitions. Floodsite project report T32-04-01.

Sarabandi P, Kiremidjian A (2007) Development of algorithms or building inventory compilation through remote sensing and statistical inferencing. The John A. Blume Earthquake Engineering Center, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, USA, Report No. 158, 419p

Sarabandi P, Kiremidjian A, Eguchi R T, Adams BJ (2008) Building inventory compilation for disaster management: Application of remote sensing and statistical modeling. Technical Report Series MCEER-08-0025, MCEER: Buffalo.

Saradjian MR, Akhoondzadeh M (2011) Thermal anomalies detection before strong earthquakes (M >6.0) using interquartile, wavelet and Kalman filter methods. Natural Hazards and Earth System Sciences 11:1099-1108

Saraf AK (2000) IRS-1C-PAN Depicts Chamoli earthquake induced landslides in Garhwal Himalayas, India. Int J Remote Sens 21:2345-2352

Saraf AK, Choudhury S (2005) Thermal Remote Sensing Technique in the Study of Pre-Earthquake Thermal Anomalies. J Ind Geophys Union 9(3):197-207

Saraf AK, Rawat V, Choudhury S, Dasgupta S, Das J (2009) Advances in understanding of the mechanism for generation of earthquake thermal precursors detected by satellites. International Journal of Applied Earth Observation and Geoinformation 11(6):373-379

Sarewitz D, Pielke R, Keykah M (2003) Vulnerability and risk; some thoughts from a political and policy perspective risk analysis. Risk Ana 23(4):805-810

Schmitt A, Wessel B, Roth A (2010) Curvelet-based change detection on SAR images for natural disaster mapping. Photogrammetrie Fernerkundung Geoinformation 6:463-474

Schneiderbauer S, Ehrlich D (2004) Risk, hazard and people's vulnerability to natural hazards. A review of definitions, concepts and data. Joint Research Centre, European Commission, EUR 21410

Schneiderbauer S, Ehrlich D (2006) Social levels and hazard (in)dependence in determining vulnerability. In: Birkmann J (ed) Measuring vulnerability to natural hazards. United Nations University Press, New York, pp. 78-102

Schölkopf B, Smola A (2002) Learning with Kernels. MIT Press, Cambridge (MA)

Schölkopf B, Smola AJ (2004) A tutorial on support vector regression. Statistics and Computing 14:199-222

Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC (2001) Estimating the support of a high dimensional distribution. Neural Computation 13(7):1443-1471

Schölkopf B, Williamson RC, Smola A, Shawe-Taylor J (1999) Support vector method for novelty detection, Proc. Adv. Neural Inf. Process. Syst., Denver, CO, vol. 12, pp. 582–588

Sengara IW, Suarjana M, Beetham D, Corby N, Edwards M, Griffith M, Wehner M, Weller R (2010) The 30th September 2009 West Sumatra Earthquake. Padang Region Damage Survey. Geoscience Australia, Record 2010/44, 201pp.

Sertel E, Kaya S, Curran PJ (2007) Use of semivariograms to identify earthquake damage in an urban area. IEEE Transactions on Geoscience and remote sensing 45(6):1590-1594

SERTIT (2012) SErvice Régional de Traitement d'Image et de Télédétection. Available online at http://sertit.u-strasbg.fr/english/en_welcome.htm Accessed 11 Jul 2012

SERVIR (2012) NASA Missions: SERVIR. Available online at http://www.nasa.gov/mission_pages/servir/index.html Accessed 11 Jul 2012

Setiadi N, Taubenböck H, Raupp S, Birkmann J (2010) Integrating socio-economic data in spatial analysis: An exposure analysis method for planning urban risk mitigation. 15th International Conference on Urban Planning and Regional Development in the Information Society, Vienna, Austria. GeoMultimedia 2010, pp. 367-374.

