Sitzungsber. Abt. II (2003) 212: 3-26

# Sitzungsberichte

Mathematisch-naturwissenschaftliche Klasse Abt. II Mathematische, Physikalische und Technische Wissenschaften

© Österreichische Akademie der Wissenschaften 2004 Printed in Austria

# Adaptation by Stochastic Down-Scaling: Preliminary Experiments for a Stochastic Limited Area Analysis and Forecasting Model

By

# Franz Huber-Pock, Christoph Kress, Georg Skoda, and Andreas Tiesner

(Vorgelegt in der Sitzung der math.-nat. Klasse am 23. Januar 2003 durch das k. M. Peter Steinhauser)

# Abstract

The idea of *down-scaling* by means of a statistical approach linking large-scale atmospheric circulation patterns with smaller scales has been applied and described on several occasions by climatologists working with GCM climate simulations using canonical correlation analysis (CCA). Following this basic idea we tried to construct a *"stochastic regional model"* with very high resolution, which is driven by a larger scale NWP-parent-model and the stochastic coupling with orography and the physical influence of the landscape. Unlike the methods of climatologists, we use statistical interpolation (*SI*) in the inverse direction to what is common in atmospheric analysis as a basic method of *down-scaling*: Analysed or forecasted variable-values are interpolated from grid-points of the NWP-model to the point of observation stations.

The "perceptible weather", which is of the highest interest to the general public, happens in the meso- $\beta$  and meso- $\gamma$  scales. Unfortunately NWP-models with very high resolution do not always produce sufficient results and time-spans in which predictability is useful is limited, especially in areas with complex topography. Therefore we wondered if it was possible to find a good adaptation of the grid-point values of the large scale NWP-model to points in the smaller scale and at the same time to transfer higher predictability in time from this large scale NWP-model into the smaller scale of the topographic complex area: The results show a good adaptation. In addition, in 1995 and 1996, the statistical parameters RMSE (root mean square error),

RV (reduction of variance) and ACC (anomaly correlation coefficient) demonstrate that for the variables surface pressure, temperature, relative humidity, and the horizontal components of the wind, a significantly higher predictability time-span can be reached for *stochastic down-scaling* than with high resolution deterministic models. Based on these results we are positive that this specific method of *SI* can be made operational in a regional model.

### Zusammenfassung

Die Idee, mit Hilfe einer statistischen Näherung großräumige atmosphärische Zirkulationsfelder mit solchen des kleineren Scales zu verbinden, wurde von Klimaforschern, die mit GCM-Modellen Klimasimulationen durchführen, mit Hilfe der Kanonischen Korrelationsanalyse (CCA) mehrfach angewendet und beschrieben. Dieser Grundidee folgend wird versucht, ein "stochastisches Regionalmodell" mit sehr hoher Auflösung zu konstruieren, welches von einem NWP-Muttermodell und der stochastischen Kopplung mit der Orographie und dem physikalischen Einfluss der Landschaft angetrieben wird. Im Gegensatz zu den Klimatologen verwenden wir als Basismethode des *down-scalings* die statistische Interpolation (*SI*) in der umgekehrten Richtung, als es in der atmosphärischen Analyse üblich ist: Analysierte oder vorhergesagte Werte werden von den Gitterpunkten des NWP-Modells an den Ort der Beobachtungsstationen interpoliert.

Das "fühlbare Wetter", welches das Hauptinteresse des Vorhersagekonsumenten findet, läuft im meso- $\beta$  und meso- $\gamma$ -Scale ab. Unglücklicherweise erbringen NWP-Modelle mit sehr hoher Auflösung nicht immer zufriedenstellende Resultate und weisen nur begrenzte Vorhersagbarkeitszeitspannen auf, insbesondere im topographisch komplexen Gelände. Deshalb stellten wir uns die Frage: Ist es möglich, eine gute Adaptierung der Gitterpunktswerte des großräumigen NWP-Modells im kleinen Scale zu finden und gleichzeitig die höhere Vorhersagbarkeit des großräumigen NWP-Muttermodells in den kleinen Scale eines topographisch komplex strukturierten Geländes zu transferieren? Die Resultate zeigen eine gute Adaptation und die statistischen Parameter RMSE (Wurzel aus dem mittleren quadratischen Fehler), RV (Reduktion der Varianz) und ACC (Anomalie-Korrelationskoeffizient) für die Jahre 1995 und 1996 demonstrieren, dass mit dem stochastischen down-scaling eine signifikant höhere Vorhersagbarkeitszeitspanne als mit deterministischen Modellen hoher Auflösung erreicht werden kann. Auf der Basis dieser Ergebnisse wird man bestärkt, diese spezielle Methode einer SI in einem Regionalmodell operationell einzusetzen.

