Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
43rd NOAA Annual Climate Diagnostics and Prediction Workshop
Santa Barbara, CA, 23-25 October 2018

Improve CFS Week 3-4 Precipitation and 2 Meter Air Temperature Forecasts with Neural Network Techniques

Yun Fan[1], Chung-Yu Wu[1], Jon Gottschalck[1], and Vladimir Krasnopolsky[2]
[1]*Climate Prediction Center, NOAA/NWS/NCEP*
[2]*Environmental Modeling Center, NOAA/NWS/NCEP*

1. Introduction

The public demand for sub-seasonal forecasts have been steadily increasing in recent years, primarily driven by many industries, such as water management, agriculture, transportation, commerce and insurance etc., to prepare for and reduce risk from damaging meteorological events well in advance. Numerical forecasts on the Week 3-4 time scale are relatively new and to be one of the most challenging and difficult forecast time scales. Past forecast efforts have been focused on the short term weather forecasts out to 7-10 days and operational short term climate outlooks from month to several seasons. There is a clear forecast gap between the Week 3 and 4.

In 2016, the National Oceanic and Atmospheric Administration (NOAA) initiated the efforts to improve its capability for the Weeks 3 and 4 extended range forecasts. Covering this extended-range Week 3~4 forecasts will enable NOAA to provide seamless S2S forecasts to the public for protecting life and property. So far, the Week 3 ~4 forecast skills from direct dynamical forecast models are much lower than the short range forecasts, such as 1~7 days and the Week 1~2 forecasts. In this study, the deep machine learning (*i.e.* Neural Network – NN) techniques are proposed to explore, test, evaluate, and eventually implement a reliable statistical post processing method utilizing model derived fields to improve the original NOAA CFS Week 3-4 time range model precipitation (P) and 2 meter air temperature (T2m) forecasts.

2. Methodology and data

Usually statistical post processing of model outputs is based on a reasonable assumption that there is a relationship between target variables (predictands) (e.g. observed weather and climate elements) and input variables (predictors) (e.g. the NWP model forecast variables). In a very generic symbolic way, this relationship can be written as:

$$Z = M(X); \quad X \in \Re^n, \; Z \in \Re^m \tag{1}$$

where X is a input vector composed of model forecast variables or predictors, Z is a output vector composed of observed meteorological elements or predictands, n is the dimensionality of the vector X (or input space), and m is the dimensionality of the vector Z (or output space). M denotes the mapping (relationship between the two vectors) that relates vectors X and Z.

Since both model forecast variables (predictors) and observations (predictands) contain errors in their data representations, a statistical approximation of the mapping Eq.(1) can be written as,

$$Y = M_s(X) \tag{2}$$

here vector Y can be considered as a vector of corrected model variables X and M_s isa statistical approximation for the mapping M in Eq (1).

The NN techniques are very flexible and convenient mathematical/statistical tools that can allow users to approximate different complicated nonlinear input/output relationships/mappings, by using statistical deep machine learning algorithms (Krasnopolsky 2013). The simplest NN approximations use a family of analytical functions like:

Correspondence to: Yun Fan, Climate Prediction Center, NOAA/NWS/NCEP, 5830 University Research Court, College Park, MD 20740; E-mail: yun.fan@noaa.gov

$$y_q = NN(X, \boldsymbol{a}, \boldsymbol{b}) = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot t_j; \qquad q = 1,2, \dots, m \tag{3}$$

where

$$t_j = F(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i) = \tanh(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i); \tag{4}$$

here x_i and y_q are components of the input and output vectors X and Y, respectively, $\boldsymbol{a}$ and $\boldsymbol{b}$ are NN weights, n and m are the numbers of inputs and outputs respectively, and k is the number of the nonlinear basis activation function t_j (or hidden neurons). Here the hyperbolic tangent is used as an activation function. Eq. (3) is a mapping, which can approximate any continuous or almost continuous (with final discontinuities) mapping (Krasnopolsky 2013). A pictographic representation of the entire NN was shown in Fig.1 and the connections (arrows) correspond to the NN weights.

