

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 2/23/2006		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 5/1/2002 - 12/31/2005	
4. TITLE AND SUBTITLE Interpreting, Improving, and Augmenting Multi-Model Ensembles				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER N00014-02-1-0473	
				5c. PROGRAM ELEMENT NUMBER 05PR04967-01	
6. AUTHOR(S) Hansen, James A.				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Massachusetts Institute of Technology 77 Massachusetts Avenue Cambridge, MA 02139				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research Ballston Centre Tower One 800 North Quincy Street Arlington, VA 22217-5660				10. SPONSOR/MONITOR'S ACRONYM(S) ONR	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT No Limitations DISTRIBUTION STATEMENT A Approved for Public Release Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Multi-Model numerical weather prediction (NWP) ensembles outperform even the world's most sophisticated single-model ensemble prediction system. An ensemble whose members are the control forecasts of a number of operational NWP centers, the poor-man's multi-model (PM MM) ensemble, samples from both initial condition space and model space. This proposal addresses two methods for intelligently adding new ensemble members to multi-model ensembles. The first method utilizes multi-model ensemble forecasts in a lagged average manner. Ensemble sizes are increased by combining ensemble forecasts launched at different times but valid at the same verification time. To avoid the limitations inherent in combining ensemble forecasts that utilize different numbers of observations, an ensemble transform Kalman filter (ET KF) is utilized to incorporate observations into existing ensemble forecasts without re-running any NWP models. The transformation conditions all ensemble members on the same number of observations. The second method is to use a simplified model to mimic the behavior of the PM MM ensemble. Parametric perturbations are used to produce states for the simplified model that populate the tails of the PM MM ensemble, increasing the PM MM ensemble spread					
15. SUBJECT TERMS Multi-Model Ensembles					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Mike Corcoran, OSP MIT
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code) 617-258-6685

Abstract: Hansen Grant#N00014-02-1-0473

Multi-Model numerical weather prediction (NWP) ensembles outperform even the world's most sophisticated single-model ensemble prediction system. An ensemble whose members are the control forecasts of a number of operational NWP centers, the poor-man's multi-model (PM MM) ensemble, samples from both initial condition space and model space. This proposal addresses two methods for intelligently adding new ensemble members to multi-model ensembles.

The first method utilizes multi-model ensemble forecasts in a lagged average manner. Ensemble sizes are increased by combining ensemble forecasts launched at different times but valid at the same verification time. To avoid the limitations inherent in combining ensemble forecasts that utilize different numbers of observations, an ensemble transform Kalman filter (ET KF) is utilized to incorporate observations into existing ensemble forecasts without re-running any NWP models. The transformation conditions all ensemble members on the same number of observations.

The second method is to use a simplified model to mimic the behavior of the PM MM ensemble. Parametric perturbations are used to produce states for the simplified model that populate the tails of the PM MM ensemble, increasing the PM MM ensemble spread and its ability to bound truth. In addition, inspection of the dominant parametric perturbations utilized by the simplified model will provide information about the difference between PM MM ensemble members and about the features necessary in any new model development.

FINAL REPORT

Interpreting, Improving, and Augmenting Multi-Model Ensembles

James A. Hansen

MIT, EAPS, 54-1616, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Phone: (617) 452-3382 Fax: (617) 253-8298 Email: jhansen@mit.edu

Award Number: N000140210473

1. LONG-TERM GOALS

To develop methods to intelligently add new ensemble members to multi-model ensemble forecasts, to maximally exploit existing multi-model ensemble forecasts, and to diagnose model inadequacies and differences through multi-model ensemble forecasts.

2. OBJECTIVES

The two primary objectives of this project were:

1. Exploiting existing multi-model ensemble analyses and forecasts
Extract as much information as possible from the analyses and forecasts currently available from different operational Numerical Weather Prediction (NWP) centers.
2. Use of a single model structure to augment and interpret multi-model results
Adjust the parameters of a single, simplified model to mimic the output of the more complex NWP models, and exploit the resulting parametric information for ensemble augmentation and for interpretation of the differences between the models making up the ensemble.

3. APPROACH

3.1 Exploiting existing multi-model ensemble analyses and forecasts

It is operationally impossible to maintain a multi-model development, data assimilation, and forecasting system at a single NWP center. This motivates extracting as much information as possible from the analyses and forecasts currently available from different operational NWP centers. This collection of analyses and forecasts from different NWP centers is denoted the poor man's multi-model (PM MM) ensemble. Because the existing PM MM ensemble has few members (there are only a handful of operational NWP centers around the world), methods for extracting as much information as possible from the ensemble are of interest. To increase the effective ensemble size without adding additional models, this project explored implementing a transformed lagged ensemble forecasting (TLEF) technique where forecasts launched at different times are combined at common verification times. Because forecasts at longer leads lack the observational information available to short lead

forecasts, the ensemble transform Kalman filter (ET KF) (Bishop *et al*, 2001) was utilized to incorporate observations into existing ensemble forecasts. The *state-dependent* scaling and *rotation* employed by the TLEF results in a significant improvement to traditional lagged ensemble methodologies where, at best, only a climatological weighting is applied to older forecasts. In the TLEF approach, each forecasts' ensemble member is conditioned on the same amount of information, regardless of its lead time.

3.2 Use of a single model structure to augment and interpret multi-model results

Designing an atmospheric GCM from scratch with the aim of optimally augmenting an existing PM MM ensemble is far beyond the scope of this project. Instead, a single (simple) model structure was used to model the output of the more complex PM MM ensemble members. A given set of PM MM ensemble analyses were used to determine the parameter perturbations necessary in the simple model to produce ensemble forecasts constrained to lie in the subspace spanned by the PM MM ensemble forecasts. The existing PM MM ensemble were augmented by perturbing the simple model's parameters in the direction of these "parametric singular vectors" and produce model states that expand the PM MM ensemble distribution. In addition, insight into the difference between the models in the PM MM ensemble is gained by examining the required parametric perturbations.

The focus of the second objective evolved during the course of the project. Parameter estimation is a key aspect of the objective, and the proper interpretation of multi-parameter estimation under the influence of structural model error is an under-developed field. Experiments were carried out an effort to understand the extent to which distributions of probabilistically tuned parameter values provide information about a model's inadequacies. These experiments guided the application of the probabilistic parameter estimation approach to the tuning of parameterizations found in complex models of the atmosphere.

