Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks

Francesco Uboldi,a* Cristian Lussanab and Marta Salvatib

a Novate Milanese, Milano, Italy
b ARPA Lombardia, Milano, Italy

ABSTRACT: An objective analysis technique is applied to a local, high-resolution meteorological observation network in the presence of complex topography. The choice of optimal interpolation (OI) makes it possible to implement a standard spatial interpolation algorithm efficiently. At the same time OI constitutes a basis to develop, in perspective, a full multivariate data assimilation scheme. In the absence of a background model field, a simple and effective de-trending procedure is implemented. Three-dimensional correlation functions are used to account for the orographic distribution of observing stations. Minimum-scale correlation parameters are estimated by means of the integral data influence (IDI) field. Hourly analysis fields of temperature and relative humidity are routinely produced at the Regional Weather Service of Lombardia. The analysis maps show significant informational content even in the presence of strong gradients and infrequent meteorological situations. Quantitative evaluation of the analysis fields is performed by systematically computing their cross validation (CV) scores and by estimating the analysis bias. Further developments concern the implementation of an automatic quality control procedure and the improvement of error covariance estimation. Copyright © 2008 Royal Meteorological Society

KEY WORDS objective analysis; data assimilation; local weather services; observation influence; cross validation

Received 5 September 2007; Revised 4 March 2008; Accepted 7 March 2008

1. Introduction

In many areas of the world, departmental environmental agencies or other sub-national institutions manage local meteorological observation networks. These are characterized by high resolution, both in space and in time, and a high degree of automation. In principle, they constitute a rich source of meteorological information, but they present specific problems that have limited their exploitation and that need to be addressed.

One issue is the non-uniform quality of local observational networks which are often managed by organizations whose primary function might not be meteorology. These observations are not, in general, included in the Global Observation System, so observation methods and quality control procedures might not fulfil WMO standards (WMO, 1996). Even if data centres comply with quality control requirements, observations from local networks can be affected by errors of representativity. This can be due to various reasons, including practical constraints (such as funding, observation sites chosen with non-meteorological priorities, administrative boundaries), and the density of the observational network itself. The representativity of observations from local networks is rarely studied, so the consequences on analysed fields of their inhomogeneous spatial distribution are seldom controlled. In fact, most of the objective analysis techniques were developed in the context of Numerical Weather Prediction (NWP) and data assimilation in centres that, until recently, focussed on larger dynamic scales (synoptic and meso-α). With good reason the major NWP centres have generally chosen to discard these data.

Recently, however, hydrostatic global models have reached very high resolutions (about 20 km for the ECMWF model since 2006). Operational non-hydrostatic models nowadays have a resolution that is even higher than that of local meteorological networks (about 2.2 km for the German Lokal Modell in 2007). Even so, and in spite of the achievements of data assimilation methods at larger scales, it is not yet a common practice to merge dense observational networks and local models. Objective difficulties arise from differences between model topography and real topography, from approximations in the parameterization of surface and boundary layer processes, and from the dynamic behaviour of short and fast (meso-γ and convective) scales. A further practical difficulty is that model fields are not always available, with their full resolution and frequency, to the small organizations that manage high resolution meteorological networks.
Any group that manages an observational network may, however, desire to produce analysis fields, even if it does not operationally run a model into which data can be assimilated. A spatial analysis procedure can, in principle, extract from the network’s measurements its full informational content and filter sub-scale noise from observations. It also allows development of robust data quality control procedures. Besides, analysis maps are required for a variety of operational and practical applications, such as weather forecasting, forecast verification, environmental monitoring, land and water management, and fire prevention. On the other hand, the high resolution of non-hydrostatic models itself may require high resolution maps of surface variables, both for a detailed definition of surface parameters, and for model forecast verification. For all these purposes, local high-resolution observational networks may provide a source of useful information, with the condition that the quality of the data is reliably controlled.

This work describes an optimal interpolation (OI)-based analysis scheme, suitable for high-resolution local networks. Reliable analysis maps can be obtained even in the absence of a model-based first guess, by using a background field obtained from the observations themselves through a de-trending procedure. The scheme is applied to temperature and relative humidity observations from the meteorological network of Lombardia, located in the southern side of the Alps (Figure 1). For these variables the analysis method proved to be quite effective. Moreover, the analysis scheme represents a solid step towards implementing a complete multivariate data assimilation scheme, possibly including all observations and model dynamics.

This article is organized as follows. Section 2 describes the interpolation scheme. Section 3 describes the implementation choices made to interpolate surface meteorological observations from Lombardia’s high-resolution network. In Section 4 the method is applied to two very different test cases: a thermal inversion with persistent fog over the Po Plain and a north foehn event with strong thermal gradients. A quantitative evaluation of the interpolation method, with its cross validation (CV) score and a discussion of bias, is presented in Section 5.

2. Interpolation method

2.1. Optimal interpolation

The role of spatial interpolation algorithms in data assimilation is extensively described in various books on the subject (Daley, 1991; Kalnay, 2003). The method used in the present work is an implementation of OI (Gandin, 1963) synthetically described here with its known relation to other methods.

The analysis field of a meteorological variable may be represented by its values at appropriate locations, typically at the nodes of a regular grid that covers an area of interest. These \( I \) values are stored as components of the analysis state vector \( \mathbf{x}^a \). It is assumed that two types of information are available. The first consists of \( M \) observations, the components of the vector \( \mathbf{y}^o \). The second is the background (or first-guess) field, available at grid points, \( \mathbf{x}^b \), and at the observation locations (stations) \( \mathbf{y}^b \). The possibility that an independent background field is not available is considered in Section 3.1, where a de-trending technique is presented.

The OI analysis is obtained by a linear relation between the analysis increment \( \mathbf{x}^a - \mathbf{x}^b \) and the innovation \( \mathbf{y}^o - \mathbf{y}^b \):

\[
\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{y}^b)
\]

where \( \mathbf{K} = \mathbf{G}(\mathbf{S} + \mathbf{R})^{-1} \) is expressed by means of the error covariance matrices

\[
\mathbf{G} = \left\langle (\mathbf{x}^o - \mathbf{x}^\prime) (\mathbf{y}^b - \mathbf{y}^\prime)^T \right\rangle
\]

\[
\mathbf{S} = \left\langle (\mathbf{y}^b - \mathbf{y}^\prime)(\mathbf{y}^o - \mathbf{y}^\prime)^T \right\rangle
\]

\[
\mathbf{R} = \left\langle (\mathbf{y}^o - \mathbf{y}^\prime)(\mathbf{y}^o - \mathbf{y}^\prime)^T \right\rangle
\]

where \( \mathbf{x}^\prime \) is the unknown ‘true’ state vector (so that \( \mathbf{x}^a - \mathbf{x}^\prime \) is the analysis error vector) and the angular brackets \( \langle \rangle \) represent the expectation value with respect to an appropriately defined statistical ensemble. Each component of the \( (I, M) \) matrix \( \mathbf{G} \) is the covariance between the background error at a grid point and the background error at a station point; each component of the \( (M, M) \) matrix \( \mathbf{S} \) is the background error covariance between a couple of station points; and the \( (M, M) \) matrix \( \mathbf{R} \) is the observation error covariance matrix.