Shafique M, van der Meijde M, Ullah S (2011) Regolith modeling and its relation to earthquake induced building damage: A remote sensing approach. Journal of Asian Earth Sciences 42(1-2):65-75

Shafique M, van der Meijde M, van der Werff HMA (2012) Evaluation of remote sensing-based seismic site characterization using earthquake damage data. Terra Nova 24(2):123-129

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 23rd ACM national conference (ACM ’68). ACM, New York, NY, USA, 27-29 August

Shevade SK, Keerthi SS, Bhattacharyya C, Murthy KRK (2000) Improvements to the SMO algorithm for SVM regression. IEEE Transactions on Neural Networks 11:1188-1193

Sinadinovski C, Edwards M, Corby N, Milne M, Dale K, Dhu T, Jones A, McPherson A, Jones T, Gray D, Robinson D, White J (2005) Earthquake risk. Natural hazard risk in Perth, WA, GeoCat No. 63527.

Singh RP (2010) Satellite observations of the Wenchuan Earthquake, 12 May 2008. Int J Remote Sens 31(10):3335-3339

Sirmacek B, Taubenböck H, Reinartz P, Ehlers M (2012) Performance evaluation for 3-D city model generation of six different DSMs from air- and spaceborne sensors. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5:59–70

Sitharam TG, Anbazhagan P, Ganesha Raj K (2006) Use of remote sensing and seismotectonic parameters for seismic hazard analysis of Bangalore. Natural Hazards and Earth System Sciences 6:927-939

Sithole G (2001) Filtering of laser altimetry data using a slope adaptive filter. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIV (3/W4): 203-210

Sithole G, Vosselman G (2004) Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS J. Photogramm. 59:85-101

Smit B, Wandel J (2006) Adaptation, adaptive capacity and vulnerability. Global Environmental Change 16(3):282-292

Smith K, Petley DN (2009) Environmental hazards - Assessing risk and reducing disaster. Routhledge

Smith W (2005) The challenge of earthquake risk. Seismological Research Letters 76(4):415-416

Smola AJ, Schölkopf B (1998) A tutorial on support vector regression, NeuroCOLT2 Technical Report Series - NC2-TR-1998-030

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression, Statistics and Computing 14:199-222

Stein RS, Barka AA, Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. Geophysical Journal International, 128:594-604

Steiniger S, Lange T, Burghardt D, Weibel R (2008) An approach for the classification of urban building structures based on discriminant analysis techniques. Transactions in GIS 12 (1):31-59

Stramondo S, Bignami C, Chini M, Pierdicca N, Tertulliani A (2006) Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies. Int J Remote Sens 27 (20):4433-4447

Stramondo S, Moro M, Tolomei C, Cinti FR, Doumaz F (2005) InSAR surface displacement field and fault modelling for the 2003 Bam earthquake (southeastern Iran). Journal of Geodynamics 40:347-353

Stramondo S, Saroli M, Tolomei C, Moro M, Doumaz F, Pesci A, Loddo F, Baldi P, Boschi, E (2007) Surface movements in Bologna (Po Plain-Italy) detected by multitemporal DInSAR. Remote Sensing of Environment 110:304-316

Strunz G, Post J, Zosseder K, Wegscheider S, Mück M, Riedlinger T, Mehl H, Dech S, Birkmann J, Gebert N, Harjono H, Anwar HZ, Sumaryono, Khomarudin RM, Muhari A (2011) Tsunami risk assessment in Indonesia. Natural Hazards and Earth System Sciences 11:67-82

Stumpf A, Lachiche N, Malet JP, Kerle N, Puissant A (2014) Active Learning in the Spatial Domain for Remote Sensing Image Classification. IEEE Transactions on Geoscience and Remote Sensing 52(5):2492 – 2507.

