

## A One-dimensional Ensemble Forecast and Assimilation System for Fog Prediction

M. D. MÜLLER,<sup>1</sup> C. SCHMUTZ,<sup>2</sup> and E. PARLOW<sup>1</sup>

*Abstract*—A probabilistic fog forecast system was designed based on two high resolution numerical 1-D models called COBEL and PAFOG. The 1-D models are coupled to several 3-D numerical weather prediction models and thus are able to consider the effects of advection. To deal with the large uncertainty inherent to fog forecasts, a whole ensemble of 1-D 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 3-D models. The design of the ensemble scheme computes members that should fairly well represent the uncertainty of the current meteorological regime. Verification for an entire fog season reveals the importance of advection in complex terrain. The skill of 1-D fog forecasts is significantly improved if advection is considered. Thus the probabilistic forecast system has the potential to support the forecaster and therefore to provide more accurate fog forecasts.

**Key words:** Fog, one-dimensional, ensemble prediction, assimilation, model coupling, advection, verification.

### 1. Introduction

Reductions in visibility have an important impact on the capacity of an airport. Different initiatives by the International Civil Aviation Organization (ICAO) and also by Eurocontrol endeavor to address this problem. In Europe the Meteorological support for Air Traffic Management Group (METATMG) has, as one of its tasks, to study the possibilities to improve visibility and runway visual range (RVR) forecasts (METATMG, 2005). A large industry project (SESAR) under the lead of Eurocontrol intends to increase the air traffic capacity in Europe in the coming years by factors on different levels (EUROCONTROL, 2006), meteorology along with visibility forecast is one of the subtasks of the project. Imperfect visibility

---

<sup>1</sup> Institute of Meteorology, Climatology & Remote Sensing, University of Basel, Klingelbergstr. 27, Basel, Switzerland. E-mail: mathias.mueller@unibas.ch

<sup>2</sup> Eidgenössisches Departement des Innern EDI, Bundesamt für Meteorologie und Klimatologie MeteoSchweiz, CH-8058, Zürich-Flughafen, Schweiz.

forecast always has an adverse impact on capacity regulated airports. It results in an overload when actual capacity is lower than expected, and in a capacity loss when conditions and the actual capacity are better than expected. In 2003 Swiss International Airlines estimated the accumulated delay due to one hour of erroneous visibility forecast in the morning for Zürich airport to 1400 minutes throughout the day (Werner Suhner, Swiss Int. Airlines, pers. comm., 2003).

Fog and visibility forecasts are provided in the Terminal Area Forecast (TAF) code for the so-called terminal area (airport). It is current practice to provide point forecast for such locations and the two main approaches are based either on statistical methods or on numerical models. Statistical approaches combine long records of site-specific observations with forecast variables generally provided by numerical weather prediction. A human forecast is, to a large extent, based on experience, and thus a subjective system of pattern recognition and climatological knowledge. For this work the focus lies on the second approach, using dedicated models for fog prediction without the need for long statistical training records.

The formation and dispersion of fog is the result of a 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 3-D model, is also needed. In complex topography cold air outflow and pooling as well as advection in the heterogeneous landscape become very important. Once the fog has formed there are further influences:

- longwave radiative cooling at fog top,
- fog microphysics,
- shortwave radiation.

Starting with the work of ZDUNKOWSKI and NIELSEN (1969) some of the above listed processes were included in fog models. In this early model there was no parameterization for the sedimentation of liquid water nor turbulence exchange coefficients. The latter were introduced by ZDUNKOWSKI and BARR (1972). An even more sophisticated model was developed by BROWN and ROACH (1976) and further

refined by TURTON and BROWN (1987), including new formulations for turbulence exchange coefficients in 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. BERGOT and GUÉDALIA (1994b) illustrated the importance of advection terms and their role in fog formation and evolution. BALLARD *et al.* (1991) used a numerical weather prediction model to simulate sea fog and pointed out the importance of initial conditions and vertical resolution. GOLDING (1993) found that the development of local nocturnal winds in complex terrain often determines the location and timing of fog formation. SIEBERT *et al.* (1992a), SIEBERT *et al.* (1992b) and VON GLASOW and 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. Currently parameterized versions of the detailed 1-D fog microphysics models are incorporated in 3-D models and are able to improve visibility forecasts (GULTEPE *et al.*, 2006; GULTEPE and MILBRANDT, 2007).

In this work an ensemble forecast system based on numerical prediction models is developed and tested. 1-D models are coupled to 3-D models so that they also can be used for advection fog and valley fog situations and are thus suitable for an operational forecasting system.

## 2. Site-specific Characteristics and Available Data

Zürich airport (432 m a.s.l.) is located in a small basin north of the City of Zürich, Switzerland. It is surrounded by hills with an average altitude of 500 m to 600 m a.s.l. Some higher hills are between 700 m and 850 m a.s.l. In the past the airport area used to be a moorland with marsh. Fog often forms in the area of the airport. Table 3 reveals that the season with the highest occurrence of fog is between September and March. SCHMUTZ *et al.* (2004) also show that the highest frequency of low visibility (here defined to be < 800 m) during the day is observed between 5 a.m. and 10 a.m. local time from December to February.

A broad spectrum of observations, listed in Table 1, is available at the airport. Furthermore a virtual profile can be constructed from standard height observations made by stations located on nearby hilltops (Table 2). For high accuracy of data assimilation it is best to directly measure profiles of temperature and humidity up to a height of approximately 2 km, which is the upper boundary of the models. 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.

Table 1  
*Available observations at and in the vicinity of the Zürich airport*

Parameter
visibility (spatially aggregated)
precipitation rate
precipitation type
soil temperature (-0.05, -0.1, -0.2, -1.0 m)
soil moisture (-0.01, -0.02, -0.1, -0.25, -0.6, -0.98 m)
“surface” temperature (0.02 m)
temperature (0.5 m)
temperature (all heights of virtual profile)
temperature profile (MTP-5)
wind profile
humidity (all heights of virtual profile)
u-wind (all heights of virtual profile)
v-wind (all heights of virtual profile)
cloud base
cloud cover
longwave radiation LW↓
longwave radiation LW↑
shortwave radiation SW↓
Radiosonde from Payerne (150 km away)

For temperature it was possible to use an MTP-5 microwave profiler (KADYGROV and PICK, 1998) but nothing comparable was available for humidity. In order to obtain entire profiles of temperature and humidity, an assimilation scheme is necessary to initialize the forecast models.

