The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

SMR/1849-1

Conference and School on Predictability of Natural Disasters for our Planet in Danger. A System View; Theory, Models, Data Analysis

25 June - 6 July, 2007

Sources of uncertainty

Roberto Buizza
European Centre for Medium-Range Weather Forecasts, Reading, UK

Sources of uncertainty

Roberto Buizza

European Centre for Medium-Range Weather Forecasts
(http://www.ecmwf.int/staff/roberto_buizza/)

Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

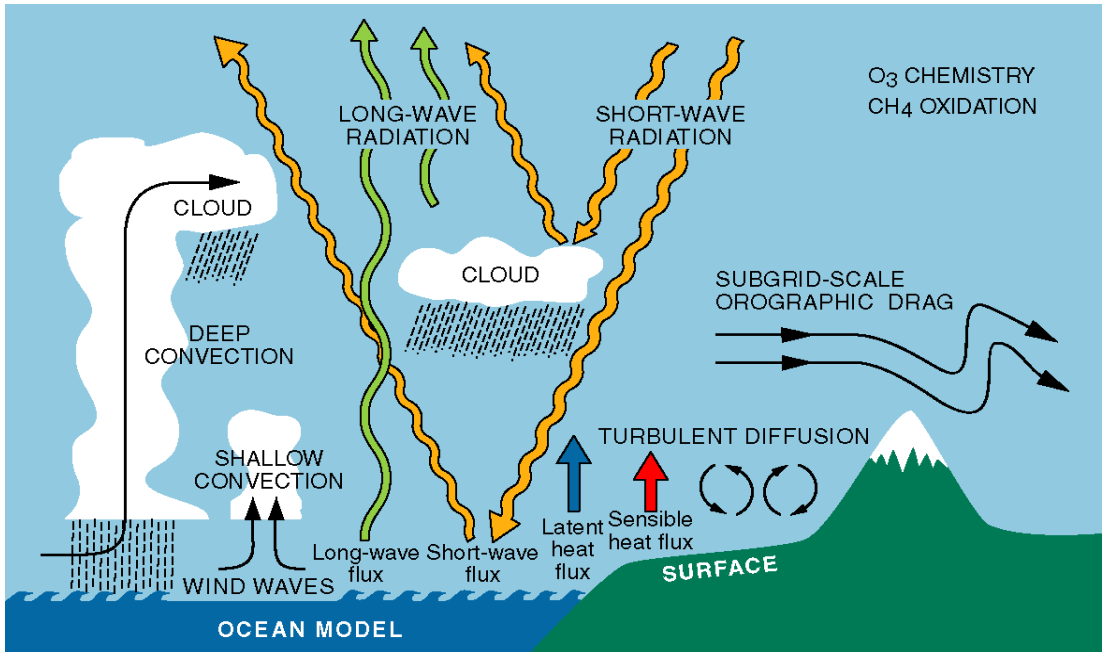
1. The ECMWF NWP Model

The behavior of the atmosphere is governed by a set of physical laws which express how the air moves, the process of heating and cooling, the role of moisture, and so on.

Interactions between the atmosphere and the underlying land and ocean are important in determining the weather.

ECMWF MODEL / ASSIMILATION SYSTEM

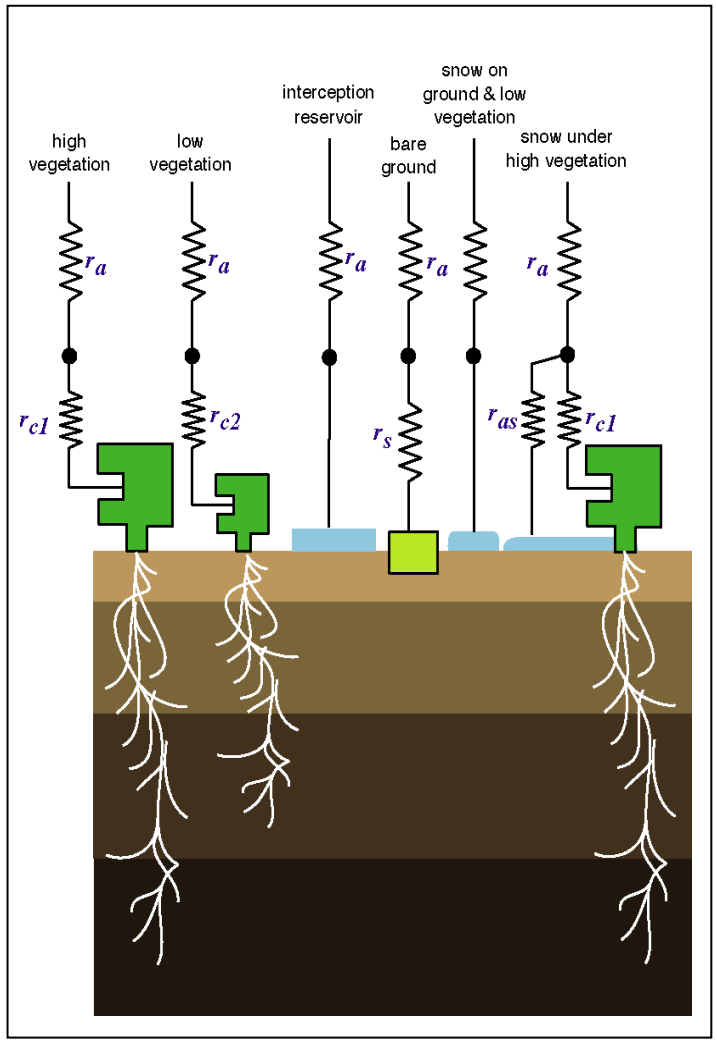
ATMOSPHERE	STRATOSPHERE	DYNAMICS–RADIATION–SIMPLIFIED CHEMISTRY		
	TROPOSPHERE	DYNAMICS–RADIATION–CLOUDS–ENERGY & WATER CYCLE		
OCEAN / LAND		OCEAN	LAND HYDROSPHERE	LAND BIOSPHERE
		OCEAN SURFACE WAVES OCEAN CIRCULATION SIMPLIFIED SEA ICE	SNOW ON LAND SOIL MOISTURE FREEZING	LAND SURFACE PROCESSES SOIL MOISTURE PROCESSES SIMPLIFIED VEGETATION



1. Surface processes

Surface processes determine the sources and sinks of temperature and moisture (in terms of sensible and latent heat fluxes) at the lowest boundary of the atmosphere. As a consequence, over land, they define the state of the ground (warm, cold, freezing, dry or wet) and whether falling rain or snow precipitation will remain or subsequently disappear.

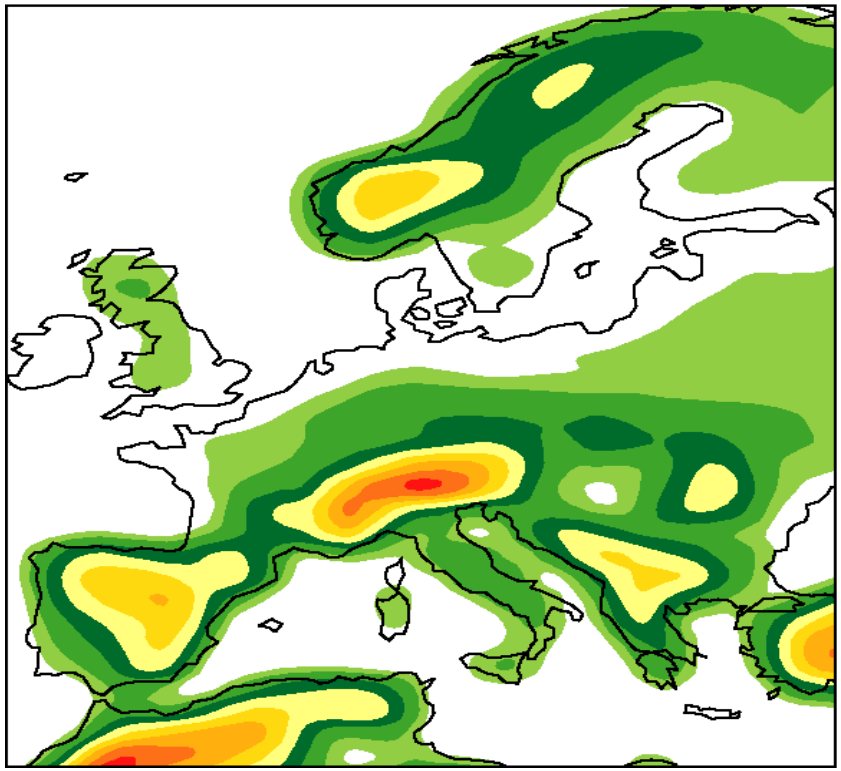
It is the surface characteristics, such as the 2-m temperature, humidity and wind, that are some of the most important (and most difficult) to predict, since, after all, the lowest 2 metres is exactly the part of the atmosphere in which we spend most of our lives!



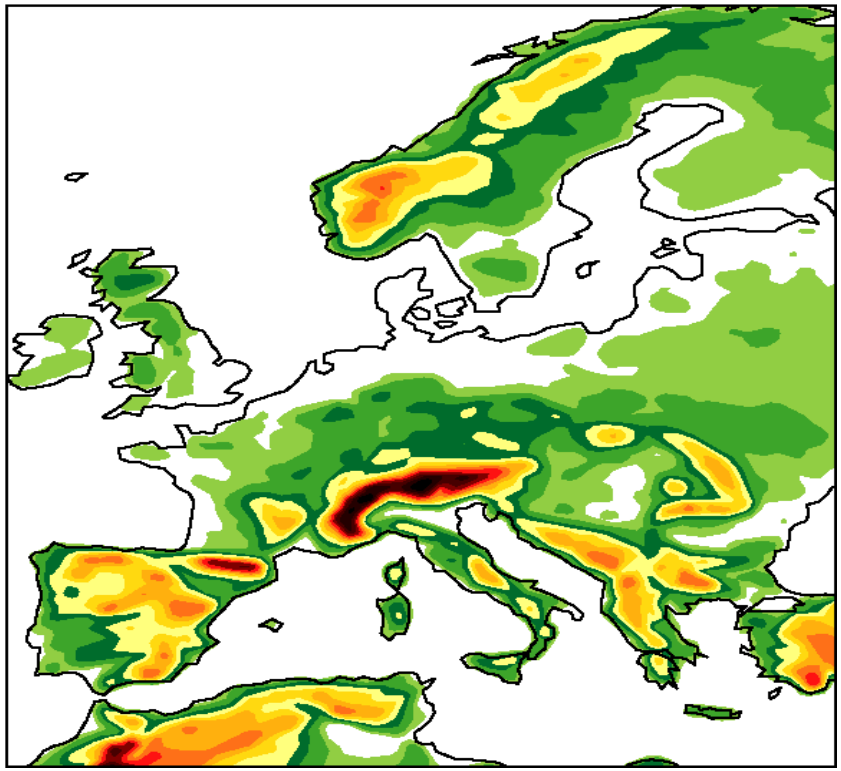
1. The importance of spatial resolution

A high spatial resolution is needed to achieve an accurate representation of the system physical processes. Similarly, the representation of the orography becomes more realistic with increased horizontal resolution.

Representation of the orography at T159 resolution



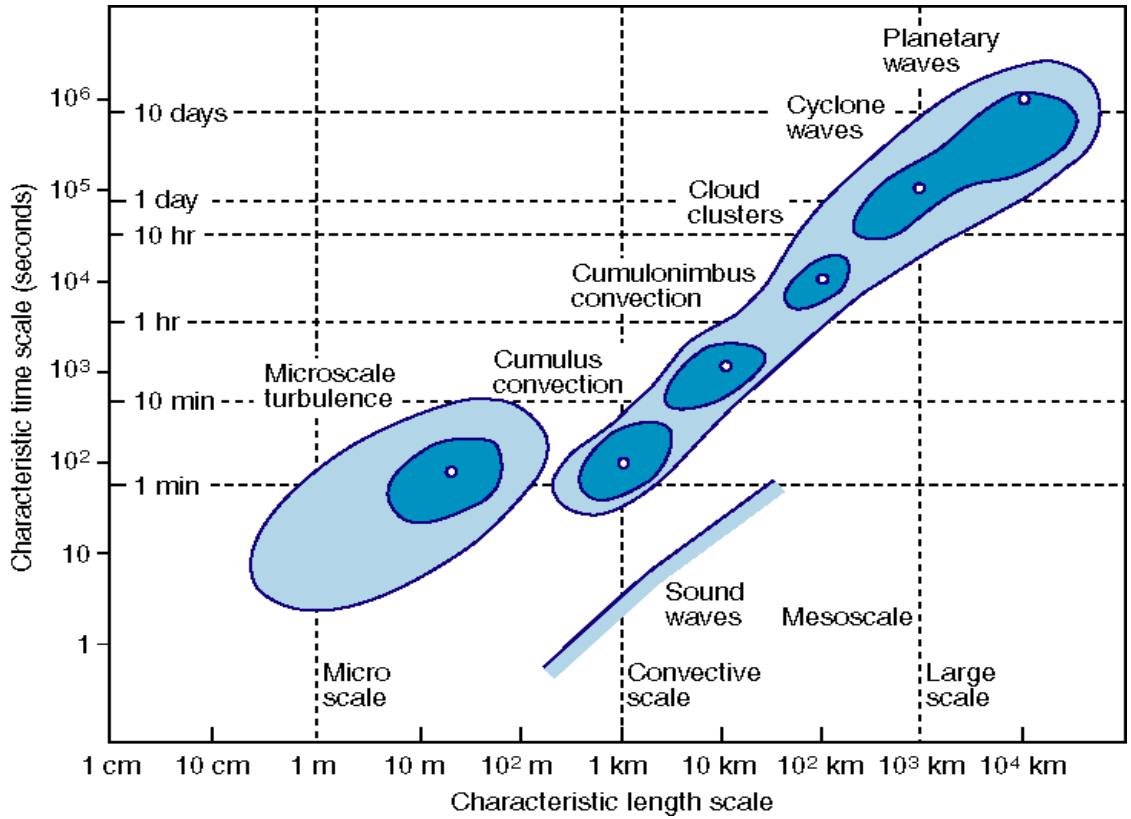
Representation of the orography at T639 resolution



1. Parameterization of physical processes

Each physical process has a characteristic spatial and temporal scale. Many processes occur on a spatial scale smaller than the model grid.

For example, over land a 40km square may include different types of vegetation, bare soil or buildings. Each type, for example, reflects the incoming solar radiation and affects moist processes in a different way.

