Horizontal structure of height-forecast errors over the southern part of South America

By CAROLINA VERA*

University of Buenos Aires, Argentina

(Received 10 March 1993; revised 18 January 1994)

SUMMARY

A study of the height-forecast error autocovariances was made for a simple intermittent data-assimilation system (SADI) applied over the southern part of South America. The main objective of this study was to determine from the data the height-forecast error autocorrelation, and to provide new estimations of the forecast- and observation-error standard deviations. In performing this derivation several significant differences were found between the estimates obtained with a simple forecast model in an area which is data sparse, and those derived for data-rich regions and advanced forecast models. The forecast-error correlations were fitted to the Fourier–Bessel series and also to a negative square exponential (NSE) function at the six model levels. The fitted Fourier–Bessel series was able to represent more accurately the wavelength range resolved by the analysis scheme than the fitted NSE function. New estimations of forecast and observation errors have been obtained from the Fourier–Bessel expansion. As the representative errors are much larger because of the coarse resolution of the network and the low-order truncation of the fit, the observation errors are greater than those used for the northern hemisphere analysis systems. Results also show that the forecast errors are dominated by the synoptic-scale components in all the levels considered. An evaluation of the impact of using the new set of horizontal covariances was done and it mainly showed that the analysis done with the fitted Fourier–Bessel series better represents the data than both the original SADI system and an analysis system using a fitted NSE function.

1. INTRODUCTION

An intermittent data-assimilation system (SADI) was developed by Vera et al. (1990) at the Department of Atmospheric Sciences of the University of Buenos Aires. SADI uses a multivariate three-dimensional statistical interpolation analysis scheme, and the height-forecast error horizontal correlations are modelled using the negative square exponential (NSE) function. SADI is substantially different from other data-assimilation systems because it uses a quasi-geostrophic prediction model, and the forecast first-guess is for 24 hours. Thus it is necessary to determine a statistical structure of forecast errors specifically for the SADI in Argentina, rather than using other approximations already in use for the northern hemisphere.

Thiébaux (1980) shows that the accuracy of the analyses based on optimum interpolation (OI) schemes is strongly dependent on the horizontal correlation function shape. She points out that although the differences between several horizontal correlation function shapes could be small, the accuracy of the corresponding analyses is highly different, not only for the height but also for wind components. Thiébaux et al. (1986) suggests that wind horizontal correlation functions obtained geostrophically could be responsible for such differences in analysis accuracy. Balgovind et al. (1983) show that the geostrophic wind horizontal correlation functions resulting from differentiating similar height horizontal autocorrelation functions can have significant differences. Thiébaux et al. (1986) points out disadvantages of the NSE function used in several analysis systems (e.g. Lorenc 1981; DiMego 1988) as well as in the SADI system.

Rutherford (1972), Hollingsworth and Lonnberg (1986) and Lonnberg and Hollingsworth (1986) have developed a comprehensive three-dimensional description of the covariance structure of the height-forecast errors. They showed that a height horizontal correlation function based on a Fourier–Bessel series provides a spectral representation

* Corresponding address: Department of Atmospheric Sciences, University of Buenos Aires, Centro de Investigaciones del Mar y la Atmosfera (CIMA/CONICET), Buenos Aires, Argentina.
of the forecast-error statistics over the band of wavelengths resolved by the model and the radiosonde network. This kind of representation of the horizontal forecast-errors correlation is used in the European Centre for Medium-range Weather Forecasts (ECMWF) global data-assimilation system.

In this paper the statistical structure of height-forecast errors are studied by comparing forecast values with verifying radiosonde data over Argentina and Chile. The knowledge of that statistical structure allows new estimations of the forecast- and observational-error standard deviations as well as a good determination of a mathematical function that represents the height-error autocorrelation function. To improve the correlation representation in the SADI analysis scheme, both the NSE and the Fourier-Bessel series were used to fit data in this paper.

Although Hollingsworth and Lonnberg (1986) assumed that height-forecast error correlations are horizontally homogeneous, Balgovind et al. (1983) suggested that there is a strong latitudinal variation due to the dependence of forecast-error structure on the Rossby deformation radius. Thiébaux et al. (1986) also obtained a longitudinal variation in the forecast-error structure of different regions within the same latitudinal band. In this paper, a discussion of the validity of homogeneity and isotropy conditions over the extratropical South American region is also presented.

