Backward Fokker-Planck equation for evaluating model predictability

Peter C Chu
Naval Postgraduate School
Monterey, CA 93943
pcchu@nps.edu,
https://faculty.nps.edu/pcchu

Meteorological Institute of Belgium, Brussel, Belgium, 16 March 2004.

References

- Chu, P.C., L.M. Ivanov, T. M. Margolina, and O.V. Melnichenko, On probabilistic stability of an atmospheric model to various amplitude perturbations. Journal of the Atmospheric Sciences, 59, 2860-2873.
- Chu, P.C., L.M. Ivanov, C.W. Fan, 2002: Backward Fokker-Planck equation for determining model valid prediction period. Journal of Geophysical Research, 107, C6, 10.1029/2001JC000879.
- Chu, P.C., L. Ivanov, L. Kantha, O. Melnichenko, and Y. Poberezhny, 2002: Power law decay in model predictability skill. Geophysical Research Letters, 29 (15), 10.1029/2002GLO14891.
- Chu, P.C., L.M. Ivanov, L.H. Kantha, T.M. Margolina, and O.M. Melnichenko, and Y.A, Poberenzhny, 2004: Lagrangian predictabilty of high-resolution regional ocean models. Nonlinear Processes in Geophysics, 11, 47-66.

This question should be answered before running any model

- How long is an ocean (or atmospheric) model valid once being integrated from its initial state?

Physical Reality

- $\boldsymbol{Y}$
- Physical Law: $d\boldsymbol{Y}/dt = \boldsymbol{h}(\boldsymbol{y}, t)$
- Initial Condition: $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$

Atmospheric & Oceanic Models

- ***X*** is the prediction of **Y**
- $d\boldsymbol{X}/dt = \boldsymbol{f}(\boldsymbol{X}, t) + q(t)\,\boldsymbol{X}$
- Initial Condition: $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$
- Stochastic Forcing:
- $\langle q(t) \rangle = 0$
- $\langle q(t)q(t') \rangle = q^2\delta(t-t')$

Model Error

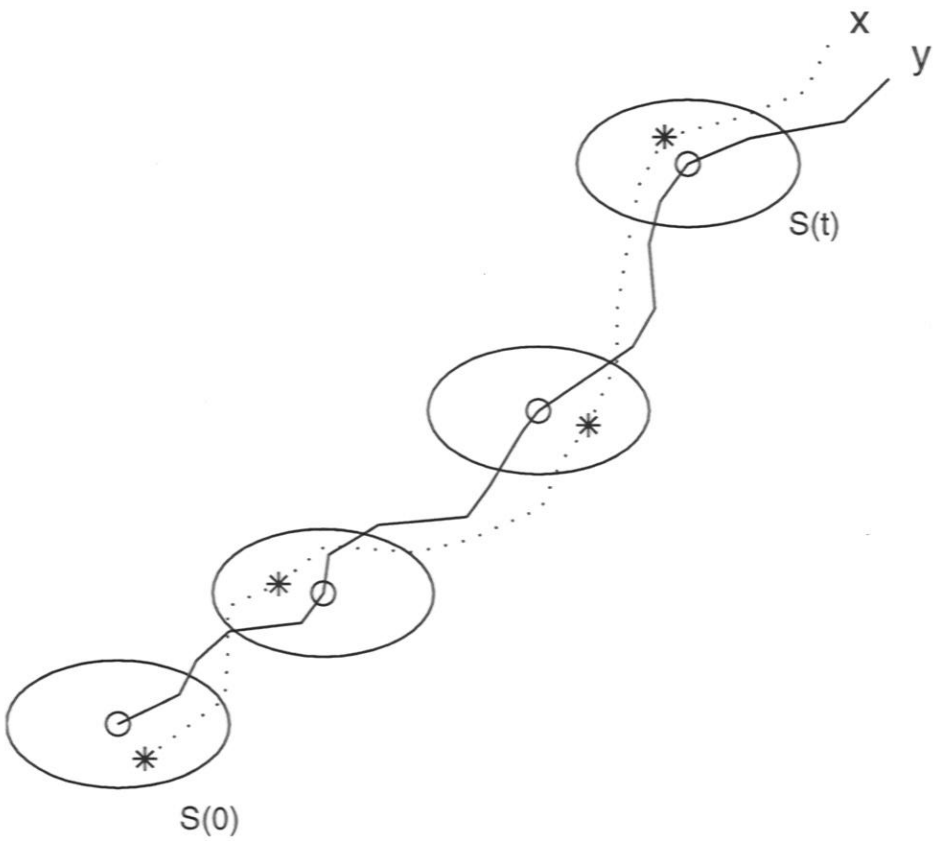
- $Z = X - Y$
- Initial: $Z_0 = X_0 - Y_0$

Valid Prediction Period (VPP)

- VPP is defined as the time period when the prediction error first exceeds a pre-determined criterion (i.e., the tolerance level ε).

- VPP is the First-Passage Time

VPP



Predictability Tube

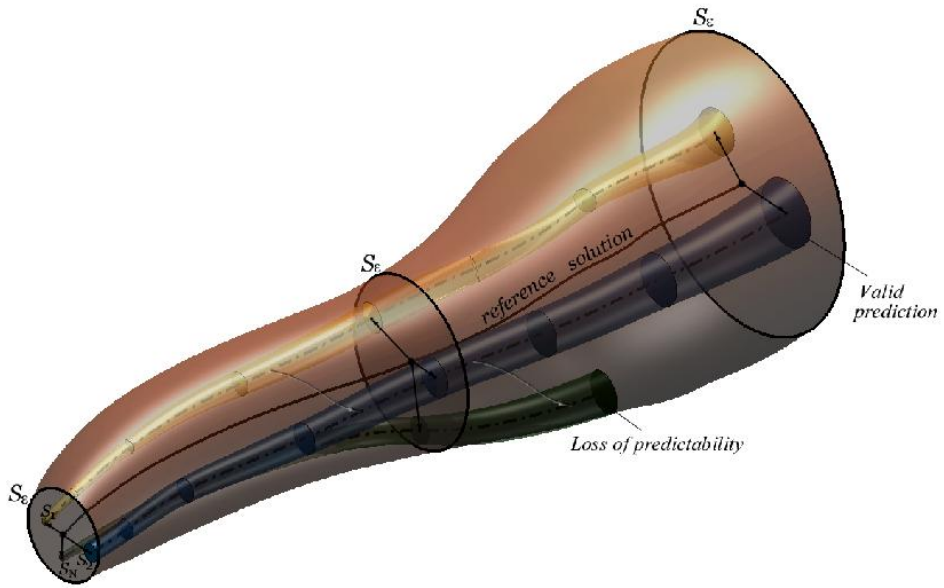
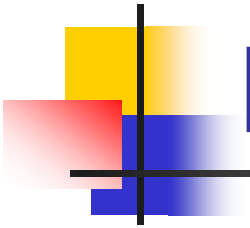


Fig. 31

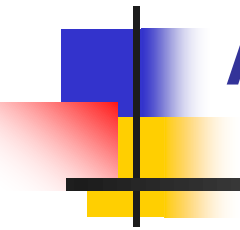
Conditional Probability Density Function

- Initial Error: z_0
- $(t - t_0)$? Random Variable
- Conditional PDF of $(t - t_0)$ with given z_0 ?
  - $P[(t - t_0) | z_0]$

