EFFECTS OF ASSIMILATION OF SURFACE OBSERWATIONS ON 24-HOUR LM FORECAST OVER POLAND

Joanna Linkowska
Blazej Krzeminski

Introduction

The aim of the study was to verify the hypothesis that spatial preprocessing of synoptic data prior to application of their assimilation by nudging technique may result in better forecast. The standard LM forecast was verified against synoptic data from selected stations and in selected areas of Poland. Bias and RMS error of this forecast was compared with corresponding errors of LM model runs with point nudging of synop observations of surface pressure, dew point, 2m temperature. Then the data from all the stations on the Poland's territory were optimally interpolated to regular mesh corresponding to LM domain and used as the super observations for nudging purposes. The time window of nudging was -1 hr to + 3 hr in both cases. The effects of point and OI interpolated synop data nudging on forecast error are discussed.

Settings

Test was held during the period 01.08.2003 – 21.08.2003. We set up a small domain centered over Poland with 14km grid spacing. The rotated coordinates of the lower left corner of the domain are $\lambda = -8.5°$, $\varphi = 0.625°$. Coordinates of the upper right corner are $\lambda = 0.125°$, $\varphi = 10.5°$. Each day 3 model runs were performed:

1. Forecast using initial state and boundary conditions interpolated from GME.
2. Forecast using initial state from GME, boundary conditions from GME and nudging to the synop stations over Poland during first 3.5 hours of model run.
3. Forecast using initial state from GME, boundary conditions from GME and nudging to the regular fields generated by the Optimal Interpolation of synops (described below).

We used synoptic data from –1,0,+1,+2,+3 hours relative to model initial time.

Model Configuration:

Domain size:	80 x 70 gridpoints
Horizontal Grid Spacing:	0.125° (~14 km)
Number of Layers:	35
Time Step and Integration Scheme:	80 sec., 3 time level split-explicit
Forecast Range:	24 h
Initial Time of Model Runs:	00 UTC
Lateral Boundary Conditions:	Interpolated from GME, at 3h intervals
Initial State:	Interpolated from GME
Model Version Running:	lm_f90 2.19
Hardware:	SGI ORIGIN 3800, using 4 processors

Default values of GME and LM control variables were used, except of those related to the nudging.

Optimal Interpolation

Optimal interpolation (OI) used by the authors is a classical OI, which is linear and optimal in the minimal RMS error sense. Interpolated values of a given field $f_0(x, y)$ at a point with coordinates (x, y) are calculated using the linear combination of real data from n measurement sites (synop stations):

$$f_0 = \left(1 - \sum_{i=1}^{n} p_i\right)\langle f_0 \rangle + \sum_{i=1}^{n} p_i f_i ,$$

where : f_i - real values at the stations, p_i - corresponding weights determining the contributions from f_i to the values at given grid point - f_0, $\langle f_0 \rangle$ jest the mean value at the location (x, y).

If we normalize the weights

$$\sum_{i=1}^{n} p_i = 1$$

the evaluation of $f_0(x, y)$ requires the estimation of corresponding weights p_i, which are found from the RMS error minimalization criterion.

$$E = \overline{\left(f_0' - \sum_{i=1}^{n} p_i f_i'\right)^2}$$

where f_0' is the deviation of f_0 from the mean.

If we assume that the errors are not correlated with the observation data, and the mean observation errors are uncorrelated and can be neglected, by differentiating the above equations against the weights p_i we obtain the set of equations in the form:

$$\sum_{j=1}^{n} \mu_{ij} p_j + \lambda_i = \mu_{0i} \qquad i = 1,...,n$$

where: μ_{0i} are the correlations coefficients of the field f_0 and the value of this field at the station i, μ_{ij} are the correlation coefficients between field values at the stations i and j, λ_i are so called Lagrange'a multipliers, the requirement of weights normalization closes this set of (n+1) equations.

