Reprint 1064

Objective Consensus Forecast: A site-specific multi-model consensus forecast in the Hong Kong Observatory

S.T. Chan & Jeffrey C.W. Lee

27th Guangdong-Hong Kong-Macao Seminar on

Meteorological Science and Technology,

Shaoguan, Guangdong, 9-10 January 2013

Objective Consensus Forecast: A site-specific multi-model consensus forecast system of the Hong Kong Observatory

Chan Sai-tick Lee Chi-wai, Jeffrey
Hong Kong Observatory

Abstract

This paper presents the development of Objective Consensus Forecast (OCF), a new site-specific multi-model consensus forecast system of the Hong Kong Observatory (HKO). OCF is a past-performance weighted multi-model consensus forecast system based on the outputs of a number of global numerical models from ECMWF, JMA and NCEP, as well as Meso-NHM, the operational regional model run at HKO. It is found that the forecasts generated by OCF generally outperform those from the member models. OCF has been put on operational trial since the beginning of the winter of 2011 in HKO.

客观集成预报 – 香港天文台的定点多模式集成预报系统

陈世倜 李子维
香港天文台

摘要

本文简介香港天文台新近发展的一个定点多模式集成预报系统 － 客观集成预报(Objective Consensus Forecast，简称 OCF)。OCF 是一个以过去表现定比重、以多个包括来自欧洲中期天气预报中心（ECMWF）、日本气象厅(JMA)、美国国家环境预测中心(NCEP)的全球数值预报模式输出，加上香港天文台的业务区域模式(Meso-NHM)输出作为基础的集成预报系统。该预报系统的预测普遍比单一模式的预测准确。OCF 于 2011 年的冬季开始在香港天文台作业务试行。

1 Introduction

In the past decades, there has been significant improvement in the accuracy of numerical weather prediction (NWP) due to continual improvement in the model physics, model resolution, and the data assimilation process. However, the accuracy of direct model outputs (DMO) from NWP is still insufficient for direct use in the public weather forecast, due to the mismatch in representativeness between the model grid points and actual locations they represent. In this regard, post-processing of the model outputs is generally needed [1]. Kalman filtering (KF) is a common method for such purpose as it is adaptive (i.e. there is no need to re-train the post-processing algorithm upon model upgrades), and it requires a relatively short training period for quick adaptation to changes in the synoptic conditions [2].

It is generally recognised that the forecast from a multi-model ensemble is usually better than that from a single model [3]. Due to the differences in the parameterization schemes adopted in various models, individual models usually have situation-dependent biases. A multi-model ensemble delivers better performance by partially removing these biases through averaging.

Following on the idea of a similar forecast system developed by the Bureau of Meteorology, Australia [4], the Hong Kong Observatory (HKO) combined the above approaches and developed the Objective Consensus Forecast (OCF). OCF is a site-specific, Kalman-filtered and past-performance weighted multi-model consensus forecast. Currently, 5 weather elements are available in the system, namely, temperature (T), dew point (T_d), relative humidity (RH), wind speed (v_s) and wind direction (v_d). Forecasts for these weather elements are generated for a number of locations at automatic weather stations (AWS) within Hong Kong.

In this paper, the formulation of the OCF and its performance are presented. In Section 2, the model data inputs and post-processing methodology deployed for OCF are described. The performance of OCF for various weather elements, in comparison with that of the member models is presented in Section 3. Section 4 gives the summary, along with a discussion on the planned enhancements of the system in future.

2 Methodology and model data of OCF

Currently, outputs from five different NWP models are used in OCF, including:

(i) Meso-NHM,

(ii) ECMWF deterministic model,

(iii) JMA GSM,

(iv) NCEP GFS,

(v) ECMWF Ensemble Prediction System (EPS).

Meso-NHM is the operational limited area model of HKO [5]. It has a horizontal resolution of 10 km and it provides forecast up to 72 hours ahead. The boundary conditions for Meso-NHM are obtained from the half-degree grid data of JMA GSM. For ECMWF EPS, the forecasts from all 51 ensemble members are averaged, i.e. the simple ensemble mean, for injection into the OCF. The current version of the OCF updates twice a day based on the model outputs initialized at 00Z and 12Z respectively.

Out of the five weather elements generated by OCF, T, T_d, v_s and v_d are post-processed according to the following steps:

(i) Spatial interpolation from the model grid to the locations of the AWSs.

(ii) Temporal interpolation to hourly data.

(iii) Applying the KF on the interpolated data.

(iv) Combining the Kalman-filtered data based on their past 30 days performance.

In step (i), DMO from various models are first interpolated to the locations of the AWS using bilinear interpolation. The various model grids and locations of the AWS are shown in Fig. 1.

Apart from Meso-NHM in which the hourly forecast data are available, all other model data come in a time step of either 3 or 6 hours. These 3-hourly or 6-hourly data are then interpolated into hourly data in step (ii) using linear interpolation. KF can then be applied on the hourly forecast data against the hourly observations from AWS in step (iii). For this purpose, the KF module from the statistical software R (http://www.r-project.org) is used in OCF. For T and T_d, a one-dimensional KF is employed whereas a two-dimensional KF is used for v_s and v_d, with KF applied on the zonal and meridional wind. The value of T_d after KF is capped by the corresponding value of T to ensure internal consistency.

The final step in OCF [step (iv)] is to combine the forecasts from various models into a single consensus forecast. This is achieved by the performance weighted-averaging (PWA) procedure, in which the weight of the forecast from a model is inversely proportional to its mean absolute error in the last 30 days [4]. While each member model of the OCF has a different forecast range (Table 1), the consensus forecast is constructed by taking all available forecasts from the member models. Currently, a total of four consensus forecasts based on different combinations of models have been routinely generated by OCF (Table 2).

Forecast for the last element, viz. RH, is derived from T and T_d following literature [6].

The interactive client-side interface of the OCF was developed using the Google Web Toolkit (GWT, https://developers.google.com/web-toolkit/). The whole system is composed of four parts:

(i) A central database for the storage of forecast data.
(ii) Background processes for routine update of the database.
(iii) A client-side javascript-based user-interface running on web browsers.
(iv) A server-side program to handle requests from clients and to retrieve data from the database.

Users can customize the display of outputs from OCF by selecting the forecast range, location, weather element, and the model data to be displayed (Fig. 2). Verification of past predictions from OCF can also be conducted through the same interface according to user-selected verification metric and data time range.

