26

A STATISTICAL EVALUATION OF HYDROLOGIC FORECASTING ON THE MISSOURI RIVER FROM 1983 TO 2013

A. Juliann Meyer*
National Weather Service, Pleasant Hill, MO

Noreen O. Schwein
National Weather Service, Lenexa, KS (retired)

Lee W. Larson
National Weather Service, Prairie Village, KS (retired)

1. INTRODUCTION

The National Weather Service (NWS) has undergone a dramatic evolution in the field of hydrology since the Organic Act of 1890, which created the Weather Bureau. Early hydrologic forecasting to meet the responsibility of “gauging and reporting on rivers” was based mainly on empirical rules and personal knowledge of river basins. By the late 1930s, empirical rules gave way to the use of physical data such as rainfall and snow melt data, the antecedent precipitation index (API) models, the development and application of the unit hydrograph, and improved routing techniques. During the period from the mid-1940s to late 1970s, River Forecast Centers (RFCs) were formed to concentrate on the problem of river and flood forecasting and to develop and refine hydrologic forecasting procedures and techniques. The latter half of the 20th century and the first decade of the 21st century have been characterized by improved hydrologic models and the infusion of science and technology to provide improved river forecasts. Rapidly changing computer and systems technology, along with evolving conceptual hydrologic models and forecast procedures, posed considerable challenges and opportunities for RFC hydrologists.

The Missouri Basin River Forecast Center (MBRFC) located in Pleasant Hill, Missouri is responsible for providing river forecast services for the entire Missouri River basin and St. Mary River basin in Montana. The Missouri River is the longest river in the United States with a length of about 3767 km (2341 miles). The headwaters begin along the Rocky Mountains in Montana and the river flows east and south to its confluence with the Mississippi River near St. Louis, Missouri. The Missouri River drains about 1,372,694 km^2 (530,000 mi^2) in the heartland of America; covering all or part of ten states and part of Canada.

For the past three decades, MBRFC has archived river forecasts and observations for the Missouri River and using that data, has calculated forecast verification statistics to assess forecast accuracy. Verification of river forecasts is completed daily on a continuous basis for a number of individual locations (12 mainstem forecast points). A significant historical database has therefore been accumulated over the last 31 years for the mainstem river forecasts. This database of observed and forecast stages is utilized to calculate verification statistics and to identify and explain any observable long term forecasting trends (e.g., improvements over time).

This topic was addressed in an earlier paper (Larson and Schwein, 2002) for selected daily forecast locations on the Missouri and Mississippi rivers for the time period 1980 through 2000. This paper focuses on the Missouri River and updates the analysis through the year 2013. It also includes the topic of long term flows in the Missouri basin with regard to forecasting errors.

2. PROCEDURES AND DATA

River forecasts are generated daily for the Missouri River mainstem by MBRFC forecasters who also produce flood forecasts during high water. Each day, forecasters issue forecasts for twelve locations on the Missouri River valid 24, 48, and 72 hours into the future. Table 1 lists and Figure 1 shows the individual forecast points on the mainstem Missouri River.

This study uses the 24, 48, and 72-hour lead time forecasts, which have been issued on a daily basis over the entire archived period since 1983. From 1983 through 2000 only the routine morning forecasts were archived and available for this study. Those forecasts were issued in 24-hour time steps until the late 1990s. Starting in 2001 the RFC began archiving all forecasts that were issued throughout a given day and the time step of the forecasts changed from 24-hour time step to 6-hour time step. Figure 2 shows the sample size for each year of the study period.

---

* *Corresponding author address:* Juliann Meyer, NWS/Missouri Basin RFC, 1803 N 7 Hwy, Pleasant Hill MO 64080; e-mail: julie.meyer@noaa.gov

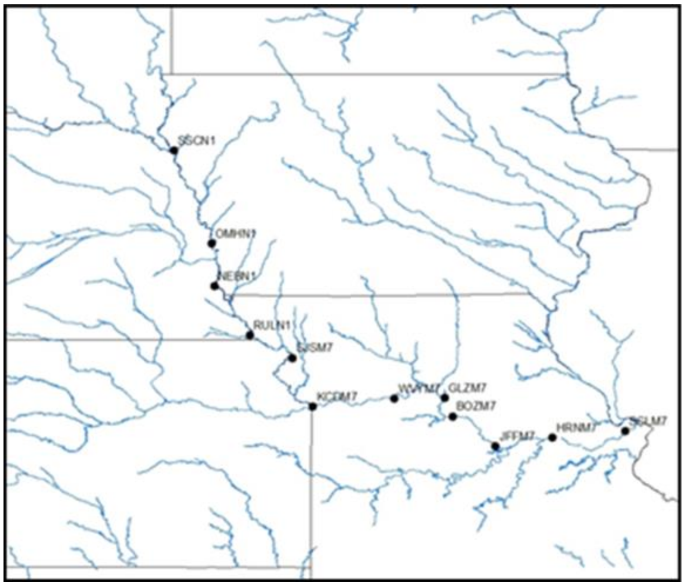

FIG. 1. Missouri River Gage Location Map

Missouri River Daily Forecast Locations
Sioux City, Iowa (SSCN1)
Omaha, Nebraska (OMHN1)
Nebraska City, Nebraska (NEBN1)
Rulo, Nebraska (RULN1)
Saint Joseph, Missouri (SJSM7)
Kansas City, Missouri (KCDM7)
Waverly, Missouri (WVYM7)
Glasgow, Missouri (GLZM7)
Boonville, Missouri (BOZM7)
Jefferson City, Missouri (JFFM7)
Hermann, Missouri (HRNM7)
Saint Charles, Missouri (SCLM7)