The paper is organized in the following way: in section 2 the SADI system is described; in section 3 the data used are given; in section 4 the proposed function based on the Fourier-Bessel series is presented; in section 5 the empirical correlations and the fitting by the Fourier-Bessel series are discussed; and in section 6 a validation of the fits is described. Finally the conclusions are given in section 7.

2. THE ASSIMILATION SYSTEM

SADI was developed to be used by the numerical weather-prediction centres in Argentina. It has three main stages: pre-analysis, objective analysis and forecast model.

The pre-analysis procedure is based on one developed by DiMego (1988) for the National Meteorological Center (NMC) regional data-assimilation system. All the observations available in the forecast domain are checked by two quality-control procedures. The first is a gross check control where the observation values are subtracted from the corresponding forecast ones. The resulting differences are compared with a limit value dependent on both the forecast-error standard deviation and the latitude. The second procedure is a consistency check that compares neighbouring observations minus forecast differences among themselves. That check, called ‘buddy check’, is based on one developed by DiMego (1988) for the NMC regional analysis system.

The objective analysis scheme was developed by Vera and Nuñez (1987) and is based on the optimum interpolation theory (Gandin 1963). The interpolation is multivariate and three-dimensional, using height, wind and thickness observations to analyse height fields. Its characteristics are briefly presented here. In this scheme the horizontal height-forecast error correlations are modelled by an NSE function. Originally, before the present study was performed, the NSE function parameters were directly adopted from the values used by DiMego (1988) in the NMC regional analysis scheme. The wind-forecast error correlations are derived from the stream-function autocorrelation function (Lorenz 1981), neglecting the potential-velocity component. The cross-correlation between stream function and height-forecast errors is defined as a function of the vertical level and the latitude to allow an uncoupling between mass and wind errors in tropical latitudes and also at the surface. The observational errors are assumed to be uncorrelated among different instruments, and also among different variables. A horizontal observational-
error correlation is only considered for thickness data derived from satellite temperature soundings. The data selection procedure is essentially the same as DiMego (1988).

In SADI a five-layer baroclinic quasi-geostrophic model is used. This model was first developed by Trenberth (1973) and was adapted to be used over the South America region by Possia et al. (1987). In the absence of friction and diabatic heating, the model conserves quasi-geostrophic potential vorticity and makes predictions of height fields at 850, 700, 500, 300 and 200 hPa. The forecast at 1000 hPa is derived diagnosisly. In this version, the model includes a smoothed South America orography and is initialized with climatological moisture. The integration area is approximately between 25°S and 73°S and 110°W and 25°W. A stereographic polar projection is used with the tangent plane at 60°S where the grid distance is 300 km.

3. DATA

The data used in this paper correspond to the differences among the radiosonde height data and 24-hour forecasts valid at 12 GMT from 1 June to 31 August 1986, over the southern part of South America. The differences were obtained on the six model forecast levels, over all the radiosonde stations located in the model grid. A station was included if at least 60% of the data were available and accepted by the quality-control system. The mean observation-minus-forecast difference was removed for each station to eliminate instrumental bias and mean forecast errors. The ensemble-mean variance was defined as the mean of all the station variances.

4. MODELLING OF THE HEIGHT-FORECAST ERROR COVARIANCES

(a) Proposed correlation model

If the forecast errors are homogeneous and isotropic (the validity of such assumptions will be discussed in section 5) the correlation \( F \) of two horizontal points is only a function of the separation \( r \) between them.

Following the general theory (Buell 1972; Rutherford 1972; Hollingsworth and Lonnberg 1986), \( F(r) \) is the autocorrelation function of a homogeneous random process if, and only if, it can be expressed by its Fourier transform, and this transform is positive for all wave numbers. In two dimensions the Fourier transform can be written as a Hankel transform

\[
F(r) = \int_0^\infty u f_h(u) J_0(u r) \, du
\]  

where \( r \) is a distance and \( u \) is a wave number and

\[
f_h(u) = \int_0^r r F(r) J_0(u r) \, dr
\]  

is the zero-order transform of \( F(r) \), and \( J_0 \) is the zero-order first-type Bessel function. This kind of representation has the advantage of producing negligible amplitudes far from the reference point, whereas the Fourier transform does not decrease with distance.

