Influences of surface layer modelling approaches on model performance

Dissertation zur Erlangung
des Doktorgrades der Naturwissenschaften
im Department Geowissenschaften der
Universität Hamburg

Vorgelegt von
Hinnerk Ries
aus
Hamburg

Hamburg
2010

Als Dissertation angenommen vom
Department Geowissenschaften der Universität Hamburg

Auf Grund der Gutachten von:
Prof. Dr. K. Heinke Schlünzen und
Prof. Dr. Martin Claußen

Hamburg, den 7. Juli 2010

Prof. Dr. Jürgen Oßenbrügge
Leiter des Department Geowissenschaften

Zusammenfassung

In dieser Arbeit wird der Einfluss der Modellierung der Prandtl-Schicht auf die Entwicklung atmosphärischer Phänomene untersucht. Die Hauptzielsetzung ist, gut geeignete Modelleinstellungen für die Prandtl-Schicht zu spezifizieren. Die Spezifikationen basieren auf Sensitivitätsstudien zur Anzahl atmosphärischer vertikaler Schichten, zur Wahl des Landoberflächenmodells und zu verschiedenen Oberflächenparametern für eine Vielzahl von atmosphärischen Situationen. Zu einem gewissen Grad werden alle mesoskaligen Systeme dem synoptisch-skaligen Antrieb ausgesetzt. Da in dieser Arbeit nur realistische Szenarien untersucht werden, wird die Modellgüte in Relation zur Stärke der synoptisch skaligen Situation bewertet. Außerdem wird die Modellgüte des antreibenden Modells und die Stärke des Antreibens durch Anfangs- und Randwerte berücksichtigt. Folglich ist eine weitere Zielsetzung dieser Arbeit, eine Empfehlung über eine passende Antriebsstärke in Abhängigkeit von der Stärke der synoptisch skaligen Situation zu geben.

Die Auswirkungen der verwendeten Anzahl von atmosphärischen vertikalen Schichten und des benutzten Landoberflächenmodells werden mit zwei unterschiedlichen Modelleinstellungen des Mesoskaligen Modells der fünften Generation der Pennsylvania State University und dem National Center for Atmospheric Research (MM5) für die Bucht von Valencia untersucht. In dieser Region ist die vorherrschende mesoskalige Zirkulation der Land-See-Wind. Modelleinstellung M1 ist durch 34 atmosphärische vertikale Sigma-Schichten, kombiniert mit dem Fünfschicht Land-Oberflächenmodell (5L LSM) gekennzeichnet. Modelleinstellung M2 hat 52 Schichten und das Noah Land-Oberflächenmodell (Noah LSM) wird verwendet. Ergebnisse von 71 Vorhersagen aus 2006 werden an Hand von Messungen oberflächennaher meteorologischer Standardparameter evaluiert. Mit der Modelleinstellung M2 verbessert sich die Genauigkeit für alle meteorologischen Parameter mit Ausnahme des Drucks. Eine ausführliche Analyse der Resultate für 4 Tage, die mit allen Kombinationen der vertikalen Auflösungen und des LSMs simuliert werden, zeigt, dass die Zunahme der vertikalen Auflösungen für die verbesserte Vorhersage verantwortlich ist. Das Noah LSM verbessert generell die Vorhersagegüte der Temperatur, verringert aber die Güte aller anderen Parameter, besonders für Tage mit See-Wind.

Nach dieser allgemeinen Auswertung zweier Landoberflächenmodelle wird eine ausführlichere Untersuchung über den Einfluss einzelner Oberflächenparameter durchgeführt. Die Meereis Eigenschaften Meereis-Verteilung, Rauhigkeit, Temperatur und Wärmeleitfähigkeit werden innerhalb des Bereiches der Beobachtungsungenauigkeit variiert. Ihre Auswirkungen werden an Hand eines sich aufeisig bewegenden Trogs in der Fram Straße vom 7 März 2002 untersucht. Die Situation wird mit dem mesoskaligen Transport- und Strömungsmodell METRAS simuliert. Als großskaliger Antrieb werden Reanalysen des "Europäischen Zentrums für mittelfristige Wettervorhersage" (ECMWF) in zwei unterschiedlichen Antriebsstärken verwendet. Die Reanalysen

zeigen jedoch einige Mängel in der Darstellung der bodennahen Inversion und der Meereis-Verteilung. Um die ECMWF Reanalysen geeignet auf das METRAS Gitter zu interpolieren, wird deshalb eine neue Vertikalinterpolation entwickelt. Die Testfälle werden anhand von Flugzeugmessungen bewertet, die innerhalb des "Fram Straßen-Zyklonen Experiments 2002" durchgeführt worden sind. Die Erkentnisse werden mit Hilfe des Durchzugs einer Zyklonenfamilie am 13. März 2002 auf andere als synoptischskalige Situationen verallgemeinert.

Den stärksten Einfluss auf die Güte eines atmosphärischen Ausschnittsmodells hat die Stärke des großskaligen Antriebs über die offenen Ränder des Modellgebietes. Die wichtigste Meereis Eigenschaft ist die Verteilung, gefolgt vom Einfluss der Meereis Temperatur und der Rauhigkeit. Die Wärmeleitfähigkeit des Meereises spielt nur eine untergeordnete Rolle.

Als Verbindung zwischen Oberflächeneigenschaften und oberflächennahen Prozessen sind die turbulenten Flüsse von zentraler Bedeutung. Sie werden für eine Situation mit abeisiger Strömung in der Fram Straße für den 15. März 2002 untersucht, indem die simulierten Oberflächenflüsse mit den Flüssen verglichen werden, die in Flugzeughöhe gemessen wurden. Mit ihrer vergleichbar guten Simulationsqualität ist ein starker Hinweis gegeben, dass die meteorologischen Standardparameter auf Grund von realistischen oberflächennahen Prozessen realistisch simuliert werden und folglich die Oberflächenparameter gut gewält sind.

Abstract

In this thesis the impact of surface layer modelling on the development of atmospheric phenomena is investigated. The main objective is to determine well suited model setups for the surface layer. The specifications are based on sensitivity studies of the number of atmospheric vertical levels, on the choice of the land surface model and on various ground parameters for a variety of atmospheric situations. To some extent mesoscale systems are always exposed to the synoptic scale forcing. As in this thesis only real-case scenarios are investigated, the model performance is assessed in relation to the dominating scale of the meteorological situation, its representation in the driving model and its consideration by the mesoscale model in form of initial values and boundary value nudging. Thus, another objective of this thesis is to give a recommendation on an appropriate nudging strength in dependence of the dominating scale of the meteorological situation.

The impact of the number of atmospheric vertical levels and of the land surface model used is investigated with two different setups of the fifth-generation Pennsylvania State University–NCAR Mesoscale Model (MM5) for the bay of Valencia. In this region the predominant mesoscale circulation is the land-see-breeze. Setup M1 is characterised by 34 atmospheric vertical sigma levels combined with a five-layer land surface model (5L LSM). M2 has 52 levels and the Noah LSM is used. Results of 71 forecasts of 2006 are evaluated with a focus on the accuracy of meteorological surface data. The accuracy improves for all meteorological parameters except pressure with the M2 setup. Detailed analysis of results for 4 days simulated with all combinations of vertical resolutions and LSMs shows that the increase of vertical resolution is responsible for the improved forecast performance. The Noah LSM generally improves temperature performance, but reduces the performance of all other parameters, especially for days with sea-breeze conditions.

After this general evaluation of two land surface models, a more detailed investigation is performed on the influence of single surface parameters. The sea-ice characteristics sea-ice distribution, roughness, temperature and thermal conductivity are varied within the range of observational uncertainty. Their impacts are studied by investigating an on-ice moving trough in the Fram Strait for 7th March 2002. The situation is simulated with the mesoscale transport and flow model METRAS. The forcing is taken from "European Centre for Medium-Range Weather Forecasts" (ECMWF) reanalyses and is applied with two different nudging intensities. The reanalyses, however, show some deficiencies in the representation of near-ground inversions and of the sea-ice distribution. A new vertical interpolation is therefore developed to properly transfer ECMWF reanalyses to the METRAS grid. The test cases are evaluated against aircraft measurements taken within the "Fram Strait Cyclone Experiment 2002". The findings are generalised to other than synoptic scale

situations by simulating the passage of a cyclone family through the same domain on 13th March 2002.

The nudging intensity shows the strongest effect on the limited area model performance. The most important sea-ice characteristic is the sea-ice distribution followed by sea-ice temperature and roughness. Thermal conductivity of the sea ice played only a minor role.

As the link between surface characteristics and near-ground processes the turbulent fluxes are of primary importance. They are investigated for an off-ice flow situation in the Fram Strait for 15th March 2002 by assessing the simulated surface fluxes with the fluxes measured at aircraft height. With their comparably good simulation quality, a strong hint is given that the standard meteorological parameters are simulated realistically as consequence of realistic near-ground processes and hence well chosen surface parameters.

Contents

1. Introduction

Land surfaces play an important role for the development of mesoscale systems because their heterogeneities sometimes lead to horizontal flux divergencies of heat, momentum or radiation. Heterogeneities of the heat fluxes are able to induce mesoscale circulations, if spatial structure, amplitude and mean wind velocity are adequate (Raasch and Harbusch, 2001). A mesoscale circulation is the more pronounced the larger the surface pattern and the horizontal heat flux divergence is. Lower wind speeds favour the development, as horizontal mixing of up- and downdrafts as well as surface friction are reduced. Wind directions parallel to the axes of heterogeneities, e.g. coastlines, are also favourable, as organization of up- and downdrafts is enhanced.

In mesoscale models land surfaces are represented by numerous thermodynamic and dynamic parameters, e.g. ground temperature, -wetness, vegetation or surface roughness. Most of the surface parameters are strongly influenced by processes at the ground, such as storage or release of heat and water, change of roughness by deformation (e.g. sea ice), production of water waves, or change in albedo by freezing or change in vegetation. The time scales of the processes at the ground differ from years to minutes. The development of the rapidly changing surface parameters is therefore simulated by coupled land surface models of different complexity. However, initialization of some surface parameters is still difficult, as observations exist only at coarse temporal or spatial resolution and might additionally contain some uncertainty.

The aim of this thesis is to give an estimation of the influences exhibited by surface heterogeneities compared to the synoptic scale forcing on mesoscale model performances. Surface heterogeneities resulting from spatially differing ground characteristics and parameterizations applied are assessed with regard to their impact on model performance. This shall help to estimate the uncertainties introduced by insufficient knowledge of the ground parameters or uncertain parameterizations.

A sensitivity study on the choice of the land surface model is performed with the mesoscale model MM5 for the bay of Valencia in Chapter 2. This region is characterized by land-ocean-contrasts that lead to land-sea-breezes in many of the investigated 71 forecasts. As a next step the impacts of single ground parameters are investigated

under the Arctic conditions of the Fram Strait. The domain is characterized by sea-ice-ocean-contrasts and sensitivity studies on the characteristics sea-ice distribution, sea-ice temperature, sea-ice thermal conductivity and sea-ice surface roughness are performed in Chapter 4. The model applied for this region is the mesoscale transport and fluid model METRAS.

According to their forcing, mesoscale atmospheric phenomena can be categorized into two classes (Browning, 1980). The first class consists of systems forced by surface heterogeneities. The second class consists of synoptically induced mesoscale systems. However, the categorization is not necessarily sharp, as systems of the second class can be modified by surface heterogeneities. Especially in real case scenarios, mesoscale systems are always to some extent exposed to the synoptic scale forcing. Therefore, the influence of synoptic scale forcing on the mesoscale model performance is investigated by applying differently strong nudging. Nudging is a realization of nesting the mesoscale model into a larger scale and coarser resolved analysis. The high resolved solution is pulled towards time dependent boundary values provided by the coarser resolved analysis (Bungert, 2008). The influences of the nudging strength and the surface heterogeneities are compared in Chapters 4 and 5 to give a judgement of their relative importance in dependence of the synoptic scale situation.

In addition to the investigations on the influence of surface characteristics and nudging strength, this work shall also contribute to the development of a validation strategy. Not alone the both complementary skill scores, the bias sensitive hit rate and the pattern sensitive Gandin-Murphy skill score (Appendices B and C), but also the numerous kinds of statistical plots provide an opportunity to evaluate large numbers of test cases. A large number of test cases is quite often necessary to achieve some statistical reliability.

The direct link between the simulation of the near-ground meteorology and the representation of the ground properties are the vertical, near-surface turbulent fluxes that exchange heat, moisture or momentum. The connection between the turbulent fluxes of momentum, sensible heat, and latent heat and meteorological parameters in METRAS simulations as well as in reality is investigated for an off-ice flow and a cyclone family in the Fram Strait in Chapter 5.

Regional peculiarities of the Arctic, e.g. shallow inversions and sea-ice distributions strongly varying in time require a modified nudging, which on the one hand assumes near-ground inversions in the vertical interpolation, and, on the other hand disregards surface temperatures of the driving model, as they are often affected by imprecise sea-ice distributions. The effects of the nudging adapted to Arctic conditions is presented in Chapter 3 with the same synoptic scale dominated trough passage in the Fram Strait as investigated in Chapter 4. Chapters 2 and 4 are published papers (Ries

and Schlünzen, 2009; Ries et al., 2010). In Chapter 6 conclusions are drawn and an outlook to future work is given.

2. Evaluation of a mesoscale model with different surface parameterizations and vertical resolutions for the bay of Valencia

Since the chapter has already been published, the notation may sometimes deviate from the one used in the other chapters of this thesis. The published version is in US English, the spelling thus differs from the chapters where British English is used.

Summary Two different setups of the fifth-generation Pennsylvania State University NCAR Mesoscale Model (MM5) are used and the results of 71 forecasts are evaluated with a focus on the accuracy of meteorological surface data including wind, temperature, dewpoint temperature, and pressure. The forecasts, which cover the Valencia coastal region with a $3-km$ grid, are nested into two coarser grids of 9- and $27-km$ mesh size. The integration time is 18 h. The ability to correctly simulate the sea breeze plays a major role for this area. Different model setups are used: in the M1 simulations 34 atmospheric vertical sigma levels combined with a five-layer land surface model (5L LSM) are applied for the first half of the year 2006. In total, 52 atmospheric vertical sigma levels and the Noah LSM are used for the second half of 2006 (setup M2). The accuracy improves for all meteorological parameters except pressure with the M2 setup. However, negative biases of temperature, dewpoint temperature, and wind speed worsen. Detailed analysis of results for 4 days simulated with all combinations of vertical resolutions and LSMs shows that the increase of vertical resolution is responsible for the improved forecast performance. The Noah LSM generally improves temperature performance, but reduces the performance of all other parameters, especially for days with sea-breeze conditions. The Noah LSM performs better than the 5L LSM scheme only for a day driven by synoptic changes.

2.1. Introduction

The goal of this paper is to evaluate the performance of two setups of the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) for a complex coastal area and to give recommendations on a suitable setup to be used for short-term regatta forecasts. The region studied is the coastal area of Valencia (Spain), where the 2007 America's Cup took place. The simulation results analyzed were produced as wind forecasts for the German America's Cup Team in their training phase 2006. They used the forecasts for the daily planning of the training and for regatta tactics. The daily planning included the choice of the sails to carry on board. Because of the model's horizontal resolution of $3\,km$ the tactical aid of the forecasts was limited to the choice of the side while beating upwind. The forecasts were delivered in text form and as vector wind maps.

The wind regimes of the studied region (Fig. 2.1) are considerably influenced by mesoscale circulations resulting from thermal differences. The region consists of an orographically strongly structured coast with an about 1000-m-higher interior. The sea breeze, as the most important thermally driven wind regime during the day, is characterized by a pronounced wind shift as the sea breeze front moves inland. This circulation interacts with mountain–valley wind circulations that result in quite local flow features (Fock and Schlünzen 2009, manuscript submitted to *Int. J. Climatol.*).

Low-level jets that are bound to the zone of high baroclinicity of the front sometimes occur (Kottmeier et al., 2000). Focus of the evaluation performed in this paper is on the wind forecast. A wide range of synoptical situations are investigated. In total, 71 forecasts unequally distributed over all seasons of 2006 are compared to routine measurements. However, the seasonal influence is reduced by concentrating on days with sea-breeze circulation potential. The forecasts are divided into two sets of setups. They differ in the number of atmospheric vertical levels and the land surface model (LSM) used. These two model properties were chosen, because they primarily influence what characterizes and drives the sea breeze: vertical structure and the energy balance at ground. For the first half of 2006, 34 atmospheric vertical sigma levels and the 5-layer soil temperature model with constant moisture availability (5L LSM; Dudhia (1996)) were used. For the second half of 2006, 52 atmospheric vertical sigma levels and the Noah LSM are used. The Noah LSM is more sophisticated in its treatment of thermodynamics and hydrology (Chen and Dudhia, 2001a) and therefore is expected to capture the comparatively large soil moisture variability in semiarid climates better. As shown by Seneviratne et al. (2006), the coupling of soil moisture and atmospheric temperature is exceptionally strong for the Mediterranean. Therefore, the choice of the LSM should have a strong impact on the forecast quality.
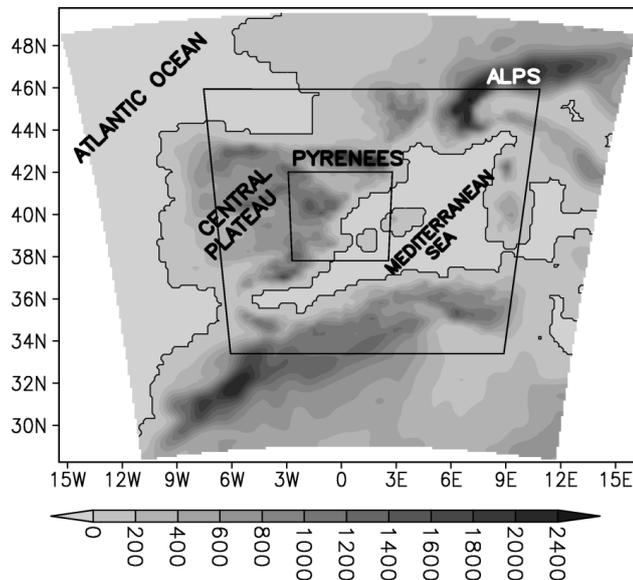
Figure 2.1.: Geographic location of DOM1 (shaded, $27 - km$ resolution), DOM2 (intermediate box, $9 - km$ resolution), and DOM3 (innermost box, $3 - km$ resolution). Location of DOM3 as used from 13 Jun 2006 onward. Shading and contour line show the orography (m MSL) and coastline as used for DOM1.

Previous studies on the evaluation and the influence of the used LSM on MM5 results (Chen and Dudhia, 2001b; Miao et al., 2007; Tonnesen and Coauthors, 2005; Olerud and Sims, 2003), and of the number of vertical levels (Beran et al., 2005; Colby, 2004) were carried out. Several MM5 studies have concentrated on sea-breeze development (Colle et al., 2003; Hong and Pan, 1996; Warner et al., 2003; Yang et al., 2005). However, only few (e.g., Hong (2003)) have performed a long-term evaluation not focusing on certain meteorological conditions. We expand the analyses to higher vertical resolution (except two case studies of Beran et al. (2005)) and base them on many cases in the forecast mode.

In section 2.2 the model setup is discussed and in section 2.3 the meteorological situations are described in detail. The evaluation method is given in section 2.4 and in section 2.5 the results are given. Conclusions are drawn in section 2.6.

2.2. Model setup

The nonhydrostatic version of the MM5, version 3.7 (Grell et al., 1995), was used for the forecasts. Three nests, centered approximately at $38° N$, $0° E$ and with mesh sizes of 27, 9, and 3 km, and an increasing number of grid points (85, 154, and 154 in both longitudinal and latitudinal directions) are applied (Fig. 2.1). The domains are named DOM1 ($27 - km$ resolution), DOM2 ($9 - km$ resolution), and DOM3 ($3 - km$ resolution) hereafter. The time step of DOM3 is 9 s. Interactions of the meteorological fields between the domains are accounted for by two-way nesting, using one-point feedback with a smoother– desmoother. We used the Nonhydrostatic Mesoscale Model (NMM) dynamics solver. The physical parameterizations selected are the medium-range forecast scheme (MRF) for turbulence parameterization (Hong and Pan, 1996); the Reisner graupel scheme for cloud microphysical processes (Reisner et al., 1998); the National Center for Atmospheric Research (NCAR) Community Climate System Model, version 2 (CCM2) for radiation (Hack et al., 1993); and a cumulus parameterization was applied for DOM1 and DOM2 (Kain–Fritsch 2; Kain (2004)). All schemes are commonly used (e.g., Wang et al. (2007); Foy et al. (2006)).

DOM1 is initialized and 3-hourly forced by meteorological fields derived from the Global Forecast System (GFS) with 1° resolution (Kalnay et al., 1990). The sea surface temperature (SST) is fixed to the daily mean field of the GFS forecast. All simulations start for 0000 UTC and last 18 h. The first 3 h are neglected as spinup time, results for 0300–1800 UTC are evaluated.

Two different model setups are used during the discontinuous forecasting period. From 1 February to 1 June 2006, 34 levels with the lowest atmospheric level ap-

proximately $36\,m$ AGL and an enhanced resolution in the lower troposphere (9 levels below $1000\,m$ over the sea), and the 5L LSM (Dudhia, 1996) are applied (M1 setup hereafter).

In the 5L LSM, the skin and soil temperature are diagnosed on basis of the energy budget at the surface including radiation, sensible and latent heat flux, and ground heat flux (Deardorff, 1978). The ground heat flux depends on the skin temperature, thermal inertia, soil moisture, and deep soil temperature. Thermal inertia and soil moisture are seasonally fixed values depending on the land use. The land use is derived from the Global Land Cover Characteristics (GLCC) datasets measured by the U.S. Geological Survey (USGS; see online at http:// edc2.usgs.gov/glcc/) in 1992 (Dudhia, 1996). The daily averaged skin temperature obtained from GFS forecasts is used as deep soil temperature in a depth of $47\,cm$. Vegetation effects are not explicitly treated.

Starting 1 June 2006 until 12 December, 52 levels with the lowest level approximately $6\,m$ AGL (27 levels below $1000\,m$ over the sea) and the Noah LSM (Chen and Dudhia, 2001a) are applied (M2 setup hereafter). The Noah LSM predicts soil temperature and soil moisture in four layers in dependence of land use. It considers a canopy layer and determines the surface skin temperature with a single linearized surface energy balance equation. For the canopy layer several vegetation properties are also derived from the GLCC landuse dataset. The Noah LSM uses GLCC datasets with soil characteristics and vegetation fraction mainly for the description of the hydrology. Soil moisture and soil temperature are initialized by actual fields for two layers with depths of 0–$10\,cm$ and 10–$200\,cm$ also obtained from GFS forecasts. DOM3 of M2 is shifted approximately $0.8°$ to the north and $0.4°$ to the west compared to M1. To separately investigate LSM and vertical resolution impacts, four cases are recomputed in all possible combinations of the two LSMs and the two choices of vertical level numbers.

2.3. Meteorological conditions and comparison data

The forecasts were performed for the coastal region of eastern Spain. The region is orographically strongly structured with the main valleys of rivers Ebro, Turia, and Jucar. The interior has mountains with a height of more than $1000\,m$ (Fig. 2.2).

During summer the large-scale pressure pattern of the area is dominated by the Azores high pressure system so that Atlantic frontal systems are mostly moving more to the north, to the European Alps. The anticyclonic circulation sometimes leads to a near-ground jet that is situated between the Pyrenees and the European

8
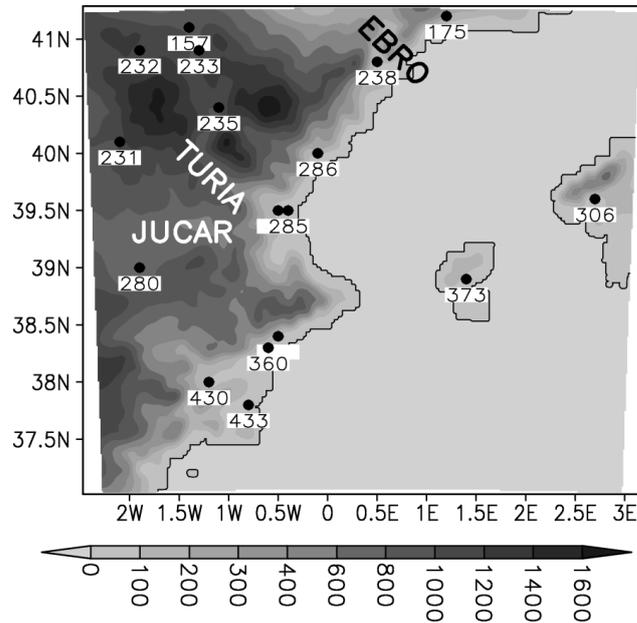
Figure 2.2.: Orography shaded (m MSL) with the three main valleys of rivers Ebro, Turia, and Jucar in DOM3. Dots mark stations of the Spanish national weather service at Valencia Airport with number 285. Selected stations are at least three grid points away from the lateral boundary of DOM3.

Alps into the Gulf of Lyons. It is associated with the Tramontana and Mistral winds (Gangoiti et al., 2001). This large-scale structure is modified by diurnal mesoscale circulations such as sea breezes, valley winds, and katabatic flows. These circulations result from the strong solar insolation causing large coastal temperature gradients. The circulations are additionally affected by the mountainous orography. Over the Iberian Peninsula the associated convective systems often accumulate to shortlived thermal lows with compensatory subsidence areas over the Mediterranean Sea (Millan et al., 1997; Perez-Landa et al., 2006). The usually northerly Tramontana and Mistral winds are often deflected to the east by these subsidence areas. Thunderstorms develop, if the convection is supported by advected cold air and/or small troughs aloft, or by cold fronts (Romero et al., 2001).

During winter, when the Azores high pressure system weakens, frontal systems are able to penetrate from the Atlantic Ocean, and Iberia is one of the main routes for them (Trigo et al., 1999). Then mesoscale cyclones also form south of the European Alps. Latent heat release and lee cyclogenesis are the principle mechanisms as continental or modified maritime-polar air is advected over the warm Mediterranean Sea (Millan et al., 2002; Trigo et al., 1999). The weaker mesoscale activity and increased importance of synoptic-scale cyclones in winter is reflected in the pressure time series in Fig. 2.3. The high frequent and small pressure changes are predominant from May

to October, while low-frequent and large pressure changes occur during the rest of the year.

The different impacts of larger-scale and mesoscale weather influences are well reflected in synoptic station measurements (e.g., for pressure as presented in Fig. 2.3). Therefore, model evaluation will be based on the routine measurements from the National Weather Service. As a common practice, the pressure of stations located more than $750\,m$ MSL is not reduced to sea level pressure and disregarded in the evaluation. Measurements below $980\,hPa$ are classified as faulty, because no low pressure systems with these pressure values were occurring. Wind measurements at wind speeds below $1\,m\,s^{-1}$ are neglected in the evaluation as the measurement uncertainty, especially for the wind direction, is too high. Note that the wind velocity is originally given in discrete knots. This implies an uncertainty of $\pm 0.5\,kt$,
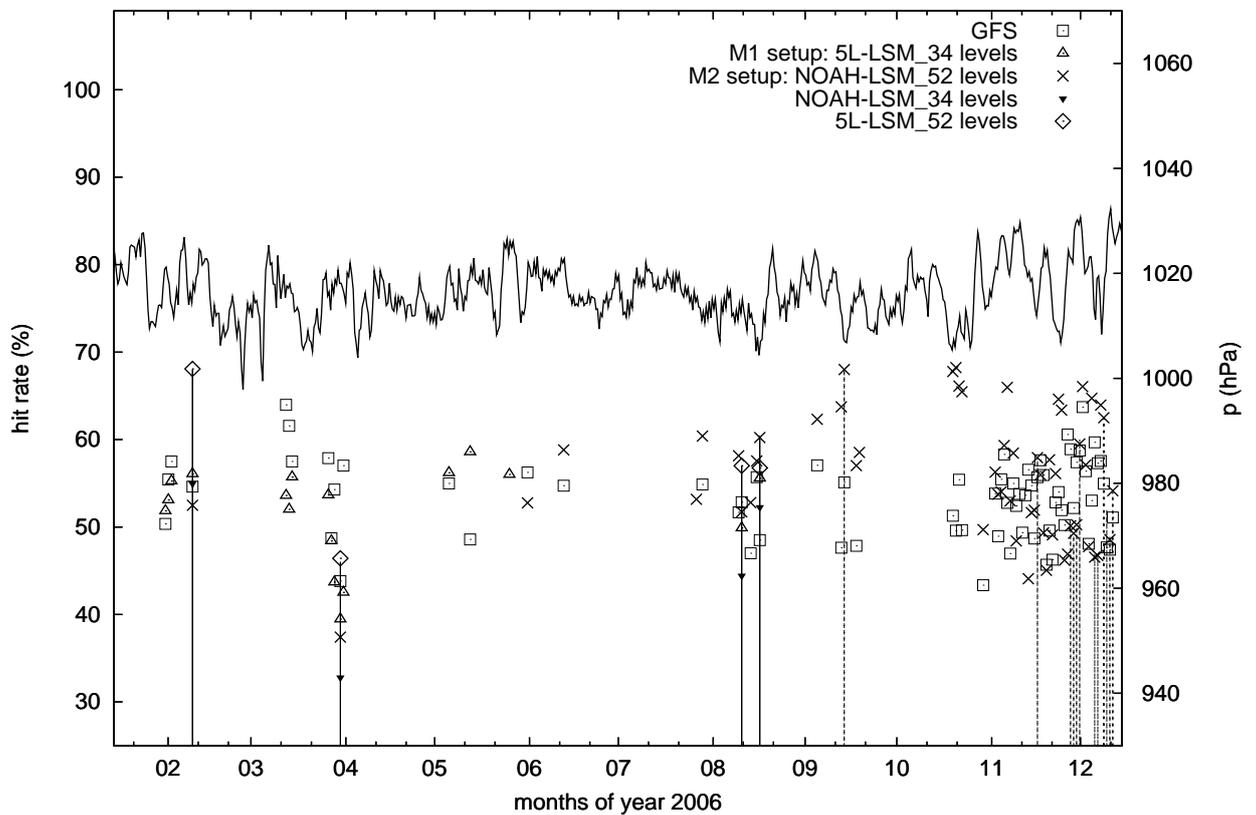
Figure 2.3.: Averaged daily hit rates HA (symbols) as defined in appendix B. Resimulation days are marked with lines. Days that are omitted in the long-term evaluation because of all-day instable stratification over the area are marked with dashed lines. Sea level pressure measured in Valencia as time series with scaling on the right y axis.

equivalent to $\pm 0.26 \, m \, s^{-1}$. To meet the desired accuracy of $\pm 1 \, m \, s^{-1}$ (section 2.4), the simulation has to be precise by $\pm 0.74 \, m \, s^{-1}$; otherwise, it is considered to be faulty.

2.4. Evaluation method

For the long-term evaluation of different model setups (section 2.5.1) the seasonal influence is reduced by concentrating on days for which ocean temperature is lower than land temperature and thus sea-breeze conditions prevail. Days with unstable stratification over the sea lasting the whole day are treated separately and marked in Fig. 2.3 (dashed lines). Instability over the sea disables sinking, which is necessary for a sea-breeze circulation. The days are identified by using the difference of air temperature (T) and SST as a measure for stratification. The days are identified by Fock (2007) and later on named "all-day unstable."

Only those measurement sites were considered that lie within the central simulation area covered by DOM3 before and after 1 June 2006. This guarantees the use of the same data collective. The three outermost grid points are disregarded to avoid the influences of the lateral boundaries. The first 3 h of each simulation are also disregarded as spinup time.

Horizontal interpolation, weighted by the inverse distance over the four nearest grid points, is performed to interpolate the model results to measurement points. In vertical direction the simulated wind speed is logarithmically extrapolated assuming a constant flux layer and neglecting stratification. Thus, the simulated friction velocity is taken into account in both setups. However, due to the greater distance the extrapolation error is larger for the 34 level setup.

For stations below $750 \, m$ MSL the simulated pressure is reduced to sea level with the same formula as used for the measurements (DeutscherWetterdienst, 1998). For the evaluation of T and dewpoint temperature (T_d) the simulated values from the lowest level are taken without vertical extrapolation as their profiles are highly variable near the ground. Therefore, the root-mean-square errors in these quantities are larger for the simulations with 34 levels, where the lowest sigma level is located approximately $36 \, m$ AGL. Also, discrepancies between simulated and real orography contribute to errors in T and T_d, as these quantities considerably depend on the height MSL.

For comparing the performance of our model runs to those of Cox et al. (1998), Miao et al. (2007), Zhong and Fast (2003), Perez-Landa et al. (2006), Schlünzen and Katzfey (2003), and Bogenschutz (2004), an average hit rate per model setup (HA) and per meteorological parameter (H) are calculated. The same model accuracy A as given in Cox et al. (1998) is used (see Table 2.1). The definition of the hit rate H

given in appendix B; definitions of all other performance measures used in this paper are given in appendix A. We calculate the mean error (ME), the root-mean-square error (RMSE), the Peirce skill score (PSS), and the correlation coefficient r for T, T_d, and p. For the wind we also calculate the RMSE of the vector wind difference (VWD) and the index of agreement (IOA). For wind direction, only RMSE and hit rate (H) are computed.

To get a better insight into the forecast performance with respect to individual meteorological parameters, conditional quantile plots are presented for DOM3 with the M1 and with the M2 model setup (Figs. 2.4 - 2.6). These figures show selected percentiles of simulated versus observed values. Furthermore, histograms of the frequency of the simulated intervals are given. From the quantile plots, simulation qualities can easily be assessed by determining the bias (constant deviation between median and the 1:1 diagonal representing the perfect forecast), conservativeness or progressiveness (whether the slope of the median is bigger or smaller than one), and the forecast error variability (deviation of small and large percentiles from median).

The sensitivity of the model on the changes of the LSM and the number of levels is separately and jointly investigated for 4 days by means of time series, conserved variable plots, and vertical profiles. The days 10 February, 31 March, and 11 August of year 2006 were sea-breeze days. Here 17 August was selected, because it was a summer season day similar to 11 August, but with large-scale changes preventing the sea-breeze development. The meteorological situations will be described in detail in section 2.5.3. The above selection was made after a first evaluation had been performed (section 2.5.1) to separately investigate the impacts of LSM and resolution.

2.5. LSM and resolution impact on model results

2.5.1. Long-term statistical evaluation

The average daily hit rates for all simulations and configurations are shown in Fig. 2.3. All-day unstable days over the sea (Fock, 2007) are treated separately and are marked

Table 2.1.: Desired accuracy range of model results and accepted measurement values.

