INTERACTIVE FORECAST PREPARATION POSITION PAPER 4

Quality of Model-based Initial Forecasts: Comparisons to Observations

Stuart Wier

Forecast Systems Laboratory

Boulder, Colorado

September 2, 1997; March 27, 1998

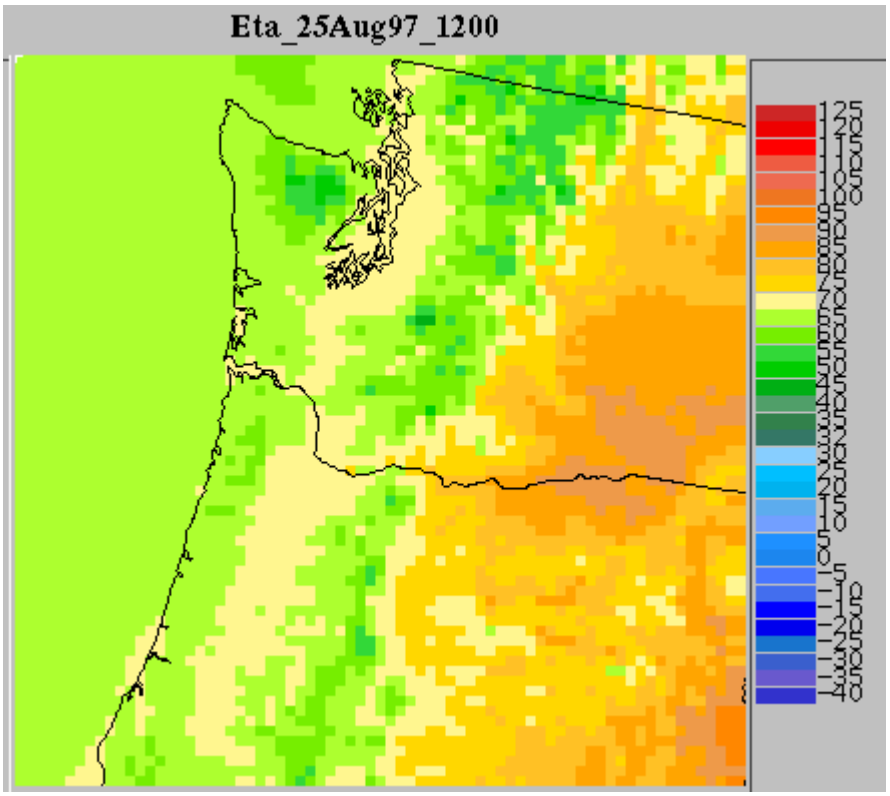
SUMMARY

Surface weather values used in public forecasts were derived from NCEP's numerical atmospheric models output. These forecasted values were compared to observations. The measures of fit indicate that in most cases this model-based initial forecast technique from NCEP models is acceptable as useful guidance for forecast preparation. Surface weather made from the Eta models and RUC are clearly the best guides to coming weather among the models tested.

