

Visualization and Quantitative Assessment to Verify Precipitation Forecasts Associated with Mid-Latitude Cyclones across the Eastern United States

Williams, J., Hung, M. C. and Wu, Y. H.

Department of Humanities and Social Science, Northwest Missouri State University, Valk Center 113, 800 University Drive, Maryville, MO 64468, USA

E-mail: jawps23@gmail.com, mhung@nwmissouri.edu, ywu@nwmissouri.edu

DOI: <https://doi.org/10.52939/ijg.v17i3.1889>

Abstract

Analysis was conducted to verify forecast against observation precipitation associated with mid-latitude cyclones over the Eastern US in winter and spring 2013 using Geographic Information Systems (GIS). The forecast data are day two 24-hour Quantitative Precipitation Forecasts (QPF) produced by the Global Forecast System (GFS) model. The analysis methods produced categorical geographic error maps of hits, misses and false alarms in spatial relation to the mid-latitude cyclones and traditional verification scores for each day. A hypothesis test was also performed to determine if the GFS mean forecast precipitation over the study area is significantly different from the mean observed precipitation during mid-latitude cyclones. The spatial verification maps, as an analytical and visualization tool, provided evidences on geographical relationship between correct predictions (hits and correct negatives) and incorrect predictions (misses and false alarms). Working together with quantitative scores and hypothesis test, spatial verification maps reveal that the GFS model has a tendency to over forecast precipitation coverage associated with mid-latitude cyclones over the Eastern US and often moves the mid-latitude cyclones too fast.

1. Introduction

Weather forecasts affect our lives in many ways, from global scale (such as crop productivity and food security), to regional scale (such as extreme weather warming and flight schedule), to local scale (such as transportation options and sport activities). Forecasts were made by National Weather Service (NWS) using various models and broadcast to the public through various media. Verification of the Numerical Weather Prediction (NWP) models is a key factor in the forecast process. Forecasters use the verification information in their forecast decision process (Ebert and McBride, 2000). Model verification statistics are used in determining whether model equations need to be updated to improve the model performance (Feldmann et al., 2015, Mohanty et al., 2018 and Rehman et al., 2018). This decision is highly dependent on the Quantitative Precipitation Forecast (QPF) skills from the model (Hamill, 1999).

However, traditional verification techniques lack answers to many questions about the spatial relationship of model errors to geographic areas and weather patterns (Gilleland et al., 2010). The results of current verification methods are mainly quantitative, providing skill numbers (Goddard et al., 2003) to assess the quality of a NWP model with the forecast and observation grids sometimes

displayed side by side (Zhu, 2007, Ebert and McBride, 2000 and Strassberg and Settelmaier, 2010). Over the last decade, many new verification techniques have been developed to not only take into account new observational data used in verification, but also to quantify model performance in regards to spatial errors (Wernli et al., 2009, Gilleland et al., 2010 and Casati et al., 2008).

Geographic Information Systems (GIS) are used to process and display spatial information to provide a greater understanding of data and assist in decision making processes. Using GIS to verify NWP models and display model error spatially can provide a greater understanding of model error as it relates to weather systems (Strassberg and Settelmaier, 2010). GIS is becoming increasingly applicable to meteorological problems as data interoperability improves and spatial data from various sources are made readily available (Stellman et al., 2009, Wilhelmi and Betancourt, 2005, Yuan, 2005 and Li et al., 2015). Maps were produced by overlapping ground truth on interpolated surface. By closely examining these maps, they reported valuable information that was not available from traditional quantitative assessment.

2. Traditional Verification of Weather Forecast

Ebert and McBride (2000) proposed an object-oriented approach in the verification of NWP QPFs. The objects of verification were defined as contiguous rain areas (CRA's), and they reported three types of errors (displacement errors, pattern errors, and rain volume errors) for QPFs. The verification was performed on 24-hour rainfall forecast totals from an NWP Model in Australia from 1995-1999. The results showed 50% of the errors were in a displacement of the rain area, 45% in the spatial weather pattern, and only 5% in the actual volume of rain. Casati et al., (2004) presented a new technique to determine forecast skill as a function of the intensity and size of precipitation areas, based on the theory that precipitation are largely impacted by the presence of weather patterns. Three-hour precipitation forecast totals with 3-hour lead times were evaluated over England, and the results showed a loss of forecasting skill was mainly due to small horizontal displacements (< 40 km) of intense precipitation events (> 4 mm/h).

Feidas et al., (2007) proposed a GIS model, Precipitation Forecasts Evaluator (PFE), to perform qualitative and quantitative verification of precipitation forecasts from the NWP Bologna Limited Area Model (BOLAM) in Greece. The PFE was run for a 2-day historical heavy rain event that occurred in 1999, and provided a difference grid between the forecast and observation 3-hour total precipitation amounts as well as a final contingency map revealing hits, misses, and false alarms spatially. Settlemaier (2009) utilized GIS to display Temperature NWS forecast low temperature error compared to high resolution temperature observation analysis spatially across a 5 x 5 km grid. Strassberg and Settlemaier (2010) proposed a GIS model to create difference grids between the low temperatures observed and both the NWP model and NWS generated forecasts. The 5km Continental United States (CONUS) forecast grids from the National Digital Forecast Database (NDFD) and the observation grids are downloaded to create such difference grids.

3. Research Objectives

Research showed much of the error associated with QPFs is based on spatial displacements of rain areas. GIS can provide a fast and efficient way to quantify the errors on large datasets as well as visualize and display them in a way that forecasters can better understand the errors in relation to weather patterns and improve upon them. Therefore, this research was conducted to accomplish the following objectives:

- (1) Produce categorical spatial error maps of hits, misses and false alarms alongside traditional skill numbers of the day two GFS 24-hour QPF associated with mid-latitude cyclones.
- (2) Determine if the day two GFS 24-hour QPF means differ significantly from the observed 24-hour precipitation totals in association with mid-latitude cyclones over the Eastern US.
- (3) Research the feasibility of using an adaptable automated GIS model for the production of near real-time spatial error maps on NWP forecast grids in daily operations.

4. Data and Methods

4.1 Data Collection

Mid-latitude cyclones can impact large areas of the U.S. with significant flooding and other hazardous weather conditions as the clash of drier cooler air moving in behind the system mixes with warm moist air ahead of it. (Ahrens, 2000). Mid-latitude cyclones were selected as the weather pattern in this research and the Eastern U.S. as the study area. The time frame of the study data is over the winter and spring (January to May) of 2013, which included 32 suitable forecast days. The day two 24 hour QPF was selected for verification in this study.

