The Forecast Skill Horizon

Roberto Buizza, Martin Leutbecher and Martin Janousek

ECMWF, Shinfield Park, RG2 9AX, Reading, UK
Outline

1. The two key questions that we will discuss
2. The data, the metric and the Forecast Skill Horizon diagram
3. Scale dependency of the skill horizon
4. Can we expect to further extend the forecast skill horizon in the future?
5. Conclusions
The two key questions we will discuss

1. Does the forecast skill horizon depend on the spatio-temporal scale of the event we are aiming to predict?

Local, instantaneous wind-speed

Weekly-mean, regional temperature anomaly

Monthly-mean, continental-scale rain anomaly

<table>
<thead>
<tr>
<th>Scale</th>
<th>Distance (km)</th>
<th>Time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local, instantaneous wind-speed</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Weekly-mean, regional temperature anomaly</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Monthly-mean, continental-scale rain anomaly</td>
<td>1000</td>
<td>10000</td>
</tr>
</tbody>
</table>

0.1 10 1 10 100 1000 10000 km

0.1 1 10 100 1000 days
The two key questions we will discuss

2. Considering grid-point values, how far is the forecast skill horizon, and can we expect to further extend it?

‘.. the range of predictability is about 16.8 days ..’
‘.. (there is) little hope for those who would extend the two-week goal to one month ..’
Outline

1. The two key questions that we will discussed
2. The data, the metric and the Forecast Skill Horizon diagram
3. Scale dependency of the skill horizon
4. Can we expect to further extend the forecast skill horizon in the future?
5. Conclusions
The forecasts: ensemble-based probabilistic

- Forecasts are ensemble-based, probabilistic to take into account observation (IC, DA, ..) and model uncertainties

- Ensemble-based probabilistic forecasts are more accurate and reliable than single-based forecasts, and are more consistent (i.e. consecutive forecasts jump less)

- All forecasts represent average values over a 4-dimensional space-time volume: even an instantaneous, local value represents an implicit spatial and temporal average
The data and the metric

- Forecasts are from the operational ECMWF medium-range ensemble (ENS)
- Accuracy is measured using the Continuous Ranked Probability Score against analyses
- Skill is measured by comparing the CRPS of ENS and of a climatological ensemble
Buizza & Leutbecher (QJRMS, 2015; BL15) showed that for the instantaneous and local (grid-point) probabilistic prediction of Z500 over NH, the FSH is about 22 days.
Outline

1. The two key questions that we will discuss
2. The data, the metric and the Forecast Skill Horizon diagram
3. Scale dependency of the skill horizon
4. Can we expect to further extend the forecast skill horizon in the future?
5. Conclusions
‘Predictability in the midst of chaos’ (Shukla, 1998)

‘… It is now clear that certain aspects of the climate system have far more predictability than was previously recognized. It also should be recognized that some aspects of the climate system will always be difficult to predict.’

‘… 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, even the extratropical circulation over some regions, especially over the Pacific–North American sector, is predictable.’

Predictability in the Midst of Chaos: A Scientific Basis for Climate Forecasting
J. Shukla, et al.
Science 282, 728 (1998):
DOI: 10.1126/science.282.5389.728
‘Noise and music’ (Hoskins 2013)

‘... despite the prevalence of chaos and turbulence in weather and climate, 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.'
Scale-dependency of the FSH

We have considered 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
Scale-dependency of the FSH

We have considered 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
Scale-dependency of the FSH

BL15 evaluated the skill of ENS fcs for 1 year and showed that the FSH is longer for large-scale, low-frequency fields.

`z500hPa, Northern Extra-tropics`

- Continuous Ranked Probability Score
- 2012070200-2013070800 (107)

2Δt average
4-day average (d-2;d+2)
8-day average (d-4,d+4)
A schematic of the scale dependency of the FSH

- ENS grid spacing
- 1000 days
- 1 year
- 1 season
- 100 days
- 1 month
- 10 days
- 1 day

Skill range:
- Unresolved
- No skill

Spatial scale (km):
- From finer to larger scales

Spatial scale (km):
- 1
- 10
- 100
- 1000
- 10000

- Monthly average sea-surface temperature (e.g. El Niño)
- Monthly average 2-metre temperature, mean sea level pressure
- Teleconnection indices
- Upper-level fields
- Surface fields
- Rainfall
- Extremes
Outline

