COMPARING NUMERICAL MODEL'S DAYS 3, 4, 5 AND 6

JERROLD A. LA RUE
RETIRED-NATIONAL WEATHER SERVICE

CONTENTS

1. ABSTRACT

Forecasters throughout the United States (US) are routinely preparing public and special weather forecasts extending up to 7 days into the future and alluding to even longer time periods. They are basing these forecasts on numerical products, principally the 500 mb forecasts, emanating from a variety of sources. Forecasters may make their decision on which model or models to use because of recent successes of a model, remembered successes of a model with a certain weather pattern, a consensus of two or more favored models, or a models consistently forecasting the same pattern on successive forecasts. On the other hand they may eliminate a model because of a possible error in initialization, a seemingly inconsistent and hard to follow progression of features, a recent poor forecast, or divergent patterns by all the models. The forecaster may choose to not use any of the models but to revert to generalities, persistence or climatology . No comprehensive verification system has been generated to evaluate the usefulness of the forecasts to forecasters for the extended time periods. Forecasters have described some erroneous forecast patterns on various models but these are based on experience, are subjective and concern the first three forecast days (Funk, Ted 2001). There have been objective verifications of derived parameters such as temperatures, relative humidity and precipitation, but again mostly for shorter range forecasts. Otherwise, the weather centers producing the numerical products verify some parameters mainly on a global basis, using statistics such as bias and standard deviation (NCEP 2001)(Environment Canada 2001). There appears to be a lack of a system objectively verifying the medium range numerical models using parameters used by forecasters. This paper continues the development of a verification system to evaluate the 500 mb forecasts for the US area, and which could be adapted to verify other Numerical Weather Products (NWP) (La Rue 2000). The data in this paper was gathered by manpower, but the system could be modified for automated computerization providing up to date statistics regarding the credibility of the models. Verification results for four models comprising 100 forecast days are presented.

2. PROCEDURE

Forecasts in the US are prepared during four time periods. The two principal forecast preparation times are from 2 to 5 am local time (LT) and from 2 to 5 pm LT. The forecasts in mid morning and in the evening are mainly a review of the earlier forecasts and are updated as necessary. The late afternoon forecast from 2 to 5 pm LT requires moving the first day of the extended to the second day of the regular forecast and adding a new day for the last day of the extended period. To clarify forecast days as used in this paper, forecast Day 1 is the first day in the regular forecast, Day 2 is the second day, Day 3 is the first day of the extended portion, Day 4 is the second day of the extended and so on. The afternoon forecaster has guidance products from a number of numerical models, some based on the preceding 1200 UTC data and some from 0000 UTC data that is nearly 24 hours old by the time of use by the forecasters. This study grades the 500 mb numerical charts that are most likely to be used by the forecaster. Consideration of numerical models for inclusion in the study was that they were based on the most recent data ,1200 UTC, and were available as guidance for the afternoon forecasts. Those models using 1200 UTC data were the US National Weather Service's (NWS) Aviation Extended Model (AVN), US Navy's Operational Global Atmospheric Prediction System (NOGAP or NOGAPS) and the European Center For Medium Range Weather Forecasts (ECMWF but referred to as EURO). The EURO may not be available until about 0130 UTC which is well past the afternoon forecast LT dead line. The inclusion of the US NWS's Medium Range Forecast (MRF), using the 0000 UTC data base, was due to it's wide usage and the availability of forecasts verifying at the same time, 1200 UTC, as the other models. Models not included were: 1.) the United Kingdom Model (UK) while based on 1200 UTC data is only available to 72 hours on the Internet. (The UK Forecast Center and the National Centers for Environmental Prediction (NCEP) denied knowledge of the availability of the longer range charts for the U.S. area.) And 2.) the Environmental Canada's Global Environmental Multi-Scale Model which is based on 0000 UTC data and forecast verification times were not compatible with the times of the study.