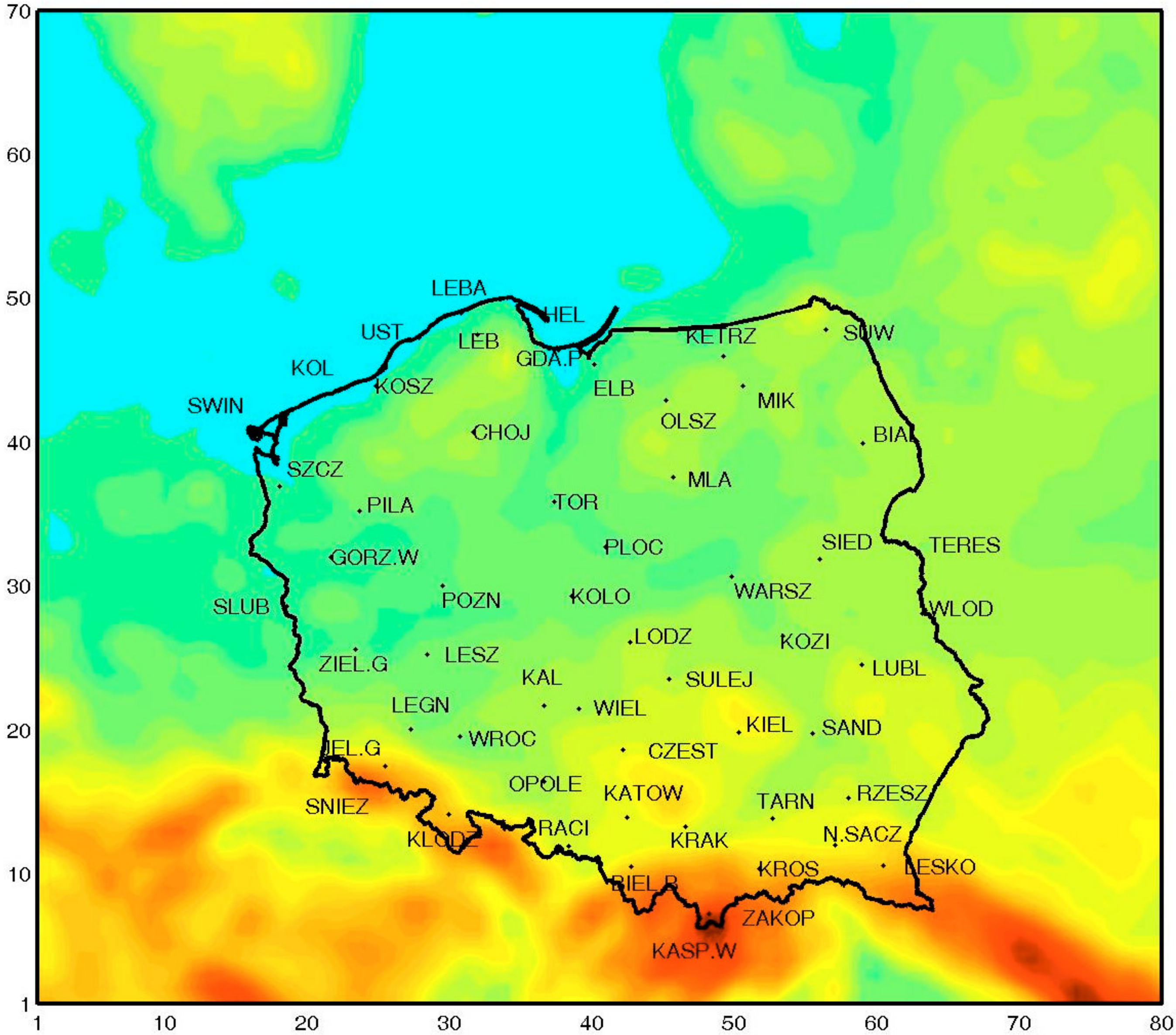
The critical stage of OI is the selection of the adequate shape of the correlation function μ as the function of ρ that is the distance between the stations. Basing on best fit to the long term mean values of correlation between the stations, the following empirical best fit exponential relation was used:

$$\mu(\rho) = (1 - a\rho)\exp(-a\rho)$$

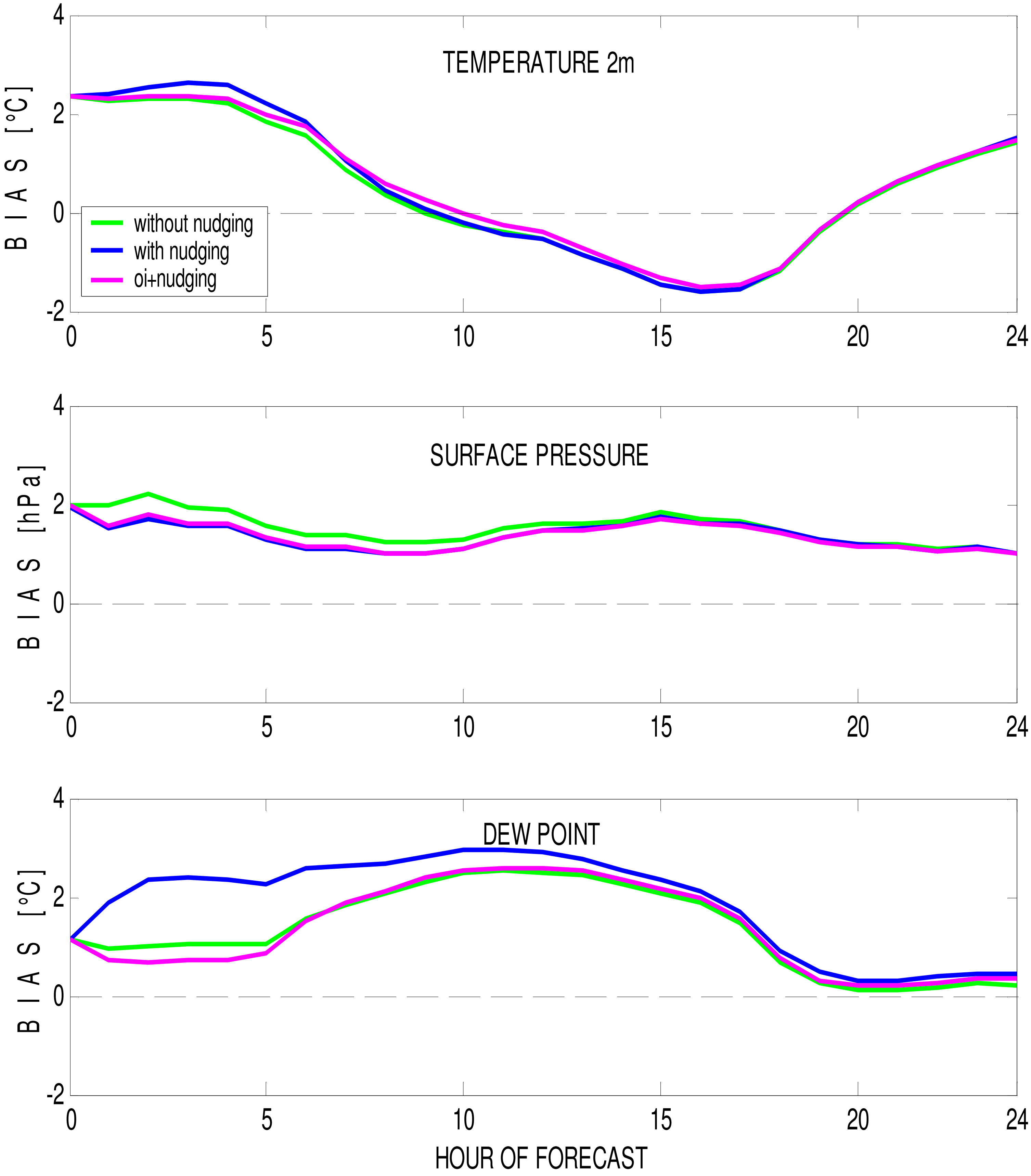
where a is empirical coefficient with the dimension of inverse of the distance between the points under consideration.

In our study we assume the correlation function to have a form given by the formulae:

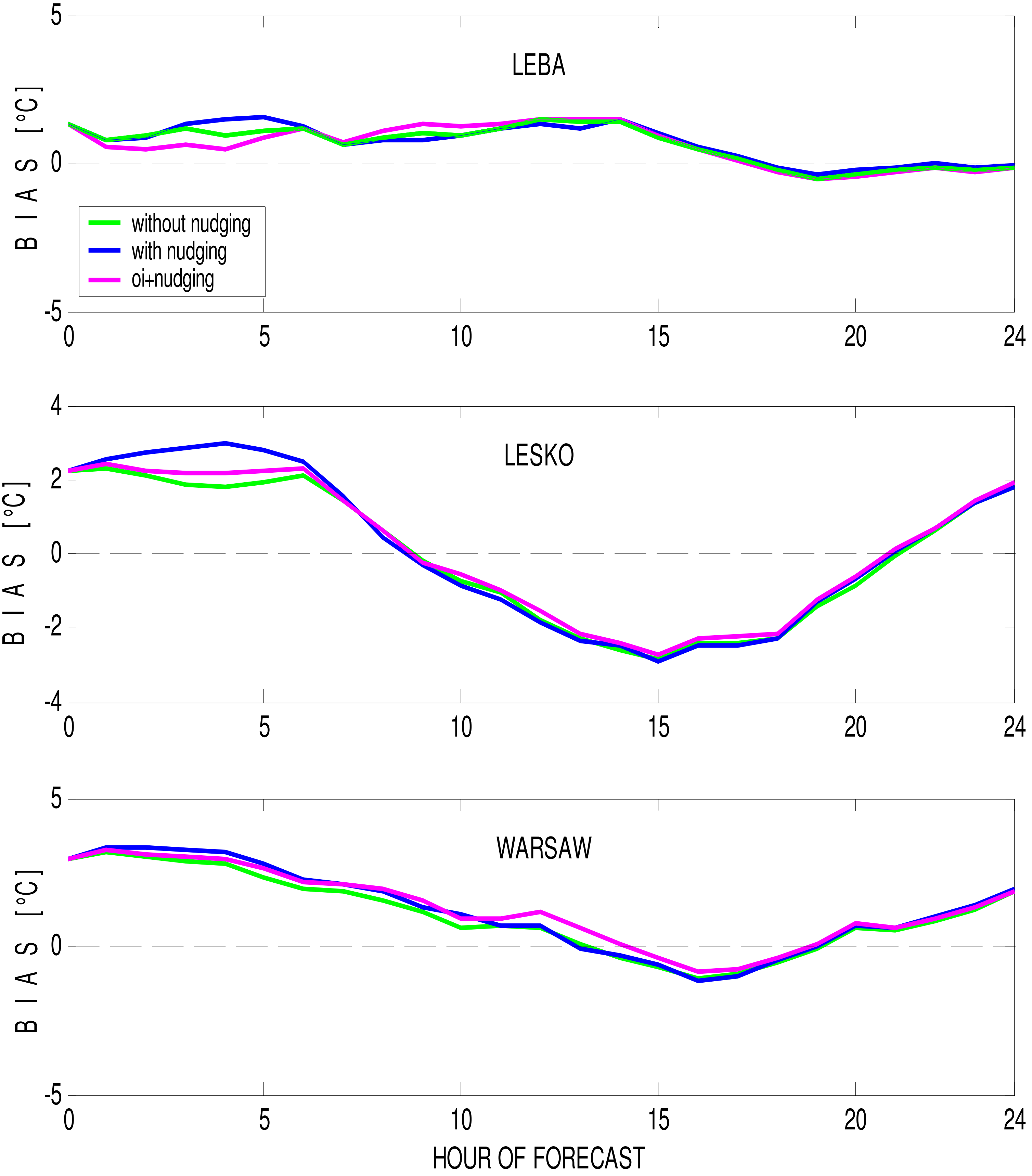
$$\mu(\rho) = \sum_{n=1}^{n=6}\left(\frac{(-\rho)^n}{n!}\exp(\rho)\right)$$