1. The two key questions that we will discuss
2. The data, the metric and the Forecast Skill Horizon diagram
3. Scale dependency of the skill horizon
4. Can we expect to further extend the forecast skill horizon in the future?
5. Conclusions
The continuous improvement of ENS fcs of ~2 days/decade

500hPa geopotential
Continuous ranked probability skill score
NHem Extratropics (lat 20.0 to 90.0, lon -180.0 to 180.0)

+3d/22y
+5d/22y
The error growth equation

To investigate predictability and the evolution of the forecast skill we will use the error growth equation (Lorenz 1982, Dalcher and Kalnay 1987, Simmons et al 1995, Buizza 2010):

\[
\frac{dE}{dt} = (a \cdot E + S)(1 - \frac{E}{E_\infty})
\]

- E: average error
- a: rate of growth of forecast error
- S: model deficiency
- \(E_\infty\): asymptotic level
Error growth equation: analytical solution

\[
\frac{dE}{dt} = (a \cdot E + S)(1 - \frac{E}{E_\infty})
\]

- \(E\): average error
- \(a\): rate of growth of forecast error
- \(S\): model deficiency
- \(E_0\): initial-time error
- \(E_\infty\): asymptotic level

\[
\eta(t) \equiv \frac{E(t)}{E_\infty} = 1 - \frac{1 + \frac{S}{a \cdot E_\infty}}{1 + C_1 \cdot e^{C_2 \cdot t}}
\]

\[
C_1 = \frac{E_0 + \frac{S}{a}}{E_\infty - E_0}
\]

\[
C_2 = a + \frac{S}{E_\infty}
\]
The error growth eq. parametrizes well the CRPS evolution

The error-growth equation describes in general very well the time evolution of the CRPS of ENS probabilistic forecasts for Z500 over NH and SH winters and summers.

For all years from 1995 to 2017, every 2 years, coefficients have been estimated using best linear fit between d1 and d10.
How do the error-growth coefficients evolve?

Results indicate that:

- $E(0)$ indicates that the initial-time error has been decreasing
- $a$ indicates that in the short-term, the error has been growing faster
- $S$ indicates that the model-error term has been decreasing
Can we gauge the potential gains linked to ICs?
Can we expect to extend further the forecast skill horizon?

We can extrapolate to 2030 (e.g. for Z500 over NH):

- **E(0):** assume it will decrease as between $<01-05>$ and $<13-17>$ ($1.2>0.55>..>0.25$)
- **a:** assume it will increase as between $<01-05>$ and $<13-17>$ ($0.33>0.37>..>0.42$)
- **S:** assume it will decrease as between $<01-05>$ and $<13-17>$ ($2.7>1.3>..>0.6$)
Can we expect to extend further the forecast skill horizon?

**summer**

CRPS normalized - NH Z500 JJA

- ave(01-03-05)
- ave (13-15-17)
- 2030 estr

**winter**

CRPS normalized - NH Z500 DJF

- ave(01-03-05)
- ave (13-15-17)
- 2030 estr
Can we expect to extend further the forecast skill horizon?

Results based on Z500 over NH indicate that if we are able to continue with the same rate of improvements, we can keep increasing the forecast skill in the short- and medium-range.
Looking ahead, how important is to improve the PDF model?

Estimated gains in CRPS (NH Z500 DJF) due to IC and MOD improvements

Estimated gains in CRPS (NH Z500 JJA) due to IC and MOD improvements

(model+ENS config)
Outline

1. The two key questions that we will discuss
2. The data, the metric and the Forecast Skill Horizon diagram
3. Scale dependency of the skill horizon
4. Can we expect to further extend the forecast skill horizon in the future?
5. Conclusions
How did we manage to extend the FSH beyond 2 weeks?

By improving our models and initialisation schemes so that all scales relevant to predict phenomena with a scale \((X,T)\) are included, ‘well’ simulated and initialized. The forecast skill horizon is determined by the competition between:

- Errors propagating from the poorly initialized scales (mainly smaller/faster?), and
- Predictive signal propagating from the better initialized scales (mainly wider/slower?)