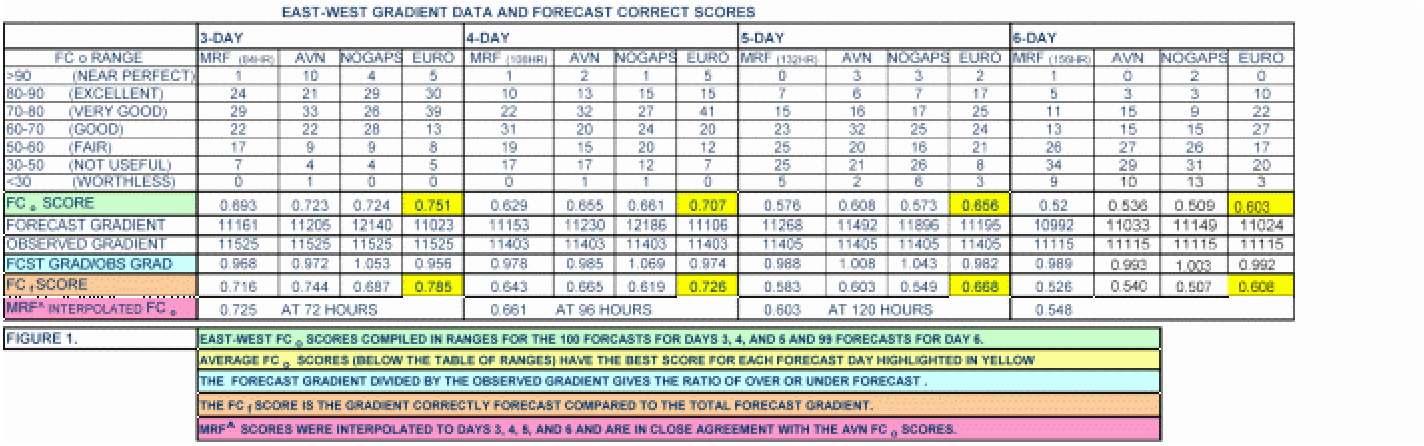
A grid was created over the US which included an area off each coast and was near the Northern and Southern borders. Intersections of 30, 40 and 50 degrees North Latitudes at 70, 85, 100, 115 and 130 West Longitude provided a rather course grid of 15 points. Appendix 2 shows an initial analysis from grid point data compared with a numerical initial analysis. Each day, 500 mb heights in decameters (dam), forecast by each model for the 4 forecast days, were collected for the 15 points. The forecasts were obtained from the Internet and preferred charts were those that had map details including Latitude and Longitude lines. The preferred Internet address for the MRF was the US National Oceanic and Atmospheric Administration's (NOAA) Air Research Laboratory (ARL) http://www.arl.noaa.gov/ready/mrfanim.html with the back up being the US Navy's http://152.80.49.210/PUBLIC/WXMAP/GLOBAL/. The preferred addresses for the AVN were Edwards Air Force Base's http://www.edwards.af.mil/weather/avnmodel.htm for Day 3 and Day 5 forecasts, http://www.emc.ncep.noaa.gov/forecasts/ forecasts for the Day 4 land grid points and http://sgi62.wwb.noaa.gov:8080/STATS/MAPS.html for the Day 4 points over the ocean and all Day 6 grid points. This latter address presents global maps which were enlarged 250 times to help interpolate data. It also has archived charts for several months which were used as a backup to the NCEP address. It should be noted that this was the only source for the 6 Day AVN forecast, and it also has a Day 7 AVN forecast. The preferred source for the

NOGAP model was http://152.80.49.210/PUBLIC/WXMAP/GLOBAL/. It also has links to forecasts for a few prior days which could be used as back up at times. http://weather.unisys.com/ecmwf/ecmwf_500p_4panel.html was the preferred address for the EURO model. Once in a while, other addresses were used and an excellent listing of sources for numerical internet products is http://atmos.umd.edu/~wxchair/models.html. A goal was set of 100 forecasts for each of the models for each of the 4 days. If a forecast was missing for any day by one model, then forecasts for that day by the other models were deleted and an additional forecast day was substituted. During the compilation of the data, it was discovered that one Day 6 forecast was erroneous and unrecoverable so that forecast day was deleted for all models and there are only 99 forecasts for Day 6. The verifying chart was the AVN Initial and the preferred source was the Edwards AFB address or if unavailable, the US Navy address, both given above.

The 500 mb heights to the nearest dam were interpolated from the charts and entered on a prepared form. This form allowed rechecking of data should an inconsistency appear. It was then entered in a spreadsheet which performed all of the computations. The difference in forecast heights between adjacent East to West (E-W) points and North to South (N-S) points produces the 500 mb height gradients, Gf between points. The E-W gradients are a measure of the N-S wind component and the N-S gradients are a measure of the E-W wind. The difference computations proceeded from East to West and from North to South. This meant positive values between E-W points indicated a southerly wind component and negative values indicated a northerly wind component, while positive values between N-S points represented westerly winds and negative values indicated an easterly wind component. The data is presented in two sections, one for the E-W gradients, and the other for the N-S gradients. The sum of the absolute values of the E-W height differences between points and of the N-S height differences was computed for each forecast giving the forecast gradients denoted as (abs)Gf. The observed 500 mb heights as provided by the initial analysis of the AVN model allowed the same differencing computations so as to arrive at the observed gradients, Go, and the summation of the absolute values, (abs)Go. The gradients between each pair of forecast points and each matching pair of observed points were compared for value and sign. The forecast gradient that was correct, Gc, is the absolute value in which Gf matched Go in sign and was the smaller of Gf or Go, including zero if either were zero. If Gf and Go were not of the same sign, then Gc was also zero. Gc by definition is either zero or a positive number. A term called the Forecast Correct, FC_o, was obtained by dividing the sum of Gc by the sum of (abs)Go. Since FC_o is not affected by the total forecast gradient, it might be increased by over forecasting with no penalty. A second Forecast Correct Score, FC_f, was developed by dividing the sum of Gc by the sum of (abs)Gf. If multiplied by 100, the Scores produce the percentage of Go that was, on average, correctly forecast in the first case, and the percentage of Gf that would verify as correct in the second case. Since the forecaster is studying forecasts, the second score, FC_f, is probably the more useful and is the one that predominates in the following tables and charts. Mathematical derivations of the formula are in Appendix 1.