3 Performance evaluation of OCF

The system is still evolving. New models and weather elements are being added to OCF from time to time. For example, temperature forecast has been available on OCF since the very beginning while dew point and RH were added to the system in February 2012. The newest consensus forecast, namely, OCF-D which comprises all five NWP models, was only introduced in late March 2012. Yet the data availability from ECMWF EPS has not been very stable, in particular for the wind forecast data which have been missing for some major tropical cyclone events during the year. Because of this, the following discussion on verification results mainly focuses on the consensus forecasts constructed from the other four models, viz. Meso-NHM, ECMWF deterministic model, JMA GSM and NCEP GFS. As the performance of OCF-A, OCF-B and OCF-C are close to each other, only the results from OCF-C are presented in this paper for sake of clarity.

Performance statistics for three AWS stations, namely HKO, R2C (Hong Kong International Airport) and TKL (Ta Kwu Ling) (Fig. 1) have been chosen for discussion in this paper. HKO, located in the urban area, is representative in the sense that both the official temperature and RH forecasts issued for Hong Kong make reference to this station. R2C, located at the western part of Hong Kong next to the mouth of the Pearl River, is characterized by the prominent land-sea breeze circulation of the wind flow. TKL is situated near the northern border of Hong Kong and is representative of inland stations that are sheltered from strong winds.

3.1 Temperature

The verification results of the temperature forecasts for HKO based on the 12 UTC model runs during one whole year period from 1 December 2011 – 31 November 2012 are shown in Fig. 3. Comparing OCF-C with DMO of the four member models in terms of root mean square error (RMSE) [Fig. 3(a)], the improvement by OCF-C (the yellow line) is prominent across the whole forecast range, particularly during the afternoon periods, corresponding to the peaks near 2 p.m. local time at T+18h, T+42h, …, and T+234h.

Part of the improvement achieved in the consensus forecast can be attributed to the post-processing procedure by KF and the remaining improvement to the merging procedure through PWA. To gauge the effectiveness of the merging procedure, the skill score of each member model after KF against OCF-C, is computed and plotted in Fig. 3(b) according to the following formula:

$$Skill\ score = \frac{RMSE_{OCF} - RMSE_{\mathrm{mod}\,el}}{RMSE_{OCF}}$$

Negative score indicates that OCF achieves skill improvement through the merging procedure while positive score indicates no skill improvement. It is easy to see that OCF-C achieved better performance than the Kalman filtered forecasts from Meso-NHM, JMA GSM and NCEP GFS, and also better than ECMWF deterministic model in the early part and the later part of the forecast period. Between T+76h and T+161h, ECMWF after KF outperformed OCF-C but the differences are small. It is noted that ECMWF generally outperformed other member models with a large margin during this period, which may render the merging procedure relatively ineffective.

The lead by ECMWF gradually diminished from T+161h onward and OCF-C outperformed ECMWF again from then on.

Since Meso-NHM forecasts are only available up to T+72h, OCF-C becomes a three-member consensus after T+72h. The number of members in OCF-C continues to decrease with ECMWF deterministic model becoming the only available model from T+217h to T+240h, i.e. OCF-C becomes equivalent to the ECMWF forecast. Therefore, the skill score of the ECMWF forecast drops to zero from T+217h onward.

Despite the fact that the forecast data from JMA GSM are only available up to T+84h for the 00 UTC model runs, observations similar to the above are noted from the results based on the 00 UTC model runs (results omitted).

3.2 Relative humidity

The verification results for RH, instead of T_d are shown in Fig. 4, considering that the former is a more relevant quantity to the public. However, due to data availability issue, the verification time period is reduced to 1 March – 30 November 2012. Similar observations as for temperature forecast are noted in Fig. 4(a), where improvement by OCF-C over DMO was seen over the whole forecast period and the improvement was most prominent during the afternoon periods, partly due to the significant improvement also seen in the corresponding temperature forecasts. For skill comparison as shown in Fig. 4(b), it is obvious that OCF-C outperformed all individual member models including ECMWF deterministic model, the best member model among all, by about 5% or more over the most of the forecast period.

3.3 Wind speed and direction

The verification results of wind speed and wind direction forecasts for R2C during 1 December 2011 – 30 November 2012 are plotted in Fig. 5 and Fig. 6 respectively. For wind speed, OCF-C successfully corrected part of the errors in DMO of member models and the correction seems to be increasingly effective with increasing forecast hours [Fig. 5(a)]. After the merging process, OCF-C achieved yet better skill in general than all member models after KF, though the improvement was relatively mild and less than 5%. For wind direction, the improvement over DMO was also observed but not as obvious as for the other elements, which suggests that the model forecasts after KF may not be able to effectively capture the wind shifts associated with the land-sea breeze circulation in the diurnal cycle.

Comparing the skill of OCF-C with the model forecasts after KF [Fig. 6(b)], the former performed better than Meso-NHM and NCEP GFS but when compared with JMA GSM and ECMWF deterministic model, the results are mixed with JMA GSM performing better than OCF-C in the early part of the forecast range and ECMWF becoming the best performer at times in the latter part.

For TKL, an inland station sheltered from strong winds, the improvement of OCF over DMO in the wind speed forecast is significant due to the effective reduction of systematic positive biases in the model forecasts through KF [Fig. 7(a)]. OCF-C achieved further improvement through the merging procedure except during the very first part of the forecast range when ECMWF deterministic model has gained some advantage over the others. As shown in Fig. 8(a), the skill improvement by the consensus forecast over DMO for wind direction was more prominent when compared to R2C, and OCF-C also showed generally improved skills over the member models after KF.

3.4 OCF with ECMWF EPS

An additional test has been conducted by re-entering the temperature forecast data from ECMWF EPS to OCF so as to study the performance of the consensus forecast (OCF-D) with the ensemble mean of EPS data included. For the same data period as presented in Fig. 3, the skill score relative to OCF-D for all member models are plotted in Fig. 9. Further improvement was achieved by OCF-D over OCF-C with the former now delivering better performance than ECMWF deterministic model. It is interesting to note, however, that, ECMWF EPS was the most skillful model among all members and it has even outperformed OCF-D for the whole except the first 45 hours of the forecast range.

While a full evaluation of the performance of ECMWF EPS forecasts is needed, the above test suggests that simple inclusion of all available models in OCF may not necessarily produce the best consensus forecast.

4 Summary and discussion

The Hong Kong Observatory has developed a multi-model consensus forecast system named as OCF. Combining model outputs with different spatial and temporal resolutions, the system is able to generate hourly forecasts for a number of key weather elements at different locations within Hong Kong a week ahead or even longer. Verification of the consensus forecasts from OCF shows that the system can effectively remove the systematic biases observed in DMO through the post-processing procedure based on KF. The forecast accuracy can be further improved in general through the merging procedure according to the past performance of the member models, despite the fact that ECWMF deterministic model and EPS are indeed the dominating members in the consensus forecasts with their superior performance over other member models.

