Quantification d'Incertitude dans l'Atmosphère et l'Océan

Olivier Talagrand
Laboratoire de Météorologie Dynamique/IPSL
École Normale Supérieure, Paris, France

avec des contributions de Guillem Candille, Laurent Descamps et Mohamed Jardak

FORUM CEA DAM

Les méthodes de quantification des incertitudes

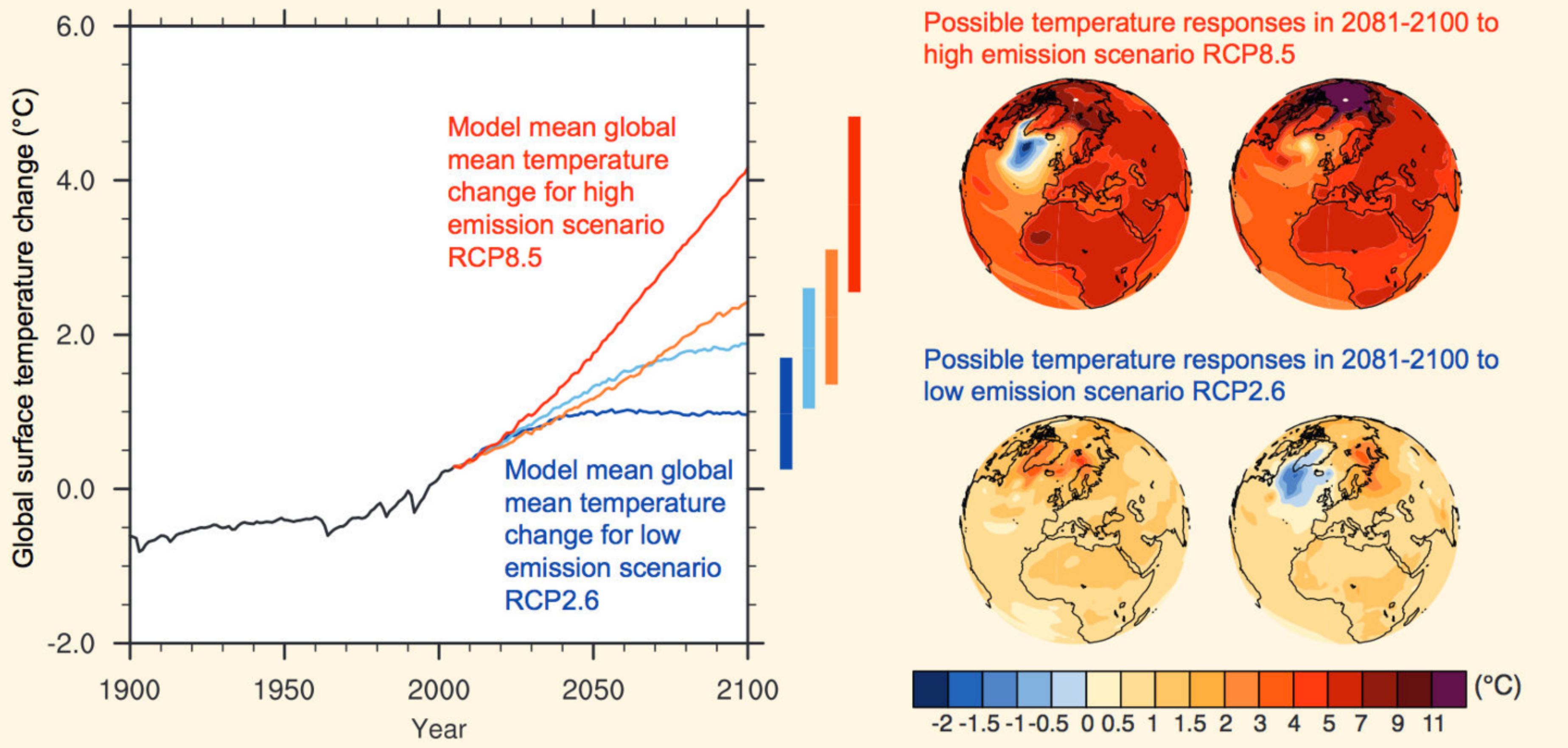
Très Grand Centre de Calcul, Bruyères-le-Châtel

1 Octobre 2014

Figure 6 Hurricane Katrina mean-sea-level-pressure (MSLP) analysis for 12 UTC of 29 August 2005 and t+84h high-resolution and EPS forecasts started at 00 UTC of 26 August:

1st row: 1st panel: MSLP analysis for 12 UTC of 29 Aug
2nd panel: MSLP t+84h T_L511L60 forecast started at 00 UTC of 26 Aug
3rd panel: MSLP t+84h EPS-control T_L255L40 forecast started at 00 UTC of 26 Aug
Other rows: 50 EPS-perturbed T_L255L40 forecast started at 00 UTC of 26 Aug.

The contour interval is 5 hPa, with shading patters for MSLP values lower than 990 hPa.



FAQ 12.1, Figure 1 | Global mean temperature change averaged across all Coupled Model Intercomparison Project Phase 5 (CMIP5) models (relative to 1986–2005) for the four Representative Concentration Pathway (RCP) scenarios: RCP2.6 (dark blue), RCP4.5 (light blue), RCP6.0 (orange) and RCP8.5 (red); 32, 42, 25 and 39 models were used respectively for these 4 scenarios. *Likely* ranges for global temperature change by the end of the 21st century are indicated by vertical bars. Note that these ranges apply to the difference between two 20-year means, 2081–2100 relative to 1986–2005, which accounts for the bars being centred at a smaller value than the end point of the annual trajectories. For the highest (RCP8.5) and lowest (RCP2.6) scenario, illustrative maps of surface temperature change at the end of the 21st century (2081–2100 relative to 1986–2005) are shown for two CMIP5 models. These models are chosen to show a rather broad range of response, but this particular set is not representative of any measure of model response uncertainty.

Fifth Assessment Report, Intergovernmental Panel on Climate Change, October 2013

- The case of Assimilation of Observations

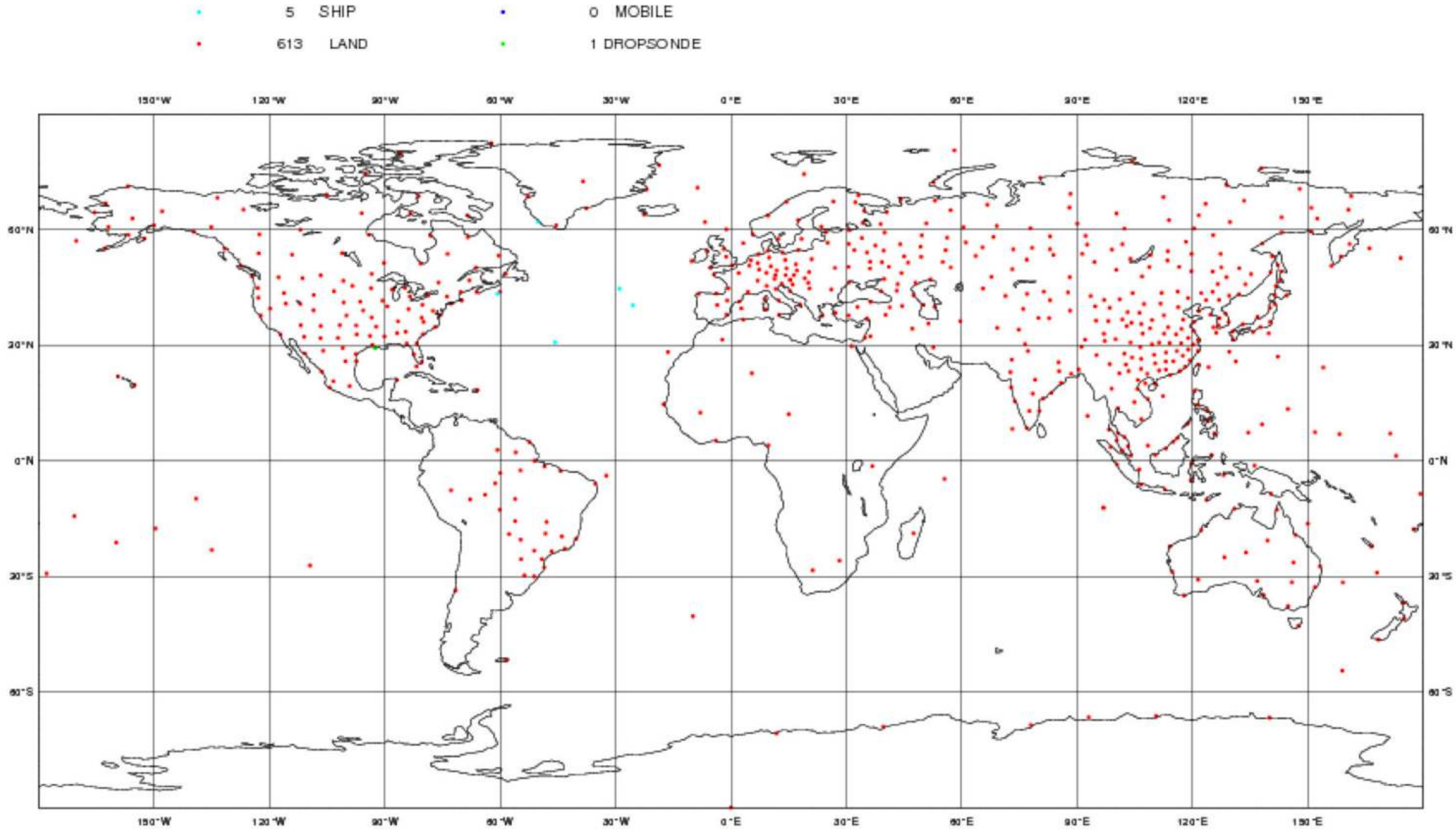
- The case of Prediction. Ensemble Prediction

Assimilation of Observations originated from the need of defining initial conditions for Numerical Weather Predictions

