Numerical Modelling, Predictability and Data Assimilation in Weather, Ocean and Climate

A Symposium honouring the legacy of Anna Trevisan

Bologna, Italy 17-20 October, 2017

Multiple-scale error growth and data assimilation in convection-resolving models

Francesco Uboldi

ARIANET Srl, Milan, Italy

http://www.aria-net.it

e-mail: f.uboldi@aria-net.it

Outline of the presentation

- Introduction
- Scientific turning points in my collaboration with Anna Trevisan on the development of the Assimilation in the Unstable Subspace (AUS).
  - (~1999-2004): 40-points Lorenz model, adaptive observations
  - (~2002-2005): primitive-equation isopycnal ocean model MICOM
- Multiple-scale error growth and data assimilation in convection-resolving models
  - (~2010-2013): non-hydrostatic, convection-resolving model MOLOCH

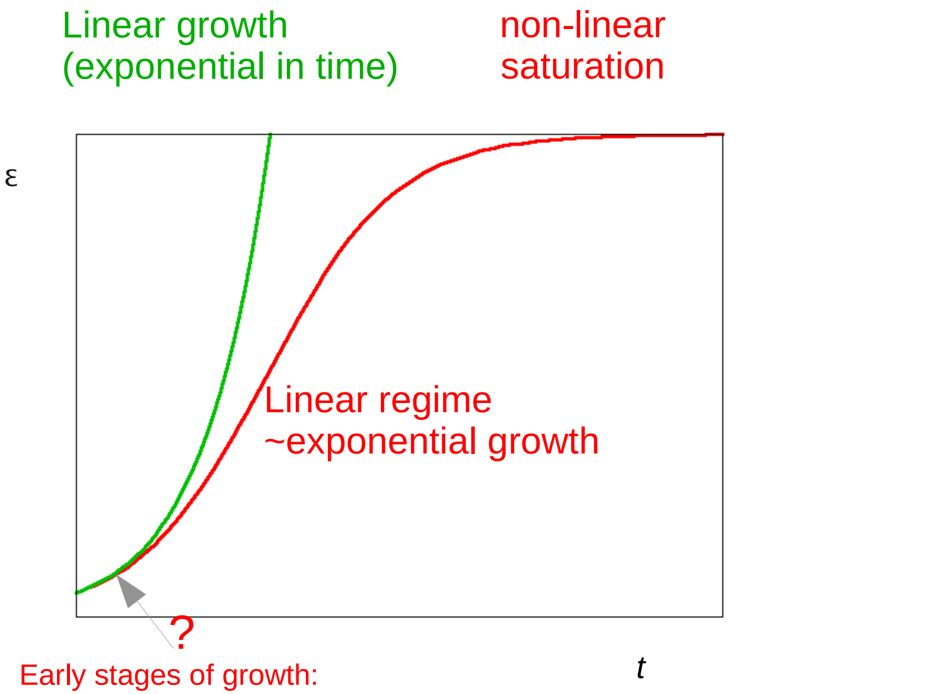
Error growth and unstable directions

Non-linear growth of small perturbations

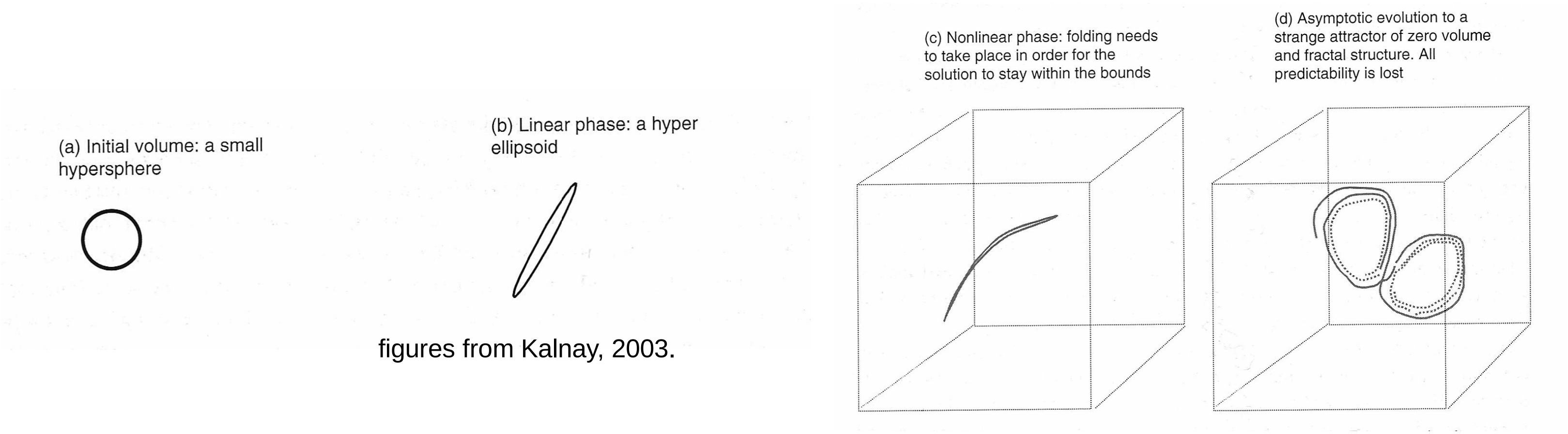
1) Linear regime: exponential growth

2) End of linear regime

3) Nonlinear Saturation



Early stages of growth:
- Decrease (along stable directions);
- Super-exponential growth (non orthogonality of unstable directions)



figures from Kalnay, 2003.

Unstable directions = **Lyapunov vectors (LV) with positive exponents**

- **LVs characterise perturbation growth in the linear regime**
- LVs (evolve with the tangent linear dynamics and) are co-variant with the phase flow
- LVs are sorted by decreasing growth exponent: **the first LV is the most unstable**.
- LVs are not orthogonal.
- REMARK that, even if an orthogonalization is often used to keep linear independence:
    - Lyapunov exponents
    - The first LV (the most unstable)
    - The sequence of subspaces spanned by LVs

    DO NOT depend on the choice of the scalar product!

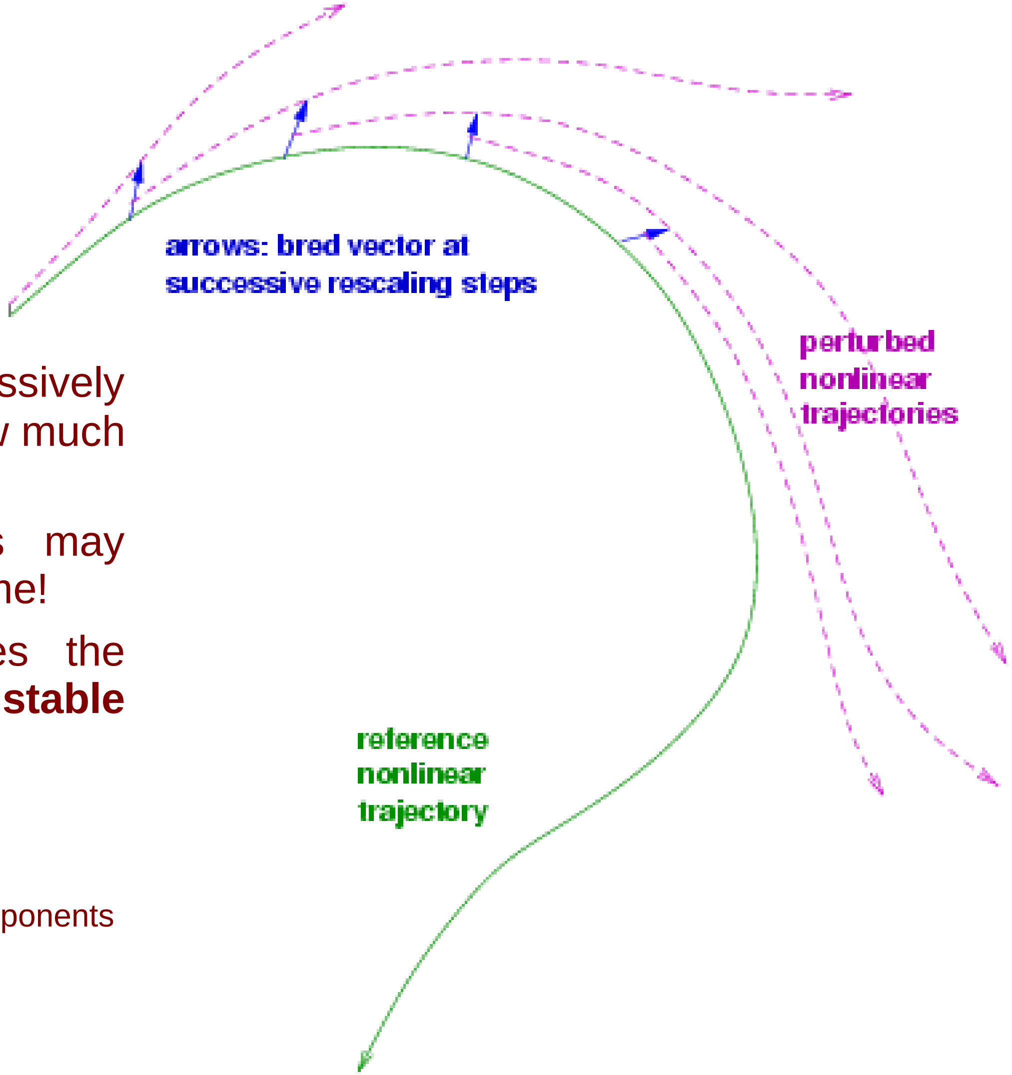
    ⇒ **The subspace sequence is a local property of the attractor and characterizes its local geometry**
- Oseledec, 1968; Benettin et al., 1980; Legras and Vautard, 1997; Trevisan and Pancotti, 1998; Kalnay, 2003 book; Wolfe and Samelson 2007, Ginelli et al. 2007... Lucarini 2017

Characterize linear regime growth

Approximate Lyapunov vectors: **Breeding** (Toth and Kalnay, 1993)

- **small** initial perturbation;
- **nonlinear** integration of control and of perturbed state;
- **frequent rescaling** of the growing perturbation to **impose linear growth** (along the non-linear trajectory)

time 0: perturbed state ← control state + perturbation
time *t*: perturbation ← perturbed state - control state



- Initially independent perturbations progressively collapse onto one direction, the 1st LV (in how much time?) $\propto 1/(\lambda_2 - \lambda_1)$
- **MANY** initially independent perturbations may collapse onto **FEW** directions in a **SHORT** time!
- Each **bred vector** progressively acquires the structure of **a linear combination of the unstable directions**

- Initial coefficients of the linear combination unknown
- How much time: depends on differences between growth exponents
- (orthogonalization: only to keep bred vectors independent)

Lorenz 40-points model

Non-linear. Chaotic.

1 spatial dimension, periodic domain
“latitude circle”

$$\frac{dx_i}{dt}=\left(x_{i+1}-x_{i-2}\right)x_{i-1}-x_i+F$$

Model error: F'=7.6 ≠ F=8.0

Fixed observations on “land”: i= 21-40
1 adaptive observation on “ocean” i= 1-20
Observation error: random with σ = 0.2

JOURNAL OF THE ATMOSPHERIC SCIENCES

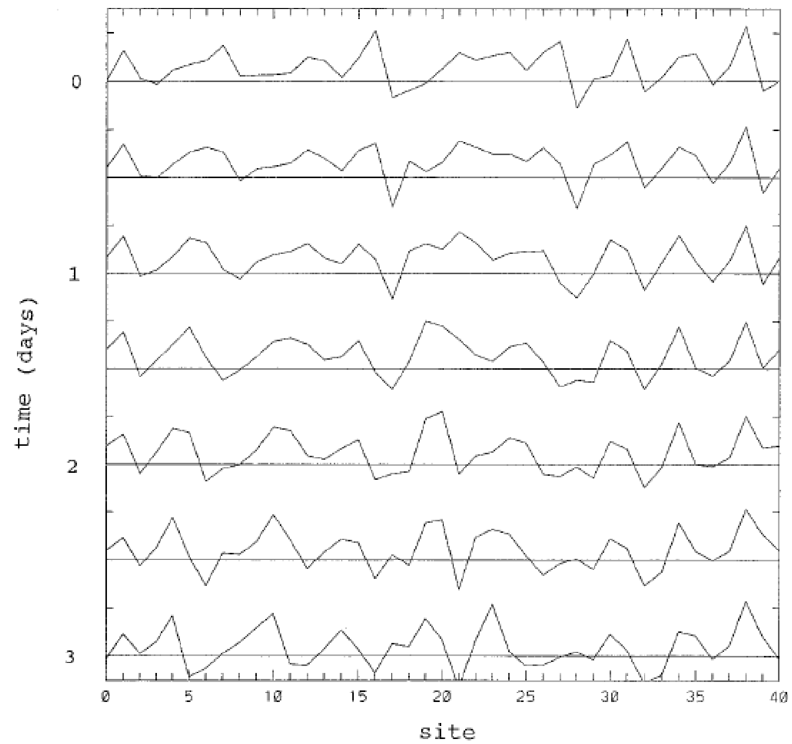


FIG. 2. Longitudinal profiles of X_j as in Fig. 1 but at 12-h intervals, and with the profile that would follow the initial profile in Fig. 1 by 3 years used as the new initial profile. Horizontal lines are zero lines. Interval between zero lines and short marks at left and right is 10.0 units.