The system has been put into trial operation in the winter of 2011 in HKO and it has since become an indispensible tool for the weather forecasters. It will continue to evolve with the incorporation of new model data as they become available and the planned introduction of additional weather elements including cloud cover and precipitation. Work will also be conducted on the searching of an optimal combination of models for the construction of more accurate consensus forecasts, taking into account the superb performance delivered by the ECMWF models.

The OCF has indeed opened the possibility of providing location- and time-specific weather forecasts for specialized users as well as the members of public in future, supplementing the current official forecasts issued by HKO which are only generic forecasts applicable to the whole territory and to one whole day.

5 Acknowledgements

The authors would like to thank Mr. H.W. Leung and Mr. K.C. Lee for their contributions in developing the system and preparing the datasets and figures used in this paper.

References

[1] Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting, *J. Appl. Meteor.,* **11**, 1203-1211

[2] Monache L.D., T. Nipen, Y. Liu, G. Roux, and R. Stull, 2011: Kalman Filter and Analog Schemes to Postprocess Numerical Weather Predictions, *Mon. Wea. Rev.,* **139**, 3554-3570

[3] Clemen, R. T., 1989: Combining forecasts: A review and annotated bibliography. *Int. J. Forecasting,* **5**, 559-583

[4] Woodcock, F., and C. Engel, 2005: Operational consensus forecasts. Wea. Forecasting, **20**, 101-111.

[5] Wong, W.K., 2010: Development of operational rapid update non-hydrostatic

NWP and data assimilation systems in the Hong Kong Observatory. The 3rd International Workshop on Prevention and Mitigation of Meteorological Disasters in Southeast Asia, 1-4 March 2010, Beppu, Japan.

[6] W.T. Wong, 1989: Comparison of algorithms for the computation of the thermodynamic properties of moist air. HKO Technical Note (Local) No. 51. (Available from http://www.hko.gov.hk/publica/tnl/tnl051.pdf.)

Table 1 Model data used in the OCF.

Model	Forecast Range	Temporal Resolution	Spatial Resolution
Meso-NHM	T+72h	1 hour	0.1^{o} (Fig. 1a)
ECMWF deterministic model	T+240h	3 hours up to T+144h 6 hours afterwards	0.125^{o} (Fig. 1b)
ECMWF EPS	T+240h	3 hours up to T+144h 6 hours afterwards	0.25^{o} (Fig. 1c)
JMA GSM	T+84h for 00Z T+216h for 12Z	3 hours up to T+84h 6 hours afterwards	0.25^{o} (Fig. 1c)
NCEP GFS	T+192h	3 hours	0.5^{o} (Fig. 1c)

Table 2 Four different consensus forecasts available from the OCF.

Abbreviation	Model combination
OCF-A	Meso-NHM + ECMWF deterministic model + JMA
OCF-B	ECMWF deterministic model + JMA + NCEP
OCF-C	Meso-NHM + ECMWF deterministic model + JMA + NCEP
OCF-D	Meso-NHM + ECMWF deterministic model + JMA + NCEP + ECMWF EPS (simple ensemble mean)