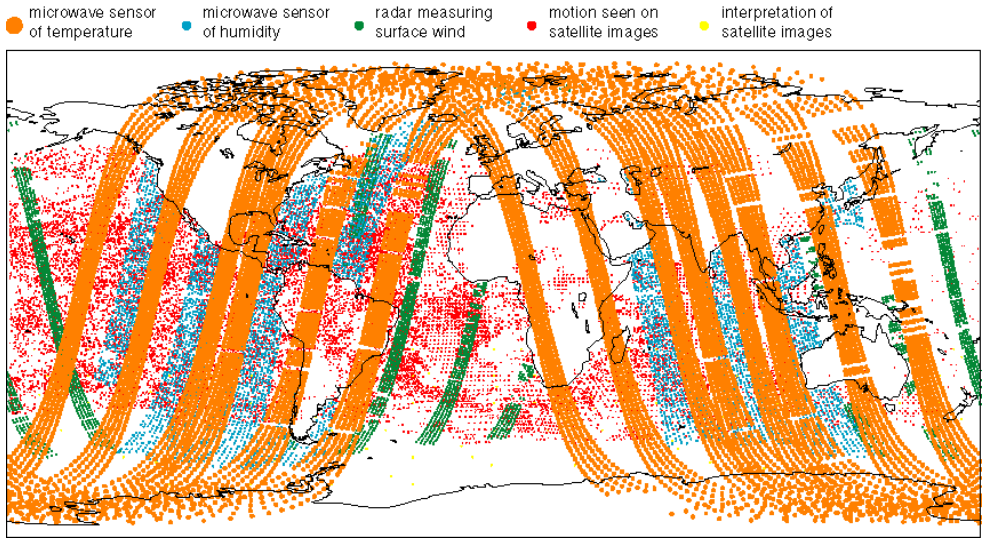


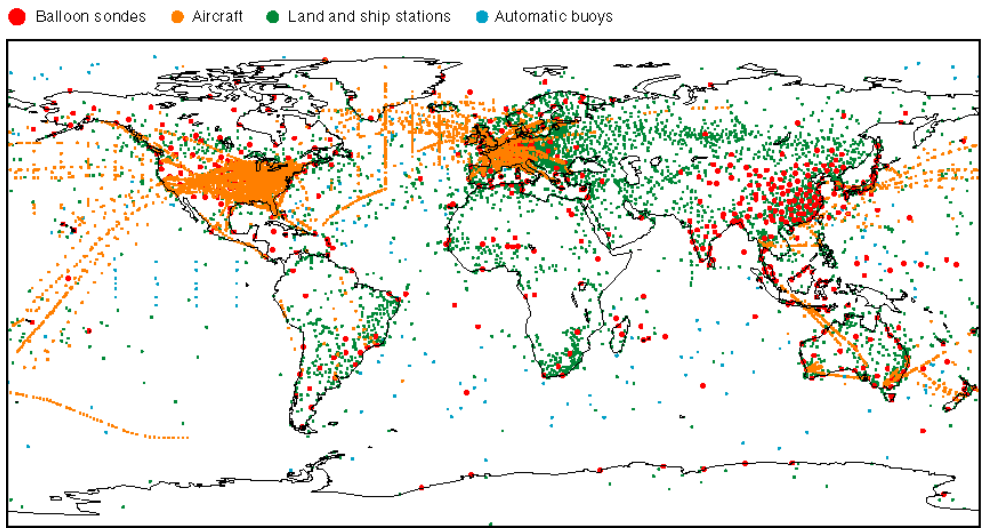
1. Starting a NWP: the initial conditions

To make accurate forecasts it is important to know the current weather:

- observations covering the whole globe are continuously downloaded and fed into the system;
- about 600,000 observations are processed every 12 hours;
- complex assimilation procedures are used to optimally define the initial state of the system.

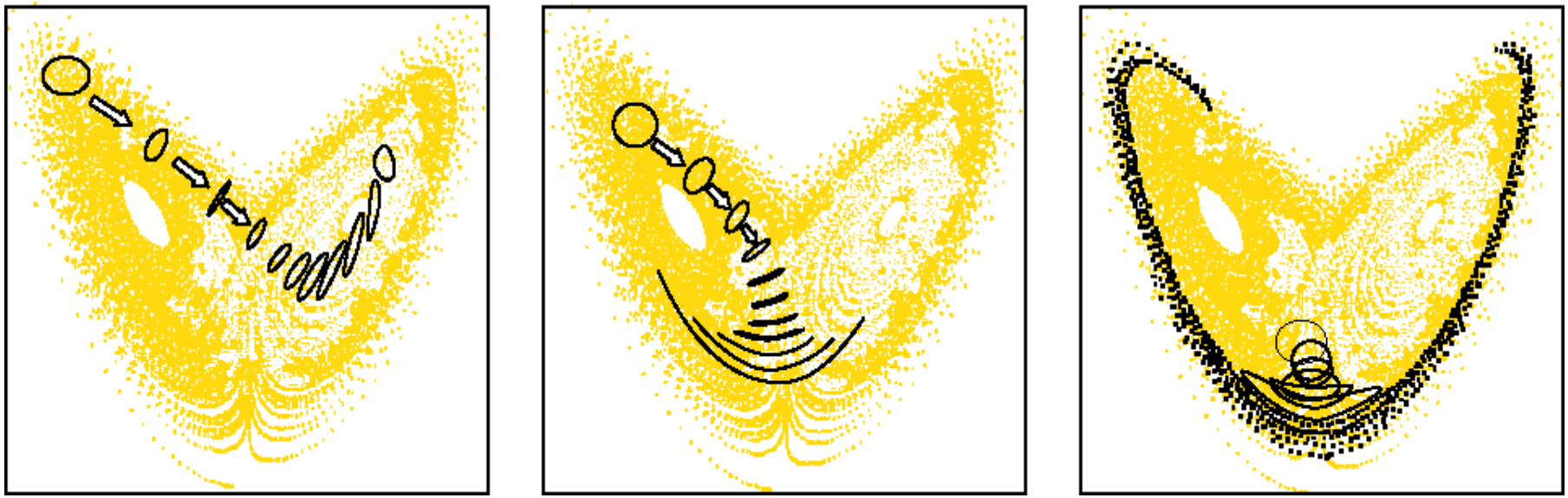
Unfortunately, very few observations are taken in some regions of the world (e.g. polar caps, oceans).





1. Sources of fc errors: initial and model uncertainties

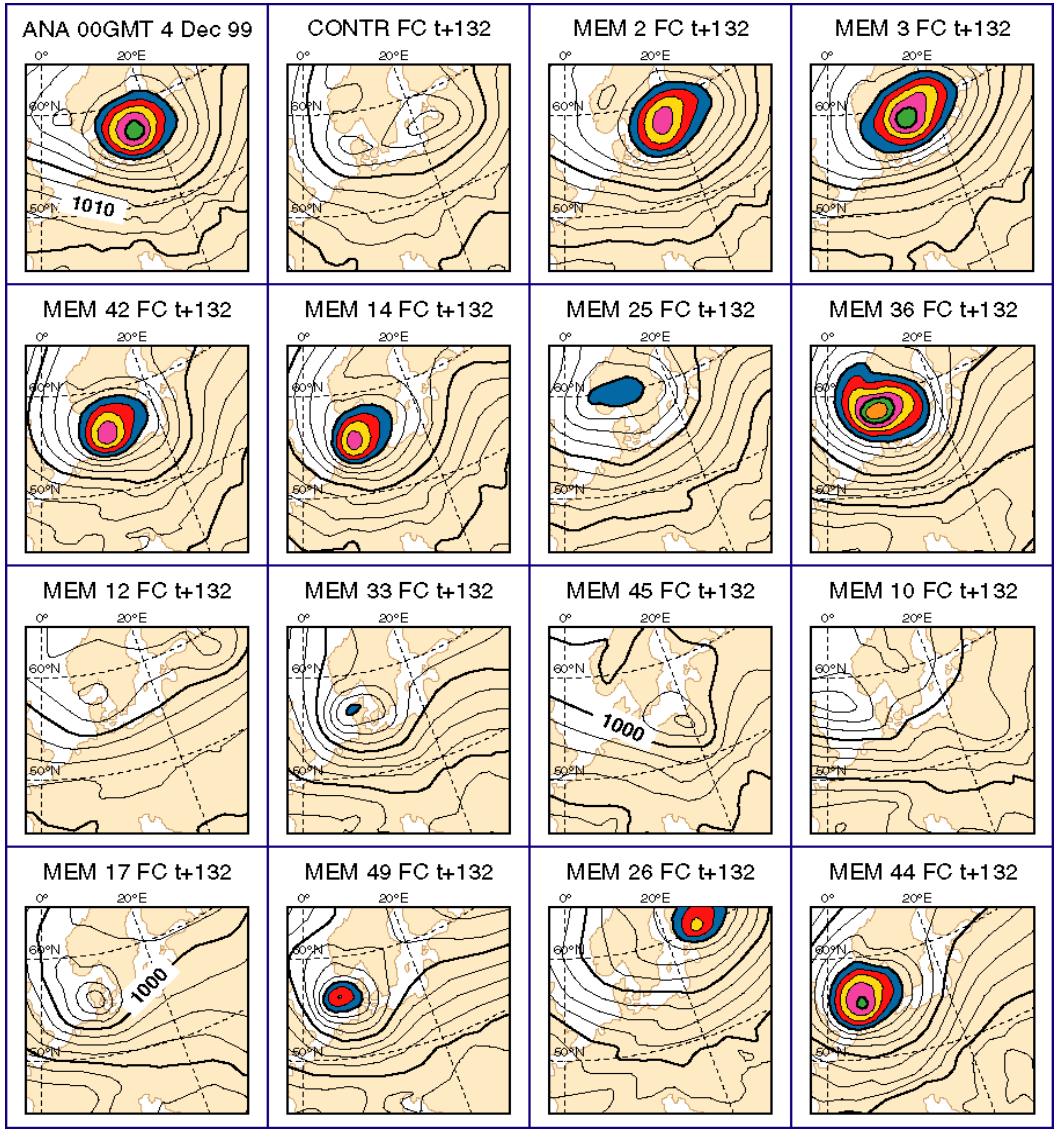
Weather forecasts lose skill because of the growth of errors in the initial conditions (initial uncertainties) and because numerical models describe the laws of physics only approximately (model uncertainties). As a further complication, predictability (i.e. error growth) is flow dependent. The Lorenz 3D chaos model illustrates this.



1. The atmosphere exhibits a chaotic behavior

A dynamical system shows a chaotic behavior if most orbits exhibit sensitivity to initial conditions, i.e. if most orbits that pass close to each other at some point do not remain close to it as time progresses.

This figure shows the verifying analysis (top-left) and 15 132-hour forecasts of mean-sea-level pressure started from slightly different initial conditions (i.e. from initially very close points).



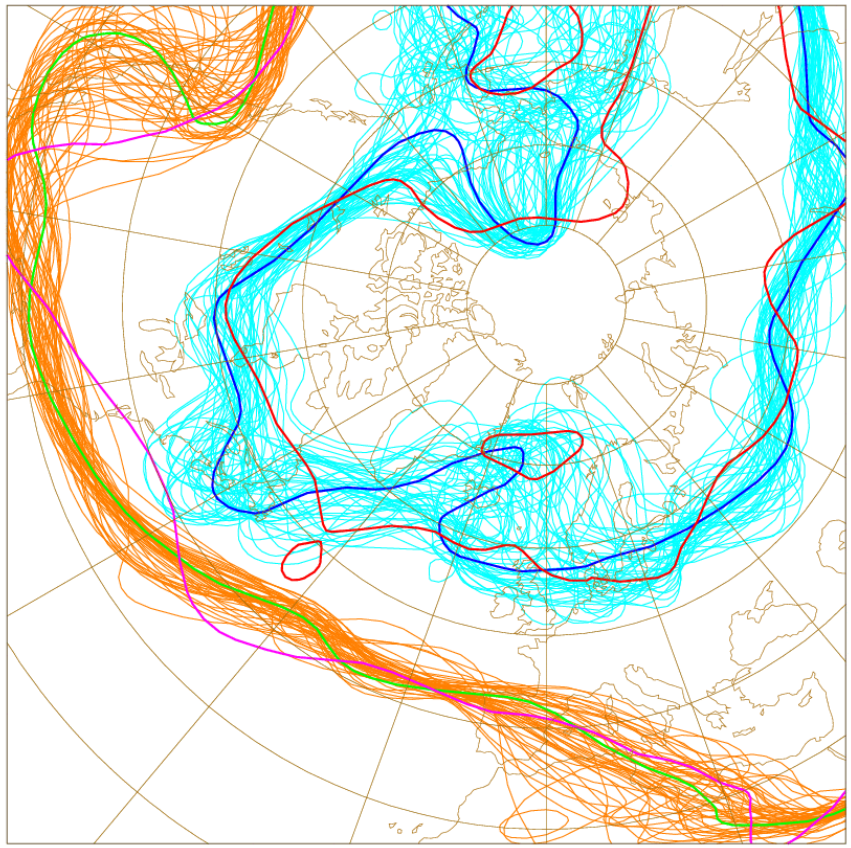
Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

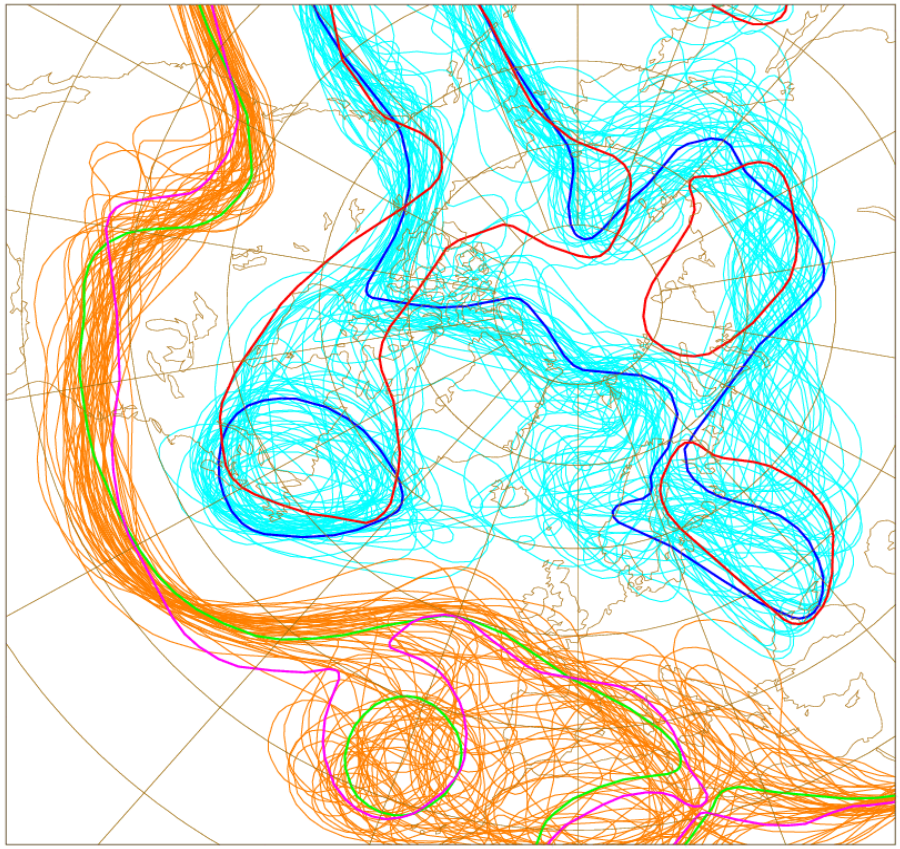
2. Predictability is flow dependent: spaghetti plots

The degree of mixing of Z500 isolines is an index of low/high perturbation growth.

500z d:1997-02-09 12:00:00 fc+120h cl:od exp:1
AN red/purple - CON blue/green - iso=5200-5700



500z d:1997-03-13 12:00:00 fc+120h cl:od exp:1
AN red/purple - CON blue/green - iso=5200-5700



Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

3. The probabilistic approach to NWP

A complete description of the weather prediction problem can be stated in terms of the time evolution of an appropriate probability density function (PDF). Ensemble prediction based on a finite number of deterministic integration appears to be the only feasible method to predict the PDF beyond the range of linear growth.