	$T(°C)$	$T_d(°C)$	$p(hPa)$	$ff(m\ s^{-1})$	$dd(°)$
Accuracy	±2	±2	±1.7	±1	±30
Threshold			$980 < p$	$1 < ff$	$1 < ff$, $0° \leq dd \leq 360°$

(a)

(b)
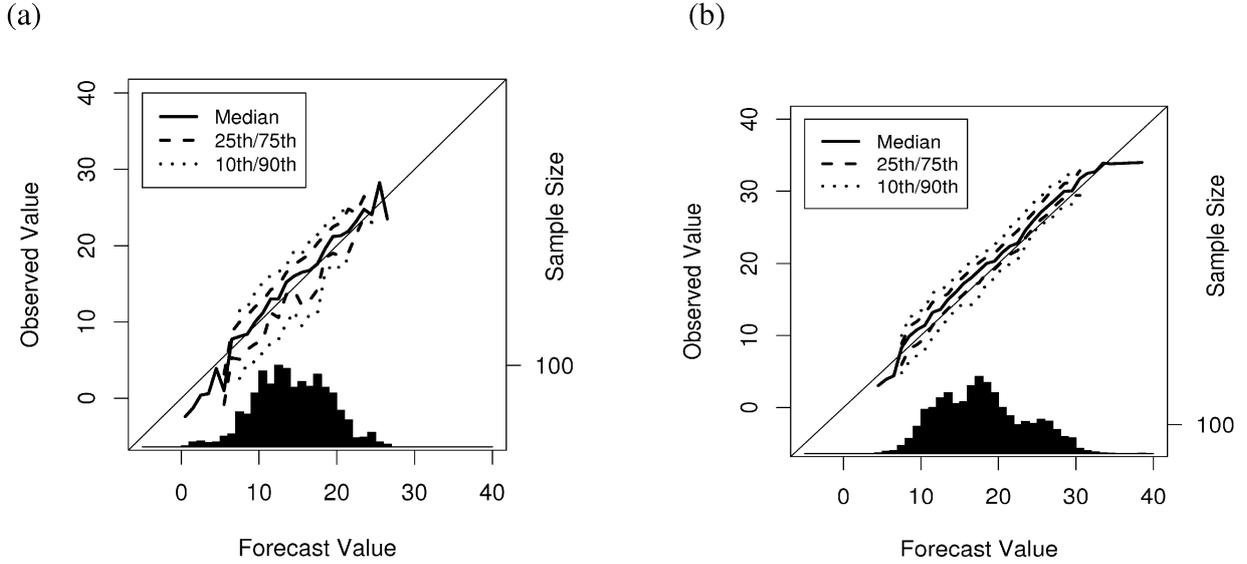
Figure 2.4.: Conditional quantile plots for temperature ($°C$) for the (a) M1 model setup used in the first half of the year and for the (b) M2 model setup used in the second half year. The histogram indicates the number counts per interval of the forecast values in $1 - K$ resolution with the scaling on the right y axis. The solid line connects the medians of the observations related to the particular forecast intervals. The dashed (dotted) lines mark the 25th (75th) quantiles and 10th (90th) quantiles.

in Fig. 2.3.

The hit rates range from 30% to 70%. The model performance does not improve significantly from the M1 setup to that of M2. For the same type of meteorological conditions the average hit rate over the 18 simulations with M1 setup is 52% versus 56% over 46 simulations with M2 setup. The improvement of the MM5 forecasts relative to the GFS forcing data is not evident (55% and 53%, respectively, as average hit rates over 71 simulations for the same days). However, a lack of improved model performance with increasing resolution may be a result from the point to point evaluation used here. This does not take into account the "double penalty" of observed but- not-forecasted and forecasted-but-not-observed features that grows with increasing resolution (Ebert, 2008). All-day unstable days only occurred in autumn and early winter and thus were simulated with the M2 setup. However, the different meteorological situations do not generally change the model performance (Fig. 2.3). Nonetheless, the further evaluation will focus on the other days to use a more homogeneous dataset.

The conditional quantile plots (Figs. 2.4 - 2.6) and Table 2.2 summarize the change in model performance due to the change from the M1 to the M2 model setup by

(a)

(b)
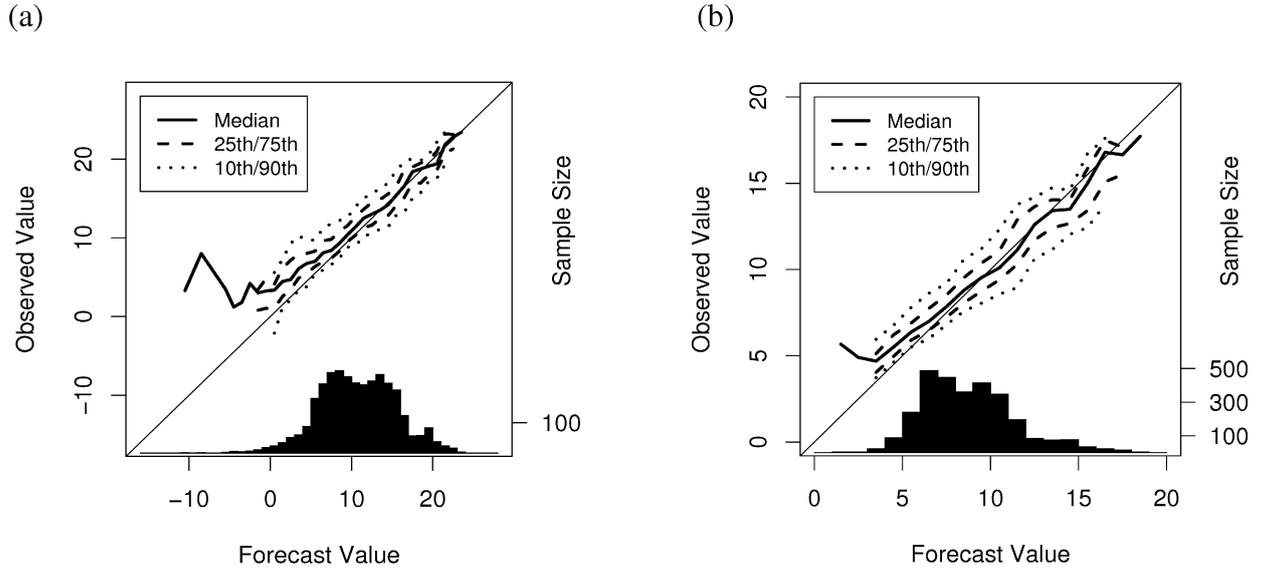
Figure 2.5.: Conditional quantile plots for (a) dewpoint temperature ($°C$) and (b) specific humidity ($g\,kg^{-1}$) for the M2 model setup used in the second half year. The histogram indicates the number counts per interval of the forecast values in (a) $1\,K$ or (b) $1\,g\,kg^{-1}$ resolution with the scaling on the right y axis. The solid line connects the medians of the observations, related to the particular forecast intervals. The dashed (dotted) lines mark the 25th (75th) quantiles and 10th (90th) quantiles.

(a)

(b)
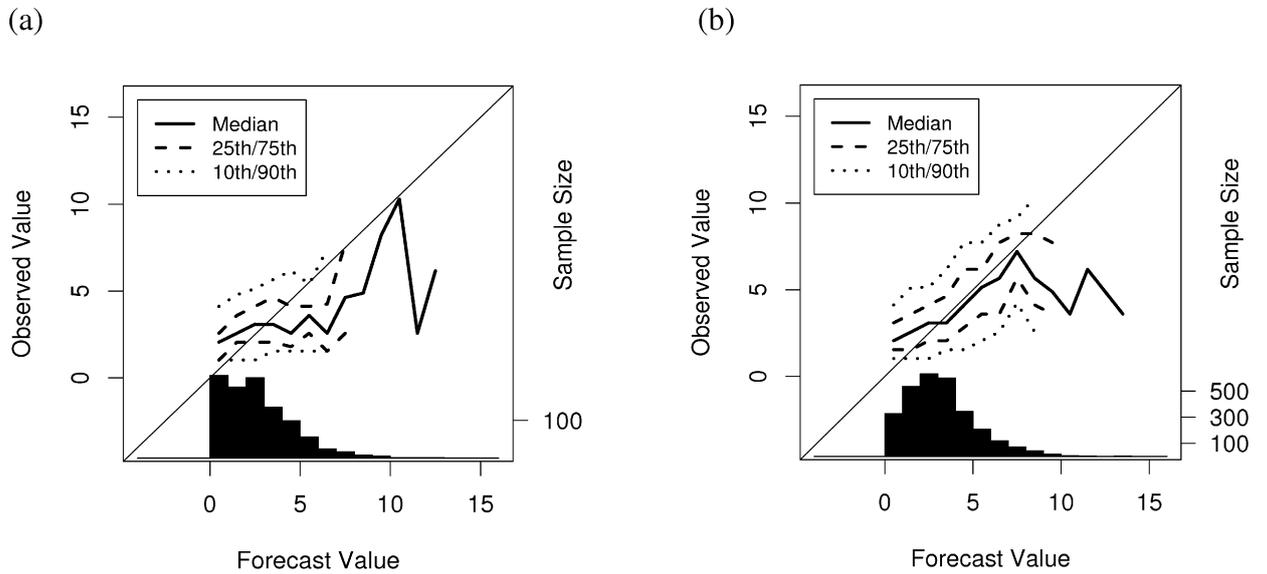
Figure 2.6.: As in Fig. 2.4, but for wind velocity. Forecast intervals of $1\,m\,s^{-1}$.

14

means of selected statistical parameters. A cold bias of approximately $1\,K$ results from the M2 setup for T values exceeding $8°C$ (Fig. 2.4b). This error is also reported in the literature for similar parameterization options (Foy et al., 2006) as well as for very different ones (Zhong and Fast, 2003). Zhong and Fast (2003) identify too high sensible heat fluxes and too low net radiation during the day, while at night the longwave radiation loss is simulated too large. These statements cannot be confirmed and not be rejected in the present investigation, as measurement data of the radiation are not available. Since simulated T_d and specific humidity also have an increasing negative bias from the M1 ($-0.55\,K$) to the M2 setup ($-0.94\,K$; Fig. 2.5 and Table 2.2), the T bias cannot only be explained by too large humidity values (Fig. 2.5). It can also not be attributed to the vertical levels used for evaluation since the vertical resolution is better for M2 setup than for M1 setup. All other statistical values improve for T and T_d when using the M2 setup and not the M1 setup.

The conditional quantile plots for wind velocity (Figs. 2.6a,b) reveal that the model especially overestimates the high wind speeds. The M2 setup considerably reduces this progressiveness. For M2 the overestimation in wind speed forecast is only found for wind velocities above $8\,m\,s^{-1}$. However, the scatter remains large. The negative bias remains similar in the M1 ($-0.4\,m\,s^{-1}$; Table 2.3) and M2 setup ($-0.5\,m\,s^{-1}$). The wind direction is better simulated for the M2 setup, which is also reflected in the VWD.

The pressure has a negative bias of approximately $-1\,hPa$ over the whole forecast range (Table 2.2). The correlation coefficients are high with 0.93 and 0.98 for M1 and M2, respectively. Therefore, the pressure patterns are believed to be well represented. The level (bias) is, however, not better simulated with the M2 setup. In addition the general closeness to measured data as reflected by the hit rate is better for the M1 setup. All in all the pressure performance is worse with the M2 setup.

2.5.2. Comparison to other mesoscale model simulations

The received performance values are compared to other simulations found in the literature, which were mostly performed as case studies either for similar regions or with similar models (Tables 2.4 and 2.5). Only the Weather Research and Forecast Model (WRF) simulations (Bogenschutz, 2004) performance measures were also achieved in forecast mode (M9). It is somewhat unfair, if not impracticable at all, to compare the statistical measures one to one, as the case studies setup is often adjusted for the specific case. However, the ranges of the statistical measures give a hint to the "typical ranges" of model results.

The MM5 simulations given in Cox et al. (1998) for a very coarse grid show a lower performance than M1 and M2 simulations for all parameters (Tables 2.4 and 2.5).

Table 2.2.: Model skills of the model setups M1 and M2 for temperature (T), dewpoint temperature (T_d), and mean sea level pressure (p, only stations below 750 m MSL). M1 is used from 1 Feb to 26 May 2006; M2 is used from 1 Jun 2006 to 12 Dec 2006; N represents the number of pairs used for the calculation of ME, RMSE, H, PSS, and r. Note that the number of compared data differs. Boldface values indicate the best performance.

	$T(°C)$						$T_d(°C)$						$p(hPa)$					
	ME	RMSE	H	PSS	r	N	ME	RMSE	H	PSS	r	N	ME	RMSE	H	PSS	r	N
M1	**-0.23**	3.43	41	.13	.83	1152	**-.55**	2.90	55	.20	.79	1160	**-0.90**	**1.40**	**86**	.29	.93	998
M2	-1.11	**2.41**	**58**	**.23**	**.94**	3290	-.94	**2.53**	**64**	**.28**	**.87**	3289	-1.08	1.57	78	**.34**	**.98**	2913

Table 2.3.: As in Table 2.2, but for wind velocity (ff) and wind direction (dd). IOA denotes the index of agreement and VWD is the vector wind difference.

	$ff(m\,s^{-1})$						$dd(°)$		
	ME	RMSE	IOA	VWD	H	N	RMSE	H	N
M1	**-0.40**	2.20	.60	3.36	36	967	51.6	54	967
M2	-0.50	**2.16**	**.68**	**3.14**	**40**	2905	**45.0**	**60**	2905

Table 2.4.: As in Table 2.2, but for model skills of simulations found in literature. M3 = MM5 (Cox et al., 1998), $46-km$ grid spacing, average over various simulations from different regions; M4 = MM5 (Miao et al., 2007), $2-km$ grid spacing, sea-breeze simulation in Sweden; M5 = MM5 (Zhong and Fast, 2003), $0.5-km$ grid spacing, simulation of locally forced synoptic situations in Salt Lake City Valley; M6 = RAMS (Perez-Landa et al., 2006), $1.5-km$ grid spacing, sea-breeze simulation in the Valencia coastal region; M7 = DARLAM (Schlünzen and Katzfey, 2003), $5-km$ grid spacing, simulation of a locally forced synoptic situation in eastern Germany; M8 = METRAS (Schlünzen and Katzfey, 2003), $4-km$ grid spacing, subgrid-scale flux aggregation with blending height for LSM, simulation of a locally forced synoptic situation in eastern Germany; M9 = WRF, $5-km$ grid spacing daily forecasts (1200–2400 UTC) for the southeast United States during 1 Jun–13 Aug 2003.

	$T(°C)$				$T_d(°C)$				$p(hPa)$			
	ME	RMSE	H	N	ME	RMSE	H	N	ME	RMSE	H	N
M3			31				44				46	
M4	-.60	2.70		2488					0.4	1.1		112
M5	-.70	2.17		1495								
M6	.40	1.49		300								
M7	-.54	2.55	68	381	2.17	5.06	35	381	-.42	.75	99	368
M8	.12	2.13	73	381	-.28	1.84	78	381	0	.95	94	368
M9	-.70	2.46		32000	0.20	0.70		32000	-.11	0.45		32000

The higher-resolution MM5 simulations (M4 and M5) show a similar skill to that of M1 and M2 for T and a better skill for pressure, especially when compared with M1. Humidity performance is not comparable as information is missing. Performance for wind speed is also in the same range for all MM5 simulations (Tables 2.3 and 2.5).

Regarding the impact of the LSMs and the number of levels the authors come to very different conclusions. While Miao et al. (2007) judge the Noah LSM better than the 5L LSM, Tonnesen and Coauthors (2005) do not rank these two LSMs. Zhong and Fast (2003) state that the results worsen with the Noah LSM when compared

Table 2.5.: As in Table 2.3, but for model skills for simulations found in literature.

	$ff(m\,s^{-1})$						$dd(°)$		
	ME	RMSE	IOA	VWD	H	N	RMSE	H	N
M3					22			36	
M4	0.40	2.10		3.1		3732			
M5	0.46	1.57					66.7		
M6	**0.36**	1.54	0.79	**2.18**					300
M7	-1.01	1.69			37	312	42.2	**64**	229
M8	0.91	**1.43**			**57**	312	**33.63**		229
M9		1.84				32000			

with the 5L LSM. For different vertical resolutions, Colby (2004) finds no substantial improvement by employing more than 24 levels to capture the sea breeze in New England.

Intercomparing the performance of MM5 (M1 to M5; Tables 2.2–2.4) to other models reveals that MM5 is for several parameters not the best performing model. However, the number of comparison data is much smaller for the Regional Atmospheric Modeling System (RAMS), Division of Atmospheric Research Limited- Area Model (DARLAM), and Mesoscale Transport and Fluid Model (METRAS) (M6–M8). In addition, the setup might have been adjusted for the specific case. For example, Schlünzen and Katzfey (2003) investigate the impact of LSM, horizontal resolution, and initial and boundary data and find a hit rate; for example, for METRAS of 29% to 75% for T, of 29% to 79% for T_d, of 51% to 58% for wind speed (ff), of 43% to 64% for wind direction (dd), and of 63% to 96% for p. With the exception of wind speed MM5 results lie within these ranges. Comparing results of M2 (Tables 2.2 and 2.3) to the WRF forecasts (M9 in Tables 2.4 and 2.5) clearly shows the improvements achieved with the MM5 successor model WRF. WRF performs better for all parameters except T. In comparison to all other model results including the case studies, WRF performs best for T_d.

2.5.3. Analysis of impact of LSM and vertical resolution for selected cases

To separately identify the influences of LSM and vertical resolution, 4 days (10 February, 21 March, 11 August, and 17 August) were resimulated in all possible combinations. They represent one winter day (10 February), the day with the worst overall performance (21 March), and two typical summer days. The first three cases are connected with a potential sea-breeze development. The last case was dominated

by a frontal passage. The performance of the 16 resimulations is summarized with respect to the individual meteorological parameters in Table 2.6.

Most of the statistical parameters show the same tendency as the parameters of M1 and M2 model setup (Tables 2.2 and 2.3). Therefore, this small subset of simulations may also be regarded as being representative with respect to the performance of the setups M1 (5L LSM/34 levels) and M2 (Noah LSM/52 levels). Sorting the individual cases in Table 2.6 with respect to average hit rates HA (defined in appendix B), as indicated by the column "rank," gives strong hints to the main reasons for enhanced or worsened performance. The biggest positive impact has the increased number of levels. This is especially true for wind speed, less so for wind direction.

For Noah LSM, three of the four cases are better for all parameters with the higher vertical resolution (except 10 February). For the 5L LSM all cases improve with higher vertical resolution, which is not as pronounced for the frontal case (17 August). For the lower vertical resolution the 5L LSM always performs better than Noah LSM, while for the high resolution (52 levels) this is not true for the frontal day (17 August) and not true for T on 21 March.

For the days with sea-breeze conditions (10 February, 21 March, and 11 August) the combination of 5L LSM and 52 levels gives the best results for all parameters except T. For the day with frontal passage (17 August), the Noah LSM with 52 levels performs better. Some deeper insight into the model performance and possible reasons for this difference in performance are gained by a detailed synoptic interpretation, which is given in the following.

10 FEBRUARY 2006

A weak $500-hPa$ ridge was situated at the eastern Atlantic Ocean with an embedded aged cyclone at Madeira (central pressure of $1005\,hPa$). Its occluded front reached the southwestern part of Iberia. Iberia was situated at a col between this cyclone and another more active one over central Europe, with associated anticyclones over the British Isles and Northern Africa. The large-scale pressure gradient was very weak. The system remained quasi-stationary for the day. The sounding from 1200 UTC (Fig. 2.7) at Murcia station (430 in Fig. 2.2) shows a slightly stable atmosphere up to $800\,m$ with an inversion above, remaining from the night before. Scattered clouds covered the region of DOM3 at midday, but were too thin to inhibit the development of the sea-breeze circulation. Figure 2.8 shows the typical wind time series for a sea breeze: light offshore winds in the morning turn to onshore and accelerate in the afternoon.

Table 2.6.: Model skills of all setups of the resimulated days. Number of compared values is approximately 70. The best performances per day are marked by boldface values. Ranking on the basis of average hit rates HA as defined in appendix B.

Date	LSM	Levels	$H(T)$	$ME[T(°C)]$	$H(T_d)$	$ME[T_d(°C)]$	$H(p)$	$H(ff)$	$H(dd)$	$VWD(ms^{-1})$	HA	Rank
10 Feb	5L	34	47	**0.39**	55	-0.80	92	49	61	3.1	60	2
	Noah	52	45	-1.05	29	-2.76	92	61	43	2.9	52	4
	Noah	34	**50**	0.45	36	-2.32	94	47	53	3.2	55	3
	5L	52	49	-0.83	**61**	**0.16**	**94**	**78**	**67**	**2.5**	68	1
31 Mar	5L	34	53	0.47	34	-1.45	16	32	45	2.9	36	3
	Noah	52	**64**	**-0.25**	23	-4.14	7	44	46	3.0	37	2
	Noah	34	52	0.76	22	-4.24	5	30	**52**	3.2	33	4
	5L	52	62	-0.41	**53**	**0.28**	**21**	48	45	**2.8**	46	1
11 Aug	5L	34	26	-2.27	41	-0.70	80	43	65	2.6	50	3
	Noah	52	**47**	**-1.96**	41	-1.01	79	43	57	2.3	52	2
	Noah	34	23	-2.28	39	-1.85	80	25	61	3.1	44	4
	5L	52	37	-2.52	45	**0.43**	**84**	45	**71**	**2.0**	57	1
17 Aug	5L	34	38	-1.99	56	**-0.48**	93	27	**68**	3.8	56	3
	Noah	52	**55**	**-1.56**	56	-0.79	89	50	65	**3.3**	62	1
	Noah	34	42	-1.58	51	-1.65	92	19	63	4.0	52	4
	5L	52	42	-2.08	43	1.40	**95**	49	62	3.4	57	2
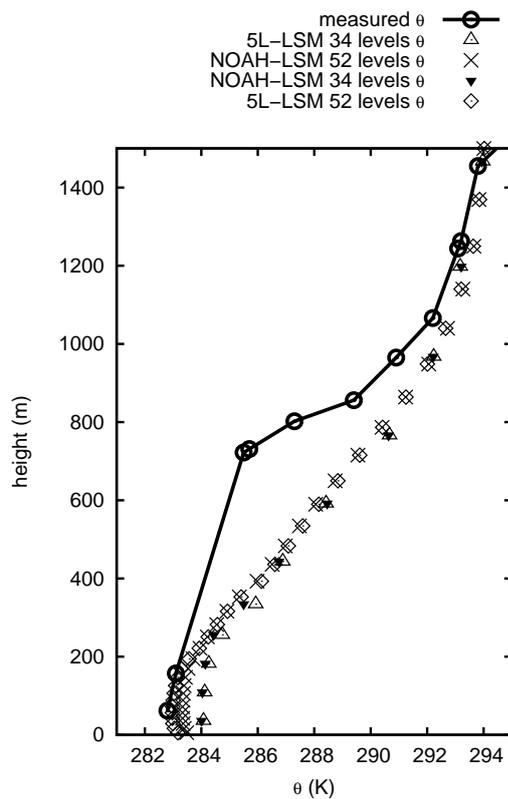
Figure 2.7.: Observed and modeled vertical profiles of potential temperature at Murcia at 1200 UTC 10 Feb 2006.
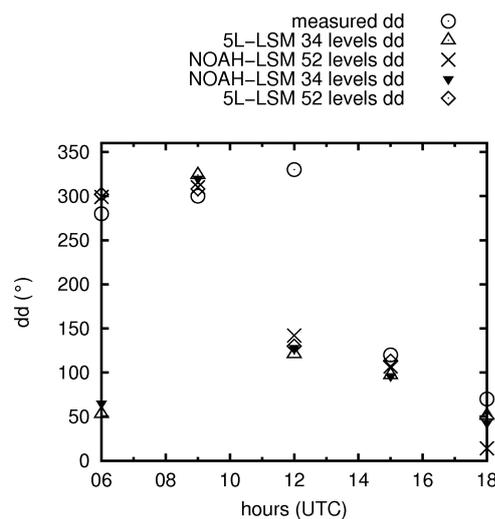
(a)   (b)
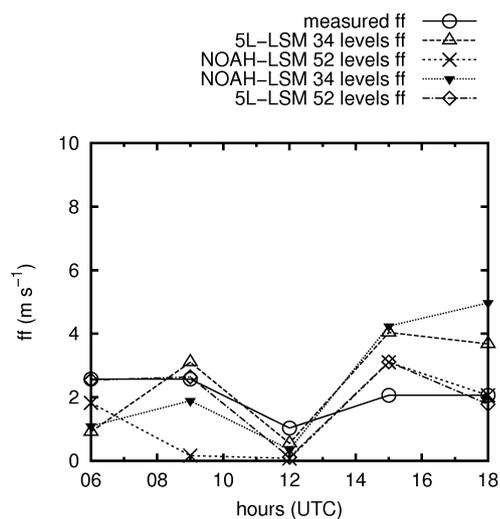
Figure 2.8.: Observed and modeled time series of (a) wind direction and (b) velocity at Valencia Airport for 10 Feb 2006 with $3 - km$ resolution.

None of the setups simulated strong enough vertical mixing, so that the neutrally stratified layer is confined to the lowest $200\,m$ (Fig. 2.7). Nevertheless, the sea breeze is captured well, especially from the setup with 5L LSM and 52 levels (Fig. 2.8). The sea-breeze diurnal cycle is well visible in temperature and humidity by using conserved variable plots (Perez-Landa et al., 2006). The measurements for Valencia (Fig. 2.9) show warming until 1500 UTC. A sharp decrease in heating and a further increase in specific humidity indicate the passage of the sea-breeze front 3 h later than the wind shift. Again the setup with 5L LSM and 52 levels performs best. Both simulations of the Noah LSM have a dry bias, which exists also for 31 March (Fig. 2.11) and 11 August (Fig. 2.13). This dry bias is also evident from the analysis of the M1 and for lower humidities for the M2 setup (Fig. 2.5).
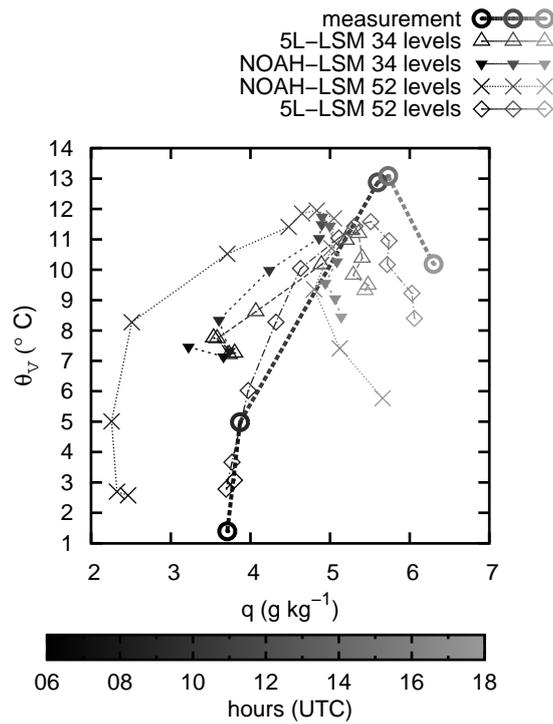
Figure 2.9.: Development of virtual potential temperature vs specific humidity in 1-hourly resolution for the simulations and 3-hourly resolution for the measurement. First symbol (thick) corresponds to 0600 UTC. Observations and simulation results of all setups for Valencia Airport for the daytime period on 10 Feb 2006.
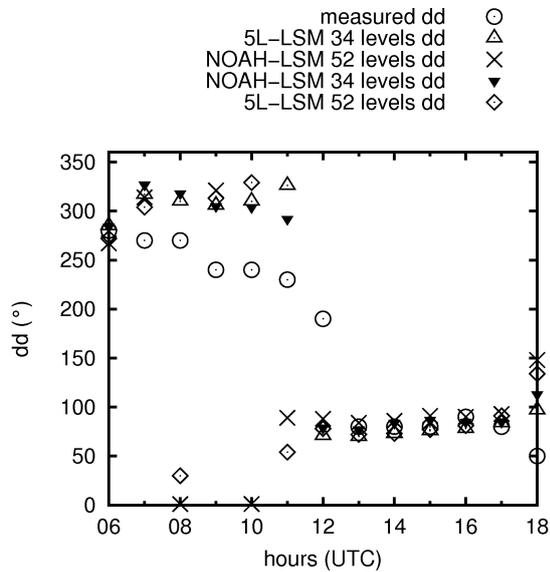
(a)   (b)
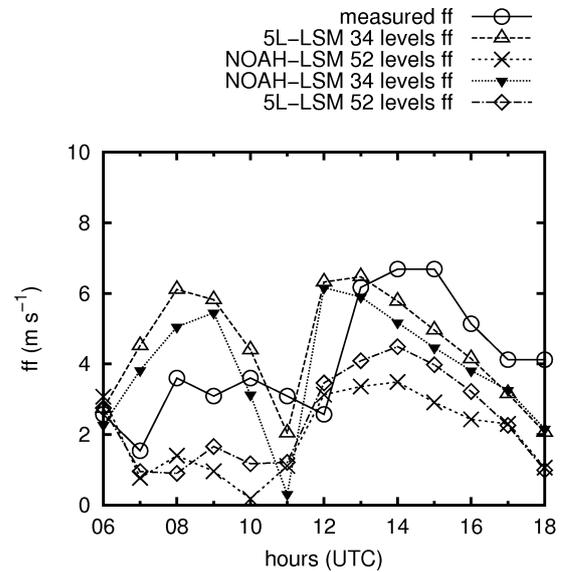
Figure 2.10.: As in Fig. 2.8, but for 31 Mar 2006 and with hourly resolution.

31 MARCH 2006

At this day the polar front was located just north of Iberia in the southwest–northeast direction. A weak surface high pressure system without large pressure gradients lay over Iberia. At 1200 UTC the boundary layer was unstably stratified in the lowest $200\,m$, neutral above (not shown). Only thin clouds were visible from satellite. In these perfect conditions a strong sea breeze developed, the wind direction changed at 1200 UTC (Fig. 2.10). The conserved variables plot (Fig. 2.11) shows about constant humidities and increasing temperatures until the front passage and then decreasing temperatures with slightly larger humidities after 1200 UTC. This may result from stronger vertical mixing due to a less stable stratification compared to 10 February 2006. The simulations with 52 levels capture the development reasonably well (apart from a humidity offset).

11 AUGUST 2006

A near-ground ridge stretched from an anticyclone reaching up to $500\,hPa$ over the North Atlantic to the Mediterranean. Over the Strait of Gibraltar lay a very weak surface low. The day was cloud free. The lowest $1000\,m$ were neutrally stratified at night, so that an onshore wind developed with the aid of the surface low at 0800 UTC.
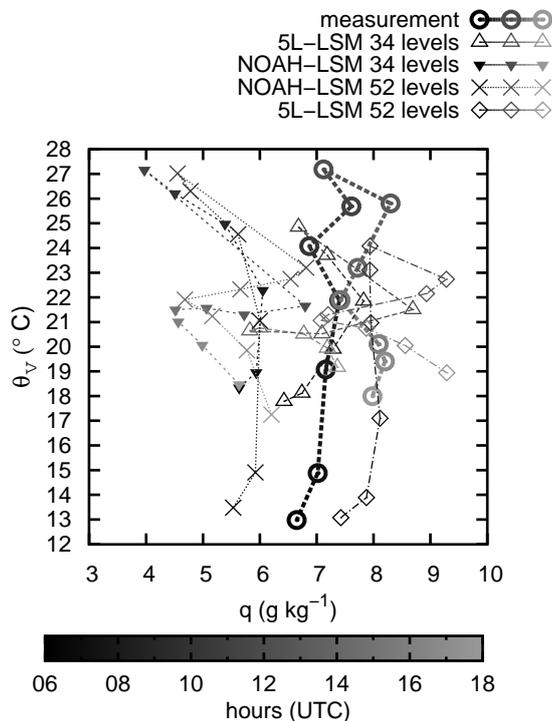
Figure 2.11.: As in Fig. 2.9, but for 31 Mar 2006 and with hourly resolution of all data.

The lowest $150\,m$ were unstably stratified at 1200 UTC (Fig. 2.12a), which is best captured when using 52 levels. Only the Noah LSM captures the profile values reasonably well above $150\,m$. The setups with the 5L LSM simulate the near-ground humidity well (Fig. 2.12b), however, all simulations are too moist above $150\,m$. High clouds were wrongly simulated in the southern part of DOM3, which suppressed the observed warming (Fig. 2.13). This behavior is also reflected by the hit rates and ME in Table 2.6. For wind the single point comparisons lead to two conclusions: the profile from Murcia (Fig. 2.14) indicates the Noah LSM simulations to be too weak, and the time series for Valencia (Fig. 2.15) indicates the 52 levels simulations to be too weak. However, comparisons for all surface sites favor the 5L 52-levels' setup, using $H(ff)$, $H(dd)$, and VWD as performance measures (Table 2.6).

17 AUGUST 2006

A trough in $500\,hPa$, directed north–south with an embedded cyclone reached into the Gulf of Biskaya. Cloud bands of the associated cold front crossed Spain at
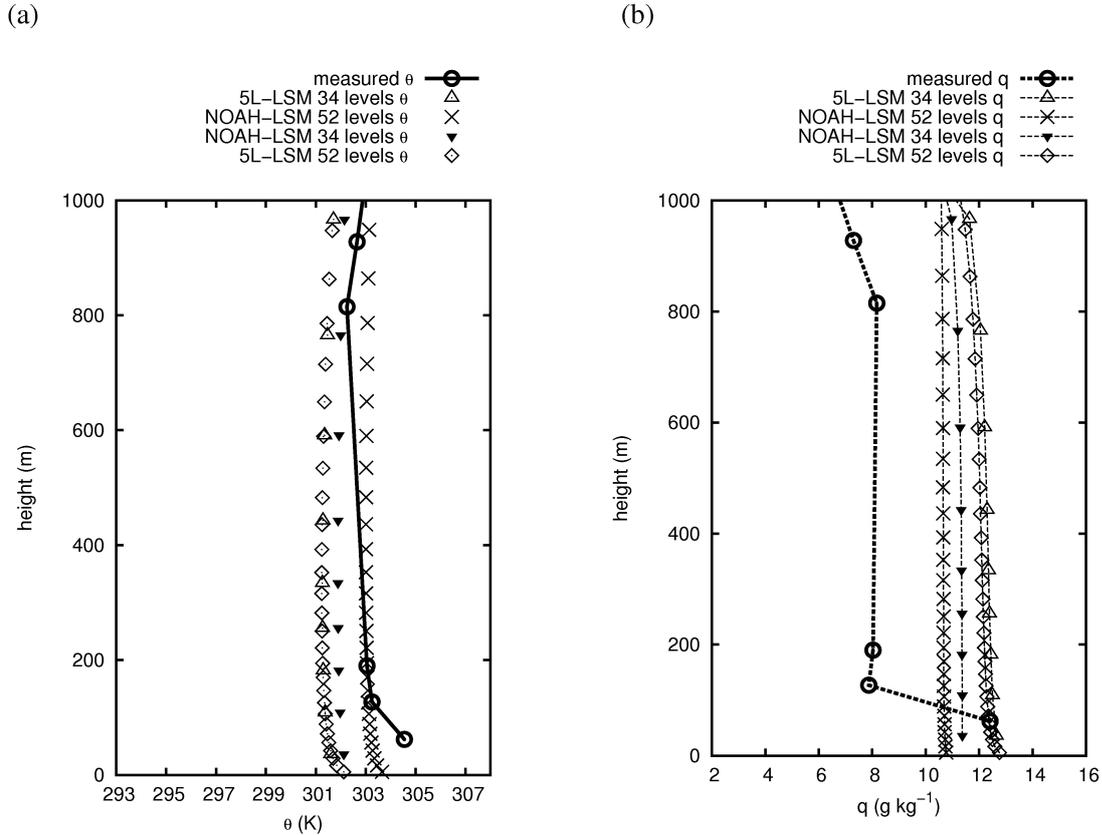
Figure 2.12.: Observed and modeled vertical profiles of (a) potential temperature and (b) specific humidity at Murcia at 1200 UTC 11 Aug 2006.

midday. Moist air was advected in $700\ hPa$. Near-ground wind backed from the northwest $5\ m\,s^{-1}$ to the southwest $3\ m\,s^{-1}$ until 1000 UTC (Fig. 2.16) and veered afterward, accelerating to a maximum wind velocity of $10\ m\,s^{-1}$ at 1400 UTC. The front was too weak to show a marked minimum in the pressure time series. The only additional hint to the timing of frontal passage is the development of T. Cooling started at 1100 UTC (Fig. 2.17). The timing of the maximum T is captured well from all simulations except that with 5L 52 levels. However, all of them miss the diurnal changes. All simulations show a too high humidity during the afternoon. The wind simulation performs well except for the 5L 52-levels' setup. It misses the wind maximum at 1400 UTC (Fig. 2.16). Concerning the thermodynamics, the bad performance of the 5L 52-levels' setup is well represented by the statistical measures (Table 2.6). As for 11 August 2006, the overall performance of the wind simulation can hardly be assessed by a single point comparison. Both HA and VWD (Table 2.6) favor the 52-level simulations. This result was not to be expected by the single time series from Valencia (Fig. 2.16).
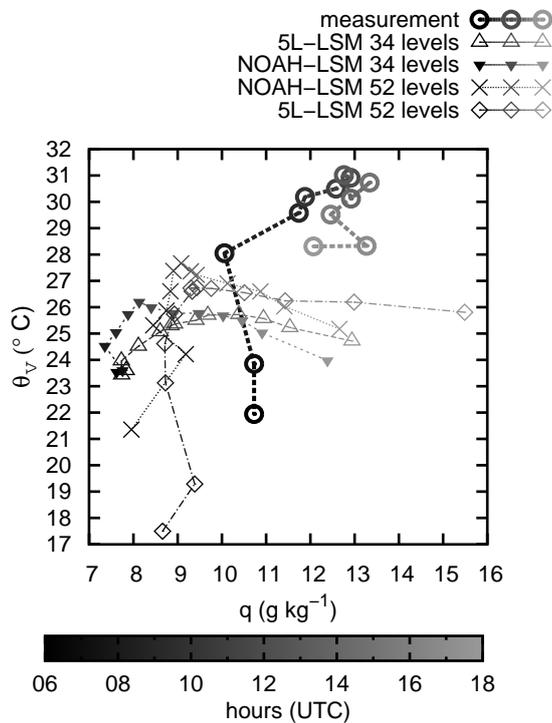
Figure 2.13.: As in Fig. 2.9, but for 11 Aug 2006 and with hourly resolution of all data.