The forecast data used were the 12 Coordinated Universal Time (UTC) GFS model runs, downloaded from the University Cooperation for Atmospheric Research (UCAR) Data Server, Thematic Real-time Environmental Distributed Data Server (THREDDS) (<https://www.ncei.noaa.gov/data/global-forecast-system/access/historical/analysis/>). The 3-hour QPF (mm) product was selected on the server for the GFS forecast through 48 hours over a set bounding box and downloaded as a NetCDF (network Common Data Form) file for each forecast day verified. The data is in 0.5 degree resolution in the World Geodetic System of 1984 (WGS84) datum. The observation data used were the Daily Precipitation Analysis Grids interpolated from over 30,000 quality controlled (QC) reports and rain gauges across the globe from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) Precipitation Working Group (https://ftp.cpc.ncep.noaa.gov/GIS/GRADS_-GIS/GeoTIFF/GLB_DLY_PREC/). An interpolation method using KF-based CMORPH (Kalman filter-based Climate Prediction Center (CPC) morphing) technique is used with consideration of elevation to provide a continuous gridded surface of observed precipitation from the stations and rain gauges (Xie et al., 2010 and Joyce and Xie, 2011). The data is in GeoTiff format with 0.5 x 0.5 degree global resolution. The grids are observed 24 hour precipitation accumulation

amounts in mm/hr from 12 UTC the day prior to 12 UTC on the valid day.

4.2 Data Preparation

A study area boundary was created to extract states from East of the Colorado Rockies to the Mid Atlantic States. The boundary of the states extracted was used as the raster analysis mask in ArcMap's geoprocessing environment settings. Re-projection was done to ensure consistency between forecast and observation data.

The 32 GFS forecasts were prepared for verification next. Since each QPF valid time from the GFS is from 3 hours prior, eight QPF forecasts were extracted from 24 to 48 hours from the model 12 UTC initialization time on each day to represent the day two 24 hour QPF. The valid times for each 3 hour QPF extracted were 15 UTC, 18 UTC, 21 UTC, 00 UTC, 03 UTC, 06 UTC, 09 UTC, and 12 UTC for each NetCDF file. After the extraction of the 8 files for each forecast date, the raster layers were exported to a grid file in order to run geospatial processes on the datasets. These eight 3-hour forecast grids were then added together using the addition property for raster grids and masked by the study area boundary in the geoprocessing environment. The observation dataset preparation was minimal since it was in the same projection and resolution of the forecast data, and already in a GIS ready raster format.

4.3 Error Maps for Visualization and Verification

Figure 1 shows major procedures in creating maps. The first step in producing error maps was to reclassify the forecast and observation grids into rain and no rain categories. For the reclassification, any cell with more than 0.254 millimeters of total forecast or observed precipitation was placed into the rain category, anything less was placed into the no rain category, since a trace of rainfall is considered anything less than 0.01 inches or 0.254 millimeters (Mekis, 2005). During the reclassification, each cell in the grid would be assigned a value of 1 (Rain), or 0 (No Rain).

The second step was to combine the resultant forecast and observed Rain/No Rain, or binary (1/0) grids to create the error grid for each verification day. The grids were compared cell by cell for forecast and observed rain to deduce hits, misses, false alarms, or correct negatives. Possible combinations were [1,1] representing hits, [1,0] representing false alarms, [0,1] representing misses, and [0,0] representing correct negatives, where the first number represents the forecast and the second number represents the observation. These categories are needed to distinguish whether the model has a

tendency to under forecast or over forecast precipitation. The resultant error grid will contain four different values of 0 - 3 for these categories. The third step was to finalize error maps in a way which enables the reader to determine the significance of the model hits or misses in relation to the weather system by displaying hits, misses and false alarms greater than or less than one inch. Both the observed and forecast 24 hour total precipitation layers were reclassified into 3 categories: < Trace, Trace to 1 inch, and > 1 inch. This final map not only gives the user an idea of where the low pressure system was and where the heaviest precipitation fell, but also where the model errors are.

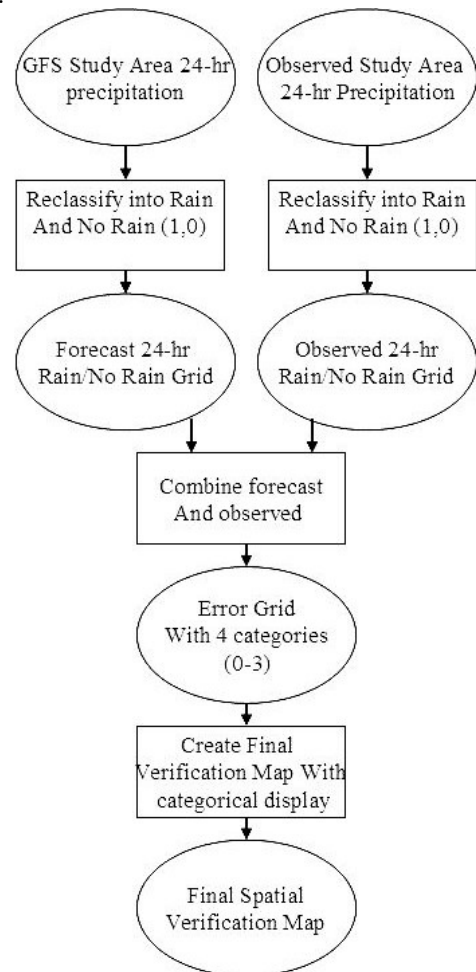


Figure 1: Major procedures in creating spatial verification maps

4.4 Traditional Descriptive QPF Scores

While these error maps provide an excellent understanding of model error, statistical quantitative verification is necessary to get an objective view on forecast performance.

Table 1: Contingency Analysis Table used to derive Accuracy, BIAS, ETS and FAR

Forecast	Observed			Total
	Rain	No rain	Total	
Rain	Hits	False alarms	Forecast yes	
No rain	Misses	Correct negatives	Forecast no	
Total	Observed yes	Observed no	Total	

The most common and preferred way is with a two by two contingency table (as seen in Table 1). The two by two contingency tables produce four outcomes: hits, false alarms, misses and correct negatives (Zhu, 2007, JWGFVR, 2012 and Hamill, 1999). From the contingency tables four verification skills were derived, shown in equations 1 through 4 below. Accuracy (equation 1) is the fraction of correct forecasts, but does not take into account hits from random chance. BIAS (equation 2) is the model tendency to under-forecast (BIAS<1) or over-forecast (BIAS>1) precipitation. Equitable Threat Score (ETS) (equation 3) is the fraction of forecast events that were correctly predicted taking into account hits from random chance. False Alarm Rate (FAR) (equation 4) is the ratio of false alarms to hits plus false alarms (JWGFVR, 2012).

