Nat. Hazards Earth Syst. Sci. Discuss., 3, C1154–C1165, 2015
www.nat-hazards-earth-syst-sci-discuss.net/3/C1154/2015/

Natural Hazards and Earth System Sciences
Open Access
Discussions

Interactive comment on "An approach to build an event set of European wind storms based on ECMWF EPS" *by* R. Osinski et al.

R. Osinski et al.

robert.osinski@met.fu-berlin.de

Received and published: 29 June 2015

Major corrections:

Several different configurations of the EPS were operational during the period chosen for analysis with resolutions varying from $T63$ to T_L 399. Some of the plots/tables exclude data from the $T63$ configuration without explanation (Table 2, Fig 3). Why is this?

As the available space for the paper is limited, we decided to present only parts of the results. The results of Table 2 and Figure 3, which you have mentioned, are added to this answer, see Figure 1 and Table 1 below. They will also be added to the manuscript.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

$T63$ (12 hourly)	50.9 (54.0)	0.70 (0.65)	41.2 (43.8)
T_L159 (12 hourly)	45.9 (49.3)	0.75 (0.79)	49.2 (50.4)
T_L255 (6 hourly)	47.6 (45.0)	0.71 (0.76)	41.4 (42.0)
T_L399 (6 hourly)	47.9 (50.5)	0.74 (0.74)	42.0 (42.6)

Table 1. (Table 2 revised) Average storm properties of EPS (ERA-Interim).

The ratio between the size of the ensemble database and ERA-Interim database is mentioned (in passing) for the first time on p1244 (300 times as the EPS has 50 members and runs for 6 days). Does this mean that at any given time, forecasts with lead times out to 6 days are taken that verify at that time and the analysis time is ignored. But if so this wouldn't work with tracked storms - don't you just have the 50 ensemble forecasts starting at any given time to compare to the ERA-Interim analysis/forecast sequence starting at the same time? Also, given that the ensemble is initialized twice daily since Nov 2000 shouldn't the ratio be greater than this or is only one of the initialization times considered? The generation of the ensemble storms dataset needs to be more clearly specified in section 2.2 (or an additional section added such that section 2.2 contains the specification of the ECMWF EPS and the new section the use of this EPS dataset to create the storms dataset).

The tracking algorithm is applied on each single EPS forecast, thus up to twice 50 members per day. As described in section 4.3 from lines 18 on page 1240, the first ten days of each forecast were used in this study because of reduced horizontal resolution after this period. Including storms not completely within a forecast run would lead to a biased sample, in particular an overrepresentation of storms with (seemingly) short duration. Allowing the inclusion of storms first appearing towards the end of the 10 day period would produce an increased likelihood of storms existing beyond the end of the 10 day forecast period. Again, this would lead to an incorrect sample of storms for statistical purposes. For this reason, we use the six days window, finding that the amount of incomplete storms in the sample is in the order of some per mill, which we find acceptable.

As the storm frequency in the EPS is comparable to the one from ERA-Interim, we have a six days times fifty member, meaning 300 times larger sample compared to ERA-Interim.

This estimation disregards the existence of 00UTC EPS initializations in addition to 12UTC initializations since March 2003. This information will be added to table 1. We wanted to be sure to avoid an overrepresentation of the period 2003 to 2010 in the statistics and thus use only the 12UTC initializations in section 5.5. For the period with 00UTC initializations available, the sample for this period is indeed 600 times larger than the ERA interim sample. This is now mentioned in the text at the end of section 4.3.

The question of the independence of the storms in the ensemble dataset needs also to be discussed. Obviously for short lead time forecasts the storms will be very similar to the storms in the reanalysis whereas for longer lead times they can be very different (as indicated by the increase in the number of 'pure' storms with lead time in Fig 11). Hence the statistics of the ensemble derived dataset will not be the same as those of a reanalysis dataset with the equivalent number of days. I think this (and the implications) should be discussed in the paper.

The dependence of events in between different member is an important question for the application of the dataset, and this is why we investigate 'pure' EPS storms in section 5.1. We already demonstrate the difference of short and longer lead times with respect to the variability of ensemble members using the storm "Emma" as an example. We also show in this chapter that the statistics of the EPS storms are comparable to ERA-Interim, using the mentioned six days window. The storm frequency keeps constant during the forecasts, as well as the distribution of storm severity, of the storm field size as well as of the storm duration (last 3 not shown). Still, the reviewer's observation is correct that we find more storms rather close to the observed ones than more distant or even “pure” when considering the whole ensemble. The sample is thus influenced by the observed storms, as it is (and should be) influenced by the observed development of the climate system in the period considered. This must be taken into account when using the EPS sample. If independence from observations is a requirement, CGCM runs may be the better choice. We included this now in the discussion of our results and in the conclusion.

