NWS-CR-TA-88-42

CRH SSD
NOVEMBER 1988

CENTRAL REGION TECHNICAL ATTACHMENT 88-42

A LOOK AT MAXIMUM TEMPERATURE BIAS IN MOS

David R. Tucek
National Weather Service Office
South Bend, Indiana

1. Introduction

Model Output Statistics (MOS) temperature data has been useful and necessary guidance to forecasters. Verification scores have been based on forecasters improvement over this guidance. Changes in MOS temperature predictions from calendar day to daytime maximums has diminished forecasters improvement over guidance. A study at WSFO Indianapolis found improvements over MOS by local Indiana offices decreased approximately 0.5 degrees after this change.

This study attempts to determine the ability of MOS to forecast the direction of maximum temperature change from one day to the next for warming and cooling events. Additionally, warm and cold biases were identified on a monthly and seasonal basis. Results from data for South Bend indicated this approach has some value, particularly in months of seasonal change. Accounting for local effects, similar conclusions may be drawn elsewhere.

2. Data and Method

MOS 24-hour maximum temperature forecasts for South Bend from March 1985 to July 1988 were used in this study. The data was selected from 00Z guidance only. January and February 1987 data were missing.

Criteria for selecting data was based on a change in forecast maximum temperature from one day to the next. If consecutive 24-hour forecasts indicated a maximum temperature decrease from one day to the next, the case was selected as a cooling event, or cool-down. This event would then be compared to the observed change for the same period. Warming events, or warm-ups, were evaluated similarly. Situations with no change forecast were not evaluated.

The selected data was analyzed on a monthly and seasonal basis. This was subdivided into all cases meeting criteria, cases with consecutive daily maximum temperature forecast changes three degrees or less, and cases with day-to-day changes ten degrees or more. The small and large day-to-day differences were chosen to see how well MOS handled slight and significant forecast weather changes. These situations were also examined because forecasters often tend to select temperatures near guidance during small and large weather changes.

3. Results

The overall ability of MOS to correctly forecast the direction of change was evaluated first. Of the cases meeting criteria for change, nearly 80 percent correctly forecast the direction of change. MOS forecast the direction of change better in spring and autumn, the transition seasons, than in summer or winter. MOS was most accurate in March and October. Overall, MOS forecast the direction of change quite well.

Cases of opposing trend were examined briefly as well. Many resulted when day-to-day changes were small. For example, MOS forecast a warm-up of a degree or two, when an observed cool-down of a degree or two actually occurred. These cases were 'casualties' of the analysis criteria and not necessarily poor forecasts. There were situations, however, in which MOS appeared to be forecasting change which opposed the observed trend. Here, MOS appeared to incorrectly forecast the trend as the result of under/over forecasting cloud amount, forecasting/not forecasting precipitation, or unaccounted for influences such as early season snowcover, ground wetter/drier than climatology, or winds off Lake Michigan. Knowledge of these conditions could improve a forecaster's selection of the 'best' temperature for a given event.

The data was examined further to determine the tendency of MOS to underforecast or overforecast change. Underforecasts were cases where cool-downs were not cool enough, warm-ups not warm enough. Overforecasts meant cool-downs were too cool and warm-ups were too warm. The results are presented in Table A. Of the 924 cases, 521 (or 56 percent) were underforecast, 322 (or 35 percent) were overforecast, the remaining nine percent were correctly forecast. Cool-downs were more than twice as likely to be underforecast than overforecast. Warming events were underforecast slightly more than overforecast.

On a monthly and seasonal basis, statistics were similar. In spring, nearly two in three changes were underforecast. Warm-ups were twice as likely to be underforecast than overforecast. This was true for all spring months. Likewise, cooling events were underforecast in nearly three of four cases for all three months in spring. Average absolute errors for spring were 3.4 degrees. The error tabulated most frequently in spring was two degrees while 90 percent of the errors ranged from zero to six degrees.

