Probabilistic Weather Forecasting via Bayesian Model Averaging

Adrian E. Raftery
University of Washington
raftery@u.washington.edu
www.stat.washington.edu/raftery

Joint work with Tilmann Gneiting
with contributions by Veronica Berrocal, Chris Fraley, Yulia Gel and McLean Sloughter
In collaboration with Cliff Mass, Susan Joslyn, and Jeff Baars
Supported by NSF and the ONR MURI Program

NWS Visit
University of Washington
May 27, 2009

Basic Idea

Basic Idea

- Forecast ensembles:

Basic Idea

- Forecast ensembles:
    - The dominant approach to probabilistic forecasting

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface

Basic Idea

- Forecast ensembles:
    - The dominant approach to probabilistic forecasting
    - They contain useful information (spread-skill relationship)
    - BUT they tend to be underdispersed, especially at the surface
    - uncalibrated

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
  - currently for the Pacific Northwest,

Basic Idea

- Forecast ensembles:
    - The dominant approach to probabilistic forecasting
    - They contain useful information (spread-skill relationship)
    - BUT they tend to be underdispersed, especially at the surface
    - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
    - based on an ensemble
    - for temperature, PoP, quantitative precip, wind speeds
    - and potentially other parameters
    - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
    - currently for the Pacific Northwest,
    - temperature and precip

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
  - currently for the Pacific Northwest,
  - temperature and precip
- UW can work with NWS to extend it to

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
  - currently for the Pacific Northwest,
  - temperature and precip
- UW can work with NWS to extend it to
  - other parameters

Basic Idea

- Forecast ensembles:
    - The dominant approach to probabilistic forecasting
    - They contain useful information (spread-skill relationship)
    - BUT they tend to be underdispersed, especially at the surface
    - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
    - based on an ensemble
    - for temperature, PoP, quantitative precip, wind speeds
    - and potentially other parameters
    - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
    - currently for the Pacific Northwest,
    - temperature and precip
- UW can work with NWS to extend it to
    - other parameters
    - aloft as well as surface

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
  - currently for the Pacific Northwest,
  - temperature and precip
- UW can work with NWS to extend it to
  - other parameters
  - aloft as well as surface
  - the nation (and beyond)

Basic Idea

- Forecast ensembles:
  - The dominant approach to probabilistic forecasting
  - They contain useful information (spread-skill relationship)
  - BUT they tend to be underdispersed, especially at the surface
  - uncalibrated
- Bayesian Model Averaging (BMA) provides calibrated and sharp probabilistic forecasts
  - based on an ensemble
  - for temperature, PoP, quantitative precip, wind speeds
  - and potentially other parameters
  - for entire weather fields and multiple parameters simultaneously (desirable for aviation)
- BMA is the basis for `Probcast`, the first operational PDF-based calibrated probabilistic forecasting website
  - currently for the Pacific Northwest,
  - temperature and precip
- UW can work with NWS to extend it to
  - other parameters
  - aloft as well as surface
  - the nation (and beyond)
  - $\Longrightarrow$ the 4D probabilistic forecasting cube

Outline

Outline

- Bayesian model averaging

Outline

- Bayesian model averaging
- BMA for exchangeable ensembles

Outline

- Bayesian model averaging
- BMA for exchangeable ensembles
- BMA for precipitation

Outline

- Bayesian model averaging
- BMA for exchangeable ensembles
- BMA for precipitation
- Local BMA

Outline

- Bayesian model averaging
- BMA for exchangeable ensembles
- BMA for precipitation
- Local BMA
- BMA for weather fields

Bayesian Model Averaging for Ensembles

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.
- BMA model for temperature:

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.
- BMA model for temperature:
  - Let y be the verifying value and $\tilde{y}_k$ be the kth forecast from the ensemble.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.
- BMA model for temperature:
  - Let y be the verifying value and $\tilde{y}_k$ be the kth forecast from the ensemble.
  - The model is:

$$p(y|\tilde{y}_1, \ldots, \tilde{y}_K) = \sum_{k=1}^{K} w_k N(a_k + b_k \tilde{y}_k, \sigma^2)$$

where $w_k \geq 0$ and $\sum_{k=1}^{K} w_k = 1$.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.
- BMA model for temperature:
  - Let y be the verifying value and $\tilde{y}_k$ be the kth forecast from the ensemble.
  - The model is:

$$p(y|\tilde{y}_1, \ldots, \tilde{y}_K) = \sum_{k=1}^{K} w_k N(a_k + b_k \tilde{y}_k, \sigma^2)$$

where $w_k \geq 0$ and $\sum_{k=1}^{K} w_k = 1$.

- The model is estimated from a training set of recent data at stations by maximum likelihood using the EM algorithm.

Bayesian Model Averaging for Ensembles

- The overall (BMA) forecast probability distribution is a mixture of distributions, each one centered on one of the forecasts after bias correction.
- The weights are the estimated probabilities of the models, and reflect the relative predictive performance of the models during a training period.
- The BMA point or deterministic forecast is just a weighted average of the forecasts in the ensemble.
- BMA model for temperature:
  - Let y be the verifying value and $\tilde{y}_k$ be the kth forecast from the ensemble.
  - The model is:

$$p(y|\tilde{y}_1, \ldots, \tilde{y}_K) = \sum_{k=1}^{K} w_k N(a_k + b_k \tilde{y}_k, \sigma^2)$$

  where $w_k \geq 0$ and $\sum_{k=1}^{K} w_k = 1$.
- The model is estimated from a training set of recent data at stations by maximum likelihood using the EM algorithm.
- Good results with a 25-day "moving window" training period.