Currently, the ECMWF operational suite includes every day:

- a single deterministic forecast run at high resolution:
  - ***Day 0 to 10: T_L799L91, ~25km, 91 levels***
- 51 15-day forecasts run at lower resolution:
  - ***Day 0 to 10: T_L399L62, ~60km, 62 levels***
  - ***Day 10 to 15: T_L255L62, ~60km, 62 levels***

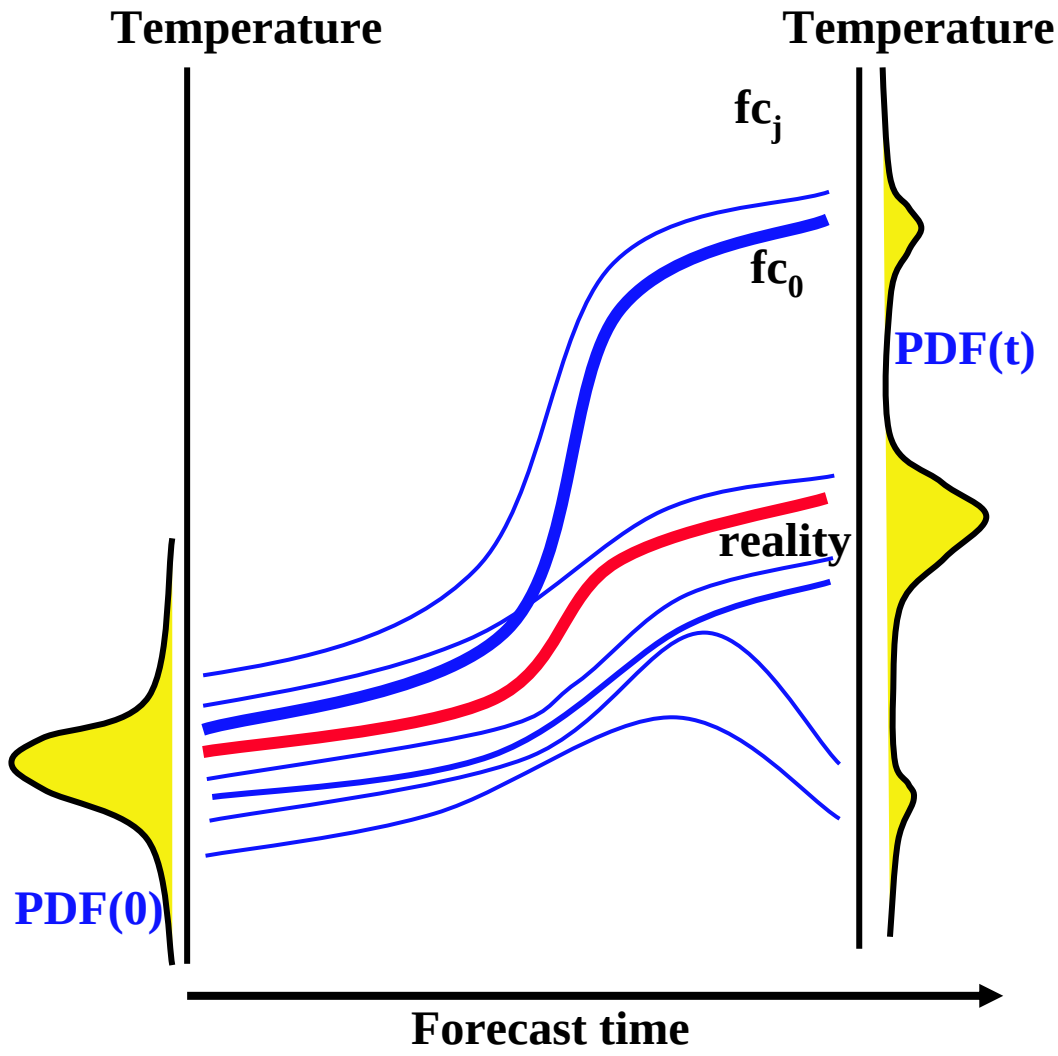
The 51 forecasts constitute the ECMWF Ensemble Prediction System. The first version of the EPS was implemented operationally in December 1992. The current version of the EPS simulates both initial and model uncertainties.

3. Schematic of ensemble prediction

Two are the main sources of error growth: initial and model uncertainties.

Predictability is flow dependent.

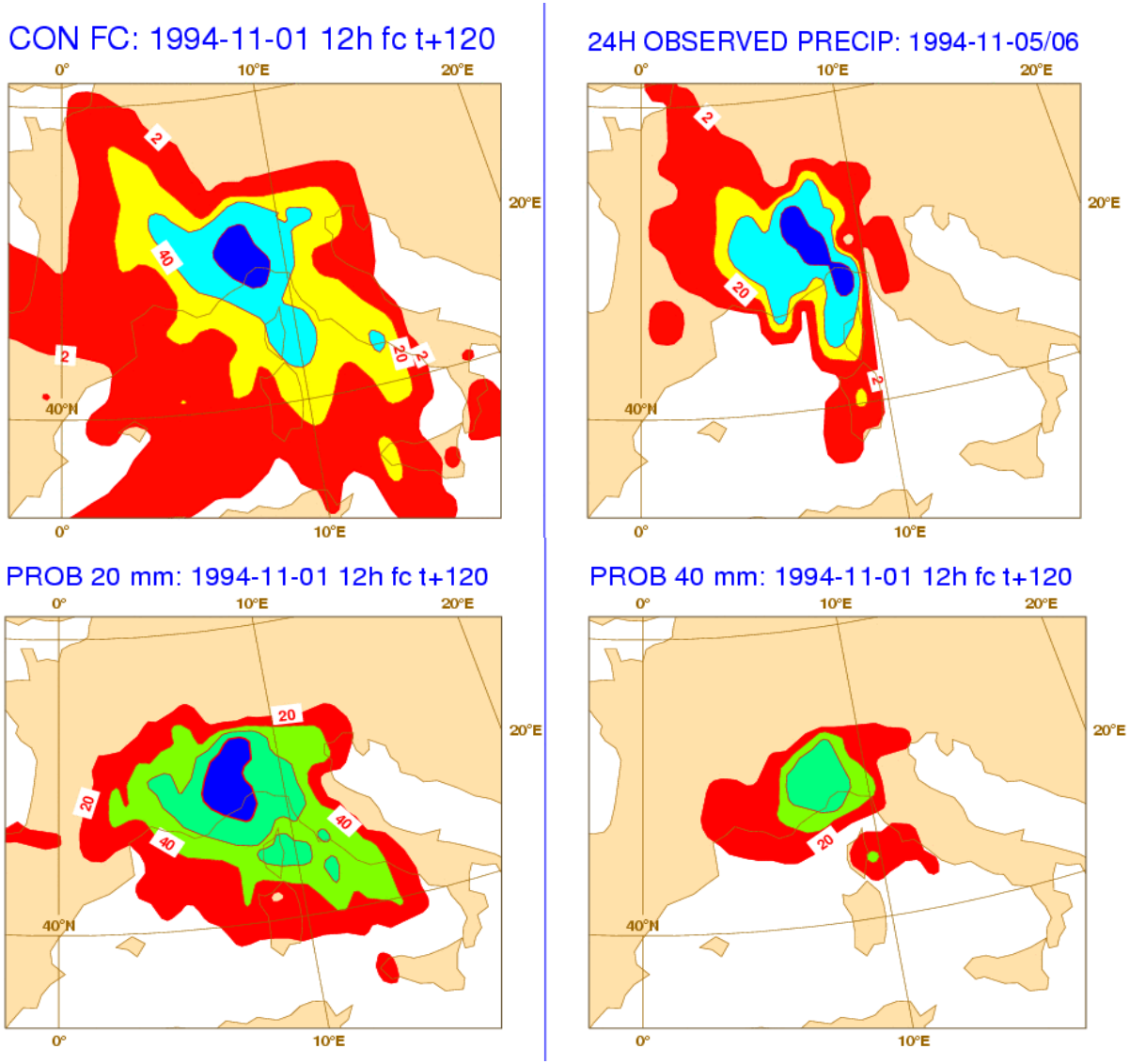
A complete description of weather prediction can be stated in terms of an appropriate probability density function (PDF). Ensemble prediction based on a finite number of deterministic integration appears to be the only feasible method to predict the PDF beyond the range of linear growth.



3. What does it mean to 'predict the PDF time evolution'?

The EPS can be used to estimate the probability of occurrence of any weather event.

Floods over Piemonte (Italy), 6 Nov 94 (top right panel). The forecast skill of the single deterministic forecast given by the EPS control (top left) can be assessed by EPS probability forecasts (bottom panels).

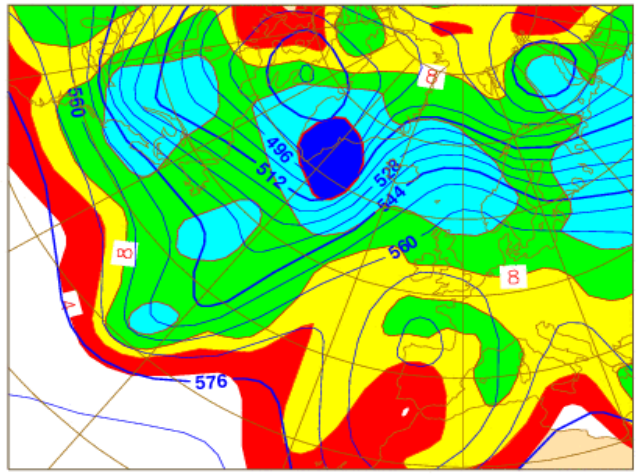


3. What does it mean to 'predict the PDF time evolution'?

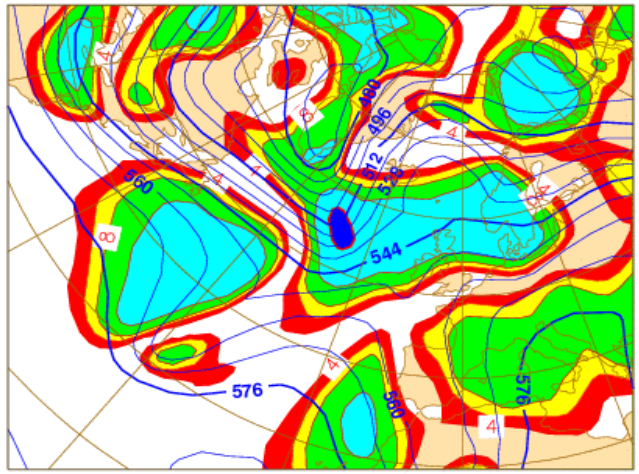
The ensemble spread around the control forecast can be used to identify areas of potential large control-forecast error.

These figures show the 5-day control forecast and ensemble spread (left) and the verifying analysis and the control error (right) for forecasts started 18 January 1997 (top) and 1998 (bottom).

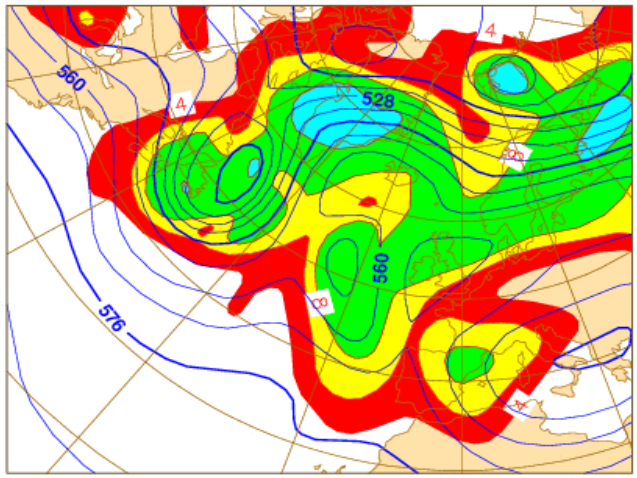
CON+SP - Z500 1997-01-18 12h fc t+120



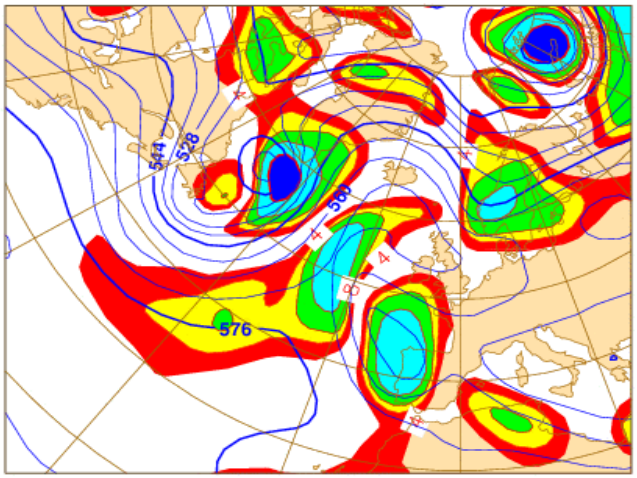
ANA+ERRCON - Z500 1997-01-23 12h fc



CON+SP - Z500 1998-01-18 12h fc t+120



ANA+ERRCON - Z500 1998-01-23 12h fc



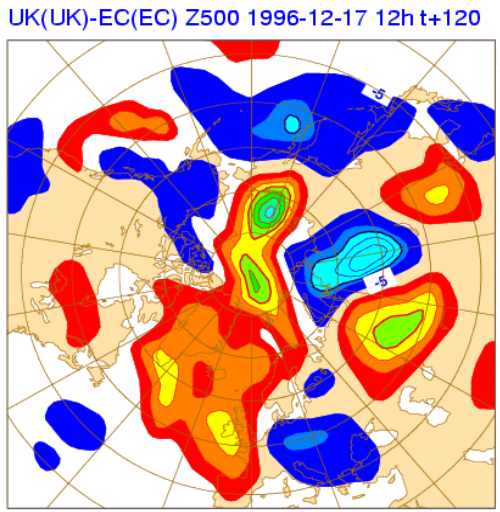
Outline

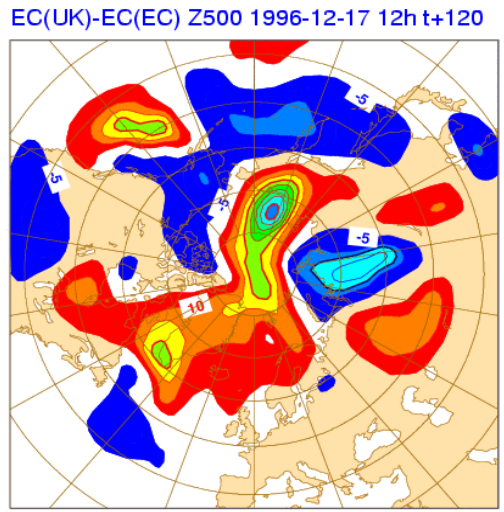
1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

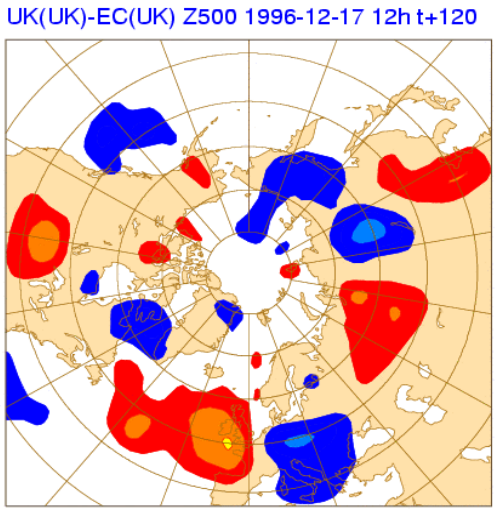
4. What should an ensemble prediction system simulate?

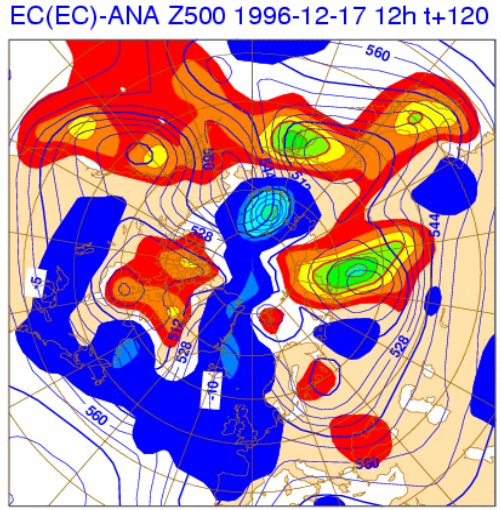
What is the relative contribution of initial and model uncertainties to forecast error?

Richardson (1998, QJRMS) have compared forecasts run with two models (UKMO and ECMWF) starting from either the UKMO or the ECMWF ICs. Results have indicated that initial differences explains most of the differences between ECMWF-from-ECMWF-ICs and UKMO-from-UKMO-ICs forecasts.







