UNCLASSIFIED

A Practical Method for Calibration of Ensemble Spread for Representation of Meteorological Uncertainty in Atmospheric Transport and Dispersion Models

Walter Kolczynski, Jr.

Dave Stauffer

Sue Ellen Haupt

Chemical and Biological Information Systems Conference & Exhibition
11 January 2007
Austin, TX

Overview

- Goals
- Motivation
- Methodology
- Linear Variance Calibration (UUE,VVE,UVE)
- Linear Covariance Calibration
- Covariance/Distance Relation (SLE)
- Conclusions
- Future Work

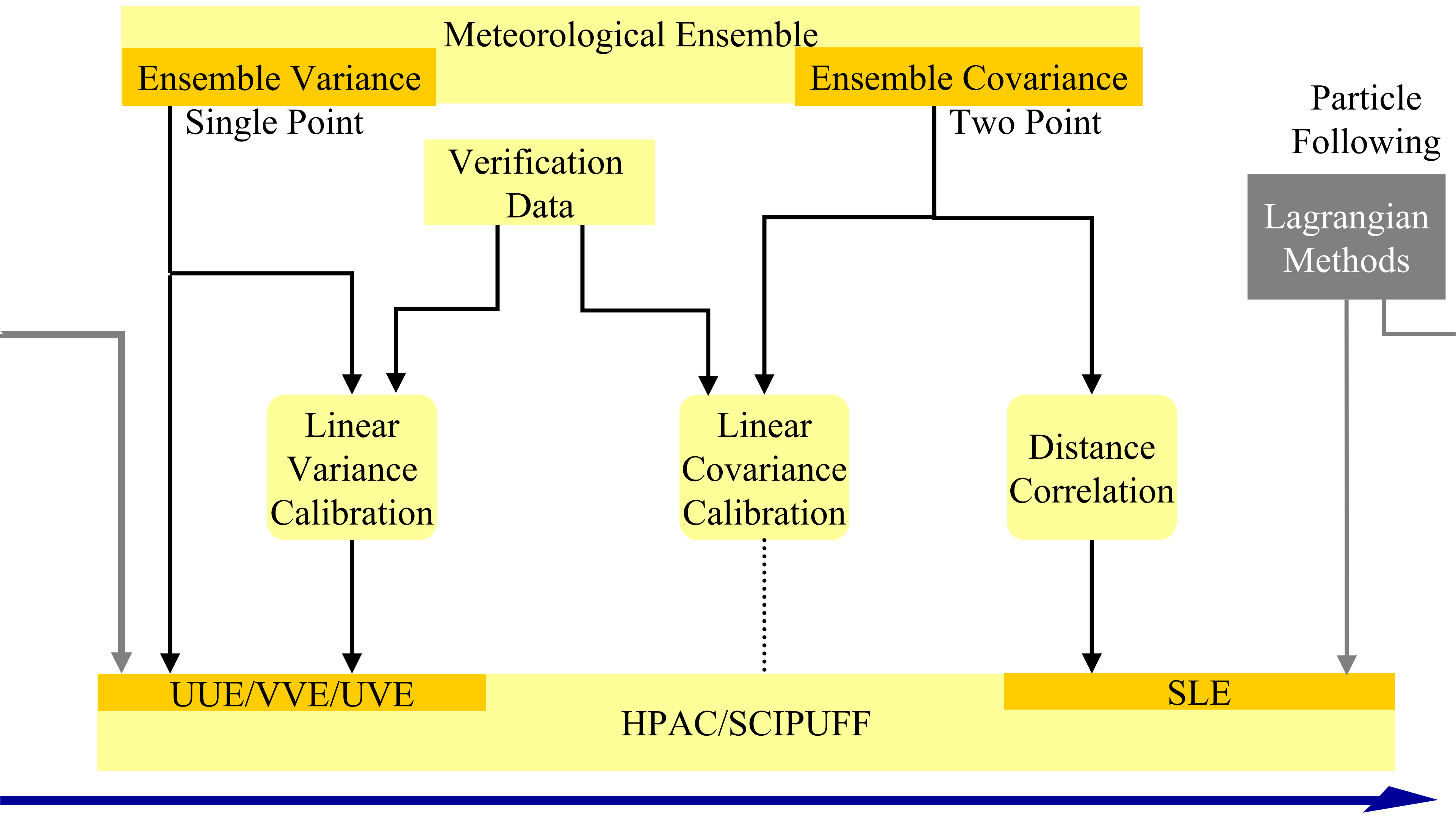
Goals

- Use an ensemble of MET models to provide HPAC/SCIPUFF with MET uncertainty information to account for uncertainty in AT&D computations
- Study applicability of a new efficient linear calibration method to compute these MET uncertainty inputs to HPAC/SCIPUFF

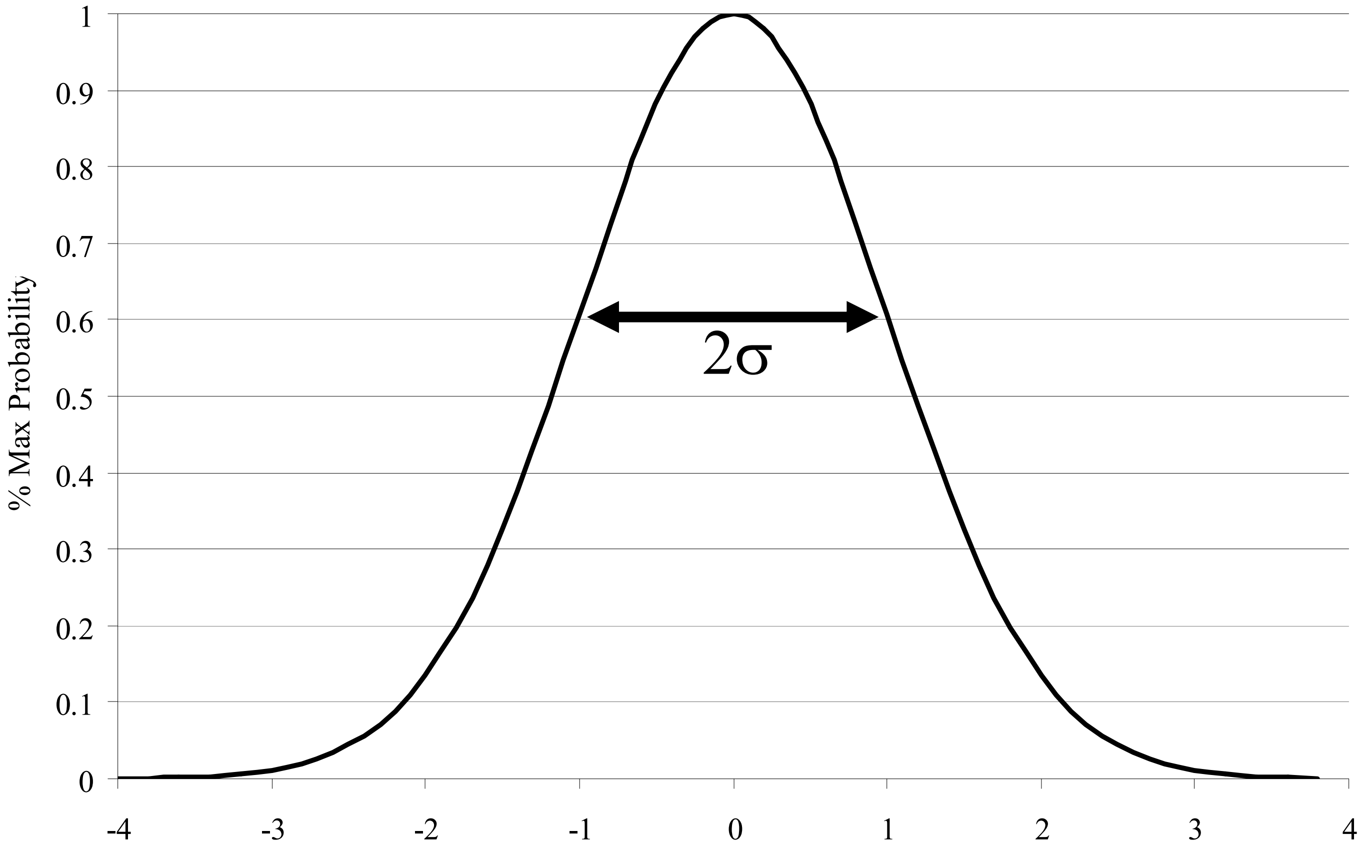
Motivation

- The variability and correlation of MET errors have important implications to AT&D predictions
- MET uncertainty information can already be input to SCIPUFF through wind variance matrices (UUE, VVE, UVE) and the Lagrangian length scale (SLE)
- Running an ensemble of AT&D models based on the MET ensemble to represent the uncertainty may not be practical for operations
- When an ensemble of AT&D models is not possible, an efficient way to pass MET uncertainty information from the MET ensemble into the single AT&D model solution is needed

