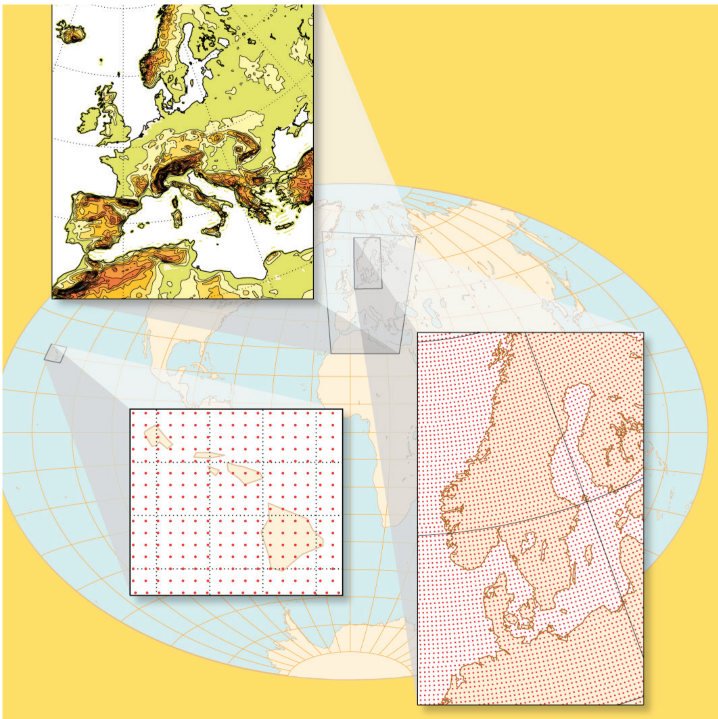
ECMWF Feature article

from Newsletter Number 108 – Summer 2006

METEOROLOGY

The ECMWF Variable Resolution Ensemble Prediction System (VAREPS)



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/st10ye392d

This article appeared in the Meteorology section of ECMWF Newsletter No. 108 – Summer 2006, pp. 14–20.

The ECMWF Variable Resolution Ensemble Prediction System (VAREPS)

Roberto Buizza, Jean-Raymond Bidlot, Nils Wedi, Manuel Fuentes, Mats Hamrud, Graham Holt, Tim Palmer, Frederic Vitart

The Ensemble Prediction System (EPS) has been part of the ECMWF operational suite since December 1992. At that time, the EPS was based on 33 forecasts produced with a T63L19 (spectral triangular truncation T63 with 19 vertical levels) resolution version of the ECMWF model (*Molteni et al.,* 1996). The initial uncertainties were simulated by starting 32 members from perturbed initial conditions defined by T21L31 perturbations which are rapidly-growing during the first 36 hours of the forecast range (the singular vectors, see *Buizza & Palmer,* 1995).

Since December 1992, the EPS has been upgraded several times. During these years, the EPS has used the same model version as the data assimilation and forecast system, benefiting from all the changes made. Some of these changes included substantial modifications of the EPS configuration, designed to improve both the simulation of initial and model uncertainties. It is worth identifying a few of them.

- In 1994 the optimisation time interval of the singular vectors was extended to 48 hours.
- In 1995 the resolution of the singular vectors was increased to T42L31.
- In 1996 the system was upgraded to a 51-member T159L31 system (spectral triangular truncation T159 with linear grid; *Buizza et al.,* 1998), with T42L31 singular vectors.
- In 1998 initial uncertainties due to perturbations that had grown during the 48 hours previous to the starting time (evolved singular vectors, *Barkmeijer et al.,* 1999) were included, and a scheme to simulate model uncertainties due to random model error in the parametrized physical processes was introduced (*Buizza et al.,* 1999). EPS wave forecasts became available following the introduction of the coupled atmosphere-wave model in the forecast model (*Saetra & Bidlot,* 2002, *Janssen et al.,* 2005).
- In 2000, following the resolution increase of the ECMWF data-assimilation and high-resolution systems from T319L31 to T511L60, the EPS resolution was upgraded to T255L40 (*Buizza et al.,* 2003), with T42L40 singular vectors. The wave model resolution was increased to a grid spacing of the order of 110 km.
- In 2002 tropical perturbations were added to the system (*Barkmeijer et al.,* 2001).
- In 2004 the Gaussian sampling method for generating the EPS initial perturbations using singular vectors was implemented (*Ehrendorfer & Beck,* 2003).
- On 1 February 2006, following another resolution increase of the ECMWF data-assimilation and high-resolution systems to T799L90, the EPS resolution was further increased to T399L62 (see the article by *Untch et al.* in this Newsletter), with T42L62 singular vectors. The wave model spectral resolution was increased to 30 frequencies and 24 directions respectively without any change to its horizontal resolution.

