National Weather Association, *Electronic Journal of Operational Meteorology*, 2010-EJ03

Skillful Utilization of the GFS Ensemble MOS Temperature Guidance

DAVID R. EVERSOLE

National Weather Service, Mobile, Alabama

(Manuscript received 22 November 2009; in final form 9 April 2010)

ABSTRACT

Global Forecast System (GFS) model output statistic (MOS) temperatures (GMOS) and GFS ensemble MOS average temperatures (EMOS) and maxima and minima were analyzed to determine when the EMOS outperformed the GMOS. Three major groups were identified as opportunities for the EMOS to outperform: when the GMOS was equal to either the maximum or minimum of the GFS ensemble MOS temperature (Group H/L), a second in which the GMOS was within one degree (F) of the maximum or minimum of the GFS ensemble MOS (Group +1/-1), and a third which contained the remainder of the data set (Group Rest). An algorithm was developed to evaluate each of the three main groups subdivided by forecast period, month and degrees (F) per standard deviation. Group H/L identified most of the situations in which the EMOS had higher skill. EMOS higher skill tended to be found in mid to long forecast ranges, with seasons and degrees per standard deviation also having a strong influence. Overall, the algorithm produced a 15.2% improvement in Root Mean Squared temperature error over the GMOS when the EMOS was utilized in lieu of the GMOS.

1. Introduction

The Global Forecast System (GFS) ensembles and resulting Model Output Statistics (MOS) guidance products have well served National Weather Service (NWS) forecasters for some time. While the ensemble data have assisted with the diagnosis of a given weather pattern evolution, it has been somewhat difficult to ascertain exactly when and under what circumstances the ensemble MOS average temperature (EMOS) can best be used over other available forms of guidance. Past research indicates EMOS tends to

Corresponding author address: David R. Eversole, National Weather Service, 8400 Airport Blvd. Bldg. 11, Mobile, AL 36608
E-mail: david.eversole@noaa.gov

have higher skill than GFS MOS temperature (GMOS) for days five through seven. Marz (2004) examined the performance of the GMOS and EMOS and found EMOS provided a slight improvement over GMOS from days five through seven. Grumm et al. (2005) found that while the GMOS had higher skill than the EMOS through 96 hours, the EMOS and GMOS were of comparable skill after 120 hours. Personal experience has revealed that significant skill may be gleaned from EMOS usage especially when considering cases in which the GMOS is equal to or near the extremes of the ensemble MOS. Maloney et al. (2010) and MDL (2008) provide descriptions of the GFS MOS and GFS ensemble MOS, respectively.

2. Data Source

An extensive analysis of the GMOS and EMOS was performed from January 2008 through October 2009 with data collected for the southeastern United States including 118 stations (Table 1, Figure 2), yielding about one million pairs of guidance data and observed highs and lows. The EMOS and GMOS temperatures were taken from ensemble guidance summary messages which also contain statistics on the 21 ensemble members such as the maximum and minimum temperature and degrees per standard deviation (Equations 1 and 2). An example of this product is shown in Table 2, and for brevity, it just contains the temperature guidance. Hourly and synoptic observations were collected to calculate the overnight minimum temperatures from 7:00 pm to 8:00 am (LST) and the daytime maximum temperatures from 7:00 am to 7:00 pm (LST).

$$Standard\ Deviation\ (\sigma) = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{x_i}-\bar{\mathrm{x}})^2}{\mathrm{n}}} \qquad \textbf{(1)}$$

$\text{sample element} = \mathrm{x_i}, \text{sample mean} = \bar{\mathrm{x}} \text{ and sample size} = \mathrm{n}$

$$Degrees\ per\ unit\ Standard\ Deviation = \frac{|EMOSmax| - |EMOSmin|}{\sigma} \quad \textbf{(2)}$$

standard deviation = σ, highest ensemble temperature = *EMOSmax*,

and lowest ensemble temperature = *EMOSmin*

3. Analysis

From previous operational experience of known opportunities for which the EMOS has shown skill over the GMOS, the GMOS and EMOS were divided into three groups: one in which the GMOS was equal to the maxima or minima of the ensemble MOS (Group H/L), a second in which the GMOS was within one degree of the maxima and minima (Group +1/-1), and a third which included the remainder of the data (Group Rest). The skill for when GMOS was at or near the ensemble MOS maxima or minima was found to rapidly diminish from Group H/L to Group +1/-1. No tangible increase in skill was found from creating groups for when the GMOS was within two or more degrees of the ensemble MOS maxima or minima. Consequently, including all of these into Group Rest was found to produce better results.