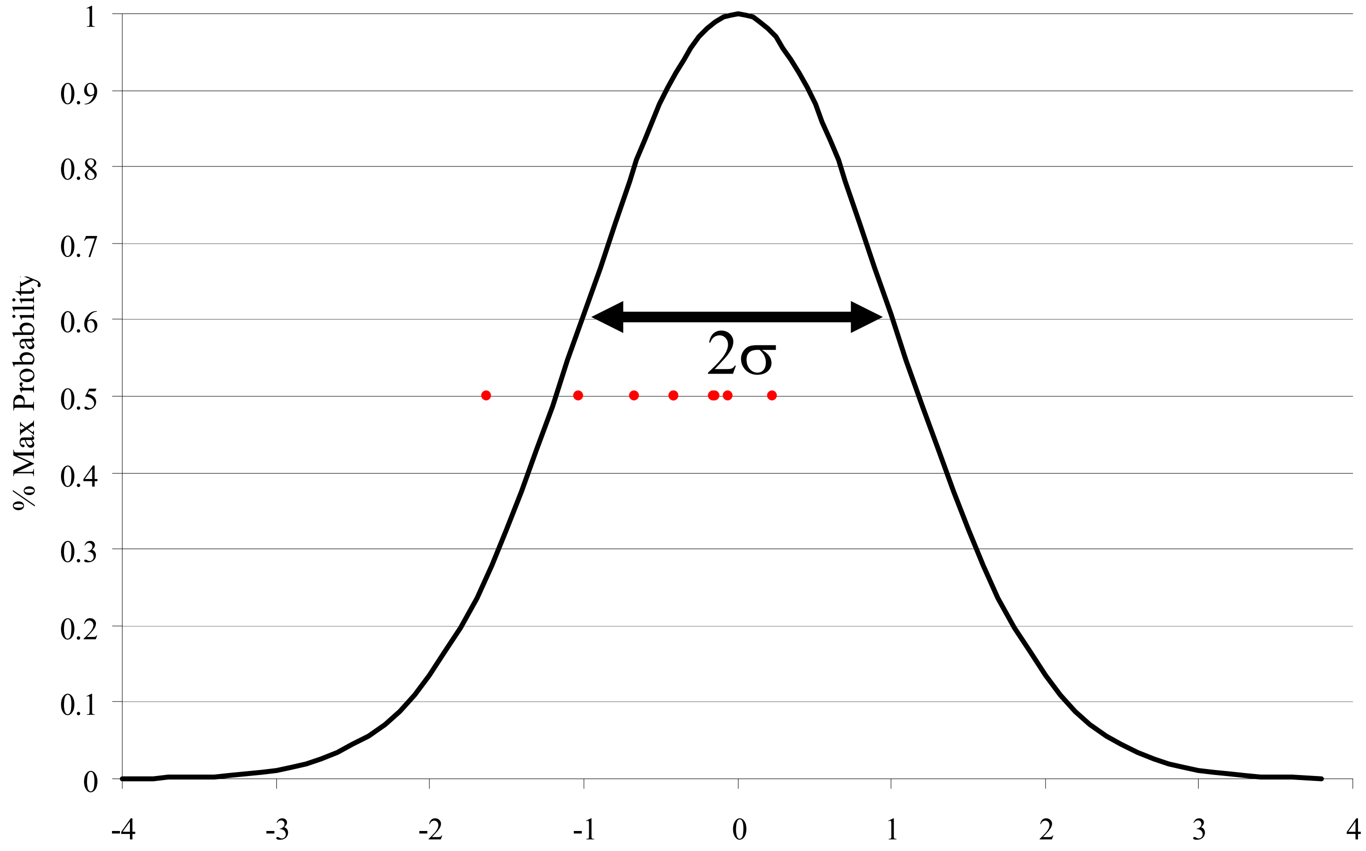
Uncertainty Information Pathways



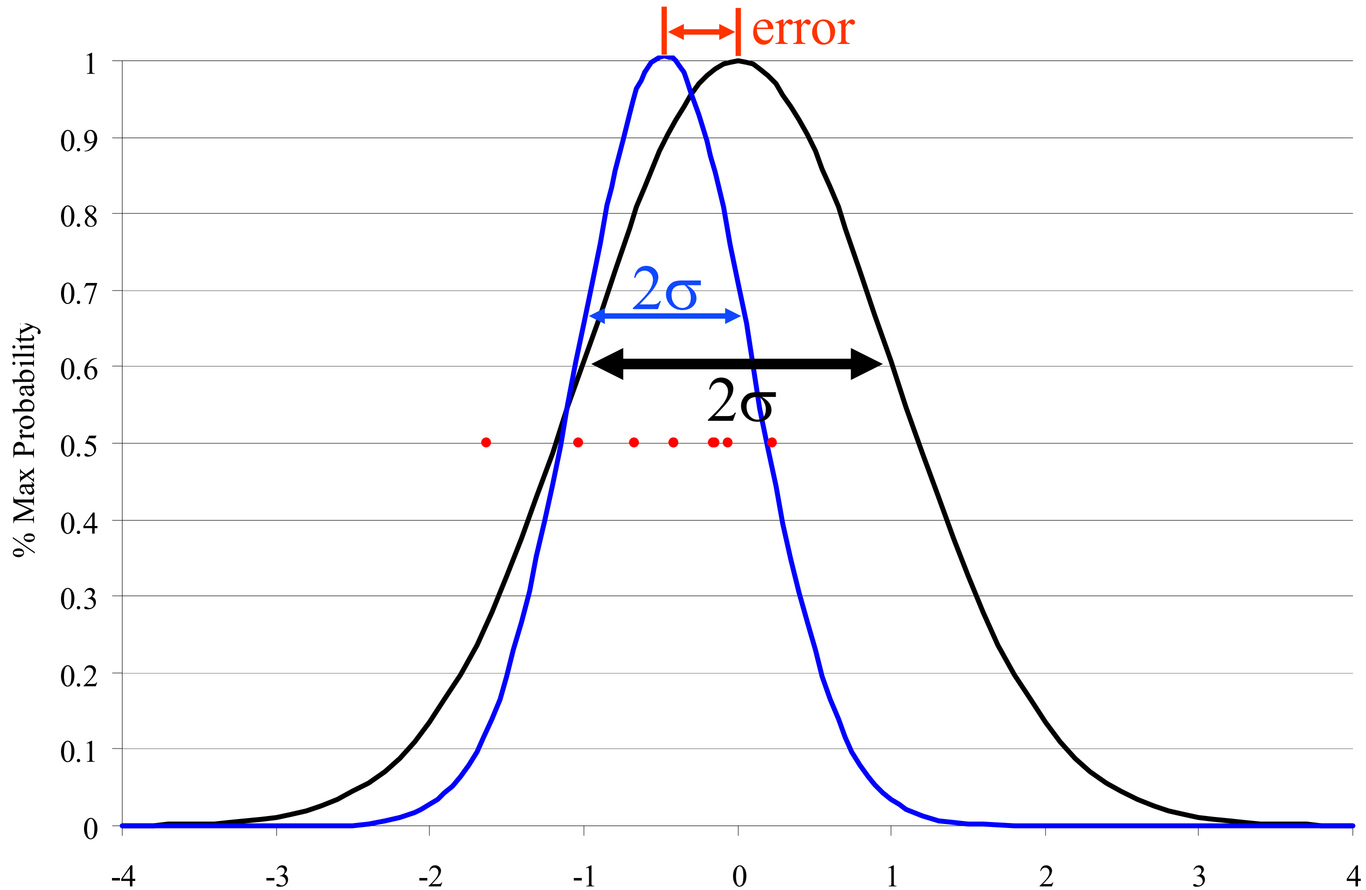
Motivation



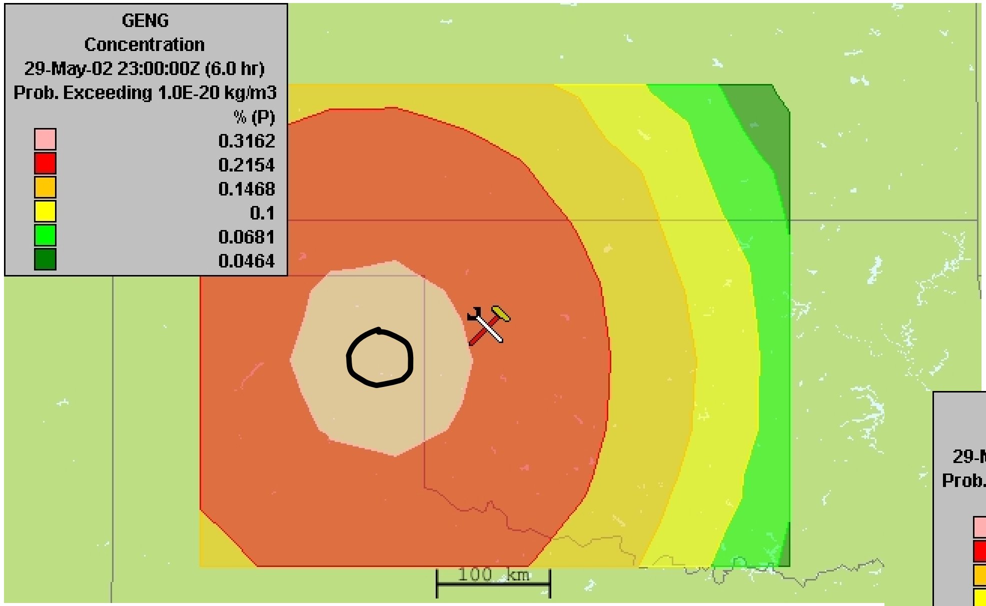
Motivation



Motivation

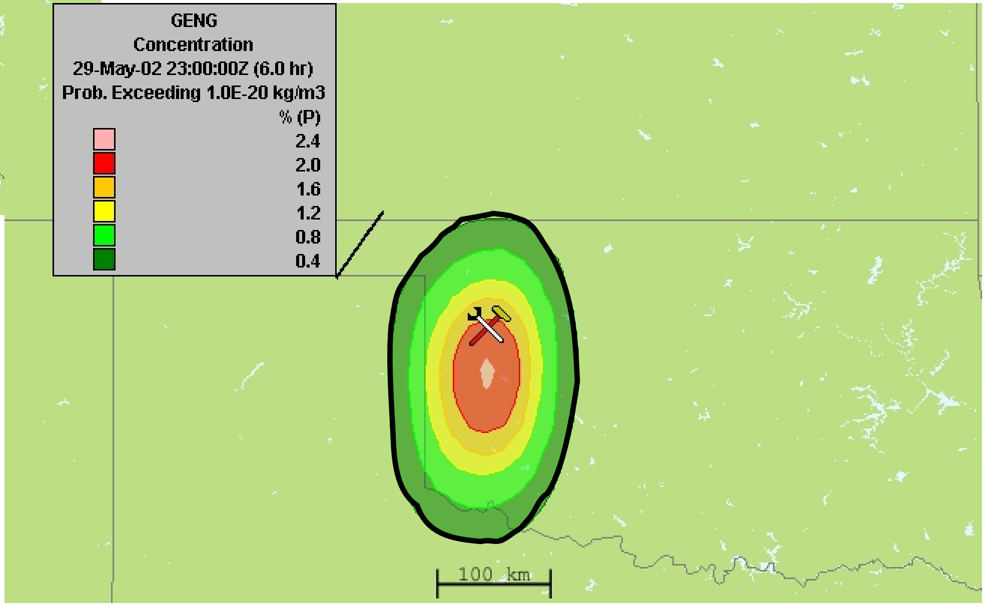

error
2σ
2σ
% Max Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

Motivation



UUE = 100 m^2/s^2

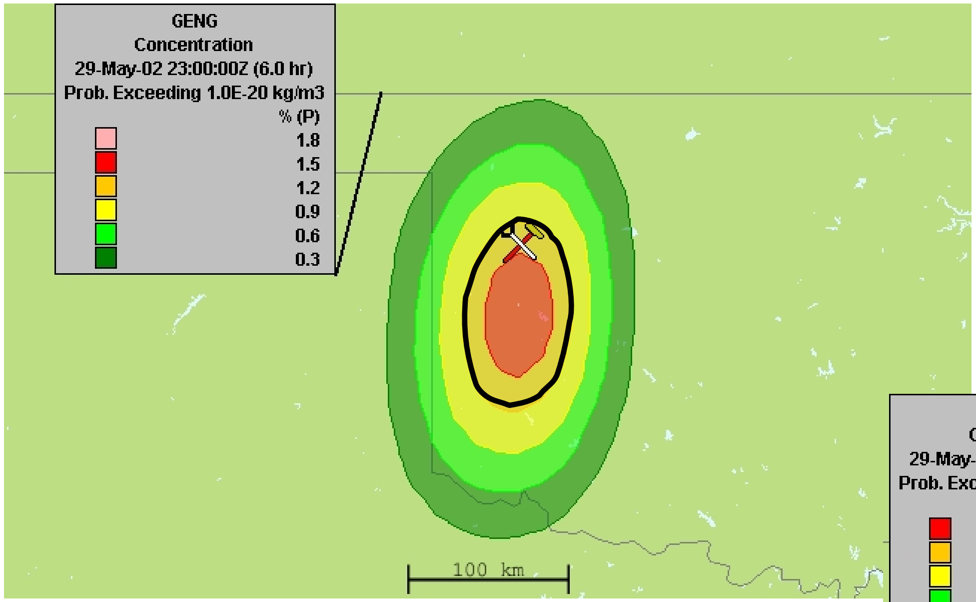
Constant SLE = 109 km



Probability of concentration Greater than 10^{-20} kg/m^3

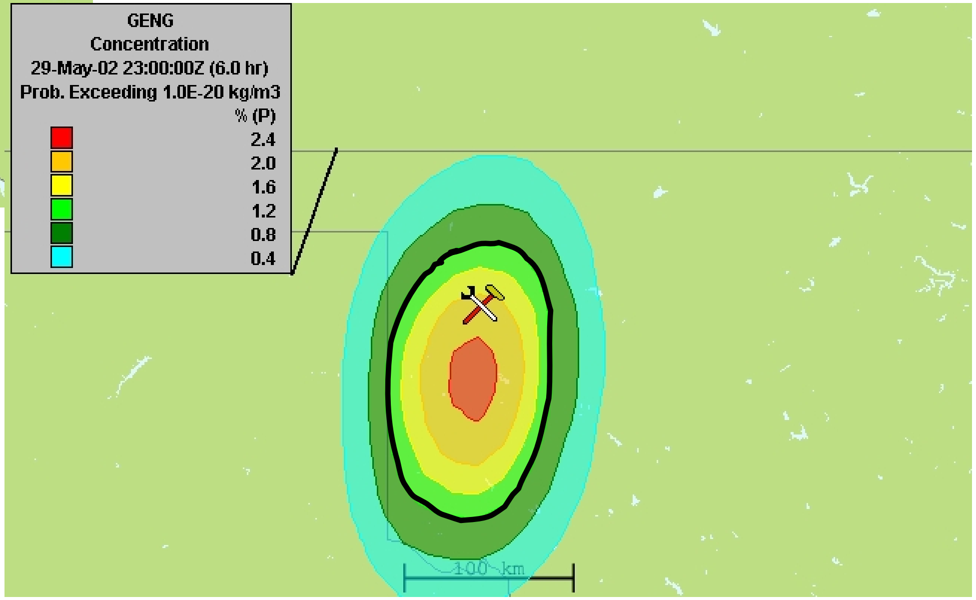
UUE = 1 m^2/s^2

Motivation



SLE = 27 km

Constant UUE = 0.4 m^2/s^2



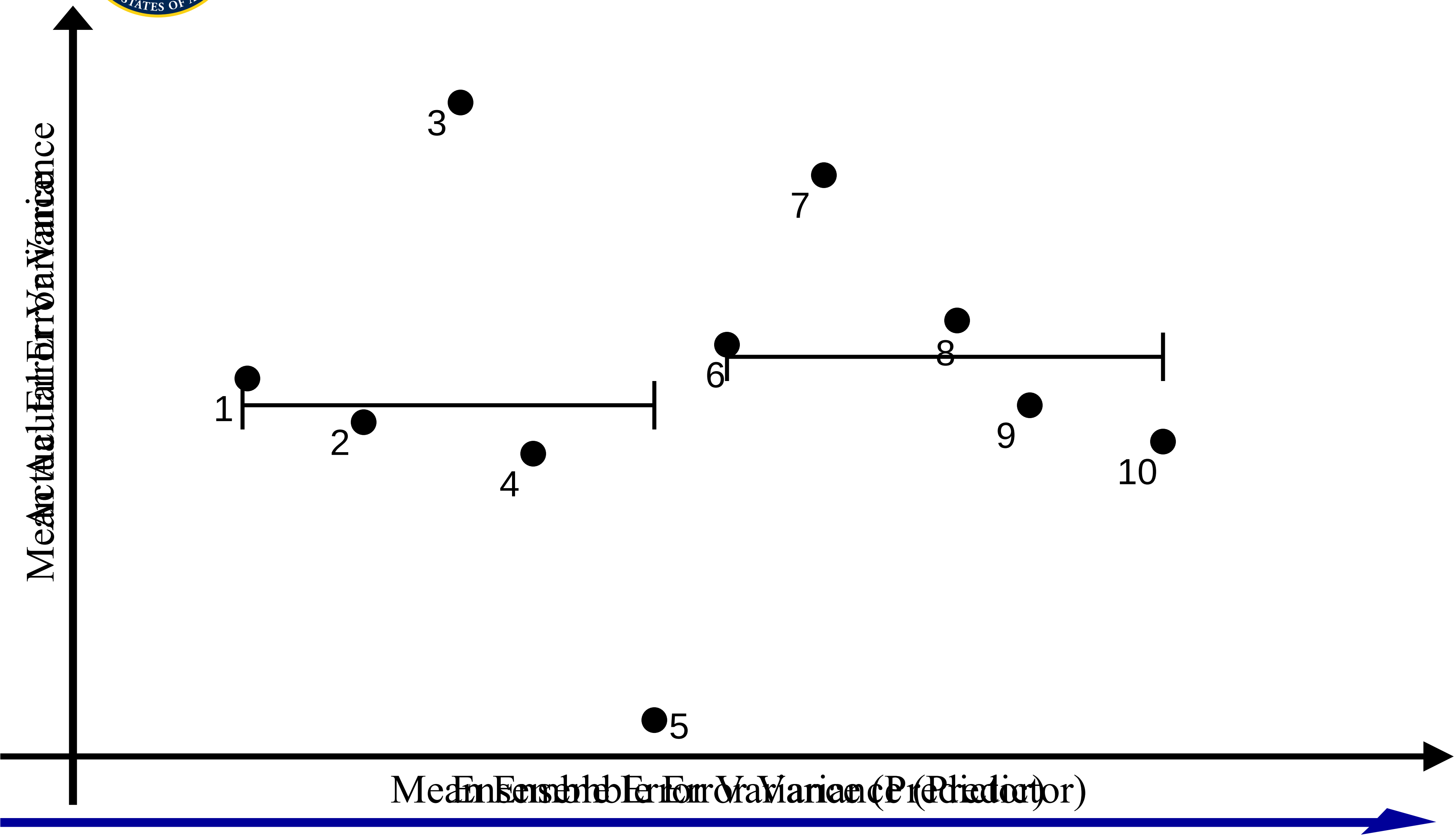
Probability of concentration Greater than 10^{-20} kg/m^3

SLE = 164 km

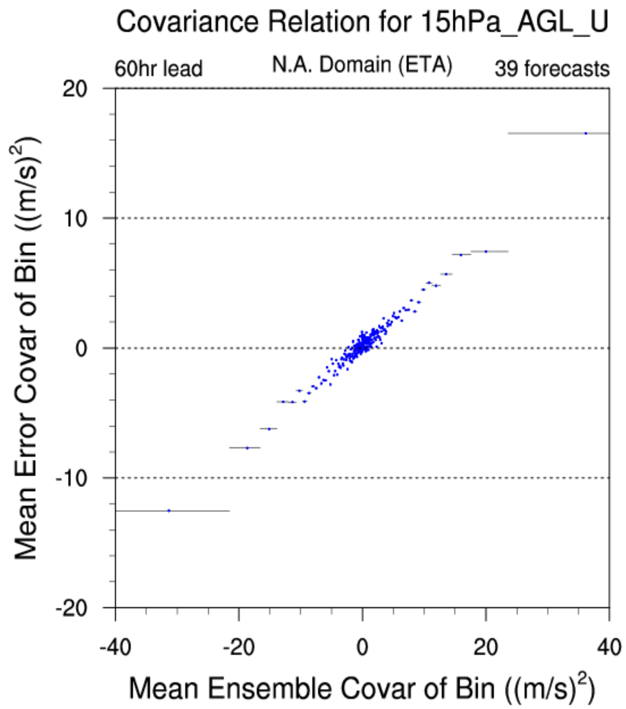
Methodology

- Use bootstrap sampling
- Bin results based on the predictor value, following Roulston (2005)
- Analyze plots for a simple relationship (linear) to be used as a calibration to ensemble data
- Utilize existing available ensemble data (SREF-ETA) to assess the promise of the technique

Methodology – Binning Procedure



Methodology – Binning Procedure



Points binned into groups of 1000

Methodology – SREF Ensemble Data

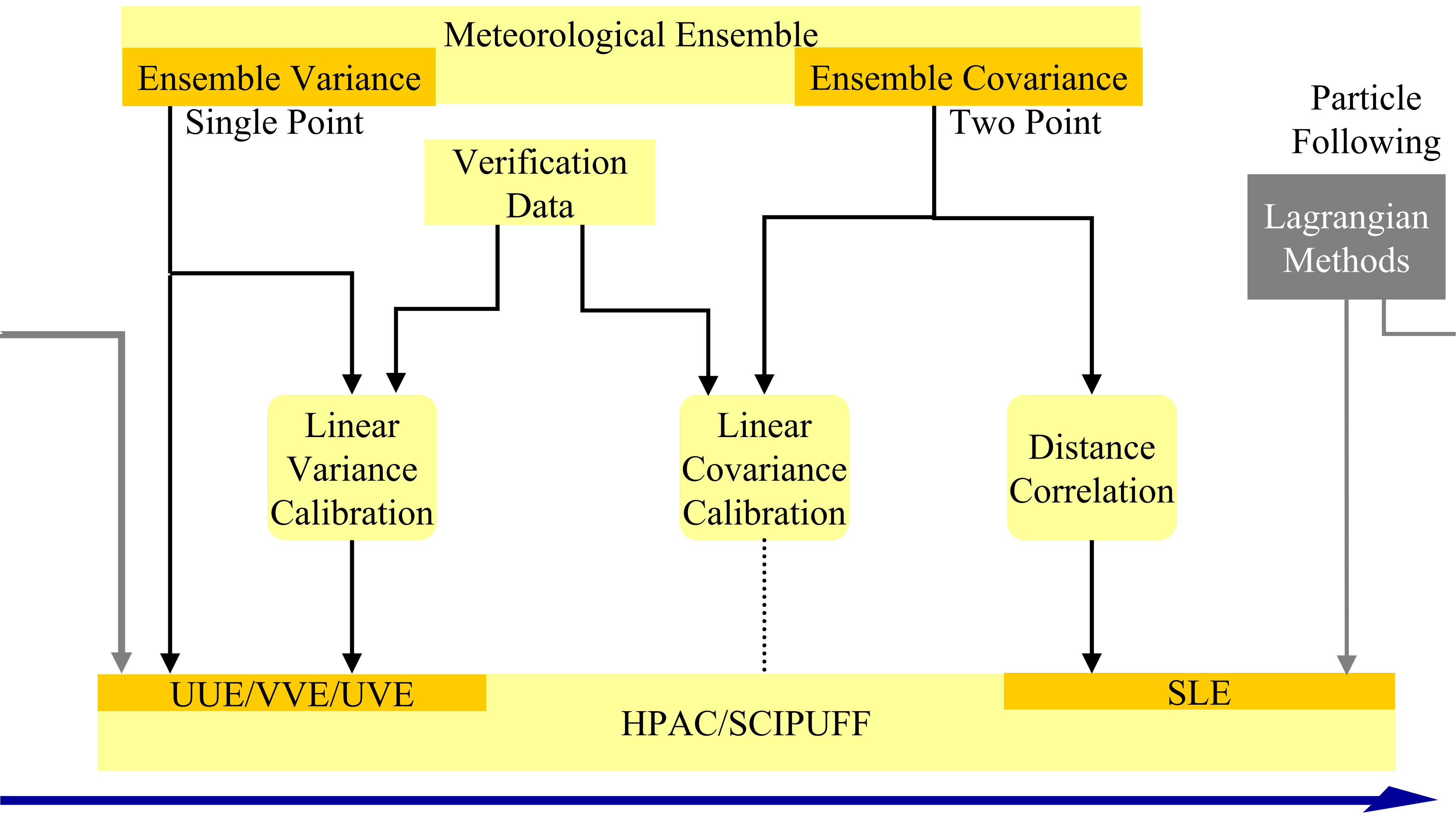
- 25 August to 15 September 2004 (22 days, 44 ensemble sets)
- Two runs per day (09 UTC and 21UTC)
- Forecasts for 12, 24, 36, 48 and 60 hrs considered
- 10 ETA members (32-km resolution)
- 0-hr forecast of ETA-ctl1 used as verification
- U and V winds “15 hPa AGL” (~150 m) used

Methodology – SREF ETA Members

Case Name	Convection	Microphysics	Breeding IC
Eta_ctl1	BMJ	OpFer	-
Eta_ctl2	KF	OpFer	-
Eta_n1	BMJ	OpFer	Eta_ctl1
Eta_n2	KF	OpFer	Eta_ctl2
Eta_n3	BMJ-SAT	OpFer	Eta_ctl1
Eta_n4	KF-DET	ExFer	Eta_ctl2
Eta_p1	BMJ	OpFer	Eta_ctl1
Eta_p2	KF	OpFer	Eta_ctl2
Eta_p3	BMJ-SAT	OpFer	Eta_ctl1
Eta_p4	KF-DET	ExFer	Eta_ctl2

BMJ: Betts-Miller-Janic KF: Kain-Fritsch SAT: Sat. Profile DET: Full Detrainment
OpFer: Operational Ferrier Micro ExFer: Experimental Ferrier Micro

Linear Variance Calibration



Linear Variance Calibration

$$EVar\left(s(ij)\right) = \frac{1}{N}\sum_{m=1}^{N}\left(s_m(ij) - \overline{s(ij)}\right)^2$$

$$AVar\left(s(ij)\right) = \left(s_v(ij) - \overline{s(ij)}\right)^2$$

$s(ij)$ is the value of a scalar field at point (i,j)

$EVar\left(s(ij)\right)$ is the Ensemble Variance of s at point (i,j)

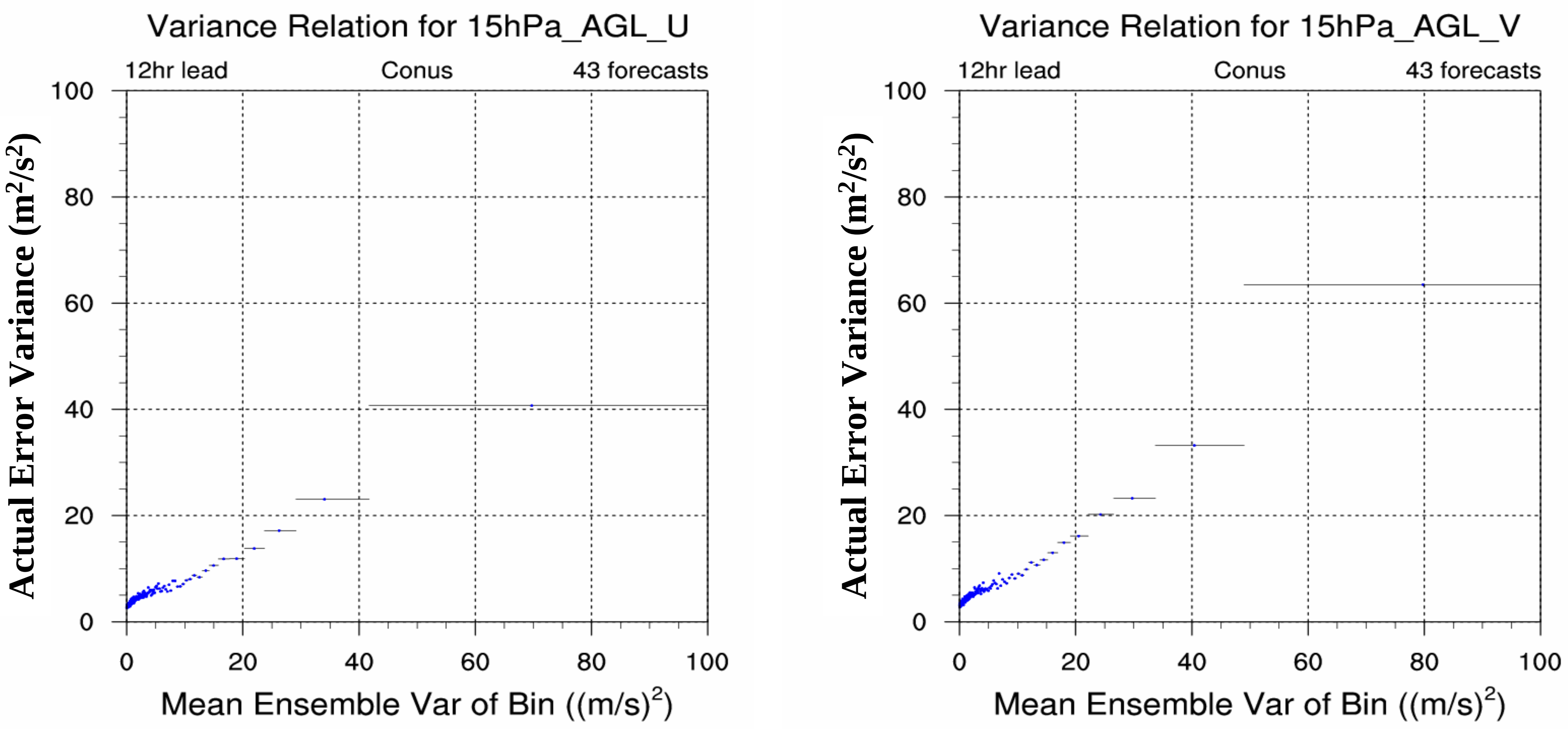
$AVar\left(s(ij)\right)$ is the Actual Variance of s at point (i,j)

N is the number of ensemble members

s_m is the scalar value of a single ensemble member

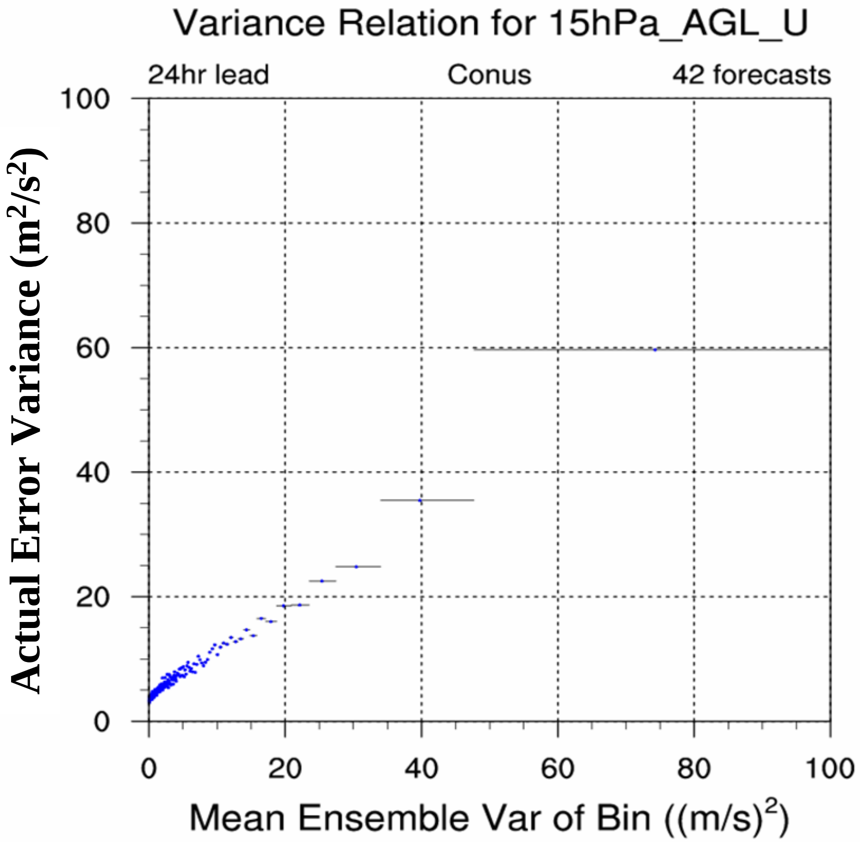
s_v is the scalar value of the verification