Lorenz, E. and K. Emanuel, 1998. *J. Atmos. Sci*, **55**, 399-414

Scalar “targeted” observation located where the bred vector is maximum

Forecast state vector

forecast estimate of observation

Analysis state vector

current “mature” bred vector

bred vector at obs location = = Max component of bred vector

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{e}\,\frac{y^o - y^f}{e_o}$$

i: generic state component index

i_o : location of “targeted” obs = index of the maximum component of the bred vector

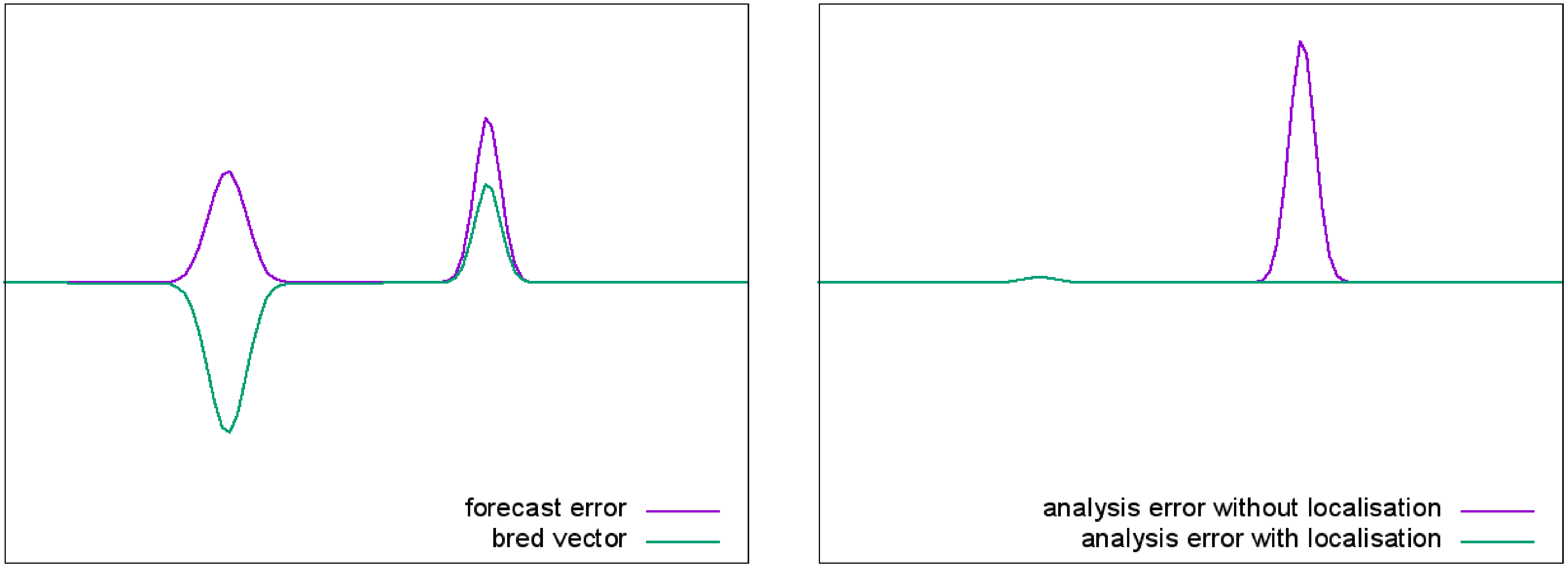
$$x^a(i) = x^f(i) + e(i)\,\frac{y^o - x(i_o)}{e(i_o)}$$

- The forecast error, evolving non-linearly, but growing in the linear regime, takes a spatial shape determined by the shape of the unstable structures.
- So does the bred vector, forced to grow linearly by frequent rescaling.
- ⇒Near the maximum of the bred vector, the forecast error has nearly the same shape

Trick: regionalisation (localisation) by a ***wide*** Gaussian modulating function:

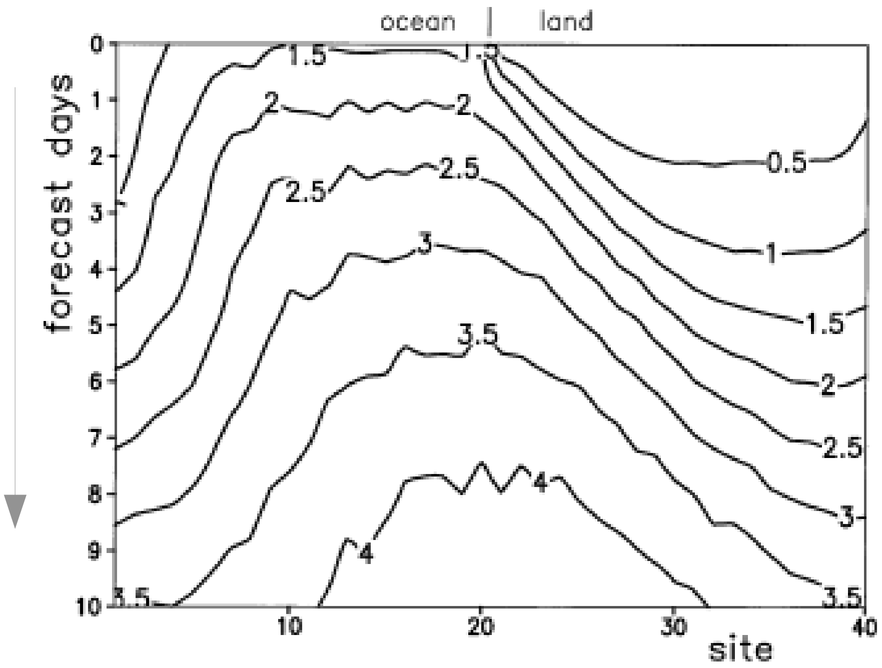
$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{g} * \mathbf{e}\,\frac{y^o - y^f}{e_o}$$

$$x^a(i) = x^f(i) + g(i)\,e(i)\,\frac{y^o - x(i_o)}{e(i_o)}$$

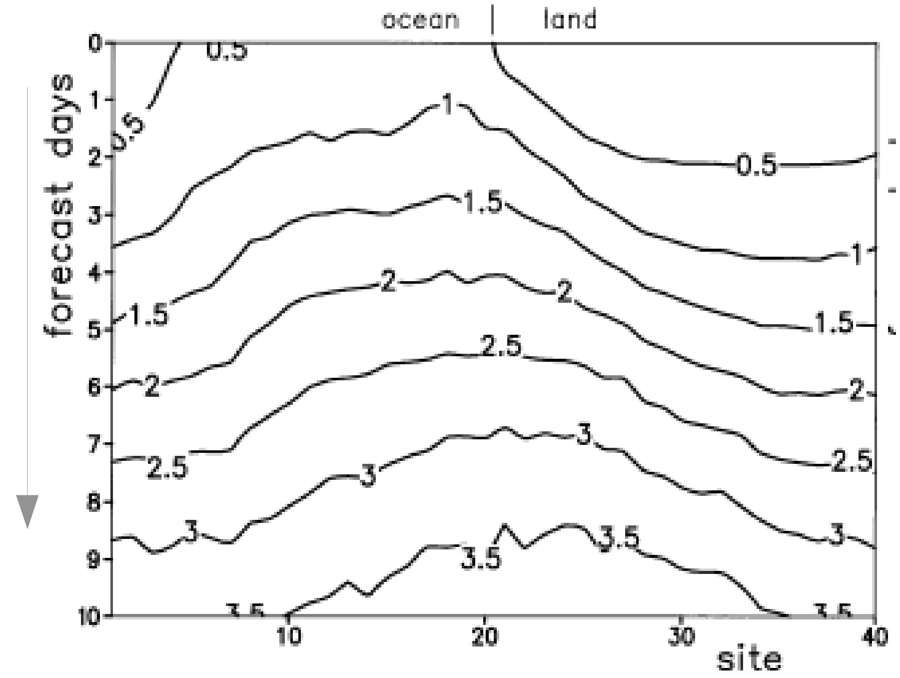


Two or more independent unstable structures can be active in both the forecast error field and in the bred vectors.
The unstable structures are the same in the forecast error and in the bred vectors.
But they may be combined differently, in particular with opposite signs.

Localisation is needed, unless all unstable Lyapunov vectors are estimated and used in the assimilation.



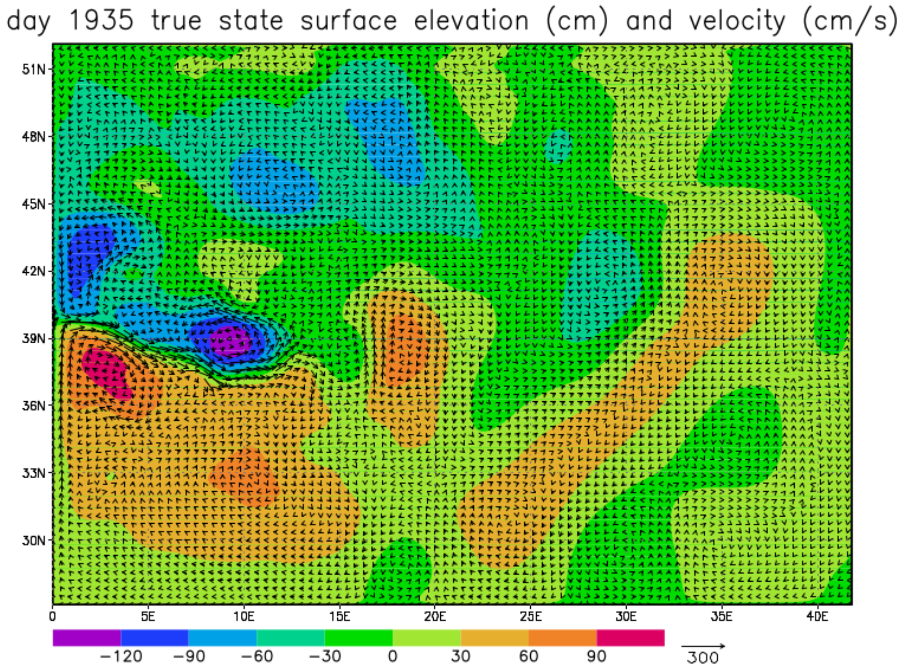
Previous result

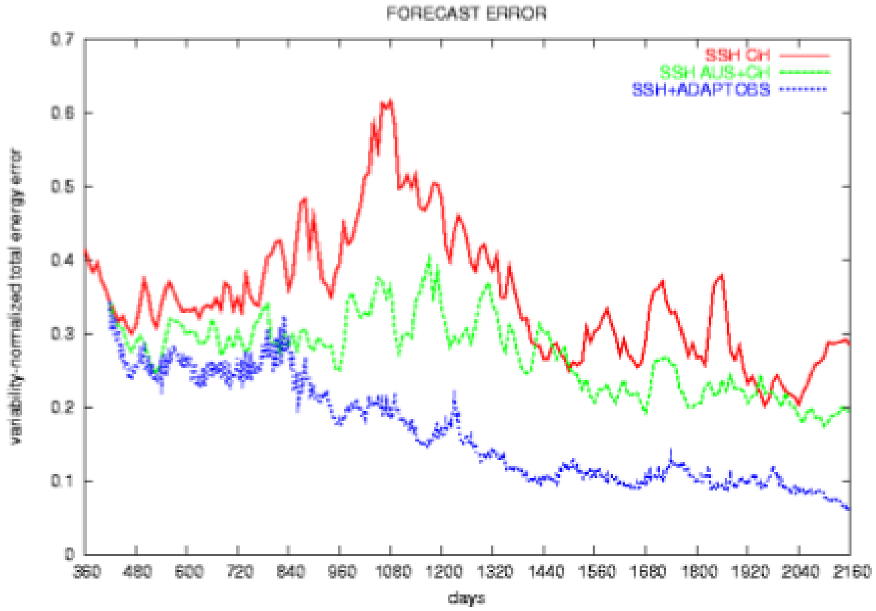


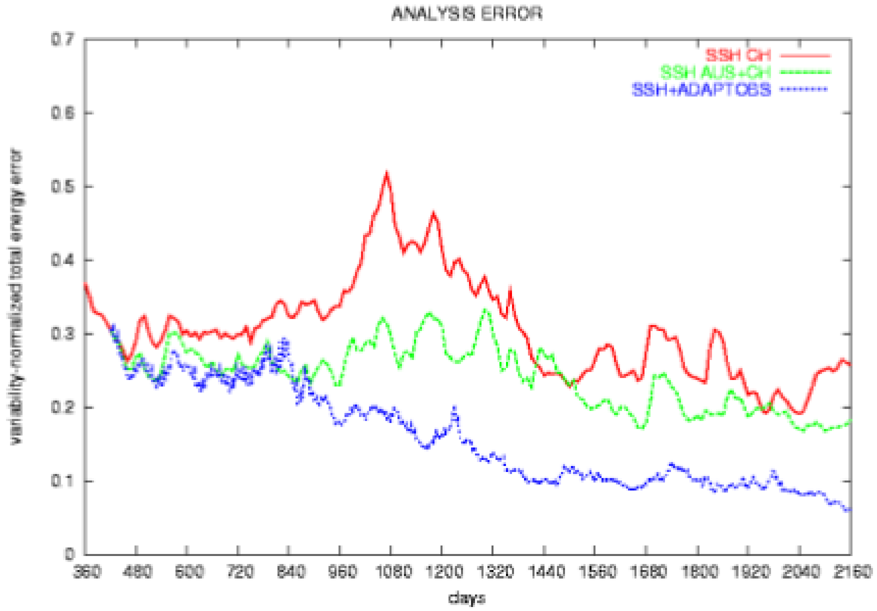
“AUS first sketch”

Trevisan, A. and F. Uboldi, 2004. *J. Atmos. Sci*, **61**, 103-113

Primitive equation ocean model
Isopycnal MICOM (Bleck, 1978)
4-layer double-gyre box