The OI analysis on station points is

\[
\mathbf{y}^a = \mathbf{y}^b + \mathbf{S}(\mathbf{S} + \mathbf{R})^{-1}(\mathbf{y}^o - \mathbf{y}^\prime)
\]

where \( \mathbf{W} = \mathbf{S}(\mathbf{S} + \mathbf{R})^{-1} \) is the influence matrix, which linearly relates the station analysis increment to the innovation. It is worth remarking that, just as the ‘true’ state is unknown, so are all the covariance matrices: \( \mathbf{G}, \mathbf{S}, \) and \( \mathbf{R} \). The estimates of these matrices determine the characteristics of the analysis field.

For the purpose of spatial interpolation, the components of the matrices \( \mathbf{G} \) and \( \mathbf{S} \) are specified by means of analytical correlation functions. The parameters of the correlation function may be chosen to fit the available statistics on the field variabilities and errors. In the literature and in the interpolation practice, many different types of functions are used to model the error correlations, mostly combinations of exponentials, Gaussian and polynomials. In the applications presented here, Gaussian functions are used (Section 3.2).

For point-wise observations taken by different instruments, it is a widely used approximation to assume that
R is a diagonal matrix, i.e. observational errors at different station locations are uncorrelated. Observed values, besides the instrumental error, may be affected by representativity error, due to small scales and local phenomena that cannot be represented by the state vector. This means, on the one hand, that if these phenomena are of interest for the final user, the chosen representation of the field is inadequate. On the other hand, if the chosen representation of the field is considered adequate, then the small scales have no interest: they should be then considered as noise and filtered out. In this case, it is appropriate to include an estimate of the representativity error covariance in the observation error covariance \( R \).

2.2. Optimal interpolation and other methods

In this paragraph, the interpolation algorithm is presented briefly in the more general context of data assimilation methods. In this way, a univariate spatial interpolation algorithm may, in perspective, become a basic component of an advanced multivariate data assimilation scheme, able to account for physical and dynamical constraints. Moreover, even for the limited purpose of spatial interpolation, the advanced development of data assimilation methods can be exploited to obtain robust tools for the implementation of objective analysis techniques and the quantitative evaluation of the produced fields.

The formal equivalence between OI and other interpolation algorithms such as kriging and smoothing splines (Wahba and Wendelberger, 1980; Steinackert et al., 2000) is well known (Lorenc, 1986; Daley, 1991), as is the convergence to OI of iterative methods such as successive corrections (SCs) (Bratseth, 1986; Daley, 1991; Uboldi et al., 2000). In order to discuss the role of OI in data assimilation further, the concept of observation operator needs to be introduced.

Analysis and background fields \( x^b \) and \( x^b \) are model states. An observation operator is a function \( H \) which, starting from the background model state \( x^b \), computes the observational estimates \( y^b = H(x^b) \) that can be directly compared with the observations \( y^o \). Generally speaking, the \( I \) state variables and the \( M \) observations may be very different. They are referred to different spatial locations; they may have different physical dimensions; and may be referred to different times in the case of four-dimensional algorithms. As a consequence, the observation operator \( H \) may be very complicated. It may be a non-linear function and it may involve integrals, thus permitting the assimilation of remote sensing observations. The concept of observation operator is fully exploited in variational algorithms, and it is explicitly employed in Kalman-filter-based sequential algorithms. When explicit use is made of an observation operator \( H \), the OI analysis becomes

\[
\begin{align*}
\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}H^T(H\mathbf{B}H^T + \mathbf{R})^{-1}(\mathbf{y}^o - H(\mathbf{x}^b))
\end{align*}
\]

where \( H \) is the Jacobian matrix of the vector function \( H \), i.e. the matrix containing its first derivatives with respect to the state components: \( \mathbf{H}_{m,i} = \partial H_m / \partial x_i \). Equation (5) also represents a linearized (or incremental) 3D-Var analysis (Daley, 1991; Kalnay, 2003). The matrix \( \mathbf{B} \) is the background error covariance matrix, defined in a way analogous to Equation (3):

\[
\mathbf{B} = \left( (\mathbf{x}^b - \mathbf{x}') (\mathbf{x}^b - \mathbf{x}')^T \right)
\]

The components of the \((I, I)\) symmetric matrix \( \mathbf{B} \) are defined between couples of state variables. By comparing Equation (5) with Equation (2), it can be seen that \( \mathbf{H} \) is used to transform the background error covariances:

\[
\mathbf{G} = \mathbf{B}H^T \quad (7)
\]

\[
\mathbf{S} = \mathbf{H}\mathbf{B}\mathbf{H}^T \quad (8)
\]

As a consequence, the relation between the influence matrix \( \mathbf{W} \) and the gain matrix \( \mathbf{K} \) is

\[
\mathbf{W} = \mathbf{H}\mathbf{K} \quad (9)
\]

If the concept of observation operator is extended to include a model integration, Equation (5) may also represent a linearized (or incremental) 4D-Var, assimilating data that are distributed in space and in time (Uboldi and Kamachi, 2000; Bennett, 2002; Kalnay, 2003). The analysis step in a Kalman filter also has the same form as Equation (5), with the important difference that the stationary matrix \( \mathbf{B} \) used here is replaced with the dynamically evolving forecast error covariance matrix \( \mathbf{P}' \). The representativity error may be formally interpreted as the observation operator error.

2.3. Integral data influence and cross validation

This section introduces some useful diagnostic and validation tools that are used in Section 3.2 and in Section 5: the integral data influence (IDI) field and the CV score.