Two Approaches to Obtain PDF of VPP



- Analytical (Backward Fokker-Planck Equation)
- Practical (Optimum Spectral Analysis)



Analytical Approach

Backward Fokker-Planck Equation

Backward Fokker-Planck Equation

$$\frac{\partial P}{\partial t} - \left[\mathbf{f}(\mathbf{z}_0, t)\right]\frac{\partial P}{\partial \mathbf{z}_0} - \frac{1}{2} q^2 \mathbf{z}_0{}^2 \frac{\partial^2 P}{\partial \mathbf{z}_0 \partial \mathbf{z}_0} = 0$$

Moments

$$\tau_1(\mathbf{z}_0) = \int_{t_0}^{\infty} P(t_0, \mathbf{z}_0, t - t_0)(t - t_0)dt$$

$$\tau_2(\mathbf{z}_0) = \int_{t_0}^{\infty} P(t_0, \mathbf{z}_0, t - t_0)(t - t_0)^2 dt$$

Example: One Dimensional Error (Nicolis 1992), Population (ecology), or General Production (economics) Models

- 1D Dynamical System (Maximum Growing Maniford of Lorenz System)

$$\frac{d\xi}{dt} = (\sigma - g\xi^2) + v(t)\xi, \qquad 0 \le \xi < \infty$$

$$\langle v(t) \rangle = 0, \qquad \langle v(t)v(t') \rangle = q^2 \delta(t - t').$$

$$\sigma = 0.64, \qquad g = 0.3, \qquad q^2 = 0.2.$$

Mean and Variance of VPP

$$(\sigma\xi_0 - g\xi_0^2)\frac{d\tau_1}{d\xi_0} + \frac{q^2\xi_0^2}{2}\frac{d^2\tau_1}{d\xi_0^2} = -1$$

$$(\sigma\xi_0 - g\xi_0^2)\frac{d\tau_2}{d\xi_0} + \frac{q^2\xi_0^2}{2}\frac{d^2\tau_2}{d\xi_0^2} = -2\tau_1$$

$$\tau_1 = 0, \qquad \tau_2 = 0 \qquad \text{for} \quad \xi_0 = \varepsilon .$$

$$\frac{d\tau_1}{d\xi_0} = 0, \qquad \frac{d\tau_2}{d\xi_0} = 0 \qquad \text{for} \quad \xi_0 = \xi_{\text{noise}}.$$

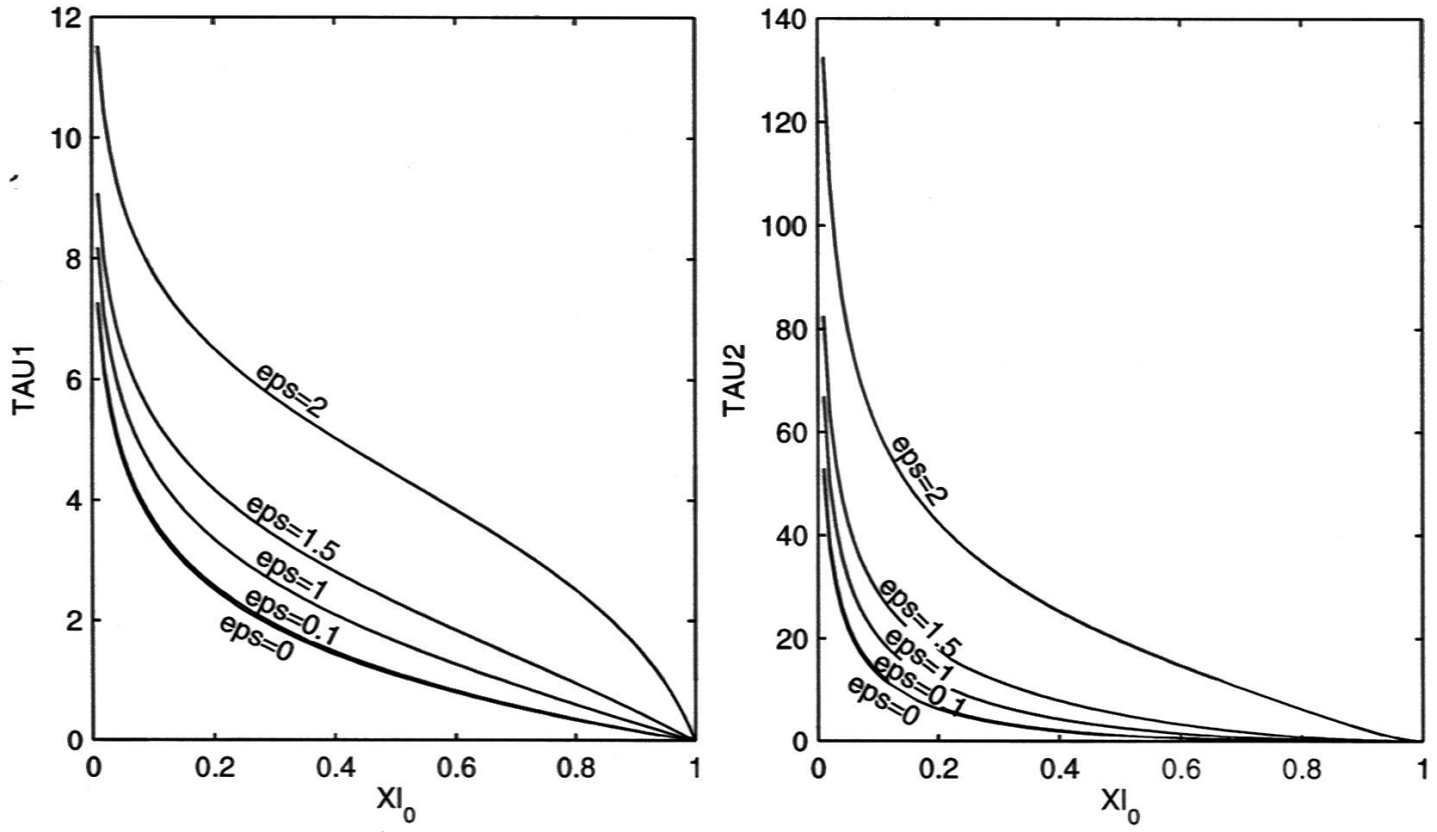
Analytical Solutions

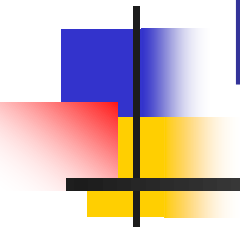
$$\tau_1(\bar{\xi}_0, \bar{\xi}_{noise}, \varepsilon) = \frac{2}{q^2} \int_{\bar{\xi}_0}^{1} y^{-\frac{2\sigma}{q^2}} \exp\left(\frac{2\varepsilon g}{q^2} y\right) \left[\int_{\bar{\xi}_{noise}}^{y} x^{\frac{2\sigma}{q^2}-2} \exp\left(-\frac{2\varepsilon g}{q^2} x\right) dx \right] dy$$

$$\tau_2(\bar{\xi}_0, \bar{\xi}_{noise}, \varepsilon) = \frac{4}{q^2} \int_{\bar{\xi}_0}^{1} y^{-\frac{2\sigma}{q^2}} \exp\left(\frac{2\varepsilon g}{q^2} y\right) \left[\int_{\bar{\xi}_{noise}}^{y} \tau_1(x) x^{\frac{2\sigma}{q^2}-2} \exp\left(-\frac{2\varepsilon g}{q^2} x\right) dx \right] dy$$

$$\bar{\xi}_0 = \xi_0 / \varepsilon, \qquad \bar{\xi}_{noise} = \xi_{noise} / \varepsilon$$