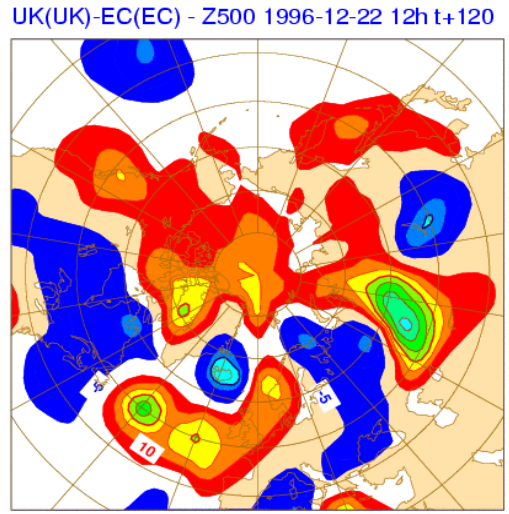


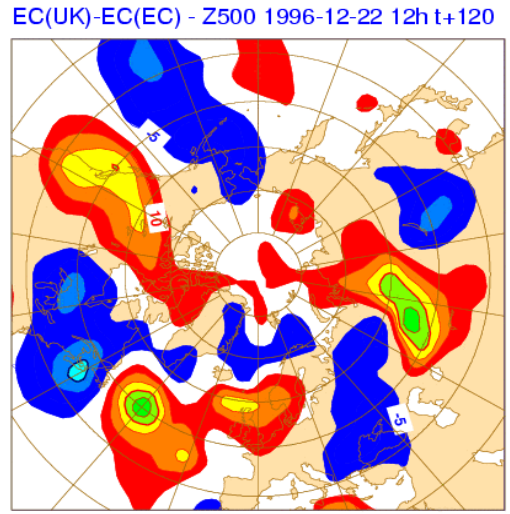
4. Initial uncertainties have a dominant effect

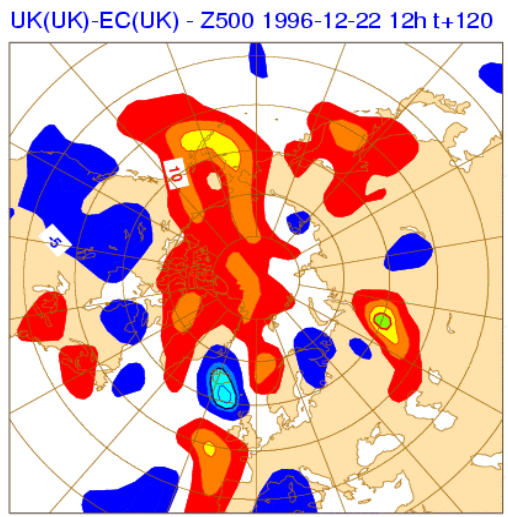
This figure shows the difference between 3 120-hour forecasts: UK(UK) (i.e. UK-from-UK-ICs) and EC(EC) (top left), EC(UK) and EC(EC) (top right), UK(UK) and EC(UK) (bottom left).

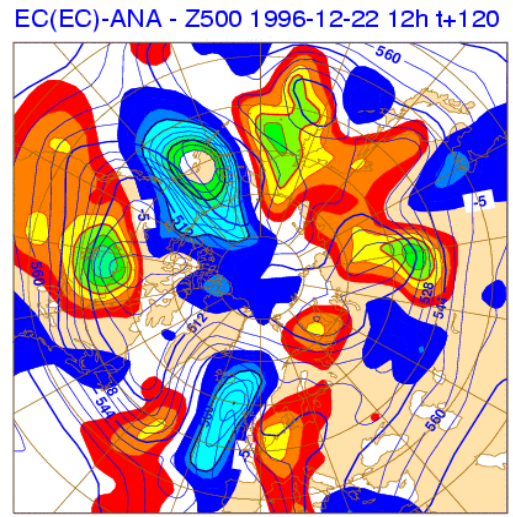
The error of the EC(EC) forecast is also shown (bottom left). Initial differences contributes more than model differences to forecast divergence. This suggests that initial uncertainties contributes more than model approximations to error growth during the first 3-5 forecast days.

How should an ensemble prediction system simulate initial uncertainties?









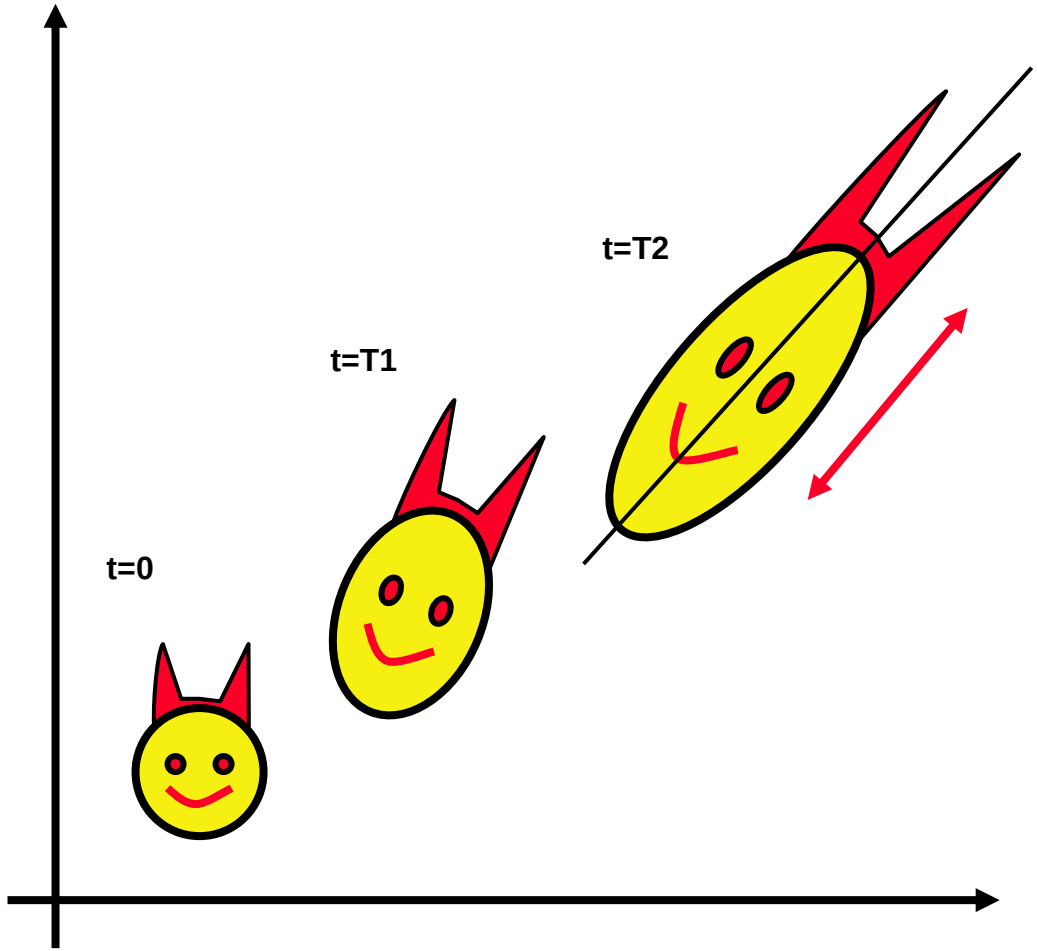
Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

5. How should initial uncertainties be defined?

Perturbations pointing along different axes in the phase-space of the system are characterized by different amplification rates. As a consequence, the initial PDF is stretched principally along directions of maximum growth.

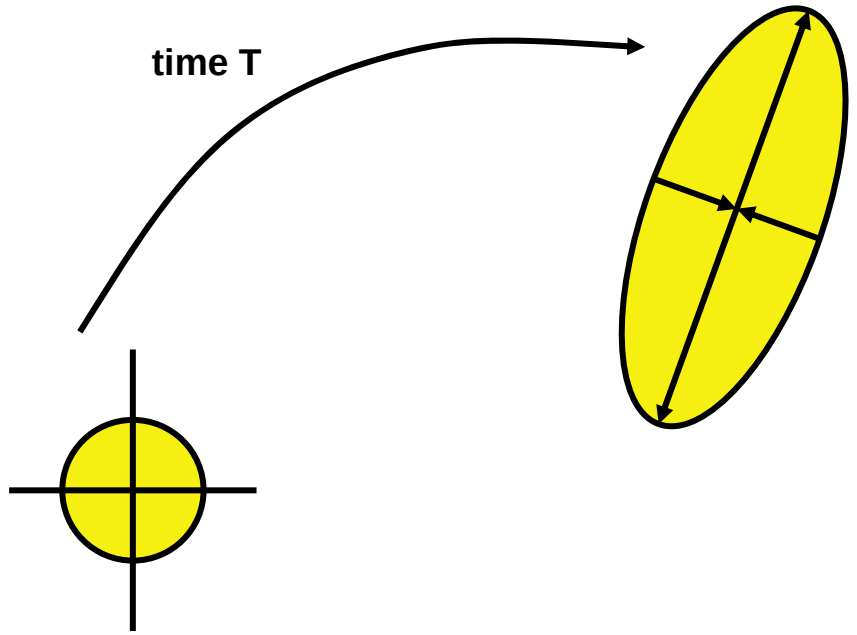
The component of an initial perturbation pointing along a direction of maximum growth amplifies more than a component along another direction (*Buizza et al* 1997).



5. Definition of the initial perturbations

To formalize the problem of the computation of the directions of maximum growth an inner product (metric) should be defined to be able to 'measure' growth.

The metric that has been used at ECMWF in the ensemble system is a total energy, defined as a function of vorticity, divergence, temperature and surface pressure.



$$< x; E_{TE} y >= \frac{1}{2}\iint (\nabla\Delta^{-1}\zeta_x \cdot \nabla\Delta^{-1}\zeta_y + \nabla\Delta^{-1}D_x \cdot \nabla\Delta^{-1}D_y + \frac{C_p}{T_r}T_x T_y) d\Sigma \frac{\partial p}{\partial \eta} d\eta$$

$$+ \int (R_d \frac{T_r}{p_r} \ln \pi_x \ln \pi_y) d\Sigma$$

5. Asymptotic and finite-time instabilities

Farrell (1982) studying perturbations' growth in baroclinic flows notices that the long-time asymptotic behavior is dominated by normal modes, but that there are other perturbations that amplify more than the most unstable normal mode over a finite time interval.

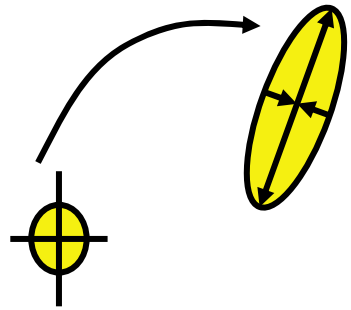
Farrell (1989) showed that perturbations with the fastest growth over a finite time interval could be identified solving an eigenvalue problem of the product of the tangent forward and adjoint model propagators. This result supported earlier conclusions by *Lorenz* (1965).

Calculations of perturbations growing over finite-time interval intervals have been performed, for example, by *Borges & Hartmann* (1992) using a barotropic model, *Molteni & Palmer* (1993) with a quasi-geostrophic 3-level model, and by *Buizza et al* (1993) with a primitive equation model.

5. Singular vectors (see appendix for more details)

The problem of the computation of the directions of maximum growth of a time evolving trajectory reduces to the computation of the singular vectors of $K=E^{1/2}LE_0^{-1/2}$, i.e. to solving the following eigenvalue problem:

$$E_0^{-1/2}L^*ELE_0^{-1/2}\nu = \sigma^2\upsilon$$

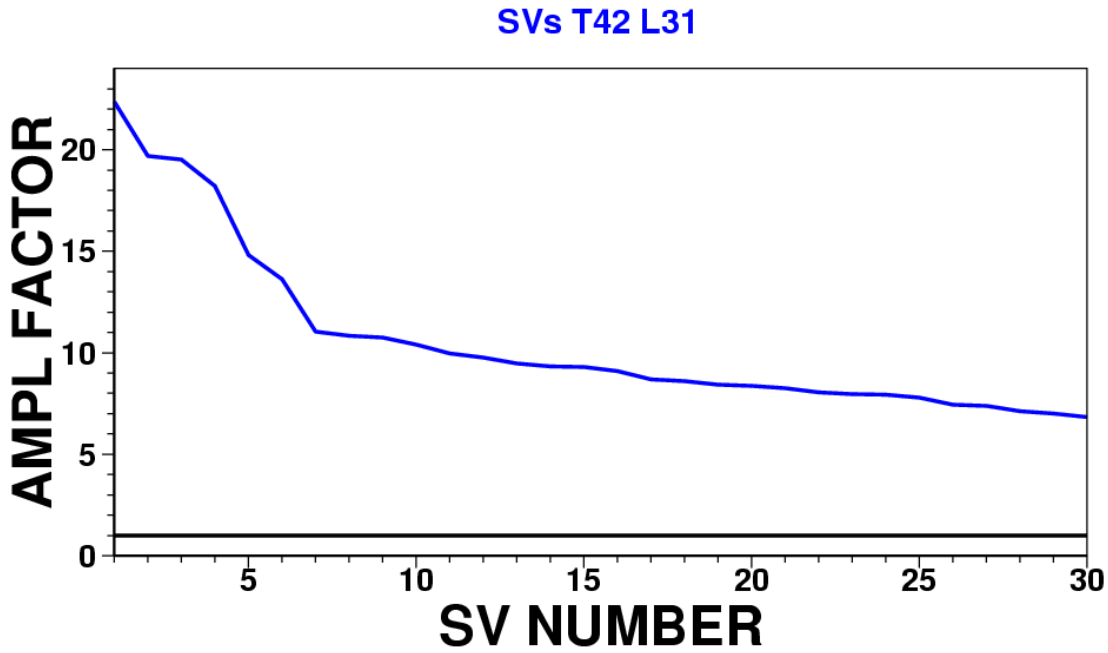


where:

- E_0 and E are the initial and final time metrics
- $L(t,0)$ is the linear propagator, and L* its adjoint
- The trajectory is time-evolving trajectory
- t is the optimization time interval

5. SVs' example: singular values for 18-20 Jan 1997

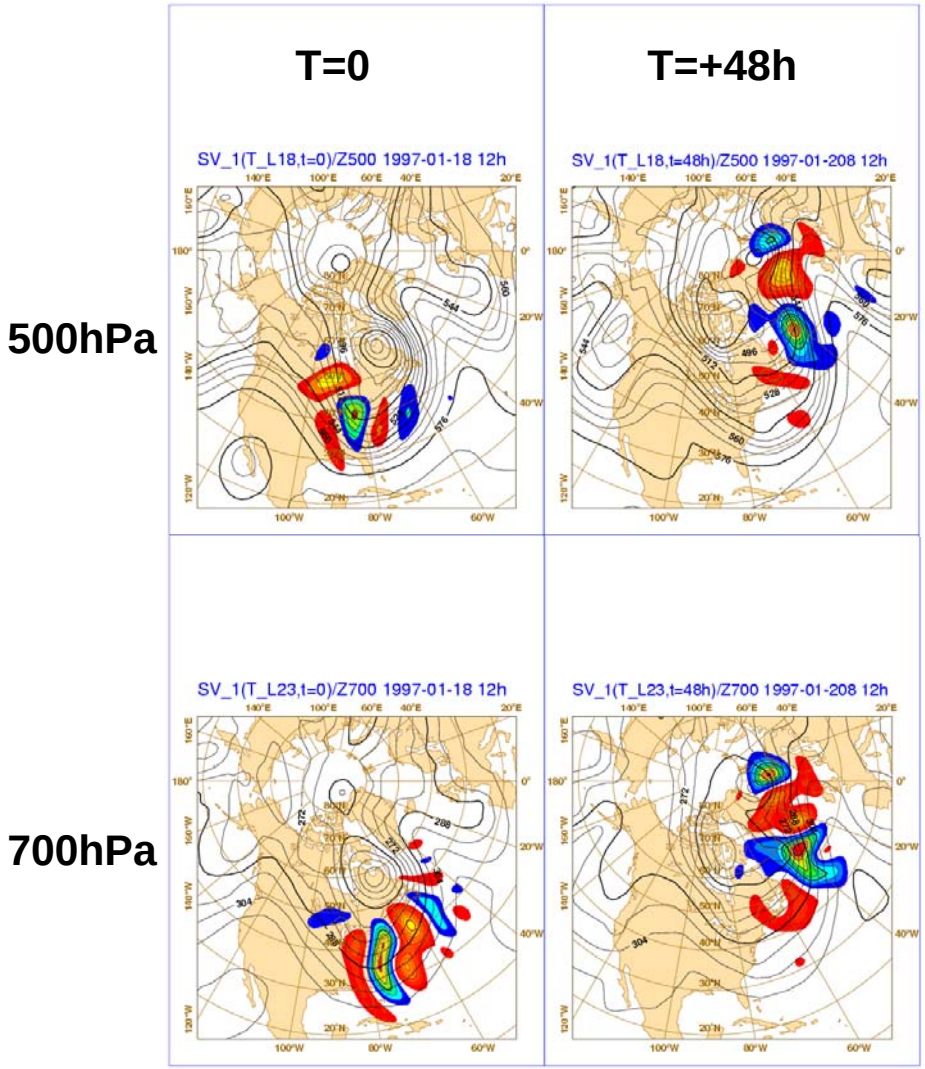
This figure shows the amplification rate (i.e. the singular value) of the leading 30 unstable singular vectors growing between 18 and 20 January 1997. The SVs were computed at the resolution T42L31 and were used to generate the EPS initial conditions.