These three groups were further subdivided again according to three variables. Two of these, forecast period and degrees per standard deviation, have been recognized as measures by which EMOS skill varies. Much can be inferred from the degrees per standard deviation, such as the general uncertainty of the weather pattern evolution and the degree of 'meridionality' of the upper level wind flow pattern. The third variable is the month of the year as skill was found to vary seasonally. These divisions resulted in about 3800 sample sets. An algorithm was developed to analyze each of these to identify those where the EMOS showed statistically significant skill over the GMOS. The flowchart in Figure 1 describes how the algorithm processes the GMOS and EMOS for each sample set.

While the divisions by forecast period and month yielded relatively equal sample sets, there was the potential for large differences in the sample sizes with differing values for degrees per standard deviation. The sample sizes tend to be larger for small values of degrees per standard deviation with progressively smaller sample sizes for correspondingly larger values of degrees per standard deviation, which are less common. The performance of the GMOS and EMOS in each of these samples was measured by computing the root mean square error (RMSE) for sample size n of the verified temperature forecast per Equation 3, and the number of times the EMOS had a smaller absolute error than the GMOS in percent (EMOSwin) as seen in Equation 4 for sample size n.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{GMOS^i} - \mathrm{EMOS^i})^2}{\mathrm{n}}} \qquad \textbf{(3)}$$

$$EMOSwin = 100 * \frac{\sum_{i=1}^{n}|EMOS - Observed| < |GMOS - Observed|}{n} \qquad \textbf{(4)}$$

The best algorithm performance was found with EMOS RMSE of at least 5% less than the GMOS RMSE and the EMOSwin of at least 51%. The EMOSwin was used in addition to the RMSE performance measurement since several very favorable scores for the EMOS average may misleadingly sway the RMSE comparison, especially for smaller sample sizes. For all cases in which the EMOS met the minimum RMSE and EMOSwin criteria (with a minimum sample size of eight pairs), a *Wilcoxon* non-parametric test for two matched samples was utilized at the 95% confidence level to determine if the EMOS was statistically different from the GMOS. The technique for the *Wilcoxon* test is described in Winkler and Hays (1975). The *Wilcoxon* test was used in lieu of parametric statistical tests such as the *t-test* due to the inherently non-Gaussian bimodal nature of the H/L and +1/-1 groups, and the tendency to depart from a Gaussian distribution for higher values for degrees per standard deviation in the Rest group.

There were instances in which the EMOS RMSE, EMOSwin and minimum sample size criteria were met, but the *Wilcoxon* test was not sufficient to pass at the 95% confidence level. In such cases an additional step was performed by the algorithm which uses the larger annual dataset for the three main data groups. If the *Wilcoxon* test for the larger dataset passes, then the algorithm allowed the case in question to be used. This is not common and affected only about 0.3% of the sample subsets.

4. Results

Overall, the algorithm was fairly conservative and identified just over 5% of the total database in which the EMOS was superior to the GMOS on an annual basis. Less stringent settings for the minimum RMSE and EMOSwin will allow for the algorithm to identify more cases but at the expense of overall improvement over the GMOS. Of the identified cases, Group H/L accounted for about 81% of the identified cases (about 4% of the total database) which seems logical as this group is made up of instances where the

GMOS was equal to the ensemble MOS maxima or minima. This relationship of success with outliers (instances in which the GMOS is equal to the ensemble MOS maxima or minima) in the GMOS breaks down quickly though as the near outlier Group +1/-1 accounts for only 7% of identified cases with Group Rest comprising the remaining 13%.

Significant seasonal differences were found and were the reason for subdividing the database by month as seen in Table 3. The bulk of the identified cases occurred in May through July (47.2%) and the period from October through January (36.8%). For the warmer months, the improvements tended to be associated with daytime high temperatures in which the GMOS was at or near the ensemble MOS minimum whereas the cooler months tended to be concentrated in later forecast periods for both daytime highs and nighttime lows.

Seasonal transition months such as March through April and August through September tended to have the least identified cases. It is not clear what contributes to the decreased EMOS skill during this period and more research is required to explore this, but it is surmised the lower EMOS skill is due to the verifying weather pattern leaning towards a warm or cold solution, rather than a more moderate solution represented by the ensemble average. The EMOS may represent a less plausible solution in such situations rather than choosing either the warm or cold solution, unless the uncertainty is so great the forecaster prefers to resort to other techniques. It rests upon the forecaster to determine which model solution seems reasonable for a given pattern, but results presented here will hopefully assist with the recognition of instances where the EMOS will outperform the GMOS.

