Numerical simulation of fog and radiation in complex terrain

Results from COST-722 / 1-DEFOP

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel
von
Mathias D. Müller
aus Sirnach/Horben (TG)
Basel, 2006

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von Prof. Dr. Eberhard Parlow und PD Dr. Mathias Rotach

Basel, den 24 Januar 2006

Prof. Dr. Hans-Jakob Wirz

Acknowledgements

First, I would like to thank Prof. Dr. Eberhard Parlow who initiated this work and gave me the opportunity to work in his motivating team. I am very thankful for the freedom he gave me to follow my scientific interests and ideas while at the same time supporting these activities with an excellent environment. His didactic skills are an inspiration for presenting complicated issues to the scientific community and students.

I am very thankful to Dr. Mathias Rotach for all the interesting discussions and of course for reviewing this thesis. The far reaching scientific support on all aspects of numerical modeling provided by Prof. Dr. Zavisa Janjic from NOAA/NCEP were crucial for this work. For the opportunity to meet the scientific staff of NOAA/NCEP and endless discussions during my two weeks visit in Washington D.C. I am deeply grateful. I also want to thank Gordana Janjic for her hospitality. Matthew Pyle (NOAA/NCEP) and Michael Hagemann (University of Basel) gave great support in the early stages of setting up the Eta and NMM models. I also want to thank NOAA/NCEP for sharing the source code of their operational numerical weather prediction models.

I would like to thank Dr. Roland Vogt for his "all-inclusive" support. Besides his valuable expertise he passed on to me his scientific enthusiasm. He let me participate in many field campaigns, which besides being great social events also opened a modelers eye for the real world meteorology.

I would like to further thank Prof. Dr. Dieter Scherer (TU Berlin) for the successful scientific collaboration on radiation in complex terrain and for his hospitality in Berlin. I am also indebted to Dr. Ismail Gultepe (Environment Canada) for the great collaboration on fog and visibility.

I would like to express my thorough thank to Dr. Christoph Schmutz (MeteoSwiss) for managing the research project, his scientific input and for organizing difficult tasks like installing meteorological sensors right next to the main runway at Zürich Unique airport. Lionel Peyraud (MeteoSwiss) provided very valuable input from an operational forecasting point of view regarding practicality and user requirements of fog forecasts. Furthermore I am thankful to Simon Rohling, Emanuele Zala and Marco Arpagaus from MeteoSwiss for setting up the framework to exchange operational forecasts and observational data between MeteoSwiss and our University.

I am grateful to Météo-France and especially Dr. Thierry Bergot (Météo-France / CNRM) for the source code of an early version of COBEL and all the related support.

Prof. Dr. Andreas Bott and Matthieu Masbou (University of Bonn) invited me to start the development of a 3D fog model. For their hospitality and the excellent scientific collaboration I am deeply grateful. I also appreciate the fact that they shared the source code of the PAFOG 1D model. Jan Cermak from the University of Marburg generously provided satellite data needed for forecast verification. For the help in developing the variational data assimilation I would like to thank Dr. Ross Bannister from the Data Assimilation Research Centre in Reading / UK.

I would like to further thank my colleagues working at the Institute of Meteorology, Climatology and Remote Sensing at the University of Basel for their support, assistance and contributions to the present work: Dr. Andreas Christen for endless discussions and for sharing an apartment with a "mad scientist", Irene Lehner for being such a great friend and fueling my brain cells with sweets, Gergely Rigo for the help with satellite data, Jan Eitel for the vivid time in our office, Thomas Kleiber for his surveys of competitive research, Günter Bing and Josette Pefferli-Stocky for an excellent computer and administrative support, respectively.

I am indebted to our university's computing center (URZ) for providing ten thousands of CPU hours on the Beowulf cluster and especially to the support of Martin Jacquot, who always found a solution for my very special demands and complicated setups.

This research was funded by the Swiss Federal Office for Education and Science (Grant C03.0024).

Finally I would like to thoroughly thank and dedicate this work to my parents who enabled that I could always follow my interests.

Abstract

Two high resolution numerical 1D models, namely COBEL and PAFOG, have been adapted to compute a probabilistic fog forecast. Major modifications were made to the COBEL model. It was coupled to the NOAH land surface model to take into account the effects of soil and vegetation and furthermore a parameterization of precipitation was added. To deal with the large uncertainty inherent to fog forecasts, a whole ensemble of 1D runs is computed using the two different numerical models and a set of different initial conditions in combination with distinct boundary conditions. Initial conditions are obtained from variational data assimilation, which optimally combines observations with a first guess taken from operational 3D models. The design of the ensemble scheme computes members that should fairly well represent the uncertainty of the current meteorological regime. Verification reveals that the probabilistic forecast can significantly improve the current methods used at Zürich Unique airport.

The complex topography in Switzerland further complicates fog forecasting. In order to simulate processes like advection, cold air drainage flows and cold air pooling, the NMM 3D model of NOAA/NCEP is modified and extended with detailed fog microphysics. The resulting 3D fog model runs at a horizontal resolution of 1 km and a vertical resolution comparable to the 1D models. First results look very promising and are able to reproduce the spatial distribution of fog as it is seen by satellite.

With increasing horizontal resolution of numerical weather prediction models, topographical effects on radiation gain importance. With a newly developed parameterization it is possible to consider slope angle, aspect angle, shadows and restricted sky view on the subgrid scale and with negligible computational costs. Verification reveals that RMS and mean error of 2 m temperature forecasts are generally improved by 0.5 to 1 K.

Contents

List of Figures

List of Tables

List of symbols and abbreviations

Symbol	Variable	Unit
A_z	area under ROC curve	
$\mathbf{B}$	covariance matrix of background errors	
C	condensation rate by air mass unit	$\mathrm{kgkg^{-1}s^{-1}}$
c_p	specific heat of air	$\mathrm{Jkg^{-3}K^{-1}}$
D	droplet diameter	m
$D_{c,0}$	mean droplet diameter	m
D_v	water vapor diffusivity	$\mathrm{m^2s^{-1}}$
$\mathbf{E}$	Eigenvector matrix	
${}^{\uparrow}E_l$	upwelling long-wave radiation	$\mathrm{Wm^{-2}}$
${}^{\downarrow}E_l$	downwelling long-wave radiation	$\mathrm{Wm^{-2}}$
${}^{\uparrow}E_s$	upwelling short-wave radiation	$\mathrm{Wm^{-2}}$
${}^{\downarrow}E_{s,diff}$	diffuse downwelling short-wave radiation	$\mathrm{Wm^{-2}}$
${}^{\downarrow}E_{s,dir}$	downwelling short-wave radiation	$\mathrm{Wm^{-2}}$
e_s	saturation vapor pressure over a plane water surface	hPa
E_t	turbulent kinetic energy	
f	Coriolis parameter	$\mathrm{s^{-1}}$
f_{cor}	short-wave radiation correction factor in complex terrain	$fraction$
F_r	net radiative flux	$\mathrm{Wm^{-2}}$
$\mathbf{H}$	linearized observation operator	
$\mathbf{I}$	Identity matrix	
K	thermal conductivity	$\mathrm{Jm^{-1}s^{-1}K^{-1}}$
K_z	exchange coefficient for vertical turbulent diffusion	
k	vertical index used in discretization	
L	latent heat of evaporation	$\mathrm{Jkg^{-1}}$
N_c	total droplet number concentration	$\mathrm{m^{-3}}$
$\mathbf{P}$	scaling matrix	
q	mixing ratio	$\mathrm{kgkg^{-1}}$
q_c	specific cloud water content	$\mathrm{kgkg^{-1}}$
q_l	liquid water content	$\mathrm{kgkg^{-1}}$
q_s	saturation mixing ratio	$\mathrm{kgkg^{-1}}$
$\mathbf{R}$	covariance matrix of observation errors	
R_e	Reynolds number	
Ri	Richardson number	
R_w	specific gas constant for moist air	$\mathrm{Jkg^{-1}K^{-1}}$
$\bar{S}$	mean supersaturation	$fraction$
T	air temperature	K
$\mathbf{T}$	T-transform matrix	
T_s	skin temperature	K
$\mathbf{U}$	U-transform matrix	
u	east-west component of wind vector	$\mathrm{ms^{-1}}$
u_g	east-west component of geostrophic wind vector	$\mathrm{ms^{-1}}$
v	north-south component of wind vector	$\mathrm{ms^{-1}}$
v_g	north-south component of geostrophic wind vector	$\mathrm{ms^{-1}}$
w	vertical component of wind vector	$\mathrm{ms^{-1}}$
$\vec{x}$	model state vector	
$\vec{x'}$	perturbation state vector for the model	
$\vec{x_b}$	background model state vector	
$\vec{x_f}$	forecast state vector	
$\vec{x_t}$	true model state vector	
$\vec{y}$	vector of observations	

Symbol	Variable	Unit
$\vec{y'}$	perturbation state vector for the observations	
z_0	roughness length	m
α	angle between the unit vector of the solar direct beam and the normal vector of the surface	$^\circ$
ϵ	emissivity	$fraction$
θ	potential temperature	K
θ_N	slope angle	$^\circ$
θ_S	sun elevation angle	$^\circ$
κ	dissipation rate of turbulent kinetic energy	m^2s^{-3}
λ	Eigenvalues	
ρ_w	density of water	kgm^{-3}
ρ	air density	kgm^{-3}
σ_c	dispersion parameter of droplet size distribution	$fraction$
ϕ_N	slope aspect angle	$^\circ$
ϕ_S	sun azimuth angle	$^\circ$
aLMo	Alpine Model of MeteoSwiss	
COBEL	Couche Brouillard Eau Liquide 1D fog model	
COBEL-NOAH	COBEL coupled to land surface model	
COST	European cooperation in the field of scientific and technical research	
DEM	digital elevation model	
ECMWF	European Centre for Medium-Range Weather Forecasts	
F	false alarm rate	
FAR	false alarm ratio	
GFS	Global Forecast System	
H	hit rate	
HSS	Heidke skill score	
LST	local standard time	
MIFOG	MIcrophysical FOG 1D model	
NCEP	National Centers for Environmental Prediction	
NDVI	normalized difference vegetation index	
NMM-22	Nonhydrostatic Mesoscale Model (22 km resolution)	
NMM-4	Nonhydrostatic Mesoscale Model (4 km resolution)	
NMM-2	Nonhydrostatic Mesoscale Model (2 km resolution)	
NOAA	National Oceanic and Atmospheric Administration	
MESO-ETA	Mesoscale Eta Model	
MeteoSwiss	Swiss Federal Office of Meteorology and Climatology	
MSG	Meteosat Second Generation	
MTP-5	Microwave Temperature Profiler	
PAFOG	PArameterized FOG 1D model	
PC	proportion of correct forecasts	
RMS	root mean square	
ROC	relative operating characteristic	
UTC	coordinated universal time	

1 Introduction

Computing power has now reached a performance level that allows for high resolution numerical weather prediction. Today's forecast models are run at a grid spacing of a few kilometers and are thus capable to resolve mesoscale phenomena and produce weather forecast with incredible spatial detail. Over the last few years I developed a high resolution semi-operational forecast system for Europe and the Alpine region in particular. In cooperation with NOAA/NCEP, the Nonhydrostatic Mesoscale Model (NMM) has been adapted to and tested in the complex topography of the Alps. Every 12 hours weather predictions up to three days are computed and the results are made available on the internet. These are indeed the highest resolution forecasts available. The modeling work on the small scales has caught my fascination and it has become a personal goal to provide and further improve high resolution weather forecasting. In this study I try to address two common problems related to this high resolution numerical weather prediction, namely fog and radiation in complex terrain.

For the safe and economically efficient flight operation the so called aeronautical users, namely air traffic management authorities or airlines depend heavily on the accurate forecasts of fog and visibility. In Switzerland, the major airports are prone to fog development and low visibility conditions from approximately September to March. Due to the central air traffic flow management in Europe non-accurate visibility predictions for an airport result in too low or too high frequencies compared to the actual capacities on the airports. The delays caused by an inaccurate visibility prediction are substantial and costly. In the year 2003, the airline Swiss assessed the impact of such a single event of one hour reduced traffic due to an inaccurate fog prediction to approximately 50000 Euro (Christoph Schmutz, personal communication 2005).

The need for a reliable fog forecast has lead to the development of various forecasting and nowcasting methods based on numerical forecast models and statistics. The most common statistical methods are based on regression, neural networks or decision trees to forecast fog, visibility and low clouds. Statistical methods are typically driven by an atmospheric profile, which is almost always taken from an operational 3D forecast model, since radiosonde data are not available at most airports. The forecast variable is then related either to this profile or to both the profile and the latest surface observation. Apart from the input data from each forecast, the key requirement is for historical data to establish the statistical relationship. It is crucial that the characteristics of the forecast do not change once the statistical relationship is established, a requirement difficult to fulfill with always improving 3D forecast models. Most systems appear to have used about three years of historical data. However, this is likely to be too short to adequately calibrate predictions of dense fog which occurs only rarely in most locations. In Austria a Model Output Statistics (MOS) system was developed based on ten years of historical data (Golding 2002). Such long data sets seem necessary for statistical regression and neural networks. Historical data from a 3D model can be avoided if forecast variables are related only to variables observed at the surface and if it is further assumed that 3D model forecasts are representative for the observations. According to a survey conducted by the COST-722 action (Golding 2002), these considerations relate to the statistical systems used in Austria, Finland, France, Germany, Italy, Netherlands and Spain. More complex systems use a multi-step process, for instance the profile is first corrected and then a Perfect-Prog method estimates visibility. For Expert Systems, Decision Tree or Score Table approaches, the requirement for historical data is likely to be much greater, since each step in the process must be calibrated. Another problem with statistical methods is the difficulty to transfer the algorithm to another location. Besides the necessary historical data, often local knowledge was used in the development and many adaptations are necessary. Because of such deficiencies inherent to statistical

methods, it seems useful to base a fog forecast model on physical principles.

In the framework of COST-722, I developed a 1D ensemble forecast system for fog. Even though 1D models only lead to point forecasts and cannot take advection into account, they have the advantage of a relatively short running time and that they can represent some processes in better detail because of a very high vertical resolution and possibly more sophisticated microphysical parameterizations. A 3D model, which considers spatial interactions in the form of advection could be run for all weather situations. However, highly sophisticated 3D models may have a rather long running time, so that especially for operational purposes a balance has to be struck between running time requirements and the accuracy of representation of physical processes. Ensemble forecasting is a suitable alternative to deterministic model forecasts. With this technique, predictions from an ensemble of slightly different initial conditions and/or various versions of models are computed to improve the accuracy of the forecast through averaging the various forecasts, which eliminates non-predictable components. The ensemble also provides reliable information on forecast uncertainties from the diversity amongst ensemble members. Unfortunately 3D probabilistic forecast exceed current computational limits. In the framework of this project, 3D deterministic forecasts at a horizontal resolution of 2 km, using the Nonyhdrostatic Mesocale Model (NMM), were carried out in order to evaluate the fog forecasting potential. The rather discouraging outcome lead to the development of a more detailed cloud microphysics module for NMM and a 1D fog forecasting system. The latter was developed for operational purpose and it has been running in such context since November 2004, producing daily forecasts for Zürich Unique airport.

Outline

Two state of the art 1D models were chosen for the fog forecast system. They are described with all the modifications made in Chapter 2. The model was then tested in a comparison project (Chapter 3). Accurate specification of initial and boundary conditions is a crucial point of the fog forecast process (Chapter 4). This expresses the need for profile measurements of temperature and humidity. But launching radiosondes at an airport is a rather impossible undertake for obvious flight safety reasons. To address the problem, a 1D data assimilation system was developed, which optimally combines all available measurements with a first guess obtained from different mesoscale models (Chapter 5). The latter are also used to include effects of spatial heterogeneity into the 1D model. In practice this is like a one way nesting of the 1D column into the 3D model. Unfortunately fog forecasting is a threshold problem where very small errors in predicted temperature or humidity might decide whether condensation takes place or not. To deal with such uncertainty, an ensemble prediction system was programmed to produce probabilistic fog forecasts for Zürich Unique airport. The verification results, summarized in Chapter 6, indicate that inclusion of advection is the key point. Only by doing so, skillful forecasts are possible.

Chapter 7 is devoted to the interaction of radiation with complex terrain. So far this was not considered in numerical weather prediction. In fact, radiation was assumed to be absorbed by a horizontal surface, thus neglecting the effects of slope and aspect angle, limited sky view and shadows cast by the topography. A new parameterization for these effects significantly modifies the available energy and improves the temperature forecast.

The last Chapter is an outlook to the future of fog forecasting. Unsatisfying results with the high resolution semi operational 3D model reveal that fog forecasting is not just a matter of resolution. More detailed microphysics are introduced into the 3D model. The computation of condensation/evaporation as well as the sedimentation of cloud water can be modeled more accurately when a prognostic equation for the total droplet number concentration is solved. The 3D framework allows to explicitly model

cold air outflow and ponding and solves the problems of advection inherent to the 1D model simulations. This modified 3D model produced very promising first results.

2 1D numerical models

The formation and dispersion of fog is the result of complex interaction between thermodynamic and dynamical processes. Duynkerke (1990) identified the most important factors for fog formation to be:

- cooling of moist air by radiative flux divergence
- mixing of heat and moisture
- vegetation
- horizontal and vertical wind
- heat and moisture transport in soil
- advection
- topographic effects,

where atmospheric conditions, location and season decide upon the relative importance of each factor. The presence of clouds increases the incoming longwave radiation at ground level and thus reduces the longwave radiative cooling at the surface, which has great influence on fog formation. Therefore a good cloud forecast, computed by a 3D model, is also needed. In complex topography cold air outflow and pooling as well as advection in the heterogenous landscape become very important. Once the fog has formed there are further influences:

- longwave radiative cooling at fog top
- gravitational droplet settling
- fog microphysics
- shortwave radiation

Starting with the work of Zdunkowski & Nielsen (1969) some of the above listed processes were included in newly developed fog models. In this early model there was no parameterization for the sedimentation of liquid water nor exchange coefficients. The latter were introduced by Zdunkowski & Barr (1972). An even more sophisticated model was developed by Brown & Roach (1976) and further refined by Turton & Brown (1987) including new formulations for exchange coefficients of the nocturnal boundary layer. A similar model was also used by Musson-Genon (1987) for his quantitative comparison between computed and observed fog evolution. Very detailed microphysics was introduced by Brown (1980) and further refined in a new model by Bott et al. (1990), who also introduced a sophisticated treatment of radiation. Siebert et al. (1992*a*), Siebert et al. (1992*b*) and von Glasow & Bott (1999) finally added a module to resolve small and tall vegetation on a high resolution grid. In these later models, the evolution of the droplet size distribution and cloud condensation nuclei is explicitly resolved, but even today such an approach is computationally very expensive.

For the 1D approach of this study two numerical models without spectral microphysics are used to compute an ensemble forecast. The COBEL model was significantly modified and coupled to the NOAH land surface model. The second model, called PAFOG was taken as it is and only the input/output handling was changed so that it could be integrated into the ensemble system. A short description of the models and modifications is now given.

2.1 Brief description of COBEL

The COBEL model (Couche Brouillard Eau Liquide) was originally derived from the 1D model of the nocturnal boundary layer, developed by the Laboratoire d'Aérologie of the Toulouse University (Estournel 1988). It has been used to predict fog events over the past years at Paris Charles de Gaulle airport (Bergot et al. 2005), a site in very flat topography. The governing equations are:

$$\frac{\partial u}{\partial t} = f(v - v_g) - \frac{\partial}{\partial z}\overline{w'u'} \tag{1}$$

$$\frac{\partial v}{\partial t} = -f(u - u_g) - \frac{\partial}{\partial z}\overline{w'v'} \tag{2}$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\overline{w'\theta'} + \frac{\theta}{T}\left(\frac{1}{\rho c_p}\frac{\partial F_r}{\partial z} + \frac{L}{c_p}C\right) + \text{ Adv } \theta \tag{3}$$

$$\frac{\partial q}{\partial t} = -\frac{\partial}{\partial z}\overline{w'q'} - C + \text{ Adv } q \tag{4}$$

where ρ is air density, u and v are the horizontal wind components, T the temperature, θ the potential temperature, q the mixing ratio and f the Coriolis parameter. Turbulent vertical fluxes of α are denoted $\overline{w'\alpha'}$, F_r is the net radiative flux, C the condensation rate by airmass unit, c_p the specific heat of air at constant pressure, L the latent heat of evaporation. Equations (1-4) support external forcing terms for the geostrophic wind (u_g, v_g) and horizontal temperature advection (Advθ) as well as humidity advection (Adv$q$) which is present in a baroclinic atmosphere.

The most important point of a 1D fog model besides the cloud microphysics is the turbulence closure. The main problem here is that it has to cope with stable or very stable environments. In COBEL, a 1.5 order closure with the following prognostic equation for turbulent kinetic energy E_t is used:

$$\frac{\partial E_t}{\partial t} = -\frac{\partial}{\partial z}\overline{w'E_t'} - \overline{u'w'}\frac{\partial u}{\partial z} - \overline{v'w'}\frac{\partial v}{\partial z} + \frac{g}{T}\overline{w'\theta'} - \kappa \tag{5}$$

where g is gravitational acceleration and κ the dissipation rate of kinetic energy. The turbulent fluxes are related to the vertical local gradients with the help of an exchange coefficient:

$$\overline{w'\alpha'} = -K_\alpha \frac{\partial \alpha}{\partial z} \quad \text{for} \quad \alpha = u, v, \theta, q \tag{6}$$

The main problem is now the specification of the turbulent diffusion coefficients K_α, which are given in terms of local stability dependent mixing length l_α and turbulent kinetic energy E_t:

$$K_\alpha = 0.4 l_\alpha E_t^{1/2} \tag{7}$$

Note that there is a nice prognostic equation for E_t but the mixing length needed in [7] requires parameterizations, and this is the main weakness of the turbulence closure. Depending on the stability a different formulation for mixing length is used. For stable stratification the relation of Estournel & Guedalia (1987) is applied:

$$l_\alpha = l_n(1 - 5 \text{ Ri }), \text{ Ri } \leq 0.16 \tag{8}$$

$$l_\alpha = l_n(1 + 41 \text{ Ri })^{-0.84}, \text{ Ri } > 0.16 \tag{9}$$

where Ri is the Richardson number and l_n is the neutral mixing length according to Delage (1974):

$$l_n = \frac{kz}{1 + (kz/G_n)} \tag{10}$$

with the von Kármán constant k and $G_n = 4 \cdot 10^{-4} u_g f^{-1}$

For unstable layers, mixing length is computed according to Bougeault & André (1986). Here the distance a parcel can ascend (L_{up}) and descend (L_{down}) against buoyancy using its kinetic energy is computed first:

$$\int_{z}^{z+L_{up}} \beta[\theta(z') - \theta(z)]dz' = E_t(z) \tag{11}$$

$$\int_{z-L_{down}}^{z} \beta[\theta(z') - \theta(z)]dz' = E_t(z) \tag{12}$$

where β is the buoyancy coefficient g/T. The mixing length can now be assigned:

$$l_\alpha = \min\left(L_{down}, L_{up}\right) \tag{13}$$

In the absence of liquid water, stability is simply diagnosed using the virtual potential temperature gradient. However when liquid water is present an expression derived from the formulation of the Brunt-Väisälä frequency after Durran & Klemp (1982) is used:

$$\left[\frac{T}{\theta}\frac{\partial\theta}{\partial z} + \Gamma_m - \frac{g}{c_p}\right]\left[1 + \frac{L_{q,sat}(T)}{RT}\right] - \frac{T}{1+q_w}\frac{\partial q_w}{\partial z} \begin{pmatrix} <0 \rightarrow \text{ unstable} \\ =0 \rightarrow \text{ neutral} \\ >0 \rightarrow \text{ stable} \end{pmatrix} \tag{14}$$

where Γ_m is the moist adiabatic lapse rate.

The microphysics are treated rather simple with an implicit bulk-type parameterization. If the mixing ratio q exceeds the saturation value q_{sat}, the excess condenses as liquid water q_l. The equilibrium mixing ratio q^* and temperature T^* are calculated iteratively conserving the total energy and total water content:

$$\begin{aligned} c_pT + Lq &= c_pT^* + Lq^* \\ q^* &= q_{sat}(T^*) \end{aligned} \tag{15}$$

Basically if water condenses, it increases the temperature and thus the saturation mixing ratio, so that the slightly warmer air can maintain a bit more water vapor, which results in evaporation and cooling of the air. The cycle repeats until equilibrium is reached. Since no information about the cloud droplets is available, the determination of the gravitational settling flux G is rather difficult and a simple parameterization of Brown & Roach (1976) has to be used:

$$G = v_i q_l \tag{16}$$

The settling velocity v_i is kept constant at a value of 0.016 m s^{-1} according to the case study of Bergot & Gudalia (1994*a*).

Using the microphysics [15,16] for condensation/evaporation (C) and the gravitational settling (G) the prognostic equation for liquid water is defined:

$$\frac{\partial q_l}{\partial t} = \frac{\partial}{\partial z}\left(K_z\frac{\partial q_l}{\partial z}\right) + \frac{\partial G}{\partial z} + C - P \tag{17}$$

where P is autoconversion of cloud water into precipitation, which is described in the next section. Finally visibility estimates are obtained using the relation of Kunkel (1984):

$$\text{vis} = \frac{3.9}{144.7(\rho q_l)^{0.88}} \tag{18}$$

The computationally most expensive part of the model is the radiation code. Net shortwave radiation is computed at each vertical level using a version of the parameterization presented by Fouquart & Bonnel (1984). It is a monospectral model

between 0.24 and 4 μm which takes into account the absorption by water vapor, carbon dioxide, ozone and cloud droplets. For the longwave radiation (4-100 μm) the scheme of Vehil et al. (1989) is used. Computations are done for 232 spectral ranges at every vertical level by taking into account the emission and absorption by water (liquid and vapor), and CO_2, as well as by the Earth's surface.

The remaining parts of the model have been changed and are presented below in more detail.

2.2 COBEL model modifications

2.2.1 Parameterization of precipitation

Many simulations showed an increase in liquid water content of stratus clouds beyond values of 0.6 g m^{-3}. In fact, the stratus clouds sometimes became completely impermeable for shortwave radiation even at noon. In order to remove this excess of liquid water from the clouds, a simple but robust and frequently used parameterization for precipitation was introduced into the model. It was developed by Kessler (1969) and considers autoconversion, accretion, evaporation of rain and computations of mean fall speeds. In COBEL, the rainwater is allowed to evaporate if a grid box is not saturated and thus can alter the temperature and humidity. Rainwater descends through the column based on the computed fall speed in a grid box and finally enters the land surface scheme. There is no direct interaction of rainwater and radiation. Humidity q, cloud water content q_c and rainwater q_r, form the following system of equations:

$$\begin{aligned} \frac{dq}{dt} &= Q_{revp} - Q_{cnevp} \\ \frac{dq_c}{dt} &= Q_{cnevp} - Q_{aut} - Q_{acc} \\ \frac{dq_r}{dt} &= Q_{acc} + Q_{aut} - Q_{revp} + Q_{sed} \end{aligned} \tag{19}$$

where Q_{revp} is the rate of rainwater evaporation, Q_{cnevp} the rate of condensation ($Q_{cnevp} > 0$) or evaporation between q and q_c, Q_{aut} the autoconversion rate, Q_{acc} the accretion rate and Q_{sed} the rate of change due to sedimentation fluxes into and out of a gird cell. Note that all water contents are volumetric.

The parameterization assumes that the droplet size distribution is of the Marschall-Palmer type:

$$n(D) = N_0 e^{-\Lambda D} \tag{20}$$

where D is the droplet diameter, N_0 is an empirically determined constant, taken to be equal to 0.8 10^7 m^{-4} and Λ is a function of the total amount of rain water (Houze (1993) p.104). The rates of autoconversion and collection of cloud droplets by precipitation are:

$$Q_{aut} = k_1(q_c - a) \tag{21}$$

$$Q_{acc} = k_2' q_c q_r^{7/8} e^c \tag{22}$$

$$c = \frac{g}{287T}\frac{z}{2}\ , \quad k_2' = k_2 E N_0^{1/8}$$

where z is the height above ground, k_1 is the autoconversion rate or reciprocal of "conversion time" of cloud water and a the autoconversion threshold ($a = 0.5$ g m^{-3}) below which no precipitation forms. Following Ghosh & Jonas (1998) a representative value for continental clouds of $k_1 = 5.46\ 10^{-4}$s^{-1} is used. Note that the autoconversion rate, is linearly related to the cloud liquid water content. The accretion rate needs specification of constants $k_2 = 6.96\ 10^{-4}$ and the collision-coalescence efficiency E, which is set to be less than unity ($E = 0.9$), since the clouds consist of

small droplets under low stratus and fog conditions. If the air in a grid cell is not saturated, evaporation reduces the precipitation water content:

$$Q_{revp} = k_3 N_0^{7/20}(q_{sat} - q)q_r^{13/20} \tag{23}$$

where the subscript sat denotes saturation and $k_3 = 1.93\ 10^{-6}$. Finally the rainwater settles with terminal velocity v_t:

$$v_t = 38.8 N_0^{-1/8} q_r^{1/8} e^c \tag{24}$$

The process of accretion is faster than autoconversion, thus cloud water depletion due to accretion is done before autoconversion. The divergence of the precipitation sedimentation flux is discretized using an upstream formulation:

$$q_r^{t+1}(k) = q_r^t(k) - \frac{\Delta t_{sed}}{\Delta z}(-v_t(k+1)q_r^t(k+1) + v_t(k)q_r^t(k)) \tag{25}$$

Due to high vertical resolution in COBEL, special care has to be taken when solving (24) in order to not violate the Courant condition. Thus, if the fastest droplet moves more than half the grid spacing, the fundamental time step is split into smaller sedimentation time steps Δt_{sed}.

2.2.2 Coupling COBEL with the NOAH Land Surface Model

The initial version of COBEL used a simple treatment of surface-atmosphere exchanges and was developed for bare soil only. A detailed description of the old scheme is given by Bergot & Gudalia (1994*a*), but computed sensible and latent heat fluxes were often of unrealistic magnitude and numerically unstable. The new approach uses the NOAH Land Surface Model (LSM) to compute the energy balance of the canopy layer and the evolution of temperature and humidity in the soil. The NOAH LSM has a long heritage and originated from Mahrt & Pan (1984), Mahrt & Ek (1984) and Mahrt & Mahrt (1987). Since then several major improvement were done (Chen et al. 1997, Ek et al. 2003) and it is currently used in the NCEP realtime Land Data Assimilation System. The NOAH LSM is a multi-layer soil model with an explicit seasonal cycle of vegetation and sophisticated treatment of snowpack physics. The prognostic equations are for soil moisture and temperature in each soil layer, skin temperature, snowpack water, and canopy water. A single skin temperature is obtained from a single surface energy balance equation for the combined ground/vegetation/snowpack surface. The evaporation in the energy balance is a combination of direct evaporation from bare soil, transpiration through plant canopy and canopy water evaporation. The weighting of these three evaporation components is controlled by a specified monthly annual cycle of NDVI-based green vegetation fraction. The vegetation resistance treatment is similar to that used in in the French ISBA model (Noilhan & Mahfouf 1996). The soil hydraulics is governed by Darcy's Law, including both a hydraulic conductivity and diffusivity, which vary with soil moisture content.

Coupling the soil to the atmosphere is a rather difficult task and the strategy chosen is a flux coupling. Therefore LSM fluxes are transferred as lower boundary conditions into the vertical turbulent diffusion scheme of COBEL. The model dynamics, shortwave radiation and the LSM use the same time-step of 10 seconds, whereas the computationally expensive longwave radiation scheme is called every 10 minutes. Effective canopy skin temperature T_s is used in the computation of longwave emission. Since the radiative scheme is called less frequently than the dynamics and the LSM, special care has to be taken in order to conserve energy. To achieve this, skin temperature is averaged over intermediate time-steps so that the mean flux emitted from the surface, as calculated by the LSM is transferred to the radiative scheme. The averaging procedure was proposed by Polcher et al. (1998) and is given by equations (26) and (27):

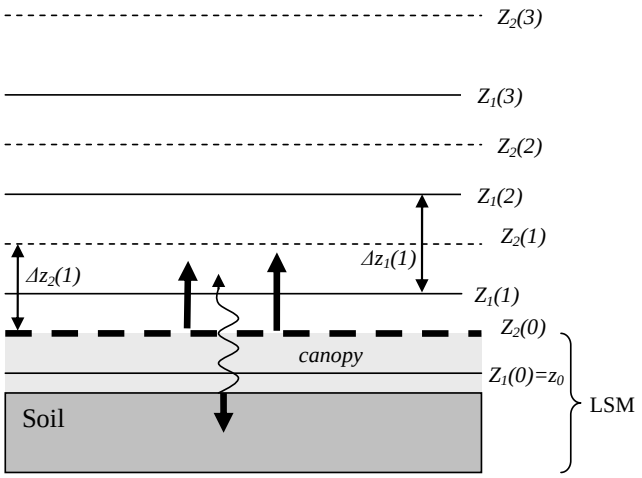

Figure 1: Vertically staggered atmospheric grid of COBEL. Temperature, humidity, wind and cloud water are defined on the Z_1 grid, fluxes are defined on the Z_2 grid. At the lower flux boundary $Z_2(0)$, sensible and latent heat fluxes computed within the LSM are introduced into the atmosphere. At z_0 wind speed is set to 0. Model level $Z_1(1)$ is used to force the LSM.

$$\overline{\epsilon}^{n+1} = \frac{n\overline{\epsilon}^n + \epsilon}{n+1} \tag{26}$$

$$\overline{T_s^4}^{n+1} = \frac{1}{n+1}(n\frac{\overline{\epsilon}^n}{\overline{\epsilon}^{n+1}}\overline{T_s^4}^n + \frac{\epsilon}{\overline{\epsilon}^{n+1}}T_s^4) \tag{27}$$

where ϵ is the emissivity, n is the time-step at which radiative temperature from the LSM is available between two calls to the radiation scheme and $\overline{\epsilon}^n$ is the mean value of emissivity from n time steps. Note that the averaging is linear for ϵ and in the 4^{th} power of T_s, which is basically computing a mean flux.

Figure 1 illustrates the vertical grid and level definitions, relevant for the surface-atmosphere coupling. A crucial part of the coupling is the computation of exchange coefficients for the canopy layer. They are calculated inside the LSM using wind speed, temperature and humidity of the first COBEL layer $Z_1(1)$ as external parameters. Since parameters are from the previous time step, the method realizes an explicit time scheme. It would also be possible to drive the LSM with exchange coefficients obtained from the vertical diffusion parameterization of COBEL. Both methods have their advantages and problems. For the first method, a flux discontinuity could arise which unrealistically accumulates or dissipates heat and moisture in the lowest level. However, the LSM has full control over the exchange coefficients and iterations with recomputations of stability functions, needed for fast convergence of the surface energy balance, are possible. Furthermore surface heterogeneities can be considered easily, since a single grid box atmospheric forcing translates into different fluxes determined by surface properties. Technically, for each surface type the LSM has to compute the fluxes and T_s and area weighted average values are fed back to COBEL. Mathematically the coupling is achieved by altering the coefficients $A(1)$, $B(1)$ and $C(1)$ of the tri-diagonal matrix ψ used in the implicit treatment of vertical diffusion. $A(1)$ is set to zero and the turbulent fluxes are introduced as source term into $Y(1)$. This flux formulation also establishes a convenient framework to test the atmospheric part of the model with fluxes obtained from field measurements. Thus avoiding model

deficiencies related to canopy layer processes if the model is run for past cases. Note that in a 1-D model like COBEL, the only resolved transport process is turbulent diffusion, which is governed by the turbulence closure scheme used to compute K. Horizontal and vertical advection enter the model dynamics as external forcings, and have to be computed by a driving 3-D model. From a theoretical point of view it seems rather inconsistent to include these advections in a 1-D model, as the model's fundamental assumption is a barotropic homogenous environment, where no gradients can develop in the first place. Thus if advections are observed at the fog forecasting site, the need for a 3-D fog model becomes obvious.

