

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

By HINNERK RIES^{1*}, K. HEINKE SCHLÜNZEN¹, BURGHARD BRÜMMER¹,
MARTIN CLAUSSEN^{1,2} and GERD MÜLLER¹, ¹ZMAW, Meteorological Institute, University of Hamburg,
Bundesstrasse 55, Hamburg 20146, Germany; ²Max Planck Institute for Meteorology, Bundesstrasse 53, Hamburg
20146, Germany

(Manuscript received 11 December 2009; in final form 29 March 2010)

ABSTRACT

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 fluid 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 homogenization 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.

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 inter-annual 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 warm front, passing westwards 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

*Corresponding author.

e-mail: hinnerk.ries@zmaw.de

DOI: 10.1111/j.1600-0870.2010.00451.x

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 characterizing 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). Parametrization 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; Brümmner 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 2, the meteorological situation and the experimental setup are described in detail. In Section 3, the model setup and forcing are discussed. Results are given in Section 4. Conclusions are drawn in Section 5.

2. Comparison data and experiment

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. 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. 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 characterized by 5/8 to 8/8 strato cumulus with a clearly marked

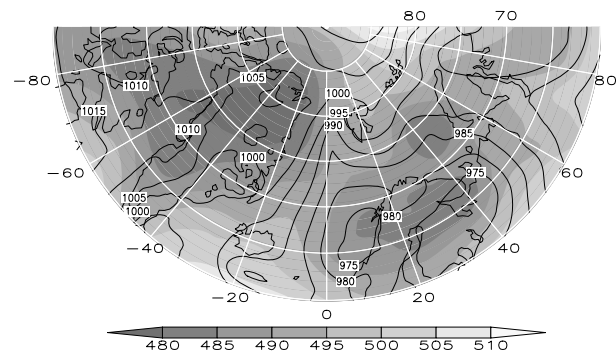


Fig. 1. ECMWF reanalysis of 500 hPa geopotential heights (gpdm) (shaded) and mean sea level pressure (hPa) (contours) for 0 UTC, 7 March 2002.

cloud base and the cloud top lower than on the western side (Brümmner et al., 2005). The profiles over the ocean at 78.84°N , 7.87°E (denoted P1 in Figs. 2, 3 and 4) and over the sea ice at 80.6°N , 5.16°W (denoted P7 in Figs. 2, 3 and 4) both show a well mixed boundary layer (Fig. 3). 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. 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. 3b). The veering of the geostrophic wind with height from N to SE in 900 m at P7 denotes warm air advection (Fig. 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. 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. 6).

2.2. Comparison data: aircraft, ship, buoys

The METRAS model simulations for the trough passage described in Section 2.1 are evaluated against measurements gathered within FRAMZY 2002 (Brümmner 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. 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

Fig. 2. Wind vector (m s^{-1}) (arrows), surface isobars (hPa) (black solid) and temperature ($^{\circ}\text{C}$) (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).

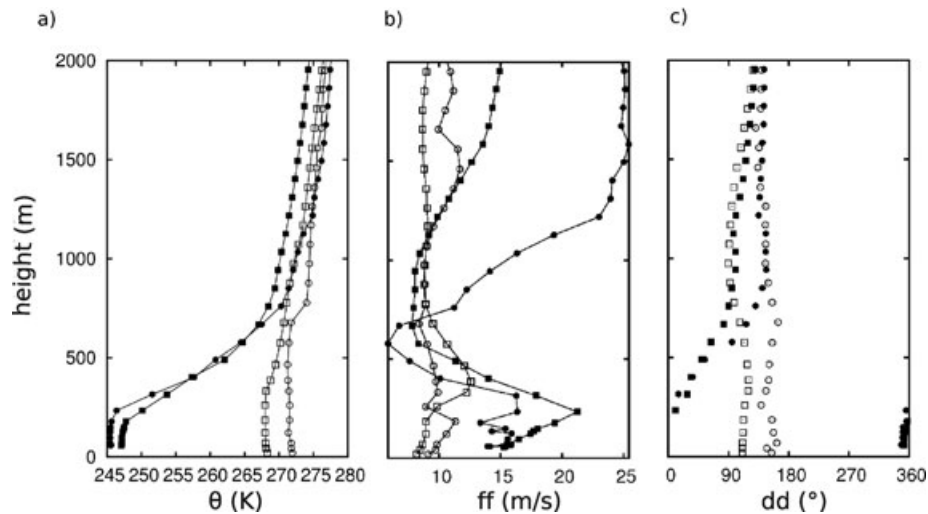
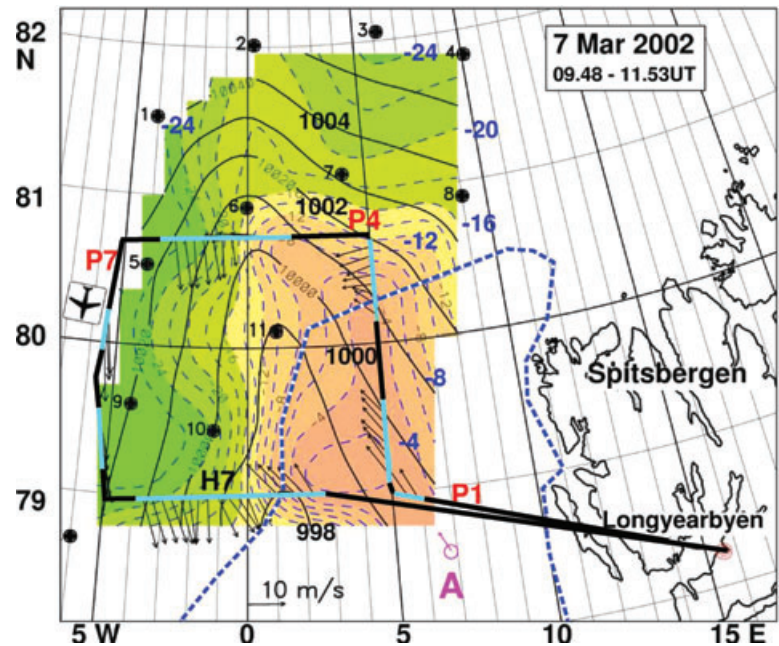