One of the motivations for this research was the hypothesis that EMOS should exhibit its greatest value when the model solution envelope is widest (i.e., when the evolution of the weather pattern was most uncertain). The results showed that while the latter is true to some degree, it is usually not the case for short-term forecast periods (days 1-3) or with extremely high values of degrees per standard deviation. Paradoxically, there tends to be a higher degree of skill for the EMOS average at small values of degrees per standard deviation, especially during summer months. Summer values for degrees per standard deviation tend to be smaller due to the lack of large-scale air mass changes, and the significant skill is likely due to the GMOS producing a solution which differs significantly from most of the EMOS members. Another interesting finding was the algorithm was unable to find a statistically valid case in which the EMOS average outperformed the GMOS when the degrees per standard deviation were greater than 7. Bimodal solutions or disparity among the EMOS members are likely the cause, but this could also be due to a lack of sufficient data in these rarely occurring high values of degrees per standard deviation.

Performance matrices were created for each of the three main groups (Group H/L, Group +1/-1 and Group Rest) subdivided by forecast period, degrees per standard deviation and month. Annual performance matrices (see Table 4) show where each of the groups were identified to have skill in a matrix of degrees per standard deviation by forecast period. The monthly performance matrices (see Table 6) are arranged similarly. A visual inspection of the monthly matrices tends to show clustering of EMOS skill in the later forecast periods for values for degrees per standard deviation of 2 to 6 degrees in the cooler months and for small values of degrees per standard deviation (1 to 2 degrees) during the warmer months, which follows upon earlier discussion. Another example of the advantage to using monthly statistics can be seen by comparing the annual and monthly matrices. The annual matrices show EMOS skill generally only with the higher values for degrees per standard deviation and miss the skill found in the higher resolution monthly data such as during the summer months.

As the performance matrices were developed from the entire dataset of the 118 stations, the skill of the performance matrices were evaluated individually for each of the 118 stations by treating each as a small sample of the large dataset. The overall improvement for each station is shown in parentheses in Table 1 and a geographic representation shown in Figure 2. The improvement averages 15.2% and ranges from a high of 26% at Springfield, Missouri (KSGF) to a low of 5% at New Orleans, Louisiana (KNEW). Interestingly, no clear pattern is evident in Figure 2 which suggests that geography is not a significant factor in determining cases in which the EMOS is superior to the GMOS. This also portends the performance matrices could be applied to stations over a wider geographic area.

The algorithm ingests the current GMOS and EMOS and creates a table for temperature guidance in the forecast area of responsibility by utilizing the performance matrices. An example of the table is shown in Table 5 and is known as the "smart guidance product." In this product, whenever the EMOS average is used in lieu of the GMOS, an asterisk is placed beside the number alerting the forecaster to the change. In this example, daytime highs in forecast period 13 are nearly all identified as having higher EMOS skill. Such instances should stand out to the forecaster as an opportunity to provide a better forecast than one obtained from the GMOS alone. Also, experience with this product indicates that when the algorithm selects only a few EMOS temperatures for a given forecast period, the forecaster will need to exercise his or her judgment on whether to use the EMOS selections at all, for the sake of consistency. The reason is that sometimes the EMOS temperatures differ significantly from the bulk of the other neighboring sites where the GMOS temperature was selected for use. Mixing the EMOS and GMOS solutions in this manner may result in areal discontinuities that do not make meteorological sense and conflict with the overall meteorological integrity of the field. On the other hand, if these few EMOS temperatures are grouped in a region of the forecast area, this may be a reasonable solution.

5. Conclusion

The monthly and annual matrices represent the culmination of this research as they contain the specifically identified cases in which the EMOS has higher skill, with an average skill of 15.2%. These can in turn be used in the forecast process to improve upon temperature verification through the skillful use of the EMOS average temperature guidance. The algorithm creates a table that accomplishes this for the forecaster. This "smart guidance product" can be generated daily upon receipt of the latest EMOS and

GMOS and represents the application of this EMOS research. It is hoped the use of such a table will improve verification scores through the skillful application of the EMOS.