Dependence of tau1 & tau2 on Initial Condition Error (ξ_0/ε)





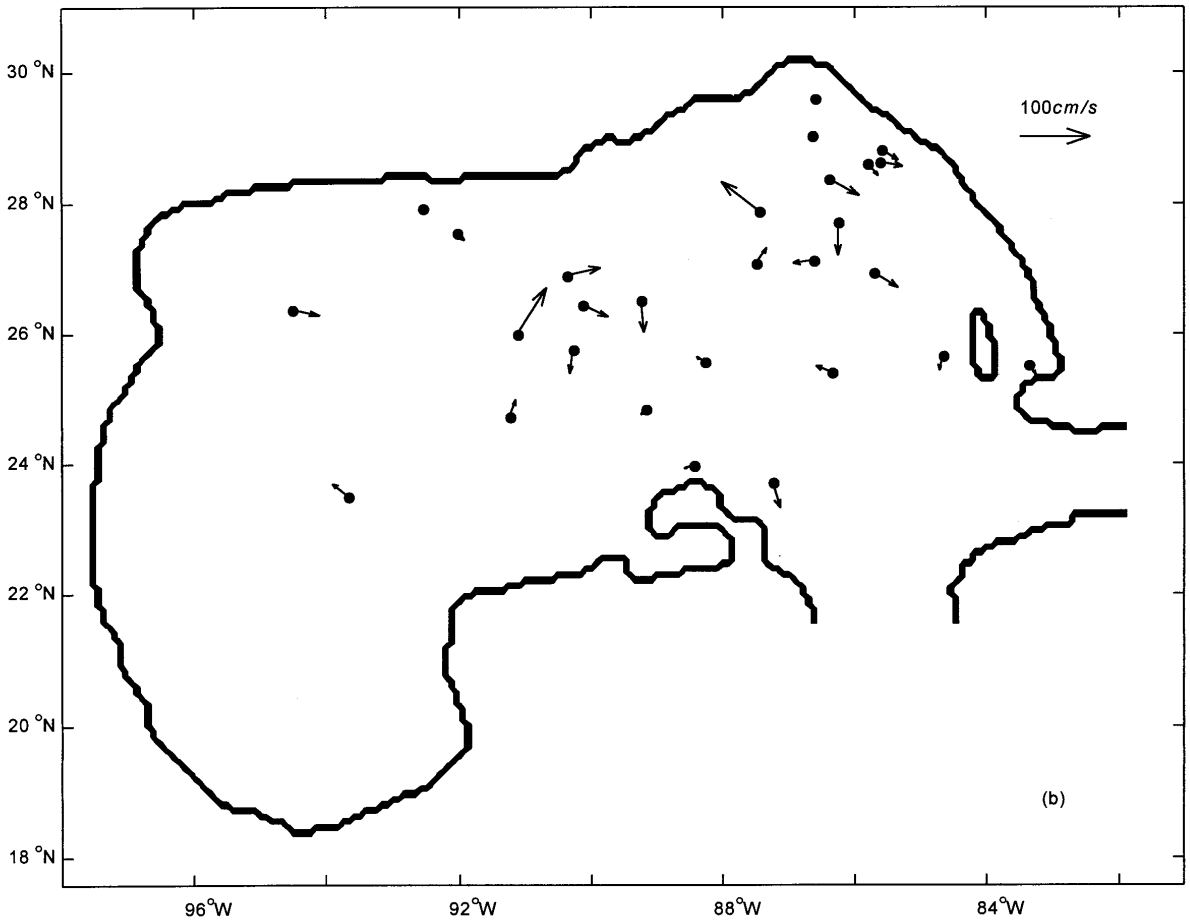
Practical Approach

Optimum Spectral Decomposition (OSD)

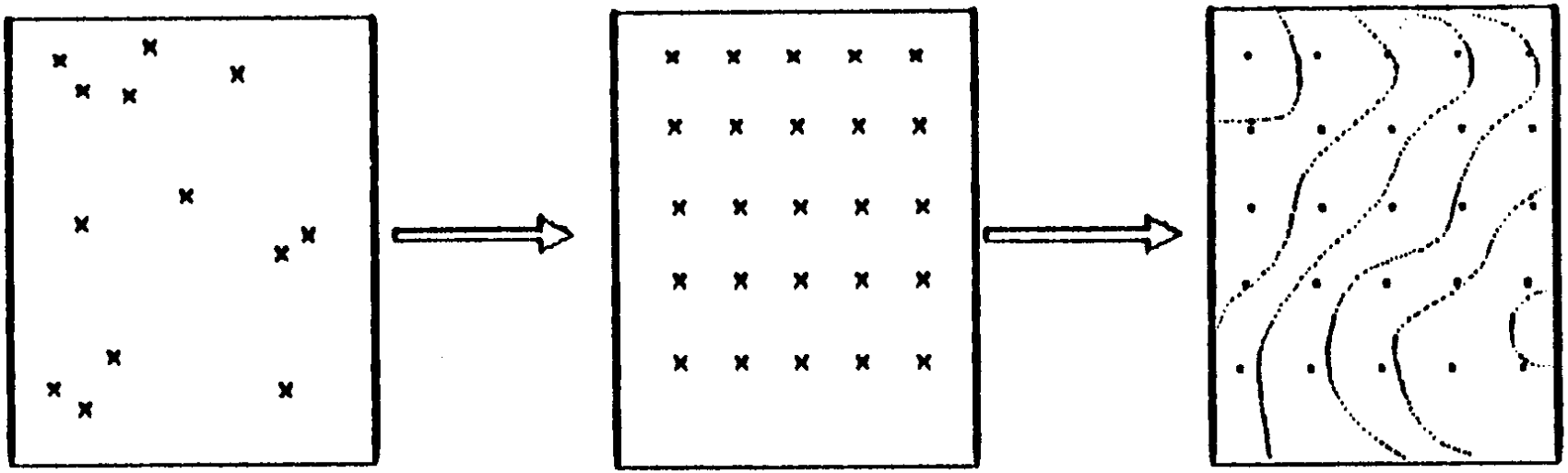
OSD Method

- Chu, P.C., L.M. Ivanov, T.P. Korzhova, T.M. Margolina, and O.M. Melnichenko, 2003a: Analysis of sparse and noisy ocean current data using flow decomposition. Part 1: Theory. Journal of Atmospheric and Oceanic Technology, 20 (4), 478-491.
- Chu, P.C., L.M. Ivanov, T.P. Korzhova, T.M. Margolina, and O.M. Melnichenko, 2003b: Analysis of sparse and noisy ocean current data using flow decomposition. Part 2: Application to Eulerian and Lagrangian data. Journal of Atmospheric and Oceanic Technology, 20 (4), 492-512.

