# Improving Complex Models Through Stochastic Parameterization and Information Theory

**Andrew J. Majda**
Morse Professor of Arts and Sciences
Department of Mathematics and
Center for Atmosphere Ocean Science (CAOS)
Courant Institute of Mathematical Sciences
New York University

Collaborators:
I) Stochastic Parameterization of Convection: Frenkel, Khouider
II) Model Error, Information Theory: Gershgorin, Giannakis
III) Filter, Data Assimilation, Judicious Model Error: Harlim, Branicki, Gershgorin

Funded by Office of Naval Research

# Three Parts for Lecture: All parts use stochastic parameterization in different ways to cope with model errors

I. Using the stochastic multicloud model to improve deterministic tropical convective parameterization:

Khouider Majda *JAS* 2006-2009 (deterministic multicloud model)
Khouider, Biello, Majda, *Comm. Math. Sci.* 2010
Frenkel, Majda, Khouider *JAS* 2011 (submitted)

II. Improving Model Fidelity, Sensitivity and Long Range Forecasting:

Majda and Gershgorin *PNAS* 2010, 2011(A),2011(B)
Giannakis and Majda, *J. Climate*, I, II, 2011 (submitted)

III. Judicious Model Errors for Filtering Turbulent Signals:

Harlim and Majda *MWR*, 2010
Majda, Harlim, Gershgorin *DCDS*, 2010
Majda, Harlim book, *Camb. U. Press*, 2011
Gershgorin and Majda *JCP*, 2011
Branicki, Gershgorin, Majda, *JCP* 2011 (submitted)
Keating, Majda, Smith, *NWR* 2011 (submitted)

All publications are available at Majda's NYU faculty website

# Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example

Yevgeniy Frenkel (Courant Institute)
Andrew Majda (Courant Institute)
Boualem Khouider (University of Victoria)

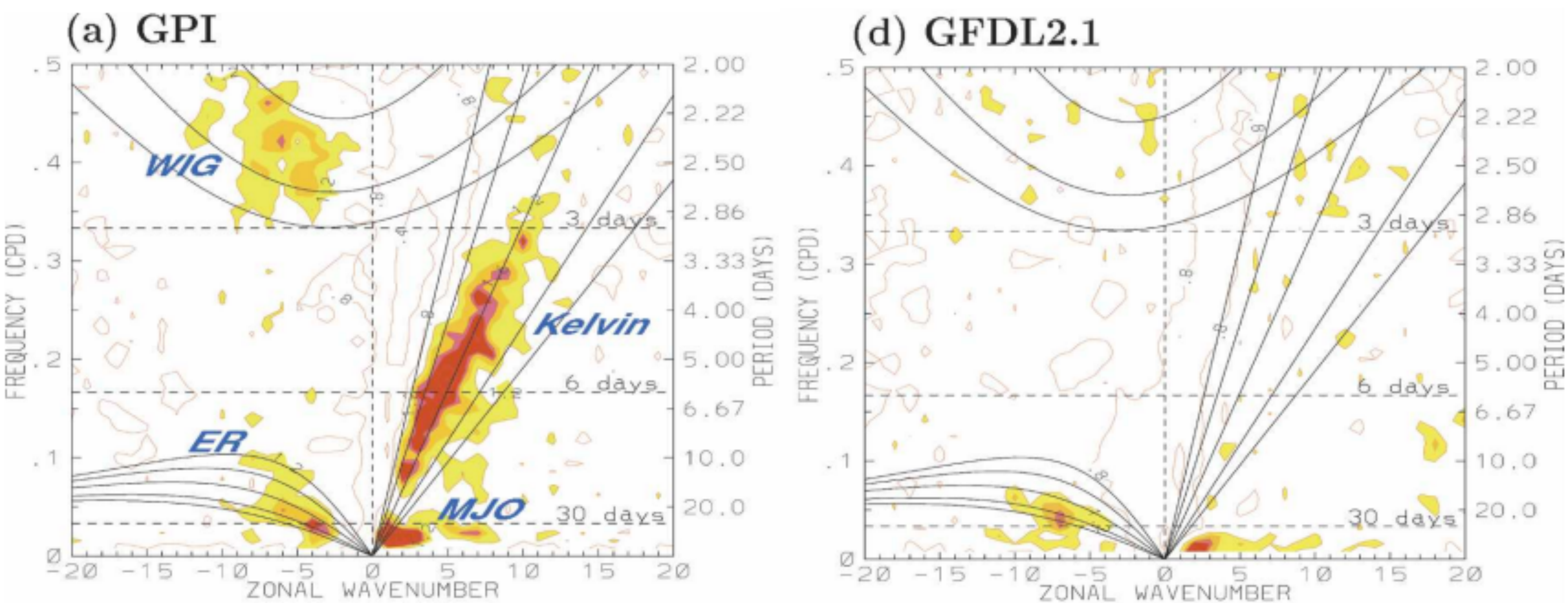
Submitted to *J. Atmos. Sci.*

# Introduction

- Current operational GCMs poorly represent the variability associated with tropical convection
- GCM convective parameterizations often fail to capture the highly intermittent organized structures of the convectively coupled waves
- Superparameterization and CRM approaches are still too computationally expensive to apply to climate forecast problems
- Stochastic convective parameterization is computationally inexpensive way to address the issue of missing variability in tropics.

# Convectively coupled waves and MJO

Observations | Typical GCM



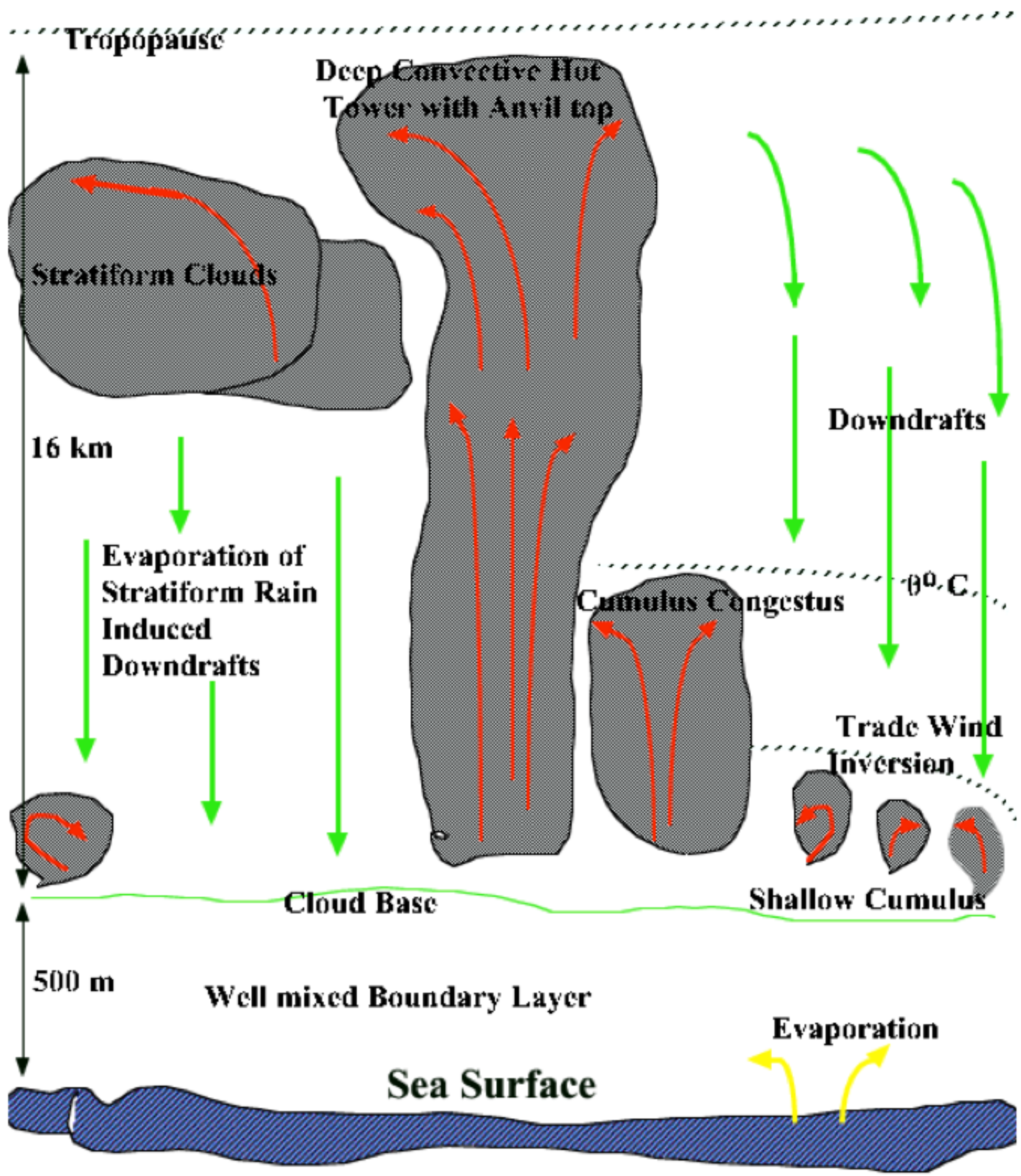

Lin *et. al* 2006

# Introduction

- A promising approach is to use a stochastic lattice to represent subgrid variability: Majda Khouider 2002, Khouider et al. 2003, Majda et al. 2008
- The stochastic multicloud model was introduced by Khouider Biello and Majda in 2010 (hereafter KBM10) in context of paradigm two-baroclinic modes single column model.
- This stochastic parameterization is based on a Markov chain lattice model where each lattice site is either occupied by a cloud of a certain type (congestus, deep or stratiform) or it is a clear sky site.
- The convective elements interact with the large scale environment and with each other through convectively available potential energy (CAPE) and middle troposphere dryness.
- Spatial interactions are ignored, and the resulting coarse grained stochastic process is computationally inexpensive to evolve via Gillespie algorithm.
- A modified version of KBM10 model is used here to study flows above the equator without rotation effects.

# Table of contents

# Multicloud model




(Khouider & Majda 2008)

# Dynamical Core of the multicloud model

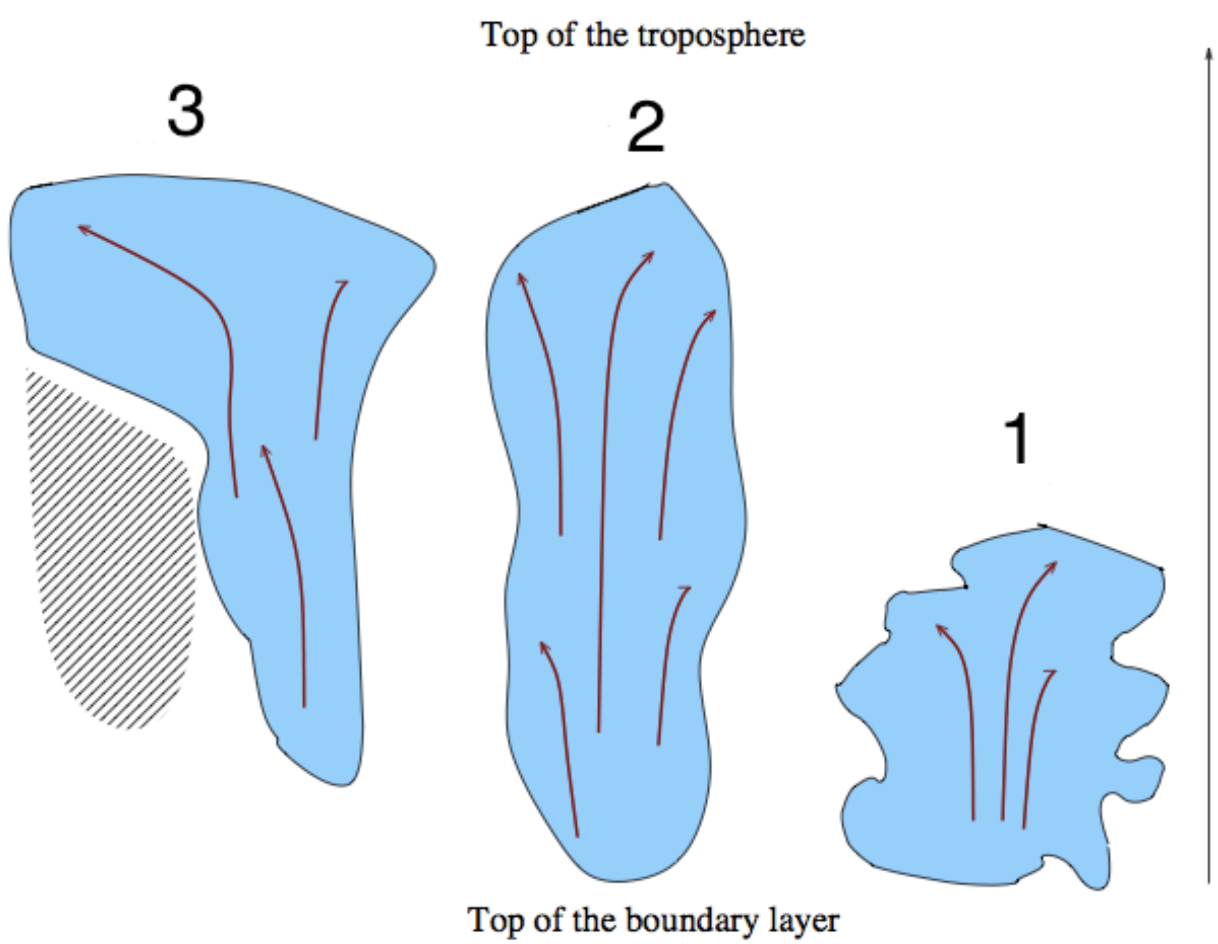
The dynamical core of the model consists of two forced and coupled shallow-water systems for first two baroclinic modes of potential temperature and zonal velocity.

| | |
|---|---|
| Momentum, $j$st mode, $j = 1, 2$ | $\partial_t u_j - \partial_x \theta_j = -C_d u_0 u_j - \frac{1}{\tau_R} u_j$ |
| Potential temperature, 1st mode | $\partial_t \theta_1 - \partial_x u_1 = P - Q_{R,1}^0 - \tau_D^{-1} \theta_1$ |
| Potential temperature, 2nd mode | $\partial_t \theta_2 - \frac{1}{4} \partial_x u_1 = H_c - H_s - Q_{R,2}^0 - \tau_D^{-1} \theta_2$ |

The precipitation $P = H_d + \xi_s H_s + \xi_c H_c$ allows for the contribution of deep convective as well as stratiform and congestus rain.

For simplicity we remove congestus rain by letting $\xi_c = 0$ and set parameter $\xi_s$ so that at RCE 40 percent of rain comes from stratiform clouds

# Stochastic lattice

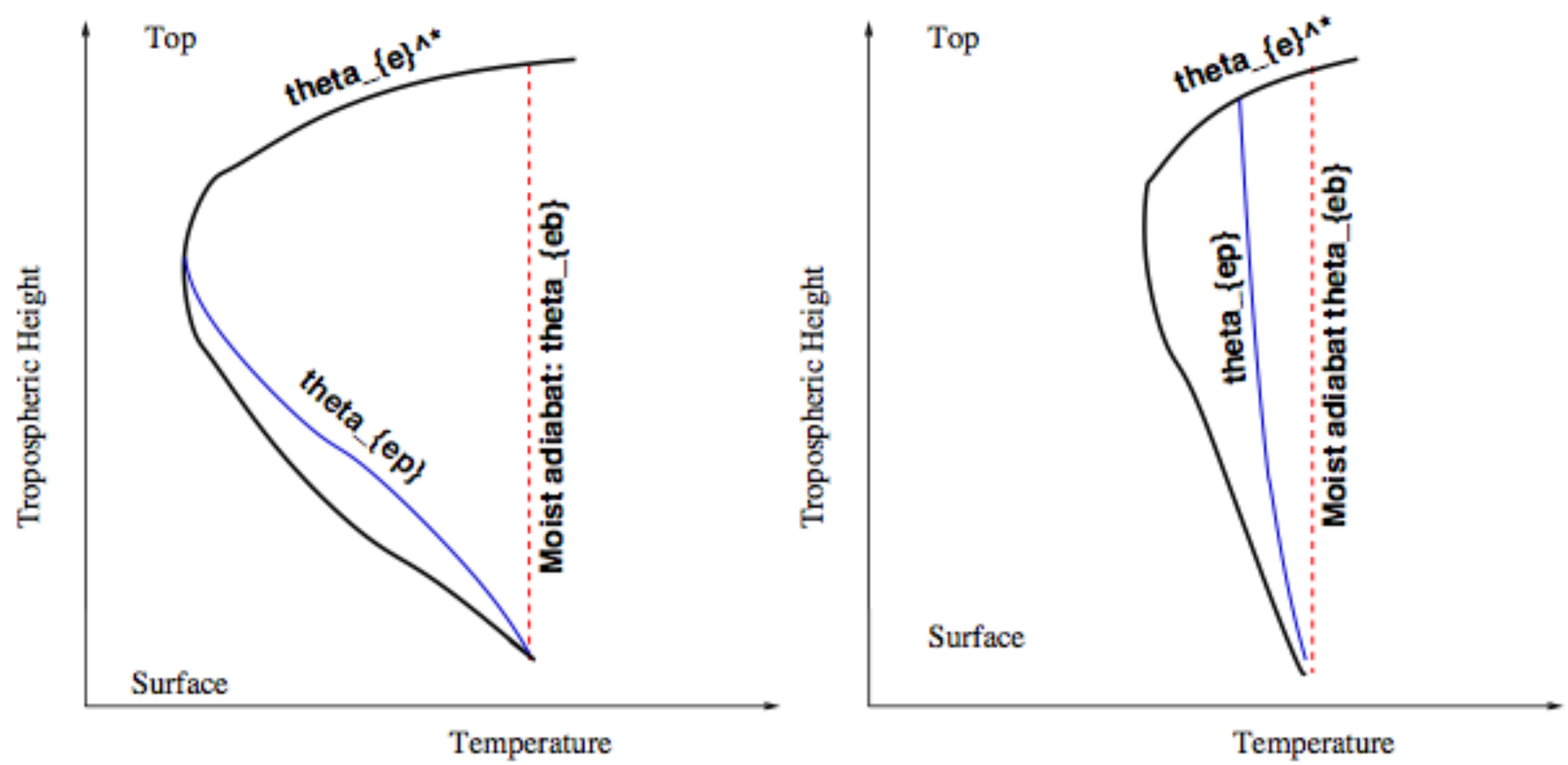



| | | | |
|---|---|---|---|
| 0 | 1 | 0 | 2 |
| 2 | 3 | 2 | 0 |
| 0 | 0 | 1 | 0 |
| 1 | 1 | 0 | 3 |

The number of convective sites is set to $30^2$ for the experiments on the 40 km grid

# Transition rates are defined through three atmospheric qualities with scaling parameters

- Dryness of Atmosphere $D = \frac{\theta_{eb} - \theta_{em}}{D_0}$




- Scaled low level CAPE $C_l = \frac{CAPE_l}{CAPE_0}$
- Scaled CAPE $C = \frac{CAPE}{CAPE_0}$ (Note that $CAPE_0$ can be viewed as "activation" energy)

