2nd THORPEX Intl. Sci. Symp. (STISS)

4–8 Dec. 2006, Landshut, Bavaria

Dynamical Systems, Sequential Estimation, and Estimating Parameters

Michael Ghil

Ecole Normale Supérieure, Paris, and
University of California, Los Angeles

Joint work with

D. Kondrashov and J. D. Neelin, UCLA; C.-J. Sun, NASA Goddard; A. Carrassi, U. of Ferrara; A. Trevisan, ISAC-CNR, Bologna; F. Uboldi, Milano; and many others: please see http://www.atmos.ucla.edu/tcd/

Outline

- Data in meteorology and oceanography
  - *in situ &* remotely sensed
- Basic ideas, data types, & issues
  - how to combine data with models
  - transfer of information
    - between variables & regions
  - stability of the fcst.–assimilation cycle
  - filters & smoothers
- Parameter estimation
  - model parameters
  - noise parameters – at & below grid scale
- Subgrid-scale parameterizations
  - deterministic (“classic”)
  - stochastic – “dynamics” & “physics”
- Novel areas of application
  - space physics
  - shock waves in solids
  - macroeconomics
- Concluding remarks

Main issues

- The solid earth stays put to be observed, the atmosphere, the oceans, & many other things, do not.
- Two types of information:

  - direct → observations, and

  - indirect → dynamics (from past observations);

  both have errors.
- Combine the two in (an) optimal way(s)
- Advanced data assimilation methods provide such ways:

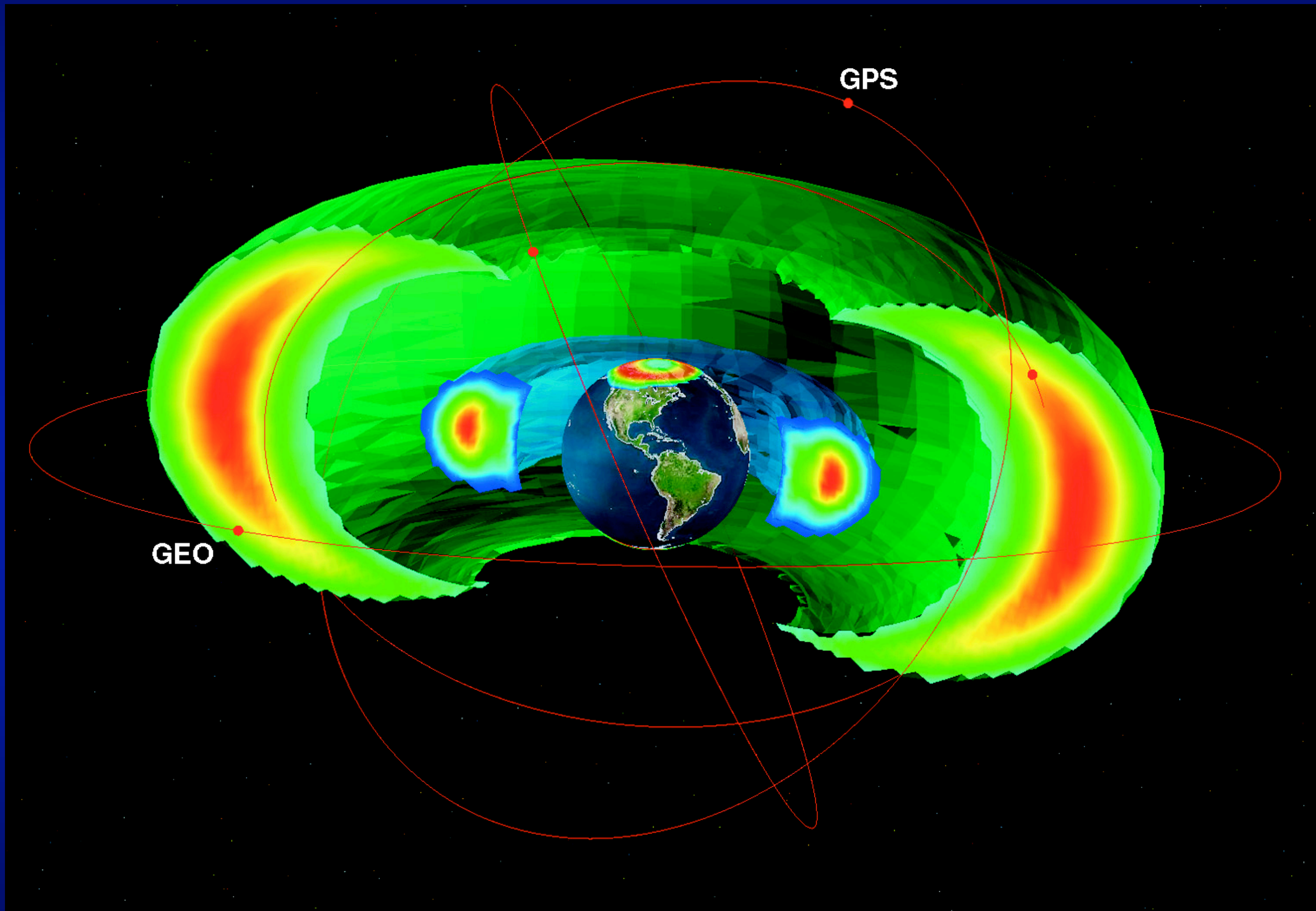
  - sequential estimation → the Kalman filter(s), and

  - control theory → the adjoint method(s)
- The two types of methods are essentially equivalent for simple linear systems (the duality principle)

Main issues (continued)

- Their performance differs for large nonlinear systems in:
  - accuracy, and
  - computational efficiency
- Study optimal combination(s), as well as improvements over currently operational methods (OI, 4-D Var, PSAS, EnKF).

Space physics data



Space platforms in Earth's magnetosphere

Extended Kalman Filter (EKF)

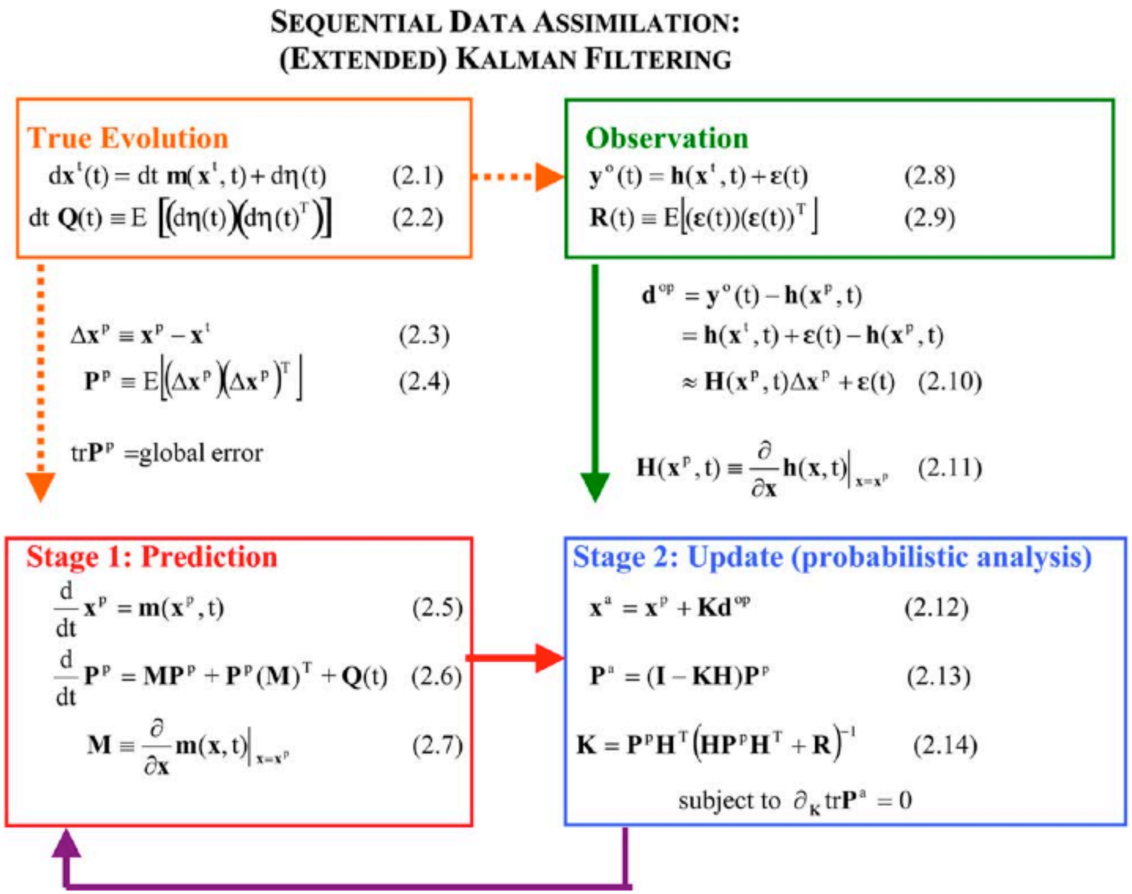


Fig. 1. A flow-chart representation of the EKF method (see Table 1 for definitions of the symbols).

Basic concepts: barotropic model

Shallow-water equations in 1-D, linearized about $(U, 0, \Phi)$, $fU = -\Phi_y$

$U = 20\ \mathrm{ms}^{-1}$, $f = 10^{-4}\mathrm{s}^{-1}$, $\Phi = gH$, $H \approx 3$ km.