Acknowledgements. Thanks to Jeffrey Medlin, SOO, National Weather Service Mobile, AL for his insights and suggestions, and thanks to the reviewers of this research for their guidance.

REFERENCES

Grumm, R.H., J. Villani, R. Hart 2005: An Evaluation of Ensemble MOS Temperature Forecasts from the Medium Range Ensemble Forecast System. *CD 21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction, Washington, DC,* American Meteorological Society, Paper 11A.6. Available online at http://ams.confex.com/ams/pdfpapers/94512.pdf.

Maloney, J. C., K. K. Gilbert, M. N. Baker, and P. E. Shafer, 2010: GFS-based MOS guidance – The extended-range alphanumeric messages from the 0000/1200 UTC forecast cycles. *MDL Technical Procedures Bulletin* No. 2010-01, NOAA, U.S. Dept. of Commerce, 12 pp.

Marz, L.C. 2004: A Comparison of MEX Guidance vs. Ensemble-Based Temperature Forecasts. *National Weather Service Southern Region Technical Attachment* SR/SSD 2004-07. Available online at http://www.srh.noaa.gov/mrx/research/marzstudy.pdf.

MDL, cited 2008: Description of Global Ensemble MOS Message. Available online at http://www.weather.gov/mdl/synop/enstxt.php.

Winkler, R.L., and Hays, W.L. 1975: Statistics: Probability, Inference, and Decision (2nd ed.), New York: Holt, Rinehart, and Winston.

TABLES AND FIGURES

KAEX Alexandria, LA (14%)
KAIZ Kaiser, MO (15%)
KANB Anniston, AL (12%)
KARA New Iberia, LA (11%)
KARG Walnut Ridge, AR (15%)
KASD Slidell, LA (7%)
KAUO Auburn, AL (11%)
KBAD Barksdale AFB, LA (18%)
KBFM Mobile Downtown, AL (14%)
KBHM Birmingham, AL (17%)
KBIX Biloxi, MS (9%)
KBLV Belleville, IL (14%)
KBPK Mountain Home, AR (13%)
KBPT Beaumont, TX (10%)
KBTR Baton Rouge, LA (12%)
KBVX Batesville, AR (12%)
KBWG Bowling Green, KY (8%)
KCBM Columbus AFB, MS (13%)
KCDJ Chillicothe, MO (13%)
KCEW Crestview, FL (17%)
KCGI Cape Girardeau, MO (13%)
KCOU Columbia, MO (16%)
KCPS Cahokia, IL (17%)
KCSV Crossville, TN (9%)
KDCU Decatur, AL (15%)
KDEC Decatur, IL (18%)
KDHN Dothan, AL (14%)
KDMO Sedalia, MO (17%)
KDTN Shreveport Dwntwn, LA (18%)
KDTS Destin, FL (17%)
KEET Alabaster, AL (14%)
KEHR Henderson, KY (18%)
KELD El Dorado, AR (16%)
KESF Alexandria Esle, LA (16%)
KEVV Evansville, IN (17%)
KFAM Farmington, MO (17%)
KFLP Flippin, AR (18%)
KFSM Fort Smith, AR (12%)
KFYV Fayetteville, AR (15%)
KGAD Gadsden, AL (12%)
KGGG Longview, TX (19%)
KGLH Greenville, MS (18%)
KGPT Gulfport, MS (8%)
KGWO Greenwood, MS (12%)
KGZH Evergreen, AL (20%)
KHEZ Natchez, MS (11%)
KHKA Blytheville Muni Aprt, AR (9%)
KHNB Huntingburg, IN (15%)
KHOP Fort Campbell, KY (11%)
KHOT Hot Springs, AR (12%)
KHRO Harrison, AR (10%)
KHRT Hurlburt Field, FL (12%)
KHSV Huntsville, AL (16%)
KIXD Olathe New Century, KS (18%)
KJAN Jackson, MS (18%)
KJBR Jonesboro, AR (19%)
KJEF Jefferson City, MO (16%)
KJLN Joplin, MO (14%)
KLCH Lake Charles Rgnl, LA (14%)
KLFK Lufkin, TX (16%)
KLFT Lafayette, LA (17%)
KLIT Little Rock, AR (14%)
KLLQ Monticello, AR (21%)
KLWV Lawrenceville, IL (20%)
KMAI Marianna, FL (12%)
KMCB McComb, MS (25%)
KMCI Kansas City Intl, MO (15%)
KMDH Carbondale, IL (19%)
KMEI Meridian, MS (13%)
KMEM Memphis, TN (17%)
KMGM Montgomery, AL (18%)
KMKC Kansas City, MO (15%)
KMKL Jackson, TN (21%)
KMLU Monroe, LA (18%)
KMOB Mobile, AL (9%)
KMSL Muscle Shoals, AL (14%)
KMSY New Orleans, LA (5%)
KMVN Mount Vernon, IL (14%)
KMWT Mount Ida, AR (15%)
KMXF Maxwell AFB, AL (7%)
KNEW New Orleans Lakefrnt, LA (7%)
KNMM Meridian NAS, MS (25%)
KNPA Pensacola NAS, FL (20%)
KNSE Whiting Field NAS, FL (24%)