Fig. 3. Observed (dots) and modelled (rectangles) vertical profiles of a) potential temperature Θ , b) wind speed, and c) wind direction at locations P1 (open symbols) and P7 (full symbols) on 7 March 2002 9:50 UTC and 10:55 UTC, respectively. For positions of profiles P1, P7 see Fig. 2.

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. In addition,

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.

3. Model set-up and forcing

The test cases are simulated with METRAS. METRAS is a non-hydrostatic mesoscale model which employs the Boussinesq-approximation, the anelastic approximation and a domain-constant Coriolis parameter. For Arctic regions the model has

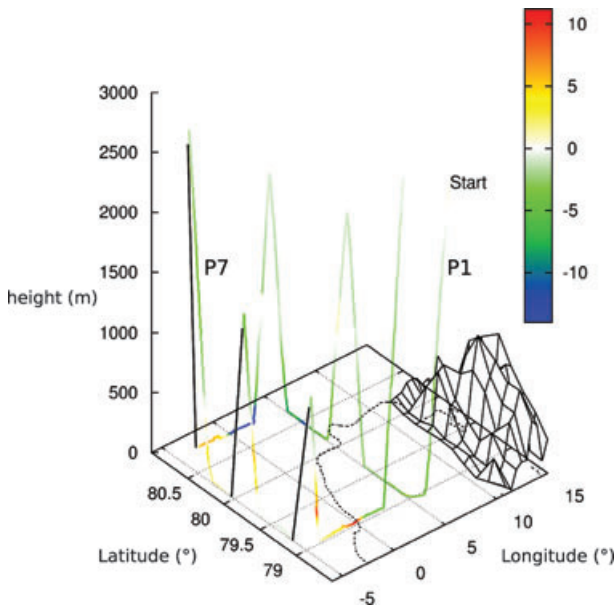


Fig. 4. Flight profile flown counter-clockwise on 7 March 2002 09:47 UTC to 11:53 UTC with temperature deviation of test case M-E from measurements (K) as indicated by the colour bar. Black profile where measurement failed. The dotted line marks the sea-ice edge as measured on 7 March 2002 at 18 UTC, Spitzbergen is shown schematically.

been applied for real case studies (Dierer et al., 2005; Vihma et al., 2003; Lüpkes and Schlünzen, 1996) as well as for process studies (Lüpkes et al., 2008a,b). For turbulence parametrization, 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 centred differences) used ensures horizontal diffusion. Parametrizations 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) (eq. 1)

$$\frac{\partial T_s}{\partial t} = \frac{2\sqrt{\pi}k_s}{\nu_s h_\theta} H_A - 2\pi \frac{T_s - T_h}{\tau_1}. \quad (1)$$

The first term on the right-hand 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 temper-

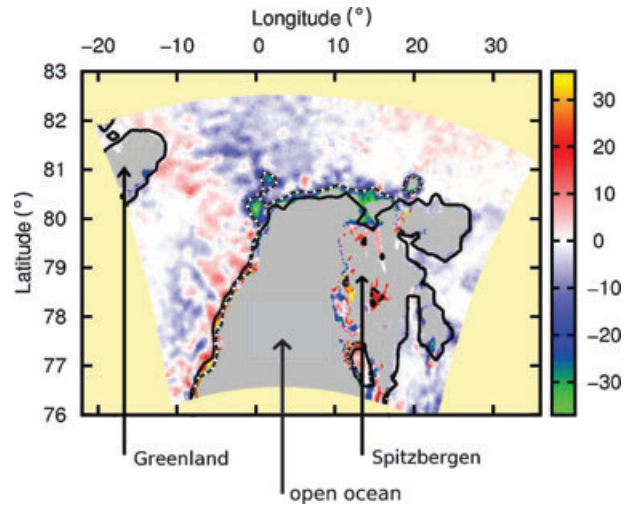


Fig. 5. Difference between observed sea-ice distributions of 7 March 2002 18 UTC and 6 March 2002 14 UTC. Areas coloured in yellow are outside of the METRAS model domain. The colour bar indicates differences in percentage of sea-ice cover; note that the range of differences is not fully resolved by the 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).

ature wave (h_θ). The values used for sea ice are specified in Table 1. The depth of the temperature wave depends on thermal diffusivity and the time span τ_1 , here 1 day: $h_\theta = \sqrt{k_s \tau_1}$, yielding $h_\theta = 0.26$ m. The second term tends to restore T_s within 1 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 13 501 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 (Fig. 5). All simulations start for 5 March 2002, 1800 UTC, and are interpreted 40 h after model initialization 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 2.2).

In this 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

Fig. 6. Histograms of measured sea-ice concentrations of METRAS grid cells for the whole domain.

