

Comparison of individual and combined ANN models for prediction of air and dew point temperature

Karthik Nadig · Walter Potter · Gerrit Hoogenboom · Ronald McClendon

Published online: 3 February 2013
© Springer Science+Business Media New York 2013

Abstract Predicted air and dew point temperatures can be valuable in decision making in many areas including protecting crops from damage, avoiding heat stress on animals and humans, and in planning related to energy management. Current web-based artificial neural network (ANN) models on the Automated Environment Monitoring Network (AEMN) in Georgia predict hourly air and dew point temperature for twelve prediction horizons, using 24 models. The observed air temperature may approach the observed dew point temperature, but never goes below it. Current web based ANN models have prediction errors which, when the air and dew point temperatures are close, may cause air temperature to be predicted below the dew point temperature. Herein this error is referred to as a prediction anomaly. The goal of this research was to improve the prediction accuracy of existing air and dew point temperature ANN models by combining the two weather variables into a single ANN model for each prediction horizon. The objectives of this study were to reduce the mean absolute error (MAE) of prediction and to reduce the number of prediction anomalies.

The combined models produced a reduction in the air temperature MAE for ten of twelve prediction horizons with an average reduction in MAE of 1.93 %. The combined models produced a reduction in the dew point temperature MAE for only six of twelve prediction horizons with essentially no average decrease in MAE. However, the combined models showed a marked reduction in prediction anomalies for all twelve prediction horizons with an average reduction of 34.1 %. The reduction in prediction anomalies ranged from 4.6 % at the one-hour horizon to 60.5 % at the eleven-hour horizon.

Keywords Air temperature · Dew point temperature · Artificial neural networks · Decision support system · Frost protection · Weather modeling

1 Introduction

Dew point temperature is the temperature at which humid air, under constant barometric pressure, will cause the water vapor to condense into liquid water. The dew point is the saturation temperature at which water vapor forms droplets on a solid surface. Hence, observed dew point temperature is always less than or equal to the air temperature. Air and dew point temperature predictions could be used to prepare for events such as frost, freeze and heat stress. Frost occurs when water vapor in the air is deposited on a solid surface as ice without turning into liquid water during the transition [24]. A freeze event occurs when the temperature drops below the freezing point of water. Frost damage is caused by the sharp ice crystals which form on the surface of the leaves. The crystals damage the cuticle and epidermis of the leaves, making the plant vulnerable to a further decrease in

K. Nadig (✉)
Institute for Artificial Intelligence, The University of Georgia,
Athens, GA 30602, USA
e-mail: kvnadig@uga.edu

W. Potter
Computer Science, The University of Georgia, Athens,
GA 30602, USA
e-mail: potter@uga.edu

G. Hoogenboom
AgWeatherNet, Washington State University, Prosser, WA, USA

R. McClendon
College of Engineering, The University of Georgia, Athens, GA,
USA

air temperature. Frost damage is mainly due to crystallization of liquid inside the individual cells [24]. Freeze damage occurs when the temperature remains below 0 °C and the amount of damage depends on the length of the freeze event. However, frost damage is usually noticeable due to visible physical impact on the plants. Freeze damage mainly impacts tender parts of the plants such as buds and shoots. This damage may not be evident immediately [24]. Heat stress occurs when the body of a human or animal becomes overheated and the body is unable to regulate the temperature to cool down [7, 20]. Severe cases of heat stress can cause heat stroke and could lead to death if proper and immediate treatment is not provided [10]. Occurrence of heat stress could be estimated using dew point temperature [28].

Low temperature conditions reduce crop yields due to frost damage to leaves and fruit, which could severely affect fruit crops such as blueberries and peaches. The duration of the extreme temperatures also determines the severity of damage that could be caused by the event. Predicting weather variables gives sufficient time to minimize the loss of crops through the use of preventive measures such as orchard heaters and irrigation [12]. Occurrence of frost or freeze events during the growing season can be severely damaging to the agricultural crops [2, 6, 21, 27, 35]. In Spring of 2002, a large area of blueberry and peach crops in South Georgia was destroyed due to unusually severe and unexpected low temperature conditions. In early April of 2007, 50 % of Georgia's peach crop and 87 % of blueberry crop were lost due to frost [8, 34]. As a result of this late frost event, the values of fruits and nuts were down by \$65 million for the year 2007 [3]. Crop managers can minimize these damages by using orchard heaters, irrigation or wind machines if they are given a warning with sufficient time. Accurate weather prediction thus plays a crucial role in managing crops. The National Weather Service (NWS) is the main provider of weather forecasts in the USA and currently issues three levels of frost and freeze warnings [23]. However, NWS collects weather data from urban areas and does not collect weather data from many of the rural areas where the weather conditions have direct impact on the agricultural production. The University of Georgia Automated Environmental Monitoring Network (AEMN) was created for the purpose of collecting weather data in the regions that are not covered by the NWS [13].

Dew point temperature can be used to estimate the amount of moisture in the air, near-surface humidity, evapotranspiration, relative humidity, and frost. Each of these can have an effect on crop production. Plants in arid regions which do not receive frequent rainfall rely on the dew formation. The dew point temperature could also give an insight into the long-term climatic changes [26].

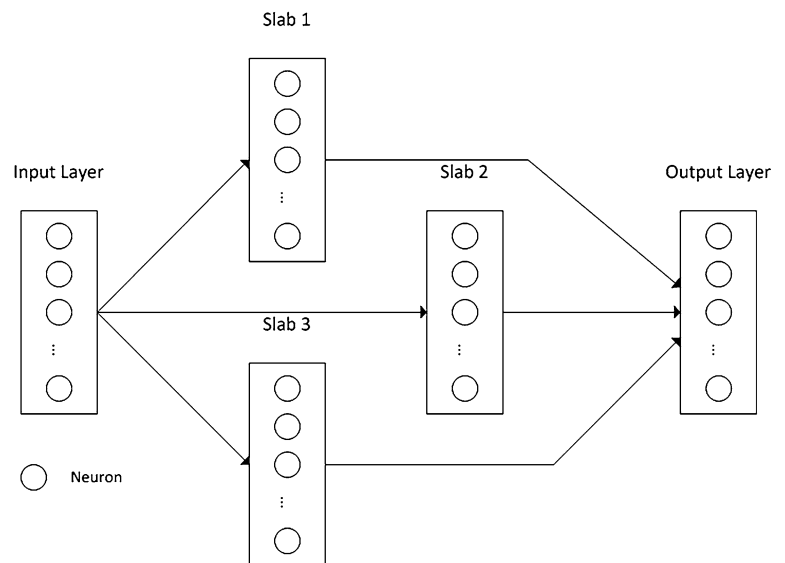
Irrigation is a common method of frost protection, in which a layer of ice forms on the flowers which insu-

lates peach and blue berry blossoms from damages due to frost. To effectively apply these preventive measures, managers need accurate predictions about weather events several hours in advance. A prediction that fails to indicate the occurrence of a frost event could lead to extensive damages to crops. Similarly, a false frost event prediction could cause economic loss due to the cost of unnecessary frost damage prevention procedures.

Prediction of air and dew point temperature relies on prior observations of weather variables, such as air temperature, relative humidity, rainfall, wind speed, solar radiation, vapor pressure, vapor pressure deficit. The infrastructure needed to record the observations is provided by the UGAAEMN, which takes observations every second and calculates an average or total every 15 minutes [14, 16]. This network of weather stations collects data and records essential information required to keep track of the weather conditions, or perform analysis and prediction [15]. The Georgia AEMN network currently has over 80 sites and the data for all the sites are accumulated at a central server located at Griffin, Georgia. Previous research has determined the variables needed for prediction of air temperature and dew point temperature, amount of historical data needed for accurate prediction and other such dependencies [30, 31]. The accumulated data are used as inputs to currently deployed individual ANN models to predict air and dew point temperature for the next twelve hours at an hourly interval. The accumulated data are parsed and the values of the weather variables are extracted from the data which includes current and prior values. This process is applied to both air and dew point temperature models for each of the twelve horizons. The predictions of air temperature and dew point temperature from the 24 models are updated every 15 minutes and made available on the AEMN website www.georgiaweather.net [14].

