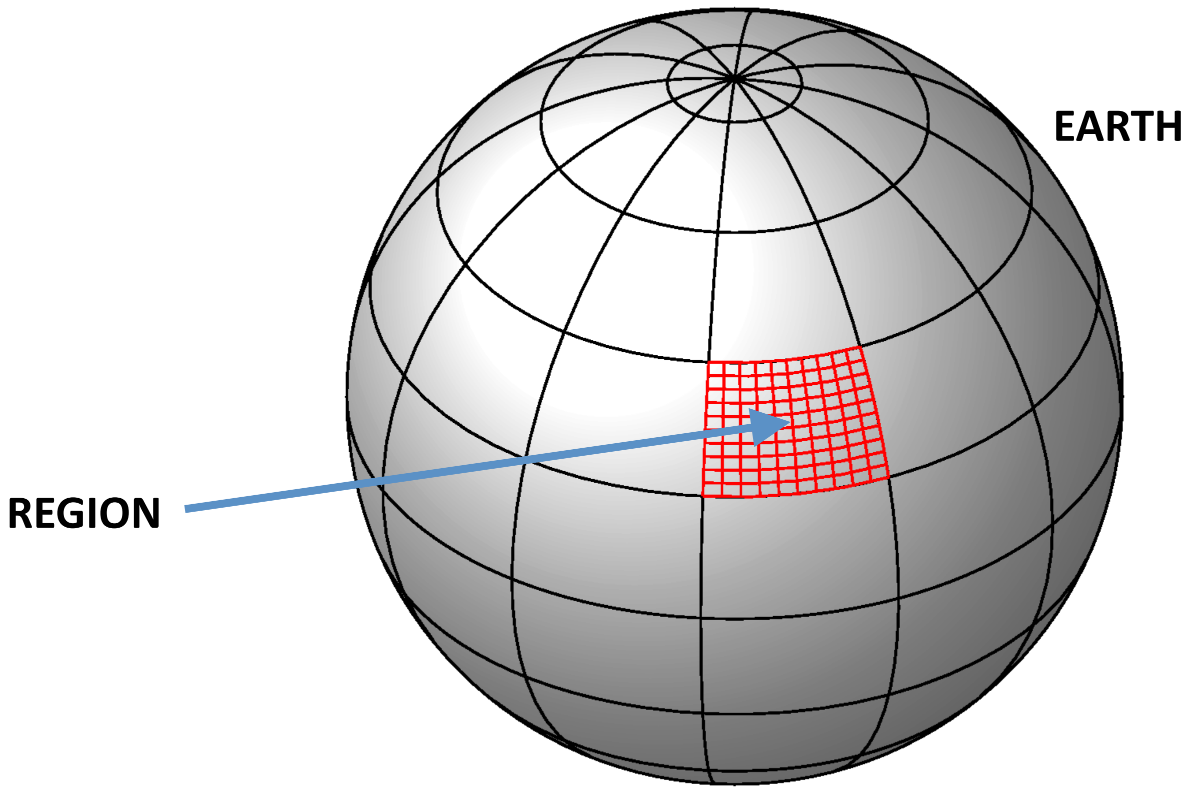
USING A LIMITED AREA MODEL TO ENHANCE GLOBAL ANALYSES

Edward Ott
University of Maryland

COWORKERS

Young-noh Yoon (UMD)
Istvan Szunyogh (TEX A&M)
Brian Hunt (UMD)