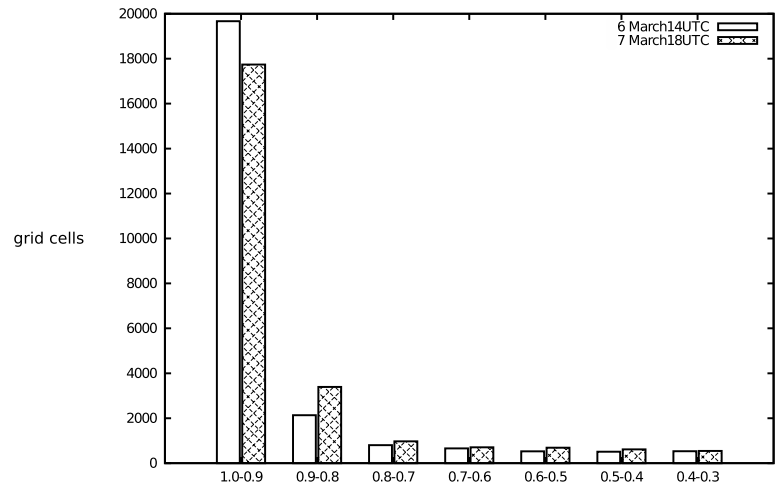


Table 1. Surface model parameters for sea ice

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)
$T_h(\text{ice})$	-9°C (^a), -25°C (^b)	^a Putkonen (1998), ^b Brümmer et al. (2005)
ν_s	$1 \text{ J m}^{-1} \text{ s}^{-1} \text{ K}^{-1}$ (snow), $2.025 \text{ J m}^{-1} \text{ s}^{-1} \text{ K}^{-1}$ (sea ice)	Pielke (2002)
k_s	$0.8 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$	Pielke (2002)
Albedo	72% (pure white snow)	Pielke (2002)

as constant parameters. These parameters have to be prescribed on basis of measurements (Table 1).

In the frame of the sensitivity study, the parameters, z_0 , representing sea-ice roughness, $T_h(\text{ice})$, 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 1. All these values lie in the reasonable range that is covered by the measurements as mentioned in Table 1. Only the extreme roughnesses exaggerate realistic areal averages, because 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 $T_h(\text{ice}) = -9^\circ\text{C}$ 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 $T_h(\text{ice})$. 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.

3.1. Forcing data and their assimilation

To prescribe the large-scale situation, four different data sets 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.

3.1.1. ECMWF. The large-scale meteorological condition is taken from ECMWF six hourly re-analyses 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. 2)

$$\Psi_f = \Psi_m + \delta(\Psi_l - \Psi_m), \quad (2)$$

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. 3)

$$\delta = \nu \Delta t. \quad (3)$$

Within the first half of the initialization phase of 110 min, ν increases from 0 to 0.001 s^{-1} (or 0.01 s^{-1}). These values are spatially constant. In the second half of the initialization 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 eq. (4):

$$\nu(i) = \nu_0(1 - \tanh(a_f i)). \quad (4)$$

Here, i is the distance from the lateral boundary in grid points and ν_0 is set to 0.001 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 four gridpoints and four gridpoints, respectively. The combination of $\nu_0 = 0.001 \text{ s}^{-1}$ and $a_f = 0.4$ is used as weak forcing. The combination of $\nu_0 = 0.01 \text{ s}^{-1}$ 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 (Bohnstengel, 2009).

3.1.2. 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 $1/12^\circ$ 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. 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 upwelling—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.

3.1.3. Satellite observations. The fractional sea-ice cover is calculated from the 'Special Sensor Microwave/Imager'-radiometer (SSM/I) measurements (Brümmer et al., 2008) with

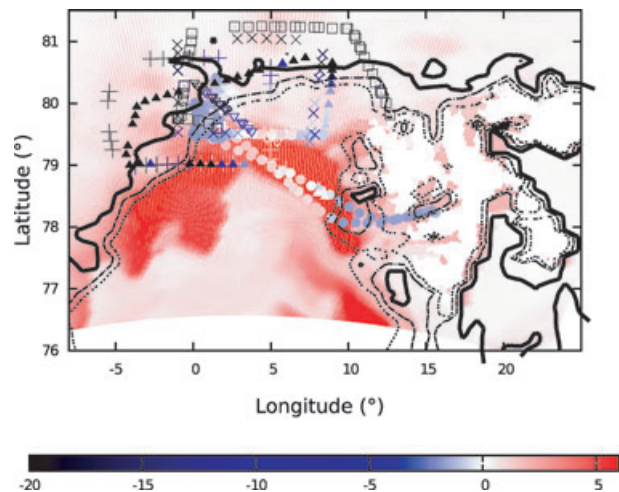


Fig. 7. Water temperatures, surface temperatures ($^\circ\text{C}$) and sea-ice edge. Water temperatures from 5 m depth simulated by NAOSIM for 6 March 2002 0 UTC (shaded) and measured by RV Aranda in 1.5 m depth from 05 March 2002 to 15 March 2002 (shaded dots). Surface temperatures measured by aircraft 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.

the ARTIST sea-ice algorithm (Kaleschke et al., 2001). SSM/I scans are available approximately every 2 h 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. 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 pre-processing and model initialization.

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 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 3.1.1). 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 2). The sensitivity study is organized 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 3.

An additional test case M-N, with an 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.

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 A, values for desired simulation accuracy are given in Table 6) are computed

Table 2. Setup of all test cases

Test case	Characteristic				
	Forcing Weak: $a_f = 0.4$, $v_0 = 0.001s^{-1}$ Strong: $a_f = 0.2, v_0 = 0.01s^{-1}$	Roughness $z_0(\text{mm})$	Temperature $T_h(\text{ice})(^\circ\text{C})$	Thermal conductivity $\nu_s (\text{Jm}^{-1}\text{s}^{-1}\text{K}^{-1})$	Distribution 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 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		

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.