The ANN models were initially designed to predict air temperature between Winter and early Spring [19]. The inputs to the air temperature ANNs included five weather variables: temperature, relative humidity, wind speed, solar radiation and rainfall. The models were developed using data from the first 100 days of the year. The observations were also restricted to those in which the temperature at the time of prediction was less than or equal to 20 °C. In addition to the current values for each observation, prior data for 24 hours spaced at one hour intervals were also included. Hourly first difference terms for the current and prior weather variables were also included. The models needed 24 hours of prior observations, time of day, seasonal terms as input, and the models had 120 neurons in the hidden layer, and they predicted air temperature for a particular horizon [18]. The ANN models were improved by Smith et al. [32] by adding the time of day and day of year as a part of the input, after transforming them using fuzzy logic membership

Fig. 1 Ward Style ANN model [19]



functions, with a resulting decrease in MAE ranging from 6 % to 14 % for the twelve prediction horizons. However their models were only designed to predict between Winter and early Spring and had the same limitations of the previous models. A model was developed for each of the twelve prediction horizons. Air temperature ANN models were developed to predict year-round by Smith et al. [31] and implemented on the AEMN website. These ANN models were used to generate short term air temperature predictions by the AEMN. The ANNs were based on Ward-style network architecture [33] and were trained using the error back-propagation (EBP) [11]. The input layer of the model consisted of 258 neurons for inputs. The hidden layer of the model consisted of 120 nodes equally distributed among the three slabs with hyperbolic tangent, Gaussian and inverse Gaussian as the activation functions, Fig. 1. The prediction accuracy of the year-round models was comparable to the previous Winter models, yet was developed to predict air temperature throughout the year. They found that unanticipated cooling events were the most significant obstacle with the year-round models. They varied several ANN parameters, such as the activation function of outputs, number of hours of prior data, additional values and rate of change for observations at 15 min intervals and the data scaling ranges for both input and output. However, no improvement in accuracy was produced. Bagging and boosting only slightly improved the accuracy, but at a high computational cost [31].

Support Vector Machine (SVM) based regression models were also developed to predict air temperature and the accuracy was compared with the existing ANNs [4]. For a reduced training set with 300,000 patterns, the SVM models were slightly more accurate than the ANN models. However, the ANN models predicted more accurately when the number of training patterns was increased to 1.5 million.

ANN models for twelve hour prediction of dew point temperature were developed by Shank et al. [30] and are included on the AEMN website. Inputs to the dew point ANNs included the same weather variables used in the existing air temperature models, plus weather variables vapor pressure and vapor pressure deficit, and their hourly rates of change. The models were developed similarly to those developed by Smith et al. [31]. Ensemble artificial neural network were developed by Shank et al. [29] to improve the prediction accuracy. Approximately four years of weather data were available that included the additional weather variables [30]. The twelve ANN models to predict dew point were implemented on the AEMN site similar to the air temperature models.

A fuzzy expert system, Georgia Extreme-weather Neural-network Informed Expert (GENIE), was developed to interpret the predicted air temperature, predicted dew point temperature and the observed wind speed in order to generate frost and freeze warning levels [5]. The numeric warning levels generate by GENIE provide higher granularity than the textual warnings provided by the National Weather Service (NWS). A web interface was developed for GENIE to provide a convenient means of access to the warnings.

The air temperature and dew point temperature models have prediction errors measured in terms of mean absolute error (MAE). The MAE for the twelve air temperature models varies between 0.516 °C and 1.873 °C. Similarly, the MAE for dew point temperature models varies between 0.508 °C and 2.081 °C. Under high relative humidity conditions, the observed air temperature will approach the observed dew point temperature, but never goes below it. Under these conditions, the predicted air temperature will frequently drop below the predicted dew point temperature, herein referred to as a prediction anomaly. Further improvement of the ANN models for air and dew point temperature

Table 1 Locations used for training, selection and evaluation datasets

Model development 2002, 2003, 2004, 2005, 2007, 2009		Model evaluation 2006, 2008, 2010
Training set	Selection set	Evaluation set
ATLANTA	ALMA	ALAPAHA
BRUNSWICK	ARABI	ALPHARETTA
CALLAWAY	BLEDSON	ARLINGTON
COVINGTON	BOWEN	ATTAPULGUS
DALLAS	DEMPSEY	BLUE RIDGE
DAWSON	DIXIE	BYROMVILLE
DEARING	EATONTON	CAIRO
DULUTH	GEORGETOWN	CALHOUN
HOMERVILLE	GRIFFIN	CAMILLA
OAKWOOD	HOWARD	CLARKS HILL
SHELLMAN	JEFFERSONVILLE	CORDELE
TIFTON	LAFAYETTE	DANIELSVILLE
TIGER	PLAINS	DOUGLAS
WOODBINE	SPARTA	ELLIJAY
–	TENNILLE	HHERC ^a
–	–	MOULTRIE
–	–	NAHUNTA
–	–	NEWTON
–	–	ODUM
–	–	OSSABAW
–	–	SASSER
–	–	SAVANNAH
–	–	VALDOSTA
–	–	VIDALIA

^aHooks-Hanner Environmental Resource Center (HHERC)

involves not only reducing the MAE but also reducing the number of prediction anomalies.

The architecture of the ANNs plays a key role in the prediction capabilities of the ANNs. Previous research explored the ANN parameters such as the nodes in the hidden layers, varying the number of inputs, and larger datasets. Another possible ANN architecture parameter is the number of outputs. Impact on each output according to the inputs for ANN based models was examined by Gevrey et al. [9], and the influence of outputs on learning in ANN models was examined by Narendra and Mukhopadhyay [22]. The current implemented air and dew point temperature ANN models have only one output. Additional outputs could be different weather variable or other prediction horizons. A model that predicts more than one weather variable is herein referred to as the combined model. Such a model could predict air and dew point temperature for a single or multiple prediction horizons. Predicting multiple values in a single model provides an opportunity for interaction among the outputs.

The goal of this research was to improve the prediction accuracy of air temperature and dew point temperature ANN models by developing combined models to predict both air and dew point temperature for each prediction horizon. The research objectives are as follows: (1) to determine if MAE

for predicted air temperature and predicted dew point temperature are reduced for the combined model in comparison with the individual models, and (2) to determine if the number of occurrences of the prediction anomaly can be reduced using the combined models. The hypothesis is that by predicting air and dew point temperature for a single prediction horizon in a single combined model, the prediction MAE may be reduced and the number of prediction anomalies may decrease. The air and dew point temperature are related weather variables, thus predicting both in a single model might aid in reducing the MAE and decreasing the number of prediction anomalies. These anomalies do not occur in the observed data; thus, training the combined models might decrease the number of prediction anomalies.

