Taming the butterfly effect to reach sub-seasonal and seasonal predictability

Roberto Buizza
Scuola Universitaria Superiore Sant’Anna Pisa
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
Weather prediction is an initial value problem

The $j$-th forecast starting from data/time $(d,0)$, is given by the time integration

$$e_j(d,T) = e_j(d,0) + \int_0^T [A(e_j,t) + P(e_j,t) + \delta P_j(e_j,t)]dt$$

of the model equations starting from the $j$-th initial conditions

$$e_j(d,0) = e_0(d,0) + de_j(d,0)$$

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

The perturbed model tendency is defined at each grid point by

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

where $r_j(t; \Phi, \lambda)$ is a random number.
Models are based on the fluid eq.

\[ \frac{dv}{dt} = -2\Omega x v - \frac{1}{\varrho} \nabla p + \mathbf{g} + P\nu \]

\[ \frac{dT}{dt} = \frac{RT\omega}{c_p p_s \sigma} + P_T \]

\[ \frac{dq}{dt} = P_q \]

\[ \frac{dp_s}{dt} = P_s \left( \nabla \cdot \mathbf{v} + \frac{d}{d\sigma} \frac{d}{dt} \right) \]

\[ \frac{d\phi}{d\sigma} = -\frac{RT}{\sigma} \]

These terms simulate the impact on the state variables of the physical processes (e.g. radiation, moist processes, turbulence, impact of sub-grid scale processes, ..).
Proper initialization is essential to go from obs to fcs

- Approximately 20 million observations
- Data assimilation
- First guess
- Initial conditions
- Model next step
- Intermediate results
- Forecast
- Archive

The atmosphere is divided into about one billion cells each having associated parameters such as temperature, pressure, and wind direction.
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
The atmosphere is a chaotic system

*Ed Lorenz* (1969): 3-d model for a two-dimensional fluid layer uniformly warmed from below and cooled from above.

\[
\begin{align*}
\frac{dx}{dt} &= \sigma(y - x) \\
\frac{dy}{dt} &= x(q - z) - y \\
\frac{dz}{dt} &= xy - \beta z
\end{align*}
\]
1969, Lorenz: the range of predictability is ~2 weeks

‘.. the range of predictability (defined as the time interval within which the errors in prediction do not exceed some pre-chosen magnitude) is about **16.8 days**’

‘.. (there is) **little hope** for those who would extend the two-week goal to one month’
Sensitivities: butterflies and hawkmoths

The ‘butterfly effect’: sensitive dependence to initial condition errors, or in other words a ‘chaotic behaviour’ (Lorenz 1963, JAS)

The ‘hawkmoth effect’: sensitive dependence to model approximations (Frigg et al 2014, POS)

How can we generate skilful fcs taking into account initial and model uncertainties, and the fact that complex models show chaotic behaviours?
From a deterministic to a probabilistic thinking

How can we move forward and go past 2 weeks?

We could **reframe the forecast problem in probabilistic terms.**

ECMWF MONTHLY FORECASTS
FORECAST BASED 22/08/2019 00UTC

NINO3.4 SST anomaly plume
ECMWF forecast from 1 Feb 2018
Monthly mean anomalies relative to NCEP OND 1981-2010 climatology

---

**Sant’Anna**
School of Advanced Studies – Pisa

**Roberto Buizza – ECMWF Annual Seminar 2019**
A probabilistic approach leads to more valuable info

Consider users that need to decide to take an action to protect against a loss. For them, it is important to discriminate between the occurrence and non-occurrence of events.

Ensemble-based probabilistic forecasts discriminate better than single, deterministic ones.
Probabilistic forecasts are more consistent

For an effective management of weather risk, consistency between consecutive forecasts valid for the same verification time.

Ensemble-based, probabilistic forecasts, are more consistent than deterministic fcs.

(Zsoter et al 2009)
Probabilistic forecasts identify predictable events

Furthermore, probabilistic forecasts allow to identify predictable events.
From a deterministic to a probabilistic thinking

How could we estimate and evolve probabilities?
One possibility would be to integrate a Liouville eq.

