

EVALUATION OF NUMERICAL WEATHER PREDICTION SOLAR IRRADIANCE FORECASTS IN THE US

Richard Perez
ASRC, Albany, NY,
Perez@asrc.albany.edu

Mark Beauharnois
ASRC, Albany, NY
mark@asrc..albany.edu

Karl Hemker, Jr.
ASRC, Albany, NY
kmh1@asrc..albany.edu

Sergey Kivalov
ASRC, Albany, NY,
skivalov@asrc.albany.edu

Elke Lorenz
Universitat Oldenburg, Germany
elke.lorenz@uni-oldenburg.de

Sophie Pelland
CanmetENERGY, Natural Resources
Canada, QC, Canada
spelland@nrcan.gc.ca

Jim Schlemmer
ASRC, Albany, NY
Jim@asrc.cestm.albany.edu

Glenn Van Knowe
AWSTruepower, Albany, NY
gvanknowe@meso.com

ABSTRACT

We evaluate the performance of seven global irradiance (GHI) forecast models. The models are based directly or indirectly on numerical weather prediction (NWP). They are validated against one year of high quality measurements from the 7-station SURFRAD network representing distinct climatic environments in the US, and benchmarked against measured persistence.

1. INTRODUCTION

Two common operational approaches to solar radiation forecasting include (1) numerical weather prediction (NWP) models that infer local cloud formation probability -- hence, indirectly, transmitted radiation -- through the dynamic modeling of the atmosphere; and (2) models using the immediate recorded history -- from satellite remote sensing or ground measurements -- to infer the motion of clouds and project their impact in the future. Earlier contributions by some of the authors have shown that satellite-derived cloud motion tend to be best up to 3-5 hours ahead depending on location (e.g., [1, 2]).

In this paper we focus our attention on the NWP models which are the most appropriate for day-ahead and multi-day solar radiation forecasts.

2. FORECAST MODELS

The models considered for this evaluation include:

1. The Global Environmental Multiscale (GEM) model from Environment Canada in its regional deterministic configuration [3]
2. An application of the European Center for Medium Range Weather Forecasts (ECMWF) model [4]
3. A research version of the Weather Research Forecast (WRF) model [5] used as part of an operational air quality forecasting program [6]
4. An other version of the WRF model [5]
5. A proprietary mesoscale model, the MASS model [7]
6. The Advanced multiscale Regional Prediction System (ARPS) model [8]
7. A model based on cloud cover predictions from the US National Forecast Data Base (NDFD) [1]

The first two models are global (planetary) Numerical Weather Prediction (NWP) systems, respectively ECMWF and GEM

The native time step of the global models and their ground resolution are respectively 7.5 minutes and ~15 km for the GEM model and 3 hours and 25 km for the ECMWF model. The GEM solar forecasts were de-archived at an hourly time step and post-processed by taking an average of the

irradiance forecasts over a square region centered on the location of each station (ref Pelland et al.). The size of this square grid was optimized for each station by selecting a size that minimized forecast root mean square error during a one year training period prior to the evaluation period used here.

The ECMWF site-specific hourly data prepared for the present analysis are obtained via time interpolation of the 3-hourly global clear sky indices. No additional training to ground measured data was applied.

Models 3-6 are mesoscale models that use global weather models as an input to define regional boundary conditions, but add high resolution terrain and other features to produce higher resolution forecasts. In all cases analyzed here, the global weather model input is NOAA's GFS model [9]. The GFS model dataset used for this project has a time resolution of three hours and a nominal ground resolution of one by one degree (i.e., ~ 80x100km in the considered latitude range). All the mesoscale models produce hourly output with a final ground resolution of 5 km.

The three mesoscale models run by a commercial company considered here (WRF, MASS and ARPS) are tested in two operational modes: with, and without Model Output Statistics (MOS) post processing. The MOS process consists in integrating ongoing local [irradiance] measurements, when available, to correct localized errors from the numerical weather prediction process. This is a common operational forecasting practice: taking advantage of ongoing local surface and upper air measurements to deliver better forecasts.