UW Ensemble Bayesian Model Averaging

User's Guide

Param: Max 2m Temp (24–48 hrs)

Valid for 24 hours ending at:

Wed May 27, 2009 5 PM

Jump to new date

Toggle Contour Lines OFF

Plot Size: Big Medium Small

Units: Celsius Fahrenheit

Grid Forecast:

Deterministic

Upper bound of interval .9

Lower bound of interval .1

Half-width of interval

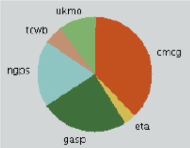
Prob. param exceeds threshold 0

Greater than Less than

Probability Distribution:

Latitude: 47.68049 Longitude: -122.2244

Retrieve Data



BMA Weights

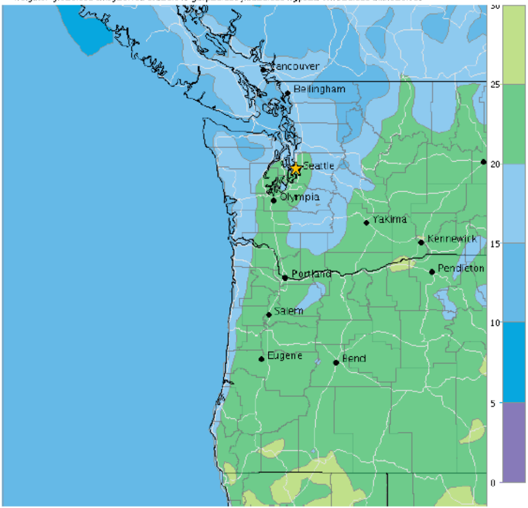
Forecasts Error: NORMAL: 2.16

BMA Forecast Verification

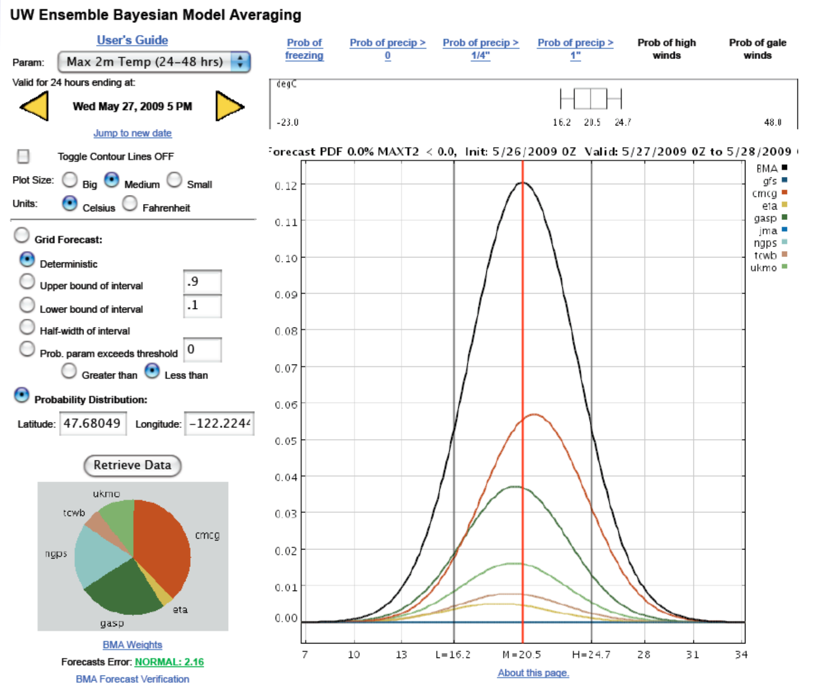
Prob of freezing | Prob of precip > 0 | Prob of precip > 1/4" | Prob of precip > 1" | Prob of high winds | Prob of gale winds

BMA MAXT2 Forecast Init: 5/26/09 12:00 AM UTC (5/25/09 5:00 PM) Valid: 5/26/09 5:00 PM to 5/27/09 5:00 PM PT

Weights: gfs=0.000 cmcg=0.462 eta=0.041 gasp=0.302 jma=0.000 ngps=[illegible] tcwb=0.063 ukmo=0.131



About this page.



UW Ensemble Bayesian Model Averaging
User's Guide
Param: Max 2m Temp (24–48 hrs)
Valid for 24 hours ending at:
Wed May 27, 2009 5 PM
Jump to new date
Toggle Contour Lines OFF
Plot Size: Big Medium Small
Units: Celsius Fahrenheit
Grid Forecast:
Deterministic
Upper bound of interval .9
Lower bound of interval .1
Half-width of interval
Prob. param exceeds threshold 0
Greater than Less than
Probability Distribution:
Latitude: 47.68049 Longitude: -122.2244
Retrieve Data
ukmo
tcwb
cmcg
ngps
eta
gasp
BMA Weights
Forecasts Error: NORMAL: 2.16
BMA Forecast Verification
Prob of freezing
Prob of precip > 0
Prob of precip > 1/4"
Prob of precip > 1"
Prob of high winds
Prob of gale winds
degC
-23.0
16.2 20.5 24.7
48.0
Forecast PDF 0.0% MAXT2 < 0.0, Init: 5/26/2009 0Z Valid: 5/27/2009 0Z to 5/28/2009
BMA
gfs
cmcg
eta
gasp
jma
ngps
tcwb
ukmo
0.12
0.11
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
7 10 13 L=16.2 M=20.5 H=24.7 28 31 34
About this page.

