18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Verification of 1 km ensemble wind predictions

Katzfey, J.J. [1] **J. McGregor**[1]**, M. Thatcher**[1] **and B. Ebert**[2]

[1] *Centre for Australian Weather and Climate Research (CAWCR), CSIRO Marine and Atmospheric Research (CMAR), PMB #1, Aspendale, Victoria, Australia*
Email: jack.katzfey@csiro.au
[2] *Centre for Australian Weather and Climate Research (CAWCR)*

Abstract: The America's Cup is a sailboat race over a 5 km course that lasts about 2.5 hours. Accurate prediction of the wind speed and direction is essential for this race. Not only are the mean wind speed and direction important, but spatial and temporal variations greater than 1 knot (0.5 m/s) and 5 degrees are also significant.

Ensembles of high-resolution numerical weather predictions were performed down to 1 km resolution over a two-year period for the Mediterranean waters near Valencia, Spain. These ensembles were based upon different initial conditions, different atmospheric models and different large-scale predictions. An array of buoys was also placed around the race course, with 3 m masts providing 15 second wind speed and direction observations. The winds in this area are dominated by afternoon sea breezes, influenced by the local orography and modulated by the large-scale weather pattern, both at the surface and the lower atmosphere.

Various verification techniques were used to assess the accuracy of the forecast ensemble members, as well as the usefulness of the ensembles in predicting forecast skill/uncertainty. The verification techniques needed to account for the local and large-scale influences on the winds, as well as the relative contribution of the un-forced, less predictable mesoscale motions. Traditional and neighbourhood (multi-scale) verification approaches provided information about the space and time scales for which the forecasts were skillful.

Keywords: *Numerical weather prediction, ensemble, verification*

1. INTRODUCTION

Validation of high-resolution wind forecasts is becoming more important as the resolution of atmospheric models becomes finer. Instead of simply ensuring similarity with the observed large-scale weather patterns, it is possible to generate fine details, which may have particular importance to some applications. Often, the observational weather station network is insufficient to capture these fine-scale features, making the validation of the high-resolution forecast more challenging.

In this study, high-resolution ensemble atmospheric model predictions for winds over the Mediterranean Sea near Valencia, Spain are described. These forecasts were generated in order to support the Swiss sailing team, Alinghi, in its defence of the America's Cup in 2007. The model and observational data sources provide a unique opportunity to test various verification techniques. A description of the forecast requirements, the dense network of instruments that was set up to provide detailed observations around the race course and the model runs completed are given below, followed by results of the various verification techniques. Finally, the strengths and weaknesses of the various techniques are discussed.

2. FORECAST REQUIREMENTS

There were several stages involved in producing high-resolution wind forecasts for the America's Cup. In the morning a general reading of the forecast conditions was prepared for the team briefing at 11 am. This was based on the daily model forecasts, providing an overview of the ensemble predictions for both wind speed and direction for the race course throughout the day. It included an hour-by-hour forecast of the wind, as well as a general forecast for other factors such as rain and temperature. Particular focus was on times when any changes might be expected. All this information was used by the team to configure the boat for optimal performance.

Following the morning briefing, the weather team went out to the race area on weather boats, which were anchored near the race course, to ensure that the communication and observing systems were working properly and to compare and update the forecast with the observations every half hour. Note, however, that the observations prior to the development of the sea breeze were not necessarily indicative of conditions later in the day.

A key limitation in forecasting for an America's Cup race is that the race boat can only communicate with the outside world up to five minutes before the start of the race. Because of this rule, the last updated forecast by the weather team had to be communicated to the race boat approximately 6 minutes before the start of the race. The forecast, based on both the ensemble model predictions and the observations, included the expected wind speed and direction trends during the first leg of the race (about 25 minutes duration), along with any expected wind speed and direction changes and trends for the whole race, which lasted roughly 2.5 hours. Some indication of the confidence of our predictions and the magnitude of anticipated variations was also given. Since small time- and space-scale wind shifts are inherently unpredictable, as they are related to boundary-layer turbulence, the forecast gave the mean wind and its range of variability in order to assess the expected possible shifts.