The NDFD does not produce irradiances per se, but cloud amount forecasts that extend up to seven days ahead with a ground resolution of ~ 5 km and a time resolution of 3 hours up to three days ahead and six hours beyond that. The NDFD is also based on the GFS global model. GFS forecasts are individually processed by regional NOAA offices using mesoscale models and local observations and gridded nationally into the NDFD. The forecast cloud amounts are modeled into irradiance using an approach developed by Perez et al. [1]

All forecasts are set to nominally originate at 0000 UTC. In addition, some of the models are also tested with an origination time of 1200UTC.

3. VALIDATION

3.1 Validation measurements

Validation measurements consist of hourly integrated global horizontal irradiance GHI recorded for a one year period –

May 1st, 2009, through April 30th, 2010 – at each SURFRAD network location (Table 1).

Some of the models were only processed for a subset of the SURFRAD sites. The commercial mesoscale models could only be run at Desert Rock, Goodwin Creek and Penn State, while the research WRF model was only available for Goodwin Street and Penn State.

All models were processed to deliver data up to 48 hours ahead (next day and two day forecasts). The ECMWF forecasts were processed up to three days, and the NDFD up to seven days.

TABLE 1

Station	Latitude	Longitude	Elevation	Climate
Goodwin Creek	34.25	89.87	98 m	subtropical
Desert Rock	36.63	116.02	1007 m	Arid
Bondville	40.05	88.37	213 m	Continental
Boulder	40.13	105.24	1689 m	Semi-arid
Penn State	40.72	77.93	376 m	humid continental
Sioux Falls	43.73	96.62	473 m	Continental
Fort Peck	48.31	105.10	634 m	Continental

3. 2 Model Benchmarks

Results are benchmarked against measured persistence. Measured persistence consists of projecting current measured conditions in the future while accounting for solar geometry changes. Since we are intercomparing NWP models originating nominally at 0:00Z, hence designed for next/subsequent day forecast, the benchmark persistence is obtained by determining the daily global clear sky index K_t^* from the last available day and projecting this index for all subsequent forecast days/hours.

The satellite irradiance model developed by Perez et al., and used in the NSRDB [10] and SolarAnywhere [11] is used as a background reference to gauge the performance of the forecast models.

3.3 Validation Metrics

The validation benchmarks include root mean square, mean absolute and mean bias errors (RMSE, MAE, and MBE) as well as the Kolmogorov-Smirnov Integral (KSI) metric that quantifies the ability of a model to reproduce observed statistical distributions (e.g., see [1]).

The RMSE is often considered as the most important model validation metric because it quantifies a model's ability to accurately predict changing conditions. However because it is based on the square of model-measure differences, it can sometimes be overly influenced by a few distant outliers.

The MAE – the mean of the differences' absolute values -- is a useful complement to the RMSE that is effective at quantifying the tightness of the measure-model scatter plot near the 1-to-1 line.

Because several issues have been raised in the definition of the KSI statistics benchmark as defined in MESOR and IEA task 36 [re], notably the strong dependence of the critical value on the number of data samples, we decided to use a more robust interpretation of the KSI metric that simply describes the integrated absolute difference between the modeled and measured normalized cumulative distributions.

4. RESULTS

Tables 1 to 4 report respectively the relative RMSE, MAE, MBE and KSI for all sites, all models and forecast time horizons.

The results from Tables 1-4 are summarized in Figs 1-4, representing respectively all-site composite RMSE, MAE, MBE, and KSI. In addition, Figure 5 illustrates the all-site RMSE of daily mean GHI.

The all-site composite error was approximated by simple averaging of individual sites' errors. For the models that were only available for a subset of the sites, an estimate of composite value was prorated by first taking the ratio of the considered model's metric and the mean of all the other models for the said subset of sites and multiplying by the ratio of the other models' all site composite to their mean for the subset of sites.

5. DISCUSSION

Interestingly, all the GFS-based models, including NDFD, have similar accuracies as quantified by the RMSE and MAE metric. Models based on the European or Canadian global weather simulations tend do deliver better RMSE results. This is in line with the results of similar benchmarking exercises in Europe and Canada where the best forecasts were obtained from (post-processing of) global numerical weather prediction models (ref Lorenz et al, 2009).

As shown by Pelland [re], the combination of ECMWF and GEM (obtained by simple averaging of their forecasts) tends to do slightly better than individual models.