For a variable over a limited range, the integral (1) can be expressed in a Fourier–Bessel series

\[
F(r) = \sum_{l=0}^{N} A(k_l) J_0 \left( k_l \frac{r}{D} \right)
\]
where \(k_i\) are the wave numbers of the Bessel function expansion that are determined (Hildebrand 1976) by the requirement of vanishing radial derivative of \(J_0(k_i)\) at \(r = D\). The zero-order term of the expansion corresponds to the forecast errors that are perfectly correlated over the domain. This term hereafter will be called the large-scale component while the other terms will be denoted the synoptic components.

\(A(k_i)\) can be obtained by multiplying both sides of Eq. (3) by \(rJ_0(k_ir/D)\), then integrating over the range \((0, D)\), and after assuming that the integral of the infinite sum is equivalent to the sum of the integrals, Eq. (3) reduces to

\[
A(k_i) = \int_0^1 rF(r)J_0\left(\frac{k_i r}{D}\right) \frac{dr}{D^2J_0^2(k_i)}
\]

by virtue of the orthogonality of the \(J_0\) functions.

Modelling the corresponding horizontal structures for wind-forecast errors is impossible for our system, because the wind is not a forecast variable of the SADI quasi-geostrophic model. Thus the horizontal correlations for wind-forecast errors were geostrophically obtained from the height horizontal autocorrelations (see the appendix).

(b) Estimation of the forecast- and observation-error standard deviations

Assuming that forecast and observation errors are not correlated at the station location, the local standard deviation over each of the stations can be written as

\[
E_i^f = \sqrt{\langle (P_i - O_i)^2 \rangle} = \sqrt{(E_i^f)^2 + (E_i^o)^2}
\]

where \(P_i\) and \(O_i\) are the forecast and observation values over the station \(l\) respectively; \(E_i^f\) and \(E_i^o\) are the forecast- and observation-error standard deviations respectively.

The forecast-error correlation between two points \(l\) and \(j\) is

\[
F_{lj} = \frac{\langle (P_l - O_l)(P_j - O_j) \rangle}{E_l^f E_j^f}.
\]

If \(F_{lj}\) is extrapolated to zero separation, an estimate of forecast- and observation-error standard deviations can be obtained from

\[
F_{lj} = \frac{(E_i^f)^2}{(E_i^f)^2 + (E_i^o)^2}.
\]

This kind of estimation was widely used by several authors (Gandin 1963; Rutherford 1972; Hollingsworth and Lonnberg 1986; Lonnberg and Hollingsworth 1986). It assumes uncorrelated and uniform observation errors representing not only the instrumental errors but also contributions from real small-scale motions not resolved by the analysis grid (representativeness errors).

(c) The covariance determination

Before modelling the statistical structure of the forecast errors it is necessary to determine empirical correlation values for all pairs of stations available. To reduce the heterogeneity and the computational time, correlation data are then averaged over certain distance ranges or 'bins' (Hollingsworth and Lonnberg 1986; Thibaux et al. 1986). The Fisher \(z\)-transform is used to compute those averages (Fisher 1921), because it makes the distribution of correlation coefficients close to Gaussian.

It is necessary to select boundary conditions applied to Eq. (3) to determine the wave numbers \(k_i\). As mentioned before, it is convenient to impose the condition that \(F_{lj}\)
should vanish at a distance \( D \). With this boundary condition the zero term in the expansion (3) is constant and independent of \( r \), representing the mean correlation value over the domain (Hildebrand 1976). Hollingsworth and Lonnberg (1986) point out that for a scalar variable as the height, the local analysis can respond to data information in the range between the upper limit of the resolved scales (dependent on \( D \)) and the lower limit which is determined by the highest wave-number term included in the expansion and by the data density. The information contained in scales larger than the upper limit will be mainly projected in the constant term. As will be seen later, the height correlations are small at distances longer than 1500 km. Since the SADI data-selection procedure has a search radius also of 1500 km I compute correlations over twice that distance. Thus a value of \( D = 3000 \) km was used to compute the expansion terms. Following Hollingsworth and Lonnberg (1986), the quantity \( L_i = 2\pi D / k_i \) provides a convenient definition of the scale of each horizontal component (Table 1). It is also useful to define the horizontal characteristic length-scale for the correlation function. For the Fourier–Bessel series this is

\[
L_F^2 = \left[ \frac{F}{\nabla^2 F} \right]_{r=0} = D^2 \frac{\sum_{i=0}^{N} A_i}{\sum_{i=0}^{N} k_i^2 A_i}.
\]

(8)

This corresponds to the spectrally weighted average of the horizontal scale of the components of \( F \).