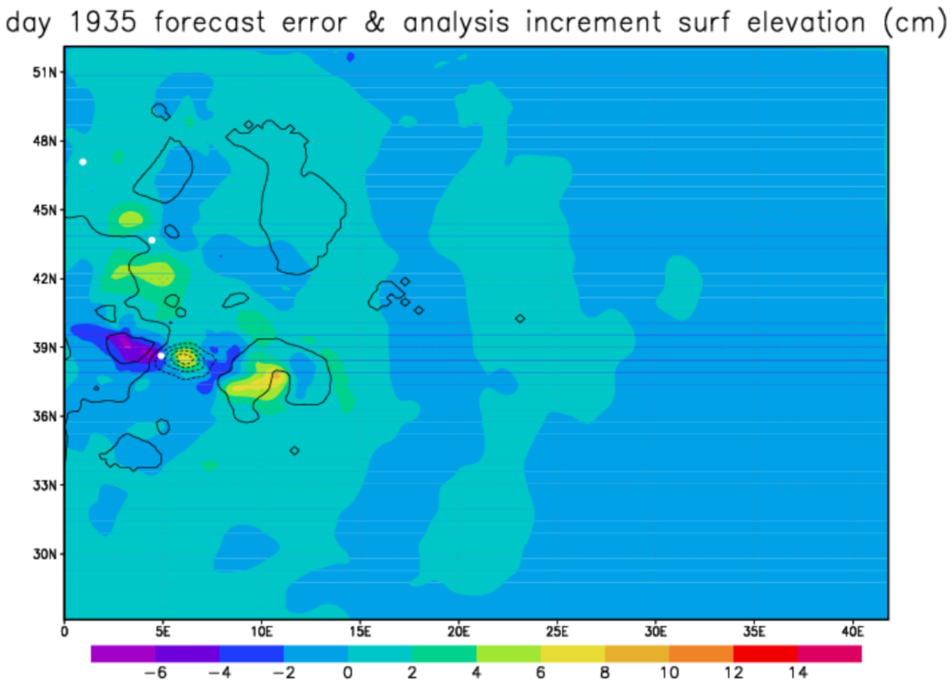


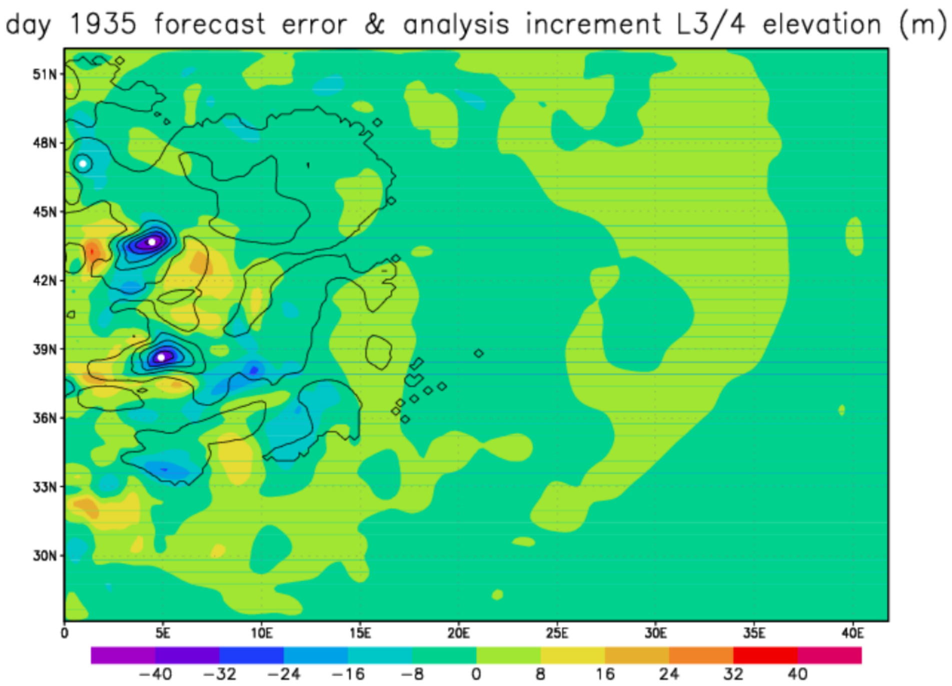




Complex vertical structure of the unstable structures present in the forecast error

- The forecast error field (shaded) differs at the **surface** and in **deep layers**
- White dots: horizontal position of 3 targeted observations
- Each observation is **a single scalar**: layer interface elevation.
- In this case, **all 3 obs are deep**: L3/L4
- **The same 3-dimensional structure is present in the forecast error and in the forced bred vector**
- **The analysis increment (contour) corrects the forecast error at the surface too**



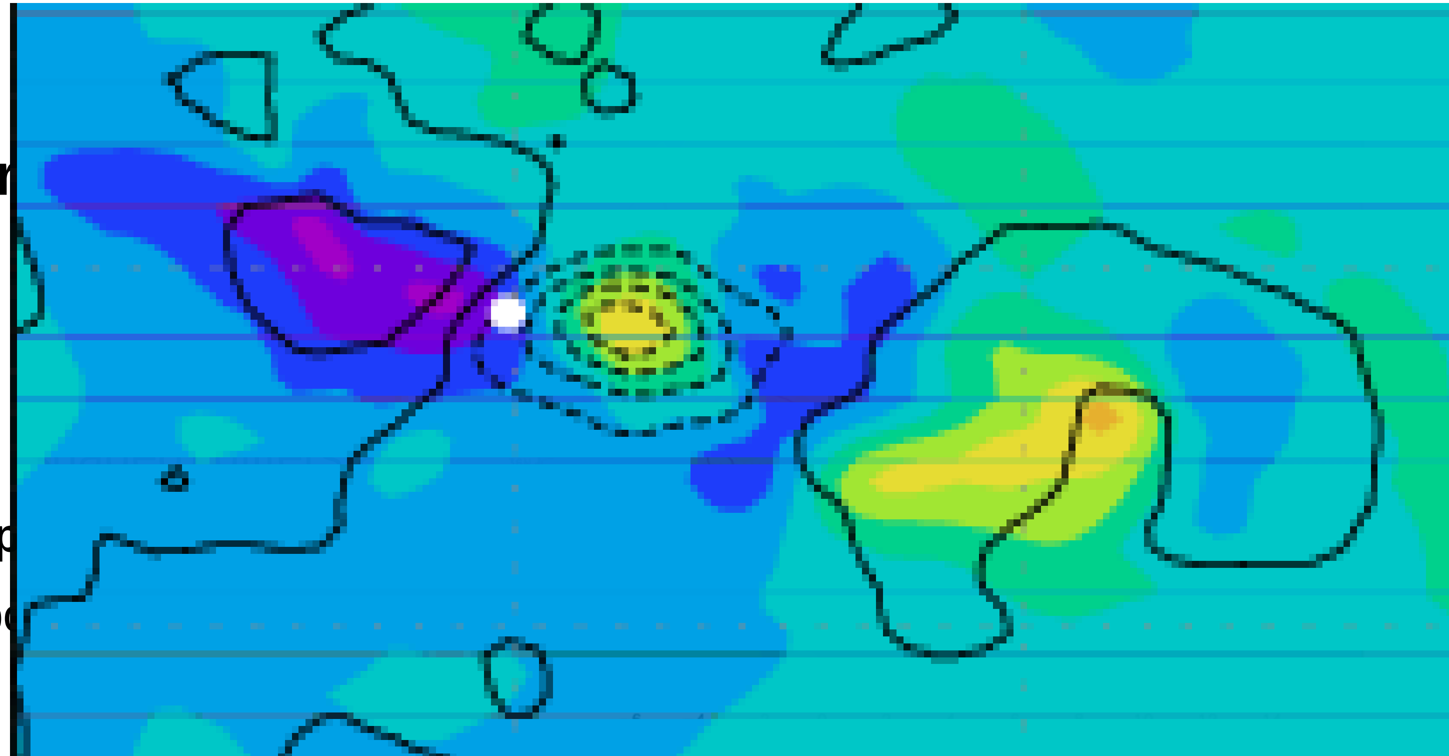


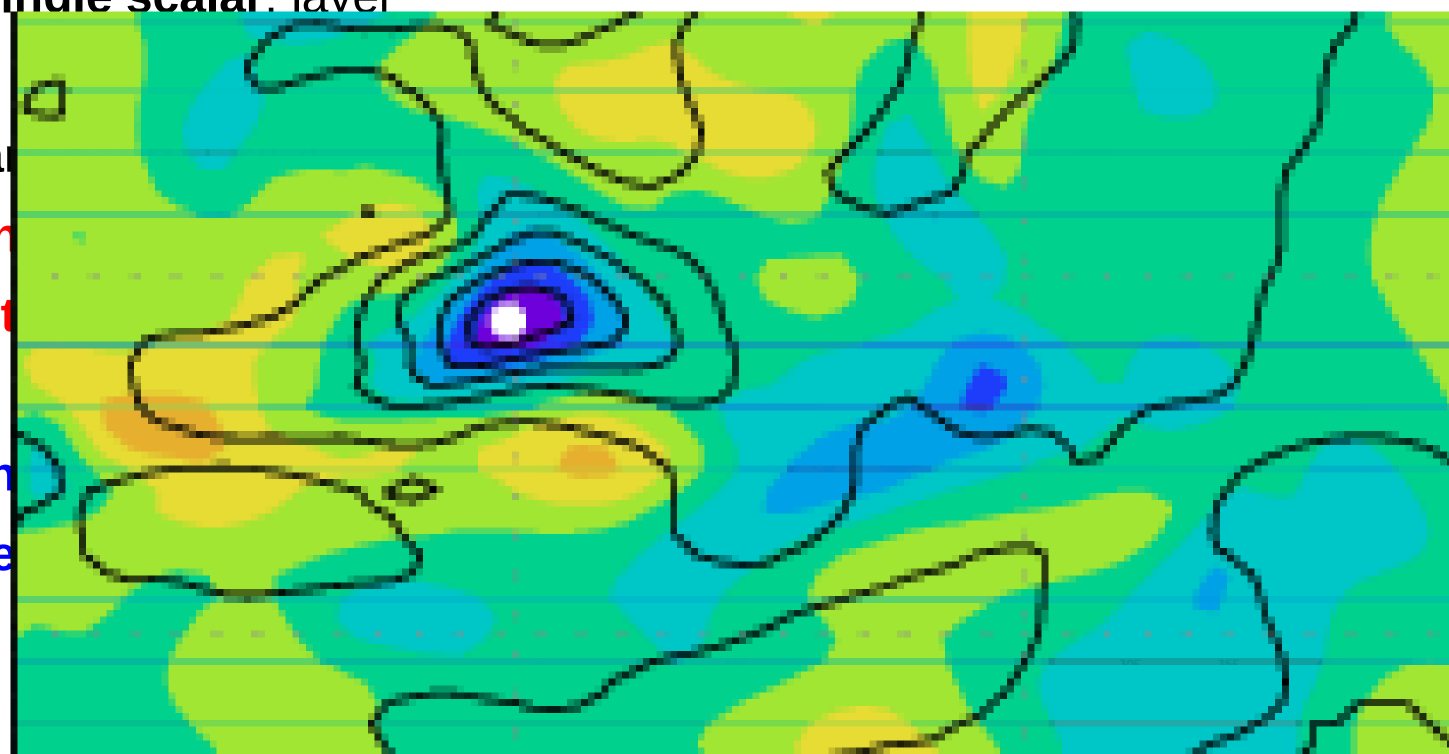
Complex vertical str

unstable structures

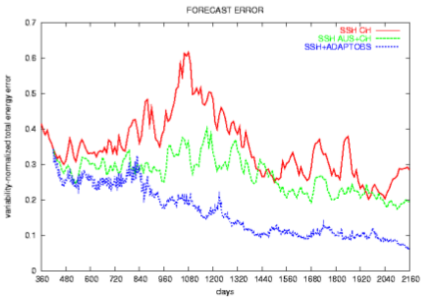
the forecast error

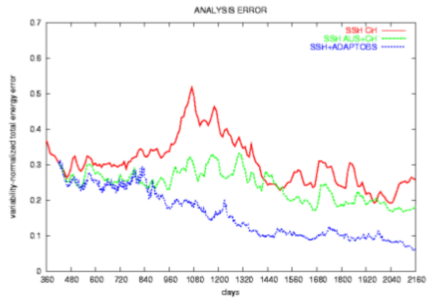
- The forecast error field the **surface** and in **deep**
- White dots: horizontal p targeted observations
- Each observation is **a single scalar**: layer interface elevation.
- In this case, **all 3 obs a**
- **The same 3-dimension present in the forecast forced bred vector**
- **The analysis incremen corrects the forecast e surface too**

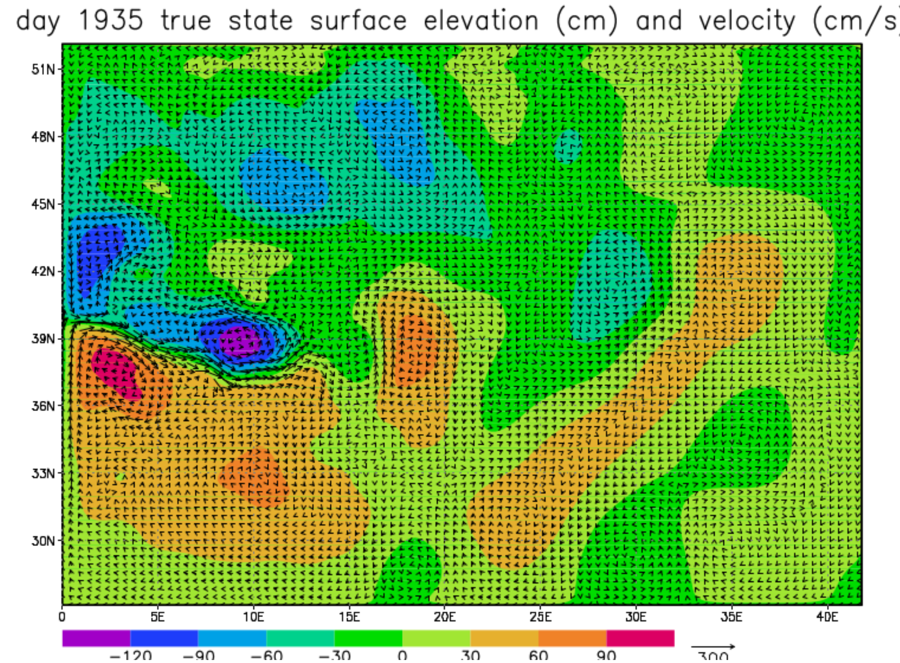




Primitive equation ocean model
Isopycnal MICOM
4-layer double-gyre box







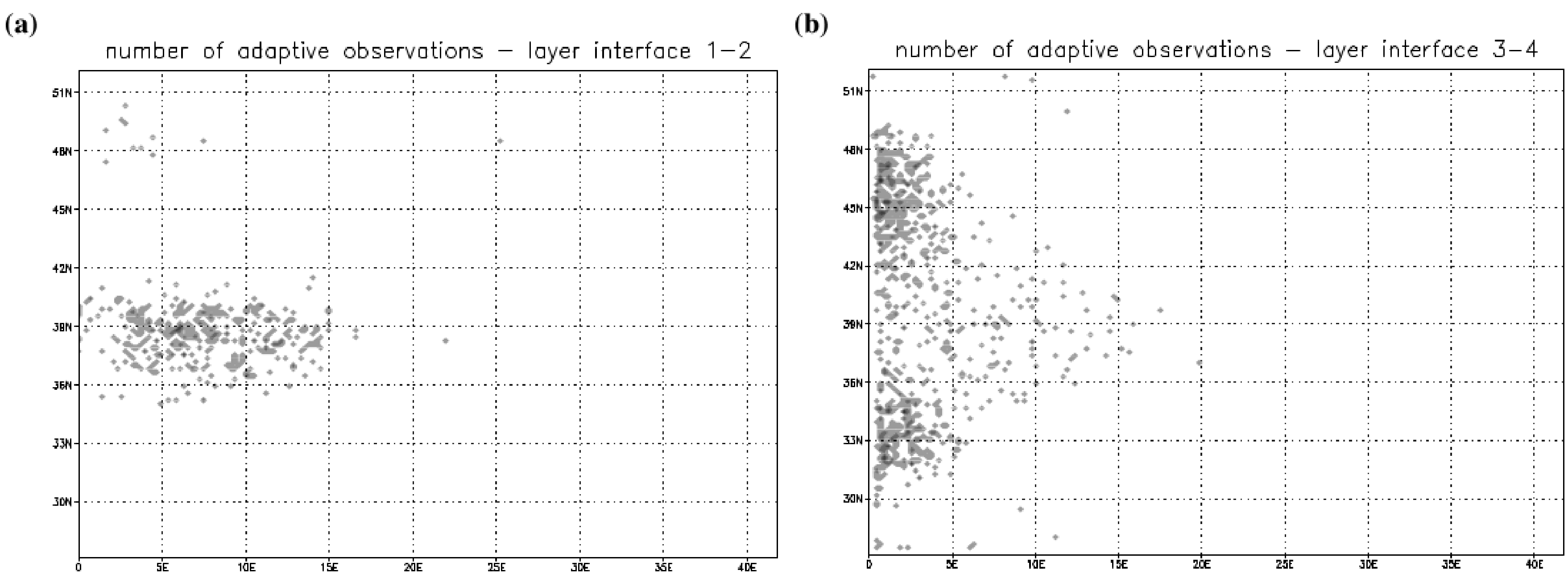


Fig. 2. Number of adaptive observations per gridpoint: **(a)** interface between layers 1 and 2; **(b)** interface between layers 3 and 4. Contouring values are set to: 0.1, 1.1, 2.1, 3.1.