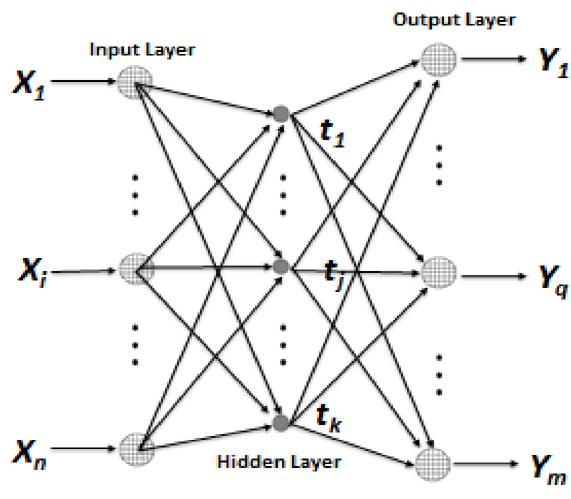

Fig.1 The simplest NN with one hidden layer and linear neurons in the output layer.

To find coefficients a_{qj} and b_{ji} in NN Eq. (3, 4), an error function, E, is created,

$$E = \frac{1}{N} \sum_{i=1}^{N} [Z_i - NN(X_i)]^2 \tag{5}$$

where vector Z_i is composed of observed weather elements, and N is the total number of paired records included in the training data set. Then, the error function (5) is minimized to obtain an optimal set of all coefficients a_{qj} and b_{ji} via a simplified version of the procedure known as the back propagation training algorithm. The back propagation algorithm searches for minimum of error (or cost) function in weight space through the steepest (gradient) descent method. It partitions the final total cost to each of the single neuron in the network and repeatedly adjusts the weights of neurons whose cost is high, and back propagate the error through the entire network from output to its inputs.

The data set used for predictors here is the bias corrected Week 3~4 forecast total precipitation (P) and mean 2 meter (T2m) temperature etc., and from the NOAA Climate Forecast System (CFS) (Saha *et al.* 2006, 2014) for period Jan. 01, 1999 to Dec. 31, 2018. The data domain used in this study covers the Conterminous US (CONUS) only, has 1x1 degree spatial resolution and on daily temporal resolution initialized at 4 different times (00Z, 06Z, 12Z and 18Z) per day.

The data set used for correspondent target variables (predictands) are the observed P from the gauge based daily CPC Unified Precipitation Analysis and the observed T2m from the Global Telecommunications System (GTS) based daily maximum and minimum 2 meter temperature analysis (Chen *et al.* 2008, Shi personal communication, Fan *et al.* 2008). Both the above observed P and T2m are converted to two weekly total and two weekly mean, and re-gridded to the same spatial-temporal resolutions as the above predictors.

3. Results

It is well known that the forecasts for the Week 3-4 time scale is one of toughest forecast areas and the skills are very low in general. In this study, an open question wanted to be addressed is if the deep machine learning techniques used here can add additional values to improve the targeted forecasts in the Week 3-4 time scale, when compared with the benchmark multiple linear regression (MLR) tools and also the bias corrected CFS the Week 3~4 forecasts as the inputs (or predictors).

Several tests have been conducted and the results indicate that using ensemble mean from 4 initial time (00Z, 06Z, 12Z and 18Z) the resultant Week 3~4 NN P and T2m forecasts in general are better than the results from using the CFS P and T2m forecasts on the individual initial time. In the following part of this paper, the main focus will be all on a more beneficial NN configuration that the entire NN training and testing at all locations will be done simultaneously in just one same training cycle.