4. WORK COMPLETED

4. 1 Exploiting existing multi-model ensemble analyses and forecasts

The work to exploit existing multi-model ensemble analyses and forecasts was shown to be productive for simple models, for output from the ECMWF ensemble prediction system (EPS), and for a multi-model collection of ensemble forecasts. Single model results describing the technique are in press, results from the ECMWF EPS are complete, and experiments on PM MM ensembles provided by operational NWP centers are currently being carried out. Proof-of-concept results motivated the UK Meteorological office to explore implementing the technique for use in their observation targeting studies. This has lead to further collaboration that aims to assess whether the technique can be used at the regional level to increase limited area model ensemble sizes, and whether the technique can be used to generate larger ensembles of boundary conditions to drive the limited area models. The approach and some representative results are presented below.

The TLEF is based on the ensemble transform Kalman filter (ETKF) of Bishop *et al* (2001). A transformation based on the ETKF is performed on old ensemble forecasts each time new observations become available. The information contained in the transformation is then propagated forward in the

ensemble subspace under the assumption of linear error evolution to alter the ensemble forecasts. The ensemble forecast itself provides all the information necessary to propagate the impact of the observations forward in time; re-running the numerical model is not necessary. This process results in ensemble forecasts that are directly comparable at any verification time; all ensemble members have been conditioned upon the same number of observations. This approach effectively increases the forecast ensemble size without adding any new model runs. The most recent ensemble forecast will provide an estimate of $p(\mathbf{x}'(\tau_{ver} | \mathbf{Y}(0)))$: the probability of the true state of the system at τ_{ver} given all observations up to and including those at $t = 0$. The ensemble forecast launched one observation time prior to the most recent provides an estimate of $p(\mathbf{x}'(\tau_{ver} | \mathbf{Y}(-\tau_{obs})))$, a different distribution from that which is desired. The aim of the TLEF is to transform the samples from $p(\mathbf{x}'(\tau_{ver} | \mathbf{Y}(-\tau_{obs})))$ so that they become samples from $p(\mathbf{x}'(\tau_{ver} | \mathbf{Y}(0)))$ and can be used to augment the most recent ensemble forecast. Ensembles can be transformed multiple times. Each time new observations become available, previously transformed ensembles can undergo another transformation. The number of times an ensemble forecast can be transformed is controlled by the validity of the underlying linearity assumption, and the level of model error.

The ETKF has been applied to the targeted observation problem (Bishop *et al*, 2001), can be used for ensemble generation (Wang and Bishop, 2003), and can be used for data assimilation as well. In this work we use the ETKF to propagate observational influence forward in time. The TLEF uses the ETKF to update a forecast mean and ensemble perturbations when observations become available, and then propagates those changes forward in the ensemble subspace under the assumptions that they evolve linearly. Essentially, the transformation matrix that is used at the observation time to transform forecast ensemble perturbations to analysis ensemble perturbations is applied to each successive forecast time.

The ensemble transform Kalman filter is an approximation of the extended Kalman filter (Jazwinski, 1970). Details on the ETKF can be found in Bishop *et al* (2001), but a brief description is given below.

In the context of the ETKF, a minimum error variance state estimate is given by

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{H} \bar{\mathbf{x}}^f),$$

where $\mathbf{H}$ is a (linear) operator that maps from model space to observation space, $\mathbf{R}$ is the observational error covariance matrix, $\mathbf{y}$ is an observation, and $\bar{\mathbf{x}}^f$ is the ensemble mean forecast. The forecast error covariance matrix, $\mathbf{P}^f$, is estimated from the distribution of ensemble members:

$\mathbf{P}^f = \mathbf{X}^f \mathbf{X}^{fT}$. The $\mathbf{X}^f$ is a matrix whose rows contain the normalized ensemble forecast perturbations, $\frac{1}{\sqrt{K-1}}(x_i - \bar{x}^f)$, where K is the ensemble size, and i represents a particular ensemble member.

This equation updates the mean of the ensemble forecast. The analytic expression for the uncertainty associated with this estimate is given by

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{P}^f.$$

Herein lies the crux of the ETKF: the normalized forecast ensemble perturbations, $\mathbf{X}^f$, are *linearly transformed* so that the resulting covariance exactly matches the one given by the equation above:

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{P}^f = \mathbf{X}^f \mathbf{T} \mathbf{T}^T \mathbf{X}^f.$$

Once equation this equation is solved for $\mathbf{T}$ (see Bishop *et al*, 2001 for details), the normalized ensemble analysis perturbations are given by $\mathbf{X}^a = \mathbf{X}^f \mathbf{T}$. The normalization is removed, and the columns of the unnormalized $\mathbf{X}^a$ are added to the minimum error variance estimate of the state, $\bar{\mathbf{x}}^a$, to produce the new analysis ensemble.

To obtain a feel for how the TLEF works in this framework, consider the ensemble perturbations. We obtain the analysis perturbations through $\mathbf{X}^a(0) = \mathbf{X}^f(0)\mathbf{T}(0)$. These are perturbations, and can be propagated forward to future times by using a linear uncertainty propagator, $\mathbf{M}(0, \tau)$, such that

$$\mathbf{X}_{\text{new}}^f(\tau) = \mathbf{M}(0, \tau) \mathbf{X}^a(0).$$

But we have an expression for $\mathbf{X}^a$ in terms of $\mathbf{T}$, and after substituting we obtain

$$\mathbf{X}_{\text{new}}^f(\tau) = \mathbf{M}(0, \tau) \mathbf{X}^f(0) \mathbf{T}(0).$$

If we believe that the ensemble perturbations evolve linearly in the ensemble subspace, then

$$\mathbf{M}(0, \tau) \mathbf{X}^f(0) = \mathbf{X}^f(\tau)$$

and our new perturbations are nothing more than

$$\mathbf{X}_{\text{new}}^f(\tau) = \mathbf{X}^f(\tau) \mathbf{T}(0).$$

All one has to do to propagate the influences of the observations forward is operate on the normalized ensemble perturbations at time $t = \tau$ with the transformation matrix obtained at the observation time, $\mathbf{T}(0)$. There in an analogous approach to transforming the ensemble mean forecast. This implies that the ensemble forecast launched at the previous observation time can be transformed using the most recent observation, and the resulting forecast ensemble perturbations added to either the transformed mean or the mean of the most recent ensemble forecast. This doubles the size of the ensemble. Older ensemble forecasts can be transformed multiple times and used to further increase the ensemble size.