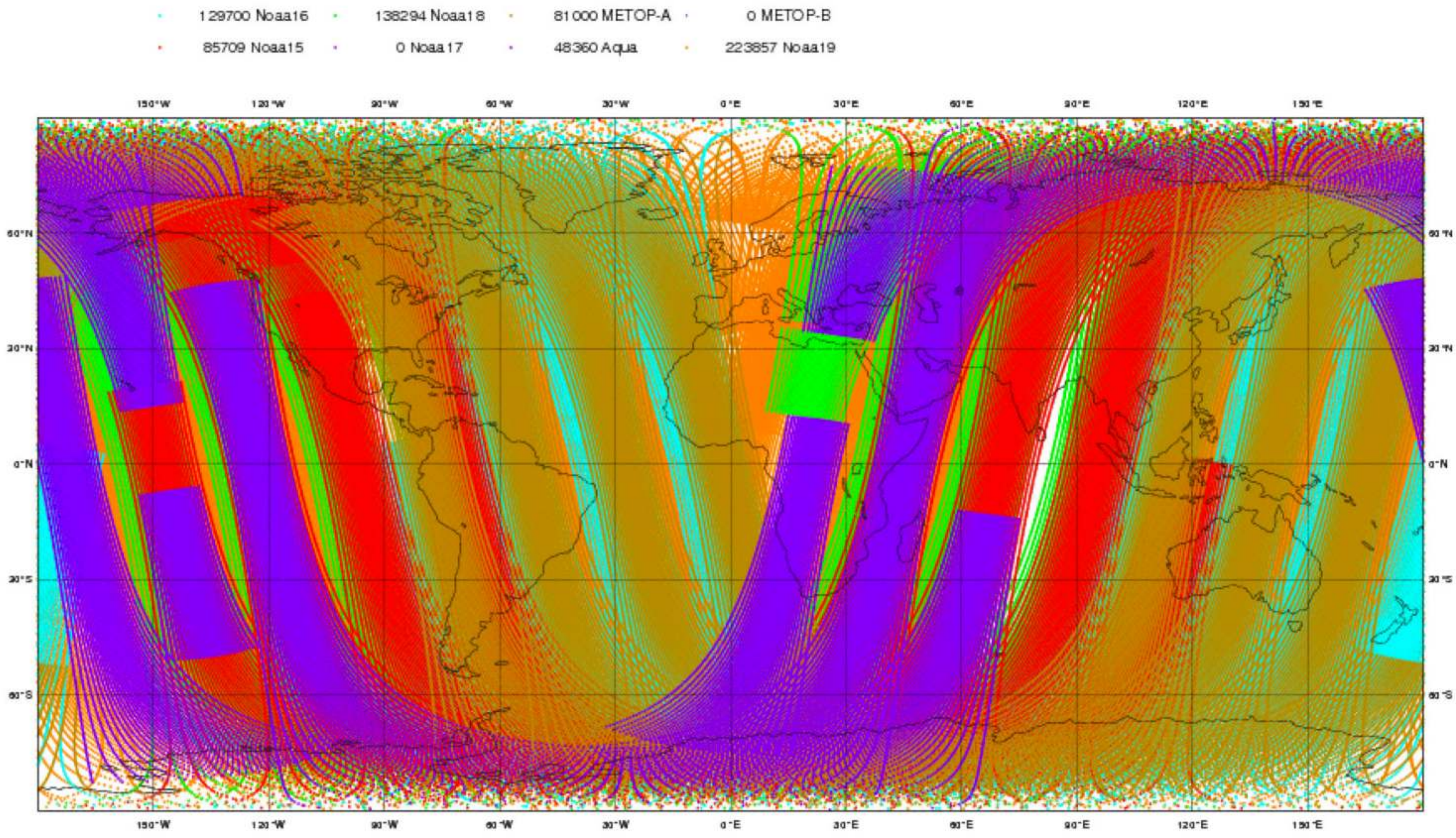
ECMWF Data Coverage (All obs DA) - Temp

16/Apr/2014; 00 UTC

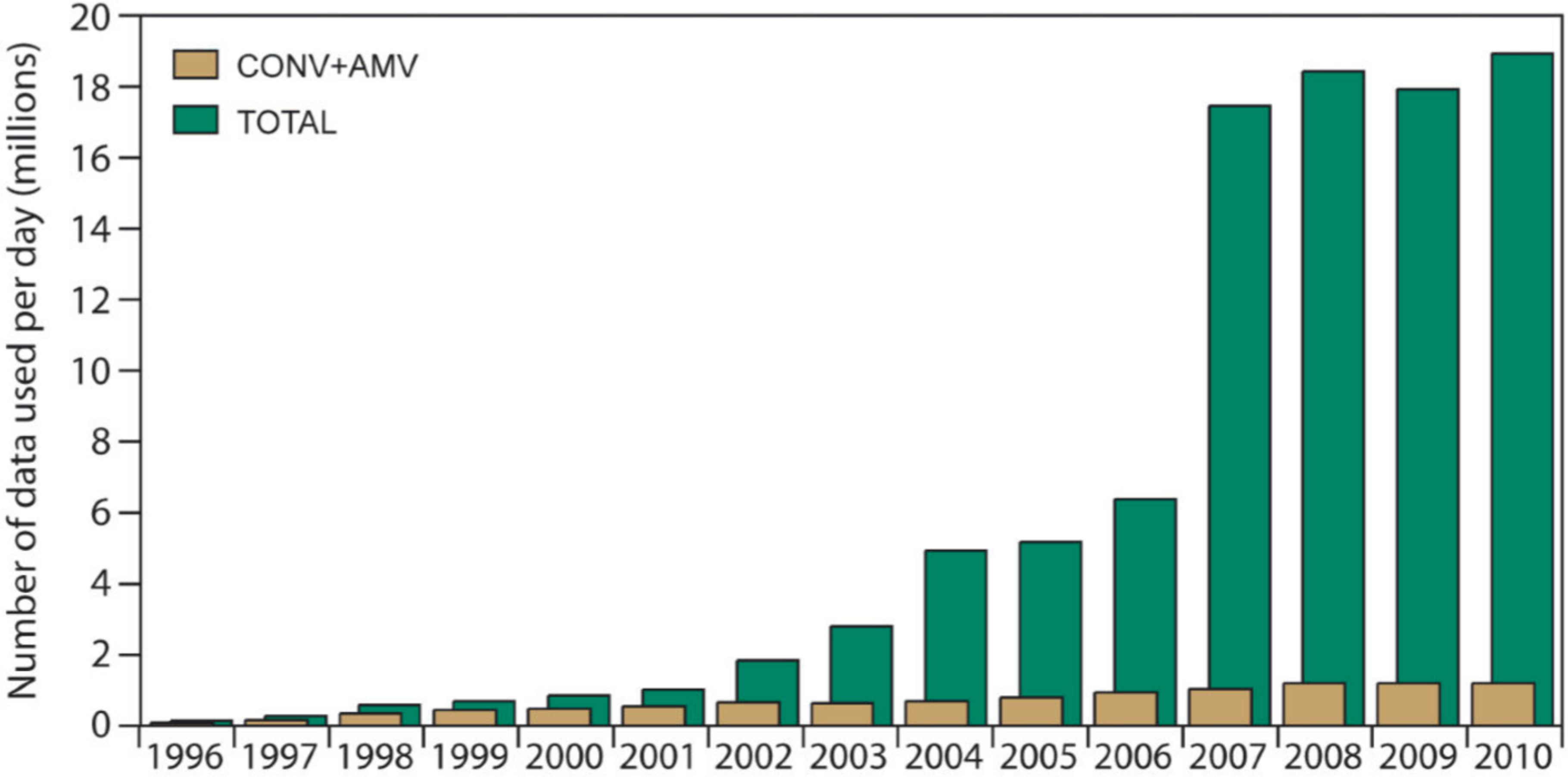
Total number of obs = 619



ECMWF Data Coverage (All obs DA) - AMSU-A
16/Apr/2014; 00 UTC
Total number of obs = 706920



ECMWF



Value as of early 2013 : around 25 millions per day

Physical laws governing the flow

- Conservation of mass

  $D\rho/Dt + \rho \, \mathrm{div}\underline{U} = 0$

- Conservation of energy

  $De/Dt - (p/\rho^2) \, D\rho/Dt = Q$

- Conservation of momentum

  $D\underline{U}/Dt + (1/\rho) \, \underline{\mathrm{grad}p} - \underline{g} + 2 \, \underline{\Omega} \wedge \underline{U} = \underline{F}$

- Equation of state

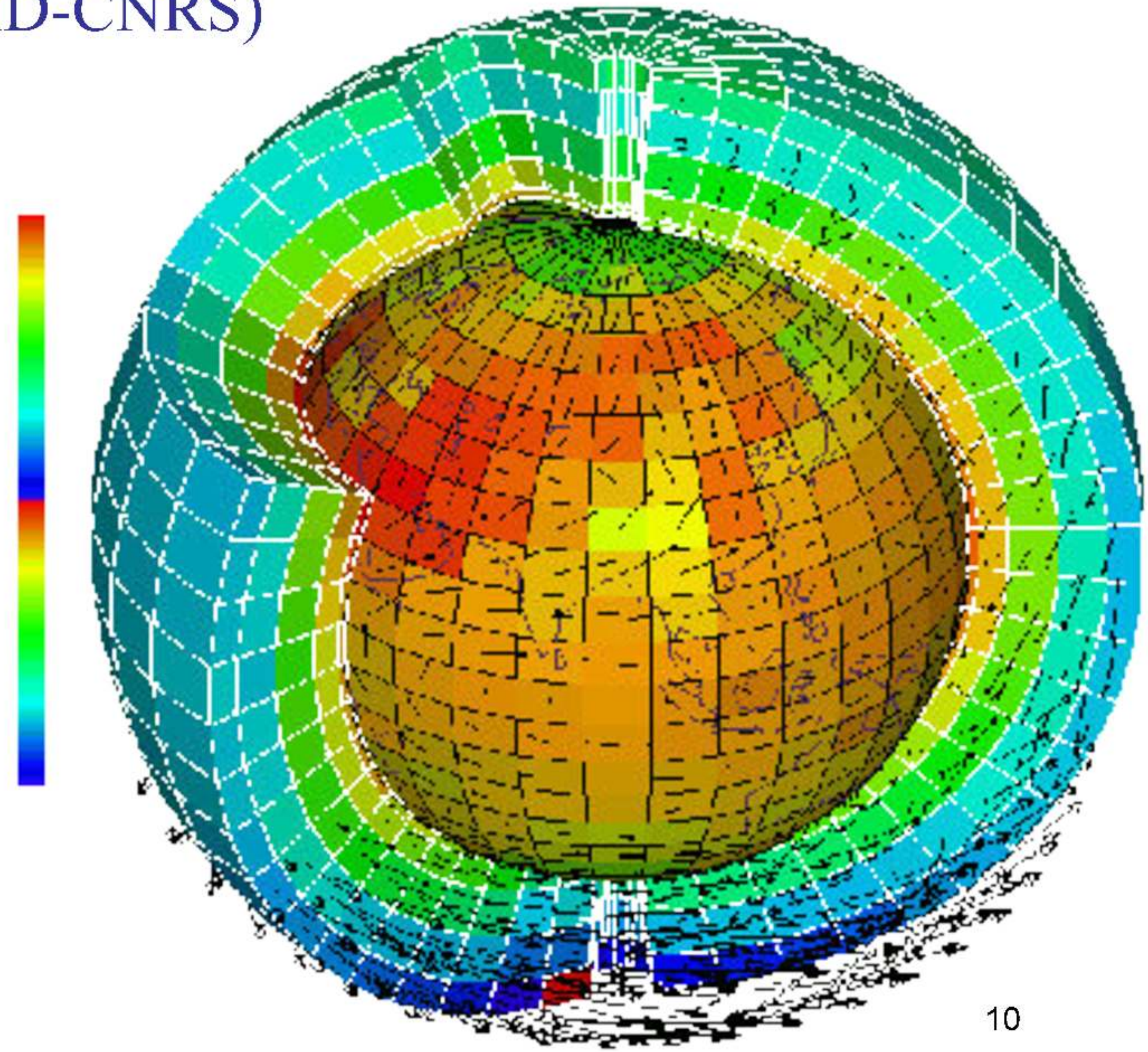
  $f(p, \rho, e) = 0$ $\quad (p/\rho = rT, \, e = C_v T)$

- Conservation of mass of secondary components (water in the atmosphere, salt in the ocean, chemical species, ...)

  $Dq/Dt + q \, \mathrm{div}\underline{U} = S$

Physical laws available in practice in the form of a discretized (and necessarily imperfect) numerical model

Schematic of of an atmospheric model (L. Fairhead /LMD-CNRS)



Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT, Reading, GB)

(European Centre for Medium-range Weather Forecasts, ECMWF)

En avril 2014 :

Troncature triangulaire T1279 (résolution horizontale ≈ 16 kilomètres)

137 niveaux dans la direction verticale (0 - 80 km)

Dimension du vecteur d'état correspondant ≈ 2,3 10^9

Pas de discrétisation temporelle : 10 minutes

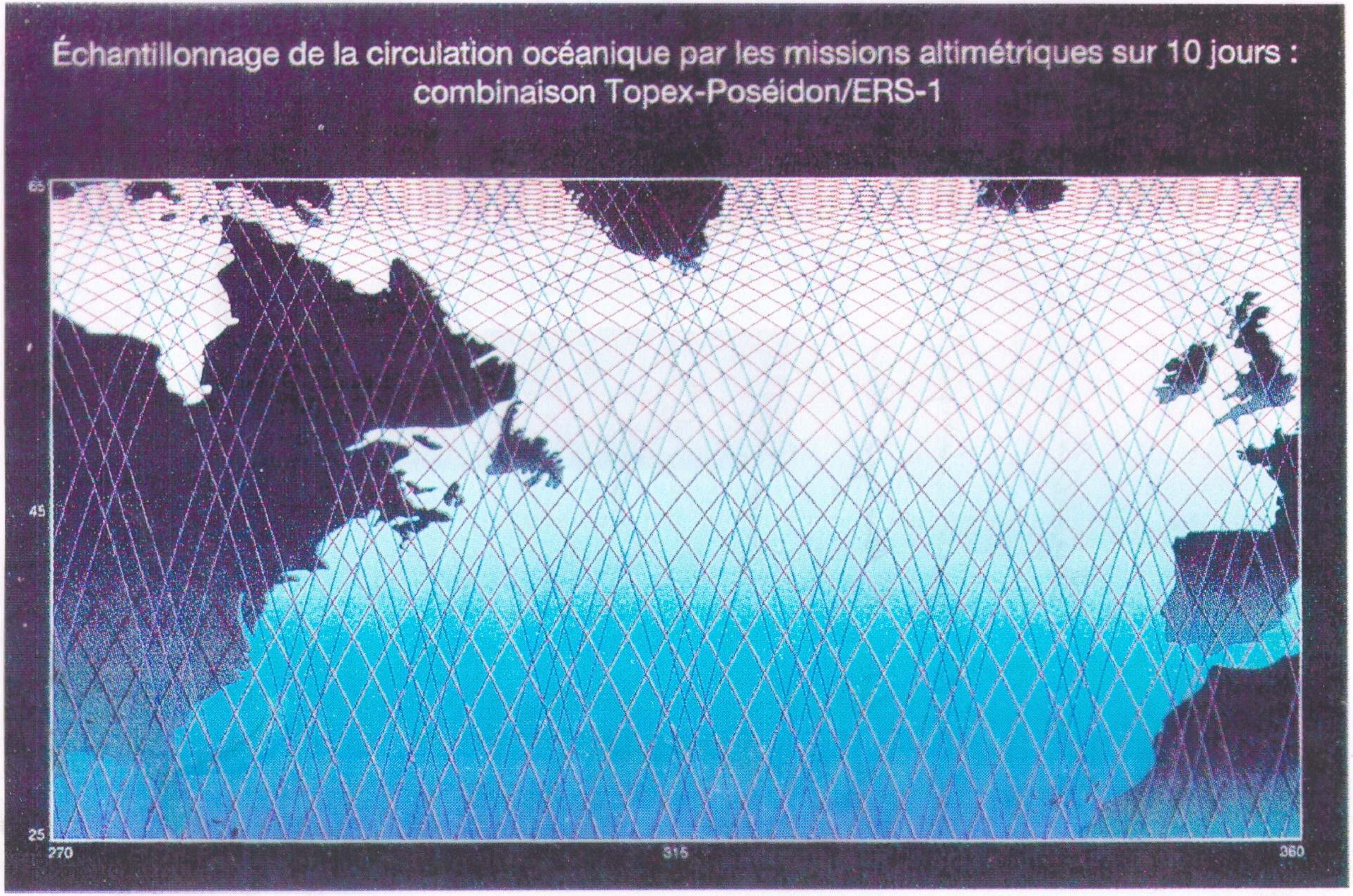
Assimilation of Observations

Purpose of assimilation : reconstruct as accurately as possible the state of the atmospheric or oceanic flow, using all available appropriate information. The latter essentially consists of

- The observations proper, which vary in nature, resolution and accuracy, and are distributed more or less regularly in space and time.

- The physical laws governing the evolution of the flow, available in practice in the form of a discretized, and necessarily approximate, numerical model.

- ‘Asymptotic’ properties of the flow, such as, *e. g.*, geostrophic balance of middle latitudes. Although they basically are necessary consequences of the physical laws which govern the flow, these properties can usefully be explicitly introduced in the assimilation process.



