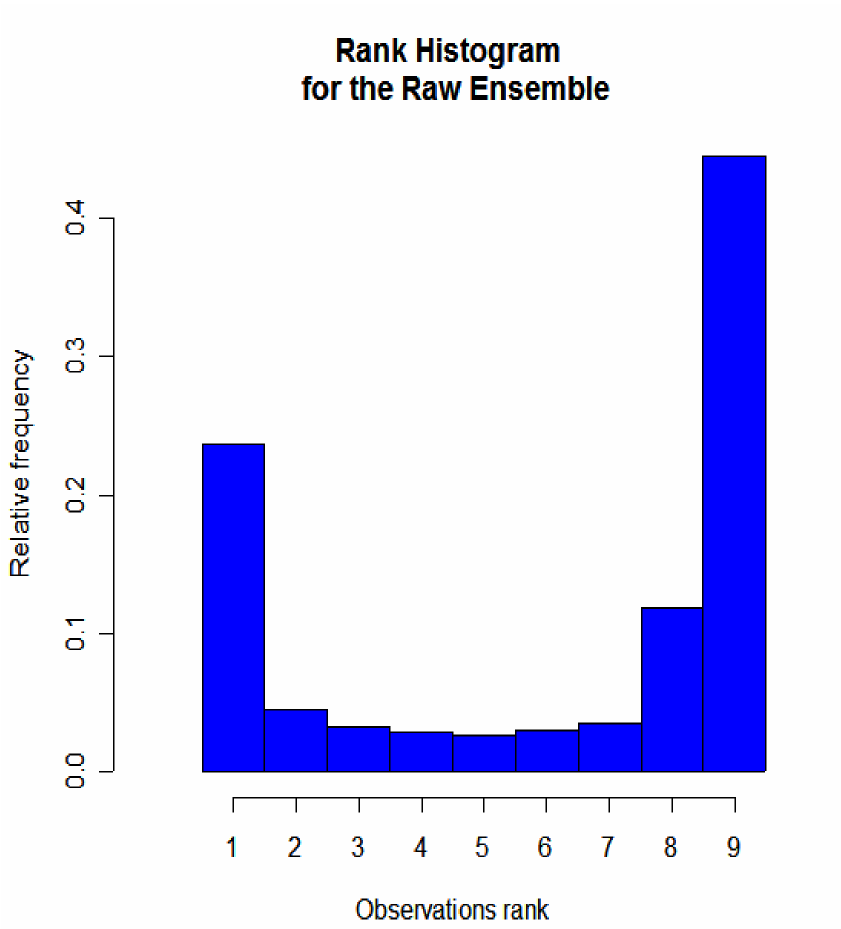
The Geostatistical Output Perturbation (GOP) Method
and
Bayesian Dressing: Statistical Ensembles for Weather Fields

Veronica J. Berrocal, Adrian E. Raftery,
Tilmann Gneiting and Yulia Gel

Department of Statistics, University of Washington

DOD Multidisciplinarity University Research Inititiative (MURI)
Program administered by the Office of Naval Research
Under grant N000 14-01-10745

Dynamical ensembles



Domain: US Pacific Northwest

- Period: 10/31/2002 – 03/31/2004
- Forecasts:
  - 12-km MM5
  - 2-m temperature at 48 h
  - NCEP AVN, CMC GEM, NCP ETA, BoM GASP, JMA, FNMOC NGPS, TCWB and UKMO for initial and lateral boundary conditions
  - Bi-linearly interpolated to observation locations
- Observations:
  - 2-m temperature from mesoscale network

Limitations of dynamical ensembles

- Ensemble members tend to agree with each other more than with the observed weather quantity (*i.e. ensembles tend to be underdispersive)*
- As a consequence, ensembles are not calibrated