The most recent change is the first of a three-phase upgrading process that will lead to the implementation of the ECMWF Variable Resolution Ensemble Prediction System (VAREPS). This is designed to increase the ensemble resolution in the early forecast range and to extend the forecast range covered by the ensemble system initially to 15 days and eventually to one month. The planned merger of the medium-range ensemble and the monthly operational system is going to be carried out in three phases.

- ***Phase 1 (February 2006):*** resolution increase of the 10-day EPS from T255L40 to T399L62.
- ***Phase 2 (planned for the second half of 2006):*** extension of the forecast range to 15 days using VAREPS, with T399L62 (day 0-10) and T255L62 (day 9-15).
- ***Phase 3 (planned for 2007):*** weekly extension of VAREPS to one month, with a T255L62 atmospheric resolution and ocean coupling introduced at day 10 (the precise configuration of this final stage of VAREPS is still to be finalized).

Only the first two phases are discussed here: the phase-3 extension to one month will be discussed in a forthcoming article.

 doi:10.21957/st10ye392d

The rationale behind a variable resolution approach

VAREPS aims to provide better predictions of small-scale, severe-weather events in the early forecast range, and skilful large-scale guidance in the medium forecast range. The strategy used to achieve these goals is (a) to resolve small-scales up to the forecast time when they are predictable and their inclusion has a positive impact on the forecast accuracy, and (b) not to resolve them later in the forecast range when including them has a smaller, less detectable impact. This strategy leads to a more cost-efficient use of the computer resources, with most of them used in the early forecast range to resolve the small but still predicable scales. It is worth noting that a similar approach to ensemble prediction is not new, since it has been used at the National Centers for Environmental Prediction (NCEP, Washington) since inception of their ensemble prediction system (*Toth & Kalnay,* 1997).

The planned operational configuration

Technically, each VAREPS member will be generated by a two-leg forecast:

- ***leg-1:*** T399L62, from day 0 to day 10.
- ***leg-2:*** T255L62, from day 9 to day 15.

The horizontal resolution of the wave model stays unchanged (~110 km), however *leg-1* is now run with the same spectral resolution as the deterministic forecast (30 frequencies and 24 directions). The second leg reverts to 25 frequencies and 12 directions.

VAREPS will also include two other constant-resolution forecasts for calibration/validation purposes: a 15-day T399L62 forecast and a 15-day T255L62 forecast (these two extra forecasts will be added to the VAREPS suite following users' requests; data from these will be accessible from MARS in *stream* = ENFO as *type* = CV, *number* = 1, 2).

Key VAREPS technical characteristics

Users should be aware of three key VAREPS technical characteristics.

- ***Leg-2 initial conditions*** – Each *leg-2* forecast starts from a *leg-1* day-9 forecast (see Figure 1), interpolated at the T255L62 resolution (in other words, the *leg-2* initial state is defined by a *leg-1* forecast instead of analysis fields for all the state-vector variables). The 24-hour overlap period has been introduced to reduce the impact on the fields that are more sensitive to the truncation from the high to the low resolution (e.g. convective and large scale precipitation). High-resolution wave spectra are smoothed out to the lower spectral resolution of the second leg.

- ***Accumulated fields*** – Accumulated fields are accumulated from the start of the *leg-1* forecast. In the *leg-2* forecast, to accumulate from the start of *leg-1*, once the *leg-2* forecast reaches the end of the overlap period (24-hour, i.e. day-10 if counted from the beginning of the *leg-1* forecast), the accumulated fields are overwritten by the *leg-1* 10-day forecast fields interpolated onto the T255 reduced Gaussian grid.

- ***FDB and MARS streams ENFO and EFOV*** – In the Field Data Base (FDB) and the Meteorological Archival and Retrieval System (MARS), *leg-1* forecasts from day 0 to day 10, and *leg-2* forecasts from day 10 to day 15 are written in the MARS stream ENFO (Ensemble Forecast stream), while *leg-2* forecasts from day 9 to day 10 are written in the new MARS stream EFOV (Ensemble Forecast Overalp stream). The *leg-1* 10-day forecast fields interpolated on the T255 reduced Gaussian grid are archived in the overlap stream, so that they can retrieved if needed (e.g. to correctly compute accumulated fields across the truncation forecast step). Similarly, ensemble wave fields are written in, respectively, streams WAEF and WEOV.

For a more detailed description of how to compute accumulated fields across the truncation forecast step (i.e. after forecast day 10), the reader is referred to the document *"Computation of accumulated fields in VAREPS",* accessible from the ECMWF web site at: www.ecmwf.int/products/data/operational_system/evolution/evolution_2006.html

These set-ups ensure that only users interested in using VAREPS forecast for accumulated fields after forecast day 10 need to take care when constructing fields accumulated between two forecast steps that include the truncation step.