S. Louvel, Doctoral Dissertation, 1999

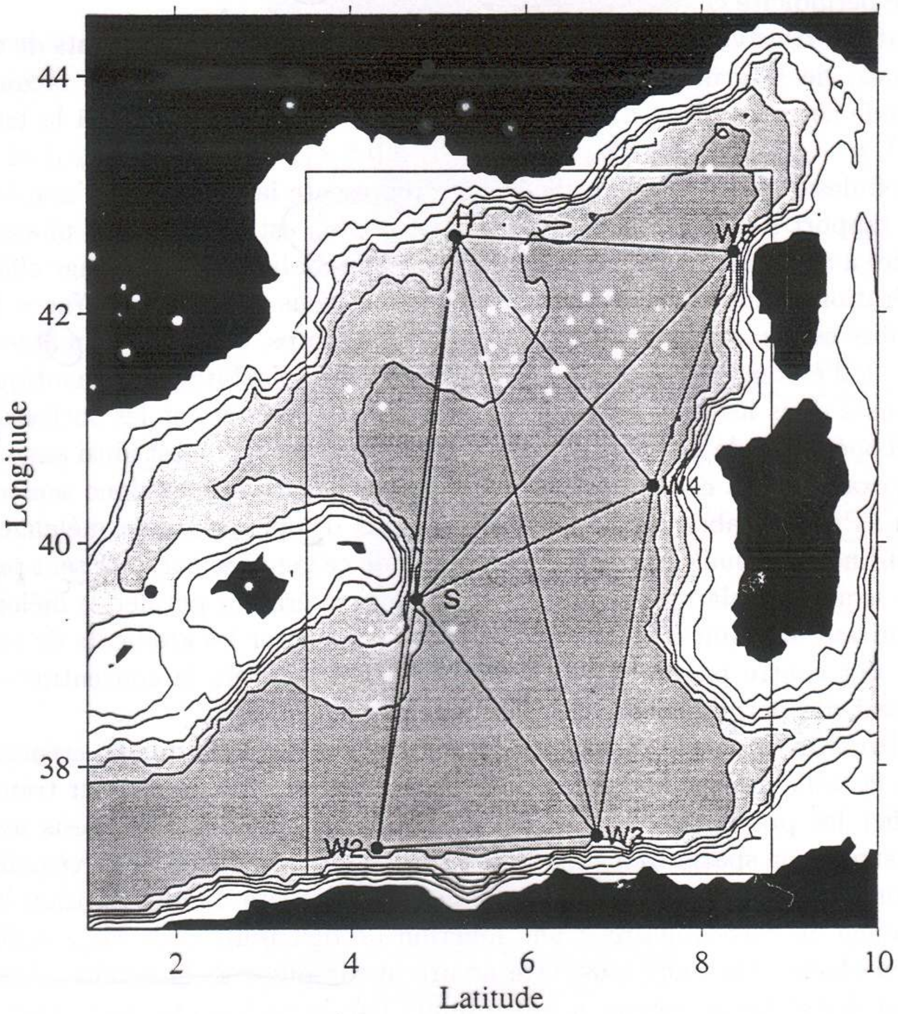


FIG. 1 – *Bassin méditerranéen occidental: réseau d'observation tomographique de l'expérience Thétis 2 et limites du domaine spatial utilisé pour les expériences numériques d'assimilation.*

E. Rémy, Doctoral Dissertation, 1999

Assimilation is one of many '*inverse problems*' encountered in many fields of science and technology

- solid Earth geophysics
- plasma physics
- 'nondestructive' probing
- navigation (spacecraft, aircraft, ….)
- …

Solution most often (if not always) based on Bayesian, or probabilistic, estimation. 'Equations' are fundamentally the same.

Difficulties specific to assimilation of meteorological observations :

- Very large numerical dimensions ($n \approx 10^6$-10^9 parameters to be estimated, $p \approx$ 1-3.10^7 observations per 24-hour period). Difficulty aggravated in Numerical Weather Prediction by the need for the forecast to be ready in time.

- Non-trivial, actually chaotic, underlying dynamics

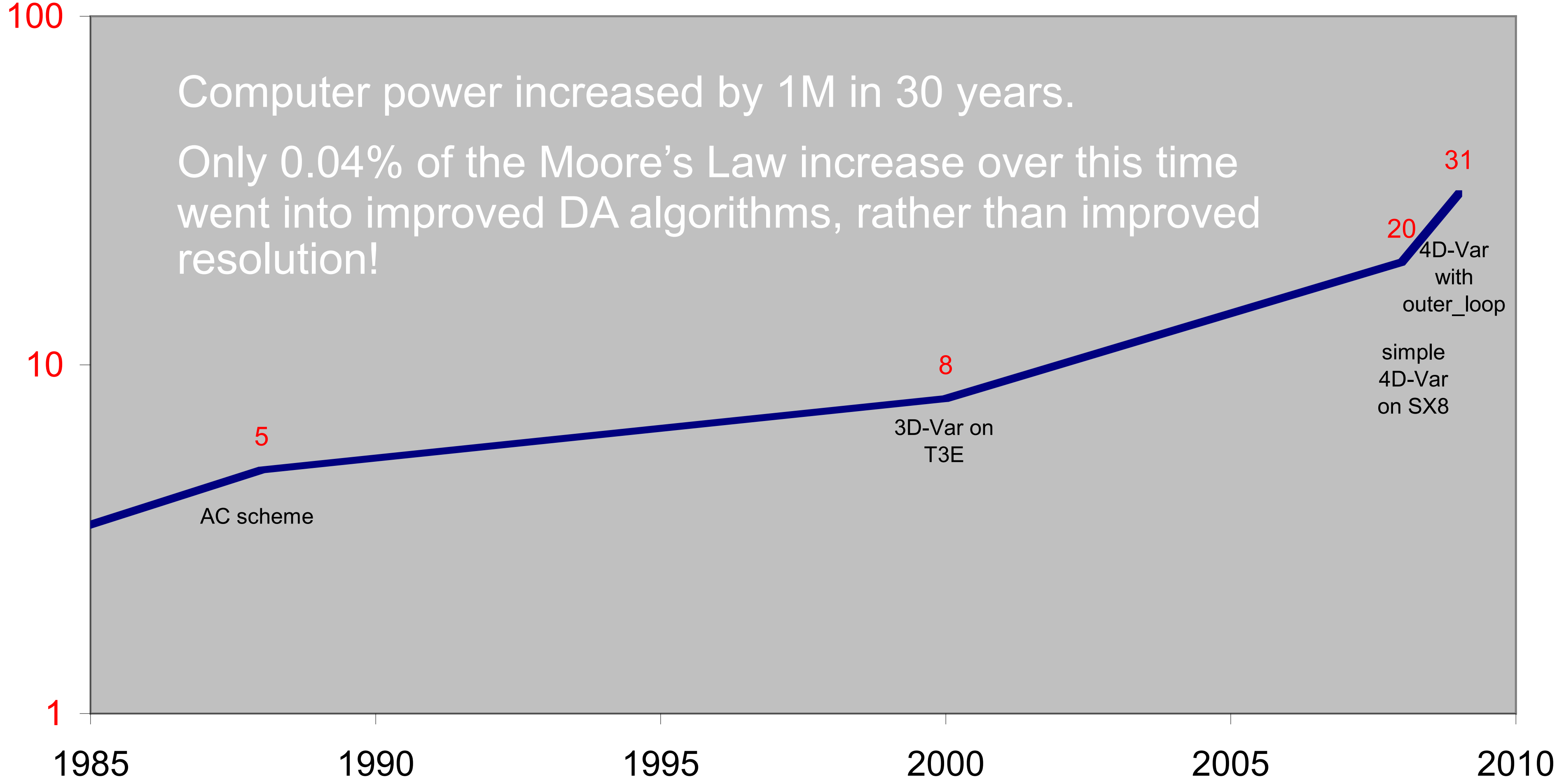
Both observations and 'model' are affected with some uncertainty $\Rightarrow$ uncertainty on the estimate.

For some reason, uncertainty is conveniently described by probability distributions (don't know too well why, but it works; see, *e.g.* Jaynes, 2007, *Probability Theory: The Logic of Science,* Cambridge University Press).

Assimilation is a problem in bayesian estimation.

Determine the conditional probability distribution for the state of the system, knowing everything we know (see Tarantola, A., 2005*, Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM).

ratio of supercomputer costs: 1 day's assimilation / 1 day forecast



Courtesy A. Lorenc

Bayesian Estimation

Data of the form

$$z = \Gamma x + \zeta, \qquad \zeta \sim \mathcal{N}[\mu, S]$$

Known data vector z belongs to *data space* $\mathcal{D}$, $dim\mathcal{D} = m$,
Unknown state vector x belongs to *state space* $\mathcal{X}$, $dim\mathcal{X} = n$
Γ known (mxn)-matrix, ζ unknown 'error'

Then conditional probability distribution is

$$P(x \mid z) = \mathcal{N}[x^a, P^a]$$

where

$$x^a = (\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T S^{-1} [z - \mu]$$
$$P^a = (\Gamma^T S^{-1} \Gamma)^{-1}$$

Determinacy condition : rank$\Gamma = n$. Requires $m \geq n$.

Variational form.

Conditional expectation x^a minimizes following scalar *objective function*, defined on state space $\mathcal{X}$

$$\xi \in \mathcal{X} \rightarrow \mathcal{J}(\xi) \equiv (1/2)\,[\Gamma\xi - (z-\mu)]^{\mathrm{T}}\, S^{-1}\, [\Gamma\xi - (z-\mu)]$$

$$P^a = [\partial^2 \mathcal{J} / \partial \xi^2]^{-1}$$

If data still of the form

$$z = \Gamma x + \zeta,$$

but 'error' ζ , which still has expectation μ and covariance S, is not Gaussian, expressions

$$x^a = (\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T S^{-1} [z - \mu]$$
$$P^a = (\Gamma^T S^{-1} \Gamma)^{-1}$$

do not achieve Bayesian estimation, but define least-variance linear estimate of x from z (*Best Linear Unbiased Estimator, BLUE*), and associated estimation error covariance matrix.

Sequential Assimilation

- Observation vector at time k

$$y_k = H_k x_k + \varepsilon_k \qquad k = 0, \dots, K$$

$$E(\varepsilon_k) = 0 \;\; ; \;\; E(\varepsilon_k \varepsilon_j^{\mathrm{T}}) = R_k \, \delta_{kj}$$

H_k linear

- Evolution equation

$$x_{k+1} = M_k x_k + \eta_k \qquad k = 0, \dots, K\text{-}1$$

$$E(\eta_k) = 0 \;\; ; \;\; E(\eta_k \eta_j^{\mathrm{T}}) = Q_k \, \delta_{kj}$$

M_k linear

- $E(\eta_k \varepsilon_j^{\mathrm{T}}) = 0$ (errors uncorrelated in time)

At time k, background x^b_k and associated error covariance matrix P^b_k known

- Analysis step

$$x^a_k = x^b_k + P^b_k H_k^{\mathrm{T}} [H_k P^b_k H_k^{\mathrm{T}} + R_k]^{-1} (y_k - H_k x^b_k)$$
$$P^a_k = P^b_k - P^b_k H_k^{\mathrm{T}} [H_k P^b_k H_k^{\mathrm{T}} + R_k]^{-1} H_k P^b_k$$

- Forecast step

$$x^b_{k+1} = M_k x^a_k$$
$$P^b_{k+1} = M_k P^a_k M_k^{\mathrm{T}} + Q_k$$

Kalman filter (KF, Kalman, 1960)

Must be started from some initial estimate (x^b_0, P^b_0)

Variational form (exactly equivalent to Kalman filter in linear case)

$(\xi_0, \xi_1, \ldots, \xi_K) \rightarrow$

$\mathcal{J}(\xi_0, \xi_1, \ldots, \xi_K)$

$$= (1/2)\,(x_0^b - \xi_0)^T\,[P_0^b]^{-1}\,(x_0^b - \xi_0)$$

$$+ (1/2)\,\Sigma_{k=0,\ldots,K}[y_k - H_k\xi_k]^T\,R_k^{-1}\,[y_k - H_k\xi_k]$$

$$+ (1/2)\,\Sigma_{k=0,\ldots,K-1}[\xi_{k+1} - M_k\xi_k]^T\,Q_k^{-1}\,[\xi_{k+1} - M_k\xi_k]$$

***Four-Dimensional Variational Assimilation* (*4D-Var*)**

Commonly used in *strong constraint* form (Q_k=0)

Can include nonlinear M_k and/or H_k.

Heuristic extensions to (not too strongly) nonlinear cases :

- ***Ensemble Kalman Filter*** (***EnKF***, *Evensen and others*)

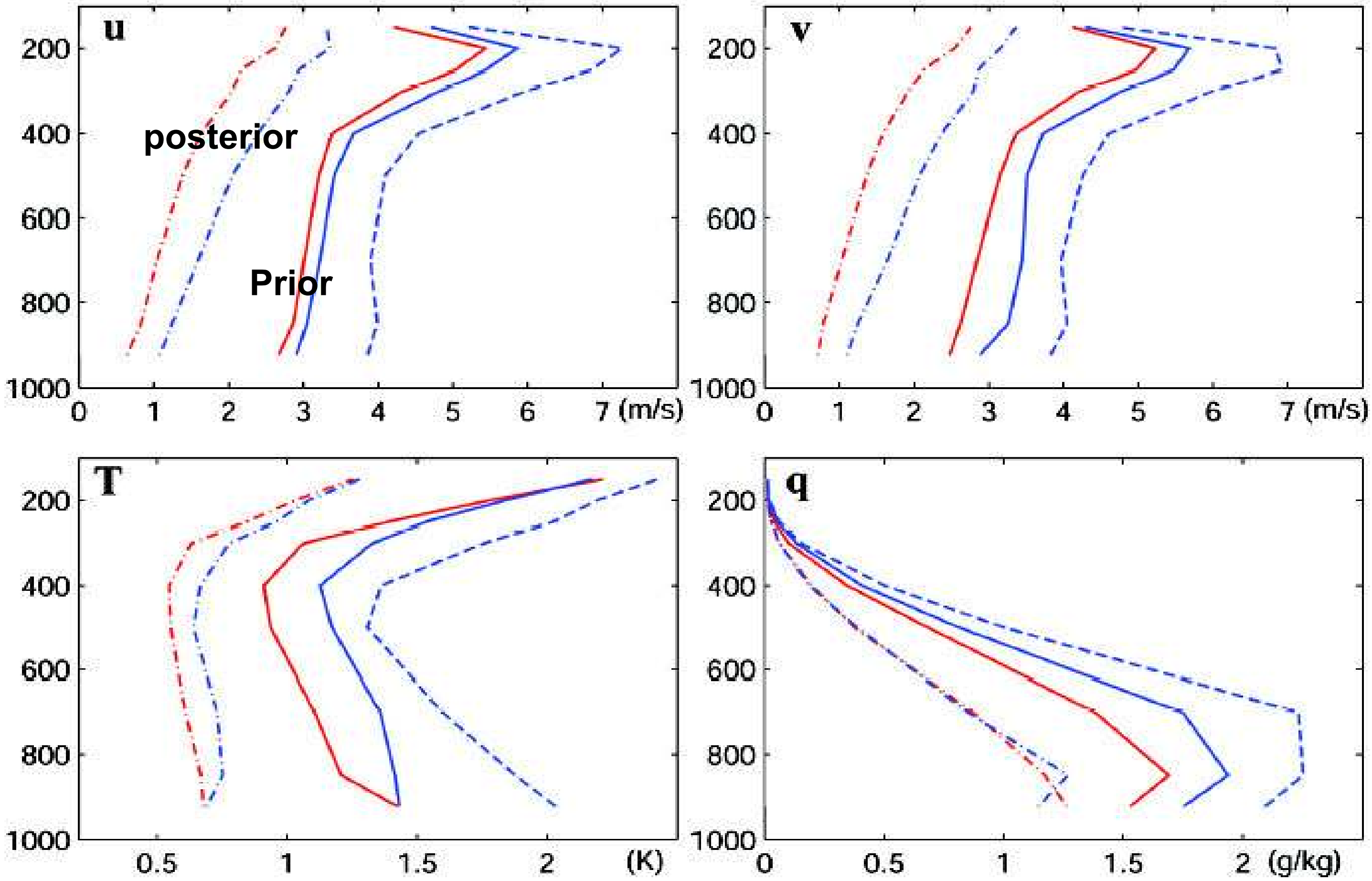
  Uncertainty is represented, not by a covariance matrix, but by an ensemble of point estimates in state space that are meant to sample the conditional probability distribution for the state of the system (dimension $N \approx O(10\text{-}100)$).