REGIONAL MODELS



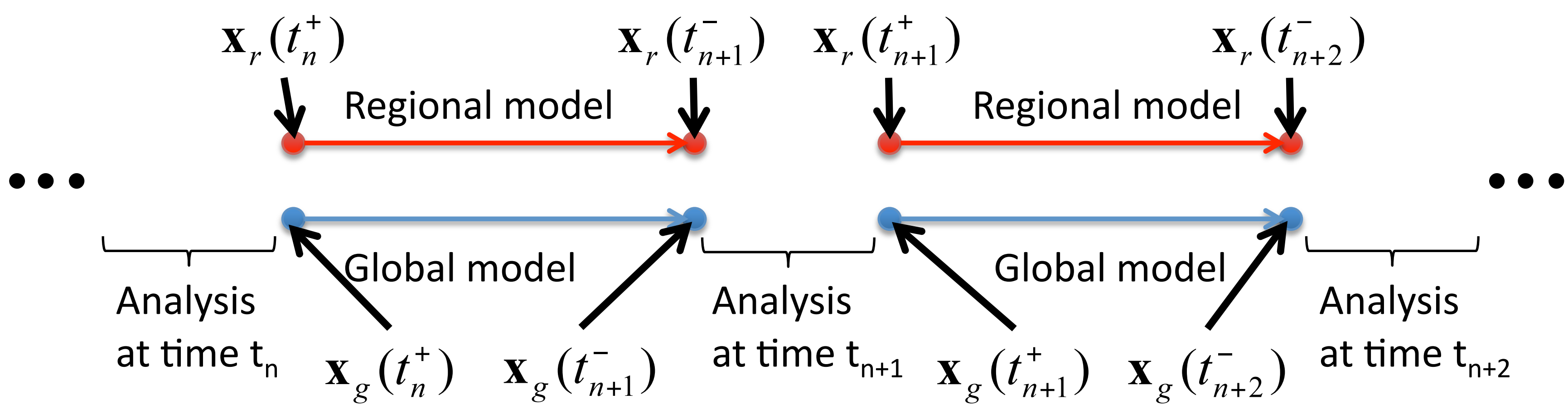
GLOBAL MODEL → REGIONAL MODEL
BOUNDARIES

REGIONAL MODEL : SUBSTANTIALLY HIGHER RESOLUTION
THAN THE GLOBAL MODEL

THE ISSUE OF SPATIAL RESOLUTION

- **A true atmospheric state is spatially continuous.**
- **The model state is a spatial discretization that is hypothesized to give approximate information on the true state.**
- **The model cannot fully represent processes associated with scales that are unresolved by its grid.**
- **The effect of subgrid scale processes on larger scales can only be represented statistically.**
- **Statistical models of subgrid scale are typically *ad hoc*. (E.g., the Reynolds decomposition/averaging procedure for subgrid scale modeling requires an assumption like a closure hypothesis.)**
- **Some examples of small spatial scale processes that are often modeled ('parametrized') are turbulent diffusion, orographic drag, clouds, etc.**
- **Implication : Higher resolution models have the potential to give a more accurate representation of atmospheric dynamics.**

FORECAST CYCLES



$\mathbf{x}_r(t)$ and $\mathbf{x}_g(t)$: **The spatially discretized state variables (temp, pressure, velocity, etc.) at time t**

Global model : $\mathbf{x}_g(t_{n+1}^-) = \mathrm{g}[\mathbf{x}_g(t_n^+)]$

Regional model : $\mathbf{x}_r(t_{n+1}^-) = \mathrm{r}[\mathbf{x}_r(t_n^+), \mathbf{x}_g(t) \text{ for } t_n^+ \le t < t_{n+1}^-]$

- **In evolution from t_n^+ to t_{n+1}^- global affects regional but not vice versa.**
- **In analysis we can consider global ⇄ regional.**

COMMON CURRENT PRACTICE IN REGIONAL DATA ASSIMILATION

At each forecast cycle the initial condition for the global model is obtained independent of the regional model state and the initial condition for the regional is usually estimated in one of two ways :

(A) Interpolation of the global analysis onto the finer scale regional model grid.

(B) A separate data assimilation specifically designed to produce initial conditions for the regional model.

Motivation: Most weather prediction centers that do regional prediction have access to global analyses, but cannot simultaneously do analysis/initialization/running of the global code. (There are a few exceptions; e.g., NCEP.)

COUPLING OF GLOBAL AND REGIONAL ANALYSES

- Here we assume that global and regional analyses and initializations can be done together (a possibility at large centers like NCEP).
- We formulate and test a candidate technique for doing this*.
- Our technique is a generalization of the Local Ensemble Transform Kalman Filter (LETKF) described in the Thursday morning talk by Szunyogh. (An important point is that ensemble techniques provide a natural way of propagating uncertainty in the global-model-determined boundary conditions to the local model.)
- Motivation: One might expect that the global and regional analyses would both benefit from information exchange between them.
- Main conclusion: Both global and regional forecasts are improved.

*Reference: Y. Yoon, B. R. Hunt, E. O., I. Szunyogh, "Regional Ensemble Data Assimilation Using a Joint State Method", (arXiv).

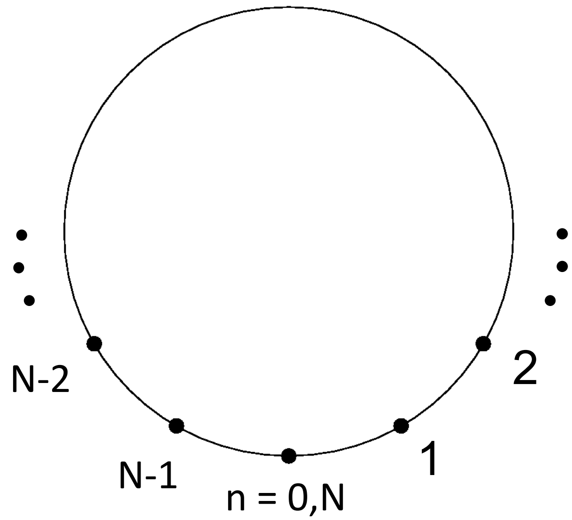
OUTLINE OF THE REST OF THIS TALK

- **First, we describe simple 1D 'toy' models introduced by Lorenz (2005)*, which we will use to illustrate and test the problem we address.**
- **Second, we briefly describe the LETKF data assimilation method.**
- **Next, we specify a proposed procedure for data assimilation simultaneously using state results from the global and the regional state time intergrations, and treating the two on an equal footing.**
- **Finally, we report numerical tests and comparisons between our assimilation technique [(global)⇄(regional)] and previous common practice [(global)→(regional)].**

***E. N. Lorenz, "Designing Chaotic Models", J. Atm. Sci. 62, 1574 (2005).**

LORENZ'S (2005) MODELS OF ATMOSPHERIC CHAOS

Evolution equations for a scalar $Z_n(t)$, where n = 0, 1, 2, …, N-1 represents points on a one dimensional spatial grid with periodic boundary conditions, $Z_0(t) = Z_n(t)$.



- **Lorenz devised three models of successively increasing sophistication, which we call Lorenz model 1, Lorenz model 2, and Lorenz model 3.**
- **These models were intended by Lorenz to be used for testing ideas related to data assimilation, model error, error propagation, etc.**

Lorenz model 1 : shows wave propagation, but Z_n varies rapidly with n.
Lorenz model 2 : shows wave propagation with Z_n varying smoothly with n.
Lorenz model 3 : shows small spatial scale activity on top of smooth waves.