# Transition rates and time scales

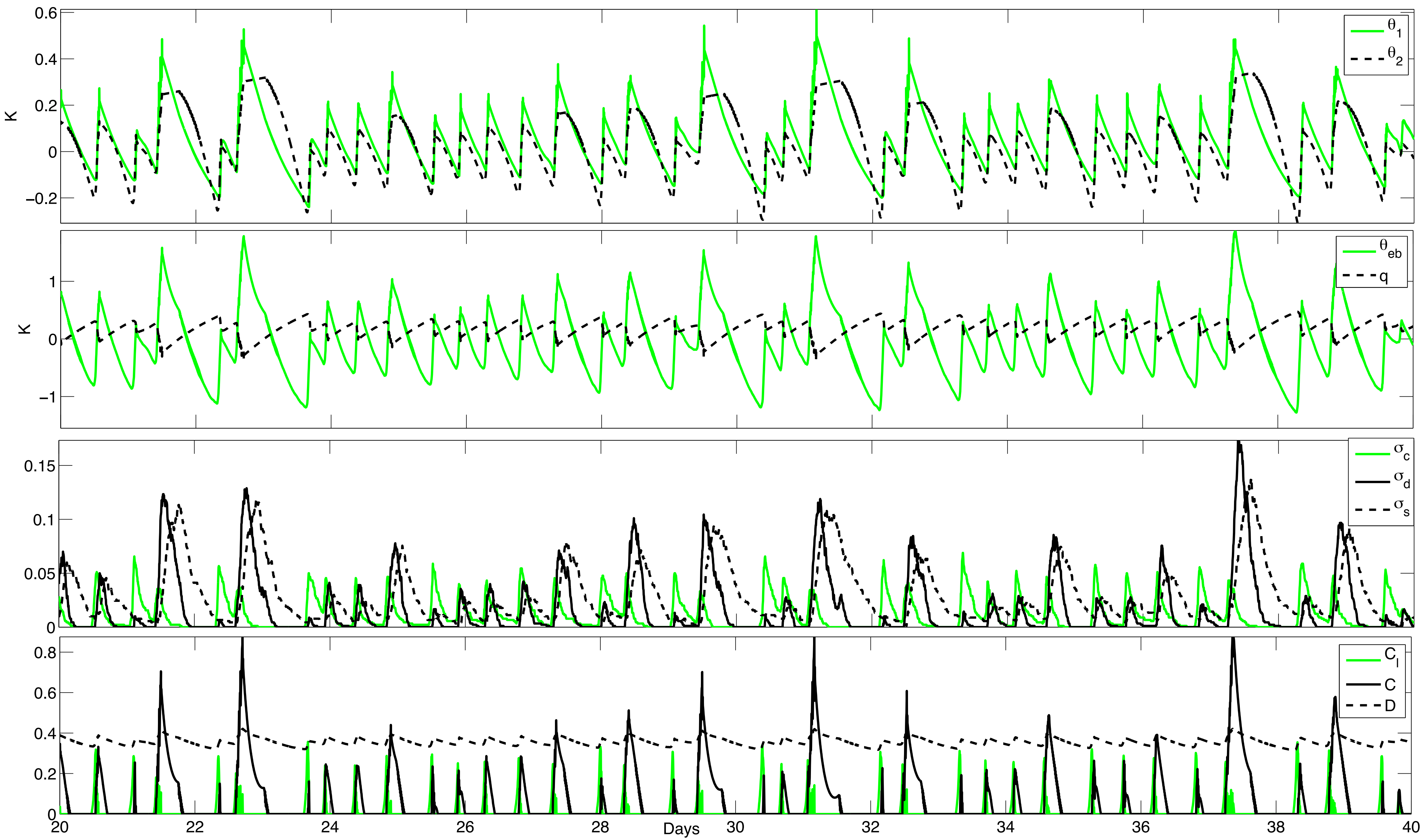
Creation of congestus clouds requires dry atmosphere and abundance of low level CAPE. Note the parity of the congestus creation and decay time scales

| Transition | Transition Rate | Time scale(h) |
|---|---|---|
| Formation of congest | $R_{01} = \frac{1}{\tau_{01}}\Gamma(C_l)\Gamma(D)$ | $\tau_{01}=1\tau_{grid}$ |
| Decay of congestus | $R_{10} = \frac{1}{\tau_{10}}\Gamma(D)$ | $\tau_{10}=1\tau_{grid}$ |
| Conversion of congest to deep | $R_{12} = \frac{1}{\tau_{12}}\Gamma(C)(1-\Gamma(D))$ | $\tau_{12}=1\tau_{grid}$ |
| Formation of deep | $R_{02} = \frac{1}{\tau_{02}}\Gamma(C)(1-\Gamma(D))$ | $\tau_{02}=3\tau_{grid}$ |
| Conversion of deep to stratiform | $R_{23} = \frac{1}{\tau_{23}}$ | $\tau_{23}=3\tau_{grid}$ |
| Decay of deep | $R_{20} = \frac{1}{\tau_{20}}(1-\Gamma(C))$ | $\tau_{20}=3\tau_{grid}$ |
| Decay of stratiform | $R_{30} = \frac{1}{\tau_{30}}$ | $\tau_{30}=5\tau_{grid}$ |

$\Gamma(x) = 1 - e^{-x}$ for $x > 0$ and 0 otherwise

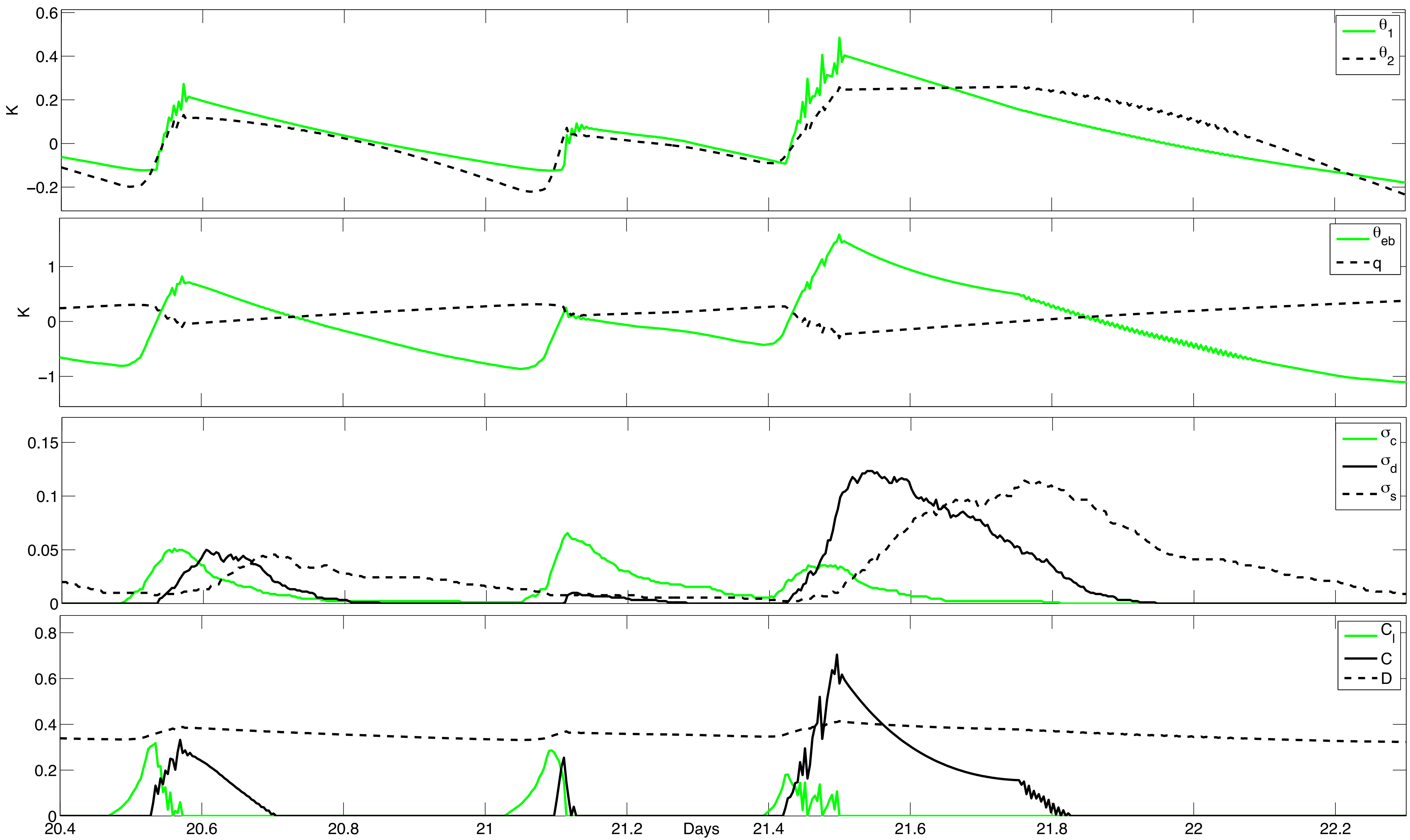
# Intermittent solution of single column simulation

The time series shows intermittent patterns of large and small convective events. Both small and large convective events follow the congestus to deep to stratiform pattern.



# Intermittent solution of single column simulation (closeup)

Smaller congestus cloud heavy convective events precondition atmosphere for large convective events dominated by the direct clear sky to deep convection transitions



# Walker type circulation simulations (outline)

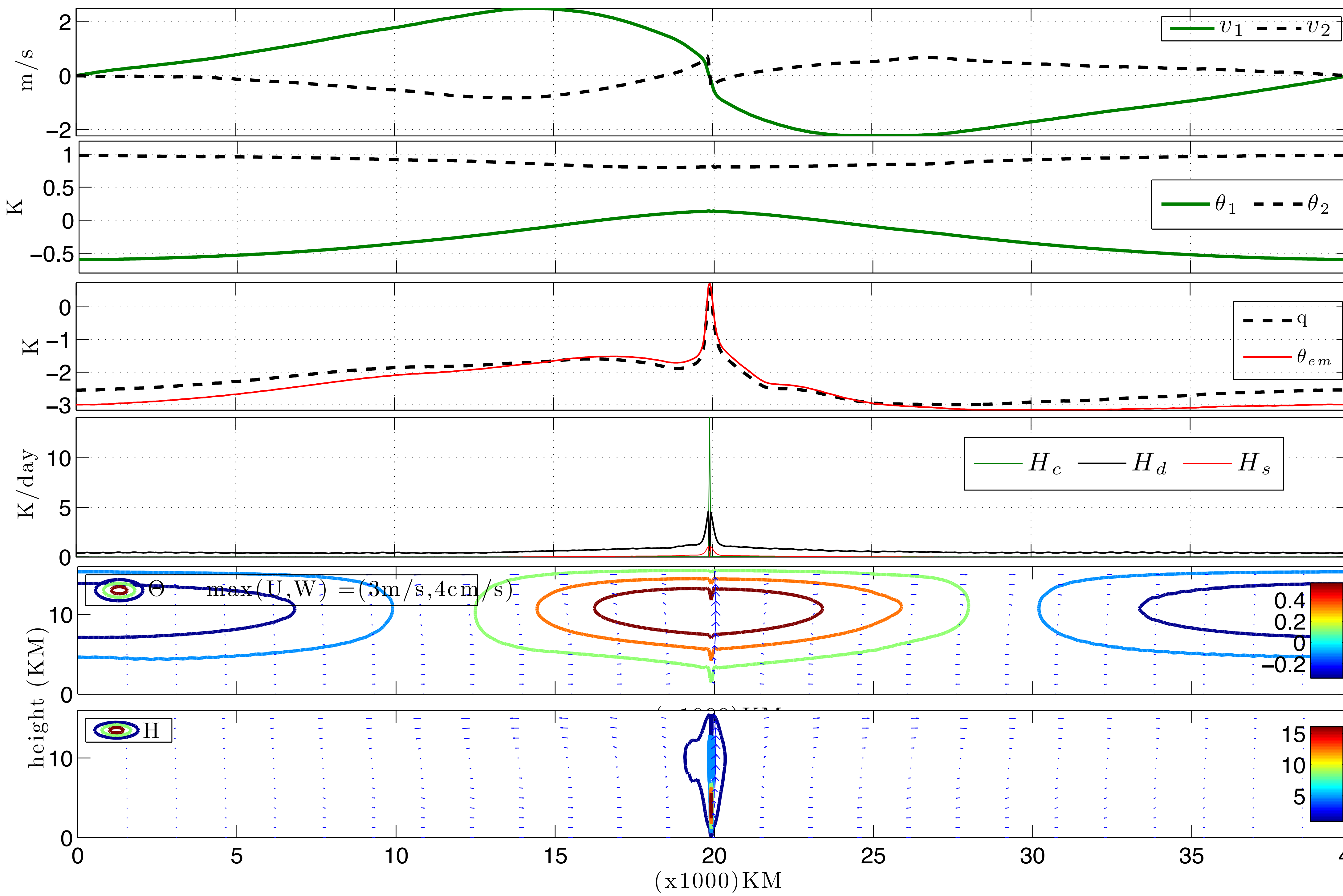
Paradigm analog of Walker circulation in deterministic GCMs

Walker circulation in stochastic multicloud model with moderate resolution

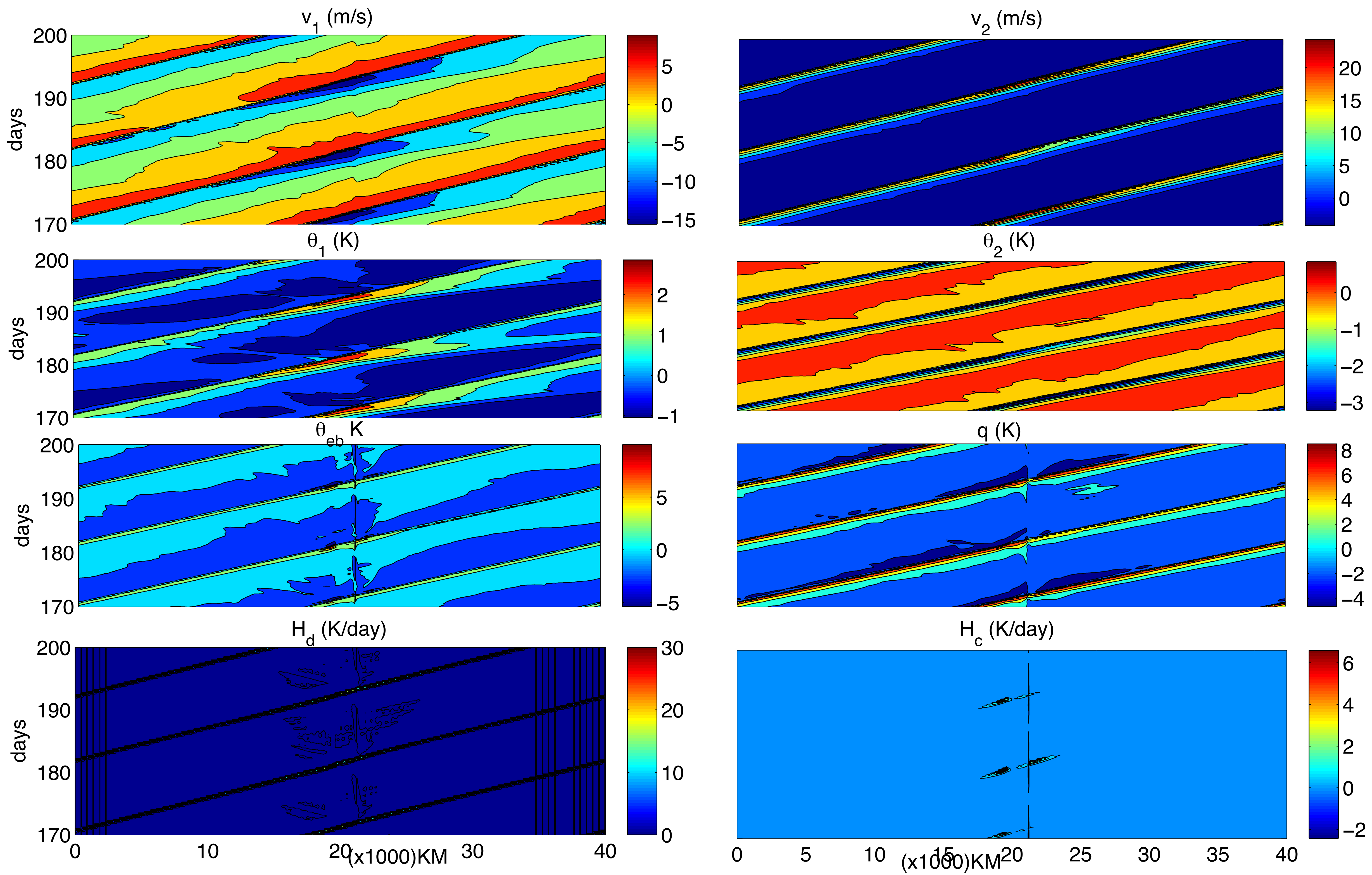
Walker circulation in stochastic multicloud model with coarse resolution

Quantitative comparison of the variability

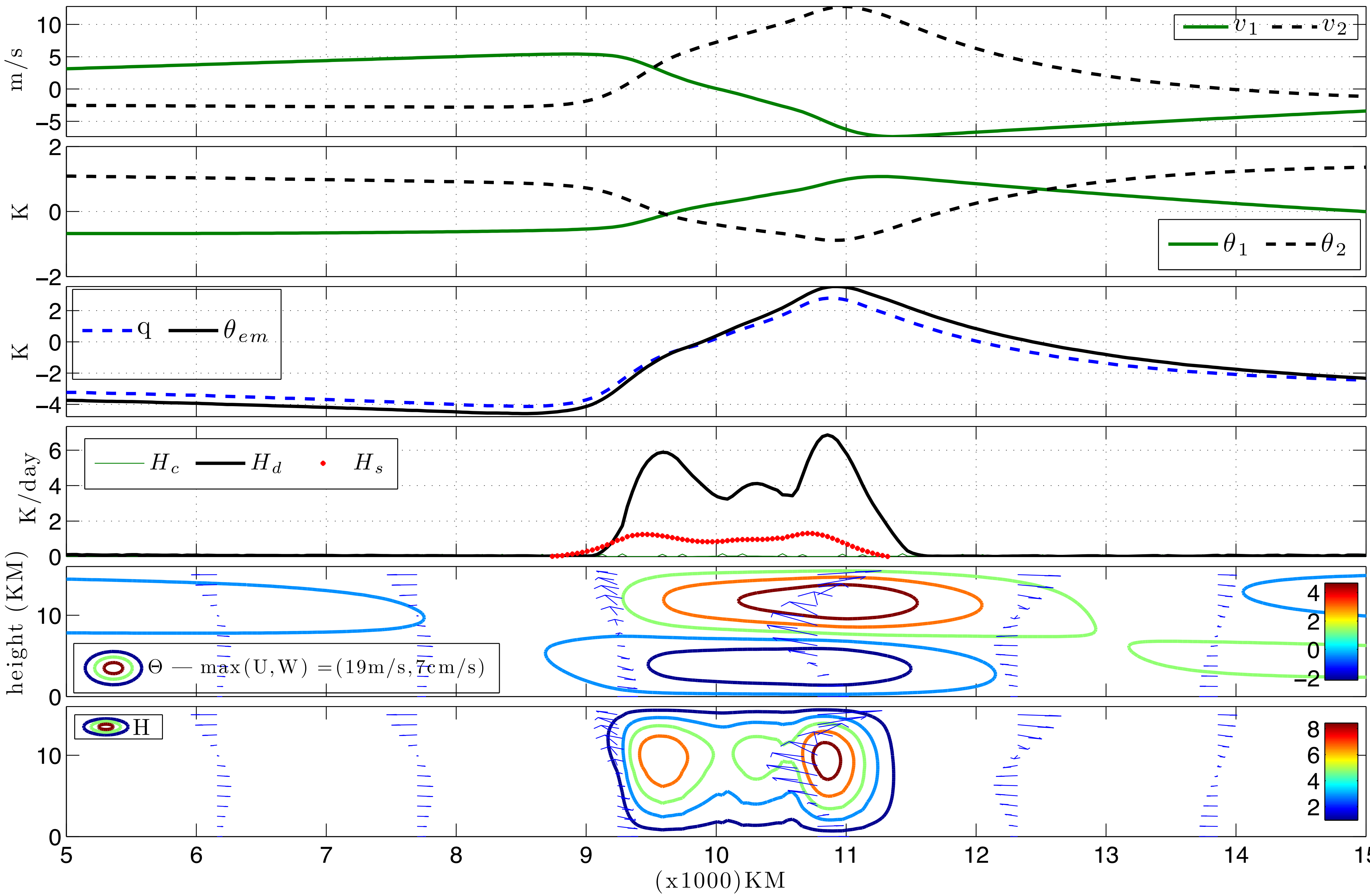
# Deterministic SST gradient induced Walker type circulation (mean)