To illustrate the flux coupling, the turbulent diffusion equation for temperature is used as a surrogate for all scalar variables. On the staggered grid having k layers, the discretization yields:

$$\frac{T^{t+1}(k) - T^t(k)}{\Delta t} = \frac{K_z(k)\frac{T^{t+1}(k+1) - T^{t+1}(k)}{\Delta z_1(k)} - K_z(k-1)\frac{T^{t+1}(k) - T^{t+1}(k-1)}{\Delta z_1(k-1)}}{\Delta z_2(k)} \tag{28}$$

The turbulent diffusion coefficients K_z are computed using a 1.5 order scheme and thus are functions of stability dependent mixing lengths l and the turbulent kinetic energy. Under the assumption that the sensible heat flux F_{sh} computed by the LSM varies only little in the time interval Δt, equation (28) for the first atmospheric level can be written as (29):

$$\frac{T^{t+1}(1) - T^t(1)}{\Delta t} = \frac{K_z(1)\frac{T^{t+1}(2) - T^{t+1}(1)}{\Delta z_1(1)} - F_{sh}^t}{\Delta z_2(1)} \tag{29}$$

which is equivalent to (30):

$$\underbrace{T^t(1) - \frac{\Delta t}{\Delta z_2(1)} F_{sh}^t}_{Y(1)} = \underbrace{\left(-\frac{\Delta t}{\Delta z_1(1)\Delta z_2(1)} K_z(1)\right)}_{C(1)} T^{t+1}(2) + \underbrace{\left(1 + \frac{\Delta t}{\Delta z_1(1)\Delta z_2(1)} K_z(1)\right)}_{B(1)} T^{t+1}(1) \tag{30}$$

where $C(1)$,$B(1)$ and $A(1) = 0$ are coefficients of the tri-diagonal matrix ψ, and the following system has to be solved for T^{t+1}:

$$\begin{pmatrix} \vdots \\ Y^t(k+1) \\ Y^t(k) \\ Y^t(k-1) \\ \vdots \end{pmatrix} = \psi \begin{pmatrix} \vdots \\ T^{t+1}(k+1) \\ T^{t+1}(k) \\ T^{t+1}(k-1) \\ \vdots \end{pmatrix} \tag{31}$$

$$\psi = \begin{pmatrix} \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & & & \\ & C(k) & B(k) & A(k) & & \\ & & \cdot & \cdot & \cdot & \\ & & & C(1) & B(1) & \end{pmatrix}$$

where matrix Inversion of ψ is carried out using the LU-decomposition algorithm of Press et al. (1998)

Now this works on paper, but how about the implementation? It turns out that the high resolution in the atmosphere, the first layer being only 0.5 m thick, requires an extremely fine resolution in the soil. Indeed six layers are packed into the uppermost 10 cm, starting with a thickness of only 1.5 mm. By doing so it is guaranteed that

the soil responds quickly to any atmospheric forcing, thus keeping the temperature gradient between the first soil and atmospheric layer relatively small. In Figure 2 it can be seen how good the modeled skin temperature fits the surface brightness temperature computed from the observed longwave emission. It has to be noted that this was a very difficult situation with weak winds of about 0.5 m s^{-1} and thus very little turbulence. If the first soil layer is 10 cm thick in this case, a common resolution in numerical weather prediction models, a strong cold bias of over 4 K develops in 2 m air temperature. If we look again at Figure 2 we notice a very strong temperature gradient in the first 8 cm of soil, which can only be resolved at high resolution. Furthermore at coarse resolution the skin temperature and the temperature of the first soil layer can be very different, which causes problems in the computation of the soil heat flux. Figure 2 also illustrates the effect of fog formation on soil temperatures, which start to rise after formation. This reduces the temperature gradient since absorption of longwave emission from the fog at the surface turns the former sink of energy into a source. Unfortunately fog onset was predicted two hours too early, explaining the offset in brightness temperature. The observation taken at 5 cm does not fit very well to the model, but one has to keep in mind that on the rather rough and vegetated soil surface the actual depth of the uppermost observation remains a mystery.

2.3 Brief description of PAFOG

The 1D model for PArameterized FOG (PAFOG) was derived from the detailed spectral microphysical model MIFOG (Bott et al. 1990, 1989). PAFOG consists of four modules, namely the dynamic module, the microphysical module, the radiation code, and a module for low vegetation. The main difference to COBEL is the parameterized but still very detailed cloud microphysics module that allows to compute the total droplet number concentration.

Since no major changes to the 1D model PAFOG were done, an up to date description can be found in Bott & Trautmann (2002) and references mentioned therein. Details on the parameterized cloud microphysics are discussed in a 3D context and the reader is referred to section 8.4.

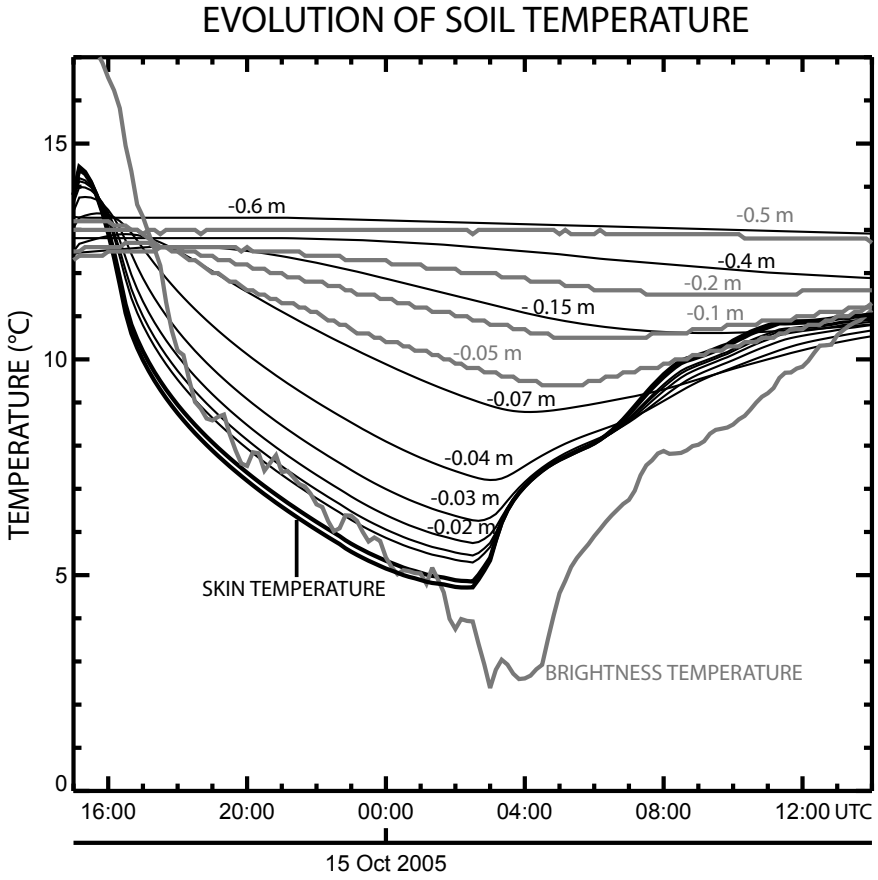

Figure 2: Evolution of modeled (black lines) and observed (thick gray lines) soil temperatures at Zürich Unique airport. Numbers indicate the corresponding soil depth. The corresponding fog forecast as well as surface observations are shown in Figure 22 in the ensemble prediction Section (5.5.1).

3 1D Model comparison

In the framework of COST-722 a 1D model comparison project was carried out. Two sites, namely Paris Charles de Gaulle airport in France and Lindenberg in Germany were chosen due to the availability of good profile measurements for model initialization. Furthermore these sites are located in flat and very smooth topography, respectively. With very little spatial heterogeneities, the two sites are ideal for 1D models. The comparison experiment is thus also a good test of the COBEL-NOAH model.

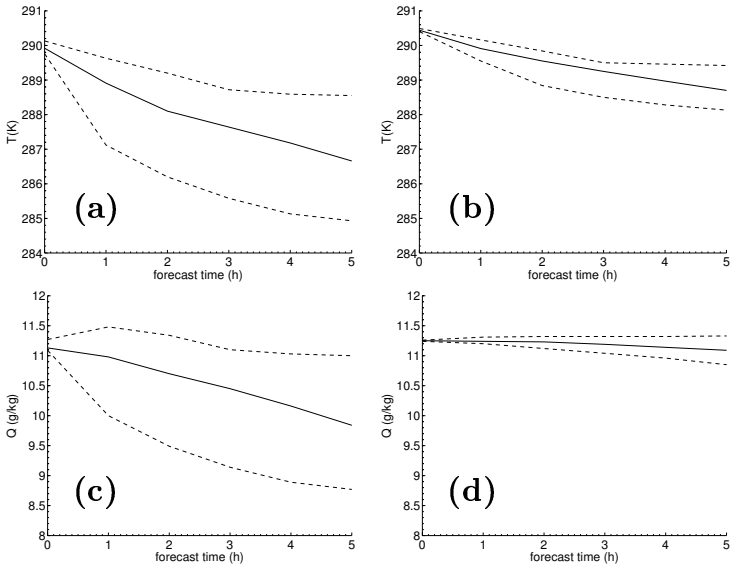

Figure 3: Dispersion between the models for the 1800 UTC run of the fog case of 1-2 October 2003: the mean (plain), minimum and maximum value (dashed) from the different models for temperature at 2 m (a), temperature at 45 m (b), specific humidity at 2 m (c) and specific humidity at 45 m (d). Reproduced from Bergot et al. (2005).

3.1 The Paris-CDG experiment

Six different numerical 1D models, listed in Table 2 participated in the comparison project. A description of the models is given in Bergot et al. (2005, 2006). The vertical resolutions used are very different and this proves to be a cause for significant differences in the simulations. Two cases were simulated, a true fog case (1-2 October 2003) and another case (11-12 October 2003) where relative humidity almost reached 100 % but no fog formation was observed. In order to demonstrate model differences we focus on the actual fog event. For comparison purposes, all models were initialized with the same atmospheric profile data and soil conditions. No participant, except France, knew the actual development of the fog case. For every case, four simulations starting at 1800 UTC, 2100 UTC, 0000 UTC and 0300 UTC, respectively, were computed by each model.

In the following we look at the differences between simulated parameters that develop in the first few hours of the simulation. After 3 hours of simulation, temperature differences of 3 K and humidity differences of 2.5 g kg^{-1} are found at 2 m above ground, when the models are initialized at 1800 UTC (Figure 3). If the models are initialized at 0000 UTC, with liquid water already present, the dispersion is

Table 2: Models of the Paris Charles de Gaulle comparison. The vertical resolution is summarized by the number of levels in the lowest 50 m ($N < 50$ m) and in the lowest 200 m ($N < 200$ m), respectively.

Model	Institution	$N < 50$ m	$N < 200$ m
Hirlam-Isba	I.N.M., Spain	1	3
Hirlam	D.M.I., Denmark	13	20
COBEL-Isba	Météo-France, France	13	20
COBEL-NOAH	Univ. of Basel, Switzerland	18	30
tBM	Analysen & Konzepte, Switzerland	2	7
MesoNH-Isba	Univ. Balears., Spain	50	89

significantly smaller (Figure 4). This is mainly due to the longwave emission of the fog, which keeps the surface at an almost constant temperature.

The classical radiation fog produces a well mixed temperature profile inside the fog with a strong inversion at the top, which can be seen in 4 profiles in Figure 5. The temperature of the mixed layer is regulated by the land surface scheme, or more specifically the temperature at the time of fog formation. After fog formation the temperature of the fog column is mainly regulated by the net radiative loss at the top of the fog layer. However, the presence of clouds can significantly reduce this net loss. Within the fog, the net radiative flux is about zero, because the radiative cooling/heating of droplets produces a thermodynamic equilibrium (Figure 6). But as long as the net radiative flux at the top is not zero this equilibrium is never quite reached. At night the fog column thus keeps cooling. Basically the fog layer acts like one thick layer, that reacts slower to changes the higher it grows since the same energy is distributed over a larger volume. The reduced cooling rates of the fog layer can be seen by comparing Figures 3a and 4a. Temperature profiles of all models for a 3 h forecast, initialized at 0000 UTC, are shown in Figure 5. The classical mixed layer shape is only reproduced by the 4 high resolution models. The coarse grid cannot resolve the vertical structure and resulting physical processes and is thus unsuitable for fog prediction. In summary the results reveal a rather large dispersion in predicted temperature and humidity (Bergot et al. 2005). Differences are largest close to the surface and almost disappear above 100 m. Even though a large dispersion in predicted meteorological parameters was observed, the fog forecasting ability of the high resolution models was rather good. This is surprising but could mean that the actual atmospheric conditions were very favorable for fog formation. Indeed, the relative humidity was already close to saturation at 1800 UTC and the wind speed was observed to be around 1.5 m s^{-1}, so that it does not seem very difficult to forecast fog. Unfortunately, the observations cannot be reproduced here due to restrictive data regulations of Météo-France.

3.2 The Lindenberg experiment

For the Lindenberg experiment two continuous periods were modeled. In the first period from 24-29 September 1998 fog was observed on 4 days and in the second period from 15-17 October no fog had formed. The conditions during both periods were nearly ideal - anticyclonic with almost no wind and thus very little advection. Unfortunately only three different 1D models participated. Namely COBEL-NOAH and tBM from the Paris Charles de Gaulle experiment and PAFOG (see Section) which was run by the University of Bonn. Similar to the Paris-CDG experiment, all models were initialized with the same atmospheric and soil data, but only one run, initialized at 1800 UTC was carried out by every model for every day. Besides the

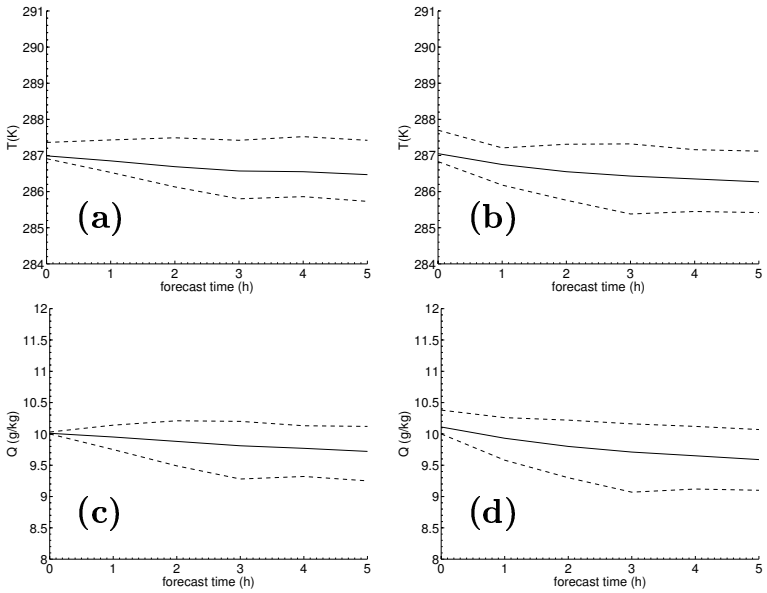

Figure 4: Summary result from the Paris-CDG model comparison study for 1D fog prediction models. 2m (a) and c)) and 45m (b) and d)) variables as a function of forecast time. Shown are the mean of all 6 models (solid lines) and minmum/maximum values (dashed lines), respectively. All models initialized at 0000 UTC for the fog case of 1-2 October 2003. Reproduced from Bergot et al. (2005).

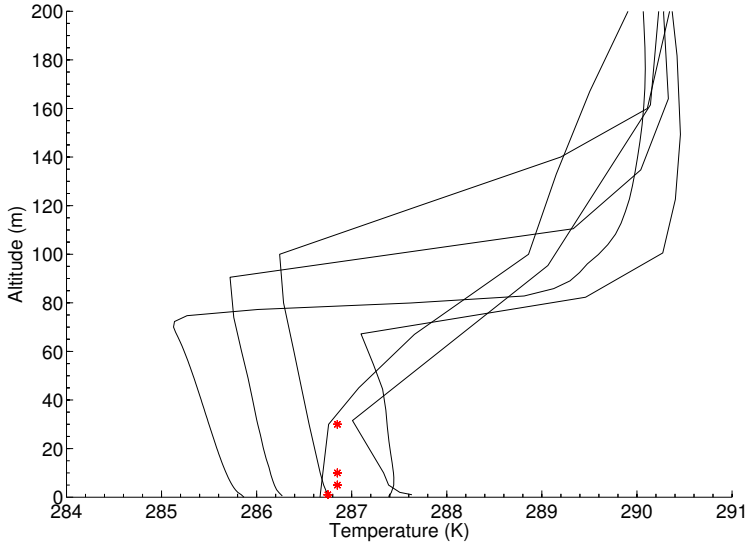

Figure 5: Summary result from the Paris-CDG model comparison study for 1D fog prediction models. Vertical profiles of temperature forecasted by 6 different models at 0300 UTC (3h forecast) of the fog case of 1-2 October 2003. Observations from the tower are indicated with crosses. Reproduced from Bergot et al. (2005).

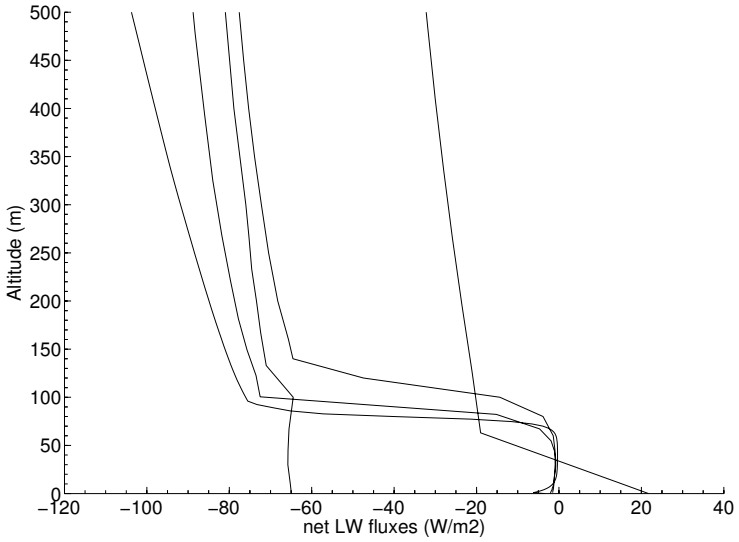

Figure 6: Summary result from the Paris-CDG model comparison study for 1D fog prediction models. Profiles of net longwave radiation flux from 6 different models at 0300 UTC (3h forecast) of the fog case of 1-2 October 2003. Reproduced from Bergot et al. (2005).

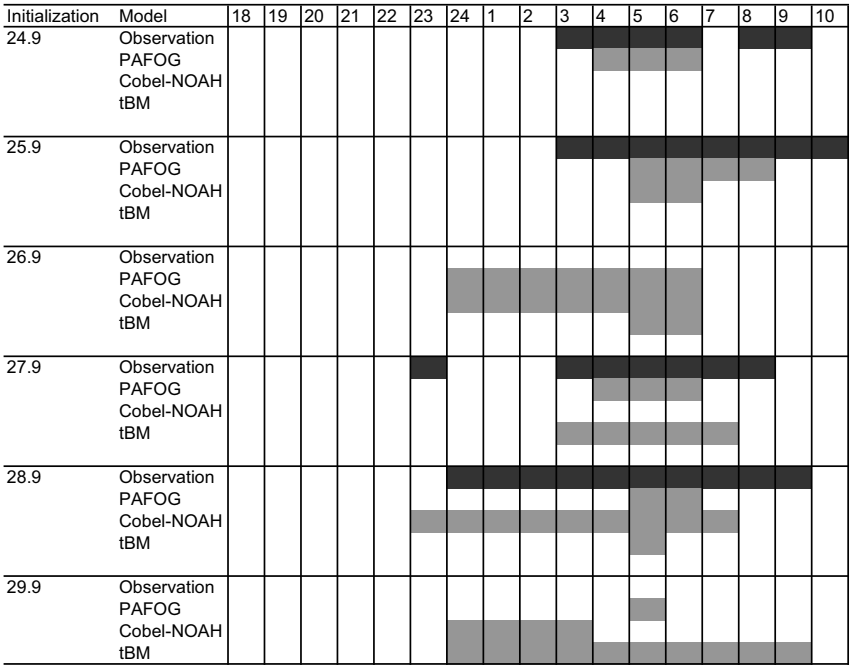

Figure 7: Simulated and observed fog occurrences from 24-30 September 1998 for the Lindenberg experiment.

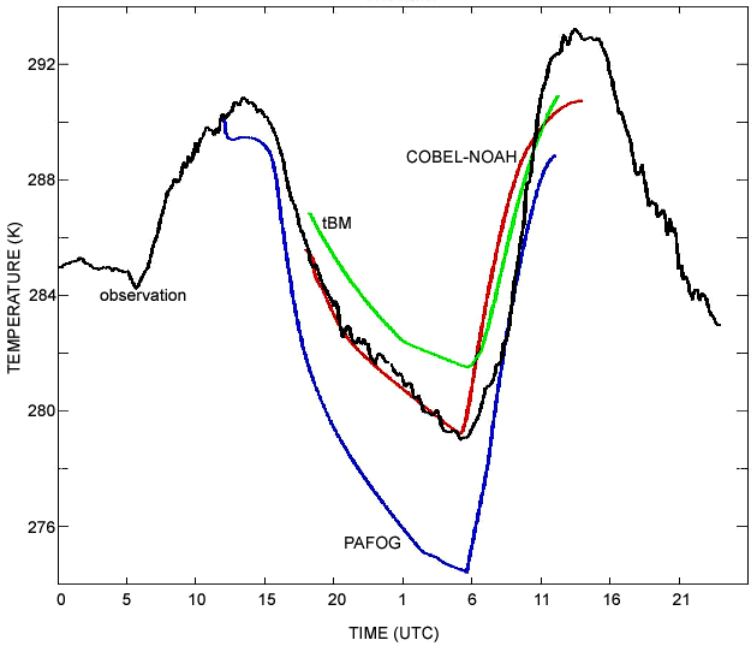

Figure 8: Evolution of observed and modeled 2 m temperature at Lindenberg for the case of 24 September 1998.

initial conditions, no more information about the meteorological development of the cases, not even if fog formed, was known by the participants.

Figure 7 summarizes observed and modeled fog for the first period on an hourly basis. The fog forecasts from PAFOG achieved significant skill, followed by Cobel-NOAH and finally tBM. For 26 and 29 September most models produced false alarms in the form of long fog periods. In the second period however, no model simulated fog occurrence. It has to be noted that tBM never produced fog but low stratus instead. The cloud typically started to grow downward from around 50 m above ground. For the summary in Figure 7 these low stratus clouds were assumed to be fog.

In Figure 8 the evolution of 2 m temperature for the fog case initialized on 24 September is shown. Even though COBEL-NOAH has an almost perfect temperature forecast in this case, the fog forecast of the generally cold biased PAFOG model was better according to Figure 7. The reason is simply that the dewpoint temperature was not reached without the cold bias since predicted humidities were very similar.

In this experiment the forecasted profiles were also compared to radiosonde data. Surprisingly there were not just differences between the models but all modeled profiles were sometimes far away from the observed truth after just 6 hours into the forecast. The differences were around 3 K for temperature and 2 g kg^{-1} for humidity. Since errors are found throughout the whole profile and not just close to the ground they are caused by advection. This is quite a remarkable result since it outlines the importance of advection under almost ideal conditions with weak wind and very little topography. Therefore even for simple terrain a 3D model might be needed to resolve advection on the grid.

Due to the high skill of the PAFOG forecast this model was also included into the ensemble forecast system of Zürich Unique airport.

3.3 Conclusions from the comparison experiments

The comparison demonstrated that fog forecasting is possible if initial conditions are well known, and the topographic and synoptic conditions give rise to no or only little advection. The different models show significant spread after only 2 hours of simulation. To model radiation fog that is generally not very thick, a high vertical resolution is necessary. Even if different models agree about occurrence or non-occurrence under almost ideal conditions, a correct forecast is not guaranteed. This might be related to the fact that some advection is always present. In fact, for certain cases all modeled profiles were significantly different from radiosonde data after just 6 hours of simulation.

4 Initial data and boundary conditions for a 1D forecast model

Prognostic numerical models compute the temporal evolution from a specified initial state. This initial state can be based on assumptions, another forecast valid at initialization time or favorably on observations. As previous work has shown, fog forecasts are very sensitive to initial conditions (Musson-Genon 1987, Fitzjarrald & Lala 1989, Ballard et al. 1991, Bergot & Gudalia 1994*a*). Especially humidity has to be known with a precision of about 0.5 g kg$^-$1 but different cooling conditions might increase or decrease this value. Fog that forms at the end of the night is most sensitive to initial conditions since a very small error decides if fog is formed or not. In reality there are often not enough observations available to properly define the initial state, which is especially true for 3D models. With complex methods of data assimilation, it is however possible to obtain initial conditions for every grid point.

In the following, a discussion and an overview of available observations that were used for model development and data assimilation at Zürich Unique airport are outlined. Furthermore, boundary conditions and their importance for the 1D model are presented. In the next chapter, the actual derivation of initial conditions with variational data assimilation is explained and extended to the case of ensemble forecasting.

4.1 Observations

For the 1D ensemble forecasts at Zürich Unique airport, several observations, listed in Table 3, were set up or already in place. Since the model is run in semi-operational mode, data had to be collected and transferred over the internet almost in realtime. To access data from all the different instruments in an easy way, a database was developed using the IDL programming language. For storage the platform independent, self documented NetCDF file format was chosen.

Table 3: Available observations around Zürich Unique airport. Details about the virtual profile are given in Table 4.

Parameter	Unit
visibility (spatially aggregated)	m
precipitation rate	mm/h
precipitation type	*code*
soil temperature (-0.05, -0.1,-0.2,-1.0 m)	K
volumetric soil moisture (-0.01, -0.02, -0.1, -0.25, -0.6, -0.98 m)	$fraction$
"surface" temperature (0.02 m)	K
temperature (0.5 m)	K
temperature (all heights of virtual profile)	K
temperature profile (MTP-5)	K
wind profile	m/s
humidity (all heights of virtual profile)	%
u-wind (all heights of virtual profile)	m/s
v-wind (all heights of virtual profile)	m/s
cloud base	m
cloud cover	$fraction$
radiation LW$\downarrow$	W/m^2
radiation LW$\uparrow$	W/m^2
radiation SW$\downarrow$	W/m^2

Some observations were only used in the development phase of the COBEL-NOAH

model. Especially helpful for the surface-atmosphere coupling were longwave emission data of the surface to validate the computed skin temperature. It turned out that a high vertical resolution in the soil, with the first layer only 1.5 mm thick, is needed for the model to respond quickly enough to an atmospheric forcing and thus avoid excessive cooling. Temperature and humidity in the atmosphere and in the soil are used in the assimilation scheme to initialize the model. Soil humidity is however not very reliable, as crosschecking with observed precipitation revealed. Soil moisture values are thus taken from the 3D models. The sensitivity to soil moisture is rather small since very little evaporation takes place during the night. It does however play a role in the morning, but then the water deposited by the sedimentation scheme basically saturated the first thin soil layers, so that the initial value is not that important. Wind observations were not very helpful and are not used. The reason is that the 1D model requires a representative geostrophic wind for the entire range of the simulation and not a snapshot of the current turbulent wind profile.

In terms of fog modeling it is best to directly measure profiles of temperature and humidity up to a height of approximately 2 km. This can be easily achieved with a radiosonde. However at an airport this is not allowed for safety reasons and operating costs are also very high. Remote sensing techniques could be used as a surrogate but they are in general less accurate and rather costly. For temperature it was possible to use an MTP-5 microwave profiler but nothing comparable was available for humidity. Thus a sophisticated data assimilation system, described in Chapter 5 had to be developed to obtain vertical profiles based on sparse and not very representative observations.

4.1.1 MTP-5 and the virtual profile

Temperature profiles for Zürich Unique airport can be obtained in two ways, either based on a virtual profile constructed from stations located on nearby mountains or from the MTP-5 microwave profiler. The first method has the problem that all observations are taken in close proximity to the ground instead of several hundred meters above ground so that they are especially error prone under calm conditions. Even though the stations are relatively close to the airport, they still spread over a considerable volume. For windy conditions the influence of the surface is smaller and the stations are more representative. Unfortunately, fog normally forms under calm conditions. On the other hand, the temperature sensors of the stations provide direct and relatively accurate measurements. The indirect remote sensing method of the microwave profiler is less accurate, but provides an in-situ measurement. A drawback of this indirect method are relatively smooth vertical profiles. To compare the two observational systems, mean and root mean square deviations of the virtual profile from the MTP-5 profile at the corresponding height were computed at a temporal resolution of 10 minutes. MTP-5 data were linearly interpolated to the height above sea level where the station measurements took place. Considered are the statistics for each day of the observation period from 26 October 2004 until 11 April 2005. In Table 4 the RMS and mean deviations for the whole observation period are listed. Also indicated is the number of ground based observations with negative or positive deviations from the MTP-5. Overall the virtual profile is not too bad but of course deviates more as altitude increases. Also the number of cases where the stations measured lower temperatures than the MTP-5 increases with altitude, which shows the dominant role of radiative cooling being most pronounced close to the surface. In summertime, increasing positive deviations with higher altitude might be observed due to the much stronger solar radiation and longer daytime. Unfortunately no MTP-5 data are available for the warm season. It is expected that the virtual profile is most reliable when the net radiation is close to zero. This time should be used for model initialization. However to spin up the model, especially the turbulence, it is favorable to have some more turbulent conditions and therefore the model should be initialized

Table 4: Statistics of temperature deviations between the MTP-5 and the virtual temperature profile.

Station	Height (m a.s.l.)	RMS (K)	Mean (K)	negative	positive
Bühlhof	520	0.77	0.44	586	2745
Gubrist	640	0.79	0.28	955	2453
Zürichberg	730	0.90	0.37	922	2486
Lägeren	870	1.21	-0.56	2137	1194

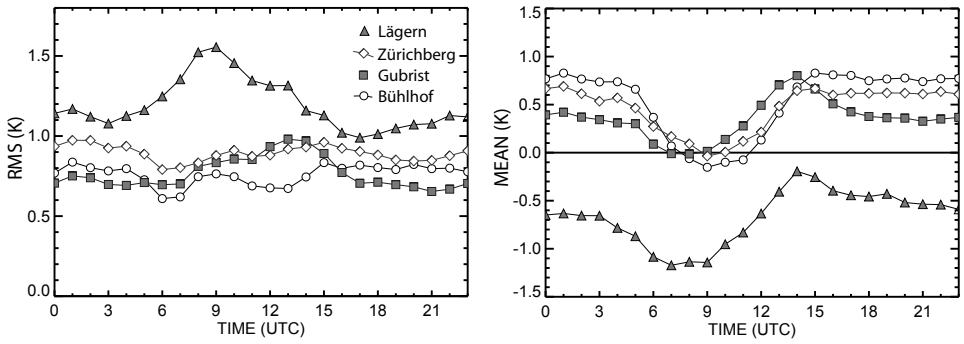

Figure 9: Mean daily course of RMS and mean deviation of temperature measured at the same height above sea level by MTP-5 and the surface stations.

a bit earlier, e.g. at 1500 UTC. Figure 9 shows the mean over all days for RMS and mean deviation as a function of time. As can be seen for Lägeren, the station located at highest elevation, the mean daily course of RMS deviation reaches a minimum around 1700 UTC. For the lower stations the daily RMS amplitude is not very distinct. However, for the mean error the situation is different. All stations show a bump in the morning. This is caused by the transition of the stable nocturnal boundary layer to a better mixed layer, which is seen differently by the surface stations and the microwave profiler. Hereby the surface stations might overestimate the inversion or the profiler underestimate it. But lets note that the morning is not really of interest for fog modeling, since the model will not be initialized at that time. It seems tempting to initialize a simulation in the morning and compute the time of fog dissipation, because a short range forecast should be more accurate. But for a complex numerical model, the initial conditions are even more complicated to derive if fog is present, since a vertical distribution of liquid water has to be defined. But in general, profiles of liquid water and humidity are not observed.

4.2 External forcing of a 1D model

It is common practice (Bergot & Gudalia 1994*b*) to drive a 1D model with temporal boundary conditions, even though this violates the fundamental underlying assumption for the model dynamics of a barotropic atmosphere. Such a procedure is used to extend the applicability of the model to more likely weather situations. Especially in complex terrain, advection plays an important role and is present even under synoptically calm situations in the form of cold air drainage flows.

Boundary conditions are a difficult subject to deal with. Processes like advection of temperature, humidity and wind can significantly modify the current state and also its evolution. For example, a warm air advection aloft increases the stability

and decreases the turbulent exchange, thus changing the temporal evolution. The presence of very weak vertical motions is often responsible for the formation of low stratus clouds. But with a 1D-model it is not possible to compute vertical velocities so that they have to be specified or derived from a 3D model. However, the extreme sensitivity makes it unpractical to specify any vertical velocities (see Section 4.2.2).

Another external forcing is radiation, which is significantly modified by clouds. Due to the limited vertical extent of the 1D model, most clouds are above the model atmosphere. Increasing the vertical extent does not solve the problem, since clouds mostly drift and do not form locally. But the effect of higher clouds can be included using the downward radiation computed by a 3D model at the top boundary height of the 1D model. If the radiation at a certain height is not available from the 3D-model, it can be quantified using a radiation model and predicted cloud cover for medium and high clouds. Note that the radiation model does not have to be extremely sophisticated and computationally expensive, since the cloud forecast from the 3D model can only provide an estimate. In this study, the radiation was computed using predicted cloud cover of medium and high clouds and the radiation code of the PAFOG model.

Under ideal, horizontally homogeneous conditions, for which and only for which, a 1D model is theoretically valid, it is not necessary to specify any time varying boundary conditions. In reality however, there are always stronger or weaker horizontal advection of temperature, humidity or wind, which a 1D model is unable to predict due to the lack of horizontal gradients. Only for cases with weak external forcing, a 1D model is likely to do a decent forecast.

4.2.1 The advection term

Advection of a as defined by (32) is specified as an external tendency. It is computed using centered finite differences [33] about the point of interest, which requires four additional vertical columns from the 3D-model. Figure 10 illustrates the naming convention within the grid point array extracted from every 3D model. Note that eight neighbors are extracted so that interpolation of rotated grids onto northward and eastward oriented axes is possible. Now this numerical solution is not very accurate, but we have to bear in mind that the solution of the large scale advection problem was already done by the 3D model. To make worthwhile a more complex derivation of advection tendencies, e.g. with a semi-Lagrangian approach, it should be done for every time step of the 3D model. The temporal resolution of the output from operational models is however very limited and often one hour or even worse. Thus it has to be assumed that advection tendencies computed every hour can be interpolated in time. This kind of temporal discretization ultimately places the limit on the accuracy of the computed advection terms (Dunlop & Clark 1997).

$$\frac{\partial a}{\partial t} = -\vec{v}\vec{\nabla}a \tag{32}$$

$$\frac{\Delta a_{i,j}}{\Delta t} = -u_{i,j}\frac{a_{i+1,j} - a_{i-1,j}}{2\Delta x} - v_{i,j}\frac{a_{i,j+1} - a_{i,j-1}}{2\Delta y} \tag{33}$$

In order to not disturb the evolution of the lower boundary layer, advection was set to zero in the lowest 100 m above ground. Remember that the development of the temperature inversion in the 1D model is simulated fairly well, and this feature should be maintained.

Another question to be raised is how skillful the forecast of advection is. The validation seems to be very difficult and cannot be done in a direct way without enormous expense, since the required observational setup would have to include several profile measurements at different locations in the area of interest at the same time. Instead advection estimates from different 3D models validating at the same time

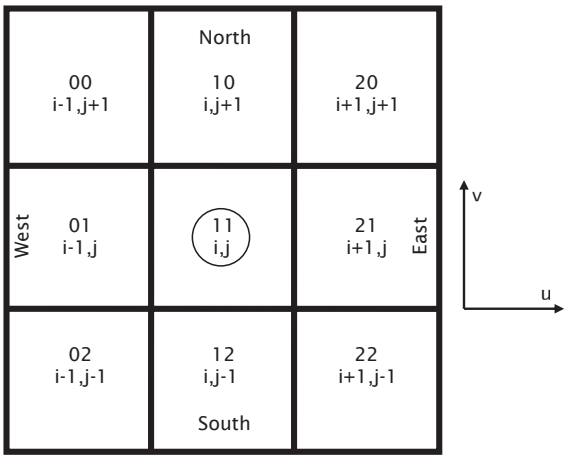

Figure 10: Naming convention for grid points extracted from a 3D model.

and location are compared for every hour and every vertical level within the lowest 1500 m above ground. The examined time period started 1 December 2004 and ended 30 April 2005. The models considered were NMM-22, NMM-4, NMM-2 and the aLMo. In Figure 11 the mean and standard deviation over the different models are shown, where every dot represents one model level at one time. The tendencies for humidity and temperature are per hour. The idea is that advection computed by different models will show similar patterns, as expressed by mean values larger than the standard deviation, if it can be estimated accurately with a model. However, it can be seen that mean values are often significantly less than the standard deviation so that there is basically no clear signal for the advection term. This suggest that advection could also be neglected or treated as a stochastic process without deteriorating the forecast skill, but saving a lot of cumbersome data extraction and setup modifications regarding the 3D model.

In summary, advection has a large impact on the 1D-model, because it directly alters the temperature and humidity profile by the advective tendency. Furthermore advection can be quite large according to forecasts from different 3D models, but there is also a large uncertainty in the predicted advection. The verification will show that inclusion of advection, despite all the problems mentioned, still significantly improves the fog forecast skill (6.4.1).