In summary, we conclude that the Noah LSM delivers too low humidities near the ground at days with sea-breeze conditions. There is another effect that is observed in all cases: the enlarged number of levels improves the near-ground temperature and humidity profiles and thereby the T and T_d values (e.g., at 10 February and 11 August). It also improves wind speed. However, the unrealistic formation of clouds caused by the wrong structure of the humidity profiles at 11 August and 17 August could not be prevented by any of the setups. The dependence of the simulated sea breeze on thermodynamics could subjectively be shown on a pointwise basis. The good simulation of the diurnal temperature cycle at 10 February and 31 March leads to good simulations of the sea breeze and related wind speeds. The unrealistic shading on 11 August disturbs the temperature cycle and subsequently leads to a too weak sea breeze. For 17 August results suggest that the wrong humidity simulations are mostly relevant for the overall weak model performance of the 5L scheme in combination with 52 levels.
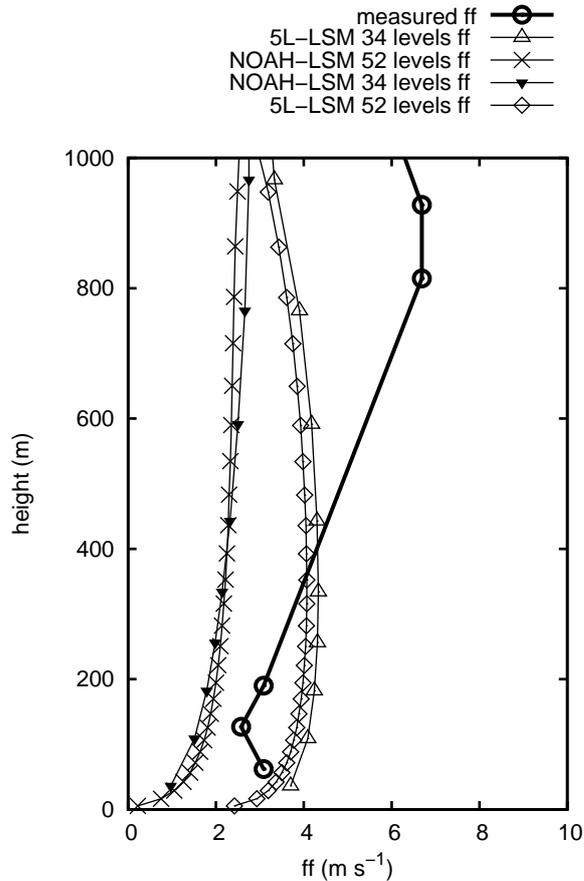
Figure 2.14.: Observed and modeled vertical profiles of wind velocity at Murcia at 1200 UTC 11 Aug 2006.

2.6. Conclusions of Chapter 2

We have evaluated operational forecasts for 71 days of 2006 with 64 days of similar meteorology. Sixteen days were resimulated. The performance of two setups that differ in the number of levels and the used LSM were compared. The evaluation was performed on statistical basis employing standard measures based on point-to-point comparison of surface data. The resimulations were evaluated in more detail by additionally comparing vertical profiles.

For the 3-km simulations the different setups M1 (5L LSM, 34 levels) and M2 (Noah LSM, 52 levels) do not show a completely consistent improvement in model performance. While setup M1 shows a smaller ME than setup M2, the hit rates are higher in general for M2 (the only exception being pressure). However, the improvement of the hit rates for T and T_d by 17% and 9% (Table 2.2), respectively, partly
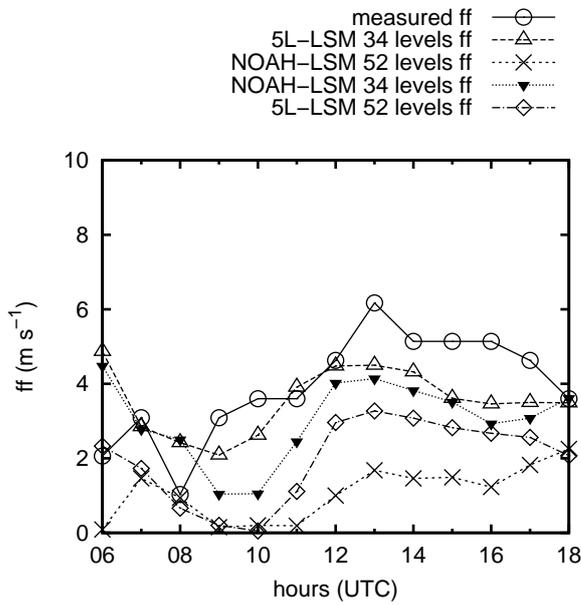
Figure 2.15.: Observed and modeled time series of wind velocity at Valencia Airport for 11 Aug 2006.
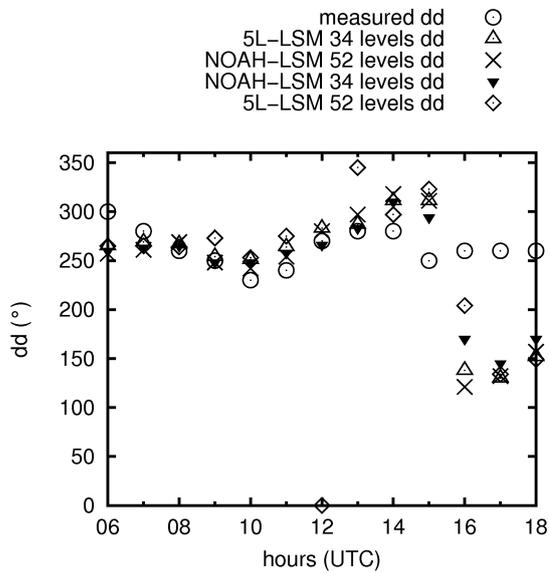
(a)   (b)
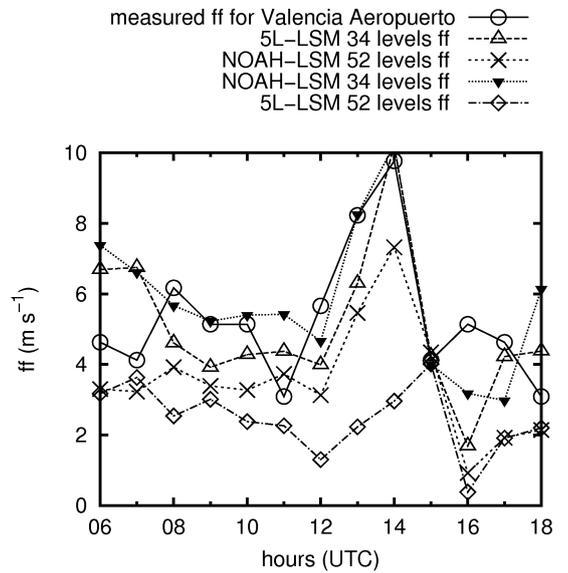
Figure 2.16.: As in Fig. 2.8, but for 17 Aug 2006 and with hourly resolution.
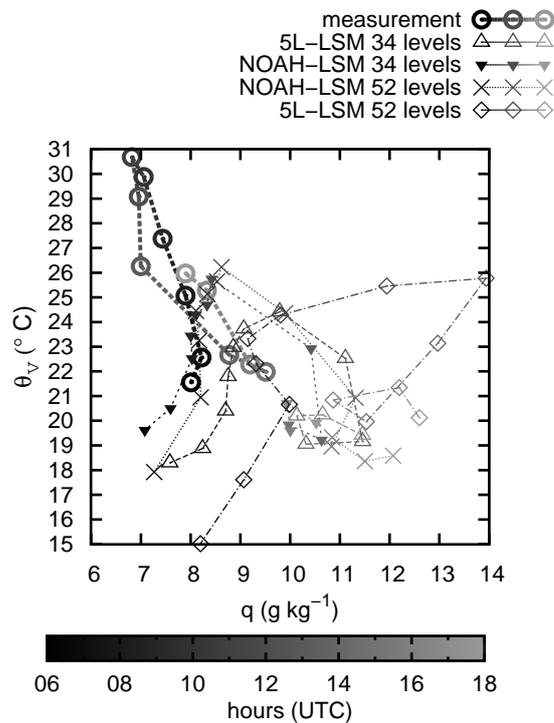
Figure 2.17.: As in Fig. 2.9, but for 17 Aug 2006 and with hourly resolution.

results from the reduced discrepancy in height between measurement and simulation due to the improved vertical resolution in M2 and thus has no model-physical meaning. This is also true for the improvement found in wind forecast when using the M2 setup. However, the assessment is not consistently promoting the M1 or the M2 setup. The analysis of the separate effects of LSM and vertical resolution reveals that the use of the more sophisticated Noah LSM shows less realistic results for all parameters except T including a near-ground dry bias for the three resimulated sea-breeze days. Also Zhong and Fast (2003) conclude that the simple 5L LSM performs better than the Noah LSM, in their case because of better simulated surface temperatures. We can only speculate about the reasons. Millan et al. (2002) point out that inaccurate specifications of parameters, such as soil temperature and moisture, texture, soil categories, and vegetation activity may degrade the performance of sophisticated LSMs beyond that of simple LSMs. There remains to be checked, whether the Noah LSM performed worse than the simple 5L LSM because of inappropriate input data that we used from the GFS and from tabulated values coming with the model. However, more realistic input data are hardly available for routine forecasts.

The performance improved for most parameters by the increased vertical resolution. Especially near-ground unstable vertical stratification, one of the main forcing features for sea-breeze development, could be captured much more realistically. The increased vertical resolution is especially relevant for the wind values. These consistently improve with increasing vertical resolution. Therefore, if a proper wind forecast is a forecast objective, a high-resolution vertical grid should be employed.

The influences of the LSM found in this paper remain relevant, as the MM5's successor model WRF uses the same LSMs. Also the numerics leading to better results when enhancing vertical resolution are the same in both models. For simulations of a sea breeze in a dry coastal area like Spain it is recommended from our results to use the 5L LSM. It also is important to use 52 levels, which improved all parameters compared to the 34 levels' results. Since the differences to Noah LSM with 52 levels are small, it is probably best to apply the 5L LSM with 52 levels to generally receive the best performing model.

3. Adaption of large scale forcing to include atmospheric boundary layer effects

For the Fram Strait two synoptic situations are simulated with METRAS (Section 4.3 gives an overview of the model). For METRAS model applications with a resolution of $7\,km$ the large scale meteorological conditions in high latitudes are most adequately prescribed by ECMWF reanalyses (Persson and Grazzini, 2005). They have shown good results for data sparse regions in high latitudes (Birnbaum et al., 2006) and have been used to force mesoscale models in high latitudes before (Wacker et al., 2009). Bohnenstengel (2010) forces METRAS simulations over Eastern Germany with ECMWF reanalyses. Six hourly data on ECMWF-pressure-levels of wind, temperature, specific humidity and cloud water are used as initial and lateral forcing values. The nudging technique is used in METRAS to assimilate the analysis data. The method is explained in detail in Section 4.3.1. The present chapter addresses several difficulties related to the vertical interpolation from the ECMWF grid to the METRAS grid. Some of them are especially obvious in the investigated region:

- Firstly, the spectral formulation of the ECMWF model causes negative surface heights offshore of steep coasts (Fig. 3.1a). Artificial levelling is necessary for the vertical interpolation to the METRAS grid, but affects the layer thickness between two pressure levels.

- Secondly, the sea-ice distribution, one of the most important topographic features of the Arctic, might lack small scale features. An imprecise sea-ice distribution influences the near-ground temperature, which would otherwise be of important additional information for the vertical interpolation.

- Thirdly, the coarse vertical resolution of the driving model sometimes leads to the omission of near-ground inversions over sea ice, which are characteristic for the Arctic. Fig. 3.2 shows simulated profiles in comparison to measured by the German research aircraft Falcon-20 (Brümmer et al., 2005) for 7th March 2002, 09:48 UTC to 11:54 UTC. The simulated profiles lack the inversion strength.
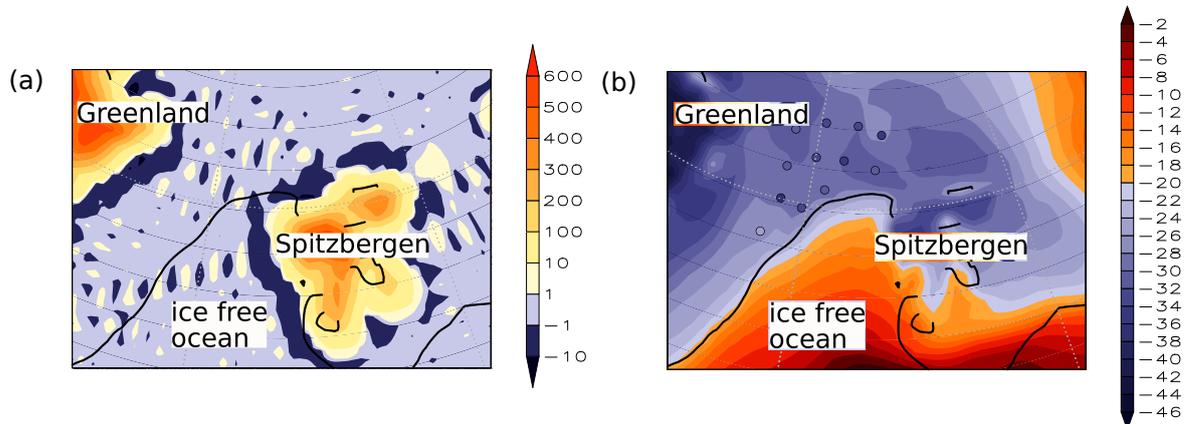
Figure 3.1.: Surface heights [m] (a) and 2-m-potentieal-temperatures [°C] for 5th March 2002, 18 UTC (b) of ECMWF forcing data shaded as indicated by the respective colour bar. The sea-ice edge is indicated as isoline of 50% sea-ice concentration. (b) additionally shows buoys as dots with measurements in the same colour coding.

- Fourthly, ECMWF data that are available on pressure levels sometimes occur below the ECMWF-orography. These data are wrong and need to be masked out.

- Fifthly, steep orography is smoothed due to ECMWF's relative coarse resolution of T341. As a consequence the surface heights of both models deviate, with the METRAS surface being higher than the ECMWF surface at crests and lower in narrow valleys.

Besides the original strategy for the vertical interpolation (OVI), in the following section also a vertical interpolation from the ECMWF grid to the METRAS grid will be presented that has been adapted to the Arctic region (AVI). The impact of the varied nudging obtained by the two vertical interpolations will be presented in Section 3.2 with two simulations of the development of an on-ice moving trough.

3.1. Vertical interpolation

For the OVI the close-to-surface vertical interpolation from the ECMWF grid to the METRAS grid is performed in three steps:

- Firstly, at least three additional auxiliary levels are introduced between the ECMWF-surface and the ECMWF-level of 925 hPa. Where ECMWF-pressure
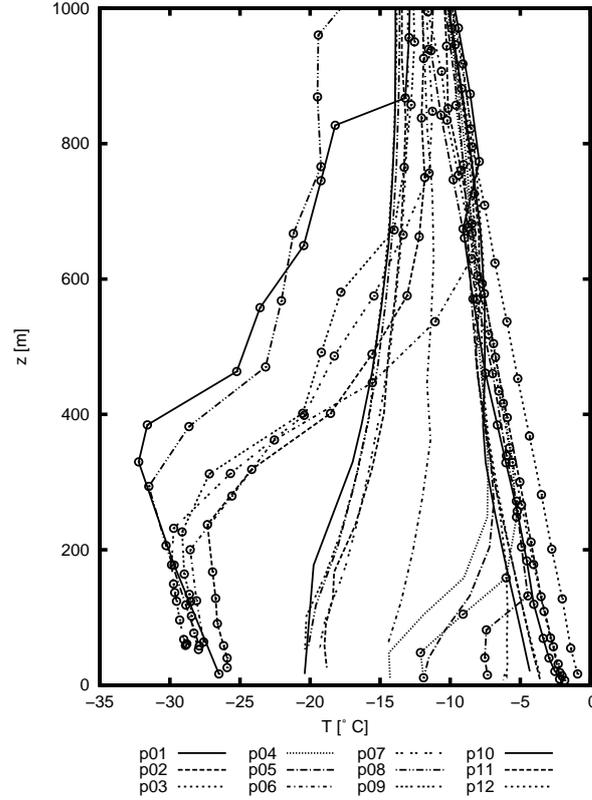
Figure 3.2.: Profiles measured by aircraft Falcon (lines with circles) within the model domain at 7th March 2002, 09:48 UTC to 11:54 UTC and the corresponding ECMWF profiles (in corresponding line style) of 12 UTC the same day. ECMWF profiles are spatially linearly interpolated to the locations of the respective profiles flown.

levels are lower than the ECMWF-orography they are omitted and the number of additional auxiliary levels is increased by their number. The height of the lowest auxiliary level is $10\,m$ above the ECMWF-surface. The remaining auxiliary levels are equally distributed between $10\,m$ and the height of the lowest above surface ECMWF-pressure-level. The meteorological fields are interpolated onto the auxiliary levels. The 2-m-temperature is omitted, because a negative influence of inconsistent land-use distributions may occur. In the Arctic this mainly means differences in the sea-ice distribution. Instead, the temperature is extrapolated from the lowest pressure level above the orography to the auxiliary levels with help of the barometric mean temperature T_{bar} (Eq. 3.1).

$$T_{bar} = -\frac{g(h_2 - h_1)}{R(ln(\frac{p_2}{p_1}))} \quad (3.1)$$

T_{bar} is calculated from the layer thickness applying the hydrostatic equation with the density replaced by the ideal gas law. h_2 denotes the height of the lowest ECMWF pressure level above the orography with pressure p_2, h_1 and p_1 the surface values. The layer thickness may be reduced due to the artificial levelling applied to negative sea-surface heights. The reduction of 10 m, as occurring east of Greenland, yields a reduction of T_{bar} of approximately 3 K. The effect becomes obvious if comparing the locations of the originally negative surface heights (Fig. 3.1a) with the locations of the exceptionally low 10-m-forcing-temperatures on the METRAS grid around Spitzbergen, east of Greenland and at cold tongues directed southward from the sea-ice edge (Fig. 3.3a).
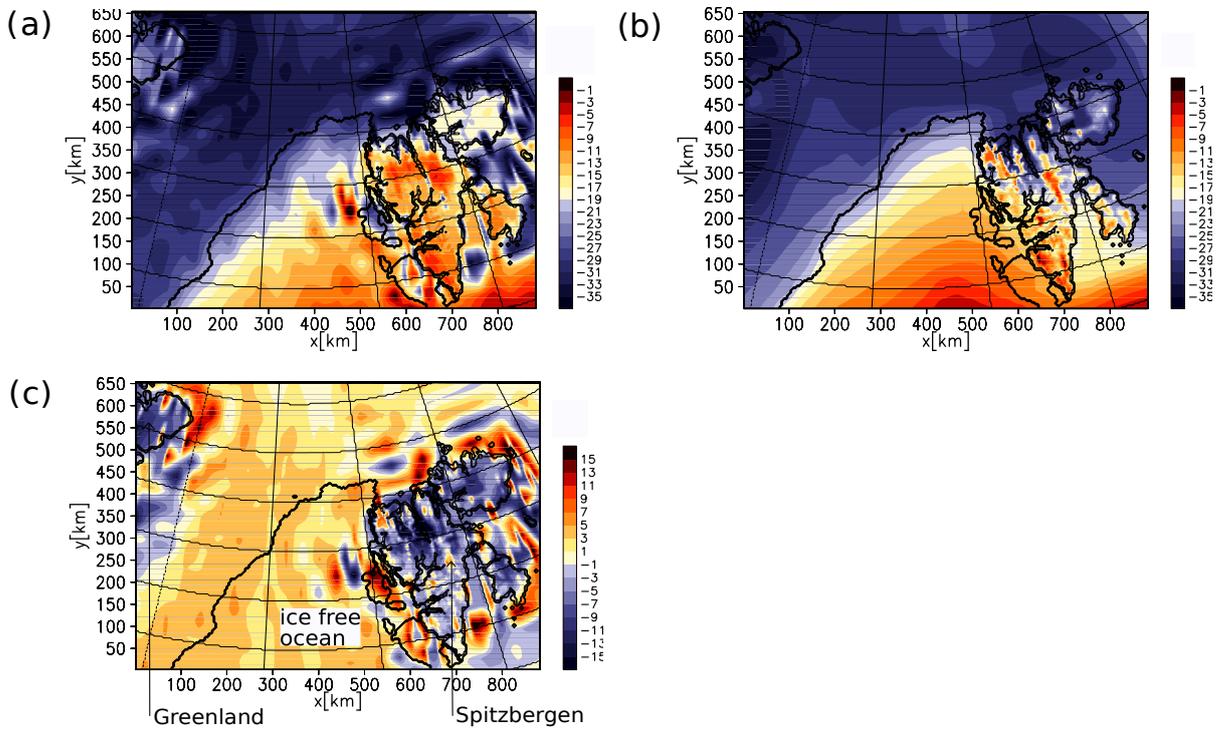
Figure 3.3.: METRAS model domain with $10\,m$ potential temperatures used as nudging ((a) and (b)) and their differences (c). (a) is derived by the original interpolation using the lowest ECMWF pressure level above orography and T_{bar} for 5th March 2002, 18 UTC, with the colour bar in °C. (b) is derived by the alternative interpolation using the ECMWF 2-m-potential-temperature for the same date and with the same colour coding. (c) shows the difference of (b)-(a) with the colour bar indicating K; note that the range of differences is not fully resolved by the colour scale. The sea-ice edge is indicated as isoline of 50% sea-ice concentration as observed on 6 March 2002 14 UTC.

- Secondly, a bilinear horizontal interpolation to the METRAS grid is performed.

- Thirdly, the forcing fields are vertically interpolated from the surface-, auxiliary- and pressure-levels onto the METRAS levels. Surface forcing-fields are only used if their height does not exceed the METRAS surface by more than $10\,m$.

Encouraged by ECMWF's relatively well located sea-ice edge (compare Fig. 3.1b with 3.3a) and the 2-m-temperature field being in good agreement with buoy measurements (Fig. 3.1b), a vertical interpolation is implemented and tested that is adapted to the Arctic region (AVI). It differs in the first of the three steps explained above.

Instead of extrapolating the temperature from the lowest pressure level with the help of T_{bar}, the temperatures of the auxiliary levels are interpolated from the 2-m-temperature and the lowest pressure level above ECMWF orography. An additional change is performed to AVI to account for near-ground inversions. They frequently occur over sea ice, but are sometimes underestimated on the model levels of ECMWF analyses (Fig. 3.2). To still capture near-ground inversions, a temperature profile is constructed that consumes all thermal energy $E_{th-bar} = \int_{h_1}^{h_2} c_p \cdot T_{bar}\,dz$ associated to the barometric mean temperature defined in Eq. 3.1. E_{th-bar} is compared with the thermal energy associated to the mean temperature from linear interpolation $E_{th-lin} = \int_{h_1}^{h_2} c_p \cdot T_{lin}\,dz$, where $T_{lin} = (T_2 + T_1)/2$ is the mean temperature with T_1 the surface temperature and T_2 the temperature at the lowest ECMWF pressure level above the orography (3.4).

If E_{th-bar} is higher than E_{th-lin} the extra thermal energy $E_{th-extra} = E_{th-bar} - E_{th-lin}$ (grey area in Fig. 3.4) is redistributed in a triangle assuming a near-ground inversion and a well mixed layer above (striped area in Fig. 3.4). The inversion temperature T_3 can be derived as follows. The grey area A_1 representing the extra thermal energy (Fig. 3.4) to be redistributed is given by Eq. 3.2.

$$A_1 = (T_{bar} - T_2)h - (T_1 - T_2)\frac{h}{2} \tag{3.2}$$

$h = h_2 - h_1$. The area A_3 after redistribution is given by Eq. 3.3.

$$A_3 = (T_3 - T_2)h - (T_3 - T_2)\frac{h - h_3}{2} - (T_3 - T_1)\frac{h_3}{2} - (T_1 - T_2)\frac{h}{2} \tag{3.3}$$

The demand of A_1 and A_3 to be equal leads to Eq. 3.4.

$$T_3 = 2T_{bar} - T_2 + T_2\frac{h_3}{h} - T_1\frac{h_3}{h} \tag{3.4}$$

T_3 can also be expressed with the dry adiabate $\gamma_d = -\frac{g}{c_p}$ as Eq. 3.5.

$$T_3 = T_2 - \gamma_d(h - h_3) \tag{3.5}$$
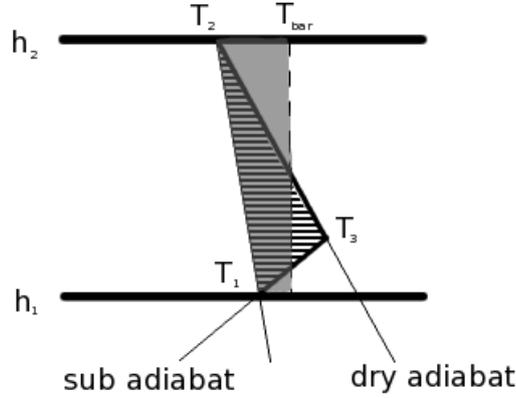
Figure 3.4.: Schematic diagram: construction of an inversion- and energy-preserving temperature profile. h_1 denotes the surface, h_2 the height of the lowest ECMWF pressure level above the orography. T_1 and T_2 are the corresponding temperatures prescribed by the ECMWF. The inversion temperature T_3 is constructed. Grey area A_1 denotes the extra thermal energy that is redistributed as indicated by the striped area (A_2).

Insertion in 3.4 leads to 3.6,

$$T_3 = 2T_{bar} - T_1 + (T_2 - T_1)\frac{T_3 - T_2}{h\gamma_d} \tag{3.6}$$

and finally leads to Eq. 3.7:

$$T_3 = \frac{(2T_{bar} - T_1)\gamma_d h - (T_2 - T_1)T_2}{\gamma_d h - (T_2 - T_1)} \tag{3.7}$$

Descending from the top of the layer, the temperature rises adiabatically closely towards the point, where all thermal energy is consumed. This inversion temperature T_3 is given in Eq. 3.7. From T_3 the temperature decreases sub adiabatically to the 2-m-temperature taken as ground level temperature, so that the triangular area built by the sides of dry adiabate, sub adiabate and linear temperature profile between the prescribed temperatures at 2 m and at the top of the layer equals the afore mentioned thermal energy.

As a result the temperature at the 10 m level is much smoother over the ocean (Fig. 3.3b) with the before mentioned locations of exceptionally low 10-m-forcing-temperatures well reduced (Fig. 3.3c). Most of the changes are in orographically structured areas. The forcing fields in the inner part of the model domain are only of influence during the initialization; for this the differences are shown in Fig. 3.3.

Only at this time the whole three dimensional field is imposed on the mesoscale solution. After the initialization is finished, forcing values are only used close to the lateral and upper boundaries. For this synoptical situation the effects on the lateral nudging are expected to be weak and confined to the north western and eastern model boundaries.

3.2. Impact of the vertical interpolation on model performance

The impact of the vertical interpolation on mesoscale model performance is demonstrated with the synoptic situation of an on-ice moving trough. Two METRAS simulations are performed. Both start for 5th March 2002, 18 UTC, when the trough is yet situated south east of the model domain, and last until 7th March 2002, 24 UTC. For a more in-depth discussion of the synoptic situation see Section 4.2.1. Test case M-A (set-up is specified in Section 4.3.2) is nudged with the alternatively interpolated forcing. Test case M-P has the same set-up but is nudged with the originally interpolated forcing. The impact will be evaluated at three time steps. The first is 5th March 2002, 19 UTC, one hour after model initialization. At this time the three dimensional homogeneous nudging is maximal (detailed explanation of the nudging procedure is given in 4.3.1). The second is 7th March 2002, 12 UTC, when aircraft based measurements are available. At that time the trough is situated approximately at the Greenwich Meridian. The last time step is 7th March 2002, 18 UTC, when a profile of a radiosonde launched from research vessel Aranda, positioned near the sea-ice edge is available.

The intended effect of the alternative interpolation in contrast to the original interpolation is shown with potential temperature profiles of the differently interpolated ECMWF data for 7th March 2002, 18 UTC (Fig. 3.5). In comparison to the radiosonde launched from RV Aranda at 79.28°N, 3.54°E, 680 m west to the ECMWF profile and 50 km east of the sea-ice edge, the alternative interpolation reproduces the observed inversion between 800 m and 1100 m slightly better than the original interpolation. However, the profiles of the associated METRAS simulations do not directly react on the forcing profiles, as the influence of the nudging in the inner domain is weak.

Fig. 3.6 shows the changes in stability from the original interpolation to the alternative interpolation for 5th March 2002, 19 UTC (a) and 7th March 2002, 12 UTC (b) with the change of the nudging potential temperature gradient between 50 m above ground and the surface. Unexpectedly, on 5th March 2002 at 19 UTC, the time when the three dimensional homogeneous forcing is strongest, the lower atmosphere is less
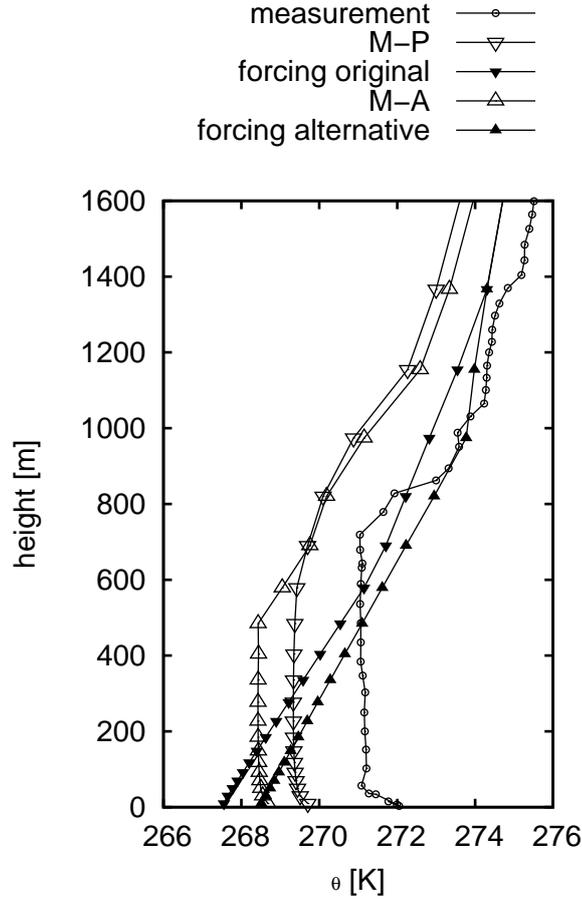
Figure 3.5.: Vertical profiles of potential temperature from observation, as used as nudging and of the corresponding METRAS results (M-A is forced with AVI ECMWF, M-P is forced with OVI ECMWF) for 7th March 2002, 18 UTC.

stable over most parts of the sea ice and more stable over the open ocean with the new scheme. Only at the later date the stratification is modified as expected with the atmosphere becoming more stable over the eastern half of the sea ice behind the trough, and less stable over the open ocean.

From Fig. 3.7 it can be seen, that the consideration of the 2-m-temperature in the AVI has a stronger effect on the stability than the relocation of the inversion. The figure shows temperature changes from OVI to AVI at the ground compared to the absolute changes $50\,m$ above ground:

$$\frac{\Theta_{h=2m,AVI} - \Theta_{h=2m,OVI}}{|\Theta_{h=50m,AVI} - \Theta_{h=50m,OVI}|} \tag{3.8}$$

Quotients with magnitude larger than 1 indicate a stronger change in surface temperature compared to the upper level temperature change. Areas with the quotient

larger 1 are destabilized with the AVI due to warming at the surface and areas smaller -1 are stabilized due to cooling at the surface.
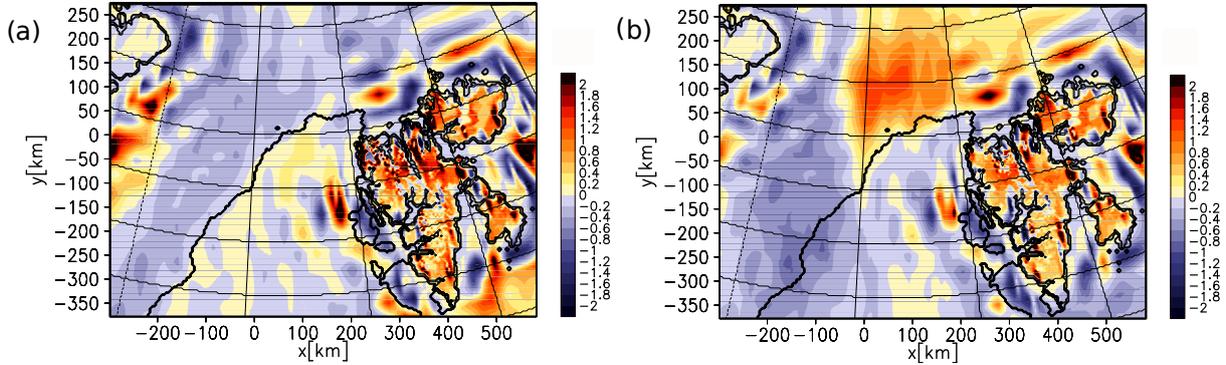
Figure 3.6.: Changes in the forcing potential temperature gradient between $50\,m$ above ground and the surface from the original interpolation to the alternative interpolation $((\Theta_{h=50m} - \Theta_{h=0m})_{AVI} - (\Theta_{h=50m} - \Theta_{h=0m})_{OVI})$ for 5th March 2002, 19 UTC (a) and 7th March 2002, 12 UTC (b). The changes [K] are shaded as indicated by the colour bars.
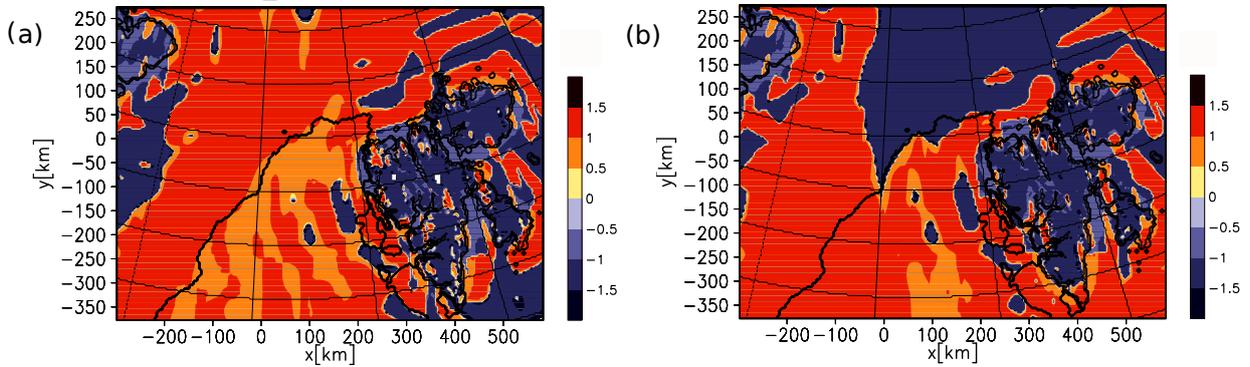
Figure 3.7.: Temperature changes from OVI to AVI at the ground compared to the absolute changes $50\,m$ above ground (Eq. 3.8) for 5th March 2002, 19 UTC (a) and 7th March 2002, 12 UTC (b). Shading as indicated by the colour bars.