# 1. Introduction

The modification of large-scale synoptic processes influenced by mountainous landscapes represents an essential aspect of successful weather forecasting in countries with a complex orographical structure. Many previous experiments (LORENZ, 1969; FLEMING, 1971) showed that the predictability of small scales, which influence the weather in mountainous areas, is very limited. In spite of the enormous progress of deterministic NWP-models with increasingly high space-time resolution in the last decades, the results are still not satisfying for orographical, complexly structured landscapes. Bearing in mind that the physics of the "actual" weather happens in the subsynoptic scale, it seems reasonable to aspire a local interpretation of numerical forecasting products. Commonly there are two efficient methods for local interpretation, namely "dynamical" and "statistical" adaptation. Relevant experiments are reported in literature for the first group of methods (MUSSON-GENON, 1989; MARAIS and MUSSON-GENON, 1992). Dynamical adaptation methods resemble very simplified deterministic models with extremely high local resolution. Unfortunately the results show a very limited predictability of only 6 to 12 hours. In order to work with deterministic LAMs or mesoscale models we need useful solutions for the following problems:

- a) the LBCs (Lateral boundary conditions)
- b) the initialisation and
- c) the efficient simulation of the physical processes in the small scale by relevant parameterisation.

Deterministic models with high resolution (LAMs) may nowadays and in the nearer future resolve insufficiently the weather in the small scale. Additionally, they exhibit a very limited predictability as it is demonstrated in relevant literature (LORENZ, 1969; FLEMING, 1971; STULL, 1985; FRAEDRICH and ZIEHMANN-SCHLUMBOHM, 1994).

Sceptics justifiably question whether deterministic models will make possible the simulation of the relevant dynamical and physical processes in the small scale of orographic landscapes in the near future. Additionally, there is hardly any numerical analysis scheme which is able to resolve meso- $\beta$  or meso- $\gamma$  scale structures in mountainous areas near topography (STEINACKER, HÄBERLI, and PÖTTSCHACHER, 2000).

*Down-scaling* can be considered as a method of accessing of scale structures in space and time, which are below the scales, which are resolved by numerical analysis-, weather-, or climate-forecast models.

Predictability is the capability to forecast variables or discrete weather phenomena in a practical, meaningful way in time (LJAPUNOV-time) or with a certain precision in space. In both predictability-aspects

- 1) the time aspect (LJAPUNOV-time)
- 2) the spatial precision

high-resolution NWP-models are not as successful as was hoped or expected.

Therefore we wondered: Would it be possible to find a good adaptation of the grid-point values of the large scale NWP-model on points in the smaller scale and at the same time to transfer higher predictability in time from this large scale NWP-parent-model into the smaller scale of the topographic complex area?

The idea of *down-scaling* by means of statistical approach for linking large-scale atmospheric circulation patterns with small scales has been described and applied by climatologists working with GCM climate simulations in recent times (V. STORCH et al., 1993; V. STORCH and ZORITA, 1994; GYALISTRAS et al., 1994). Climate models show useful information only for spatial scales above a threshold, which may be in the scale of several grid-distances. However, users of data from climate simulations need information on much lower spatial scales. Therefore *down-scaling* can be considered as an essential method to provide useful information. The climatologists use canonical correlation analysis (CCA) or use empirical orthogonal functions (EOF) to describe the coherent simultaneous variations between different patterns of atmospheric variables (EHRENDORFER. 1987). STORCH et al. (1993) constructed a statistical regression model between Iberian rainfalls and North Atlantic surface pressure. The method of the climatologists cited above, like our method, is based upon a *perfect prog*-approach: The statistical relationship is first established between the large scale and the local observations and then applied unmodifiedly to predict changes in the local small-scale variables from any large-scale observation, analysis or forecasted values. A somewhat similar situation can be found for weather forecasting and analysis in orographical areas, because of the small scale of the relevant dynamical and physical weather processes. We started our experiment of statistical adaptation by stochastic downscaling on the basis of such considerations. Our method is different to the method of climatologists and can be considered as statistical interpolation (SI) in the inverse direction as it is common in atmospheric analysis: We interpolate from the grid-points of the NWPmodel to the points of observation.

The basic philosophy of our stochastic adaptation can be described as follows: For every grid box of a larger scale NWP-model, which is located above a topographic structured area, and for different meteorological elements it is possible to calculate the characteristic statistical relationships (correlations) for a large sample between grid point values of the NWP forecast or analysis model and the observations situated in the grid box. Correspondingly, the analysis and forecast values at the grid points and observations, valid at the same time, are used. From the relevant correlations between different gridpoints and between grid-points and observation stations the relevant interpolation coefficients for interpolation from grid-point to observation station can be ascertained. The method can be used for a high resolution analysis or forecast, if the analysis or forecast variables are available at the grid points of the parent-model.