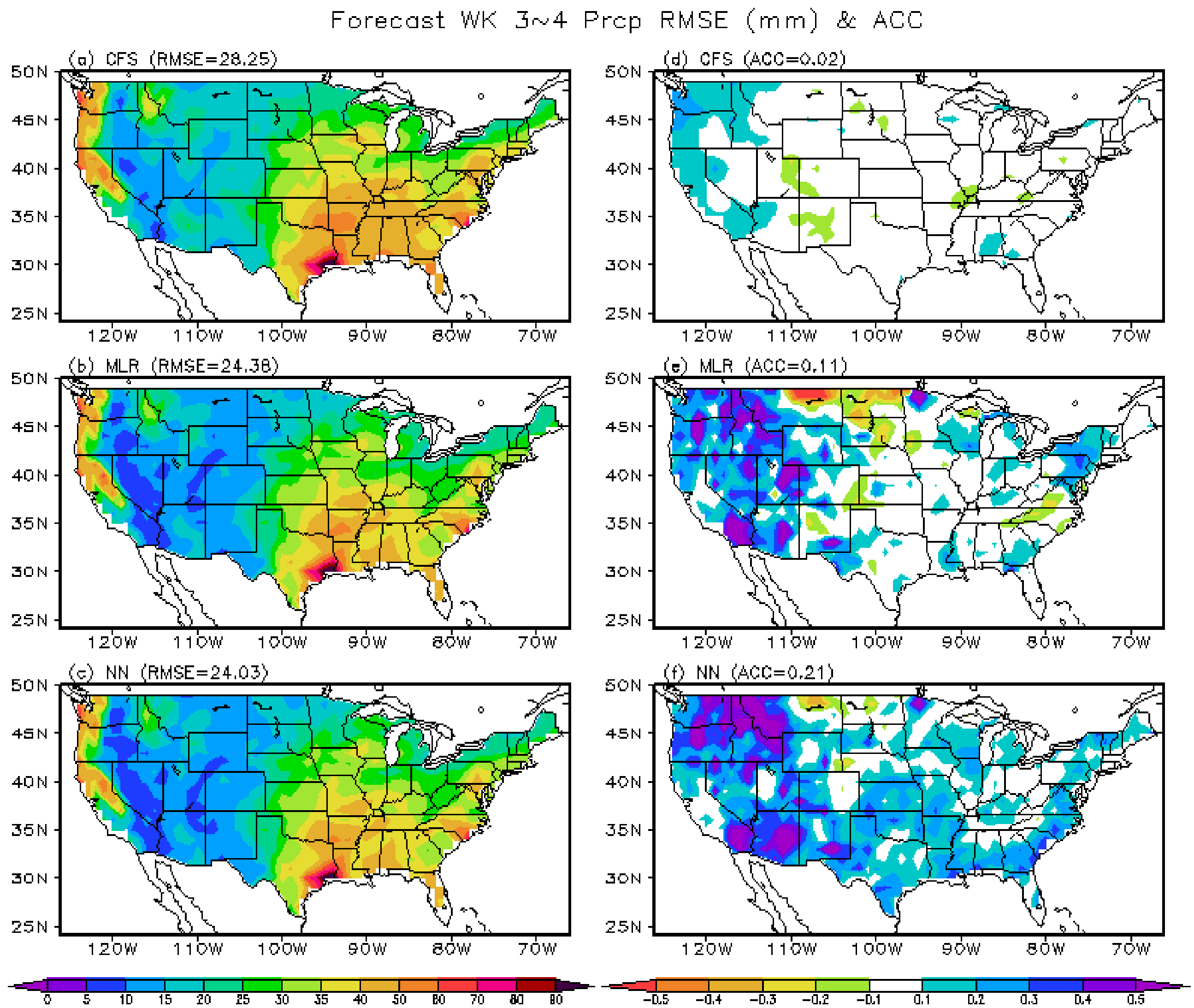

Fig.2 The RMSE and ACC of daily Week 3~4 total P by (a, d) Bias corrected CFS ensemble mean forecasts, (b, e) MLR forecasts, (c, f) NN forecasts, against correspondent observation for period of Jan.1 2017 to Oct. 31, 2018.

