ECMWF extreme forecast index for water vapor transport: a forecast tool for atmospheric rivers and extreme precipitation

Lavers, D. A. ORCID: https://orcid.org/0000-0002-7947-3737, Pappenberger, F. ORCID: https://orcid.org/0000-0003-1766-2898, Richardson, D. S. and Zsoter, E. (2016) ECMWF extreme forecast index for water vapor transport: a forecast tool for atmospheric rivers and extreme precipitation. Geophysical Research Letters, 43 (22). pp. 11852-11858. ISSN 0094-8276 doi: https://doi.org/10.1002/2016GL071320 Available at https://centaur.reading.ac.uk/106884/

AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2016GL071320

Key Points:
- Evaluate the ECMWF Extreme Forecast Index (EFI) for water vapor transport (IVT) across Europe
- IVT EFI is better in capturing extreme rainfall in week 2 forecasts that are initialized in a positive North Atlantic Oscillation (NAO)
- Precipitation EFI is better at capturing extreme rainfall in forecasts initialized in a negative NAO and at shorter lead times

Correspondence to:
D. A. Lavers,
david.lavers@ecmwf.int

Citation:
Lavers, D. A., F. Pappenberger, D. S. Richardson, and E. Zsoter (2016), ECMWF Extreme Forecast Index for water vapor transport: A forecast tool for atmospheric rivers and extreme precipitation, *Geophys. Res. Lett.*, *43*, 11,852–11,858, doi:10.1002/2016GL071320.

Received 23 SEP 2016
Accepted 22 OCT 2016
Accepted article online 26 OCT 2016
Published online 17 NOV 2016

ECMWF Extreme Forecast Index for water vapor transport: A forecast tool for atmospheric rivers and extreme precipitation

David A. Lavers[1], Florian Pappenberger[1], David S. Richardson[1], and Ervin Zsoter[1]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK

Abstract In winter, heavy precipitation and floods along the west coasts of midlatitude continents are largely caused by intense water vapor transport (integrated vapor transport (IVT)) within the atmospheric river of extratropical cyclones. This study builds on previous findings that showed that forecasts of IVT have higher predictability than precipitation, by applying and evaluating the European Centre for Medium-Range Weather Forecasts Extreme Forecast Index (EFI) for IVT in ensemble forecasts during three winters across Europe. We show that the IVT EFI is more able (than the precipitation EFI) to capture extreme precipitation in forecast week 2 during forecasts initialized in a positive North Atlantic Oscillation (NAO) phase; conversely, the precipitation EFI is better during the negative NAO phase and at shorter leads. An IVT EFI example for storm Desmond in December 2015 highlights its potential to identify upcoming hydrometeorological extremes, which may prove useful to the user and forecasting communities.

1. Introduction

Intense horizontal water vapor transport within the atmospheric river of extratropical cyclones is responsible for extreme winter precipitation and flooding along the west coasts of midlatitude landmasses [e.g., *Ralph et al.*, 2006; *Lavers et al.*, 2011]. With this strong connection between vertically integrated horizontal water vapor transport (integrated vapor transport (IVT)) and extreme hydrological events, recent research has assessed the predictability of IVT and precipitation to determine whether there is potential to use IVT over precipitation forecasts as a way to provide earlier awareness of upcoming extreme events. In essence, can the IVT yield a higher estimate of predictability of an extreme event? Results using forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble prediction system showed that the IVT did have higher predictability than precipitation during winter 2013/2014 over Europe [*Lavers et al.*, 2014], and similar findings were reported in the National Centers for Environmental Prediction global ensemble reforecasts over western North America [*Lavers et al.*, 2016]. These studies suggest the potential for using IVT as a forecast diagnostic to provide earlier awareness of hydrometeorological extremes.

In *Lavers et al.* [2014], an example of a possible forecast tool based on IVT was given for the heavy precipitation across northwest Europe on 24 December 2013. The tool employed was the ECMWF Extreme Forecast Index (EFI) [*Lalaurette*, 2003; *Zsoter*, 2006; *Zsoter et al.*, 2014], a product that determines how the ensemble forecasts' probability distribution differs from that of the model climate, and the results showed that the IVT EFI had potential to provide earlier awareness of the extreme on 24 December 2013. The aim of the current study, which builds on this initial EFI finding and the previous predictability assessments, is to apply and verify the EFI for IVT in ECMWF ensemble forecasts during three winters (2013/2014, 2014/2015, and 2015/2016) across western Europe. Through comparison with the EFI for precipitation forecasts, it will be possible to ascertain if there is any usefulness in monitoring the IVT EFI to provide earlier awareness of extreme hydrometeorological events.