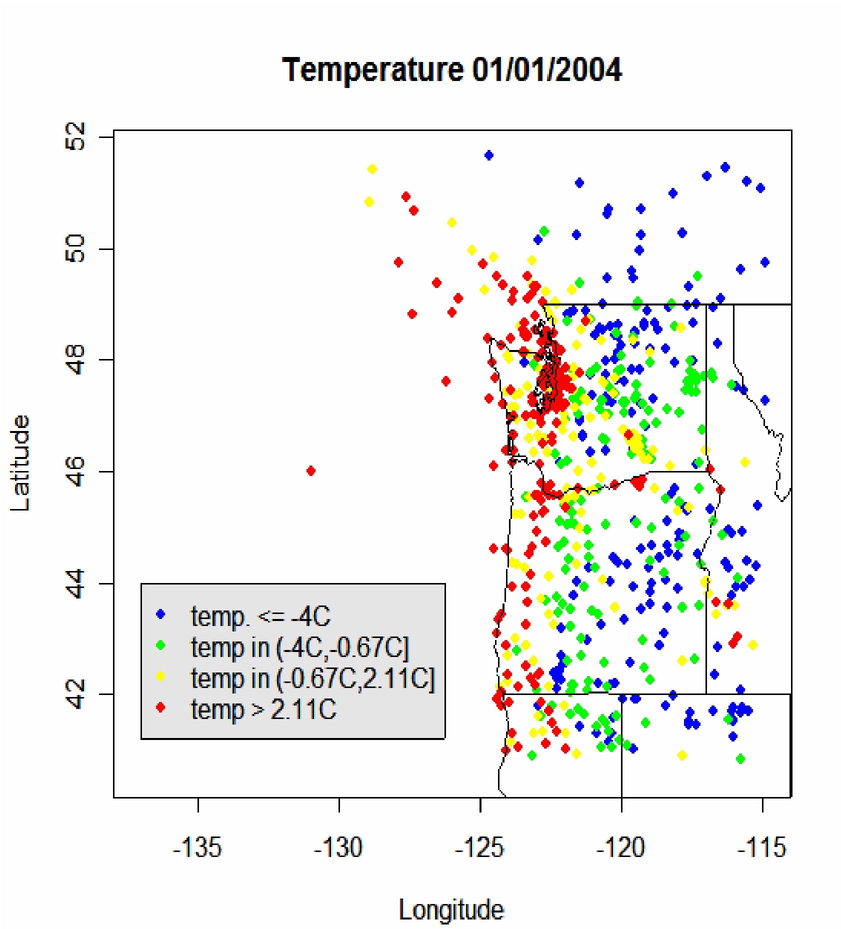
- Statistical methods to postprocess ensembles (example: MOS, Ensemble MOS, Bayesian Model Averaging)

The Geostatistical Output Perturbation (GOP) Method and the Bayesian Dressing Method

- Statistical methods to postprocess ensembles and create statistical ensembles of weather fields

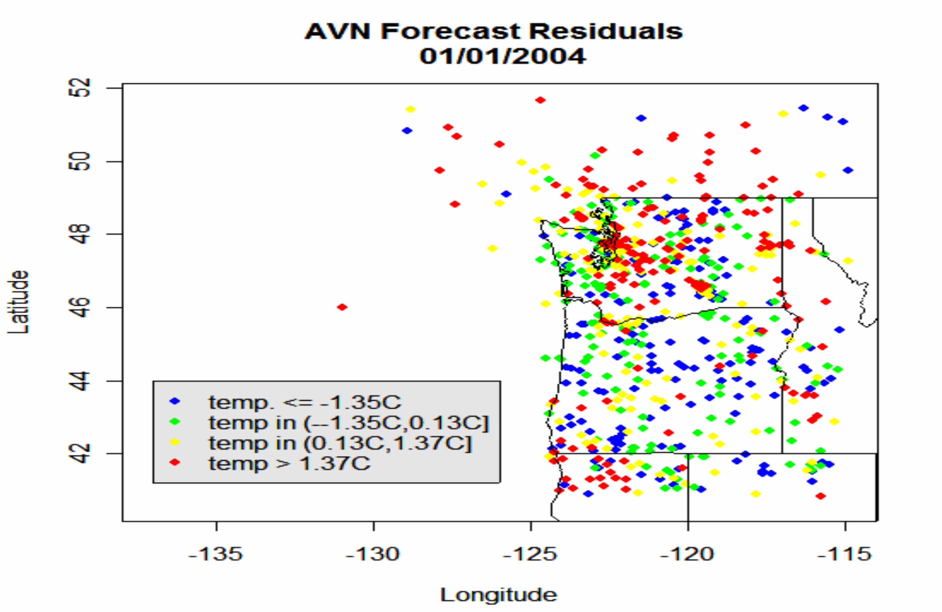
- Common features:
  - Inexpensive and quick way to create new ensemble
  - Apply to whole weather fields simultaneously
  - Preserve the field's spatial correlation structure
  - Perturb the output rather than the input