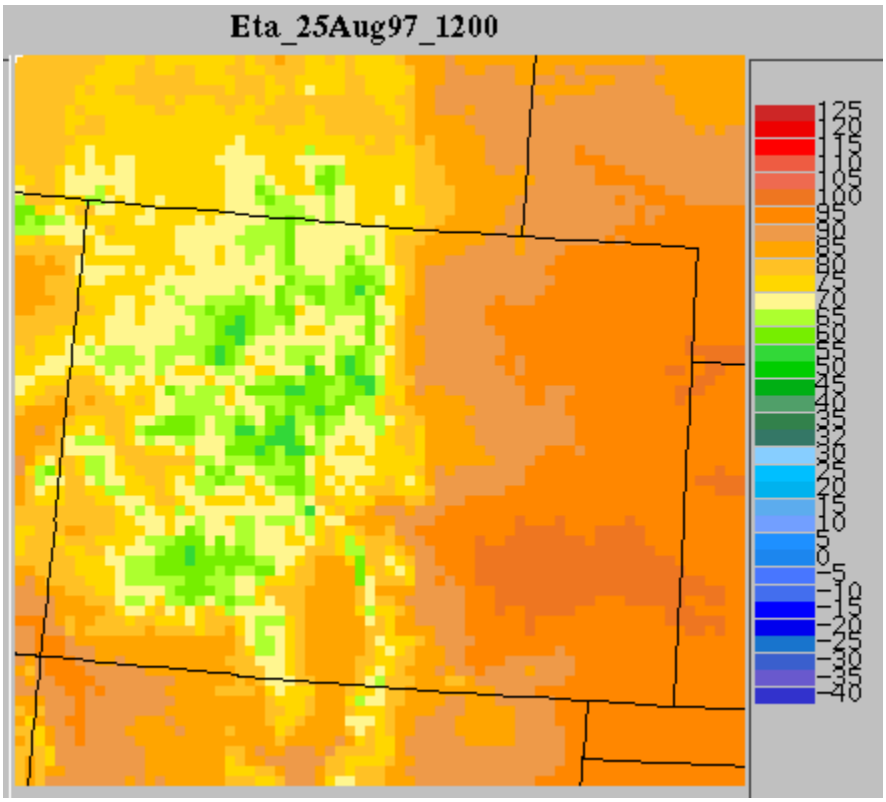
Derivation of initial surface forecasts from NCEP numerical atmospheric models is described in IFP Position Paper 3. The initial surface forecast values made by this process will be used as a starting point in the NWS Interactive Forecast Preparation system. This approach conveys high quality model output, in the form of variables actually used in forecasts, directly to grids that can go into or even make forecasts.

Graphic examples of the forecasts are available. Here are the daily maximum temperatures for August 25, 1997 for:

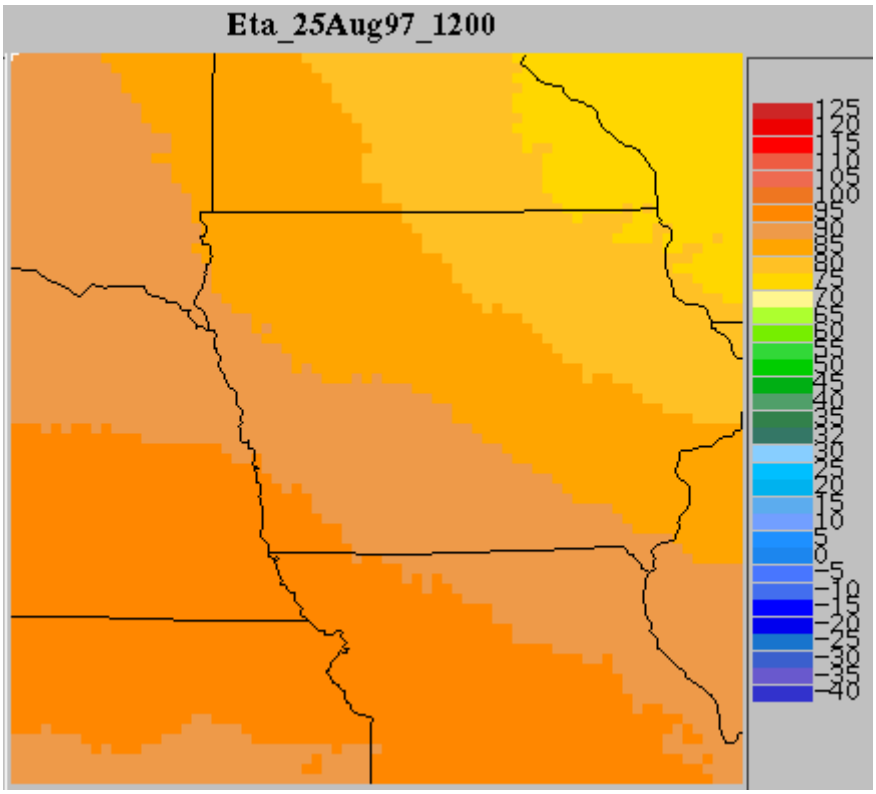
Oregon and Washington



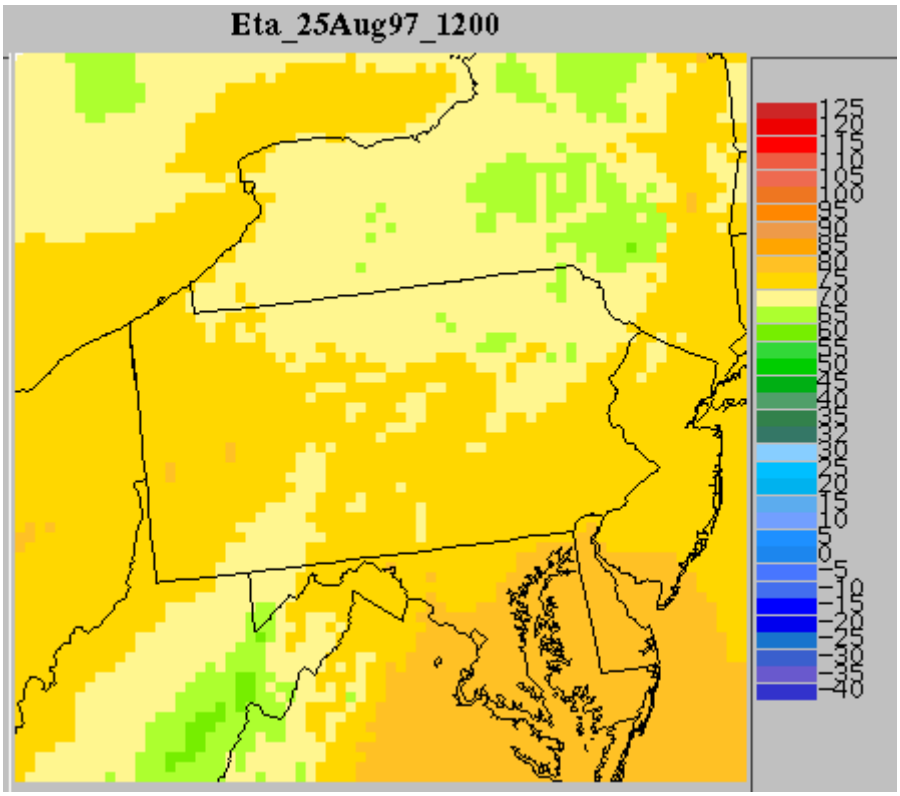
Colorado



Iowa



Washington D.C. to New York



These images demonstrate the high precision and detail possible from current atmospheric models. Note the high spatial resolution, and the correct variations in temperature expected for elevation as well as for atmospheric conditions and proximity to water. The remaining question is whether these automatic model-based surface forecasts are accurate.

Tests of forecast quality start by computing forecasted weather values at the location of observation stations. Differences (forecast value - observation) are computed from forecasted values and observations for the same locations and times. The differences are used to compute measures of fit of the forecasts to actual conditions, both for each individual location, and as overall averages. The overall results are reported here.

Data

Presently we use the NCEP AVN, Eta, MesoEta, NGM and RUC models. The output used is from the files transmitted to AWIPS offices via the SBN. (Later in 1998 the Eta (48 km) and MesoEta (29 km) model runs will merge into one model with a spatial resolution intermediate between the current Eta and MesoEta. RUC will be superceded by the improved RUC-2.)

Forecast values are computed at 109 locations in the lower 48 U.S. states. Metar observations from the same locations are stored for comparison. The test locations are scattered pretty evenly

across the lower 48 states and range in elevation from near sea level to 2900 meters. A list of the stations used is available from the author.