Due to some difficulties already mentioned, in obtaining estimates of advection, it was also tried to specify a total tendency, rather than pure advection. Therefore, the total hourly change of humidity and temperature in the profile of the 3D model was computed. Of course this change is not solely caused by advection but by all processes. Because at night the planetary boundary layer is not very thick, and surface fluxes loose influence with increasing altitude rather quickly, this forcing is mainly based on advection. For the simulations the total tendency was set to zero in the lowest 100 m and then linearly increases to its full value in the following 200 m. Thereby the 1D-model is undisturbed by boundary conditions in the lower levels and can develop its own boundary layer. The use of the total tendency has some other benefit, originating in the turbulent diffusion. Because good wind forecasts at higher levels are not possible with a 1D model, the resulting turbulent vertical mixing is not very reliable. Note that in 1D there are no horizontal gradients that are necessary for wind computation, and often a nocturnal jet of unrealistic magnitude is developed. Such problems are not present in a 3D framework so that wind forecasts and the resulting vertical exchange are more reliable. However, as verification will reveal, the

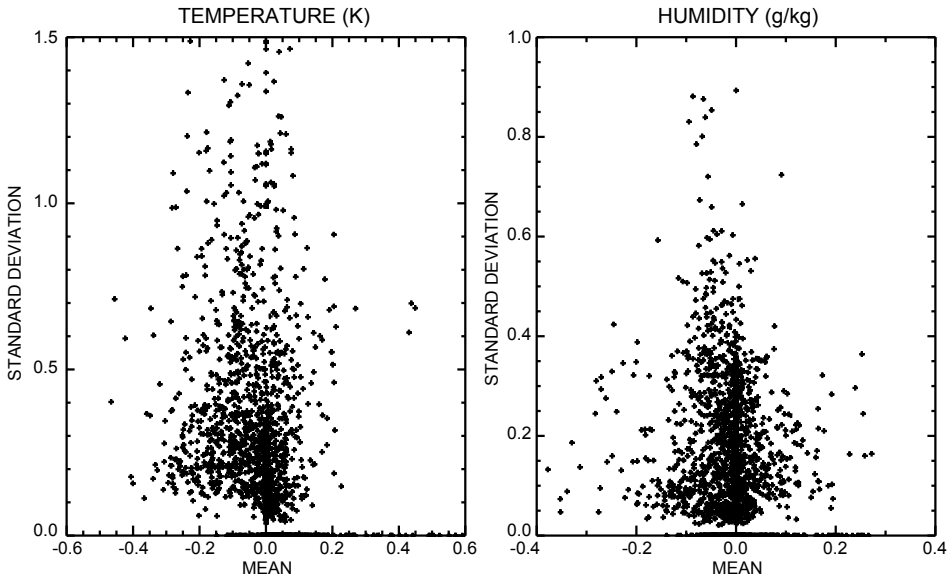

Figure 11: Hourly advection estimates for the period from 1 December 2004 - 30 April 2005 at Zürich Unique airport from different mesoscale models.

runs with the total rate of change are worse than purely advection based runs. The details can be found in the verification section (6.4.1).

4.2.2 Vertical velocity

Vertical velocities are important for fog and cloud formation due to the resulting adiabatic temperature change. Yet the dynamics of a 1D model make it impossible to compute vertical motion inside the model and it has to be specified as an external forcing. Unfortunately these velocities are of such a small magnitude that they cannot be measured but have to be extracted from a 3D model. Now the vertical velocities in the 3D models are not very reliable and often very noisy as well. The magnitudes might not be realistic and the physics packages of a 3D model need to be tuned to work with these estimates. For example the nonhydrostatic MESO-ETA model has relatively low vertical velocities when run with step mountain coordinates and higher vertical velocites when run with sigma coordinates. Experiments with different models used in the COST-722 showed that 1D models are so sensitive to vertical motion that its inclusion is of no practical use indeed. The behavior is illustrated in Figure 12, where the fog event starting on 11 November 2004 is modeled with COBEL-NOAH using no vertical velocity, an upward motion of 0.1 cm/s and a downward motion of 0.3 cm/s, respectively. As can be seen in the case of no vertical motion, the fog event is predicted quite well. However, even a very small vertical forcing severely limits fog formation. In other cases vertical motion also resulted in unrealistically strong fog growth. Another example is given for the 1D PAFOG model in Figure 13, where a low stratus is exposed to weak downward motion. As can be seen the stratus remains relatively close to the ground and almost dissolves at 0500 UTC if vertical motion is enforced. In summary, due to the high sensitivity of the 1D model to vertical motion and the difficulties in obtaining a representative value, it is best to neglect vertical motion in a 1D fog model.

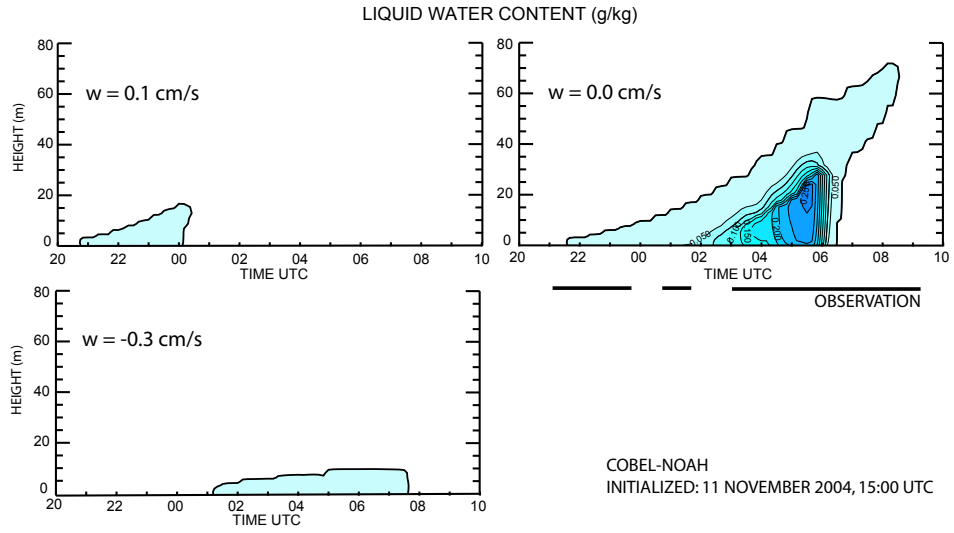

Figure 12: Sensitivity of COBEL-NOAH to vertical motion for the fog event starting on 11 November 2004 at Zürich Unique airport.

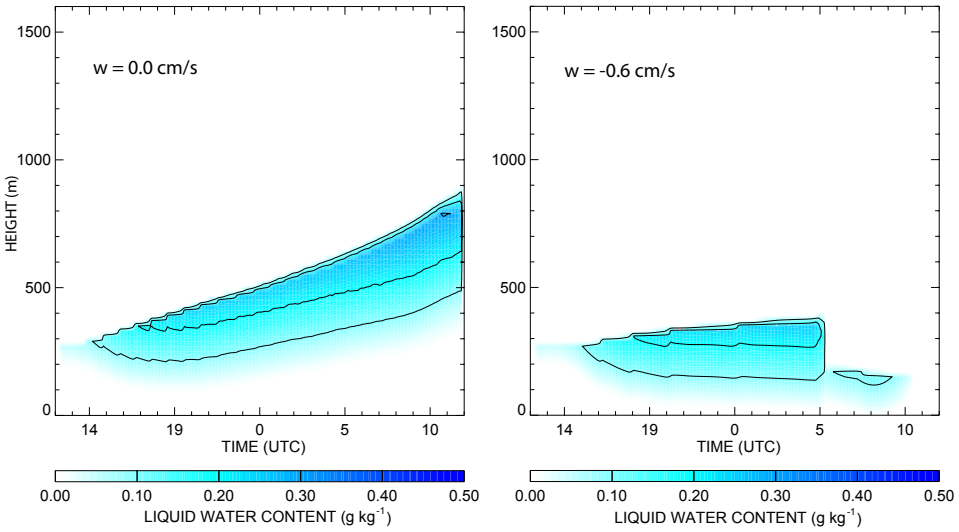

Figure 13: Sensitivity of PAFOG to vertical motion for a low stratus case.

5 Variational data assimilation and ensemble forecasting

Data assimilation has grown to a large and important field of research over the past years. This is because differential equations of any forecast model solve the temporal evolution from a given initial state but cannot derive the initial state itself. Basically the quality of the forecast is strongly dependent on the initial state it started from. Many articles and books have been devoted to the subject of data assimilation. The articles of (Ghil & Malanotte-Rizzoli 1991, Talagrand 1990, Lorenc 1986, Ingleby 2001, Schlatter 2000), the technical report of Bannister (2003) and lecture series of ECMWF (Bouttier & Courtier 1999, Fisher 2002*a*,*b*,*c*) are very useful to understand the complicated subject. Also the books of Daley (1991) and Kalnay (2003) are a valuable reference. In the following, a brief description of variational data assimilation is given, which includes basic theory but also some technical details, necessary to develop a data assimilation system. The focus lies on one dimensional variational assimilation (1D-VAR), as it was developed for Zürich Unique airport. 3D-VAR and 4D-VAR systems are based on the same theory, which is extended in the case of 4D-VAR, but mainly the larger size of the system introduces a manifold of technical problems and solution strategies outlined mainly in technical reports of the the above references. Similar to numerical modeling, data assimilation looks nice and clean as written equations on paper, but becomes really messy when it has to be implemented and actually solved numerically. I therefore try to avoid most aspects of implementation and numerical solution and refer the interested reader to the source code.

5.1 Understanding variational data assimilation

The process of approximating the true state of a physical system at a given time is called analysis. An analysis which combines time distributed observations and a dynamic model is called assimilation or data assimilation (Daley 1991). The assimilation has to know the statistical error properties of observations and of the numerical model in order to give each information source the proper weight. Under the assumption that errors can be modeled by a Gaussian unbiased distribution, and thus entirely specified by a covariance matrix, variational assimilation becomes the problem of minimizing the cost function [34]. Detailed mathematical derivations of the cost function are given in Hólm (2003) and Kalnay (2003).

$$J(\vec{x}) = \frac{1}{2}(\vec{x} - \vec{x_b})^T \mathbf{B}^{-1}(\vec{x} - \vec{x_b}) + \frac{1}{2}(\vec{y} - \mathbf{H}\vec{x})^T \mathbf{R}^{-1}(\vec{y} - \mathbf{H}\vec{x}) \tag{34}$$

Here the following notation, proposed by Ide et al. (1997), and hypothesis are used:
$\vec{x}$ model state (dimension n)
$\vec{x_t}$ true model state (dimension n)
$\vec{x_b}$ background model state (dimension n)
$\vec{y}$ vector of observations (dimension p)
$\mathbf{H}$ linearized observation operator (from dimension n to p)
$\mathbf{B}$ covariance matrix of background errors ($\vec{x_b} - \vec{x_t}$) (dimension $n \times n$)
$\mathbf{R}$ covariance matrix of observation errors ($\vec{y} - \mathbf{H}\vec{x_t}$) (dimension $p \times p$)

- Unbiased errors: the expectation of observation and background errors is zero: $\overline{\vec{x_b} - \vec{x_t}} = \overline{\vec{y} - \mathbf{H}\vec{x_t}} = 0$
- Uncorrelated errors: observation and background errors are mutually uncorrelated: $(\vec{x_b} - \vec{x_t})(\vec{y} - \mathbf{H}\vec{y_t})^T = 0$

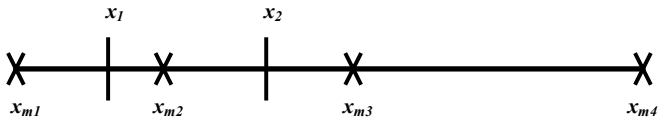

Figure 14: A simple case with four model grid points and two observations.

- Linear analysis: the analysis defined by corrections to the background state depends linearly on background observation departures
- Optimal analysis: the analysis state is as close as possible to the true state in an r.m.s. sense (i.e. it is a minimum variance estimate)

So basically once **H**, **B** and **R** are determined and the observations $\vec{y}$ as well as a first guess for $\vec{x}$ are available, the values in $\vec{x}$ are modified. Every modification results in costs in a least-squares sense that are quantified by a scalar value, which is the result of Eqn. 34. The values of the state vector where $J(\vec{x})$ has its minimum is the analysis.

The first thing to do when setting up an assimilation scheme is defining a state vector $\vec{x}$ which will store the information of the system. The values of $\vec{x}$ may start with the temperature of the lower left corner on the lowest model layer followed by the other gird points in this layer before proceeding to the next vertical layer. After the last temperature value, the first humidity value could follow and so on until all needed variables on all grid points are stored. The background state $\vec{x_b}$ will be the first guess for the assimilation and is normally a previous model forecast valid at the time of analysis. The true state of the system $\vec{x_t}$ is unknown but the assimilation tries to find the best estimate of it, called the analysis $\vec{x_a}$. To carry out an analysis some observed values are needed and stored in an observation vector $\vec{y}$. The storage structure of $\vec{y}$ should also be carefully organized, e.g. it could contain all available measurements from different instruments at a particular location before proceeding to the next location. Unfortunately the number and locations of measurements will not correspond to the the state vector $\vec{x}$. There might be locations where several measurements from different instruments are available, and others where no observations are available. Even worse some observations may only be a proxy of a variable in the state vector. It is thus necessary to introduce an observation operator **H** which must be linearized and expressed as a matrix. Multiplying by **H** on the left, as it is done in the cost function [34], transforms model state values of $\vec{x}$ into synthetic observation values corresponding to elements in $\vec{y}$ and the adjoint $\mathbf{H}^T$ transforms from observation points back to grid points. The design of **H** can be quite difficult and e.g. include a linearization of a radiative transfer model as done by Janisková et al. (2002) or Eyre (2002) to obtain radiances measured by a satellite from modeled vertical profiles of temperature and humidity. In the current framework of considered observations, such complicated transitions do not have to be included and **H** just represents a linear interpolation of model points to observation locations. Thus, every row of **H** has no more than two nonzero elements. If $\vec{x}$ has 4 elements and observations are taken at two different locations x_1 and x_2 as illustrated in Figure (14), then the coefficients of **H** for linear interpolation are:

$$\mathbf{H} = \begin{pmatrix} \frac{x_{m2}-x_1}{x_{m2}-x_{m1}} & \frac{x_1-x_{m1}}{x_{m2}-x_{m1}} & 0 & 0 \\ 0 & \frac{x_{m3}-x_2}{x_{m3}-x_{m2}} & \frac{x_2-x_{m2}}{x_{m3}-x_{m2}} & 0 \end{pmatrix} \tag{35}$$

To actually perform an analysis, the differences between the background state and the observations, given by $\vec{y} - \mathbf{H}\vec{x}$ have to be analyzed and minimized through an iterative process. As can be seen, the cost function [34] has two terms, where the

first one penalizes departures of the analysis from the background and the second one does so for departures from observed values. The magic of data assimilation now lies in the treatment of uncertainties inherent to observations and the background state, which means a model for errors is introduced. Mathematically this is done by error covariance matrices $\mathbf{R}$ for the observations and $\mathbf{B}$ for the background state. Let us denote the background errors as $\vec{\epsilon_b} = \vec{x_b} - \vec{x_t}$ and the observation errors as $\vec{\epsilon_o} = \vec{y} - \mathbf{H}\vec{x_t}$. The symmetric error covariance matrices are then obtained by multiplying the vectors with their transpose:

$$\mathbf{B} = \overline{(\vec{\epsilon_b} - \overline{\vec{\epsilon_b}})(\vec{\epsilon_b} - \overline{\vec{\epsilon_b}})^T} \tag{36}$$

$$\mathbf{R} = \overline{(\vec{\epsilon_o} - \overline{\vec{\epsilon_o}})(\vec{\epsilon_o} - \overline{\vec{\epsilon_o}})^T} \tag{37}$$

In the case of 3 elements in $\vec{x}$ and the corresponding backround errors (e_1, e_2, e_3) this yields:

$$\mathbf{B} = \begin{pmatrix} var(e_1) & cov(e_1, e_2) & cov(e_1, e_3) \\ cov(e_1, e_2) & var(e_2) & cov(e_2, e_3) \\ cov(e_1, e_3) & cov(e_2, e_3) & var(e_3) \end{pmatrix} \tag{38}$$

The problem now is, that $\vec{\epsilon_b}$ and $\vec{\epsilon_o}$ both contain the true state of the system which is unknown. But before considering ways of determining $\mathbf{B}$ and $\mathbf{R}$ it is helpful to shed some light on their meanings. $\mathbf{B}$ is a square symmetric matrix of dimension $n \times n$, where n is the dimension of the model state vector $\vec{x}$. In 3D-VAR the model state vector can contain about 10^7 elements, yielding 10^{14} elements for $\mathbf{B}$, which exceeds the memory capacity of supercomputers. With the control variable transform, the size of $\mathbf{B}$ can be enormously reduced into smaller parts (see Bannister (2003)). The procedure is outlined later for the case of 1D-VAR. It might be helpful to see that the off-diagonal terms in [38] can be transformed into correlations by

$$\rho(e_i, e_j) = \frac{cov(e_i, e_j)}{\sqrt{var(e_i)var(e_j)}} \tag{39}$$

Thus the background error covariance matrices describe correlations between different locations and different variables. Observation errors are often assumed to be uncorrelated, reducing $\mathbf{R}$ to a diagonal matrix. Background error correlations are of major importance for data assimilation. They are responsible for information spreading, information smoothing and for physical balance. The spreading of information is necessary to obtain analysis increments for a limited area of points surrounding the observation. It fills the holes in data-spares areas. On the other hand, in data-dense areas the correlations in $\mathbf{B}$ control the amount of smoothing. For example under anticyclonic conditions horizontal correlations are desirable to spread and average the observations, whereas for inversions or fronts the smoothing should be kept relatively small. Balance properties are imposed by correlations between variables. The large scale atmosphere is mainly hydrostatic and geostrophic, thus a single temperature observation can modify not just the temperature of surrounding grid points, but also the wind field. The correlations and variances determine the relative amplitude of the assimilation increments for the various model fields.

5.2 Methods for estimating background error covariances

The quality of the assimilation relies on an accurate estimation of $\mathbf{B}$. This is a difficult task since it cannot be observed directly and hence has to be estimated in a statistical sense. Simplifying assumptions of homogeneity and stationarity have to be made. The later can be avoided by using ensemble Kalman filtering introduced by Evensen (1994), where an ensemble of data assimilations is used to estimate a situation dependent background error covariance (Houtekamer & Mitchell 1998, 2001, Hamill &

Snyder 2000, Hamill et al. 2001, Anderson 2001). If stationarity is assumed, a simple and reliable estimation method is the observational or Hollingworth-Lönnberg method (Hollingsworth et al. 1986, Hollingsworth & Lönnberg 1986). Given a dense and large enough observing network that provides uncorrelated discrete information on many scales, it is possible to determine background and observation errors. Observation minus background covariance statistics are computed for different distances from the observing point. At zero separation, averaged information about observation and background errors are obtained, which need to be separated by using the interpolated background error correlation tendency. The background error correlation should go to zero for large separations and increase towards the observing point.

Another widely used method that does not depend on measurements at all is the "NMC" or NCEP Method introduced by Parrish & Derber (1992).

$$\mathbf{B} \approx \alpha \frac{1}{n} \sum_{i=1}^{n} \left((\vec{x_f}(t_1) - \vec{x_f}(t_0))(\vec{x_f}(t_1) - \vec{x_f}(t_0))^T \right) \tag{40}$$

where $\vec{x_f}$ represents the forecast state vector and α is an empirical scaling factor. As can be seen in [40] $\mathbf{B}$ is estimated as the average over n differences between two short-range model forecasts verifying at the same time. Normally $t_1 = 48$h, $t_0 = 24$h and about 50 different forecasts representative for the season are used. In 3D-VAR spatial inhomogeneities can also be taken into account following e.g. the methodology of Wu et al. (2002).

The cost function can also be written in incremental form, so that the minimization process tries to find the optimal increment to be added to the background state (Courtier et al. (1994)). We can define perturbation quantities $\vec{x'}$ and $\vec{y'}$ with the help of a linearization state, which will be the background $\vec{x_b}$.

$$\vec{x'} = \vec{x} - \vec{x_b} \tag{41}$$

$$\vec{y'} = \vec{y} - \mathbf{H}\vec{x_b} \tag{42}$$

Thus the incremental form of the cost function becomes:

$$J(\vec{x'}) = \frac{1}{2}(\vec{x'}^T \mathbf{B}^{-1} \vec{x'}) + \frac{1}{2}(\vec{y'} - \mathbf{H}\vec{x'})^T \mathbf{R}^{-1} (\vec{y'} - \mathbf{H}\vec{x'}) \tag{43}$$

The incremental form has the advantage that any imbalance introduced through the analysis procedure is limited to the increments that are added to the balanced background. So because the increments are of smaller magnitude than the background state the balance is disturbed only little. During the iterative minimization process, the values of $\vec{x'}$ are varied to find the minimal value of the cost function J. As can be seen in [34,43] this involves the inversion of $\mathbf{B}$, which is normally a large and badly conditioned matrix. It is however not obvious to see this bad conditioning, but any attempt to solve [34, 43] is very likely to fail. The shape of the cost function can illustrate this for a two variable space of $\vec{x'}$.

In Figure 15 isolines of J are sketched with the optimal analysis being located in the center. A minimization algorithm now starts somewhere on the outer lines and tries to reach the center by proceeding in a down gradient direction. This process is much easier in regions of small curvature than in regions with large curvature, where the algorithm proceeds in a zig-zag path, a phenomenon called the narrow valley effect. A small curvature represents a mode of the data assimilation system with high variance and thus little information.

Mathematically the shape of the cost function is described by the second derivatives of J with respect to the elements of the control vector $\vec{x}$, the so called Hessian $\mathbf{A_x}$. The derivation of the Hessian, given in Bouttier & Courtier (1999) yields:

$$\frac{\partial^2 J(x)}{\partial x^2} = \mathbf{A_x} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \tag{44}$$

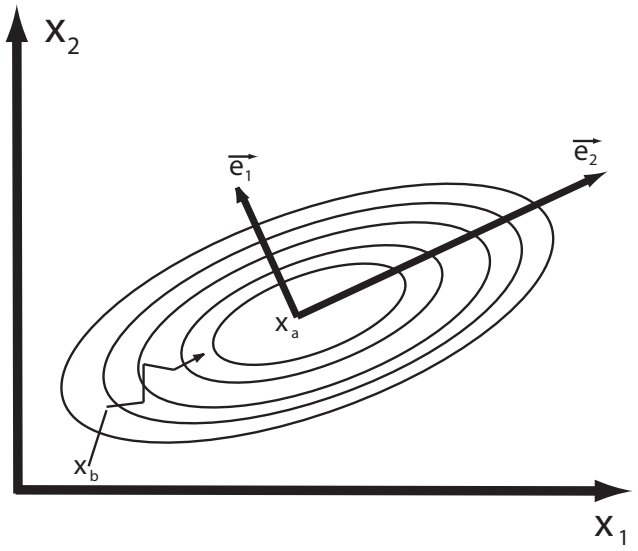

Figure 15: Graphical illustration of the cost function with iso-J curves. The isolines are narrow ellipses and the minimum, which corresponds to the ideal analysis, lies in the center.

and we see that the shape is independent of the actual values of the state vectors $\vec{x}$ and $\vec{y}$. In Figure 15 the principal axes of J are the eigenvectors of $\mathbf{A_x}$ and the curvature of J along these directions are given by the eigenvalues. Normally the principal axes are different from the axes given by the state vector, which means they represent no meteorological variable at a particular location but some sort of mixture between variables and locations.

The conditioning number of J is defined as the ratio between the largest and the smallest eigenvalue of the Hessian. Bad conditioned matrices have a large conditioning number, giving rise to slow convergence and error prone results of the minimization. For a perfectly conditioned matrix the number is unity, which corresponds to a circle of J-Isolines in the above example.

To be able to solve the assimilation problem it is necessary to precondition $\mathbf{B}$ and reduce the number of elements, which is done by the so called control variable transform. The aim is to get rid of $\mathbf{B}$ in the cost function, which can be stated as follows:

$$\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} = \mathbf{I} \tag{45}$$

It is thus necessary to transform our model variables (actually perturbations) $\vec{x'}$ into control variables $\vec{v'}$, the so called T-Transform and vice versa using the so called U-Transform:

$$\vec{x'} = \mathbf{U}\vec{v'} \tag{46}$$

$$\vec{v'} = \mathbf{T}\vec{x'} \tag{47}$$

In order to appreciate this, substitute [46] into the cost function [43],which can then be formulated in $\vec{v'}$-space:

$$J(\vec{v'}) = \frac{1}{2}((\mathbf{U}\vec{v'})^T\mathbf{B}^{-1}\mathbf{U}\vec{v'}) + \frac{1}{2}(\vec{y'} - \mathbf{HU}\vec{v'})^T R^{-1}(\vec{y'} - \mathbf{HU}\vec{v'}) \tag{48}$$

$$J(\vec{v'}) = \frac{1}{2}(\vec{v'}^T\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U}\vec{v'}) + \frac{1}{2}(\vec{y'} - \mathbf{HU}\vec{v'})^T R^{-1}(\vec{y'} - \mathbf{HU}\vec{v'}) \tag{49}$$

with the help of [45], $\mathbf{B}$ can be eliminated:

$$J(\vec{v'}) = \frac{1}{2}(\vec{v'}^T \vec{v'}) + \frac{1}{2}(\vec{y'} - \mathbf{HU}\vec{v'})^T R^{-1}(\vec{y'} - \mathbf{HU}\vec{v'}) \tag{50}$$

The cost function does not contain the largest matrix anymore and the complexity of the problem has been significantly reduced. Since the minimization now works on the control vector $\vec{v'}$, the final analysis has to be converted back to $\vec{x}$-space:

$$\vec{x_a} = \vec{x_b} + \mathbf{U}\vec{v'} \tag{51}$$

During the minimization process, the cost function and the gradient of [50] as derived by Bouttier & Courtier (1999) and transformed into the incremental form in $\vec{v'}$-space [52] have to be evaluated during every step of the minimization.

$$\nabla_{\vec{v'}} J(\vec{v'}) = \left(\frac{\partial J(\vec{v'})}{\partial \vec{v'}} \right)^T = \begin{pmatrix} \frac{\partial J}{\partial v'_1} \\ \dots \\ \frac{\partial J}{\partial v'_n} \end{pmatrix} = \vec{v'} - \mathbf{U}^T \mathbf{H}^T \mathbf{R}^{-1}(\vec{y'} - \mathbf{HU}\vec{v'}) \tag{52}$$

In the current implementation, the Broyden-Fletcher-Goldfarb-Shanno variant of the Davidon-Fletcher-Powell method, as described in Section 10.7 in Numerical Recipies in C (Press et al. 1998), is used to compute the minimization of the cost function with the help of its gradient. Equations [50] and [52] now describe the assimilation, however the control variable transformations have to be found first. In order to do this, the background error covariance matrix $\mathbf{B}$ has to be diagonalized, which can easily be done with the small size of $\mathbf{B}$ in the case of 1D-VAR. The decompositon of $\mathbf{B}$ into the diagonal matrix of eigenvalues λ and the matrix $\mathbf{E}$ composed of the corresponding eigenvectors is expressed as:

$$\mathbf{B} = \mathbf{E}\lambda\mathbf{E}^{-1} \tag{53}$$

This most general form has the problem that it may destroy the physical meaning of $\mathbf{B}$. For example, if the error covariances represent different volumes, as it is the case for a model grid with variable vertical resolution, all elements are weighted equally and not in relation to their volume. To avoid this, a diagonal matrix $\mathbf{P}$ of grid box volumes is used for scaling. Technical details are given in Bannister (2002).

$$\mathbf{BPE} = \mathbf{E}\lambda \tag{54}$$

Since $\mathbf{P}$ is a symmetric matrix it is useful to decompose it into its square roots:

$$\mathbf{P} = \mathbf{P}^{T/2}\mathbf{P}^{1/2} \tag{55}$$

Now lets substitute this decomposition into [54] and also we multiply on the left both sides of [54] with $\mathbf{P^{1/2}}$ to obtain

$$\mathbf{P}^{1/2}\mathbf{B}\mathbf{P}^{T/2}\mathbf{P}^{1/2}\mathbf{E} = \mathbf{P}^{1/2}\mathbf{E}\lambda \tag{56}$$

which is equivalent to

$$\mathbf{B'F} = \mathbf{F}\lambda \tag{57}$$

where [58] are the eigenvectors of the transformed $\mathbf{B}$ matrix [59]

$$\mathbf{F} = \mathbf{P}^{1/2}\mathbf{E} \tag{58}$$

$$\mathbf{B'} = \mathbf{P}^{1/2}\mathbf{B}\mathbf{P}^{T/2} \tag{59}$$

After the eigendecomposition has been completed, using standard mathematical library routines, the **U**-transform can be computed following Bannister (2003). Because the background error covariance matrix is symmetric also $\mathbf{B}'$ is symmetric so that the corresponding eigenvectors are orthogonal:

$$\mathbf{F}\mathbf{F}^T = I \quad \text{and} \quad \mathbf{F}^T\mathbf{F} = I \tag{60}$$

so that

$$\mathbf{F}^T = \mathbf{F}^{-1} \tag{61}$$

This will be useful now and also in the computation of the adjoint and inverse of the U-transform. To find the U-transform, the eigenvalue problem [57] has to be brought into a form similar to [45]. To do this [57] is first multiplied on the left with $\mathbf{F}^T$ on both sides:

$$\mathbf{F}^T\mathbf{B}'\mathbf{F} = \lambda \tag{62}$$

Furthermore the diagonal eigenvector matrix λ is expressed as $\lambda^{1/2}\lambda^{1/2}$ and we multiply with the inverse on both sides:

$$\lambda^{-1/2}\mathbf{F}^T\mathbf{B}'\mathbf{F}\lambda^{-1/2} = \lambda^{-1/2}\lambda^{1/2}\lambda^{1/2}\lambda^{-1/2} \tag{63}$$

Finally [63] is inverted and after substitution of [59] the form containing $\mathbf{U}$ and $\mathbf{U}^T$ as in [45] is obtained:

$$\lambda^{1/2}\mathbf{F}^T\mathbf{P}^{-T/2}\mathbf{B}^{-1}\mathbf{P}^{-1/2}\mathbf{F}\lambda^{1/2} = I \tag{64}$$

where the right part is the U-transform:

$$\mathbf{U} = \mathbf{P}^{-1/2}\mathbf{F}\lambda^{1/2} \tag{65}$$

and the left part is the adjoint:

$$\mathbf{U}^T = \lambda^{1/2}\mathbf{F}^T\mathbf{P}^{-T/2} \tag{66}$$

Finally the T-transform is the inverse of the U-transform:

$$\mathbf{T} = \lambda^{-\mathbf{1/2}}\mathbf{F}^T\mathbf{P}^{1/2} \tag{67}$$

The T-transform can convert model variables to new parameters which have mutually uncorrelated errors. This can be thought of as doing a Fourier decomposition, but the projection is onto empirical orthogonal functions (EOFs) rather than plane waves. The EOFs are found from the eigendecomposition of **B** as shown above. So we have to change our view of **B**. It was used to represent error variances at different vertical levels and covariances between levels. Now, with the T-Transform, **B** consists of variances of the EOFs that are uncorrelated and thus have no covariances, which makes the problem much simpler. It is funny that the T-Transform is not needed to solve the assimilation, but the U-Transform and the transposed U-Transform are needed many times in the computation of the cost function and and its gradient, respectively. So it might be hard to picture the T-transform, but even harder for the inverse, the U-transform, which does the same but in the other direction. Finally note that the variances of the vertical EOFs are the eigenvalues.

Figure 16 shows an example of the eigenvector matrix **E** and eigenvalues λ for the estimated temperature error covariance matrix of the NMM-22 model at 00 UTC. Each eigenvector is in a column and each row represents each variable. The eigenvectors are sorted in decreasing order corresponding to the size of the eigenvalues λ from left to right. As can be seen, most of the information is carried in a few leading eigenvectors and the problem could be further simplified by cutting off eigenvectors on the right.

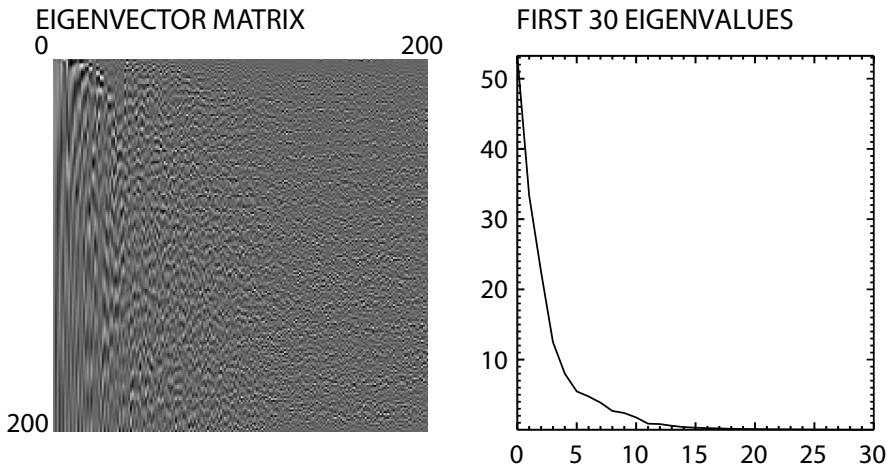

Figure 16: Example of an eigenvector matrix **E** with 200 vectors and the first 30 corresponding eigenvalues λ.

5.3 Estimating the observational error covariance matrix

Similar to the background error covariance matrix, **R** specifies errors of the observational system. The error is mainly caused by representativeness problems of the observation in model space and only secondly an instrumental characteristic. In contrast to **B**, correlations are assumed to be zero. This seems reasonable for observations made with different instruments but may be invalid for observations done with the same platform such as a radiosonde or satellite measurements. Especially, complicated preprocessing, expressed by **H**, might require model data which are correlated so that also the derived observations become correlated. Furthermore, representativeness errors are correlated as soon as observations are dense compared to the model resolution. However, in practice the correlations are difficult to estimate and can create problems in the numerical solution of the assimilation. Hence a diagonal structure for **R** is used in most assimilation schemes. This can be partly achieved by minimizing correlations by avoiding unnecessary preprocessing of observations and thinning of dense data.

Concerning the assimilation at Zürich Unique airport, the most difficult part is putting a label on the representativeness of radiosonde data recorded in Payerne. Because the latter is about 150 km away from Zürich, the lower part of the sounding is expected to be rather unreliable. To quantify the similarity between the two locations, model profiles from high resolution numerical weather prediction were analyzed. It is believed that the NMM model, run at 2 km resolution is able to capture most spatial differences between Payerne and Zürich. Thus, for the time from October 2004 to March 2005, correlations for humidity as well as temperature, were computed for all vertical layers, respectively. The result is shown in Figure 17 for radiosonde ascent time of 1200 UTC.

For the virtual profile of temperature and humidity the errors increase with increasing station height. This expresses the fact that the station located at the surface can only provide an estimate of the thermodynamic state of the free atmosphere at that height. Table 5 lists the errors used in the assimilation scheme. For the virtual temperature profile the values could be derived from comparison with MTP-5 data as described in Section 4.1.1.

It has to be noted that new observations, like temperature reports from landing aircrafts, can easily be included in the assimilation system. All what has to be done is

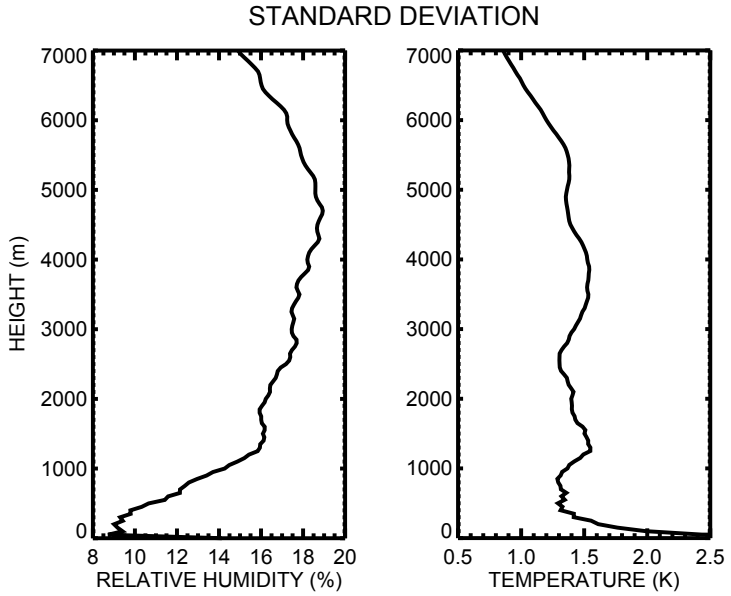

Figure 17: Standard deviations from the observational error covariance matrix $\mathbf{R}$ of the radiosonde in Payerne.

Table 5: Error variance of operational systems used in the assimilation for Zürich Unique airport.

Station	Height (m a.s.l.)	Temperature error	Humidity error
Bühlhof	520	0.04	5
Gubrist	640	0.8	10
Zürichberg	730	0.9	28
Lägeren	870	1.21	35
MTP-5	all levels	0.5	

specification of the observational error variances, which are inserted as new elements into $\mathbf{R}$. They need to be at the right place according to the storage logic used to build the observational vector $\vec{y}$, which then also has the new observations inserted. Finally $\mathbf{H}$ has to be extended to provide model estimates for the new observations.