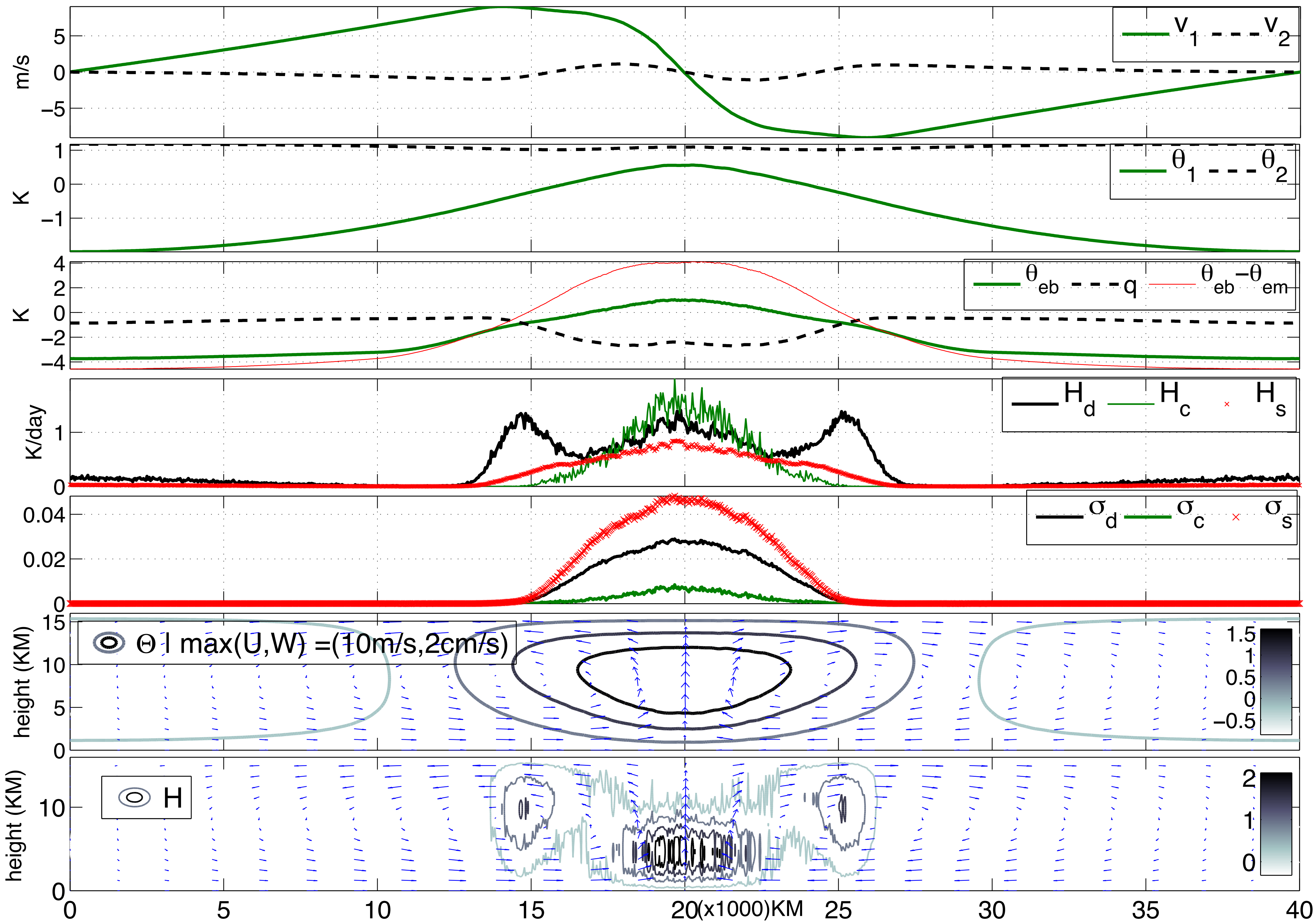
# Deterministic SST gradient induced Walker type circulation (deviations from the mean)



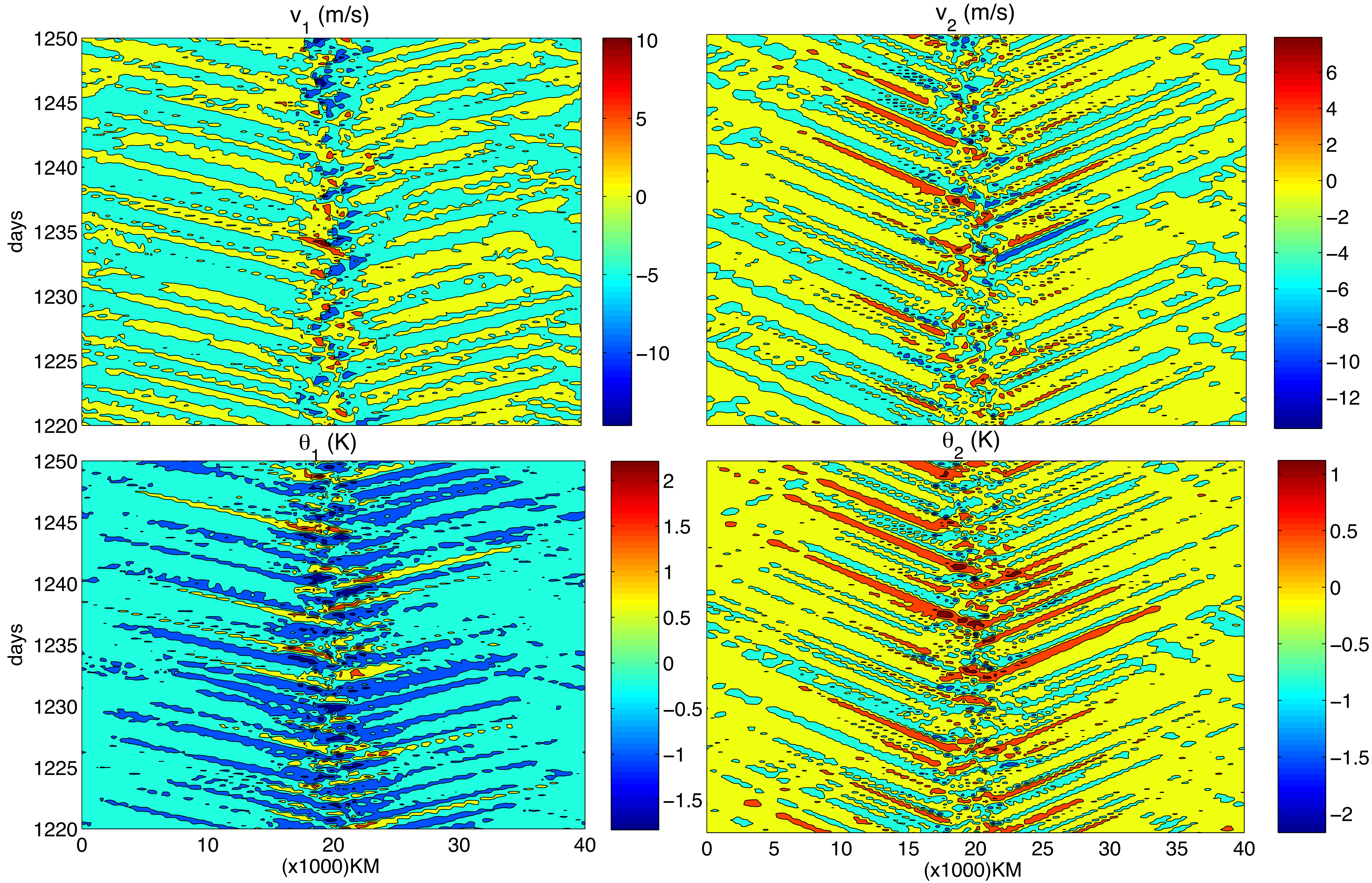
# Eastward propagating waves in the suboptimal parameter regime deterministic simulation



# SST gradient induced Walker type circulation (mean)

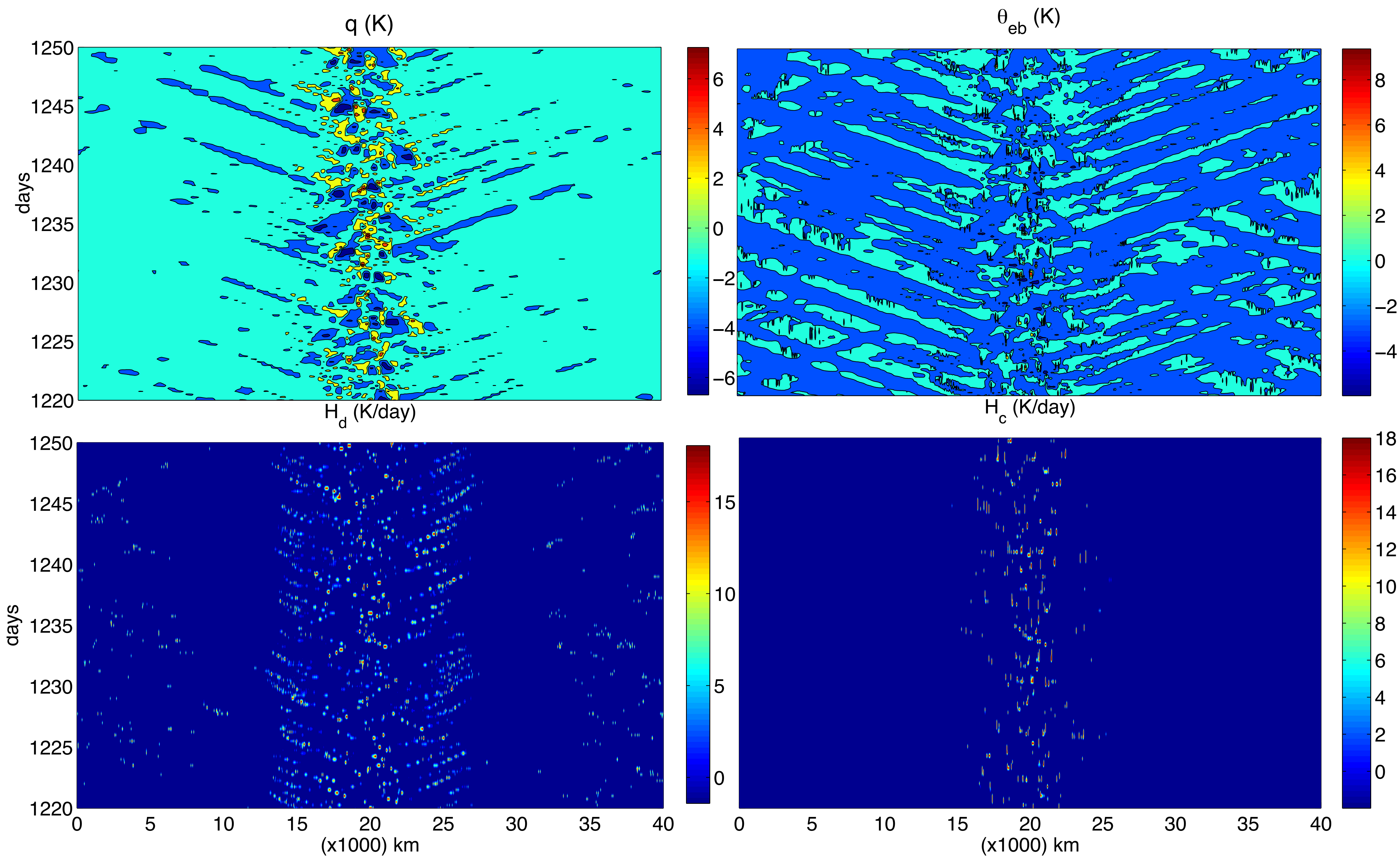


# SST gradient induced Walker type circulation (deviations )



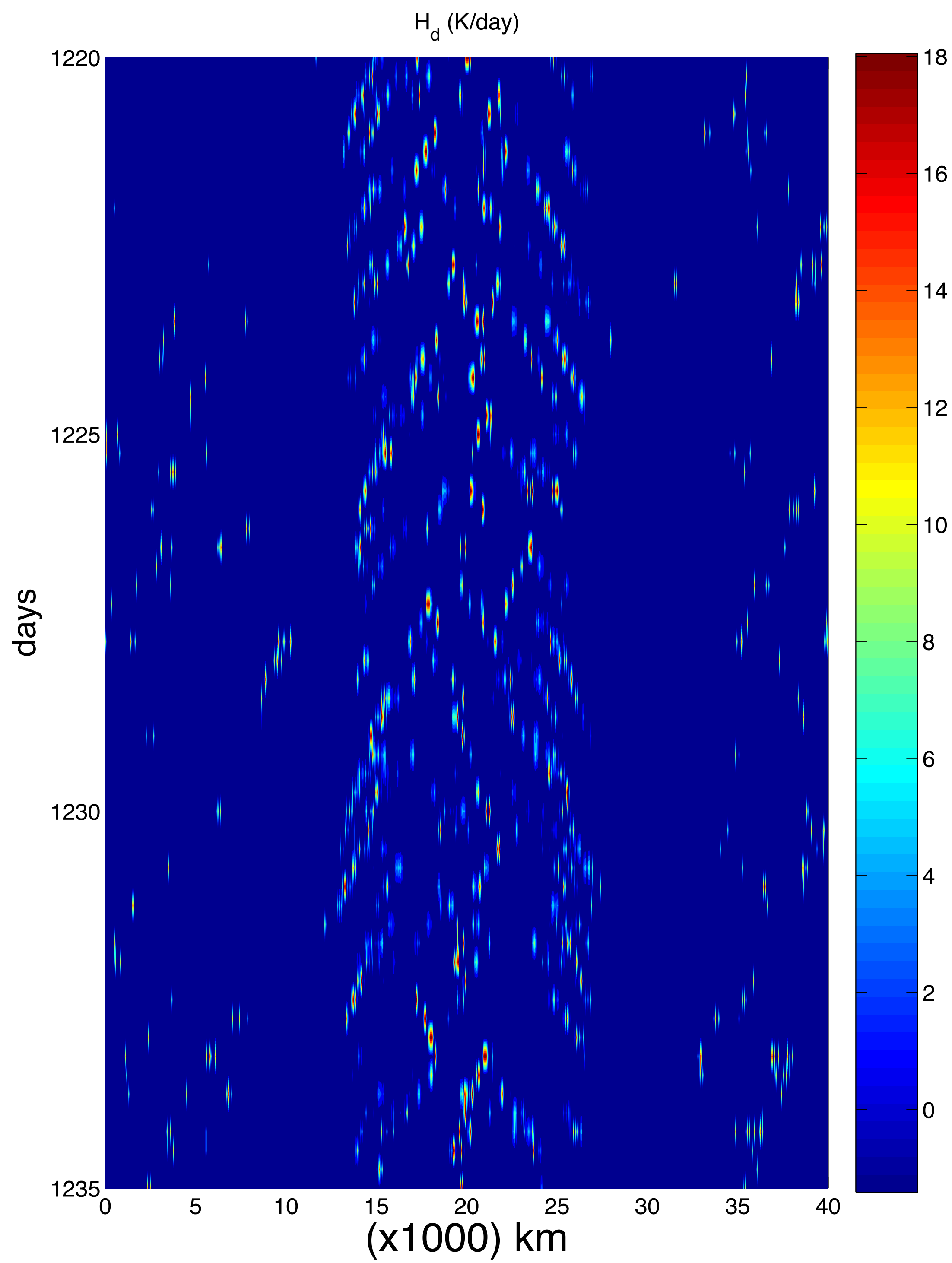
# SST gradient induced Walker type circulation (deviations)

Congestus cloud decks are localized to the center region of the warm pool.

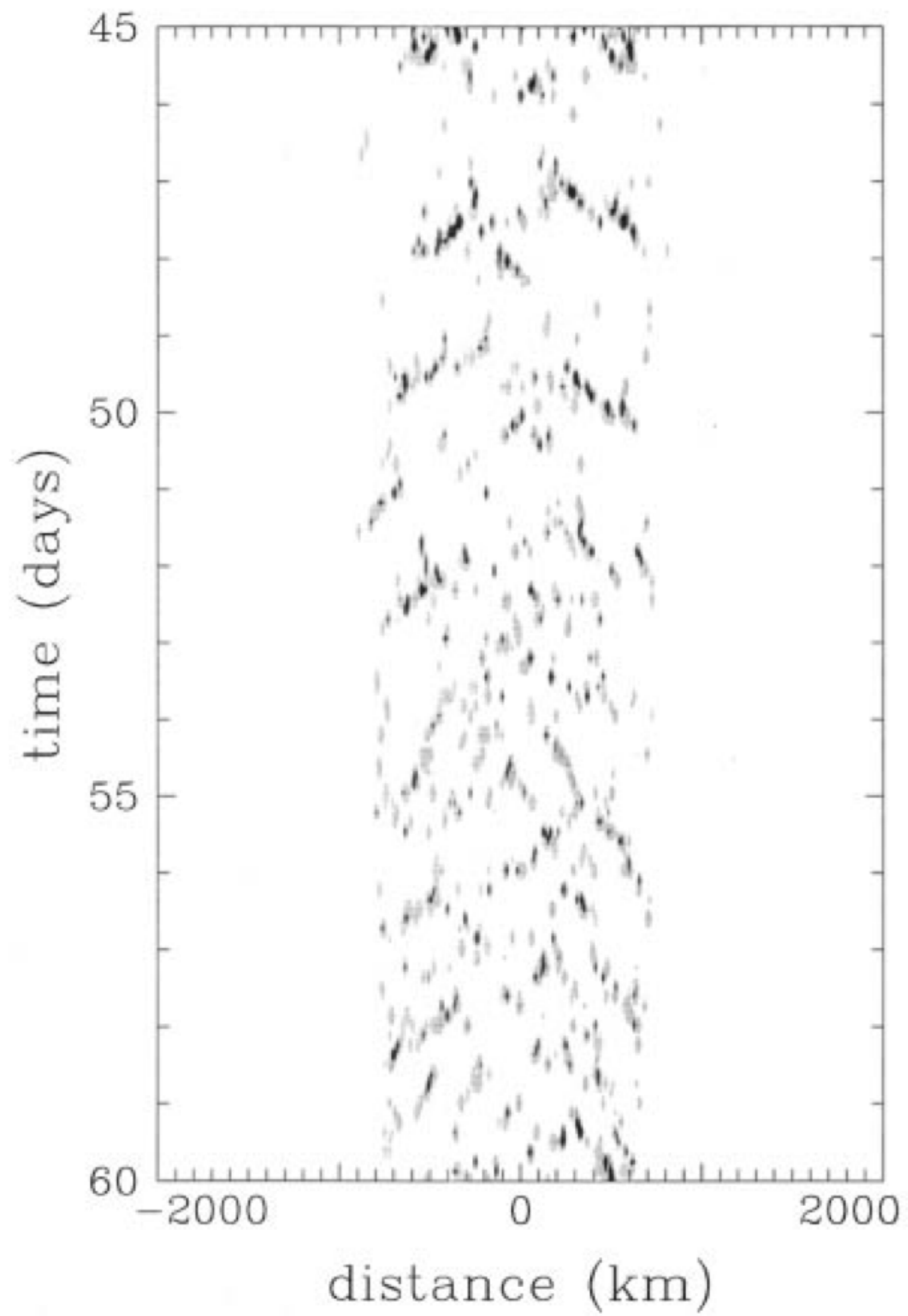


# Precipitation Profiles

Stochastic multicloud model

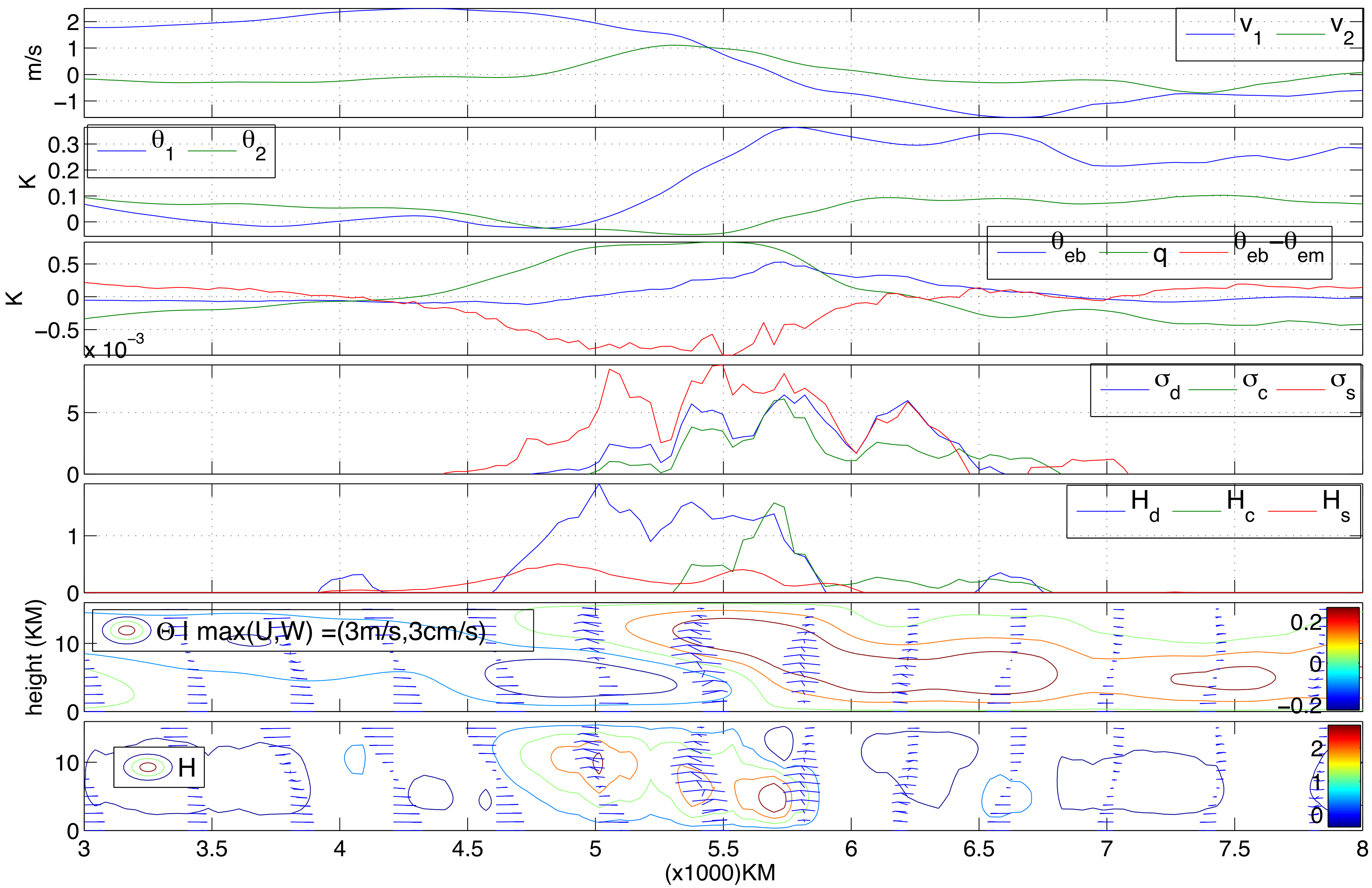



CRM ( Grabowski et. al 2000)



# Structure of a eastward propagating wave born on the boundary of the warm pool

Both cloud fractions and heating fields follow the congestus to deep to stratiform pattern. This results in characteristic tilt of the heating field.