Once the covariance matrices are chosen, the IDI field is defined as the analysis field obtained when all observed values are set to 1 and all background values are set to 0. By doing this in Equation (2) and in Equation (4), the IDI field can be defined both on station points, \( y^{\text{IDM}} \), and on grid points, \( x^{\text{IDM}} \) (Figure 3). The IDI field can be formally defined from the influence matrix \( \mathbf{W} \) (Appendix A), but it has a straightforward interpretation. The observational information is that the field is uniform and it has a value 1, while the background information is that the field is uniform but it has a value 0. In station dense areas, the analysis field is almost uniform and close to 1. Since the correlation function approaches zero for increasing distance, far from all stations the analysis is dominated by the background field, resulting in an approximately uniform zero field. In intermediate areas the analysed value is between 0 and 1, but gradients are present. Values larger than 1 are possible and they are due to strong non-uniformities in the data distribution. The IDI field may be used to study the effects on the analysis of removing
or adding stations; hence, it represents a useful tool for optimizing an observational network, as long as the covariance estimates are considered significant.

The CV analysis, \( \tilde{y}_m^a \), is defined as the analysis estimate obtained for the \( m \)th observation by using all other observations, but without using the \( m \)th observation itself. On the basis of the CV analysis, the CV score (Wahba and Wendelberger, 1980; Daley, 1991) is defined as

\[
CV_{\text{score}} = \sqrt{\frac{1}{M} \sum_{m=1}^{M} (\tilde{y}_m^a - y_m^o)^2} \tag{10}
\]

The CV score represents an estimate of the analysis error based on the idea that each observation is used as an independent verification of the analysis field. The estimate is conservative because of the implicit degradation of the local resolution of the observational network. The CV analysis \( \tilde{y}_m^a \) is also useful for quality control tests on the \( m \)th observation.

CV and data influence can be combined by defining the CV-IDI parameter \( y_{m}^{\text{IDI}} \) in a way analogous to the CV analysis, so that the difference between isolated stations and densely observed areas is enhanced. In fact, for a completely isolated station (i.e. having zero correlation with all the other stations), \( y_m^{\text{IDI}} = 0 \) and \( y_m^{\text{IDI}} = \sigma_{b,m}^2 / (\sigma_{b,m}^2 + \sigma_{o,m}^2) \). Conversely, \( y_m^{\text{IDI}} = 1 \) implies \( y_m^o = 1 \). Details of the relations between IDI, the CV analysis and other diagnostic tools are discussed in Appendix A.

The direct knowledge of the components of the inverse matrix \((S + R)^{-1}\) and, in particular, of its diagonal elements allows efficient computation of the CV score. A data quality control procedure (work currently in progress) can also make use of the CV analysis \( \tilde{y}_m^a \) at each station. For these reasons a technique, described in Appendix B, has been implemented for updating the inverse matrix \((S + R)^{-1}\) when new data become available or some data are occasionally missing, replacing a previous version of the interpolation procedure, in which an iterative method was used at each analysis time to solve the linear system.

3. Implementation of the analysis method for temperature and relative humidity fields

The OI scheme described above is applied to hourly observations of surface temperature and relative humidity collected by Lombardia’s local meteorological network of automatic weather stations.

The observational network is managed by the Regional Weather Service, operational since 2004, and has been created by merging previously existing networks, not all of them designed for meteorological applications (main components are air quality and hydrology networks). About 270 stations measure temperature and precipitation, and about half of them also measure other parameters (pressure, relative humidity, wind speed and direction, solar radiation). This seemingly high density of the network, though, may be misleading. The station distribution is not homogeneous, with several sites located in urban areas. Non-standardized observation sites often lead to low-quality observations (i.e. high representativity error). Moreover, the analysis area is on the southern slope of the Alpine range and it is characterized by very complex terrain; hence the vertical distribution must also be considered: the elevation of observation sites ranges from 0 to 3000 m above mean sea level. Figure 1 shows station sites and Lombardia’s orography.

The general characteristics of the observation network and the availability of an appropriate first-guess field strongly influence the choices that must be made in implementing the interpolation scheme, namely, the background \( \chi^b \) and \( y^b \) and the error correlation matrices \( G, S, R \).

3.1. Pseudo-background field built from the observations (de-trending)

The use of a background or first guess field is one of the strong points of OI schemes, which in general aim to the optimal merging (in a statistical sense) of information from different sources. A model field is often used for this purpose, because it represents a source of information that is consistent with the atmospheric conditions and/or it is often used as a control field. In this work, we decided to use Lombardia’s observed temperature and relative humidity fields as a pseudo-background field. The reason is that not only the geographical location of Lombardia, longitude 8.5 to 11.5°E, latitude 44.6 to 45.7°N, is available, but also the general characteristics of the observation network and the availability of an appropriate first-guess field strongly influence the choices that must be made in implementing the interpolation scheme, namely, the background \( \chi^b \) and \( y^b \) and the error correlation matrices \( G, S, R \).
dynamics. This work presents a different possibility that can be used if a model-based analysis is undesired (as it may be in verification studies) or impossible (if model fields are unavailable). Besides, an analysis scheme which only makes use of observations can provide insight on the network characteristics. This knowledge can then be profitably used in the integration of observations with other sources of information.

The choice presented here is to build, from the data themselves, a ‘pseudo’ background field, meant to represent the main spatial trends present in the field and detected by the observations. The background field is calculated as a linear function of the spatial coordinates, by performing a least-square minimization at each observation time. The scheme allows for a change in the vertical derivative; this has a particular relevance for the case of thermal inversions that frequently occur on the Po Plain.

The vertical dependence of an atmospheric field resulting from measurements taken at observing stations located at the surface cannot be considered, in general, an information on the vertical structure of the free atmosphere. On the other hand, when an important thermal inversion is present over the area it does have an influence on observations taken near the surface, particularly on those taken at stations located on the orography facing a wide valley such as the Po Plain. This can easily be seen in Figure 2, showing the temperature observations and the background estimates on station points as a function of station elevation above sea level (on the y axis) for the two application cases discussed in Section 4. In both cases a thermal inversion (of different nature) is evident. At higher elevations, the vertical dependence is very clear in Figure 2(b) (north foehn case), whereas it appears more dispersed in Figure 2(a) (persistent fog case).

If high-resolution topographic information is available for the area, the de-trending scheme allows the choice of any grid size which can be customized to the intended use of the analysis. The fields presented here are computed on a 174 × 177 grid, with 1.5 km resolution. The grid point elevation has been regularly sampled from a 250 m digital elevation model without any smoothing. The choice is made to minimize the difference between grid point and station elevation, while, at the same time, retaining a reasonable grid size in terms of computational time. As a consequence, the grid orography is very discontinuous. For the present application of the analysis maps, which is purely diagnostic, the choice appears justified. If the interpolated fields are needed as input in numerical integrations then different choices must be made.

It is worth remarking that the grid resolution does not correspond to the analysis scale resolution, which is determined by the network density distribution, by the properties of the analysed variable, and by the choices made in specifying the error covariance matrices.