$$\text{Accuracy} = \frac{\text{hits} + \text{correct negatives}}{\text{total}} \quad \text{Equation 1}$$

$$\text{BIAS} = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \quad \text{Equation 2}$$

$$\text{ETS} = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}} \quad \text{Equation 3}$$

where:

$$\text{hits}_{\text{random}} = \frac{(\text{hits} + \text{misses})(\text{hits} + \text{false alarms})}{\text{total}}$$

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad \text{Equation 4}$$

4.5 Hypothesis Testing

Verification of precipitation forecasts using formal hypothesis testing to infer whether NWP model forecasts are significantly different from each other or from the observations are rarely performed (Hamill, 1999). In this research, the two sample difference of means test was used to infer whether the GFS forecast precipitation means are significantly different from the observed

precipitation means. The difference of means test was performed as a non-directional test with the null hypothesis stating the mean of the forecast precipitation totals is not significantly different from the mean of the observed precipitation totals during mid-latitude cyclones. The alternative hypothesis is that they are significantly different. The test used the normal (Z) distribution since there are at least 30 samples for both and the variances are known, and the confidence level will be set at 95%. The two sample variances are assumed unequal, due to the significantly larger variance in the forecast sample compared with the observation sample, therefore the Separate Variance Estimate (SVE) is used in the Z test (McGrew and Monroe, 2000). In order to calculate the test statistic, the sample mean and variance was calculated for both the forecast and observed 32 samples. Using the confidence level of 0.95, the Z critical for this two-tailed test with 32 degrees of freedom is 1.96 (McGrew and Monroe, 2000), therefore, the test statistic needs to be less than 1.96 or greater than -1.96 for the null hypothesis to be accepted. If it does not fall within those thresholds the null hypothesis is rejected and the alternative hypothesis is accepted with 95% confidence.

5. Spatial Verification Maps

The spatial verification maps produced gives meteorologists an indication of forecast performance associated with mid-latitude cyclones not only for each sample verification day but also an overview for the entire study period. An understanding of model forecast errors in spatial relation to weather systems will give forecasters knowledge of where to improve upon the models during these systems. An example spatial verification map consists of three maps, a forecast map, an observation map, and a spatial error map, as shown in Figure 2. The spatial verification map visualizes the location of incorrect predictions (misses and false alarms) in relation to correct predictions (hits and correct negatives), such spatial relationship is not available from traditional verification scores, yet it is an essential piece of information in weather forecast, as well as suggestions for improving existing forecast models.

For example in Figure 2, one can observe a trend of misses concentrating on the west side (back side) of the storm and false alarms on the east side (front side), implicating that the model prediction moves the storm too fast. Traditional verification scores can only report the ratio between correct predictions and incorrect predictions, they do not show the locations where these incorrect predictions are. Attention should be directed to the overall look of these spatial verification maps. The pixelated look is because of the cell or grid size of data, not of the image quality. Observation data was interpolated to 0.5 x 0.5 degree grid, and forecast data follows the grid size for consistency. Figure 3 displays the 6 selected spatial verification maps during the 2013 winter and spring season of mid-latitude cyclones. Looking at the dates and verification maps presented in this research, from Jan 13th to Jan 18th

meteorologists can deduce the precipitation was associated with a mid-latitude cyclone slowly moving east then stalling out over the southeast US, then moving eastward again. Analyzing the forecast error, the false alarm areas occur along the entire leading edge of the area of precipitation, with the forecast misses, though at a lesser extent, occurring along the back edge of the precipitation. This tells the meteorologist that the model is moving the mid-latitude cyclone too fast. If these verification maps were available in near-real time operationally, the forecaster could use this information to adjust their forecast grids, and move not only the precipitation associated with the mid-latitude cyclone, but also temperature, humidity and wind changes all slightly backward to the extent the model was off for the next day when producing their forecast grids.