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

\(\text{slave} \leftrightarrow \text{Phenomena} \leftrightarrow \text{wider longer-time}\)

\(\text{free} \rightarrow \text{External forcing} \rightarrow \text{wider longer-time}\)

(Hoskins 2012, QJRMS)
How did we manage to extend the FSH beyond 2 weeks?

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.

(Buizza and Leutbecher 2015, QJRMS)
Changes in convection improves tropical precipitation ...

PDF of daily (24-hour accumulated) precipitation

Progression in tropical precipitation modelling as a result of convection improvements (entrainment / detrainment and closure formulations)

(From Peter Bechtold)
... and allows a more realistic MJO propagation ...

Progression in MJO modelling as a result of convection improvements (entrainment / detrainment and closure formulations).

(From Frederic Vitart)
… leading to a more realistic and skilful MJO fcs up to w3

Considering the MJO, the ECMWF system is capable to predict it up to about 25 days.

(From B. Wang and J.-Y. Lee)
**Better MJO prediction > higher skill over Europe**

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).

(From Frederic Vitart)
How did we extend the FSH beyond 2 weeks? 3D-ocean ...

For example, the coupling of the atmosphere to a 3-dimensional ocean model has led to ENS improvements (results are based on 80 ENS, starting on 1st Feb/May/Aug/Nov 1989-2008).

MJO Bivariate Correlation

Observed SSTs

Persisted SSTs

Coupled to NEMO

(F Vitart)
How did we extend the FSH beyond 2 weeks? ICs …

A proper initialisation played a key role.

(F Vitart, M Alonso-Balmaseda)
Is there evidence that we can improve further? Sea-ice..

Ruggieri et al (2016, JGR) have shown that sea-ice concentration anomalies in the Barents-Kara seas in late autumn/early winter has an impact on weather variability over the Euro-Atlantic sector in late winter.
Is there evidence that we can improve further? Aerosols ..

Preliminary results show a bias reduction from using an interactive aerosols on meteorological fields (winds and precipitation). More prominent impact can be detected over the Indian Ocean.

(A Benedetti, F Vitart)
Conclusions: Q1

1. Does the forecast skill horizon depend on the spatio-temporal scale of the event we are aiming to predict?

Yes, there is a clear evidence of this, e.g. from ECMWF ENS monthly forecasts and from results published in the literature.
Conclusions: Q2

2. Considering grid-point values, how far is the forecast skill horizon, and can we expect to further extend it?

Yes, if we manage to keep improving the model and the estimation of the initial states, we should be able to extend the forecast skill horizon by **0.6-1.4 days in the next decade**.
Thank you for your attention ...
Extra slides ...
The ECMWF suites (Sept 2017)

- Atmosphere grids: \(T_{CO}\) (cubic octahedral Gaussian reduced grid) or \(T_L\) (Gaussian linear grid)
- Ocean grid: ORCA (tri-polar grid)

**Model components**
- atmosphere
- land
- waves
- sea-ice
- 3D-ocean

**Initial conditions**
- EDA-25 4DV
  - 18km - TCo639L137
- HRES 4DV
  - 9km – TCo1279L137
- ORAS5-5 3DV-FGAT
  - 0.25 degrees z75 – ORCA

**Forecasts**
- HRES FC
  - 9km – TCo1279L137 (d0-10)
- ENS-51 FC
  - 18km – TCo639L91 (d0-15)
  - 36km – TCo319L91 (d15-46)
- SEAS5-51 FC
  - 36km – TCo319L91 (m0-7)

**Model components**
- atmosphere
- land
- waves
- sea-ice
- 3D-ocean
# The ECMWF suites (Sept 2017)

<table>
<thead>
<tr>
<th>Operational suite</th>
<th>Uncertainty sources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
</tr>
<tr>
<td>HRES - 9km</td>
<td>T(_{\text{CO}})1279 L137 (0-10d)</td>
</tr>
<tr>
<td>4DVAR- 9km</td>
<td>T(_{\text{CO}})1279 (inner T(_L)255/319/399) L137 (t0)</td>
</tr>
<tr>
<td>EDA – 18km</td>
<td>25 members: T(_{\text{CO}})639 L137 (t0)</td>
</tr>
<tr>
<td>ENS – 18km</td>
<td>51 members: T(_{\text{CO}})639 L91 (0-15d)</td>
</tr>
<tr>
<td></td>
<td>T(_{\text{CO}})319 (~36km) L91 (15-46d)</td>
</tr>
<tr>
<td></td>
<td>- 3D-ocean: NEMO ORCA25z75</td>
</tr>
<tr>
<td>SEAS5 – 36km</td>
<td>51 members: T(_{\text{CO}})319L91 (1-7/13m)</td>
</tr>
<tr>
<td></td>
<td>- 3D-ocean: NEMO ORCA25z75</td>
</tr>
</tbody>
</table>