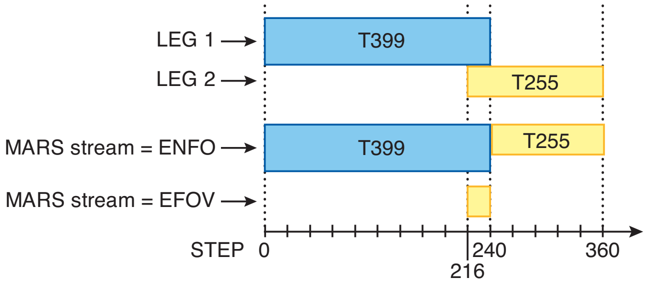


Figure 1 Schematic of the two-leg VAREPS planned for operational implementation, with MARS data streams ENFO and EFOV.

Expected average impact of introducing VAREPS

To assess the impact of the introduction of VAREPS, ensembles run with an earlier VAREPS configuration with a day-7 truncation, a 13-day forecast length and 40 vertical levels have been compared with two constant-resolution ensemble configurations.

- *T255:* T255L40(day 0–13), with a 2700 second time step (this was the EPS configuration operational before 1 February 2006).
- *VAREPS:* T399L40(day 0–7) with a 1800 second time step and T255L40(day 6–13) with a 2700 second time step.
- *T319:* T319L40(day 0–13) with a 1800 second time step.

The second and the third configurations require ~3.5 times the computing requirements of the first configuration. Hereafter, the average performance of these configurations in providing probabilistic predictions of 500 hPa geopotential height, 850 hPa temperature and total precipitation anomalies over the Northern Hemisphere are compared. Apart from the resolution, these ensembles used the same model cycle, started from the same analysis, had the same set of initial perturbations and were based on 50 perturbed plus 1 unperturbed forecast.

Verification: T255(day 0–13) EPS versus VAREPS T399(day 0–7)+T255(day 7–13)

Figure 2 shows the 60-case average area under the relative operating characteristic curve and the Brier Skill Score for the probabilistic prediction of total precipitation in excess of 10 mm over 12 hours, for the T255 EPS and VAREPS. The forecasts are verified against a proxy of observed precipitation defined by the 24-hour forecast of the operational, high-resolution system. These 60 cases span a five-year period, and include both severe and non-severe event cases (in selecting these cases care was taken not to introduce any bias in the sample). This figure also shows the value of the rank-sum Mann-Whitney-Wilcoxon (RMW) significance test (computed using a bootstrapping technique): this test measures the probability that the distributions of scores for the systems may come from the same overall population. For example, RMW values of 10% indicate that there is a 10% chance that the distributions of the two scores coincide. Figure 2 shows that VAREPS has higher average scores than T255 up to forecast day 7 for the 10 mm/12 h threshold, with RMW values below 10% in the first case and 20% in the second one. Results also indicate that after the truncation step the difference between the two systems is not statistically significant.

Figure 3 shows the 60-case average area under the relative operating characteristic curve and the Brier Skill Score for the probabilistic prediction of positive 850 hPa temperature and 500 hPa geopotential height anomalies, for the T255 EPS and VAREPS, verified against the ECMWF analysis. Results indicate that the difference between these two systems in terms of the prediction of these two other variables still favours the VAREPS, but the RMW test has values below 20% only up to forecast day 5.

It is worth pointing out that the area under the relative operating characteristic for the prediction of both 850 hPa temperature and 500 hPa geopotential height stays above 0.7 for the whole forecast range. This suggests that VAREPS can provide valuable probabilistic forecasts beyond 10 days (note that the current operational EPS stops at day 10).

 doi:10.21957/st10ye392d

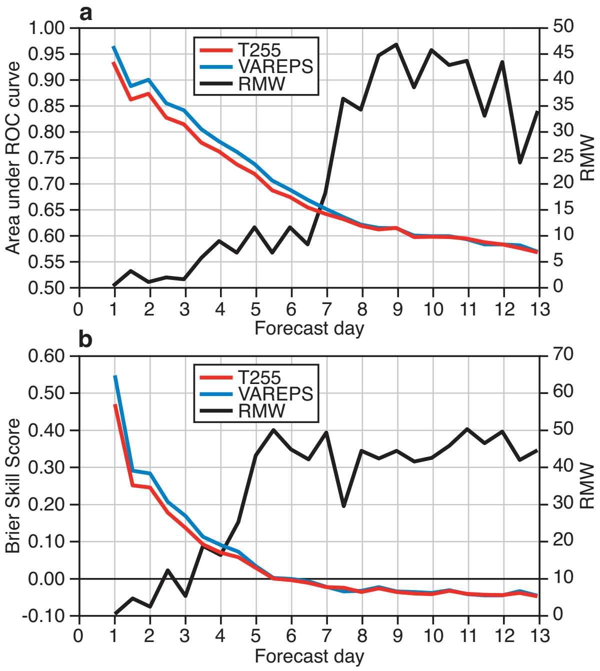


Figure 2 (a) 60-case average area under the relative operating characteristic (ROC) curve for the probabilistic prediction of total precipitation in excess of 10 mm/12 h over the Northern Hemisphere for T255 EPS (red line, left axis) and VAREPS (blue line, left axis), and the value of the rank-sum Mann-Whitney-Wilcoxon significance test (RMW, black line, right axis). (b) As (a) but for the Brier Skill Score, computed against climatology.