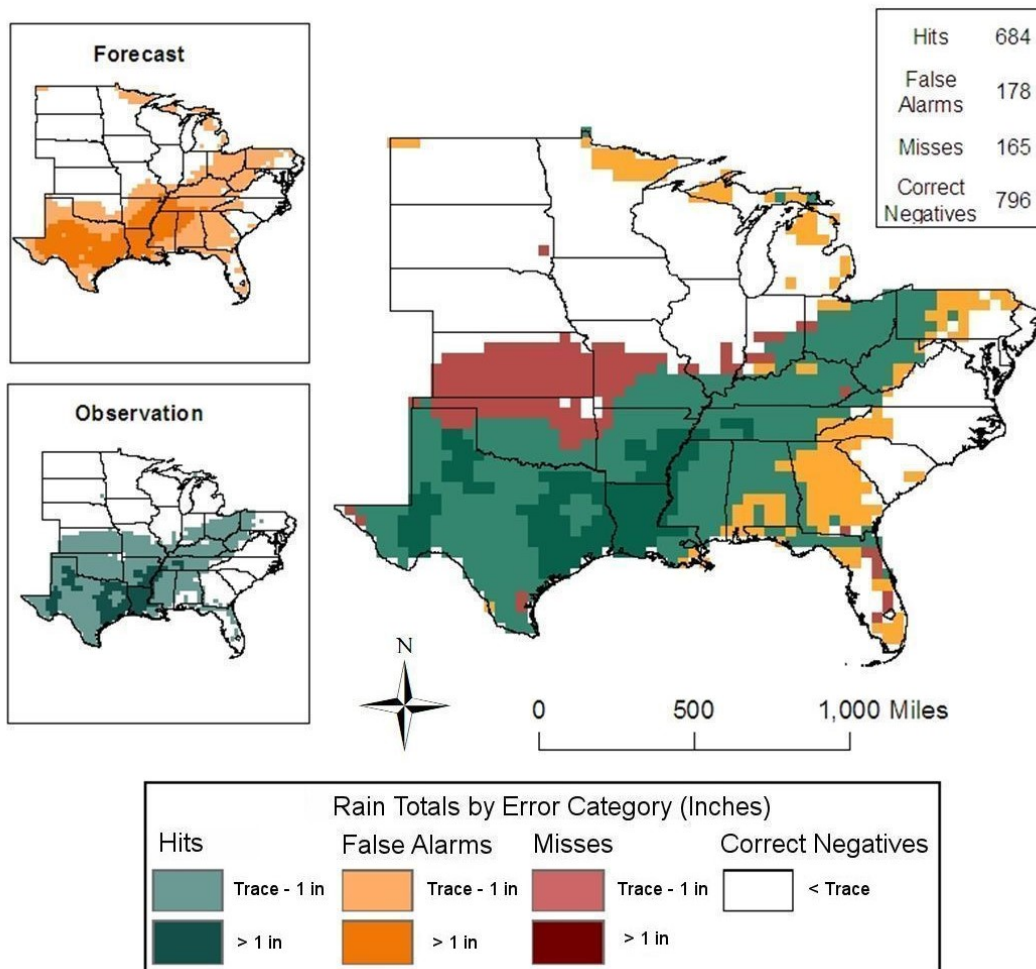


Figure 2: Spatial verification map for Jan 10, 2013

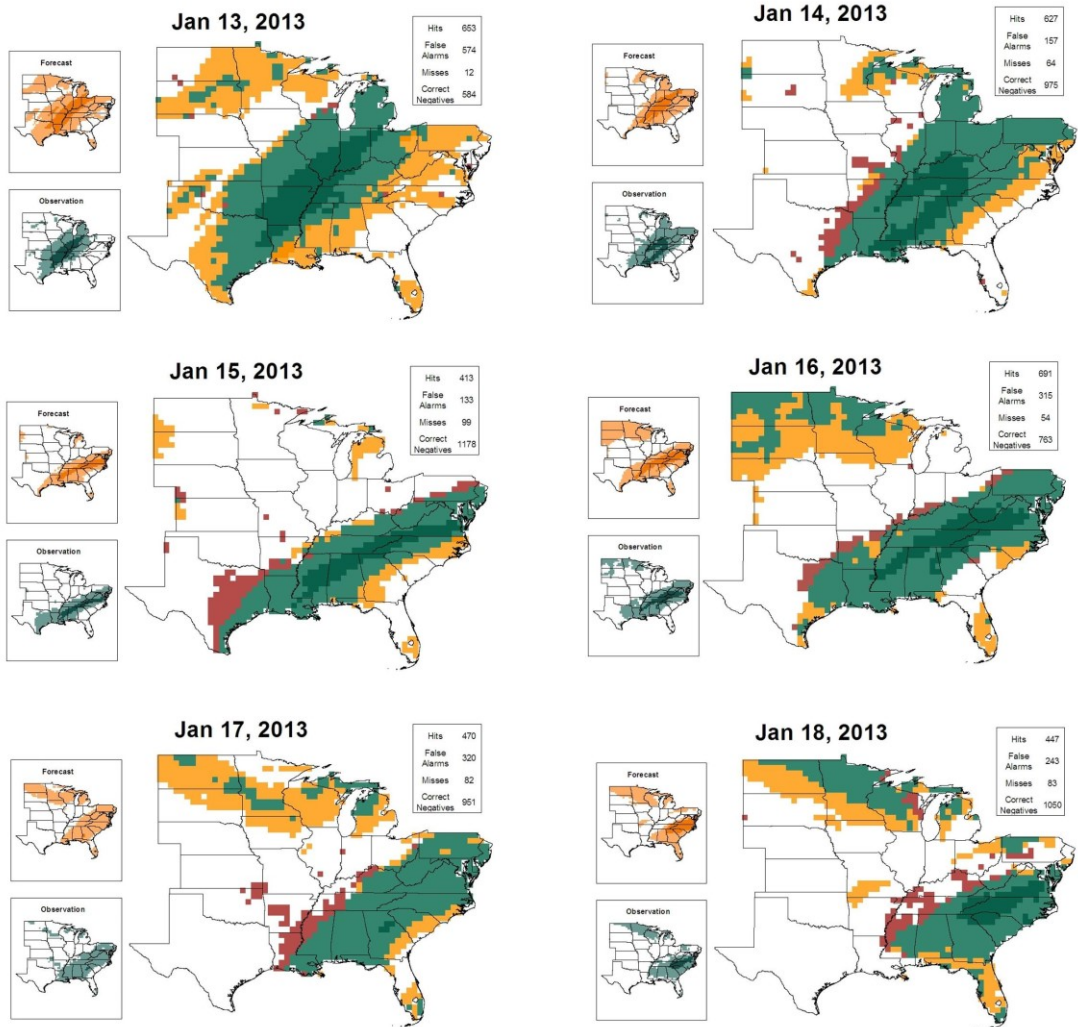


Figure 3: 6 selected spatial verification maps during the 2013 winter and spring season

The information deduced from these spatial maps over the 32 sample days could also be used to infer behavior of the GFS during mid-latitude cyclones over the Eastern US. Looking at these sample verification maps in Figure 3, there are many more false alarms than misses for precipitation associated with these systems. The GFS forecasts a larger spatial coverage of precipitation for every one of these sample days than actually occurred. Such pattern can be observed from other maps that were not shown in Figure 3. There is a tendency that the spatial coverage of precipitation forecasted is much greater than the spatial coverage where precipitation actually fell, with only a few misses. Based on the spatial verification maps from this research, one can infer that the GFS tends to over forecast precipitation coverage associated with mid-latitude cyclones over the Eastern US, and often moves cold fronts too fast. A closer look at Figure 3, one may

observe some spatial pattern between intensity and inaccurate forecasts. Taking Jan. 13, 2013 in Figure 3 as example, there is a linear heavy precipitation (dark green area) in the center of the storm. This dark green indicates areas with more than 1 inch of precipitation. Light green (the entire storm) is areas with observed precipitation. By looking at Jan. 13 through Jan. 18, it is clear the entire storm is moving eastward, and on the back of the storm, there is another storm (closer to Canada border), not so well organized and separated from the said storm. If one sees the heavy precipitation area as an axis of the storm, one can observe that the storm is wider on Jan. 13 and Jan. 14, and narrower on Jan. 15 through Jan. 18. Interestingly, on Jan. 13 and Jan. 14 (when the storm is wider), there are more orange (false alarms) than red (misses), counting only cases associated with the said storm, not the separated storm closer to Canada border.