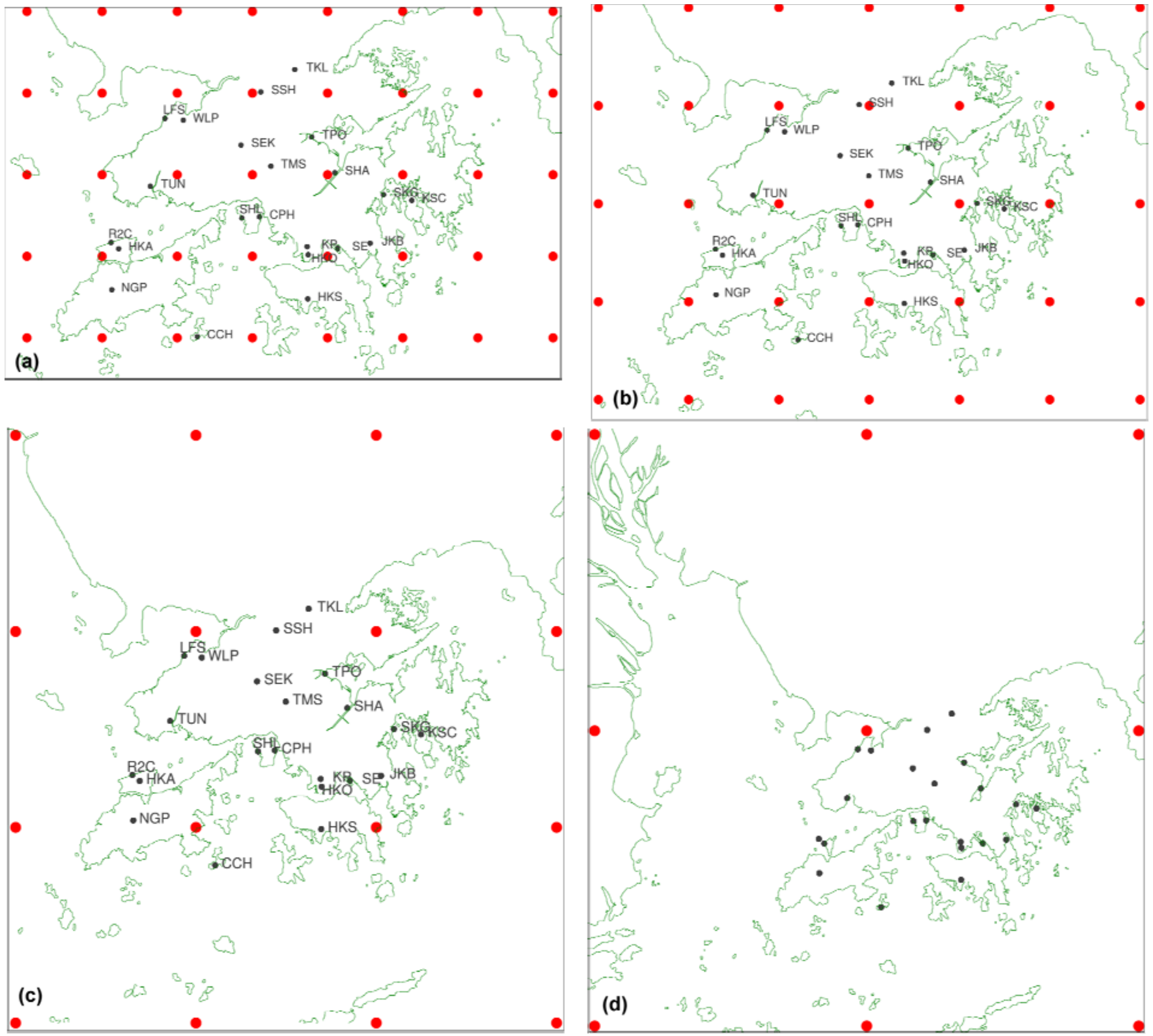


Fig. 1 AWS locations (black dots) and model grids (red dots) used for compiling the consensus forecasts in OCF. (a) Meso-NHM with resolution of 0.1^{o}; (b) ECMWF deterministic model with resolution of 0.125^{o}; (c) JMA GSM and ECWMF EPS with resolution of 0.25^{o}; and (d) NCEP GFS with resolution of 0.25^{o}.

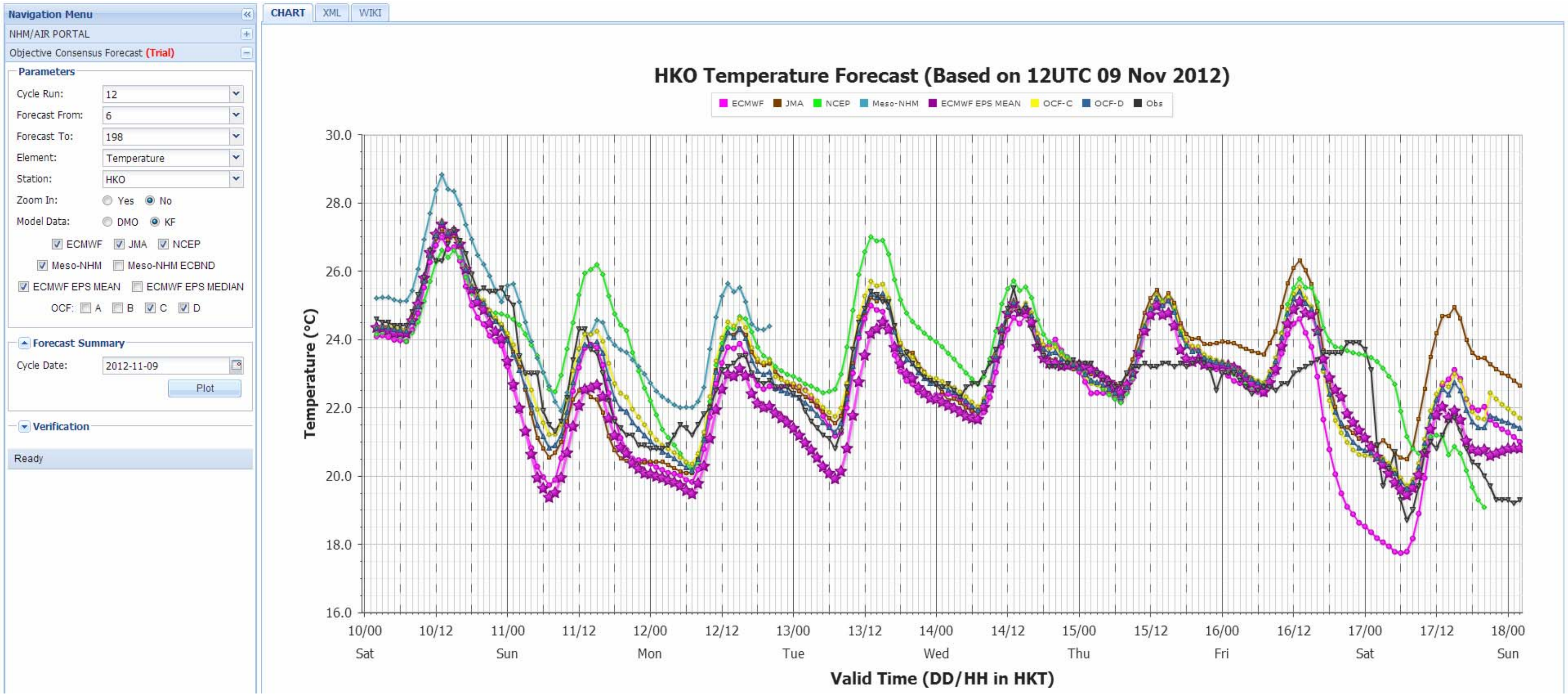


Fig. 2 User interface of OCF.