Figure 2 shows that overall the root mean square errors (RMSE) and anomaly correlation coefficients (ACC) of the bias corrected ensemble mean CFS P forecasts improved by the NN are better than that of the forecast results obtained from the benchmark multiple linear regression method for most locations. Here the NN training period is from 01/01/1991 to 12/31/2016 (6575 day records). The period of 01/01/2017 to 10/31/2018 (~670 day records) is used as independent verification period. The above results indicate both the NN and the MLR methods improved the bias corrected CFS Week 3~4 P forecasts, especially the forecast skills in quite parts of the western CONUS are encouraging (ACC over 0.4 or 0.5). The NN forecasts show clearly better forecast skills than the MLR forecasts over most locations in term of the RMSE and ACC. It may also indicate that the NN corrections which take into account of the non-linearity, pattern relationship and co-variability impacts are important for improving P forecasts.

Same as precipitation, Figure 3 shows that the root mean square errors (RMSE) and anomaly correlation coefficients (ACC) of the bias corrected ensemble mean CFS T2m forecasts by the NN are better than that of the forecast results obtained from both CFS and the benchmark MLR method for most locations. The above results indicate both the NN and the MLR methods improved the bias corrected CFS Week 3~4 T2m forecasts, especially the forecast skills in large parts of the southwestern CONUS and the eastern half of the CONUS are quite encouraging (ACC over 0.4 or 0.5). The NN forecasts show clearly better forecast skills than the MLR