### 3. The Ensemble Forecast System

Fog forecasting is a threshold problem, namely a small difference in temperature and/or humidity determines if condensation occurs and thus fog forms or not. Keeping in mind the difficulties of providing good temperature and humidity

Table 2

*Statistics of temperature deviations between the MTP-5 and the stations of the virtual temperature profile. Height in m above sea level (m a.s.l.), also indicated is the number of 10 min intervals with positive or negative temperature deviation from the MTP-5 profile*

Station	Height (m a.s.l.)	RMS (K)	Mean (K)	negative	positive
Zürich airport	432	0	0	0	
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

Table 3

*Monthly frequencies (%) of low visibility (vis) as observed between 1993 and 2002 at Zürich airport. (SCHMUTZ et al. 2004)*

	vis <800 m	vis <1500 m
Jan.	6.0	8.8
Feb.	2.5	4.1
Mar.	1.4	2.2
Apr.	0.9	1.2
May.	1.0	1.2
Jun.	0.4	0.6
Jul.	0.4	0.6
Aug.	0.9	1.3
Sep.	4.0	4.9
Oct.	7.2	8.7
Nov.	5.4	7.6
Dec.	3.5	5.2

forecasts and the dependence of these forecasts on initial conditions, a deterministic forecast incorporates rather large uncertainties. This problem can be addressed with ensemble forecasting, which not necessarily provides a more accurate forecast for all conditions although it can provide a likelihood of fog occurrence and thus also inform about the predictability of a particular situation.

The developed ensemble forecast system consists of an assimilation system to generate a set of initial conditions based on prior forecasts and current observations, two distinct 1-D forecast models and a post-processor.

### *Data Assimilation Strategy*

An important part of every ensemble forecast is the derivation of a set of initial conditions representative of the current uncertainty of the initial state.

The process of data assimilation optimally combines observations with a first guess or background estimate. In this case data have to be assimilated by a 1-D model, and it seems natural to use a previous forecast of that model as background state. There are however several reasons for using the 3-D model forecasts, having resolutions between 2 and 7 km, as a first guess. In 1-D it is not possible to simulate horizontal gradients which are responsible for advection and wind so that the 1-D model cannot simulate its own background state needed in the data assimilation process. Basically the 1-D model is unaware of changes in temperature, humidity and wind caused by advection so that large errors develop over time. Since the 1-D model is operated at a location where observed humidity and wind is based on measurements taken only a few meters above ground, the assimilation process cannot correct the background state in most parts of the vertical profile. Another big problem is the limited vertical extent of the 1-D model, which currently simulates the lowest 2000 m of the atmosphere. Thus mid-and high-altitude clouds and their effects on radiative

fluxes are not within the domain of the 1-D models. To simulate the background state with a 1-D model it would thus be necessary to run it in an “assimilation mode” which incorporates information about advection and radiative fluxes from 3-D models. Unfortunately, due to post-processing time constraints and large data amounts, a 3-D model output is only available with very limited temporal resolution of 1 hour or even worse. The accuracy of tendencies derived from these temporally heavily discretized data is rather limited (DUNLOP and CLARK 1997). To overcome these problems related to the strong dependence of the 1-D assimilation model on 3-D data, vertical columns of the 3-D models are used directly as background terms in the assimilation process. A problem is that close to the surface the 3-D profiles are less detailed than the 1-D forecasts, due to a coarser vertical resolution. However, considering the whole profile, the surface layer contains only slight energy and the high resolution 1-D model adjusts the surface layer profiles in a relatively short time. Also the assimilation will correct a suboptimal surface layer background from the 3-D model, using all the observations. 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 3-D model is able to do skillful forecasts for several days, where the most recent one is not necessarily the best and the skill of surface variables such as temperature, pressure and wind is generally similar within the first 48 hours as can be seen e.g., in BERNARDET *et al.* (2005) or LIN *et al.* (2005). 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.

We therefore decided to base our initial conditions on available forecasts computed from different 3-D models. Currently the aLMO (STEPPELER *et al.*, 2003) of MeteoSwiss, running at 7 km resolution, as well as the semi-operational forecasts with the Nonhydrostatic Mesoscale Model (NMM) (JANJIC *et al.*, 2001; JANJIC, 2003) at resolutions of 22, 4 and 2 km from the University of Basel are available for the study area of Zürich airport. The different 3-D models driven by distinct global assimilations, with varying resolutions as well as different initialization times produce a spread in forecast fields which are used as initial conditions by the 1-D models. This spread is dependent on the current weather situation.

Certainly four different forecasts from four different models do not produce a large set of initial conditions. To increase the number of 3-D forecasts, all the runs valid at the same time but initialized at different times are 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. This procedure is summarized in Figure 1. Note that an individual assimilation is computed from every 3-D run using the corresponding error covariance matrix  $\mathbf{B}$ , as derived in the next section. The number of members is finally doubled by using two different numerical 1-D models to integrate all initial conditions.

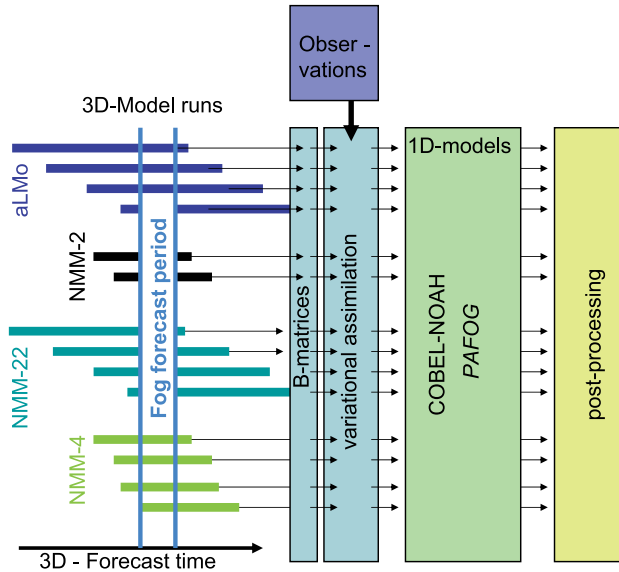


Figure 1

Schematic of the 1-D ensemble prediction system. Every 3-D run provides initial conditions that are used as a background for an individual variational assimilation. Since the 3-D models also provide boundary conditions, a 3-D run has to cover the entire 1-D fog forecast period.

### Data Assimilation System

Data assimilation combines observations and a dynamic model by using specified 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 (1) (KALNAY, 2003).

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

Here the following notation proposed by IDE *et al.* (1997), is 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  $p \times p$ ).

In order 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.

$$\mathbf{U}^T \mathbf{B}^{-1} \mathbf{U} = \mathbf{I}. \quad (2)$$

It is thus necessary to transform the 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:

$$\vec{x} = \mathbf{U} \vec{v}, \quad (3)$$

$$\vec{v} = \mathbf{T} \vec{x}. \quad (4)$$

By doing so the cost function to be evaluated becomes

$$J(\vec{v}) = \frac{1}{2} (\vec{v}^T \vec{v}) + \frac{1}{2} (\vec{y} - \mathbf{H} \mathbf{U} \vec{v})^T R^{-1} (\vec{y} - \mathbf{H} \mathbf{U} \vec{v}). \quad (5)$$

During the minimization process, the cost function and the gradient of (5), as derived by BOUTTIER and COURTIER (1999) and transformed into the incremental form in  $\vec{v}$ -space (6), have to be evaluated during every step of the minimization.

$$\nabla_{\vec{v}} J(\vec{v}) = \vec{v} - \mathbf{U}^T \mathbf{H}^T \mathbf{R}^{-1} (\vec{y} - \mathbf{H} \mathbf{U} \vec{v}). \quad (6)$$

In the current implementation, the Broyden-Fletcher-Goldfarb-Shanno variant of the Davidon-Fletcher-Powell method, as described in PRESS *et al.* (1988), is used to compute the minimization of the cost function with the help of its gradient.