Stumpf A, Kerle N (2011) Object-oriented mapping of landslides using random forests. Remote Sens. Environ. 115(10):2564-2577

Suzuki S (2012) ALOS Program Status. Presentation held at meeting JAXA and DLR on 13 Jul 2012, Oberpfaffenhofen, Germany

Suzuki T, Miyoshi D, Meguro J, Amano Y, Hashizume T, Sato K, Takiguchi J (2008) Real-time hazard map generation using small unmanned aerial vehicle. Proceedings of SICE Annual Conference, Tokyo (Japan), 20-22 Aug 2008

T

Tahayt A, Feigl K, Mourabit T, Rigo A, Reilinger R, McClusky S, Fadil A, Berthier E, Dorbath L, Serroukh M, Gomez F, Ben Sari D (2009) The Al Hoceima (Morocco) earthquake of 24 February 2004, analysis and interpretation of data from ENVISAT ASAR and SPOT5 validated by ground-based observations. Remote Sensing of Environment 113:306-316

Takano T, Maeda T (2009) Experiment and theoretical study of earthquake detection capability by means of microwave passive sensors on a satellite. IEEE Geosci Remote Sens Lett 6:107-111

Tang A, Wen A (2009) An intelligent simulation system for earthquake disaster assessment. Computers and Geosciences 35:871-879

Taubenböck H (2011) The vulnerability of a city - diagnosis from a bird's eye view. Mörner N-A (ed) The Tsunami Threat - Research and Technology. InTech, Croatia, pp 107-128

Taubenböck H, Esch T, Felbier, A, Wiesner M, Roth A, Dech S (2012) Monitoring urbanization in mega cities from space. Remote Sensing of Environment 117:162-176

Taubenböck H, Geiß C (2012) Vulnerability and resilience research – a critical perspective. Atlas of Vulnerability and Resilience Research: Case studies for Germany, Austria and Schwitzerland, to be published.

Taubenböck H, Goseberg N, Lämmel G, Setiadi N, Schlurmann T, Nagel K, Siegert F, Birkmann J, Traub KP, Dech S, Keuck V, Lehmann F, Strunz G, Klüpfel H (2013) Risk reduction at the « Last-Mile » : an attempt to turn science into action by the example of Padang, Indonesia. Natural Hazards 65(1):915-945

Taubenböck H, Goseberg N, Setiadi N, Lämmel G, Moder F, Oczipka M, Klüpfel H, Wahl R, Schlurmann T, Strunz G, Birkmann J, Nagel K, Siegert F, Lehmann F, dech S, Gress A, Klein R (2009b) “Last-Mile” preparation for a potential disaster – Interdisciplinary approach towards tsunami early warning and an evacuation information system for the coastal city of Padang, Indonesia. Natural Hazards and Earth System Sciences 9:1509-1528

Taubenböck H, Klotz M, Wurm M, Schmieder J, Wagner B, Wooster M, Esch T, Dech S (2013) Delineation of Central Business Districts in mega city regions using remotely sensed data. Remote Sens. Environ 136:386-401

Taubenböck H, Post J, Roth A, Zosseder K, Strunz G, Dech S (2008) A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing. Natural Hazards and Earth System Sciences 8:409-420

Taubenböck H, Post J, Kiefl R, Roth A, Ismail FA, Strunz G, Dech S (2008a) Risk and vulnerability assessment to tsunami hazard using very high resolution satellite data, in Proceedings of the EARSeL Joint Workshop: Remote Sensing: New Challenges of high resolution, Carsten J. (editor), Bochum, Germany

Taubenböck H, Post J, Roth A, Strunz G, Kiefl R, Dech S, Ismail F (2008b) Multi-scale assessment of population distribution utilizing remotely sensed data – The case study Padang, West Sumatra, Indonesia, International Conference on Tsunami Warning (ICTW), Bali, Indonesia, 12-14 November 2008

Taubenböck H, Roth A, Dech S (2007) Linking structural urban characteristics derived from high resolution satellite data to population distribution. In: Coors V, Rumor M, Fendel E, Zlatanova S (eds) Urban and Regional Data Management. Taylor&Francis Group, London, pp 35-45

Taubenböck H, Roth A, Dech S, Mehl H, Münich JC, Stempniewski L, Zschau J (2009a) Assessing building vulnerability using synergistically remote sensing and civil engineering. In: Kreck A, Rumor M, Zlatanova S, Fendel E (eds) Urban and Regional Data Management. Taylor&Francis Group, London, pp 287-300