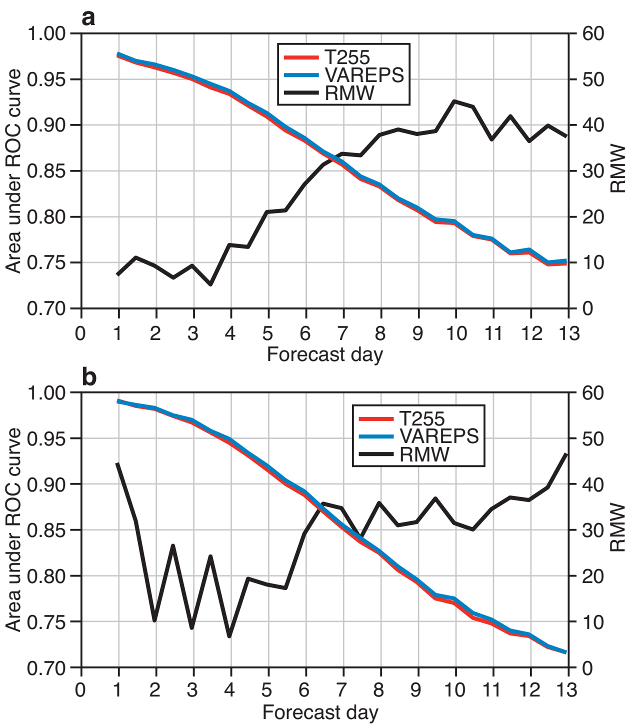


Figure 3 (a) 60-case average area under the relative operating characteristic (ROC) curve for the probabilistic prediction of positive 850 hPa temperature anomalies over the Northern Hemisphere for T255 EPS (red line, left axis) and VAREPS (blue line, left axis), and the value of the rank-sum Mann-Whitney-Wilcoxon significance test (RMW, black line, right axis). (b) As (a) but for the probabilistic prediction of positive 500 hPa geopotential height anomalies.

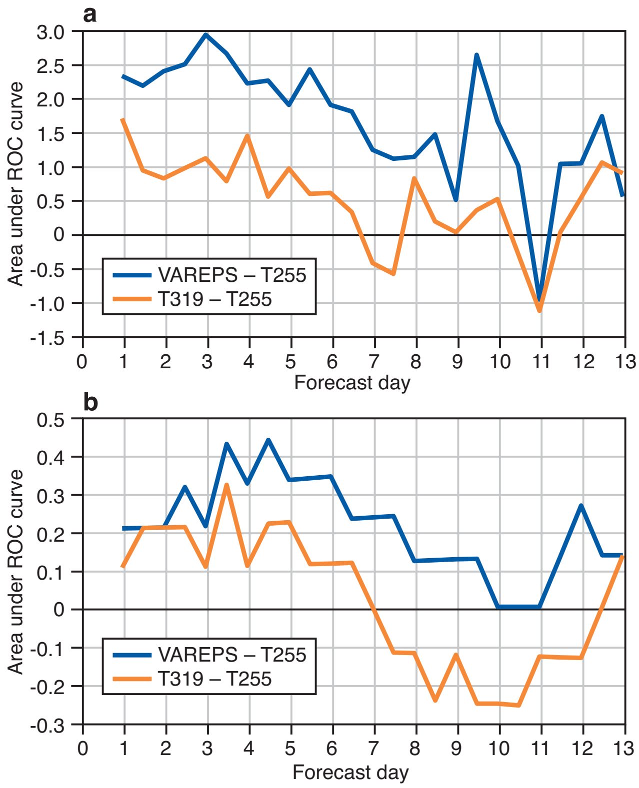


Figure 4 (a) Differences of 45-case average of the area under the relative operating characteristic (ROC) curve for the probabilistic prediction of total precipitation in excess of 10 mm/12 h over the Northern Hemisphere between VAREPS and T255 EPS (blue line) and between T319 system and T255 EPS (yellow line).
(b) As (a) but for the probabilistic prediction of positive 850 hPa anomalies.

