

**Multi-Model Ensemble Wake Vortex
Prediction**

Stephan Körner

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Physik der Atmosphäre
Oberpfaffenhofen

Dissertation an der
Rheinisch-Westfälischen
Technischen Hochschule Aachen



Herausgeber

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Bibliotheks- und
Informationswesen
D-51170 Köln
Porz-Wahnheide
Linder Höhe
D-51147 Köln

Telefon

(0 22 03) 6 01 - 44 44

Telefax

(0 22 03) 6 01 - 47 47

Als Manuskript gedruckt.
Abdruck oder sonstige Verwendung
nur nach Absprache mit dem DLR gestattet.

ISSN 1434-8454
D 82 (Diss. RWTH Aachen University, 2017)

Wake vortices, prediction, Multi-Model Ensemble, Direct Ensemble Averaging, Reliability Ensemble Averaging, Bayesian Model Averaging, wake vortex measurements

Stephan KÖRNER

Institut für Physik der Atmosphäre, DLR, Oberpfaffenhofen

Multi-Model Ensemble Wake Vortex Prediction

Doctoral Thesis Rheinisch-Westfälische Technische Hochschule Aachen

DLR-Forschungsbericht 2017-44, 2017, 155 pages, 96 figs., 33 tabs., 162 refs, 36,00 €

As a response to lift, a complex flow pattern is shed from the wings of an aircraft that evolves into a pair of counter-rotating vortices. Due to the rolling momentum that the vortices may induce and the forces that they can exert they pose a potential hazard to following air traffic. To avoid dangerous incidents fast-time wake vortex models have been developed in the past that predict the vortex position and strength dependent on aircraft parameters and ambient conditions. This thesis investigates the capability to further enhance wake vortex forecasts by combining various independent models in a Multi-Model Ensemble, which has shown to increase the deterministic forecast skill and generate reliable probabilistic predictions in other applications. Therefore the fast-time models of NASA (APA 3.2, APA 3.4, TDP 2.1) and DLR (D2P) are combined by the methods Direct Ensemble Averaging, Reliability Ensemble Averaging and Bayesian Model Averaging. The ensemble output is validated with wake vortex measurements and compared to the respective best model.

Wirbelschleppen, Vorhersage, Multi-Model Ensemble, Direct Ensemble Averaging, Reliability Ensemble Averaging, Bayesian Model Averaging, Wirbelschleppenmessungen

(Published in English)

Stephan KÖRNER

Institut für Physik der Atmosphäre, DLR, Oberpfaffenhofen

Multi-Model Ensemble Wirbelschleppenvorhersagen

Dissertation Rheinisch-Westfälische Technische Hochschule Aachen

DLR-Forschungsbericht 2017-44, 2017, 155 Seiten, 96 Bilder, 33 Tabellen, 162 Literaturstellen, 36.00 € zzgl. MwSt.

Als Folge des Auftriebs entsteht hinter den Tragflächen eines Flugzeugs ein komplexes Strömungsfeld, welches sich im Weiteren zu einem gegensinnig rotierenden Wirbelpaar formiert. Aufgrund des Rollmomentes, welches die Wirbel induzieren können, stellen sie eine potentielle Gefahr für den nachfolgenden Luftverkehr dar. Um gefährliche Zwischenfälle zu vermeiden, wurden in der Vergangenheit Wirbelschleppenmodelle entwickelt, welche in Abhängigkeit der meteorologischen Bedingungen und der Flugzeugparameter die Position und Stärke der Wirbel vorhersagen können. Diese Arbeit untersucht, ob die Vorhersage durch Kombination verschiedener unabhängiger Modelle in einem Multi-Model Ensemble weiter verbessert werden kann. Dieser Ansatz konnte bei anderen Anwendungen die Genauigkeit der deterministischen Vorhersagen erhöhen und verlässliche probabilistische Vorhersagen generieren. Dazu werden Modelle der NASA (APA 3.2, APA 3.4, TDP 2.1) und des DLR (D2P) mit den Methoden Direct Ensemble Averaging, Reliability Ensemble Averaging und Bayesian Model Averaging kombiniert. Die Ensemble-Vorhersagen werden anhand von Wirbelschleppenmessungen validiert und dem jeweils besten Modell gegenübergestellt.

Forschungsbericht 2017-44

Multi-Model Ensemble Wake Vortex Prediction

Stephan Körner

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Physik der Atmosphäre
Oberpfaffenhofen

Dissertation an der
Rheinisch-Westfälischen
Technischen Hochschule Aachen

155 Seiten
96 Bilder
33 Tabellen
162 Literaturstellen



DLR

Deutsches Zentrum
für Luft- und Raumfahrt

Multi-Model Ensemble Wake Vortex Prediction

Multi-Model Ensemble Wirbelschleppenvorhersagen

Von der Fakultät für Maschinenwesen der Rheinisch-Westfälischen Technischen Hochschule
Aachen zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften
genehmigte Dissertation

vorgelegt von

Stephan Körner

Berichter: Univ.-Prof. Dr.-Ing. Eike Stumpf
apl. Univ.-Prof. Dr.-Ing. habil. Christian Breitsamter

Tag der mündlichen Prüfung: 25.07.2017

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Für meinen Großvater.

Eidesstattliche Erklärung

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Dissertation mit dem Titel "Multi-Model Ensemble Wake Vortex Prediction" selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen. Zudem versichere ich Eides statt, dass ich kein erfolgloses Promotionsverfahren bzw. Promotionsstudium zum Dr.-Ing/ Dr.rer.nat. durchgeführt habe.

Abstract

As a response to lift, a complex flow pattern is shed from the wings of an aircraft that evolves into a pair of counter-rotating vortices. Their behavior is strongly influenced by the atmosphere and further underlies a complex interaction with the ground. Due to the rolling momentum that the vortices may induce and the forces that they can exert they pose a potential hazard to following air traffic, especially along the glide path and in ground proximity. To avoid dangerous incidents, separation regulations based on aircraft mass exist that is currently under revision in the context of RECAT. However, air traffic is expected to further increase in the future, making wake vortex encounters more likely. For this reason fast-time wake vortex models have been developed in the past that predict the vortex position and strength dependent on the aircraft parameters and the ambient conditions in order to avoid dangerous situations. Furthermore, they may reduce flight delays at congested airports as they allow adapting overly conservative separations tactically under certain conditions. However, forecasts include uncertainties that originate from inaccurate and highly variable initial and ambient conditions as well as from inadequate understanding and simplification of the underlying physics. As a consequence, deterministic forecasts are to be complemented by probabilistic envelopes.

In the field of meteorology such envelopes are frequently computed by combining the forecasts of multiple independent models to quantify the model uncertainty. This approach, also known as Multi-Model Ensemble (MME), has not only shown to increase the reliability of the probabilistic forecast but also to enhance its deterministic skill. For this reason the capability of the MME approach to improve both deterministic and probabilistic wake vortex forecasts is assessed and existing approaches are further developed in this thesis. Three of the ensemble members have been provided in the frame of a NASA-DLR cooperation. The employed models comprise APA 3.2, APA 3.4 and TDP 2.1 (NASA) and are combined with the DLR model D2P. Several MME approaches are examined, comprising the Direct Ensemble Average, the Reliability Ensemble Averaging (REA) and the Bayesian Model Average (BMA), where the BMA outperforms the other ensemble methods. It is found that the model deviations from observations increase with time and that only a temporal treatment of model uncertainty enables reliable probabilistic envelopes at any vortex age.

Data collected during campaigns at the airports of Munich, Frankfurt, special airport Oberpfaffenhofen, Memphis and Dallas serve to train and to evaluate the ensemble predictions. This research concludes that a well balanced MME based on the BMA may indeed outperform the best individual model predictions in a deterministic manner by up to 3.3%. Additionally, it is demonstrated that the BMA approach is also beneficial to probabilistic forecasts where the agreement between selected and observed probability level deviates by less than 4%. This demonstrates that the forecast uncertainty can be adequately predicted.

Moreover, this thesis aims at a further analysis of wake vortex physics, at confirming findings from Large-Eddy Simulations and at further improving the forecast of the DLR model D2P.

For this purpose newly captured wake vortex measurements that comprise a variety of vortex generation altitudes are evaluated. Evidence is found that the vortex-ground interaction already starts at higher altitudes than expected. It is also demonstrated that vortex decay is accelerated for vortices generated at low altitudes. Eventually, it is shown that the decay of the lee vortex is enhanced by increasing crosswind.

Kurzfassung

Als Folge des Auftriebs entsteht hinter den Tragflächen eines Flugzeugs ein komplexes Strömungsfeld, welches sich zu einem gegensinnig rotierenden Wirbelpaar entwickelt. Das Verhalten der beiden Wirbel hängt stark von den atmosphärischen Bedingungen ab und ist zusätzlich durch eine komplexe Interaktion mit dem Boden geprägt. Aufgrund des Rollmomentes, welches die Wirbel induzieren können, stellen sie besonders im Gleitpfad und in Bodennähe eine potentielle Gefahr für den nachfolgenden Luftverkehr dar. Um gefährliche Zwischenfälle zu vermeiden wurden Standards für die Staffeln an Flughäfen eingeführt, welche momentan im Rahmen von RECAT überarbeitet werden. Allerdings wird erwartet, dass der weltweite Flugverkehr weiter steigen und damit die Wahrscheinlichkeit von Wirbelschleppeneinflügen erhöht wird. Deshalb wurden in der Vergangenheit Wirbelschleppenmodelle entwickelt, welche in Abhängigkeit der meteorologischen Bedingungen und der Flugzeugparameter die Position und Stärke der Wirbel vorhersagen können. Ihre Prognosen können sowohl zu einer erhöhten Sicherheit als auch zu geringeren Verspätungen auf überfüllten Flughäfen beitragen, indem sie unter bestimmten Bedingungen eine Anpassung nicht ausreichender oder übermäßig konservativer Stafflungsabstände ermöglichen. Die Vorhersagen beinhalten allerdings Unsicherheiten, welche sowohl von nicht ausreichendem Verständnis der Physik als auch von stark schwankenden Anfangs- und Umgebungsbedingungen stammen. Aus diesem Grund ist es sinnvoll deterministische Prognosen mit probabilistischen Einhüllenden zu ergänzen.

Für deren Berechnung werden in der Meteorologie oft die Vorhersagen verschiedener unabhängiger Einzelmodelle kombiniert um so die Modellunsicherheit zu quantifizieren. Dieser Ansatz wird auch als Multi-Model Ensemble (MME) bezeichnet. Er hat bewiesen, dass er sowohl die Verlässlichkeit der probabilistischen Vorhersagen als auch die Genauigkeit der deterministischen Prognosen steigern kann.

Aus diesem Grund wird in dieser Dissertation untersucht, ob die deterministischen und probabilistischen Prognosen durch speziell angepasste MME-Ansätze verbessert werden können. Die im Rahmen einer Kooperation zwischen NASA und DLR zur Verfügung gestellten drei Amerikanischen Wirbelschleppenmodelle APA 3.2, APA 3.4 und TDP 2.1 werden dafür mit dem DLR Modell D2P kombiniert. Als verschiedene Ansätze werden der Direct Ensemble Average (DEA), das Reliability Ensemble Averaging (REA) und das Bayesian Model Average (BMA) untersucht. Als am geeignetsten stellt sich dabei der BMA-Ansatz heraus. Es zeigt sich, dass die Abweichungen der Modellvorhersagen von den Beobachtungen mit dem Wirbelalter zunehmen, sodass nur eine zeitabhängige Ensemble-Parametrisierung für durchgehend alle Wirbelalter verlässliche Prognosen liefert.

Um das Ensemble zu trainieren und zu testen werden Wirbelmessungen von den Flughäfen München, Frankfurt, Oberpfaffenhofen, Memphis und Dallas genutzt. Die Untersuchungen zeigen, dass ein gut ausbalanciertes BMA-Ensemble tatsächlich im Schnitt um 3.3% bessere deterministische Vorhersagen als das beste Einzelmodell treffen kann. Außerdem wird aufgezeigt,

dass durch den BMA Ansatz probabilistische Vorhersagen erzeugt werden können, deren Eintreten zu weniger als 4% vom beobachteten Wirbelverhalten abweicht.

Des Weiteren beschäftigt sich diese Thesis mit der Analyse der Wirbelphysik anhand neu ausgewerteter Daten, mit deren Vergleich mit Large-Eddy Simulationen und mit der Verbesserung der Vorhersagen des DLR Modells D2P. Dazu werden die gemessenen Landungen in verschiedene Wirbelgenerierungshöhen eingeteilt. Daraus ergeben sich Hinweise, dass die Interaktion der Wirbel mit dem Boden bereits in größeren Höhen als bisher vermutet beginnt. Außerdem wird gezeigt, dass der Wirbelzerfall niedrig erzeugter Wirbel schneller erfolgt als bei höher erzeugten. Zudem wird bestätigt, dass der Lee-Wirbel einen schnelleren Zerfall als der Luv-Wirbel aufweist.

Contents

Abstract	vii
Kurzfassung	ix
Nomenclature	xv
1. Introduction	1
1.1. Motivation	1
1.2. State of the Art	3
1.2.1. Separation standards	3
1.2.2. Wake vortex alleviation	4
1.2.3. Wake vortex prediction	5
1.2.4. Wake vortex detection	7
1.2.5. Encounter severity analysis	8
1.2.6. Ensemble forecasts	9
1.3. Context and Goals of this Thesis	10
1.4. Overview	11
2. Wake Vortex Principles	13
2.1. Helmholtz Theorems	13
2.2. Vortex System of a Wing	13
2.3. Wake Vortex Properties	15
2.3.1. Basic quantities	15
2.3.2. Non-dimensional quantities	16
2.4. Idealized Two-Dimensional Vortices	17
2.4.1. Decaying potential vortex	17
2.4.2. Lamb-Oseen model	18
2.4.3. Rosenhead-Burnham-Hallock vortex model	18
2.5. Wake Vortex Evolution Stages	18
2.6. Wake Vortex Formation	20
2.6.1. Vortex roll-up	20
2.6.2. Vortex merging	20
2.7. Wake Vortex Descent	21
2.7.1. Out-of-ground effect	21
2.7.2. In- and near-ground effect	22
2.8. Wake Vortex Decay	23
2.8.1. Out-of-ground effect	23
2.8.2. In- and near-ground effect	24

3. Field Measurement Data	27
3.1. Measurement Equipment	27
3.1.1. LiDAR	27
3.1.2. Sonic Detection and Ranging (SoDAR), Radio Acoustic Sounding (RASS) and Ultrasonic Anemometer (USA) System	28
3.2. Campaigns	28
3.2.1. WakeFRA	28
3.2.2. WakeMUC	29
3.2.3. WakeOP	30
3.2.4. MEM95	30
3.2.5. DFW97	30
3.2.6. DEN03	30
3.3. Distribution of Aircraft and Meteorological Parameters	31
4. Wake Vortex Prediction	33
4.1. Skill Metrics	33
4.2. The Fast-Time Models	33
4.2.1. Common model features	33
4.2.2. D2P	34
4.2.3. APA	37
4.2.4. TDP	39
4.2.5. Necessary adaptations to the model output	39
4.3. Model Evaluation	40
4.3.1. Initial conditions	40
4.3.2. Sensitivity analysis	41
4.3.3. Performance	47
5. Multi-Model Ensemble Methods	51
5.1. Principles	51
5.1.1. Conditions for the MME success	53
5.1.2. Qualification of the available wake vortex models	53
5.1.3. Spread-error correlation	55
5.2. Best Member Selection	56
5.2.1. WakeMUC, WakeFRA, WakeOP	56
5.2.2. MEM95, DFW97	58
5.2.3. Reasons for absence of expert models	58
5.3. Direct Ensemble Average (DEA)	60
5.4. Reliability Ensemble Averaging (REA)	60
5.4.1. Natural variability	61
5.4.2. Weighting of R_B and R_D	65
5.4.3. Potential	65
5.5. Bayesian Model Averaging (BMA)	66
5.5.1. Maximum-likelihood estimation of ensemble parameters	67
5.5.2. Distribution formulation	70
5.5.3. Temporal development of model uncertainty	71

5.5.4. Potential	73
5.6. Monte Carlo Simulation (MCS)	76
5.7. Lagged Average Forecasting (LAF)	76
6. Results	77
6.1. Evaluation of the LiDAR Wake Vortex Measurements	77
6.1.1. Impact of initial height on descent and decay	77
6.1.2. Impact of wind speed and direction on descent and decay	79
6.1.3. Comparison of the WakeMUC measurements with D2P	81
6.2. Assessment of the Ensemble Performance	85
6.2.1. Deterministic Performance	86
6.2.2. Probabilistic Performance	95
6.2.3. Spread-error correlation	100
7. Discussion	105
7.1. LiDAR Measurement Evaluation	105
7.2. Multi-Model Ensemble Wake Vortex Prediction	105
7.2.1. Deterministic benefit of the ensemble approach	105
7.2.2. Probabilistic benefit of the ensemble approach	107
7.2.3. Skill distribution	115
7.2.4. Effort and benefit	116
7.2.5. Concept of operations	116
8. Conclusion and Outlook	123
A. Appendix	127
A.1. Model Correlation	127
A.2. Model Performance	128
A.3. LiDAR Measurement Evaluation	130
A.4. Rare But Dangerous	132
A.5. Extraordinary Measured Landings	135
A.6. Combined Confidence Areas for Vortex Position	140
Bibliography	143
Acknowledgments	155

Nomenclature

A	event A
A_{D2P}	tuning constant in D2P
A_O	area of vortex oval
b	vortex spacing, m
B_i	bias
B	event B
C	turbulence proportionality constant
C_1	constant in APA model to describe vortex decay
C_{umin}	minimum speed coefficient
dA	surface segment
ds	curve segment
D_i	absolute distance to ensemble mean
\tilde{f}	ensemble forecast
f_+	upper ensemble uncertainty limit
f_-	lower ensemble uncertainty limit
f_i	forecast of i th model
F_L	lifting force, N
g	gravitational acceleration, m/s ²
$g_i(y f_i)$	probability density function of forecast
I	Impulse
K	parameter in APA model to describe vortex spacing
l_0	integral length scale of the turbulence, m
L	likelihood function
m	weighting factor for $R_{B,i}$
M_i	statistical model
M_c	parameter in APA model to describe vortex decay
n	weighting factor for $R_{D,i}$
N	Brunt-Väisälä frequency, 1/s
nv	natural variability
$p(y)$	probability of forecast y
p_{obs}	observed coverage
p_{nom}	nominal forecast probability
$P(A)$	probability that A occurs
$P(A B)$	probability that A occurs, given B
$P(B)$	probability that B occurs
q	turbulence velocity, m/s
r	radius, m

r_c	vortex core radius, m
r_l	lower radius
r_u	upper radius
R	mean radius
$R_{B,i}$	reliability factor (considering model bias)
$R_{D,i}$	reliability factor (considering model convergence)
R_i	total reliability factor
$s_{i,p}$	relative skill measure
s_i	relative skill measure, averaged over all parameters
s_l	load factor
t	vortex age, s
t_g	time at which vortices reach minimum height, s
T_c	time at which a catastrophic demise event occurs, s
T_1	age of the vortices at $t = 0$ s in D2P, s
T_2	time at which rapid decay starts in D2P, s
u	headwind, m/s
u_{ac}	aircraft airspeed, m/s
ud_{DES}	descent speed increment, m/s
u_{stall}	stall speed, m/s
v	crosswind, m/s
V	flow velocity, m/s
V_∞	undisturbed flow velocity
V_θ	tangential velocity, m/s
w	vortex descent speed, m/s
w_i	weight
y	lateral position of vortex core, m
y^T	training data
z	vertical position of vortex core, m
$\hat{z}_{i,t}^{(j)}$	unobserved quantity

Greek letters

α	constant in APA model to describe vortex spacing
β	vortex core diffusion parameter
Γ	circulation, m ² /s
$d\Gamma_D$	decay contribution from turbulence diffusion in TDP, m ² /s
$d\Gamma_L$	decay contribution from linking instability in TDP, m ² /s
$d\Gamma_S$	decay contribution from stratification in TDP, m ² /s
$\tilde{\delta}_f$	ensemble forecast uncertainty
ϵ	eddy dissipation rate, m ² /s ³
θ	potential temperature, K
μ	distribution mean
$\tilde{\mu}$	distribution mean in preconditioning step
ν	kinematic viscosity, m ² /s
ν_1	effective viscosity, m ² /s
ν_2	effective viscosity in rapid decay phase, m ² /s

ρ	fluid density, kg/m ³
$\tilde{\rho}$	collective reliability
σ	standard deviation
$\tilde{\sigma}$	standard deviation in preconditioning step
σ_{obs}	variability observed in measurements
σ_{err}	LiDAR measurement error
σ_{nv}	natural variability of vortices
ω	vorticity, 1/s

Subscript

e	ensemble approach number
i	model number
I	number of models
n	number of observations
o	observation
p	parameter
t^*	vortex-age dependent treatment
0	initial value

Superscript

(j)	iteration number
*	normalized quantity
'	fluctuation velocity

Abbreviations

APA	AVOSS Prediction Algorithm
AVOSS	Aircraft Vortex Spacing System
AWIATOR	Aircraft Wing with Advanced Technology Operation
BMA	Bayesian Model Averaging
CAT	Clear Air Turbulence
DEA	Direct Ensemble Average
DFS	Deutsche Flugsicherung
DLR	Deutsches Zentrum für Luft- und Raumfahrt
D2P	Deterministic Two-Phase wake vortex model
EDR	Eddy Dissipation Rate
EM	Expectation-Maximization
FAA	Federal Aviation Administration
HH	<i>heavy</i> leader, <i>heavy</i> follower
HM	<i>heavy</i> leader, <i>medium</i> follower
ICE	Initial Condition Ensemble
IFR	Instrument Flight Rules
IGE	In-Ground-Effect
LES	Large Eddy Simulation
LiDAR	Light Detection and Ranging
LOS	Line Of Sight
MCS	Monte Carlo Simulation

MLW	Maximum Landing Weight
MME	Multi-Model Ensemble
NASA	National Aeronautics and Space Administration
NATS	National Air Traffic Services
NGE	Near-Ground-Effect
OEW	Operational Empty Weight
OGE	Out-of-Ground-Effect
PDF	Probability Density Function
PPE	Perturbed Physics Ensemble
prt	port
P2P	Probabilistic Two-Phase wake vortex model
RANS	Reynolds-Averaged-Navier-Stokes
RASS	Radio Acoustic Sounding System
RCR	Roll Control Ratio
REA	Reliability Ensemble Average
RMSE	Root Mean Square Error
RS	Radar Separation
SHAPE	Simplified Hazard Area Prediction
SoDAR	Sonic Detection And Ranging
stb	starboard
TASS	Terminal Area Simulation System
TBO	Trajectory Based Operations
TBS	Time Based Separation
TDP	TASS Driven Algorithms for Wake Prediction
TKE	Turbulent Kinetic Energy
USA	Ultrasonic Anemometer
UTC	Coordinated Universal Time
VFR	Visual Flight Rules
WEAA	Wake Encounter Avoidance and Advisory
WSVBS	Wirbelschleppen Vorhersage- und Beobachtungssystem (Wake Vortex Prediction and Monitoring System)
WVAS	Wake Vortex Advisory System

1. Introduction

1.1 Motivation

Steadily increasing air traffic is a challenge for air traffic controllers, airports, air carriers, and pilots. Safe operations may be threatened by wake vortices [1] which can be encountered en-route, during departures and arrivals, and particularly in ground proximity, limiting the capacity of runways and posing a potential risk to aircraft. Generated at the wings of aircraft as a response to lift wake turbulence may induce a hazardous rolling momentum or vertical force to any follower. Dangerous situations can occur when wake vortices hover in the glide path corridor, drift to parallel runways and live longer than expected. Especially at low altitudes during final approach, where the pilots do not have much time to take countermeasures to a vortex-induced rolling momentum, it is crucial to ensure that the vortices of the preceding aircraft have been transported out of the glide path or have decayed to an appropriate strength. Both decay and transport are affected by atmospheric parameters, such as wind conditions, thermal stratification, shear layers, and turbulence. As global air traffic is estimated to double in the next 15 years [2] with regional growth rates over 4% (see Table 1.1) the risk to encounter wake vortices both in terminal area but also in-flight [3] is increased.

As regulatory wake vortex separation distances are one of the bottlenecks of airport capacity, the air-traffic organizations decided to recategorize the separation standards. So far four wake categories, depending on the mass of an aircraft, have been distinguished in the ICAO standards. In the RECAT-EU project it is now split into six clusters, depending on both mass and wing span. Increasing the number of categories allows to decrease the aircraft separations while maintaining the existing safety standards. A similar recategorization has already been implemented by the FAA in the USA at the airports of Memphis, Louisville, Cincinnati and Atlanta with further airports to follow [4]. This progression emphasizes the need for more precise tools to adjust aircraft separation tactically in order to avoid airport congestion.

Large-Eddy Simulations (LES) [5–8], wind tunnel [9, 10] and water-towing laboratory experiments [11, 12] may contribute to a better understanding of the underlying physics and destruction mechanisms of wake vortices. Nevertheless, to increase the safety at airports and in-flight as well as to enable a tactical adaption of the separations, fast-time forecasts are required.

Various fast-time wake vortex models have been developed in the past to achieve a tactical capacity and safety gain [13–20]. These are simplified models, which are often trained with measurement data or LES output in order to calibrate the model forecasts.

Depending on the ambient conditions, the forecast quality of the wake vortex models varies due to different underlying physics and due to the diverse data that the models have been trained with. The evaluation of wake vortex forecasts, based on observations by LiDAR (Light Detection And Ranging), reveals significant uncertainty in the output of deterministic fore-

casts. Firstly, the stochastic and intermittent behavior of the atmospheric turbulence renders the required input parameters uncertain. Secondly, the turbulent vortex decay and interaction with the ground, during which three-dimensional structures can be generated that cannot be resolved by the models, gives rise to the so-called natural variability of wake vortex behavior. To cover these uncertainties, deterministic forecasts are not sufficient. For this purpose probabilistic predictions, providing uncertainty envelopes for various uncertainty levels, are required in practice.

Several studies, many of them conducted in the field of meteorology, demonstrated that in such situations the combination of various independent models may be beneficial to increase the prediction skill. This approach is known as Multi-Model Ensemble (MME). It is not only advantageous for the deterministic forecast skill, but it also enables the estimation of probabilistic envelopes for the forecasted parameters. The rationale of the MME is to incorporate the model uncertainty in the projections, based on the spread of prognoses between the models.

For this reason the models D2P [15, 16, 18], APA 3.2, APA 3.4 [14], and TDP 2.1 [17, 21] have been provided in the context of a NASA-DLR cooperation to investigate the capabilities of a MME to improve the deterministic wake vortex forecast and provide robust probabilistic projections. However, the temporal forecast horizon of weather forecasts, where the MME is common, is on the order of days, while it is on the order of a few minutes for wake vortex forecasts. The success of applying this approach to wake vortex projections for this reason is not self-evident. Therefore, the ensemble skill is to be compared to the performance of its individual members.

To ensure that the results are statistically significant, a sufficiently large amount of wake vortex field measurements is required. For this purpose the WakeMUC campaign [22] is evaluated and used for the development of the MME. Additionally, data from field experiments at Frankfurt (WakeFRA) [23], special airport Oberpfaffenhofen (WakeOP) [22], Dallas (DFW97) and Memphis (MEM95) [24] airport are employed. Eventually, it shall be assessed whether the MME approach bears the chance of covering rare but dangerous wake vortex cases under the assumption that the ensemble members are experts on different fields.

Table 1.1.: Estimated average annual growth of air traffic (number of direct flights) between the years 2012 and 2035 [25].

region	annual growth [%]
North-Atlantic	1.0
Mid-Atlantic	1.9
South-Atlantic	2.1
Europe	1.3
North-Africa	3.8
Southern Africa	3.4
Middle-East	3.6
Asia/Pacific	4.1

1.2 State of the Art

1.2.1 Separation standards

Today aircraft separation regulations follow the ICAO rules which are based on the weight ratio of the leading and following aircraft. Three weight classes exist: *Heavy*, *Medium* and *Light*. Additionally, special separation must be followed for the Airbus A380. The corresponding separation standards in nautical miles are listed in Table 1.2. If no wake vortex separation is required, radar separation of either 3 NM, or 2.5 NM is applied [1]. The reduced radar separation may only be applied if a radar with appropriate range resolution is available, if the runway occupancy time does not exceed 50 s, if braking action is reported as good, if weather and runway conditions are appropriate and if the approach speeds are monitored by the controller [26].

Table 1.2.: ICAO aircraft separation in nautical miles (NM) during approach and departure [27]. RS denotes the radar separation which is either 3 NM or 2.5 NM. The weights correspond to the maximum take-off masses.

Following aircraft	Leading aircraft			
	A380-800	Heavy	Medium	Light
A380-800 (ca. 560.000 kg)	RS	RS	RS	RS
Heavy (≥ 136.000 kg)	6	4	RS	RS
Medium (≥ 7.000 kg, < 136.000 kg)	7	5	RS	RS
Light (< 7.000 kg)	8	6	5	RS

As these standards are empirical and believed to be over-conservative [1] a recategorization has been elaborated to increase the runway throughput by more precise separations while the safety levels are maintained [27]. Therefore, the *Heavy* and *Medium* categories are split into *Upper* and *Lower* within the context of RECAT-EU. Additionally, the category *Super Heavy* is introduced, which contains the A380 and the AN-124 aircraft. Furthermore, the weight classes were revised, but in contrast to the ICAO standards the categorization of single aircraft types does not solely depend on the maximum take-off weight but also on the wingspan (Table 1.3). Table 1.4 shows the consequential new separation distances. It is estimated that these new standards may achieve throughput benefits of up to 5% during traffic peaks [27]. The first airports at which these six new categories have been applied (since 22nd of March 2016) were LFPG/Paris Charles de Gaulle and LFPB/Paris Le Bourget [28].

Besides, NATS implemented Time-Based-Separation (TBS) at the London Heathrow airport [30]. Strong headwinds lead to decreased speed over ground and increased temporal separations, such that airport capacity is reduced. TBS ensures that the time between the landings, rather than the distance, stays constant to maintain the capacity. NATS found that headwind accelerates vortex dissipation so that the safety levels can be maintained likewise [31].

Table 1.3.: RECAT-EU subcategorization [27].

category	MTOW	subcategory	wingspan
Heavy	≥ 100.000 kg	Super	> 72 m
		Upper	> 60 m, < 72 m
		Lower/Upper	> 52 m, < 60 m
Medium	> 15.000 kg, < 100.000 kg	Lower	< 52 m
		Upper	> 32 m
		Lower	< 32 m
Light	≤ 15.000 kg	–	–

Table 1.4.: RECAT-EU aircraft separation scheme for approach and departure with distances in nautical miles (NM) [27, 29]. RS denotes the radar separation which is either 3 NM or 2.5 NM.

Following aircraft	Leading aircraft					
	A	B	C	D	E	F
	Super Heavy	Upper Heavy	Lower Heavy	Upper Medium	Lower Medium	Light
A Super Heavy	3	RS	RS	RS	RS	RS
B Upper Heavy	4	3	RS	RS	RS	RS
C Lower Heavy	5	4	3	RS	RS	RS
D Upper Medium	5	4	3	RS	RS	RS
E Lower Medium	6	5	4	RS	RS	RS
F Light	8	7	6	5	4	3

1.2.2 Wake vortex alleviation

Independent of the current separations standards, various measures have been assessed to further reduce the likelihood of wake vortex encounters. One of them is the reduction of the vortex strength. This is especially interesting for aircraft manufacturers who want to design heavy aircraft without increasing the separations required at airports [1].

Gerz et al. [1] distinguish two basic strategies. The first aims at producing less harmful vortices. Therefore, Stumpf [6] accomplished LES simulations to investigate a vortex-topology that results in reduced rolling moments. He employs a sequence of numerical simulations that includes the modification of the aircraft geometry such that the vortex-induced rolling momentum can be reduced down to roughly 50% within a distance of 2.89 NM for an Airbus A340.

Another strategy rests upon the acceleration of vortex decay by introducing disturbances. Constructive measures serving those purposes could be active control surfaces on the wing or passive wake generators. A ground based strategy of wake vortex mitigation has been investigated by Holzäpfel et al. [32]. They performed LES simulations [7] and accomplished field measurements [32] that confirmed that so-called plate lines may reduce the longevity of wake vortices. The plates, being installed at the beginning of the runway, trigger disturbances that travel along the vortex and accelerate its decay. Nevertheless, the alleviation of wake

vortices is far from being resolved despite extensive research [1], which justifies the existence of fast-time wake vortex forecasts.

1.2.3 Wake vortex prediction

The DLR fast-time wake vortex model D2P and the models operated by NASA, namely APA 3.2, APA 3.4 and TDP 2.1, provide predictions of vortex position and strength. They are introduced in more detail in the course of this work as they form the basis of the ensemble. With exception of D2P all of these models offer deterministic forecasts exclusively. By contrast, D2P is the deterministic version of P2P, which generates probabilistic envelopes from a combination of analytical considerations and fits to observations. Other approaches seek to utilize the model sensitivity by performing Monte Carlo Simulations (MCS), in which either the initial conditions or physical model parameters are varied in order to model the forecast uncertainty. The probabilistic wake vortex model (PVM) [33], developed by the Université catholique de Louvain (UCL), combines both. By employing boot-strap methods DeVisscher et al. claim to achieve short simulation times that are even applicable to real-time operational systems. Its deterministic core is formed by the Deterministic Vortex Model (DVM) [19]. It models the vortex-ground interaction very detailed by regarding the vortex induced vorticity at relatively small scales.

Another wake vortex model, currently under development at North West Research Associates (NWRA), is the Vortex algorithm Including Parametrized Entrainment Results (VIPER) and includes entrainment and detrainment physics [34]. In this manner they attempt to treat the conservation laws for mass, momentum and angular momentum properly. Further models, such as the STL (based on the APA model) and the Linear Circulation Decay Model (LIN) exist [34].

The applicability of wake vortex forecasts has been assessed at Dallas Ft. Worth International Airport during July 2000 where NASA applied its Aircraft Vortex Spacing System (AVOSS) [14, 35, 36] to a single runway. The AVOSS generates wake vortex predictions on the basis of wind, turbulence, temperature profiles and glide path adherence statistics in order to calculate dynamic aircraft separations. For verification, wake vortex sensors monitored whether the vortices behaved as predicted [36]. The increase in mean runway capacity was estimated to be 6%, corresponding to a capacity gain of 2 aircraft per hour at a potential arrival rate of 31 aircraft per hour for a single runway with default spacing [35, 36].

A similar approach, the wake vortex prediction and monitoring system WSVBS (see Figure 1.1) by DLR, has been validated with two parallel runways at Frankfurt airport [37]. It employs predictions of the probabilistic two phase wake vortex model (P2P) by DLR for multiple gates along the glide path. The meteorological conditions are captured with a SoDAR/RASS system and an ultra sonic anemometer (USA) in the vicinity of the threshold. The rest of the glide slope is covered by numerical weather predictions. In a first step P2P predicts vortex position and strength based on the meteorological conditions as well as on the glide path adherence statistics (FLIP) [37] in each gate. In addition, a safety area of elliptical shape is generated by expanding the upper and lower bounds of the position and strength of the vortices predicted by P2P to ensure undisturbed flight according to the Simplified Hazard Area Prediction (SHAPE) method. SHAPE predicts the safety critical areas, based on the required roll control power

of the follower aircraft to counteract the wake vortices [38]. In a next step safe temporal aircraft separations are determined, which are subsequently translated into dynamic separations with support of the DLR arrival manager (AMAN) [37]. To guarantee safe operations a LiDAR system, that is able to capture the position and strenghts of the vortices, monitors the correctness of the wake vortex predictions at the lower altitudes of the glide path [37].

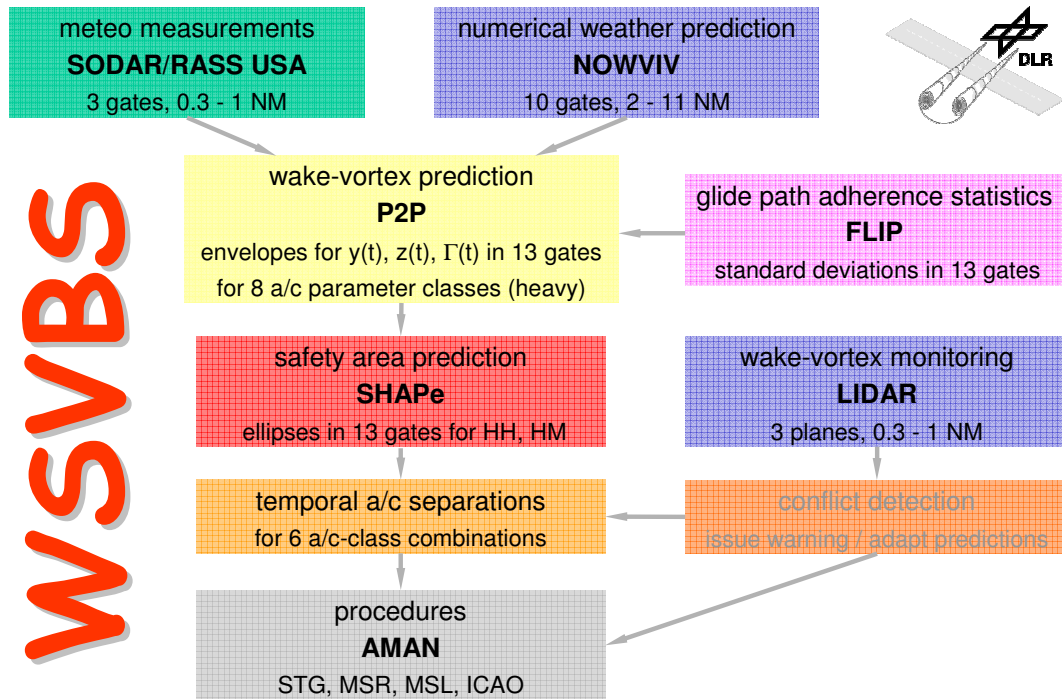


Figure 1.1.: Flowchart for WSVBS [37] (reprinted with permission of Frank Holzäpfel).

The assessment of the WSVBS during a period of 66 days between 18/12/06 and 28/02/07 at Frankfurt airport demonstrated that aircraft separations could have been reduced in 75% of the landings compared to the ICAO standards [39]. Gerz et al. [39] showed that dynamical separations may reduce the delays while the ICAO standards accumulate delayed flights that can only be resolved in the late evening hours [39]. For this reason they consider the improved capacity as tactical in the sense that punctuality can be increased under favorable weather conditions.

In this context the Wake Vortex Scenarios Simulation (WakeScene) [40] package is to be mentioned. It offers a freely selectable perspective view of trajectories of vortex generating aircraft together with the positions of the port and starboard vortex in vertical gates along the flight track [40]. The wake vortex behavior is predicted by D2P in terms of circulation, position, core radius and attitude of the vortex axis [40] in dependency of the aircraft and meteorological parameters. To define a hazard area two options exist. Either a circle of 50 m radius around the vortex position is generated, or the simplified area hazard prediction (SHAPe) is applied [40]. The package is capable of performing Monte Carlo simulations as well as dedicated parameter variations [40]. Thereby, WakeScene allows to assess the probability of wake encounters for different air traffic scenarios and can be applied for sensitivity analysis, optimization of air

traffic control procedures, assessment of new wake vortex advisory systems or to elaborate a novel aircraft separation matrix [40]. A similar package, the Wake4D platform [33], has been developed by the Université catholique de Louvain (UCL). It is based on the earlier mentioned DVM and PVM and allows to analyze approach and departure patterns in a similar manner.

Although the pilots certainly have enough time to regain control of the aircraft at high altitudes in-flight, structural damage or personal injuries cannot be ruled out. En-route, aircraft are restricted to certain flight altitudes, called Flight Levels (FL) that are given in hft (hundred feet). Under Instrument Flight Rules (IFR) the vertical spacing between these levels amounts to 1000 ft, starting with FL 40 [41]. According to the semi-circular flight rule, odd thousands are reserved for aircraft with headings from 0° to 179° while even thousands are reserved for aircraft with headings from 180° to 359° [41].

Although wake vortices generated en-route naturally descend towards lower flight levels, the risk of encountering wake vortices appears low at first glance, given the vertical spacing of the flight levels. Wake vortices usually descend by at most 900 ft [42] which is due to their limited lifetime as a consequence of natural decay, caused by viscosity and decay enhancing effects such as turbulence and baroclinic vorticity. Nevertheless, 73 wake-turbulence incidents have been reported at upper levels (defined as flight levels above 270 hft in Europe) between July 2009 and July 2012 [43]. Schumann and Sharman [43] found based on simulations that most encounters occur on nearly parallel flight routes during descent (46%). Only 0.4% of the encounters are observed when both aircraft are flying at constant altitudes [43]. However, future plans such as the Trajectory Based Operations (TBO) incorporate the compression of the air space [44], rendering encounters more likely.

To reduce the risk of wake vortex encounters in-flight, the Wake Encounter Avoidance and Advisory system (WEAA) [45] aims at visualizing the predicted wake vortex position of surrounding aircraft on the navigation display in order to warn the pilots in case of conflict with the planned flight track. The predictions are performed by an airborne version of P2P which employs offline Monte Carlo simulations to accomplish the probabilistic envelopes from which a volume is derived where the vortices are likely to be encountered. With the meteorological data being crucial to wake vortex forecasts substantial effort is undertaken to derive wind speed, temperature and turbulence from various sources and interpolate them to the flight track of the vortex generating aircraft [45]. During test flights the DLR research aircraft Falcon served as wake-vortex generator, while the DLR A320 ATRA intentionally encountered the vortices at the positions predicted by the WEAA system [45]. Depending on the source of the meteorological input, hit rates between 65% and 90% were achieved.

1.2.4 Wake vortex detection

A future possibility to avoid wake vortex encounters might be to scan the intended flight track in front of the aircraft by means of an airborne LiDAR [46]. Originally intended to enable the avoidance of Clear Air Turbulence (CAT), the system developed in the DELICAT project measures density fluctuations with a UV LiDAR to quantify vertical velocities. Mounted to a Cessna Citation from the Dutch Aerospace Lab NLR the system had its maiden flight in 2013 during a flight campaign to measure natural CAT events. Although the test of the instrumentation was successful with all systems performing well, it turned out to be difficult

to localize major CAT events despite available forecasts [47]. For this reason further validation flights proved to be necessary before it can be applied to detect wake vortices in the future.

1.2.5 Encounter severity analysis

The severity of a wake vortex encounter and the aircraft reaction depend on the vortex strength, on the size of the following aircraft as well as on the distance to the vortex core and the encounter angle. Luckner et al. [48] distinguish between the following aircraft reactions that may be superimposed:

- Roll motions (A)
- Vertical motions (up (B) and down (C))
- Yawing motions (D).

Figure 1.2 illustrates the regions where these aircraft reactions occur, given parallel flight routes of the leading and following aircraft.

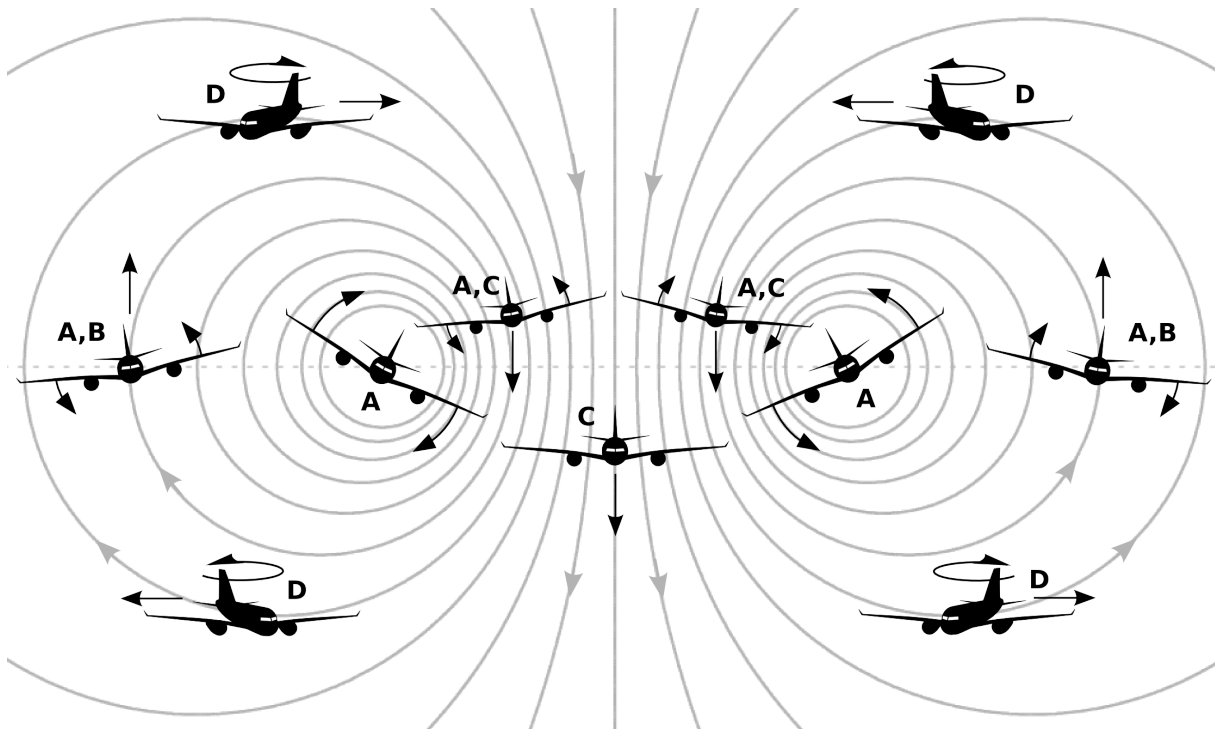


Figure 1.2.: Dominant wake encounter reactions (black). In the background the stream lines of the vortex pair seen from a static viewpoint are illustrated (see Lamb [49]). For a quantitative analysis see Luckner et al. [48].

The larger the encounter angle Ψ between aircraft trajectory and vortex (see Figure 1.3) the less rolling momentum is experienced by the aircraft. Instead more and more vertical forces act upon the aircraft [48]. To further assess the aircraft reaction Bieniek and Luckner

[50] performed flight simulations with different encounter angles for encounters with straight and wavy vortices as well as with vortex rings, based on analytical wake vortex models. In a simulator study airline pilots flew manual landing approaches during which straight and wavy vortex encounters were simulated [50]. Due to the varying effective encounter angles (which depends on where the aircraft penetrates the deformed vortex) the pilot hazard ratings exhibited a large spread [50]. Vortex rings, generating lower bank angles, were rated less hazardous than both wavy and straight vortices, while wavy vortices generated the largest bank angles [50].

As no final conclusions could be drawn from this study, further effort has been undertaken to understand whether deformed or straight vortices are more hazardous. Therefore Vechtel [51] performed encounter simulations of the DLR ATRA aircraft on the basis of an idealized analytical model for straight and LES fields for deformed vortices. His simulations revealed that the encounters with deformed vortices are weaker and that the maximum aircraft response is smaller compared to encounters with straight vortices on average. Furthermore, it was found that wavy vortices lead to fewer strong encounters than straight vortices.

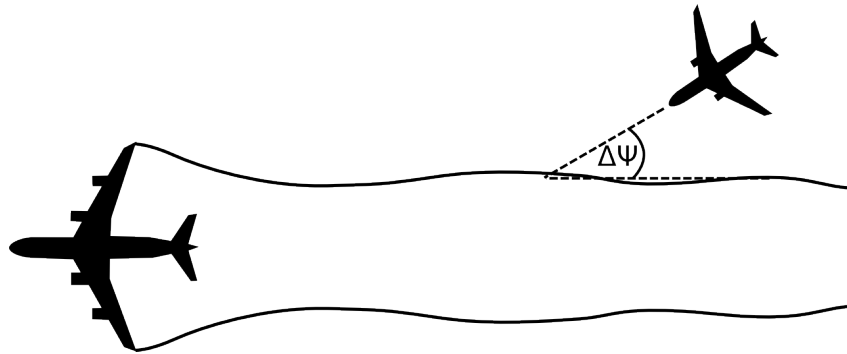


Figure 1.3.: Wake encounter angle Ψ . With increasing Ψ weaker rolling momentum and stronger vertical forces are experienced by the following aircraft.

1.2.6 Ensemble forecasts

Already back in the times when weather predictions were less machine-aided than today, even skilled forecasters reached different results. Today a variety of models exist which are based on different physical aspects and deliver a variety of forecasts. The information on which model to trust can be based on validation by means of measurements. However, in an early study Sanders [52] showed that the mean of multiple forecasters, which he called consensus, is more skillful than that of the most skillful individual. This was confirmed in later studies [53–55] that evaluated the same method for an enlarged set of climatological measurements and for further locations. Eventually, studies were conducted that attempted to quantify the uncertainty of climate change forecasts by connecting the model agreement (or ensemble spread) with prediction reliability [56–59].

Having been most common in climate modeling [57] and weather forecasting, the ensemble approach was soon transferred to further fields of application, such as medicine [60], biology [61] or agriculture [62]. Ever since the ensemble approach became popular it has been argued

where its success originates from [63] and whether it can indeed outperform single model forecasts [64]. The conclusions obtained from these discussions suggest that the ensemble success is conditional on the coverage of the solution space by the ensemble members.

Various ensemble types have been developed, each covering different uncertainty types. Initial Condition Ensembles (ICE) [65–67] assume that the initial conditions are imprecise. For this reason they average forecasts obtained from perturbed starting values. Likewise Perturbed Physics Ensembles (PPE) [68] accomplish various simulations based on different model parameters under the conjecture that they are uncertain. Finally, the Multi-Model Ensemble (MME) [64, 68] involves various independent models to allow for the model uncertainty. Krishnamurti et al. [69] demonstrated that appropriate prior weighting of the models may reach higher skill than forecasts based on simple averaging. Evaluating the ensemble performance in a probabilistic manner, Palmer et al. [70] point out that the MME achieves significantly higher skill and higher reliability than a single model ensemble that is achieved by perturbed initial conditions.

Different methods to generate a MME, such as the Direct Ensemble Average (DEA), the Reliability Ensemble Averaging (REA) [58] and the Bayesian Model Averaging (BMA) [71] demonstrated added-value for weather and climate forecasts. Further approaches incorporate neuronal networks [72] or Bayesian hierarchical models [73]. Today, met offices around the world [74–76] routinely provide ensemble forecasts that are based on Monte Carlo Simulations, Multi-Model Ensembles or a combination of both.

1.3 Context and Goals of this Thesis

At present, no multi-model ensemble capability exists for wake vortex predictions. The forecast uncertainties are derived from static and dynamic uncertainty allowances and training with measurement data [15, 16, 18], or estimated from perturbed initial conditions [33, 77] and from perturbed physics [33]. However, they do not incorporate model uncertainty yet. In this thesis the latter is addressed by combining the forecasts of structurally independent wake vortex models in order to increase the forecast skill and reliability. Achieving consistently credible and robust probabilistic envelopes that incorporate the uncertainty of the NASA and DLR models offers further prospect to increase the acceptance of wake vortex forecasts. Therefore, various MME approaches are investigated, adapted and eventually applied to the available wake vortex models. Furthermore, the development of model uncertainty is assessed and utilized to accurately compute probabilistic envelopes that are reliable independently of the vortex age. In this manner the probabilistic coverage of rare but dangerous cases, that occur when vortices remain in the glide path, is believed to be enhanced.

Besides, the latest LiDAR wake vortex measurements enable an increased understanding of vortex physics and offer the opportunity to further develop D2P. In particular the effects of cross- and headwind as well as the impact of the initial altitude on vortex transport and decay are investigated. The goals can be summarized as follows:

- Deeper understanding of wake vortex physics based on the latest measurement campaigns.
- Increased deterministic wake vortex forecast skill.

- High reliability and accuracy of the probabilistic forecasts.
- Enhanced coverage of rare but dangerous cases.

1.4 Overview

In Chapter 2 the fundamental theory behind the generation and evolution of wake vortices is discussed. This includes formulations already found more than one century ago, simplified vortex models, and results from the latest Large-Eddy Simulations (LES).

In Chapter 3 the measurement equipment and the captured wake vortex data, used for development and evaluation of the MME, are introduced. In addition an evaluation of the field measurement campaigns is presented to further elucidate the influence of atmospheric parameters on wake vortex behavior based on a just recently made available sufficiently large dataset.

In a next step the models that have been exchanged within the context of a cooperation between NASA and DLR are described in Chapter 4. To emphasize their differences a sensitivity analysis, premised on various meteorological conditions as input, is presented.

The basic concept of the Multi-Model Ensemble technique and its success factors are introduced in Chapter 5. Furthermore, different potentially promising methods to generate an ensemble are described.

In Chapter 6 at first the evaluation of the LiDAR measurements is illustrated. In the second part, the results achieved by the introduced ensemble approaches are presented. Their performance is evaluated in a deterministic and probabilistic manner and compared to the respectively best performing member for both the DLR and NASA measurements.

Finally, these results are discussed in Chapter 7, followed by the conclusions drawn from the findings and the outlook in Chapter 8.

2. Wake Vortex Principles

2.1 Helmholtz Theorems

A vortex is defined as a concentric distribution of vorticity [78], with the latter being formulated as

$$\omega = \nabla \times u. \quad (2.1)$$

In the following a *vortex line* denotes the curve in a fluid that is tangent to the local vorticity vector [78]. Furthermore, a *vortex tube* describes a tubular surface that encloses various vortex lines in a flow with nontrivial vorticity [78]. More than one century ago Hermann von Helmholtz formulated theorems describing the nature of vortices in inviscid flow [78].

1. Vortex lines move with the fluid.
2. The strength of a vortex tube is constant along its length.
3. A vortex tube cannot end within the fluid. It must either end at a solid boundary or form a closed loop.
4. The strength of a vortex filament remains constant in time.

The first theorem can be easily observed in LiDAR measurements where vortices are transported with the velocity of the wind if they are out of reach of the ground. The third theorem can frequently be observed when wake vortices, visualized by the contrails of aircraft, form vortex rings which prevents them from having free ends [79]. In ground proximity they may link with the ground instead of forming closed rings [80]. The second and the fourth theorem, however, cannot be observed in measurements, as inviscid flow does only exist in theory. Viscous effects introduce temporal decay due to friction, resulting in circulation gradients along the vortex line.

2.2 Vortex System of a Wing

The generation of lift can be explained by several complementary theories. A very common one is the lifting line theory, introduced by Prandtl [81] and Lanchester [82] at the beginning of the 20th century, when wake vortices did not cause capacity issues yet. They associated the lift along a wing of finite span with the circulation around the airfoil which is also termed bound vortex (Figure 2.1). As a consequence of its decreasing strength towards the wing tips vorticity is shed off and finally merges into the tip vortices. According to the third Helmholtz

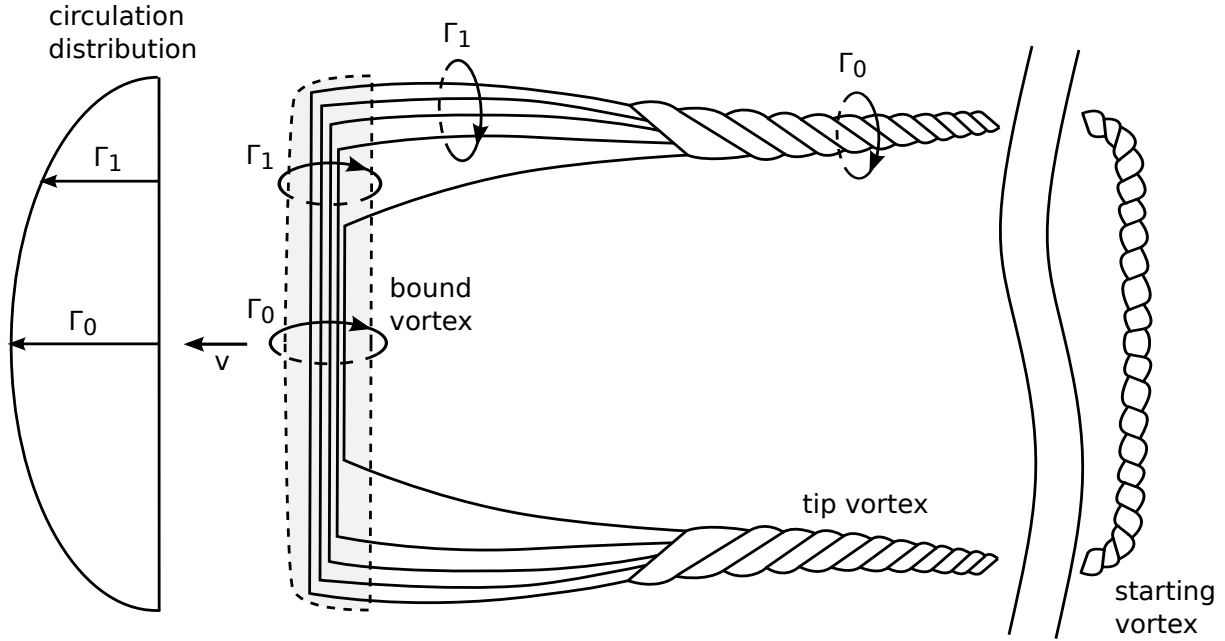


Figure 2.1.: Vortices generated by a moving wing with upward directed lift, viewed from above. cf. [78].

law a vortex cannot end within the fluid, which is why both vortices are connected by the starting vortex. Figure 2.2 shows that the flow only leaves the trailing edge smoothly in the presence of the bound vortex, such that vertical impulse and thus lift is generated.

The strength of the bound vortex is characterized by the circulation Γ that is defined by the integral of the tangential velocity v_θ along a closed curve s . According to the Stokes theorem it can be equally expressed by the integral of vorticity ω over the surface A [83].

$$\Gamma = \oint_C V_\theta ds = \int_A \omega dA \quad (2.2)$$

The resulting lifting force F_L can be directly connected with the strength Γ of the bound vortex according to the Kutta-Joukowski theorem for either one vortex filament or the whole wing by [78]

$$dF_L = \rho V_\infty \Gamma(y) dy \quad \text{and} \quad F_L = \int_{-B/2}^{B/2} \rho V_\infty \Gamma(y) dy, \quad (2.3)$$

where y denotes the span-wise direction, B the wingspan, ρ the fluid density and V_∞ the undisturbed flow velocity. The decreasing strength of the bound vortex towards the wing tips leads to vortex filaments detaching from the trailing edge and forming a vortex sheet [78]. The stronger the gradient of the bound vortex strength, the stronger the circulation of the trailing vortex filament that is shed off [78]. Eventually, further downstream the vortex sheet rolls up from each half of the wing to form two counter-rotating vortices [78].

The resulting flow resistance is also referred to as induced drag [78]. Prandtl [81] concluded that the induced drag of a wing of infinite span would account to zero. Furthermore, he found

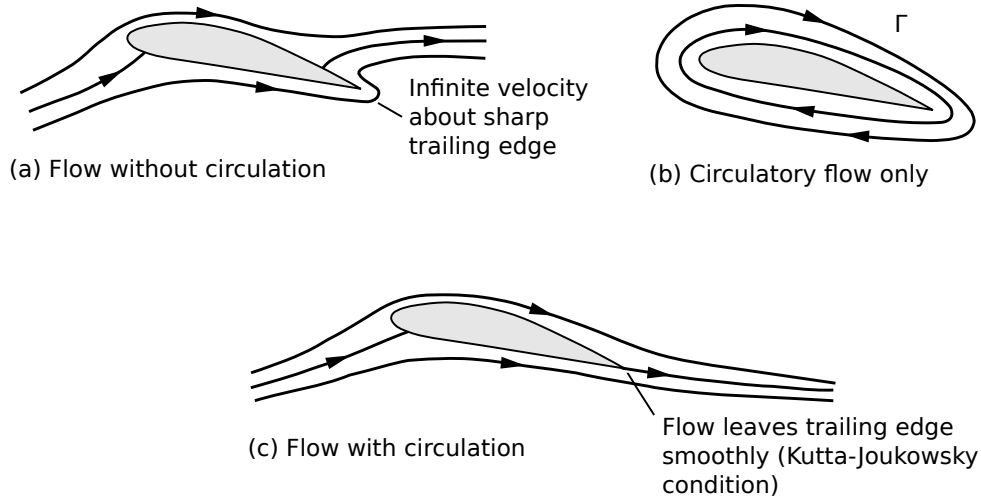


Figure 2.2.: Formation of the bound vortex around an airfoil cf. [85].

that finite wings with elliptical distributions of lift show the best ratio of lift and drag. Thus elliptical lift distributions are often the basis of theoretical considerations. In reality, the lift distribution of modern aircraft together with the generation of vorticity is more complex. The reasons for this are the fuselage and attached parts, such as engine nacelles or flaps, which generate further vortical structures [84], or structural considerations to increase lift close to the wing roots in order to optimize the wing's weight. After the vorticity sheet has rolled up, a pair of counter-rotating vortices is formed. The properties of the vortex pair as well as the processes of vortex evolution and decay are discussed in more detail in the following.

2.3 Wake Vortex Properties

2.3.1 Basic quantities

Under the assumption that all forces acting on the aircraft are in equilibrium ($L = Mg$), the initial circulation of the fully developed vortices is described by

$$\Gamma_0 = \frac{Mg}{\rho s_l B V}, \quad (2.4)$$

where M indicates aircraft mass, g gravitational acceleration, ρ air density, s_l the spanwise load factor, B wingspan, and V velocity [1]. The load factor s_l depends on the circulation distribution of the wing and can be computed by [1]

$$s_l = \frac{2}{B} \int_0^{B/2} \frac{\Gamma(y)}{\Gamma_0} dy. \quad (2.5)$$

For elliptically loaded wings this yields $s_l = \pi/4$, but it is often observed that also wings with other distributions than that have very similar load factors [1]. The initial separation b_0 turns out to equal $s_l B$ and is thus for elliptical circulation distributions $\pi/4 B$ [1].

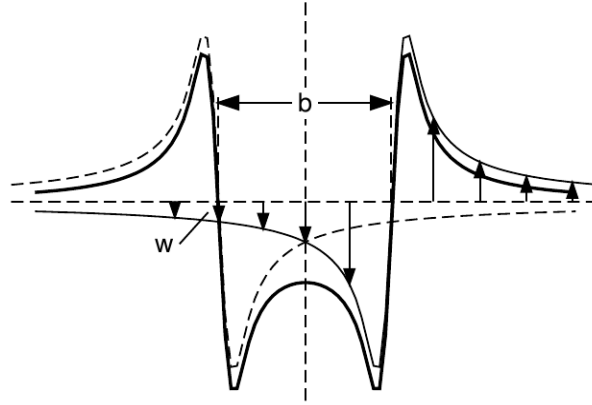


Figure 2.3.: Spanwise vertical velocity for a counter-rotating pair of Lamb-Oseen vortices separated by b . The bold line indicates the resulting vertical velocities of the descending vortex oval [86].

When two vortices exist their velocity fields overlap which results in a mutually induced downwards directed velocity (see Figure 2.3). Given that both vortices decay at the same rate and that the vortices can be approximated by a potential vortex at the radius $r = b$, the descend speed w of the vortex pair can be expressed by

$$w = \frac{\Gamma}{2\pi b}, \quad (2.6)$$

where b denotes the current vortex spacing [1]. To compare the behavior of vortices generated by aircraft of different sizes, the reference time scale t_0 is introduced [1]. It describes the time during which the vortex pair descends one initial vortex spacing b_0 .

$$t_0 = 2\pi \frac{b_0^2}{\Gamma_0} = \frac{b_0}{w_0} \quad (2.7)$$

2.3.2 Non-dimensional quantities

To compare the vortex behavior for aircraft of various sizes, dimensionless parameters are employed. It is found that vortex transport scales with the initial vortex spacing b_0 , while the vortex strength is normalized by the initial circulation Γ_0 . Time is normalized by t_0 and henceforth termed vortex age t^* .

$$z^* = \frac{z}{b_0}, \quad y^* = \frac{y}{b_0}, \quad \Gamma^* = \frac{\Gamma}{\Gamma_0}, \quad t^* = \frac{t}{t_0} \quad (2.8)$$

The impact of the meteorological conditions on wake vortex behavior also depends on the aircraft and wake vortex properties, such that wind speed u and v , the Brunt-Väisälä frequency N (see Equation (2.18)) and the Eddy Dissipation Rate (EDR) ϵ are treated non-dimensionally likewise.

$$u^* = \frac{u}{w_0} \quad , \quad v^* = \frac{v}{w_0} \quad , \quad N^* = N \cdot t_0 \quad , \quad \epsilon^* = \frac{\sqrt[3]{\epsilon b_0}}{w_0} \quad (2.9)$$

2.4 Idealized Two-Dimensional Vortices

To study the behavior of wake vortices, both measurements and numerical simulations are accomplished. In many cases the latter are initialized using vortex models. Vortex models enable analytical solutions of wake vortex transport and decay. The comparison of the tangential velocity profiles obtained from numerical simulations and analytical models shows that depending on the vortex age the identity of the best fitting model changes [87]. While the Lamb-Oseen model shows excellent accordance with the simulations for $t^* = 0.4$, the Rosenhead-Burnham-Hallock model offers better agreement for later vortex ages. The vortex core radius r_0 , constituting the radius at which the maximum tangential velocity is observed, has been found to vary significantly [88] and has not yet been conclusively investigated. Measurements [88] suggest that it may be as small as 1% of the wingspan between 0 and $3.5t_0$ on average. In other studies the relation $r_c/b_0 = 0.0625$ is applied [89, 90]. For a more comprehensive study of existing vortex models it can be referred to Gerz et al. [1].