Taubenböck H, Wiesner M, Felbier A, Marconcini M, Esch T, Dech S (2014) New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data Appl. Geogr. 47:137-153

Taubenböck H, Wurm M, Setiadi N, Gebert N, Roth A, Strunz G, Birkmann J, Dech S (2009c) Integrating remote sensing and social science – The correlation of urban morphology with socioeconomic parameters, Urban Remote Sensing Joint Event, Shanghai, China

Teimouri M, Delavar MR, Kolyaie S, Chavoshi SH, Kiavarz Moghaddam H (2008) A SDSS-based earthquake damage assessment for emergency response: case study in Bam. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (B8), 451-456.

Theilen-Willige B (2010) Detection of local site conditions influencing earthquake shaking and secondary effects in Southwest-Haiti using remote sensing and GIS-methods. Natural Hazards and Earth System Sciences 10:1183-1196

Thywissen K (2006) Core terminology of disaster reduction: A compartive glossary. In: Birkmann J (ed) Measuring vulnerability to natural hazards. United Nations University Press, New York, pp 448 – 496

Tigges J, Lakes T, Hostert P (2013) Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. Remote Sensing of Environment 136:66-75

Timmermann P (1981) Vulnerability, resilience and the collapse of society. No. 1 in Environmental Monograph. Institute for Environmental Studies, University of Toronto, Canada

Tralli DM, Blom RG, Fielding EJ, Donnellan A, Evans DL (2007) Conceptual case for assimilating interferometric synthetic aperture radar data into the HAZUS-MH earthquake module. IEEE Transactions on geoscience and remote sensing 45:1595-1604

Tralli DM, Blom RG, Zlotnicki V, Donnellan A, Evans DE (2005) Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards. ISPRS Journal of Photogrammetry and Remote Sensing 59:185-198

Trianni G, Gamba P (2009) Fast damage mapping in case of earthquakes using multitemporal SAR data. Journal of Real-Time Image Processing 4:195-203

Trimble (2012) Trimble eCognition Developer 8.8 reference book. Munich, Germany: Trimble Germany GmbH

Tronin AA (2010) Satellite Remote Sensing in Seismology. A Review. Remote Sensing 2:124-150

Tronin AA (1996) Satellite thermal survey—a new tool for the study of seismoactive regions. Int J Remote Sens 17(8):1439-1455

Tronin AA (2006) Remote sensing and earthquakes: A review. Physics and Chemistry of the Earth 31:138-142

Tsai YB, Liu JY, Ma KF, Yen HY, Chen KS, Chen YI, Lee CP (2006) Precursory phenomena associated with the 1999 Chi-Chi earthquake in Taiwan as identified under iSTEP Program. Phys Chem Earth 31:365-377

Tucker BE (2013) Reducing earthquake risk. Science, 341(6150):1070-1072

Tucker C J (2004) NASA's global orthorectified Landsat data set. Photogrammetric Engineering and Remote Sensing 70:313-322

Tuia D, Copa L, Kanevski M, Munoz-Mari J (2011) A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. IEEE Journal of Selected Topics in Signal Processing 5(3):606-617

Tuia D, Ratle F, Pacifici F, Kanevski MF, Emery WJ (2009) Active Learning Methods for Remote Sensing Image Classification. IEEE Trans. Geosci. Remote Sens. 47 (7):2218-2232

Turker M, Cetinkaya B (2005) Automatic detection of earthquake-damaged buildings using DEMs created from pre- and post-earthquake stereo aerial photographs. Int J Remote Sens 26(4):823-832

Turker M, San BT (2004) Detection of collapsed buildings caused by the 1999 Izmit, Turkey earthquake through digital analysis of post-event aerial photographs. Int J Remote Sens 25(21):4701-4714

Turker M, Sumer E (2008) Building-based damage detection due to earthquake using the watershed segmentation of the post-event aerial images. Int J Remote Sens 29(11):3073-3089