  Forecast step. Ensemble is evolved in time through the full model (eliminates any need for linear hypothesis as to the temporal evolution).

  Analysis step. Uses formulæ for the *BLUE*

Month-long Performance of EnKF vs. 3Dvar with WRF

— EnKF — 3DVar (prior, solid; posterior, dotted)



Better performance of EnKF than 3DVar also seen in both 12-h forecast and posterior analysis in terms of root-mean square difference averaged over the entire month

(Meng and Zhang 2007c, MWR, in review)

Many variants exist for Ensemble Kalman Filter, which has become, together with Variational Assimilation, one of the two powerful classes of algorithms used in numerical modeling of the atmospheric and oceanic flow.

Dimension $N \approx O(10\text{-}100)$. Stability of filter requires N to be at least as large as number of unstable directions in the system (Hoang, Trevisan and Palatella).

Bayesian properties of Ensemble Kalman Filter ?

Very little is known.

Le Gland *et al*. (2011). In the linear and gaussian case, the discrete pdf defined by the filter, in the limit of infinite sample size N, tends to the bayesian gaussian pdf.

No result for finite size (note that ensemble elements are not mutually independent)

In the nonlinear case, the discrete pdf tends to a limit which is in general not the bayesian pdf.

Exact bayesian estimation ?

Particle filters

Predicted ensemble at time t : $\{x^b{}_n, n = 1, \ldots, N\}$, each element with its own weight (probability) $P(x^b{}_n)$

Observation vector at same time : $y = Hx + \varepsilon$

Bayes' formula

$$P(x^b{}_n|y) \sim P(y|x^b{}_n)\, P(x^b{}_n)$$

Defines updating of weights

Bayes' formula

$$P(x^b{}_n|y) \sim P(y|x^b{}_n)\ P(x^b{}_n)$$

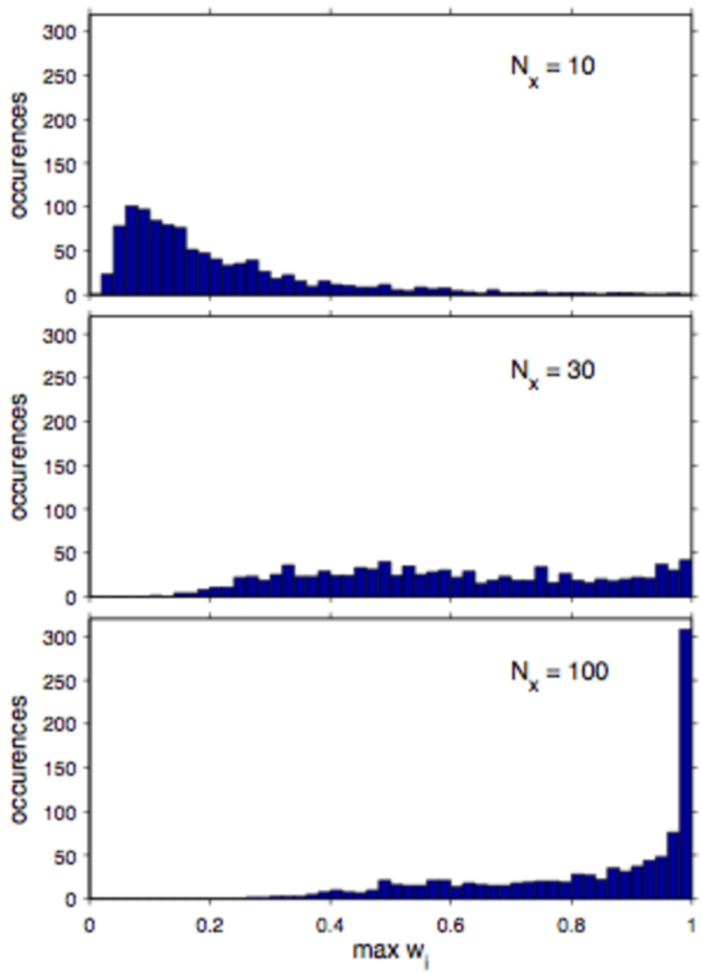
Defines updating of weights; particles are not modified. Asymptotically converges to bayesian pdf. Very easy to implement.

Observed fact. For large state dimension, ensemble tends to collapse.

Problem originates in the 'curse of dimensionality' Large dimension pdf's are very diffuse, so that very few particles (if any) are present in areas where conditional probability ('*likelihood*') $P(y|x)$ is large.

Behavior of $\max w^i$

- $N_e = 10^3$; $N_x = 10, 30, 100$; 10^3 realizations



average squared error of posterior mean = 5.5

... = 25

... = 127

C. Snyder, http://www.cawcr.gov.au/staff/pxs/wmoda5/Oral/Snyder.pdf

Alternative possibilities (review in van Leeuwen, 2009, *Mon. Wea. Rev.*, 4089-4114)

Importance Sampling.

Use a *proposal density* that is closer to the new observations than the density defined by the predicted particles (for instance the density defined by EnKF, after the latter has used the new observations). Independence between observations is then lost in the computation of likelihood $P(y|x)$ (or is it not ?)

In particular, *Guided Sequential Importance Sampling* (van Leeuwen, 2002). Idea : use observations performed at time k to resample ensemble at some timestep anterior to k, or 'nudge' integration between times k-1 and k towards observation at time k.

Particle filters are actively studied

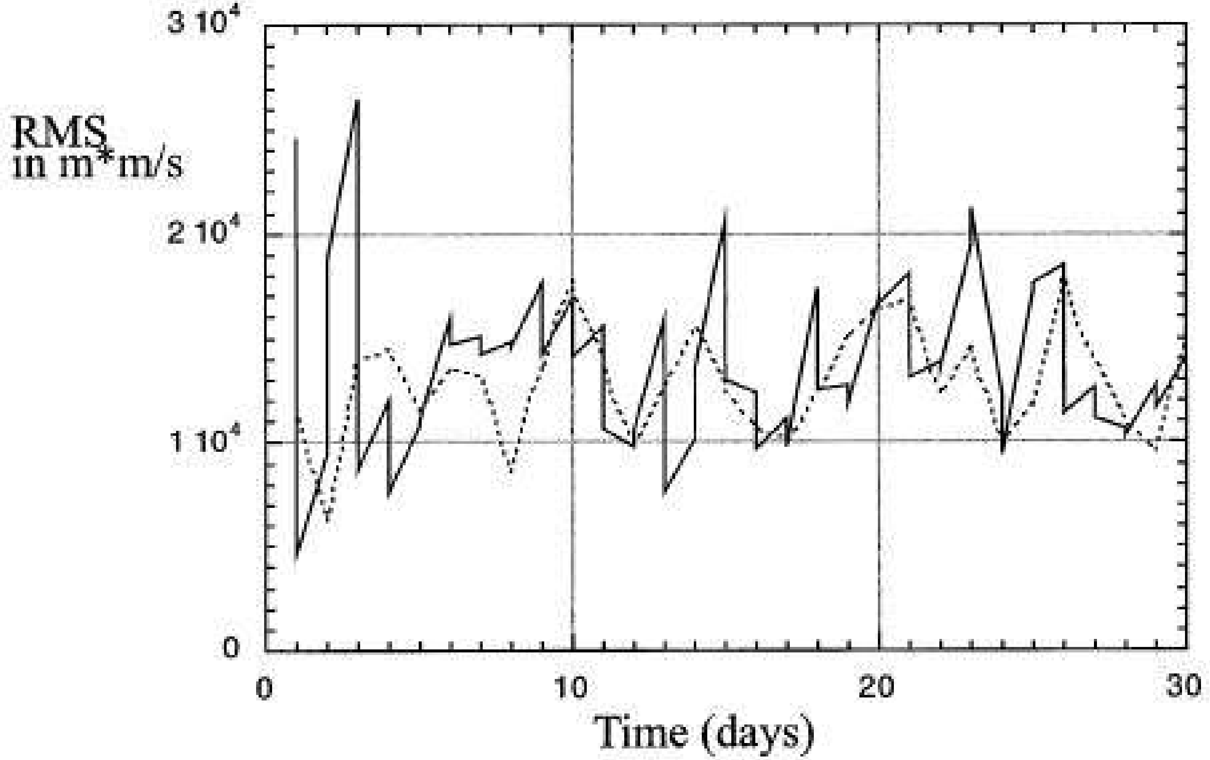


FIG. 12. Comparison of rms error ($m^2\ s^{-1}$) between ensemble mean and independent observations (dotted line) and the std dev in the ensemble (solid line). The excellent agreement shows that the SIRF is working correctly.

van Leeuwen, 2003, *Mon. Wea. Rev.*, **131**, 2071-2084

Ensemble Variational Assimilation (work done with Mohamed Jardak)

Data of the form

$$z = \Gamma x + \zeta, \qquad \zeta \sim \mathcal{N}[\mu, S]$$

Conditional probability distribution is

$$P(x \mid z) = \mathcal{N}[x^a, P^a]$$

with

$$x^a = (\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T S^{-1} [z - \mu]$$

$$P^a = (\Gamma^T S^{-1} \Gamma)^{-1}$$

Ready recipe for determining Monte-Carlo sample of conditional pdf $P(x \mid z)$:

- Perturb data vector z according to its own error probability distribution

$$z \rightarrow z' = z + \delta, \qquad \delta \sim \mathcal{N}[0, S]$$

and compute

$$x'^{a} = (\Gamma^{\mathrm{T}} S^{-1} \Gamma)^{-1} \Gamma^{\mathrm{T}} S^{-1} [z' - \mu]$$

x'^{a} is distributed according to $\mathcal{N}[x^{a}, P^{a}]$

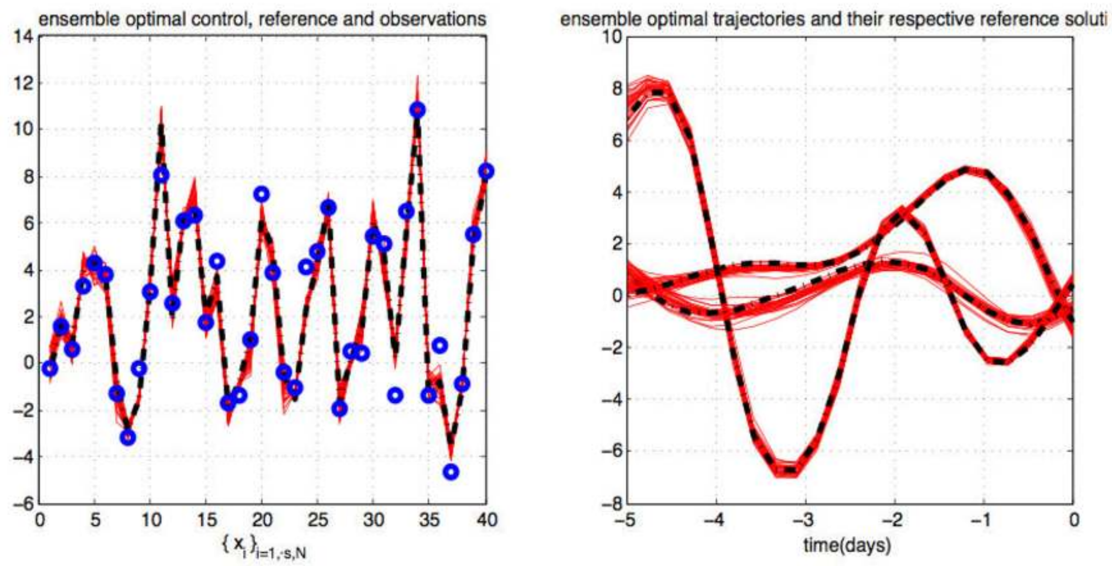
Ensemble Variational Assimilation (*EnsVar*) implements that algorithm, the expectations $x^{\prime a}$ being computed by standard variational assimilation (optimization)

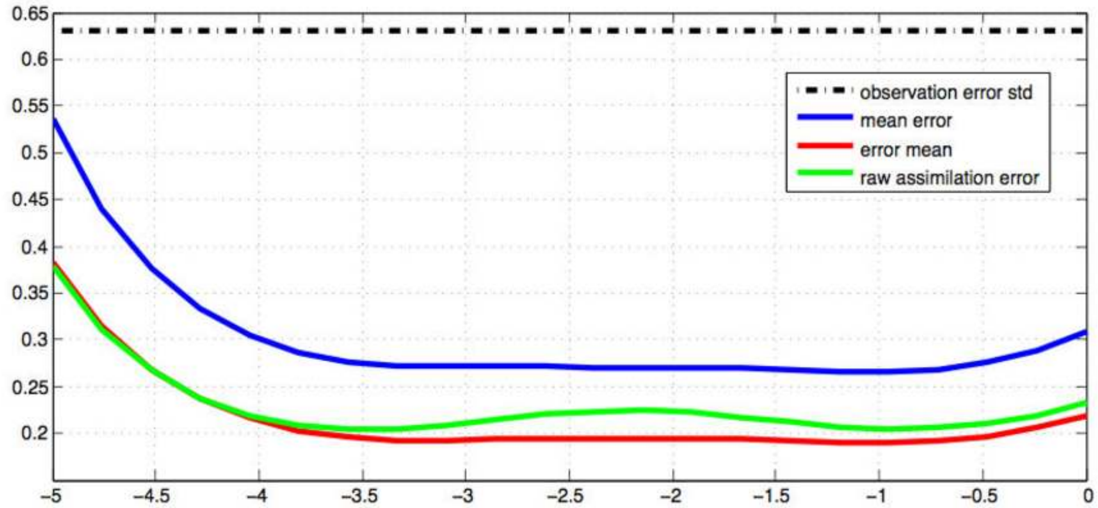
The Lorenz96 model

- Forward model

$$\frac{dx_k}{dt} = (x_{k+1} - x_{k-2})x_{k-1} - x_k + F \quad \text{for} \quad k = 1, \cdots, N$$

- Set-up parameters :
  1. the index k is cyclic so that $x_{k-N} = x_{k+N} = x_k$.
  2. $F = 8$, external driving force.
  3. $-x_k$, a damping term.
  4. $N = 40$, the system size.
  5. $Nens = 30$, number of ensemble members.
  6. $\frac{1}{\lambda_{max}} \simeq 2.5days, \lambda_{max}$ the largest Lyapunov exponent.
  7. $\Delta t = 0.05 = 6hours$, the time step.
  8. frequency of observations : every 12 hours.
  9. number of realizations : 9000 realizations.