3. EAST-WEST GRADIENT DATA

Figure 1 is a table summarizing the results derived from the East-West gradients. The FC_o Scores were sorted into ranges for each model for each day. A worded description was added to each range to indicate the probable value of the forecast range to the forecaster. The numbers in each range are also the percent of the cases in that range as there were 100 cases for each forecast (99 cases for Day 6). When FC_o is as small as .50, only half of the observed gradient was correctly forecast and it is reasonable to assume the forecast is of no value to the forecaster. Using this and a perfect 1.00 as pegs to grade Scores, those in the 90s would be nearly perfect, those in the 80s excellent, etc. The FC_o Scores for each model for each period are below the ranges, and the best score is highlighted in yellow. The FC_f is also given and is located below the totals of Gf and Go and below the results of dividing the Gf by Go. This latter computation gives the ratio of under or over forecasting and shows that, except for the NOGAPS model and the AVN model for Day 5, all models consistently under forecast the gradients by a small amount. As a result, those FC_f Scores were a bit better than the FC_o Scores. The final row gives an FC_o Score for the MRF which was interpolated time-wise to obtain probable scores for the same forecast times as the other models. It should be noted with no surprise that the interpolated MRF FC_o Score and the AVN FC_o Scores are very similar.



EAST-WEST GRADIENT DATA AND FORECAST CORRECT SCORES

	3-DAY				4-DAY				5-DAY				6-DAY			
FC o RANGE	MRF (84HR)	AVN	NOGAPS	EURO	MRF (108HR)	AVN	NOGAPS	EURO	MRF (132HR)	AVN	NOGAPS	EURO	MRF (156HR)	AVN	NOGAPS	EURO
>90 (NEAR PERFECT)	1	10	4	5	1	2	1	5	0	3	3	2	1	0	2	0
80-90 (EXCELLENT)	24	21	29	30	10	13	15	15	7	6	7	17	5	3	3	10
70-80 (VERY GOOD)	29	33	28	39	22	32	27	41	15	16	17	25	11	15	9	22
60-70 (GOOD)	22	22	28	13	31	20	24	20	23	32	25	24	13	15	15	27
50-60 (FAIR)	17	9	9	8	19	15	20	12	25	20	16	21	28	27	26	17
30-50 (NOT USEFUL)	7	4	4	5	17	17	12	7	25	21	26	8	34	29	31	20
<30 (WORTHLESS)	0	1	0	0	0	1	1	0	5	2	6	3	9	10	13	3
FC_o SCORE	0.693	0.723	0.724	0.751	0.629	0.655	0.661	0.707	0.576	0.608	0.573	0.656	0.52	0.536	0.509	0.603
FORECAST GRADIENT	11161	11206	12140	11023	11153	11230	12186	11106	11268	11492	11896	11196	10992	11033	11149	11024
OBSERVED GRADIENT	11525	11525	11525	11525	11403	11403	11403	11403	11405	11405	11405	11405	11115	11115	11115	11115
FCST GRAD/OBS GRAD	0.968	0.972	1.053	0.956	0.978	0.985	1.069	0.974	0.988	1.008	1.043	0.982	0.989	0.993	1.003	0.992
FC_f SCORE	0.716	0.744	0.687	0.785	0.643	0.665	0.619	0.726	0.583	0.603	0.549	0.668	0.526	0.540	0.507	0.608
MRF^A INTERPOLATED FC_o	0.725	AT 72 HOURS			0.661	AT 96 HOURS			0.603	AT 120 HOURS			0.548			

FIGURE 1.

EAST-WEST FC_o SCORES COMPILED IN RANGES FOR THE 100 FORCASTS FOR DAYS 3, 4, AND 5 AND 99 FORECASTS FOR DAY 6.

AVERAGE FC_o SCORES (BELOW THE TABLE OF RANGES) HAVE THE BEST SCORE FOR EACH FORECAST DAY HIGHLIGHTED IN YELLOW

THE FORECAST GRADIENT DIVIDED BY THE OBSERVED GRADIENT GIVES THE RATIO OF OVER OR UNDER FORECAST .

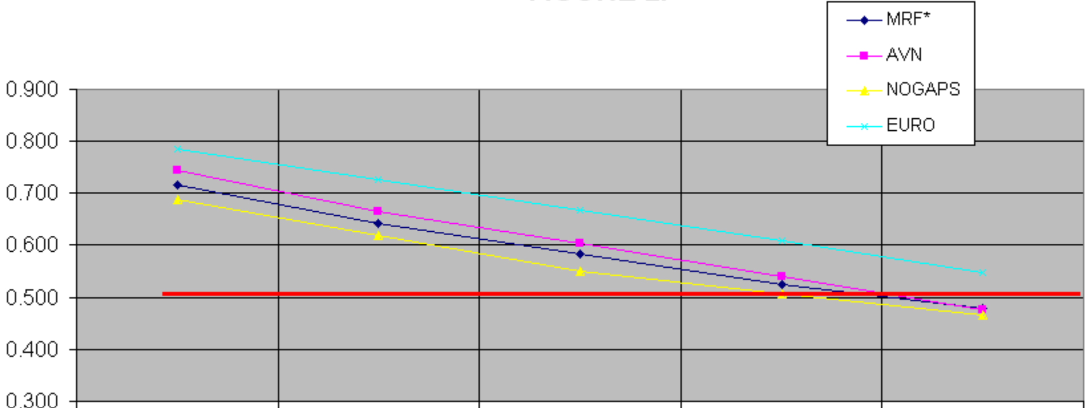
THE FC_f SCORE IS THE GRADIENT CORRECTLY FORECAST COMPARED TO THE TOTAL FORECAST GRADIENT.