The MOS versions of the mesoscale models do of course very well in terms of mean bias (MBE) and statistical distribution (KSI) since they are, in effect, calibrated in real time with ground measurements. The initial (non MOS)

radiation schemes accuracy is not as good, pointing to shortcomings with the selected mesoscale model's radiation algorithms. This confirms an earlier US evaluation of the WRF model by some of the authors [14].

The strength of the MOS process is apparent in the daily RMSE (shown in Figure 5) where the mesoscale models tends to perform relatively better with respect to the other non-MOS models, although not to the level of ECMWF or GEMS.

In subsequent work it would be interesting to consider versions of the ECMWF and GEM forecasts that incorporate bias removal methods such as those that have been developed by Lorenz [15] and Pelland [16]. These approaches would reduce forecast RMSE further and avoid a systematic tendency to over or under forecast. They were not considered here since they require ongoing training using past data within the evaluation period.

6. ACKNOWLEDGEMENT

This work was supported by NREL under contract no. AGJ04029001 as a contribution to IEA SHC Task 36

7. REFERENCE

1. Perez, R., S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renne and T. Hoff, (2010): Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy* (in press).
2. Heinemann, D., Lorenz, E., Girodo, M.: 2006, 'Solar irradiance forecasting for the management of solar energy systems', *Solar 2006*, Denver, CO (USA), 07.07.2006
3. Mailhot J, Bélair S, Lefaivre L, Bilodeau B, Desgagné M, Girard C, Glazer A, Leduc AM, Méthot A, Patoine A, Plante A, Rahill A, Robinson T, Talbot D, Tremblay A, Vaillancourt PA and Zadra A, 2006. The 15-km version of the Canadian regional forecast system. *Atmosphere-Ocean*, 44, 133-149
4. Lorenz, E; Hurka, J; Heinemann, D, Beyer H.G.: 'Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems', *IEEE Journal of Special Topics in Earth Observations and Remote Sensing*, 2, 2–10 (2009).
5. Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A Description of the Advanced Research WRF Version 2, NCAR TECHNICAL NOTE: NCAR/TN-468+STR, National Center for Atmospheric Research, Boulder, Colorado, USA

6. Air Quality Forecast Modeling System (<http://asrc.albany.edu/research/aqf/>)
7. MASS model reference
8. Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Physics.*, 76, 134-165.
9. GFS reference
10. 11. NSRDB, (1998-2005): http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/
11. 9. Solar Anywhere, (2010): www.solaranywhere.com
12. International Energy Agency, Solar Heating and Cooling Programme, Task 36 subtask A, model benchmarking (2011)
13. Pelland, S., (2011): Personal communication
14. J. Remund, R. Perez and E. Lorenz , (2008): Comparison of Solar Radiation Forecasts for the USA. Proc. 23rd European Photovoltaic Solar Energy Conference, Valencia, Spain
15. Pelland, S., Gallanis, G., Kallos, G 2011: Solar and Photovoltaic Forecasting through Post-Processing of the Global Environmental Multiscale Numerical Weather Prediction Model. (In preparation)
16. Lorenz, E., Remund, J., Müller, S.C., Traunmüller, W., Steinmauer, G., Pozo, D., Ruiz-Arias, J.A., Lara Fanego, V., Ramirez, L., Romeo, M. G., Kurz, C., Pomares, L. M., Guerrero, C.G., 2009: Benchmarking of Different Approaches to Forecast Solar Irradiance, Proceedings of the 24th European Photovoltaic Solar Energy Conference, 21-25 September 2009, Hamburg,