2.4.1 Decaying potential vortex

To describe the behavior and interaction of vortices the formulation of the velocity field of a single vortex is required. Therefore, the tangential steady-state velocity distribution of a potential vortex ($\text{rot}(\vec{v}) = 0$) which is initiated by a rotating cylinder [78] is introduced as

$$V_\theta(r) = \frac{\Gamma}{2\pi r}. \quad (2.10)$$

In a viscous flow with kinematic viscosity ν the fluid will slow down due to viscous diffusion. To derive the temporal development of the velocity field of a line vortex the incompressible Navier-Stokes equation, neglecting the source term, serves as a basis [78].

$$\frac{\partial V}{\partial t} = \nu \nabla^2 V \quad (2.11)$$

With appropriate initial and boundary conditions the solution [78] can then be formulated with

$$V_\theta(r, t) = \frac{\Gamma}{2\pi r} \left[1 - \exp\left(\frac{-r^2}{4\nu t}\right) \right]. \quad (2.12)$$

The temporal evolution of the circulation distribution [78], derived from $\oint V(r, t) ds = \int_0^{2\pi} V_\theta r d\theta$, can be expressed by

$$\Gamma(r, t) = \Gamma_0 \left[1 - \exp\left(\frac{-r^2}{4\nu t}\right) \right] \quad (2.13)$$

2.4.2 Lamb-Oseen model

The Lamb-Oseen model is also an exact solution for the temporal development of a two-dimensional tangential velocity field of a potential vortex. However, in contrast to the above presented solution it incorporates the temporal development of the vortex core due to viscous diffusion [10] with

$$r_c^2(t) = r_{c,\text{init}}^2 + 4\beta\nu(t - t_{\text{init}}) \quad \text{and} \quad \frac{dr_c^2}{dt} = 4\beta\nu. \quad (2.14)$$

Equation (2.12) and Equation (2.13) can thus be formulated by [10]

$$V_\theta(r, t) = \frac{\Gamma}{2\pi r} \left[1 - \exp\left(-\beta \frac{r^2}{r_c(t)^2}\right) \right] \quad \text{and} \quad (2.15)$$

$$\Gamma(r, t) = \Gamma_0 \left[1 - \exp\left(-\beta \frac{r^2}{r_c(t)^2}\right) \right]. \quad (2.16)$$

With the condition $dV_\theta/dr = 0$ for $r = r_c$ this yields $\beta = 1.25643$ [10].

2.4.3 Rosenhead-Burnham-Hallock vortex model

In comparison to the Lamb-Oseen vortex, the Rosenhead-Burnham-Hallock vortex exhibits smaller velocities at the core radius. It is described by [89]

$$V_\theta(r, t) = \frac{\Gamma_0}{2\pi r} \frac{r^2}{r^2 + r_c(t)^2}. \quad (2.17)$$

Figure 2.4 shows that both the Lamb-Oseen and the Rosenhead-Burnham-Hallock vortex blend to a potential vortex for radii larger than r_c .

2.5 Wake Vortex Evolution Stages

The evolution of wake vortices from formation to decay can be characterized depending on the distance to the vortex generating aircraft (Figure 2.5). Depending on the lift configuration and the ambient conditions the extension of the regions may vary. In the first stage vorticity generated by the aircraft is shed off and merges. The extended near field is characterized by the roll-up and merging of the vorticity which finally results in two coherent counter-rotating vortices. In the following mid and far field the behavior of the vortices is affected by mutual velocity induction and the interaction with the environment while the vortex strength does not considerably decrease [10]. Finally, the vortices start to decay at an increased rate in the decay region [10]. The processes within those stages are discussed in the following sections.

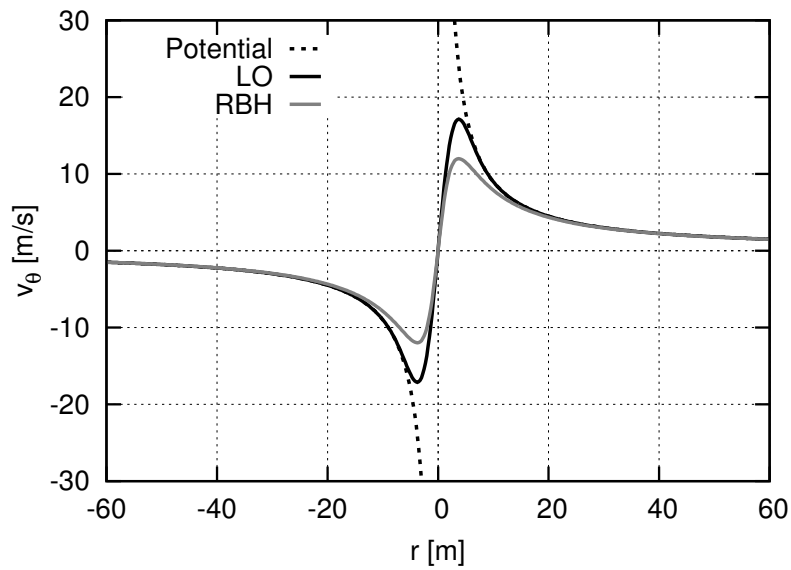


Figure 2.4.: Velocity against radius for a potential, Lamb-Oseen and Rosenhead-Burnham-Hallock vortex with $b_0 = 60$ m, $r_c = 3.75$ m and $\Gamma_0 = 565$ m²/s (corresponding to a B747 as in Ahmad et al. [89]).

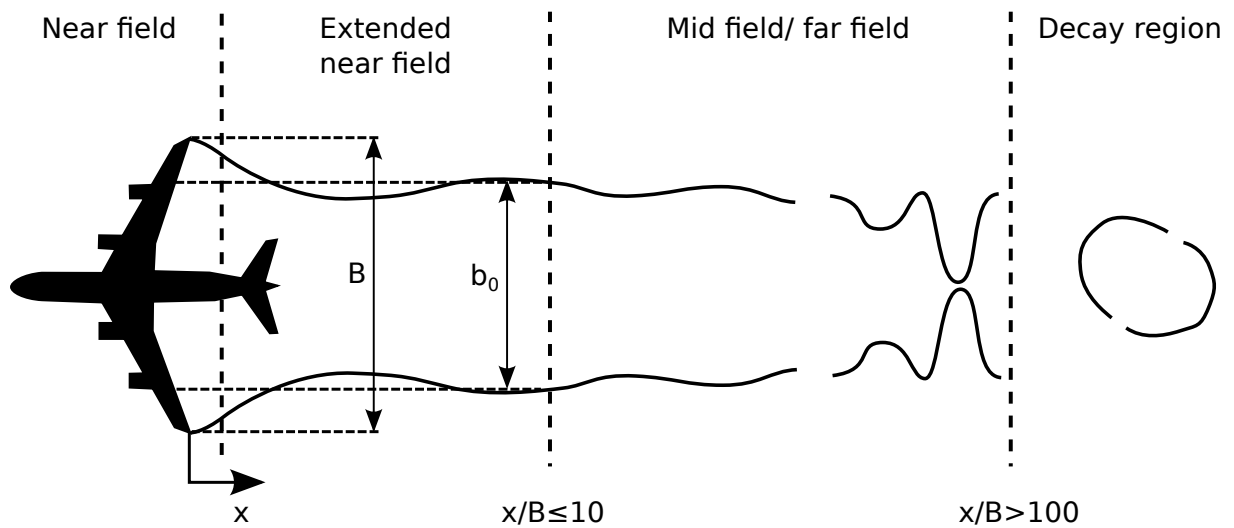


Figure 2.5.: Wake vortex evolution stages. cf. [10].

2.6 Wake Vortex Formation

2.6.1 Vortex roll-up

Within the near field the vorticity sheet is shed off the wing and the main vortex structures are formed. In high-lift configuration the generation of major vortices occurs at the wing tips and at the outer edges of the flaps. Measurements [10] and LES simulations [91, 92] imply that also the engine nacelles and the junction between wing and fuselage contribute to the complex vortex pattern behind the aircraft. Depending on the aircraft type, the sense of rotation of the wing tip vortex (WTV), the outboard nacelle vortex (ONV), the outboard flap vortex (OFV), the inboard engine nacelle vortex (INV) and the wing-fuselage vortex (WFV) may vary. For the wind tunnel model employed in the experimental study of Breitsamter [10] (half model of a four-engine transport aircraft and twin aisle configuration) all vortices but the wing-fuselage vortex have the same sense of rotation as the wing tip vortex, which is caused by the change of circulation gradient at the junction. Figure 2.6 shows the vortex roll-up achieved by a coupled RANS/LES simulation for the long-range aircraft model used in the European research project Aircraft Wing with Advanced Technology Operation (AWIATOR) with four engines and in high-lift configuration [8].

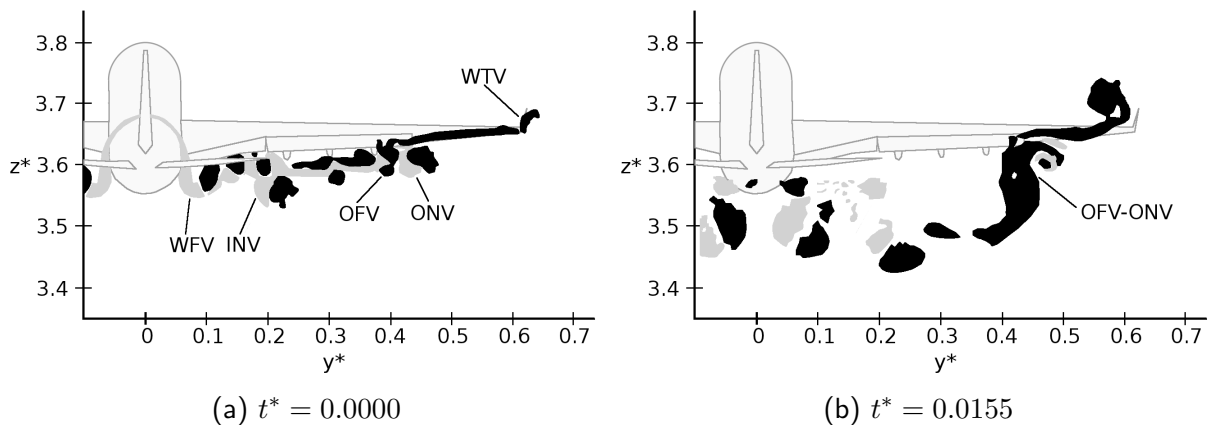


Figure 2.6.: Qualitative roll-up and merging of the vorticity generated at the wing tip (WTV), the outboard nacelle (ONV), outboard flap (OFV), the inboard engine nacelle (INV) and the wing-fuselage (WFV) for the long-range aircraft AWIATOR model. Gray contours represent clockwise and black represent counter-clockwise rotations. cf. [8]

2.6.2 Vortex merging

The extended near field is characterized by the merging of the co-rotating vortices displayed in Figure 2.7 and extends downstream over a distance of roughly 10 wingspans [10]. The merging of co-rotating vortices can be divided into three stages according to Meunier et al. [12]. Initially the vortices are axisymmetric. When the core size increases due to viscosity, the vortices start to influence each other and their shape becomes elliptic. The second stage begins when the

core size reaches a critical value in relation to the separation distance. Both vortices can still be distinguished in this phase. Meunier et al. [12] argue that the merging itself is a convective process as it is independent of the Reynolds number. In a third stage the spirals are increasingly entangled until an axisymmetrical vortex has developed. Also counter-rotating vortices may merge, triggered by instabilities which have been studied by Laporte et al. [93].

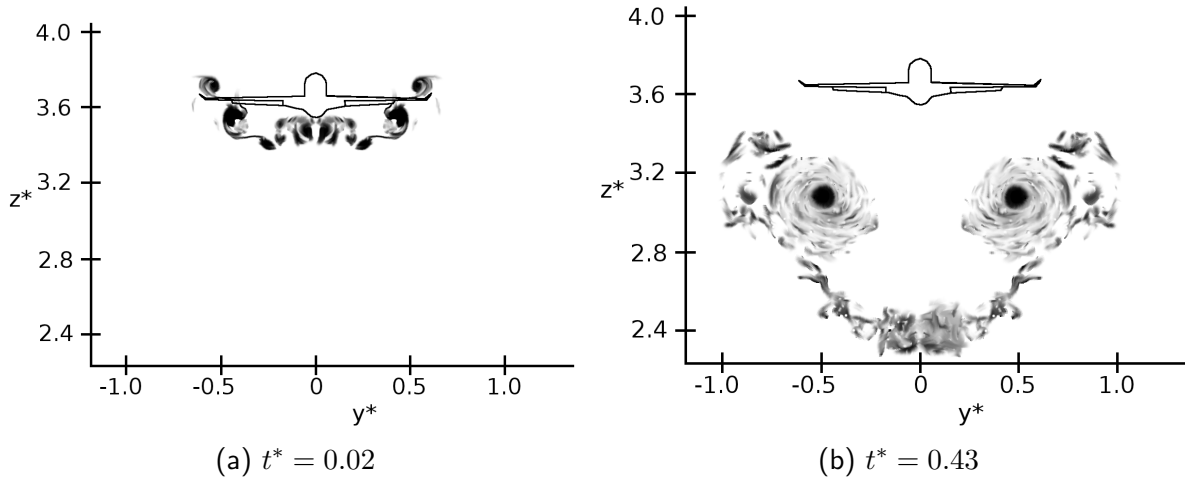


Figure 2.7.: Qualitative development of the vorticity magnitude in the extended near field. In (b) the flap and wing tip vortices have already merged. cf. [8]

2.7 Wake Vortex Descent

2.7.1 Out-of-ground effect

The fully developed vortices induce each other a mutual descent velocity according to the law of Biot-Savart. As long the vortices are not tilted the horizontal velocities of both vortices compensate one another, so that no change in lateral separation occurs without the influence of the ground. Meanwhile, both vortices are advected with crosswind speed [78] according to the 1st Helmholtz theorem. If the vortices are generated in ground proximity, their behavior is more complex and further discussed in Section 2.7.2.

Stratification The descent speed w of the vortices given in Section 2.3.1 is valid for an unstratified atmosphere under the absence of shear. Investigations showed that buoyancy effects caused by thermal stratification may slow down or even reverse the descent [94]. During its descent the air descending with the vortex pair is compressed due to increasing ambient pressure at lower altitudes [94]. Under the assumption that this process is adiabatic this results in a warming of the fluid. In a stably stratified environment a density difference between the in- and outside of the vortex oval is generated which results in a buoyancy force [94]. To describe the resulting oscillation the Brunt-Väisälä frequency N is employed. N describes the oscillation frequency of a fluid parcel that is displaced from its equilibrium [95].

$$N = \sqrt{\frac{g}{\theta} \frac{d\theta}{dz}} \quad (2.18)$$

Crosswind shear Beside the stratification, it was found that also vertical crosswind shear gradients may affect wake vortex descent substantially. Both experiments [96] and simulations [97, 98] as well as analytical considerations [99] demonstrate that a positive shear gradient accelerates the descent of the luff vortex, whereas a negative shear gradient has the same effect on the lee vortex. Due to the different descent rates the vortex pair may tilt, additionally affecting the lateral transport.

2.7.2 In- and near-ground effect

The risk to encounter wake vortices is increased at low flight altitudes during approach and landing. 78% of the reported wake vortex incidents in the United States between 1983 and 2000 occurred below 200 feet [100]. This can be referred to both the fact that all approaching aircraft follow each other on the same path, namely the glide slope, and to the interaction of vortices with the ground. For these reasons wake vortex behavior in vicinity of the ground deserves special discussion.

Figure 2.8 displays the principle of wake vortex behavior for an A320 aircraft in the glide slope from roll-up to decay, starting at an initial altitude of $z_0 = 2b_0$. After the two counter-rotating vortices have evolved they induce each other a descend velocity and start to separate while being advected by crosswind. Holzäpfel and Steen found in measurements at Frankfurt airport that the minimum height amounts on average to $0.53b_0$ for the luff and $0.62b_0$ for the lee vortex [18]. Similar behavior is found by Stephan et al. [101] in LES.

Increasing separation When the vortices have approached the ground close enough for inviscid phenomena to play a role, they start to separate. This can be modeled by introducing so-called image vortices that satisfy the slip boundary condition on the wall [102]. They are computed by mirroring the real vortices at the ground, such that they have the same strength but opposite vorticity. As the vortex separation and advection superimpose, they may compensate one another at crosswind speeds of approximately $v^* = 0.5$ [18]. This results in potentially dangerous situations when the vortices hover in the glide path after rebound.

Rebound Harvey and Perry [103] found that the vortex trajectory near ground did not coincide with what would have been expected according to inviscid theory. Instead of leveling out, the vortices rose again. They explained this phenomenon by secondary vortices that circulate the main (primary) vortices and exhibit vorticity of opposite sense, resulting in an upward directed mutual velocity induction. Their theory was supported by simulations that indeed showed eruptions of secondary vorticity structures from the boundary layer, induced by the primary vortices [101, 102, 104, 105]. The coherent detachment of redistributed vorticity is supported for the downwind (lee) vortex, where the sign of crosswind shear and vorticity is the same, and attenuated for the upwind (luff) where they have opposite signs [102]. This leads to a more pronounced rebound

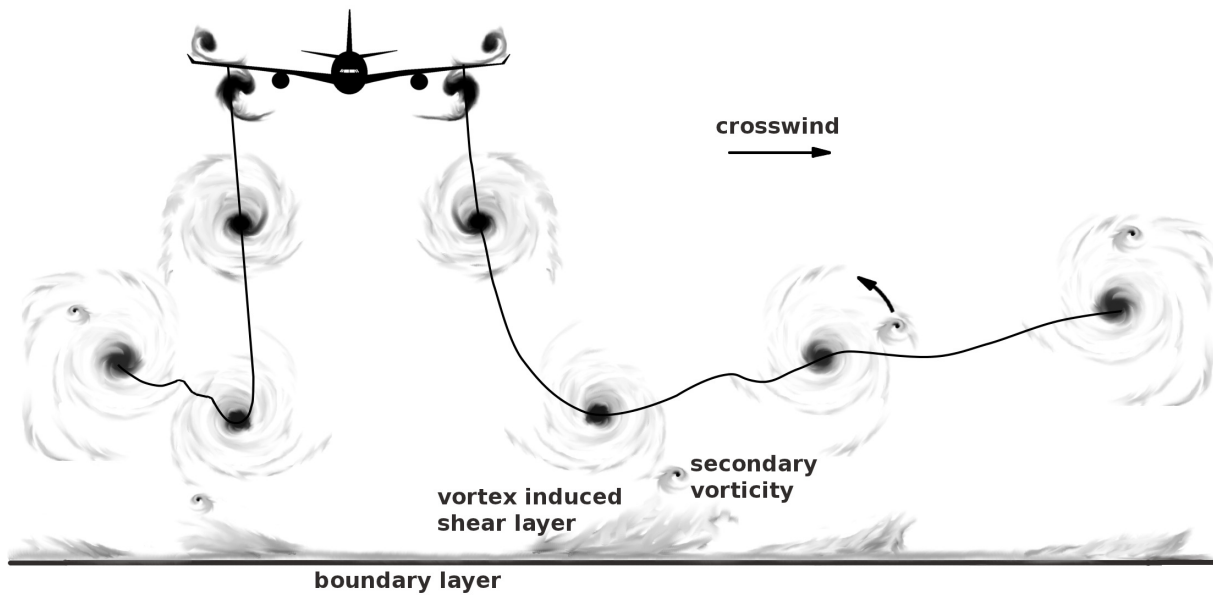


Figure 2.8.: Principle of evolution and descent of the most dominant vortices in ground proximity, displayed within a plane perpendicular to the flight path. The evolution is based on LES simulations of Misaka et al. [8] while the trajectory is derived from a D2P forecast with a crosswind speed of 0.43 m/s at flight altitude $2b_0$ of an A320 aircraft in the glide slope.

of the lee vortex [102] and furthermore shortens its lifespan due to an increased level of turbulence [101, 106]. Analytical considerations of Doligalski [107] imply that vortex rebound is suppressed if the crosswind speed exceeds $4w_0$, which has also been observed in DNS [102] and LES [22]. The maximum descent height of the lee vortex, however, increases continuously with increasing crosswind speed [22].

2.8 Wake Vortex Decay

2.8.1 Out-of-ground effect

The decay in the mid and far field is mainly driven by dissipative effects. The tangential velocities in the core start to continuously decrease due to internal diffusion (or growth of core radius) [108], whereas the circulation at larger radii is barely reduced [106]. In the following decay phase the dissolution is accelerated, which is caused by instabilities that are promoted by the atmosphere. The most prominent factor triggering decay is coherent secondary vorticity that can be caused by turbulence, shear or stratification. Interacting with the primary vortices, the secondary vortices consume energy and cause deformations [106]. These and further instability mechanisms are discussed in the following.

Turbulence Turbulence may be introduced by the flow around the aircraft (e.g. vorticity behind exposed parts such as engine nacelles or landing gear and the fuselage), by the

engine jets or by the environment. Besides, the interaction between both vortices that is connected with vortex spacing has large impact on vortex longevity, with a reduction of normalized vortex spacing of two accelerating decay by a factor of four [106]. In turn, increasing the core radii shortens the lifetime only slightly [106]. Turbulence is characterized by the Turbulent Kinetic Energy (TKE) or the Eddy Dissipation Rate (EDR, ϵ). The first is defined as

$$\text{TKE} = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}), \quad (2.19)$$

where u' , v' and w' denote the fluctuation velocities in the three directions in space respectively. EDR is the rate at which TKE is dissipated by breaking the eddies in smaller and smaller eddies [109]. It is given in m^2/s^3 by

$$\epsilon = \frac{C \cdot \text{TKE}^{3/2}}{l_0}, \quad (2.20)$$

with C denoting the proportionality constant of order one and l_0 denoting the integral length scale of the turbulence [109].

Crow instability This phenomenon, theoretically described by Crow in 1970 [110], is a sinusoidal instability that eventually results in the formation of vortex rings and is frequently observed in contrails. The time when the linking occurs is a function of turbulence [111]. Hennemann et al. [79] performed LES simulations, investigating the descent and decay of such deformed vortices. Their results reveal that a third decay stage exists in which the vortex rings mainly undergo diffusion. Hence, they may have substantial lifetimes while descending at reduced rates. Furthermore, simulations and observations reveal that the vortex rings may split and subsequently link again [112].

Baroclinic vorticity While the vortex oval descends in a stably stratified environment, it produces baroclinic vorticity [108, 113, 114] that contributes to the dissolution of vortex coherence and accelerates the decay of the primary vortices. It is generated by baroclinic torque at the border of the adiabatically heated vortex oval that is itself caused by the non-alignment of the density and pressure gradient between vortex oval and environment. This results in forces with different points of origin that act in opposite directions.

Vortex bursting and stretching This phenomenon, also occasionally visualized by contrails in funnel or pancake-shaped manner, is characterized by regions in the vortex that contract while others expand [115]. The LES studies of Misaka et al. [112] suggest that this is due to secondary vortical structures that travel along the vortex axis but do almost not affect circulation.

2.8.2 In- and near-ground effect

Additional to the above mentioned mechanisms, the vortex-ground interaction may trigger further decay phenomena. Concerning decay, the LES results of Proctor et al. suggest that

the onset of rapid decay in ground proximity starts with a time offset of $0.25t_0$ after the vortices have reached the lowest point at t_g^* [80] and start to increasingly interact with the ground. Destruction mechanisms that potentially occur by the interaction with the ground are as follows.

Secondary vorticity The interaction of the secondary vortices with the primary vortices introduces further turbulence and hence promotes the annihilation of vorticity [101]. This well-known effect has been described in detail by Holzäpfel et al. [106].

End effects LES show that when the aircraft has touched down, so-called end effects may occur [7]. Spiral disturbances are caused by a sudden reduction of circulation when hitting an obstacle or after touch down [7]. These disturbance correspond to a pressure increase inside the core, travel along the vortex and enhance its decay.

Ground linking According to the 3rd Helmholtz theorem a vortex cannot have free ends. As a consequence the vortex tube starts to link with the ground [80]. While Proctor et al. [80] state that the linked vortices rapidly dissipate after linking, the LES results of Stephan et al. [7] suggest that the ground-linking may stabilize the primary vortices which prevents them from further dissolving such that a certain level of circulation is maintained.

3. Field Measurement Data

Simulations and laboratory experiments can give valuable insights in wake vortex behavior, although they always imply simplifications. Nevertheless, they cannot substitute wake vortex field measurements as the complex behavior of the atmosphere can only be studied in real environments. With remote sensing technologies the strength and position of the vortices can be quantified in the free atmosphere. In addition, the environmental conditions in the course of the campaign must be captured. Together with the aircraft specifications they are the basis of wake vortex forecasts. This chapter presents the measurement equipment, the measurement campaigns and based on them an evaluation of wake vortex behavior.

3.1 Measurement Equipment

3.1.1 LiDAR

The rationale of LiDAR (Light Detection And Ranging) systems is based on the aerosol backscatter of a laser beam. If the aerosols are in motion, the Doppler shift of frequency between the emitted and backscattered light can be used to determine their line-of-sight (LOS) velocity. A plane perpendicular to the flight direction is scanned within a few seconds to capture the aerosol movement. By evaluating the LOS-velocity spectra, the vortex core position and the vortex circulation can be derived [116]. The maximum LOS velocity is assumed to be equal to the tangential velocity V_θ . By averaging both vortex sides, the descent speed and drifting velocity can be compensated for the evaluation of the circulation [116]. The wake vortex field-measurements of DLR were accomplished by a 2 μm pulsed laser. The advantage of using pulsed signals compared to continuous-waves is the increased range of more than 1 km and the determination of the vortex position from the time delay of the backscattered signal [116]. The sensitivity of the LiDAR is determined by the wave length of the emitted light [116].

While the circulation of a single vortex converges for large radii (Equation 2.2), the neighboring vortex in a vortex pair affects the circulation evaluation. For this reason Holzäpfel et al. suggest to average the circulations derived for various radii with an increment of 1 m [83].

$$\Gamma_{r_l-r_u} = \frac{1}{n_r} \sum_{i=l}^u \Gamma(r_i) \quad (3.1)$$

The upper (r_u) and lower (r_l) averaging boundaries must be chosen such that neither the circulation is significantly underestimated nor the influence of the neighboring vortex becomes too large (see Figure 2.3). However, this method may lead to strongly overestimated initial circulations as long as the roll-up process is not yet completed. Holzäpfel et al. [83] found that

this can be attributed to the secondary vortices being interpreted as high tangential velocities on large radii. Depending on the aircraft wingspan and vortex spacing b_0 , the circulation is distinguished between Γ_{5-15m} and Γ_{3-8m} in the data captured and evaluated by DLR. The data provided by NASA always comprises Γ_{5-15m} averages.

In a study of Köpp et al. [117] the accuracy of wake vortices measured by LiDAR was evaluated utilizing the 2- μm pulsed Doppler LiDAR from DLR and the 10- μm continuous wave Doppler LiDARs from ONERA and QinetiQ. They found that the errors in terms of standard deviations were $\sigma_{err,z} = 9 \text{ m}$, $\sigma_{err,y} = 12 \text{ m}$ and $\sigma_{err,\Gamma} = 13 \text{ m}^2/\text{s}$. Usually the tracking of the vortices and the calculation of their strength is possible until a normalized circulation between 0.6 to 0.4 is reached.

3.1.2 Sonic Detection and Ranging (SoDAR), Radio Acoustic Sounding (RASS) and Ultrasonic Anemometer (USA) System

To capture the atmospheric conditions during the evolution of the wake vortices usually a remote sensing system is employed that consists of a Sonic Detection and Ranging (SoDAR) and a Radio Acoustic Sounding System (RASS) [118]. While the SoDAR uses sonic waves to measure profiles of wind direction and speed, the RASS system sends out radio waves to capture virtual temperature profiles [119]. The basic principle of both systems is the Doppler effect.

Optimally it is desired to know the atmospheric conditions directly at the position where the wake vortices are located. However, in practice this is not applicable as the instrumentation cannot be installed directly under the glide slope in runway proximity and because aircraft induced velocities would falsify the measurements, wherefore they must be accomplished at a certain distance. The accuracies that can be achieved for wind speed measurements are $\sigma_{ws} = \pm 0.5 \text{ m/s}$ below 5 m/s, and $\sigma_{ws} = \pm 10\%$ between 5 and 35 m/s [120]. The wind direction can be measured with an accuracy of $\sigma_{wd} = \pm 5^\circ$ [120].

In addition, an ultra-sonic anemometer (USA) is utilized to estimate wind speed and the turbulence level from the longitudinal velocity spectrum at a height of 10 m.

3.2 Campaigns

This section is commenced by a brief introduction of the campaigns which is followed by a discussion of the distributions of the initial altitudes, initial vortex separation and meteorological conditions in Section 3.3. Headwind u is negative, while tailwind is considered to be positive. Positive crosswinds v occur if they are directed from left to right from the pilot's perspective and negative if otherwise.

3.2.1 WakeFRA

WakeFRA has been conducted in 2004 at Frankfurt airport from August to December [18]. 292 landings of heavy aircraft have been captured by LiDAR. Additionally, a METEK SoDAR with a RASS extension provided 10-minute averaged vertical profiles of the three wind components, vertical fluctuation velocity, and virtual temperature with a vertical resolution of 20 m. The

measurements offer high quality and have thus been used to train P2P during its development in the past. The 2- μm LiDAR was positioned south-east of the runways with a lateral offset of 373.5 m to runway 25L and 891.5 m to runway 25R such that landings on both could be captured. The vertical profiles of *EDR* were also derived from LiDAR with a temporal averaging of 5 minutes. A detailed description is given by Holzäpfel and Steen [18].

3.2.2 WakeMUC

In 2011 the WakeMUC campaign [22] has been accomplished at Munich airport. Between 11th of March and 12th of May 2011 LiDAR measurements of 779 landings of heavy and medium type aircraft have been collected on runway 26L [22]. The LiDAR was positioned at a distance of 850 m northerly of the runway center line. To measure the meteorologic data, the same system as for WakeFRA has been employed. However, *EDR* was not measured by LiDAR which is why only *TKE* is available. During the 12 measurement days the angle of the laser beam in reference to an observation direction perpendicular to the aircraft path has been adjusted multiple times. This way, landings with various vortex generation heights could be captured [22]. Before utilizing the data for the evaluation of the models, a quality check reduced the number of landings to 374, whereas the whole set was employed to evaluate wake vortex physics. As the models rely on *EDR* rather than on *TKE*, the relation found by Donaldson and Bilanin [121] is used.

$$\text{EDR} = \epsilon = \begin{cases} \frac{1}{8} \frac{q^3}{0.65z} & \text{for } 0 < z < 169 \text{ m} \\ \frac{1}{8} \frac{q^3}{110 \text{ m}} & \text{for } z > 169 \text{ m} \end{cases} \quad (3.2)$$

with

$$q = \sqrt{u'^2 + v'^2 + w'^2} = \sqrt{2 \text{ TKE}}, \quad (3.3)$$

where u' , v' and w' denote the fluctuation velocities in the three directions in space respectively.

Unfortunately, neither aircraft type, initial position nor weight were recorded during the campaign. As a consequence, initial vortex position and circulation need to be taken from the initial LiDAR measurements. While the initial aircraft height and lateral position are derived from linear extrapolation of the first data points, the initial circulation is computed from the first luff and lee vortex measurement as an average. In addition, the time stamps of the LiDAR measurements were corrupted and could not be assigned to the time stamps of the meteorological data. As a consequence, they were corrected by shifting the time stamps by mapping landings with large measured circulation to landings of heavy aircraft derived from STANLY Track, a service provided by the DFS to track approaches and departures at various airports. The time shifts found to derive the correct UTC amount to -51 minutes for cases before March 30th 2011 and -75 minutes after. Nevertheless, it was not possible to assign the correct aircraft types by using this method.

3.2.3 WakeOP

Conducted in 2013 at special airport Oberpfaffenhofen, 74 overflights of the research aircraft HALO were accomplished in order to demonstrate the capability of the plate lines to enhance wake vortex decay [22]. In this study only the 31 overflights without plates are utilized. The ambient conditions were measured at three different heights: by the HALO nose boom at flight altitude, by the 10 m ultra-sonic anemometer mast and by two wind and temperature sensors at 2 m height. This offers very accurate wind data very close to where the vortices have been generated at flight altitude but biased measurements at lower altitudes where the instruments may have been affected by the aircraft induced flow.

3.2.4 MEM95

In the mid 1990s several wake vortex field measurements were sponsored under the Aircraft Vortex Spacing System (AVOSS) project [122]. One of them was the MEM95 dataset that has been gathered in 1995 at Memphis airport which is the base of FedEx. During the campaign 305 landings and the corresponding meteorological conditions have been captured. Wind and temperature profiles were derived from a SoDAR/RASS system and were complemented by a 150 foot tower that measured wind direction, temperature and humidity at five elevations [122]. The meteorological conditions were accomplished between runway 36L and 36R and were situated at a distance of about 2 km from the closest LiDAR measurement site [24]. Note that at the southernmost measurement site trees are situated beneath the glide path which may impact wake vortex behavior.

3.2.5 DFW97

In 1997 another NASA wake vortex field experiment was conducted, this time at Dallas/Fort Worth International Airport [123]. The wake measurements were accomplished with a 2 μm LiDAR by NASA and by a 10.6 μm continuous wave LiDAR from the Lincoln Labs. To measure the atmospheric properties at different heights, two meteorological towers were installed. In addition, a SoDAR/RASS system was deployed. Turbulence was measured by ultrasonic anemometers (10 Hz).

3.2.6 DEN03

In September 2003 NASA conducted a field experiment at Denver airport with the objective to evaluate wake measurements using acoustic sensors [124]. Nevertheless, the trajectories of the wake vortices and their circulation were measured by pulsed and continuous wave LiDARs. The temperature profiles were measured with a microwave radiometer (MTP5) while wind profiles were derived from pulsed LiDAR measurements. In addition, one minute average winds were measured by towers that were equipped with propeller anemometers at 7 m, 14.6 m, and 32.3 m above ground and an ultrasonic anemometer (10 Hz) at 7 m. Profiles of *EDR* were obtained from the LiDAR measurements using spatial structure functions [125]. This dataset has not been used to assess the ensemble because it was made available too late. Nevertheless, in the Appendix the performance achieved by the individual models is listed.

3.3 Distribution of Aircraft and Meteorological Parameters

Between the campaigns the distributions of the initial conditions, such as aircraft size and initial altitudes, and the meteorological conditions may differ significantly. Figure 3.1 depicts the aircraft initial condition distribution for each set of measured landings. While mainly landings with initial heights at about $1 b_0$ have been captured during WakeOP and WakeFRA, WakeMUC comprises wake vortices generated between 0.7 and $4 b_0$. In comparison, the campaigns contributed by NASA cover higher initial heights. During DFW97 the vortex generating aircraft were at altitudes between 2 and $5 b_0$, with some landings even below $0.5 b_0$. Memphis covers a larger range, starting from 0.7 and up to more than $10 b_0$. Eventually, the DEN03 dataset contains the largest initial heights with z_0 from 5 to $11 b_0$. Concerning the initial vortex spacing b_0 , Figure 3.1 right reveals that all campaigns mostly involve medium-sized aircraft, with exception of WakeFRA, where the larger part of the dataset is composed of heavy aircraft.

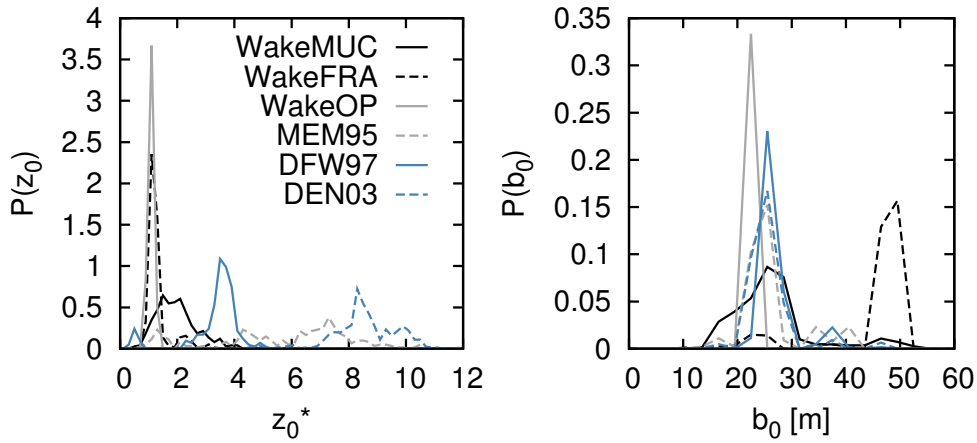


Figure 3.1.: Distribution of initial height z_0 and initial wake vortex spacing b_0 for each campaign.

Figure 3.2 features the headwind and crosswind conditions for all campaigns. As the NASA models do not include headwind in their initial conditions, u is only available for the DLR campaigns. The left panel indicates that during normal airport operation aircraft rarely land under tailwind conditions. And even if they do the winds are weak. The reason for head- and tailwind being equally distributed for WakeOP is that HALO always had the same heading during the conducted traffic patterns, with no distinguished wind direction. As far as crosswind is concerned the calmest conditions can be found during WakeOP and DFW97, whereas the crosswinds are relatively strong for WakeFRA and DEN03. Nevertheless, DEN03 does not exhibit the largest turbulence levels (Figure 3.3 left) which could be referred to the high initial heights in this campaign where the ground has less impact. Instead, WakeMUC is the campaign that offers the highest EDR values, followed by WakeFRA. It is worth to mention that the TKE values (Figure 3.3 right) show very similar distributions for the latter two campaigns, while the EDR distributions deviate notably. It seems that the computation of EDR according

to Donaldson and Bilanin [121] from *TKE* results in larger values compared to measurements. With respect to the Brunt-Väisälä frequency, it is found that the largest stratification can be observed in the WakeMUC campaign.

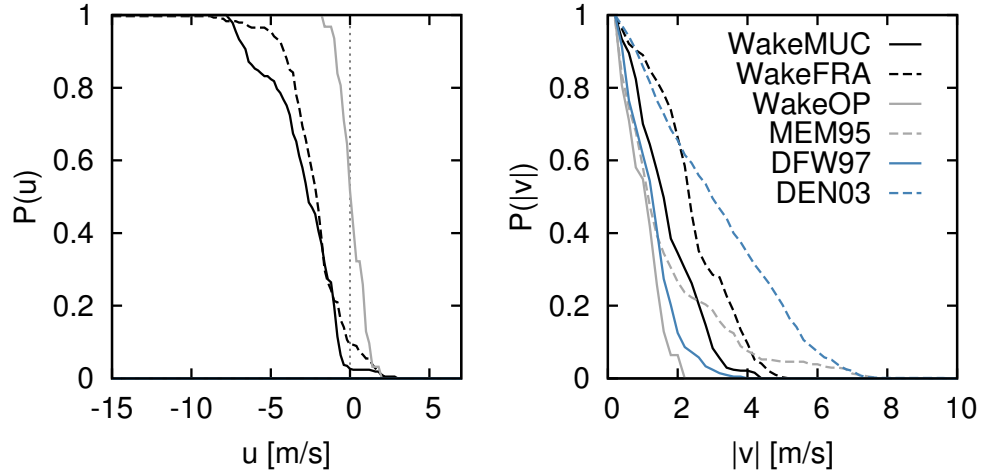


Figure 3.2.: Cumulative distribution of the averaged head- and crosswind from ground to initial height.

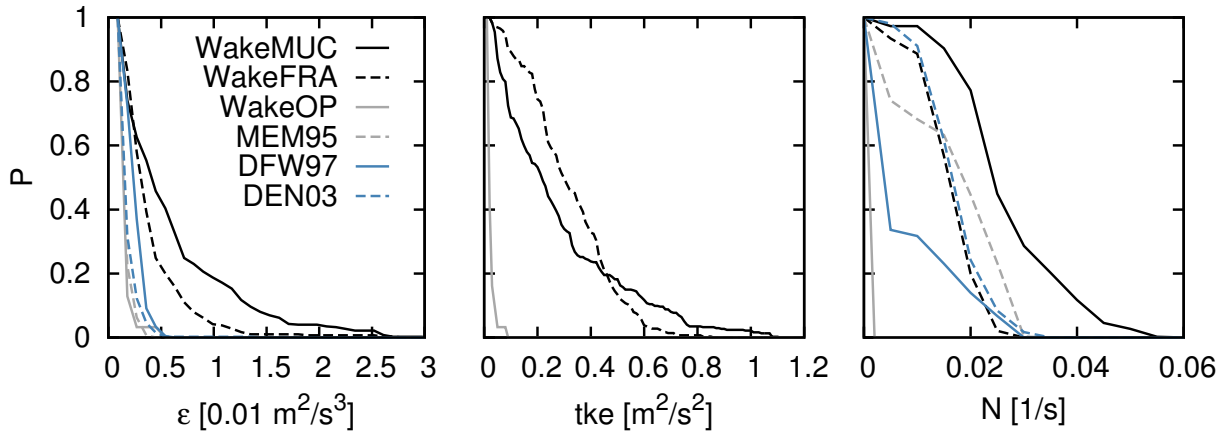


Figure 3.3.: Cumulative distribution of ϵ , *TKE* and N , averaged from ground to initial height.

4. Wake Vortex Prediction

This chapter is devoted to the properties and the performance of the employed wake vortex models from NASA and DLR. For the performance evaluation measurements from the campaigns WakeMUC, WakeFRA, WakeOP, MEM95 and DFW97 are employed.

4.1 Skill Metrics

To evaluate the model performance, the predictions are compared with wake vortex measurements. Various skill metrics are employed that are discussed in this section.

Error The error err_i in this thesis is referred to as the difference between the forecast f_i and a single observation o_i at a given vortex age.

$$err_i = f_i - o_i \quad (4.1)$$

Root-mean-square error (RMSE) The *RMSE* considers outliers more than the Mean Absolute Error (*MAE*) as the errors are treated quadratically via

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (f_i - o_i)^2}{n}} \quad (4.2)$$

Bias The bias specifies the model deviation from the measurement including its sign. In this thesis, models with negative (positive) bias underestimate (overestimate) the observation on average if compared to measurements.

$$bias = \frac{1}{n} \sum_{i=1}^n (f_i - o_i) \quad (4.3)$$

In this thesis, the above introduced metrics are computed for single landings that usually comprise several measurements. Eventually, to evaluate the overall model performance the *median* error of all simulated landings is derived. Compared to the *mean*, the *median* is more robust regarding extreme error values.

4.2 The Fast-Time Models

4.2.1 Common model features

The ensemble is based on the forecasts of four different models which all predict the trajectory and circulation development of the vortices in a plane perpendicular to the aircraft flight path.

Their forecasts are based on aircraft side on input parameters such as weight, speed, wingspan and initial position. Concerning the ambient conditions their predictions rely on vertical profiles of crosswind speed, headwind speed, temperature and atmospheric turbulence, with the latter being represented by the eddy dissipation rate ϵ . For the initial state of the vortices the models assume that all vorticity created due to lift is rolled up into a pair of counter-rotating vortices, with luff and lee vortex decaying at the same rate. All models treat the wake vortex parameters non-dimensionally according to the convention presented in Section 2.3.2.

Further commonalities can be found between the lateral vortex transport that is driven by the crosswind. Likewise, the lateral divergence in ground proximity is modeled by image vortices, which are created by mirroring the actual vortices at the ground. The interaction of real and image vortices causes the vortices to depart from each other as it would be expected according to inviscid theory.

Differences can be found regarding the parametrization of vortex decay and descent, the parametrization of the effects of the atmospheric conditions and the parametrization of the interaction with the ground.

4.2.2 D2P

Both D2P and P2P are wake vortex models developed at DLR [15]. Vortex decay in D2P is based on an analytical solution of the Navier-Stokes equations for the decaying potential vortex (see Equation (2.13)). It has been extended to model the turbulent decay of a pair of Lamb-Oseen vortices and split into a diffusion and into a rapid decay phase. Circulation is calculated as a 5-15 m average. The first phase is described by

$$\Gamma_{5-15}^*(t^*) = \frac{1}{11} \sum_{r=5m}^{15m} \left[A_{D2P} - \exp \left(\frac{-r^{*2}}{4\nu_1^*(t^* - T_1^*)} \right) \right], \quad (4.4)$$

where A_{D2P} , ν_1^* and T_1^* are constants with $-T_1^*$ corresponding to the age of the vortices at $t^* = 0$ [15]. A_{D2P} is used to adjust $\Gamma_{5-15}^*(t^* = 0)$ [15] and $\nu_1^* = 1.78 \cdot 10^{-3}$ ($\nu_1 = 0.16$ m²/s) is termed the effective viscosity. In the second phase a second term with T_2^* , the onset time of rapid decay, and ν_2^* , the adjusted effective viscosity, is introduced. For simplicity this is formulated without averaging over different radii by

$$\Gamma_{5-15}^*(t^*) = A_{D2P} - \exp \left(\frac{-R^{*2}}{\nu_1^*(t^* - T_1^*)} \right) - \exp \left(\frac{-R^{*2}}{\nu_2^*(t^* - T_2^*)} \right), \quad (4.5)$$

with the mean radius R^* approximately corresponding to the mean value of 10 m [15]. Both T_2^* and ν_2^* are determined in dependency of the meteorological conditions as presented in Holzäpfel 2003 [15].

Concerning vortex descent, the relation $w^* = \Gamma^*$ only holds if Γ^* represents the circulation at the neighboring vortex [15]. This is not valid here due to the use of 5-15 m Γ^* averages [15]. Therefore, the descent speed is calculated according to the self-similar velocity profile of a potential vortex as a function of vortex spacing b and core radius r_c , adapted to the behavior observed in LES [15].

$$w^* = 1 - \exp\left(\frac{-1.257b^2}{r_c^2}\right) \quad (4.6)$$

The core radius is then derived from Γ_{5-15}^* that is formulated according to

$$\Gamma_{5-15}^* = \frac{1}{11} \sum_{r=5m}^{15m} \left[1 - \exp\left(\frac{-1.257r^2}{r_c^2}\right) \right]. \quad (4.7)$$

The comparison with LES data showed that these equations are a useful approximation with an effective vortex spacing of $b = 0.4b_0$ [15]. In strongly stratified environments additional terms that depend on the Brunt-Väisälä frequency N are introduced in order to allow vortex rebound.

When the vortices approach ground they start to diverge laterally which is modeled by introducing image vortices. Secondary vorticity that detaches from ground is modeled using point vortices that orbit the primary vortices. Another set of secondary image vortices orbits around the respective image vortices. Crosswind encourages the detachment of secondary vorticity for the lee vortex and suppresses it for the luff vortex [18]. The strength of the secondary vortices is adapted according to crosswind speed [18]. Measurements from WakeMUC showed that the lee vortex decays on average slightly faster than the luff vortex [22]. This effect is, however, disregarded in the model physics. The impact of headwind is considered by D2P by adjusting vortex age and generation height. In comparison to calm situations headwind advects younger vortex segments into the observation plane [16].

The impact of shear on vortex descent is neglected in D2P as it is believed that the vortex-shear-layer interaction is extremely sensitive to a number of shear layer parameters [16, 126]. Below vortex generation altitudes of b_0 vortex decay is assumed to be faster the lower the vortices are generated, rather than being affected by turbulence or stratification [18].

P2P is the probabilistic version of D2P and provides envelopes with the default uncertainty levels 3σ and 2σ [16, 18]. The concept behind the probabilistic envelopes is based on three layers. In a first step, three model runs are accomplished with varied decay parameters. Additionally to the resulting envelopes, fixed uncertainty allowances are added for vortex location and strength. In the second step, dynamic uncertainty allowances are superimposed that consider stochastic vortex deformation and transport driven by environmental turbulence and wind shear.

In a final step, the probabilistic envelopes for vortex position and strength are calibrated employing field measurement data. For this purpose unbounded Johnson distributions are fitted to the error Probability Density Function (PDFs) generated by the deviations between measured vortex behavior and the predicted probabilistic bounds achieved by steps one and two. Uncertainty levels can be selected arbitrarily from the predicted Johnson distributions and are adjusted by default to either 2σ or 3σ . Pronounced shear increases the widths of the uncertainty allowances.

Further development

During the evaluation of the field measurement campaigns it became evident that the vortex rebound in D2P was exaggerated. One of the reasons for this is that up to this point the

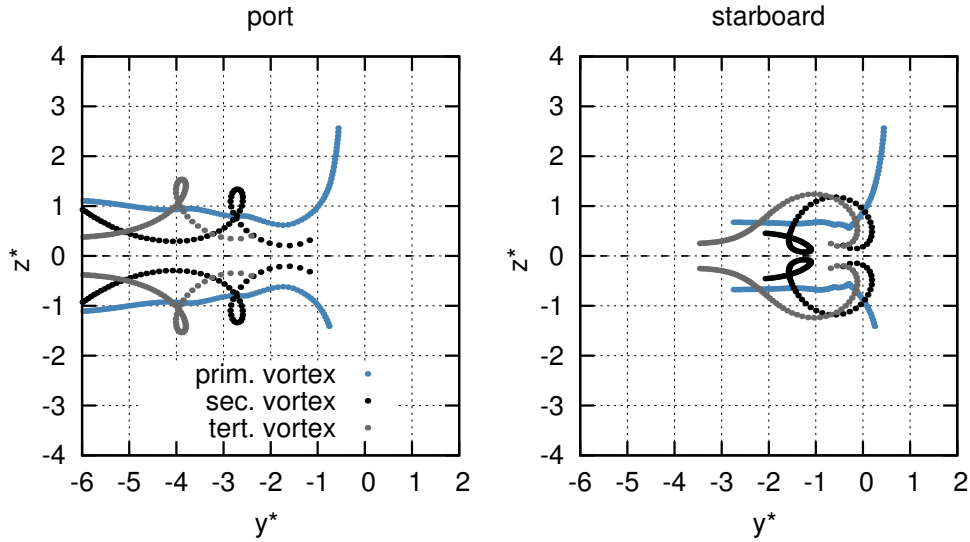


Figure 4.1.: Trajectories of the primary, secondary and tertiary vortices and their respective image vortices exemplarily for a case with crosswind from the right.

strength of the secondary and tertiary vortices did only decrease proportionally to the primary vortices and did not depend on the distance between vortex and shear layer. However, in theory one would expect that the secondary vortices detaching from the shear layer become weaker the higher the primary vortices. For this reason the optimal magnitude of weakening is explored in a parameter study. As a result, the secondary vortices are weakened by 30% after they have absolved their first orbit around the primary vortices. In addition, the tertiary vortices are weakened by 30% when they are generated. Consequently, the maximum strength of the secondary vortices amounts initially to 40% of the strength of the primary vortex and is reduced to 28% after their first orbit - the same strength as the tertiary vortices. The resulting vortex trajectories are exemplarily depicted in Figure 4.1 for a case with a mean crosswind of $v^* = 0.6$.

Table 4.1 lists the RMS deviations for the DLR measurement campaigns with and without weakening of the secondary and tertiary vorticity. As the Γ^* -forecast is only indirectly influenced by the adaptations, its RMS deviations remain nearly unmodified. For the assessment with the WakeMUC measurements 374 landings of high quality were employed. While the achieved improvement is substantial for the y_{lee}^* -forecast, it is minor for the z^* -prediction. Significant improvements for the z^* -prediction can only be revealed when restricting the WakeMUC data to $z_0^* < 1.5$. This can be readily explained by two opposing trends. While the legacy version of D2P overpredicts vortex rebound, it overestimates vortex descent starting from high altitudes. The latter trend is even increased by the described adaptations and compensates the improvements achieved at lower altitudes. Potentially, the reasons for overpredicted vortex descent could be related to the far-field of the vortices that shed off secondary vorticity from the ground earlier than expected. This trend is further discussed in the results section. Further decrease of the strength of secondary and tertiary vorticity leads to an increase of the RMS deviation for z^* while y^* is further improved. Obviously the reasons for the decreased RMS deviations for the y^* -prediction are not directly related to the improved z^* -forecast.

Table 4.1.: Comparison of the median RMS errors achieved with and without the adaption of the strength of secondary and tertiary vortices.

campaign	$\Gamma_{\text{sec/tert}}$ factor	RMSE Γ^*		RMSE y^*		RMSE z^*	
		luff	lee	luff	lee	luff	lee
WakeMUC	1	0.113	0.112	0.690	0.550	0.184	0.203
	0.7	0.113	0.112	0.690	0.540	0.180	0.203
WakeMUC, $z_0^* < 1.5$	1	0.111	0.118	0.629	0.517	0.163	0.193
	0.7	0.110	0.117	0.617	0.497	0.151	0.177
	0.5	0.110	0.117	0.614	0.494	0.157	0.181
WakeFRA	1	0.114	0.103	0.472	0.516	0.136	0.169
	0.7	0.115	0.103	0.475	0.504	0.125	0.160
WakeOP	1	0.090	0.097	0.685	0.535	0.182	0.233
	0.7	0.090	0.097	0.647	0.549	0.161	0.178

With regard to WakeFRA also significant improvements can be achieved for z^* . However, note that for this campaign the initial conditions were not determined as accurately as at the time when the parametrization of the ground effect has been developed with this data [18]. Instead, the focus was laid on estimating prediction skills closer to practice. For this reason the improvements could be less if applied to the original training data. For WakeOP the decrease of RMS deviations for z^* are even more significant than for WakeMUC. Conclusively, the suggested adaptations prove to be successful for all evaluated DLR campaigns.

4.2.3 APA

APA, the AVOSS (Aircraft Vortex Spacing System) Prediction Algorithm discriminates between Out-of-Ground Effect (OGE), Near-Ground Effect (NGE) and In-Ground Effect (IGE) [14]. If the vortices are OGE the model utilizes a decay and transport model, such as developed by Sarpkaya with two vortices descending at equivalent rates [13, 21]. His equations are based on the rate of change of impulse in dependency of the buoyancy force due to stratification which includes the area of the vortex oval A_O and the force due to the rate of change of circulation [14, 127].

$$\frac{dI}{dt} = 2\pi\rho \frac{d(b^2w)}{dt} = -\rho A_O N^2 z - \rho \frac{d(b\Gamma)}{dt} \quad (4.8)$$

On the basis of the MEM95 and DFW97 field data it is then assumed that the decay of a vortex pair can be described by $e^{-C_1 t^*/T_c^*}$, with T_c^* being the time at which a catastrophic demise event such as Crow instability or vortex bursting occurs that depends on the eddy-dissipation rate ϵ [127]. Additionally, the effect of stratification is included via the Brunt-Väisälä frequency N [14].

$$\Gamma^* = \exp \left(-(C_1 + 0.25N^*(z)^2) \frac{t^*}{T_c^*(\epsilon^*)} \right). \quad (4.9)$$

The vortex descent speed w can then be formulated by combining Equations (4.8) and (4.9)

to

$$\frac{dw}{dt} = -C_2 \int_{z_0}^z N(z)^2 dz - 2w \left(\frac{1}{b} \frac{db}{dt} \right) + \frac{b_0}{b} \left[M_c \frac{w_0^2}{b_0} - w_0 \left(\frac{1}{b} \frac{db}{dt} \right) \right] e^{-M_c t^*}, \quad (4.10)$$

with M_c being defined as

$$M_c = \frac{C_1 + 0.25N^{*2}}{T_c^*}. \quad (4.11)$$

The descent speeds observed in measurements are faster than the calculated descents despite agreeing circulations. This is assumed to be related to Crow instability and is modeled in APA by a slow decrease of effective vortex separation initially [13].

$$\frac{b}{b_0} = \left(\frac{1 + \alpha\epsilon^*}{1 - \alpha\epsilon^*} \right) \left(\frac{1 - \alpha\epsilon^* e^{t^*/T_c^*}}{1 + \alpha\epsilon^* e^{t^*/T_c^*}} \right) \quad \text{for } t^* < T_c^* \quad (4.12)$$

The parameter α is set 0.5. After T_c^2 , vortex rings or other deformations may occur, such that assuming organized descent is no longer valid [13]. Therefore, the following expression is used, where $K = 5/\epsilon^*$ and $\alpha = 0.5$:

$$\frac{b}{b_0} = \frac{b(T_c^*)}{b_0} \left\{ 1 - \frac{2e\alpha\epsilon^*/K}{1 - (e\alpha\epsilon^*)^2} \left[1 - \exp \left(-K \frac{t^* - T_c^*}{T_c^*} \right) \right] \right\} \quad \text{for } t^* > T_c^* \quad (4.13)$$

At altitudes below $1.5z_0$ (NGE) the algorithm no longer employs the above introduced Sarpkaya model but the equations given by Donaldson and Bilanin [121], while applying the same decay rate that occurs at the transition from OGE. At this stage image vortices are introduced [128]. The equations are given exemplarily for one vortex only but are further described in Robins and Delisi 2002 [14]. The subscripts 1 and 2 denote the port and starboard vortices, whereas 3 and 4 describe their image vortices.

$$\frac{dz_1}{dt} = \frac{1}{2\pi} \sum_{i=2}^4 \frac{\Gamma_i(t)(y_i - y_1)}{r_{i1}^2} \quad (4.14)$$

$$\frac{dy_1}{dt} = -\frac{1}{2\pi} \sum_{i=2}^4 \frac{\Gamma_i(t)(z_i - z_1)}{r_{i1}^2} + v(z_1) \quad (4.15)$$

Even closer to the ground at $0.6b_0$ secondary vortices are generated together with their image vortices [14, 128] (D2P uses a similar formulation). The secondary vortices are unconstrained, and their positions are calculated by the interacting vortex dynamics [128]. Further, crosswind shear can affect the distance between primary and secondary vortices for the luff and lee vortex differently, which may lead to an asymmetric rebound. The decay is then computed with a formula derived from a Terminal Area Simulation System (TASS) LES study [80].

The difference between the two deployed APA versions lies in the approach for the OGE region. In APA 3.4 the effect of buoyancy triggered by ambient stratification in the OGE phase

is reduced. In contrast to APA 3.2, the newer version does not include the possibility to model the linking process that occurs when the vortices experience Crow instabilities by decreasing the vortex separation b as a function of time. Instead the vortices are separated by $b = b_0$ throughout the OGE phase [34]. Furthermore, the vortices do not separate in the OGE region, that is $b = b_0$ throughout the OGE phase [34]. Together with divergent decay characteristics the results of both versions are well-distinguishable.

4.2.4 TDP

TDAMP retains the IGE and NGE formulation of APA [34]. However, the transition between OGE and NGE is lowered to $1.0z_0$ as the model's OGE already contains NGE physics [34]. Additionally, the OGE module differs from the APA formulation. The motivation for the development of TDAMP (TASS Driven Algorithms for Wake Prediction) were the results from the TASS [129] which showed that stable stratification causes vortex circulation to decay more rapidly at large radii than near the core ($r \leq 15$ m) [17]. Therefore the TDAMP model uses separate prognostic equations for vortex descent rate and circulation in OGE which are applied separately to the port and starboard vortices. For vortex transport the circulation at $r = b_0$ is modeled by

$$\frac{d\Gamma^*}{dt^*} = \frac{d\Gamma_L^*}{dt^*} + \frac{d\Gamma_D^*}{dt^*} + \frac{d\Gamma_S^*}{dt^*}, \quad (4.16)$$

where the terms on the right-hand side are contributions from linking instability (Γ_L), turbulence diffusion (Γ_D), and stratification (Γ_S). They are described in more detail by Proctor et al. [17]. In a similar manner the average circulation between 10 and 15 m is calculated by

$$\frac{d\Gamma^*}{dt^*} = \frac{d\Gamma_{SS}^*}{dt^*} + \frac{d\Gamma_D^*}{dt^*} + \frac{d\Gamma_S^*}{dt^*}. \quad (4.17)$$

Here Γ_{SS} represents the change in average circulation from short- and long-wave instabilities [17]. It is based on a hyperbolic tangent function with tuning parameters that are derived from TASS results [17].

This version of TDAMP considers the effects of crosswind shear on vortex descent rate and thus allows the prediction of vortex tilt and change in lateral separation due to crosswind above $1.0z_0$ (OGE) [21, 98]. Therefore, the term $d\Gamma_{sh}^*$ is added to Equation (4.16) that depends on the second derivative of ambient crosswind [17]. A negative shear gradient causes the luff vortex to descend faster whereas a positive shear gradient accelerates the descent of the lee vortex (see Fig. 4.2).

4.2.5 Necessary adaptations to the model output

In order to prepare the models for their use within the ensemble their output data needs to be adjusted. Firstly, the forecasts are continued even after the vortices already have decayed. While the vortices are assumed to stay at a constant altitude and to be further advected by the wind, circulation is extrapolated, such that Γ may take unphysical negative values. Otherwise, the uncertainty envelopes would converge to values other than zero as they would be centered at zero in Γ -direction after some point. The latter is however considered unrealistic as this

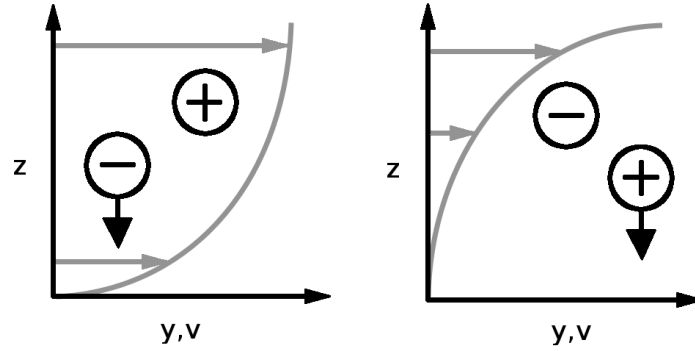


Figure 4.2.: Influence of shear on descent in TDP 2.1. The port vortex is illustrated by the - symbol (negative rotational direction), whereas the + symbol represents the starboard vortex (positive rotational direction), assuming a right-handed coordinate system. The left panel displays the impact of $\partial^2 v / \partial z^2 < 0$ and the right of $\partial^2 v / \partial z^2 > 0$, c.f. [98].

implies that even after endless time the probability to encounter vortices is still given. Secondly, the model output is mapped to a temporal grid with bin width $0.1 t_0$.

4.3 Model Evaluation

4.3.1 Initial conditions

The initial conditions of the NASA data are consistently given in form of initial vortex descent speed, wingspan, weight, aircraft position and speed as well as air density. In contrast, the DLR data partly requires the estimation of the starting values from LiDAR data.

WakeMUC

In the WakeMUC campaign the aircraft types are unfortunately unknown (see Section 3.2.2), such that the initial conditions are extrapolated from the first two LiDAR measurements for y_0^* and z_0^* . To derive Γ_0^* the first luff and lee measurements are averaged under the assumption that the vortex age given by the LiDAR is trustable.

WakeFRA

The aircraft types and the landing speed in WakeFRA are known. The wingspan and the Maximum Landing Weight (MLW) are derived from a table, consisting of theoretical values. Although employing the MLW seems overly conservative, it turns out that utilizing lower values would significantly underestimate the circulation in this campaign. If the aircraft speed is unknown, a speed of $u_{ac} = 70$ m/s is assumed. The initial values y_0^* and z_0^* are derived by extrapolation in the same manner as in WakeMUC.

WakeOP

In WakeOP the initial altitudes, the aircraft speeds and aircraft masses are available. y_0^* is derived from the measurements. Extrapolation of y_0^* does not deliver satisfying results due to partly fluctuating lateral vortex positions in the beginning, such that the initial value of y_0 is averaged from the first luff and lee measurement.