Results for 2007 ($^\circ C$)

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble		

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	
BMA for UWME		

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	
BMA for UWME	2.15	

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

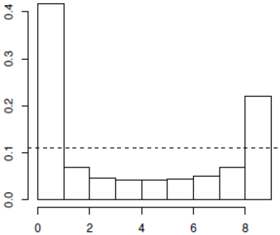
	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Results for 2007 ($^\circ C$)

(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram
for raw ensemble

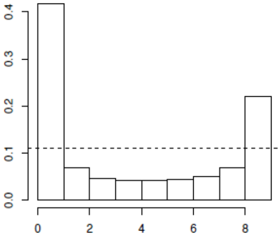


Results for 2007 ($^\circ C$)

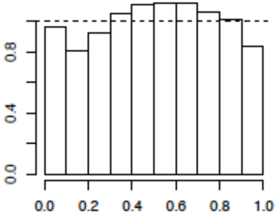
(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram for raw ensemble



PIT histogram for BMA

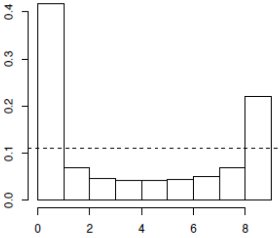


Results for 2007 ($^\circ C$)

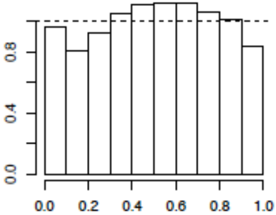
(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram for raw ensemble



PIT histogram for BMA



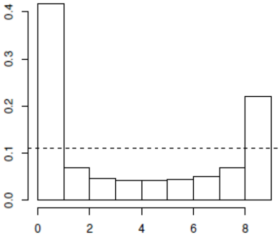
- BMA better calibrated and more accurate than the raw ensemble

Results for 2007 ($^\circ C$)

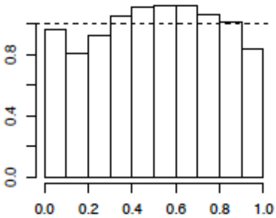
(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram for raw ensemble



PIT histogram for BMA



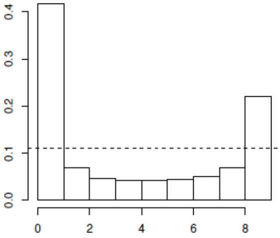
- BMA better calibrated and more accurate than the raw ensemble
- Similar results for other times (2000-2009) and places:

Results for 2007 ($^\circ C$)

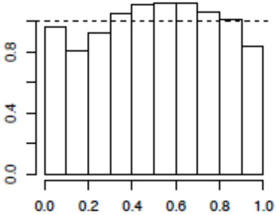
(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram for raw ensemble



PIT histogram for BMA



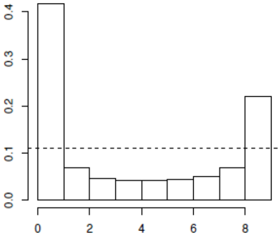
- BMA better calibrated and more accurate than the raw ensemble
- Similar results for other times (2000-2009) and places:
    - Canada (Wilson et al 2007 [MSC])

Results for 2007 ($^\circ C$)

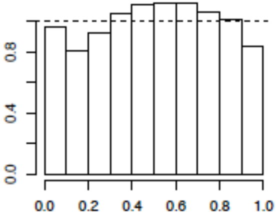
(24hr forecasts of 2m temperature at ASOS stations and buoys)

	MAE	CRPS
Raw Ensemble	2.31	1.96
BMA for UWME	2.15	1.55

Verification rank histogram for raw ensemble



PIT histogram for BMA



- BMA better calibrated and more accurate than the raw ensemble
- Similar results for other times (2000-2009) and places:
  - Canada (Wilson et al 2007 [MSC])
  - Netherlands/Europe (KNMI)

BMA for Ensembles with Exchangeable Members

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
    - But often ensembles have subsets of exchangeable members:

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
  - But often ensembles have subsets of exchangeable members:
  - ECMWF: 2 groups: control (1); singular vector perturbations (50)

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
  - But often ensembles have subsets of exchangeable members:
  - ECMWF: 2 groups: control (1); singular vector perturbations (50)
  - NCEP 2006 SREF: 21 members divided into 13 groups

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
  - But often ensembles have subsets of exchangeable members:
  - ECMWF: 2 groups: control (1); singular vector perturbations (50)
  - NCEP 2006 SREF: 21 members divided into 13 groups
- Exchangeable BMA forces the weights for exchangeable members to be the same.

BMA for Ensembles with Exchangeable Members

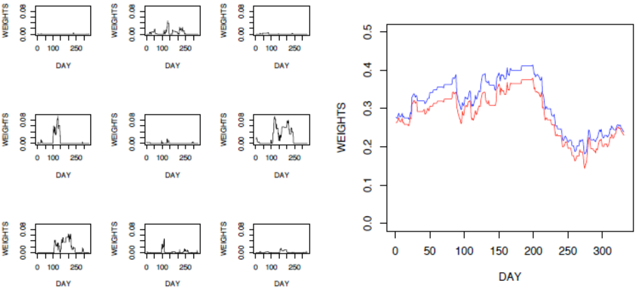
- "Traditional" BMA gives a different weight to each ensemble member
  - But often ensembles have subsets of exchangeable members:
  - ECMWF: 2 groups: control (1); singular vector perturbations (50)
  - NCEP 2006 SREF: 21 members divided into 13 groups
- Exchangeable BMA forces the weights for exchangeable members to be the same.
- Example: 89-member ensemble combining UWME (8 individual members) and UW-EnKF (experimental: 80 exchangeable members).