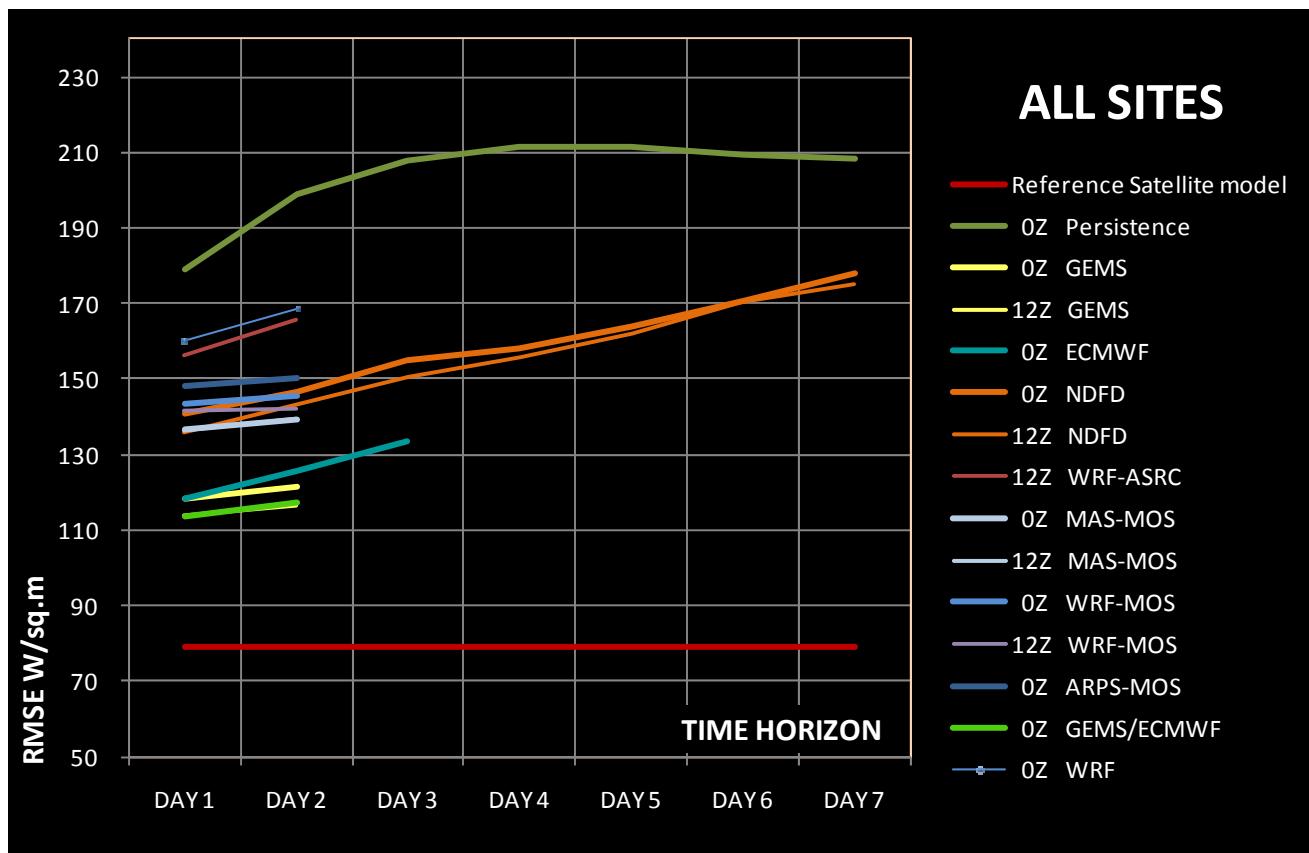


Figure 1 Composite RMSE as a function of prediction time horizon

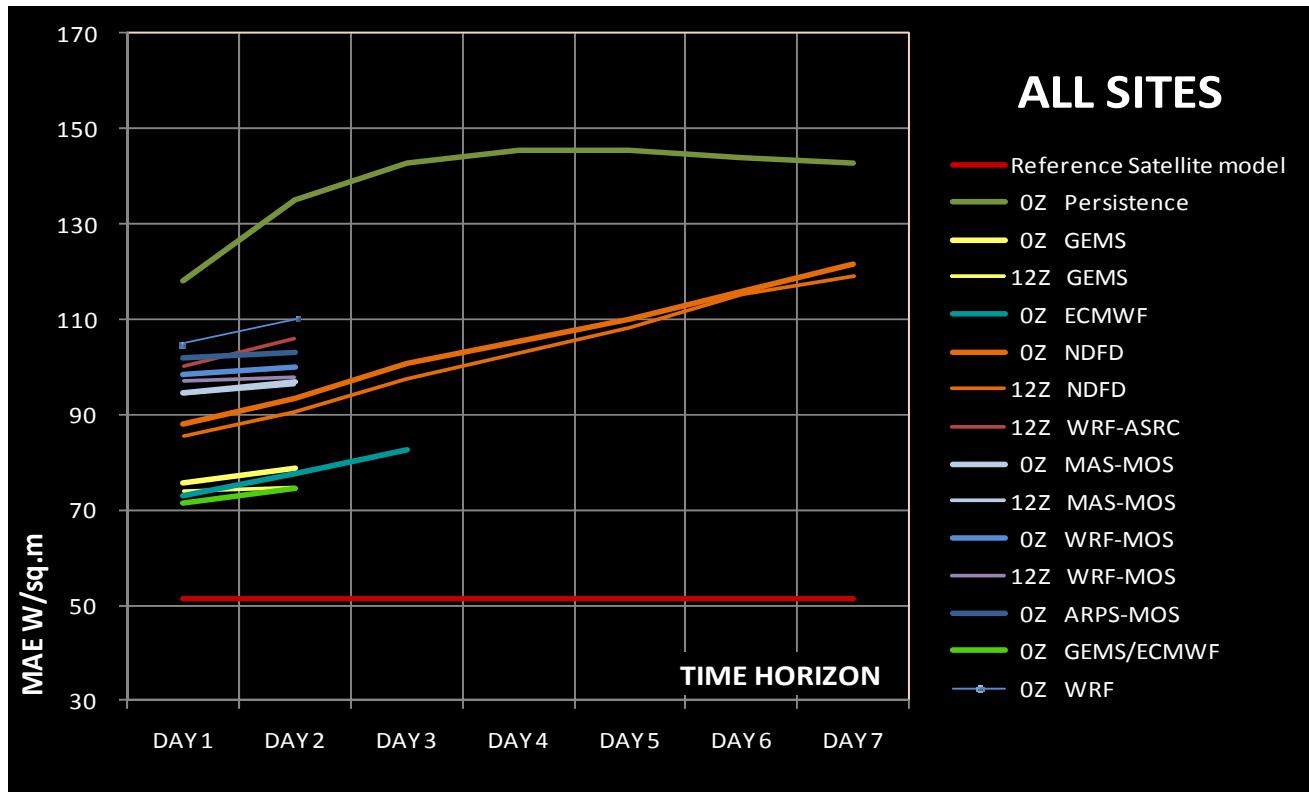