Figure 1 demonstrates the utility of the TLEF in the simple model and the multi-model ensemble forecasting framework. The simple model utilized is a 100-d version of the Lorenz (1996) system. Panel a) in figure 1 shows how a TLEF forecast is an improvement over a pre-transformed forecast. The heavy solid line gives the ensemble mean RMSE with respect to truth as a function of lead. This forecast was launched at $t = -2$ model days, and run forward to $t = +6$ model days. Negative numbers indicate times before the most recent observation, taken at $t = 0$. The dotted line is the RMSE that results from applying the TLEF to the original forecast once a new observation becomes available at $t = -1.75$ days. Note that it is an improvement over the original forecast at all leads shown (it eventually becomes worse at around $t = 11$ days). The dash-dot line is the RMSE that results from applying the TLEF to the ensemble represented by the dotted line once a new observation becomes available at $t = -1.5$ days. It too is an improvement over the ensemble it transformed for all leads shown. Each additional line represents subsequent transformations. The thin solid line is the RMSE that results after the original forecast (thick solid line) has been transformed eight times. An investigation of the ensemble statistics shows that they have correct second moments.

Panel b) in figure 1 shows how a TLEF forecast is an improvement over the pre-transformed forecast when a multi-model ensemble is constructed using forecasts from the UKMO, ECMWF, DWD, NCEP, and MeteoFrance. The black line is the ensemble mean RMSE with respect to the ECMWF analysis for the original forecast launched at $t = -4$ days. The blue line gives the ensemble mean RMSE of the

forecast transformed by the observations that become available at $t = -3.5$ days. The red line shows that ensemble mean RMSE that result from transforming the previously transformed ensemble once observations become available at $t = -3$ days; the red line represents an ensemble forecast that has been transformed twice. Note that transformed forecasts become worse than their untransformed counterparts after approximately 4 days. This indicates a breakdown of the assumptions implicit in the TLEF, including linearity, a perfect model, and an ensemble that spans all relevant subspaces.

4.2 Use of a single model structure to augment and interpret multi-model results

Fundamental to the ensemble-based approach to parameter estimation is the concept that whenever one's model is structurally inadequate, one should not expect to obtain a fixed set of parameter values. Instead, one should expect the parameters to change as a function of time in order to offset the state-dependent nature of structural inadequacy. The investigation of the resulting distribution of tuned parameter values is beneficial for two reasons: 1) It provides a basis for the stochastic modeling of model error, and 2) It can provide information about how the model behaves and what might be wrong with the model.

4.2.1 Stochastic differential equations

To demonstrate point 1, an experiment was carried out where the "truth" was defined to be a set of nonlinear stochastic differential equations (SDEs) with multiplicative noise. The system was modeled using a set of ordinary differential equations (ODEs) with a function form identical to the SDEs. The scalar parameters of the ODEs corresponded to stochastic parameters in the SDEs. Observations were drawn from the SDEs and ensemble-based parameter estimation was performed using the ODEs. The aim was to use ensemble-based parameter estimation techniques to try and uncover the correct form of the stochasticity in the SDEs. It was found that the resulting tuned parametric distributions provided no information about the true stochasticity of the SDEs, but they provided an excellent means of modeling the stochasticity of the SDEs. It was found that if one sampled from the parametric distribution during forecasting in the same manner as the distribution was determined, then statistically correct second moment ensemble forecast statistics were achieved. That is, the correct way to sample from the distribution was to make a random draw, integrate the model forward for six hours holding the parameter constant, and then making another random draw. This result is shown graphically in Figure 2. Plotted is the standard deviation of the ratio of the ensemble mean forecast error to the ensemble spread as a function of forecast lead. An ensemble forecasting system with correct second moment statistics will have a value of one. The green and black curves take on this value of one. The curves are for the method of sampling described above, and for a perfect model, respectively.

4.2.2 A parameterization of boundary layer clouds

A goal of the stated second objective of this project was to understand the ways in which models are wrong. Rather than applying to a simplified model in an effort to mimic a more complex model (as originally proposed), the choice was made to instead apply the approach to important components of complex models: their parameterizations. The aim became to produce the time-varying distribution of parameter values used by a parameterization that provide the best fit to observations. The resulting distributions are analyzed to better understand the limitations of the parameterizations, and ultimately will be applied to the ensemble forecasting problem as the basis for stochastic parameterizations.

I began by applying the technique to the Higher Order Closure (HOC) boundary layer cloud and turbulence parameterization of Golaz and Larson (2002). Rather than observations, the parameters of the parameterization were tuned so that the parameterization best mimicked the statistics of a cloud resolving, large eddy simulation (LES). Figure 3 plots the state-space results of an ensemble approach to parameter estimation. Panel a) plots cloud fraction, panel b) plots liquid water mixing ratio, and panel c) plots third-order moment of vertical velocity. Thin blue lines are the model states associated with tuned parameter values, and the thick red line represents truth (as defined by the LES). The model does an excellent job reproducing the observations in this case.

Figure 4 plots the correlation matrix associated with the ensemble of tuned parameter values and one element of the model state. The lower right quadrant shows the correlation between fit parameter values and provides information about how the model has to configure itself to best mimic observations. The upper left quadrant shows the correlation between errors in the tuned state of the model (cloud fraction, in this case). The upper right and lower left quadrants show the **cross-correlation** between model parameters and model state. These quadrants provide quantitative information about how the parameters have changed in order to best fit the state, and provide significant insight into the limitations of the parameterization. The black dashed lines serve only to guide the eye.

Figure 5 provides a concrete example of the utility of the approach for diagnosing model inadequacy. The parameters of the model were tuned for two different atmospheric cases: trade wind cumulus and nighttime stratus. The distribution of parameter values for each are shown as blue lines and red lines, respectively. Note that for one of the parameters in particular, very different parameter values are needed to best fit the observations. The offending parameter is closely associated with the skewness of distributions of vertical velocity (very important for correctly modeling cumulus clouds). The model struggles to generate enough skewness to mimic the cumulus case, providing information to the model developers to help them improve their model.

4.2.3 NOGAPS convection parameterization

Wanting to work with parameterizations in an operational model, I leveraged this YIP to obtain funding for a six month visit to NRL Monterey. During my visit I implemented the ensemble-based parameter estimation technique using a single column version of NOGAPS. I have focused on the Emanuel convection scheme (Emanuel, 1991), but also invested the effort necessary to also focus on the Webster gravity wave drag parameterization (Webster et al, 2002), and the Kershaw convective momentum transport parameterization (Kershaw, 1997). An example of results is given in figure 6. Plotted is the 2d distribution of parameter values associated with the fractional area of unsaturated downdrafts (sigd) and the fractional of precipitation falling outside of a cloud (sigs) resulting from a perfect model experiment. The fact that all ensemble members did not converge to the correct parameter value indicates an important dynamic relationship between the two parameters. While the parameters are used throughout the parameterization, this figure indicates that their dominant contribution is to an equation governing precipitation. In that equation, sigd and sigs exist as a product, and the curve in figure 6 is consistent with a relationship such as $\text{sigd} \times \text{sigs} = \text{constant}$.