# Mean circulation strength and variability of heating fields for the stochastic and deterministic parameterizations

- Stochastic multicloud model outperforms its deterministic counterpart, providing higher variability with more realistic convective structures.
- Stochastic model can be scaled to coarse grid in a manner that preserves the variability and statistical structure of the coherent features.

| Model | grid (km) | $\tau_{grid}$ | $n$ | max $(U, W)$ | $std(H_d)$ K/Day | std$(H_c)$ K/Day |
|---|---|---|---|---|---|---|
| Stochastic | 40 | 1 | $30^2$ | $(10m/s, 2cm/s)$ | 2.14 | 2.83 |
| Stochastic | 160 | 1 | $120^2$ | $(12m/s, 3cm/s)$ | 1.34 | 1.89 |
| Stochastic | 160 | 1 | $30^2$ | $(12m/s, 3cm/s)$ | 1.67 | 2.41 |
| Stochastic | 160 | 3 | $30^2$ | $(12m/s, 4cm/s)$ | 1.80 | 2.21 |
| Stochastic | 160 | 4 | $30^2$ | $(12m/s, 6cm/s)$ | 1.96 | 2.07 |
| Stochastic | 160 | 5 | $30^2$ | $(11m/s, 5cm/s)$ | 2.09 | 1.80 |
| Stochastic | 160 | 16 | $30^2$ | $(10m/s, 3cm/s)$ | 0.49 | 0.89 |
| Determin. | 40 | - | - | $(4m/s, 4cm/s)$ | 0.97 | 0.14 |
| Determin. | 160 | - | - | $(5m/s, 4cm/s)$ | 0.55 | 0.14 |

Free tropospheric moisture equation is identical to KBM10 except for the addition of moisture convergence term

| | |
|---|---|
| Free tropospheric moisture | $\partial_t q + \frac{\partial}{\partial x}[(u_1 + \tilde{\alpha} u_2)q + \tilde{Q}(u_1 + \tilde{\lambda} u_2)] = -P + \frac{D}{H_T}$ |
| Boundary layer equivalent potential temperature | $\partial_t \theta_{eb} = \frac{1}{h_b}(E - D)$ |
| Downdrafts | $D = m_0[1 + \mu(H_s - H_c)/Q^0_{R,1}]^+ \triangle_m \theta_e$ |
| Sea surface evaporation flux | $E/h_b = \tau_e^{-1}(\theta^\star_{eb} - \theta_{eb})$ |

For simulations with spatial SST gradient $\theta^*_{eb}(x) = 5\cos\left(\frac{4\pi x}{40000}\right)$ in central 20 000 km of the domain and -5K outside of the region

# Convective heating closures

Congestus heating closure $\qquad H_c = \sigma_c \frac{\bar{\alpha}\alpha_c}{H_m}\sqrt{CAPE_l^+}$

Deep heating closure $\qquad H_d = \left(\sigma_d \bar{Q} + \frac{1}{\tau_c(\sigma_d)}(a_1\theta_{eb} + a_2 q - a_0(\theta_1 + \gamma_2\theta_2))\right)^+$

Stratiform heating closure $\qquad H_s = \alpha_s[\sigma_s \bar{Q} + \frac{1}{\tau_c(\sigma_s)}(a_1\theta_{eb} + a_2 q - a_0(\theta_1 + \gamma_2\theta_2))]^+$

Maximum energy available for deep convection $\qquad CAPE = \bar{CAPE} + R(\theta_{eb} - \gamma(\theta_1 + \gamma_2\theta_2))$

Maximum energy available for congestus convection $\qquad CAPE_l = \bar{CAPE} + R(\theta_{eb} - \gamma(\theta_1 + \gamma_2'\theta_2))$

$$\tau_c(\sigma_X) = \frac{\bar{\sigma}_x}{\sigma_x}\tau_c^0$$

# Conclusions

- The stochastic model dramatically improves the variability of tropical convection compared to the conventional moderate and coarse resolution paradigm GCM parameterizations.
- This increase in variability comes from intermittent coherent structures such as synoptic and mesoscale convective systems, analogs of squall lines and convectively coupled waves seen in nature whose representation is improved by the stochastic parameterization.
- Furthermore, simulations with sea surface temperature (SST) gradient yield realistic mean Walker-cell circulation with plausible high variability.
- An additional feature of the present stochastic parameterization is a natural scaling of the model from moderate to coarse grids which preserves the variability and statistical structure of the coherent features.

# Quantifying Uncertainty in Climate Change Science: Empirical Information Theory, Fluctuation Dissipation


- A. Majda and B. Gershgorin, Quantifying Uncertainty in Climate Change Science through Empirical Information Theory, PNAS 107, p. 14958 (2010)
- A. Majda, R. Abramov, and B. Gershgorin, High Skill in Low Frequency Climate Response through Fluctuation Dissipation Theorems despite Structural Instability, PNAS 107, p. 581 (2010)
- A. Majda and B. Gershgorin, Improving Model Fidelity and Sensitivity for Complex Systems through Empirical Information Theory, PNAS, in press (2011)
- A. Majda and B. Gershgorin, The Link between Statistical Equilibrium Fidelity and Forecasting Skill for Complex Systems with Model Error, PNAS, in press (2011)
- A. Majda, B. Gershgorin, and Y. Yuan, Low Frequency Climate Response and Fluctuation Dissipation Theorems: Theory and Practice, JAS 67, p. 1186 (2010)

# Practical questions in climate change science

- How will the mean temperature change if the heating from the sun increases?
- How will the variance of the temperature respond to the changes of $CO_2$ concentration?
- How will the mean velocity profile in the ocean behave if the salinity starts changing?
- How will the mean temperature in April change if the heating in January decreases?

# Quantifying Uncertainty in Climate Change Science through Empirical Information Theory

Quantifying the uncertainty for the present climate and the predictions of climate change in the suite of imperfect Atmosphere Ocean Science (AOS) computer models is a central issue in climate change science.
Basic questions:

- A How to measure the skill of a given model in reproducing the present climate and predicting the future climate in an unbiased fashion?

- B How to make the best possible estimate of climate sensitivity to changes in external or internal parameters by utilizing the imperfect knowledge available of the present climate?
What are the most dangerous parameters for climate change given uncertain knowledge of the present climate?

- C How do coarse-grained measurements of different functionals of the present climate affect the assessments in A), B)?
What are the weights which should be assigned to different functionals of the present climate as targets to improve the performance of the imperfect AOS models?
Which new functionals of the present climate should be observed in order to improve the assessments in A), B)?

Difficulty: Don't know dynamics for actual climate!

# Empirical Information Theory

Jaynes 1957

Majda, Abramov, Grote 2005 AMS

Majda, Wang 2006, Cambridge Press

## Empirical information theory and climate science

With a subset of variables $\vec{u} \in \mathbb{R}^N$ and a family of measurement functionals $\vec{E}_L(\vec{u}) = (E_j(\vec{u}))$, $1 \leq j \leq L$, for the present climate, $\underline{\text{empirical information theory builds the least biased}}$ $\underline{\text{probability measure } \pi_L(\vec{u}) \text{ consistent with the } L \text{ measurements}}$ of the present climate, $\overline{\vec{E}}_L$.

The natural way to measure the lack of information in one probability density, $q(\vec{u})$, compared with the true probability density, $p(\vec{u})$, is through the relative entropy, $\mathcal{P}(p, q)$, given by

$$\mathcal{P}(p, q) = \int p \ln \left( \frac{p}{q} \right).$$

This functional on probability densities has two attractive features as a metric for climate change science:

1) $\mathcal{P}(p, q) \geq 0$ with equality if and only if $p = q$,
2) $\mathcal{P}(p, q)$ is invariant under general nonlinear changes of variables.

$\mathcal{P}(\pi, \pi_L)$ precisely quantifies the intrinsic error in using the $L$ measurements of the present climate, $\overline{\vec{E}}_L$.

An AOS model for the present climate is described by $\pi^M(\vec{u})$, intrinsic model error in the climate statistics is given by

$$\mathcal{P}(\pi, \pi^M).$$

Consider a class of imperfect models, $\mathcal{M}$, for the climate, the best climate model for the coarse-grained variable $\vec{u}$ is the $M_* \in \mathcal{M}$ so that the true climate has the smallest additional information beyond the modelled climate distribution $\pi^{M_*}(\vec{u})$, i.e.,

$$\mathcal{P}(\pi, \pi^{M_*}) = \min_{M \in \mathcal{M}} \mathcal{P}(\pi, \pi^M).$$

Also, actual improvements in a given climate model with distribution $\pi^M(\vec{u})$ either through higher resolution or improved parameterization resulting in a new $\pi^M_{post}(\vec{u})$ should result in improved information for the actual climate, so that

$$\mathcal{P}(\pi, \pi^M_{post}) \le \mathcal{P}(\pi, \pi^M),$$

otherwise, objectively, the model has not been improved compared with the original climate model.

**<u>Fact1</u>** :
$$\mathcal{P}(\pi, \pi_{L'}^{M}) = \mathcal{P}(\pi, \pi_L) + \mathcal{P}(\pi_L, \pi_{L'}^{M})$$
$$= (\mathcal{S}(\pi_L) - \mathcal{S}(\pi)) + \mathcal{P}(\pi_L, \pi_{L'}^{M}) \text{ for } L' \leq L.$$

The unbiased intrinsic error in the finite number of climate measurements in of the actual climate is exactly the entropy difference. With **<u>Fact1</u>** and a fixed family of $L$ measurements of the actual climate, the <u>optimization principles can be computed</u> explicitly <u>by replacing the unknown</u> $\pi$ <u>by</u> the hypothetically <u>known</u> $\pi_L$ in these formulas so that for example, $\pi^{M_*}$ is calculated by

$$\mathcal{P}(\pi_L, \pi_{L'}^{M_*}) = \min_{M \in \mathcal{M}} \mathcal{P}(\pi_L, \pi_{L'}^{M}).$$

# Algorithms for effective calculation of the empirical metrics for climate uncertainty

Practical setup for calibration of contemporary AOS models: climate measurements and model measurements involve only mean and covariance of $\vec{u}$ so that $\pi_L$ is Gaussian with climate mean $\overline{\vec{u}}$ and covariance $R$ while $\pi^M$ is Gaussian with model mean $\overline{\vec{u}}_M$ and covariance $R_M$.
$\mathcal{P}(\pi_L, \pi^M)$ has the explicit formula:

$$\mathcal{P}(\pi_L, \pi^M) = \left[\frac{1}{2}(\overline{\vec{u}} - \overline{\vec{u}}_M)^*(R_M)^{-1}(\overline{\vec{u}} - \overline{\vec{u}}_M)\right]$$
$$+ \left[-\frac{1}{2}\log\det(RR_M^{-1}) + \frac{1}{2}(\mathrm{tr}(RR_M^{-1}) - N)\right].$$

First term is the signal, reflecting the model error in the mean but weighted by the inverse of the model covariance, $R_M^{-1}$, while the second term, the dispersion, involves only the model error covariance ratio, $RR_M^{-1}$.
This intrinsic metric is invariant under any (linear) change of variables which maps Gaussian distributions to Gaussians and the signal and dispersion terms are individually invariant under these transformations.

Non-Gaussian statistics: Kleeman (2002), Majda Kleeman Cai (2002), Haven Majda Abramov (2005), Abramov (2006-7-9)

# A simple example with an intrinsic barrier for improving model sensitivity

Perfect model:

$$\frac{du}{dt} = au + v + F,$$
$$\frac{dv}{dt} = qu + Av + \sigma \dot{W},$$

smooth Gausssian measure if $\quad a + A < 0, \quad aA - q > 0.$

Imperfect model:
$$\frac{du_M}{dt} = -\gamma_M u_M + F_M + \sigma_M \dot{W}_M.$$

# Climate Fidelity for Imperfect Model

$$\frac{F_M}{\gamma_M} = -\frac{AF}{aA - q}, \qquad \frac{\sigma_M^2}{2\gamma_M} = \frac{\sigma^2}{2(a + A)(aA - q)} \equiv E.$$

Response to change in forcing:

$$\delta u = -\frac{A}{aA - q}\delta F, \qquad \delta u_M = \frac{1}{\gamma_M}\delta F,$$

# Information Model Error

With perfect model fidelity

$$\mathcal{P}(\pi_\delta, \pi_\delta^M) = \frac{1}{2} E^{-1} \left| -\frac{A}{aA - q} - \frac{1}{\gamma_M} \right|^2 |\delta F|^2.$$

In this situation with A>0, the attempt to minimize the information theoretic model error is futile because no finite minimum over $\gamma_M$ is achived and necessarily $\gamma_M \to \infty$ in the approach to the minimum - intrinsic barrier.

With A<0, minimize the lack of information in the sensitivity:

$$\gamma_M^* = -A^{-1}(aA - q), \quad A < 0,$$

Capture fidelity and sensitivity!

# Empirical theory for finding the most dangerous climate change directions from the present climate

Consider a family of parameters $\vec{\lambda} \in \mathbb{R}^p$ with $\pi_{\vec{\lambda}}$ the true climate that occurs; $\vec{\lambda}$ — external parameters, changes in forcing, internal variability, change in dissipation.

The most dangerous perturbed climate is the one with largest uncertainty of the present climate

$$\mathcal{P}(\pi_{\vec{\lambda}_*}, \pi) = \max_{\vec{\lambda} \in \mathbb{R}^p} \mathcal{P}(\pi_{\vec{\lambda}}, \pi).$$

$$\pi_{\vec{\lambda}}\Big|_{\vec{\lambda}=0} = \pi, \text{ then for small values of } \vec{\lambda}:$$

$$\mathcal{P}(\pi_{\vec{\lambda}}, \pi) = \vec{\lambda} \cdot I(\pi)\vec{\lambda} + O(|\vec{\lambda}|^3).$$

Fisher information:

$$\vec{\lambda} \cdot I(\pi)\vec{\lambda} = \int \frac{(\vec{\lambda} \cdot \nabla_{\vec{\lambda}}\pi)^2}{\pi}.$$

<u>Fact 2</u>: The most dangerous climate change direction occurs along the unit direction $\vec{e}^{\,*}_{\pi} \in \mathbb{R}^p$ which is associated with the largest eigenvalue, $\lambda^*_{\pi}$, of the above quadratic form.

# Exactly solvable test models for climate change science

$U(t) = \bar{U}(t) + U'(t),$ zonal jet, seasonal cycle

$v(x,t),$ turbulent Rossby waves

$T(x,t) + \alpha y = \text{``}T\text{''},$ passive tracer with mean gradient ($CO_2$, CO, etc)

$$\frac{dU}{dt} = -\gamma(U - \bar{U}(t)) + \sigma \dot{W},$$ similar equation for each Fourier mode of $v$

$$\frac{\partial T}{\partial t} + U(t)\frac{\partial T}{\partial x} = -\alpha v(x,t) + \kappa \frac{\partial^2 T}{\partial x^2},$$

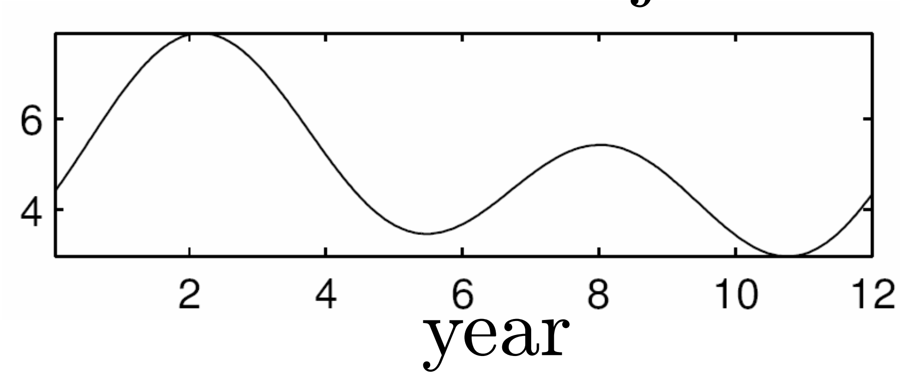
Statistically exactly solvable

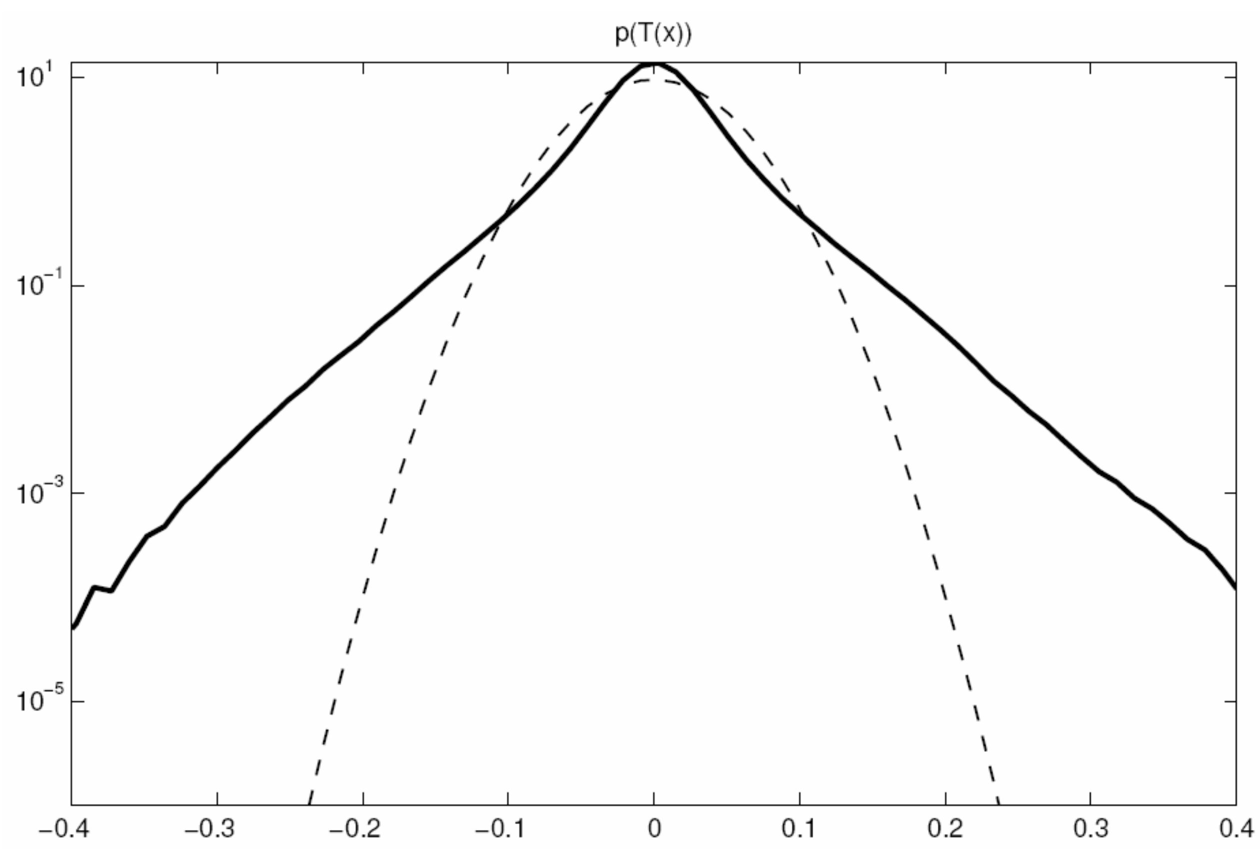
Gershgorin, Majda (2010),
Bourlioux, Majda (2002)