Weather Elements Tested

Presently model-based initialization forecasts the following weather elements for comparison to the metar observations:

- Temperature (every 3 or 6 hours)
- Max temp (in some cases)
- Dewpoint
- Wind Speed
- Wind Gust Speed
- Wind Direction
- Total Precipitation amount (liquid) (3 or 6 hours)
- Total cloud coverage (percent)
- Ceiling - height of base above ground level
- Visibility (results not reported here)

Time Periods covered by the Models, and Model Grid Resolution

Table 1 shows time periods covered by each kind of model, forecast time step or interval between forecasts, and model grid resolution.

RUC actually uses a 1-hour time step but only every 3rd hour is presently being transmitted via the SBN, and that is what was tested.

Table 1: Model Time Coverage and Grid Resolution

Model	Time Coverage	Time Step	Model's Grid Resolution
AVN	0 - 72 hours	6 hours	80 km
Eta	0 - 48 hours	6 hours	48 km; 32 km (1998)
MesoEta	0 - 33 hours	3 hours	29 km
NGM	0 - 48 hours	6 hours	80 km
RUC	0 - 12 hours	3 hours	60 km

Note that all the Eta data was distributed via the SBN on AWIPS CONUS 211 grids, which have an 80 km resolution.

Results of Tests.
Overall Results

Overall results, including measures of fit of the forecasts to observations, for all observation stations, times of day, and forecast projections, are shown in Tables 2 through 5.

The table entries are described as follows.

No. of Cases is how many forecast-observation differences were used. Some parameters have far fewer reproted observations than others. For example, min temp is reported much less ofter than max temp, which is of course reported less often than hourly temperature.

Ave. Algebraic Error or Bias is the average of all forecast minus observation differences (bias), including sign; it is the average bias between all forecasts and observations. It can be positive or negative. It indicates a systmatic shift between forecasts and observations. Forecasts with zero average bias can still have significant errors.

Ave. Absolute Error is the average SIZE of all forecast-observation differences, regardless of sign; it gives an indication of the typical size of the error of forecast from actual conditions.

RMS Error or Root Mean Square error is an indication of error which includes both the effect of bias and scatter or error size. It is usually the largest error of the measures used to test fit to observations.

Table 2a: Fit of Eta-48-based surface Forecast values to Observations.
May, June, July 1997 (at 0, 6, 12 and 18 Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature, degrees Fahrenheit				
6 to 48 hours	71658	-0.0015	3.9	5.10*
Dewpoint, degrees F				
6 to 48 hours	70739	0.16	4.0	5.61
Wind Direction, degrees				
6 to 48 hours	59508	7.0	43.4	60.6
Wind Speed, knots				
6 to 48 hours	63572	-1.0	3.0	4.00
Wind Gust Speed, knots				
6 to 48 hours	2945	-1.2	6.7	8.98
QPF, 6 hr, inches				
6 to 48 hours	22216	0.07	0.10	0.18
Total Cloud Cover, percent				
6 to 48 hours	35714	5.6	34.5	41.8

Percent of forecasted temperatures in error by > 10.0 degrees: 5.6 %
Percent of forecasted temperatures in error by < 5.0 degrees: 61.2 %
Percent of forecasted dew points in error by > 10.0 degrees: 7.3 %
Percent of forecasted dew points in error by < 5.0 degrees: 62.2 %
Percent of forecasted wind speeds in error by > 5 knots: 18.0 %
Percent of forecasted wind directions in error by > 45 degrees 34.3 %
Percent of 6-hour precip amount fcsts in error by > 1/4 inch: 10.5 %
Percent of 6-hour precip amount fcsts in error by > 1/2 inch: 1.9 %