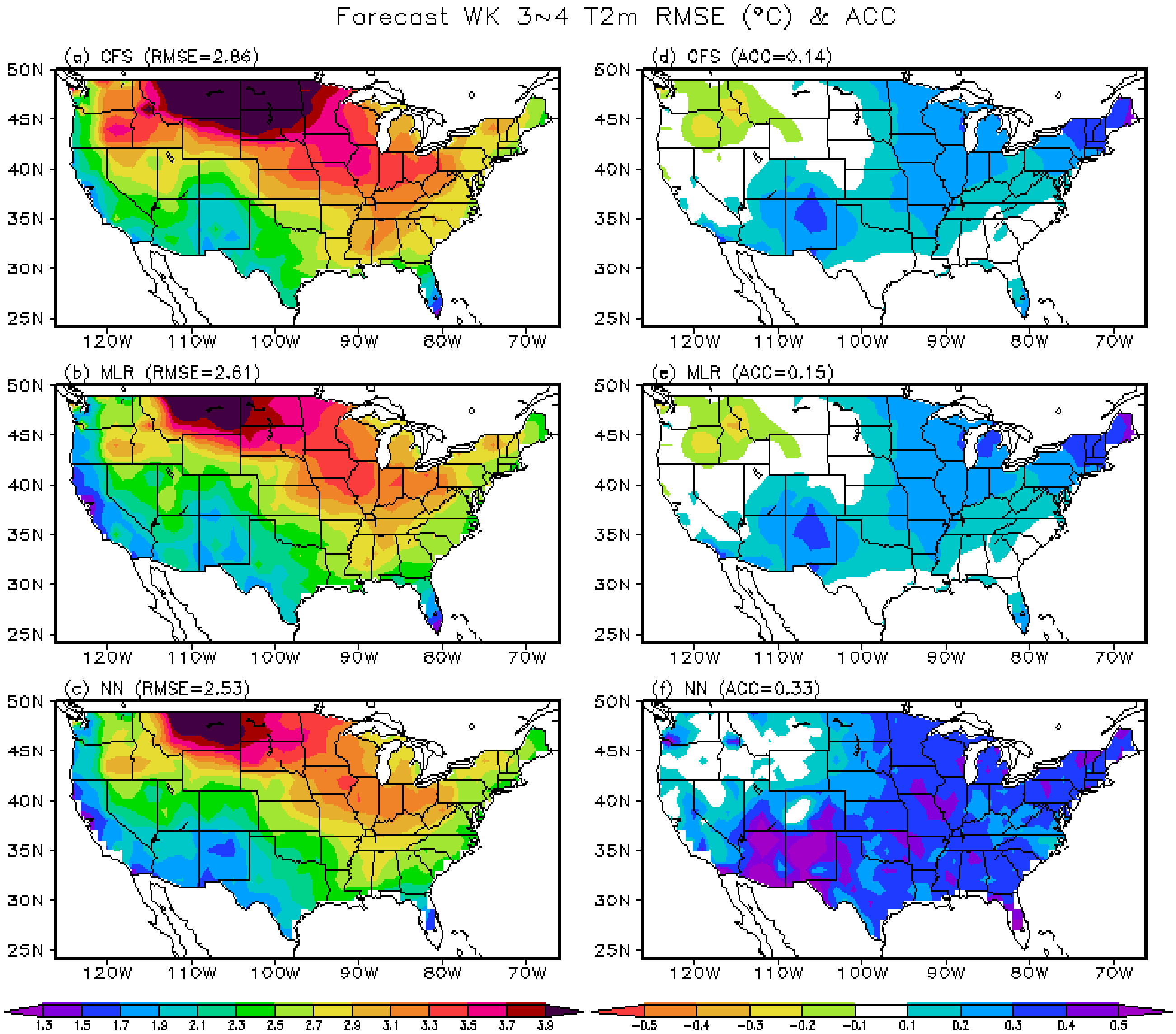

Fig. 3 The RMSE and ACC of daily Week 3~4 T2m by (a, d) Bias corrected CFS ensemble mean forecasts, (b,e) MLR forecasts, (c,f) NN-A forecasts, against daily observation for period of Jan.1 2017 to Oct. 31, 2018.

forecasts over most locations in term of the RMSE and ACC. It may also represent that the impacts of the non-linearity, pattern relationship and co-variability are also very important for the T2m correction.

Checking the overall forecast performance of three (CFS, MLR and NN) forecasts over the 2017-2018 two years verification period, both the MLR and the NN constantly beat the CFS. But the NN forecasts did much better job than the MLR forecasts in many aspects. Figure 4 depicts that the examples of the observed P and T2m anomalies, together with the correspondent Week 3-4 CFS, MLR and NN forecast anomalies. In both cases, the NN techniques show very impressive ability to reverse the wrong P and T2m forecast patterns.

4. Conclusion

In this study, the artificial neural network (deep machine learning) techniques are used to improve the NCEP CFS Week 3~4 P and T2m forecasts. Benefiting from the great advance in machine learning in recent years, the NN techniques show some advantages over traditional statistical methods (*e.g.* multiple linear regressions): such as flexible algorithm that can account for complicated linear and non-linear relationships, spatial dependency and co-variability *etc.*, at the same time is able to handle big data easily. Those learned statistical patterns and relationships from the NN training processes then are used by the NN to make the corrected forecasts.

Better data representation is very important for the NN training. The EOF analysis indicates that the CFS is very good at predicting large-scale patterns and low frequency variations in observed precipitation, rather than at capturing those highly parameterized and unresolved processes in precipitation. Better data representation may be archived by using ensemble means to increase the explained percentage of the total variance and to reduce noise in the data.

Although the improvement on the Week 3~4 precipitation and T2m is very encouraging, the overall forecast skill (in terms of RMSE and ACC) for the Week 3~4 precipitation and 2m air temperature predictions is still not great. Further studies are definitely needed.