3. OBSERVATIONS

In order to provide weather observations to all the teams participating in this America's Cup series, a Meteorological Data Service (MDS), was set up by race management. This consisted of 21 buoys with 3 m masts located on the water, distributed mostly on the race course but also further afield (see Figure 1). There were also several land-based stations and weather boats gathering data. The observations were primarily wind speed and direction, recorded every 15 seconds. At certain buoys, pressure, temperature and relative humidity were also recorded every minute. The MDS also provided a wind profiler, located near the race course, and satellite and radar images to assess the amount of clouds and any precipitation in the area.

4. MODELLING

A suite of nine high-resolution forecasts was set up in order to assess the range of possible wind conditions over the race course during the race. The high-resolution forecasts were based upon different initial conditions, different large-scale forecasts, and different models. The primary model used to generate the forecasts was the Conformal-Cubic Atmospheric Model (CCAM) from CSIRO (McGregor, 2003, McGregor and Dix, 2008). The WRF (Wang et al., 2004) and RAMS (Pielke et al., 1992) models were also set up locally, and the MDS provided 5 km forecasts from the HIRLAM model to all teams. Table 1 provides a summary of the forecasts completed each day.

For the CCAM forecasts, the analyses used were the US NCEP GFS analysis and the Canadian CMC analysis. Both were provided on a global 1° lat-long grid. The analysis data were interpolated to the CCAM model 60 km grid, and a five-day forecast was then completed. The 8 km grid CCAM forecasts were run for three days using far-field nudging from the 60 km forecasts. The two-day 1 km CCAM forecasts were also completed using far-field nudging from the 8 km forecast. Two sets of CCAM forecasts were generated using the global forecast from either the GFS or CMC global forecasts as large-scale forcing for the 8 km CCAM forecasts. The MDS also provided a subset of the ECMWF global analysis. This data was merged with the GFS global analyses to provide initial conditions for the 60 km CCAM forecasts. The last set of CCAM forecasts was based upon using different model settings of the CCAM model initialized using the GFS analysis.

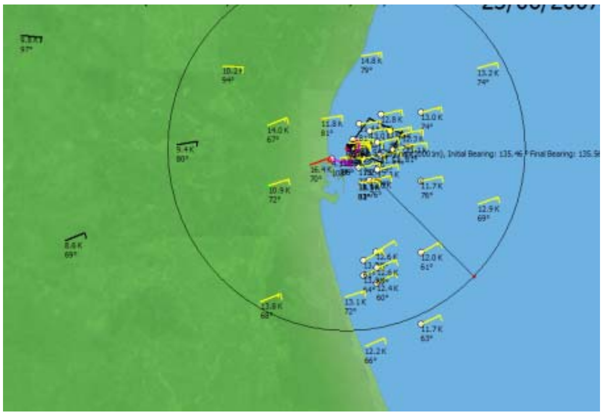

Figure 1. Distribution of wind observations in Valencia, Spain on 23 June, 2007. Below each wind observation is wind speed (knots) and wind direction (degrees). The circle has 10 km radius.

Table 1. Range of simulations completed each day indicating model, data and various resolutions completed. IC=Inital Conditions, BC=Boundary Conditions

Initial UTC	Model	IC	BC	Resolutions			
00/06/12/18	CCAM	GFS		60 km	8 km	1 km	
00/12	CCAM	CMC		60 km	8 km	1 km (CMC1)	
00/12	CCAM	GFS	GFSf		8 km	1 km	
00/12	CCAM	CMC	CMCf		8 km	1 km	
00/12	CCAM	ECMWF/GFS		60 km	8 km	1 km	
00/12	CCAM w/ alt. setting	GFS		60 km	8 km	1 km	
00/06/12	HIRLAM	HIRLAM			5 km		
00/12	WRF	GFS	GFSf	64 km	16 km	4 km	
00/12	RAMS	GFS	GFSf	64 km	16 km	4 km	1 km

The HIRLAM forecasts were generated by the Spanish meteorology service. The WRF and the RAMS model forecasts were initialized from the GFS analysis and forecasts (which provided lateral boundary conditions for the outermost grid). The resolutions used were 64 km, 16 km grid, and 4 km. The RAMS model was also run at 1 km.