Model waves are analogous to Rossby waves:
v_{group} is eastward, v_{phase} is westward

Model 3

$$\frac{dZ_n}{dt} = \underbrace{[X,X]_{K,n} + b^2[Y,Y]_{1,n} + c[Y,X]_{1,n}}_{\text{Nonlinear interactions}} \underbrace{- X_n - bY_n}_{\text{Damping}} \underbrace{+ F}_{\text{Forcing}}$$

$$X_n = \sum_{i=-I}^{I} (\alpha(I) - \beta(I)\,|\,i\,|) Z_{n+i}$$ ('smooth' part of Z_n)

$$Y_n = Z_n - X_n$$ (small spatial scale fluctuating component of Z_n)

$$[X,Y]_{K,n} = \left(\text{sum of quadratic terms } X_l\, Y_m \text{ over a spatial range of l and m points} \right)$$

Model 2

$$\text{Model 3} \xrightarrow{I=1} \text{Model 2}$$

For $I = 1, X_n = Z_n, Y_n = 0$ (Small scale activity is absent.)

$$dZ_n / dt = [Z,Z]_{K,n} - Z_n + F$$

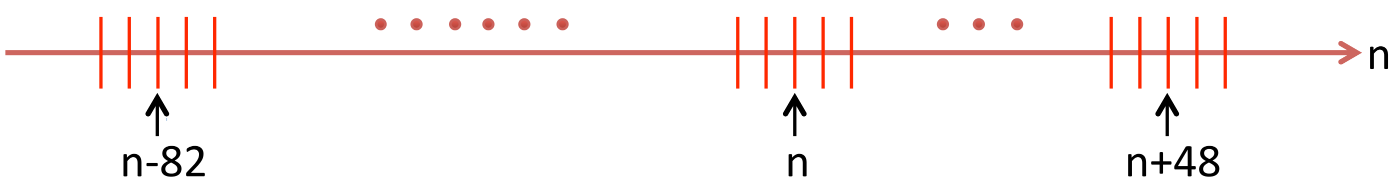
Model 1

$$\text{Model 2} \xrightarrow{K=1} \text{Model 1}$$

$$[Z,Z]_{1,n} = -Z_{n-2}Z_{n+1} + Z_{n-1}Z_{n+1}$$

$$dZ_n / dt = -Z_{n-2}Z_{n+1} + Z_{n-1}Z_{n+1} - Z_n + F$$

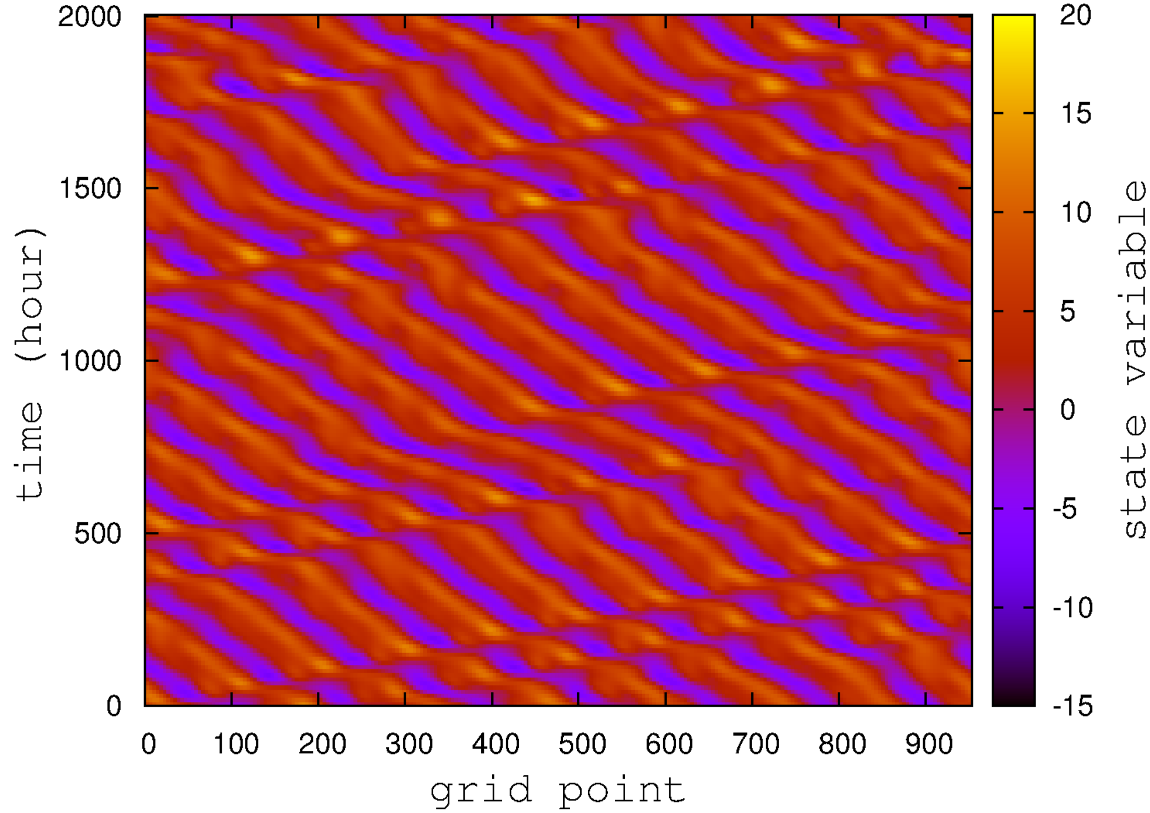
Note that this equation is directional. 'Information' flows more to the right than to the left. This is also true for models 2 and 3. E.g., for model 3 our numerics use K=32; for which $[X,Y]_{K,n}$ involves products X_mY_l for grid points (l,m) in the range [n-82, n+48].