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. Here, the ‘‘NMC’’ or NCEP Method (PARRISH and DERBER, 1992) is used which is independent of measurements.

$$\mathbf{B} \approx \alpha \frac{1}{n} \sum_{i=1}^n \left( (\vec{x}_f(t_1) - \vec{x}_f(t_0)) (\vec{x}_f(t_1) - \vec{x}_f(t_0))^T \right), \quad (7)$$

where  $\vec{x}_f$  represents the forecast state vector and  $\alpha$  is an empirical scaling factor. As can be seen in (7)  $\mathbf{B}$  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.

The estimates of  $\mathbf{B}$  obtained for different 3-D models using the NMC-Method (7) and corresponding correlations are shown in Figure 2. The statistics are based on the vertical profiles of the 3-D models from October 2004 to February 2005. In Figure 2, it is evident that variances are largest close to the surface, where 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



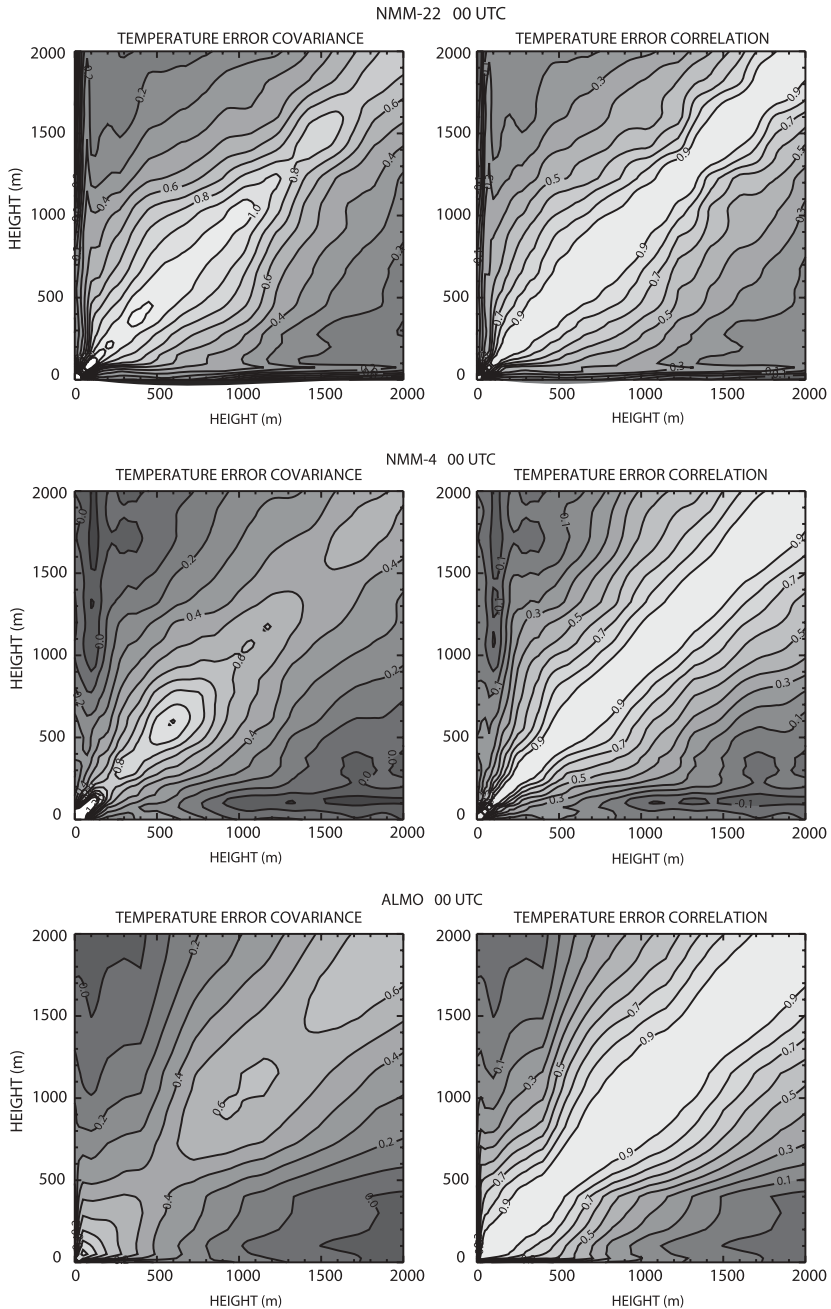


Figure 2  
Temperature error covariance and error correlation matrices for the winter season 2004/2005 at 00 UTC for 3 different 3-D models.

and an unreliable forecast for that layer does not pose a problem for data assimilation. For the two resolutions of NMM, a second 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 larger. Note that the structure of the error covariance is different throughout the day due to the diurnal evolution of the boundary layer, which was also analyzed by HACKER and SNYDER (2005).

Similar to the background error covariance matrix,  $\mathbf{R}$  specifies errors of the observational system. The error is mainly caused by representativeness problems of the observation in model space and only secondly on instrumental characteristic. In contrast to  $\mathbf{B}$ , correlations are assumed to be zero. Concerning the assimilation at Zürich airport, the most difficult part of  $\mathbf{R}$  is assessing 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 unrepresentative. To quantify the similarity between the two locations, model profiles from high resolution numerical weather prediction were analyzed. According to the amount of resolved topography, 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 3 for radiosonde ascent time of 1200 UTC.

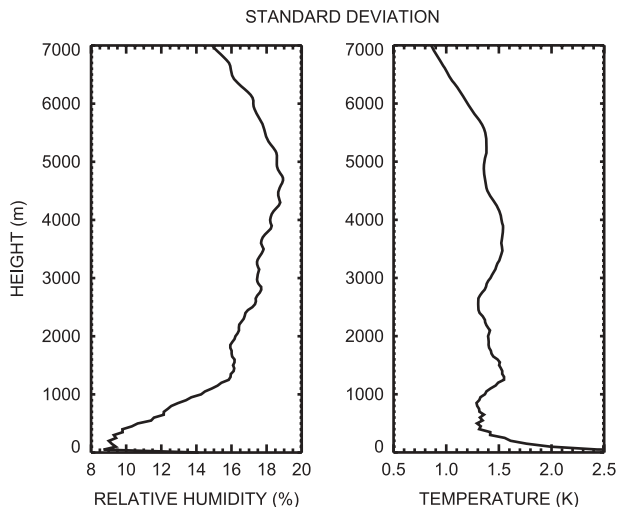


Figure 3

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

The virtual profile of temperature and humidity is problematic in 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 when they reflect more the local surface layer conditions than the free atmosphere. To quantify the errors of the virtual profile, mean and root-mean-squares 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 MTP-5 observation period from 26 October, 2004 until 11 April, 2005. In Table 2 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. As can be seen, the virtual profile 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 important role of radiative cooling of the surface. Thus the stations located at the surface can only provide an estimate of the thermodynamic state of the free atmosphere at that height, nonetheless there is still useful information which can be exploited by data assimilation.