TurkStat (2014) Turkish Statistical Institute. URL: http://rapor.tuik.gov.tr/reports/rwservlet?adnksdb2&ENVID=adnksdb2Env&report=wa_buyukbelediye.RDF&p_il1=34&p_kod=2&p_yil=2013&p_dil=1&desformat=html. (Accessed 10 July, 2014)

Tucker BE (2013) Reducing earthquake risk. Science 341(6150):1070-1072

Turner BL, Kasperson R, Matson P, McCarthy J, Corell R, Christensen L, Eckley N, Kasperson J, Luers A, Martello M, Polsky C, pulsipher A, Schiller A (2003) A framework for vulnerability analysis in sustainability science. Proceedings of the National Academy of Sciences of the United States of America 100(14):8074 – 8079

Tyc G, Tulip J, Schulten D, Krischke M, Oxfort M (2005) The RapidEye mission design. Acta Astronautica 56:213-219

U

UN (1994) Report of the world conference on natural disaster reduction. Held at Yokohama, Japan, 23–27 May 1994

UN/ISDR (2004) Living with risk: A global review of disaster reduction initiatives. United Nations/International Strategy for Disaster Reduction, Geneva, Switzerland, UN Publications

UN/ISDR (2006) NGOs & Disaster Risk Reduction: A Preliminary Review of Initiatives and Progress Made. Background Paper for a Consulative Meeting on a "Global Network of NGOs for Community Resilience to Disasters", Geneva, Switzerland, 25-26 October 2006

UN/ISDR (2011) United Nations International Strategy for Disaster Reduction. Available online at http://www.unisdr.org/ Accessed 17 Oct 2011

United Nations (2011) The 2010 revision. Population Division. Department of Economic and Social Affairs. United Nations Publications

United Nations (2012) The 2011 revision. Population Division. Department of Economic and Social Affairs. United Nations Publications

UNDP (2004) Reducing disaster risk - A challenge for development - A Global report, United Nations Development Programme, New York, USA

UNDP/ERRRP (2009) Earthquake vulnerability profile and preparedness plan of Hetauda municipality. In: Joshi A (ed) Katmandu: United Nations Development Programme/Earthquake Risk Reduction and Recovery Preparedness Programme for Nepal

UNDRO (1979) Natural Disasters and Vulnerability Analysis. Report of Expert Group Meeting, Geneva, Switzerland, 9-12 July 1979

UNOCHA (2012) United Nations Office for the Coordination of Humanitarian Affairs. Available online at http://www.unocha.org/ Accessed 13 Jul 2012

UNOOSA (2011) United Nations office for outer space affairs. Available online at http://www.oosa.unvienna.org/ Accessed 17 Oct 2011

UNOSAT (2011) Unitar's operational satellite applications programme. Available online at http://www.unitar.org/unosat/ Accessed 19 Sep 2011

UN SPIDER (2011) United Nations platform for space-based information for disaster management and emergency response. Available online at http://www.un-spider.org/ Accessed 17 Oct 2011

UR (2012) Understanding Risk - Innovation in Disaster Risk Assessment. Available online at http://www.understandrisk.org/ur/node/4889 Accessed 17 Oct 2011

USGS (2011) Earthquakes with 1,000 or More Deaths since 1900. Available online at http://earthquake.usgs.gov/earthquakes/world/world_deaths.php Accessed 21 Sep 2011

Ushahidi (2012) Ushahidi Platform for information collection, visualization and interactive mapping. Available online at http://ushahidi.com/index.php/products/ushahidi-platform Accessed 12 Jul 2012

V

van den Broek B, Kiefl R, Riedlinger T, Scholte K, Granica K, Gutjahr K, Stephenne N, Binet R, de la Cruz A (2009) Rapid mapping and damage assessment. In: Jasani B, Pesaresi M, Schneiderbauer S, Zeug G (eds) Remote sensing from space. Supporting international peace and security. Springer, New York, pp 261 - 286

van Dillen S (2004) Different choices: assessing vulnerability in South Indian village. Studien zur geographischen Entwicklungsforschung, Verlag für Entwicklungspolitik, vol 9, Saarbrücken, Germany

van Puymbroeck N, Michel R, Binet R, Avouac JP, Taboury J (2000) Measuring earthquakes from optical satellite images. Appl Opt Inf Process 39:1-14

Van Hulse J, Khoshgoftaar TM, Napolitano A, Wald R (2009) Feature selection with high-dimensional imbalanced data, in Proceedings of the 2009 IEEE International Conference on Data Mining Workshops, ICDMW '09, pp. 507–514, Washington, DC, USA. IEEE Computer Society.

