ECMWF ACTIVITIES FOR IMPROVED HURRICANE FORECASTS

L. Magnusson, J.-R. Bidlot, M. Bonavita, A. R. Brown, P. A. Browne, G. De Chiara, M. Dahoui, S. T. K. Lang, T. McNally, K. S. Mogensen, F. Pappenberger, F. Prates, F. Rabier, D. S. Richardson, F. Vitart, and S. Malardel

ECMWF global forecasts of tropical cyclone tracks have improved greatly in the past decades, and addressing the challenge of reliable intensity forecasts is an active research area.

Tropical cyclones are some of the most devastating natural hazards. They usually impact a large area, causing damage through strong winds and heavy rain. They cause even greater harm through damage from storm surges, flooding, landslides, or tornadoes. To reduce this damage, early and reliable warnings are needed from weather centers worldwide including the World Meteorological Organization (WMO) regional specialized meteorological centers (RSMCs) and tropical cyclone warning centers (TCWCs). These centers provide forecasts of the future location and intensity of the cyclones. This information has considerable economic value; the total loss during the 2017 Atlantic hurricane season reached $220 billion (Munich RE 2018).

Global forecasts provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) are regularly used by RSMCs as well as commercial enterprises. These forecasts are used in their raw format, are used as boundary conditions for limited-area models, and/or form the basis for statistical postprocessing. Several of the national meteorological services in ECMWF member states operate in parts of the world where tropical cyclones occur. This became very clear with the case of Hurricane Irma that severely affected French, Dutch, and British overseas territories in the Caribbean.

The basic forecast quantities for predicting tropical cyclones are the track of the cyclone and its intensity, determined by either the maximum sustained wind
seasonal scales. The deterministic high-resolution ING SYSTEM IN 2017.

the progress on the current strategy benefits tropical cyclone forecasting, using the three major hurricanes the ECMWF system. We demonstrate how of tropical cyclone intensity forecasting, using the...result reports acknowledge the tropical cyclone tracks from ECMWF as world leading both in deterministic (Yamaguchi et al. 2017; Cangialosi 2018) and probabilistic applications (Yamaguchi et al. 2015). However, ECMWF forecasts for tropical cyclone intensity still have a low level of skill, even if the ECMWF model is nowadays able to predict strong intensities in terms of central pressure for mature cyclones (Rodwell et al. 2012). Cangialosi (2018) stated that maximum wind forecasts from ECMWF for the 2017 Atlantic hurricane season were not skillful. The challenge from the ECMWF perspective is that the global model configuration cannot be tailored to perfectly suit tropical cyclones, since the model physics must perform well for all regions on the globe. Furthermore, the current horizontal resolution of global models (approximately 10 km) is still believed to limit the ability to predict intense cyclones due to the sharp gradient in the core.

ECMWF's strategy (ECMWF 2016) highlights the aim to have a high-resolution ensemble forecasting system, with a model that includes several different components of the Earth system. To do this, data assimilation for all individual system components is needed together with an optimal use of available observations. One main objective is to be able to predict high-impact weather two weeks in advance. In this paper, we review recent changes and highlight the future challenges that are applicable across all global and regional forecasting systems to improve the skill of tropical cyclone intensity forecasting, using the example of the ECMWF system. We demonstrate how the progress on the current strategy benefits tropical cyclone forecasting, using the three major hurricanes of 2017 (Harvey, Irma, and Maria) as examples to illustrate both the current capability and the benefits of recent and future developments.

THE ECMWF OPERATIONAL FORECASTING SYSTEM IN 2017. ECMWF produces global forecasts aimed at medium-range to subseasonal and seasonal scales. The deterministic high-resolution forecast (HRES) uses 9-km horizontal resolution and 137 vertical levels (of which 20 are below 1,000 m) and runs twice a day out to a lead time of 10 days. The ensemble (ENS) uses 18-km horizontal resolution and 91 vertical levels and runs out to 15 days twice a day. Twice a week (Monday and Thursday) the ensemble forecasts are extended to 45 days at 36-km resolution to provide forecasts on the subseasonal time scale. Once a month, seasonal forecasts are produced with 7-month lead time. The atmospheric component of the forecasting system is the ECMWF Integrated Forecasting System (IFS) model, which during the summer of 2017 was based on model version 43r3 (ECMWF 2017). In 2017, all systems but HRES were coupled to the Nucleus for European Modelling of the Ocean (NEMO; Madec 2008) ocean model with 0.25° resolution (note that in June 2018 HRES also became coupled to the same interactive ocean). The ocean coupling is described in Mogensen et al. (2017, and references therein).

The initial conditions are currently created separately for the atmosphere and ocean. The atmospheric initial conditions are produced with a four-dimensional variational data assimilation (4D-Var) system (Rabier et al. 2000). To provide background error statistics a 25-member ensemble of 4D-Vars [ensemble of data assimilations (EDA)] is run with a lower (18 km) resolution (Bonavita et al. 2012). The EDA members are also used to initialize the ensemble forecast (see below). The ocean initial conditions are provided by the ECMWF OCEAN5 operational three-dimensional variational data assimilation system (Zuo et al. 2018).