4.2.4 Probabilistic tropical cyclone track forecasts

Parameter estimation has many applications. While at NRL MRY I worked with Jim Goerss to explore the application of parameter estimation ideas to the problem of probabilistic Tropical Cyclone (TC) track forecasts. Goerss' work to predict TC track error takes a regression approach to relate predictors such as multi-model ensemble spread, initial hurricane intensity, and latitudinal displacement to the realized track error (Goerss, 2006). For every storm, Goerss produces an error radius relative to the multi-model ensemble mean that is tuned to bound the 70% contour interval: the verifying forecast is expected to fall within that bound 7 times out of 10. While demonstrably superior to current operational approaches, Goerss' approach is limited in by it's implicit Gaussianity assumption; the errors are, by construction, assumed to be isotropic relative to the ensemble mean.

To address this limitation we explored a kernel dressing methodology that permits probabilistic track forecasts of arbitrary functional form. The approach is to center a Gaussian distribution on top of each multi-model track forecast and tune the spread of the Gaussian and its weight (subject to all weights summing to 1) so that a probabilistic measure of forecast quality is optimized. In the current work we

chose to optimize ignorance, $Ign = -\sum_{i=1}^N \log(p_{ver,i})$, where N is the number of forecast realizations, and

p_{ver} is the forecast probability of the verifying observation. Minimizing ignorance encourages the production of forecast PDFs that maximize the forecast probability of the verification. In the kernel density approach, the forecast PDF is obtained by summing the probabilities associated with each of the Gaussian kernels. The optimization condition that that sum of the weights of the kernels be 1 means that the total probability is guaranteed to be 1. In this application the parameters of interest aren't part of a physical model, but instead they are the spreads and weights (and eventually the biases and covariances) of the kernel Gaussians.

An example of the types of forecasts obtained by the kernel dressing algorithm is given in figure 7. Tuned using track forecasts and verification over the period 2000-2004 for all forecasts that included the NCEP, NOGAPS, GFDL, and UKMO forecasts, an example is given for a 72hr forecast of hurricane Wilma. The blue squares are the multi-model ensemble forecast of the NCEP, GFDL, NOGAPS, and UKMO models, the magenta triangle is the ensemble mean, and the green circle is the verifying observation. The black circle is the 70% isopleths of the Goerss probabilistic forecast and the red curve is the 70% isopleths of the kernel dressing probabilistic forecast. It was clear at this point in Wilma's development that the storm would track over southern Florida, but there was great uncertainty in its speed. The isotropic assumption implicit in the Goerss method is unable to capture this uncertainty in speed, but the kernel density method captures the along-track spread.

Objective verification of the two methods shows that both are equally reliable, and that the Goerss method is less ignorant at short forecast leads while the kernel dressing method is less ignorant at longer forecast leads. Extensions to the kernel dressing method are planned to attempt to incorporate the non-ensemble information utilized by the Goerss method. The intension is to submit a proposal to the joint hurricane testbed facility in an effort to transition the results of the kernel dressing algorithm to operations.

5. RESULTS

- The TLEF was applied to output from idealized models, the ECMWF EPS, and a multi-model ensemble made up of operational models, and successfully increased the forecast ensemble size with a minimal increase in computational cost.
- Ensemble-based approaches to parameter estimation can be used both to identify structural model inadequacies (make the model better) and to provide a basis for stochastic perturbations that are capable of modeling any remaining model inadequacy (make ensemble forecasts better).
- Ensemble-based approaches to parameter estimation have been successfully applied to a boundary layer cloud parameterization (HOC) and to the Emanuel convection parameterization in the context of the NOGAPs NWP model.
- Parameter estimation has been successfully applied to the generation of skillful continuous probabilistic forecast distributions of TC track.

6. IMPACT/APPLICATIONS

The demonstrated ability of the TLEF to increase global model ensemble sizes will be leveraged for the purpose of forcing large limited area ensemble forecasts. The PIs work on the tropical observability and predictability problem will require large ensemble forecasts to be run for the purpose of ensemble-based data assimilation. The TLEF will enable each limited area ensemble member to have a unique, dynamically consistent boundary condition. The PIs parameter estimation work will also be leveraged for the tropical observability and predictability work in an effort to obtain the best possible forecast model of the tropical atmosphere.

The PI implemented a methodology and in situ software infrastructure for use by NRL MRY scientists. The infrastructure can be used to improve any parameterization utilized by NOGAPS as well as provide the basis for stochastic parameterizations.

The work on the production of probabilistic TC track forecasts will be extended with the aim of its being implemented as an official form of “guidance of guidance” available to hurricane forecasters.

7. TRANSITIONS

Currently none in the Navy sense, but work is underway to move towards implementing the TLEF for use in providing a large ensemble of boundary conditions for the Joint Ensemble Forecasting System (JEFS).

8. RELATED PROJECTS

The PI is associated with an NSF-funded project that aims to address the impact of model inadequacy in data assimilation and forecasting using a single model structure. Model inadequacy insights gained during the NSF project will be applied to the current project.

The PI has recently learned that a joint proposal with Greg Hakim at the University of Washington (Tropical Observability and Predictability) will be funded by ONR. This project will utilize the TLEF for the generation of large numbers of ensemble boundary conditions for limited area model ensemble forecasts. It will also utilize the parameter estimation work to produce the best forecast model possible of the tropical atmosphere.

9. REFERENCES

Bishop, C.H, B.J. Etherton, and S.J. Majumdar (2001) Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, vol. 129, pages 420-436.

Emanuel, K.A. (1991) A Scheme for Representing Cumulus Convection in Large-Scale Models. *Journal of the Atmospheric Sciences*, **48**, 2313-2335.

Goerss, J. (2006) Prediction of consensus tropical cyclone track forecast error. Submitted to *Monthly Weather Review*.

Golaz, J.C., V.E. Larson, and W.R. Cotton (2002) A PDF-based model for boundary layer clouds. Part I: Method and model description. *Journal of the Atmospheric Sciences*, **59**, 3552-3572.

Kershaw, R. and D. Gregory (1997) Parameterization of momentum transport by convection. I: Theory and cloud modeling results. *Quarterly Journal of the Royal Meteorological Society*, **123**, 1133-1151.

Lorenz, E.N. (1996) Predictability – A Problem Partly Solved. ECMWF Seminar Proceedings, Shinfield Park, Reading, RG2 9AX, UK.