2 Methodology

2.1 Data sets

The AEMN measures weather variables each second and then stores the averages or totals of the values every 15 minutes. Data from the initial sites were available from 1991. However, the data selected for this research were from 2002 to 2010 because the dew point temperature observations were initiated in 2002. Routines were developed to perform error checking on the raw data to remove missing variables, incomplete records, instrument malfunctions and erroneous records. The data were partitioned into a model development set and a model evaluation set. The two sets were chosen so that they were mutually exclusive of years and locations, as shown in Table 1. The model development set was further partitioned into a training set and a selection set. The training set was used to train the ANN models using resilient propagation [25] to adjust the ANN weights. The selection set was presented to the ANNs in feed forward mode only to choose the model with the lowest MAE. The chosen model was treated as the final model for a given prediction horizon. The training set and selection set were mutually exclusive by locations. The model evaluation set was presented to the ANNs in feed forward mode and the resulting MAE was used as a metric to compare with other models. This method of partitioning datasets and training was used to avoid over-training the models. The training set consisted of 297,974 patterns, the selection set had 306,972 patterns and the evaluation set included 507,347 patterns. This would be the strongest evaluation of the models, since the final ANNs will be evaluated with data from sites and years which were not used in model development. The approach to partitioning the data was modified from the partitioning used to develop the current web-based ANN temperature models in order to include the additional sites and years. The current web-based ANN air temperature models were

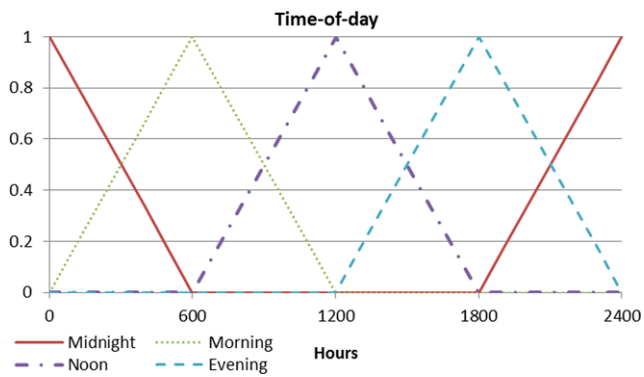


Fig. 2 Fuzzy membership functions for the cyclic time of the day input variable

developed with data from 1997 to 2005 and over 20 locations [31]. The current web-based ANN dew point temperature models were developed with data from 2002 to 2005, with over 20 locations [30]. Additional years from 2006 to 2010 and locations were included herein to provide for more robust models.

The data partitions were subjected to constraints to ensure a fair distribution of patterns. The locations and years involved in the partitioning were chosen to minimize the difference in the range and average air temperatures between model development and evaluation. The years and locations were distributed among the sets until the difference in the range and average air temperatures among the three sets was minimized. This was done to ensure that the model development set and the model evaluation set were representative of the population.

Input and output patterns were generated from each of the datasets by transforming and scaling the data. The input patterns consisted of current and 24-hours of prior hourly values of air temperature, relative humidity, rainfall, wind speed, solar radiation, vapor pressure, vapor pressure deficit and their hourly rates of change. Each input pattern also had the time-of-day and day-of-year cyclic values obtained by using triangular fuzzy membership functions, and were shown in [32] to cause considerable reduction to the MAE. The membership functions to convert the time-of-day input variable are shown in Fig. 2 and the membership functions to convert the day-of-year input variable are shown in Fig. 3. The pattern included observed air and dew point temperature as targets for the twelve prediction horizons. All the values in the pattern were scaled to the range $[-0.9, 0.9]$ since the domain of operation of the activation functions used in the ANNs was in the range $[-1, 1]$. The range $[-0.9, 0.9]$ was chosen since it captured the range of values for each of the variables and transformed them to the domain of the activation functions. Inverse scaling was used to transform the value of prediction from the domain of the activation function to the domain of the predicted variable. There were a total of 358 inputs and two output values per pattern. Based on

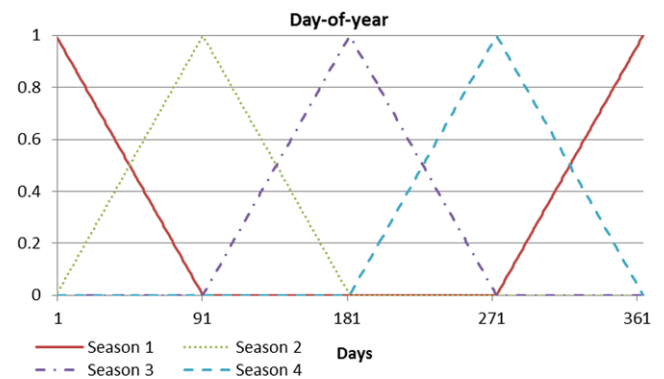


Fig. 3 Fuzzy membership functions to convert the day of year input variable to seasons

prior research, the “air temperature only” model produced the lowest MAE with 258 inputs and adding other weather variables did not show any improvement [19, 31]. Similarly, the dew point temperature model produced the lowest MAE with 358 inputs [30]. Therefore, the previously determined number of inputs was used herein for comparison. Patterns were generated for each of the twelve prediction horizons and the three data partitions.

2.2 Model development

All models were developed using the Ward-style network architecture [33], consisting of a three layered neural network with input, hidden and output layers, Fig. 1. The input layer consisted of neurons with linear activation functions. The hidden layer consisted of 120 neurons in three equally sized slabs of 40 neurons. The neurons in the slabs had hyperbolic tangent, Gaussian and inverse Gaussian activation functions. The output layer neurons used the symmetric sigmoid activation function. Only the number of inputs or outputs varied with the models.

Ten instances of each model were trained. Each instance is an ANN whose initial weights were selected randomly. Although all the model instances were presented with the same set of patterns, the order was randomized. This provided the training algorithm a different starting point for each instance. Thus the training process took a different path while searching for the set of weights which could minimize the prediction error and improve the accuracy of the model. Selecting among multiple instances made it more likely that the training algorithm will approach the optimal set of ANN weights [31].

Resilient propagation (R-Prop) training algorithm was used to train the models [25]. R-Prop tends to converge faster and is more stable in comparison with the back-propagation algorithm [1, 17]. R-Prop was used in its basic configuration. It is similar to the back-propagation algorithm except that the error update in R-Prop is dependent on a constant value. This constant value exists for each synapse and

Table 2 Comparison of the individual ANN models for air and dew point temperature with combined models, evaluation dataset

Horizons	Air temperature MAE °C			Dew point temperature MAE °C		
	Individual	Combined	% Reduction in MAE	Individual	Combined	% Reduction in MAE
1	0.889	0.846	4.84	0.818	0.842	-2.93
2	0.959	0.942	1.77	0.890	0.880	1.12
3	1.170	1.080	7.69	1.073	1.081	-0.75
4	1.262	1.256	0.48	1.261	1.264	-0.24
5	1.413	1.383	2.12	1.424	1.416	0.56
6	1.565	1.543	1.41	1.600	1.585	0.94
7	1.617	1.621	-0.25	1.711	1.720	-0.53
8	1.707	1.705	0.12	1.836	1.809	1.47
9	1.825	1.787	2.08	1.954	1.965	-0.56
10	1.848	1.830	0.97	2.052	2.030	1.07
11	1.916	1.918	-0.10	2.155	2.151	0.19
12	2.016	1.975	2.03	2.247	2.276	-1.29
Models	2/12	10/12	(Avg.) 1.93	6/12	6/12	(Avg.) -0.08

is updated by increasing or decreasing the constant based on the sign of the error value. If the sign of the error value changes frequently, then the magnitude of the constant value is decreased. If the sign does not change, then the magnitude of the constant value is increased. The sign of the error is used to determine the direction of change. Resilient propagation does not depend on the derivative of the activation function [25]. EnCog 3.0.0.0 (runtime v2.0.50727) package was used to develop and train the models. The unit training step or learning event begins with the model presented with input values from the training dataset. The values are passed forward through the model to obtain a prediction. The predicted values are compared with the measured values. The difference in the observed and predicted values is the error that will be used to adjust the weights as it is propagated backwards from the output layer to input layer.