2. Data and Methods

2.1. ECMWF Ensemble Forecasts and Reforecasts

The control and 50 perturbed ensemble members (out to forecast day 10) from the ECMWF ensemble prediction system were retrieved from the ECMWF Meteorological Archival and Retrieval System (MARS) for the 00:00 UTC initialization for three extended winter seasons (November to February) of 2013/2014, 2014/2015, and 2015/2016 (361 forecasts). Daily total surface precipitation accumulated at 00:00 UTC was

retrieved, and the specific humidity q and the zonal and meridional (u and v, respectively) winds at 300, 400, 500, 700, 850, 925, and 1000 hPa were retrieved at 00:00 UTC and 12:00 UTC. The fields were analyzed on a regular Gaussian grid corresponding to about 0.28° × 0.28°. The daily averaged (using 00:00 UTC, 12:00 UTC, and 00:00 UTC of the next day) vertically integrated horizontal zonal and meridional water vapor transports were then calculated in an Eulerian framework [e.g., *Neiman et al.*, 2008] and combined into the total water vapor transport (IVT).

At ECMWF, reforecasts are generated to produce a model climate with which to compare the current ensemble forecasts. In the current operational system (as existed in winter 2015/2016), these reforecasts have 11 ensemble members and are run from every Monday and Thursday for the past 20 years (e.g., the reforecasts run from Thursday 3 December 2015 have dates of 3 December 1995, 3 December 1996, up to 3 December 2014). In winter 2015/2016, the model climate was built by using the closest reforecast date (to the forecast of interest) and four dates either side, in turn providing a climate of 1980 members (9 dates × 20 years × 11 members). In the preceding two winters, the reforecasts had five ensemble members run from every Thursday for the past 20 years. In these winters, the model climate composed of the nearest reforecast date (to the forecast of interest) and two either side, yielding a climate of 500 members (5 dates × 20 years × 5 members). For all reforecasts, as with the forecasts, the daily total surface precipitation accumulated at 00:00 UTC was retrieved, and the total water vapor transport (IVT) was calculated.

In this study, we only consider forecasts out to day 10 primarily because of the horizontal resolution degradation that occurred at day 10 in the forecast system that was operational at the time. Future work will consider beyond forecast day 10 as the current operational ECMWF system now has a constant resolution out to forecast day 15.

2.2. ECMWF Extreme Forecast Index (EFI)

The EFI [*Lalaurette*, 2003; *Zsoter*, 2006; *Zsoter et al.*, 2014] was developed to compare the forecasts' probability distribution with that of the model climate (i.e., the reforecasts), thus revealing the extremeness of a forecast.

The EFI is calculated as

$$\mathrm{EFI} = \frac{2}{\pi}\int_0^1 \frac{p - F(p)}{\sqrt{p(1-p)}}\,dp$$

where $F(p)$ is the proportion of ensemble members that lie below the pth percentile of the model climate.

The EFI was calculated for IVT and precipitation on each forecast day (1–10) for all 361 winter forecasts. Values of the index range from −1 to 1, with −1 indicating extremely low and 1 indicating extremely high values with respect to the model climate. A skill assessment was then undertaken on the IVT and precipitation EFI forecasts to see how well they captured extreme precipitation events using all (4314) land grid points (points with ≥ 50% land) across western Europe (70.1°N–35.0°N, 11.0°W–14.9°E).

2.3. Observed Precipitation and Forecast Verification

Daily observed precipitation generated by the European Flood Awareness System (EFAS) [*Bartholmes et al.*, 2009; *Thielen et al.*, 2009] during the three winters was used as the verification data set. Also, an EFAS precipitation climatology using all November–February days from 1990–2014 was built, with extreme precipitation events herein considered as any exceedances over the 99th percentile of climatology. The precipitation was available at a 5 km resolution and was then interpolated to the ensemble grid (0.28° × 0.28°) by using first-order conservative regridding.