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
    - But often ensembles have subsets of exchangeable members:
    - ECMWF: 2 groups: control (1); singular vector perturbations (50)
    - NCEP 2006 SREF: 21 members divided into 13 groups
- Exchangeable BMA forces the weights for exchangeable members to be the same.
- Example: 89-member ensemble combining UWME (8 individual members) and UW-EnKF (experimental: 80 exchangeable members).

BMA for Ensembles with Exchangeable Members

- "Traditional" BMA gives a different weight to each ensemble member
  - But often ensembles have subsets of exchangeable members:
  - ECMWF: 2 groups: control (1); singular vector perturbations (50)
  - NCEP 2006 SREF: 21 members divided into 13 groups
- Exchangeable BMA forces the weights for exchangeable members to be the same.
- Example: 89-member ensemble combining UWME (8 individual members) and UW-EnKF (experimental: 80 exchangeable members).



Exchangeable BMA: Results for 2007

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)		
UW EnKF (80 exchangeable)		
Combined (89)		

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	
UW EnKF (80 exchangeable)		
Combined (89)		

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	
UW EnKF (80 exchangeable)	3.32	
Combined (89)		

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	
UW EnKF (80 exchangeable)	3.32	
Combined (89)		

- UW EnKF worse than UW ME (experimental)

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	
UW EnKF (80 exchangeable)	3.32	
Combined (89)	3.25	

- UW EnKF worse than UW ME (experimental)

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	
UW EnKF (80 exchangeable)	3.32	
Combined (89)	3.25	

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	2.15
UW EnKF (80 exchangeable)	3.32	
Combined (89)	3.25	

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	2.15
UW EnKF (80 exchangeable)	3.32	2.49
Combined (89)	3.25	

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	2.15
UW EnKF (80 exchangeable)	3.32	2.49
Combined (89)	3.25	2.09

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	2.15
UW EnKF (80 exchangeable)	3.32	2.49
Combined (89)	3.25	2.09

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME
- BMA improves all 3 ensembles

Exchangeable BMA: Results for 2007

	MAE	
	Raw	BMA
UW ME (8)	2.31	2.15
UW EnKF (80 exchangeable)	3.32	2.49
Combined (89)	3.25	2.09

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME
- BMA improves all 3 ensembles
- With BMA, combined ensemble *better* than UW ME alone!

Exchangeable BMA: Results for 2007

	MAE		CRPS	
	Raw	BMA	Raw	BMA
UW ME (8)	2.31	2.15	1.96	1.55
UW EnKF (80 exchangeable)	3.32	2.49	2.84	1.76
Combined (89)	3.25	2.09	2.64	1.48

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME
- BMA improves all 3 ensembles
- With BMA, combined ensemble *better* than UW ME alone!

Exchangeable BMA: Results for 2007

	MAE		CRPS	
	Raw	BMA	Raw	BMA
UW ME (8)	2.31	2.15	1.96	1.55
UW EnKF (80 exchangeable)	3.32	2.49	2.84	1.76
Combined (89)	3.25	2.09	2.64	1.48

- UW EnKF worse than UW ME (experimental)
- Combined *raw* ensemble worse than UW ME
- BMA improves all 3 ensembles
- With BMA, combined ensemble *better* than UW ME alone!
- Same conclusion with MAE (deterministic) and CRPS (probabilistic)

BMA for Precipitation

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of
  - a point mass at zero, whose probability is specified by logistic regression given the forecast

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of
  - a point mass at zero, whose probability is specified by logistic regression given the forecast
  - a gamma distribution, whose mean and variance depend on the forecast

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of
  - a point mass at zero, whose probability is specified by logistic regression given the forecast
  - a gamma distribution, whose mean and variance depend on the forecast
- We then proceed with maximum likelihood estimation as before

BMA for Precipitation

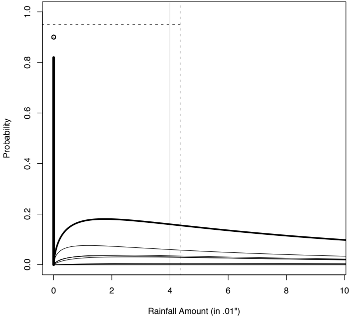
- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of
  - a point mass at zero, whose probability is specified by logistic regression given the forecast
  - a gamma distribution, whose mean and variance depend on the forecast
- We then proceed with maximum likelihood estimation as before
- Recently extended to wind speeds:

BMA for Precipitation

- The normal-based BMA used for temperature doesn't work for precipitation because:
  - precip has a nonzero probability of being exactly zero
  - it is constrained to be nonnegative
  - its distribution, given that it is not zero, is strongly skewed
- For probabilistic forecasting of precipitation, we replace the normal distribution by a mixture of
  - a point mass at zero, whose probability is specified by logistic regression given the forecast
  - a gamma distribution, whose mean and variance depend on the forecast
- We then proceed with maximum likelihood estimation as before
- Recently extended to wind speeds:
  - Zero component not needed in the Pacific Northwest

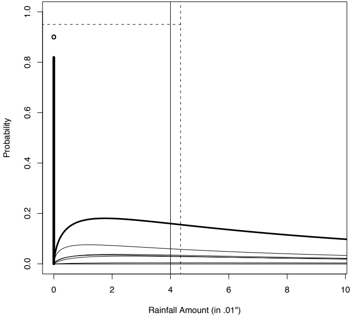
BMA Predictive Distributions for Precipitation

