Calibrated Probabilistic Mesoscale Weather Field Forecasting: The Geostatistical Output Perturbation (GOP) Method

Yulia Gel,

Department of Statistics, George Washington University

Joint work with Adrian Raftery and Tilmann Gneiting,

Department of Statistics, University of Washington

Supported by the DOD ONR Multidisciplinary University Research Initiative (MURI)

Criteria

- **Calibrated forecasts:**
  - **if we define a probability interval, such as a 90$\%$ probability interval, then on average in the long run, 90$\%$ of such intervals contain the true value;**
- **Sharp forecasts:**
  - **if the distribution is more concentrated than the forecast distribution from climatology alone.**

Outline

1. The basic idea
2. Data analysis
3. Direct method for simulations of statistical forecast ensembles
4. Verification statistics
5. Future development
6. Combining global and local grid-based bias correction methods

$$\boxed{X \rightarrow M \rightarrow Y}$$

X are model inputs; Y are values to be forecast,

M is a mesoscale numerical prediction model (MM5)

(MM5 model is developed at the National Center for Atmospheric Research, Boulder, CO, USA, http://www.mmm.ucar.edu/mm5/mm5-home.html,)

Available Data

- **Current initial conditions from a big synoptic model on a grid, $\tilde{X}$.**
- **Historic information on $\tilde{X}$, also on a grid.**
- **Current MM5 forecasts on a grid, based on $\tilde{X}$, $\tilde{Y}$.**
- **Current direct observations, irregularly spaced, $\dot{X}$.**
- **Historic information about direct observations at stations, irregularly spaced, $\hat{X}$.**
- **Historic information on MM5 forecast $\tilde{Y}$, on a grid, $\hat{Y}$.**

The basic approach

Our main framework will be that of **Bayesian melding** (*Poole and Raftery, 2000, JASA*). This provides a way of combining up to 4 sources of information:

- prior information about inputs;
- prior information about outputs;
- likelihoods (data) about inputs;
- likelihoods (data) about outputs;

Our goal is to produce calibrated probabilistic forecasts of the kind of two-dimensional images that operational forecasters look at, with the whole image being calibrated, rather than just the individual forecasts that make it up.

Basic idea: For now, we focus only on a single predictive distribution of Y, $q(Y|\tilde{Y})$, and generate **statistical ensembles of forecasts** from $q(Y|\tilde{Y})$.

Alternative: Form a single ”prior” distribution, $\pi(X|\tilde{X})$, and generate **statistical ensembles of initializations** from $\pi(X|\tilde{X})$.

Plan for the direct method

To define $q(Y|\tilde{Y})$ we use the model

$$\omega(s,t) = \hat{Y}_{bc}(s,t) - \hat{X}(s,t),$$

- $\hat{Y}_{bc}(s,t)$ are the historic bias-corrected MM5 forecasts of a meteorological variable at the spatial point $s \in \Re^2$ (station location), verifying at time t, at a given forecast lag;
- $\hat{X}(s,t)$ are the corresponding historic observations at station locations;
- $\omega(s,t)$ is a spatio-temporal random process with correlated values (**random error**);
- $\omega(s,t) \sim \mathrm{MVN}\left(0, \Omega_\omega\right)$ with the covariance function Ω_ω of some parametric form.

For simplicity. At this stage, for simplicity we are modelling only the spatial correlation in $\omega(s,t)$ and ignoring the temporal correlation, because the spatial correlation is what counts for getting calibrated images.

We assume that $\omega(s,t)$ follows the exponential spatial variogram model,

$$\frac{1}{2}\,\text{Var}\left(w(s_1,t)-w(s_2,t)\right)=\rho+\sigma^2\left(1-e^{-||s_1-s_2||r}\right)$$

for $s_1 \neq s_2$ where $||\cdot||$ is the Euclidean norm.

In geostatistical terminology,

- ρ is called the **nugget** effect and is usually thought of as the measurement error variance of observations,
- $\rho+\sigma^2$ is the marginal variance of $w(s,t)$ and is called the **sill**,
- r is a **range** parameter. It is interpreted as follows. The error process $w(s,t)$ can be viewed as a sum of two component processes: measurement error (viewed as spatially uncorrelated), and continuous spatial variation. The spatial correlation of the continuous spatial variation component process at distance d is e^{-dr}.