All models were trained using the training dataset, to avoid under-training, until the change in error was less than 0.01 %. After training the models, the ten instances were presented with the selection set in feed forward mode to obtain the selection set MAEs. The instance with the lowest selection set MAE was chosen. That completed the model development part of the process. The selected models were then presented with the evaluation set once in feed forward mode to obtain the evaluation set MAE.

Individual ANNs to predict air temperature and dew point temperature were developed using the partitioned data as a base line for the prediction accuracy in terms of MAE. These individual models correspond to the existing air temperature and dew point temperature ANN models currently implemented on the AEMN. However, they take advantage of additional years and sites. The individual models have only one output since they were designed to predict either air temperature or dew point temperature. Each model was developed to predict a single weather variable for a single horizon. Hence, there were a total of 24 models, twelve for

air temperature and twelve for dew point temperature. The individual air temperature model had 258 input neurons and one output neuron based on Smith et al. [31]. The dew point temperature model had 358 input neurons and one output neuron, based on Shank et al. [30]. In this research, only the combined model that predicts air and dew point temperature for a single prediction horizon is considered. The combined model had 358 input neurons and the two outputs predicted air temperature and dew point temperature. The MAEs for individual and combined models were calculated and compared.

While obtaining the evaluation set MAEs for both individual and combined models, the occurrence of prediction anomalies was recorded. The predicted air temperature and the predicted dew point temperature obtained from the individual models of corresponding prediction horizons were compared. The number of instances where the predicted air temperature was lower than the predicted dew point temperature was calculated for each prediction horizon. Similarly, the number of instances of the anomaly was computed for the combined model. The number of instances of anomaly for the individual and combined models was calculated and compared. The day of the year was used to compare prediction anomalies by seasons. The count of prediction anomalies was grouped by seasons and prediction horizons to analyze the seasonal variation in the ANN models.

3 Results and discussion

The air temperature and dew point temperature MAE values were obtained by presenting the evaluation dataset to the individual and combined ANNs in feed forward mode only as shown in Table 2. For air temperature, ten of twelve combined models produced an MAE lower than the individual models. The combined model showed an average reduction

Fig. 4 Comparison of air temperature MAEs for prediction horizons, individual and combined models, model evaluation set

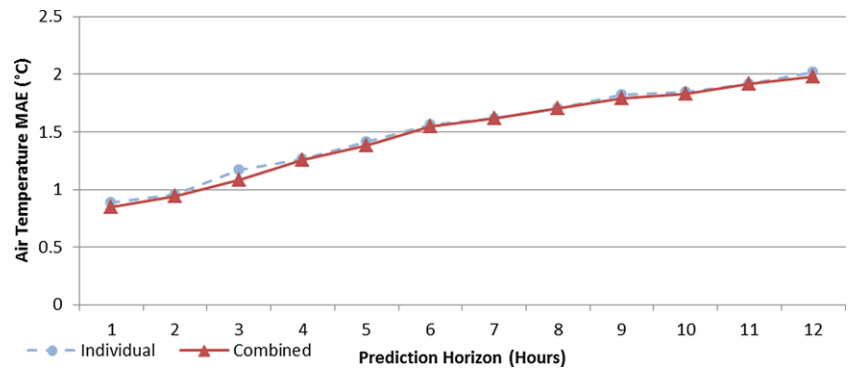
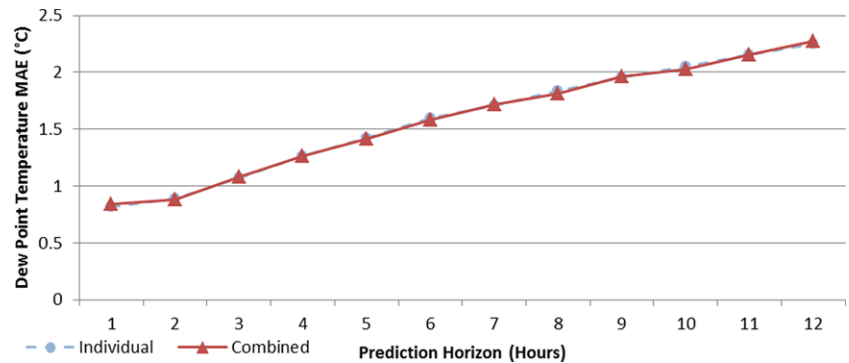


Fig. 5 Comparison of dew point temperature MAEs for prediction horizons, individual and combined models, model evaluation dataset



in MAE for air temperature by 1.93 %. The two prediction horizons in which the individual model provided a lower MAE were the seven and eleven hour horizons. Figure 4 shows the generally observed trend of increasing MAEs for longer horizons, with a slight decrease in MAEs for air temperature prediction from the combined model. This suggests that the combined model was able to predict the air temperature with a lower MAE when predicting both air temperature and dew point temperature. From Table 2, six of twelve combined models predicted the dew point temperature with lower MAE than the individual dew point temperature models. The combined model produced a slight increase in MAE and the average increase in MAE was 0.08 %. Both the individual and the combined models maintained the expected trend of increasing MAE with longer horizons as shown in Fig. 5. From Table 2, of the ten combined models with lower MAE for air temperature, five also produced a lower MAE for dew point temperature in comparison with individual models.

Second approach to assess the accuracy of the ANNs was performed by determining the number of prediction anomalies found with individual and combined models, as shown in Table 3. The models were presented with the evaluation set in feed forward mode to count the number of prediction anomalies. The evaluation set had 507,347 patterns, and the number of prediction anomalies for the worst case was 6.55 % of the evaluation set patterns. Twelve of twelve combined ANNs model showed a reduction in the number of prediction anomalies over the individual models. The av-

Table 3 Number of prediction anomalies for individual and combined model, model evaluation dataset

Horizons	Individual		Combined		% Reduction
	#	%	#	%	
1	33252	6.6	31723	6.3	4.6
2	21161	4.2	12857	2.5	39.24
3	14367	2.8	7331	1.4	48.97
4	18775	3.7	11142	2.2	40.66
5	17885	3.5	11394	2.2	36.29
6	16593	3.3	11989	2.4	27.75
7	19345	3.8	15054	3.0	22.18
8	15595	3.1	13292	2.6	14.77
9	11845	2.3	8281	1.6	30.09
10	12815	2.5	9061	1.8	29.29
11	21833	4.3	8619	1.7	60.52
12	12676	2.5	5723	1.1	54.85
Total	216142		146466		(Average) 34.10

Note: The prediction anomalies were obtained after presenting the models with all 507,347 patterns from the evaluation set

erage reduction in prediction anomalies for the combined model was 34.1 %, and the reduction ranged from 4.6 % at the one-hour horizon to 60.5 % at the eleven-hour horizon. Figure 6 shows a comparison between the number of prediction anomalies found in individual and combined models for each prediction horizon. The highest number of prediction anomalies occurred at the one hour horizon for both individual and combined models. The combined models provided

Fig. 6 Comparison of number of prediction anomalies for individual and combined models, model evaluation dataset

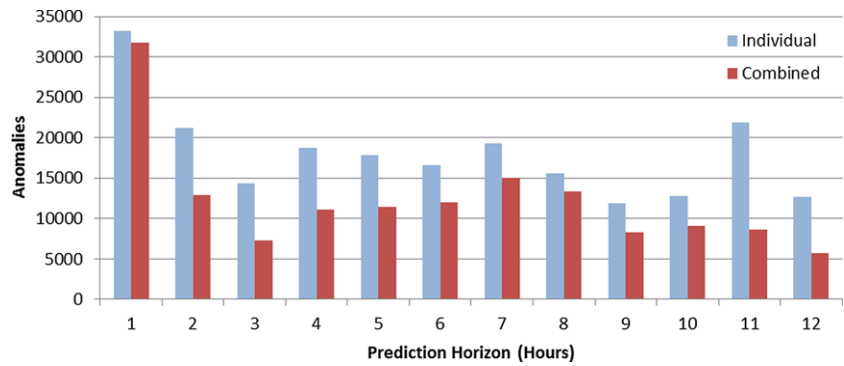


Fig. 7 Comparison of anomalies by season, summed over the twelve prediction horizons, model evaluation dataset