To evaluate the skill of the EFI forecasts in capturing extreme precipitation (>99th percentile), we evaluate the probability of detection (or hit rate, H), defined as the number of forecast hits with respect to the total hits and misses, and the probability of false detection (false alarm rate, F), defined as the ratio of false alarms to correct negatives and false alarms. These hit and false alarm rates were assessed in terms of relative operating characteristic (ROC) curves for EFI thresholds ranging from 0 to 1; the ROC area or score was also calculated. The ROC area ranges from 0 to 1, with areas >0.5 representing a skillful forecast. In the assessment herein, we focus on the region 70.1°N–35.0°N, 11.0°W–14.9°E because this part of Europe in winter is most influenced by the strong IVT associated with atmospheric rivers.

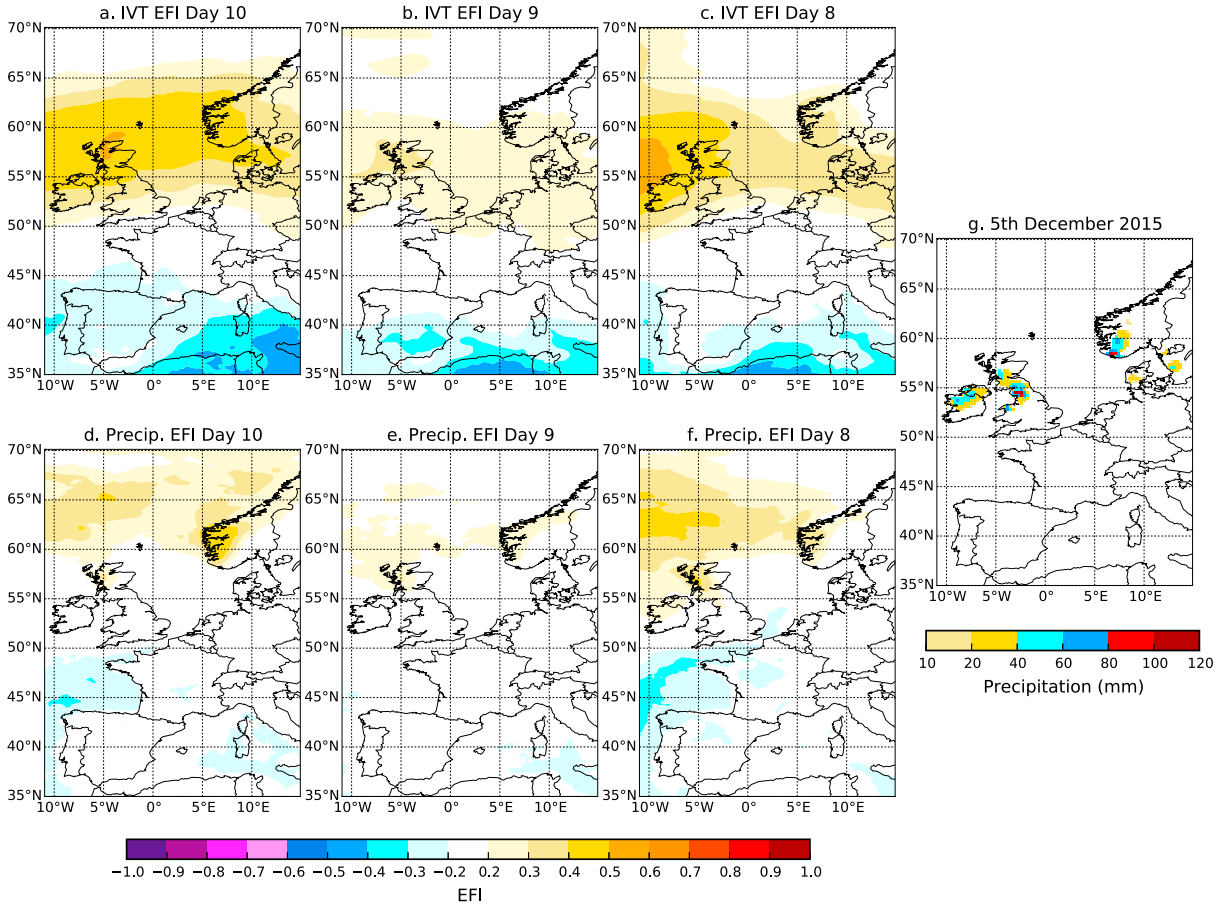

Figure 1. Examples of EFI fields in forecasts initialized at 00:00 UTC on 26–28 November 2015 (forecast days 10, 9, and 8, respectively) valid for storm Desmond on 5 December 2015 for (a–c) IVT and (d–f) precipitation. (g) The observed EFAS precipitation that exceeded the 99th percentile of climatology on 5 December 2015.