Observational Data (Sparse and Noisy)



Most Popular Method for Data Analysis: Optimum Interpolation (OI)



Three Necessary Conditions For OI Method

- (1) Given first guess field
- (2) Given autocorrelation functions
- (3) Low noise-to-signal ratio

Ocean velocity data

- (1) First guess field (?)
- (2) Unknown autocorrelation function
- (3) High noise-to-signal ratio

OSD Method

$$c(\mathbf{x}, z_k, t) = A_0(z_k, t) + \sum_{m=1}^{M} A_m(z_k, t)\ \Psi_m(\mathbf{x}, z_k),$$

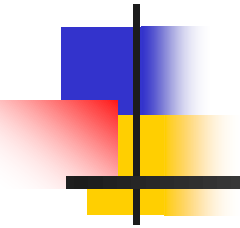
Two approaches to obtain basis functions

- EOF
- Eigenfunctions of Laplace Operator

$$\nabla_h^2 \Psi_m = -\lambda_m \Psi_m, \quad \Psi_m|_\Gamma = 0, \quad m = 1, 2, \ldots, M$$

Optimal Mode Truncation

$$J_{emp} = J(\tilde{A}_1,\ldots,\tilde{A}_M,P,M) = \frac{1}{P}\sum_{j=1}^{P}\left(c^{(j)} - \sum_{m=1}^{M}\tilde{A}_m(z,t)\Psi_m^{(j)}(\mathbf{x},z)\right)^2 \rightarrow \min .$$



OSD is a perfect method to process ocean velocity data

Flow Decomposition

- 2 D Flow (Helmholtz)

$$\mathbf{u}_H = \mathbf{r} \times \nabla_H A_1 + \nabla_H A_3$$

- 3D Flow (Toroidal & Poloidal): Very popular in astrophysics

$$\mathbf{u} = \mathbf{r} \times \nabla A_1 + \mathbf{r} A_2 + \nabla A_3$$

3D Incompressible Flow

-
- When $\nabla \cdot \boldsymbol{u} = 0$
- We have

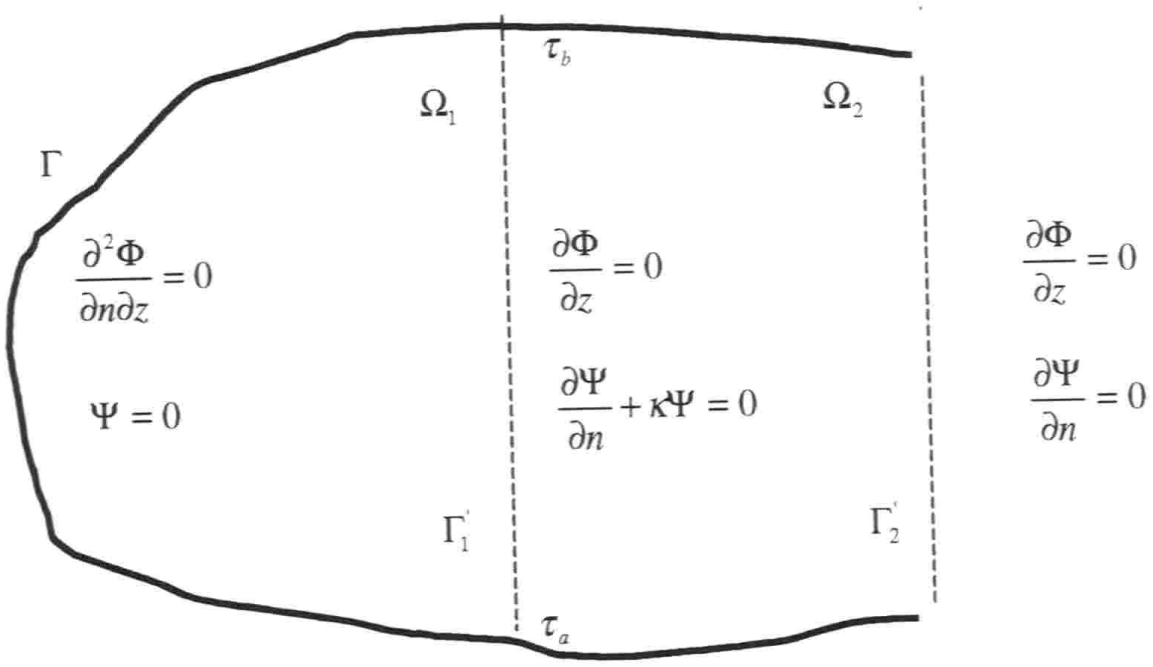
$$\mathbf{u} = \boldsymbol{\nabla} \times (\mathbf{r}\Psi) + \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times (\mathbf{r}\Phi).$$

Flow Decomposition

$$u = \frac{\partial \Psi}{\partial y} + \frac{\partial^2 \Phi}{\partial x \partial z}, \qquad v = -\frac{\partial \Psi}{\partial x} + \frac{\partial^2 \Phi}{\partial y \partial z},$$

- $\nabla^2 \Psi = -\zeta$,
- ζ *is relative vorticity*
- $\nabla^2 \Phi = -w$

Boundary Conditions



Basis Functions

$$\Psi(x, y, z, t^\circ) = \sum_{k=1}^{\infty} a_k(z, t^\circ)\Psi_k(x, y, z, \kappa^\circ),$$

$$\frac{\partial \Phi(x, y, z, t^\circ)}{\partial z} = \sum_{m=1}^{\infty} b_m(z, t^\circ)\Phi_m(x, y, z),$$

Flow Reconstruction

$$u_{KM} = \sum_{k=1}^{K} a_k(z,t^\circ)\frac{\partial \Psi_k(x,y,z,\kappa^\circ)}{\partial y} + \sum_{m=1}^{M} b_m(z,t^\circ)\frac{\partial \Phi_m(x,y,z)}{\partial x},$$

$$v_{KM} = -\sum_{k=1}^{K} a_k(z,t^\circ)\frac{\partial \Psi_k(x,y,z,\kappa^\circ)}{\partial x} + \sum_{m=1}^{M} b_m(z,t^\circ)\frac{\partial \Phi_m(x,y,z)}{\partial y}$$

Several Comments on OSD Method

- Reconstruction is a useful tool for processing real-time velocity data with short duration and limited-area sampling.
- The scheme can handle highly noisy data.
- The scheme is model independent.
- The scheme can be used for velocity data assimilation.