<table>
<thead>
<tr>
<th>( i )</th>
<th>( k_i )</th>
<th>( L_i ) (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>( \infty )</td>
</tr>
<tr>
<td>1</td>
<td>3.8</td>
<td>4957</td>
</tr>
<tr>
<td>2</td>
<td>7.0</td>
<td>2695</td>
</tr>
<tr>
<td>3</td>
<td>10.2</td>
<td>1847</td>
</tr>
<tr>
<td>4</td>
<td>13.3</td>
<td>1420</td>
</tr>
<tr>
<td>5</td>
<td>16.5</td>
<td>1143</td>
</tr>
<tr>
<td>6</td>
<td>19.6</td>
<td>961</td>
</tr>
<tr>
<td>7</td>
<td>22.8</td>
<td>829</td>
</tr>
<tr>
<td>8</td>
<td>25.9</td>
<td>729</td>
</tr>
</tbody>
</table>

5. Analysis of SADI Height-Forecast Errors

(a) Horizontal height autocorrelations

For each isobaric level the height-forecast error correlations with the original SADI system were computed between each station pair. Figure 1 shows forecast-error correlations, for four of the six model forecast levels, as a function of the distance (in kilometres). As expected, the correlations show a decay as distance increases. The volumes of data used here are not as large as those presented in other papers (Thiébaux et al. 1986; Andersson et al. 1986; Lonnberg and Hollingsworth 1986) owing to the sparse radiosonde network available in southern South America. It can be seen that the dispersion of the points is minimum at 500 hPa.
Figure 1. Forecast-error autocorrelation for height versus distance at (a) 850 hPa, (b) 700 hPa, (c) 500 hPa and (d) 300 hPa.
The dispersion can be due to either observational noise or to anisotropic and inhomogeneous characteristics in the height-forecast error statistical structures. Figures 2(b) and 2(c) show one-point correlation values considering two base points at different latitudes, as well as the corresponding fit of Fourier-Bessel expansions. It is apparent that the correlation decays faster with distance for the higher-latitude base point. Figures 2(a) and 2(b) show one-point correlations for two base points at different longitudes, indicating that the eastern and western station points also have different decays with

![Correlation Graphs](image)

Figure 2. One-point forecast-error autocorrelation for height versus distance at 500 hPa centred on (a) 33°S, 72°W, (b) 34°S, 58°W and (c) 27°S, 59°W, together with the corresponding fit of Fourier-Bessel expansions.
distance. The correlations of the station located at the west slope of the Andes seem to
decay to a value greater than zero at long distances. The reason for this characteristic
behaviour will be discussed in the next subsection. It is apparent that the forecast-error
structure over the region has both latitudinal and longitudinal inhomogeneities.

In order to assess whether the dispersion of the points in Fig. 1 is due to anisotropic
forecast-error correlations, I computed one-point correlations of the 24-hour forecast
differences with the ECMWF analysis. Figure 3 shows an example of such an auto-
correlation field, indicating a clear anisotropy associated with the presence of the Andes.

However, since the radiosonde network over South America is very sparse, I am
forced to adopt a homogeneous and isotropic correlation model. But the fact that these
assumptions are not well supported by the data will affect the analysis made with a
Fourier–Bessel series.

(b) Forecast-error correlation fit

Using Fisher’s z transform, for each model level, the horizontal correlations were
averaged each 200 km and then fitted by a Fourier–Bessel series. In order to compare
the proposed correlation model with the one previously used in SADI, a fit by an NSE
function was also done. Computations with 100 km bins did not show much sensitivity
to the size of the bin.