Consider an N-dimensional system, whose evolution is described by:

\[
\dot{X} = \Phi(X(t), t) \quad X(t = 0) = X_0
\]

The \textbf{Liouville Eq. (LE)} is the continuity eq. for the pdf of the state vector \(X(t)\):

\[
\frac{\partial g(X, t)}{\partial t} + \sum_{k=1}^{N} \frac{\partial}{\partial X_k} \left[ g(X, t)X_k(\dot{X}, t) \right] = 0
\]

The LE is an inhomogeneous quasi-linear (linear in the first derivatives of \(g\)) eq. with dependent variable \(g(X, t)\) and independent variables \((X, t)\). The LE solution depend on the system equations (1) \((\text{Ehrendorfer 1995, MWR})\).
The Liouville equation was integrated for a 3D system

*Ehrendorfer* (1995) applied the LE to simple 3-dimensional system, to compute the evolution of the pdf $\varrho$.

![Graphs showing marginal pdfs for A and G](image)

**Fig. 2.** Marginal pdfs for (a) $A$ and (b) $G$, derived by numerical integration over phase space of the analytical solution (2.14) of the LE for the 3D periodic model (2.1), plotted at certain selected times (see legend). Initial marginals are marked bold. Note that in panel (a) [(b)] the top of the marginal at $t = 50$ [$t = 0$] has been omitted to enhance the vertical resolution.
Could we integrate the LE for a high-dim system?

Ehrendorfer (1995):

- “… the LE is central to the issue of how initial-state uncertainties and model errors affect the skill of numerical weather forecasts ..”

- By considering realistic systems, he concluded that “the high dimensionality of the phase space encountered in the case of realistic meteorological models seems to prohibit this approach”
The process revisited: from obs to fcs via ENS

Ensembles should aim to simulate all sources of errors.

- Observation errors
- ICs’ uncertainties

Model uncertainties
Sensitivity to initial ($\mathbb{I}$) and model ($\mathbb{M}$) uncertainties

Focus in the early 1990s was on estimating initial uncertainties.

An estimate of the relative role of initial and model uncertainties came from Harrison et al (1999), who compared:
- ECMWF-from-ECMWF-ICs [EC(EC)]
- EC(UK)
- UK(UK) and
- UK(EC).

They concluded that up to fc day 5, initial differences dominated.
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
### Simulation of the initial PDF: 3 ‘classes’, 9 methods

#### Three main classes:

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Lagged</td>
<td>Based on the hypothesis that time-lagged analyses have the statistics of analysis errors</td>
<td>- Lagged Average Forecast</td>
</tr>
<tr>
<td>b) Kalman</td>
<td>Inspired by the Kalman Filter</td>
<td>- Ensemble Kalman Filter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Ensemble Transformed Kalman Filter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- ET with Rescaling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Ensemble Data Assimilation</td>
</tr>
<tr>
<td>c) Reduced sampling</td>
<td>Inspired by the analysis cycle and trying to identify leading error-growth directions</td>
<td>- Bred vectors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Singular vectors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- EOF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- STOCH</td>
</tr>
</tbody>
</table>

---

Sant’Anna
School of Advanced Studies – Pisa

Roberto Buizza – ECMWF Annual Seminar 2019
## Simulation of the initial PDF: 9 methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Driving idea</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAF</td>
<td>Lagged Average Fcs: Differences between analyses approximate analysis errors</td>
</tr>
<tr>
<td>EnKF</td>
<td>Ensemble Kalman Filter: An approximation of the Kalman Filter that increases accuracy as ensemble size increases; can take model uncertainties into account</td>
</tr>
<tr>
<td>ETKF</td>
<td>Ens. Transformed KF: BVs transformed using ETKF ideas</td>
</tr>
<tr>
<td>ETR</td>
<td>ET with Rescaling: An extension of breeding, via ETKF plus rescaling</td>
</tr>
<tr>
<td>EDA</td>
<td>Ensemble of Data Assimilation: Uses ideas from the EnKF, but with each ensemble member being generated by an independent 4D-Var</td>
</tr>
<tr>
<td>BV</td>
<td>Bred Vectors: Mimic the analysis cycle</td>
</tr>
<tr>
<td>SV</td>
<td>Singular Vectors: Assumes that the analysis error components fastest growth over a finite time interval are the most relevant</td>
</tr>
<tr>
<td>EOF</td>
<td>Emp. Orth. Functions: BVs transformed using an EOF method</td>
</tr>
<tr>
<td>STOC</td>
<td>Stochastic Scheme: Initial perturbations are generated using a stochastic scheme</td>
</tr>
</tbody>
</table>
An example of initial spread (10 Jan 2013)
Does it matter how the initial PDF is estimated?