The data assimilation system is a dynamical system FORCED by the assimilation of observations

$$\mathbf{x}^f_{k+1} = M \circ \mathbf{x}^a_k$$

Forecast (nonlinear model integration)

$$\mathbf{x}^a_{k+1} = (\mathbf{I} - \mathbf{K}\, H \circ)\, \mathbf{x}^f_{k+1} + \mathbf{K}\, \mathbf{y}^o_{k+1}$$

Analysis

$$\mathbf{x}^a_{k+1} = (\mathbf{I} - \mathbf{K}\, H \circ)\, M \circ \mathbf{x}^a_k + \mathbf{K}\, \mathbf{y}^o_{k+1}$$

$$\delta\, \mathbf{x}^a_{k+1} = (\mathbf{I} - \mathbf{K}\mathbf{H})\, \mathbf{M}\, \delta\, \mathbf{x}^a_k$$

Perturbations

$$\boldsymbol{\eta}^a_{k+1} = (\mathbf{I} - \mathbf{K}\mathbf{H})\, \mathbf{M}\, \boldsymbol{\eta}^a_k + (\mathbf{I} - \mathbf{K}\mathbf{H})\, \boldsymbol{\eta}^M_{k+1} + \mathbf{K}\, \boldsymbol{\varepsilon}^o_{k+1}$$

Errors (at first order)

growth

decrease (depending on $\mathbf{K}$)

System forced by data assimilation: **BDAS = B**reeding on the **D**ata **A**ssimilation **S**ystem

All perturbed states assimilate the same observations, with the same assimilation scheme as the control state.

Instabilities grow during free evolution and are partly suppressed at each analysis stage.

1) The resulting bred vector are composed of instabilities that survived the analysis steps

2) The overall growth rate should be smaller than that of the free system

Assimilation in the unstable subspace (AUS)

An important component of forecast error belongs to the unstable subspace.

$$\boldsymbol{\eta}^f = \mathbf{E}\boldsymbol{\gamma} + \boldsymbol{\xi} \qquad \mathbf{E} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_N]$$

Confine the analysis increment into the unstable subspace:

$$\mathbf{P}^f \simeq \mathbf{E}\boldsymbol{\Gamma}\mathbf{E}^{\mathrm{T}}$$

Γ: prior covariance of unstable components of forecast error

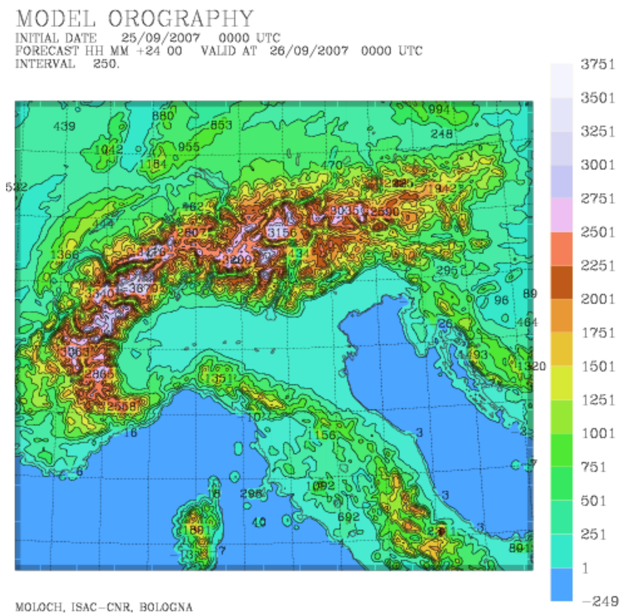
$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{E}\boldsymbol{\Gamma}(\mathbf{H}\mathbf{E})^{\mathrm{T}}\left[(\mathbf{H}\mathbf{E})\boldsymbol{\Gamma}(\mathbf{H}\mathbf{E})^{\mathrm{T}} + \mathbf{R}\right]^{-1}\left[\mathbf{y}^o - H(\mathbf{x}^f)\right]$$

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{E}\left[\boldsymbol{\Gamma}^{-1} + (\mathbf{H}\mathbf{E})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{E})\right]^{-1}(\mathbf{H}\mathbf{E})^{\mathrm{T}}\mathbf{R}^{-1}\left[\mathbf{y}^o - H(\mathbf{x}^f)\right]$$

How many LVs? Dimension of the unstable subspace = number of non-negative exponents

1) Estimate LV ← computational cost
2) Breeding: each Bred Vector is ~ a linear combination of unstable structures
3) If possible, compute as many BVs the number of unstable LVs
4) Otherwise: frequent analyses, localization, periodical reseeding,...

LIMITED-AREA MODELS AND BOUNDARY FORCING



- A limited area model is **forced** by lateral boundary conditions
- The **boundary forcing** has a **stabilizing effect**
- The same system is stabler in a smaller area
- Stable case: the boundary forcing determines the evolution
- In **unstable** cases, stability can be obtained by assimilating observations, then reducing errors
- Number and frequency of observations necessary to control the system depend on number of unstable directions and their growth rates

Stabilizing effects of different kinds of forcing:

- **Boundary:** forces the system trajectory to **approach** that of the **external model**
- **Data assimilation:** forces the system trajectory to **approach reality**

Multiple scale instabilities in non-hydrostatic, convection resolving systems

Convection-resolving system: **MOLOCH**

Non-hydrostatic, convection-resolving model developed at CNR -ISAC, Bologna
Malguzzi et al., JGR-Atmospheres, 2006; Davolio et al., MAP, 2007; 2009; ...

For this work:

- Resolution ~**2.2 km**.
- Domain: northern Italy Alps, part of Ligurian and Adriatic seas (Mediterranean).
- Initial and boundary conditions: BOLAM (hydrostatic LAM) and GFS.
- Control trajectory: simulation of a real case: (26 September 2007). Intense convective precipitation over the Venice area (north-eastern Italy). Scattered convection during the night, frontal-forced, organized convection in the day

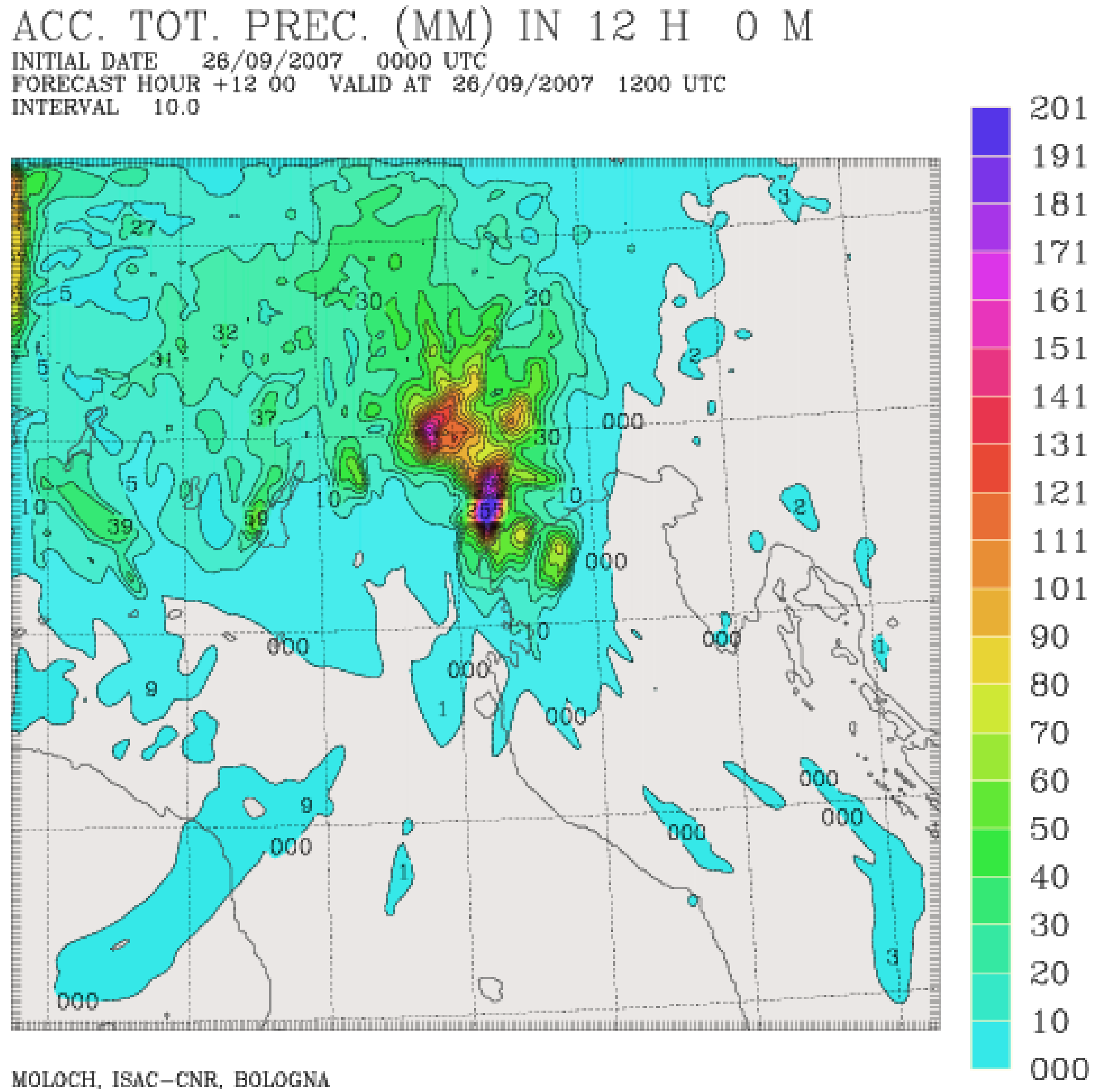


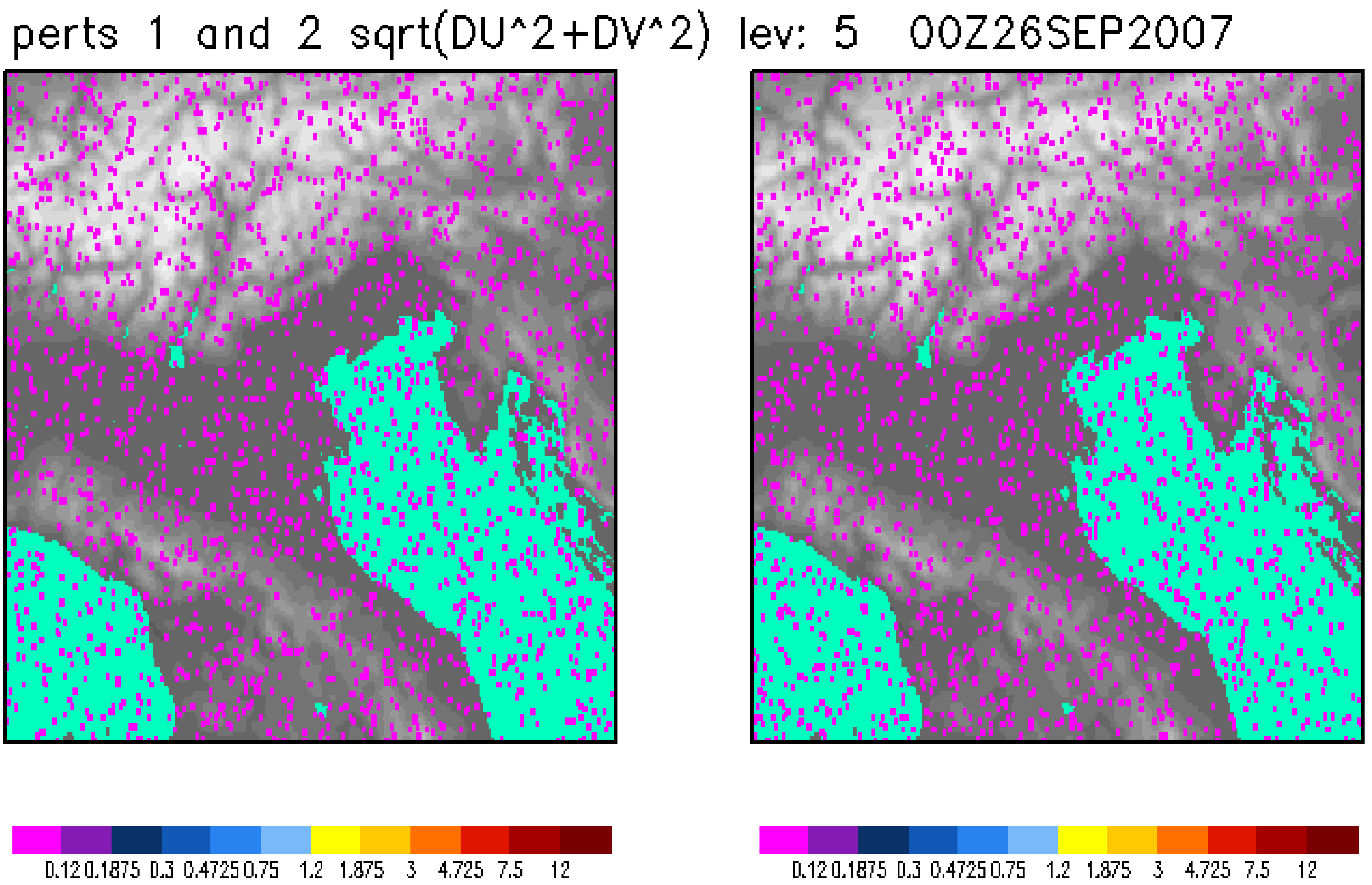
Figure: Total precipitation accumulated fro 00 to 12 UTC in control trajectory.

