

TECHNISCHE UNIVERSITÄT MÜNCHEN
FAKULTÄT FÜR LUFTFAHRT, RAUMFAHRT UND GEODÄSIE
LEHRSTUHL FÜR METHODIK DER FERNERKUNDUNG

EARTH ORIENTED SPACE SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Analysis of Airborne Infrared Measurements to Estimate Localized Methane Emissions

Author:
Alexey Kupriyanov
(Matrikel-Nr. 03709771)

Supervisor:
Dr.rer.nat. Franz Schreier
Deutsches Zentrum für Luft- und Raumfahrt
Institut für Methodik der Fernerkundung

Assessor:
PD Dr.-Ing. habil. Adrian Doicu
DLR – Institut für Methodik der Fernerkundung
TUM – Lehrstuhl für Methodik der Fernerkundung

Munich – December 1, 2020

Statement of authorship

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

Munich
1December 2020

(Place, date)

Author (signature)

Abstract

Context. Methane is an essential atmospheric gas representing one fifth of the whole radiative forcing gases. The understanding of the behavior of methane is crucial for studying the climate change. Atmospheric methane concentration evolution through the last decades demonstrates a steady increasing trend and industrial sources of this gas, e.g. methane shafts are one of the sources of emissions into the atmosphere.

Aims. To make an analysis of the airborne Hyper Spectral camera (HySpex) infrared measurements over one of the largest methane sources in Europe – Upper-Silesian region in Poland in order to estimate localized emissions.

Methods. Four methods were introduced in this thesis based on the ratios and differences of the radiances from different wavenumber intervals. The focus of the work was concentrated at the methane absorption interval in short-wave infrared range (around 6000 cm^{-1} or 1.6 μm). Besides this, least-squares methods were used to estimate geophysical state parameters, i.e. ground albedo values and methane abundance enhancements from the observed data.

Results. One out of four presented methods show the presence of the methane plume over one of the sensed shafts from the various measurement altitudes: 1520 m and 2900 m. Moreover, two more weak methane sources were detected over another shaft. Comparison between the outputs of simulations with different changing parameters were done, e.g. different molecular absorption line parameter databases, diverse altitudes for the top of atmosphere (ToA) selection, divergent type of atmospheric datasets, distinct empirical albedo types, various water concentrations in the atmosphere etc. Finally, the contour plots with least-squares fitted albedo values and methane scaling factor over one of the detected plumes were demonstrated.

Conclusions. This work has shown the ability of detecting methane plume with an airborne sensor with high spatial (1.4 m), but low spectral (4.75 nm) resolution. One of the suggested methods can detect plumes with ~10-20 meters wide. Comparing the results of this thesis with the outcomes of other authors who also used the HySpex family of airborne sensor for methane plume detection, but with a totally different processing approach, one can conclude that the accuracy of the outputs in both cases is comparable. Also, there is a big field for further improvements that are mentioned in this work.

Key words: HySpex, infrared, methane, detection, plume

Acknowledgments

First of all, I would like to express my sincere gratitude to everyone who made this Master's work possible.

I would like to thank the following people who have helped me undertake this research:

- Head of Atmospheric Processors department at DLR – Prof. Dr. Thomas Trautmann, for accepting my application for the work on this thesis and for his interesting lectures at university about the introduction into atmospheric remote sensing.
- My supervisor from DLR side – Dr.rer.nat. Franz Schreier, for his guidance during the thesis, for his detailed explanations of physical and mathematical aspects and patience.
- Assessor of this thesis from TUM side – PD Dr.-Ing.habil. Adrian Doicu, for his work as the assessor of my thesis.
- PhD student at DLR – Philipp Hochstaffl for useful comments and remarks during this work and the essential help with some issues in Python coding.
- Dr. Dmitry Efremenko (DLR) – for his informative seminars at my university program which first introduced me to atmospheric physics.

To conclude, I want to express the biggest thanks to my family, my parents Dmitry and Marina, who set me off on the road to this MSc a long time ago. I can not forget to thank my friends and my girlfriend for all the unconditional support during this very intense Master's program.

Table of contents

List of illustrations

1. Introduction

1.1. Methane gas in Earth's climate

Methane (CH_4) is one of the most important anthropogenic greenhouse gases (GHG) which drives the global climate change. Jain Atul et al. (2000) mentioned that methane has a global warming potential (GWP) which is 84 times larger than CO_2 and represents one fifth of the whole radiative forcing gases. This gas also related to the group of so-called Essential Climate Variables (ECVs) which leads to global warming. Without any doubts the world community of scientists is convinced that global warming and climate change will have dramatic consequences, such as melting of glaciers, sea level rise, higher wildfire rates, more frequent severe weather events and etc. This is why it is especially important to understand the data, referring to the concentration of gasses in the atmosphere. It is also crucial to monitor the spatial and temporal distributions and evolutions of gas concentrations to improve our knowledge about anthropogenic and natural sources and sinks. Figure 1 depicts the flow of methane from sources into the atmosphere and the sinks, that consume methane.

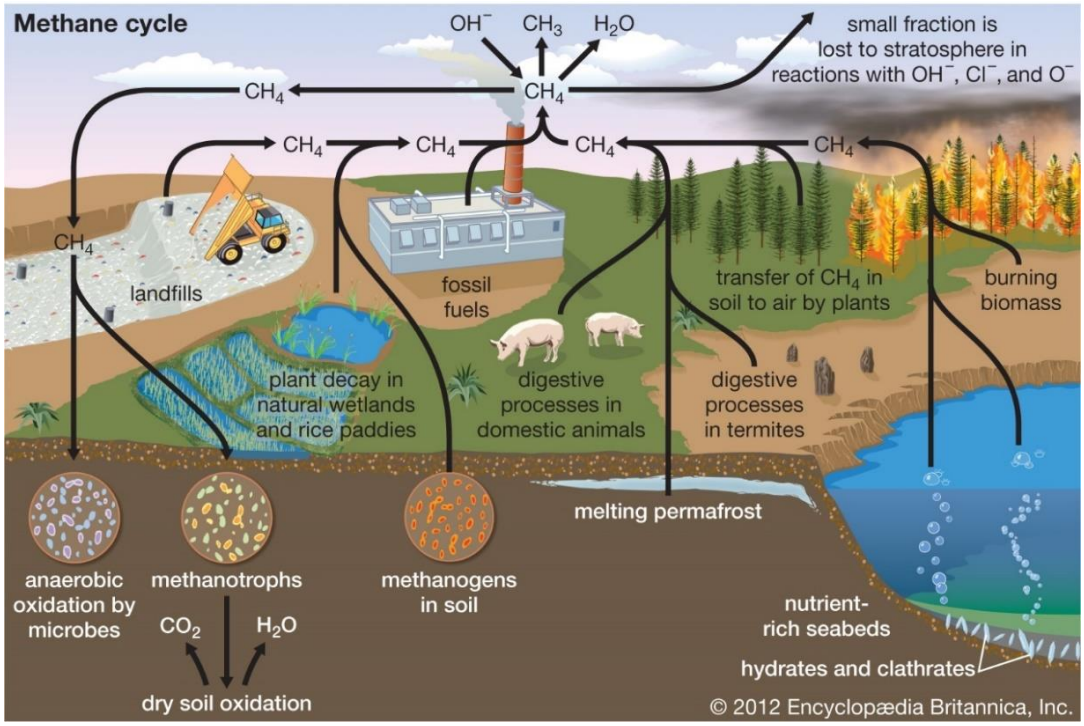


Figure 1 Global methane cycle diagram [Encyclopedia Britannica, 2012]

The sources of methane can be divided into natural and anthropogenic.

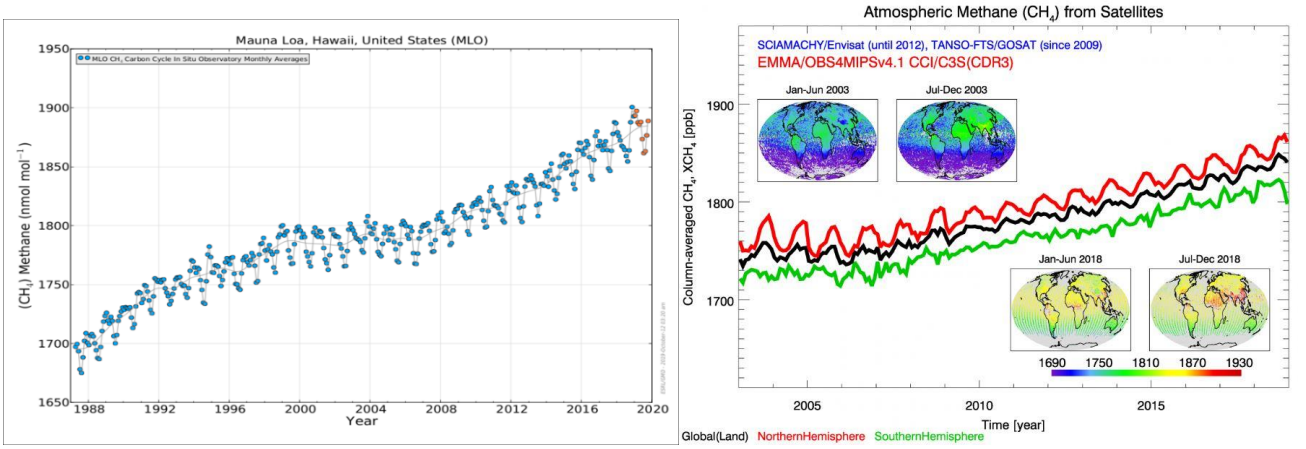
- **Natural sources of methane:**
  - **Wetlands/ rice paddies** – methane is produced due to the warm and moist environments
  - **Soils** – can be also source of methane due to the presence of the methanogens
  - **Permafrost/ glaciers/ ice cores** – during melting as global temperature rise, these sources slowly liberate trapped gas.
  - **Plants** – allow for direct travelling of methane from the soil into the atmosphere through the roots and leaves. Plants can be also direct producers of the methane.
  - **Methane hydrates** – located around the world on the nutrient-rich seafloors
  - **Forest fires** – burning biomass releases methane into the atmosphere
- **Anthropogenic sources of atmospheric methane:**
  - **Landfills** – significant source of methane due to decaying organic matter and anaerobic conditions

- **Thermal power plants** - exhaust gases from power plants and coal-fired thermal power plants
- **Farm animals** – digestive processes in domestic animals are also origin of methane

- **Removal processes (sink of the gas):**
  - **Methanotrophs in soils** – major sink for methane through the methanotrophic bacteria (it uses methane as a source of energy). Reaction equation of dry soil oxidation:
    $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$
  - **Anaerobic oxidation** by microbes in soil.
  - **Hydroxyl radical** (OH^-) - the largest sink for methane in the atmosphere, moreover it is one of the most significant sources of water vapor in the upper atmosphere. Reaction equation: $CH_4 + OH^- \rightarrow CH_3 + H_2O$
  - **Chlorine radical** $(Cl^\bullet)$ & **Oxide radical** $(O^\bullet)$ - free chlorine and oxide reacts with methane in the atmosphere. Reaction equation, e.g. with chlorine radical:
    $CH_4 + Cl^\bullet \rightarrow CH_3^\bullet + HCl$

Global anthropogenic methane emissions in 2019 according to the [www.globalmethane.org] were estimated as 9.390 million metric tons. Approximately 10% of these emissions came from coal mining where methane released from the coal and surrounding rock strata due to mining activities.

According to the temporal trends of methane concentration in the atmosphere, long term measurements by NOAA showed that the current concentration of methane in the atmosphere nearly tripled since pre-industrial times and reached the highest value in at least 800.000 years. Remarkably methane concentration is higher in Northern hemisphere. This is because the majority of natural and anthropogenic sources are located there.



(a) (b)

Figure 2 (a) – Atmospheric methane concentration evolution above Hawaii from 1987 to 2019 [Dlugokencky et al., (2020)]; (b) – An overview of global methane levels from 2003 to 2019 [Greenhouse gases CCI (www.cci.esa.int)]

Figure 2 (a) shows the atmospheric methane concentration evolution in Mauna Loa, Hawaii (USA), measured by NOAA. It clearly shows a long-term increase trend of CH_4 concentration, that started since pre-industrial times, but then a "temporary stop" of this growth took place in 2000-2007. Scientists do not fully understand the reasons for that "temporary stop". One of the reasons could be reduced industrial emissions and drought in wetland areas during that years.

Figure 2 (b) are the combination of the measurements of ESA's Envisat satellite, Japanese GOSat and NASA's AIRS Obs4MIPs. On this figure it is clearly seen, that methane concentrations are increasing in the last decade

despite several global efforts to reduce anthropogenic emissions. The colored lines show the average amount of atmospheric methane in different hemispheres and worldwide. Methane concentrations were roughly constant for some years around 2000, but after that have been increasing about 0.3% per year.

Another important trend on this graph are seasonal fluctuations, for example higher concentrations of methane in the south-east Asian region (i.e. India and China and others) during the second half of the year. Wetlands and rice paddies, located majorly in these regions, with warm and humid climate, are the main source of methane emissions. Other parts of the world, for ex. the Tropics, the USA and some parts of Russia have a similar seasonal behavior.

1.2. Radiative transfer in the atmosphere

One of the most challenging tasks of the remote sensing of the atmosphere is the quantitative estimation of all parts that included into it. The retrieval of information about trace gases or any other atmospheric parameters (e.g. temperature) is based on knowledge about absorption, emission and scattering of electromagnetic radiation in the atmosphere. Figure 3 depicts the atmospheric propagation effects.

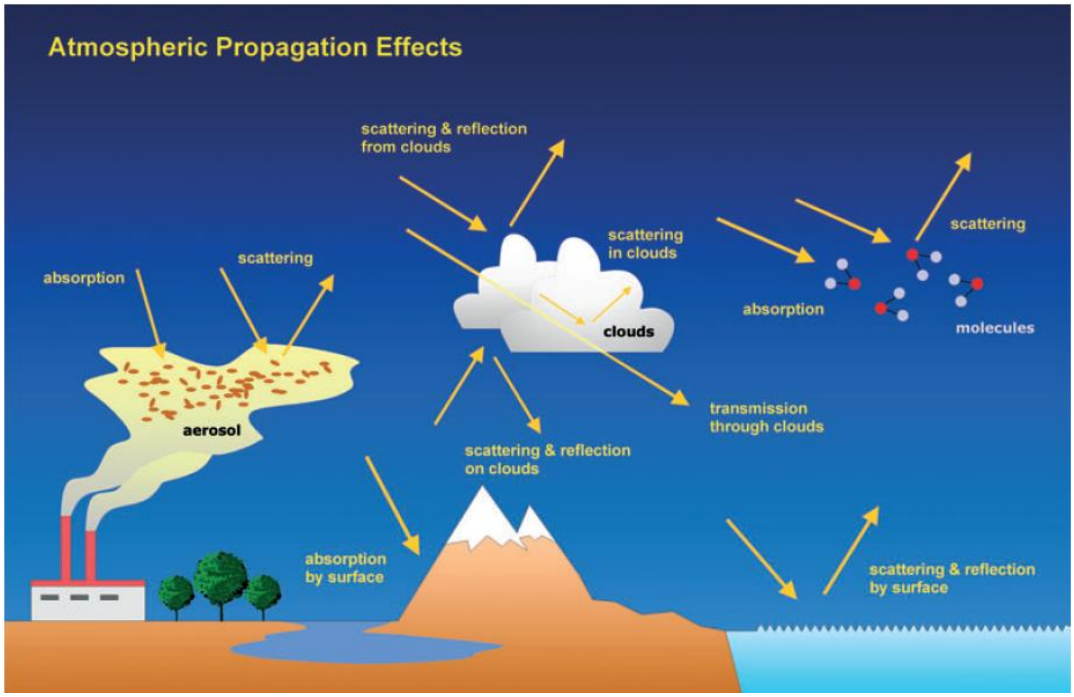


Figure 3 Scheme of the relevant interactions of light with the Earth's atmosphere and surface [Gottwald et al. (2011)]

Gottwald et al. (2011) highlighted in their book the following effects, that impact to the radiative transfer through the atmosphere:

- Scattering by air molecules (Rayleigh/ Raman scattering),
- Scattering and absorption by aerosol and cloud particles (Mie scattering),
- Absorption and emission by air molecules,
- Refraction due to the density gradient,
- Reflection form Earth's surface.

Emission of the Earth's surface and the atmosphere can be observed in the "thermal" IR (λ >3 μm). They have spectral distributions described by the Planck function corresponding to the local temperature i.e. of Earth's surface or temperature distribution with altitude. Special term "emissivity", which is wavelength dependent, was explained in Burrows et al. (2011), it was introduced because neither Earth's surface nor atmosphere are perfect black bodies.

It is important to measure first in the laboratory characteristics of absorption and emission of each molecule in order to predict the variation of the molecular spectra as a function of temperature, pressure etc. Schreier et al. (2006) mentioned, that laboratory spectroscopy in support of atmospheric remote sensing has significantly helped with creating several molecular spectroscopic databases of line parameters (i.e. line positions, line strengths etc.) or absorption cross sections.

Different molecules from all gases contributes into the absorption depending on the wavenumber. The following figure demonstrates the total transmission of the atmosphere (upper graph), water molecules transmission (middle graph) and main carbon gases transmission (lower graph) in the thermal infrared (TIR) and in the short-wave infrared (SWIR) range as a function of transmission from the wavenumbers.

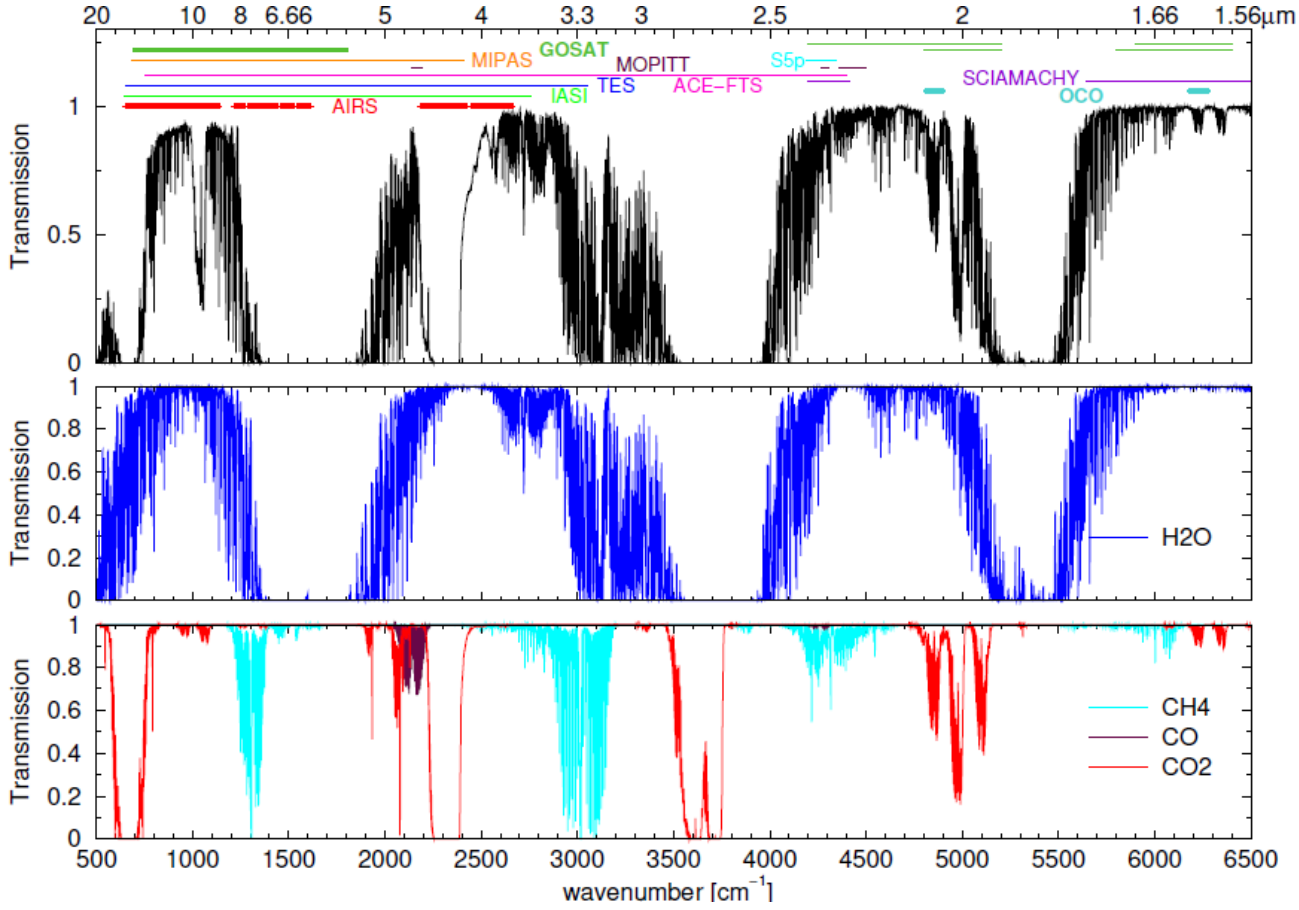


Figure 4 Total atmosphere transmission (upper graph), water molecules transmission (middle graph) and main carbon gases transmission (lower graph) in TIR and SWIR as a function of the wavenumbers [Schreier F.; personal communication]

1.3. Literature overview

Before introducing the research objective, let us have a look what was done so far in the research field and which problems are mentioned by the authors.

The history of remote sensing satellites begins with the appearance of the first artificial Earth satellites. Humanity has tried to understand the processes taking place in Earth system via using satellites because they allow contactless, remote observations that has a great coverage. The development history from early remote sensing satellites to the latest Earth observation missions is described by Burrows et al. (2011). In this book

authors also notice, that since pre-industrial times, as the population grows significantly, the required amount of energy also increased drastically. Hence, widespread use of coal etc., leads to significant air pollutions on local, regional and global scales. This and other anthropogenic activities are changing now the entire atmosphere, which in turn leads to global climate change. Global climate change resulting from man's activity is currently receiving much public attention and has been recognized as a threat to both the biosphere and mankind. Burrows et al. (2011) also highlight the importance for global tropospheric measurements of the greenhouse gases. The problem in a nowadays that a ground-based monitoring network are sparse and an accurate knowledge of the regional sources and sinks of pollutants is needed. This can be achieved by using a combination of spaceborne, airborne and ground-based measurements.

According to the estimation of the various gases such as methane (CH_4), carbon dioxide (CO_2), carbon monoxide (CO) etc. for these purposes nadir observations in SWIR range by different sensors can be used. One of these sensors was Scanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) – one of the ten ESA's Environmental Satellite (ENVISAT) instruments.

Among the publications that were under consideration, there were two, which focus on the retrievals from SCIAMACHY by Gimeno Garcia et al. (2011) and another one by Hochstaffl et al. (2018). In both works BIRRA (Beer InfraRed Retrieval Algorithm) was used for the operational analysis of carbon monoxide (CO) vertical column densities.

Gimeno Garcia et al. (2011) mentioned that BIRRA is being prepared for methane retrievals. It is also important to note, that for nadir modelling in this article, refraction was only taken into account for the Sun-Earth path element, whereas for the Earth-satellite (observer) with Observer Zenith Angle $(OZA) \leq 30^o$ refraction was neglected. The scattering was completely neglected in the forward model. In order to take into account scattering and refraction particularly at 2.3 μm band, authors recommend select methane as a proxy because it has strong absorption lines across the CO spectral fitting window. Moreover CH_4 is a well-mixed gas with long life time in the atmosphere and hence it has quite homogeneous global distributions.

The terms forward model and inverse model, that are intensively used in retrieval techniques were discussed in Schreier (2019). Forward model is used simulate the radiance field, which is observed by the instrument or sensor at their location. The radiance spectrum depends in a non-linear manner on the various atmospheric parameters. Inverse model is a kind of "way back", i.e. retrieval of the atmospheric parameters from the observed radiance spectra.

Besides computing the forward model, Rodgers (2000) noticed, that one has to also design the model for computing weighting functions and Jacobians of the forward model. He also mentioned, that if the absorption coefficient is independent of the wavelength or wavenumber, one can solve radiative transfer equation (RTE) only once, while in general case, it varies strongly and rapidly with the wavelength or wavenumber. Hence, the RTE has to be solved again and again and these make calculations extremely complicated.

Gimeno Garcia et al. (2011) mentioned that sometimes aerosol and cloud scattering may have a significant contribution to the intensity. Liuzzy et al. (2017) presented code in their work, that is able to calculate both clear and cloudy sky radiances and also their Jacobians with respect to the desired geophysical parameter. In the latest publication by Kumar et al. (2020) the authors introduce a method that is able to solve problems, which usually occur in hyperspectral data: human errors due to plumes small pixel-footprint signature, e.g. omission of the plume's presence etc. The proposed Hyperspectral Mask-RCNN uses signal processing and deep neural networks to address these limitations.

Coming back to the importance of methane, Thompson et al. (2017) highlights that methane has much stronger global warming potential (GWP) than an equivalent quantity of carbon dioxide. The main challenge according to the authors is to distinguish CH_4 sources from background noise. This work describes a dataset that comes from NASA's Airborne Visible InfraRed Imaging Spectrometer-Next Generation (AVIRIS-NG), airborne

spectrometer. The usage of airborne observations was explained by much better spatial resolution on comparison to the space observer. Finally, more than 250 distinct sources of methane were found in the Four Corners area (USA), which can drastically improve the knowledge about fugitive sources.

The following articles can be combined into a group dedicated to retrieval of local methane emissions from various airborne sensors which is close to the topic of this Master thesis.

Thorpe et al. (2014) specifically noted that between 60%-70% of all worldwide emissions are presently anthropogenic and the majority of anthropogenic CH_4 emissions are from the energy sector. In their work by using high spatial and moderate spectral resolution data from Airborne Visible/Infrared Spectrometer (AVIRIS) two quantitative retrieval techniques were evaluated in order to estimate methane enhancement in concentrated plumes over marine and terrestrial locations from the flight altitude 8.9 km. The usage of airborne sensor authors explained by convenient combination of large image footprints and fine spatial resolution which are well suited for mapping local methane plumes. So, the first technique presented by the authors is an iterative maximum a posteriori differential optical absorption spectroscopy (IMAP-DOAS) showed good results over the ocean scene with natural methane emissions from Coal Oil Point (COP) seep field near Santa Barbara (California). The retrieval precision errors over the lowest atmospheric layer (up to 1.04 km) were between 0.31 and 0.61 ppm CH_4, which corresponds to about 30-60 ppm error for a 10 m thick plume. However, the authors noted that IMAP-DOAS results for a terrestrial area were negatively impacted by the underlying differences of the land cover. Second, a hybrid approach based on singular value decomposition (SVD) that was effective for terrestrial scenes because it can better account for spectral variability in surface albedos. By this approach the maximum near surface enhancement above plume was by 8.45 ppm larger than the background. Over COP the authors present maximum enhancement of 2.85 ppm above background. One of the conclusions according to this method the authors mentioned about averaging: the maximum concentration will increase when scaled for smaller atmospheric column. In their work the authors also commented a cluster-tuned matched filter (CTMF) that was able to map methane plumes over marine and terrestrial scenes but does not directly quantify gas concentrations in contrast to the IMAP-DOAS and hybrid SVD approaches.

Krings et al. (2011) used a non-imaging passive spectrometer system called Methane Airborne Mapper (MAMAP) comprising two spectrometers in SWIR and NIR bands for column-averaged methane observations. Data acquisition was done on the altitude 1250 m and the retrieval precision of the measured column in comparison to the background is typically $\leq 1\%$. In comparison to the AVIRIS instrument, MAMAP has higher CH_4 sensitivity but limited to flying transects across local gas plumes due to the small field of view (FOV). In their work Krings et al. (2011) implemented two different inversion approaches: based on optimal estimation scheme in order to fit Gaussian plume models and another approach using a simple Gaussian integral method. Since the standard DOAS technique assumes independence of the absorption cross section with the height, for the processing MAMAP data a modified version of the Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) was used. WFM-DOAS takes into account the pressure and temperature dependency. For the retrieval the authors used solely constant albedo of 0.18 assuming a Lambertian reflector and no spectral dependency. In their study the authors named wind as a key parameter because it is directly influence to the error on the emission rates. In order to provide results stable against variations special threshold criterium was used during the averaging: each burst of the 10 single measurements has been accepted as an average if more than half of the measurements passed root mean square (rms) threshold criteria. Another interesting conclusion were done by the authors about flight track. There was an assumption that there is no systematic error on the inversion result due to the flight track. This assumption is valid for the Gaussian plume model fit but not for the integral method. The final conclusion is the following: MAMAP sensor has the ability to quantify point source emission rates from power plants.

Humpage et al. (2018) introduced another airborne SWIR spectrometer for remote sensing of greenhouse gases – GreenHouse gas Observations of the Stratosphere and Troposphere (GHOST). This instrument uses an

original design comprising optical fibre inputs along with a single diffraction grating and detector array which allows to observe simultaneously carbon dioxide and methane absorptions in 4 different bands. This allows to save space and weight in comparison to other instruments which use separate gratings and detectors for each band. The authors declared that due to benefits achieved in spatial coverage, overlapping spectral ranges and comparable spectral resolutions, GHOST is a unique sensor that can be used in airborne platforms with special efficiency over marine areas, where ground-based measurements are not available.

The article by French research group Nesme et al. (2020) presented an algorithm, developed for the airborne Hyper Spectral camera (HySpex-NEO), that was able successfully to detect industrial methane plume and quantify emission source. They validated their results in a campaign over Lacq industry (France) and then apply their algorithm to the AVIRIS sensor installed on a satellite. The core idea of their algorithm is based on estimation of the transmission of the observed methane plume in order to estimate its integrated concentration pixel by pixel. In general, their algorithm can be divided in 2 steps: Preprocessing – including k-means clustering and Detection – using Cluster-Tuned Matched Filter (CTMF) algorithm. A bit detailed look to k-means clustering and CTMF algorithm is done in the following 2 articles.

Funk et al. (2001) in addition to explaining how the CTMF algorithm works gave an arguments about the importance of using k-means clustering before it. K-means clustering algorithm was chosen by the authors for several reasons, i.e. it is computational efficiency and simplicity, it provides direct way to minimize the within-class variance, maximize within-class correlation and moreover it improves the performance of matched filters. In the step when pixels assigned to clusters the authors recommended to use Mahalanobis distance instead of Euclidean because it reduces the number of potential false alarms. Finally, the authors compare three different clustering variations but the main point from their work is that clustering together with clutter matched filtering can dramatically improve detection of the weak signals, e.g. methane plumes.