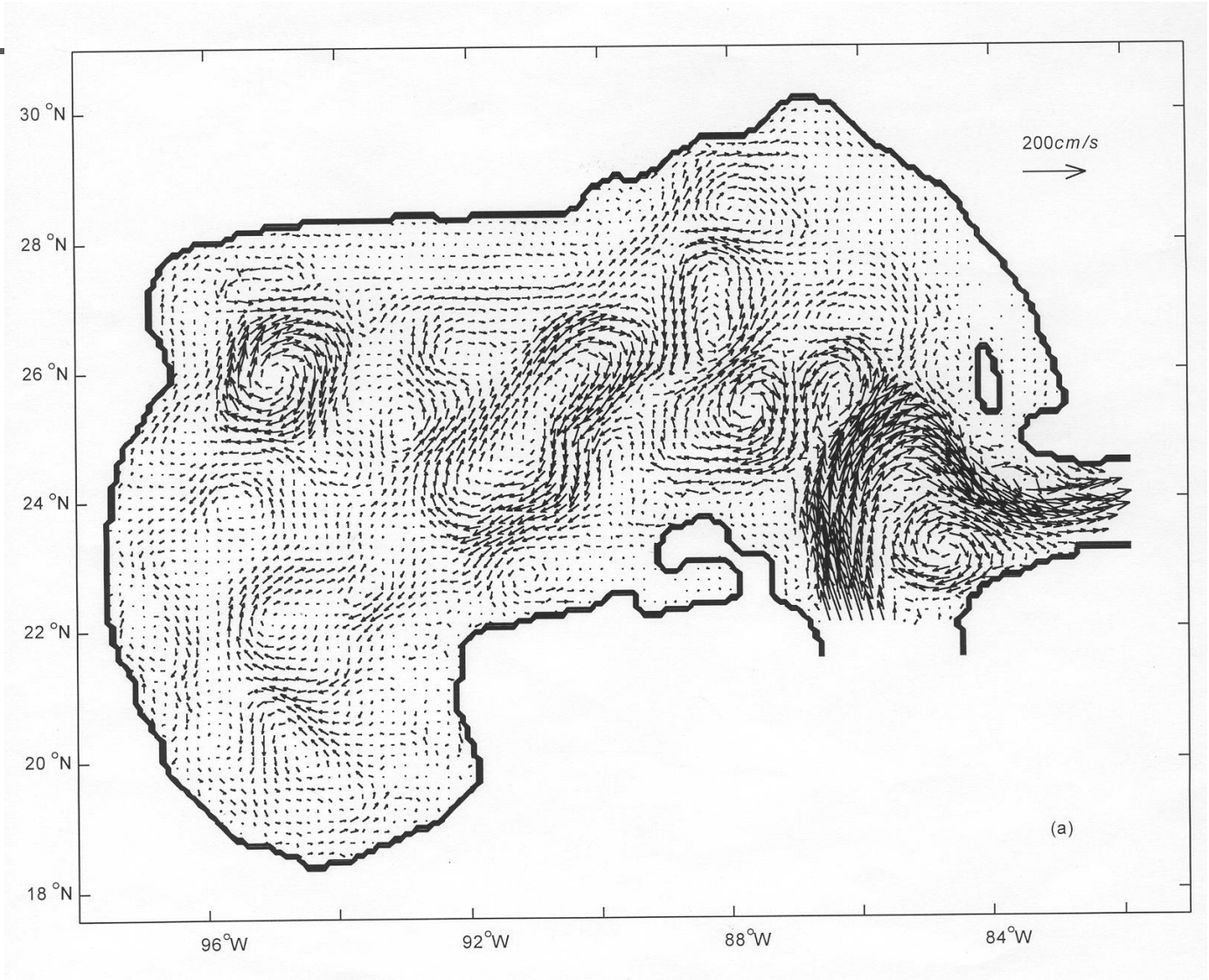
Gulf of Mexico Prediction System

Example for Practical Approach

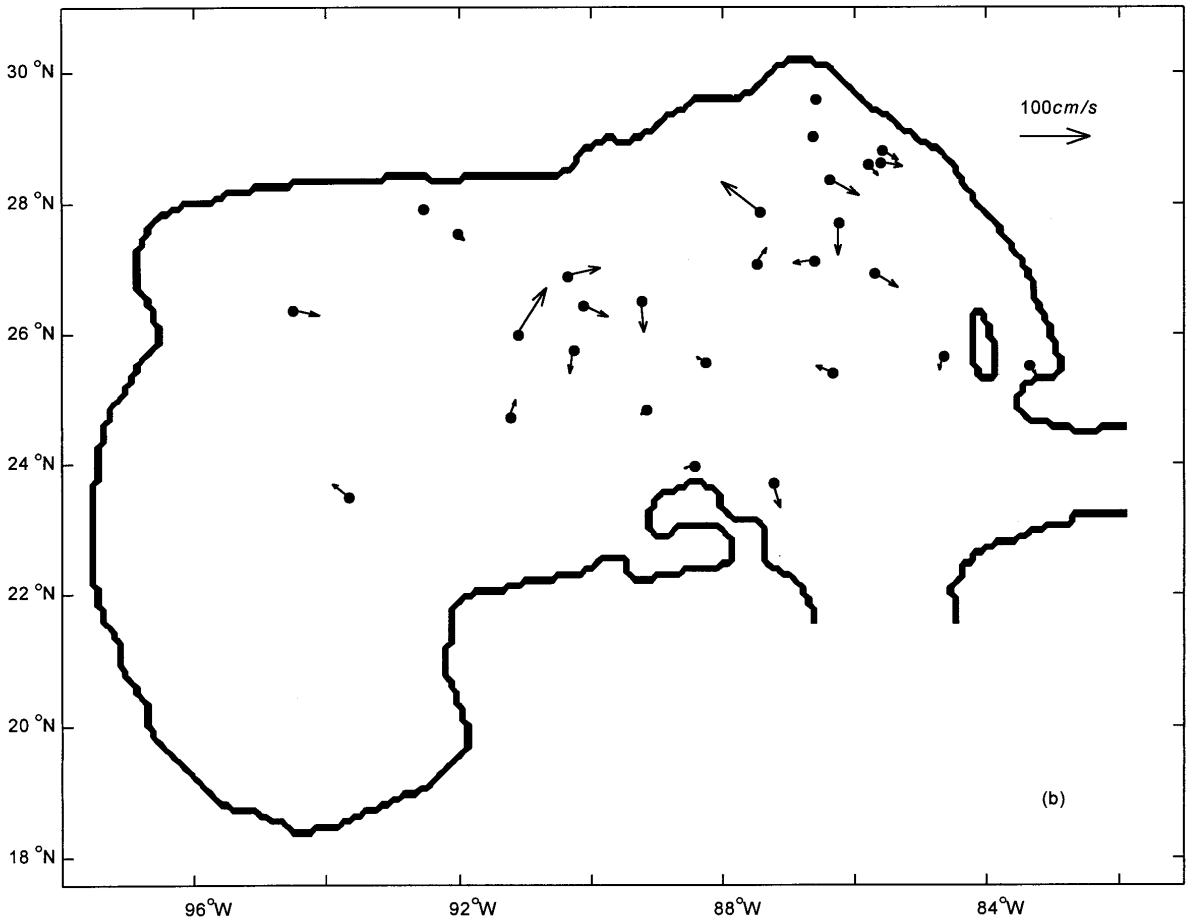
Gulf of Mexico Forecast System

- University of Colorado Version of POM
- 1/12^{o} Resolution
- Real-Time SSH Data (TOPEX, ESA ERS-1/2) Assimilated
- Real Time SST Data (MCSST, NOAA AVHRR) Assimilated
- Six Months Four-Times Daily Data From July 9, 1998 for Verification