http://www.isac.cnr.it/dinamica/projects/forecasts/

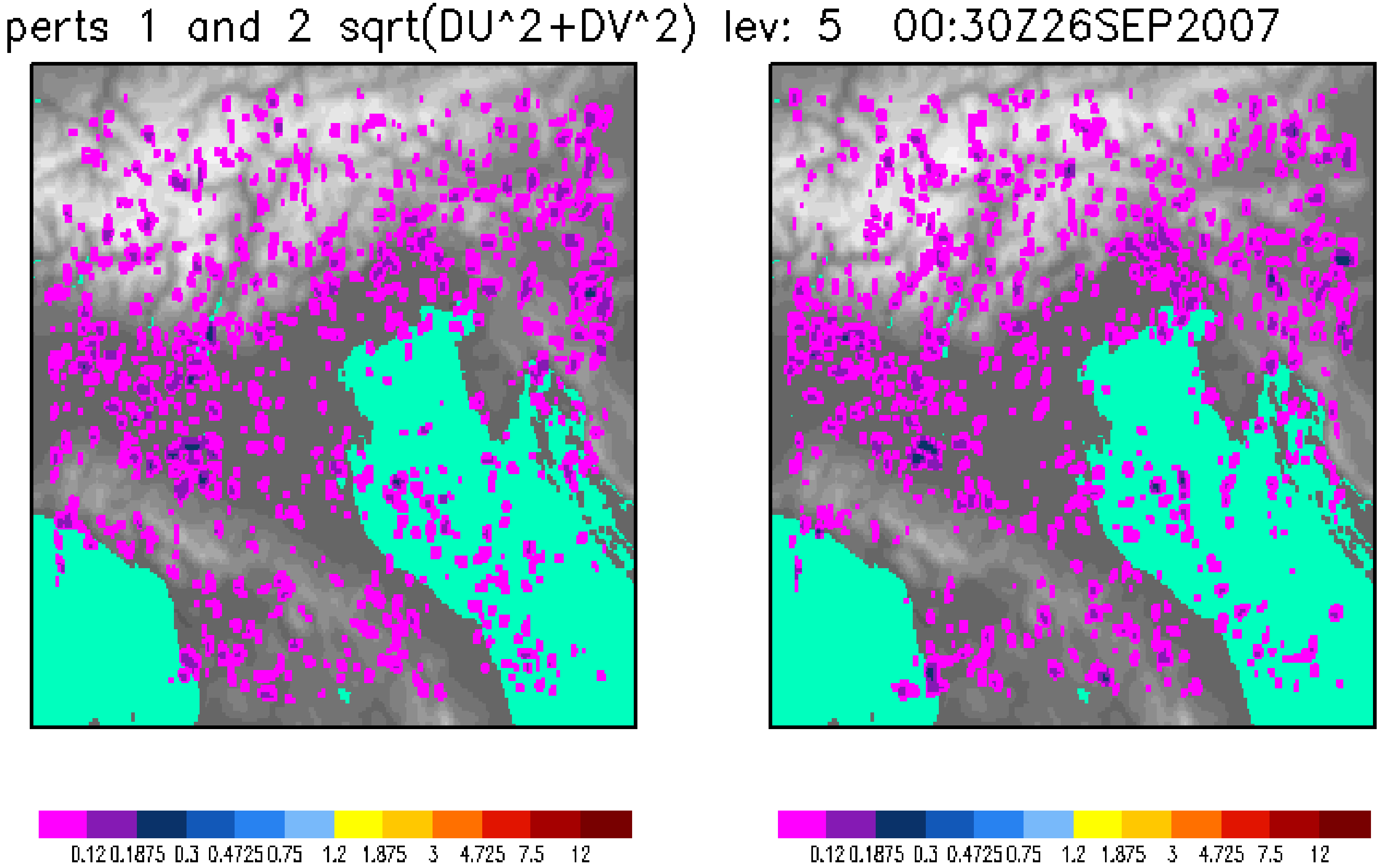
Two small, independent, randomly generated perturbations. Each variable scaled with its variability.

Breeding perturbations rescaled every 5 minutes, so that RMS of level 5 (~925hPa over sea) horizontal velocity is 0.05 m s^{-1}.

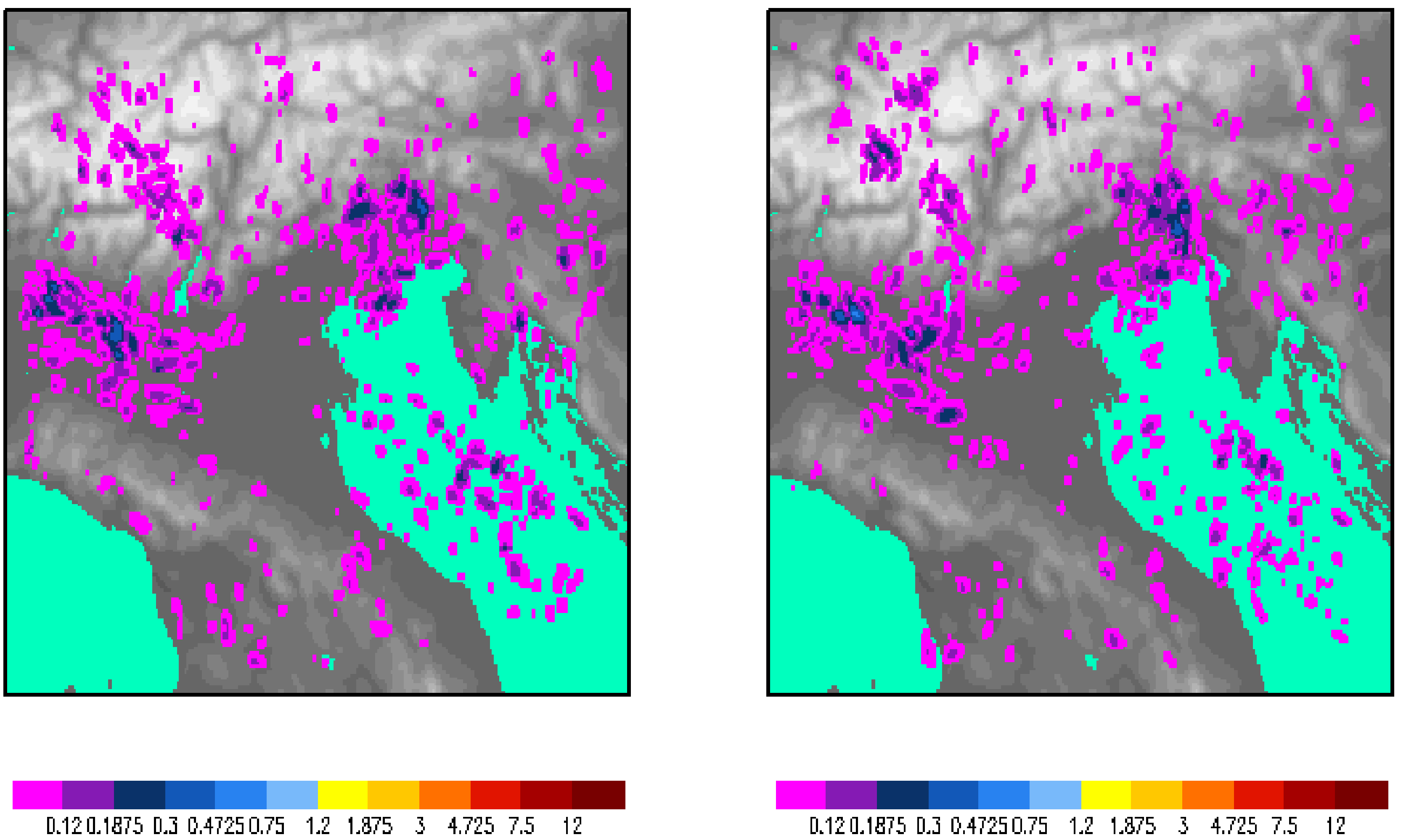
Instabilities related to fastest and smallest dynamical scales.



Module of wind vector difference (perturbed state – control state) at level 5 : $\sqrt{\delta u^2+\delta v^2}$

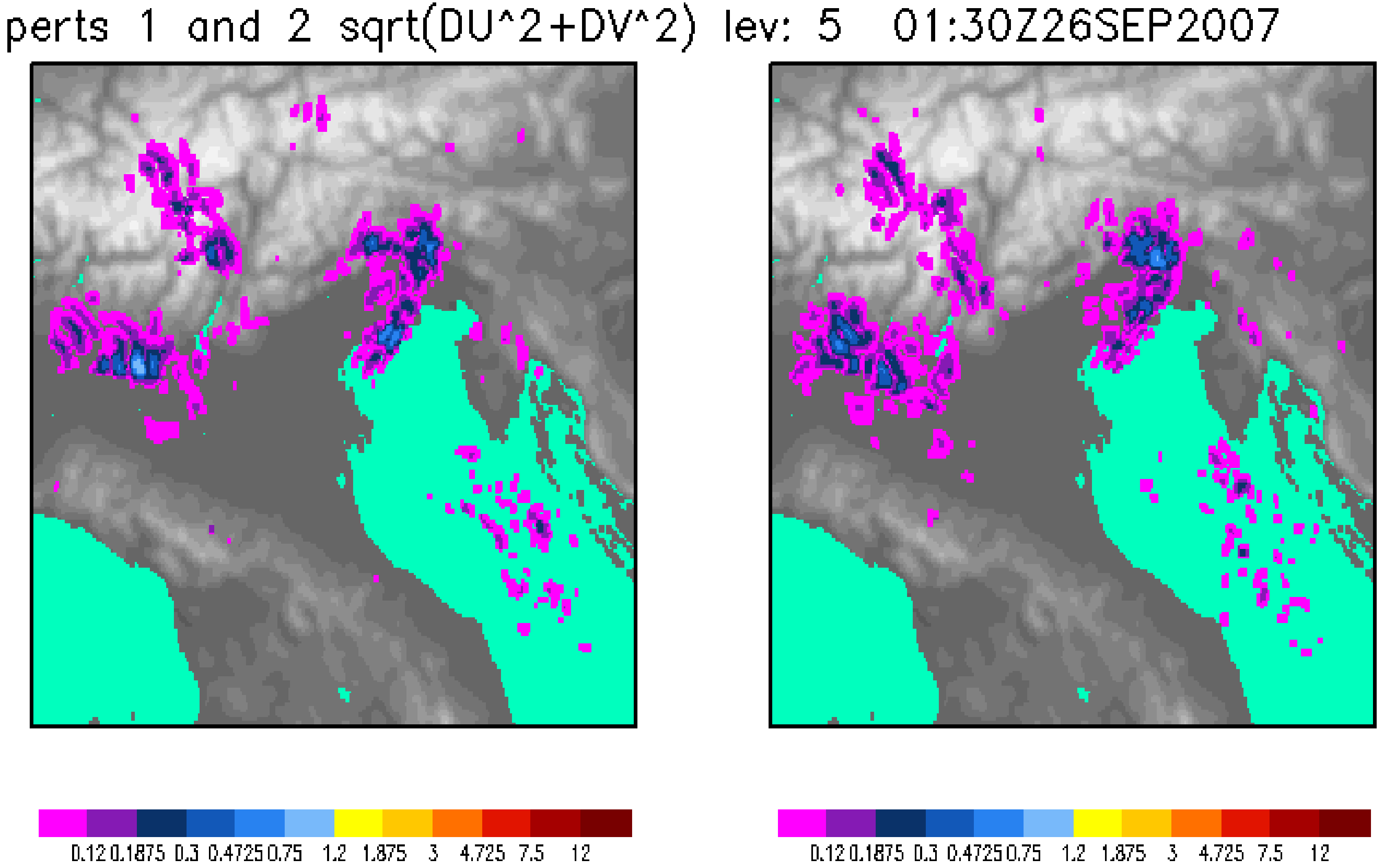

perts 1 and 2 sqrt(DU^2+DV^2) lev: 5 00:30Z26SEP2007
0.12 0.1875 0.3 0.4725 0.75 1.2 1.875 3 4.725 7.5 12
0.12 0.1875 0.3 0.4725 0.75 1.2 1.875 3 4.725 7.5 12

perts 1 and 2 sqrt(DU^2+DV^2) lev: 5 01Z26SEP2007



After 1h30 the bred vectors show organized and similar spatial structures, localized in dynamical active areas (intense winds and convective precipitation)

- Bred vectors quickly get organized and show spatially coherent structures.
- Small perturbation growth in the linear regime is not immediately suppressed by the strong non-linear processes of moist convection thermodynamics.
- Different structures for different re-normalization amplitudes and frequencies



Non-linear evolution from this time on: DOUBLING TIME ~ LINEARITY TIME ~ 2h ~ 2.5h