The simulation results valid along the aircraft track (Fig. 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) (Appendix B). 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

normalized 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 centred pattern root mean square difference Δc (eq. 5) is the measure of error inherent of the Taylor diagram

$$\Delta c = \left(\frac{1}{N} \sum_{n=1}^N [(f_n - \bar{f}) - (r_n - \bar{r})]^2 \right)^{0.5}, \quad (5)$$

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

4. Simulation results

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) are given in Fig. 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 (Figs. 2 and 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 and 2 hPa, whereas the simulation shows 3 and 1 hPa. The temperature increase measured at 79°N from the western boundary to the temperature maximum is 23 K. The

Table 4. 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 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
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
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
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
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
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
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
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
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
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
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
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
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
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
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

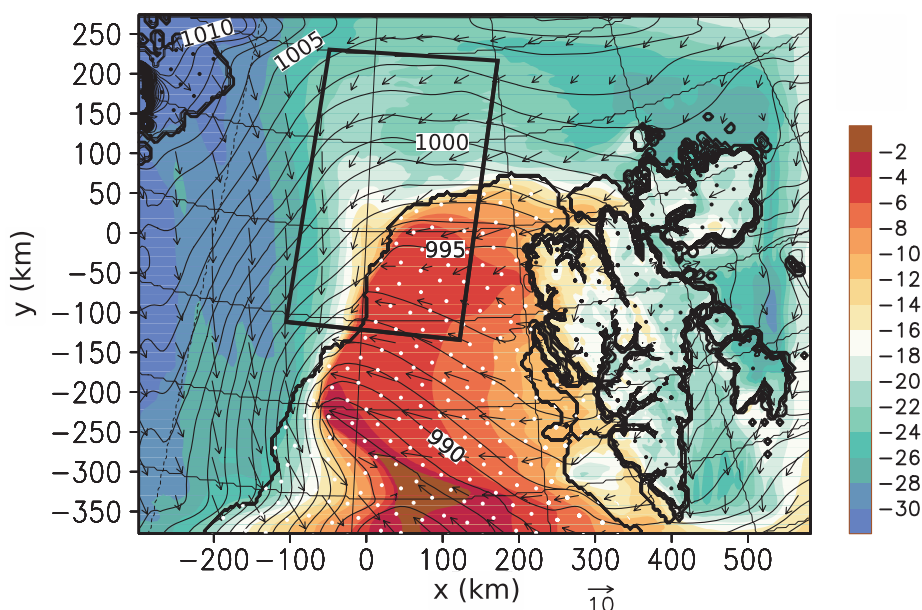


Fig. 8. Simulation results of M-E for 7 March 2002 11 UTC. Temperature ($^{\circ}\text{C}$) at 10 m above ground as indicated by the colour bar, surface pressure (hPa) as isobars, and wind field (m s^{-1}) at 10 m above ground as vectors. Sea-ice edge from 7 March 2002 18 UTC. Open ocean and land surfaces are identified by white and black dots, respectively. Frame marks the approximate area shown in Fig. 2.

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), 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. 3). On the cold side (P7 in Fig. 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. 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. 2 and 8).

A visual impression of the result bandwidth caused by the uncertainties in the initial data is given in Fig. 9. Differences between the two test cases set up most contrarily (M-B and M-E, see Table 2 for set-ups) are illustrated. With respect to HA, they are the least and best performing test cases (Table 4). 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. Although 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. 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 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 5 shows the overall change of evaluation measures for the experiments defined in Table 3. The scores HA and GA agree on the impact of the initialization characteristics. The forcing is most relevant which indicates that the synoptic situation is dominating the mesoscale structures. The least important initialization 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 4.2, 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

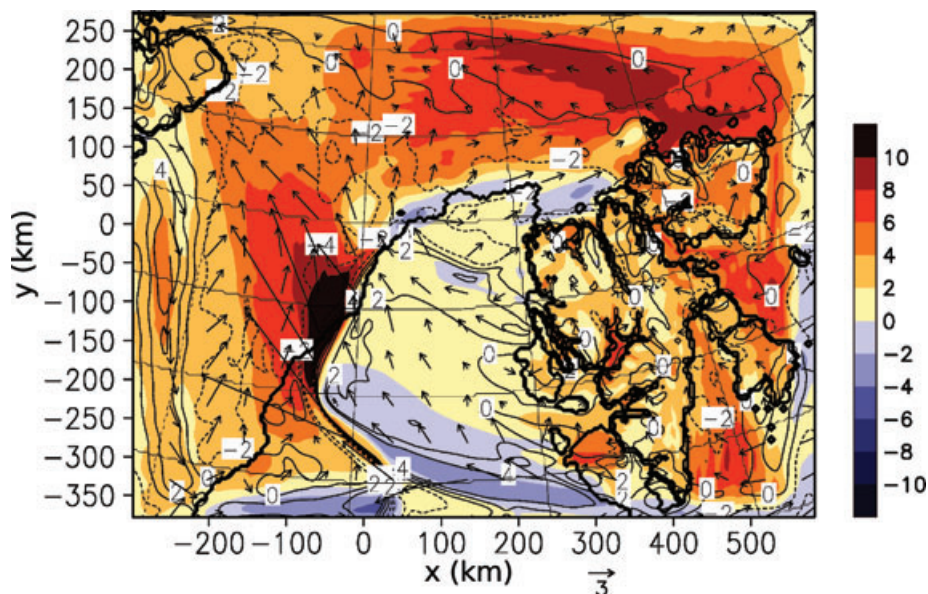


Fig. 9. Difference fields of temperature (K) for M-B to M-E at 10 m height above ground for 7 March 2002 11 UTC. The colour bar denotes temperature differences. Vectors and isotaches (m s^{-1}) denote differences in the wind field. Sea-ice edge from 6 March 2002 14 UTC.