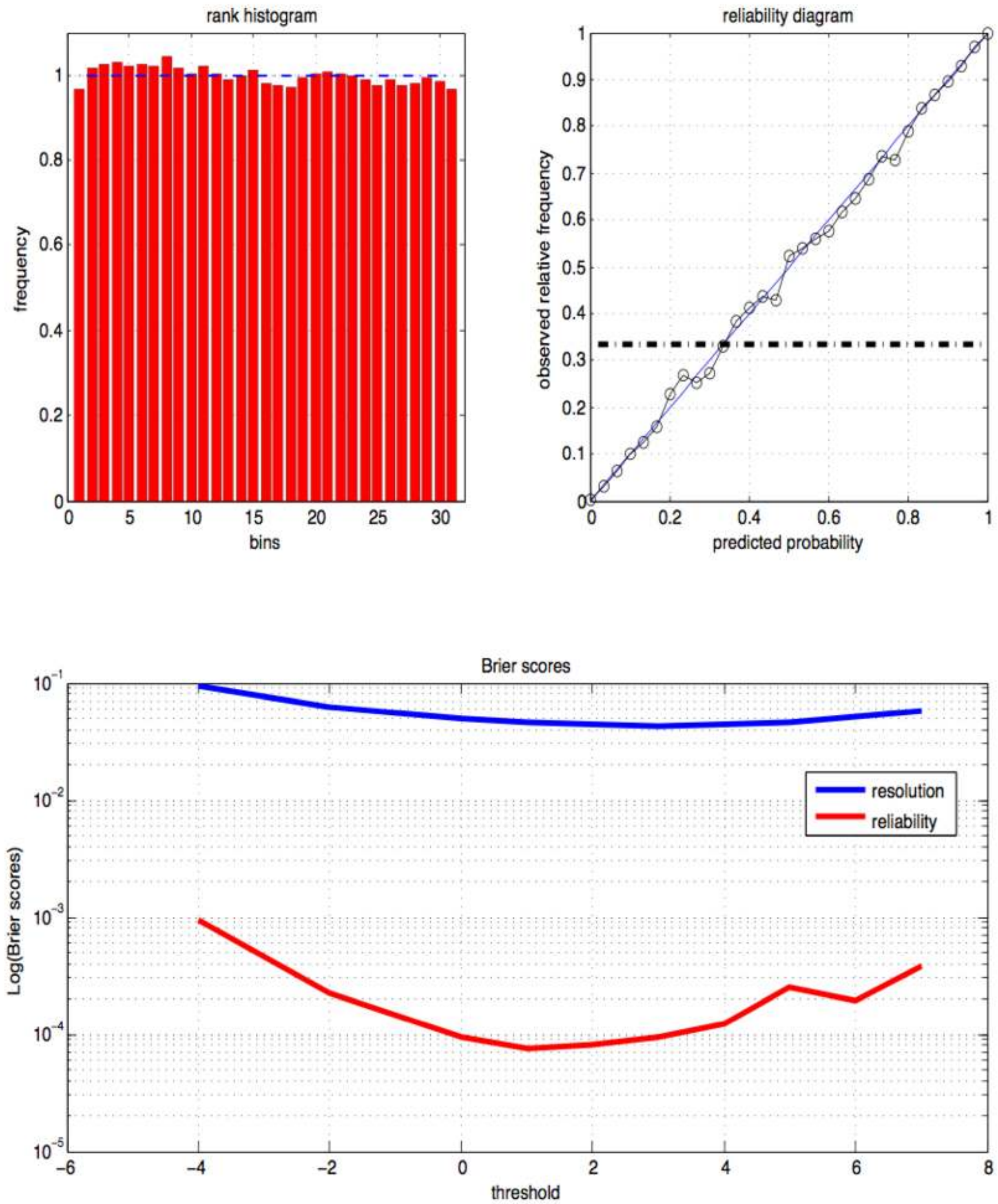
Nonlinear Lorenz'96. 5 days

Evaluation ?

There is no general objective criterion for Bayesianity, and we use instead the weaker property of *reliability*.

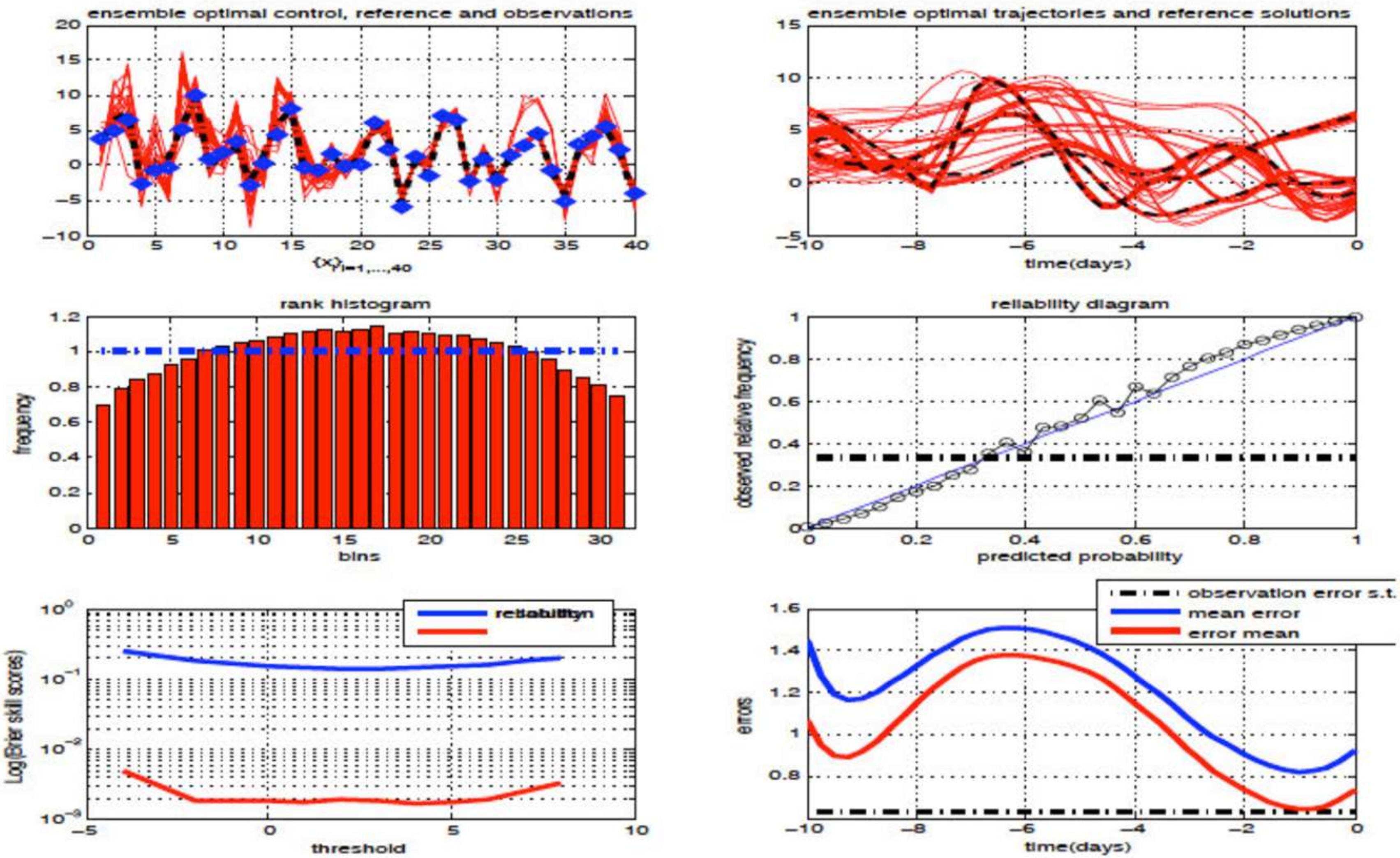
Reliability is statistical consistency between predicted probabilities of occurrence, and observed frequencies of occurrence (it rains with frequency 40% in the circumstances when I have predicted 40% probability for the occurrence of rain). More generally, for any probability distribution F, observed reality is distributed with frequency distribution F in the circumstances when I have predicted F.

Bayesianity implies reliability, the converse not being true.