A sample of a daily high-resolution ensemble forecast for wind speed and direction over the race course for 3 July, 2007 is shown in Figure 2. The time period is from 8 am to 7 pm local time (LT). Note the very light winds of less than 6 knots prior to around 10 or 11 am LT. Typically, the morning winds were offshore from an overnight land breeze. As can be seen from the wind direction graphs, the wind shifted to the south-east around noon and the wind speed began to increase as the sea breeze developed. Most forecasts showed a consistent wind direction of around 135° between 1 pm and 4 pm LT. Later in the day the wind shifted to a more easterly direction, and even northeasterly in some of the model forecasts. The strengths of the breeze varied markedly between the various forecasts, ranging from around 13 to 21 knots. Again, most models showed a decrease in wind speed after 5 pm LT as the wind trended towards the northeast. The observations showed winds from 140° at 16 knots in the afternoon, with a sharp drop in speed and a shift to the northeast around 4:30 pm LT.

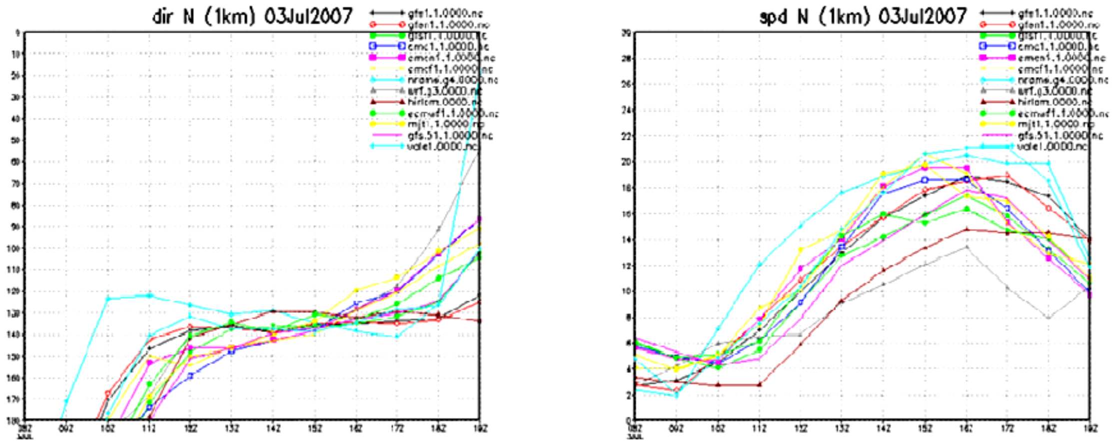

Figure 2. Sample 1 km ensemble forecasts for a) wind direction (degrees) and b) wind speed (knots) for forecasts initialized at 2 am LT, 3 July 2007.

5. VERIFICATION

The initial verification was done using the CCAM model initialized with the analyses for one-day forecasts generated for the years 1998-2003 for the months of April through September. The frequency distributions of the model forecast wind conditions were compared to observations at the Port of Valencia (see Figure 3) for the six months analysed for the period between 12 pm and 6 pm LT. The wind speed distribution shows close alignment of the model and observed values, with a slight over-forecast of lighter winds and an