Table 2b: Fit of Eta-32-based surface Forecast values to Observations.
1998 Feb 18 - March 6 (at 0, 6, 12 and 18 Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature, degrees Fahrenheit				
6 to 48 hours	24616	-1.0	3.7	4.8*
Dewpoint, degrees F				
6 to 48 hours	24084	2.5	4.3	5.5
Wind Direction, degrees				
6 to 48 hours	21274	3.3	36.2	53.0
Wind Speed, knots				
6 to 48 hours	22138	-1.0	3.1	4.2
Wind Gust Speed, knots				
6 to 48 hours	2642	2.3	7.6	9.6
QPF, 6 hr, inches				
6 to 48 hours	11744	0.05	0.05	0.11
Total Cloud Cover, percent				
6 to 48 hours	21269	-4.8	35.0	45.7
Ceiling height, meters above ground				
6 to 48 hours	3039	-62	265	1336

Percent of forecasted temperatures in error by > 10.0 degrees: 4.4 %
Percent of forecasted temperatures in error by < 5.0 degrees: 72.8 %
Percent of forecasted dew points in error by > 10.0 degrees: 6.4 %
Percent of forecasted dew points in error by < 5.0 degrees: 66.7 %
Percent of forecasted wind speeds in error by > 5 knots: 19.4 %
Percent of forecasted wind directions in error by > 45 degrees 26.6 %
Percent of 6-hour precip amount fcsts in error by > 1/4 inch: 4.8 %
Percent of 6-hour precip amount fcsts in error by > 1/2 inch: 1.3 %

Table 3: Fit of RUC-based surface Forecast values to Observations.
August 21 - Sept 2 1997 (at 0, 3, 6, 9, 12, 15, and 18 Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature, degrees Fahrenheit				
3 to 12 hours	41303	1.7	4.1	5.45*
Dewpoint, degrees F				
3 to 12 hours	41263	-1.6	4.1	5.54
Wind Direction, degrees				
3 to 12 hours	31698	6.0	44.5	61.3
Wind Speed, knots				
3 to 12 hours	34694	-0.8	2.6	3.56
Wind Gust Speed, knots				
3 to 12 hours	1792	-0.1	6.4	9.2
QPF, 3 hr, inches				
3 to 12 hours	13898	0.02	0.02	0.03
Total Cloud Cover, percent				
3 to 12 hours	31280	16.1	36.0	44.3
Ceiling, meters above ground level				
3 to 12 hours	526	-371.	669.	1053.

Percent of forecasted temperatures in error by > 10.0 degrees: 7.1 %
Percent of forecasted temperatures in error by < 5.0 degrees: 69.7 %
Percent of forecasted dew points in error by > 10.0 degrees: 7.5 %
Percent of forecasted dew points in error by < 5.0 degrees: 69.9 %
Percent of forecasted wind speeds in error by > 5 knots: 12.7 %
Percent of forecasted wind directions in error by > 45 degrees: 35.7 %
Percent of 3-hour precip amount fcsts in error by > 1/4 inch: 0.2 %
Percent of 3-hour precip amount fcsts in error by > 1/2 inch: 0.0 %

*Note that the forecasted temperatures verified here have an average RMS error of 5.10 for Eta (6 to 48 hours) in the warm season, 4.8 for Eta in the cool season, and 5.45 for RUC (3 to 12 hours). These are averages for all 109 stations across the U.S. Compare this to the *best* average RMS errors of 5.90 for TMAX and 5.56 for TMIN reported in a study of pubic media forecasts (CH9, days 1 and 2), "A Study of Five-day Weather Forecasts from Denver TV Stations and Newspapers," Brooks Martner and Marcia Politovich, pages J 53 to J 60, in *the Proceedings of the 16th Conference on Weather Analysis and Forecasting* AMS, Phoenix, (January 1998).

The Eta-derived surface temperature average RMS error at Broomfield Colorado (KBJC) near Denver was 5.45 degrees Fahrenheit.

While these two results cannot be directly compared (they cover different time periods, and one test was for max and min temps and the other for temps at several hours during the day), this result nevertheless strongly supports the view that actual surface temperatures generated from numerical atmosphere models are a very good source of forecast guidance.