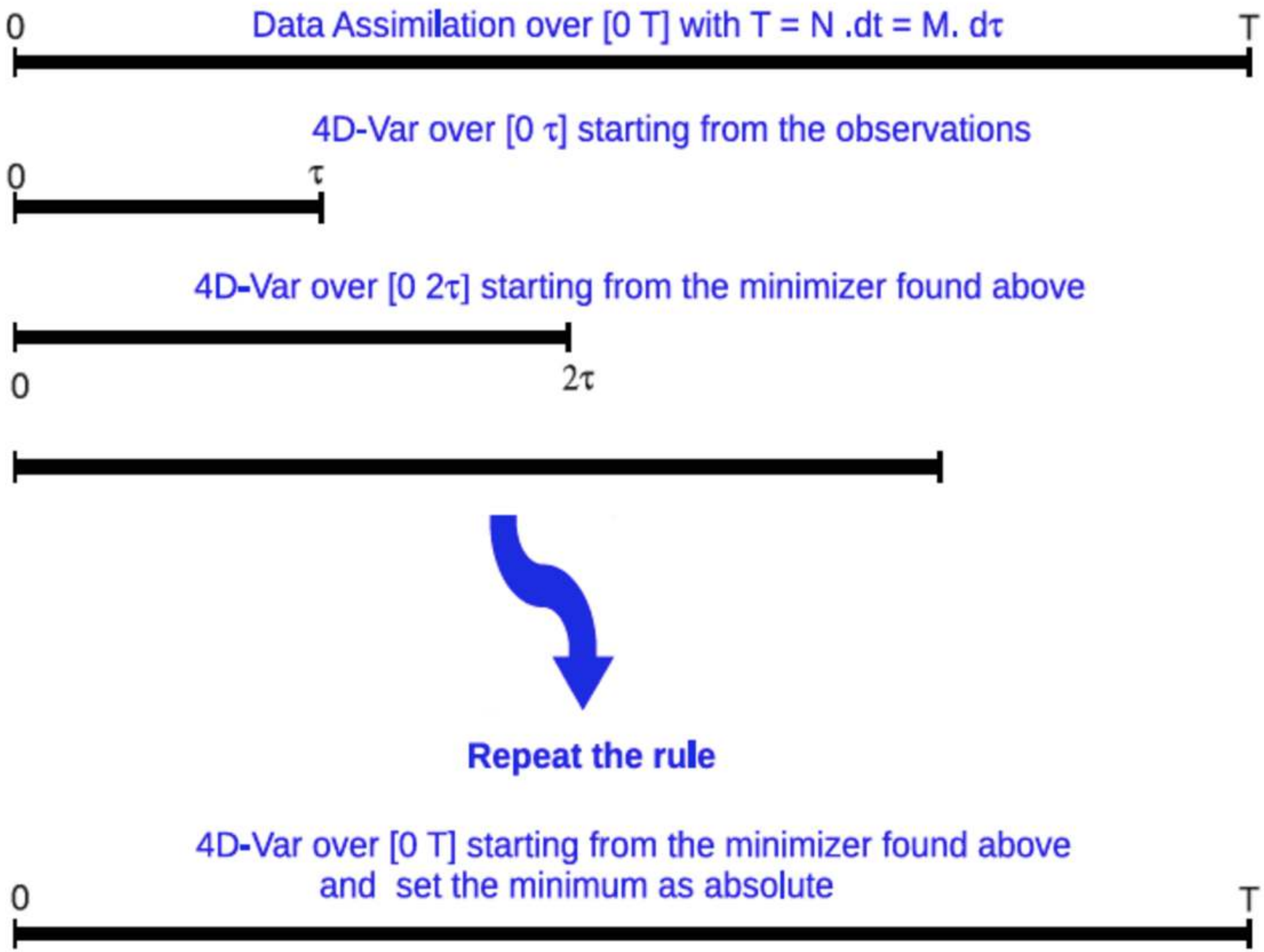


Nonlinear Lorenz’96. 5 days

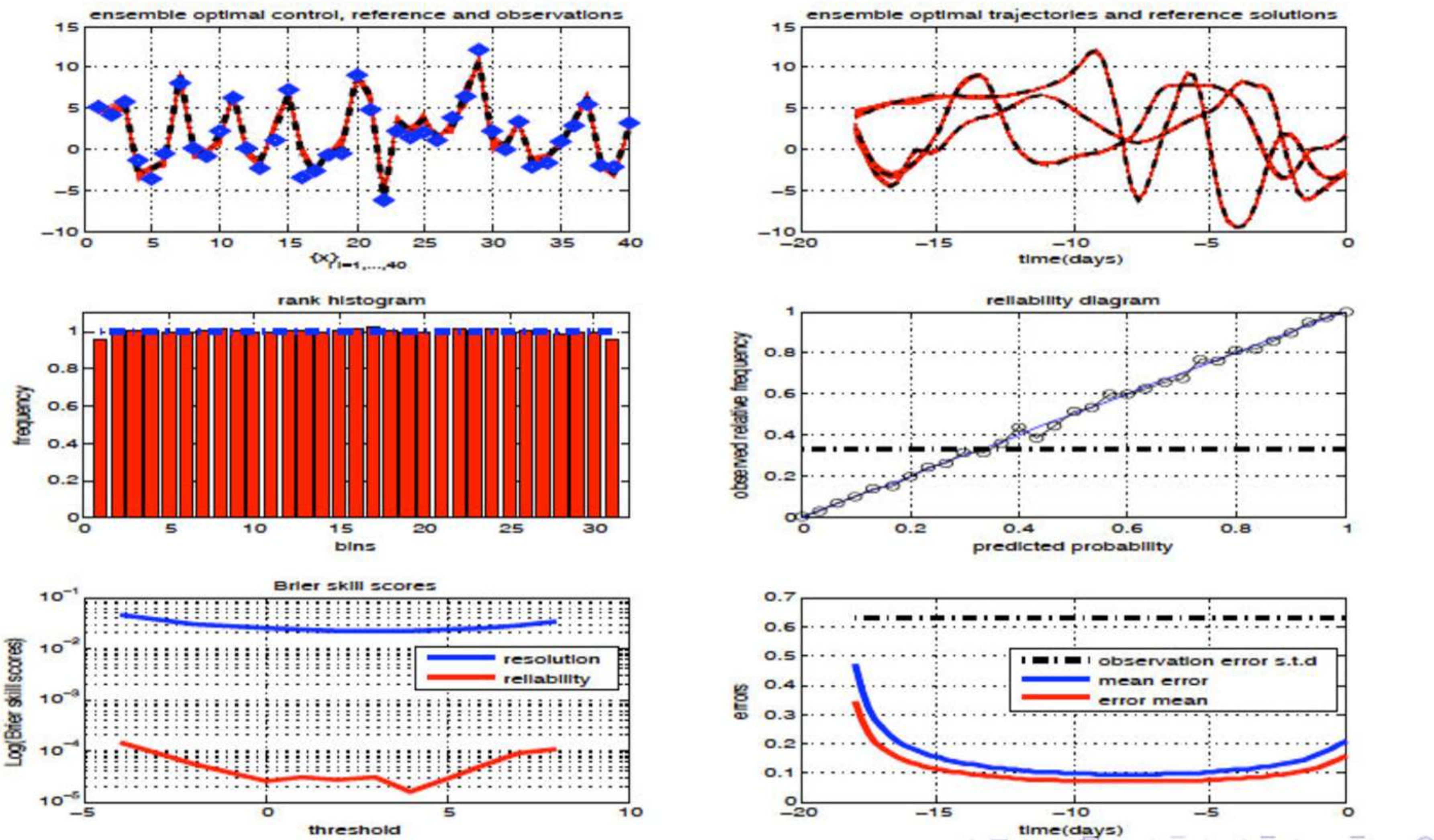
EnsVar : the non-linear Lorenz96 model (10 days ≃ 2 TU)



Quasi-Static Variational Assimilation (QSVA)



EnsVar : the non-linear Lorenz96 model 18 days with QSVA



- Results are independent of the Gaussian character of the observation errors (trials have been made with various probability distributions)

- Ensembles produced by EnsVar are very close to Gaussian, even in strongly nonlinear cases.

- Comparison *Ensemble Kalman Filter* (*EnKF*) and *Particle Filters* (*PF*)

  Both of these algorithms being sequential, comparison is fair only at end of assimilation window

  Results produced by ENSVar are at least as good as results produced by EnKF or PF

DA procedure / *method*	*Assimilation*	Forecasting
EnsVAR	0.2193510	1.49403506
EnKF	0.2449690	1.67176110
PF	0.7579790	2.62461295

RMS errors in the ensemble means at the end of 5-day assimilations and 5-day forecasts

Weak constraint EnsVar

- define the objective function.

$$\mathfrak{J}(x,\eta_1,\eta_2,\cdots.\eta_{N-1},\eta_N) = \frac{1}{2}\left\{(x-x_b)^T \mathrm{B}^{-1}(x-x_b)\right\} \quad +$$

$$\frac{1}{2}\sum_{i=0}^{N}\left\{(y_i - H_i(x_i))^T \mathrm{R}_{\mathrm{i}}^{-1}(y_i - H_i(x_i))\right\} + \frac{1}{2}\sum_{i=1}^{N}\eta_i^T \mathrm{Q}_i^{-1}\eta_i$$

1. B background error covariance matrix and R observation error covariance matrix.
2. Q model error covariance matrix.
3. $H : \mathbb{R}^{state} \longrightarrow \mathbb{R}^{obs}$ observation operator.
4. x_b background state vector and y_i observation vector at time $t = t_i$.
5. η_i model error vector at $t = t_i$ with $x(t_i) = \mathfrak{M}_{t_i \leftarrow t_{i-1}}(x(t_{i-1})) + \eta_i$

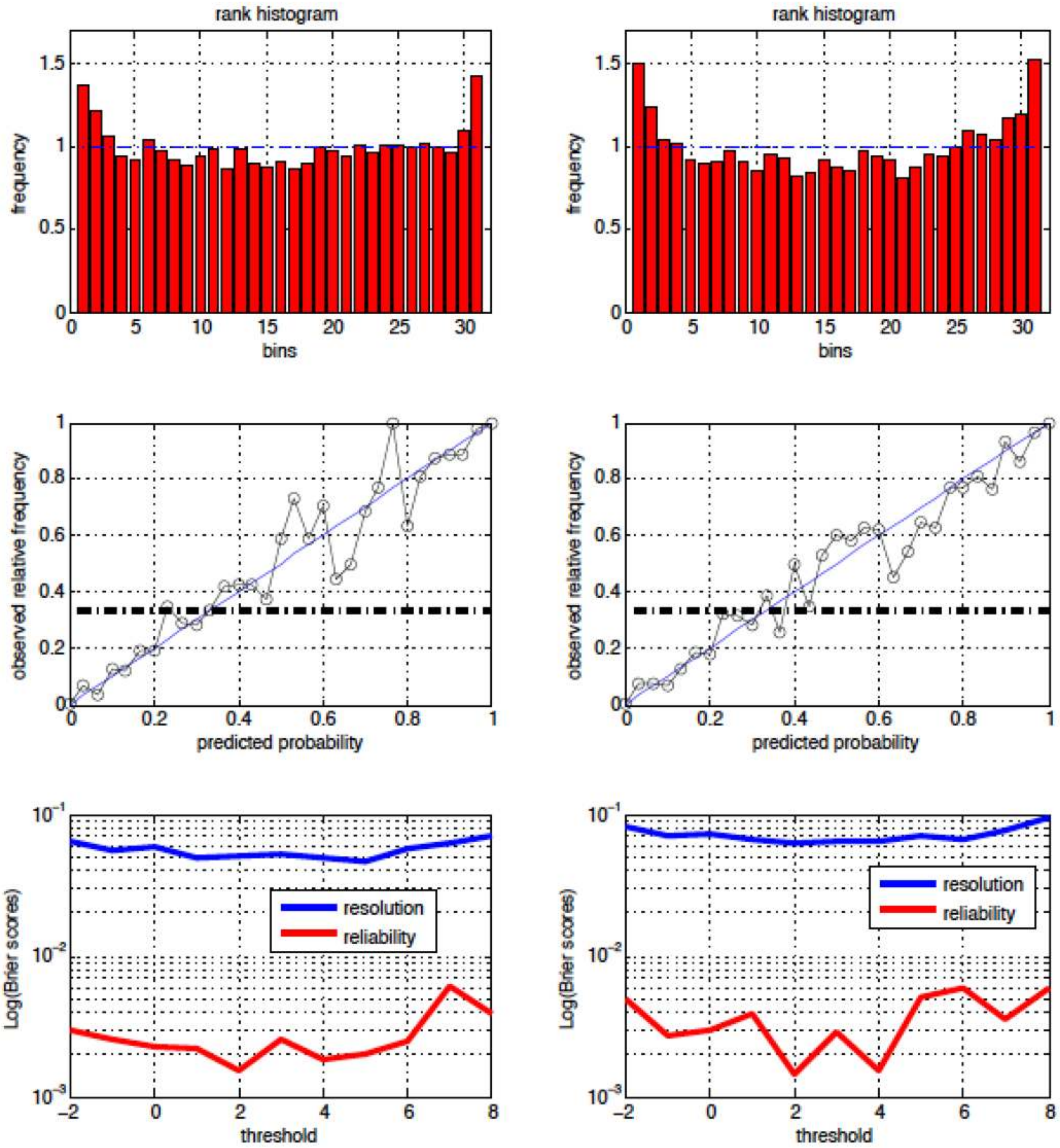
- find the optimal control variable $(x_0^{opt}, \eta_1^{opt}, \eta_2^{opt}, \cdots .\eta_N^{opt})$ and the optimal trajectory x^{opt}.

$$(x_0^{opt}, \eta_1^{opt}, \eta_2^{opt}, \cdots .\eta_N^{opt}) = \min_{x,\eta_1,\eta_2,\cdots.\eta_N \in \mathfrak{A}} \mathfrak{J}(x,\eta_1,\eta_2,\cdots.\eta_N)$$

$$x_i^{opt} = \mathfrak{M}_{t_i \leftarrow t_{i-1}}(\mathfrak{M}_{t_{i-1}\leftarrow t_{i-2}} \cdot\cdot (\mathfrak{M}_{t_2 \leftarrow t_1}(\mathfrak{M}_{t_1 \leftarrow t_0}(x_0^{opt}) + \eta_1^{opt}) + \eta_2^{opt}) \cdot\cdot + \eta_{i-1}^{opt}) + \eta_i^{opt}$$

Weak constraint
Q = 0.001

Left column: EnsVAR
Right column: EnKF



Summary

Summary

- Under non-linearity and non-Gaussianity the EnsVar is a reliable and consistent ensemble estimator (provided the QSVA is used for long DA windows) .
- EnsVar is at least as good an estimator as EnKF and PF.
- Similar results have been obtained for the Kuramuto-Sivashinsky model.

Ensembles obtained are Gaussian, even if errors in data are not

Produces Monte-Carlo sample of (probably not) bayesian pdf

EnsVar : Pros and cons

Pros

- Easy to implement when having a 4D-Var code
- Highly parallelizable
- No problems with algorithm stability (i.e. no ensemble collapse, no need for localization and inflation, no need for weight resampling)
- Propagates information in both ways and takes into account temporally correlated errors

Cons

- Costly (Nens 4D-Var assimilations).
- Empirical.
- Cycling of the process (work in progress).

Assimilation, which originated from the need of defining initial conditions for numerical weather forecasts, has progressively extended to many diverse applications

- Oceanography
- Atmospheric chemistry (both troposphere and stratosphere)
- Oceanic biogeochemistry
- Ground hydrology
- Terrestrial biosphere and vegetation cover
- Glaciology
- Magnetism (both planetary and stellar)
- Plate tectonics
- Planetary atmospheres (Mars, …)
- Reassimilation of past observations (mostly for climatological purposes, ECMWF, NCEP/NCAR)
- Palaeoclimatology
- Identification of source of tracers
- Parameter identification
- *A priori* evaluation of anticipated new instruments
- Definition of observing systems *(Observing Systems Simulation Experiments)*
- Validation of models
- Sensitivity studies (adjoints)
- …

It has now become a major tool of numerical environmental science

Ensemble Prediction

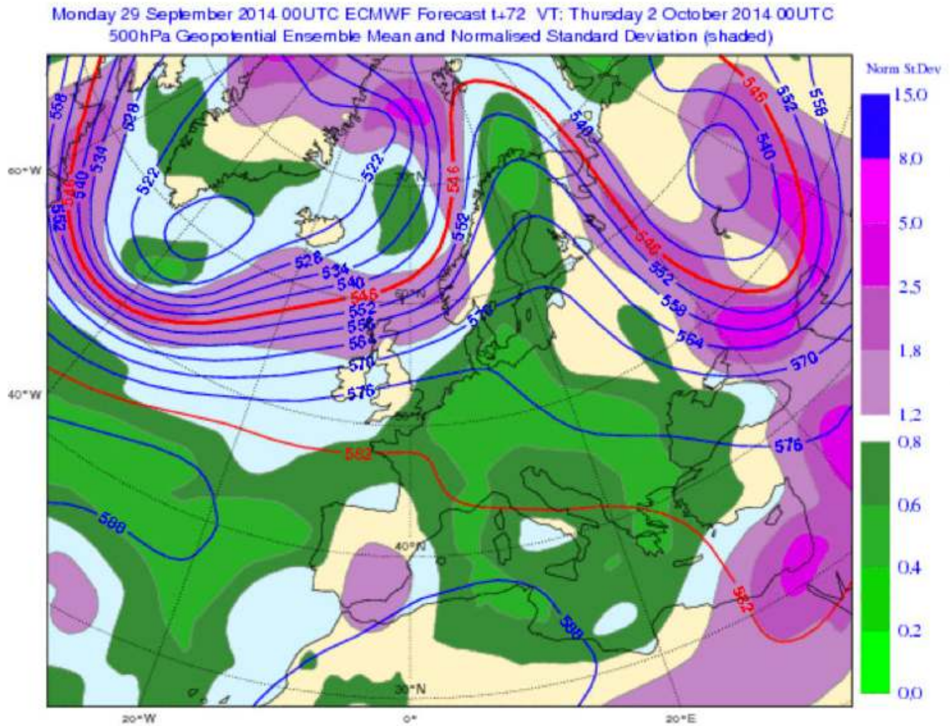
Ensemble Prediction Systems (*EPS*s) are implemented daily in a number of meteorological services. Typically, N model states are chosen at initial time, which are meant to sample the initial uncertainty on the state of the atmosphere (or to sample the components of the uncertainty that matter most). These N states are evolved with the model equations, thereby producing N estimates at any forecast time. In some systems, perturbations are added in the course of the integrations, in order to simulate the effect of errors in the model, considered as a simulator of the atmosphere.