5.4 Assimilation strategy for a 1D model

Data assimilation optimally combines observations with a first guess or background estimate. In this case data have to be assimilated for a 1D model, and it seems natural to use a previous forecast of that model as background state. There are however several reasons for using the 3D model forecasts, having resolutions between 2 and 7 km, as a first guess. In 1D it is not possible to simulate horizontal gradients which are responsible for advection and wind. Even though the latter is modified due to turbulent diffusion and might be quite accurate close to the ground, the upper levels are a different story and often subject to overestimation of nocturnal jets. Thus, nonlocal changes have to be introduced in form of tendencies computed on the 3D model using gradients and wind. It is favorable to take tendencies instead of computing the advection based on the 1D model wind due to the weakness of 1D wind forecasts. But the tendencies computed on the 3D model are not very accurate. This

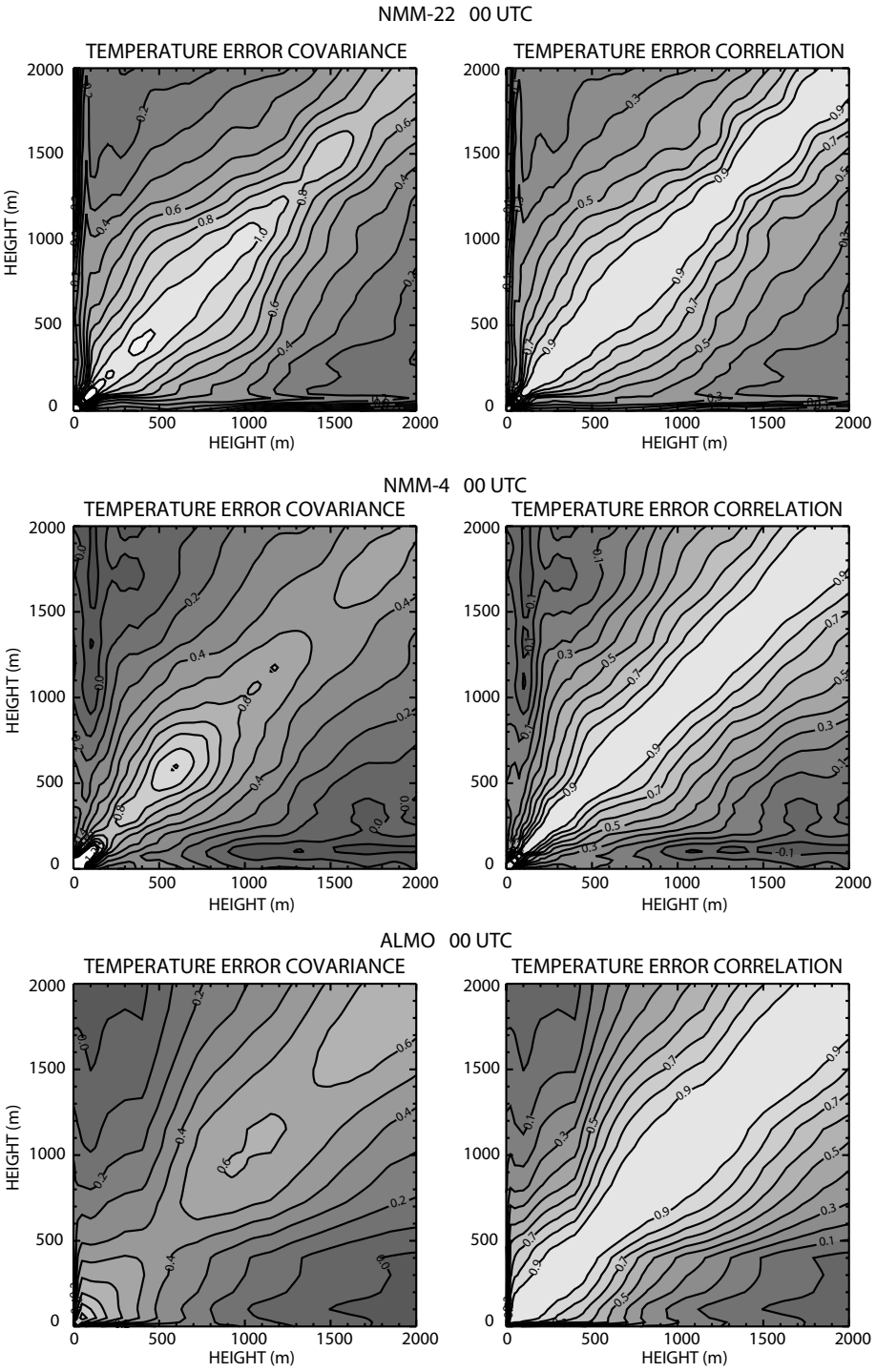

Figure 18: Temperature error covariance and error correlation matrices for the winter season 2004/2005 at 00 UTC for 3 different 3D models.

is because 3D model output is normally only available with very limited temporal resolution of 1 hour or even worse. Another big problem is the limited vertical extent of the 1D model, which currently simulates the lowest 1500 m of the atmosphere. Thus most clouds are not part of the model and the radiative fluxes at the model top have to be supplied from a 3D model as well. Therefore the 1D model is heavily dependent on data that have to be derived from 3D model output of low temporal resolution. So if instead a column of the high resolution 3D model is taken as first guess, it might be less accurate close to the surface, but the higher levels are obtained directly. The 1D model background could also be computed in realtime with a special assimilation run, which integrates new observations every 10 minutes. Unfortunately, observations of temperature, humidity, wind and radiation are available only at the surface, and advection at higher levels is unrelated to them and can introduce large errors over time. To prevent a drift, the model has to be corrected with measurements covering the whole column. Therefore, a radiosonde would be the preferred choice, but at the airport such data are not available. Regardless which model provided the first guess, a drift correction is carried out during variational assimilation. Radiosonde data from Payerne, some 150 km away, in combination with an error covariance matrix, that takes into account the spatial correlations, provide the necessary means. Considering the whole profile, the surface layer contains not much energy and the high resolution 1D model adjusts the surface layer profiles in a relatively short time. Also the assimilation will correct a suboptimal surface layer background from the 3D model, using all the observations. Furthermore, due to large error variances and abundant observations close to the ground, the assimilation normally gives little weight to the background in the surface layer. Another important point is the fact that a 3D model is able to do skillful forecasts for several days, where the most recent one is not necessarily the best. Thus, several forecasts initialized in the past, but valid at the same time, can be used as first guess for the generation of ensemble members. Note that determination of the background error covariance matrix $\mathbf{B}$, using the NMC method, also requires a model able to do skillful time integrations of 48 hours or more. Such long forecasts are impossible with a 1D model, making the determination of $\mathbf{B}$ a very difficult task.

The estimates of $\mathbf{B}$ obtained for different models using the NMC-Method [40] and corresponding correlations, computed with [39] are shown in Figure 18. The statistics are based on the vertical profiles of the 3D models from October 2004 to February 2005. Of course these are overall statistics including different synoptic situations. It would also be possible to derive specific $\mathbf{B}$ matrices for particular synoptic conditions, but this requires either a sophisticated automated classification of the situation for which an assimilation has to be done, or an experienced human forecaster. Rather than using error prone classifications and a set of static $\mathbf{B}$ matrices for which it might be difficult to obtain enough cases to be statistically representative, extended Kalman filtering could be used in the future. In this approach a whole ensemble of assimilations is carried out to estimate the forecast error covariance (Evensen 1994, Hamill & Snyder 2000, Houtekamer & Mitchell 1998, 2001, Anderson 2001). In particular the ensemble of initial conditions from the previous cycle is integrated with the model and the error covariance for the current cycle is estimated from the members of the ensemble forecast. However, the large number of time integrations with the 3D model make such an approach computationally very expensive.

If we now have a look at the estimates of $\mathbf{B}$ shown in Figure 18, it is evident that variances are largest close to the surface, where most of the energy fluxes are transformed and a small change in e.g. predicted cloud cover results in a large temperature difference. This means that in this region, the background term will have relatively little influence compared to the observations. Fortunately most observations are available close to the surface and an unreliable forecast for that layer does not pose a problem for data assimilation. For the two resolutions of NMM, a second

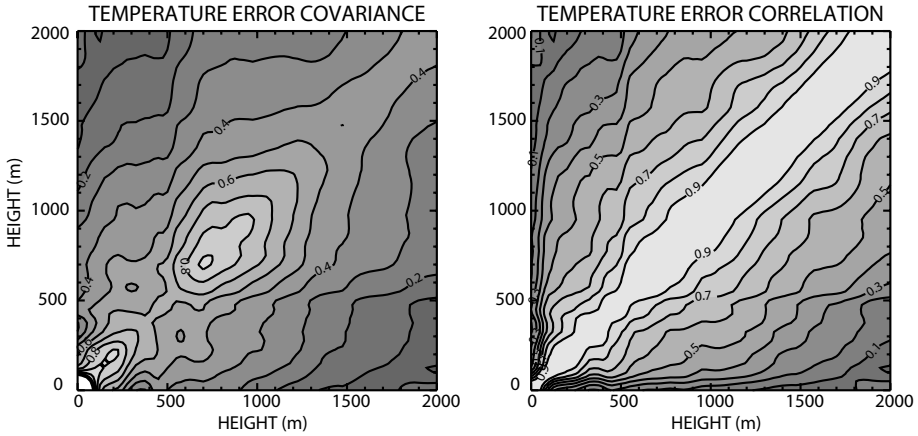

Figure 19: NMM-4 temperature error covariance and error correlation matrices for the winter season 2004/2005 at 1400 UTC.

maximum can be found around 500 m above ground which is not present in the aLMo. The correlations between vertical layers generally increase from ground level to the region of maximum variance. In the aLMo the vertical layers are less independent of each other than in the NMM, which means that the spread and smoothing of information during the assimilation process is bigger. The structure of the error covariance is different throughout the day due to the daily evolution of the boundary layer. The temporal evolution is rather smooth and thus not shown in detail. As an example, **B** for NMM-4 at 1400 UTC is shown in Figure 19. When compared to 00 UTC, the secondary maximum of error variance is found at higher altitude due to the generally thicker boundary layer in the afternoon.

As an example of a successful minimization of the cost function, assimilated temperature and humidity profiles for 28 November 2004 at 1200 UTC are presented in Figure 20. The background estimate is a 21 hours forecast of NMM-2. The temperature profile of the MTP-5 and the radiosonde of Payerene agree quite well, indicating an inversion with a lower bound between 200 and 300 m above ground. In the model background this inversion is rather week and some zig-zag temperature structure is present in the lowest 500 m which corresponds well with the pattern shown by the stations. The assimilated profile combines dominant features of the structures and has its absolute values close to the different observations. For humidity, no profiler data are available, and the surface stations do not provide a representative value for upper levels, since the surface is the main source of moisture. Furthermore the effects of radiative cooling and heating are strongest close to the surface, resulting in a temperature error which also influences relative humidity. The observational error standard deviation for station based humidities is thus specified to be as large as 35 % at the highest station. As for temperature, the benefits of variational assimilation are well manifested. The assimilated humidity profile inherits the structure of the NMM-2 model forecast but is drawn towards the absolute values of the radiosonde.

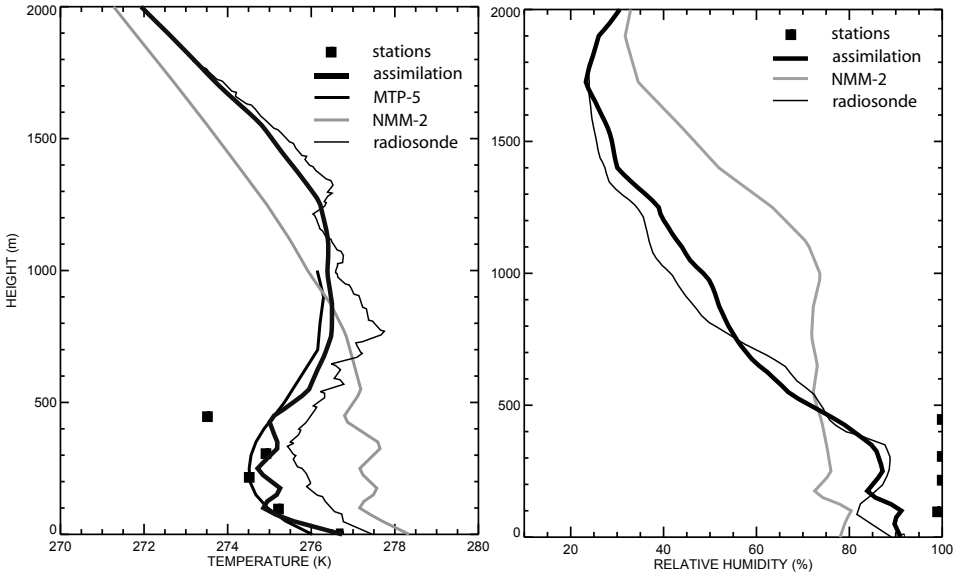

Figure 20: Assimilated temperature and humidity profiles for 28 November 2004 at 1200 UTC.

5.5 The ensemble forecast system

The forecast skill of an atmospheric model depends on the accuracy of the initial conditions, the quality of the model and on the current meteorological conditions itself (Lorenz 1963*a*,*b*, 1965). Lorenz showed that any nonlinear system with instabilities has limited predictability and that it is necessary to account for the stochastic nature of the evolution of the atmosphere.

The first method to deal with the uncertainty was developed by Epstein (1969) who introduced stochastic-dynamic forecasting. His method is a simplification of the infinite Monte Carlo ensemble but still requires about half as much model integrations as degrees of freedom, which makes it completely unfeasible even for a simple weather forecast model with millions of degrees of freedom. Leith (1974) found that Monte Carlo forecasting can be done with few members if forecast errors can be estimated and used in the derivation of initial conditions.

As an alternative to Monte Carlo forecasting, lagged average forecasting was proposed by Hoffman & Kalnay (1983). With this method the current forecast as well as forecasts initialized earlier are combined to form an ensemble. From the older forecasts a previous forecast error can be computed and used for the generation of initial conditions. So far the members of different age are all weighted equally by the operational centers, since it is very difficult to derive weights.

An extension to lagged average forecasting is scaled lagged average forecasting and was introduced by Ebisuzaki & Kalnay (1991). Therein a linear scaling of error growth during the first 2-3 days is used and the number of ensemble members is doubled because perturbations are not just added, but also subtracted from the analysis.

Houtekamer et al. (1996) and Houtekamer & Mitchell (1998) developed an ensemble forecasting system based on running a whole ensemble of data assimilation systems to create initial conditions. Every data assimilation system adds random errors to the observations. The errors added are in addition to the observational error specified with the error covariance matrix $\mathbf{R}$. Furthermore different parameters in the physical parameterizations of the model are used in different ensembles.

A very costly and work intensive way to derive initial conditions is by introducing singular vectors, which are basically the perturbations with maximal energy growth in the forecast. The computation of the singular vectors requires the tangent linear model and its adjoint. In practice the linear model is obtained by linearizing the nonlinear forecast model which is an enormous amount of work. To obtain the singular vectors many forward and backward integrations with the linearized model and its adjoint are needed, which makes the procedure computationally very expensive. This approach is used at ECMWF (Molteni et al. 1996, Molteni & Palmer 1993, Buizza 1997, Buizza et al. 1997).

Very popular at the moment is the multisystem approach, also called "poor man's ensemble". It turns out that an ensemble average of deterministic operational global forecasts from different operational centers is far more skilful than the best individual forecast (Fritsch et al. 2000). Hou et al. (2001) showed that this also applies to short-range regional models. Krishnamurti et al. (2000) further improved the quality of the ensemble forecast by removing systematic model errors by means of regression. He calls this ensemble a "superensemble".

The 1D ensemble system developed for Zürich Unique airport uses the numerical model COBEL (Bergot & Gudalia 1994*a*,*b*), which has been coupled to the NOAH land surface model (Chen et al. 1997, Ek et al. 2003) as described in Section 2.2.2. Furthermore the 1D PAFOG model is run (Bott et al. 1989, Bott & Trautmann 2002), which integrates the same initial conditions as COBEL. The derivation of initial conditions in the 1D ensemble system is outlined in the next section.

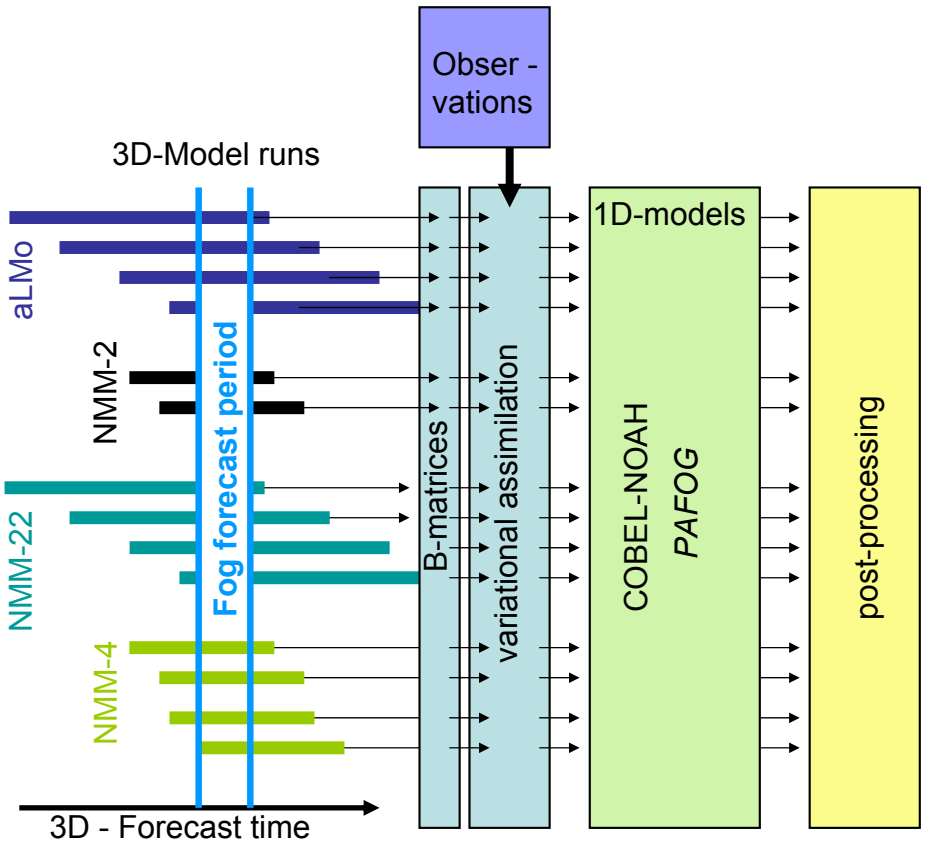

Figure 21: Schematic of the 1D ensemble prediction system. Every 3D run provides initial conditions that are used as a background for an individual variational assimilation.

5.5.1 Initial conditions for ensemble members

An important part of every ensemble forecast is the derivation of initial conditions. They should be representative for the current uncertainty of the initial state. Perturbations added to the control analysis should thus be the leading eigenvectors of the analysis error covariance (Ehrendorfer & Tribbia 1997) and also include our uncertainty about model deficiencies. Due to limited computational resources and time constraints, it is impossible to do Monte Carlo simulations. Therefore a relatively small set of initial conditions has to be derived that covers the variational range a Monte Carlo simulation would produce. With few ensemble members this is very difficult to achieve and remains always an approximation. The most common methods used were summarized in the preceding section.

For the 1D model, a new and somewhat different ensemble strategy was developed. It combines ideas from ensembles based on multiple data assimilation, lagged average forecasting and the multisystem ensemble approach.

The main problem of a 1D forecast is the assumed horizontal homogeneity and the resulting dependence from a 3D model to include the effects of heterogeneity. The 1D model on its own is thus unable to produce an independent skillful forecast under most conditions. The only exception are situations with almost perfect horizontal homogeneity. Also the generation of the ensemble members should have computational costs comparable to the time integration. Here, instead of adding deviations

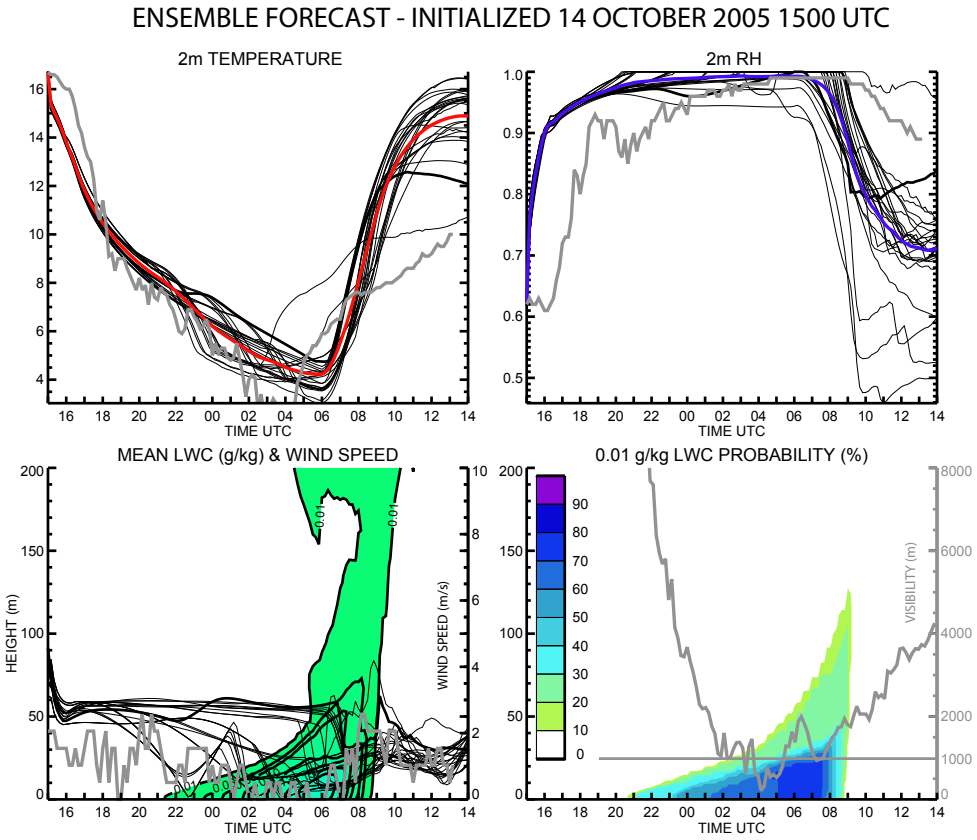

Figure 22: 1D ensemble prediction of the fog event from 14-15 October 2005. The first two panels show computed temperature and humidity at 2 m height for each member (thin lines), the ensemble mean (thick colored line) as well as the corresponding observations (thick gray line). In the lower left panel the ensemble mean liquid water content is contoured together with observed (thick gray line) and modeled wind speed. The last panel indicates the probability that a liquid water content of 0.01 g kg^{-1} is exceeded. Figures were taken from the developed semi-operational system.

to a so called control run to obtain a set of initial conditions, a whole bunch of different control runs is used. It is interesting to note that adding a deviation to the control run, by definition, makes the best estimate of the true state worse, so that the resulting time integration is expected to have less skill. For these reasons the initial conditions are based on available forecasts computed with different 3D models. Currently the aLMo of MeteoSwiss, running at 7 km resolution, as well as the operational forecasts with the Nonhydrostatic Mesoscale Model (NMM) at resolutions of 22, 4 and 2 km are used. The NMM forecasts are very distinct due to the dynamical core (Janjic et al. 2001, Janjic 2003), the resolution, but also due to differences in physical parameterizations. For example the model with 2 km resolution uses unparameterized convection and a newly developed prototype of a land surface model developed by Zavisa Janjic from NOAA/NCEP. The dynamical core of the 22 km version is hydrostatic and the higher resolutions are nonhydrostatic. The aLMo has a different dynamical core and also other physical parameterizations. It is initialized with ECMWF data, whereas the NMM is driven with the Global Forecast System (GFS). Now to increase the number of 3D forecasts, all the runs valid at the same time, but initialized at different times are also used. This can be done since the most

recent run does not have to be the most skillful. Of course this is not true for runs that are several days old and initial conditions are not taken if they are older than two days. It has to be noted that other models could be easily included if available.

The different 3D models driven by distinct global assimilations, their varying resolutions as well as different initialization times produce a spread in initial conditions which is dependent on the current weather situation. Experiments at NOAA/NCEP revealed that initial conditions based on lagged average forecasting, scaled lagged average forecasting and on the forecast differences between forecasts verifying at the same initial time grew much faster than Monte Carlo perturbations with similar size and statistical distribution (Kalnay & Toth 1996). If the atmosphere is very predictable, e.g. under stable anticyclonic conditions the forecasts are more similar than during westerly flow situations. However the 3D models have limited skill and the boundary layer and local influences are represented with different degrees of accuracy. Hence, such forecasts should not be used directly and have to be combined with local observations. The forecasts are thus used as a first guess in the variational assimilation described in at the beginning of this Section. Hereby, profiles of temperature and humidity are assimilated separately for every 3D model background using observations from a temperature profiler, surface observations on nearby mountains and data from a radiosonde, located about 150 km away from the airport. The model error covariance Matrix **B** needed in the variational assimilation was derived separately for every 3D model using the NMC-method. In Figure 21 a schematic of the ensemble forecast system is given. Note how the 1D forecast time range is cut out of the 3D forecasts so that it is possible to extract initial and temporal boundary conditions.

Postprocessing finally aggregates the individual forecasts and computes probabilities of liquid water occurrence as well as mean values of predicted variables. Furthermore graphical output is generated for every individual member as well as for aggregated information. In the last step a web page is compiled to allow for platform independent access and easy navigation through hundreds of plots. Figure 22 gives an example of an ensemble forecast for the fog event during the night of 14-15 October 2005. As can be seen in the upper left panel, the cooling during the night is predicted rather good, but because the modeled fog disappears around 0800 UTC the temperature rises too fast in the morning. Note how the temperature forecasts from different members slowly diverge. The probability for a liquid water content above 0.01 g kg^{-1} is over 70 % and indeed fog formed that night. The timing however was not perfect according to the visibility observations shown in the lower right panel.

Since the assimilation of humidity has to operate with few and not very representative observations, the resulting humidity profile is a rather uncertain estimate. To address this issue, a humidity ensemble is generated by simply making the assimilated profile drier and moister, respectively. A variation of 10 % in relative humidity was taken. To avoid supersaturation, the humidity profile cannot exceed a relative humidity of 99 % anywhere. The verification presented in section 6.4.3 reveals that the moister profiles reach the highest skill, thus suggesting that the assimilation is too dry.

5.5.2 Boundary conditions for ensemble members

Specification of the boundary conditions for the ensemble members is done as outlined in section 4.2. It was explained that the advection term, which represents the boundary conditions for the 1D forecast, is difficult to determine and only a crude estimate of reality. It is thus reasonable to further increase the number of ensemble members with different temporal boundary conditions. This means that the same initial conditions in combination with different external forcings during the time integration are used. Such a procedure has not been used elsewhere, mainly because advection is computed internally by the 3D model. In the implementation, the number of members is tripled. The same initial conditions are integrated using no advection, the total

tendency and the mesoscale advection. Currently, initial conditions and boundary conditions are taken from the same 3D forecast run. It is however possible to further increase the number of ensemble members by mixing initial and boundary conditions derived from different 3D runs.

Note that in order to define an external forcing, the 3D model run needs to temporally cover the whole period of the 1D forecast. Of course the boundary conditions cannot be improved with data assimilation because no observations of the future atmospheric conditions are available.

6 Verification of the 1D ensemble fog forecast

Now this is the moment we have all waited for. Was it worth all the trouble? This section tries to answer this but to raise suspense even more, some details about the verification method are outlined first. It should also be noted that a model cannot be verified but only falsified, and thus the term model validation would be more appropriate. However the term verification is used here to be consistent with current literature.

6.1 Verification methods for deterministic forecasts

For categorical forecasts, having only two possible outcomes, a simple contingency table can be defined (Table 6). In our case the binary event is fog or no fog and threshold values for visibility and liquid water content will have to be used to define a fog event.

Table 6: Contingency table for categorical forecasts of a binary event. The numbers of observations in each category are denoted by a,b,c and d, respectively.

forecast	observed		
	yes	no	total
yes	a (hit)	b (false alarm)	a+b
no	c (miss)	d (correct rejection)	c+d
total	a+c	b+d	a+b+c+d

The most basic descriptive measure based on the contingency table is the base rate s or event probability. It is solely a characteristic of the observations and not of the forecast model. Unfortunately many performance measures are in some way dependent on the base rate, so that the pure natural variability may cause different skill even though the forecast model has not changed. Using the variables introduced in Table 6, the base rate can be computed as

$$s = \frac{a+c}{a+b+c+d} \tag{68}$$

For the verification period chosen the base rate is very low. Only 2 % for a visibility threshold of 500 m, 5 % for 1000 m and 9 % for 1500 m. So fog is a very rare event.

Several performance measures can be defined based on the contingency table. The hit rate H (Swets 1986), also called probability of detection, is the proportion of occurrences that were correctly forecast.

$$H = \frac{a}{a+c} \tag{69}$$

Since a good forecast should have high hit rates and a low number of false alarms, the hit rate alone is insufficient for measuring forecast skill. The false alarm rate F (Swets 1986) is the proportion of non-occurrences that were incorrectly forecast by the model. It must be distinguished from the false alarm ratio FAR (Donaldson et al. 1975), which is the conditional probability of of a false alarm conditioned on the event being forecast. FAR is reduced if more events are observed and is thus not a very reliable performance measure. Note that the number of forecast events can be controlled by changing the forecast probability threshold needed to define a fog event. A high probability threshold will reduce the false alarm rate.

$$F = \frac{b}{b+d} \tag{70}$$

$$\text{FAR} = \frac{b}{a+b} \tag{71}$$

The proportion of correct forecasts PC (Finley 1884) is also expressed in percent and then called percent correct. Note that very frequent or very rare events score with a high PC if occurrence or non-occurrence, respectively, is always predicted.

$$\text{PC} = \frac{a+d}{a+b+c+d} \tag{72}$$

Heidke (1926) adjusted PC to account for the proportion of forecasts that would have been correct by chance. His Heidke Skill Score (HSS) defines hits due to chance E as the event relative frequency multiplied by the number of event forecasts.

$$\text{HSS} = \frac{\text{PC} - E}{1-E} \quad \text{with random prognosis} \quad E \tag{73}$$

$$E = \left(\frac{a+c}{a+b+c+d}\right)\left(\frac{a+b}{a+b+c+d}\right) + \left(\frac{b+d}{a+b+c+d}\right)\left(\frac{c+d}{a+b+c+d}\right)$$

The HSS is an equitable score, because constantly forecasting occurrence or non-occurrence results in zero skill. HSS can have values from -1 to 1, where 1 corresponds to a perfect forecast, -1 is a perfect, but wrongly calibrated forecast and 0 is the no skill level. It has to be noted that the HSS is highly dependent on the base rate and threshold probability so that it is a rather unreliable performance measure. This will be seen later when HSS and ROC are compared. There are many other measures of skill that can be derived from the contingency table. A very good review of verification concepts and skill scores is given in Jolliffe & Stephenson (2003). For our purpose the above mentioned scores together with the signal detection methods mentioned in the next section are a valuable verification tool to asses the quality of the forecast system.

6.2 Verification of probabilistic forecasts

The performance measures mentioned so far cannot directly asses the quality of probabilistic forecasts. The development of a general framework for verification by (Murphy & Winkler 1987) together with the introduction of methods from signal detection theory by Mason (1980, 1982, 1989) allow for the verification of probabilistic forecasts. Basically it is necessary to first transform the probability forecast into a set of binary yes/no forecasts using a whole sequence of probability threshold in the range 0 to 1. An event is forecast if the specified probability threshold is exceeded. In that way the contingency table for every threshold value, as well as all the categorical skill measures can be computed. Now this seems to become a rather messy procedure with hundreds of verification scores to keep track of. But fortunately there is a powerful signal detection tool called the *relative operating characteristic* (ROC) which is obtained by plotting the hit rate H versus the false alarm rate F for each possible decision probability threshold. The ROC distinguishes between the decision threshold and the intrinsic discrimination capacity of the forecast system. Low probability thresholds result in both, high hit rates but also high false alarm rates. These points of the ROC are located in the upper right corner of the ROC diagram. A forecast model with perfect discrimination has a ROC curve that rises from (0/0) along the hit rate axis. Note that the diagonal has no skill and if the ROC curve is below the diagonal, the system is wrongly calibrated, because it again requires skill to be that bad. In practice the ROC curve can only be approximated for a certain number of probability thresholds. A resolution of 1 % in the probability requires at least an ensemble of 100 members. Every threshold probability gives a point on the ROC, and neighboring points are then connected to form the ROC curve. If a threshold probability of the ensemble forecast is exceeded it makes a modeled fog event. But

what can we learn from a particular ROC curve? It illustrates the false alarms and hit rates the system has for a particular forecast probability. So the threshold for good detection and relatively few false alarms can be found. However, depending on the costs of false alarms or missed events another probability will be chosen.

The most popular overall skill measure in this context is the area under the ROC, typically denoted A_z, which would be unity for a perfect system and 0.5 for a no skill system. For values of A_z smaller than 0.5 the corresponding ROC curve lies below the diagonal, indicating the same level of discrimination ability as if it was symmetrically above the diagonal but wrongly calibrated in this case.

The verification has to be carried out with fog events that are not directly predicted but need to be defined e.g. based on threshold values for observed visibility and modeled liquid water content. This opens a new field of problems, namely how to automatically detect a fog event.

6.3 Detection of fog events

In terms of observations, the aggregated visibility estimate derived from the different instruments installed around the runways is used. In that way an area is probed rather than a single point, which is more reliable. By definition the visibility has to be below 1000 m for fog, but in the verification also other thresholds are used to account for some uncertainty in the observations and representativeness. From the model output, visibility and liquid water content can be analyzed. But visibility itself is not a prognostic variable and is derived using statistical relations that include liquid water content and eventually also the droplet number concentration. For the purpose of verification, a fog event is classified based on the presence of liquid water, rather than low visibility. This is because visibility might be wrongly derived from the prognostic variables and thus introduce another source of error. For fog the modeled liquid water content has to exceed 0.01 g kg^{-1}. Note that also higher thresholds were used but resulted in worse verification scores.

All verification is done on a temporally aggregated resolution of one hour. Fog is classified as such, when the modeled liquid water content threshold is exceeded or the observed visibility lies below the threshold value for at least 10 min.

For aviation purposes it is of course very important to know visibility since it makes a big difference if a pilot can see 1000 m or only 50 m. However this is far too high a goal at the moment for a numerical model. In fact just predicting occurrence or non-occurrence of fog with formation and clearance times accurate to one hour is a very difficult task in a complex location such as Zürich Unique airport.

6.4 Verification results for Zürich Unique airport

In this section the verification results of fog forecasts at Zürich Unique airport obtained from the 1D ensemble forecast system are presented. The verification period begins on 1 November 2004, when all observations needed for data assimilation and verification (see Table 3) were available on a regular basis, and it ends on 30 April 2005. Verification will address the importance of advection and uncertainties in the humidity assimilation, the impact of different initialization times as well as effects of the driving 3D model. For every initialization time an ensemble consists of around 30 members and over 50'000 runs were done in total, so that the statistical significance of this verification should be fulfilled. All verification is done on a temporally aggregated resolution of one hour in the time window from 03-11 UTC. This time window is of great importance for an airport and also the time of likely fog occurrence. The discrimination between fog and no fog for model and observation data is done as described in the preceding Section.

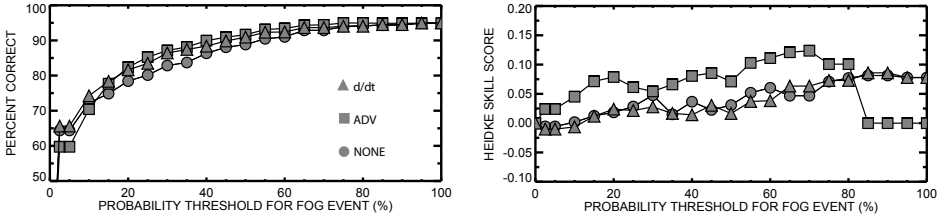

Figure 23: Percent Correct and Heidke Skill Score for fog occurrence simulated with the 1D ensemble if different derivations of advection are used; ADV = pure advection, d/dt = total rate of change in profile, CONTROL = no advection. Verification was done on an hourly basis between 03-11 UTC from 1 November 2004 until 30 April 2005 at Zürich Unique airport. The observational visibility threshold for a fog event was 1000 m and all simulations were initialized at 1500 UTC.

