Predictability and data assimilation issues in multiple-scale convection-resolving systems

Francesco UBOLDI

Milan, Italy, *uboldi@magritte.it*

MET-NO, Oslo, Norway

23 March 2017

Predictability and data assimilation issues in multiple-scale convection-resolving systems

Topics

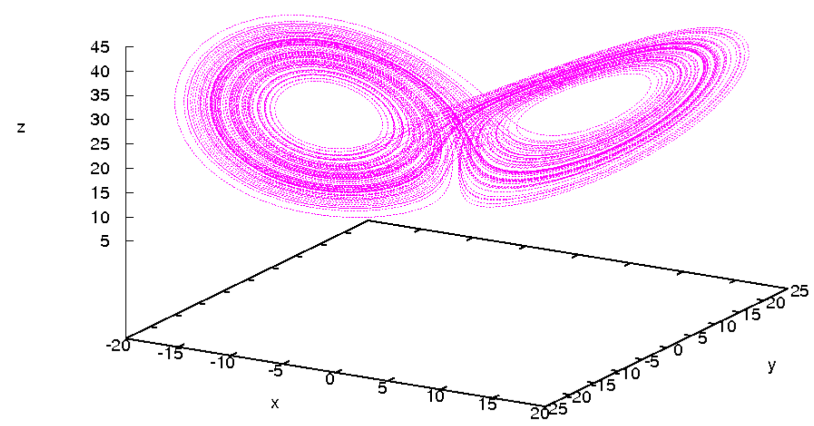
1. Error growth and unstable directions.
2. Assimilation in the unstable subspace (AUS)
3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

1. Error growth and unstable directions

Attractors and sensitivity to initial conditions

Non linear dynamical system: sensitivity to initial condition

A small perturbation of the initial condition originates a trajectory which progressively detaches from the unperturbed trajectory. After a sufficiently long evolution time, the two states (on the perturbed and on the unperturbed trajectory) are as far away from each other as two states randomly chosen among all those that are physically possible. **They cannot go indefinitely far from each other!**



1. Error growth and unstable directions

Non-linear growth of small perturbations

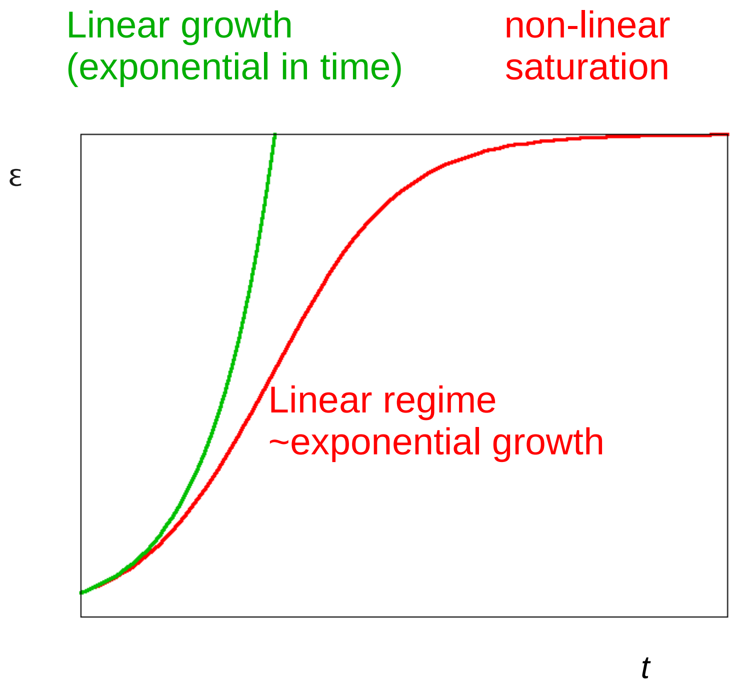
1) Early stages of growth:

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

1) Linear regime: exponential growth

2) End of linear regime

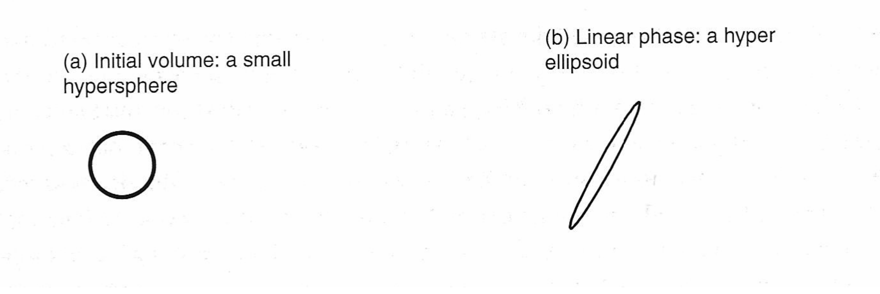
3) Saturation

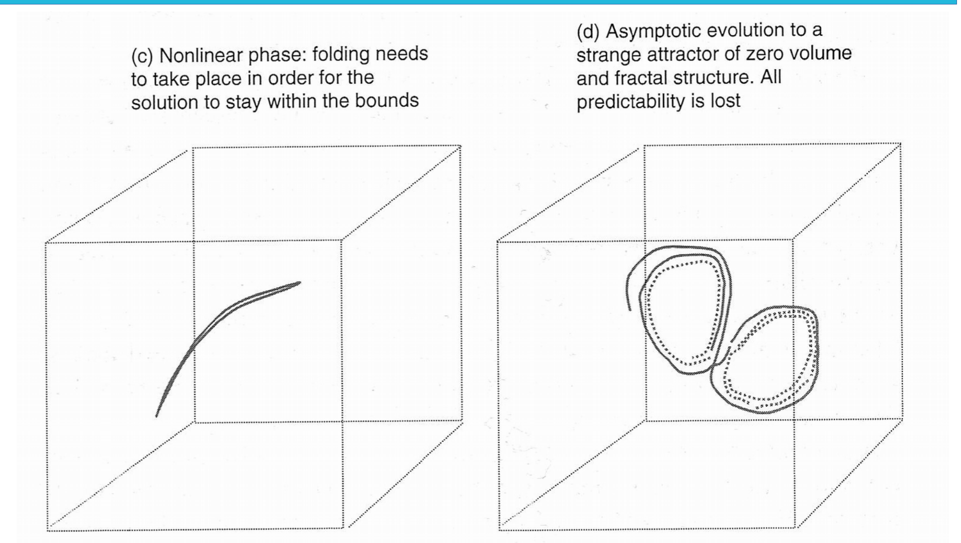


1. Error growth and unstable directions

Non-linear growth of small perturbations

- Linear regime: a small (hyper-) sphere, a “cloud” of possible initial states initially evolves into an (hyper-) ellipsoid, **extending along the unstable directions**, and **shrinking along the stable directions**.
- **Each direction** is characterized by its ***growth rate***: positive for unstable directions, negative for stable directions.
- In dissipative systems, shrinking prevails: (hyper-) volumes decrease.
- At the end of the linear regime the ellipsoid deforms: trajectories follow the shape of the attractor.
- As time proceeds, the initial “cloud” progressively distributes over the whole attractor.





Figures from: Kalnay, 2003.

1. Error growth and unstable directions

TL dynamics: **unstable directions** = **Lyapunov vectors with positive exponents**

- (Oseledec, 1968) global Lyapunov exponents: average on the whole attractor of local *growth rates* – time limit along a trajectory ergodically visiting the whole attractor.
- Benettin et al. (1980): method to compute Lyapunov exponents: frequent Gram-Schmidt ortho-normalization of *n* independent perturbations:
  - maintains the growth in the linear regime
  - prevents that all directions collapse onto the most unstable one
- **Lyapunov** characteristic vectors are “co-variant” with the phase flow
- NO dependence on scalar product:
  - **Lyapunov *exponents***
  - The **first Lyapunov vector**
  - The **subspaces spanned by Lyapunov vectors**

  DO NOT depend on the choice of the scalar product
- ⇒ **Lyapunov** characteristic vectors ARE NOT ORTHOGONAL

Legras and Vautard 1997; Trevisan and Pancotti 1998; Kalnay 2003; Wolfe and Samelson 2007; Ginelli et al. 2007

2. Assimilation in the unstable subspace (AUS)

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

Confine the analysis increment into the unstable subspace:

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

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

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

How many LVs?: dimension of the unstable subspace (of the forced system) =

= number of 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 analysis, **localization**...

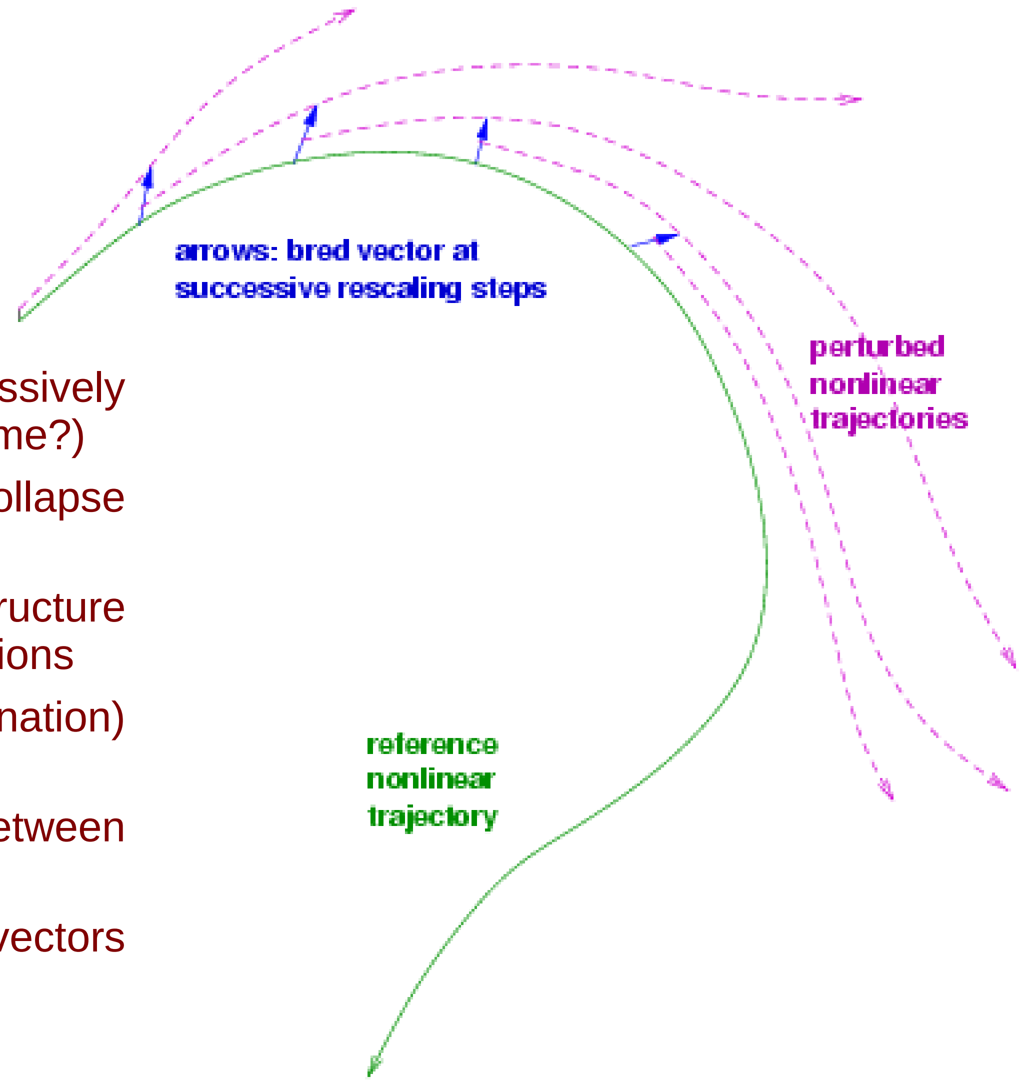
AUS successfully applied to many systems of different complexity, and compared to EKF, 4d-Var... References: *http://www.magritte.it/francesco.uboldi/bdasaus_papers.html*