5. Leading singular vector for 18-20 Jan 1997

This figure shows the most unstable singular vector growing between 18 and 20 Jan 1997. Left (right) panels show the SV at initial and final (i.e. +48h) time.

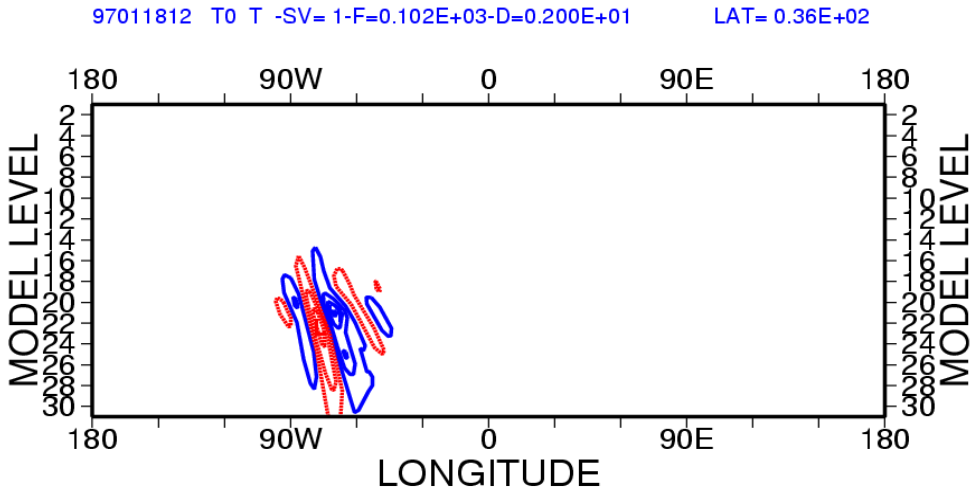
The top panels show the SV T at model level 18 (~500hPa, shading) and the Z500 analysis; the bottom panels the SV T at model level 23 (~700hPa, shading). The contour interval is 8dam for Z, and 0.01 (0.05) deg for T at initial (final) time (the SV is normalized to have unit total energy norm at initial time).

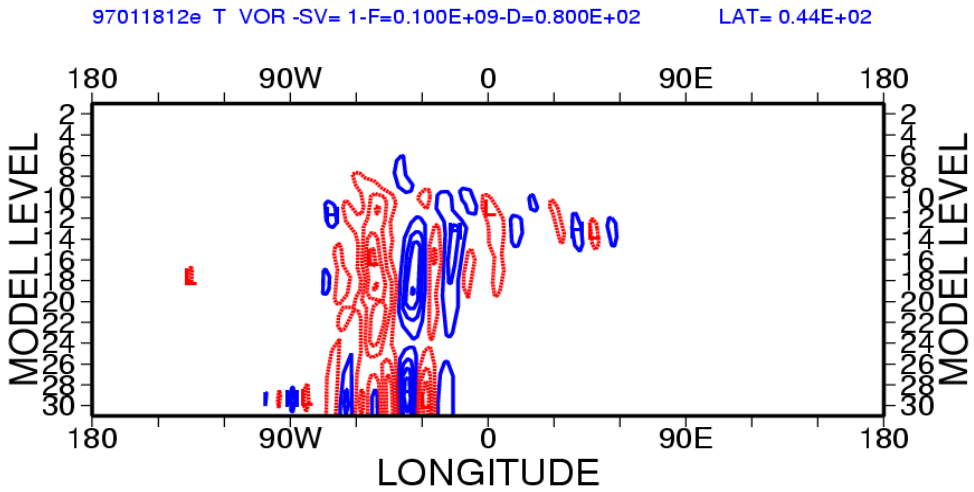


5. Vertical cross section of SV 1 for 18-20 Jan '97

This figure shows, for SV 1, the vertical cross section of the T component at initial time (top, for 36N) and of the vorticity component at final time (bottom, for 44N). The two cross sections have been taken along the parallel where the SV had maximum amplitude.

Note the strong initial tilt, suggesting baroclinic instability, and the final time more barotropic-type structure. Note that T is shown at initial time and vor at final time because the initial time SV has a strong potential energy part.

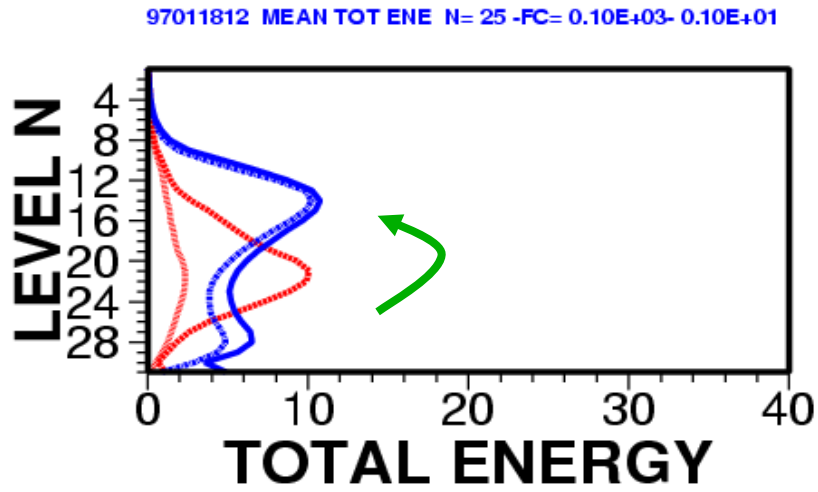


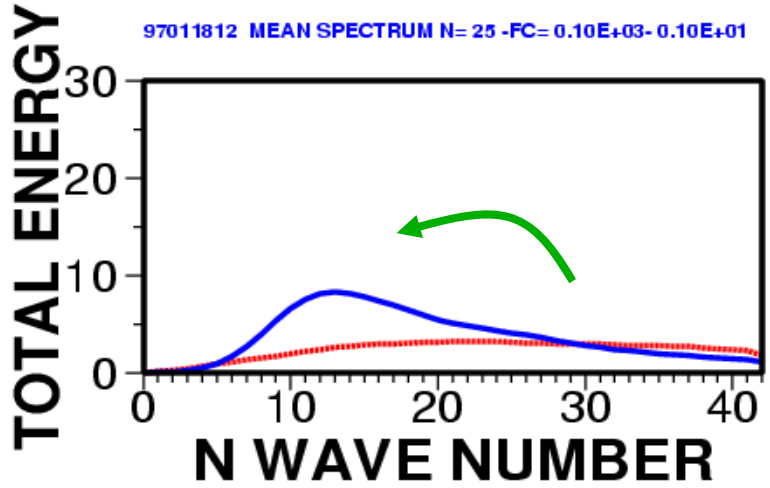


5. Average energy distribution for 18-20 Jan 1997

The top figure shows the SV1:25 average vertical distribution at initial time of the kinetic (red dotted, x100) and total (red solid, x100) energy, and the corresponding final time distributions (blue).

The bottom figure shows the SV1:25 average total energy spectrum at initial (red solid, x100) and at final time (blue solid). Note the SV typical upward and upscale energy transfer/growth, and the transformation from initial potential to mainly final kinetic energy.



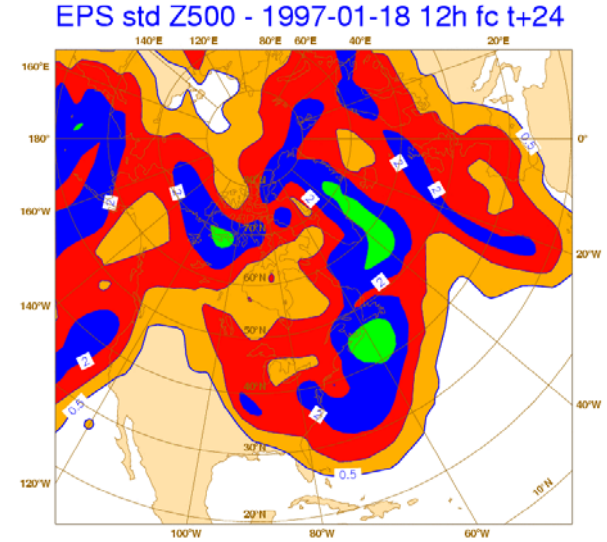


5. SVs' and Eady index for 18-20 Jan 1997

The top panel shows the t+24h average root-mean-square (rms) amplitude (in terms of Z500) of the first 25 singular vectors growing between 18 and 20 January 1997.

The bottom panel shows the 18-20 January 1997 average Eady index. The contour isolines are 0.5dam for the SV's rms amplitude and 0.5d^{-1} for the Eady index.

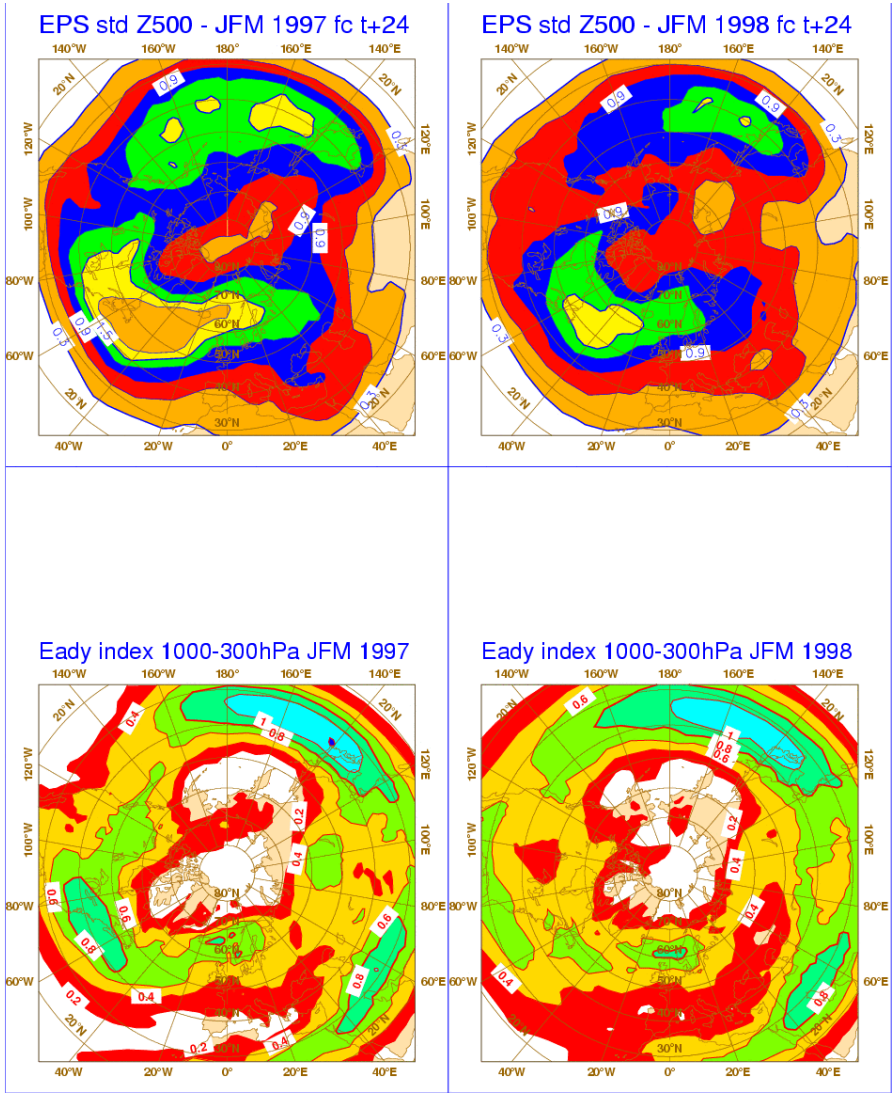
Results indicate a good correspondence between areas of SV concentration and of maximum value of the Eady index.



5. NH SVs and Eady index - JFM '97 & '98

The top panels show the average t+24h root-mean-square amplitude (in terms of Z500, ci=0.3dam) of the first 25 singular vectors during JFM 1997 (left) and 1998 (right) over the NH.

The bottom panels show the average Eady index computed between 1000 and 300 hPa (ci=0.2d^{-1}). Results indicate a good agreement between areas of large Eady index and high SV concentration.



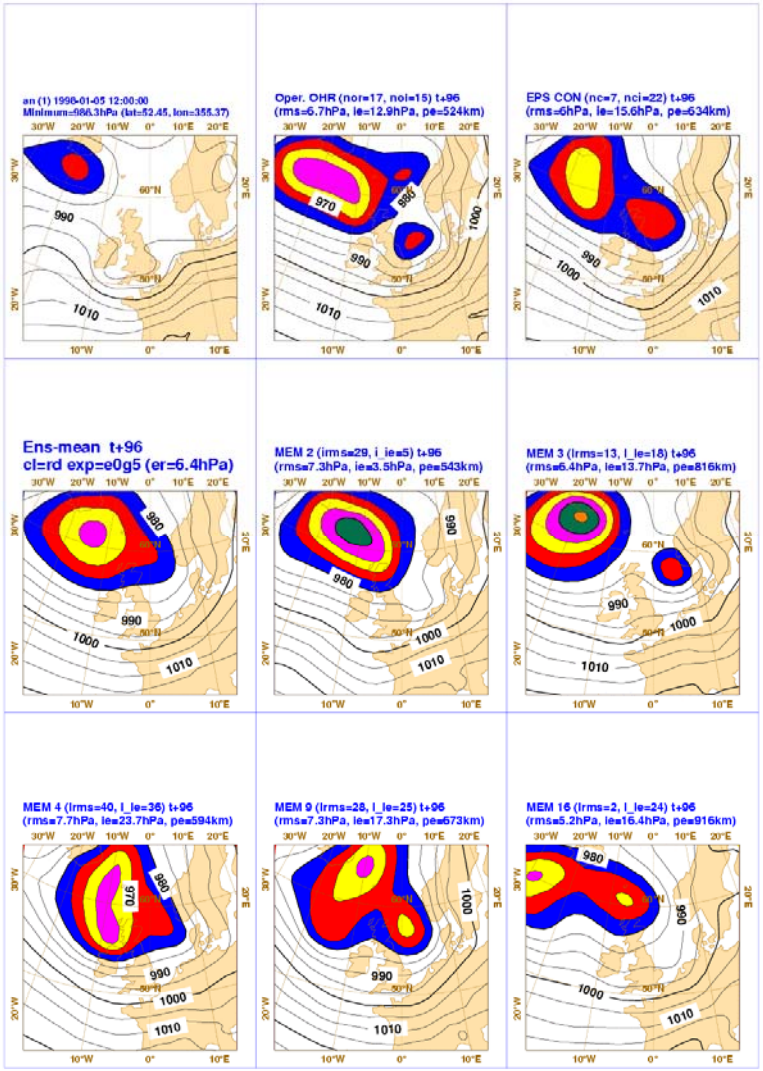
Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

6. Simulation of model uncertainties

The fact that numerical models describe the laws of physics only approximately (model uncertainties) contributes to the growth of forecast errors.

This figure shows the effect of small random perturbations added to the tendencies due to the parameterized physical processes on a 4d forecast (MSLP).



6. Stochastic physics: the rationale

The stochastic scheme (named 'stochastic physics') currently used in the ECMWF ensemble system has been designed such that:

- It is simple and 'robust' to model changes
- It simulates the sort of random errors in parameterized forcing which are coherent among the different parameterization models (moist-processes, radiation, turbulence, ..). A way to take this into account is to apply the stochastic forcing on the total tendency.
- The random numbers that are used to perturb the model tendencies have a space-time correlation designed to emulate the coherence of the model tendencies due to parameterized physical processes (such a coherence represents, e.g., the space and time scales associated with organized convection).
- It should improve the model climate.

6. The EPS with perturbed physics

Each ensemble member evolution is given by the time integration

$$e_j(T) = \int_{t=0}^{T} [A(e_j,t) + P(e_j,t) + \delta P_j(e_j,t)]dt$$

of the perturbed model equations.

The model tendency perturbation is defined at each grid point by

$$\delta P_j(\lambda,\phi,p) = r_j(\lambda,\phi) P_j(\lambda,\phi,p)$$

where r(x) is a random number.