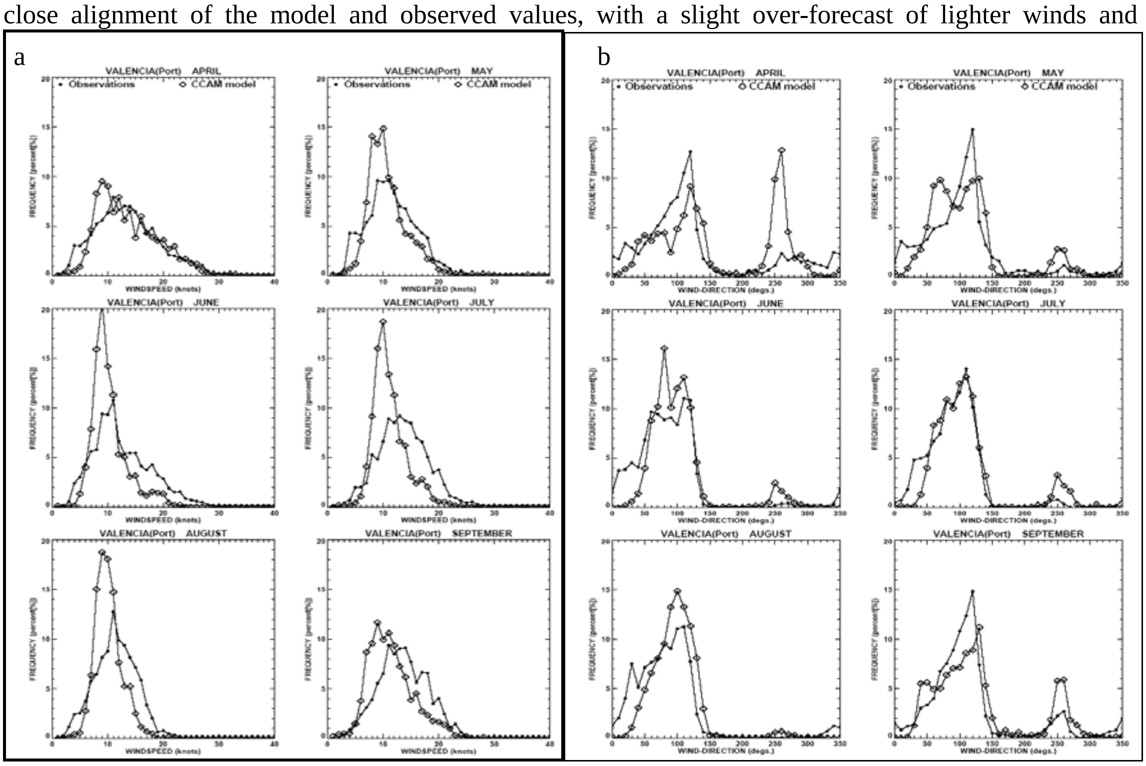

Figure 3. Frequency distribution of a) wind speed (knots) and b) wind direction (degrees). Observations line has closed dots, model has open circles.

underprediction of heavier winds. Note that the wind distribution for April was broader, indicative of the more frequent heavier wind conditions.

The wind direction distribution again shows close correspondence between the model and the observations. The predominant wind direction, related to the sea breeze, is from east to southeast for these months. There is a second peak for westerly winds in the frequency distribution. The model tends to slightly overpredict the frequency of these events compared to the observations. Also note that April has more westerly wind events. Partly based upon this analysis, Alinghi picked Valencia as the venue for the last America's Cup. Comparing the frequency distributions shows that the general characteristics of the forecasts match the observations for the various wind conditions experienced in Valencia during these months.

Bias and mean absolute error are commonly used as verification techniques to assess model performance. Verification using these techniques for a select number of the high-resolution model forecasts for the summer of 2006 is presented in Figure 4. The Figure shows the bias (top two graphs) and mean absolute error (bottom two graphs). On the left are the graphs for wind speed, while on the right are the graphs for wind direction. The verification is a comparison of the nearest model grid point to each buoy on the race course. In general, the CMC 1 km forecasts (CMC1, see in Table 1) had the smallest bias and mean absolute error for wind speed. Wind direction bias for all models is between 0-5°. The mean absolute error is between 10-15°. Again, the CMC1 forecast tends to have the smallest mean absolute error for wind direction. These verification techniques indicate general accuracy of the various forecasts, how it varies over time, and the relative performance of the various ensemble members.