Ranjan et al. (2017) presented in their work a new scheme of k-means clustering of hyperspectral images. Their algorithm includes 3 steps: Principal Component Analysis (PCA) for dimensionality reduction, then apply k-means clustering to the reduced features and the last step – training the clusters separately by multi-class support vector machine (M-SVM). The authors declared that this algorithm gives better results and is more time effective.

1.4. Research objective

After introducing the overview through the whole research field, it is possible to formulate the research objective of this work. One can notice that the aim of most of the tropospheric works – to hold kind of analysis e.g.: detect certain gases, quantify gas concentrations or produce column densities and concentrations profiles of gases and aerosols etc. This work is focusing on retrievals of methane (CH_4) from shafts by the airborne Hyper Spectral camera (HySpex) over the south-west Poland in the Upper-Silesian region.

Objective: Analyze the HySpex airborne sensor IR measurements over methane plumes in order to estimate localized gas emissions.

During the current research work the following targets were highlighted:

1) Hold a series of preliminary simulations in Python for Computational ATmospheric Spectroscopy (Py4CAts) package for understanding main molecules strength and cross sections, total transmission behavior with altitude, comparison of different HITRAN database versions etc.
2) Perform a simulation of the observed radiance with and without ground albedo. Compare modelled radiances with empirical and constant albedo with HySpex measurements samples.
3) Compare modelled observed radiance with normal, doubled and tenfold methane concentration in the atmosphere at different resolutions.
4) Select proper wavenumber interval for methane plume retrieval.
5) Select the most prominent methane shaft for plume retrieval.
6) Implement various methods for methane plume retrieval and compare their results.
7) Implement least-squares fitting of the albedo values and methane concentration scaling factor.
8) Conclude and discuss obtained results and further perspectives of the work.

The main question that should be answered during this work is the following: is it possible to estimate localized methane emissions above plumes from industrial sources, areas with elevated concentrations, using a low-resolution spectrometer (HySpex)? The following tradeoff should be clarified: can data with high spatial resolution compensate low spectral resolution?

This Master thesis is organized as follows: chapter 1 ‘Introduction’ demonstrates the importance of methane in Earth’s climate as well as a literature overview and formulation of research objective. In chapter 2 ‘HySpex sensor system’ all essential technical details, i.e. working principle, instrument resolution etc. are presented. The thesis continues with a chapter 3 ‘Theory and Methods’ where definition of the main variables, mathematical formulation of linear and nonlinear least-squares problems as well as important theory for implemented four methods are done. In chapter 4 ‘Py4CAtS results’ result graphs of various simulations are demonstrated and commented. In the chapter 5 ‘Methane plume retrieval’ explanation of selection wavenumber intervals with methane (1.6 μm and 2.3 μm) as well as their comparison are done. In this chapter the outputs of four different methods for plume detection presented. Chapter 6 ‘Py4CAtS methane plume radiance simulations’ introduces the comparisons between observed and modelled spectrums in across-track and along-track directions also in this chapter least-squares fitted albedo values and methane concentration scaling factor presented as a contour plots and the comparison between atmospheric datasets MLS and CIRA. Finally, the thesis ends up with chapters 7 and 8 where discussion of the obtained results and conclusions are presented as well as the following ones list of acronyms, code snippets and references.

2. HySpex sensor system

2.1. CoMet campaign and measurement principle

HySpex – Hyper Spectral camera is operated by the Remote Sensing Technology Institute (IMF) at the German Aerospace Center (DLR). Detailed introduction into working principle of this sensor as well as system overview and data processing chain was done by Köhler (2016). All further sections in this chapter 2, except sections 2.4 and 2.5, are referenced to that article.

The HySpex instrument is certified on two DLR research aircrafts: a Cessna 208 Grand Caravan (D-FDLR) and a Dornier 228-212 (D-CFFU). The measurement data that are used in this thesis was acquired by the Dornier 228-212 during a flight to measure emissions from coal mine shafts with methane usage in the Upper Silesian region (south-west Poland). This area is one of the largest methane sources in Europe with the annual methane emissions according to the Kedzior et al. (2019) around 950 million metric tons. The gas is liberated there as firedamp from numerous anthracite mines. The measurement flight was a part of CoMet campaign for HALO held by IPA (DLR).

Figure 5 Cessna Grand Caravan measure emissions. [German Aerospace Center (2020) www.dlr.de]

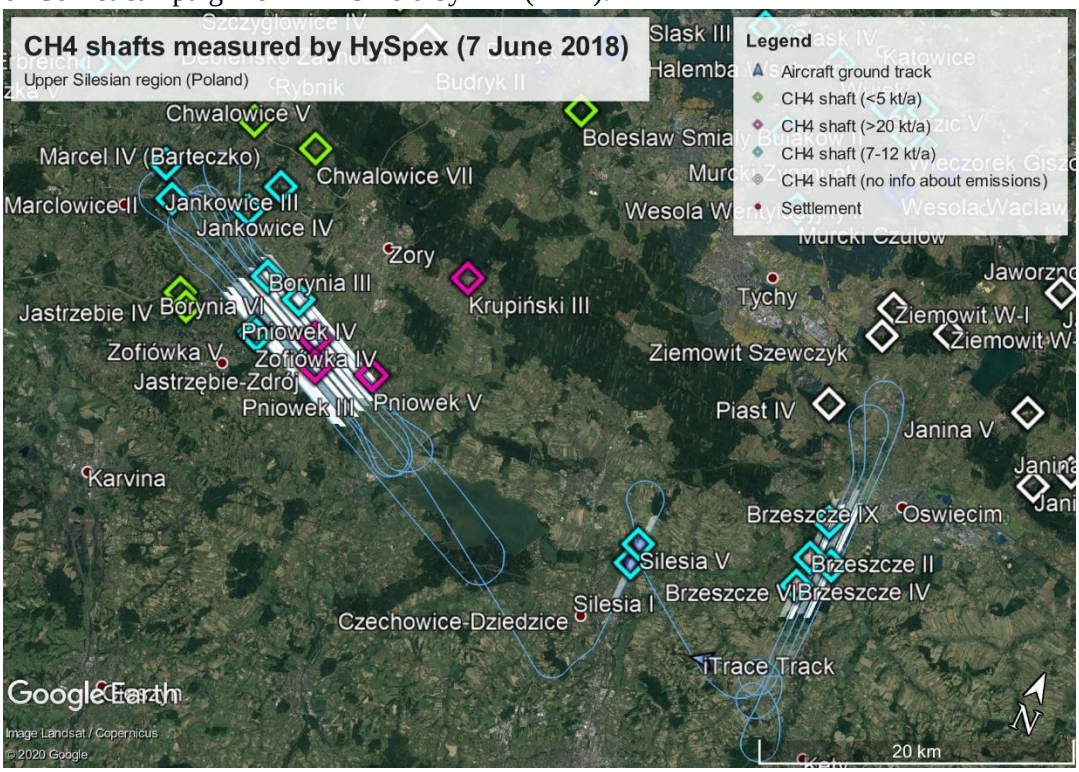


Figure 6 Methane shafts measured by Dornier 228-212 with activated HySpex sensor (white lines)

2.2. Spectrometer system

The HySpex instrument belongs to the group of pushbroom or along-track scanners. The schematic representation of data acquisition by this type of scanners is shown in Figure 7. This means that a line of sensors arranged perpendicular to the flight direction of the satellite.

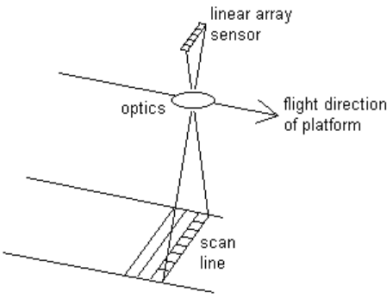


Figure 7 Scheme of data acquisition of pushbroom scanners [Butler et al. (1988)]

The HySpex spectrometer system consists of two different pushbroom cameras that covers VNIR (416-992 nm) and SWIR (968-2498 nm) spectral domains. The working principle of a HySpex instrument is depicted on Figure 8.

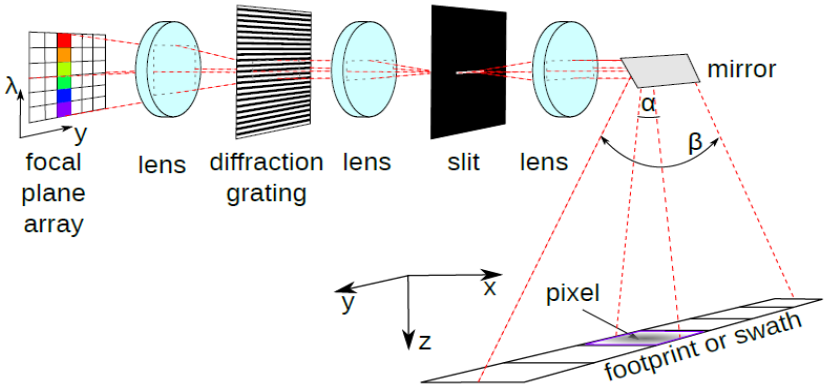


Figure 8 Working principle of a HySpex instrument.

The spectrum of the purple shaded pixel is mapped to the colored column on the focal plane array. The movement of the spacecraft is done in the x direction and allows to make the combination of consecutive footprints. Variable β stands for field of view (FOV) determined by the swath width and variable α - for instantaneous FOV related to the size of individual pixel. Radiation that comes from the individual pixel is decomposed by grating into its spectral components. Due to this step, the one-dimensional FOV β is mapped into two dimensional (λ, y) intensity distribution in the focal plane. Each image that is acquired by focal plane array is called a frame.

The HySpex VNIR-1600 consists of silicon charge coupled device (CCD) detector, that is normally used in cameras for measuring intensity distribution on focal plane array and covering 160 channels. The spectral resolution ranges from 3.5 nm at nadir to roughly 6 nm at the outer edge of the swath. Remarkably that the size of pixels in VNIR-1600 sensor are not square, but rectangular. This is since along track instantaneous FOV are approximately twice as large than across track instantaneous FOV.

The HySpex SWIR-320m-e include a mercury cadmium telluride (MCT) detector with 256 channels with a spectral resolution varying from 6-12 nm. In comparison to the VNIR-1600, the SWIR-320m-e pixels are square in shape. In the overlap region of two consecutive swaths each of 320 pixels from SWIR covers approximately 4 pixels of VNIR in across-track and 2 pixels of VNIR in along track directions. Köhler (2016) emphasizes that both cameras (sensors) can be equipped with a removable FOV expander lens, making the HySpex suitable for various purposes. The HySpex instrument data that was used in this work was acquired exactly from the camera SWIR-320m-e.

Figure 9 shows the experimental setup of the HySpex sensor system as well as its real appearance on board of the plane. On figure depicted: VNIR-1600 and SWIR-320m-e cameras (sensors), both facing nadir (z direction), iTrace RT – high precision ADCS for georeferencing acquired data with internal measurement unit (IMU), PAV 30 – stabilizing platform for reducing aircraft motion with an aid from Litton LN-200 IMU and finally Copper Weights – specifically designed weights which purpose is to move center of gravity and all inertial moments into the operational envelope of the PAV 30.

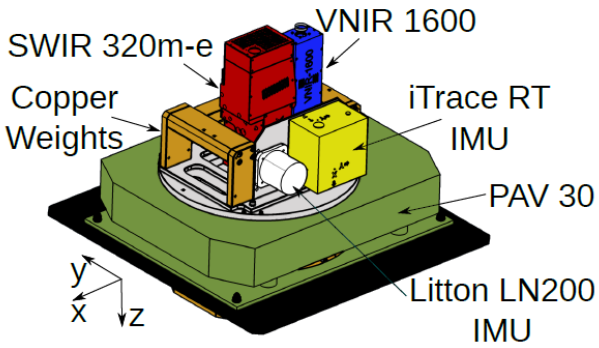


Figure 9 (on the left) – experimental setup of the HySpex sensor system; (on the right) – HySpex sensor (on the foreground) and operating rack during the research measurement flight on 7 June 2018 [German Aerospace Center, 2018]

The schematic representation of the geometry of the observations as well as the allowable regions of the Sun/Solar Zenith (SZA) (red area) and Observer Zenith (OZA) (green area) angles depicted on Figure 10. The sunlight which passed through the atmosphere and then been reflected by Earth's surface onto the observer (HySpex airborne instrument).

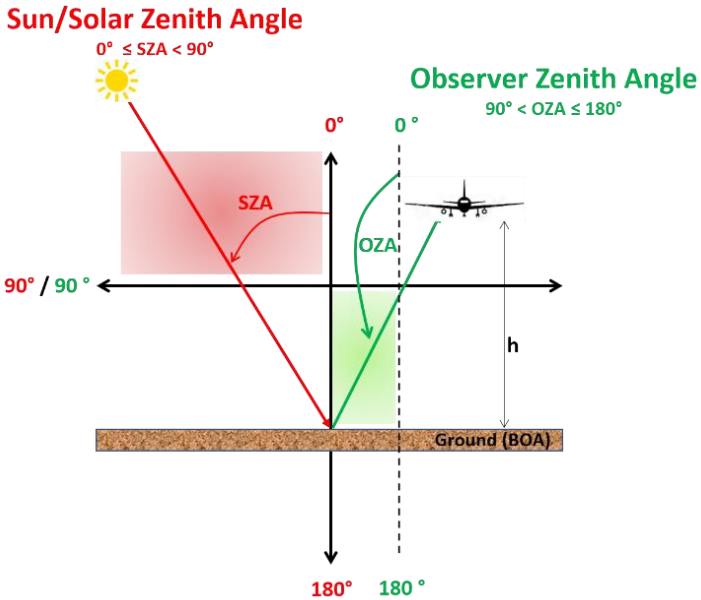


Figure 10 Geometry of the observations. Allowable regions of Sun/Solar Zenith Angle (SZA) and Observer Zenith Angle (OZA).

2.3. Data processing

HySpex data processing chain organized referred to as levels:

- **Level 0** – raw data
- **Level 1A** – raw spectra with synchronized navigation data for each frame
- **Level 1B** – at-sensor radiance including navigation data for georeferencing
- **Level 1C** – orthorectified radiance
- **Level 2A** – orthorectified surface reflectance

The main idea of pre-processing on levels 0 and 1A can be formulated as synchronization of navigation data and acquired HySpex images. On the levels 1A and 1B takes place conversion from raw digital numbers recorded by the sensor to the calibrated at-sensor radiance, including a system correction for several sensor artefacts.

Several algorithms and softwares are used for further steps i.e. achieving high co-registration accuracy and mitigation the radiative effects of gases and clouds on the image data. Detailed description of all these algorithms and softwares are beyond the scope of this work but can be found in Köhler (2016).

2.4. HySpex camera SWIR-320m-e resolution

In infrared spectroscopy both wavelength and wavenumbers are used by research community. In this section connection between wavenumbers and wavelength is introduced.
The relation between the wavelength and the wavenumber is the following:

$$\nu = \frac{1}{\lambda} \tag{1}$$

The resolution $(\triangle\lambda)$ of the instrument can be defined as full-width half maximum (FWHM) or half-width half maximum (HWHM) of the spectral response function. The only difference between these two approaches is the factor of 2. In the short-wave infrared domain FWHM resolution equals to $\triangle\lambda = 9.5$ nm or $\triangle\lambda = 9.5 \cdot 10^{-7}$ cm for the CH_4 window at 6000 cm^{-1} (spectral channel 117) in the vicinity of center of geometric pixels. This value was selected in accordance with the full-width half maximum (FWHM) map for the HySpex camera SWIR-320m-e depicted on figure below.

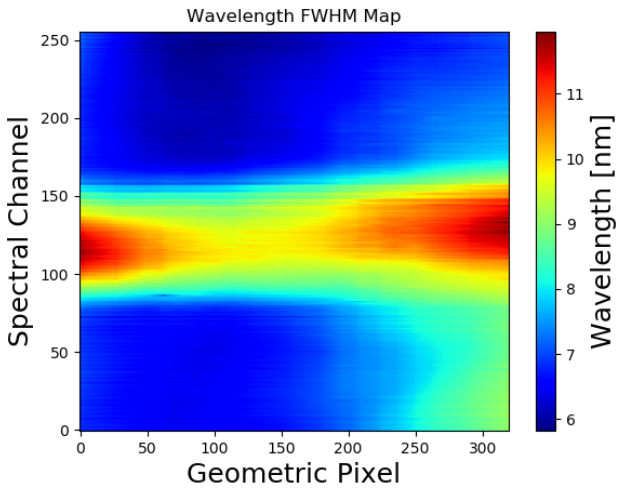


Figure 11 Wavelength FWHM map for HySpex camera SWIR-320m-e [Köhler C.; personal communication]

Figure 11 shows the wavelength full-width half maximum map of the HySpex camera SWIR-320m-e. It is important to note that 256 spectral channels correspond to the wavenumber or wavelength range covered by the HySpex instrument. In other words, 1-st spectral channel corresponds to the wavelength at 2.5 μm and 256-th spectral channel to the wavelength at 0.97 μm. Also, the across-track width of the HySpex swath is 320 pixels and one can see how the spectral resolution changing on different parts of the image. More clear and straightforward representation of the spectral resolution in the wavenumber and wavelength domains along the central geometric pixel with the index 160 is depicted on Figure 12.

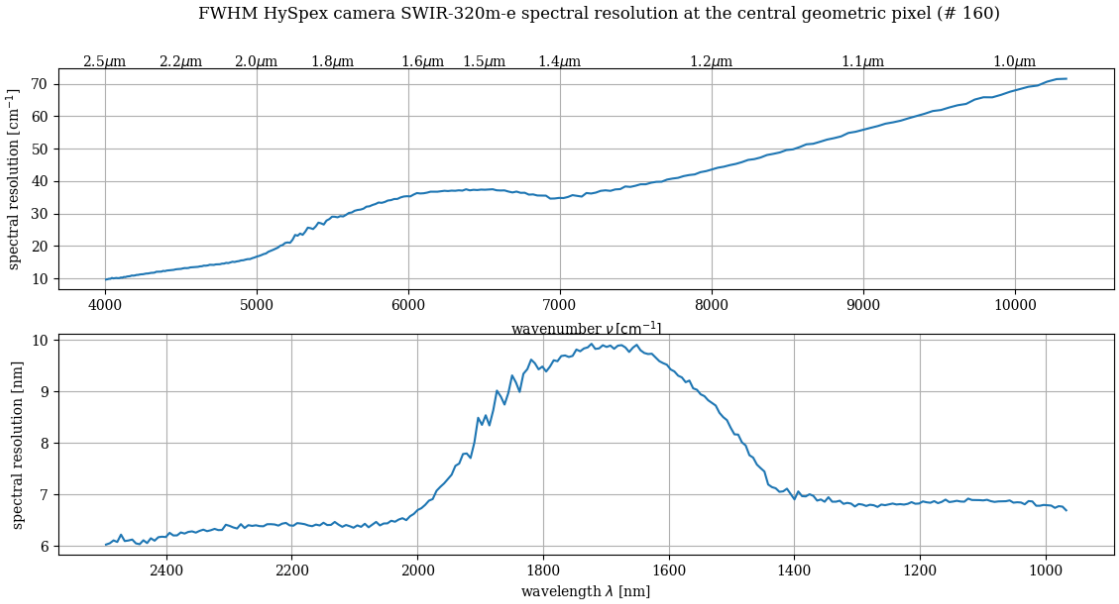


Figure 12 FWHM spectral resolution of the HySpex camera SWIR-320m-e at the central geometric pixel (#160) in the wavenumber and wavelength domains (upper graph) and only in wavelength domain (lower graph)

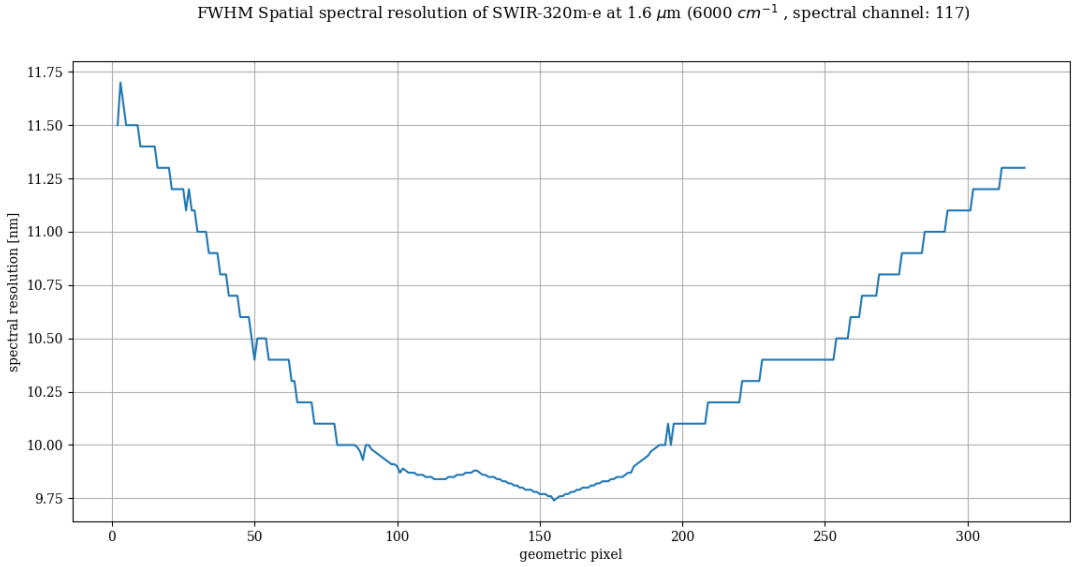


Figure 13 FWHM spatial spectral resolution of the HySpex camera SWIR-320m-e at 1.6 μm (~ 6000 cm^{-1}), which corresponds to the spectral channel 117.

The choice of this particular interval 1.6 μm is not made by chance, since methane has absorption bands in this interval, at 2.3 μm and in other intervals. More detailed explanation and comparison of these two intervals are presented in chapter 5 ‘Methane plume retrieval’.

The resolution in terms of wavenumber can be calculated by the formula:

$$\Delta\nu = \left|\frac{\partial\nu}{\partial\lambda}\right|\Delta\lambda = \frac{\Delta\lambda}{\lambda^2}$$

where the wavelength corresponding to methane can be calculated via the formula:

$$\lambda = \frac{1}{\nu} = \frac{1}{6000\ \text{cm}^{-1}} = 1.6\times10^{-4}\ \text{cm} = 1.6\ \mu\text{m}$$

Putting all values together will lead to the instrument resolution in terms of wavenumber:

$$\Delta\nu = \frac{9.5\cdot10^{-7}\ \text{cm}}{\left(1.6\cdot10^{-4}\right)^2\ \text{cm}^2} = 37.1\ \text{cm}^{-1}$$

Hence, the half-width half maximum (HWHM) resolution in the wavenumber domain is half of the calculated value or 18.5 cm^{-1} and in the wavelength domain – 4.75 nm.

2.5. Measurements illustration

Figure 14 demonstrates HySpex airborne measurements that were done on the altitude 1520 m (lines 1-9) and on 2900 m (lines 10-13) over the Upper Silesian region (south-west Poland) by SWIR-320m-e camera. One can notice that the density of measurement points decreases with the increasing of the wavenumbers. This is since HySpex instrument has a fixed sampling width in wavelength domain, but in the wavenumber domain it is changing, according to the Equation 1 from previous section. Hence, we have denser measurements at the small wavenumbers (left side of the graph) and more coarse measurements in high wavenumbers (right side of the graph).

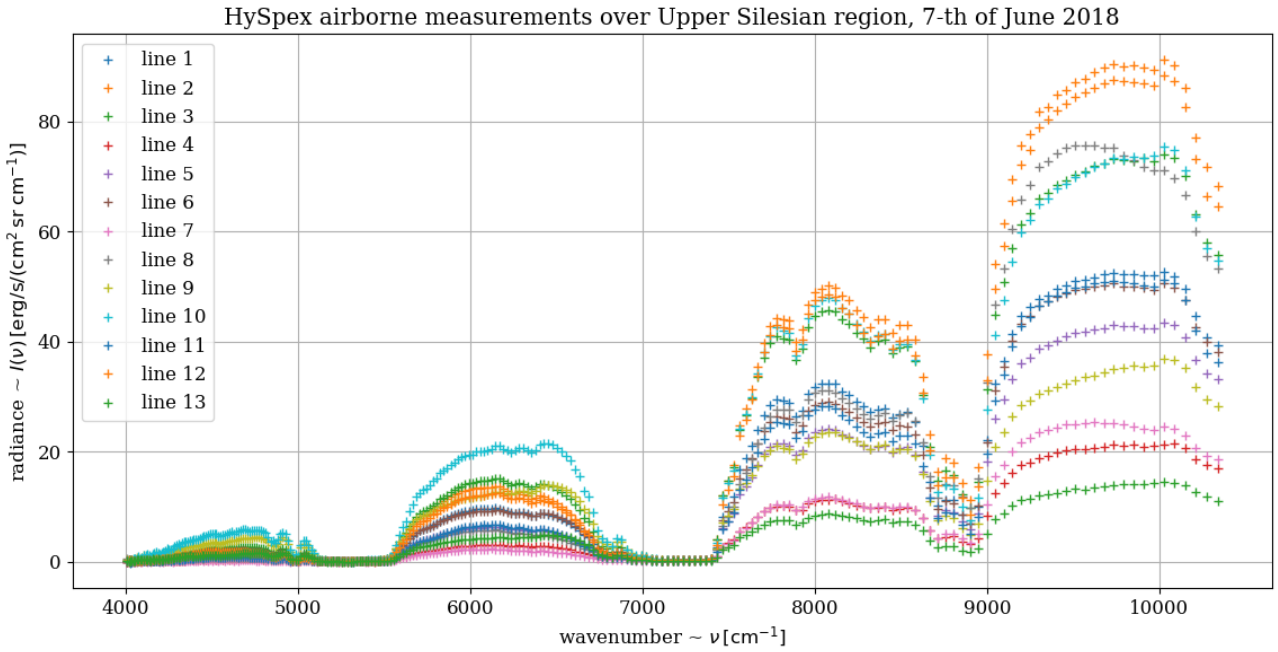


Figure 14 Plotted measurements of HySpex airborne sensor on the very first pixel of the spatial matrix (bottom left corner) on different lines (altitude of the flight for the lines 1-9: 1520 m; altitude of the flight for the lines 10-13: 2900 m) through the all wavenumbers covered by HySpex.

As it is clearly seen from the above graph and taking into account Figure 4 with the total transmission through the atmosphere and transmission of the particular gases, there is a prominent methane detection ‘peak’ of the

radiance roughly around 1.6 μm (~6000 cm^{-1}). The smaller but also promising wavenumber interval for CH_4 retrieval is in the vicinity of 2.3 μm (~4347 cm^{-1}). In the 'peaks' corresponding to higher wavenumbers ~7500 - 8500 cm^{-1} and from ~9000 cm^{-1} scattering as well as the temperature dependence should be taken into account. These intervals correspond to the near infrared (NIR) range which is out of the scope of this work.

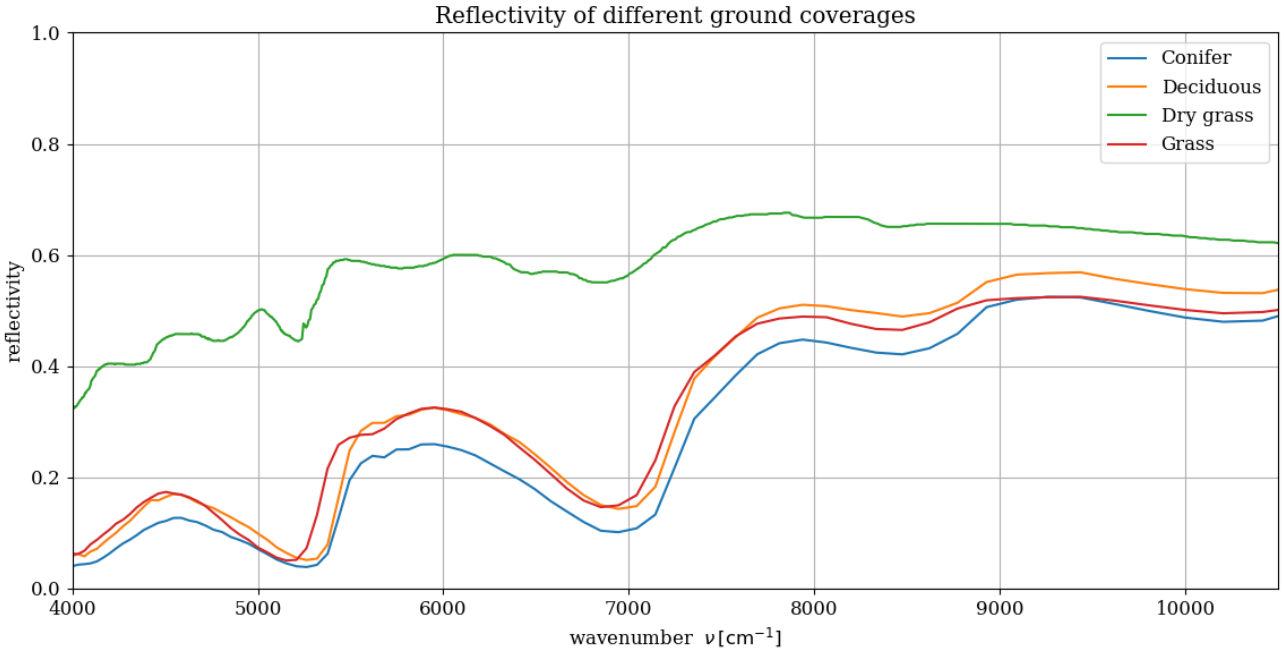


Figure 15 Comparison of reflectivity for different ground coverages in the HySpex spectral range

Taking into account Figure 15 one can note that the intervals corresponding to the minima of the observed radiance on Figure 14 correlate with the minor reflection in that spectral regions. Reflectivity behavior of various ground coverages are also one of the peak sources on the observed spectrum on Figure 14.

2.6. Prospects for space application

After numerous measurements in the IMF calibration laboratory, HySpex is a well characterized high precision instrument suited for reference analysis and feasibility studies for Earth observation applications. HySpex can be used for validation of measurements from novel remote sensing satellite missions over land and water. In particular this sensor is used to explore the potential of Earth observation applications for future German satellite mission EnMAP. In other words, HySpex plays a role as a simulator for the upcoming EnMAP mission. Also, third party users are interested in the HySpex data through the user service OpAiRS. This service allows users to get an access to the IMF expertise gathered for more than 20 years of airborne research from several sensors including HySpex. Moreover, the transnational access program of the European Facility for Airborne Research (EuFAR) includes HySpex instrument as one of its sensors.