MRF^A SCORES WERE INTERPOLATED TO DAYS 3, 4, 5, AND 6 AND ARE IN CLOSE AGREEMENT WITH THE AVN FC_o SCORES.

(click image to enlarge)

Figure 2 is a chart of the FC_f Scores. The Figure 1 data was interpolated to add a seventh forecast day since some extended forecasts go that far in advance. Figure 2 is important as it also provides a time-line for forecasting from Day 3 through Day 7. The forecasts in this study all verified at 1200 UTC of each of the forecast days. 1200 UTC ranges from 0400 to 0800 LT in the Continental US which in all cases is in the morning hours. The daytime forecast period extends, approximately, from the center of each column to the line separating the columns and the nighttime forecast from those lines to the center of the next column. For example, the AVN FC_f Score of .744 on DAY 3 morning, has decreased to .700 by that evening. By the evening of DAY 6, all models except the EURO have FC_f Scores of .500 or less. The FC_f Score of .500 on the chart is in bold red denoting the score at which the forecasts become valueless. Not only does the EURO appear to have value at the morning of DAY 7, further interpolation indicates an FC_o Score of about .518 that evening.

FIGURE 2.



	3-DAY	4-DAY	5-DAY	6-DAY	(7-DAY)
MRF*	0.716	0.643	0.583	0.526	0.479
AVN	0.744	0.665	0.603	0.540	0.477
NOGAPS	0.687	0.619	0.549	0.507	0.465
EURO	0.785	0.726	0.668	0.608	0.548

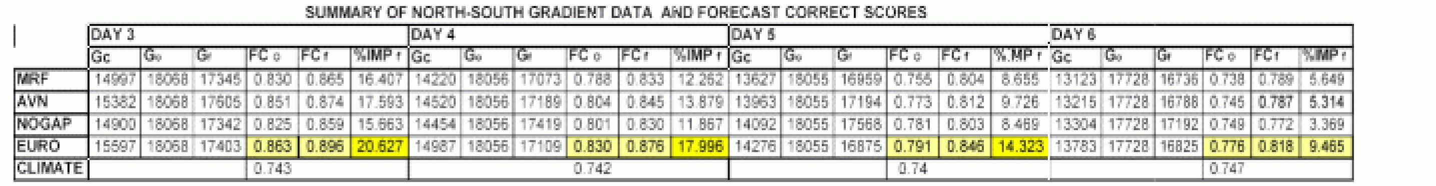
EAST- WEST FCf SCORES

There is no question but that the EURO Model performed the best in the East-West gradients for both FC Scores. Unfortunately it is not available in time for the afternoon forecast cycle. The AVN model is slightly better than the NOGAPS and the older MRF.. By Day-6, all forecasts available for the afternoon extended forecast package were at or near the .50 line. This strongly suggests the Day-6 forecast prepared during the afternoon cycle lean towards climatology as should any day or days past Day-6. When the EURO model becomes available in the evening, Day-6 could be amended beneficially, and even Day 7 might be improved on average.

An examination of individual model scores as they proceeded from day 3 to day 6 showed only a little correlation from one day to the next. DAY-3 could be quite low but scores might improve markedly for the rest of the forecast days. Or a score could be quite high for DAY3 and quickly deteriorate by the fourth or fifth day. There was some similarity between models scoring well or poorly for the same forecast series. It appeared that some forecast situations were more difficult for all of the models, but perhaps only for one or two of the forecast days. It may well be that the variability in scores was due to the 15 degree interval between east-west grid points and a shorter grid length would not only decrease the variability in the scores but allow a better comparison of the different model's ability to deal with different weather systems. The reader is urged to refer to Appendix 2 which critically examines the grid system.

4. NORTH-SOUTH GRADIENT DATA

The grid used in the study was along 50 North, 40 North and 30 North Latitude. The N-S gradient is predominately negative meaning the wind flow is usually from West to East. There were a few closed upper lows South of 50N during the study as well as upper ridges that extended North of the more positively or negatively tilted upper troughs. In the case of the E-W gradients one would expect them to be about equal or the total southerly flow would about equal the total northerly flow. This allowed logical assumptions to be made about the FC Scores for the E-W gradients. These assumptions can not be made concerning the N-S FC Scores. This is not to say the Scores are invalid, only that the Scores are meaningless except for allowing a comparison between models. To measure the usefulness of the Scores to forecasters, it was decided to establish a base score using the average gradient which will be denoted as the "Climate FC" Score in the study. The average gradient between 50 and 40 North was 17.3 dam and between 40 and 30 North was 18.3 dam. These values were inserted into the spreadsheet as forecasts and Climate FC Scores were obtained for each forecast period. It should be noted that Go and Gf are the same so that the Climate $FC_{f \text{ and } o}$ Scores are the same. A comparison of the Go and Gf scores in Figure 3, show that all of the models under forecast the N-S gradient on all four Days. The NOGAP model forecast the N-S gradient to increase on each successive day reaching about 97 % of Go on Day 6. The other models generally decreased Gf each day reaching about 95% of Go by Day 6. The FC Scores of the models that are equal to or are less that this Climate FC Score are considered to have no skill. By Day 6 the FC_o scores of the MRF and the AVN are less than the Climate FC Score. Since the Forecasters are studying forecasts, the percentage improvement of the forecast model's FC_f over the Climate FC scores are included in Figure 3 and are plotted against time in figure 4.