Table 2: Storm width and cases of miss and false alarm

Date	Storm width (km)	Case of miss	Case of false alarm
Jan. 13	822	8	312
Jan. 14	852	52	89
Jan. 15	473	84	70
Jan. 16	706	52	52
Jan. 17	667	59	55

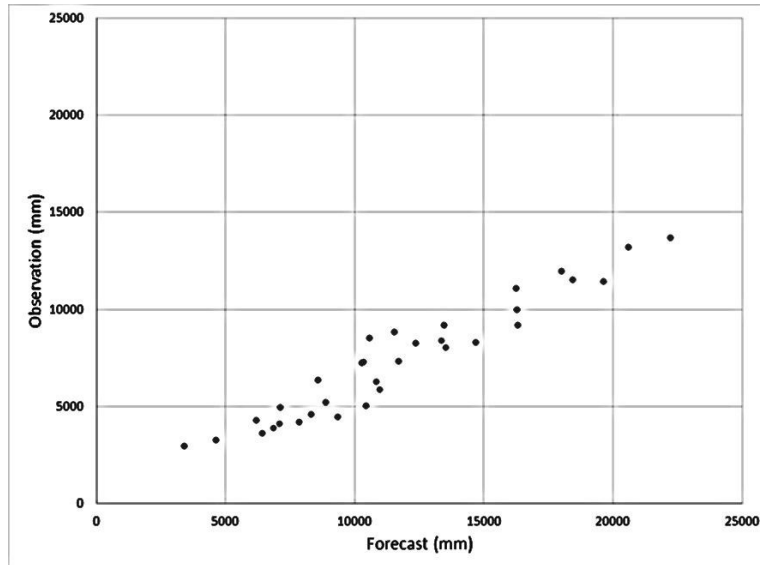


Figure 4: Scatter plot between precipitation data from forecast and observation

However, on Jan. 15 through Jan. 18 (when the storm is narrower), there are more (or equal) red (misses) than orange (false alarms). Such spatial pattern is summarized in Table 2. The storm width on Jan. 13 and Jan. 14 is 822 km and 852 km, respectively, but only 473 km, 706 km, and 667 km for Jan. 15, Jan. 16, and Jan. 17, respectively. On Jan. 13 and Jan. 14, there are more cases of false alarm than miss, but on Jan. 15 through Jan. 17, there are more cases of miss than false alarm. Such spatial pattern or information is not available from traditional verification scores, explained in the next section. In Table 2, there are only five days, Jan. 13 through Jan. 17. The storm on Jan. 18 is moving out to Atlantic Ocean. The front side (east side) of the storm is missing out because of study area boundary. By patterns observed on other dates, front side of the storm is where false alarms are supposed to occur. Because front side is missing on Jan. 18 and to avoid biased observation, Jan. 18 is not summarized in Table 2.

6. Traditional Verification Scores

The objective GFS forecast skill numbers shown in Table 3 can be used along side the spatial verification maps for each day, if the spatial verification maps were generated in real time

operationally. This combines quantitative objective verification with spatial subjective verification, giving the forecaster more information on how much and where the model is off, and model tendencies in association with weather patterns and geographic location. Based on the descriptive results in Table 3 and what each forecast skill number means, the GFS model was anywhere from 62% to 88% accurate in correctly predicting rain or no rain for each of the 32 sample days. For every single day, the model was biased towards over forecasting precipitation coverage with the false alarm percentage ranging from 19% to 55%. The ETS ranged from 0.19 to 0.60, likely due to the events displaying a wide range in hits due to random chance because some of the sample days had larger areas of precipitation coverage than others. Over the entire 32 sample days associated with mid-latitude cyclones over the Eastern US, the model was 76% accurate, with the inaccuracies biased towards over forecasting precipitation occurrence. The False Alarm Rate was 35% overall, and the ETS was 0.35, which can be compared with the ETS of other models associated with mid-latitude cyclones. Additionally Figure 4 shows the scatter plot between forecast and observation data.

Table 3: QPF Verification Scores and Total Precipitation Amounts for each sample day

DATE	Accuracy	BIAS	ETS	FAR	Forecast (mm)	Observation (mm)
10-Jan	0.81	1.02	0.45	0.21	20604	13193
11-Jan	0.82	1.20	0.41	0.20	16330	9193
12-Jan	0.81	1.28	0.43	0.25	7089	4105
13-Jan	0.68	1.85	0.26	0.47	18453	11523
14-Jan	0.88	1.13	0.60	0.20	11525	8826
15-Jan	0.87	1.07	0.53	0.24	8570	6381
16-Jan	0.8	1.35	0.43	0.31	12355	8263
17-Jan	0.78	1.43	0.36	0.41	3407	2976
18-Jan	0.82	1.30	0.43	0.35	9357	4462
20-Jan	0.77	1.43	0.39	0.34	10564	8533
24-Apr	0.65	1.82	0.22	0.48	8894	5213
25-Apr	0.71	1.52	0.28	0.38	7841	4209
27-Apr	0.72	1.90	0.26	0.55	6201	4309
28-Apr	0.8	1.49	0.44	0.35	10281	7230
29-Apr	0.77	1.37	0.37	0.32	13458	9207
3-May	0.74	1.55	0.33	0.37	13526	8034
4-May	0.88	1.18	0.61	0.19	16276	9970
5-May	0.83	1.28	0.50	0.27	10345	7275
6-May	0.76	1.76	0.34	0.47	7111	4963
7-May	0.73	1.81	0.31	0.48	4614	3260
8-May	0.62	1.96	0.18	0.55	6856	3870
9-May	0.68	1.60	0.24	0.42	10444	5020
10-May	0.78	1.26	0.35	0.25	16243	11077
11-May	0.67	1.49	0.19	0.38	14702	8307
16-May	0.67	1.85	0.21	0.55	6425	3642
17-May	0.72	1.41	0.30	0.38	8306	4617
18-May	0.71	1.48	0.28	0.39	10845	6275
19-May	0.76	1.20	0.33	0.26	13346	8415
20-May	0.76	1.35	0.31	0.29	22219	13674
21-May	0.72	1.40	0.23	0.32	18006	11977
22-May	0.73	1.30	0.22	0.28	19635	11432
23-May	0.8	1.27	0.43	0.27	10955	5897
Average	0.76	1.45	0.35	0.35	11712	7354

The correlation coefficient is 0.961953, which implies a positive and strong relationship between the forecast data and observation data. As one may notice from figure 4, forecasted precipitation is higher than observed precipitation, implies over forecast. This information supports what was found in the overall spatial verification maps, and gives forecasters additional support in making decisions about forecasting the amount or at least spatial coverage of precipitation associated with mid-latitude cyclones.

7. Hypothesis Test Results

Lastly, a hypothesis test was performed on the precipitation amounts for the study area on both the forecast and observation. Table 3 also shows the total amount of precipitation in millimeters forecast and observed over the entire study area for each sample day. The mean total precipitation amount is higher for the forecasts, which supports the results of the spatial maps and descriptive statistical scores showing the GFS with a tendency to over-forecast precipitation.