6.4.1 Importance of advection

Advection might be insignificant in flat terrain, but will certainly collect its toll in a location such as Zürich Unique airport. As there are some problems in estimating the amount of advection, two methods were proposed to do so in section 4.2.1 and the results are indeed very different. Nevertheless we also want to look at what happens if advection is not considered at all. Now, the percentage of correct forecasts seems to be a good measure of skill at first, and is shown in Figure 23. Obviously there is not much difference whether advection was derived using gradients and wind (ADV), computed as a total temporal change in the 3D column (d/dt) or was even neglected (NONE). But what is seen here is mainly the inability of the skill score for an unlikely event such as fog. Recall that the base rate is only 5 % when visibility has to be below 1000 m for a fog event. A probability of 100 % in the ensemble prediction was never achieved in the period considered, so this value is equivalent to never forecast any fog in which case 95 % of all forecasts are correct. Which seems great, but there is no skill required for this and it is about the same as saying it will not rain tomorrow in the Sahara desert. However, with higher base rates the PC can be a suitable verification score. In our case the increasing PC with increasing probability thresholds reflects the decline in false alarms. When the HSS is examined the ADV members score a bit higher, even though there is basically no skill. But also the HSS suffers from the base rate effect and for a real assessment of skill, ROC curves have to be considered. In Figure 24 these are given for different thresholds of observed visibility. Now it is evident that inclusion of advection significantly improves the forecast, if it is derived with [32], using gradients and wind. The other way of estimating advection as the total rate of change in the 3D column did generally not improve the performance, which means that local processes in the column overshadow advection. Especially the relatively coarse vertical resolution of the 3D model does not allow an accurate simulation of these local processes. But all possible inaccuracies are transferred into the 1D model when the total rate of change is used to determine advection.

The fact that inclusion of advection improves the forecast demonstrates the importance of advection. But because advection offends against the supposition of a 1D model (horizontal homogeneity), it is very questionable to use a 1D model in such cases. One might think of other tricks and experiments and eventually also include local knowledge and experience for the derivation of advection. This might further improve the forecast skill, but it seems more reasonable to seek a general solution, which would be a 3D fog model able to compute spatial interactions. A possible way to go is the 3D model described in the outlook of this work.

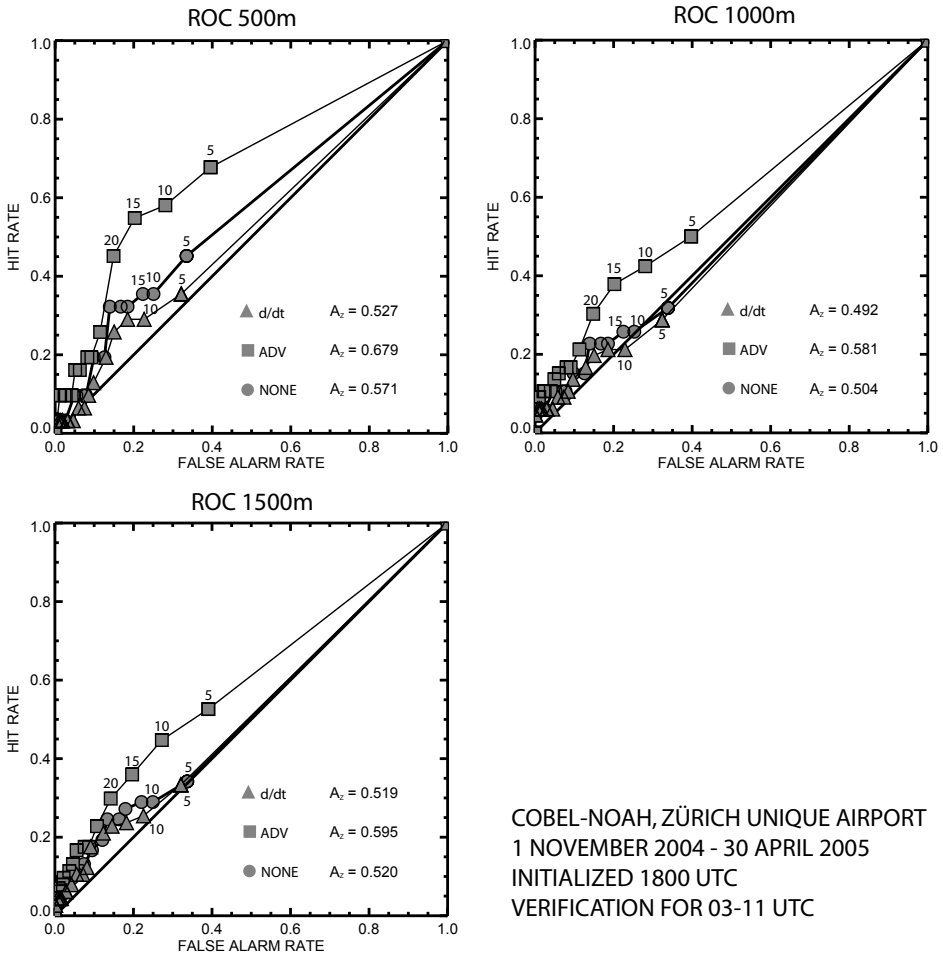

Figure 24: ROC for fog occurrence in the 1D ensemble for different types of computed advection; ADV = pure advection, d/dt = total rate of change in profile, CONTROL = no advection. Visibilities below 500 m, 1000 m and 1500 m were used as observational thresholds of an observed fog event, respectively. Numbers above the symbols indicate the ensemble forecast probability in percent that has to be exceeded in order to be classified as a modeled fog event. A_z indicates the area under the ROC.

6.4.2 Different initialization time

The forecast skill of every model decreases with increasing forecast length. Thus the later we initialize, the better our forecast should be for a given time window. This is what theory says, but it has to be considered that for example radiosonde data are only available at 0000 and 1200 UTC and that the model also needs some spin-up, which is generally easier in a mixed than in a stable boundary layer. Furthermore if there is already liquid water present at model start, it cannot be properly initialized due to the lack of liquid water content measurements. The time window we will use for verification stretches from 0300 to 1100 UTC, which is not only the time of likely fog occurrence but also very important for an airport.

In Figure 25 the ROC curves corresponding to different initialization times are shown. For the observations a visibility threshold of 1000 m is used and the advection is as outlined in the previous section. Obviously the 1500 UTC initialization has the highest skill. Starting at 1500 UTC, the model is able to spin-up and simulate the entire night with the cooling of the surface layer. Also the planetary boundary layer is generally well mixed, producing simple profiles of temperature and humidity. The latter allows for a more accurate initialization since the virtual profile is quite representative at that time (Figure 9). Radiosonde data after all do not seem to be very useful, primarily because they are available either too early (1200 UTC) or too late (0000 UTC). But again the importance of advection has to be pointed out, because basically at all times the members considering advection reach higher skill scores.

According to the ROC a low forecast probability of about 15 % has quite some skill. Even though the hit rate is only about 60 % the false alarm rate is significantly lower at 30 %. This does not seem very convincing, but since low visibilities are very difficult to forecast in the daily operations, such a performance might potentially increase the quality of the forecasts (Christoph Schmutz - MeteoSwiss, personal communication 2005).

6.4.3 Humidity profile

Even though data assimilation gives a good estimate of the temperature profile, the thermodynamic state of the atmosphere is not accurately defined without a reasonable humidity profile. But the latter causes some trouble, because the assimilation simply does not have a reliable data source to work on. Therefore the effects on forecast quality related to the uncertainties in the humidity profile were examined by deriving the whole set of ensemble members again, but this time with 10 % increased and decreased relative humidity profiles, respectively. By doing so every humidity class still has the same number of about 30 ensemble members. Note that the relative humidity was not allowed to exceeded 99 % to avoid an artificial creation of fog at the beginning. The result is summarized in Figure 26, where every panel represents a different visibility threshold used in the classification of observed fog events. If observed visibility has to be below 500 m to be considered as a fog event, there is not much difference between the control run and the two deviations, but as soon as the threshold, and thus also the number of events, is increased, a more humid profile yields better forecasts. This indicates that the assimilation is often too dry. If for example the 10 % probability is examined and the points are mentally interconnected to form a curve, the resulting slope is steeper than the diagonal. Thus the hit rate grows faster than the false alarm rate with increasing humidity, finally making a better forecast.

6.4.4 PAFOG and multi-model ensembles

For maximum comparability PAFOG is verified exactly the same way as COBEL-NOAH. Recall that PAFOG does not support the specification of advection. Never-

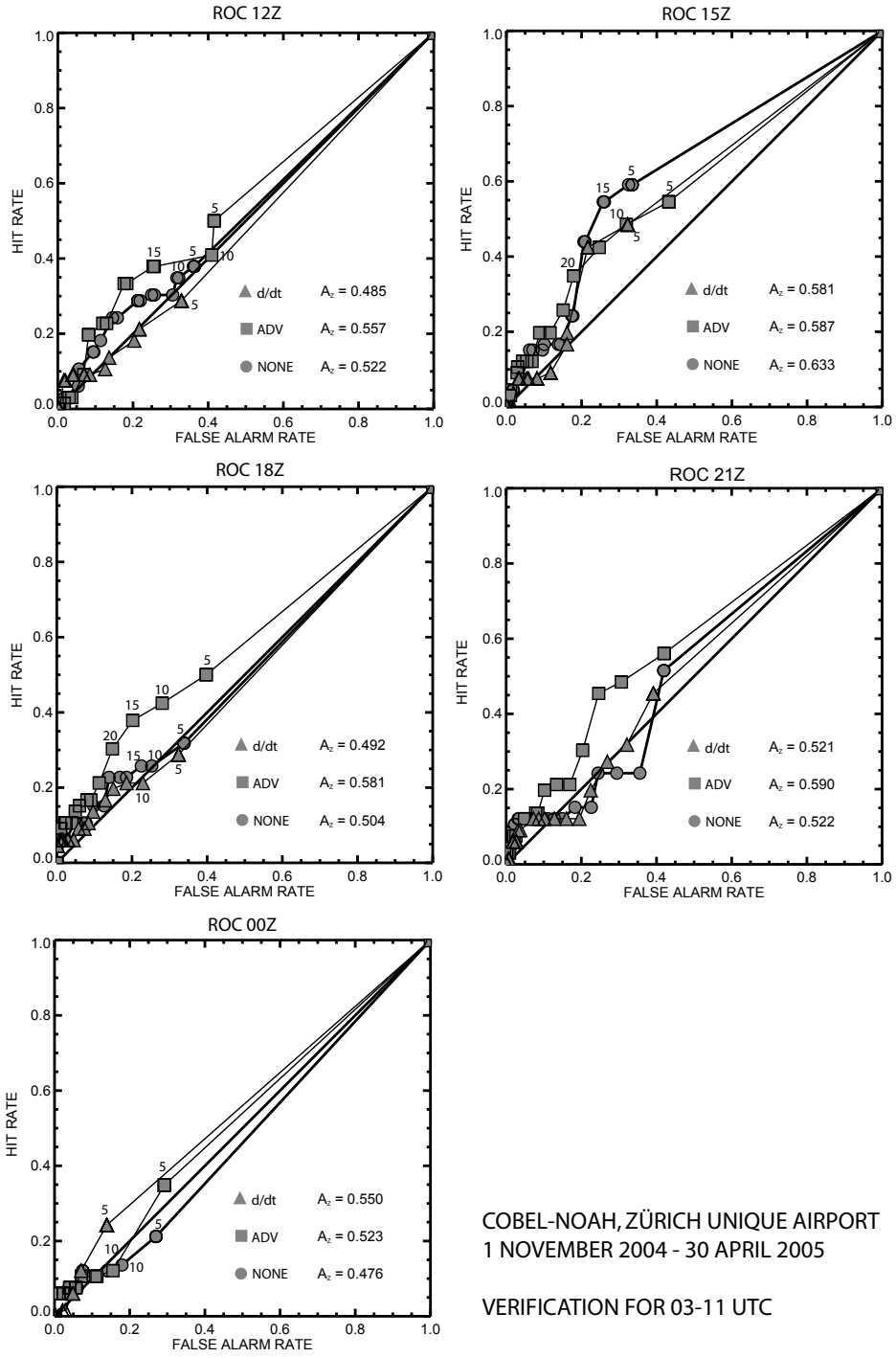

Figure 25: ROC for fog occurrence in the 1D ensemble for different types of computed advection and initialization times. Numbers above the symbols indicate the ensemble forecast probability in percent that has to be exceeded in order to be classified as a modeled fog event. The observational visibility threshold for a fog event was 1000 m. A_z indicates the area under the ROC.

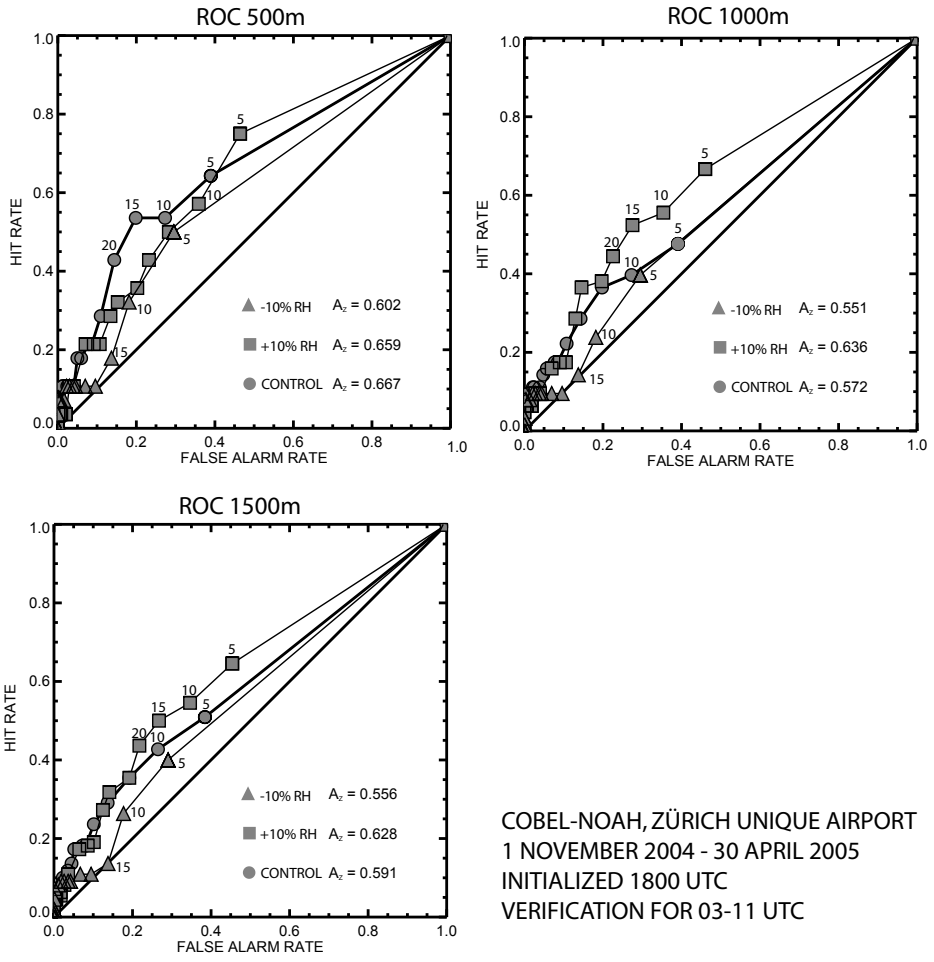

Figure 26: ROC for fog occurrence in the 1D ensemble for initializations with different relative humidity profiles. Visibilities below 500 m, 1000 m and 1500 m were used as observational thresholds of an observed fog event, respectively. Numbers above the symbols indicate the ensemble forecast probability in percent that has to be exceeded in order to be classified as a modeled fog event. The observational visibility threshold for a fog event was 1000 m. A_z indicates the area under the ROC.

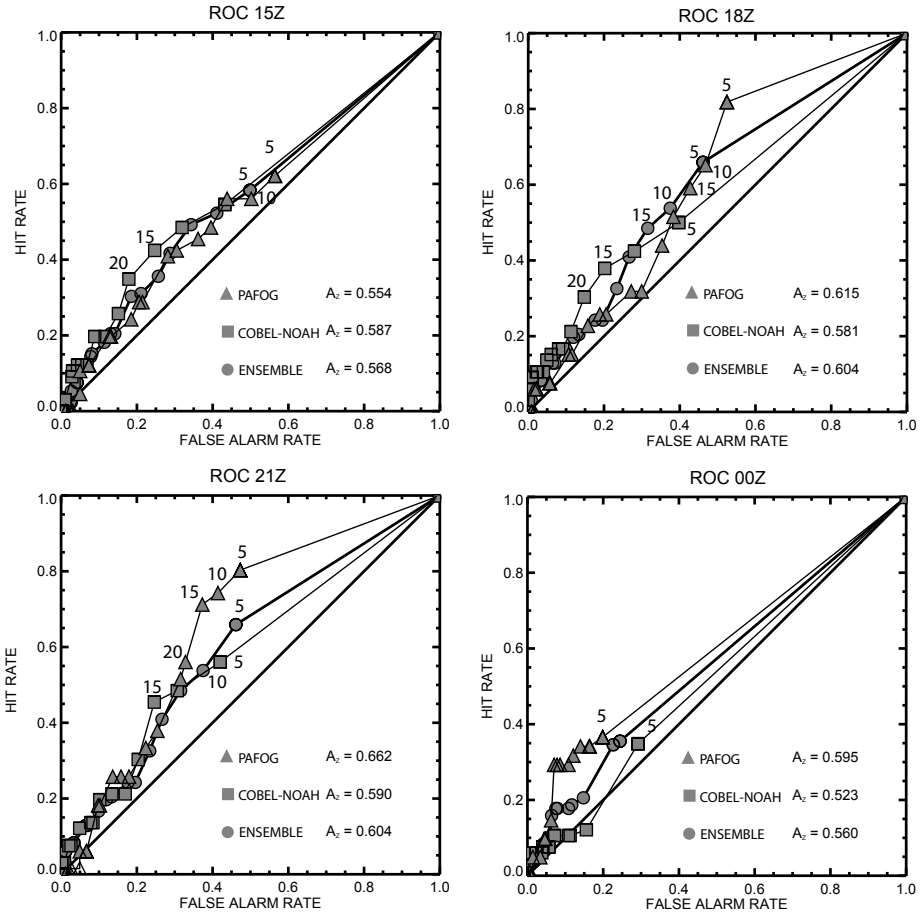

Figure 27: ROC for fog occurrence in the 1D ensemble for PAFOG, COBEL-NOAH and the multi-model ensemble at different initializations times. A visibility of 1000 m was used as observational thresholds of an observed fog event. Numbers above the symbols indicate the ensemble forecast probability in percent that has to be exceeded in order to be classified as a modeled fog event. Verification was done on an hourly basis between 03-11 UTC from 1 November 2004 until 30 April 2005 at Zürich Unique airport. A_z indicates the area under the ROC.

theless there are boundary conditions with a temporal resolution of one hour, which specify cloud cover above 2000 m. The verification scores are shown in Figure 27 for different initialization times. For comparison purposes the ROC curves of COBEL-NOAH are also shown. Furthermore the ROC of the mulit-model ensemble consisting of COBEL-NOAH and PAFOG was computed. From the COBEL-NOAH ensemble only the members with advection (ADV) are considered, since they are the most skillful according to Figure 24. If we now look at the skill score A_z of PAFOG we notice that unlike COBEL-NOAH the skill increases with later initialization time. This is related to a faster model spin-up of PAFOG. However for the 00 UTC initialization all models have less skill, because at that time fog has already formed in most cases. Note that an existing fog layer cannot be properly initialized due to the lack of observations. According to the ROC it are the low forecast probabilities that have the most skill and higher probabilities are not very useful for forecasting, as it is the case for COBEL-NOAH. Especially the 2100 UTC initialization has a remarkable discrimination between hit and false alarm rates when forecast probabilities between 5 and 15 % are used.

Interestingly the multi-model ensemble is not that different from each individual model. In fact it is rarely better than any individual model and most of the time either COBEL-NOAH or PAFOG reach a higher level of discrimination. Therefore the ensemble system should not just provide the results from the multi-model ensemble but also the individual model ensembles. Especially because PAFOG outperforms COBEL-NOAH with later initialization time, a forecaster can give more weight to PAFOG for late model initializations.

Finally a last comment relating to the skill discussion on the ROC curve. With an increasing probability threshold the skill decreases because the probability of detection approaches the false alarm rate for the models used. It is also important to note that fog is a relatively rare event. The consequence is that if e.g. false alarm rate and hit rate are equal, the absolute number of detected and false alarm forecasts is not equal, as can be seen from the mathematical definitions [69, 70]. With the low base rate for fog, the absolute number of false alarm forecasts will be bigger than the absolute number of detected fog events.

6.5 How could the forecast be improved?

The verification clearly demonstrated the importance of advection. Therefore application of a 1D model is theoretically wrong and in practice of questionable nature. Nevertheless a significantly higher skill than with currently used forecast methods was achieved. Improvements can be expected with better initialization. Mainly for humidity, available observations are insufficient and skill scores from the humidity ensemble underline the need for more measurements. However the observations required are vertical profiles, so that expensive remote sensing instruments are the only option for an airport. The models can always be improved to run e.g. with more complex microphysics or a better turbulence scheme. For the latter one has to consider that a 1D model cannot predict the wind vector because horizontal gradients are nonexistent. Therefore the outcome of any turbulence scheme is governed by the externally supplied wind speed. This can be a fine tool to perform some tests if the model is driven with observations, but for forecasting this seems not very promising. Hence, if the goal is forecasting, a model should require as little external forcing as possible and simulate any spatial interactions. An improvement from the model side is thus seen in the development of a 3D model.

7 A grid and subgrid scale radiation parameterization of topographic effects for mesoscale weather forecast models

This is the article of Müller & Scherer (2005), which was published in *Monthly Weather Review*.

Complex topography significantly modifies radiation fluxes at the Earth's surface. As spatial resolution of mesoscale weather forecast models increases, terrain effects on radiation fluxes induced by slope aspect, slope angle, sky view factor and shadowing also gain importance. A radiation parameterization scheme is hence designed to better represent these topographic influences to improve weather forecasts.

The grid and subgrid scale radiation parameterization scheme allows computation of radiation fluxes for each weather forecast model grid cell by considering arbitrarily fine resolved topography without degrading the model's computational performance. The proposed scheme directly computes mean fluxes for each model grid cell based on flux computations at full spatial resolution of a digital elevation model covering the model domain. Thus the scheme does not require a problematic computation of averaged topographic properties such as aspect angles. Furthermore the scheme has a non-local computation of sky view restriction and shadowing effects.

Case studies with the Nonhydrostatic Mesoscale Model (NMM) at resolutions of 4 and 2 km respectively and the parameterization based on a 1 km resolved elevation model, showed that effects of this parameterization are significant, and result in better temperature forecasts in complex terrain. RMS and mean error of 2 m temperature forecasts are generally improved by 0.5 to 1 K. At night, the consideration of restricted sky view leads to a temperature increase between 0.5 and 1.5 K along valleys. During clear sky daytime, this warming is of the same magnitude for grid cells containing slopes exposed to the sun. Under overcast conditions, RMS error is reduced by 0.2 to 0.5 K. In wintertime shadows reduce temperatures in valleys by 0.5 to 3 K during daytime.

7.1 Introduction

Radiation, the main source and sink of energy at the Earth's surface, is significantly influenced by topography. Local slope and aspect angles considerably modify the amount and daily course of downwelling short-wave radiation (e.g. Whiteman et al. (1989)). For instance, a time lag for the maximum downwelling short-wave radiation of about two hours, and intensities increased by 200 W m^{-2} were observed on inclined surfaces in an alpine valley (Matzinger et al. 2003). Simulated diurnal averages of downwelling short-wave radiation showed variations as large as 450 Wm^{-2} for the Tekapo watershed (Oliphant et al. 2003). Sky view restriction increases downwelling longwave radiation and generates spatial variability of diffuse radiation (Matzinger et al. (2003), Dubayah & Loechel (1997)). Scherer & Parlow (1994) have thoroughly demonstrated the consequences of terrain induced modifications of radiation fluxes on the energy balance, and hence on snowmelt and snow-hydrology for a drainage basin in Svalbard. In their study they showed that specific topographic conditions led to significantly increased solar radiation on the west-exposed slopes of the studied drainage basin in NW-Spitsbergen, which is the main reason for intensified snowmelt and subsequent meltwater runoff measured in the field. The parameterization scheme used in their spatially distributed local scale radiation model forms the basis of the one developed in our study, but it has to be noted that many other models for the treatment of radiation in complex terrain have been developed over the last years (e.g. Dozier (1980), Dozier (1989), Duguay (1993), Dubayah & Rich (1995), Kumar et al. (1997)). Our parameterization is specially tailored for mesoscale models, in a way that it deals with coarse grid resolutions used in mesoscale weather forecast

Table 7: Topographic effects considered by mesoscale models used for numerical weather prediction. The term shadow is used for shadows cast by surrounding terrain, not self-shading of a grid cell, which is considered within slope and aspect.

Model	Considered topographic influence on radiation
ARPS	slope, aspect, shadow
Bolam (Bologna limited area model)	none
Hirlam (High Resolution Limited Area Model)	none
LM / aLMo (Lokal-Modell)	none
MASS (Mesoscale Atmospheric Simulation System)	slope, ascpect
MC2 (Mesoscale Compressible Community model)	none
Meso ETA	none
Meso-NH	slope, aspect
MM5 (Fifth-generation Pennsylvania State UniversityNational Center for Atmospheric Research Mesoscale Model)	slope, aspect
NMM (Nonhydrostatic Mesoscale Model)	none
NMS	none
RAMS (Regional Atmospheric Modeling System)	slope, aspect
RUC (Rapid Update Cycle)	none
UK Unified Model	none
WRF (Weather Research and Forecast model)	none

models, severe computational cost restrictions as well as ease of implementation and portability.

To the knowledge of the authors, the majority of mesoscale weather forecast models only consider a few or none of topographic influences on radiation. A survey of considered topographic effects for popular mesoscale models in operational use is given in Table 7. It can be seen that most models only include slope aspect and slope angle, which in fact are the most important effects, as mentioned by Oliphant et al. (2003). However, at typical grid resolutions of a few kilometers, such effects decrease due to the flattening of slopes. Dubayah et al. (1990) analyzed topographic modulation of clear sky irradiance using a parameterization very similar to the one of this study. They found that variance and spatial autocorrelation of simulated radiation changed with sun angle and grid spacing. As grid spacing increased, variance decreased and spatial autocorrelation increased. Recently, shadow effects were included into the Advanced Regional Prediction System (ARPS), and its importance was demonstrated (Colette et al. 2003). Mesoscale weather forecast models are usually running at spatial resolutions of a few kilometers, thus resolving larger valleys on the grid. However, sloping surfaces are flattened, and topographic details are not resolved, so that radiation computation based on the grid scale topography has little effect. The proposed parameterization scheme considers topographic effects on radiative fluxes by using an arbitrarily fine resolved topography. Thus the scheme's accuracy is not restricted by the spatial resolution of the weather forecast model grid but only by the one of an available digital elevation model (DEM). If the DEM used for radiation computations has a higher resolution than the model grid, the model grid cell radiation fluxes are representative means based on the higher resolution fluxes as computed using the DEM grid.

Treatment of radiation in the described way seems to be computationally expensive, but most computations can be done prior to model execution, leaving only a few multiplications for the weather forecast model during time integration, so that

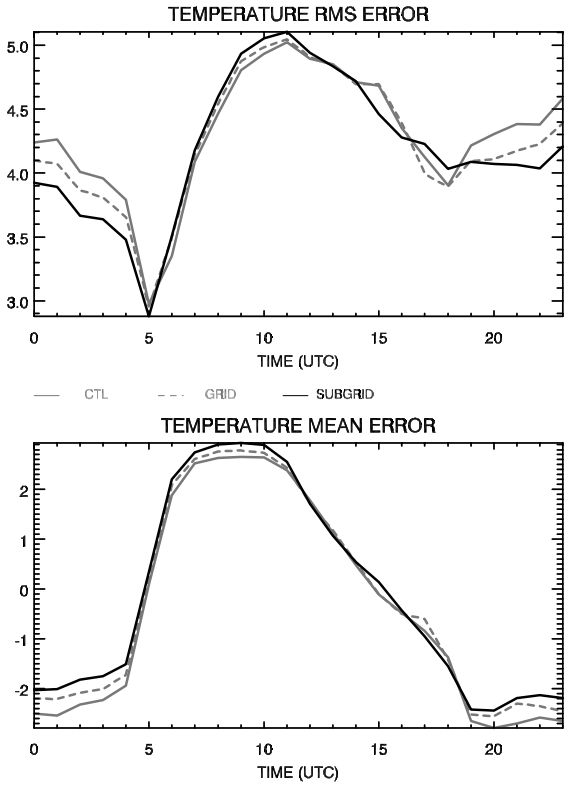

Figure 28: Verification of 2 m temperatures for alpine stations on 21 June 2003. The NMM was run at 4 km resolution using no topographic radiation parameterization (CTL), the parameterization based on the model topography (GRID) and the parameterization based on a 1 km DEM (SUBGRID), respectively. Mean and RMS errors are given in Kelvin.

computational costs are negligible.

In the first part of this article we derive the parameterization scheme and present the preprocessing steps. In the second part we demonstrate the skill improvement for air temperature forecasts due to the new parameterization scheme under clear sky summer and winter conditions, under overcast conditions, as well as for a whole month of strongly varying cloud conditions. Therefore, the scheme was implemented into the Nonhydrostatic Mesoscale Model (NMM) of NOAA/NCEP (Janjic et al. (2001) and Janjic (2003)).

7.2 Parameterization scheme and data

7.2.1 Mathematical description

The parameterization affects five radiation fluxes at ground level, namely direct downwelling short-wave ${}^{\downarrow}E_{s,dir}$, diffuse downwelling short-wave ${}^{\downarrow}E_{s,diff}$, upwelling short-wave ${}^{\uparrow}E_s$, upwelling long-wave ${}^{\uparrow}E_l$, and downwelling long-wave ${}^{\downarrow}E_l$. These fluxes are computed by the mesoscale forecast model considering the altitude of each grid cell as provided by a DEM of same spatial resolution as the model grid. The same symbols but marked by $*$ are used for the parameterized radiation fluxes. Depending on the details of the respective model, radiative transfer computations consider the actual atmospheric conditions, but often assuming flat terrain, i.e. the surface of each grid cell is horizontally oriented, and also disregarding neighborhood effects like

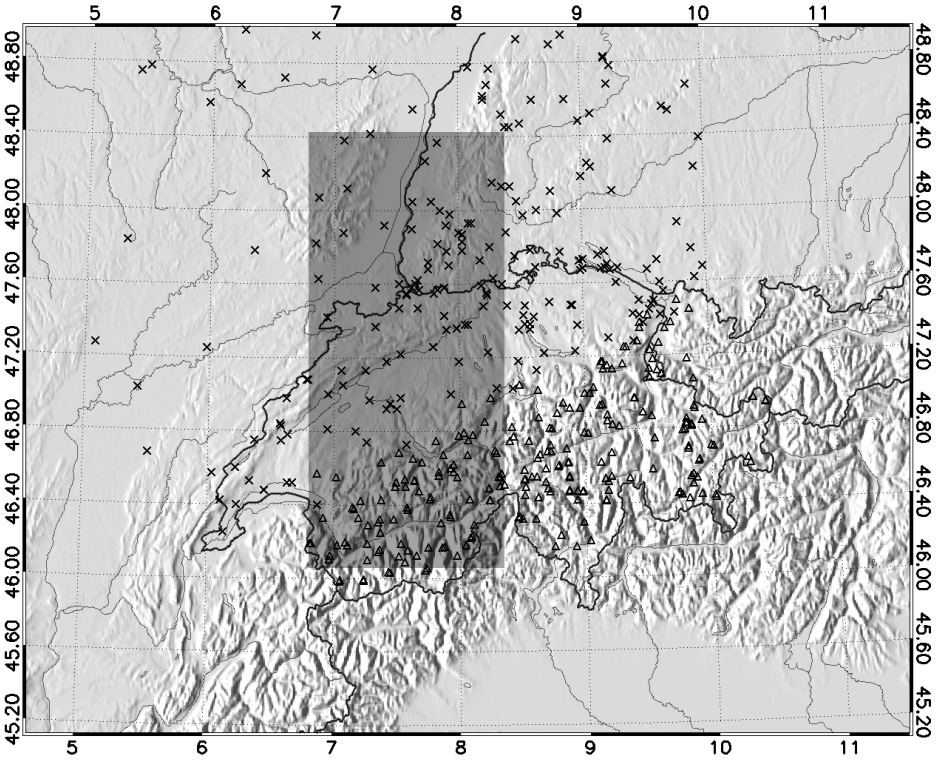

Figure 29: Map of model domain showing locations of alpine (Δ) and non-alpine (x) measurement stations. The dark shaded region indicates the subset used in Figures 30 and 36 to 38. Latitude and longitude coordinates are given on the borders. The shaded topography was generated from the 1 km GTOPO30 DEM used for the parameterization.

sky view restrictions and shadowing. The NMM, used in this study, computes shortwave radiation absorption, reflection and scattering in the model atmosphere based on Lacis & Hansen (1974). Absorption for water vapor, O_3 and CO_2 are computed separately and over single bands in the UV/visible and NIR portions of the solar spectrum, each representing 50 % of incoming solar energy. The basic surface albedo taken from climatology is modified to take into account the actual local conditions at the surface. The parameterization for long-wave radiative transfer in clear-sky conditions was developed at GFDL (Fels & Schwarzkopf (1975), Schwarzkopf & Fels (1985), Schwarzkopf & Fels (1991)). Radiation interactions with clouds are computed for each vertical layer (Harshvardhan et al. (1989), Hong et al. (1998), Slingo (1987), Xu & Randall (1996)). It has to be noted that the proposed parameterization of topographic effects takes radiative fluxes as externally computed input, which allows easy implementation in different mesoscale models.

The dependence of $|{}^{\downarrow}E_{s,dir}|$ on slope angle α was formulated by Kondratyev (1977) with help of the geometry factor $\cos(\alpha)$:

$$\cos\alpha = \cos\theta_N \sin\theta_S + \sin\theta_N \cos\theta_S \cos(\phi_S - \phi_N) \tag{74}$$

where α is the angle between the unit vector of the solar direct beam and the normal vector of the surface, ϕ_N is the slope aspect, and θ_N is the slope angle. The position of the sun is given by the sun elevation angle θ_S and sun azimuth angle ϕ_S.

A correction factor f_{cor} is derived for ${}^{\downarrow}E_{s,dir}$ combining the effects of slope angle

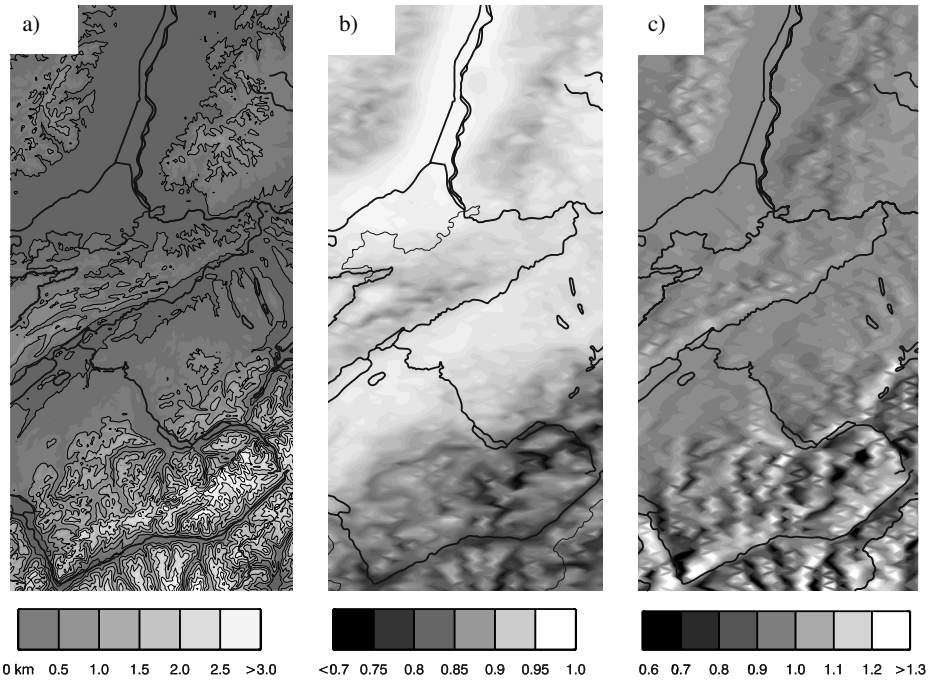

Figure 30: a) Topography at 1 km resolution with height contours every 500 m for the dark shaded area in Figure 28. The most complex topography is found in the South and is part of the Swiss Alps. b) Sky view factor aggregated to the 2 km NMM Grid. A significant reduction of sky view appears within alpine valleys in the South (lower part). c) Direct short-wave radiation correction factor f_{cor} for 0900 LST 22 June 2003, on the 2 km NMM grid.

and aspect, geometric enlargement of a sloping surface, and shadowing. By using [74], the effective direct downwelling short-wave radiation ${}^{\downarrow}E^{*}_{s,dir}$ on an inclined surface can be computed as:

$$
\begin{aligned}
{}^{\downarrow}E^{*}_{s,dir} = {}^{\downarrow}E_{s,dir} \left(\frac{1}{\sin\theta_S}\right)\left(\frac{1}{\cos\theta_N}\right) \text{ mask }_{shadow} \\
[\cos\theta_N \sin\theta_S + \sin\theta_N \cos\theta_S \cos(\phi_S - \phi_N)]
\end{aligned}
\tag{75}
$$

The geometric surface enlargement - that is, the ratio between the actual area of a sloping surface and its projected area as given by the spatial resolution of the underlying DEM (Scherer & Parlow 1994) - is given by the third factor in [75]. Note that surface enlargement can only be applied for the computation of ${}^{\downarrow}E^{*}_{s,dir}$, since it is the only flux showing directional dependence by definition.