Ensemble forecasts are run with the aim of estimating a range of possible future states (e.g., Leutbecher and Palmer 2008). At ECMWF, the ensemble consists of 50 perturbed forecasts and one unperturbed member with the same resolution (control member). The ensemble is generated by applying initial perturbations based on a combination of EDA and singular vector perturbations (Buizza 2008) and model uncertainties represented by the stochastically perturbed parameterization tendency (SPPT) scheme and the stochastic kinetic energy backscatter (SKEB) scheme [see Leutbecher et al. (2017) for details]. To increase the spread of the ENS tropical cyclone track forecasts, singular vectors are targeted on each reported tropical cyclone (Barkmeijer et al. 2001). For a detailed discussion of the impact of the different perturbation methods and their relative contributions to the ENS tropical cyclone track and intensity forecasts spread, see Lang et al. (2012).
PREDICTIVE SKILL FOR HARVEY, IRMA, AND MARIA. In August and September 2017, three major hurricanes (Harvey, Irma, and Maria) hit the Caribbean and the United States. Each of them is documented in reports from the National Hurricane Center [Harvey by Blake and Zelinsky (2018), Irma by Cangialosi et al. (2018), and Maria by Pasch et al. (2018)], and we refer to these reports for details. In this article, we will use as verification the cyclone intensity and positions determined and broadcast by the National Hurricane Center on the Global Teleconnection Network (GTS), and later put into the best track database (we will refer to the dataset as BestTrack). However, one needs to be aware of the caveats for this product, for example, regarding the intensity and also the position of weak cyclones. [See, e.g., Velden et al. (2006) and Trahan and Sparling (2012) for further details.]

Harvey made landfall in Texas on 26 August (25 August local time) and, subsequently, the cyclone became quasi-stationary and produced persistent rainfall for 3–5 days. The rainfall totals reached more than 1,000 mm in the worst-affected areas around Houston, where unprecedented flooding occurred. Irma hit several countries along its path in the Caribbean Sea. On 5 and 6 September, the category 5 cyclone (on the Saffir–Simpson scale) made its first landfall on some of the Leeward Islands, where the islands of Barbuda, Saint Barthelemy, and Sint Maarten/Saint Martin were hit worst. The cyclone later hit the Virgin Islands, the Turks and Caicos Islands, and the Bahamas. Irma made landfall on Cuba on 8/9 September as a category 5 cyclone and made its final landfall on the southern tip of Florida on 10 September. Maria was classified as a tropical storm on 17 November, explosively intensified just before hitting Dominica, and kept its intensity until 20 November, when it hit Puerto Rico and its extreme winds and rainfall both caused severe damage.

To illustrate the ensemble performance, we show forecasts issued 6 days before landfall in Texas at 0000 UTC 26 August for Harvey (Fig. 1a), on the Leeward Islands at 0000 UTC 6 September for Irma (Fig. 1b), and on Puerto Rico at 0000 UTC 20 September for Maria (Fig. 1c). The gray lines show the 10-day track from each ensemble member and the symbols
indicate the position and intensity close to the observed landfall time for the BestTrack (hourglass symbol) and ensemble members (squares). The black line shows the path of the cyclone in the verification during its lifetime, and the circle symbols indicate the observed intensity every 6 h. To summarize the ensemble’s performance, Fig. 1d shows the percentage of ensemble members that passed within a 300-km radius of the BestTrack landfall (within 2 days) as a function of lead time. The plot includes results averaged over all observed cyclone positions every 12 h during its lifetime (solid line) and for the observation points in Figs. 1a–c (dashed lines). Figure 1d also includes the corresponding result for the landfall of Irma in Florida (dashed, pink line).

Looking into the medium-range performance, Harvey was the least predictable of the three cyclones, both in terms of the landfall in Texas and along the full track, partly because of its very weak structure before reaching the Gulf of Mexico. In the forecast for 0000 UTC 20 August (Fig. 1a), the cyclone was predicted to enter the Gulf of Mexico by most of the ensemble members but with landfall farther to the south compared to the verification in almost all members. Only a few members had a central pressure below 1,005 hPa (blue squares) and none was near being a category 4 hurricane as in the verification.

In the case of Maria, many ensemble members predicted the occurrence of a cyclone north of the Caribbean in the medium range, which was for example the case for the forecast at 0000 UTC 14 September in Fig. 1c. Note that this forecast was initialized 3 days before the cyclone was recognized as a tropical storm, and the plotted cyclones are selected by being within ±10° of the observed location on 16 September. These forecast cyclones are not necessarily associated with the same disturbance that became Maria, but may have developed from another disturbance present at the initial time of the forecast; most of them passed north of the formation point of Maria. Regarding the intensity of Maria, the ensemble predicted much weaker cyclones than the verification and only one member predicted a central pressure below 990 hPa (yellow square). Later ECMWF forecasts had problems capturing the propagation speed over the Atlantic for Maria (not shown).

Irma appears to be the most predictable of these three cyclones for both the observation points highlighted in Figs. 1a–c and the overall statistics presented in Fig. 1d. The passage of the Leeward Islands (dashed red) seems to have been more predictable than the landfall in Florida (dashed pink). The forecast from 0000 UTC 31 August in Fig. 1b had the ensemble plume centered on the verification point, and although being far from the observed intensity (916 hPa), two members predicted a central pressure below 965 hPa (red squares).