2. Assimilation in the unstable subspace (AUS)

Breeding

- **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** (in how much time?)
- **MANY** initially independent perturbations collapse onto **FEW** directions in a **SHORT** time!
- A **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: to keep bred vectors independent)

2. Assimilation in the unstable subspace (AUS)

Data assimilation system

$$\mathbf{x}_{k+1}^{f} = M \circ \mathbf{x}_{k}^{a}$$

Forecast (nonlinear model integration)

$$\mathbf{x}_{k+1}^{a} = (\mathbf{I} - \mathbf{K} H \circ)\, \mathbf{x}_{k+1}^{f} + \mathbf{K}\, \mathbf{y}_{k+1}^{o}$$

Analysis

$$\Downarrow$$

$$\mathbf{x}_{k+1}^{a} = (\mathbf{I} - \mathbf{K} H \circ)\, M \circ \mathbf{x}_{k}^{a} + \mathbf{K}\, \mathbf{y}_{k+1}^{o}$$

It is a dynamical system FORCED by the assimilation of observations

Errors (at first order):

$$\boldsymbol{\eta}_{k+1}^{a} = \boxed{(\mathbf{I} - \mathbf{K}\mathbf{H})}\,\boxed{\mathbf{M}}\,\boldsymbol{\eta}_{k}^{a} + (\mathbf{I} - \mathbf{K}\mathbf{H})\,\boldsymbol{\eta}_{k+1}^{M} + \mathbf{K}\,\boldsymbol{\varepsilon}_{k+1}^{o}$$

Perturbations (at first order):

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

decrease (depending on K, assimilation goal)

growth

2. Assimilation in the unstable subspace (AUS)

Breeding on the data assimilation system (BDAS)

Freely evolving system: standard breeding

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

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

$$\boldsymbol{\eta}^f_{k+1} = \mathbf{M}\boldsymbol{\eta}^a_k + \boldsymbol{\eta}^M_{k+1}$$

System forced by data assimilation: BDAS "Breeding on the Data Assimilation System"

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

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

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

BDAS: 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

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

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 observation, 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:

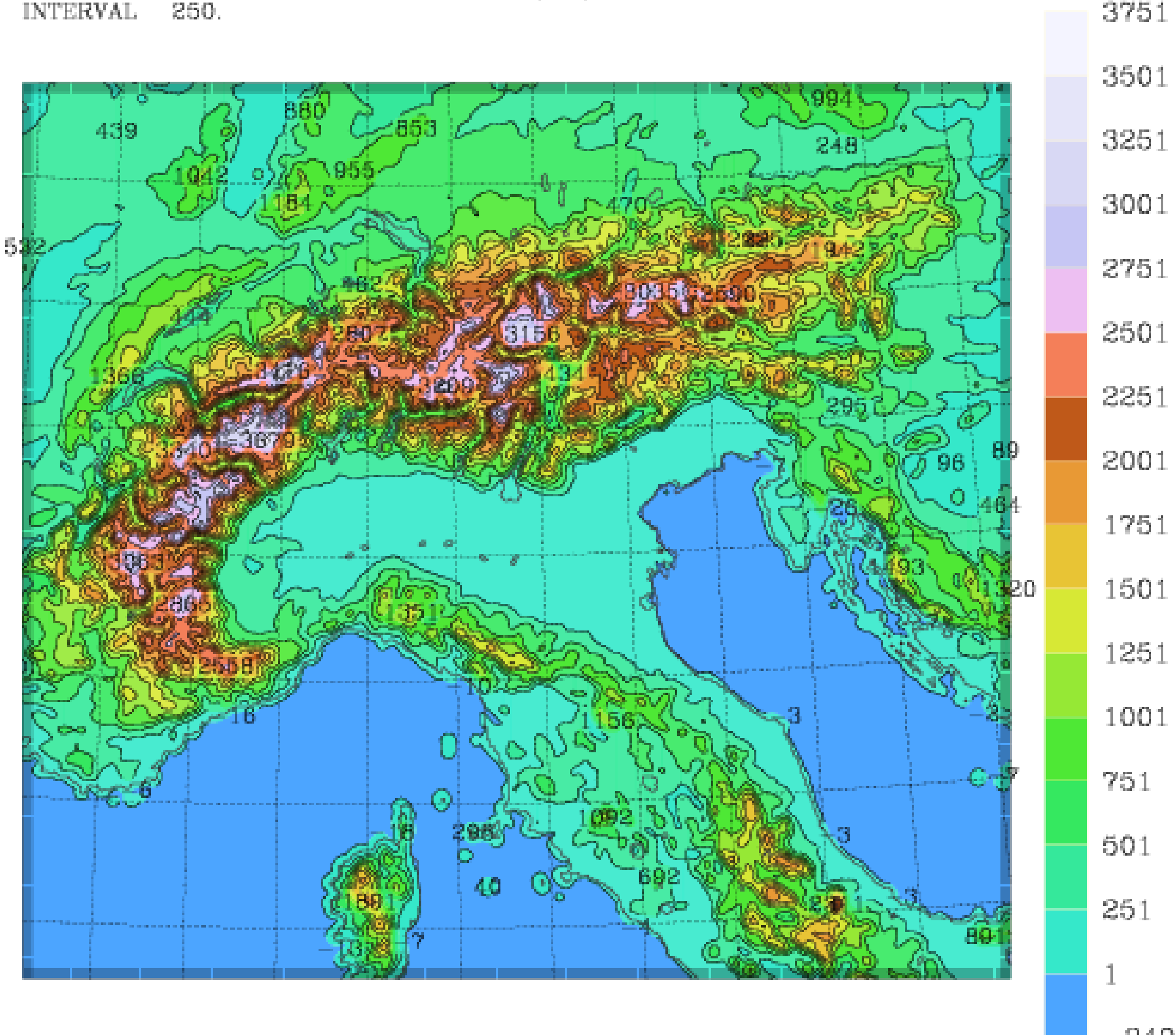
- **Boundary:** forces the system trajectory to approach that of the **external model**
- **Data assimilation:** forces the system trajectory to approach **reality**

MODEL OROGRAPHY

INITIAL DATE 25/09/2007 0000 UTC

FORECAST HH MM +24 00 VALID AT 26/09/2007 0000 UTC

INTERVAL 250.



MOLOCH, ISAC−CNR, BOLOGNA

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

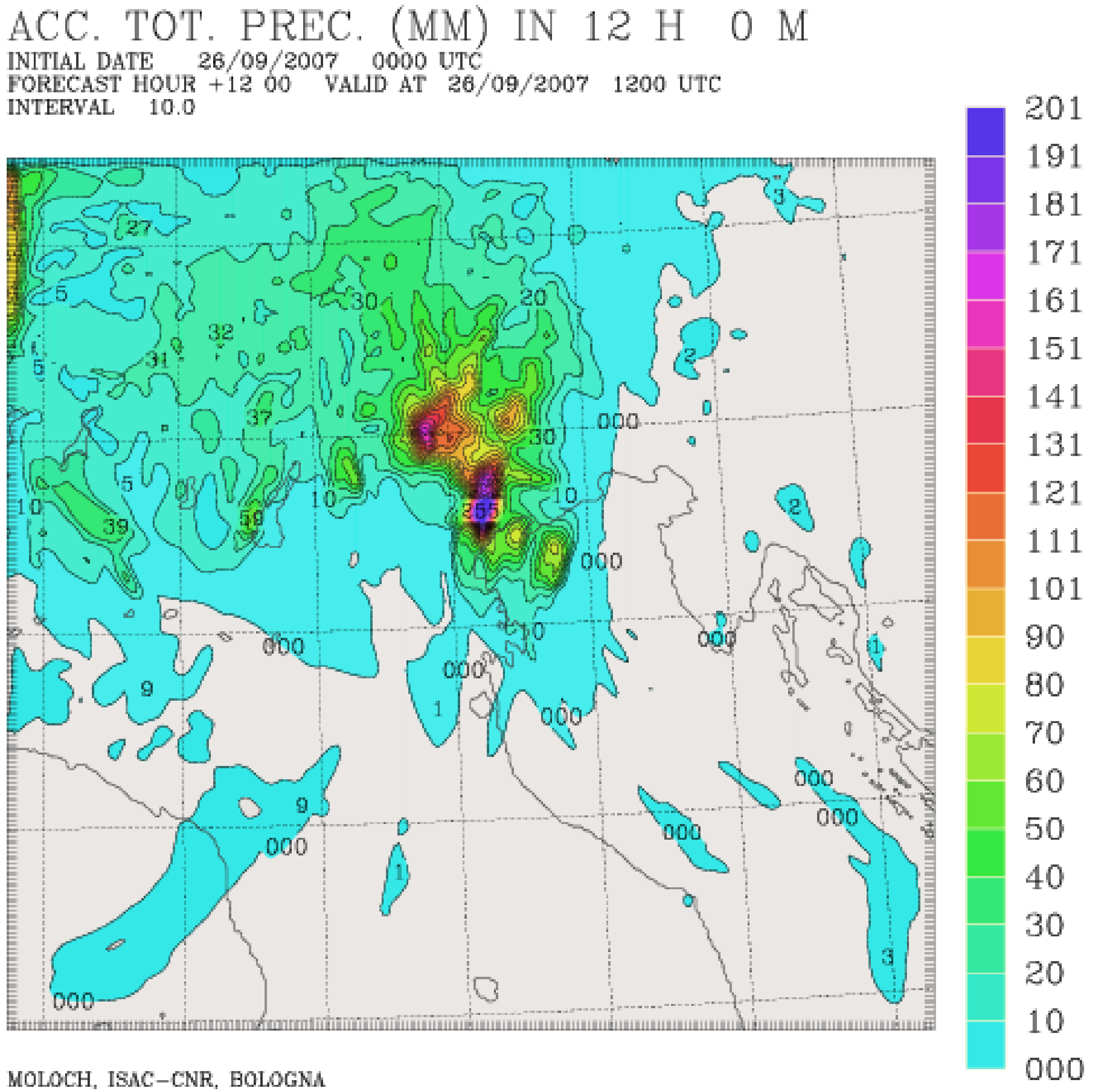
Convection-resolving system: MOLOCH

MOLOCH: non-hydrostatic, convection-resolving model developed at CNR -ISAC, Bologna

For this work:

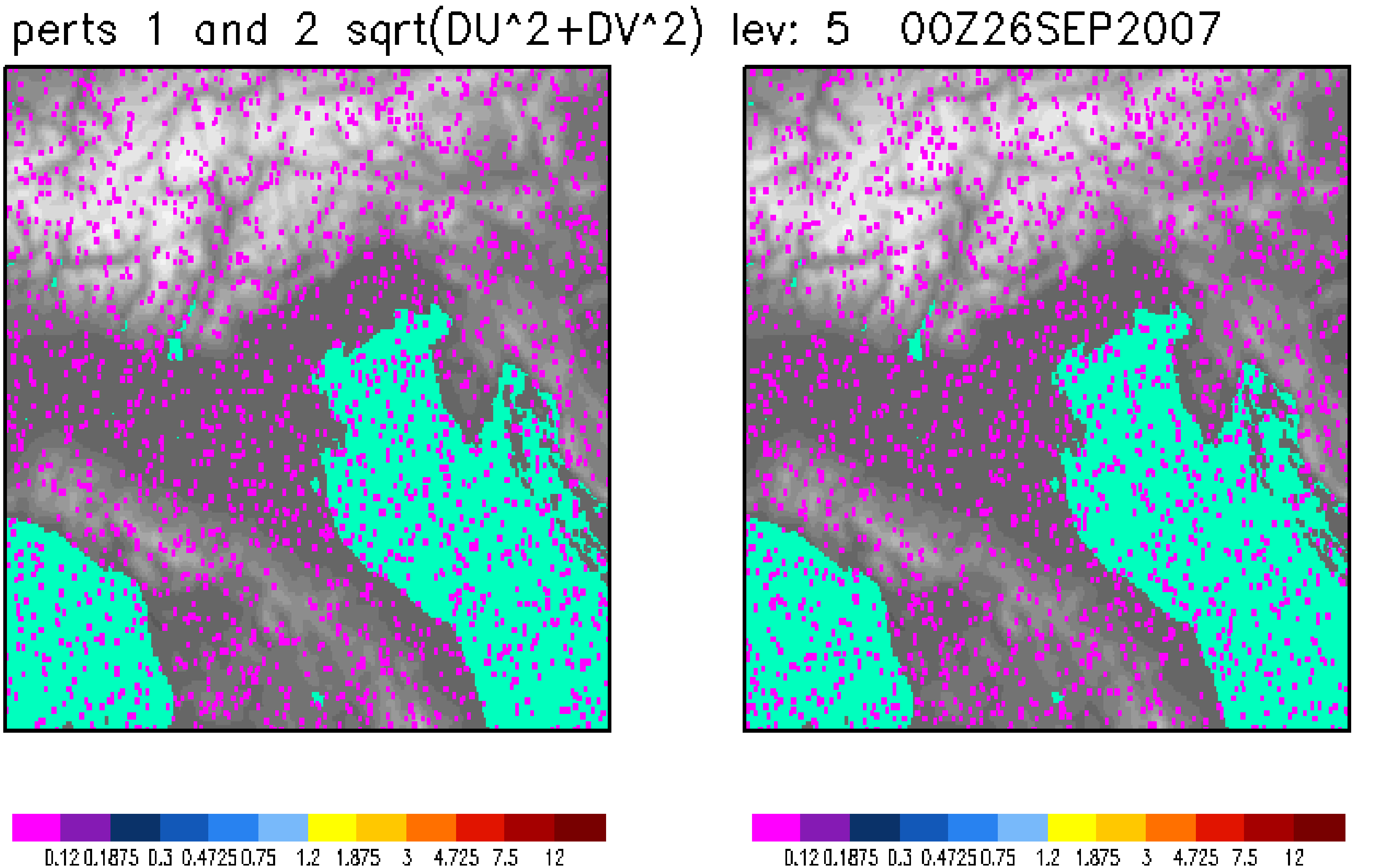
- Resolution ~**2.3 km**.
- Domain: northern Italy Alps, part of Ligurian and Adriatic seas (Mediterranean).
- Initial and boundary conditions: BOLAM (hydrostatic LAM) and GFS.
- Control trajectory: simulation of a real case: (26 September 2007). Intense convective precipitation over the Venice area (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.

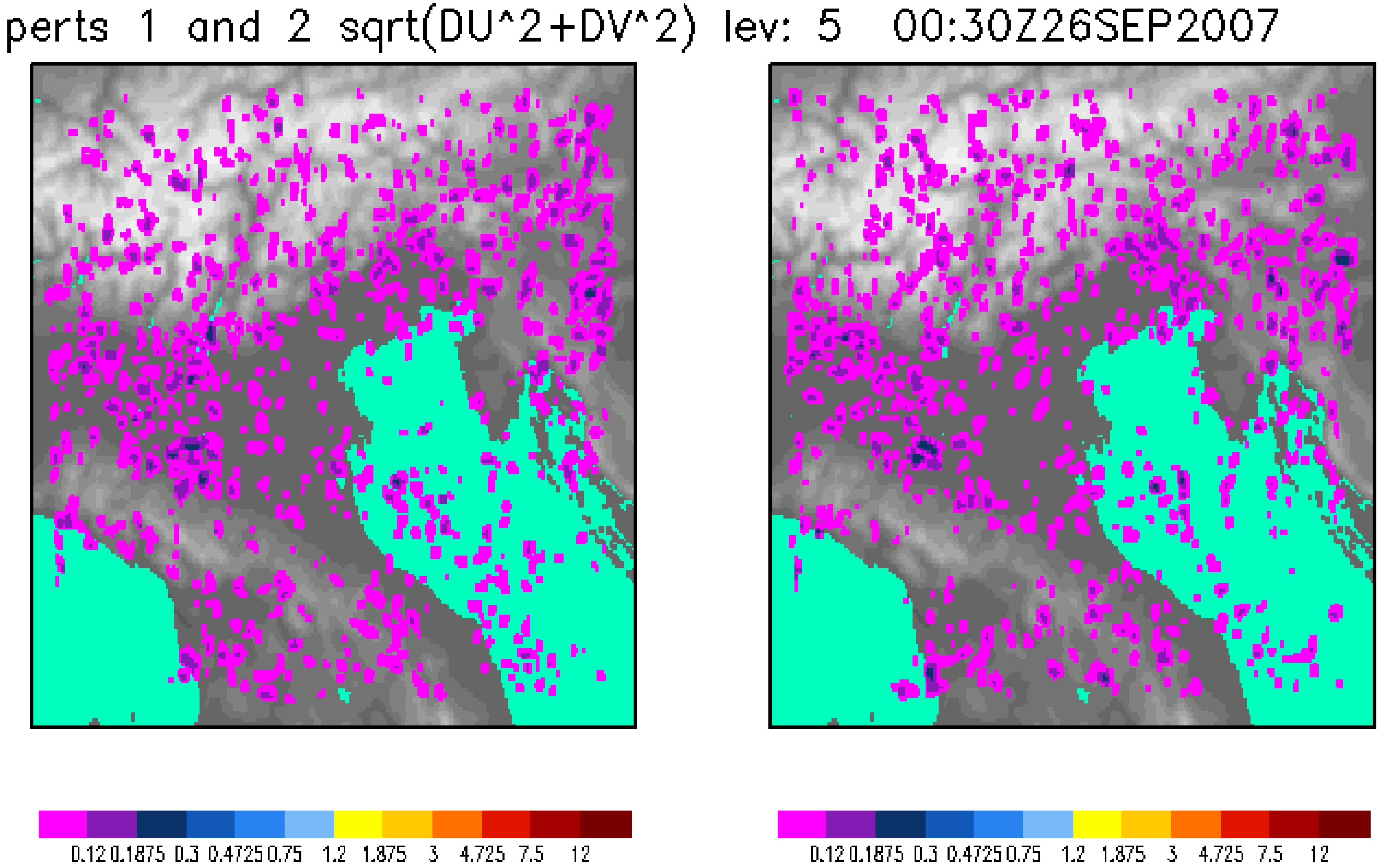


MOLOCH: *Malguzzi et al., JGR-Atmospheres, 2006; Davolio et al., MAP, 2007; 2009;* **http://www.isac.cnr.it/dinamica/projects/forecasts/**

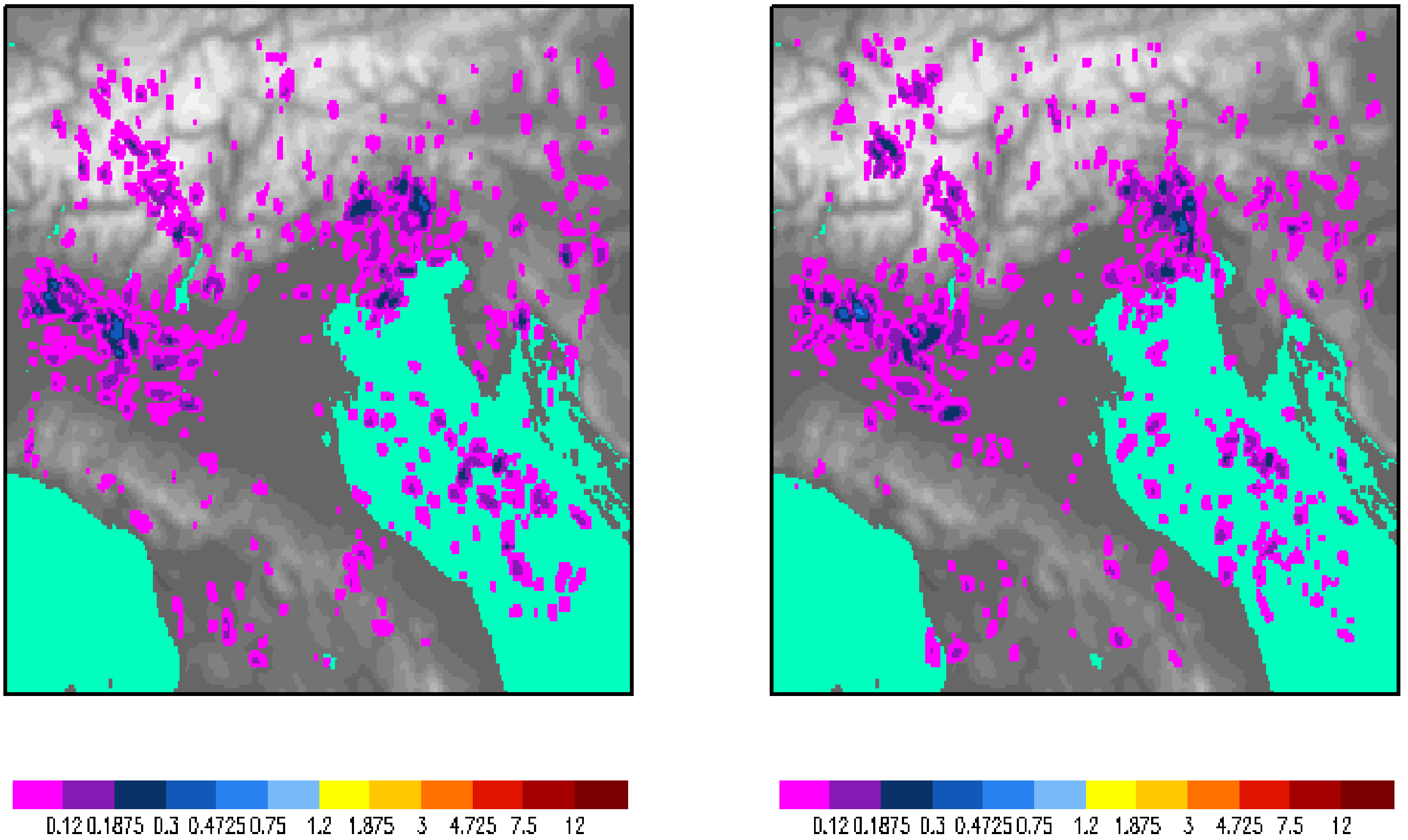
Two small, independent, randomly generated perturbations. Each variable scaled with its variability.
Breeding perturbations rescaled every 5 minutes, so that RMS of level 5 (~925hPa over sea) horizontal velocity is 0.05 m s^{-1}.
Instabilities related to fastest and smallest dynamical scales.