Table 5. Average changes of HA and GA between experiment pairs as listed in Table 3. For the roughness experiment averaging is only performed on pairs that use the extreme roughness lengths

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

was used with more time steps to update the boundary values. An interval of 1 h compared to 6 h 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, personal communication, 2008).

4.2. Influences on temperature

Two main features characterize the measured time series of temperature along the horizontal flight legs (filled dots, Fig. 10): the

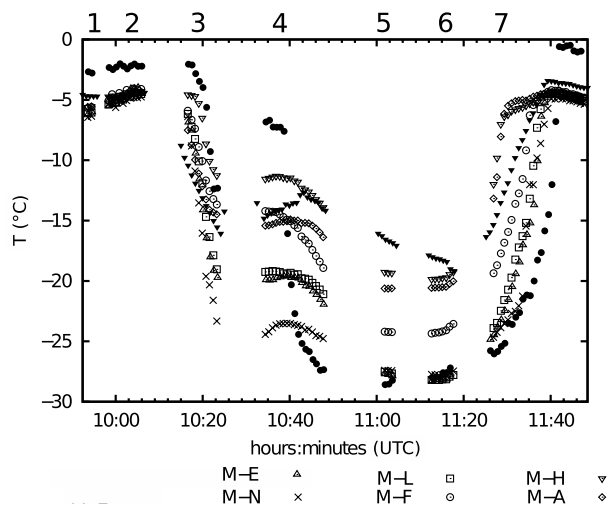


Fig. 10. Observed (full dot), from ECMWF reanalyses (down-pointing triangle) and modelled (other symbols) time series of temperature along the horizontal flight legs (numbering given on top) for 7 March 2002. Legs one and two are over the ocean. The sea-ice edge is crossed at 10:22 UTC, the trough at 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 for locations.

Table 6. Desired accuracy A of model results and accepted measurement values

	Temperature ($^{\circ}\text{C}$)	Specific humidity (g kg^{-1})	Pressure (hPa)	Wind speed (m s^{-1})	Wind direction ($^{\circ}$)
Accuracy	± 0.5	± 0.2	± 1.5	± 1.4	± 20
Threshold				$1 < \text{wind speed}$	$1 < \text{wind speed}$

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. 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 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 DISTRIBUTION 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 normalized standard deviations of 1.25–1.8 with correlation coefficients of 0.45–0.55 (Fig. 11). With the later sea-ice map from 7 March 2002 18 UTC, which is approximately 7 h later than the measurement time, the correlation coefficients improve to the range of 0.61–0.63. The normalized 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 coefficients 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 normalized standard deviation) that primarily result from a wrong sea-ice cover distribution remain unchanged. The more homogeneous distribution of the later sea-

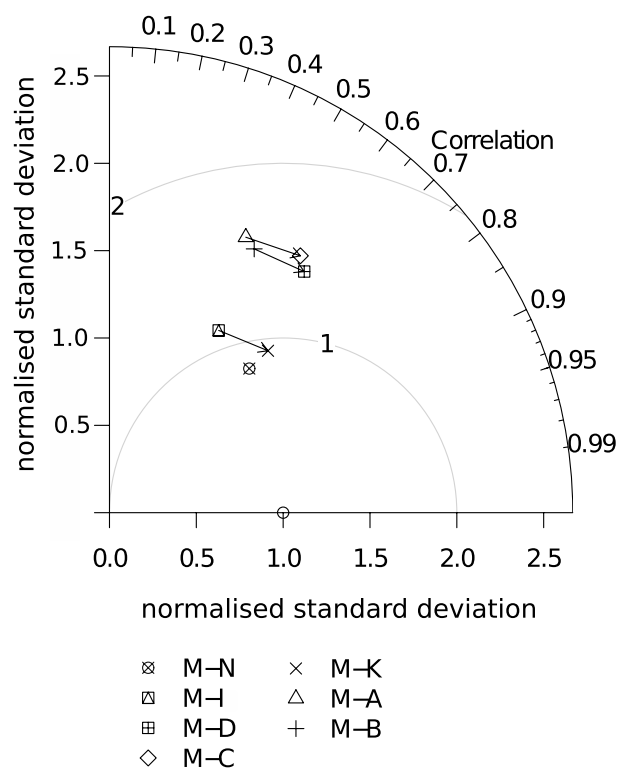


Fig. 11. Taylor diagram of temperatures for members of DISTRIBUTION experiment. The normalized standard deviation is given by the radial distance from origin. The correlation is given by the azimuthal angle. The Δc between the measurement (circle at correlation equal one and normalized standard deviation 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.

ice concentration (Fig. 6) has no homogenizing effect on the standard deviation of the temperature time series. Only the artificially homogenized sea-ice map (M-N, any sea-ice fraction set to 100%) leads to a reduction of normalized 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 h 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. 12. The amplitude error of the simulations performed with the sea-ice temperature of -9°C is reduced by initializing the sea-ice with -25°C . Only the test cases with the lower sea-ice temperature simulate a realistic cooling over sea ice as already visible in Fig. 10. Phase errors are nearly the same, because the correlation coefficient is more or less unchanged.