4.3.2 Sensitivity analysis

To compare the model's performance under various ambient conditions, a sensitivity study is conducted. Therefore, the meteorological parameters v , N and ϵ are adjusted to a set of different values following the exercise conducted at WakeNET New Orleans [130]. The parameters of the vortex generating aircraft are $b_0 = 50.6$ m, $u_{ac} = 70$ m/s and $\Gamma_0 = 500$ m²/s. When wind is non-zero a logarithmic wind profile is employed according to Figure 4.3 left with the wind speed at 10 m ($0.2b_0$) serving as reference. Additionally, wind profiles with inverted shear are included that demonstrate the susceptibility of the NASA models to shear (see Figure 4.3 right). The examined dimensional meteorological input parameters and their respective non-dimensional values are listed in Table 4.2. In the following, non-dimensional results are discussed. By design, the NASA models continue to predict the last vortex position value after the vortices have decayed. However, for the sake of clarity this is not displayed in the following figures. An overview of the model sensitivity in general is given in Tables 4.4 and 4.3. While D2P is not listed to be sensitive to shear, P2P widens its envelopes under shear conditions [16].

The model characteristics in OGE are illustrated by Figure 4.4. Under calm conditions (no stratification, no turbulence) the longest lifetime is predicted by TDP 2.1 with about $15t_0$ (more than 480 s). Very similar decay characteristics can be observed for D2P and APA 3.2. In all forecasts the vortices reach less than $1b_0$ above ground, with D2P predicting the fastest descent. The rebound of APA 3.2 is the weakest in this case. Concerning the lateral separation in ground proximity D2P shows the strongest and APA 3.2 the weakest divergence of the vortices.

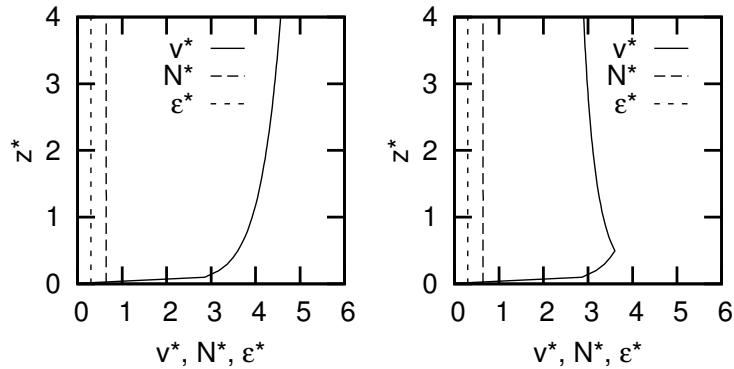


Figure 4.3.: Exemplary meteorological profiles with logarithmic wind profile (a) and inverted shear that starts from $0.4b_0$ (b). In both cases the crosswind v^* is equal to $3.2w_0$ (5 m/s) at a height of $0.2b_0$ (10 m).

Table 4.2.: Dimensional and non-dimensional meteorological parameters used for the sensitivity analysis.

v	v^*	N	N^*	ϵ	ϵ^*
1 m/s	0.64	0.02 1/s	0.64	0.002 m ² /s ³	0.296
2.5 m/s	1.60	0.04 1/s	1.28		
3.0 m/s	1.92				
5 m/s	3.20				

Table 4.3.: Overview of the model sensitivity in OGE. d denotes dependency, (d) weak dependency and i indirect dependency.

sensitivity of z^* to						sensitivity of Γ^* to					
model	v	u	ϵ	N	sh	model	v	u	ϵ	N	sh
TDP 2.1	d	-	i	d	d	TDP 2.1	(d)	-	d	d	(d)
APA 3.4	-	-	i	d	-	APA 3.4	-	-	d	d	-
APA 3.2	-	-	i	d	-	APA 3.2	-	-	d	d	-
D2P	-	(d)	i	d	-	D2P	-	(d)	d	d	-

The second row indicates that all models are sensitive to turbulence as the decay rates are increased. In the case of APA 3.4 and TDP this results in an increased vortex lifetime, although a circulation of $\Gamma^* = 0.4$ is reached earlier than under calm conditions. As the vortices in all models are weak or have decayed before they reach ground, no pronounced vortex-ground interaction can be observed. Thus, the lateral separation is constant, with exception of APA 3.2. In this model, vortex descent and circulation are coupled which is why vortex convergence caused by three-dimensional structures is used to achieve constant descent speed even after a catastrophic demise event occurs, such as Crow instability [127].

In the third row, stratification is increased and turbulence is set zero. This has the effect that the vortex lifetime is further reduced, even without the influence of turbulence. However, the onset of rapid decay of APA 3.4 is delayed compared to the previous case which is why the vortices reach ground and subsequently separate. Both D2P and TDP 2.1 incorporate the influence of stratification on vortex descent caused by buoyancy, that is both models predict vortex rebound after reaching the lowest point. This trend is continued when stratification is further increased (fourth row). For strong stratification, vortex lifetime and descent deviate substantially from the other models.

Table 4.4.: Overview of the model sensitivity in NGE and IGE.

sensitivity of z^* to						sensitivity of Γ^* to					
model	v	u	ϵ	N	sh	model	v	u	ϵ	N	sh
TDP 2.1	d	-	i	d	(d)	TDP 2.1	(d)	-	d	d	(d)
APA 3.4	(d)	-	i	d	(d)	APA 3.4	-	-	d	d	-
APA 3.2	(d)	-	i	d	(d)	APA 3.2	-	-	d	d	-
D2P	d	(d)	-	-	-	D2P	-	(d)	-	-	-

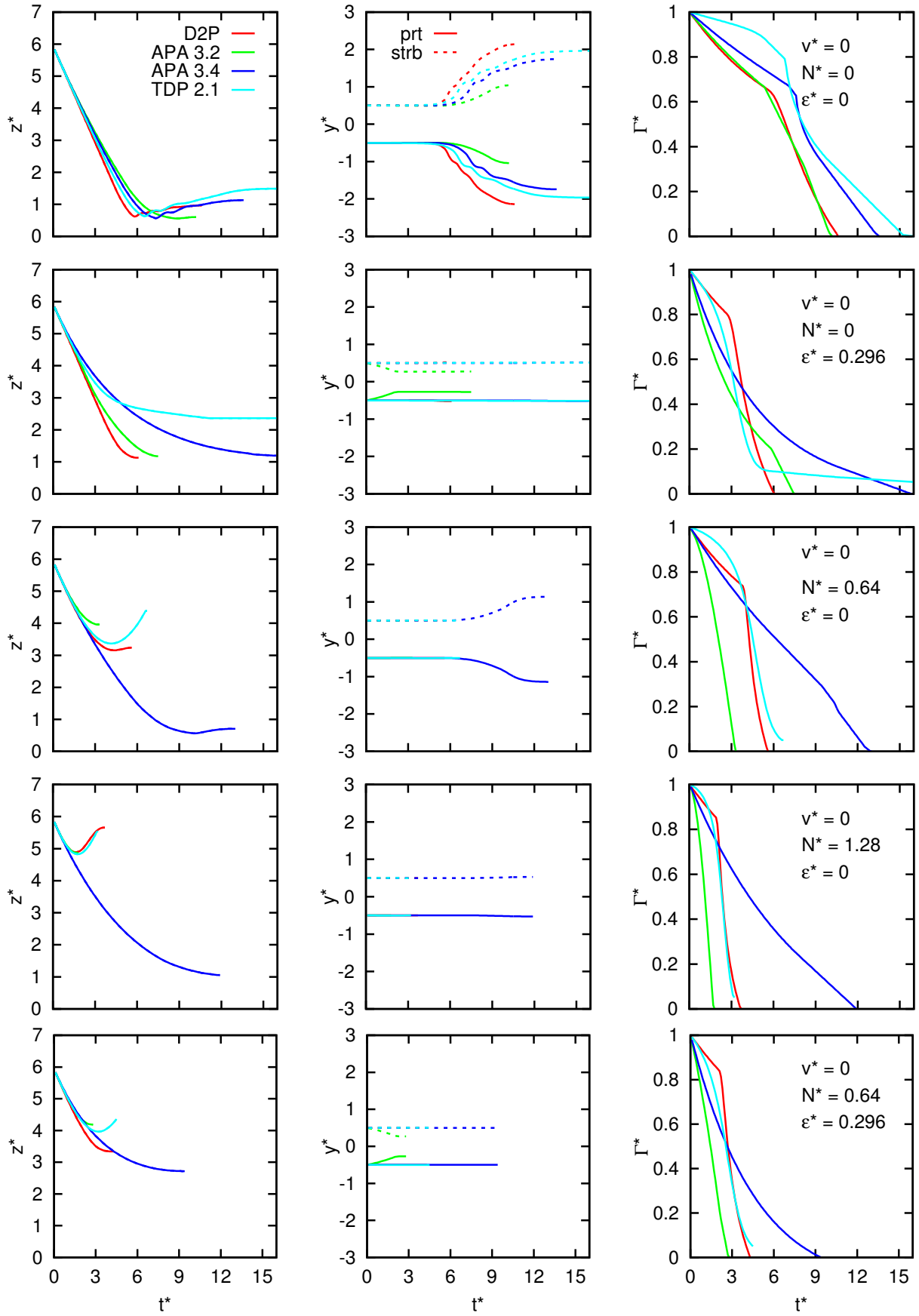


Figure 4.4.: Model sensitivity OGE.

The last row shows the forecasts for combined stratification and turbulence. Compared to the cases with either turbulence or stratification (second and third row) all models predict faster decay when turbulence and stratification are present at the same time. Above $1.0z_0$, the TDP model considers the impact of shear on vortex descent (see Figure 4.5). Positive shear (positive wind speeds provided) leads to a faster descent of the luff vortex. The other models show such asymmetry only after rebound.

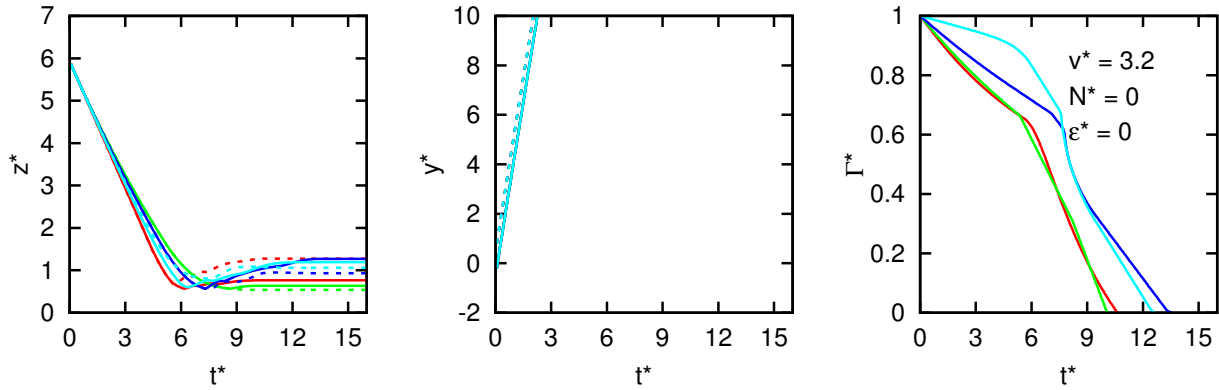


Figure 4.5.: Model sensitivity in OGE with crosswind.

Near ground (NGE phase), the model behavior is dominated by the vortex-ground interaction that is connected with crosswind and shear. For this reason the emphasis is put on the impact of crosswind strength and profile. Figure 4.6 indicates that rapid decay sets in faster than in OGE for all models. However, in contrast to the other models D2P is insensitive to stratification and turbulence in proximity to ground (compare first and second row). Holzäpfel and Steen [18] argued that the sensitivity of vortex decay to these effects is minor compared to the interaction with the ground. It is thus neglected in D2P, where decay is rather accelerated by decreasing initial altitudes. All models show no or almost no sensitivity to crosswind strength for vortex decay. In absence of crosswind the luff and lee vortices rise symmetrically in all models (second row). As soon as crosswind sets in, an asymmetry of vortex rebound can be observed which is increased with rising wind speed (fourth row). D2P assumes that this behavior, observed in measurements, is due to different strengths of secondary vortices of the luff and the lee vortex. With increasing crosswind the secondary luff vortex is weakened, whereas the secondary lee vortex is strengthened such that the rebound of the primary lee vortex is more pronounced [18, 90].

The common IGE formulation of both TDP and APA also allows asymmetric rebound, as shear may affect the distances of the secondary vortices for the up- and downwind vortex differently. However, the observed asymmetries are less distinctive for the APA models than for TDP. The analysis reveals, that the observed rebound of the NASA models is contrary to the D2P forecast. If shear is, however, inverted (see Figure 4.3) with unchanged wind direction, the lee vortex rises stronger than the luff vortex such that the trends conform with D2P, although ascribed to different effects. If shear is totally absent (constant wind profile) the rebound predicted by the NASA models is symmetrical, whereas D2P still predicts a stronger lee vortex rebound. The lateral separation is predicted very similarly by the NASA models. Differences can be attributed to different rebound characteristics. D2P, however,

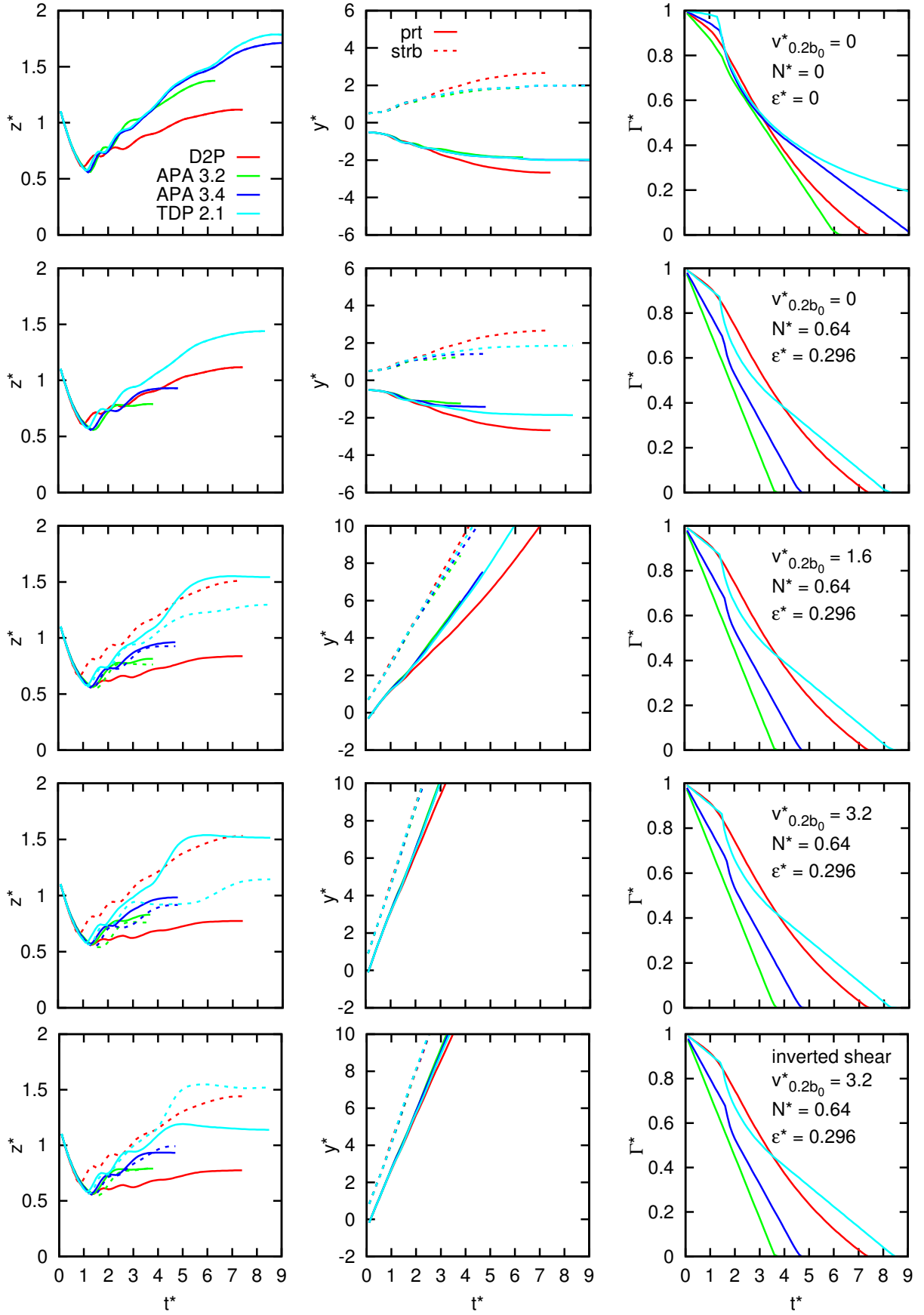


Figure 4.6.: Model sensitivity NGE.

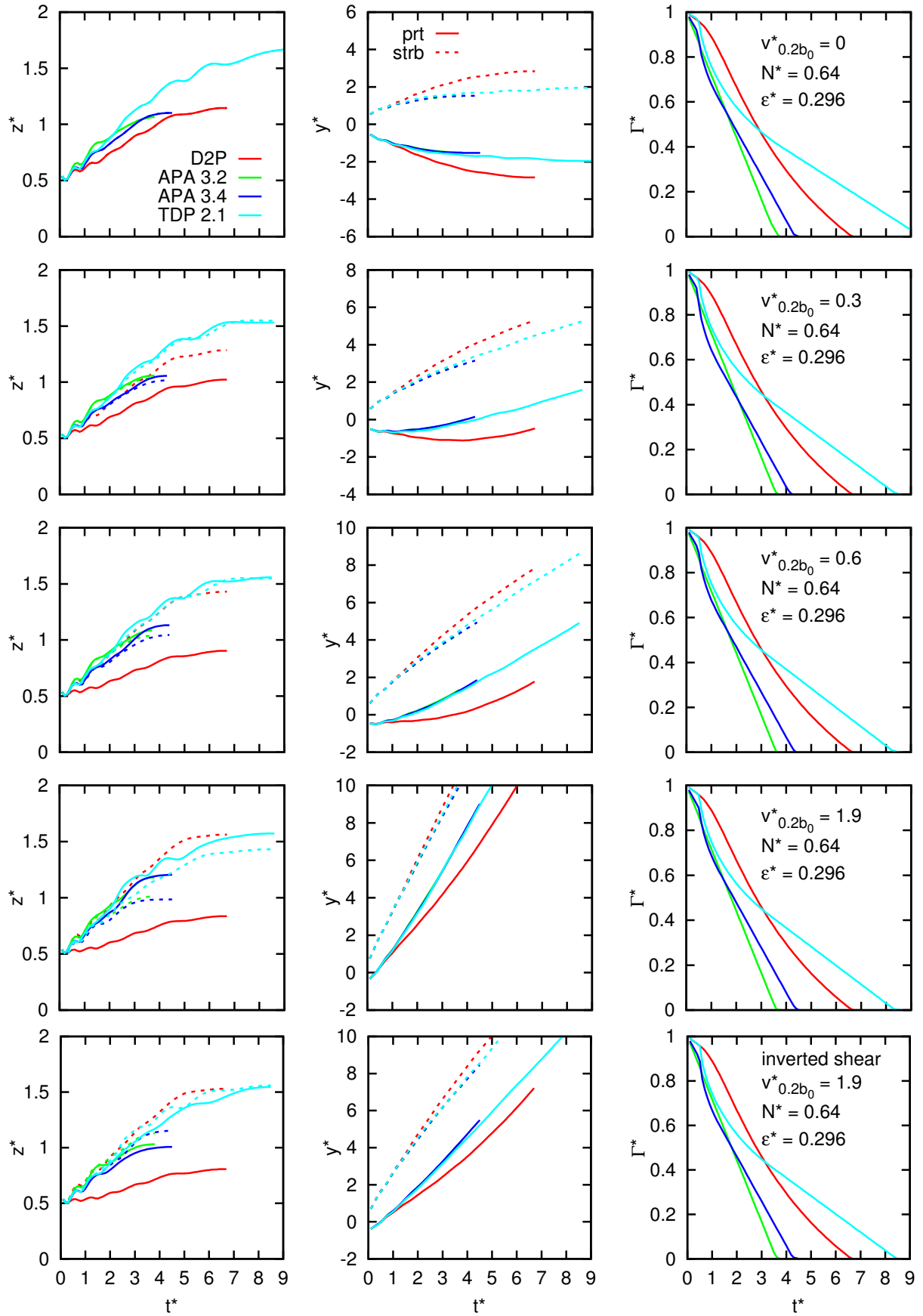


Figure 4.7.: Model sensitivity IGE.

shows a significantly stronger vortex divergence and furthermore a clear asymmetry in lateral vortex transport. This can be ascribed to the strong asymmetry in rebound that firstly leads to different lateral velocity induction by the image vortices and secondly exposes them to different wind speeds due to the wind profile.

Even closer to the ground (here $0.5b_0$) rapid decay starts before $1t_0$ in all models (see Figure 4.7). Increasing crosswind has the same effect on decay as in NGE. In any model, the vortices do not descend but start to rise immediately after they have been generated with no pronounced descent being visible. While the trend of D2P and the APA models concerning rebound is comparable with what has been observed in NGE, the asymmetry of TDP 2.1 is clearly less pronounced. Moreover, the NASA models show the same reaction to inverted shear as in NGE.

4.3.3 Performance

While the DLR dataset concentrates on landings in IGE and NGE the NASA dataset comprises for the most part OGE measurements. For this reason the model performance is evaluated separately for the WakeOP/ WakeFRA/ WakeMUC and the MEM95/ DFW97 dataset.

Error distributions

Although the RMSE and bias are valuable quantities to characterize the model skill they cannot fully describe their performance. Complementarily, the error distribution illustrates the distribution of the model deviations. On the basis of the WakeMUC/ WakeFRA/ WakeOP dataset, Figure 4.8 illustrates that the errors do not necessarily follow Gaussian but rather leptokurtic distributions for z^* and y_{luff}^* whereas the Γ^* forecast can be assumed to be normally distributed. Beside the kurtosis, the distributions are distinguished by their skewness which can be caused by a bias in the wind speed data (y_{luff}^*), the impenetrable ground that makes negative errors less likely (z_{luff}^*) or by a bias in the model physics as the model rebound or the lateral drift is under- or overestimated. Furthermore, the illustration reveals that the error distributions for the y_{lee}^* forecast nearly coincide, whereas D2P differs significantly from the other models for y_{luff}^* . Larger differences between the models are observed for z^* and Γ^* .

When evaluating OGE data by employing the NASA dataset (see Figure 4.9) significant differences are only revealed for the Γ^* forecast. For z^* they are very similar while they almost coincide for the y^* forecast. Reasons for the discrepancy between both datasets may be referred to different initial altitudes, meteorological conditions, orographics as well as measurement data quality. As an example, the less pronounced interaction with the ground, due to higher initial heights, results in less skewed and shifted distributions.

RMSE and bias

The RMSE deviations of the respective models are illustrated in Figure 4.10 for the DLR dataset and in Figure 4.11 for the NASA dataset. In comparison, the models perform on average better for the DLR datasets. Furthermore, the models are more similar in their performance for the NASA dataset with observations in OGE which is in line with the conclusions from the error distributions in the previous section. Figure 4.12 and Figure 4.13 depict the model bias,

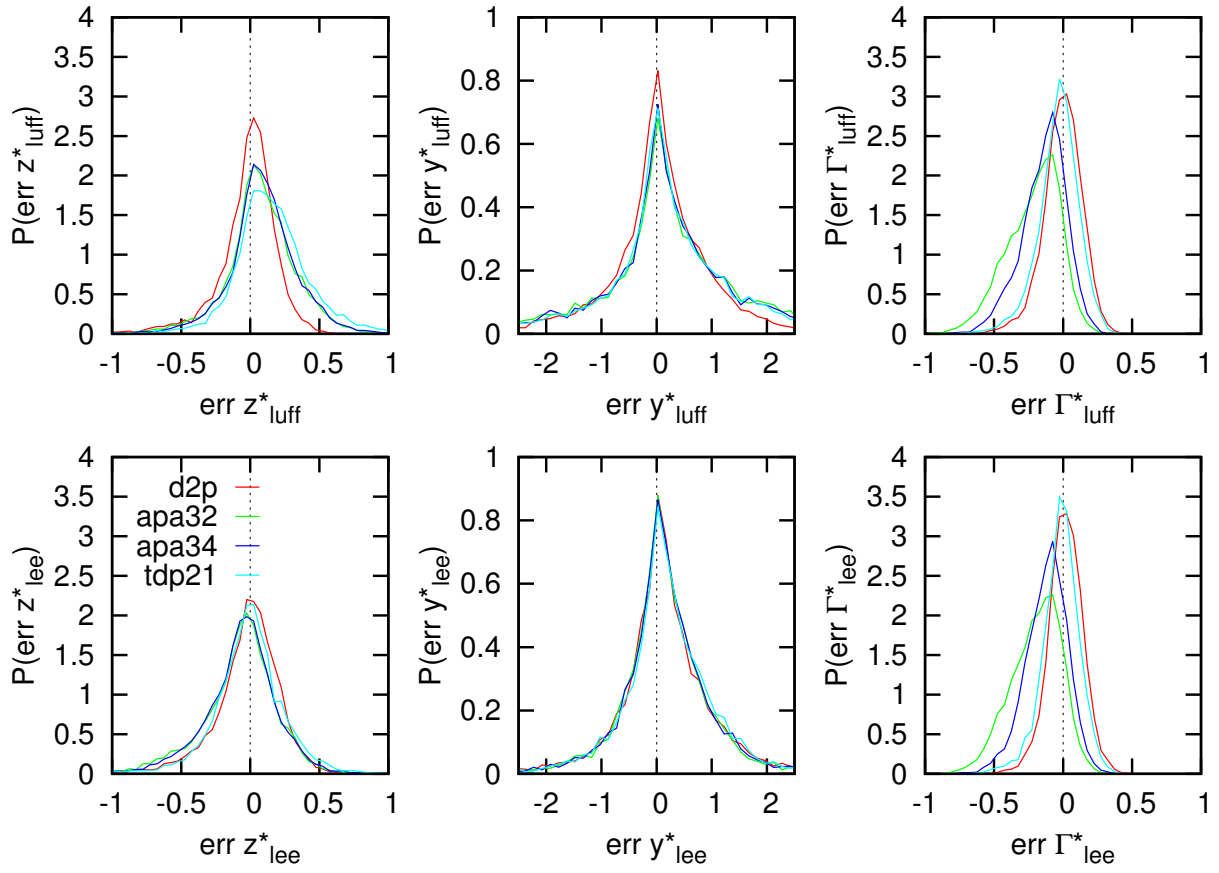


Figure 4.8.: Distribution of model deviations for the WakeOP, WakeFRA and WakeMUC campaign.

revealing that the sign of the error may vary significantly for different campaigns for the same reasons listed above. As a result it is hard to identify a single best model for both RMSE and bias as the performance varies significantly for different campaigns and initial heights.

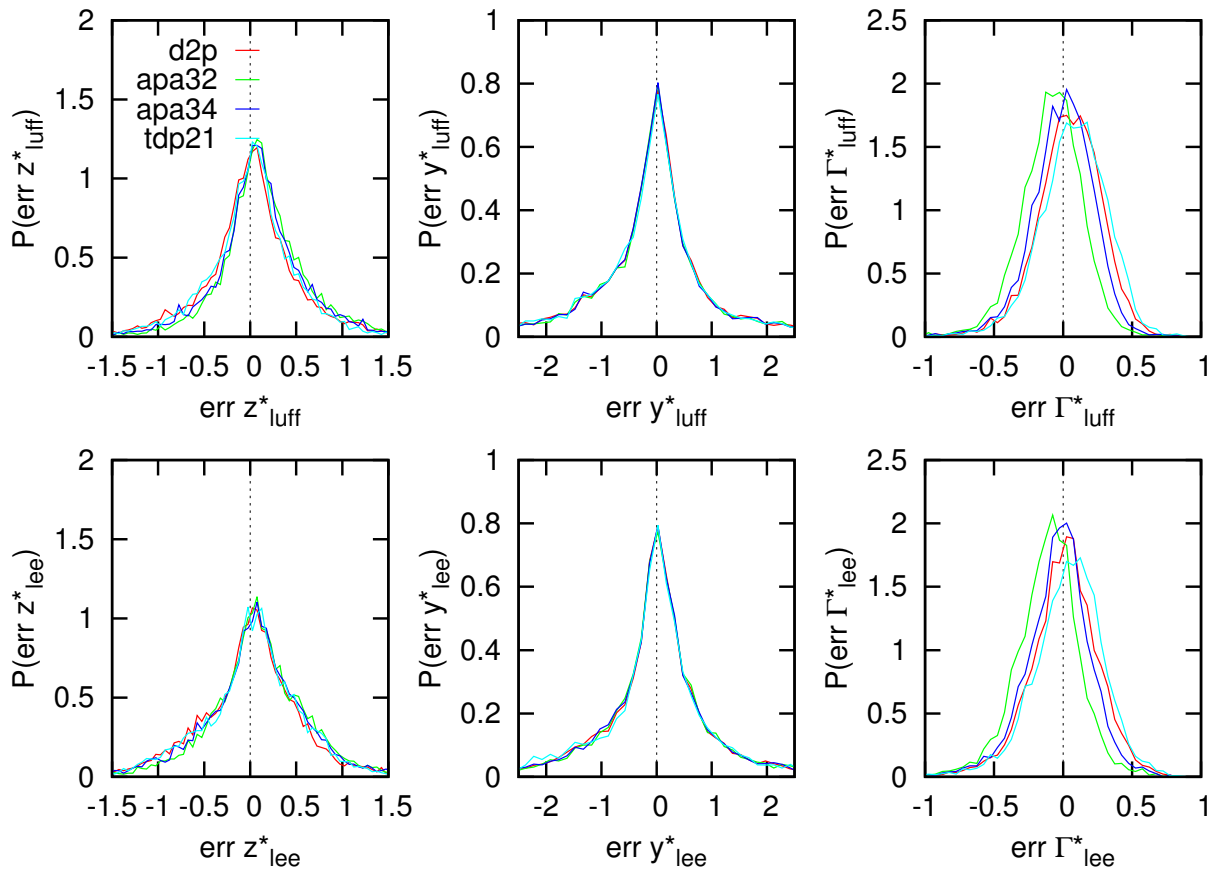


Figure 4.9.: Distribution of model deviations for 513 landings of the DFW97 and MEM95 campaign.

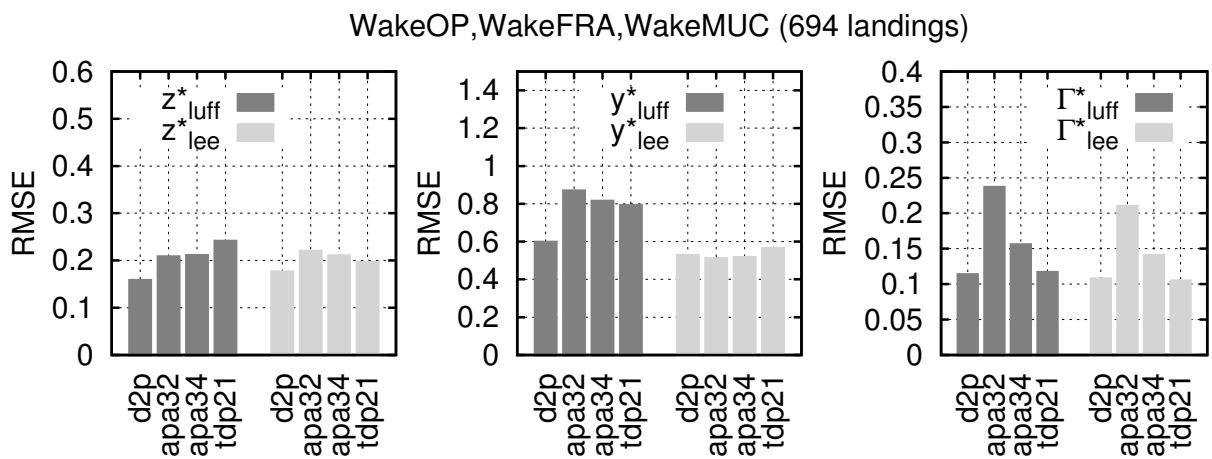


Figure 4.10.: Median model RMSE for the WakeMUC/ WakeFRA/ WakeOP dataset.

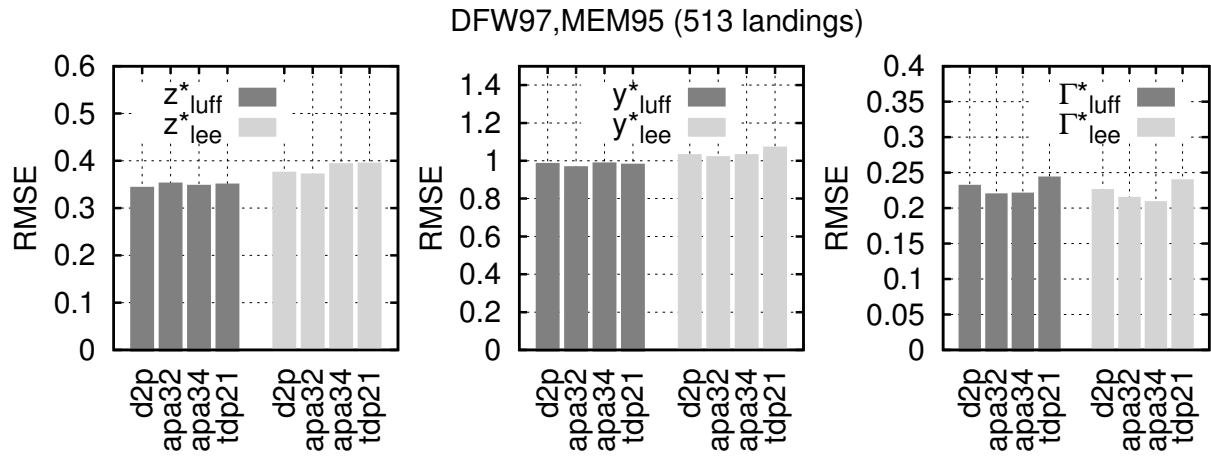


Figure 4.11.: Median model RMSE for the DFW97/ MEM95 dataset.

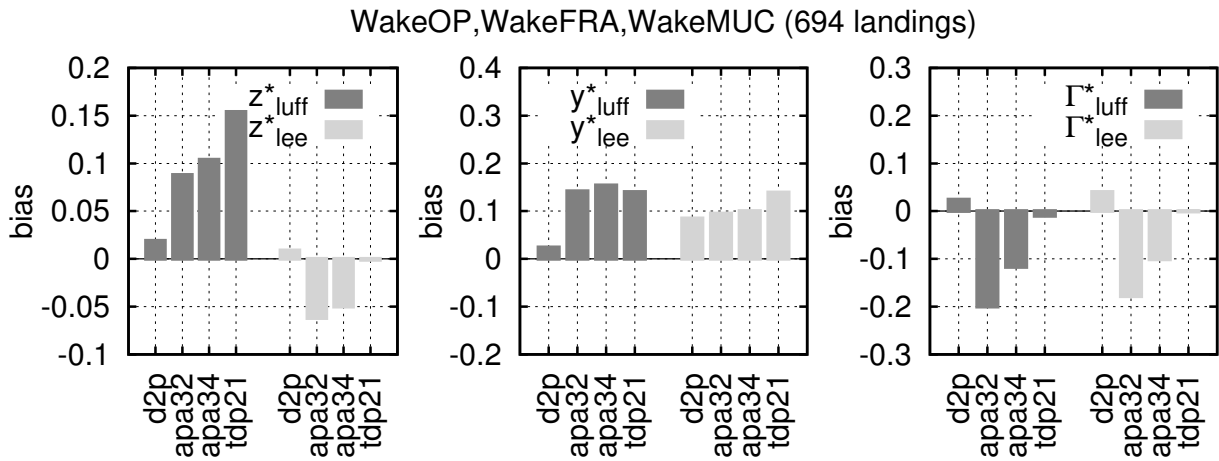


Figure 4.12.: Median model bias for the WakeMUC/ WakeFRA/ WakeOP dataset.

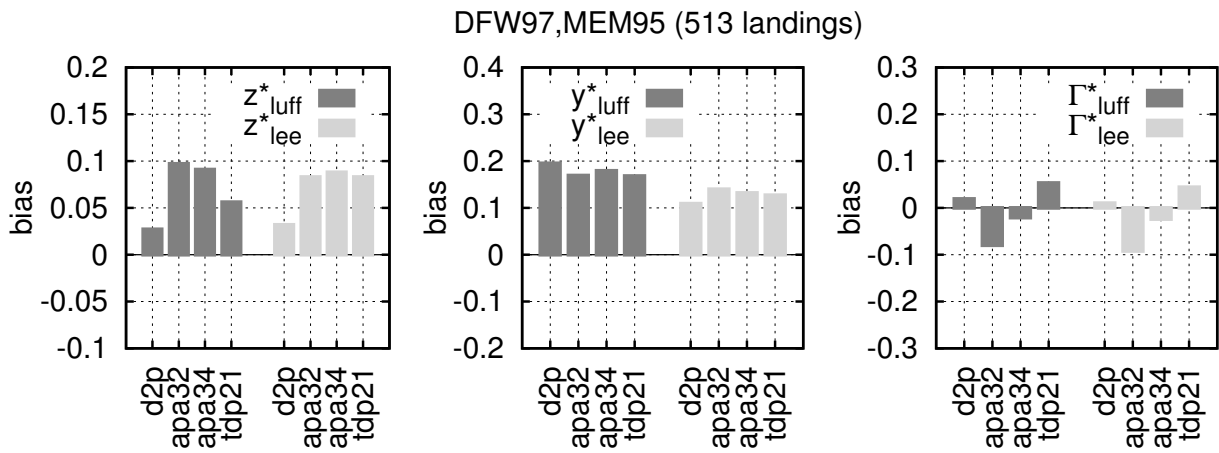


Figure 4.13.: Median model bias for the DFW97/ MEM95 dataset.

5. Multi-Model Ensemble Methods

Naturally, an ensemble forecast can exhibit better prediction skills than the worst-performing model among all members. However, the question how an ensemble can possibly outperform the best model is not straightforward and is thus discussed in this chapter. In the following, the ensemble models are also termed *members*. Several methods to set up a MME are considered, among them as most promising the Reliability Ensemble Averaging (REA) method [58] and the Bayesian Model Averaging (BMA) [71]. These two methods have in common that they weight the models on the basis of a training dataset in advance. As reference the Direct Ensemble Average (DEA) is employed. Eventually, Monte-Carlo Simulations and Lagged Average Forecasting, and the reasons why they were discarded are discussed. The methods described below only constitute the most promising in connection with wake vortex forecasts, although further variations of the Bayesian approach [131–134] and neural network methodologies [72] exist. Despite they might also achieve good results, they were discarded in favor of the DEA, REA and BMA approach.

5.1 Principles

Predictions are usually derived from models that either fully resolve or parametrize the relevant processes [68]. Errors in the projections may arise from manifold uncertainty sources which can be decomposed in the following groups.

Initial condition uncertainties Commonly the initial conditions are derived from measurements that exhibit uncertainties or they are based on assumptions of unknown accuracy. This may result in forecasts that diverge from the actual state right from the start. For example, a wake vortex forecast can be flawed by an inaccurately given aircraft mass or position. But also wind, deviating from the aircraft experienced conditions, may fall in this category.

Ambient condition uncertainties This includes measurement uncertainty caused by the instruments themselves and not only in the beginning but rather over the whole forecast horizon. For wake vortex forecasts an offset between the location of the instruments and the vortex generation may cause such uncertainty. Additionally, the temporal resolutions of the environmental condition measurements are limited. During the evolution of the simulated phenomena the ambient conditions may change, which may lead to large errors if the model forecast is not assimilated. In application to wake vortex forecasts this would be equivalent to the gusts that strongly deviate from the crosswind mean. Not being resolved they cannot be considered deterministically.

Model uncertainties They stem from the fact that not all physical aspects of the environmental conditions and vortex evolution are resolved in the models [135], either because their scale is too small or the physics is simplified or not yet fully understood. Furthermore, the uncertainty in the model parameters causes errors in the projections.

To address these uncertainties a variety of ensemble types have been developed. Various definitions and terms exist in literature but to clearly distinguish the different groups in this study the following convention is applied:

Initial Condition Ensembles (ICE) To allow for the initial condition uncertainty imposed by observational errors their influence on the simulation must be studied [65, 66]. By perturbing the initial conditions, for example by employing Monte Carlo Simulations, a variety of results can be achieved with one and the same model [67]. If the model behaves non-linearly, the ensemble mean may deviate from the unperturbed simulation. Moreover, a forecast PDF can be derived from which a probabilistic envelope can be generated.

Perturbed Physics Ensembles (PPE) The model parameters also exhibit uncertainties that can be accounted to the model uncertainty. By generating various ensemble members by parameter perturbations, these uncertainties can be considered [68].

Multi-Model Ensembles (MME) Employing multiple independent models, this ensemble type covers model uncertainty generated by simplifications or different and possibly incomplete assumptions about the underlying physics. The MME can be further extended by perturbing the initial conditions to allow for both the initial condition and model uncertainty. However, if not explicitly stated, the term Multi-Model Ensemble is used here for the combination of multiple models only.

In this study the focus lies on the MME. The reason is the increased run time of the ICE that collides with the time critical application of wake vortex forecasts. The PPE is discarded at this point as well, because it is assumed that the multiple available models should already cover model uncertainty sufficiently.

While the ICE has proven to be successful in general as it considers non-deterministic initial conditions, the reason to employ the MME can be found elsewhere. The key factor for its success lies in the varying performance of each of its members as a consequence of model uncertainty. Certainly, employing the best member instead of relying on the model consensus is tempting. However, Hagedorn et al. [63] argue that the identity of the best member varies under different conditions and is therefore difficult to identify. A Multi-Model Ensemble (MME) takes this fact into account by accepting that any model can be the best in a certain situation. An important consideration when discussing the superiority of the MME is that it may not facilitate large improvements compared to the single best model in any individual case. The study of Hagedorn et al. [63] rather points out the main advantage is the consistently better performance compared to the best member over a large range of conditions.

The superiority of the MME has been proven many times [63, 64, 136–138] and led to the ensemble approach becoming operational in met-offices all over the world. The simplest approach to combine several ensemble members is the direct ensemble average (also called

arithmetic ensemble mean), where no weights are employed. More sophisticated methods give higher weights to more skillful models and introduce further parameters that calibrate the ensemble prediction according to the environmental conditions. Here, the challenge is to estimate the skill of each model correctly a priori. To estimate the likelihood of the prediction, probabilistic ensemble forecasts utilize the ensemble spread.

5.1.1 Conditions for the MME success

Concerning the composition of the ensemble the study of Fritsch et al. [139] comes to the conclusion that the superiority of the ensemble approach is not to be owed to simple error cancellation. It can rather be attributed to the fact that models based on different physics and numerics play an important role in generating the full spectrum of possible solutions. This implies that a successful ensemble must be based on independent models. The potential of a MME to improve the forecast skill is believed to be high when the members are experts on different areas. In this context it is important to state that even poor-performing models can contribute to the ensemble skill [63]. However, models with consistent low performance compared to the other members will not add skill. For this reason it must be ensured that only models are chosen that outperform the other members at least sometime.

5.1.2 Qualification of the available wake vortex models

To find evidence that the available models fulfill the condition of independence, their correlation is evaluated. The dependency between two parameters f and r , evaluated for n cases, can be expressed by the correlation coefficient R introduced by Pearson [140] via

$$R = \frac{\text{Cov}(f; r)}{\sqrt{\text{Cov}(f; f)\text{Cov}(r; r)}} = \frac{\frac{1}{n} \sum_{i=1}^n (f_i - \bar{f})(r_i - \bar{r})}{\sigma_f \sigma_r}, \quad (5.1)$$

where σ_f denotes the standard deviation of parameter f and where \bar{f} describes its mean value. Here, f and r constitute the forecasts of two individual models. If the models are linearly correlated the values of R are equal to 1 while for uncorrelated models R would equal 0. To evaluate the level of independence of the models among each other R is calculated for a dataset comprising 694 landings from WakeOP, WakeFRA and WakeMUC. The results are depicted in Figure 5.1 and additionally listed in Table A.1-A.5 in the appendix. They demonstrate that the correlation among the models fluctuates between 0.900 and 0.986 for z^* , and between 0.886 and 0.953 for Γ^* . Concerning the prediction of the lateral vortex position y^* , higher correlations that vary between 0.976 and 0.998 occur. As APA 3.2 and APA 3.4 constitute two different versions of the same model, their correlations are the highest among all models and for all parameters. In contrast, the least correlation occurs between D2P and TDP 2.1 for z_{luf}^* , whilst it is D2P and APA 3.2 for z_{lee}^* . For y^* the correlation is the smallest between D2P and TDP 2.1. Eventually, the least linear correlation for the Γ^* -forecast can be found between D2P and APA 3.2. The OGE simulations of APA 3.4 and APA 3.2 are lower correlated than in ground proximity (see Figure 5.2). Nevertheless, on average the correlations are slightly higher compared to IGE and OGE.

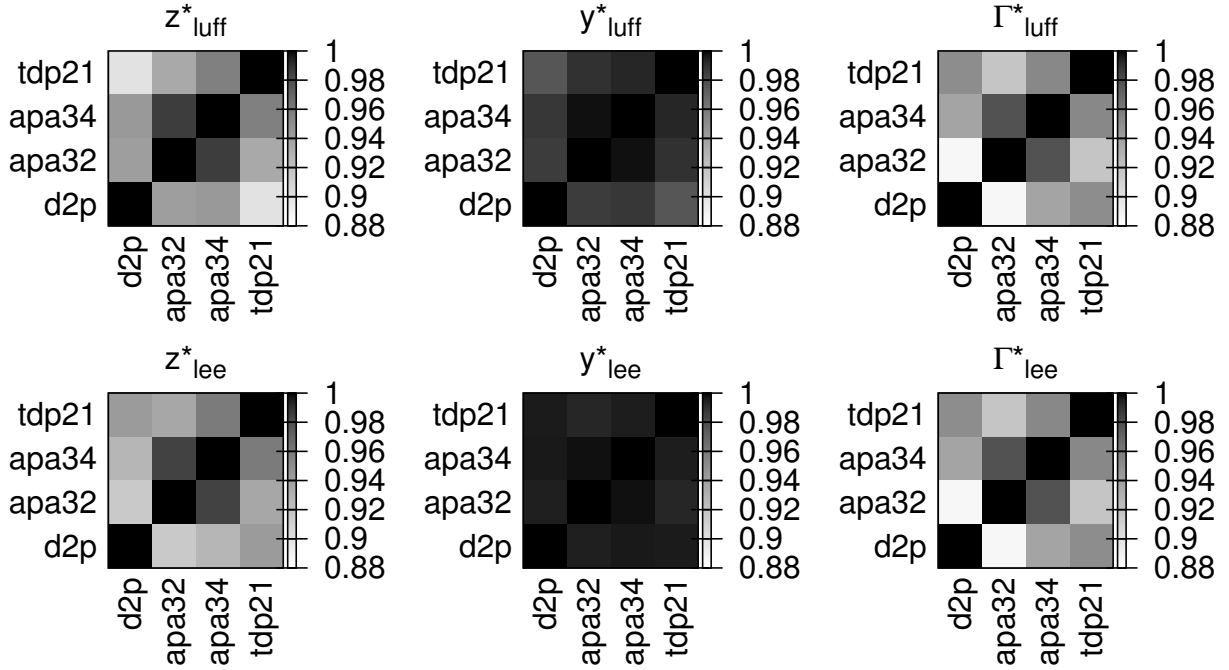
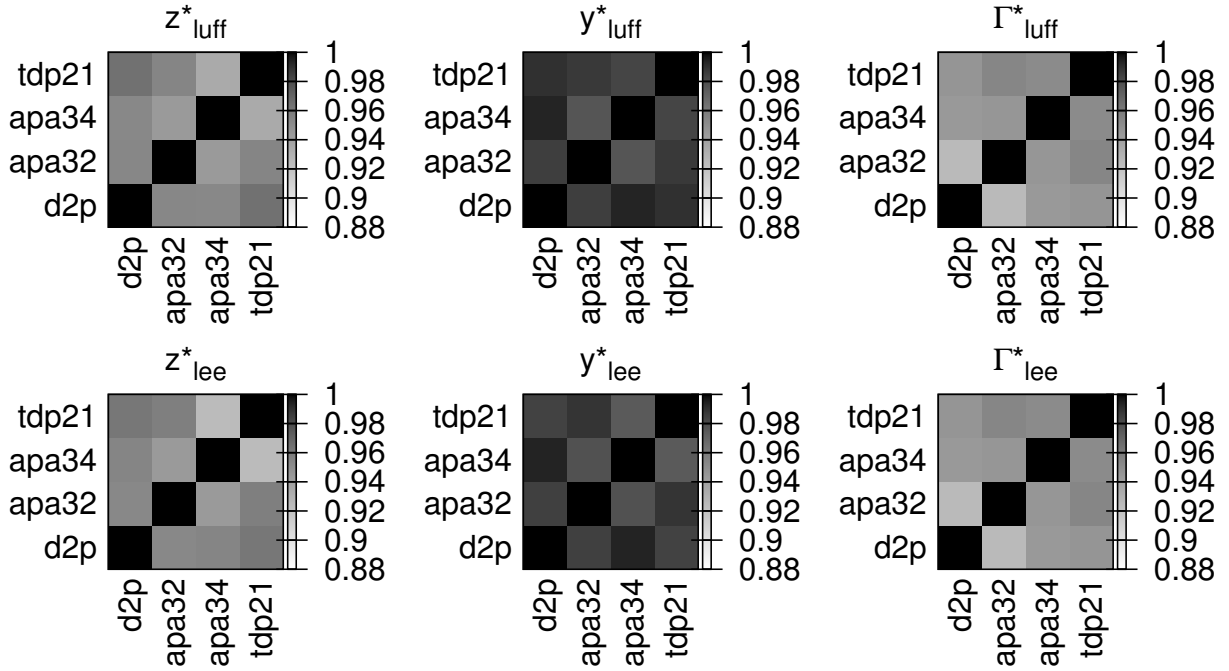
Figure 5.1.: Correlation coefficients R for the WakeMUC/ WakeFRA/ WakeOP sample.Figure 5.2.: Correlation coefficients R for the DFW97/MEM95 dataset.

Figure 5.3 displays the forecasts for the WakeMUC/ WakeFRA/ WakeOP dataset of the least correlated models against each other for the luff parameters, respectively. The left panel reveals the increased deviations between the D2P and TDP 2.1 forecast for z_{luff} in ground proximity, whereas the agreement is larger at higher altitudes. Note that the correlation is

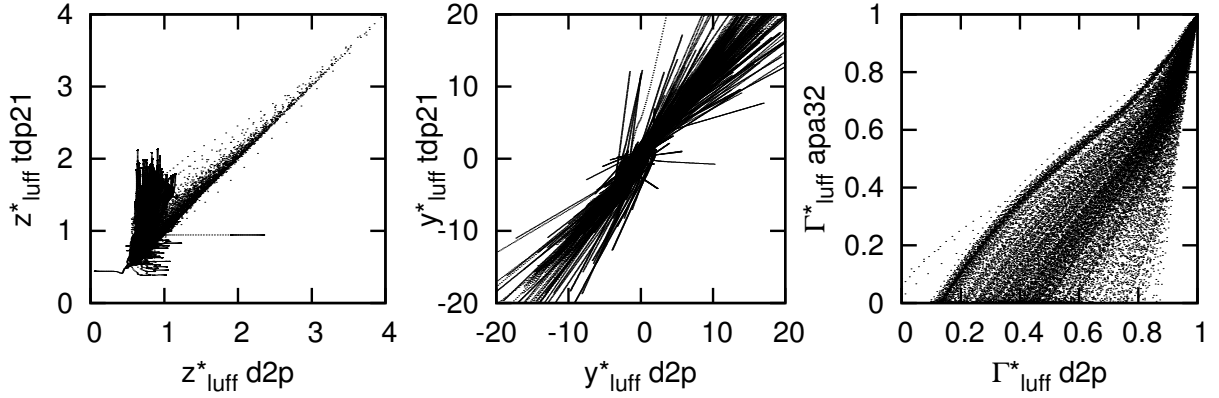


Figure 5.3.: Model forecasts against each other of the respectively least correlated models for the prediction of z_{luff}^* , y_{luff}^* and Γ^* for WakeOP, WakeFRA and WakeMUC (694 landings).

a measure of linear dependency and not related to the gradient of the line that would be achieved by linear regression.

These results suggest that structural model differences concerning the forecast of the vertical position and of decay exist, even if they are partly small. Especially at low altitudes and decreasing circulation the physics of the models seem to differ substantially. As far as the lateral vortex position is concerned the model correlation is higher. This implies that a forecast skill improvement is less likely for y^* than it is for z^* and Γ^* with respect to the findings of Fritsch et al. [139].

5.1.3 Spread-error correlation

Predicting the forecast skill and error distribution is desirable for two reasons. Firstly, the information whether a prediction is reliable or unreliable can support the decision making process. And secondly, on the basis of the forecast deviation distribution probabilistic envelopes can be generated. Therefore, the dispersion of the ensemble members, usually referred to as ensemble spread, can be exploited to indicate the forecast skill [141, 142]. Intuitively, large spread is associated with high uncertainty in the prediction. In this context Whitaker and Loughé [143] revealed that spread which is close to the climatological mean has little predictive value for an ensemble generated by initial condition perturbations. By contrast spread that departs from the climatological error distribution is more suitable as predictor of skill. They suggest that the correlation of spread and error should be related to the magnitude of spread variability which they confirmed for an operational ensemble. Note that the study only discusses ensemble members generated from perturbed initial conditions such that the source of spread is based on the initial condition uncertainty. In turn, in a MME the spread originates from model uncertainty. Giorgi and Mearns [58] nevertheless connect the spread of multiple models with the forecast reliability.

5.2 Best Member Selection

Hagedorn et al. [63] stated that the task to identify a single best model is complicated due to its changing identity. Nevertheless, it might be possible to identify certain conditions under which one model may be exposed as an expert. For this purpose the correlation between the meteorological conditions and model error is of particular interest in order to answer the question if it could be possible to always pick the best model depending on the ambient conditions (in the following termed *expert model*).

5.2.1 WakeMUC, WakeFRA, WakeOP

Figure 5.4 displays the RMS deviation of the z^* -forecast against z_0 and the meteorological conditions that have an impact on wake vortex descent for the WakeMUC/ WakeFRA/ WakeOP dataset.

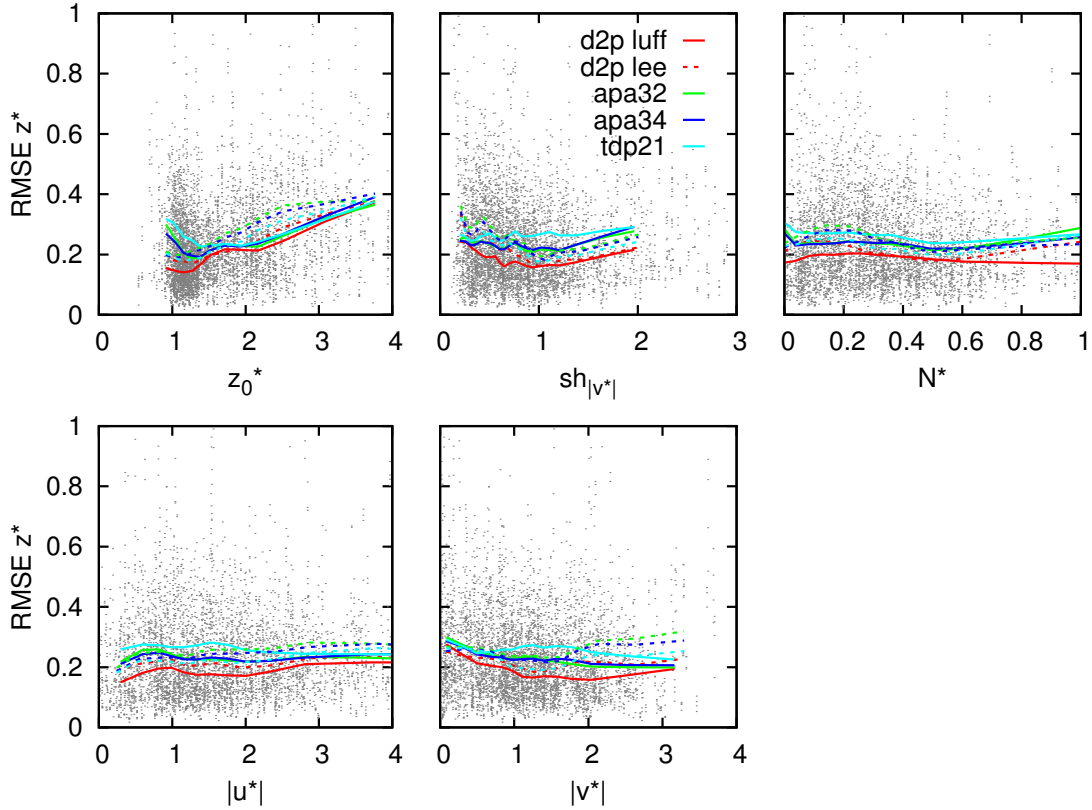


Figure 5.4.: Correlation of $RMSE\ z^*$ and the ambient conditions experienced by the vortices during their descent. The solid line corresponds to the luff and the dashed line to the lee vortex.

In all panels the average developments are mostly parallel, with the order of the models being the same. This indicates that on average it is on average always the same model that achieves the least deviations, independently from the ambient condition. Yet, D2P seems to perform even better for the luff vortex at low initial heights, as well as at increasing crosswind speeds

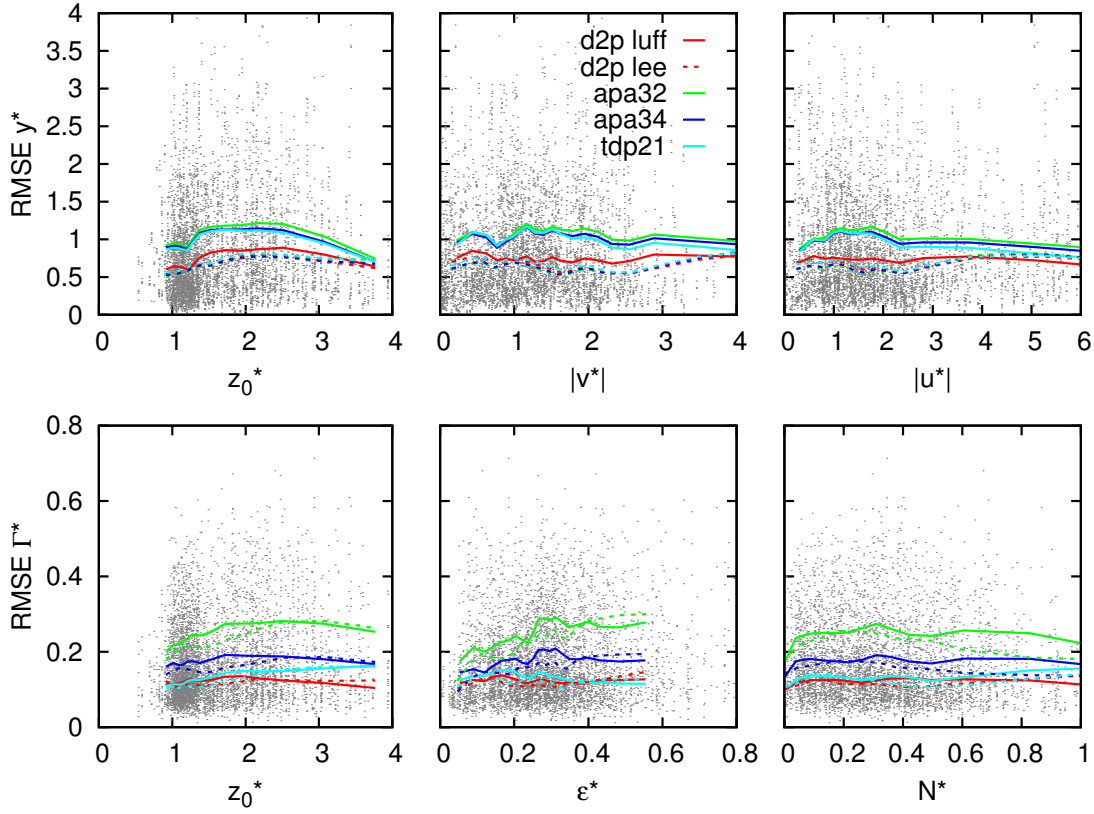


Figure 5.5.: Correlation of both $\text{RMSE } y^*$ and $\text{RMSE } \Gamma^*$ with the ambient conditions experienced by the vortices during their descent. The solid line corresponds to the luff and the dashed line to the lee vortex.

and shear. Although it would have been expected to identify TDP as expert model for shear, this could not be confirmed. To the contrary, it performs worst for the luff vortex in this dataset. This is also observed for the derivative of shear.

The first row of Figure 5.5 depicts the model RMS errors for the y^* -prediction against the most influential parameters, namely initial height, crosswind and headwind speed. It is again observed that D2P shows significant better performance for low initial heights and stronger crosswinds than the other models for the luff vortex. Moreover, the average RMS error demonstrates that the performance of all models is very similar for the lee vortex where the curves nearly intersect. In addition, the best performing member is constantly the same.

The second row of Figure 5.5 illustrates the RMS deviation of the Γ^* -forecast against initial altitude, turbulence level and stratification. Here, the curves of D2P and TDP coincide. This indicates that TDP would be the model of choice for low initial heights and strong turbulence, whereas it would be reversed for D2P. However, the differences in model performance are rather small between both models.

Conclusively, the findings point out that even if an expert model can be identified, it is the same model that performs best on average anyway. An exception is the Γ forecast, where the differences are, however, negligible. More meaningful results might be achieved by assessing multiple parameters at a time (e.g. strong crosswind with pronounced shear). However,

considering the size of the dataset this is hardly possible. At this point it is therefore not beneficial to choose one model over another in specific situations.

5.2.2 MEM95, DFW97

Figure 5.6 and Figure 5.7 depict the model RMSE against the meteorological conditions for the DFW97/ MEM95 dataset, where headwind is not provided in the input data. As in the evaluation of the WakeMUC/ WakeFRA/ WakeOP dataset no expert models can be identified. The fact that the trends are different in the DLR campaigns can be attributed to the distance between the measurement site of the LiDAR and the meteorologic instrumentation and the larger variability in the wake vortex measurements.

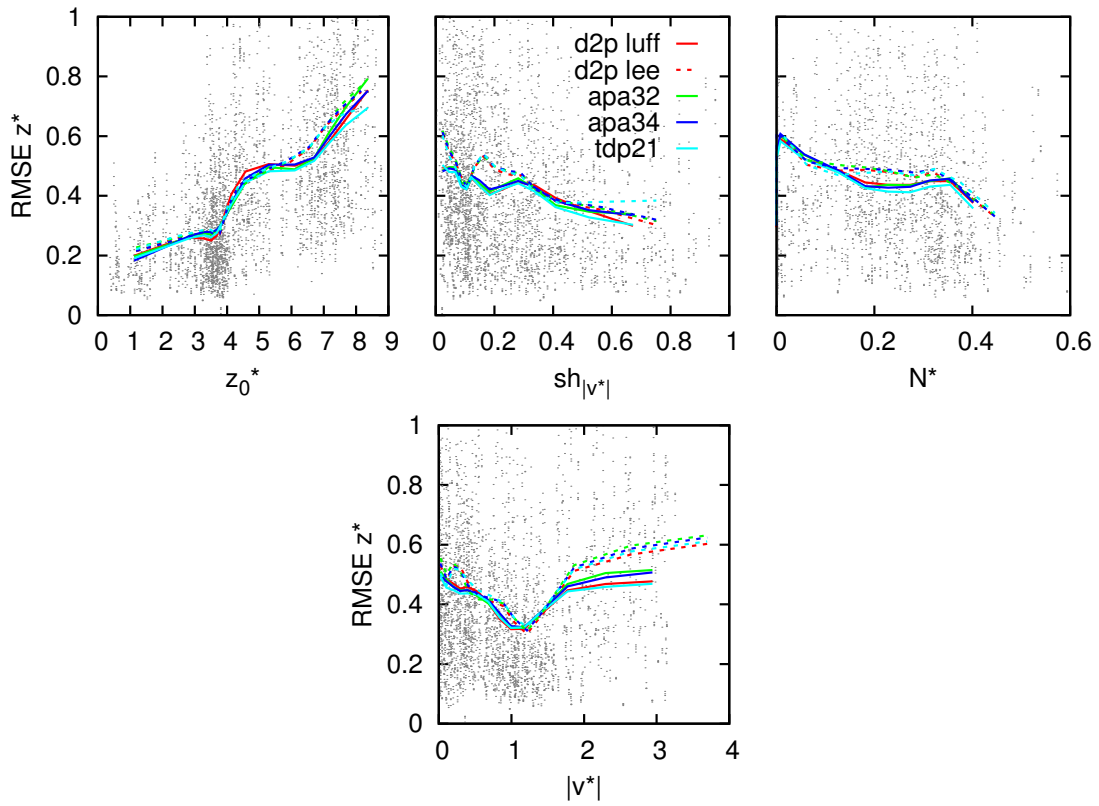


Figure 5.6.: Correlation of $RMSE z^*$ and the ambient conditions experienced by the vortices during their descent for the DFW97/ MEM95 campaign. The solid line corresponds to the luff and the dashed line to the lee vortex.