- **We will think of Lorenz model 3 as simulating a 'true' spatiotemporal dynamics that includes small scale processes and whose dynamics we want to forecast.**
- **We will think of Lorenz model 2 with fewer grid points as a coarse grid model of the dynamics of model 3, and we will use model 2 as a global forecast model for the 'true' spatiotemporal dynamics of Lorenz model 3.**

LEFTWARD MOVING WAVE PHASE VELOCITIES

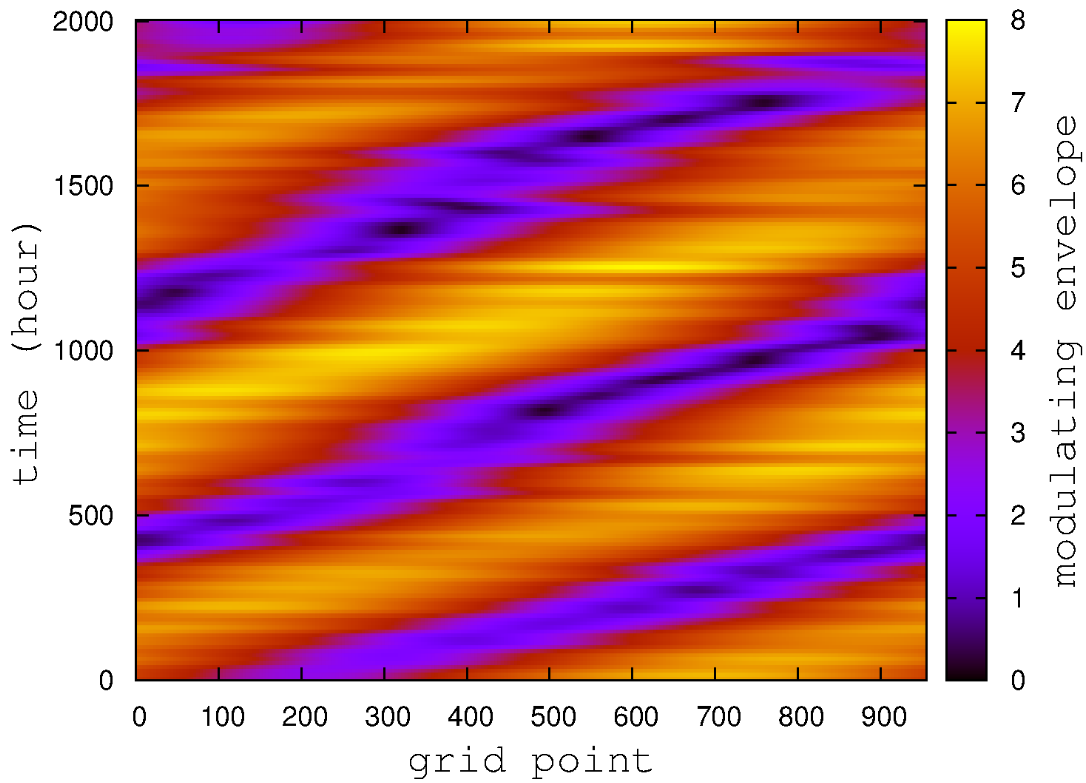
Lorenz model 3, N = 960, K = 32, b = 10, c = 2.5, F = 15, I =12.
$Z_n(t)$ versus n (horizontal) and t (vertical).
Red : high values, Blue : low values



Y. Yoon, E. Ott, and I. Szunyogh, "On the Propagation of Information and the Use of Localization in Ensemble Kalman Filtering", J. Atmos. Sci. 67, 3823 (2010).

ENVELOPE PROPAGATION

We view the spatiotemporal behavior seen in the previous slide as being due to a spectrum relatively short wavelength waves that are amplitude modulated by a longer scale envelope. We apply an envelope extraction technique* to the data for $Z_n(t)$:

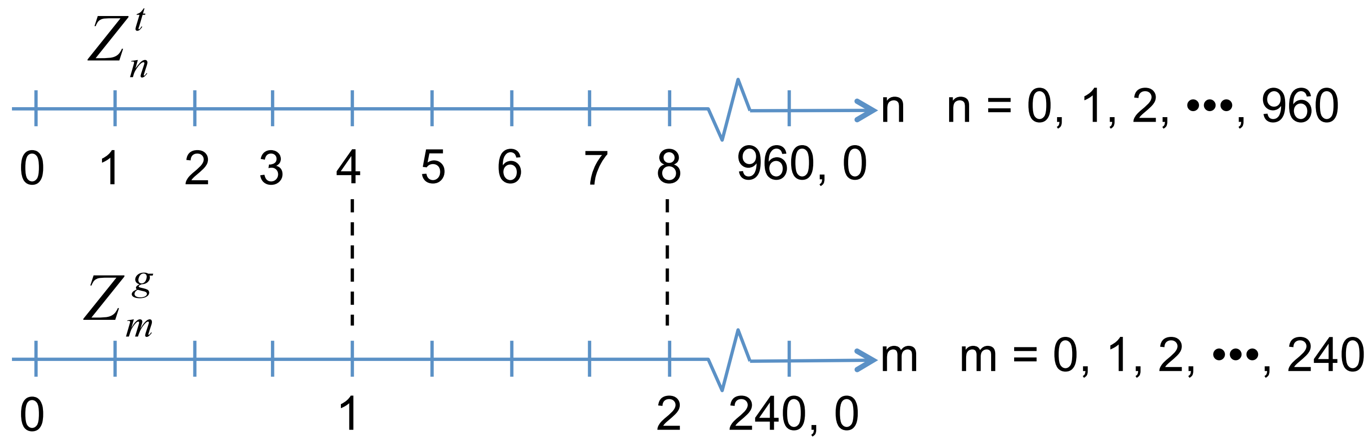


Envelope propagation and hence group velocity, is to the right.