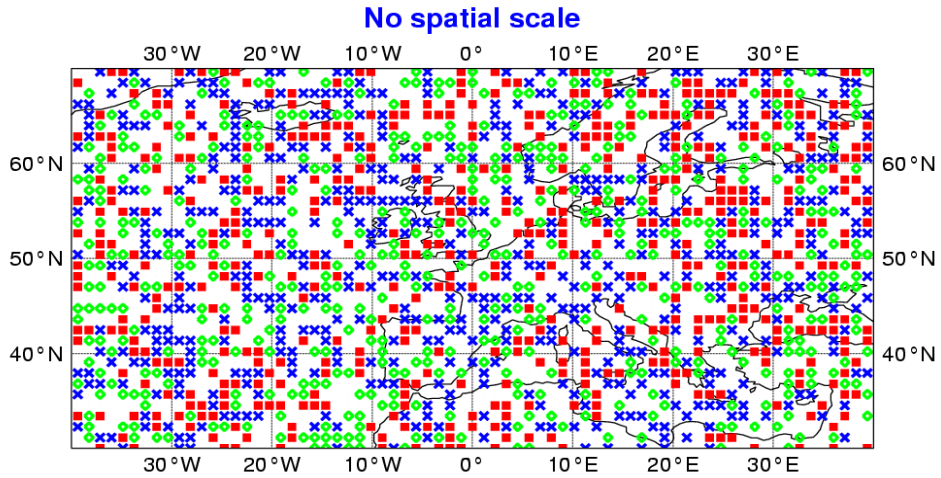
6. Selection of random numbers

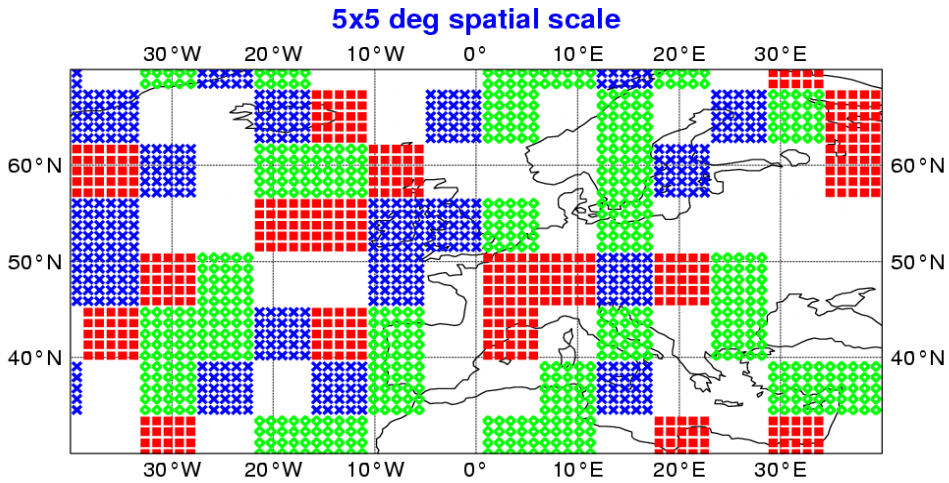
Random numbers can be selected with different spatial correlation scales.

The top figure shows the case when different random numbers are used at each grid point.

The bottom figure shows the case when the same random number is used inside 5-degree boxes. In this case the numbers have been selected inside the interval $-0.5 \leq r(x) \leq 0.5$:

- blue is for $-0.5 \leq r(x) \leq -0.3$;
- green is for $-0.1 \leq r(x) \leq 0.1$;
- red is for $0.3 \leq r(x) \leq 0.5$.

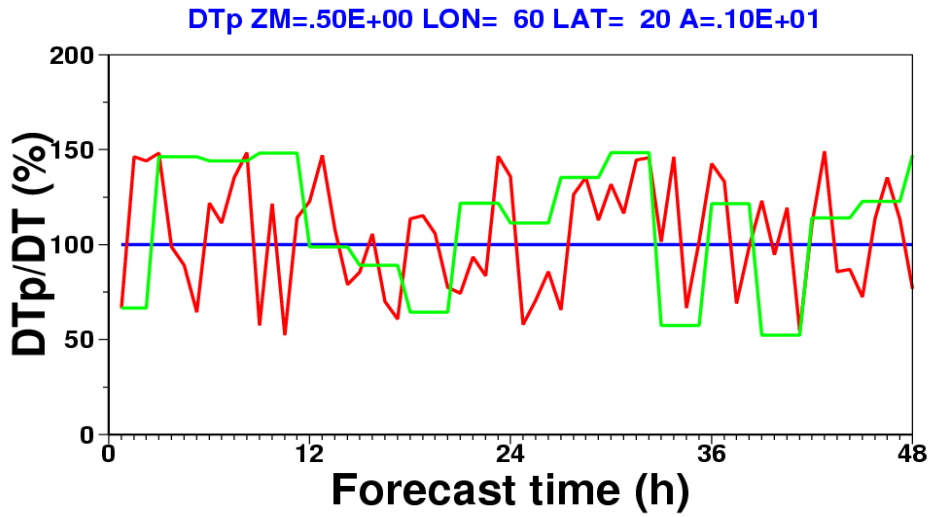


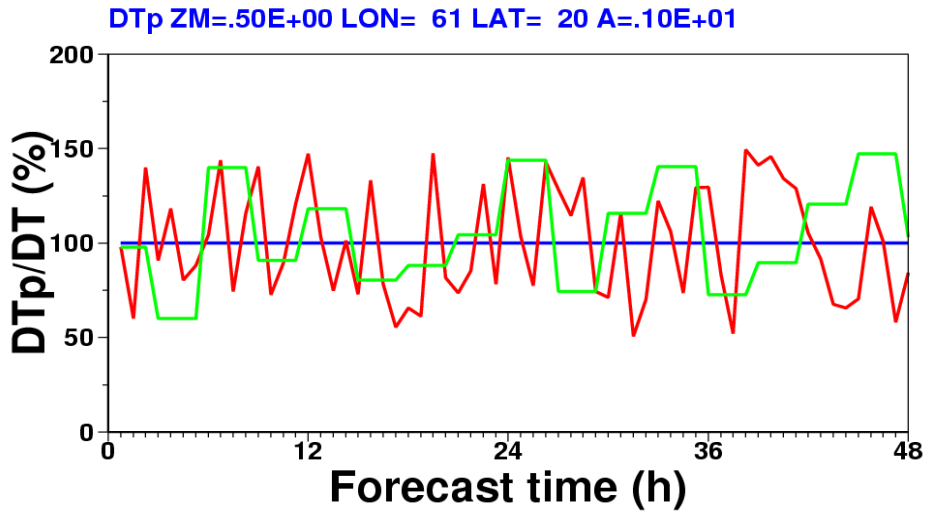


6. Selection of random numbers

Random numbers can be selected with different temporal correlation scales.

These two figures shows the effect on the amplitude of the perturbation tendency for two adjacent boxes when the random numbers are re-selected every time step (red) or every 4 time steps (green). In this case the numbers have been selected inside the interval $-0.5 \leq r(x) \leq 0.5$.

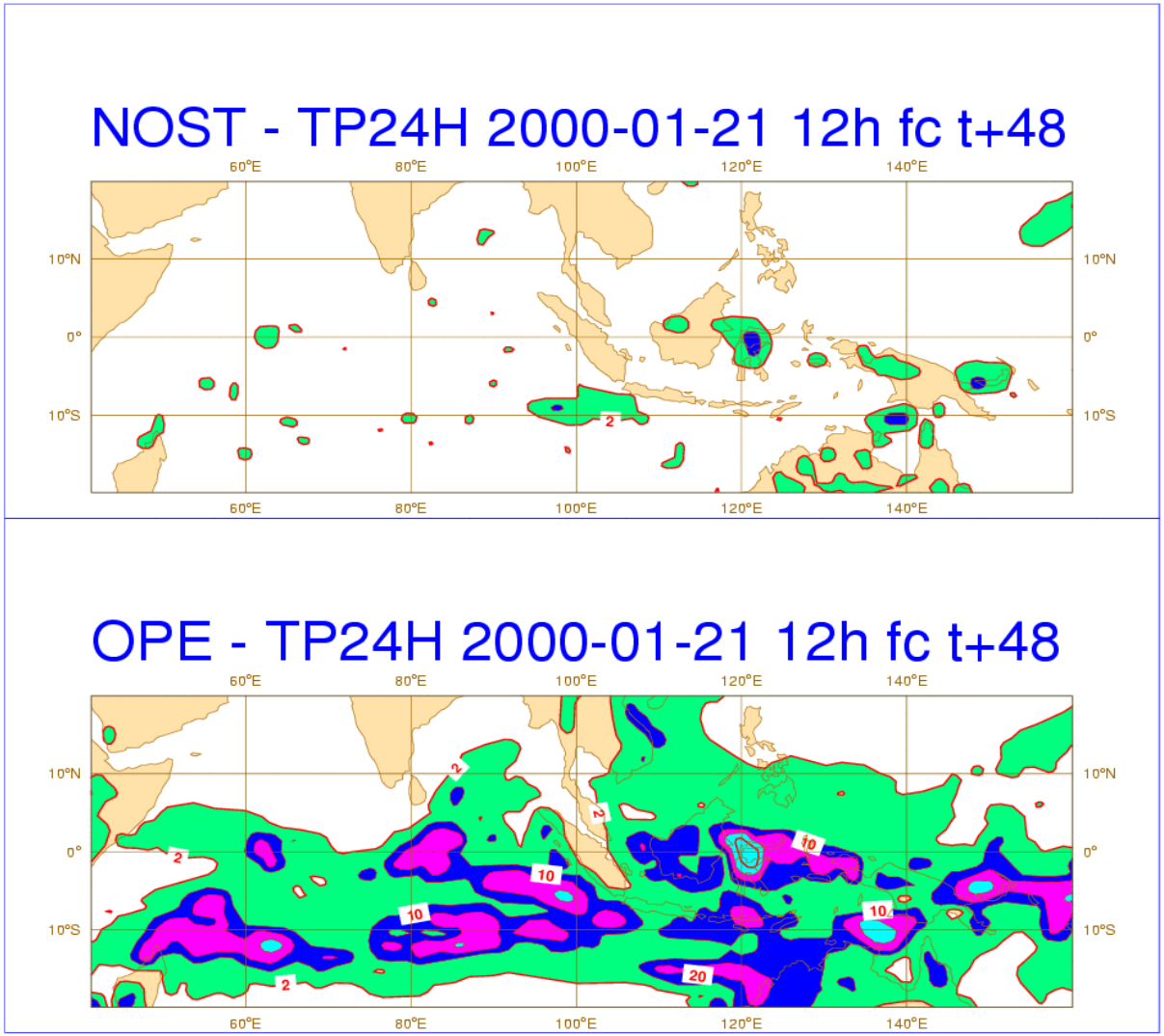




6. Stochastic physics as a source of spread in the tropics

Currently, stochastic physics is the dominant source of ensemble spread in the tropical region (the SVs are optimized to grow north of 30N and south of 30S).

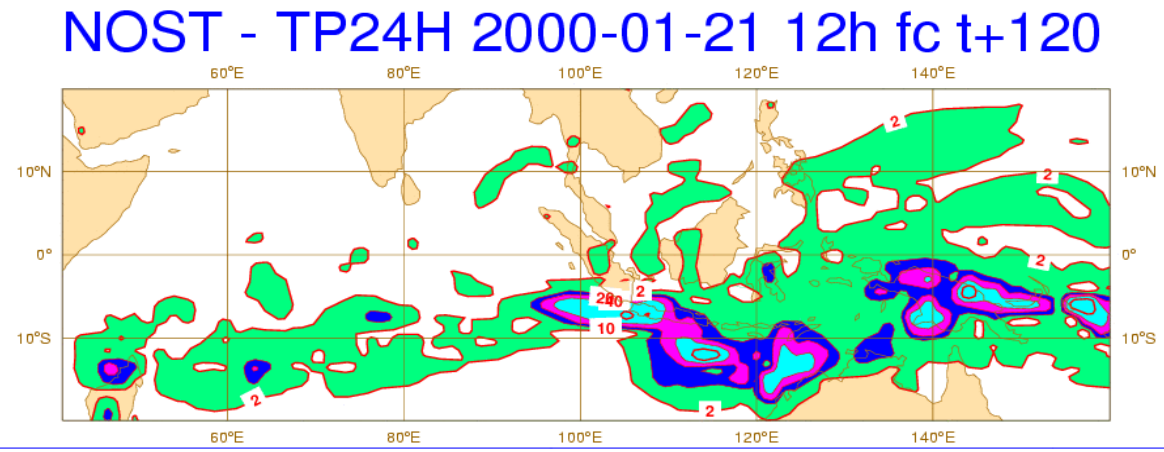
This can be seen, e.g., by comparing the 48-h ensemble spread in terms of precipitation in a tropical region for ensembles run without (top) and with (bottom) stochastic physics.

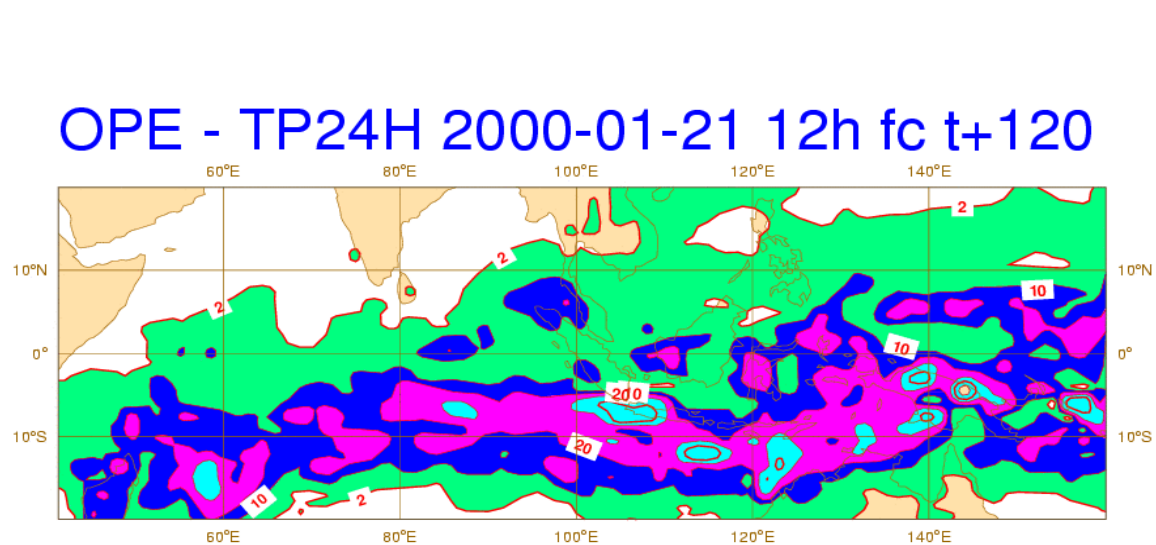


6. Stochastic physics as a source of spread in the tropics

After 5-days of time integration, initial perturbations initiated north of 30N and south of 30S induce divergence also in the tropical region.

Thus, the difference between the ensemble spread in terms of precipitation for ensembles run without (top) and with (bottom) stochastic physics is smaller.