5.2.3 Reasons for absence of expert models

Given the sensitivity analysis in Chapter 4.3.2 this result is astonishing at first glance. However, the models being based on different physical equations does not guarantee their superiority under certain conditions. Nevertheless, in the first place at least TDP would have been expected to act as an expert in shear situations, especially in campaigns with higher vortex generation altitudes. Although the predicted behavior has been observed in laboratory experiments and LES

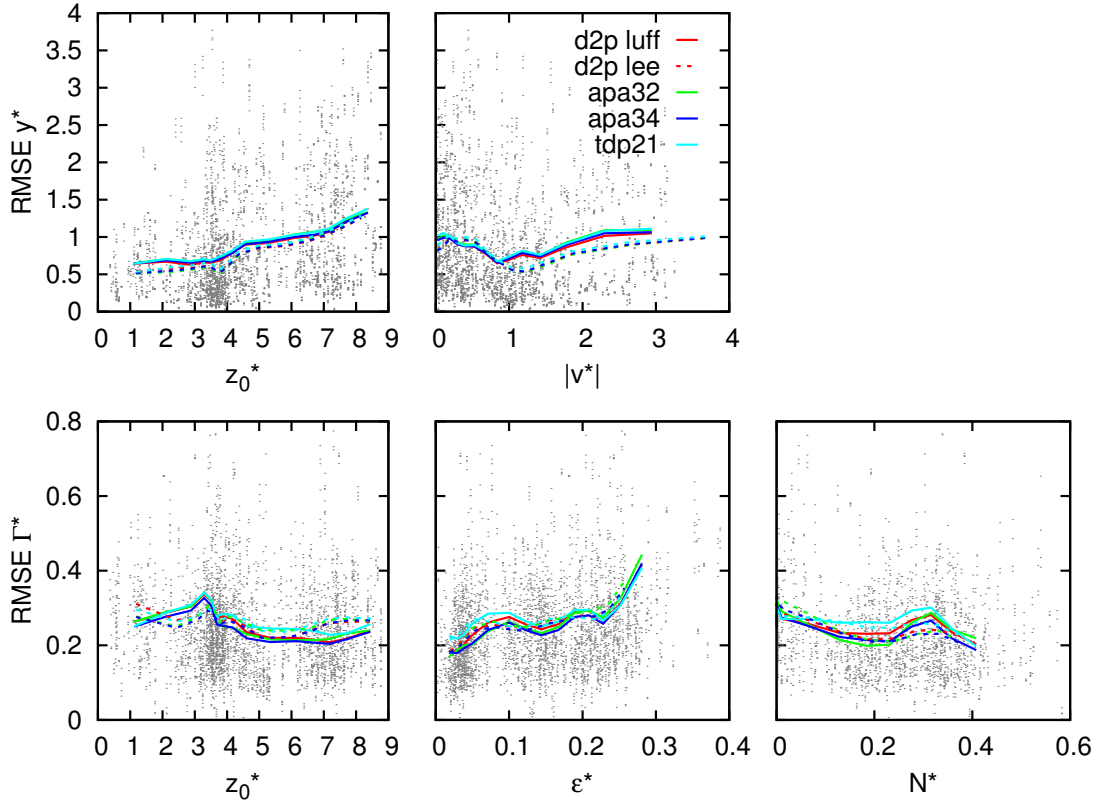


Figure 5.7.: Correlation of both $\text{RMSE } y^*$ and $\text{RMSE } \Gamma^*$ with the ambient conditions experienced by the vortices during their descent for the DFW97/ MEM95 campaign. The solid line corresponds to the luff and the dashed line to the lee vortex.

simulations [96, 98], it seems that the meteorological input from field measurements, usually averaged over 10 minutes, is not accurate enough to produce robust deterministic predictions in consideration of shear (see references [144] and [16]). Furthermore, the natural variability of the vortices and their interaction with the ground might dilute environmental impacts, such as observed by Holzäpfel and Steen [18] for turbulence induced decay in ground proximity. Finally, it cannot be ruled out that the model skill does not depend on single meteorological parameters only, but rather on the interplay of several. With regard to the limited size of measurements this cannot be proved for the available datasets.

Even smaller differences between the models are observed for the NASA dataset. This can be attributed to the imprecise input data due to different vortex generation and ambient condition measurement sites, as well as due to the impact of the trees beneath the glide slope. Inaccurate input data obscures the model performance differences.

5.3 Direct Ensemble Average (DEA)

The simplest method to combine several model forecasts is to average them without employing weights (see eqn.(5.2)). This neglects the strengths of individual ensemble members, but it can help to reduce the influence of outlying predictions. Here it is used as a baseline to compare the skills of the DEA with the other ensemble approaches.

$$\tilde{f} = \frac{\sum_{i=1}^I f_i}{n} \quad (5.2)$$

5.4 Reliability Ensemble Averaging (REA)

This method was developed by Giorgi and Mearns (2002) [58] and applied to atmosphere-ocean general circulation models to predict the climate change. In their approach a model is classified as reliable if both its bias B_i (performance criterion) and its distance D_i to the ensemble mean (divergence criterion) are within the scope of what they call natural variability. They found that this approach allows a reduction of the uncertainty range as the influence of poorly performing models is decreased compared to the DEA. The ensemble mean \tilde{f} can be expressed in a general form by

$$\tilde{f} = \frac{\sum_{i=1}^I R_i f_i}{\sum_i R_i}, \quad (5.3)$$

where R_i denotes the general reliability factor and f_i the individual forecasts for each model i . The reliability factor is defined as

$$R_i = [(R_{B,i})^m \cdot (R_{D,i})^n]^{1/(m \cdot n)} \quad (5.4)$$

and can be decomposed in two specific factors. $R_{B,i}$ is a factor that describes the model reliability as a function of its bias B_i (with $B_i = \text{forecast} - \text{observation}$) and $R_{D,i}$ characterizes the model reliability in terms of distance D_i between an individual model forecast and the REA average prediction (with $D_i = \text{forecast} - \text{ensemble mean}$). $R_{D,i}$ can thus be understood as a measure of the degree of convergence [58]. The natural variability nv represents variations of the observed parameters that cannot be predicted by the models (e.g. turbulent fluctuations). If B_i or D_i are less than the natural variability, the model is considered as reliable and it obtains the highest possible weight ($R_{D,i} = 1$ or $R_{B,i} = 1$ or both) because it cannot naturally perform better (see Equation (5.5) and (5.14)).

$$R_{B,i} = \min \left(\frac{nv}{|B_i|}, 1 \right) \quad (5.5)$$

$$R_{D,i} = \min \left(\frac{nv}{|D_i|}, 1 \right) \quad (5.6)$$

The distance $D_{T,i}$ is calculated iteratively with the direct ensemble average as first guess. The exponents n and m in Equation 5.4 weight the influence of $R_{D,i}$ and $R_{B,i}$ on the ensemble

prediction. $R_{B,i}$ is constant for all vortex predictions and determined a priori from a test dataset, whereas $R_{D,i}$ varies depending on the model spread in each individual prediction. In order to estimate the uncertainty of the ensemble forecast Giorgi and Mearns [58] introduce $\tilde{\delta}_f$. It quantifies the weighted average of the model deviations from the ensemble mean:

$$\tilde{\delta}_f = \sqrt{\frac{\sum_{i=1}^I R_i (f_i - \tilde{f})^2}{\sum_{i=1}^I R_i}}. \quad (5.7)$$

That is, the larger the ensemble spread, the larger the uncertainty. We obtain upper and lower uncertainty limits by adding and subtracting $\tilde{\delta}_f$ from the ensemble mean.

$$f_{\pm} = \tilde{f} \pm \tilde{\delta}_f \quad (5.8)$$

Note that the probabilistic envelopes, however, do not naturally provide uncertainty levels based on previous observations. Nevertheless, probability levels can be estimated under the assumption that the forecasts are distributed following a Gaussian PDF [58]. Then $\tilde{\delta}_f$ can be viewed as a standard deviation and accordingly the envelopes represent a 68.3% confidence interval.

5.4.1 Natural variability

In this study the natural variability is treated as a model resolution limit. It describes the range for which a model forecast can be considered reliable. In our application the model resolution is limited by [145]:

- stochastically decaying vortices (ground linking, Crow instability, vortex bursting)
- turbulent deformation
- variability of environmental conditions
- uncertainty of initial conditions

To identify the impact of stochastically decaying vortices and the effect of large scale turbulence on vortex variability, wake vortex measurement data is used. The variations that are found in measurements under essentially the same ambient conditions can be attributed to the wake vortex variability and are quantified by the deviations of the observations from the measurement derived mean. The observed variability is given by the standard deviation σ_{obs} for z^* and Γ^* using non-dimensional data. As the LiDAR measurements are afflicted with errors, the observed variability σ_{obs} contains both the measurement error σ_{err} and the actual wake vortex variability σ_{nv} [145]:

$$\sigma_{obs} = \sqrt{\sigma_{err}^2 + \sigma_{nv}^2} \quad (5.9)$$

For the LiDAR measurement errors the findings of Köpp et al. (2005) [117] serve as a reference. They used simultaneous measurements with different LiDAR systems of the same vortex pair to derive standard deviations:

- $\sigma_{err,z} = 9 \text{ m}$
- $\sigma_{err,y} = 13 \text{ m}$
- $\sigma_{err,\Gamma} = 13 \text{ m}^2/\text{s}$.

However, finding a significant number of cases featuring identical ambient conditions is not possible for the available dataset. Therefore, the wake behavior is clustered in bins, which represent intervals of various initial or ambient conditions. Either crosswind (for σ_{obs,z^*}) or turbulence bins (for σ_{obs,Γ^*}) are employed to compute the observed variability within given bins of vortex generation altitudes. Stratification is ignored here because its impact is minor in ground proximity compared to the vortex ground interaction [146]. Certainly, the natural vortex variability also increases with time. Given the size of the dataset this is ignored here.

In order to compute σ_{obs,z^*} and σ_{obs,Γ^*} we utilize 1252 landings of the WakeMUC, WakeFRA and WakeOP campaigns. Figure 5.8 displays the measured wake vortex variability for z^* , y^* and Γ^* for specific crosswind and turbulence bins and model runs for the upper and lower bin limit. The model derived sensitivities (here from D2P as a reference) to the chosen bin widths are small for z^* and Γ^* which implies that they can be neglected and that the bins are small enough for this study. However, this does not hold for y^* which is why nv_{y^*} must be derived differently [145].

The standard deviations listed in Table 5.1 indicate that larger variability can be observed for larger z_0^* where wake vortex behavior is less restricted by the ground (see Holzäpfel and Steen [18]). Thus σ_{obs,z^*} is calculated with the fits depicted in Figure 5.9. No clear correlation between crosswind strength and observed variability can be found for the considered bins. Furthermore the results reveal that the variability of the lee vortex is higher compared to the luff vortex [145].

Holzäpfel and Steen [18] show that there is a correlation between turbulence and vortex decay in ground proximity. Nevertheless, they state that its significance is minor. This also applies to the respective observed variability that shows to be insensitive to the width of the turbulence bins that are employed. Table 5.2 indicates that the observed variability of vortex strength during decay is neither sensitive to turbulence nor to the initial height z_0^* .

Tables 5.1 and 5.2 also indicate that σ_{obs,z^*} can be on the order of the non-dimensional measurement error, normalized by b_0 . This suggests that either the measurement errors determined by Köpp et al. [117] are too conservative or that there is no vertical vortex variability. However, LES indicate that wake vortices can exhibit significant intrinsic variability [79] which implies that the estimated LiDAR error is too conservative. Thus nv_{z^*} and nv_{Γ^*} are calculated by [145]:

$$nv = \begin{cases} \sqrt{\sigma_{obs}^2 - \sigma_{err}^2} & \text{if } \sigma_{obs} > \sigma_{err} \\ \sigma_{obs} & \text{otherwise} \end{cases} \quad (5.10)$$

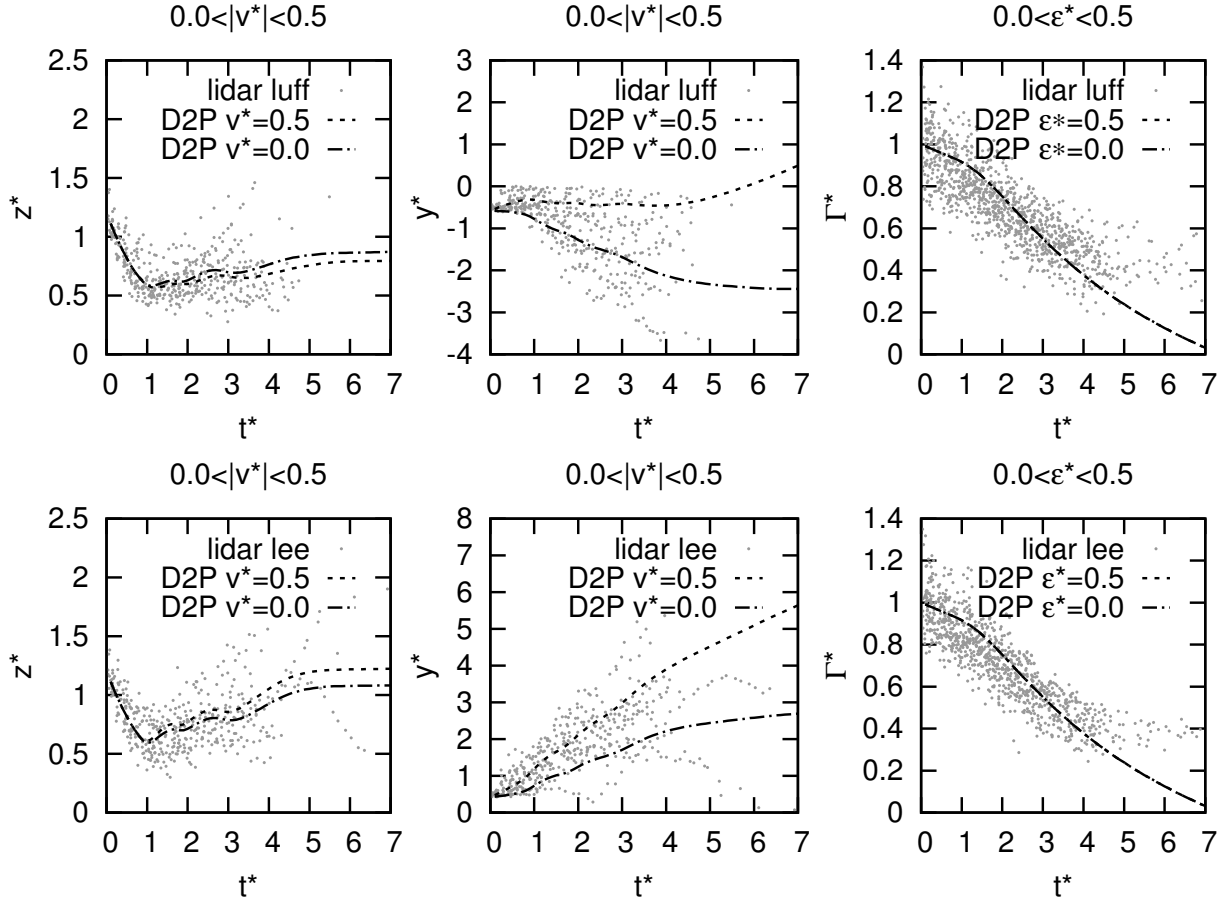


Figure 5.8.: Wake vortex variability for $1.0 < z_0^* < 1.2$ for luff (above) and lee (below) [145].

The above described method to derive nv is applicable to z^* and Γ^* . However, the sensitivity of the lateral position to crosswind is high (see Figure 5.8), which implies that the crosswind bins must be smaller to derive the natural variability. As the number of observations in our dataset is limited, this method is not applicable to nv_{y^*} [145]. Instead of employing LiDAR data, nv_{y^*} is calculated from the turbulence velocity q , which is a measure of ambient turbulence [121].

$$q = \sqrt{u_x'^2 + u_y'^2 + u_z'^2} \quad (5.11)$$

Furthermore, it is assumed that q can be considered as a superimposed propagation velocity [15]. As a result, the standard deviation of the vortex displacement can be determined in dependency of vortex age [145]:

$$nv_y^* = \sigma_{nv_{y^*}} = \frac{1}{\sqrt{3}} \cdot q^* \cdot t^*. \quad (5.12)$$

However, the assumption made in Equation (5.12) yields a natural variability that is mostly larger than both B_i and D_i . Although the physics of the models do not differ much for the y -prediction there are differences that can be assigned to different vortex lifetimes and descend

Table 5.1.: Observed variability for z^* , with number of employed data points in brackets that are used for the weighted mean [145].

z_0^*	$0.0 < v^* < 0.5$		$0.5 < v^* < 1.0$		$1.0 < v^* < 1.5$	
	σ_{obs,z_{luff}^*}	σ_{obs,z_{lee}^*}	σ_{obs,z_{luff}^*}	σ_{obs,z_{lee}^*}	σ_{obs,z_{luff}^*}	σ_{obs,z_{lee}^*}
0.8 - 1.0	0.194 (251)	0.311 (305)	0.228 (207)	0.267 (200)	0.168 (196)	0.197 (208)
1.0 - 1.2	0.172 (428)	0.196 (427)	0.181 (299)	0.252 (290)	0.194 (380)	0.197 (330)
1.2 - 1.4	0.172 (207)	0.267 (210)	0.183 (296)	0.189 (252)	0.180 (199)	0.195 (194)
1.4 - 1.6	0.240 (118)	0.223 (122)	0.240 (234)	0.225 (210)	0.197 (195)	0.207 (176)
1.6 - 1.8	0.194 (126)	0.304 (161)	0.201 (131)	0.239 (120)	0.208 (149)	0.226 (114)
1.8 - 2.0	0.246 (77)	0.127 (73)	0.277 (186)	0.211 (171)	0.217 (210)	0.324 (182)
2.0 - 2.2	0.272 (83)	0.327 (84)	0.200 (127)	0.196 (121)	0.226 (181)	0.217 (161)
2.2 - 2.4	0.280 (42)	0.325 (46)	0.206 (114)	0.343 (94)	0.248 (177)	0.269 (155)
mean	0.198	0.254	0.230	0.235	0.202	0.225

Table 5.2.: Observed variability for Γ^* , with number of employed data points in brackets that are used for the weighted mean [145] .

z_0^*	$0.0 < \epsilon^* < 0.5$		$0.5 < \epsilon^* < 1.0$	
	$\sigma_{obs,\Gamma_{luff}^*}$	$\sigma_{obs,\Gamma_{lee}^*}$	$\sigma_{obs,\Gamma_{luff}^*}$	$\sigma_{obs,\Gamma_{lee}^*}$
0.8 - 1.0	0.086 (162)	0.113 (167)	0.097 (88)	0.097 (83)
1.0 - 1.2	0.108 (791)	0.101 (627)	0.082 (121)	0.100 (129)
1.2 - 1.4	0.119 (412)	0.112 (387)	0.116 (124)	0.108 (122)
1.4 - 1.6	0.101 (313)	0.113 (319)	0.090 (111)	0.094 (115)
1.6 - 1.8	0.111 (255)	0.094 (278)	0.100 (77)	0.086 (78)
1.8 - 2.0	0.095 (366)	0.106 (359)	0.080 (95)	0.097 (88)
2.0 - 2.2	0.101 (305)	0.103 (311)	0.098 (79)	0.091 (94)
2.2 - 2.4	0.096 (260)	0.095 (256)	0.082 (38)	0.111 (48)
mean	0.104	0.104	0.094	0.098

trajectories. By weighting all models equally ($R_i = 1$) these small differences are neglected. It shows that decreasing nv_{y^*} by 4/5 clearly increases the ensemble skill of the y -prediction. Although using standard deviations for the natural variability is successful for the z - and Γ -forecast, it fails for the prediction of y . The prediction skill of the lateral position can only be increased by the REA ensemble approach if we reduce nv_{y^*} to 1/5th (see Equation (5.13)) of the value we would expect based on physical considerations (see Equation (5.12)) [145].

$$nv_y^* = \frac{1}{\sqrt{3}} \cdot q^* \cdot t^* \cdot \frac{1}{5} \quad (5.13)$$

It is found that computing both reliability factors in dependency of vortex age can be beneficial for the ensemble performance. Therefore, $R_{B,i}$ is computed vortex-age binned before being interpolated, while $R_{D,i}$ is derived for each time step before being smoothed. The weights achieved if nv is derived in this manner for a single landing are illustrated in Figure 5.10. The low R_B factors for the y -forecast can be referred to the fact that nv_{y^*} is

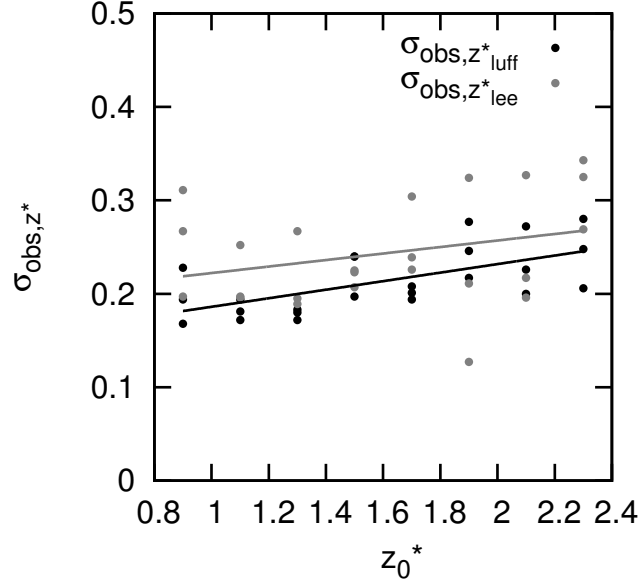


Figure 5.9.: z_0^* dependency of σ_{nv} for the examined vortex generation height and crosswind bins in Table 5.1 [145].

small due to the factor $1/5$ (see Equation (5.13)). The weights of APA 3.2 and APA 3.4 are very similar for the z - and y -forecasts, as the calculation of the reliability factors does not take into account that they are based on the same equations with different parameters. For this reason the reliability factors of APA 3.2 are artificially set 0 for z and y . It is found that this improves the ensemble performance considerably [145].

5.4.2 Weighting of R_B and R_D

The bias does not differ drastically between the models, such that the differences of $R_{B,i}$ are rather small. However, it turns out that a non-linear weighting that promotes models with small biases stronger has positive impact on the ensemble performance. One can achieve this by employing exponents n and m that are less than 1. For a better understanding of the weighting of R_B and R_D , Equation (5.4) can be put differently according to

$$R_i = [(R_{B,i})^{1/n} \cdot (R_{D,i})^{1/m}]. \quad (5.14)$$

A large weighting exponent m leads to weak weighting of $R_{D,i}$ as $1/m$ converges towards 0. By contrast, if m is smaller than 1 the exponent $1/m$ takes values larger than 1 and thus introduces non-linearity. Of course, the same holds for n . It turns out that the best results are achieved for $n = 0.3$ and $m = 1$ [145].

5.4.3 Potential

The REA is in comparison to the DEA more sophisticated as it assesses the model performance a-priori. For this reason the improvements achieved by the ensemble are expected to be larger. Moreover, this approach is chosen due to its capability to provide probabilistic envelopes that

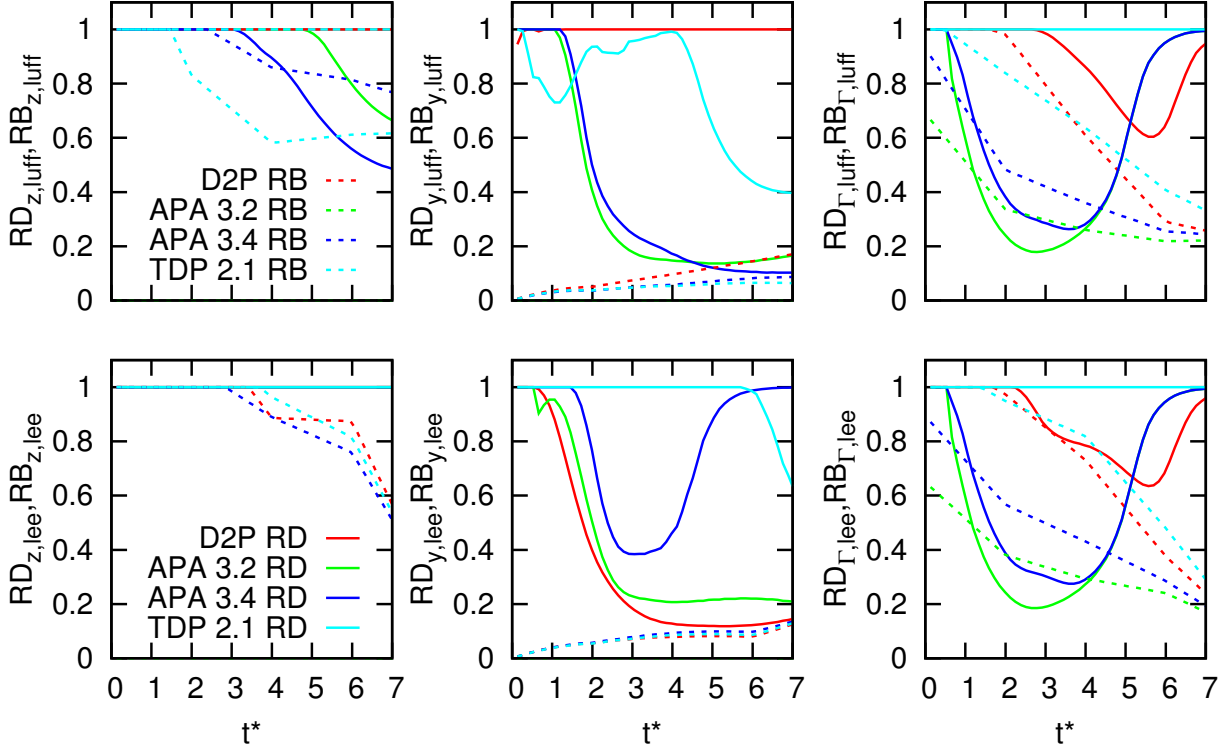


Figure 5.10.: REA weights for a single landing from the WakeMUC/ WakeFRA/ WakeOP sample. While $R_{B,i}$ does not change after being calculated in the training, $R_{D,i}$ varies between single landings.

are based on model spread. At the same time the natural variability ensures that only models with self-imposed poor skill are degraded, leaving measurement errors and wake vortex behavior that cannot be modeled yet disregarded.

5.5 Bayesian Model Averaging (BMA)

Model forecasts are connected with uncertainties with respect to initial and environmental conditions, stochastic vortex behavior and model physics. Further uncertainty is ascribed to the assessment of the model performance against measurement data which itself is subject to measurement uncertainties. Calculating uncertainty allowances is one way to handle them. Another way is to include them a priori, which constitutes the essence of the Bayesian Model Averaging [71]. This approach quantifies the prediction uncertainties by dressing each model forecast with a Probability Density Function (PDF). An ensemble forecast can then be generated by weighting the PDFs according to the respective model performance and summing them up.

This can be formulated by the law of total probability. It describes the probability that event B occurs in dependency of the likelihood of event A_i , given that B resides in the same probability space. $P(B)$ is the probability that B occurs, $P(A_i)$ denotes the probability that A_i occurs and $P(B|A_i)$ describes the probability that B occurs given A_i .

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(A_i)P(B|A_i) \quad (5.15)$$

Raftery et al. [71] apply this to an ensemble forecast. In Equation (5.16) y denotes the quantity to be forecasted and y^T denotes the training data for I different statistical models M_1, \dots, M_I . The forecast PDF is then given by [71]

$$p(y) = \sum_{i=1}^I p(M_i|y^T)p(y|M_i). \quad (5.16)$$

Here $p(y|M_i)$ is the forecast PDF of M_i alone and $p(M_i|y^T)$ denotes the a priori probability of model M_i being correct given the training data [71]. The resulting PDF $p(y)$ is given by the sum of the individual model PDFs, weighted by their a priori probabilities [71]. Applying this to dynamical models, the ensemble PDF $p(y)$ quantifies the uncertainty about the best member in our ensemble [71]. Each of the forecasts f_1, \dots, f_I by the models 1 to I has an a priori probability w_i of being the best member among the ensemble and a conditional PDF $g_i(y|f_i)$. Centered at the individual deterministic model forecasts, the width of the PDF can be viewed as the accuracy of the forecast, given that it is the best forecast in the ensemble. The individual w_i 's are weights that sum up to one. The ensemble PDF is then defined by

$$p(y|f_1, \dots, f_I) = \sum_{i=1}^I w_i g_i(y|f_i), \quad (5.17)$$

which constitutes a weighted sum of PDFs [71]. Figure 5.11 displays the weighted PDFs at $t^* = 2$ where the individual PDFs $g_i(y|f_i)$ are centered at the respective deterministic model forecast. In addition, the envelope that comprises the 90% probability level of the MME is displayed. The deterministic ensemble forecasts \tilde{f} equals the expectation of the ensemble PDF $p(y)$ [71] and is calculated according to Equation (5.18). Deriving \tilde{f} from the global maximum of $p(y)$ instead results in impaired performance.

$$\tilde{f} = w_i \cdot f_i. \quad (5.18)$$

5.5.1 Maximum-likelihood estimation of ensemble parameters

To determine w_i and the standard deviation σ of the PDF g_i it is useful to employ a maximum likelihood method [147] to combine various PDFs such that their sum matches a given distribution best. Deriving the ensemble parameters directly, by counting how often each model is the best among the ensemble and by computing the standard deviation of the model forecasts, showed to be inferior to such an optimization method.

The maximum-likelihood estimation approach is based on the idea that only a random sample from an unknown population is available [147]. The goal of this method is to estimate the properties of a PDF for a population that has most likely generated the sample [147]. Therefore, a PDF formulation must be assumed in the first place (e.g. a normal distribution). Exemplarily, the likelihood function for a sample of observed data y_o at observation site or

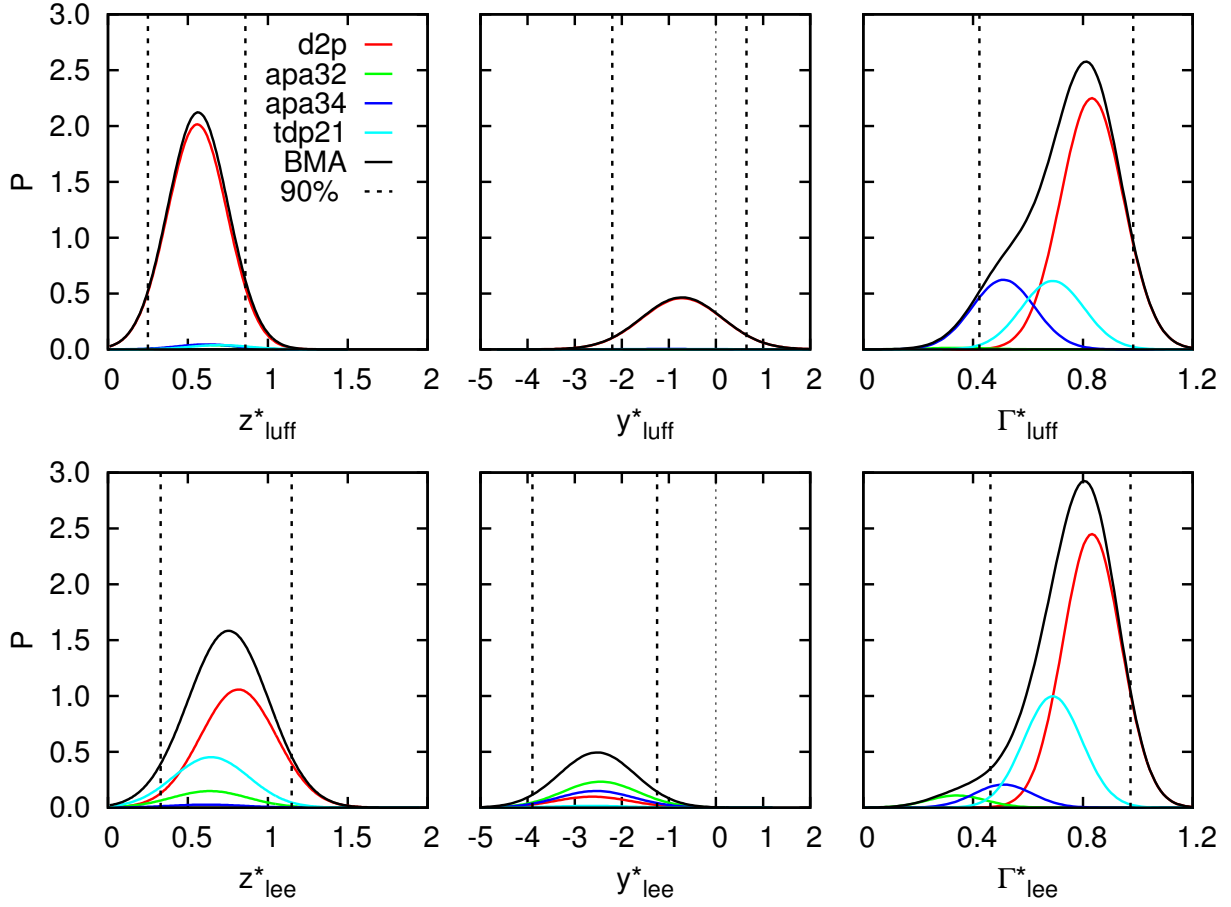


Figure 5.11.: Ensemble forecast at $t^* = 2$ with the PDFs of the individual model predictions. The black lines show the ensemble PDF and the dotted vertical lines denote the 90% uncertainty envelope.

time o , that is to be approximated by a single model (the PDF $g(\mu, \sigma^2)$ with mean μ), is

$$L(\mu, \sigma) = \prod_{o=1}^n g(y_o | \mu, \sigma). \quad (5.19)$$

It describes the probability of the training data y_o to be estimated as a function of μ and σ [71] (Figure 5.12a). Finding the maximum of L as a function of μ and σ is equivalent to finding the optimal PDF that describes the population from which the sample was extracted. Note, that this method is sensitive to the starting values, as the likelihood function may exhibit more than one maximum [71].

In the ensemble more than one model is available (in contrast to the example before), where each is represented by a PDF with given mean f_i . The function that is to be maximized for the MME is [71]

$$\log(L(w_i, f_i, \sigma)) = \sum_{o=1}^n \log \left(\sum_{i=1}^I w_i \cdot g_i(y_o | f_{i,o}, \sigma) \right). \quad (5.20)$$

The logarithm is applied such that the likelihood function can be expressed by sums instead of products to simplify the calculation. Here, $w_i g_i(y_o|f_{i,o}, \sigma)$ is the weighted a priori likelihood that observation y_o is predicted correctly by forecast f_i for a given standard deviation σ , which is depicted in Figure 5.12b. As an analytical maximization is not possible here the Expectation-Maximization (EM) algorithm [148,149] is utilized to find the optimal ensemble properties from the training dataset [71]. It determines the parameters of a weighted sum of PDFs that have most likely produced the available distribution iteratively, with the current iteration denoted as j . Explained more descriptively, the goal is to approximate the PDF of the model deviations of the ensemble forecast in the most realistic way, based on all available model forecasts for the training data. For a single model this problem would be solved easily by directly dressing it with its PDF of model deviations derived from observations. For multiple models, however, this is achieved by determining w_i and σ such that the ensemble PDF reproduces on average the distribution of model deviations for the whole dataset in an optimal way. Approximating PDFs as a combination of various individual PDFs is also known as finite mixture modeling [150].

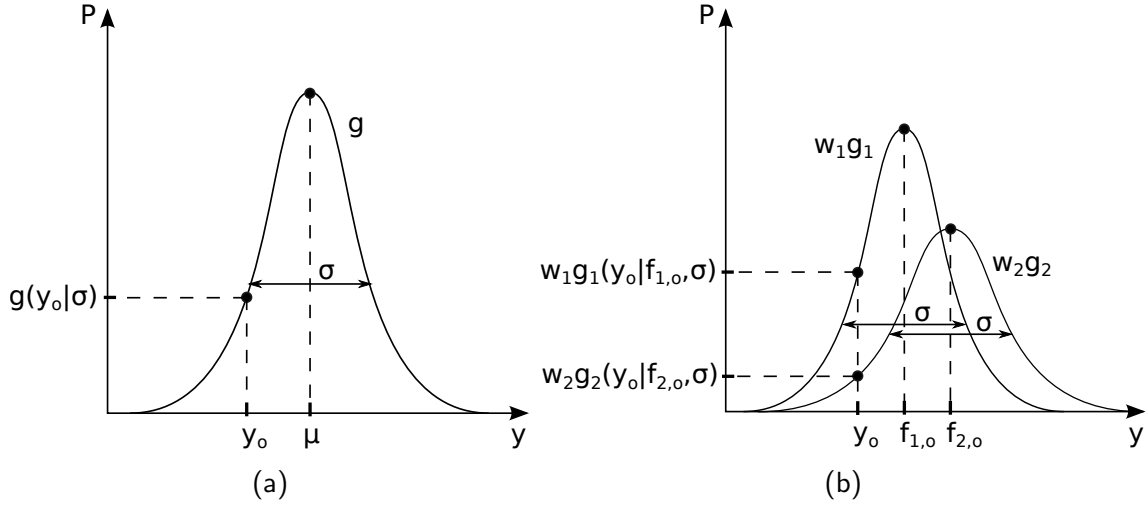


Figure 5.12.: Computation of $g(y_o|\mu, \sigma)$ (a) and $w_i g_i(y_o|f_{i,o}, \sigma)$ (b).

At first the so-called unobserved quantities $\tilde{z}_{i,t}^{(j)}$ must be computed in the expectation step, employing an initial guess for w_i and σ in the first iteration. Under the assumption that the identity of the best model cannot be determined directly, the unobserved quantities are a measure of which model i offers the best forecast f_i for a given observation at time t [71]. Here, the best forecast does not simply depend on the RMSE but on w_i and σ . Given the case that model one achieves the best forecast at a specific time, $\tilde{z}_{1,t}^{(j)}$ eventually approaches the value one, while it is zero for the other models. However, during the iterations of the EM-algorithm $\tilde{z}_{i,t}^{(j)}$ may also take values other than zero or one [71] and does not necessarily converge to either of the values. It is calculated according to

$$\tilde{z}_{i,t}^{(j)} = \frac{w_i^{(j-1)} g(y_t|f_{i,t}, \sigma^{(j-1)})}{\sum_{i=1}^I w_i^{(j-1)} g(y_t|f_{i,t}, \sigma^{(j-1)})}. \quad (5.21)$$

In the maximization step $\tilde{z}_{i,t}^{(j)}$ is utilized to determine the individual model weights $w_i^{(j)}$ and

the variance $\sigma^{2(j)}$ with n being the number of total observations.

$$w_i^{(j)} = \frac{1}{n} \sum_{t=1}^n \tilde{z}_{i,t}^{(j)} \quad (5.22)$$

$$\sigma^{2(j)} = \frac{1}{n} \sum_{t=1}^n \sum_{i=1}^I \tilde{z}_{i,t}^{(j)} (y_t - f_{i,t})^2 \quad (5.23)$$

The iteration can be stopped when the parameters are iterated to convergence. In our case 200 iterations proved to be sufficient such that the change of $w_i^{(j)}$ is below 0.5% eventually. Note that σ is not distinguished between the individual models as it represents the standard deviation of the respectively best model, which can be any of the members. Implementing model dependent standard deviations σ_i is de facto possible but turns out to impair the performance significantly due to highly bi-modal ensemble forecasts, caused by single models with small σ_i [151].

5.5.2 Distribution formulation

If the distributions of the deviations of the measurements from the deterministic predictions are unknown, normal distributions are a good estimate to generate the ensemble forecast [151]. However, it turns out that the unweighted mean model error distributions are rather leptokurtic as they exhibit higher peaks and fatter tails than Gaussian distributions, especially for the y -forecast (Figure 5.13) [151]. This can be explained by the intermittent nature of wind [152], manifested in the form of gusts. R  th et al. relate the thereof resulting fat tails in the turbulence distribution to so-called phase correlations [153]. Not being resolved by the measurement-derived 10 minute wind averages that are employed as input data, gusts can cause large errors due to misjudged drifting velocities. Further, the wind measured by the instrumentation and the wind sensed by the vortices may deviate substantially depending on the spatial distance between measurement device and vortex [151]. As wind also affects the interaction of the vortices with the ground, the PDF of model deviations for the z^* -forecast is non-Gaussian likewise. In addition, the model deviation PDFs turn out to be skewed and not to be centered at zero [151]. In order to enable the ensemble to achieve leptokurtic forecast distributions, the individual model PDFs that compose the ensemble forecast must be leptokurtic likewise. Therefore, the training strategy is as follows [151]:

1. Determine the formulation of the non-Gaussian mean model error PDF in the preconditioning step
2. Use the in step one calculated PDF formulation as baseline to calculate the ensemble parameters σ and w_i as described in Section 5.5.1

Fits for step one are here obtained by applying the EM algorithm to four single Gaussian PDFs with the standard deviations $\tilde{\sigma}_{1,2,3,4}$, and mean $\tilde{\mu}_{1,2,3,4} = 0$ (see Fig. 5.14). The standard deviations $\tilde{\sigma}_{1,2,3,4}$ have a fixed ratio in dependency of $\tilde{\sigma}$, such that each of the four PDFs may approximate a different part of the mean model error distribution found in the training data. Exemplarily, the PDF defined by $\tilde{\sigma}_1$ is intended to model the tails, while $\tilde{\sigma}_4$ increases the

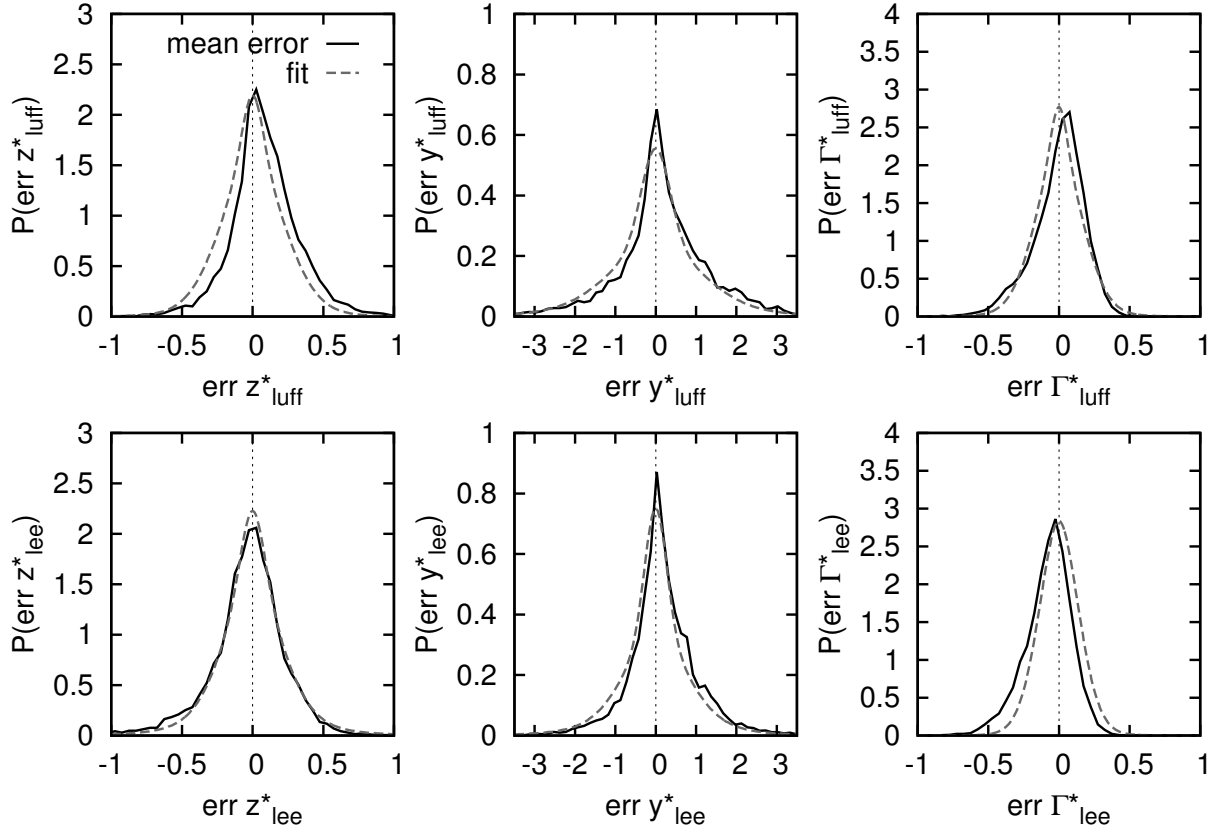


Figure 5.13.: Average distributions of model deviations (all models) and their fits for the training dataset consisting of landings from WakeMUC, WakeFRA and WakeOP [151].

curtosis in the middle part. The algorithm then determines $\tilde{\sigma}$ and the weights $\tilde{w}_{1,2,3,4}$ such that the best fit is achieved under the given constraints. It is important to mention that it is not intended to find a perfect fit here. Although sophisticated fits can be easily achieved by this technique by choosing $\tilde{\mu}_1 \neq \tilde{\mu}_2 \neq \tilde{\mu}_3 \neq \tilde{\mu}_4$, the goal is rather to find a PDF that approximates the kurtosis [151]. The rest will be managed by the training described in Section 5.5.1, where the individual model forecasts are incorporated. The resulting fits are depicted in Figure 5.13 for the WakeMUC/ WakeFRA/ WakeOP training sample [151]. Applying leptokurtic instead of Gaussian distributions eventually delivers better probabilistic prediction results as the tails of the PDFs are approximated more precisely [151]. This procedure can once more be understood as finite mixture modeling [150].

5.5.3 Temporal development of model uncertainty

In the original approach [71] the uncertainty is assumed to be constant over time. However, Figure 5.15 clearly indicates that the error is increasing with advancing vortex age for all parameters and regardless of the model. While the steepest error increase occurs below one t_0 , the growth progresses more slowly afterwards. Qualitatively, the error development is similar among all models featuring an increase for all parameters. Nevertheless, the gradients of the

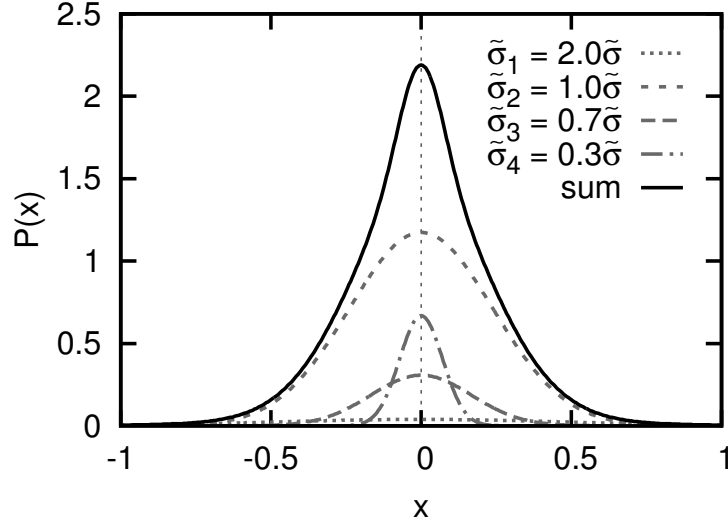


Figure 5.14.: Method to approximate the PDFs by employing the EM-algorithm to four single normal distributions with fixed relations of $\tilde{\sigma}_i$, with $\tilde{\mu} = 0$ and with suitable weights \tilde{w}_i . Instead of three, as in a previous study [151], four PDFs are utilized for an enhanced fit.

linear fits are different where D2P shows the lowest for the WakeMUC/ WakeFRA/ WakeOP training dataset, except for y_{lee} . The average value in the first time bin can be understood as initial condition uncertainty and has about the same value for all models. It can be noted that the NASA models exhibit a large performance difference between the forecast of y_{luff}^* and y_{lee}^* . This tendency is exclusively observed in ground proximity and must thus be related to the vortex-ground interaction.

Given these results, constant standard deviations are a simplification, leading to overly conservative envelopes for early and to non-conservative bounds for later vortex ages. Note, that in the approach with constant parameters the impact of the early conservative and the late non-conservative envelopes may compensate each other such that on average the probabilistic results seem satisfactory nevertheless.

The temporal increase in model error can be considered by applying the maximum likelihood method described before to vortex age bins. To avoid overfitting of the ensemble to the training data both the weights and the standard deviations are approximated linearly instead of interpolating. As w_i and σ are dependent (see Equation 5.22 and 5.23), both parameters are treated vortex-age dependently. Figure 5.16 and Figure 5.17 display the standard deviations derived in this manner for the DLR and NASA measurements, respectively. Only measurements below $5 t_0$, where the data coverage is sufficient, are evaluated. σ clearly increases linearly for z^* and y^* for all evaluated datasets. While σ_{Γ^*} increases in the DLR data as well, it decreases for the NASA training sample. Although a square fit would be more suitable for σ_{Γ^*} for the WakeMUC/ WakeFRA/ WakeOP test sample, a linear approximation is expected to be the best compromise as the order of the best fit deviates between different datasets.

Figure 5.18 and Figure 5.19 depict the corresponding model weights. While clear trends are observed for the WakeMUC/ WakeFRA/ WakeOP training sample, they remain less clear for the MEM95/ DFW97 training set. Nevertheless, for the sake of simplicity and robustness a

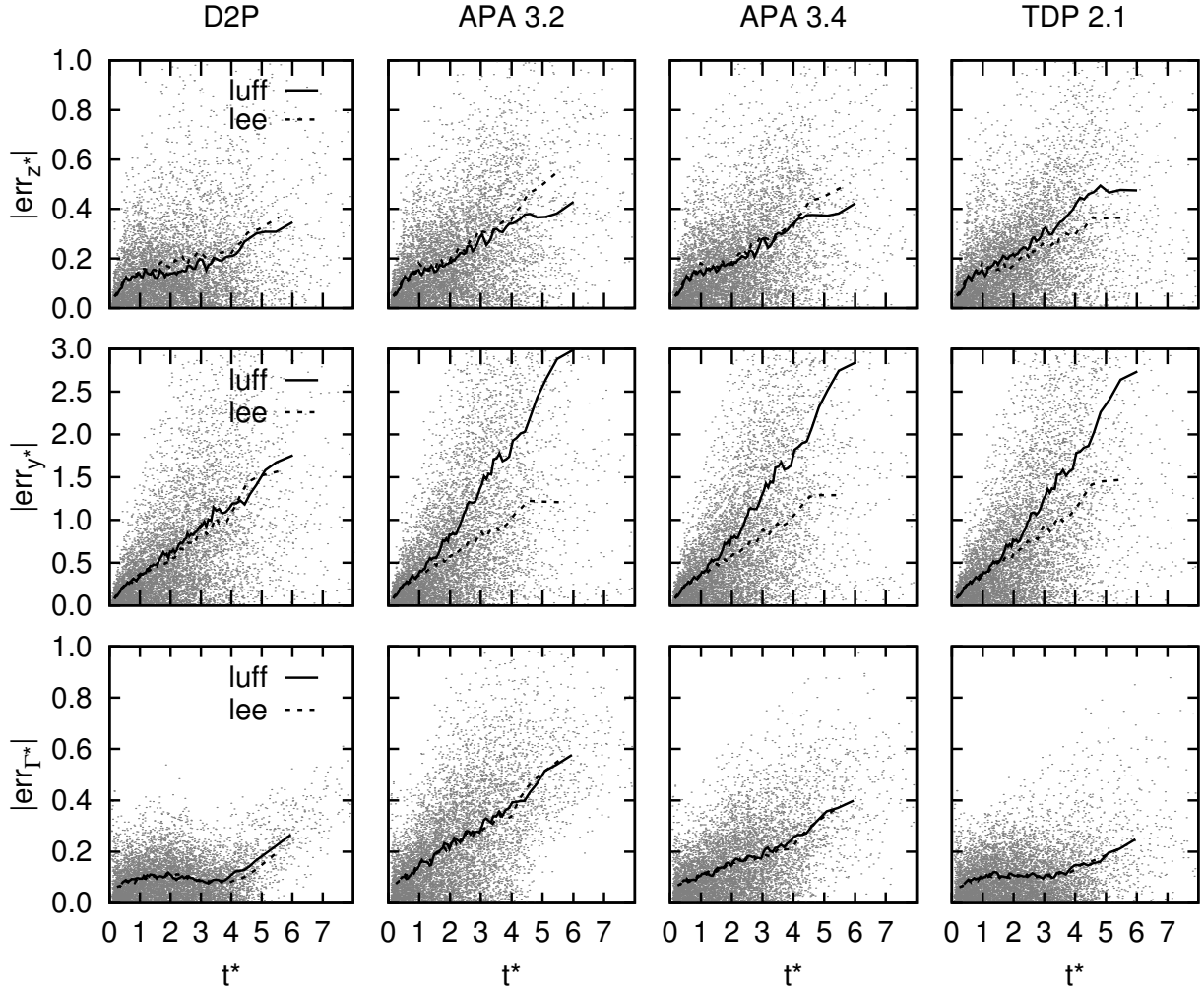


Figure 5.15.: Temporal development of the model uncertainty in terms of error magnitude against vortex age for the WakeMUC/ WakeFRA/ WakeOP training sample.

linear approximation is employed again. Note that the constraints $\sigma, w_i > 0$ and $\sum_{i=1}^I w_i = 1$ must be satisfied when applying the parameters eventually.

Also note that the shapes of the error distributions differ depending on vortex age. Both Gaussian and leptokurtic PDFs may occur during different stages of the wake vortex lifetime. It is found that the PDFs for early vortex ages are more leptokurtic while they turn more Gaussian towards the end of the vortex lifetime. As a consequence the approximation of the model deviation PDFs by the EM algorithm must be applied vortex-age-binned as well. As a consequence, the parameters are approximated linearly once more.

5.5.4 Potential

This approach has been chosen as it enables uncertainty envelopes with arbitrary uncertainty levels. Being based on model deviation PDFs, this method directly addresses the uncertainties derived from the assessment of observations. This way the BMA does not only include model

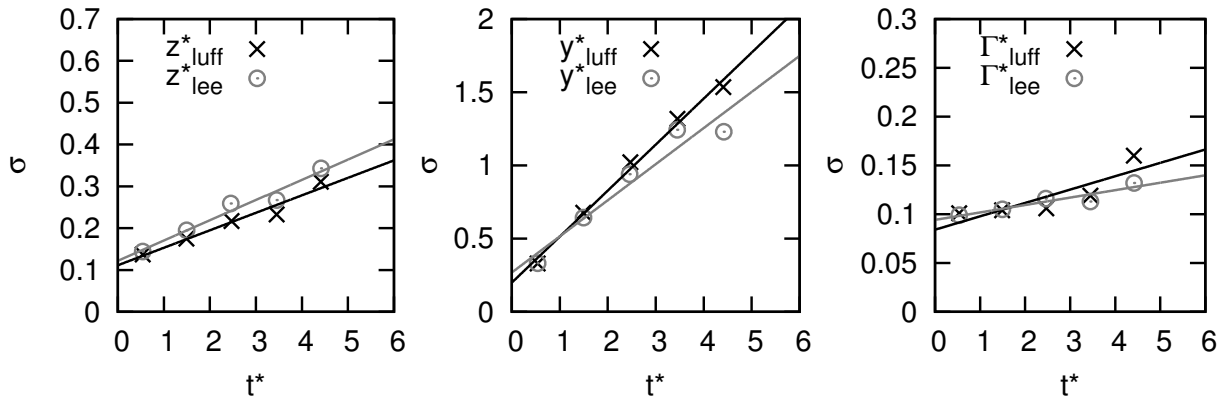


Figure 5.16.: Temporal development of σ if computed by maximum likelihood based on the test dataset from WakeMUC, WakeFRA and WakeOP.

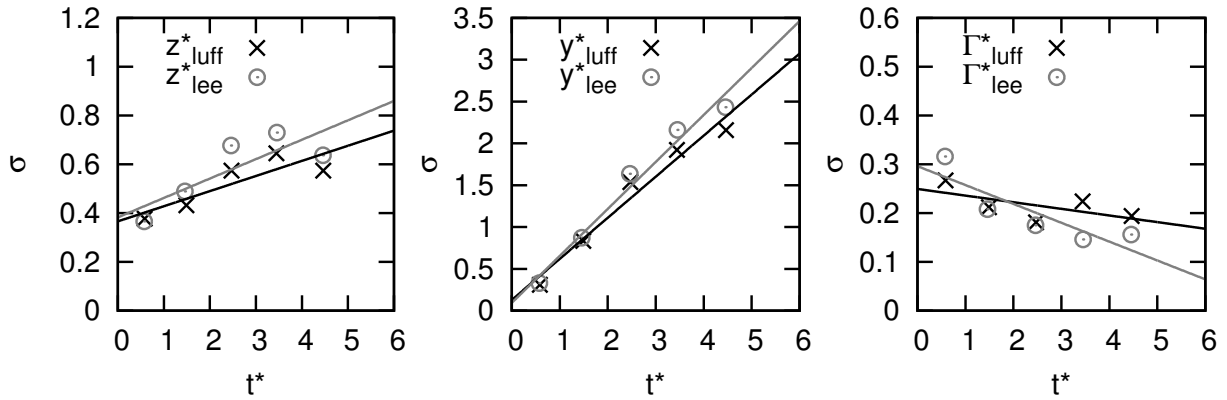


Figure 5.17.: Temporal development of σ if computed by maximum likelihood based on the test dataset from MEM95 and DFW97.

uncertainty, but if applied vortex-age dependently also initial condition and ambient condition uncertainty. This offers the opportunity to incorporate the vortex-age dependence of the model errors which is believed to further increase the reliability of the uncertainty envelopes. In comparison to the REA the BMA applies different skill measures to the models, such that both approaches are expected to perform differently.

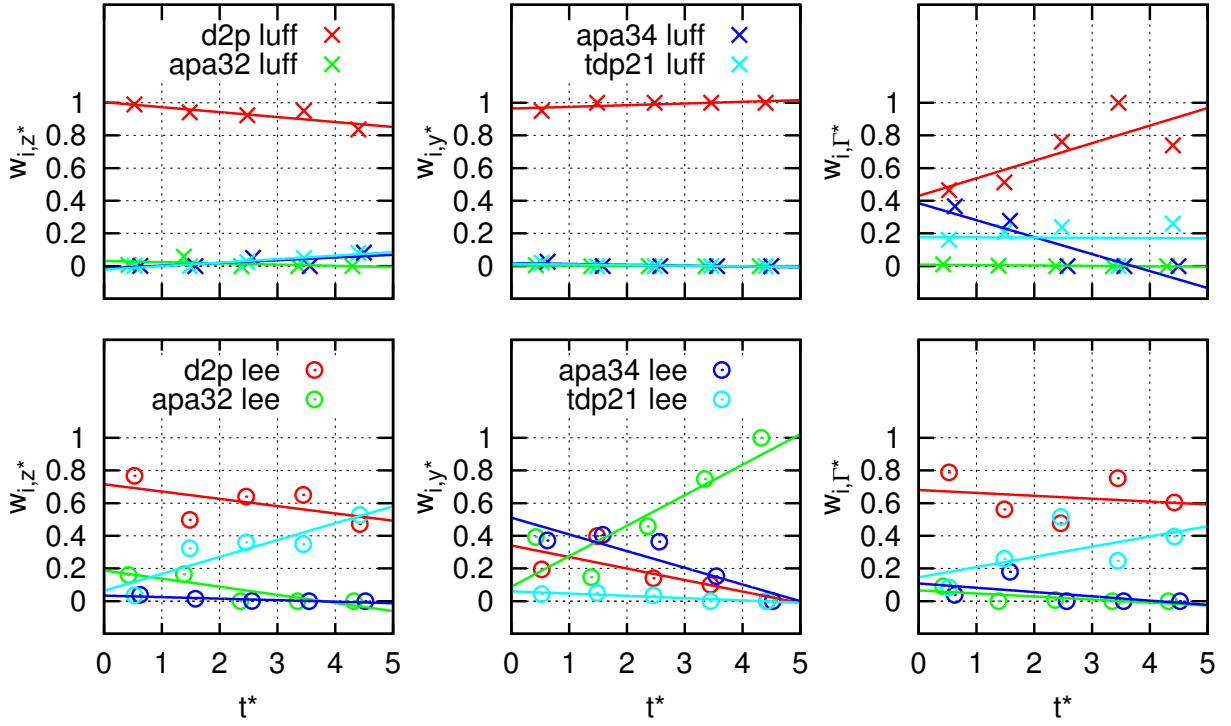


Figure 5.18.: Temporal development of w_i if computed by maximum likelihood based on the test dataset from WakeMUC, WakeFRA and WakeOP.

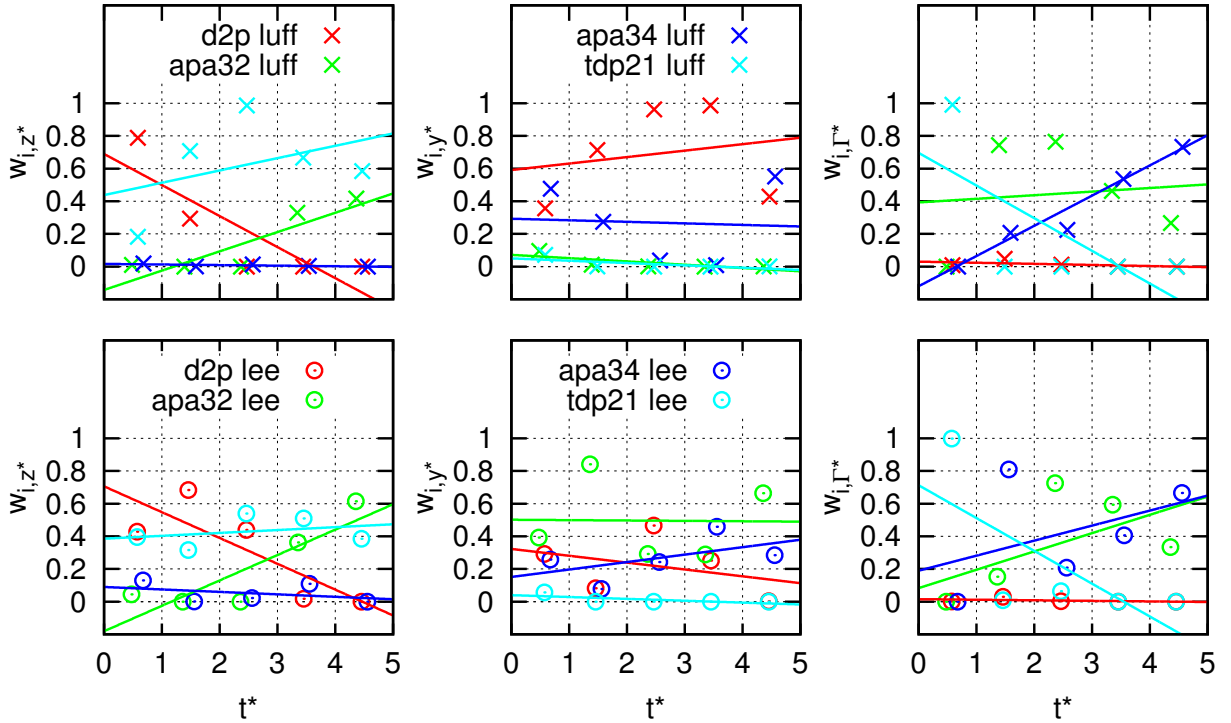


Figure 5.19.: Temporal development of w_i if computed by maximum likelihood based on the test dataset from MEM95 and DFW97.