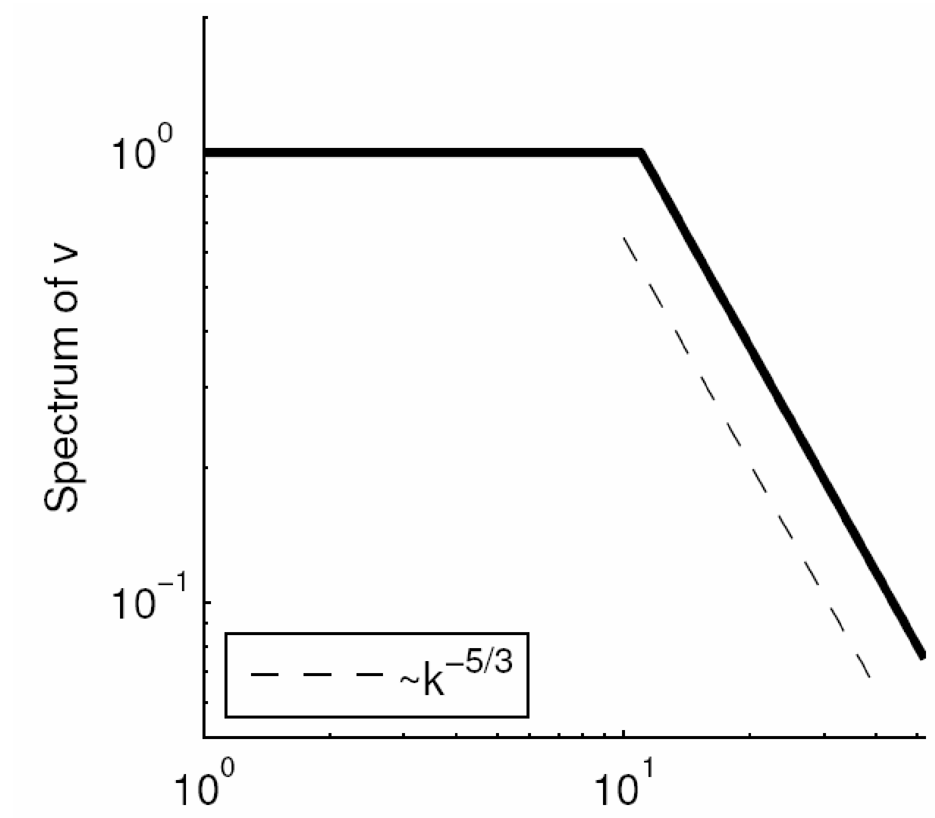
$U^M, \; v^M, \; T^M,$ solutions with Model Error

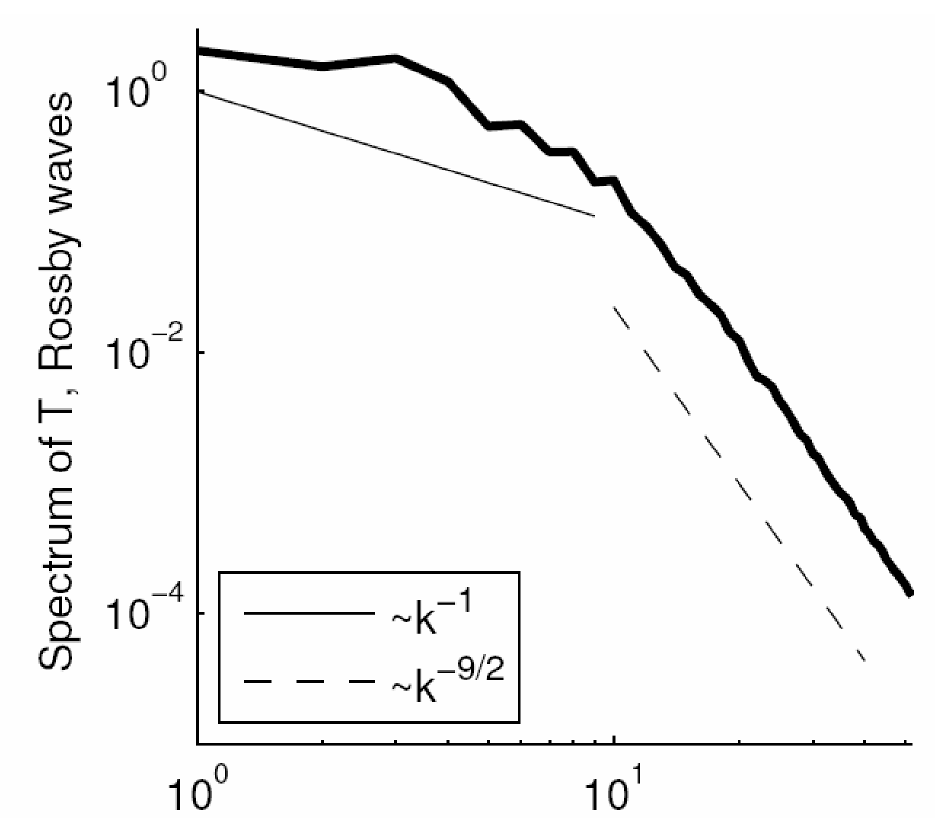
Mimic GCM: increase damping, $\gamma_M$, eddy duffusivity for $T^M$

Mean zonal jet



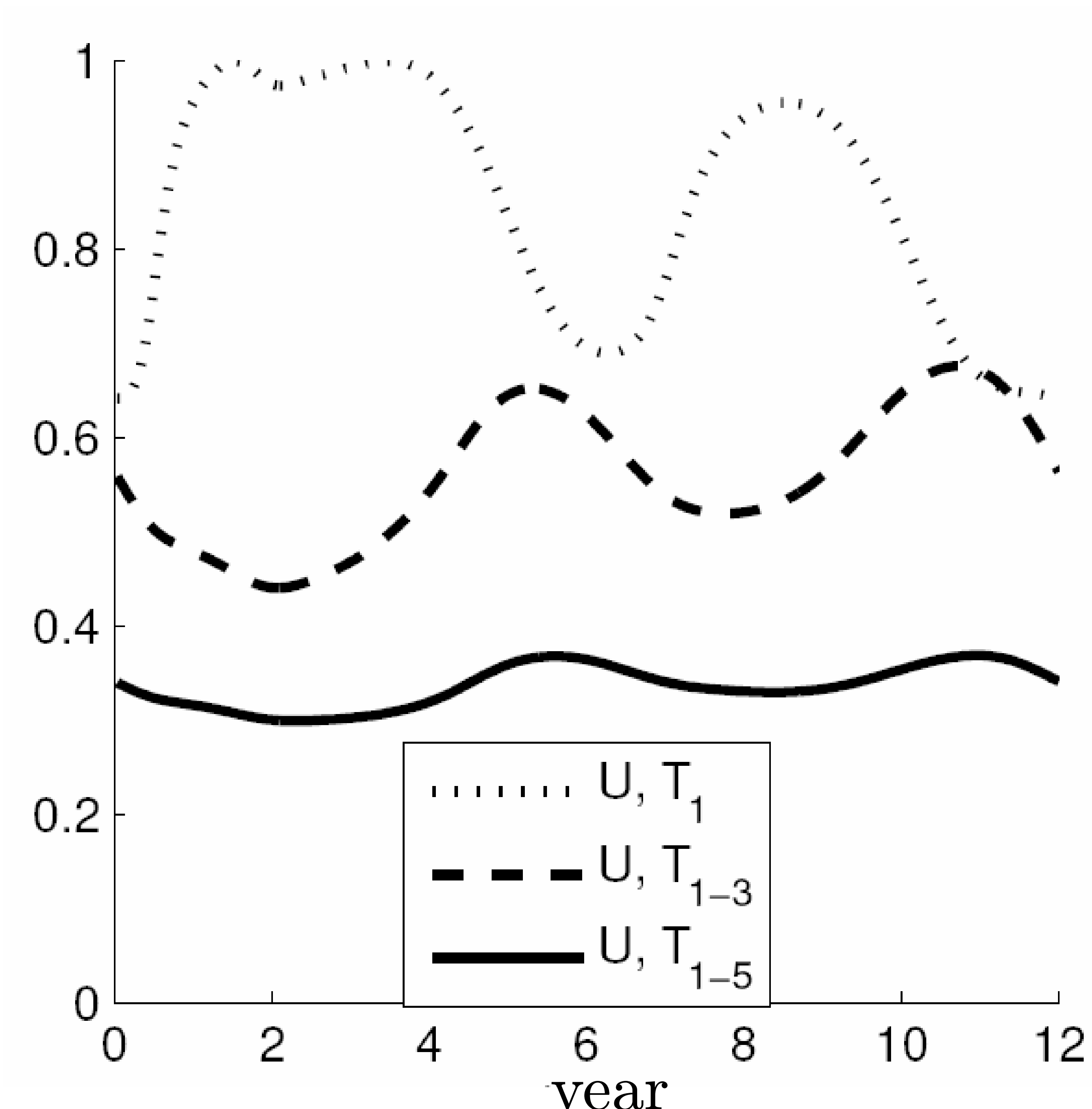




Pdf for $T$ like atmospheric tracers in observations, Neelin et al (2010)

Fraction of the signal part in the total lack of information $\mathcal{P}$



# Stochastic Model Parameterization

eddy diffusivity — space-time white noise

$$\frac{\partial T_M}{\partial t} + \bar{U}_M(t)\frac{\partial T_M}{\partial x} = -\alpha v_M(x,t) + (\kappa + \kappa_M)\frac{\partial^2 T_M}{\partial x^2} - d_T T_M + \sigma_T \dot{W}(x,t).$$

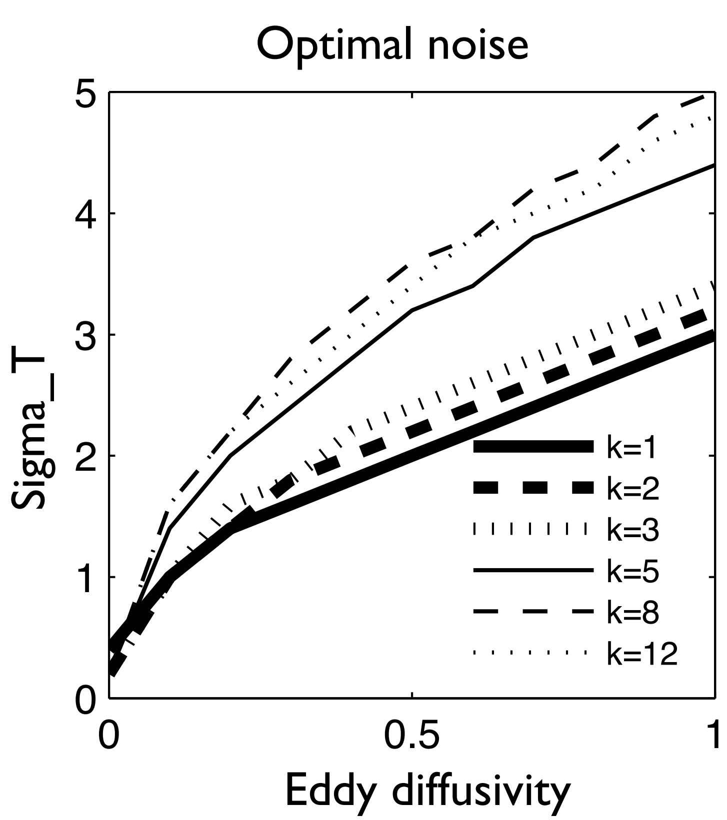
## Systematic information-theoretic improvement with optimal stochastic forcing:

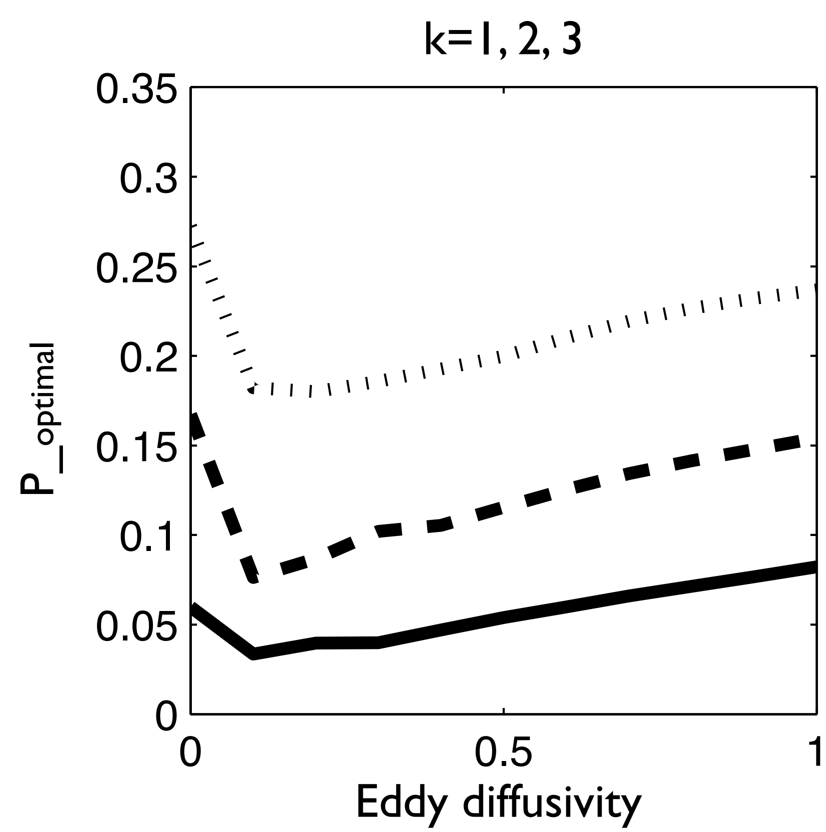
1) Optimal noise increases as model error eddy-diffusivity increases

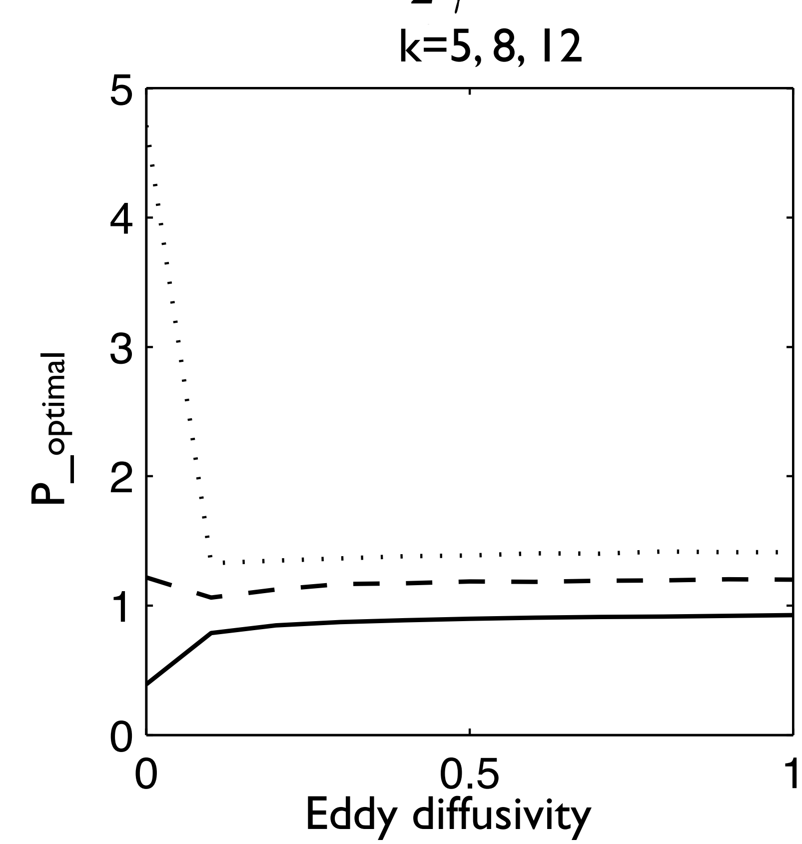
2) Larger noise is needed for refined coarse-grainings

3) There is a significant information gain in the dispersion with optimal stochastic parameterization vs deterministic one

4) Optimal value with the smallest information is at $\kappa_M = 0.1\kappa_M^*$ where $\kappa_M^* = \frac{\sigma^2}{2\gamma^2}$



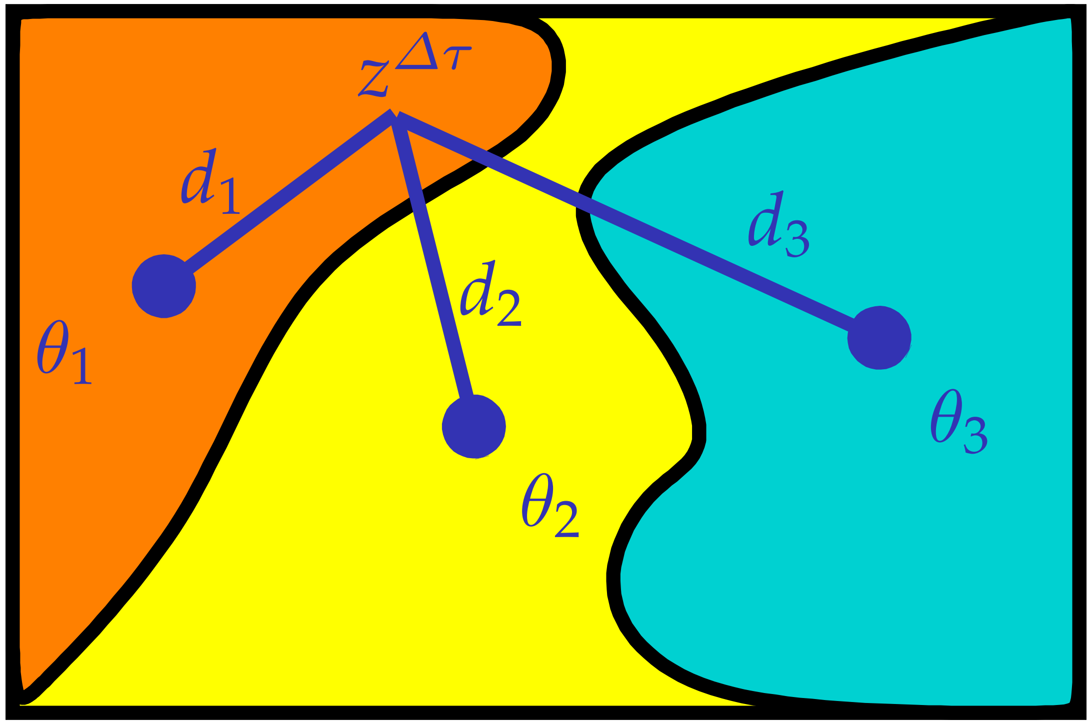



# Revealing long-range predictability and model error through coarse-grained partitions of phase space

- Giannakis & Majda (2011), Quantifying the predictive skill in long-range forecasting. Part I: Coarse-grained predictions in a simple ocean model, submitted to *J. Climate*.
- Giannakis & Majda (2011), Quantifying the predictive skill in long-range forecasting. Part II: Model error in coarse-grained Markov models with application to ocean-circulation regimes, submitted to *J. Climate*.
- Giannakis, Majda & Horenko (2011), Information theory, model error, and predictive skill of stochastic models for complex nonlinear systems, submitted to *Physica D*.