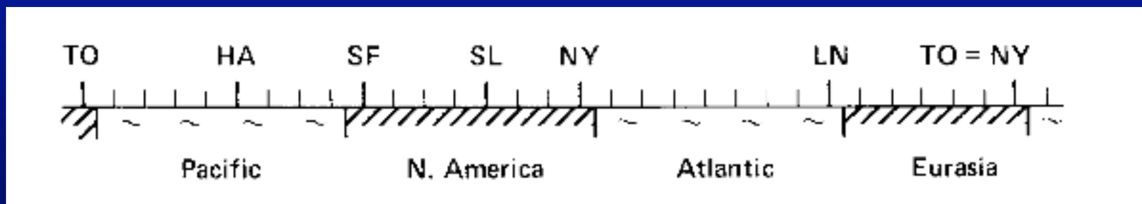
$$u_t + Uu_x + \phi_x - fv = 0$$

$$v_t + Uv_x + fu = 0$$

$$\phi_t + U\phi_x + \Phi u_x - fUv = 0$$

PDE system discretized by finite differences, periodic B. C.

$\mathbf{H}_k$: observations at synoptic times, over land only.



Ghil *et al*. (1981), Cohn & Dee (Ph.D. theses, 1982 & 1983), etc.

Conventional network

Relative weight of observational *vs.* model errors

$P_\infty = QR/[Q + (1 - \Psi^2)R]$

(a) $Q = 0 \Rightarrow P_\infty = 0$

(b) $Q \neq 0 \Rightarrow$ (i), (ii) and (iii):

(i) "good" observations

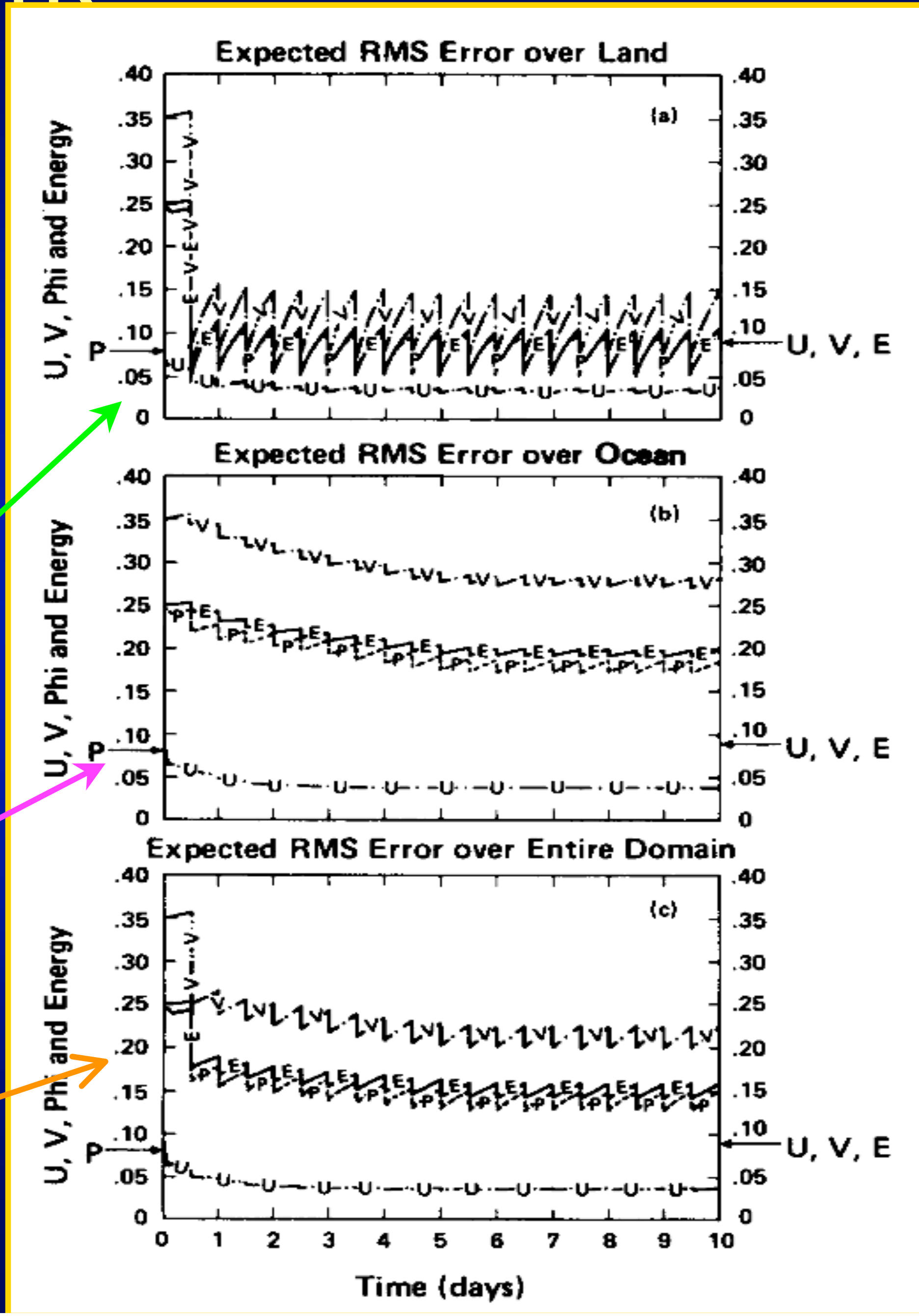
$R << Q \Rightarrow P_\infty \approx R$;

(ii) "poor" observations

$R >> Q \Rightarrow P_\infty \approx Q/(1 - \Psi^2)$;

(iii) always (provided $\Psi^2 < 1$)

$P_\infty \leq \min\{R, Q/(1 - \Psi^2)\}$.



Advection of information

Upper panel (NoSat):

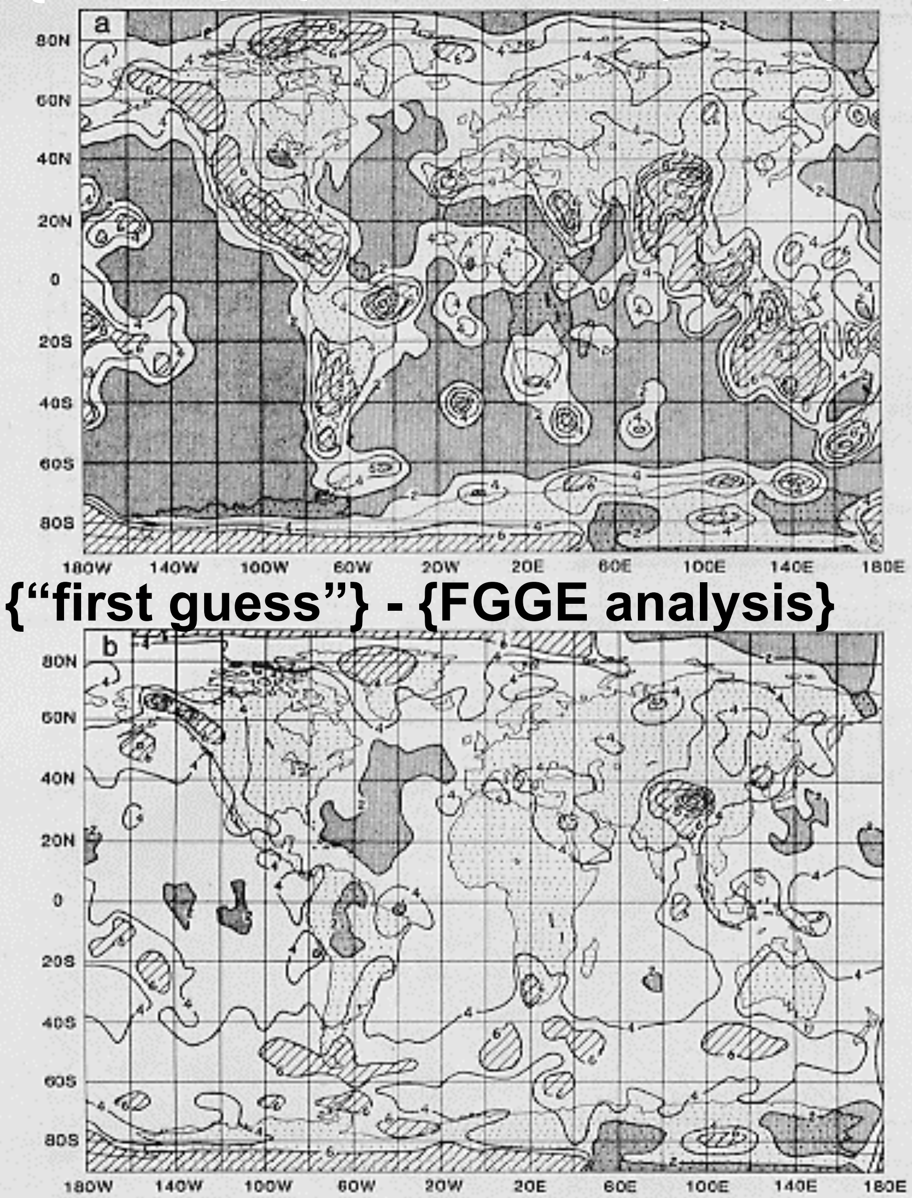
Errors advected off the ocean

Lower panel (Sat):

Errors drastically reduced, as info. now comes in, off the ocean

{6h fcst} - {conventional (NoSat)}

ϕ_{300}



ϕ_{300}

FIG. 5. The rms difference between the 6 h forecast of the 300 mb geopotential height field and the analysis for the period 5–21 January 1979. Contour interval is 20 m. a) Rms difference between the NOSAT analysis and forecast. b) Rms difference between the FGGE analysis and forecast.