5.6 Monte Carlo Simulation (MCS)

In the course of this work it has often been argued whether the initial condition uncertainty must be taken into account directly by introducing perturbations. Already invented as early as in the 1930s [154] this approach has various applications of which one is weather forecasting. Lorenz stated in 1963 that “slightly differing initial states can evolve into considerably different states” [155]. Under the assumption that the true state of the atmosphere is hard to capture, the MCS is used to find all possible forecast developments by perturbing the initial conditions. Therefore, a set of initial condition samples is drawn from a probability density $p(y)$ [154], describing the initial condition uncertainty, from which a set of simulations is generated. The spread of the simulations can then be utilized to compute probabilistic envelopes. It is important to provide a sufficient number of perturbed initial conditions in order to achieve convergence of the uncertainty bounds. The drawback of this approach is the increasing simulation time due to the increased number of necessary simulations. This makes it not only expensive but also difficult to implement for time-sensitive applications. Nevertheless it is employed by NASA in Körner et al. [156] due to theoretical interest. Furthermore, it has already been applied to wake vortex forecasts before by DeVisscher et al. [33], using a single model and limiting the number of runs by employing a bootstrap resampling technique.

5.7 Lagged Average Forecasting (LAF)

The LAF [157,158] also addresses initial condition uncertainty. However, instead of introducing random perturbations the LAF is initialized with current and past observations such that the simulations start with time lags [157]. The rationale of this approach is to allow for the uncertainty of the time at which an event will occur rather than for the initial condition uncertainty. An optimal model would be able to predict the observation that is used in the following initialization step correctly. Hoffman and Kalnay [157] showed the superiority of this approach to Monte Carlo forecasting and simple single-model predictions. However, the time horizons they utilize for their predictions comprises 25 days, and the time lag between the observations used as initial values amounts to six hours. Assuming that wake vortices may live on average as long as $12 t_0$ for calm conditions (see Figure 4.4) this would correspond to time lags of about $0.12 t_0$. For an aircraft with $t_0=20$ s (e.g. A320) this would mean that the initialization observations needed to be separated by 2.4 s. However, such a high temporal resolution of meteorological measurements can seldom be provided for wake vortex measurements. Given this, the LAF is not suitable for the application to wake vortex forecasts.

6. Results

Before the ensemble skill is evaluated by applying the DEA, REA and BMA approaches to the captured LiDAR data, the available wake vortex measurements are employed to further study vortex physics. In a next step the deterministic and probabilistic performance of the wake vortex ensemble is presented.

6.1 Evaluation of the LiDAR Wake Vortex Measurements

6.1.1 Impact of initial height on descent and decay

The WakeMUC data comprises landings with various vortex generation heights, which is why the z_0^* dependence of wake vortex behavior can be evaluated. The findings presented here have been published in Holzäpfel et al. [32] in a similar manner. However, the number of landings and the initial conditions slightly changed in comparison. Additionally, the D2P model has been further developed as described in Section 4.2.2.

Figure 6.1 depicts the measured vertical wake vortex position z^* , the vortex spacing b^* and the circulation Γ^* in scatter plots for 907 landings. Additionally, the plots include the averages of wake vortex behavior for various generation altitudes for the luff and lee vortex separately. As the LiDAR is usually not able to capture incoherent vortices which occur below $0.4\Gamma_0$, the temporal Γ^* averages are biased towards later vortex ages for low circulations. For this reason a Γ^* -binned average is provided additionally to the t^* -bins. The curves are derived from bins containing a fixed number of measurements.

The upper left panel in Figure 6.1 indicates that the minimum vortex altitude increases from $0.5b_0^*$ to $1b_0^*$ with increasing vortex generation height. Possibly the interaction with the ground does not only depend on the altitude as expected so far, but also on time as diffusion processes are involved. Secondary vortices might detach after the far field of a vortex has acted long enough upon the boundary layer [32]. Besides, the results suggest that the lee vortex descends slower and initiates its rebound earlier while rising faster after having reached the lowest point [32]. This is in line with the results from LES [101, 102, 104, 105] and can be referred to the production of secondary vorticity in the shear layer at the ground which is supported by crosswind. The upper right panel indicates that the vortices diverge faster than it is assumed by theory. In Figure 6.1 b^* is determined by

$$b^* = \sqrt{z^{*2} + y^{*2}}. \quad (6.1)$$

This supports the theory that the vortex-ground interaction starts earlier than expected. D2P assumes that the lateral separation starts when the vortices have reached $1.5 b_0$ above ground. However, even the vortices generated within the highest altitude interval show lateral

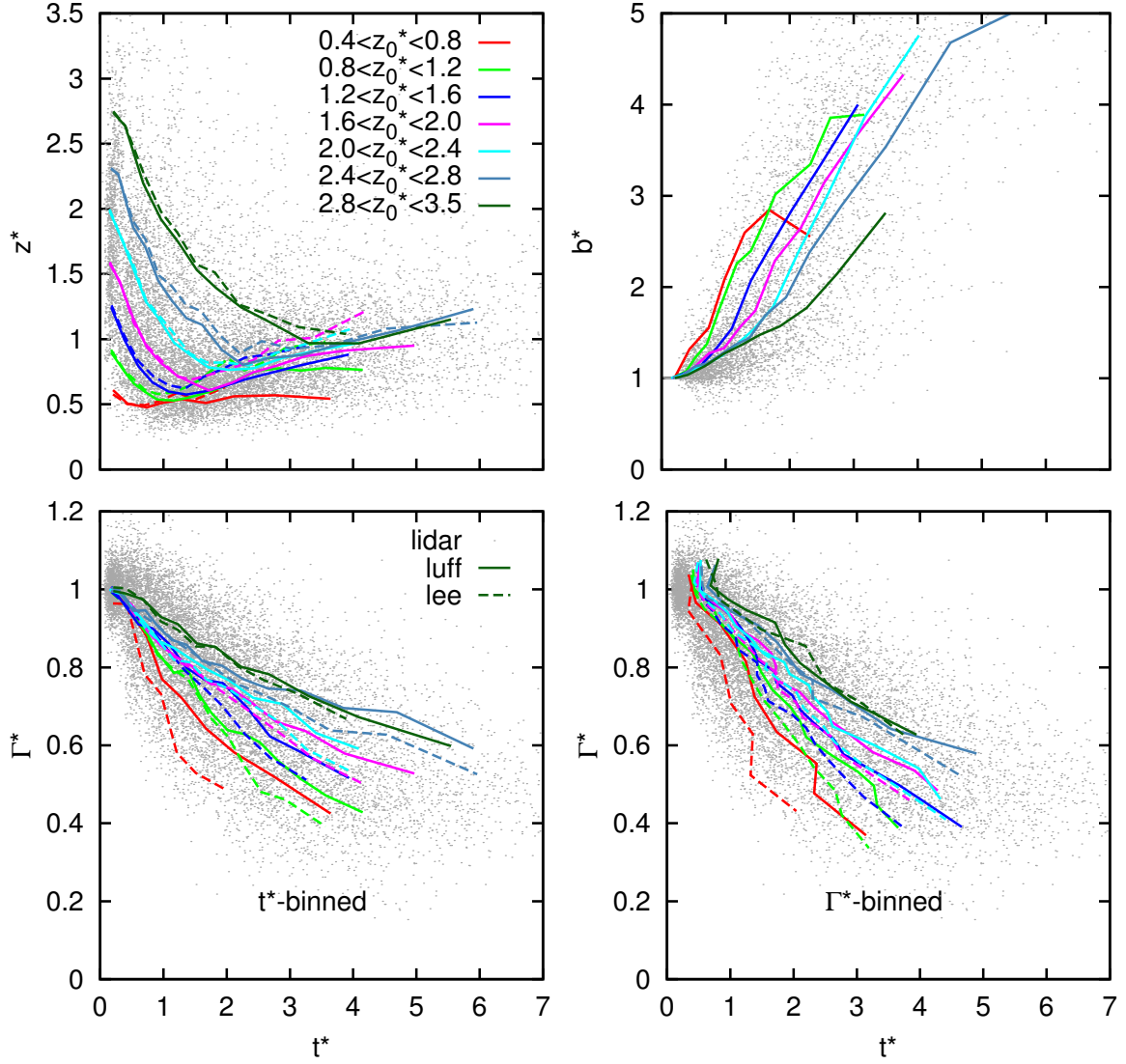


Figure 6.1.: Average wake vortex transport and decay for WakeMUC (911 landings) for various intervals of vortex generation altitude.

divergence at early stages. Firstly, the lower panels in Figure 6.1 point out that the decay rate is ordered according to the vortex generation height, with higher generated vortices living on average longer [32]. The life time of the on average longest-lived luff vortex for $z_0^* = 0.4-0.8$ is more than halved compared to $z_0^* = 2.4-2.8$ [32]. This can possibly be ascribed to either the increased turbulence level close to the ground or to the end effects that are expected to reach vortex segments closer to the touchdown point earlier [32]. Secondly, the results demonstrate without exception that the lee vortex decays on average faster than the luff vortex [32]. Holzäpfel and Steen [18] expected this trend as a consequence of stronger secondary vorticity on the lee side, but they could not detect it in the WakeFRA LiDAR data. The here presented findings and results from LES [101] support their hypothesis. Comparing the different methods to derive the average developments for Γ^* it is obvious that the curves level less out at large vortex-ages for the Γ^* -binned approach.

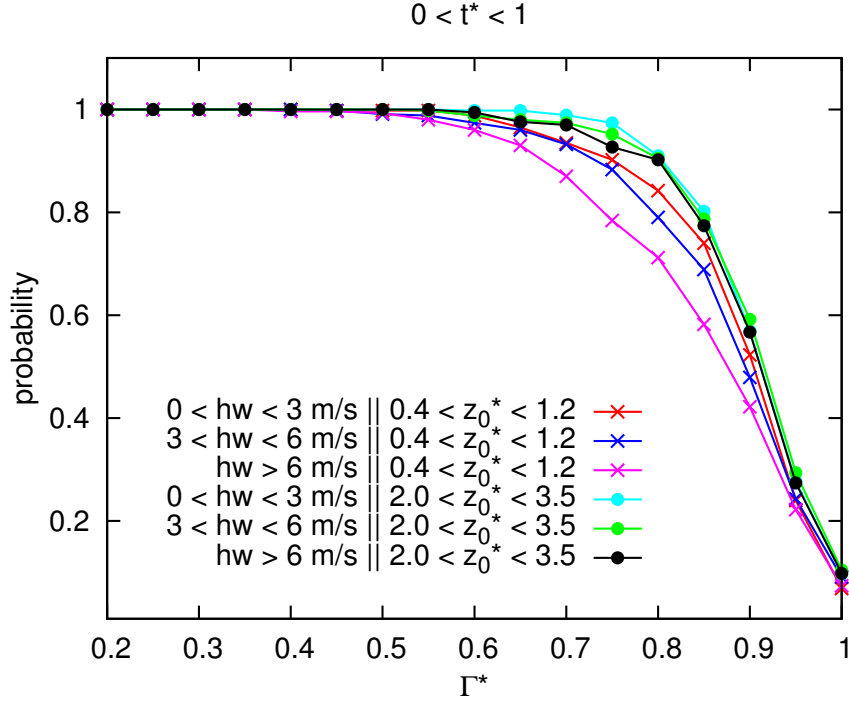


Figure 6.2.: Cumulative distribution of vortex strength for $t^* < 1$ and various initial height intervals for the WakeMUC dataset (reprinted from Holzäpfel et al. [159]).

To find further evidence for the existence of the so-called end effects [7] mentioned in Chapter 2.8.2, the decay rate for various height intervals is evaluated. Under the assumption that these effects travel with a designated velocity and start from the touchdown point of the aircraft, faster decay rates would be expected for lower altitudes [159]. For this reason the cumulative distribution of vortex strength is evaluated for headwind speeds between 0-3 m/s, 3-6 m/s and for velocities larger than 6 m/s, and for the two intervals of vortex generation altitudes $0.4-1.2 b_0^*$ and $2.0-3.5 b_0^*$. The results in Figure 6.2 demonstrate that the stronger the headwind the less likely it is to observe strong vortices [159]. This trend is less significant at larger initial altitudes where the disturbances have to travel larger distances from the touchdown point. Another explanation for this effect could be headwind-induced turbulence which is increased close to ground (boundary layer) [159].

6.1.2 Impact of wind speed and direction on descent and decay

Holzäpfel et al. [90] concluded from LES results that not only crosswind but also headwind may have an impact on wake vortex behavior which they concluded from LES results. In order to find further evidence and assess the LES capabilities, the LiDAR data captured during the WakeMUC, WakeFRA and WakeOP campaigns is evaluated with regard to head- and crosswind influence. Figure 6.3 shows the head- and crosswind which the vortices were exposed to on their descent trajectory. The linear fits demonstrate that strong headwind does not necessarily imply strong crosswind. Therefore, evaluating the impact of either head- or crosswind without excluding the other is appropriate.

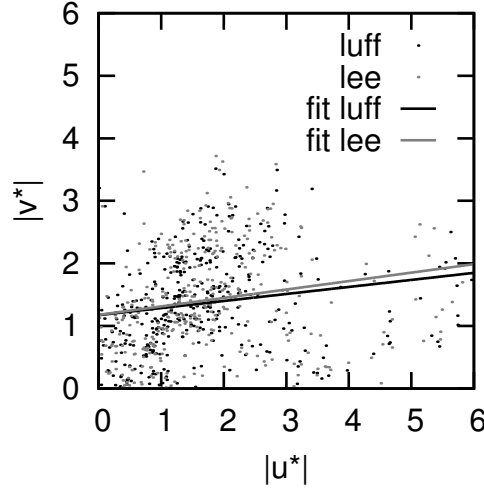


Figure 6.3.: Correlation of headwind u^* and crosswind v^* that the vortices were exposed to for $z_0^* < 1.5$.

Figure 6.4 displays the LiDAR measurement derived minimum (before rebound) and maximum vortex height (after rebound) above ground against head- and crosswind. Additionally, the respective trends and LES results that comprise simulations with different ratios of head- and crosswind are included.

First of all, the trends suggest that headwind has only minor impact on descent and rebound. A possible explanation is that headwind induced vorticity is oriented orthogonal to the vortex induced secondary vorticity and thus has no impact on wake vortex descent or rebound. Furthermore, the trends confirm that the rebound of the luff vortex is attenuated in the presence of crosswind (decreasing maximum descent and rebound height), while it is supported for the lee vortex (increasing maximum descent and rebound height). Both analytical considerations [107] and the depicted LES results imply that the detachment of secondary vorticity, triggered by the luff vortex, and the subsequent rebound is suppressed for wind speeds above $4w_0$. As the data coverage for wind speeds above $3w_0$ is poor, an exact value cannot finally be determined here. Extrapolation of the trend derived from the LiDAR measurements, however, suggests that this effect only occurs above $5w_0$.

In general the trends are qualitatively well reflected by the LES. While the maximum descent height conforms very well with the LiDAR data, the rebound height seems to be too high in the LES. One possible explanation for this discrepancy might be the LiDAR only being able to track the vortices until they have reached a circulation of about 50% of the initial value on average. This also involves that their maximum altitude cannot finally be determined. Therefore Figure 6.4 also depicts the vertical vortex position found in LES at the time the circulation reaches $0.5\Gamma_0^*$, which turns out to be closer to the LiDAR fit.

Figure 6.5 depicts the time $t_{0.6\Gamma_0^*}$, at which the vortices reach 60% of the initial circulation, against normalized head- and crosswind speed. The value 60% is chosen as most vortices can still be tracked by the LiDAR at that stage. The results suggest that headwind accelerates decay to the same extent for both the luff and the lee vortices (parallel trends). While crosswind affects the luff vortex in a similar way as headwind, it enhances decay additionally for the lee vortex where the secondary vortices are expected to be stronger. In comparison to

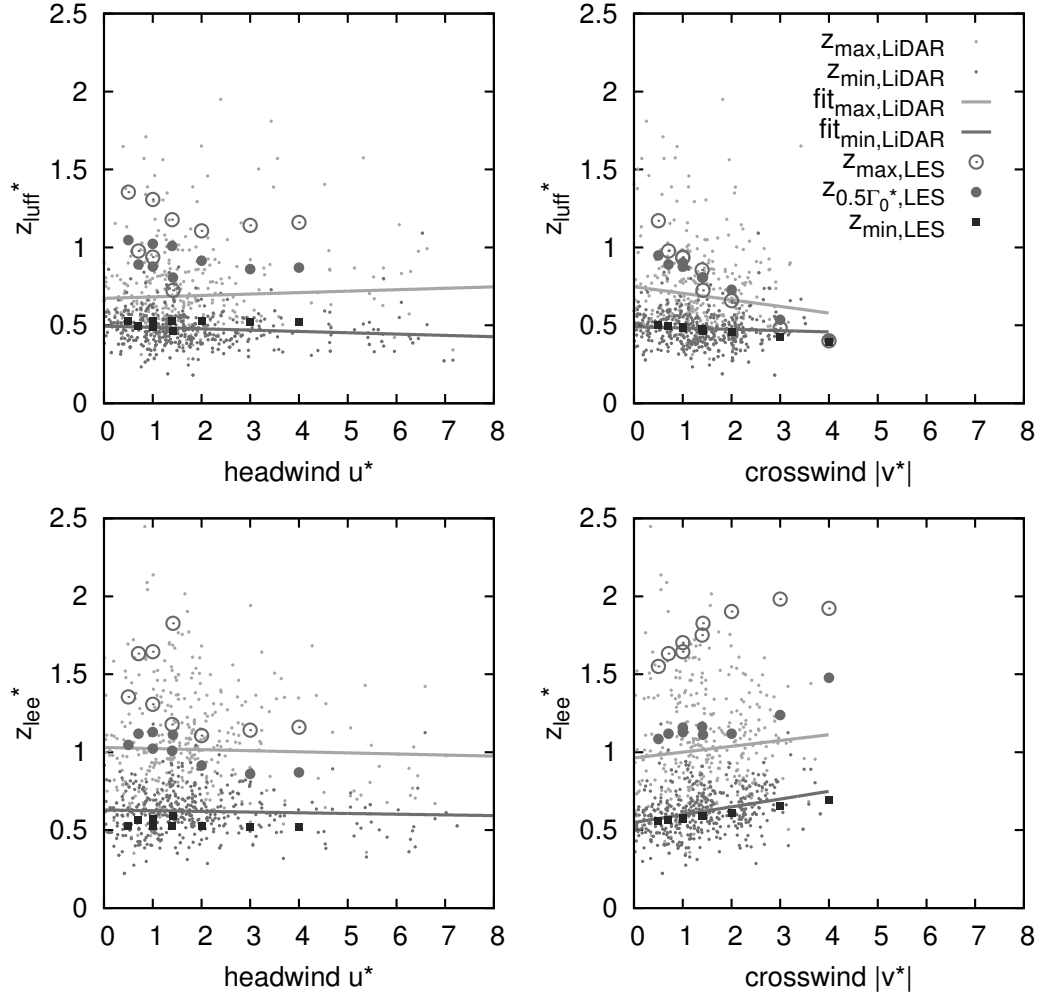


Figure 6.4.: Correlation of average wind exposure, and wake vortex descent and rebound for $z_0^* < 1.5$ for 392 landings of WakeMUC, WakeFRA and WakeOP. The subscript $0.5\Gamma_0^*$ indicates the vortex height at the time when the circulation reaches $0.5\Gamma_0$. The LES results have been published in [90].

the LES results the crosswind impact on the lee vortex conforms very well with the measured trend, whereas the trends are contrary for the luff vortex. This can be well explained by the occurrence of ground linking of the luff vortex in the LES cases for crosswind speeds of 3 and 4 w_0 , such that the decay characteristics deviate strongly from the lee vortex [90]. For LES cases with $v^* = 0$ the effect of headwind on decay is the same for both luff and lee vortices.

6.1.3 Comparison of the WakeMUC measurements with D2P

Figure 6.6 and Figure 6.7 illustrate the comparison of the LiDAR measured mean and the averages of the D2P predictions (version with updated strength of secondary and tertiary vortices) of the individual cases. The initial altitude intervals are the same as in Figure 6.1 and consist of 0.4–0.8, 0.8–1.2, 1.2–1.6, 1.6–2.0, 2.0–2.4, 2.4–2.8 and 2.8–3.5 z_0 . The results

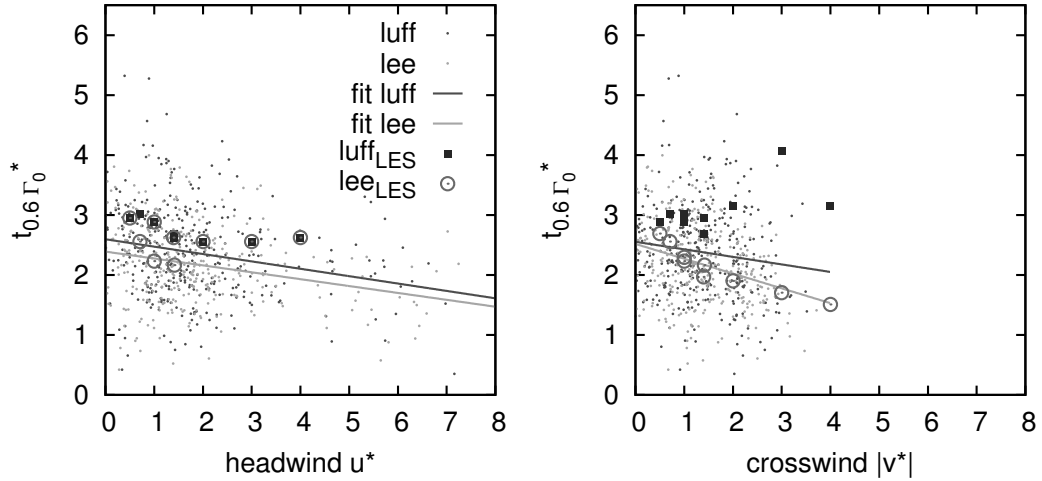


Figure 6.5.: Correlation of average wind exposure, and wake vortex decay for $z_0^* < 1.5$ for 392 landings of WakeMUC, WakeFRA and WakeOP. The LES results have been published in [90].

for the lowest z_0 -interval indicate that rebound in D2P is still too strong. However, note that only a small amount of data is available (11 landings) for this initial altitude range such that the significance cannot be guaranteed. For the intervals $0.8 - 1.2$ and $1.2 - 1.6$ z_0 the rebound is on average qualitatively well-predicted. Eventually, the plots for the succeeding z_0 intervals reveal that D2P overestimates the maximum descent distance. This goes together with a less pronounced rebound of the lee vortex in the measurements. Likewise a very good accordance is found in the interval $0.4 - 0.8$ z_0 for the vortex separation b . With increasing vortex generation height the vortex separation sets in both earlier and stronger in the LiDAR data than in the D2P predictions. This trend has been observed before in Section 6.1.2. One hypothesis for the described discrepancies between measurements and D2P-predictions is that the vortex-ground interaction already begins at larger heights than so far assumed. This could be due to the temporal development of the vortex far-field as a result of diffusive processes. Given this, not only the vortex height but also the time that a vortex remains at a specific altitude above ground would have an impact on its descent.

Moreover, the results demonstrate that D2P predicts significantly longer lived vortices than what is measured for the lowest evaluated z_0 interval. With increasing vortex generation altitude this discrepancy is reduced. One reason for this trend might be the existence of the before mentioned end effects which enhance decay especially in close ground proximity and which are not yet fully considered in D2P in the here employed version. In the DLR project L-bows a parametrization of the end effects has been developed for D2P that agrees much better with observations. The same evaluation is performed for the WakeFRA and WakeOP campaign. For both datasets the D2P model performs very well. The results can be found in the appendix (see Chapter A.3). Note that the here employed initial conditions for the WakeFRA dataset are different from the original training of D2P [15] which leads to different results.

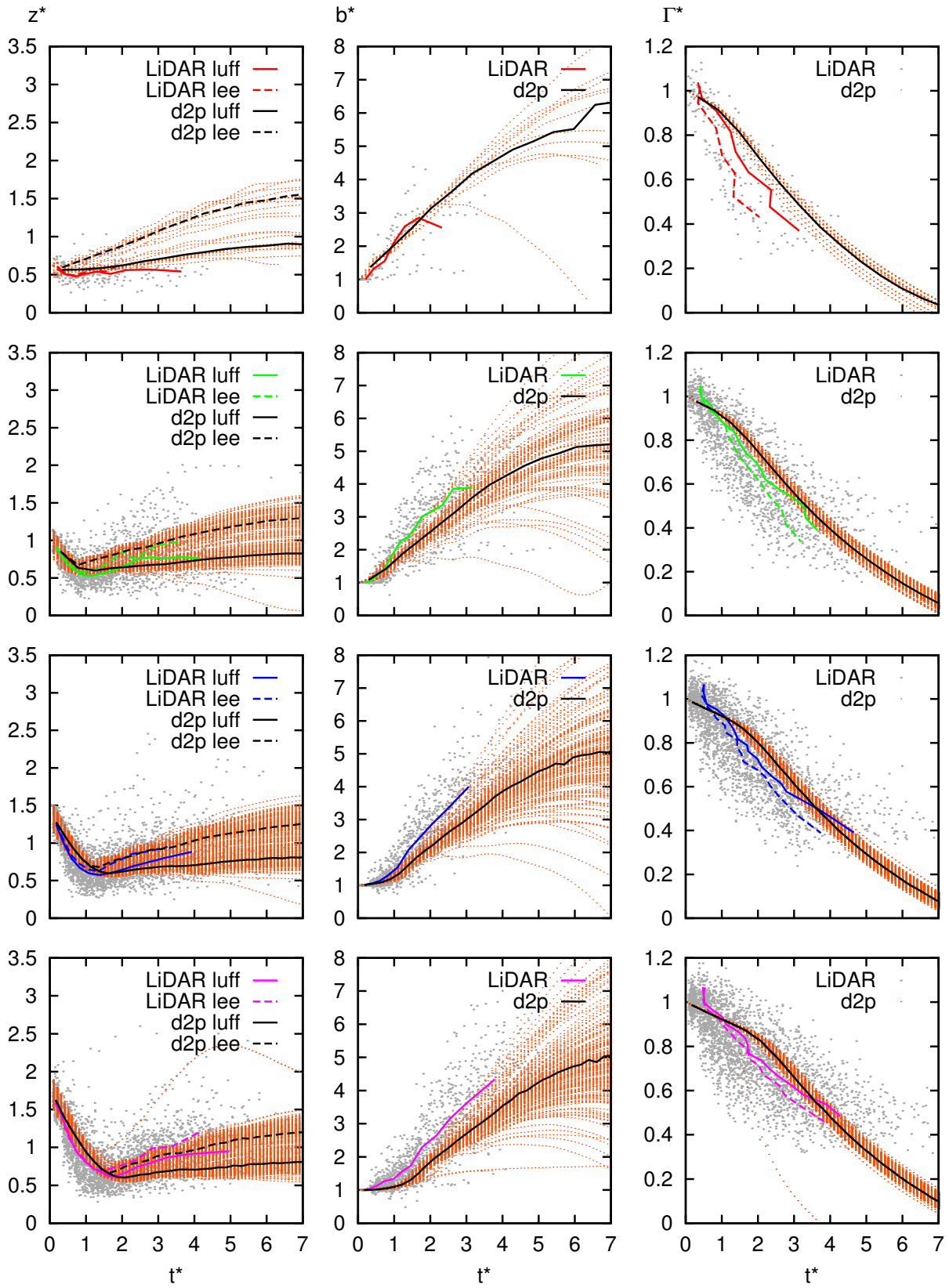


Figure 6.6.: Comparison of mean D2P prediction and the LiDAR mean for various height intervals.

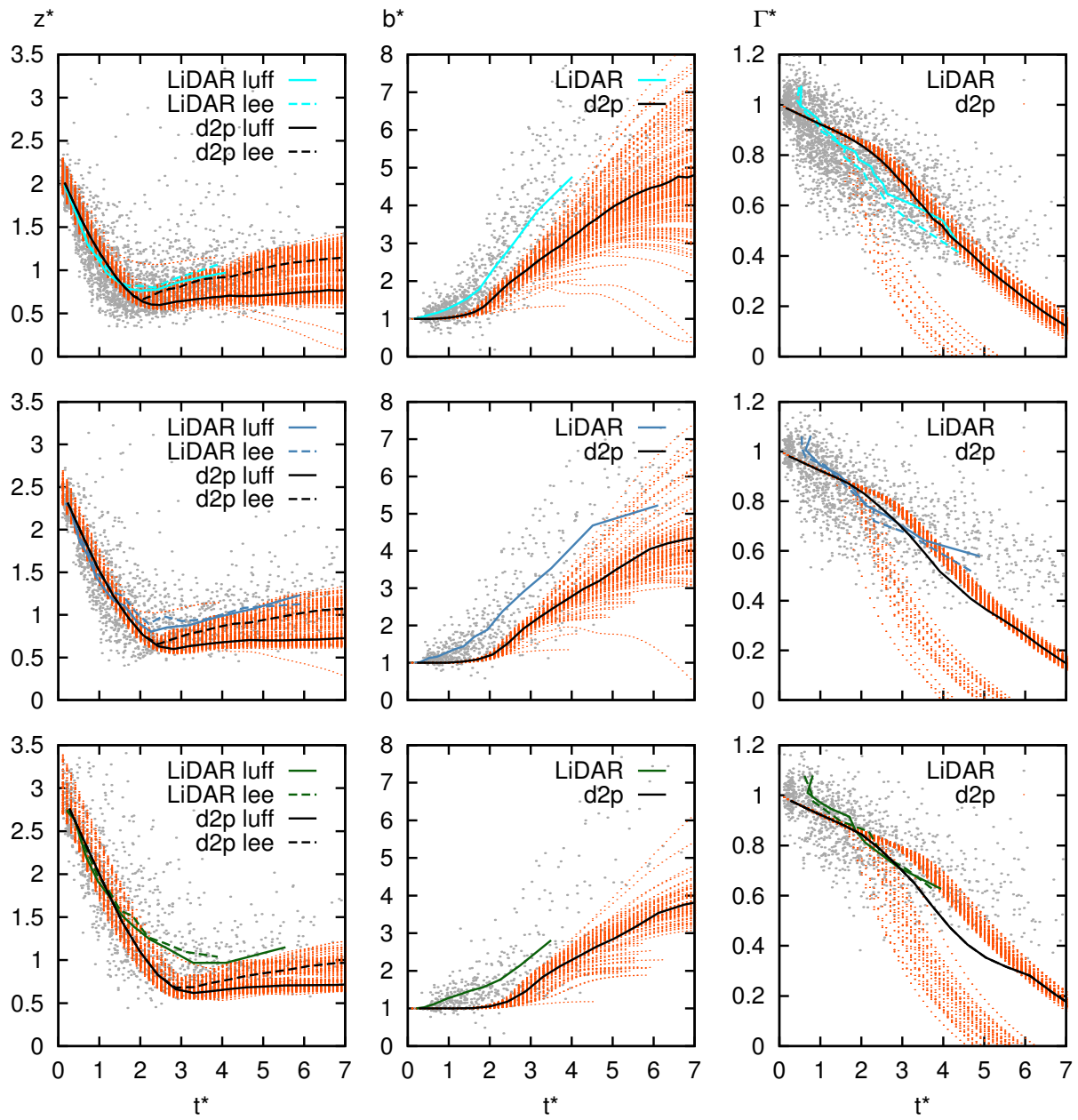


Figure 6.7.: Comparison of mean D2P prediction and the LiDAR mean for various height intervals.

6.2 Assessment of the Ensemble Performance

The BMA and REA are employed based on both a vortex age dependent and independent training approach for comparison. In the following, the subscript t^* denotes approaches with vortex age dependently treated ensemble parameters. The campaigns of NASA and DLR are assessed separately, as they differ concerning the orography, the LiDAR systems, the algorithms to calculate the circulations and the positions of the instruments to measure the meteorological conditions. The application to a combined dataset turned out to be less successful, although the ensemble still achieved at least RMSE deviations less than the second best model for all parameters.

Before the ensemble is applied the datasets are split into training and test samples. To provide realistic operational conditions the training sample contains landings from the first part of each of the campaigns, whereas the test sample consists of the second part. For the WakeMUC, WakeFRA and WakeOP dataset the ensemble is trained with 335 landings before being applied to a test sample of 359 cases. In case of the second dataset that consists of DFW95 and MEM95 data, the training sample comprises 250 landings, while the test dataset consists of 263. Figure 6.8 demonstrates that the meteorological conditions of the training and test sample of the WakeMUC/ WakeFRA/ WakeOP dataset resemble each other. In contrast, the differences are larger for the MEM95/ DFW97 dataset which is illustrated by Figure 6.9. As the NASA data does not contain headwind, the distribution of u cannot be plotted.

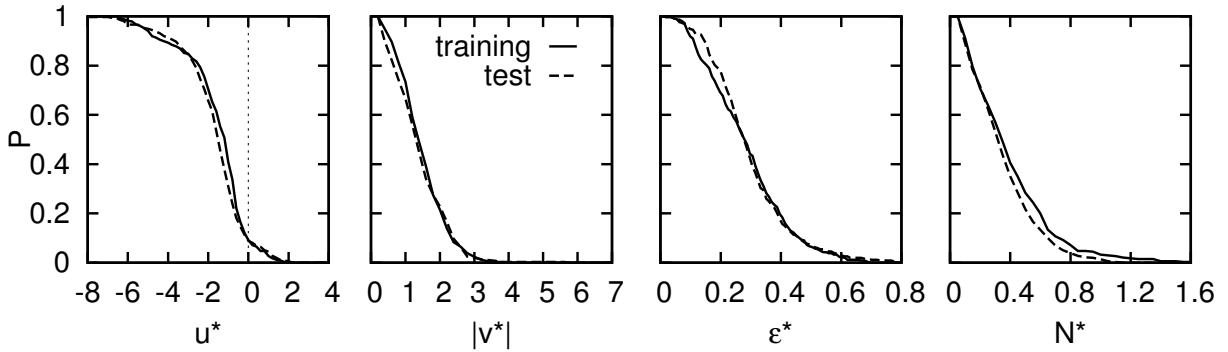


Figure 6.8.: Cumulative distribution of the meteorological conditions that the vortices have been exposed to for the training and test sample of the WakeMUC/ WakeFRA/ WakeOP dataset.

Exemplary, the forecasts of the REA_{t^*} and BMA_{t^*} and its members are depicted in Figure 6.10 and Figure 6.11, respectively, for a landing of a Boeing 747. Additionally, the 10 minute averages of the meteorological conditions are displayed. The crosswind (v^*) profile reveals that there is negative shear above an altitude of $1.3 b_0$. Furthermore, the atmosphere is weakly stably stratified while turbulence is on a medium level. While the ensemble means of the y^* and Γ^* forecasts do not deviate significantly between the REA and BMA, larger differences can be observed for the z^* -prediction.

However, the most significant disagreement between both methods is revealed by the probabilistic envelopes that are depicted for the 2σ (95.4%) of the REA and the 95% probability

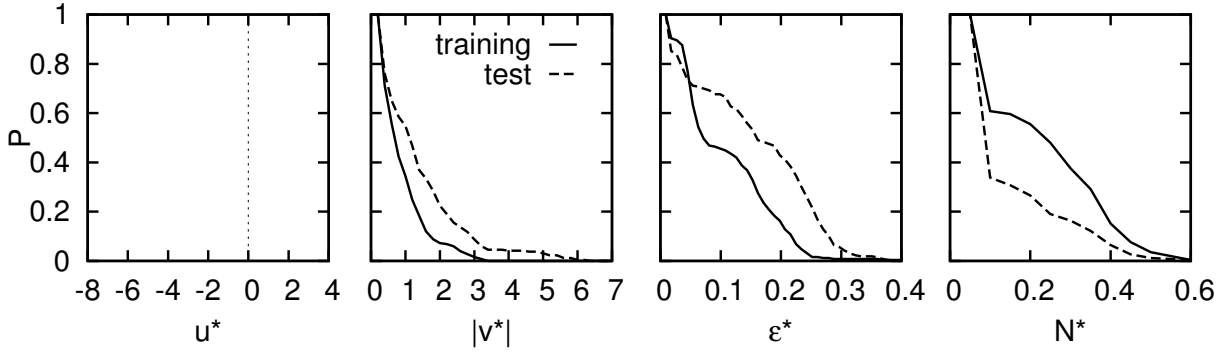


Figure 6.9.: Cumulative distribution of the meteorological conditions that the vortices have been exposed to for the training and test sample of the MEM95/ DFW97 dataset. In this dataset no headwind is provided.

level of the BMA. As the REA computes its uncertainty bounds based on model spread, the envelopes are much narrower than the BMA envelopes. For parameters where model spread is close to zero, such as for y_{lee}^* , the envelope can merely be distinguished from the ensemble mean. For the depicted landing the probability level of the REA uncertainty bounds are thus non-conservative and do not seem credible. It is peculiar that single model forecasts may lie outside of the envelope. Adding the initial condition uncertainty on top would increase the envelope width. However, this still would not address the ambient condition uncertainty. In contrast the BMA envelopes are not only composed of model uncertainty but also of initial and ambient condition uncertainty. Additionally, the stringent vortex age dependent treatment of the ensemble parameters leads to a stronger temporal increase of uncertainty. The BMA envelopes thus seem much more credible in comparison. For this reason the focus of the evaluation of the probabilistic approach in Section 6.2.2 lies on the evaluation of the BMA approach, whereas the deterministic performance is also evaluated for the DEA and REA approaches.

6.2.1 Deterministic Performance

WakeMUC, WakeFRA and WakeOP

Table 6.1 lists the median RMSE for the ensemble approaches and the individual wake vortex models when applied to the test dataset. While the DEA approach is not able to achieve smaller RMS deviations than the best model for any of the parameters, the other ensemble methods yield significantly better outcomes. Both BMA approaches are capable of achieving smaller RMS errors than any of the models for four out of six parameters. The two REA methods achieve this only for two out of six parameters. Further, the results indicate that the vortex age dependent treatment of the ensemble parameters yields on average smaller RMS deviations than the methods that rely on constant ensemble parameters. Obviously the model performance varies with vortex age. It seems that the more information about the models is added, the larger the potential benefit achieved by the ensemble.

The BMA_{t^*} achieves on average smaller errors, larger skill enhancements and smaller skill impairments than the REA_{t^*} for all parameters. The improvements by the BMA_{t^*} reach up to 11.4% (Γ_{luff}^*) and the maximum impairment does not exceed 1.9% (z_{luff}^*) (0.2% for y_{luff}^*).

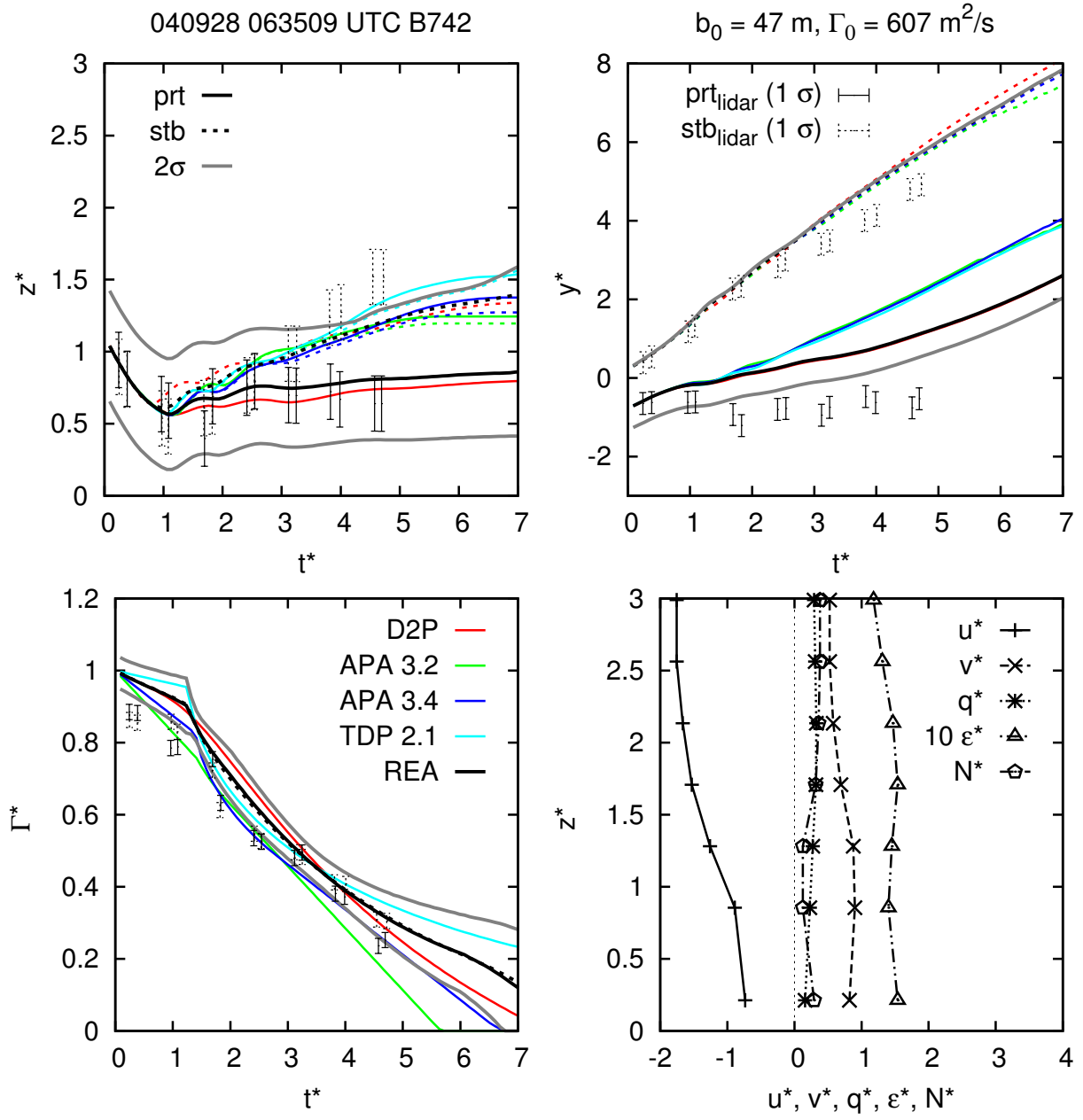


Figure 6.10.: REA_{t^*} forecast. The solid lines denote the forecast for the port vortex and the dashed lines the prediction for the starboard vortex.

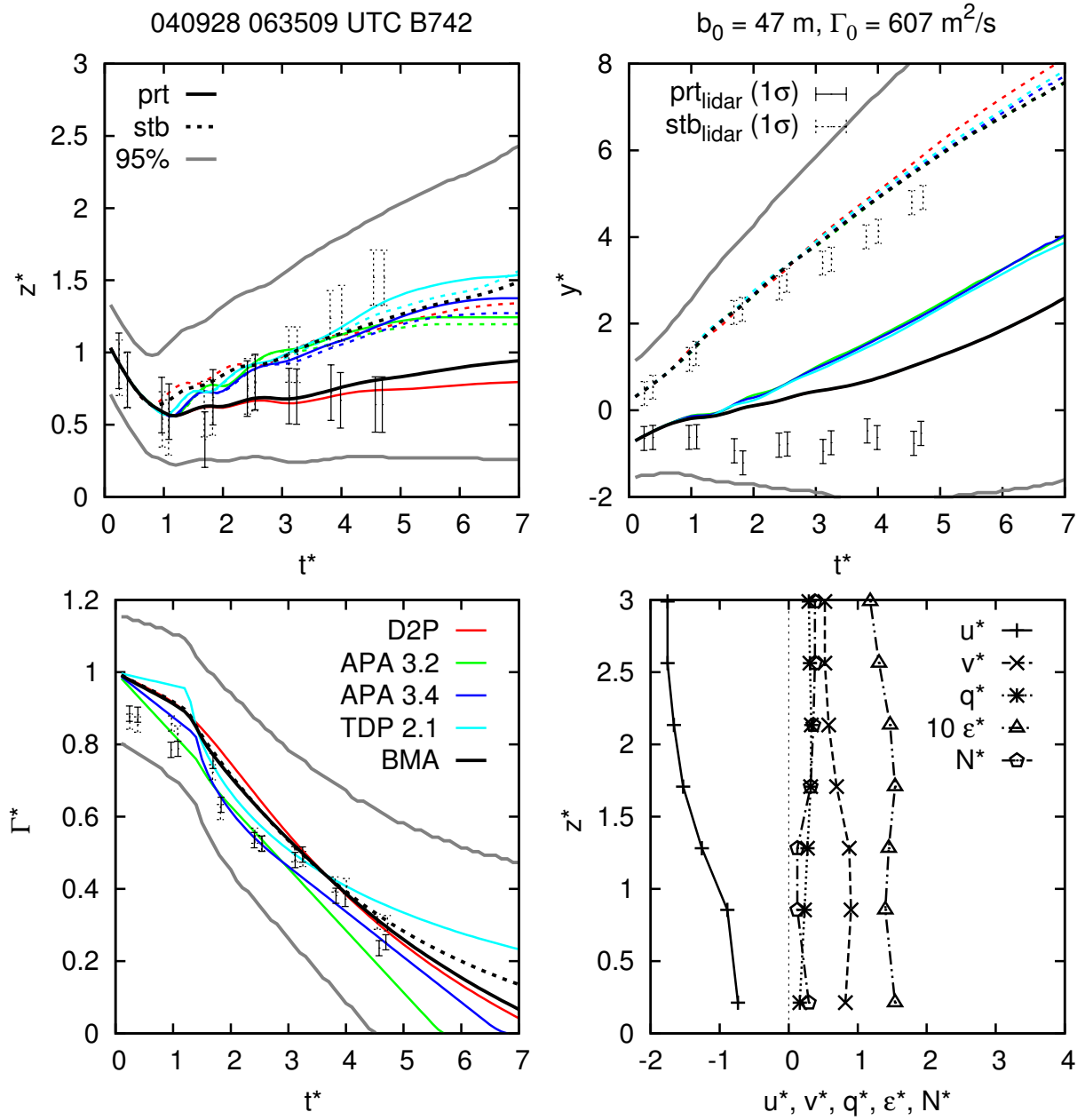


Figure 6.11.: BMA_{t^*} forecast. The solid lines denote the forecast for the port vortex and the dashed lines the prediction for the starboard vortex.

The largest improvement accomplished by the REA_{t^*} amounts to 6.14% (Γ_{luff}^*), whereas the impairment can be as high as 15.4% (z_{luff}^*). These findings reveal the superiority of the BMA approach in comparison to the DEA and REA methods in this dataset.

It is interesting that the ensemble is in general able to improve the forecast of the lateral position. The fact that the lateral transport is governed to a large part by crosswind would imply that the model forecasts do not deviate much from each other. However, it seems that in ground proximity the differences in parametrization are large enough to achieve a sufficient widening of the solution space.

Table 6.1.: Median RMS deviations between deterministic prediction and measurement data for the ensemble and its members for the WakeMUC, WakeFRA and WakeOP test sample. The smallest RMS errors are underlined.

model	RMSE Γ^*		RMSE y^*		RMSE z^*	
	luff	lee	luff	lee	luff	lee
DEA	0.132	0.116	0.752	0.518	0.200	0.191
REA	0.108	0.099	0.691	0.529	0.182	0.188
REA_{t^*}	0.107	0.101	0.640	0.520	0.180	0.186
BMA	0.110	0.100	<u>0.606</u>	0.511	0.161	0.176
BMA_{t^*}	<u>0.101</u>	<u>0.098</u>	0.608	<u>0.506</u>	0.159	<u>0.174</u>
TDP 2.1	0.116	0.105	0.790	0.552	0.241	0.192
APA 3.4	0.156	0.142	0.841	0.525	0.207	0.207
APA 3.2	0.248	0.218	0.915	0.513	0.209	0.221
D2P	0.114	0.103	<u>0.606</u>	0.523	<u>0.156</u>	0.177

Instead of further comparing the model and ensemble performance based on single parameters, the overall relative skill measure s_i is introduced. It is calculated by the relation of the average RMS deviation between the MME and model i , averaged over all parameters p . Therefore, s_i expresses to which extent the ensemble could improve the RMSE, if it competed with a single model that was chosen to predict all six parameters. Negative skill values indicate that the ensemble performs better than the respective individual model. As shown in Section 5.2, there is no clear correlation between model RMSE and the ambient conditions, such that the skill is computed independently.

$$s_i = \frac{1}{6} \sum_{p=1}^6 \frac{RMSE_{MME,p}}{RMSE_{i,p}} - 1 \quad (6.2)$$

Figure 6.12 depicts the model skill relative to the ensemble approaches for the WakeMUC/ WakeFRA/ WakeOP test dataset. Due to their superiority only the vortex-age dependent approaches as well as the DEA as baseline are presented. The results stress that among the three methods only the BMA_{t^*} is able to outperform D2P (the best model for this dataset) on average with all parameters considered. It achieves RMSE values that are on average by 3.3% smaller. The largest improvement compared to a single model is also achieved by the BMA_{t^*} with 32.4% compared to APA 3.2 alone.

When evaluating the bias, it becomes apparent that the ensemble approach is less successful

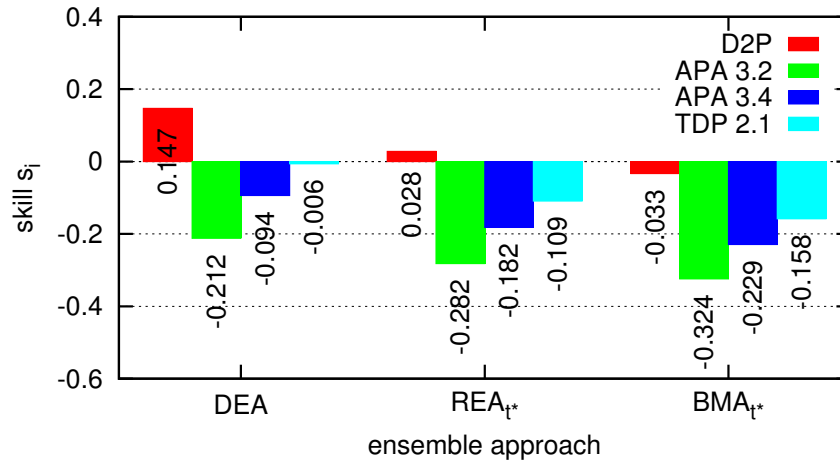


Figure 6.12.: Model skill relative to the respective ensemble approach for the WakeMUC/ WakeFRA/ WakeOP test dataset.

compared to the RMS deviations (Table 6.2). Only for the Γ_{luff}^* - and y_{luff}^* -forecasts smaller biases can be achieved by the BMA_{t^*} . It is peculiar that both REA approaches, that base their weight calculation on the model bias, are not able to yield better results. Nevertheless, both the REA and BMA approaches are capable of outperforming the DEA.

Table 6.2.: Median bias (=model-observation) for the ensemble and its members for the WakeMUC, WakeFRA and WakeOP test sample. The smallest biases are underlined.

model	bias Γ^*		bias y^*		bias z^*	
	luff	lee	luff	lee	luff	lee
DEA	-0.080	-0.066	0.079	0.102	0.092	-0.035
REA	-0.005	0.005	0.048	0.090	0.074	-0.023
REA $_{t^*}$	0.001	0.005	-0.027	<u>0.068</u>	0.067	-0.018
BMA	0.013	0.018	-0.003	0.102	0.043	-0.008
BMA $_{t^*}$	<u>0.000</u>	0.014	<u>0.001</u>	0.111	0.035	-0.004
TDP 2.1	-0.006	<u>-0.002</u>	0.104	0.151	0.146	-0.012
APA 3.4	-0.122	-0.105	0.109	0.110	0.101	-0.057
APA 3.2	-0.214	-0.182	0.133	0.105	0.091	-0.075
D2P	0.027	0.033	-0.003	0.080	<u>0.022</u>	<u>0.000</u>

To assess whether extreme errors can be reduced by the ensemble approach, the 90th percentile of the RMSE is evaluated and listed in Table 6.3. The results indicate that large errors are only considerably reduced by the BMA_{t^*} projections, where five out of six parameters exhibit smaller 90th percentile RMS deviations than the respective best individual member for the respective parameter. Compared to D2P the number is even increased to six.

Table 6.3.: 90th percentile RMSE for the ensemble and its members for the WakeMUC/ WakeFRA/ WakeOP test dataset. The smallest 90% RMS errors are underlined.

model	90% RMSE Γ^*		90% RMSE y^*		90% RMSE z^*	
	luff	lee	luff	lee	luff	lee
DEA	0.272	0.243	1.895	1.142	0.377	0.363
REA	0.209	0.191	1.814	1.211	0.360	0.353
REA _{t*}	0.210	0.188	1.811	1.188	0.360	0.351
BMA	0.206	0.183	1.505	1.170	<u>0.351</u>	0.341
BMA _{t*}	<u>0.197</u>	<u>0.178</u>	<u>1.504</u>	1.155	0.354	<u>0.331</u>
TDP 2.1	0.226	0.208	1.998	1.285	0.442	0.405
APA 3.4	0.338	0.308	2.046	1.149	0.412	0.412
APA 3.2	0.476	0.451	2.152	<u>1.147</u>	0.419	0.443
D2P	0.204	0.186	1.505	1.220	0.368	0.332

DFW97 and MEM95

Provided by NASA, this dataset comprises to a good part OGE data. Applied to these approaches the RMSE improvements by the ensemble are not as distinctive as for the DLR data, as shown in Table 6.4. Notably, the DEA and REA approaches achieve results that are closer to the BMA outcomes and in parts even better. The DEA is even able to achieve the same RMS deviation as the best member for Γ_{lee}^* and even better skill for the z_{luff}^* prediction. Further, it is rather surprising that the REA_{t*} performs better than the BMA_{t*} for z_{luff}^* and Γ_{luff}^* , given the performance results in Chapter 6.2.1. Nevertheless, the BMA_{t*} is still the best choice as it achieves again for four out of six parameters smaller RMS deviations than any of the models (one out six for the REA, three out of six for the REA_{t*} and two out of six for the BMA). Again, the approaches with temporal treatment of the ensemble parameters are superior to the methods with constant parameters. In this dataset the largest enhancement achieved by the BMA_{t*} compared to the best member amounts to 2.2% (z_{luff}^*), and the highest impairment to 2.6% (y_{lee}^*). The maximum improvement of the REA_{t*} is with 2.5% (z_{luff}^*) even a little higher, but so is its largest impairment with 3.0% (y_{lee}^*).

The skill factors s_i are once more calculated according to Equation 6.2 and illustrated by Figure 6.13. The results stress that the BMA_{t*} is again the only model to outperform all individual models on average concerning all parameters. With RMS errors that are by 1.6% smaller than the deviations achieved by APA 3.4 (the best model in this dataset) the improvements are however smaller than for the DLR data. Further, the maximum improvement compared to a single model is with 5.3% enhanced performance relative to TDP 2.1 not as significant as for the WakeMUC/ WakeFRA/ WakeOP dataset.

In accordance with the ensemble performance for the DLR dataset, the bias cannot be significantly reduced on average. The ensemble may only achieve the best bias for the z_{lee}^* -forecast (Table 6.5). Table 6.6 lists the 90th percentile RMSE deviations for the DFW97/ MEM97 test dataset. The results underline the findings from the DLR data that it is the BMA_{t*} that achieves the best performance among all ensemble approaches. For three out of six parameters it may achieve smaller 90th percentile RMS deviations than the best performing

Table 6.4.: Median RMS deviations between deterministic prediction and measurement data for the ensemble and its members for the DFW97 and MEM95 test sample. The smallest RMS errors are underlined.

model	RMSE Γ^*		RMSE y^*		RMSE z^*	
	luff	lee	luff	lee	luff	lee
DEA	0.219	0.240	0.517	0.490	0.319	0.367
REA	0.216	0.240	0.517	0.483	0.318	0.368
REA _{t*}	<u>0.208</u>	0.237	0.517	0.484	<u>0.315</u>	0.367
BMA	0.219	0.241	0.504	0.484	0.324	0.363
BMA _{t*}	0.214	<u>0.236</u>	<u>0.496</u>	0.482	0.316	<u>0.362</u>
TDP 2.1	0.238	0.252	0.520	0.494	0.331	0.373
APA 3.4	0.211	0.240	0.530	0.487	0.323	0.364
APA 3.2	0.222	0.248	0.535	<u>0.470</u>	0.325	0.364
D2P	0.234	0.245	0.506	0.484	0.328	0.372

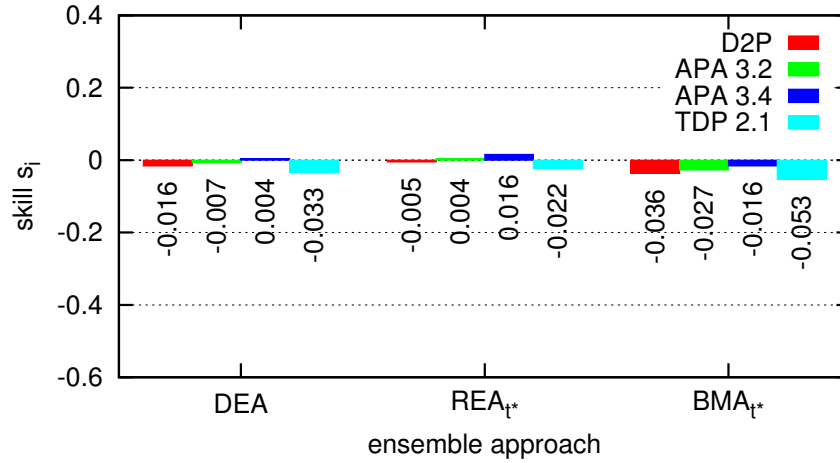


Figure 6.13.: Model skill relative to the respective ensemble approach for the DFW97/ MEM95 test dataset.

individual model. Compared to D2P it even achieves better performances for all parameters except for y_{luff} .

Statistical significance and robustness

To provide evidence that the number of simulated cases is sufficient to draw conclusions about the ensemble performance for all predicted parameters, the parameter skill $s_{i,p}$ is computed consistently according to

$$s_{i,p} = \frac{\text{RMSE}_{\text{MME},p}}{\text{RMSE}_{i,p}} - 1, \quad (6.3)$$

based on the average RMS deviations of the previous landings. Figure 6.14 depicts the skill $s_{i,p}$ against the number of already simulated cases for each model for the WakeMUC/ WakeFRA/

Table 6.5.: Median bias (=model-observation) for the ensemble and its members for the DFW97/ MEM95 test sample. The smallest biases are underlined.

model	bias Γ^*		bias y^*		bias z^*	
	luff	lee	luff	lee	luff	lee
DEA	-0.023	-0.068	-0.056	0.004	0.034	0.013
REA	-0.029	-0.077	-0.053	-0.002	0.034	0.010
REA _{t*}	-0.039	-0.086	-0.054	-0.002	0.033	0.011
BMA	-0.025	-0.053	-0.057	0.007	0.026	0.010
BMA _{t*}	-0.043	-0.067	-0.059	0.009	0.027	<u>0.009</u>
TDP 2.1	0.033	<u>-0.005</u>	<u>-0.044</u>	<u>-0.001</u>	0.039	0.023
APA 3.4	-0.046	-0.092	-0.061	0.003	0.055	0.027
APA 3.2	-0.112	-0.145	-0.051	0.013	0.063	0.020
D2P	<u>0.013</u>	-0.033	-0.051	0.002	<u>-0.009</u>	-0.011

Table 6.6.: 90th percentile RMSE for the ensemble and its members for the DFW97/ MEM95 test dataset. The smallest 90% RMSE deviations are underlined.

model	90 % RMSE Γ^*		90 % RMSE y^*		90 % RMSE z^*	
	luff	lee	luff	lee	luff	lee
DEA	0.417	0.435	1.561	1.363	0.777	0.873
REA	0.415	0.434	1.528	1.339	0.768	0.873
REA _{t*}	0.409	0.441	1.535	1.340	0.767	0.873
BMA	0.408	0.443	1.487	1.321	<u>0.766</u>	0.870
BMA _{t*}	<u>0.402</u>	<u>0.430</u>	1.489	1.319	0.767	0.864
TDP 2.1	0.419	0.432	1.609	1.397	0.799	0.876
APA 3.4	0.406	0.437	1.542	1.342	0.825	0.889
APA 3.2	0.433	0.464	1.594	<u>1.290</u>	0.833	<u>0.859</u>
D2P	0.415	0.440	<u>1.484</u>	1.337	0.802	0.875

WakeOP test sample for the BMA_{t*} method. The curves demonstrate that a dataset larger than 200 cases leads to adequately converging skills. Fluctuations cannot be totally ruled out, as the ensemble performs differently in the course of the measurement days.

To demonstrate that the above results are robust, firstly the test and training dataset are swapped and secondly totally different sets are generated from the whole dataset by mixing the cases randomly. For this study only the DLR data is employed. Figure 6.15 and Figure 6.16 both show enhanced ensemble performance compared to the baseline training. While the BMA_{t*} approach achieves on average 3.3% smaller RMS errors than the best model, it reaches 4.7% and 3.5% for the swapped and mixed datasets, respectively. Furthermore, the performance ranking of the different ensemble methods stays the same, giving proof for the robustness of the presented methods.

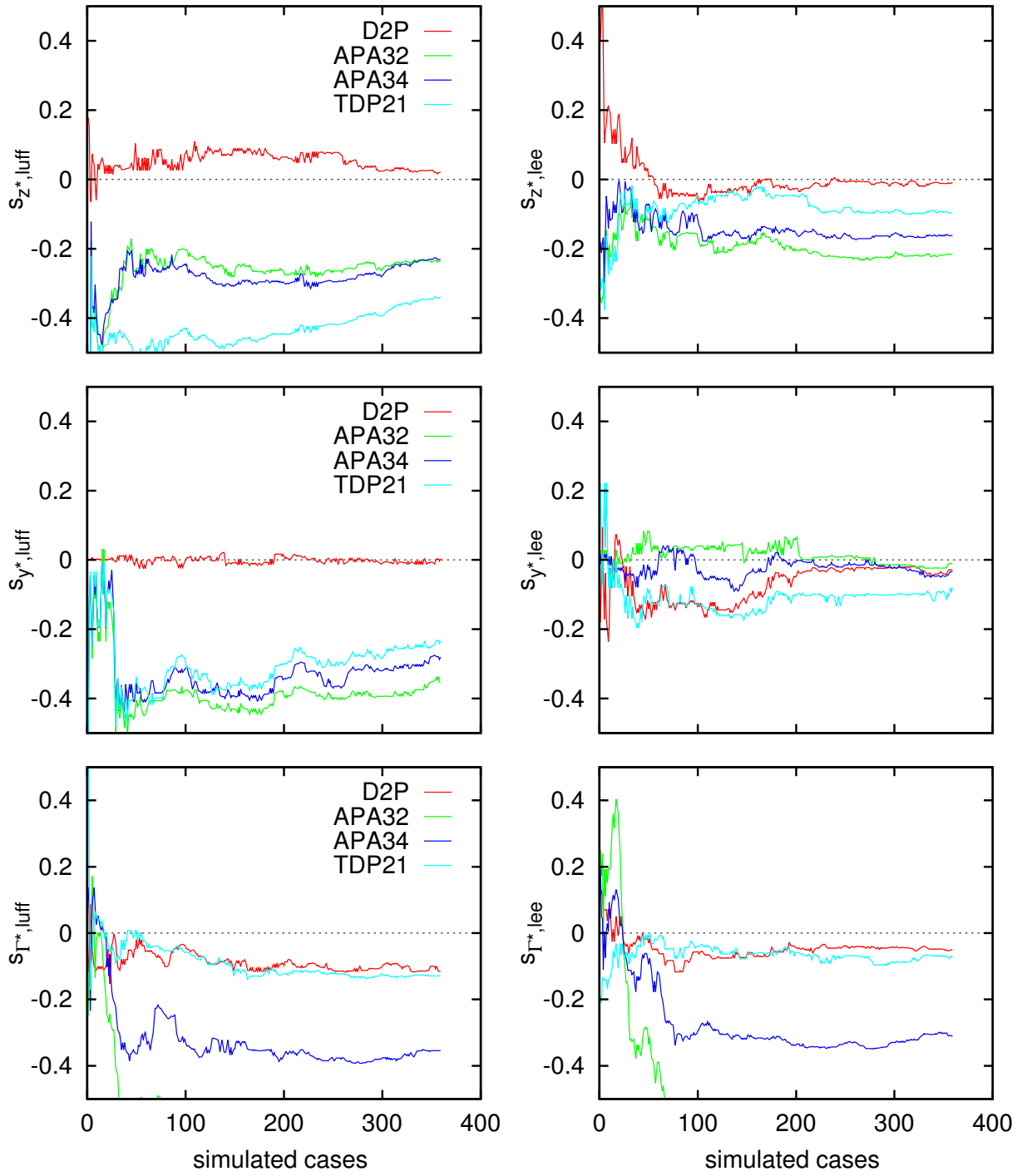


Figure 6.14.: Individual relative parameter skill, computed according to Equation 6.3, taking into account the preceding simulated landings. Illustrated exemplarily for the WakeMUC/ WakeFRA/ WakeOP test sample for the BMA_{t^*} .