TABLE 1. Missouri River Daily Forecast Locations

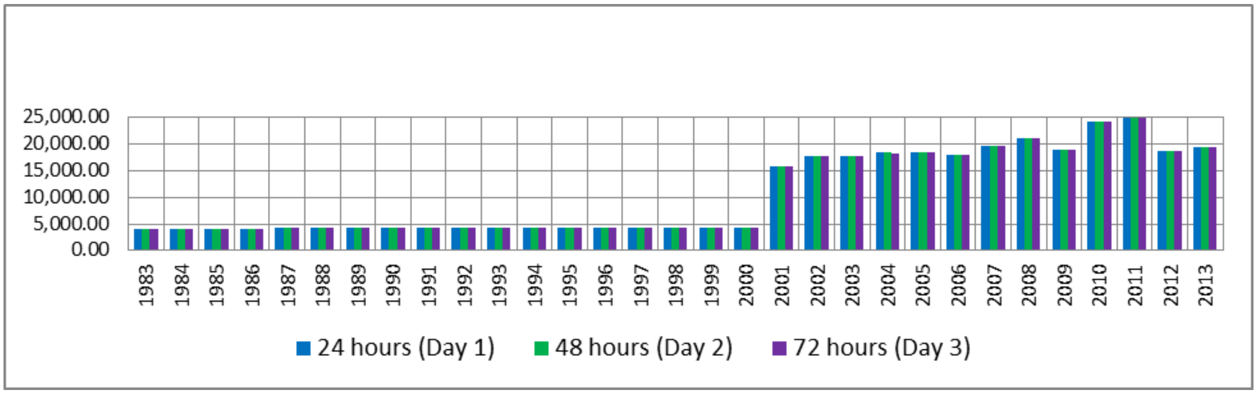

FIG. 2. Sample Size by Lead Time and by Year

Over the 1983-2013 time period encompassing approximately 990,000 individual forecasts, RFC staff has used the NWS interactive verification program (IVP) to calculate performance metrics. MS Excel 2010 was used to calculate a statistic not available in IVP. The IVP application was used to compare individual stage forecasts as described above, to the actual observed stage at the locations in question for the appropriate time period. For example, a 24-hour forecast for Hermann, Missouri was compared to the actual stage observed at Hermann 24 hours after the initializing time period (e.g., 1200 UTC). The 48-hour forecast was compared to the actual stage observed at Hermann 48 hours after the forecast initial time and so on. The paired data was used to calculate forecast error (forecast stage versus observed stage values). The forecast error values for each time period were then used to calculate yearly and all years combined statistics including mean absolute error (MAE), mean error (ME), and standard deviation (StdDev). It is reasonable to assume that, over time, forecaster skills and capabilities improve in hydrologic modeling, calibration, data handling, and precipitation forecasting for headwater and tributary rivers and thus, the mainstem forecasts also improve. The mainstem river forecasts can be regarded as a summary of all the forecasting capabilities and skills employed on the tributaries, which ultimately produce the flow and forecasts on the Missouri River. Therefore, the long-term statistics on the mainstem river forecasts are good indicators as to whether overall hydrologic forecasting skills are improving across the basin.

3. RESULTS

One of the most basic ways to look at forecast accuracy is with the use of a simple scatter plot. A forecast vs. observation scatter plot allows the user to see trends in the forecasts and corresponding observations. Figures 3-6 show scatter plots for the twelve Missouri River daily forecast locations. Figure 3 displays all the forecast-observed stage data pairs for the entire 72-hour lead time period. The scatter plots shown in Figures 4-6 show results for each individual lead time period, 24 hours, 48 hours and 72 hours, respectively. These scatter plots in all cases, whether looking at all lead time data combined or each lead time separately, show a slight under-forecasting bias. Figures 4-6 also show that as the forecast lead time increases, the spread around the "perfect forecast line" (45 degree yellow line) increases, as well. Looking at the accuracy error statistics MAE and ME will allow us to see how large the bias and errors are.