KOCH Nacogdoches, TX (19%)	KSPI Springfield, IL (13%)
KOJC Olathe, KS (17%)	KSTJ St. Joseph, MO (21%)
KOZR Ozark, AL (22%)	KSTL St. Louis, MO (14%)
KPAM Panama City Tyndall, FL (19%)	KSUS Spirit of St Louis, MO (14%)
KPBF Pine Bluff, AR (17%)	KSZL Whiteman AFB, MO (17%)
KPFN Panama City, FL (10%)	KTBN Fort Leonard Wood, MO (16%)
KPIB Pine Belt, MS (20%)	KTCL Tuscaloosa, AL (13%)
KPNS Pensacola, FL (16%)	KTOI Troy, AL (12%)
KPOE Fort Polk, LA (14%)	KTUP Tupelo, MS (21%)
KPOF Poplar Bluff, MO (14%)	KTVR Vicksburg, MS (10%)
KPQL Pascagoula, MS (20%)	KTXK Texarkana, AR (13%)
KRUE Russellville, AR (12%)	KUIN Quincy, IL (12%)
KSET St. Charles, MO (16%)	KUNO West Plains, MO (18%)
KSGF Springfield, MO (26%)	KVBT Bentonville, AR (13%)
KSGT Stuttgart, AR (22%)	KVIH Rolla, MO (16%)
KSHV Shreveport, LA (16%)	KVPS Valparaiso, FL (13%)
KSLG Siloam Springs, AR (13%)	KXNA NW Arkansas Aprt, AR (13%)

Table 1. Stations used in this study with algorithm RMSE percent improvement of EMOS over GMOS for each station. See Figure 2 for geographic references.

```
KMOB   ENSEMBLE MOS GUIDANCE   11/04/2008  0000 UTC 30.68  -88.25
FHR  24| 36  48| 60  72| 84  96|108 120|132 144|156 168|180 192
TUE  04| WED 05| THU 06| FRI 07| SAT 08| SUN 09| MON 10| TUE 11 CLIMO
X/N
GFS  77| 53  79| 55  81| 59  75| 50  77| 50  76| 51  72| 54  74 50 71
AVG  74| 52  74| 53  74| 56  75| 57  72| 47  71| 46  69| 48  68
STD   0|  0   1|  0   1|  1   1|  3   1|  3   1|  1   1|  2   1
HI   77| 54  79| 55  81| 59  78| 61  77| 56  76| 51  73| 54  74
LOW  73| 51  74| 52  73| 55  74| 50  70| 43  67| 43  66| 45  66
```

Table 2. Example of combined GMOS and EMOS guidance product. “GFS” is the GMOS, “AVG” is the EMOS average temperature, and “STD” is the degrees per standard deviation. This type of EMOS summary message was produced by software developed by Timothy Barker (NWS Weather Forecast Office, Boise, ID), a description of which can be found at:
http://www.mdl.nws.noaa.gov/~applications/LAD/generalappinfoout.php3?appnum=2525

January	7.6%	May	9.4%	September	2.9%
February	5.5%	June	15%	October	10.6%
March	2.6%	July	22.8%	November	11%
April	1%	August	4%	December	7.6%

Table 3. Algorithm identified cases by month.

DpStdv	0	1	2	3	4	5	6	7
Pd 7					H	HR	R	
Pd 8						H		
Pd 9						HFR	HR	
Pd 12							HR	R
Pd 13			H					
Pd 15			H		H	FR	R	

Table 4. Annual algorithm performance matrices by degrees per standard deviation (DpStdv) and forecast period. “H” is Group H/L, “F” is Group +1/-1, and “R” is Group Rest.