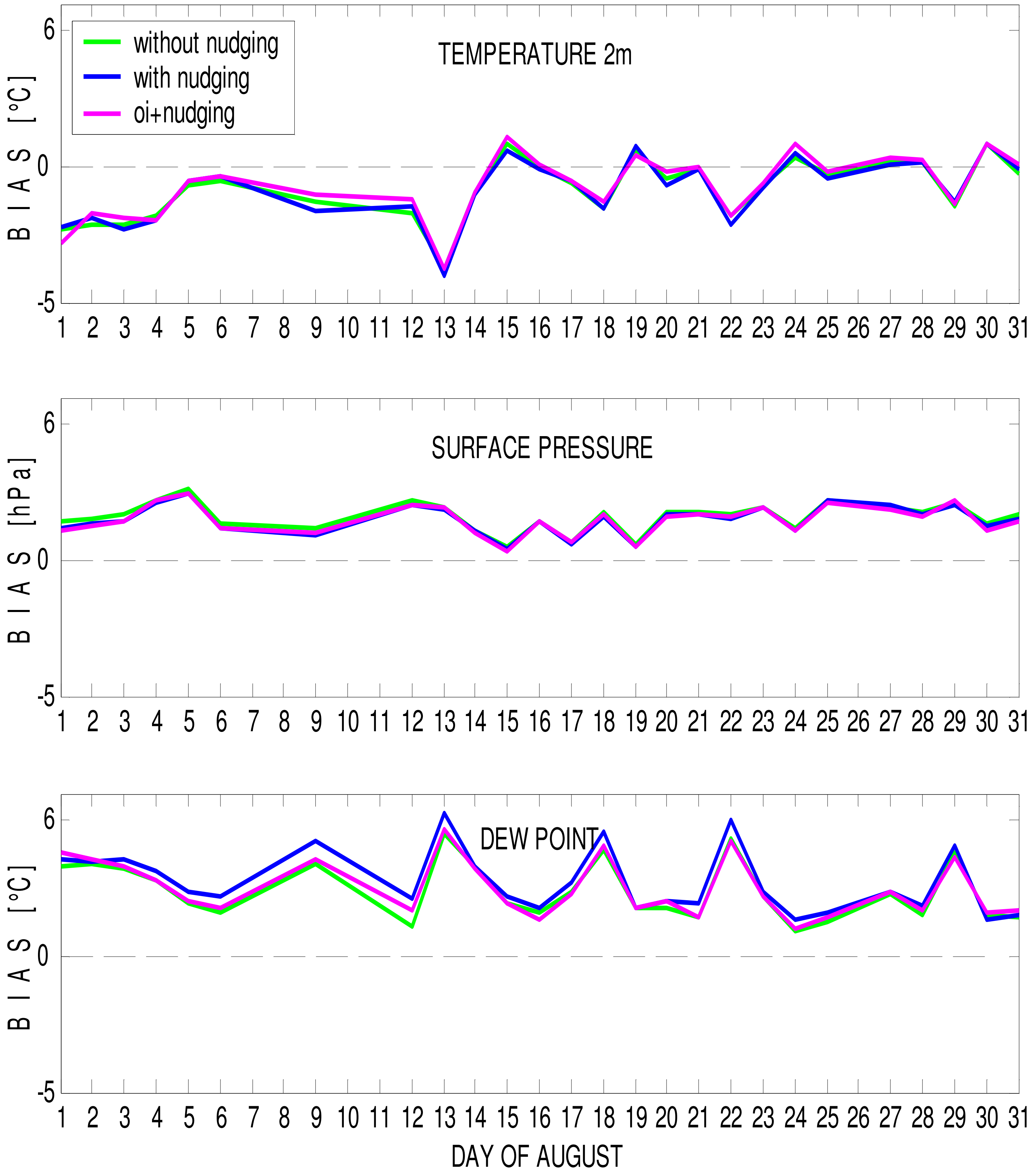
LOCATIONS OF POLISH SYNOP STATIONS



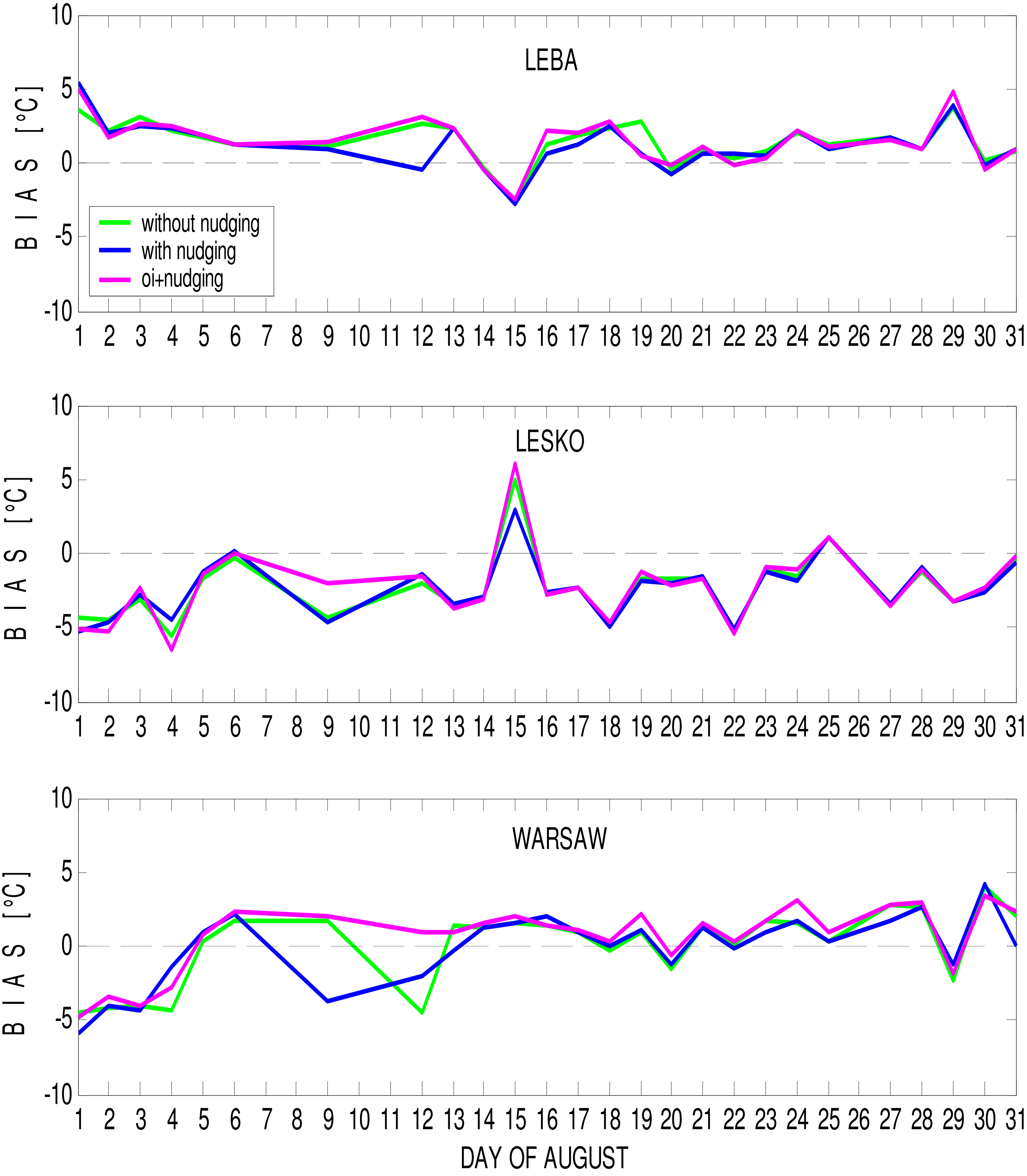
MEAN DIURNAL CYCLE OF BIAS FOR POLAND
(01.08.2003 –31.08.2003)