Table 4a: Fit of AVN-based surface forecast values to observations.
Warm Season: May 1997 (at 0, 6, 12, 18Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature (degrees Fahrenheit)				
6 to 72 hours	25361	-2.5	4.9	6.19
Dewpoint (degrees F)				
6 to 72 hours	24596	-0.8	5.2	6.89
Wind Direction, degrees				
6 to 72 hours	21808	11.4	48.1	66.0
Wind Speed, knots				
6 to 72 hours	23080	4.1	5.7	7.06
QPF, 6 hr, inches				
6 to 72 hours	9438	0.04	0.08	

Percent of forecasted temperatures in error by > 10.0 degrees: 9.8 %
Percent of forecasted temperatures in error by < 5.0 degrees: 58.0 %
Percent of forecasted dew points in error by > 10.0 degrees: 13.5 %
Percent of forecasted dew points in error by < 5.0 degrees: 59.8 %
Percent of forecasted wind speeds in error by > 5 knots: 48.6 %
Percent of forecasted wind directions in error by > 45 degrees: 38.4 %
Percent of 3-hour precip amount fcsts in error by > 1/4 inch: 5.0 %
Percent of 3-hour precip amount fcsts in error by > 1/2 inch: 1.3 %

Table 4b: Fit of AVN-based surface forecast values to observations.
Cold Season: Jan 28 - March 13 1998 (at 0, 6, 12, 18Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature (degrees Fahrenheit)				
6 to 72 hours	49880	.30	4.6	6.03
Daily Max temp, (degrees Fahrenheit)				
6 to 72 hours	5619	-2.58	5.46	7.28
Dewpoint (degrees Fahrenheit)				
6 to 72 hours	49458	1.22	5.66	7.35
Wind Direction, degrees				
6 to 72 hours	43040	13.5	47.0	64.7
Wind Speed, knots				
6 to 72 hours	44768	5.3	6.4	8.15
QPF, 6 hr, inches				
6 to 72 hours	15865	0.08	0.08	0.15
Sky cloud coverage, total, percent				
6 to 72 hours	42320	12.7	36.5	50.0
Ceiling height, meters above ground				
6 to 72 hours	8664	-672	786	1650

Percent of forecasted temperatures in error by > 10.0 degrees: 9.3 %
Percent of forecasted temperatures in error by < 5.0 degrees: 63.3 %
Percent of forecasted dew points in error by > 10.0 degrees: 15 %
Percent of forecasted dew points in error by < 5.0 degrees: 55 %
Percent of forecasted wind speeds in error by > 5 knots: 53.1 %
Percent of forecasted wind directions in error by > 45 degrees: 37.6 %
Percent of 6-hour precip amount fcsts in error by > 1/4 inch: 9.3 %
Percent of 6-hour precip amount fcsts in error by > 1/2 inch: 2.4 %

Table 5a: Fit of NGM-based surface forecast values to observations.
May 1997 (at 0, 6, 12, 18Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature (degrees Fahrenheit)				
6 to 72 hours	21017	0.6	5.4	6.93
Dewpoint (degrees F)				
6 to 72 hours	20373	-4.7	5.9	7.36
Wind Direction, degrees				
6 to 72 hours	18087	13.0	46.0	63.5
Wind Speed, knots				
6 to 72 hours	19175	4.4	6.4	7.95
QPF, 6 hr, inches				
6 to 72 hours	4548	0.05	0.10	

Percent of forecasted temperatures in error by > 10.0 degrees: 13.7 %
Percent of forecasted temperatures in error by < 5.0 degrees: 55.5 %
Percent of forecasted dew points in error by > 10.0 degrees: 13.7 %
Percent of forecasted dew points in error by < 5.0 degrees: 55.5 %
Percent of forecasted wind speeds in error by > 5 knots: 51.5 %
Percent of forecasted wind directions in error by > 45 degrees: 36.5 %
Percent of 3-hour precip amount fcsts in error by > 1/4 inch: 9.5 %
Percent of 3-hour precip amount fcsts in error by > 1/2 inch: 2.1 %