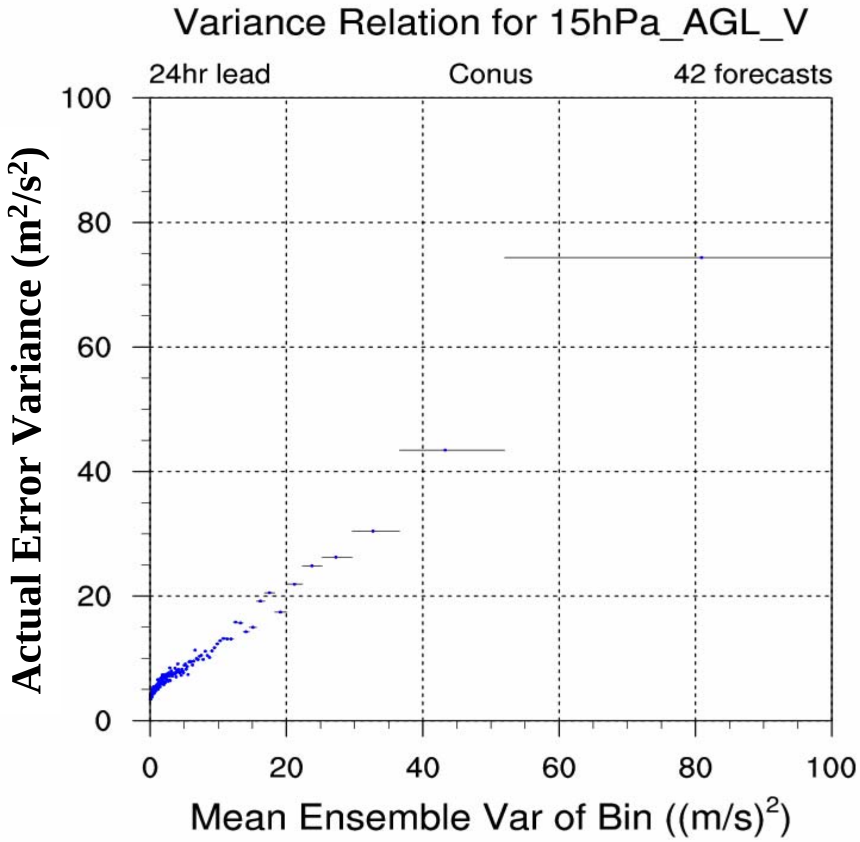
Linear Variance Calibration



12 Hour Forecast

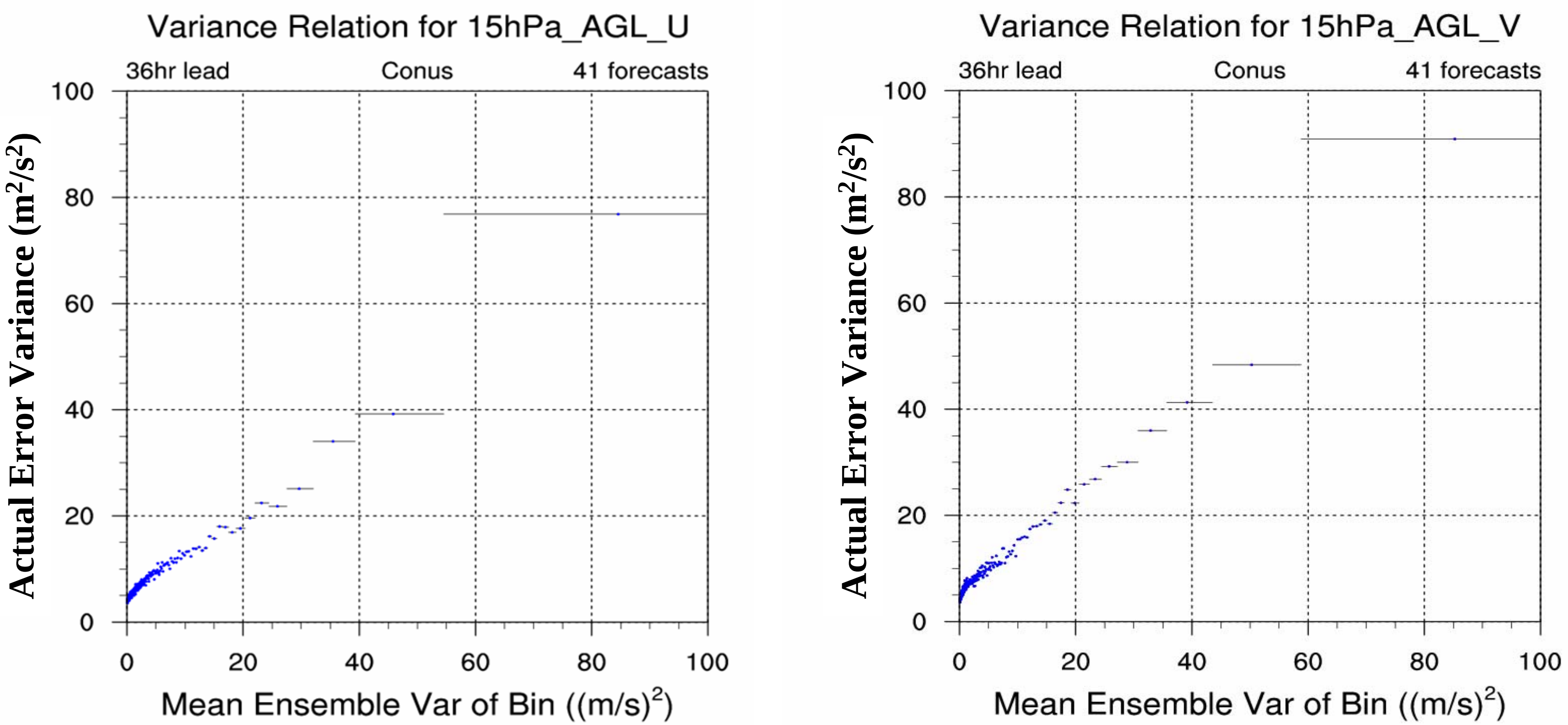
Linear Variance Calibration





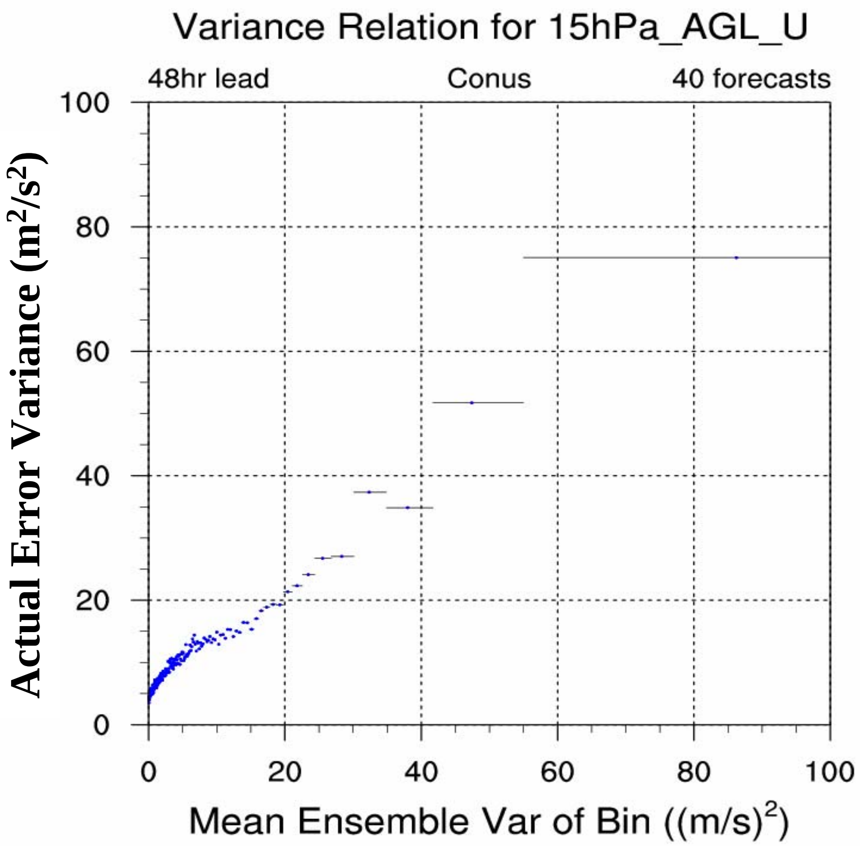
24 Hour Forecast

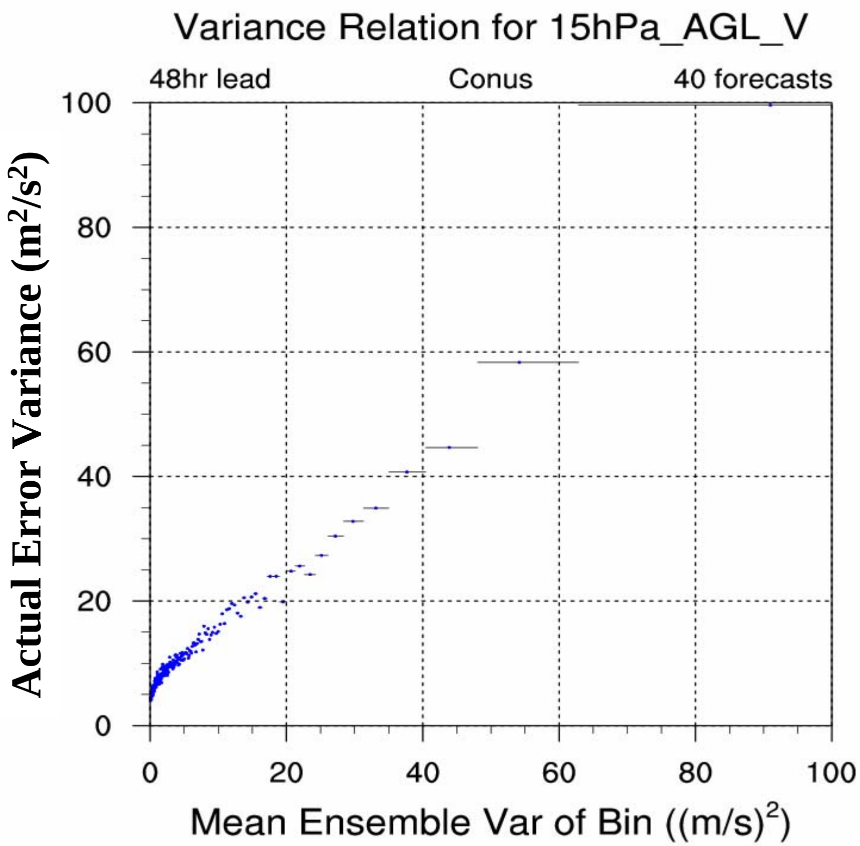
Linear Variance Calibration



36 Hour Forecast

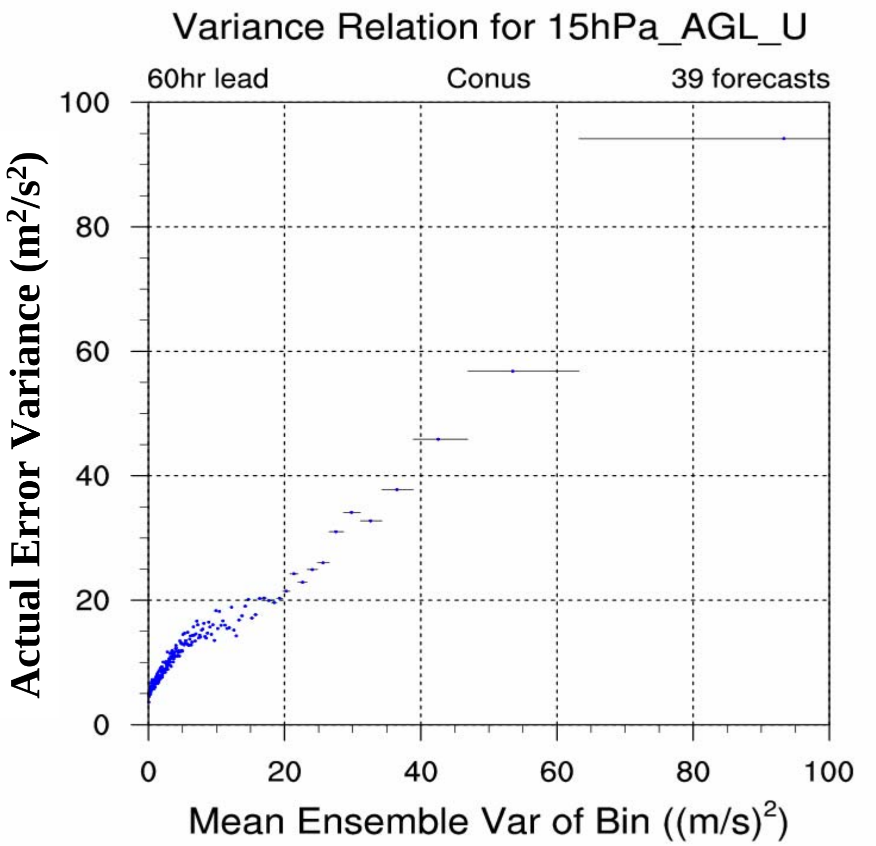
Linear Variance Calibration

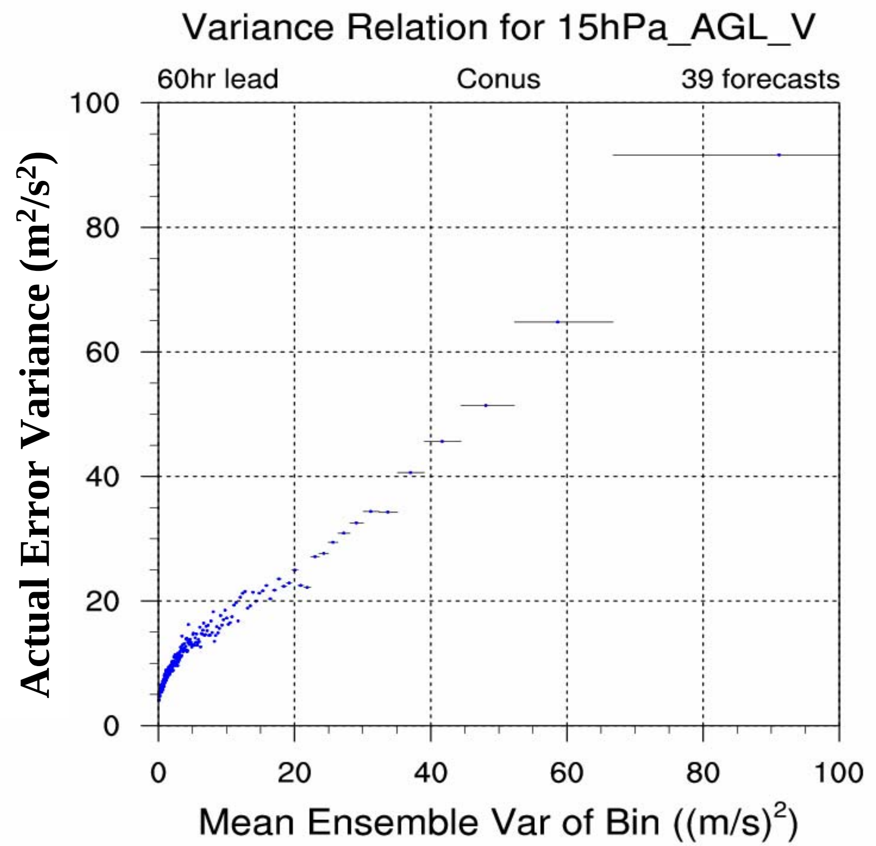




48 Hour Forecast

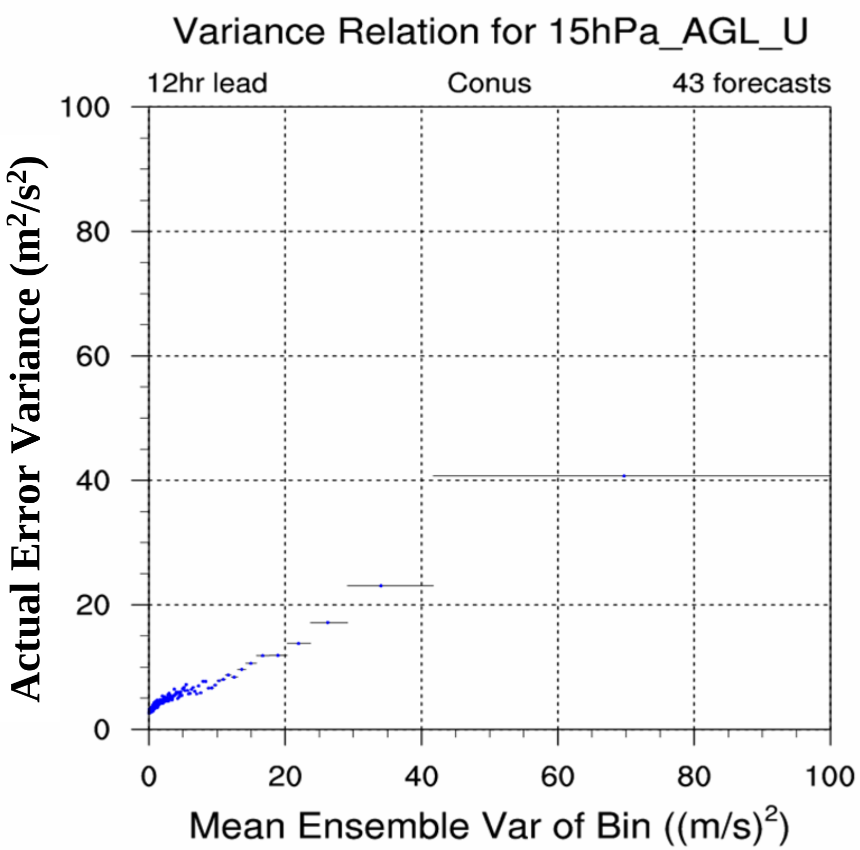
Linear Variance Calibration

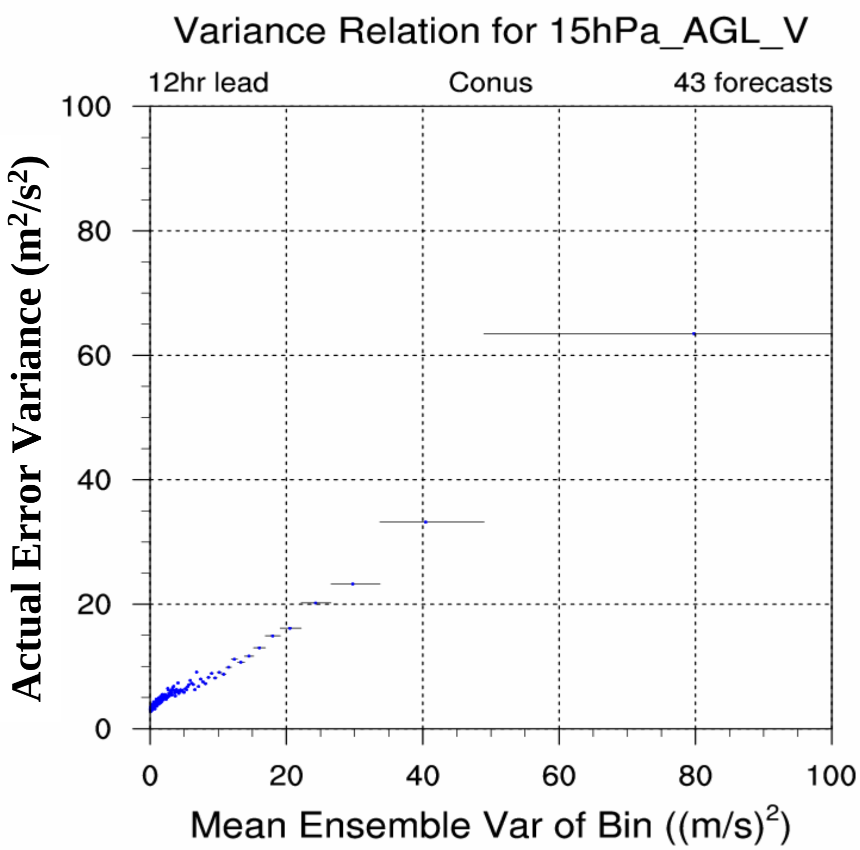




60 Hour Forecast

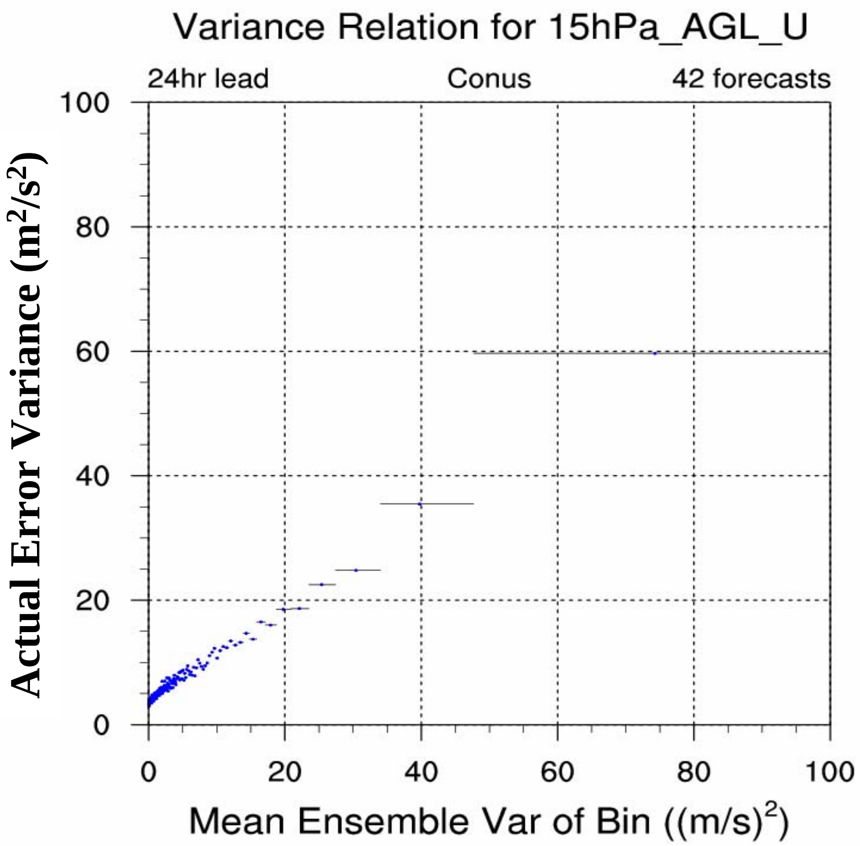
Linear Variance Calibration

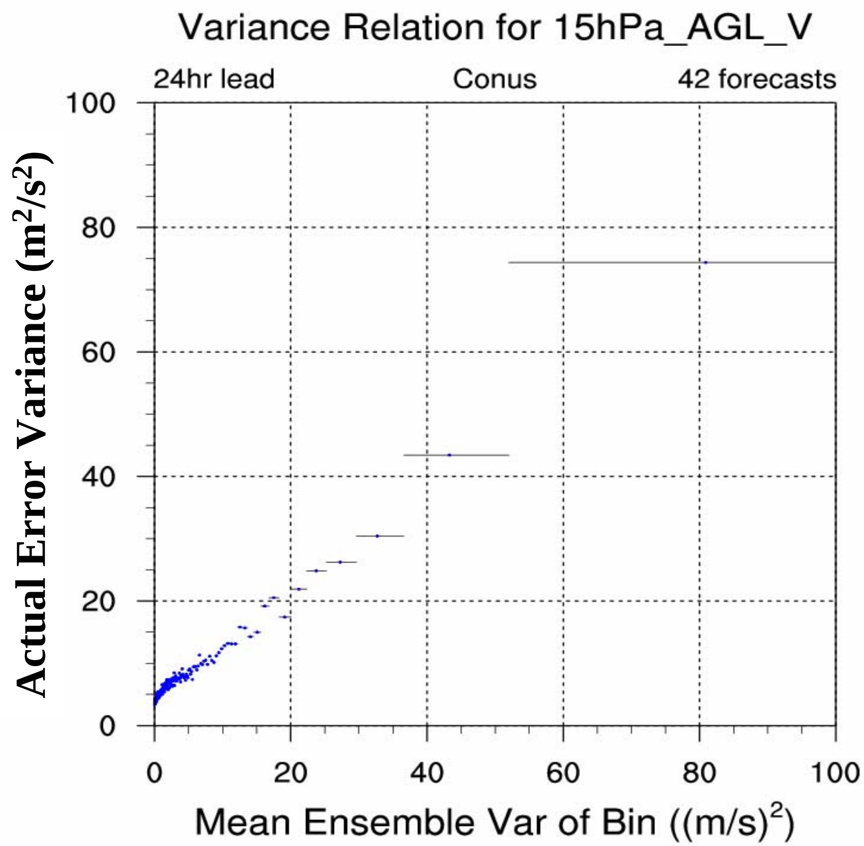




12 Hour Forecast

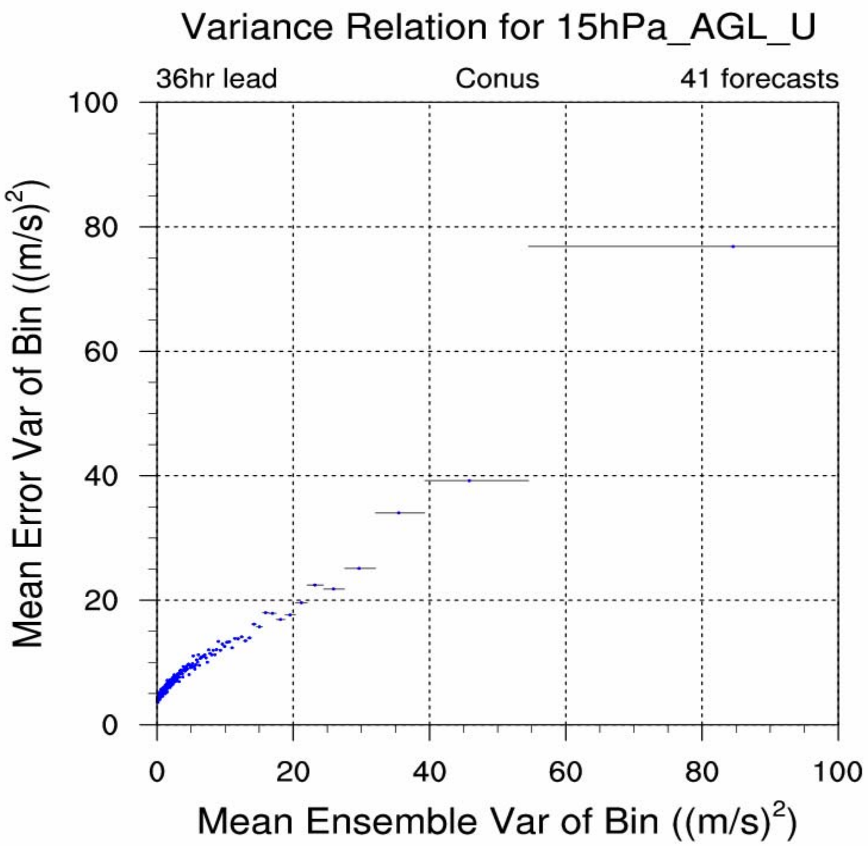
Linear Variance Calibration

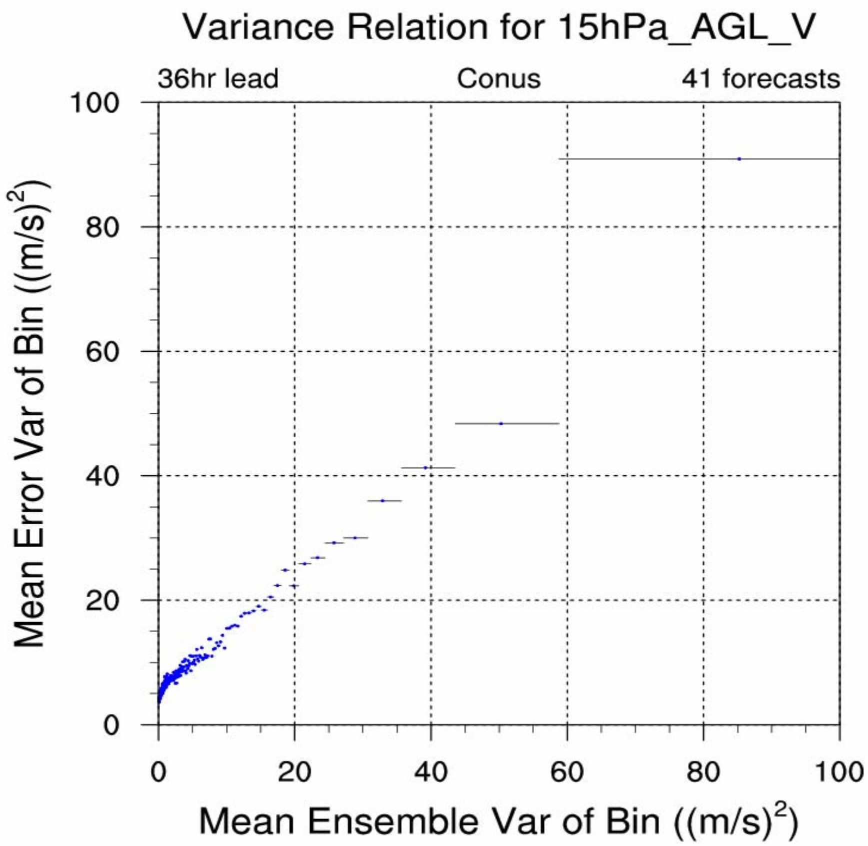