BMA Predictive Distributions for Precipitation

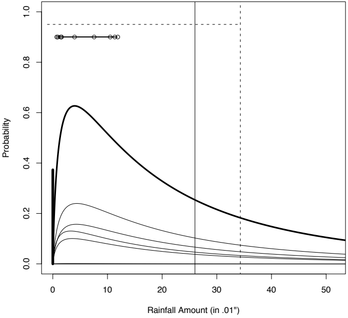


Renton, 19th May, 2003

BMA Predictive Distributions for Precipitation



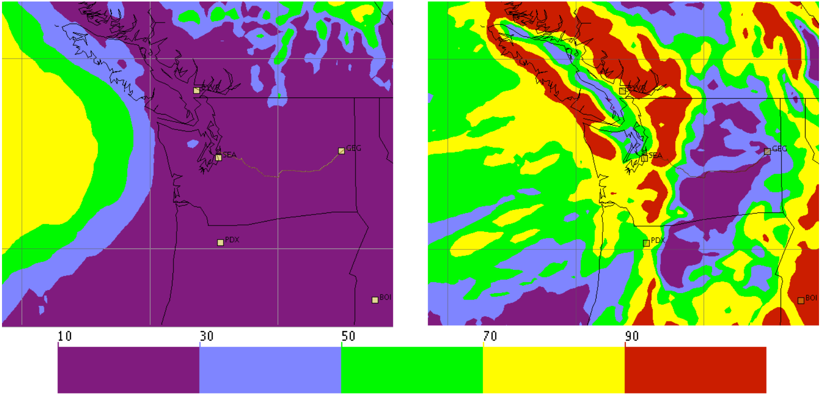
Renton, 19th May, 2003



Station KPWT, 26th January, 2003

BMA Probability of Precipitation

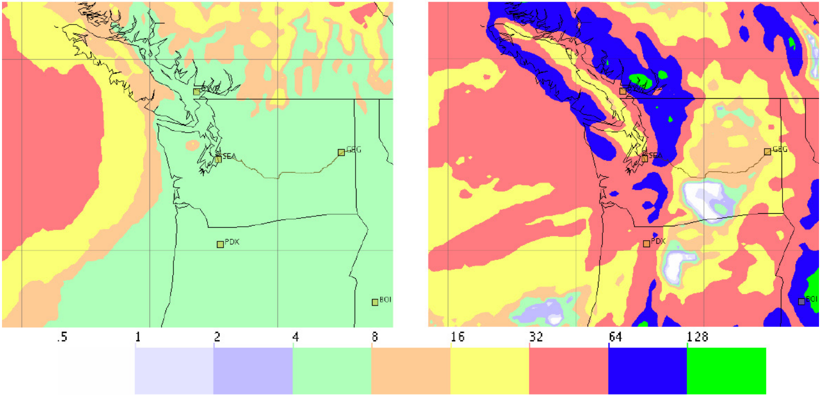
BMA Probability of Precipitation



(a) 19th May, 2003 (b) 26th January, 2003

BMA 90% Upper Bound forecast

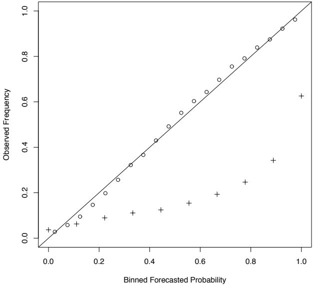
BMA 90% Upper Bound forecast



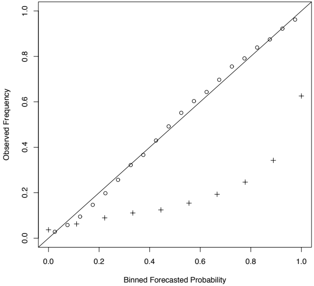
(a) 19th May, 2003 (b) 26th January, 2003

Calibration of Forecasts of the *Probability* of Precipitation

Calibration of Forecasts of the *Probability* of Precipitation

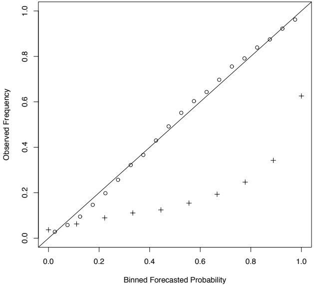


Calibration of Forecasts of the *Probability* of Precipitation



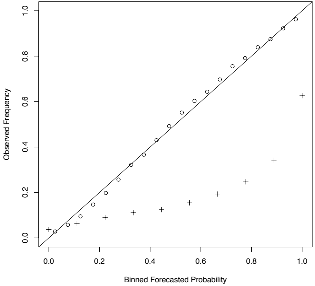
- x-axis shows the forecast probability of precipitation (PoP)

Calibration of Forecasts of the *Probability* of Precipitation



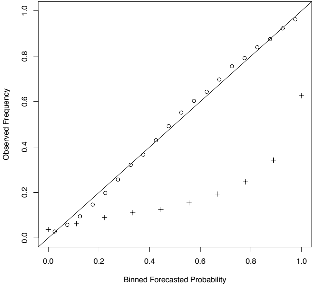
- x-axis shows the forecast probability of precipitation (PoP)
- y-axis shows the observed relative frequency of precipitation, based on 2 years of data, 2003–2004 (100K obs)