Halem, Kalnay, Baker & Atlas

(*BAMS*, 1982)

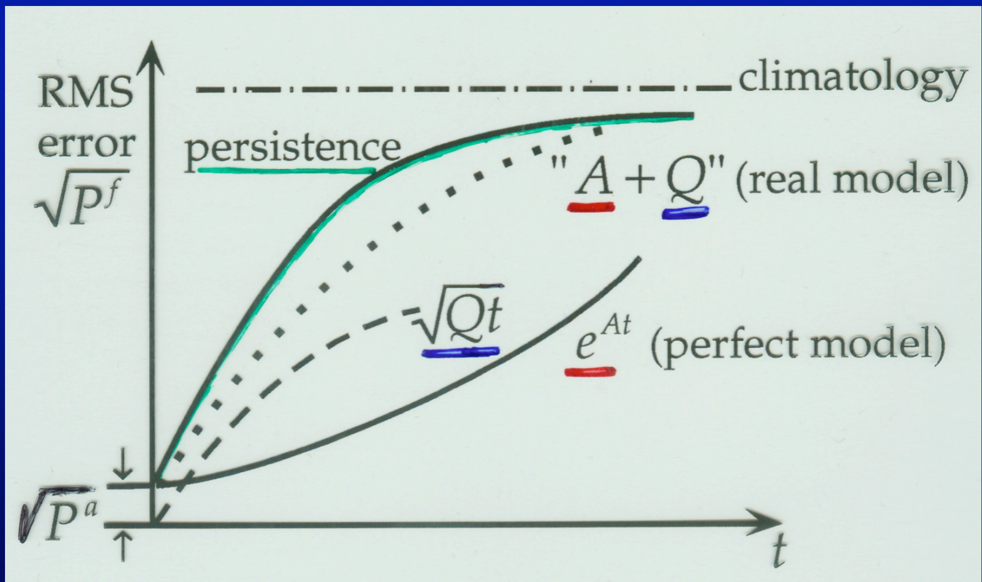
Outline

- Data in meteorology and oceanography
  - *in situ &* remotely sensed
- ➤ Basic ideas, data types, & issues
  - how to combine data with models
  - stability of the fcst.–assimilation cycle
  - filters & smoothers
- Parameter estimation
  - model parameters
  - noise parameters – at & below grid scale
- Subgrid-scale parameterizations
  - deterministic ("classic")
  - stochastic – "dynamics" & "physics"
- Novel areas of application
  - space physics
  - shock waves in solids
  - macroeconomics
- Concluding remarks

Error components in forecast–analysis cycle

$$\underbrace{P^f}_{\text{first-guess error}} \cong \underbrace{P^a}_{\text{analysis error}} + \Delta t(\underbrace{2AP^a}_{\text{id. twins error growth}} + \underbrace{Q}_{\text{modeling error}})$$

$$(\Psi = e^{A\Delta t} \cong \underline{1 + A\Delta t})$$



The relative contributions to error growth of

- **analysis error**
- **intrinsic error growth**
- **modeling error (stochastic?)**

Assimilation of observations: Stability considerations

Free-System Dynamics (sequential-discrete formulation): ***Standard breeding***

forecast state; model integration from a previous analysis

$$\mathbf{x}^f_{n+1} = M(\mathbf{x}^a_n)$$

Corresponding perturbative (tangent linear) equation

$$\delta\mathbf{x}^f_{n+1} = \mathbf{M}\,\delta\mathbf{x}^a_n$$

Observationally Forced System Dynamics (sequential-discrete formulation): ***BDAS***

If observations are available and we assimilate them:

Evolutive equation of the system, subject to forcing by the assimilated data

$$\mathbf{x}^a_{n+1} = \left[\mathbf{I} - \mathbf{K}H \circ\right] M(\mathbf{x}^a_n) + \mathbf{K}\mathbf{y}^o_{n+1}$$

Corresponding perturbative (tangent linear) equation, if the same observations are assimilated in the perturbed trajectories as in the control solution

$$\delta\mathbf{x}^a_{n+1} = \left[\mathbf{I} - \mathbf{K}\mathbf{H}\right]\mathbf{M}\,\delta\mathbf{x}^a_n$$

- The matrix (**I – KH**) is expected, in general, to have a stabilizing effect;
- the free-system instabilities, which dominate the forecast step error growth, can be reduced during the analysis step.

Joint work with A. Carrassi, A. Trevisan & F. Uboldi

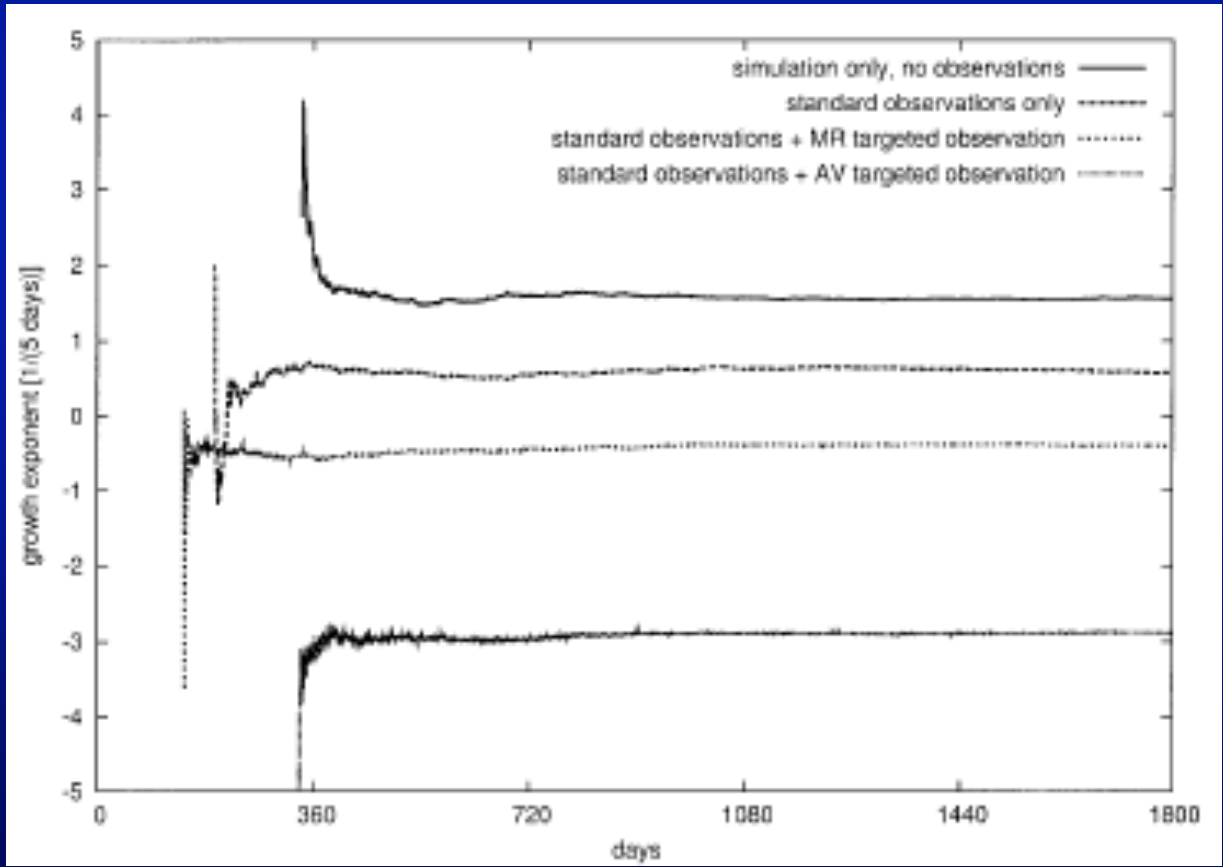
Stabilization of the forecast–assimilation system – I

Assimilation experiment with a low-order chaotic model

- Periodic 40-variable Lorenz (1996) model;
- Assimilation algorithms: replacement (Trevisan & Uboldi, 2004), replacement + one adaptive obs'n located by multiple replication (Lorenz, 1996), replacement + one adaptive obs'n located by BDAS and assimilated by AUS (Trevisan & Uboldi, 2004).