Meteorological forecast skill generally shows strong dependencies on spectral aspects of the atmospheric flow. The larger the scale of an atmospheric system is, the more predictable it normally is. The decrease in predictive skill starts to affect small baroclinic systems or fronts around D+2, whole cyclonic systems around D+4 and finally the long planetary waves around D + 6. These large scale systems relate to the general weather type and the systems skill in predicting them means that the forecaster will be able to make useful forecasts up to a week ahead most of the time, although the timing of cyclones or fronts might be in error (ECMWF, 1999). The deterministic predictability of atmospheric structures (BOER, 1984) is strongly dependent on the spectral wavenumber "n": The higher "n" is, the more details are described by the model but the lesser is the overall precision of the calculated results (DALCHER and KALNAY, 1987). Due to the short life-times of the inherent meso- $\beta$ -structures (BOER, 1994; STULL, 1985) deterministic meso-scale modelling usually crashes after a time span of approximately 48 hours.

We will show that the time-span of successful predictions can be enlarged by our *down-scaling method* for the small scales, too.

Our stochastic model has some major advantages:

- a) No difficulties with the LBCs. These problems do not exist for global parent-models or are solved in the larger scale parent-model itself.
- b) The initialisation is applied to the parent-model because the *stochastic model* is a product of high resolution model interpretation in the small scale. The *down-scaled* fields are not used as initial fields for the next time step of integration, because the parent-model prepares the grid-point values for the next time step. These are linked by statistical relationships to the small scale fields.
- c) A special parameterisation is not necessary, as the physical processes are implicitly included in the statistical relations. A separation of these relations according to the seasons of the year in order to consider the changes of "diabatic forcing" is possible. We think that the installation of a "learning file" for future improvements of the model is very important.

For our *down-scaling* procedure a variant of statistical interpolation (*SI*) is used and will be described in more detail in Section 2. It is not necessary to assume isotropy and homogeneity for the statistical relationships as it is usually done in conventional SI, because we always use the same grid-points and observation stations. A very dense network of observation stations is available for Austria.

The values of the variables surface pressure, 2 m-temperature, 10 m-winds and 2 m-humidity at the grid points of the parent-model have been correlated between different grid points and between the grid points and the observed values at more than 50 stations for a large sample (1993–1995) to determine the relevant covariances, correlation matrices and correlation vectors. All observation stations are situated at the ground-level  $\eta = 1$ . Unfortunately, only 2 radio-sounding stations are available for the free atmosphere. The calculation of the correlation between grid points on surface, 1000, 850 and 700 hPa and the observation stations is pursued regardless of the height above sea level, thus warranting the 3-dimensional influence of the parent-model. From the inversion of the correlation stations inside a special grid box were received as a result.

# 2. An Univariate Algorithm of Statistical Interpolation (SI) for Down-Scaling of Meteorological Elements

In the following section *stochastic down-scaling* to the coordinatelevel  $\eta = 1$  is introduced. With reference to DALEY (1991) we can write the SI-interpolation algorithm

$$(A_i - B_i) = \sum_{k=1}^{K} W_{ik} (O_k - B_k)$$
(1)

whereby

- *K* the number of available observation stations inside a certain influence radius
- $A_i$  the analysed value at the grid point *i*
- $B_i$  the "background" or "first guess" value at the gridpoint *i*
- $O_k$  the observed value at the observation station
- $B_k$  the "background" or "first guess" value at the observation station k
- $W_{ik}$  the "a posteriori" interpolation coefficient, which has to be determined.

The term in the brackets on the right side is commonly referred to as "innovation increment", the difference on the left is referred to as "analysis increment". The grid points I resemble a regular mesh of points and the points K represent irregularly distributed observation stations. In the common form of SI the  $B_k$  are mostly forecasted values of a NWP-model. In our method they are values of the climatic mean.  $B_k$  can be calculated by the so-called "forward interpolation" from grid points to the observation stations.

In our method of *down-scaling* the algorithm has to be inverted. We calculate the analysis increment at the observation stations with the help of an innovation increment  $(A_i - B_i)$  or  $(F_i - B_i)$  at the grid points of the larger scaled NWP model. The  $A_i$  are values from a larger scale objective analysis and  $F_i$  are forecasted values from a NWP model in the large scale.

The "inverse algorithm" can be written as follows:

$$F_k - B_k = \sum_{i=1}^{I} W_{ik}^* (F_i - B_i)$$
(2)

with

 $F_k$  the required forecasted value at the observation station k

 $F_i$  the forecasted value of the NWP-model at the gridpoint *i* 

*I* the number of used grid points

 $W_{ik}^*$  the interpolation coefficient of the inverse variant.