Calibration of Forecasts of the *Probability* of Precipitation



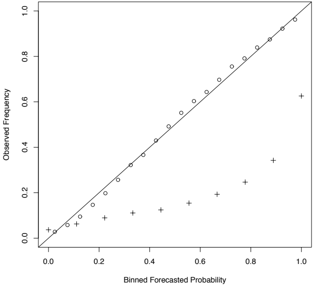
- x-axis shows the forecast probability of precipitation (PoP)
- y-axis shows the observed relative frequency of precipitation, based on 2 years of data, 2003–2004 (100K obs)
- Thus a good PoP forecast would be on the diagonal (solid line)

Calibration of Forecasts of the *Probability* of Precipitation



- *x*-axis shows the forecast probability of precipitation (PoP)
- *y*-axis shows the observed relative frequency of precipitation, based on 2 years of data, 2003–2004 (100K obs)
- Thus a good PoP forecast would be on the diagonal (solid line)
- Crosses show the proportion of the ensemble members that predict precipitation, i.e. the raw ensemble PoP forecast. Poorly calibrated

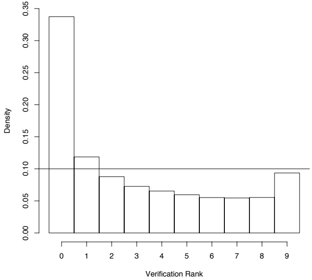
Calibration of Forecasts of the *Probability* of Precipitation



- x-axis shows the forecast probability of precipitation (PoP)
- y-axis shows the observed relative frequency of precipitation, based on 2 years of data, 2003–2004 (100K obs)
- Thus a good PoP forecast would be on the diagonal (solid line)
- Crosses show the proportion of the ensemble members that predict precipitation, i.e. the raw ensemble PoP forecast. Poorly calibrated
- Circles show the BMA PoP forecast. Much better.

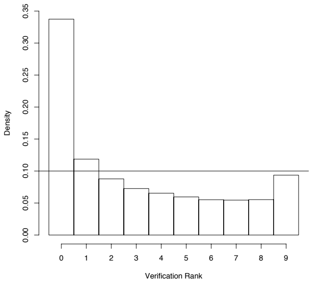
Calibration of Forecasts of the *Amount* of Precipitation

Calibration of Forecasts of the *Amount* of Precipitation



Verification rank histogram
for ensemble forecast

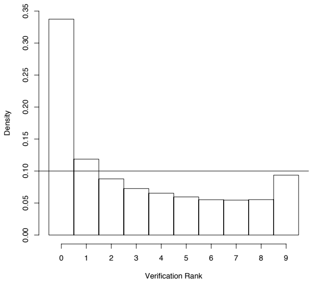
Calibration of Forecasts of the *Amount* of Precipitation



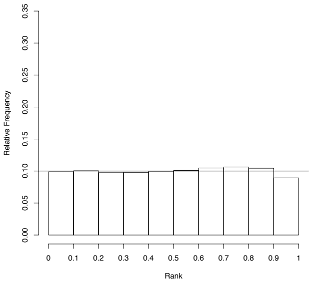
Verification rank histogram
for ensemble forecast

- The raw ensemble is poorly calibrated

Calibration of Forecasts of the *Amount* of Precipitation



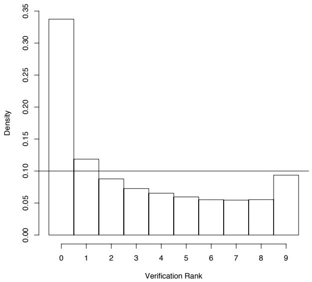
Verification rank histogram for ensemble forecast



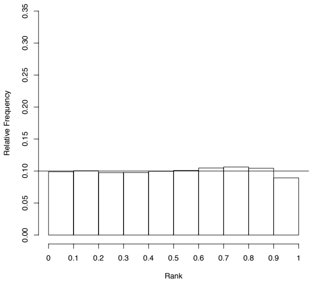
PIT histogram for BMA forecast distribution

- The raw ensemble is poorly calibrated

Calibration of Forecasts of the *Amount* of Precipitation



Verification rank histogram for ensemble forecast



PIT histogram for BMA forecast distribution

- The raw ensemble is poorly calibrated
- The BMA forecast distribution is much better calibrated

Local BMA

Local BMA

- "Traditional" BMA estimates its parameters over the whole forecast domain (in our case the Pacific Northwest)

Local BMA

- "Traditional" BMA estimates its parameters over the whole forecast domain (in our case the Pacific Northwest)
- But relative performance of models ($\Rightarrow$ BMA weights) and forecast error variance ($\Rightarrow$ BMA variance) can vary over the domain

Local BMA

- "Traditional" BMA estimates its parameters over the whole forecast domain (in our case the Pacific Northwest)
- But relative performance of models ($\Rightarrow$ BMA weights) and forecast error variance ($\Rightarrow$ BMA variance) can vary over the domain
- Local BMA estimates BMA for each gridpoint separately, using only observations at stations that are similar to the gridpoint in terms of location, elevation and land use.