### *Boundary Conditions*

The specification of boundary conditions is used to extend the applicability of the 1-D model to more heterogeneous environments and different synoptic 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. Advection of  $a$  as defined by (8) is specified as an external tendency. It is computed using centered finite differences (9) about the point of interest  $(i, j)$ , which requires four additional vertical columns from the 3-D model.

$$\frac{\partial a}{\partial t} = -\vec{v} \cdot \vec{\nabla} a, \quad (8)$$

$$\frac{\Delta a_{i,j}}{\Delta t} = -u_{i,j} \frac{a_{i+1,j} - a_{i-1,j}}{2\Delta x} - v_{i,j} \frac{a_{i,j+1} - a_{i,j-1}}{2\Delta y}. \quad (9)$$

Another possibility in deriving advection is the use of a total tendency, rather than pure advection. Therefore, the total hourly change of humidity and temperature in the profile of the 3-D model is computed. Of course this change is not solely caused by advection but by all processes, like turbulent mixing, radiative cooling or phase changes. The dominating process depends on the actual situation, but the fog modeling system, which has to produce daily forecasts, needs to be evaluated and tested for all situations in order to be of operational use. This method is beneficial in that it requires only one column of 3-D model data at the point of interest and no neighboring columns. Furthermore temporal discretization errors in the case of

dominating advection are minimized since advection was computed internally by the 3-D model for every time step. Temporal discretization is a problem (DUNLOP and CLARK, 1997) when advection is computed using (9), since operational model output is available only on an hourly basis and not every time step.

Besides advection of temperature and humidity, radiation is another important boundary condition. It is significantly modified by clouds. Due to the limited vertical extent of the 1-D model, some clouds are above the model domain. However the effect of higher clouds can be included using the downward radiation computed by a 3-D model at the top boundary height of the 1-D model. If the radiation at a certain height is not available from the 3-D 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 3-D 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.

The advection term, which represents the boundary conditions for the 1-D 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. In the current implementation, the number of members is tripled. The same initial conditions are used in combination with no advection, the total tendency and the mesoscale advection. Currently, initial conditions and boundary conditions are taken from the same 3-D forecast run. It is however possible to further increase the number of ensemble members by mixing initial and boundary conditions derived from different 3-D runs.

### *Numerical 1-D Models*

The ensemble system uses two numerical 1-D models for fog prediction, namely, COBEL-NOAH and PAFOG.

The COBEL model (Couche Brouillard Eau Liquide) was originally derived from the 1-D model of the nocturnal boundary layer, developed by the Laboratoire d'Aérodologie of the Paul Sabatier University in Toulouse (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. A detailed description of the model together with some case studies is given in BERGOT and GUÉDALIA (1994a,b). However major modifications were made to the model (MÜELLER, 2006). It was coupled to the NOAH land surface model (NOAH-LSM) using an explicit flux coupling. The NOAH-LSM has a long heritage and originated from MAHRT and PAN (1984), MAHRT and EK (1984) and PAN and MAHRT (1987). Since then several major improvements were made (CHEN *et al.*, 1997; EK *et al.*, 2003) and it is currently used in the NCEP realtime Land Data Assimilation System. In our version of

COBEL, which we call COBEL-NOAH, the LSM computes the energy balance of the canopy layer and the evolution of temperature and humidity in the soil. Furthermore a parameterization of precipitation by KESSLER (1969) which considers autoconversion, accretion, evaporation of rain and computations of mean fall speeds was implemented into COBEL-NOAH.

The other model used in the ensemble system is the 1-D model for PArAmeterized FOG (PAFOG) and 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-NOAH is the considerably more detailed, but still parameterized, cloud microphysics module that allows to compute the total droplet number concentration. It is therefore possible to realistically compute droplet size-dependent sedimentation and supersaturation controlling condensation and evaporation. No major changes to the 1-D model PAFOG were done, so that an up-to-date description of the model can be found in BOTT and TRAUTMANN (2002) and references mentioned therein.

### *Post-processing*

Post-processing finally aggregates the individual forecast members 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. Figure 4 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 well, 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  $0.01 \text{ g kg}^{-1}$  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.

## *4. Results*

### *Method of Verification*

The verification is carried out using fog events, which are defined based on threshold values for observed visibility and modeled liquid water content. 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 other thresholds are also used to account for uncertainty in the observations and representativeness, since observed visibility at the

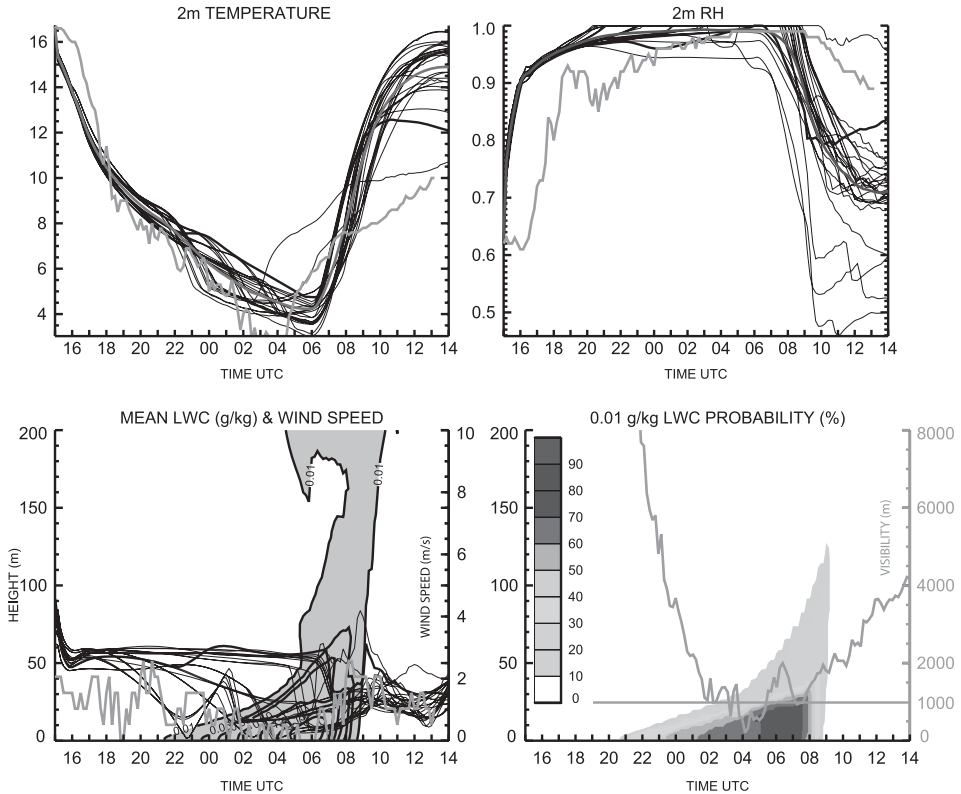


Figure 4

1-D 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 (white 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  $0.01 \text{ g kg}^{-1}$  is exceeded.

different runways of the airport often disagree due to the patchiness/inhomogeneity of fog as well as due to observational difficulties. 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  $0.01 \text{ g kg}^{-1}$ , which is a clear signal in the model output for the presence of liquid water. Note that higher thresholds also were used but resulted in worse verification scores.