Spatial structure of weather fields



- 2-m temperature from mesoscale network observed on Jan. 1st, 2004
- Temperatures are represented by different colors using the quartiles of the distribution of 2-m temperature observed on Jan. 1st, 2004
- Spatial pattern in the data: locations closer together have similar temperatures

How spatial structure is used in the GOP method (Gel, Raftery, Gneiting, JASA, 2004)



True weather quantity = *Bias-adjusted forecast weather quantity* + *error*

- The spatial structure of weather fields is estimated using forecast errors and historical data
- New forecast errors are simulated using the estimated spatial structure
- New ensemble members are generated adding the simulated forecast errors to the bias-adjusted forecast

The GOP "Ensemble"

Traditional Ensemble	GOP "Ensemble"

Initial conditions

Ensemble Member 1

Ensemble Member 2

.
.
.
.

Ensemble Member m

Ensemble Member 1

Ensemble Member 2

.
.
.

Ensemble Member N

Limitations of the GOP Method

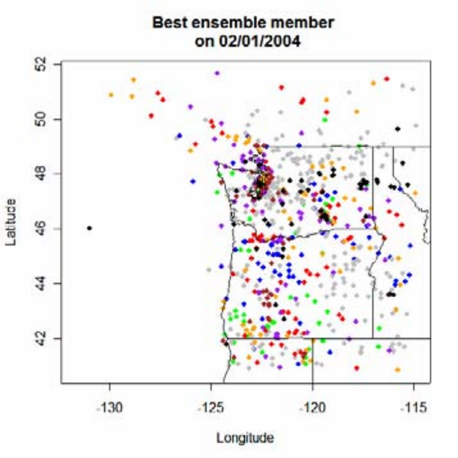
- It applies to only one member of the ensemble
- It does not take into account the change in relative skill of the member of the original dynamical ensemble

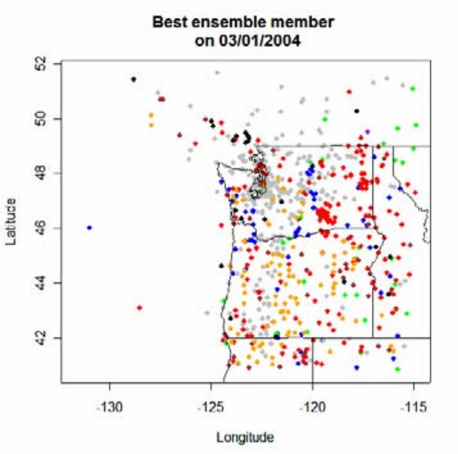
- We need to take into account of ensemble information contained in the original dynamical ensemble

Bayesian Model Averaging (BMA) in probabilistic weather forecasting
(Raftery et al., 2005)

- Based on the idea that in any given dynamical ensemble there is always a "best" model but we do not know which one it is.
- The true weather quantity is expressed as a weighted average of bias-adjusted forecasts where the weights reflect the relative skill of each ensemble members.
- The relative skill of each member can change over time.
- Forecasts are made at individual locations. The locations are considered mutually independent, and the spatial structure of the weather field is ignored.

Which one is the best ensemble member?





- 48 h forecasts of 2-m temperature initialized on February 1st and March 1st, 2004
- The ensemble member with the smallest forecast error is displayed at each location
- Colors:

(blue)	**NCEP AVN**	(magenta)	**JMA**
(green)	**CMC GEM**	(brown)	**FNMOC NGPD**
(gray)	**NCEP ETA**	(orange)	**TCWB**
(red)	**BoM GASP**	(black)	**UKMO**

The Bayesian Dressing Method

Bayesian Model Averaging

(it takes into account the ensemble information)

+

The Geostatistical Output Perturbation method

(it takes into account the spatial structure of the weather field)

Bayesian Dressing

The Bayesian Dressing Method:

- combines the GOP method and BMA
- is marginally BMA, at each location
- is jointly like BMA, “dressed” with a spatial structure

The Bayesian Dressing "Ensemble"

Original Ensemble of m members

Bayesian Dressing "Ensemble" of N members

Ensemble Member 1 — **total:** $N \cdot w_1$

New Ensemble Member
New Ensemble Member
.
.
.
New Ensemble Member

.
.
.
.
.