24 Hour Forecast

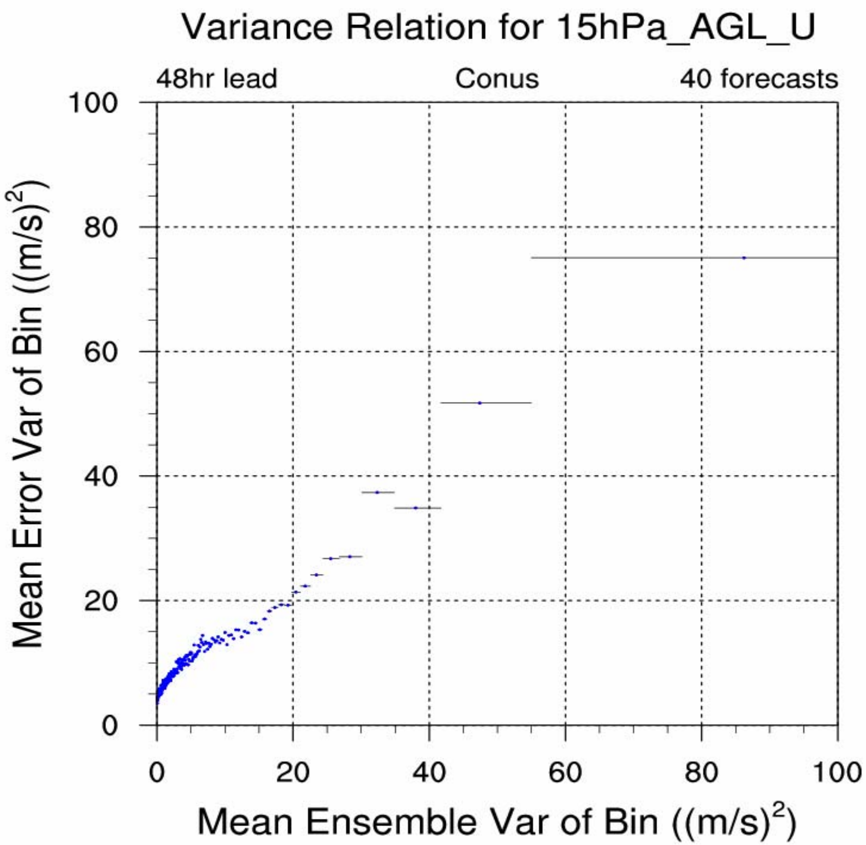
Linear Variance Calibration

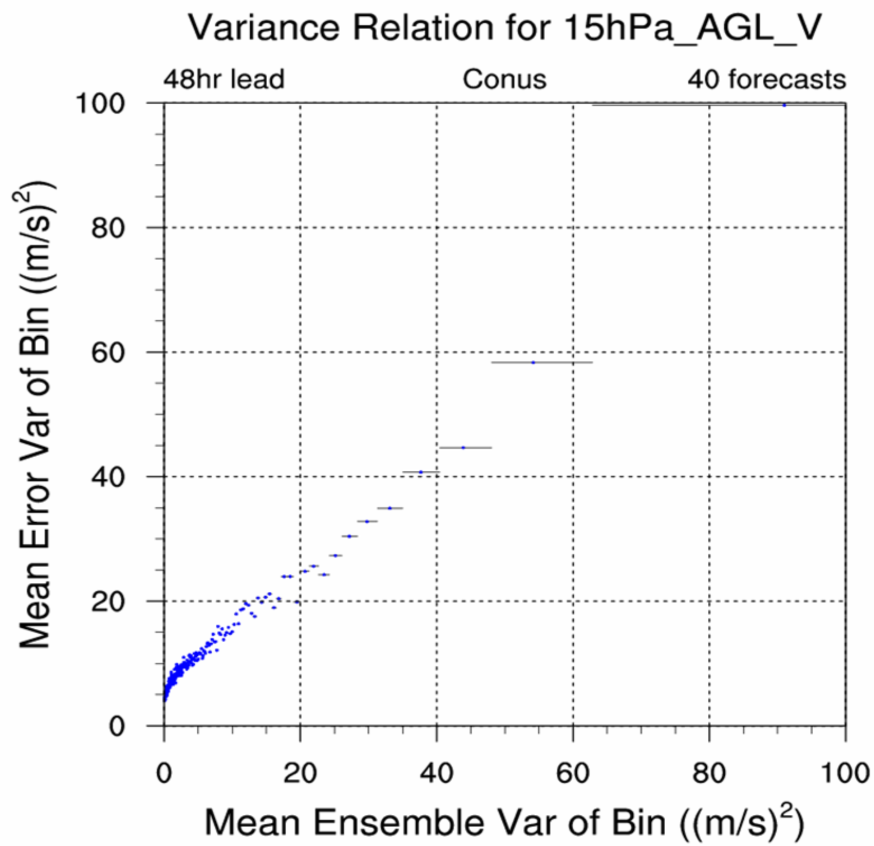




36 Hour Forecast

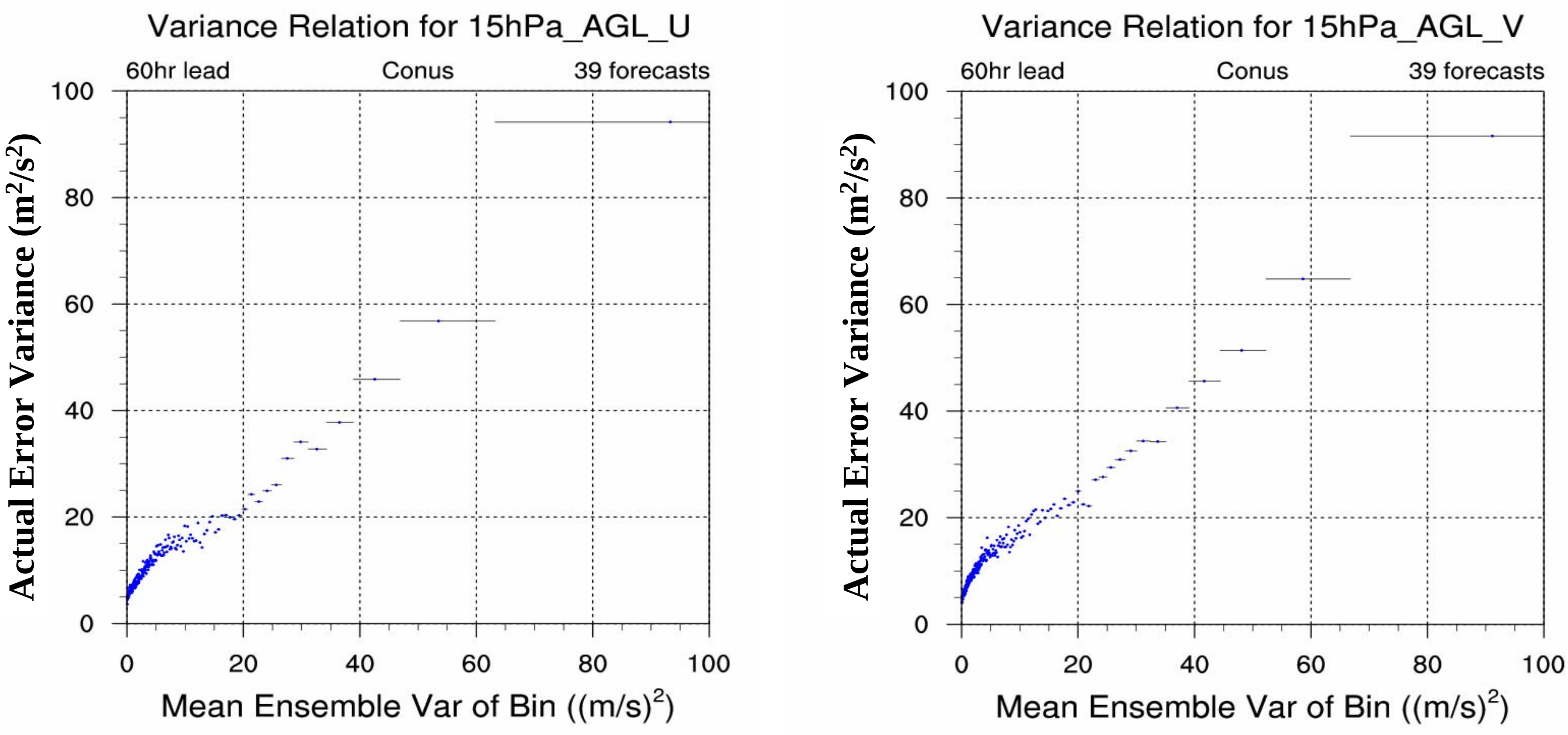
Linear Variance Calibration





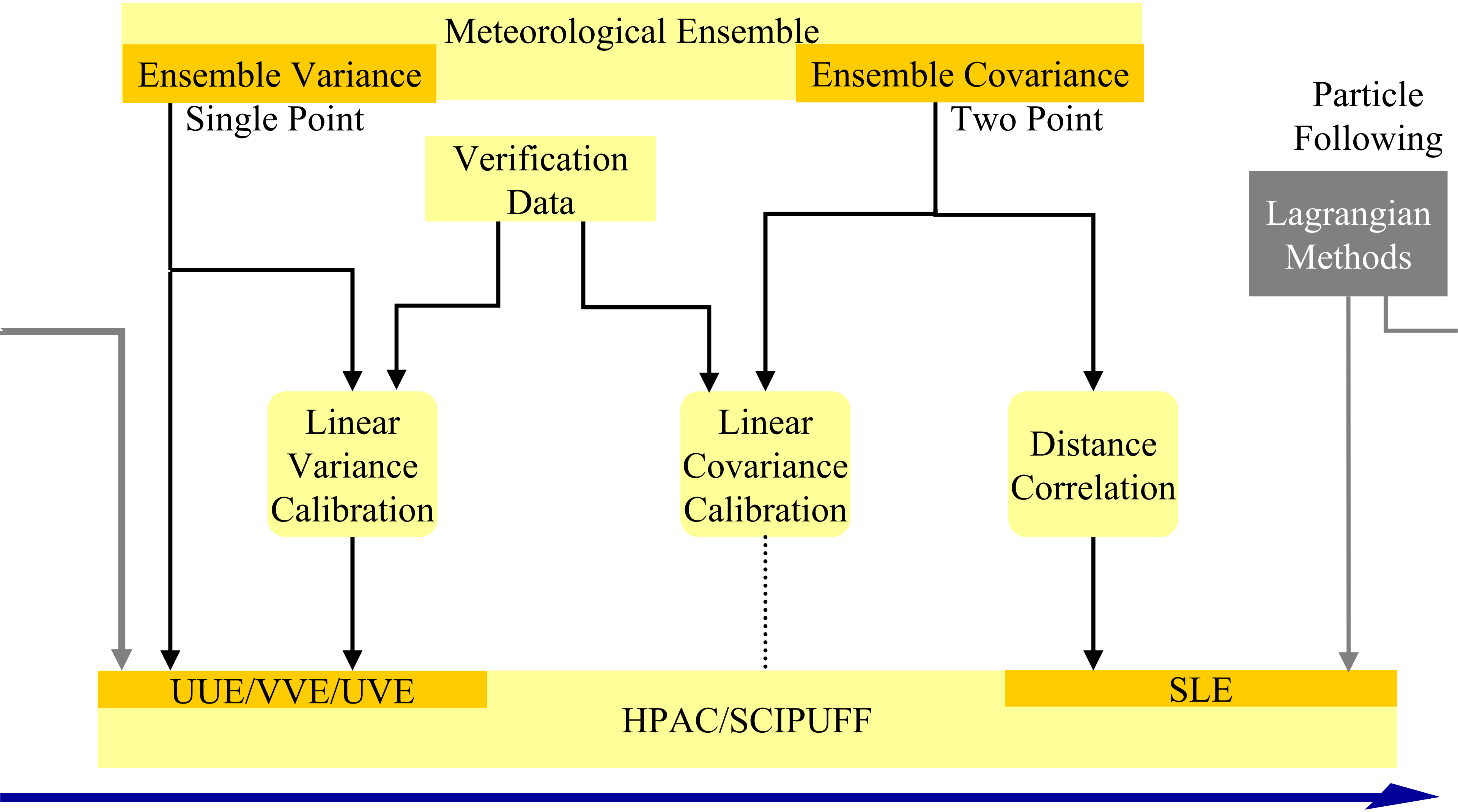
48 Hour Forecast

Linear Variance Calibration

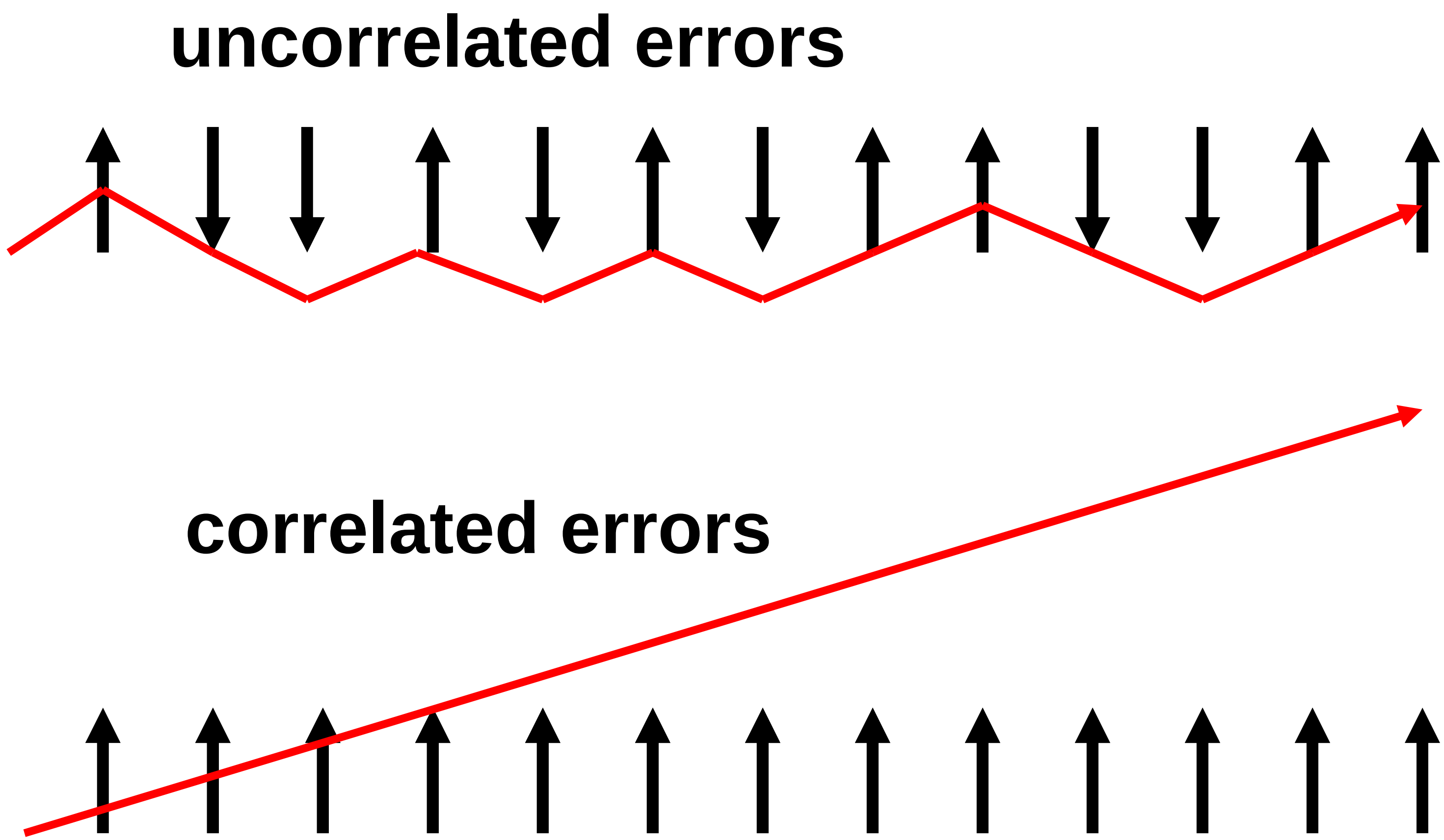


60 Hour Forecast

Linear Covariance Calibration



Linear Covariance Calibration



(Roulston 2005b)

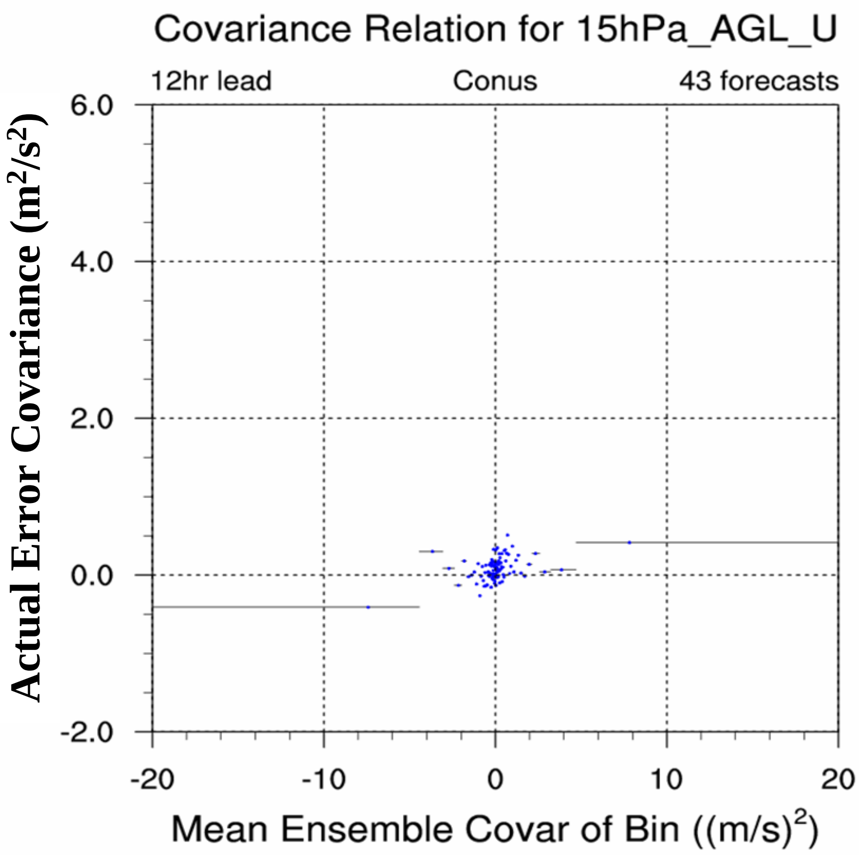
Linear Covariance Calibration

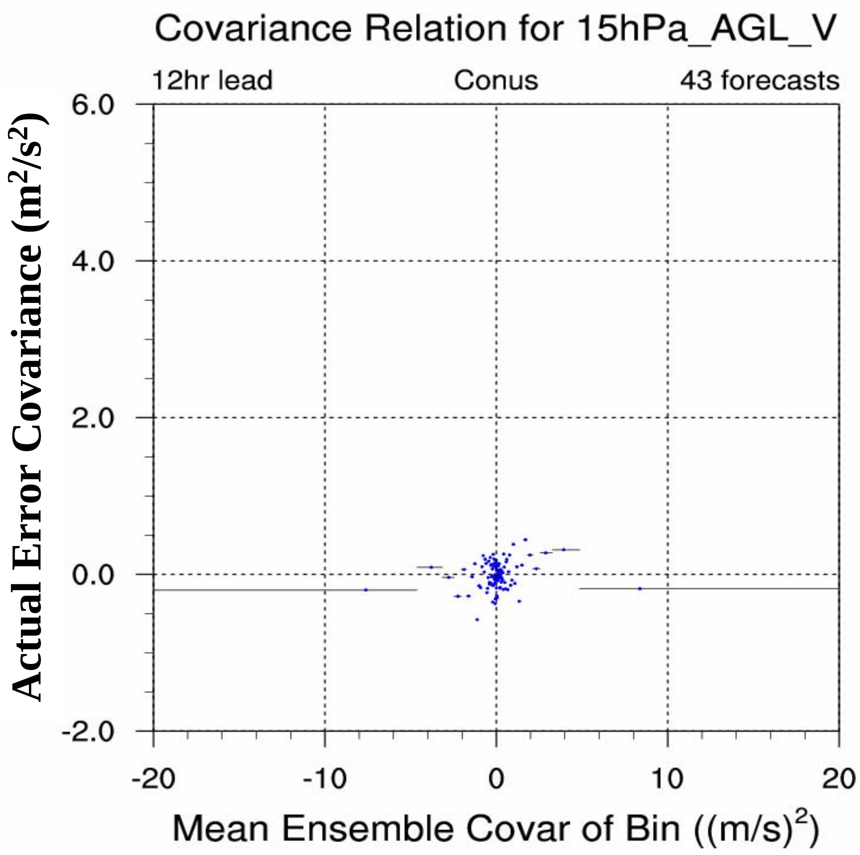
$$ECov\left(s(ij,kl)\right)=\frac{1}{N}\sum_{m=1}^{N}\left(s_m(ij)-\overline{s(ij)}\right)\left(s_m(kl)-\overline{s(kl)}\right)$$

$$ACov\left(s(ij,kl)\right)=\left(s_v(ij)-\overline{s(ij)}\right)\left(s_v(kl)-\overline{s(kl)}\right)$$