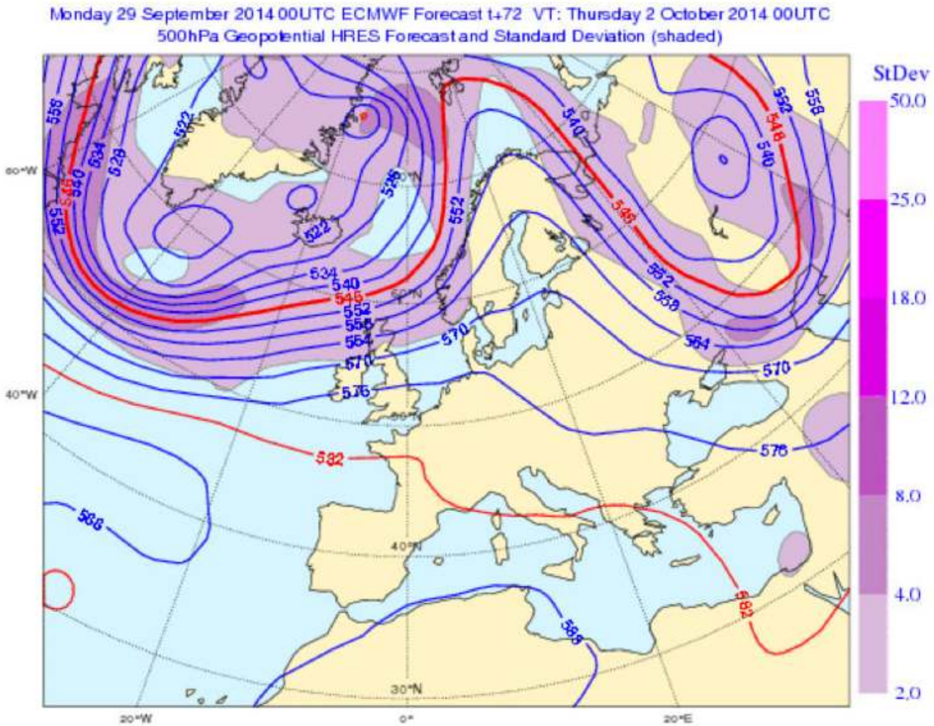
In present *EPS*s, N lies in the range 10-50.

There exist both global and local Ensemble Prediction Systems.

Ensemble Prediction

As part of *The Observing System Research and Predictability Experiment* (*THORPEX*), the *THORPEX Interactive Grand Global Ensemble* (*TIGGE*) consists of the (global) forecast ensembles produced by ten meteorological services (see http://tigge.ecmwf.int/models.html). The total number of forecasts is about 300. The forecasts are accessible within a few days of production.





CEPMMT, Reading, UK

Richardson *et al.*, 2013,
Evaluations of ECMWF forecasts,
…

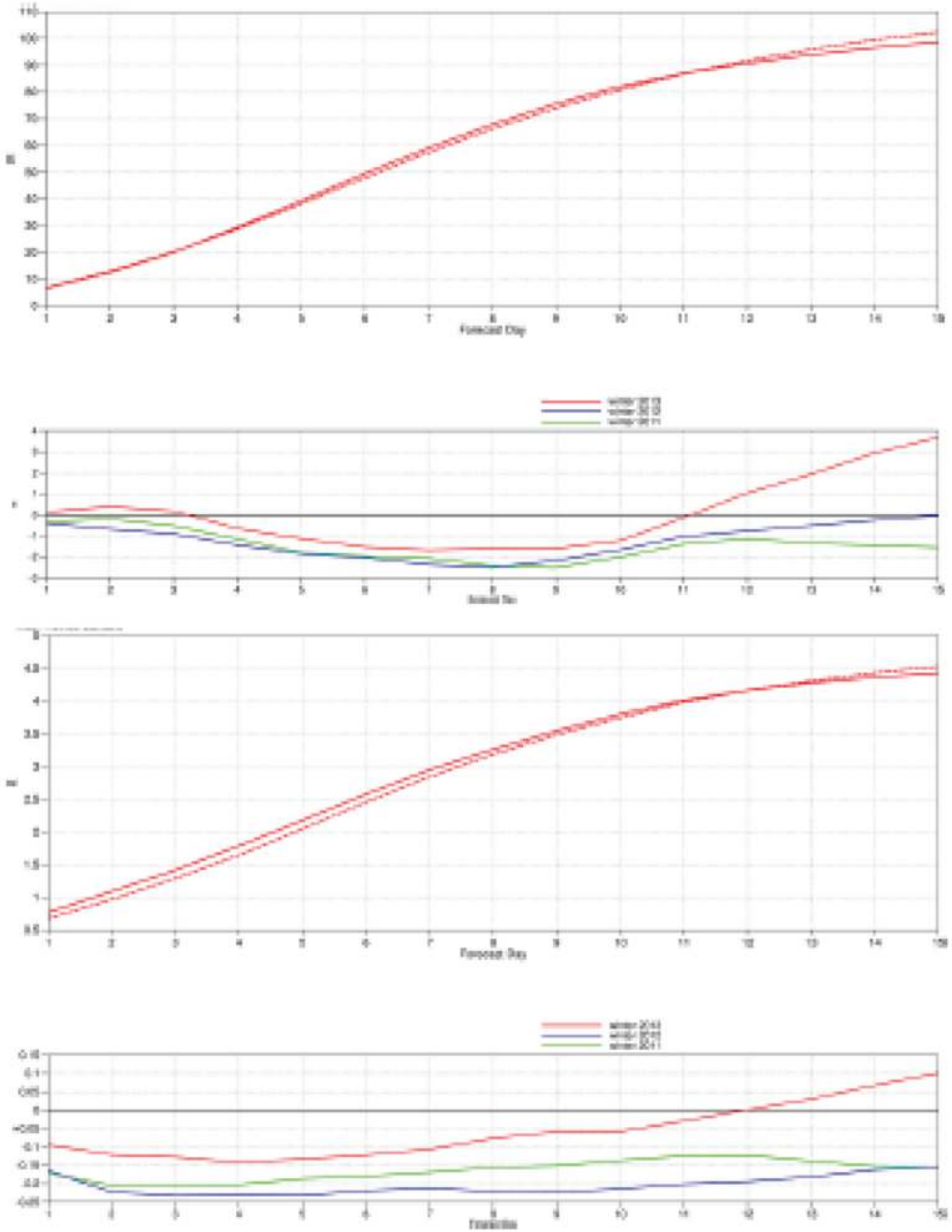


Figure 10: Ensemble spread (standard deviation, dashed lines) and RMS error of ensemble-mean (solid lines) for winter 2012–2013 (upper figure in each panel), and differences of ensemble spread and RMS error of ensemble mean for last three winter seasons (lower figure in each panel, negative values indicate spread is too small); plots are for 500 hPa geopotential (top) and 850 hPa temperature (bottom) over the extratropical northern hemisphere for forecast days 1 to 15.

Richardson *et al.*, 2013, *Evaluations of ECMWF forecasts*, …

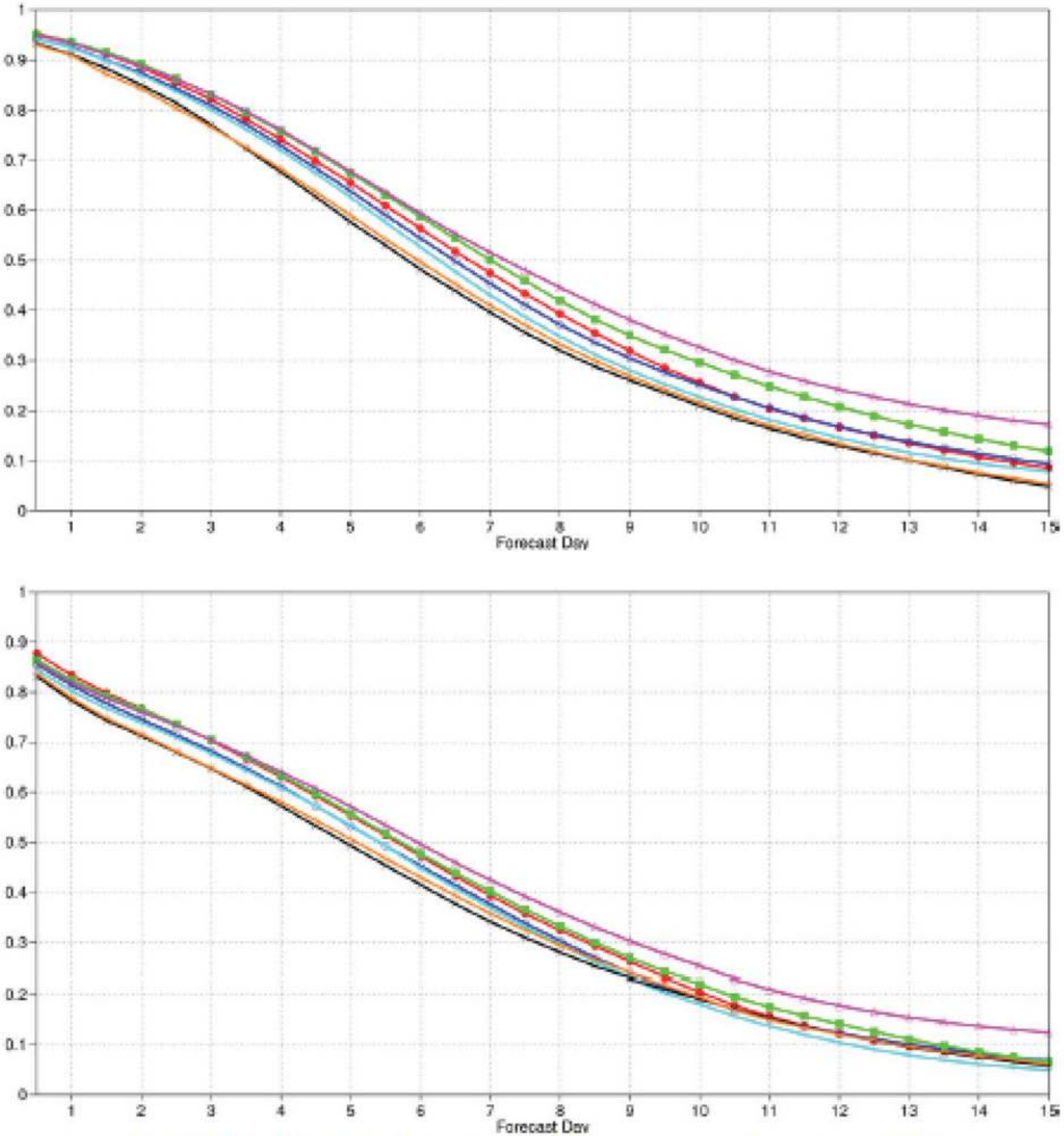
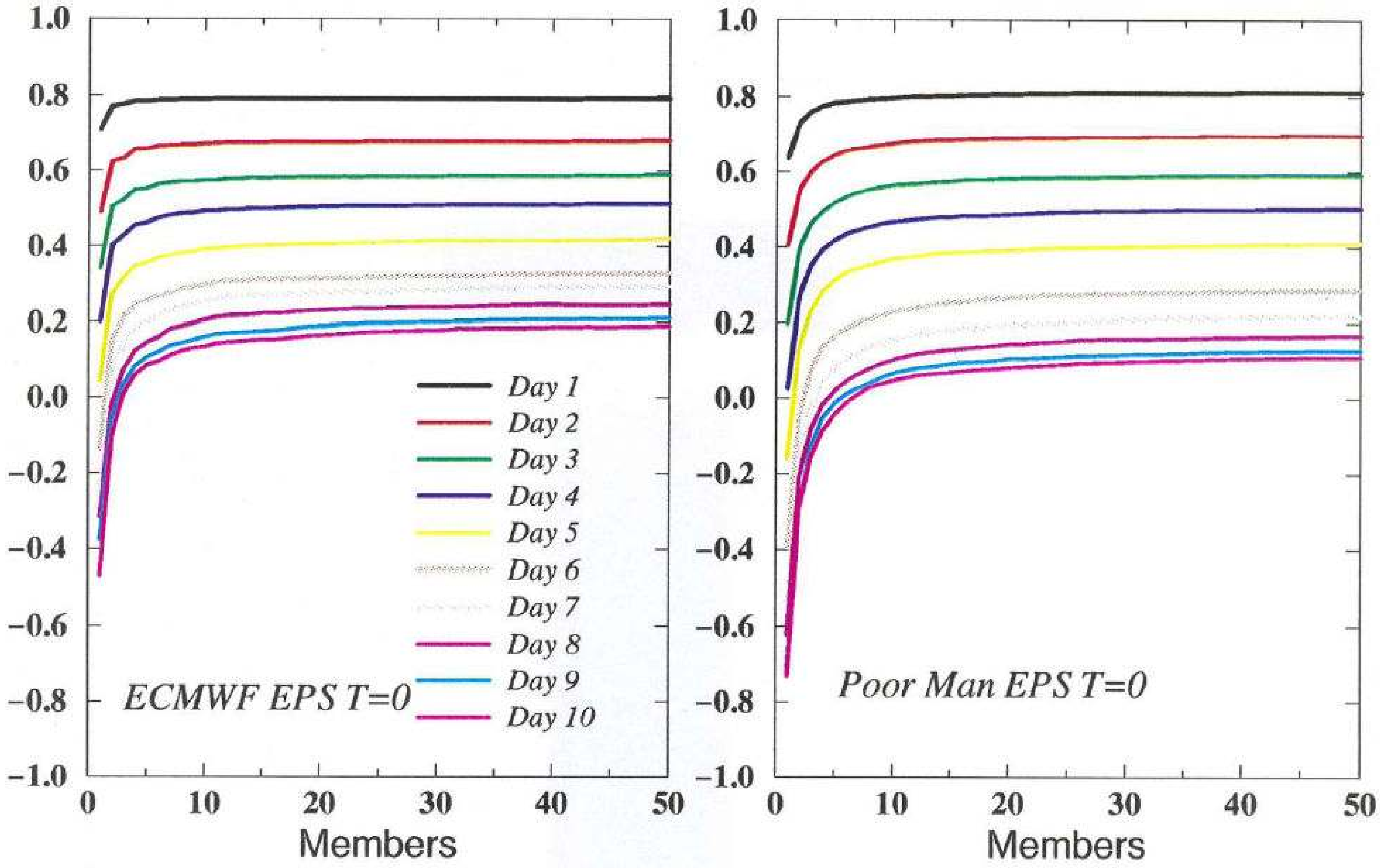


Figure 11: CPRSS for 500 hPa height (top) and 850 hPa temperature (bottom) ensemble forecasts for winter (December–February) over the extratropical northern hemisphere. Skill from the ensemble day 1–15 forecasts is shown for winters 2012–13 (red), 2011–12 (blue), 2010–11 (green), 2009–10 (magenta), 2008–09 (cyan), 2007–08 (black) and 2006–07 (orange).

Size of Ensembles ?

Given the choice, is it better to improve the quality of the forecast model, or to increase the size of the predicted ensembles ?

- o Observed fact : in ensemble prediction, present scores saturate for ensemble size N in the range 30-50.