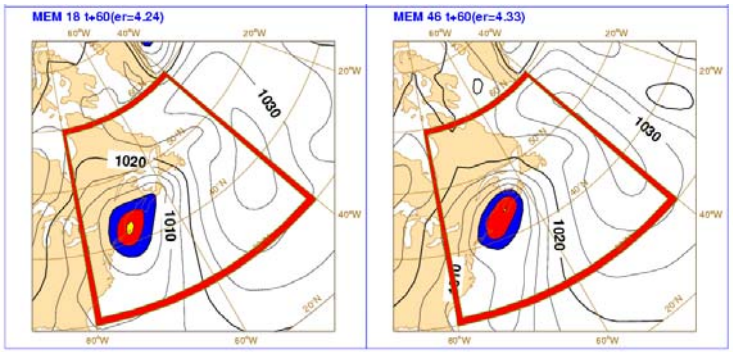




6. Impact of stochastic physics: the US storm case

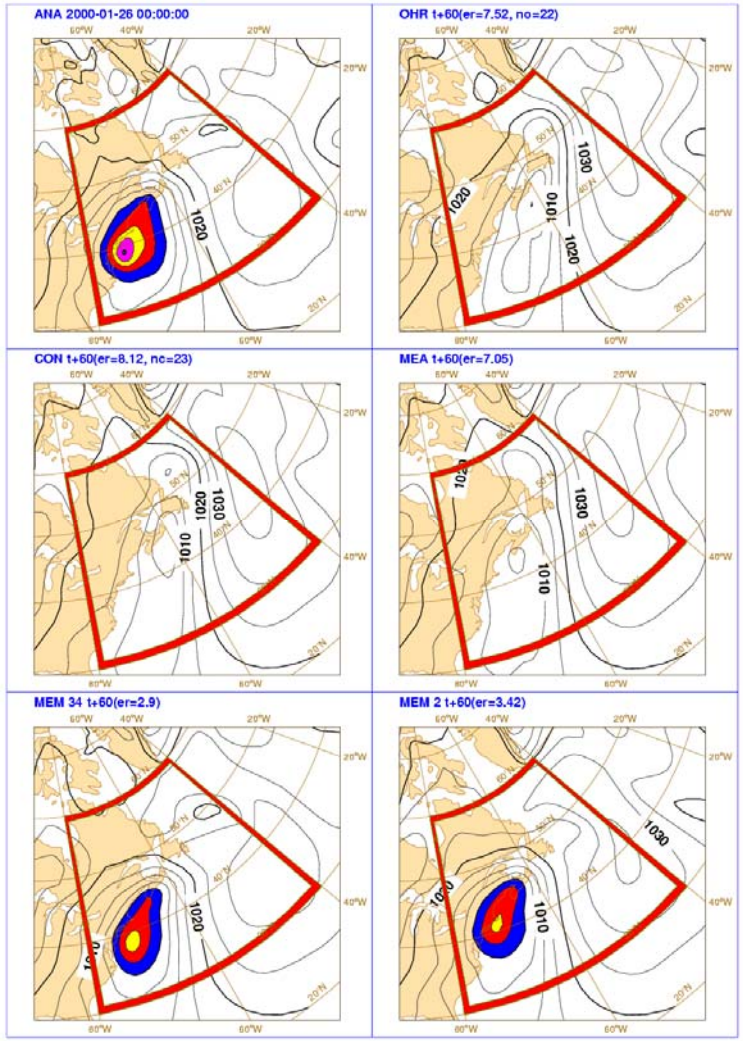
Between 25 and 26 Jan 2000 a very intense storm affected the US East Coast. Sensitivity experiments have indicated that stochastic physics had a key positive role in producing some skilful members.

These figures show the MSLP analysis for 26 Jan 00UTC (top left) and t+60h ensemble forecasts (started on 23 Jan at 12UTC) run with and without stochastic physics.



NOST

OPE



Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

7. The current ECMWF Ensemble Prediction System

The Ensemble Prediction System consists of 51 forecasts run with variable resolution:

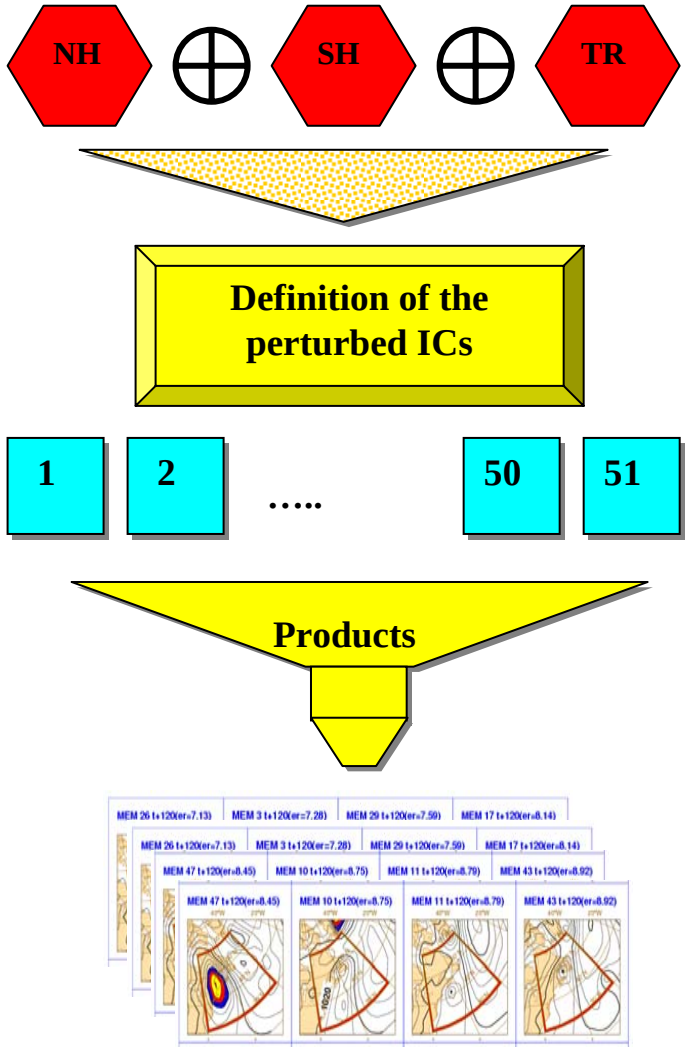
- T_L399L62 (~60km, 62 levels) from day 0 to 10
- T_L255L62 (~80km, 62 levels) from day 10 to 15

[1,5,7,8,13,11,15].

The EPS is run twice a-day, at 00 and 12 UTC.

Initial uncertainties are simulated by perturbing the unperturbed analyses with a combination of T42L62 singular vectors, computed to optimize total energy growth over a 48h time interval (OTI).

Model uncertainties are simulated by adding stochastic perturbations to the tendencies due to parameterized physical processes.



7. The ECMWF Ensemble Prediction System

Each ensemble member evolution is given by the time integration

$$e_j(T) = \int_{t=0}^{T} [A(e_j,t) + P(e_j,t) + \delta P_j(e_j,t)]dt$$

of perturbed model equations starting from perturbed initial conditions

$$e_j(d) = e_0(d) + de_j(d)$$

$$de_j(d) = \sum_{area} \sum_{k=1}^{N_{SV}} [\alpha_{j,k} \cdot SV_k(d,0) + \beta_{j,k} \cdot SV_k(d-2,+2d)]$$

The model tendency perturbation is defined at each grid point by

$$\delta P_j(\lambda,\phi,p) = r_j(\lambda,\phi) P_j(\lambda,\phi,p)$$

where r(x) is a random number.

7. Since May '94 the EPS configuration changed 15 times

Between Dec 1992 and Sep 2006 the ECMWF system changed several times: ~50 model cycles (which included changes in the model and DA system) were implemented, and the EPS configuration was modified 15 times.

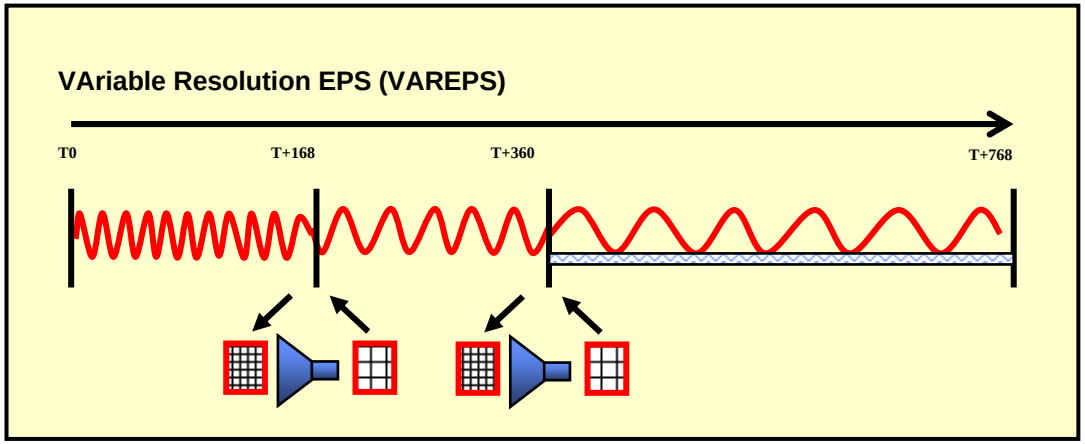
Date	Description	Singular Vectors's characteristics						Forecast characteristics				
		HRES	VRES	OTI	Target area	EVO SVs	sampl	HRES	VRES	Tend	#	Mod Unc
Dec 1992	**Oper Impl**	T21	L19	36h	globe	NO	simm	T63	L19	10d	33	NO
Feb 1993	**SV LPO**	"	"	"	NHx	"	"	"	"	"	"	"
Aug 1994	**SV OTI**	"	"	48h	"	"	"	"	"	"	"	"
Mar 1995	**SV hor resol**	T42	"	"	"	"	"	"	"	"	"	"
Mar 1996	**NH+SH SV**	"	"	"	(NH+SH)x	"	"	"	"	"	"	"
Dec 1996	**resol/mem**	"	L31	"	"	"	"	TL159	L31	"	51	"
Mar 1998	**EVO SV**	"	"	"	"	YES	"	"	"	"	"	"
Oct 1998	**Stoch Ph**	"	"	"	"	"	"	"	"	"	"	YES
Oct 1999	**ver resol**	"	L40	"	"	"	"	"	L40	"	"	"
Nov 2000	**FC hor resol**	"	"	"	"	"	"	TL255	"	"	"	"
Jan 2002	**TC SVs**	"	"	"	(NH+SH)x+TC	"	"	"	"	"	"	"
Sep 2004	**sampling**	"	L40	"	"	"	Gauss	"	"		"	"
Jun 2005	**rev sampl**	"	"	"	"	"	"	"	"		"	"
Feb 2006	**resolution**	"	L62	"	"	"	"	TL399	L62	10d	"	"
Sep 2006	**VAREPS**	T42	L62	48h	(NH+SH)x+TC	YES	Gauss	TL399(0-10)+TL255(10-15)	L62	15d	51	YES

7. The variable resolution ensemble system (VAREPS)

The key idea behind VAREPS is to resolve small-scales in the forecast up to the forecast range when resolving them improves the forecast, but not to resolve them when unpredictable.

VAREPS aims to increase the value of the current EPS in two ways:

- in the short range, by providing more skilful predictions of the small scales
- in the medium-range, by extending the range of skilful products to 15 days



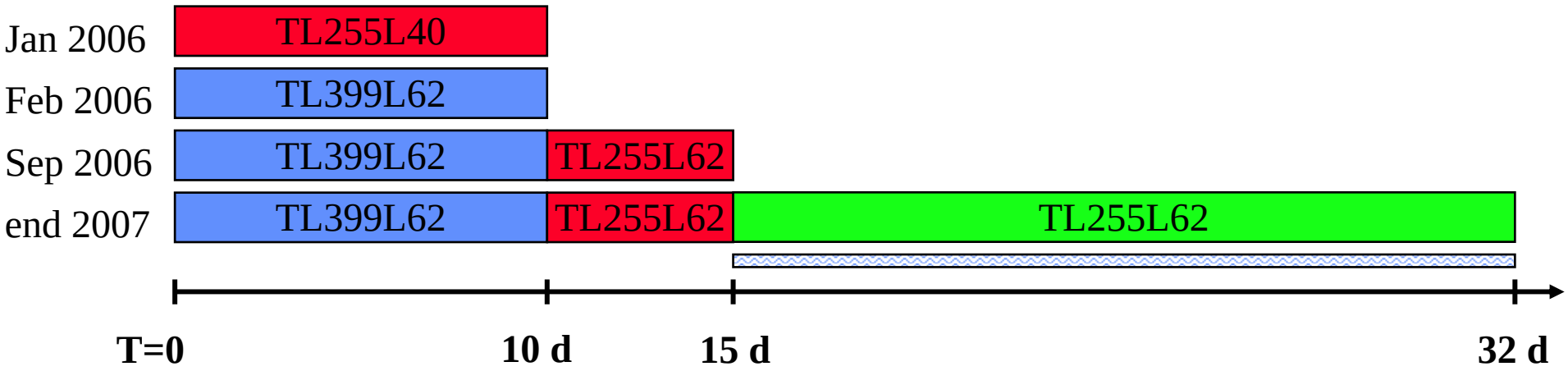
7. VAREPS

Results have indicated that:

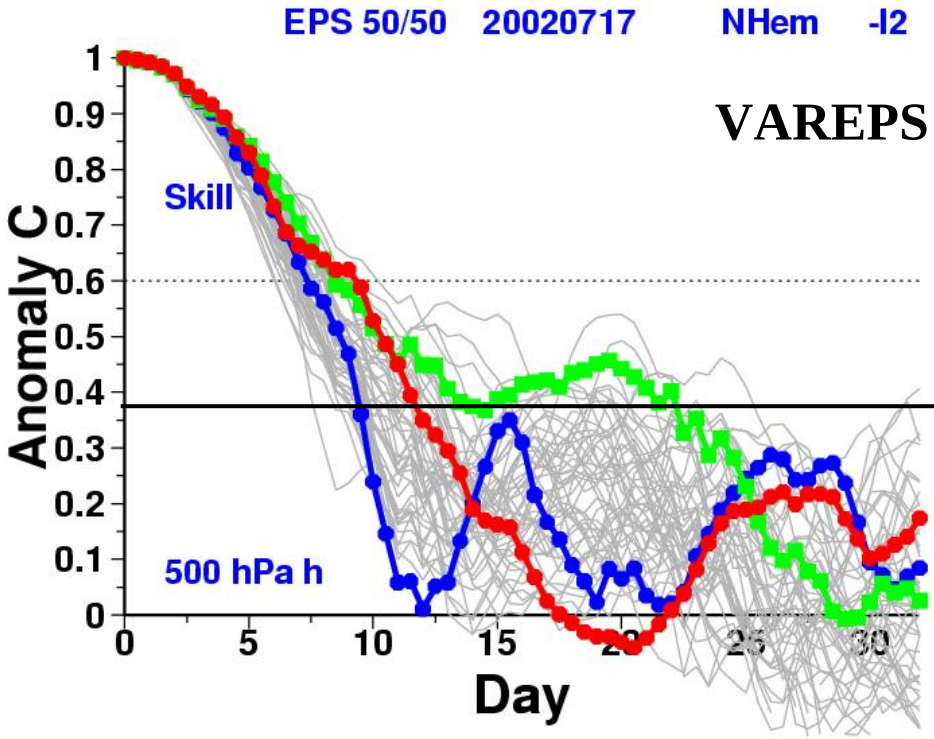
- In the short-range, increasing the EPS resolution improves the average skill, in particular in cases of extreme weather events (hurricanes, small-scale vortices, wind, intense precipitation, ..)
- In the long-range, the impact of increasing resolution can still be detected, but it is less evident
- These results suggest that, given a limited amount of computing resources, it is more valuable (i.e. cost effective) to use most of them in the short-range
- The EPS benefits from better starting from a better analysis