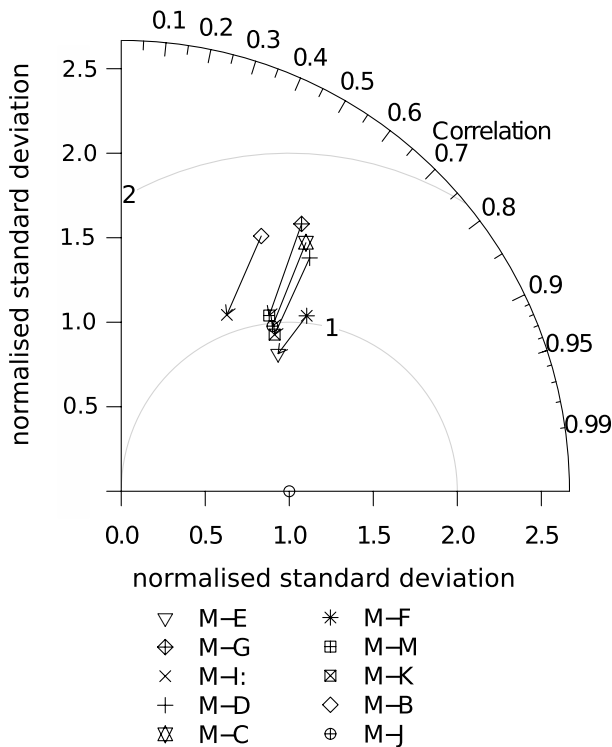


Fig. 12. As Fig. 11 but for the TEMPERATURE experiment. Arrows point from realizations with sea-ice temperature = -9°C to realizations with the sea-ice temperature = -25°C .

The changes in model performance found for the experiments FORCING, ROUGHNESS 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.

4.3. Influences on near-ground wind

As already indicated by the ranking of the averaged delta HA and GA (Table 5), 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. 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 to -2.1 m s^{-1} for the test cases with $z_0 = 27 \text{ mm}$ (M-B, M-D, M-I, M-K; see Table 4) to the range of -0.5 to -1.2 m s^{-1} for the test cases with $z_0 = 5.9 \text{ mm}$ (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 $z_0 = 1 \text{ mm}$. In the test cases M-L,

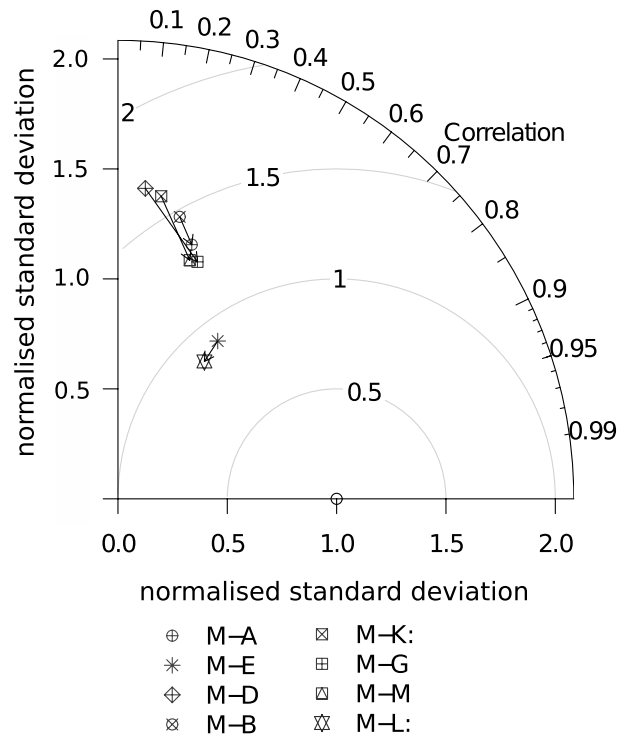


Fig. 13. As Fig. 11 but of wind speeds for members of ROUGHNESS experiment. Arrows point from realizations with higher roughness to realizations with lower roughness.

M-M, M-G the negative bias ranges from -0.1 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. 14a). The dipole structure at $x = -80 \text{ km}$, $y = -200 \text{ km}$ in Fig. 14a) 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 ($z_0 = 1 \text{ mm}$) with M-D ($z_0 = 27 \text{ mm}$), the off-ice flow west of the trough increases by more than 4 m s^{-1} (Fig. 14b). 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 .

5. Conclusions

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

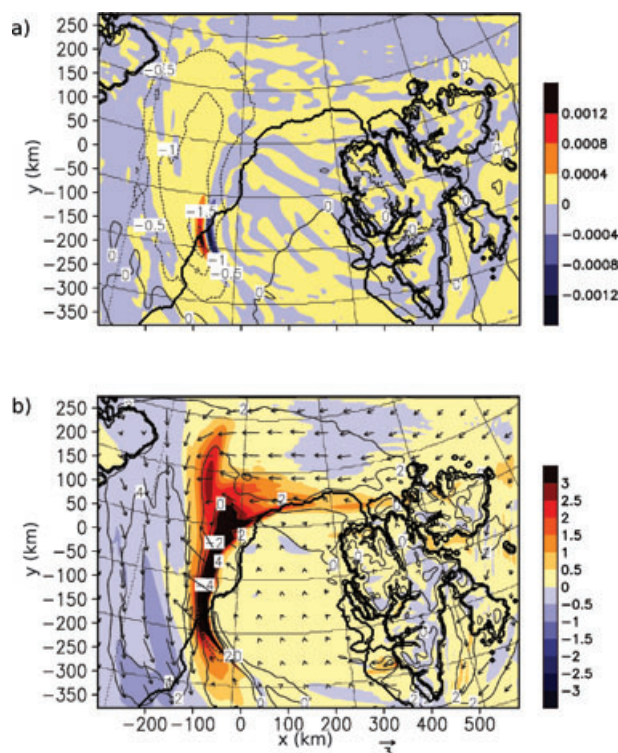


Fig. 14. Horizontal cross-section at 10 m above sea level of simulated differences between M-G ($z_0 = 1$ mm) and M-D ($z_0 = 27$ mm) of (a) mean sea level pressure (hPa, labelled contours) and vorticity (s^{-1} , shaded), (b) wind vector (ms^{-1} , scaling with respect to legend vector), wind speed (ms^{-1} , labelled contours) and temperature (K, shaded).