Equal-cost comparison: T319(day 0–13) versus T399(day 0–7)+T255(day 7–13)

An assessment has been made of the relative improvement (compared to the T255 EPS) of two ensemble configurations that require the same amount of computing resources to be completed: VAREPS and a contant-resolution T319 ensemble system. Figure 4 shows the percentage differences between average values (computed for 45 of the 60 cases shown in Figures 2 and 3) of the area under the relative operating characteristic for total precipitation in excess of 10 mm/12 h and positive 850 hPa temperature anomalies. Positive/negative relative differences mean that VAREPS/T319 outperforms/ underperforms the T255 EPS.

Overall, results indicate first of all that both VAREPS and T319 outperform the T255 EPS, and, although the difference between the VAREPS and the T319 performances is small, that VAREPS is associated with a larger relative improvement than T319.

Impact of increased resolution in the short-range for selected cases

The results discussed so far suggest that VAREPS is, on average, a better system than the T255 ensemble that was operational up to the end of January 2006. The average differences are small but statistically significant with a RMW value below 20% up to forecast day 7. Results indicate also that VAREPS is to be preferred to a constant-resolution, equal cost T319 ensemble. The average results have also indicated that the differences are more detectable in the early forecast range, and especially if one considers fields characterized by small-scale features such as total precipitation.

Two synoptic cases are now discussed to illustrate the positive impact of increasing the resolution in the early forecast range from T255 to T399 in severe weather events.

Hurricane Katrina (29 August 2005)

The first case is very recent: Hurricane Katrina, one of the strongest storms of the last 100 years. Katrina started to develop as a tropical depression on 23 August south-east of the Bahamas, reached category 5 on 28 August and category 4 when it landed on the 29th. At landfall, close to New Orleans, sustained winds of more than 220 km/h were detected.

Figure 5 shows the intensity error (IE) and position error (D) of mean-sea-level-pressure (MSLP) minima predictions by the ensembles members of the T255 EPS, T319 system and VAREPS, with an 84, 72, 60 and 48 hour time lead. Ensemble forecasts have been clustered in three categories, accordingly to the intensity

 doi:10.21957/st10ye392d

and position errors: (*IE*<5 hPa, *D*<100 km), (*IE*<15 hPa, *D*<200 km) and (*IE*<30 hPa, *D*<300 km), with the first category identifying forecasts with very small errors. Accordingly to this accuracy measure, the T399L40 VAREPS has the highest number for all forecast ranges and for all categories apart for the T+60 h forecast for the category (*IE*<5 hPa, *D*<100 km).

As a consequence of the more accurate development and intensification of the hurricane in each ensemble member, significant wave height (SWH) probabilistic forecasts for the Gulf of Mexico are more accurate in the T399L40 VAREPS. This can be seen, for example, by comparing the 84-hour probability forecasts of SWH in excess of 8 m (Figure 6). The T255 system gives no probability of SWH exceeding 8 m and the T319 system gives a 2–5% probability, while the T399L40 VAREPS gives a 10–20% probability correctly located in the area where SWH exceeded 8 m in the ECMWF operational analysis. Similar differences are detected by comparing probabilistic forecasts for earlier forecast ranges.

In the case of Katrina, the highest resolution T399L40 VAREPS rightly intensified the hurricane development, thus improving probabilistic predictions of other surface variables such as wind speed and SWH. But it is worth mentioning that this is not because the T399L40 model systematically intensifies cyclonic developments. For example, in the case of Hurricane Stan, a system that caused severe damage and loss of life in Guatemala because of a land-slide induced by the intense precipitation, the T399L40 VAREPS forecasts outperformed the T255L40 and T319L40 forecasts mostly by positioning more accurately the area affected by the intense precipitation, rather than in the intensification of the cyclone.