In order to estimate how many terms should be included in the Fourier–Bessel
series, and avoid overfitting, I first present, in Table 1, the wavelengths for different
terms corresponding to $D = 3000$ km. The fact that the distance between stations is about
300 km, a distance also representative of the grid size, suggests that terms higher than
$i = 5$ should not be included, which have wavelengths of about four times that distance.
Since this could still represent overfitting given the sparsity of the data, I repeated the
estimation of the coefficients using fewer terms in the series, stopping the expansion at
the first negative coefficient or when the power started to increase with wave number.
Figure 4 shows the power-percentage distribution using both six terms and the lower-
order fitting. The fact that the power percentage of the large-scale component increases
as the truncation is increased can be expected from the expression for the normalized
variance spectrum derived from (1), which is $A_i J_0^2(k)$. Since $J_0^2 = 1$ for wave number
zero and it decreases rapidly with wave number, a truncation at lower wave numbers
will redistribute the power most strongly into the $A_6$ contribution. Also Fig. 4 shows that
the total power increases as the truncation is increased, suggesting that the $J_0$’s are not
exactly orthogonal on the grid defined by the data points, as pointed out by Rutherford
(1972) and Hollingsworth and Lonnberg (1986).

Figure 5 shows that the original SADI NSE function decays faster with distance than
the fitted correlations. This behaviour is to be expected because the original function
was applied in a data-assimilation system using primitive-equation models and 6-hour
forecast first guesses. The data and the fitted correlations also present a vertical variation
of the decay with distance, suggesting it is not desirable to use the same correlation
representation for all the vertical levels.

Note that at 850 and 700 hPa the averaged correlations and their fitted functions
tend to a value greater than zero at long distances, whereas at higher levels the correlations
tend to zero. This is probably due to model deficiencies that introduce large-scale forecast
ersors. It is clear that the SADI regional model, being quasi-geostrophic, cannot handle
well the orography since the Andes are represented as lower and wider than in reality.
In addition, the quasi-geostrophic model cannot simulate well the land–sea contrast.
Such model deficiencies should be expected to affect the lower levels more than the mid
levels, as observed in Fig. 5.
Figure 3. One-point 24-hour forecast-minus-ECMWF analysis difference autocorrelation for height at (a) 850 hPa and (b) 500 hPa.
Figure 4. Power-percentage distribution of the Fourier–Bessel expansion for two different truncations at (a) 850 hPa, (b) 700 hPa, (c) 500 hPa and (d) 300 hPa.

The fit by an NSE function is less close to the averaged correlations than the one by the Fourier–Bessel series, which is to be expected given the larger number of degrees of freedom of the latter. For example, to fit the 500 hPa data the NSE uses two parameters whereas the Fourier–Bessel series uses four. Also, at all levels, the zero intercept of the NSE function is smaller, indicating that shorter scales are not resolved by this function and so treated as horizontally uncorrelated observational errors (Eq. 7). This can also be seen in Fig. 6, which presents the vertical profile of the horizontal characteristic length-scale for both the fitted Fourier–Bessel series and the NSE function. The NSE
Figure 5. Averaged forecast-error autocorrelation for height, the original SADI NSE function and the fits by both a Fourier-Bessel series and an NSE function at (a) 850 hPa, (b) 700 hPa, (c) 500 hPa and (d) 300 hPa.
resolves less well the higher wave numbers so that its estimated correlation length-scale is larger than for the Fourier–Bessel fit. As mentioned above, both methods result in larger correlation length-scales at the lowest levels due to model deficiencies. The fact that correlation length-scale increases again at 200 hPa, the top level of the model, is probably also due to model deficiencies associated with the presence of an artificial lid.

The autocorrelation functions of the wind-forecast errors and cross-correlation functions of wind and height are geostrophically derived from the height-forecast error autocorrelation function following Lorenc (1981) and Buell (1972). The derivations of the wind correlation functions are presented in the appendix and the set of equations (A.4) gives the expressions for the horizontal correlations based on the Fourier–Bessel series. As an example, Fig. 7 shows both the autocorrelations for height and \( u \) component and also the cross-correlations between height and \( v \) component based on the original SADI NSE function, the Fourier–Bessel series and the NSE function. They are presented for 500 hPa with a cross-correlation value between stream function and height equal to 1. The set of correlations based on the Fourier–Bessel series shows weaker values of both autocorrelations and cross-correlations than the corresponding ones obtained with the SADI NSE function. A comparison between the set of correlations based on the Fourier–Bessel series and the corresponding set based on the fitted NSE function reveals stronger wind autocorrelations associated with the NSE function while the cross-correlations derived from the Fourier–Bessel series are stronger.