We also used a daily index of the North Atlantic Oscillation (NAO) [*Ferranti et al.*, 2015]. The NAO is a measure of the large-scale circulation across the European-North Atlantic sector, and it was used as a way to evaluate forecast skill in different atmospheric states.

3. Results and Discussion

As an example of the EFI, in Figure 1 we present EFI fields valid for storm Desmond on 5 December 2015 by using forecasts initialized at 00:00 UTC on 26–28 November 2015 (i.e., forecast days 10, 9, and 8, respectively). For the IVT EFI in Figures 1a–1c, a broad area indicating the potential for extreme IVT is found, albeit with weaker values on forecast day 9 (Figure 1b); in contrast, however, there is a weaker and smaller spatial signal present for the precipitation EFI (Figures 1d–1f). By comparing these EFI maps with the regions where observed precipitation on 5 December 2015 exceeded the 99th percentile of climatology (Figure 1g), qualitatively, the ability of the EFI to capture the extreme precipitation in the north-western British Isles and southern Norway can be ascertained. This comparison shows that a signal for extreme precipitation was more clearly seen in the IVT EFI and that the precipitation EFI first exhibited a signal over the worst affected region—northwest England—on forecast day 8. This suggests, qualitatively, that the IVT EFI product could have been beneficial in raising awareness of this extreme hydrological episode.

To determine the usefulness of the IVT compared to the precipitation EFI, we firstly consider all 361 forecasts over the 2013/2014, 2014/2015, and 2015/2016 winters. Figure 2 displays the ROC curves for IVT (black) and precipitation (grey) on forecast days 4, 6, 8, and 10. It is evident on days 4, 6, and 8 that the precipitation EFI is more able to discriminate between the extreme precipitation events and nonevents, as highlighted by the precipitation ROC curve generally being closer to the top left corner, and hence having higher ROC areas.

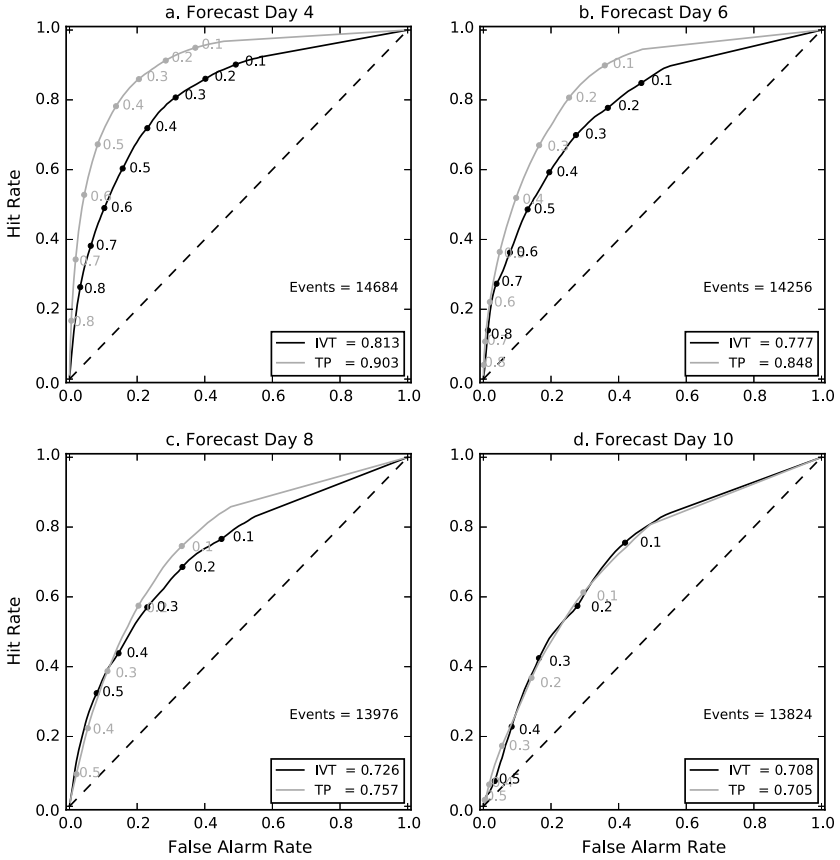

Figure 2. The relative operating characteristic (ROC) curves for IVT (black lines) and precipitation (grey lines) for all 361 winter forecasts on forecast days (a) 4, (b) 6, (c) 8, and (d) 10. The ROC areas are provided in the legends, and the number of extreme precipitation events is also given.