The treatment of shadows is done by introducing a binary shadow mask:

$$
\text{mask }_{shadow} = \begin{array}{ll} 0 & \theta_S < \theta_{h,\phi_S} \\ 1 & \text{elsewhere} \end{array}
\tag{76}
$$

where θ_{h,ϕ_S} is the horizon angle towards the sun azimuth angle ϕ_S. The great advantage of this approach is its high computational performance, since only a logical operator has to be applied to each grid cell instead of ray tracing or radiosity computations. The time consuming part of the computation, i.e., the derivation of the horizon angles, is carried out as part of the preprocessing. According to [75] it is possible to compute ${}^{\downarrow}E^{*}_{s,dir}$ by simply multiplying ${}^{\downarrow}E_{s,dir}$ with a correction factor

f_{cor}:

$$^{\downarrow}E^{*}_{s,dir} =^{\downarrow} E_{s,dir} f_{cor} \tag{77}$$

where, after simplification of [75]:

$$f_{cor} = \text{ mask }_{shadow} \left[1 + \frac{\tan\theta_N}{\tan\theta_S} \cos(\phi_S - \phi_N) \right] \tag{78}$$

Here, $^{\downarrow}E_{s,dir}$ is computed by the mesoscale model, and is representative for a single model grid cell usually covering an area of several square kilometers. Computing topographic parameters from the mesoscale model topography, as needed by [75], would be a crude approximation producing values for f_{cor} close to 1. A better approach is to compute the radiative fluxes based on a high-resolution DEM covering the whole model domain. Because of the multiplicative nature of [77] it is possible to simply compute a mean f_{cor} and multiply it with mean $^{\downarrow}E_{s,dir}$ to obtain mean $^{\downarrow}E^{*}_{s,dir}$. Figure 28 illustrates the benefits of computing subgrid fluxes based on a 1 km resolved DEM rather than using the 4 km resolved mesoscale model topography. Shown are verifications of 2 m temperature forecasts for clear sky conditions on 21 June 2003. The grid scale resolved parameterization lies between the unparameterized control run and the subgrid parameterized run. Thus the sugbrid computation of fluxes doubles the effect of a gridscale parameterization in the current configuration. Both parameterizations have the same memory requirements and computational costs during time integration. The verification procedure is described later in Section 3a.

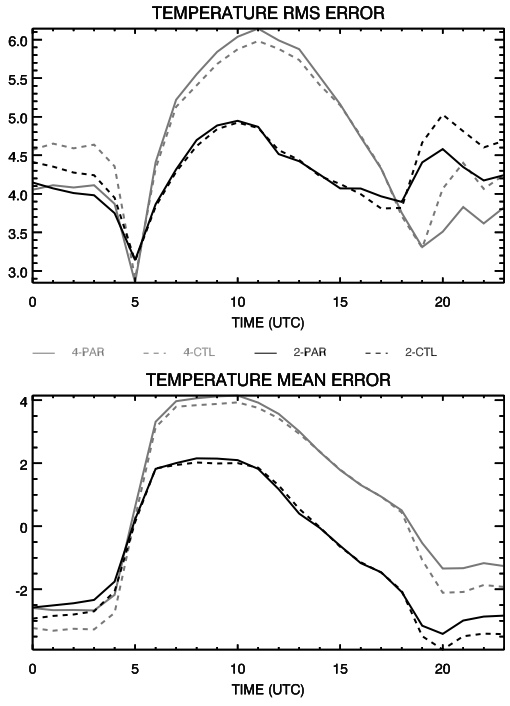

Figure 31: Verification for alpine stations for 22 June 2003. Shown are unparameterized control runs (CTL) and parameterized runs (PAR). Grey curves are the 4 km resolution NMM simulation and black curves are the 2 km resolution NMM simulation. A perfect model has an RMS and mean error of 0 K. LST is UTC+2.

To save computational resources during model execution, mean f_{cor} can be computed diagnostically prior to model execution for each hour t of the model run period

based on all n grid points of the DEM that cover a mesoscale model grid cell:

$$\overline{f_{cor}^{t}} = \frac{1}{n} \sum_{i=1}^{n} f_{cor}^{t,i} \tag{79}$$

The parameterization also considers topographic effects on ${}^{\downarrow}E_{s,diff}$ by restricted sky view. Effective diffuse downwelling short-wave radiation ${}^{\downarrow}E_{s,diff}^{*}$ is approximated by:

$${}^{\downarrow}E_{s,diff}^{*} = {}^{\downarrow}E_{s,diff} f_{sky} + {}^{\uparrow}E_{s}(1 - f_{sky}) \tag{80}$$

where f_{sky} is the sky view factor. The last term in [80], which is the part of downwelling short-wave radiation reflected by adjacent areas, assumes homogenous albedo equal to the one of the model grid cell. This is a simplification, since the reflected short-wave radiation may originate from a neighboring grid cell having a different albedo. Furthermore, albedo depends on the relative incident angle, which can vary across the entire range at different locations and times throughout the day. However, the effect of this simplification is usually small, and computational performance would significantly decrease by a more sophisticated treatment.

The flux ${}^{\downarrow}E_{l}$ may be modified by radiation emitted by adjacent areas. In analogy to ${}^{\downarrow}E_{s,diff}^{*}$, this effect is considered by

$${}^{\downarrow}E_{l}^{*} = {}^{\downarrow}E_{l} f_{sky} + {}^{\uparrow}E_{l}(1 - f_{sky}) \tag{81}$$

Usually, the surface is warmer than the sky, thus ${}^{\downarrow}E_{l}$ in complex topography is increased in most cases by the parameterization given by [81] in areas of restricted sky view.

7.2.2 Preprocessing

The required topographic parameters are derived as spatially distributed properties from a DEM of arbitrary spatial resolution or coordinate system. Due to their stationarity they need to be computed only once. The computations are following the scheme described by (Scherer & Parlow 1994). The only modification was done with respect to DEMs that use non-metric coordinate systems like the GTOPO30 data set used in our study.

Since $\theta_{h,\phi}$ and f_{sky} are non-local properties, they have to be computed considering a neighboring area for each grid element of suitable horizontal extent. In our study, we used an area of 40 km by 40 km, where the grid element, for which the properties are to be computed, is located in the center. The grid elements of the DEM within this area are resampled from the original grid to a Cartesian grid using a local orthographic projection, also taking into account the earth's curvature. The procedure is thus applicable for any model domain of the earth.

The variables ϕ_N and θ_N are local properties, and are derived from the resampled neighborhood by locally fitting a biquadratic surface through the central grid element and its eight neighbors. To save computational and disk storage resources, $\theta_{h,\phi}$ is computed for 24 discrete azimuth angles ϕ, that is, using an azimuthal step width of 15°. For other azimuth directions the corresponding horizon angle can be obtained with sufficient accuracy by linear interpolation, and f_{sky} is computed from $\theta_{h,\phi}$ after Dozier & Marks (1987):

$$f_{sky} = 1 - \frac{1}{24} \left[\sum_{j=0}^{23} \sin(\theta_{h,j15^{\circ}}) \right] \tag{82}$$

For the dark shaded subset in Figure 29, the spatial distribution of f_{sky} is shown in Figure 30b. Darker shades indicate smaller values of f_{sky}. The corresponding topography is contoured in Figure 30a. In regions of complex terrain, f_{sky} is significantly

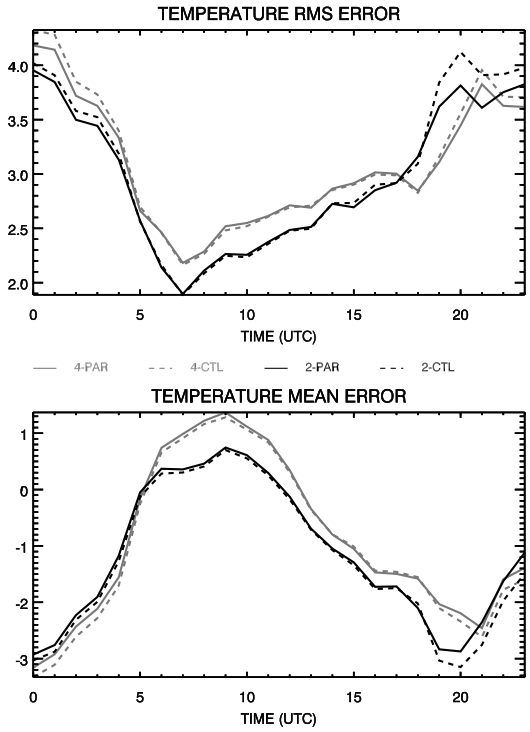

Figure 32: Same as Figure 31 for non-alpine stations.

reduced, and can be as low as 0.7. Note that f_{sky}, computed from 1 km resolved GTOPO30 data, were aggregated to the 2 km grid of NMM.

Temporal discretization of the time dependent parameters can be achieved with sufficient accuracy using a resolution of one hour. To save computational resources during time integration of the mesoscale model, the computation of θ_S, ϕ_S, mask $_{shadow}$ and f_{cor} for each grid cell is carried out prior to model execution in a preprocessor, as well as the subsequent aggregation of f_{cor} to the model grid. Because of the preprocessor, the only computational task left for the mesoscale model itself is application of Eqs. [77], [80] and [81].

Figure 30c shows the spatial distribution of the direct short-wave correction factor f_{cor} for a morning situation. It is evident that eastward sloping surfaces have larger f_{cor} values than westward sloping surfaces, where f_{cor} is smaller than 1. On horizontal surfaces, f_{cor} becomes unity. At noon, southward sloping surfaces have higher f_{cor} and in the evening westward sloping surfaces are subject to higher values of f_{cor}.

7.2.3 Model area and measurement data

Case studies were carried out in a domain shown in Figure 29, covering Switzerland and surroundings, thus including terrain ranging from very complex in the Swiss Alps with elevations over 4000 m a.s.l. to very flat in the Rhine plane of southern Germany. Figure 30a shows the topography for the dark shaded subset of the model domain indicated in Figure 29. Data of about 400 automatic weather stations, half of those located in the Swiss Alps, is available for the entire model domain, thus yielding a statistically significant number of observations. From all stations hourly mean temperatures could be obtained, allowing a comparison with model data at a temporal resolution of one hour.

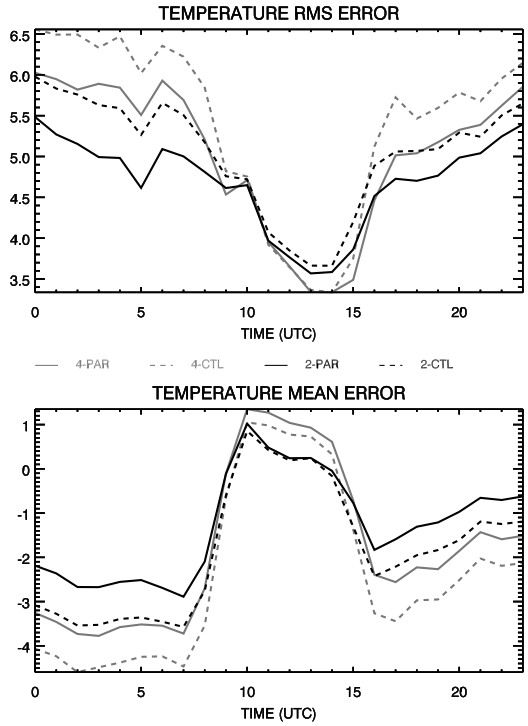

Figure 33: Same as Figure 31, but for a wintertime clear sky situation on 24 December 2003. LST is UTC+1.

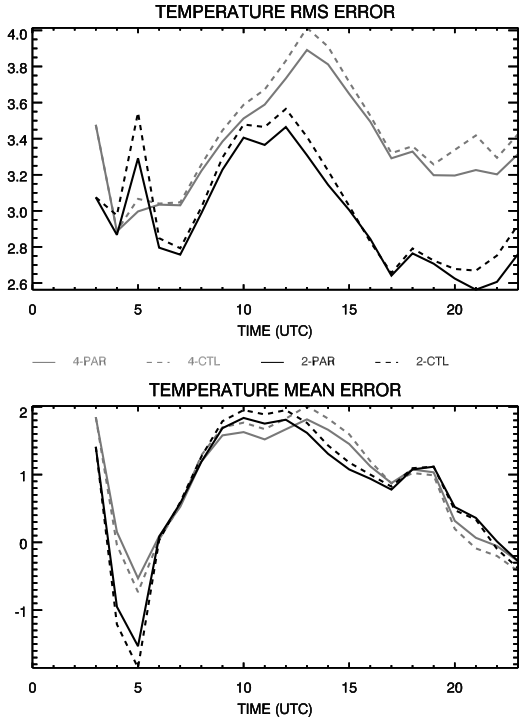

Figure 34: Same as Figure 31, but for an overcast situation on 20 October 2003. LST is UTC+2.

7.3 Results

The proposed radiation parameterization was implemented into the Nonhydrostatic Mesoscale Model (NMM) of Janjic et al. (2001) and Janjic (2003). NMM is used for operational weather forecasts at NOAA/NCEP. Effects of the parameterization are analyzed using modeled and observed air temperatures at 2 m above ground. Air temperature is strongly correlated to radiation, and is measured at every observing site. In addition, air temperature is an important forecast variable that is expected to be improved by application of the parameterization scheme.

7.3.1 Verification results

Results for clear sky conditions on 22 June 2003 and 25 December 2003, as well as completely overcast conditions on 20 October 2003 are presented separately to discuss the different effects of the parameterization. In order to evaluate the benefits of the parameterization for operational weather forecasts under different synoptic conditions, parallel runs where carried out for September 2004. All situations were simulated with full model physics but with an unparameterized topographic radiation control run (CTL) and with the new radiation parameterization (PAR) at spatial resolutions of 4 km as well as 2 km. Forecast initialization was done using NOAA 1° resolution GFS data and one way nesting via a 22 km resolution NMM run covering Europe.

As expected, the effects of the radiation parameterization are most pronounced in complex terrain (cf. Whiteman et al. (1989); Colette et al. (2003); Chow et al. (2004)). Thus, available observations are divided into an alpine and a non-alpine group as shown in Figure 29. The alpine group consists of all stations within the Alps and represents measurements from the most complex topography. The non-alpine group consists of the remaining stations. Verification scores are computed on an hourly basis by comparing all measured temperatures with the corresponding modeled temperatures. Note that temperature forecasts for stations were taken from the closest model grid point without further processing. Due to the very complex small scale topography of the Alps, elevation differences between model grid cells and actual station heights are around 400 m for alpine stations, which contributes to the large mean and RMS temperature errors. However these values are typical for temperature forecasts of the Alps. The high resolution model (aLMo), running at the Swiss national weather service (MeteoSwiss) obtains RMS errors, for stations located above 1500 m a.s.l, of 2.5 to 3.5 K in summer and 4 to 6 K in winter, respectively (Schubiger 2001, F. Schubiger 2004, personal communication).

In Figure 31 and 32 verification results of 22 June 2003 are shown for alpine and non-alpine stations, respectively. It can be seen that for both resolutions the new parameterization has most impact at night and of course is more pronounced for alpine stations, located in complex terrain. The positive effect at night is caused by reduced sky view considered in the parameterization scheme. Long-wave radiative loss of energy is reduced in valleys yielding warmer temperatures with less negative mean errors and smaller RMS errors. Thus the parameterization weakens the cold bias, which together with an underestimation of nocturnal ${}^{\downarrow}E_l$, is also observed in the operational models, run at NOAA/NCEP (Z. Janjic 2004, personal communication). During daytime the parameterization seems to have little influence. At 4 km resolution the parameterization slightly increases the forecast error during daytime. This is mainly a problem of forecast verification and model resolution. The parameterization computes a representative mean of net radiation for a model grid cell including slopes and plains which covers 16 or 4 km^2, respectively. Measurement stations, however, are mostly located in flat areas of a grid cell, so that radiation and resulting temperature computed for a flat area, as it is done in the unparameterized version, are more accurate for this particular part of the grid cell, but not for the whole area of the grid cell. For example, if a grid cell contains large slopes facing the

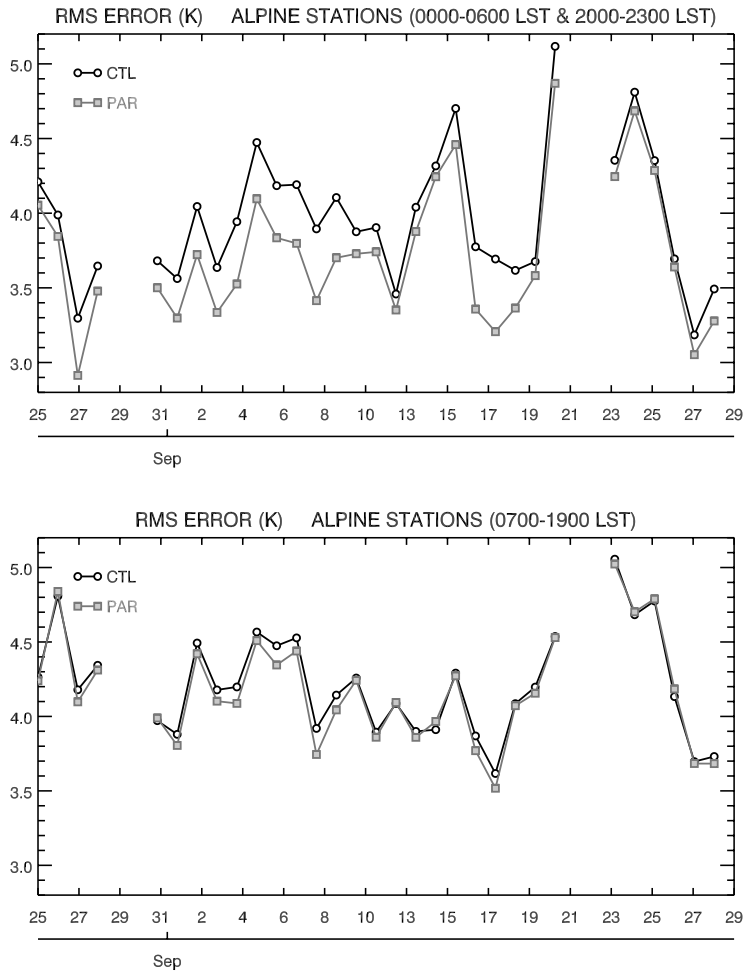

Figure 35: Verification of 2 m temperatures for alpine stations under different synoptic conditions in August and September 2004. Shown are unparameterized control run (CTL) and parameterized run (PAR) for night (upper panel) and daytime (lower panel). The two gaps are due to missing initial conditions. The mesoscale model (NMM) was run at a resolution of 4 km.

sun, the parameterization yields more short-wave radiation that is absorbed, and thus higher temperatures for the whole grid cell. But the station located in the flat terrain part of the grid cell measures a lower temperature, which is not representative for the entire grid cell in complex terrain. At higher resolutions, observed and modeled areas correspond better, and the verification of parameterized and unparameterized runs converge, which is evident from the 2 km model resolution. Reduced sky view exists on the valley bottom as well as on sloping surfaces, thus measurement stations in valleys capture the phenomenon, and the above mentioned verification problem is less severe at night.

Another clear sky situation is shown in Figure 33. Those wintertime verification results are very similar to the summer situation, with most RMS and mean error reduction at night. Again the benefit of the parameterization is the same for both resolutions of the mesoscale model.

Completely overcast conditions on the entire domain prevailed on 20 October 2003. After initialization the parameterized and control runs diverge only little when

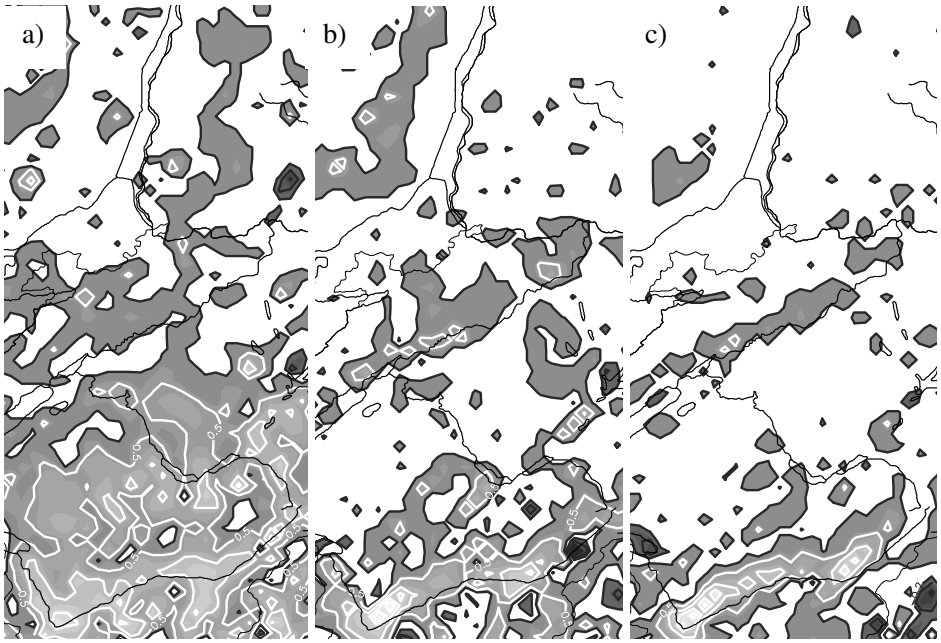

Figure 36: Modeled 2 m height temperature differences (4PAR-4CTL) for the clear sky situation on 22 June 2003 at 4 km model resolution. White and black contour lines inside the shaded areas are 0.5 K increments. White contours and bright filling indicates a warming due to the parameterization, whereas darker filling and black contours indicate a cooling effect. White fields, surrounded by black contours are differences between 0 and 0.2 K and thus regions where the parameterization has negligible effect. a) Differences at 0500 LST. b) Differences at 0900 LST. c) Differences at 1200 LST.

compared to the clear sky situation. The parameterization has only a small positive impact, but both during day- and nighttime, which is caused by inclusion of the sky view factor. The importance of sky view obstruction under overcast conditions and the resulting spatial variability has been demonstrated by Dubayah & Loechel (1997) for a portion of the Rocky Mountains. It has to be noted that this situation is a rarely occurring "worst case" scenario, since for large model domains there are almost always regions and time windows with no or very little clouds, where the parameterization has more effect. The mean error in Figure 34 shows that during daytime the improvement is due to cooler temperatures in the parameterized run. This can be explained by looking at equation [80], where ${}^{\downarrow}E_{s,diff}$ is reduced by multiplication with the sky view factor, and where, under overcast conditions, the second term is only dependent on ${}^{\downarrow}E_{s,diff}$ and albedo. Thus the net effect is a decrease of incoming short-wave radiation. Longwave downwelling radiation is increased due to restricted sky view, but less than under clear sky conditions. This is due to the presence of relatively warm clouds that decrease the contrast between cold clear sky temperatures and warmer terrain (cf. Matzinger et al. (2003)). The shortwave part of the parameterization dominates during daytime, whereas at night, only the longwave part, with its above mentioned reduced effect, is present. A verification problem associated with direct downwelling shortwave radiation does not appear under cloudy conditions.

Finally, the parameterization was tested under strongly varying synoptic conditions, such as stable high pressure with warm temperatures, many passages of fronts and even snowfall in the Alps, as well as unstable situations with development of local

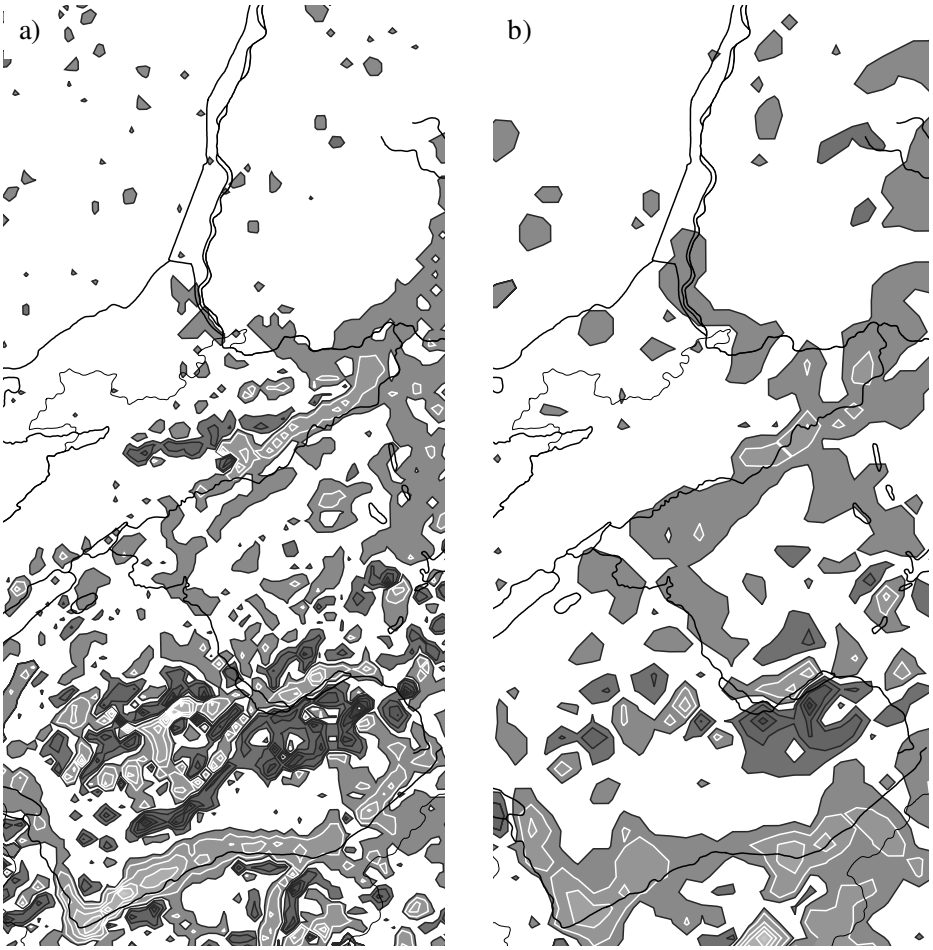

Figure 37: Modeled 2 m height temperature differences (PAR-CTL) for the clear sky situation at 1100 LST on 25 December 2003 for 2 km and 4 km model resolution, respectively. The upper left quadrant was cloudy. Contours are as specified in Figure 36. a) NMM simulation with 2 km resolution. b) NMM simulation with 4 km resolution.

thunderstorms. The chosen period starts on 25 August and ends on 29 September 2004. Figure 35 shows RMS errors computed for day- and nighttime, respectively. The verification is done using hourly data between 0700 and 1900 LST for daytime and the remaining 11 hours for nighttime, by comparing all measured temperatures with the corresponding modeled temperatures. As can be seen, the parameterization reduces the temperature error throughout the time period, especially at nighttime. During daytime, the verification is again less representative, and parameterization effects are thus much smaller.

7.3.2 Spatial impacts of the parameterization on modeled air temperature

Local deviations of the temperature field caused by a more detailed treatment of radiation are larger than the spatially integrated verification results suggest. Figures 36 to 38 illustrate spatial impacts of the parameterization for selected conditions and times. Depicted are differences between the parameterized and the control run for

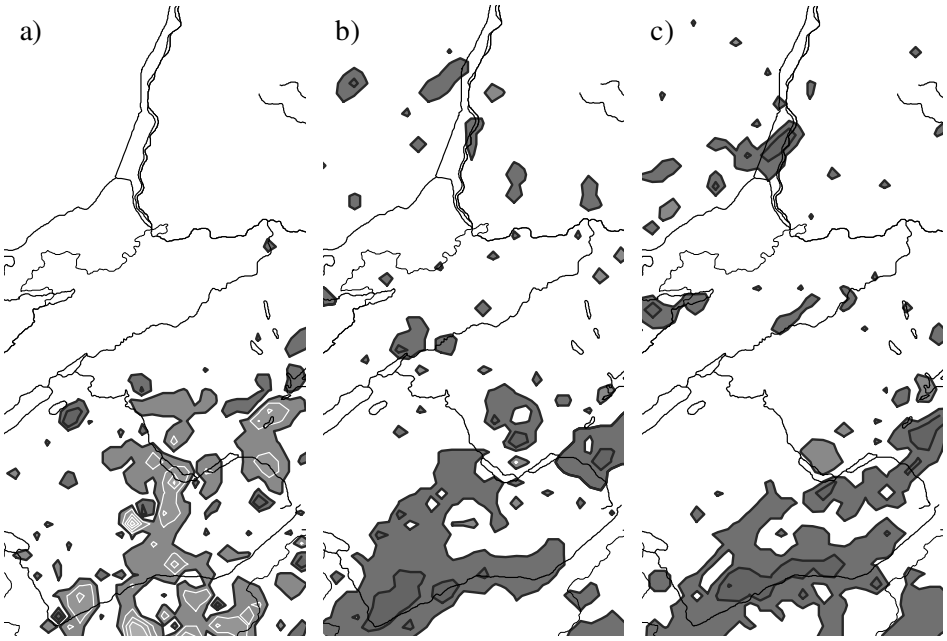

Figure 38: Modeled 2 m height temperature differences (4PAR-4CTL) for overcast conditions on 20 October 2003 at 4 km model resolution. Contours are as specified in Figure 36. a) Differences at 0100 LST. b) Differences at 1400 LST. c) Differences at 1600 LST.

clear sky conditions on 22 June 2003 and 25 December 2003, as well as overcast conditions on 20 October 2003, respectively. For clarity, only a small spatial subset marked as dark shaded area in Figure 29 is used here. In the clear sky situation of Figures 36b and 36c the differences are mainly caused by ${}^{\downarrow}E^{*}_{s,dir}$. Higher values of f_{cor} in Figure 30c correspond well with warmer areas in Figure 36b. The correlation is, however, not perfect, since temperature is also controlled by dynamic processes such as turbulent exchange and advection. The spatial correspondence between temperature and f_{cor} exists for the other hours of the day, depicting the sun's movement.

In Figure 37 clear sky winter conditions at 1100 LST 25 December are shown for 2 and 4 km model resolution, respectively. In addition to warmer regions, likewise modeled for the summer situation, significantly cooler temperatures can also be found in valleys. The cooling is in the range of 0.5 to 3 K and caused by shadows, which are much longer in wintertime. It can be seen that the 4 km resolution develops the main patterns of the higher resolution run. However, a lot of detail in smaller valleys is missing, and the effects of the parameterization are weaker at the coarser resolution.

At night, when short-wave radiation is not present, restricted sky view (Figure 30b) leads to reduced net loss of longwave radiation, which results in warmer air temperatures (Figure 36a). As for shortwave radiation, the effects are best seen in the Alps. Mainly along the big valleys, where sky view factors are generally smaller, temperature increased more than 1 K.

For overcast conditions, spatial patterns are caused by the sky view factor. Figure 38 shows temperature differences at 1400, 1600 and 0100 LST on 20-21 October 2003, respectively. As can be seen by comparison with Figure 30b, the parameterization leads to the already discussed cooling during the day in regions of restricted sky view. At night, a warming in regions of restricted sky view is observed. It is of smaller magnitude than under clear sky conditions due to the presence of relatively warm clouds, as mentioned before. Note that advection smooths the spatial patterns.

7.4 Conclusions

The proposed parameterization allows simulation of important radiation effects associated with complex topography at negligible computational costs. Verifications with 2 m air temperatures demonstrate a general improvement of 0.5 to 1 K in RMS and mean error, respectively. Especially at night, the consideration of restricted sky view leads to higher air temperatures in complex terrain. Along valleys the nighttime warming is between 0.5 and 1.5 K. During clear sky daytime, this warming is of the same magnitude for grid cells containing slopes exposed to the sun. The parameterization also shows improvement under overcast conditions, where the nighttime warming was about half that of clear sky conditions. During an overcast day, the parameterization cools the air in areas of restricted sky view by 0.2 to 0.7 K. In wintertime with lower sun elevation, shadows reduce temperatures in valleys by 0.5 to 3 K during daytime. At a resolution of 4 km the parameterization has significant impacts, which are even more pronounced on the finer 2 km grid.

Higher temperatures on slopes facing the sun may play an important role in triggering convective processes, and thus influence precipitation patterns as well. But this will have to be analyzed in the future.

7.5 Acknowledgements

The authors want to express their gratitude to NOAA/NCEP for providing the NMM source code. The scientific discussion with Zavisa Janjic of NCEP and all his help concerning the NMM were crucial for this research. This work substantially benefited from the comments and suggestions made by the anonymous reviewers. The support of Matthew Pyle of NCEP was very helpful. Thanks to Martin Jacquot of the University of Basel for reconfiguring the Beowulf cluster to run NMM. The authors also want to thank SLF and MeteoSwiss for providing measurement data.

8 Outlook - a 3D fog model

This chapter is thought to be an outlook to the future of fog forecasting. So far many problems of 1D models have been pointed out and several tricks are used to extend the applicability of 1D models to more complex situations. However, to model the processes in complex topography, it seems inevitable to sooner or later extend a 3D model in such a way that it can do fog prediction. This is of course a long term project because many issues like cloud microphysics, turbulence or problems related to high vertical and horizontal resolution, have to be solved. More complex microphysics also requires more complex initialization with properties currently not measured or assimilated, like the concentration and chemical composition of aerosols. It will not be sufficient to develop only a forecast model but also an accurate data assimilation for high resolution.

Here I want to present a first step towards 3D fog prediction, which becomes possible due to the outstanding dynamical core of the NMM developed by Zavisa Janjic at NOAA/NCEP.

8.1 Requirements for 3D fog forecasting

A 3D fog model needs detailed cloud microphysics and should be able to run at very high resolutions. Horizontal resolution is important to simulate cold air drainage flows and cold air ponding in complex topography. The growth of the surface inversion, eventually followed by fog formation requires also a high vertical resolution in the boundary layer. The dynamical core of the NMM seems to be very suitable, since it can run at horizontal resolutions as fine as 100 m (Janjic 2003) and because it is computationally very efficient. The latter is very important if forecasts have to be done in an operational context. The condensation/evaporation scheme as well as computation of droplet settling are essential to model fog. However these processes are represented in a highly parameterized manner in the NMM, as well as in all other major weather forecast models in operational use listed in Table 8. What is used are very simple bulk adjustment schemes for condensation and evaporation. With this kind of adjustment, condensation and evaporation are separated by a prescribed relative humidity threshold value. For high resolution models this is usually 100 %. For coarse resolutions the threshold is decreased, because the large grid box volume is often only partly cloudy in reality, and these clouds have to be predicted by the model, particularly because of the interaction with radiation. The condensation/evaporation and the related change in temperature and saturation vapor pressure is solved using different strategies. For example in the MM5 model, if supersaturation exists an adjustment step is applied that converts water vapor to liquid water and warms the air. This is done with the equation of Soong & Ogura (1973) that uses a single step to approximately reach 100 % relative humidity. A Taylor series expansion is used to derive this equation, where only the first-order terms are retained. By doing so the amount of condensation necessary to obtain 100 % relative humidity is only approximated. Other models, like the NMM use an iterative procedure (Asai 1965) similar to COBEL which is more accurate but also computationally more expensive. A statistic approach with a PDF and height dependent humidity thresholds is used in the UK MetOffice model (Wilson & Ballard 1999, Smith 1990).

But there are also higher degrees of sophistication. From Köhler theory (Köhler 1936) it is known that the critical supersaturation for a droplet depends on its size and chemical composition. Bigger size and higher salt concentration decrease the critical supersaturation. Therefore droplets can exist at relative humidities below 100 % and larger droplets still grow at humidities where smaller ones evaporate. If these processes have to be modeled, it is necessary to solve the droplet growth equations for several droplet size bins, as it was done in some sophisticated 1D models (Brown 1980, Bott et al. 1990). Even though such schemes have been implemented and tested in some of

Table 8: Cloud condensation schemes of operationally used numerical weather prediction models.

Model	Condensation scheme
ARPS	Soong & Ogura (1973)
Hirlam	Sundqvist (1978)
LM / ALMO	iterative bulk adjustment (Doms et al. 2005)
MC2	Sundqvist et al. (1989)
Meso ETA	Asai (1965)
MM5	all schemes follow Soong & Ogura (1973)
NMM	Asai (1965)
RAMS	direct, non-iterative implicit algorithm
RUC	as MM5 Brown et al. (1998)
UK UM	Wilson & Ballard (1999), Smith (1990)
WRF	different schemes but all threshold based

the models mentioned in Table 8, they are computationally too expensive to be used for actual weather forecasting. For our 3D model a parameterization of the Köhler theory is used, so that only a total droplet number concentration, but not different size bins are needed. The microphysics still have a high degree of sophistication and are computationally much cheaper.