***A. Zimin, I. Szunyogh, D. J. Patil, B. R. Hunt, and E. Ott, "Extracting Envelopes of Rossby Wave Packets", Mon. Wea. Rev. 131, 1011 (2003).**

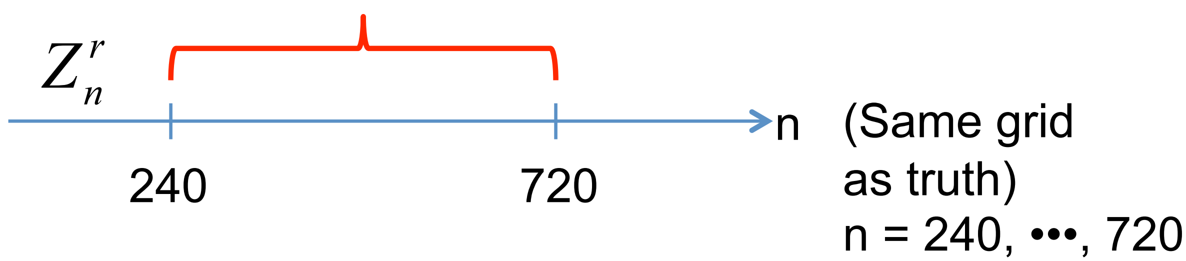
TRUTH
Lorenz model 3

GLOBAL MODEL
Lorenz model 2
(1/4 Resolution)



m = n/4: A global model grid point is at every 4th truth grid point.

REGIONAL MODEL
Lorenz model 3
(same as truth)



PARAMETERS

The parameter F are the same in all three models. The parameters b and c are the same in the true and regional models.
K (truth) = K (regional) = 4 K (Global) = 32
I (truth) = I (regional) = 12

COUPLING THE REGIONAL MODEL TO THE GLOBAL MODEL

Problem:

dZ^r_n/dt depends on bracket terms $[\bullet, \bullet]_{K,n}$ and X^r_n, which are nonlocal (i.e., they depend on state variables in a finite neighborhood around n). Thus for n too close to the subregion boundaries, we require X, Y and Z values at points outside the subregion in order to evaluate dZ^r_n/dt and evolve $Z^r_n(t)$ in time.

What we do:

We take the global state $Z^g_m(t)$, interpolate it from the coarse m-grid onto the finer n-grid, and insert these interpolated values for values required by the regional model in the evaluation of $[\bullet, \bullet]_{K,n}$ and X^r_n. (This essentially plays the role of boundary conditions on the regional model.)

SIMULATED OBSERVATIONS

15 simulated observations are made every '6 hours'.
They are evenly spaced starting at n = 0