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

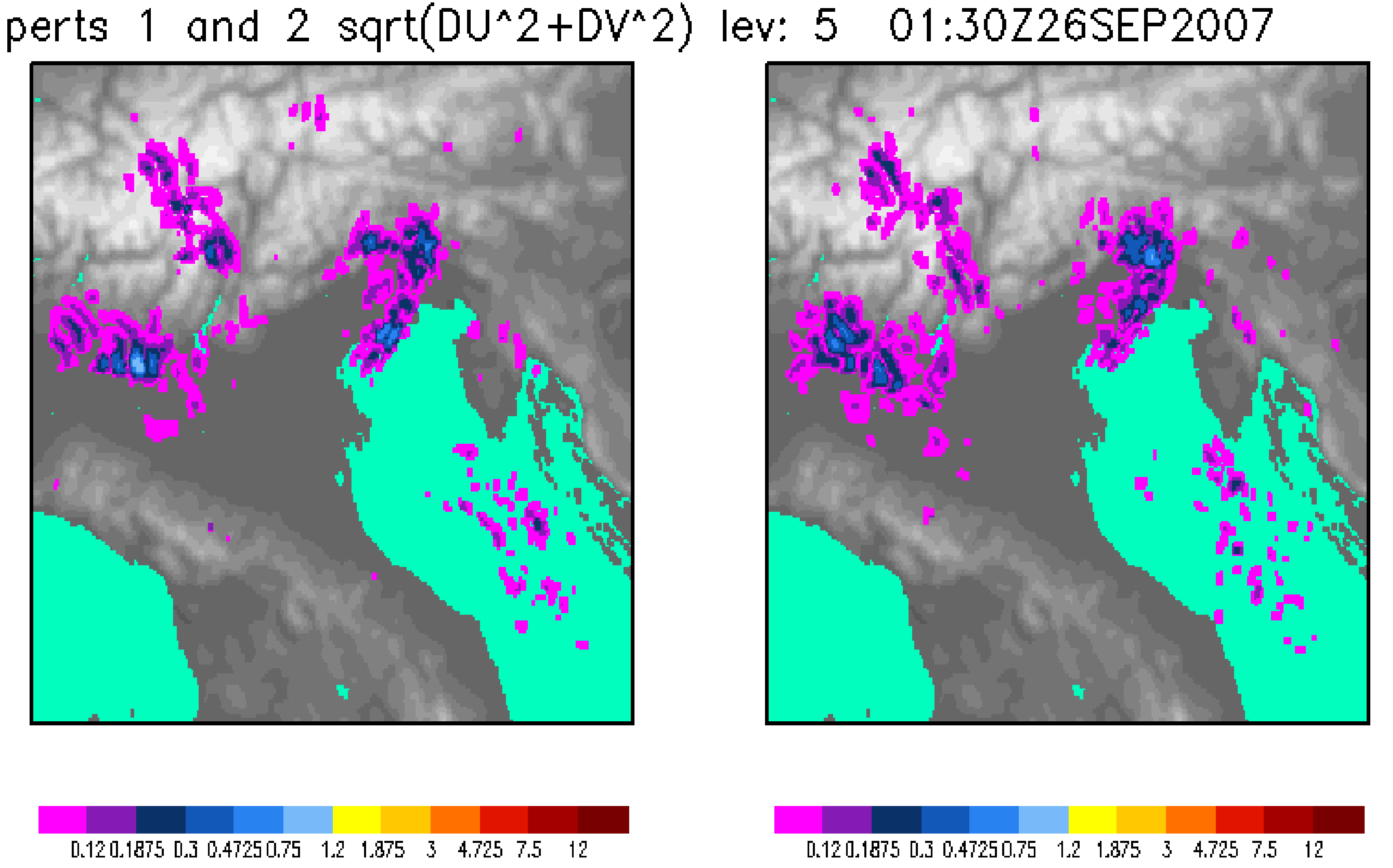

perts 1 and 2 sqrt(DU^2+DV^2) lev: 5 00:30Z26SEP2007
0.12 0.1875 0.3 0.4725 0.75 1.2 1.875 3 4.725 7.5 12
0.12 0.1875 0.3 0.4725 0.75 1.2 1.875 3 4.725 7.5 12

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



3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

Doubling time $\boldsymbol{T_D}$ estimated during the linear regime of growth for the most unstable modes:

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

Linearity time $\boldsymbol{T_{lin}}$ = duration of linear regime of growth
The linear regime is estimated to end when anti-correlation is lost between to two perturbations initially having the same direction and opposite orientation, each undergoing non-linear free evolution

When T_{lin} is shorter than T_D, errors may reach saturation so quickly that one can hardly see a linear growth phase — **after non-linear saturation, a reduction of the initial error cannot guarantee a reduction of forecast error**

(Hohenegger and Schär, 2007)

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

SMALL AND FAST SCALE INSTABILITIES:

Rescaling frequency: every 5 min, rescaling amplitude 0.05 m s^{-1} (UV level 5)

Scattered convection: $T_{lin} \simeq T_D \simeq 2.5\,h$

Organized, mesoscale front driven convection: $T_{lin} > 2\,T_D \simeq 4.0\,h$

The second convection episode, though more intense, is in fact more predictable, because of the leading role of the mesoscale (i. e. larger scale) forcing.

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

SMALL AND FAST SCALE INSTABILITIES:

Rescaling frequency: every 5 min, rescaling amplitude 0.05 m s^{-1} (UV level 5)

Scattered convection: $T_{lin} \simeq T_D \simeq 2.5\,h$

Organized, mesoscale front driven convection: $T_{lin} > 2\,T_D \simeq 4.0\,h$

The second convection episode, though more intense, is in fact more predictable, because of the leading role of the mesoscale (i. e. larger scale) forcing.

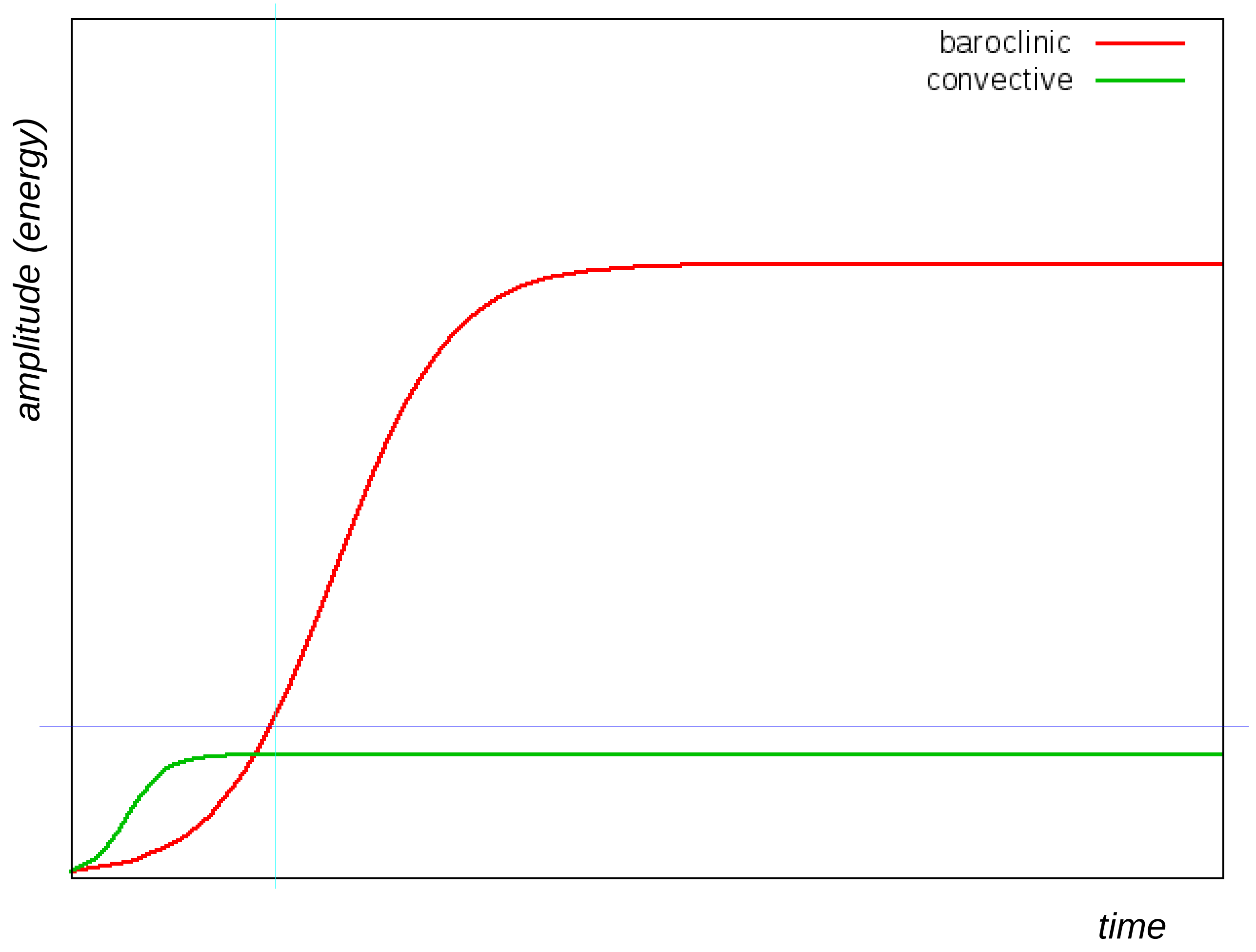
Much smaller than present-day analysis error!

- With a **larger rescaling amplitude** and appropriately lower rescaling frequency, such **fast instabilities saturate**.
- Saturated small-scale instabilities are not responsible for further error growth, but they affect (non-linear interaction) larger scales (which are in their linear regime): these are responsible for error growth at a level comparable with analysis error.

Two experiments with larger amplitudes:
“LARGE” ~ analysis error ← initial state from the external hydrostatic model ~2 m s^{-1} (UV level 8 ~850 hPa)

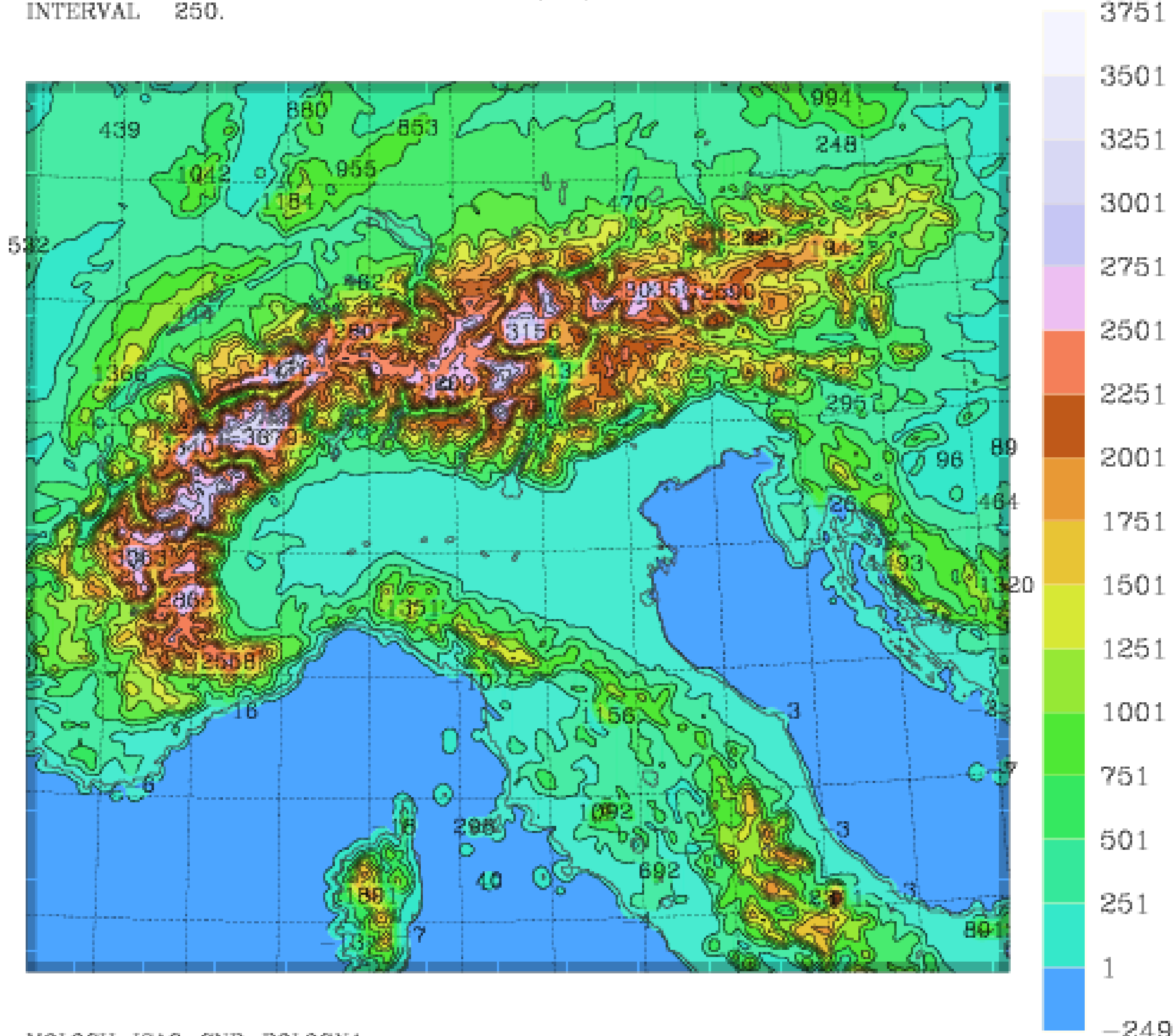
“SMALL” ~0.1 of “LARGE”, but still about 1 order of magnitude larger than the smallest (above)