\(T\(_{\text{CO}}\)\) – Cubic octahedral Gaussian reduced grid  
\(T\(_L\)\) – Gaussian linear grid
The quality of our forecasts: HRES 500 hPa geop height

HRES headline score: anomaly correlation for the 500 hPa height over the northern hemisphere.

Predictability gains of ~ 0.8 day/decade

2007-2017: +0.8d HRES Z500 NH
The quality of our forecasts: ENS Temp at 850 hPa

Results indicate that the reliability and accuracy of the ECMWF medium-range ensemble (ENS) has been increasing continuously, with gains of 1.5-2.0 days/decade.

850hPa temperature
Lead time of Continuous ranked probability skill score reaching 25%
NHem Extratropics (lat 20.0 to 90.0, lon -180.0 to 180.0)

2007-2017:
+0.8d HRES Z500 NH
+1.5d ENS T850 NH

Predictability gains of ~ 1.5 day/decade
The quality of our forecasts: HRES precipitation forecast

Supplementary score for HRES forecasts: 1-SEEPS (Stable Equitable Error in Probability Space) score for 24-h precipitation totals in the northern extra-tropics (verified against synop observations).

total precipitation
1-SEEPS
Extratropics (lat -90 to 30.0 and 30.0 to 90, lon -180.0 to 180.0)

2007-2017:
+0.8d HRES Z500 NH
+1.5d ENS T850 NH
+0.9d HRES TP NH

Predictability gains of ~ 0.9 days/decade
The quality of our forecasts: ENS precipitation forecasts

Supplementary score for ENS forecasts: Continuous Ranked probability Skill Score (CRPSS) for 24-h precipitation totals in the northern extra-tropics (verified against synop observations).

2007-2017:
+0.8d HRES Z500 NH
+1.5d ENS T850 NH
+0.9d HRES TP NH
+1.7d ENS TP NH
We have been improving also the prediction of surface weather

Improvements can be detected not only in the upper levels but also in surface weather variables such as cloud cover, wind and 2-m temperature.

Improvements are due to a mix of upgrades in model, data assimilation, use of observations and resolution.

Everything contributes!!
We have been improving also the prediction of surface weather.

Improvements can be detected not only in the upper levels but also in surface weather variables such as cloud cover, wind and 2-m temperature.

Improvements are due to a mix of upgrades in model, data assimilation, use of observations and resolution.

Everything contributes!!
We have been improving also the prediction of surface weather

Improvements can be detected not only in the upper levels but also in surface weather variables such as cloud cover, wind and 2-m temperature.

Improvements are due to a mix of upgrades in model, data assimilation, use of observations and resolution.

Everything contributes!!
We have been improving also the prediction of surface weather

Improvements can be detected not only in the upper levels but also in surface weather variables such as cloud cover, wind and 2-m temperature.

Improvements are due to a mix of upgrades in model, data assimilation, use of observations and resolution.

Everything contributes!!
Since 2002 we have also been producing monthly forecasts

Comparison of the forecast lead-time (in days) when the prediction of the MJO reaches 0.6 correlation (orange) and 0.5 correlation (yellow).

The data are from the Sub-seasonal to Seasonal (S2S) WWRP/WCRP WMO project.
Prognostic sea ice fraction

41R2: sea ice fraction fixed - 43R1: sea ice fraction predicted with NEMO/LIM2

The figures show the Reduction in RMSE of ensemble-mean forecast of sea ice fraction for 10-day forecast.
ENS MJO

- MJO: Nice improvement in MJO skill scores
- Correlations are higher (for both PC1 and PC2, but mostly PC1)
- RMS error is also reduced
- MJO spread is improved which makes it closer to RMS error
Aerosol impacts on monthly forecasts (summer)

**Caveat:** May start dates only 2003-2015 period

Observed (prescribed) Fire emission

(A Benedetti, F Vitart)