GMOS/EMOS TEMPERATURE GUIDANCE PRODUCT FOR 11/29/2008 0000Z

DATE	29	SUN	30	MON	01	TUE	02	WED	03	THU	04	FRI	05	SAT	06
KDTS	72	60	64	45	58	38	59	46	67	52	65	50	66	55	69
KGZH	70	54	59	42	48	32	56	36	64	44	61*	40	58*	48	68
KMEI	66	50	55	35	47	32	55	39	64	43	56	39	57*	46	64
KMGM	69	54	57	41	50	33	53	36	64	44	59	38	57*	46	66
KMOB	73	55	63	40	57	34	59	42	70	49	62*	46	59*	52	69
KNPA	72	59	65	42	55	37	56	46	70	52	64	46	62*	53	68
KPNS	73	58	64	42	55	35	59	43	66	50	63	48	61*	44*	68
KVPS	72	59	64	43	54	36	58	42	64	50	64*	48	62*	54	68

* means EMOS guidance was used due to superior performance.

Otherwise GMOS guidance was used due to superior performance.

Table 5. Example of Smart Guidance Product. Where the EMOS temperature guidance was used in lieu of the GMOS, the temperature is followed by an asterisk.

January

DpStdv	0	1	2	3	4	5	6	7
Pd 7					HF	R	R	
Pd 8						H		
Pd 9								R
Pd 10				H	HR			
Pd 11		H	H	H	H			
Pd 12			H	F			H	R
Pd 13						H		
Pd 14						H		
Pd 15					H	R	R	

February

DpStdv	0	1	2	3	4	5	6	7
Pd 8								R
Pd 9						F		
Pd 11					HF			
Pd 12							R	
Pd 13		H	H					
Pd 15	H	H	H		H	R		

March

DpStdv	0	1	2	3	4	5	6	7
Pd 5					H			
Pd 7					R	H		
Pd 8					F			
Pd 9				F		HFR		
Pd 10								R
Pd 12							H	
Pd 13			H					
Pd 15						HFR	R	

April

DpStdv	0	1	2	3	4	5	6	7
Pd 7					H	R		
Pd 8						R		
Pd 9				H				
Pd 11							R	
Pd 12	H							
Pd 14	H			F				
Pd 15				H			R	

May

DpStdv	0	1	2	3	4	5	6	7
Pd 2			F					
Pd 4				H				
Pd 5				R				
Pd 7					H			
Pd 9					R			
Pd 10		H			HFR			
Pd 11				H	H			
Pd 12			F		R			
Pd 13			H	F	R			
Pd 15		H	H	H				

June

DpStdv	0	1	2	3	4	5	6	7
Pd 2		H						
Pd 4		H						
Pd 5		H						
Pd 6		HF	F					
Pd 7		H						
Pd 9		H						
Pd 11		H						
Pd 15		H						

July

DpStdv	0	1	2	3	4	5	6	7
Pd 2		H						
Pd 3			F					
Pd 4		H						
Pd 5		H						
Pd 7		H	F					
Pd 8			H					
Pd 9		H	H	H				
Pd 10			H					
Pd 11		H						
Pd 13			H					
Pd 15		H	H					

August

DpStdv	0	1	2	3	4	5	6	7
Pd 2		H						
Pd 4		H						
Pd 8		H	HF					
Pd 10			HF					
Pd 12			F					
Pd 13			F					
Pd 14			F					

September

DpStdv	0	1	2	3	4	5	6	7
Pd 2		H						
Pd 7				H	H			
Pd 9				H	R			
Pd 11				HR				
Pd 13			H	H				
Pd 15			HFR					

October

DpStdv	0	1	2	3	4	5	6	7
Pd 1			H					
Pd 2			H					
Pd 8			R					
Pd 9	H			HF		R	R	
Pd 10		F	HFR		R			
Pd 11						R	R	
Pd 12		H					R	
Pd 13	H	H	H					
Pd 14			HF					
Pd 15		H	HF	H	H		R	

November

DpStdv	0	1	2	3	4	5	6	7
Pd 5			H	H				
Pd 6					HR			
Pd 7			HF					
Pd 8	H	H	HF	H		R		
Pd 9			HF					
Pd 10	H			H	H			
Pd 11		H		H				
Pd 12		HF	HF	H				
Pd 13		H	H	F	R	H		
Pd 14	HF	H			H			
Pd 15						R		

December

DpStdv	0	1	2	3	4	5	6	7
Pd 3				R	R			
Pd 5		H	H		R	R	R	
Pd 7				H	F	HR	R	
Pd 8				H	R	H	H	
Pd 9					HR	HR	HR	R
Pd 11				H	H		R	
Pd 12							R	
Pd 14			F		F			
Pd 15			HF	H	H			

Table 6. Monthly algorithm performance matrices by degrees per standard deviation (DpStdv) and forecast period. "H" is Group H/L, "F" is Group +1/-1, and "R" is Group Rest.