BDAS: Breeding on the Data Assimilation System

AUS: Assimilation in the Unstable Subspace



Trevisan & Uboldi (*JAS*, 2004)

Stabilization of the forecast–assimilation system – II

Assimilation experiment with the
40-variable Lorenz (1996) model

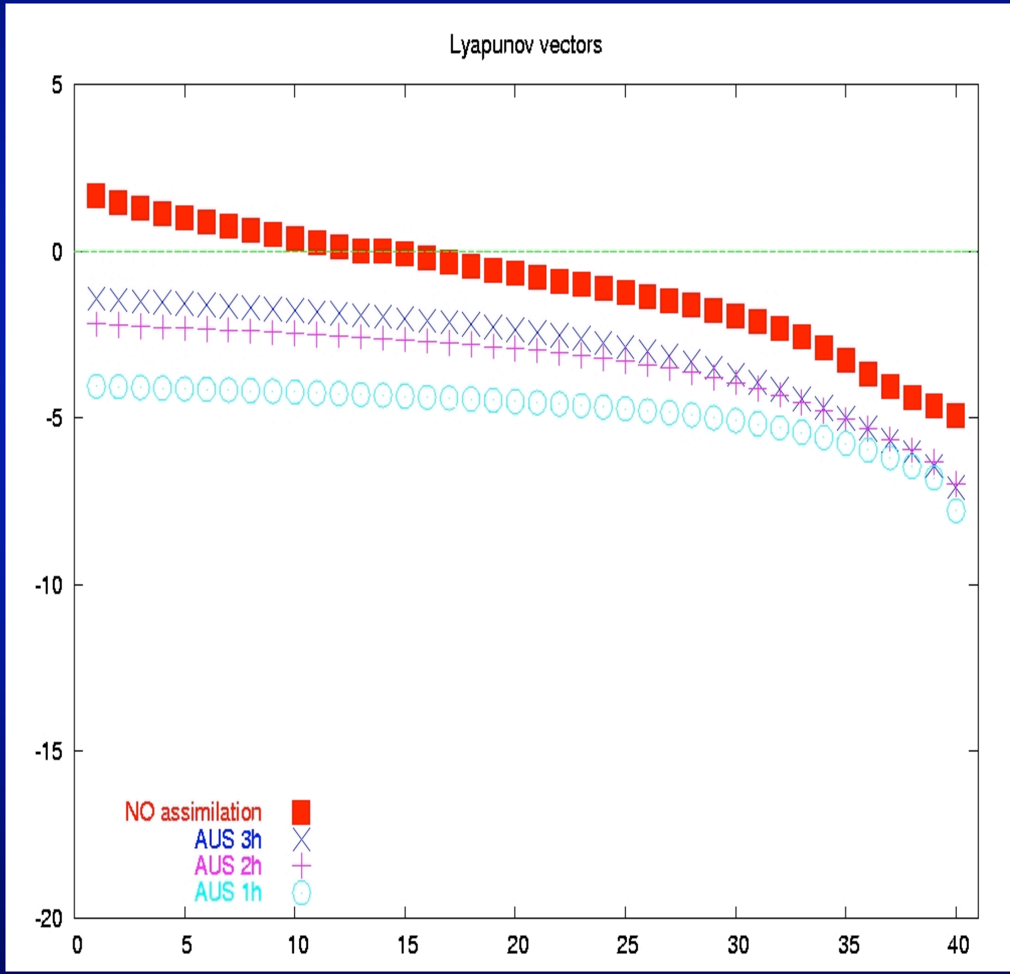
Spectrum of Lyapunov exponents:

Red: free system

Dark blue: AUS with 3-hr updates

Purple: AUS with 2-hr updates

Light blue: AUS with 1-hr updates



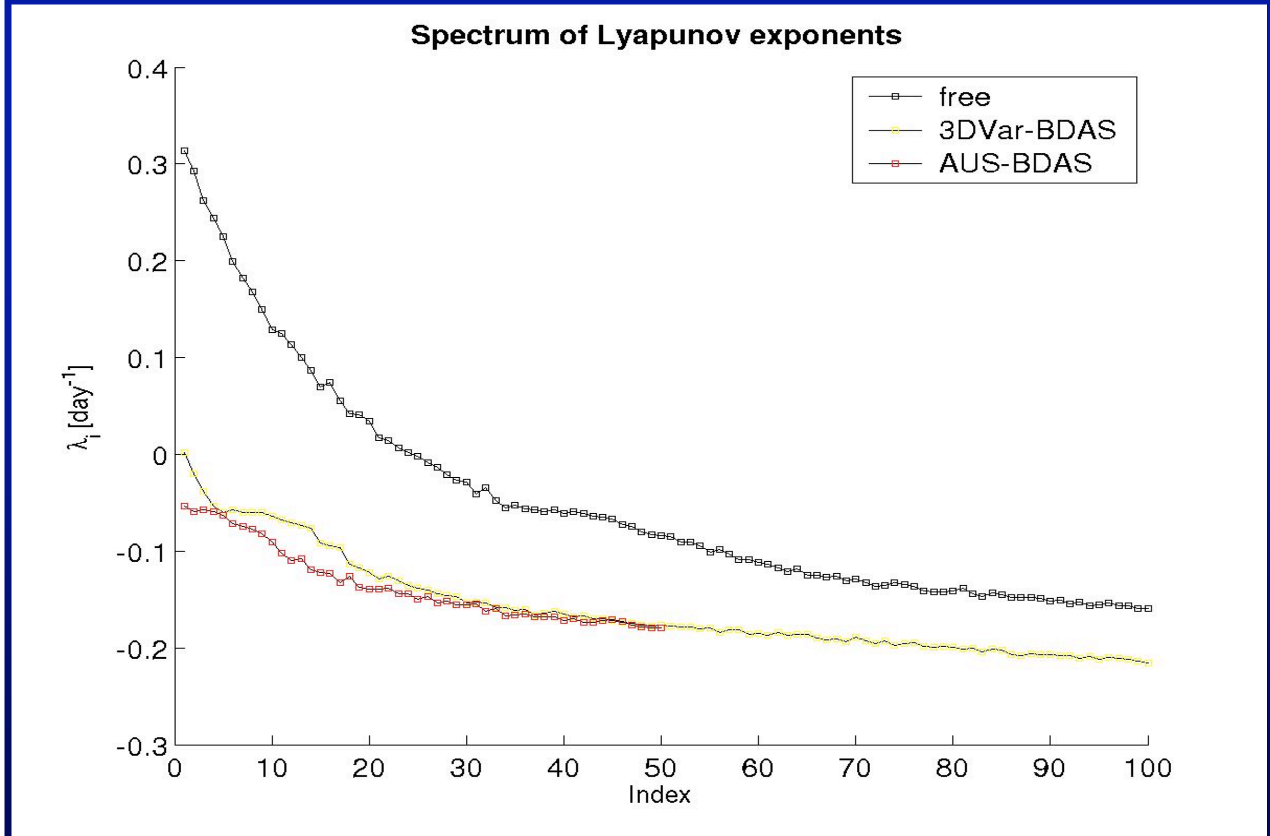
Carrassi, Ghil, Trevisan & Uboldi, 2006, submitted

Stabilization of the forecast–assimilation system – III

Assimilation experiment with an intermediate atmospheric circulation model

- 64-longitudinal x 32-latitudinal x 5 levels periodic channel QG-model (Rotunno & Bao, 1996)
- Perfect-model assumption
- Assimilation algorithms: 3-DVar (Morss, 2001); AUS (Uboldi *et al.,* 2005; Carrassi *et al*., 2006)

Observational forcing ⇒ Unstable subspace reduction



➤ ***Free System***

Leading exponent:
$\lambda_{max} \approx 0.31$ days^{-1};
Doubling time ≈ 2.2 days;
Number of positive exponents:
$N^+ = 24$;
Kaplan-Yorke dimension ≈ 65.02.

➤ ***3-DVar–BDAS***

Leading exponent:
$\lambda_{max} \approx 6\text{x}10^{-3}$ days^{-1};

➤ ***AUS–BDAS***

Leading exponent:
$\lambda_{max} \approx -0.52\text{x}10^{-3}$ days^{-1}

Outline

- Data in meteorology and oceanography
  - *in situ &* remotely sensed
- Basic ideas, data types, & issues
  - how to combine data with models
  - filters & smoothers
  - stability of the fcst.-assimilation cycle
- ➤ Parameter estimation
  - model parameters
  - noise parameters – at & below grid scale
- Subgrid-scale parameterizations
  - deterministic ("classic")
  - stochastic – "dynamics" & "physics"
- Novel areas of application
  - space physics
  - shock waves in solids
  - macroeconomics
- Concluding remarks

Parameter Estimation

a) Dynamical model

$dx/dt = M(x, \mu) + \eta(t)$

$y^o = H(x) + \varepsilon(t)$

Simple (EKF) idea – augmented state vector

$d\mu/dt = 0$, $X = (x^T, \mu^T)^T$

b) Statistical model

$L(\rho)\eta = w(t)$, L – AR(MA) model, $\rho = (\rho_1, \rho_2, \ldots.\ \rho_M)$

Examples: 1) Dee *et al*. (*IEEE*, 1985) – estimate a few parameters in the covariance matrix $Q = E(\eta, \eta^T)$; also the bias $\langle\eta\rangle = E\eta$;

2) POPs - Hasselmann (1982, Tellus); Penland (1989, *MWR*; 1996, *Physica D*); Penland & Ghil (1993, *MWR*)

3) $dx/dt = M(x, \mu) + \eta$: Estimate both M & Q from data (Dee, 1995, *QJ*), Nonlinear approach: Empirical mode reduction (Kravtsov *et al*., 2005, Kondrashov *et al*., 2005)

Estimating noise – I

$Q_1 = Q_{slow}$, $Q_2 = Q_{fast}$, $Q_3 = 0$;

$R_1 = 0$, $R_2 = 0$, $R_3 = R$;

$Q = \sum \alpha_i Q_i$; $R = \sum \alpha_i R_i$;

$\alpha(0) = (6.0, 4.0, 4.5)^T$;