Vapnik VN (1995) The nature of statistical learning theory, Springer, New York, 187 pp

Vapnik VN (1998) Statistical learning theory, Wiley Series on Adaptive and Learning Systems, Wiley, New York

VDV (2012) Virtual Disaster Viewer. Available online at http://vdv.mceer.buffalo.edu/vdv/select_event.php Accessed 21 Sep 2011

Verrelst J, Muñoz-Marí J, Alonso L, Delegido J, Rivera JP, Camps-Valls G, Moreno J (2012) Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. Remote Sensing of Environment 118:127-139

Verstappen HT (1995) Aerospace technology and natural disaster reduction. Advances in Space Research 15(11):3-15

Villagrán De Léon JC (2006) Vulnerability: A conceptual and methodological review. SOURCE - Studies of the University: Research, Counsel, Education; UNU Institute for Environment and Human Security (4)

Voigt S, Kemper T, Riedlinger T, Kiefl R, Scholte K, Mehl H (2007a) Satellite image analysis for disaster and crisis-management support. IEEE Transactions on Geoscience and Remote Sensing 45(6):1520-1528

Volpi M, Tuia D, Bovolo F, Kanevski M, Bruzzone L (2013) Supervised change detection in VHR images using contextual information and support vector machines. Int. J. Appl. Earth Obs. 20:77–85

Vosselman G (2000) Slope based filtering of laser altimetry data. Int. Arch. Photogramm. Remote Sens. B4(33):958-964

Vrijling JK, van Hengel W, Houben RJ (1995) A framework for risk evaluation. Journal of Hazardous Materials 43:245-261

Vu TT, Ban Y (2010) Context-based mapping of damaged buildings from high-resolution optical satellite images. Int J Remote Sens 31(13):3411-3425

Vu TT, Matsuoka M, Yamazaki F (2004) LiDAR based change detection of buildings in dense urban areas. In: IEEE International Geoscience and Remote Sensing Symposium, September 2004, Anchorage, AK, USA, pp 3413-3416

Vu TT, Matsuoka M, Yamazaki F (2005) Detection and animation of damage using very high-resolution satellite data following the 2003 Bam, Iran, Earthquake. Earthquake Spectra 21:319-327

W

Wackernagel H (2003) Multivariate Geostatistics – An Introduction with Applications. 3rd Ed., Berlin, Germany: Springer-Verlag, 2003

Walker RT (2006) A remote sensing study of active folding and faulting in southern Kerman province, S.E. Iran. Journal of Structural Geology 28:654-668

Walton M, Bebbington AJ, Dani AA, deHaan A (2008) Institutional pathways to equity: addressing inequality traps. The World Bank, Washington D. C

Wang C, Zhang H, Wu F, Zhang B, Tang Y, Wu H, Wen X, Yan D (2009a). Disaster phenomena of Wenchuan earthquake in high resolution airborne synthetic aperture radar images. Journal of Applied Remote Sensing 3: 031690 (26 May 2009). doi: 10.1117/1.3154558

Wang FT, Zhou Y, Wang SX, Liu WL, Wei CJ, Han Y (2011) Investigation and assessment of damage in earthquake in Yushu, Qinghai based on multi-spectral remote sensing. Spectroscopy and spectral Analysis 31(4):1047-1051

Wang Z, Xu Q, Xu B, Zhang W (2009b) Emergency aero-photo survey after the 5.12 Wenchuan earthquake, China. Science in China Series E-Technological Sciences 52:835-843

WB/UN (2010) Natural hazards, unnatural disasters: the economics of effective prevention. World Bank/United Nations Washington D.C.