3.2. High resolution topography and three-dimensional correlations

It is assumed that the observation error covariance matrix \( R \) is diagonal and all observation errors have the same variance \( \sigma_o^2 \):

\[
R = \sigma_o^2 I
\]  
(11)

where \( I \) is the identity matrix. A diagonal matrix implies the assumption, reasonable for point-wise measurements, that different observation sites have independent errors. On the other hand, assuming the same error variance for all observations is quite a drastic simplification: a more realistic representation of observational error is currently under study.

The function of horizontal and vertical distances that is used to estimate the background error correlation is

\[
\gamma(d, \Delta z) = e^{-\frac{1}{2} \left( \frac{d}{D_h^2} \right)^2} e^{-\frac{1}{2} \left( \frac{\Delta z}{D_z^2} \right)^2}
\]  
(12)

where \( d \) is the horizontal distance between the two points, and \( \Delta z \) is the difference between their elevations above sea level. \( D_h \) and \( D_z \) are the de-correlation distances in the horizontal direction and vertical direction, respectively. If the background error variance \( \sigma_b^2 \) is assumed to

![Figure 2](https://www.interscience.wiley.com/ma)}
be uniform, then the background error correlation matrices are defined as $\tilde{G} \equiv \sigma_\theta^2 G$ and $\tilde{S} \equiv \sigma_\epsilon^2 S$. The analysis is then obtained from Equation (2) as

$$\tilde{x} = \tilde{x}^b + \tilde{G} \left( \tilde{S} + \varepsilon^2 \mathbf{I} \right)^{-1} (y^o - y^b) \quad (13)$$

where the scalar $\varepsilon^2 = \sigma_\epsilon^2 / \sigma_\theta^2$ is the ratio between the background and the observation error variances. In this way the components of the gain matrix, $\mathbf{K} = \tilde{G} \left( \tilde{S} + \varepsilon^2 \mathbf{I} \right)^{-1}$, only depend on the three parameters $D_h$, $D_z$, and $\varepsilon^2$. From the definition of $\varepsilon^2$, it is clear that assuming $\varepsilon^2 = 0$ implies assuming perfect observations, hence exact interpolation. On the other hand, setting $\varepsilon^2 > 1$ implies a greater confidence in the background field rather than in the observations.

The OI algorithm may also be seen as a numerical low-pass filter, whose cut-off wavelength depends on the parameters $D_h$, $D_z$, and $\varepsilon^2$. This is discussed in more detail, in an ideal case, in Appendix C.

The values of the scale parameters of the correlation functions, $D_h$ and $D_z$, must be set large enough to filter out short scales that cannot be resolved by the observational network. Intuitively, $D_h$ and $D_z$ cannot be chosen too small with respect to the typical distances that characterize the data distribution. Above this lower bound, the values of $D_h$ and $D_z$ can be chosen in order to retain the relevant scales for the meteorological variable under consideration.

The non-dimensional IDI field, defined in Section 2.3 and shown in Figure 3, is very useful in choosing the minimum values for these scale parameters. In fact, given a set of parameter values $D_h$, $D_z$, and $\varepsilon^2$, the IDI field only depends on the data distribution and may change in time only because of missing data.

The values $D_h = 20$ km, $D_z = 500$ m, and $\varepsilon^2 = 0.5$ have been chosen for the temperature field so that the IDI field appears rather uniform and above 0.9 in the Plain and it is only marginally noisy in the mountain areas. $D_h$ and $D_z$ have been chosen by progressively increasing their values until an acceptable IDI field is obtained. Given the filtering properties of the interpolation, described in Appendix C for a continuous data distribution, an homogeneous IDI field can be interpreted as an indication that the non-uniformity existing in the data distribution only marginally affects the filter. The choice of larger $D_h$ and $D_z$ values would essentially increase the cut-off wavelength and would filter out scales that can actually be detected by the observational network (though with some deformation).

Figure 4 shows how a decrease in these values corresponds to a drastic increase in non-uniformity, while increasing the scale lengths by the same amount leads to smaller differences in the IDI field uniformity.

The trade-off realized for the value of $\varepsilon^2$ is also shown in Figure 4. Smaller $\varepsilon^2$ values imply more weight to the observations, and correspond to a larger IDI, but with maxima (above 1) and gradients. Larger $\varepsilon^2$ values imply more weight to the background field, smaller IDI values, and a more uniform field.

The analysis of the IDI field is intended to estimate a lower bound, imposed by the station spatial distribution, for scales that can be detected and resolved by the observational network. Above these minimum scales, however, it is possible to optimize the values of $D_h$, $D_z$, and $\varepsilon^2$ by means of statistical parameter estimation methods (Dee et al., 1999; Dee and Da Silva, 1999; Tarantola, 2005). These methods, by making use of the actual observations over the period chosen to calculate expectation values, allow optimization of correlation scales with respect to spatial scales typically present in the real atmospheric fields. Work on statistical parameter estimation is under way for temperature and it is planned for other variables. Preliminary results (not shown) seem to confirm the values chosen for temperature by means of the IDI field, with a remarkable variability for $\varepsilon^2$ that may perhaps be attributed to the representativity component of observational error.

### 3.3. Relative humidity

The procedure used for relative humidity analysis is as follows. Since a temperature observation is always available in correspondence to each relative humidity observation (Figure 1), the ‘observed’ dew-point temperature can always be obtained. These dew-point temperature observations are homogeneous to temperature and can be similarly treated. In particular, sharp vertical derivative changes also occur for dew-point temperature profiles in the free atmosphere, affecting surface measurements in orographic areas. A background field is then calculated from dew-point temperature observations, using the procedure described for temperature in Section 3.1. The subsequent interpolation is carried out for dew-point...
Figure 4. Temperature observations on 20 January 2006, 1200 h (1100 UTC). IDI fields (non-dimensional) obtained by varying the parameters with respect to the chosen values $D_h = 20 \text{ km}$, $D_z = 500 \text{ m}$, $\varepsilon^2 = 0.5$ (grey scale as in Figure 3): (a) $D_h = 15 \text{ km}$; (b) $D_h = 25 \text{ km}$; (c) $\varepsilon^2 = 0.3$; (d) $\varepsilon^2 = 0.7$; (e) $D_z = 400 \text{ m}$; (f) $D_z = 600 \text{ m}$. All fields are masked out below the value 0.5.

temperature, with error covariance matrices estimated as described in Section 3.2. In particular, the parameter values are determined by making use of the OI filtering properties and of the IDI field: the resulting values are $D_h = 30 \text{ km}$, $D_z = 600 \text{ m}$, and $\varepsilon^2 = 0.5$. These values are larger than those found for temperature because the relative humidity observation distribution is coarser, particularly in the Alpine area. A saturation check is performed both on the background and on the analysis field, by using the known temperature analysis and allowing for over-saturation up to a maximum relative humidity of 103%. The dew-point temperature and temperature...
analysis fields are then used to compute a relative humidity value on each grid point. Some small-scale details, unresolved by the hygrometer network, may appear in relative humidity maps thanks to the higher resolution of temperature observations.