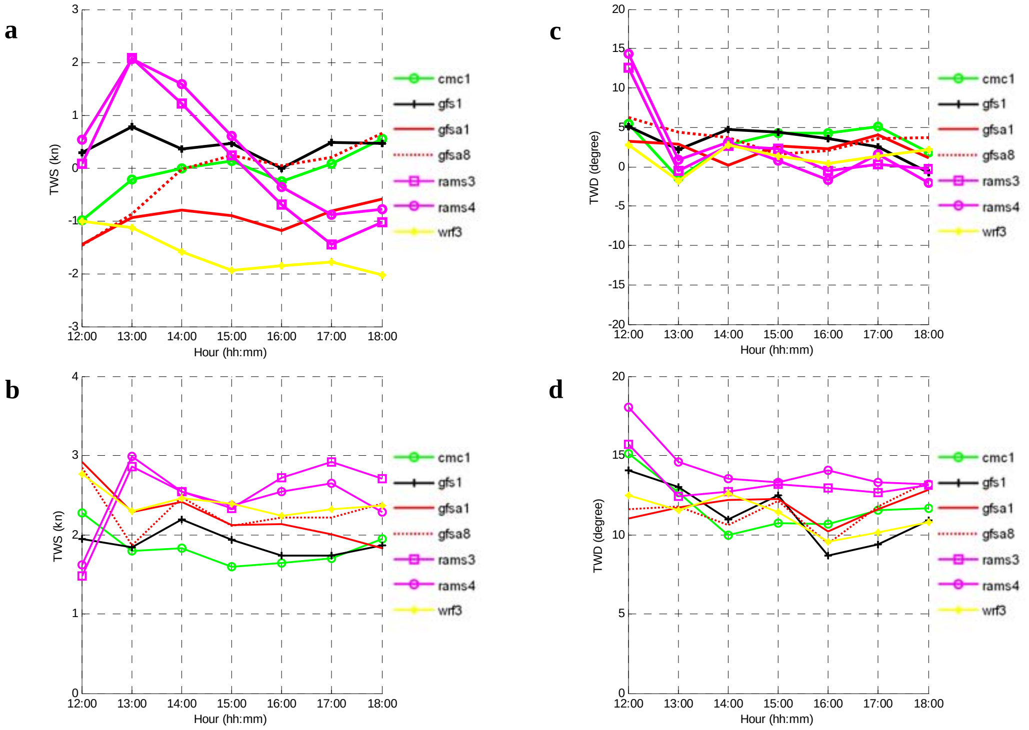

Figure 4. Verification of wind speed (knots) and wind direction (degrees) for various model forecasts for May-Aug 2006 for Noon-6 pm local time. a) wind speed bias , b) mean absolute error of wind speed, c) wind direction bias, d) mean absolute error of wind direction.

Users of numerical forecasts often interpret the model output in a probabilistic way, rather than taking the predictions strictly at face value. This approach acknowledges that forecast weather features are rarely exactly correct but are frequently close enough to be of value. An evaluation methodology known as "neighbourhood verification" provides information on how far in space (and possibly time) the user can expect to find forecast values that match the observation, by evaluating forecasts within a space/time neighbourhood surrounding each observation. Ebert (2008, 2009) reviewed a number of neighbourhood methods and used them to evaluate high-resolution precipitation forecasts.

The *conditional square root of RPS* (CSRR) described by Germann and Zawadzki (2004) is a neighbourhood method that is particularly relevant for measuring the usefulness of high-resolution forecasts of wind speed and direction. Based on the rank probability score (RPS) used to verify probabilistic forecasts, this approach measures the accuracy and sharpness of the distribution of forecast values in the neighbourhood. Given a set of wind speed or direction categories, the number of forecast values in the neighbourhood in each category is an estimate of the probability of that category. The square root of the RPS is like a standard error of the estimated category. (For rainfall verification, Germann and Zawadzki (2004) normalised the score by the observed rain fraction; for wind verification, this step is unnecessary.)

To illustrate the methodology we apply the CSRR to evaluate an individual forecast shown in Figure 5. The forecast shows the sea breeze front aligned along the coast, whereas the observations indicate that the sea breeze has already penetrated inland by this time. The categories used for verifying wind speeds were 1, 2, 5, 10, 15, 20, 25, 30, 40, and 50 knots; for verifying wind direction eighteen 20º bins were used. Fig. 6 shows that the CSRR is optimised at a horizontal scale of 0.19º for wind speed and 0.07º for wind direction, the latter being approximately the error in the location of the sea breeze front. The small values of CSRR for wind speed indicate that the forecast values were off by only about half a category, whereas the average wind direction distribution was in error by roughly two categories (40º angle).