Then the predictive distribution of Y given $\tilde{Y}$ is defined

$$q(Y|\tilde{Y}) \sim N\left(\tilde{Y}_{bc}, \Omega_\omega\right)$$

where Ω_ω is the covariance matrix of $\omega(s)$,
$\Omega_\omega = \rho + \sigma^2 e^{-dr}$**.**

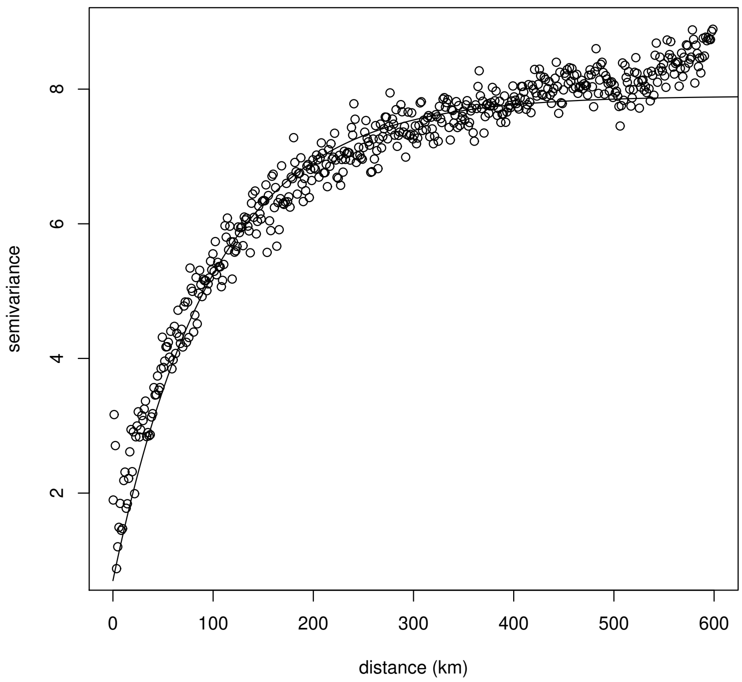
Plan:

- Postprocess MM5 forecasts, i.e. remove bias.
- Estimate the parameters ρ, σ^2 and r of the covariance function Ω_ω by semivariogram model fitting.
- Simulate statistical forecast ensembles from $q(Y|\tilde{Y})$ as realizations of the gaussian random field with the current bias corrected MM5 forecast $\tilde{Y}_{bc}$ as a mean and the covariance function Ω_ω.

Example. MM5AVN 48 hours ahead forecast of surface temperature (T2) initialized at 00Z (0GMT). Variogram of $\omega(s)$. Analysis is based on 102 days from January to June 2000.

Semivariogram

$$\frac{1}{2}\,\text{Var}\left(w(s_1,t)-w(s_2,t)\right)=\rho+\sigma^2\left(1-e^{-||s_1-s_2||r}\right)$$



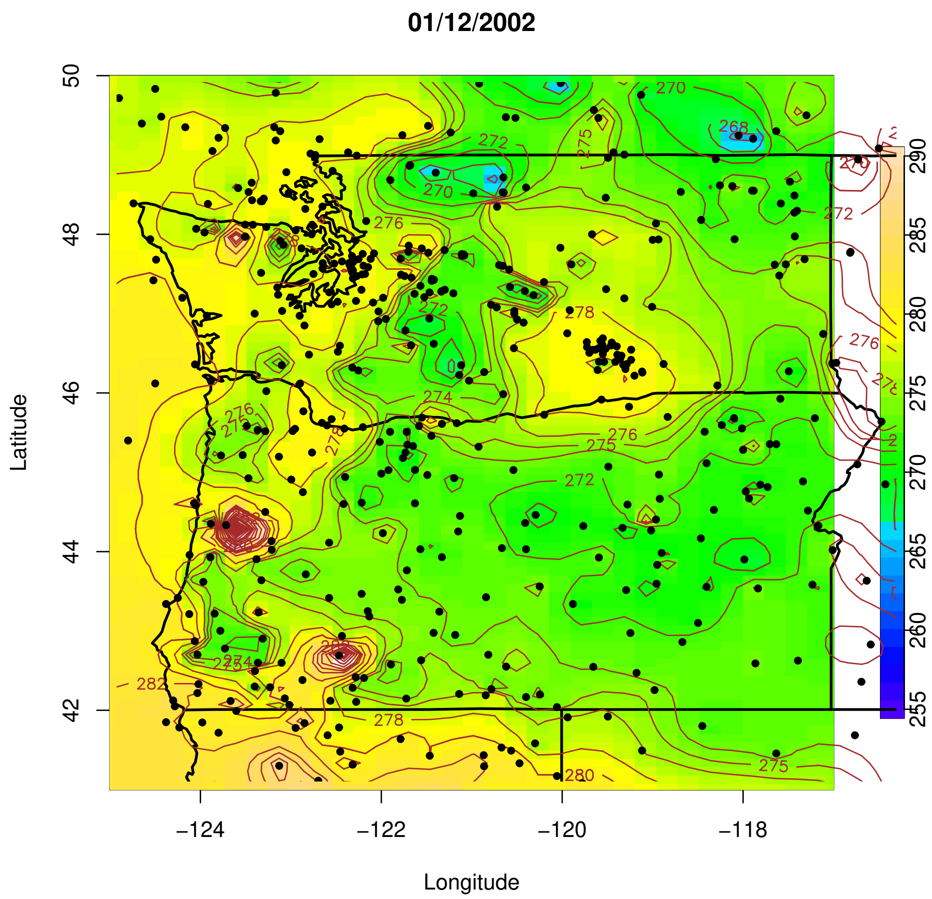
Covariance parameters for ω:

the range $\theta_\varepsilon = 0.0088$,

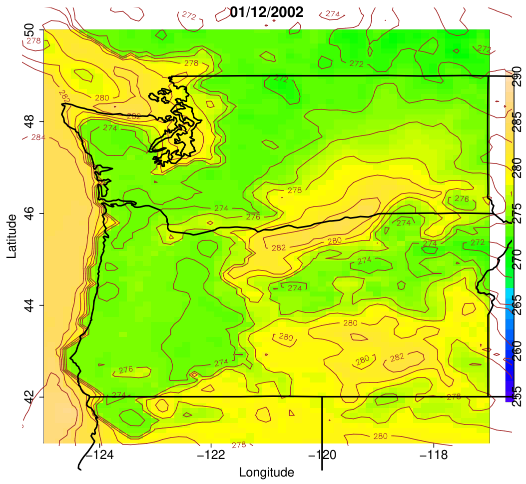
the sill $\sigma_\varepsilon^2 = 7.20$,

the nugget $= 0.51$.

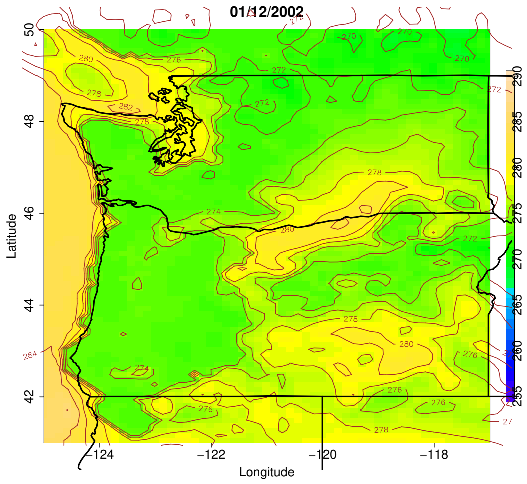
Out-of-sample verification. Observed temperature on Jan 12, 2002 at 0 hour GMT about 500 stations.



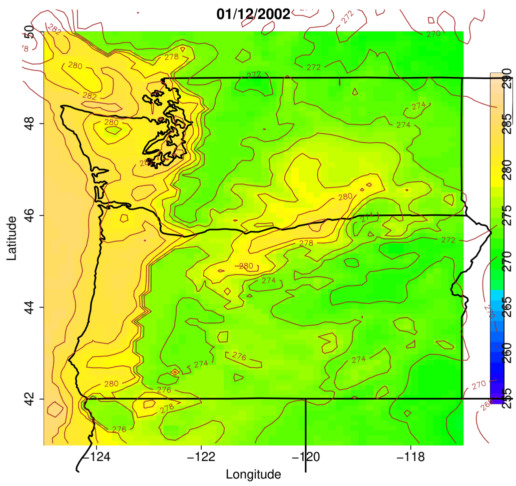
Temperature on January 12, 2002 at 0 hour GMT



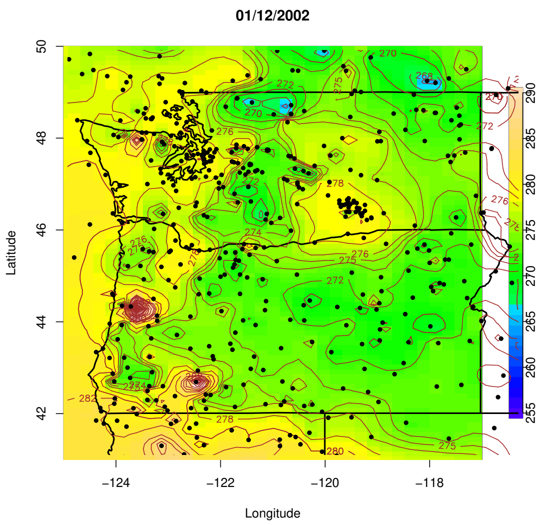
(a) The gridded MM5 forecast (init. on 10.01.2002 at 00GMT, 48 hours ahead), $\tilde{Y}$.