Ensemble Member m — **total:** $N \cdot w_m$

New Ensemble Member
New Ensemble Member
.
.
.
New Ensemble Member

An example:
Application of the Bayesian Dressing Method

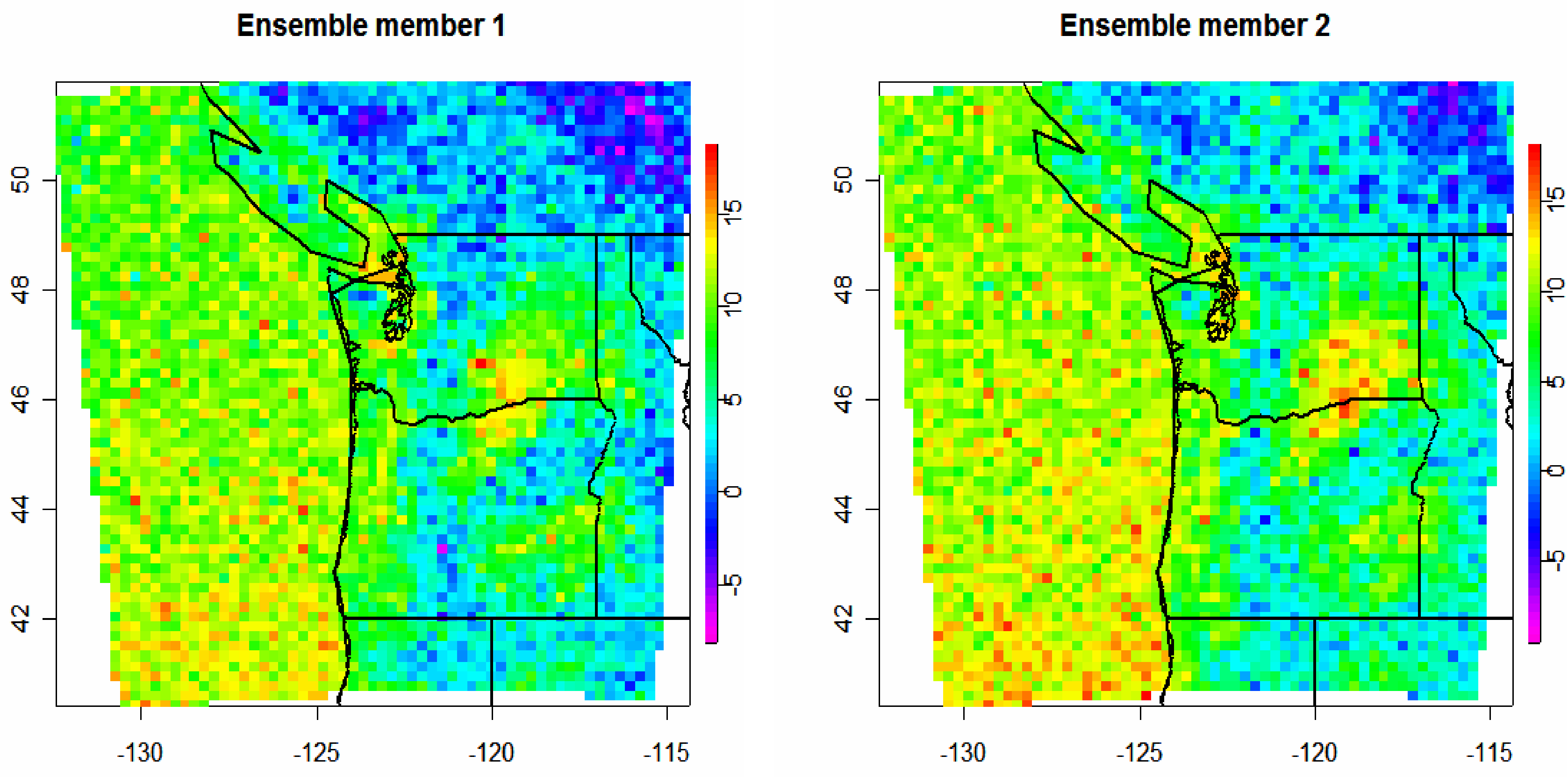
- Data
  - Domain: US Pacific Northwest
  - Period:
    - 10/31/2002-04/15/2003: Historical data used to determine the spatial structure of the weather field
    - 01/01/2004 – 03/31/2004: Initialization period for forecasts and observations
  - Forecasts:
    - 12-km MM5
    - 2-m temperature at 48 h
    - NCEP AVN, CMC GEM, NCEP ETA, BoM GAPS, JMA, FNMOC NGPS, TCWB, and UKMO for initial and lateral boundary conditions
    - Bi-linearly interpolated to observation locations
  - Observations:
    - 2-m temperature from mesoscale network