Local BMA

- "Traditional" BMA estimates its parameters over the whole forecast domain (in our case the Pacific Northwest)
- But relative performance of models ($\Rightarrow$ BMA weights) and forecast error variance ($\Rightarrow$ BMA variance) can vary over the domain
- Local BMA estimates BMA for each gridpoint separately, using only observations at stations that are similar to the gridpoint in terms of location, elevation and land use.
- Much better performance on average than global BMA

Local BMA

- "Traditional" BMA estimates its parameters over the whole forecast domain (in our case the Pacific Northwest)
- But relative performance of models ($\Rightarrow$ BMA weights) and forecast error variance ($\Rightarrow$ BMA variance) can vary over the domain
- Local BMA estimates BMA for each gridpoint separately, using only observations at stations that are similar to the gridpoint in terms of location, elevation and land use.
- Much better performance on average than global BMA
- $\Longrightarrow$ feasible to use BMA for the nation

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
  - The functional is the minimum temperature over a spatial area

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
    - The functional is the minimum temperature over a spatial area
    - This helps decide whether to pretreat the road with chemicals

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
  - The functional is the minimum temperature over a spatial area
  - This helps decide whether to pretreat the road with chemicals
- Basic idea: Produce a statistical ensemble of forecasts of entire weather fields by perturbing the *outputs* from the model (i.e. the forecasts), rather than the inputs.

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
  - The functional is the minimum temperature over a spatial area
  - This helps decide whether to pretreat the road with chemicals
- Basic idea: Produce a statistical ensemble of forecasts of entire weather fields by perturbing the *outputs* from the model (i.e. the forecasts), rather than the inputs.
- For the simplest case of just one forecast (no ensemble):

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
  - The functional is the minimum temperature over a spatial area
  - This helps decide whether to pretreat the road with chemicals
- Basic idea: Produce a statistical ensemble of forecasts of entire weather fields by perturbing the *outputs* from the model (i.e. the forecasts), rather than the inputs.
- For the simplest case of just one forecast (no ensemble):
  - A spatial geostatistical model is used for the forecast errors

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
  - The functional is the minimum temperature over a spatial area
  - This helps decide whether to pretreat the road with chemicals
- Basic idea: Produce a statistical ensemble of forecasts of entire weather fields by perturbing the *outputs* from the model (i.e. the forecasts), rather than the inputs.
- For the simplest case of just one forecast (no ensemble):
  - A spatial geostatistical model is used for the forecast errors
  - A fast and exact simulation method is used to generate a statistical ensemble of forecasts (the circulant embedding method)

Probabilistic Forecasting of Entire Weather Fields

A Very High-Dimensional Quantity of Interest (10,000 dimensions)

- Important for forecasting functionals of a weather field
- Desirable for route planning in aviation
- Example: What is the probability that there will be freezing precipitation somewhere on the I-90 freeway in Washington State?
    - The functional is the minimum temperature over a spatial area
    - This helps decide whether to pretreat the road with chemicals
- Basic idea: Produce a statistical ensemble of forecasts of entire weather fields by perturbing the *outputs* from the model (i.e. the forecasts), rather than the inputs.
- For the simplest case of just one forecast (no ensemble):
    - A spatial geostatistical model is used for the forecast errors
    - A fast and exact simulation method is used to generate a statistical ensemble of forecasts (the circulant embedding method)
    - The result: the Geostatistical Output Perturbation (GOP) Method

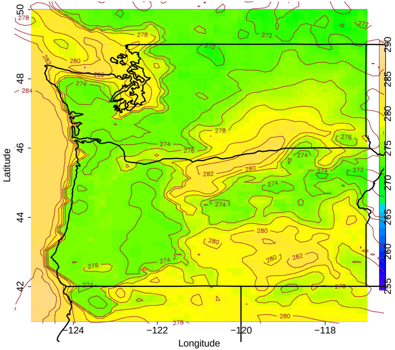
Example

Example

Gridded Forecast for January 12, 2002

Example

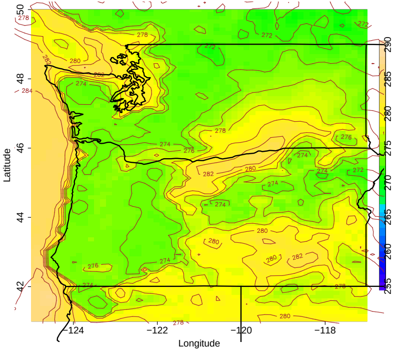
Gridded Forecast for January 12, 2002



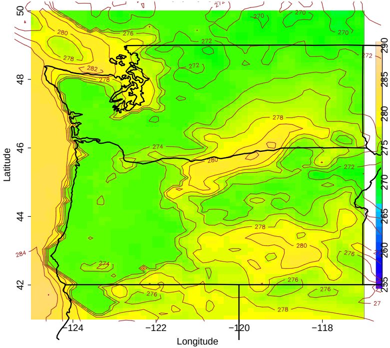
Gridded forecast

Example

Gridded Forecast for January 12, 2002



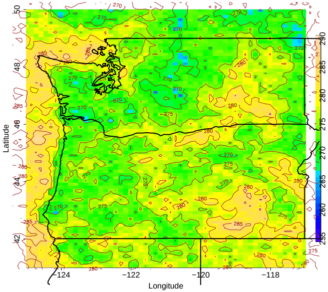
Gridded forecast



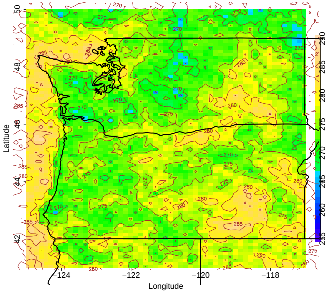
Bias-corrected

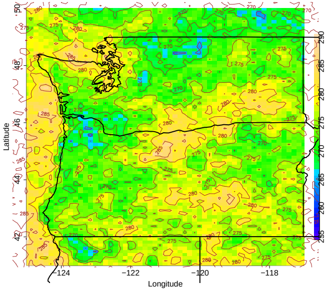
Sample from the Forecast Predictive Distribution