$s(ij)$	is the value of a scalar field at point (i,j)
$ECov\left(s(ij,kl)\right)$	is the Ensemble Covariance of s between (i,j) and (k.l)
$AVar\left(s(kl)\right)$	is the Actual Covariance of s between (i,j) and (k,l)
N	is the number of ensemble members
s_m	is the scalar value of a single ensemble member
s_v	is the scalar value of the verification

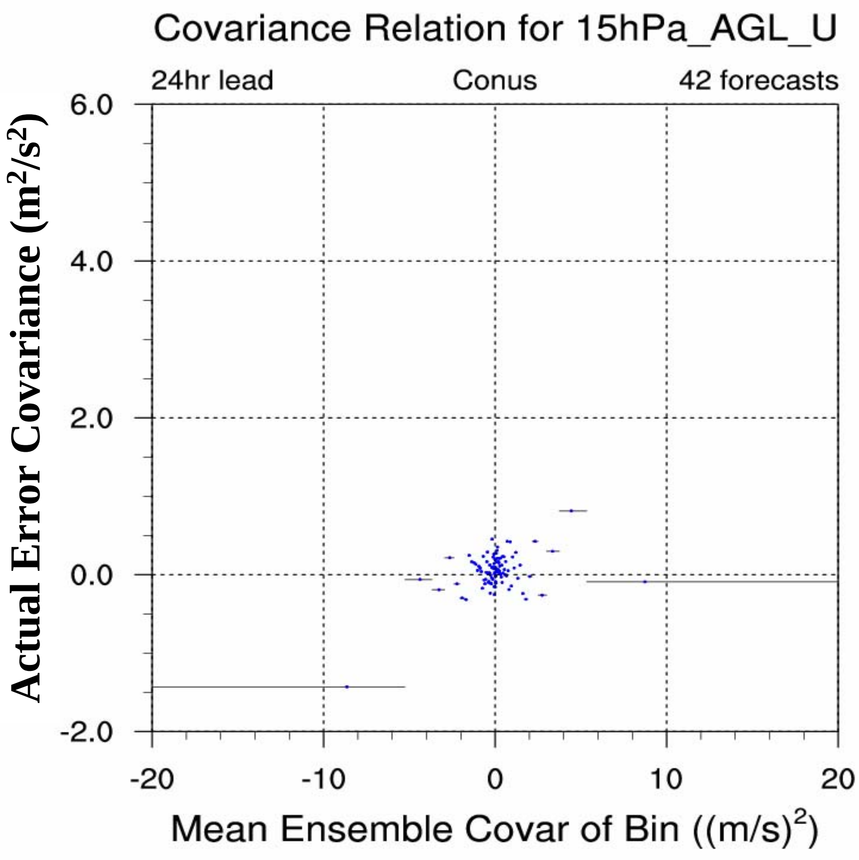
Linear Covariance Calibration

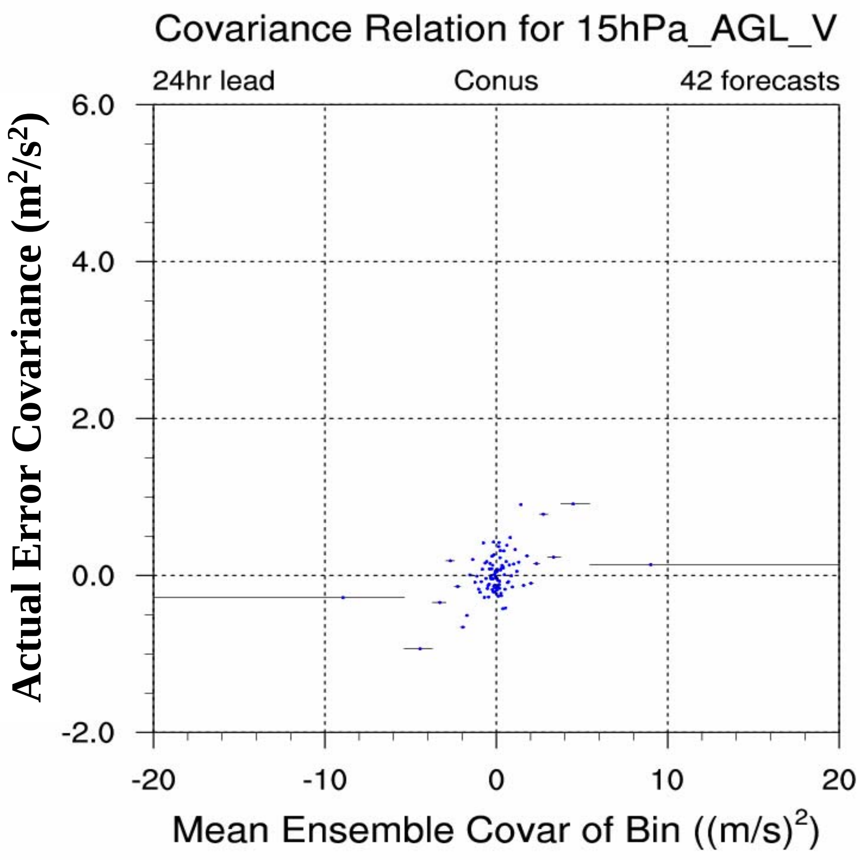




12 Hour Forecast

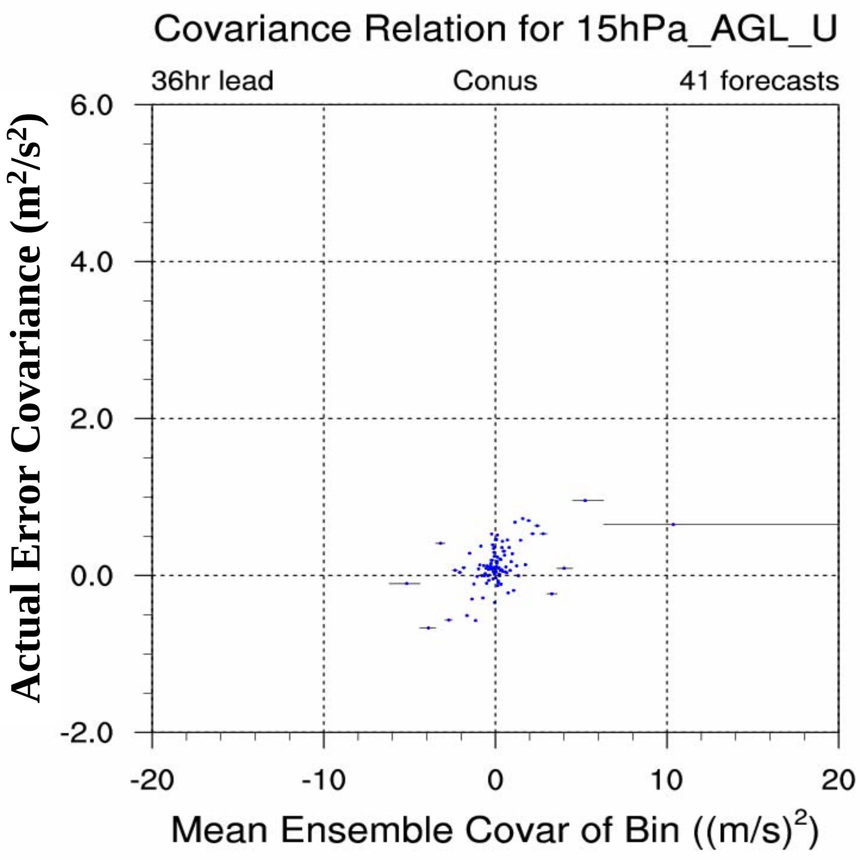
Linear Covariance Calibration

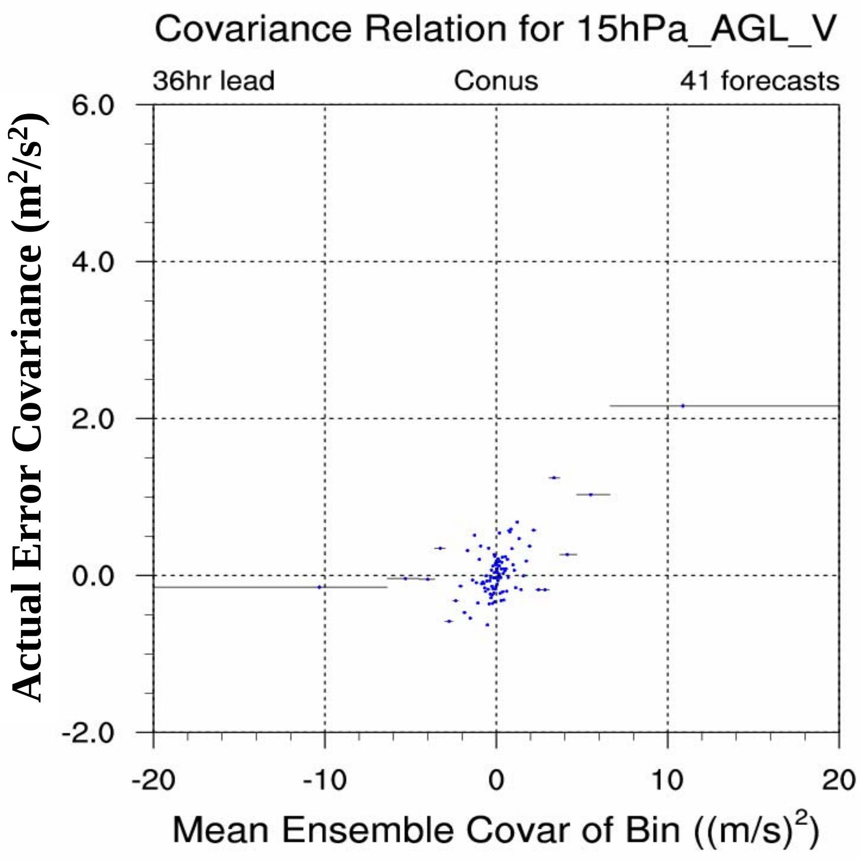




24 Hour Forecast

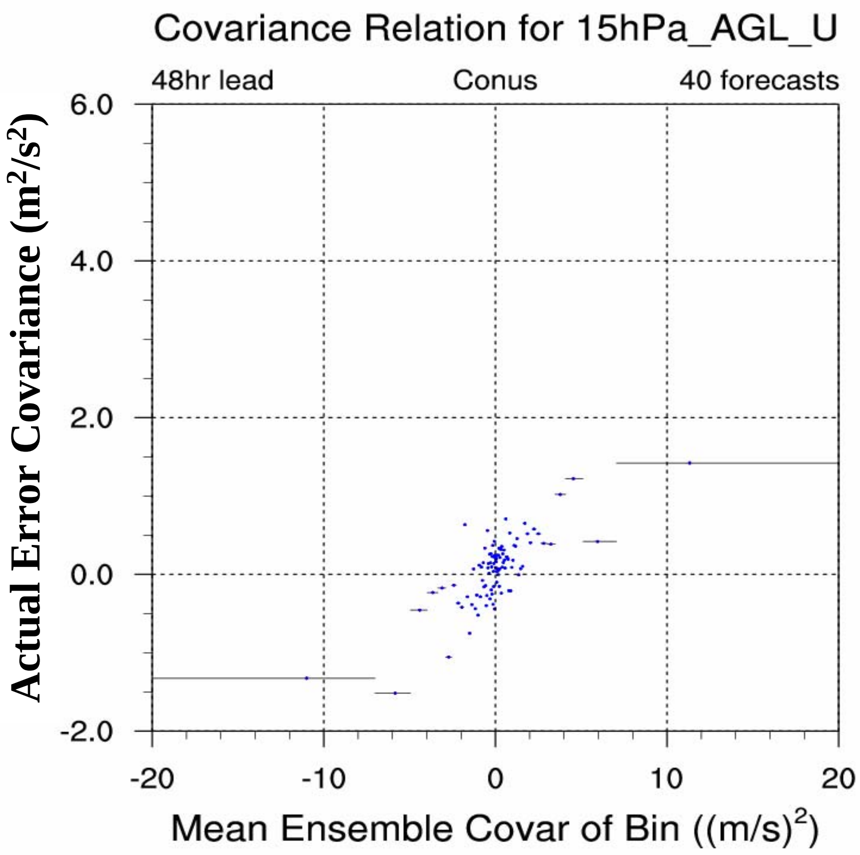
Linear Covariance Calibration

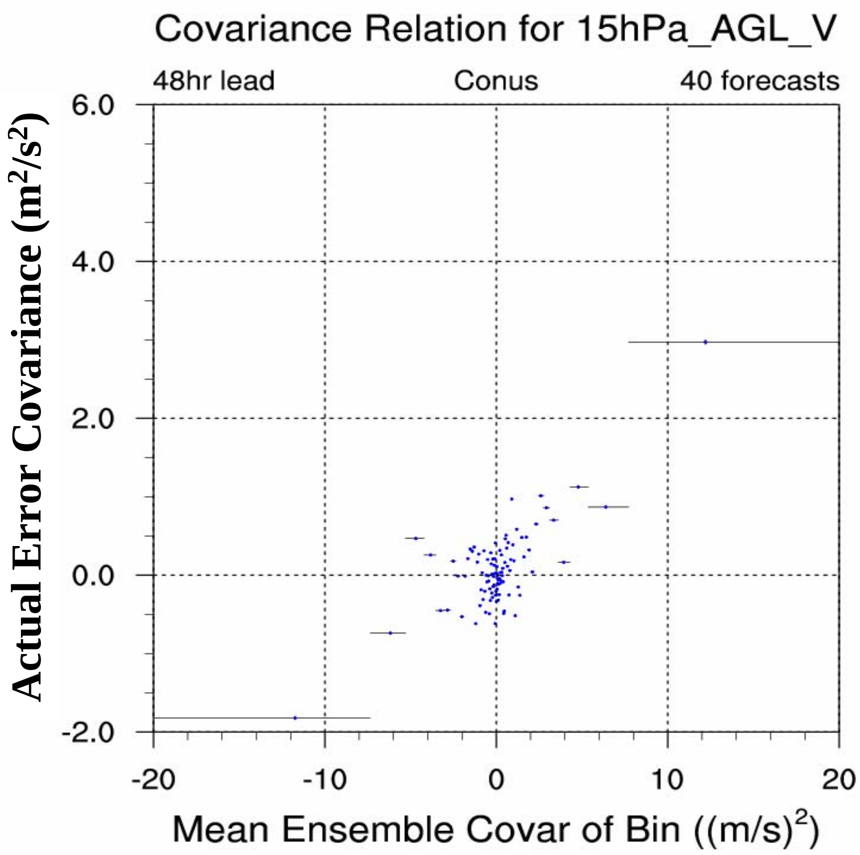




36 Hour Forecast

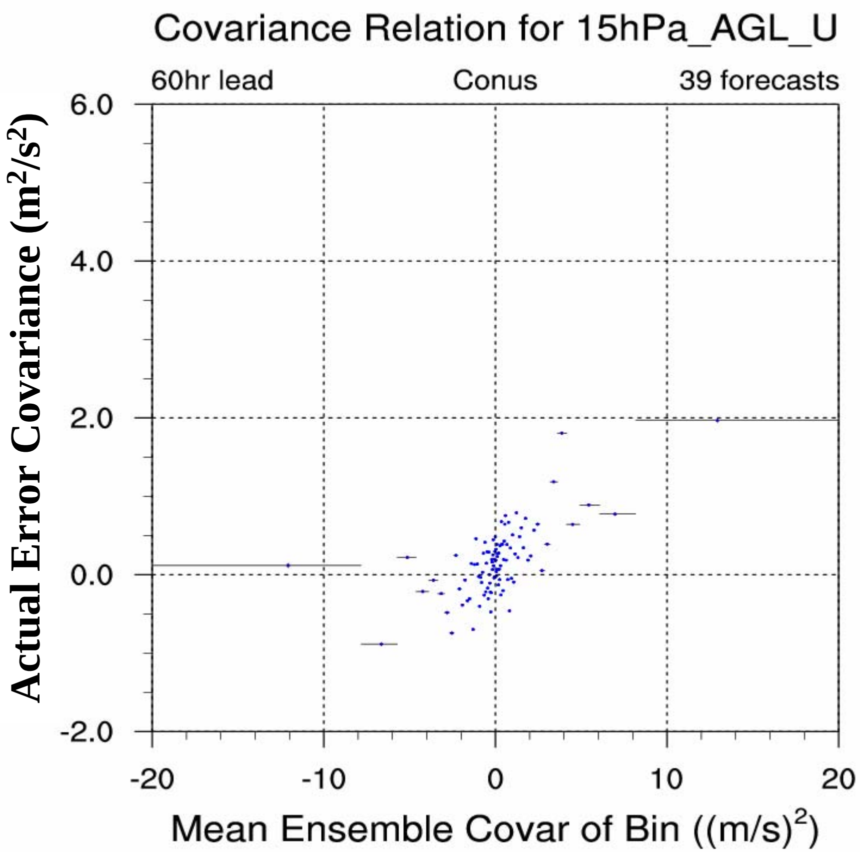
Linear Covariance Calibration

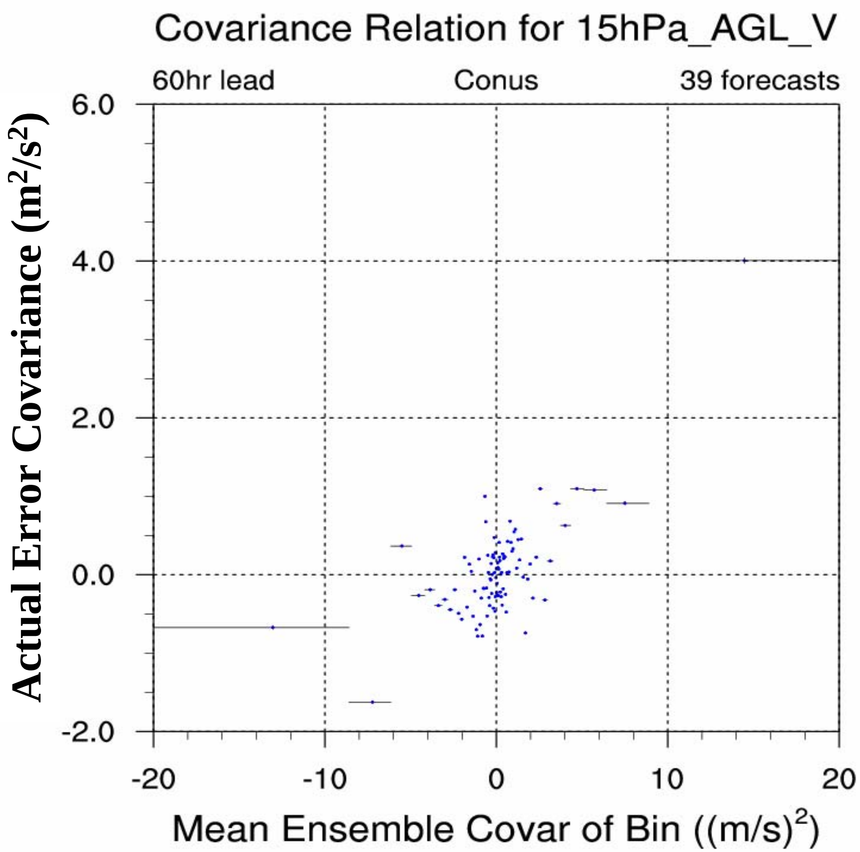




48 Hour Forecast

Linear Covariance Calibration





60 Hour Forecast

Linear Covariance Calibration

$$ECov\left(s(ij,kl)\right) = \frac{1}{N}\sum_{m=1}^{N}\left(s_m(ij) - \overline{s(ij)}\right)\left(s_m(kl) - \overline{s(kl)}\right)$$

$$ECov\left(s(ij,kl)\right) = \frac{1}{N}\sum_{m=1}^{N}\left(s_m(ij) - s_c(ij)\right)\left(s_m(kl) - s_c(kl)\right)$$

$$ACov\left(s(ij,kl)\right) = \left(s_v(ij) - \overline{s(ij)}\right)\left(s_v(kl) - \overline{s(kl)}\right)$$

$$ACov\left(s(ij,kl)\right) = \left(s_v(ij) - s_c(ij)\right)\left(s_v(kl) - s_c(kl)\right)$$

$s(ij)$ is the value of a scalar field at point (i,j)

$ECov\left(s(ij,kl)\right)$ is the Ensemble Covariance of s between (i,j) and (k.l)

$AVar\left(s(kl)\right)$ is the Actual Covariance of s between (i,j) and (k,l)