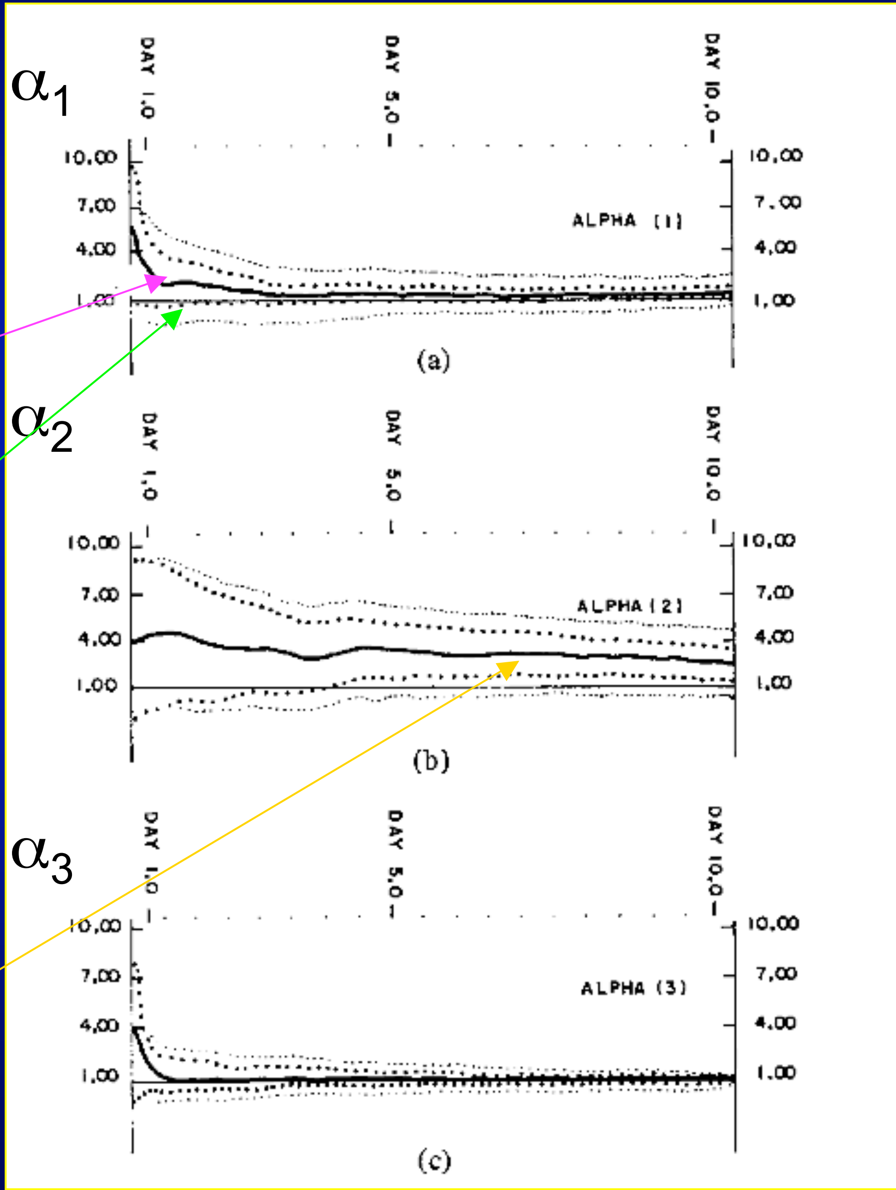
$Q(0) = 25 * I$.

estimated

true ($\alpha = 1$)

Dee *et al*. (1985, *IEEE Trans. Autom. Control*, **AC-30**)

Poor convergence for Q_{fast}?



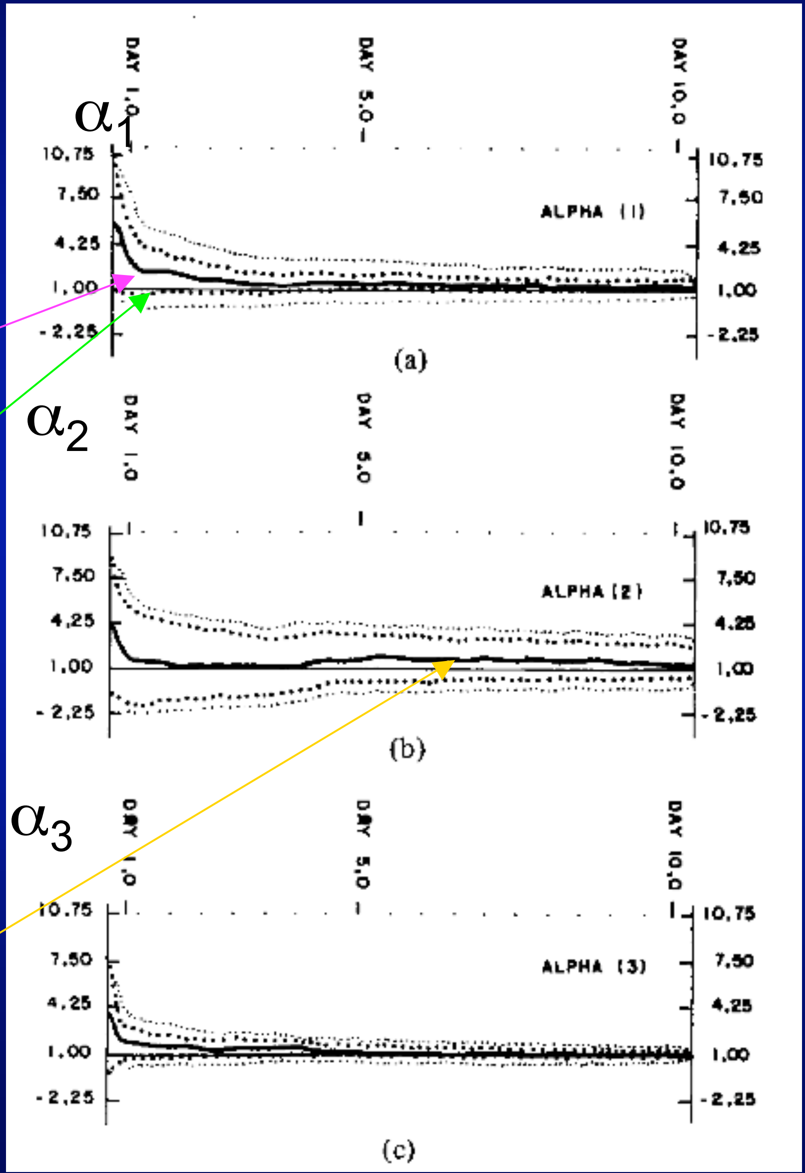
Estimating noise – II

Same choice of $\alpha(0)$, Q_i, and R_i but

$$\Theta(0) = 25 * \begin{bmatrix} 1 & 0.8 & 0 \\ 0.8 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

estimated

true ($\alpha = 1$)



Dee *et al.* (1985, *IEEE Trans. Autom. Control*, **AC-30**)

Good convergence for Q_{fast}!

Sequential parameter estimation

- "**State augmentation**" method – uncertain parameters are treated as additional state variables.
- Example: one unknown parameter μ

$$\bar{x}_k = \begin{pmatrix} x_k \\ \mu_k \end{pmatrix} = \begin{pmatrix} F(x_{k-1}, \mu_{k-1}) \\ \mu_{k-1} \end{pmatrix} + \begin{pmatrix} \epsilon_k \\ \epsilon^{\mu}_{k-1} \end{pmatrix}$$

$$y^o_k = \begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ \mu_k \end{pmatrix} + \epsilon^0 = \bar{H}\bar{x}_k + \epsilon^0$$

$$\bar{x}^a_k = \bar{x}^f_k + \bar{K}(y^o_k - \bar{H}\bar{x}^f_k); \quad \bar{K} = \bar{P}^f \bar{H}^T (\bar{H}\bar{P}^f \bar{H}^T + R)^{-1}$$

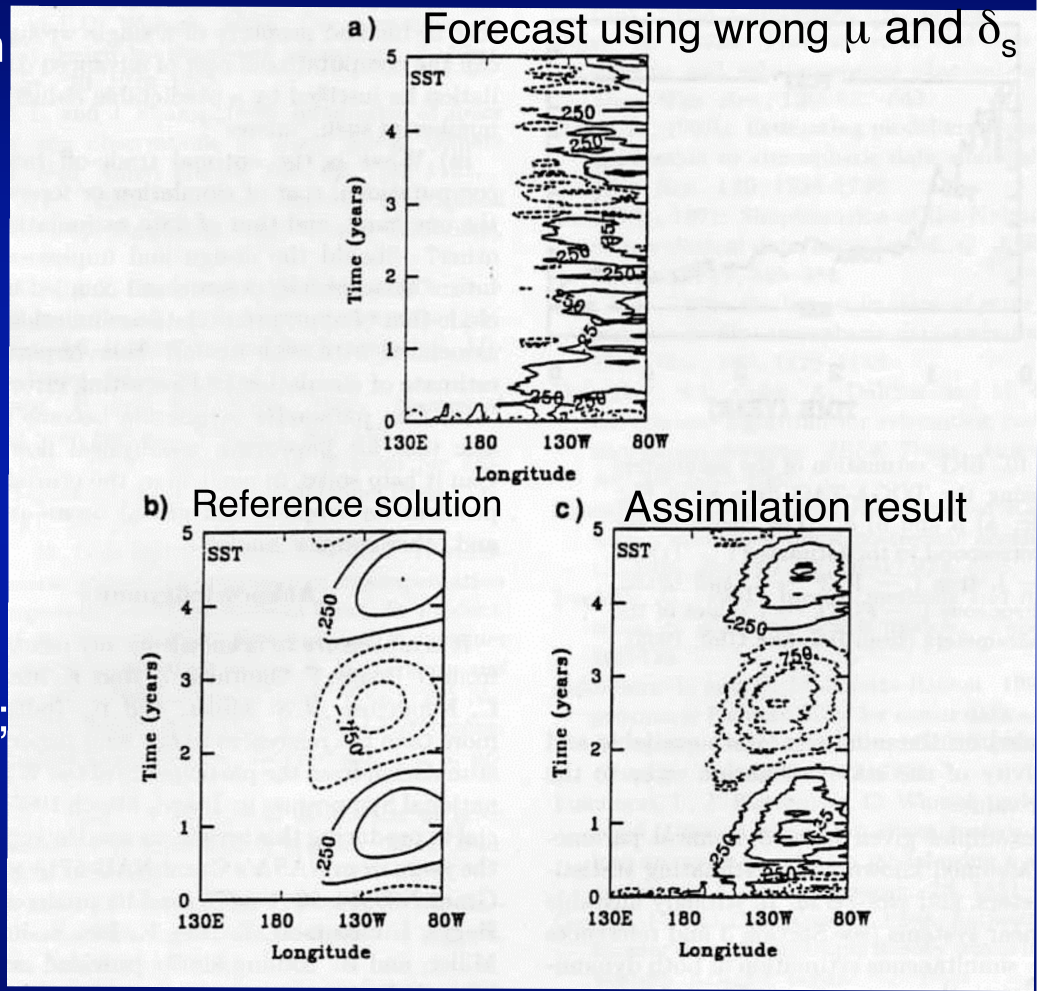
- **The parameters are not directly observable, but** the **cross-covariances** drive parameter changes from innovations of the state:

$$\bar{P}^f = \begin{pmatrix} P^f_{xx} & P^f_{x\mu} \\ P^f_{\mu x} & P^f_{\mu\mu} \end{pmatrix}; \quad \bar{K} = \begin{pmatrix} P^f_{xx} H^T \\ P^f_{\mu x} H^T \end{pmatrix} (H P^f_{xx} H^T + R)^{-1}$$