For all results presented next, the verification period begins on 1 November, 2004 and ends on 30 April, 2005. Verification addresses the importance of advection and

uncertainties in the humidity assimilation, the impact of different initialization times as well as effects of the driving 3-D model. For every initialization time an ensemble consists of around 30 members, depending on the availability of output from semi-operational 3-D models, 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. 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 the verification of probabilistic forecasts 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. The *relative operating characteristic* (ROC) is obtained by plotting the hit rate versus the false alarm rate 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, whereas the diagonal indicates no skill. A popular overall skill measure 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.

### *Importance of Advection*

Advection might be insignificant in flat terrain, but will be important in a location such as Zürich airport. As there are some problems in estimating the amount of advection, two methods were proposed in section 3 and the results are indeed very different. Nevertheless we also want to look at what happens if advection is not considered at all. In Figure 5 ROC curves are given for different thresholds of observed visibility. It is evident that inclusion of advection significantly improves the forecast, if it is derived with (8), using gradients and wind. The other way of estimating advection as the total rate of change in the 3-D 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 3-D model does not allow an accurate simulation of these local processes. However all possible inaccuracies are transferred into the 1-D model when the total rate of change is used to determine advection.

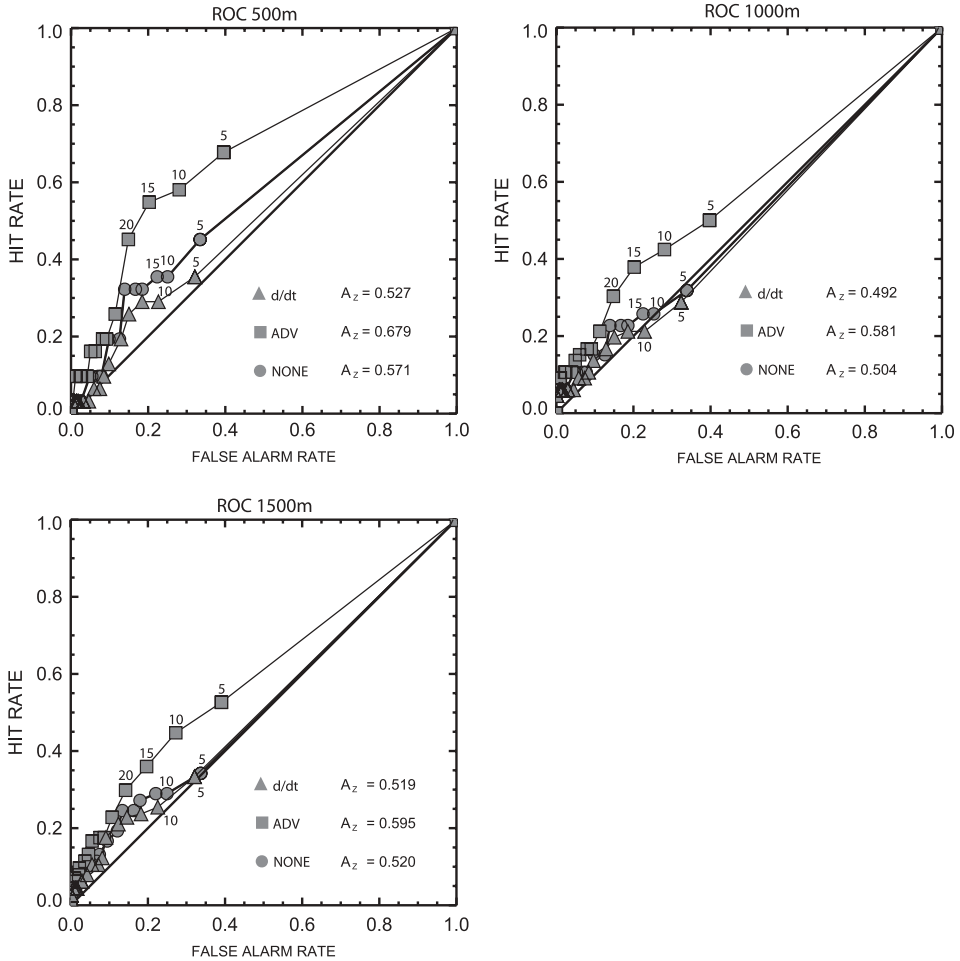


Figure 5

ROC for fog occurrence in the 1-D 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.

### *Different Initialization Times*

After an initial spin-up time, the forecast skill of a model decreases with increasing forecast length. Thus the later we initialize, but not beyond the spin-up time, the better our forecast should be for a given time window. However, 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. 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.

In Figure 6 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. 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 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 provide valuable hints for the synoptic forecaster about the most likely time of fog formation and dissipation.

### *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. This test does not intend to find the best humidity threshold but to detect a possible bias in the assimilation. By doing so each humidity class still has the same number of about 30 ensemble members. Note that the relative humidity was not allowed to exceed 99%, to avoid an artificial creation of fog at the beginning. The result is summarized in Figure 7, 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 as can be seen by higher  $A_z$  for the increased humidity forecasts.