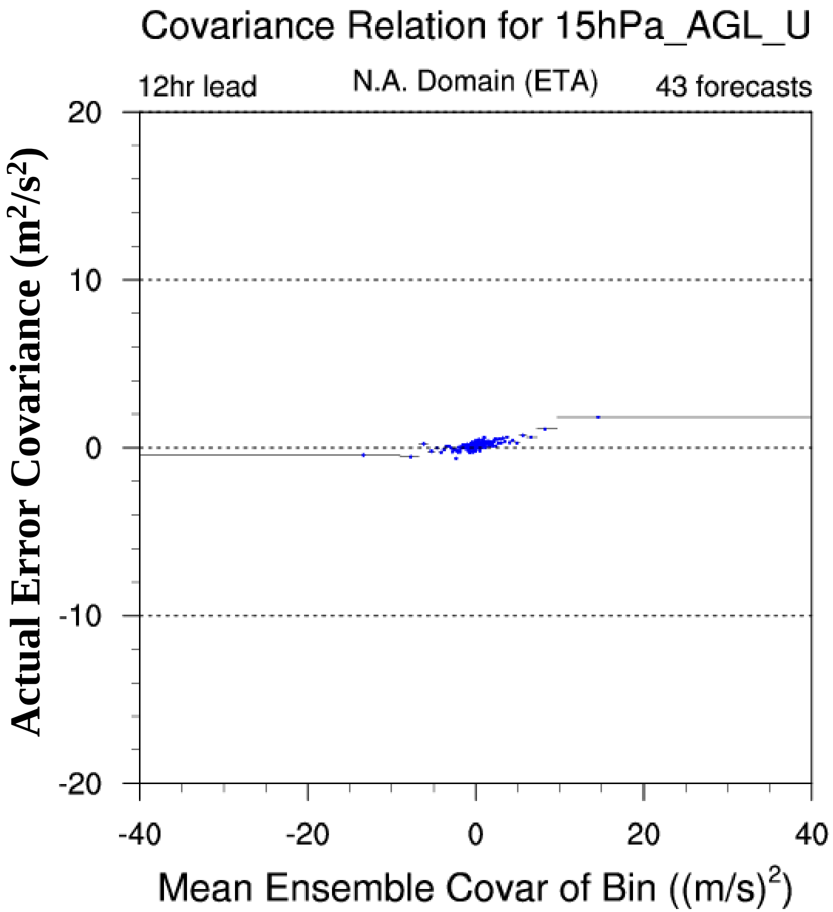
N is the number of ensemble members

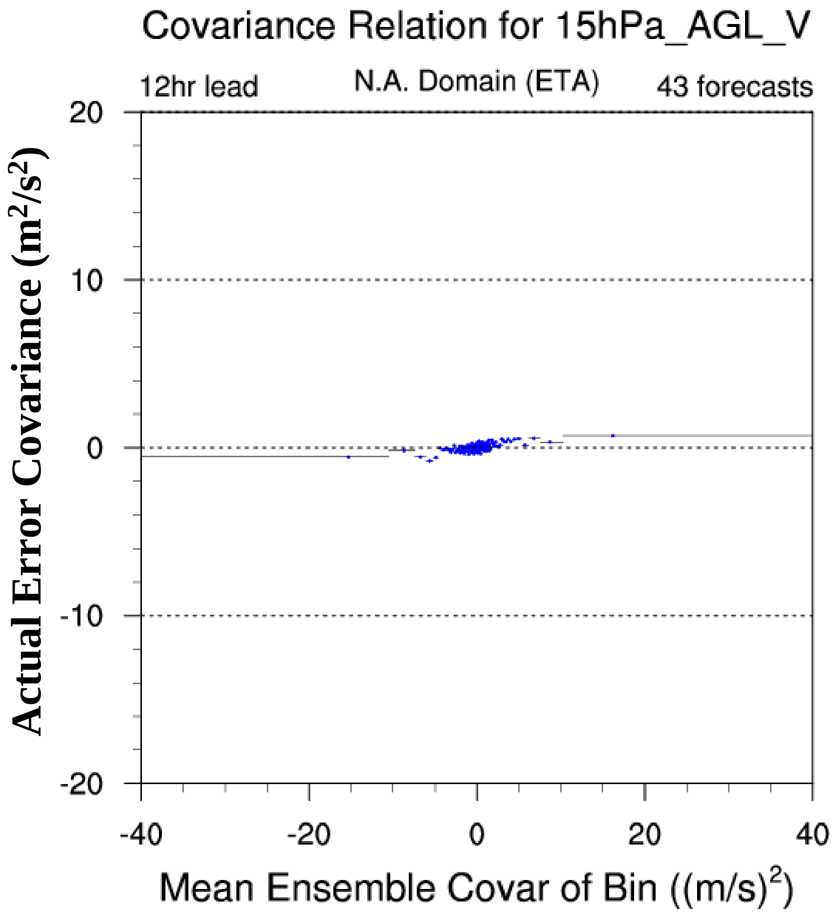
S_c is the scalar value of the control member