3. Theory and Methods

3.1. Radiative transfer in the SWIR

The Planck function $B_\lambda(T)$ gives the radiance emitted by a blackbody at temperature T. Here are the expressions for the Planck function in wavelength (λ) and wavenumber (ν):

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5\left(\exp\left(hc/kT\lambda\right)-1\right)}$$

$$B_\nu(T) = \frac{2h\nu^3c^2}{\exp\left(h\nu c/kT\right)-1}$$

It is important to note, that the relation between the different expressions follows from ν:
$B_\lambda d\lambda = B_\nu d\nu$ with $\lambda = 1/\nu$.
The equation describing the observed radiance looks like:

$$\Im(\nu) = \frac{r}{\pi}\cdot\cos(SZA)\cdot\underbrace{\mathcal{T}_\uparrow(\nu)\mathcal{T}_\downarrow(\nu)}_{\mathcal{T}-transmission}\cdot\Im_{Sun}(\nu). \qquad (2)$$

In this equation wavenumber (ν), Earth's albedo (r), upwelling transmission from BoA to airborne observer $\left(\mathcal{T}_\uparrow(\nu)\right)$, downwelling transmission from ToA to BoA $\left(\mathcal{T}_\downarrow(\nu)\right)$ and the incoming solar radiation at the top of the atmosphere $\left(\Im_{Sun}(\nu)\right)$.
Incoming solar radiation can be represented in the following equation:

$$\Im_{Sun}(\nu) \approx B_\nu(T_{Sun}(K))\cdot\pi\left(\frac{R_{Sun}}{d_{Earth-Sun}}\right)^2$$

where the Planck function in wavenumber notation $\left(B_\nu(T_{Sun}(K))\right)$ and the term $\pi\left(\frac{R_{Sun}}{d_{Earth-Sun}}\right)^2$ stands for conversion between irradiance to radiance.
The first step of a simulation is to calculate the absorption cross section (xs) for a particular molecule m:

$$k_m(\nu,p,T) = \sum_l S_{ml}(T)\,g(\nu;\hat{\nu}_{ml},\gamma_{ml}(p,T)) \qquad (3)$$

where the wavenumber (ν), the sum comprises all relevant lines characterized by position $(\hat{\nu}_{ml})$, line strength $\left(S_{ml}\right)$ and broadening parameter $\left(\gamma_{ml}\right)$. The summation is done over all relevant lines with corresponding line strengths, wavenumbers and broadening parameters.

Generally speaking, the absorption cross section for a particular molecule is a measure of probability of an absorption process. It is an ability of molecule to absorb a photon of a particular wavelength.

The line strength $S(T)$ at temperature T can be described as follows:

$$S(T) = S(T_0)\frac{Q(T_0)}{Q(T)}\frac{\exp(-E_i/k_BT)}{\exp(-E_i/k_BT_0)}\frac{1-\exp(-hc\hat{\nu}/k_BT)}{1-\exp(-hc\hat{\nu}/k_BT_0)} \qquad (4)$$

where $S(T_0)$ – reference strength, E_i – lower state energy, $Q(T)$ – product of rotational and vibrational partition functions, c – speed of light, h – Planck constant and k_B – Boltzmann constant.

Absorption coefficient (*ac*) – is the sum of all molecular cross sections, scaled by the molecular number densities (n_m). In general (*ac*) depends on altitude *z* and the corresponding pressure and temperature. Absorption coefficient described by the following formula:

$$\alpha(\nu, z, p, T) = \sum_m n_m(z) k_m(\nu, p(z), T(z)) . \quad (5)$$

Optical depth (*od*) – refers to the mathematical operation of integration of the absorption coefficient (*ac*) along line of sight (LOS). If we consider a plane-parallel atmosphere and a vertical path from z_{low} to z_{high}:

$$\tau(\nu) = \int_{z_{low}}^{z_{high}} \alpha(\nu, z, p, T) dz . \quad (6)$$

Transmission $(\mathcal{T}(\nu))$ upwelling or downwelling, in general for slant paths can be calculated in the following way:

$$\mathcal{T}(\nu) = e^{-\tau(\nu)/\cos(\theta)} \quad (7)$$

where θ - can be Sun/Solar Zenith Angle (SZA) for the downwelling transmission or Observer Zenith Angle (OZA) for the upwelling transmission respectively.

Sun/Solar Zenith Angle (SZA) and Observer Zenith Angle (OZA) are important variables for radiative transfer. Both angles can be considered in the corresponding transmissions for slant paths and moreover SZA was also reckon in the equation describing observed radiance (Equation 2).

- ***Maximum radiance***:

Calculation of the maximum and low radiance implemented for the wavenumber 6000 cm^{-1}. Maximum radiance will take place, when we have weak absorption (small optical depth $(\tau = 0) \rightarrow e^{-\tau} = e^0 = 1$). It means that downwelling and upwelling transmissions are close to 1, hence their multiplication $\mathcal{T}_{\uparrow}(\nu)\mathcal{T}_{\downarrow}(\nu) \cong 1$. Maximum radiance corresponds to the highest albedo value (*r*=1). According to the [www.suncalc.org] the Sun/Solar Zenith Angle on the day of the measurements (7-th June 2018) over Upper-Silesian region around noon was about 30°. Hence for theoretical estimation we will assume $SZA = 30°$.

Sun radiance at the top of the atmosphere will be $\Im_{\text{Sun}} = 55.9\ \text{erg/s/(cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$. And the maximum radiance by Equation (2) will be equal to $\Im_{\text{max}} = 15.4\ \text{erg/s/(cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$.

- ***Low radiance***:

This case will take place, when we have strong absorption, or in other words large optical depth. The restriction that we have for the term $e^{-\tau}$ is: $0 \le e^{-\tau} \le 1$, then assuming $\tau = 1 \rightarrow e^{-\tau} = e^{-1} = 0.37$. We have both, upwelling and downwelling transmissions, hence we have to multiply transmissions for these two paths: $e^{-1} \cdot e^{-1} = 0.135$.

Low radiance calculated in the same manner as maximum and its value is equal to $\Im_{\text{low}} = 0.99\ \text{erg/s/(cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$.

3.2. Py4CAtS package

In the article 'Py4CAts – PYthon for Computational ATmospheric Spectroscopy', Schreier et al. (2019) introduce a Python re-implementation of the Fortran Generic Atmospheric Radiation Line-by-line Infrared Code (GARLIC). A good introduction of GARLIC implementation was done in the article 'GARLIC a general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation' by Schreier et al. (2014). The Py4CAts package allows to make individual steps of an infrared or microwave radiative transfer computation, which are implemented in separate scripts and corresponding functions.

Figure 16 shows the general working scheme of the Py4CAts package, including the following functions:

- **higstract**: extract (select) lines of relevant molecules in the spectral range of interest;
- **lbl2xs**: compute line-by-line (lbl) cross sections for given pressure(s) and temperature(s): Equation (3);
- **xs2ac**: multiply cross sections with number densities and sum over all molecules: Equation (5);
- **ac2od**: integrate absorption coefficients along the line-of-sight through the atmosphere to the vertical optical depth, Equation (6);
- **od2ri**: solve Schwarzschild equation, i.e., integrate the Planck function vs. optical depth along the line-of-sight through atmosphere (assuming a plane-parallel, non-scattering atmosphere in local thermal equilibrium). One should consider that this step does not needed in SWIR because thermal emission of surface and atmosphere is negligible. Instead of this function one can compute the radiance as described in Equation 2.

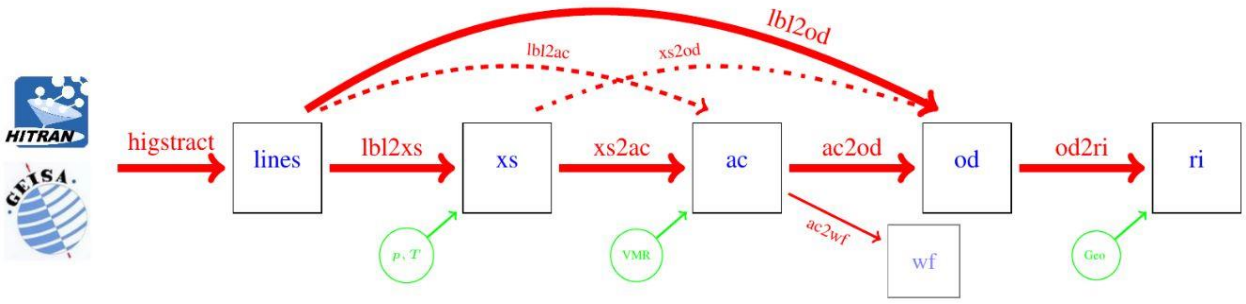


Figure 16 The general working scheme of the Py4CAtS package. [Schreier et al., 2019].

From HITRAN/GEISA databases via cross sections (*xs*), absorption coefficients (*ac*), optical depths (*od*) to radiation intensity (*ri*). It is important to note, that cross sections are temperature and pressure dependent, absorption coefficients depend on composition (VMR) and radiation intensity depends on the geometry of observations (Geo).

3.3. Methods description for detecting pixels with enhanced methane

Methane gas has less absorption abilities at short wave infrared (SWIR) range in comparison to carbon dioxide gas (Figure 4). Hence, its detection is a challenging task. In chapter 5 'Methane plume retrieval' the explanation of selection of particular wavenumber intervals with methane presence as well as selection of methane shafts and result graphs of four various methods are presented, while in this section an introduction to the physical processes standing behind those methods take place. The core idea of all four methods was to be able to 'see' or detect pixels with enhanced methane presence over a strong radiance from the ground.

A few common terms are actively used in the description of the methods, e.g. 'transparent interval' stands for the radiance from the wavenumber interval 6477-6557 cm^{-1}. According to Figure 4 one can see that the total transmission through the atmosphere is almost equal to 1 in this wavenumber window.

Term 'methane interval' stands for the radiance from the wavenumber interval 5991-6014 cm^{-1}. According to Figure 4 the radiance at this wavenumber interval stands for the strongest methane absorption at 1.6 μm.

Term 'spectrally averaged' means simply the calculation of the mean value of the radiance in the corresponding intervals, either methane interval or transparent atmosphere, e.g.:

```
#1.6 mue CH4 data lines (from HySpex Line 11):
data1= (paw11[:,:,117]) #5991-5992 cm-1 (left border of enhanced CH4)
data2= (paw11[:,:,116]) #6012-6014 cm-1 (right border of enhanced CH4)
mean1= (1/2)*(data1+data2)
```

In this code fragment *paw11* stands for the variable with the measurements from the HySpex ground coverage line 11, *[:, :, 117]* or *[:, :, 116]* corresponds to the all along-track pixels (1-st dimension), all across-track pixels (2-nd dimension) from the spectral channels 117 and 116 respectively. As it was already mentioned in section 2.4, 256 spectral channels correspond to the wavenumber or wavelength range covered by the HySpex instrument.

3.3.1. Method I. Difference of the radiances from methane and transparent intervals

The idea of the first method can be represented in the in following formula notation:

$$\Im \text{ (spectrally averaged } CH_4 \text{ interval)} - \Im \text{ (spectrally averaged transparent interval)}$$

The idea of subtraction can be explained as following: the spectral radiance from the ground presented in both wavenumber intervals: in methane and in transparent intervals. Hence, any change of the surface albedo, e.g. from conifer to grass etc. will have an impact on both windows. That means that subtraction should eliminate this common effect and might indicate a plume. It is important to note here that this method works in assumption that reflectivity or ground albedo values do not change dramatically from the methane to the transparent interval. According to Figure 15 this hypothesis works for the 'nearby' intervals, i.e. for methane interval at 5991-6014 cm^{-1} and transparent interval at 6477-6557 cm^{-1}.

3.3.2. Method II. Ratio of the radiances from methane and transparent intervals

The second method introduces the idea of the division of the radiances from spectrally averaged methane interval and spectrally averaged transparent atmosphere interval:

$$\frac{\Im \text{ (spectrally averaged } CH_4 \text{ interval)}}{\Im \text{ (spectrally averaged transparent interval)}}$$

The idea of the division can be explained in the similar manner with the same assumptions as a subtraction in the previous method. Due to the presence of the radiance from the ground in both intervals, the impact of changing surface albedo will be in both windows. The subtraction should eliminate this common effect and might indicate a plume.

3.3.3. Method III. Ratio of the radiances from methane and transparent intervals with 2D convolution

In the third method, similar idea of the ratios of the radiances as in the second method was applied. The only difference here was not only the spectral averaging of methane interval, as in the first two methods, but also spatial averaging. For the spatial averaging 2D convolution over the matrices for calculation of the element-wise mean values over 5×5 neighborhood pixels were implemented. In this method the following algorithm was applied to the two situations:

- Ratio of the spatially and spectrally averaged radiance from methane interval and only spectrally averaged radiance from transparent interval:

$$\frac{\Im\,(\text{spatially } [5\times 5 \text{ pixels}] \text{ and spectrally averaged } CH_4 \text{ interval})}{\Im\,(\text{spectrally averaged transparent interval})}$$

- Ratio of the spatially and spectrally averaged radiance from methane interval and only spectrally averaged radiance from methane interval:

$$\frac{\Im\,(\text{spatially } [5\times 5 \text{ pixels}] \text{ and spectrally averaged } CH_4 \text{ interval})}{\Im\,(\text{spectrally averaged } CH_4 \text{ interval})}$$

The idea behind these two slightly different approaches is to avoid possible spatial ‘artefacts’ or outliers in the radiance appearing at some pixels. For this purpose, 2D convolution of the element-wise average calculation through the 5×5 neighbor pixels was implemented. This method was also applied with 3×3 neighbor pixels averaging, but the results with the larger mask (5×5) are more pronounced and contrasting.

3.3.4. Method IV. Ratio of absorbing and transparent residuals

The fourth method also exploits the idea of the calculation of the ratio of the radiances, but here the main target is to separate slowly varying parts of the absorption that are caused by ground albedo or aerosol extinction from molecular absorption cross sections that vary rapidly with wavenumber.

A good visualization of the theory standing behind this method could be Figure 17, which is the fragment from the HySpex modelled spectra from dry grass empirical albedo with normal and enhanced methane concentration in the atmosphere with three randomly chosen samples of the measurements. The significant difference in this particular method is that methane interval was extended to 5950-6050 cm^{-1} (yellow area) and two groups of data points were processed separately.

The first group of ‘inner’ data points includes the data which is the closest to the strong absorption of methane (data points marked by green rectangle).

The second group of the points – ‘outer’ data points are corresponding to the small absorption. At this example ‘outer’ data points are located on the right and left from the inner points, highlighted by blue circles, but this is not necessary to be the case.

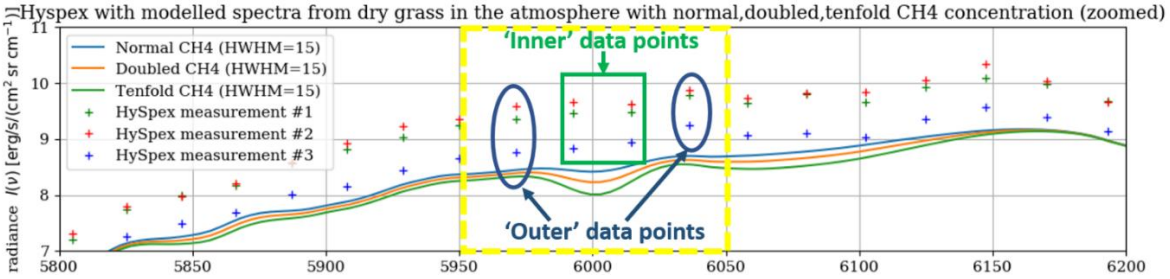


Figure 17 Visualization of method IV theory on the graph with the radiance from dry grass (empirical albedo) with normal, doubled and tenfold methane concentration in the atmosphere (solid lines) with three randomly chosen samples of HySpex measurements (crosses).

The main idea of this method is to create a zero order polynomial *p(y)*, by using least-squares, for the outer group of data points and then plot the ratio of the residuals from this polynomial to the inner and outer group of data points (green and blue dotted lines):

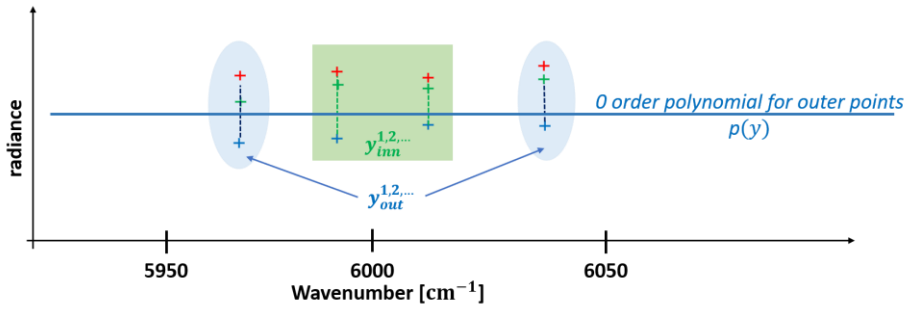


Figure 18 Simplified schematic visualization of three randomly chosen samples of HySpex measurements and polynomial for the outer points (blue line) with the corresponding residuals to it from inner points (green dashed lines) and outer points (blue dashed lines).

In order to catch most of the low frequency components in the spectrum a zero degree polynomial is fitted to pixels where no molecular absorption is taking place. The L2-norm of this 'outer' data points polynomial fit (which is the least squares minimum) represents the denominator in the ratio (see below).

In a next step the numerator is determined based on the L2-norm of the so called 'inner' data points (pixels where molecular absorption takes place) to that polynomial. For background scenes (no plume) the output ratio will be rather constant while for different concentrations of the target molecule the ratio will become larger due to increased norm of the residuals of the numerator and hence indicate the plume by its enhanced absorption.

In the formula notation the idea of this method can be represented as following:
Residuals (z) of the difference between polynomial $p(y)$ and inner data points $y_{inn}^{1,2,\dots}$ as well as the residuals (e) of the difference between polynomial $p(y)$ and outer data points $y_{out}^{1,2,\dots}$ can be calculated as following:

$$z = \sqrt{(p(y) - y^{1}_{inn}) + (p(y) - y^{2}_{inn}) + \dots}$$

$$e = \sqrt{(p(y) - y^{1}_{out}) + (p(y) - y^{2}_{out}) + \dots}$$

With a general comparison between these two residuals z and e, residuals between polynomial and inner data points *(z)* is bigger than the residuals between polynomial and outer data points *(e)* since the polynomial was created based on outer data points.

The final result of this method calculated as the ratio of the inner residuals *(z)* to the outer residuals *(e)*:

$$result = \frac{z}{e}.$$

3.4. Least-squares problems

A good introduction to the least-squares problems was done by Kahaner et al. (1989). Lets suppose the function $b(t)$ represented by given data (t_i, b_i), $i = 1,...,m$ such that $b_i = b(t_i)$. The authors assume that some model for the data is given and observation can be represented as sum of model and error. Since there is a minimization procedure of an error, one can use term optimization. Optimization denotes either minimization or maximization of a function and take place in all least-squares problems.

3.4.1. Linear least-squares problems

Assume that the model for the data has the form: $b_i \approx x_1\phi_1(t_i) + x_2\phi_2(t_i) + ... + x_n\phi_n(t_i)$, where $\phi_j(t)$ - given model functions and x_j -parameters of the model. This model is called linear because it consists from a linear combination of the model functions. However, the model functions by itself can be nonlinear. In some cases, more than one independent variable can be involved into the model.

By introducing the matrix consisting of model functions as $A_{ij} = \phi_j(t_i)$, the model for the data could be written in the form: $b \approx Ax$ or $b - Ax \approx 0$, where the left part, difference – is called vector of residuals. The target is to find parameters of the model that will make the residuals as small as possible. A common approach for this is solving $\min_x \sum_{j=1}^{m} \left[(b - Ax)_j \right]^2$. This is called least squares data fitting since we are minimizing the sum of squares. On Figure 19 it is depicted graphical interpretation of the least-squares techniques, as minimization of the vertical distances from the data points to the model. This is based on the assumption that all errors in the approximation correspond to errors in the observations.

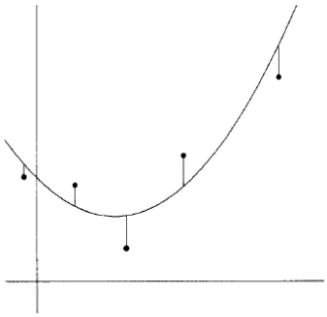


Figure 19 Graphical interpretation of the least-squares techniques – minimization of vertical distances from data points to the model.

3.4.2. Nonlinear least-squares problems

Nonlinear least-squares (NLS) problems have a model for the data $b(t)$ as a mixture of linear and nonlinear parameters. A common example for NLS problem with two parameters could be the following: $\min_{x_1,x_2} F(x_1,x_2) \equiv \sum_{j=1}^{m}\left[b(t_j) - x_1 e^{x_2 t_j}\right]^2$, where the model is a linear function of the parameter x_1, but nonlinear of the parameter x_2. System of nonlinear equations in the variables *x*:

$$\begin{cases} \dfrac{\partial}{\partial x_1} F\left(x_1, x_2, ..., x_n\right)\Big|_{x=x^*} = 0, \\ \dfrac{\partial}{\partial x_2} F\left(x_1, x_2, ..., x_n\right)\Big|_{x=x^*} = 0, \\ \vdots \\ \dfrac{\partial}{\partial x_n} F\left(x_1, x_2, ..., x_n\right)\Big|_{x=x^*} = 0. \end{cases}$$

shows that first derivatives of the objective function *F(x)* are zero in case if *x** minimizes *F(x)* and there are no constraints. As it was mentioned in Kahaner et al. (1989), techniques for solving nonlinear equations are similar to the techniques for solving optimization problems. However, despite of techniques for solving nonlinear equations, optimization method can guarantee convergence to a local minimum (maximum), moreover optimization can provide savings in memory storage and arithmetic. Finally, in case of constraints on the variables *x*, it is harder to manage with them in an algorithm for nonlinear equations in comparison with unconstrained optimization procedure.

There are many different optimizations methods, starting from one-dimensional optimizations, such as Quasi-Newton methods, Fibonacci search etc. and multi-dimensional optimizations i.e. method of steepest descent. Detailed description of all these methods is beyond the scope of this work.

3.4.3. Separable nonlinear least-squares problem

Separable nonlinear least-squares (SNLS) problem is a special class of nonlinear least-squares (NLS) problems. Kaufman et al. (1978) summarized the procedure of solving SNLS problems in the following way: use the explicit coupling between the linear and nonlinear variables in order to define a new NLS problem in the nonlinear variables only whose solution is the solution to the original problem. According to Gharibi et al. (2011) SNLS is widely used in areas such as environmental sciences, neural networks, numerical analysis, mechanical systems etc. Separable nonlinear least-squares (SNLS) problem implemented in the Beer InfraRed Retrieval Algorithm (BIRRA) which is introduced in the next section.

3.5. BIRRA as inverse problem solver

So far, we operate the term of forward model, where the spectra, observed at the sensor, retrieved from atmospheric parameters. Based on this approach theoretical values (maximum and minimum) of modelled radiance where calculated in section 3.1 and practical results, modelled spectrums, based on forward model are presented in the chapter 4 “Py4CatS results”.

Beer InfraRed Retrieval Algorithm works on principle of inverse model, according to Hochstaffl et al. (2018) inverse problem is the retrieval of trace gas concentrations or temperatures from radiance or transmission spectra. This problem is typically solved by least-squares algorithms, which were introduced in the section 3.4, in order to provide more accurate results without rough assumptions.

BIRRA which can only operate in nadir geometries, consists of GARLIC (Generic Atmospheric Radiation Line-by-line Infrared Code) which can operate in all geometries, i.e. nadir, limb etc. on forward model and coupled into a least-squares inversion algorithm for retrieval of trace gases in the SWIR interval.

Due to the fact that BIRRA refers to the SNLS problem, here the state vector $x = (\eta, \beta)$ consists of nonlinear (η) and linear (β) parameters. This separation is important because it allows to implement a separable least-squares fit in order to estimate unknown variables. These unknown variables include: molecular scaling factor(s), half width at half maximum (HWHM), the wavelength shift of the instrumental slit function, surface albedo (with linear interpolated coefficients) and the optional baseline correction. The model for the least-squares problem looks like: $y_i = \sum_{j=1}^{n} \phi_{i,j}(\eta)\beta_j \qquad i = 1, 2, ..., m$

Where *n* – the number of linear parameters, *m* – the number of observations (spectral pixels) in the window chosen for retrieval. It is important to note that in separable least-squares approach model functions $\phi_{i,j}(\eta)$ depend only on the nonlinear parameters (η). For any given nonlinear parameter η we can define a matrix $\Phi(\eta)$ that consists of model functions $\phi_{i,j}(\eta)$:

$$\Phi(\eta) = \begin{pmatrix} \phi_{1,1}(\eta) & \phi_{1,2}(\eta) & \cdots & \phi_{1,n}(\eta) \\ \phi_{2,1}(\eta) & \phi_{2,2}(\eta) & \cdots & \phi_{2,n}(\eta) \\ \vdots & \vdots & \therefore & \vdots \\ \phi_{m,j}(\eta) & \phi_{m,2}(\eta) & \cdots & \phi_{m,n}(\eta) \end{pmatrix}$$

In order to find the result as good as possible we have to minimize through both variables η and β, in other words, the difference between model function computed by GARLIC and measurement vector from HySpex sensor. Minimization through the β (inner minimization) is a least-squares problem: $L(\hat{\beta}, \hat{\eta}) = \min_{\eta} \left\{ \min_{\beta} \left\{ \left\| y - \Phi(\eta)\beta \right\|^2 \right\} \right\}$. Since the fact that only nonlinear parameters η required in prior knowledge (initial estimate), the overall minimization problem, described above, can be represented only in terms of nonlinear parameters: $L(\hat{\beta}, \hat{\eta}) = \min_{\eta} \left\{ \left\| y - \Phi(\eta)\Phi^{\dagger}(\eta)y \right\|^2 \right\}$, where $\Phi^{\dagger}(\eta)$ stands for generalized inverse: $\Phi^{\dagger}(\eta) = \left[\Phi^T(\eta)\Phi(\eta) \right]^{-1} \Phi^T(\eta)$.

Linear parameters β are not required in the initial estimate hence the Jacobian matrix is reduced which improves the computing times.

Prior information, e. g. scale factors of the optical depth of the CH_4 molecule can be taken from the for ex. Air Force Geophysical Laboratory (AFGL) atmospheric constituent profiles by Anderson et al. (1986).

4. Py4CAtS results

This chapter present results of different simulations with the Py4CAts package exploiting the data from HITRAN (High Resolution Transmission database) mostly from the version 2008. According to the information from official database website [www.hitran.org], HITRAN is a collection of spectroscopic parameters that different computer codes can use for predictions and simulations of the transmission and emission of light in the atmosphere. HITRAN is a long-running project that started in the late 1960s by the Air Force Cambridge Research Laboratories (AFCRL) for the purpose of detailed knowledge of the infrared properties of the atmosphere.

The HITRAN database includes both theoretical and experimental values for millions of transitions for different molecules. The data that are available from HITRAN include: line-by-line, absorption cross sections, collision induced absorption, additional data that are used in radiative transfer calculations and many other molecule parameters.

The HITRAN database is updated every 4 years by adding the data of new molecules that were not yet considered, adding the data of weak lines that were not considered for existing molecules, improving the existing data by more accurate values etc. In this chapter the result graphs were plotted based on mainly HITRAN 2008, but few graphs show the comparison of HITRAN 1986, HITRAN 2000 and HITRAN 2008. There is a dramatic difference between HITRAN'86 and HITRAN 2008, e.g. HITRAN 1986 database has information on 28 molecules with around 300.000 lines (each line represents a transition between different molecules states, e.g during emission or absorption of photon), while HITRAN 2008 database has information on 42 molecules with 2.713.968 lines.

Focusing in the vicinity of 1.6 μm interval (~5400 - 7000 cm^{-1}), HITRAN'86 has 7818 lines of 5 main molecules ($H_2O, CO_2, CH_4, CO, O_2$), while HITRAN 2008 has 53.571 lines of 7 molecules (molecules from HITRAN'86 + O_3, N_2O).

4.1. Preliminary simulations

After extracting lines of molecules (CO_2, CH_4, H_2O) in the SWIR spectral range, Py4CAtS was used for displaying line strength (Equation 4) vs. position by default. In the legend of the graph, as a third number, one can also see the number of lines for each of the molecule.

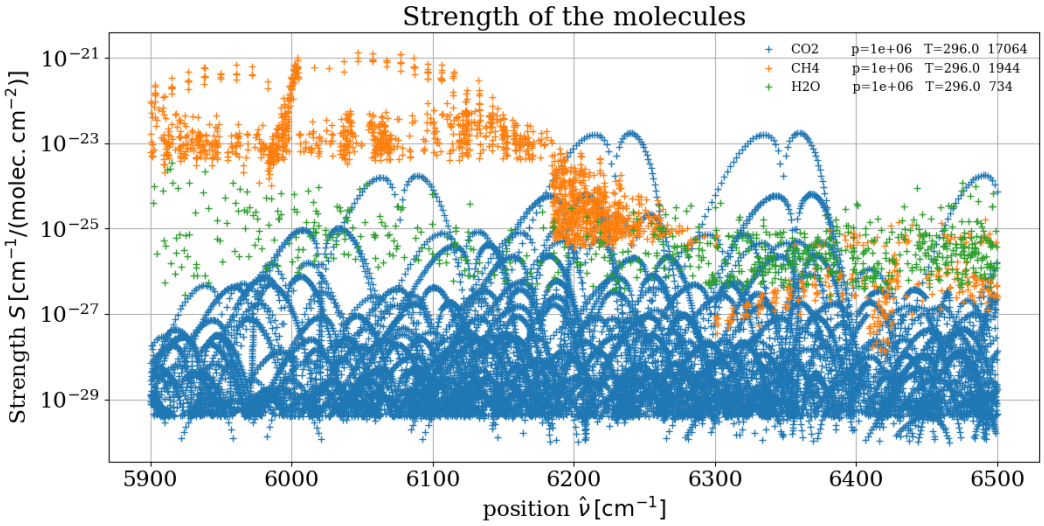


Figure 20 Strength of molecules CO_2, CH_4, H_2O in the SWIR range, HITRAN 2008.