Figure 2 Composite MAE as a function of prediction time horizon

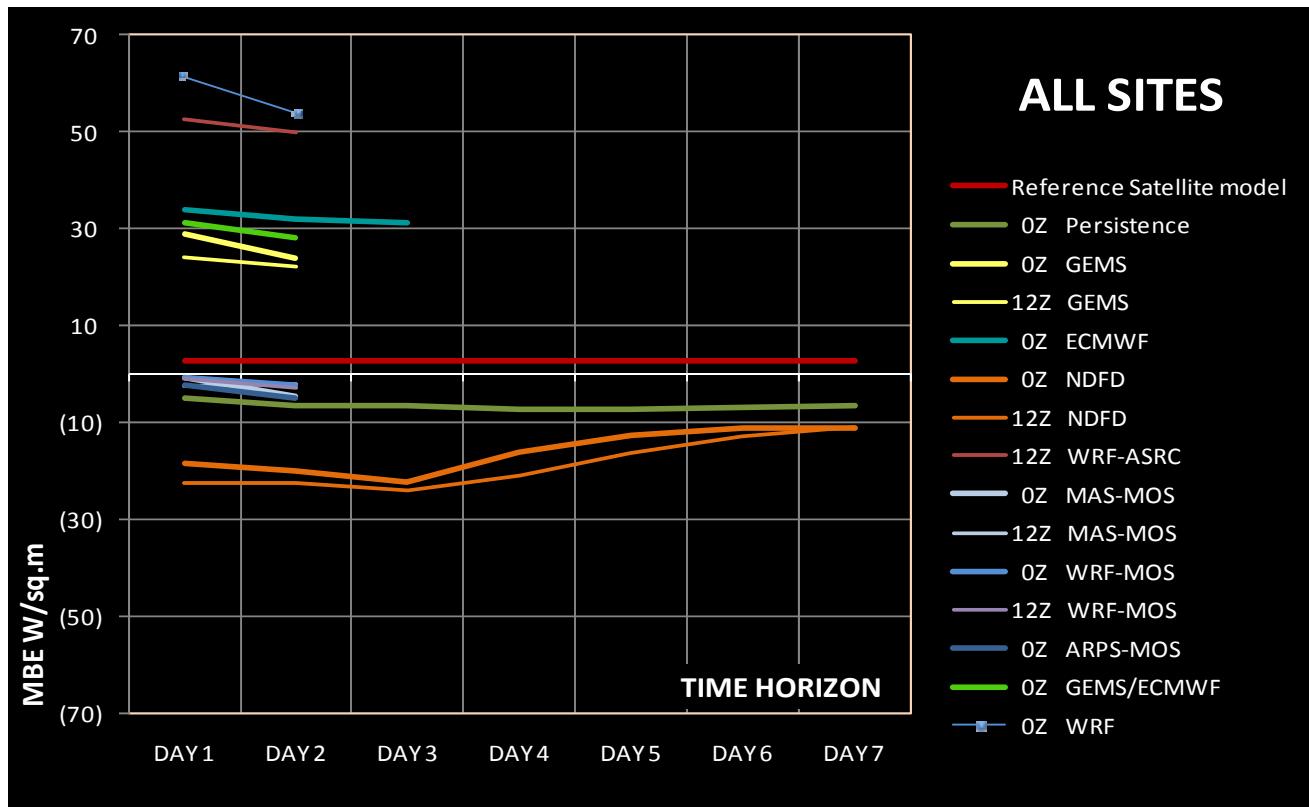