SUMMARY OF NORTH-SOUTH GRADIENT DATA AND FORECAST CORRECT SCORES

	DAY 3						DAY 4					
	Gc	G_o	G_f	FC_o	FC_f	$\%IMP_f$	Gc	G_o	G_f	FC_o	FC_f	$\%IMP_f$
MRF	14997	18068	17345	0.830	0.865	16.407	14220	18056	17073	0.788	0.833	12.262
AVN	15382	18068	17605	0.851	0.874	17.593	14520	18056	17189	0.804	0.845	13.879
NOGAP	14900	18068	17342	0.825	0.859	15.663	14454	18056	17419	0.801	0.830	11.867
EURO	15597	18068	17403	0.863	0.896	20.627	14987	18056	17109	0.830	0.876	17.996
CLIMATE				0.743						0.742		

	DAY 5						DAY 6					
	Gc	G_o	G_f	FC_o	FC_f	$\%.MP_f$	Gc	G_o	G_f	FC_o	FC_f	$\%IMP_f$
MRF	13627	18055	16959	0.755	0.804	8.655	13123	17728	16736	0.738	0.789	5.649
AVN	13963	18055	17194	0.773	0.812	9.726	13215	17728	16788	0.745	0.787	5.314
NOGAP	14092	18055	17568	0.781	0.803	8.469	13304	17728	17192	0.749	0.772	3.369
EURO	14276	18055	16875	0.791	0.846	14.323	13783	17728	16825	0.776	0.818	9.465
CLIMATE				0.74						0.747		

(click on image to enlarge)

FIGURE 3. Gc is the gradient correctly forecast, Go is the observed gradient and Gf is the total forecast gradient. FC_o is Gc divided by G_o and FC_f is Gc divided by Gf.. CLIMATE is the FC Score using the average N-S gradient during this study. $\%IMP_f$ is the improvement of the FC_f scores of the models over the CLIMATE FC Score

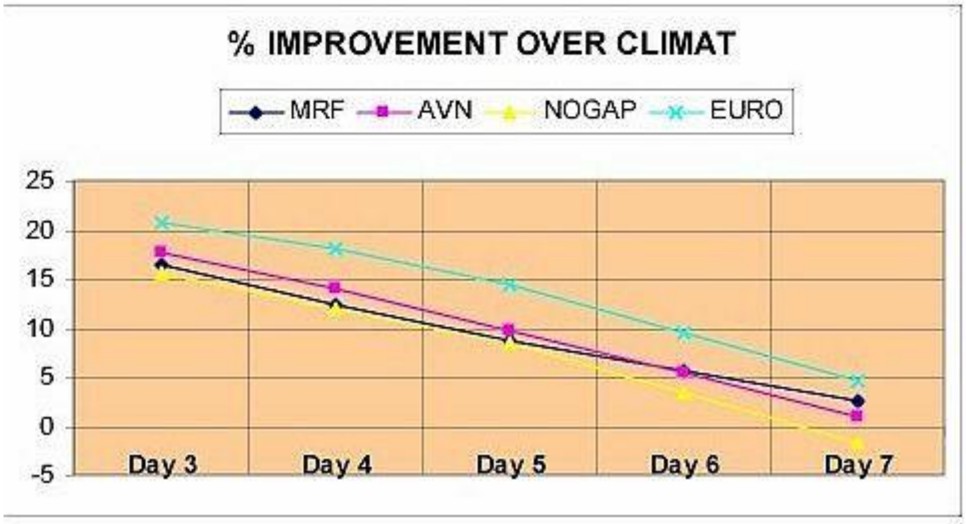


FIGURE 4. The percent improvement of the model's FC_f over the CLIMATE FC is plotted against
time with day 7 values estimated using the persistence decline from day 5 to day 6.
The 0.0% FC_f improvement indicates no improvement over CLIMATE.

It is obvious that the EURO Model's forecasts were the best in every category for the N-S gradients. The average CLIMATE gradient was 35.6 dam from 50 degrees to 30 degrees North Latitude along each of the Longitude lines used in the study. A 20 percent improvement amounts to 7.12 dam, 15 percent is 5.34 dam, 10 percent is 3.56 dam and 5 percent is only 1.78 dam. The 5 percent improvement is the probable point at which forecasts are no longer useful. That point is reached on Day 6 by all models but the EURO, which reaches it on Day 7.

5. CONCLUSIONS

The EURO Numerical Model was revised in November of 2000 (ECMWF 2001). It shows a significant improvement in the comparative FC Scores it received over those in the previous paper by La Rue (2000). The EURO was significantly superior to the other models for both E-W and N-S gradients and for all time periods. The Extended AVN is a new Model but it is essentially the MRF based on 1200 UTC data. It might be expected that the MRF would be superior to the AVN since it collects and digests data over a longer time. It was thought, with reason at the time, more data becomes available world wide, given more time for collection. This delay in data cut-off time was initiated, like maybe in 1960, but with the vast improvement in communications and data sources, it may be past time to review that policy. The AVN is a little better than the MRF for the verifying times used in the study. However, if the MRF is interpolated so that times-to-verification are the same, then the two are about equal, which presents a strong argument against the longer data cut-off time for the MRF. The NOGAPS lagged a bit in the scores and especially when compared to forecast gradients as it tended to over forecast gradients.