Sample from the Forecast Predictive Distribution

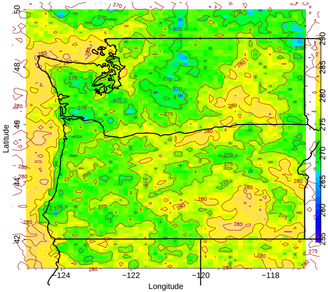


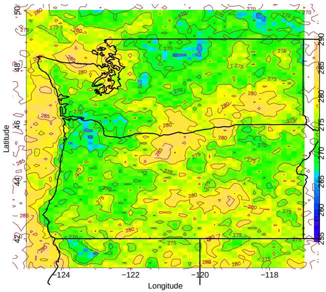
Sample from the Forecast Predictive Distribution

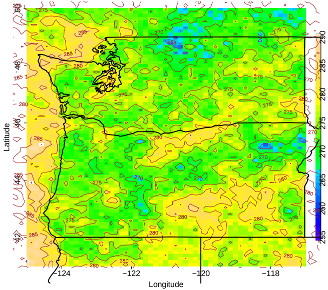




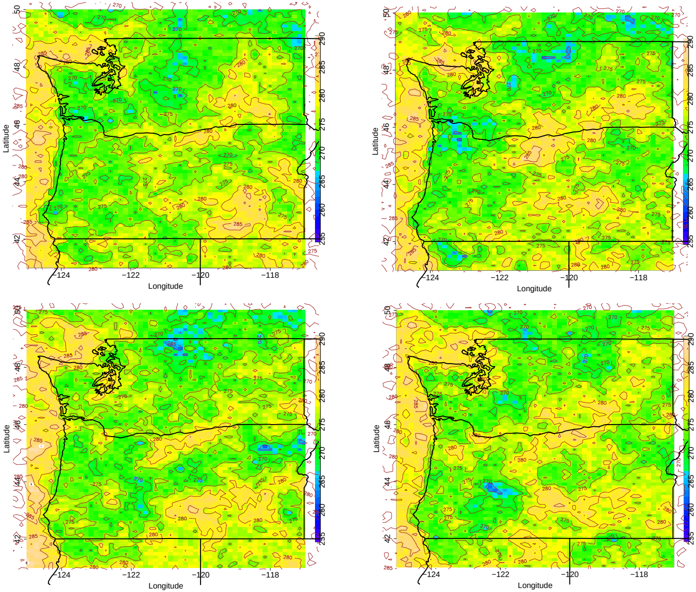
Sample from the Forecast Predictive Distribution







Sample from the Forecast Predictive Distribution



Some References

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.

- Similar in practice to BMA, giving similar results.

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.

- Similar in practice to BMA, giving similar results.
- Not so similar in concept

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.

- Similar in practice to BMA, giving similar results.
- Not so similar in concept

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.
- Similar in practice to BMA, giving similar results.
- Not so similar in concept

Overview paper:

Some References

MWR = Monthly Weather Review
JASA = Journal of the American Statistical Association

BMA for temperature:

- Raftery, Gneiting et al (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *MWR* 133: 1155–1174.
- Wilson, Beauregard, Raftery, Verret (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *MWR* 135: 1364–1385.

BMA for precip and wind:

- Sloughter, Raftery, Gneiting, Fraley (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *MWR* 135: 3209–3220.
- Sloughter, Gneiting, Raftery (2009). Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging. *JASA*, to appear.

Probabilistic forecasting of weather fields:

- Gel, Raftery, Gneiting (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method. *JASA* 99: 575-590.
- Berrocal, Raftery, Gneiting (2007). Combining spatial statistical and ensemble information in probabilistic weather forecasts. *MWR* 135: 1386–1402.
- Berrocal, Raftery, Gneiting (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Ann. Appl. Stat.* 2: 1170–1193.

EMOS:

- Gneiting, Raftery et al (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *MWR* 133: 1098–1118.
- Similar in practice to BMA, giving similar results.
- Not so similar in concept

Overview paper:

- Gneiting and Raftery (2005). Weather forecasting with ensemble methods. *Science* 310: 248–249.

Summary

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu
- We can work with NWS to develop probabilistic forecasts for

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu
- We can work with NWS to develop probabilistic forecasts for
  - other parameters relevant to aviation

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu
- We can work with NWS to develop probabilistic forecasts for
  - other parameters relevant to aviation
  - all levels of the atmosphere

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu
- We can work with NWS to develop probabilistic forecasts for
  - other parameters relevant to aviation
  - all levels of the atmosphere
  - the nation (and beyond)

Summary

- Forecast ensembles show a spread-skill relationship, but still tend to be underdispersed
- Bayesian model averaging is a statistical way of getting sharp calibrated probabilistic forecasts from an ensemble, that honor the spread-skill relationship
- In experiments with forecasting temperature, precip and wind, BMA has consistently been calibrated, sharp, and has given good deterministic forecasts
- BMA has been extended to provide forecasts of entire meteorological fields (Spatial BMA)
- Free R packages: EnsembleBMA, ProbForecastGOP
- Web sites: www.stat.washington.edu/raftery/Research/dsm.html
  www.stat.washington.edu/MURI
  www.probcast.washington.edu
  bma.apl.washington.edu
- We can work with NWS to develop probabilistic forecasts for
  - other parameters relevant to aviation
  - all levels of the atmosphere
  - the nation (and beyond)
  - to produce the 4-d probabilistic forecasting cube