Weidner U, Förstner W (1995) Towards automatic building extraction from high resolution digital elevation models. ISPRS J. Photogramm. 50(4):38-49

Wendleder A, Wessel B, Roth A, Breunig M, Martin K, Wagenbrenner S (2012) TanDEM-X Water Indication Mask: Generation and First Evaluation Results. IEEE J. Sel. Topics Appl. Earth Observ. 6(1):171-179

Weng Q Quattrochi DA (2006) Urban remote sensing, CRC Press/Taylor and Francis, Boca Raton, FL, USA, 432 pp.

Wessel B, Fritz T, Busche T, Bräutigam B, Krieger G, Schättler B, Zink M (2013) TanDEM-X Ground Segment DEM Products Specification Document. Earth Observation Center, https://tandemx-science.dlr.de/, TD-GS-PS-0021, Issue 2.0, 22.04.2013

Weston J, Ferreira AMG, Funning GJ (2012) Systematic comparisons of earthquake source models determined using InSAR and seismic data. Tectonophysics 532-535:61-81

Weszka JS, Dyer CR, Rosenfeld A (1976) A comparative study of texture measures for terrain classification, IEEE Transactions on Systems, Manufacturing and Cybernetics SMC-6:2269–2285

White G, Kates R, Burton I (2001) Knowing better and losing even more: the use of knowledge in hazards management. Environmental Hazards 3:81–92

White P, Pelling M, Sen K, Seddon D, Russell S, Few R (2005) Disaster risk reduction: a development concern. A scoping study on links between disaster risk reduction, poverty and development. Department for International Development

Whitman RV (1973) Damage probability matrices for prototype buildings. Structures Publication. Department of Civil Engineering, Massachusetts Institute of Technology, Boston, Massachusetts, USA

Wieland M, Pittore M, Parolai S, Zschau J, Moldobekov B, Begaliev U (2012) Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-source imaging. Soil Dynamics and Earthquake engineering 36:70-83

Williams R (2006) Generalized ordered logit/partial proportional odds models for ordinal dependent variables. The Stata Journal 6(1):58-82

Wisner B (2004) Assessment of capability and vulnerability. In: Bankoff G, Frerks G, Hilhorst D (eds) Mapping vulnerability. Disasters, Development & People. Earthscan, London, pp 183-205

Wisner B (2007) Regions at risk or people at risk? Wie natürlich sind "Naturkatastrophen"? Geographische Rundschau 59(10):256-276

Wisner B, Blaikie P, Cannon T, Davis I (2003) At risk: natural hazards, people's vulnerability and disasters. 2nd edn. Routledge, London & New York

Wolf S (2011) Vulnerability and risk: comparing assessment approaches. Natural Hazards. doi: 10.1007/s11069-011-9968-4

Wu SS, Qiu X, Usery EL, Wang L (2008) Using geometrical, textural, and contextual information of land parcels for classification of detailed urban land use. Annals of the Association of American Geographers 99:76-98

Wurm M, Taubenböck H, Schardt M, Esch T, Dech S (2011) Object-based image information fusion using multisensor earth observation data over urban areas. International Journal of Image and Data Fusion, 2(2): 121-147

Wurm M, Taubenböck H, Roth A, Dech S (2009) Urban structuring using multisensoral remote sensing data. Urban Remote Sensing Event, Shanghai, China

Wyss M (2004) Real-time prediction of earthquake casualties. In: Malzahn D, Plapp T (eds) Disasters and society – From hazard assessment to risk reduction. Logos Publishers, Conference proceedings, Karlsruhe, Germany, pp 165-173, 26-27 July 2004

Wyss M, Wang R, Zschau J, Xia Y (2006) Earthquake loss estimates in near-real time. EOS Transactions American Geophysical Union 87(44):477-492

Wyss M (2012) Mapping Seismic Risk: The current crisis. Taubenböck H, Post J, Strunz G (eds.) Remote Sensing contributing to mapping earthquake vulnerability and effects. Special Issue in Natural Hazards. To be published.