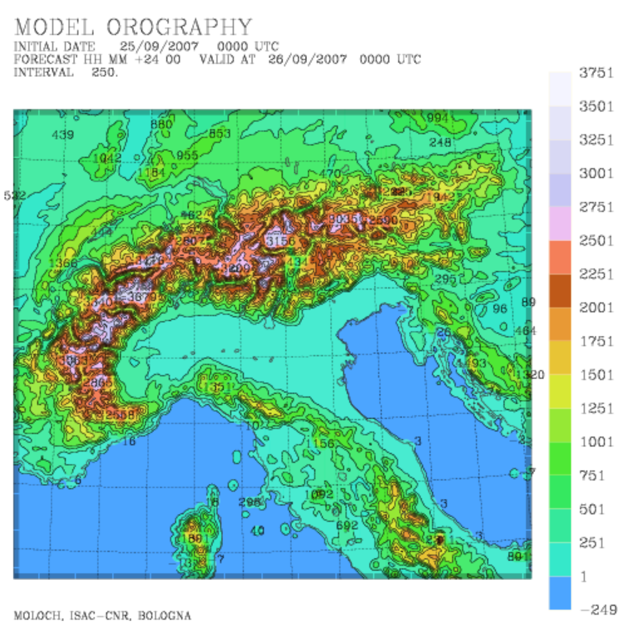
Multiple scale instabilities in non-hydrostatic, convection resolving systems

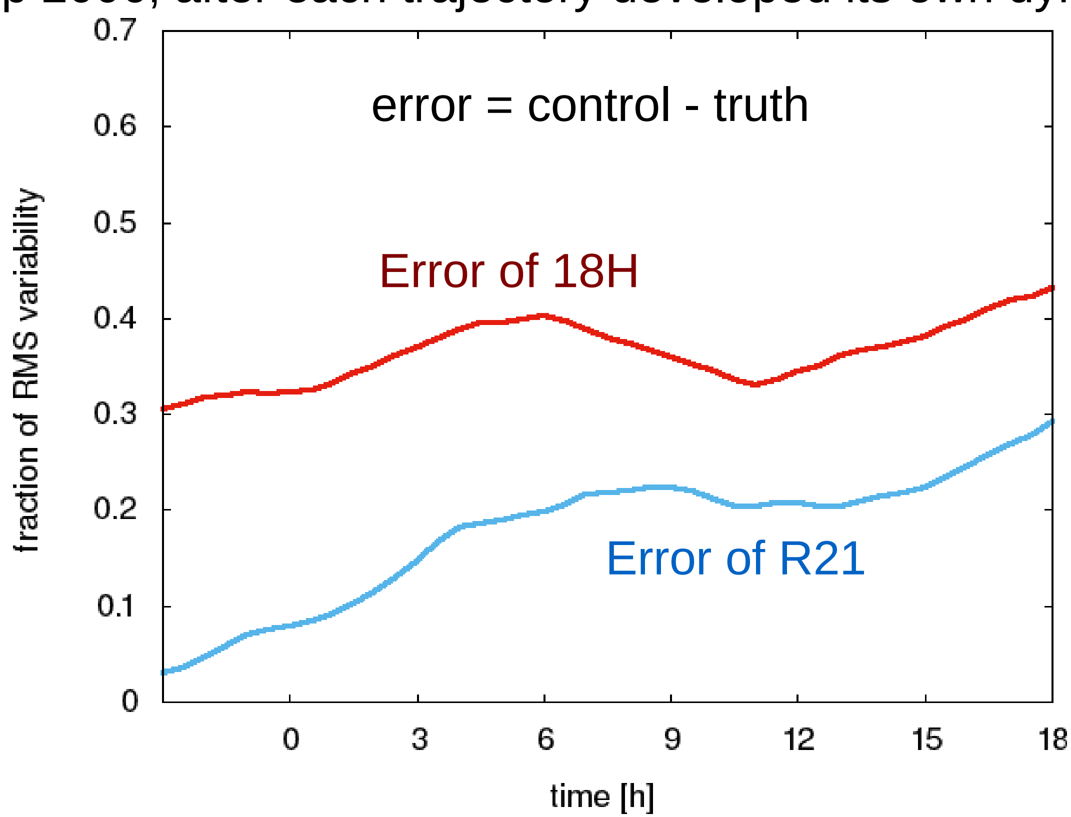
True trajectory: model trajectory from 21h of 25 Sep 2006 to 18h of 26 Sep 2006
Initial condition from external model

Control trajectory “**18H**” for **LARGE initial error, SLOW instabilities**:
initial condition from external model at 18h of 25 Sep 2006
– initial error: 0.4 °C 1.9 m/s at 1500hPa; 0.25 °C, 1.8 m/s at 500hPa

Control trajectory “**R21**” for **SMALL initial error, FAST instabilities**: same as 18H, but
error rescaled at 21h of 25 Sep 2006 so that (R21 -TRUTH) = 0.1 (18H - TRUTH)

Experiments start at 00h of 26 Sep 2006, after each trajectory developed its own dynamics





Multiple scale instabilities in non-hydrostatic, convection resolving systems

Lines: growth rates of non-linear trajectories forecast errors

with **LARGE initial amplitude (18H)**

and **SMALL initial amplitude (R21)**

$$e^{\lambda T_D} = 2 \qquad T_D = \frac{\ln(2)}{\lambda}$$

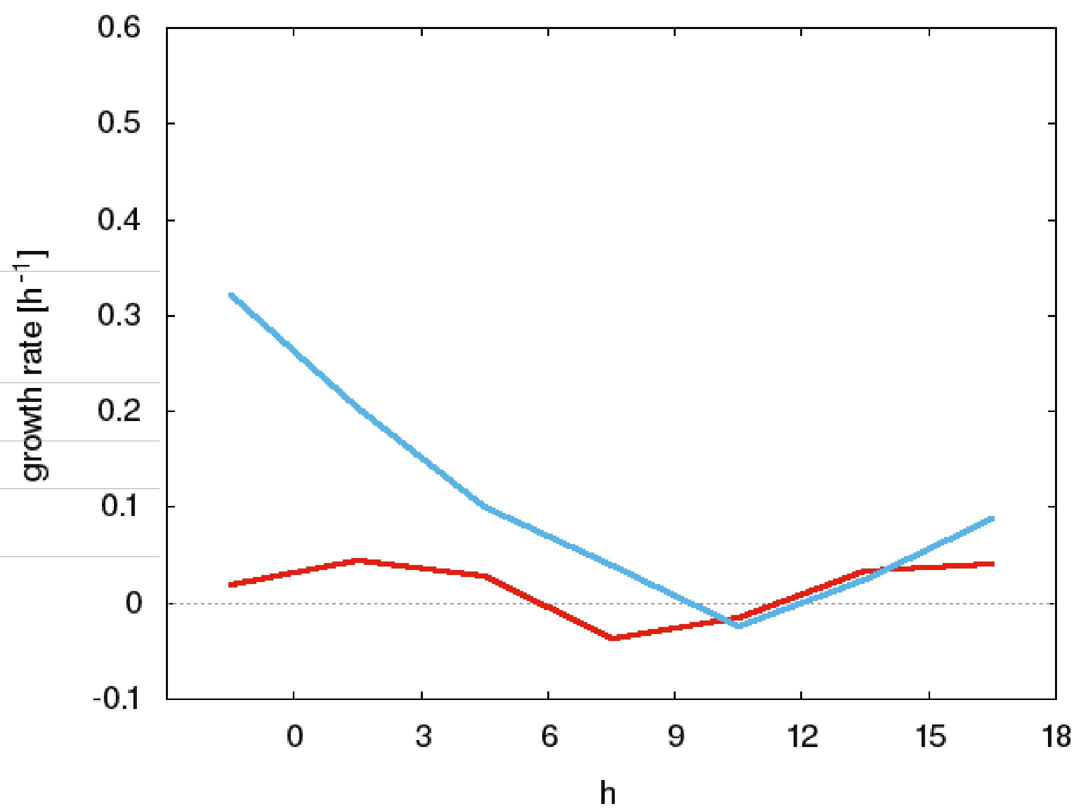
$$T_D \simeq 2\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.35$$

$$T_D \simeq 3\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.23$$

$$T_D \simeq 4\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.17$$

$$T_D \simeq 6\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.12$$

$$T_D \simeq 12\,\mathrm{h} \Leftrightarrow \lambda = 0.06$$



Multiple scale instabilities in non-hydrostatic, convection resolving systems

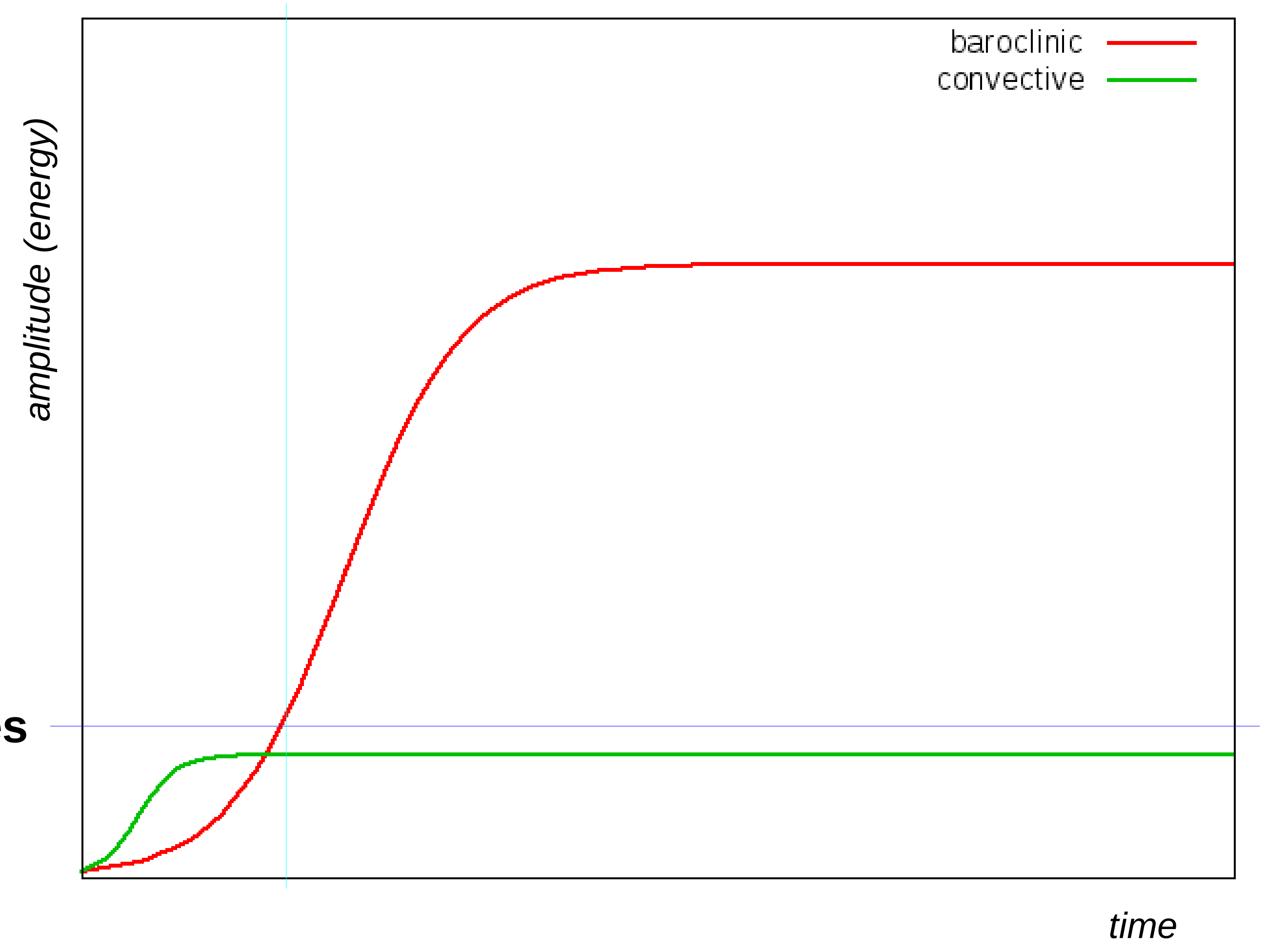
Characterization of errors

- When **errors** are very **small**, they **grow very fast,** T_d**~2.5h–7h** : convective scale instability

- **Larger errors grow more slowly,** T_d**~10h–14h**

- When error is large there also are **non-growing error components**:
  - Saturated small-scale fast instabilities
  - Larger-scale error structures present in an initial condition from a larger-scale hydrostatic model

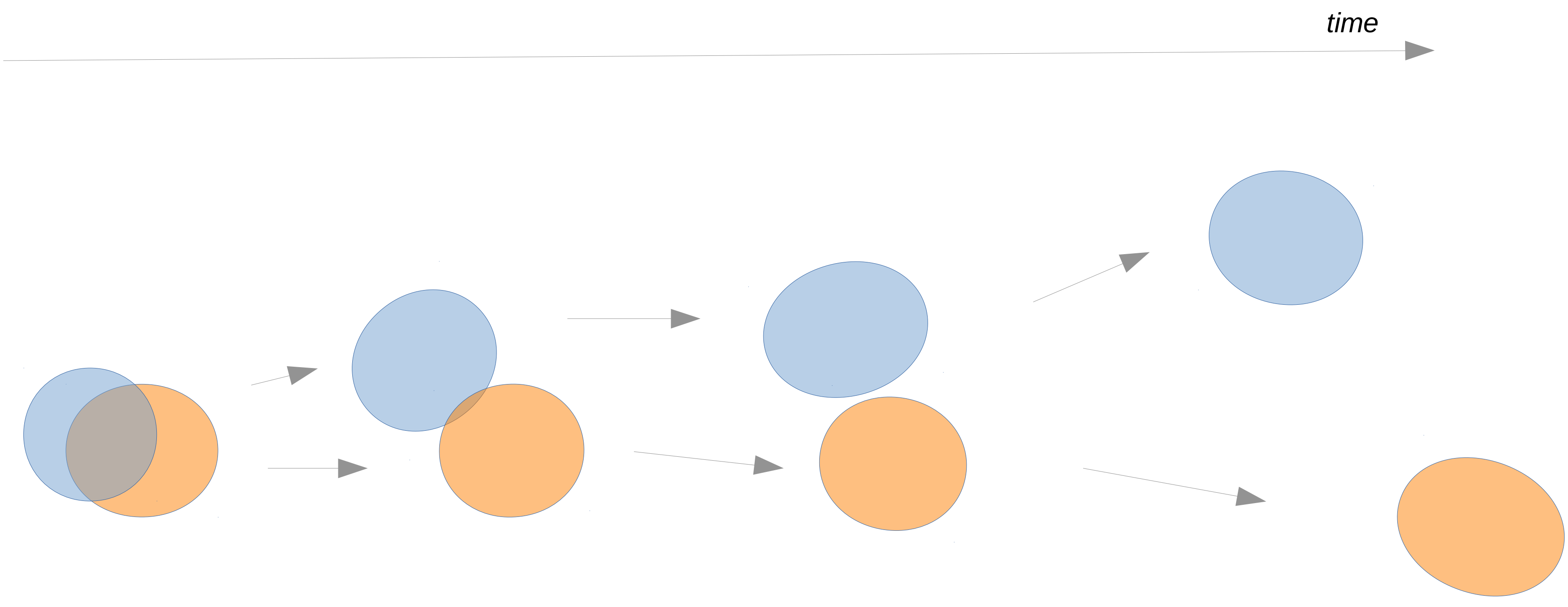
Breeding enables selection of instabilities relevant for forecast errors of a given typical amplitude (Toth and Kalnay,1997)



Non-linear Breeding
filters small and fast scales

Schematic example of small-scale non-linear saturation

A localized small-scale signal evolves differently in two model runs — or in a numerical forecast and in reality — about the same extension and intensity, but different center location, drifting apart.