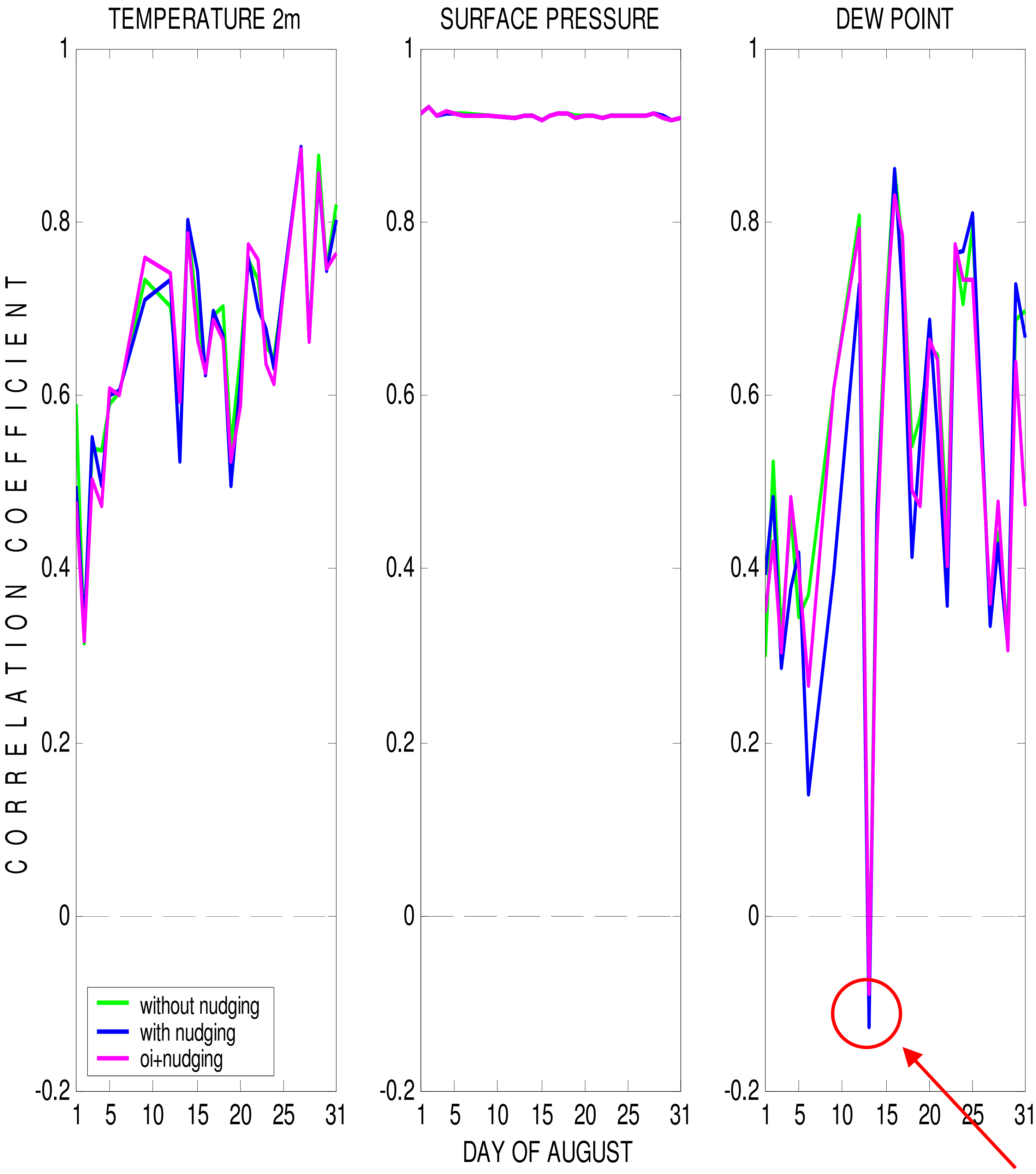
MEAN DIURNAL CYCLE OF BIAS OF TEMPERATURE AT 2m
FOR THREE SELECTED STATIONS (FORECAST VS. STATIONS)