 $F_i$  can be replaced by a value of a larger scale SI analysis  $A_i$ . In this case  $F_k$  has to be exchanged by  $A_k$ .

Unfortunately the "true" value of each meteorological element is unknown, but we are convinced that the observed values, in spite of instrumental errors, are close to the true value. Therefore we assume that the "true" value is best approximated by  $O_k$ . After subtracting the observed value  $O_k$  on both sides of (2) we receive

$$F_k - O_k = B_k - O_k + \sum_{i=1}^{I} W_{ik}^* (F_i - B_i).$$
(3)

After squaring both sides of (3), applying the expectation operator  $\langle \cdots \rangle$  and finding interpolation-coefficients  $W_{ik}^*$ , which minimise the resulting relation (see the derivation in DALEY (1991)), we receive a relation in matrix form

$$\mathbf{C}\vec{\mathbf{w}}_{\mathbf{k}}^{*} = \vec{\mathbf{c}}_{\mathbf{k}},\tag{4}$$

whereby C resembles a  $I \times I$  symmetric covariance matrix,  $\vec{\mathbf{w}}_k^*$  and  $\vec{\mathbf{c}}_k$  resemble column vectors of the weights and of the covariance

between "first guess" error at the grid point *i* and the observation station *k*. A 3-dimensional grid box made up by I = 16 grid points has been used.

DALEY (1991) demonstrated that for the above algorithm (1), a normalised variant can also be derived. In analogy to this procedure one arrives after some algebra at the relation

$$\vec{\omega}_k = (\rho_B^{il})^{-1} \vec{\rho}_B^{ik} \tag{5}$$

whereby  $(\rho_B^{il})^{-1}$  is the inverse of a correlation matrix between "first guess" errors at the *I* grid-points with the elements

$$R_{il} = \frac{\langle (F_i - B_i) \cdot (F_l - B_l) \rangle}{\left[ \langle (F_i - B_i)^2 \rangle \langle (F_l - B_l)^2 \rangle \right]^{1/2}}.$$
(6)

 $F_l$  are forecasted,  $B_l$  "first-guess" and  $A_l$  analysis values at gridpoint l.  $\vec{\rho}_B^{ik}$  resembles a correlation vector between the station k and grid point i with the components

$$R_{ik} = \frac{\langle (F_i - B_i) \cdot (O_k - B_k) \rangle}{\left[ \langle (F_i - B_i)^2 \rangle \langle (O_k - B_k)^2 \rangle \right]^{1/2}}.$$
(7)

As soon as the optimal interpolation coefficients  $\vec{\omega}_k$  are known the values  $F_k$  at the observation stations can be calculated in analogy to DALEY with the normalised formula

$$F_{k} = B_{k} + \left[ \left\langle \left( O_{k} - B_{k} \right)^{2} \right\rangle \right]^{1/2} \sum_{i=1}^{I} \frac{\omega_{ik}(F_{i} - B_{i})}{\left[ \left\langle \left( F_{i} - B_{i} \right)^{2} \right\rangle \right]^{1/2}}.$$
 (8)

For the solution of (8) and the determination of the optimal interpolation coefficients  $\omega_{ik}$  it is necessary to ascertain the relevant statistics (6) and (7) for the years 1993–1995 for a sample of more than 50 observation stations for every day 12 UTC. We dealt with the calculation of the correlation matrices, which we handled according to a *perfect prog*-procedure, which is defined by WILKS (1995): "Development of *perfect prog*-regression equations, in that observed predictors are used to specify observed predictands". Therefore we always used analysis values  $A_i$  and  $A_l$  for  $F_i$  and  $F_l$  in (6) and (7), assuming that  $A_i$  and  $A_l$  are very near to the true values. This procedure however seems to be a reasonable way of finding the best statistical structures and the optimal  $\omega_{ik}$ . We calculated the statistics for every grid box of the operational ECMWF-mesh, which covers Austria. We expected that the impact of the NWP model (ECMWF)



Fig. 2.1. Grid-box with I = 16 gridpoints. Inside of the box are k observing stations on the Model level  $\eta = 1$  for evaluating the correlation matrix with elements (6) and correlation vector with components (7)

T213/L31 or T319/L60) would be reproduced by the formula (8) and the modifications by the orographic landscape would be reproduced by the correlation vector of equation (5) with the components (7) (see Fig. 2.1).

We intended to separate the statistics according to the seasons of the year to examine the changes of the diabatic forcing. The installation of a "learning file" for future improvement is highly recommended. In our method the values  $A_k$  or  $F_k$  can be *down-scaled* exactly only to locations, of which observation values are available. We consider the fact that the interpolation weights must be calculated once and for all as a great advantage, contrary to the common *SI*, where observation stations are only temporarily available. In our procedure the same data points (grid points) can always be used in equation (8).