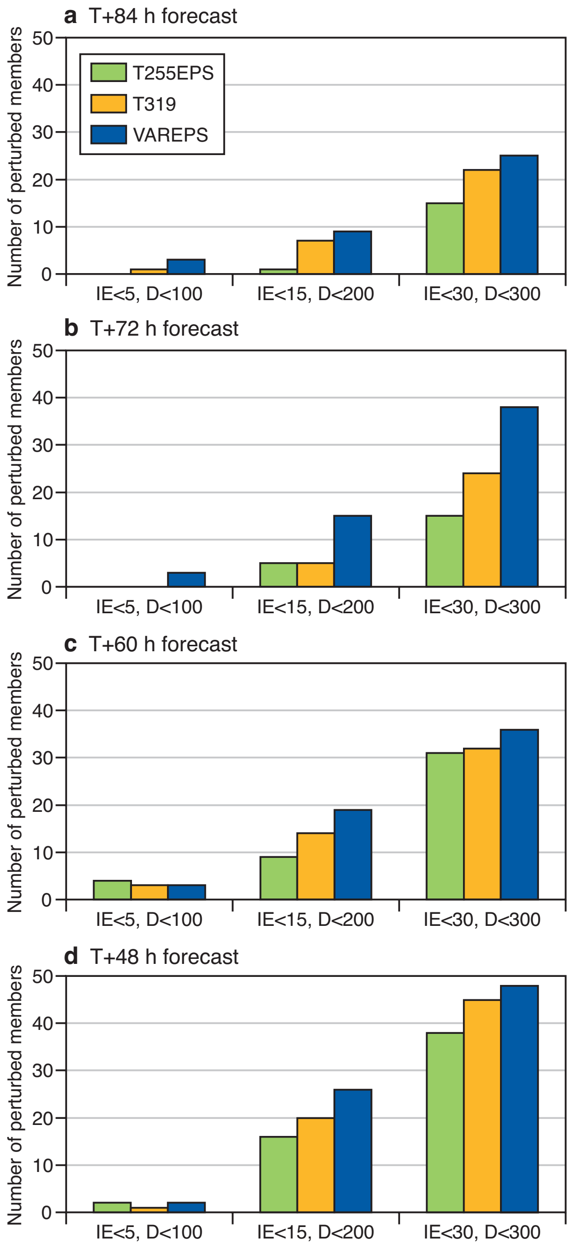


Figure 5 Mean-sea-level-pressure (MSLP) intensity and position error statistics for Hurricane Katrina for the T255L40 operational EPS, T319L40 system and T399L40 VAREPS forecasts valid for 12 UTC on 29 August 2005 using (a) T+84 hour, (b) T+72 hour, (c) T+60 hour and (d) T+48 hour forecasts. "IE<X, D<Y" refers to forecasts with intensity error less than X hPa and position error less than Y km (e.g. IE<5, D<100 indicates is the number of forecasts with intensity error less than 5 hPa and position error less than 100 km. Forecasts have all been verified against the operational TL511L60 analysis.

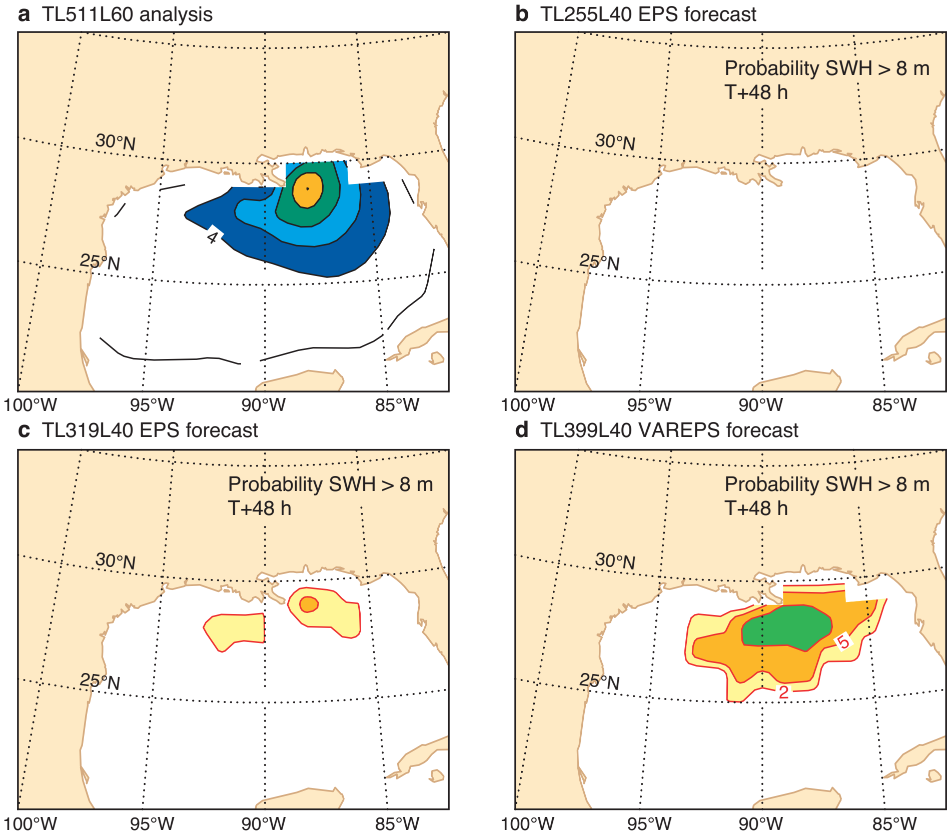


Figure 6 Forecasts of significant wave height (SWH) for Hurricane Katrina. (a) SWH from the operational T511L60 analysis valid at 12 UTC on 29 August 2005 (contour interval 2 m). (b) T+84 hour forecast of the probability of SWH higher than 8 m from the T255L40 operational EPS. (c) As (b) but from the T319L40 EPS. (d) As (b) but from the T399L40 VAREPS. Contour isolines for probabilities are 2, 5 and 10%.