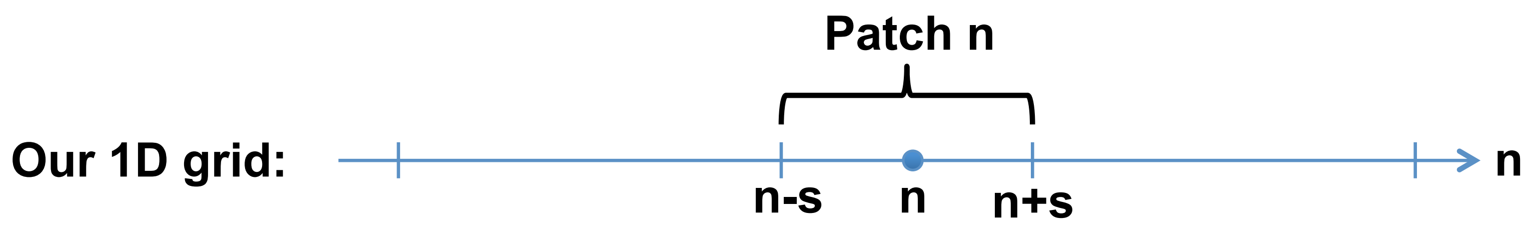
n = 0, 64, 128, •••, 896

$$O_n = Z_n^t + \left(\text{Gaussian random number With variance one} \right)$$

NOW WE WANT TO DO DATA ASSIMILATION AND FORECASTING

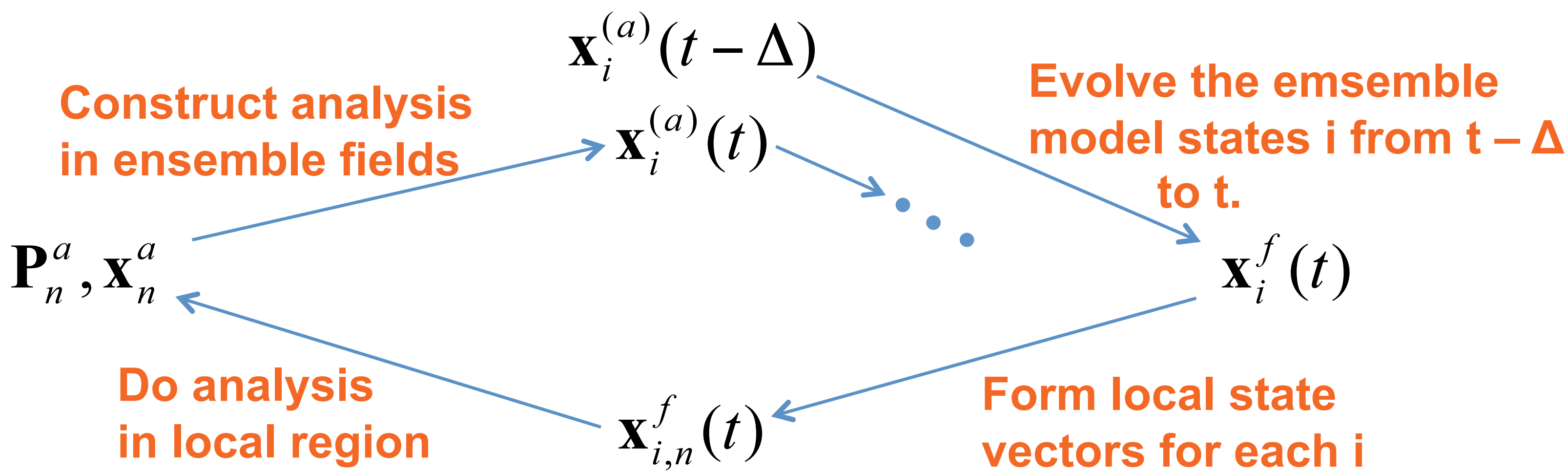
But first a brief review of the local Ensemble Kalman Filter as applied to a global model alone (i.e., without consideration of a coupled regional model).

REVIEW OF LOCAL THE ENSEMBLE KALMAN FILTER



We will use s = 40 corresponding to 2s+1 = 81 grid points.

i = Ensemble member, $x_i(t)$ = State of ensemble member i



$$J_n(\mathbf{x}_n) = (\mathbf{x}_n - \bar{\mathbf{x}}_n^f)^{\mathrm{T}} (\mathbf{P}_n^f)^{-1} (\mathbf{x}_n - \bar{\mathbf{x}}_n^f) + [o_n - \mathbf{H}_n(\mathbf{x}_n)]^{\mathrm{T}} \mathbf{R}^{-1} [o_n - \mathbf{H}_n(\mathbf{x}_n)]$$

Refs: E. Ott et al., Tellus A (2004); B. Hunt, E. Kostelich, I. Szunyogh, Tellus A (2007).

COUPLED GLOBAL/REGIONAL ANALYSIS

$x_{n,i}^g$ = patch n global state (the values of $Z_{m,i}^g$ in patch n) for ensemble member i
$x_{n,i}^r$ = patch n regional state (values of $Z_{n,i}^r$ in patch n; ≤ 2s+1 = 81 components)

Joint state for ensemble member i: $\mathbf{x}_{n,i} = \begin{bmatrix} \mathbf{x}_{n,i}^g \\ \mathbf{x}_{n,i}^r \end{bmatrix}$

Do local analysis in each patch n using $x_{n,i}$ with

$$J_n(\mathbf{x}_n) = (x_n - \bar{x}_n^f)^{\mathrm{T}}(\mathrm{P}_n^f)^{-1}(x_n - \bar{x}_n^f)$$
$$+ [\mathbf{y}_n - \mathbf{H}_n(\mathbf{x}_n)]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y}_n - \mathbf{H}_n(\mathbf{x}_n)]$$
$$+ \kappa[\mathbf{G}_n^g(\mathbf{x}_n^g) - \mathbf{G}_n^r(\mathbf{x}_n^r)]^{\mathrm{T}}[\mathbf{G}_n^g(\mathbf{x}_n^g) - \mathbf{G}_n^r(\mathbf{x}_n^r)]$$

$\mathbf{G}_n^g(\mathbf{x}_n^g) =$ **Vector of global state values at patch n points that are on both the global and regional grid**

$\mathbf{G}_n^r(\mathbf{x}_n^r) =$ **Vector of regional state values at the same points**

$$[\mathbf{H}_n(\mathbf{x}_n)]_i = \begin{cases} (1-\lambda)[\mathbf{x}^g]_{j(i)} + \lambda[\mathbf{x}^r]_{j(i)} & \text{(for observations at points } j(i) \text{ in regional model domain)} \\ [\mathbf{x}^g]_{j(i)} & \text{(for observations outside the regional model domain)} \end{cases}$$

$j(i) =$ location of the i[th] observation in patch n

METHOD FOR COMPARISON

- Do global assimilation separately (via LETKF) and independently from the regional state (as in common current practice).
- Do LETKF assimilations on the regional model.
- For both methods we do some smoothing of the initial regional analysis state so that there is not an abrupt discontinuity between global and regional values at the regional boundaries.

TESTS

We used 40 ensemble members.