# Strategies for phase-space partitioning

$n$-dim. space of initial data




Each cluster is characterized by its centroid, $\theta_k$.

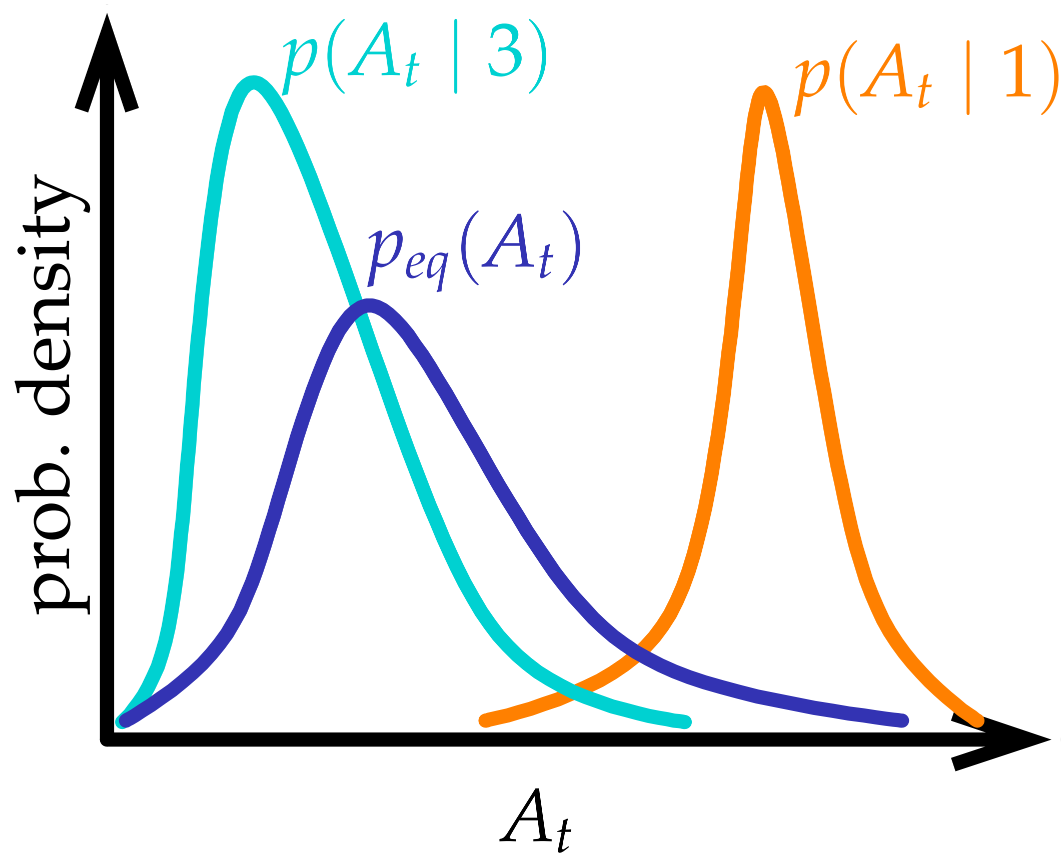
1. Collect observations $z(t)$ over a time window $\Delta\tau$ and compute the average,

$$z^{\Delta\tau} = \frac{1}{\Delta\tau} \int_{\Delta\tau} dt\, z(t).$$

2. Set $S$ equal to the cluster that lies closest to $z^{\Delta\tau}$, i.e.,

$$S = \operatorname*{argmin}_k d_k, \quad d_k = \|z^{\Delta\tau} - \theta_k\|.$$

# The predictive information content in a partition




Predictive skill given that the initial data lie in the $k$-th cluster:

$$\mathcal{D}_t^k = \mathcal{P}(p_t^k, p_{eq}), \quad p_t^k(A) = p(A_t \mid S = k).$$

"Super-ensemble" measure of skill:

$$\mathcal{D}_t = \sum_{k=1}^{K} \pi_k D_t^k, \quad \pi_k = p(S = k).$$

## Interpretation

- $\mathcal{D}_t$ is equal to the mutual information $I(A_t; S)$ between the coarse-grained initial data $S$ and the value $A_t$ of the prediction observable at time $t$.
- $\mathcal{D}_t$ vanishes if and only if $S$ and $A_t$ are statistically independent; namely, in the $t \to \infty$ limit.

# Mathematical Strategies for Filtering Turbulent Dynamical Systems

**By Andrew J. Majda**
Morse Professor of Arts and Sciences
Center for Atmosphere/Ocean Science (CAOS)
Courant Institute of Mathematical Sciences (CIMS)
New York University

**Main Collaborators:**
John Harlim, North Carolina State University
Boris Gershgorin, CIMS, New York University (postdoc)
Marcus Grote, University of Basel

Supported by National Science Foundation, Office of Naval Research and DARPA

## Modern Applied Modus Operandi

**Theory**: Important mathematical guidelines
Qualitative Exactly Solvable Models

**Novel Algorithms:**
**Applications to Real Problems in Science/Engineering**

**General Refs for Talk:** Research/Expository

*A. Majda, J. Harlim, and B. Gershgorin "Mathematical Strategies for Filtering Turbulent Dynamical System" 2010, Dis. Cont. Dyn. Sys., 27, pp 441-486*

**Introductory Graduate Text**

*A. Majda and J. Harlim, "Mathematical Strategies for Real Time Filtering of Turbulent Signals in Complex Systems," Cambridge University Press (2011)*

## Exactly Solvable Test Models and NEKF Algorithms

Prototype Test Problems which are Nonlinear yet exactly solvable statistically for filtering multiple time scale systems

Examples: Gravity Waves, Moisture, and Large Scale Flow in Tropics or Mesoscale, Tracking hazardous pollutants in real time from partial observations

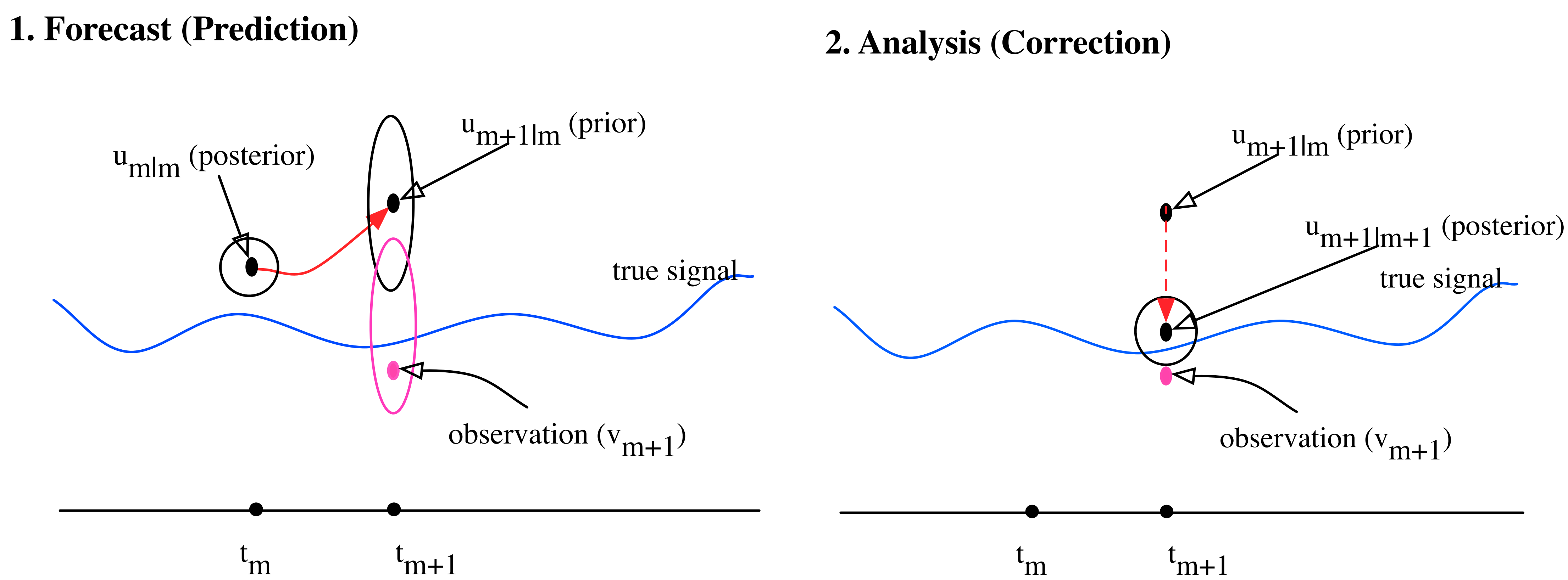
**References:**

1. B. Gershgorin and A. Majda, 2008, "**A nonlinear test model for filtering slow-fast systems**," *Comm. Math. Sci.*, 6, 3, pp. 611 – 649

2. B. Gershgorin and A. Majda, 2010, "**Filtering a nonlinear slow-fast system with strong fast forcing**," *Comm. Math. Sci.* 8, 1, pp. 67 – 92

3. B. Gershgorin and A. Majda, 2011, "**Filtering a statistically exactly solvable test model for turbulent tracers from partial observations**," *J. Comp. Phys*, Vol. 230, February 2011, pp 1602-1638

# What is filtering?




The correction step is an application of Bayesian update

$$p(u_{m+1|m+1}) \equiv p(u_{m+1|m}|v_{m+1}) \sim p(u_{m+1|m})p(v_{m+1}|u_{m+1|m})$$

Kalman filter formula produces the optimal unbiased posterior mean and covariance by assuming linear model and Gaussian observations and forecasts errors.

## Theoretical and Computational Issues:

- Handling nonlinearity! Why not particle filter? Convergence requires ensemble size that grows exponentially with respect to the ensemble spread relative to observation errors rather than to the state dimension per se(Bengtsson, Bickel, and Li 2008).
- How to handle large system? Perhaps $N = 10^6$ state variables (e.g., 200 km resolved Global Weather Model)
- Where is the computational burden? Propagating covariance matrix of size $N \times N$ (6N minutes = 300,000 hours).
- Some successful strategies: Ensemble Kalman filters (ETKF of Bishop et al. 2001, EAKF of Anderson 2001). Each involves computing singular value decomposition (SVD).
- However, these accurate filters are not immune from "catastrophic filter divergence" (diverge beyond machine infinity) when observations are sparse, even when the true signal is a dissipative system with "absorbing ball property".

# Mean Stochastic Model

The prototype one-mode stochastic mean model

$$du(t) = \Big[(-\bar{\gamma} + \mathrm{i}\omega)u(t) + F(t)\Big]dt + \sigma dW(t)$$

where one fits the parameters using climatological statistical quantities such as the energy spectrum and correlation time.

This "poor-man" strategy is discussed in Harlim and Majda Nonlinearity 2008, Comm. Math. Sci. 2010.

# Stochastic Parameterized Extended Kalman Filter:

We consider the following canonical model that accounts additive and multiplicative biases:

$$
\begin{aligned}
du(t) &= \Big[(-\gamma(t) + i\omega)u(t) + F(t) + b(t)\Big]dt + \sigma dW(t) \\
db(t) &= (-\gamma_b + i\omega_b)b(t)dt + \sigma_b dW_b(t) \\
d\gamma(t) &= -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t)
\end{aligned}
$$

We find stochastic parameters $\{\gamma_b, \omega_b, \sigma_b, d_\gamma, \sigma_\gamma\}$ that are robust for high filter skill beyond the MSM and in many occasions comparable to the perfectly specified filter model.

This special form has exactly solvable nonlinear solutions and moments and we do not need any linearization as in the standard EKF.

**References:**

a) Test Models for Improving Filtering with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):1-31, 2010.

b) Improving Filtering and Prediction of Spatially Extended Turbulent Systems with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):32-57, 2010.

c) Filtering Turbulent Sparsely Observed Geophysical Flows (with A.J. Majda), Monthly Weather Review, 138(4): 1050-1083, 2010.

**Review Article:** Mathematical Strategies for Filtering Turbulent Dynamical Systems (with B. Gershgorin and A.J. Majda), DCDS-A: 27(2), 441-486, 2010.

**Ch 13 of:** Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems (with A.J. Majda), Cambridge University Press (in preparation), 2010.

# Test model for true signal

Consider the following SDE

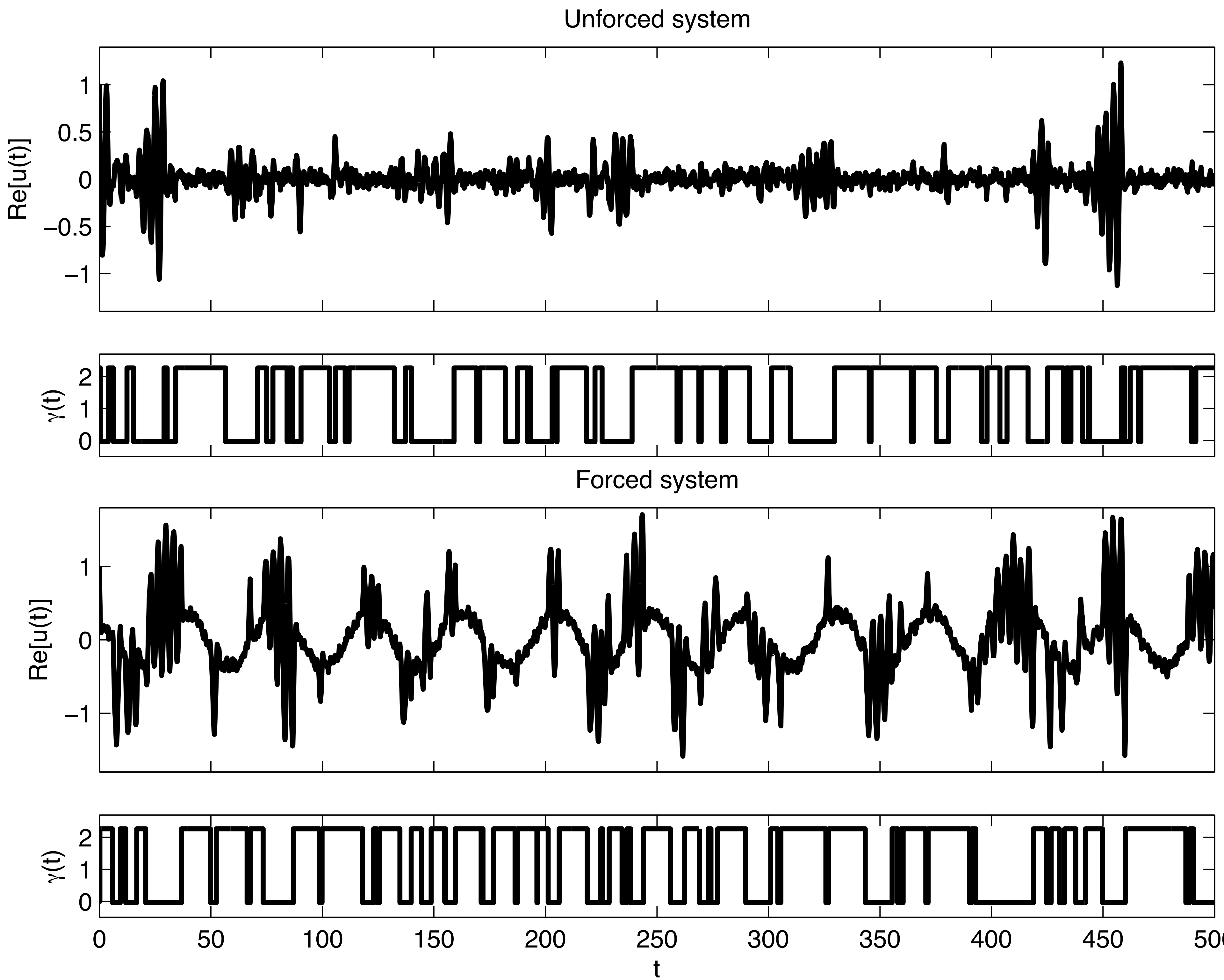
$$\frac{du(t)}{dt} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma \dot{W}(t) + f(t)$$

as a test model for filtering with model error.

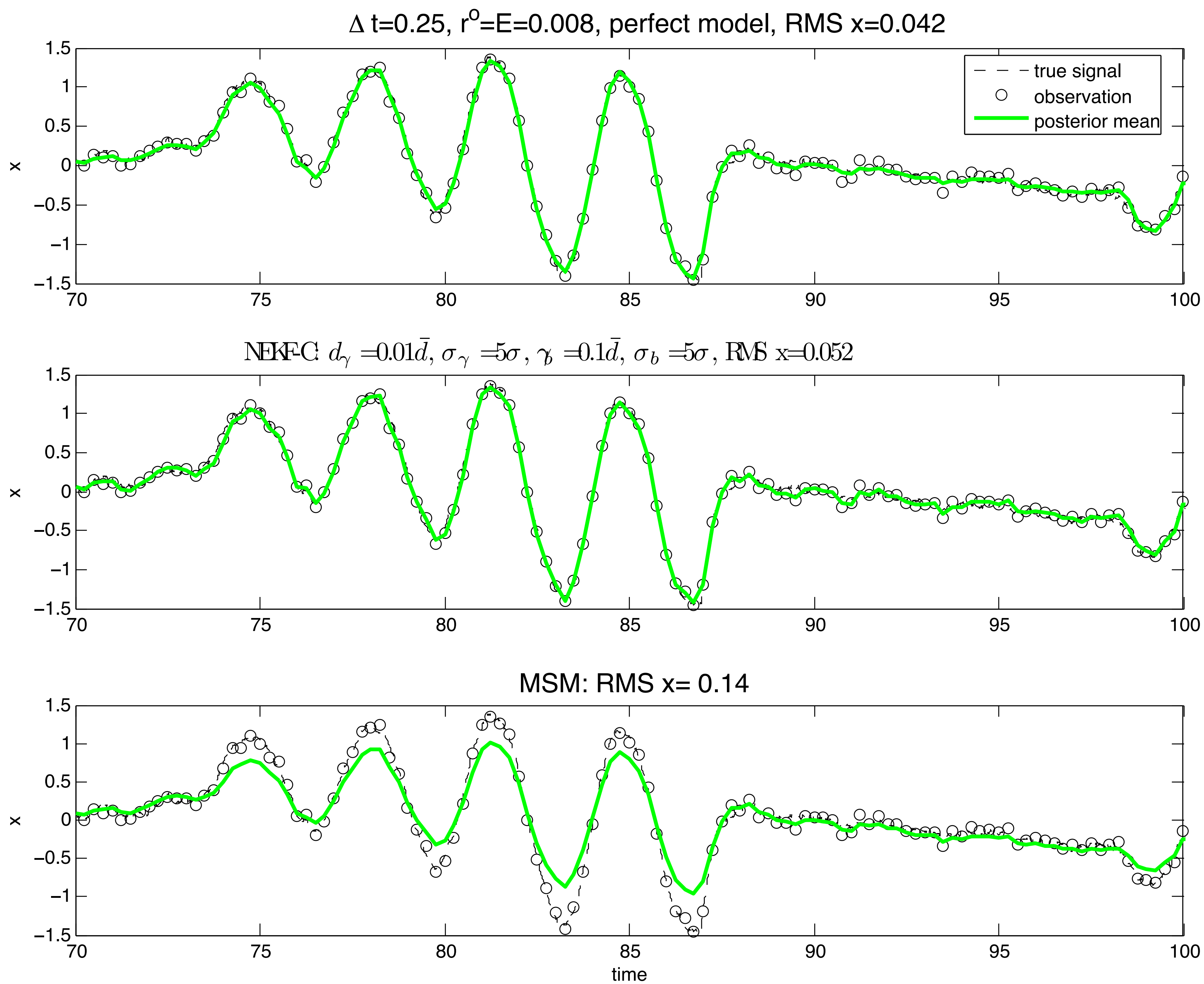
To generate significant model errors as well as to mimic intermittent chaotic instability as often occurs in nature, we allow $\gamma(t)$ to switch between stable ($\gamma > 0$) and unstable ($\gamma < 0$) regimes according to a two-state Markov jump process.
Assume the following observation model:

$$v_m = u(t_m) + \sigma_m^o, \quad \sigma_m^o \sim \mathcal{N}(0, r^o). \tag{1}$$