The presence of Irma in the Atlantic was even predicted with high confidence by the extended-range forecast system more than a week before the genesis of the cyclone. Such predictability is not unseen, and forecast verification based on a large reforecast dataset covering the past 20 years for a weekly strike probability product suggests some skill up to week 3 (Vitart and Molteni 2010), with predictability mainly originating from the strong modulation by the

| Table 1. Satellite observations used at ECMWF. |
|-----------------------------------------------|
| Measurement | Satellites | Primary information |
|-------------|------------|---------------------|
| Passive infrared radiances | Meteorological Operational satellites (MetOp), National Oceanic and Atmospheric Administration (NOAA), Earth Observing System (EOS), Meteorological Satellite (Meteosat), Geostationary Operational Environmental Satellite (GOES), Multifunctional Transport Satellite (Himawari) | Temperature, humidity, and wind (via 4D tracing) |
| Passive microwave radiances | MetOp, NOAA, Aqua, Defense Meteorological Satellite Program (DMSP), Global Precipitation Measurement (GPM), Global Change Observation Mission (GCOM), Sondeur Atmosphérique du Profil d’Humidité Intertropicale par Radiométrie (SAPHIR), Fengyun (FY) | Temperature, humidity, and wind (via 4D tracing) |
| Atmospheric motion vectors | Meteosat, GOES, Himawari, MetOp, Suomi National Polar-Orbiting Partnership (SNPP), Aqua, NOAA | Wind |
| GPS radio signal phase delays | MetOp, Constellation Observing System for Meteorology, Ionoosphere and Climate (COSMIC), TerraSAR-X, TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X), FY | Temperature |
| Scatterometers | MetOp | Sea surface winds |
| Altimeters | Jason, Satellite with Argos and Altika (SARAL), CryoSat | Sea surface height and significant wave height |
Madden–Julian oscillation and other tropical waves, which can be predicted up to 4–5 weeks in advance (Vitart 2017). All three cases experienced a rapid intensification just before the landfalls illustrated in Fig. 1 (cf. Fig. 3 for HRES forecasts). The HRES forecasts issued 3 days before the maximum intensity of Harvey and Maria had central pressure errors of more than 50 hPa, which were the two worst intensity forecasts of the full 2017 tropical cyclone season. Predicting rapid intensification is still challenging and is an area of intense research (e.g., Emanuel and Zhang 2016). In particular, the timing of the intensification seems to have low predictability (Zhang and Tao 2013; Judt and Chen 2016). The errors are expected to be larger in global models than in limited-area models due to the coarser resolution, but with global models entering the resolution domain below 10 km, intensity forecasts are receiving increased attention. For the case of Irma, the intensification was captured to some degree, and model sensitivities for this phase will be further explored in this paper.

ASSIMILATION OF OBSERVATIONS AROUND TROPICAL CYCLONES. One of the main challenges in assimilating tropical cyclones is the lack of direct observations in the vicinity of the active environment inside and around these systems. Available conventional observations (from buoys, aircraft, and ships) are usually limited to a few scattered surface pressure observations and sometimes drop sondes from reconnaissance flights. ECMWF does not make any use of tropical cyclone reports from RSMCs in the data assimilation, as is the practice at several other global prediction centers (e.g., Heming 2016). Instead, satellites provide critical observations for forecasting, especially in the early stage of tropical cyclone development.

ECMWF uses satellite observations from a wide variety of sophisticated sensors flying on more than 25 dedicated meteorological spacecraft (described in Table 1). This complementary information from multiple satellite instruments is brought together by the 4D-Var system to build a comprehensive thermodynamic picture of the cyclone and its environment and to provide the best possible initial conditions for predicting its future evolution. For skillful medium-range predictions of tropical cyclones, not only is information needed in the vicinity of tropical cyclones, but also in remote regions that later can affect the steering winds, as in the case of Tropical Cyclone Sandy in 2012 (McNally et al. 2014).

The impact of satellite data for Irma has been investigated by retrospectively rerunning the forecasting and assimilation system with satellite observations deliberately withheld, to simulate the effect of a major absence of satellite data. The experiment is designed so that no satellite data were used in the data assimilation for the 7 days prior to the initialization. The assimilation experiments, including one with all data used, were run with a reduced (30 km) resolution. The comparison between ensemble forecasts initialized from analyses with and without satellites for the arrival of Irma in the Leeward Islands on 6 September (Fig. 2) shows that the forecasts launched from initial conditions with full use of all available satellite data have a very clear signal of an intense cyclone 6 days ahead of time. In contrast, the ensemble made without satellite observations had, albeit still present in most members, much weaker cyclones and only 1 member was stronger than 1,005 hPa compared to 41 members for the operational configuration. A clear difference

---

**Fig. 2.** The 15-day ensemble forecast tracks (gray) and observed track (black), as well as symbols (ensemble members, squares; BestTrack, hourglass) representing the central pressure in 6-day forecasts, for Irma initialized at 0000 UTC 31 Aug (a) using the full set of observations in the assimilation and (b) with all satellite data withdrawn. The circles indicate the intensity along the observed track. Note that the ensembles are initialized from analyses with lower resolution than the operational configuration in Fig. 1b.
Later seen for the ensembles regarding the approach toward the United States. Even if this experiment is extreme in the sense of omitting satellite data, it could be seen as a worst-case scenario where several satellites are out of action simultaneously.