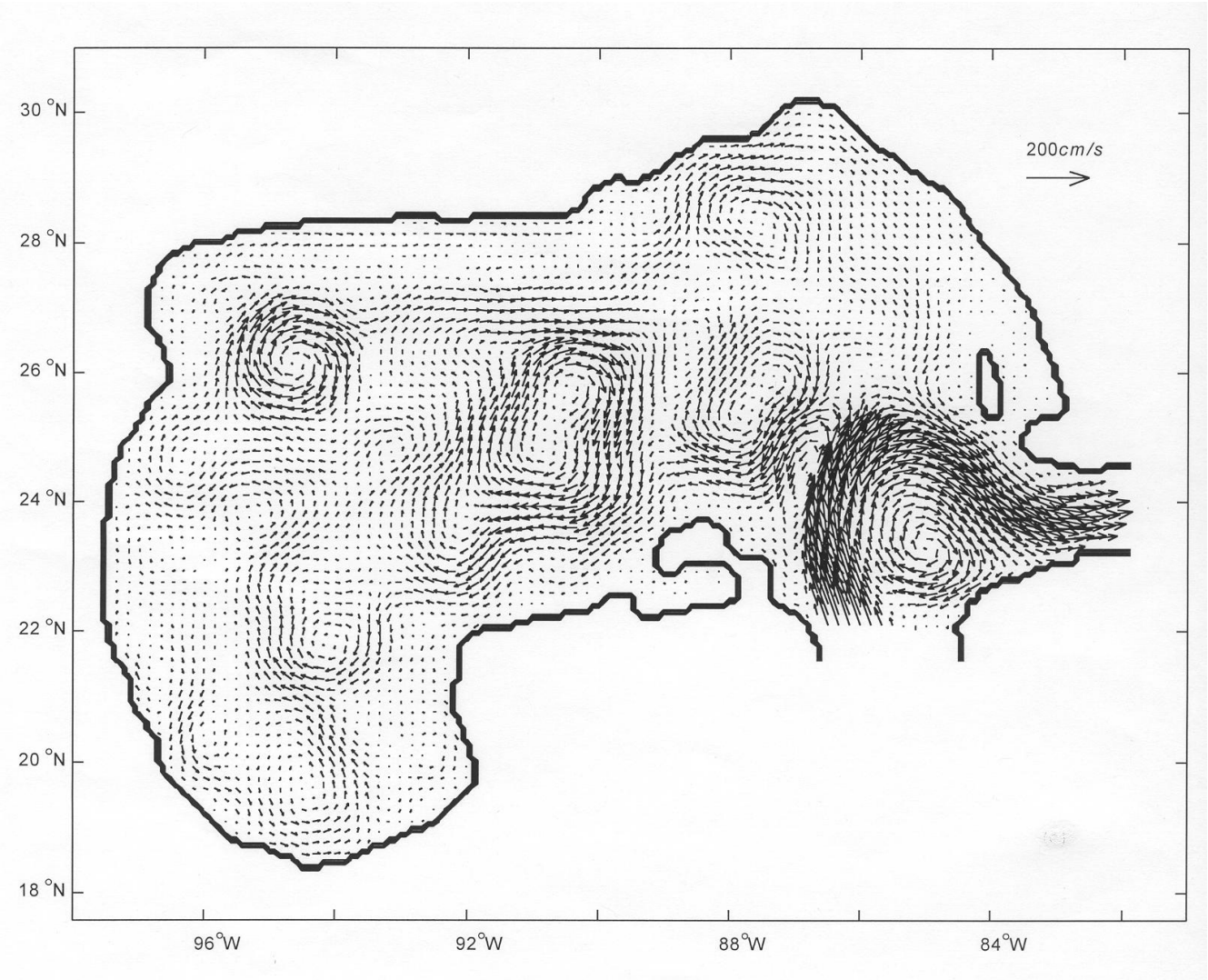
Model Generated Velocity Vectors at 50 m on 00:00 July 9, 1998



(Observational) Drifter Data at 50 m on 00:00 July 9, 1998



Reconstructed Drift Data at 50 m on 00:00 July 9, 1998 Using the OSD Method (Chu et al. 2002 a, b, JTECH)



Error Mean and Variance

Error Mean $$L_1 = \langle \mathbf{z} \rangle$$

Error Variance $$L_2 = \left\langle \left(\mathbf{z} - \langle \mathbf{z} \rangle\right)^t \left(\mathbf{z} - \langle \mathbf{z} \rangle\right) \right\rangle$$

Exponential Error Growth

$$L_1 \propto e^{\sigma t}, \qquad L_2 \propto e^{\omega t},$$

Classical Linear Theory

No Long-Term Predictability

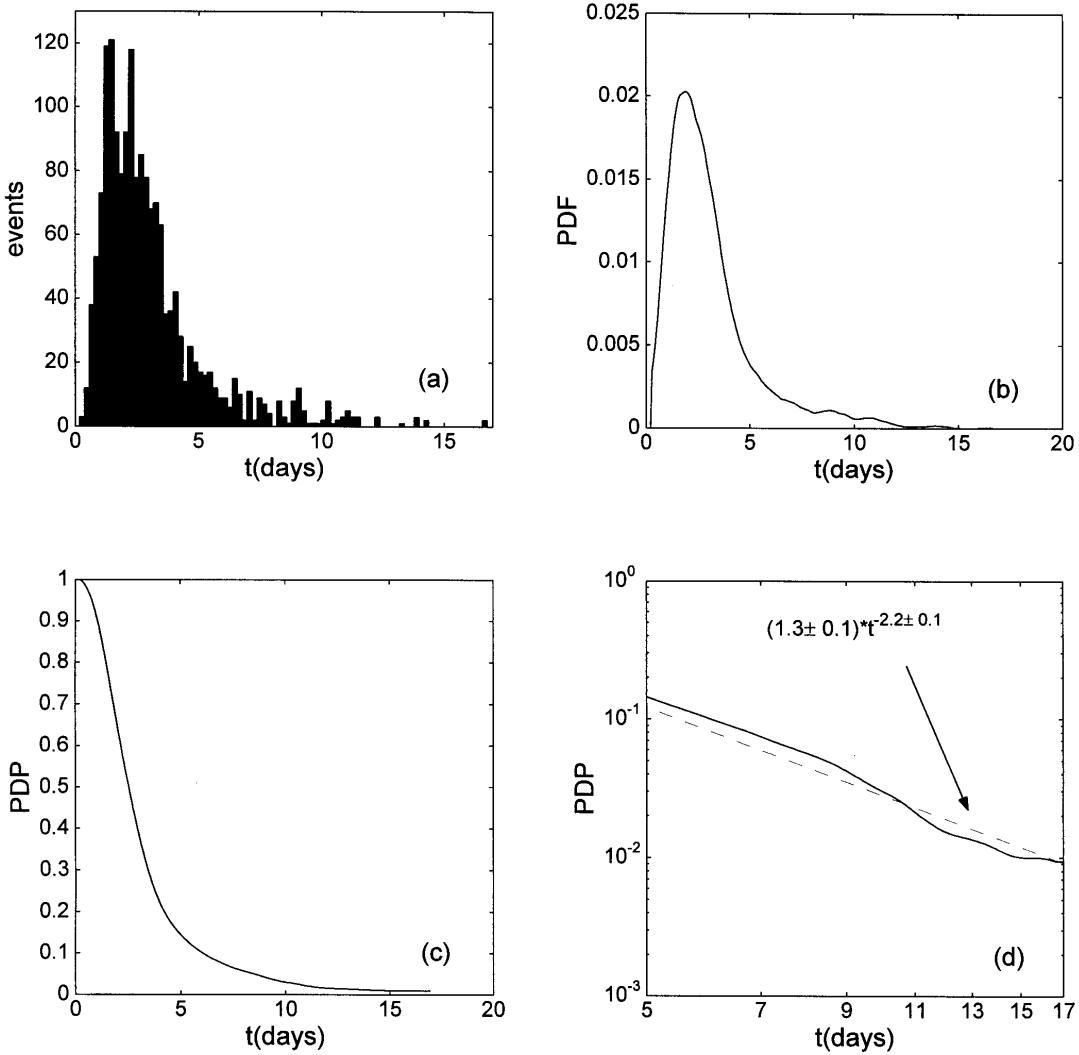
Power Law

$$L_1 \propto t^{\alpha}, \qquad L_2 \propto t^{\beta},$$

$$P(t_0, \mathbf{z}_0, \varepsilon, t - t_0) \sim t^{-\gamma} \quad \text{for large } t.$$