Bayesian Dressing "Ensemble" of Forecasts

- Probabilistic 48 h forecasts of 2-m temperature in the US Pacific Northwest verifying in the period from 01/03/2004 to 04/02/2004 at 00Z
- For each day, we have generated a 20-member Bayesian Dressing "Ensemble" of 48 h forecasts
- In each ensemble, members were obtained by perturbing one of the original 8 members of the UW MM5 ensembles according to the BMA weights

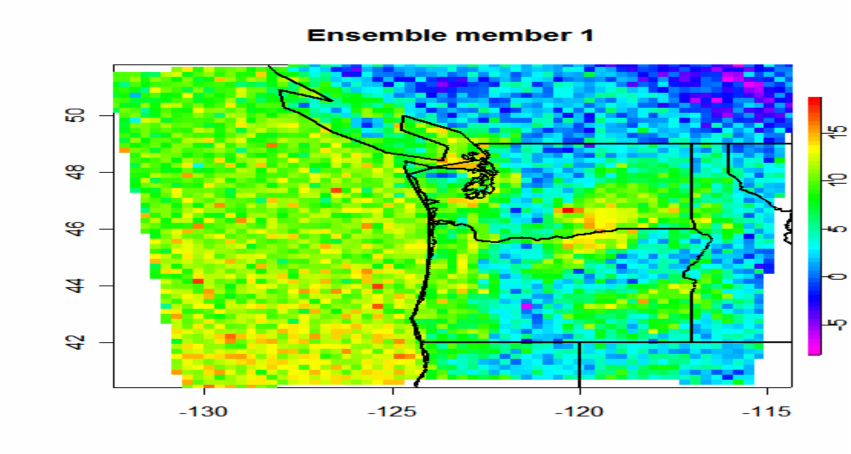
Bayesian Dressing "Ensemble" of Forecasts

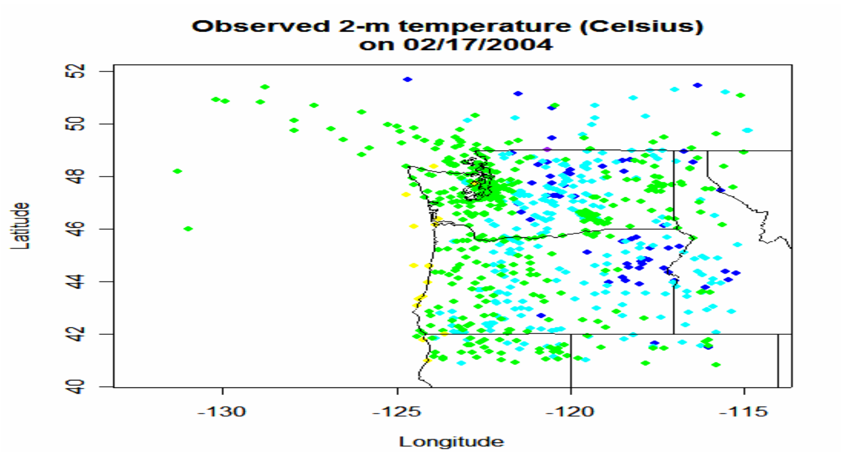
2-m temperature (degree Celsius)– verifying on 02/17/2004



How good is the Bayesian Dressing "Ensemble" of Forecasts?