(c) Estimation of forecast and observation errors

Figure 8 shows the root-mean-square forecast-minus-observation differences for the considered period, denoted as height total error. The squared total error should be equal to the sum of the forecast and observation variances (Eq. 5), but Fig. 8 shows that the old SADI estimations of the total errors are lower than the errors obtained from the
Figure 7. Horizontal prediction error autocorrelations for height ($\phi$), $u$ and the cross-correlation between $\phi$ and $v$ component at 500 hPa based on (a) the original SADI NSE function, (b) the Fourier–Bessel series, and (c) the fitted NSE function.

actual data at all levels. Thus it is necessary to obtain a new estimation of forecast and observation errors.

Using Eqs. (5) and (7) new estimations of both forecast and observation errors are obtained. Figure 9 shows that forecast errors estimated from the Fourier–Bessel expansion have a vertical variation similar to that of the total errors. They are also larger than
Figure 8. Height total error and the old SADI estimation.

Figure 9. Height total error and the new estimation of forecast and observation errors from the Fourier–Bessel series fit.
the observational errors at all vertical levels. The estimated observational errors increase slowly with height, as could be expected, but are larger than those used for the northern hemisphere analysis systems (Lonnberg and Hollingsworth 1986; Mitchell et al. 1990). This can be understood when we consider that these errors include not only random observational errors, which are similar throughout the world rawinsonde network, but also the representativeness errors, much larger for the coarse resolution SADI system and the low-order truncation of the fit.

In Fig. 10 we separate the forecast error into large-scale (wave number zero) and synoptic components. At all levels the contribution of the synoptic-scale part of the forecast errors is dominated by errors in the position and amplitude of the longer synoptic waves.

![Figure 10. Height-forecast error and both large-scale and synoptic component contributions.](image)

6. **VALIDATION OF THE FITS**

In this section the SADI analyses made using both the Fourier-Bessel series and NSE function fits are evaluated over the southern part of South America. This analysis verification was done withholding an observation, performing the analysis at its location, and comparing the analysed and observed values through the computation of the root-mean-square analysis-minus-observation differences \((AO)\). This methodology is an important tool to evaluate the analysis quality and it was used by several authors (Gandin 1963; Thiébaux 1975, 1977, 1980; Seaman and Hutchinson 1985; Vera 1993). Vera (1994) pointed out that the best analysis would not only be the one that is best verified by the observations, but also that one which provides comparable empirical and estimated analysis errors. As in this verification system there is no correlation between observation and analysis errors, the empirical mean-square analysis errors, \((AO)^2\) obtained from the verification method gives

\[
(AO)^2 = (E^o)^2 + (E^a)^2
\]  

(9)
where $E^o$ is the observation error estimated from the correlation fits and $E^a$ is the analysis errors estimated by the statistical interpolation.

Using the same data set presented in section 3, the analysed values were computed over all rawinsonde locations using the original SADI forecast-error covariances and those obtained from both the Fourier–Bessel series and the NSE function fits. As an example, the analysis verification results for 850 hPa will be described. Figure 11 shows that the $AO$ values obtained with the original SADI system are larger than the corresponding root-mean-square prediction-minus-observation differences ($PO$) while the $AO$ values obtained for the analysis made with the Fourier–Bessel series fit are smaller than both of them in most of the rawinsonde locations. The $AO$ values obtained for the NSE function fit (not shown) are much larger, probably due (as shown in subsection 5(b)) to larger observational errors and a larger correlation characteristic length-scale which give less weight to the sparse and few observations available in the region.