For 5th March 2002, 19 UTC, test cases M-P and M-A show the same development of the stability, as described for the nudging (compare Fig. 3.8a to Fig. 3.6a). For the METRAS simulations the stability is computed as gradient of the potential temperature difference between the local inversion height and 10 m above ground divided

by its vertical distance. For the later time step, the METRAS test cases do not reproduce the changed stability in the forcing (compare Fig. 3.8b to Fig. 3.6b). Differences exist at the front due to different positioning, over Spitzbergen and at the lateral boundaries, where the nudging is active.
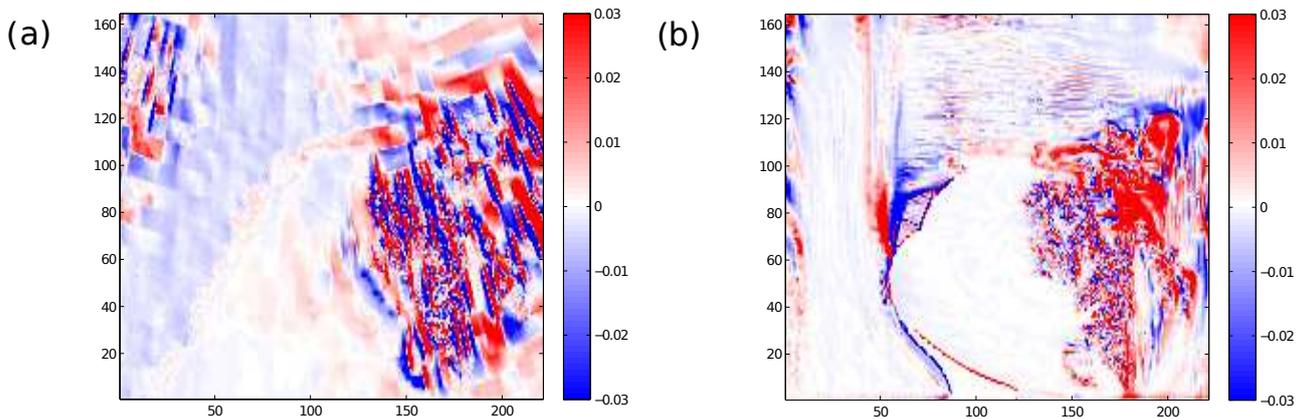
Figure 3.8.: Differences of METRAS potential temperature gradients between AVI and OVI for 5th March 2002, 19 UTC (a) and 7th March 2002, 12 UTC (b). The differences of gradients [K/m] are shaded as indicated by the colour bars. The gradients are computed as the differences between the potential temperatures at the local inversion heights and the potential temperatures at $10\,m$ above ground, divided by its vertical distance apart.

The performance of the model is evaluated near the ground with meteorological standard variables against measurements performed by the aircraft Falcon. As the flight track is located in the vicinity of the trough (Fig. 4.2 in Section 4.2.1), also the evaluation is strongly focussed on the model's ability to simulate the trough properly. Via baroclinicity the trough is strongly linked to the area of the strongest temperature gradient. Due to the stronger consideration of the 2-m-temperature in the simulation M-A, the temperature gradient is stronger bound to the sea-ice edge than in M-P (Fig. 3.9a, b). Also the difference plot of vorticity, pressure and horizontal wind (Fig. 3.9 c)) indicates, that the trough simulated with M-A remains further east, as the observed trough does. The pressure decrease to the trough is larger in M-A and better agreeing with the observations. Due to the better positioning of the trough, the Gandin-Murphy skill score, averaged over near-ground temperature, humidity, pressure and wind speed (GA, Section 4.3.3) improves with M-A (Table 3.1). The mean errors (ME) and root mean square errors (RMSE) of temperature, specific humidity and wind (also vector wind difference VWD) improve with M-A. Only the hit rates (H, Section 4.3.3) provide a contradictory picture. However, due to their

small numbers they are not considered reliable.

The omission of ECMWF's too low T_{bar} due to the artificial levelling has an effect on the temperature simulated by METRAS, where the layer thickness was reduced near the lateral boundaries. Comparatively warm air tongues in M-A exist north east of Greenland and east of Spitzbergen (compare Fig. 3.3c with Fig. 3.9b). But as they do not reach into the area observed by aircraft Falcon, it cannot be judged if they are realistic or not.

In conclusion, with AVI the atmospheric stability is stronger influenced by the near-ground temperature than by the relocation of the inversion. The use of ECMWF's 2-m-temperature is advantageous, because too cold near-water temperatures off-shore steep coasts can be avoided (Fig 3.3).

Table 3.1.: Model skills of test cases M-P and M-A for the horizontal legs (below 100 m above ground), averaged over 40 s running mean. N represents the number of samples used for the calculation of ME, RMSE, H, and VWD (explanations given in appendices A and B). HA and GA are the average hit rates and Gandin-Murphy skill scores of temperature, specific humidity, pressure, and wind speed.

N=6663	temperature [°C]			specific humidity [g kg⁻¹]			pressure [hPa]			wind speed [m s⁻¹]				wind direction [°]		average	
	ME	RMSE	H	ME	RMSE	H	ME	RMSE	H	ME	RMSE	VWD	H	RMSE	H	HA	GA
M-P	5.7	11.8	2	0.4	1.1	6	-4.7	5.2	5	-1.6	3.4	12.1	23	59	40	9	0.135
M-A	4.5	10.3	1	0.2	0.9	4	-5.4	5.7	1	-1.2	3.4	10.8	38	51	28	11	0.175
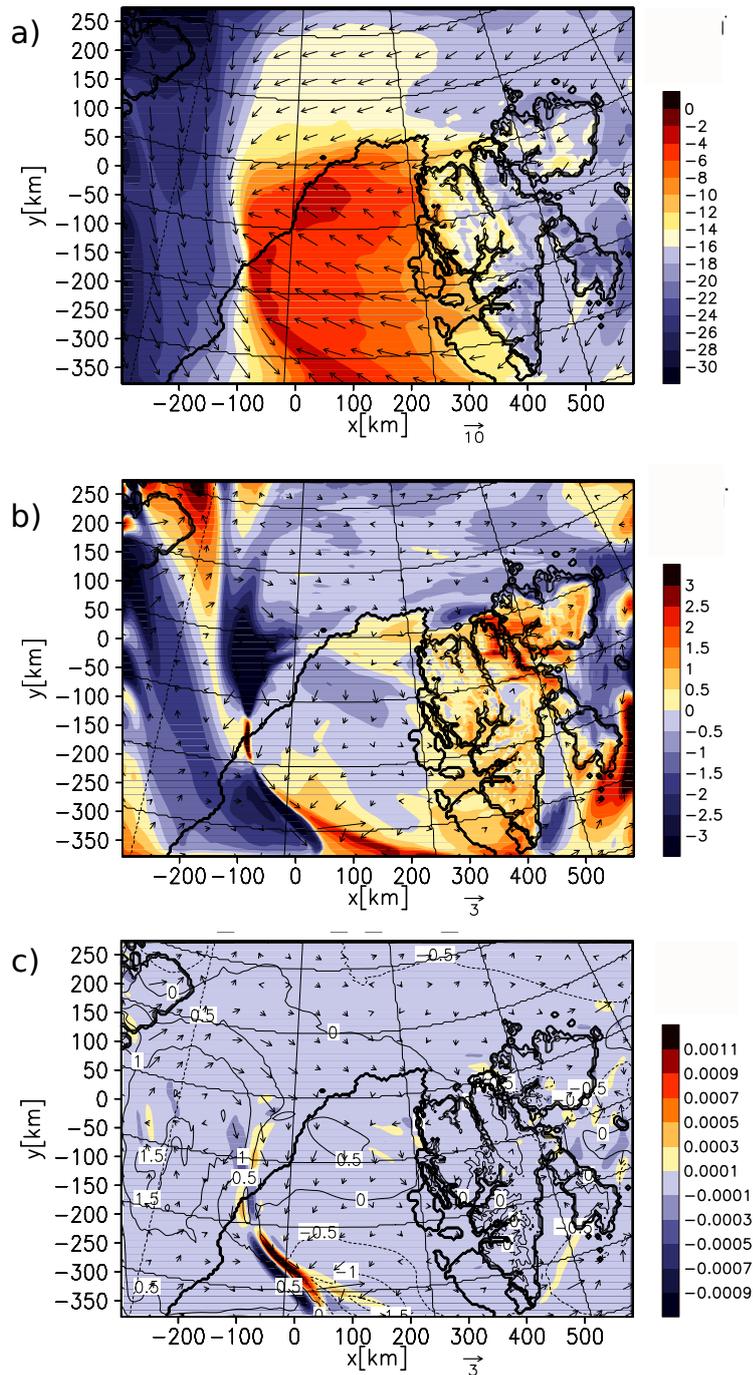
Figure 3.9.: Simulation results for 7th March 2002, 12 UTC for (a): M-P, temperature [°C] as indicated by the colour bar and wind field [m/s] as vectors at $10\,m$ above ground. (b): Difference fields of temperature and wind for M-P - M-A at the same time and height. The colour bar denotes temperature differences [K]. Vectors [m/s] denote differences in the wind field. (c): Difference fields of mean sea level pressure (hPa, labelled contours), and vorticity (s^{-1}, shaded) and wind field (m/s, vectors) at $10\,m$ above ground.

4. Impact of surface parameter uncertainties on the development of a trough in the Fram Strait region

Since the chapter is accepted for publication, the notation may deviate from the one used in the other chapters of this thesis. For purposes of comparison to Section 5.3.3, Table 4.5 is expanded by hit rates and Gandin-Murphy Skill Scores averaged over temperature humidity and wind speed.

Summary

The impacts of the sea-ice characteristics distribution, roughness, temperature and thermal conductivity on an on-ice moving trough in the Fram Strait on 7 March 2002 are investigated. The situation is simulated with the mesoscale transport and flow model METRAS and the named characteristics are varied within the range of observational uncertainty. The test cases are evaluated against aircraft measurements performed within the "Fram Strait Cyclone Experiment 2002". The model's sensitivity on the changes in sea-ice characteristics is quantified by statistical means. The strongest impacts on the near-ground temperature are found from sea-ice temperature, manifesting as an overall bias, and the positioning of the sea-ice edge, manifesting as a phase error. Only higher than natural homogenisation of the sea-ice cover leads to some reduction of the amplitude error. A reduction of the sea-ice surface roughness is performed by applying an unrealistically small roughness length of $z_0 = 1\,mm$. This reduces the negative wind speed bias, enhances the advection of contrasting air masses and improves the frontal sharpness. The thermal conductivity has the smallest influence. The lateral forcing taken from "European Centre for Medium-Range Weather Forecasts" (ECMWF) reanalyses shows the strongest effect on the limited area model performance.

4.1. Introduction

The sea-ice export through the Fram Strait plays an important role for the thermohaline circulation by transporting large amounts of freshwater into the North Atlantic Ocean. Changes in the sea-ice export occur not only at interannual to decadal time scales but also at very short scales from several hours to a few days. This large variability is mainly due to synoptic scale and mesoscale disturbances passing through the Fram Strait. The development of these atmospheric disturbances is influenced by the sea-ice characteristics (Dierer and Schlünzen, 2005). Via turbulent momentum- and heat fluxes mesoscale structures like fronts and polar cyclones are linked to surface characteristics such as roughness or sea-ice distribution. Furthermore, the atmospheric drag coefficient plays an important role as relation of the surface stress on the sea ice to the wind speed just above the sea ice. Resultingly the wind is the principal predictor of the sea-ice drift (Thorndike and Colony, 1982). For the atmospheric disturbances the vertical wind shear, defined by the drag coefficient, is of primary importance.

Several studies have been performed on atmospheric processes in the marginal ice zone. One of which describes a weak warm front observed over the Fram Strait in spring 1989 (Rasmussen et al., 1997). In their study the baroclinicity was strong enough to form a mesoscale vortex. Also Brümmer and Höber (1999) reported a

45

warm front, passing westward through the Fram Strait, that starts to develop a cyclone. From measurements and 2D–simulations with METRAS Vihma et al. (2003) identified the sea-ice roughness to play a most important role in the formation of an internal boundary layer in on-ice flow. Valkonen et al. (2008) found the sea-ice concentration to play an important role for the near-ground temperature and wind field during a study in the Weddell Sea. For the simulation of a mesoscale cyclone passage with dynamic changes in the sea-ice distribution a strong influence on the heat fluxes has been found when compared to simulations with constant sea-ice distribution (Dierer and Schlünzen, 2005).

Here the development of an on-ice moving trough in the Fram Strait on 7 March 2002 is investigated. The findings of Vihma et al. (2003) and Valkonen et al. (2008) also suggest a major influence of sea-ice roughness and sea-ice concentration for this type of synoptic situations. In addition, the main thermal sea-ice properties, temperature and thermal conductivity, are assumed to have an important influence on the trough development over sea ice. These characteristics are highly variable in space. The characteristics distribution and roughness also vary strongly in time. Measurements of sea ice are predominantly performed by remote sensing. If used as characterising parameters for model simulations, their uncertainty is further enlarged due to the still limited spatial and temporal resolution. In our model simulations the named sea-ice characteristics are varied within the range of observational uncertainty and their importance on the model performance is estimated. The simulations are performed with the mesoscale transport and fluid model METRAS (Schlünzen, 1990; Lüpkes and Schlünzen, 1996). Parameterisation options that have been proven to deliver most realistic results for high latitude applications have been chosen (Lüpkes and Birnbaum, 2005; Dierer et al., 2005). The model treats the sea ice as invariant with respect to position and the characteristics mentioned (except sea-ice temperature). METRAS is nested into "European Centre for Medium-Range Weather Forecasts" (ECMWF) reanalyses. The model's sensitivity on the lateral forcing is investigated with two different nudging coefficients.

The test cases are evaluated against aircraft measurements that have been performed during the "Fram Strait Cyclone Experiment 2002" (FRAMZY 2002; Brümmer et al., 2005). An assessment with regard to the characteristics' influence on the model performance is given on basis of statistical measures. The best performing combination is eventually identified.

In Section 4.2 the meteorological situation and the experimental setup are described in detail. In Section 4.3 the model setup and forcing are discussed. Results are given in Section 4.4. Conclusions are drawn in Section 4.5.

46

4.2. Comparison data and experiment

4.2.1. Synoptic situation: trough passage on 7 March 2002

The synoptic situation investigated is characterized by an on-ice moving trough. A cyclone over Scandinavia was corresponding to a low at the $500\,hPa$ level and was associated with the Greenland anticyclone. Accordingly, the geostrophic flow along the isohypsic contours of the $500\,hPa$ surface over the Fram Strait was south-easterly (Fig. 4.1). Within this flow a north-south directed trough crossed the Fram Strait from east to west reaching the Greenwich meridian (approximately the western edge of the sea ice) at 11 UTC on 7 March 2002 (Fig. 4.2). In the planetary boundary layer the trough had distinct frontal contrasts with a horizontal temperature difference of $25\,K$ over $100\,km$, a humidity mixing ratio difference of $3\,g\,kg^{-1}$ and a surface pressure drop of about $3\,hPa$ over the $100\,km$ distance. On the trough's cold western side a near-ground northerly flow of about $15\,m\,s^{-1}$, a cloud cover of 8/8 stratus with snow showers and low visibility prevailed. The warmer eastern side showed near-ground south-easterly to easterly winds of $10\,m\,s^{-1}$. The cloud situation was characterised by 5/8 to 8/8 strato cumulus with a clearly marked cloud base and the cloud top lower than on the western side (Brümmer et al., 2005). The profiles over the ocean at $78.84°N$, $7.87°E$ (denoted P1 in Figs. 4.2, 4.3 and 4.4) and over the sea ice at $80.6°N$, $5.16°W$ (denoted P7 in Figs. 4.2, 4.3 and 4.4) both show a well mixed boundary layer (Fig. 4.3).
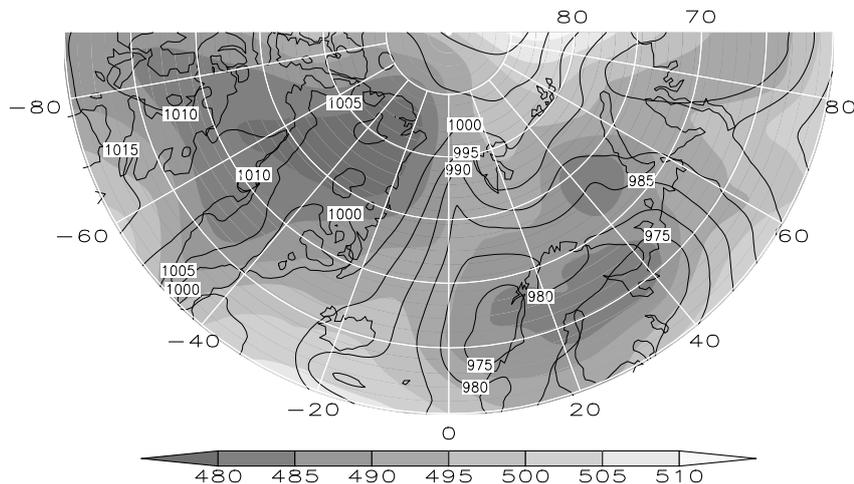
Figure 4.1.: ECMWF reanalysis of $500\,hPa$ geopotential heights (gpdm) (shaded) and mean sea level pressure (hPa) (contours) for 0 UTC, 7 March 2002.
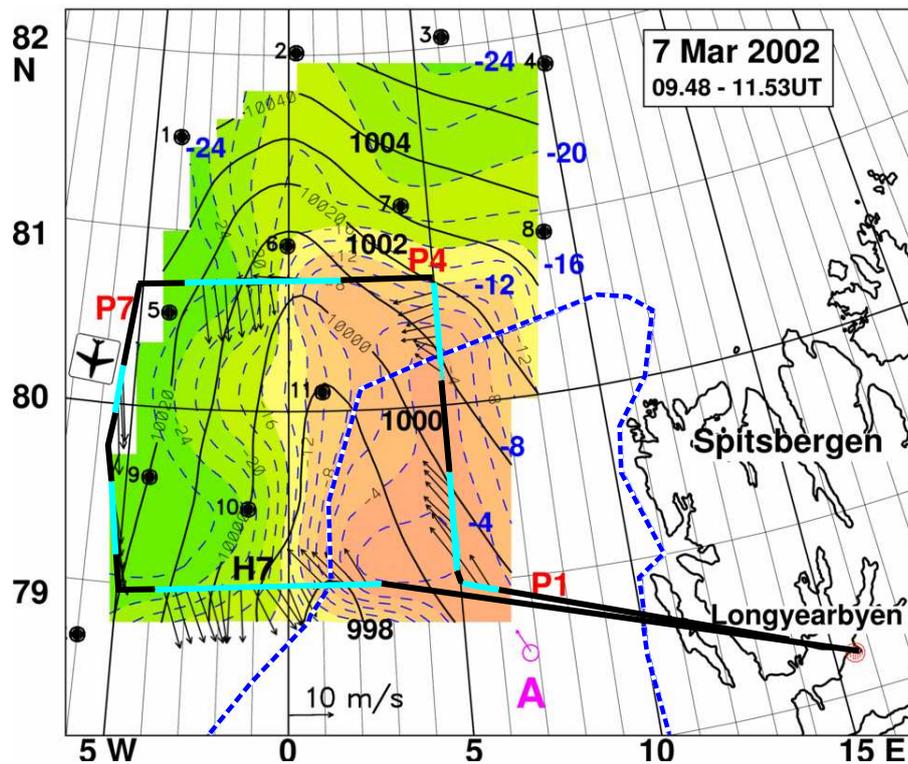
Figure 4.2.: Wind vector ($m\,s^{-1}$) (arrows), surface isobars (hPa) (black solid) and temperature (°C) (blue long dashed and colours) as measured during low-level flight tracks (thick light blue lines) by Falcon aircraft, by ARGOS ice buoys (numbered black points) and by RV Aranda (A) during FRAMZY 2002 on 7 March 2002 around 11 UTC. The wind vector is averaged over a flight distance of $7\,km$. Also indicated is the entire flight pattern (black solid) and the sea-ice edge (blue dashed line). From Brümmer et al. (2005).

The inversion height is higher over the ocean ($700\,m$) than over sea ice ($250\,m$). The surface isotherms indicate a well defined low level baroclinic zone over the sea ice with the thermal wind in the opposite direction than the near-ground flow (Fig. 4.2). In the free atmosphere this causes a decrease of the geostrophic wind with height and the formation of a weak low level jet just above the inversion (P7 in Fig. 4.3b)). The veering of the geostrophic wind with height from N to SE in $900\,m$ at P7 denotes warm air advection (Fig. 4.3c)).

Primarily due to the atmospheric forcing the northern part of the sea-ice edge shifted approximately $30\,km$ north, while the south-western sea-ice edge shifted approximately $30\,km$ east (Fig. 4.5). For the atmospheric development the sea-ice edge is an important factor but the distribution of the sea-ice concentration is just as important (Dierer et al., 2005). Due to the trough's impacts the high concentrations of sea ice are reduced, resulting in a more homogeneous distribution (Fig. 4.6).
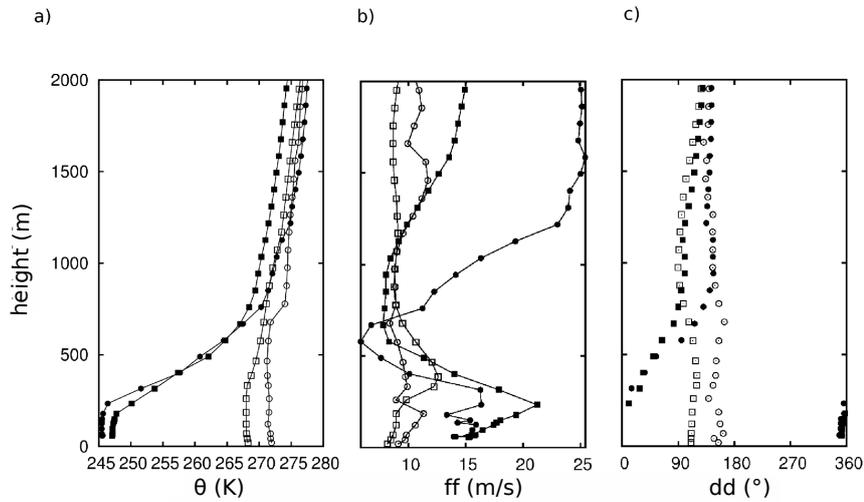
Figure 4.3.: Observed (dots) and modelled (rectangles) vertical profiles of a) potential temperature Θ, b) wind speed, and c) wind direction at locations P1 (open symbols) and P7 (full symbols) on 7 March 2002 9:50 UTC and 10:55 UTC, respectively. For positions of profiles P1, P7 see Fig. 4.2.
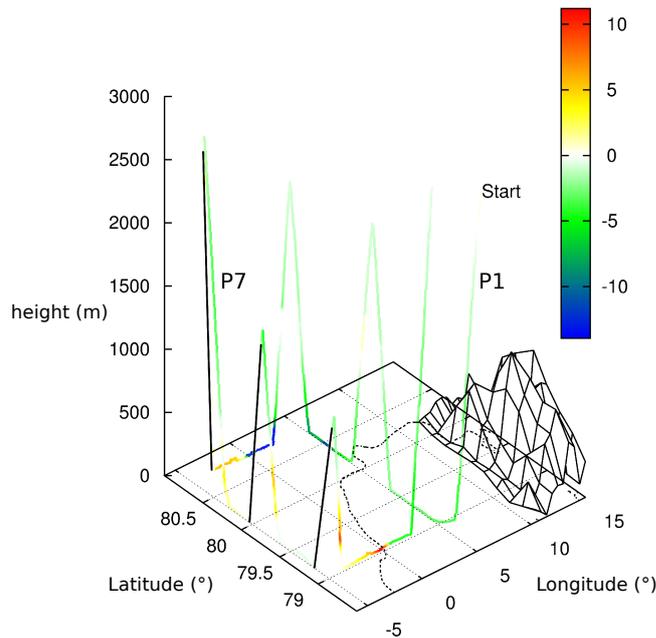
Figure 4.4.: Flight profile flown counter-clockwise on 7 March 2002 09:47 UTC to 11:53 UTC with temperature deviation of test case M-E from measurements (K) as indicated by the colour bar. Black profile where measurement failed. The dotted line marks the sea-ice edge as measured on 7 March 2002 at 18 UTC, Spitzbergen is shown schematically.
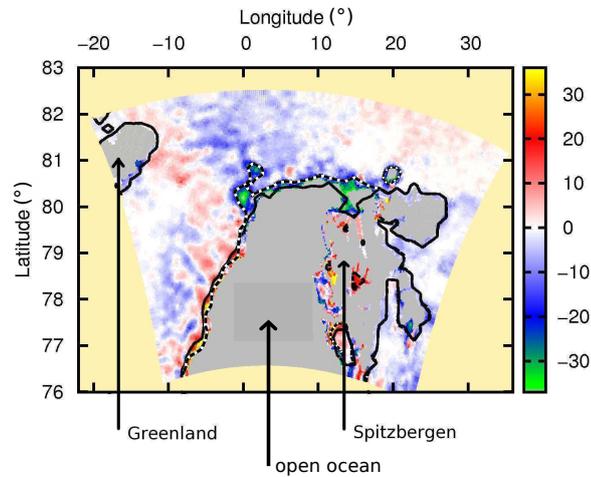
Figure 4.5.: Difference between observed sea-ice distributions of 7 March 2002 18 UTC and 6 March 2002 14 UTC. Areas coloured in yellow are outside of the METRAS model domain. The colour bar indicates differences in percentage of sea-ice cover; note that the range of differences is not fully resolved by the colour scale. Sea-ice edges are indicated as isolines of 75 % sea-ice concentration at the beginning (continuous line) and at the end (dashed line) of the difference interval. Areas with sea-ice concentrations below 30 % on one of the two dates are disregarded to account for uncertainties of the observations (indicated by area shaded in grey).
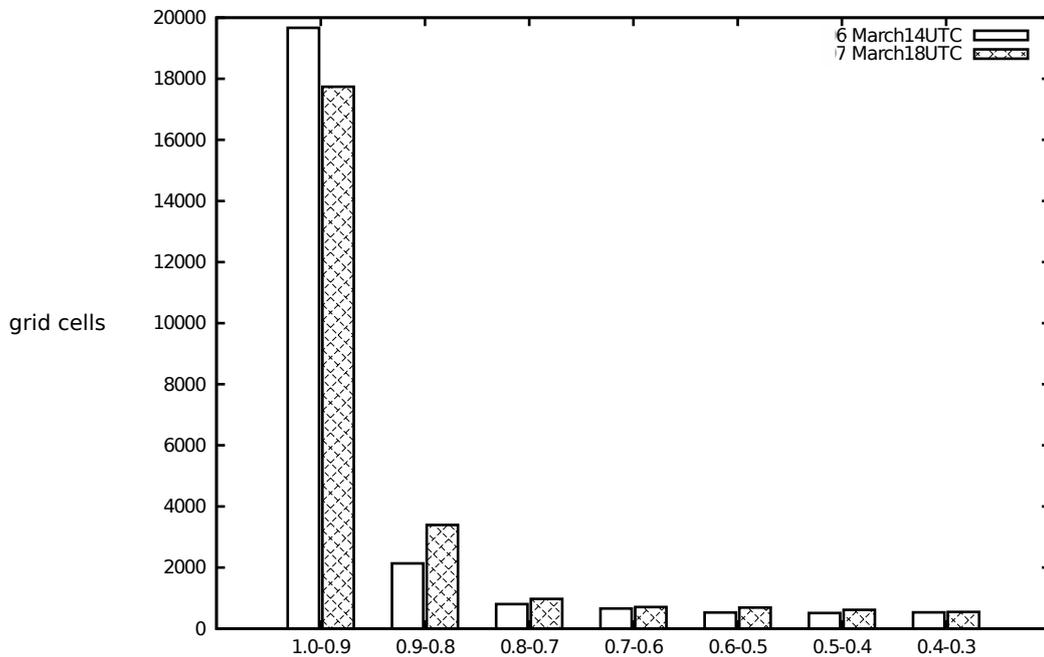
Figure 4.6.: Histograms of measured sea-ice concentrations of METRAS grid cells for the whole domain.

4.2.2. Comparison data: aircraft, ship, buoys

The METRAS model simulations for the trough passage described in section 4.2.1 are evaluated against measurements gathered within FRAMZY 2002 (Brümmer et al., 2005). This field experiment on cyclones in the Fram Strait and their impact on sea ice began on 25 February 2002 and lasted 4 weeks. Measurements were performed amongst others with the aircraft Falcon, the research vessel Aranda and sea-ice buoys. The trough crossed over the simulation domain on 7 March 2002 shortly before noon and was measured, using aircraft Falcon, from 09:48 UTC to 11:53 UTC along the path shown in Fig. 4.2.

The meteorological parameters considered here and measured from the aircraft are temperature, specific humidity, mixing ratio, pressure and the horizontal wind vector. All measurements were recorded at a sampling rate of $100\,Hz$. For the given aircraft speed of $100\,m\,s^{-1}$ a spatial resolution of $1\,m$ results. The used sensor types and the specifications are given in Brümmer et al. (2005).

The airborne measurements are complemented by hourly recordings from an array of buoys on the sea ice. The primary intention for deploying the buoys was to determine the sea-ice drift. Apart from measuring the position, all buoys were equipped with pressure and temperature sensors. Due to the special location of the temperature sensors inside the buoys and the unknown height of the sensor above the surface the measurements were only used for a qualitative evaluation. Additionally, one buoy was measuring wind speed and wind direction. However, this measurement is uncertain and thus not included. None of these measurements were provided to the ECMWF, so that the ECMWF reanalyses represent a completely independent interpretation of the meteorological situation.

4.3. Model set-up and forcing

The test cases are simulated with METRAS. METRAS is a nonhydrostatic mesoscale model which employs the Boussinesq-approximation, the anelastic approximation and a domain-constant Coriolis parameter. For Arctic regions the model has been applied for real case studies (Dierer et al., 2005; Viehma et al., 2003; Lüpkes and Schlünzen, 1996) as well as for process studies (Lüpkes et al., 2008a; Lüpkes et al., 2008b). For turbulence parameterisation the similarity theory is used below $10\,m$ height. Above $10\,m$, the counter-gradient scheme of Lüpkes and Schlünzen (1996) is used for convective conditions, and the mixing length scheme is used in the case of stable stratification. Horizontal diffusion is included to a sufficient amount by the numerical scheme (upstream) for temperature and humidity. For wind the seven

point filter (Adams-Bashforth with centered differences) used ensures horizontal diffusion. Parameterisations selected are the Kessler scheme for cloud microphysical processes and a two stream approximation scheme for radiation. Surface subgrid-scale characteristics are considered in the model. All surface fluxes are calculated by using averaged subgrid-scale surface fluxes applying the blending height concept (von Salzen et al., 1996).

For METRAS the temperature at the surface (T_S) is predicted by the force-restore method (Deardorff, 1978), see Eq. 4.1.

$$\frac{\partial T_S}{\partial t} = \frac{2\sqrt{\pi}k_s}{\nu_s h_\theta}H_A - 2\pi\frac{T_S - T_h}{\tau_1} \tag{4.1}$$

The first term on the right side describes the absorption of heat provided by the sum of all atmospheric fluxes (H_A) by the ground and thus forces changes in T_S. Changes in surface temperature depend on the ground properties thermal diffusivity (k_s), thermal conductivity (ν_s), and depth of the diurnal temperature wave (h_θ). The values used for sea ice are specified in Table 4.1. The depth of the temperature wave depends on thermal diffusivity and the time span τ_1, here one day: $h_\theta = \sqrt{k_s\tau_1}$, yielding $h_\theta = 0.26\,m$. The second term tends to restore T_S within one day to the temperature T_h valid in the depth h.

METRAS is run with a $4\,km$ horizontal resolution and 33 vertical levels. Close-to-surface resolution is $20\,m$ with the lowest model level at $10\,m$. The resolution increases above $80\,m$ with height by up to $20\,\%$ from grid level to grid level and is approximately $200\,m$ at $1000\,m$ above ground. The uppermost model level is at $13501\,m$. The domain covers a region of $900\,km$ (west-east) by $650\,km$ with Spitzbergen on the easterly side and the north eastern tip of Greenland in the north-west

Table 4.1.: Surface model parameters for sea ice.

parameter	value	source
z_0	$1\,mm$ (fused pancake sea ice[a]), $5.9\,mm$ (strongly deformed sea ice[b]), $27\,mm$ (very rough multi year sea ice[c])	[a], [c]: Guest and Davidson (1991) [b]: Lüpkes and Birnbaum (2005)
$T_h(ice)$	$-9°$C ([a]), $-25°$C ([b])	[a]: Putkonen (1998), [b]: Brümmer et al. (2005)
ν_s	$1 Jm^{-1}s^{-1}K^{-1}$ (snow), $2.025 Jm^{-1}s^{-1}K^{-1}$ (sea ice)	Pielke (2002)
k_s	$0.8 \times 10^{-6}m^2s^{-1}$	Pielke (2002)
albedo	$72\,\%$ (pure white snow)	Pielke (2002)

(Fig. 4.5). All simulations start for 5 March 2002, 1800 UTC, and are interpreted 40 hours after model initialisation for the period lasting from 7 March 09:45 UTC to 12 UTC. For this period aircraft data are available from the FRAMZY 2002 field experiment (section 4.2.2).

In the present investigation the ocean partly consists of sea ice and its surface characteristics such as distribution, roughness, temperature, thermal conductivity and albedo as well as its extent vary in time. However, assuming the time dependency of the sea-ice extent and the named sea-ice characteristics mainly on larger than daily time scales, the land surface model is also applied for sea ice. Except for sea-ice surface temperature and sea-ice water content the land surface model treats the sea-ice characteristics as constant parameters. These parameters have to be prescribed on basis of measurements (Table 4.1).

In the frame of the sensitivity study the parameters, z_0, representing sea-ice roughness, $T_h(ice)$, representing the mean sea-ice temperature in the layer of $10\,cm$ to $2\,m$, and ν_s, thermal conductivity, are varied stepwise between the values given in Table 4.1. All these values lie in the reasonable range that is covered by the measurements as mentioned in Table 4.1. Only the extreme roughnesses exaggerate realistic areal averages, since in reality the sea-ice characteristics in the Fram Strait strongly change in zonal direction. The absence of an isolating snow cover in the land surface model and k_s and ν_s set to the values of pure sea ice justify the use of a sea-ice temperature representing a layer nearer to the surface. Besides $T_h(ice) = -9°C$ as measured in the specified depth, the approximate mean surface temperature of $-25°C$ as measured by aircraft Falcon (Brümmer et al., 2005) is also used as $T_h(ice)$. With regard to the unknown snow height both thermal conductivities, that for sea ice and that for snow are used. The albedo of thin ice sheets grows with the sun's zenith angle. As the minimal zenith angle for the domain during the simulation phase is 84°, this fact is accounted for by setting the albedo to the relative large value of pure white snow.

4.3.1. Forcing data and their assimilation

To prescribe the large-scale situation, four different datasets are used. For the atmospheric conditions reanalyses provided by the ECMWF are taken. The sea surface temperature is derived from North Atlantic/Arctic Ocean Sea Ice Model (NAOSIM) results. Topography data are taken from "United States Geological Survey" (USGS; http://edc2.usgs.gov/glcc/) in a resolution of 3". Satellite derived sea-ice cover data are used to determine the sea-ice distribution. The data sources and the data transfer to METRAS are explained in the following.