(b) The predictive mean, $a + b\tilde{Y}$.



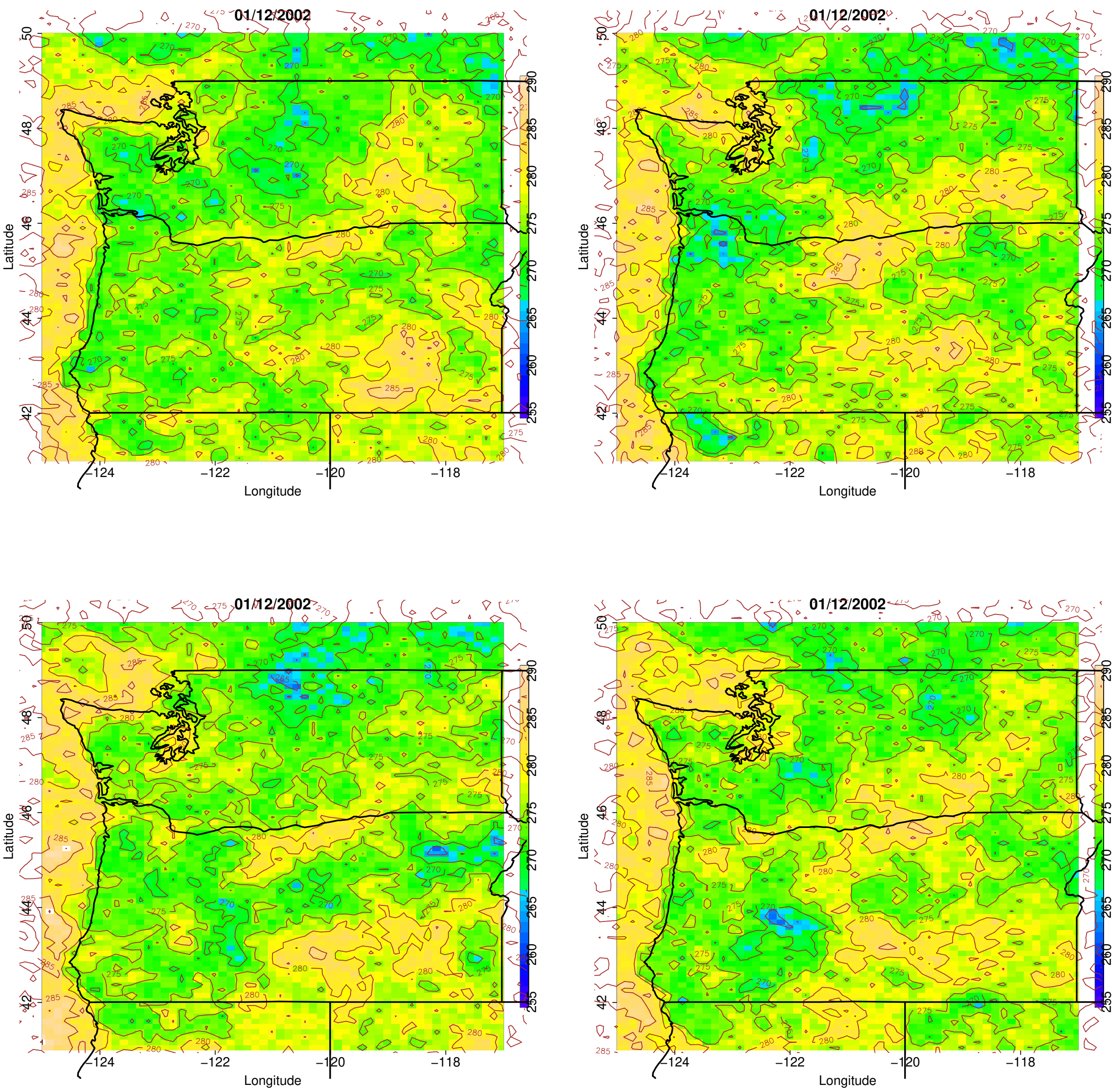
(c) Current estimation of T (the gridded MM5 output) on 12.01.2002 at 03GMT.