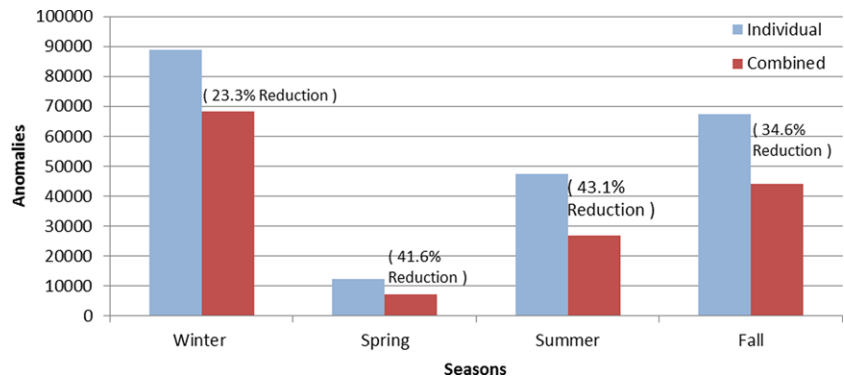


Table 4 Comparison of number of anomalies in individual (indv.) and combined (comb.) models by seasons, model evaluation dataset

Horizons	Winter		Spring		Summer		Fall	
	Indv.	Comb.	Indv.	Comb.	Indv.	Comb.	Indv.	Comb.
1	17570	21310	2610	1581	5565	4842	7507	3990
2	9168	7570	1002	575	5060	1433	5931	3279
3	6814	3985	1109	402	2037	1233	4407	1711
4	6267	4490	1296	133	5619	2911	5593	3608
5	7709	6068	554	899	5405	1347	4217	3080
6	6897	4797	1373	808	4231	2176	4092	4208
7	6975	5262	705	322	5239	4283	6426	5187
8	4795	3883	433	815	4103	4073	6264	4521
9	5571	2714	977	310	1971	1415	3326	3842
10	6345	3201	527	919	2009	1623	3934	3318
11	7106	2399	987	396	5037	1297	8703	4527
12	3704	2515	863	107	1138	333	6971	2768
Total	88921	68194	12436	7267	47414	26966	67371	44039
	1/12	11/12	3/12	9/12	0/12	12/12	2/12	10/12

Note: The prediction anomalies were obtained after presenting the models with all 507,347 patterns from the evaluation set

a slight reduction in the number of prediction anomalies for the one-hour horizon. Other horizons showed a marked reduction in the number of prediction anomalies.

The models were further examined to compare the occurrence of prediction anomalies by season. The prediction anomalies were classified into the following seasons: Winter (Dec–Feb), Spring (Mar–May), Summer (Jun–Aug) and Fall (Sep–Nov). Figure 7 compares the prediction anomalies for individual and combined models for all twelve prediction horizons, classified by seasons. The highest number of prediction anomalies occurred in the Winter season, with reduc-

tion in prediction anomalies of 23.3 %. The lowest number of prediction anomalies occurred in the Spring season, with a reduction of 41.6 %. In Summer, the combined models produced the highest reduction in the number of prediction anomalies with a value of 43.1 %. In Fall, the number of prediction anomalies was slightly higher than the number of prediction anomalies produced during both Spring and Summer, and the combined models produced a reduction in the number of anomalies by 34.6 %. The number of prediction anomalies for the individual and combined ANNs by season and horizons is shown in Table 4. In Winter, eleven of

Table 5 Distribution of prediction anomalies by severity of error, combined models, model evaluation dataset

Horizons (°C) (Hr) T _{DP} - T _A	1	2	3	4	5	6	7	8	9	10	11	12
	0.00 - 0.25	10034	7164	4439	7536	5814	5830	6283	6008	4742	4306	3967
0.25 - 0.50	7678	3176	1705	2563	2694	3036	3996	3510	2183	2419	2334	1520
0.50 - 0.75	5279	1515	710	712	1258	1534	2279	2009	769	1177	1188	547
0.75 - 1.00	3645	631	267	204	631	732	1153	996	244	544	577	167
% severity 0.00 - 1.00	5.25	2.46	1.40	2.17	2.05	2.19	2.70	2.47	1.56	1.66	1.59	1.10
1.00 - 1.25	2320	267	117	76	327	391	616	454	166	292	261	66
1.25 - 1.50	1424	75	44	35	241	205	313	175	81	155	137	28
1.50 - 1.75	780	19	22	11	126	114	161	75	29	77	62	20
1.75 - 2.00	372	8	10	4	97	67	99	41	31	42	39	8
% severity 0.00 - 2.00	6.22	2.53	1.44	2.20	2.21	2.35	2.94	2.62	1.63	1.78	1.69	1.13
2.00 - 2.25	137	2	1	1	68	37	59	11	16	15	21	1
2.25 - 2.50	32	0	4	0	39	18	38	8	9	17	10	0
2.50 - 2.75	16	0	6	0	26	11	20	2	3	4	13	0
2.75 - 3.00	5	0	2	0	19	8	11	2	7	7	4	0
% severity 0.00 - 3.00	6.25	2.53	1.44	2.20	2.24	2.36	2.96	2.62	1.63	1.78	1.70	1.13
3.00 - 3.25	1	0	2	0	19	0	9	1	1	2	3	0
3.25 - 3.50	0	0	2	0	14	4	5	0	0	3	0	0
3.50 - 3.75	0	0	0	0	6	1	3	0	0	0	3	0
3.75 - 4.00	0	0	0	0	8	1	3	0	0	1	0	0
4.00 - 4.25	0	0	0	0	6	0	0	0	0	0	0	0
4.25 - 4.50	0	0	0	0	0	0	5	0	0	0	0	0
4.50 - 4.75	0	0	0	0	1	0	0	0	0	0	0	0
4.75 - 5.00	0	0	0	0	0	0	1	0	0	0	0	0
Total	31723	12857	7331	11142	11394	11989	15054	13292	8281	9061	8619	5723

Note: The prediction anomalies were obtained after presenting the models with all 507,347 patterns from the evaluation set

twelve combined models reduced the number of prediction anomalies in comparison with the individual models. During the Spring season only nine of twelve combined models showed reduction in the number of prediction anomalies. Consistent reduction was produced in Summer where twelve of twelve combined models showed reduction and in Fall ten of twelve combined models showed reduction. The total row shows the sum of the number of prediction anomalies that occurred for all prediction horizons for each season using the evaluation set.

The anomalies from the combined models were further analyzed to determine the extent to which the air temperature prediction dropped below the predicted dew point temperature or severity. The severity of the prediction anomaly was classified using increments of 0.25 °C as shown in Table 5. Green indicates few or no prediction anomalies. Yellow indicates that the number of prediction anomalies is in the range greater than 95 and less than or equal to 2000. Red or orange indicates that the number of prediction anomalies is greater than 2000. Table 5 shows that a large portion of the anomalies are between 0 °C and 1.5 °C. The combined models produced a total of 146,466 prediction anomalies.