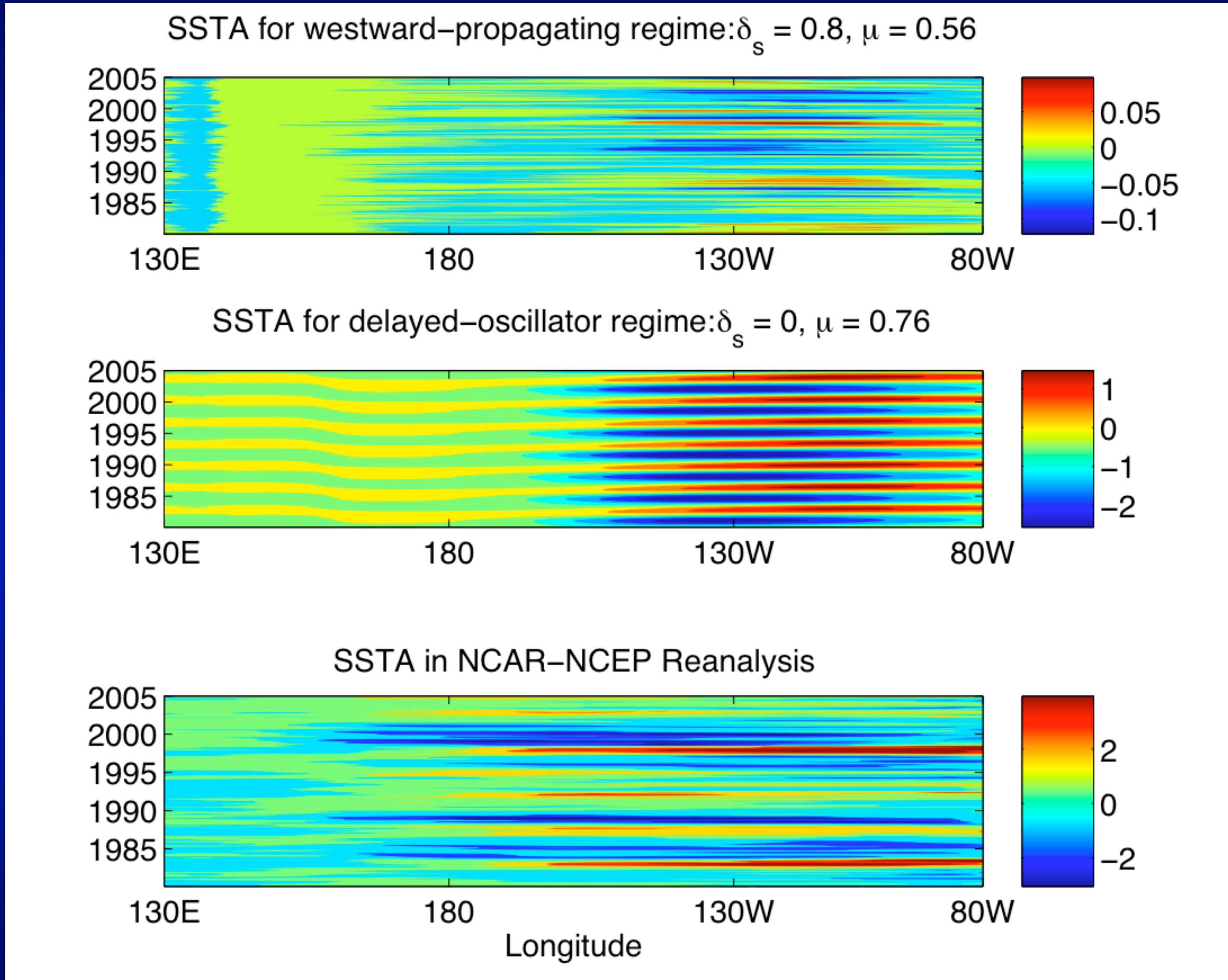
- Parameter estimation is always a **nonlinear problem,** even if the model is **linear** in terms of the model state: use **Extended Kalman Filter (EKF).**

Parameter estimation for coupled O-A system

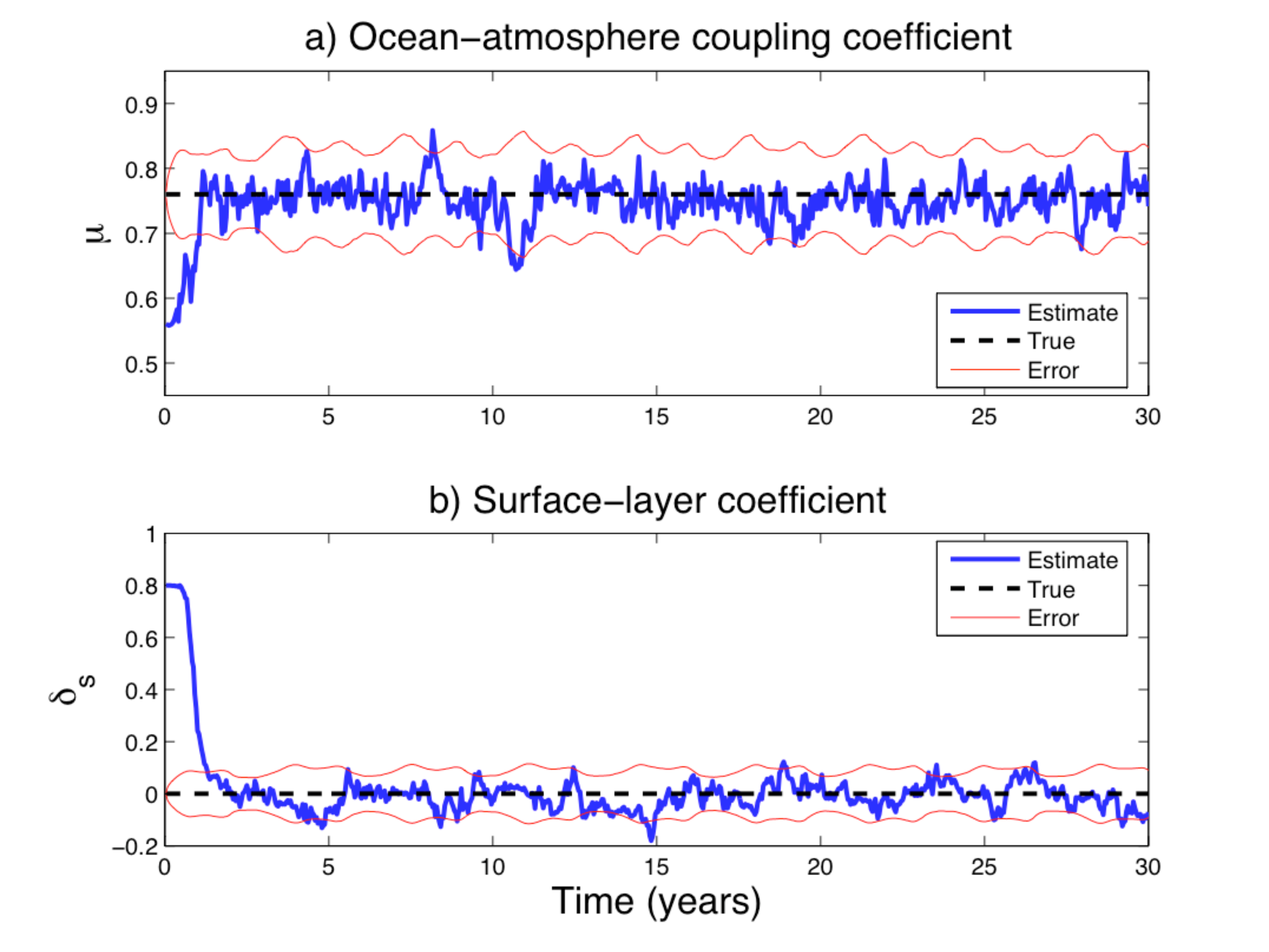
- Intermediate coupled model (ICM: Jin & Neelin, *JAS*, 1993)
- Estimate the state vector $W = (T', h, u, v)$, along with the coupling parameter μ and surface-layer coefficient δ_s by assimilating data from a single meridional section.
- The ICM model has errors in its initial state, in the wind stress forcing & in the parameters.
- M. Ghil (1997, *JMSJ*); Hao & Ghil (1995, *Proc. WMO Symp. DA Tokyo*); Sun *et al.* (2002, *MWR*).
- *Current work with D. Kondrashov, J.D. Neelin, & C.-j. Sun.*