Non-linear Breeding
filter small and fast scales

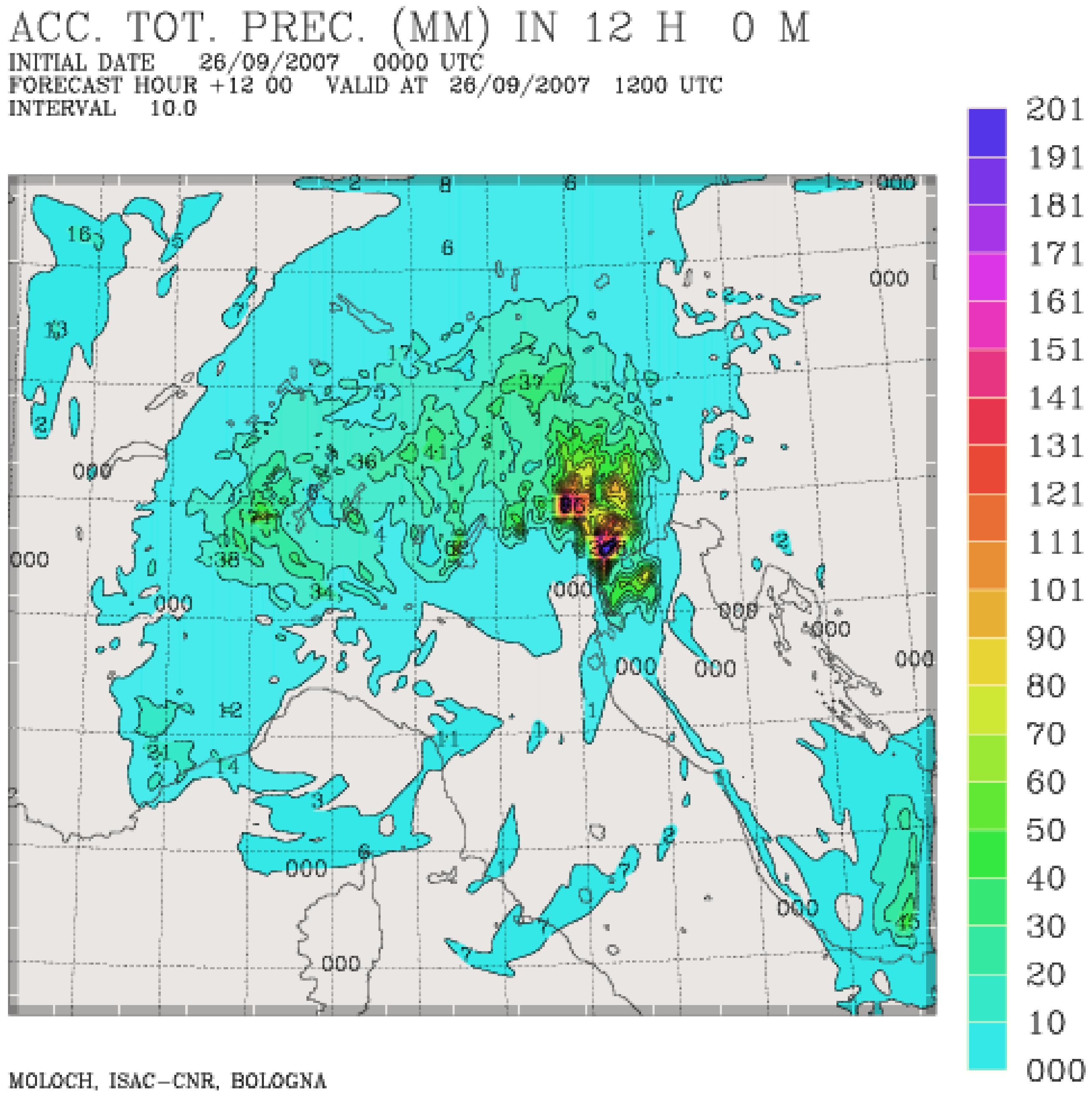


MODEL OROGRAPHY

INITIAL DATE 25/09/2007 0000 UTC
FORECAST HH MM +24 00 VALID AT 26/09/2007 0000 UTC
INTERVAL 250.



MOLOCH, ISAC−CNR, BOLOGNA



Accumulated precipitation from 0000 to 1200 UTC in the control trajectory

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

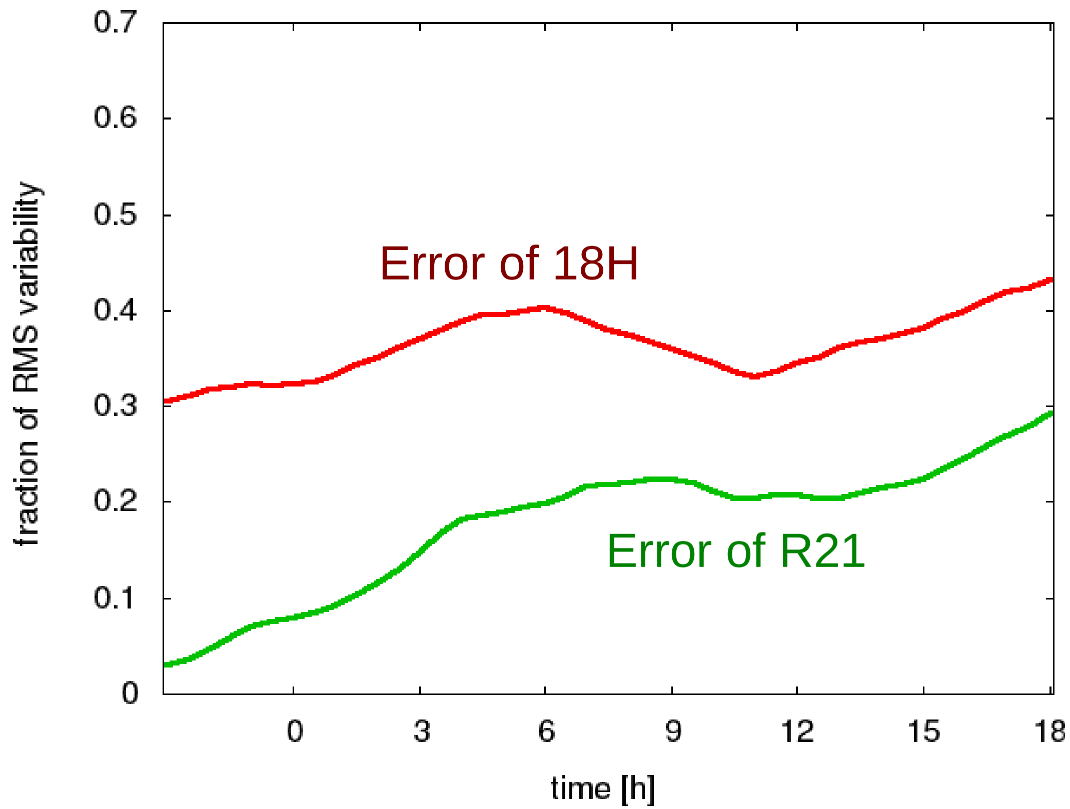
True trajectory: model trajectory from 21h of 25 Sep 2006 to 18h of 26 Sep 2006
Initial condition from external model

Control trajectory “**18H**” for **LARGE BV amplitude, SLOW instabilities**:
initial condition from external model at 18h of 25 Sep 2006

Control trajectory “**R21**” for **SMALL BV amplitude, 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

error = control - truth

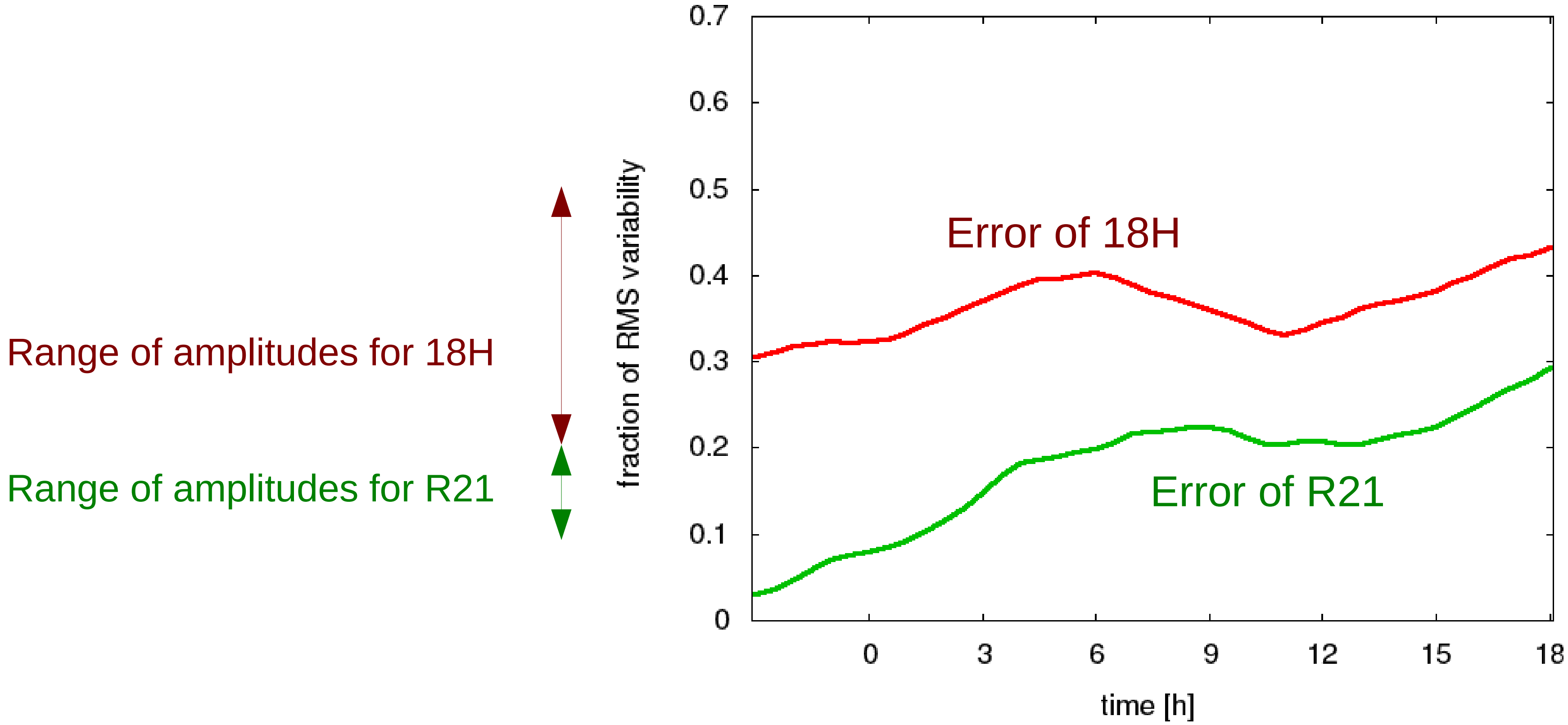


3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

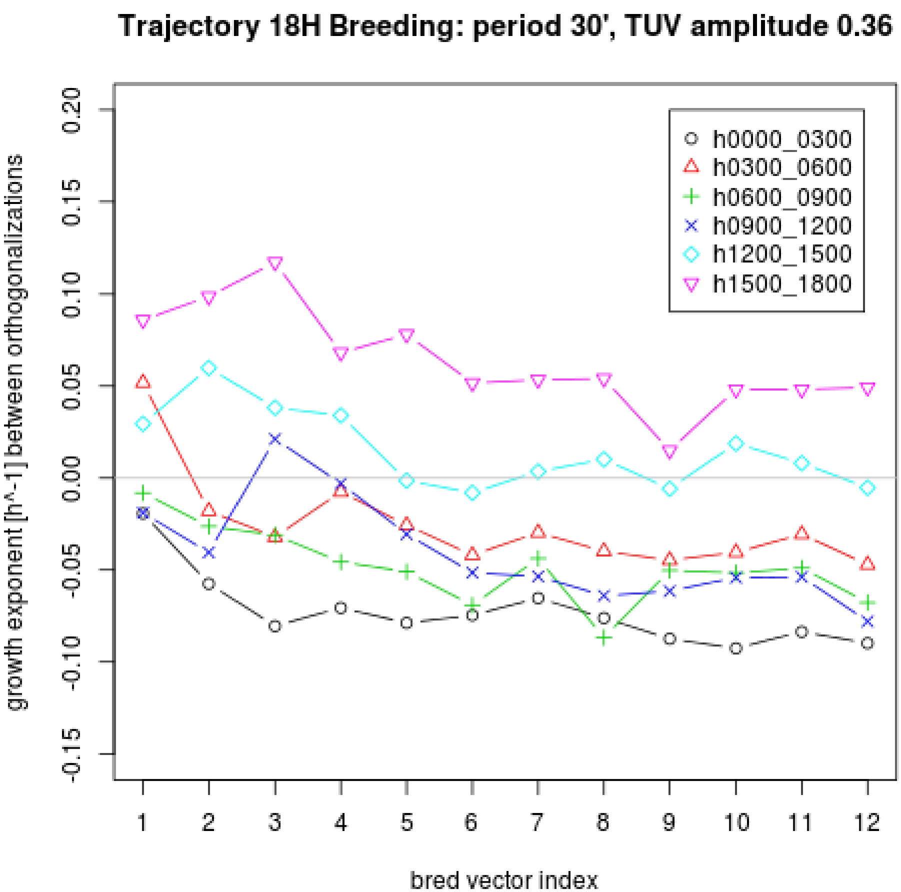
Orthogonalization required, scalar product needed:
sum of component products T and U,V normalized with their variabilities