X

Xiong, B., Oude Elberink S, Vosselmann G (2014) A graph edit dictionary for correcting errors in roof topology graphs reconstructed from point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 93:227-242

Xu W, Dong R, Wang X, Ouyang Z, Li Z, Xiao Y, Zhang J (2009) Impact of China's May 12 earthquake on Giant Panda habitat in Wenchuan County. Journal of Applied Remote Sensing, 3: 031655 (26 May 2009). doi: 10.1117/1.3153916

Y

Yamazaki F, Kouch K (2006) Automated damage detection of buildings from high-resolution satellite images. Paper presented at the First European Conference on Earthquake Engineering and Seismology, September 2006, Geneva, Switzerland, pp 3-8

Yamazaki F, Matsuoka M (2007) Remote sensing technologies in post-disaster damage assessment. Journal of Earthquake and Tsunami 1(3):193-210

Yang Y, Guo G (2010) Studying the thermal anomaly before the Zhangbei earthquake with MTSAT and meteorological data. International Journal of Remote Sensing 31:2783-2791

Yano Y, Yamazaki F (2006) Building damage detection of the 2003 Bam, Iran earthquake using Quickbird images based on object-based classification. Proceedings, Asian Conference on Remote Sensing, Ulaanbaatar, Mongolia

Yen J-Y, Chen K-S, Chang C-P, Boerner W-M (2008) Evaluation of earthquake potential and surface deformation by differential interferometry. Remote Sensing of Environment 112:782-795

Yodmani S (2001) Disaster Risk Management and Vulnerability Reduction: Protecting the Poor. Paper presented at the Asia Pacific Forum on Poverty; Organized by the Asian Development Bank

Yong A, Hough SE, Abrams MJ, Cox HM, Wills CJ, Simila GW (2008a) Site Characterization Using Integrated Imaging Analysis Methods on Satellite Data of the Islamabad, Pakistan, Region. Bulletin of the Seismological Society of America 98(6):2679-2693

Yong A, Hough SE, Abrams MJ, Wills CJ (2008b) Preliminary results for a semi-automated quantification of site effects using geomorphometry and ASTER satellite data for Mozambique, Pakistan, Turkey. J. Earth Syst. Sci. 117(S2):797-808

Yu H, Zhao Y, Ma Y, Sun Y, Zhang H, Yang S, Luo Y (2011) A remote sensing-based analysis on the impact of Wenchuan Earthquake on the core value of World Nature Heritage Sichuan Giant Panda Sanctuary. Journal of Mountain Science 8:458-465

Yurur MT (2006) The positive temperature anomaly as detected by Landsat TM data in the eastern Marmara Sea (Turkey): possible link with the 1999 Umit earthquake. International Journal of Remote Sensing 27:1205-1218

Z

Zeng J, Zhu ZY, Zhang JL, Ouyang TP, Qiu SF, Zou Y, Zeng T (2012) Social vulnerability assessment of natural hazards on county-scale using high spatial resolution satellite imagery: a case study in the Luogang district of Guangzhou, South China. Environ Earth Sci 65:173-182

Zhang K, Chen SC, Whitman D, Shyu ML, Yan J, Zhang C (2003) A Progressive Morphological Filter for Removing Nonground Measurements From Airborne LIDAR Data. IEEE Trans. Geosci. Remote Sens. 41(4):872-882

Zhang J, Zhou C, Xu K, Watanabe M (2002) Flood disaster monitoring and evaluation in China. Global Environmental Change Part B: Environmental Hazards 4:33-43

Zhang K, Ng AHM, Ge L, Dong Y, Rizos C (2010) Multi-path PALSAR interferometric observations of the 2008 magnitude 8.0 Wenchuan Earthquake. International Journal of Remote Sensing 31:3449-3463

Zschau J, Isikara M, Ergünay O, Yalcin M N, Erdik M. (2002) Towards an earthquake early warning system for themegacity Istanbul. Early Warning Systems for Natural Disaster Reduction, edited by Zschau J and Küppers A, pp.433-440, Springer.