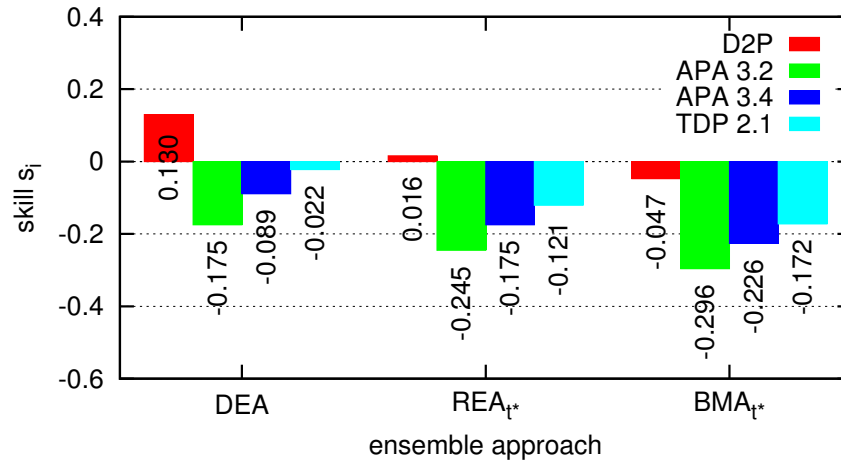


Figure 6.15.: Model skill relative to the respective ensemble approach if trained by the test and applied to the training sample for the WakeMUC/ WakeFRA/ WakeOP dataset.

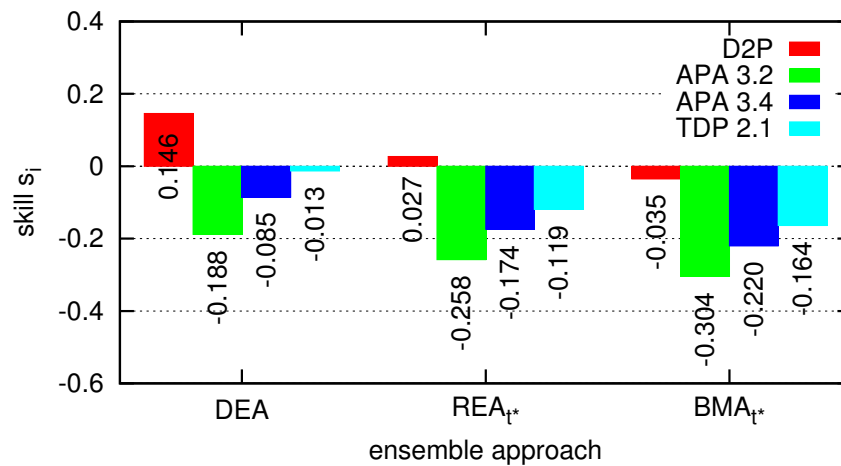


Figure 6.16.: Model skill relative to the respective ensemble approach for the WakeMUC/ WakeFRA/ WakeOP mixed test dataset.

6.2.2 Probabilistic Performance

As mentioned earlier, the REA envelopes do not seem sufficiently credible if applied to wake vortex forecasts (compare Figure 6.10 and Figure 6.11). For this reason only the probabilistic forecasts of the BMA approach are discussed in the following.

WakeMUC, WakeFRA and WakeOP

Figure 6.17 and Figure 6.18 depict a probabilistic BMA forecast for both constant and vortex age dependent ensemble parameters, separately for the luff and the lee vortex. Firstly, it can be noted that the deterministic forecasts do not deviate much between both approaches,

whereas larger differences are observed for the probabilistic envelopes. While the uncertainty allowances of the BMA_{t^*} approach are smaller than those of the BMA at early vortex ages, they are wider for later vortex ages, especially for the z^* and y^* forecasts. Secondly, for both the BMA and BMA_{t^*} the width of the uncertainty envelope of the z^* -forecast is on the order of the measurement uncertainty which indicates the good forecast quality. For the mostly crosswind driven lateral transport, however, the LiDAR measurement uncertainty is small compared to the envelope width. This underlines that the forecast uncertainty is mostly caused by inaccurate wind measurements. Likewise, the ratio of envelope width and LiDAR measurement uncertainty for the Γ^* -forecast is large and points out that uncertainties arise from the natural variability of wake vortices.

Figure 6.19 illustrates the probabilistic performance of the ensemble for the training (a) and the test sample (b) in terms of coverage of the measured vortex behavior for the BMA approach with constant ensemble parameters in a so-called reliability diagram. Accordingly, the same is depicted in Figure 6.20 for the BMA_{t^*} envelope. Perfect probabilistic prediction skill (the highest reliability) is achieved when all points are lying on the diagonal line.

The deviations from the optimum if applied to the training dataset can be accounted to a good part to the constraints under which the training algorithm has to find the optimal properties. Firstly, the employed model PDFs are the same among all models and secondly they are symmetrical. Besides, the temporal linear approximation of the model weights and the standard deviations is a simplification that may lead to inaccuracies. And thirdly, for higher probability levels fewer data points are available. Given this, it is especially hard to find a solution that accurately describes both the tails and the middle part of the distributions. Nevertheless, further calibration of the σ -values may increase the prediction quality. Note that in the original approach by Raftery et al. [71] a bias correction is applied in the training period which might further increase the probabilistic performance. Given the limited size of the test dataset it could not be guaranteed that a bias-correction would improve model performance for other data.

As wake vortex predictions are safety relevant, well-dispersed forecasts, leading to more conservative probability levels, are more desirable than overconfident ensemble predictions. If normal distributions were employed instead of fitting the kurtosis of the error distributions in advance (see Section 5.5.2), the observed coverage would increase for lower selected forecast probability levels. In turn it would decrease for the operationally relevant higher nominal forecast probability levels, reducing the safety level. The accordance between the performance for the training and the test dataset in Figure 6.19 and Figure 6.20 is an indicator whether all possible situations could be covered in the training phase.

It is remarkable that both the BMA and the BMA_{t^*} approach achieve very similar performance for both the training and test sample if the whole vortex longevity is assessed. For the training data in the left panels of Figure 6.19 and Figure 6.20, the observed coverage never lies more than 3% above or beneath the nominal forecast probability. As expected the deviations are larger if applied to the test sample. Here, the observed coverage may be as high as 7% above the nominal forecast probability, while it does not lie more than 4% beneath the optimum. A small advantage of the BMA_{t^*} is observed.

However, the benefit of the temporal approach can be emphasized by assessing late vortex ages separately, which is illustrated in Figure 6.21 for $t^* > 2$ exemplarily for the training sample.

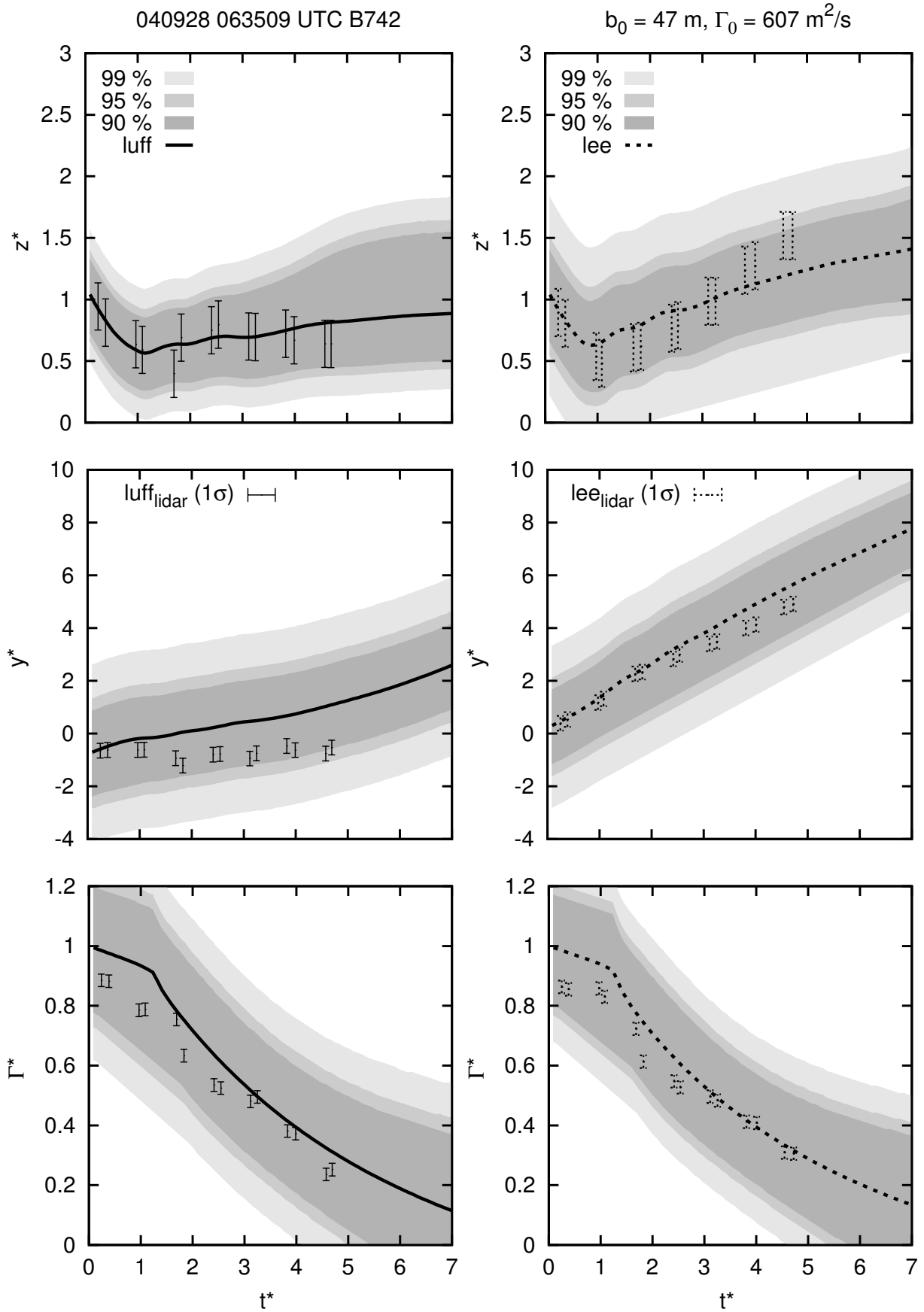


Figure 6.17.: BMA forecast with constant ensemble parameters. The predictions are shown separately for each parameter.

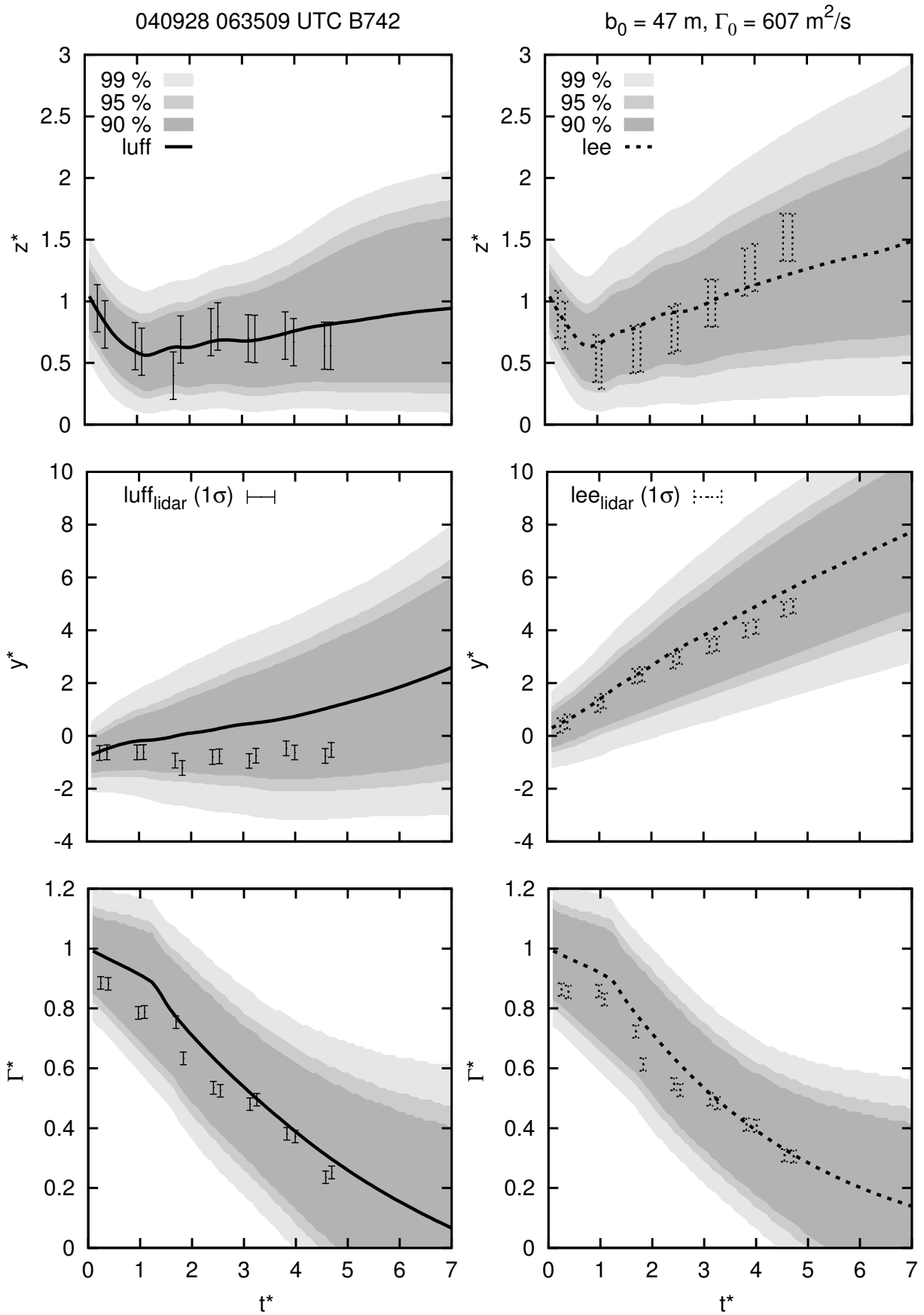


Figure 6.18.: BMA forecast with temporal dependency of σ . The predictions are shown separately for each parameter.

The BMA_{t^*} -envelope is able to achieve adequate coverage (right panel), whereas the BMA (left panel) with constant ensemble parameters is not conservative enough for all parameters but z_{luff} and Γ_{lee} . If vortex ages earlier than $t^* = 2$ are evaluated, the performance of the BMA_{t^*} does not change much while the BMA envelopes are too conservative.

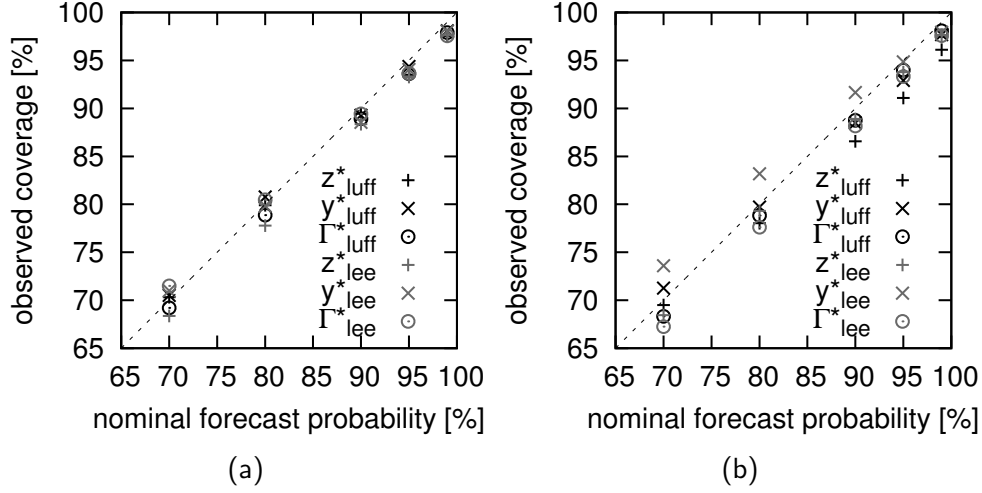


Figure 6.19.: Reliability of the probabilistic BMA envelopes (constant ensemble parameters) for the training (a) and the test dataset (b) (WakeMUC/ WakeFRA/ WakeOP).

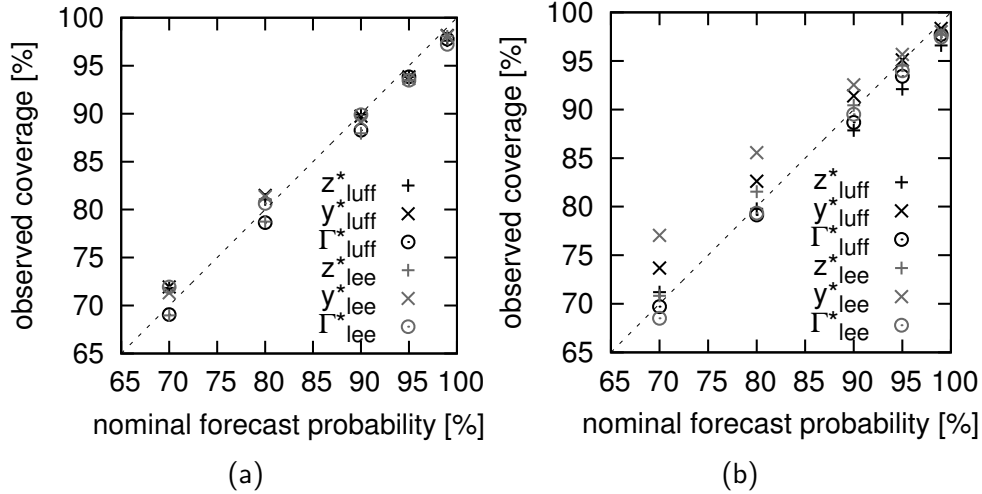


Figure 6.20.: Reliability of the probabilistic BMA_{t^*} envelopes for the training (a) and the test dataset (b) (WakeMUC/ WakeFRA/ WakeOP).

DFW97 and MEM95

The previous results demonstrated that the BMA_{t^*} approach achieves reliable probabilistic forecasts at any vortex age as it allows for the growing forecast uncertainty. For this reason only the results of the BMA_{t^*} method are shown for the DFW97/ MEM95 dataset.

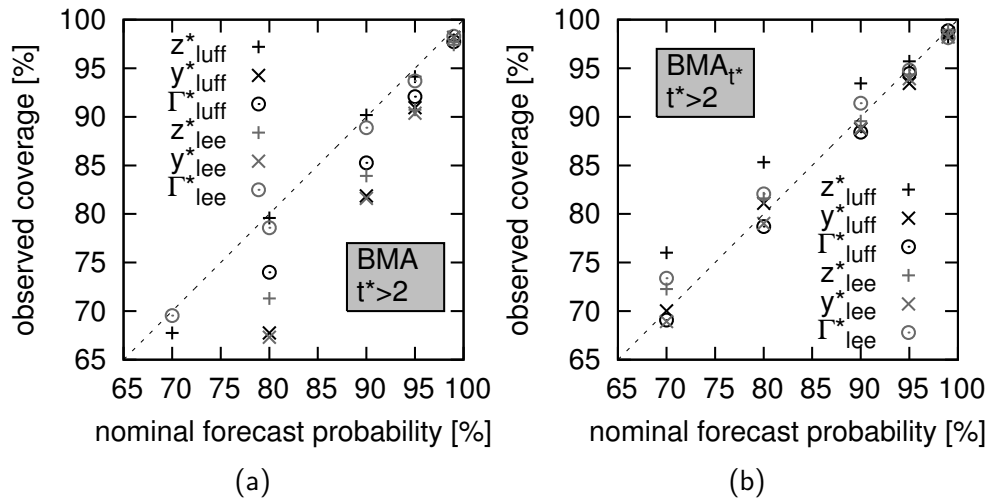


Figure 6.21.: Reliability of the probabilistic ensemble envelopes of the BMA (a) and BMA_{t^*} approaches (b) for the training sample and for $t^* > 2$ (WakeMUC/ WakeFRA/ WakeOP).

Figure 6.22 illustrates the BMA_{t^*} forecast for a single landing. The ratio of LiDAR measurement uncertainty and ensemble envelope width again reveals that the LiDAR uncertainty is on the order of the forecast uncertainty for the z^* -forecast. For y^* and Γ^* the LiDAR measurement uncertainty is considerably smaller than the ensemble uncertainty. This confirms the findings in Section 6.2.2.

Compared to the WakeMUC/ WakeFRA/ WakeOP dataset the widening of the y^* -envelope is significantly larger which indicates larger uncertainties in the wind measurements in the NASA campaigns. It is striking that the uncertainty of the Γ^* forecast is decreasing with vortex age for this dataset. Obviously the circulation data exhibits larger variations for early vortex ages. Additionally, a discrepancy between the observed (higher rates) and predicted decay rates (lower rates) can often be noted at earlier vortex ages. The Γ_{luff}^* -forecast in Figure 6.22 illustrates this observation. The conclusion that the natural vortex circulation variability decreases with vortex age would be incorrect.

Figure 6.23 depicts the probabilistic performance for both the training (a) and the test dataset (b), consisting of the NASA measurements. It is remarkable that, although DLR and NASA data differ in vortex generation altitude and quality, the achieved probabilistic results resemble one another. This indicates the qualification of the here presented training approach. While the coverage does not deviate from the optimum by more than 4% for the training data, it may be up to 10% too conservative for the test dataset. Moreover, the observed coverage never comes more than 4% below the nominal probability.

6.2.3 Spread-error correlation

In Section 5.1.3 the correlation between spread and forecast error has been theoretically discussed. Here, it is assessed whether the ensemble spread is indeed qualified to predict the

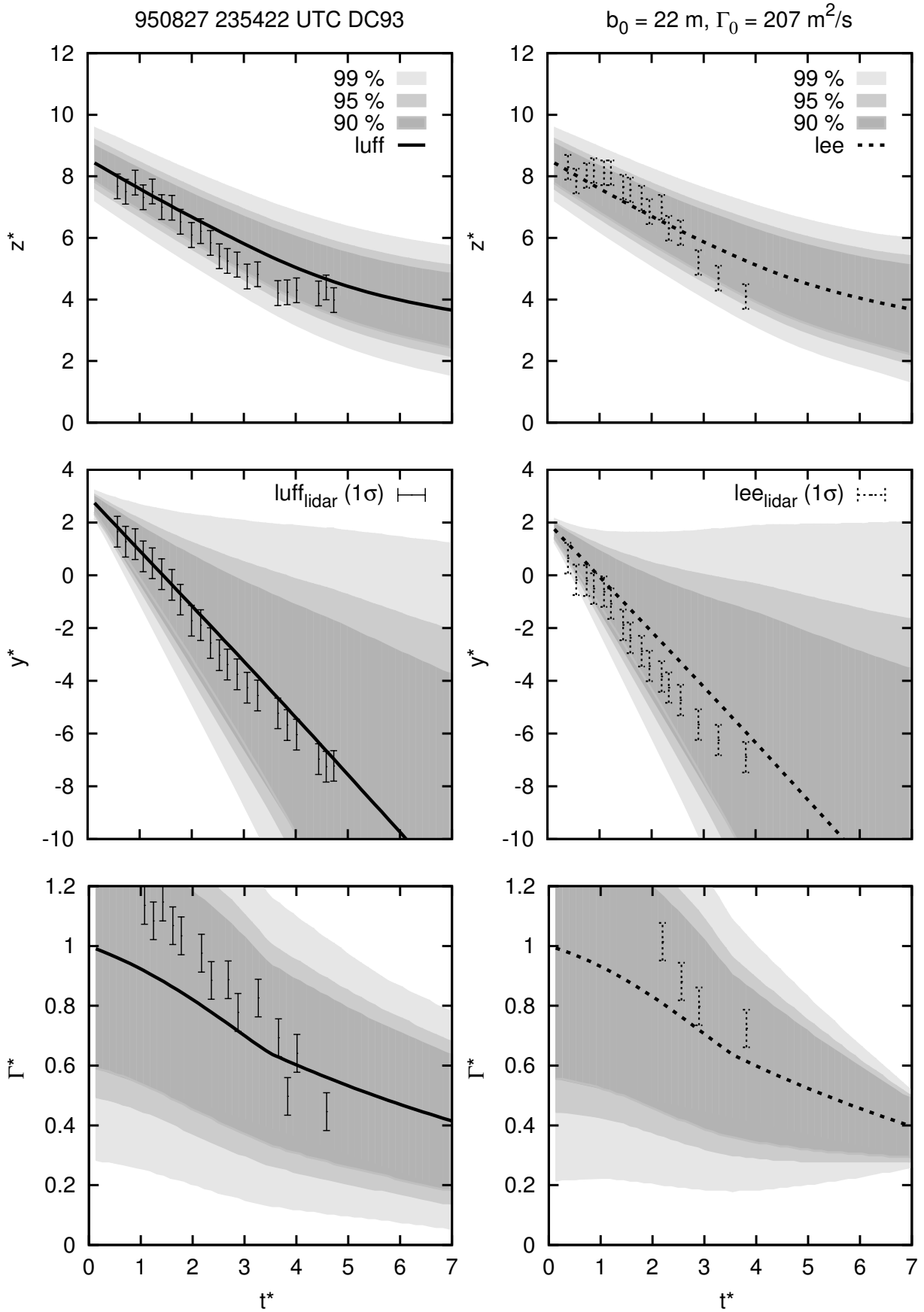


Figure 6.22.: BMA forecast with temporal dependency of σ . The predictions are given separately for each parameter.

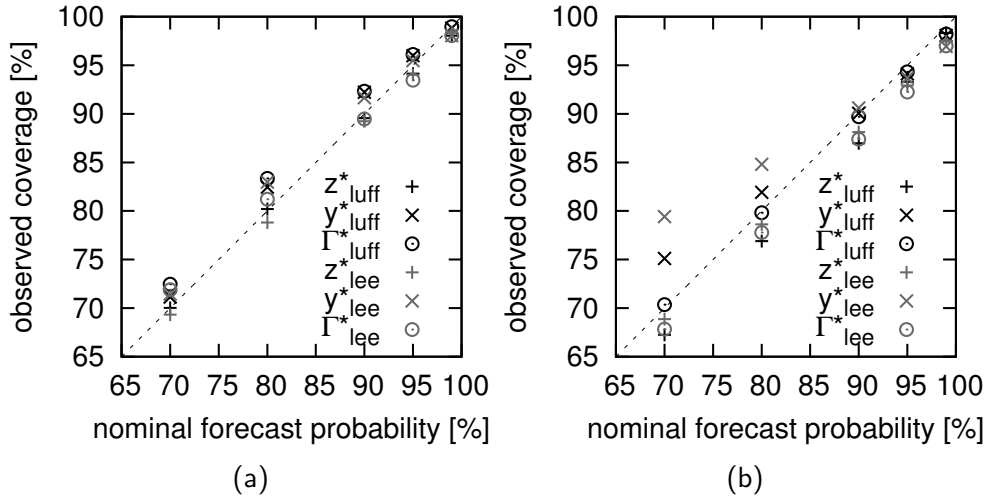


Figure 6.23.: Reliability of the BMA_{t^*} envelope for the training (a) and the test sample (b) (MEM95/ DFW97).

wake vortex forecast quality. Therefore, the weighted ensemble spread σ_{MME} of the BMA_{t^*} approach is computed according to

$$\sigma_{MME} = \sqrt{\frac{1}{n} \sum_{i=1}^I \sum_{t=1}^n w_i (f_{i,t} - \tilde{f}_t)^2}. \quad (6.4)$$

Figure 6.24 demonstrates that larger σ_{MME} results in increased RMSE for all parameters in the WakeMUC/ WakeOP/ WakeFRA test sample with a very strong relation for the z_{lee}^* and y_{lee}^* prediction. This indicates that the ensemble is well calibrated. Similar results are obtained for the DFW97/ MEM95 dataset (see Fig. 6.25) regarding the vortex position. However, contrary to the DLR data the results suggest that increasing Γ^* -spread indicates more accurate predictions. This might be attributed to the large variations in the measured vortex strength in the NASA measurements, making the ensemble spread less meaningful if compared to the LiDAR data. Conclusively, the BMA_{t^*} envelopes proved to be qualified to predict the forecast quality, with exception of Γ^* for the MEM95/ DFW97 test sample. For the latter data it can be argued that the unexpected characteristics are rather related to the measurement data, respectively the LiDAR data processing algorithm, than to the MME approach.

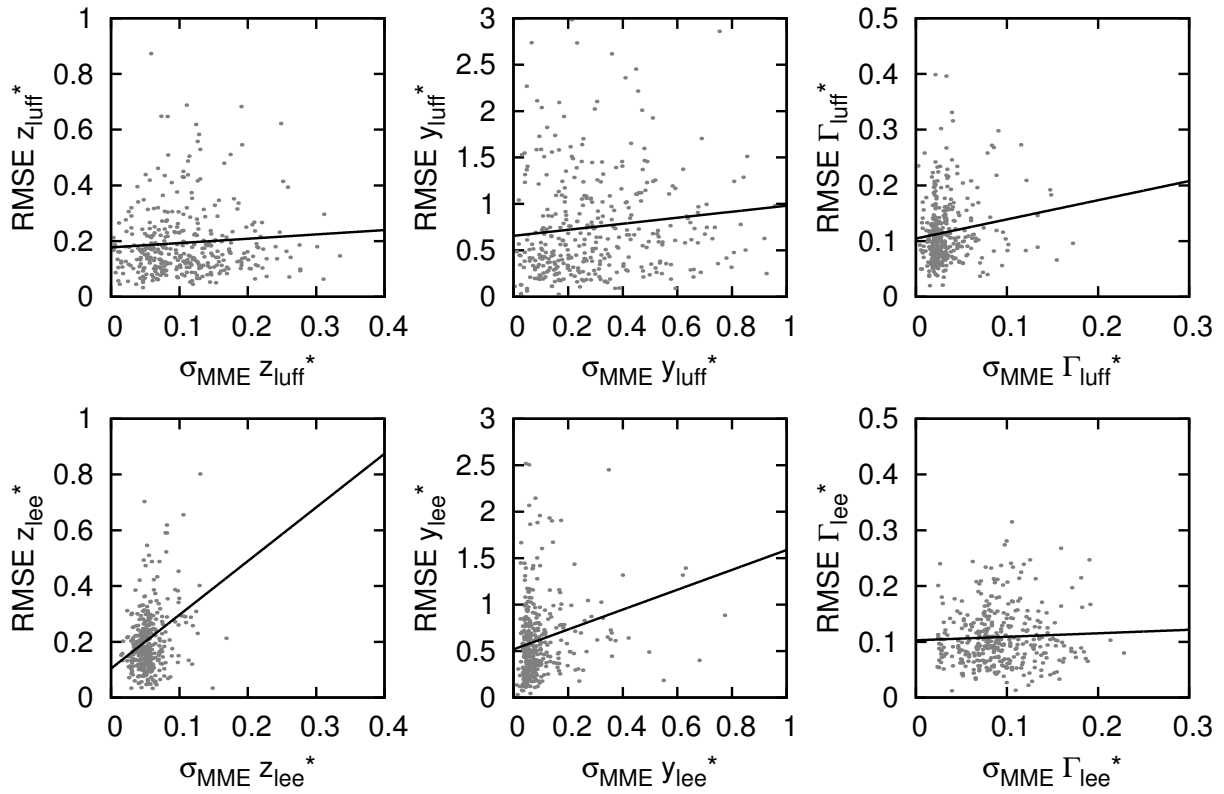


Figure 6.24.: Ensemble RMSE against ensemble spread in terms of standard deviation σ_{MME} for the WakeMUC/WakeFRA/WakeOP test sample.

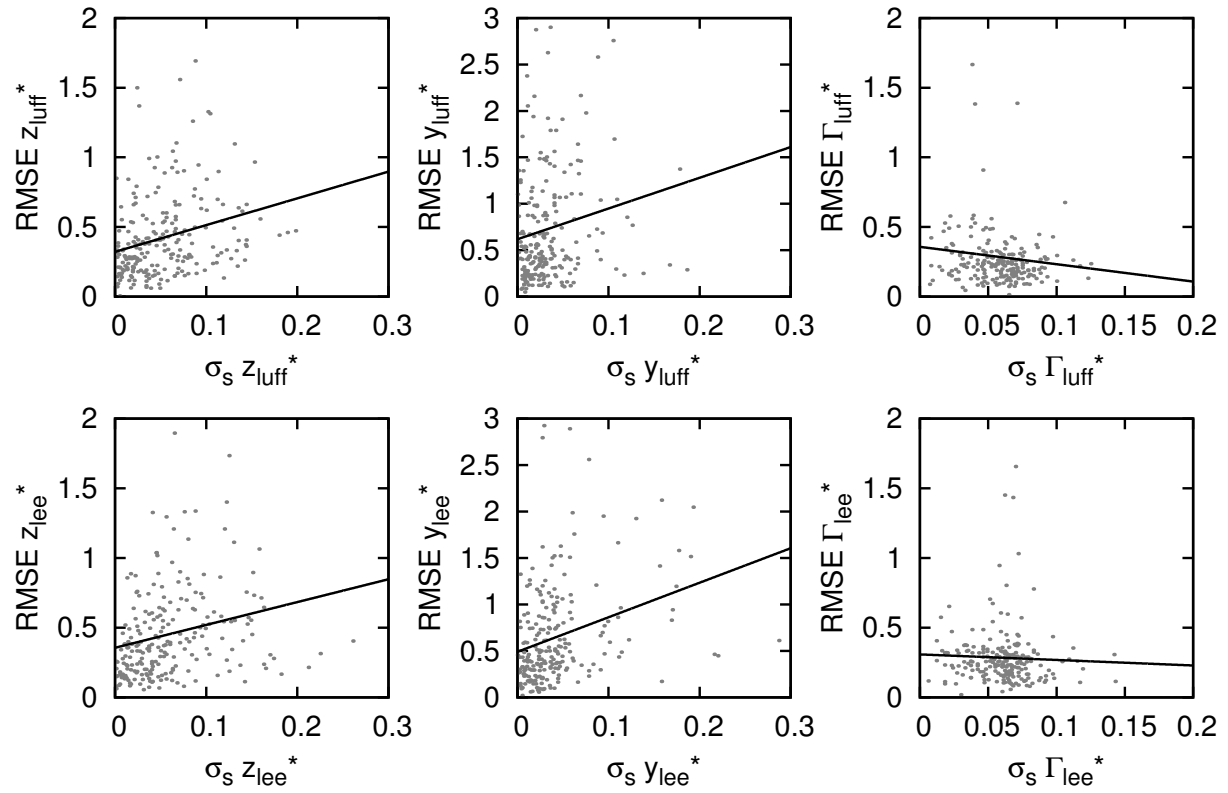


Figure 6.25.: Ensemble RMSE against ensemble spread in terms of standard deviation σ_{MME} for the DFW97/MEM95 test sample.

7. Discussion

7.1 LiDAR Measurement Evaluation

The presented evaluation of the LiDAR measurements indicates that wake vortex forecasts do not yet incorporate the interaction with the ground to its whole extent. Considering the temporal evolution of the far-field, which is a result of diffusion, might further increase the model prediction skill. Besides, the results suggest that the effect of headwind on decay of both vortices is similar to the impact of crosswind on the luff vortex. Whether this is caused by end effects or headwind-induced turbulence could not conclusively be determined. Although not as distinctive as the accelerated decay of the lee vortex that is caused by crosswind, there is a clear trend that could be implemented in model forecasts. NATS also investigated the impact of headwind on vortex dissipation in the context of time-based separation and came to the conclusion that the vortices decay more quickly in strong headwind conditions [30].

7.2 Multi-Model Ensemble Wake Vortex Prediction

7.2.1 Deterministic benefit of the ensemble approach

To compare the ensemble approaches more directly, the skill according to Equation 6.2 is now calculated only based on the best performing approach, the BMA_{t^*} , and related to both the individual models and the other ensemble approaches. In the following the BMA and REA approach with constant ensemble parameters are disregarded, as they showed to be inferior to the temporal approach in the previous chapter. The relative skill measure s_i is listed in Table 7.1. It indicates that the BMA_{t^*} ensemble outperforms the best member (in this case D2P) by 3.3% for the DLR data. Compared to the other models it is even consistently able to achieve RMSE deviations that are by more than 15% smaller. The REA_{t^*} and the DEA achieve smaller skill values. The results indicate that the REA_{t^*} is by 5.7% (0.5%) inferior to the BMA_{t^*} while the DEA achieves 15.0% (2.0%) worse RMS deviations.

For the NASA test dataset, the best model with regard to all parameters is APA 3.4. The skill measure indicates that the ensemble exhibits RMSE deviations that are by 1.6% smaller which is less than the values achieved in IGE and NGE. Nevertheless, the ensemble outperforms the worst model by 5.3%. Again the BMA_{t^*} approach outperforms the other ensemble approaches, although the differences are smaller than for the DLR data.

The results demonstrate that the ensemble provides added value for the deterministic forecast independently of which dataset is used, given that only one of the models can be chosen. Larger skill improvements seem to be achievable closer to the ground. Additionally, the measurement data quality might have an impact on the ensemble skill. Also note, that the identity

Table 7.1.: Relative skill improvement for both the DLR and NASA test sample. Negative skill stands for better performance of the BMA_{t^*} compared to the individual model (or to the REA_t^* and the DEA).

model	skill s_i	
	DLR campaigns	NASA campaigns
BMA_{t^*}	0.000	0.000
REA_{t^*}	-0,057	-0,005
DEA	-0.150	-0.020
TDP 2.1	-0.158	-0.053
APA 3.4	-0.229	-0.016
APA 3.2	-0.324	-0.027
D2P	-0.033	-0.036

of the best model is, however, not known before the ensemble is applied. Under the wrong assumption that D2P would achieve the best results for the NASA dataset, the ensemble would still achieve by 3.6% smaller RMSE deviations. It is found that mixing the DLR and NASA dataset for training and testing of the ensemble is not as effective as treating the datasets separately. This is due to the significantly different model forecast quality for both datasets, resulting in notably different model weights.

To answer the question why the potential forecast skill improvements achieved by the ensemble approach are not larger, the analysis of the natural variability in Section 5.4.1 must be compared with the RMSE deviations. This reveals that the model predictions are already on the order of the observed variability caused by gusts, stochastic vortex decay and measurement uncertainty that cannot be accounted for by the models. This leads to the conclusion that the potential of the models to be enhanced is small.

The improvements of single parameters achieved by the original BMA approach, compared to the best bias-corrected model, do not exceed 3% [71]. This implies that the presented improvements for wake vortex forecasts, given all parameters are considered at once, are on the same order as the benefits achieved for temperature and precipitation predictions [71].

Reasons for the superiority of the BMA over the REA method

The reason for the BMA method being more robust and effective than the REA approach is the basic concept of the weight calculation. Instead of calculating a-priori determined weights from an absolute mean value, it is rather derived from the number of times a model achieves the best forecast compared to the others. This allows that the weights of very similar performing models are yet distinguishable, in favor of the slightly better skilled model. In practice this means that the weights of APA 3.2 and APA 3.4 may differ substantially. However, both models exhibit very similar bias values such that their REA weights do not deviate much. Because of the similarity of their forecasts this would be equivalent to adding the forecast of one model twice, or to doubling its weight. This is the reason why APA 3.2 must be neglected for the REA z^* - and y^* -forecasts. For the BMA forecasts none of the models has to be degraded artificially.

7.2.2 Probabilistic benefit of the ensemble approach

Comparison with P2P

To demonstrate the added value of the ensemble approach, a comparison with P2P is conducted in a probabilistic manner. Note that the dynamic uncertainty allowances in P2P sum up over time which introduces vortex age dependency. However, the model calibration with training data is vortex age independent, because there was not enough measurement data available at the time when this method was developed. In contrast, the new Multi-Model Ensemble approach directly models the temporal error growth based on a comparison of the models with observations to accomplish vortex age dependent uncertainty envelopes.

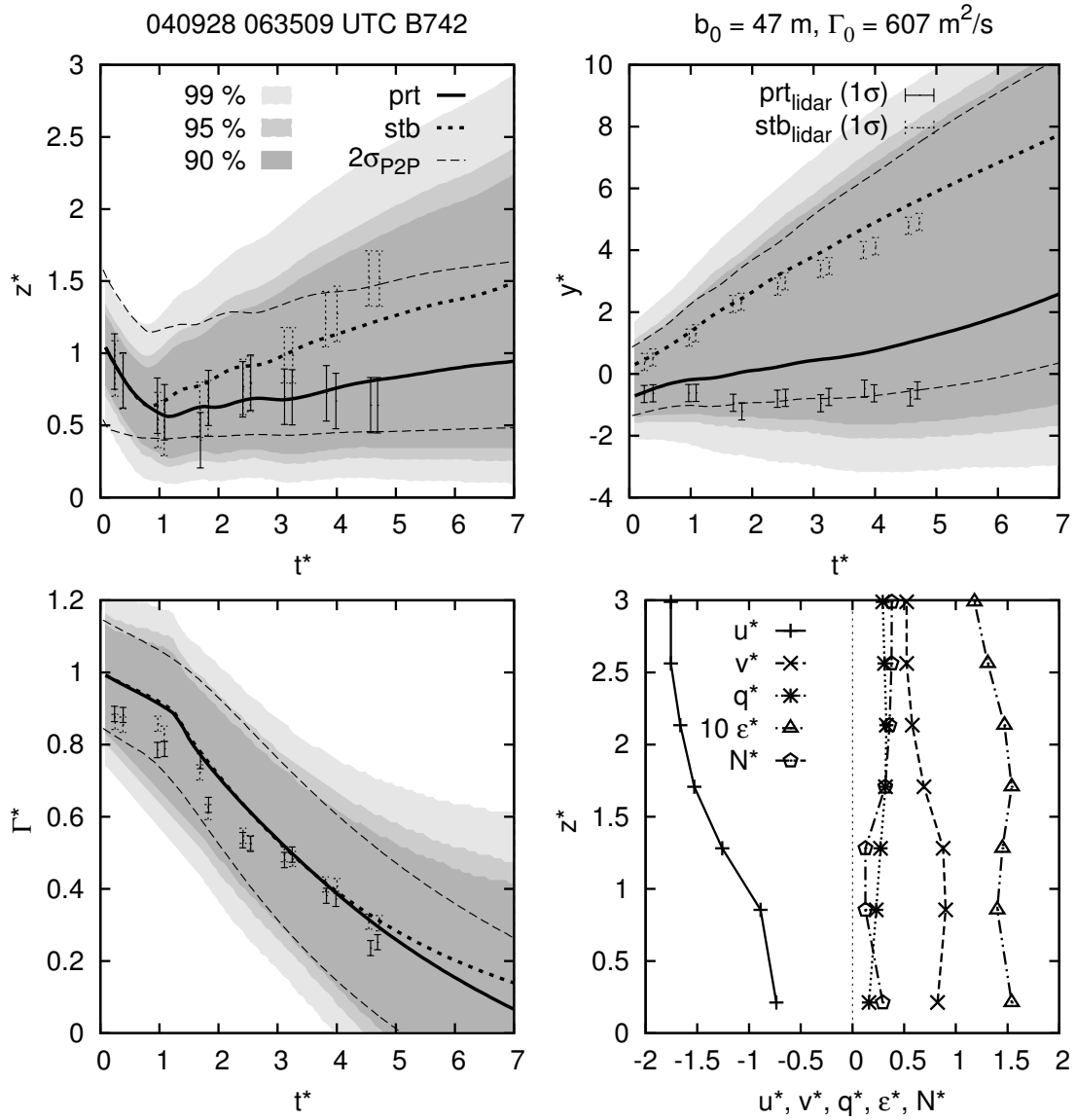


Figure 7.1.: Comparison of probabilistic P2P and BMA_{t*} forecast. The projections for luff and lee vortices are combined here.

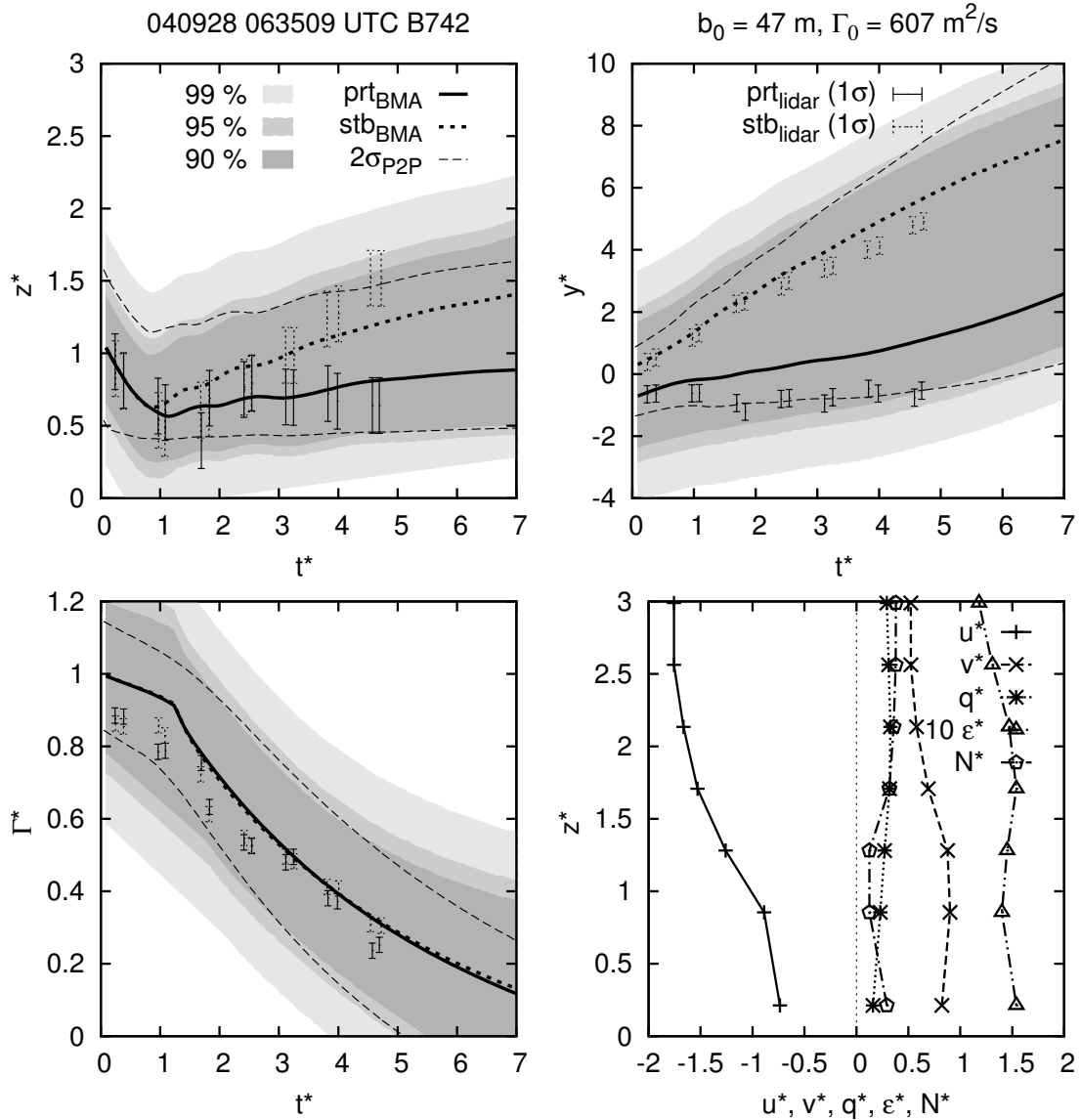


Figure 7.2.: Comparison of probabilistic P2P and BMA forecast without vortex age dependent ensemble parameters. The projections for luff and lee vortices are combined here.

To assess this approach the ensemble envelopes shall be compared with the P2P uncertainty bounds. Note that wake vortex predictions based on the MME should employ separate vortex habitation areas for port and starboard vortex in order to not to be overly conservative. On the other hand, the combination of uncertainty allowances of both vortices gives room to decrease the uncertainty allowances for the vortex pair and thus potential capacity increases of a Wake Vortex Advisory System (WVAS). In the previous section, individual envelopes for the luff and lee vortex were computed. However, P2P chooses the more conservative of both as upper and lower bound. For this reason the same is applied to the ensemble for this comparison which results in more conservative probabilistic forecasts as shown in Fig. 7.1. Additionally, the P2P envelope for the 2σ probability level is depicted. While a large difference of the

temporal development can be observed between P2P and the ensemble for the z^* forecast, the deviations are less for the y^* and Γ^* forecast. Fig. 7.2 depicts the BMA forecast for the same landing but without vortex age dependent ensemble parameters where the difference to the P2P forecast is even larger for y^* , whereas it is smaller regarding z^* .

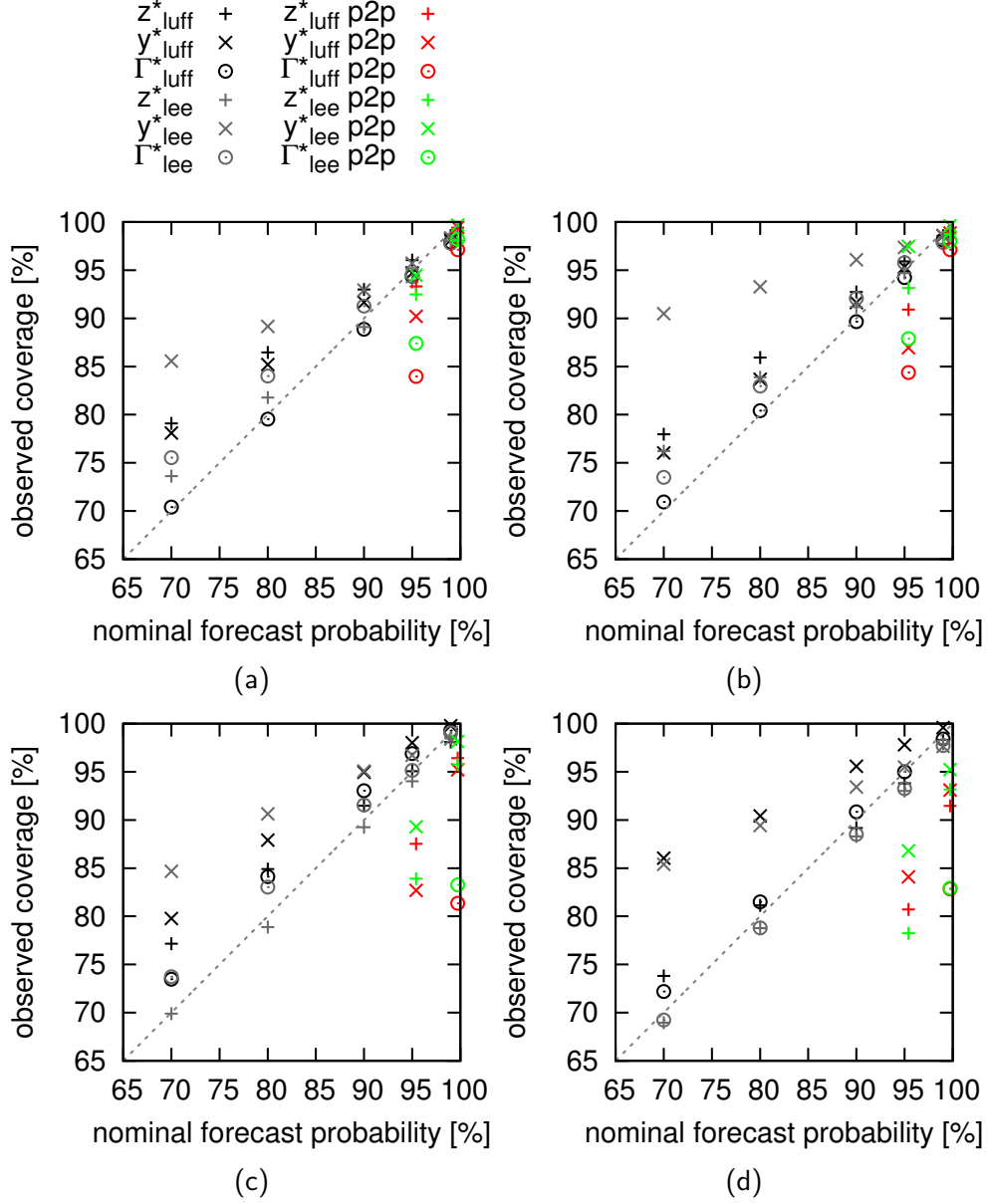


Figure 7.3.: Reliability of the conservative ensemble and P2P envelopes for the training (a) and the test sample (b) of the WakeMUC/ WakeFRA/ WakeOP dataset. Accordingly, the second row displays the performance for the training (c) and the test dataset (d) of the MEM95/ DFW97 campaigns.

Fig. 7.3 depicts the observed coverage for the training (a) and test dataset (b) of the WakeMUC, WakeFRA and WakeOP dataset as well as for the MEM95 and DFW97 samples ((c) and (d)) for the conservative uncertainty envelopes. The combined ensemble predictions

are as expected more conservative than the separate predictions for luff and lee for both the DLR and NASA data. In contrast, the P2P predictions are shifted towards the non-conservative direction. For the DLR data in the first row, especially the 2σ envelope (95.4%) deviates strongly, with the Γ^* -forecast having an offset of more than 5%. The 3σ forecast (99.73%) comes closer to the optimum but is still not able to outperform the ensemble prediction on average. This applies to both the training and the test dataset.

The same trend can be observed for the DFW97 and MEM95 dataset which is displayed in the second row. Note that the deviations of P2P are even larger here. This can be accounted to the fact that P2P has been trained with WakeToul data, where the variability of the vortices may deviate from what is observed in the NASA measurements.

The differences to the P2P forecasts must also be partly accounted to the fact that it has been trained with high quality datasets with thoughtfully derived initial conditions (partly from aircraft data and partly from LiDAR measurements). In contrast, in the here employed training dataset the first measurements are extrapolated to derive the possibly most accurate initial conditions for the ensemble members. Furthermore, the initial conditions were partly derived from theoretical values instead of employing LiDAR or aircraft data. Naturally, this results in different model uncertainty. Theoretically, if the ensemble was trained with and applied to the datasets P2P has been trained with, utilizing the same initial conditions, no difference should be notable.

Nevertheless, this analysis highlights the added value of the probabilistic BMA_{t^*} envelopes. Firstly, the superiority of the ensemble is caused by the training with similar data that is also used for the performance assessment. Note that this could also be done for the P2P model. Secondly, the ensemble envelopes are vortex age dependent. And thirdly the ensemble envelopes allow for the model uncertainty, which increases their robustness if applied to a dataset that the ensemble has not been trained with.

Performance for rare but dangerous cases

One of the questions that was intended to be answered by this thesis is whether the ensemble approach is able to cover wake vortex behavior that is seldom observed but may result in dangerous situations for following aircraft. For this reason, landings are investigated where the vortices hover in a square with a side length of b_0 , surrounding the glide path while being observed until at least $3t_0$. Among 665 landings in total, 17 cases with such dangerous vortex behavior could be identified in the WakeMUC and WakeFRA dataset. While in 13 cases a dangerous situation is correctly foreseen by the models, for 4 landings wake vortex behavior is not sufficiently predicted by the deterministic forecasts. Two different situations are distinguished: (a) the vortex rebound is stronger than expected and (b) the vortices are stronger than predicted.

Figure 7.4 and Figure 7.5 depict the ensemble forecasts for vortices with stronger rebound than the members expect. In both cases the γ^* - and Γ^* -envelopes are well-predicted, whereas the vortices partly exceed the 99% envelope for the z^* -forecast. Further, at both landings wake vortex encounters are likely if the following aircraft performs a long landing. In Figure 7.6 and Figure 7.7 decay in the models progresses faster than in the observations. This is partly due to an underestimated initial circulation but also due to a slower decay rate. In both cases all forecasted parameters are within the 99% envelopes. While in the first case a vortex

encounter is likely, a vortex would only be encountered in the second case if the flight track of the following aircraft was below the glide path. In 3 of those 4 cases the probabilistic coverage is superior to the P2P envelopes if the 2σ and the 95% envelope are compared.

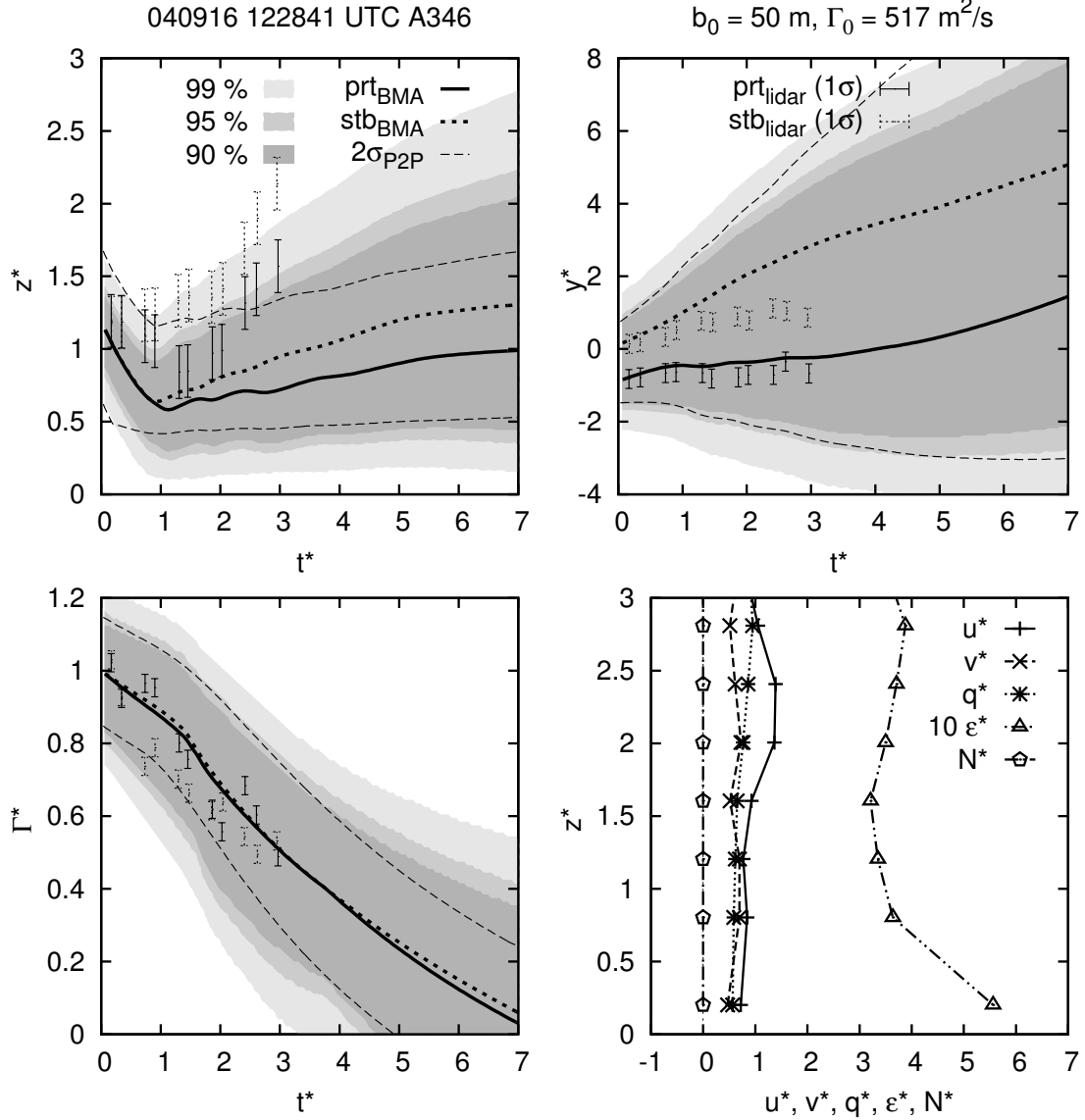


Figure 7.4.: BMA _{t^*} forecast with combined envelopes for the luff and lee vortex. The lee vortex remains close to the glide path.

Comparison with probabilistically dressed D2P

In order to identify the benefit of employing multiple models more clearly, the training is now restricted to D2P only. The weights for all other models are forced to be zero. This can be understood as a probabilistic dressing of D2P alone. In Figure 7.8 the probabilistic performances of both the ensemble and the dressed D2P model are depicted for the training (a) and the test sample (b). The deviations from the optimum (diagonal line) demonstrate

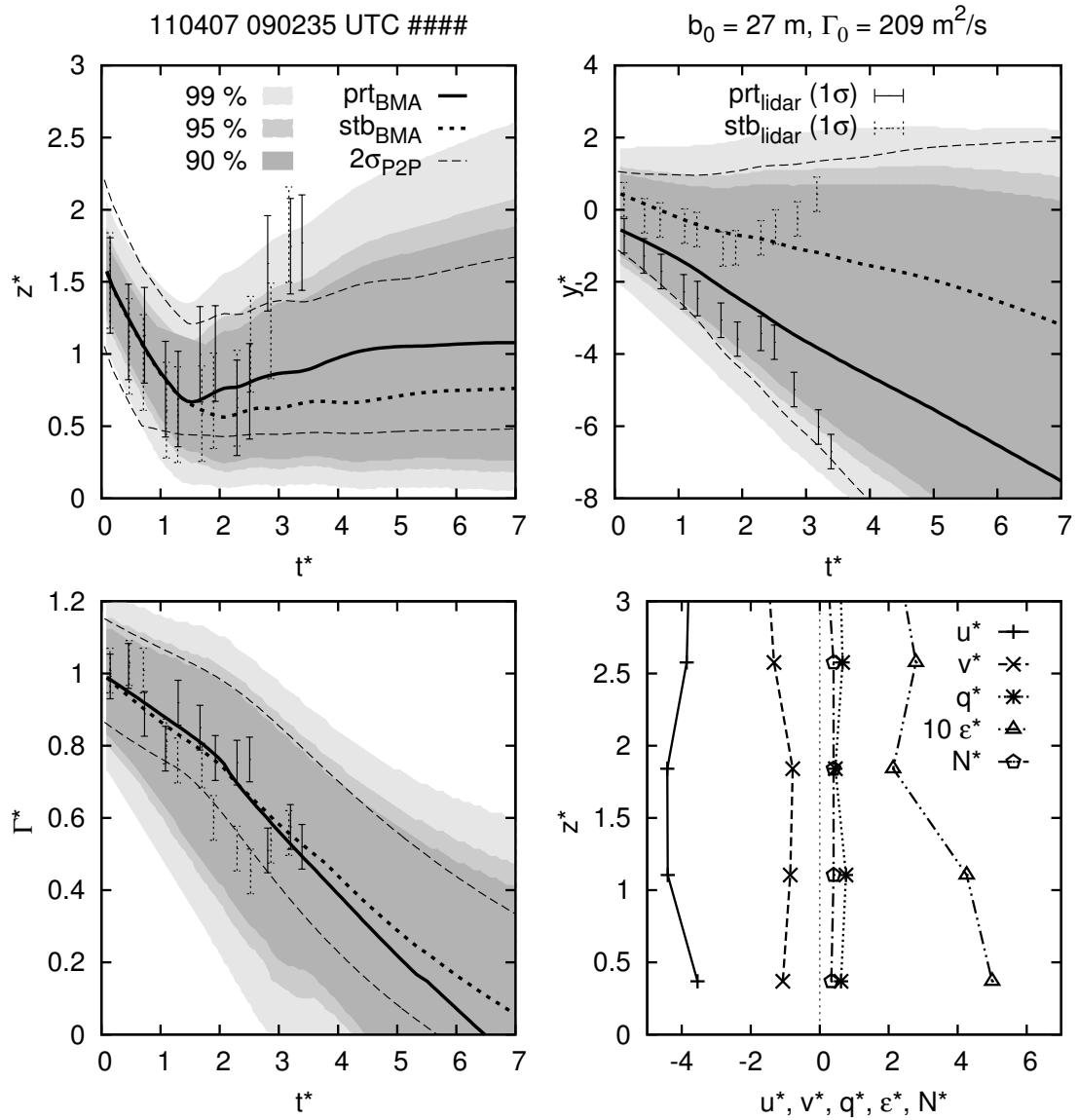


Figure 7.5.: BMA _{t^*} forecast with combined envelopes for the luff and lee vortex. The luff vortex hovers in the glide path.