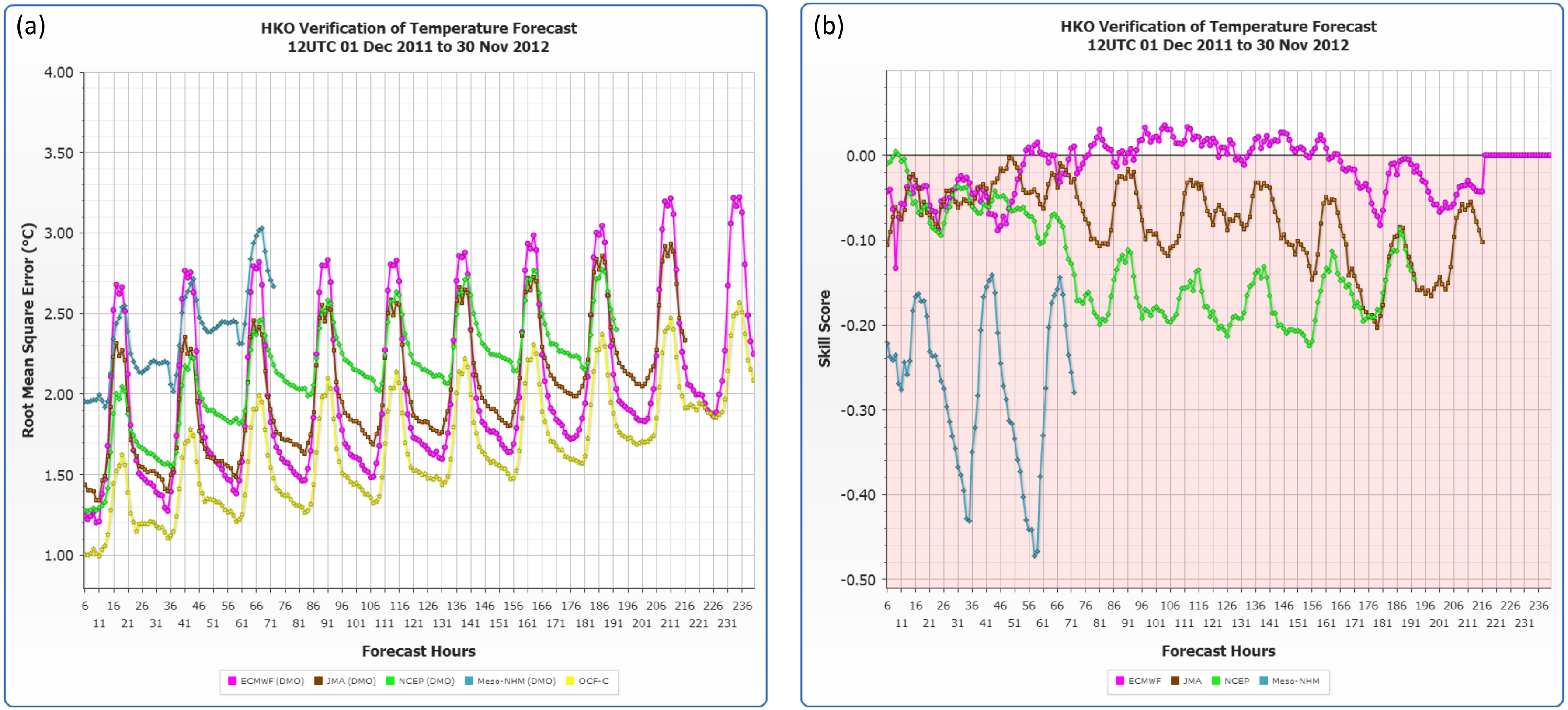
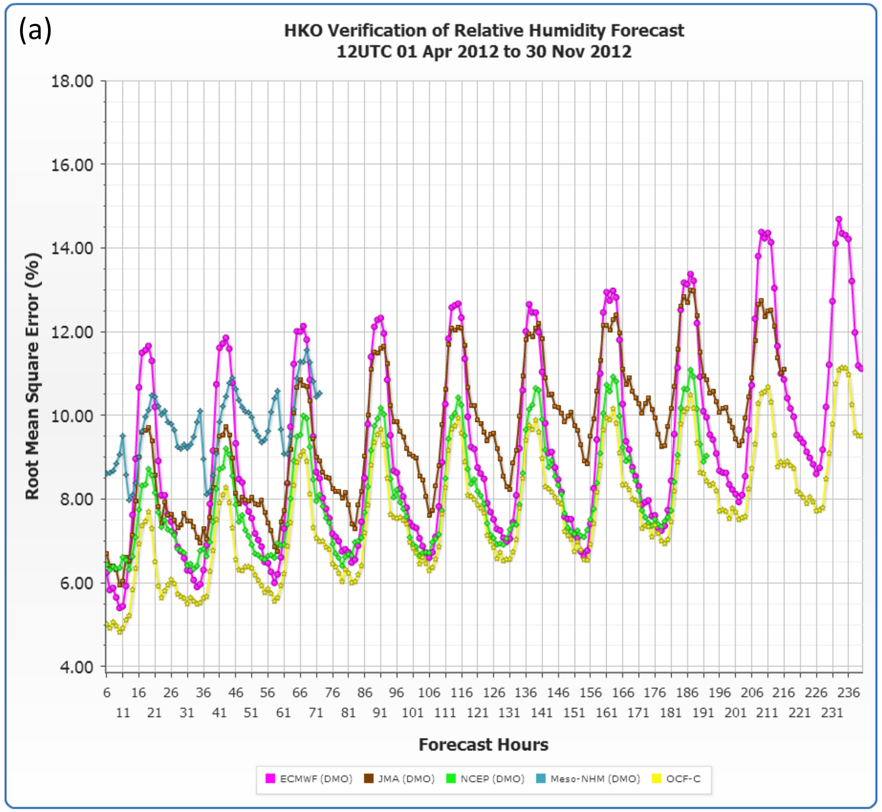


Fig. 3 Verification of temperature forecasts for HKO based on the 12 UTC model runs during 1 December 2011 – 30 November 2012. (a) RMSE versus forecast hours of OCF-C and DMO of the four member models, namely Meso-NHM, ECMWF deterministic model, JMA GSM and NCEP GFS; (b) skill score versus forecast hours of Kalman filtered forecasts of the four member models relative to OCF-C.



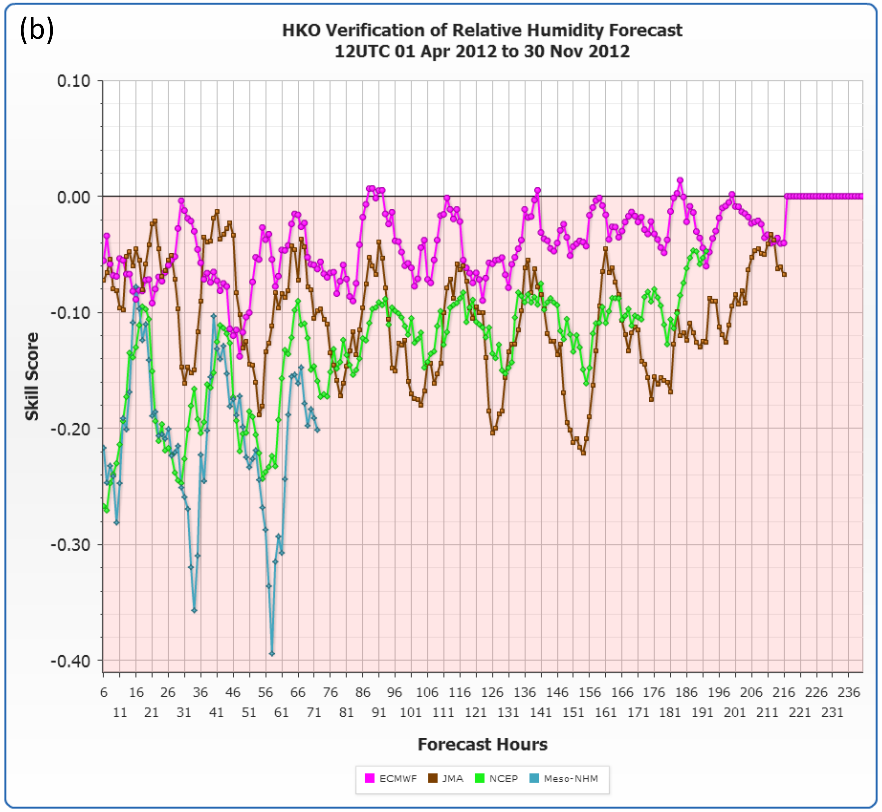


Fig. 4 Same as Fig. 3, except for verification of RH forecasts during 1 March – 31 November 2012.