It matters, especially in the short forecast range (say day 0-3), up to the time when the role of the initial uncertainties is dominant.

RMSE(ens-mean)
STD

(Buizza, Houtekamer, Toth, …, MWR 2005)
How can we deal with the long-range underdispersion?

**Initial perturbations alone proved not enough to generate reliable ensembles. Independently on the method, all ensembles were under-dispersive.**

This is when people started testing methods that would simulate model uncertainties.

(Buizza, Houtekamer, Toth, …, MWR 2005)
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
How can we simulate the effect of model approx?

\[ \frac{dv}{dt} = -2\Omega x v - \frac{1}{\varrho} \nabla p + g + P_v \]

\[ \frac{dT}{dt} = \frac{RT\omega}{c_p p_s \sigma} + P_T \]

\[ \frac{dq}{dt} = p_q \]

\[ \frac{dp_s}{dt} = p_s \left( \nabla \cdot v + \frac{d}{d\sigma} \frac{d}{dt} \right) \]

\[ \frac{d\phi}{d\sigma} = -\frac{RT}{\sigma} \]
And how can we account for the unresolved scales?

(from Hoskins 2012, QJRMS)
Stochastic schemes to simulate model error

The idea was to include perturbation terms in the r.h.s. of the model equations:

\[ e_j(d,T) = e_j(d,0) + \int_0^T [A(e_j,t) + P(e_j,t) + \delta P_j(e_j,t)] dt \]

\[ \delta P_j(\lambda, \varphi, p) = r_j(\lambda, \varphi)P_j(\lambda, \varphi, p) + F_\Psi(\lambda, \varphi, p) \]

SPPT: Stochastically Perturbed Parameterized Tendencies
(to represent uncertainty associated with parameterisations)

SKEB: Stochastic Kinetic Energy Backscatter
(to represent unresolved upscale energy transfer)
SPPT schemes can act on different scales

Perturbations with different spatial and temporal correlations can be injected in the model. They can be designed to simulate the effect of coherent errors linked to processes with different characteristic scales.

(M Leutbecher)
Backscatter schemes

These schemes simulate the lost upscale energy cascade by estimating the numerically dissipated energy and projecting it back onto the larger scales as a stochastic forcing (Berner et al. 2009; Shutts, 2005):

\[
F_\Psi(\lambda, \mu, z, t) = \left( b_R \overline{D_{tot}}(\lambda, \mu, z, t) \right)^{1/2} F^*_\Psi(\lambda, \mu, z, t)
\]

- **“Total dissipation rate”** estimates energy lost due to numerical methods (explicit diffusion and semi-Lagrangian advection) and kinetic energy production due to sub-grid deep convection.
- **“Pattern generator”** evolves 3D random fields with an AR-1 process
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
We said that a way to move forward, to manage the sensitivities and go past the 2 weeks thought to be the predictability limit, was to reframe the forecast problem in probabilistic terms. We said that ensemble methods could be used to estimate and evolve probabilities.

Do we have evidence that using ensemble-based probabilistic forecasts we can now predict events with certain scales (which ones?) past 2 weeks?
‘Predictability in the midst of chaos’ (Shukla, 1998)

‘… certain aspects of the climate system have far more predictability than was previously recognized. …’

‘… it should be possible to predict the large-scale tropical circulation and rainfall for as long as the ocean temperature can be predicted. If changes in tropical Pacific sea-surface temperature are quite large, the extratropical circulation over some regions should be predictable.’
Large scales: ‘noise and music’ (Hoskins 2013)

‘... despite the prevalence of chaos and turbulence, the optimistic notion has been developed that there could be predictive power on all time-scales …’

‘... On all scales, there are phenomena and external conditions that may give predictability …’
Scales’ interactions, errors and signals propagation

- Tropical convection influences organized convection (e.g. the Madden-Julian Oscillation, MJO), and the MJO propagates and interacts with El Nino.
- The MJO can affect extra-tropical, low-frequency phenomena such as blocking, and blocking can influence synoptic scales, fronts.

\[(X_S, T_S) \quad (X, T) \quad (X_L, T_L)\]

- XTR fronts
- XTR blockings
- TR convection
- TR organized convection
- MJO
- El Nino Southern Oscillation
- Solar radiation
Let's investigate the scale-dependency of Forecast Skill

Consider increasingly coarser fields:

- Spatially: spectrally truncated from T120 (180km) to T60 (360km), T15, T7, T3
- Temporally: from $2\Delta t$ (40 minutes) to 1, 2, 4 and 8 day averages

![2Δt - T120 (180km)](image1)

![2Δt - T7 (3000km)](image2)

![8d - T120 (180km)](image3)
Consider ENS fcs and measure skill with CRPS

ENS forecast probabilities are compared with observations (a very narrow Gaussian). A climatological distribution is used as a reference fc.