S_m is the scalar value of a single ensemble member

S_v is the scalar value of the verification

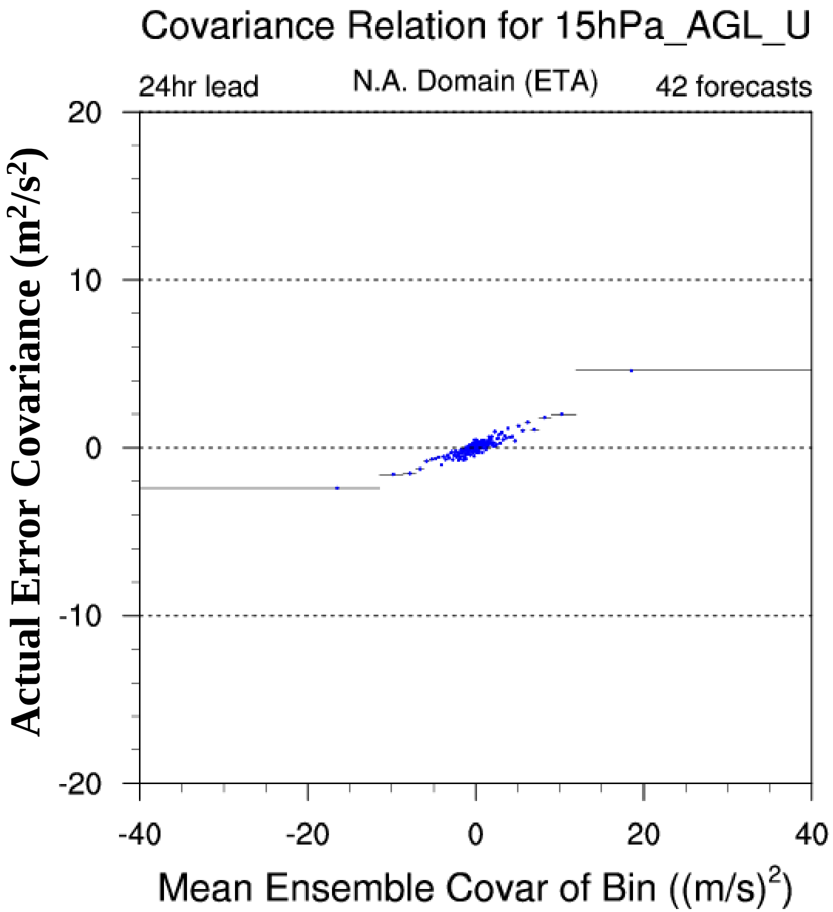
Linear Covariance Calibration - Control

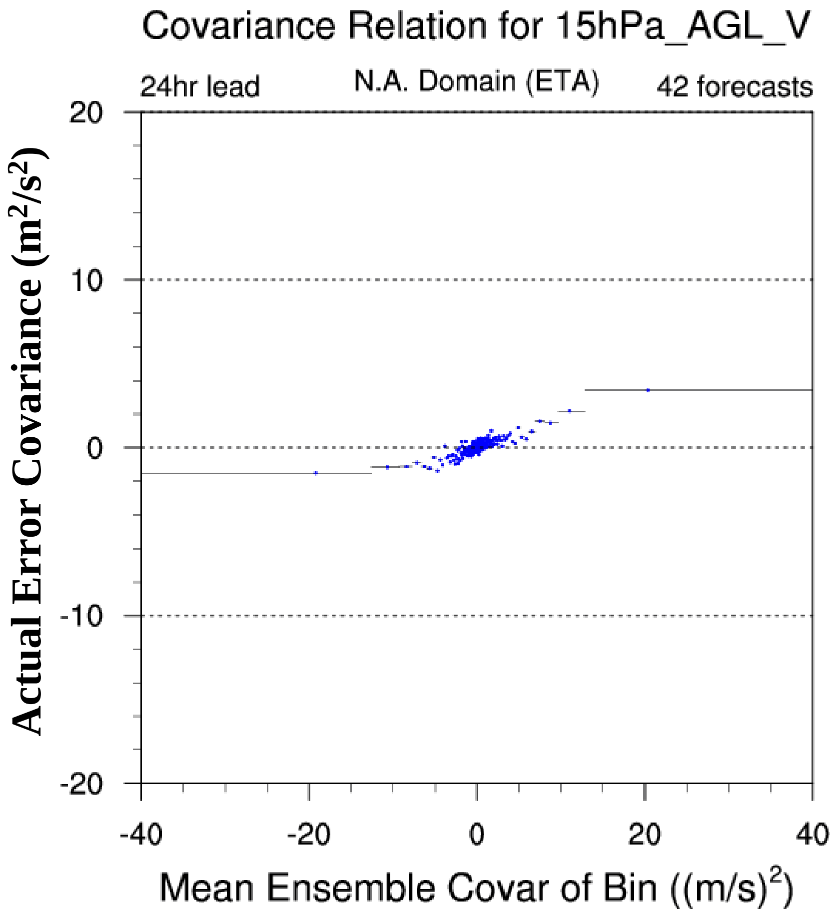




12 Hour Forecast

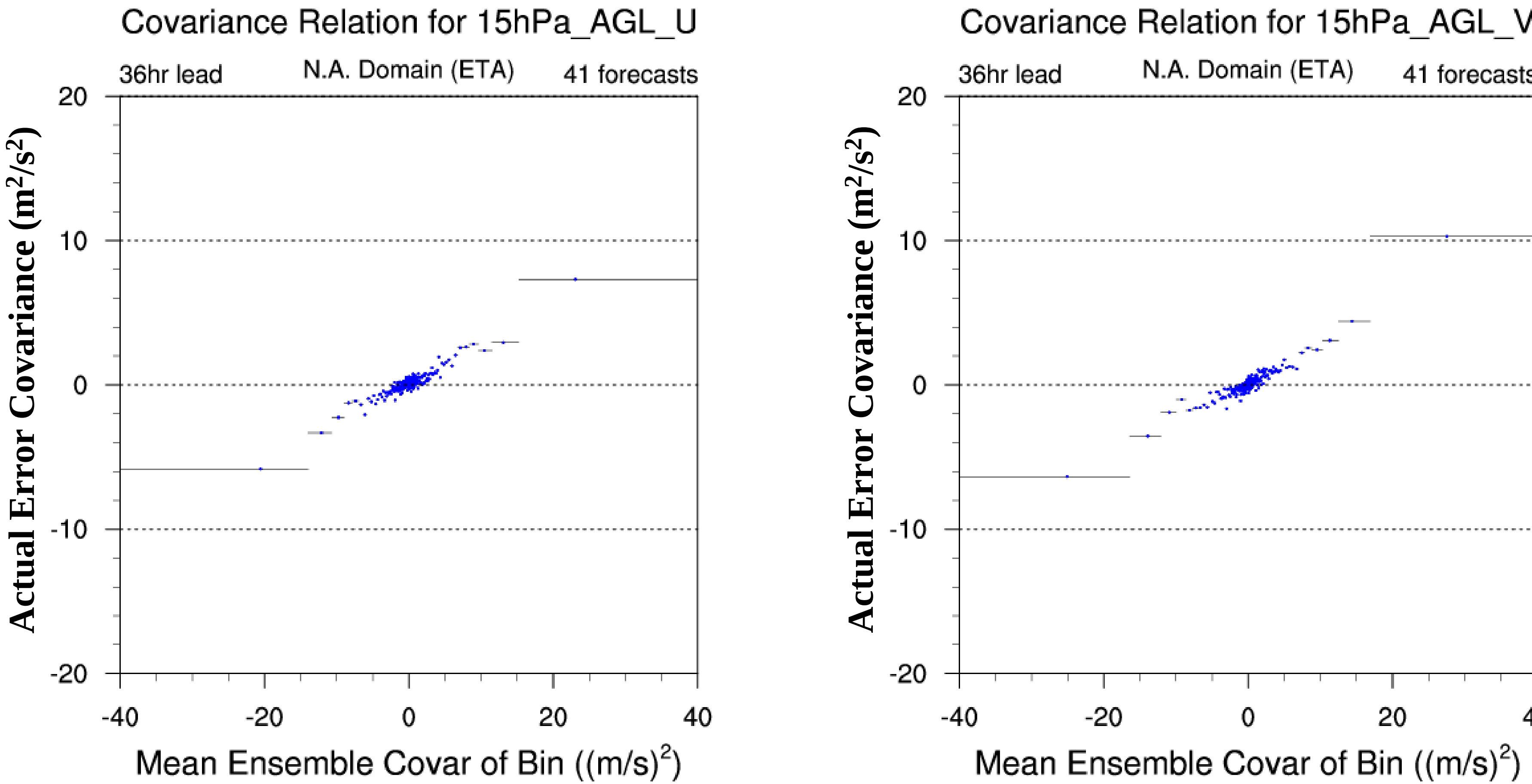
Linear Covariance Calibration - Control





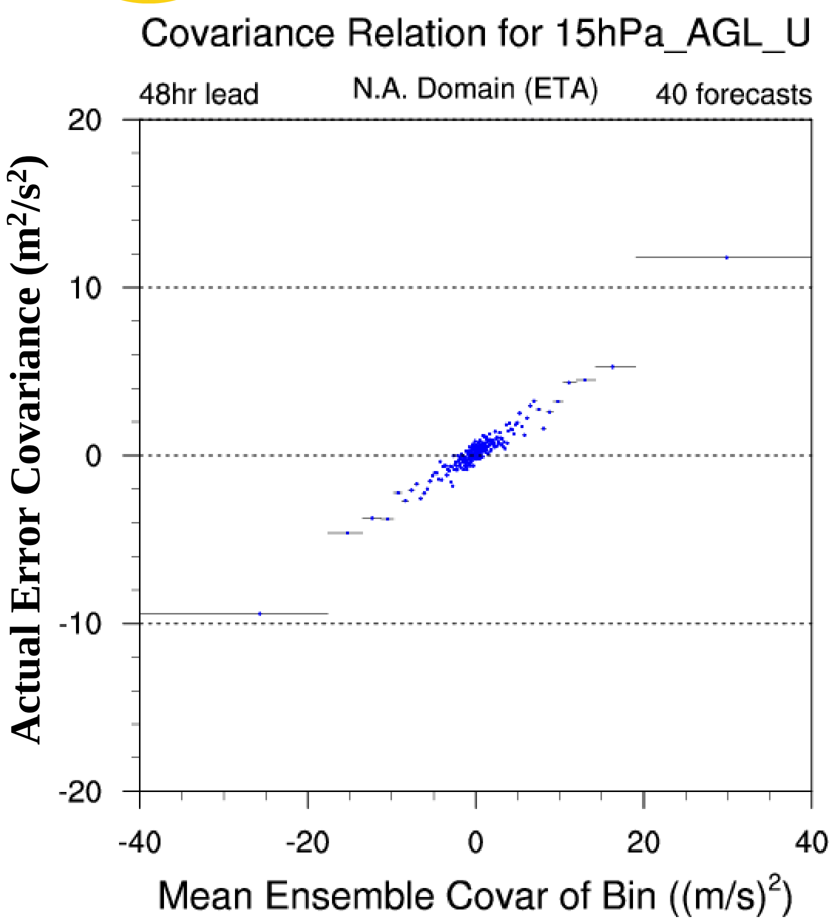
24 Hour Forecast

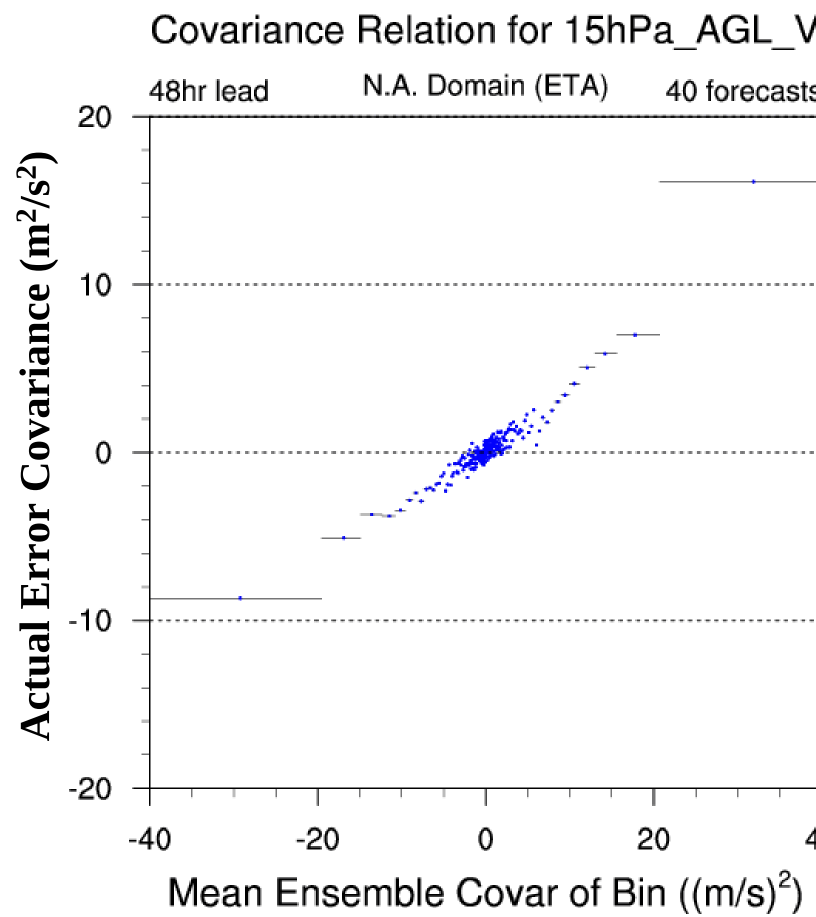
Linear Covariance Calibration - Control



36 Hour Forecast

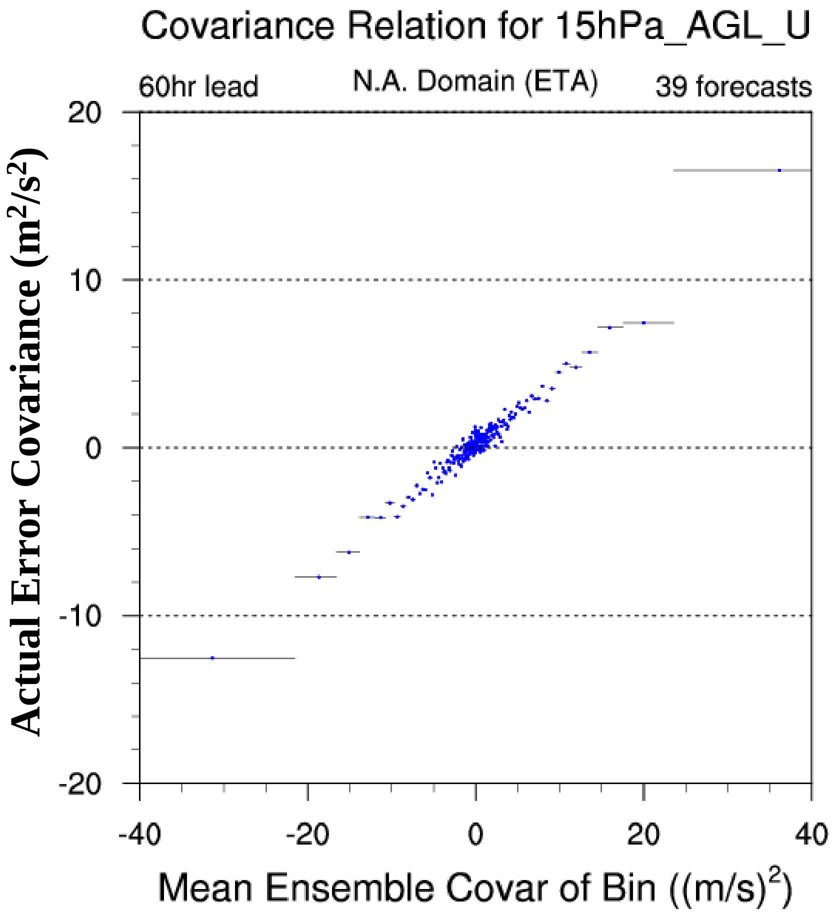
Linear Covariance Calibration - Control

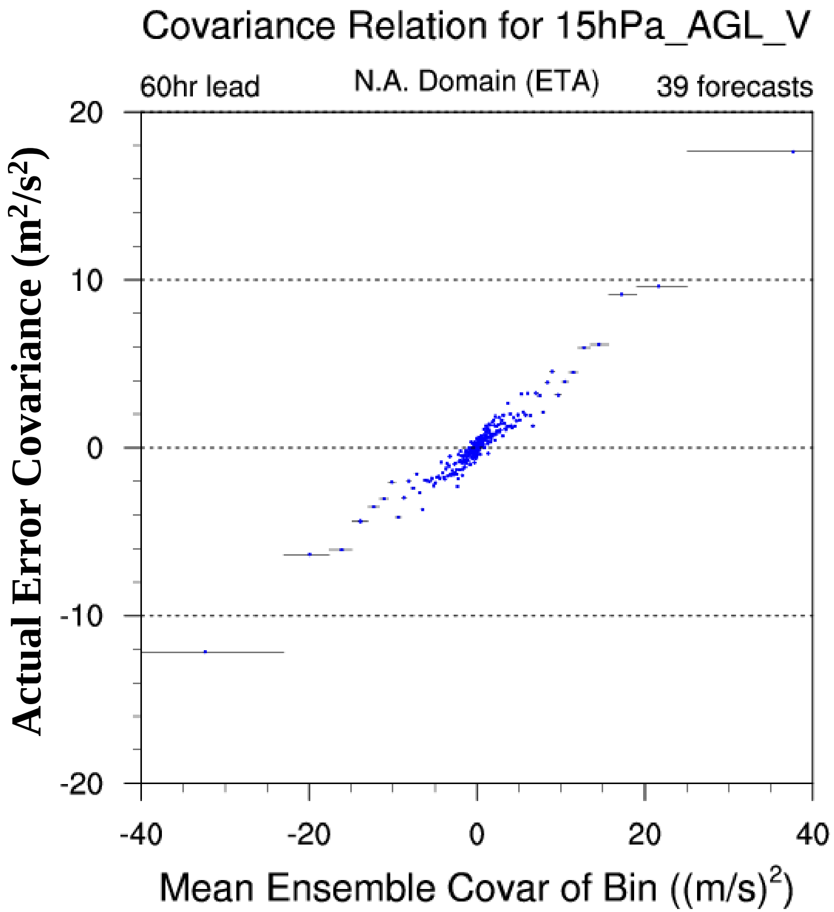




48 Hour Forecast

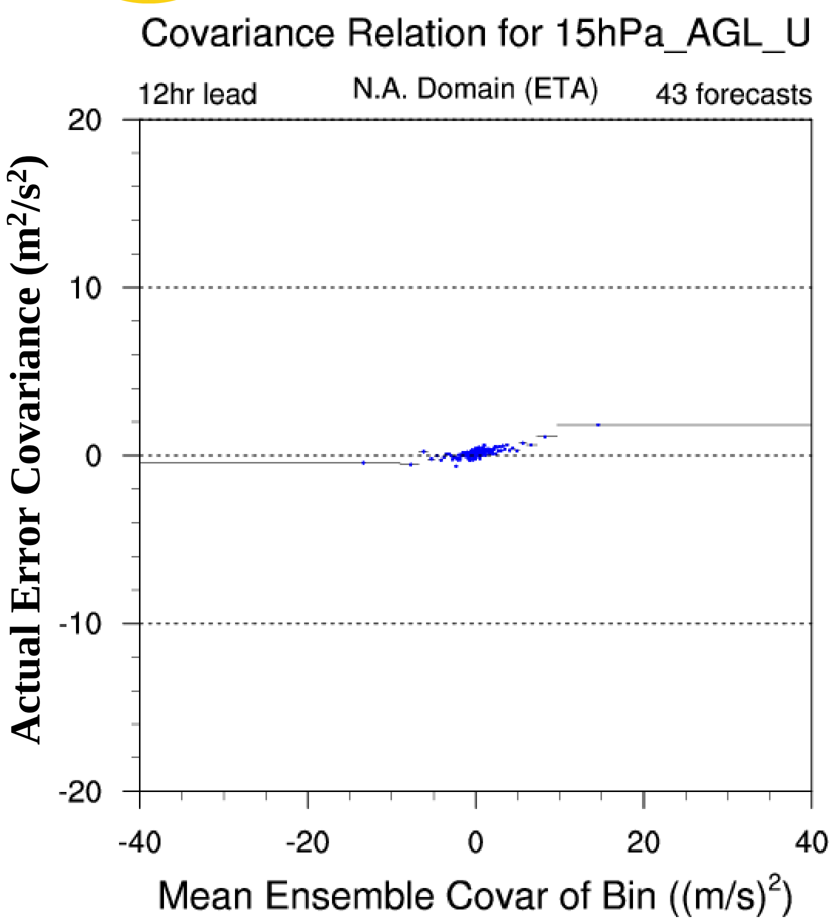
Linear Covariance Calibration - Control

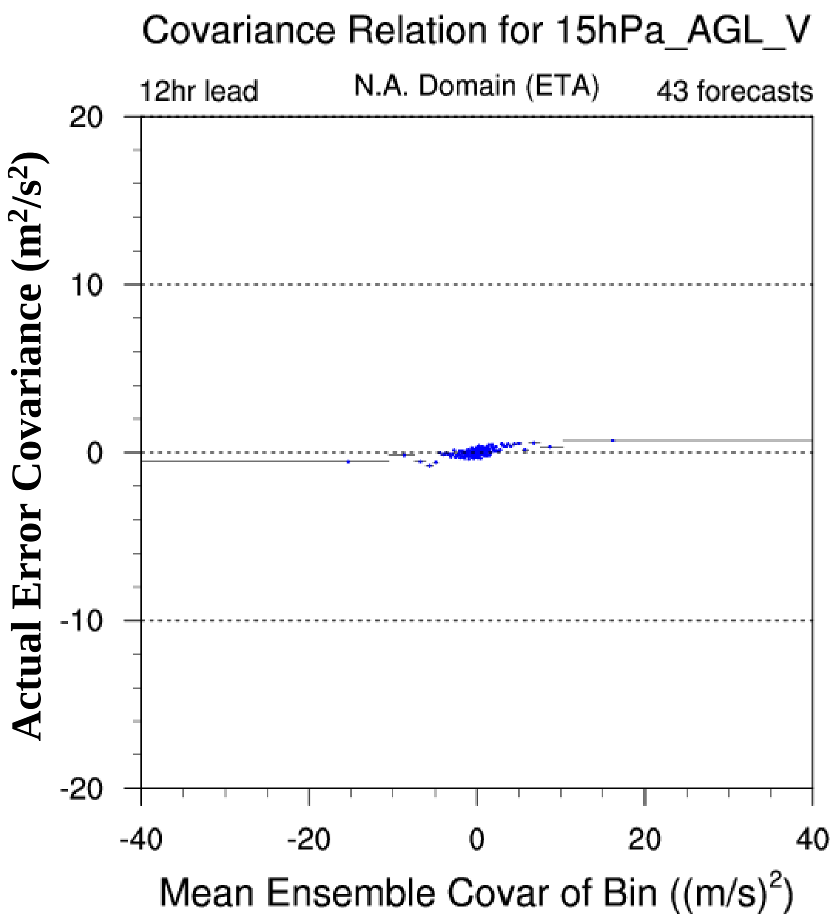




60 Hour Forecast

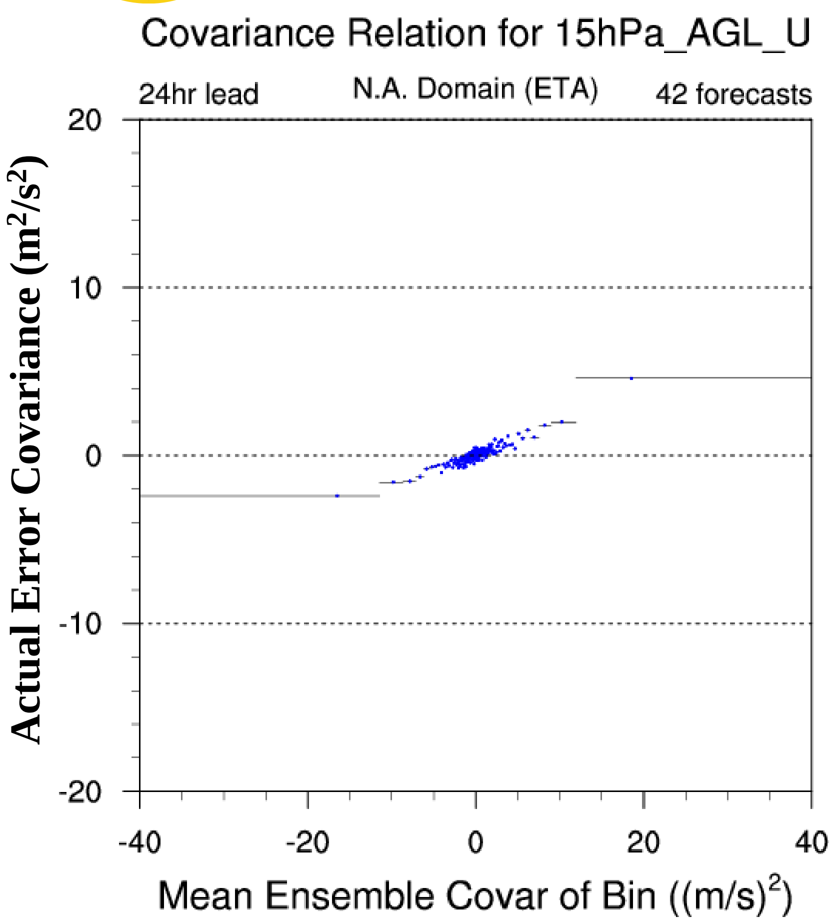
Linear Covariance Calibration - Control

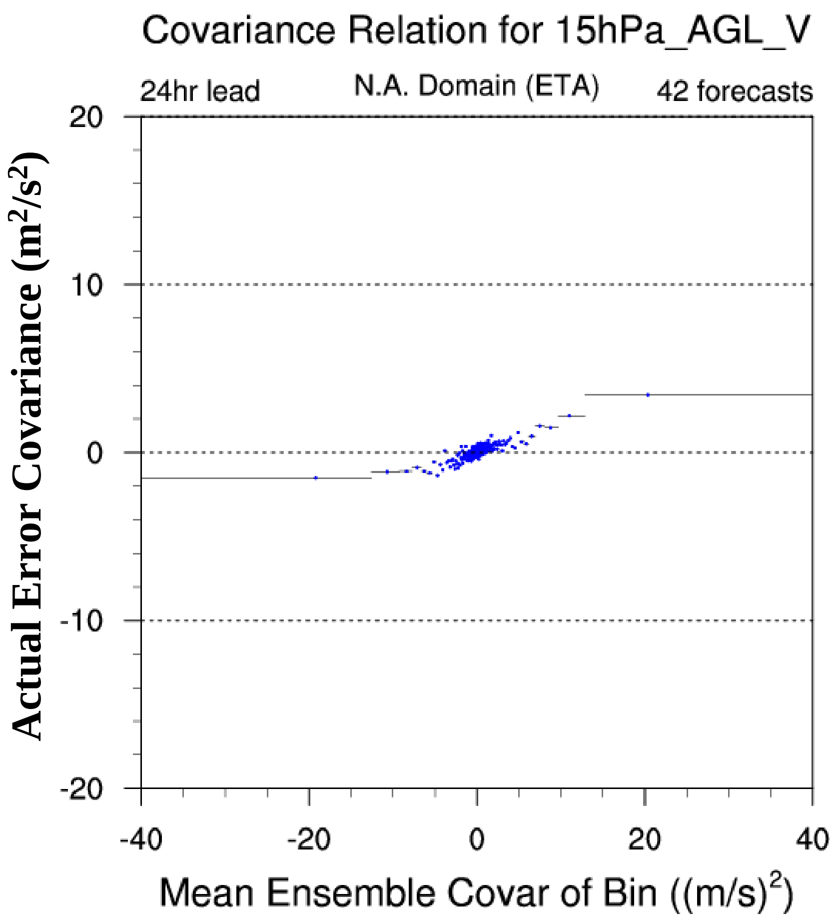




12 Hour Forecast

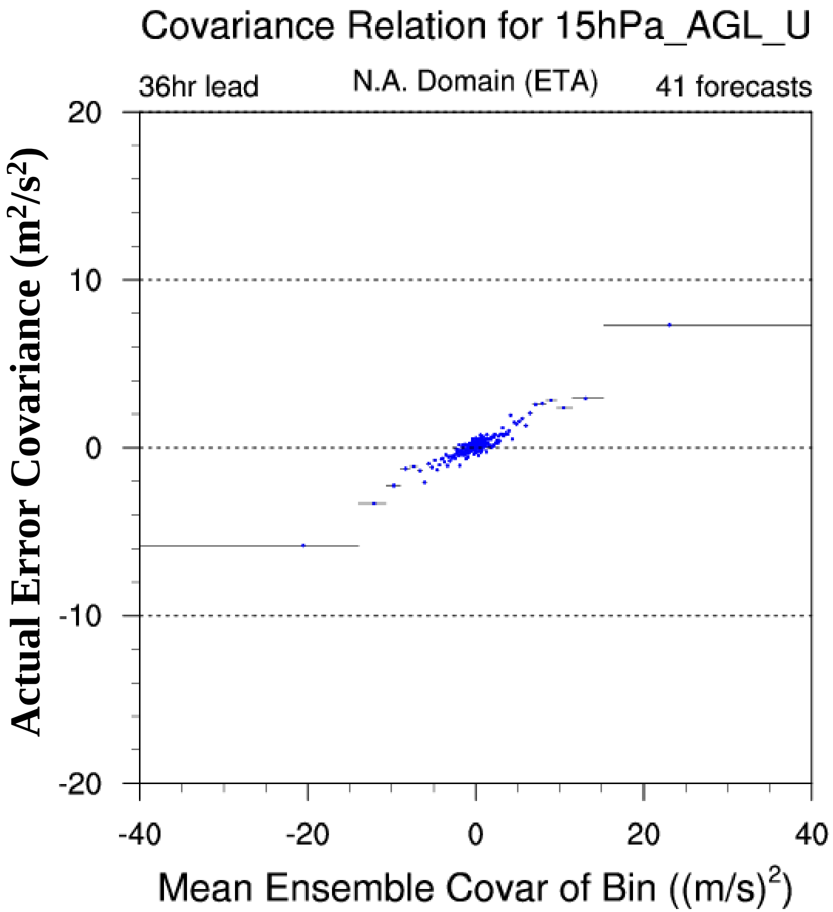
Linear Covariance Calibration - Control

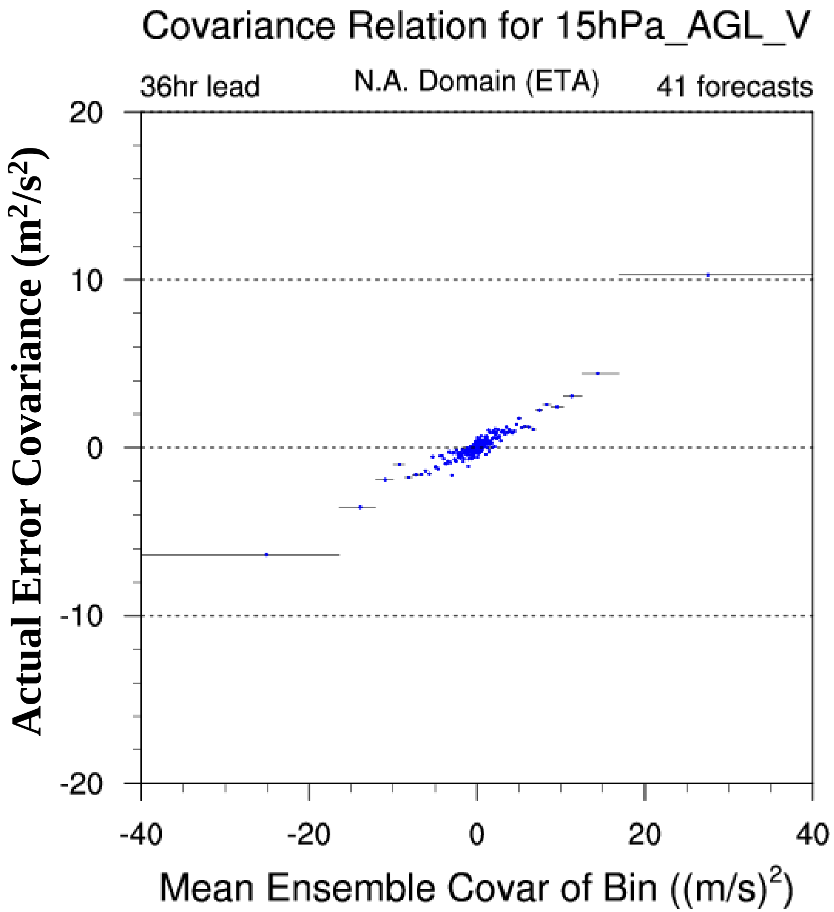




24 Hour Forecast

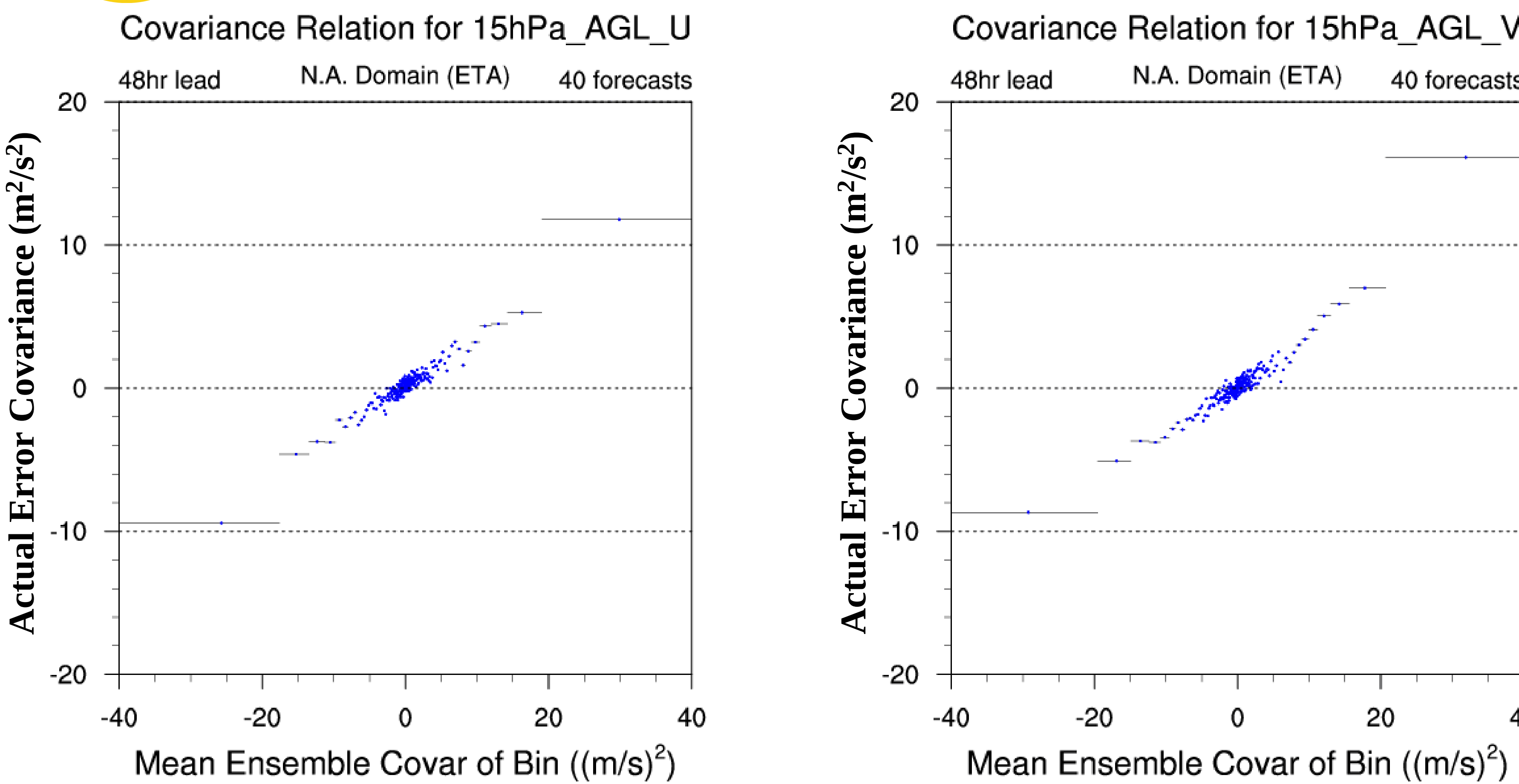
Linear Covariance Calibration - Control





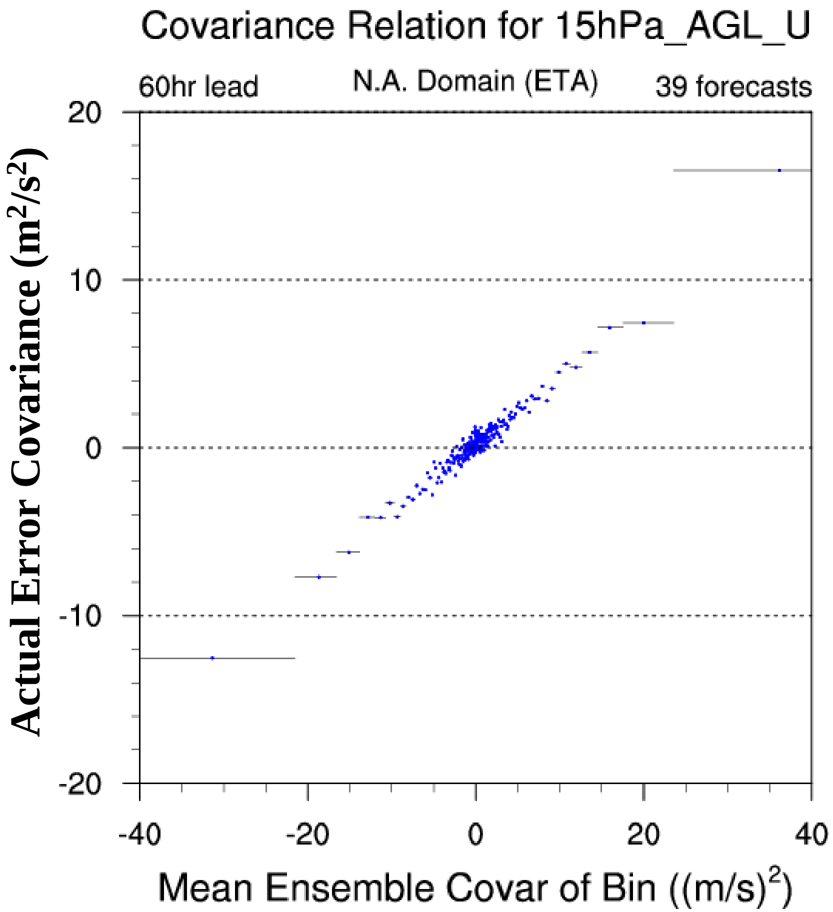
36 Hour Forecast

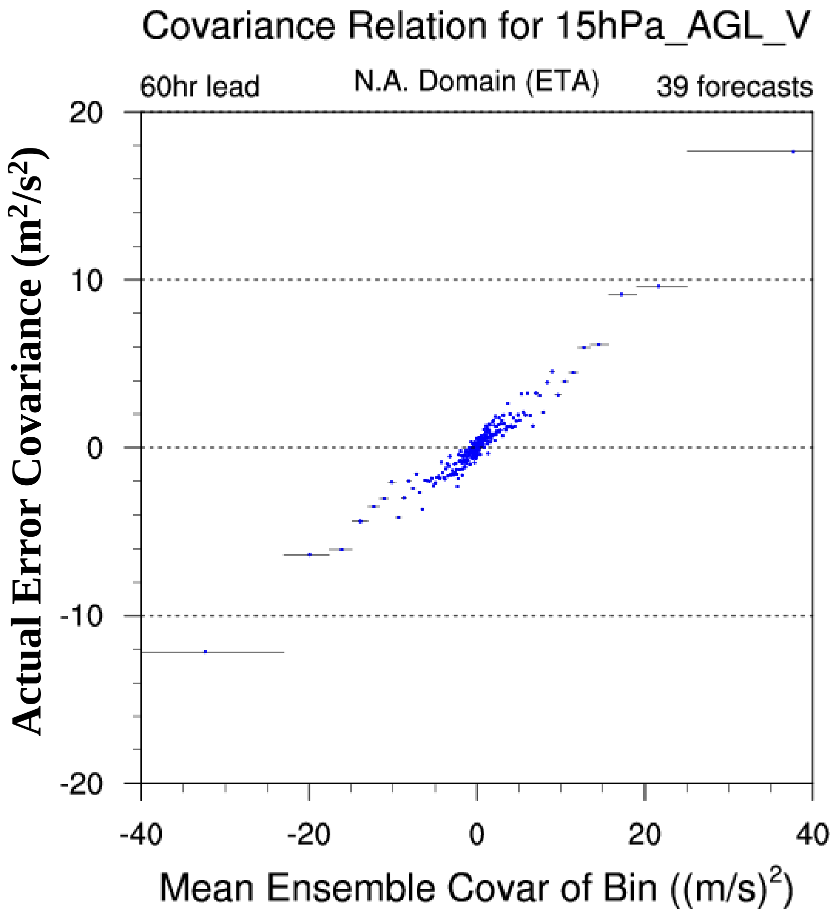
Linear Covariance Calibration - Control



48 Hour Forecast

Linear Covariance Calibration - Control



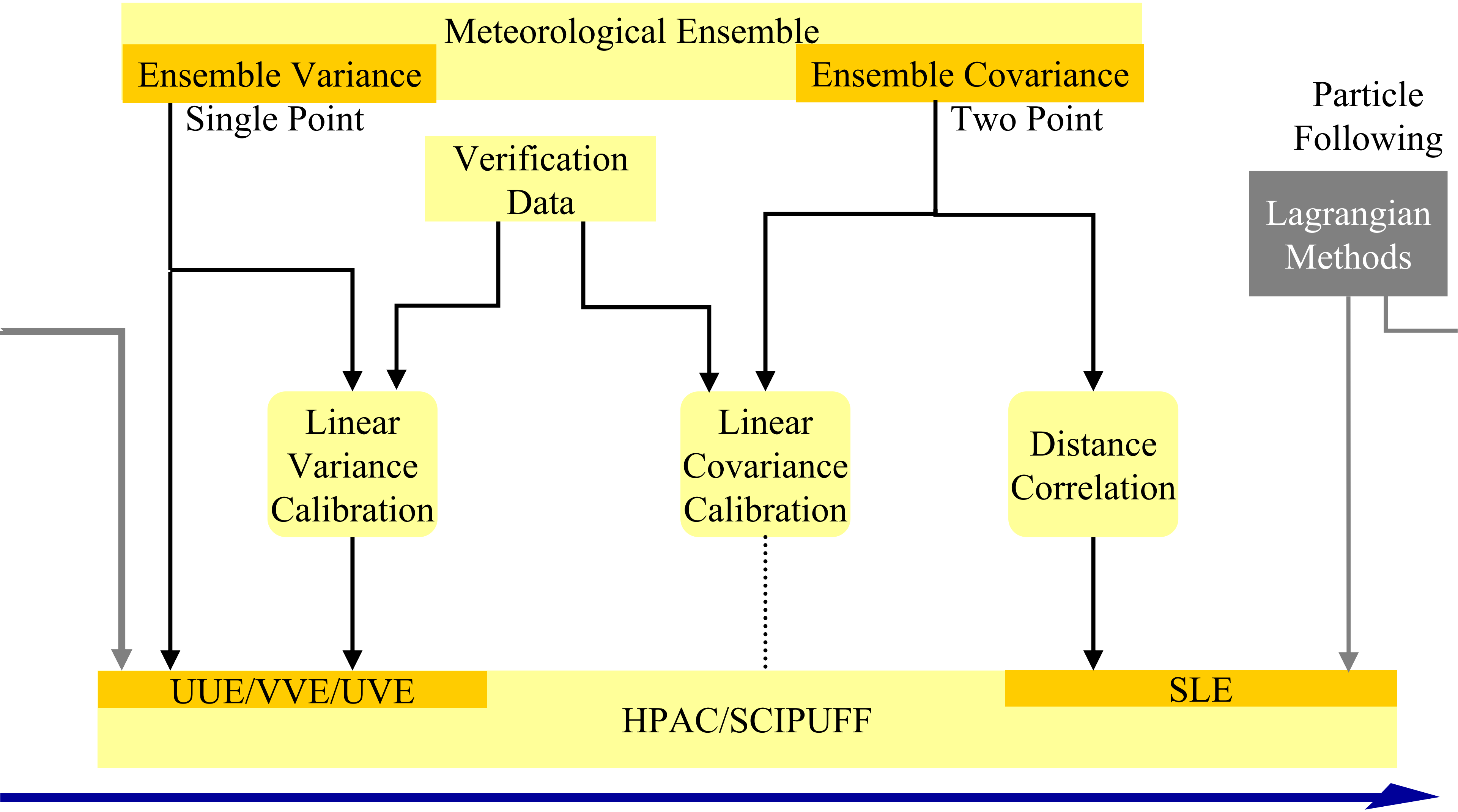


60 Hour Forecast

Linear Covariance Calibration

- Linear Covariance Calibration may work, but there is currently no direct route for ingesting covariance information in SCIPUFF
- However, spatial variability is related to the Lagrangian time scale, so perhaps we could find a way to use covariance information that CAN be used for SCIPUFF…