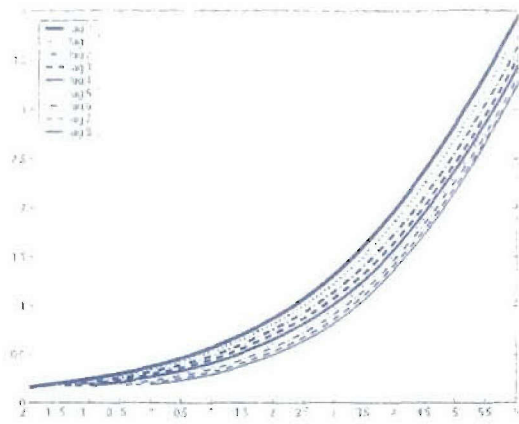
Wang, X. and C.H. Bishop (2003) A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes. *Journal of the Atmospheric Sciences*, vol. 60, pages 1140-1158.

Webster, S., A. R. Brown, D. R. Cameron, and C. P Jones, (2002) Improvements to the representation of orography in the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, **129**, 1989-2010.

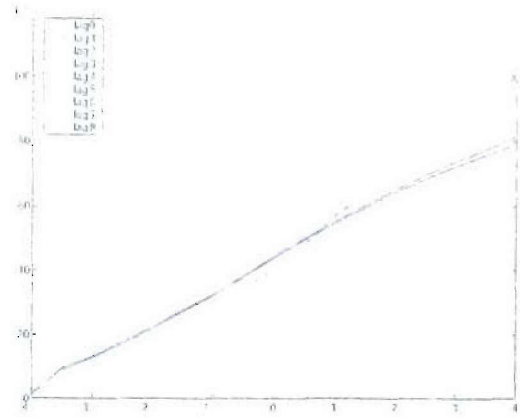
10. PUBLICATIONS

Gombos, D., J.A. Hansen, J. Du, and J. McQueen (2006) Theory and Applications of the Minimum Spanning Tree Rank Histogram. *Monthly Weather Review*. [submitted]

- Moskaitis, J., and J.A. Hansen (2006) Deterministic forecasting and verification: A busted system? *Weather and Forecasting*. [submitted]
- Lawrence, A., and J.A. Hansen (2005) A Transformed Lagged Ensemble Forecasting Technique for Increasing Ensemble size. *Monthly Weather Review*. [submitted]
- Hansen, J.A. and C. Penland (2005) Parameter estimation: To what extent can data assimilation techniques correctly uncover stochasticity? *Physica D*. [submitted]
- Hansen, J.A. and C. Penland (2005) Efficient approximate techniques for integrating stochastic differential equations. *Monthly Weather Review*. [accepted]
- Lawson, W.G. and J.A. Hansen (2005) Alignment error models and ensemble-based data assimilation, *Monthly Weather Review*, 133, 6, 1687-1709. [published]
- Khade, V.M. and J.A. Hansen (2004) State dependent predictability: Impact of uncertainty dynamics, uncertainty structure and model inadequacies. *Nonlinear Processes in Geophysics*, **11**, 351-362. [published]
- Smith, L.A. and Hansen, J.A. (2003) Extending the Limits of Forecast Verification with the MST, *Monthly Weather Review*, **132**, 6, 1522-1528. [published]



a)



b)

Figure 1: Example of the utility of the TLEF. Panel a) plots ensemble mean RMSE as a function of number of transformations in a simple model (Lorenz, 1996), and panel b) plots ensemble mean RMSE as a function of number of transformations from a multi-model NWP ensemble using ECMWF, UKMO, NCEP, DWD, and MeteoFrance model output.

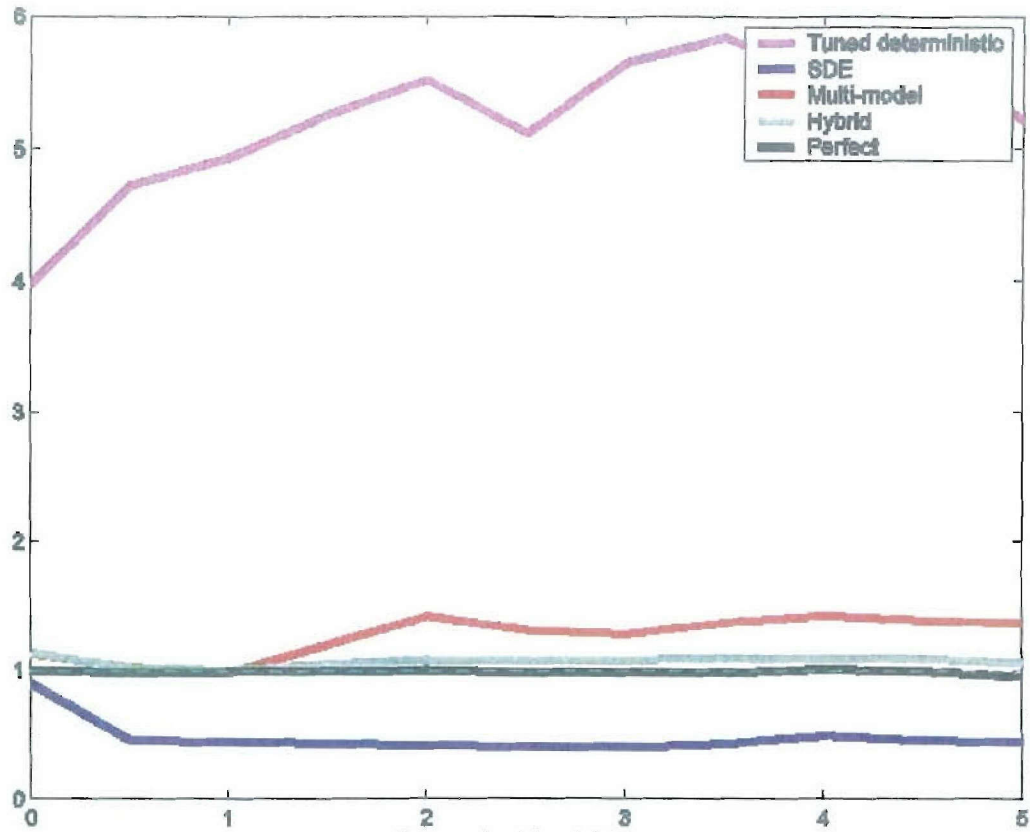


Figure 2: A plot of the standard deviation of the ratio of ensemble mean error to ensemble mean spread as a function of forecast lead time. A value of 1 means that the second moment forecast statistics are correct. Four different methods to account for model inadequacy using a tuned distribution of parameter values are plotted along with results from a perfect model. It is found that by sampling from the tuned distribution in exactly the same manner as the distribution was generated (i.e. parameter values are held constant for six hours before being changed), then nearly perfect second moment forecast statistics are produced, as shown by the green curve.