By default, cross sections will be computed for the reference pressure and temperature that are used in the line parameter database, e.g. $p_0 = 1013.25\,\text{mb}$ and $T_0 = 296\,\text{K}$. It is also possible to use ready-made files from Air Force Geophysical Laboratory (AFGL) datasets with different types of atmosphere, e.g. midlatitude summer (MLS) or midlatitude winter (MLW) etc. A detailed comparison of different atmospheres is presented in the section 4.3.

Figure 21 demonstrates computed molecular cross sections for given temperature and pressures (Equation 3).

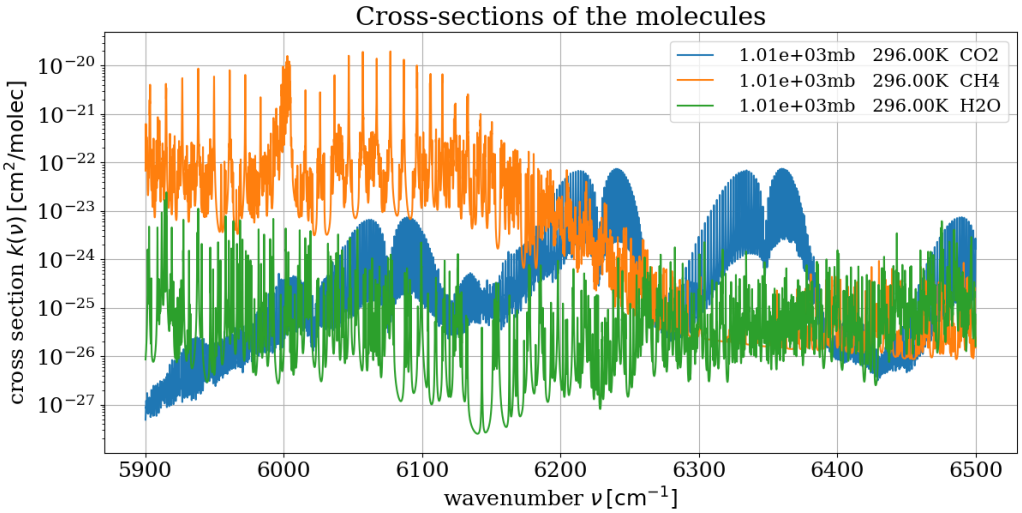


Figure 21 Cross sections of molecules CO_2, CH_4, H_2O in SWIR range at the BoA, HITRAN 2008.

Figure 21 depicts, that the first level of BoA, the maximum of cross section of the molecule CH_4 (orange curve) is in the 6000 – 6100 cm^{-1} interval. Water molecule H_2O (green curve) compared to CO_2 (blue curve) and CH_4 does not clearly pronounced with strong band.

After calculation of the absorption coefficients (Equation 5) it is possible to find optical depths (Equation 6) as an integration of the absorption coefficients along the path through the atmosphere. Figure 22 demonstrates the summation of all layers optical depths or in other words total transmission of the atmosphere up to 80 km (Equation 7). Detailed explanation of selecting exactly 80 km as the top of the atmosphere (ToA) will be done in the section 4.3. On this figure one can also see the comparison of the monochromatic transmission and the transmission convolved with box and Gaussian functions, both HWHM = 1.0 cm^{-1}. Convolution operation is a mathematical operation on two functions that creates a third function expressing how the shape of one function modified by anotheone. Convolution is widely used in this work for more general and clarity view of particular functions.

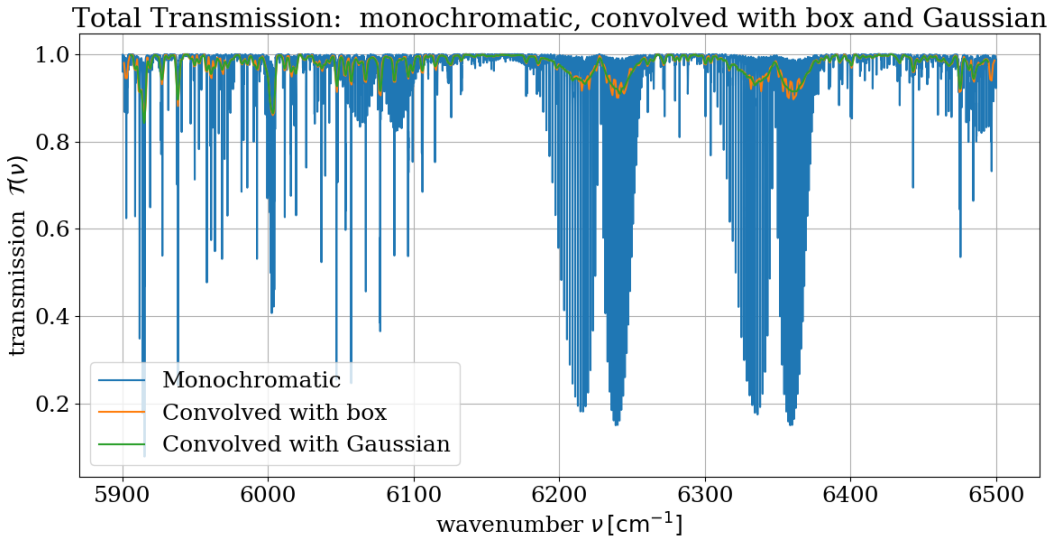


Figure 22 Total transmission of the molecules: monochromatic, convolved with box and Gaussian functions, both HWHM = 1.0 cm^{-1}, HITRAN 2008

4.2. Comparison of total transmissions for different HITRAN versions

As it was already introduced at the beginning of chapter 4, there is a dramatic difference between different HITRAN database versions, e.g. HITRAN'86 and HITRAN 2008. The following graph depicts the total transmission, convolved with Gaussian function (HWHM = 1.0 cm^{-1}) for clarity, through the whole atmosphere up to 80 km at midlatitude summer atmosphere from different HITRAN versions.

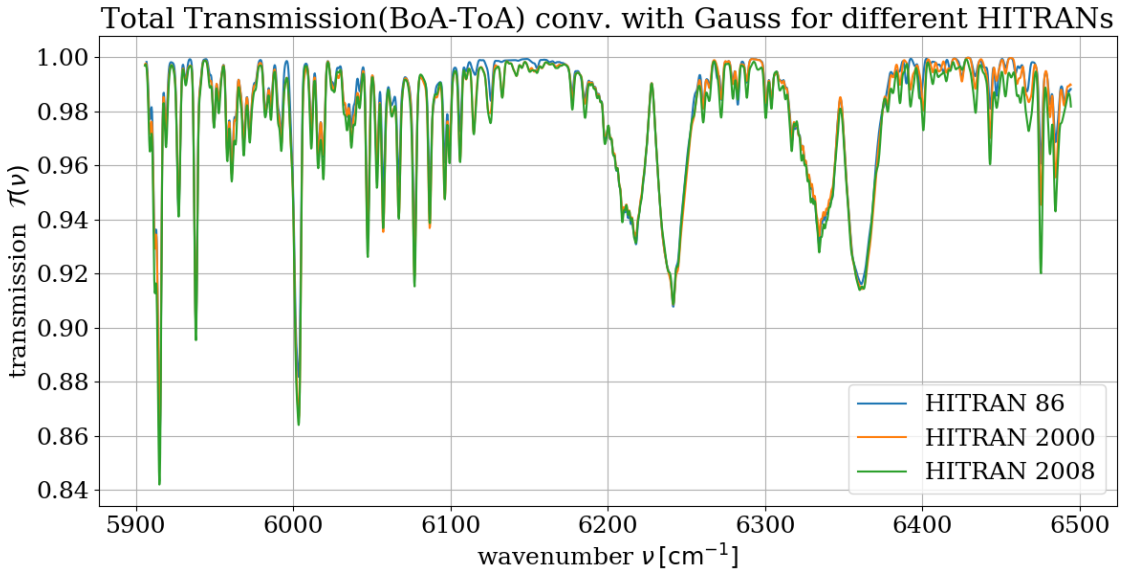


Figure 23 Comparison of total transmission (BoA-ToA) convolved with Gaussian function, HWHM = 1.0 cm^{-1} from different versions of HITRAN.

As it can be clearly seen from Figure 23, total transmissions from different HITRAN versions exactly at the observed wavenumber interval are almost identical, however for more accurate and precise results the newer version of the database (HITRAN 2008) were used in further simulations.

4.3. Impact of the atmosphere

According to the general scheme of the measurement process, the light that is coming from the Sun goes through the whole atmosphere, partly reflected by Earth's surface, partly absorbed by particular gases and then detected by HySpex airborne sensor at altitude (~1500 or 2900 m). Since the first part of its way the light is going through the whole atmosphere it is important to know the ToA altitude or in other words at which altitude the absorption of the incoming light is so neglectable small and hence transmission is almost equal to 1. Figure 24 demonstrates the comparison of the transmissions from different altitudes up to 100 km, convolved with Gaussian function (HWHM = 10.0 cm^{-1}) for clarity at midlatitude summer atmosphere.

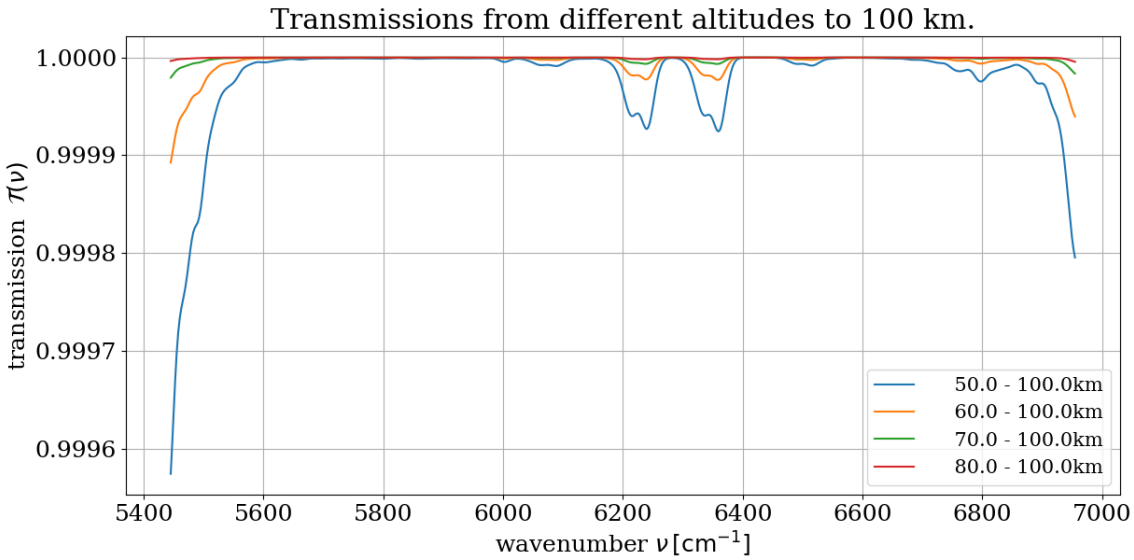


Figure 24 The difference between transmissions from the different altitudes to 100 km in MLS, convolved with Gaussian function (HWHM =10.0 cm^{-1}).

Figure 24 clearly shows that the transmission between the altitudes 80-100 km is almost 1. Hence the last layer, which can influence on the transmission is 80 km and therefore we can select it as a ToA. It is important to note that this value for ToA is only valid for 1.6 μm spectral interval and a single vertical path geometry. For other geometries, for ex. slant path, when the Sun is slightly above horizon, the value for ToA might differ.

As it was already mentioned in previous sections, the selection of the proper seasonal and geographical type of the atmosphere is an important step which should not be forgotten, since it is directly influencing on the accuracy of final results. Following Figure 25 demonstrates the comparison the optical depths on the first layers of the atmosphere for midlatitude summer (MLS) and midlatitude winter (MLW), convolved with Gaussian function (HWHM = 5.0 cm^{-1}) for clarity.

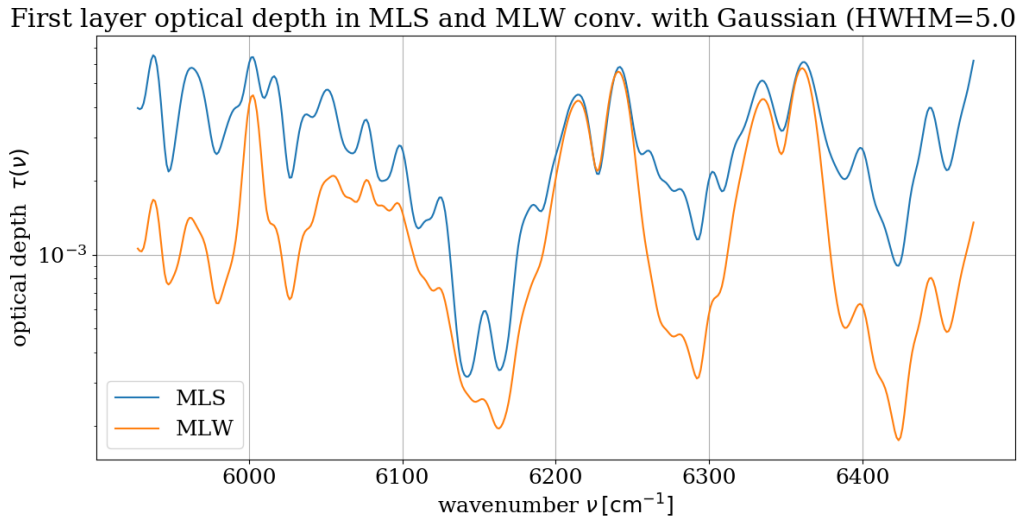


Figure 25 First layer of optical depth (*od*) in midlatitude summer (MLS) and midlatitude winter (MLW), convolved with Gaussian function (HWHM = 5.0 cm^{-1}).

Figure 25 shows, that there is a difference in optical depths. It is important to note that there is a small difference in pressure in summer and winter. However, the difference in humidity in turn is due to the difference in the

temperature in summer and winter. Winter atmosphere is much more humid than summer one in the mid-latitudes. Since the HySpex measurements that are used in this work, were done in June, midlatitude summer atmosphere was selected for further simulations.

4.4. Interpolation of albedo

Term interpolation in this section explains the following idea. Typically, albedo is a smooth function of the wavelength or wavenumber or in other words it does not change significantly with wavelength or wavenumber. While molecular absorption changes rapidly with wavelength or wavenumber. Hence empirical albedo data are usually given on a coarse wavenumber grid, whereas molecular absorption spectra are defined on thousands or millions of wavelength or wavenumber grid points. In order to 'combine' these data: empirical albedo and molecular absorption spectra, both have to be given on the same grid of wavelength or wavenumber. Therefore, interpolation of albedo is mandatory.

In order to understand and be able to describe accurately HySpex observed spectra (Figure 14) the knowledge of surface albedo is crucial. The importance of the albedo can be demonstrated on Figure 26 that shows the radiances of the main molecules without albedo.

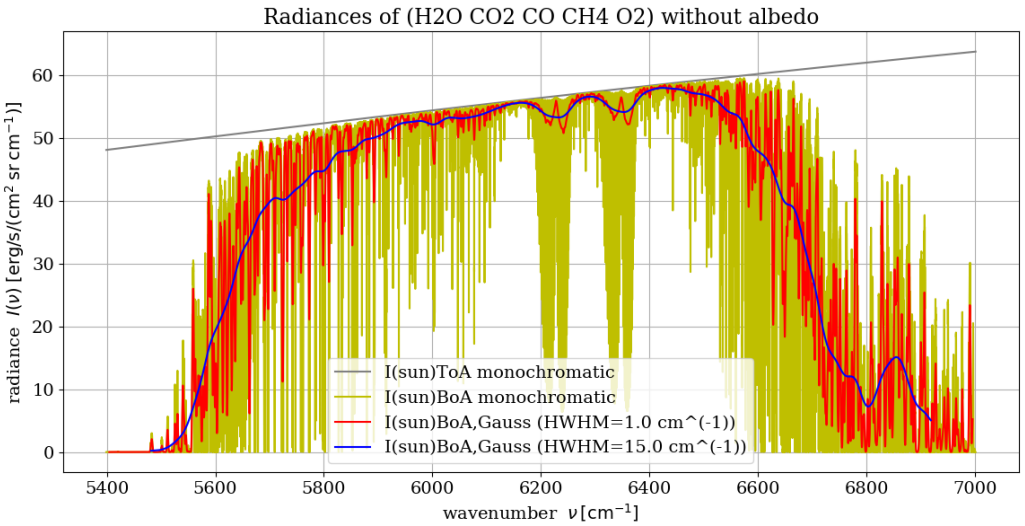


Figure 26 Radiances of the main molecules without albedo, MLS, HITRAN 2008.

On Figure 26 one can see the downwelling radiation from Sun to Earth's surface and maximum of simulated radiance in case of 'ideal' albedo (r=1) ~50-55 erg/s/(cm$^2 \cdot$sr$\cdot$cm^{-1}) is much higher than theoretically calculated value of the maximum radiance at section 3.1, which was equal to $\Im_{max} = 15.4$ erg/s/(cm$^2 \cdot$sr$\cdot$cm^{-1}). Hence, the importance of the ground albedo can not be ignored.

HySpex 2018 campaign in the Upper Silesian region (Poland) was looking for the CH_4 emissions from coal mines. As far as we do not exactly know the type of surface coverage, whether it is only "artificial", i.e. concrete, metal etc. or only "natural", i.e. grass, different types of forests etc. or a mix of both "artificial" and "natural". In order clarify this topic we have to take the interpolated empirical albedo data into account and then compare results for different types of ground coverage with the HySpex measurements.

After knowing empirical ground albedo values at the wavenumbers of interest interval (Figure 15), it is possible to calculate theoretical radiances that can be detected by the HySpex airborne sensor from different types of the ground coverage.

Figure 27 demonstrates modelled radiances (solid lines), calculated by Equation 2, with empirical albedos from conifer, deciduous, dry grass and grass and the modelled radiance with constant albedo (r=0.5) that are seen at the observer located at the altitude 1500m. On this figure one can also see three randomly chosen samples of HySpex measurements (crosses). The modelled radiances are convolved with Gaussian function (HWHM = 15.0 cm^{-1}) in order to be in almost the same resolution level as HySpex SWIR-320m-e camera (HWHM = 18.5 cm^{-1}) from the section 2.4.

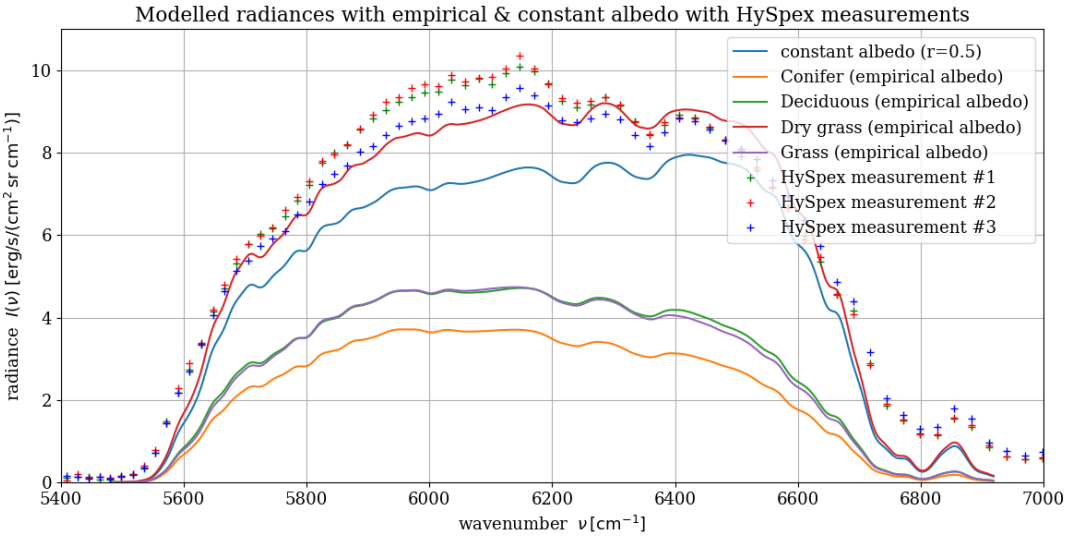


Figure 27 Comparison of the radiances in case of empirical albedo and constant albedo (r =0.5) with 3 HySpex randomly chosen samples of measurements. MLS, all functions convolved with Gaussian function (HWHM = 15 cm^{-1}), HITRAN 2008.

As it is clearly seen on Figure 27, from the listed types of ground coverage, the radiance flux from the dry grass with empirical albedo is the closest to these particular randomly chosen samples of the HySpex measurements.

Following Figure 28 demonstrates the difference between different resolution levels. Simulated theoretical radiance from dry grass ground coverage from previous graph plotted in three resolutions: monochromatic signal, medium resolution (HWHM = 1.0 cm^{-1}) and coarse resolution (HWHM = 15.0 cm^{-1}), which is the closest to the real HySpex SWIR-320m-e camera resolution.

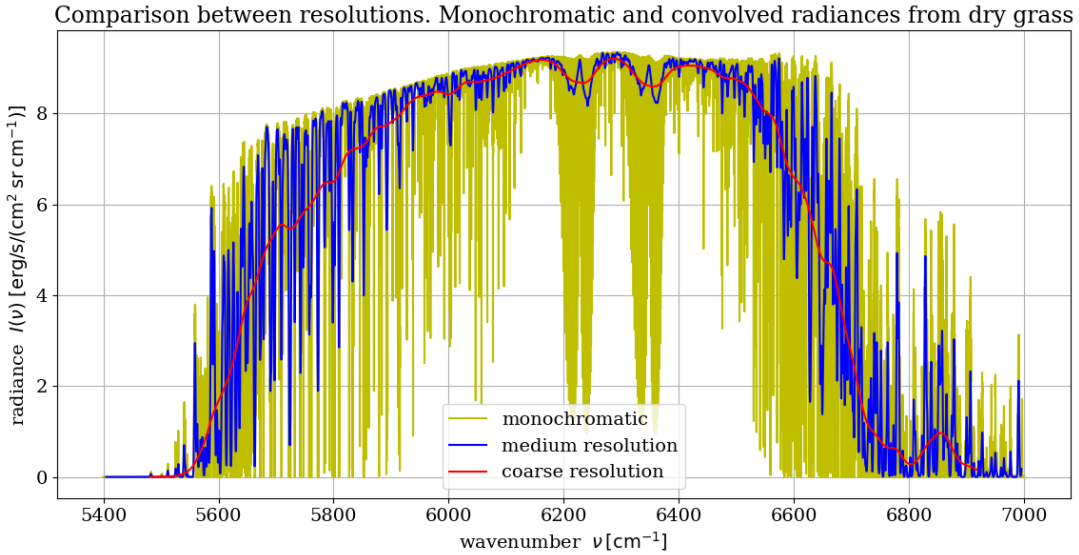


Figure 28 Comparison of the monochromatic and convolved with Gaussian function radiances seen at the aircraft (1500 m) from the dry grass (empirical albedo). Medium resolution stands for HWHM = 1 cm^{-1} and coarse resolution for HWHM = 15 cm^{-1}, MLS, HITRAN 2008.

Following Figure 29 demonstrates comparison between modelled spectrums with dry grass empirical albedo from different HITRAN datasets versions together with three randomly chosen HySpex measurements.

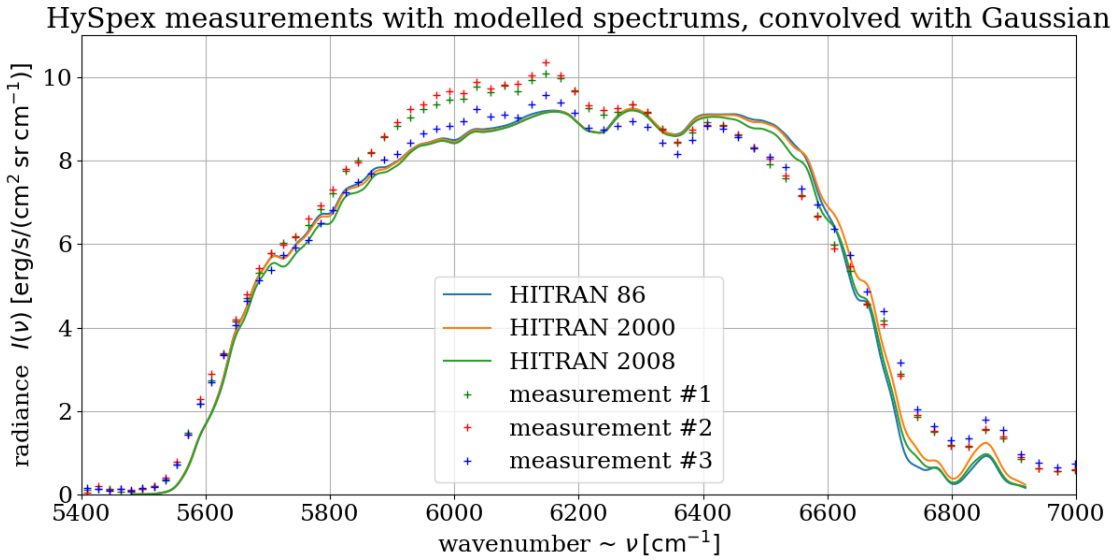


Figure 29 Three HySpex randomly chosen samples of measurements with modelled spectrums from dry grass (empirical albedo) from different versions of HITRAN, convolved with Gaussian function (HWHM=15 cm^{-1}).

Figure 29 shows that there is no significant difference in the modelled spectrums with empirical albedo from dry grass depending on various versions of HITRAN databases, particularly in the vicinity of 1.6 μm interval. However, for more accurate simulations HITRAN 2008 version will be exploited in further simulations with enhanced concentrations of methane in the atmosphere.

4.5. Normal and enhanced methane concentration in the atmosphere

This section demonstrates results of the simulations of the observing radiances in case of normal and enhanced (doubled and tenfold) methane concentrations in the atmosphere. Again, similar to the previous graphs, dry grass albedo was selected for this modelling radiances detected at 1500 m altitude. Simulated radiances here are also convolved with a Gaussian function (HWHM=15 cm^{-1}) and compared with three randomly chosen samples of the HySpex measurements.

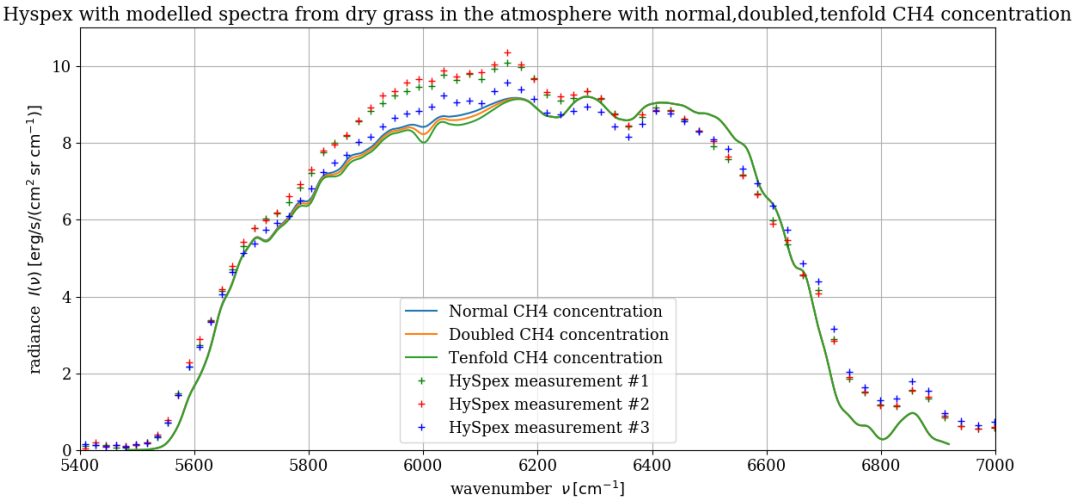


Figure 30 Modelled spectra from the dry grass empirical albedo in atmosphere with normal, doubled and tenfold concentration of methane with 3 HySpex randomly chosen samples of measurements. MLS, all functions convolved with Gaussian function (HWHM = 15 cm^{-1}), HITRAN 2008.

According to figure above one can only detect the difference between three kind of atmospheres at relatively small wavenumber 'window' ~ 5900-6100 cm^{-1}.

Following Figure 31 shows 'zoomed' at 6000 cm^{-1} area of the simulated radiances with coarse (HWHM = 15 cm^{-1}) and medium (HWHM = 1.0 cm^{-1}) resolutions.

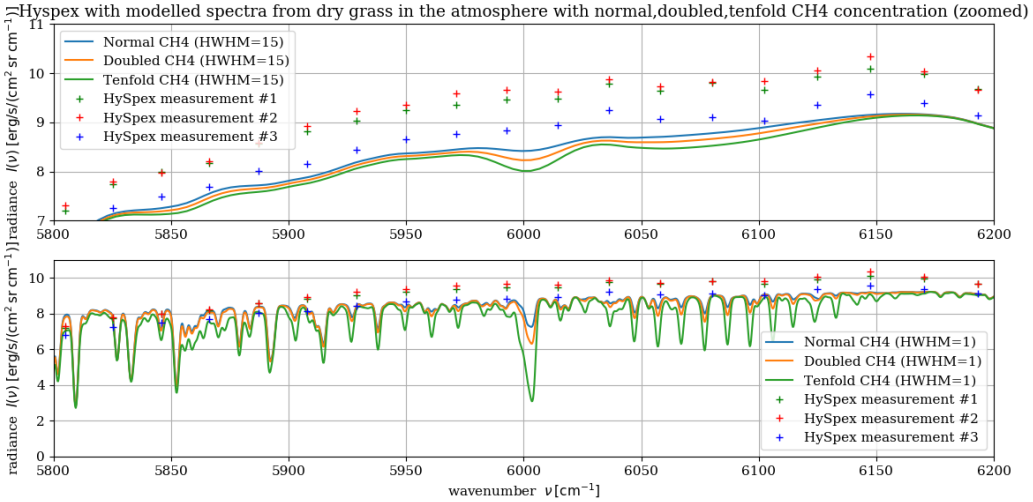


Figure 31 Radiance from the dry grass (empirical albedo) in atmosphere with normal, doubled and tenfold concentration of methane with 3 HySpex randomly chosen samples of measurements. MLS (zoomed interval 5800-6200 cm^{-1}). Upper graph: all functions convolved with Gaussian function (HWHM = 15 cm^{-1}); Lower graph: all functions convolved with Gaussian function (HWHM = 1 cm^{-1}), HITRAN 2008.

It is obvious that the absorption at roughly 6000 cm^{-1} looks more noticeable at better resolution (lower graph) than on the coarse resolution (upper graph). However, it is worth noting that HySpex SWIR-320m-e camera resolution even a bit worse (HWHM = 18.5 cm^{-1}) than depicted on the upper graph.