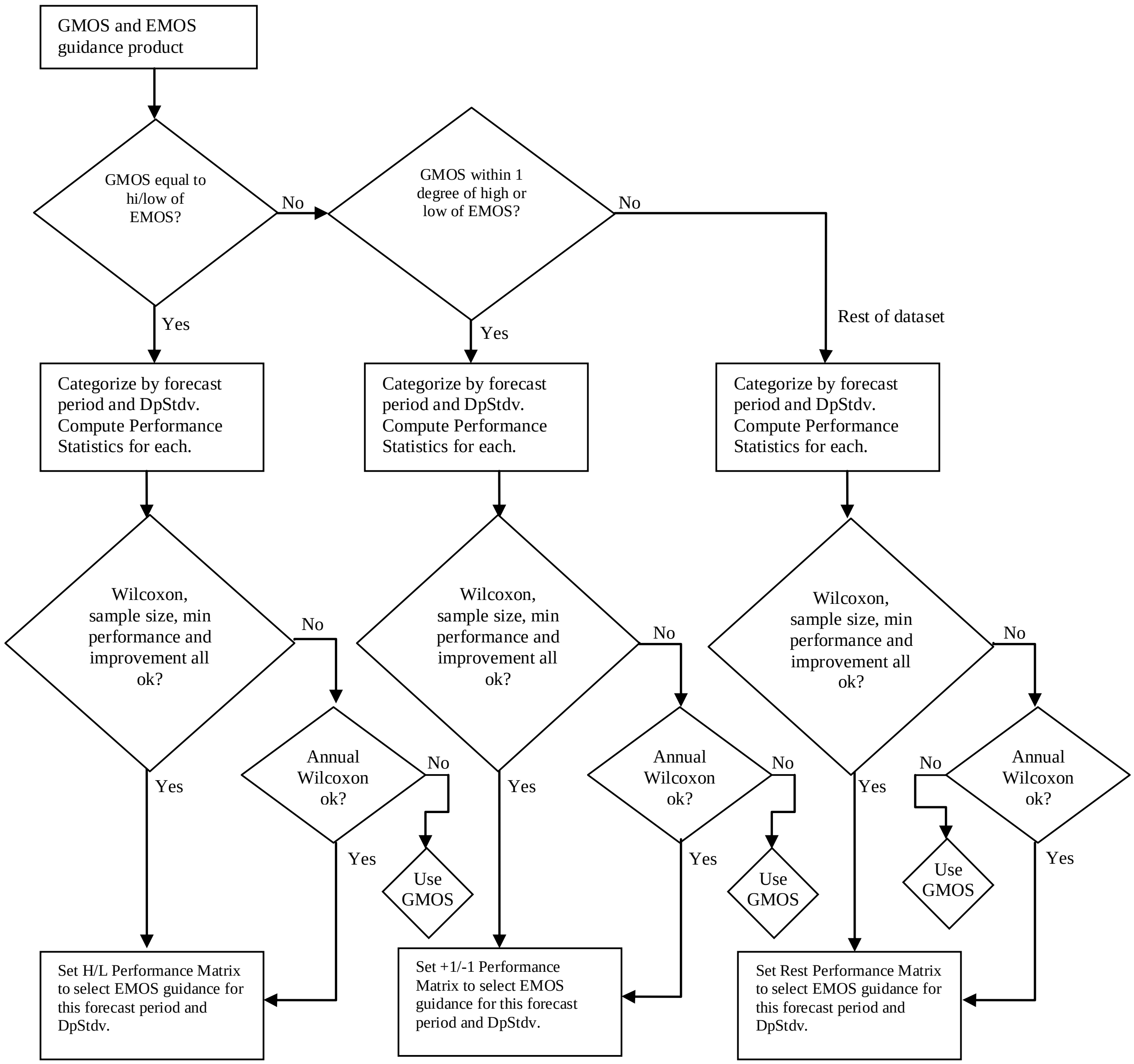

Figure 1. Algorithm flowchart. The algorithm ingests paired GMOS and EMOS data and subdivides each by Group H/L, Group +1/-1 and Group Rest. The *Wilcoxon* non parametric statistical test (Winkler and Hays, 1975) for paired samples was conducted in addition to parameters for minimum RMSE improvement and EMOSwin. A minimum sample size was also used. The results in turn are used to create the monthly and annual performance matrices. DpStdv is the degrees per standard deviation.