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 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 idealized 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 ($T_h(\text{ice}) = -25^\circ\text{C}$), and a sea-ice cover that is similar to the observed. Therefore the late sea-ice distribution with a discrepancy of only 7 h to the observation period produces better results than the earlier one at 21 h 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 $T_h(\text{ice}) = -25^\circ\text{C}$ 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 $T_h(\text{ice}) = -25^\circ\text{C}$ are recommended for early spring situations.

There remains to examine whether for less synoptic scale situations, for example 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.

6. Acknowledgments

The authors thank Lars Kaleschke, Gunnar Spreen and Stefan Kern (Institute for Oceanography, University of Hamburg) for providing and discussing satellite observations of the sea-ice cover. The authors also thank Rüdiger Gerdes (Alfred-Wegener-Institute, Bremerhaven) for providing NAOSIM results for the sea-surface temperature and Arne Kriegsmann (Meteorological Institute, University of Hamburg) for making them readable.

This research was funded by the German Science Foundation (DFG) under grant SFB 512-E6 ‘Interactions between mesoscale cyclones and sea ice in the Fram Strait’.

7. Appendix A: Hitrate

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)

$$H = \frac{100}{N} \sum_{i=1}^N n_i \text{ with } n_i = \begin{cases} 1 & \text{for } |F_i - O_i| < A, \\ 0 & \text{for } |F_i - O_i| > A. \end{cases} \quad (\text{A1})$$

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 Schl nzen and Katzfey (2003) and Ries and Schl nzen (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 normalized 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.

8. Appendix B: 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 categorized 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 (B1) so that G ranges from -1 (anti-skilled

forecast) via 0 (no skill forecast) to 1 (maximum skill forecast).

$$G = \sum_{i=1}^I \sum_{j=1}^J p(y_i, o_j) s_{ij}. \quad (\text{B1})$$

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

$$p(o_1)s_{11} + p(o_2)s_{12} = 0 \quad (\text{B2})$$

and

$$p(o_1)s_{21} + p(o_2)s_{22} = 0 \quad (\text{B3})$$

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

$$p(o_1)s_{11} + p(o_2)s_{22} = 1. \quad (\text{B4})$$

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

$$s_{11} = \frac{p(o_2)}{p(o_1)} \quad (\text{B5})$$

and

$$s_{22} = \frac{p(o_1)}{p(o_2)}. \quad (\text{B6})$$

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. For example for the forecast of event 3 (third row) in a 4×4 -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

$$wo_{m:n} = \frac{1 - \sum_{r=m}^n p(o_r)}{\sum_{r=m}^n p(o_r)}. \quad (\text{B7})$$

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 4×4 -category joint distribution table the following weights result:

$$s_{33} = \frac{1}{3} (w_{02:4} + w_{03:4} + w_{01:3}) \quad (\text{B8})$$

$$s_{34} = \frac{1}{3} (w_{02:4} + w_{03:4} - 1) \quad (\text{B9})$$

$$s_{32} = \frac{1}{3} (w_{02:4} - 1 + w_{01:3}) \quad (\text{B10})$$

$$s_{31} = \frac{1}{3} (-1 - 1 + w_{01:3}) \quad (\text{B11})$$

Assuming the observations to be equally distributed over all categories results in progressively lower winning odds as the bets become less sharp ($m \ll n$ in eq. B7). 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. B7].

Three requirements have to be fulfilled by the categorization 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. 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, although the dependency of G on the number of classes is weak for the profiles investigated here.