ECMWF

The large scale meteorological condition is taken from ECMWF six hourly reanalyses with 0.375° horizontal resolution (Persson and Grazzini, 2005). ECMWF wind-, temperature-, specific humidity- and cloud water-fields are used as initial and lateral boundary data. The forced solution Ψ_f is obtained by applying a variable weighting factor δ (Eq. 4.2).

$$\Psi_f = \Psi_m + \delta(\Psi_l - \Psi_m) \tag{4.2}$$

Ψ_m denotes the mesoscale model solution and Ψ_l the solution of the driving model. The weighting factor is linked to the nudging coefficient ν with the time step of the model (Δt, Eq. 4.3).

$$\delta = \nu \Delta t \tag{4.3}$$

Within the first half of the initialisation phase of 110 minutes, ν increases from $0\,s^{-1}$ to $0.001\,s^{-1}$ (or $0.01\,s^{-1}$). These values are spatially constant. In the second half of the initialisation phase ν reduces to $0\,s^{-1}$ in the inner domain. During the rest of the simulation time the values for ν are kept constant and nudging only acts at the lateral boundary grid points following equation 4.4.

$$\nu(i) = \nu_0(1 - tanh(a_f \cdot i)) \tag{4.4}$$

Here, i is the distance from the lateral boundary in grid points and ν_0 is set to $0.001\,s^{-1}$ or $0.01\,s^{-1}$. For a_f set to 0.4 or 0.2, ν reduces by about one order of magnitude over the outermost 4 gridpoints and 8 gridpoints, respectively. The combination of $\nu_0 = 0.001\,s^{-1}$ and $a_f = 0.4$ is used as weak forcing. The combination of $\nu_0 = 0.01\,s^{-1}$ and $a_f = 0.2$ is used as strong forcing. A detailed description of this nudging technique is given in Davies (1976). The temporal interpolation is performed linearly (Bungert, 2008). The spatial interpolation of the forcing fields given at pressure levels to the height-dependent grid used in METRAS is mainly performed linearly (Bohnenstengel, 2010).

NAOSIM

NAOSIM is based on the Geophysical Fluid Dynamics Laboratory modular ocean model (Pacanowski, 1995). A description of the changes is given in Karcher et al. (2003). The data are given in 1/12° resolution (Karcher et al., 2005). The vertical temperature profile in the upper ocean is set constant, as frequently observed in the well mixed surface layer. Therefore NAOSIM's water temperature field, valid in the depth of $5\,m$, is used to force the sea surface temperature of METRAS.

The horizontal structure of the simulated temperature field well reflects the temperature structure measured from ship and aircraft (Fig. 4.7). However, the simulated

differences of $8\,K$ within the model area and maximum temperatures of $6\,°C$ at the sea-ice edge are relatively high. The big differences are caused by convective up-welling - triggered by atmospheric cooling - of too warm water from greater depths. The deep water is too warm, because of weak cooling by mixing on its way from the Iceland-Faroe Ridge to the Fram Strait, especially at the Greenland Gyre (personal communication, Filip Hacker, Alfred Wegener Institute 2007). Therefore, we confine the maximum sea surface temperature to $3\,°C$, which is approximately the maximum temperature measured by ship and aircraft during the simulation phase.

The original data are brought to the METRAS uniform grid by firstly rotating to geographical coordinates and secondly interpolating by weighting the four nearest neighbours with their inverse distances. Land points are disregarded. If all four nearest neighbours are land points, the value of the next nearest non land point is taken. At rugged coastlines this procedure may result in longer than necessary water ways for the extrapolation.
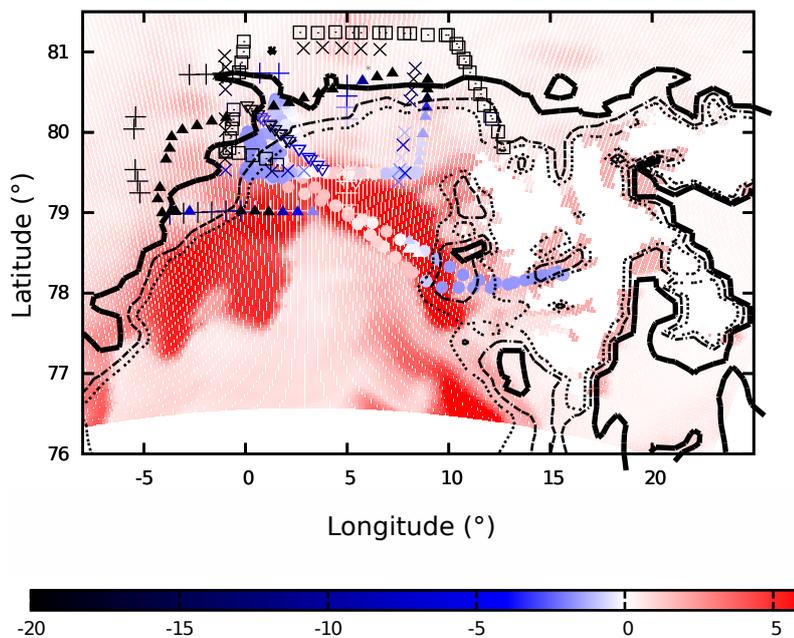
Figure 4.7.: Water temperatures, surface temperatures (°C) and sea-ice edge. Water temperatures from $5\,m$ depth simulated by NAOSIM for 6 March 2002 0 UTC (shaded) and measured by RV Aranda in $1.5\,m$ depth from 5 March 2002 to 15 March 2002 (shaded dots). Surface temperatures measured by aircraft Falcon for 5 March 2002 (shaded box), 7 March 2002 (shaded plus), 10 March 2002 (shaded down-pointing triangle), 13 March 2002 (shaded saltire), 15 March 2002 (shaded full triangles). All symbols have the same colour scale. Surface temperature reaches $-37\,°C$. Continuous, dash-dotted and dotted lines show isolines of $90\,\%$, $60\,\%$ and $30\,\%$ sea-ice concentration as observed by satellite on 6 March 2002 at 14 UTC.

Satellite observations

The fractional sea-ice cover is calculated from the "Special Sensor Microwave/Imager"-radiometer (SSM/I) measurements (Brümmer et al., 2008) with the ARTIST sea-ice algorithm (Kaleschke et al., 2001). SSM/I scans are available approximately every 2 hours during daylight hours, but rarely cover the whole simulation domain. For the time span investigated only two complete scans are available on 6 March 2002 at 14 UTC and on 7 March 2002 at 18 UTC. The interpolation from its original field of adjacent scans with $12.5\,km$ footprint to the METRAS uniform grid with $4\,km$ spacing is performed by linearly weighting the four nearest neighbours with their inverse distances. Because the ARTIST sea-ice algorithm does not work correctly on land, SSM/I scans containing coastlines have to be masked out before interpolation. The pattern of sea-ice cover obtained for 6 March 2002 14 UTC agrees well with the pattern of surface temperatures measured by aircraft Falcon before the trough passage on 7 March 2002 (Fig. 4.7). Surface temperatures in the range between the mean sea-ice surface temperature of $-25°$C and the minimal sea surface temperature of $-1.8\,°$C result from interpolation over the instrument's footprint on partly sea-ice covered ocean. Test cases are also simulated for the sea-ice map of 7 March 2002 18 UTC which is processed analogously. For sea-ice concentrations below $30\,\%$ SSM/I measurement uncertainties lie in the same range as the measurement itself (Spreen et al., 2008). Nevertheless these uncertainties are not considered during preprocessing and model initialization.

4.3.2. Experimental setup

All test cases are simulated for the period of 5 March 2002 18 UTC to 7 March 2002 24 UTC. The sensitivity of the simulated trough development during the aircraft measurement on the intensity of lateral forcing and several sea-ice characteristics is investigated. These are roughness, sea-ice temperature, thermal conductivity and sea-ice concentration. The first three are varied by changing the respective parameters to any of the values given in Table 4.1. The sea-ice concentration is varied by applying sea-ice maps from two different dates (6 March 2002 14 UTC or 7 March 2002 18 UTC). The forcing is varied by applying two different magnitudes of lateral forcing (section 4.3.1). From all possible combinations of variations 13 test cases, named "M-A", "M-B" to "M-M" are executed within the frame of the sensitivity study (Table 4.2). The sensitivity study is organised in five model experiments named "DISTRIBUTION", "ROUGHNESS", "FORCING", "TEMPERATURE" and "CONDUCTIVITY". Each test case is grouped into these experiments if a complementary test case exists whose setup only differs in the parameter the respective experiment is dealing with. The grouping leads to pairs of complementary test cases which are specified in Table 4.3. An additional test case M-N, with an

56

artificial sea-ice concentration does not directly contribute to the sensitivity study. The artificial sea-ice map uses the sea-ice edge taken from 6 March 2002 14 UTC and every grid cell's sea-ice concentration is set to 100 % if any sea ice is diagnosed within the corresponding grid cell.

4.3.3. Evaluation Method

The general evaluation of the atmosphere simulation considers the observed and modelled meteorological standard parameters temperature, specific humidity, pressure and wind at horizontal flight legs approximately $15\,m$ above ground. The mean error (ME), root mean square error (RMSE) as used in Ries and Schlünzen (2009) and the hit rate H (Appendix B, values for desired simulation accuracy A are given in Table 4.4) are computed for temperature, specific humidity and pressure. For wind velocity the vector wind difference (VWD) is additionally calculated as used in Ries and Schlünzen (2009), for wind direction only RMSE and H. GA and HA are G and H averaged over the meteorological parameters temperature, pressure, specific humidity and wind speed. The modelled stability is discussed qualitatively based on vertical profiles of potential temperature, wind direction and wind speed. For the assessment of simulated trough positioning we concentrate on the near-ground wind- and temperature fields.

Table 4.2.: Setup of all test cases.

test case	characteristic				
	forcing	roughness	temperature	thermal conductivity	distribution
	weak: $a_f = 0.4$, $\nu_0 = 0.001 s^{-1}$ strong: $a_f = 0.2, \nu_0 = 0.01 s^{-1}$	$z_0\,(mm)$	$T_h(ice)\,(^\circ C)$	$\nu_s\,(Jm^{-1}s^{-1}K^{-1})$	early: 6 March 2002 14 UTC late: 7 March 2002 18 UTC
M-A	weak	5.9	-9	1	early
M-B	weak	27	-9	1	early
M-C	weak	5.9	-9	1	late
M-D	weak	27	-9	1	late
M-E	strong	5.9	-25	1	late
M-F	strong	5.9	-9	1	late
M-G	weak	1.0	-9	1	late
M-H	weak	5.9	-9	2.025	late
M-I	weak	27	-25	1	early
M-J	weak	5.9	-25	1	late
M-K	weak	27	-25	1	late
M-L	strong	1.0	-25	1	late
M-M	weak	1.0	-25	1	late
M-N	weak	5.9	-25	1	filled to 100 %

Table 4.3.: Pairs of test cases grouped into experiments, that are built to determine the model sensitivity.

experiment	FORCING	ROUGHNESS	TEMPERATURE	CONDUCTIVITY	DISTRIBUTION
	M-M/M-L	M-D/M-G	M-D/M-K	M-C/M-H	M-B/M-D
	M-J/M-E	M-K/M-M	M-G/M-M		M-A/M-C
	M-C/M-F	M-E/M-L	M-C/M-J		M-I/M-K
		M-B/M-A	M-B/M-I		
			M-F/M-E		

Table 4.4.: Desired accuracy A of model results and accepted measurement values.

	temperature (°C)	specific humidity ($g\,kg^{-1}$)	pressure (hPa)	wind speed ($m\,s^{-1}$)	wind direction (°)
Accuracy	±0.5	±0.2	±1.5	±1.4	±20
Threshold				1<wind speed	1<wind speed

The simulation results valid along the aircraft track (Fig. 4.4) are extracted on-line. During integration the simulation results of the eight surrounding grid points are linearly interpolated to the aircraft position. The timestep nearest to the measurement time is used. A $40\,s$ running mean is applied to the measurements, taken during horizontal flight legs flown at a speed of $100\,m\,s^{-1}$, to obtain a spatial representativeness similar to the simulations at $4\,km$ grid spacing.

The test cases are assessed and the model's sensitivity is measured by means of the statistical measures mentioned above and the Gandin-Murphy skill score (Gandin and Murphy, 1992). The Gandin-Murphy skill score (G) is used with the weights as proposed by Gerrity (1992), see Appendix C. G reflects very well differences in the patterns and is not bias sensitive as, for example H. The term "pattern" encompasses phenomena on a spatial and temporal scale and hence the pattern correspondence is affected by errors in standard deviation, amplitude and phase. These types of errors are visible in the Taylor diagram (Taylor, 2001) which uses correlation and standard deviation as coordinates. Pure errors in standard deviation are primarily reflected by a reduction of correlation. Pure amplitude errors deviate the normalised standard deviation from one but leave the correlation unchanged. Pure phase errors are reflected by a reduction of correlation but leave the standard deviation unchanged.

The centered pattern root mean square difference Δc (Eq. 4.5) is the measure of error inherent of the Taylor diagram,

$$\Delta c = \left(\frac{1}{N} \sum_{n=1}^{N} \left[\left(f_n - \bar{f} \right) - \left(r_n - \bar{r} \right) \right]^2 \right)^{0.5} \tag{4.5}$$

where N is the number of samples, f_n a single simulated value and r_n is a single observed value.

4.4. Simulation results

4.4.1. Typical result features

For a qualitative impression the surface pressure and near-ground fields of wind and temperature of the best performing test case (M-E; HA=24 %, Table 4.5) are given in Fig. 4.8. The trough is the most characteristic feature of this simulation. Its positioning and sharpness strongly influence the advection and therefore are of primary influence on the quantitative evaluation of temperature- and wind speed-simulations.

Results of test case M-E show a principally correct positioning of the trough (Fig. 4.2 and Fig. 4.8). However, a detailed comparison to the measurements reveals that the northerly extension of the trough as apparent from isobars is underestimated. The pressure increase from the trough centre to the western and eastern edges of the measured field at $79°N$ is similar. It reaches for the measured field $4\,hPa$ and $2\,hPa$, whereas the simulation shows $3\,hPa$ and $1\,hPa$. The temperature increase measured at $79°N$ from the western boundary to the temperature maximum is $23\,K$. The simulated increase is $21\,K$. Further north, differences in east-west temperature gradients become larger, mostly because the simulated trough is positioned further south. As the temperature differences are largest at the ground (Fig. 4.4), the vertical temperature stratification is also affected. Even though the spatial representativeness between the simulations at $4\,km$ grid spacing and the measured vertical profiles is not given, some qualitative assessments can be made. On the warm eastern side of the trough the inversion height is simulated about $300\,m$ too low (P1 in Fig. 4.3). On the cold side (P7 in Fig. 4.3) the inversion height of $250\,m$ is very well simulated. Unfortunately, the atmospheric stability is slightly underestimated. In the height of $600\,m$ the sloped trough crosses the $5°W$ meridian. Above the trough the temperature is simulated too low due to the underestimation of frontal contrasts (P7 in Fig. 4.3). For the same reason the wind is shifting its direction only gradually and at low wind speeds instead of an abrupt shift at the trough with high wind speeds on both sides (see wind vectors at $80.75°N$, $0°E$ and $79°N$, $1°E$ in Figs. 4.2 and 4.8). A visual impression of the result bandwidth caused by the uncertainties in the initial data is given in Fig. 4.9. Differences between the two test cases set up most contrarily (M-B and M-E, see Table 4.2 for setups) are illustrated. With respect to HA, they are the least and best performing test cases (Table 4.5). Most striking is the higher temperature over sea ice resulting from the higher sea-ice temperatures used to initialize case M-B. This results in a further proceeded trough at the western sea-ice edge. While the northerly flow over the sea ice on the western side of the trough is reduced in M-B, the advection of maritime air on the eastern side is enhanced.
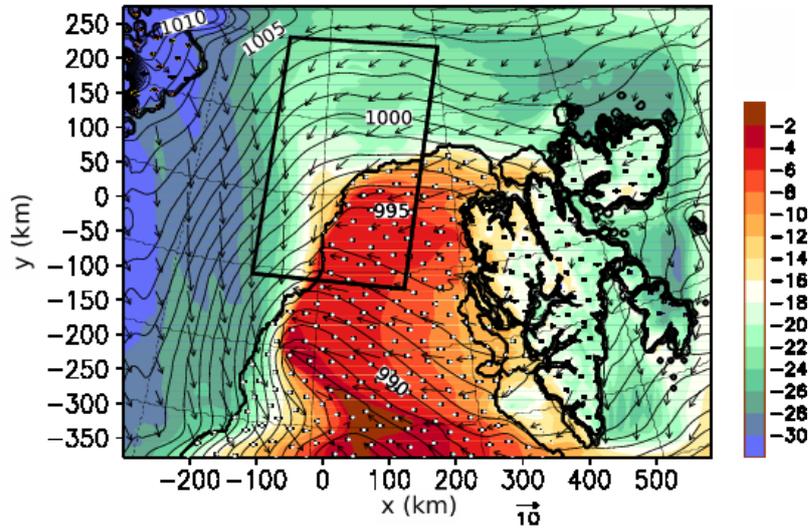
Figure 4.8.: Simulation results of M-E for 7 March 2002 11 UTC. Temperature (°C) at $10\,m$ above ground as indicated by the colour bar, surface pressure (hPa) as isobars, and wind field $(m\,s^{-1})$ at $10\,m$ above ground as vectors. Sea-ice edge from 7 March 2002 18 UTC. Open ocean and land surfaces are identified by white and black dots, respectively. Frame marks the approximate area shown in Fig. 4.2.
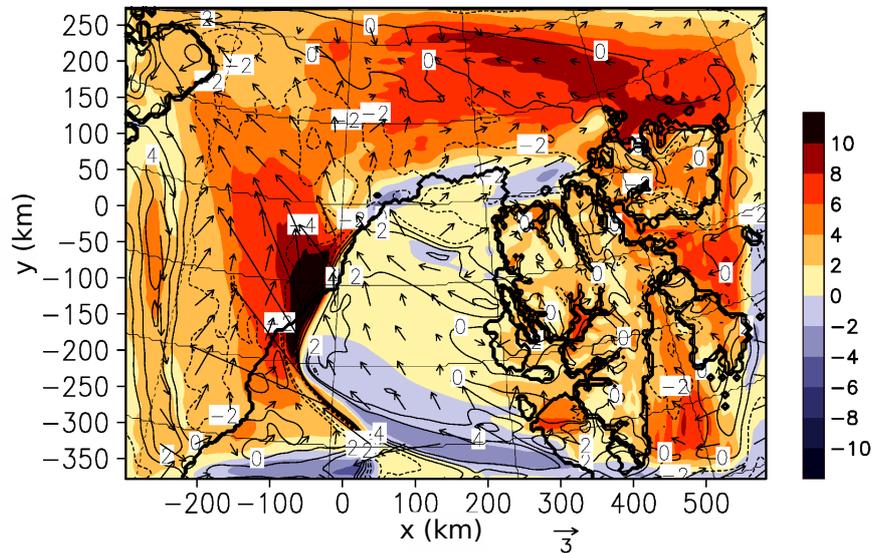
Figure 4.9.: Difference fields of temperature (K) for M-B - M-E at $10\,m$ height above ground for 7 March 2002 11 UTC. The colour bar denotes temperature differences. Vectors and isotaches $(m\,s^{-1})$ denote differences in the wind field. Sea-ice edge from 6 March 2002 14 UTC.

Table 4.5.: Model skills of all test cases for the horizontal legs (below 100 m above ground). N represents the number of samples used for the calculation of ME, RMSE, H, and VWD. HA and GA are the average hit rates and Gandin-Murphy skill scores. For purposes of comparison to section 5.3.3, HA(T,q,ff) and GA(T,q,ff) are only averaged over the respective parameters. For details see text.

N=6663	temperature (°C)			specific humidity (g kg⁻¹)			pressure (hPa)			wind speed (m s⁻¹)				wind direction (°)		average			
	ME	RMSE	H	ME	RMSE	H	ME	RMSE	H	ME	RMSE	VWD	H	RMSE	H	HA	GA	HA(T,q,ff)	GA(T,q,ff)
M-A	4.5	10.3	1	0.2	0.9	4	-5.4	5.7	1	-1.2	3.4	10.8	38	69.7	23	11	0.175	15	0.14
M-B	3.9	9.9	1	0.2	0.9	4	-5.1	5.5	1	-1.9	3.7	9.9	31	65.5	21	9	0.183	12	0.14
M-C	5.3	9.8	1	0.3	0.8	7	-5.5	5.9	0	-1.0	3.4	10.4	36	67.8	27	11	0.217	15	0.20
M-D	4.7	9.3	1	0.3	0.8	7	-5.3	5.6	1	-1.7	3.7	9.5	33	63.3	24	10	0.243	14	0.20
M-E	-0.4	6.8	13	-0.1	0.7	24	-2.8	3.3	23	-0.7	3.1	7.4	34	48.6	24	24	0.307	24	0.39
M-F	2.9	7.7	1	0.1	0.7	19	-3.2	3.6	22	-0.5	3.2	8.0	31	52.0	24	18	0.200	17	0.34
M-G	5.9	10.4	2	0.4	0.9	9	-5.8	6.1	0	-0.2	3.1	11.5	30	72.9	33	10	0.190	14	0.24
M-H	5.9	10.2	2	0.4	0.9	8	-5.7	6.0	0	-1.0	3.4	10.6	37	69.2	28	12	0.209	16	0.20
M-I	-0.7	9.4	2	-0.2	0.9	26	-4.6	5.0	6	-2.1	3.7	8.7	24	58.0	19	14	0.139	17	0.22
M-J	1.6	7.8	1	0.0	0.7	19	-5.0	5.3	2	-1.4	3.5	9.0	36	59.5	20	14	0.177	18	0.27
M-K	0.8	7.4	1	-0.0	0.7	19	-4.8	5.1	4	-2.0	3.8	8.4	30	56.2	21	14	0.188	17	0.27
M-L	0.2	6.9	12	-0.1	0.7	21	-3.0	3.5	23	-0.1	3.3	7.9	23	50.1	21	20	0.342	19	0.38
M-M	2.3	8.2	1	0.1	0.8	18	-5.2	5.5	1	-0.8	3.3	9.8	28	63.4	23	12	0.266	16	0.30
M-N	-1.8	7.8	11	-0.2	0.7	39	-4.8	5.1	4	-1.6	3.6	8.3	33	55.2	19	22	0.214	28	0.32

The influence of the different sea-ice maps is visible in the lower temperatures at the northern sea-ice edge. In that region M-B has higher sea-ice concentrations.

Table 4.5 gives an overview of the near-ground model performance by means of ME, RMSE, H, VWD, HA, and GA of the most important meteorological parameters. Evident is the negative wind velocity bias caused by the wider than observed trough. The negative pressure bias does not influence the simulated flow field.

Table 4.6 shows the overall change of evaluation measures for the experiments defined in Table 4.3. The scores HA and GA agree on the impact of the initialisation characteristics. The forcing is most relevant which indicates that the synoptic situation is dominating the mesoscale structures. The least important initialisation characteristic is the sea-ice thermal conductivity. The determination of the most relevant sea-ice characteristic depends on the choice of the skill score. Even though for the bias-sensitive HA the sea-ice temperature is of primary importance, the pattern-sensitive GA ranks the sea-ice distribution highest. As will be shown in section 4.4.2, the most important feature of the sea-ice distribution is the position of the ice edge but not the distribution of sea-ice concentration.

In comparison to a similar simulation conducted by Dierer et al. (2005) the test cases presented here perform worse. Dierer et al. (2005) used METRAS-MESIM with $7\,km$ horizontal grid spacing and nudged into results of the regional model (REMO). Their results for a meteorological situation with synoptic scale character (18 April 1999) prove a higher HA over the same meteorological parameters of $43\,\%$. However, the situation was dominated by a synoptic scale cyclone without fronts in the simulation domain. In addition a higher resolved forcing data set was used with more time steps to update the boundary values. An interval of 1 hour compared to 6 hours from ECMWF was considered. As shown by Bungert (2008), too large update intervals can deteriorate the model performance. Using the method suggested by Bungert (2008) and applying it to the current case it is found that the frequency of forcing data is too low (Bungert, pers. communication, 2008).

Table 4.6.: Average changes of HA and GA between experiment pairs as listed in Table 4.3. For the roughness experiment averaging is only performed on pairs that use the extreme roughness lengthes.

experiment	averaged delta HA	averaged delta GA
DISTRIBUTION	0.15	0.050
ROUGHNESS	-1.02	0.013
FORCING	7.98	0.063
TEMPERATURE	3.65	0.009
CONDUCTIVITY	0.70	-0.008

4.4.2. Influences on temperature

Two main features characterize the measured time series of temperature along the horizontal flight legs (filled dots, Fig. 4.10): the change at the ice edges (legs 3 and 7) and the temperature drop in the trough (leg 4). The six test cases shown in Fig. 4.10 (M-E, M-N, M-L, M-F, M-H, and M-A) exemplarily demonstrate the impact of the sea-ice characteristics (not filled symbols) and artificially closed sea-ice cover (M-N, crosses) on model results. All test cases miss one of the two main features. A large spread between the test cases of approximately $10\,K$ exists over the sea ice. An offset of the test cases with higher sea-ice temperature is most obvious at legs 5 and 6 in the cold sector over the inner sea ice. These test cases also simulate the temperature rise at the trough and the ice edge of leg 7 too early. A further westerly positioned trough and/or a more westerly positioned sea-ice edge can be made responsible for this too early increase (not shown). On the other hand, the test cases with the lower sea-ice temperature (M-L, M-E, M-N) show a severely strong temperature drop at the ice edge of leg 3. By initializing with the lower sea-ice temperature, the warm sector over the sea ice is cooled too much. As a consequence of the low temperatures over the ice the temperature drop at the trough is underestimated. The test case with artificially
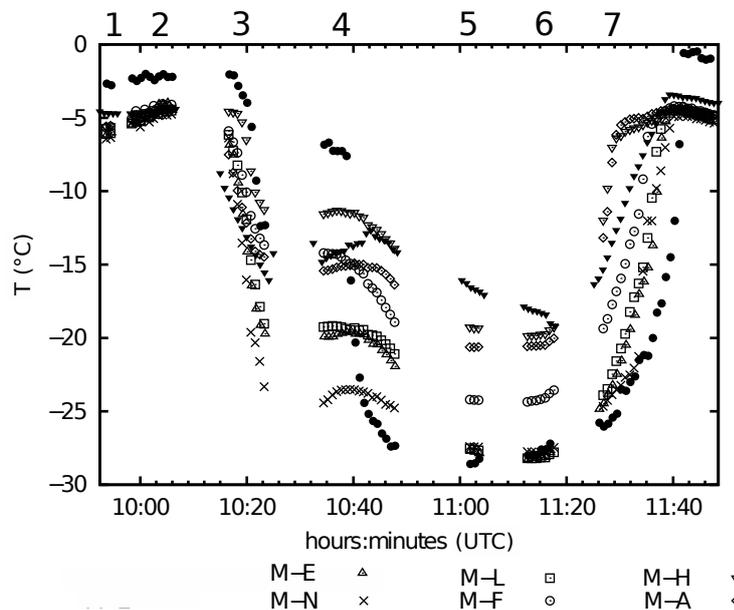
Figure 4.10.: Observed (full dot), from ECMWF reananlyses (down-pointing triangle) and modelled (other symbols) time series of temperature along the horizontal flight legs (numbering given on top) for 7 March 2002. Legs one and two are over the ocean. The sea-ice edge is crossed at 10:22 UTC, the trough at 10:40 UTC. Legs five and six are flown over sea-ice. Both the sea-ice edge and the trough are again crossed at 11:40 UTC. See Fig. 4.4 for locations.

closed sea-ice cover (M-N) shows an even larger temperature gradient perpendicular to the sea-ice edge. The reduced spatial extent of the marginal ice zone due to the artificial covering with sea ice leads to a more abrupt transition from open ocean to closed sea ice. Consequently, the surface sensible heat fluxes change suddenly and are responsible for the faster temperature changes.

All test cases are too cold over the open ocean (beginning and end of time series), indicating that the alleged too warm water temperatures from NAOSIM could be well suited. The forcing data (filled down-pointing triangle) show a good timing and a realistic temperature drop and temperature increase at the ice edges. However, temperature levels are incorrect and the trough is not visible from the temperature time series. The warm bias over the inner sea ice is largest for the forcing data.

The influence of the sea-ice coverage on the temperature is summarized in the DIS-TRIBUTION experiment. The values for the Taylor diagram are calculated for the test cases in comparison to the aircraft data (horizontal legs). The test cases with the sea-ice map from 6 March 2002 14 UTC (M-A, M-B, M-I) result in normalised standard deviations of 1.25 to 1.8 with correlation coefficients of 0.45 to 0.55 (Fig. 4.11).

With the later sea-ice map from 7 March 2002 18 UTC, which is approximately 7 hours later than the measurement time, the correlation coefficients improve to the range of 0.61 to 0.63. The normalised standard deviations are only minorly reduced hinting at a reduction of phase errors. The primary sources for phase errors in the present test case are a misplaced trough or sea-ice edge. The higher correlation co-efficients with test cases simulated with the later sea-ice map give a strong hint that the later sea-ice edge is more adequate. Amplitude errors (visible in normalised standard deviation) that primarily result from a wrong sea-ice cover distribution remain unchanged. The more homogeneous distribution of the later sea-ice concentration (Fig. 4.6) has no homogenizing effect on the standard deviation of the temperature time series. Only the artificially homogenised sea-ice map (M-N, any sea-ice fraction set to 100 %) leads to a reduction of normalised standard deviation. Hence, an influence of the sea-ice concentration exists, but for this situation its influence is much smaller than the influence of the relocation of the sea-ice edge. The fact that the better results, measured in terms of Δc, are achieved with a sea-ice map valid 7 hours later than the evaluation period suggests that the boundary layer structure quickly adopts to the underlying sea-ice conditions.

The model's skill in simulating near-ground temperature is affected strongest by the sea-ice temperature. The results for experiment TEMPERATURE are summarized in Fig. 4.12. The amplitude error of the simulations performed with the sea-ice
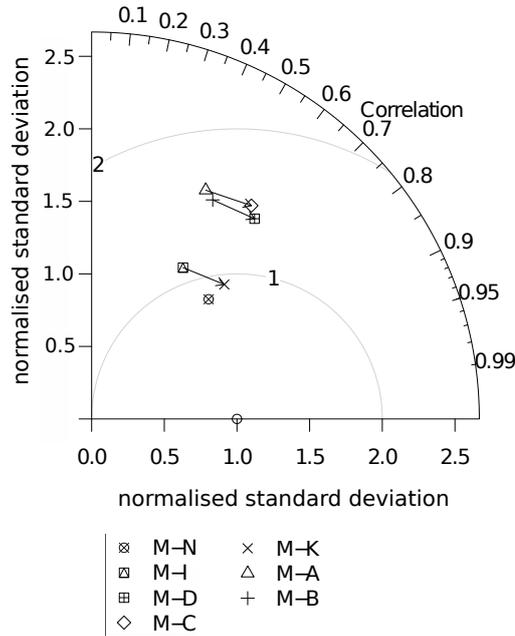
Figure 4.11.: Taylor diagram of temperatures for members of DISTRIBUTION experiment. The normalised standard deviation is given by the radial distance from origin. The correlation is given by the azimuthal angle. The Δc (circle between the measurement at correlation equal one and normalised standard deviation equal one) and the test cases is indicated by grey circles. Arrows point from realizations with the early sea-ice distribution to realizations with the late sea-ice distribution.

temperature of $T_h(ice) - 9°\text{C}$ is reduced by initialising the sea-ice with $-25°\text{C}$. Only the test cases with the lower sea-ice temperature simulate a realistic cooling over sea ice as already visible in Fig. 4.10. Phase errors are nearly the same, since the correlation coefficient is more or less unchanged.

The changes in model performance found for the experiments FORCING, ROUGH-NESS and CONDUCTIVITY are much smaller. The stronger forcing reduces errors in standard deviation in temperature by approximately $\Delta c = 0.3$. The reduction of the surface roughness leads to successively higher errors in standard deviation, though always less than $\Delta c = 0.3$. The amplitude error is enlarged by approximately $\Delta c = 0.2$ for the higher sea-ice thermal conductivity.
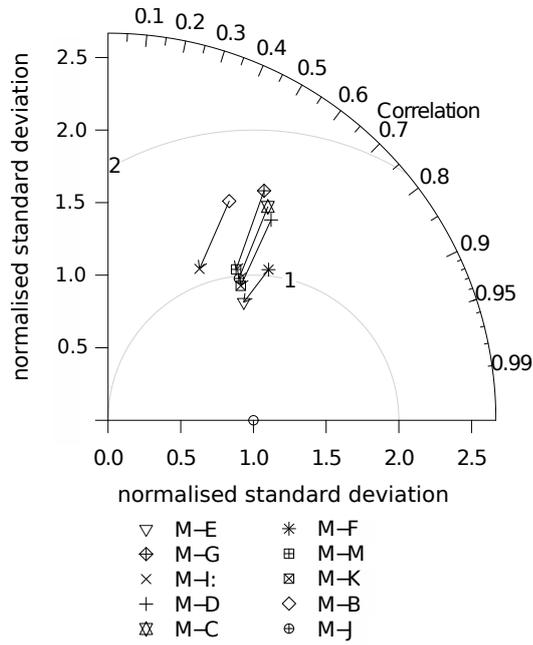
Figure 4.12.: As Fig. 4.11 but for the TEMPERATURE experiment. Arrows point from realizations with sea-ice temperature $= -9°$C to realizations with the sea-ice temperature $= -25°$C.
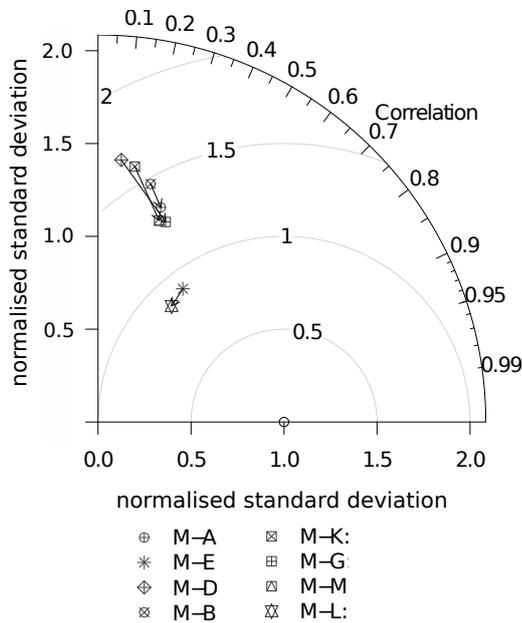
Figure 4.13.: As Fig. 4.11 but of wind speeds for members of ROUGHNESS experiment. Arrows point from realizations with higher roughness to realizations with lower roughness.