## 3. Results

## 3.1. Observations and Down-Scaled Variable Values

We calculated *down-scaled values* for all days 12 h UTC of the years 1995 and 1996. In the following demonstration only the values of January and July 1995 are shown as examples for the surface pressure (hPa) and the 2 m relative humidity for station Vienna, Hohe Warte (11035), the 2 m temperature of the stations Innsbruck (11120) and



comparison observation - analysis, station 11035



number of day, 1995

Fig. 5.1. Surface pressure, January and Jury 1995 for Vienna-Hone watte

Patscherkofel (11126) and the components of the 10 m horizontal wind for the station Patscherkofel (11126).

In the following Figs. 3.1–3.5 the abscissa indicates the day in the year and the ordinate indicates the air pressure in hPa, the relative humidity in %, the temperature in degrees Celsius and the wind components in m/sec. The solid curve indicates the measured value of the element in question at the observation-station (Synop), the broken line indicates the *down-scaled* value and the dotted line is the result of simple linear interpolation from the four nearest gridpoints of the ECMWF-model to the observation-station. In Fig. 3.1 the results of *down-scaling* the sea level pressure to the observation-station Vienna. Hohe Warte, is demonstrated.

Figure 3.2 shows the measurements of 2 m relative humidity for Vienna, Hohe Warte, in comparison with *down-scaled* values. Here

#### Adaptation by Stochastic Down-Scaling



comparison observation - analysis, station 11035





Fig. 3.2. Relative humidity for Vienna Hohe Warte. January and July 1995

the results are not as good as with surface pressure. This is not surprising since the distribution of relative humidity can be considered as more or less discontinuous.

In Figs. 3.3 and 3.4 one can see measured and down-scaled temperatures for Innsbruck and Patscherkofel. The agreement of measurements and *down-scaling* is very good on condition that it is in an area with very complex topography. In Figs. 3.5 and 3.6 one will find the down-scaled adaptations for the 10 m horizontal wind components u and v.

The down-scaling of the components of the wind is not as successful compared to temperature and pressure, but generally a satisfactory adaptation can be obtained.

To ensure that statistical interpolation (SI) is the right algorithm for our down-scaling method we compared SI-results with the

comparison observation - analysis, station 11120





Fig. 3.3. Temperature for Innsbruck. January and July 1995

results of simple linear interpolation. For an estimation of the variance shares to explain the variation of the time series and the absolute and relative deviation between *down-scaled* and measured values see Table 3.1 below. It indicates a significant improvement of the punctual *down-scaled* forecast-values when using *SI* as opposed to linear interpolation. *SI* is significantly better in nearly all cases.

# 3.2. Statistical Evaluation

The statistical part contains more data in comparison to Chapter 3.1., e.g. from the time span 1991 to 1995 and independent prognoses for the year 1996.

#### Adaptation by Stochastic Down-Scaling



comparison observation - analysis, station 11126

- Synop

182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 number of day, 1995

----- inear interpolation

# 3.2.1. Characteristic Measurements

The table below contains a range of basic statistical parameters. In this table, the prognosis by means of statistical interpolation (SI) is compared to the linear interpolation (Lin). The statistical measured numbers refer to the mean values from about 60 stations.

# 3.2.2. Comparison of Different Forecast Variants

3.2.2.1. RMSE (Root Mean Square Error) in the Groups of Data for the Calculation of Interpolation Coefficients (Omega Values)

The RMSE indicates the average error of a prognosis as compared to the actual measured value. When calculating prognoses for forecast lead times of 24 to 240 hours, using the method of down-scaling

down-scaled value Fig. 3.4. Temperature for Patscherkofel. January and July 1995



Fig. 3.5. Horizontal 10 m wind component u for Patscherkofel. January and July 1995

presented above, the initial data, i.e. the data of measurement stations can be calculated. The RMSE values have been determined for the following variants:

- 1: Calculation of the prognosis with coefficients (omega value) from the entire set of data
- 2: Calculation of the prognosis separated in summer and winter coefficients
- 3: Calculation of the prognosis just with summer coefficients
- 4: Calculation of the prognosis just with winter coefficients
- 5: Calculation of the prognosis with a combination of summer and winter coefficients with weighting between the main season coefficients by a moving average procedure for spring and autumn

The following serve as reference values:

- ref-clim: Calculation of the prognosis only from climate values
- ref-lin: Calculation of the prognosis with linear interpolation

comparison observation - analysis, station 11126



comparison observation - analysis, station 11126

Fig. 3.6. Horizontal 10 m wind component v for Patscherkofel. January and July 1995

After evaluating the RMSE values of the different variants, it becomes obvious that the error increases as expected with increased time steps. The reference prognosis by means of linear interpolation from the gridpoints does, however, show the same trend, so that this behaviour can be considered as a property of the basic NWP model. It is, however, essential that on the whole, the total error is smaller with *down-scaling*, whereby the most useful option is the weighted prognosis, i.e. the one using a combination of summer and winter coefficients.