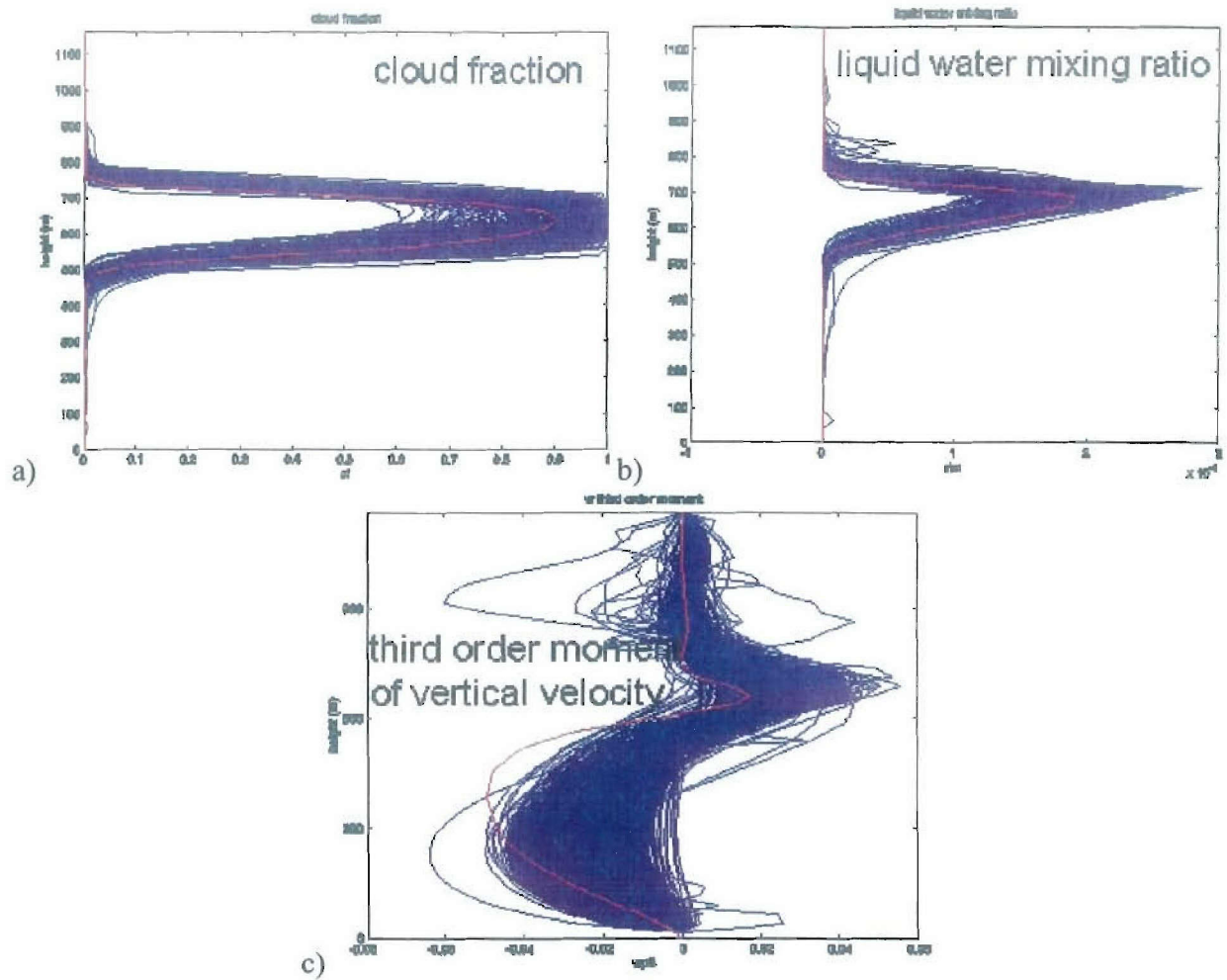


Figure 3: The ensemble-based parameter estimation scheme is applied to the boundary layer cloud and turbulence parameterization of Golaz and Larson (2002), called HOC. Model parameters were tuned so that the resulting model states were driven as close as possible to the states produced by a cloud resolving large eddy simulation. State space spaghetti plots generated by running the tuned parameter values through the HOC model are plotted in blue, while the “true” LES state is given in red. Panel a) plots the profile of cloud fraction, panel b) the profile of liquid water mixing ratio, and panel c) the profile of the third order moment of vertical velocity. Good fits are achieved in this case.