~linear regime:
RMS error exponentially grows in time as the two signals separate

Distance exceeds size: end of linear regime

Non-linear saturation:
The two signals are completely separated: as they drift further away, the RMS error does not substantially grow anymore.

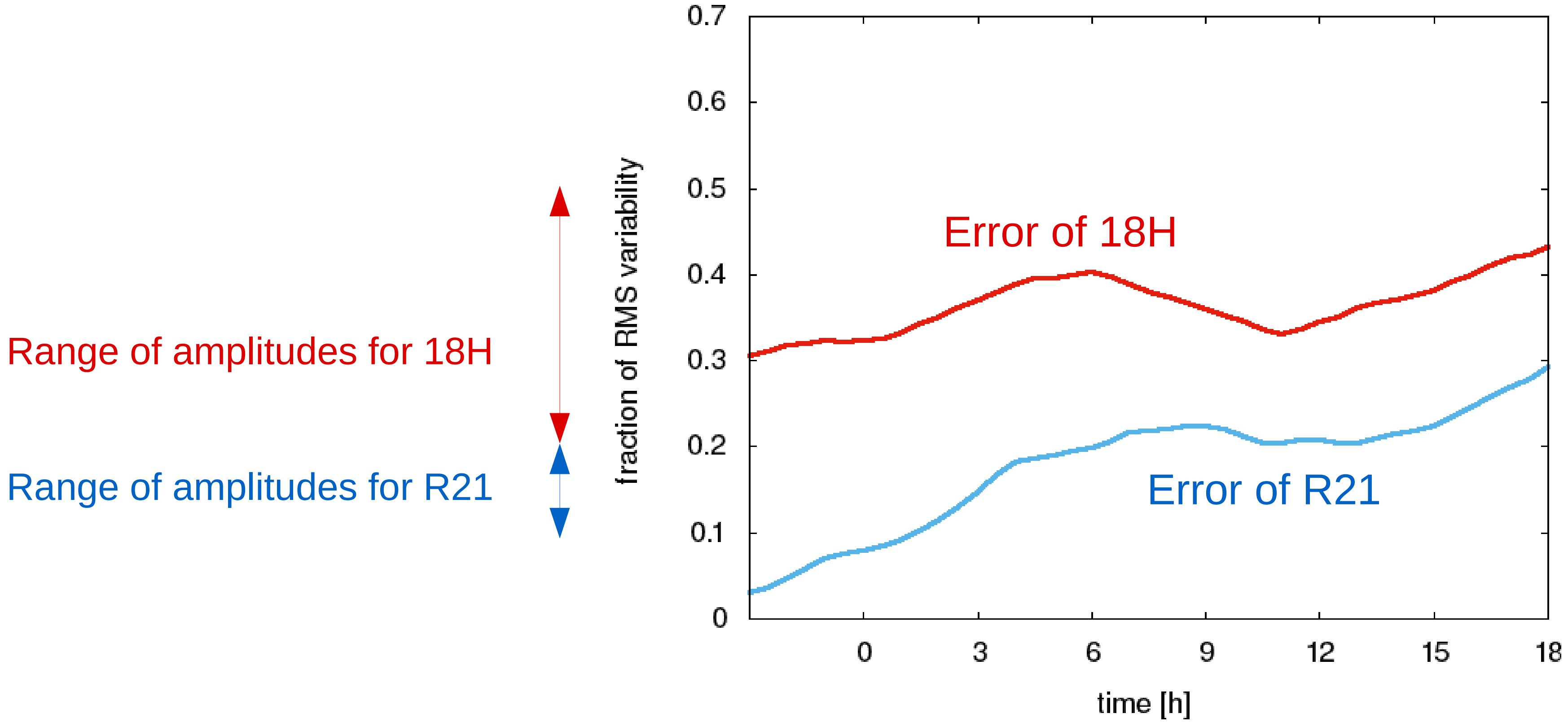
The large scale environment of the small scale signal is still predictable

Multiple scale instabilities in non-hydrostatic, convection resolving systems

Scalar product needed for orthogonalization:
sum of component products T and U,V normalized with their variabilities

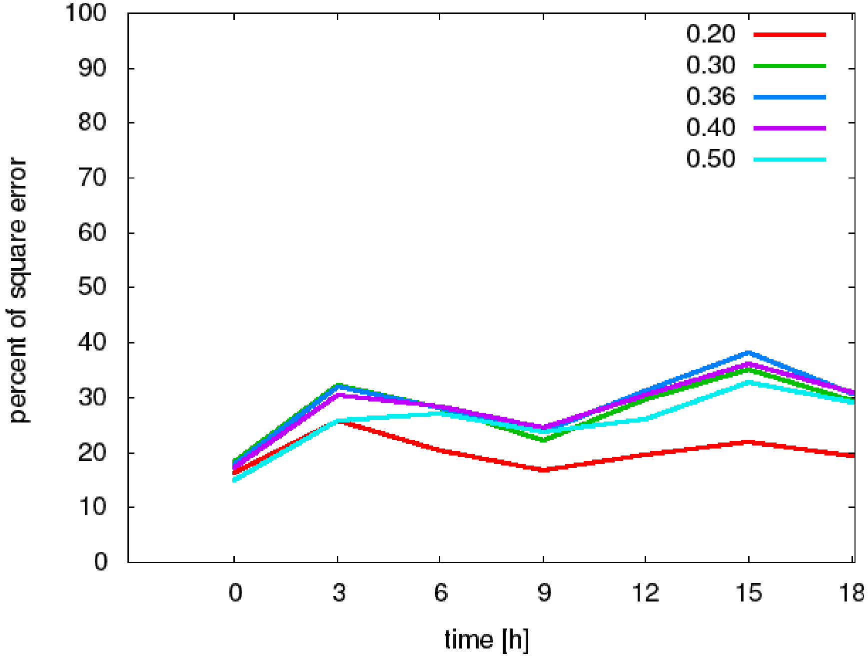
Breeding
LARGE – 18H : rescaling every 30 min — large(r)-scale slow instabilities

SMALL – R21 : rescaling every 15 min — small-scale fast instabilities



Multiple scale instabilities in non-hydrostatic, convection resolving systems

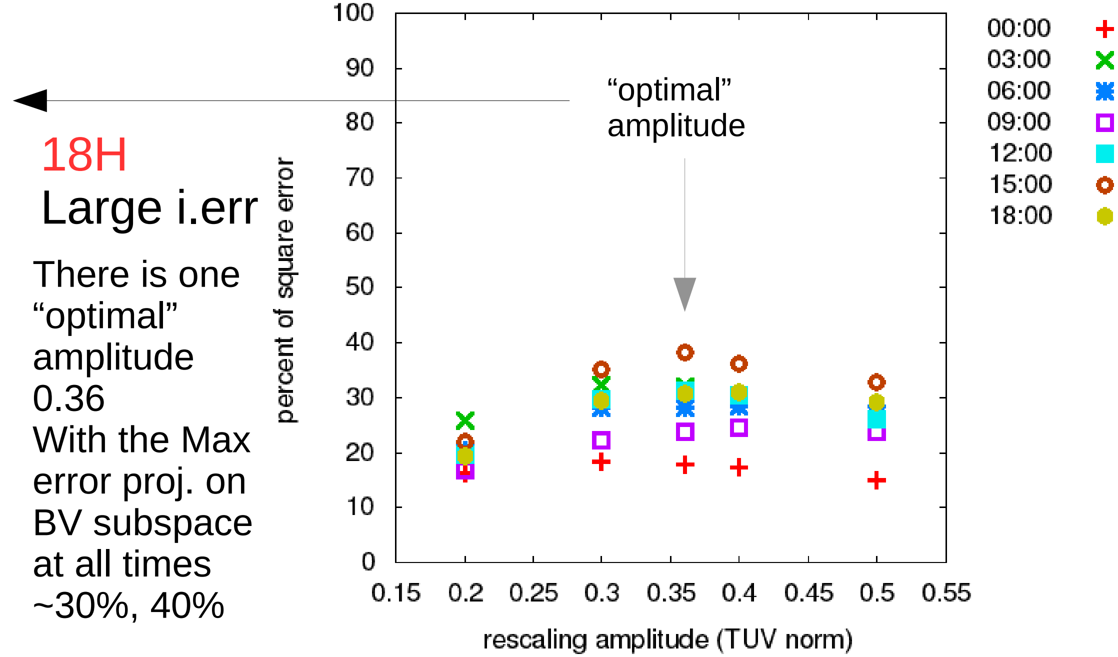
Square norm of error orthogonal projection onto **12**-BV subspace

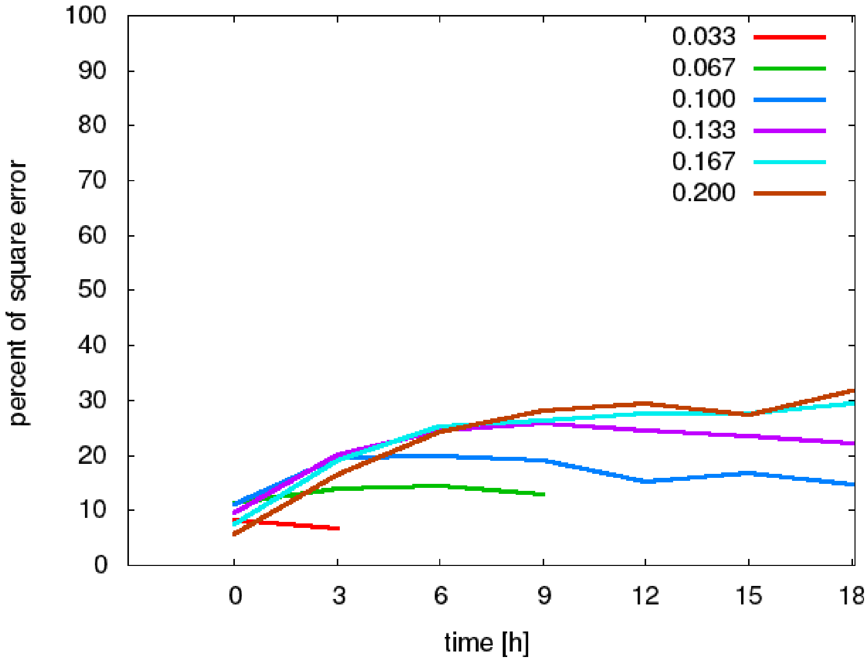


18H

Large i.err

There is one "optimal" amplitude 0.36
With the Max error proj. on BV subspace at all times ~30%, 40%



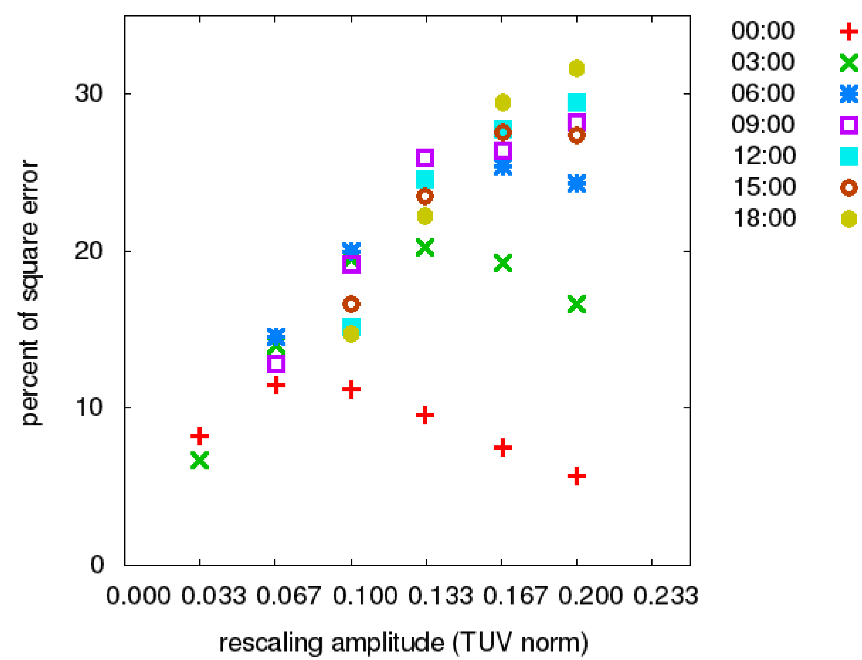


R21

Small i.err.