Note that at shorter forecast lead times, similar results are found as on forecast day 4 (not shown). The reason for the poorer performance of the IVT EFI is due in part to the broad scale nature of water vapor fluxes (often associated with atmospheric rivers) compared to the smaller scale characteristics of precipitation. This property tends to result in the IVT EFI not only capturing an event but also a large swath around it, resulting in a relatively high false alarm rate. Figure 2d shows that it is only on forecast day 10 that there is evidence in the ROC curve and area for a slight benefit in using the IVT EFI (over the precipitation EFI) to identify extreme precipitation.

Given that forecasts initialized in different atmospheric states can have varying levels of predictability and skill, we secondly conditioned the EFI verification on the NAO index on forecast day 1. The results conditioned on the 90 top ranked NAO states (i.e., 25% most positive) and the 90 bottom ranked NAO states (i.e., 25% most negative) are shown for forecast days 9 and 10 in Figures 3a and 3b, respectively. It is clearly seen that during the most positive NAO states, the IVT EFI is more useful at detecting extreme events than the precipitation EFI (cf. the solid black and grey lines in Figures 3a and 3b). Conversely, during the 90 bottom ranked NAO states, as shown by the dash-dotted lines in Figures 3a and 3b, the precipitation EFI is more able to discriminate the extreme events than the IVT EFI. These results suggest that the most useful product to monitor to anticipate extreme precipitation events depends on the large-scale circulation, as defined by the NAO. The benefit of using the IVT EFI during the top 90 ranked NAO states is thought to arise because during this phase extratropical cyclone activity and atmospheric rivers are more common, meaning that the IVT EFI may better capture the processes responsible for extreme precipitation. Conversely, it is probable that the extreme precipitation during the bottom ranked NAO states is less related to intense IVT, as extratropical cyclone and atmospheric river activity are less common in this phase.

We summarize the NAO-conditioned results by assessing the change in ROC area with forecast day (Figure 3c). This highlights the findings presented in Figures 3a and 3b and reveals that during week 1, similar