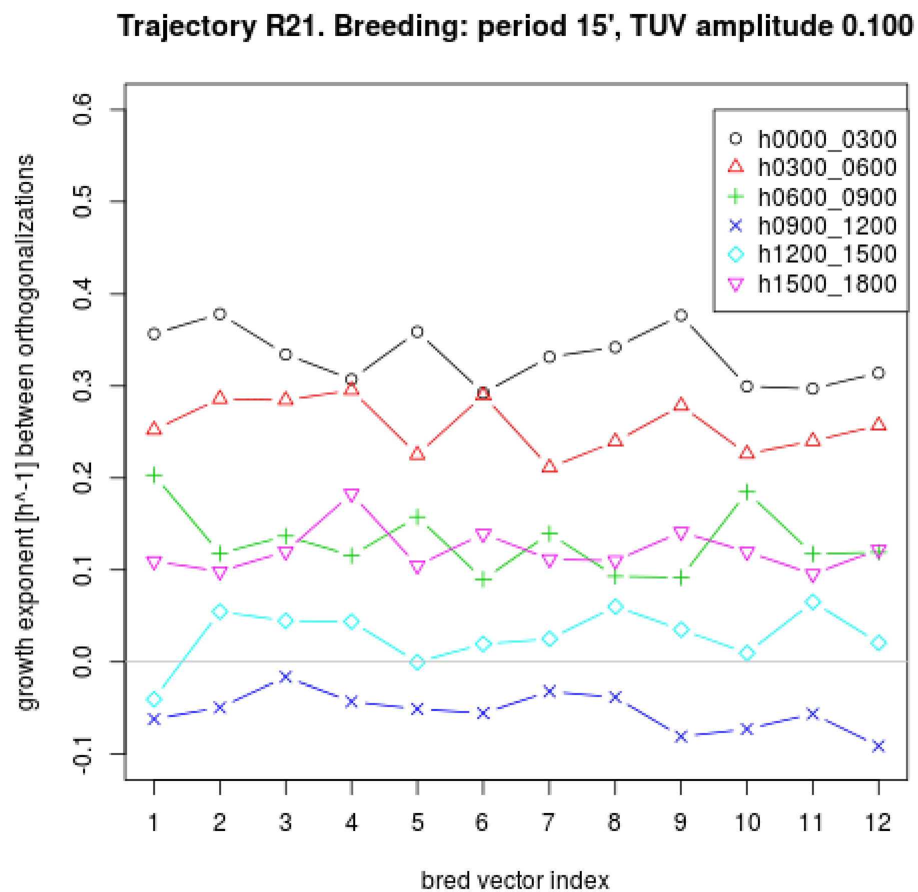
Breeding
LARGE – 18H : rescaling every 30 min (several amplitude values)

SMALL – R21 : rescaling every 15 min (several amplitude values)



3. Multiple scale instabilities in non-hydrostatic, convection resolving systems





BOTH: time variability (different curves) reflects time variability of forecast error:
larger BV growth rates when respective forecast error increases
Number of positive growth rates → error complexity (independent growth directions)

18H: growth rate decreases with BV index
All positive only 15:00-18:00 (slower scales dominant), few positive otherwise.

R21: growth rate does not decrease with BV index
All positive always except 09:00-12:00

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

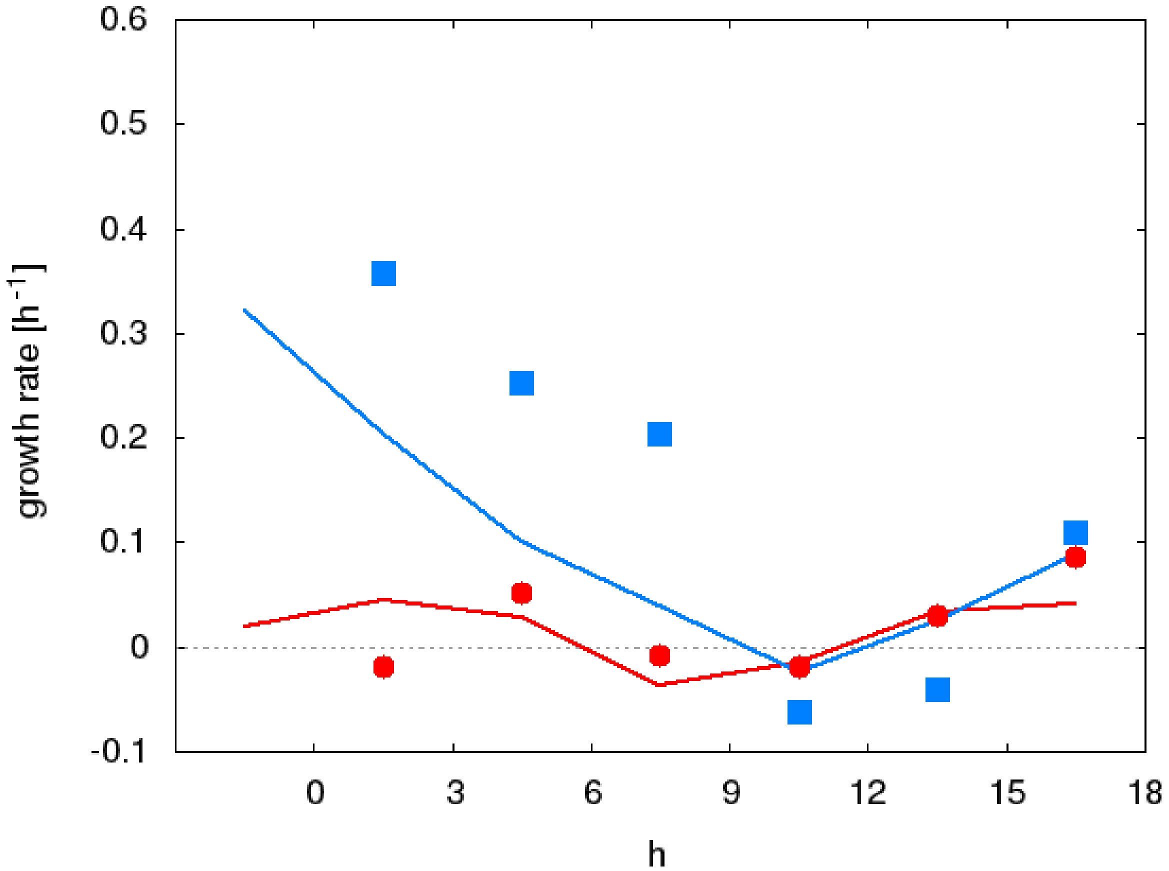
$T_D \simeq 2\,\mathrm{h} \Leftrightarrow \lambda \simeq 0.35$

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

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

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

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

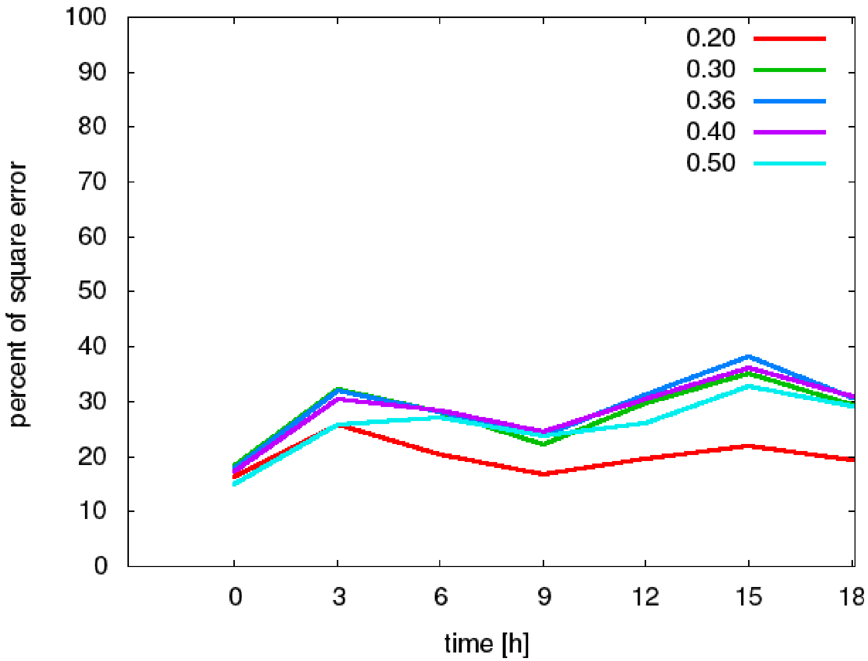


Growth rates of forecast error (non-linear) trajectories **18H** and **R21** (lines)
and their first BV (kept linear by rescaling at 0.36TUV 30min, 0.10TUV 15min)

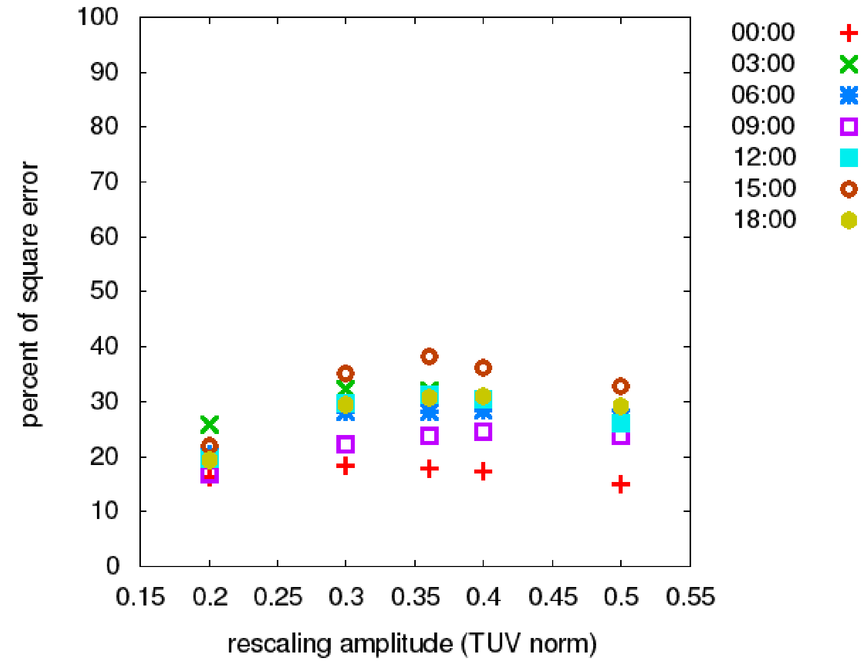
Correspondence: BVs contain the same instabilities as the forecast error