To perform the two sample difference of means hypothesis test, the two samples need to follow a normal distribution curve. To ensure this, a graph was created overlaying the observed and forecast precipitation histograms in Figure 5. The Z test for two sample difference of means was performed, and the Z value was 4.32, which is greater than the Z critical of 1.96. Based on the calculated z test statistic being larger than Z-critical, the null

hypothesis stating that the forecast precipitation mean is not significantly different from the observed precipitation mean associated with mid-latitude cyclones over the Eastern US, is rejected with 95% confidence. The Alternative hypothesis is accepted with 95% confidence, which states there is a significant difference between the forecast and observed precipitation means associated with mid-latitude cyclones over the Eastern US. These results stress the importance of model verification in the forecast process and human forecasts. They also show there is still much room for improvement with the GFS forecast model during mid-latitude cyclones over the Eastern U.S. The more information forecasters can gain from model verification results, the more likely they are to improve upon the model errors in making their forecasts. Using model verification maps and statistics gives meteorologists a spatial understanding of model errors to greatly improve upon, reducing forecast errors and thereby reducing the impacts that adverse weather has on the community.

8. Developing the Automated Verification Model

In order to save time and minimize human error, as well as produce methods that could potentially be run automatically and repetitively for daily forecast operations, two GIS models were created in ArcMaps model builder to semi-automatically run through the majority of the analysis.

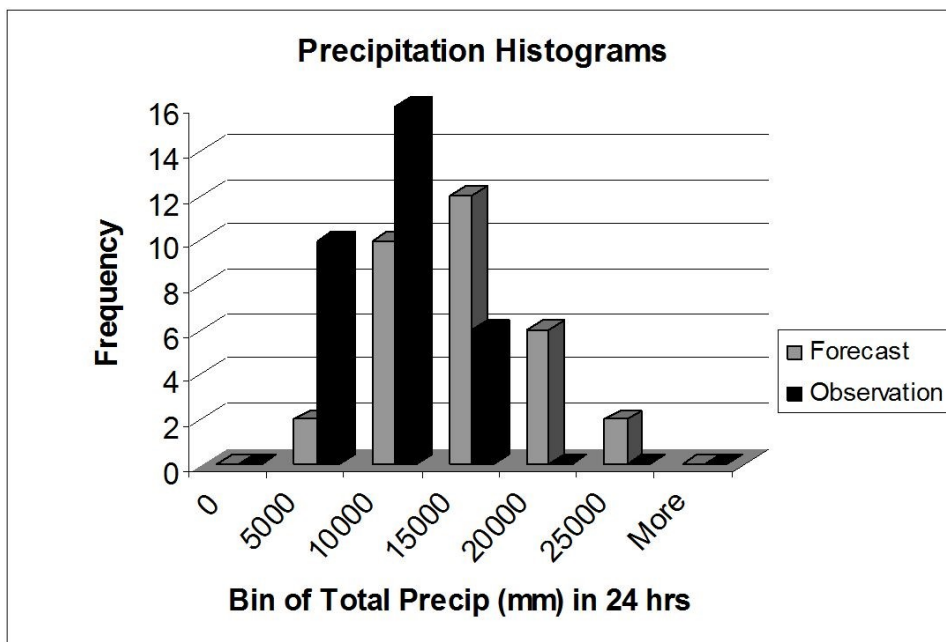


Figure 5: Forecast and observation total precipitation histograms for the 32 verification samples

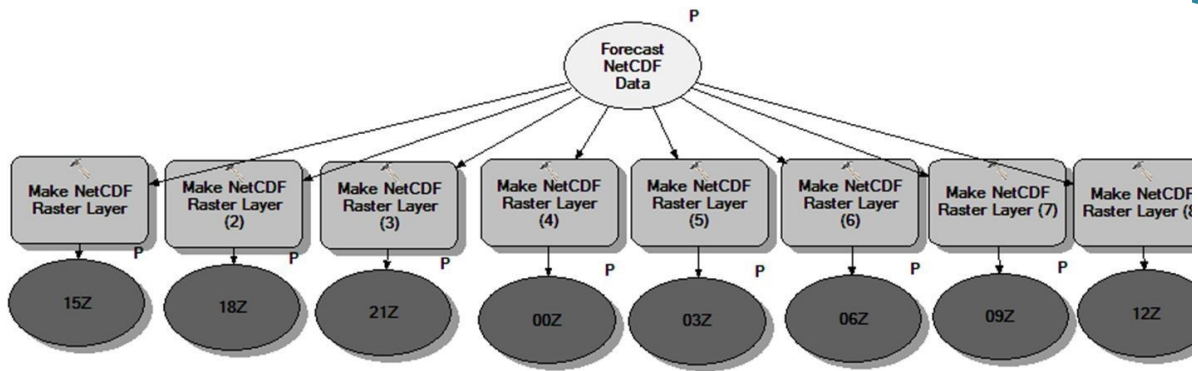


Figure 6: GIS model to extract 3-hour forecast grids from NetCDF dataset

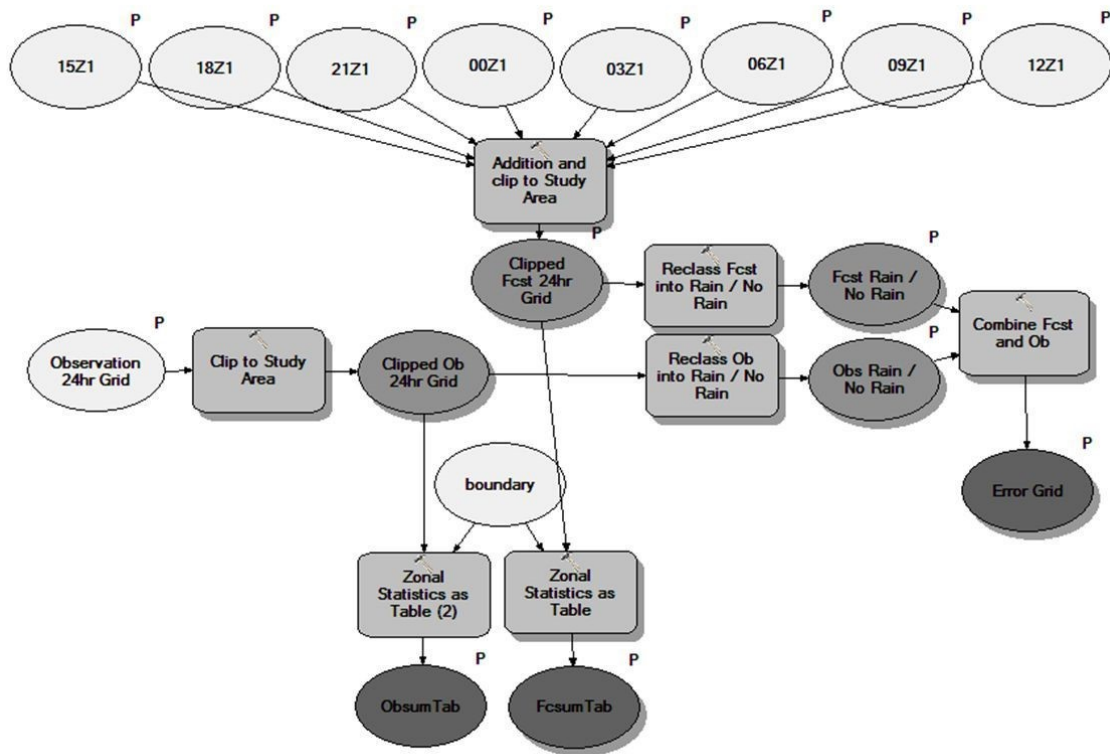


Figure 7: GIS model to create error grid and total precipitation statistics over the study area