4.4.3. Influences on near-ground wind

As already indicated by the ranking of the averaged delta HA and -GA (Table 4.6), the forcing has also a large influence on the performance of near-ground wind speed. The standard deviation of the near-ground wind speed is reduced by approximately $\Delta c = 0.5$ with the stronger forcing. Furthermore, the sea-ice roughness plays an important role as noted in Fig. 4.13. With respect to Δc, correlation and standard deviation, the test cases with smaller roughness give better results. The absolute value of the negative mean bias reduces from $-1.7\,m\,s^{-1}$ to $-2.1\,m\,s^{-1}$ for the test cases with $z_0 = 27\,mm$ (M-B, M-D, M-I, M-K; see Table 4.5) to the range of $-0.5\,m\,s^{-1}$ to $-1.2\,m\,s^{-1}$ for the test cases with $z_0 = 5.9\,mm$ (M-A, M-C, M-E, M-F, M-h and M-J). Further reduction of the negative bias is achieved by applying an even shorter roughness length of $z_0 = 1\,mm$. In the test cases M-L, M-M, M-G the negative bias ranges from $-0.1\,m\,s^{-1}$ to $-0.8\,m\,s^{-1}$. However, H and VWD do not show a substantial improvement, possibly due to a slightly worse positioning of the trough for the test cases with low surface roughness (Fig. 4.14 a)). The dipole structure at $x = -80\,km$, $y = -200\,km$ in Fig. 4.14 a) is caused by a zonal displacement between the troughs of approximately $20\,km$. The enhancement of the wind speed is a consequence of improvements of the flow field over sea ice and the trough structure. When comparing M-G ($z_0 = 1\,mm$) with M-D ($z_0 = 27\,mm$), the off-ice flow west of the trough increases by more than $4\,m\,s^{-1}$ (Fig. 4.14 b)). The increase results in a slightly enhanced cold air advection. On the backside of the trough the on-ice flow is also increased, which results in a much stronger warm air advection. The cross frontal temperature gradient improves by approximately $5\,K$ and is thus better simulated in case M-G. As an eventual consequence of the smaller roughness, the lifting and therefore the vorticity at the trough is enhanced. The trough is found to have deepened by up to $-1.5\,hPa$.
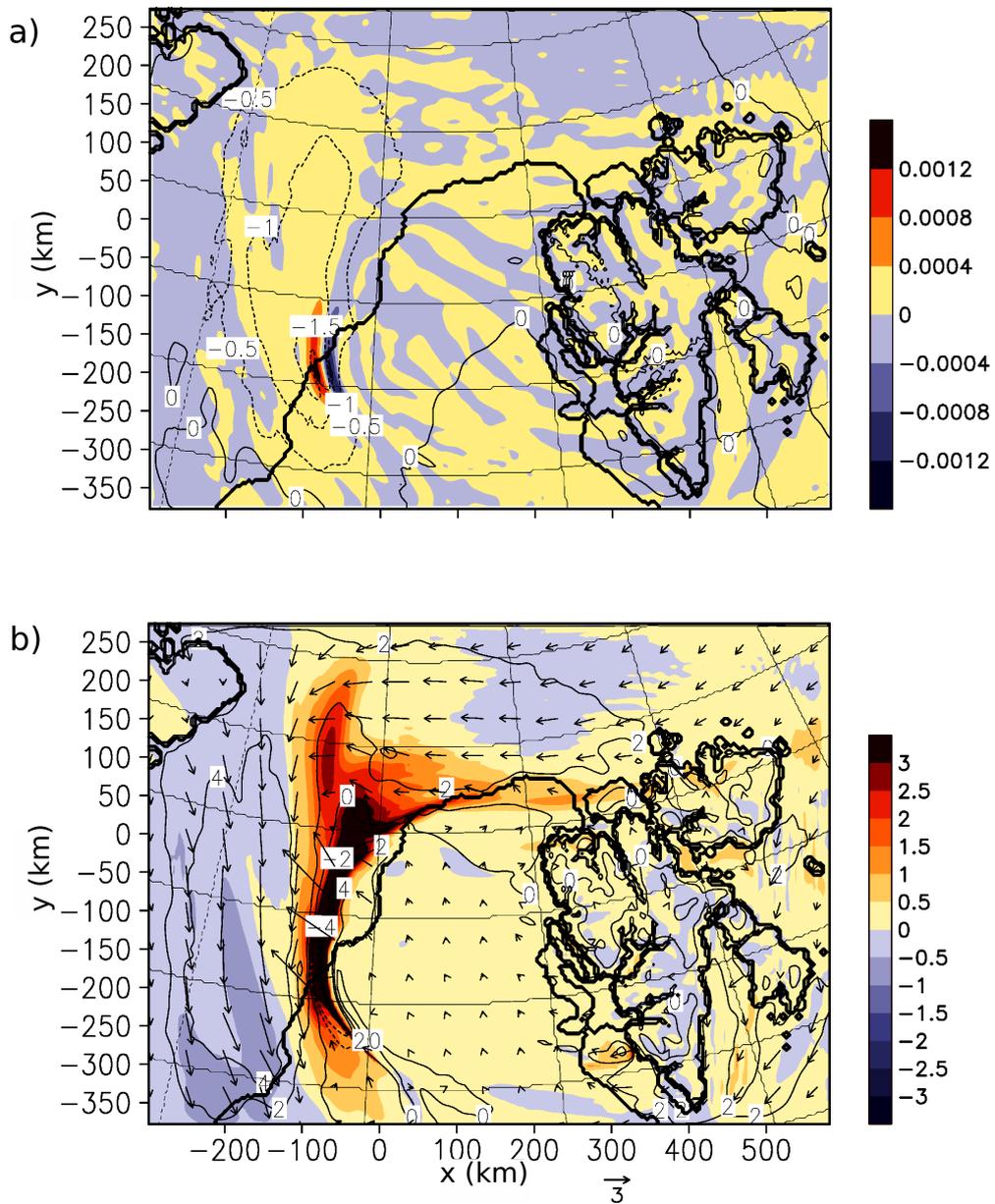
Figure 4.14.: Horizontal cross-section at $10\,m$ above sea level of simulated differences between M-G ($z_0 = 1$ mm) and M-D ($z_0 = 27$ mm) of a): mean sea level pressure (hPa, labelled contours) and vorticity (s^{-1}, shaded), b): wind vector (ms^{-1}, scaling with respect to legend vector), wind speed (ms^{-1}, labelled contours) and temperature (K, shaded).

4.5. Conclusions of Chapter 4

A trough passage through the Fram Strait has been simulated with 14 different model set-ups. The results have been evaluated for 7 March 2002. The test cases were simulated with the standard version of METRAS, using fixed sea-ice distributions during the simulations. For METRAS the sea-ice characteristics roughness, temperature, thermal conductivity and distribution were varied within the range of observational uncertainty. One test case was performed with an artificial sea-ice distribution having all water areas within the sea-ice closed (M-N, any sea-ice concentration set to 100%). In addition to the surface parameters, the strength of the lateral forcing was varied.

The evaluation was based on comparisons with measurements from near-ground aircraft flights. The evaluation was performed on statistical basis employing standard measures. The trough development was discussed for selected test cases using standard meteorological parameters. The hit rates and Gandin-Murphy skill scores, averaged over temperature, pressure, specific humidity and wind speed indicate the strength of the lateral forcing to have the largest influence on the overall performance. However, only little is published on a meaningful selection of the nudging coefficients (Deng and Stauffer, 2006; Bungert, 2008). Typical values range from $0.0001\,s^{-1}$ to $0.001\,s^{-1}$ and the forcing that was chosen might be too strong. On the other hand, all statistical measures except hit rate of wind speed improve with the stronger forcing. This indicates that the investigated meteorological situation is dominated by the synoptic scale.

Under the investigated meteorological situation changes in the sea-ice cover have revealed that for the temperature simulation the location of the sea-ice edge is more important than the homogeneity of the sea-ice concentration. Only the idealised test case M-N shows an atmospherical response to the enlarged homogeneity. M-N performs well, because it keeps the temperature gradient at the sea-ice edge unrealistically large, which happens to coincide with the positioning of large parts of the trough during the measurement. The best simulation results are achieved with low sea-ice surface roughness ($z_0 = 1\,mm$ or $z_0 = 5.9\,mm$), low sea-ice temperature ($T_h(ice) = -25°C$), and a sea-ice cover that is similar to the observed. Therefore the late sea-ice distribution with a discrepancy of only 7 hours to the observation period produces better results than the earlier one at 21 hours prior to the observation period. The wind speed and the cold air advection are highly reduced by the surface roughness of $z_0 = 27\,mm$. Only the low sea-ice temperature of $T_h(ice) = -25°C$ enables the near-ground air over sea ice on the forefront side to cool sufficiently. The increased sea-ice thermal diffusivity slightly worsens model performance in the present case but is of little relevance for the trough development.

The conclusions drawn for the present case are assumed to be transferable only on other atmospherical situations of large scale nature. Thermodynamic influences of the sea-ice distribution may be of greater relevance if the atmospheric situation is less dominated by synoptic scale features. Nevertheless, if synoptic scale features dominate, the position of the sea-ice edge, the sea-ice temperature and the sea-ice roughness should be well known for mesoscale simulations. This means that timely satellite data are very important and the data need to be used and updated in the simulations. If no better information is available, a roughness length of $z_0 = 5.9 \, mm$ or lower and a sea-ice temperature of $T_h(ice) = -25°\mathrm{C}$ are recommended for early spring situations.

There remains to examine whether for less synoptic scale situations e.g. cyclone families, the forcing has a similarly large influence. The too much tilted trough could have been caused by the too low temporal and spatial resolution of the forcing data. Future studies should use more frequent data.

5. Model performance under mesoscale and synoptic scale Arctic conditions

5.1. Introduction

Two questions arise from the conclusions drawn in chapter 4:

1) Can the statement regarding a good setup of sea-ice characteristics be generalised to other than synoptic scale situations?

2) In which way does the mesoscale model performance depend on the synoptic-scale forcing (strong wind, variable wind, weak wind)?

In order to answer these questions some of the model experiments on sensitivity that were performed in chapter 4 are repeated here for two further meteorological situations. The first situation on 13 March 2002 is dominated by two mesoscale cyclones (phase A). The second situation from 15 March 2002 is characterised by a mostly undisturbed, homogeneous off ice flow and is therefore of larger scale nature (phase B). Since both situations are separated by only one day, they are integrated in a single simulation phase without reinitialization.

The impact of the surface characteristics on near-ground processes depends on the large scale meteorological situation. For meteorological situations that are only weakly forced by the synoptic scale, surface characteristics favouring convection and barotropic instability become more important. Therefore, the choice of surface parameters recommended in chapter 4 still may be inappropriate for phase A.

As the link between surface characteristics and near-ground processes the fluxes are of primary importance, they will be investigated for phase B by assessing the simulated surface fluxes with the fluxes measured at aircraft height. As the off-ice flow during phase B is not disturbed by smaller than synoptic scale systems, this phase is an ideal situation to investigate the simulated fluxes that lead to modifications of vertical profiles of temperature, humidity and wind. The measured and simulated dependency of the fluxes on the underlying sea-ice coverage is compared with idealised simulations of Lüpkes and Birnbaum (2005) that include form drag at floe edges.

Mesoscale situations like phase A, where a family of two cyclones passes the Fram Strait on 13 March 2002 are the applications for which mesoscale models can add most value (Lange and Waldl, 2001). The performance improvement compared to the ECMWF is expected to be larger for phase A than for the synoptic scale situation of phase B and of 7 March 2002 (see chapter 4). On the other hand, compared to phase A, the stronger nudging is expected to have an equally strong and positive impact on the synoptic scale situation of phase B, as seen for the synoptic scale situation of 7 March 2002. As a consequence the nudging strength should be chosen dependent on the meteorological situation.

In section 5.2 the meteorological situations of the experiment and the FRAMZY 2002-measurements are described in detail. With the focus on the flux simulations the model setup is briefly repeated in section 5.3. Specifications of the model experimental setup and the evaluation method are also given there. Results are presented in section 5.4. The relation between the scale of the meteorological situation, the nudging and the mesoscale model performance is discussed there. A recommendation for an appropriate nudging strength in dependence on the scale of the meteorological situation is given. Conclusions are drawn in section 5.5.

5.2. Data

5.2.1. Synoptic situation, phase A: mesocyclone at the sea-ice edge on 13 March 2002

In a high pressure area in the Fram Strait with weak near-ground wind conditions the first of three mesocyclones developed at the sea-ice edge at 76° N on 12 March 2002 12 UTC. In the southerly flow on the backside of the ridge at $500\,hPa$ in the North Atlantic Ocean it was steered northeast-wards along the sea-ice edge. The mesocyclone is first detectable in the sea level pressure field of the ECMWF reanalysis for 13 March 2002 00 UTC (Fig. 5.1a). At the time of the first flight mission on 13 March at 8:52 UTC to 11:02 UTC, the mesocyclone has moved to 80.5° N and has crossed the sea-ice edge (Fig. 5.2a). On its eastern side (end of leg 1 and leg 2 in Fig. 5.2a) a weak warm front with a decrease in temperature of $8\,K$ over $90\,km$ (Fig. 5.3a, end of leg 1 and leg 2) and a shift in wind direction from southwest to east-northeast was observed. Due to reduced sea-ice cover, a warm and humid air tongue had developed on the mesocyclone's western side, see leg 4 and leg 5 in Fig. 5.2a and 5.4a) for sea-ice concentration. The reduction of sea-ice concentration is probably caused by the mesocyclone. Fog and haze and 8/8 stratus cloud coverage were reported from almost all horizontal flight legs except legs seven and eight. Strong easterly winds were observed in the northern half of the cyclone, while the wind was very weak

southwest of the cyclone. The second cyclone of the cyclone family was about to
enter into the simulation domain at the southerly boundary but did not affect the
area covered by the measurements.
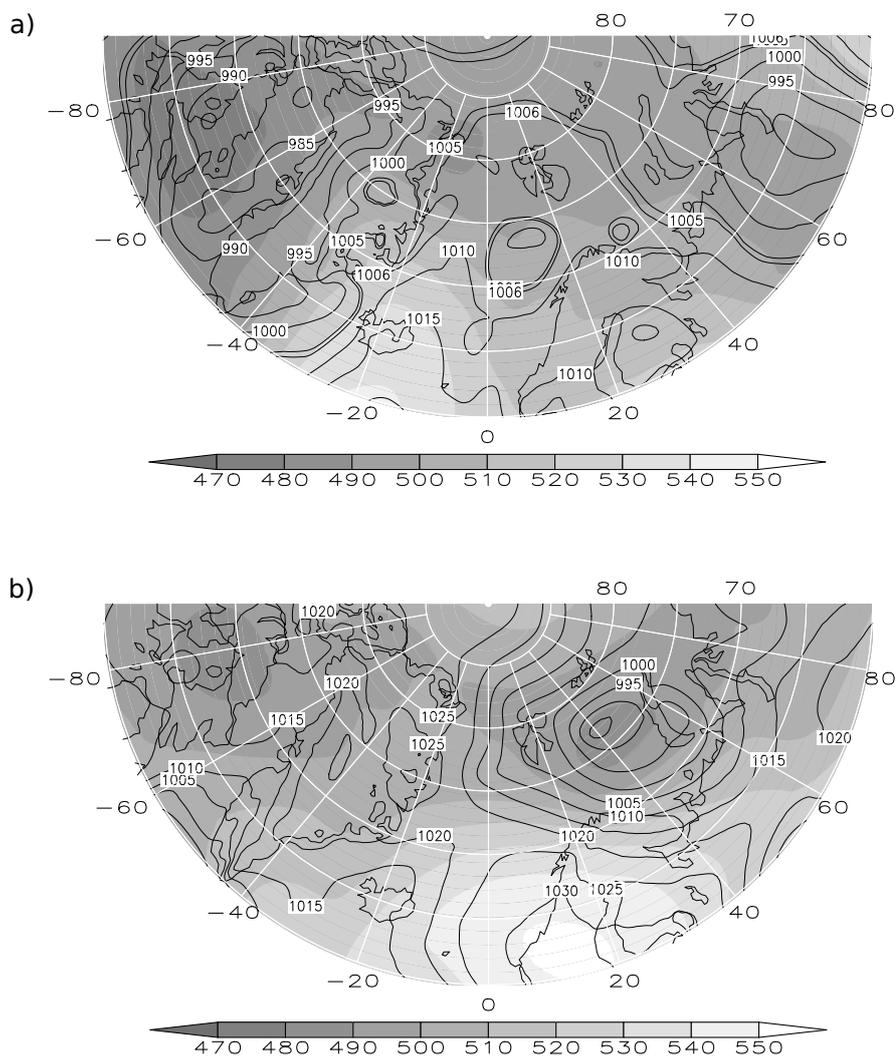
Figure 5.1.: ECMWF reanalyses of $500\,hPa$ geopotential heights [gpdm] (shaded)
and mean sea level pressure [hPa] (contours with $5\,hPa$ increment) for (a) phase A,
13 March 2002 00 UTC (With additional $1006\,hPa$ isobar; the mesocyclone is located
at 0° longitude and 80° N.) and (b) phase B, 15 March 2002 12 UTC.
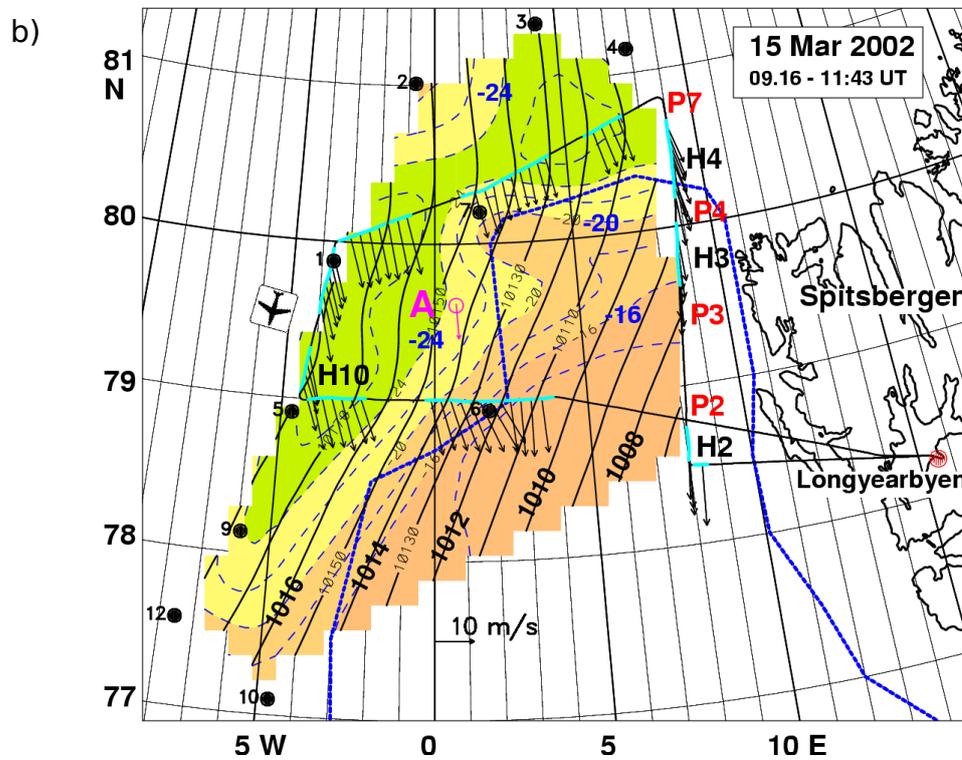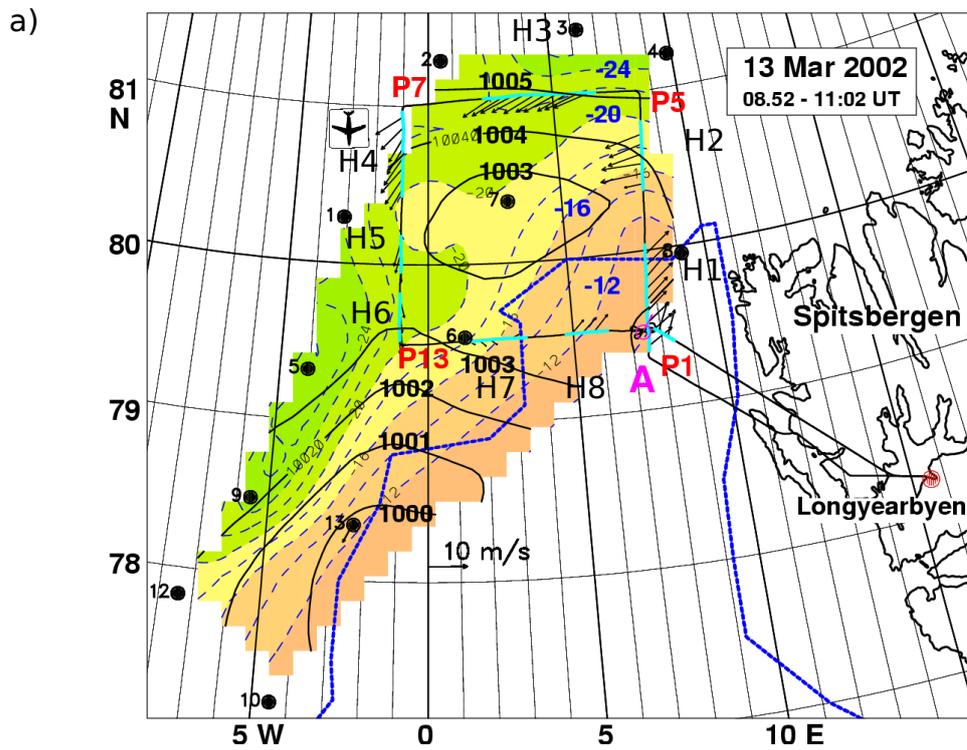
Figure 5.2.: As Fig. 4.2 but for 13 March 2002 around 10 UTC (a)) and 15 March 2002 around 11 UTC (b). From Brümmer et al. (2005).
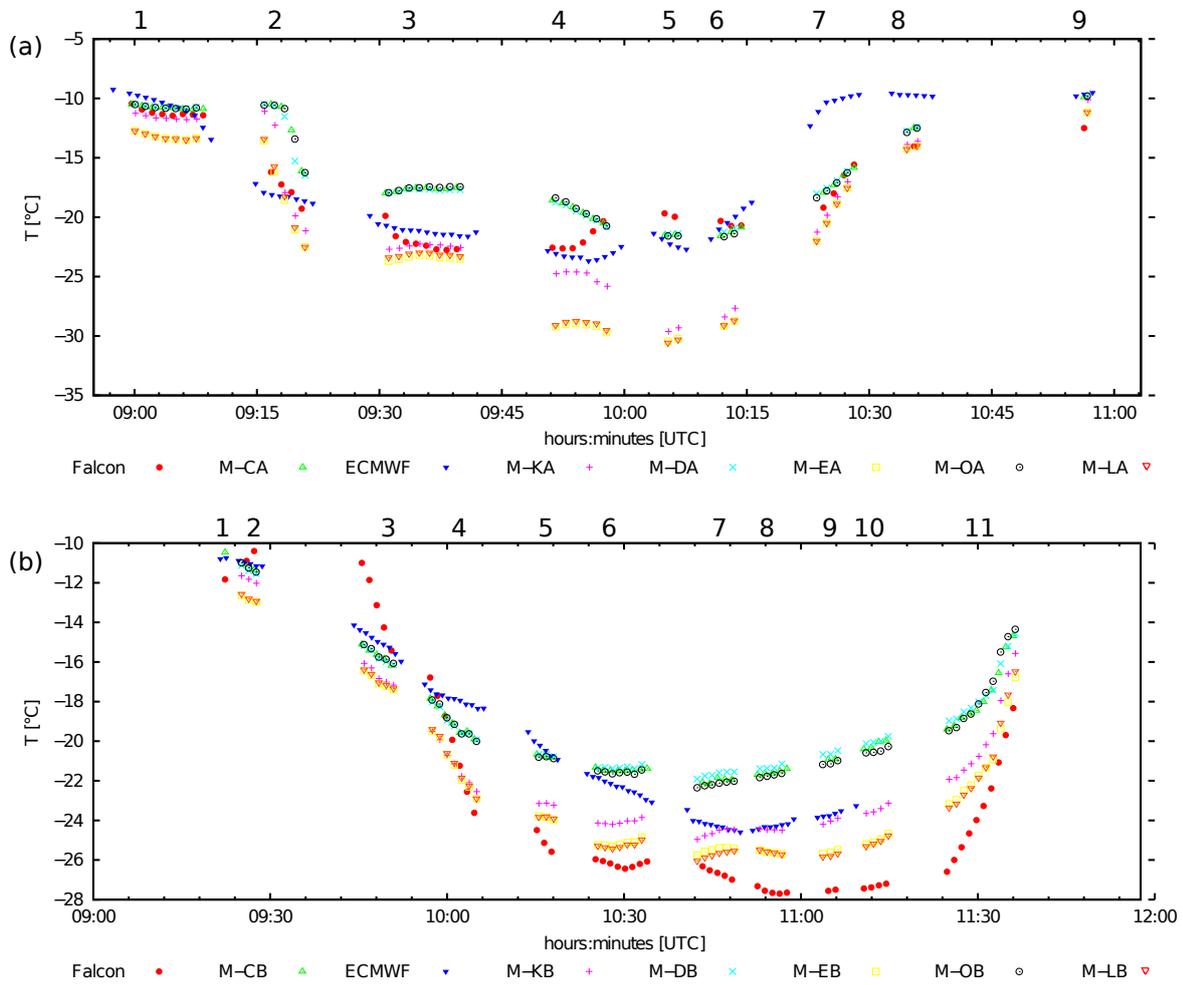
Figure 5.3.: Observed and modelled time series of temperature along the horizontal flight legs. The leg numbers are given above the figures. (a) shows 13 March 2002: Leg one is mainly over the ocean and crosses the sea-ice edge in the end. The warm front is crossed until the end of leg two. Leg three is flown in the cold sector over sea-ice. Legs four and five touch a warm air tongue. Legs six and seven are dominated by the approximation to the ice edge. See Fig. 5.2a for locations. (b) shows 15 March 2002: The temperature is mainly governed by the distance to the sea-ice edge. See Fig. 5.2b for locations.
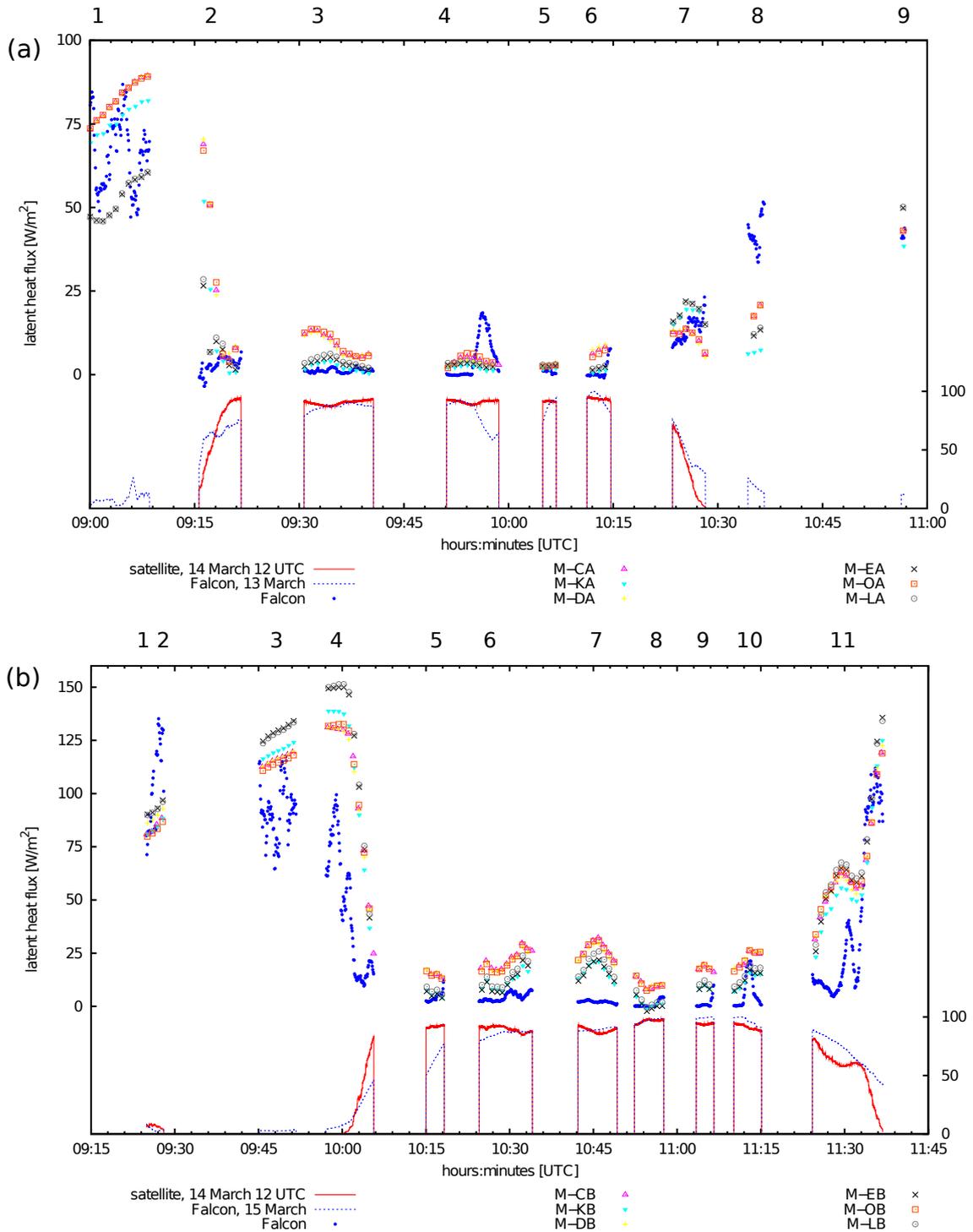
Figure 5.4.: Observed and modelled time series of latent heat flux along the horizontal flight legs for phase A (a) and for phase B (b). The leg numbers are given above the figures. To show the spread of the measured fluxes they are not smoothed with a running mean but every 10th value (for the simulation every 110th) of the fluxes computed at 2 Hz is plotted. Sea-ice coverage as derived from satellite measurements and as derived from surface temperature measurements by aircraft with lines.

5.2.2. Synoptic situation, phase B: off-ice flow on 15 March 2002

With the weakening of the upper air ridge over the North Atlantic Ocean, a high reaching cyclone developed in the Barents Sea (Fig. 5.1b) that led to strong northerly winds in the Fram Strait on 15 March 2002 12 UTC (Fig. 5.2b). Taking the surface pressure gradient of the ECMWF simulation as a proxy for the dominating scale of the meteorological situation it can be said that the meteorological situation in the Fram Strait is much more dominated by the large scale during phase B than during phase A. Cloudless but hazy conditions prevailed over ice. In the off-ice flow over open water cloud streets and extremely intense sea smoke developed (Brümmer et al., 2005). The measurements of potential temperature, specific humidity and wind from profiles P7, P4, P3 and P2 at the eastern side of the flight pattern show the transition from an Arctic boundary layer to a maritime boundary layer (Fig. 5.5). The temperature below the inversion was warming by $16\,K$, while the inversion height rose by $440\,m$ and weakened (Fig. 5.5a). Near ground specific humidity rose from $0.3\,g/kg$ (P7 in Fig. 5.5b), which was even less than over the inversion, to $1.3\,g/kg$ over the open ocean (P2). The inversion was strong enough to develop a low level jet over the sea ice (P7 in Figs. 5.5c and d). The jet immediately dissipated over the marginal ice zone to a, compared to the speed above, moderately increased speed inside the boundary layer. Given sufficient distance from the comparatively rough marginal ice zone the wind speed profile over the open ocean was enhanced (P2). The mean wind direction in the boundary layer veered by 20°from P7 to P2.

In comparison to the sea-ice conditions observed on 14 March 2002 at 12 UTC the sea-ice edge east of 5° E has shifted approximately $50\,km$ northward (Fig. 5.2). Compared to sea-ice conditions from 13 March 2002 many new ridges with approximately 3-4 m height were reported.

5.2.3. Mesoscale data

The same measurement devices as used for the evaluation of the simulation of the trough passage on 7 March 2002 (section 4.2.2) have been used (FRAMZY 2002; Brümmer et al. (2005)). Two flights with patterns comparable to that of 7 March 2002 (Fig. 4.4) were performed; the first on 13 March 2002 from 8:52 UTC to 11:02 UTC measuring the first cyclone of a cyclone family moving through the Fram Strait on this and the following day. The second flight took place on 15 March 2002 from 9:16 UTC to 11:43 UTC in off-ice flow.

In addition to the meteorological quantities considered earlier (section 4.2.2), also the turbulent fluxes of momentum (τ), sensible heat (H) and latent heat (E) are computed. With the eddy correlation method mean fluxes over $80\,s$ are calculated for
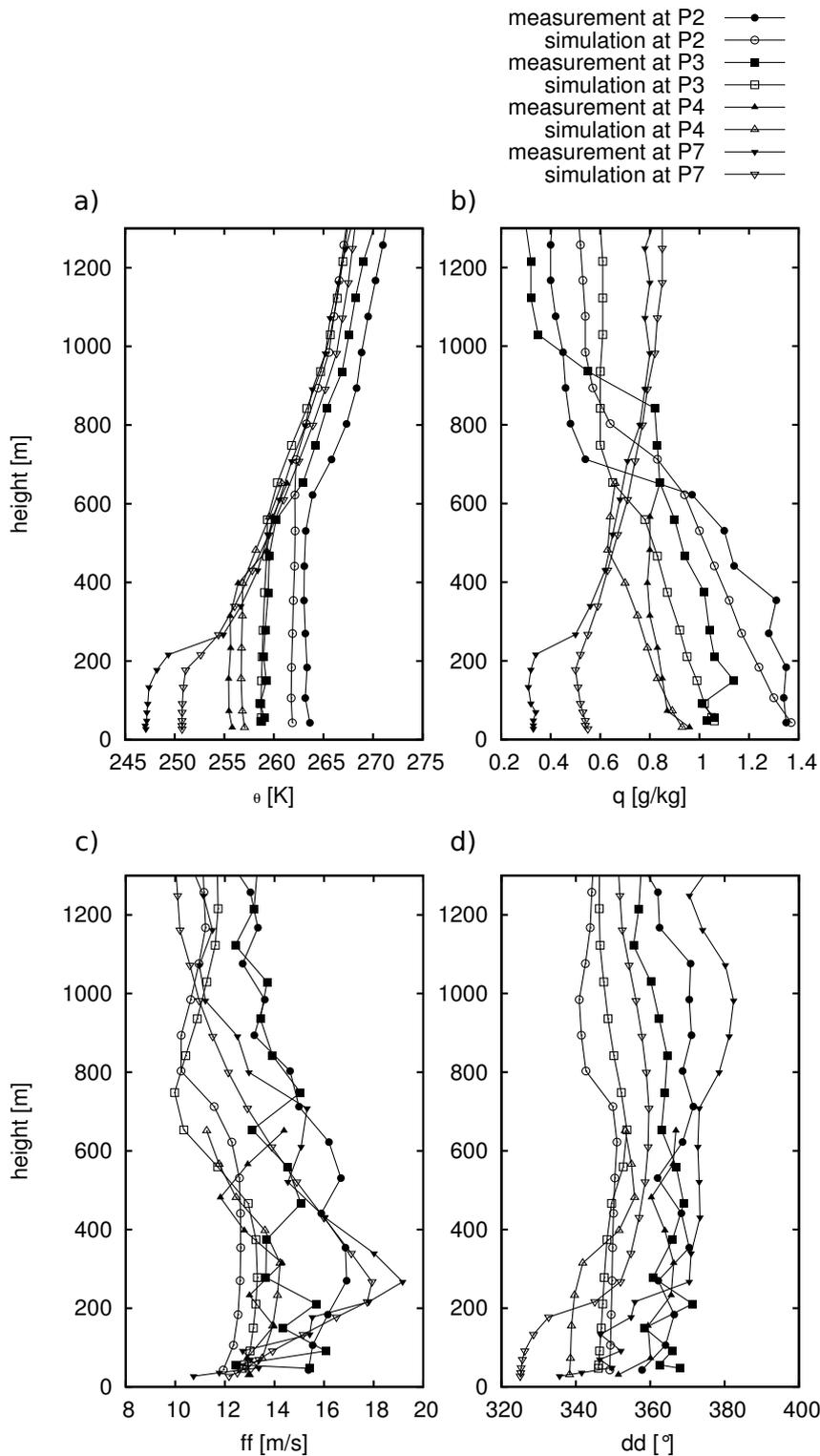
Figure 5.5.: Observed and modelled vertical profiles of a) potential temperature θ, (b) specific humidity q, (c) wind speed ff, and (d) wind direction dd at locations P2, P3, P4 and P7 (see Fig 5.2a for locations) on 15 March 2002 between 9:16 UTC and 11:43 UTC.