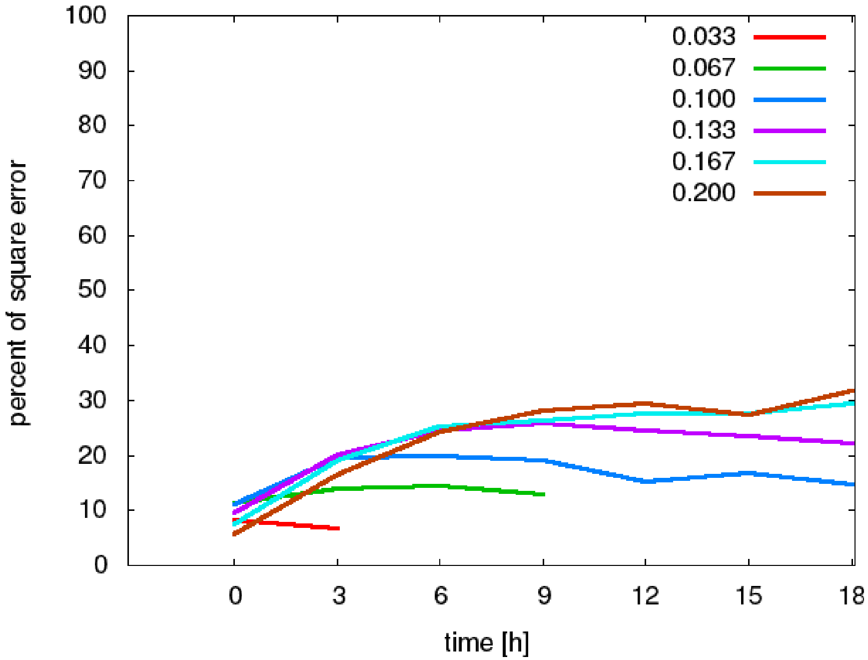
3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

Square norm of error orthogonal projection onto 12-BV subspace

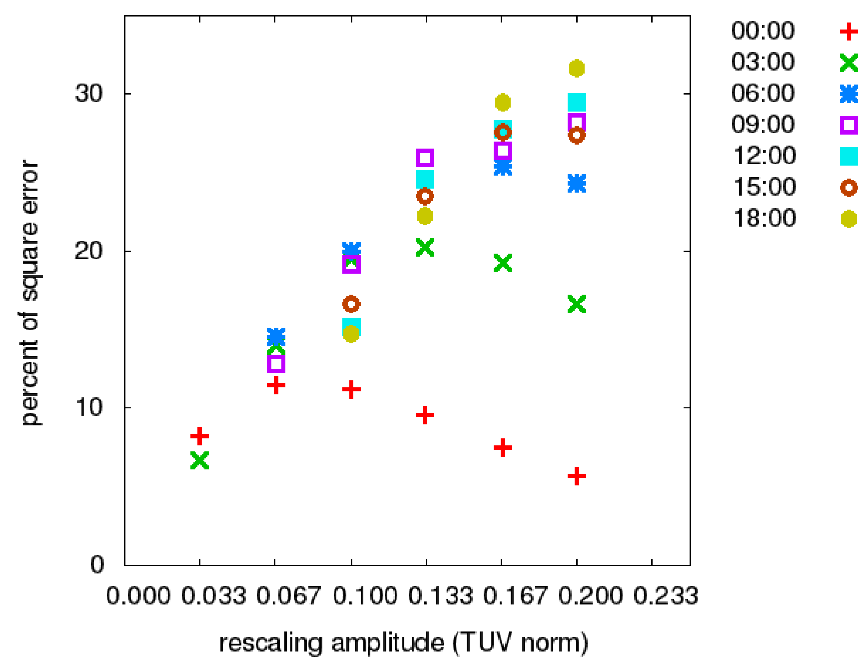


18H





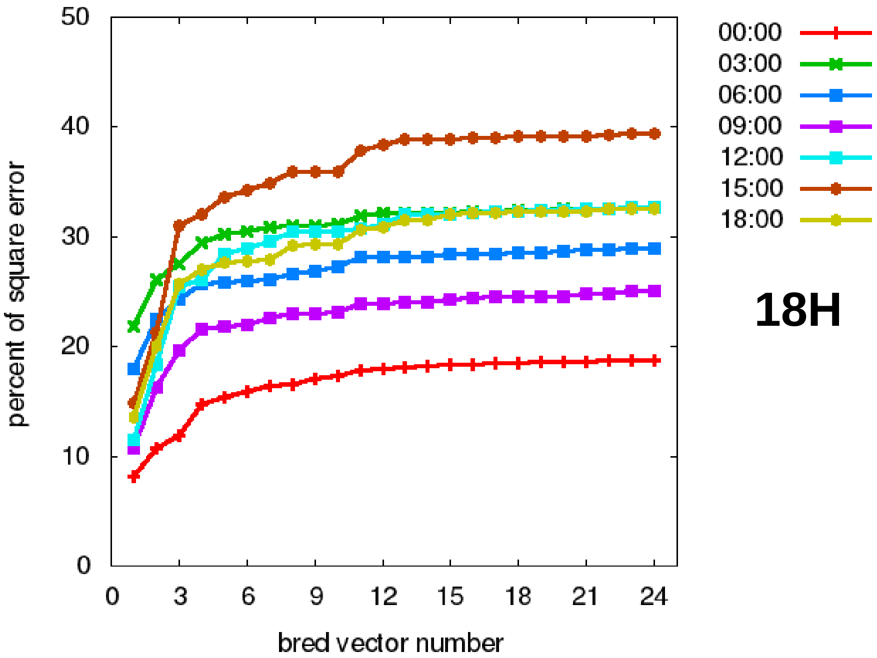
R21



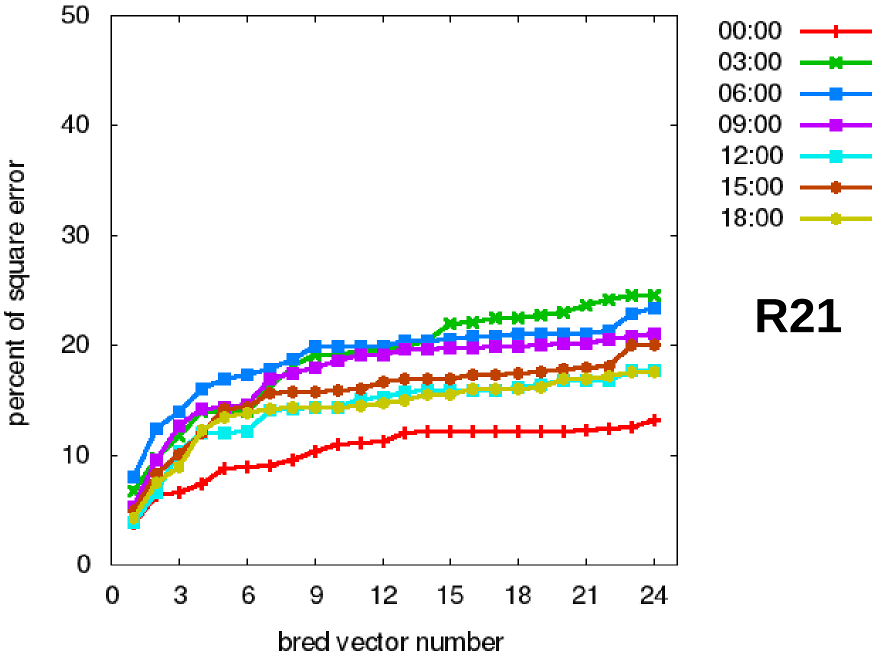
3. Multiple scale instabilities in non-hydrostatic, convection resolving systems

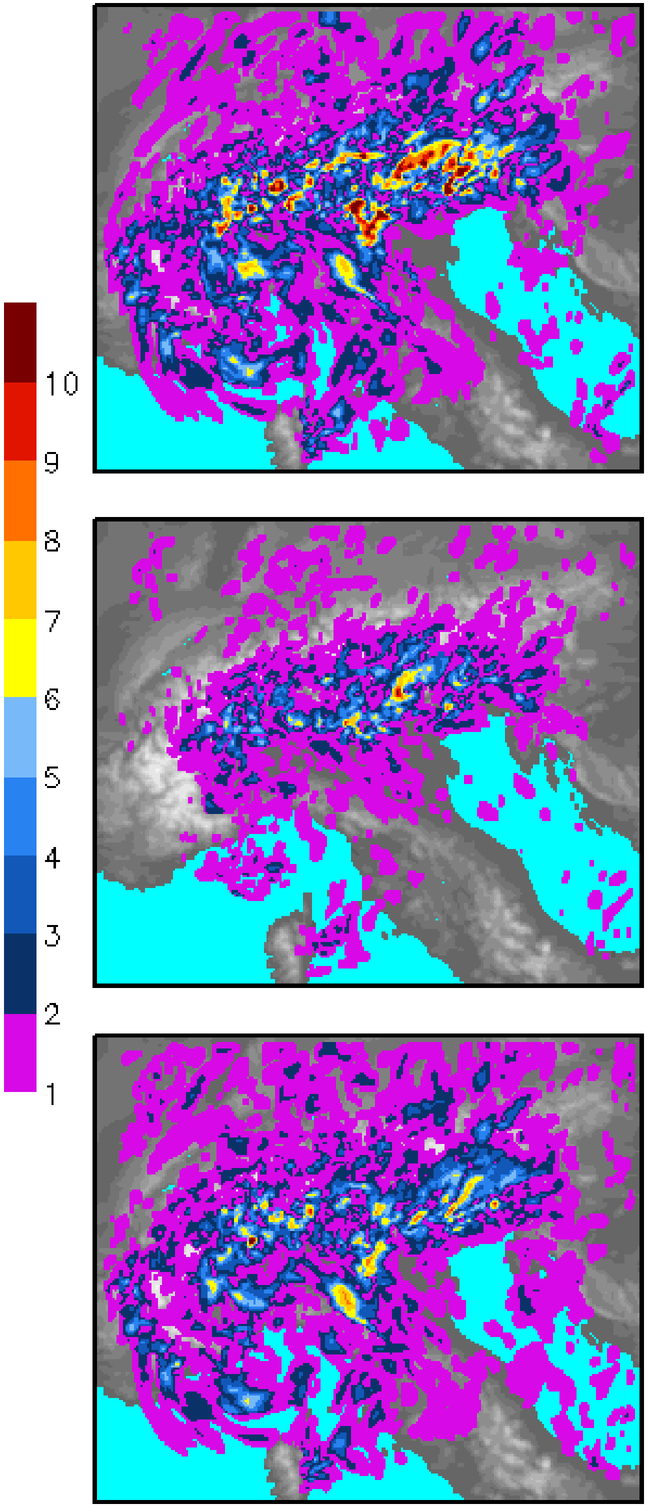
Square norm of error orthogonal projection onto BV subspaces: 1 to 24BVs

18H:
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:
Slow regular increase:
Many Bvs (more than 24) determine each a small amount of error fraction
Many independent instabilities