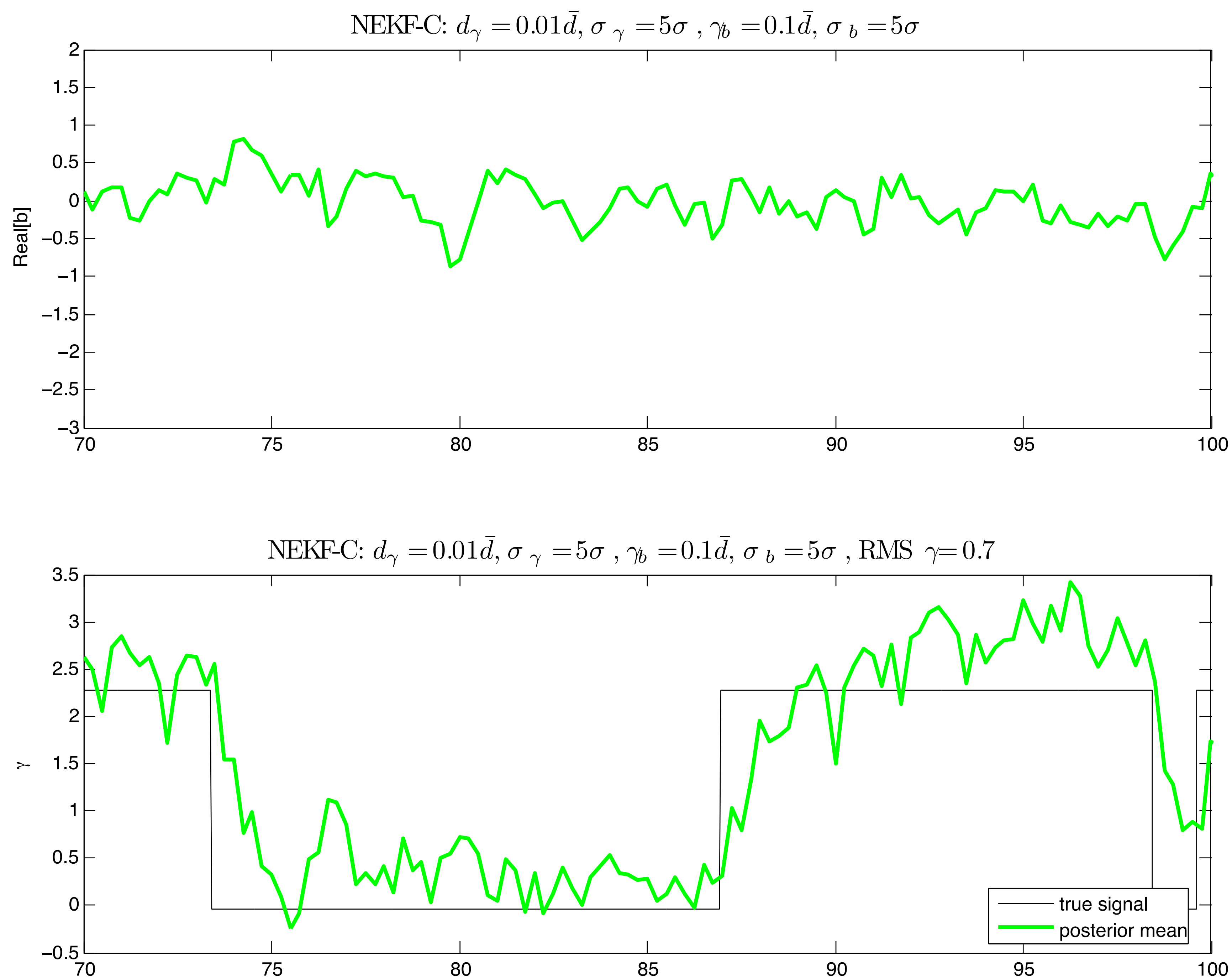
# True Signals for Unforced and Forced cases



# One mode demonstration of the filtered solution: observed mode



# One mode demonstration of the filtered solution: unobserved parameters



# Canonical Spatially Extended Turbulent Systems

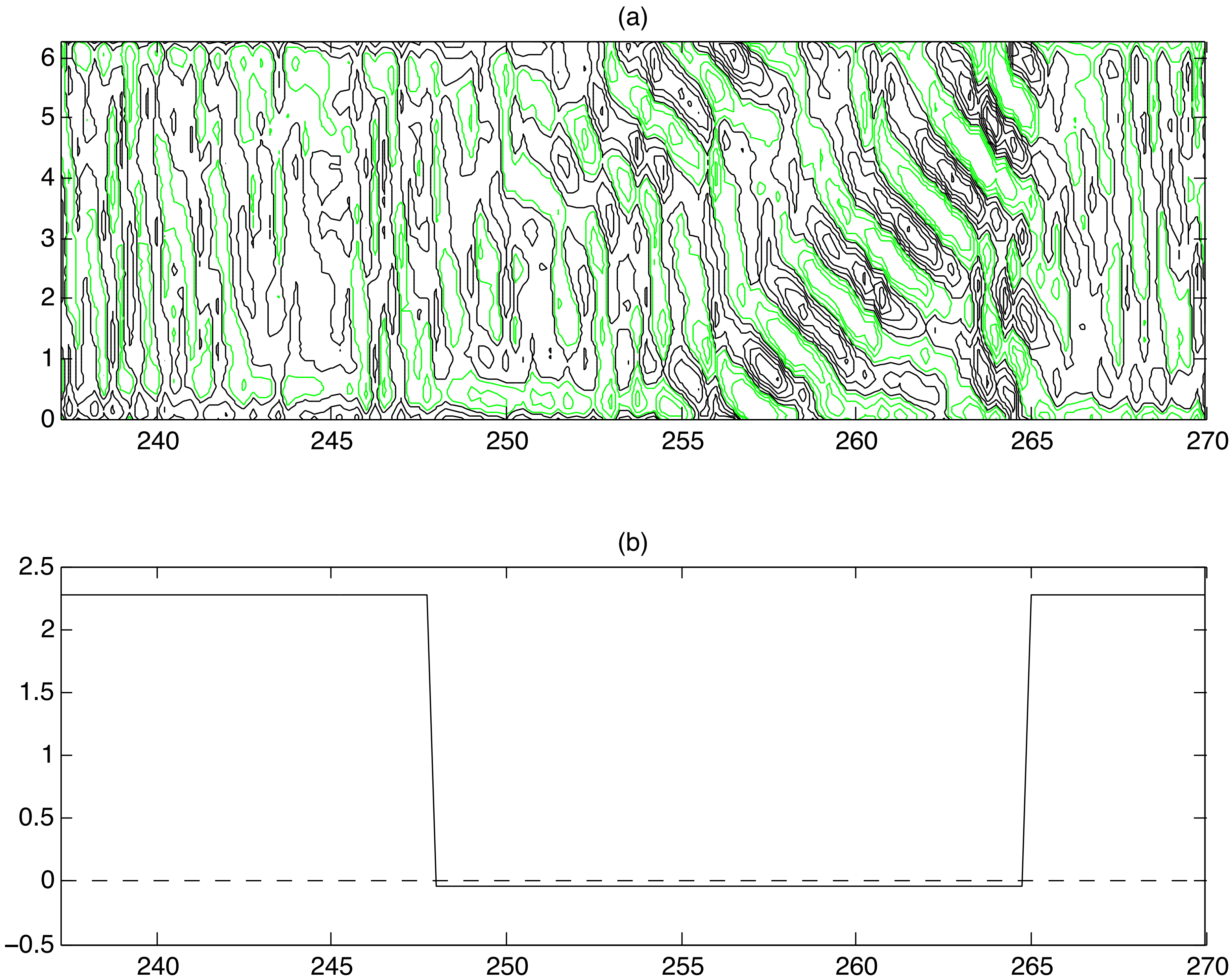
We consider a stochastic PDE with time-dependent damping Langevin equation for the first five Fourier modes, i.e.,

$$\frac{du_k(t)}{dt} = -\gamma_k(t)u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k = 1, \ldots, 5,$$

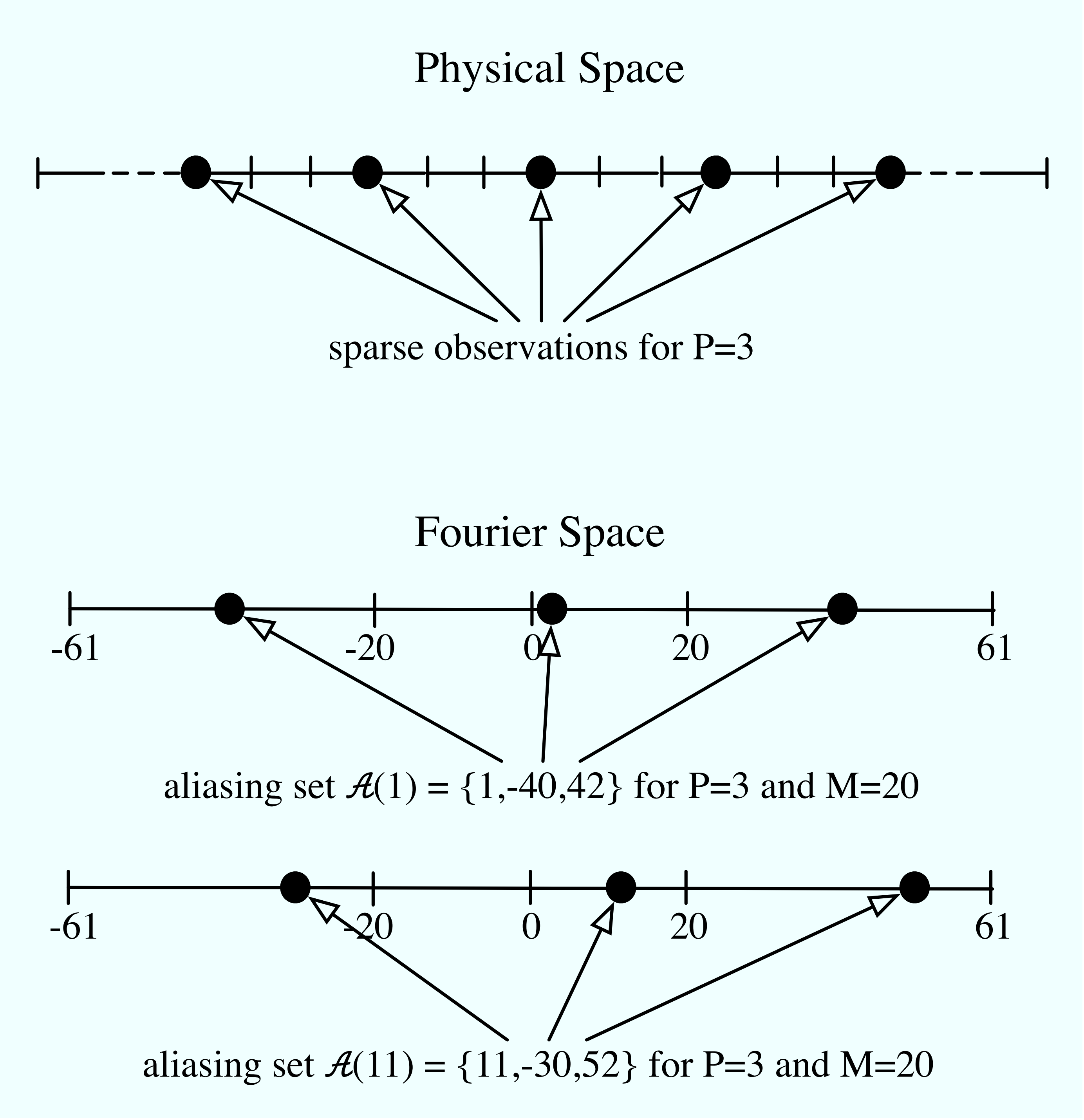
and linear Langevin equation with constant damping $\bar{d}$ for modes $k > 5$,

$$\frac{du_k(t)}{dt} = -\bar{d}u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k > 5.$$

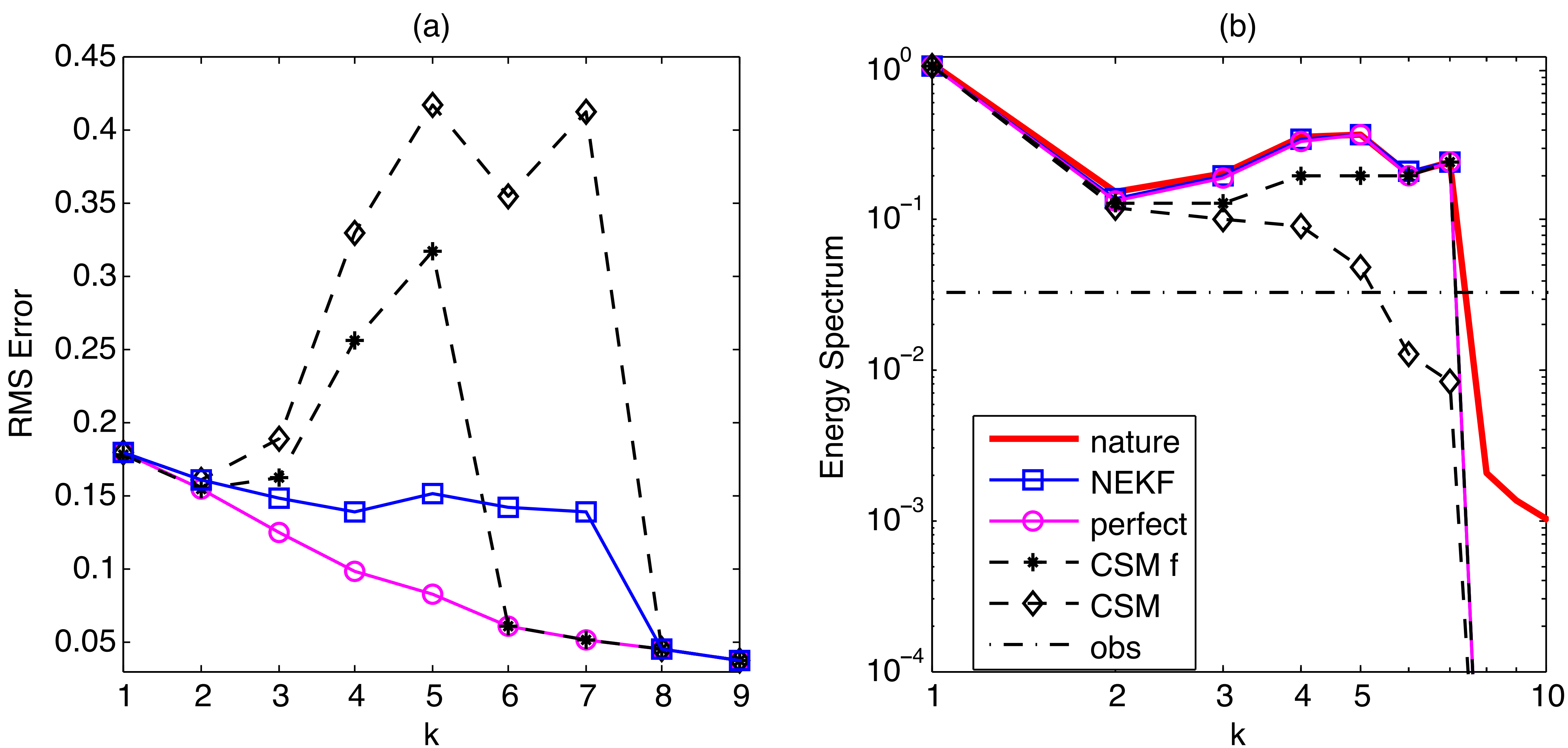
# Turbulent barotropic Rossby wave equation: $\omega_k = -\beta/k, E_k = k^{-3}$



Example: 123 grid pts (61 modes) but only 41 observations (20 modes) available



# Incorrectly specified forcings, observed only 15 observations of 105 grid points



# Canonical Model for Midlatitude Geophysical Flows:

The dynamical equations for the perturbed variables are:

$$\begin{aligned} \frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} + \nu\nabla^8 q_1 &= 0 \\ \frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial x} + \nu\nabla^8 q_2 + \kappa\nabla^2\psi_2 &= 0 \end{aligned}$$

where $q_j$ is the quasi-geostrophic potential vorticity given as

$$q_j = \beta y + \nabla^2\psi_j + \frac{k_d^2}{2}(\psi_{3-j} - \psi_j)$$

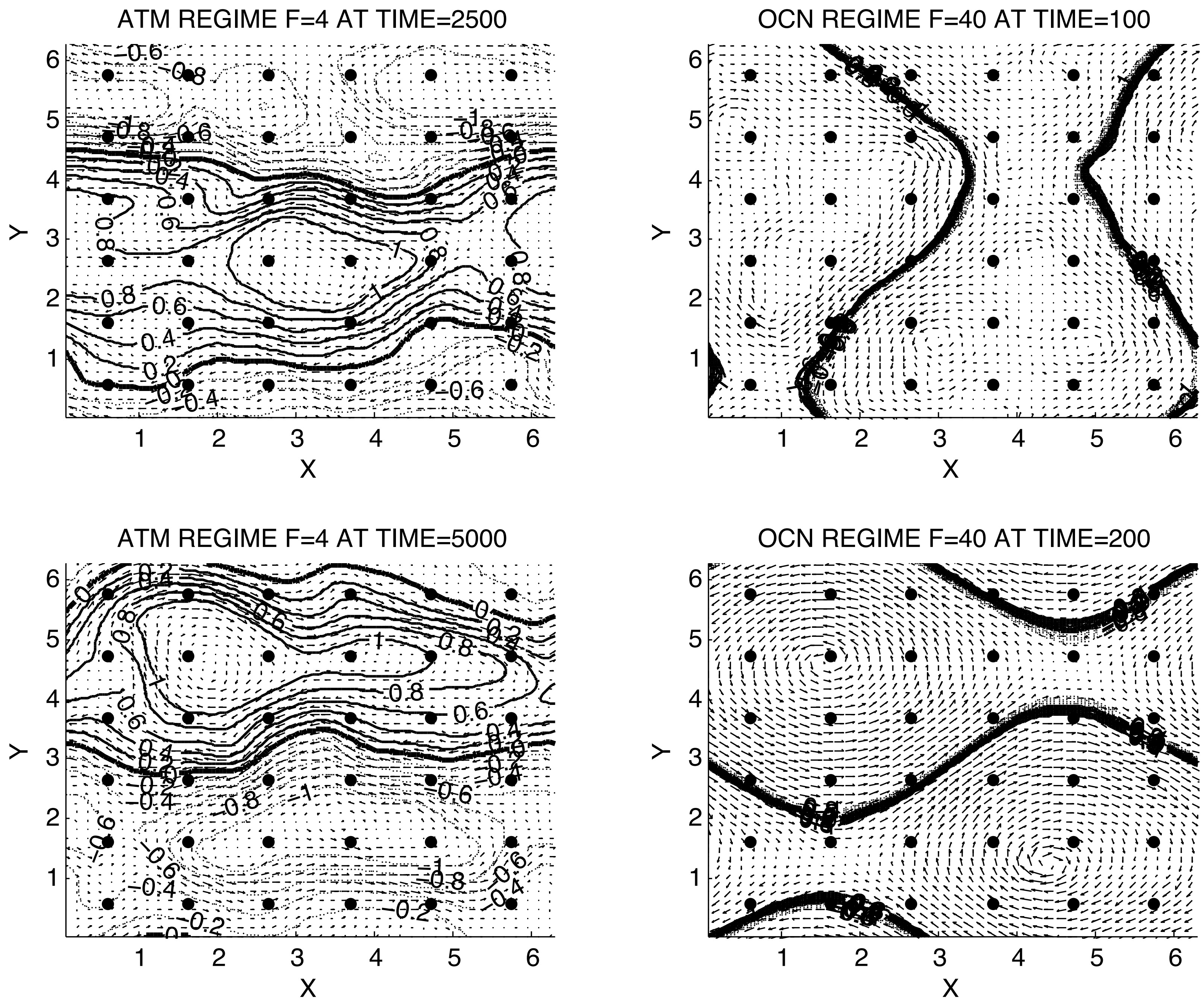
with $\vec{u} = \nabla^\perp\psi$, $k_d = \sqrt{8}/L_d$.

In the two-layer case, the barotropic vertical and baroclinic modes are defined as $\psi_b = (\psi_1 + \psi_2)/2$ and $\psi_c = (\psi_1 - \psi_2)/2$, respectively.
Notice that the barotropic mode dynamical equation,

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b$$
$$+ \left( J(\psi_c, q_c) + U \frac{\partial \nabla^2 \psi_c}{\partial x} - \kappa \nabla^2 \psi_c \right) = 0$$

is numerically stiff when $k_d^2$ is large (ocean case).