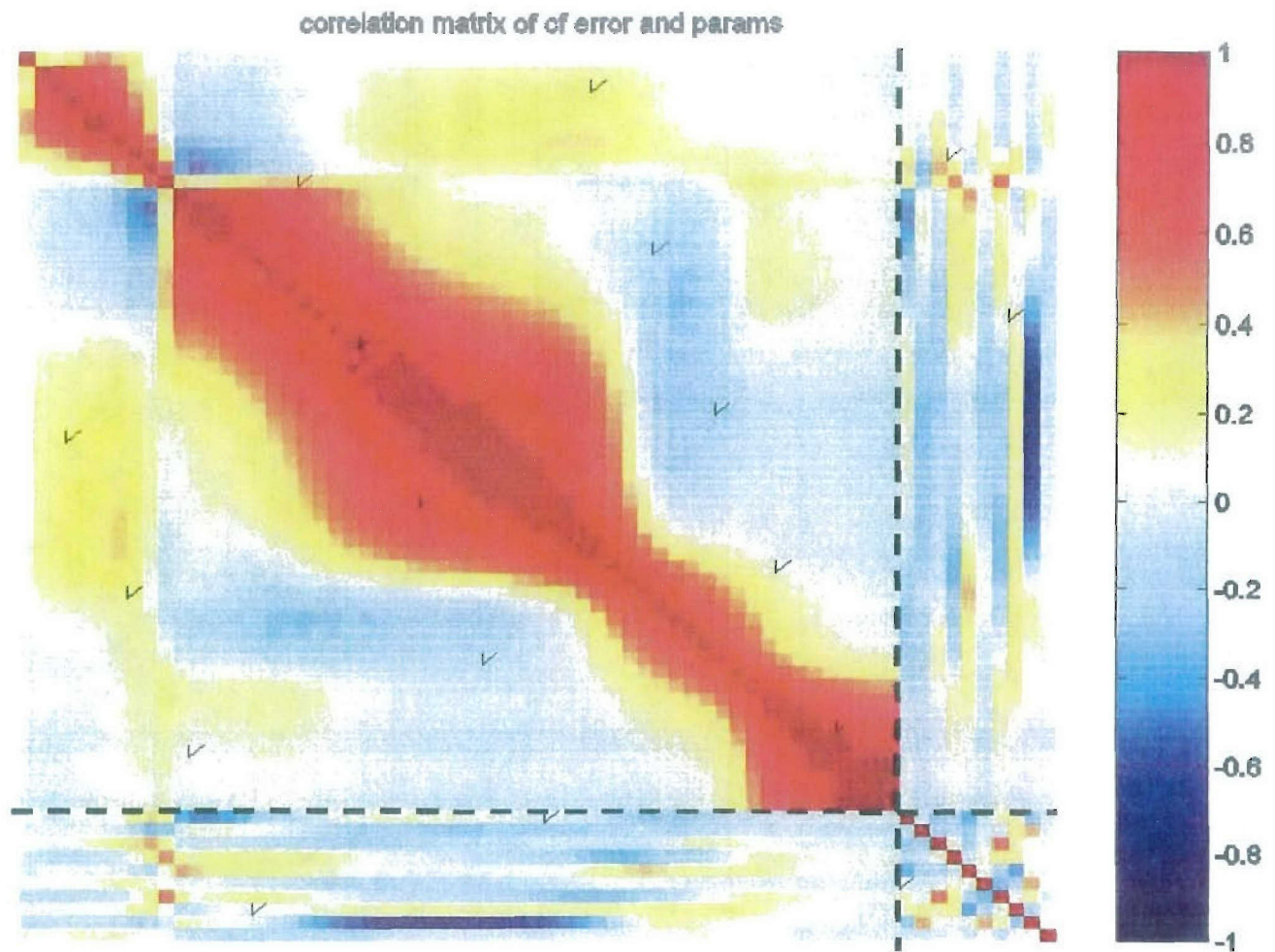


Figure 4: Correlation matrix relating the ensemble of tuned parameter values to the associated profiles of cloud fraction error. The upper left quadrant plots the cloud fraction error correlation matrix, and the lower right plots the parameter correlation matrix (note the significant off-diagonal structure). The upper right and lower left quadrants describe how the parameter distribution correlates with the cloud fraction error distribution. These types of diagnostics can be of great value to model developers.

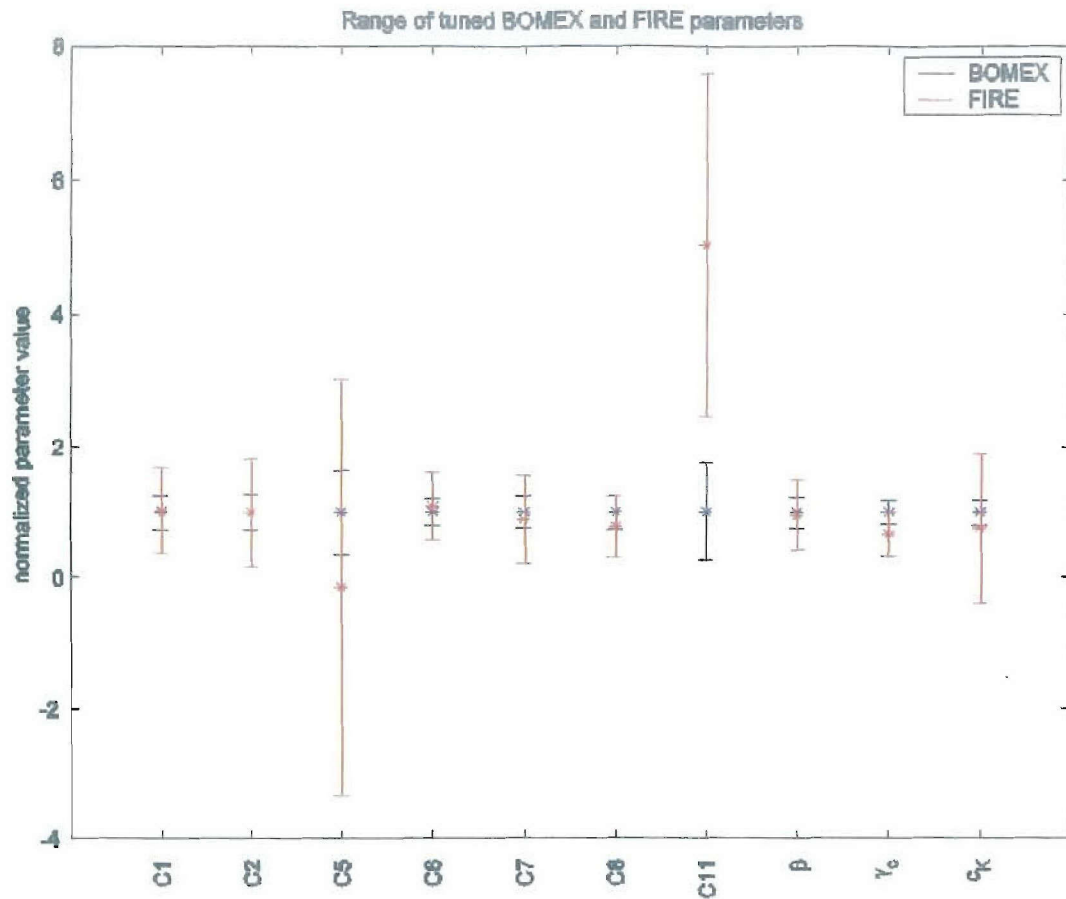


Figure 5: Plot of the marginal distributions of tuned parameters for two different forcing cases: trade wind cumulus (BOMEX) in blue, and nighttime stratocumulus (FIRE) in red. Note the significant difference in the C11 parameter. C11 helps control the magnitude of the skewness of the distribution of vertical wind in the model gridbox. The model developers have used this information to focus their efforts on improving the representation of the third order moment of vertical velocity in the model.