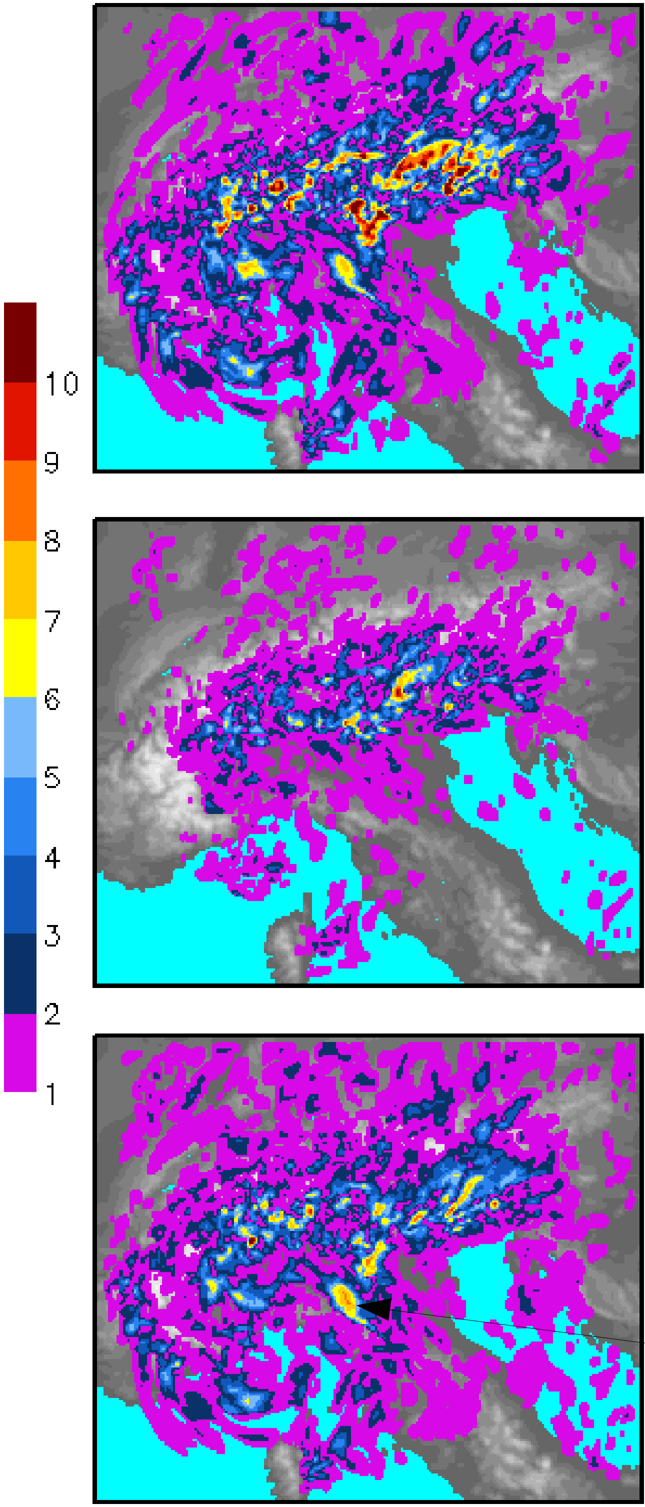
Forecast error
(norm of horizontal velocity vector at level 15 ~700hPa)

Forecast error component on 12-BV subspace

- Contains important features present in the forecast error
- Not all of them

Forecast error component orthogonal to the 12-BV subspace

- Some important structures disappear
- Large (amplitude and extension) structures still present
- Some are non-growing (from hydrostatic initial condition)

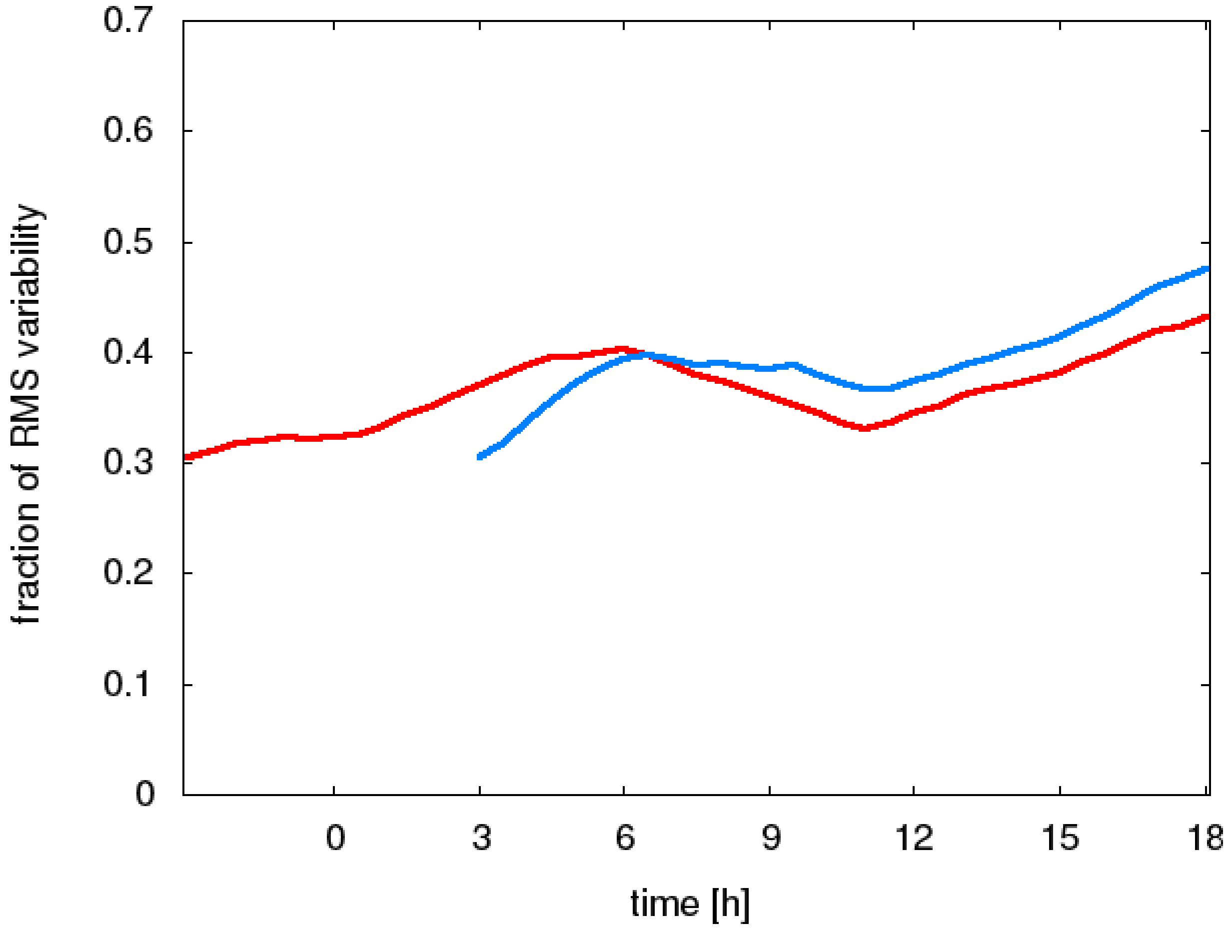


Forecast error
(norm of horizontal velocity vector at level 15 ~700hPa)

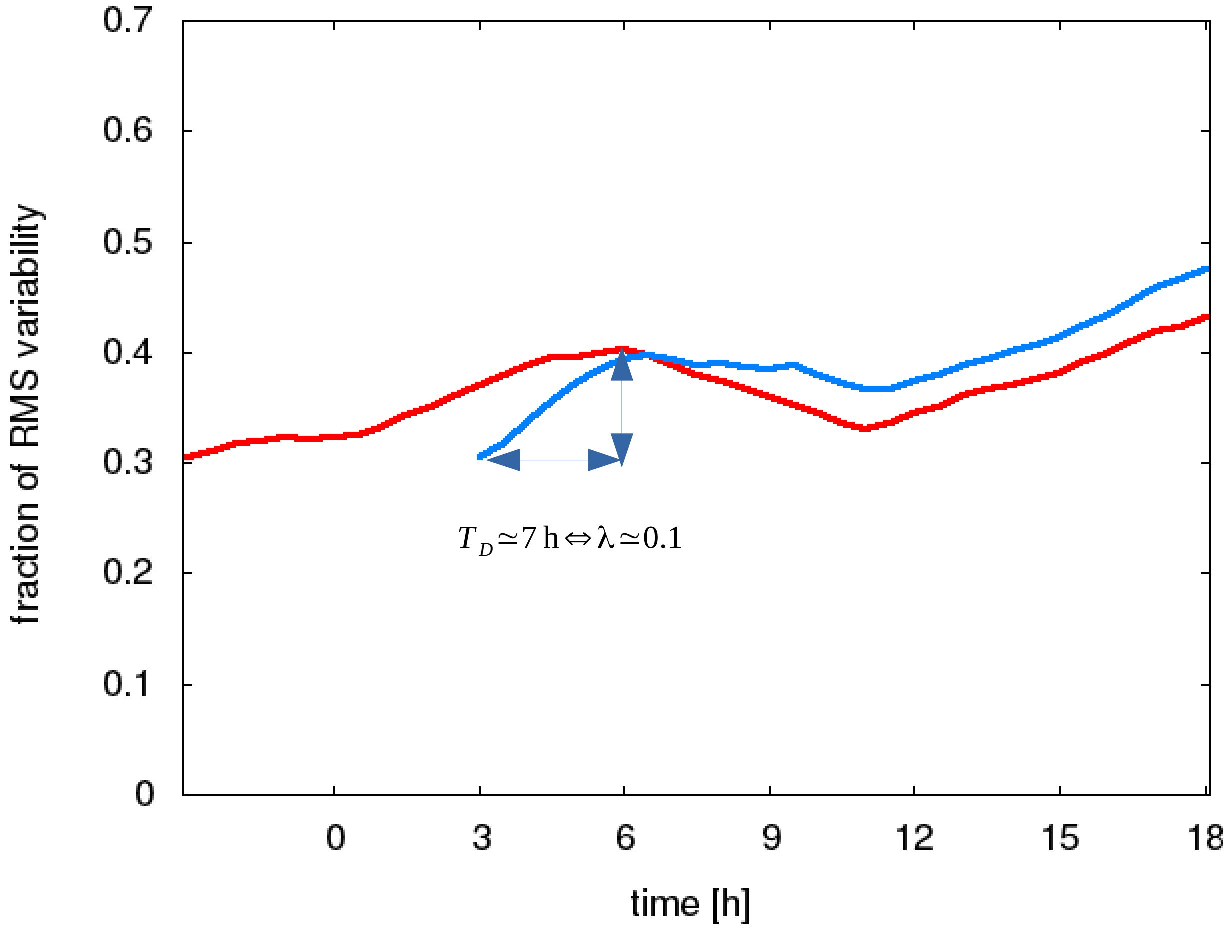
Forecast error component on 12-BV subspace
- Contains important features present in the forecast error
- Not all of them

Forecast error component orthogonal to the 12-BV subspace
- Some important structures disappear
- Large (amplitude and extension) structures still present
- Some are non-growing (from hydrostatic initial condition)

3. Multiple scale instabilities in non-hydrostatic, convection resolving systems



3. Multiple scale instabilities in non-hydrostatic, convection resolving systems



CONCLUSIONS 1/3

- Quantitative study of instabilities representative of forecast error evolution at different (time and space) scales
- 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**
- Non-growing error components:
  - Saturated small scale, fast instabilities
  - larger scale error structures present in an initial condition coming from a larger-scale hydrostatic model

- The **breeding** technique **enables selection of instabilities** relevant for forecast errors of a given typical amplitude:
  - BVs amplitude of **about the order of the analysis error** – **slow** instabilities
  - BVs amplitude about **1/10** of the order of the analysis error – **small, fast** instabilities

CONCLUSIONS 2/3

- **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, not much increasing 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.
    - Unstable subspace of convective scale has very large dimension

CONCLUSIONS 3/3

- When we **decrease** much the **analysis error** (by improving DA) we **activate convective scale instability**! Then we need more and more members in an ensemble forecast or ensemble-based assimilation
- At the level of **present-day analysis error**, though:
  - Instabilities **do not grow that fast** (errors at fast convective scale are saturated)
  - There are **not many independent growth directions**
  - **Relevant portion of forecast error** can be captured by BVs and eliminated by DA
  - **Rapid growth** is expected **after the analysis**
- What to do:
  - Frequent analysis (every hour or so)
  - **DO NOT restart from** external, larger-scale, **hydrostatic** model, **initial condition** – use it for boundary conditions only: make use of DA to control the trajectory
  - Localization techniques ?!?
  - ...Multiple scale breeding /ensemble ?!?