# 3.2.2.2. RMSE (Root Mean Square Error) in Relation to Different Time Spans of the Calculation of Interpolation Coefficients (Omega Values)

An examination of the RMSE for the calculation of the coefficients in relation to different starting time spans shows that the time span has got little impact on the RMSE values. In the example above, a

Table 3.1. Statistica humidity	l parameters (me and air pressure	ean, minimum, e. The calculate	maximum, stand d numbers descr	ard deviation, v ibe the differer	/ariance) when concertion values observing	omparing SI wil vation minus pr	th <i>Lin</i> for temper ognosis	ature, relative
	Mean of all differences observation minus prognosis <i>Lin</i>	Mean of all differences observation minus prognosis <i>SI</i>	Minimum of all differences observation minus prognosis <i>Lin</i>	Minimum of all differences observation minus prognosis SI	Maximum of all differences observation minus prognosis <i>Lin</i>	Maximum of all differences observation minus prognosis SI	Standard- deviation of all differences observation minus prognosis <i>Lin</i>	Standard- deviation of all differences observation minus prognosis <i>SI</i>
Temperature °C								
Mean	-1.4	-0.6	-13.5	-8.1	6.7	6.9	3.2	2.2
Mininum	-6.1	-1.6	-25.0	-13.5	1.2	3.5	2.2	1.4
Maximum	11.2	0.6	-2.5	-3.5	19.6	10.2	5.4	3.2
Standard-deviation	3.1	0.5	3.2	2.2	3.4	1.5	0.7	0.4
Variance	9.3	0.2	10.1	5.0	11.3	2.4	0.4	0.1
Relative Humidity ?	%							
Mean	5.7	-1.4	-29.3	-36.7	49.8	32.9	13.8	12.0
Mininum	-8.8	-11.5	-52.5	-58.8	28.5	19.4	9.6	8.5
Maximum	26.8	5.9	-13.6	-24.8	74.5	52.5	24.5	16.3
Standard-deviation	7.5	3.5	8.5	6.8	13.0	7.5	3.2	1.9
Variance	56.5	12.5	76.6	46.0	168.2	55.5	10.1	3.6
Air pressure hPa								
Mean	87	0.7	81.2	-3.3	96.8	5.5	2.4	0.9
Minimum	16.3	-0.7	13.2	-20.7	19.2	1.5	0.8	0.5
Maximum	323.2	6.9	305.8	3.8	348.8	9.0	6.9	3.8
Standard-deviation	72.4	1.2	70.8	4.7	77.2	8.8	1.5	0.5
Variance	5242.3	1.3	5018.4	21.7	5965.3	76.6	2.2	0.3

18

F. Huber-Pock et al.



Fig. 3.7. Comparison of different mean-values of RMSE (references ref-clim and ref-lin, 1 to 5) for the temperature (°C) in 1996

period of three years turns out to be most favourable for the calculation of the coefficient for the year 1996.

For the year 1995, not shown here, a four year period was more favourable. In general, we think that a four year period is the most favourable time span. The differences mentioned above can be explained by an ECMWF-model change which took place in 1996, and by that the prognosis was changed since it depended on the basic parent-model. From Fig. 3.8 it also becomes clear that the differences





- 1: Prognosis of the temperature of the year 1996 with coefficients (omega values) from 1991 to 1995
- 2: Prognosis of the temperature of the year 1996 with coefficients (omega values) from 1992 to 1995
- 3: Prognosis of the temperature of the year 1996 with coefficients (omega values) from 1993 to 1995
- 4: Prognosis of the temperature of the year 1996 with coefficients (omega values) from 1991 to 1994

of the RMSE values of different starting time spans decreases with increasing forecast lead times. In the next paragraph we continue with some predictability considerations.

#### 3.3. Predictability Considerations

In this section we demonstrate different error measures or skill-scores of the predictability of our methodology to answer the question, if it is possible to transfer higher predictability in time from a large scale NWP-parent-model into the smaller scale of a topographic complex area by a *down-scaling process*.