It would also be helpful to include in the introduction a clear statement of purpose for the presented research i.e. the aim of the research or questions to be answered.

The purpose of our study is the exploration of the ECMWF EPS for producing an enhanced statistics of storms under observed climate conditions. It shall be shown that the statistics of the EPS storms can be used to extend the observed climatology, and thus arrive at a more reliable statistics. Such climatology is still intended to be close to the observed development of climate conditions, and must be distinguished from alternative approaches such as climate simulations for present day GHG and solar forcing, for example, which allow the models to produce windstorms largely independent from the observed development of weather and climate in the time period considered. Summarizing, our study intends to describe a possibility for producing storm statistics which are still very close to the observed climate.

In the sense of an event set, we do not expect complete independency but just variations of storms, as it is done e.g. for stochastic event sets out of a fixed historical sample. Finally, the way how you select events for the construction of an event set, is a question which pretty much depends on the specific purpose of the event set under consideration and is not further discussed in this paper.

ERA-Interim data: section 2.1 states that 6 hourly reanalysis data is used. I think the reanalysis dataset contains analyses at 0 and 12 UTC, and 6 hourly forecasts out to 36 hours. So, are the analyses and 6 hour forecasts used to comprise the ERA-Interim storms dataset?

The ERA-Interim reanalysis from the ECMWF mars archive at 00, 06, 12 and 18UTC were used. We do not consider the associated forecasts.

Minor corrections:

p1235, L27: I don't understand the sentence 'As ERA-Interim data is only available with 6 hourly resolution, the EPS data with 3 h resolution were used in subsets of 6 hourly resolution again'. What does the 'again' mean?

EPS data are available in 3h, 6h and 12h temporal resolution. In case of the 12h resolution, both datasets, the EPS and ERA-Interim were used at the 00UTC and 12UTC time steps. Data in 6h temporal resolution were used at 00UTC, 06UTC, 12UTC and 18UTC, both EPS and ERA-Interim. The newer EPS data are partly available in 3h resolution. The 'again' should mean that these EPS data were also used in 6h resolution, as the ERA-Interim reanalysis is only available in 6h. Another reason is the fact that not ten days are available in 3h resolution, only up to the first 144h. It may be possible to adapt the tracking algorithm to changing temporal resolutions to detect the storms. A difficulty would remain and is the calculation of the SSI, because this index is integrated over time, so that the values in 3h resolution would be systematically higher than in 6h resolution. A direct comparison with ERA-Interim would be difficult, and a correction is complicated as we are interested in extreme events.

p1237, L16. What is the evidence (e.g. a reference) that difference resolutions of the EPS system produce difference wind speed biases?

A model with a coarser resolution represents an average of a larger area of each individual grid cell than a model with a fine resolution. Orographic effects can be better captured if the resolution of the model is finer. This leads to different wind distributions. These changes can for example be seen in the 98th percentiles which are increasing with finer resolution. In fig. 2 it is shown that the upper tail of the distribution also differs. This has the consequence that SSI values are not directly comparable between the different EPS resolutions. For this reason a quantile-quantile mapping was used.

NHESSD

3, C1154–C1165, 2015

We will add these explanations to the paragraph.

p1238, L9: It would be helpful to add a sentence explaining what a quantile-quantile mapping is.

A quantile-quantile mapping is a standard method for a bias correction and changes a dataset in the way that its distribution is afterwards the same one as the one of a reference dataset. The method chosen is empirical and based on the comparison and adjustment of quantiles. It compares the i-th quantile of two datasets. If the wind speed for the i-th quantile in the EPS is different from the one in ERA-Interim, than the value in the EPS is set to the ERA-Interim value. For further details, see for example Maraun (2013).

p1240, L8: Here a 'notable feature' of storm Emma is described. What interpretation and importance do you attach to this notable feature? Is it 'meaningful' or just a result of random chance.

This example was chosen intentionally to demonstrate that the reality is one particular representation out of plenty of possible realizations. The track of storm Emma lies at the outer end of tracks represented by the EPS. It is an extreme example. For the majority of storms, ERA-Interim lies somewhere close to the middle of the EPS tracks. If the southward shift were systematic, fig. 12 would also show a southward shift of the tracks in the EPS.

The example of storm Emma visualizes that for short lead times, the storm realizations are nearby. Longer lead times show larger differences between the members. This is a general feature in the EPS, but the degree of difference and its temporal evolution will depend on how well the forecast system is able to predict the particular weather situation.