Firenze flood (4 November 1966): The famous 'Alluvione di Firenze'

The second case is an historical one, the flood of North-Eastern and Central Italy of November 1966. This flood event is known as "l'alluvione di Firenze del '66", since Firenze was the most famous Italian city affected by it. As one of the most severe over Europe, this flood caused severe damage to the historical towns of Florence and Venice, disruption in the Po' Valley and in Tuscany, including loss of lives.

Figure 7 shows the T+48 to T+72 hour probabilistic prediction of total precipitation in excess of 75 and 150 mm given by the T255 EPS (Figures 7(a) and 7(c)) and the T399 VAREPS (Figures 7(b) and 7(d)) valid for the 24-hour period starting at 12 UTC on 3 November. These probability maps can be compared with the proxy for precipitation verification given by a T511L60 forecast started at 12 UTC on 3 November (Figure 7(e)). It is worth mentioning that this proxy field represents rather accurately the overall pattern of the observed precipitation field, but underestimates the maximum values (during the verification period, maximum values of between 200 and 400 mm were observed in Tuscany, and between 300 and 700 mm in North-Eastern Italy.

Figure 7 shows that higher probability values are predicted by the T399 VAREPS both over Tuscany and North-Eastern Italy in the areas where intense precipitation was detected. It is interesting to point out that the T399 VAREPS gives also a 40-60% probability that precipitation could exceed 150 mm over North-Eastern Italy, correctly indicating that this area was going to be affected by the most intense rainfall.