4.6. Downwelling, upwelling and effective radiation

Following figure depicts three different radiations, convolved with Gaussian function (HWHM=1.0 cm^{-1}) for clarity. In general upwelling radiation has four components:

- The downwelling (solar) radiation, attenuated along the path from top of the atmosphere (ToA) to bottom of the atmosphere (BoA) and reflected at the surface.
- The thermal emission of the atmosphere downwelling to the Earth's surface and reflected towards the airborne observer, i.e. the atmospheric layers are also black bodies.
- The thermal emission of the surface, e.g. the surface as a blackbody.
- The thermal emission of the atmosphere between Earth's surface and the airborne observer which is also described as a blackbody. The scattering effect has been neglected at this particular component.

In the short-wave infrared (SWIR) or near infrared (NIR) contributions by thermal emission can be neglected. Upwelling radiation (orange curve) on Figure 32 include only thermal emissions and hence equal to zero. Effective radiation is the radiation seen by airborne observer at certain altitude. This radiation is calculated by Equation 2, by taking into account empirical albedo of dry grass.

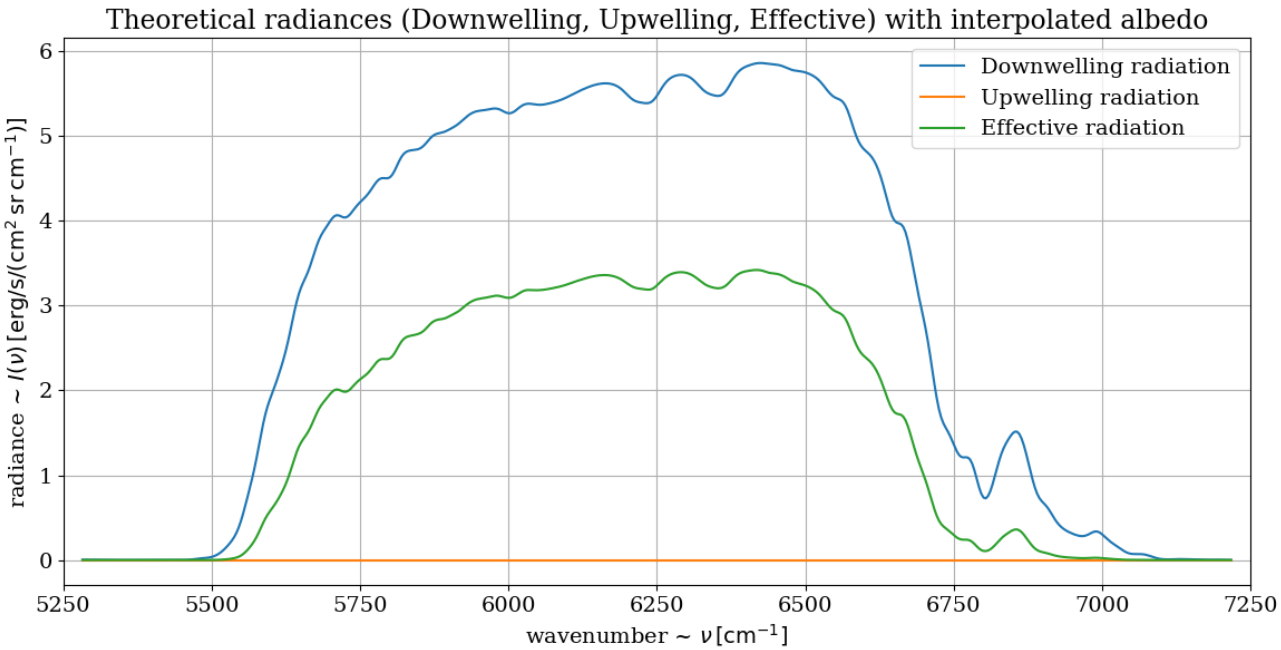


Figure 32 Theoretical radiances: downwelling, upwelling and effective with interpolated (empirical) dry grass albedo, all functions convolved with Gaussian function (HWHM = 1 cm^{-1}), HITRAN 2008.

Next figure demonstrates the difference between upwelling radiances detected by airborne observer at the altitude 1500 m with different observer zenith angles (OZA) in the midlatitude summer and winter atmospheres.

According to Figure 10 OZA = 180° stands for nadir viewing sensor. It is important to note, that the Sun Zenith Angle (SZA) for the Upper Silesian region in June is approximately equal 30 degrees.

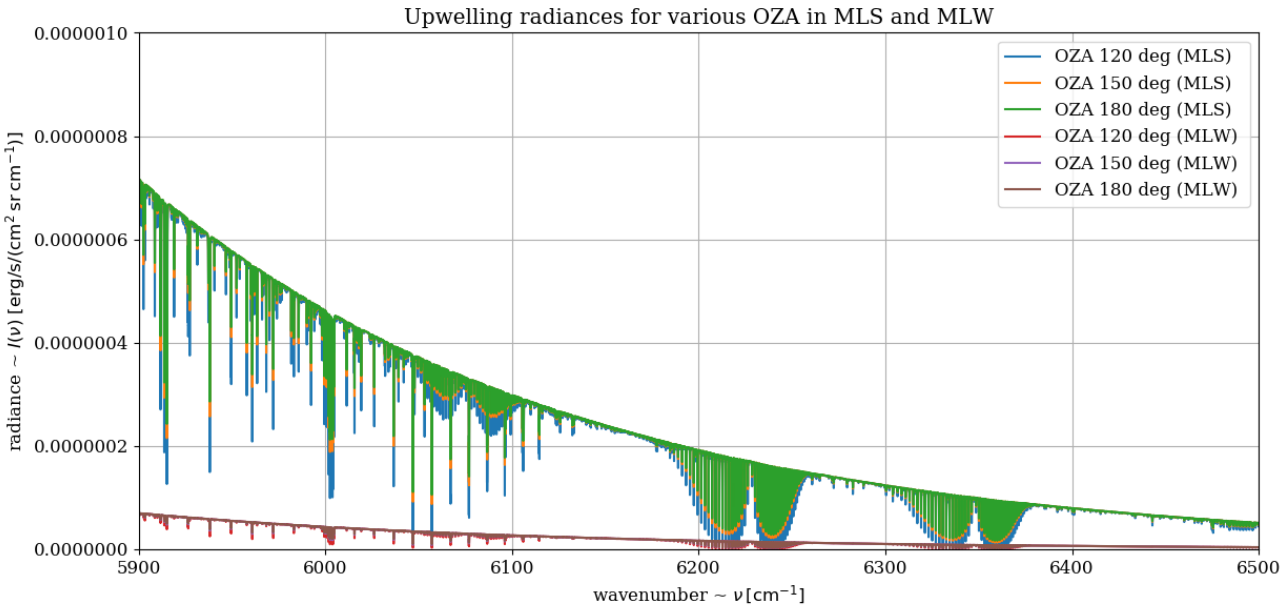


Figure 33 Upwelling radiances for different Observer Zenith Angles (OZA) in midlatitude summer and winter (MLS and MLW).

Figure 33 shows that upwelling radiance in different OZA are significantly different during midlatitude summer and winter with a minor difference in the observing angles at one atmosphere type.

5. Methane plume retrieval

The first step in successful methane plume retrieval is an accurate and faithful representation of the wavenumber intervals, where methane is represented in the atmosphere. HySpex instrument covers the wavenumber interval from 4000 cm^{-1} to 10275 cm^{-1}. Through this extensive wavenumber coverage, one has to select specific wavenumber intervals with the presence of the methane. The reasoning for choosing the specific wavenumber intervals is presented in the next section. Also in this chapter presented the arguments for the selection of particular methane shaft for plume retrieval as well as four methods for its detection.

5.1. Wavenumber intervals with methane (2.3 μm and 1.6 μm)

According to Figure 4 in the short wave infrared (SWIR) interval that covered by the HySpex instrument, one can see noticeable methane transmission descent through the atmosphere in the vicinity of 2.3 μm(~4347 cm^{-1}) and minor absorption of this gas at around 1.6 μm (~6000-6250 cm^{-1}).

Another critical point for successful methane detection is the knowledge about ground surface albedos or in other words the reflectivity of different ground coverages. The corresponding graph can be found in the section 2.5 Figure 15. According to this graph one can see that on average the reflectivity from all ground coverages at 1.6 μm are higher by about 25% than reflectivity from the same coverages at 2.3 μm. From Figure 14 one can see that the radiance at around 1.6 μm (~6000-6250 cm^{-1}) is greater in 2-5 times than the radiance at 2.3 μm (~4347 cm^{-1}).

To sum up, here we have a kind of compromise between moderately strong methane absorption at 2.3 μm, but lower reflectivity from ground coverages and hence minor radiance detected by HySpex and between a weaker methane absorption at 1.6 μm, but higher surface reflectivity at this range and hence higher radiance (by 3-4 times), detecting by observing sensor.

A good visualization of the abovementioned conclusions can be the following Figure 34. On this figure one can see the simulations for the atmosphere (solid lines) at 1.6 μm (upper graph) and 2.3 μm (lower graph) with the normal, doubled and tenfold methane concentration. As it was already mentioned the radiance detected by HySpex camera (crosses) at 1.6 μm is higher (by 3-4 times) than at 2.3 μm, but from the other hand the absorption presented at 2.3 μm extends over a larger wavenumber interval (~4150 - 4550 cm^{-1}) than in comparison to 1.6 μm interval (~5990 - 6020 cm^{-1}). The critical and most important point in the wavenumber interval selection for methane plume retrieval is the ground albedos behavior in the corresponding intervals. At 1.6 μm ground albedos, and hence the radiance detected by HySpex, do not changing dramatically in comparison to the 2.3 μm. That means that 2.3 μm wavenumber interval can not be used for methane plume retrieval instead of stronger methane presence there.

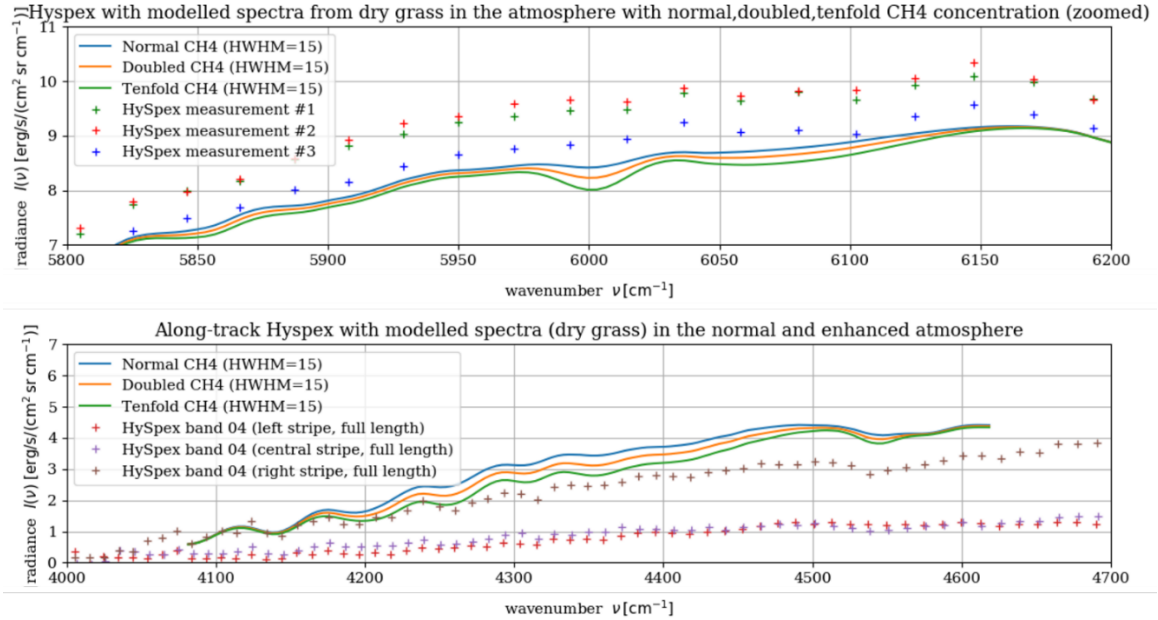


Figure 34 HySpex sampling measurements with the radiance for simulated atmosphere with normal, doubled and tenfold methane concentration at 1.6 µm (upper graph) and 2.3 µm (lower graph). Simulated radiance convolved with Gaussian function at HWHM = 15 cm^{-1}

5.2. Selection of methane shaft

In the section 2.1 it was already introduced that the research measurement flight with HySpex sensor onboard was a part of CoMet campaign. On the 7-th of June 2018 a lot of methane shafts (Figure 6) with different values of annual emissions were sensed. The table below summarizes the data of the location of the corresponding methane shafts by the covered HySpex lines.

Table 1 HySpex lines coverage over emitting CH_4 shafts in the Upper Silesian region (Poland)

Name of CH4 shaft	Sequential number of HySpex line →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	altitude of measurements →	~1520 m									~2900 m				~1620 m				
	Emissions (kt/a) ↓																		
Zofiowka IV	9.83	•									•								
Zofiowka V	9.83	•									•								
Pniowek III	20.09		•					•					•						
Pniowek IV	20.09				•		•							•					
Pniowek V	20.09									•		•							
Borynia III	5.58						•					•		•					
Borynia VI	5.58								•	•		•							
Silesia I	7.13														•				
Silesia V	7.13														•				
Brzeszcze II	9.63															•		•	
Brzeszcze IV	9.63															•			
Brzeszcze VI	9.63																		•
Brzeszcze IX	9.63															•		•	

Based on shaft's operator data of hourly methane emissions (Galkowski M., Swolkien J. and Fiehn A.; personal communication) and according to the analysis of methane emissions from air shafts of the Jastrzebska coal mining company in 2017-2019 by Badura et al. (2020), it was found that the most promising shaft for methane plume detection is Pniowek V, covered by HySpex lines 9 and 11 (highlighted by green color).

5.3. Raw HySpex SWIR-320m-e data over shaft Pniowek V

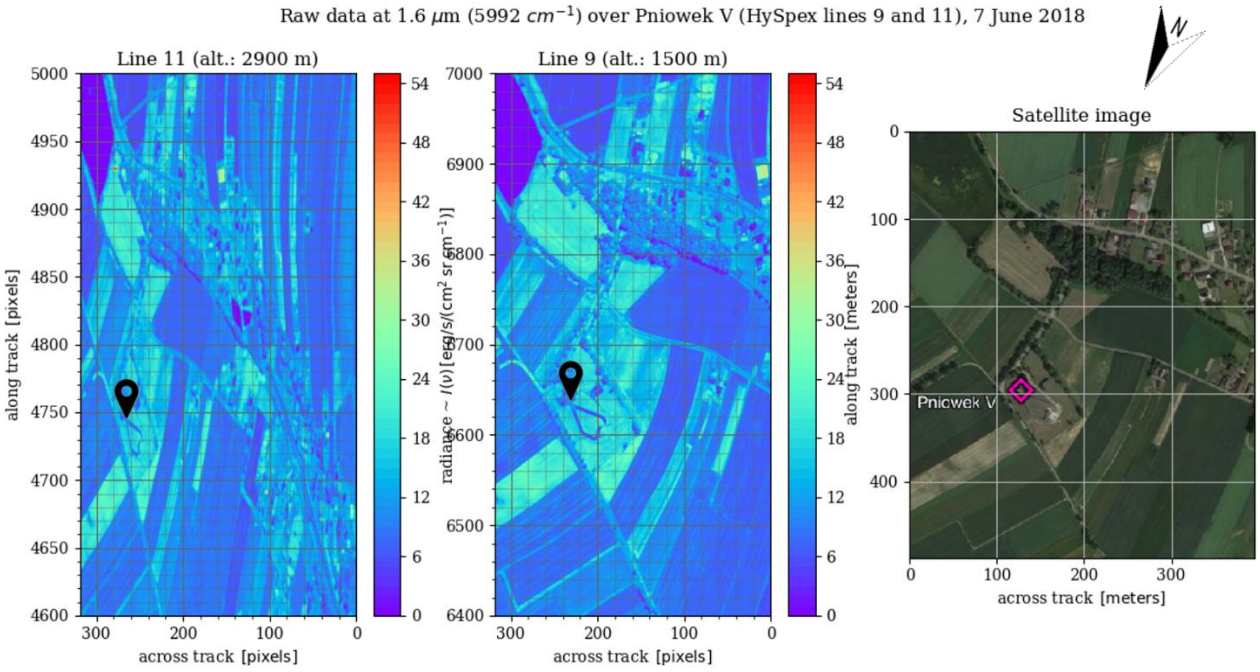


Figure 35 Raw data of HySpex SWIR-320m-e camera over methane shaft Pniowek V from the HySpex lines 9 and 11 at 1.6 μm (5992 cm^{-1}), 7-th June 2018

Figure 35 shows the HySpex raw data from two different HySpex lines 9 and 11 over shaft Pniowek V selected for methane plume retrieval. The position of methane shaft is marked by the corresponding symbol.

As it was already mentioned in chapter 3 the core idea of all four methods for methane retrieval was to get rid of the ground reflectivity patterns that are clearly seen at Figure 35, in order to be able to detect the plume itself.

The following Figure 36 demonstrates the radiance behavior detected by HySpex camera over Pniowek V at the different parts of the lines (swaths) 9 and 11.

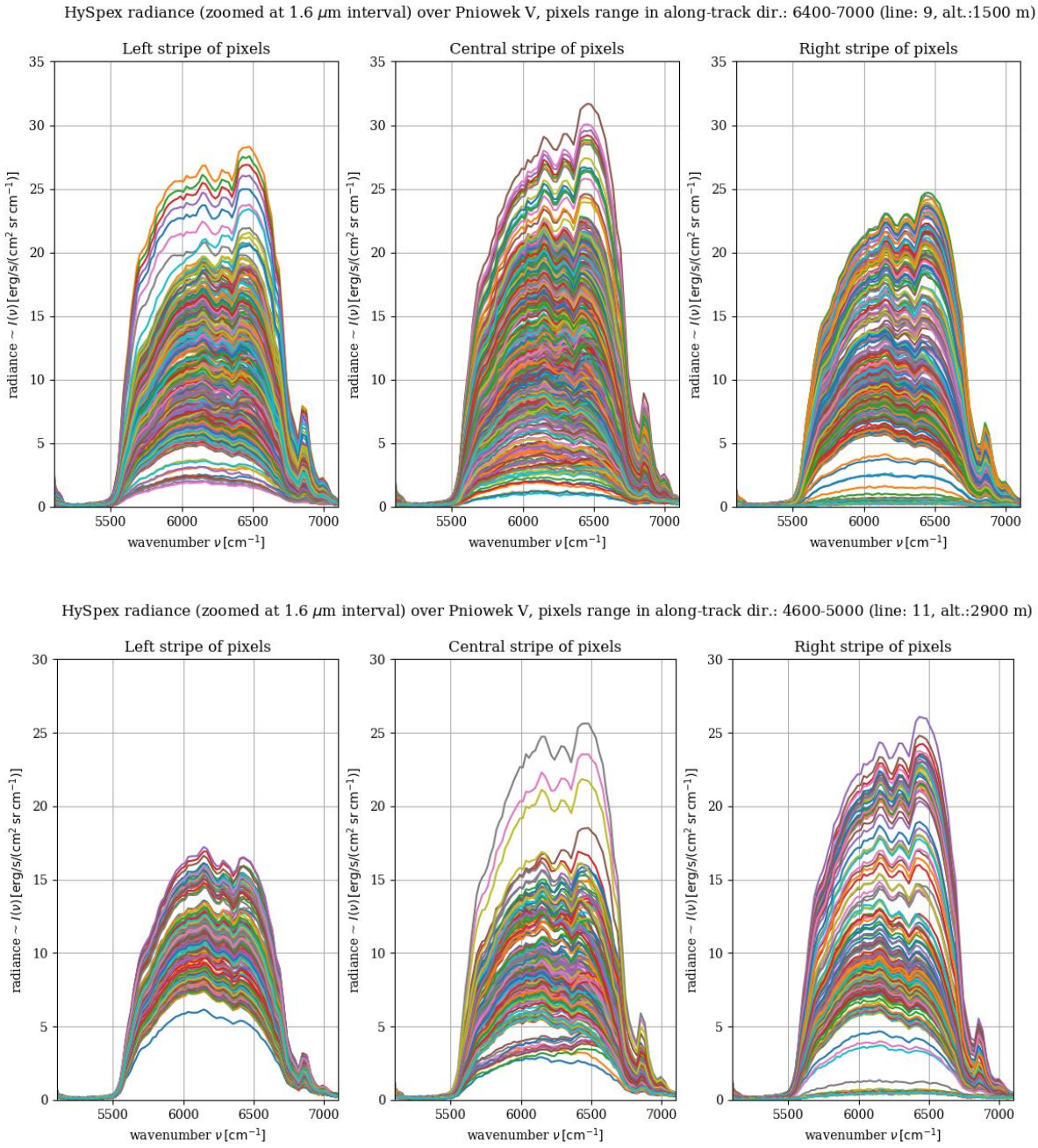


Figure 36 Radiance detected by HySpex camera at 1.6 μm interval over methane shaft Pniowek V from the HySpex line 9 (upper graphs) and 11 (lower graphs) at different parts of the swaths, 7-th June 2018

5.4. Method I. Difference of the radiances from methane and transparent intervals

The following graph demonstrates the radiance after subtraction of the spectrally averaged transparent atmosphere from the wavenumber interval with methane for the HySpex lines 9 and 11. The position of methane shaft is marked by the corresponding symbol.

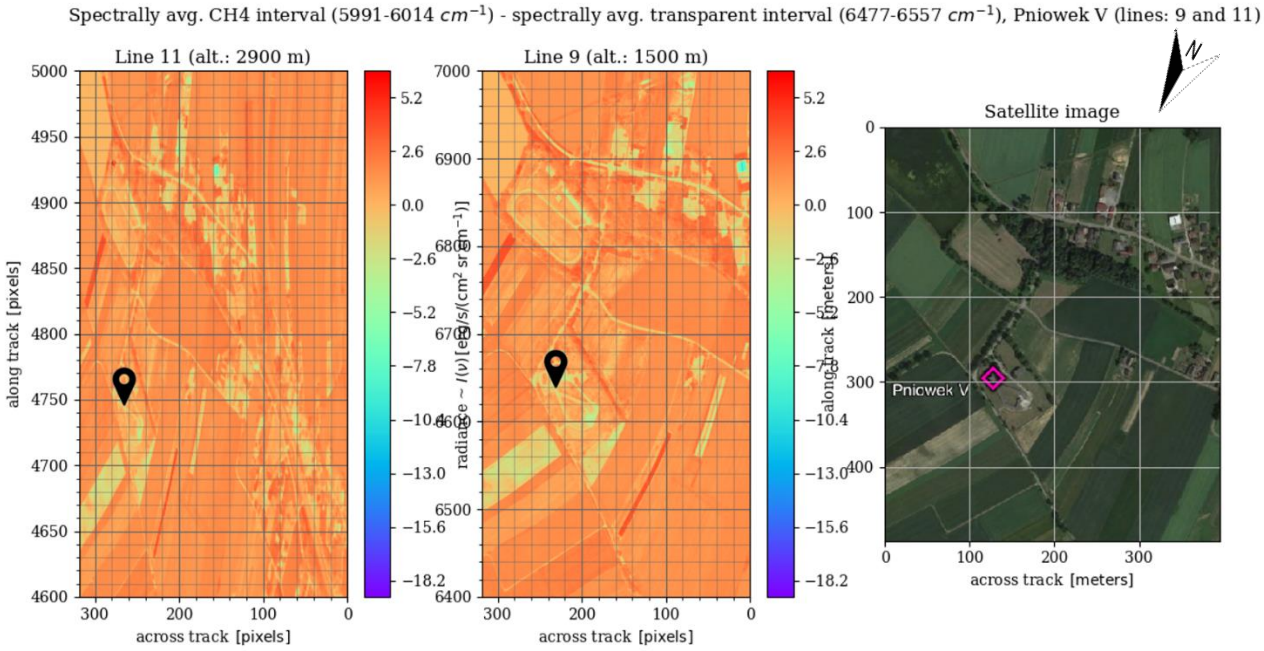


Figure 37 Subtraction of the spectrally averaged transparent interval (6477-6557 cm^{-1}) radiance from methane wavenumber interval (5991-6014 cm^{-1}) radiance at the HySpex lines 9 and 11

Methane plume from this method over methane shaft Pniowek V is not detected neither on the HySpex line 9, nor on the line 11.

5.5. Method II. Ratio of the radiances from methane and transparent intervals

Since we are looking of the ratio of two radiances, the values in the color bars are unitless. The position of methane shaft is marked by the corresponding symbol.

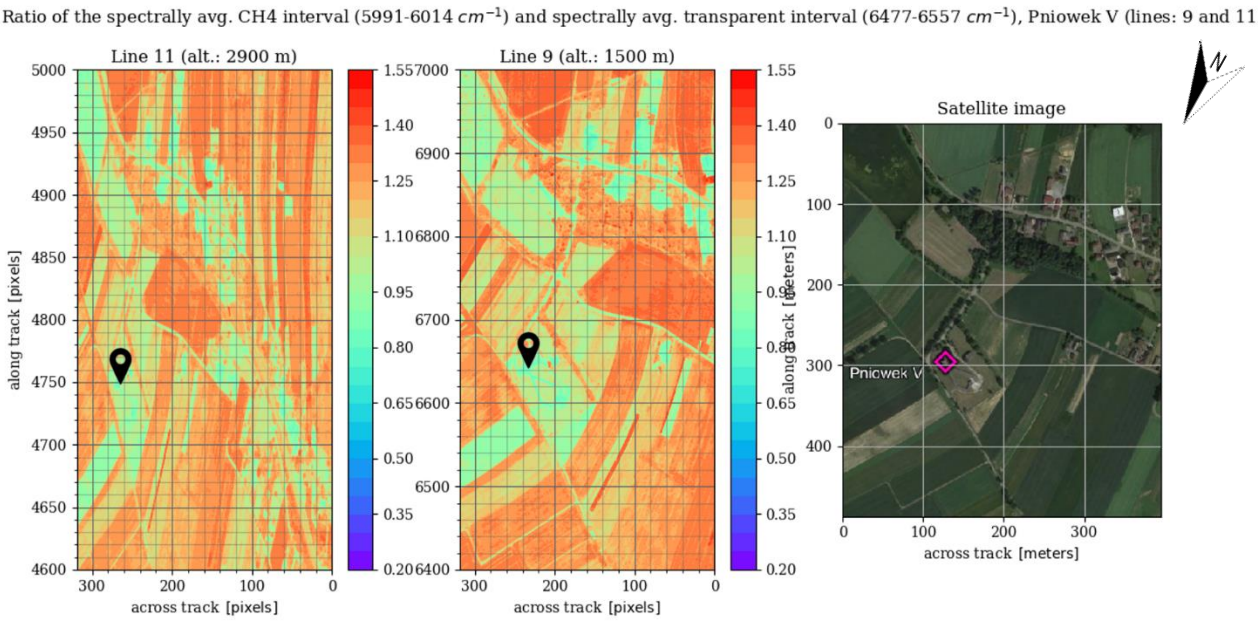


Figure 38 Ratio of the spectrally averaged methane interval (5991-6014 cm^{-1}) radiance and spectrally averaged transparent interval (6477-6557 cm^{-1}) radiance at the HySpex lines 9 and 11.

Methane plume from this method over methane shaft Pniowek V is not detected neither on the HySpex line 9, nor on the line 11.

5.6. Method III. Ratio of the radiances from methane and transparent intervals with 2D convolution

Figure 39 introduces the result of the division between spatially (5×5 pixels) and spectrally averaged methane interval and only spectrally averaged transparent atmosphere interval. Since we are looking for the ratio of two radiances, the values in the color bars are unitless. The position of methane shaft is marked by the corresponding symbol.

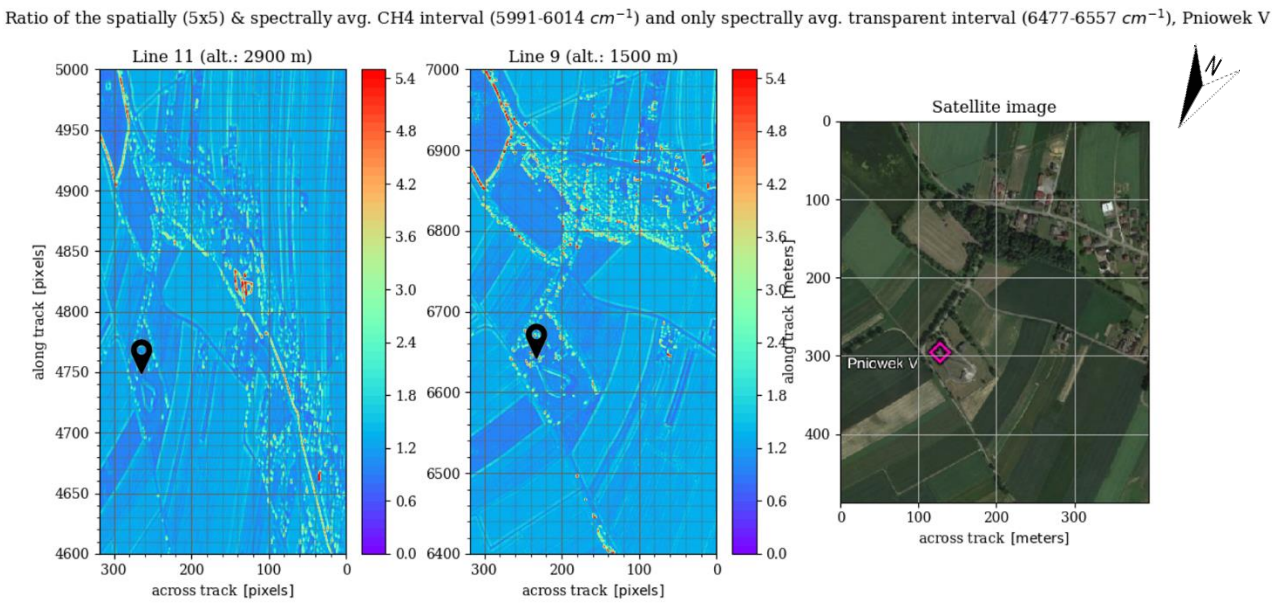


Figure 39 Ratio of the spatially (5×5 pixels) and spectrally averaged methane interval (5991-6014 cm^{-1}) radiance and only spectrally averaged transparent interval (6477-6557 cm^{-1}) radiance at the HySpex lines 9 and 11.