78

every $0.5\,s$ from the instantaneous $100\,Hz$ measurements of wind velocity, temperature and specific humidity (Eqs. 5.1, 5.2 and 5.3). To establish comparable spatial representativeness between the simulations at 4 km grid spacing and the calculated fluxes, a 40 seconds running mean is applied to the time series of measured fluxes. As mentioned in section 5.3.1 the constants for the c_p and ρ used for the calculation of the fluxes differ between measurements and simulations. Correction factors of $5\,\%$ are applied.

$$\tau = -\rho\,\overline{u'w'} \tag{5.1}$$

$$H = -\rho c_p\,\overline{\theta'w'} \tag{5.2}$$

$$E = -\rho l_v\,\overline{q'w'} \tag{5.3}$$

$\overline{u'w'}$, $\overline{\theta'w'}$ and $\overline{q'w'}$ are the covariances of the turbulent fluctuations of horizontal wind speed (u') and vertical wind speed (w'), of potential temperature (θ') and w', and of specific humidity (q') and w'. The turbulent fluctuations are computed as differences of the instantaneous values to the linearly detrended mean value, averaged over $80\,s$. The air density is set to $\rho = 1.31\,kg/m^3$, the specific heat at constant pressure to $c_p = 1005\,J(kgK)^{-1}$ and the specific heat of vaporization $l_v = 2500\,Jkg^{-1}$.

In section 5.4.2 atmospheric fluxes obtained from measurements are compared with simulated surface fluxes. As a prerequisite the existence of a constant flux layer up to the measurement height is required. This requirement is only fulfilled over homogeneous terrain and in homogeneous atmospheric conditions. Otherwise a mean vertical velocity may occur, which leads to an additional advective flux in vertical direction. In case of an advective flux, the turbulent atmospheric flux would not equal the flux through the lower boundary of the atmosphere any more. While the measurement height is for the most flight legs below the inversion height and also the large scale atmospheric conditions of phase B can be considered homogeneous, the sea ice, especially the marginal sea-ice zone, cannot be considered homogeneous. Several approaches have been made to transform the coordinate system in a way that causes to vanish a mean vertical velocity (Griessbaum and Schmidt, 2009; Wu et al., 2005) for heterogeneous terrain. For this comparison, however, simply a long averaging interval of 80 s (equivalent to $8\,km$) is applied and it is relied on its homogenizing effect. Areas of stable stratification that occur intermittently are disregarded.

5.3. Model Experiment

5.3.1. Model setup

The model setup and nudging is identical to that described in section 4.3. However, simulations performed within this experiment are limited to a single thermal conductivity (that of snow, see Table 5.1) and to a single sea-ice distribution (14 March 2002 12 UTC). For the location of the sea-ice edge see Fig. 5.6.

Water temperatures simulated by NAOSIM for $5\,m$ depth for 13 March 2002, 00 UTC, are used as surface values (for details see section 4.3.1).

Surface subgrid-scale characteristics are considered and the turbulent surface fluxes of momentum (τ_s), sensible heat (H_s) and latent heat (E_s) are calculated from the flux-averaged (section 4.3) surface-layer scales for velocity (u_*), temperature (T_*) and humidity (q_*) (Eqs. 5.4, 5.5 and 5.6).

$$\tau_s = \rho_{0,surf} \left(-u_*\right) u_* \tag{5.4}$$

$$H_s = \rho_{0,surf}\, c_p\, u_*\, \Theta_* \tag{5.5}$$

$$E_s = \rho_{0,surf}\, l_v\, u_*\, q_* \tag{5.6}$$

As the thermodynamic surface-layer scales depend on respective quantities derived from the surface energy budget equation (Eq. 4.1) and the friction velocity depends on the roughness, they are the parameters by which information about surface characteristics is passed on to the atmosphere. The simulated large scale density of humid air at the surface $\rho_{0,surf} = 1.38\,kg/m^3$ and specific heat of dry air $c_p = 1006\,J(kgK)^{-1}$ deviate from the values used for calculating measured fluxes. The simulation time is extended by 12 hours to 56 hours and the simulations are started for 12 March

Table 5.1.: Surface model parameters for sea ice.

parameter	value	source
z_0	5.9 mm (strongly deformed sea ice[a]), 27 mm (very rough multi year sea ice[b])	[a]: Lüpkes and Birnbaum (2005) [b]: Guest and Davidson (1991)
$T_h(ice)$	-9 °C ([a]), -25°C ([b])	[a]: Putkonen (1998), [b]: Brümmer et al. (2005)
ν_s	$1 Jm^{-1}s^{-1}K^{-1}$ (snow)	Pielke (2002)
k_s	$0.8 \times 10^{-6} m^2 s^{-1}$	Pielke (2002)
albedo	72 % (pure white snow)	Pielke (2002)
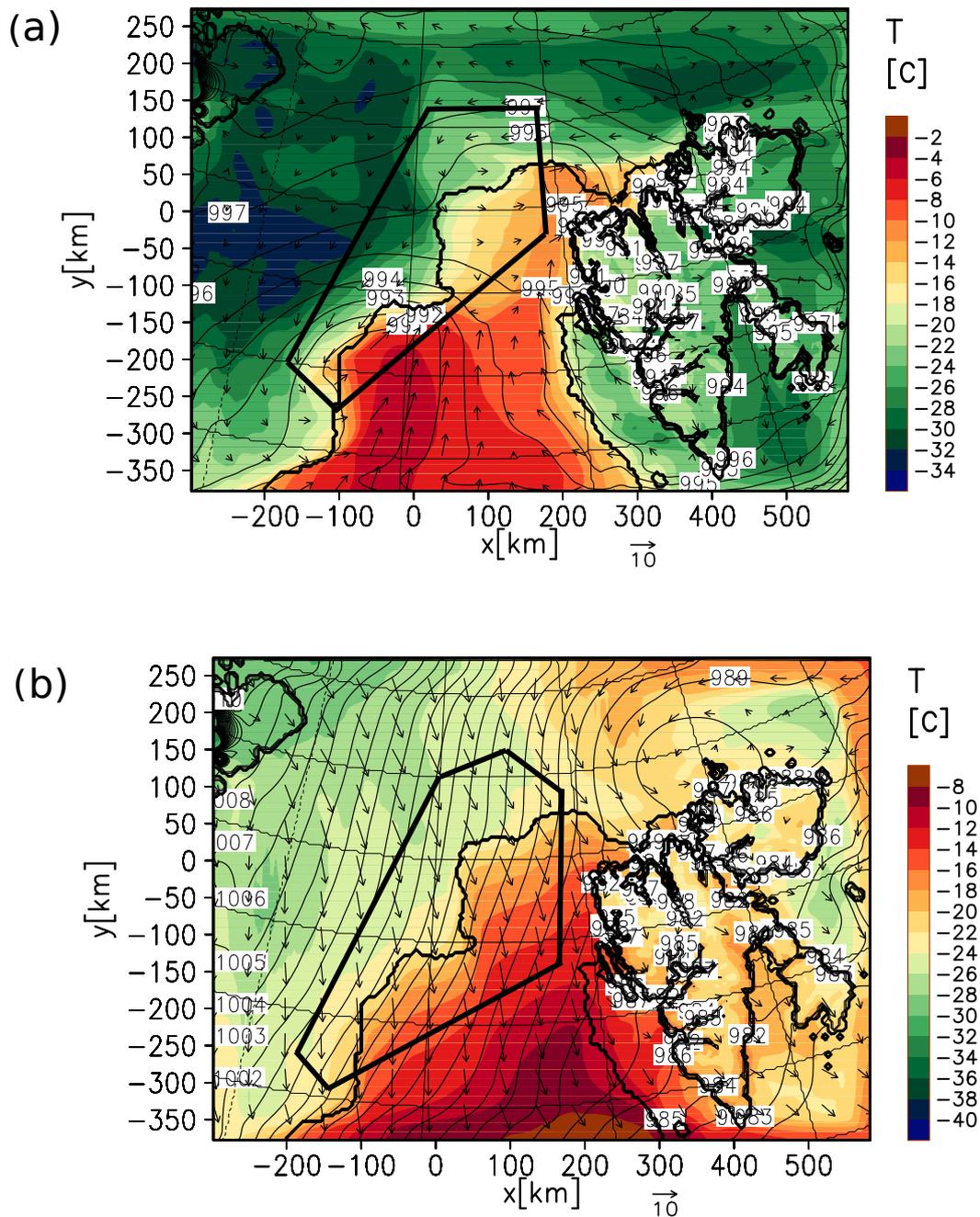
Figure 5.6.: Temperature [°C] at $10\,m$ above ground as indicated by the colour bars, surface pressure [hPa] as isobars, and wind field [$m\,s^{-1}$] at $10\,m$ above ground as vectors. Sea-ice edge from 14 March 2002 12 UTC. Frames mark the approximate areas shown in Fig. 5.2. (a) shows simulation results of M-LA for 13 March 2002 10 UTC, (b) of M-LB for 15 March 2002 10 UTC.

2002, 18 UTC, so that each simulation encompasses both synoptic situations (phases A and B). The extended simulation time inevitably leads to a reduction of model performance for the later phase B with off ice flow. The error growth depends on many factors e.g. the synoptic- and the geographic situation, and on the nudging strength and model accuracy of the driving model so that an universal estimation is not possible.

5.3.2. Experimental setup

All simulations are performed for the period of 12 March 2002 18 UTC to 15 March 2002 12 UTC. The model sensitivity on the intensity of lateral nudging and two sea-ice characteristics, namely sea-ice roughness and sea-ice temperature, is investigated. The characteristics are varied by changing the respective parameters to any of the values given in Table 5.1. All other sea-ice characteristics that influence the surface energy budget are specified there, too. The nudging is varied by applying two different magnitudes of lateral forcing (section 4.3.1). From all possible combinations of variations six are selected to be simulated. With respect to the meteorological situations during the simulation time, each of the simulations is subdivided into the phases A and B resulting in 12 test cases, named "M-CA", "M-DA", "M-EA", "M-KA", "M-LA", "M-OB", "M-CB", "M-DB", "M-EB", "M-KB", "M-LB" and "M-OB" with the last character of the test case name indicating the simulation phase (Table 5.2). The middle character corresponds to naming of test cases in subsection 4.3.2.

The sensitivity study is organised in three model experiments named "ROUGHNESS", "FORCING*" and "TEMPERATURE". Each test case is grouped into these experiments if a complementary test case exists whose setup only differs in the parameter the respective experiment is dealing with. Only for the FORCING* experiment differences in roughness length are allowed. As will be shown in section 5.4.1, the influence of the roughness length is low. The grouping leads to pairs of complementary test cases which are specified in Table 5.3.

5.3.3. Evaluation Method

The evaluation method described in section 4.3.3 is applied to the standard parameters. As average Gandin-Murphy Skill Score only the individual scores of temperature, specific humidity and wind speed will subsequently be used (GA(T,q,ff)). The score of pressure has been omitted because it has become apparent that G(pressure) randomly shows perfect scores. Due to the strong correlation of pressure and height, even small changes in flight altitude lead to a strongly imprinted but artificial pattern in the time series of pressure on horizontal legs. As on the one hand the correct

Table 5.2.: Setup of all test cases. The last character of the test case name indicates the simulation phase.

test case	characteristic		
	forcing	roughness	temperature
	weak: $a_f = 0.4$, $\delta_0 = 0.001 s^{-1}$ strong: $a_f = 0.2, \delta_0 = 0.01 s^{-1}$	$z_0[mm]$	$T_h(ice)[°C]$
M-CA, M-CB	weak	5.9	-9
M-DA, M-DB	weak	27	-9
M-EA, M-EB	strong	5.9	-25
M-KA, M-KB	weak	27	-25
M-LA, M-LB	strong	1	-25
M-OA, M-OB	weak	1	-9

Table 5.3.: Pairs of test cases grouped into experiments, that are built to determine the model sensitivity.

experiment	ROUGHNESS	TEMPERATURE	FORCING*
	M-OA/M-DA, M-OB/M-DB	M-DA/M-KA, M-DB/M-KB	M-KA/M-EA, M-KB/M-EB
	M-CA/M-DA, M-CB/M-DB		
	M-LA/M-EA, M-LB/M-EB		

reduction of pressure with height is easily simulated by the model and on the other hand deviations on the mean flight level are rare, G(pressure) erroneously assumes that the model perfectly simulates exceptional events and the simulation of pressure is assessed very positively. This problem only occurs if different flight legs are intercompared. Therefore, for the inter-comparison of simulations for phases A and B with the simulations for 7 March 2002, two additional columns are introduced in Table 4.5 that specify GA(T,q,ff) and HA(T,q,ff). For reasons of consistency the average hit rate subsequently will also just be averaged over the individual scores of temperature, specific humidity and wind speed HA(T,q,ff). Since two thermodynamic quantities and one dynamical quantity remain in the averaged scores they are still assumed to represent the performance of the whole simulated meteorology.

In addition to the standard parameters also the turbulent fluxes of momentum, sensible heat and latent heat are investigated. As the aircraft flew on its horizontal legs always below the inversion, the fluxes calculated from measurements in the height of the aircraft can be compared to the simulated surface fluxes. For that purpose, the simulated surface fluxes valid along the aircraft track are extracted on-line. During integration the simulated surface fluxes of the four surrounding grid points are linearly interpolated to the aircraft position.

Due to the isolating effect of the sea ice, heat fluxes are expected to be lower over sea-ice surfaces than over open ocean. Indeed, for both phases A and B a strong anti-correlation between the satellite-measured sea-ice coverage and the latent heat flux measured by the aircraft Falcon exist (phase A:-0.85, phase B: -0.89). For the two-dimensional sea-ice concentration satellite measurements are used, but the sea-ice concentration can also be estimated from the infra-red radiation thermometer KT-19 (Gerbush et al., 2008), that was mounted on the aircraft and pointing downward. With a field-of-view of 2° the sampling surface on the horizontal flight legs was approximately $0.2\,m^2$. The time series of surface temperature, averaged with a 40 s running mean, is scaled so that the minimal observed temperature corresponds to $100\,\%$ ice concentration. The ice concentration thus obtained (Fig. 5.4, dashed line) has a correlation coefficient of 0.95 to the sea-ice concentration obtained from satellite (Section 4.3.1). The anti-correlation between the surface-temperature-derived sea-ice coverage and the latent heat flux is 0,04 and 0.06 higher than that with the satellite-measured sea-ice coverage for phases A and B, respectively. The reduced discrepancy in time between the flux measurement and the ice detection can be named as reason for the higher correlation. For the investigation of relations between measured fluxes and surface cover the surface-temperature-derived sea-ice coverage will be used.

5.4. Simulation results

5.4.1. Typical result features

For a qualitative comparison the surface pressure and near-ground fields of wind and temperature of test case M-L are given (compare Fig. 5.6 with Fig. 5.2). This test case performs best on average over all three situations of phase A, phase B and the trough passage of 7 March 2002 (see Table 5.4 and section 4). Table 5.4 gives an overview on the near-ground model performance by means of ME, RMSE, H, VWD, HA(T,q,ff), and GA(T,q,ff) of the most important meteorological parameters. For M-LA and M-LB average hit rates of HA(T,q,ff)=29% and HA(T,q,ff)=54%, respectively are reached. On average over HA(T,q,ff) and GA(T,q,ff), however, simulation results for phase A are similar good as for phase B and even better than the simulation results presented in chapter 4.

With respect to the surface pressure field the positioning and the strength of the trough are well simulated for phase A (compare Fig. 5.6a with Fig. 5.2a). The mesocyclone is separated from the trough by less than 1 hPa pressure increase. No closed isobar appears in the simulated surface pressure field. However, with the observed pressure increase less than the discretization interval for the isobars, it remains unclear if a similar pressure increase is simulated. Also the temperature gradient as well as the warm core west of the low pressure centre are well reproduced by the model. Wind speeds appear to be simulated lower than observed around the mesocyclone.

The dominating east-west pressure gradient of phase B is well simulated in strength and direction (compare Fig. 5.6b with Fig. 5.2b). The observed pressure increases of $9\,hPa$ and $11\,hPa$ from the eastern to the western edges of the measured field at $80°N$ and $79°N$, respectively, are simulated very similar. However, the isobars derived from observations veer against the north direction by an angle of approximately $20°$, whereas the simulation shows only $10°$. The west-east temperature increase measured at $79°N$ of $13\,K$ is well simulated ($12\,K$), whereas the temperature reduction from south to north along the eastern flight legs is underestimated by approximately $5\,K$ (Fig. 5.5a). The larger portion of the underestimation originates from the temperature overestimation over the sea ice. As upper air temperatures are in better agreement, the inversion strength is underestimated over ice by maximal $4\,K$. Inversion heights are in good agreement. The profiles of specific humidity over the open ocean are simulated slightly drier than observed within the boundary layer and slightly more humid above. The observed low level jet over the sea ice is simulated nearly as pronounced (P7 in Figs. 5.5c and d). Also the dissipation of the jet over the marginal ice zone is well reproduced. Only the re-development of the voluminous wind speed profile over the open ocean is underestimated by the simulation

by approximately $4\,m\,s^{-1}$. The veering of the mean wind in the boundary layer by $20°$ from P7 to P2 is well represented.

Table 5.5 shows the overall change of the evaluation measures HA(T,q,ff) and GA(T,q,ff) for the experiments defined in Table 5.3. The scores agree on the impact of the initialization characteristics. Least important is the roughness. It is thanks to this relative unimportance that the FORCING*-experiment with varying roughnesses can be justified. The sea-ice temperature has medium strong influence on the results. The nudging is most relevant with a positive impact of the stronger nudging for phase B.

With the corresponding meteorological situation dominated by the larger scale (section 5.2.2), the stronger nudged mesoscale simulation gives an added value of 19 percentage points compared to the model performance of the driving model ($\Delta\,HA(T, q, ff)$ of MB-E in Table 5.6). The added value is only small ($\Delta\,HA(T, q, ff)$ of MB-K = 2 percentage points) if the mesoscale model is weakly nudged. On the other hand a negative impact of the stronger nudging exists for phase A where mesoscale structures are more pronounced (M-EA - M-KA in Table 5.5). Whereas weak nudging leads to an added value of the mesoscale model simulation of 10 percentage points ($\Delta\,HA(T, q, ff)$ of MA-K in Table 5.6), strong nudging under mesoscale dominated meteorological conditions even reduces the mesoscale model performance compared to the model performance of the driving model. These findings are supported also by comparison of results for 7 March 2002 (Table 4.5, last two columns in section 4.4.1). For the corresponding large scale dominated meteorological situation also better results are achieved with the stronger nudging (HA(T,q,ff)(M-L) - HA(T,q,ff)(M-M) = 3 percentage points). With respect to the performance of the driving model an added value of 12 percentage points and 9 percentage points for the strong and the weak nudging, respectively, are achieved ($\Delta\,HA(T, q, ff)$ of M-L and of M-M in Table 5.6). Except for phase A, relations between GA(T,q,ff) allow the same conclusions as drawn for HA(T,q,ff)

5.4.2. Influences on fluxes

Due to the more homogeneous flow field, the principles leading to modulations of the fluxes with location can be named with greater certainty for phase B. Caused by the long fetch over the sea-ice that has insulated the air against the comparatively warm water, a continental air mass has been formed by cooling and drying. With the turbulent latent and sensible heat fluxes as main mechanisms, an air mass conversion from a continental Arctic air mass to a maritime air mass takes place in the off ice flow over the water. The strength of the turbulent heat fluxes primarily depends on the air-surface temperature- and humidity differences. They are maximal at or a few kilometres downstream the sea-ice edge (legs 4 in Figs. 5.7b and 5.4b), where on the

Table 5.4.: Model skills of all test cases for the horizontal legs (below 100 m above ground) of both flights A (13 March 2002) and B (15 March 2002). N(A) and N(B) represent the numbers of samples used for the calculation of ME, RMSE, H, G and VWD. HA(T,q,ff) and GA(T,q,ff) are the average hit rates and Gandin-Murphy skill scores of temperature, specific humidity and wind speed. For details see text.

N(A)=5465 N(B)=7761	temperature [°C]				specific humidity [g kg⁻¹]				pressure [hPa]		wind speed [m s⁻¹]					wind direction [°]		average	
	ME	RMSE	H	G	ME	RMSE	H	G	ME	RMSE	ME	RMSE	VWD	H	G	RMSE	H	HA(T,q,ff)	GA(T,q,ff)
M-CA	2.1	3.1	21	0.65	0.1	0.2	51	0.31	-8.1	8.1	-2.9	3.8	6.3	32	0.38	71.0	40	35	0.45
M-DA	2.1	3.0	21	0.69	0.1	0.2	54	0.31	-8.1	8.1	-3.2	4.1	6.3	28	0.34	74.3	42	34	0.45
M-EA	-3.3	4.7	5	0.68	0.0	0.2	61	0.63	-6.6	6.7	-3.7	4.3	5.3	10	0.22	47.6	41	25	0.51
M-KA	-1.7	3.4	34	0.82	0.0	0.2	81	0.44	-7.8	7.8	-3.4	4.2	6.0	22	0.28	62.3	54	46	0.51
M-LA	-3.2	4.6	11	0.68	0.0	0.2	62	0.61	-6.7	6.7	-3.5	4.0	5.1	15	0.27	46.9	41	29	0.52
M-OA	2.1	3.2	19	0.62	0.1	0.2	49	0.28	-8.1	8.2	-2.4	3.4	6.5	44	0.43	70.1	36	37	0.44
M-CB	3.8	4.8	6	0.47	0.3	0.3	21	0.77	-17.0	17.1	-0.3	2.3	3.4	66	0.06	19.7	96	31	0.43
M-DB	3.9	5.0	6	0.45	0.2	0.3	23	0.55	-16.9	17.1	-1.2	2.8	3.8	37	-0.03	18.6	91	22	0.32
M-EB	0.6	2.0	3	0.71	0.1	0.2	80	0.66	-14.9	15.0	-0.3	2.3	4.3	63	0.10	21.2	72	49	0.49
M-KB	1.5	2.8	3	0.61	0.1	0.2	64	0.69	-16.5	16.7	-1.7	3.1	4.2	28	-0.09	20.1	81	32	0.40
M-LB	0.6	2.0	4	0.72	0.1	0.2	81	0.67	-14.9	15.0	0.6	2.1	4.0	78	0.10	19.8	74	54	0.50
M-OB	3.8	4.7	6	0.47	0.3	0.3	20	0.54	-17.0	17.1	0.5	2.1	3.2	75	0.06	17.8	97	34	0.36

Table 5.5.: Changes of HA(T,q,ff) and GA(T,q,ff) for both phases, averaged for the experiments ROUGHNESS and TEMPERATURE, between experiment pairs as listed in Table 5.3. For the ROUGHNESS experiment averaging is only performed on pairs that use the extreme roughness lengthes.

experiment	delta HA(T,q,ff)	delta GA(T,q,ff)
ROUGHNESS	7.3	0.01
TEMPERATURE	10.2	0.05
FORCING*	M-EA - M-KA=-21	M-EA - M-KA=0
	M-EB - M-KB=17	M-EB - M-KB=0.09

one hand the air temperature has not yet risen significantly and on the other hand only few sea-ice floes remain. Over sea ice, changes in surface temperature due to changes in sea-ice concentration contribute much more to the air-surface temperature differences than changes in the air temperature. Therefore a strong and nearly linear correlation between sea-ice concentration and the strength of heat fluxes exists (Fig. 4a in Lüpkes and Birnbaum (2005)). The obvious overestimation (e.g. M-LB in Fig. 5.7) of sensible heat fluxes depends on sea-ice concentration, too (Fig. 5.8a). The lower the sea-ice concentration the higher is the overestimation. Higher sea-ice temperatures lead to a reduction, presumably because less convection is triggered due to reduced horizontal surface temperature gradients (Fig. 5.8b).

Table 5.6.: Model skills of ECMWF forcing for the horizontal legs (below 100 m above ground) of phase A (F-A, 13 March 2002), phase B (F-B, 15 March 2002) and of test case from 7 March 2002 (F-7Mar, section 4.4.1). N represents the number of samples used for the calculation of ME, RMSE, H, G and VWD. HA(T,q,ff) and GA(T,q,ff) are the average hit rates and Gandin-Murphy skill scores of temperature, specific humidity and wind speed. $\Delta HA(T, q, ff)$ and $\Delta GA(T, q, ff)$ specify the change in model performance of the named simulations compared to the respective nudging. For details see text.

422678	T [°C]				q [g/kg]				ff [ms-1]					dd [deg]		average		added value	
	ME	RMSE	H	G	ME	RMSE	H	G	ME	RMSE	VWD	H	G	RMSE	H	HA(T,q,ff)	GA(T,q,ff)	$\Delta HA(T,q,ff)$	$\Delta GA(T,q,ff)$
F-7Mar	3.9	9.5	1	0.27	0.1	0.8	3	0.20	-3.8	5.1	7.2	18	-0.06	59	67	7	0.14	M-L: +12, M-M: +9	M-L: 0.24, M-M: 0.16
F-B	3.2	3.8	2	0.71	0.2	0.2	51	0.77	-1.0	2.1	2.8	37	-0.20	8.3	99	30	0.43	ME-B: +19, MK-B +2:	ME-B: 0.06, MK-B: -0.0
F-A	0.7	2.8	12	0.74	0.0	0.2	83	0.45	-3.1	4.5	6.9	15	-0.22	85	16	36	0.32	ME-A:-11, MK-A: +10	ME-A: 0.19, MK-A: 0.19
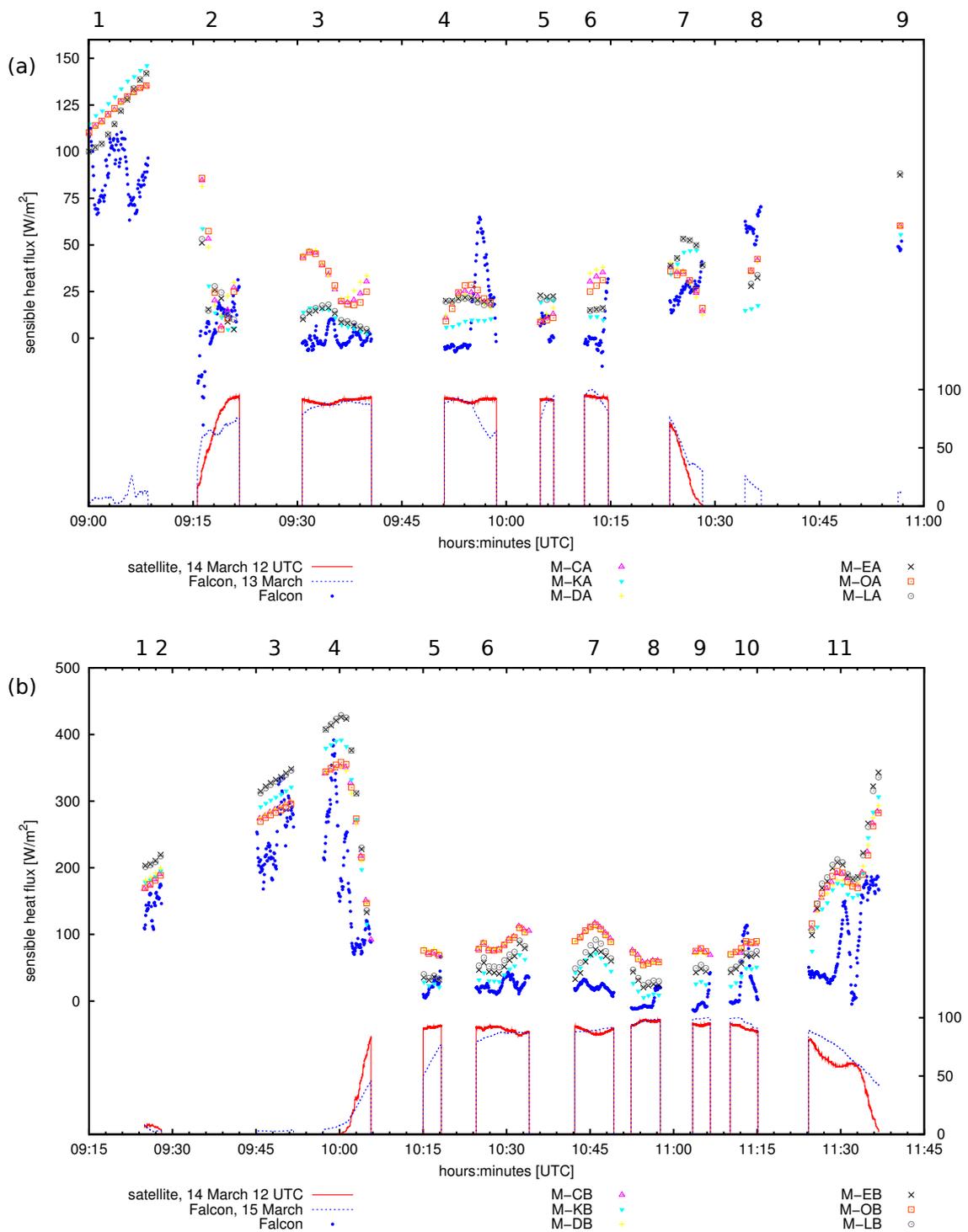
Figure 5.7.: Same as Fig. 5.4 but for sensible heat flux.
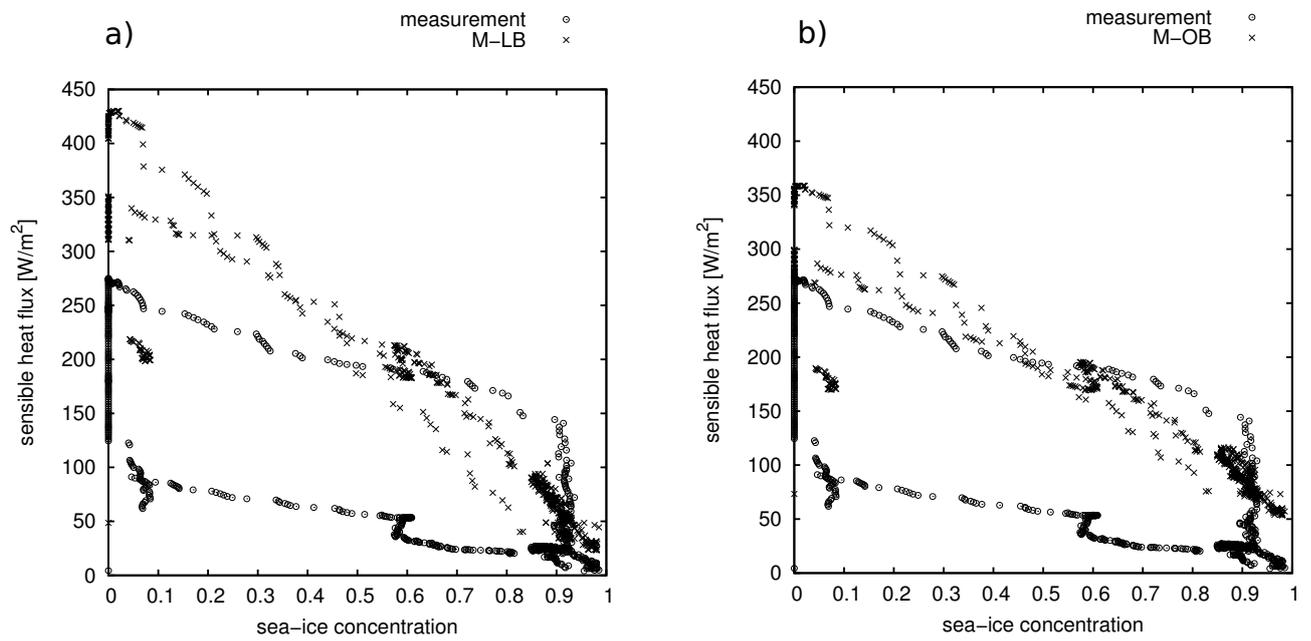
Figure 5.8.: Observed and modelled heat fluxes of phase B as function of sea-ice concentration under unstable conditions.
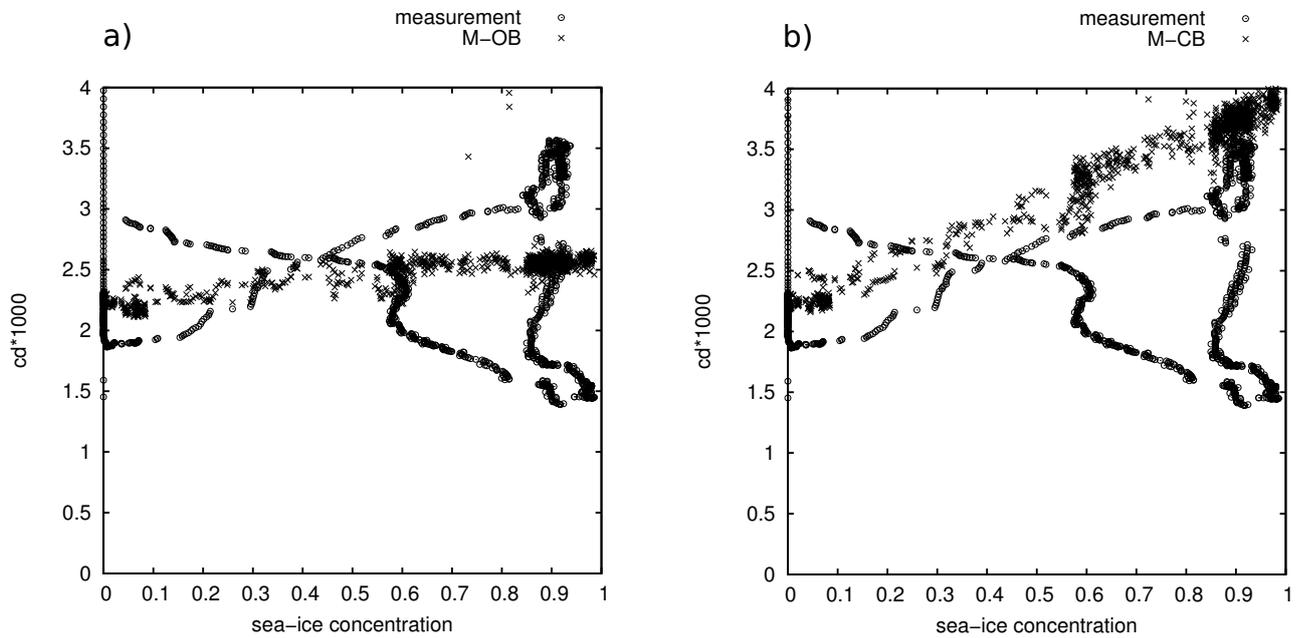
Figure 5.9.: Observed and modelled drag coefficients at aircraft height of phase B as function of sea-ice concentration under unstable conditions.

Even though the surface roughness is usually higher for sea ice than for water, for the momentum fluxes no linear correlation to the sea-ice concentration exists. An additional contribution of the form drag at floe edges and pressure and shear ridges to the friction velocity leads to increased momentum fluxes in the marginal sea-ice zone (Lüpkes and Birnbaum, 2005) but also in the inner sea ice. Lüpkes and Birnbaum (2005) parameterized the dependency of the floe edges on the sea-ice concentration for the marginal sea-ice zone only. According to that a maximal form drag is achieved at an intermediate sea-ice concentration. As the domain investigated here encompasses not only the marginal sea-ice zone but also the inner sea ice, the functional relationship with a maximum friction velocity at 50% sea-ice concentration is not clearly visible in the measurements (Fig. 5.9 a). To enable comparability to Fig. 3 of Lüpkes and Birnbaum (2005), the drag coefficient $c_d = (u_*/u)^2$ at aircraft height rather than u_* is plotted. Around 50% sea-ice concentration at least no minimal drag coefficients are observed. Much higher drag coefficients, however, exist in the inner sea ice, presumably due to pressure ridges. As the distribution of pressure and shear ridges is hardly detectable (Dierking, 1995), the approach of METRAS with a linear dependency of the friction velocity on the sea-ice concentration leads to overall good results. Larger roughness lengthes than $z_0 = 1\,mm$ lead to an overestimation of the drag coefficients (Fig. 5.9 b).

Consequently, M-OB, the simulation with the most realistically simulated heat- and momentum fluxes shows the best performance of all simulations with weak nudging. For phase B the alleged too low surface roughness of $z_0 = 1\,mm$ and the comparatively high sea-ice temperature of $T_h(ice) = -9°C$ describe the sea-ice properties most realistically.