3.4. Discussion

The interpolation scheme is applied to hourly-averaged observations, as all stations in the network record and report at least at hourly intervals. For temperature it has been used as an operational tool, monitored daily, since January 2006. Relative humidity analysis maps have been produced operationally since February 2007. The interpolation scheme has proved to be robust and sensitive to details and has correctly described major mesoscale features of temperature and relative humidity fields in the area.

There are known limits in the choices made, the main being the unavailability of an independent background field. The background field built from the data is however effective in estimating the main trends present over the area. Considering the relatively small size of the area under consideration, these may be approximately attributed to the synoptic and meso-α scales. A better representation of the larger scales, such as that present in a model field for example, would certainly improve the quality of the analysis. However, as long as a model field is not available, imposing a more complicated spatial dependence on the background field would be rather arbitrary. It must be stressed, though, that such a simple linear dependence on spatial coordinates would be inadequate, even for de-trending purposes, in the case of a larger area.

When a pseudo-background field is built from the data, the assumption that background and observational error are independent may be put into question. However the larger scales present in the background field and those (shorter) resolved in the analysis field may be considered uncorrelated. Parrish and Derber (1992) used a covariance matrix diagonal in the spectral space, thus assuming de-correlation between different scales; they obtained a non-diagonal background error covariance matrix after transformation to the grid point space.

4. Application examples in strong gradient cases

This section shows two examples of temperature and relative humidity analysis obtained through the OI scheme in different synoptic situations. The cases discussed here represent a good test for the scheme for the presence of strong temperature and humidity gradients across the area. Resolving gradients and fronts poses a challenge for all analysis techniques and, in fact, also for observational networks. In both cases the phenomena could be investigated from independent information (satellite images, soundings, synoptic analysis, SYNOP reports, surface observation network).

The quantitative diagnostic tools available to estimate analysis error variance and bias and the results of their application to temperature analysis are presented in Section 5.

4.1. Strong ground based temperature inversion

Figure 5 refers to a winter case of high pressure subsidence causing persistent fog on the Po Plain, 20 January 2006. The fog, of radiative origin, is associated with a marked ground based inversion, apparent also from the 0600 UTC Linate radio sounding and corroborated by SYNOP reports (not shown). Figure 5(a) shows the temperature analysis field at 1200 h (1100 UTC, local time is UTC + 1 h). An intense temperature gradient is present in the plain area. Figure 5(c) shows the relative humidity analysis field at the same time. Figure 5(b) the corresponding METEOSAT satellite image in the HRV (High Resolution Visible) channel is displayed: this is an informational source that is completely independent, and shows the presence of dense fog or low-level stratus clouds in the plain area. The correspondence between the cloudy area of Figure 5(b) and the cold, humid area defined by the strong temperature gradient in Figure 5(a) and by the relative humidity field in Figure 5(c) is quite remarkable. The interpolation scheme takes advantage,
in this case, of the high observation density in the plain, particularly for relative humidity. Moreover, the pseudo-background field, allowing for a vertical derivative change, is particularly well suited to describe situations where a main thermal inversion is present.

The CV score for the temperature analysis field shown in Figure 5 is 2.43 °C. This value, higher than the average value of 1.49 °C (the CV score distribution is shown in Figure 8), is determined by contributions from stations located in the Alpine area. In fact, if the CV score calculation is restricted to the plain and to the orographic area immediately facing the plain (thus including the gradient and the warm belt appearing in Figure 5(a), a value very close to the average is obtained, 1.61 °C. On the one hand, as it is apparent in Figure 2(a), the observed temperature values appear quite dispersed for elevations higher than 500 m: the background field, in this case, can only be considered a rough approximation for the temperature field in the Alpine region. On the other hand, many observations are missing, almost all relative to stations located in the mountain area: only 206 stations provided useful measurements, while, for comparison, 242 observations are used for the map of Figure 6 (274 temperature stations are nominally active, August 2007). This lack of data limits the correction that the interpolation can perform to the background field. In conclusion, the temperature analysis field is quite satisfactory in the Po Plain and its immediate surroundings, but it cannot be considered very accurate in the Alpine areas in this particular case. Presumably, the same is also true for relative humidity, due to the smaller number of humidity observations and to the dependence of the relative humidity field on the temperature analysis.

4.2. North foehn

Figures 6 and 7 refer to a completely different meteorological situation. In this case, 12 March 2006, the western part of Lombardia is under the influence of intense northerly winds (caused by a meridional intensification of the jet stream) which assumed the character of (north) foehn during the last part of the night and in the early morning. Heating due to adiabatic compression, a relatively frequent phenomenon in the mountain valleys in the north-western part of Lombardia (Valchiavenna), was also detected in some western plain areas in this case. In the eastern part of the analysis domain, the high-level circulation assumed a cyclonic character. Very low temperature values were observed in the mountain areas to the northeast and light snowfall occurred during the morning in the south-eastern plain.

The temperature analysis field influenced by all these phenomena is shown in Figure 6. The symbols on the map mark the locations of the four observing stations whose measurements are plotted versus time in Figure 7. In the time plots of Figure 7(a)–(c), it is easy to see the foehn onset, marked by the sudden increase in wind and temperature and the strong decrease in relative humidity.

In Figure 6, mild temperatures (up to +9 °C) are evident in the western areas affected by the subsidence warming (foehn): Valchiavenna, where the symbol marks Samolaco, Figure 7(a); the λ-shaped Como Lake; the western plain, where symbols mark Minoprio and Castello, respectively in Figure 7(b) and (c). It is worth noting that in Castello’s time plot, Figure 7(c), foehn onsets about 2 h after Figure 6’s analysis time. In fact, in the map, a cold area still surrounds the observing station: the progressive foehn expansion can be clearly seen through the sequence of hourly analysis maps (not shown).

At the same time, the map of Figure 6 shows very low temperatures (below −15 °C) in the north-eastern mountains (Alta Valtellina) and a cold area (about 0 °C) in the south-eastern plain, where a symbol marks Mantova, Figure 7(d). This last time plot shows very different weather compared to the other stations of the same figure: weak winds and high relative humidity almost constant during the whole morning.