The number of prediction anomalies with severity greater than 1 °C was 11395, or 0.2 % of the number of predictions in the evaluation set for all prediction horizons. The two-hour, four-hour and twelve-hour horizon models did not generate any anomalies greater than 2.25 °C. The one-hour, eight-hour and nine-hour horizon models did not generate any anomalies greater than 3.25 °C. Only 105 of the 146,466 prediction anomalies, for the combined models across all twelve prediction horizons, had severity greater than 3 °C. Table 6 shows a similar analysis of the individual models. Green indicates few or no prediction anomalies. Yellow indicates that the number of prediction anomalies is in the range greater than 95 and less than or equal to 2000. Red or orange indicates that the number of prediction anomalies is greater than 2000. The individual models showed higher prevalence of prediction anomalies, produced a total of 216,142 prediction anomalies. Approximately 10.7 %, or 23067 prediction anomalies, of the total number of prediction anomalies across all twelve prediction horizons had severity greater than 1 °C. However, 310 of the 216,142 prediction anomalies, for the combined models across all twelve prediction horizons, had severity greater than 3 °C. The combined

Table 6 Distribution of prediction anomalies by severity of error, individual models, model evaluation dataset

Horizons (°C) (Hr) $T_{DP} - T_A$	Horizons											
	1	2	3	4	5	6	7	8	9	10	11	12
0.00 - 0.25	11105	10227	9349	8464	7660	7635	7626	6080	5334	4836	6366	5941
0.25 - 0.50	8125	5816	2517	5065	4682	4741	5243	4112	2531	2994	4757	2567
0.50 - 0.75	5583	2853	1168	2596	2572	1357	3168	2529	1990	1740	3472	1513
0.75 - 1.00	3725	1276	973	1311	1380	1422	1550	1399	1029	1108	2445	1143
% severity 0.00 - 1.00	5.62	3.98	2.76	3.44	3.21	2.99	3.47	2.78	2.15	2.10	3.36	2.20
1.00 - 1.25	2220	555	257	658	770	851	861	656	551	716	1659	752
1.25 - 1.50	1250	231	83	327	405	417	432	342	222	474	1148	571
1.50 - 1.75	634	118	15	185	198	115	237	183	109	332	654	116
1.75 - 2.00	295	46	4	108	88	42	108	110	61	220	461	51
% severity 0.00 - 2.00	6.49	4.16	2.83	3.69	3.50	3.27	3.79	3.04	2.33	2.45	4.13	2.49
2.00 - 2.25	158	27	0	38	52	9	65	68	12	144	321	11
2.25 - 2.50	79	5	1	14	26	3	24	48	3	100	176	7
2.50 - 2.75	32	4	0	8	23	0	17	20	0	67	128	3
2.75 - 3.00	16	1	0	1	10	1	6	16	2	30	82	1
% severity 0.00 - 3.00	6.55	4.17	2.83	3.70	3.52	3.27	3.81	3.07	2.33	2.52	4.27	2.50
3.00 - 3.25	12	2	0	0	5	0	3	15	1	15	52	0
3.25 - 3.50	6	0	0	0	4	0	2	9	0	10	37	0
3.50 - 3.75	5	0	0	0	3	0	2	3	0	8	28	0
3.75 - 4.00	7	0	0	0	0	0	1	4	0	8	12	0
4.00 - 4.25	0	0	0	0	2	0	0	1	0	4	13	0
4.25 - 4.50	0	0	0	0	5	0	0	0	0	3	8	0
4.50 - 4.75	0	0	0	0	0	0	0	0	0	2	3	0
4.75 - 5.00	0	0	0	0	0	0	0	0	0	1	5	0
5.00 - 5.25	0	0	0	0	0	0	0	0	0	0	1	0
5.25 - 5.50	0	0	0	0	0	0	0	0	0	1	1	0
5.50 - 5.75	0	0	0	0	0	0	0	0	0	0	2	0
5.75 - 6.00	0	0	0	0	0	0	0	0	0	1	0	0
6.00 - 6.75	0	0	0	0	0	0	0	0	0	3	0	0
Total	33252	21161	14367	18775	17885	16593	19345	15595	11845	12815	21833	12676

Note: The prediction anomalies were obtained after presenting the models with all 507,347 patterns from the evaluation set

models showed an overall reduction of 32.2 % in the total number of prediction anomalies across all twelve horizons. The combined models produced similar MAEs as the individual models and the combined models considerably reduced the prediction anomalies. Based on these two metrics, the results suggests that combined models are more accurate than individual models.

Figure 8 shows the scatter plot for the air temperature predictions using the combined model for the prediction horizons of one, three, six, nine and twelve. As expected, the observed scatter about the 1:1 line increases and the R2 value decreases as the prediction horizon increases. The plot for the one hour horizon, Fig. 8a, shows a narrower scatter, which indicates a large portion of the predicted air temperatures are in close proximity to the observed value. The regression line had a slope of 0.96, the intercept was

0.69 and the R2 was 0.98. At low observed temperatures the model tends to over-predict and at high observed temperatures the model tends to under-predict. This trend is observed in other horizons as well. The scatter plot for three-hour horizon model, shown in Fig. 8b, the regression line had a slope of 0.97, the intercept was 0.63 and the R2 was 0.97. The scatter plot for the six-hour model had a slightly greater distribution about the 1:1 line as compared with the plot for the three-hour horizon model, as shown in Fig. 8c. The regression line had a slope of 0.94, and the intercept was 1.21, and the R2 was 0.95. The nine hour scatter plot had observably greater distribution about the 1:1 line, as shown in Fig. 8d. The regression line had a slope of 0.91, and the intercept was 1.6, and the R2 was 0.92. This was expected as the MAEs of the six and nine models are higher than that of the one-hour model. Lastly, the twelve