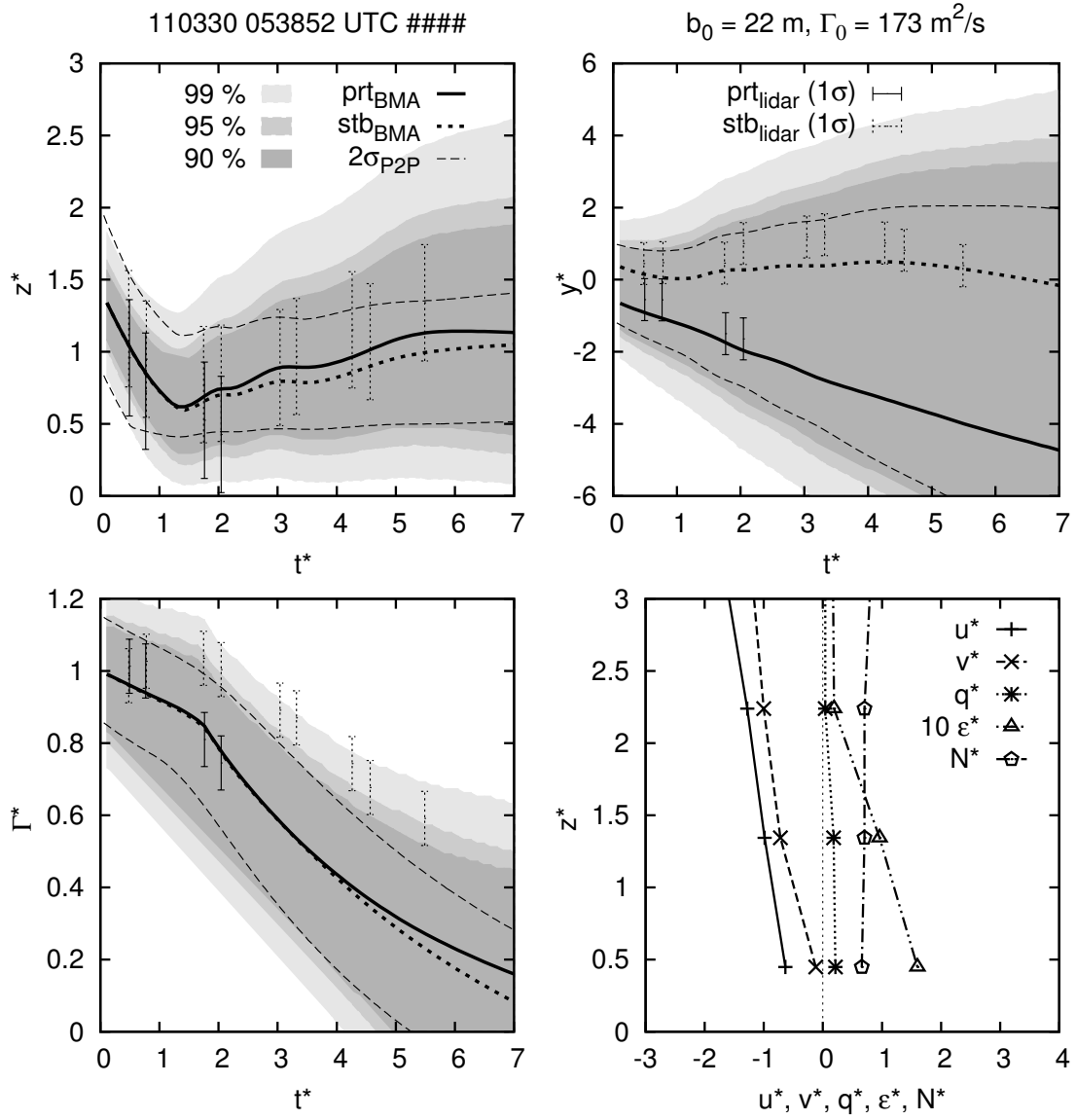


Figure 7.6.: BMA_{t^*} forecast with combined envelopes for the luff and lee vortex. The vortices remain stronger than predicted.

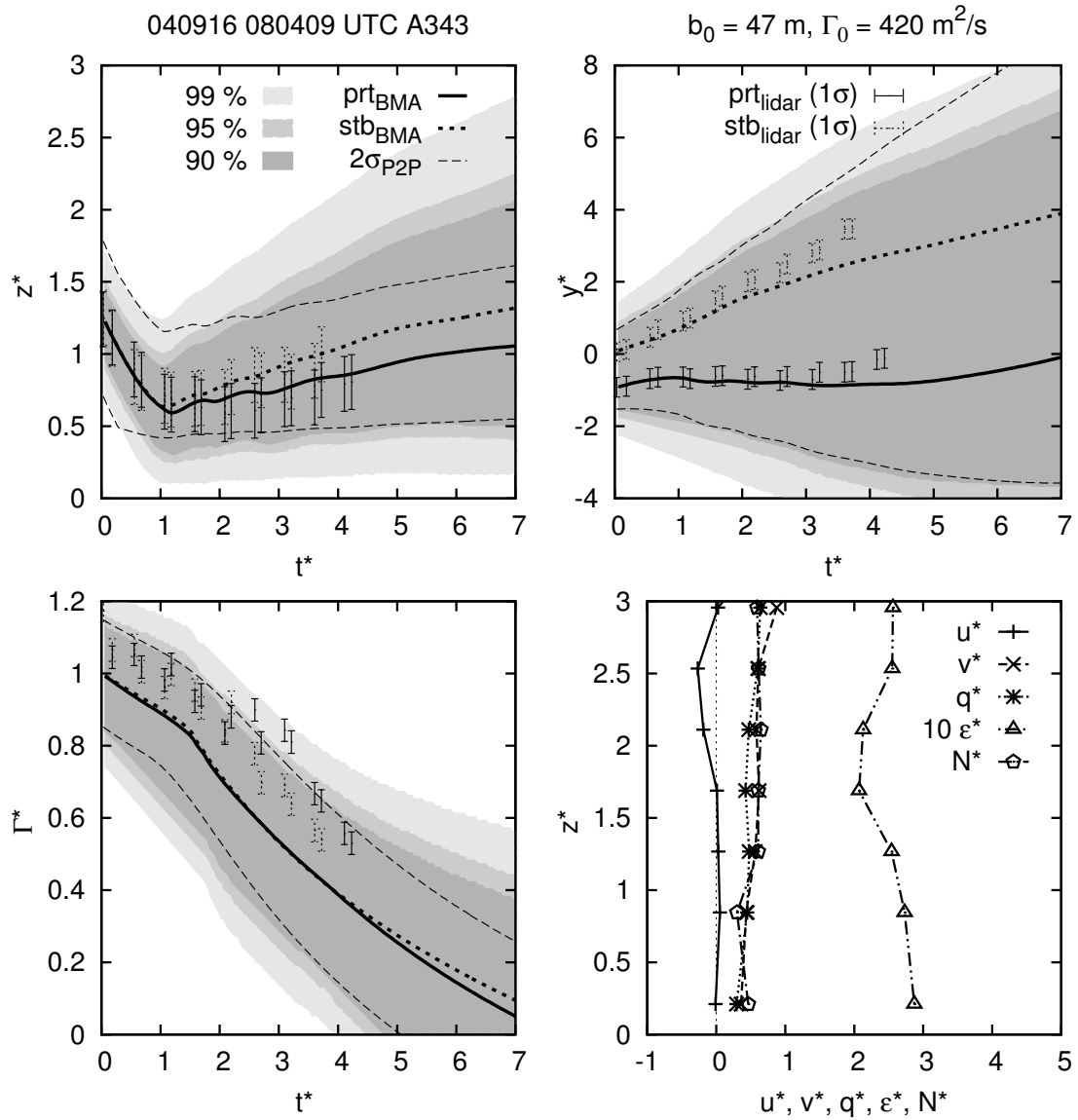


Figure 7.7.: BMA_{t^*} forecast with combined envelopes for the luff and lee vortex. The initial circulation is underestimated and the vortices remain stronger than predicted.

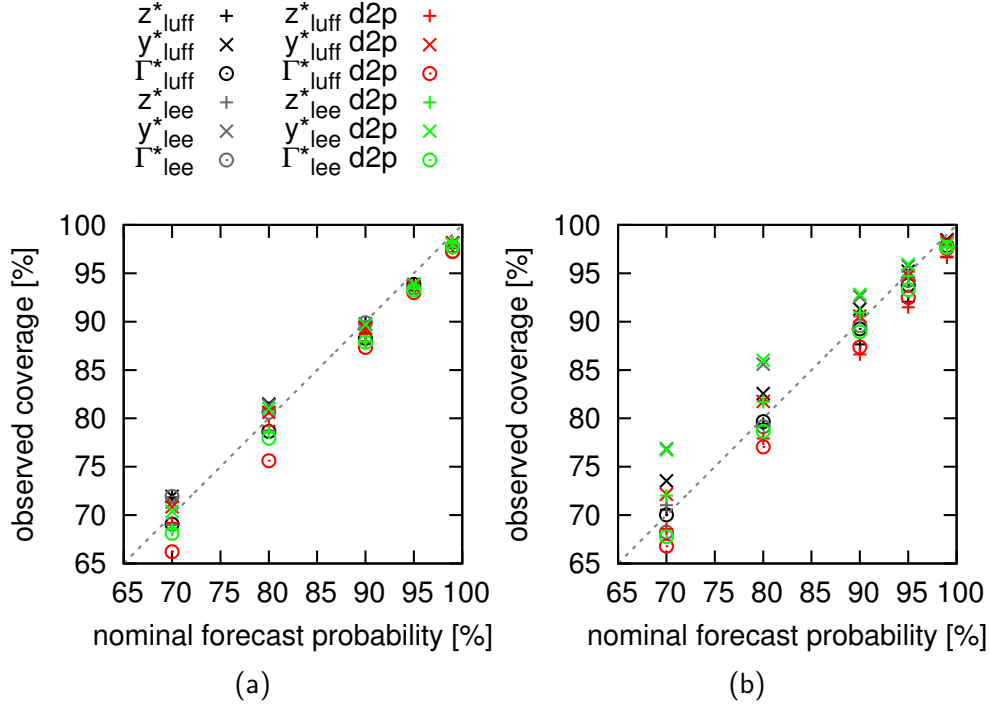


Figure 7.8.: Reliability of D2P if dressed by an envelope computed in the same vortex-age dependent manner as the ensemble envelope. Illustrated for the training (a) and the test dataset (b) of the WakeMUC/ WakeFRA/ WakeOP sample and compared to the ensemble results.

that the MME approach is advantageous for both approximating the overall model deviation distribution for the training data and for the application to the test sample. Although the dressed D2P forecasts do not depart much from the ensemble for the 99% level, the trend is towards higher deviations for lower probabilities.

Table 7.2 lists the probabilities for the 99% and 95% envelope depicted in Figure 7.8 (b) for a more thorough comparison. The outcomes reveal that the MME approach achieves for the 95% bound better coverage (closer to the nominal forecast probability) than the dressed D2P model for all parameters but z_{lee}^* . For the 99% bound the difference is less significant but yet to be observed with the dressed D2P forecast achieving better probabilistic reliability for the y_{luff}^* - and z_{luff}^* -forecast. This stresses that multiple models achieve on average better probabilistic reliability for the operationally relevant 99% and 95% uncertainty bounds than a single model, although the differences are small.

7.2.3 Skill distribution

In order to analyze how much the ensemble is able to improve single parameters in single forecasts Figure 7.9 and Figure 7.10 illustrate the cumulative distribution of the skill factors (see Equation 6.3) for the test samples for the DLR and NASA data and for each model separately. They describe the probability that values smaller than shown on the x-axis occur. On the one hand the results reveal that for the DLR data improvements of over 50% are possible (negative

Table 7.2.: Nominal and observed probability for the WakeMUC/ WakeFRA/ WakeOP test sample.

model	p_{nom} [%]	$p_{\text{obs}} \Gamma^*$ [%]		$p_{\text{obs}} y^*$ [%]		$p_{\text{obs}} z^*$ [%]	
		luff	lee	luff	lee	luff	lee
BMA _{t*}	95.0	93.8	94.2	92.1	95.8	92.1	94.6
	99.0	97.8	97.7	96.6	98.1	96.6	97.8
D2P (dressed)	95.0	92.5	93.3	91.5	95.9	91.5	94.7
	99.0	97.6	97.7	96.7	98.0	96.7	97.7

skill values). On the other hand the ensemble may also impair the forecast by 150% in single cases (positive skill values). Note that this may be related to single models with extremely small RMSE deviations such that the skill $s_{i,p}$ can take very large values. Nevertheless, it is peculiar that this can only be observed for impairments (positive values) and not for improvements (negative values). This can be noted for all models and is related to the fact that s is limited towards negative values by -1 whereas it is unlimited towards positive values. Yet the curves intersect the $P=0.5$ line at skill values lower than or close to 0, indicating that the ensemble success is based rather on many moderate than on few extreme skill improvements when compared to the respective best model.

7.2.4 Effort and benefit

With regard to the above presented evaluations, the advantages of the ensemble approach lie in the better deterministic average performance in consideration of all parameters. Improvements compared to all models can be achieved - even compared to the best performing member. As deterministic forecasts are operationally not feasible, the second focus lies on the probabilistic skill. Combining the models in a manner such that a deterministic and probabilistic benefit is achieved is connected with substantial effort. However, the presented BMA_{t*} approach, including the training, has been automated such that for future application no further effort has to be expended into manual computation of ensemble parameters. While the RMS error could be diminished by at least 3% for D2P on average, the probabilistic benefit appears to be larger. To a large part this can be accounted to the maximum-likelihood estimation of the ensemble parameters and their vortex age dependent treatment that can also be applied to a single model. Nevertheless, it becomes apparent that the multi-model combination achieves enhanced reliability compared to a single dressed model, although the superiority for the operationally relevant 95% and 99% envelopes is less than 1%.

7.2.5 Concept of operations

Implementation into the WSVBS

To implement the ensemble into the WSVBS the envelopes for z and y must be combined such that the probability of a vortex within a closed area in the z - y -space is defined. The WSVBS

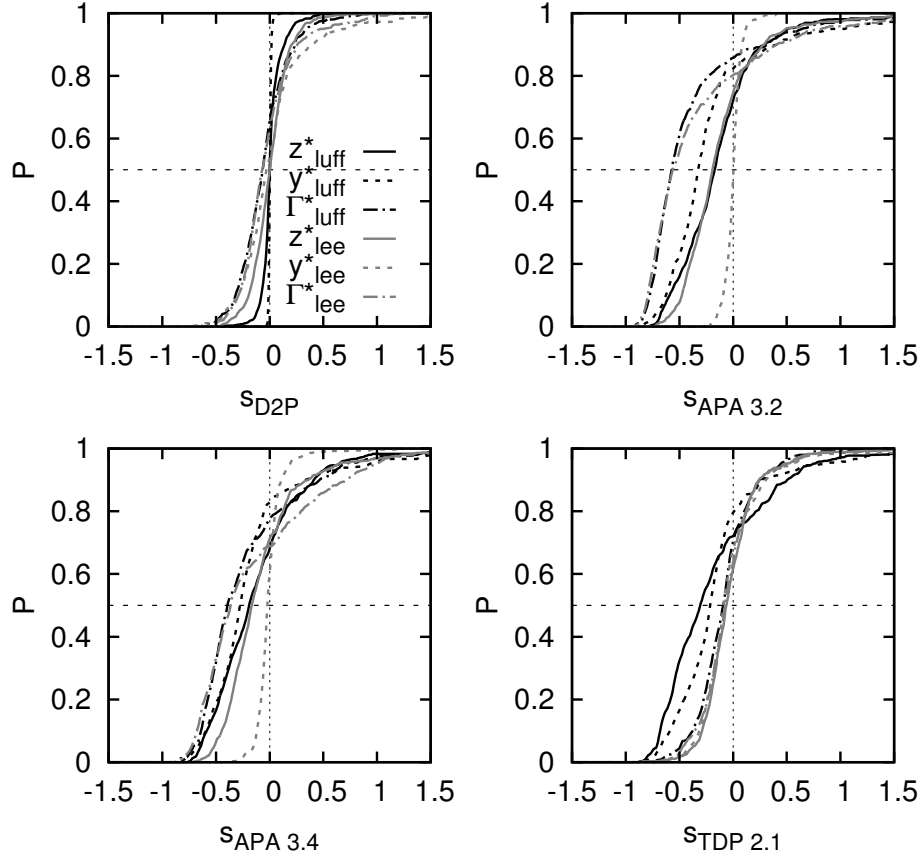


Figure 7.9.: Cumulative ensemble skill distribution (WakeMUC/ WakeFRA/ WakeOP).

computes this vortex area within 15 planes along the glide path, based on numerical weather prediction data for the 10 upper and on measurements for the lowest 5 planes. The distance between the so-called gates is reduced according to Figure 7.11 in order to properly resolve the interaction of the vortices with the ground [37]. The upper gates are separated by 1 NM.

The initial circulation of the vortices is computed individually in each gate, as the airspeed is variable along the glide path. For the approach below 3000 ft it calculates for jet and turboprop aircraft as

$$u_{ac} = C_{u_{min}} \cdot u_{stall} + ud_{DES}, \quad (7.1)$$

with the stall speed u_{stall} , the descent speed increment ud_{DES} , and the minimum speed coefficient $C_{u_{min}} = 1.3$ [160]. u_{stall} is a function of flight configuration and is obtained from the Base of Aircraft Data (BADA) [160] that is provided by Eurocontrol. ud_{DES} depends on the flight altitude and is 10 kt below 1500 ft, 20 kt below 2000 ft and 50 kt below 3000 ft [160]. For the calculation of the initial circulation two runs with different weight assumptions are performed. In the first run, the Maximum Landing Weight (MLW) is employed for maximum circulation. In the second run, the Operational Empty Weight (OEW) plus the fuel weight for one hour of flight plus the weight of 10% of the maximum amount of passengers as input is applied for the minimum circulation [161].

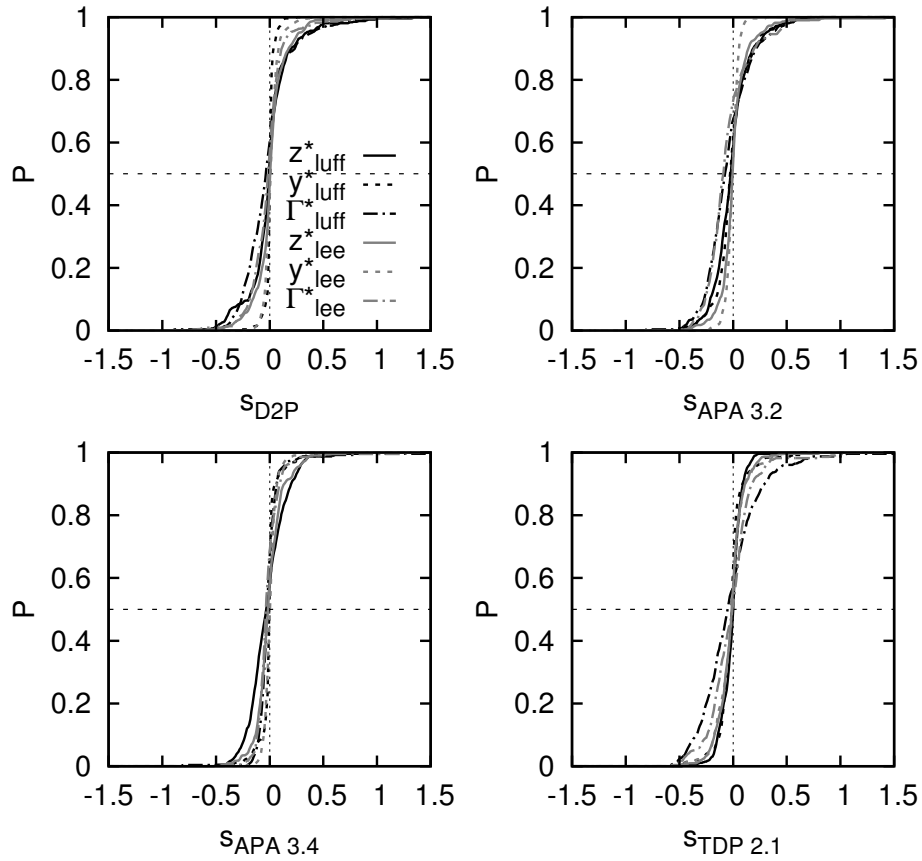


Figure 7.10.: Cumulative ensemble skill distribution (DFW97/ MEM95).

To model the advection of the vortices caused by head- or tailwind properly, the whole gate is transported according to the windspeed. This must be considered by switching off the headwind impact on vortex transport in the models (here only D2P considers headwind) when implementing the ensemble in the WSVBS. However, the computation of the lateral vortex transport is still handled by the models.

When operated with P2P the vortex area is computed using the upper and lower bounds as well as the left and right bounds of the envelopes to define the elliptic vortex area (1 or 2σ). It represents the area within any of the two vortices may be encountered under a given probability.

In a next step, the elliptical approach corridor (1 or 2σ) is added. It is derived from studies of the arrival flight tracks at the airports of St. Louis, Atlanta and Chicago below a distance of 3.3 NM and at larger distances from the flight path adherence statistics (FLIP) [161]. The corridor dimensions decrease when approaching the runway [161].

With this information SHAPE (Simplified area HAZard PrEdiction) computes simplified hazard areas based on the Roll Control Ratio (RCR) for a given aircraft pairing. The RCR is the ratio of the exerted rolling moment of the vortices and the maximum available roll control momentum achievable by aileron deflection [37, 38]. The shapes in the WSVBS represent areas with RCRs of 0.2 or more [37]. Values less than that are considered nonhazardous. The hazard areas shrink due to vortex decay [37], predicted deterministically by the SHAPE model

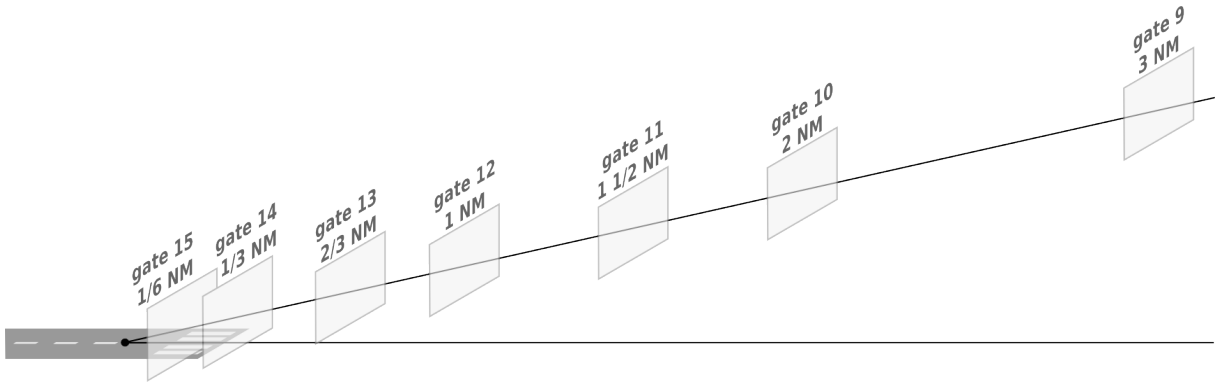
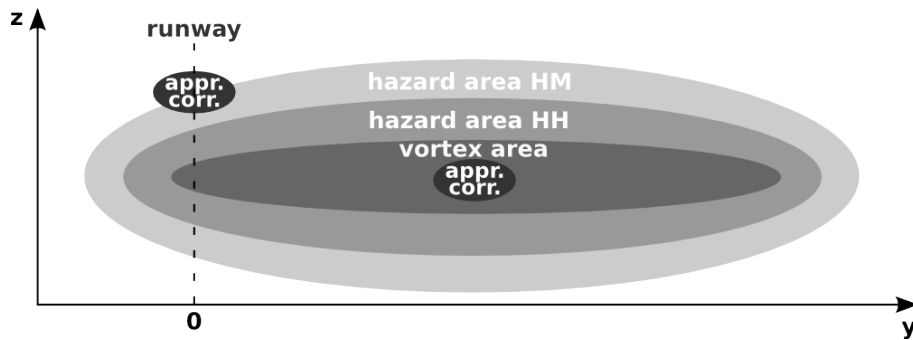


Figure 7.11.: Gate topology of the WSVBS c.f. [37]

Figure 7.12.: Exemplary hazard areas and approach corridor for *heavy* leader with *medium* follower as well as *heavy* leader and *heavy* follower [37].

for the conservative aircraft mass assumption.

The gates stay blocked until the approach corridor does not intersect the SHAPe area anymore [37]. Figure 7.12 illustrates exemplarily the hazard areas for two different aircraft pairings (*heavy* leader, *medium* follower (HM) and *heavy* leader, *heavy* follower (HH)) in a specific gate at a specific time. For the HM pairing the hazard area is still large enough to overlap with the approach corridor such that the gate is still blocked. In contrast, the area for the HH pairing does not intersect the approach corridor and the gate can be cleared. SHAPe is run with both the minimum and maximum circulation predictions and the more conservative separation time is chosen.

With the meteorological measurements conducted by a SoDAR/ RASS system and the numerical weather predictions being updated every 10 minutes, the WSVBS can predict minimum separation times for a 60 minutes horizon for planning purpose [37]. Short-term it can be employed for a 20 minutes horizon as advisory system.

Capacity and safety gain

As stressed earlier the BMA_{t^*} ensemble envelopes widen stronger with increasing vortex age than the P2P bounds, albeit being narrower right after vortex generation. Most probably this will rather contribute to larger than to smaller separations, because for the separation relevant vortex ages the envelopes of the BMA_{t^*} ensemble may be wider. Still, the shape of

the hazard areas can be formulated in a more sophisticated way by the BMA than by P2P, as multivariate PDFs can be employed in the future (see Appendix A.6). Further, the circulation can be predicted more accurately by the ensemble, affecting the temporal development of the hazard areas. Which of the discussed effects dominates must be investigated in a further study. Therefore, the WSVBS must be operated with the ensemble and applied to the traffic mix and the meteorological data of a major airport as demonstrated in Gerz et al. [39]. It can be expected that the benefit for a closely-spaced parallel runway system, where the vortices of the adjacent runway may affect the airport operations in crosswind situations, will be larger than for single runway operations [39].

Beside a potential capacity gain, wake vortex predictions may also increase safety under conditions where the current separation standards are not conservative enough. Figure 7.13 depicts a landing where the luff vortex still exhibits more than 20% of the initial circulation ($95.6 \text{ m}^2/\text{s}$) at a vortex age of 135 s. Given the standard ICAO separation rules were applied, a *medium* follower (75 m/s approach speed) would already pass the prediction plane after 123 seconds in the best case. The ensemble forecast recognizes this situation as critical and could advise the controller to increase separation rather than to decrease it.

Training

In practice the employment of the ensemble at an airport will require a preceding training period. Due to varying measurement equipment quality, different orography and traffic mix the combination of both the DLR and the NASA wake vortex measurements for training, the ensemble did not achieve significant improvements. For this reason it is suggested to train the ensemble with reliable data captured under similar conditions (traffic mix, orography, climatology) prevailing at the location of employment. It turned out that the larger the training dataset the better the distributions of the model deviations could be approximated which increased the probabilistic performance. For optimized performance in all gates along the glide path it might be beneficial to train the ensemble not only vortex-age dependently, but also vortex generation altitude dependently when implemented into the WSVBS. Further, the uncertainty of the meteorological input is different for the numerical weather predictions and the measurements. This can be taken into account by distinguishing the source of the meteorological input data in the training.

The original BMA approach incorporated a continuous training, which means that measurements are provided on a regular basis such that a fixed number of the latest verifications can be consistently applied to update the ensemble parameters. Given the case that a LiDAR and a SoDAR/RASS system are installed at the airport, with the raw data processing being fast and accurate enough, a training of similar manner could be applied to the ensemble wake vortex prediction. However, the runtime of the not yet optimized training routine would have to be reduced. As airports are not yet equipped with such instrumentation, another option would be to train the ensemble at the end of each day, week or even month.

Runtime

Due to the four employed models running serially, the runtime is naturally increased. Additionally, the BMA code itself must be executed which includes the time needed for data

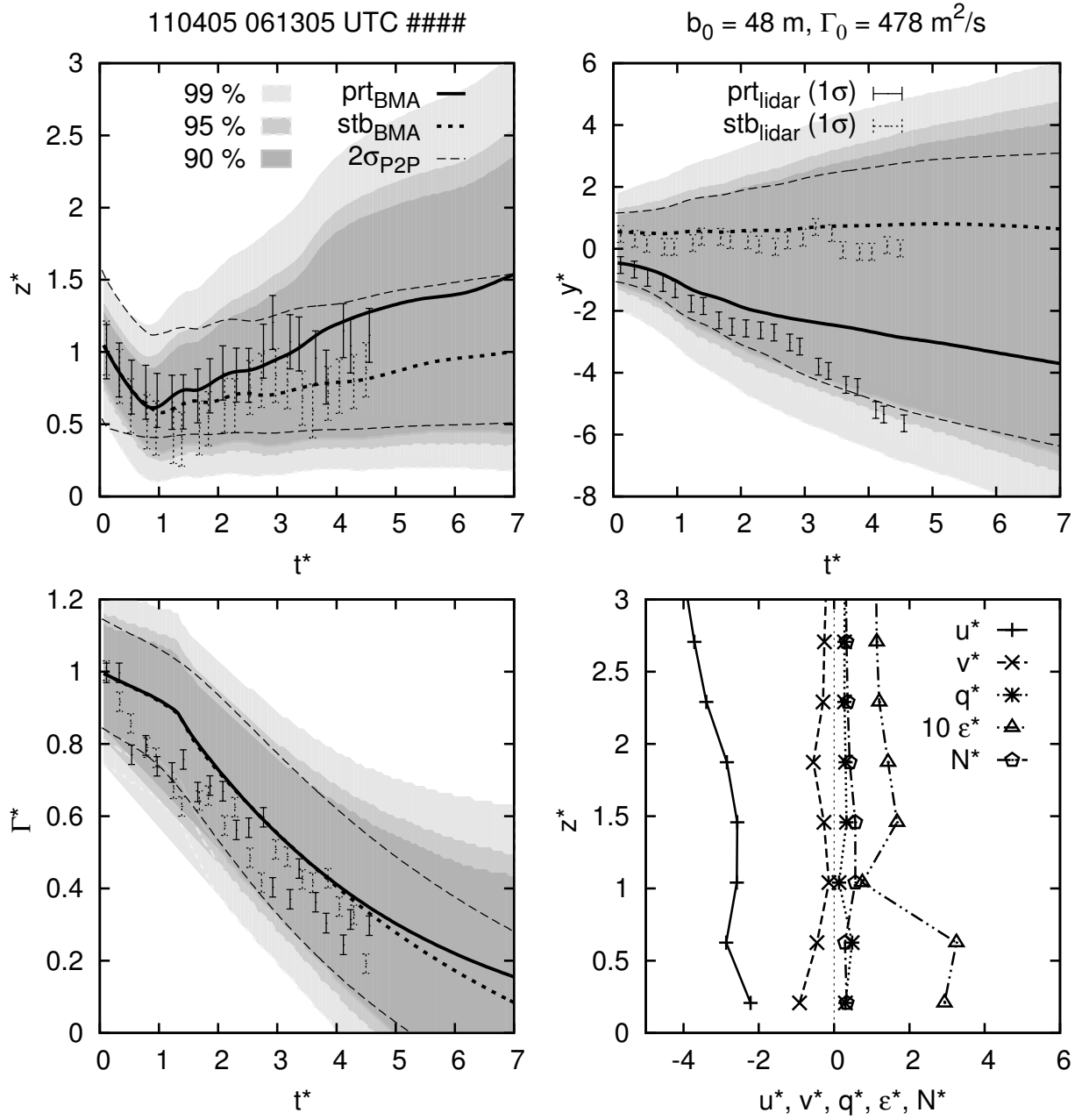


Figure 7.13.: BMA_{t^*} forecasts for a long-lived vortex that remains in the glide path until it cannot be tracked anymore after 135 s.

processing. The whole BMA ensemble run takes about 0.5 s if executed on a local Linux PC on a single core with 3.16 GHz and 3.8 GB of usable memory. Although the BMA code has been runtime optimized in parts there is still room for improvement concerning in- and output operations. Table 7.3 lists the runtime of the individual models, of the data processing and of the BMA_{t^*} code in total. It indicates that the processing of the forecasts constitutes the largest part of the total runtime rather than the individual model predictions. For lower spatial PDF resolutions the runtime can be further reduced to achieve adequate prediction times for the 15 gates in the WSVBS. Further, a run of the training routine with vortex-age dependent ensemble parameters takes between 10 and 15 minutes for the employed training samples and depends on their sizes.

Table 7.3.: Runtime of the ensemble and its members for a single WakeFRA case in IGE (averaged over 3 consecutive runs).

task	runtime [s]
TDP 2.1	0.0108
APA 3.4	0.0085
APA 3.2	0.0081
D2P	0.0112
processing	0.4587
BMA_{t^*} total	0.4973

8. Conclusion and Outlook

This thesis investigates the capability of a Multi-Model Ensemble to improve both the deterministic and probabilistic forecast of wake vortex position and circulation. The ensemble is based on wake vortex models that have been provided in the context of a NASA-DLR cooperation and comprise D2P, APA 3.2, APA 3.4 and TDP 2.1. Three different ensemble approaches, namely the Direct Ensemble Average (DEA), the Reliability Ensemble Averaging (REA) [58] and the Bayesian Model Average (BMA) [71], have been assessed. For training and evaluation of the ensemble methods, wake vortex campaigns accomplished by NASA (DFW97, MEM95) and DLR (WakeMUC, WakeFRA, WakeOP) are employed. Therefore, the training datasets and the samples for evaluation are of roughly equal size. The assessment of the individual models demonstrates that no models with outstanding expertise under certain ambient conditions are available. The only exception is D2P, which performs for the most part better for low vortex generation altitudes in the presence of wind.

While the DEA computes the ensemble average as simple arithmetic mean, the REA and BMA put effort into weighting the models according to their performance in advance. The REA employs two distinctive reliability factors for each model where the first one ($R_{B,i}$) is calculated from the model bias. The second reliability factor $R_{D,i}$ is derived from the distance of the individual forecast to the ensemble mean and is computed iteratively. The larger the deviation from the ensemble mean, the smaller $R_{D,i}$. The overall reliability factor then calculates as $R_{D,i} \cdot R_{B,i}$. Given the case that the bias or the deviation from the ensemble mean are smaller than the so-called natural variability, the reliability factors are set to one and the respective model is considered to be reliable. The observed natural variability is evaluated based on LiDAR data.

The BMA approach computes the ensemble forecast as a weighted sum of Probability Density Functions (PDFs), with the weights and standard deviations being determined by maximum likelihood estimation in a training phase [71]. Choosing the ensemble parameters directly, based on the RMSE deviations and by counting how often each model achieves the best forecast, is inferior to such an optimization method. The evaluation of the model errors suggests that the deviation distributions are rather leptokurtic than Gaussian which may be related to the fat tails found in the distributions of turbulence data [153]. By applying an additional maximum likelihood estimation, the kurtosis of the model deviation PDFs can be approximated more accurately. Further analysis of the model errors reveals that the model error grows with increasing vortex age. While the REA only allows for the temporal error growth in terms of increasing model spread, the adapted BMA approach considers the growing uncertainties directly by employing vortex age dependent standard deviations, derived in a training period.

It turns out that the BMA approach is superior to both the DEA and the REA ensemble method and that it may improve the deterministic forecast, in comparison to the respective

best model, by 3.3% if evaluated for the DLR data (IGE and NGE) and 1.6% for the NASA measurements (OGE). In direct comparison, the DEA, which constitutes the natural baseline of a MME, performs by 15.0% (2.0%) worse than the BMA approach, while the REA is by 5.7% (0.5%) inferior (results for NASA campaigns in brackets). Furthermore, it is found that the vortex-age dependent calculation of ensemble parameters is beneficial on average for both the REA and the BMA. The improvements achieved are on the same order of what has been published by Raftery et al. [71]. It is hypothesized that the reason why the possible skill increase is not larger might be the high forecast quality of the models that is already within the order of the natural vortex variability. While the vortex age dependent approach reveals a distinct benefit for the RMSE, no advantage could be found with regard to the model bias.

The assessment of the probabilistic ensemble prediction skill reveals that the uncertainty bounds provided by the REA are not as reliable in terms of forecast and observed probability as the envelopes computed by the BMA. This can be attributed to the fact that the REA envelopes consider the model uncertainty only. Adding initial condition and ambient condition uncertainties is potentially possible but complicates calibration of the envelopes. For this reason, the BMA approach, which does not only incorporate model but also initial condition and ambient condition uncertainty, is chosen to be further pursued. Further analysis of the relation between weighted ensemble spread and ensemble error of the BMA points out that large errors are more likely to occur when the model forecasts deviate strongly from one another. The spread-error correlation emphasizes the high probabilistic skill of the BMA ensemble.

The analysis of the probabilistic BMA performance suggests, that the here presented temporal training approach is beneficial. While the approach with constant ensemble parameters is only reliable as an average over the whole vortex lifespan, the temporal approach achieves consistently reliable forecasts at all vortex ages. The deviations between the selected nominal forecast probability and the observed coverage are less than 4% towards the non-conservative direction for all evaluated campaigns. If the conservative bound of the respective luff and lee vortex is chosen, this value can be decreased towards 2%.

In addition, the uncertainty envelopes outperform the probabilistic prediction skill of the P2P model, which is the probabilistic version of D2P. Two main reasons can be identified for this. Firstly, the ensemble is trained with similar data before it is applied to the datasets and is capable of adapting to the prevailing conditions. And secondly, the ensemble applies vortex age dependent uncertainties to allow for the error growth with vortex age.

The initially stated expectation to enhance the coverage of rare but dangerous cases could be reached, which is mainly related to the fact that the ensemble envelopes are wider than the P2P bounds for late vortex ages. Four landings were identified where deterministic model predictions would have been insufficient. In comparison to the P2P bounds, the ensemble envelopes achieve better coverage in those cases. By adding further ensemble members that are experts under certain ambient conditions, it is believed that the performance could be further increased.

Note that the BMA envelopes are data-derived, such that dynamic uncertainties are only generated through the varying behavior during different ambient conditions. The advantage of this approach is that the envelopes can be easily adapted to any measurement campaign. The disadvantage is that even if the initial and ambient condition uncertainty were known a-priori for each individual forecast they could not be considered.

It is stressed that dressing D2P alone by an envelope derived in the same manner as for the ensemble by maximum-likelihood yields on average slightly worse reliability for the 99% and 95% uncertainty bounds. For lower probability levels this trend is even more significant, emphasizing the benefit of employing multiple models.

Although the runtime of wake vortex forecasts increases by a large factor, it is with 0.5 s on a local Linux PC still fast. Note, that a runtime optimization has not yet been fully carried out such that there is potential for improvement. The presented methodology allows to predict wake vortex behavior probabilistically and more reliable than in the past. As the uncertainty envelopes are wider but more accurate than the P2P envelopes, especially for later vortex ages, it can be concluded that the benefit is rather the increase of safety than the increase of capacity.

Additionally, this thesis analyzes the vortex-ground interaction based on the WakeMUC campaign. Different vortex generation altitudes are available due to multiple LiDAR azimuth angles during the campaign. The evaluation points out that vortices generated at altitudes above $1.5z_0$ do not approach the ground as close as the expected altitude of $0.5z_0$. The conclusion drawn from this behavior is that the vortex-ground interaction begins already at higher altitudes as it has been assumed so far.

Furthermore, it is demonstrated that the decay of the lee vortex, where the secondary vorticity is believed to be stronger, is enhanced as it has been suggested by Holzäpfel and Steen [18]. It is further found, that the closer to the ground the vortices are generated, the faster they decay. Firstly, this can be attributed to end effects, constituting disturbances that travel along the vortex core, starting from the touchdown point. Secondly, also the increased turbulence in ground proximity may play a role.

The evaluation of the WakeMUC LiDAR measurements furthermore illuminates the influence of head- and crosswind speed on vortex descent and decay. The results suggest that the headwind influence on vortex rebound is minor compared to crosswind. However, both head- and crosswind enhance decay with increasing strength. In addition, the direct comparison of LiDAR mean and D2P mean demonstrates that the rebound for vortices generated between 0.8 and $1.5 z_0$ is well modeled, which emphasizes the success of the adaptations made to D2P concerning the strength of the secondary and tertiary vortices. For higher initial altitudes D2P overestimates the maximum vortex descent distance, which confirms the trend described above and offers the capability for further model improvement. This is in line with the underestimation of vortex divergence that increases with increasing vortex generation altitude.

For the future, a comparison with the performance of the envelopes accomplished by Monte-Carlo simulations shall be further investigated. Another interesting aspect is the implementation of the training method to single models, which has already been briefly assessed in this work. Especially the application of the EM algorithm, to approximate the model deviation distributions in a vortex-age-binned manner, may help to increase the reliability of probabilistic forecasts of single models. Additionally, it might be beneficial to combine the here presented MME approach with a Monte-Carlo simulation by perturbing the initial conditions of each single model. This way the initial condition uncertainty could be covered more thoroughly and the dependency of the envelopes from measurement data would further decrease. The latter is believed to further increase the reliability if the ensemble is applied to a dataset substantially

different from the training sample.

With regard to operational feasibility it will be necessary to adjust the probability levels for vortex pairs instead of for single vortices in order to fit to the requirements of a Wake Vortex Advisory System (WVAS). This also involves combining the envelopes of the z - and y -forecast in order to define hazard areas in flight direction. To fully assess the operational benefit of employing a MME, it will be implemented in the Wake Vortex Prediction and Monitoring System (WSVBS) of DLR and tested offline with the traffic mix of a major airport as input.

Note that the here presented models are still being further developed. This offers the prospect that in the future expert models might be available that contribute to better ensemble wake vortex predictions not only in general, but also in rare but dangerous situations at airports.

A. Appendix

A.1 Model Correlation

The tables in the following list the correlation R between the different wake vortex models, which has been introduced in Section 5.1.2 and illustrated by figures.

Table A.1.: Matrix of correlation coefficients R for z_{luff}^* .

$R z_{luff}^*$	D2P	APA 3.2	APA 3.4	TDP 2.1
D2P	1.000	0.942	0.944	0.900
APA 3.2	0.942	1.000	0.986	0.935
APA 3.4	0.944	0.986	1.000	0.957
TDP 2.1	0.900	0.935	0.957	1.000

Table A.2.: Matrix of correlation coefficients R for z_{lee}^* .

$R z_{lee}^*$	D2P	APA 3.2	APA 3.4	TDP 2.1
D2P	1.000	0.916	0.928	0.943
APA 3.2	0.916	1.000	0.984	0.937
APA 3.4	0.928	0.984	1.000	0.960
TDP 2.1	0.943	0.937	0.960	1.000

Table A.3.: Matrix of correlation coefficients R for y_{luff}^* .

$R y_{luff}^*$	D2P	APA 3.2	APA 3.4	TDP 2.1
D2P	1.000	0.986	0.988	0.976
APA 3.2	0.986	1.000	0.998	0.989
APA 3.4	0.988	0.998	1.000	0.993
TDP 2.1	0.976	0.989	0.993	1.000

Table A.4.: Matrix of correlation coefficients R for y_{lee}^* .

$R y_{lee}^*$	D2P	APA 3.2	APA 3.4	TDP 2.1
D2P	1.000	0.995	0.996	0.996
APA 3.2	0.995	1.000	0.998	0.993
APA 3.4	0.996	0.998	1.000	0.995
TDP 2.1	0.996	0.993	0.995	1.000

Table A.5.: Matrix of correlation coefficients R for Γ^* (no distinction of luff and lee).

Γ^*	D2P	APA 3.2	APA 3.4	TDP 2.1
D2P	1.000	0.886	0.938	0.950
APA 3.2	0.886	1.000	0.978	0.919
APA 3.4	0.938	0.978	1.000	0.953
TDP 2.1	0.950	0.919	0.953	1.000

A.2 Model Performance

These tables list the performance of the individual models for the complete WakeMUC/ WakeFRA/ WakeOP dataset and the DFW97/ MEM95 dataset. Furthermore, the performance is given for the campaigns individually.

Table A.6.: Median root-mean square errors of the models for 694 selected high quality cases (WakeMUC/ WakeFRA/ WakeOP).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.116	0.104	0.789	0.562	0.240	0.195
APA 3.4	0.155	0.140	0.812	0.514	0.210	0.209
APA 3.2	0.238	0.209	0.866	0.508	0.207	0.219
D2P	0.113	0.107	0.591	0.526	0.158	0.176

Table A.7.: Median root-mean square errors of the models (complete DFW97/ MEM95 dataset, 513 landings)

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.242	0.240	0.635	0.644	0.361	0.354
APA 3.4	0.205	0.220	0.592	0.605	0.364	0.361
APA 3.2	0.201	0.235	0.598	0.590	0.356	0.364
D2P	0.222	0.232	0.609	0.603	0.354	0.364

Table A.8.: Median model bias of the models for 694 landings of the WakeOP, WakeFRA and WakeMUC campaign.

model	bias					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	-0.014	-0.005	-0.179	0.239	0.149	-0.004
APA 3.4	-0.121	-0.106	-0.170	0.199	0.100	-0.054
APA 3.2	-0.195	-0.180	-0.152	-0.205	0.082	-0.065
D2P	0.024	0.041	-0.088	0.202	0.021	0.007

Table A.9.: Median model bias of the models for 513 landings of the DFW97 and MEM95 campaign.

model	bias					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.074	0.020	-0.042	-0.021	0.049	0.075
APA 3.4	0.000	-0.057	-0.046	-0.015	0.068	0.104
APA 3.2	-0.058	-0.124	-0.028	-0.014	0.080	0.100
D2P	0.048	-0.012	-0.044	-0.009	0.018	0.047

Table A.10.: Median RMS deviation between model forecast and observation for the Wake-MUC campaign (374 high-quality landings).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.118	0.110	0.827	0.598	0.228	0.215
APA 3.4	0.158	0.144	0.875	0.552	0.228	0.250
APA 3.2	0.253	0.219	0.937	0.554	0.224	0.248
D2P	0.113	0.112	0.691	0.540	0.181	0.199

Table A.11.: Median RMS deviation between model forecast and observation for the WakeFRA campaign (291 landings).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.117	0.101	0.740	0.526	0.243	0.157
APA 3.4	0.159	0.145	0.781	0.457	0.179	0.171
APA 3.2	0.236	0.211	0.829	0.434	0.175	0.172
D2P	0.115	0.103	0.476	0.505	0.125	0.159

Table A.12.: Median RMS deviation between model forecast and observation for the WakeOP campaign (31 landings).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.098	0.097	0.910	0.543	0.364	0.224
APA 3.4	0.101	0.093	0.865	0.525	0.334	0.200
APA 3.2	0.106	0.077	0.878	0.506	0.367	0.225
D2P	0.090	0.097	0.647	0.549	0.161	0.178

Table A.13.: Median RMS deviation between model forecast and observation for the DFW97 campaign (208 landings).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.240	0.246	0.495	0.382	0.249	0.244
APA 3.4	0.203	0.227	0.489	0.371	0.253	0.233
APA 3.2	0.207	0.227	0.513	0.356	0.245	0.235
D2P	0.242	0.251	0.500	0.378	0.221	0.223

Table A.14.: Median RMS deviation between model forecast and observation for the MEM95 campaign (305 landings).

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.234	0.240	0.740	0.818	0.427	0.503
APA 3.4	0.207	0.214	0.735	0.790	0.451	0.512
APA 3.2	0.213	0.230	0.731	0.785	0.435	0.521
D2P	0.220	0.221	0.724	0.787	0.432	0.494

Table A.15.: Median RMS deviation between model forecast and observation for the DEN03 campaign (775 landings) with ϵ derived using structure functions.

model	RMSE					
	Γ_{luff}^*	Γ_{lee}^*	y_{luff}^*	y_{lee}^*	z_{luff}^*	z_{lee}^*
TDP 2.1	0.296	0.302	1.067	1.217	0.500	0.506
APA 3.3	0.219	0.227	1.060	1.228	0.505	0.522
APA 3.2	0.185	0.185	1.042	1.186	0.516	0.527
D2P	0.262	0.269	1.063	1.231	0.519	0.518

A.3 LiDAR Measurement Evaluation

The figures below show the comparison between LiDAR mean and D2P forecast for various initial altitude intervals for the WakeFRA and WakeOP campaign.

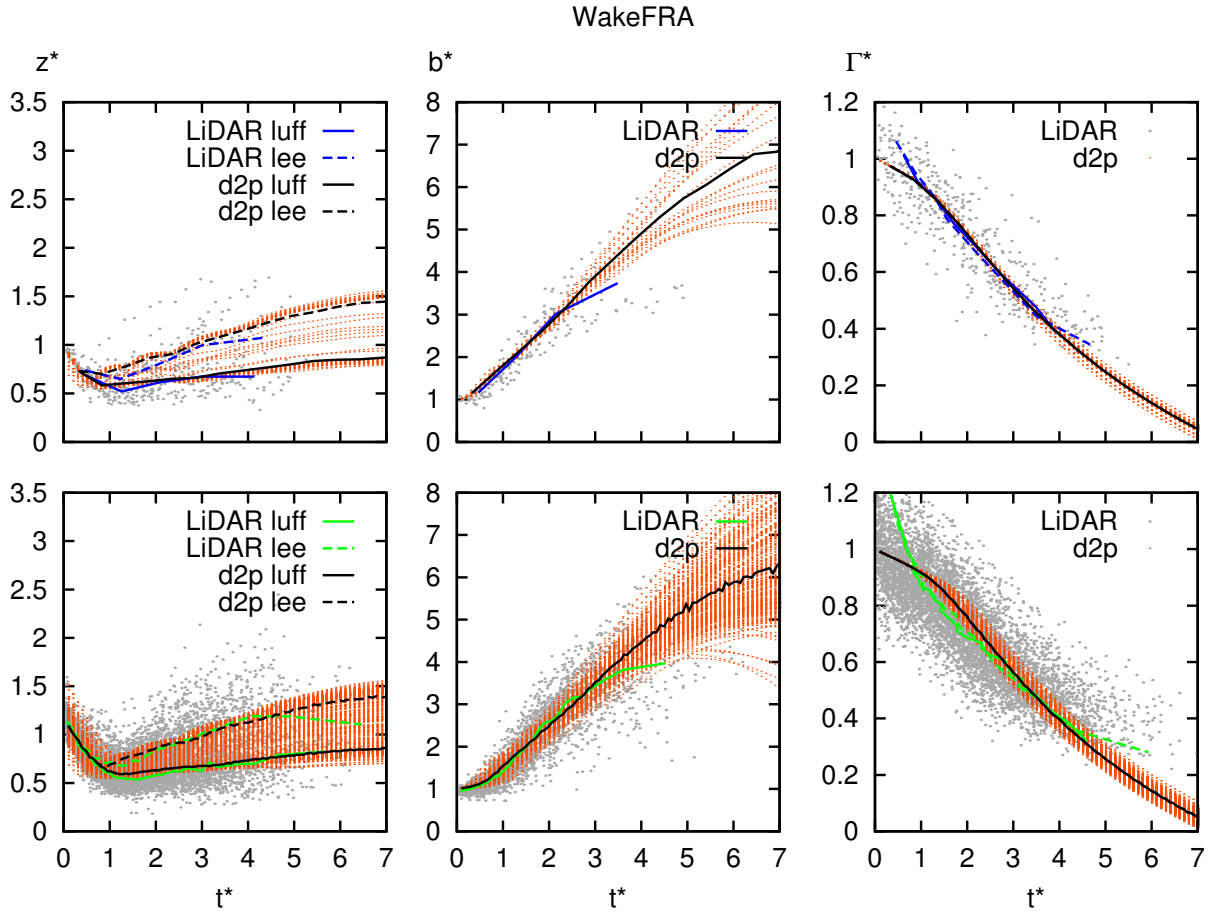


Figure A.1.: Comparison of mean D2P prediction and the LiDAR mean for WakeFRA.

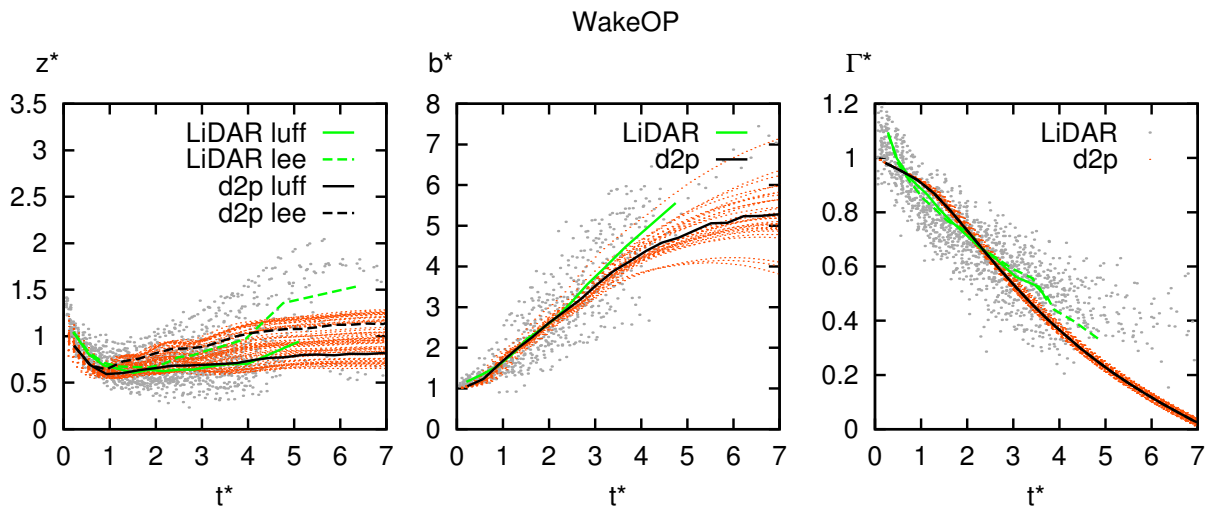


Figure A.2.: Comparison of mean D2P prediction and the LiDAR mean for WakeOP.

A.4 Rare But Dangerous

The figures below illustrate the rare but dangerous cases evaluated in Section 7.2.2, including the single model forecasts.

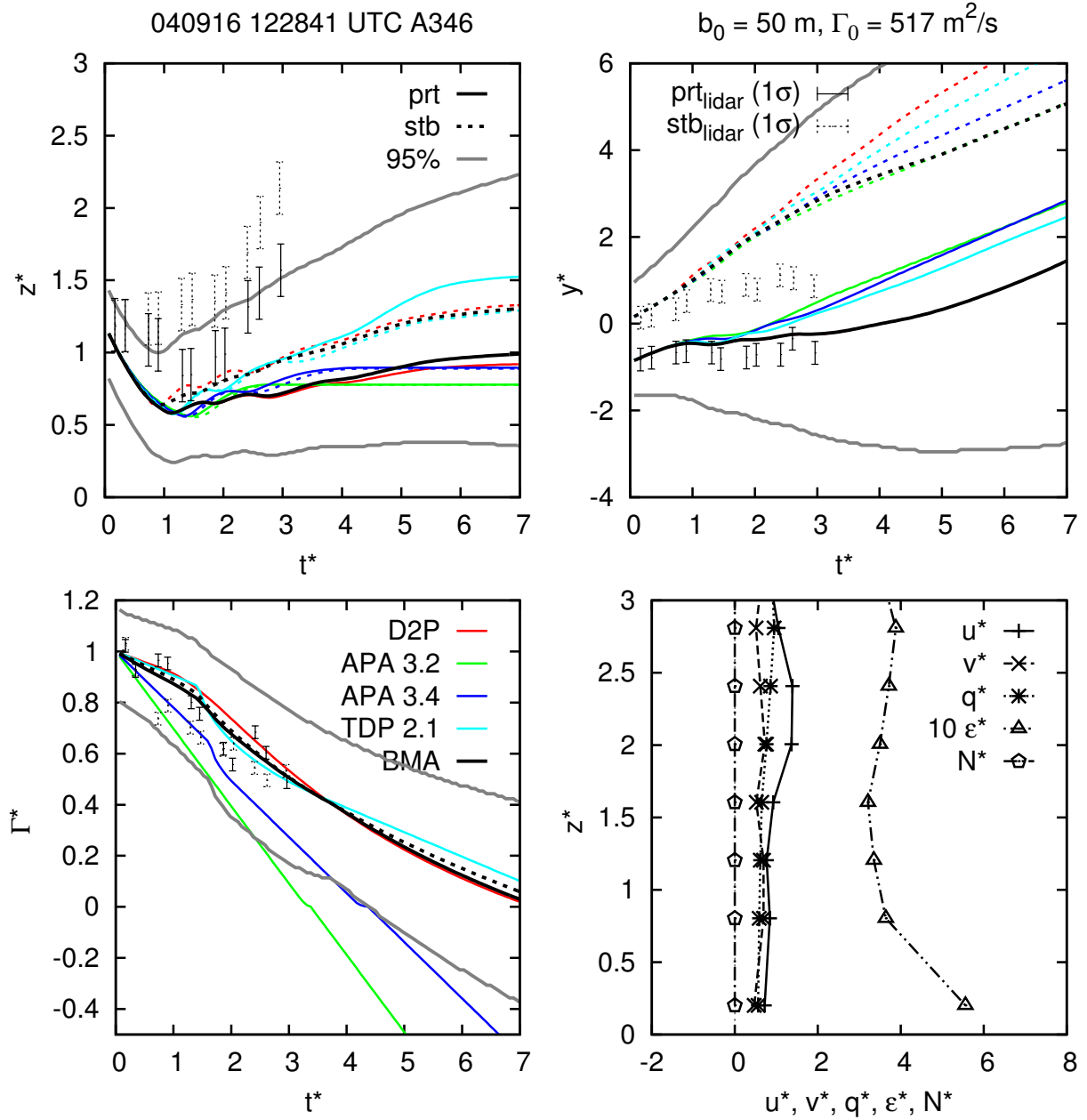


Figure A.3.: Long-lived vortex that remains in the glide path until it cannot be tracked anymore after 97 s.

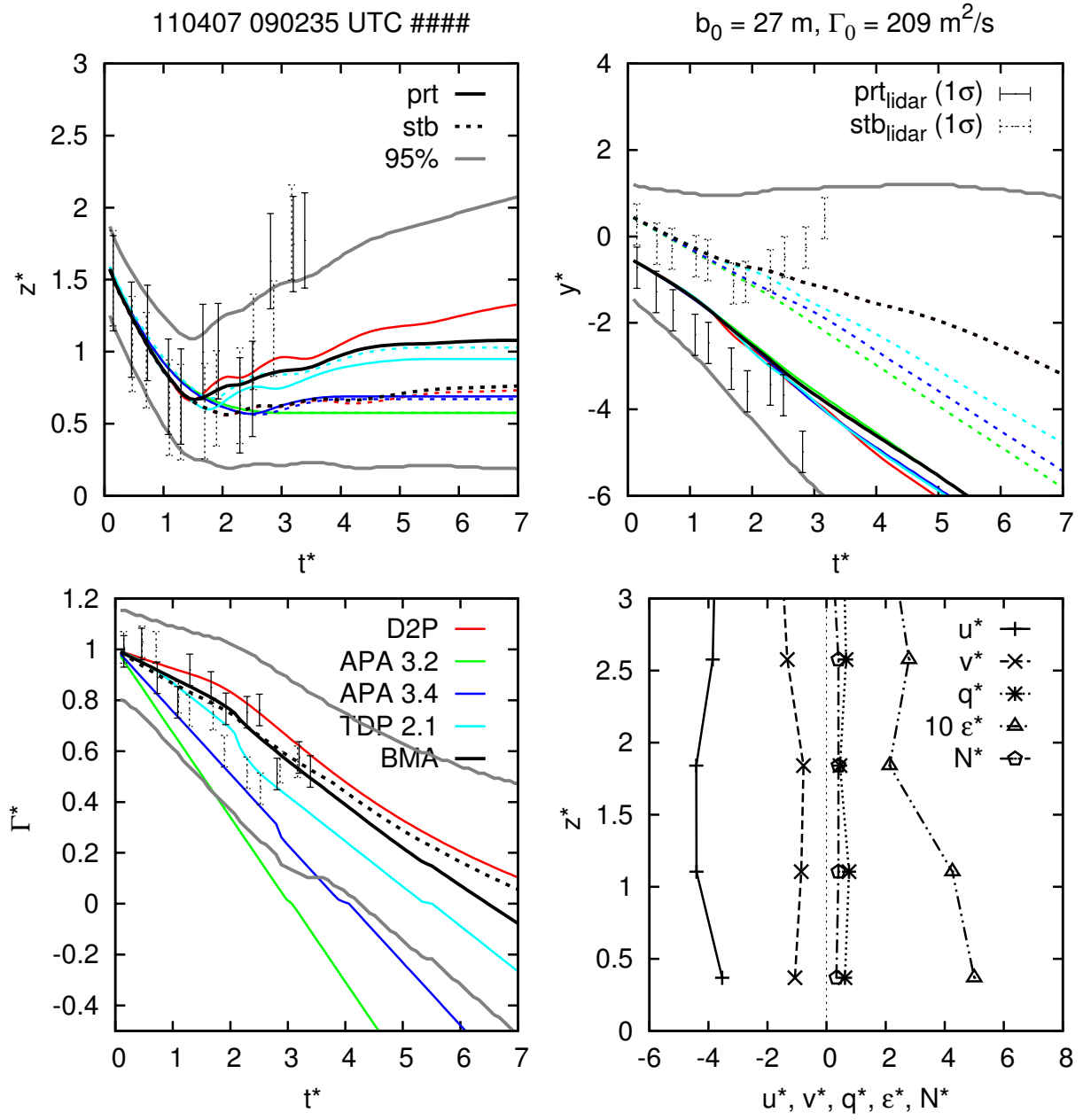


Figure A.4.: Long-lived vortex that remains in the glide path until it cannot be tracked anymore after 77 s.

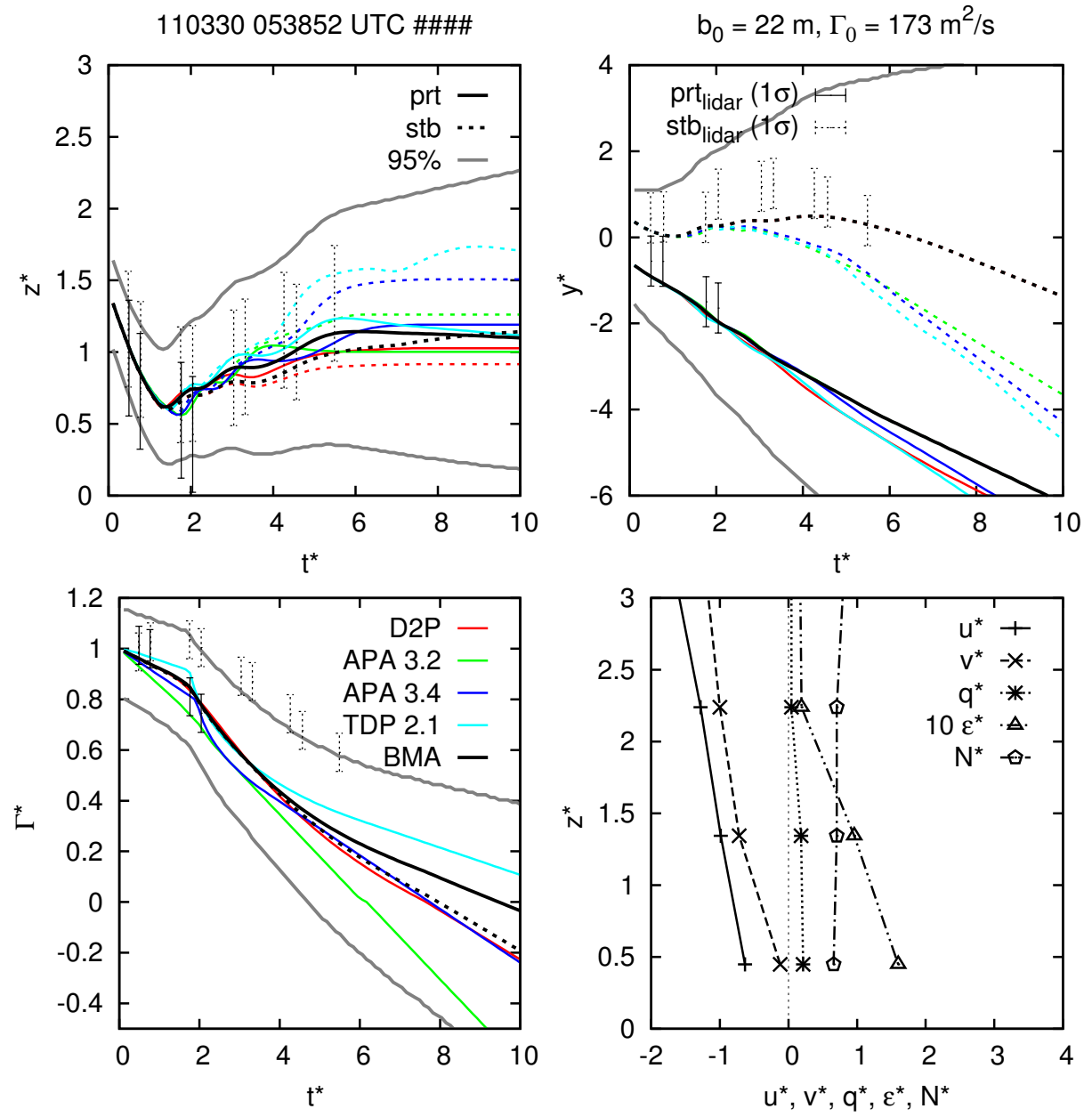


Figure A.5.: Long-lived vortex that remains in the glide path until it cannot be tracked anymore after 97 s.