Figure 3 Composite MBE as a function of prediction time horizon

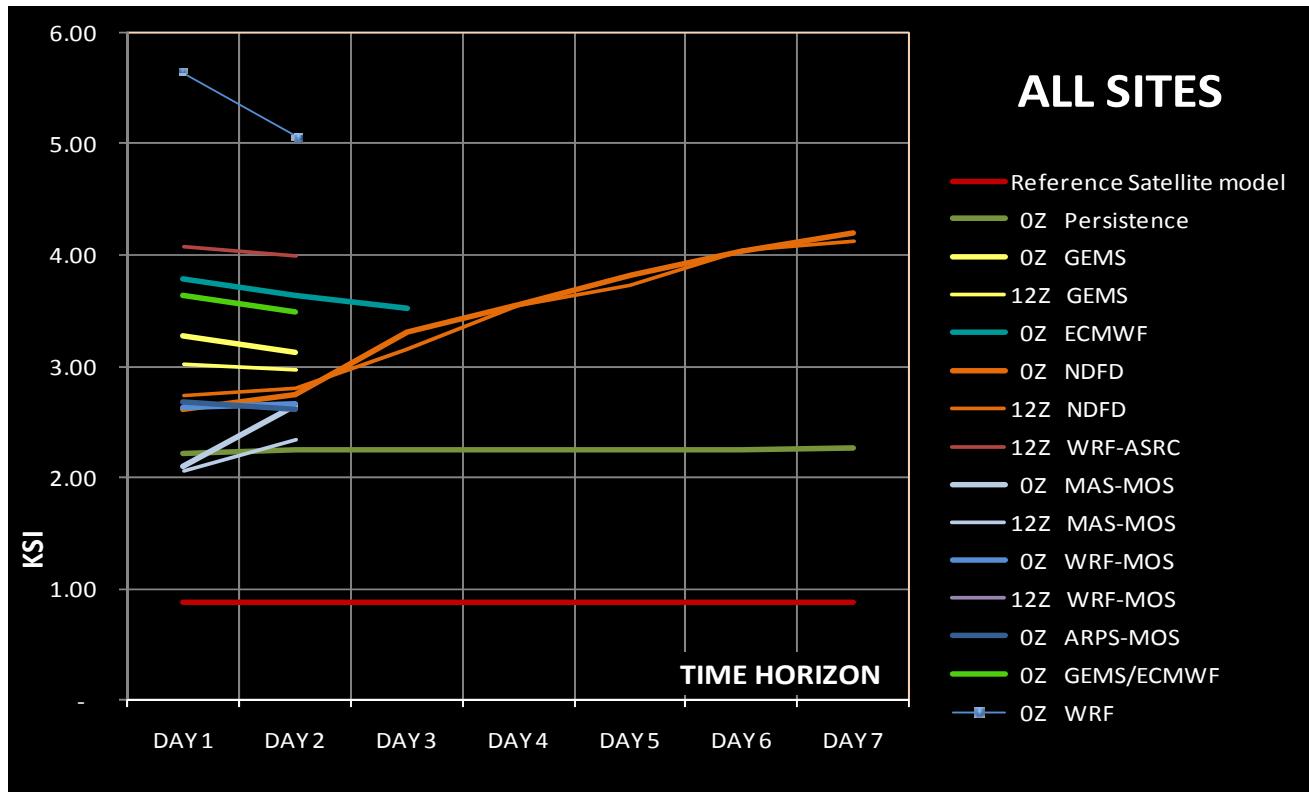


Figure 4 Composite KSI*100 as a function of