References

- Bohnenstengel, S. I. 2009. Can a simple locality index be used to improve mesoscale model forecasts? Department Geowissenschaften, Universität Hamburg. Ph.D. thesis.
- Brümmer, B. and Höber, H. 1999. A mesoscale cyclone over the Fram Strait and its effects on sea ice. *J. Geophys. Res.-Atmos.* **104**(D16), 19 085–19 098.
- Brümmer, B., Launiainen, J., Müller, G. and Schröder, D. 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., Schröder, D., Müller, G., Spreen, G., Jahnke-Bornemann, A. and co-authors. 2008. Impact of a Fram Strait cyclone on ice edge, drift, divergence, and concentration: possibilities and limits of an observational analysis. *J. Geophys. Res.-Oceans* **113**(C12). doi: 10.1029/2007JC004149.
- Bungert, U. 2008. Einfluss der nestung auf die ergebnisse meteorologischer modelle. Department Geowissenschaften, Universität Hamburg. Ph.D. thesis.
- Davies, H. 1976. Lateral boundary formulation for multilevel prediction models. *Q. J. Roy. Meteorol. Soc.* **102**(432), 405–418.
- Deardorff, J. 1978. Efficient prediction of ground surface-temperature and moisture, with inclusion of a layer of vegetation. *J. Geophys. Res.-Oceans. Atmos.* **83**(NC4), 1889–1903.
- Deng, A. and Stauffer, D. 2006. On improving 4-km mesoscale model simulations. *J. Appl. Meteorol. Climatol.* **45**(3), 361–381.
- Dierer, S. and Schlünzen, K. H. 2005. Influence parameters for a polar mesocyclone development. *Meteorol. Zeitschr.* **14**(6), 781–792. doi: 10.1127/0941-2948/2005/0077.
- Dierer, S., Schlünzen, K. H., Birnbaum, G., Brümmer, B. and Müller, G. 2005. Atmosphere-sea ice interactions during a cyclone passage investigated by using model simulations and measurements. *Monthly Weather Rev.* **133**(12), 3678–3692.
- Gandin, L. and Murphy, A. 1992. Equitable skill scores for categorical forecasts. *Monthly Weather Rev.* **120**(2), 361–370.
- Gerrity, J. 1992. A note on Gandin and Murphy equitable skill score. *Monthly Weather Rev.* **120**(11), 2709–2712.
- Guest, P. and Davidson, K. 1991. The aerodynamic roughness of different types of sea ice. *J. Geophys. Res.-Oceans* **96**(C3), 4709–4721.
- Kaleschke, L., Lüpkes, C., Vihma, T., Haarpaintner, J., Bochert, A. and co-authors. 2001. SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis. *Can. J. Remote Sens.* **27**(5), 526–537.
- Karcher, M., Gerdes, R., Kauker, F. and Köberle, C. 2003. Arctic warming: evolution and spreading of the 1990s warm event in the Nordic seas and the Arctic Ocean. *J. Geophys. Res.-Oceans* **108**(C2). doi: 10.1029/2001JC001265.
- Karcher, M., Gerdes, R., Kauker, F., Köberle, C. and Yashayaev, I. 2005. Arctic Ocean change heralds North Atlantic freshening. *Geophys. Res. Lett.* **32**(21). doi: 10.1029/2005GL023861.
- Lüpkes, C. and Birnbaum, G. 2005. Surface drag in the arctic marginal sea-ice zone: a comparison of different parameterisation concepts. *Bound.-Layer Meteorol.* **117**(2), 179–211. doi: 10.1007/s10546-005-1445-8.
- Lüpkes, C. and Schlünzen, K. H. 1996. Modelling the arctic convective boundary-layer with different turbulence parameterizations. *Bound.-Layer Meteorol.* **79**(1), 107–130. URL: <http://dx.doi.org/10.1007/BF00120077>
- Lüpkes, C., Gryanik, V. M., Witha, B., Gryschka, M., Raasch, S. and co-authors. 2008a. Modeling convection over arctic leads with LES and a non-eddy-resolving microscale model. *J. Geophys. Res.-Oceans* **113**(C9). doi: 10.1029/2007JC004099.
- Lüpkes, C., Vihma, T., Birnbaum, G. and Wacker, U. 2008b. Influence of leads in sea ice on the temperature of the atmospheric boundary layer during polar night. *Geophys. Res. Lett.* **35**(3). doi: 10.1029/2007GL032461.
- Pacanowski, R. C. 1995. Mom 2 Documentation, User's Guide and Reference Manual. Technical Report 3, Princeton University, Princeton, NJ, USA.
- Persson, A. and Grazzini, F. 2005. User Guide to ECMWF Forecast Products. Technical Report Meteorological Bulletin M3.2, European Centre for Medium-Range Weather Forecasts.

- Pielke, R. A. 2002. *Mesoscale Meteorological Modeling* Pielke, R. A. Academic Press, Academic Press, USA.
- Putkonen, J. 1998. Soil thermal properties and heat transfer processes near Ny-Ålesund, northwestern Spitsbergen, Svalbard. *Polar Res.* **17**(2), 165–179.
- Rasmussen, E., Guest, P. and Davidson, K. 1997. Synoptic and mesoscale atmospheric features over the ice-covered portion of the Fram Strait in spring. *J. Geophys. Res.-Atmos.* **102**(D12), 13975–13986. XXIst General Assembly of the International-Union-of-Geodesy-and-Geophysics, BOULDER, CO, JUL 02-14, 1995.
- Ries, H. and Schlünzen, K. H. 2009. Evaluation of a mesoscale model with different surface parameterizations and vertical resolutions for the Bay of Valencia. *Monthly Weather Rev.* **137**(8), 2646–2661. doi: 10.1175/2009MWR2836.1.
- Schlünzen, K. H. 1990. Numerical studies on the inland penetration of sea breeze fronts at a coastline with tidally flooded mudflats. *Beitr. Phys. Atmos.* **63**, 243–256.
- Schlünzen, K. H. and Katzfey, J. 2003. Relevance of sub-grid-scale land-use effects for mesoscale models. *Tellus. Series A, Dynam. Meteorol. Oceanogr.* **55**, 232–246.
- Spreen, G., Kaleschke, L. and Heygster, G. 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res.-Oceans* **113**(C2). doi: 10.1029/2005JC003384.
- Taylor, K. 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.-Atmos.* **106**(D7), 7183–7192.
- Thorndike, A. and Colony, R. 1982. Sea ice motion in response to geostrophic winds. *J. Geophys. Res.-Oceans. Atmos.* **87**(NC8), 5845–5852.
- Valkonen, T., Vihma, T. and Doble, M. 2008. Mesoscale modeling of the atmosphere over Antarctic sea ice: a late-autumn case study. *Monthly Weather Rev.* **136**(4), 1457–1474. doi: 10.1175/2007MWR2242.1.
- Vihma, T., Hartmann, J. and Lüpkes, C. 2003. A case study of an on-ice air flow over the Arctic marginal sea-ice zone. *Bound.-Layer Meteorol.* **107**(1), 189–217.
- von Salzen, K., Claussen, M. and Schlünzen, K. H. 1996. Application of the concept of blending height to the calculation of surface fluxes in a mesoscale model. *Meteorol. Zeitschr.* **5**, 60–66.