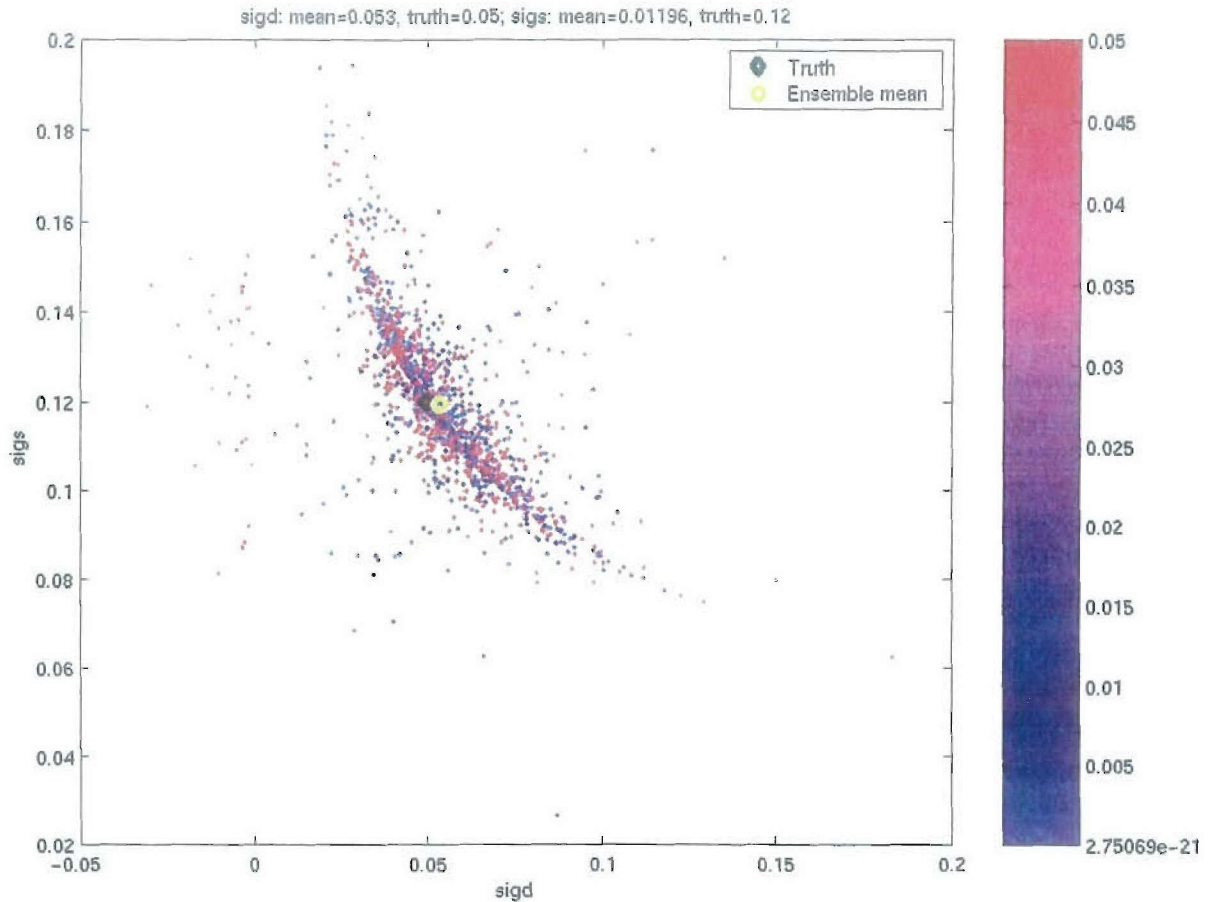


Figure 6: Preliminary perfect model results from ensemble-based parameter estimation in the Emanuel convection scheme as implemented in NOGAPS. The sigd parameter controls the fractional area of unsaturated downdrafts, while the sigs parameter controls the fraction of precipitation that falls outside of the parameterized clouds. Each point represents the parameter values for a single ensemble member, and points are colored by cost function value. Even though this is a perfect model experiment, all ensemble members do not converge to the correct parameter combination due to the structure of cost function space. Upon investigation, it is found that the curve in the figure is consistent with the dynamic relationship between the two parameters in one of the parameterization's equations: $\text{sigd} \cdot \text{sigs} = \text{constant}$. This demonstrates a powerful diagnostic quality of the ensemble approach to parameter estimation.

Tuned IGN: 11.6754, base IGN: 11.3941

72hr lead, storm 24, dtg 5102106

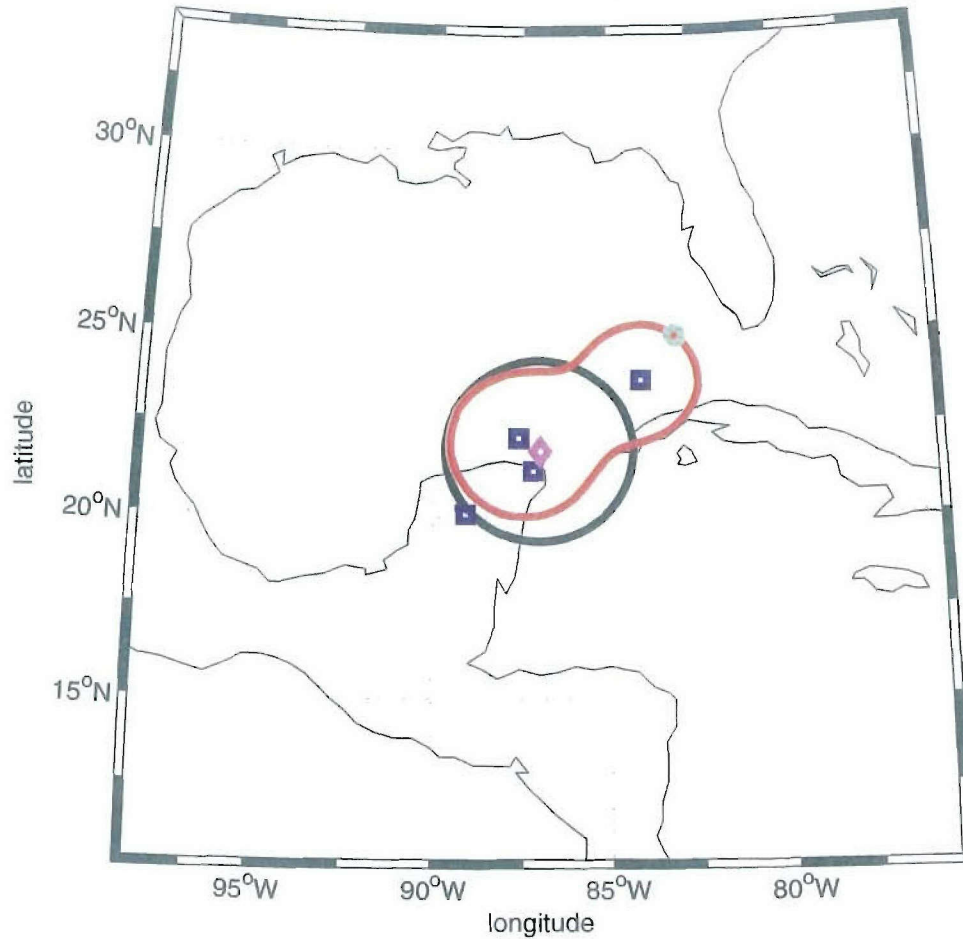


Figure 7: An example of a probabilistic track forecast produced by a multi-model kernel dressing technique for a 72hr forecast of Wilma (2005). The black circle is the 70% isopleths of uncertainty provided by the Goerss method. The red curve is the 70% isopleths provided by the kernel dressing method. The blue squares are the individual multi-model ensemble track forecasts, the magenta diamond is the ensemble mean, and the green circle is the verifying observation. Note that the kernel dressing method captures the error in the speed of the storm, spreading probability along the forecast track.