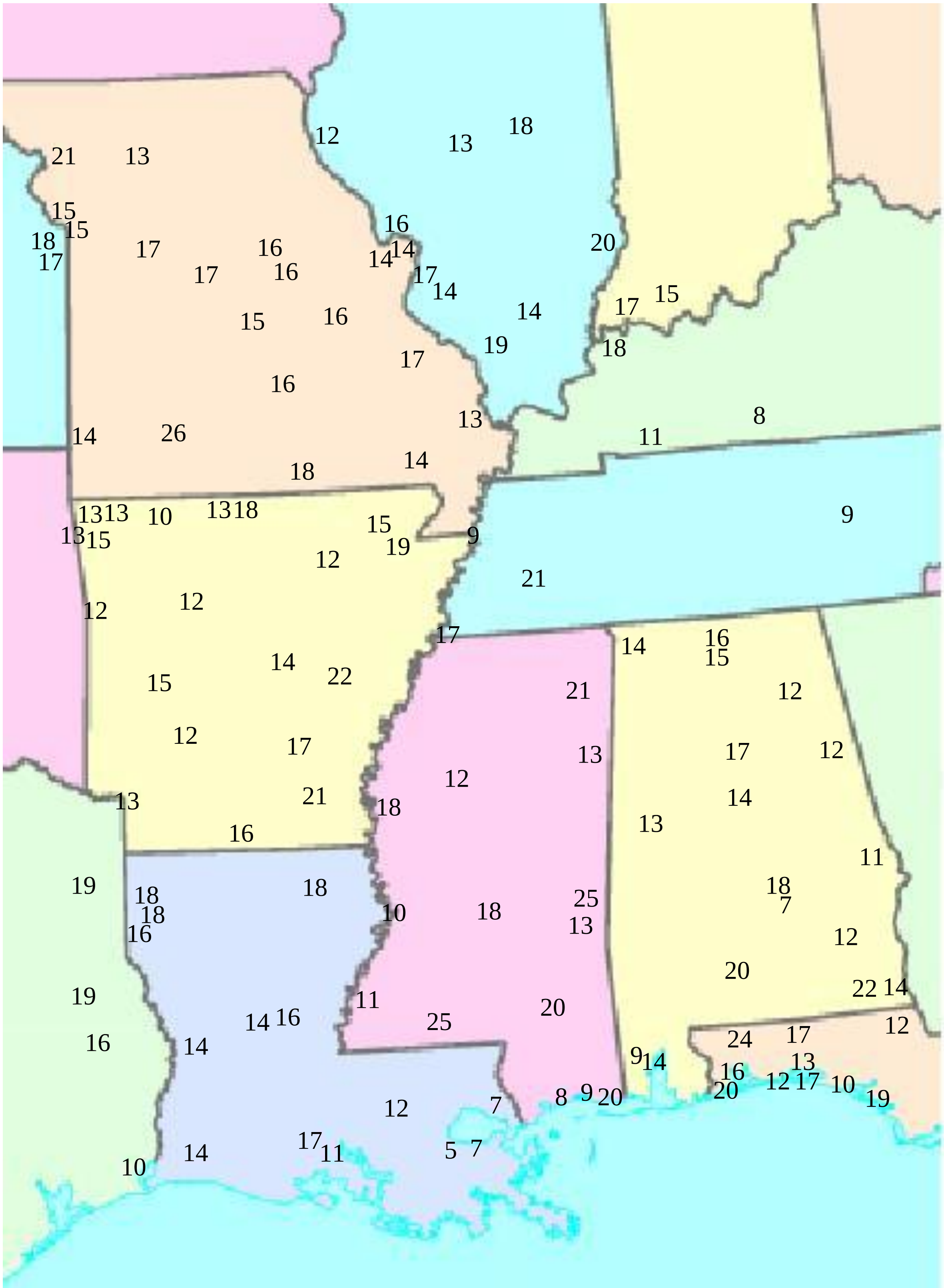

Figure 2. Algorithm RMSE Percent Improvement of EMOS over GMOS.