- Accuracy is measured using the Continuous Ranked Probability Score
- A forecast is skilful if CRPS(ENS fc) < CRPS(climatological ensemble)
The skill horizon for Z500 over NH is at ~22 days.

Results indicate that for grid-point forecasts of Z500 over NH, the forecast skill horizon is about 22 days.
Forecast skill is scale-dependent

The skill horizon is even longer for larger-scale, lower-frequency phenomena.
Weekly-average large scale anomalies (July 2019)

22/07-28/07 2019
Weekly-average large scale anomalies (July 2019)

Ensemble forecasts issued up to 4 weeks before the event, of weekly-average 2-meter temperature forecasts for the week 22/07 al 28/07 2019.

+2 weeks

+3 weeks
Seasonal SST anomalies in El Nino area (2019)
Seasonal SST anomalies in El Nino area (2019)

Example of a 1-year ensemble forecast issued on 1 May 2018, of sea surface temperature in the tropical Pacific, in the El Nino 3.4 area.
Predictability depends on the scale of the phenomena
How did we get here? Predictable signals versus errors

Predictable signals propagate from the better-initialized and more predictable scales (‘mainly’ the large scales, the slowly evolving components) to the less predictable (small/fast) scales.

Errors propagate from poorly initialized scales (‘mainly’ the smaller scales) thus reducing the predictive skill.

(R Buizza and M Leutbecher, QJRMS 2015)
How did we get here? Impact of IC and/or model improvements

Better initialization, i.e. smaller analysis’ errors, leads to improved forecasts

Model improvements, inclusion in the models of relevant processes (properly initialized) and of model error schemes, lead to improved forecasts
Ex. 1: the role of the 3D-ocean …

The coupling of the atmosphere to the NEMO 3-dimensional ocean led to better MJO fcs (results based on 80 ENS, starting 1st F/M/A/N 1989-2008).

MJO Bivariate Correlation

- Observed SSTs
- Persisted SSTs
- Coupled to NEMO

(F Vitart)
Ex. 1: the role of the 3D-ocean if properly initialized …

Clearly, this required a proper ocean initialisation

(F Vitart, M Alonso-Balmaseda)
Ex 2: improved physics led to more realistic TP, …

PDF of daily (24-hour accumulated) precipitation

Progression in tropical precipitation modelling as a result of convection improvements (entrainment / detrainment and closure formulations)
… which led to more realistic MJO fcs, …

Progression in MJO modelling as a result of convection improvements (entrainment / detrainment and closure formulations).
... and to higher skill with an active MJO in the ICs

The skill of d19-25 PR(2mT>Upp3) forecasts is higher if there is an active MJO in the ICs (results based on 45 cases, 1989-2008).

\[
\text{MJO in ICs BSS(NH)=0.04} \\
\text{NO MJO in ICs BSS(NH)=-0.06}
\]

\[
\text{MJO in ICs BSS(EU)=0.03} \\
\text{NO MJO in ICs BSS(EU)=-0.09}
\]

(F Vitart)
Outline

1. The process leading to sub-seasonal/seasonal prediction
2. Sensitivities and the value of an ensemble approach
3. The estimation of the initial PDF
4. Physical processes and the estimation of model uncertainties
5. Error growth, scales’ interactions and predictability
6. Conclusions
Conclusions …

I hope I have given you evidence that we have managed to tame the butterfly and hawkmoth effects, by improving models, better initializing them, moving to a probabilistic approach, and developing reliable ensembles.
Conclusions and few open questions

1. **PDF evolution**: are ensemble methods the best way to estimate the probability density function of forecast states?

2. **Initial conditions**: as we move towards more complex Earth-system models, which is the best way to initialize Earth-system ensembles?

3. **Model error and sub-grid scale processes**: which is the best way to simulate them? Do we know the model error statistics?

4. **Scales’ interactions and predictability**: Why is the skill of seasonal forecasts over Europe so low?

*Thank you for your attention…*