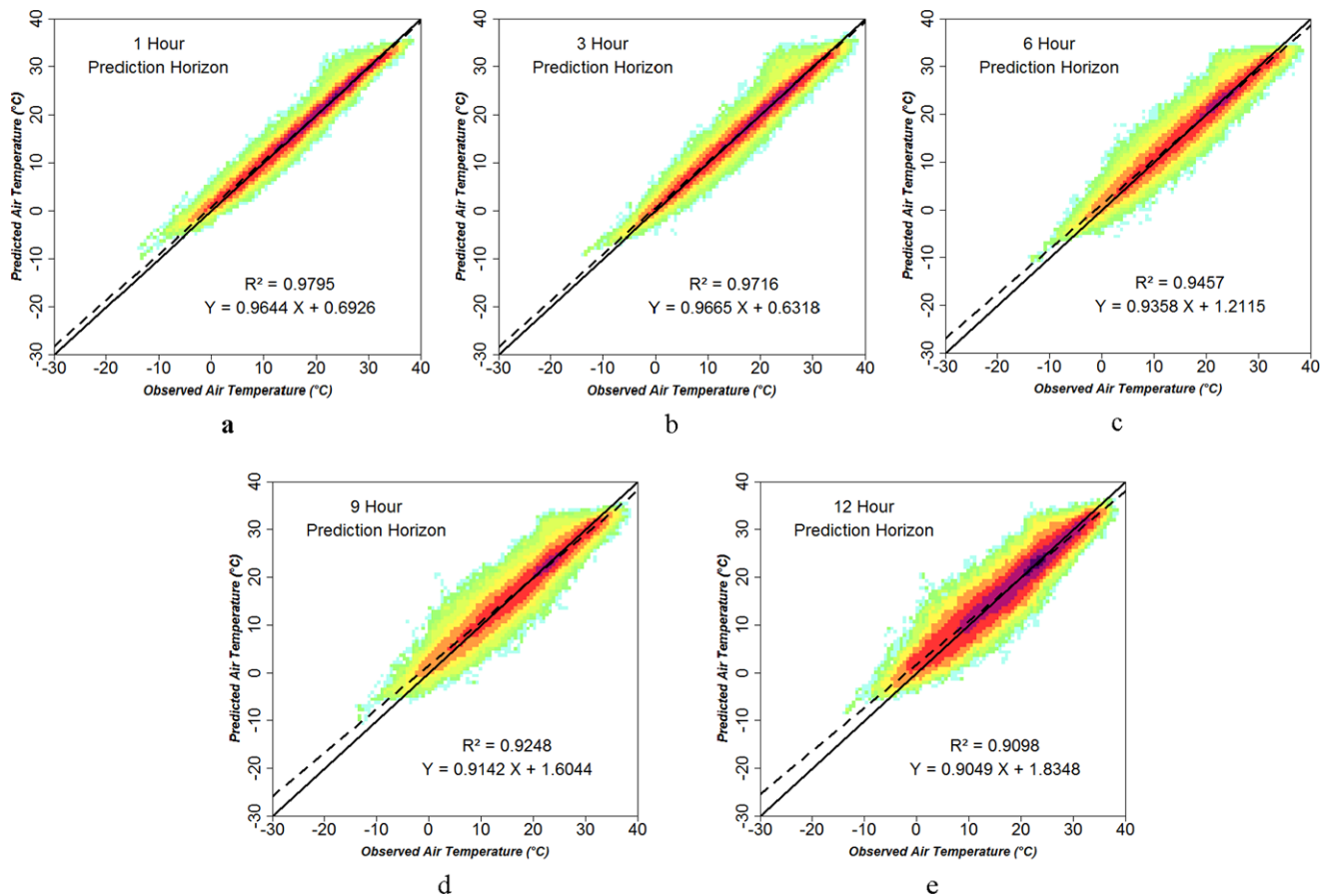


Fig. 8 Predicted and observed air temperature, combined model, model evaluation dataset

hour scatter plot had markedly greater distribution about the 1:1 line, as shown in Fig. 8e, and it was the highest distribution among the other horizons. The regression line had a slope of 0.9, and the intercept was 1.8, and the R^2 was 0.91.

Dew point temperature predictions from the combined model were used to generate the scatter plots shown in Fig. 9, for the prediction horizons one, three, six, nine and twelve. The low observed temperatures were over-predicted and high observed temperatures were under-predicted. The plots show a similar trend that was observed in air temperature scatter plots. The one-hour horizon plots, Fig. 9a, shows a narrow dense region but a much greater distribution about the 1:1 for the low density region as compared with the scatter plot of one-hour horizon air temperature model shown in Fig. 8a. In the scatter plot for one-hour horizon dew point temperature model, as shown in Fig. 9a, the regression line had a slope of 0.97, the intercept was 0.4 and the R^2 was 0.98. In the scatter plot for three-hour horizon model, as shown in Fig. 9b, the dense region was similar to dew point temperature scatter plot for the one-hour horizon model, but showed greater distribution about the 1:1 line for the low density region. The regression line had a slope of

0.97, the intercept was 0.26 and the R^2 was 0.97, for the three-hour horizon model. Both the scatter plots of one and three hour horizons for dew point temperature are narrow compared to the six, nine and twelve hour horizons scatter plots. The scatter plot for the six-hour model, as shown in Fig. 9c, had greater distribution about the 1:1 line as compared to that of the three-hour model. The regression line had a slope of 0.94, the intercept was 0.64 and the R^2 was 0.94. The scatter plots for the nine and six hour model were similar in the distribution about the 1:1 line. However, the nine-hour model, as shown in Fig. 9d, had slightly greater distribution about the 1:1 line and the regression line had a slope of 0.91, the intercept was 0.95 and the R^2 was 0.91. Lastly, the scatter plot for the twelve-hour model, as shown in Fig. 9e, had the highest distribution about the 1:1 line. The slope of 0.88 was the lowest and the intercept of 1.23, was the highest among the other horizons. The R^2 for the twelve-hour horizon model was 0.88, which was lowest among the other horizons. The dew point temperature scatter plots had greater distribution about the 1:1 line as compared with the corresponding air temperature scatter plots, possibly due to higher MAEs of the dew point temperature models.

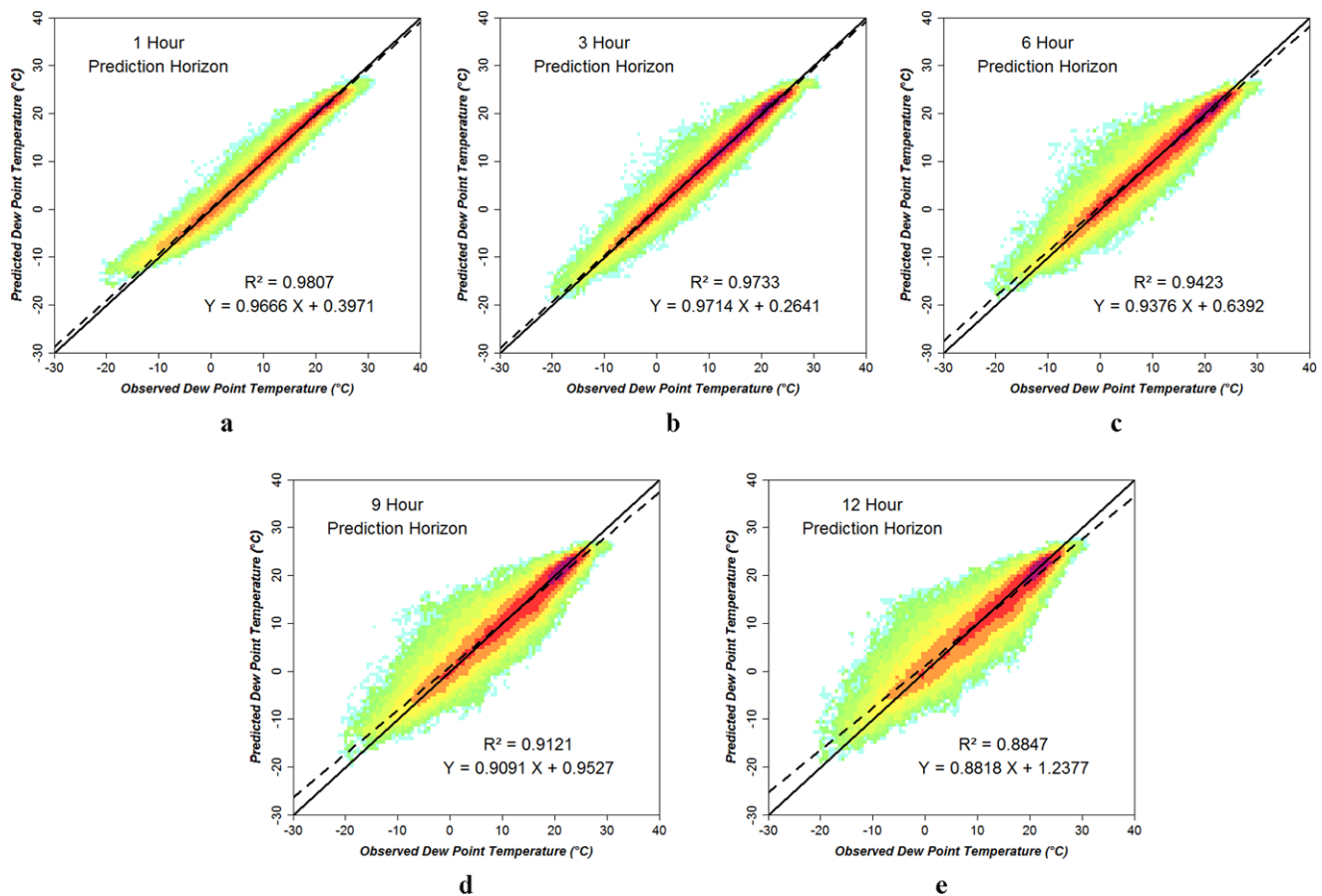


Fig. 9 Predicted and observed dew point temperature, combined model, model evaluation dataset

4 Summary and conclusions

Combined air temperature and dew point temperature models were developed for the twelve prediction horizons. The air temperature predictions from the combined model showed reduction in MAE for ten of twelve prediction horizons over the corresponding individual models, with a corresponding average reduction in MAE of 1.9%. The dew point temperature predictions from the combined model showed reduction in MAE for six of twelve prediction horizons over the corresponding individual models. However, averaging over the twelve prediction horizons showed that there was essentially no difference in MAE for the dew point temperature predictions. The combined models showed a marked reduction in the number of prediction anomalies as compared with the individual models. Also, experiments showed that the anomalies occurred most often in Winter and least frequently in the Spring season for individual and combined models. The combined models reduced prediction anomalies for each season, with reduction ranging from 23.3% in Winter to 43.1% in Summer.