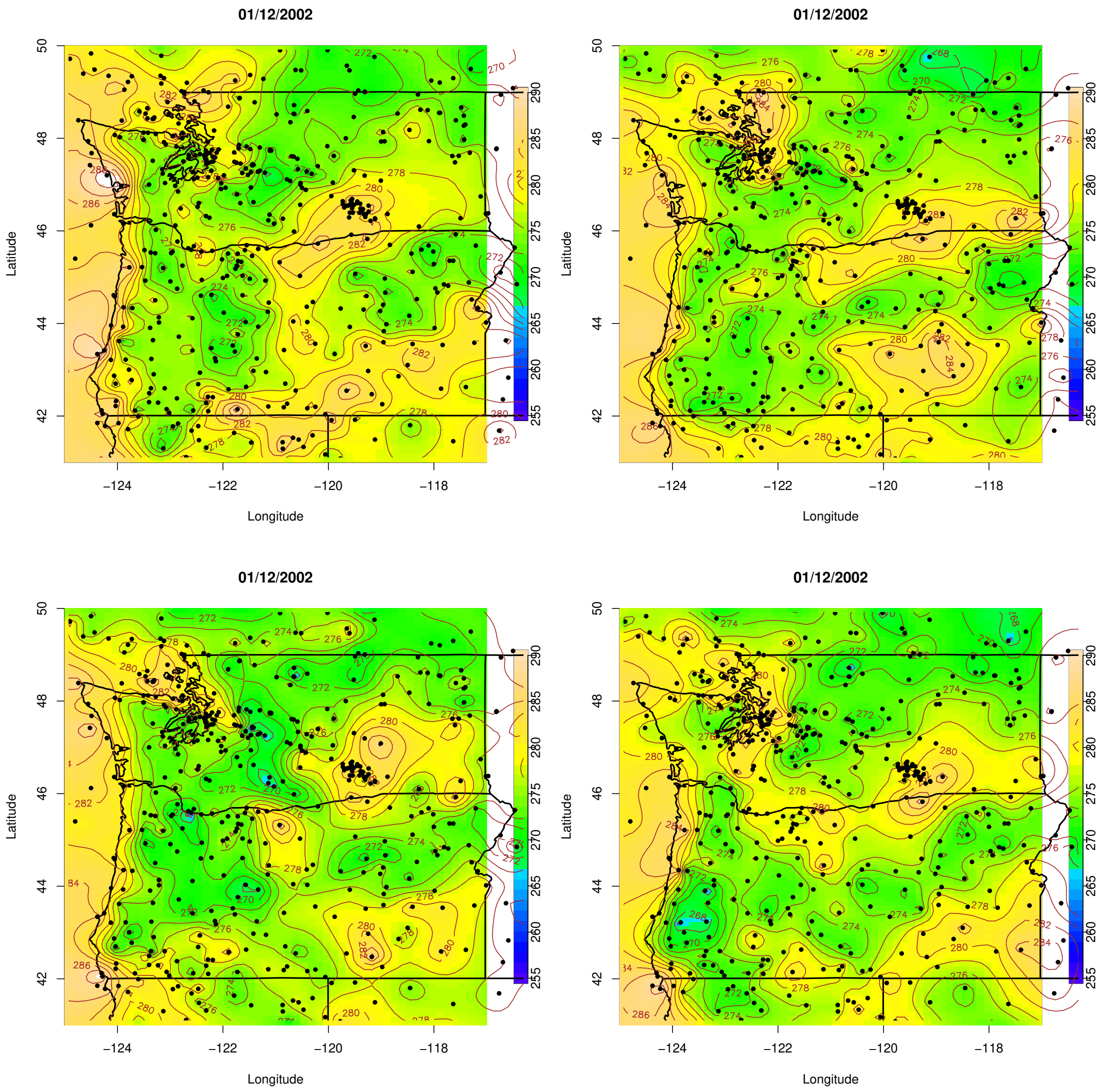
(d) Observed temperature at 500 stations) on 12.01.2002 at 00GMT

.

Simulated ensemble of forecasts on a grid. Temperature on January 12, 2002 at 0 hour GMT



Map of the simulated forecasts interpolated by kriging from the station locations. January 12, 2002 at 0 hour GMT Coverage is 90.8% for 90% PI. Actual R^2 between F and obs is 0.66.