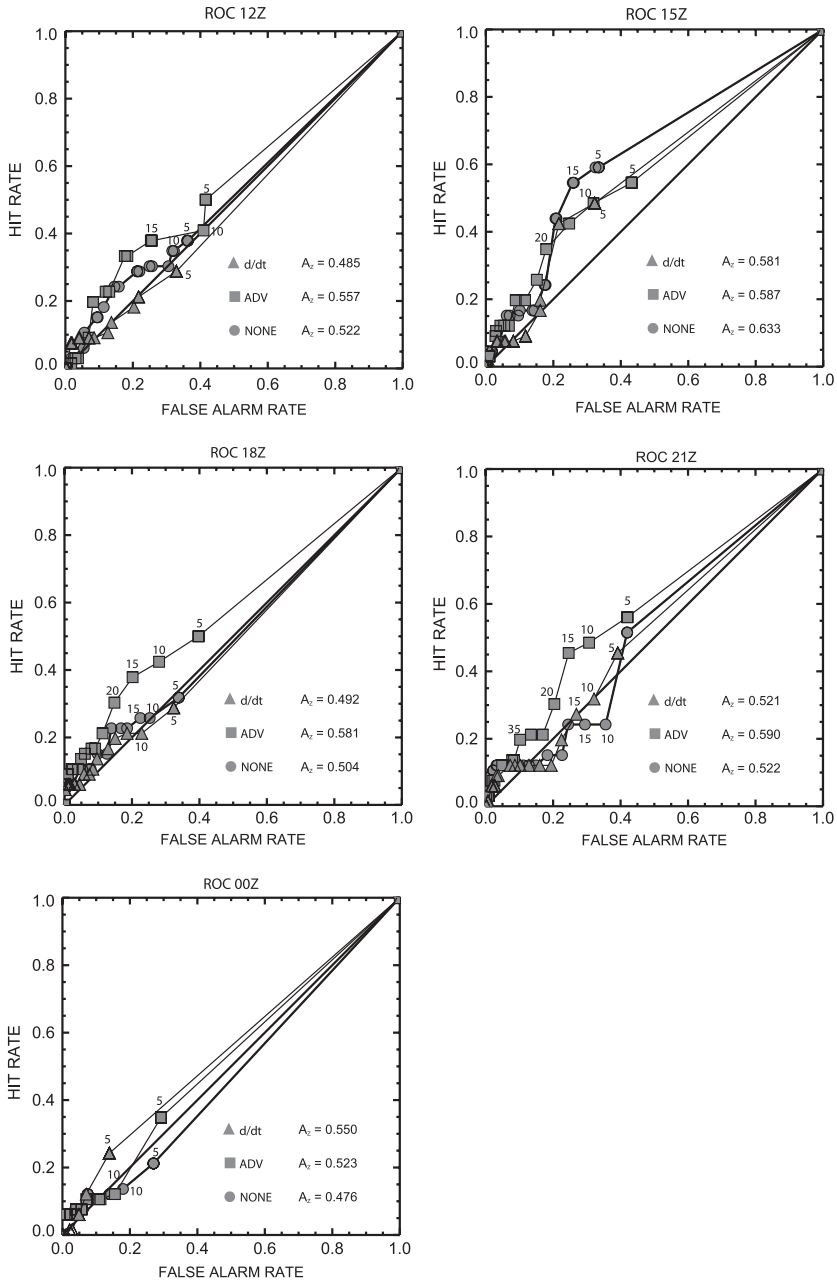


Figure 6

ROC for fog occurrence in the 1-D 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.

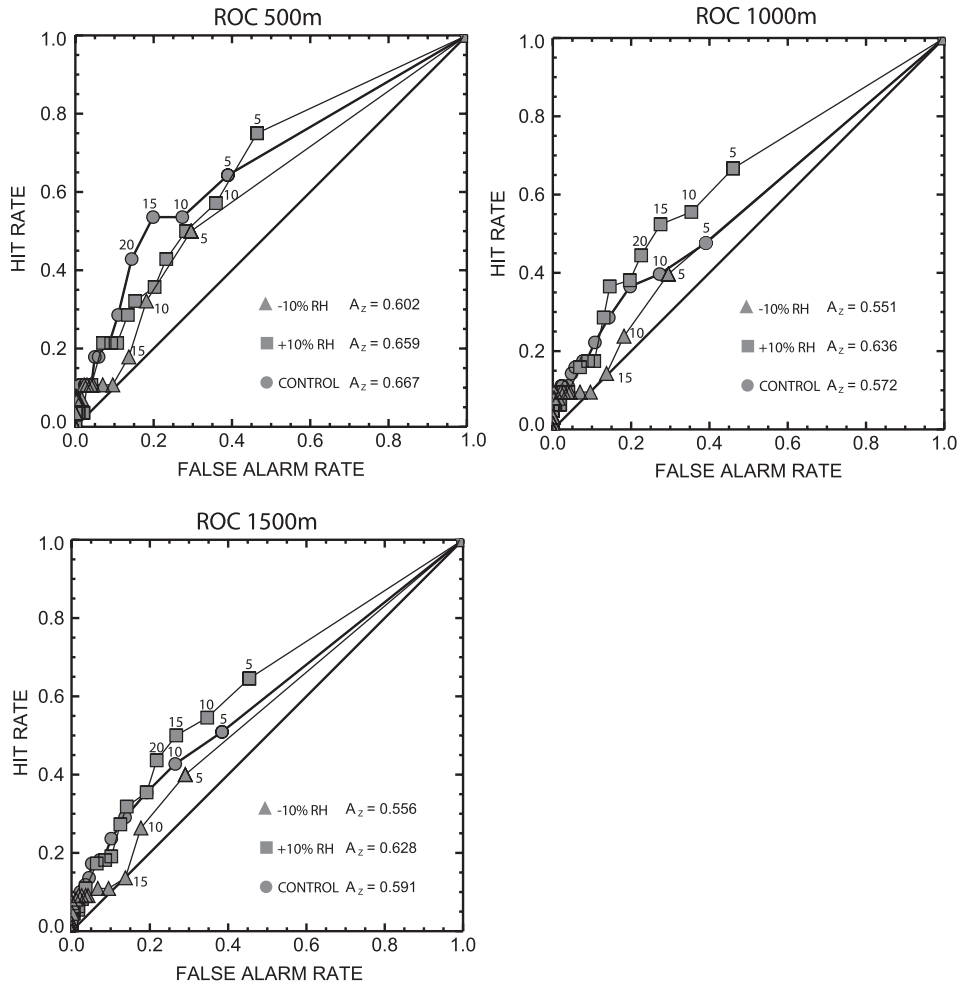


Figure 7

ROC for fog occurrence in the 1-D 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.

### Verification of PAFOG and the Multi-model Ensemble

For maximum comparability PAFOG is verified the same way as COBEL-NOAH. The verification scores are shown in Figure 8 for different initialization times. For comparison purposes the ROC curves of COBEL-NOAH are also shown. Furthermore the ROC of the multi-model ensemble consisting of COBEL-NOAH and PAFOG was computed. From the COBEL-NOAH ensemble only the members