Long-Term Predictability May Occur

Statistical Characteristics of VPP for zero initial error and 55 km tolerance level (Non-Gaussion)

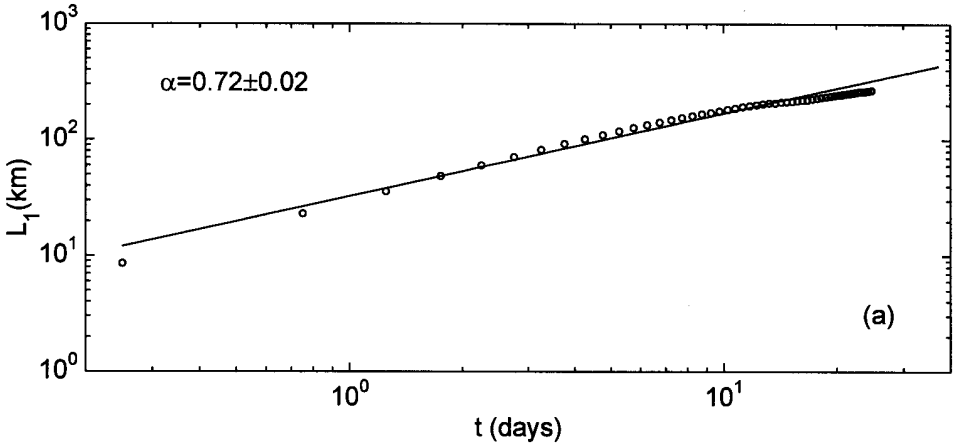


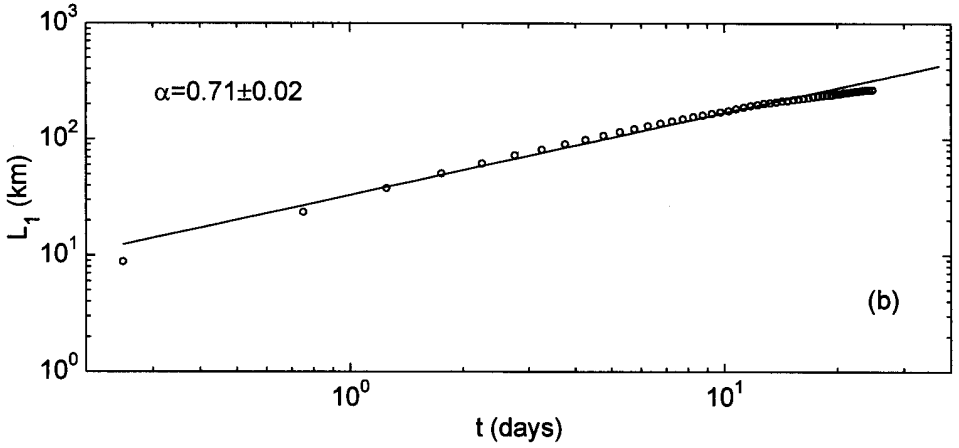
Scaling behavior of the mean error (L_1) growth for initial error levels:

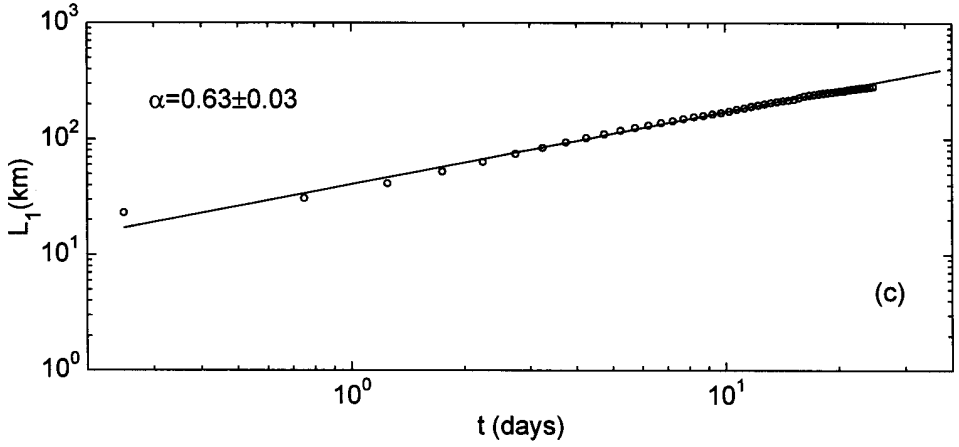
(a) 0

(b) 2.2 km

(c) 22 km







Scaling behavior of the

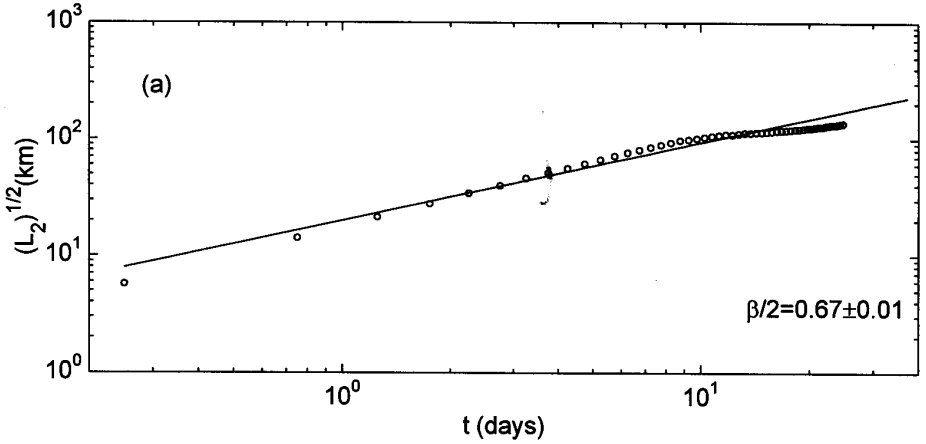
Error variance (L_2) growth

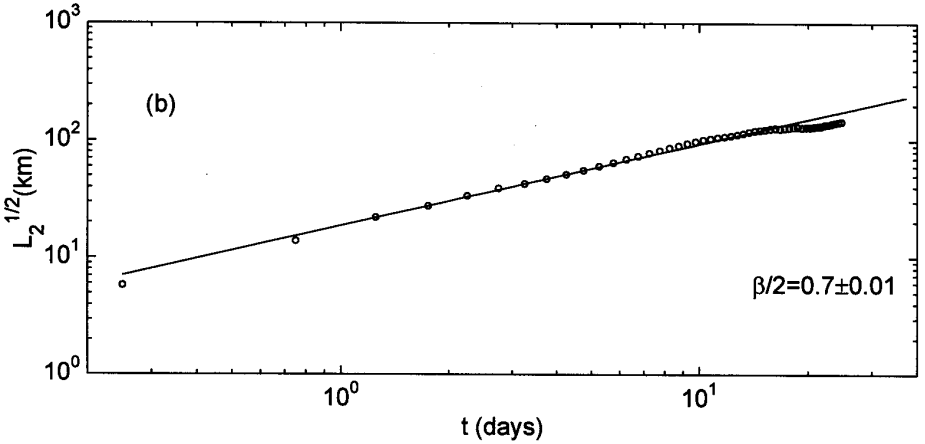
for initial error levels:

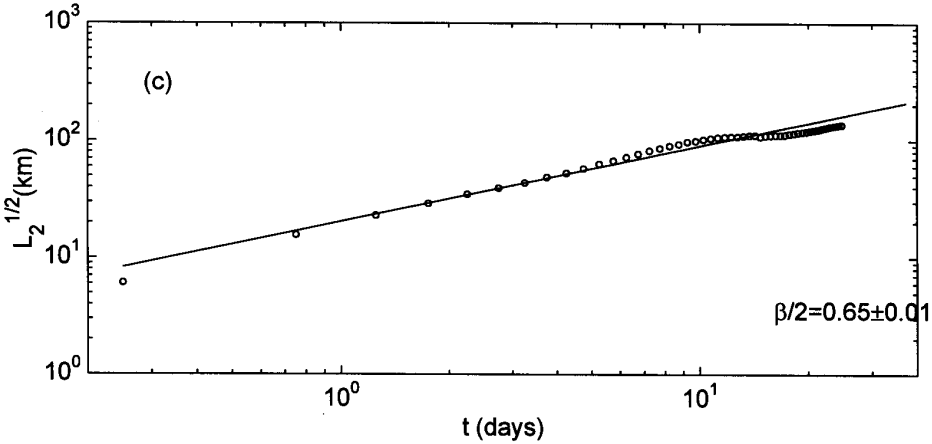
(a) 0

(b) 2.2 km

(c) 22 km

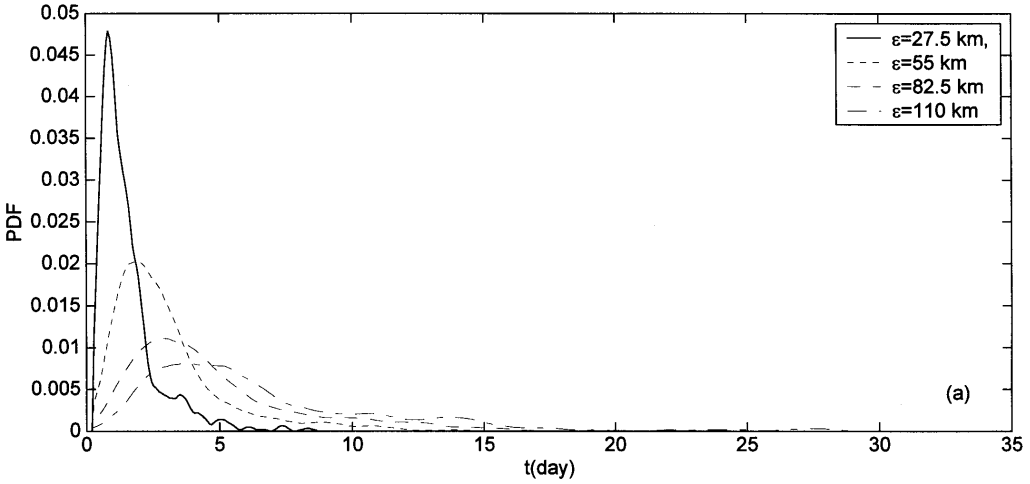


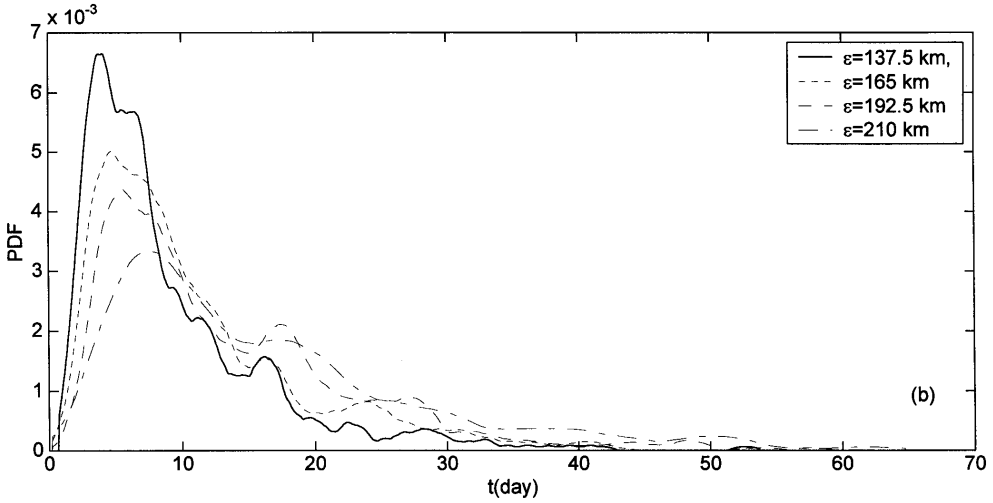




Probability Density Function of VPP calculated with different tolerance levels

Non-Gaussian distribution with long tail toward large values of VPP (Long-term Predictability)





Predictability Tube

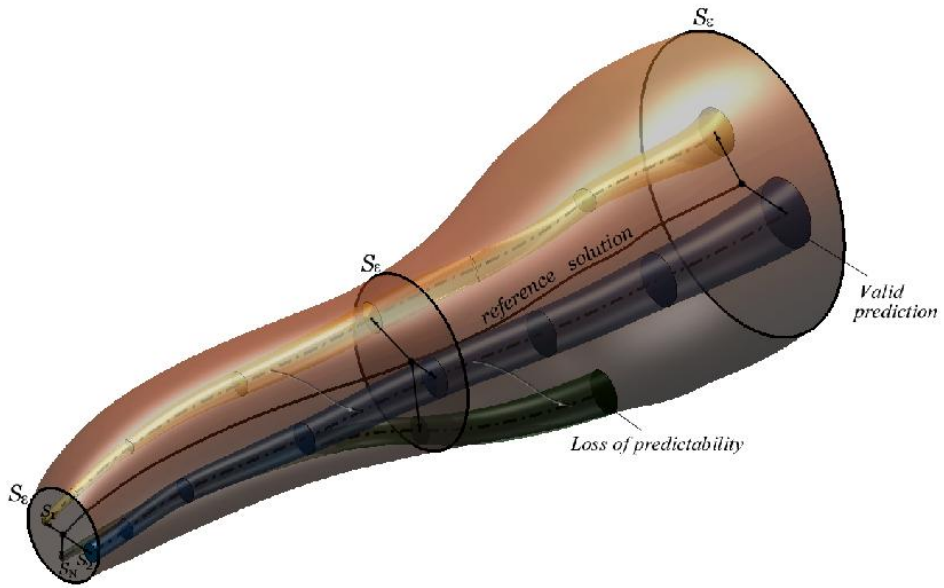


Fig. 31

Conclusions

- (1) VPP is an effective prediction skill measure (scalar).
- (2) Backward Fokker-Planck equation is a useful tool for predictability study.
- (3) Stochastic-Dynamic Modeling