Summer errors tabulated most frequently were the smallest for any season at one degree. The absolute error for summer was also smallest at 2.6 degrees; 90 percent of the errors ranged from zero to five degrees. June warm-ups were three times as likely to be underforecast than overforecast. July results were insignificant. August cool-downs were twice as likely to be underforecast than overforecast. The drought summer of 1988 was interesting. For June and July, over three-fourths of the warming events were underforecast while two-thirds of the cool-downs were overforecast. The short term drought was broken in August and results were less significant. These statistics suggest soil moisture and climatological temperatures had a negative impact on MOS guidance.

Table A. All Cases

WU: warm-ups CD: cool-downs
+ : overforecast - : underforecast
# : number of cases

Month Season	Total Cases	#WU+	#WU-	#CD+	#CD-	Absolute Error in Degrees F	90% in Range Deg F
Jan	47	13	10	6	15	3.22	0-06
Feb	41	9	7	4	19	3.91	0-07
Mar	101	17	37	10	30	3.46	0-07
Apr	102	17	39	11	27	3.49	0-07
May	95	15	32	10	27	3.26	0-06
Jun	93	11	32	18	20	2.88	0-07
Jul	95	21	31	19	17	2.43	0-05
Aug	66	22	12	7	18	2.36	0-04
Sep	65	14	12	7	23	3.11	0-05
Oct	79	18	20	10	26	3.09	0-06
Nov	72	28	6	5	27	4.73	0-12
Dec	68	23	11	7	23	3.20	0-05
Sp	298	49	108	31	84	3.40	0-06
Su	254	54	75	44	55	2.56	0-05
Au	216	60	38	22	76	3.64	0-07
Wi	156	45	28	17	57	3.44	0-06
Total	924	208	249	114	272	3.26	0-06

NOTE: Total Cases = WU+ + WU- + CD+ + CD- + Exact Fcst

Autumn cases were underforecast slightly more often than overforecast. Cool-downs were underforecast about 70 percent of the time; warm-ups were overforecast a little over half the time. Autumn exhibited the largest range of errors with 90 percent falling within the range of zero to seven degrees. The absolute error was 3.6 degrees. Cool-downs in September were three times as likely to be underforecast than overforecast. Nearly 70 percent of October cool-downs were underforecast. Over 75 percent of November cool-downs were underforecast while the same percentage of warm-ups were overforecast. It appeared that in November, when climatological transition toward winter is greatest, MOS exhibited a warm bias.

Winter cases were underforecast 56 percent of the time. Forty percent were overforecast with nearly 75 percent of these occurring during warming events. Cool-downs were three times as likely to be underforecast than overforecast. On a monthly basis, two in three December warm-ups were overforecast; January and February warm-ups were overforecast slightly more often than underforecast. Cool-downs in December, January and February were underforecast 70, 71, and 76 percent of the time, respectively. The average absolute error for winter was 3.4 degrees; 90 percent ranged from zero to six degrees.

Overall, autumn and winter exhibited a warm bias for both warming and cooling events. Spring and summer had a warm bias for cooling events and a cold bias in warming events.

Further analysis of small day-to-day changes (three degrees or less) and large day-to-day changes (ten degrees or more) yielded similar statistics in most cases. As a result, only significant results will be presented with seasonal data in Table B.

Table B: Same as A except small changes (top) and large (bottom)
Small: Temperatures three degrees F or less;
Large: Ten degrees F or more

Season	Total Cases	#WU+	#WU-	#CD+	#CD-	Absolute Error in Degrees F	90% in Range Deg F
Sp	64	9	30	5	14	2.83	0-06
Su	92	21	33	11	15	2.22	0-04
Au	58	12	14	8	20	3.12	0-08
Wi	43	17	1	5	16	2.70	0-05
Sp	62	11	15	6	26	3.90	0-07
Su	11	1	1	3	5	4.38	0-10
Au	33	8	3	3	15	3.34	0-06
Wi	30	6	8	5	10	3.40	0-09

For small daily changes, winter was the only season to overforecast more often than underforecast (51 percent to 40 percent). Warm-ups were overforecast 85 percent of the time, cool-downs were underforecast 70 percent of the time. In spring, almost 70 percent of the cases were underforecast. About three in four warm-ups and two in three cool-downs were underforecast. Summer yielded no significant conclusions. Autumn results were not very impressive either; overforecasts occurred a little more often than underforecasts. In autumn, cool-downs were underforecast about 70 percent of the time; warm-ups were overforecast as often as underforecast. The range of errors in autumn was the largest of any season.