Parameters to be set

(1) Variance inflation factor ø: $\mathbf{P}_n^f \rightarrow (1+\phi)\mathbf{P}_n^f$

(2) Observation operator weighting factor λ

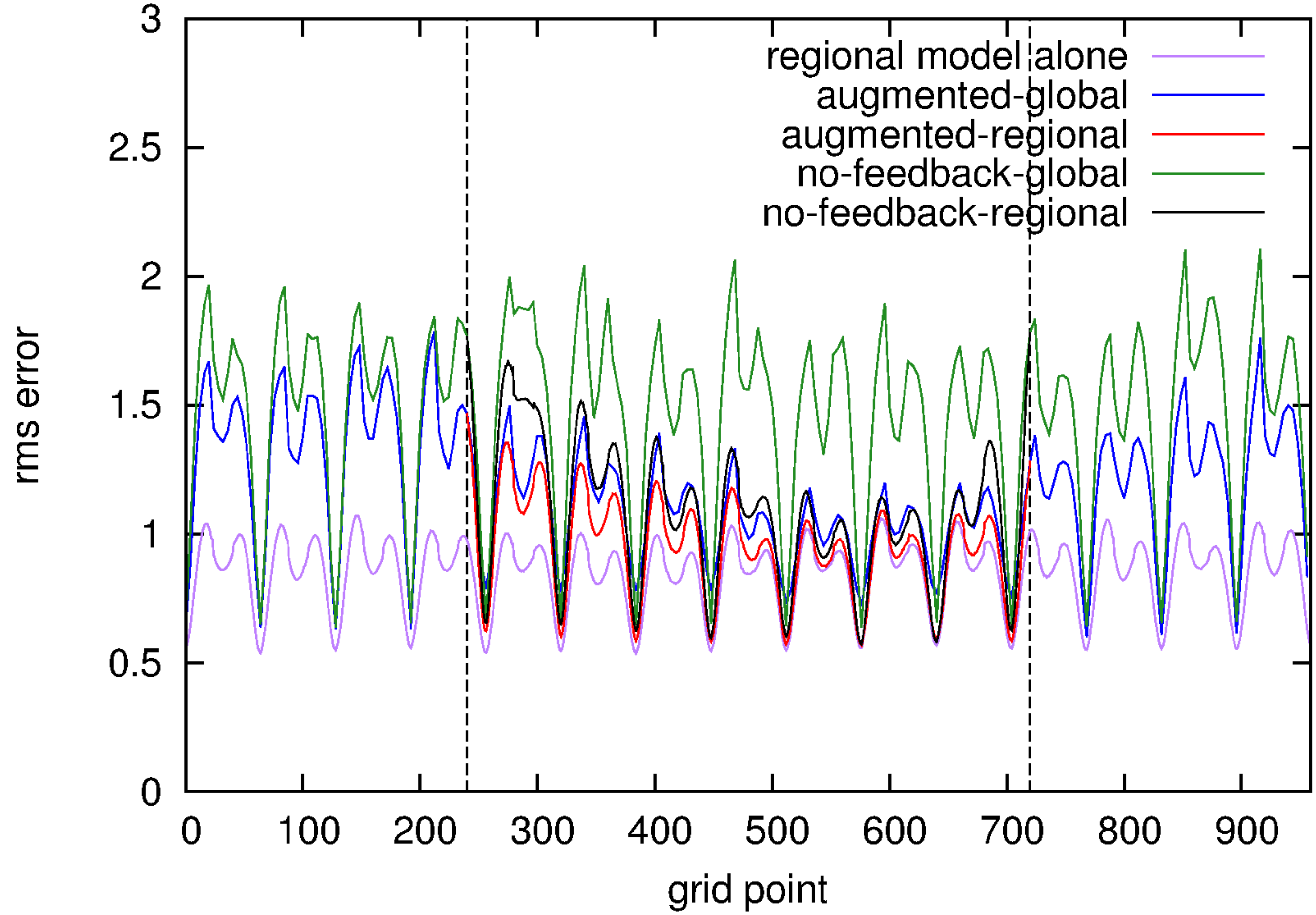
(3) Factor κ penalizing the difference between global and regional states.

We chose these parameters to minimize the analysis error:

ø (global) = 0.024, λ = 0.90

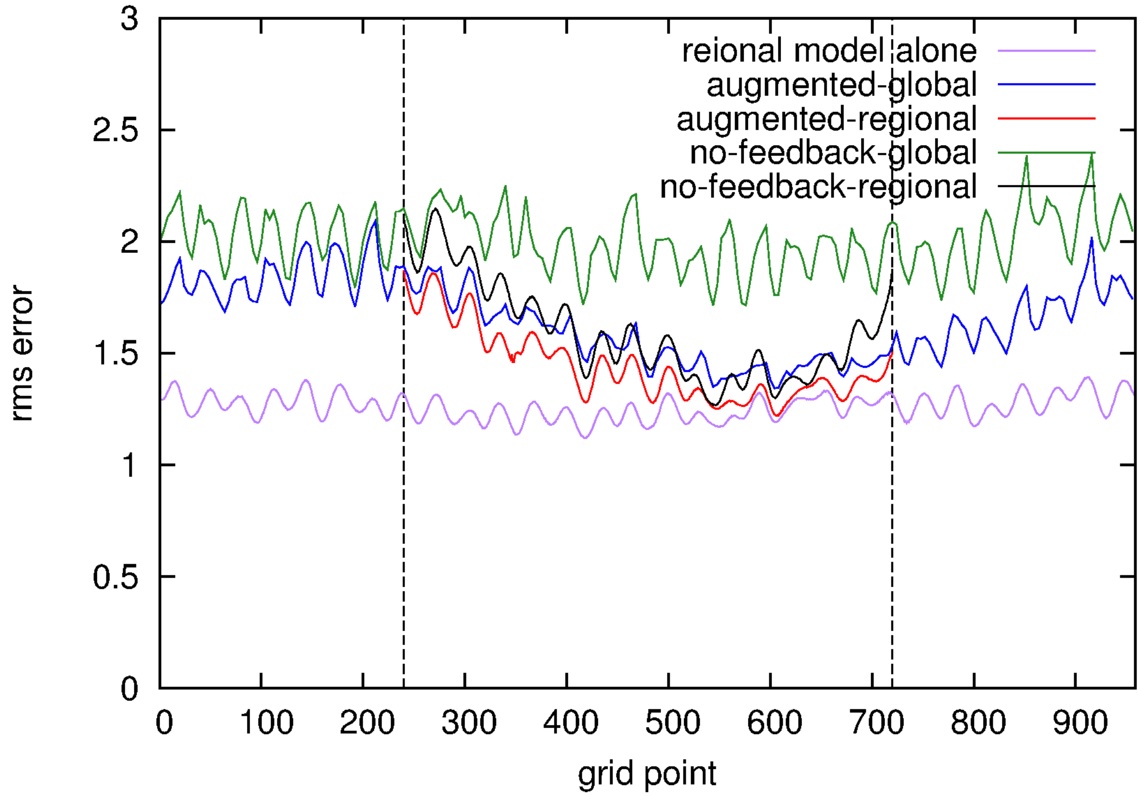
ø (regional) = 0.020, κ = 0.04

RESULTS – ANALYSIS ERROR



Dashed vertical lines show subregion boundaries (at n = 240, 720)
Purple curve = error when model 3 was used for the whole region [0, 960] for getting the analysis (perfect model scenario).

RESULTS – ERROR IN '1 DAY' FORECAST



- **For both the analysis and the forecast joint state method does better, even outside the subregion.**
- **This can be interpreted as due to**

  (Better global analysis inside subregion) → (Better forecasts outside the subregion) →

  (Better boundary conditions for the regional model) → (Better regional forecasts)
- **Global model improvements are better to the right of the subregion than to its left (v_g is to the right).**

ISSUES

- Effect of model error in the regional model.
- Tests and implementation on more realistic models and using real events.
- Will improvement persist in these future tests?
- Can the method improve the forecasting of, e.g., cyclone tracks and intensity? Etc.

SOME BENEFITS OF HIGHER RESOLUTION

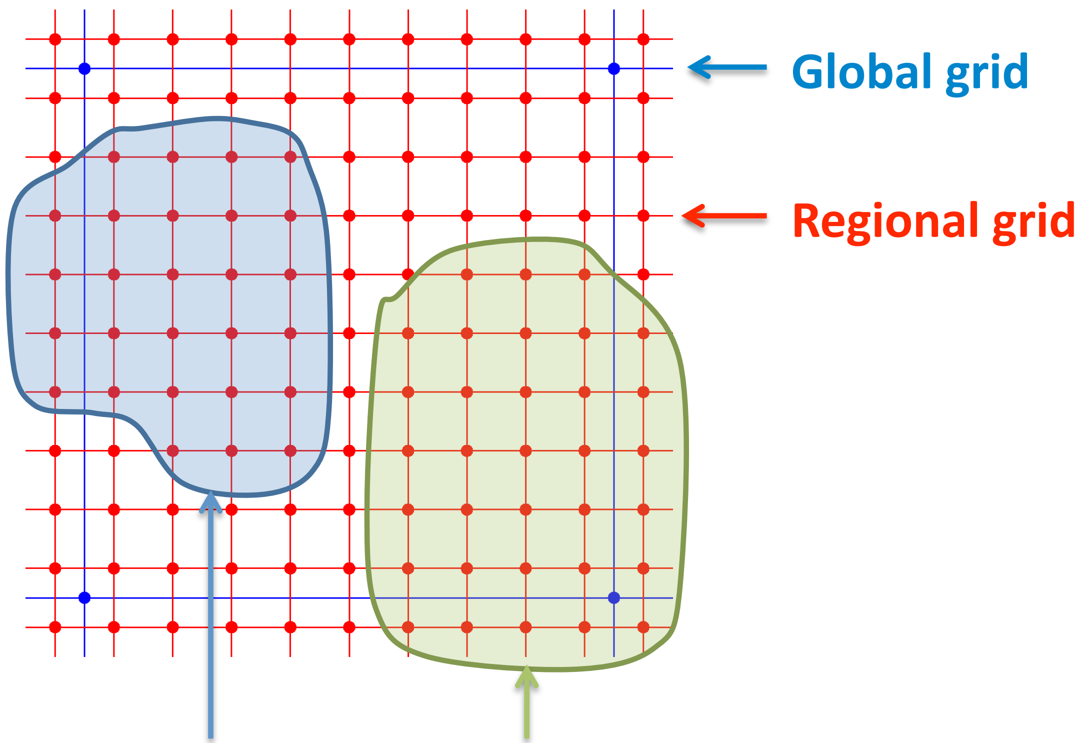
Resolving orography (e.g., changes in elevation, mountains, etc.), sharp atmospheric changes (e.g., at fronts), cyclones.

Tropical cyclones : Eye wall radius ~ 25-50 km
4-10 grid points needed to resolve eye
→ (grid spacing) ~ 5-10 km
which is significantly less than
what global models can currently do.

Ref. On tropical cyclone data assimilation with coupled regional and global models : Christina Holt and Istvan Szunyogh, MS Thesis, Tex A&M (to be published).

GOALS

Obtain more accurate, finer spatial forecasts. E.g., specify different weather predictions in contiguous regions that are unresolved by the global model.



Can we make reliably different forecasts for these two regions based on the regional model?