Table 5b: Fit of NGM-based surface forecast values to observations.
Jan 28 - March 14 1998 (at 0, 6, 12, 18Z)

Value	No. of Cases	Ave. Algebraic Error	Ave. Absolute Error	Ave. RMS Error
Temperature (degrees Fahrenheit)				
6 to 72 hours	45171	0.5	4.8	6.3
Max temp (degrees fahrenheit)				
6 to 72 hours	5387	-2.4	5.4	7.0
Dewpoint (degrees F)				
6 to 72 hours	44805	0.9	4.6	6.2
Wind Direction, degrees				
6 to 72 hours	39244	16.1	44.8	62.1
Wind Speed, knots				
6 to 72 hours	40771	5.5	6.9	8.5
Wind Gust Speed, knots				
6 to 72 hours	5302	15.4	16.5	20.2
QPF, 6 hr, inches				
6 to 72 hours	7017	0.11	0.11	0.18
Sky cloud coverage, total, percent				
6 to 72 hours	42567	24.6	36.8	51.1
Ceiling height, meters above ground				
6 to 72 hours	6797	-459	634	1335

Percent of forecasted temperatures in error by > 10.0 degrees: 10.4 %
Percent of forecasted temperatures in error by < 5.0 degrees: 61.6 %
Percent of forecasted dew points in error by > 10.0 degrees: 9.4 %
Percent of forecasted dew points in error by < 5.0 degrees: 64.2 %
Percent of forecasted wind speeds in error by > 5 knots: 56 %
Percent of forecasted wind directions in error by > 45 degrees: 35.5 %
Percent of 3-hour precip amount fcsts in error by > 1/4 inch: 11.3 %
Percent of 3-hour precip amount fcsts in error by > 1/2 inch: 2.9 %

NOTE: the measures of QPF quality used here are of value only for intercomparsion of results in this set of tests. Thorough assessments of model QPF forecasts have been done by others. See for example "Verification of RUC-2 and Eta Model Precipitation Forecast" by Barry Schwartz and Stanley Benjamin (pages J 103 to J105), and "Comparison of QPF by the 48- and by the 29-KM Eta Model" by Fedor Mesinger (pages J 106 to J 107), in *the Proceedings of the 16th Conference on Weather Analysis and Forecasting* AMS, Phoenix, (January 1998).

Quality of Fit Depending on Time of Day, and Projection Time.

The measures of fit were separated by time of day, and projection time, averaging results from all stations. The intent was to see if there were any systematic shifts in forecast quality.

Quality of Fit Depending on Time of Day.

The dependency on time of day indicates if the model fit to future weather has some kind of diurnal variation. Some cases of clear dependency on time of day are present, as is shown in Table 6 for NGM summer temperature biases. These may be indications of places for improvement in the models. NGM appears to have a thermal lag at the surface: it is too warm in the middle of the night and too cool in the afternoon. The same tendenccy is visible in the NCEP RUC output, though not so large. Early indications are that this the surface thermal lag is removed in the RUC 2 model which will be operational at NCEP in March 1998.

Table 6: Forecast surface temperature bias depending on time of day (degrees Fahrenheit)
May to September 1997

Time	AVN	Eta	NGM	RUC
0 Z	-3.1	-0.02	-0.56	2.8
3 Z				4.2
6 Z	-2.1	-0.03	4.68	4.0
9 Z				3.1
12 Z	-2.2	-0.02	1.87	2.2
15 Z				-1.0
18 Z	-2.7	-0.40	-3.49	0.1
21 Z				0.3

Table 7: Forecast surface temperature bias depending on time of day (degrees Fahrenheit)
January - March 1998

Time	AVN	Eta	NGM	RUC
0 Z	-.22	-2.09	1.08	
3 Z				.
6 Z	.31	-0.20	-0.41	
9 Z				.
12 Z	.77	-0.07	1.09	
15 Z				
18 Z	.32	-1.75	1.08	
21 Z				

Quality of Fit Depending on Projection Time.