This complex situation puts the approximations used, in particular those present in the background field, close to their limits. The field shown in Figure 6 appears quite acceptable, however, for diagnostic purposes. Its CV score is 1.68 °C, a value near the average.

5. Diagnostics

The hourly maps produced by the algorithm appear detailed and consistent with independent meteorological information. However, to understand the characteristics
of the interpolation scheme, a quantitative analysis has been carried out for temperature using CV and bias estimates.

5.1. Cross validation score

The CV score, defined in Section 2.3, represents an estimate of the analysis root-mean-square error based on the idea that each observation is used as an independent verification of the analysis field. The CV score is routinely computed for temperature analysis fields using Equation (10). The values obtained for the application cases have been presented and discussed in Section 4. A global CV score was computed for temperature analysis fields by averaging the individual hourly values over the whole year 2006, giving the mean value $CV_{\text{score}} = 1.49^\circ$C. The 2006 hourly CV score distribution is shown in Figure 8.

A systematic check of cross validation residuals $|\hat{y}_m - y_m|\ (^\circ\text{C})$ for each observation site was carried out. The box-plot of Figure 9 shows the distribution of the CV residuals (calculated in absolute value for each station) against the corresponding CV-IDI value $\hat{y}_{\text{IDI}}$ (Section 2.3). This parameter measures the data influence due to the other stations: isolated stations have $\hat{y}_{\text{IDI}} \approx 0$, while stations located in densely observed areas have $\hat{y}_{\text{IDI}} \approx 1$. In other words, the CV-IDI may be interpreted as a measure of observational coverage (dependent on the correlation estimates). It can then be seen that the probability of large errors, represented by the upper box height, decreases with increasing observational coverage. The median of the CV residuals is below 2 $^\circ$C for all areas. For data-dense areas, characterized by $\hat{y}_{\text{IDI}} > 0.85$, the median is smaller, about 1 $^\circ$C; it is however larger than zero because local scales affecting the observations are

Figure 7. Foehn and snow case, 12 March 2006. Time plot of temperature (dashed line), relative humidity (thick solid line), and wind speed (thin solid line) observed by the four stations marked in Figure 6. Left $y$-axis: temperature ($^\circ$C) and wind speed (ms$^{-1}$); right $y$-axis: relative humidity (%). Time on the $x$-axis is local time (UTC + 1). In all graphs the vertical line indicates the time corresponding to the map of Figure 6.

This figure is available in colour online at www.interscience.wiley.com/ma

Figure 8. Distribution of hourly temperature analysis CV scores for the year 2006. The mean is $CV_{\text{score}} = 1.49^\circ$C.

Figure 9. Box-plot of absolute values of CV residuals, $|\hat{y}_m - y_m|\ (^\circ\text{C})$, versus CV-IDI, $\hat{y}_{\text{IDI}}$ (non-dimensional), for hourly 2006 temperature analysis.
filtered out by the interpolation. It is worth remarking that in data-sparse areas, corresponding to small CV-IDI values \( \hat{\text{IDI}}_m < 0.45 \), the analysis is determined by the background field. In this case the analysis error can be interpreted as background error: its rather low average values support the assumptions of Section 3.1.

5.2. Analysis bias

OI provides an unbiased analysis under the hypothesis that both the observations and the background field are unbiased estimates of the atmospheric state. In practice, both background and observations can be biased (observation bias is mostly due to representativity error). The resulting analysis bias should then be estimated.

On station points, the (station dependent) background bias \( \langle \varepsilon^b \rangle \) and the observational/representativity bias \( \langle \varepsilon^o \rangle \) appear as components of the average innovation

\[
\langle d \rangle = \langle y^o - y^b \rangle = \langle \varepsilon^o \rangle - \langle \varepsilon^b \rangle
\]  

The distribution of the \( \langle d \rangle \) components evaluated on the whole 2006 dataset is practically normal (Figure 10), with mean 0.02 °C and standard deviation 1.72 °C. For most stations the average innovation is almost zero, but in some cases a background bias or representativity bias is present. A closer inspection shows features that have intuitive explanations. Systematically negative innovations are mainly located in mountain valleys, because the background estimate on stations at low elevation in valley floors is influenced by the warmer and more numerous stations in the Po Plain. On the other hand, most of the stations with positive \( \langle d \rangle \) are located in large urban areas, subject to the urban heat island effect.

Warm anomalies in urban areas and cold anomalies in mountain valleys are also visible in Figure 11, that shows the analysis increment averaged over the year 2006. This field can be written as

\[
\langle x^a - x^b \rangle = K (\langle \varepsilon^o \rangle - \langle \varepsilon^b \rangle)
\]  

If the observations and the background were both actually unbiased, this average field should be practically zero: the non-zero features can be attributed to the presence of a background or a representativity bias, reduced or amplified by the interpolation (the effect of the gain matrix \( K \)). For example, the extension of the warm anomaly observed by stations located in the urban area of Pavia (square marker in Figure 11) is amplified and extended in the south-eastern direction to a large data-void area. Such an effect is typical of isotropic correlation functions in the presence of inhomogeneous data distribution.

The urban heat island effect is presently neglected both in the background field and in the covariance estimates. In the case of large urban areas, these warm anomalies should be considered as background errors, and their correction could be achieved by including land use information in the covariance model. On the other hand, bias effects due to smaller scale features should be considered as representativity errors and filtered out. This could be obtained by means of a station-dependent \( \varepsilon^2 \).

Existing techniques to estimate and correct the background bias in data assimilation (Dee and Da Silva, 1998; Dee and Todling, 2000) will be taken into consideration in future work.

6. Conclusions

This document presents a spatial interpolation scheme based on OI and suitable for observations from high-resolution local networks. Despite the presence of important representativity errors and the complex topography of
Lombardia, the observations do provide useful mesoscale information, that is correctly recovered by the interpolation.

The resulting temperature and relative humidity analysis maps are presented in two example cases, chosen for the presence of strong gradients that stress the approximations used in the implementation. Independent observations show that the analysis maps correctly describe relevant meteorological phenomena. The accuracy of temperature analysis fields is quantitatively evaluated through the distribution of hourly CV scores over the whole year 2006 resulting in rather low values, always below 3 °C and with average of about 1.5 °C. The presence of bias in the analysis fields is evaluated by computing the average analysis increment, and it is discussed with reference to representativity bias and background bias.

The presented scheme has been running daily for over a year, producing hourly maps of temperature and relative humidity. The analysis maps, quite satisfactory in all weather situations, are used for the daily activity of weather analysis and forecasting at the Regional Weather Service of Lombardia. The quality of the hourly maps is generally analogous to that of the two examples shown (in fact it is often higher, as the CV scores of the example fields are larger than the average).