Coupled O-A Model (ICM) vs. Observations

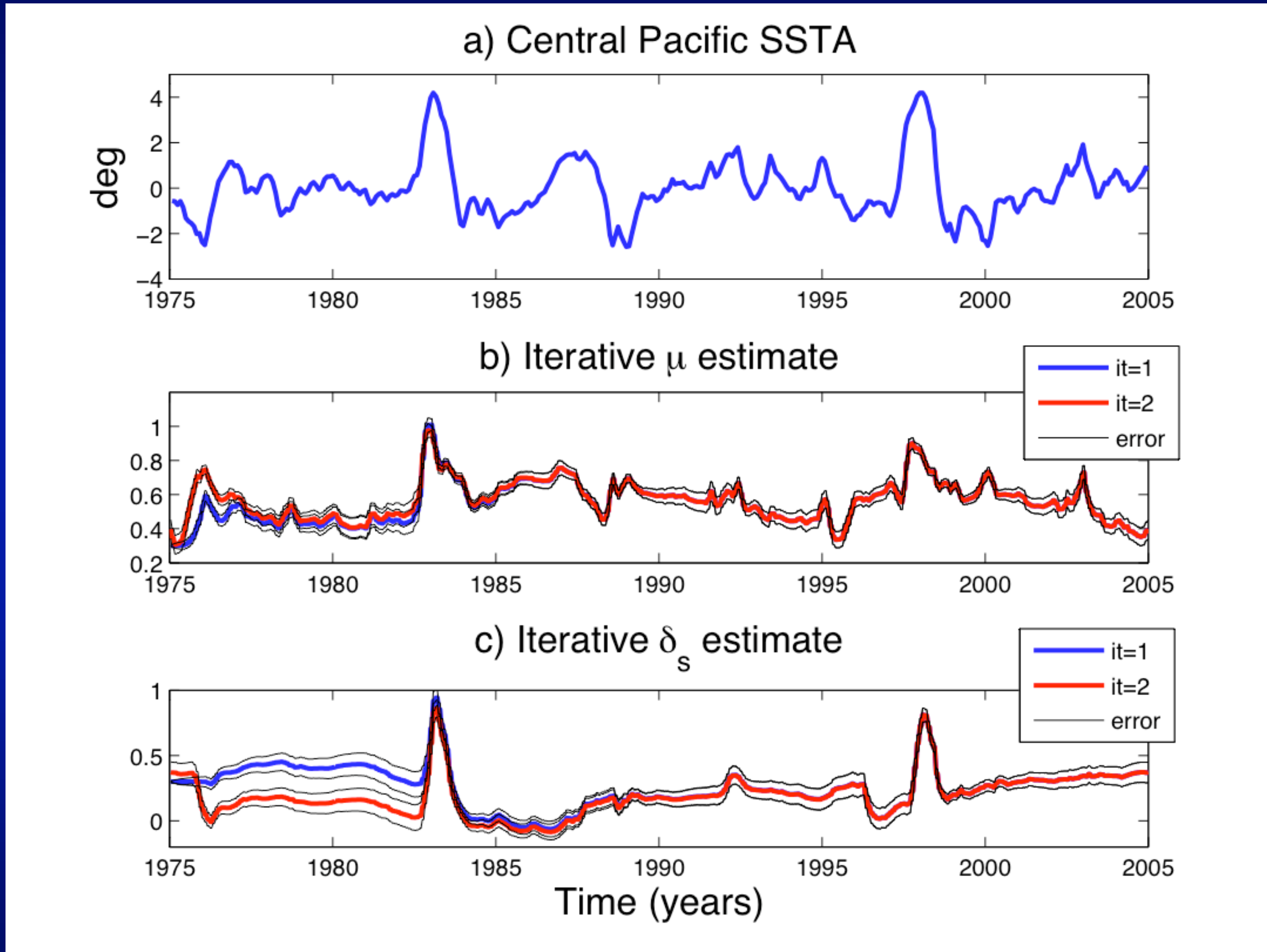


Convergence of Parameter Values – I



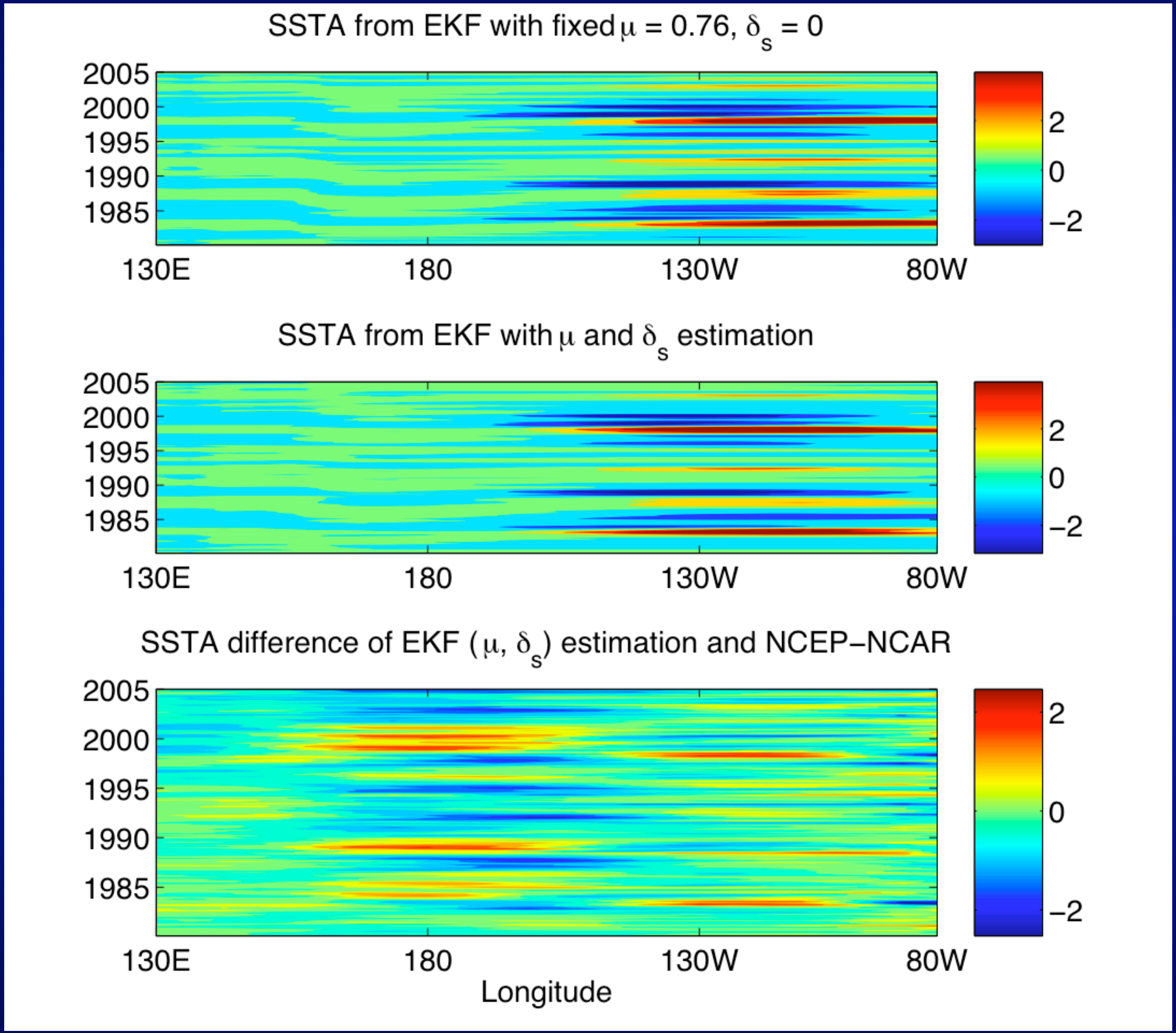
Identical-twin experiments

Convergence of Parameter Values – II



Real SSTA data

EKF results with and w/o parameter estimation



Computational advances

a) Hardware

- more computing power (CPU throughput)
- larger & faster memory (3-tier)

b) Software

- better numerical implementations of algorithms
- automatic adjoints
- block-banded, reduced-rank & other sparse-matrix algorithms
- better ensemble filters
- efficient parallelization, ….

How much DA *vs*. forecast?

- Design integrated observing–forecast–assimilation systems!

Observing system design

- **Need no more (independent) observations than *d-o-f* to be tracked:**
  - "features" (Ide & Ghil, 1997a, b, *DAO*);
  - instabilities (Todling & Ghil, 1994 + Ghil & Todling, 1996, *MWR*);
  - trade-off between mass & velocity field (Jiang & Ghil, *JPO*, 1993).
- The cost of **advanced** DA is **much** less than that of instruments & platforms:
  - at best use DA **instead** of instruments & platforms.
  - at worst use DA to determine which instruments & platforms (**advanced** OSSE)
- **Use any** observations, if forward modeling is possible (observing operator H)
  - satellite images, 4-D observations;
  - pattern recognition in observations and in phase-space statistics.

Conclusion

- No observing system without data assimilation and no assimilation without dynamics[a]
- Quote of the day: "You cannot step into the same river[b] twice[c]"
  (Heracleitus*, Trans. Basil. Phil. Soc. Miletus*, *cca.* 500 B.C.)