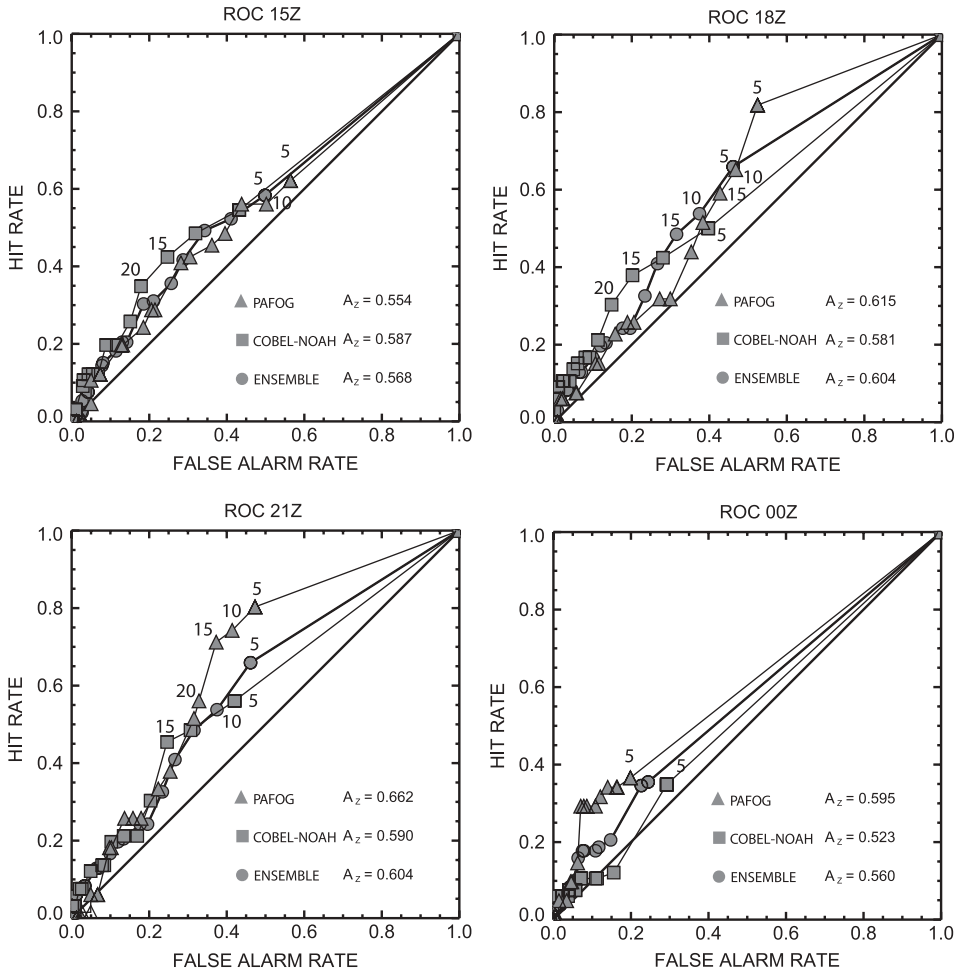


Figure 8

ROC for fog occurrence in the 1-D ensemble for PAFOG, COBEL-NOAH and the multi-model ensemble at different initialization 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 airport.  $A_2$  indicates the area under the ROC.

with advection (ADV) are considered, since they are the most skillful according to Figure 5. If we now look at the skill score  $A_2$  of PAFOG we notice that unlike COBEL-NOAH the skill increases with later initialization time. 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. The ROC reveals that the low forecast probabilities

have the most skill and higher probabilities are not very useful for forecasting. 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.

### *5. Conclusions*

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 a 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. In this study an ensemble member does not just have distinct initial conditions but also distinct boundary conditions to account for the uncertainty in advection tendencies. The variability within the ensemble members should reflect the current atmospheric uncertainty, which can be achieved by computing a separate assimilation for every member using different background estimates from different 3-D models but the same observations. To further increase the ensemble size, since the number of running operational 3-D models is limited, forecast runs valid at the same time but initialized at different times are also included. If such a series of succeeding 3-D forecasts has little variation for the initialization time of the fog model, the atmosphere is likely to be in a state of good predictability. Especially if this behavior is present in models from different operational centers.

With few observations available, variational assimilation must give considerable weight to the background derived from 3-D models. For the humidity assimilation, observational data are scarce and of poor quality, leaving considerable 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.

The same initial and boundary conditions were used by a simple (COBEL-NOAH) and by a detailed microphysics model (PAFOG). The forecast skill of the ensemble system was assessed separately for both ensembles. The COBEL-NOAH ensemble

initialized at 1800 UTC is able to increase the discrimination up to a hit rate of 60% with a false alarm rate of 30% if advection is considered or for the PAFOG ensemble, the 2100 UTC initialization achieves a hit rate of 80% with a false alarm rate of 45%. Overall the benefit of the detailed microphysics is not evident at the moment if the models are run for an entire season rather than well-defined case studies. It has to be noted that these forecasts are purely machine based. If the model results were interpreted by a human forecaster, better skills could be achieved.

The potential of a 1-D model in complex terrain is however limited by the effects resulting from spatial heterogeneity, namely advection of temperature, humidity and wind and the resulting accumulation of cold air in valleys and basins. Furthermore the radiation flux above the model domain has to be supplied to include the effects of higher clouds. If 1-D models are coupled to 3-D models it is possible to obtain estimates of these non-local effects which can be expressed as lateral and upper boundary conditions. As verification results indicate, advection should be derived using neighboring columns of the 3-D model, rather than the total change in a single column. A 1-D fog forecast is very sensitive to the forecast quality of the driving 3-D model but the benefits of the coupling are seen to be significant.

A 1-D model coupled to a 3-D model requires only a fraction of the computing power needed for a full 3-D fog simulation. Therefore it is possible to compute a probabilistic forecast for the next 18 h at a high temporal frequency, e.g., every hour on a simple workstation. Despite the large differences in computational time, preliminary results with the authors 3-D fog model, based on NMM and PAFOG, indicate that the coupled 1-D ensemble seems to be better than a full 3-D forecast in relatively simple terrain, whereas for complex terrain a single 3-D forecast can be better than the 1-D ensemble.

### *Acknowledgements*

The authors want to express their gratitude to MeteoSwiss for providing observational data and aLMo predictions. The comments and suggestions made by anonymous reviewers lead to a significant improvement of the manuscript. This work was funded by the Swiss Federal Office for Education and Science (Grant C03.0024) and COST-722.

### REFERENCES

- BALLARD, S., GOLDING, B., and SMITH, R. (1991), *Mesoscale model experimental forecasts of the Haar of northeast Scotland*, Mon. Wea. Rev. 191, 2107–2123.
- BERGOT, T., CARRER, D., NOILHAN, J., and BOUGEAULT, P. (2005), *Improved site-specific numerical prediction of fog and low clouds*, Weather Forecast 20, 627–646.