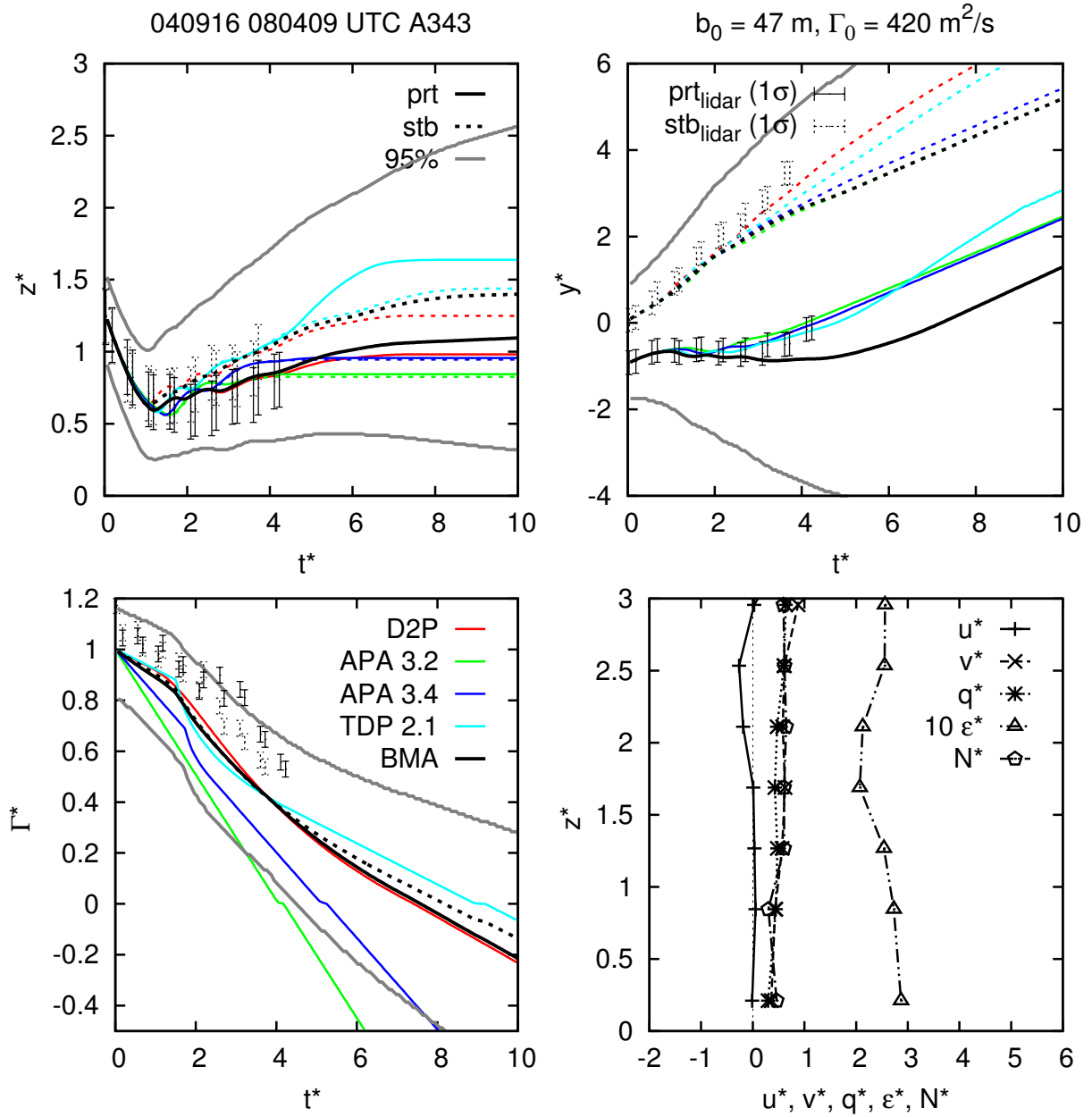


Figure A.6.: Long-lived vortex that remains in the glide path until it cannot be tracked anymore after 149 s.

A.5 Extraordinary Measured Landings

Within the WakeMUC/ WakeFRA/ WakeOP dataset various cases with extraordinary wake vortex rebound have been identified of which 4 are presented here. With exception of the first case, crosswind shear can be observed. None of the models seems to be capable of predicting such extreme behavior.

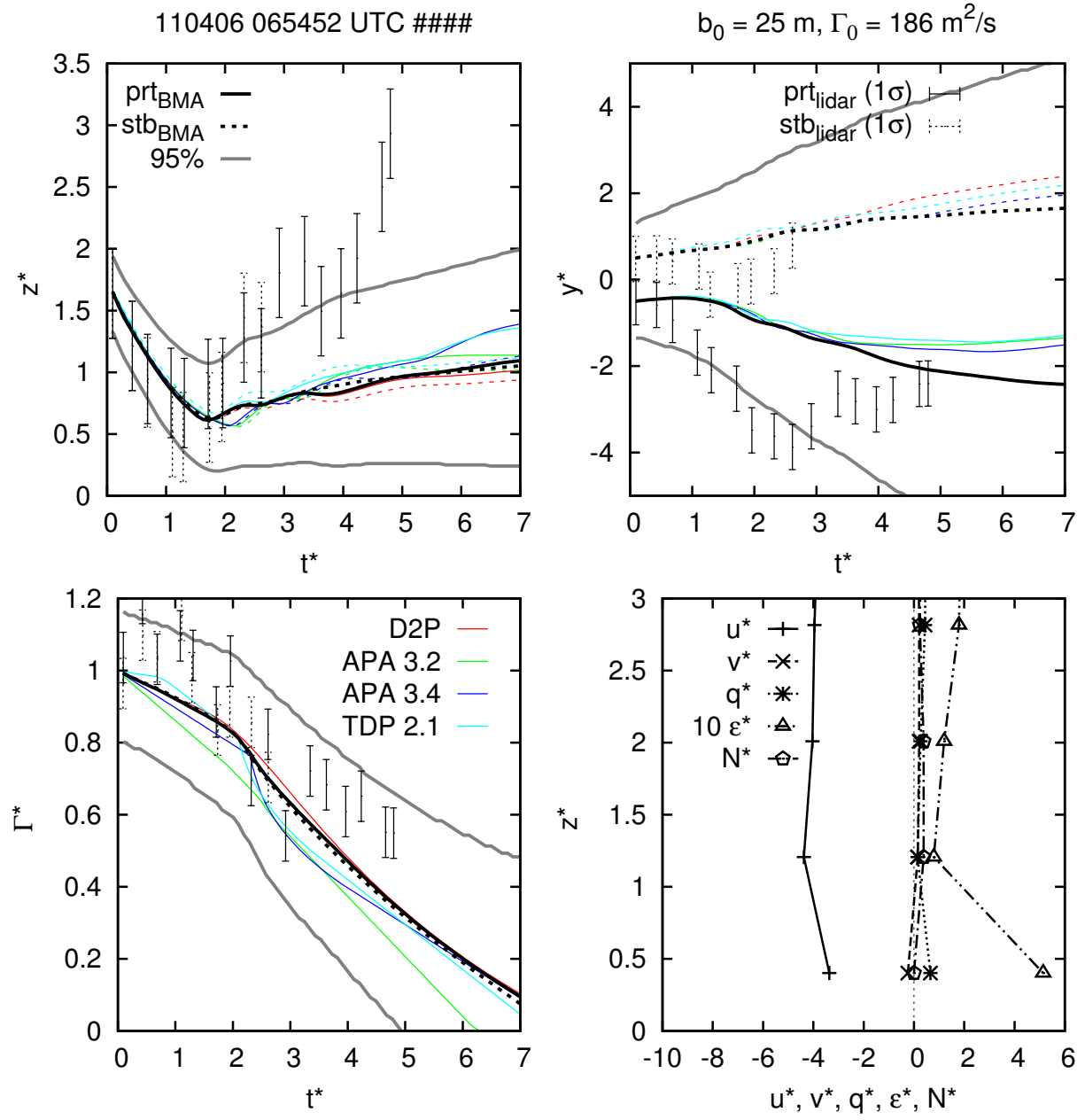


Figure A.7.: Rarely strong rebound of the port vortex, although the 10-minute average cross-wind is low. Lateral vortex transport indicates strong gust that may also be responsible for unusually strong rebound.

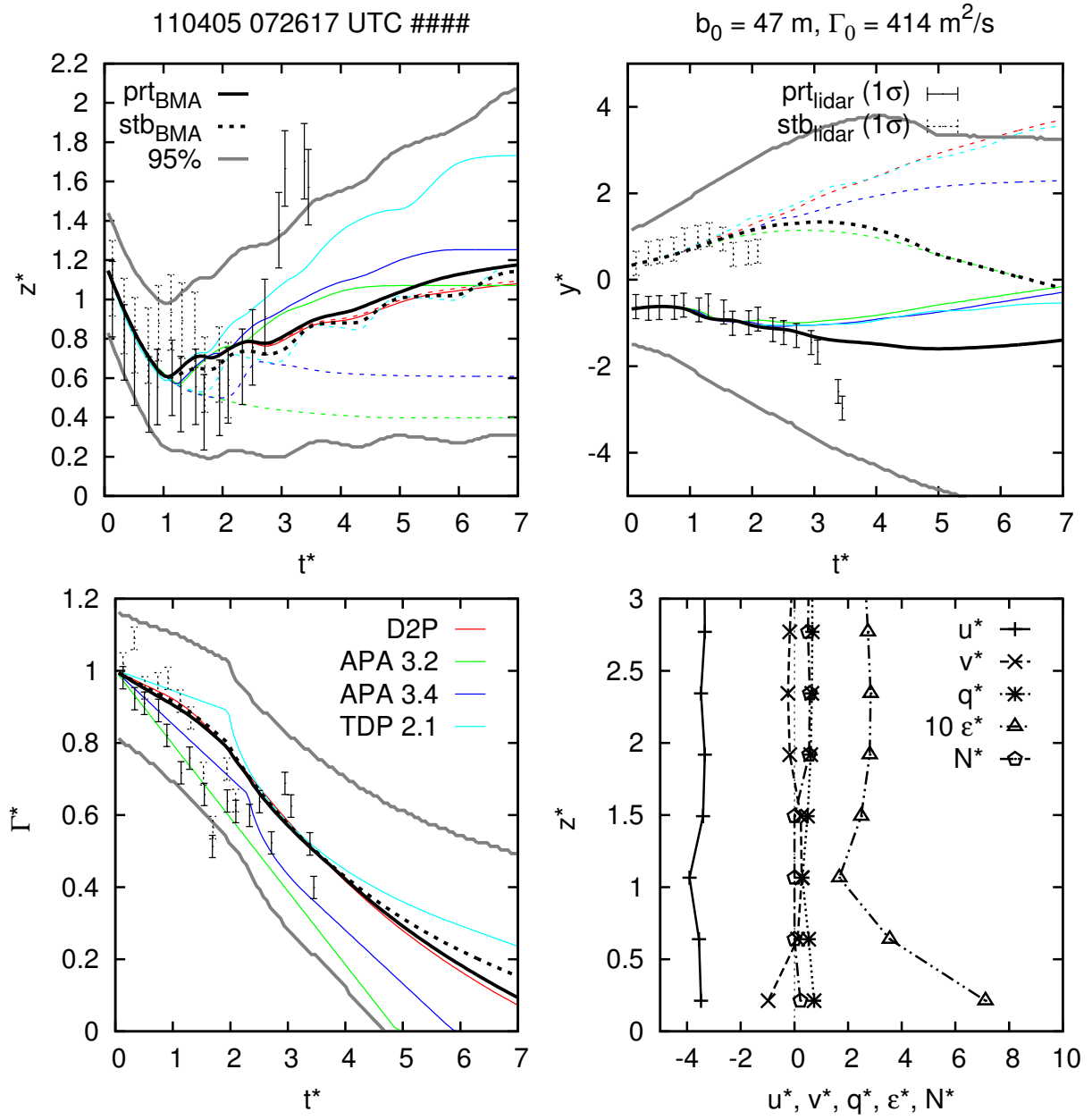


Figure A.8.: Rarely strong rebound of the port vortex, with strongly deviating model forecasts. This might either be caused by the gust that can be observed at $t^* = 3$ in the lateral transport.

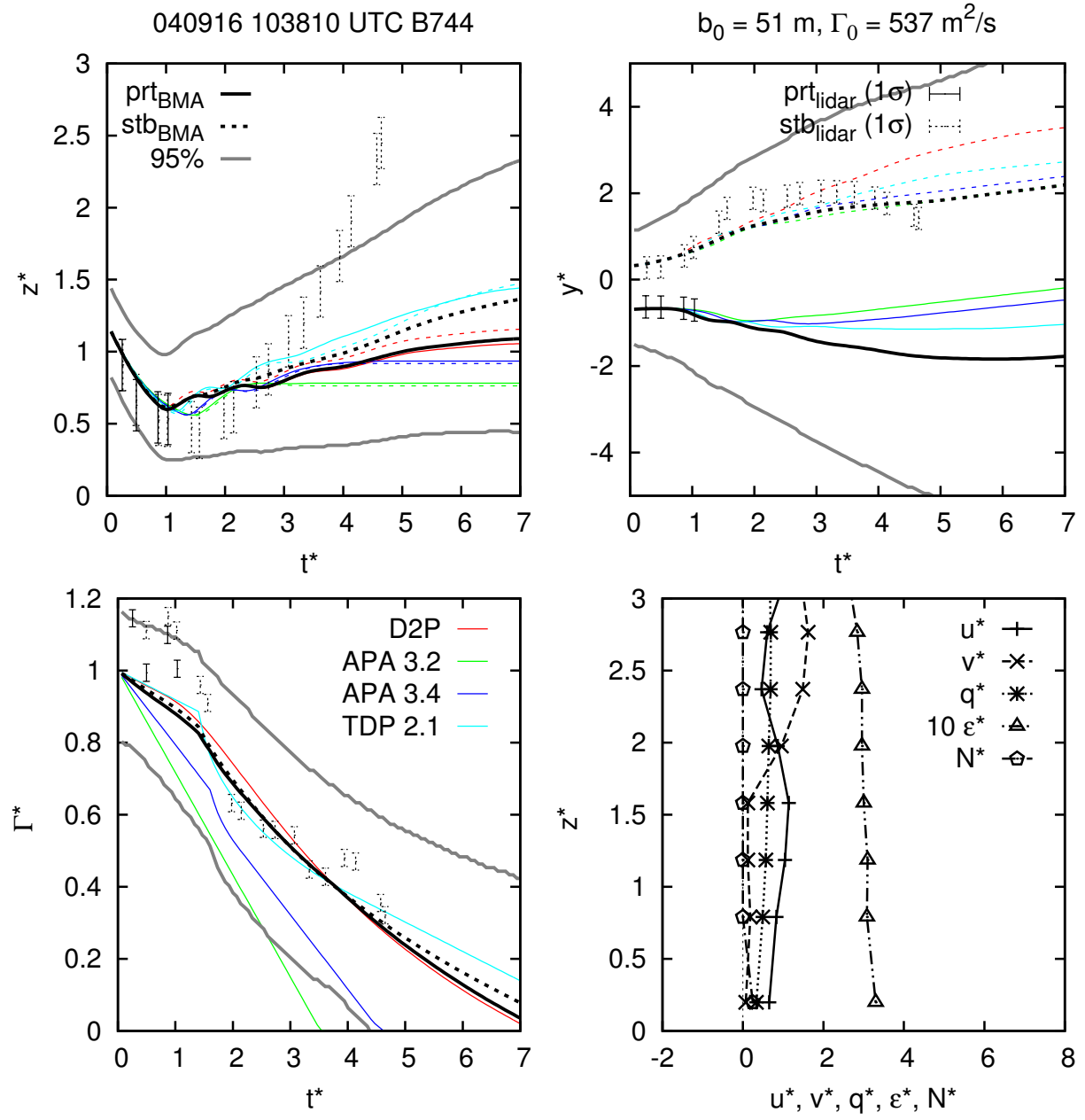


Figure A.9.: Rarely strong rebound of the starboard vortex which might be caused by the gust observed at $t^* = 1.5$. Although a crosswind shear gradient can be observed the results of Proctor et al. [97] suggest that this cannot be the reason for this behavior.

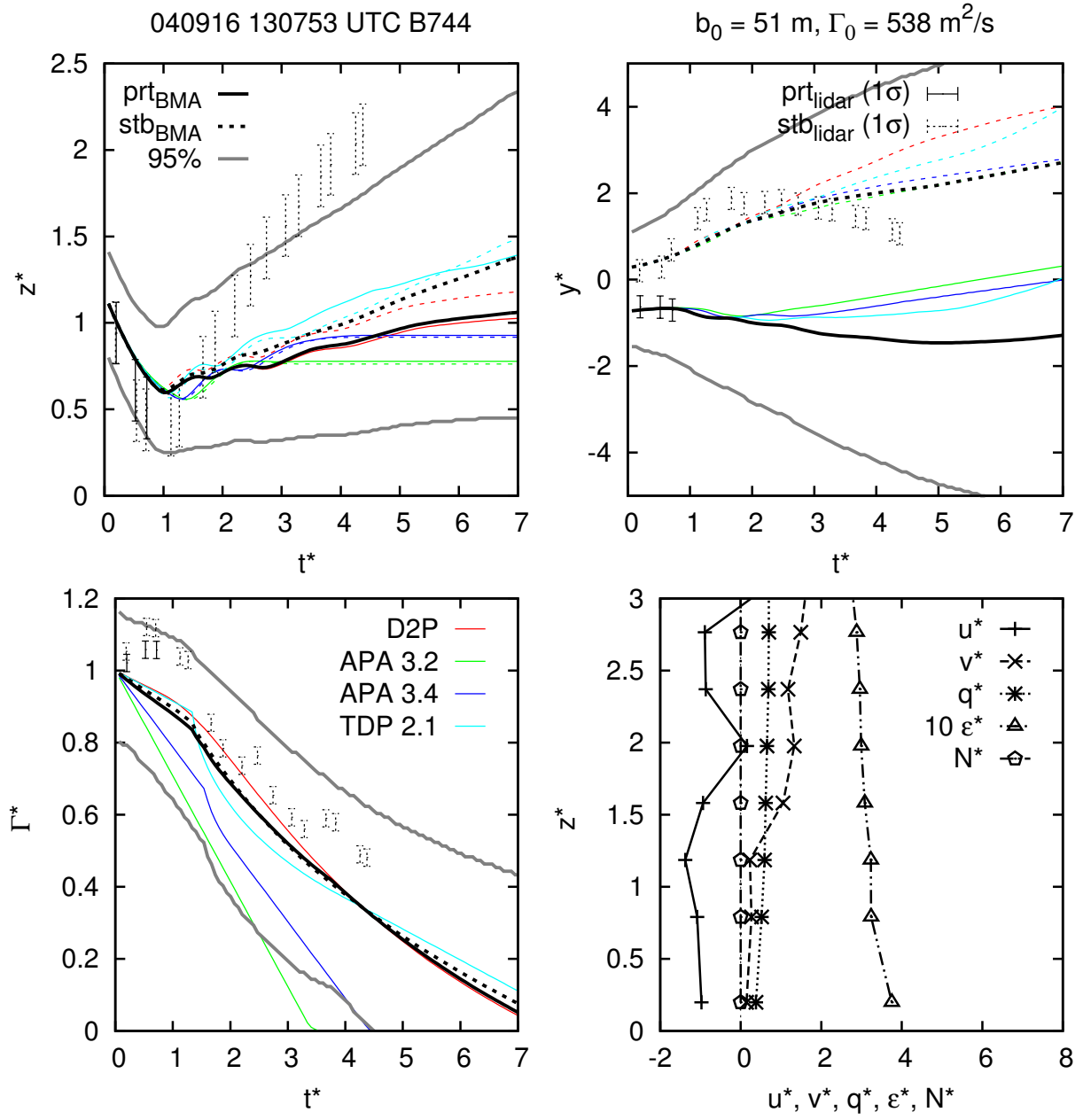


Figure A.10.: Rarely strong rebound of the starboard vortex which might be caused by the gust observed at $t^* = 1$. Although a crosswind shear gradient can be observed the results of Proctor et al. [97] suggest that this cannot be the reason for this behavior.

A.6 Combined Confidence Areas for Vortex Position

The main task of the ensemble within the WSVBS is the calculation of the vortex area. While P2P calculates elliptic areas, the BMA approach offers the opportunity to calculate more sophisticated shapes by employing multivariate PDFs. The previously presented envelopes represent the probability of encountering a single vortex in z - or y -direction individually. In practice the more interesting parameter is however the vortex position in a 2-dimensional prediction plane at a specific time that quantifies the probability to encounter any of the two vortices. As the BMA output consists of PDFs, it is straightforward to compute multivariate distributions here instead of assuming elliptic shapes as in P2P. Then, nominal forecast probabilities may be defined to minimize the risk of a vortex encounter below a certain threshold.

In a first step the joint probabilities are calculated for z_{luff}^* and y_{luff}^* , as well as for z_{lee}^* and y_{lee}^* . In the design of the one-dimensional BMA PDFs, the lateral and vertical vortex position are assumed to be independent and, therefore, the probability to find the luff or lee vortex at a certain position is given by

$$P_{luff} = P(z_{luff}^* \cap y_{luff}^*) = P(z_{luff}^*) \cdot P(y_{luff}^*) \quad (\text{A.1})$$

and

$$P_{lee} = P(z_{lee}^* \cap y_{lee}^*) = P(z_{lee}^*) \cdot P(y_{lee}^*). \quad (\text{A.2})$$

P_{luff} and P_{lee} are thus conjunctive probabilities (i.e. z_{lee}^* "and" y_{lee}^*) [162]. In general, z_{luff}^* and y_{luff}^* may be dependent, for example in the case of strong wind shear. Equations (A.1) and (A.2) result in separate joint (bivariate) PDFs for the luff and lee vortex as depicted in Figure A.11a and Figure A.11b. The normalizations $\iint P_{luff} dz^* dy^* = 1$ and $\iint P_{lee} dz^* dy^* = 1$ are retained. Note, that the contours for given probability levels are only close to elliptic because the ensemble PDFs are a sum of the individual leptokurtic model PDFs.

In a next step the final PDF to encounter any of the two vortices at a time is generated by

$$P = P(luff \cup lee) = (P_{luff} + P_{lee}) / 2. \quad (\text{A.3})$$

This can be understood as disjunction (i.e. luff "or" lee) [162]. Again $\iint P dz^* dy^* = 1$ is guaranteed. Figure A.12a illustrates the overall joint PDF that now exhibits two peaks. To find contours for specific confidence levels, the surface above the contours must be integrated until the desired probability level is reached. As depicted in Figure A.12b, the contours are tapered between both vortices at lower and may even be separated at higher probabilities. It is striking that the confidence in the prediction of the vertical position is much higher than in the lateral position, which can be attributed to a large part to the uncertainty in the wind measurements.

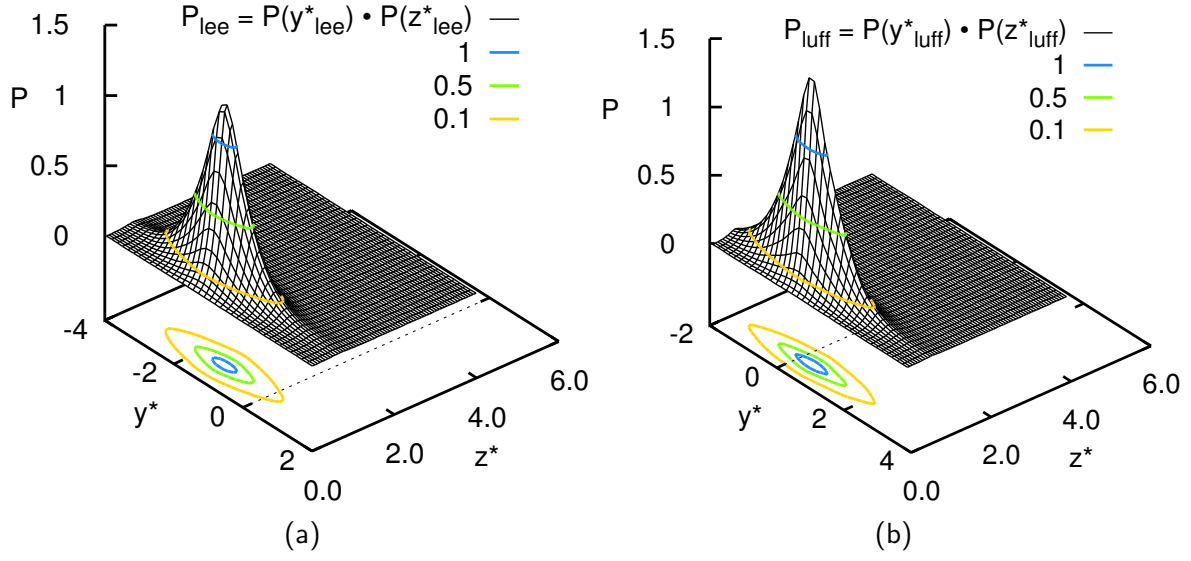


Figure A.11.: Joint PDFs for the lee (a), and the luff vortex (b) for vortices generated in IGE.

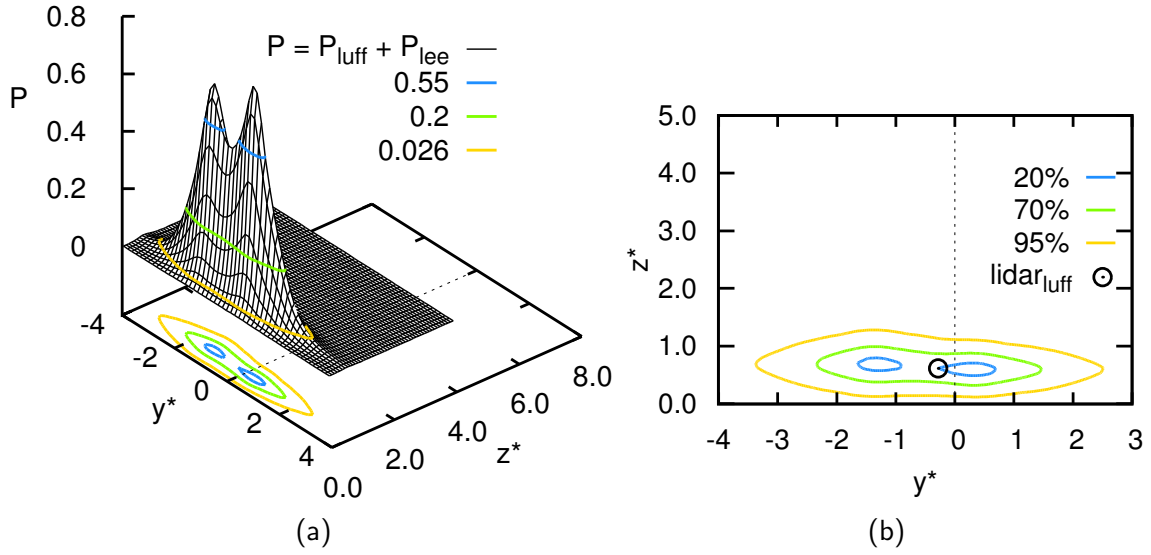


Figure A.12.: Overall joint PDF, calculated as the sum of the joint PDFs of the luff and lee vortex. P_{cut} is depicted for the 20% (0.55), 70% (0.2) and 95% (0.026) confidence level respectively (a). Further, the contours are displayed in the z - y -space exclusively (b).

Bibliography

- [1] T. Gerz, F. Holzäpfel, and D. Darracq, "Commercial aircraft wake vortices", *Progress in Aerospace Sciences*, Vol. 38, No. 3, pp. 181–208, 2002.
- [2] Airbus S.A.S., "Flying by numbers", research report D14029463 issue 2, Airbus, Blagnac Cedex, France, 2015.
- [3] T. Bauer, D. Vechtel, F. Abdelmoula, and T. Immisch, "In-flight wake encounter prediction with the wake encounter avoidance and advisory system", *AIAA Atmospheric and Space Environments Conference*, June 2014. doi:10.2514/6.2014-2333.
- [4] "Re-categorization (RECAT) of FAA wake turbulence separation categories at specific airports", Safety Alert for Operators 12007, FAA, 2013.
- [5] E. Stumpf, "Untersuchung von 4-Wirbelsystemen zur Minimierung von Wirbelschleppen und ihre Realisierung an Transportflugzeugen", Dissertation, Bibliothek der RWTH Aachen, 2003.
- [6] E. Stumpf, "Study of four-vortex aircraft wakes and layout of corresponding aircraft configurations", *Journal of aircraft*, Vol. 42, No. 3, pp. 722–730, 2005.
- [7] A. Stephan, F. Holzäpfel, and T. Misaka, "Hybrid simulation of wake-vortex evolution during landing on flat terrain and with plate line", *International Journal of Heat and Fluid Flow*, Vol. 49, pp. 18–27, 2014.
- [8] T. Misaka, F. Holzäpfel, and T. Gerz, "Large-eddy simulation of aircraft wake evolution from roll-up until vortex decay", *AIAA Journal*, Vol. 53, No. 9, pp. 2646–2670, 2015.
- [9] V. R. Corsiglia, V. J. Rossow, and D. L. Ciffone, "Experimental study of the effect of span loading on aircraft wakes", *Journal of Aircraft*, Vol. 13, No.12, pp. 968–973, 1976.
- [10] C. Breitsamter, "Wake vortex characteristics of transport aircraft", *Progress in Aerospace Sciences*, Vol. 47, No. 2, pp. 89–134, 2011.
- [11] D. L. Ciffone and K. Orloff, "Far-field wake-vortex characteristics of wings", *Journal of Aircraft*, Vol. 12, No. 5, pp. 464–470, 1975.
- [12] P. Meunier, U. Ehrenstein, T. Leweke, and M. Rossi, "A merging criterion for two-dimensional co-rotating vortices", *Physics of Fluids*, Vol. 14, No. 8, pp. 2757–2766, 2002.
- [13] T. Sarpkaya, R. Robins, and D. Delisi, "Wake-vortex eddy-dissipation model predictions compared with observations", *Journal of Aircraft*, Vol. 38, pp. 687–692, 2001.

- [14] R. E. Robins and D. P. Delisi, "NWRA AVOSS wake vortex prediction algorithm version 3.1.1", research report NASA/CR-2002-211746, Northwest Research Associates, Inc., Bellevue, Washington, USA, 2002.
- [15] F. Holzäpfel, "Probabilistic two-phase aircraft wake-vortex decay and transport model", *Journal of Aircraft*, Vol. 40, pp. 323–331, 2003.
- [16] F. Holzäpfel, "Probabilistic two-phase aircraft wake-vortex model: Further development and assessment", *AIAA Journal*, Vol. 43, No.3, pp. 700–708, 2006.
- [17] F. H. Proctor, D. W. Hamilton, and G. F. Switzer, "TASS driven algorithms for wake prediction", *44th AIAA Aerospace Sciences Meeting and Exhibit*, AIAA, No. 2006-1073, 2006.
- [18] F. Holzäpfel and M. Steen, "Aircraft wake-vortex evolution in ground proximity: Analysis and parameterization", *AIAA Journal*, Vol. 45, pp. 218–227, 2007.
- [19] I. De Visscher, L. Bricteux, and G. Winckelmans, "Aircraft vortices in stably stratified and weakly turbulent atmospheres: Simulation and modeling", *Journal of Aircraft*, Vol. 51, No. 3, pp. 551–566, 2013.
- [20] I. De Visscher, T. Lonfils, and G. Winckelmans, "Fast time modeling of ground effects on wake vortex transport and decay", *Journal of Aircraft*, Vol. 50, No. 5, pp. 1514–1525, 2013.
- [21] F. H. Proctor and D. W. Hamilton, "Evaluation of fast-time wake vortex prediction models", *AIAA paper*, Vol. 2009-344, 2009.
- [22] F. Holzäpfel, A. Stephan, N. Tchipev, T. Heel, S. Körner, and T. Misaka, "Impact of wind and obstacles on wake vortex evolution in ground proximity", in *6th AIAA Atmospheric and Space Environments Conference*, Atlanta, No. 2470, 2014.
- [23] M. Frech and F. Holzäpfel, "Skill of an aircraft wake-vortex model using weather prediction and observation", *Journal of Aircraft*, Vol. 45, pp. 461–470, 2008.
- [24] S.D.Campbell, T.J.Dasey, R.E.Freehart, R.M.Heinrichs, M.P.Matthewsa, G.H.Perras, and G.S.Rowe, "Wake vortex field measurement program at Memphis, TN data guide", research report NASA/L-2. 1997, Lincoln Laboratory, Massachusetts Institute of Technology, 1997.
- [25] Eurocontrol, "Asia pacific growing fastest - but other regions are still significant", *Skyway magazine*, No. autumn/winter, 2015.
- [26] G. De Clercq *et al.*, "Guidelines for the application of the ECAC radar separation minima", report ASM.ET1.ST18.1000-REP-01.00, Eurocontrol, 1998.
- [27] F. Rooseleer and V. Treve, "European wake turbulence categorisation and separation minima on approach and departure", report 1.1, Eurocontrol, 2015.

-
- [28] R. Doe, "Caution wake turbulence: New rules for the EU", April 2016. <http://flightservicebureau.org/caution-wake-turbulence-new-rules-for-the-eu/>.
- [29] "Implementation of the RECAT-EU wake turbulence separation scheme at Paris Charles De Gaulle, Paris-Le Bourget and Pointoise-Cormeilles-En-Vexin airports from March 22nd 2016", report A 03/16, Direction des Opérations, Service de l'Information Aéronautique, 2016.
- [30] C. Morris, J. Peters, and P. Choroba, "Validation of the time based separation concept at London Heathrow Airport", in *10th USA/Europe ATM R&D Seminar, Chicago*, pp. 10–13, 2013.
- [31] NATS, "Intelligent approach, time and distance based separation", May 2016. <http://www.nats.aero/wp-content/uploads/2015/03/TBS-Poster-online.pdf>.
- [32] F. Holzäpfel, A. Stephan, T. Heel, and S. Körner, "Enhanced wake vortex decay in ground proximity triggered by plate lines", *Aircraft Engineering and Aerospace Technology*, Vol. 88, No. 2, pp. 206–214, 2016. doi:10.1108/AEAT-02-2015-0045.
- [33] I. De Visscher, G. Winckelmans, T. Lonfils, L. Bricteux, M. Duponcheel, and N. Bourgeois, "The WAKE4D simulation platform for predicting aircraft wake vortex transport and decay, description and examples of application", *AIAA Atmospheric and Space Environments Conference*, August 2010.
- [34] M. Pruis and D. Delisi, "Assessment of fast-time wake vortex prediction models using pulsed and continuous wave lidar observations at several different airports", *AIAA paper*, June 2011.
- [35] D. A. Hinton, J. K. Charnock, and D. R. Bagwell, "Design of an aircraft vortex spacing system for airport capacity improvement", in *38th AIAA Aerospace Sciences Meeting*, No. 2000-0622, 2000.
- [36] D. K. Rutishauser and C. J. OConnor, "Aircraft wake vortex spacing system (AVOSS) performance update and validation study", 2001.
- [37] F. Holzäpfel, T. Gerz, M. Frech, A. Tafferner, F. Köpp, I. Smalikho, S. Rahm, K.-U. Hahn, and C. Schwarz, "The wake vortex prediction and monitoring system WSVBS part I: Design", *Air Traffic Control Quarterly*, Vol. 17, No. 4, p. 301, 2009.
- [38] C. W. Schwarz and K.-U. Hahn, "Full-flight simulator study for wake vortex hazard area investigation", *Aerospace Science and Technology*, Vol. 10, No. 2, pp. 136–143, 2006.
- [39] T. Gerz, F. Holzäpfel, W. Gerling, A. Scharnweber, M. Frech, K. Kober, K. Dengler, and S. Rahm, "The wake vortex prediction and monitoring system wsvbs part II: Performance and ATC integration at Frankfurt Airport", *Air Traffic Control Quarterly*, Vol. 17, No. 4, pp. 323–346, 2009.

- [40] F. Holzäpfel, M. Frech, T. Gerz, A. Tafferner, K.-U. Hahn, C. Schwarz, H.-D. Joos, B. Korn, H. Lenz, R. Luckner, *et al.*, "Aircraft wake vortex scenarios simulation package – WakeScene", *Aerospace Science and Technology*, Vol. 13, No. 1, pp. 1–11, 2009.
- [41] J. ICAO, "Rules of the air", in *Annex 2 to the Convention on International Civil Aviation, Chicago, IL*, Vol. 10, 2005.
- [42] "Wake turbulence", report DOT/FAA/RD-95/6, Department of Transportation, Federal Aviation Administration, 1995.
- [43] U. Schumann and R. Sharman, "Aircraft wake-vortex encounter analysis for upper levels", *Journal of Aircraft*, Vol. 52, No. 4, pp. 1277–1285, 2014.
- [44] S. Ramasamy, R. Sabatini, A. Gardi, and T. Kistan, "Next generation flight management system for real-time trajectory based operations", *Applied Mechanics and Materials*, Vol. 629, pp. 344–349, 2014.
- [45] I. Sölch, F. Holzäpfel, F. Abdelmoula, and D. Vechtel, "Performance of onboard wake–vortex prediction systems employing various meteorological data sources", *Journal of Aircraft*, 2016. doi:10.2514/1.C033732.
- [46] P. Vrancken, M. Wirth, D. Rempel, G. Ehret, A. Dolfi-Bouteyre, L. Lombard, T. Gaudo, D. Rees, H. Barny, and P. Rondeau, "Clear air turbulence detection and characterisation in the DELICAT airborne lidar project", in *25th International Laser RADAR conference*, 2010.
- [47] P. Vrancken, G. Ehret, H. Barny, H. Veermann, and L. Lombard, "Aeronautics LiDAR applications - airborne LiDAR detection of clear air turbulence (CAT) within the FP7 DELICAT project", in *4th EASN Association International Workshop on Flight Physics & Aircraft Design, Aachen*, 2014.
- [48] R. Luckner, G. Höhne, and M. Fuhrmann, "Hazard criteria for wake vortex encounters during approach", *Aerospace Science and Technology*, Vol. 8, No. 8, pp. 673–687, 2004.
- [49] H. Lamb, *Hydrodynamics*. Cambridge university press, 1932.
- [50] D. Bieniek and R. Luckner, "Simulation of aircraft encounters with perturbed vortices considering unsteady aerodynamic effects", *Journal of Aircraft*, Vol. 51, No. 3, pp. 705–718, 2014.
- [51] D. Vechtel, "Simulation study of wake encounters with straight and deformed vortices", *The Aeronautical Journal*, Vol. 120, No. 1226, pp. 651–674, 2016.
- [52] F. Sanders, "On subjective probability forecasting", *Journal of Applied Meteorology*, Vol. 2, No. 2, pp. 191–201, 1963.
- [53] F. Sanders, "Skill in forecasting daily temperature and precipitation: Some experimental results", *Bulletin of the American Meteorological Society*, Vol. 54, No. 11, pp. 1171–1178, 1973.

-
- [54] L. F. Bosart, "Sunya experimental results in forecasting daily temperature and precipitation", *Monthly Weather Review*, Vol. 103, No. 11, pp. 1013–1020, 1975.
- [55] J. R. Gyakum, "Experiments in temperature and precipitation forecasting for Illinois", *Weather and Forecasting*, Vol. 1, No. 1, pp. 77–88, 1986.
- [56] M. R. Allen, P. A. Stott, J. F. Mitchell, R. Schnur, and T. L. Delworth, "Quantifying the uncertainty in forecasts of anthropogenic climate change", *Nature*, Vol. 407, No. 6804, pp. 617–620, 2000.
- [57] J. Räisänen and T. Palmer, "A probability and decision-model analysis of a multimodel ensemble of climate change simulations", *Journal of Climate*, Vol. 14, No. 15, pp. 3212–3226, 2001.
- [58] F. Giorgi and L. O. Mearns, "Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "Reliability Ensemble Averaging" (REA) method", *Journal of Climate*, Vol. 15, pp. 1141–1158, 2002.
- [59] J. M. Murphy, D. M. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, "Quantification of modelling uncertainties in a large ensemble of climate change simulations", *Nature*, Vol. 430, No. 7001, pp. 768–772, 2004.
- [60] D. West, P. Mangiameli, R. Rampal, and V. West, "Ensemble strategies for a medical diagnostic decision support system: A breast cancer diagnosis application", *European Journal of Operational Research*, Vol. 162, No. 2, pp. 532–551, 2005.
- [61] H. M. Meier, H. C. Andersson, B. Arheimer, T. Blenckner, B. Chubarenko, C. Donnelly, K. Eilola, B. G. Gustafsson, A. Hansson, J. Havenhand, *et al.*, "Comparing reconstructed past variations and future projections of the baltic sea ecosystem—first results from multi-model ensemble simulations", *Environmental Research Letters*, Vol. 7, No. 3, p. 034005, 2012.
- [62] P. Martre, D. Wallach, S. Asseng, F. Ewert, J. W. Jones, R. P. Rötter, K. J. Boote, A. C. Ruane, P. J. Thorburn, D. Cammarano, *et al.*, "Multimodel ensembles of wheat growth: many models are better than one", *Global change biology*, Vol. 21, No. 2, pp. 911–925, 2015.
- [63] R. Hagedorn, F. J. DOBLAS-REYES, and T. Palmer, "The rationale behind the success of multi-model ensembles in seasonal forecasting—i. basic concept", *Tellus A*, Vol. 57, No. 3, pp. 219–233, 2005.
- [64] A. P. Weigel, M. Liniger, and C. Appenzeller, "Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?", *Quarterly Journal of the Royal Meteorological Society*, Vol. 134, No. 630, pp. 241–260, 2008.
- [65] E. N. Lorenz, "A study of the predictability of a 28-variable atmospheric model", *Tellus A*, Vol. 17, No. 3, 1965.

- [66] E. S. Epstein, "Stochastic dynamic prediction", *Tellus A*, Vol. 21, No. 6, 1969.
- [67] D. J. Stensrud, J.-W. Bao, and T. T. Warner, "Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems", *Monthly Weather Review*, Vol. 128, No. 7, pp. 2077–2107, 2000.
- [68] C. Tebaldi and R. Knutti, "The use of the multi-model ensemble in probabilistic climate projections", *The Royal Society*, Vol. 365, pp. 2053–2075, 2007.
- [69] T. Krishnamurti, C. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, "Improved weather and seasonal climate forecasts from multimodel superensemble", *Science*, Vol. 285, No. 5433, pp. 1548–1550, 1999.
- [70] T. Palmer, F. Doblas-Reyes, R. Hagedorn, A. Alessandri, S. Gualdi, U. Andersen, H. Feddersen, P. Cantelaube, J. Terres, M. Davey, *et al.*, "Development of a european multimodel ensemble system for seasonal-to-interannual prediction (DEMETER)", *Bulletin of the American Meteorological Society*, Vol. 85, No. 6, pp. 853–872, 2004.
- [71] A. E. Raftery, T. Gneiting, F. Balabdaoui, and M. Polakowski, "Using Bayesian model averaging to calibrate forecast ensembles", *Monthly Weather Review*, Vol. 133, pp. 1155–1174, 2005.
- [72] V. M. Krasnopolsky and Y. Lin, "A neural network nonlinear multimodel ensemble to improve precipitation forecasts over continental us", *Advances in Meteorology*, Vol. 2012, 2012.
- [73] A. M. Greene, L. Goddard, and U. Lall, "Probabilistic multimodel regional temperature change projections", *Journal of Climate*, Vol. 19, No. 17, pp. 4326–4343, 2006.
- [74] T. Palmer, C. Brankovic, F. Molteni, S. Tibaldi, L. Ferranti, A. Hollingsworth, U. Cubasch, and E. Klinker, "The european centre for medium-range weather forecasts (ECMWF) program on extended-range prediction", *Bulletin of the American Meteorological Society*, Vol. 71, No. 9, pp. 1317–1330, 1990.
- [75] L. S. Froude, L. Bengtsson, and K. I. Hodges, "The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems", *Monthly weather review*, Vol. 135, No. 7, pp. 2545–2567, 2007.
- [76] M. Yamaguchi, R. Sakai, M. Kyoda, T. Komori, and T. Kadowaki, "Typhoon ensemble prediction system developed at the japan meteorological agency", *Monthly Weather Review*, Vol. 137, No. 8, pp. 2592–2604, 2009.
- [77] N. A. Nash'at, R. L. VanValkenburg, R. L. Bowles, F. M. L. Duparcmeur, T. Gloude-smann, S. van Lochem, and E. Ras, "Evaluation of fast-time wake vortex models using wake encounter flight test data", in *6th AIAA Atmospheric and Space Environments Conference*, 2014.
- [78] P. K. Kundu and I. M. Cohen, *Fluid Mechanics*. San Diego, CA, USA: Elsevier Academic Press, 2004.

-
- [79] I. Hennemann and F. Holzäpfel, "Large-eddy simulation of aircraft wake vortex deformation and topology", *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, Vol. 225, 2011. doi: 10.1177/0954410011402257.
- [80] F. H. Proctor, D. W. Hamilton, and J. Han, "Wake vortex transport and decay in ground effect: Vortex linking with the ground", *AIAA paper*, January 2000.
- [81] L. Prandtl, "Tragflügeltheorie I u. II", *Nachr. vd K. Ges. d. Wissensch. zu Göttingen, Mathematisch-physikalische Klasse*, 1918.
- [82] F. Lanchester, *Aerodynamics*. Orange Street Leicester Square, London, UK: Archibald Constable & CO. LTD., 1907.
- [83] F. Holzäpfel, T. Gerz, F. Köpp, E. Stumpf, M. Harris, R. I. Young, and A. Dolfi-Bouteyre, "Strategies for circulation evaluation of aircraft wake vortices measured by lidar", *Journal of Atmospheric and Oceanic Technology*, Vol. 20, No. 8, pp. 1183–1195, 2003.
- [84] W.F.Phillips and D.O.Snyder, "Modern adaptation of Prandtl's classic lifting-line theory", *Journal of Aircraft*, Vol. 37, pp. 662–670, 2000.
- [85] T. A. Talay, *Introduction to the Aerodynamics of Flight*, Vol. 367. Scientific and Technical Information Office, National Aeronautics and Space Administration, 1975.
- [86] F. Holzäpfel, *Aircraft Wake Vortex Evolution and Prediction*. Habilitation, TU München, October 2005.
- [87] T. Misaka, S. Obayashi, A. Stephan, F. Holzäpfel, T. Gerz, and K. Nakahashi, "Numerical simulation of jet-wake vortex interaction", *AIAA Paper*, Vol. 926, p. 2014, 2014.
- [88] D.P.Delisi, G.C.Greene, R.E.Robins, D.C.Vicroy, and F.Y.Wang, "Aircraft wake vortex core size measurements", *AIAA paper*, Vol. 2003-3811, 2003.
- [89] N. A. Nash'at, F. H. Proctor, F. M. L. Duparcmeur, and D. Jacob, "Review of idealized aircraft wake vortex models", in *52nd Aerospace Sciences Meeting, AIAA SciTech*, 2014.
- [90] F. Holzäpfel, N. Tchipev, and A. Stephan, "Wind impact on single vortices and counter-rotating vortex pairs in ground proximity", *Flow, Turbulence and Combustion*, pp. 1–20, 2016.
- [91] T. Misaka, F. Holzäpfel, and T. Gerz, "Wake evolution of high-lift configuration from roll-up to vortex decay", *AIAA Paper*, Vol. 362, No. 2013, p. 11, 2013.
- [92] A. Stephan, "Wake vortices of landing aircraft", Dissertation, LMU München, 2014.
- [93] F. Laporte and T. Leweke, "Elliptic instability of counter-rotating vortices: experiment and direct numerical simulation", *AIAA journal*, Vol. 40, No. 12, pp. 2483–2494, 2002.
- [94] G. C. Greene, "An approximate model of vortex decay in the atmosphere", *Journal of Aircraft*, Vol. 23, No. 7, pp. 566–573, 1986.

- [95] D. Etling, *Theoretische Meteorologie: Eine Einführung*. Springer-Verlag, 2008.
- [96] D. P. Delisi and R. E. Robins, "Effects of crosswind shear on trailing vortex evolution", in *44th AIAA Aerospace Sciences Meeting and Exhibit*, pp. 9–12, 2006.
- [97] F. Proctor, D. Hinton, J. Han, D. Schowalter, and Y. Lin, "Two dimensional wake vortex simulations in the atmosphere: Preliminary sensitivity studies", *AIAA Paper*, No. 97-0056, pp. 6–10, 1997.
- [98] F. H. Proctor and N. Ahmad, "Crosswind shear gradient affect on wake vortices", in *3rd AIAA Atmospheric Space Environments Conference*, p. 3038, 2011.
- [99] H. Kuo, "Motions of vortices and circulating cylinder in shear flow with friction", *Journal of the Atmospheric Sciences*, Vol. 26, No. 3, pp. 390–398, 1969.
- [100] P. R. Veillette, "Data show that US wake-turbulence accidents are most frequent at low altitude and during approach and landing", *Flight Safety Digest*, Vol. 21, No. 3-4, pp. 1–47, 2002.
- [101] A. Stephan, F. Holzäpfel, and T. Misaka, "Aircraft wake-vortex decay in ground proximity-physical mechanisms and artificial enhancement", *Journal of Aircraft*, Vol. 50, No. 4, pp. 1250–1260, 2013.
- [102] A. Corjon and T. Poinso, "Behavior of wake vortices near ground", *AIAA Journal*, Vol. 35, No. 5, pp. 849–855, 1997.
- [103] H. JK and P. FJ, "Flowfield produced by trailing vortices in the vicinity of the ground", *AIAA Journal*, Vol. 9, No. 8, pp. 1659–1660, 1971.
- [104] P. Orlandi, "Vortex dipole rebound from a wall", *Physics of Fluids A: Fluid Dynamics (1989-1993)*, Vol. 2, No. 8, pp. 1429–1436, 1990.
- [105] A. Luton, S. Ragab, and D. Telionis, "Interaction of spanwise vortices with a boundary layer", *Physics of Fluids*, Vol. 7, No. 11, pp. 2757–2765, 1995.
- [106] F. Holzäpfel, T. Hofbauer, D. Darracq, H. Moet, F. Garnier, and C. F. Gago, "Analysis of wake vortex decay mechanisms in the atmosphere", *Aerospace Science and Technology*, Vol. 7, No. 4, pp. 263–275, 2003.
- [107] T. Doligalski, C. Smith, and J. Walker, "Vortex interactions with walls", *Annual Review of Fluid Mechanics*, Vol. 26, No. 1, pp. 573–616, 1994.
- [108] F. Holzäpfel, T. Gerz, and R. Baumann, "The turbulent decay of trailing vortex pairs in stably stratified environments", *Aerospace Science and Technology*, Vol. 5, No. 2, pp. 95–108, 2001.
- [109] N. A. Nash'at and F. H. Proctor, "Estimation of eddy dissipation rates from mesoscale model simulations", *AIAA paper 2012-0429*, 2012.

-
- [110] S. C. Crow, "Stability theory for a pair of trailing vortices", *AIAA journal*, Vol. 8, No. 12, pp. 2172–2179, 1970.
- [111] S. Crow and E. Bate, "Lifespan of trailing vortices in a turbulent atmosphere", *Journal of Aircraft*, Vol. 13, No. 7, pp. 476–482, 1976.
- [112] T. Misaka, F. Holzäpfel, I. Hennemann, T. Gerz, M. Manhart, and F. Schwertfirm, "Vortex bursting and tracer transport of a counter-rotating vortex pair", *Physics of Fluids*, Vol. 24, No. 2, p. 025104, 2012.
- [113] R. Scorer and L. Davenport, "Contrails and aircraft downwash", *Journal of Fluid Mechanics*, Vol. 43, No. 03, pp. 451–464, 1970.
- [114] F. Holzäpfel and T. Gerz, "Two-dimensional wake vortex physics in the stably stratified atmosphere", *Aerospace science and technology*, Vol. 3, No. 5, pp. 261–270, 1999.
- [115] P. R. Spalart, "Airplane trailing vortices", *Annual Review of Fluid Mechanics*, Vol. 30, No. 1, pp. 107–138, 1998.
- [116] F. Köpp, S. Rahm, and I. Smalikho, "Characterization of aircraft wake vortices by 2- μ m pulsed doppler lidar", *Journal of Atmospheric and Oceanic Technology*, Vol. 21, February 2004.
- [117] F. Köpp, S. Rahm, I. Smalikho, A. Dolfi, J.-P. Cariou, M. Harris, and R. I. Young, "Comparison of wake-vortex parameters measured by pulsed and continuous-wave lidars", *Journal of aircraft*, Vol. 42, No. 4, pp. 916–923, 2005.
- [118] D. Engelbart, H. Steinhagen, U. Görsdorf, J. Neisser, H. Kirtzel, and G. Peters, "First results of measurements with a newly-designed phased-array sodar with RASS", *Meteorology and Atmospheric Physics*, Vol. 71, No. 1-2, pp. 61–68, 1999.
- [119] G. Peters, H. Timmermann, and H. Hinzpeter, "Temperature sounding in the planetary boundary layer by RASS-system analysis and results", *International Journal of Remote Sensing*, Vol. 4, No. 1, pp. 49–63, 1983.
- [120] M. Frech, F. Holzäpfel, A. Tafferter, and T. Gerz, "High-resolution weather database for the terminal area of Frankfurt Airport", *Journal of Applied Meteorology and Climatology*, Vol. 46, No. 11, pp. 1913–1932, 2007.
- [121] C. Donaldson and A. Bilanin, "Vortex wakes of conventional aircraft", Tech. Rep. 204, Advisory Group for Aerospace Research and Development (AGARD), Neuilly sur Seine, France, May 1975.
- [122] D. Hinton, "An aircraft vortex spacing system (AVOSS) for dynamical wake vortex spacing criteria", report 19970014118, NASA, 1996.
- [123] T. J. Dasey, R. E. Cole, R. Heinrichs, M. Matthews, and G. Perras, "Aircraft vortex spacing system (AVOSS) initial 1997 system deployment at Dallas/Ft. Worth (DFW) airport", Tech. Rep. L-3, NASA/A-1, Massachusetts Inst. of Tech.; Lincoln Lab.; Lexington, MA United States, 1998.

- [124] R. P. Dougherty, F. Y. Wang, E. R. Booth, M. E. Watts, N. Fenichel, and R. E. D'Errico, "Aircraft wake vortex measurements at Denver International Airport", *AIAA Paper*, Vol. 2880, No. 10, 2004.
- [125] M. J. Pruis, D. P. Delisi, N. N. Ahmad, and F. H. Proctor, "Atmospheric turbulence estimates from a pulsed lidar", in *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, Vol. 10, pp. 6–2013, 2013.
- [126] D. Thomas Hofbauer and F. Holzäpfel, "Behaviour of aircraft wake vortices subjected to wind shear", in *21st AIAA Applied Aerodynamics Conference*, 2003.
- [127] T. Sarpkaya, "New model for vortex decay in the atmosphere", *Journal of Aircraft*, Vol. 37, No. 1, pp. 53–61, 2000.
- [128] R. E. Robins, D. P. Delisi, and G. C. Greene, "Development and validation of a wake vortex predictor algorithm", *AIAA Paper*, No. 98-0665, 1998.
- [129] G. F. Switzer and F. H. Proctor, "Numerical study of wake vortex behavior in turbulent domains with ambient stratification", *AIAA paper*, No. 2000-0755, 2000.
- [130] F. Holzäpfel, "Probabilistic two-phase wake vortex model." Presentation at WakeNET New Orleans, April 2004.
- [131] A. P. Dempster, "A generalization of Bayesian inference", *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 205–247, 1968.
- [132] M. I. Jordan and R. A. Jacobs, "Hierarchical mixtures of experts and the EM algorithm", *Neural computation*, Vol. 6, No. 2, pp. 181–214, 1994.
- [133] C. M. Bishop and M. Svenskn, "Bayesian hierarchical mixtures of experts", in *Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence*, pp. 57–64, Morgan Kaufmann Publishers Inc., 2002.
- [134] A. W. Robertson, U. Lall, S. E. Zebiak, and L. Goddard, "Improved combination of multiple atmospheric GCM ensembles for seasonal prediction", *Monthly Weather Review*, Vol. 132, No. 12, pp. 2732–2744, 2004.
- [135] C. Tebaldi, R. L. Smith, D. Nychka, and L. O. Mearns, "Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles", *Journal of Climate*, Vol. 18, No. 10, pp. 1524–1540, 2005.
- [136] T. N. Krishnamurti, C. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, "Multimodel ensemble forecasts for weather and seasonal climate", *Journal of Climate*, Vol. 13, No. 23, pp. 4196–4216, 2000.
- [137] C. Ziehmann, "Comparison of a single-model eps with a multi-model ensemble consisting of a few operational models", *Tellus A*, Vol. 52, No. 3, pp. 280–299, 2000.

-
- [138] R. Buizza, P. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, "A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems", *Monthly Weather Review*, Vol. 133, No. 5, pp. 1076–1097, 2005.
- [139] J. Fritsch, J. Hilliker, J. Ross, and R. Vislocky, "Model consensus", *Weather and forecasting*, Vol. 15, No. 5, pp. 571–582, 2000.
- [140] K. Pearson, "Note on regression and inheritance in the case of two parents", *Proceedings of the Royal Society of London*, pp. 240–242, 1895.
- [141] E. Kalnay and A. Dalcher, "Forecasting forecast skill", *Monthly weather review*, Vol. 115, No. 2, pp. 349–356, 1987.
- [142] T. W. Barker, "The relationship between spread and forecast error in extended-range forecasts", *Journal of climate*, Vol. 4, No. 7, pp. 733–742, 1991.
- [143] J. S. Whitaker and A. F. Lough, "The relationship between ensemble spread and ensemble mean skill", *Monthly weather review*, Vol. 126, No. 12, pp. 3292–3302, 1998.
- [144] F. Holzäpfel and R. E. Robins, "Probabilistic two-phase aircraft wake vortex model: application and assessment", *Journal of Aircraft*, Vol. 41, No. 5, pp. 1117–1126, 2004.
- [145] S. Körner and F. Holzäpfel, "Multi-model ensemble wake vortex prediction", *Aircraft Engineering and Aerospace Technology*, Vol. 88, No. 2, pp. 331–340, 2016. doi:10.1108/AEAT-02-2015-0068.
- [146] F. Holzäpfel, "Effects of environmental and aircraft parameters on wake vortex behavior", *Journal of Aircraft*, Vol. 51, No. 5, pp. 1490–1499, 2013.
- [147] I. J. Myung, "Tutorial on maximum likelihood estimation", *Journal of mathematical Psychology*, Vol. 47, No. 1, pp. 90–100, 2003.
- [148] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.
- [149] T. K. Moon, "The expectation-maximization algorithm", *Signal processing magazine, IEEE*, Vol. 13, No. 6, pp. 47–60, 1996.
- [150] G. McLachlan and D. Peel, *Finite mixture models*. John Wiley & Sons, 2004.
- [151] S. Körner, N. A. Nash'at, F. Holzäpfel, and R. L. VanValkenburg, "Multimodel ensemble methods for prediction of wake-vortex transport and decay", *Journal of Aircraft*, 2017.
- [152] Y. Pomeau and P. Manneville, "Intermittent transition to turbulence in dissipative dynamical systems", *Communications in Mathematical Physics*, Vol. 74, pp. 189–197, 1980.
- [153] C. Räth, I. Laut, and K. Schreiber, "Phase correlations in leptokurtic time series", in *Dynamics Days Europe 2015, Exeter*, 2015.

- [154] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning", *Machine learning*, Vol. 50, No. 1-2, pp. 5–43, 2003.
- [155] E. N. Lorenz, "Deterministic nonperiodic flow", *Journal of the atmospheric sciences*, Vol. 20, No. 2, pp. 130–141, 1963.
- [156] S. Körner, N. Ahmad, F. Holzäpfel, and R. L. VanValkenburg, "Multi-model ensemble wake vortex prediction", in *7th AIAA Atmospheric and Space Environments Conference*, p. 3173, 2015.
- [157] R. N. Hoffman and E. Kalnay, "Lagged average forecasting, an alternative to monte carlo forecasting", *Tellus A*, Vol. 35, No. 2, pp. 100–118, 1983.
- [158] Č. Branković, T. Palmer, F. Molteni, S. Tibaldi, and U. Cubasch, "Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting", *Quarterly Journal of the Royal Meteorological Society*, Vol. 116, No. 494, pp. 867–912, 1990.
- [159] F. Holzäpfel, A. Stephan, S. Körner, and T. Misaka, "Wake vortex evolution during approach and landing with and without plate lines", in *52nd Aerospace Sciences Meeting, National Harbor, Maryland*, 2014.
- [160] A. Nuic, "User manual for the base of aircraft data (bada) revision 3.11", *Atmosphere*, Vol. 2010, p. 001, 2010.
- [161] F. Holzäpfel, K. Dengler, T. Gerz, and C. Schwarz, "Prediction of dynamic pairwise wake vortex separations for approach and landing", *AIAA Paper*, Vol. 3037, No. 2011, p. 15, 2011.
- [162] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to probability*, Vol. 1. Athena Scientific Belmont, MA, 2002.

Acknowledgments

This thesis has been accomplished at the German Aerospace Center (DLR), Institute of Atmospheric Physics in Oberpfaffenhofen, Germany where I have been working as doctoral student. My gratitude first of all goes to my mentor Frank Holzäpfel who initiated the wake vortex NASA-DLR collaboration, who supported me continuously with his advice, reviewed my publications and who gave me the opportunity to collect experience during several international conferences and workshops. Furthermore, I would like to thank Prof. Eike Stumpf for supervising this thesis, for offering short-termed appointments and for his technical, job-related and always friendly advice. I also thank Prof. Christian Breitsamter who agreed on being my co-supervisor and contributed with his comments to the success of this project. I am also obliged to Nash'at Ahmad, Fred Proctor, Randal VanValkenburg and Fanny Limon Duparcmeur from the NASA Langley Research Center for their continuous support, for the provision of the NASA wake vortex models and measurements and for the organization of the NASA-DLR wake-vortex workshop. They also welcomed me warmly during my stay at their center for two weeks. I am also obliged to Thomas Gerz, head of our department, for his suggestions, critical comments and his support in my research at international conferences. I am also very grateful to the ideas and thoughts Ingo Sölch contributed in numerous discussions. Not to forget are my dear colleagues Anton Stephan, Martin Köhler, Dennis Stich, Simon Unterstrasser, Johannes Wagner, Martin Hagen, Ayla Augst and Mareike Kenntner who supported me with their ideas and distracted me in a welcome way during breaks. Furthermore, I very much appreciated the help of Windfried Beer with the IT.

Last but not least I want to thank my family who supported me on this way. I feel grateful to my mother who encouraged me to follow my own paths, to my father from whom I inherited the enthusiasm for natural sciences, my younger sister for not receiving her doctorate before I do and for her and Marco believing in my endurance, to Rupp for consistently providing me with aeronautical articles and to Sabine for being my personal physician. I also want to thank my grandfather who already enthused me with engineering at an early age. Further gratitude goes to Joanna for her continuous encouragement, understanding and stunning support all along.