[a]of state and errors
[b]Meandros
[c] "You cannot do so even once" (subsequent development of "flux" theory by Plato, c*ca.* 400 B.C.)
Τα παντα ρεει = Everything flows

Evolution of DA – I

TABLE I. CHARACTERISTICS OF DATA ASSIMILATION SCHEMES IN OPERATIONAL USE AT THE END OF THE 1970s[a]

Organization or country	Operational analysis methods	Analysis area	Analysis/forecast
Australia	Successive correction method (SCM)	SH[d]	12 hr
	Variational blending techniques	Regional	6 hr
Canada	Multivariate 3-D statistical interpolation	NH[d] Regional	6 hr (3 hr for the surface)
France	SCM; wind-field and mass-field balance through first guess	NH	6 hr
	Multivariate 3-D statistical interpolation	Regional	
F.R. Germany	SCM. Upper-air analyses were built up, level by level, from the surface	NH	12 hr (6 hr for the surface)
	Variational height/wind adjustment		Climatology only as preliminary fields
Japan	SCM	NH	12 hr
	Height-field analyses were corrected by wind analyses	Regional	
Sweden	Univariate 3-D statistical interpolation	NH	12 hr
	Variational height/wind adjustment	Regional	3 hr
United Kingdom	Hemispheric orthogonal polynomial method		
	Univariate statistical interpolation (repeated insertion of data)	Global	6 hr
U.S.A.	Spectral 3-D analysis	Global	
	Multivariate 3-D statistical interpolation	Global	6 hr
U.S.S.R.	2-D[c] statistical interpolation	NH	12 hr
ECMWF[b]	Multivariate 3-D statistical interpolation	Global	6 hr

[a] After Gustafsson (1981).
[b] European Centre for Medium Range Weather Forecasts.
[c] 2-D is in a horizontal plane.
[d] Southern Hemisphere and Northern Hemisphere, respectively.

Transition from "early" to "mature" phase of DA in NWP:

– no Kalman filter (Ghil *et al.*, 1981(*))

– no adjoint (Lewis & Derber, *Tellus,* 1985); Le Dimet & Talagrand (*Tellus,* 1986)

(*) Bengtsson, Ghil & Källén (Eds., 1981), *Dynamic Meteorology: Data Assimilation Methods.*

M. Ghil & P. M.-Rizzoli (*Adv. Geophys.,* 1991).

Evolution of DA – II

TABLE IV. DUALITY RELATIONSHIPS BETWEEN STOCHASTIC ESTIMATION AND DETERMINISTIC CONTROL[a]

A. Continuous (linear) Kalman Filter

System Model	$\dot{\mathbf{w}}^t(t) = F(t)\mathbf{w}^t(t) + G(t)\mathbf{b}^t(t), \quad \mathbf{b}^t(t) \sim \mathrm{N}[0, Q(t)]$
Measurement Model	$\mathbf{w}^0(t) = H(t)\mathbf{w}^t(t) + \mathbf{b}^0(t), \quad \mathbf{b}^0(t) \sim \mathrm{N}[0, R(t)]$
State estimation	$\dot{\mathbf{w}}^a(t) = F(t)\mathbf{w}^a(t) + K(t)[\mathbf{w}^0(t) - H(t)\mathbf{w}^a(t)], \quad \mathbf{w}^a(0) = \mathbf{w}^a_0$
Error covariance propagation (Riccati Equation)	$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) - K(t)R(t)K^T(t), \quad P(0) = P_0$
Kalman Gain	$K(t) = P(t)H^T(t)R^{-1}(t)$
Initial conditions	$E[\mathbf{w}^t(0)] = \mathbf{w}^a_0, \quad E\{[\mathbf{w}^t(0) - \mathbf{w}^a_0][\mathbf{w}^t(0) - \mathbf{w}^a_0]^T\} = P_0$
Assumptions	$R^{-1}(t)$ exists $E\{\mathbf{b}^t(t)[\mathbf{b}^0(t')]^T\} = 0$
Performance Index	$p^{f,a}(t) = E\{[\mathbf{w}^{f,a} - \mathbf{w}^t][\mathbf{w}^{f,a} - \mathbf{w}^t]^T\}$

B. Continuous (linear) Optimal Control

System Model	$\dot{\mathbf{w}}^t(t) = \tilde{F}(t)\mathbf{w}(t) + \tilde{H}(t)\mathbf{u}(t)$
Measurement Model	$\mathbf{w}^0(t) = \mathbf{w}(t)$ (all system variables are measured)
Performing control	$\mathbf{u}(t) = -\tilde{K}(t)\mathbf{w}(t)$
Performance propagation (Riccati Equation)	$\dot{\tilde{P}}(t) = -\tilde{F}^T(t)\tilde{P}(t) - \tilde{P}(t)\tilde{F}(t) - \tilde{Q}(t) + \tilde{P}(t)\tilde{H}(t)\tilde{K}(t)$
Control Gain	$\tilde{K}(t) = \tilde{R}^{-1}(t)\tilde{H}(t)\tilde{P}(t)$
Terminal conditions	$\mathbf{w}(t_f) = 0$ $\mathbf{P}(t_f) = \tilde{Q}_f$
Cost function	$J[\mathbf{w}, \mathbf{u}] = \mathbf{w}_f^T\tilde{Q}_f\mathbf{w}_f + \int_0^{t_f} [\mathbf{w}^T(t)\tilde{Q}(t)\mathbf{w}(t) + \mathbf{u}^T(t)\tilde{R}(t)\mathbf{u}(t)]\,dt$

C. Estimation-Control Duality

Estimation	Control
t_0 initial time	t_f final time
$\mathbf{w}(t)$ unobservable state variable of random process	$\mathbf{w}(t)$ observable state variable to be controlled
$\mathbf{w}^0(t)$ random observations	$\mathbf{u}(t)$ deterministic control
$F(t)$ dynamic matrix	$\tilde{F}^T(t)$ dynamic matrix
$Q(t)$ covariance matrix for the model errors	$\tilde{Q}(t)$ quadratic matrix defining acceptable errors on model variables
$H(t)$ effect of observations on state variables	$\tilde{H}(t)$ effect of control on state variables
$P(t)$ covariance of estimation error under optimization	$\tilde{P}(t)$ quadratic performance under optimization
$K(t)$ weighting on observation for optimal estimation	$\tilde{K}(t)$ weighting on state for optimal control

[a] (A), Kalman filter as the optimal solution for the former problem; (B), optimal solution for the latter problem; (C), equivalences between the two (after Kalman, 1960, and Gelb, 1974, Section 9.5; courtesy of R. Todling).

Cautionary note:

"Pantheistic" view of DA:

- variational ~ KF;
- 3- & 4-D Var ~ 3- & 4-D PSAS.

Fashionable to claim it's all the same but it's not:

- **God** is in **everything,**
- **but the devil is in the details.**

M. Ghil & P. M.-Rizzoli (*Adv. Geophys.*, 1991).

The DA Maturity Index of a Field

- **Pre-DA:** few data, poor models
  - The theoretician: Science is truth, don't bother me with the facts!
  - The observer/experimentalist: Don't ruin my beautiful data with your lousy model!!
- **Early DA:**
  - Better data, so-so models.
  - Stick it (the obs'ns) in – direct insertion, nudging.
- **Advanced DA:**
  - Plenty of data, fine models.
  - EKF, 4-D Var (2^{nd} duality).
- **Post-industrial DA:**

  (Satellite) images --> (weather) forecasts, climate "movies" …

General references

Bengtsson, L., M. Ghil and E. Källén (Eds.), 1981. *Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, 330 pp.

Daley, R., 1991. *Atmospheric Data Analysis*. Cambridge Univ. Press, Cambridge, U.K., 460 pp.

Ghil, M., and P. Malanotte-Rizzoli, 1991. Data assimilation in meteorology and oceanography. *Adv. Geophys*., **33**, 141–266.

Bennett, A. F., 1992. *Inverse Methods in Physical Oceanography*. Cambridge Univ. Press, 346 pp.

Malanotte-Rizzoli, P. (Ed.), 1996. *Modern Approaches to Data Assimilation in Ocean Modeling*. Elsevier, Amsterdam, 455 pp.

Wunsch, C., 1996. *The Ocean Circulation Inverse Problem*. Cambridge Univ. Press, 442 pp.

Ghil, M., K. Ide, A. F. Bennett, P. Courtier, M. Kimoto, and N. Sato (Eds.), 1997. *Data Assimilation in Meteorology and Oceanography: Theory and Practice*, Meteorological Society of Japan and Universal Academy Press, Tokyo, 496 pp.

Perec, G., 1969: *La Disparition*, Gallimard,Paris.

Parameter Estimation

a) Dynamical model

$dx/dt = M(x, \mu) + \eta(t)$

$y^o = H(x) + \varepsilon(t)$

Simple (EKF) idea – augmented state vector

$d\mu/dt = 0$, $X = (x^T, \mu^T)^T$

b) Statistical model

$L(\rho)\eta = w(t)$, L – AR(MA) model, $\rho = (\rho_1, \rho_2, ... \rho_M)$

Examples: 1) Dee *et al*. (*IEEE*, 1985) – estimate a few parameters in the covariance matrix $Q = E(\eta, \eta^T)$; also the bias $<\eta> = E\eta$;

2) POPs - Hasselmann (1982, Tellus); Penland (1989, *MWR*; 1996, *Physica D*); Penland & Ghil (1993, *MWR*)

3) $dx/dt = M(x, \mu) + \eta$: Estimate both M & Q from data (Dee, 1995, *QJ*), Nonlinear approach: Empirical mode reduction (Kravtsov *et al*., 2005, Kondrashov *et al*., 2005)