NO "optimal" amplitude

Max projection obtained for different amplitudes at different times



Multiple scale instabilities in non-hydrostatic, convection resolving systems

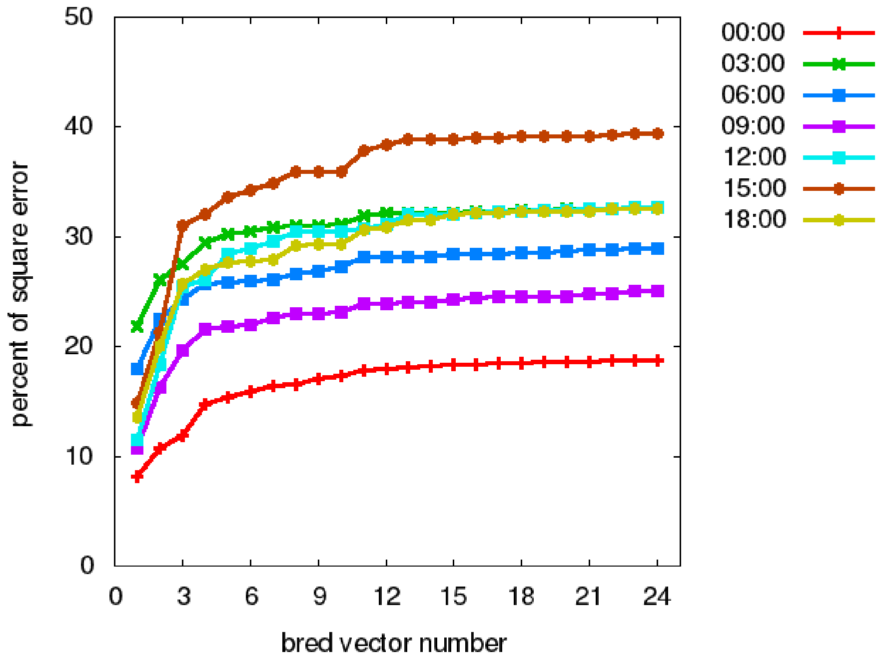
Square norm of error orthogonal projection onto BV subspaces: 1 to 24BVs

18H Large initial error amplitude

(larger, slower instabilities)

Increasing the subspace dimension is very effective at first, then the square error fraction in practice does not increase anymore.

Few BVs are sufficient to “explain” an important error portion.



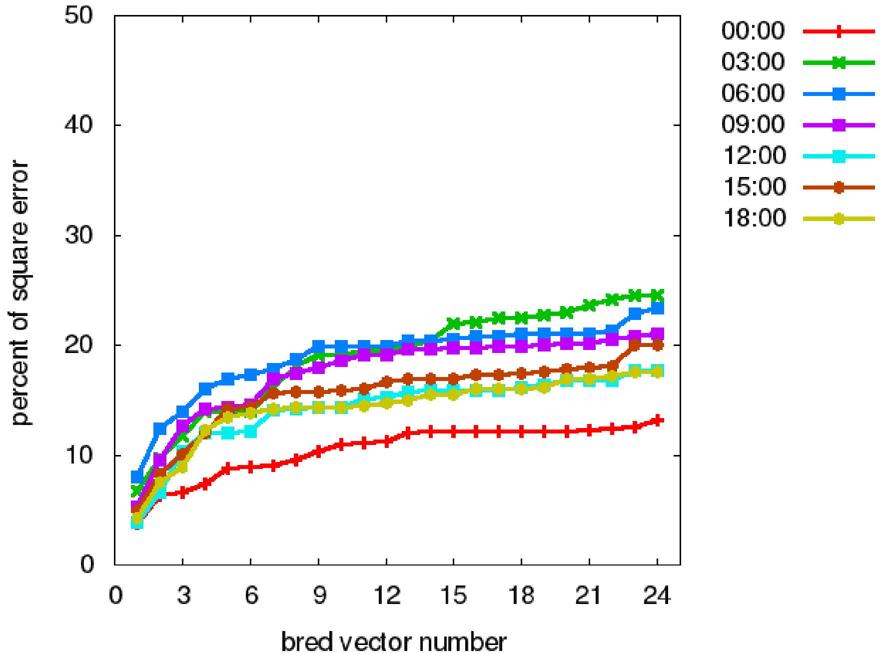
R21 Small initial error amplitude

(smaller, faster instabilities)

Slow regular increase:

Many Bvs (more than 24) determine each a small amount of error fraction

Many independent instabilities



Multiple scale instabilities in non-hydrostatic, convection resolving systems

Lines: growth rates of non-linear trajectories forecast errors

with **LARGE initial amplitude**

and **SMALL initial amplitude**

Marks: their **first** BV:

Larger amplitude, less frequent rescaling: 0.36, 30 min

Smaller amplitude, more frequent rescaling: 0.100, 15 min

Correspondence: BVs contain the same instabilities as the forecast error

$$e^{\lambda T_D} = 2$$

$$T_D = \frac{\ln(2)}{\lambda}$$

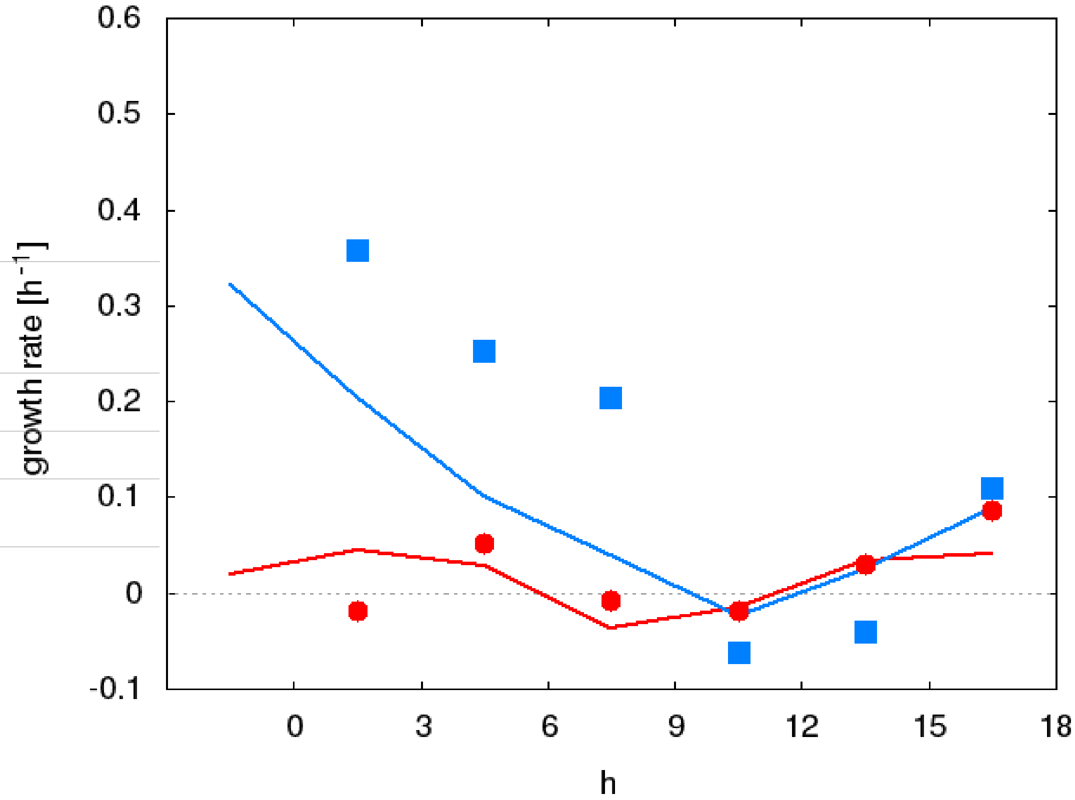
$T_D \simeq 2\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.35$

$T_D \simeq 3\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.23$

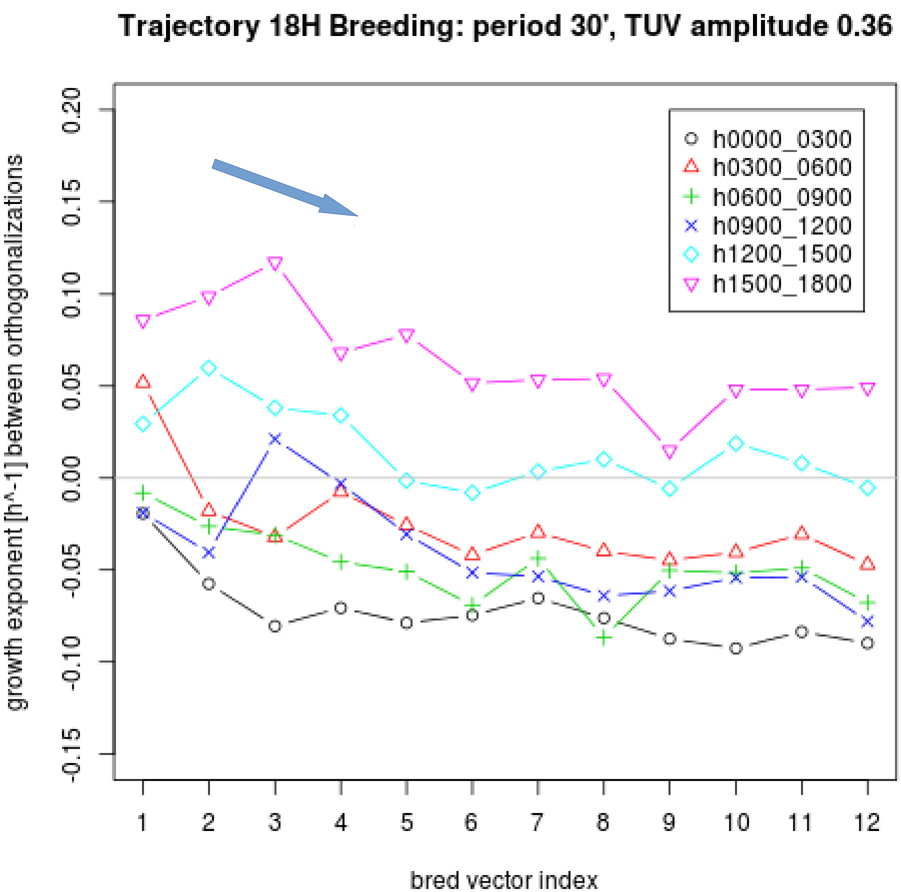
$T_D \simeq 4\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.17$

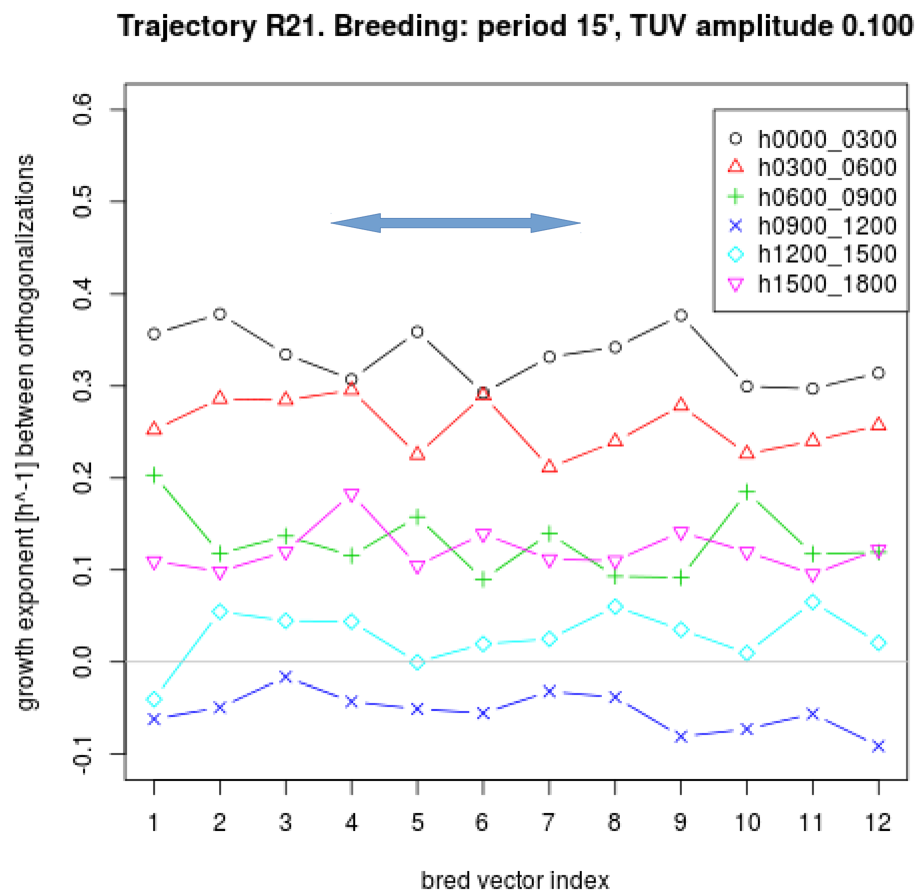
$T_D \simeq 6\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.12$

$T_D \simeq 12\,\mathrm{h} \Leftrightarrow \lambda = 0.06$



LARGE, SLOW instabilities ←*...comparatively...*→ SMALL, FAST instabilities





BOTH: time variability (different curves at different hours) OK: **larger BV growth rates** at times **when** the respective **forecast error increases**

LARGE, SLOW instabilities:
- Growth rate **decreases with BV index**
- All positive only at 15:00-18:00 (slower scales dominant), few positive otherwise.

⇒ FEW unstable directions at LARGE scale

SMALL, FAST instabilities
- Growth rate **does NOT decrease with BV index: flat spectrum**
- *Always all positive except at 09:00-12:00*

⇒ MANY unstable direction at SMALL scale

BRED VECTORS AND INSTABILITIES – RESULTS

- **BVs amplitude** of about the order of the **analysis error**:
  - Growth rate decrease with Bred Vector (BV) index
  - Doubling times 10-14 h
  - Small number of actively unstable BVs
  - Projection of error onto 12-BV subspace:
    - Most of it on the leading BVs, It does not increase much from 12 to 24 BVs

- **BVs amplitude** about **1/10** of the order of the **analysis error**:
  - The spectrum of BVs is flat: many BVs with competitive – large – growth rates
  - Doubling times 2-7 h
  - Projection of error onto 12-BV subspace:
    - Small
    - Slowly and steadily increasing with BV index: many BVs needed.
    - The unstable subspace of convective scale may have a very large dimension

Multiple scale instabilities in non-hydrostatic, convection resolving systems

WHAT TO DO

- Frequent analyses – every hour at least

- **DO NOT restart from** external, larger-scale, **hydrostatic** model, **initial condition**
  use the external model for boundary conditions only
  make use of Data Assimilation to control the trajectory

- Localization techniques
- Periodical reseeding
- Multiple scale breeding /ensemble

- Study fast small convective-scale instabilities in a much smaller domain

Multiple scale instabilities in non-hydrostatic, convection resolving systems

WHAT TO DO

- Frequent analyses – every hour at least

- **DO NOT restart from** external, larger-scale, **hydrostatic** model, **initial condition**
  use the external model for boundary conditions only
  make use of Data Assimilation to control the trajectory

- Localization techniques
- Periodical reseeding
- Multiple scale breeding /ensemble

- Study fast small convective-scale instabilities in a much smaller domain

THANK YOU FOR YOUR ATTENTION

f.uboldi@aria-net.it