 doi:10.21957/st10ye392d

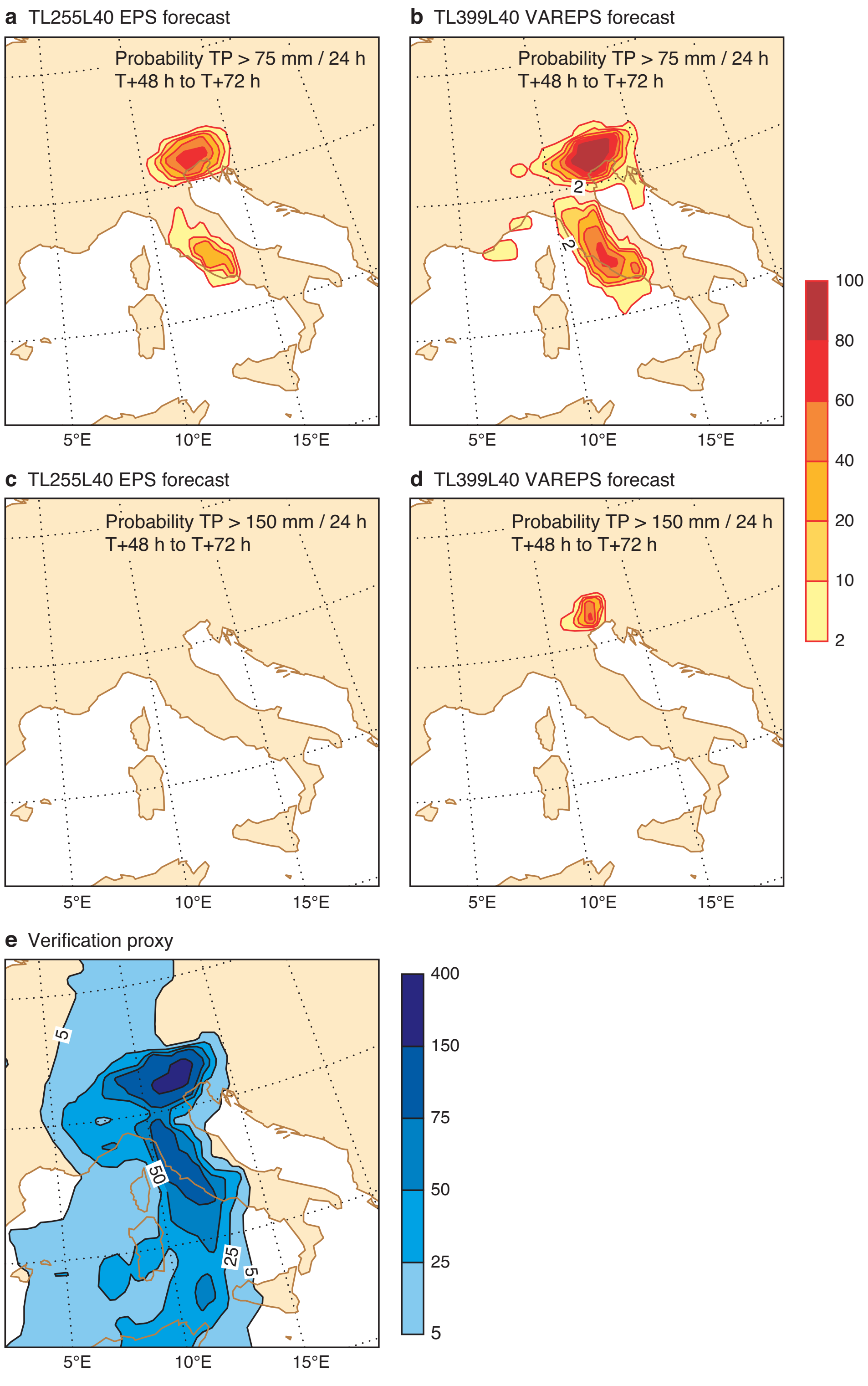


Figure 7 Probabilistic predictions of 24-hour total precipitation (TP) for the 1966 Italian flood. Forecasts started at 12 UTC on 1 November 1966 and are valid for the 24-hour period starting at 12 UTC on 3 November. (a) T+48 to T+72 hour EPS prediction of the probability of TP in excess of 75 mm/24 h. (b) As (a) but for T399 VAREPS. (c) T+48 to T+72 hour EPS prediction of the probability of TP in excess of 150 mm/24 h. (d) As (c) but for T399 VAREPS. (e) Verification proxy given by the T+24 hour T511L60 prediction of TP started at 12 UTC on 3 November. Contour isolines for probabilities are 2, 10, 20, 4, 60 and 80%, and for TP 5, 25, 50, 75, 150 and 400 mm.

Planned implementation schedule

The implementation on 1 February 2006 of the T399L62(day 0-10) ensemble prediction system completed the first of a three-phase upgrading process that will lead to the implementation of the ECMWF Variable Resolution Ensemble Prediction System (VAREPS). This is designed to increase the ensemble resolution in the early forecast range and to extend the forecast range covered by the ensemble system initially to 15 days and eventually to 32 days, following the planned merger of the medium-range and the monthly ensemble forecasting systems.

The second of this three-phase process, planned for the second half of 2006, will lead to the extension of the 00 and 12 UTC ensemble systems to 15 days using the VAREPS approach, with a T399L62 resolution up to forecast day 10 and a T255L62 resolution between forecast day 10 and 15.

VAREPS will further increase the value of the ECMWF probabilistic forecasting system, and deliver to ECMWF users more accurate predictions of small-scale, severe weather events in the early forecast range and skilful probabilistic predictions of larger scale features in the medium forecast range.

Further reading

Barkmeijer, J., R. Buizza & **T.N. Palmer,** 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.,* **125,** 2333–2351.

Barkmeijer, J., R. Buizza, T.N. Palmer, K. Puri & **J.-F. Mahfouf,** 2001: Tropical singular vectors computed with linearized diabatic physics. *Q. J. R. Meteorol. Soc.,* **127,** 685–708 (also available as *ECMWF Technical Memo. No. 297*).

Buizza, R. & **T.N. Palmer,** 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.,* **52,** 1434–1456 (also available as *ECMWF Technical Memo. No. 208*).

Buizza, R., T. Petroliagis, T.N. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons & **N. Wedi,** 1998: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Q. J. R. Meteorol. Soc.,* **124,** 1935–1960 (also available as *ECMWF Technical Memo. No. 245*).

Buizza, R., M. Miller & **T.N. Palmer,** 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.,* **125,** 2887-2908 (also available as *ECMWF Technical Memo. No. 279*).

Buizza, R., D.S. Richardson & **T.N. Palmer,** 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Q. J. R. Meteorol. Soc.,* **129,** 1269-1288 (also available as *ECMWF Technical Memo. No. 389*).

Ehrendorfer, M. & **A. Beck,** 2003: Singular vector-based multivariate sampling in ensemble prediction. *ECMWF Technical Memo. No. 416.*

Janssen P., J.-R Bidlot, S. Abdalla & **H. Hersbach,** 2005: Progress in ocean wave forecasting at ECMWF. *ECMWF Technical Memo. No. 478*.

Molteni, F., R. Buizza, T.N. Palmer & **T. Petroliagis,** 1996: The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.,* **122,** 73–119 (also available as *ECMWF Technical Memo. No. 202*).

Saetra, Ø. & **J.-R Bidlot,** 2002: Probabilistic forecasts of ocean waves. *ECMWF Newsletter No. 95,* 2–9.

Toth, Z., & **Kalnay, E.,** 1997: Ensemble Forecastign at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Untch, A., M. Miller, M. Hortal, R. Buizza, & **P. Janssen,** 2006: Towards a global meso-scale model: The high-resolution system TL799L91 and TL399L62 EPS. *ECMWF Newsletter No. 108,* 6–13.

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England

 doi:10.21957/st10ye392d