One potentially rich source for tropical cyclone intensity information is scatterometer winds, and ongoing research is aiming to improve their use in data assimilation (De Chiara et al. 2017). However, many cases are found where scatterometer observations of strong winds in the vicinity of tropical cyclones were rejected prior to the assimilation partly because of the data thinning and partly due to the variational quality control (Anderson and Jarvinen 1999). It is common practice in numerical weather prediction (NWP) systems to thin satellite observations to reduce their spatial and temporal error correlations and to decrease data density, and currently only every fourth scatterometer observation is used in each direction. This setting is currently under investigation by testing several combinations of thinning scheme and observation errors together with the configuration of the quality control for scatterometer winds, which could be beneficial in the vicinity of tropical cyclones.

Dropsondes deployed in the vicinity of tropical cyclones provide potentially valuable observations as they deliver a good vertical sampling of temperature, humidity, wind direction, and wind speed. Apart from humidity observations in the stratosphere and very-near-surface reports, the dropsonde data are automatically used by the ECMWF 4D-Var system. If used optimally, dropsonde data are expected to improve the representation of tropical cyclones (intensity and position) in the analysis and forecasts. However, assimilation of dropsonde observations is very challenging due to the extreme and rapidly changing conditions close to the cyclones. The handling of dropsondes in ECMWF 4D-Var has evolved over the years with the aim of extracting the relevant information, while protecting the system from numerical instabilities due to extreme gradients, for example, a recent recalibration of the observation quality control (Bonavita et al. 2017). A remaining problem for the data assimilation is because the current operational dropsonde reports do not contain information on the drift during the descent, which causes large differences in and around the core of the cyclone where the drift is very high. The expected change to BUFR high-resolution reports will address this issue by providing the dropsonde position at each level.

Fig. 3. Central pressure in observations (black) and analyses for (top) Harvey, (middle) Irma, and (bottom) Maria using all observations (pink) and without dropsondes (cyan), as well as forecasts initialized from different initial times from all-observation analyses (red) and without dropsondes (blue). The different initial times are distinguished by using solid- and dashed-line styles.

is later seen for the ensembles regarding the approach toward the United States. Even if this experiment is extreme in the sense of omitting satellite data, it could be seen as a worst-case scenario where several satellites are out of action simultaneously.
To assess the impact of dropsondes and the approach used to assimilate them, data-denial experiments are being conducted, targeting the full season of 2017. The results for the three cases are included in Fig. 3, which shows the evolution of the intensity in terms of central pressure around these times for verification (black), operational analysis (pink), and operational HRES forecasts initialized at different times (red). For Harvey and Maria, as well as for Irma around 1 September, the intensifications were not well predicted in the medium range or the short range by the ensemble (not shown) or HRES. Although a neutral impact is obtained for most of the forecasts, the results show a positive impact of dropsondes for the
assimilation during the intensification phase for Harvey and Maria, while no benefit was found for Irma. However, the use of the dropsondes did not improve the prediction of the onset of the rapid intensification. Future work will reveal if this is due to suboptimal usage of the information from the observations.

**IMPACT OF HORIZONTAL RESOLUTION ON FORECASTS.** Each horizontal resolution upgrade of the IFS model is associated with a better representation of the strong gradients that are characteristic of tropical cyclone circulations. As the spatial representation of the system is improved by the grid resolution, the values for the minimum of pressure get deeper and the resolved wind maxima represented at the scale of the model grid get stronger. The time resolution of the fast processes involved in the cyclone development is also improved as the time step is reduced in accordance with the horizontal resolution. Higher resolution also brings a better representation of orographic precipitation that is important for some landfalling cyclones; the impact for Hurricane Sandy (2012) was illustrated in Magnusson et al. (2014). The effects of the horizontal and time resolution are shown for two initial dates (0000 UTC 4 September, Fig. 4, left; 0000 UTC 6 September, Fig. 4, right) for Irma where we compare the operational configuration (red line) with an experimental configuration with 5-km resolution (pink) and with the time step reduced from 450 to 240 s (both configurations use the same vertical resolution). One should note that the 5-km forecasts are initialized from the operational (9 km) analysis, so that benefit from an increased resolution in the assimilation is not accounted for here. However, we find that the intensity rapidly spins up after the initialization and results in more intense cyclones with 5-km resolution. The improvement compared to the 9-km forecast is most apparent in the central pressure (Fig. 4, top), and even though the maximum wind speed (Fig. 4, middle) also increased, it is still clearly underestimated with 5-km resolution. This issue is illustrated in the wind–pressure plot in Fig. 4 (bottom), where we find too-weak winds during the intense phase of Irma. The plot also includes the statistics for all tropical cyclones during 2017 and confirms the problem in the operational HRES

Fig. 5. The (a),(c) 18-km (operational) ensemble forecast and (b),(d) 5-km ensemble forecast initialized at 1200 UTC 5 Sep. (a),(b) Strike probability for Irma passing within 120 km (shaded) and (c),(d) ensemble distribution of central pressure (box-and-whisker plot). Ensemble mean of (a),(b) the forecast tracks and (c),(d) the central pressure is included as solid lines. The observed track is shown by the black crosses in (a) and (b) and the observed central pressure by the red dots in (c) and (d).
forecasts. In contrast to most of the forecasts shown here, current operational forecasts have on average a small negative (too intense) bias in terms of pressure, but the bias is substantial (too weak) for the maximum wind speed.

