Revisiting the synoptic-scale predictability of severe European winter storms using ECMWF ensemble reforecasts

Florian Pantillon, Peter Knippertz and Ulrich Corsmeier
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
Collaborative Research Center "Waves to Weather"

Upscale error growth

Cloud-scale uncertainty

Predictability of local weather

Revisiting the synoptic-scale predictability of severe European winter storms

Florian Pantillon, Peter Knippertz and Ulrich Corsmeier
Motivation and strategy

Predictability of wind gusts in winter storms over central Europe

- Storms = destructive natural hazard
- Predictability = Multi-scale problem

- Synoptic scale → global ensemble forecasts
- Mesoscale → regional ensemble forecasts
- Turbulent scale → Doppler wind lidar observations

Synoptic scale
$O(1000 \text{ km})$

Mesoscale
$O(10-100 \text{ km})$

Turbulent scale
$O(0.1-1 \text{ km})$
Synoptic scale: model data

ECMWF ensemble reforecast

- Retrospective forecast: 20 years with homogeneous model version
- dx=30 km, 10 days, 10+1 members, no stochastic physics, 2 runs/week

Selection of storms: XWS open access catalogue (Roberts et al. 2013, NHESS)
- 52 most severe European storms 1979-2013
- Available online http://www.europeanwindstorms.org

→ 25 storms (1995-2015) x 3 forecasts/storm x 11 members/forecast
→ Comparison ERA-Interim reanalysis (retrospective analysis) dx=80km
Three metrics to assess predictability

1. Track and intensity → storm dynamics
2. Strength of wind gusts → storm impact
3. Area covered by gusts → storm warnings
Dynamics: track and intensity

1. Tracking: algorithm based on Laplacian MSLP (Pinto et al. 2005, MetZ)

2. Identification: two methods = first occurrence or maximum intensity
   \[ \rightarrow \text{divergences for lead times beyond 3 days} \]

Identified tracks of ex-hurricane Lili in the 6-day ensemble reforecast init on 22 October 1996
Results for the ensemble average

Difference reforecast – ERA-Interim
- Bias - in longitude (too slow) > day 3
- Bias + in MSLP (too weak) > day 4
→ For severe storms reaching Europe

Large variability between storms
- Strongest bias for storm Gero (2005) = explosive cyclogenesis to 948 hPa

symbol = median per storm
black curve = median per lead time
Results for individual members

Number of members with *actual* storm

- Most members until day 2-4
  → *Average not clearly defined* beyond

- At least 1 member until day 8-10
  → *Potential for early warning*
  → !!but focus on observed events (hits) without accounting for false alarms!!

![Diagram showing results for individual members](image)

1 symbol = 1 storm
black curve = median per lead time

Figure 6. Position and intensity of the storms in the ensemble reforecast as identified and compared on the day of maximum intensity: number of ensemble members predicting the storm within 10 hPa and 10 great circle (a) or 5 hPa and 5 great circle (b) as compared to ERA-Interim in minimum MSLP and position, respectively. The symbols represent the storms as given in Table 1 and the black curve shows the median of the storms per lead time.
Impact: Storm Severity Index (SSI)

\[ SSI = \left( \frac{v_{\text{max}}}{v_{98}} - 1 \right)^3 \]

(Klawa and Ulbrich 2003, NHESS; Leckebusch et al. 2007, GRL)

- \( v_{\text{max}} \): daily maximum wind gusts
- \( v_{98} \): local 98th climatological percentile (in reforecast or ERA-Interim)

→ Integral over central Europe = measure of storm severity

(a) Maximum wind gusts

(b) Storm Severity Index

Daily wind gusts and SSI for storm Lothar on 26 December 1999 in ERA-Interim
Results for the storms

Intense/extreme events
- overestimation by factor ~2
- stable with lead time

Storms
- drop order of magnitude by day 4
- large variability between storms

→ Predictability impact restricted to first 3 days
**Warnings: Extreme Forecast Index**

**Motivation** predicted < observed extremes

**Idea** measure extremes in model world (Lalaurette 2003, QJRMS; Zsoter et al. 2006)

Extreme Forecast Index (EFI)
- Uses distribution of ensemble forecast
- Gives deviation from model climate
- 0 = model climate  +/-1 = extreme

!!!many hits but also false alarms!!!
→ look for optimal threshold in EFI
→ trade-off with Heidke Skill Score (Petroliagis & Pinson 2014; Boisserie et al. 2016)

6-day reforecast of storm Lothar and analysis on 26 December 1999
Results for strong gusts

EFI to predict gusts > 98\textsuperscript{th} clim. percentile

- **Whole dataset**: skill until day 10
- **Storms**: higher skill
  \(\rightarrow\) bias when focus obs events only

Large variability between storms

- Lowest skill for **Yuma** (1997) = smallest storm in sample!
- High skill at day 10 for **Xynthia** (2010) = favourable environment?

(a) Extreme Forecast Index

![Figure 9](image)

(b) Shift of Tails

![Figure 10](image)

1 symbol = 1 storm
black curve = median per lead time
dotted curve = whole 20-year dataset

---

Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 31 March 2017

© Author(s) 2017. CC-BY 3.0 License.
Summary

Synoptic-scale predictability of severe European winter storms

- ECMWF ensemble reforecast for 25 severe European storms 1995-2015
- 3 metrics

Low predictability for cases of ET, explosive intensification, small storm

No systematic link with dynamics → larger dataset? case studies?