Methane plume from this case over methane shaft Pniowek V is not detected neither on the HySpex line 9, nor on the line 11.

The graph below (Figure 40) demonstrates the ratio of the spatially (5×5 neighbor pixels) and spectrally averaged methane interval (5991-6014 cm^{-1}) and only spectrally averaged methane interval radiance over Pniowek V. Again, like the previous cases since there is a ratio of radiances here the values in the color bars are unitless. The position of methane shaft is marked by the corresponding symbol.

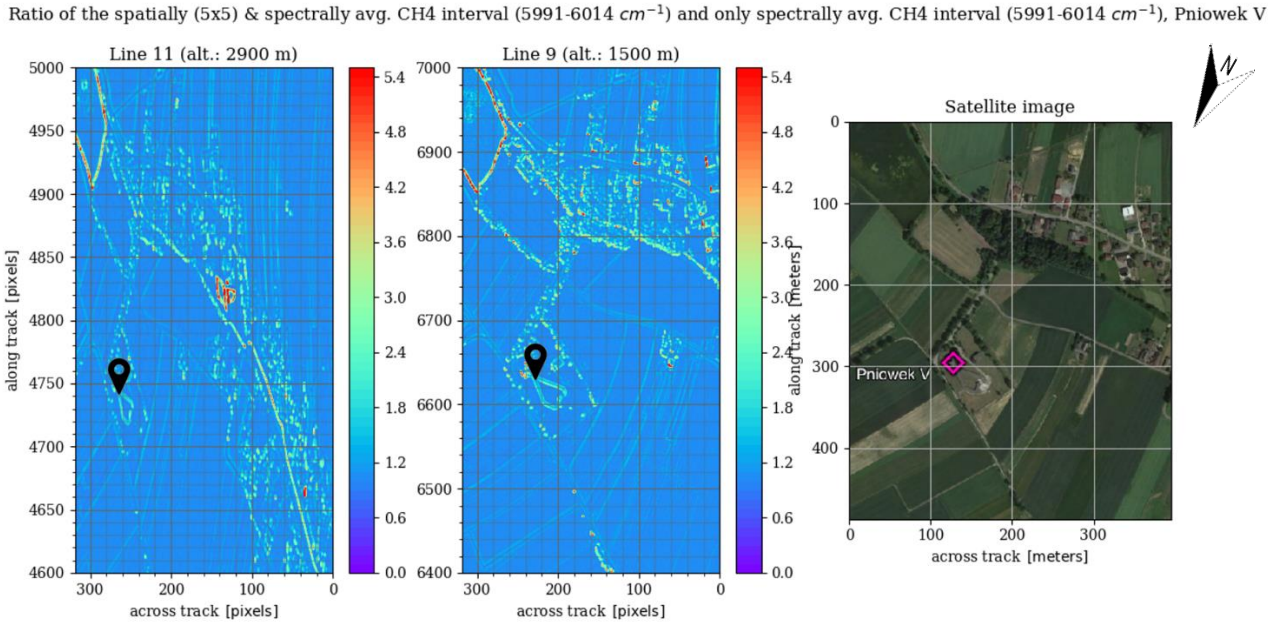


Figure 40 Ratio of the spatially (5×5 pixels) and spectrally averaged methane interval (5991-6014 cm^{-1}) radiance and only spectrally averaged methane interval (5991-6014 cm^{-1}) radiance at the HySpex lines 9 and 11.

Methane plume from this case over methane shaft Pniowek V is not detected neither on the HySpex line 9, nor on the line 11.

5.7. Method IV. Ratio of absorbing and transparent residuals

The graph below (Figure 41) demonstrates the ratio of the absorbing residuals and transparent atmosphere residuals from the polynomial based on the data points corresponding to the minimum absorption (aka transparent interval). Similar to the previous cases since there is a ratio of radiances here the values in the color bars are unitless. The position of methane shaft is marked by the corresponding symbol as well as the position of methane plume.

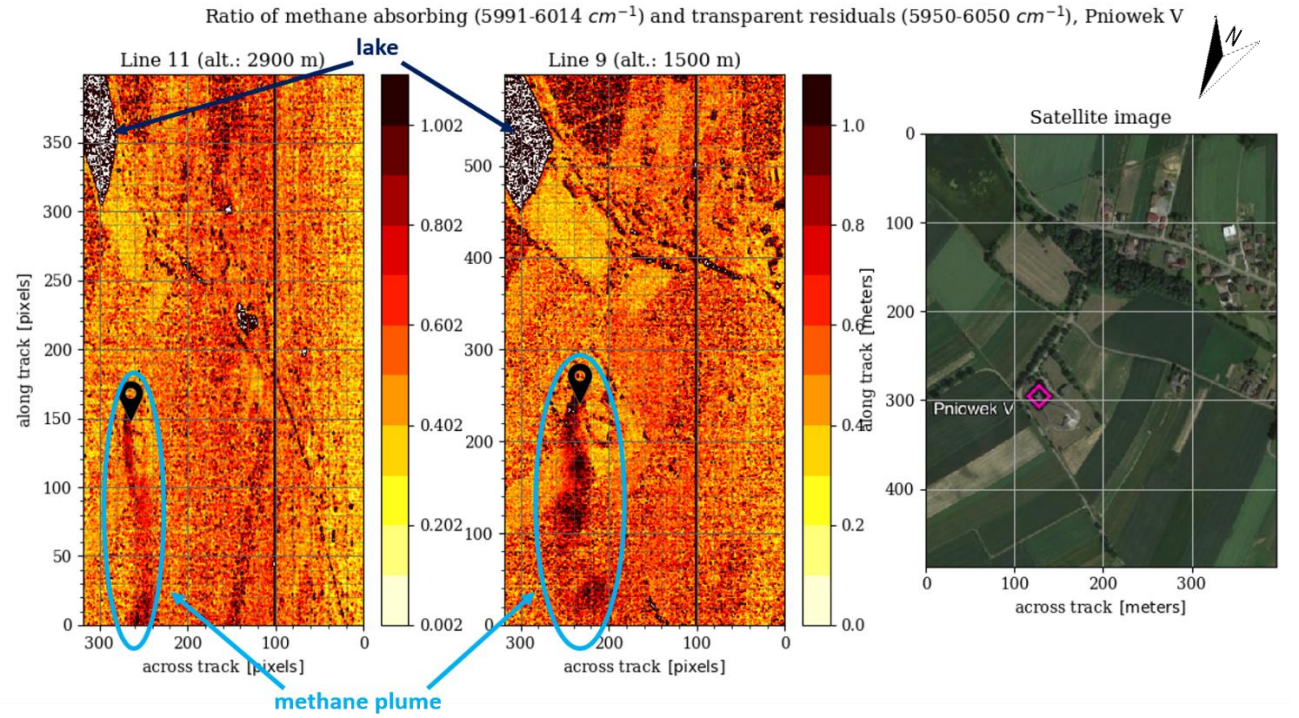


Figure 41 Detected methane plume on the ratio of the absorbing (5991-6014 cm^{-1}) and transparent (5950-6050 cm^{-1}) intervals residuals over Pniowek V at the HySpex lines 9 and 11.

According to this figure the size of the detected plume is around 200 meters long and ~40 meters wide. It is important to note that the output results are wavenumber sensitive. This means that in case of shifting methane interval to other wavenumbers, the final result graphs will not show methane plume anymore.

After detection of methane plume over the shaft Pniowek V on HySpex lines 9 and 11, this method was applied to all other HySpex lines and shafts. The following table summarizes results of the implementation of this algorithm.

Table 2 Detected (✔) and non-detected (x) methane plumes by method IV over CH_4 shafts covered by HySpex in the Upper Silesian region (Poland)

Name of CH4 shaft	Sequential number of HySpex line →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	altitude of measurements →	~1520 m									~2900 m				~1620 m				
	Emissions (kt/a) ↓																		
Zofiowka IV	9.83	x									x								
Zofiowka V	9.83	x			x						x								
Pniowek III	20.09		x					✔					x						
Pniowek IV	20.09				x		x							x					
Pniowek V	20.09									✔		✔							
Borynia III	5.58						x					x		x					
Borynia VI	5.58								x	x		x							
Silesia I	7.13														x				
Silesia V	7.13														x				
Brzeszcze II	9.63															x		x	
Brzeszcze IV	9.63															x			
Brzeszcze VI	9.63																		x
Brzeszcze IX	9.63															x		x	

As it is clearly seen from the above table, methane plumes were not detected over most of the shafts sensed by the HySpex instrument in the Upper Silesian region on 7-th of June 2018. This does not mean that most of the shafts were not working, more likely that plume was 'blown' by wind out of the HySpex sensed areas (lines). However, two more weak methane emissions were detected on the Hyspex line 7 over methane shaft Pniowek III. The following Figure 42 demonstrates the marked position of the shaft as well as the weak methane plume of 40-50 of meters long and 10-20 meters wide and another neighbor weak methane source, which probably also belongs to the complex of Pniowek III shaft.

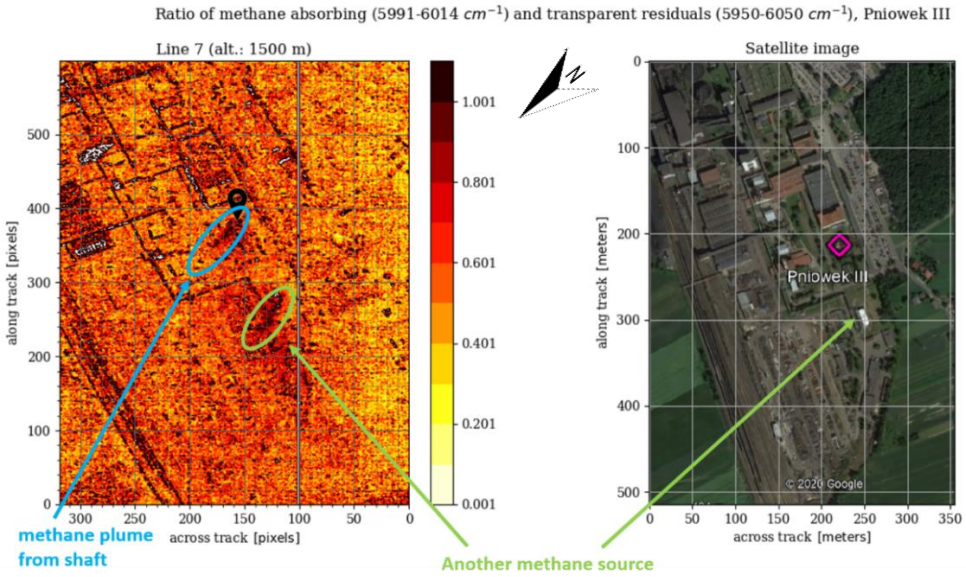


Figure 42 Detected methane plume and another methane source on the ratio of the absorbing (5991-6014 cm^{-1}) and transparent (5950-6050 cm^{-1}) intervals residuals over Pniowek III at the HySpex line 7.

6. Py4CAtS methane plume radiance simulations

6.1. Comparison between observed and modelled radiances in along-track direction

After successful detection of methane plumes on the HySpex lines 7, 9 and 11 over methane shafts Pniowek III and Pniowek V the next step is to find such parameters of the ground albedo, that after substitution them in Equation 2, the resulting modelled spectrum will be the most accurately consistent with HySpex measurements directly from the plume pixels and from the pixels far from methane shaft. For this purpose, one has to select the most prominent pixels from the plume with the highest methane concentration or in other words the 'darkest' pixels. For this procedure the methane plume from Pniowek V detected on HySpex lines 9 and 11 was selected. Figure 43 depicts the pixel coordinates of the areas with probably high methane concentration on the HySpex lines 9 and 11 respectively.

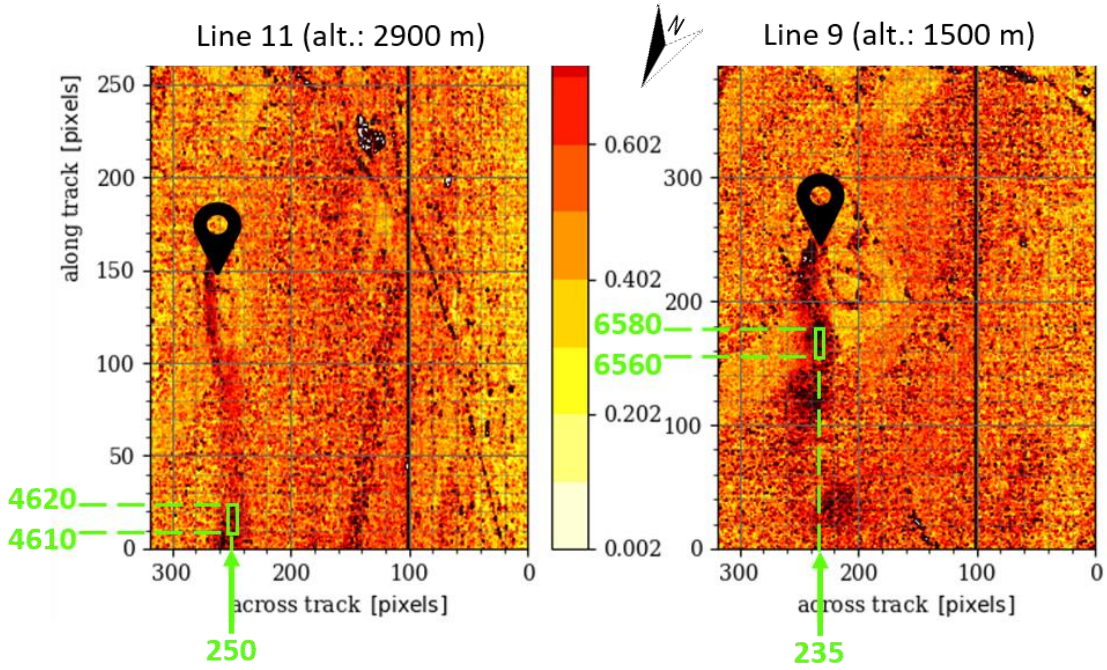


Figure 43 Zoomed fragment of the detected methane plumes from the method IV over Pniowek V at the HySpex lines 9 and 11 with marked areas for radiance modelling.

After selecting the areas with probably high methane concentration in the plume one can plot the radiance observed by HySpex at these particular regions together with the modelled radiances from various ground albedos and for comparison with the constant value of albedo (r=0.5), similar as it was done in section 4.4 'Interpolation of albedo'. The following Figure 44 demonstrates modelled radiances with empirical and constant albedo (solid lines) as well as the HySpex measurements (crosses) at the selected areas of methane plume with probably the highest methane concentration.

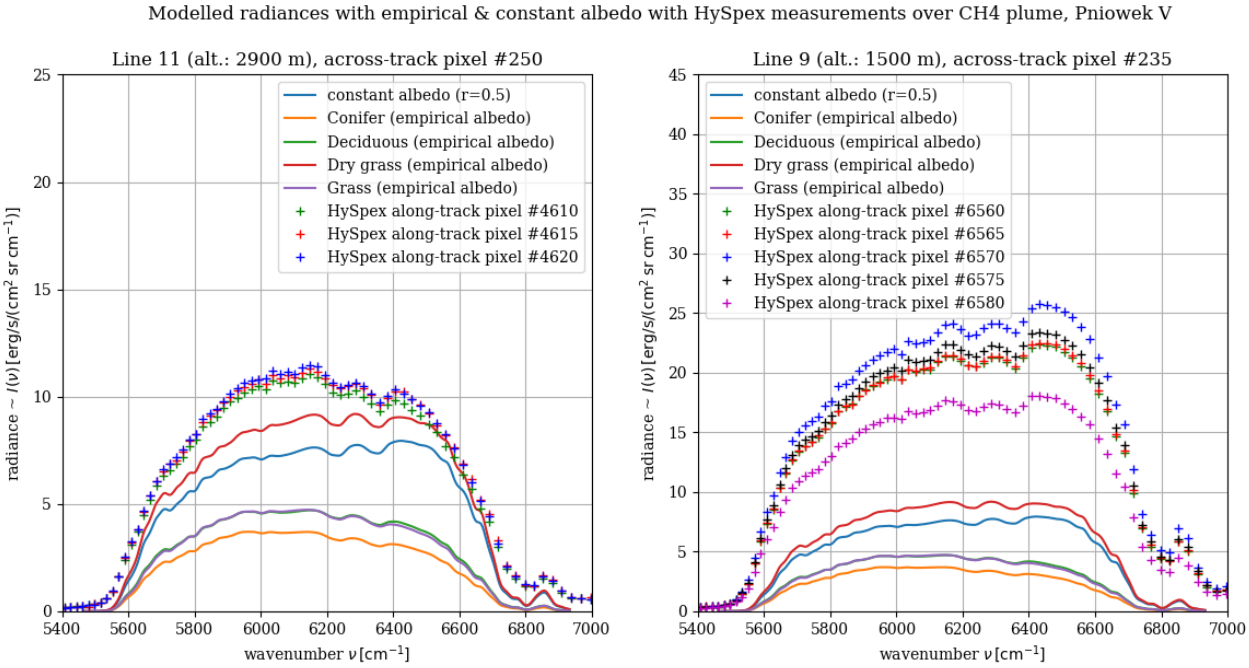


Figure 44 Modelled radiance from different ground coverages with empirical albedo, modelled radiance with constant albedo and HySpex measurements over selected areas from plumes for the HySpex lines 9 and 11.

On this figure it is clearly seen that radiance values of the HySpex measurements on line 11 almost twice smaller than on line 9. This is due to the fact that the altitude of the sensing at line 11 was almost twice higher (2900 m) than on line 9 (1520 m), hence on the lower altitude of the measurements there is less attenuation or absorption on the path from the ground surface to the airborne observer. As a consequence, the upwelling transmission is increasing for the lower altitude measurements, according to the Equation 7, hence the observed radiance on line 9 is higher than in comparison to the line 11.

6.2. Comparison between observed and modelled radiances in along-track and across-track directions

In this chapter for radiance modelling we will focus on the HySpex line 9, since the sensing altitude on this line was the lowest from all available sensor data and hence less attenuation of the upwelling radiance. Existence of methane plume means that the optical depths are high, which leads to smaller transmissions through the atmosphere and hence, less observed radiance, according to the Equation 2.

Figure 45 demonstrates zoomed region with the position of the 10×10 pixels square from methane plume from the HySpex line 9. On the right side of figure marked area also depicted on the corresponding region of the satellite image.

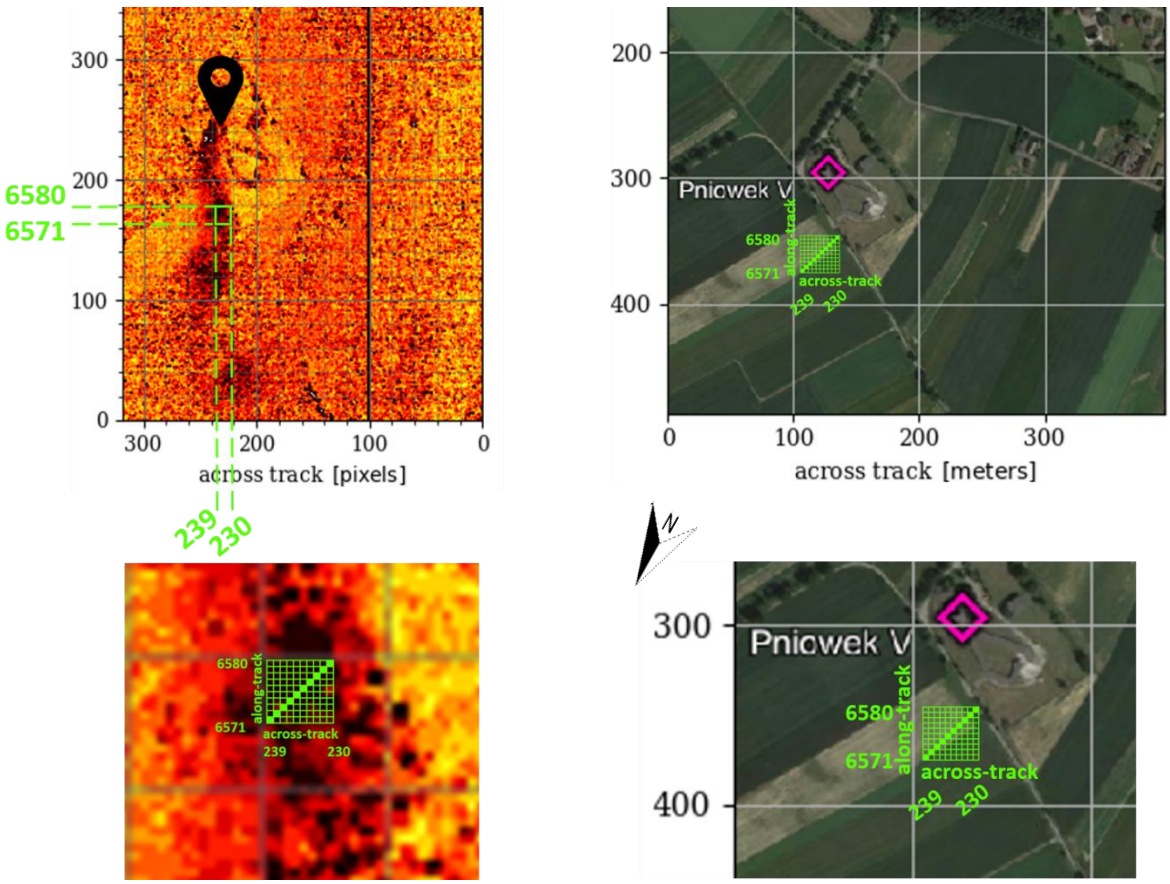


Figure 45 Zoomed region with the position of the 10×10 pixels area from methane plume from the HySpex line 9 over Pniowek V.

Following Figure 46 shows the radiances detected by the HySpex senor from the diagonal pixels highlighted by green on Figure 45 together with the modelled radiance from various constant albedos.

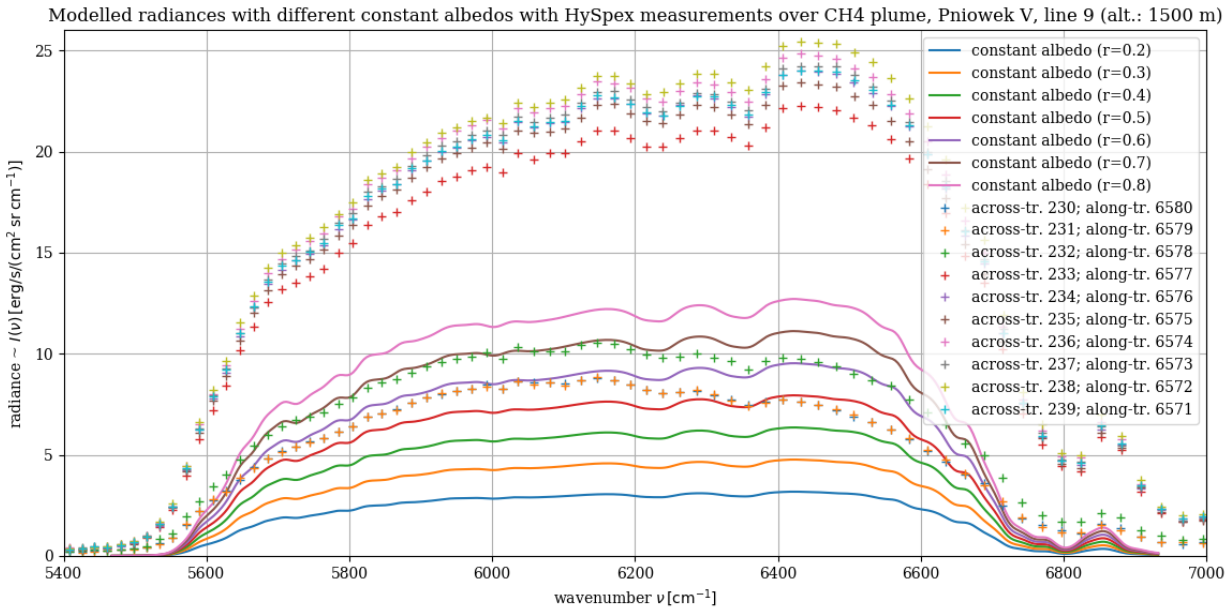


Figure 46 Modelled radiances with different constant albedos with HySpex measurements over methane plume from the HySpex line 9 over Pniowek V.

It is clearly seen on Figure 46 two groups of the radiances detected by HySpex sensor: first group with the radiance of around 10 $\text{erg/s/(cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ corresponds to the upper right area of highlighted by green pixels on Figure 45 (across track: 230-232, along-track: 6578-6580) and the second group corresponds to all other observing diagonal pixels with much higher radiance values.

6.3. Least-squares estimation of geophysical state parameters

In order to estimate the accuracy of the modelled radiance, one has to compare it with the observed HySpex measurements. For this purpose, we will consider the 5850-6150 cm^{-1} wavenumber interval. As a first step of estimation we will look at the residuals between modelled interpolated radiance and the HySpex measurements.

Since the grid of the modelled radiance and the observed radiance differ (the modelled is denser), one has to interpolate the modelled radiance to the same grid as the observed one. Figure 47 (upper graph) demonstrates the zoomed wavenumber interval from Figure 46 with the modelled radiances from albedos r=0.6 and r=0.5 (purple and dark red curves respectively), modelled radiance from albedo r=0.6 interpolated to the observed radiance grid (light red curve) and HySpex measurements from the across-track pixel 231 and along-track pixel 6579 (orange curve). On the lower graph one can see the absolute values of the residuals from the subtraction between modelled interpolated radiance with albedo r=0.6 and the HySpex data from the abovementioned pixel.

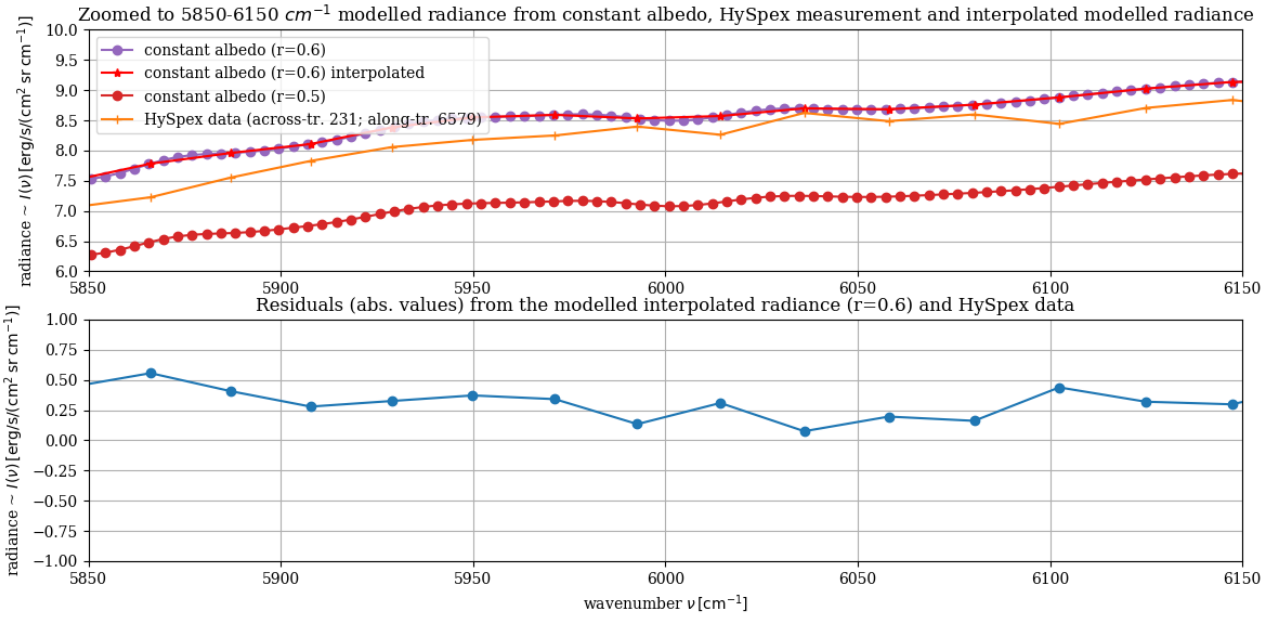


Figure 47 Upper graph: Zoomed to 5850-6150 cm^{-1} modelled radiance from constant albedo (r=0.5 and r=0.6), modelled radiance with interpolated grid points and HySpex measurements from particular pixel. Lower graph: Absolute values of the residuals from the modelled interpolated radiance (r=0.6) and HySpex data from particular pixel.

Figure 47 demonstrates that the absolute values of the residuals at the 5850-6150 cm^{-1} wavenumber interval almost do not exceed 0.5 $\text{erg/s/(cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$. However, for more accurate estimation and analysis of the observing radiance we will implement least-squares fitting of the first order polynomial presented in the next subsection.

6.3.1. Least-squares fitting of reflectivity

Measured intensity spectrum $\hat{\Im}(\nu)$ and transmission spectrum $\mathcal{T}(\nu)$ described by the mathematical forward model $\phi(x)$ according to Beer's law for a double path through the atmosphere and accounts for characteristics of the instrument itself in terms of instrument response function *S*. Evaluation of the Equation 2 yields:

$$\Phi(x) \equiv \hat{\Im}(\nu) = \frac{r(\nu)}{\pi} \cdot \cos(SZA) \cdot \Im_{Sun}(\nu) \cdot \exp\left(-\sum_m \beta_m \tau_m(\nu)\right) \otimes S(\nu, \gamma) \qquad (8)$$

where the surface reflectivity $r(\nu)$ in this subsection is modelled by a first order polynomial: $r(\nu) = r_0 + r_1\nu$. The state vector in this particular subsection stands for $x = (r_0, r_1)$ including the unknown parameters to be estimated. The half width of the instrument's response function $S(\nu, \gamma)$ is denoted as γ and can be treated independent of the wavenumber ν .

The general equation of the least-squares residuals calculation for the abovementioned modelled polynomial can be written as:

$$\chi^2 = \sum_{i=1}^{m} \left(y_{obs}(\nu_i) - \left(r_0 + r_1 \cdot \nu_i\right)\right)^2$$

where $y_{obs}(\nu_i)$ - the HySpex measurements, r_0 and r_1 - unknown albedo parameters. According to the least-squares error, the unknown parameters can be found:

$$r_0 = \frac{\langle y \rangle \langle \nu^2 \rangle - \langle y\nu \rangle \langle \nu \rangle}{\langle \nu^2 \rangle - \langle \nu \rangle^2}$$

$$r_1 = \frac{\langle y\nu \rangle - \langle y \rangle \langle \nu \rangle}{\langle \nu^2 \rangle - \langle \nu \rangle^2}$$

Following Figure 48 demonstrates comparison between least-squares fitted by first order polynomial $(r(\nu) = r_0 + r_1\nu)$ albedo values r_0 in the atmosphere with full water concentration with the satellite image over Pniowek V shaft for the HySpex line 9 with sensing altitude 1500 m. The term full water concentration means the usage of 'standard' midlatitude summer (MLS) atmospheric dataset without any further changes. The position of the shaft is marked by the corresponding symbol.