As a demonstration of the potential for ensemble forecasts in the future, Fig. 5 shows a comparison between the 18-km operational ensemble and a 5-km 51-member ensemble with 91 vertical levels for Irma initialized from the same initial conditions (operational analysis) at 1200 UTC 5 September. The track forecast from the 5-km ensemble is similar to the operational forecast (cf. the top panels in Fig. 5), although the ensemble-mean track is slightly improved. The intensity forecast, on the other hand, is very different (Figs. 5c,d). The intensity errors in terms of central pressure are substantially reduced and the observed core pressure falls within the ensemble envelope most of the time. Note also the increase in the ensemble standard deviation around 9 September. This corresponds to the uncertainty associated with Irma making landfall on Cuba.

**IMPACT OF OCEAN COUPLING.** As mentioned above, since June 2018 all ECMWF forecasts from medium range to seasonal are based on a coupled atmosphere–wave–ocean model. In the past, the initial sea surface temperature (SST) anomalies were persisted through the forecasts, which led to a fictitious infinite source of heating. By introducing an interactive ocean, this unlimited heat source is removed, preventing spurious overdeepening of the core pressure of tropical cyclones as demonstrated in Mogensen et al. (2017) for the ECMWF model. This can also be demonstrated for Irma, and Fig. 4 shows the predicted core pressure (top panels) and maximum wind speed (middle panels) for Irma for both coupled and uncoupled experiments with both current atmospheric resolution (9 km; red and blue) and an experimental setup with 5-km resolution (pink and cyan). There is not much difference in the predicted pressure (Fig. 4, top panels) comparing uncoupled and coupled runs in the first few days. However, the uncoupled (especially the 5-km) runs continue to deepen the cyclone, whereas the coupled simulations weaken the cyclone when it starts to interact with the shallow waters of the Bahama Banks, in agreement with the BestTrack estimates.

The exchange of momentum at the ocean surface is sea-state dependent and is achieved by coupling the atmosphere with an ocean wave model. The momentum exchange depends on the state of the waves (sea state) and not just the surface winds, resulting in a range of possible values for the drag coefficient over the oceans for similar wind speeds, as illustrated in Fig. 6. For very high wind speeds, the coefficient is believed to reach an asymptotic level, as one of the limiting factor is the steepness of the waves. This factor was included in the parameterization of the momentum exchange in IFS cycle 43r1 (November 2016). The impact from the change to the drag coefficient is shown in Fig. 6 (red vs black line). Note that near the area of the strongest winds, the rapidly rotating wind field moves over new patches of the ocean, in a perpetual state of young wave generation corresponding to the cluster of relative low values of the exchange coefficient for momentum (Cd).

![Fig. 6. Exchange coefficient for momentum (Cd) for all model ocean points from 5-km coupled 10-day forecasts at 0000 UTC 4 Sep with the present wave model, stratified with respect to the corresponding 10-m wind speed. Running mean from current wave model (black) and with previous (noncapped) parameterization of Cd (red).](image-url)
A bit farther outward from the storm center, the interaction between the wind and waves has more time to act, resulting in a rapid increase in momentum transfer and, hence, larger $C_d$.

The sea state also has an impact on the latent and sensible heat fluxes via the exchange coefficients for heat and moisture. From recent observation campaigns (Brut et al. 2005; Cook and Renfrew 2015), there are indications that the exchange coefficients for heat and moisture should have a stronger dependency on the wind speed (i.e., the sea state) than currently modeled. Recent sensitivity experiments have shown the potential impact on the atmosphere (Janssen and Bidlot 2018), and ongoing work will quantify the impact of this new parameterization on tropical cyclones. With this change, more heat and moisture can be extracted from the ocean, the cyclone gets deeper, and surface winds stronger. It is worth stressing that such a change would not have been feasible without an interactive ocean, as the cyclone-induced cooling of the ocean is needed to avoid further overdeepening of cyclones.

**PERSPECTIVES FOR THE FUTURE.** The skill of hurricane forecasting has continuously improved over the past few decades and the track errors in the northern Atlantic and northwestern Pacific have been halved over two decades (Yamaguchi et al. 2017). Such improvements originate from a range of sources, such as resolution upgrades, improved coupling, and data assimilation, and are expected to continue into the future. While the skill of tropical cyclone track forecasts has reached a useful level and the ensemble on average produces reliable spread (Haiden et al. 2017), the future challenge lies in reliable intensity predictions, especially in cases with rapid intensification as illustrated in this paper. The ECMWF forecasts currently have only a small negative (too intense) bias in terms of central pressure, but in cases where the central pressure of the cyclones is well simulated, the maximum wind speeds around the tropical cyclones are underestimated compared to BestTrack data. The sensitivity test in this paper with a 5-km version of the ECMWF model shows a similar problem in reproducing the observed maximum wind–central pressure relationship as for the operational 9-km version. However, it can be regarded as an open question whether the expected relationship for either 9- or 5-km resolution should be the same as between the maximum sustained wind and central pressure in the verification, considering that the model mean wind represents a gridbox average. The question of the wind–pressure relationship will challenge both the model development and evaluation communities, and additional structure diagnostics for tropical cyclones such as the horizontal dimension (e.g., radius to maximum wind and radius to outermost closed isobar) are needed (Chavas et al. 2017). Another model property that needs further verification is the modeled wind gust. The current parameterization used at ECMWF is not specifically developed for the extreme conditions under tropical cyclones and might be suboptimal within this context.