For phase A, however, the larger atmospherical inhomogeneities superpose the dependencies of fluxes on sea-ice concentration described above. But again, sensible heat fluxes are less overestimated for the higher sea-ice temperature (M-OA in Fig. 5.10) and the lower roughness length of $z_0 = 1\,mm$ leads to more realistic drag coefficients (M-OA in Fig. 5.11). In the ranking of all weakly nudged simulations, however, M-OA only manages second place. M-KA, the simulation with the roughest and coldest sea ice, simulates temperature and specific humidity much better, but wind speed worse (Table 5.4). Therefore only the low surface roughness of $z_0 = 1\,mm$ can be recommended with certainty. But, as the influence of the sea-ice temperature on the heat fluxes is low and with the much too warm temperature-time-series of simulations with a sea-ice temperature of $T_h(ice) = -9°C$ in mind (Fig. 5.3), the recommendation of section 4.5 to use a sea-ice temperature of $T_h(ice) = -25°C$ for early spring situations can be maintained.
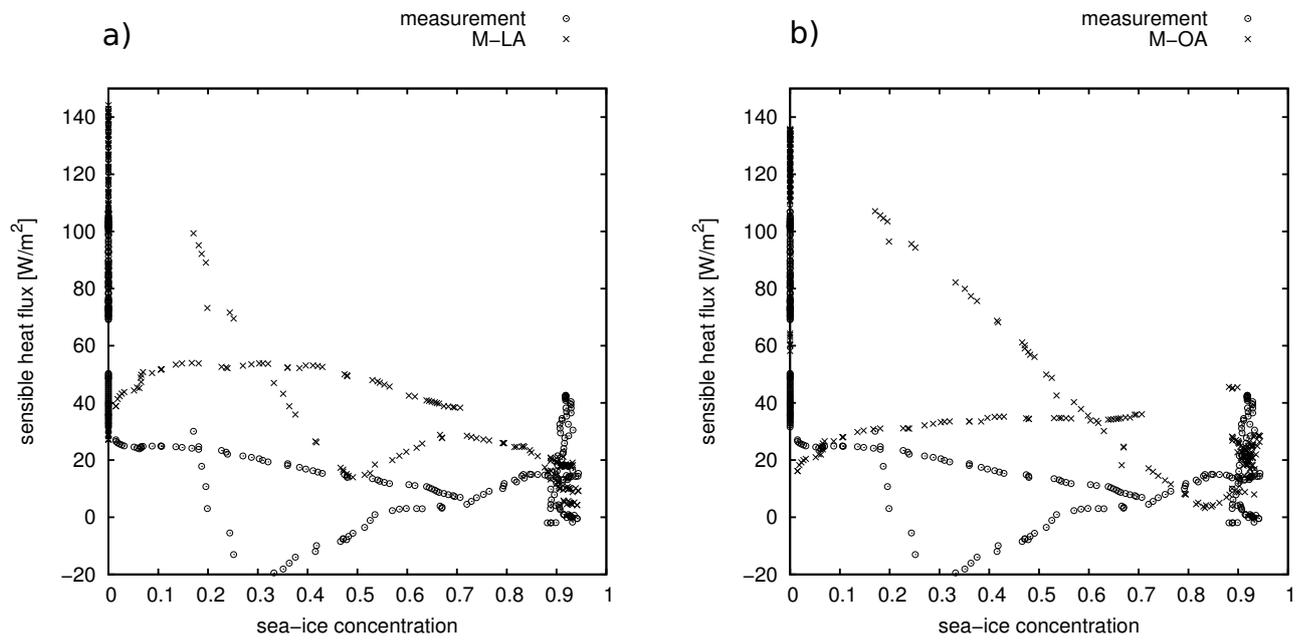
Figure 5.10.: Observed and modelled heat fluxes of phase A as function of sea-ice concentration under unstable conditions.
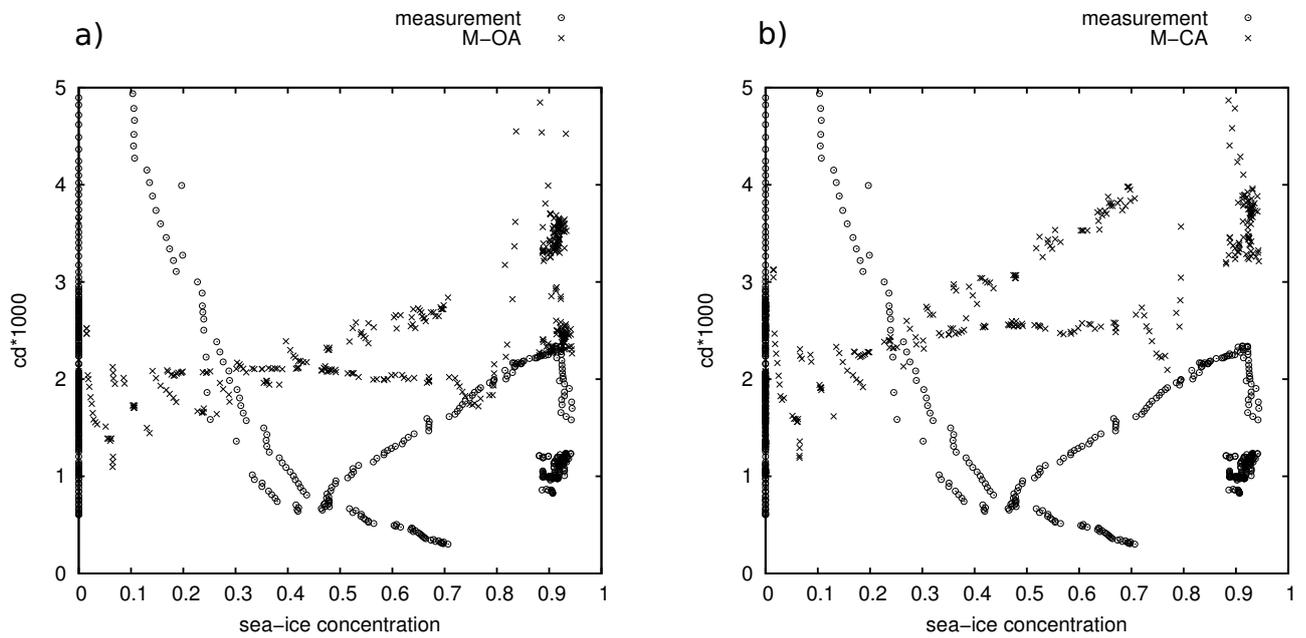
Figure 5.11.: Observed and modelled drag coefficients at aircraft height of phase A as function of sea-ice concentration under unstable conditions.

5.5. Conclusions of chapter 5

The effect of the sea-ice characteristics roughness and sea-ice temperature as well as the effect of lateral nudging strength on mesoscale model performance have been investigated with two meteorological situations. Especially the simulation performance of turbulent surface fluxes were investigated. The first meteorological situation, a mesoscale cyclone movement along the sea-ice edge and finally on-ice has been simulated with six different model set-ups for the Fram Strait. Results have been evaluated for 13 March 2002. The second situation was an off-ice flow for the same model domain on 15 March 2002. The test cases were simulated with the standard version of METRAS, using fixed sea-ice distributions during the simulations. For METRAS the sea-ice characteristics roughness and temperature were varied within the range of observational uncertainty. In addition to the surface parameters, the strength of the lateral nudging was varied.

The evaluation was based on comparisons with measurements from near-ground aircraft flights. Statistical measures were employed for meteorological standard parameters. Time series and scatter plots of selected cases sufficed for the evaluation of simulated turbulent surface fluxes. With respect to the standard meteorological parameters, simulation results for phase A are similarly good as for phase B and even better than the simulation results presented in chapter 4.

It has been shown that, in order to achieve a good model performance, the nudging strength has to be chosen in accordance to the strength of synoptic scale forcing. Too weak nudging under meteorological situations strongly dominated by the synoptic scale leads to loss of information of the driving model and hence to degradation of the mesoscale model. Too strong nudging under meteorological situations that are only weakly forced by the synoptic scale damps the smaller scale circulations which become now more important and can only be simulated by the mesoscale model.

For the synoptic scale dominated situation of phase B the most adequate sea-ice parameters could be identified on basis of assessment of the turbulent surface fluxes. With their comparably good simulation quality, a strong hint is given that the standard meteorological parameters are simulated realistically as consequence of realistic near-ground processes and hence well chosen surface parameters.

With a surface roughness of $z_0 = 1\,mm$ and a sea-ice temperature of $T_h(ice) = -9°C$, surface fluxes of simulation M-OB are in best agreement with measurements and also show the best model performance with respect to standard meteorological parameters. Simulation results could have been further improved if strong nudging was applied. It remains questionable, however, if the specified sea-ice temperature is also suitable for a meteorological situation that is less synoptically forced. For phase A, simulation M-KA with a sea-ice temperature of $T_h(ice) = -25°C$ leads to better

results of the thermodynamic quantities. In respect of near-ground temperature simulation performance the lower sea-ice temperature of $T_h(ice) = -25°C$ is better suited for early spring situations. The conclusions of chapter 4 regarding a well suited setup of sea-ice characteristics can thus be confirmed. The suspicion made there, the low sea ice temperature of $T_h(ice) = -25°C$ only showed better results, because frontal contrasts were artificially improved, has been disapproved by this case. The generalization of this setup to smaller than synoptic scale dominated situations was possible with some uncertainties. Further simulations with different combinations of sea-ice characteristics and nudging strengths are necessary to eliminate these uncertainties.

6. Conclusions and outlook

In this work several methods that are used to represent the surface layer in mesoscale models were investigated regarding their influence on mesoscale model performance. Two different land surface models, several sea-ice properties and the influence of the vertical resolution and the role of surface heterogeneities were investigated.

Four scenarios differing in domain and/or meteorological situation with in total nine model experiments have been investigated. Several different evaluation methods suitable for the respective number of test cases within a model experiment have been applied. It turned out, that the characteristic scale of the meteorological situation sometimes severely interacts with the characteristic scale of the surface heterogeneities. In the case of a trough approaching the sea ice of the Fram Strait (Chapter 4) the characteristic horizontal scales of trough and sea ice were similar and once they were located close enough, effects caused by sea-ice characteristics were not easy to distinguish from effects caused by the trough. In the evaluation of surface layer models for the Bay of Valencia a frontal passage also led to results contrary to findings valid under smaller scale dominated situations (Chapter 2). On the contrary, only for the synoptic scale dominated off ice flow from 15 March 2002 (Chapter 5) scale separation was sufficient to investigate interactions between sea-ice characteristics and turbulent fluxes.

Numerical meteorological models and observation systems have become less restricted to one specific scale. Thanks to improved computer power, mesoscale models are applied to domains large enough to simulate synoptic scale phenomena, with horizontal resolutions as small as the employed physical parameterizations allow. As more and more (satellite-) measurements with a wide range of resolutions become available, it seems sensible to investigate the discrepancy in characteristic scales between the simulated meteorological situation and the applied model. If this discrepancy consists of a large scale meteorolocical situation and a smaller scale modell, the gap might be overcome by nudging. But the nudging strength has to be chosen carefully, as the correlation between model performance and the dependence of nudging strength on the characteristic scale of the meteorological situation is strong. This relation was investigated with a model experiment in Chapters 4 and 5. The nudging had the strongest influence on the model performance of all model experiments.

By applying the MM5 to a mid latitude domain where land-see-breezes frequently develop (the Bay of Valencia, Chapter 2), it was shown that the model performance improves with the increase of the vertical resolution. Especially near-ground unstable vertical stratification, one of the main forcing features for sea-breeze development, could be captured much more realistically. The increased vertical resolution is especially relevant for the wind values. These consistently improve with increasing vertical resolution. On the other hand, the application of a more sophisticated land surface model than the default one could not fulfil the expectations. The sophisticated land surface model could not capture the comparatively large soil moisture variability in the semiarid climate of the investigated region better.

The absence of vegetation limits the number of ground processes of sea ice. Hence the investigated effects of sea-ice surface parameters could be separated easier and could be determined with great certainty. Simulations performed with METRAS showed largest sensitivity on the sea-ice temperature, followed by sea-ice roughness (Chapters 4 and 5). Thermal conductivity of the sea ice played for short term simulations only a minor role. The distribution of the sea ice turned out to be a crucial parameter. Not only its influence on model performance is strong, but it is also subject to sudden changes. Furthermore, due to the coarse temporal resolution of extensive satellite observations, the coupling of an atmosphere model to a prognostic sea-ice model is inevitable, if reliable and accurate simulations or forecasts are desired. A further property crucial for atmospheric simulations is the distribution of sea-ice pressure- and shear-ridges. They have a strong influence on turbulent fluxes. But their detection by remote sensing is still subject of research. Up to now the distribution of ridges can only roughly be parameterized in dependence of the sea-ice class. Once the distribution is known, their contribution as form drag to the total drag could be explicitly simulated.

Acknowledgements

A. Definitions of Statistical Measures

- Mean Error ME:

$$ME = \frac{1}{N} \cdot \sum_{i=1}^{N} (P_i - O_i). \tag{A.1}$$

- Root-mean-square error RMSE:

$$ME = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (P_i - O_i)^2}. \tag{A.2}$$

- Peirce skill score PSS:

$$PSS = \frac{\sum_{i=1}^{C} N_i \sum_{i=1}^{C} \{O_i\}\{P_i\}}{N - \sum_{i=1}^{C} \{O_i\}\{O_i\}}, \tag{A.3}$$

where C denotes the number of classes of a contingency table, Ni is the number of correct forecasts in class i, and braces denote the number of values of that class.

- Correlation coefficient r:

$$r = \frac{\sum_{i=1}^{N} (O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{N} (O_i - \overline{O})} \sqrt{\sum_{i=1}^{N} (P_i - \overline{P})}}. \tag{A.4}$$

- Vector wind difference VWD:

$$VWD = \sqrt{(u_O - u_P)^2 + (v_O - v_P)^2}, \tag{A.5}$$

where u_O and v_O are the components of the observed wind velocity in east and north direction, respectively. Both u_P and v_P are the components of the simulated wind velocity in east and north direction, respectively.

- Index of agreement IOA:

$$IOA = 1 - \frac{N(RMSE)^2}{\sum_{i=1}^{N} (\left| ff_{P_i} - \overline{ff_O} \right| + \left| ff_{O_i} - \overline{ff_O} \right|)^2}, \tag{A.6}$$

where ff_P denotes the simulated wind speed and ff_O denotes the observed wind speed. The overbar denotes the mean.

B. Hit rate

Individual differences between observation- and simulation-pairs are compared against fixed forecast accuracy thresholds. The fraction of pairs passing this is named hit rate (H):

$$H = \frac{100}{N} \cdot \sum_{i=1}^{N} n_i \ with \ n_i = \{1 \ for \ |F_i - O_i| < A, 0 \ for \ |F_i - O_i| > A\} \qquad (B.1)$$

N denotes the total number of comparison data, F_i a single forecast value and O_i a single observation value. The desired accuracy range of model results is denoted by A. The same desired accuracies A are used as given in Dierer et al. (2005) and applied by Schlünzen and Katzfey (2003) and Ries and Schlünzen (2009).

To receive the average hit rate HA the hit rates of the parameters pressure, temperature, specific humidity, wind speed and wind direction are weighted with the number of available comparison data and normalised by the total number of comparison data.

By discriminating against a single threshold only two classes, that of the hits and that of the misses, are built for the hit rate. Consequently, at least for one of the two classes the resolution is poor. Therefore, the desired accuracy A is usually shifted towards small values. The meaning of H then is that for the fraction H of hits during the evaluated period the simulation has a small bias, small amplitude error, high correlation and similar standard deviation (that is small RMSE). But no information on the distribution of the misses is given. That means error types cannot be distinguished by H. Especially, in case of small H it is unknown whether a bias, a phase error, an amplitude error or an error in the standard deviation is primarily responsible.

C. Gandin-Murphy skill score

The Gandin-Murphy skill score (G) equitably measures the accuracy of categorical forecasts whose values have a natural ordering (Gandin and Murphy, 1992). Equity requires the score to value constant forecasts (of just one category) and random forecasts with no skill, here set to zero. Amongst other effects, equity leads to the property that the score assigned to a correct forecast event increases as the climatological probability of the event decreases.

The basic concept is to define a matrix of weights s_{ij} for the matrix of the joint distributions of categorised forecasts and observations $p(y_i, o_j)$. y_i denotes the number of forecasts in category i. o_j denotes the number of observations in category j. $p()$ denotes the probability to encounter a certain forecast-observation combination. The weights s_{ij} must fulfil the equation for G (C.1) so that G ranges from -1 (anti-skilled forecast) via 0 (no skill forecast) to 1 (maximum skill forecast).

$$G = \sum_{i=1}^{I} \sum_{j=1}^{J} p(y_i, o_j) s_{ij} \tag{C.1}$$

The definition of the weights of the scoring matrix for a two-category joint distribution and a four-category joint distribution is presented in the following. For a two-category joint distribution with equal sized bins, the scoring matrix is built from the following three requirements:

1.) The first requirement of equity can be expressed by the equations

$$p(o_1)s_{11} + p(o_2)s_{12} = 0 \tag{C.2}$$

and

$$p(o_1)s_{21} + p(o_2)s_{22} = 0 \tag{C.3}$$

with $p(o_r)$ denoting the marginal distributions of observations (that is the sample climatology). s_{ij} denotes the weights of the scoring matrix.

2.) Demanding symmetry of the scoring matrix ($s_{12} = s_{21}$) reduces the number of weights to three so that the system is solved with only one further equation.

3.) By definition of the best possible score is

$$p(o_1)s_{11} + p(o_2)s_{22} = 1. \tag{C.4}$$

The solution of these equations reveals that for the hits the weights are equal to the winning odds based on the climatological distribution of these events:

$$s_{11} = \frac{p(o_2)}{p(o_1)} \tag{C.5}$$

and

$$s_{22} = \frac{p(o_1)}{p(o_2)}. \tag{C.6}$$

The basic principle for the construction of weights for a joint distribution with an arbitrary number of categories, as proposed by Gerrity (1992), can be understood as a set of bets with varying sharpness on the forecast to meet the correct observation. E.g. for the forecast of event 3 (third row) in a 4x4-category joint distribution table, the following bets are made:

a) The forecast is within the observation-categories 3 and 4.

b) The forecast is within the observation-categories 2 to 4.

c) The forecast is within the observation-categories 1 to 3.

odds of winning can be determined by:

$$wo_{m:n} = \frac{1 - \sum_{r=m}^{n} p(o_r)}{\sum_{r=m}^{n} p(o_r)} \tag{C.7}$$

The weights s_{ij} for each element of the scoring matrix are then the mean over all winning odds for bets won and penalties of -1 for bets lost. For the event 3 being forecasted in a 4x4-category joint distribution table the following weights result:

$$s_{33} = \frac{1}{3} \left(wo_{2:4} + wo_{3:4} + wo_{1:3} \right) \tag{C.8}$$

$$s_{34} = \frac{1}{3} \left(wo_{2:4} + wo_{3:4} - 1 \right) \tag{C.9}$$

$$s_{32} = \frac{1}{3} \left(wo_{2:4} - 1 + wo_{1:3} \right) \tag{C.10}$$

$$s_{31} = \frac{1}{3} \left(-1 - 1 + wo_{1:3} \right) \tag{C.11}$$

Assuming the observations to be equally distributed over all categories results in progressively lower winning odds as the bets become less sharp ($m << n$ in Eq. C.7). For a non-uniformly distributed climatology the winning odds also reduce as the marginal distributions involved in the bet get more populated ($p(o_r)$ in Eq. C.7).

Three requirements have to be fulfilled by the categorisation of a continuous forecast and observation time series: 1.) The number of categories should be large to ensure a proper resolution. 2.) The extreme columns of the joint distribution table have to be occupied in order to compute the weights of G. 3.) All categories must span the same range in order to not distort the climatological distribution (because the weights depend on it). 1.) and 2.) increasingly exclude each other, as the deviation between the set of forecasts and the set of observation grows. In order to gather all observation-simulation pairs, the range of values from the collective minimum of observations and simulations to the collective maximum of the both needs to be considered. However, fulfilling requirement 2) becomes more difficult. For inter model comparison the number of categories should be equal, though the dependency of G on the number of classes is weak for the profiles investigated here.

Bibliography

Beran, J., **B. Calveri**, **B. Lange** and **L. v. Bremen** (**2005**): Offshore wind modelling and forecast. 6th ERF / 15th MM5 Users' Workshop. National Center for Atmospheric Research, Boulder, Colorado.

Birnbaum, G., **R. Brauner** and **H. Ries** (**2006**): Synoptic situations causing high precipitation rates on the Antarctic plateau: observations from Kohnen Station, Dronning Maud Land. *Antarctic Science*, **18**(02), 279–288.

Bogenschutz, P. A. (**2004**): *Skill assessment and benefits on applying the new weather research and forecast model to national weather forecast operations.* Diploma thesis, The Florida State University.

Bohnenstengel, S. I. (**2010**): *Can a simple locality index be used to improve mesoscale model forecasts?* PhD thesis, in preparation, Department Geowissenschaften, Universität Hamburg.

Browning, K. A. (**1980**): Local weather forecasting. *Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, **371**(1745), 179 pp.

Brümmer, B. and **H. Höber** (**1999**): A mesoscale cyclone over the Fram Strait and its effects on sea ice. *Journal Of Geophysical Research-Atmospheres*, **104**(D16), 19 085–19 098.

Brümmer, B., **J. Launiainen**, **G. Müller** and **D. Schröder** (**2005**): FRAMZY2002. Second Field Experiment on Fram Strait Cyclones and their impact on sea ice. Field Report with Measurement Examples. *Berichte Aus Dem Zentrum Für Meeres- Und Klimaforschung*, **A37**.

Brümmer, B., **D. Schröder**, **G. Müller**, **G. Spreen**, **A. Jahnke-Bornemann** and **J. Launiainen** (**2008**): Impact of a Fram Strait cyclone on ice edge, drift, divergence, and concentration: Possibilities and limits of an observational analysis. *Journal Of Geophysical Research-Oceans*, **113**(C12), 15 pp.

Bungert, U. (**2008**): *Einfluss der Nestung auf die Ergebnisse meteorologischer Modelle.* PhD thesis, Meteorological Institute, University of Hamburg.

Chen, F. and **J. Dudhia** (**2001a**): Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 Modeling System. Part I: Model implementation and sensitivity. *Monthly Weather Review*, **129**(4), 569–585.

Chen, F. and **J. Dudhia** (**2001b**): Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 Modeling System. Part II: Preliminary model validation. *Monthly Weather Review*, **129**, 587–604.

Colby, F. P. (**2004**): Simulation of the New England sea breeze: The effekt of grid spacing. *Weather And Forecasting*, **19**, 277–286.

Colle, B., **J. Olson** and **J. Tongue** (**2003**): Multiseason verification of the MM5. Part I: Comparison with the Eta model over the central and eastern United States and impact of MM5 resolution. *Weather And Forecasting*, **18**(3), 431–457.

Cox, R., **B. Bauer** and **T. Smith** (**1998**): A mesoscale model intercomparison. *Bulletin Of The American Meteorological Society*, **79**(2), 265–283.

Davies, H. (**1976**): Lateral Boundary Formulation For Multilevel Prediction Models. *Quarterly Journal Of The Royal Meteorological Society*, **102**(432), 405–418.

Deardorff, J. (**1978**): Efficient prediction of ground surface-temperature and moisture, with inclusion of a layer of vegetation. *Journal Of Geophysical Research-Oceans And Atmospheres*, **83**(NC4), 1889–1903.

Deng, A. and **D. Stauffer** (**2006**): On improving 4-km mesoscale model simulations. *Journal Of Applied Meteorology And Climatology*, **45**(3), 361–381.

DeutscherWetterdienst (**1998**): Beobachterhandbuch für wettermeldestellen des synoptisch-klimatologischen mess- und beobachtungsnetzes.

Dierer, S. and **K. H. Schlünzen** (**2005**): Influence parameters for a polar meso-cyclone development. *Meteorologische Zeitschrift*, **14**(6), 781–792.

Dierer, S., **K. H. Schlünzen**, **G. Birnbaum**, **B. Brümmer** and **G. Müller** (**2005**): Atmosphere-sea ice interactions during a cyclone passage investigated by using model simulations and measurements. *Monthly Weather Review*, **133**(12), 3678–3692.

Dierking, W. (**1995**): Laser profiling of the ice surface-topography during the winter Weddell gyre study 1992. *Journal of Geophysical Research-Oceans*, **100**(C3), 4807–4820.

Dudhia, J. (**1996**): A multi-layer soil temperature model for mm5. Sixth PSU/NCAR Mesoscale Model User´s Workshop. Boulder, Colorado.

Ebert, E. E. (**2008**): Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications*, **15**(1), 51–64.

Fock, B. H. (**2007**): *METRAS als Grobstruktursimulationsmodell.* Diploma thesis, Faculty of Earth Sciences, University of Hamburg.

Foy, B., **L. Molina** and **M. Molina** (**2006**): Satellite-derived land surface parameters for mesoscale modelling of the Mexico City basin. *Atmoshperic Cemistry And Physics*, **6**(5), 1315–1330.

Gandin, L. and **A. Murphy** (**1992**): Equitable Skill Scores For Categorical Forecasts. *Monthly Weather Review*, **120**(2), 361–370.

Gangoiti, G., **M. Millan**, **R. Salvador** and **E. Mantilla** (**2001**): Long-range transport and re-circulation of pollutants in the western Mediterranean during the project regional cycles of air pollution in the west-central Mediterranean area, journal = Atmospheric Environment. **35**(36), 6267–6276.

Gerbush, M. R., **D. A. R. Kristovich** and **N. F. Laird** (**2008**): Mesoscale boundary layer and heat flux variations over pack ice-covered Lake Erie. *Journal of Applied Meteorology and Climatology*, **47**(2), 668–682.

Gerrity, J. (**1992**): a note on Gandin and Murphy equitable skill score. *Monthly Weather Review*, **120**(11), 2709–2712.

Grell, G., **J. Dudhia** and **D. Stauffer** (**1995**): A description of the fifth-generation Penn State/NCAR mesoscale model (MM5).

Griessbaum, F. and **A. Schmidt** (**2009**): Advanced tilt correction from flow distortion effects on turbulent CO2 fluxes in complex environments using large eddy simulation. *Quarterly Journal of the Royal Meteorological Society*, **135**(643, Part B), 1603–1613.

Guest, P. and **K. Davidson** (**1991**): The aerodynamic roughness of different types of sea ice. *Journal Of Geophysical Research-Oceans*, **96**(C3), 4709–4721.

Hack, J., **B. Boville**, **B. Briegleb**, **J. Kiehl**, **P. Rasch** and **D. Williamson** (**1993**): Description of the NCAR community climate model (CCM2).

Hong, J. (**2003**): Evaluation of the high-resolution model forecasts over the Taiwan area during GIMEX. *Weather and Forecasting*, **18**(5), 836–846.

Hong, S. and **H. Pan** (**1996**): Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Monthly Weather Review*, **124**(10), 2322–2339.

Kain, J. (**2004**): The Kain-Fritsch convective parameterization: An update. *Journal Of Applied Meteorology*, **43**(1), 170–181.

Kaleschke, L., **C. Lüpkes**, **T. Vihma**, **J. Haarpaintner**, **A. Bochert**, **J. Hartmann** and **G. Heygster** (**2001**): SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis. *Canadian Journal Of Remote Sensing*, **27**(5), 526–537.

Kalnay, E., **M. Kanamitsu** and **W. Baker** (**1990**): Global numerical weather prediction at the national-meteorological-center. *Bulletin Of The American Meteorological Society*, **71**(10), 1410–1428.

Karcher, M., **R. Gerdes**, **F. Kauker** and **C. Koberle** (**2003**): Arctic warming: Evolution and spreading of the 1990s warm event in the Nordic seas and the Arctic Ocean. *Journal Of Geophysical Research-Oceans*, **108**(C2).

Karcher, M., **R. Gerdes**, **F. Kauker**, **C. Koberle** and **I. Yashayaev** (**2005**): Arctic Ocean change heralds North Atlantic freshening. *Geophysical Research Letters*, **32**(21).

Kottmeier, C., **P. Palacio-Sese**, **N. Kalthoff**, **U. Corsmeier** and **F. Franz** (**2000**): Sea breeze and coastal jets in southeastern Spain. *International Journal Of Climatology*, **20**(14), 1791–1808.

Lange, M. and **H. Waldl** (**2001**): Assessing the uncertainty of wind power predictions with regard to specific weather situations. *Proceedings of the 2001 European Wind Energy Association Conference*, S. 695–698.

Lüpkes, C. and **G. Birnbaum** (**2005**): Surface drag in the Arctic marginal sea-ice zone: A comparison of different parameterisation concepts. *Boundary-Layer Meteorology*, **117**(2), 179–211.

Lüpkes, C., **V. M. Gryanik**, **B. Witha**, **M. Gryschka**, **S. Raasch** and **T. Gollnik** (**2008a**): Modeling convection over Arctic leads with LES and a non-eddy-resolving microscale model. *Journal Of Geophysical Research-Oceans*, **113**(C9).

Lüpkes, C. and **K. H. Schlünzen** (**1996**): Modelling the Arctic convective boundary-layer with different turbulence parameterizations. *Boundary-Layer Meteorology*, **79**(1), 107–130.

Lüpkes, C., **T. Vihma**, **G. Birnbaum** and **U. Wacker** (**2008**): Influence of leads in sea ice on the temperature of the atmospheric boundary layer during polar night. *Geophysical Research Letters*, **35**(3).

Miao, J.-F., **F. Chen** and **K. Borne** (**2007**): Evaluation and comparison of Noah and Pleim-Xiu land surface models in MM5 Using GÖTE2001 data: Spatial and temporal variations in near-surface air temperature. *Journal Of Applied Meteorology*, **46**(10), 1587–1605.

Millan, M., **R. Salvador**, **E. Mantilla** and **G. Kallos** (**1997**): Photooxidant dynamics in the Mediterranean basin in summer: Results from European research projects. *Journal Of Geophysical Research*, **102**(D7), 8811–8823.

Millan, M., **J. Sanz**, **R. Salvador** and **E. Mantilla** (**2002**): Atmospheric dynamics and ozone cycles related to nitrogen deposition in the western Mediterranean. *Environmental Pollution*, **118**(2), 167–186.

Olerud, D. and **A. Sims** (**2003**): MM5 sensitivity modeling in support of VISTAS (visibility improvement - state and tribal association). Draft.

Pacanowski, R. C. (**1995**): MOM 2 documentation, user's guide and reference manual. Tech. Rep. 3, Princeton Univ., Princeton, N. J.

Perez-Landa, G., **P. Ciais**, **M. J. Sanz**, **B. Gioli**, **F. Miglietta**, **J. L. Palau**, **G. Gangoiti** and **M. M. Millan** (**2006**): Mesoscale circulations over complex terrain in the Valencia coastal region, Spain - Part 1: Simulation of diurnal circulation regimes. *Atmoshperic Cemistry And Physics*, **6**(2), 2809–2852.

Persson, A. and **F. Grazzini** (**2005**): User guide to ECMWF forecast products. Techn. Ber. Meteorological Bulletin M3.2, European Centre for Medium-Range Weather Forecasts.

Pielke, R. A. (**2002**): *Mesoscale meteorological modeling.* Academic Press, 676 S.

Putkonen, J. (**1998**): Soil thermal properties and heat transfer processes near Ny-Alesund, northwestern Spitsbergen, Svalbard. *Polar Research*, **17**(2), 165–179.

Raasch, S. and **G. Harbusch** (**2001**): An analysis of secondary circirculations and their effects caused by small-scale surface inhomogeneities using large-eddy simulation. *Boundary-Layer Meteorology*, **101**(1), 31–59.

Rasmussen, E., **P. Guest** and **K. Davidson** (**1997**): Synoptic and mesoscale atmospheric features over the ice-covered portion of the Fram Strait in spring. *Journal Of Geophysical Research-Atmospheres*, **102**(D12), 13 975–13 986.

Reisner, J., **R. Rasmussen** and **B. R.** (**1998**): Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quarterly Journal Of The Royal Meteorological Society*, **124**(548), 1071–1107.

Ries, H. and **K. H. Schlünzen** (**2009**): Evaluation of a mesoscale model with different surface parameterizations and vertical resolutions for the bay of Valencia. *Monthly Weather Review*, **137**(8), 2646–2661.

Ries, H., **H. Schlünzen**, **B. Brümmer**, **M. Claussen** and **G. Müller** (**2010**): Impact of surface parameter uncertainties on the development of a trough in the Fram Strait region. *Tellus A*, **62**, 377 – 392.

Romero, R., **C. Doswell** and **R. Riosalido** (**2001**): Observations and fine-grid simulations of a convective outbreak in northeastern Spain: Importance of diurnal forcing and convective cold pools. *Monthly Weather Review*, **129**(9), 2157–2182.

von Salzen, K., **M. Claussen** and **K. Schlünzen** (**1996**): Application of the concept of blending height to the calculation of surface fluxes in a mesoscale model. *Meteorol. Zeitschrift*, **5**, 60–66.

Schlünzen, H. and **J. Katzfey** (**2003**): Relevance of sub-grid-scale land-use effects for mesoscale models. *Tellus. Series A, Dynamic Meteorology And Oceanography*, **55**, 232–246.

Schlünzen, K. H. (**1990**): Numerical studies on the inland penetration of sea breeze fronts at a coastline with tidally flooded mudflats. *Beitr. Phys. Atmosph.*, **63**, 243–256.

Seneviratne, S. I., **D. Lüthi**, **M. Litschi** and **C. Schär** (**2006**): Land-atmosphere coupling and climate change in Europe. *Nature*, **443**, 205–209.

Spreen, G., **L. Kaleschke** and **G. Heygster** (**2008**): Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal Of Geophysical Research-Oceans*, **113**(C2).

Taylor, K. (**2001**): Summarizing multiple aspects of model performance in a single diagram. *Journal Of Geophysical Research-Atmospheres*, **106**(D7), 7183–7192.

Thorndike, A. and **R. Colony** (**1982**): Sea ice motion in response to geostrophic winds. *Journal Of Geophysical Research-Oceans And Atmospheres*, **87**(NC8), 5845–5852.

Tonnesen, G. and **Coauthors** (**2005**): 2004 interim report for the western regional air partnership (WRAP) regional modeling center (RMC).

Trigo, I., **T. Davies** and **G. Bigg** (**1999**): Objective climatology of cyclones in the Mediterranean region. *Journal Of Climate*, **12**(6), 1685–1696.

Valkonen, T., **T. Vihma** and **M. Doble** (**2008**): Mesoscale modeling of the atmosphere over Antarctic sea ice: A late-autumn case study. *Monthly Weather Review*, **136**(4), 1457–1474.

Vihma, T., **J. Hartmann** and **C. Lüpkes** (**2003**): A case study of an on-ice air flow over the Arctic marginal sea-ice zone. *Boundary-Layer Meteorology*, **107**(1), 189–217.

Wacker, U., **H. Ries** and **U. Schättler** (**2009**): Precipitation simulation for Dronning Maud Land using the COSMO Model. *Antarctic Science*, **21**(6), 643–662.

Wang, X. M., **W. S. Lin**, **L. M. Yang**, **R. R. Deng** and **H. Lin** (**2007**): A numerical study of influences of urban land-use change on ozone distribution over the Pearl River Delta region, China. *Tellus Series B-Chemical and Physical Meteorology*, **59**(3), 633–641.

Warner, T., **B. Mapes** and **M. Xu** (**2003**): Diurnal patterns of rainfall in northwestern South America. Part II: Model simulations. *Monthly Weather Review*, **131**(5), 813–829.

Wu, J., **D. Guan**, **X. Sun**, **G. Yu**, **X. Zhao**, **S. Han** and **C. Jin** (**2005**): Eddy flux corrections for CO2 exchange in broad-leaved Korean pine mixed forest of Changbai Mountains. *Science in China Series D-Earth Sciences*, **48**(Suppl. 1), 106–115.

Yang, Y., **Y. Chen** and **F. Fujjoka** (**2005**): Numerical simulations of the island-induced circulations over the Island of Hawaii during HaRP. *Monthly Weather Review*, **133**(12), 3693–3713.

Zhong, S. and **J. Fast** (**2003**): An evaluation of the MM5, RAMS, and Meso-Eta models at subkilometer resolution using VTMX field campaign data in the Salt Lake Valley. *Monthly Weather Review*, **131**(7), 1301–1322.