A better knowledge on reliability and errors of the observational network also resulted from the studies carried out to implement the interpolation scheme. An automatic quality control procedure based on the analysis of CV residuals (Lorenc, 1984; Gandin, 1988) is currently under development and will be the subject of a separate communication. By means of the spatial consistency check, it is possible to automatically detect and discard rough errors that inevitably affect the observations. The IDI field, presented and used here to tune the correlation coefficients, also represents a useful tool for network planning and management: its changes allow quick evaluation of the effects of adding, moving, or removing observation sites.

Some limitations are necessarily present in any interpolation procedure that does not account for the atmosphere dynamics. Even so, the simplifications employed in the implementation of the interpolation scheme very seldom resulted in unrealistic features, always very local and with small amplitude. Qualitative and quantitative diagnostics indicate that, among the assumptions used in the implementation, the least realistic appear to be the use of isotropic correlation functions and the hypothesis of unbiased background and observations. Consequently, work is at present devoted to:

- estimating the background-to-observation error ratio \( e^2 \) for each station, including local effects as representativity error;
- refining the covariance estimates, with the aim of explicitly accounting for local terrain features (urban heat islands, orientation of orographic slopes, and proximity to surface water masses) to attenuate the correlation between differently characterized sites;
- studying the feasibility of a bias estimation and correction procedure (Dee and Da Silva, 1998) even though in the absence of a model background field, corrections in data sparse areas may remain difficult;
- extending the quantitative CV and bias evaluation to relative humidity.

Improvements to the analysis scheme are regularly checked by evaluating CV scores and by studying innovation and analysis increment statistics.

The realization of a full four-dimensional data assimilation system, including model dynamics and remote sensing information, may be beyond the immediate reach of small working groups with operational priorities. However, there are plans to progressively merge all the information available into a multivariate analysis scheme, by extending the interpolation procedure to other variables measured near the soil surface. Physical relations among different variables can then be included as weak or strong (Lagrangian) constraints, by implementing a variational formulation of the analysis algorithm.

**Appendix A: Observation influence**

The ‘Degrees of Freedom for Signal’ (DFS) parameter is defined as the trace of the influence matrix \( W \). It has been used with four-dimensional data assimilation schemes by Cardinali et al. (2004) and Fourrié et al. (2006), who used it to assess the impact of different observational components in the North Atlantic THORPEX Regional Campaign. From Equation (4),

\[ y^a = W y^o + (I - W)y^b \]  \hfill (A.1)

Each element of the matrix \( W \) can be interpreted as the first-order ‘sensitivity’ of a particular analysed value with respect to an observation

\[ \frac{\partial y^a_m}{\partial y^o_k} = [W]_{m,k} \]  \hfill (A.2)

Since

\[ I - W = R(S + R)^{-1} \]  \hfill (A.3)

if \( R \) is diagonal: \( R = \text{diag} (\sigma_{o,1}^2, \ldots, \sigma_{o,M}^2) \), the following relation between the \( m \)th diagonal elements is obtained

\[ 1 - [W]_{m,m} = \sigma_{o,m}^2 (S + R)^{-1}_{m,m} \]  \hfill (A.4)

These diagonal elements are useful to obtain the CV analysis. When \( R \) is diagonal one obtains

\[ (y^a_m - y^o_m)^2 = \sigma_{o,m}^2 [(S + R)^{-1}]_{m,m} (y^a_m - y^o_m) \]  \hfill (A.5)

A similar equation has been used by Cardinali et al. (2004).
The IDI for station \( m \) is re-defined here as the sum of all observational sensitivity contributions to the \( m \)th analysed value:

\[
\gamma_m^{\text{IDI}} = \sum_{k=1}^{M} \frac{\partial \gamma_m}{\partial y_k} (A.6)
\]

where all the terms \( \frac{\partial \gamma_m}{\partial y_k} \) are non-dimensional because, in the present univariate context, all observations have the same physical dimensions. By making use of Equations (A.2) and (4), the IDI parameter can be easily seen as the analysis obtained on the \( m \)th station when all observed values are set to the value 1 and all background values are set to 0, i.e. the intuitive definition given in Section 2.3.

By making use of the CV-IDI parameter \( \tilde{\gamma}_m^{\text{IDI}} \), defined in Section 2.3, a relation analogous to Equation (A.5) between IDI and the diagonal elements of the influence matrix (appearing in the DFS definition) is obtained when \( R \) is diagonal:

\[
(\gamma_m^{\text{IDI}} - 1) = (1 - [W]_{m,m}) (\tilde{\gamma}_m^{\text{IDI}} - 1) \quad (A.7)
\]

In principle, the IDI definition can be generalized to other assimilation algorithms, at least for each dimensionally homogeneous component of an observational network.

Appendix B: Updating the inverse matrix

As long as observations are taken at stations having fixed locations, and covariance estimates are stationary in time, the error covariance matrices change only because of occasionally missing data. Moreover, direct knowledge of the diagonal elements of the matrices \( (S + R)^{-1} \) and \( S(S + R)^{-1} \) is useful for diagnostic purposes and for quality control. It is then useful and efficient to compute the inverse matrix \( (S + R)^{-1} \) for a typical data set, to store it in memory, and to update it at each analysis time to account for changes in the data set. For the case of a missing datum this means to discard a row and a column of the matrix \( (S + R) \); for the case of a new datum, absent or invalid at the previous analysis time, this means inserting an additional row and column (they are inserted at the end here; they can be sorted afterwards, if needed).

Consider the block decomposition of the \((M,M)\) symmetric matrix \( S + R \):

\[
S + R = \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \quad (B.1)
\]

Here the scalar \( c = [S + R]_{M,M} \) is the last diagonal element of \( S + R \); \( b \) is a column vector of length \( M - 1 \), so that \( b^T \) is the \( M \)th row of the symmetric matrix \( S + R \) (and the transpose of its \( M \)th column); \( A \) is the \((M - 1, M - 1)\) symmetric sub-matrix obtained from \( S + R \) by taking its last column and row away. Proceeding in the same way, the inverse of \( S + R \) is

\[
(S + R)^{-1} = \begin{bmatrix} X & q \\ q^T & z \end{bmatrix} \quad (B.2)
\]

where \( X, q, \) and \( z \) have respectively the same dimensions as \( A, b, c \). Define \( I \) as the identity matrix of order \( M - 1 \), \( 0 \) as the column vector, and \( 0^T \) the corresponding row vector, composed of \((M - 1)\) zeroes. The inverse matrix relation is then

\[
\begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} X & q \\ q^T & z \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix} \quad (B.3)
\]

that is to say

\[
AX + bq^T = I \quad (B.4)
\]

\[
Aq + bz = 0 \quad (B.5)
\]

\[
b^TX + cq^T = 0^T \quad (B.6)
\]

\[
b^Tq + cz = 1 \quad (B.7)
\]

**From \( M - 1 \) to \( M \)**

In case the inverse \( A^{-1} \) of \( A \) is known and one new observation becomes available, then the inverse of order \( M \) can be computed by defining the auxiliary vector \( p \) as

\[
p = A^{-1}b \quad (B.8)
\]

and proceeding as follows:

\[
z = \frac{1}{b^Tb - c} \quad (B.9)
\]

\[
q = -zp \quad (B.10)
\]

\[
X = A^{-1} - pq^T \quad (B.11)
\]

This can be done under the condition that \( A^{-1} \) exists and that \( c \neq b^TA^{-1}b \).