The first model, as shown in Figure 6, extracts 8 3-hour raster layers from a NetCDF forecast dataset. However, with each new sample day the forecast dataset dates changed, so before running the GIS model with each new sample day the date/time of each of the 8 extraction layers were edited in the Make NetCDF Raster Layer process, and then the model was run. Once the first model is run, the 8 raster layers were manually exported from layers to grids to enable geoprocessing on them.

The second model, as shown in Figure 7, runs through the rest of the GIS processes. First, it finishes the data preparation steps to create a 24

hour forecast and observation total precipitation grid. Then it uses these grids to create the forecast and observed Rain/No Rain grids, which are combined to generate the final error grid. It also uses the 24 hour total precipitation grids in the inferential statistics process to create a sum of precipitation over the entire study area for the forecast period. For the organization of numerous verification days, a separate processing folder and ArcMap document was created for the models to be run in for each set of data. After a set of data was processed through the models the resultant datasets were moved into a folder specifying the date of the

verification sample. From our tests, once all data are downloaded to local computers, creating spatial verification maps as well as verification scores from these two GIS models took less than 5 minutes with a moderate desktop computer.

9. Discussion

Weather forecast is like spatial interpolation of soil properties in ways that they both base on current data; they both have established models; and they both generate predictions. Traditional quantitative assessment and hypothesis test can assess performance or accuracy of predictive models in general. However, detailed assessment on model performance requires further evaluations or examinations. Previous studies on spatial interpolation of soil properties (Wu et al., 2013 and Wu and Hung, 2016) utilized GIS as visualization and analytical tools, in addition to traditional quantitative assessment, and reported information that is not available from traditional quantitative assessment. Such information showed model performance by location. Therefore, users can observe where models performed better or worse than expectations. As demonstrated in this case study, spatial verification maps showed geographic relationship between incorrect predictions and correct predictions. They showed concentration of misses on the back side of the storm, while concentration of false alarms on the front side of the storm. It is an indication that the model moved the storm too fast. This information can assist not only daily weather forecast (temperature or precipitation), but also model improvement in further studies.

This research targeted specifically the verification of precipitation forecasts from the GFS model for the day two forecasts from each model run selected. However, the methodology in this research can be applied to other forecast parameters, other forecast times and time periods (3-hour, 6-hour, 24-hour, etc), other NWP models during any weather pattern and study area by changing the inputs to the model. It is possible to develop a fully automated model to produce spatial error maps in near-real time operationally with more programming in ArcMap. Scripts could be written to fill in the gaps in the model presented in this research to automatically download the forecast and observation data sets from a server or the NDFD, to export the NetCDF data into grid format, then finally to post the spatial error maps onto a local intranet or to be made available in the Advanced Weather Interactive Processing System (AWIPS) for the forecasters' viewing. Since standard GIS desktop applications are being incorporated into the

NWS operations and research (Dobson et al., 2009 and Shipley, 2005), the methodology in this research could be employed by NWS offices to produce automated spatial verification maps on gridded forecasts by both NWP models and the NWS, thus increasing the likelihood of having more accurate forecasts. Moreover, with GIS being used by more and more organizations nowadays, this project with these two developed GIS models can be replicated by other organizations to verify other forecast models, study areas, and/or time durations, and draw more reliable conclusions with larger datasets.

A weather forecast model is a complex model, involving differential equations based on fluid motion and thermodynamics. Current weather observation data (such as precipitation, air pressure, temperature, etc.) are input parameters to the model. This case study pointed out an approach to visually and spatially verify forecast accuracy in a timely manner. This could be useful information to forecasters for forecast adjustments based on most recent accuracy report. A step behind forecast adjustments, these spatial verification maps may serve as references for model adjustments. Because of the complexity in weather forecast models, a systematic examination of long-term comparisons between forecast and observation is needed for model adjustments (Geer, 2016). It is suggested that such systematic examination should involve spatial components. Inaccurate forecasts (misses and false alarms) should be mapped out and projected to each input data layer (such as a temperature data layer, a precipitation data layer, an air pressure data layer, etc.). Spatial analysis should be performed at the cell level to examine the relationship between inaccurate forecasts and input data. If strong or significant relationship exists between inaccurate forecasts and a certain input parameter or a group of input parameters, it may indicate causes for inaccurate forecasts, which can be useful information for future model adjustments.

10. Conclusions

Spatial information on forecast error and tendencies can be provided by quickly processing forecast and observation data using the spatial and statistical processing tools inherent in GIS. This research investigated the use of an adaptable model in GIS for the verification of gridded forecasts over a particular study area. The model was tested with 32 verification samples of the GFS day two 24-hour QPF in association with mid-latitude cyclones over the Eastern US to reveal the tendencies of the GFS error during mid-latitude cyclones. Thirty-two geographical error maps showing hits, misses, false

alarms and correct negatives were produced alongside verification scores generated from a two-by-two contingency analysis table of rain and no rain. The verification maps show the GFS has a tendency to over forecast precipitation coverage spatially, which matched well with the verification statistics showing model biases > 1 for all 32 days as well as overall. Several verification maps show false alarms on the east side of the system and misses on the west side, revealing the model tendency to move the mid-latitude cyclones too fast. A hypothesis test was also performed to confirm that with 95% confidence the GFS mean precipitation during mid-latitude cyclones over the Eastern US is significantly different from the observation mean. Two GIS models were proposed here to facilitate making verification maps in near real-time. Having an understanding of model error spatially associated with specific weather patterns and also in near-real time during daily weather patterns can give meteorologists more forecast opportunities allowing them to produce more accurate forecasts. More accurate forecasts give a better estimate on how much precipitation will fall, allowing more warning and preparation for the adverse weather, thereby reducing the impacts the adverse weather has on the community.