prediction time horizon

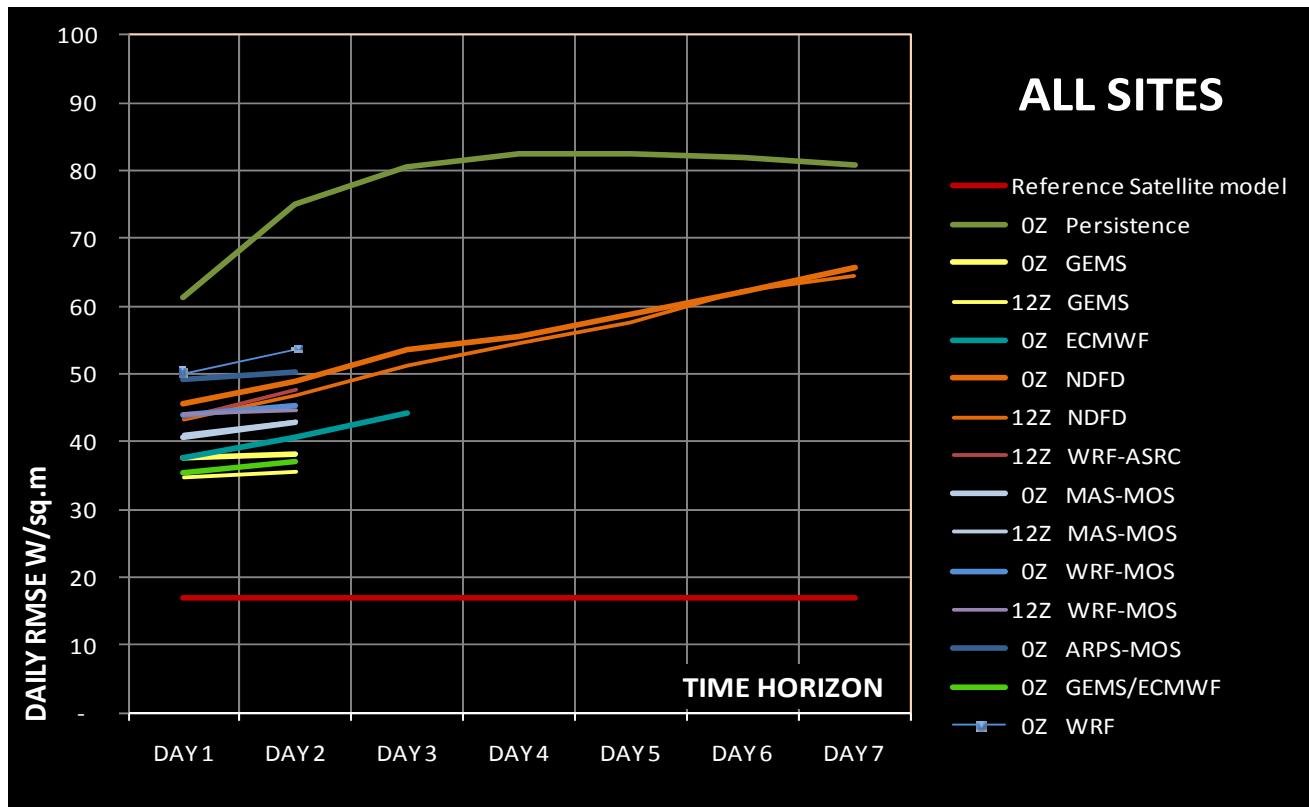


Figure 5 Composite daily RMSE as a function of forecast time horizon