Spatial verification

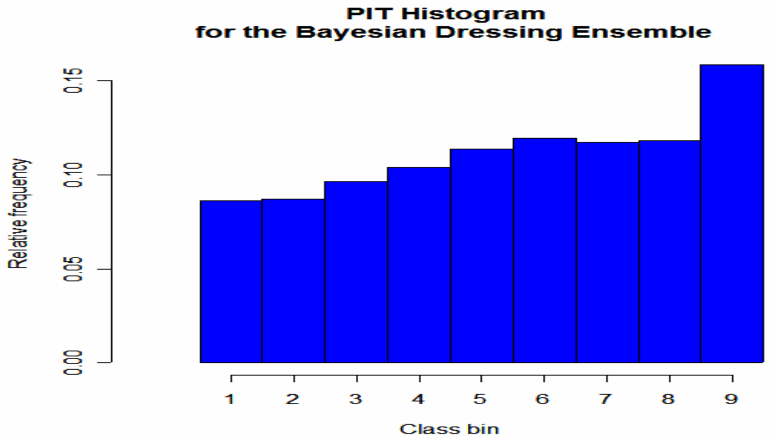


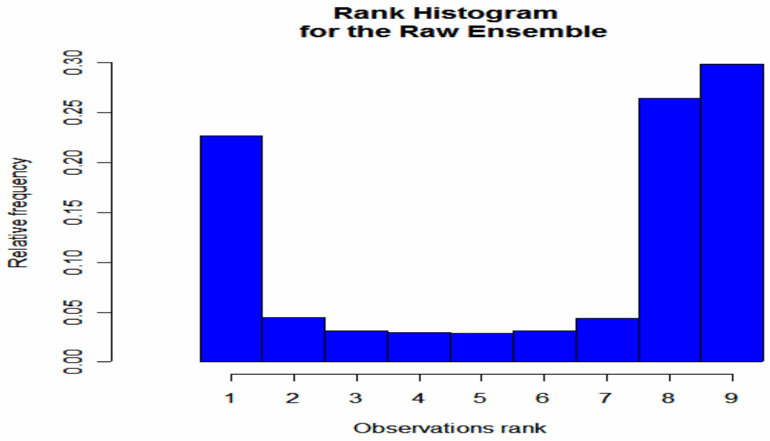


- One member of the Bayesian Dressing "Ensemble" of 48 h forecasts of 2-m temperature verifying on February 17th, 2004
- Observations of 2-m temperature (measured in degree Celsius) at weather stations from mesoscale network represented using the same color scale as for the ensemble member

How good is the Bayesian Dressing "Ensemble" of Forecasts?

Marginal verification





- Probability Integral Transform histogram of the Bayesian Dressing "Ensemble" of 48 h forecasts of 2-m temperature for the period 01/01/2004-03/31/2004

- Rank histogram of raw ensemble of 48 h forecasts of 2-m temperature for the period 01/01/2004-03/31/2004

- Flatter is better

Future work

- Apply the Bayesian Dressing method to produce probabilistic 48 h weather forecasts of 2-m temperature for the Pacific Northwest for a longer period of time (1 year)
- Apply the Bayesian Dressing method in a "regionalized" fashion by dividing the domain (US Pacific Northwest) according to land use
- Work on more formal verification methods

The `ProbForecastGOP` package

- The `ProbForecastGOP` package is an R package available on CRAN (see: www.r-project.org) that produces probabilistic weather forecasts of weather fields using the GOP method
- Input:

Date of observation	**Forecast of weather at grid points**
Observed weather quantity	**Latitude and longitude of grid points**
Latitude and longitude of weather stations	**ID of weather stations**
Forecast of weather quantity at weather stations	**Number N of ensemble members**

- Output:

  GOP "Ensemble" of Forecasts of N members

  Percentiles of weather field

Conclusions

- The GOP method is a simple method that produces calibrated and sharp probabilistic forecast of weather field
- The Bayesian Dressing method extends the GOP method to account for the changing relative skill of the ensemble members, by combining BMA and the GOP method
- The `ProbForecastGOP` package is an R package that implements the GOP method