So, the time has come to appraise what this information might mean to a forecaster. First, the afternoon cycle forecaster must shift the existing Day 3 forecast into the new regular two day forecast as Day 2. This problem is not addressed in the study. Secondly, a last day, which might be Day 5, 6 or 7, must be added to the extended forecast and each other Day in the lapsing extended forecast must be reviewed. From the stand point of guidance availability, the forecaster has recent AVN and NOGAPS forecasts, the MRF, based on data 12 hours earlier, and the EURO based on 1200 UTC data but 24 hours earlier. This study shows that, for the afternoon forecast cycle, the AVN is about the equal of the EURO which is 24 hours old. The MRF and the NOGAPS are not quite as accurate as the AVN. While it might apparently be advantageous to produce an entire extended forecast at this time, which would be inviolate for the ensuing 24 hours, this study suggest otherwise. First, the EURO arrives at about 0130 UTC which is in the range of 2030-2130 LT on the East Coast and 1730-1830 LT on the West Coast. It is too late for guidance inclusion to the afternoon forecaster. However, the evening update forecaster has the new EURO which is, on average, one forecast day superior to the guidance used that afternoon. This strongly suggests that the evening forecasters should carefully review the existing extended forecast and give consideration to amendments in light of the EURO forecasts. Amending the entire forecast package due to changes only in the extended period might seem pointless in an updating scheme which implies updating only the shorter range forecast. None-the-less, it should be considered in this day and age of effortless word processing and near instantaneous communications. If it is not policy to amend the extended during the evening update cycle, then the EURO should always be evaluated as new data during the early morning forecast cycle. The MRF model's data cut-off time of about 0830 UTC makes it available at about 0330-0430 Eastern LT ranging to 0030-0130 Pacific LT. Eastern Time Zone early morning forecasters would have to delay issuing the forecast package until after 0430 DLT and until after 0330 Eastern Standard Time (EST). Central Time Zone early morning forecasters would need to delay forecast issuance until after 0330 LT during DLT and after 0230 on Central Standard Time in order to take advantage of MRF guidance. An examination of the release time of the Area Forecasts by Eastern and Central Time Zone NWS Offices shows this is not the case by any means and, in fact in some cases, not even by Mountain Time forecasters. Media forecasters and private meteorologists preparing more local and special forecasts have the advantage in being able to make use of the MRF, even in Eastern Time Zones during DLT. It is probable that the mid-morning NWS forecasters doing the update cycle treat the MRF as history with rare amendments to the extended period. This would then mean it would not be used until the afternoon forecast cycle. The solution is to shorten the MRF data cut-off time by two to four hours to ensure it's use by early morning forecasters. The later MRF using 0000 UTC data has the benefit of 12 hours over the EURO in time-to-forecast-verification, but even so would still be about a half a day less accurate than the EURO. Referring to Figure 2 and translating the AVN decay curve to the right by a half a day, would probably be a good estimate of the decay line for the later MRF. This evaluation of a later MRF forecasts shows it to have some value into part of Day 7 and should certainly be considered along with the EURO in amending the extended forecasts by the early morning forecasters.

The emphasis of the study may seem to dwell too much on the value of the NWP 500 mb forecasts. There is no doubt that the Model guidance can be very much in error as shown by the number of cases when the E-W FC Scores were less than .50. In fact most models had scores of .50 or less 40 % of the time for Day 6. While forecasters would like to be specific for the extended period, prudence would tend to tilt the extended forecasts towards generalities and climatology more and more as the forecast time lengthens.

The author does not intend to continue the study of NWP 500 mb gradients in this manner. He would entertain suggestions and help in modifying procedures used in the study so that they could be automated. It is essential that a comparative verification of extended numerical forecasts be conducted in parameters familiar to forecasters and encompassing smaller regional areas of concern to forecasters. It is believed that this paper produces a valid evaluation of the forecast 500 mb gradients, but the models are continually changing and a continuing study is needed. Other NWP parameters are available for extended forecast guidance and they too need to be verified.

References

ECMWF, http://www.ecmwf.int/pressroom/index.html.

Environment Canada, Canadian Meteorological Centre: *Performance of the Canadian Numerical Models*, http://www.cmc.ec.gc.ca/~cmsw/cmsw/revue/html/modeles/html/liste_199912_an.html.

Funk, Ted: The NWS Louisville Science and Technology Page, http://atmos.umd.edu/~wxchair/models.html.

La Rue, Jerrold A.: Model comparisons for 60 hours to 6 days, *NWA Electronic Journal of Operational Meteorology*, ttp://www.nwas.org/ej/larue/MODEL%20COMPARISONS_newnew.htm.

NCEP, Environmental Modeling Branch, http://www.emc.ncep.noaa.gov/.

NCEP, Global Modeling Branch, *Model performance statistics (2001)*, http://sgi62.wwb.noaa.gov:8080/STATS/STATS.html.

NCEP, Hydrometeorological Prediction Center: Model Diagnostics, http://www.hpc.ncep.noaa.gov/html/model2.html#links.