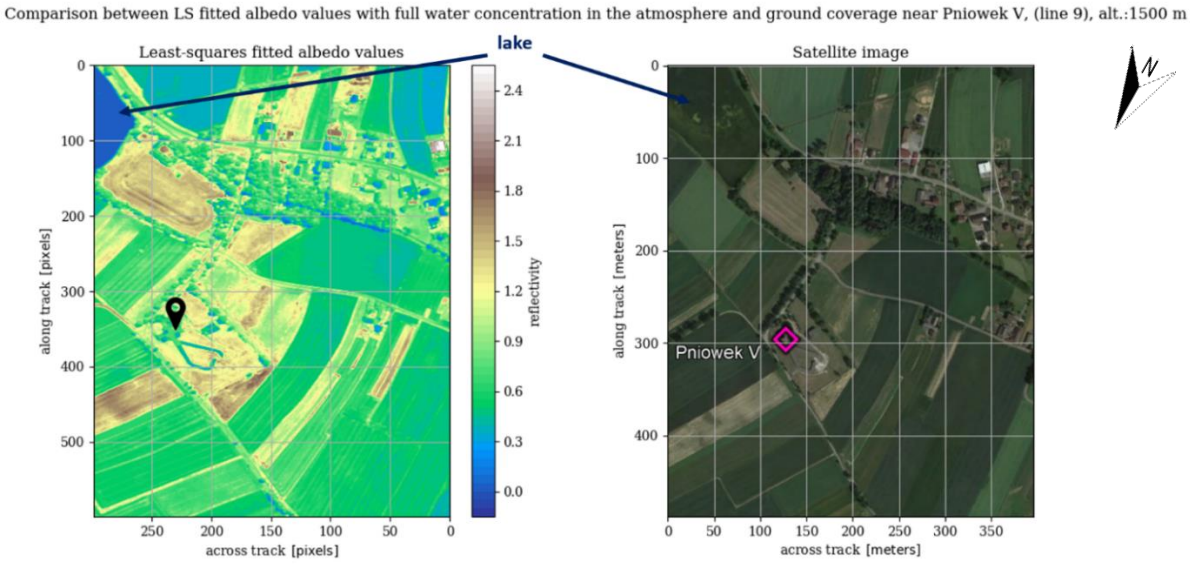


Figure 48 Comparison between least-squares fitted by 1-st order polynomial albedo values (r_0) in the atmosphere with full water concentration with ground coverage from the satellite image over Pniowek V (300×600 pixels) at HySpex line 9 (alt.: 1500 m)

As it is clearly seen from Figure 48 obtained from the least-squares fitting reflectivity values (r_0) are greater than 1 in some parts of the graph. This is physically implausible since the albedo values can only be in range [0; 1]. One of the possible solutions of this kind of problem can be more accurate description of the atmospheric conditions during the measurements. Figure 49 shows atmospheric and weather conditions over Pniowek settlement on the day of the measurements (7 July 2018) at the time slot (10:00-12:00 a.m.).

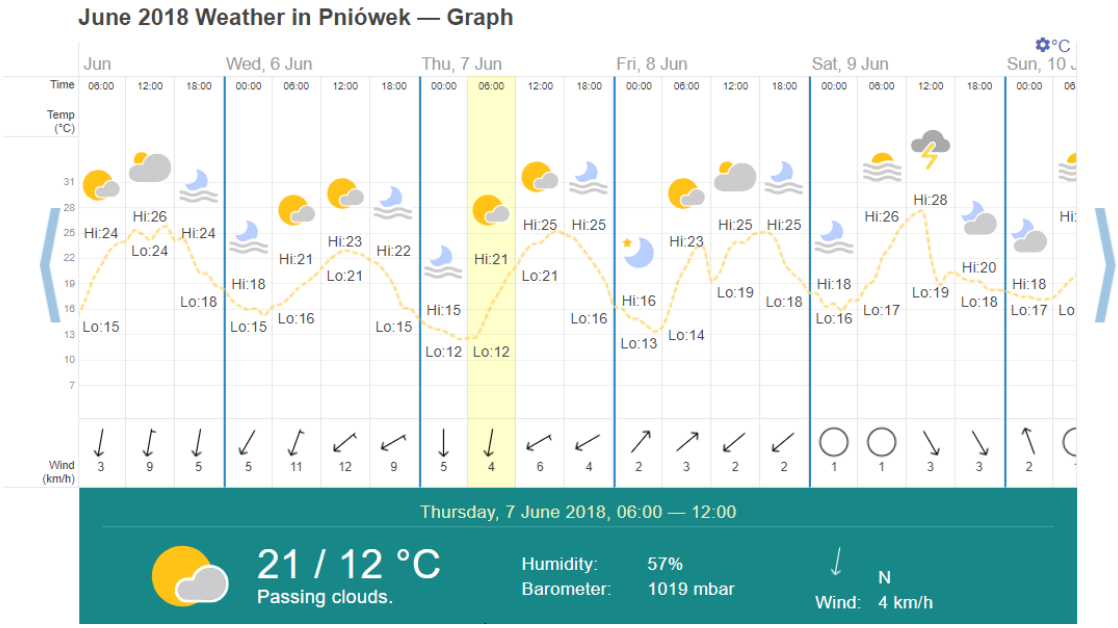


Figure 49 Weather conditions in Pniowek settlement on 7-th of June 2018 at the time slot of the measurements (10:00-12:00 a.m.) [www.timeanddate.com/weather]

According to Figure 49 the humidity of the air during the measurements was around 60%. That means that by decreasing the water concentration in the atmosphere using for the simulations we can obtain more reasonable albedo values. Therefore, in the results presented in the next subsection atmosphere with half water concentration was used.

6.3.2. Least-squares fitting of methane scaling factor with constant albedo

Another critical parameter for the accuracy of the obtained results is the optical depth. In the simulation processes used before we used the optical depth of all gases through the whole atmosphere, more precisely till the observing altitude (1500 m or 2900 m for different lines). Methane plume retrieval should scale an a priori plume according to the measurement data. So, we are only interested in the enhancement of the methane column (i.e. the plume) in the lowest part of the atmosphere, e.g. <500 m. Figure 50 shows the comparison between monochromatic optical depths (upper graph) of all gases through the whole atmosphere (blue curve), of methane through the whole atmosphere (orange curve) and methane only through the bottom layer 0-500 m (green curve). The lower graph on the same figure demonstrates the same optical depths but convolved with Gaussian function (HWHM=15 cm^{-1}) for clarity.

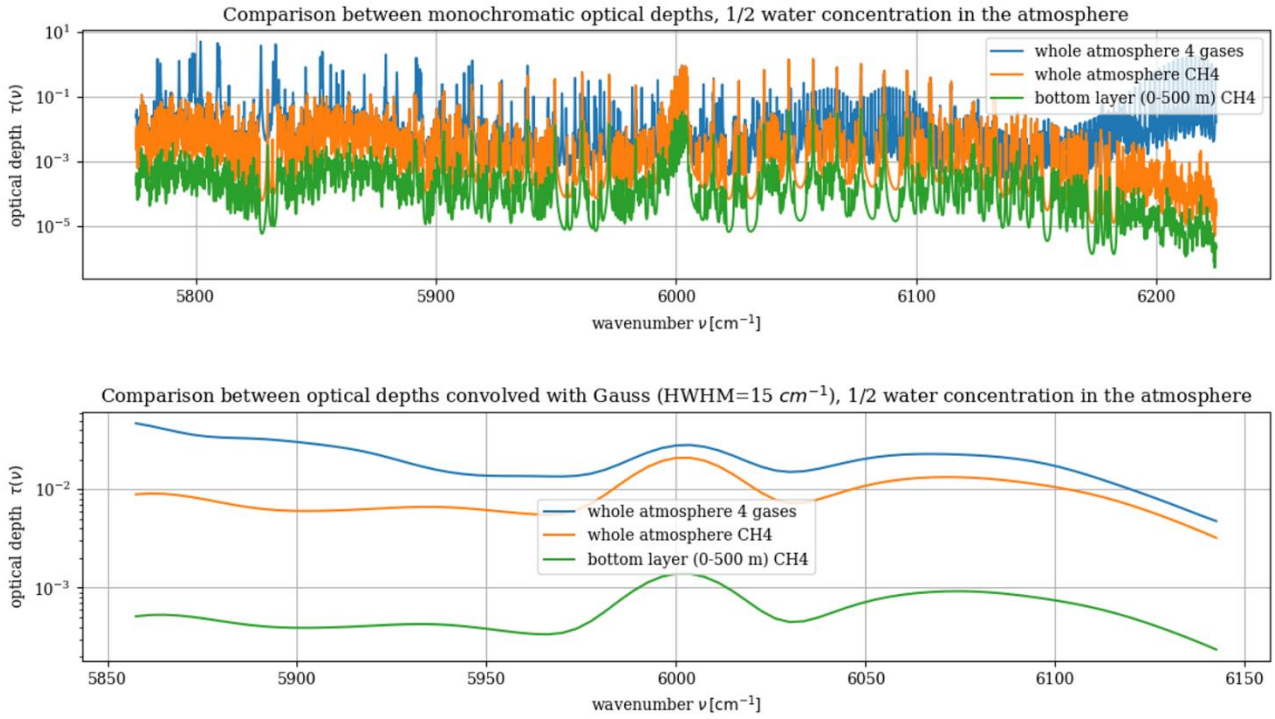


Figure 50 Comparison between monochromatic optical depths (upper graph) through the whole atmosphere with half water concentration of 4 gases (H_2O, CO_2, CH_4, CO), only methane and methane from the first layer (0-500 m) with the same optical depths, convolved with Gaussian function (lower graph) .

From Figure 50 it is clearly seen that there is a noticeable difference between for example the optical depth of the methane through the whole atmosphere (orange curve) and only through the bottom layer (green curve). For more precise results we will use the methane optical depth through the whole atmosphere in further simulations instead of optical depth through the whole atmosphere of all gases.

Coming back to the Equation (8) in current subsection we will introduce the state vector as a $x = (\beta, r)$. The molecular scaling factor(s) β_m of the absorbing methane molecule m are contained in β. As it was already mentioned before, in this subsection we are interested only in the methane optical depth. This optical depth can be interpreted as $\tau_{CH_4} = \tau_{ref} + \beta\tau_{plume}$. Where τ_{ref} -stands for the climatological optical depth (e.g. MLS) through

the whole atmosphere and τ_{plume} - stands for the extra optical depth due to enhanced methane presence, aka plume (0-500 m).

Moreover, since we are only using a small spectral range around methane absorption 6000 cm^{-1} it is assumed a constant reflectivity at this interval. Hence Equation (8) can be extend, but simplify according to:

$$\hat{\Im}(\nu) = \frac{r_0}{\pi} \cdot \cos(SZA) \cdot \Im_{Sun}(\nu) \cdot \exp\left(-\sum_m \tau_m^{ref}(\nu)\right) \times \exp\left(-\beta \tau_{plume}(\nu)\right) \otimes S(\gamma) \tag{9}$$

Before estimating the unknown parameters $x = (r_0, \beta)$ by linear least squares we first use the power-series expansions for the exponential function $\exp(\tau) := \sum_{n=0}^{\infty} \frac{r^n}{n!}$ to see that for small $\beta\tau_{plume}$ the individual terms for $n>1$ become very small and a linear approximation seems reasonable for our application. With this consideration, Equation (9) becomes:

$$\hat{\Im}(\nu) \approx \Im(\nu) = \frac{r_0}{\pi} \cdot \cos(SZA) \cdot \Im_{Sun}(\nu) \cdot \exp\left(-\sum_m \tau_m(\nu)\right) \cdot \left(1 - \beta \tau_{plume}(\nu)\right) \otimes S(\gamma) \tag{10}$$

The model for a linear least squares problem for *M* measurements and *n* unknowns is:

$$y(\nu_i) = \sum_{j=1}^{n} x_j \phi_j(\nu_i), \; i = 1, 2, ..., M \tag{11}$$

$$\phi_1 = \frac{\cos(SZA)}{\pi} \Im_{Sun} \mathcal{T} \otimes S,$$

$$\phi_2 = \frac{\cos(SZA)}{\pi} \Im_{Sun} \mathcal{T} \tau_{plume} \otimes S,$$

with $x_1 = r_0$ and $x_2 = r_0\beta$. In the definition of Equation (11) the columns are linearly independent because τ_{plume} depends on the wavenumber according to Figure 50.

Figure 51 demonstrates comparison between least-squares fitted albedo values $(x_1 = r_0)$ in the atmosphere with half water concentration with the satellite image over Pniowek V shaft for the HySpex line 9 with sensing altitude 1500 m. The position of the shaft is marked by the corresponding symbol.

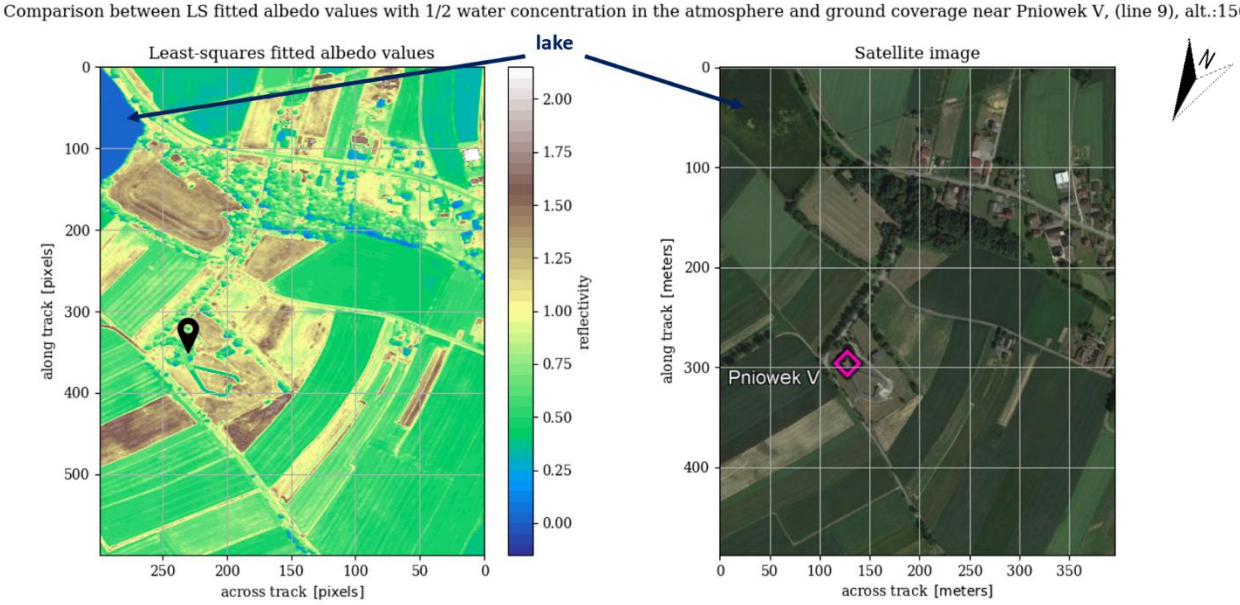


Figure 51 Comparison between least-squares fitted albedo values in the atmosphere with half water concentration with ground coverage from the satellite image over Pniowek V (300×600 pixels) at HySpex line 9 (alt.: 1500 m)

As it is clearly seen from Figure 51 the reflectivity values in this case are smaller than in Figure 48, $r_0 < 1$ for most of the pixels, but still there are some 'outliers' with reflectivity larger than 1. One of the possible reasons of physically implausible reflectivity values can be the 'mirror' model of the reflection, especially at the bright pixels. While the assumption for the reflectivity model that was used in the equation describing observed radiance (Equation 2) is true for Lambertian or diffuse reflection model.

(a) (b)

Figure 52 (a) – 'Mirror' model of the reflection; (b) – Lambertian or diffuse reflection model [NVIDIA ARC GmbH (2015)]

Following Figure 53 demonstrates comparison between least-squares fitted methane scaling factor for the lowest atmospheric layer (0-500 m) as $\beta = \frac{x_2}{r_0}$ in the atmosphere with half water concentration with the output from the method IV over Pniowek V shaft for the HySpex line 9 with sensing altitude 1500 m. The position of the shaft is marked by the corresponding symbol.

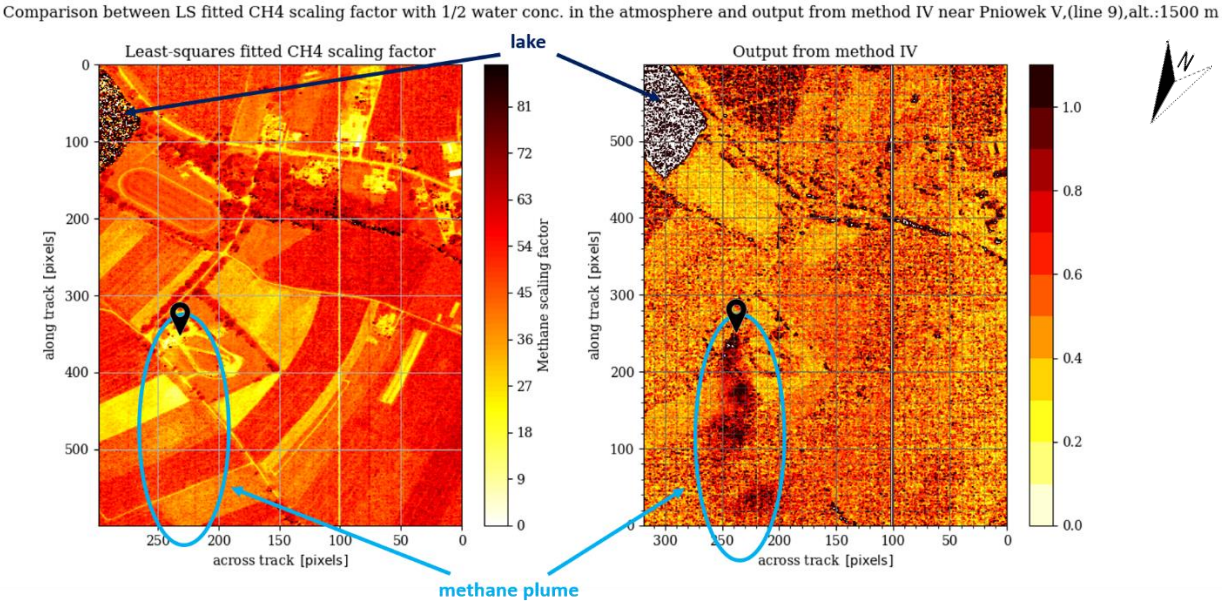


Figure 53 Comparison between least-squares fitted methane scaling factors in the atmosphere with half water concentration with the fragment of the output from the Method IV, over Pniowek V (300×600 pixels) at HySpex line 9 (alt.: 1500 m)

On the left graph of Figure 53 one can see barely noticeable trace (light yellow spots) coinciding with the position of the methane plume (dark red) on the right graph. The methane scaling factor is a unitless variable showing the enhancement of the methane concentration. This means that light yellow spots on the left graph corresponds roughly to the numbers 10 – 18. Normal methane concentration in the atmosphere till 500 m is 1.7 ppm, hence the scaling factor of 10 corresponds to the concentration of the methane of 17 ppm.

6.4. Radiance simulations with COSPAR International Reference Atmosphere (CIRA)

As it was already mentioned in the previous section one of the possible ways of improving quality of results is using more accurate models of the atmosphere. This section demonstrates the comparison between modelled radiances with the Air Force Geophysical Laboratory (AFGL) dataset 'midlatitude summer' (MLS) defined on 50 levels from 0 to 120 km that was used in all previous simulations and the dataset recommended by the Committee on Space Research (COSPAR) which is called 'COSPAR International Reference Atmosphere' (CIRA) defined on 71 levels from 0.1 to 119 km.

CIRA is a database with empirical models of atmospheric temperatures and pressures. According to [catalogue.ceda.ac.uk] the data in this database are retrieved from satellites, radiosonde and ground-based measurements with a 5-degree latitude grid and roughly 2 km vertical resolution. Geographical coordinates of methane shaft Pniowek V correspond to the 49° N. latitude, hence the closest CIRA grid point is 50° N. latitude. Moreover, CIRA dataset has the average data for each month, in contrast to the Air Force Geophysical Laboratory (AFGL) dataset which data has only two seasonal options: winter and summer. Obviously that the atmospheric data from CIRA database is much more detailed in comparison to collection of atmospheric model profiles from the AFGL dataset, e.g. midlatitude summer (MLS).

The following Figure 54 demonstrates the comparison of the temperature profiles from the original Air Force Geophysical Laboratory (AFGL) dataset, in particular midlatitude summer (MLS) - orange curve and original

data from Committee on Space Research (COSPAR), in particular COSPAR International Reference Atmosphere (CIRA) for 50° N, June (blue curve).

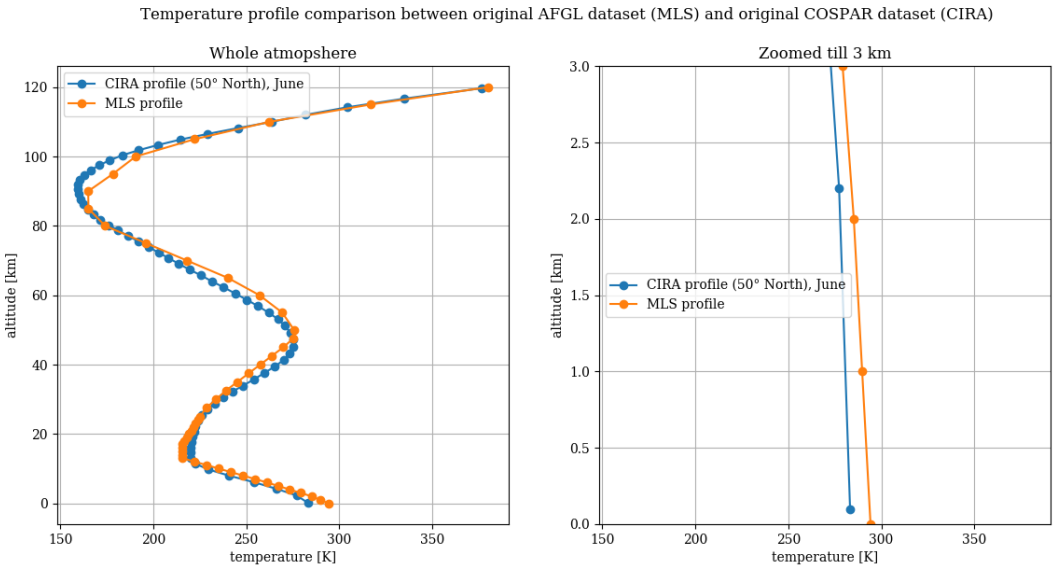


Figure 54 Comparison of the temperature profiles from the original AFGL dataset 'midlatitude summer' (MLS) and original COSPAR dataset 'COSPAR International Reference Atmosphere' (CIRA) for 50° N. latitude in June.

The following Figure 55 shows the comparison between modelled radiance for observations at 1500m with different constant albedo values with atmospheric data from MLS (dashed lines) and CIRA (solid lines) datasets.

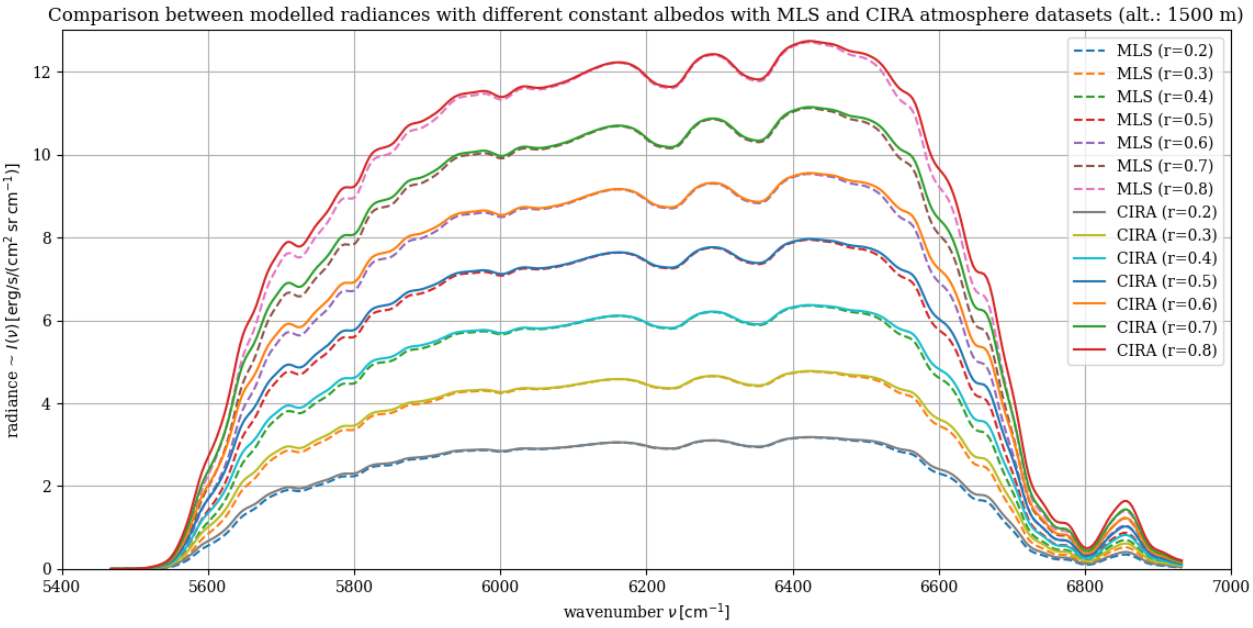


Figure 55 Comparison between modelled radiance for observations at 1500m with different constant albedos with atmospheric data from MLS and CIRA datasets.

It is clearly seen from Figure 55 that on the considering wavenumber interval in the vicinity of 6000 cm^{-1} the difference between modelled radiances based on MLS and CIRA atmospheric datasets is almost neglectable. Hence, CIRA atmospheric dataset will not dramatically improve previously obtained results.

7. Discussion

This chapter presents the analysis and interpretation of the results reported in previous chapters.

First, in chapter 4 'Py4CAtS results' output graphs from the different preparation steps of the simulations were presented. The educational tasks that were set at this stage of the work included the study of impacting various HITRAN database versions, diverse altitudes for ToA selection, divergent type of atmospheres (e.g. MLS and MLW), distinct empirical albedo types, difference in behavior of the modelled radiance with normal and enhanced methane concentration in the atmosphere etc. Generally speaking, all the data obtained in this chapter were consistent with the expected results or in other words the outcomes of simulations do not contradict the laws of atmospheric physics.

Then chapter 5 'Methane plume retrieval' presented outputs of the four methods that were suggested for plume detection. Three out of four methods did not show any noticeable trace of the plume, while thanks for the fourth method 'Ratio of absorbing and transparent residuals' one was able to see the plume of around 200 meters long and ~40 meters wide from Pniowek V methane shaft on both HySpex lines 9 (1520 m) and 11 (2900 m) and the weak plume of 40-50 meters long and 10-20 meters wide from Pniowek III shaft. The output graphs from this method are sensible to the wavenumber window. If we move considering wavenumber interval from the methane absorption range in the vicinity of 6000 cm^{-1} to e.g. 6050 cm^{-1} the final output pattern will differ and there will be no plume detected.

Chapter 6 'Py4CAtS methane plume radiance simulations' demonstrates various comparisons of the radiance simulations from constant and empirical albedo values with the HySpex observed radiance over methane plume in different spatial positions, e.g. in along-track, across-track directions and combination of both. An important conclusion can be made regarding the values of the observed radiances: the lower the measurement altitude – the more amplified the detected radiation, but from the other side, the larger methane enhancement in the atmosphere – the more detected radiation decreases. Also, in the same chapter results from the least-squares fitting of albedo values and methane scaling factor are demonstrated. The least-squares fitting was applied to two models: first with full ('standard' MLS) and optical depths of all gases till 1500 m and the second with half water concentration in the atmosphere and only methane optical depth. However, in both cases fitted reflectivity values at some pixels were larger than 1, which is physically unreasonable for the Lambertian or diffuse reflection model, but maybe reasonable for specular reflection. Another unforeseen result was regarding the methane scaling factor. On the corresponding graph one can see that a slightly noticeable trace, correlating with the position of the methane plume on the output from method IV graph, has lower values of the scaling factor compared to the surrounding landscape. In contrast to the result graph of method IV where the methane plume has a pronounced character on the HySpex line 9. Among other things in chapter 6 was presented comparison between two different atmospheric datasets MLS and CIRA as well as the modelled radiance based on these data comparison. Despite the apparent superiority of CIRA over MLS, the simulated radiances were almost the same at the methane absorption area around 6000 cm^{-1}.

Finally, it is useful to hold a comparison of the approaches and results for methane plume detection from this work with the closest related article of French research group from September 2020 by Nesme et al. (2020) who also used HySpex family airborne sensor. The HySpex-NEO sensor that was used by abovementioned research group has similar to DLR's HySpex spatial (1.4 m) and spectral resolution (6 nm for NEO and 4.75 nm for the sensor used in this work). The difference in approaches of plume detection between this thesis and French research group is that the latter one used 2.3 μm interval instead of 1.6 μm as in this work. Moreover, Nesme et al. (2020) implemented image processing algorithms (e.g. k-means clustering and CTMF) as a preprocessing step, while in this work all methods were directly operated with spectral data without any image processing steps. French research group was able to detect 60 m wide plume, while presented method IV from this work is able to detect plume with ~10-20 meters wide.

8. Conclusions

The main objective that was defined for this Master's work was focused on analyzing of the HySpex airborne infrared measurements for estimation of localized gas emissions. Methods and results that are presented in this thesis allows to detect methane plume over industrial sources in particular over methane shafts of the Upper – Silesian region. This thesis answers to the question that for the airborne methane plume detection it is possible to use data with high spatial, but low spectral resolution as the HySpex sensor has.