Verification statistics

Table 1: Empirical coverage of 66.7% and 90% prediction intervals for temperature based on 99 ensemble members.

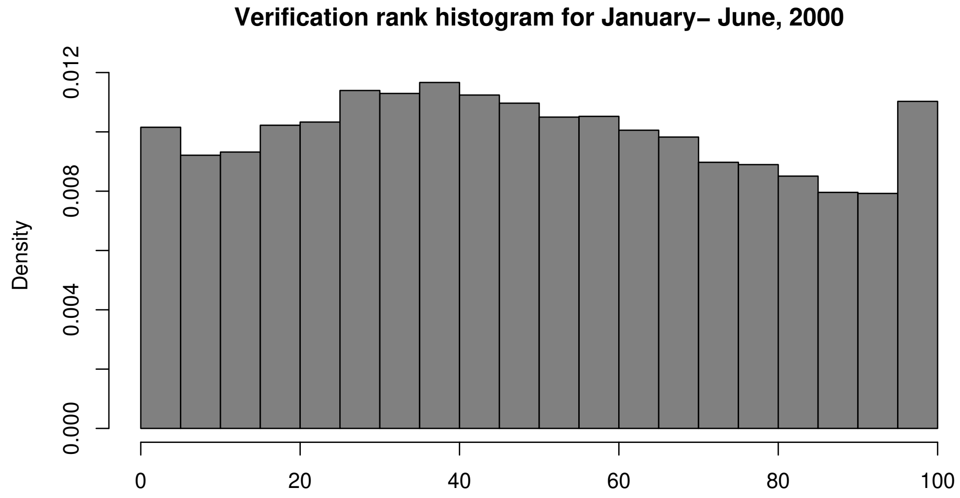
Prediction int.	Ensemble Forecast January-June, 2000	Ensemble Forecast January-June, 2001
66.7%	68.13%	67.2%
90%	90.8%	88.0%

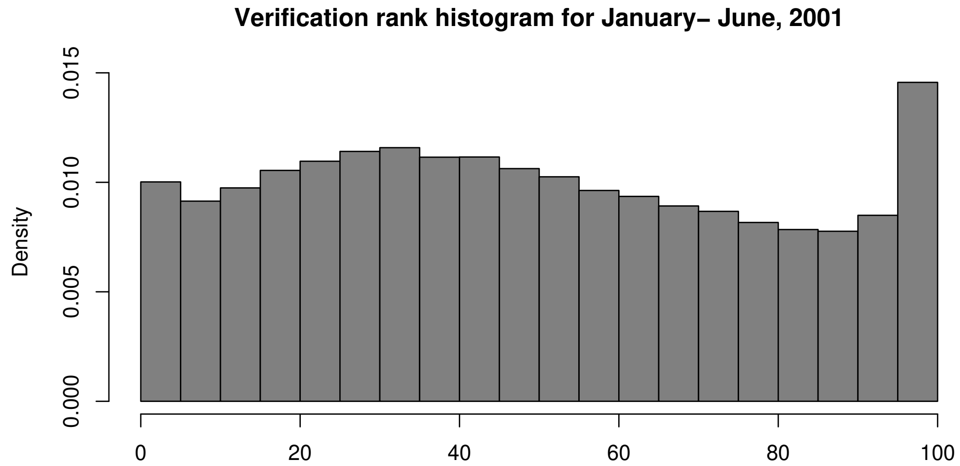
- **The prediction intervals are well calibrated.**

Table 2: Length of the prediction intervals.

Predictive intervals	66.7%	90.0%
Climatology values	17.2	28.3
Ensemble values in 2000	5.5	9.4
Ensemble values in 2001	5.5	9.1

Verification Rank Histograms





The verification rank histograms indicates a larger number of particularly high observations than anticipated by the ensemble. However, the rank histogram is much more uniform than what is typically observed in conventional ensembles and deviation is moderate.

Future directions

- <u>Combination of dynamic and statistical ensembles of forecasts</u>

- **Analysis of spatio-temporal non-stationarity:**
  - *to allow sill and range to be spatio-temporal random processes (e.g. via the Bayesian approach);*
  - *analysis of spatio-temporal separability, anisotropy (directional variogram) analysis;*

- **General form of the covariance matrix and elaboration of Gaussianity assumption**

- **Dynamic multivariate and multilevel statistical ensembles of forecasts based on multimodel initializations**