Appendix 1. Mathematical Derivations

On a hypothetical grid there is a point 1 (P1) and an adjacent point 2 (P2). (For the E-W points, P2 is west of P!, and for the N-S points, P2 is south of P1.)

A forecast has a 500 mb height at P1 of F1 and at P2 a height of F2, both in dam.

The computation for two adjacent points:

1. The forecast gradient, Gf, is the difference between F1 and F2 or F1 - F2. (If F1 < F2, then Gf is negative).

The observed 500 mb height at P1 is O1 and at P2 a height of O2 , both in dam.

2. The observed gradient, Go, is the difference between O1 and O2 or O1 - O2. (If O1 < O2, then Go is negative).

3 The forecast gradient that is correct, Gc , requires the sign of Gf = the sign of Go; then Gc = the smaller of the (abs)Gf or (abs)Go. If the sign of Gf is not the same as the sign of Go, or if either = 0, then Gc = 0. (Note that Gc is always positive.)

The computations for the E-W or the N-S segments of the grid:

4. The Forecast Correct compared to Go, FC_o, is the sum Gc / sum (abs)Go, computed separately for the E-W and the N-S direction.

5. The Forecast Correct compared to Gf, FC_f, is the sum Gc / sum (abs)Gf, computed separately for the E-W and the N-S direction.

For the forecast, the sums of Gc , (abs)Go and (abs)Gf in either the E-W or N-S directions for the entire grid are used.

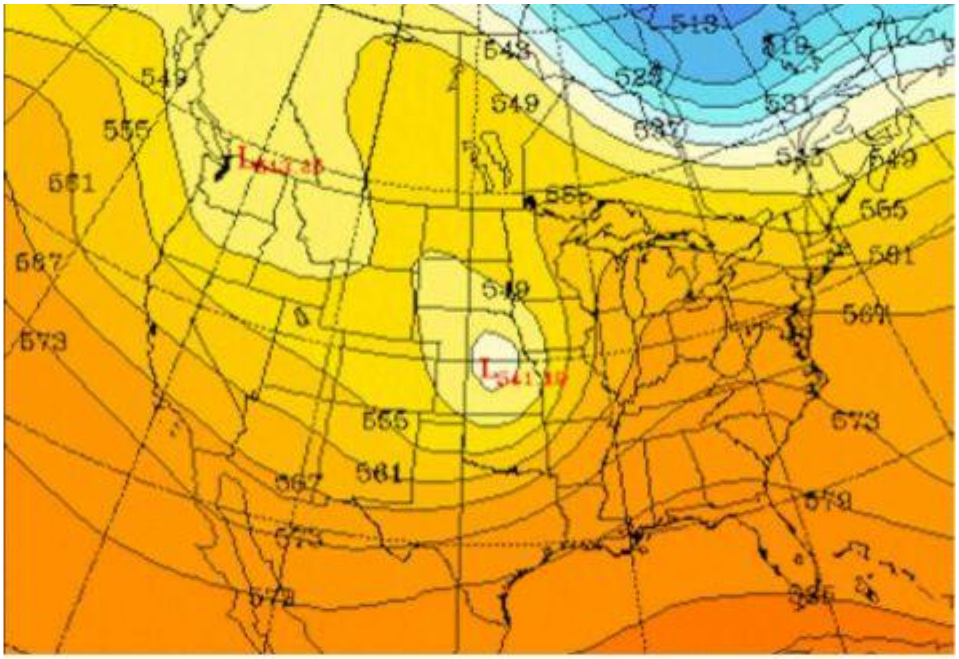
For the study, the sums of Gc , (abs)Go and (abs)Gf in either the E-W or N-S directions for all forecasts are used.

The $FC_{o \text{ or } f}$ is the S 1 Skill Score with the following differences: The S 1 Score would divide Gc by the larger of the sum (abs)Go or the sum (abs)Gf and multiply the result by 100 for each forecast. In the study, the terms were summated for the 100 cases for each model and for each forecast Day. The S 1 Score would compare the sum of Gc with the larger of the sum (abs)Go or the sum (abs)Gf for each forecast with the summation and averaging of those scores being the final S 1 Score for the study. This would allow a comparison of the models, but the value would not have any real meaning. Using FC_f and FC_o provides a measure of about how much of the forecast gradient or the ensuing observed gradient was correct.