Impact of ensemble size on Brier Skill Score
ECMWF, event $T_{850} > T_c$ Northern Hemisphere

(Talagrand *et al.*, ECMWF, 1999)

Theoretical estimate (raw Brier score)

$$B_N = B_\infty + \frac{1}{N}\int_0^1 p(1-p)g(p)dp$$

Brier score

$$B \equiv E[(p-o)^2]$$

For N-sized finite ensembles

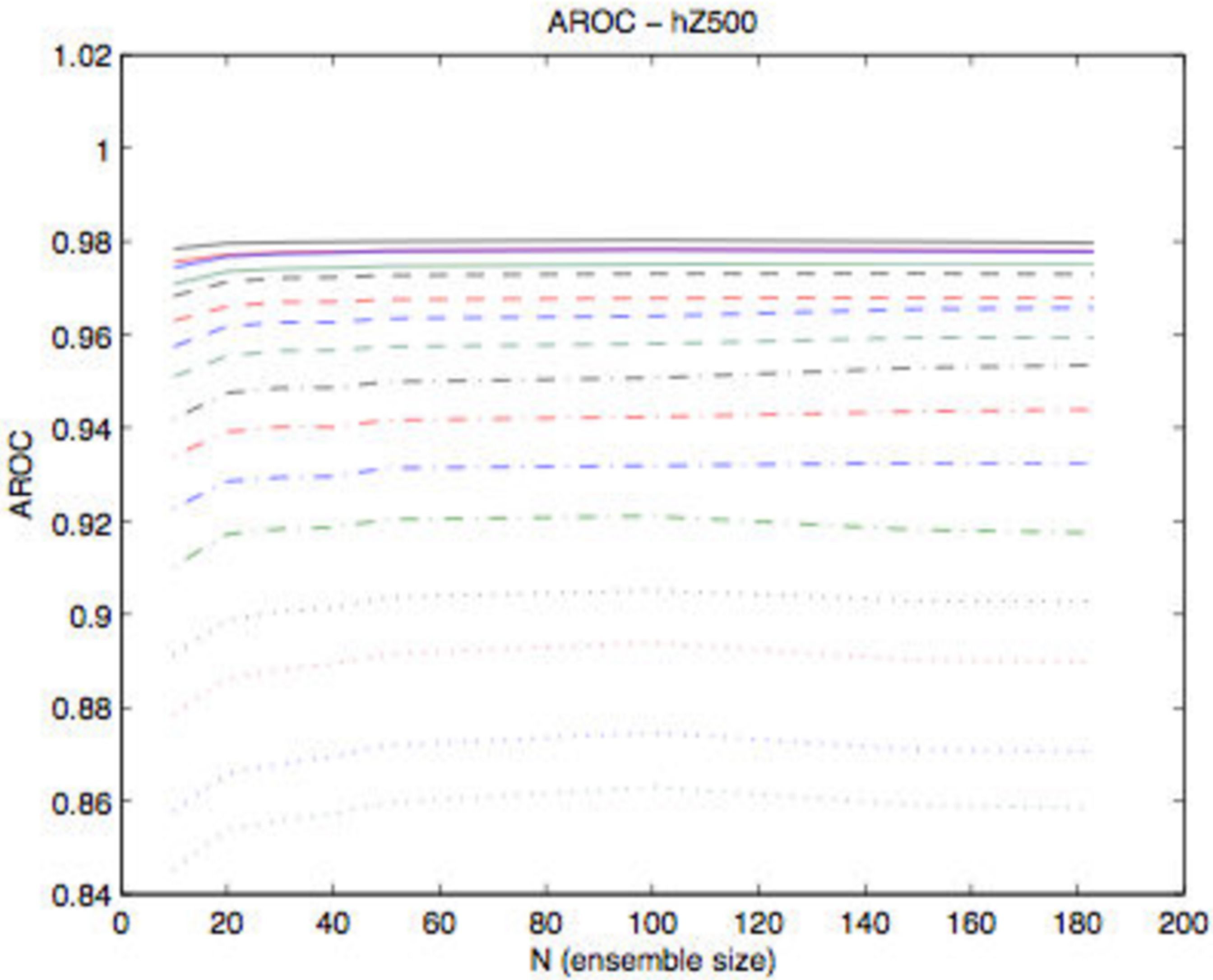
$$B_N = B_\infty + \frac{1}{N} E[p(1-p)]$$

(Talagrand *et al*., ECMWF, 1999, Richardson, *QJRMS*, 2001)

The sharper the distribution of raw predicted probabilities, the more rapid the saturation of the score.

$E[p(1-p)]$ is half the probability that a 2-member ensemble will produce a .5 probability for the occurrence of the event under consideration. Brier score for infinite ensemble size N can be determined from an infinite number of realizations of finite ensembles.

No similar formula found for reliability and resolution components separately.



TIGGE, ROC curve area, courtesy L. Descamps

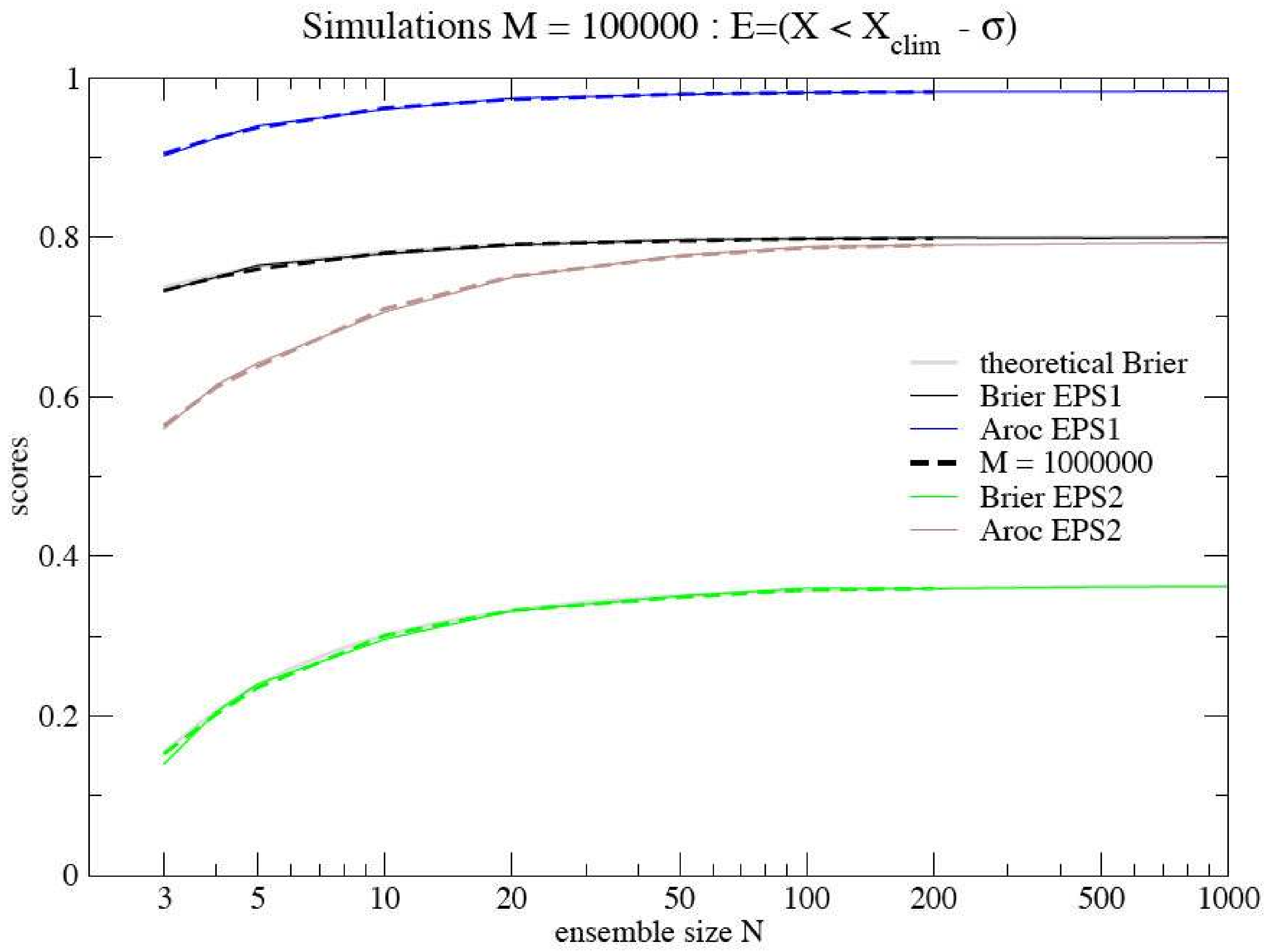
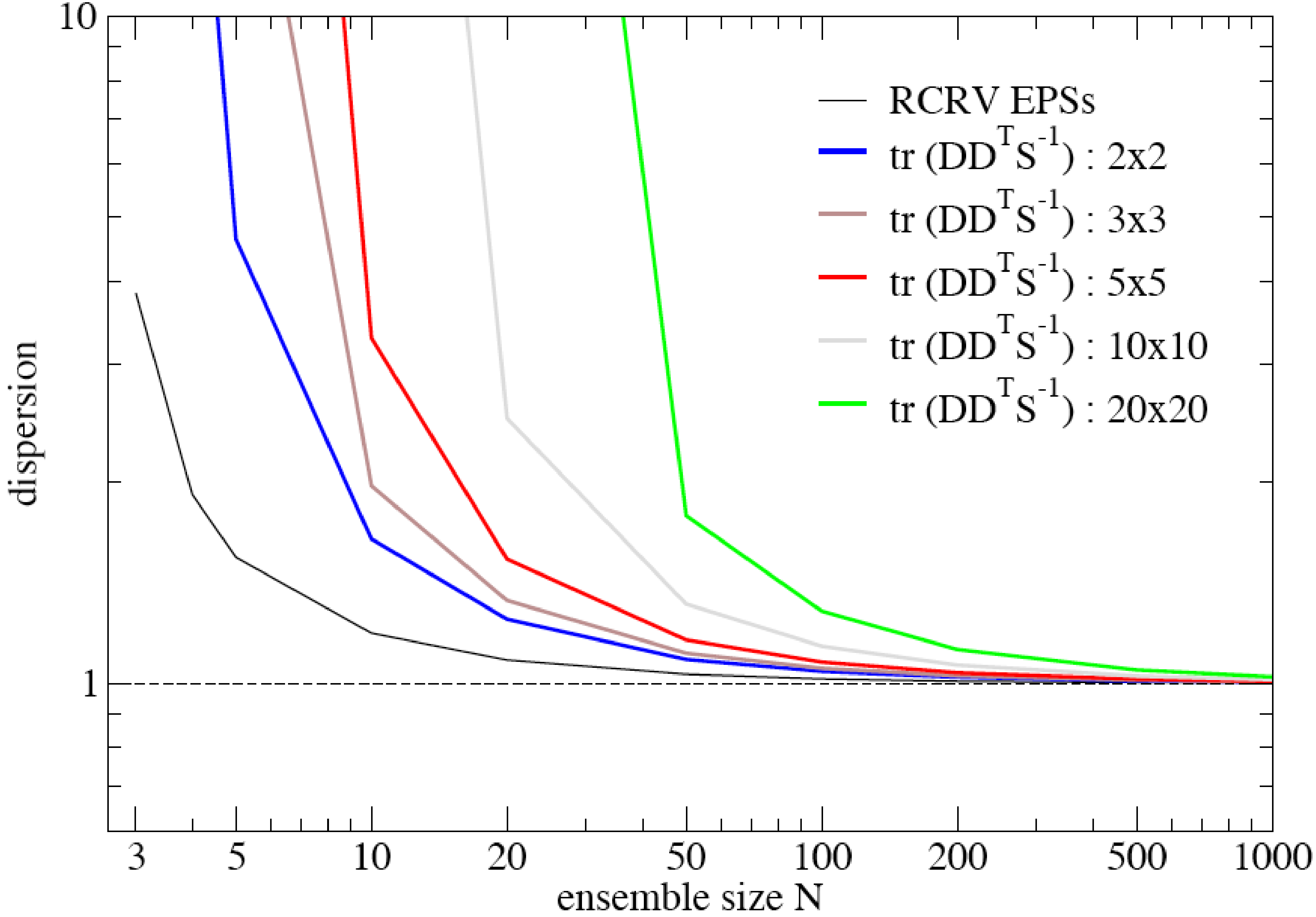


Figure 1: Impact of N on Brier Skill Score and ROC area

G. Candille, 2009

Scores saturate for ensemble sizes N of the order of a few tens. The higher the sharpness of the predicted probabilities, the more rapid the saturation.

In addition, considerations on frequencies of occurrences show that size of validation sample necessary to check reliability increases like N.



Impact of dimension of multidimensional variables (G. Candille, 2008)

Conclusion on ensemble size

Objective scores saturate in the range $N \approx$ 30-50 because it is possible in practice to evaluate only probabilisic predictions of events or of one-dimensional variables. Evaluating probabilistic predictions of multi-dimensional variables would require validating samples of inaccessible size.

Is there any point in taking larger values of N ?

Questions and Problems on Ensemble Prediction

- Quantify errors (especially model errors) more accurately

  Dispersion of forecasts performed with (exactly) the same model from different initial conditions is a statistical integrated measure of sensitivity to initial conditions. Comparison with deviation from observed reality then provides by difference a statistical integrated measure of model error (idea behind 'Lorenz curves').

- Definition of initial conditions (continuation of ensemble assimilation, or something else) ?

- Must we tend to a situation where the output of prediction will be a probability distribution ?

Uncertainties in Weather Prediction and Climate Models ? (1)

Models are generally considered as being good simulators of the atmosphere as concerns the 'dynamics', namely the effects of gravity, pressure gradient and rotation (*i. e.*, the thermodynamically reversible processes).

They are considered as much less accurate as concerns the 'physics', *i. e.*, the thermodynamically irreversible processes. This is particularly true of the water cycle in general, and of the exchanges of momentum, energy and moisture between the atmosphere and the underlying medium. Many of these processes take place at spatial (and temporal) scales that are not explicitly resolved by the model, and their description requires some form of 'parametrization'.

Uncertainties in Weather Prediction and Climate Models ? (2)

How to represent associated uncertainties ?

Deterministic models, run in succession with different values of parameters that are considered as being uncertain ?

Stochastic models, in which uncertainty is represented by random perturbations in the model equations ?

Or any intermediate approach ?

In any case, evaluation will have to be made on the basis of ensembles. And, as concerns the future evolution of climate, objective validation can be made only on past climate (either recent or less recent)

Uncertainties in Weather Prediction and Climate Models ? (3)

"*I believe that the ultimate climatic models will be stochastic, i.e. random numbers will appear somewhere in the time derivatives***"**
(Lorenz 1975)

The End !