# The 2-layer QG model with baroclinic instability



# Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta\frac{\partial \psi_b}{\partial x} + \kappa\nabla^2\psi_b + \nu\nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2\psi_b$.

**Poorman's stochastic models:** replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes.

**Discrete Fourier Transform:**

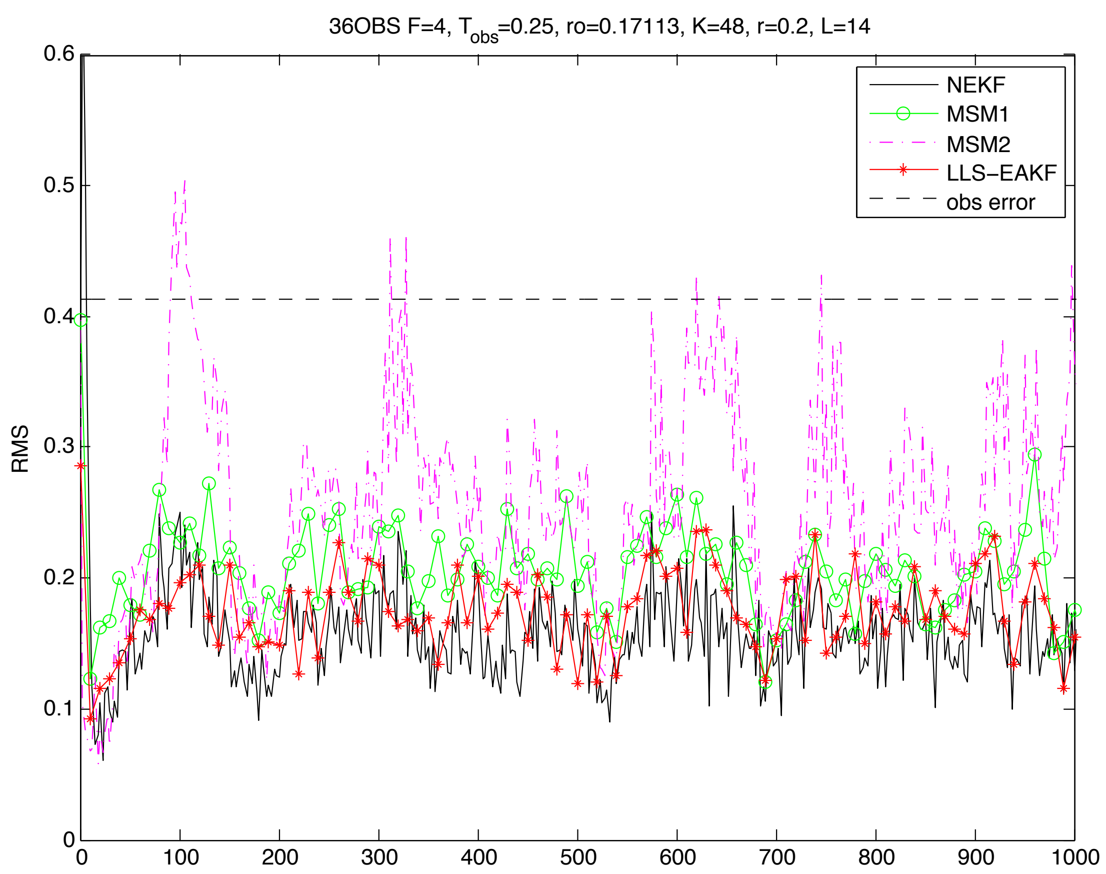
$$\psi = \sum_{k,\ell} \hat{\psi}_{k,\ell} e^{\mathrm{i}(kx+\ell y)}$$

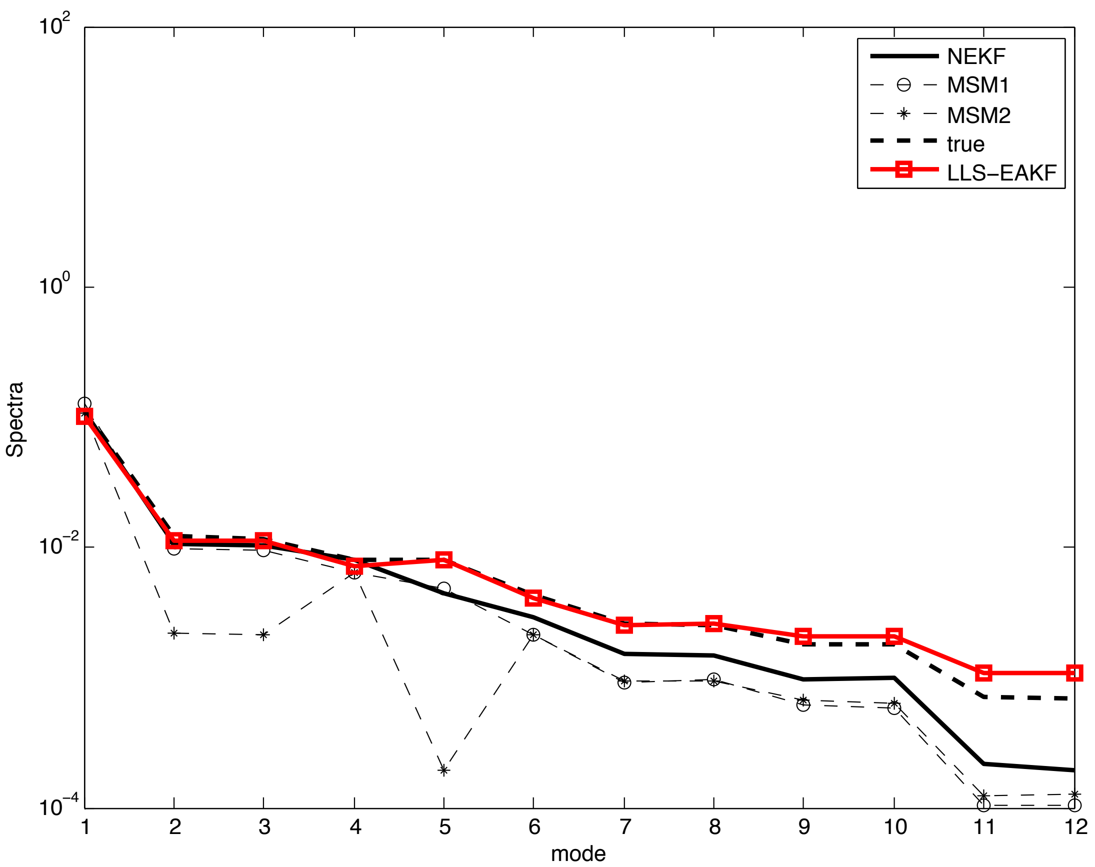
Thus, each horizontal mode has the following form

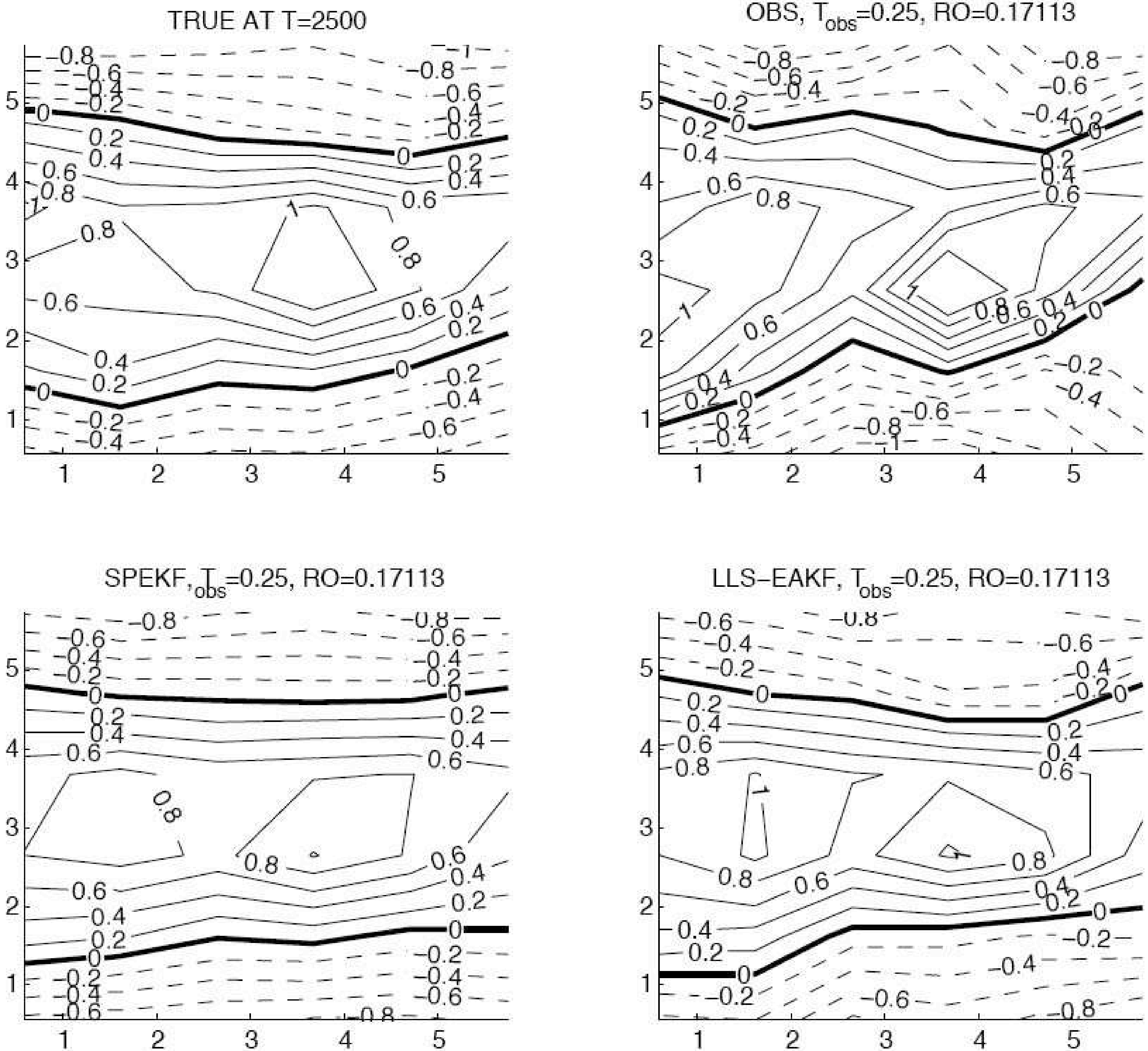
$$d\hat{\psi}(t) = (-d + \mathrm{i}\omega)\hat{\psi}(t)dt + f(t)dt + \sigma dW(t)$$

and our task is to parameterize $d, \omega, f(t), \sigma$?

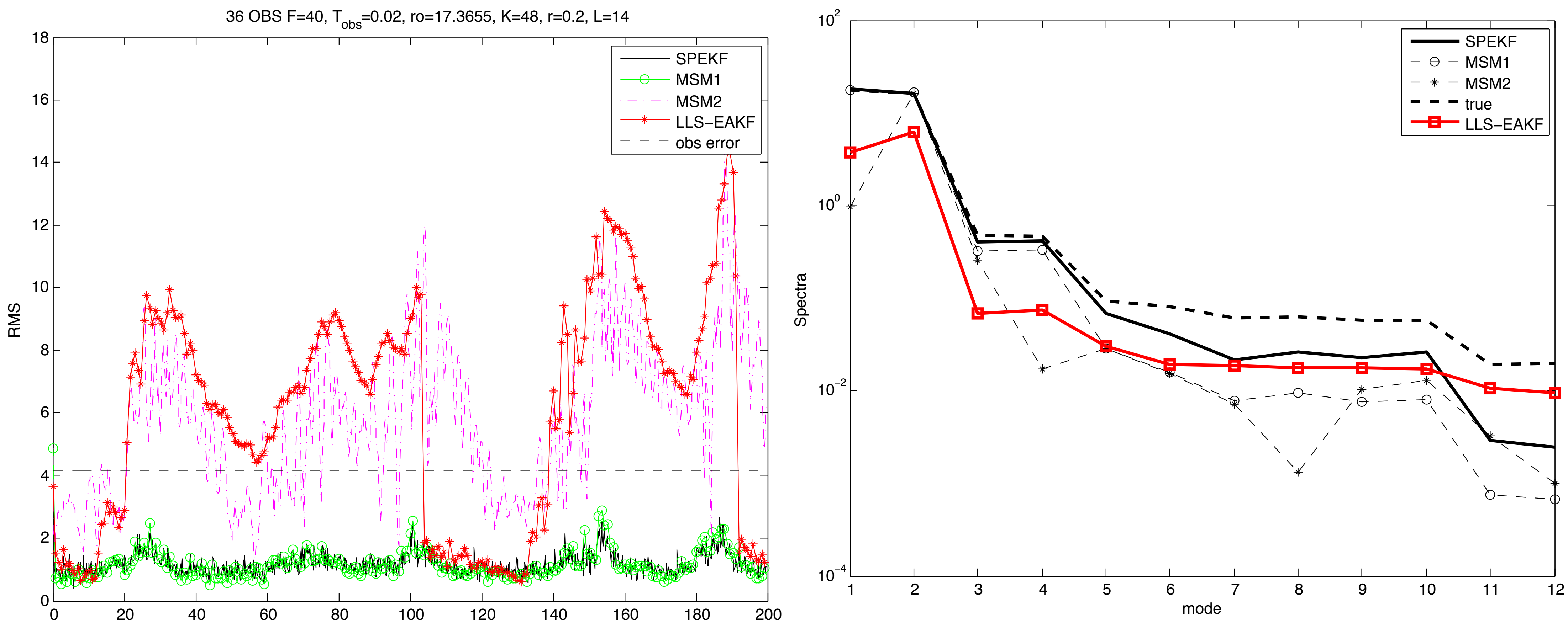
# Longer deformation radius case ("atmospheric" regime).

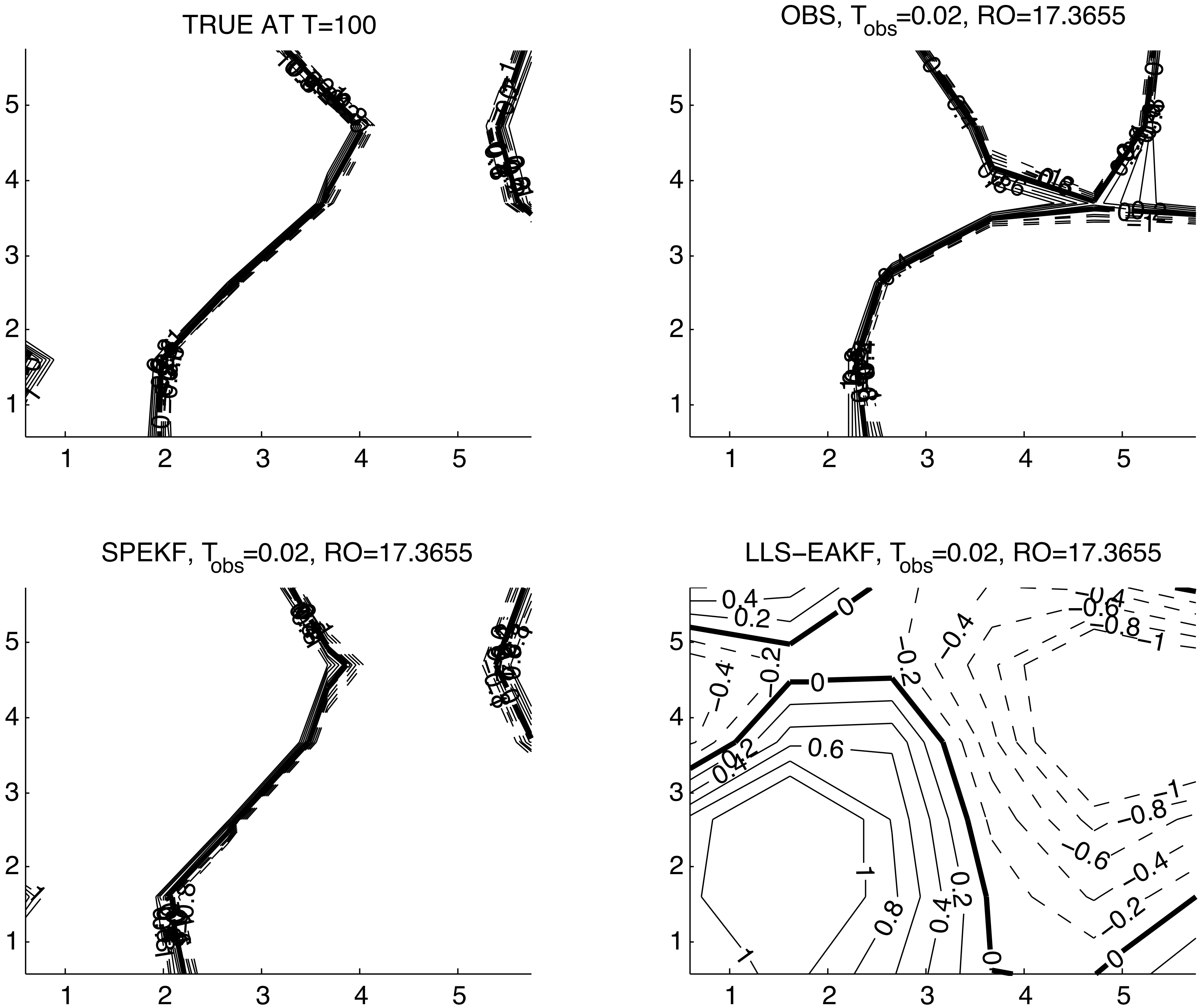



TRUE AT T=2500
OBS, $T_{obs}$=0.25, RO=0.17113
SPEKF, $T_{obs}$=0.25, RO=0.17113
LLS−EAKF, $T_{obs}$=0.25, RO=0.17113

# Shorter deformation radius case ("oceanic" regime).

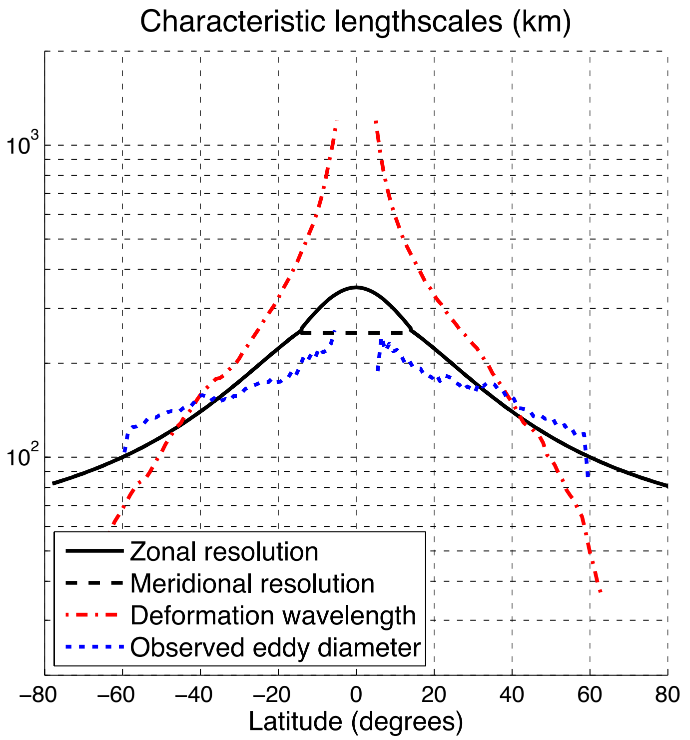



TRUE AT T=100
OBS, $T_{obs}$=0.02, RO=17.3655
SPEKF, $T_{obs}$=0.02, RO=17.3655
LLS−EAKF, $T_{obs}$=0.02, RO=17.3655
1
2
3
4
5

# Summary:

1. MSM: We introduce reduced stochastic models through replacing the nonlinearity and baroclinic components with Ornstein-Uhlenbeck process for filtering purpose. This reduced poor man's strategy is numerically very cheap and accurate in a regime when the dynamical systems is strongly chaotic and fully turbulent.
2. SPEKF: We introduce a paradigm model for "online" learning both the additive and multiplicative biases from observations beyond the MSM. This model is analytically solvable such that NO LINEARIZATION is needed when Kalman filter formula is utilized.

# Stochastic Super-resolution: Estimating turbulent heat transport in the ocean using satellite altimetry




"New methods for estimating poleward eddy heat transport using satellite altimetry"

Shane R. Keating, **Andrew J. Majda** & K. Shafer Smith

*J. Phys. Oceanogr.* 2011 (submitted)