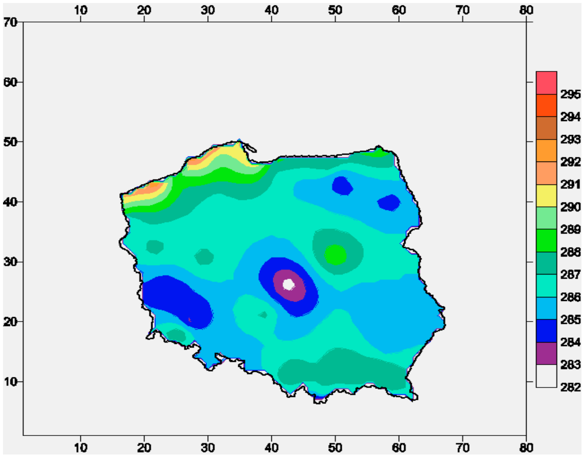
TIME COURSE OF TEMPERATURE, SURFACE PRESSURE
AND DEW POINT AT 13:00 UTC (AVERAGED FOR POLAND)



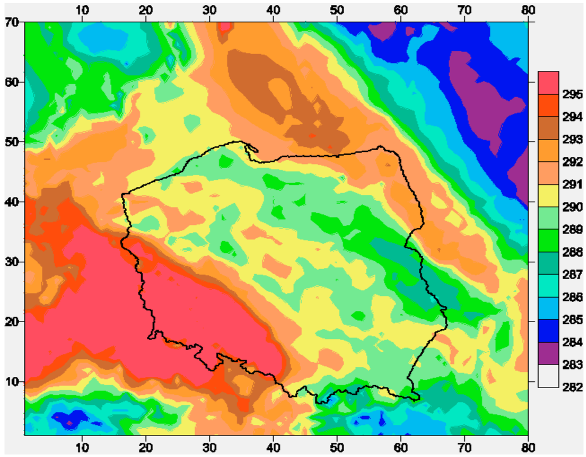
TIME COURSE OF TEMPERATURE AT 2m (13:00 UTC)
FOR THREE SELECTED STATIONS



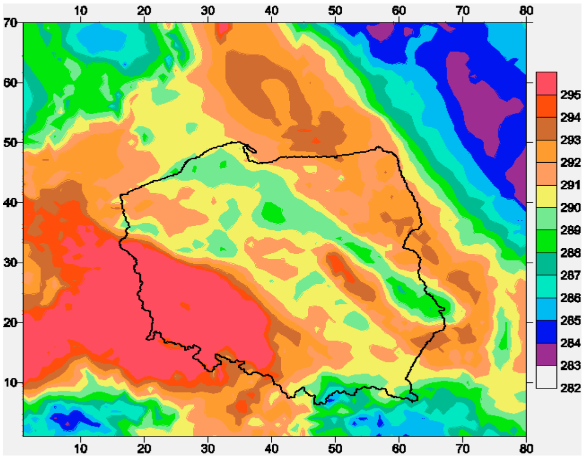
CORRELATION COEFFICIENT (FORECAST VS. OBSERVATIONS)



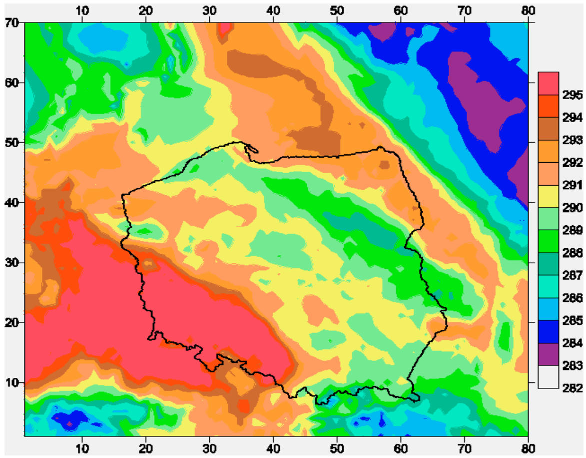
2m dewpoint temperature, interpolated from synops