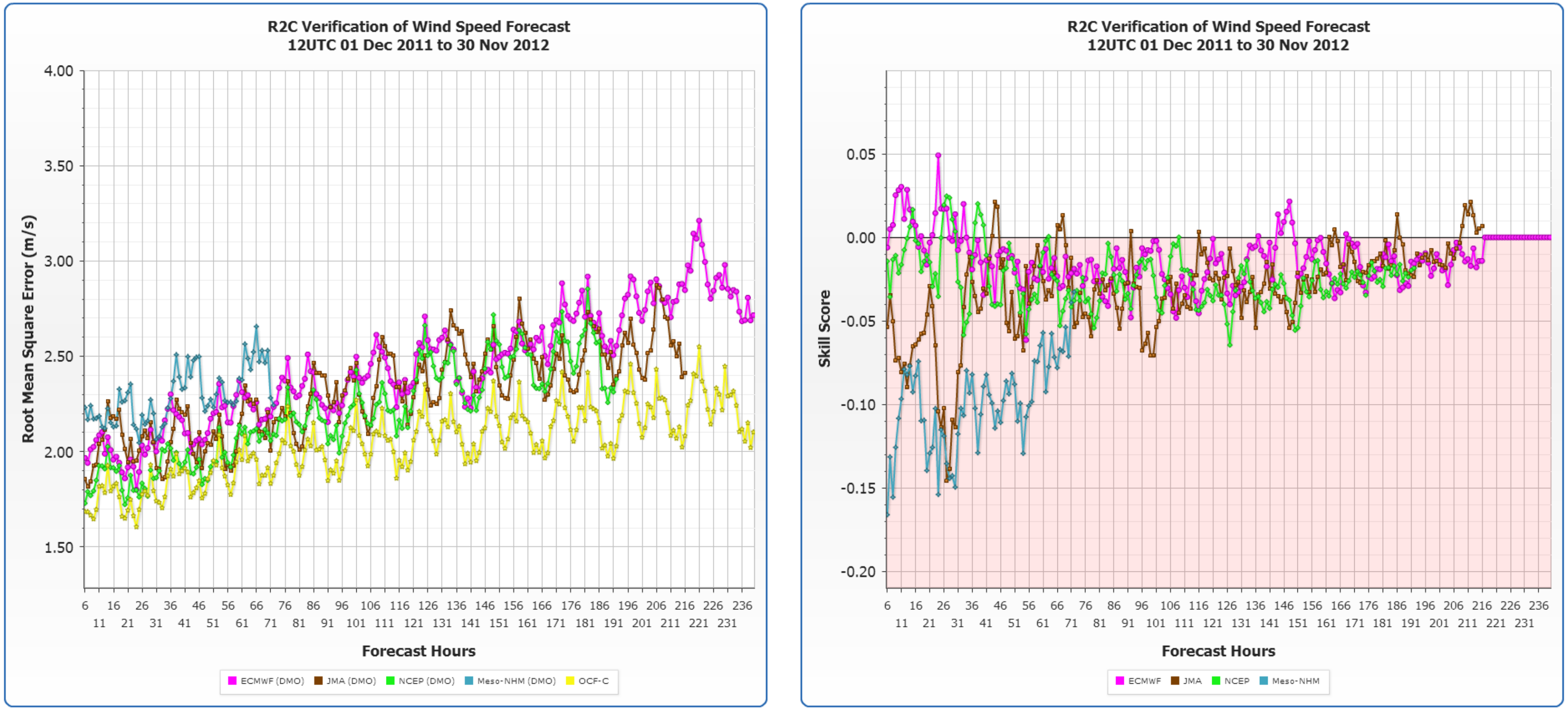


Fig. 5 Same as Fig. 3, except for verification of wind speed forecasts for R2C.