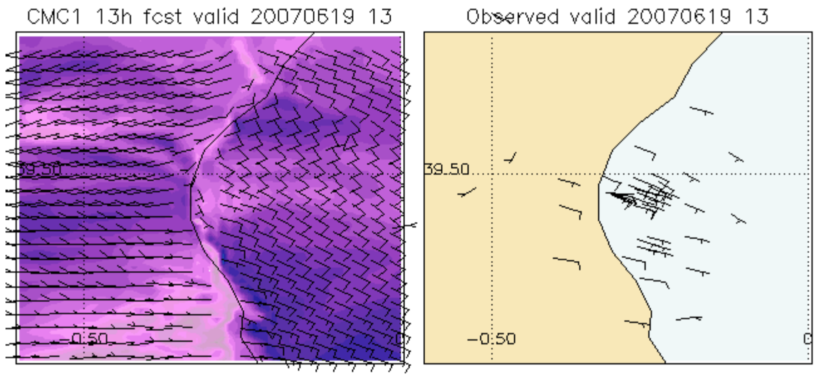

Figure 5. 13h forecast (left) and observed wind (right) at 13 UTC (3 pm LT) on 19 June 2007.

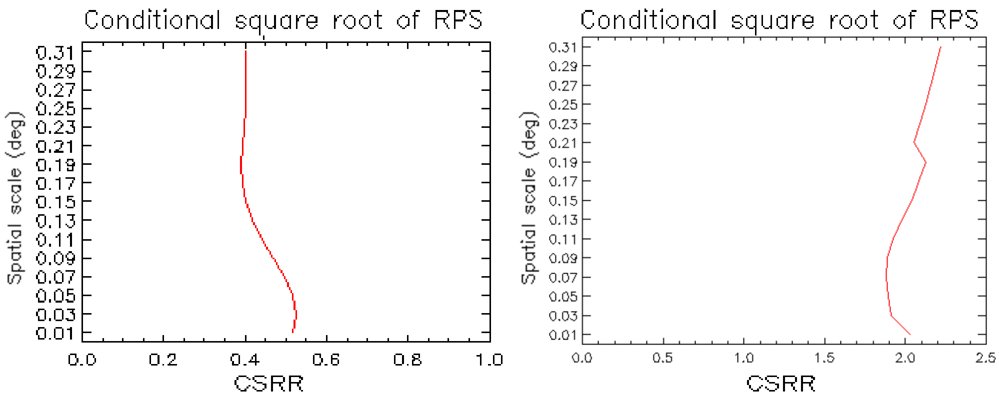

Figure 6. Neighbourhood verification of forecasts of wind speed (left) and wind direction (right) for CMC1 forecasts initialized at 00 UTC and valid at 13 UTC (3 pm LT) on 19 June 2007. The size of the neighbourhoods increases from 1x1 to 31x31 along the vertical axis.

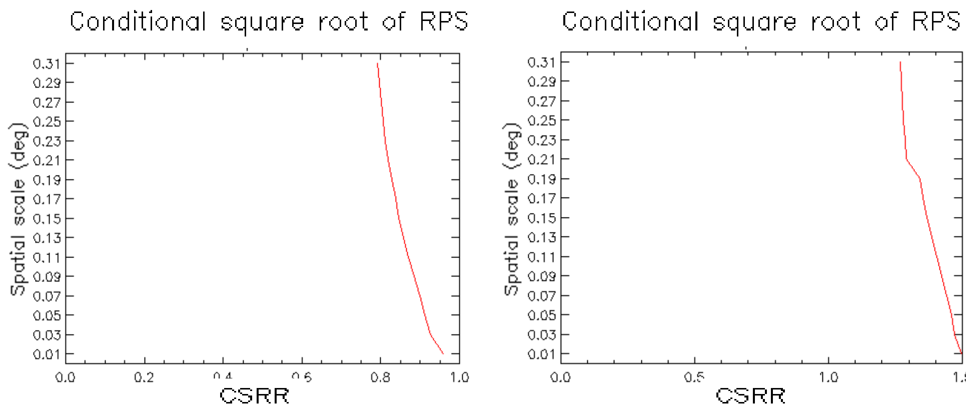

Figure 7. Neighbourhood verification of forecasts of wind speed (left) and wind direction (right) for CMC1

forecasts initialized at 00 UTC and valid at 13 UTC (3 pm LT) during June 2007.