References

- Ahrens, C. D., 2000, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. Pacific Grove: Brooks/Cole.
- Casati, B., Ross, G. and Stephenson, D. B. 2004, A New Intensity-Scale Approach for the Verification of Spatial Precipitation Forecasts. *Meteorological Applications*, Vol. 11(2): 141-154.
- Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocerich, M., Damrath, U., Ebert, E. E., Brown, B. G. and Mason, S., 2008, Forecast Verification: Current Status and Future Directions. *Meteorological Applications*, Vol. 15(1): 3-18.
- Dobson, J. G., Pierce, T. and Philips, M. 2009, Developing Geospatial Decision Support Tools for a Local NWS Office and Other Regional Decision Makers. *25th Conf. on International Interactive Information and Processing Systems (IIPS) for Meteor., Oceanography, and Hydrology*, Phoenix, AZ, American Meteorological Society. 7A.4 [Available online at https://ams.confex.com/ams/89annual/techprogram/paper_150608.htm]
- Ebert, E. and McBride, J., 2000, Verification of Precipitation in Weather Systems: Determination of Systematic Errors. *Journal of Hydrology*, Vol. 239, 179-202.
- Feidas, H., Kontos, T., Soulakellis, N. and Lagouvardos, K., 2007, A GIS Tool for the Evaluation of the Precipitation Forecasts of a Numerical Weather Prediction Model Using Satellite Data. *Computers and Geosciences*, Vol. 33, 989-1007.
- Feldmann, K., Scheuerer, M. and Thorarinsdottir, T. L., 2015, Spatial Postprocessing of Ensemble Forecasts for Temperature Using Nonhomogeneous Gaussian Regression. *Monthly Weather Review*, Vol.143(3), 955-971.
- Geer, A. J., 2016, Significance of Changes in Medium-Range Forecast Scores. *Tellus A: Dynamic Meteorology and Oceanography*, Vol. 68(1), 30229.
- Gilleland, E., Lindstrom, J. and Lindgren, F., 2010, Analyzing the Image Warp Forecast Verification Method on Precipitation Fields from the ICP. *Weather and Forecasting*, Vol. 25: 1249-1262.
- Goddard, L., Barnston, A. and Mason, S., 2003, Evaluation of the IRI's 'Net Assessment' Seasonal Climate Forecasts: 1997-2001. *Bulletin of the American Meteorological Society*, Vol. 88: 1761-1781.
- Hamill, T. M., 1999, Hypothesis Tests for Evaluating Numerical Precipitation Forecasts. *Weather and Forecasting*, Vol. 14: 155-167.
- Joint Working Group on Forecast Verification Research (JWGFVR), 2012, *Forecast Verification: Issues, Methods and FAQ*. [Available online at <http://www.cawcr.gov.au/projects/verification/>.]
- Joyce, R. J. and Xie, P., 2011, Kalman Filter-Based CMORPH. *Journal of Hydrometeorology*, Vol. 12(6), 1547-1563.
- Li, J., Hsu, K., AghaKouchak, A. and Sorooshian, S., 2015, An Object-Based Approach for Verification of Precipitation Estimation. *International Journal of Remote Sensing*, Vol. 36(2), 513-529.
- McGrew, J. C. and Monroe, C. B., 2000, *An introduction to Statistical Problem Solving in Geography*. The McGraw-Hill Companies, Inc, 254.
- Mekis, E., 2005, Adjustments for Trace Measurements in Canada. *15th Conference on Applied Climatology*, Savannah, GA, American Meteorological Society. J3.7 [Available online at https://ams.confex.com/ams/15AppClimate/techprogram/paper_92155.htm]

- Mohanty, U. C., Nageswararao, M. M., Sinha, P., Nair, A., Singh, A., Rai, R. K. and Rathore, L. S., 2018, Evaluation of Performance of Seasonal Precipitation Prediction at Regional Scale Over India. *Theoretical and Applied Climatology*, 1-20.
- Rehman, A., Chishtie, F., Qazi, W., Ghuffar, S. and Fatima, I., 2018, Evaluation of Three-Hourly TMPA Rainfall Products Using Telemetric Rain Gauge Observations at Lai Nullah Basin in Islamabad, Pakistan. *Remote Sensing*, Vol. 10(12), 2040.
- Settelmaier, J., 2009, Something Old, Something New, Something Blue: A Marriage of Weather and GIS." *ESRI Atmospheric Front*. Spring 2009: 8-10.
- Shipley, S. 2005. GIS Applications in Meteorology, or Adventures in a Parallel Universe." *Bulletin of the American Meteorological Society*, 86: 171-173.
- Stellman, K., Brice, T., Hansing, D., Foster, A., Pieper, C. and Lander, K., 2009, How Enterprise GIS is Improving the Effectiveness of the National Weather Service. *25th Conference on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Pheonix, AZ, American Meteorological Society. 7A.2 [Available online at https://ams.confex.com/ams/89annual/techprogram/paper_148642.htm]
- Strassberg, G. and Settelmaier, J., 2010, Verification of Gridded Morning Temperatures at NWS. *26th Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology*, Atlanta, GA, American Meteorological Society. 7A.6 [Available online at https://ams.confex.com/ams/90annual/techprogram/paper_160386.htm]
- Wernli, H., Hofmann, C. and Zimmer, M., 2009, Spatial Forecast Verification Methods Inter-Comparison Project: Application of the SAL technique. *Weather and forecasting*, Vol. 24: 1472-1484.
- Wilhelmi, O. V. and Betancourt, T. L., 2005, Evolution of NCAR's GIS Initiative: Demonstration of GIS Interoperability. *Bulletin of the American Meteorological Society*, Vol. 86, 176-178.
- Wu, Y. H. and Hung, M. C., 2016, Comparison of Spatial Interpolation Techniques Using Visualization and Quantitative Assessment. Chapter 2 in Hung, M.-C. (editor), *Applications of Spatial Statistics*. Rijeka, Croatia: InTech. 17-34.
- Wu, Y.-H., Hung, M.-C. and Patton, J., 2013, Assessment and Visualization of Spatial Interpolation of Soil pH Values in Farmland. *Precision Agriculture*, Vol. 14(6): 565-585.
- Xie, P. Chen, M. and Shi, W., 2010, CPC Unified Gauge-Based Analysis of Global Daily Precipitation. *24th Conference on Hydrology*, Atlanta, GA, American Meteorological Society. 2.3A [Available online at https://ams.confex.com/ams/90annual/techprogram/paper_163676.htm]
- Yuan, M., 2005, Beyond Mapping in GIS Applications to Environmental Analysis. *Bulletin of the American Meteorological Society*, Vol. 86, 169-170.
- Zhu, Y., 2007, Objective Evaluation of Global Precipitation Forecast. *International Symposium on Advances in Atmospheric Science and Information Technology*, Beijing, China. NOAA EMC, 3-8.