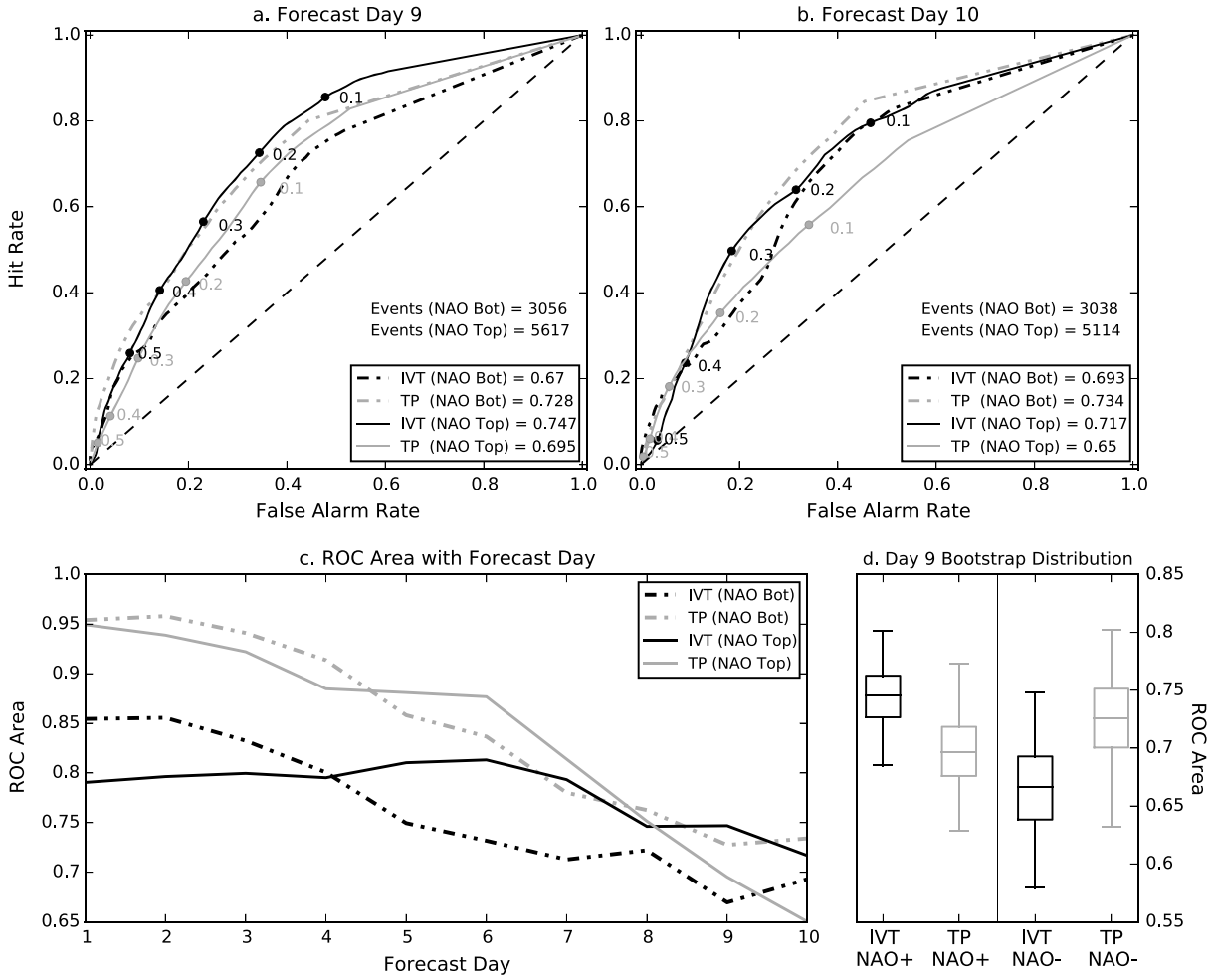

Figure 3. The relative operating characteristic (ROC) curves on forecast days (a) 9 and (b) 10 conditioned on the NAO index. The solid lines are for the 90 top ranked NAO index days, and the dash-dotted lines are for the 90 bottom ranked NAO index days (IVT in black; precipitation in grey). The ROC areas are provided in the legends, and the number of extreme precipitation events in each NAO category is also given. For clarity, ROC curve points and numbering are only given for the solid lines. (c) The change in ROC areas with forecast day. (d) The ROC area distributions on forecast day 9 calculated from a bootstrap process that was repeated 1000 times by using resampling of the 90 NAO-conditioned days with replacement. The bottom and top of the boxes correspond to the 25th and 75th percentiles, respectively; the line in the box is the median, and the whiskers represent the 1st and 99th percentiles.

results (to using all 361 forecasts) of the precipitation EFI being more able to discriminate the extreme events are found. The plot also shows that the ROC area degrades more rapidly for precipitation than for IVT, which possibly relates to the higher medium-range predictability of IVT [*Lavers et al.*, 2014]. A bootstrap procedure was further employed (the bootstrap process was repeated 1000 times by using resampling of the 90 NAO-conditioned days with replacement) to evaluate the statistical significance of the differences in ROC areas, and the results for forecast day 9 are shown in Figure 3d. As the interquartile ranges of the IVT and precipitation do not overlap, this suggests that there is statistical evidence for a difference in ROC areas and that the IVT EFI is more useful than the precipitation EFI during the positive NAO phase and less useful during the negative NAO phase. We note that changes in the ECMWF model configuration over the three winters are not considered to have had undue influence on the results because the forecasts used in the NAO-conditioned verification were fairly evenly sampled from across the three winters.