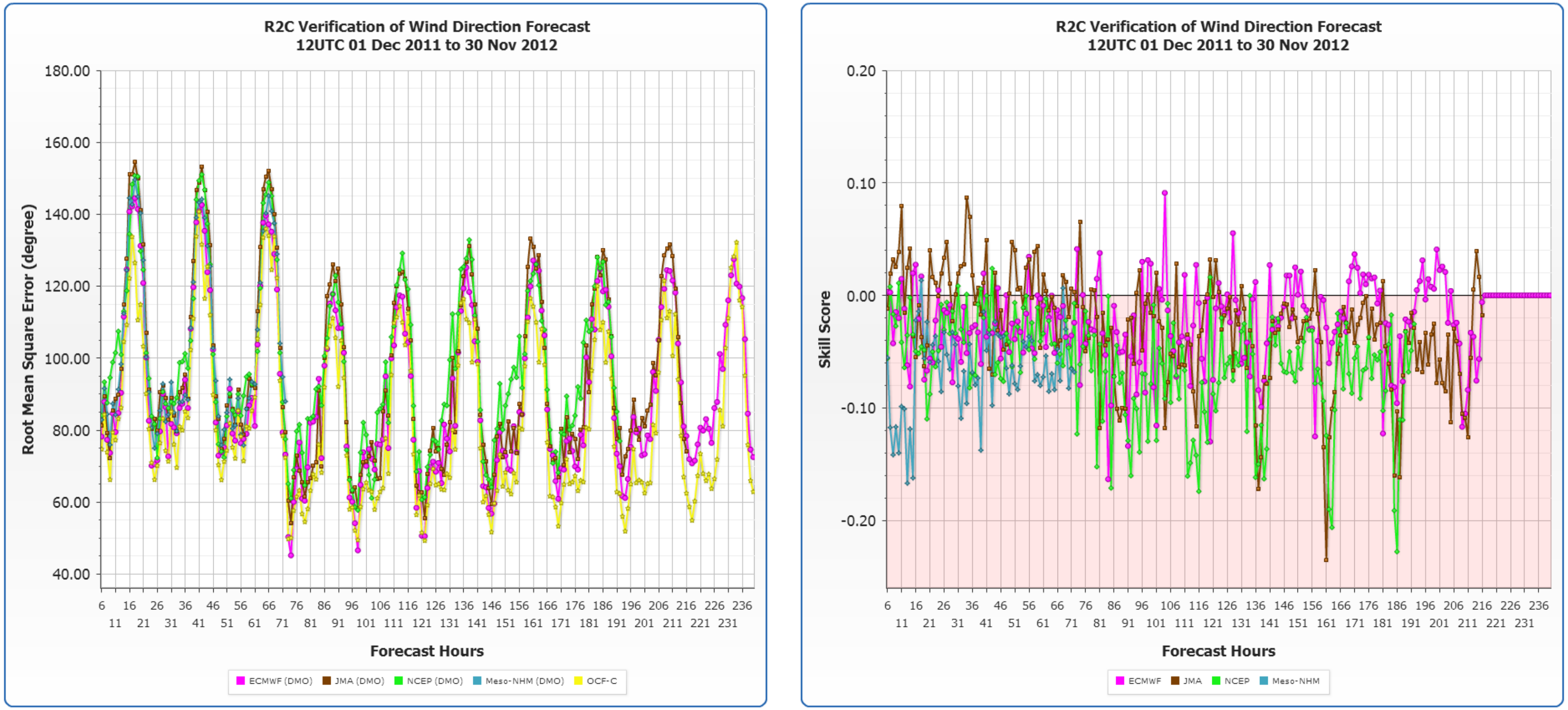


Fig. 6 Same as Fig. 3, except for verification of wind speed forecasts for R2C.

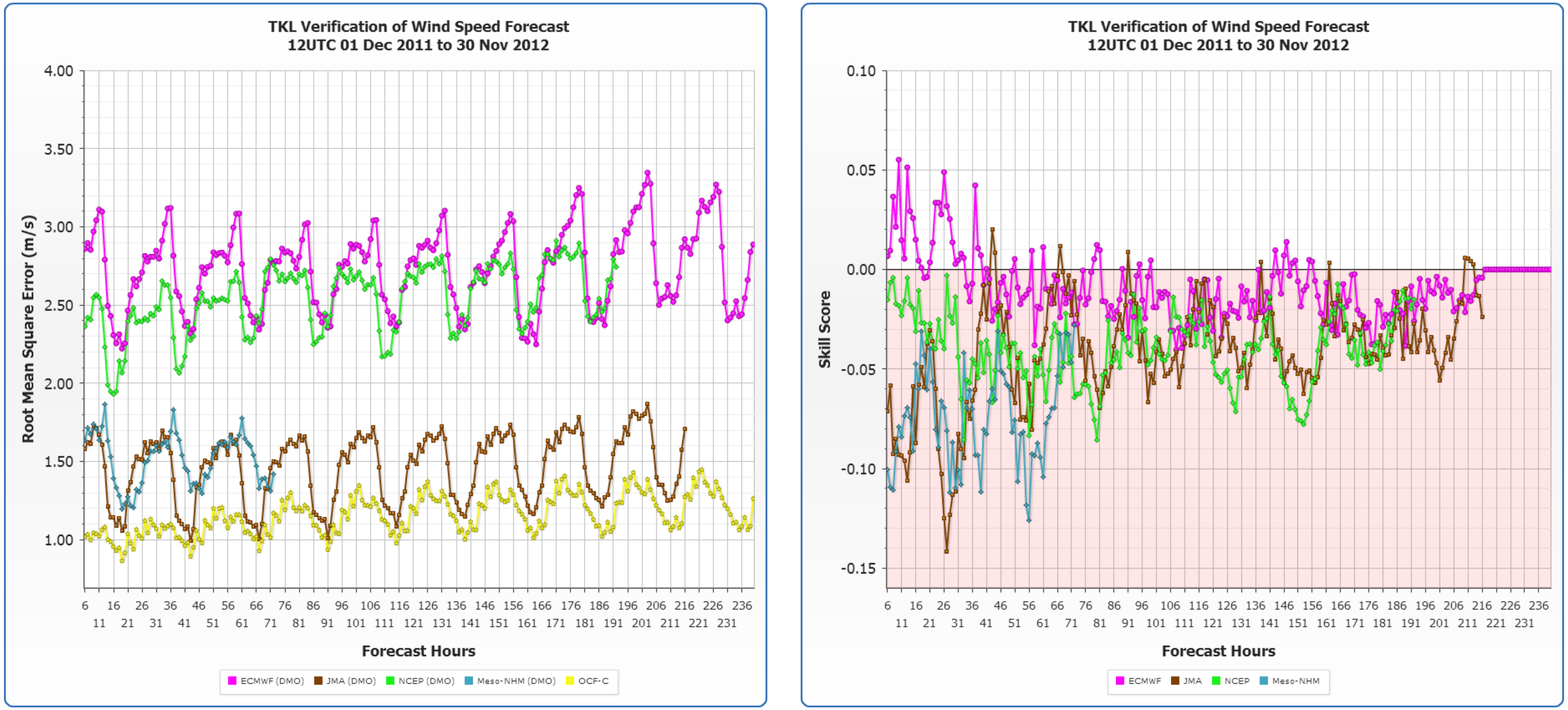
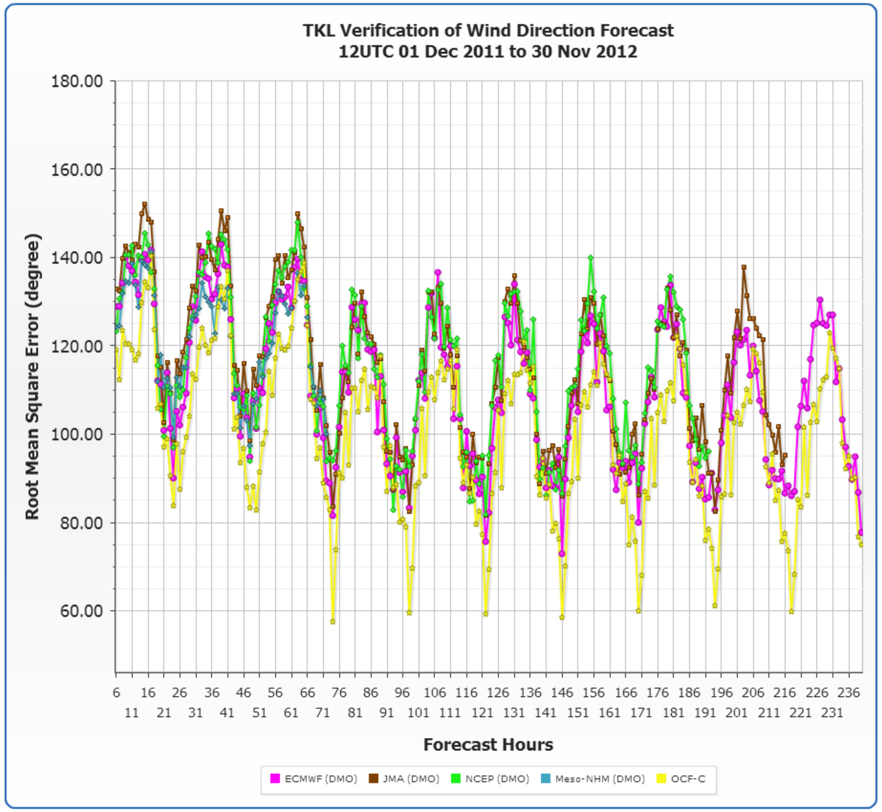


Fig. 7 Same as Fig. 3, except for verification of wind direction forecasts for TKL.