Results for large daily changes were again similar to all cases and the small day-to-day changes. In winter, underforecasts outnumbered overforecasts (18 to 11). Cooling events were underforecast twice as often as overforecast. Warm-ups were underforecast slightly more often than overforecast (eight to six). The range of errors was large (90 percent within zero to nine degrees). In spring, two-thirds of the cases were underforecast with almost 80 percent of cool-downs being underforecast. Warm-ups, as in winter, were underforecast slightly more often than overforecast. In summer, only 11 cases had large daily changes. Cool-downs accounted for eight of those with five of the eight being underforecast. Finally, in autumn, underforecasts outnumbered overforecasts (18 to 11). Warm-ups were nearly three times as likely to be overforecast than underforecast. On the other hand, cool-downs were about five times as likely to underforecast.

Again, as in all cases, these results indicated a warm bias for all seasons during cooling events. Warming events had variable results dependent on season. Statistics for these large day-to-day changes were likely to be less reliable due to the small number of cases meeting criteria.

4. Conclusion

MOS maximum temperature data for South Bend was analyzed to determine the frequency and magnitude of temperature bias for warming and cooling events. Results were tabulated on a monthly and seasonal basis. Overall, MOS correctly forecast the direction of maximum temperature change from one day to the next approximately 80 percent of the time. During autumn and winter MOS exhibited a warm bias for both warming and cooling events. During spring and summer MOS had a warm bias for cooling events and a cold bias for warming events. The transition seasons appeared to be the most reliable predictors of bias. Summer was the least reliable for predicting trends although the drought summer of 1988 (June and July only) consistently underforecast warm-ups and overforecast cool-downs. Average absolute errors ranged from 2.6 to 3.6 degrees. For all seasons, errors tabulated most frequently ranged from one to three degrees. The range of errors was largest in autumn. Subdividing all cases into small day-to-day maximum temperature change (three degrees or less) and large (ten degrees or more) daily changes yielded similar results in most cases.

One can conclude that overall, MOS is hard to beat, especially in summer and the summer months. At the same time, the results suggest that some improvements can be made. In particular, cooling events in transition seasons provide forecasters the best chance to beat MOS, and in some cases, beat MOS significantly. Questions to ask to determine how well MOS will do include: How has MOS and the LFM handled this event? Are dry bulb and dew point temperatures much lower than MOS? Are clouds forecast? In particular, are low clouds being forecast? Is precipitation occurring or expected by the forecaster even though MOS POPS are low? In South Bend, high pressure building in from Canada does not always bring fair weather, rather in fall it often produces northwest winds which bring cold air stratocumulus off of Lake Michigan. If air-water temperature differences are large enough and low level circulation is even slightly cyclonic, precipitation often occurs. MOS will usually miss the precipitation, the amount and height of clouds and, as a result, the temperature, sometimes by more than ten degrees.

Granted, South Bend has local effects not experienced by other locations, but most weather offices have some geographical feature that effects local temperatures in similar ways. If the results of this study parallel effects at other stations, then one can use this data to at least decide 'which side of MOS' to forecast on in many cases. Then, if MOS is in large error in various parameters which will effect temperature, improvements of more than a degree or two may be made.

Small daily changes forecast by MOS might be improved upon as well. This could be particularly useful for transportation purposes, for example, if the forecaster can determine the correct side of freezing to forecast on during a precipitation event. Results here, unfortunately, will not guarantee which side to choose. MOS, no doubt, is good guidance, but we can be better.

5. Acknowledgements

The author would like to thank Richard Livingston of Scientific Services Division for review of this manuscript and for providing useful comments and suggestions.

6. References

Edwards, C., 1988: MOS Guidance ... Can It Get Much Better? Central Region Technical Attachment 88-3, National Weather Service Central Region, Scientific Services Division, Kansas City, MO.

National Weather Service, 1985: Automated Daytime Maximum, Nighttime Minimum, 3-Hourly Surface Temperature, and 3-Hourly Surface Dew-Point Guidance. NWS Technical Procedures Bulletin No. 356. NWS Headquarters, Silver Springs, MD.