The further possible steps of the improvement of the results and its accuracy can be, for example:

- Application of detailed modeling of the atmosphere. Obvious trend of decreasing least-squares fitted reflectivity values with decreasing water concentration in the atmosphere was noticed in the work.
- Implementation of advanced filters in order to truncate spectral outliers on the measurement data as a preprocessing step.
- Usage of land cover classification datasets (e.g. Copernicus) for getting more accurate knowledge about what kind of empirical albedos can be used for radiance simulations.
- Coordination with methane shaft operators. Methane shafts are emitting discretely in time domain. A close exchange of information about shaft emission timetable will help to avoid the situation when the majority of the HySpex lines are without plumes due to non-emitting shafts at the moment of sensing.
- Consideration of wind datasets. In order to avoid situation when plume was 'blown' by wind out of the HySpex sensed areas (lines) one should take into account wind datasets (e.g. from LIDAR).

List of acronyms

ac	absorption coefficients
ADCS	Attitude Determination and Control System
AFCRL	Air Force Cambridge Research Laboratories
AFGL	Air Force Geophysical Laboratory
AIRS	The Atmospheric Infrared Sounder
AVIRIS-NG	Airborne Visible InfraRed Imaging Spectrometer-Next Generation
BIRRA	Beer InfraRed Retrieval Algorithm
BoA	Bottom of the Atmosphere
CCD	Charge Coupled Device
CCI	Climate Change Initiative
CIRA	COSPAR International Reference Atmosphere
CoMet	Carbon Dioxide and Methane Mission
COP	Coal Oil Point
COSPAR	Committee on Space Research
CTMF	Cluster-Tuned Matched Filter
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
ECVs	Essential Climate Variables
EnMAP	Environmental Mapping and Analysis Program
ENVISAT	Environmental Satellite
ESA	European Space Agency
EuFAR	European Facility for Airborne Research
FOV	Field of View
FWHM	Full Width at Half Maximum
GARLIC	Generic Atmospheric Radiation Line-by-line Infrared Code
GEISA	Gestion et Etude des Informations Spectrocopiques Atmospheriques (Management and Study of Atmospheric Spectroscopic Information)
GHG	GreenHouse Gas
GHOST	GreenHouse gas Observations of the Stratosphere and Troposphere
GWP	Global Warming Potential
GOSat	Greenhouse Gases Observing Satellite
HALO	High Altitude and Long Range Research Aircraft
HITRAN	High Resolution Transmission database
HWHM	Half Width at Half Maximum
HySpex	Hyper Spectral camera
IASI	Infrared Atmospheric Sounding Interferometer
IMAP-DOAS	Iterative Maximum A Posteriori Differential Optical Absorption Spectroscopy
IMF	Institut für Methodik der Fernerkundung (The Remote Sensing Technology Institute)
IMU	Internal Measurement Unit
IPA	Institut für Physik der Atmosphäre (Institute of Atmospheric Physics)
lbl	Line-by-line
LIDAR	Light Detection and Ranging
LOS	Line of Sight
MAMAP	Methane Airborne Mapper
MCT	Mercury Cadmium Telluride
MIR	Mid-infrared
MLS	Midlatitude Summer
MLW	Midlatitude Winter
M-SVM	Multi-class Support Vector Machine
NASA	National Aeronautics and Space Administration
NIR	Near Infrared
NLS	Nonlinear least-squares

NOAA	National Oceanic and Atmospheric Administration
Obs4MIPs	Observations for Model Intercomparison Projects
od	optical depths
OpAiRS	Optical Airborne Remote Sensing and Calibration Homebase
OZA	Observer Zenith Angle
PCA	Principle Component Analysis
Py4CAtS	Python for Computational ATmospheric Spectroscopy
RCNN	Regional Convolutional Neural Network
ri	radiation intensity
rms	root mean square
RTE	radiative transfer equation
SCIAMACHY	Scanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SNLS	Separable nonlinear least-squares
SRF	Spectral Response Function
SVD	Singular Value Decomposition
SWIR	Short-wave Infrared
SZA	Sun/Solar Zenith Angle
TIR	Thermal Infrared
ToA	Top of the Atmosphere
VMR	Volume Mixing Ratio
VNIR	Visible and Near-Infrared
WFM-DOAS	Weighting Function Modified Differential Optical Absorption Spectroscopy
xs	cross sections

Code samples

Code snippet for albedo interpolation

```
#Albedo interpolation.
#Code example for plotting Figure 27 “Comparison of the radiances in case of empirical
and constant albedo”. Code example only for dry grass albedo (other ground coverages -
analogically).

#Interpolation of the atmospheric grid to small altitudes (MLS):
mls=atmRead('D:/Alex/DLR/data/py4cats/atmos/20/mls.xy', zToA=80)  # ToA = 80 km
zOld_km = cgs('!km', mls['z'])               # the old grid converted to kilometers
zNew = insert(zOld_km,[1,2],[0.5,1.5])       # two extra grid points
MLS = atmRegrid(mls,zNew,'h')    # use cubic Hermite interpolation (default would be 2-
point Lagrange)

# read albedo data: given in percent!
lGrid3, rdrygrass = loadxy('D:/Alex/DLR/data/reflectances/drygras_reflectance.dat')

# now in absolute numbers
rdrygrass *= 0.01

vRange = Interval(5400,7000)
hwhm = 1.0  # spectral response function

dll = higstract('D:/Alex/DLR/data/hitran/2008/lines.txt',vRange+5*hwhm+5, 'main')
dodl = lbl2od(MLS,dll,vRange+hwhm) # monochrom in the extended interval
tod  = dod2tod(dodl)  # total atmosphere opt depth

# convolve with Gauss
todg = tod.convolve(hwhm,'G')
todgg = tod.convolve(15*hwhm,'G') #convolve with hwhm=15 (same as HySpex instrument reso
lution)

# save the corresponding wavenumber grids
uGrid = tod.grid()
vGrid = todg.grid()
wGrid = todgg.grid()

# flip data before interpolation!   (flip(xx)=xx[::-1])
#uGrid86=(x_new); 1e4/flip(lGrid1)=(x_old)
rdrygrassU = interp(uGrid, 1e4/flip(lGrid3),flip(0.01*rdrygrass))
rdrygrassV = interp(vGrid, 1e4/lGrid3[::-1],rdrygrass [::-1])  # mean 0.25
rdrygrassW = interp(wGrid, 1e4/lGrid3[::-1],rdrygrass [::-1])

T_sun=6000
rSun = 6.96e10
```

```
au=1.49598e+13
SZA = 30 #SZA=30 degrees

# monochromatic
I_sun_ToA = planck(uGrid,6000) *pi*((rSun/au)**2)
I_sun_BoA = I_sun_ToA *exp(-tod.base/cosdg(SZA))

# moderate resolution (hwhm =1.0)
I_sun_ToA_g = planck(vGrid,6000) *pi*((rSun/au)**2)
I_sun_BoA_g = I_sun_ToA_g *exp(-todg.base/cosdg(SZA))

# low resolution (hwhm=15.0)
I_sun_ToA_gg = planck(wGrid,6000) *pi*((rSun/au)**2)
I_sun_BoA_gg = I_sun_ToA_gg * exp(-todgg.base/cosdg(SZA))

zhs = 1.5  # km HySpex aircraft altitde
# define the number of layers from BoA up to the aircraft
lhs = np.searchsorted(cgs('!km',MLS['z']),zhs) # np.searchsorted(mls['z']),5e5)
# here lhs=5 because in this mls atmospheric dataset the lower atmosphere is sampled wit
h 1km steps!

od5   = dod2tod(dodl[:lhs]).regrid(len(tod))  # lower troposphere, regridded to make gri
d identical to tod grid
od5g  = od5.convolve(hwhm,'G')
od5gg = od5.convolve(15*hwhm,'G')

# radiance seen at aircraft:  drygrass
I_mono_drygrass = cosdg(SZA)/pi * rdrygrassU * I_sun_BoA * exp(-od5.base)
I_g01_drygrass  = cosdg(SZA)/pi * rdrygrassV * I_sun_BoA_g * exp(-od5g.base)
I_g15_drygrass = cosdg(SZA)/pi * rdrygrassW * I_sun_BoA_gg * exp(-od5gg.base)

# radiance seen at aircraft: all surfaces; constant albedo (r=0.5), hwhm=15
I_g15  = cosdg(SZA)/pi * 0.5 * I_sun_BoA_gg * exp(-od5gg.base)

# plotting radiances with const. and interpolated albedo with Hyspex measurements
plot(wGrid, I_g15, label='constant albedo (r=0.5)')
plot(wGrid, I_g15_conifer, label='Conifer (interpolated albedo)')
plot(wGrid, I_g15_deciduous, label='Deciduous (interpolated albedo)')
plot(wGrid, I_g15_drygrass, label='Dry grass (interpolated albedo)')
plot(wGrid, I_g15_grass, label='Grass (interpolated albedo)')
plot(wGrid1,rad1, 'g+', label='HySpex measurement #1')
plot(wGrid2,rad2, 'r+', label='HySpex measurement #2')
plot(wGrid3,rad3, 'b+', label='HySpex measurement #3')
legend()
xlabel (r'wavenumber  $\nu \rm\,[cm^{-1}]$')
xlim(5400,7000)
ylabel (r'radiance  $I(\nu) \rm \; [erg/s/(cm^2\:sr\:cm^{-1})]$')
ylim(0,11)
```

```
title ("Modelled radiances with interpolated & constant albedo with HySpex measurements"
)
```

Code snippet for plume retrieval Method I

```
#HySpex measurements
paw11 = numpy.load("E:\DLR\hyspex_measurements\pawlowice\pawlowice_scanline11_hyspex.npy
")
#Reading actual wavenumber assignment of the 320x256 detector pixels
wavenumbers = numpy.loadtxt("E:\DLR\hyspex_measurements\wavenumbers_hyspex.txt")
#1.6 mue CH4 data lines (from HySpex Line 11):
data1= (paw11[:,:,117]) #5991-5992 cm-1 (lesft border of enhanced CH4)
data2= (paw11[:,:,116]) #6012-6014 cm-1 (right border of enhanced CH4)
mean1= (1/2)*(data1+data2)

#Reference (transparent) atmosphere 1.6 mue (6477-6557 cm-1):
data3= (paw11[:,:,96])
data4= (paw11[:,:,95])
data5= (paw11[:,:,94])
data6= (paw11[:,:,93])
mean3 = (1/4)*(data3+data4+data5+data6)

#Example of the visualization of the results for HySpex line 11 (for line 9-
analogically):
fig=plt.figure(figsize=(5, 6))
fig.suptitle('Spectrally avg. CH4 interval (5991-6014 cm-1) - spectrally avg.
transparent atmosphere (6477-6557 cm-1), Pniowek V (lines: 9 and 11)')
subplot(131)
cmap2=contourf(mean1-mean3, levels=np.arange(start=-20, stop=6.1, step=0.1),
cmap='rainbow')
xlabel (r'across track $\rm\,[pixels]$')
plt.gca().invert_xaxis()
ylabel (r'along track $\rm\,[pixels]$')
ylim(4600,5000)
cbar = plt.colorbar()
cbar.set_label('radiance [erg/s/(cm^2 sr cm^(-1)]')
title('Line 11 (alt.: 2900 m.)')
# Show the major grid lines
plt.grid(b=True, which='major', color='#666666', linestyle='-')
# Show the minor grid lines
plt.minorticks_on()
plt.grid(b=True, which='minor', color='#666666', linestyle='-', alpha=0.4)
```

Code snippet for plume retrieval Method II

```
#Reading and loading of the data are done in the same way as in Method I. The only
difference is in the visualization of the results:
#Example of the visualization of the results for HySpex line 11 (for line 9-
analogically):
fig=plt.figure(figsize=(5, 6))
fig.suptitle('Ratio of the spectrally avg. CH4 interval (5991-6014 cm-1) and spectrally
avg. transparent atmosphere (6477-6557 cm-1), Pniowek V (lines: 9 and 11)')
subplot(131)
cmap2=contourf(mean1/mean3, levels=np.arange(start=0.2, stop=1.7, step=0.1),
cmap='rainbow')
xlabel (r'across track $\rm\,[pixels]$')
plt.gca().invert_xaxis()
ylabel (r'along track $\rm\,[pixels]$')
ylim(4600,5000)
cbar = plt.colorbar()
title('Line 11 (alt.: 2900 m.)')
# Show the major grid lines
plt.grid(b=True, which='major', color='#666666', linestyle='-')
# Show the minor grid lines
plt.minorticks_on()
plt.grid(b=True, which='minor', color='#666666', linestyle='-', alpha=0.4)
```

Code snippet for plume retrieval Method III

```
#this code part show the most important steps for the element-
wise spatial averaging of 2D numpy arrays using 'signal.convolve2d' function. 2D
convolution for the HySpex line 11 (for line 9-analogically)

import numpy as np
import matplotlib.pyplot as plt
import matplotlib.image as mpimg

#HySpex measurements
paw11 = numpy.load("E:\DLR\hyspex_measurements\pawlowice\pawlowice_scanline11_hyspex.npy
")

#Reading actual wavenumber assignment of the 320x256 detector pixels
wavenumbers = numpy.loadtxt("E:\DLR\hyspex_measurements\wavenumbers_hyspex.txt")

#1.6 mue CH4 data lines (from HySpex Line 11):
data1= (paw11[:,:,117]) #5991-5992 cm-1 (lesft border of enhanced CH4)
data2= (paw11[:,:,116]) #6012-6014 cm-1 (right border of enhanced CH4)
mean1= (1/2)*(data1+data2)
```

```
#Reference atmosphere 1.6 mue (6477-6557 cm-1):
data3= (paw11[:,:,96])
data4= (paw11[:,:,95])
data5= (paw11[:,:,94])
data6= (paw11[:,:,93])
mean3 = (1/4)*(data3+data4+data5+data6)

#Calculation of the element-wise mean values (2D convolution):
from scipy import signal

x = mean1 #filtered with standard radiance boundaries at 2.3 mue spectrally averaged CH4
 bands, line 11
y = np.ones((5,5)) #5*5 neighbor elements block

# We convolve x with y and then we normalize those value with another convolution
# |first convolution (sum of desired elements) |/| second convolution (quantify how many
 elements are sumed up)
r = signal.convolve2d(x,y,mode='same')/signal.convolve2d(np.ones(x.shape),y,mode='same')

#Code snippet for the visualization of the results
fig=plt.figure(figsize=(5, 6))
fig.suptitle('Ratio of the spatially (5*5) & spectrally avg. CH4 interval (5991-6014 cm-
1) and only spectrally avg. transparent atmosphere (6477-6557 cm-1), Pniowek V')
cmap2=contourf(r/mean3, levels=np.arange(start=0, stop=5.5, step=0.1), cmap='rainbow')
xlabel (r'across track $\rm\,[pixels]$')
plt.gca().invert_xaxis()
ylabel (r'along track $\rm\,[pixels]$')
ylim(4600,5000)
cbar = plt.colorbar()
title('Line 11 (alt.: 2900 m.)')
# Show the major grid lines
plt.grid(b=True, which='major', color='#666666', linestyle='-')
# Show the minor grid lines
plt.minorticks_on()
plt.grid(b=True, which='minor', color='#666666', linestyle='-', alpha=0.4)
```

Code snippet for plume retrieval Method IV

```
# PREPARATION
# define the models for the polynomial fit for the out-of-band pixels
def quadratic_model(wn,coeffs):
    """"""
    return coeffs[0] + coeffs[1]*wn + coeffs[2]*wn**2

def linear_model(wn,coeffs):
```

```
    return coeffs[0] + coeffs[1]*wn

def const_model(wn,coeffs):
    return coeffs[0] + 0.*wn

# define the (objective) function to be optimized
# which function to choose depends on the behavior of the out-of
band radiances in the range you examine
def residual_quadratic(variables, wn, data):
    """Use this if the out-of-band pixel show a quadratic dependence
    of the radiances with wavelength (check visually before anaylsing)"""

    model = quadratic_model(wn,variables)
    return data-model

def residual_linear(variables, wn, data):
    """Use this if the out-of-band pixel show a linear dependence
    of the radiances with wavelength (check visually before anaylsing)"""

    model = linear_model(wn,variables)
    return data-model

def residual_constant(variables, wn, data):
    """Use this if the out-of-band pixel show (almost) no dependency
    of the radiances with wavelength (check visually before anaylsing)"""

    model = const_model(wn,variables) # essentially the mean
    return data-model

from scipy.optimize import leastsq
from scipy import interpolate

# ANALYSIS
# read data and wavenumber assignments
data = np.load("E:\DLR\hyspex_measurements\pawlowice\pawlowice_scanline09_hyspex.npy")
wn = np.loadtxt("E:\DLR\hyspex_measurements\wavenumbers_hyspex.txt", unpack=False)

# define region of interest
bounds = [5950,6050] # wavenumber boundaries for the analysis
central_wn = 6000. # roughly the central wavenumber of the interval
alongtrack = [6400,7000] # bounds for which along track pixels to use
# apply region of interest on wavenumber dataset
wavenumber = wn[0][np.logical_and(wn[0]>bounds[0],wn[0]<bounds[1])] - central_wn
wavenumber_out = np.array([wavenumber[0],wavenumber[-1]]) # out-of-band wavenumbers
wavenumber_in = np.array([wavenumber[1],wavenumber[2]]) # in-
band wavenumbers (where CH4 is absorbing)

# auxilary variables
```

```
l = alongtrack[0]
u = alongtrack[1]
# create arrays beforehand:
residuum_in = np.zeros_like(data[l:u,:,0])
residuum_out = np.zeros_like(data[l:u,:,0])
residuum_in_out = np.zeros_like(data[l:u,:,0])
ds = np.zeros_like(data[l:u,:,0])

# separate the region of interest in out-of-band data and in-band data
dataset = data[l:u,:,np.logical_and(wn[0]>bounds[0],wn[0]<bounds[1])]
ds = dataset.mean(axis=2)
dataset_out = dataset[:,:,0:4:3] # create array with only out-of-band values
dataset_in = dataset[:,:,1:3] # create array with only in-band values
coeff_out = np.median(dataset_out, axis=2)

# fit the out-of-band pixels and determine the of the out-of-band and in-
band distance to it (2-norm)
for j in range(u-l): # along track
    for i in range(320): # cross track
        # fit coefficients for the out-of-band pixels
        leastsq_coeff_out = leastsq(residual_constant, coeff_out[j,i], args=(wavenumber_
out,dataset_out[j,i])) # constant model is used here
        # interpolate to out-of-band wavenumbers
        p = interpolate.interp1d(wavenumber_out, const_model(wavenumber_out,leastsq_coef
f_out)) # check that model matches the one that is used in the previous line
        # evaluate polynomial p at the out-of-band wavenumbers
        datamodel_out = p(wavenumber_out)
        # determine the out-of-band residual
        residuum_out[j,i] = np.linalg.norm(datamodel_out-dataset_out[j,i])
        # evaluate polynomial p at the in-band wavenumbers
        datamodel_in = p(wavenumber_in)
        # determine the in-band residual
        residuum_in[j,i] = np.linalg.norm(datamodel_in-dataset_in[j,i])

# calc the ratio which should indicate if gas absorption takes place
residuum_in_out = residuum_in/residuum_out

#VISUALIZATION
from matplotlib import pyplot as plt
plt.contourf(residuum_in_out, levels=np.arange(start=residuum_in_out.min(), stop=3*np.me
dian(residuum_in_out), step=0.1),cmap='hot_r')
plt.yticks(ticks=np.arange(0,u-l,200), labels=np.arange(l,u,200))
cbar = plt.colorbar()
```

Code snippet for least-squares fitting of albedo and methane scaling factor

```
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.image as mpimg

#Interpolation of the atmospheric grid to small altitudes (MLS):
mls=atmRead('D:/Alex/DLR/data/py4cats/atmos/20/mls.xy', zToA=80)  # ToA = 80 km
zOld_km = cgs('!km', mls['z'])               # the old grid converted to kilometers
zNew = insert(zOld_km,[1,2],[0.5,1.5])       # two extra grid points
MLS = atmRegrid(mls,zNew,'h')

#Case of atmosphere with halfed water concentration:
MLS_hlf_H2O = MLS.copy() #copy the interpolated atmopshere
MLS_hlf_H2O['H2O'] *= 0.5 #half the concentration of H2O in copied atmosphere

T_sun=6000
rSun = 6.96e10
au=1.49598e+13
SZA = 30 #SZA=30 degrees

vRange = Interval(5850,6150)
hwhm = 15.0
dll = higstract('D:/Alex/DLR/data/hitran/2008/lines.txt',vRange+5*hwhm+hwhm, 'main')
dodl = lbl2od(MLS_hlf_H2O,dll,vRange+5*hwhm) # monochrom in extended interval for gauss
tod  = dod2tod(dodl)  # total atmosphere opt depth
todg = tod.convolve(hwhm,'G')
vGrid = todg.grid()

jCross=231;  kLong=6579 #across-track and along-track coordinates of observing pixel
wGrid = numpy.loadtxt("E:\DLR\hyspex_measurements\wavenumbers_hyspex.txt")[jCross,::-1]
dataObs = numpy.load("E:\DLR\hyspex_measurements\pawlowice\pawlowice_scanline09_hyspex.n
py")
radObs = dataObs[kLong,jCross,::-1]
iLow = searchsorted(wGrid,vRange.lower)
iHigh = searchsorted(wGrid,vRange.upper)

# prepare model
radPerfect  = cosdg(SZA)/pi * 1.0 * I_sun_BoA_g * exp(-
od5g.base) # "ideal" radiance seen at aircraft: full reflectivity (r=1.0)
radMod1 = interp(wGrid[iLow:iHigh], vGrid, radPerfect)

##############################################################################
#Least-squares fitting for methane scaling factor (beta):
##############################################################################
# prepare model for phi1
dodl_CH4 = lbl2od(MLS_hlf_H2O,dll['CH4'],vRange+5*hwhm)
```

```
tod_CH4  = dod2tod(dodl_CH4)  # total methane opt depth
zhs = cgs('km',0.5)  # (km) methane plume altitde 0-500m
lhs = np.searchsorted(MLS_hlf_H2O['z'],zhs)

od_CH4 = dod2tod(dodl_CH4[:lhs]).regrid(len(tod))    #od_CH4 = dodl_CH4[0] (it is optica
l depth of CH4 on 1-st atmospheric layer)
odg_CH4 = od_CH4.convolve(hwhm,'G')
odg_CH4_rg = interp(wGrid[iLow:iHigh], vGrid, odg_CH4)

# model basis functions
phi0 = radMod1
phi1 = phi0*odg_CH4_rg
aMatrix = np.vstack([phi0, phi1]).T #creating a matrix for least-square fitting
r0, r0_beta = np.linalg.lstsq(aMatrix, radObs[iLow:iHigh])[0]

#visualization of the least-
squares fitted albedo values (r0) or CH4 scaling factor (beta) is done with contourf fun
ction
```

Code snippet of the for-loop for least-squares fitting

```
#iteration 'for loop' for area of 300x600 pixels around shaft Pniowek V:
r0_res = np.ones((300,600))
r1_res = np.ones((300,600))

jCross = 0
for jCross in range(0, 300):
    kLong = 0
    for kLong in range(6400, 7000):
        radObs = dataObs[kLong,jCross,::-1]
        iLow = searchsorted(wGrid,vRange.lower)
        iHigh = searchsorted(wGrid,vRange.upper)
        odg_CH4_rg = interp(wGrid[iLow:iHigh], vGrid, odg_CH4)
        phi0 = radMod1
        phi1 = phi0*odg_CH4_rg
        aMatrix = np.vstack([phi0, phi1]).T
        r0, r0_beta = np.linalg.lstsq(aMatrix, radObs[iLow:iHigh])[0]
        if jCross<= 300:
            r0_res[jCross, kLong-6400]= r0
            r1_res[jCross, kLong-6400]= r0_beta
        kLong=kLong+1
    jCross=jCross+1
```

References

Anderson G., S. Clough, F. Kneizys, J. Chetwynd and E. Shettle, 1986. AFGL Atmospheric Constituent Profiles (0-120 km). Technical Report TR-86-0110, AFGL: Hanscom AFB, MA, USA: 1-14.

Atmospheric methane from satellites, Retrieved 2020, Greenhouse gases CCI: http://cci.esa.int/ghg

Badura H. and Z. Lukaszczyk, 2020. Analysis of methane emissions from air shafts of the Jastrzebska coal mining company in 2017-2019. MAPE 2020, vol. 3, issue 1: 435-449.

Burrows J.P., P. Ulrich and P. Borrell, 2011. The remote sensing of tropospheric composition from space, Springer, Heidelberg: 1-152.

Butler M.J.A., M.C. Mouchot, V. Barale, C. LeBlanc, 1988. The application of remote sensing technology to marine fishers: an introductory manual. FAO Fish. Tech. Pap., 295: 23-38.

Deutsches Zentrum für Luft- und Raumfahrt/ German Aerospace Center, Retrieved 2020, Institute of Atmospheric Physics: https://www.dlr.de/content/en/institutes/institute-of-atmospheric-physics.html

Deutsches Zentrum für Luft- und Raumfahrt/ German Aerospace Center, Tracking methane with an airplane, Retrieved 2018, News/Archive: https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-12632/22039_read-52329

Dlugokencky E.J., A.M. Crotwell, K.W. Thoning and J.W. Mund, 2020. Atmospheric methane from quasi-continuous measurements at Barrow, Alaska and Mauna Loa, Hawaii, 1986-2019, Version: 2020-03, https://doi.org/10.15138/ve0c-be70, FTP path: ftp://aftp.cmdl.noaa.gov/data/greenhouse_gases/ch4/in-situ/surface/

Encyclopedia Britannica, 2012. https://www.britannica.com/science/methane-cycle

Funk C., J. Theiler, D.A. Roberts and C. Borel, 2001. Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal energy. *IEEE transactions on geoscience and remote sensing*, 39(7): 1410-1420.

Gharibi W. and O.S. Al-Mushayt, 2011. A note on separable nonlinear least squares problem. Jazan University, Jazan, KSA, DOI: 10.1109/ICFCSA.2011.19: 1-3.

Gimeno Garcia S., F. Schreier, G. Lichtenberg and S. Slijkhuis, 2011. Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmospheric Measurement Techniques*, 4: 2633-2657.

Gottwald M. and H. Bovensmann, eds., 2011. SCIAMACHY - Exploring the Changing Earth's Atmosphere. Earth and Environment Science. Springer, Dordrecht, Heidelberg, London, New York: 1-225.

Hochstaffl P., F. Schreier, G. Lichtenberg and S. Gimeno Garcia, 2018. Validation of carbon monoxide total column retrievals from SCIAMACHY observations with NDACC/TCCON ground-based measurements. *Remote Sensing*, vol.10, issue 2: 1-32.

Humpage N., H. Boesch, P.I. Palmer, A. Vick, P. Parr-Burman, M. Wells, D. Pearson, J. Strachan and N. Bezawada, 2018. Greenhouse gas observations of the stratosphere and troposphere (GHOST): an airborne shortwave-infrared spectrometer for remote sensing greenhouse gases. *Atmospheric Measurement Techniques*, 11: 5199-5222.

Jain Atul K., P. Briegleb, K. Minschwaner and J. Wuebbles, 2000. Radiative forcings and global warming potentials of 39 greenhouse gases. *Journal of Geophysical Research: Atmospheres*, 105(D16): 20773-20790.

Kahaner D., C. Moler and S. Nash, 1989. Numerical methods and software. Englewood Cliffs, Prentice Hall, New Jersey: 190-235.

Kaufman L. and V. Pereyra, 1978. A method for separable nonlinear least squares problems with separable nonlinear equality constraints, *SIAM Journal on Numerical Analysis*, vol. 15(1): 12-20.

Kedzior S. and M. Dreger, 2019. Methane occurrence, emissions and hazards in the Upper Silesian Coal Basin, Poland. *International Journal of Coal Geology;* vol. 211, ISSN: 0166-5162.

Köhler C.H., 2016. Airborne Imaging Spectrometer HySpex. *Large-scale research facilities,* 2, A93: 1-6.

Krings T., K Gerilowski, M. Buchwitz, M. Reuter, A. Tretner, J. Erzinger, D. Heinze, U. Pflüger, J.P. Burrows and H. Bovensmann, 2011. MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: retrieval algorithm and first inversions for point source emission rates. *Atmospheric Measurement Techniques,* 4: 1735-1758.

Kumar S., C. Torres, O. Ulutan, A. Ayasse, D. Roberts and B.S. Manjunath, 2020. Deep remote sensing methods for methane detection in overhead hyperspectral imagery, *IEEE Xplore*: 1776-1785.

Liuzzi G., M.G. Blasi, G. Masiello, C. Serio and S. Venafra, 2017. All-sky radiative transfer calculations for IASI and IASI-NG: the σ-IASI-as code, *AIP Conference*, 1810, 040004, published online: https://doi.org/10.1063/1.4975506

Nesme N., P.-Y. Foucher and S. Doz, 2020. Detection and quantification of industrial methane plume with the airborne HuSpex-neo camera and applications to satellite data, *XXIV ISPRS Congress*; vol. XLIII-B3-2020: 821-827.

NVIDIA ARC GmbH, Retrieved 2015: http://help.autodesk.com/cloudhelp/2016/ENU/mental-raydocs/mr_docs/shaders/layering/layering.html

Ranjan S., D. R. Nayak, K. S. Kumar, R. Dash and B. Majhi, 2017. Hyperspectral image classification: A k-means clustering based approach, *4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore: 1-7.

Rodgers C.D., 2000. Inverse methods for atmospheric sounding: Theory and practice, World Science, London: 1-253.

Rothman S., 1988. Modeling of the atmosphere. SPIE, 928: 1-231.

Schreier F., 2019. Atmosphärische IR-Fernerkundung. Strahlungstransport und Inversion, *CCG-Seminar SE 1.02*: 1-61.

Schreier, F., S. Gimeno García, P. Hedelt, M. Hess, J. Mendrok, M. Vasquez and J. Xu, 2014. GARLIC a general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 137: 29 – 50.

Schreier F., S. Gimeno Garcia, P. Hochstaffl and S. Städt, 2019. Py4CAtS – Python for Computational ATmospheric Spectroscopy. *Atmosphere*, 262(10): 1-26.

Schreier F., G. Wagner and M. Birk, 2006. Molecular absorption, DLR – Remote Sensing Technology Institute: 1-6.

Thompson D.R., A. Karpatne, I. Ebert-Uphoff, C. Frankenberg, A.K. Thorpe, B.D. Bue and R.O. Green, 2017. IS-GEO Dataset JPL-CH4-detection-2017-V1.0: A benchmark for methane source detection from imaging spectrometer data, California Institute of Technology, Jet Propulsion Laboratory:1-15.

Thorpe A.K., C. Frankenberg and D.A. Roberts, 2014. Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS. *Atmospheric Measurement Techniques,* 7: 491-506.