To improve tropical cyclone forecasts, challenges lie in several parts of the forecasting system. In the U.S. community, the problem is tackled as part of the Hurricane Forecast Improvement Project (Gall et al. 2013) initiative. From the ECMWF perspective, in data assimilation, one major challenge is the observation usage in the vicinity of tropical cyclones. Historically, most of the satellite channels have been used only in clear-sky conditions. Recent developments (Bauer et al. 2010) have led to the use of such data also in cloudy conditions (all sky), which will increase their application, for example, around tropical cyclones. Another difficulty is the very strong gradients associated with tropical cyclones, which makes the data assimilation very sensitive to small errors in the forecast used as the first guess in the assimilation. The recent introduction of flow-dependent background errors has revealed the delicacy of observation usage close to the strongest gradients in tropical cyclones (Bonavita et al. 2017). We also need to stress that to make accurate cyclone track predictions in the medium range, a high-quality analysis of the hemispheric flow pattern is needed, as errors in the initial conditions can propagate from remote regions on this time scale.

The initialization of the coupled model is currently done separately for the atmosphere and ocean. To improve the initialization, ECMWF is working on methods to initialize the coupled ocean–atmosphere model simultaneously by using “outer loop” coupling in the analysis, where the 4D-Var nonlinear integrations use the coupled model, but the minimizations are performed separately for the atmospheric and oceanic states (Laloyaux et al. 2016). Notable benefits of coupled data assimilation are 1) more consistent ocean and atmospheric states (seen by a reduction in initialization shock; Mulholland et al. 2015), 2) the ability to have a more timely SST analysis that contains a diurnal cycle, and 3) the ability for in situ ocean observations to provide atmospheric increments, particularly under the high-cloud conditions of tropical cyclones. Currently, the ability to use passive microwave and infrared data to accurately determine SST is limited in storm conditions by the presence of opaque cloud and rain. However, in a coupled system, radar measurements of surface roughness as measured by scatterometer data...
(which are unaffected by cloud and rain) can improve our knowledge of the SST below the storm by informing on the degree of mixing with subsurface ocean water. This information is important for including effects such as the cold wake behind tropical cyclones, which can affect slow-moving/stationary cyclones and even subsequent cyclones.

The representation of finer-scale processes in a model is not only gained due to an increase in resolution, but goes hand in hand with updates of the model physics and the dynamical core of the model. While experiments at ECMWF so far indicate that the current hydrostatic equations are valid for representing the main processes of the cyclone life cycle at all resolutions coarser than 2 km, the partitioning between the resolved and the parameterized convection at resolutions around 5 km and higher is known to be a challenge that may directly affect the tropical cyclone representation when the IFS does reach such resolutions. New developments (not discussed in this paper) for the parameterization of convection (scheme adjustment for the gray zone\(^1\) of deep convection), boundary layer [improvement of the marine boundary layer (BL); Kepert 2012], and cloud physics (e.g., the description of the ice phase in the cloud microphysics), as well as a better understanding of the interaction between the parameterizations and the model dynamics, are works in progress at ECMWF. It is also necessary to ensure physical stability in the model during extreme conditions such as in tropical cyclones (Diamantakis and Magnusson 2016).

As the exchanges with the underlying ocean are some of the main factors controlling the intensity of the cyclone, model development efforts are also ongoing in terms of improved coupling with the wave model and the ocean model. Aspects to be revisited in the future include, for example, the formulation of the momentum and heat fluxes at the surface of the ocean in cases of extreme winds and the interactions with aerosol or sea spray (Wu et al. 2015).

To obtain a reliable and skillful ensemble, the representation of initial and model uncertainty in ensemble forecasts is crucial. In terms of model uncertainty, it is expected that improvements will come from introducing physically more consistent model uncertainty schemes and perturbations to the numerical core (Leutbecher et al. 2017). The initial conditions of the ensemble are currently constructed by adding perturbations constructed from the EDA to a single analysis. Starting the ensemble forecasts directly from the EDA members has potential benefits (Lang et al. 2015), especially in situations with strong gradients such as in tropical cyclones. However, this will require running the EDA with comparable resolution and quality to the high-resolution analysis. It remains to be seen if this will be computationally feasible in the future. Recent results have shown that it is also possible to improve ensemble forecasts by running a small number of high-quality analyses, which are then combined with the lower-resolution EDA to initialize the ensemble forecasts (Holm et al. 2018).

Another challenge for ensemble forecasting is to find the right balance between resolution and number of ensemble members (Leutbecher 2019).