As pointed out in section 5(c), the observation errors estimated from the Fourier–Bessel fit are larger than those used in other analysis systems. This reduces the impact of the observations on the analysis and it is a problem in the data-sparse network over South America. Mitchell et al. (1990) have similar problems with the surface data and they reduce their observation errors to give more weight to these data. In order to reduce the rawinsonde observation errors, I had to take other arbitrary decisions, including more terms in the Fourier–Bessel fit. If a term of the expansion had a negative sign but explained a low percentage of the variance, the term was reset to a very small positive number and the expansion was continued (Lomnberg and Hollingsworth 1986). New estimations of both forecast and observation errors were derived from the higher-order fit. Figure 11(d) shows that the $AO$ values obtained using the Fourier–Bessel series of high-order truncation are smaller at most of the station locations. It is also encouraging that the values of $AO$ are lower or similar than the estimated observation error for 850 hPa, indicating that the analysis is drawing tighter to the data.

The validity of (9) was also verified. As expected, Fig. 12 shows that the differences between $(AO)^2$ and the estimated analysis errors are lower for the analyses made using the Fourier–Bessel series than for the original SADI system. This figure also shows similar difference values between the analyses made with the Fourier–Bessel series with both high- and low-order truncation. This last result means that using an artificial extended Fourier–Bessel series in the SADI analysis did not greatly affect the similarity between the statistical structure of the empirical analysis errors and the one estimated using both the analysis errors derived from the statistical interpolation scheme and the observation errors obtained from the forecast-error correlation fit.

7. CONCLUSIONS

A study of the height-forecast error autocovariances was made for a simple intermittent data-assimilation system applied over the southern part of South America. The main objective of this study was to determine from the data, the height-forecast error autocorrelation and to provide new estimations of the forecast- and observation-error standard deviations.

In performing this derivation, several significant differences were found between the estimates obtained with a simple forecast model in an area which is data sparse, and those derived for data-rich regions and advanced forecast models. It was found that homogeneity and isotropy conditions are not fulfilled over the region. But owing to the low number of radiosonde observations available, it was still necessary to use an isotropic and homogeneous function over all the region.
The forecast-error correlations were fitted to the Fourier–Bessel series and also to an NSE function at the six model levels. Results showed the original SADI NSE parameters adopted from northern hemisphere systems did not agree with forecast-minus-observation data. The fitted Fourier–Bessel series was able to represent more accurately the wavelength range resolved by the analysis scheme than the fitted NSE function.
Figure 12. Mean-square analysis errors $(A_0)^2$ minus estimated total analysis error differences at 850 hPa for (a) the original SADI NSE function and the Fourier-Bessel series fit of (b) two components and (c) five components.

A vertical variation of the correlation characteristic length scale $L_t$ was found. Also in all levels $L_t$ values for both fits are greater than those previously considered in SADI. An inaccurate handling of orography and the land–sea contrast by the forecast model produced greater $L_t$ values than expected in low levels.

The comparison between the derived correlations for wind previously used in the SADI multivariate analysis, with those obtained from both the Fourier–Bessel series and
the fitted NSE function gives important differences that will have a significant impact on the multivariate analysis.

New estimations of forecast and observation errors have been obtained from the Fourier–Bessel expansion. As the representative errors are much larger, because of the coarse resolution of the network and the low-order truncation of the fit, the observation errors are greater than those used for the northern hemisphere analysis systems. Results also show that the forecast errors are dominated by the synoptic-scale components in all the levels considered.

An evaluation of the impact of using the new set of horizontal covariances was done, comparing the analyses with rawinsonde observations which were withheld from the interpolation. These results mainly showed that the analysis done with the fitted Fourier–Bessel series represent better the data than both the original SADI system and an analysis system using a fitted NSE function. But by using large estimated observation errors such as those obtained, the impact on the analysis of the few and sparse data available in the region is reduced. Thus a Fourier–Bessel series with a higher-order truncation was also verified and the resulting analysis gives a better analysis fit to the data. Although the high-order truncation was not justified by the data expansion, I have been forced to adopt it to obtain a more reliable analysis in the region.

Although the analysis quality over the South American region is strongly affected by the sparsity of the network, the results obtained in this work also show that the use of a more complex forecast model would allow the analysis system to use the few observations available more efficiently.