In this research, the ANN architecture used was based on previous work by Smith et al. [31] and Shank et al.

[30]. In future research the ANN parameters such as activation functions, number of nodes in the hidden layer, and distribution of nodes between the slabs of the Ward-style model could be explored for improvement in MAE or reduction in the number of anomalies. Also, longer duration of prior data, the inclusion of other weather variables, and different resolution of input data could be explored. Alternate architectures such as recurrent neural networks, hybrid neural networks, and alternative training algorithms such as scaled conjugate gradient propagation, quick-propagation, Manhattan-propagation, Lavenberg-Marquardt algorithm and evolutionary training algorithms could be applied. The training data could be further examined so that the patterns are distributed evenly for the various temperature values. Future work could also be focused on combining the models for various prediction horizons into a single model. This could include combining a time series of one through twelve prediction horizons for air temperature or dew point temperature or both into a single model.

Acknowledgements This work was funded by a partnership between the USDA-Federal Crop Insurance Corporation through the Risk Management Agency and the University of Georgia and by state and

federal funds allocated to Georgia Agricultural Experiment Stations Hatch projects GEO00877 and GEO01654.

References

- Anastasiadis AD, Magoulas GD, Vrahatas MN (2005) New globally convergent training scheme based on the resilient propagation algorithm. *Neurocomputing* 64:253–270
- Attaway JA (1997) A history of Florida Citrus freezes. Florida Science Source, Inc, Lake Alfred
- Boatright SR, McKissick JC (2008) 2007 Georgia farm gate value report. Center for Agribusiness and Economic Development. The University of Georgia. AR-08-01
- Chevalier R, Hoogenboom G, McClendon R, Paz J (2011) Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. *Neural Comput Appl* 20:151–159
- Chevalier R, Hoogenboom G, McClendon R, Paz J (2012) A web-based fuzzy expert system for frost warnings in horticultural crops. *Environ Model Softw* 35:84–91
- Cooper WC, Young RH, Turrell FM (1964) Microclimate and physiology of citrus: their relation to cold protection. *Agric Sci Rev* 2(1):38–50
- Fauci A (2008) *Harrison's principles of internal medicine*. McGraw-Hill, New York
- Fonsah EG, Taylor KC, Funderburk F (2007) Enterprise cost analysis for middle. Technical report, AGECON-06-118
- Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol Model* 160:249–264
- Grundstein A, Ramseyer C, Zhao F, Pesses J, Akers P, Qureshi A et al (2012) A retrospective analysis of American football hyperthermia deaths in the United States. *Int J Biometeorol* 56:11–20
- Haykin S (1999) *Neural networks: a comprehensive foundation*, 2nd edn. Prentice Hall, Upper Saddle River
- Hochmuth GJ, Kostewicz S, Martin F (1993) Irrigation method and Rowcover use for strawberry freeze protection. *J Am Soc Hortic Sci* 118:575–579
- Hoogenboom G (1993) The Georgia automated environmental monitoring network. In: Hatcher K (ed) *Proceedings of the 1993 Georgia water resources conference*, The University of Georgia, Athens, pp 398–402
- Hoogenboom G (2000) The Georgia automated environmental monitoring network. In: *Preprints of the 24th conference on agricultural and forest meteorology*, pp 24–25. American Meteorological Society, Boston
- Hoogenboom G (2005) The Georgia automated environmental monitoring network: experiences with the development of a state-wide automated weather station network. In: *Proceedings of the 13th symposium on meteorological observations and instrumentation & 15th conference on applied climatology*. American Society of Meteorology, Boston
- Hoogenboom G, Coker DD, Edenfield JM, Evans DM, Fang C (2003) The Georgia automated environmental monitoring network: ten years of weather information for water resources management. In: *Proceedings of the 2003 Georgia water resources conference*
- Igel C, Hüsken M (2003) Empirical evaluation of the improved RPROP learning algorithms. *Neurocomputing* 50:105–123
- Jain A (2003) Frost prediction using artificial neural networks: a temperature prediction approach. M.S. thesis
- Jain A, McClendon RW, Hoogenboom G, Ramyaa R (2003) Prediction of frost for fruit protection using artificial neural networks. *Am Soc Agric Eng* 03:3075
- Jones GM, Stallings CC (1999) Reducing heat stress for dairy cattle. Virginia Cooperative Extension, Blacksburg
- Martolf JD, Gerber JF, Chen EY, Jackson HL, Rose AJ (1984) What do satellite and other data suggest about past and future Florida freezes. In: *Proceedings of the Florida state horticultural society*, pp 17–21
- Narendra KS, Mukhopadhyay S (1994) Adaptive control of nonlinear multivariable systems using neural networks. *Decis Control* 7(5):737–752
- Perry KB (1994) Frost/freeze protection for horticultural crops. North Carolina cooperative extension service. Leaflet No: 705-A
- Perry KB (1998) *Basics of frost and freeze protection for horticultural crops*, 8th edn. HortTechnology, Alexandria
- Riedmiller M, Braun H (1993) A direct adaptive method for faster backpropagation learning: the RPROP algorithm. In: *IEEE international conference on neural networks*, pp 586–591
- Robinson PJ (2000) Temporal trends in United States dew point temperatures. *Int J Climatol* 20(9):985–1002
- Rodrigo J (2000) Spring frosts in deciduous fruit trees—morphological damage and flower hardiness. *Sci Hortic* 85:155–173
- Sandstrom MA, Lauritsen RG, Changnon D (2004) A Central-US summer extreme dew-point climatology (1949–2000). *Phys Geogr* 25(3):191–207
- Shank DB, McClendon RW, Paz JA (2008) Ensemble artificial neural networks for prediction of dew point temperature. *Appl Artif Intell* 22(6):523–542
- Shank D, Hoogenboom G, McClendon R (2008) Dewpoint temperature prediction using artificial neural networks. *J Appl Meteorol Climatol* 47:1757
- Smith BA, Hoogenboom G, McClendon RW (2009) Artificial neural networks for automated year-round temperature prediction. *Comput Electron Agric* 68:52–61
- Smith BA, McClendon RW, Hoogenboom G (2008) Improving air temperature prediction with artificial neural networks. *Int J Comput Intell* 3(3):179–186
- Ward System Group (1993) *Manual of NeuroShell 2*. Ward System Group, Frederick
- Warmund MR, Guinan P, Fernandez G (2008) Temperatures and cold damage to small fruit crops across the Eastern United States associated with the April 2007 freeze. *HortScience* 43(6):1643–1647
- White GF, Haas HE (1975) *Assessment of research on natural hazards*. MIT Press, Cambridge