In summary, to create a reliable, high-resolution ensemble for tropical cyclones, it is essential to make the optimal use of existing observations in a sophisticated and efficient data assimilation system, and also to have a reliable estimate of the initial uncertainties. A model with suitable model physics and dynamics, with a sufficient addition of model uncertainties, is needed to simulate the evolution of the tropical cyclones. To create such a forecast, all aspects of numerical weather prediction development are needed in a joint effort. While we have demonstrated that it is possible to run a 5-km global ensemble forecast, major challenges lie in computational science to make the code efficient enough to be run operationally, both for the model and the data assimilation (www.ecmwf.int/en/about/what-we-do/scalability). If all this can be achieved, then, as we have shown, there is the potential to make major advances in the prediction of tropical cyclone tracks and intensity.

ACKNOWLEDGMENTS. We first would like to thank all the scientists at ECMWF who have contributed to the developments behind the results presented here but are not explicitly listed as coauthors. We would also like to acknowledge Dr. Nils Wedi and Prof. Sharan Majumdar, and three anonymous reviewers whose comments helped to improve this article. A special thanks to Anabel Bowen for help with preparing the graphics for this article.

REFERENCES

Anderson, E., and H. Jarvinen, 1999: Variational quality control. Quart. J. Roy. Meteor. Soc., 125, 697–722, https://doi.org/10.1002/qj.49712555416.
Barkmeijer, J., R. Buizza, T. N. Palmer, K. Puri, and J.-F. Mahfouf, 2001: Tropical singular vectors computed with linearized diabatic physics. Quart. J. Roy.

\(^1\) The gray zone refers to horizontal resolutions of about 5 km, for which the grid is still too large for the convective cells to be well sampled but too small for the statistical approach of most parameterizations to be valid.
De Chiara, G., M. Bonavita, and S. J. English, 2017: Improving the assimilation of scatterometer wind observations in global NWP. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 10, 2415–2423, https://doi.org/10.1109/JSTARS.2017.2691011.

Diamantakis, M., and L. Magnusson, 2016: Sensitivity of the ECMWF model to semi-Lagrangian departure point iterations. *Mon. Wea. Rev.*, 144, 3233–3250, https://doi.org/10.1175/MWR-D-15-0432.1.

ECMWF, 2016: ECMWF strategy 2016–2025: The strength of a common goal. ECMWF Tech. Rep., 27 pp., www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf.

Emanuel, K., and F. Zhang, 2016: On the predictability and error sources of tropical cyclone intensity forecasts. *J. Atmos. Sci.*, 73, 3739–3747, https://doi.org/10.1175/JAS-D-16-0100.1.

Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, 94, 329–343, https://doi.org/10.1175/BAMS-D-12-00071.1.

Haiden, T., M. Janousek, J. Bidlot, L. Ferranti, F. Prates, F. Vitart, P. Bauer, and D. S. Richardson, 2017: Evaluation of ECMWF forecasts, including 2016-2017 upgrades. ECMWF Tech. Memo. 817, 56 pp., www.ecmwf.int/en/elibrary/17913-evaluation-ecmwf-forecasts-including-2016-2017-upgrades.

Heming, J. T., 2016: Met Office Unified Model tropical cyclone performance following major changes to the initialization scheme and a model upgrade. *Wea. Forecasting*, 31, 1433–1449, https://doi.org/10.1175/WAF-D-16-0040.1.

Hólm, E. V., S. T. K. Lang, M. Fisher, T. Kral, and M. Bonavita, 2018: Distributed observations in meteorological ensemble data assimilation and forecasting. 21st Int. Conf. on Information Fusion, Cambridge, United Kingdom, IEEE, 92–99, https://doi.org/10.23919/ICIF.2018.8455209.

Janssen, P. A. E. M., and J.-R. Bidlot, 2018: Progress in operational wave forecasting. *Symp. on Wind Waves*, London, United Kingdom, International Union for Theoretical and Applied Mechanics, www.science.direct.com/journal/procedia-iutam/vol/26/suppl/C.

Judt, F., and S. S. Chen, 2016: Predictability and dynamics of tropical cyclone rapid intensification deduced from high-resolution stochastic ensembles. *Mon. Wea. Rev.*, 144, 4395–4420, https://doi.org/10.1175/MWR-D-15-0413.1.

Kepert, J. D., 2012: Choosing a boundary layer parameterization for tropical cyclone modeling. *Mon. Wea. Rev.*, 140, 1427–1445, https://doi.org/10.1175/MWR-D-11-00217.1.

Laloyaux, P., M. Balmaseda, D. Dee, K. Mogensen, and P. Janssen, 2016: A coupled data assimilation system
for climate reanalysis. Quart. J. Roy. Meteor. Soc., 142, 65–78, https://doi.org/10.1002/qj.2629.

Lang, S. T. K., M. Leutbecher, and S. C. Jones, 2012: Impact of perturbation methods in the ECMWF Ensemble Prediction System on tropical cyclone forecasts. Quart. J. Roy. Meteor. Soc., 138, 2030–2046, https://doi.org/10.1002/qj.1942.

——-, M. Bonavita, and M. Leutbecher, 2015: On the impact of re-centring initial conditions for ensemble forecasts. Quart. J. Roy. Meteor. Soc., 141, 2571–2581, https://doi.org/10.1002/qj.2543.

Leutbecher, M., 2019: Ensemble size: How suboptimal is less than infinity? Quart. J. Roy. Meteor. Soc., https://doi.org/10.1002/qj.3387, in press.