# 3.3.1. Reduction of Variance (RV)

The reduction of variance is a combination of RMSE values of a prognosis that has to be tested and a reference prognosis:

$$RV = 1 - \left(\frac{RMSE_{PROG}}{RMSE_{ref}}\right)^2$$
(9)



Fig. 3.9. Comparison of forecasts for different prognostic elements for 1996

- 1: Prognosis of the temperature of 1996 with coefficients (omega values) from 1991 to 1995
- 2: Prognosis of the relative humidity of 1996 with coefficients (omega values) from 1991 to 1995
- 3: Prognosis of the air pressure of 1996 with coefficients (omega values) from 1991 to 1995
- 4: Prognosis of the wind component u of 1996 with coefficients (omega values) from 1991 to 1995
- 5: Prognosis of the wind component v of 1996 with coefficients (omega values) from 1991 to 1995

The RV value in that case is positive, if the tested prognosis is better than the reference prognosis, and negative vice versa. In this case the *down-scaling method* is compared to a reference prognosis from climate values. An improvement for the different elements, such as temperature ( $t_2$ ), relative humidity ( $rh_2$ ), air pressure (ps), u-component of the wind (10 u) and v-component of the wind (10 v) is of varied quality compared to a prognosis from the climate values. Prognoses from *stochastic down-scaling* are clearly improved with temperature and air pressure, but the improvement with other elements is smaller.

The image also shows the general decrease of prognosis quality for longer forecast lead times.

# 3.3.2. The Anomaly Correlation Coefficient (ACC)

The ACC measures the correlation between forecast and analysed deviations from climatology. It evaluates the similarity between prognosis results of different time steps. If the prognosis is correct, the ACC is one, or 100%.

$$ACC = \frac{(\overline{F_j - C})(O_j - \overline{C})}{\left\{ \left[ (\overline{F_j - C})^2 - (\overline{F_j - C})^2 \right] \left[ (\overline{\overline{O_j - C}})^2 - (\overline{O_j - C})^2 \right] \right\}^{0.5}}$$
(10)

- $F_i$  = prognostic value at observation station k or gridpoint i
- $O_j = observed$  value at observation station k or analysed value at gridpoint i
- C = climatic mean at observation station k or gridpoint i

In the following, the averaged ACC values of the ECMWF prognoses (referred to as ECMWF all), the averaged ACC values of the ground values of the model (referred to as ECMWF SFC) and the values from *stochastic down-scaling* (referred to as Synop all) are being compared on the  $4 \times 55$  gridpoints of the grid for 1996.

The ACC of the average values of the *down-scaling* prognoses of all synop stations is close to the ACC values of the ECMWF data. It becomes clear that the predictability is closely linked to the model the *down-scaling* is based upon. The loss of predictability of *stochastic down-scaling* in face of the ECMWF parent-model for temperature is only marginal.



Fig. 3.10. Comparison of ACC-values for ECMWF-forecasts and *down-scaling* forecasts of 2 m surface temperature (forecast-period 1996)

We calculated the ACC for surface pressure (Fig. 3.11), for 2 m relative humidity (Fig. 3.12) and for the components of the 10 m horizontal wind (Fig. 3.13 and Fig. 3.14). As limit of practical predictability a value of ACC = 0.6 is commonly accepted. The best result we found



Fig. 3.11. Comparison of ACC-values for ECMWF-forecasts and *down-scaling* forecasts for surface air pressure (forecast-period 1996)



Fig. 3.12. Comparison of ACC-values for ECMWF-forecasts and *down-scaling* forecasts for 2 m relative humidity (forecast-period 1996)



Fig. 3.13. Comparison of ACC-values for ECMWF-forecasts and *down-scaling* forecasts for the 10 m horizontal wind component *u* (forecast-period 1996)

with temperature. The predictability is very good and expands over 10 days for the parent-model and for *down-scaled* forecasts. The loss of predictability in comparison to the NWP-model is only marginal.



Fig. 3.14. Comparison of ACC-values for ECMWF-forecasts and *down-scaling* forecasts for the 10 m horizontal wind component v (forecast-period 1996)

As usual the value of ACC = 0.6 is reached for surface pressure at about 7 days. Our *down-scaled* forecasts reach this value in approximately the same time span. It is interesting that the ECMWF-values are less satisfactory. At first this is surprising, but we checked our calculations and data very thoroughly and found the results to be valid. One possible cause could be that the ECMWF-model topography is not quite optimal, but this has no special impact on the statistical relationships from gridpoint to observation station. We also found the same good results for *down-scaling* the horizontal components of the wind. Last but not least, the relative humidity shows that the practical predictability ends for the *down-scaled* forecasts at about 96 hours. Here the parent-model is much better with a predictability-time of 144 hours. However, in any case the value of 96 hours is much better than the predictability-time of most high-resolution deterministic models.