Relating Correlation and Distance

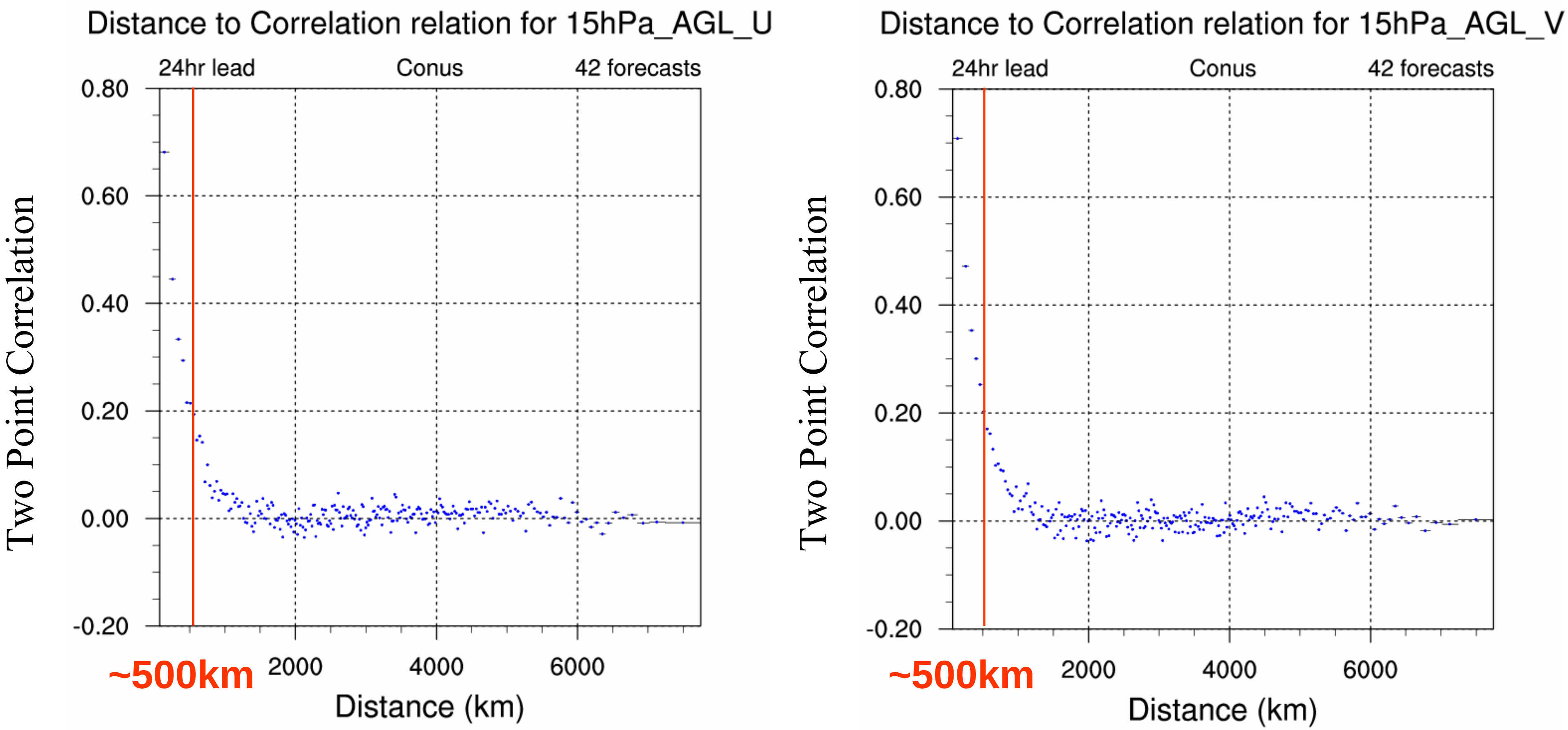


Relating Correlation and Distance

$$Corr(x,y) = \frac{Cov(x,y)}{\sigma_x \sigma_y}$$

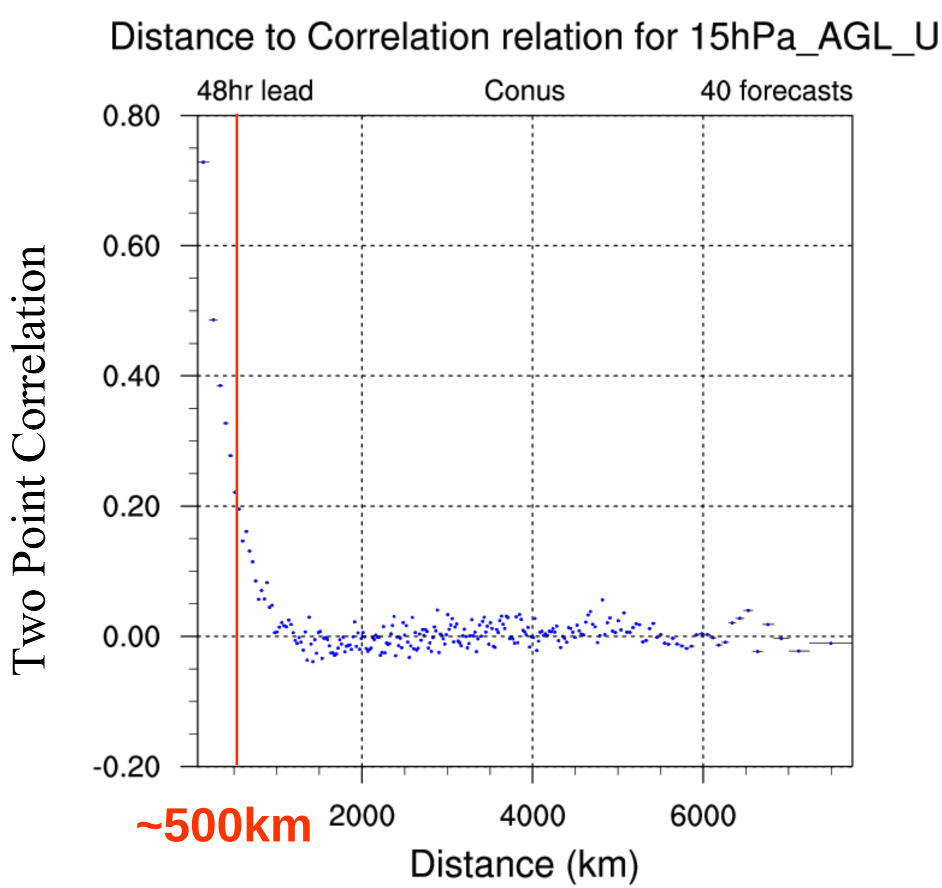
$$ECor\left(s(ij),s(kl)\right) = \frac{ECov(s(ij),s(kl))}{\sqrt{EVar\left(s(ij)\right)}\sqrt{EVar\left(s(kl)\right)}}$$

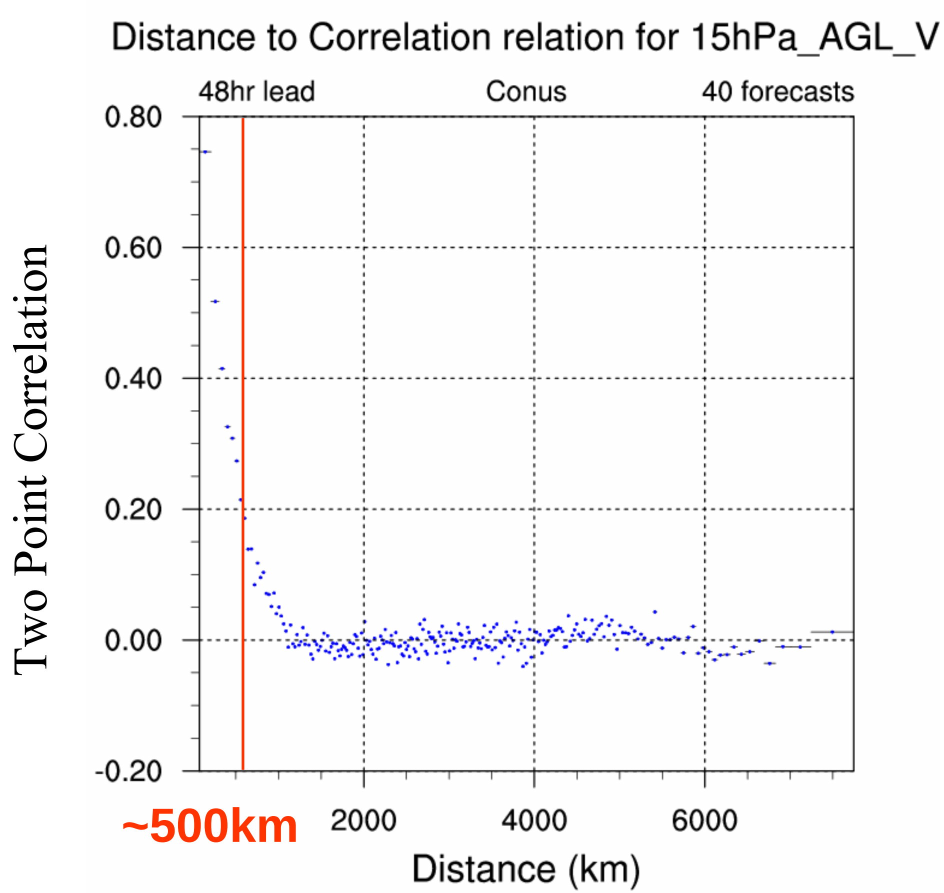
Relating Correlation and Distance



24 Hour Forecast

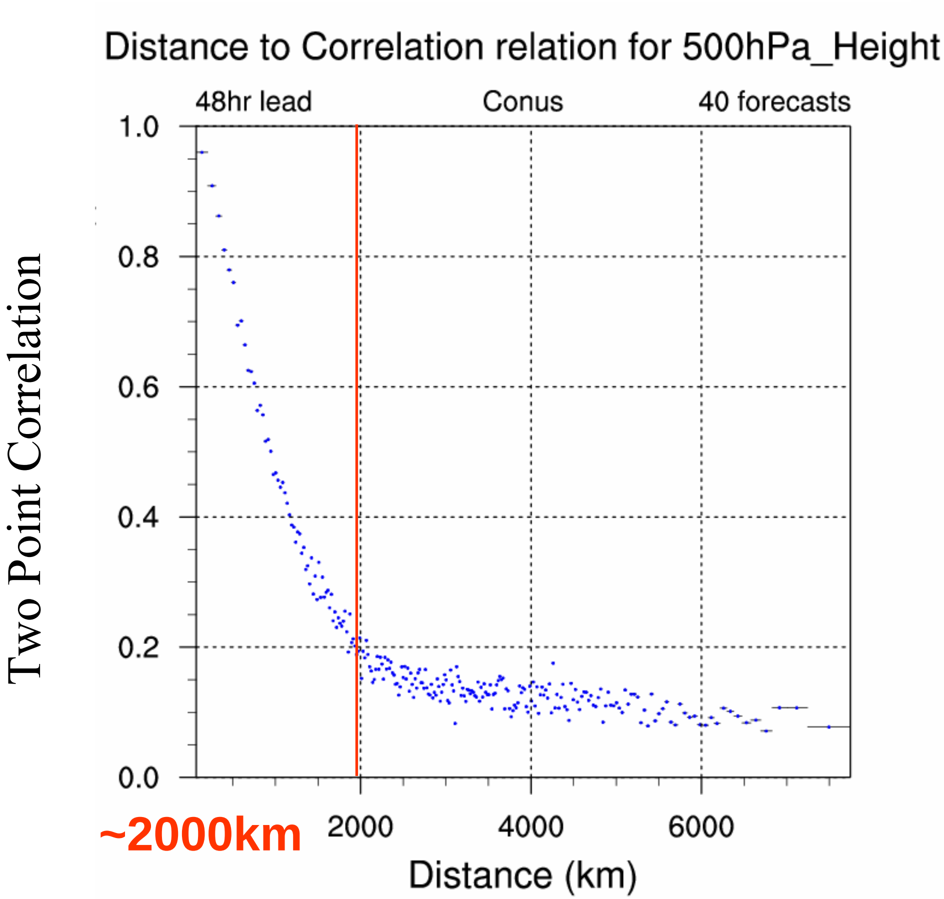
Relating Correlation and Distance

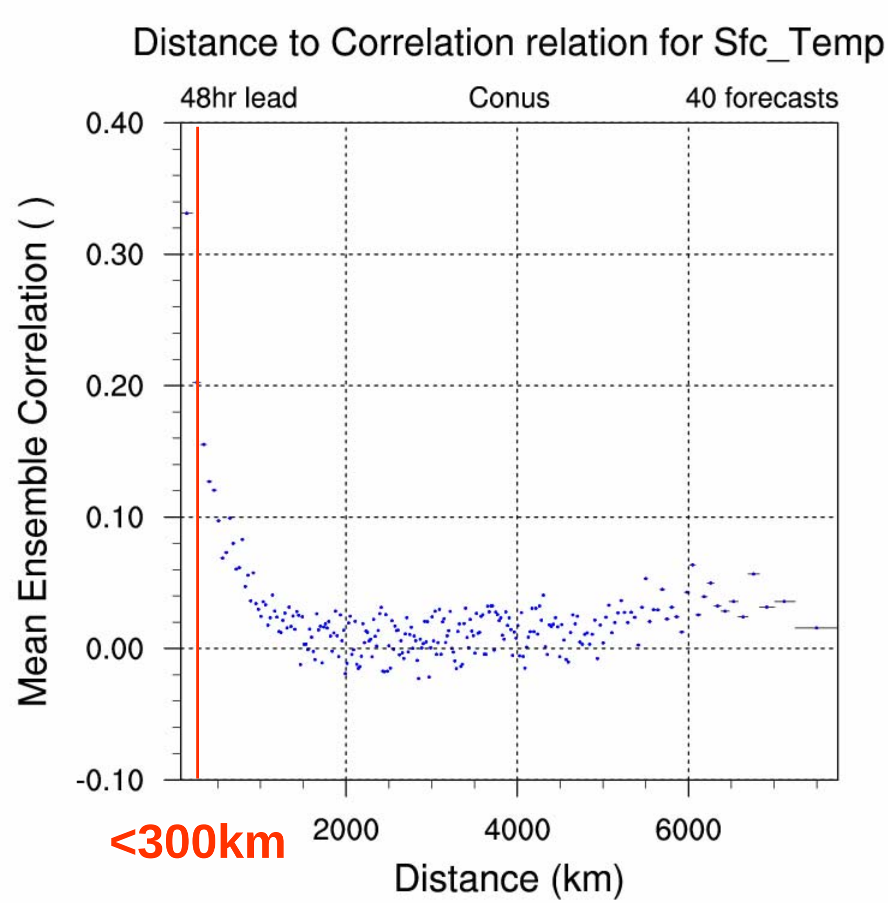




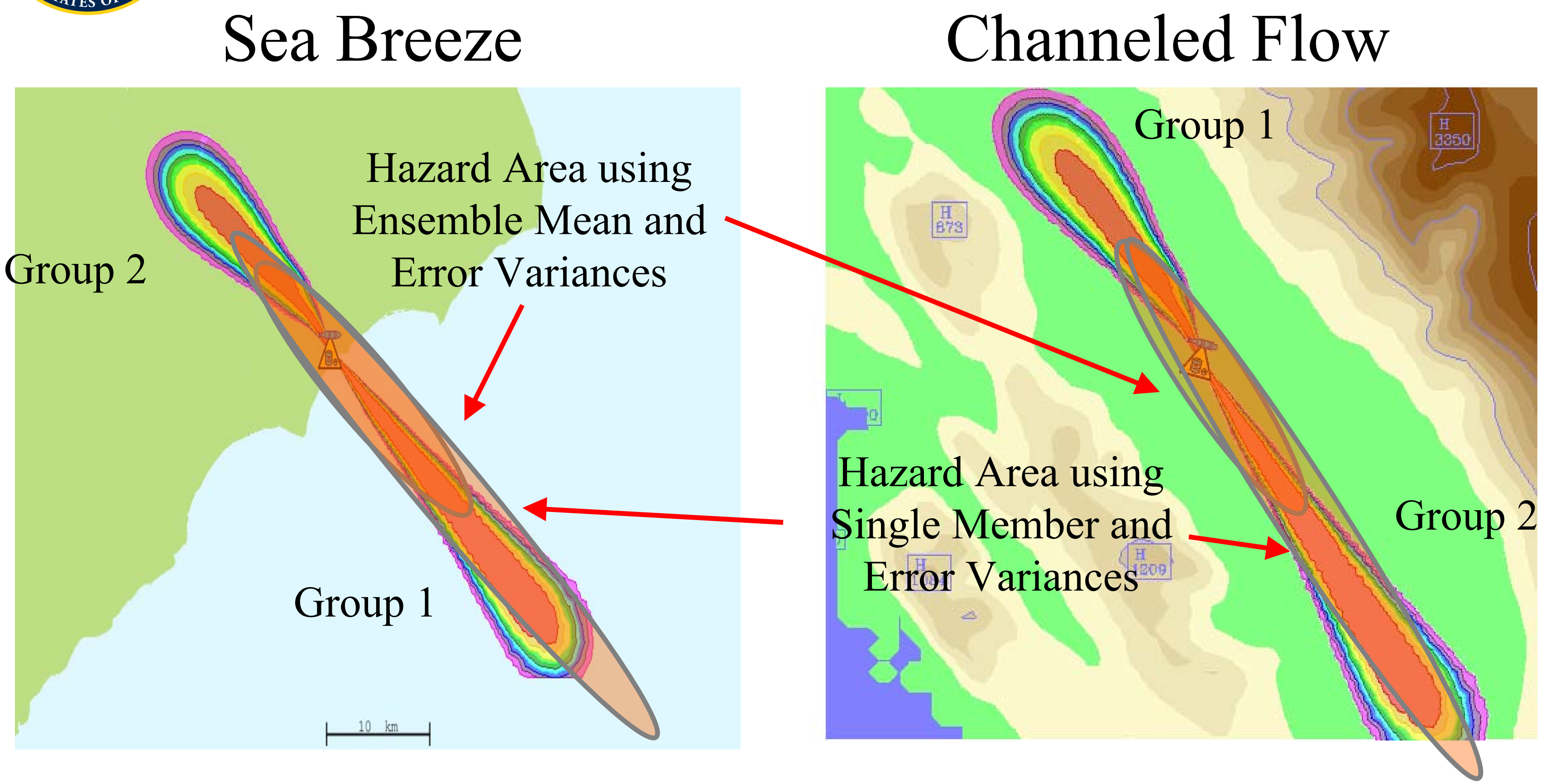
48 Hour Forecast

Relating Correlation and Distance





Limitations of a Single AT&D Run



No amount of broadening of a single plume can accurately predict the shape and extent of the hazard area for these bimodal cases! Covariance information can be combined with an AT&D ensemble to provide a more accurate prediction.

Conclusions

- The Roulston (2005) binning method allows us to recover potentially useful relationships from highly scattered data by binning similar points predictor values.
- When applied to ensemble variances, this method allows us to identify a simple, computationally inexpensive, linear relationship between the ensemble variance and actual variance that can be used to calibrate ensemble output for use in SCIPUFF.
- When applied to ensemble covariances, the binning technique reveals more diffuse and less linear plots, however the actual covariance range is much smaller than the ensemble covariance range. If a control member is used rather than an ensemble mean, a clear linear relation is recovered.
- There is a clear relationship between distance separation and ensemble spread correlation. The exact length scale depends on the correlation value you choose as a cutoff; a 0.2 cut-off yields a distance length of ~500km for 15 hPa AGL winds

Future Work

- Examine the capability of the linear calibration method with an ensemble more tuned for PBL parameters
- Explore dependence (if any) on grid resolution and domain size
- Increase the length of the training period and determine any seasonal divisions needed
- Evaluate the effectiveness for other variables and levels, including p-level vs. σ-level considerations
- Continue investigating the use of covariance/distance binning for calculating SLE
- Test implementation of the calibration by using it to calculate UUE for SCIPUFF and compare to SCIPUFF ensemble

Acknowledgements

- Mark Roulston for stressing the importance of covariance information and of calibration, and starting us down the binning path
- Ian Sykes for many discussions about SCIPUFF parameterizations of uncertainty
- Jeff McQueen et al. at NCEP for full ensemble data from SREF