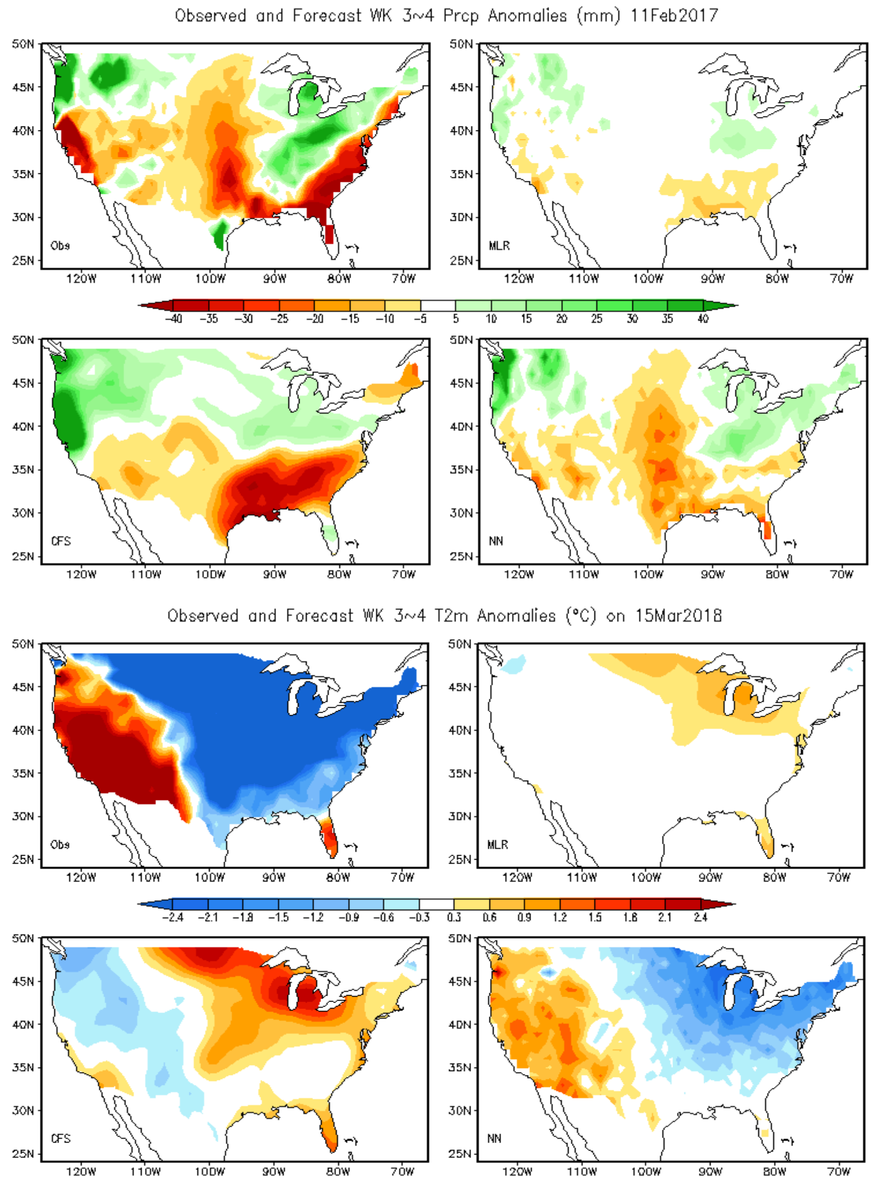

Fig. 4 The observed (Obs), CFS, MLR and NN forecast P (top 4 plots) and T2m (bottom 4 plots) week 3-4 forecast anomalies.

References

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V E. Kousky, R. Wayne Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation, *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.

Fan, Y. and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.*, **113**, D01103, doi:10.1029/2007JD008470.

Krasnopolsky, V., 2013: *The application of neural networks in the Earth system sciences. Neural network emulations for complex multidimensional mappings*. Springer, 200 pp.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483-3517, doi:10.1175/JCLI3812.1.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System Version 2. *J. Climate*, **27**, 2185–2208. doi:10.1175/JCLI-D-12-00823.1.