2m dewpoint temperature, forecast without nudging

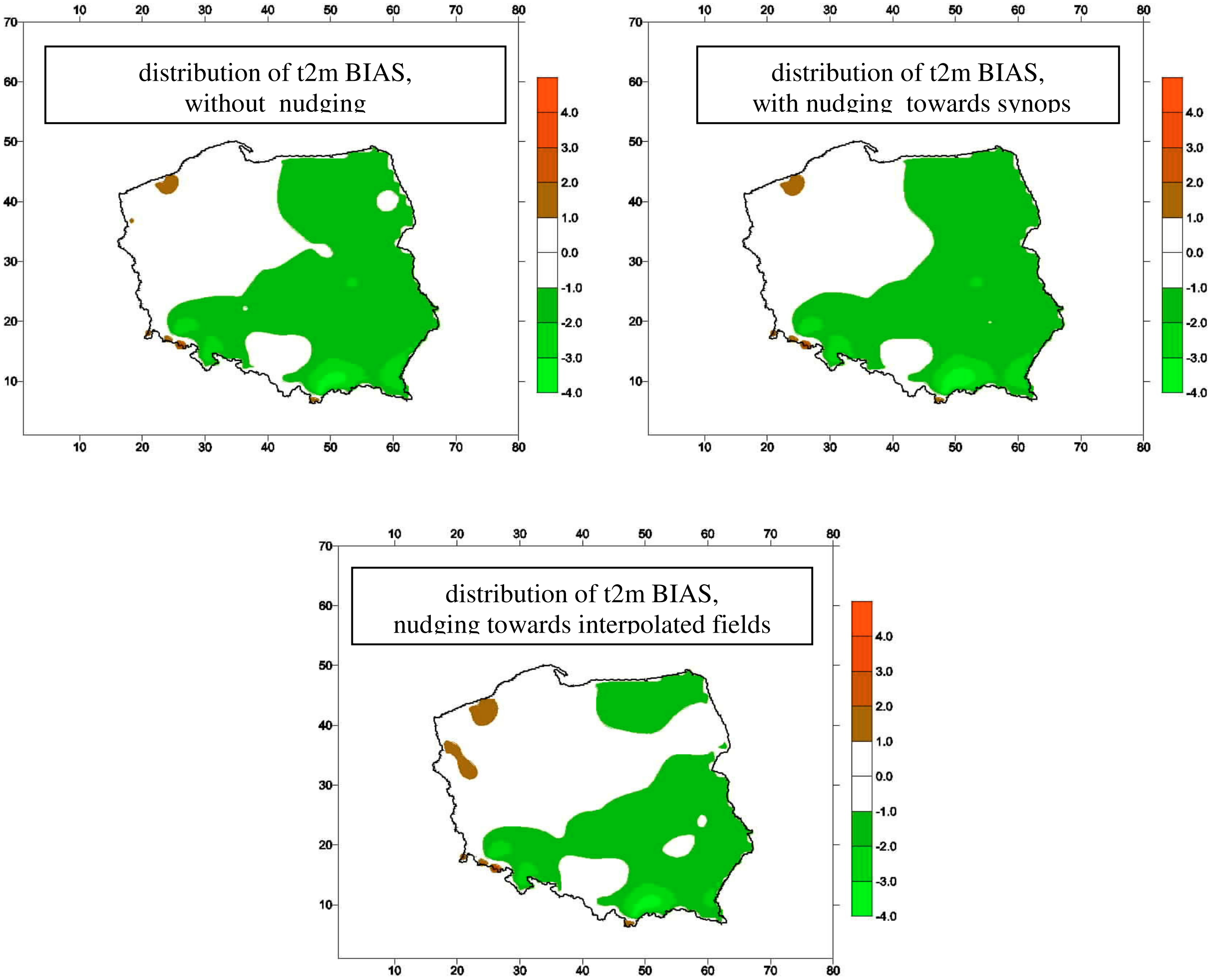


2m dewpoint temperature, forecast with nudging towards synops



2m dewpoint temperature, forecast with nudging towards interpolated fields

The case of a cold front over Poland, 13.08.2003 13h,
corresponds to the deep in corelation coefficient marked above.



Spatial distribution of averaged forecast errors for 2m temperature at 13th hour. (Biases for each station were averaged over period 01.08.2003 – 31.08.2003).

This picture shows that LM tends to systematically underestimate forecast over highlands. Slight improvement can also be seen when nudging towards interpolated fields is used.

Conclusions:

1. Nudging of synop data unexpectedly resulted in the deterioration of the forecast quality especially in the first few hours of the forecast.

2. Application of optimal interpolation prior to the assimilation gives more accurate forecasts than the nudging towards synops itself.

3. The analysis of mean diurnal bias cycles revealed that the errors reach maximal values at the 13th hour of the forecast and these errors were analyzed in detail for a period of 30 days.

4. The LM model generates better forecasts for the coastal and lowland areas then for the highlands.

5. The deep in the correlation index values which occurred on 13th of August was due to the dew point temperature forecast corresponding to the passage of the cold front across Poland.

6. Hence, the precipitation forecasts for the corresponding time window will be verified in future studies.