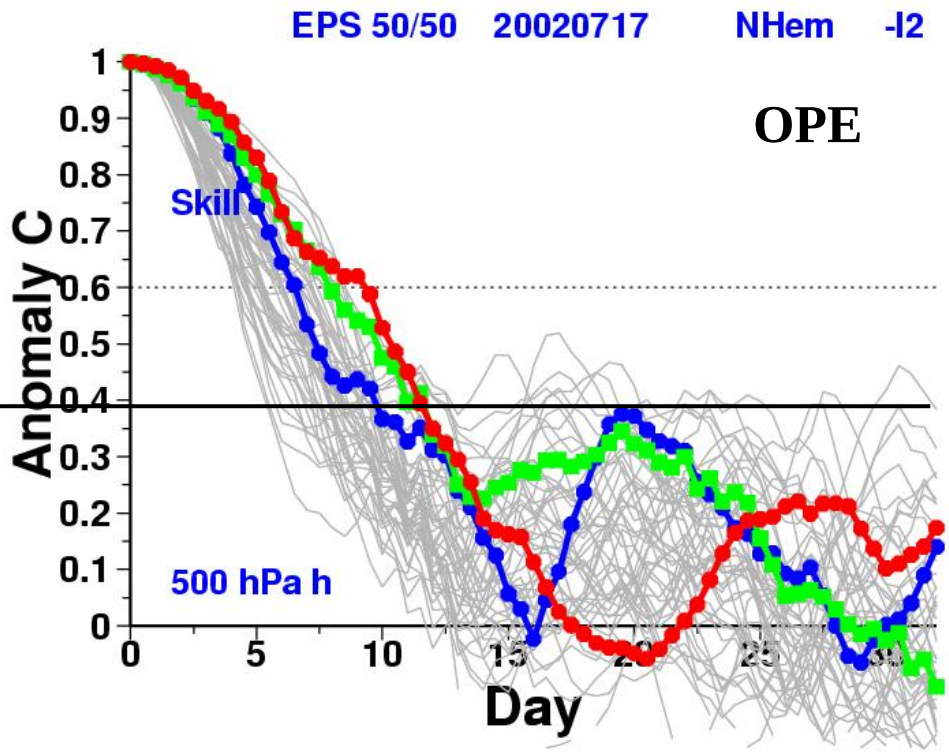
7. The future ensemble systems at ECMWF

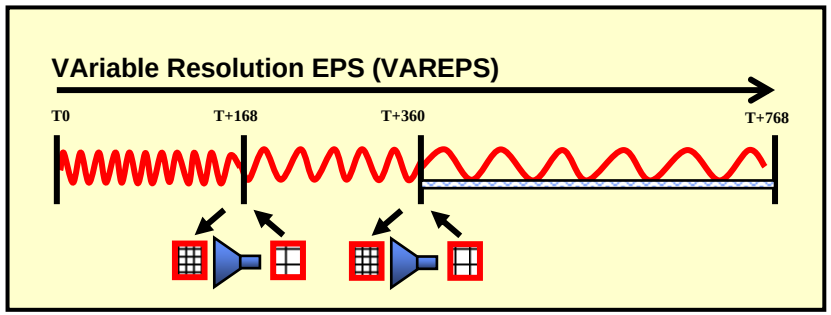
- Until 1 Feb '06, the EPS had 51 10-day forecasts at T_L255L40 resolution
- On the 1st of Feb '06, the 10-day EPS resolution was upgraded to T_L399L62
- On the 12th of Sep '06, the new Variable Resolution EPS (VAREPS) was introduced, and the ensemble forecast range was extended to 15 days
- The next change will be to link the 15-day VAREPS with the monthly forecast system will continue, with the goal to implement a seamless d0-32 VAREPS



7. First case of a 3-legs VAREPS (17 July 2002)







Outline

1. Sources of forecast errors: initial and model uncertainties
2. Flow-dependent predictability
3. The probabilistic approach to NWP
4. Ensemble prediction as a practical tool for probabilistic prediction
5. The simulation of initial uncertainties in ensemble prediction
6. The simulation of model errors in ensemble prediction
7. The ECMWF Ensemble Prediction System
8. Conclusions

8. Conclusion

- Initial and model uncertainties are the main sources of error growth. Initial uncertainties dominates during the first 3-5 forecast days. Predictability is flow dependent.
- A complete description of weather prediction can be stated in terms of an appropriate probability density function (PDF). Ensemble prediction based on a finite number of deterministic integration appears to be the only feasible method to predict the PDF beyond the range of linear growth.
- The initial error components along the directions of maximum growth contribute most to forecast error growth. These directions are identified by the leading singular vectors, and are computed by solving an eigenvalue problem.
- The EPS changed 14 times between 1 May 1994 (first day of daily production) and now. Currently, it includes 50 perturbed and 1 unperturbed 15-day forecasts with variable resolution [T_L399L62(0-10)+T_L255L62(10-15)]. Soon, once a-week the 15-day VAREPS will be extended to 32-day, with a coupled ocean from day 10.

Acknowledgements

The success of the ECMWF EPS is the result of the continuous work of ECMWF staff, consultants and visitors who had continuously improved the ECMWF model, analysis, diagnostic and technical systems, and of very successful collaborations with its member states and other international institutions.

The work of all contributors is acknowledged: *Judith Berner, Jean Bidlot, Manuel Fuentes, Mats Hamrud, Graham Holt, Martin Leutbecher, Tim Palmer, Frederic Vitart* and *Nils Wedi*, and former ECMWF staff who worked directly on the ECMWF Ensemble Prediction System: *Jan Barkmeijer, Franco Molteni, Robert Mureau, Anders Persson, Otto Pessonen, Thomas Petroliagis, David Richardson, Stefano Tibaldi*, and visitors and consultants *Bill Bourke, Mariane Coutinho, Martin Ehrendorfer, Ron Gelaro, Isla Gilmour, Dennis Hartmann, Andrea Montani, Steve Mullen, Kamal Puri, Carolyn Reynolds, Joe Tribbia*. (I hope that the list of names is complete: please forgive me if this is not the case.)

Bibliography

On optimal perturbations and singular vectors

- Borges, M., & Hartmann, D. L., 1992: Barotropic instability and optimal perturbations of observed non-zonal flows. *J. Atmos. Sci.*, 49, 335-354.
- Buizza, R., & Palmer, T. N., 1995: The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, 52, 1434-1456.
- Buizza, R., Tribbia, J., Molteni, F., & Palmer, T. N., 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus*, 45A, 388-407.
- Coutinho, M. M., Hoskins, B. J., & Buizza, R., 2004: The influence of physical processes on extratropical singular vectors. *J. Atmos. Sci*., 61, 195-209.
- Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, 39, 1663-1686.
- Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, 46, 1193-1206.
- Hoskins, B. J., Buizza, R., & Badger, J., 2000: The nature of singular vector growth and structure. *Q. J. R. Meteorol. Soc*., 126, 1565-1580.
- Lorenz, E., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17, 321-333.
- Molteni, F., & Palmer, T. N., 1993: Predictability and finite-time instability of the northern winter circulation. *Q. J. R. Meteorol. Soc.*, 119, 1088-1097.

Bibliography

On normal modes and baroclinic instability

- Birkoff & Rota, 1969: *Ordinary differential equations*. J. Wiley & sons, 366 pg.
- Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteorol.*, **4**, 135-162.
- Eady. E. T., 1949: long waves and cyclone waves. *Tellus*, **1**, 33-52.

On SVs and predictability studies

- Buizza, R., Gelaro, R., Molteni, F., & Palmer, T. N., 1997: The impact of increased resolution on predictability studies with singular vectors. *Q. J. R. Meteorol. Soc.*, **123**, 1007-1033.
- Gelaro, R., Buizza, R., Palmer, T. N., & Klinker, E., 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, **55**, 6, 1012-1037.

On the validity of the linear approximation

- Buizza, R., 1995: Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. *Q. J. R. Meteorol. Soc.*, **121**, 1705-1738.
- Gilmour, I., Smith, L. A., & Buizza, R., 2001: On the duration of the linear regime: is 24 hours a long time in weather forecasting?. *J. Atmos. Sci.*, **58**, 3525-3539 (also EC TM 328).

Bibliography

On the ECMWF Ensemble Prediction System

- Molteni, F., Buizza, R., Palmer, T. N., & Petroliagis, T., 1996: The new ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, 122, 73-119.
- Buizza, R., Miller, M., & Palmer, T. N., 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2887-2908 (also EC TM 279).
- Buizza, R., & Hollingsworth, A., 2002: Storm prediction over Europe using the ECMWF Ensemble Prediction System. *Meteorol. Appl.*, 9, 1-17.
- Buizza, R., Richardson, D. S., & Palmer, T. N., 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Q. J. R. Meteorol. Soc*.. 129, 1269-1288.
- Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., & Zhu, Y., 2005: A comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev.*, 133, 5, 1076,1097.
- Buizza, R., Bidlot, J.-R., Wedi, N., Fuentes, M., Hamrud, M., Holt, G., & Vitart, F., 2007: The new ECMWF VAREPS. *Q. J. Roy. Meteorol. Soc.*, in press (also EC TM 499).
- Leutbecher, M. & T.N. Palmer, 2007: Ensemble forecasting. *J. Comp. Phys.*, in press (also EC TM 514).
- Mullen, S., & Buizza, R., 2001: Quantitative precipitation forecasts over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*,129, 638-663.

Bibliography

On SVs and targeting adaptive observations

- Buizza, R., & Montani, A., 1999: Targeting observations using singular vectors. *J. Atmos. Sci.*, 56, 2965-2985 (also EC TM 286).
- Palmer, T. N., Gelaro, R., Barkmeijer, J., & Buizza, R., 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, 55., 6, 633-653.
- Majumdar, S., Bishop, C., Buizza, R., & Gelaro, R., 2002: A comparison of PSU-NCEP Ensemble Transformed Kalman Filter targeting guidance with ECMWF and NRL Singular Vector guidance. *Q. J. R. Meteorol. Soc.*, 128, 1269-1288.
- Majumdar, S J, Aberson, S D, Bishop, C H, Buizza, R, Peng, M, & Reynolds, C, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev*., 134, 2354-2372 (also EC TM 482).
- Reynolds, C, Peng, M, Majumdar, S J, Aberson, S D, Bishop, C H, & Buizza, R, 2007: Interpretation of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.,* in press.
- Buizza, R., Cardinali, C., Kelly, G., & Thepaut, J.-N., 2007: The value of targeted observations - Part II: the value of observations taken in singular vectors-based target areas. *Q. J. R. Meteorol. Soc*., submitted (also EC TM 512).

Appendix: singular vector definition

Hereafter, some more details on the definition of singular vectors, and their relationship with normal modes, is reported.

Inner product and norm definition

Given two state-vectors *x* and *y* expressed in terms of vorticity ζ, divergence *D*, temperature *T*, specific humidity *q* and surface pressure π, the following inner products (and the associated norms) can be defined (<..,..> is the Euclidean inner product):

- total energy inner product (no humidity term):

$$< x; E_{TE} y >= \frac{1}{2} \iint (\nabla \Delta^{-1} \zeta_x \cdot \nabla \Delta^{-1} \zeta_y + \nabla \Delta^{-1} D_x \cdot \nabla \Delta^{-1} D_y + \frac{C_p}{T_r} T_x T_y) d\Sigma \frac{\partial p}{\partial \eta} d\eta$$

$$+ \int (R_d \frac{T_r}{p_r} \ln \pi_x \ln \pi_y) d\Sigma$$

- enstrophy inner product:

$$< x; E_{Ens} y >= \frac{1}{2} \iint (\nabla \Delta^{-1} \zeta_x \cdot \nabla \Delta^{-1} \zeta_y) d\Sigma \frac{\partial p}{\partial \eta} d\eta$$

- ψ-square inner product:

$$< x; E_{\psi^2} y >= \frac{1}{2} \iint (\psi_x \psi_y) d\Sigma \frac{\partial p}{\partial \eta} d\eta$$

Inner product and norm definition

Denote by $\zeta^{n,l}$ the level-l vorticity component with total wave number n, by $D^{n,l}$ …. of a state vector x. The norm of *x* can be written in matrix form as:

$$\|x\|_{TE}^2 = \frac{1}{2}\sum_l \sum_n \begin{pmatrix} \zeta_x^{n,l} & D_x^{n,l} & T_x^{n,l} \end{pmatrix} \begin{pmatrix} -[R_a^2/n(n+1)]\delta p_l & 0 & 0 \\ 0 & -[R_a^2/n(n+1)]\delta p_l & 0 \\ 0 & 0 & [C_p/T_r]\delta p_l \end{pmatrix} \begin{pmatrix} \zeta_x^{n,l} \\ D_x^{n,l} \\ T_x^{n,l} \end{pmatrix}$$
$$+ \sum_n R_d\, T_r / p_r \ln \pi_x^n \ln \pi_x^n$$

where n is the total wave number, δp is the pressure difference between two half-levels; T_r=350deg and p_r=100kPa are reference values; R_a=6371km, R_d=287JK^{-}kg^{-1}, C_p=1004JK^{-}kg^{-1}.

Definition of the system instabilities: normal modes

Consider an N-dimensional autonomous system:

$$\frac{\partial y}{\partial t} = A(y)$$

The method most commonly applied to study the stability of a solution z of the system equations is based on normal modes, whereby small disturbances are resolved into modes which may be treated separately because each of them satisfies the system equations. The system equations are linearized around the constant solution z:

$$\frac{\partial y}{\partial t} = A_l(z)y \qquad A_l(z) = \left.\frac{\partial A(z)}{\partial z}\right|_z$$

A normal mode is a solution of the linearized equations of the form:

$$y(x,t) = f(x)e^{\lambda t}$$

Definition of the system instabilities: normal modes

By substituting the normal mode definition into the linear equations an eigenvalue problem is defined:

$$A_l(z)\,f(x) = \lambda f(x)$$

The eigenvectors with real positive eigenvalues λ identify the unstable normal modes of the systems. A system is defined asymptotically stable if and only if every eigenvalue has negative real part.

Charney (1947) and *Eady* (1949) considered idealized atmospheric flows and by applying a normal-mode stability analysis they studied the baroclinic instability mechanism and showed that the zonal mean component of realistic mid-latitude flows is unstable. The resulting exponentially growing structure proved to have length and time scales similar to observed atmospheric cyclogenesis.

Singular vector definition: the linear equations

Consider an N-dimensional autonomous system:

$$\frac{\partial y}{\partial t} = A(y)$$

Denote by *z'* a small perturbation around a time-evolving trajectory *z*:

$$\frac{\partial z'}{\partial t} = A_l(z)z' \qquad A_l(z) = \left.\frac{\partial A(z)}{\partial z}\right|_z$$

$$\frac{\partial z}{\partial t} = A(z)$$

The time evolution of the small perturbation *z'* is described to a good degree of approximation by the linearized system $A_l(z)$ defined by the trajectory. Note that the trajectory is not constant in time.

Singular vector definition: the linear propagator

The perturbation *z'* at time *t* is given by the time integration from the initial state *z'(t=0)* of the linear system:

$$z'(t) = z'_0 + \int_0^t A_l(z) d\tau$$

The solution can be written in terms of the linear propagator *L(t,0)*:

$$z'(t) = L(t,0) z'_0$$

The linear propagator is defined by the system equations and depends on the trajectory characteristics. The E-norm of the perturbation at time *t* is given by:

$$\|z'(t)\|^2 = < z'(t); Ez'(t) > = < L(t,0)z'_0; EL(t,0)z'_0 >$$

Singular vector definition: the adjoint operator

Given any two vectors *x* and *y*, the adjoint operator L^* of the linear operator *L* with respect to the Euclidean norm <..,..> is the operator that satisfies the following property:

$$< L^* x; y >=< x; Ly >$$

Using the adjojnt operator L^* the time-t E-norm of z' can be written as:

$$\left\| z'(t) \right\|^2 =< Lz_0'; ELz_0' >=< z_0'; L^* ELz_0' >$$

Singular vector definition: the problem

The problem of the computation of the directions of maximum growth can be stated as 'finding the directions in the phase-space of the system characterized by the maximum ratio between the time-t and the initial norms'. Formally, this problem reduces to an eigenvector problem:

$$\max_{x_0 \in \Sigma} \frac{\|x(t)\|_E^2}{\|x_0\|_E^2} = \max_{x_0 \in \Sigma} \frac{< x_0; L^* E L x_0 >}{< x_0; E x_0 >}$$

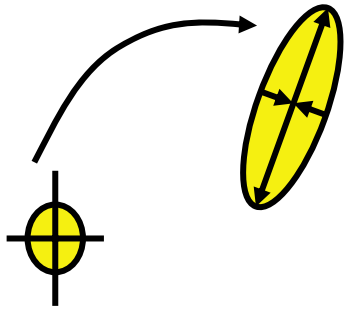
The problem can be generalized by using two different norms at initial and final time:

$$\max_{x_0 \in \Sigma} \frac{\|x(t)\|_E^2}{\|x_0\|_{E_0}^2} = \max_{x_0 \in \Sigma} \frac{< x_0; L^* E L x_0 >}{< x_0; E_0 x_0 >}$$

Singular vector definition: the eigenvalue problem

Apply the following coordinate transformation:

$$y = E_0^{1/2} x$$



Then the generalized problem reduces to:

$$\max_{x_0} \frac{\|x(t)\|_E^2}{\|x_0\|_{E_0}^2} = \max_{y_0} \frac{< y_0; E_0^{-1/2} L^* E L E_0^{-1/2} y_0 >}{< y_0; y_0 >}$$

The directions of maximum growth are defined by the following eigenvalue problem:

$$E_0^{-1/2} L^* E L E_0^{-1/2} \nu = \sigma^2 \upsilon$$