When aggregated for the full month of June 2007 (Figure 7), overall the CSRR for wind speed and direction improves as the spatial scale of the neighbourhood is increased. At the horizontal grid scale of the model (0.01º), the wind speeds were typically in error by less than 1 category, while the directions were out by about 1.5 categories (30º). By looking in a larger neighbourhood around the point of interest, the broader distribution of values gave a better prediction for both wind speed and direction. As the curves have not asymptoted at the largest scale, it is not possible to determine an optimal scale for interpreting wind forecasts in this (relatively small) domain.

6. DISCUSSION

In this study, verification of high-resolution wind forecasts used for the 2007 America's Cup is presented. These forecasts were initialized with global analyses provided by the US and Canadian meteorological services, with no local data assimilation. The high-resolution forecasts were validated against a unique dense observational network using frequency distributions, bias and mean absolute error, and neighbourhood techniques.

Validating frequency distributions gives a clear picture of whether or not the forecasts are capturing a realistic range of atmospheric phenomena. In this case, the distribution of wind direction was very accurate, while there was a slight under-prediction of the stronger wind events. However, this technique does not tell how accurate an individual forecast may be or how the accuracy varies over time. Computing biases and mean absolute errors over time for the various ensemble members gives a much clearer picture of the relative accuracy of the forecasts and the strengths and weaknesses of individual model configurations. It was also clear that some forecasts had significant time-varying biases. However, these two types of validation require a one-to-one comparison of observations with the model, so do not take into account spatial errors in the forecasts. Neighbourhood techniques can potentially quantify the accuracy of the forecasts in spatial terms, and can also quantify the impact of slightly displaced fields, but care must be taken if the spatial fields are relatively uniform.

In summary, the validation techniques tested in this study produced useful information of various kinds, but no single technique is able to completely determine forecast skill or uncertainty. Traditional validation techniques provide measures against point observations, while neighbourhood verification methods quantify the accuracy of distinct features. For this sailboat race, where determination of small-scale features and short time-scales was necessary, determining the accuracy of the forecasts on the race course over time was most important, a statistic provided by the mean absolute error. Future work is needed on validating the utility of ensemble predictions as well as the cause(s) of the errors, especially for less-predictable mesoscale features.

ACKNOWLEDGMENTS

Part of this work was completed as part of a contract between Alinghi and CSIRO.

REFERENCES

Ebert, E. E., (2008), Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteorol. Appl.*, 15, 51-64.

Ebert, E. E., (2009), A framework for neighbourhood verification of high resolution spatial forecasts. *MODSIM09, Cairns, Australia, 13-17 July 2009.*

Germann, U. & Zawadzki, I. (2004), Scale dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts. *J. Appl. Meteorol.* 43: 74-89.

McGregor, J. L. (2003), C-CAM: Geometric aspects and dynamical formulation [electronic publication]. *CSIRO Atmospheric Research Tech. Paper No. 70*, 43 pp.

McGregor, J. L., and Dix, M.R. (2008), An updated description of the Conformal-Cubic Atmospheric Model. In *High Resolution Simulation of the Atmosphere and Ocean*, eds. K. Hamilton and W. Ohfuchi, Springer, 51-76.

Pielke, R. A., Cotton, W. R., Walko, R. L. , Tremback, C. J., Lyons, W. A., Grasso, L. D., Nicholls, M. E., Moran, M. D., Wesley, D. A., Lee, T. J. and Copeland, J. H. (1992), A comprehensive meteorological modeling system – RAMS, *Meteorol. Atmos. Phys.*, *49*, 69–91.

Wang, W., Barker, D., Bruyère, C., Dudhia, J., Gill, D. and Michalakes, J. (2004), WRF Version 2 modeling system user's guide. http://www.mmm.ucar.edu/wrf/users/docs/user guide/.