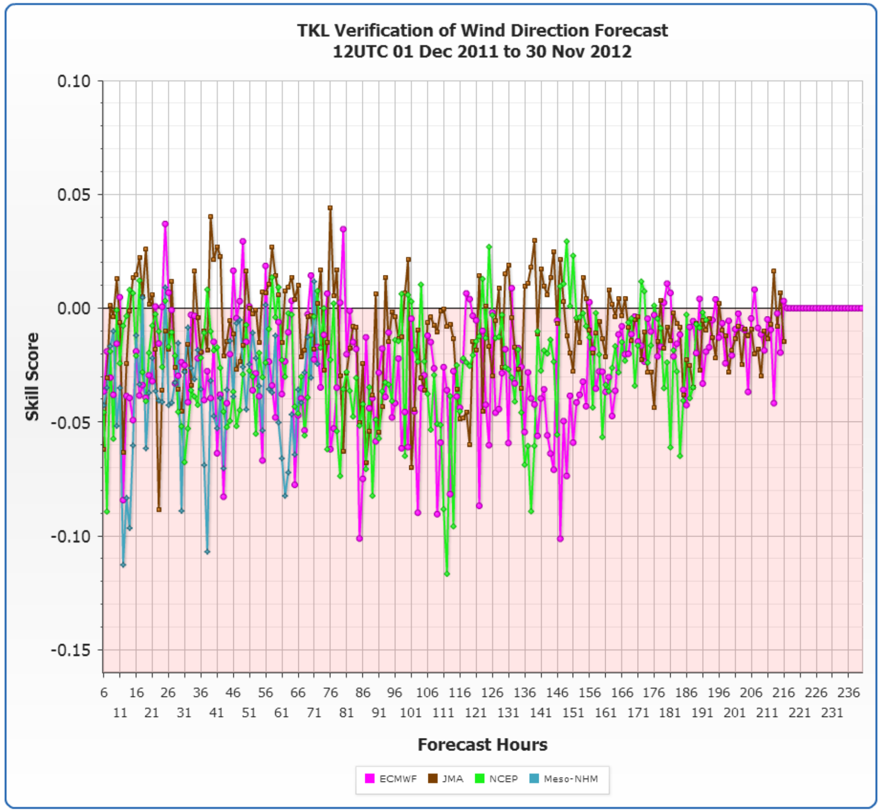


Fig. 8 Same as Fig. 3, except for verification of wind speed forecasts for TKL.

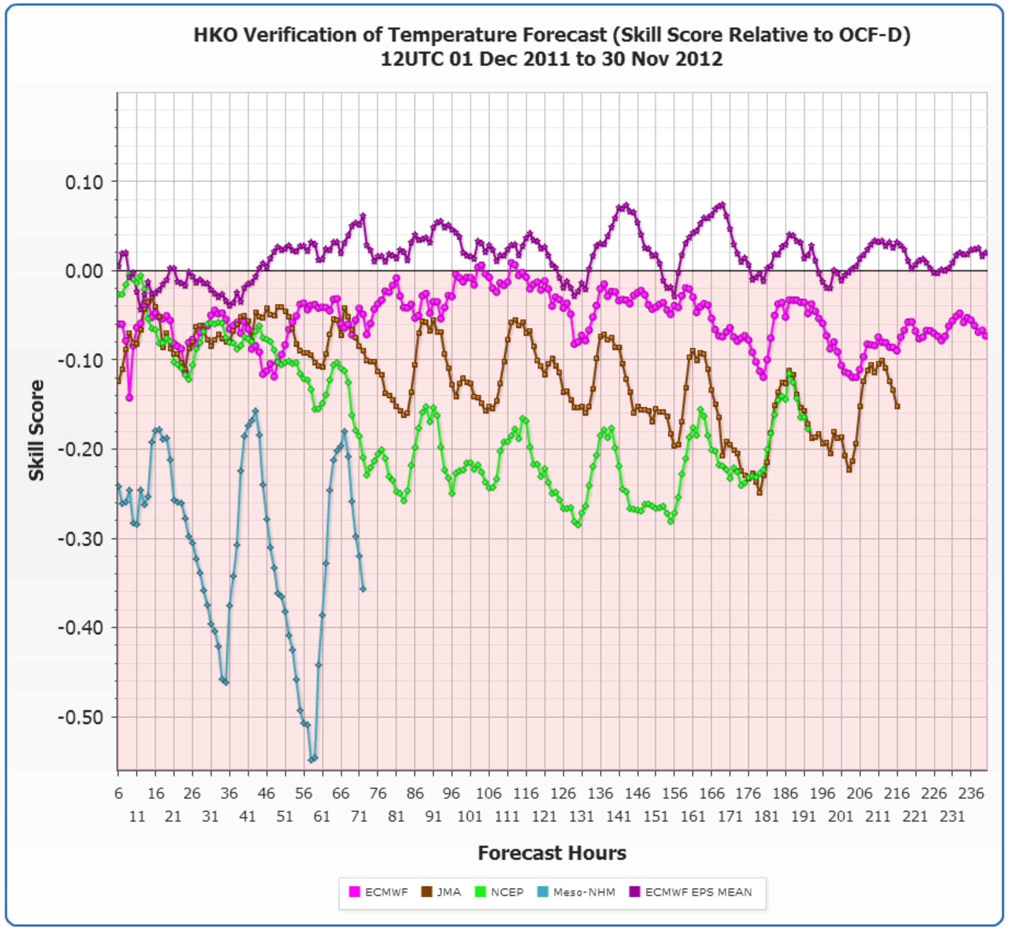


Fig. 9 Skill score relative to OCF-D of temperature forecasts for HKO based on the 12 UTC model runs during 1 December 2011 – 30 November 2012, with the inclusion of ECMWF EPS ensemble mean forecasts.