Fig. 7 shows that for a large amount of events, the EPS shows strong variability in the storm severity. This variability is based, like demonstrated for storm Emma, in different

NHESSD

3, C1154–C1165, 2015

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

spatiotemporal extents as well different intensities of the EPS representations, and is a typical feature of the EPS.

p1240, L13: You argue that the increased range of the SSI in the EPS compared to ERA-I is 'partly due to the definition of the SSI, using cubic exceedences'. I don't think this directly follows. Surely the range in the EPS is larger than that in ERA-I simply because the severity of the storms in the EPS can exceed those in ERA-I. The definition of the SSI using cubic exceedences leads to the range in SSI in the EPS being very much larger than that for ERA-I as a result of relatively small increases in windspeed and/or area in the EPS.

The reason for using the EPS for this study is based on the assumption that some of the perturbations of the forecasts should lead to storms with a higher damage potential. This is based on small increases of the wind speeds compared to a historical storm or on larger footprints/durations what would result in larger cumulative damages. Differences in the track result also in differences in the SSI due to different values of the underlying percentile. As has been mentioned by the reviewer, the consequence of the cubic is that the small differences in wind speed can lead to much larger SSI values. A different definition of the SSI (e.g. linear) would also lead to larger SSI values in the EPS, but the tail of the SSI distribution would be shorter, and a small difference in wind speed would in this case lead only to a small difference in SSI. In Klawa and Ulbrich, 2003 it was shown that storm damages can be modeled using the cubic exceedance of the 98th percentile. This study was the motivation to define the SSI (as introduced by Leckebusch et al., 2008) as it is used in the presented paper. The large range in the SSI should for this reason be connected with much higher potential storm damage.

p1241, L20: You say that the SSI values are expected to be lower for data with 12-hourly compared to 6-hourly temporal resolution. Is this simply because the time of maximum severity is more likely to be further away from one of the data times for the 12-hourly data?

NHESSD

3, C1154–C1165, 2015

The SSI is a measure, which is integrated over time. A higher temporal resolution leads to more summands over time t in the SSI formula (eq. 1). For this reason the SSI values for the 6h resolution must be higher, as they are summed over all detected 98th percentile exceedances at 00, 06, 12 and 18 UTC for a particular event. For the 12 hourly data only exceedances at 00 and 12 UTC can be taken into account. For this reason the SSI in 12h resolution is reduced by the contributions at 06UTC and 18UTC. As we are interested in the extreme events, it is hardly possible to interpolate these fields.

There is a clear daily cycle in the wind speed. The wind speed distribution has a longer upper tail at 12UTC than for 00, 06 and 18UTC. The threshold value used for the tracking is based on the entire dataset. The probability of maximum severity is the largest for this reason at 12UTC. This is also the reason for the daily cycle of starting events in Fig. 11.

p1241, L28: Add a reference for the 'Anderson-Darling test'.

Thode H.C. (2002); Testing for Normality, Marcel Dekker, New York.

Table 1: Why are there 2 temporal resolutions available for 2 of the EPS configurations?

The change in the archived temporal resolution took place in between the period given in the table. For this reason in the beginning of the period the coarser resolution is only available. The table will be extended by these informations, see below.

Fig 11: given that ERA-Interim does not include a 192+ h forecast, what does the line for ERA-Interim mean on this plot?

The continuous ERA-Interim reanalysis were used. For this reason values only at 00,06,12 and 18UTC are available. To make it possible to compare the results with the EPS, the results of ERA-Interim were repeated. For clarity again: no forecasts were used.

NHESSD

3, C1154–C1165, 2015

21 Nov 1992–9 Dec 1996	$T63$ 12 (12)	32pf+1cf	12:00UTC
10 Dec 1996–12 Jan 2000	T_L159 12	50pf+1cf	12:00UTC
13 Jan 2000-20 Nov 2000	T_L159 6 (12)	50pf+1cf	12:00UTC
21 Nov 2000–24 Mar 2003	T_L255 6	50pf+1cf	12:00UTC
25 Mar 2003–27 Jun 2005	T_L255 6	50pf+1cf	00:00 and 12:00UTC
28 Jun 2005-31 Jan 2006	T_L255 3h until 126/144h and 6 (6)	50pf+1cf	00:00 and 12:00UTC
1 Feb 2006–25 Jan 2010	T_L399 3h until 144h (6)	50pf+1cf	00:00 and 12:00UTC

Table 2. (Table 1 revised). Overview of general characteristics of the EPS (used temporal resolution) pf: perturbed forecast; cf: control forecast.

Fig 13 caption: please define 'relative grid cell affection (%)'. Is this the % of time a 'pure' storm relative to 'all storms' is seen at this grid point in the EPS? Is this for a specific lead time? In the associated text, p1244 L16, it says that this figure shows that 'Over the Atlantic the number for the "pure" EPS storms is lower than over North Africa and Eastern Europe'. The % pure storms is less over the Atlantic but is the absolute number also less? This section needs some clarifying.