——-, and T. N. Palmer, 2008: Ensemble forecasting. J. Comput. Phys., 227, 3515–3539, https://doi.org/10.1016/j.jcp.2007.02.014.

——-, and Coauthors, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. Quart. J. Roy. Meteor. Soc., 143, 2315–2339, https://doi.org/10.1002/qj.3094.

Madec, G., 2008: NEMO ocean engine. Institut Pierre-Simon Laplace Note du Pôle de Modélisation 27, 386 pp., www.nemo-ocean.eu/bibliography/documentation/.

Magnusson, L., J.-R. Bidlot, S. T. K. Lang, A. Thorpe, N. Wedi, and M. Yamaguchi, 2014: Evaluation of medium-range forecasts for Hurricane Sandy. Mon. Wea. Rev., 142, 1962–1981, https://doi.org/10.1175/MWR-D-13-00228.1.

McNally, T., M. Bonavita, and J.-N. Thepaut, 2014: The role of satellite data in the forecasting of Hurricane Sandy. Mon. Wea. Rev., 142, 634–646, https://doi.org/10.1175/MWR-D-13-00170.1.

Mogensen, K. S., L. Magnusson, and J. Bidlot, 2017: Tropical cyclone sensitivity to ocean coupling in the ECMWF Coupled Model. J. Geophys. Res. Oceans, 122, 4392–4412, https://doi.org/10.1002/2017JC012753.

Mulholland, D. P., P. Laloyaux, K. Haines, and M. A. Balmaseda, 2015: Origin and impact of initialization shocks in coupled atmosphere–ocean forecasts. Mon. Wea. Rev., 143, 4631–4644, https://doi.org/10.1175/MWR-D-15-0076.1.

Munch RE, 2018: A stormy year—Natural catastrophes 2017. Munich RE Tech. Rep., 64 pp., www.munichre.com/topics-online/en/climate-change-and-natural-disasters/natural-disasters/topics-geo-2017.html.

Pasch, R. J., A. B. Penn, and R. Berg, 2018: Hurricane Maria. National Hurricane Center Tropical Cyclone Rep. AL152017, 48 pp., www.nhc.noaa.gov/data/tcr/AL152017_Maria.pdf.

Rabier, F., H. Jarvinen, E. Klinker, J. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. Quart. J. Roy. Meteor. Soc., 126, 1143–1170, https://doi.org/10.1002/qj.49712656415.

Rodwell, M., L. Ferranti, T. Haiden, and L. Magnusson, 2015: New developments in the diagnosis and verification of high-impact weather forecasts. ECMWF Tech. Memo. 759, 44 pp., www.ecmwf.int/en/elibrary/15255-new-developments-diagnosis-and-verification-high-impact-weather-forecasts.

Trahan, S., and L. Spirling, 2012: An analysis of NCEP Tropical Cyclone Vitals and potential effects on forecasting models. Wea. Forecasting, 27, 744–756, https://doi.org/10.1175/WAF-D-11-00063.1.

Velden, C., and Coauthors, 2006: The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. Bull. Amer. Meteor. Soc., 87, 1195–1210, https://doi.org/10.1175/BAMS-87-9-1195.

Vitart, F., 2017: Madden–Julian oscillation prediction and teleconnections in the 52S database. Quart. J. Roy. Meteor. Soc., 143, 2210–2220, https://doi.org/10.1002/qj.3079.

——-, and F. Molteni, 2010: Simulation of the Madden–Julian oscillation and its teleconnections in the ECMWF forecast system. Quart. J. Roy. Meteor. Soc., 136, 842–855, https://doi.org/10.1002/qj.623.

Wu, L., A. Rutgersson, E. Sahlee, and X. G. Larsen, 2015: The impact of waves and sea spray on modelling storm track and development. Tellus, 67A, 27967, https://doi.org/10.3402/tellusa.v67.27967.

Yamaguchi, M., F. Vitart, S. T. K. Lang, L. Magnusson, R. L. Elsberry, G. Elliott, M. Kyouda, and T. Nakazawa, 2015: Global distribution of the skill of tropical cyclone activity forecasts on short- to medium-range time scales. Wea. Forecasting, 30, 1695–1709, https://doi.org/10.1175/WAF-D-14-00136.1.

——-, J. Ishida, H. Sato, and M. Nakagawa, 2017: WGNE intercomparison of tropical cyclone forecasts by operational NWP models: A quarter century and beyond. Bull. Amer. Meteor. Soc., 98, 2337–2349, https://doi.org/10.1175/BAMS-16-0133.1.

Zhang, F., and D. Tao, 2013: Effects of vertical wind shear on the predictability of tropical cyclones. J. Atmos. Sci., 70, 975–983, https://doi.org/10.1175/JAS-D-12-0133.1.

Zuo, H., M. Alonso-Balmaseda, K. Mogensen, and S. Tietscke, 2018: OCEAN5: The ECMWF ocean reanalysis system and its real-time analysis component. ECMWF Tech. Memo. 821, 44 pp., www.ecmwf.int/sites/default/files/elibrary/2018/18519-ocean5-ecmwf-ocean-reanalysis-system-and-its-real-time-analysis-component.pdf.