- BERGOT, T. and GUÉDALIA, D. (1994a), *Numerical forecasting of radiation fog. part I: Numerical model and sensitivity tests*, *Mon. Wea. Rev.* *122*, 1218–1230.
- BERGOT, T. and GUÉDALIA, D. (1994b), *Numerical forecasting of radiation fog. part II: A comparison of model simulations with several observed fog events*, *Mon. Wea. Rev.* *122*, 1231–1246.
- BERNARDET, L. R., BOGENSCHUTZ, P., SNOOK, J., and LOUGHE, A. (2005), *WRF forecast over the southeast United States: Does a larger domain lead to better results?* In 'Preprints of the 6th WRF/15th MM5 users' workshop', 2.10, National Center for Atmospheric Research, Boulder, Co.
- BOTT, A., SIEVERS, U., and 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. and TRAUTMANN, T. (2002), *PAFOG - A new efficient forecast model of radiation fog and low-level stratiform clouds*, *Atmos. Res.* *64*, 191–203.
- BOTT, A., TRAUTMANN, T., and 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.
- BOUITIER, F. and COURTIER, P. (1999), *Data assimilation concepts and methods*, Technical report, European Center for Medium Range Weather Forecast ECMWF.
- 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. and ROACH, W. T. (1976), *The physics of radiation fog. part II: A numerical study*, *Quart. J. Roy. Meteor. Soc.* *102*, 335–354.
- CHEN, F., JANJIC, Z., and MITCHELL, K. (1997), *Impact of atmospheric surfacelayer parameterization in the new landsurface scheme of the NCEP mesoscale Eta model*, *Boundary-Layer Meteor.* *85*, 391–421.
- DUNLOP, C. and CLARK, P. (1997), *Forcing the single column UM from the mesoscale model*, Technical Report 255, UK MetOffice.
- DUYNKERKE, P. G. (1991), *Radiation fog: A comparison of model simulation with detailed observations*, *Mon. Wea. Rev.* *119*, 324–341.
- EK, M. B., MITCHELL, K. E., LIN, Y., ROGERS, E., GRUNMANN, P., KOREN, V., GAYNO, G., and TARPLEY, J. D. (2003), *Implementation of NOAA land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model*, *J. Geophys. Res.* *108*, doi:10.1029/2002JD003296.
- ESTOURNEL, C. (1988), *Etude de la phase nocturne de la couche limite atmospherique*, These doctorat d'etat 1361, Université Paul Sabatier, Toulouse, France.
- EUROCONTROL (2006), *SESAR project site*, Internet: [www.eurocontrol.int](http://www.eurocontrol.int).
- GOLDING, B. W. (1993), *A study of the influence of terrain on fog development*, *Mon. Wea. Rev.* *121*, 2529–2541.
- GULTEPE, I. and MILBRANDT, J. A. (2007), *Microphysical observations and mesoscale model simulation of a warm fog case during FRAM project*, *Pure Appl. Geophys.* *164*, 6/7, this issue.
- GULTEPE, I., MÜLLER, M. D., and BOYBEYI, Z. (2006), *A new visibility parameterization for warm-fog applications in numerical weather prediction models*, *J. Appl. Meteor. Climat.* *45*(11), 1469–1480.
- HACKER, J. P. and SNYDER, C. (2005), *Ensemble kalman filter assimilation of fixed screen-height observations in a parameterized PBL*, *Mon. Wea. Rev.* *133*, 3260–3275.
- IDE, K., COURTIER, P., GHIL, M., and LORENC, A. C. (1997), *Unified notation for data assimilation: Operational, sequential and variational*, *J. Meteorol. Soc. Japan* *75*(1B), 181–189.
- JANJIC, Z. I. (2003), *A nonhydrostatic model based on a new approach*, *Meteor. Atmos. Phys.* *82*, 271–285.
- JANJIC, Z. I., GERRITY, J. P., and NICKOVIC, S. (2001), *An alternative approach to nonhydrostatic modeling*, *Mon. Wea. Rev.* *129*, 1164–1178.
- KADYGROV, E. N. and PICK, D. R. (1998), *The potential for temperature retrieval from an angular-scanning single channel microwave radiometer and some comparisons with in-situ observations*, *Meteor. Appl.* *5*, 393–404.
- KALNAY, E., *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, Cambridge 2003).
- KESSLER, E. (1969), *On the distribution and continuity of water substance in atmospheric circulations*, *Meteor. Monog.* *10*(32), 84.

- LIN, Y., COLLE, B. A., and NOVAK, D. R. (2005), *Comparison of the real-time MM5 and WRF over northeastern United States*. In 'Preprints of the 6th WRF/15th MM5 users' workshop', 3.5, National Center for Atmospheric Research, Boulder, Co.
- MAHRT, L. and EK, M. (1984), *The influence of atmospheric stability on potential evaporation*, J. Clim. Appl. Meteorol. 23, 222–234.
- MAHRT, L. and PAN, H.-L. (1984), *A two-layer model of soil hydrology*, Bound.-Layer Meteor. 29, 1–20.
- METATMG (2005), *Terms of References of the METATMG*, ICAO METG, Paris.
- MÜLLER, M. D. (2006), *Numerical simulation of fog and radiation in complex terrain*, Ph.d. Thesis, stratus 12, University of Basel.
- MUSSON-GENON, L. (1987), *Numerical simulations of a fog event with a one-dimensional boundary layer model*, Mon. Wea. Rev. 115, 592–607.
- PAN, H.-L. and MAHRT, L. (1987), *Interaction between soil hydrology and boundary-layer development*, Bound.-Layer Meteor. 38, 185–202.
- PARRISH, D. F. and DERBER, J. C. (1992), *The national meteorological center's spectral statistical-interpolation analysis system*, Mon. Wea. Rev. 120, 1747–1763.
- PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., and VETTERLING, W. T., *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, New York, 1998).
- SCHMUTZ, C., SCHMUKI, D., and ROHLING, S. (2004), *Aeronautical climatological information Zürich LSZH*, Arbeitsbericht 201, MeteoSwiss.
- SIEBERT, J., BOTT, A., and ZDUNKOWSKI, W. (1992a), *Influence of a vegetation-soil model on the simulation of radiation fog*, Beitr. Phys. Atmos. 65, 93–106.
- SIEBERT, J., BOTT, A., and ZDUNKOWSKI, W. (1992b), *A one-dimensional simulation of the interaction between land surface processes and the atmosphere*, Boundary - Layer Meteor. 59, 1–34.
- STEPPELER, J., DOMS, G., SCHÄTTLER, U., BITZER, H. W., DAMRATH, A. G., and GREGORIC, G. (2003), *Meso gamma scale forecasts by the nonhydrostatic models Im*, Meteor. Atmos. Phys. 82, 75–96.
- TURTON, J. D. and BROWN, R. (1987), *A comparison of a numerical model of radiation fog with detailed observations*, Quart. J. Roy. Meteor. Soc. 113, 37–54.
- VON GLASOW, R. and BOTT, A. (1999), *Interaction of radiation fog with tall vegetation*, Atmos. Environ. 33, 1333–1346.
- ZDUNKOWSKI, W. and BARR, A. (1972), *A radiative-convective model for the prediction of radiation fog*, Bound.-Layer Meteor. 3, 152–157.
- ZDUNKOWSKI, W. and NIELSEN, B. (1969), *A preliminary prediction analysis of radiation fog*, Pure Appl. Geophys. 19, 45–66.

(Received April 18, 2006, accepted October 25, 2006)

Published Online First: June 8, 2007

---

To access this journal online:  
[www.birkhauser.ch/pageoph](http://www.birkhauser.ch/pageoph)

---