It is the % of all detected events which are classified as 'pure' ones starting during the mentioned six days window. The idea behind the 'pure' storms is to get an additional sample of events, which are independent to the historical events. The absolute number of 'pure' events per year can be seen if combining fig. 12c with fig. 13. Then we have about 1 event over the Atlantic and about 1.5 to 2 events over central Europe. The absolute numbers will be added to the manuscript.

Fig 14 & 15 caption: The x-axes here are labeled duration but I think this is actually forecast time with the different colored lines relating to different storm durations (if I've interpreted the text correctly).

Duration as x-axes label is correct. All storms with same duration from the entire generated EPS sample (restricted to storms starting inside a 6-days window, one initialization per day, and 6-hourly EPS) are taken, regardless the lead time.

Grammatical corrections/ typographical errors:

p1233, L7: 'station-data' should not be hyphenated.
→ Corrected into station data

p1233, L 11: 'quintessence' is a seldom used English word and does not seem to e used in context here -perhaps 'implication' would be more appropriate.
→ is changed into implication

P1233, L16: too many commas and sentence doesn't make sense - rewrite.
→ will be corrected

p1233, L23: 'been become' - remove 'been'
→ changed

p1234, L2: 'was already used' -> 'was used'
→ changed

p1235: L16-18: Sentences describing the evolution of the resolution of the model and the singular vectors need re-writing for clarity. e.g. 'The horizontal resolution was increased from T63 as follows: T_L159 (12.1996), T_L255 (11.2000) ....'
→ changed

p1236, eq 1: Please define all terms in the equation (V_perc, k and A_k are not defined).
→ added

p1237, L8: Sentence beginning 'As the excess...' doesn't quite make sense. What 'can also be affected by the inhomogeneities'?
→ The distributions of wind speed are affected by the inhomogeneities arising from the different model versions of the EPS. The SSI values are based on the exceedance of the 98th percentile. Differences in the tail of the distribution lead to an incomparability of the SSI values from different periods. A quantile-quantile mapping is used to map the EPS wind distribution to the ERA-Interim distribution. The mapped distributions are used to calculate the SSI values using the excess over the 98th ERA-Interim percentile.

NHESSD
3, C1154–C1165, 2015

Interactive Comment

p1238, L24: 'steps then' -> 'steps than'
→ corrected

p1239, L4: 'valid noon' -> 'valid at noon'
→ corrected

p1239, L9: 'well visible' -> 'clearly visible'
→ corrected

p1239, L12: 'without figure' -> 'figure not shown'
→ corrected

p1240, L27: 'forecasted' -> 'forecast'
→ corrected

p1240, L28: 'for a maximum' -> 'as the maximum'?
→ corrected

p1241, L7: 'mentioned earlier' -> please be specific i.e. 'described in section 4.1'.
→ described in section 4.1 and visualized in fig. 3 b and e.

p1241, L8: Here an un-numbered bold heading is given that is virtually identical to the previous heading numbered 4.3. Either this should be a different numbered heading or be removed.
→ This should be a subsection of 4.3. Corrected

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 3, 1231, 2015.

NHESSD
3, C1154–C1165, 2015

Interactive Comment

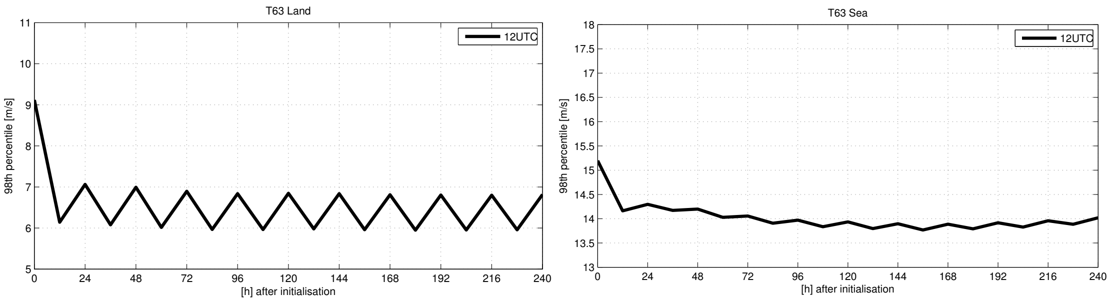

Fig. 1. (complement for Figure 3): Average of 98th percentile [m/s] for different forecast lead times (right axis, [h] after initialization): for T63 12 hourly for land (left) and sea (right) grid boxes