# 4. Conclusions

After obtaining these results of our experiments, we can be sure that statistical adaptation by *stochastic down-scaling* is useful. Our forecast experiments for the complete years 1995 and 1996 confirm this conclusion. Additionally our question in reference to the transfer of higher predictability from an NWP-parent-model into the smaller scale of an area with complex topography can be answered positively. Therefore we are encouraged to continue our examinations to turn *stochastic down-scaling* in an operational model. Cloudiness and precipitation should also be included in a next phase.

In order to make not only point-forecasts, but also forecasts of *down-scaled* field distributions on a high resolution subgrid appropriate to the complex topography, we will try to incorporate the high resolution analysis model VERA (PÖTTSCHACHER, STEINACKER, and DORNINGER, 1996) in the future. This will be possible, as soon as appropriate variable values are *down-scaled* from the parent-model of ECMWF for the observation points over Austria and the countries surrounding the Alps. After application of the VERA-analysis process high resoluted prognostic fields will result.

#### Acknowledgement

The research was supported by the Austrian "Fonds zur Förderung der wissenschaftlichen Forschung" (FWF) under contract P10556-TEC, the Central Institute for Meteorology and Geodynamics in Vienna and the Institute for Meteorology and Geophysics of the University of Vienna, Austria. A lot of useful criticism and propositions of two anonymous reviewers are gratefully acknowledged.

# References

- [1] BOER, G. J. (1984) A spectral analysis of predictability and error in an operational forecasting system. Mon. Wea. Rev. **112**: 1183–1197
- [2] BOER, G. J. (1994) Predictability regimes in atmospheric flow. Mon. Wea. Rev. 122: 2285–2295
- [3] DALCHER, A., KALNAY, E. (1987) Error growth and predictability in operational ECMWF forecasts. Tellus 39A: 474–491
- [4] DALEY, R. (1991) Atmospheric Data Analysis. Cambridge University Press, Cambridge
- [5] ECMWF (1999) User Guide to ECMWF Forecast Products. In: § 3.1. Reading
- [6] EHRENDORFER, M. (1987) A regionalisation of Austria's precipitation climate using principal component analysis. J. Climatol. 7: 71–89
- [7] FRAEDRICH, K., ZIEHMANN-SCHLUMBOHM, C. (1994) Predictability experiments with persistence forecasts in a red-noise atmosphere. Q.J.R. Met. Soc. 120: 387–428
- [8] FLEMING, R. J. (1971) On stochastic dynamic prediction II: Predictability and utility. Mon. Wea. Rev. 99: 927–938
- [9] GYALISTRAS, D., V. STORCH, H., FISCHLIN, A. (1994) Linking GCM-simulated climatic changes to the ecosystem models; Case studies of statistical downscaling in the Alps. Clim. Res. 4: 99, 167–189

- 26 F. Huber-Pock et al.: Adaptation by Stochastic Down-Scaling
- [10] LORENZ, E. N. (1969) The predictability of a flow which possesses many scales of motion. Tellus 21: 289–307
- [11] MUSSON-GENON, L. (1989) Forecasting in the vertical with a local dynamical interpretation method. Mon. Wea. Rev. 117: 29–39
- [12] MARAIS, CH., MUSSON-GENON, L. (1992) Forecasting the surface weather elements with a dynamical adaptation method using a variational technique. Mon. Wea. Rev. **120**: 1035–1049
- [13] PÖTTSCHACHER, W., STEINACKER, R., DORNINGER, M. (1996) VERA A high resolution analysis scheme for the atmosphere over complex terrain. In: Preprints 7<sup>th</sup> Int. Conf. on Mesoscale Processes. Amer. Meteor. Soc., Reading, UK, pp. 245–247
- [14] STEINACKER, R., HÄBERLI, C., PÖTSCHACHER, W. (2000) A transparent method for the analysis and quality evaluation of irregularly distributed and noisy observational data. Mon. Wea. Rev. 128: 2203–2316
- [15] V. STORCH, H., ZORITA, E., CUBASCH, U. (1993) Down-scaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. J. Climate 6: 1161–1171
- [16] V. STORCH, H., ZORITA, E. (1994) Down-scaling and weather generators: The problem of inferring user-oriented information from Climate Model Output. In: 6th Conf. on Climate Variations, Nashville, 1994. Amer. Met. Soc., Boston
- [17] STULL, B. R. (1985) Predictability and scales of motion. Bull. Amer. Met. Soc. 66: 432–436
- [18] WILKS, D. S. (1995) Statistical Methods in the Atmospheric Sciences. Academic Press, London

Authors' addresses: Prof. Dr. Franz Huber-Pock, Prof. Dr. Georg Skoda, Mag. Andreas Tiesner, Institute for Meteorology and Geophysics, University of Vienna, Althanstraße 14, 1090 Vienna, Austria; Dr. Christoph Kress, Central Institute of Meteorology and Geodynamics, Hohe Warte 38, 1190 Vienna, Austria.