ACKNOWLEDGEMENTS

This work is part of my Ph.D. thesis at the University of Buenos Aires. I am grateful to Dr Eugenia Kalnay for her encouragement and valuable counsel as my co-advisor, and to Dr Lev Gandin for very helpful discussions. The suggestions of two anonymous reviewers are also gratefully acknowledged. This work was supported by the Argentinean Consejo Nacional de Investigaciones Científicas y Técnicas under PID 3-140600-90.

APPENDIX

Wind-forecast error correlations derived from the Fourier–Bessel series

Using Helmholtz's theorem and assuming negligible velocity potential prediction errors, the horizontal autocorrelations for both $u$ and $v$ components and the horizontal cross-correlations between height and wind are derived. Like Buell (1972) and Lorenc (1981), I choose coordinates such that $u_i$ is the wind component in the direction from point $i$ to point $j$, and $u_i$ is perpendicular to it (positive to the left when looking from $i$ to $j$). Defining $\psi$ as the stream function, $\mu$ as the decoupling factor, considering $F_{\psi}$ as defined in (3), ignoring prediction error $E^p$ horizontal derivatives, and assuming that
\[ E^p_\phi = \frac{1}{f} E^p_\psi \]
\[ E^p_\psi = E^p_\psi \]
\[ L = \frac{E^p_\psi}{f E^p_\psi} \]
\[ F_{\psi j} = F_{\phi j} \]
\[ \text{corr}(\phi_i \psi_j) = \mu \]
\[ \text{corr}(\phi_i \mu_j) = \mu \text{corr}(\psi_i \mu_j) \]  \hspace{1cm} (A.1)

where
\[ \text{corr}(a, b) = \frac{\langle a, b \rangle}{\left( \langle E^p_\psi \rangle^2 \right)^{1/2}} \]  \hspace{1cm} (A.2)

yields the following prediction-error correlations
\[ \text{corr}(u_i \phi_j) = -\text{corr}(\phi_i \mu_j) = -L \frac{\partial F_{\phi j}}{\partial r_j} = L \sum_{n=0}^{N} A_n \frac{k_n}{D} J_1 \left( k_n \frac{r_j}{D} \right) \]
\[ \text{corr}(u_i \phi_j) = -\text{corr}(\phi_i \mu_j) = 0 \]  \hspace{1cm} (A.3)

If one transforms the correlations in (A.3) back to correlations for northward and eastward components used in the analysis scheme, this yields:
\[ \text{corr}(u_i \mu_j) = \text{corr}(u_i \mu_j) \cos^2 \theta + \text{corr}(u_i \mu_j) \sin^2 \theta \]
\[ \text{corr}(v_i \mu_j) = \text{corr}(u_i \mu_j) \sin^2 \theta + \text{corr}(u_i \mu_j) \cos^2 \theta \]
\[ \text{corr}(u_i \psi_j) = \text{corr}(v_i \psi_j) = [\text{corr}(u_i \mu_j) - \text{corr}(u_i \mu_j)] \sin \theta \cos \theta \]
\[ \text{corr}(u_i \phi_j) = -\text{corr}(\phi_i \mu_j) = -\text{corr}(u_i \phi_j) \sin \theta \]
\[ \text{corr}(v_i \phi_j) = -\text{corr}(\phi_i \mu_j) = \text{corr}(u_i \phi_j) \cos \theta \]  \hspace{1cm} (A.4)

where the angle \( \theta \) defines the local bearing of the points \( i \) and \( j \).


Possia, N., Nuñez, M. and Ciappeson, H. 1987 *Adaptación del modelo baroclinico filtrado para cinco niveles del Servicio Meteorológico de Nueva Zelanda, para uso operativo en el Centro Meteorológico Reg. B s. As. Anales del II Congreso Interamericano de Meteorología. Bs. As. 30/11-4/12 de 1987*


Seaman, R. S. and Hutchinson, M. F. 1985 Comparative real data tests of some objective analysis methods by withholding observations. *Austr. Meteorol. Mag.*, 33, 37–46


Trenberth, K. 1973 'A five layer numerical weather prediction model'. New Zealand Meteorological Service. Tech. Note 222


Vera, C., Possia, N. and Nuñez, M. 1990 Evaluación preliminar de un sistema de asimilación de datos intermitente en la región sur de Sudamérica. *Meteorológica*, 17, 27–32