**From \( M \) to \( M - 1 \)**

When the inverse of order \( M \) is known and the \( m \)th observation is missing, then the inverse of order \( M - 1 \), \( A^{-1} \), can be computed, if \( z_m \neq 0 \), as

\[
A^{-1} = X_m + \frac{1}{z_m} q_mq_m^T \quad (B.12)
\]

\( z_m \) is here the \( m \)th diagonal element of \( (S + R)^{-1} \); \( q_m \) is the \( m \)th row (without the diagonal element \( z_m \)); \( q_m^T \) the \( m \)th column; and \( X_m \) is what remains of the inverse matrix when the \( m \)th row and column are taken away.
Appendix C: Response function of the optimal interpolation for an ideal data distribution

The OI analysis increment, when \( R = \varepsilon^2 I \) is

\[
\delta x = G (S + \varepsilon^2 I)^{-1} d
\]  \hspace{1cm} (C.1)

It is well known that this expression can be obtained as the limit of the SC iteration (Bratseth, 1986; Daley, 1991; Pedder, 1993; Uboldi and Buzzi, 1994). In fact, the inverse matrix can be written as the sum of a series [converging when all eigenvalues of \((S + \varepsilon^2 I)\) have values between 0 and 2]:

\[
\delta x = G \sum_{j=0}^{\infty} [I - (S + \varepsilon^2 I)]^j d \hspace{1cm} (C.2)
\]

where each partial sum

\[
\delta x_n = G \sum_{j=0}^{n} [I - (S + \varepsilon^2 I)]^j d = G \sum_{j=0}^{n} s_j \hspace{1cm} (C.3)
\]

can be interpreted as an \( n \)-step SC scheme (Uboldi and Buzzi, 1994), with

\[
s_j = [I - (S + \varepsilon^2 I)] s_{j-1} \hspace{1cm} (C.4)
\]
\[
s_0 = d \hspace{1cm} (C.5)
\]

Barnes (1964, 1973) calculated the response function of the SC scheme in the idealized case of a one-dimensional continuous data distribution and Gaussian correlation functions. By proceeding in the same way, the OI response function can be calculated as the limit. In the same idealized condition, each field is a function of the space coordinate \( \xi \), and the matrix multiplication becomes a convolution:

\[
\delta x(\xi) = \sum_{j=0}^{\infty} g_j(\xi) \hspace{1cm} (C.6)
\]
\[
g_j(\xi) = \frac{1}{\sqrt{2\pi}D} \int e^{-\frac{1}{2} \left( \frac{\xi - \eta}{D} \right)^2} s_j(\eta)d\eta \hspace{1cm} (C.7)
\]
\[
s_j(\xi) = (1 - \varepsilon^2) s_{j-1}(\xi) - g_{j-1}(\xi) \hspace{1cm} (C.8)
\]
\[
s_0(\xi) = d(\xi) \hspace{1cm} (C.9)
\]

The innovation, the SC analysis at step \( n \), and the OI analysis (the limit) are then written as Fourier integrals. Respectively,

\[
d(\xi) = \frac{1}{\sqrt{2\pi}} \int e^{-ik\xi} \omega(\xi) d\xi \hspace{1cm} (C.10)
\]
\[
\delta x_n(\xi) = \frac{1}{\sqrt{2\pi}} \int e^{-ik\xi} \gamma_n(\xi) d\xi \hspace{1cm} (C.11)
\]
\[
\delta x_{OI}(\xi) = \frac{1}{\sqrt{2\pi}} \int e^{-ik\xi} \gamma_{\infty}(\xi) d\xi \hspace{1cm} (C.12)
\]

Figure 12. Response function for OI (solid), and for different SC steps: 1 (dashed), 2 (short-dashed), 5 (dotted). The SC curves approach the OI curve as the step number increases. The independent variable is wavelength \( \lambda \) in unit \( D \). The horizontal line corresponds to the limit value \( \frac{1}{1 + \varepsilon^2} \).

The response function for SC at step \( n \) is calculated by making use of the properties of the Fourier transform and of the Gaussian functions. Written as a function of wavelength \( \lambda = \frac{1}{k} \), it is

\[
\Gamma_n(\lambda; \varepsilon^2, D) \equiv \frac{\gamma_n}{\omega} = \left\{ 1 - \left[ 1 - \left( \varepsilon^2 + e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2} \right)^{\gamma+1} \right] \right\} \frac{\varepsilon^2 + e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2}}{\varepsilon^2 + e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2}} \hspace{1cm} (C.13)
\]

The limit, i.e. the response function for OI is then

\[
\Gamma_{OI}(\lambda; \varepsilon^2, D) \equiv \frac{\gamma_{\infty}}{\omega} = \frac{e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2}}{\varepsilon^2 + e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2}} = 1 - \frac{\varepsilon^2}{\varepsilon^2 + e^{-\frac{1}{2} \left( \frac{D}{\lambda} \right)^2}} \hspace{1cm} (C.14)
\]

These functions are plotted in Figure 12: they appear as approximations of a step function. The SC and OI analysis are then interpreted as realizing a low-pass filter whose cut-off wavelength is approximately \( D \).

Acknowledgements

The authors thank Maria Ranci who participated in discussions regarding the statistical evaluation of the interpolation scheme and Umberto Pellegrini for his help in the meteorological characterization of the cases presented here. The data used in the presented results were provided by ARPA Lombardia. A previous version

Copyright © 2008 Royal Meteorological Society
of the analysis technique was implemented for ARPA Piemonte and another version was recently adapted for the weather service of P. A. Trento. Part of this work has been performed within the project ‘FORALPS – Meteo-hydrological forecast and observations for improved water resource management in the Alps’, supported by the European Union through the European Regional Development Fund under the Interreg III B ‘Alpine Space’ Initiative.

References