8.2 Microphysics implementation strategy

The PAFOG microphysics, outlined in section 8.4, was incorporated into the Nonhydrostatic Mesoscale Model to improve forecasts of fog and low stratus clouds in 3D. In the NMM the condensation/evaporation algorithm of Asai (1965) is used and now replaced by detailed microphysics. It is important that the implementation does not interfere with the already existing cloud and precipitation microphysics scheme (Ferrier 2002) of the NMM. PAFOG microphysics is limited to the lower part of the atmosphere which extends from the ground up to a prescribed height, currently 1500 m. In this lower part, where fog and low stratus clouds form, the condensation and evaporation processes, as well as the settling of cloud droplets are modeled with the detailed cloud microphysics of PAFOG. Precipitation, processes including the ice phase, autoconversion, accretion and evaporation of precipitation are left unchanged and predicted with NMM microphysics. Liquid water content already is a prognostic variable in the NMM, and once formed, is transported by turbulence and advection. Now a new prognostic variable, the total droplet number concentration (N_c) is introduced into the dynamical framework. The three dimensional link of the PAFOG microphysics to the other model equations is done by the transport of N_c and liquid water. The implementation has to consider horizontal and vertical advection as well as the turbulent transport of N_c. Note that N_c fields are not smooth at all and values can be as high as 10^9, producing a field with large gradients that looks very noisy. Tests revealed that it is necessary to carry out calculations of droplet number concentration in double precision. However, to increase the speed of communication with MPI between different nodes on the parallel machine, N_c fields are converted to floating point precision before exchanging them between processors.

8.3 Spatial discretization

In terms of resolution a couple of changes are necessary. Modeling the boundary layer processes, especially the nocturnal cooling due to emission at the surface, requires a high vertical resolution. The formation of radiation fog normally begins close

to the surface and then steadily grows upward. In today's operational high resolution weather forecast models, the first layer above ground is about 30 m thick in the NMM and around 70 m in the aLMo and LM. Such coarse resolutions cannot describe the slowly growing fog. For the 3D fog model this is changed to a very high vertical resolution with about 27 layers in the lowest 1000 m and a thickness of 5 m for the first few layers. The top boundary is set to be at 50 hPa, and 18 further layers are used to reach this height. For the boundary layer, the vertical resolution is similar to what is used in 1D models. There is however a drawback related to the smaller heat capacity of thin layers. The cooling or heating at the surface affects the lowest layer first and is then communicated to upper layers via turbulent exchange processes. During the day a warming of lower layers results in a more unstable lapse rate and thus the upward transport of heat works well. Hence temperatures are relatively independent of vertical resolution as long as the surface is heated. At night, the radiative cooling of the surface increases the stability and thus deteriorates the conditions for vertical exchange that could transport heat downwards. So there is a positive feedback which produces a very strong inversion in the lowest layers, where the temperatures can be several Kelvins colder than observed. The problem could probably be solved with a sophisticated turbulence scheme that is fully three dimensional and able to produce intermittent turbulence. With current schemes, the specification of a minimal exchange coefficient could reduce the cold bias at night. However, the residual artificial diffusion in the surface layer and the layer below the inversion in the NMM would then be un-physically high. Furthermore, the longwave flux reaching the surface over night in clear sky conditions is underestimated in the NMM (Janjic, personal communication 2005). The underestimation of incoming longwave radiation is also related to topography. Due to a reduced sky view factor in complex terrain, a particular grid point receives more longwave radiation from the surrounding mountains than from the clear sky. Because the sky is normally colder than the surrounding, a reduced sky view factor increases the amount of incoming radiation. But topography also influences shortwave direct and diffuse radiation that can be parameterized. A detailed description of this radiation parameterization was given in Section 7. Consequently, it appears that the crux of the problem is that there is not enough energy reaching the surface in order to close the energy balance over night under clear sky conditions.

Concerning the resolution of the surface or constant flux layer, this layer is approximately 10 % of the PBL height, so its thickness varies from a few meters to over a hundred meters. Near the surface, the profiles in this layer are very nonlinear (approach logarithmic), and that is the reason why the NMM uses the Monin-Obukhov similarity to represent it. Further away from the surface, the profiles are more linear so that an explicit handling of turbulence is possible using finite-differences that imply piece-wise linear profiles. Considering that the surface layer is only a few meters thick under strongly stable conditions, it is clear that a thickness of 50-100 m for the first model layer does not appropriately resolve the physical processes. However, under such conditions the fluxes are small, so that the resulting error is small in absolute terms. On the other hand, very high, but still insufficient resolution within the strongly nonlinear part of the profiles also could be harmful.

Some improvements for the above may however be possible. For example, emissivity is often assumed to be 1, which is an overestimate, consequently this value could be reduced for the fog model if cold bias problems are found. For complex terrain the mentioned radiation parameterization also improves the nightly cold bias problem. More desperate measures could also artificially increase the ground flux or add a "residual" cirrus. The first actually happened in the 1D COBEL-NOAH model when the resolution of the soil layers was increased to the current setup. In calm nights the computed soil heat flux reached values between 50 to 100 W m^{-2} which is an overestimate but leads to quite accurate temperature forecasts for the stable boundary layer. In the 3D fog model the high resolution of the soil grid did not really improve

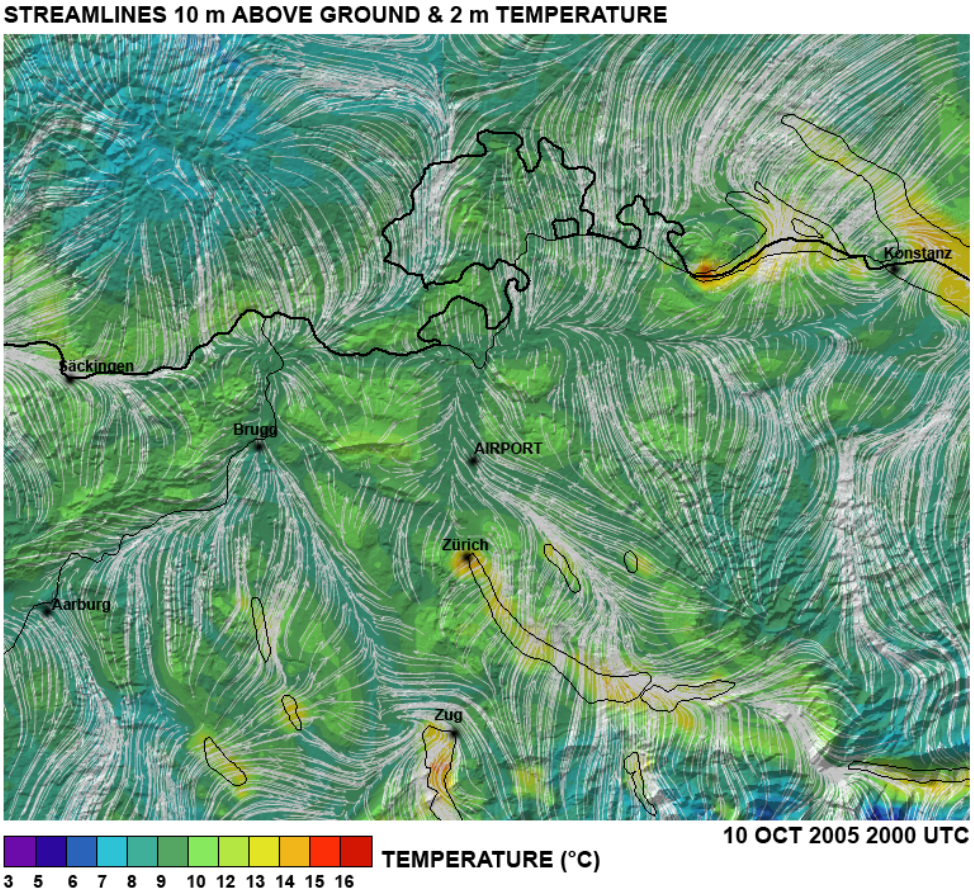

Figure 39: 3D forecast of 2 m temperature and 10 m wind for the fog case of 10 October 2005, 2000 UTC.

the cold bias problem. This is because in complex terrain cold air drainage flows keep bringing cold air into basins where fog forms. Nevertheless the thin soil layers are needed for the temperature and hydrological coupling. The effects on temperature were described in Section 2.2.2. In hydrological terms, the deposition of fog is small and can only change the water content of a thin soil layer. During the day this water is again evaporated and the first thin layers can dry, which produces vertical moisture gradients in the soil. The related moisture transport in the soil can be resolved and computed on the high resolution soil grid. With the normal resolution, which is 10 cm for the uppermost layer, the response times are too long and vertical gradients do not really develop, so that the moisture availability for evaporation does not change much during the day. This means that evaporation is purely controlled by the atmospheric conditions and no constraints due to a drying soil layer can be considered on the short term.

The horizontal resolution should also be rather high to resolve a considerable amount of topographical details. Currently 1 km is used and the effects of cold air outflow and ponding are well resolved as can be seen in Figure 39. The high horizontal resolution did not cause any problems since the dynamical core was originally designed to work with a grid spacing as fine as 100 m.

8.4 The microphysics model

The microphysical parameterization was taken from the 1D model PAFOG (Bott & Trautmann 2002) and is used to compute the cloud liquid water content in the lower part of the model atmosphere. Prognostic equations for the total number concentration N_c of cloud droplets and for the total specific cloud water content q_c are solved.

$$\frac{\partial N_c}{\partial t} = ADV(N_c) + DIF(N_c) + \left(\frac{\partial N_c}{\partial z}\right)_{sed} + s(N_c) \tag{83}$$

$$\frac{\partial q_c}{\partial t} = ADV(q_c) + DIF(q_c) + \left(\frac{\partial q_c}{\partial z}\right)_{sed} + s(q_c) \tag{84}$$

The operators ADV and DIF are computed by the dynamical framework of the NMM and stand for advection and turbulent diffusion, respectively. These two transport processes are essential for the 3D coupling of the microphysics. Their implementation for N_c into the 3D model require major modifications of source code, which, depending on the 3D model, are quite complex. Note that a new prognostic variable has to be added. The third term represents sedimentation of cloud droplets and the source-sink terms $s(N_c)$ and $s(q_c)$ describe phase changes between the gaseous and liquid phase. From the continuity equations [83,84], the following two prognostic equations are solved with the PAFOG microphysics.

$$\frac{\partial N_c}{\partial t} = \left(\frac{\partial N_c}{\partial t}\right)_{act} + \Delta(\bar{S})\left(\frac{\partial N_c}{\partial t}\right)_{eva} + \left(\frac{\partial N_c}{\partial t}\right)_{sed} \tag{85}$$

$$\frac{\partial q_c}{\partial t} = \left(\frac{\partial q_c}{\partial t}\right)_{con/eva} + \left(\frac{\partial q_c}{\partial t}\right)_{sed} \tag{86}$$

$$\Delta(\bar{S}) := \begin{cases} 1, & \text{for } \bar{S} < 0 \\ 0, & \text{for } \bar{S} \geq 0 \end{cases} \tag{87}$$

$$S = \frac{q}{q_{sat}} - 1$$

where $\bar{S}$ is the mean supersaturation. Note that due to the $\Delta(\bar{S})$ mechanism, evaporation changes N_c only if the air is not saturated. Besides the dynamic transports due to advection and turbulence, the concentration of N_c and liquid water content can be changed due to microphysical processes. The interactions are quite complex and an increase in liquid water does not necessarily change the droplet number concentration. This is because existing droplets may grow without new droplets being formed. The inverse is true for evaporation of cloud water. Sedimentation is always directed downwards and the same flux divergence may increase or decrease N_c and cloud water differently, depending on the size of settling droplets. Because larger droplets have higher settling velocities they are removed effectively by sedimentation and evaporation has little effect on the number concentration for big droplets. For small droplets the opposite is true and the evaporation algorithms gain importance. At the ground, liquid water is treated like precipitation and droplets disappear due to deposition. Before further considering the individual terms, an assumption on the droplet size distribution has to be made. In PAFOG this is a log-normal function of the form

$$dN_c = \frac{N_c}{\sqrt{2\pi}\sigma_c D} \exp\left(-\frac{1}{2\sigma_c^2}\ln^2\left(\frac{D}{D_{c,0}}\right)\right) dD \tag{88}$$

where D is the droplet diameter, $D_{c,0}$ is the mean value of D and σ_c is the dispersion parameter of the given droplet size distribution ($\sigma_c = 0.2$). The resulting size distributions for different values of σ_c are shown in Figure 41. The shape reflects the explicitly computed spectral distribution obtained from simulations with the MIFOG model (Bott et al. 1990). Note that the simple size distribution of fog is caused by the

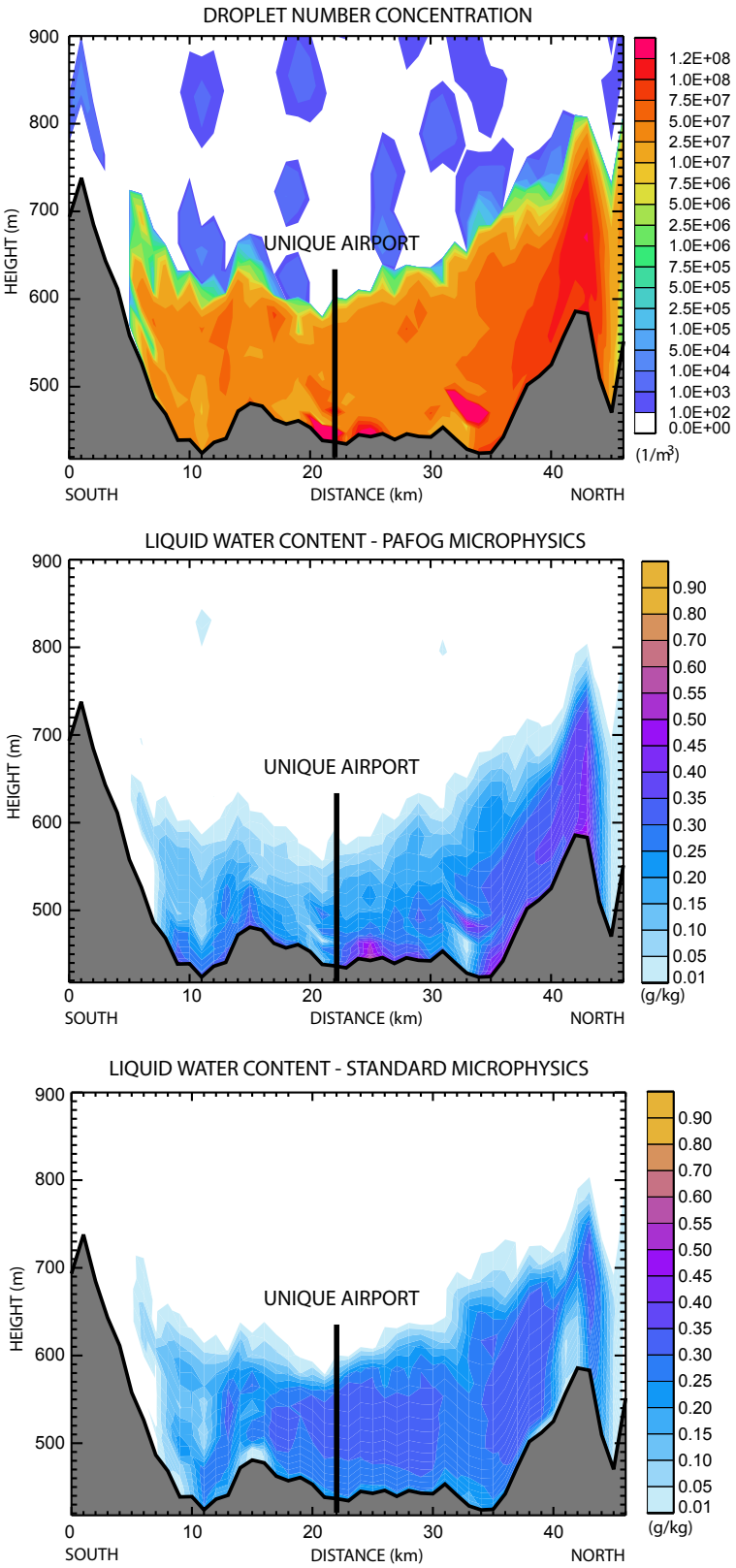

Figure 40: Cross section of droplet number concentration and liquid water content as computed with PAFOG and standard microphysics for the fog event of 28 November 2004 at 0100 UTC (10 h forecast).

nature of its formation under stable conditions due to slow cooling, which produces a relatively shallow layer of small droplets. There are only very weak vertical motions and the differences in fall speed are too small for coagulation to be important. This is very different from a convective cloud, were up- and downdrafts, coagulation and the presence of liquid and frozen water produce very complicated droplet spectra.

A second important assumption has to be made to determine the number of droplets that activate when supersaturation is reached. Note that a newly activated droplet competes with other newborns, as well as older siblings, about the available condensate so that too large a population may result in sudden death enforced by evaporation. Birth control is done using Twomey's relation (Twomey 1959). For a supersaturation S the number of activated cloud droplets N_{act} is calculated as:

$$N_{act} = CS^k \tag{89}$$

The constants k and C describe the aerosol model (maritime: $k = 0.7$, $C = 100$ cm^{-3}, continental: $k = 0.9$, $C = 3500$ cm^{-3}, this study: $k = 0.9$, $C = 10000$ cm^{-3}). Aerosols are needed condensation nuclei to produce liquid water in the form of small droplets. In clean air without aerosols, supersaturation could easily exceed 100 %. For Switzerland a typical value of $C = 10000$ cm^{-3} as in Bott & Trautmann (2002) is used, but it could be even higher in rural areas. Currently C is the same for the whole domain, but if detailed data become available it can be easily specified as a spatial field. The increase in total concentration of cloud droplets during the time step Δt represents the first term of the prognostic equation [85] and with Twomey's relation is written as:

$$\left(\frac{\partial N_c}{\partial t}\right)_{act} : N_c(t + \Delta t) = N_c(t) + max(N_{act} - N_c(t), 0) \tag{90}$$

Due to the maximum operator, N_c only increases in the case of a positive tendency in the supersaturation. Thus if the supersaturation remains unchanged or decreases no new droplets are activated and existing droplets grow. The time rate of change for the cloud droplet diameter D due to condensation or evaporation is expressed with the analytical solution of Mason (1971):

$$\frac{dD}{dt} = A\bar{f}\frac{S}{D} \tag{91}$$

where f is a ventilation coefficient and A the thermodynamic function given in [92]. Now this needs some appreciation. Even though at the beginning a size distribution was specified, it is actually allowed to evolve. The function A is defined as:

$$A = \frac{L\rho_w}{KT}\left(\frac{L}{R_wT} - 1\right) + \frac{\rho_w R_w T}{e_s(T)D_v} \tag{92}$$

here e_s is the saturation vapor pressure over a plane water surface, ρ_w the density of water, R_w the specific gas constant for moist air, D_v the water vapor diffusivity and K is the thermal conductivity. The first term considers the effects of heat conduction, and diffusion of water vapor is expressed by the second term.

In the case of evaporation (the second term of [85]), the smallest droplets disappear first and N_c decreases. It is assumed that evaporation of small droplets is fast enough for the ventilation coefficient to be negligible. The process can be modeled by first computing the critical diameter $D_{c,eva}$ of the largest droplet able to evaporate

$$D_{c,eva} = \sqrt{-2AS\Delta t} \quad (S < 0) \tag{93}$$

and then integrating from the smallest droplet to the critical diameter

$$N_c|_{eva} = \int_0^{D_{c,eva}} N_c(D)dD \tag{94}$$

therefore the change in droplet number concentration due to evaporation can be computed

$$\left(\frac{\partial N_c}{\partial t}\right)_{eva} : N_c(t+\Delta t)|_{eva} = N_c(t) - N_c|_{eva} \tag{95}$$

So far the supersaturation S was assumed to be known. For its computation the procedure of Bott & Trautmann (2002), which is based on Chaumerliac et al. (1987) and Sakakibara (1979) was used:

$$\bar{S} = -\frac{c_3}{c} - \left(S(t) + \frac{c_3}{c}\right)\left(\frac{1 - e^{c\Delta t}}{c\Delta t}\right) \tag{96}$$

where c and c_3 are assumed to be constant for the time step Δt and $c := c_1 + c_2 + c_3$

$$c_1 = -\frac{1}{q_{v,sat}}\frac{\rho_w}{\rho}\frac{\pi}{2}A\Sigma_c \tag{97}$$

$$c_2 = -\frac{L^2}{R_1 T^2 c_p}\frac{\rho_w}{\rho}\frac{\pi}{2}A\Sigma_c \tag{98}$$

$$c_3 = \left(\frac{1}{p} - \frac{L}{R_1 T^2 \rho c_p}\right)\frac{dp}{dt} \tag{99}$$

$$\Sigma_c = N_c D_{c,0} exp\left[\frac{\sigma_c^2}{2}\right] \tag{100}$$

Knowing the supersaturation, the changes in liquid water content and specific humidity are:

$$\left(\frac{\partial q_c}{\partial t}\right)_{con/eva} = \frac{\rho_w}{\rho}\frac{\pi}{2}A\bar{S}\Sigma_c \tag{101}$$

$$\left(\frac{\partial q}{\partial t}\right)_{con/eva} = \left(\frac{\partial q_c}{\partial t}\right)_{con/eva} \tag{102}$$

and with that the change in potential temperature due to phase changes can be computed by:

$$\left(\frac{\partial \theta}{\partial t}\right)_{con/eva} = -\frac{L}{c_p}\left(\frac{\partial q}{\partial t}\right)_{con/eva} \tag{103}$$

The sedimentation tendencies for cloud water and droplet number concentration appearing in the main prognostic equations [85] are:

$$\left(\frac{\partial N_c}{\partial t}\right)_{sed} = \frac{\partial S_{n,c}}{\partial z} - \frac{\partial}{\partial z}\left(N_c w\right) \tag{104}$$

$$\left(\frac{\partial q_c}{\partial t}\right)_{sed} = \frac{\partial S_{q,c}}{\partial z} - \frac{\partial}{\partial z}\left(q_c w\right) \tag{105}$$

They can be evaluated assuming the log-normal droplet size distribution [88]. The size distribution is multiplied with the size dependent settling velocities and integrated over the whole droplet spectra. The formulation of Berry & Pranger (1974) is used to express the settling velocity as a function of the Reynolds number R_e

$$v(D) = \frac{\eta R_e}{D\rho} \tag{106}$$

where

$$\eta = 1.496286 \times 10^{-6}\frac{T^{1.5}}{T + 120} \tag{107}$$

The sedimentation terms $S_{n,c}$ and $S_{q,c}$ can then be evaluated

$$S_{n,c} = \int_0^\infty N_c(D)v(D)dD = \frac{N_c v(D_{c,0})}{\sqrt{2}\sigma_c \xi} \exp\left[\frac{(k_2-1)^2)}{4\xi^2}\right] \tag{108}$$

$$S_{q,c} = \frac{1}{\rho}\int_0^\infty m(D)N_c(D)v(D)dD = \frac{N_c m(D_{c,0})v(D_{c,0})}{\sqrt{2}\sigma_c \xi \rho} \exp\left[\frac{(k_2+2)^2)}{4\xi^2}\right] \tag{109}$$

using $a_2 = 1.01338$, $a_3 = -0.0191182$ and the following parameters

$$\xi := \sqrt{\frac{1}{2\sigma_c^2} - 9a_3} \tag{110}$$

$$k_2 := 3a_2 + 6a_3 \ln(aD_{c,0}^3) \tag{111}$$

$$a := \frac{4\rho\rho_w g}{3\eta^2} \tag{112}$$

$$m(D_{c,0}) := \frac{\pi}{6}\rho_w D_{c,0}^3 \tag{113}$$

For the numerical solution of sedimentation, the positive definite advection scheme of Bott (1989) is used. Since the size distribution of the droplets is known, the sedimentation of the droplets can be computed accurately and it is not necessary to make assumptions of a representative settling velocity as it is done in simple bulk microphysics models like COBEL. The sedimentation of liquid water is a very important sink, especially for the large droplets containing a lot of water and having high settling velocities.

The effects of more detailed microphysics can best be studied in a vertical cross section (Figure 40). The top panel shows the distribution of computed droplet number concentration, a variable not part of the standard microphysics. Generally highest concentrations correspond well with high liquid water contents, shown in the second panel. But interestingly there are also many droplets just over the airport, where liquid water content is not very high. This means that the droplets must be rather small. Note that for visibility the droplet number and the liquid water concentration are important. For a constant liquid water content, visibility decreases with increasing droplet number concentration. The structure of liquid water is very different from the one obtained with standard microphysics. With PAFOG microphysics the liquid water content generally decreases with increasing height. In the standard case the liquid water distribution is rather uniform with lower concentrations close to the ground and on top. In the case of PAFOG microphysics, most of the vertical structure is caused by the sophisticated sedimentation scheme, which transports droplets of all sizes at corresponding fall speeds. With standard microphysics sedimentation kicks in after a critical liquid water concentration is exceeded and then decreases the water content uniformly due to the lack of information about droplet size. Evaporation of droplets is not very important at night.

8.5 Boundary conditions for N_c

PAFOG microphysics introduces droplet number concentration N_c as a new prognostic variable into the 3D forecast model. The main goal is to link the N_c to the liquid water content and thus allow a more precise modeling of the cloud water content. The processes involved are activation, condensation/evaporation and sedimentation. Since the droplet distribution outside the PAFOG domain is unknown, some problems arise when liquid water is transported into the domain. In particular PAFOG is unable to evaporate or sediment this water because there are no droplets related to it. As a consequence, liquid water accumulates in an environment that might have

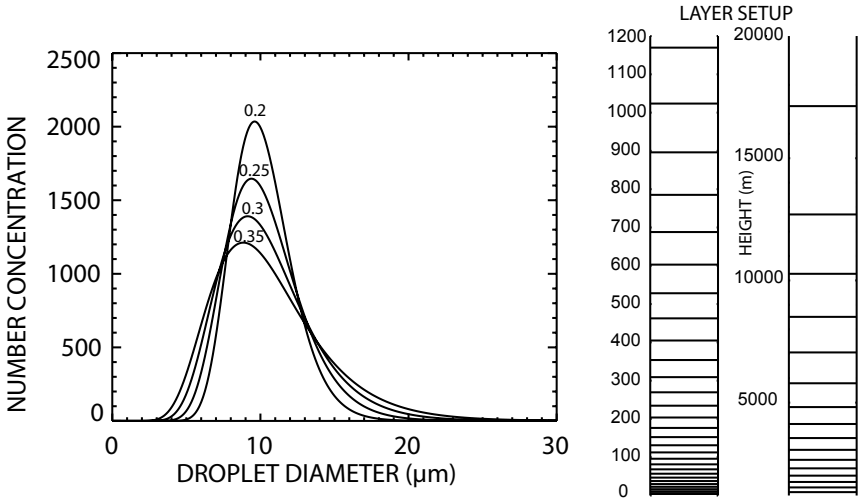

Figure 41: Some parameterized droplet size distributions with different σ used as boundary conditions (left panel) and vertical layer distribution of the 3D fog model (right panel).

relative humidities far below saturation and finally even precipitation may form. Now the sources of such invisible water are advection through the lateral boundaries as well as advection or turbulent diffusion through the top boundary, since PAFOG is restricted to the lower part of the atmosphere. To maintain the link between N_c and liquid water content, and thus control the development of fog, a boundary condition for N_c has to be formulated.

A first possible solution is forced evaporation of the water at the boundaries and condensation with PAFOG microphysics. This procedures produced too high relative humidity values for the PAFOG microphysics. Note, that normally condensation occurs over some time always keeping the relative humidity at low supersaturation. It is however impossible to reconstruct the history of past supersaturation conditions of the water at the boundaries so that the use of such an approach is not expected to produce accurate results. Furthermore the evaporation/condensation procedure of PAFOG does not allow its use with a threshold of humidity due to the relaxation in the Sakakibara parameterization.
A second, more simple, very robust and also computationally efficient approach is to relate liquid water content at the boundaries directly to N_c, assuming a lognormal droplet size distribution as defined in [88].

We can solve Equation 3 of Chaumerliac et al. (1987) for N_c

$$q_c = \frac{N_c}{\rho_a}(\frac{\pi}{6}D_{c,0}^3\rho_w)exp(\frac{9}{2}\sigma_c^2) \tag{114}$$

$$N_c = \rho_a q_c(\frac{1}{\frac{\pi}{6}D_{c,0}^3\rho_w})exp(\frac{-9}{2}\sigma_c^2) \tag{115}$$

which is used at the boundaries.

The difficulty now is to find representative values for the mean diameter and standard deviation defining the lognormal droplet size distribution. They were determined from the work of Miles et al. (2000), who presented a summary of different N_c measurement campaigns for maritime and continental clouds. It is assumed that clouds of the PAFOG domain are of continental type only. The lognormal distribution is

parameterized with an evolution of the standard deviation and a fixed mean diameter of 10 micrometers. The standard deviation σ_c is linearly interpolated with height between a value of 0.2 at 1000 m and 0.35 at the top of the PAFOG domain. Below and above this region the values of σ_c are kept constant at 0.2 and 0.35, respectively. Figure 41 summarizes the parameterization and shows some resulting droplet size distributions. Note that this procedure basically broadens the droplet size distribution with increasing height.

The boundary condition is applied at the lateral boundaries as well as at the next five layers above the PAFOG domain, so that water entering from above due to advection or turbulent diffusion contains N_c. Furthermore the boundary condition is applied at initialization when liquid water is present. It has to be noted that droplets introduced by this procedure are generated with a prescribed size distribution, but once existent they can grow or shrink, since the mean diameter of the droplet size distribution is computed with [91].

8.6 The cold bias problem

The already mentioned cold bias problem can give rise to severe over- as well as underestimation of fog. The latter might seem surprising but we have to consider that the cold bias at the surface is caused by insufficient turbulent mixing. So the lowest few meters are too cold but the upper air is consequently too warm, in fact it could just be warm enough to prevent fog formation. Depending on the humidity conditions near the ground, the reduced mixing might also keep upper levels too dry.

Zavisa Janjic (personal communication, 2005) introduced a few modifications to the moist turbulence scheme which addresse the cold bias problem. The floor value for two times the turbulent kinetic energy (q^2) was reduced from $q^2 = 0.2$ to $q^2 = 0.02$ and the minimum mixing length $l = 0.32$ m was changed to $l = 0.1$ m. With these physically more realistic values the residual exchange coefficients drop from about 0.1 to 0.01 in the case of no turbulence, i.e. when q^2 and l hit the floor values. Thus in order to send more heat toward the surface in the case of very strong ground inversions, residual exchange coefficients near the surface are set to 0.5 m^2/s. The most important change was made in the calculation of the production of turbulent kinetic energy, which now also considers the gradient of equivalent potential temperature. By directly converting some of the potential instability into turbulent kinetic energy, too unstable shallow convective profiles are corrected towards a moist adiabatic lapse rate. In the case of fog it appears to transport more heat downwards which significantly reduces the cold bias.

Several case studies revealed that temperatures are now very similar to simulations done on a coarser vertical grid. Therefore the changes solve the problem associated with the high vertical resolution. Figure 42 shows an example cross section of predicted temperature and liquid water content. A cold bias of 6 K in 2 m temperatures at Zürich Unique airport was reduced to 0.5 K using the above mentioned modifications. It can be seen how the fog resides in a relatively homogenous layer of cold air, a pattern discussed for 1D models in Section 3.1. Without the modifications, a strong surface inversion leading to the cold bias developed, which kept the fog within the lowest 20 m.

8.7 A new possibility to compute visibility

The PAFOG microphysics compute the total droplet number concentration. This variable is important for the calculation of visibility, because an increasing number of light scattering droplets decreases visibility resulting from the same liquid water content. In our article (Gultepe et al. 2005) we develop a new parameterization for visibility that considers the new prognostic variable. Detailed observations of liquid

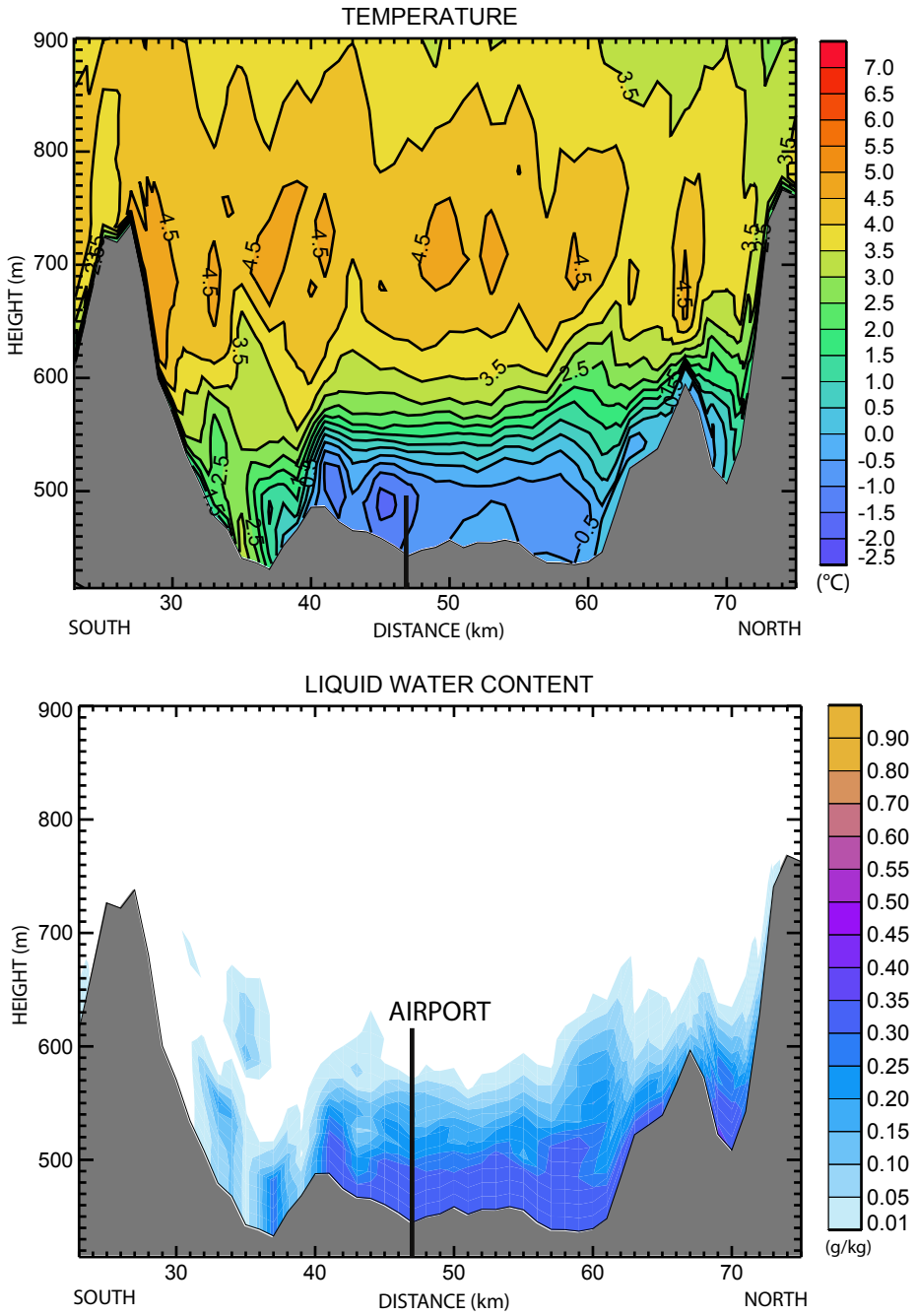

Figure 42: 12 h forecast of temperature and liquid water content for 28 November 2004 at 0300 UTC.

water content (q_c), visibility (vis) and droplet number concentrations N_c lead to the following parameterization for visibility:

$$vis = 1000 \left(1.0002 (\frac{1}{q_c N_c})^{0.6473} \right) \tag{116}$$

So with the more detailed microphysics it is also possible to better diagnose visibility. But it has to be mentioned that very accurate predictions of q_c and N_c are needed for reliable visibility forecasts, which is currently not possible. In fact the 1D model verification showed that it is very difficult to only predict occurrence or non-occurrence of fog.

8.8 Data and method for verification - a case study

Assessing the quality of the 3D fog model is rather difficult with currently available data. Ideally a spatial distribution of fog would be needed, but such data have to be derived from satellite measurements. Unfortunately satellites look at clouds from above so that the discrimination between low stratus clouds and real fog is currently not very accurate. Monitoring onset and dissipation further requires a high temporal resolution, so that Meteosat-8, also called Meteosat Second Generation (MSG) with a frequency of 15 min and a spatial resolution of around 3 km in Switzerland seems to be appropriate. NOAA/AVHRR provides a high resolution 1 km image without much geometric distortion, but there are only a handful of satellite passes a day. Of course one could also just focus on a single point, where ground observations are available, and compare one vertical column from the 3D results to assess the forecast skill for a whole season of simulations.