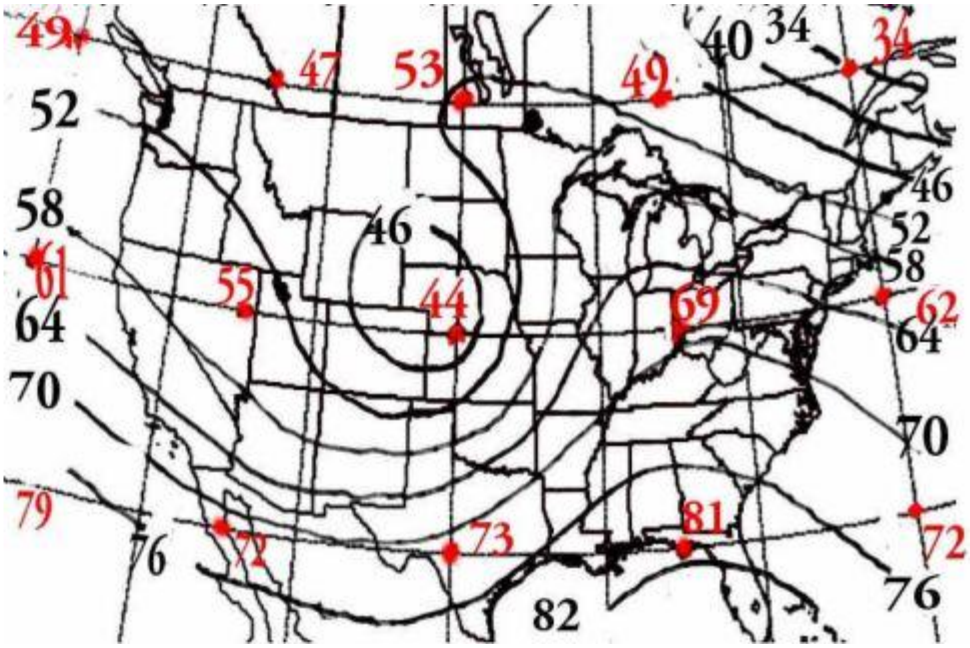
Appendix 2. Grid System

The rather course grid system in this study used intersections at Latitudes 50, 40 and 30 North and Longitudes 70, 85, 100, 115 and 130 West. Heights at 500 mb were selected for each point on the grid and the height difference gave the gradients between the points. The East-West distance between grid points at 50 North Latitude is about 655 miles, at 40 North it is about 793 miles and at 30 North about 896 miles. The difference in the span of the grid varies from about 2600 miles in the north to about 3600 miles in the south. Purely subjective reasoning suggests that wave lengths are longer with fewer small perturbations in the south than in the north. The difference in grid length distances might tend to increase the weight of southern gradients, probably improving overall scores.

The real question is: "Does the grid allow a decent evaluation of the 500 mb height field?". As a test, the initial 500 mb heights for 1200 UTC January 14, 2001 entered in the spread sheet for the study were plotted on the grid. That particular date began an interesting weather cycle and a series of maps had been saved so that the initial 500 mb MRF analysis was available. Without referring to the MRF analysis, the heights at the grid points were analyzed. Both the MRF analysis and the grid analysis are displayed for comparison. The NOAA ARL analysis was saved because it has only the 500 mb analysis and is relatively uncluttered. One draw-back of the ARL analysis is that the height contours are not necessarily at standard levels, and on this particular chart are at usual intermediate heights. Since the MRF analysis was not viewed until after the grid analysis was completed, it's analysis is at standard heights. The grid analysis was not changed to emulate the non-standard contours as that might have influenced the final grid analysis.



The above analysis is MRF from NOAA's ARL for 1200 UTC 1/14/01. Note contour heights are non-standard.



The above analysis is also for 1200 UTC 1/14/01 using the grid point values (in red).

The FC$_{\text{o or f}}$ Scores for the above grid analysis would be 1.00. The greatest error in the analysis is being as much as 7 dam too high to the Southeast of the upper low in Nebraska. The plus error of 2 to 7 dam encompasses Iowa, Eastern Nebraska, Western Illinois, Northwest Missouri, Eastern Kansas, Central Arkansas all of Oklahoma and extreme Northern Texas. Intermediate grid points at 35 North and at 92 1/2 West would have about eliminated this error. A 4 to 5 dam deficient error in the grid analysis exists over Oregon and Washington which completely missed a weak short wave along the West Coast. Again, intermediate grid points would have pretty much eliminated the error. A 3 dam error is at 50 North and 80 West which de-emphasizes a short wave that might affect Maine only. The grid initial analysis is correct over much of the area and does successfully reflect the major features. Halving the grid spacing would eliminate about all the errors but would increase the grid points from 15 to 45; thus the need for automation.

ACKNOWLEDGEMENTS

The author wishes to thank Ron Murphy for his valued comments and suggestions in reviewing the paper. Thanks also goes to Alan E. Gerard and Jeffrey Craven for their role in routing and evaluating the paper, to Stacy Bunin for Internet preparation and placing the paper on-line and to the Weather Analysis and Forecasting Committee of the NWA for sponsoring the Electronic Journal of Operational Meteorology.

BIOGRAPHY

Jerrold A. La Rue graduated in meteorology from UCLA. He entered the Weather Bureau at Peoria, IL in 1951, was transferred to Huron, SD in 1953, and to Buffalo, NY in 1955. In 1957 he moved to the National Meteorological Center (NMC) at Suitland, MD. where he was involved in analyzing and forecasting all standard level upper air charts as well as surface charts. In 1959 he was selected along with Vern Bohl to investigate, develop and adapt procedures for Quantitative Precipitation Forecasting (QPF), using NMC products. The QPF Unit was established in 1961 with Mr. La Rue and Russ Younkin as the original forecasters. In 1967 he became Chief of the Surface Analysis Branch at NMC. He was appointed Meteorologist in Charge of the Washington, DC Forecast Office in 1970. At that time and during his tenure there, the WSFO was engaged in about every forecasting task undertaken by the NWS. He has been a member of he American Meteorological Society since1953. He was one of the founders of the National Weather Association in 1975, it's first President and was Executive Director for five years. He retired from the National Weather Service in 1980 but not to a sedentary life style..

His E-Mail address is mailto:jerlarue@sonic.net