Following the potentially useful results with the IVT EFI over western Europe, in Figure 4 we present the product at a global scale at the time of storm Desmond. On forecast day 9 from 00:00 UTC on 27 November 2015 (Figure 4a), a signature of the extreme IVT associated with Desmond is found over the North Atlantic Ocean; an IVT feature is also seen approaching western North America. A global IVT EFI map is also presented for forecast day 1 (from 00:00 UTC on 5 December 2015) in Figure 4b. This map has much stronger values, which

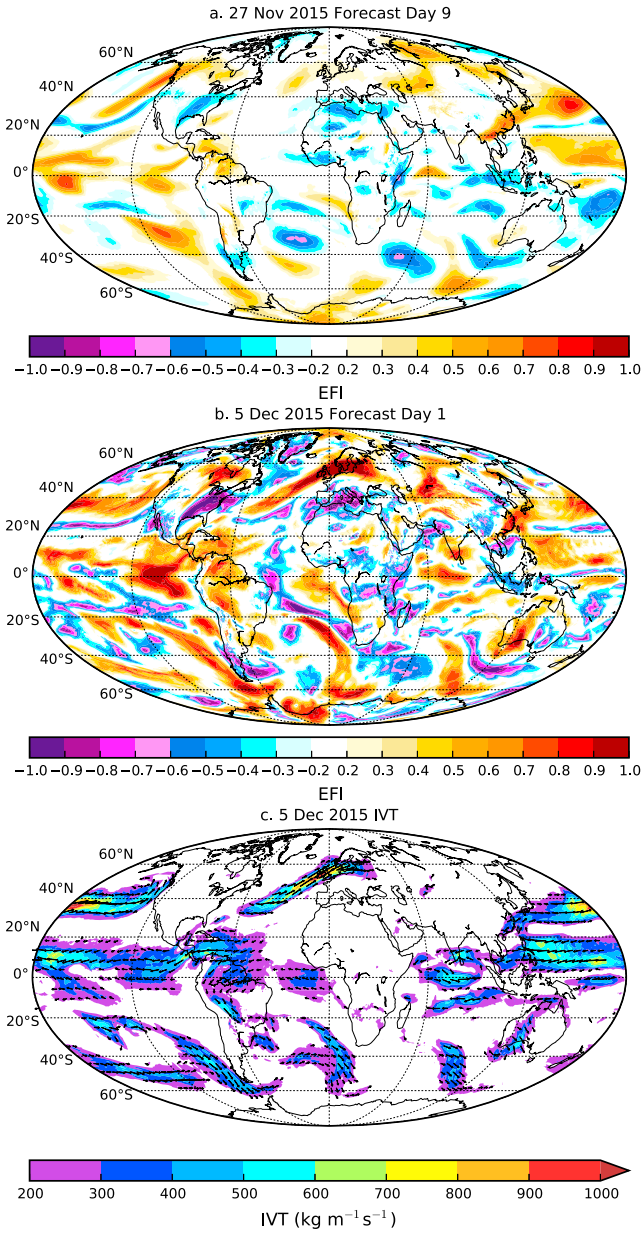

Figure 4. The IVT EFI product at a global scale valid for storm Desmond on 5 December 2015 on (a) forecast day 9 initialized at 00:00 UTC on 27 November 2015 and on (b) forecast day 1 initialized at 00:00 UTC on 5 December 2015. (c) Proxy observed daily IVT (kg m^{-1} s^{-1}) calculated by using the 0, 12, and 24 h lead times of the control forecast initialized at 00:00 UTC on 5 December 2015.

is expected at short lead times, indicating that the ensemble forecasts are more extreme with respect to the model climatology, and the atmospheric river responsible for extreme precipitation across north-western Europe is clearly visible. The somewhat similar structures in the fields on forecast days 1 and 9 are quantified by a Spearman rank correlation of 0.465. Also, a proxy for the observed daily IVT field, taken from the control forecast, shows the atmospheric river within Desmond affecting north-western Europe (Figure 4c). Although on forecast day 1 (and other short lead times) the higher false alarm rate of the IVT EFI may preclude its use compared to the precipitation EFI in forecasting the precise location of extreme precipitation, the IVT EFI does have the benefit of informing the user of the atmospheric processes behind the event, thus providing its synoptic context.