Separation by forecast projection time shows whether the forecasts become worse as the forecast is farther in the future. An example is shown in Table 8, for absolute error (error size) of forecasted surface temperature.

There is some degradation in forecast quality in temperature values derived from NGM (0 to 48 hour), and little degradation in forecast quality from AVN (6 to 72 hours) and RUC (1 to 12 hours). The Eta model is notable in having no sign of increasing errors as forecast time increases for any of the values made from it (6 to 48 hours).

Table 8: Forecast surface temperature absolute error (average error size), depending on forecast projection time (hours; degrees C) Summer 1997

Proj .	AVN	Eta	NGM	RUC
6	2.57	2.43	2.64	2.21
12	2.41	2.19	2.78	2.69
18	2.57	2.08	2.62	
24	2.54	2.21	3.39	
30	2.57	2.02	3.00	
36	2.52	2.18	3.22	
42	2.67	2.14	2.94	
48	2.60	2.16	3.48	
54	2.83			
60	2.65			
66	3.05			
72	3.31			

Comparison to MOS Quality

For purposes of comparison some similar tests of NGM MOS are shown here. These results are based on observations at 96 stations across the U.S.

Table 9: Fit of NGM MOS surface forecast values to Observations.
0Z run, October 1993 - March 1994

Value	No. of Cases	Mean Algebraic Error	Mean Absolute Error	% Errors > 10 F
Today's Max	15896	-0.2	3.3	2.1
Tonight's Min	15613	0.0	3.9	4.5
Tomorrow's Max	15871	-0.3	4.1	5.4
Tomorrow's Min	15544	-0.1	4.7	7.7
Wind Direction, degrees				
12 hr forecast	4661		21.0	
18 hr forecast	7580		23.0	
24 hr forecast	5748		26.0	
Wind Speed, knots				
12 hr forecast	4697	2.2	3.5	
18 hr forecast	7617	1.4	3.4	
24 hr forecast	5790	1.9	3.6	

From TDL Office Note 95-2, 1995.

As well as accuracy of forecasts at individual observations stations, where model-based initial values do about as well as MOS, the model-based values also have the advantage of being available on high resolution grids, providing forecast guidance for virtually any location. Compare this model-based grid of Max Temp to an equivalent grid of analysed MOS values for the same time and location. Analysed and imaged MOS provided by Ohio State

Eta-based Northern California Max T April 10

MOS Northern California Max T April 10

Summary

Forecasts of surface temperature, dewpoint, wind, precipitation, and other values were compared to observations at 109 stations across the lower 48 states of the U.S. The forecasts were automatically generated from NCEP numerical model output.

There is strong indication that model-based forecasts of surface temperature, dewpoint, relative humidity, wind, and precipitation amount are good enough for operational forecast guidance at any location in the models' domains, in the lower 48 U.S. states, and possibly in adjacent land and offshore water areas. The models also cover Alaska and Hawaii and this technique could be used there as well. The quality of forecasts has not been tested in those two states yet.

These forecasts, though tested at individual stations, may have similar validity at any other point in the models' domains. The stations used in these tests are only sample points for testing and otherwise play no part in the creation of the initial forecasts. Forecast values can be made for any location inside the models' domains.

Eta and RUC are clearly the models of choice for derivation of surface weather parameters in the projection period 3 to 48 hours. The early Eta and MesoEta models will be improved and merged in 1998, and RUC will be superceded by the improved RUC 2 model. These changes should further improve the accuracy of model-based initial surface forecast values.

A important advantage of this technique for making forecast guidance is that model improvements are instantly translated into improved surface forecast guidance. Likewise, improved algorithms (for example, for orographic or lake snow amount, visibility, clouds) can also be easily incorporated in processing. This is one of the best ways to rapidly make improvements from meteorological research directly help forecasters.

Examples of today's initial forecast guidance from Eta are displayed by the AFPS Graphical Forecast Viewer (the GFV) available on the EFT home page. The Viewer is in development and may not be available or have data at all times.