Since October 2005 the 3D fog model is run in a semi-operational framework to acquire a data set for future verification. For an area around Berlin-Lindenberg as well as Zürich Unique airport 24 hours of simulation are computed daily. The Berlin-Lindenberg domain is part of a model comparison project of COST-722. Both experiments use a model domain of 100 by 100 grid points at a resolution of 1 km and a time step of 2 seconds. This configuration requires about 55 min for a 24 h simulation if 9 Pentium-4 CPU's are used. This high efficiency of the model would make it possible to run fog forecasts even for operational purposes. With an extent of 100 km by 100 km, the model domain requires a proper nesting into a larger domain. For the Swiss experiment, the fog model was nested into the 4 km resolution NMM-4 model and for Berlin, the operational 13 km grid of NMM-22 was taken. It should be noted that a nesting with the Lokalmodell of the German Weather Service was also programmed.

In the scope of this outlook a case study for 11 October 2005 is presented to demonstrate the potential of the fog model and the applicability of MSG satellite data for model verification. For this particular case a prototype cloud classification developed by Cermak et al. (2004) was available. Cermak & Bendix (2004) mention some co-location issues in the satellite product. On the whole the geo-location of Meteosat-8 is stable, but slight changes may occur in individual scenes. One therefore has to consider that features may be found in a pixel other than the expected one in some instances, resulting in an error of around 5 km for Switzerland. With correctly geo-referenced visualizations as in Figure 43, conclusive verification can be done by visual inspection. Black shaded areas indicate fog occurrence as detected by the satellite and color shading shows the modeled liquid water content in the lowest layer. Overall the result looks very promising in terms of spatial distribution as well as temporal evolution. Especially if one keeps in mind the uncertainties and relatively coarse resolution of the satellite observations. It has to be noted that the model underestimates the fog in the upper right corner over lake Constance but this is very likely a boundary effect. There was an easterly flow entering the domain but

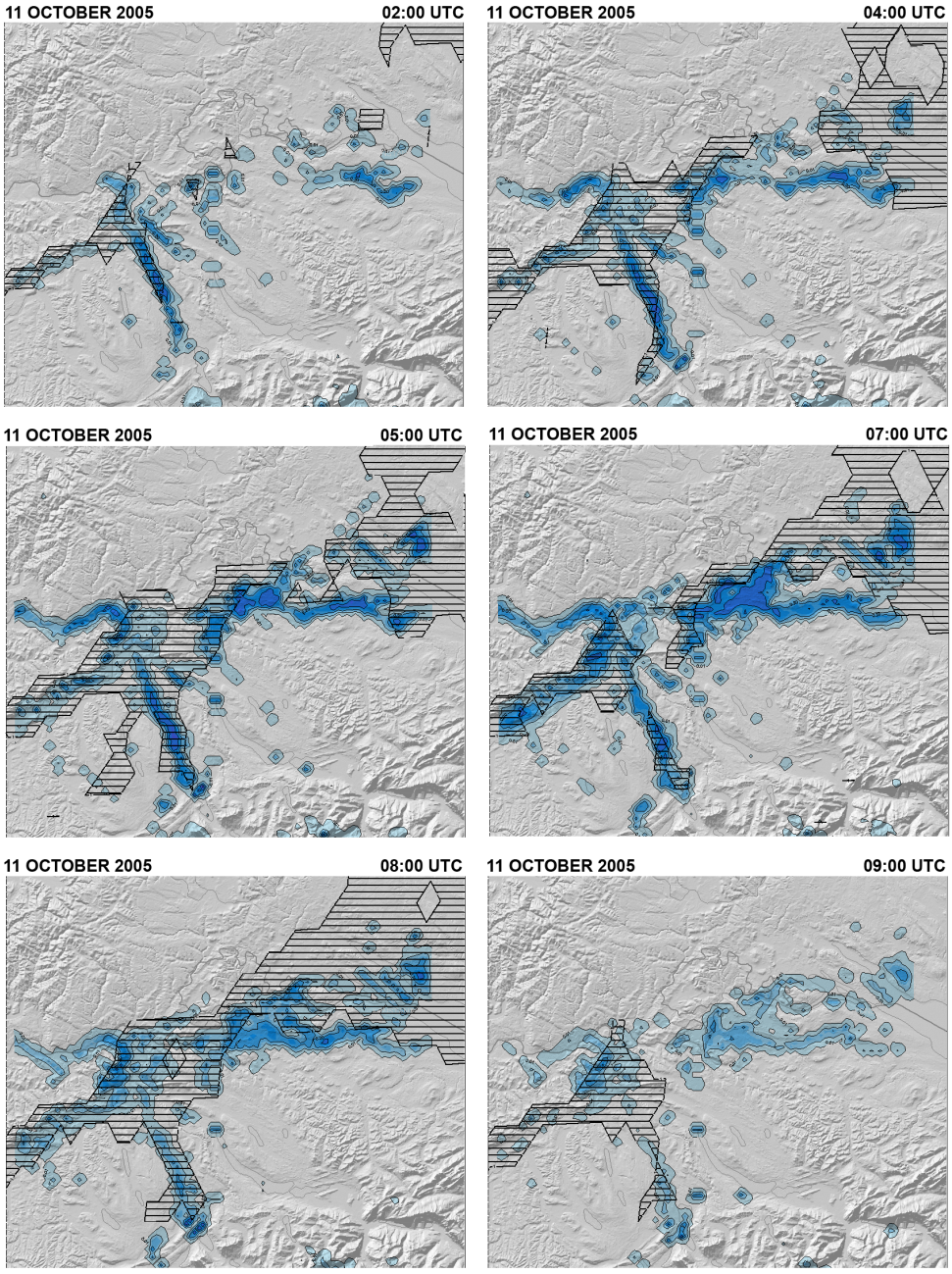

Figure 43: Spatial distribution of fog as seen by satellite (black shaded areas) and simulated by the 3D fog model (colored areas). Contour lines of modeled liquid water have a spacing of 0.1 g kg^{-1}.

the driving model did not predict any fog, so that the boundary conditions were too dry. For an objective skill assessment the model and satellite data can be interpolated to a common resolution and similarities and discrepancies can be analyzed for every pixel.

A case study is of course not very representative, but once a whole winter season has been simulated and the MSG fog product becomes available for the entire time range, more conclusive verification can be carried out. Of special interest will be the number of false alarms. So far the model seems to overestimate fog occurrence but a verification at the end of the fog season will give a definitive answer.

9 Conclusions and summary

1D ensemble fog forecast

The numerical simulation of fog is still a very difficult task. The reason is the need for an almost perfect prediction of boundary layer temperature, humidity and wind together with a detailed treatment of cloud microphysics. Small errors in temperature or humidity can have dramatic effect if the air is close to saturation. Even though there are situations with better predictability, such as days with high relative humidity at the beginning of the night, clear sky and only weak wind, the exact onset and dissipation times are still difficult to obtain from numerical simulations. Ensemble forecasting is very helpful to deal with the large quantity of uncertainties. The generation of ensemble members is a rather complicated process and involves data from different operational models and an assimilation scheme. The ensemble members should reflect the current atmospheric uncertainty, which can be achieved by computing a whole set of assimilations using background estimates from different 3D models. With few observations available, variational assimilation has to give a lot of weight to the background derived from 3D models. For the humidity assimilation, observational data are scarce and of poor quality, leaving a lot of uncertainty in the assimilated profiles. Experiments showed that artificially increasing the humidity improves the fog forecast skill. Therefore a humidity ensemble might be helpful for operational purpose. Skill scores from 180 days of simulations proof that the 1D ensemble forecast system can significantly improve fog forecasting at Zürich Unique airport. The numerical model ensemble is able to increase the discrimination up to a hit rate of 60 % with a false alarm rate of 30 % or for the 2100 UTC initialization to a hit rate of 80 % with a false alarm rate of 45 %. It has to be noted that these forecasts are purely machine based, and if the model results were interpreted by a human forecaster, even better skills could be achieved.

As good as this seems, the potential of a 1D model in complex terrain is however limited by the effects resulting from spatial heterogeneity. There are many problems if spatial inhomogeneities have to be included in a 1D model. Boundary conditions can be introduced to simulate the effects of advection and changing cloud cover can be considered by variable radiation fluxes at the model top. All in all the 1D model is heavily dependent on a 3D model supplying the boundary conditions. A 1D fog forecast is thus very sensitive to the forecast quality of the driving 3D model. Another problem in a 1D model are accurate forecasts of the mean flow, which are essential for computing turbulent exchange.

Nevertheless, the big advantage of 1D models is high computational efficiency, allowing to compute a probabilistic forecast for the next 18 h at a high temporal frequency, e.g. every hour. If driven with observations and not run in forecast mode, a 1D model can also be a very good research tool to test and develop physical parameterizations. From the underlying assumption of spatial homogeneity, the developed forecast system is expected to yield even better results for locations in relatively flat terrain.

3D fog modeling

A proper solution to deal with spatial heterogeneity is without question the development of a 3D model. A first step in this direction was made by including detailed fog microphysics into an existing 3D model. By doing so, the total droplet number concentration is introduced as a new prognostic variable, which allows to better represent the gravitational settling as well as condensation and evaporation of cloud water. During the growth phase, radiation fog steadily gains thickness, starting from a thin layer above the surface. To model this process it is necessary to have a very high vertical resolution not used in numerical weather prediction models. As a consequence,

the turbulent exchange is not adapted to such thin layers, and through nocturnal decoupling a severe cold bias may develop if no modifications are made. In order to resolve topographical features of complex terrain, a grid spacing of 1 km or even less is required. At such high resolution, the model reproduced cold air drainage flows and cold air ponding resulting in a fog of larger vertical extent than in the case of local surface cooling. Indeed, simulated fog from the 1D model was often very shallow and could not reach the vertical extent typically seen. When it comes to verification, spatial fields have to be analyzed, which expresses the need for satellite data or a very dense observational network. A case study outlines the good performance of the 3D fog model and also the potential of sophisticated satellite cloud products, which can be used for verification purposes. However, for conclusive skill assessment a whole season needs to be simulated and analyzed, which will be part of future research. Also of great importance is the specification of initial conditions, which is far more complicated than in the 1D case. Currently initial conditions are simply interpolated from an operational forecast model, but the development of a high resolution data assimilation system is expected to have significant impact.

The high computational efficiency of the NMM results in a fog model with relatively modest CPU requirements. Furthermore the 3D context allows to do time integrations of two days or even more if the model is nested into an operational forecast model. Even today it would be possible to run a 2 day forecast for Switzerland in an operational context. Especially in complex terrain the future of fog forecasting will be a 3D model. For the moment this can be a separate model running for a small region, but as the resolution of the operational weather forecast reaches 1 km, the fog module could be incorporated into an operational model.

Radiation in complex terrain

With increasing resolution, topographical effects on radiation gain importance. The smaller grid spacing results in steeper slopes and the radiation absorbing surface cannot any longer be modeled as a horizontal surface. In fact radiation interacts in a very complex way with topography. It is computationally too expensive to model processes like multiple reflections, but if only some basic features like slope angle, aspect angle, shadows and sky view restriction are considered, the resulting spatial redistribution of energy has a significant positive impact on the temperature forecast. In absolute values, a reduction of 0.5 K in RMS error and 1 K in mean error was found. The benefit is most pronounced at night, when the inclusion of sky view restriction leads to a smaller cold bias. But the radiation parameterization does not solely affect temperature. Indeed, a changed temperature can influence the wind field, clouds and even precipitation.

The developed parameterization is especially useful because it comes at no computational costs. Furthermore its effects increase with decreasing grid spacing, which is likely to happen with the steadily improving performance of computing resources. Even though the grid spacing of operational numerical weather prediction will reach 1 km in the next 5 to 10 years, there will still be a lot of topographical detail missing. But since topography is treated in a sub-grid way, higher resolution elevation data can be used to take into account small scale features not seen on the forecast model grid. Fortunately the necessary topographical information is already available. One such data set is the near globally available elevation model of the Shuttle Radar Topography Mission (SRTM), which has a resolution of about 80 m. Finally it has to be mentioned that, NOAA/NCEP as well as MeteoSwiss decided to implement the parameterization into their operational forecast models.

References

Anderson, J. L. (2001), 'An ensemble adjustment Kalman filter for data assimilation', *Monthly Weather Review* **129**, 2884–2903.

Asai, T. (1965), 'A numerical study of the air-mass transformation over the japan sea in winter', *J. Met. Soc. Japan* **43**, 1–15.

Ballard, S. P., Golding, B. W. & Smith, R. N. B. (1991), 'Mesoscale model experiment forecasts of the Haar of northeast Scotland', *Monthly Weather Review* **119**, 2107–2123.

Bannister, R. N. (2002), Assimilation of research satellite data into the Met Office 3d-Var. system: Part I: Forward models and adjoints for layer averaging, Technical Report 3, Data Assimilation Research Centre, Dept. of Meteorology, Univ. of Reading.

Bannister, R. N. (2003), On control variable transforms in the Met Office 3d and 4d Var., and a description of the proposed waveband summation transformation, Technical report, Data Assimilation Research Centre, Dept. of Meteorology, Univ. of Reading.

Bergot, T., Carrer, D., Noilhan, J. & Bougeault, P. (2005), 'Improved site-specific numerical prediction of fog and low clouds', *Weather Forecast.* **20**, 627–646.

Bergot, T., Cuxart, J., Liechti, O., Mira, T., Müller, M. D., Nielsen, N. W. & Terradellas, E. (2006), 'Intercomparison of single-column numerical models of fog prediction', *Journ. appl. Meteor.* p. in Press.

Bergot, T. & Gudalia, D. (1994*a*), 'Numerical forecasting of radiation fog. part I: Numerical model and sensitivity tests', *Monthly Weather Review* **122**, 1218–1230.

Bergot, T. & Gudalia, D. (1994*b*), 'Numerical forecasting of radiation fog. part II: A comparison of model simulations with several observed fog events', *Monthly Weather Review* **122**, 1231–1246.

Berry, E. X. & Pranger, M. P. (1974), 'Equations for calculating the terminal velocities of water drops', *J. Appl. Meteor.* **13**, 108–113.

Bott, A. (1989), 'A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes', *Monthly Weather Review* **117**, 1006–1015.

Bott, A., Sievers, U. & Zdunkowski, W. (1990), 'A radiation fog model with a detailed treatment of the interaction between radiative transfer and fog microphysics', *J. Atmos. Sci.* **47**, 2153 2166.

Bott, A. & Trautmann, T. (2002), 'PAFOG - a new efficient forecast model of radiation fog and low-level stratiform clouds', *Atmospheric Research* **64**, 191–203.

Bott, A., Trautmann, T. & Zdunkowski, W. (1989), 'A numerical model of the cloud topped planetary boundary layer: Radiation, turbulence and spectral microphysics in a marine stratus', *Quart. J. Roy. Meteor. Soc.* **122**, 635–667.

Bougeault, P. & André, J. C. (1986), 'On the stability of the three order turbulence closure for the modeling of stratocumulus-topped boundary layer', *J. Atmos. Sci.* **43**, 1574–1581.

Bouttier, F. & Courtier, P. (1999), Data assimilation concepts and methods, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Brown, J. M., Smirnova, T. & Benjamin, S. (1998), Introduction of MM5 level 4 microphysics into the RUC-2, *in* 'Preprints, 12th Conf. on Numerical Weather Prediction', Amer. Meteor. Soc., Phoenix, AZ, pp. 113–115.

Brown, R. (1980), 'A numerical study of radiation fog with an explicit formulation of the microphysics', *Quart. J. Roy. Meteor. Soc.* **106**, 781–802.

Brown, R. & Roach, W. T. (1976), 'The physics of radiation fog. part II: A numerical study', *Quart. J. Roy. Meteor. Soc.* **102**, 335–354.

Buizza, R. (1997), 'Potential forecast skill of ensemble prediction, and spread and skill distributions of the ECMWF Ensemble Prediction System', *Mon. Wea. Rev.* **125**, 99–119.

Buizza, R., Gelaro, R., Molteni, F. & Palmer, T. N. (1997), 'The impact of increased resolution on predictability studies with singular vectors', *Quart. J. Roy. Meteor. Soc.* **123**, 1007–1033.

Cermak, J. & Bendix, J. (2004), 'A novel approach to fog / low stratus detection using Meteosat 8 data', *Atmospheric Research* p. in Press.

Cermak, J., Thies, B. & Bendix, J. (2004), A new approach to fog detection using SEVIRI and MODIS data, *in* 'Proceedings of the 2004 EUMETSAT Meteorological Satellite Conference', EUMETSAT, Darmstadt, Germany, pp. 130–136.

Chaumerliac, N., Richard, E. & Pinty, J.-P. (1987), 'Sulfur scavenging in a mesoscale model with quasi-spectral microphysics: Two dimensional results for continental and maritime clouds', *J. Geophys. Res.* **92**, 3114–3126.

Chen, F., Janjic, Z. & Mitchell, K. (1997), 'Impact of atmospheric surfacelayer parameterization in the new landsurface scheme of the NCEP mesoscale Eta model', *Boundary-Layer Meteorology* **85**, 391–421.

Chow, F. K., Weigel, A. P., Street, R. L., Rotach, M. W. & Xue, M. (2004), High-resolution large-eddy simulations of the Riviera Valley: methodology and sensitivity studies, *in* '11th Conference on Mountain Meteorology and the Annual Mesoscale Alpine Program (MAP)', Preprints, American Meteorological Society, Bartlett, NH, pp. CD–ROM, 6.2.

Colette, A., Chow, F. K. & Street, R. L. (2003), 'A numerical study of inversion-layer breakup and the effects of topographic shading in idealized valleys', *J. Appl. Meteor.* **42**, 1255–1272.

Courtier, P., Thépaut, J.-N. & Hollingsworth, A. (1994), 'A strategy for operational implementation of 4D-VAR, using an incremental approach', *Quart. J. Roy. Meteor. Soc.* **120**, 1367–1387.

Daley, R. (1991), *Atmospheric data analysis*, Cambridge University Press, Cambridge.

Delage, Y. (1974), 'A numerical study of the nocturnal atmospheric boundary layer', *Quart. J. Roy. Met. Soc.* **100**, 351–364.

Doms, G., Förstner, J., Heise, E., Herzog, H.-J., Raschendorfer, M., Schrodin, R., Reinhardt, T. & Vogel, G. (2005), A description of the nonhydrostatic regional model LMpart II : Physical parameterization, Technical report, Deutscher Wetterdienst, Offenbach, Germany.

Donaldson, R. J., Dyer, R. M. & Kraus, M. J. (1975), An objective evaluator of techniques for predicting severe weather events, *in* 'Preprints, Ninth Conference on Severe Local Storms', American Meteorological Society, Norman, Oklahoma, pp. 321–326.

Dozier, J. (1980), 'A clear-sky spectral solar radiation model for snow-covered mountainous terrain', *Water Resour. Res.* **16**, 709–718.

Dozier, J. (1989), 'Spectral signature of alpine snow cover from the Landsat Thematic Mapper', *Remote Sens. Environ.* **28**, 9–22.

Dozier, J. & Marks, D. (1987), 'Snow mapping and classification from Landsat Thematic Mapper data', *Ann. Glaciol.* **9**, 97–103.

Dubayah, R., Dozier, J. & Davis, F. W. (1990), 'Topographic distribution of clear-sky radiation over the Konza Prairie, Kansas, USA', *Water Resour. Res.* **26**, 679–690.

Dubayah, R. & Loechel, S. (1997), 'Modeling topographic solar radiation using GOES data', *J. Appl. Meteor.* **36**, 141–154.

Dubayah, R. & Rich, P. M. (1995), 'Topographic solar radiation models for GIS', *Int. J. Geogr. Inf. Syst.* **9**, 405–419.

Duguay, C. R. (1993), 'Radiation modeling in mountainous terrain: Review and status', *Mountain Res. Develop.* **13**, 339–357.

Dunlop, C. & Clark, P. (1997), Forcing the single column UM from the mesoscale model, Technical Report 255, UK MetOffice.

Durran, D. R. & Klemp, J. B. (1982), 'On the effects of moisture on the Brunt-Väisälä frequency', *J. Atmos. Sci.* **39**, 2152–2158.

Duynkerke, P. G. (1990), 'Radiation fog: A comparison of model simulation with detailed observations', *Mon. Wea. Rev.* **119**, 324–341.

Ebisuzaki, W. & Kalnay, E. (1991), Ensemble experiments with a new lagged average forecasting scheme, Report 15, WMO, Research activities in atmospheric and oceanic modeling, Geneva, Switzerland.

Ehrendorfer, M. & Tribbia, J. J. (1997), 'Optimal prediction of forecast error covariances through singular vectors', *J. Atmos. Sci.* **54**, 286–313.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. & Tarpley, J. D. (2003), 'Implementation of NOAH land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model', *J. Geophys. Res.* **108**, doi:10.1029/2002JD003296.

Epstein, E. S. (1969), 'Stochastic-dynamic prediction', *Tellus* **21**, 739–759.

Estournel, C. (1988), Etude de la phase nocturne de la couche limite atmospherique, These doctorat d'etat nr. 1361, Université Paul Sabatier, Toulouse, France.

Estournel, C. & Guedalia, D. (1987), 'A new parameterization of eddy diffusivities for nocturnal boundary layer modeling', *Boundary Layer Meteorology* **39**, 191–203.

Evensen, G. (1994), 'Sequential data assimilation wiht a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics', *J. Geophys. Res.* **99**(C5), 10143–10162.

Eyre, J. R. (2002), Inversion methods for satellite sounding data, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Fels, S. B. & Schwarzkopf, M. D. (1975), 'The simplified exchange approximation: A new method for radiative transfer calculations', *Mountain Res. Develop.* **32**, 1475–1488.

Ferrier, B. (2002), A new grid-scale cloud and precipitation scheme in the NCEP Eta model, *in* 'Spring Colloquium on the Physics of Weather and Climate: Regional weather prediction modelling and predictability', Trieste, Italy.

Finley, J. P. (1884), 'Tornado predictions', *Amer. Met. J.* **1**, 85–88.

Fisher, M. (2002*a*), Assimilation techniques: 3dVar, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Fisher, M. (2002*b*), Assimilation techniques: 4dVar, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Fisher, M. (2002*c*), Assimilation techniques: Approximate kalman filters and singular vectors, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Fitzjarrald, D. R. & Lala, G. G. (1989), 'Hudson Valley fog environment', *J. Appl. Meteor.* **28**, 1303–1328.

Fouquart, Y. & Bonnel, B. (1984), 'Computations of solar heating of the Earths atmosphere: a new parameterization', *Beiträge zur Physik der Atmosphäre* **53**, 35–62.

Fritsch, J. M., Hilliker, J., Ross, J. & Vislocky, R. (2000), 'Model consensus', *Wea. Forecasting* **15**, 571–582.

Ghil, M. & Malanotte-Rizzoli, P. (1991), 'Data assimilation in meteorology and oceanography', *Adv. Geophys.* **33**, 141–266.

Ghosh, S. & Jonas, P. R. (1998), 'On the application of the classic Kessler and Berry schemes in large eddy simulation models with particular emphasis on cloud autoconversion, the onset time of precipitation and droplet evaporation', *Annales Geophysicae* **16**, 628–637.

Golding, B. W. (2002), Review of state of the art of prediction methods for fog and low cloud, *in* 'COST-722: Very short range forecasting of fog and low clouds Inventory phase on current knowledge and requirements by users and forecasters: Phase 1 Report', pp. 15–24.

Gultepe, I., Müller, M. D. & Boybeyi, Z. (2005), 'A new visibility parameterization for warm fog applications in numerical weather prediction models', *J. Appl. Meteorology* p. submitted.

Hamill, T. M. & Snyder, C. (2000), 'A hybrid ensemble Kalman filter - 3D variational analysis scheme', *Monthly Weather Review* **128**, 2905–2919.

Hamill, T. M., Whitaker, J. S. & Snyder, C. (2001), 'Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter', *Monthly Weather Review* **129**, 2776–2790.

Harshvardhan, Randall, D. A., Corsetti, T. G. & Dazlich, D. A. (1989), 'Earth radiation budget and cloudiness simulations with a general circulation model', *J. Atmos. Sci.* **46**, 1922–1942.

Heidke, P. (1926), 'Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst', *Geogr. Ann.* **8**, 301–349.

Hoffman, R. N. & Kalnay, E. (1983), 'Lagged average forecasting, an alternative to Monte Carlo forecasting', *Tellus* **35A**, 100–118.

Hollingsworth, A. & Lönnberg, P. (1986), 'The statistical structure of short-range forecast errors as determined from radiosonde data, part I: The wind field', *Tellus* **38A**, 111–136.

Hollingsworth, A., Shaw, D., Lönnberg, P., Illari, L., Arpe, K. & Simmons, A. (1986), 'Monitoring of observation and analysis quality as a data assimilation system', *Monthly Weather Review* **114**, 861–879.

Hólm, E. V. (2003), Lecture notes on assimilation algorithms, Technical report, European Center for Medium Range Weather Forecast ECMWF.

Hong, S.-Y., Juang, M.-M. H. & Zhao, Q. (1998), 'Implementation of prognostic cloud scheme for a regional spectral model', *Mon. Wea. Rev.* **126**, 2621–2639.

Hou, D., Kalnay, E. & Droegemeier, K. K. (2001), 'Objective verification of the SAMEX'98 ensemble forecast', *Mon. Wea. Rev.* **129**, 73–91.

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. & Mitchell, H. L. (1996), 'A system simulation approach to ensemble prediction', *Mon. Wea. Rev.* **124**, 1225–1242.

Houtekamer, P. L. & Mitchell, H. L. (1998), 'Data assimilation using an ensemble Kalman filter technique', *Monthly Weather Review* **126**, 796–811.

Houtekamer, P. L. & Mitchell, H. L. (2001), 'A sequential ensemble Kalman filter for atmospheric data assimilation', *Monthly Weather Review* **129**, 123–137.

Houze, R. A. (1993), *Cloud Dynamics*, Academic Press, San Diego, California.

Ide, K., Courtier, P., Ghil, M. & Lorenc, A. C. (1997), 'Unified notation for data assimilation: operational, sequential and variational', *J. Meteorol. Soc. Japan* **75**(1B), 181–189.

Ingleby, N. B. (2001), 'The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme', *Q. J. R. Meteorol. Soc.* **127**, 209–231.

Janisková, M., Mahfouf, J.-F., Morcrette, J.-J. & Chevallier, F. (2002), 'Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation.', *Quart. J. Roy. Meteor. Soc.* **128**, 1505–1527.

Janjic, Z. I. (2003), 'A nonhydrostatic model based on a new approach', *Meteorology and Atmospheric Physics* **82**, 271–285.

Janjic, Z. I., Gerrity, J. P. & Nickovic, S. (2001), 'An alternative approach to nonhydrostatic modeling', *Monthly Weather Review* **129**, 1164–1178.

Jolliffe, I. T. & Stephenson, D. B. (2003), *Forecast Verification - A practitioner's guide in atmospheric science*, Wiley, Chichester.

Kalnay, E. (2003), *Atmospheric modeling, data assimilation and predictability*, Cambridge University Press, Cambridge.

Kalnay, E. & Toth, Z., eds (1996), *The breeding method*, Proceedings of the Seminar on Predictability, ECMWF, Reading, England.

Kessler, E. (1969), 'On the distribution and continuity of water substance in atmospheric circulations', *Meteor. Monog.* **10**(32), 84.

Köhler, H. (1936), 'The nucleus in and the growth of hygroscopic droplets', *Trans. Faraday Soc.* **32**, 1152–1161.

Kondratyev, K. Y. (1977), Radiation regime of inclined surfaces, Tech. Note 152, WMO.

Krishnamurti, T. N., Kishtawal, C. M., Zang, Z., LaRow, T., Bachiochi, D., Williford, E., Gadgil, S. & Surendran, S. (2000), 'Multimodel ensemble forecast for weather and seasonal climate', *J. Climate* **13**, 4196–4216.

Kumar, L., Skidmore, A. K. & Knowles, E. (1997), 'Modeling topographic variation in solar radiation in a GIS environment', *Int. J. Geogr. Inf. Syst.* **11**, 475–497.

Kunkel, B. (1984), 'Parameterization of droplet terminal velocity and extinction coefficient in fog model', *J. Appl. Meteor.* **23**, 34–41.

Lacis, A. A. & Hansen, J. E. (1974), 'A parameterization for the absorption of solar radiation in the earth's atmosphere', *J. Atmos. Sci.* **31**, 118–133.

Leith, C. E. (1974), 'Theoretical skill of Monte Carlo forecasts', *Mon. Wea. Rev.* **102**, 409–418.

Lorenc, A. C. (1986), 'Analysis methods for numerical weather prediction', *Quart. J. Roy. Met. Soc.* **112**, 1177–1194.

Lorenz, E. N. (1963*a*), 'Deterministic nonperiodic flow', *J. Atmos. Sci.* **20**, 130–141.

Lorenz, E. N. (1963*b*), 'The predictability of hydrodynamic flow', *Trans. NY Acad. Sci. Series II* **25**, 409–432.

Lorenz, E. N. (1965), 'A study of the predictability of a 28-variable atmospheric model', *Tellus* **17**, 321–333.

Mahrt, L. & Ek, M. (1984), 'The influence of atmospheric stability on potential evaporation', *J. Clim. Appl. Meteorol.* **23**, 222–234.

Mahrt, L. & Mahrt, L. (1987), 'Interaction between soil hydrology and boundary-layer development', *Bound.-Layer Meteorol.* **38**, 185–202.

Mahrt, L. & Pan, H.-L. (1984), 'A two-layer model of soil hydrology', *Bound.-Layer Meteorol.* **29**, 1–20.

Mason, B. J. (1971), *The physics of clouds*, Oxford Press, Oxford.

Mason, I. B. (1980), Decision-theoretic evaluation of probabilistic predictions, *in* 'WMO Symposium on Probabilistic and Statistical Methods in Weather Forecasting', Nice, France, pp. 219–228.

Mason, I. B. (1982), 'A model for assesment of weather forecasts', *Austral. Met. Mag.* **30**, 291–303.

Mason, I. B. (1989), 'Dependence of the critical success index on sample climate and threshold probability', *Austral. Met. Mag.* **37**, 75–81.

Matzinger, N., Andretta, M., van Gorsel, E., Vogt, R., Ohmura, A. & Rotach, M. W. (2003), 'Surface radiation budget in an alpine valley', *Quart. J. Roy. Meteor. Soc.* **129**, 877–895.

Miles, N. L., Verlinde, J. & Clothiaux, E. E. (2000), 'Cloud droplet size distributions in low-level stratiform clouds', *Journ. Atmos. Sci.* **15**, 295–311.

Molteni, F., Buizza, R., Palmer, T. N. & Petroliagis, T. (1996), 'The new ECMWF Ensemble Prediction System: Methodology and validation', *Quart. J. Roy. Meteor. Soc.* **122**, 73–119.

Molteni, F. & Palmer, T. N. (1993), 'Predictability and finite time instability of the northern winter circulation', *Quart. J. Roy. Meteor. Soc.* **119**, 269–298.

Müller, M. D. & Scherer, D. (2005), 'A grid- and subgrid-scale radiation parameterization of topographic effects for mesoscale weather forecast models', *Mon. Wea. Rev.* **133**, 1431–1442.

Murphy, A. H. & Winkler, R. L. (1987), 'A general framework for forecast verification', *Monthly Weather Review* **115**, 1330–1338.

Musson-Genon, L. (1987), 'Numerical simulations of a fog event with a one-dimensional boundary layer model', *Monthly Weather Review* **115**, 592–607.

Noilhan, J. & Mahfouf, J.-F. (1996), 'The ISBA land surface parameterisation scheme', *Glob. Planet. Change* **13**, 145–159.

Oliphant, A. J., Spronken-Smith, R. A., Sturman, A. P. & Owens, I. F. (2003), 'Spatial variability of surface radiation fluxes in mountainous terrain', *Quart. J. Roy. Meteor. Soc.* **42**, 113–128.

Parrish, D. F. & Derber, J. C. (1992), 'The national meteorological center's spectral statistical-interpolation analysis system', *Monthly Weather Review* **120**, 1747–1763.

Polcher, J., McAvaney, B., Viterbo, P., Gaertner, M.-A., Hahmann, A., Mahfouf, J.-F., Noilhan, J., Phillips, T., Pitman, A., Schlosser, C., Schulz, J.-P., Timbal, B., Verseghy, D. & Xue., Y. (1998), 'A proposal for a general interface between land-surface schemes and general circulation models', *Global and Planetary Change* **19**, 261–276.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1998), *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York.

Sakakibara, H. (1979), 'A scheme for stable numerical computation of the condensation process with large time steps', *J. Meteorol. Soc. Japan* **57**, 349–353.

Scherer, D. & Parlow, E. (1994), 'Terrain as an important controlling factor for climatological, meteorological and hydrological processes in NW-Spitsbergen', *Z. Geomorphol. N.F.* **97**(Suppl. Bd.), 175–193.

Schlatter, T. W. (2000), 'Variational assimilation of meteorological observations in the lower atmosphere: a tutorial on how it works', *J. Atmos. Sol. Terr. Phys.* **62**, 1057–1070.

Schwarzkopf, M. D. & Fels, S. B. (1985), 'Improvements to the algorithm for computing CO2 transmissivities and cooling rates', *J. Geophys. Res.* **90**, 10541–10550.

Schwarzkopf, M. D. & Fels, S. B. (1991), 'The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes', *J. Geophys. Res.* **96**, 9075–9096.

Siebert, J., Bott, A. & Zdunkowski, W. (1992*a*), 'Influence of a vegetation-soil model on the simulation of radiation fog', *Beitr. Phys. Atmos.* **65**, 93–106.

Siebert, J., Bott, A. & Zdunkowski, W. (1992*b*), 'A one-dimensional simulation of the interaction between land surface processes and the atmosphere', *Boundary - Layer Meteorol.* **59**, 1–34.

Slingo, J. M. (1987), 'The development and verification of a cloud prediction model for the ECMWF model', *Quart. J. Roy. Meteor. Soc.* **113**, 899–927.

Smith, R. N. B. (1990), 'A scheme for predicting layer clouds and their water content in a general circulation model', *Quart. J. Roy. Meteor. Soc.* **116**, 435–460.

Soong, S.-T. & Ogura, Y. (1973), 'A comparison between axisymmetric and slab-symmetric cumulus cloud models', *J. Atmos. Sci.* **30**, 879–893.

Sundqvist, H. (1978), 'A parameterization scheme for non-convective condensation including prediction of cloud water content', *Quart. J. Roy. Meteor. Soc.* **104**, 677–690.

Sundqvist, H., Berge, E. & Kristjansson, J. E. (1989), 'Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model', *Mon. Wea. Rev.* **117**, 1641–1657.

Swets, J. A. (1986), 'Indices of discrimination of diagnostic accuracy: their ROCs and implied models', *Psychol. Bull.* **99**, 100–117.

Talagrand, O. (1990), 'Assimilation of observations, an introduction', *J. Met. Soc. Japan* **75**(Special Issue), 191–209.

Turton, J. D. & Brown, R. (1987), 'A comparison of a numerical model of radiation fog with detailed observations', *Quart. J. Roy. Meteor. Soc.* **113**, 37–54.

Twomey, S. (1959), 'The nuclei of natural cloud formation. part II: The supersaturation in natural clouds and the variation of cloud droplet concentration', *Geophys. Pura Appl.* **43**, 243–249.

Vehil, R., Monerris, J., Guedalia, D. & Sarthou, P. (1989), 'Study of the radiative effects within a fog layer', *Atmos. Res.* **23**, 179–194.

von Glasow, R. & Bott, A. (1999), 'Interaction of radiation fog with tall vegetation', *Atmos. Environ.* **33**, 1333–1346.

Whiteman, C. D., Allwine, K. J., Fritschen, L. J., Orgill, M. M. & Simpson, J. R. (1989), 'Deep valley radiation and surface energy budget microclimates. Part I: Radiation', *J. Appl. Meteor.* **28**, 414–426.

Wilson, D. R. & Ballard, S. P. (1999), 'A microphysically based precipitation scheme for the UK Meteorological Office Unified Model', *Quart. J. Roy. Meteor. Soc.* **125**, 1607–1636.

Wu, W.-S., Purser, R. J. & Parrish, D. F. (2002), 'Three-dimensional variational analysis with spatially inhomogeneous covariances', *Mon. Wea. Rev.* **130**, 2905–2916.

Xu, K.-M. & Randall, D. A. (1996), 'A semiempirical cloudiness parameterization for use in climate models', *J. Atmos. Sci.* **53**, 3084–3102.

Zdunkowski, W. & Barr, A. (1972), 'A radiative-conductive model for the prediction of radiation fog', *Bound.-Layer Meteor.* **3**, 152–157.

Zdunkowski, W. & Nielsen, B. (1969), 'A preliminary prediction analysis of radiation fog', *Pure Appl. Geophys.* **19**, 45–66.