4. Conclusions

Building on previous research that first showed the EFI for IVT and that forecasts of IVT have higher predictability than those of precipitation, the aim of this study was to evaluate the ability of the IVT EFI forecast product to capture extreme precipitation in western Europe. In so doing, we have assessed the IVT and precipitation EFI in the ECMWF ensemble prediction system for three extended winter seasons. Our results suggest that the IVT EFI is more useful in detecting extreme events during early week 2 (i.e., forecast days 9 and 10) for forecasts initialized in the most positive phase of the NAO, indicating its potential to provide earlier awareness of upcoming hydrometeorological extremes. During the positive NAO phase, extratropical cyclones are more frequent, meaning that extreme precipitation events are more likely to be associated with the IVT within the atmospheric river of the storm. At shorter lead times, however, the large-scale characteristic of water vapor fluxes can lead to a higher false alarm rate than when using the precipitation EFI, in turn reducing its usefulness. We also note that precipitation is not only caused by IVT, and the omission of key rainfall processes (e.g., cloud microphysics) in the IVT is likely to be partly responsible for this result. For forecasts initialized on the bottom ranked NAO days (i.e., most negative phase), the precipitation EFI was found to be the most useful in discriminating the extremes.

This IVT EFI product could be beneficial in other regions where the IVT within atmospheric rivers is responsible for extreme events, such as in western North America. In these areas, future research will undertake an IVT EFI forecast verification by using high-resolution observed precipitation data sets, and as in Europe, it may also be advisable to condition the forecasts on the pertinent modes of atmospheric circulation. Furthermore, in areas where high IVT is the driver of extreme precipitation and flooding, the IVT EFI could complement the precipitation EFI at shorter lead times by providing insight to users and forecasters on the atmospheric processes responsible for an event. In the future, it is expected that the IVT EFI will become an operational forecast product.

Acknowledgments
The authors gratefully acknowledge financial support from the European Union Horizon 2020 IMPREX project (grant agreement 641811). We thank Ivan Tsonevsky, Linus Magnusson, and Paul Smith for the discussions that helped this research. The data used are listed in the references and are available on the ECMWF MARS server. We also thank two anonymous reviewers for their comments.

References

Bartholmes, J. C., J. Thielen, M. H. Ramos, and S. Gentilini (2009), The European Flood Awareness System EFAS. Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, *Hydrol. Earth Syst. Sci.*, *13*, 141–153, doi:10.5194/hess-13-141-2009.

Ferranti, L., S. Corti, and M. Janousek (2015), Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector, *Q. J. R. Meteorol. Soc.*, *141*, 916–924, doi:10.1002/qj.2411.

Lalaurette, F. (2003), Early detection of abnormal weather conditions using a probabilistic extreme forecast index, *Q. J. R. Meteorol. Soc.*, *129*, 3037–3057.

Lavers, D. A., R. P. Allan, E. F. Wood, G. Villarini, D. J. Brayshaw, and A. J. Wade (2011), Winter floods in Britain are connected to atmospheric rivers, *Geophys. Res. Lett.*, *38*, L23803, doi:10.1029/2011GL049783.

Lavers, D. A., F. Pappenberger, and E. Zsoter (2014), Extending medium-range predictability of extreme hydrological events in Europe, *Nat. Commun.*, *5*, 5382, doi:10.1038/ncomms6382.

Lavers, D. A., D. E. Waliser, F. M. Ralph, and M. D. Dettinger (2016), Predictability of horizontal water vapor transport relative to precipitation: Enhancing situational awareness for forecasting western U.S. extreme precipitation and flooding, *Geophys. Res. Lett.*, *43*, 2275–2282, doi:10.1002/2016GL067765.

Neiman, P. J., F. M. Ralph, G. A. Wick, J. D. Lundquist, and M. D. Dettinger (2008), Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the West Coast of North America based on eight years of SSM/I satellite observations, *J. Hydrometeorol.*, *9*(1), 22–47.

Ralph, F. M., P. J. Neiman, G. A. Wick, S. I. Gutman, M. D. Dettinger, D. R. Cayan, and A. B. White (2006), Flooding on California's Russian River: Role of atmospheric rivers, *Geophys. Res. Lett.*, *33*, L13801, doi:10.1029/2006GL026689.

Thielen, J., J. Bartholmes, M.-H. Ramos, and A. de Roo (2009), The European Flood Alert System. Part 1: Concept and development, *Hydrol. Earth Syst. Sci.*, *13*, 125–140, doi:10.5194/hess-13-125-2009.

Zsoter, E. (2006), Recent developments in extreme weather forecasting, *ECMWF Newslett.*, *107*, 8–17.

Zsoter, E., F. Pappenberger, and D. Richardson (2014), Sensitivity of model climate to sampling configurations and the impact on the Extreme Forecast Index, *Meteorol. Appl.*, *22*, 236–247, doi:10.1002/met.1447.