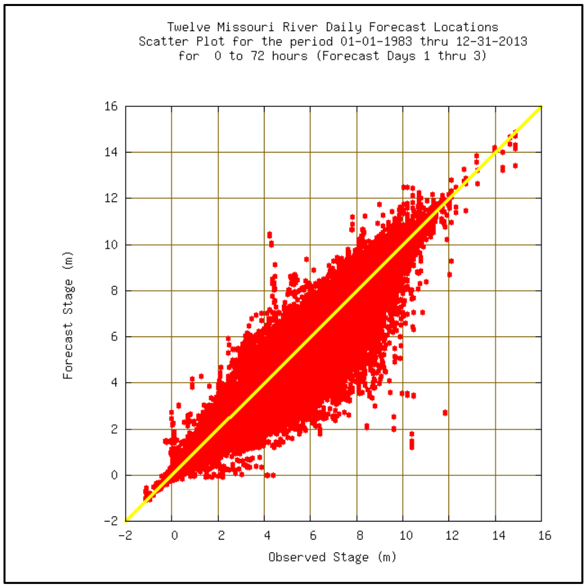

FIG. 3. Scatter Plot for twelve Missouri River daily forecast locations for entire lead time period 0 to 72 hours (forecast days 1 thru 3)

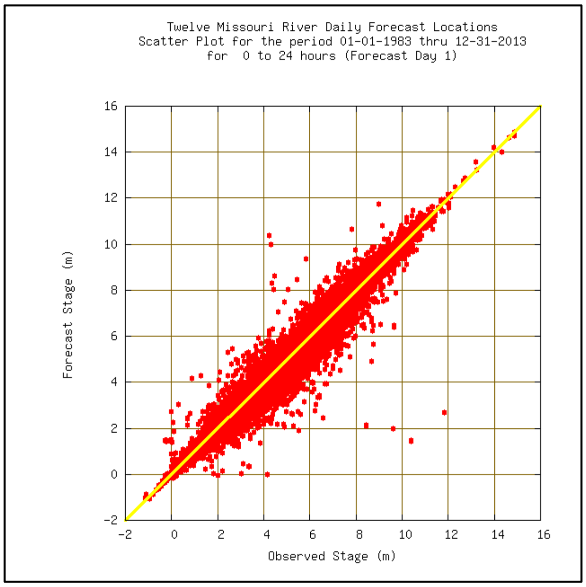

FIG 4. Scatter Plot for Twelve Missouri River Daily Forecast Locations for lead time period 0 to 24 hours (forecast day 1)

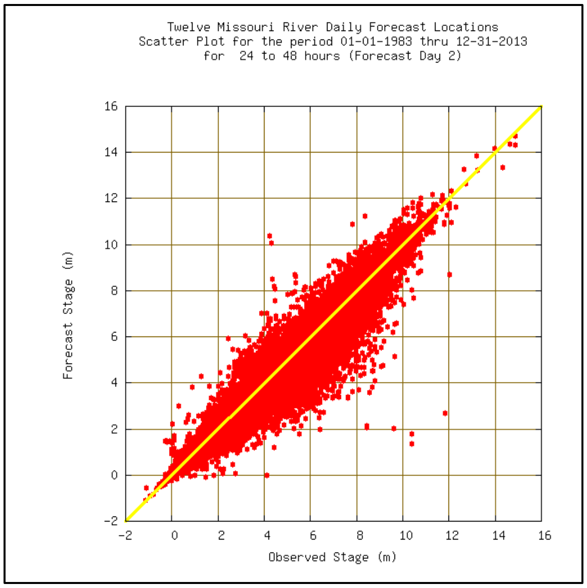

FIG. 5. Scatter Plot for twelve Missouri River daily forecast locations for lead time period 24 to 48 hours (forecast day 2)

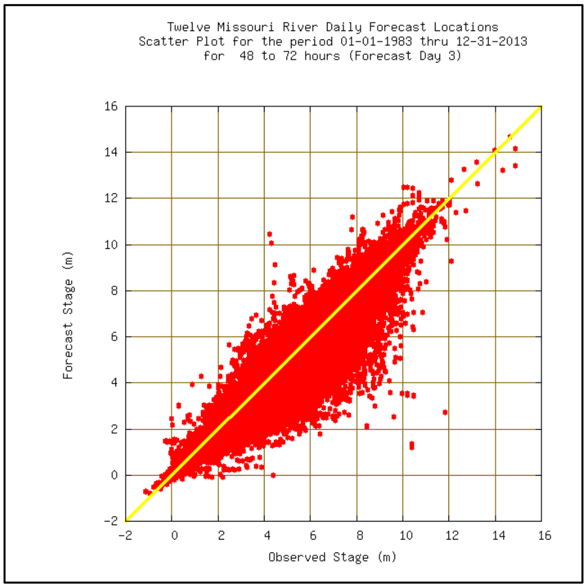

FIG. 6. Scatter Plot for twelve Missouri River daily forecast locations for lead time period 48 to 72 hours (forecast day 3)

Table 2 (based on Figures 8-10) shows a summary of average error statistics for the Missouri River over the period 1983 through 2013 (31 years) for all the 24, 48, and 72-hour forecasts issued each day. The MAE for the 24-hour forecast for the Missouri River is 0.07 meters (0.24 feet).

Average Forecast Statistics

Missouri River 1983-2013

Lead Time	MAE (meters)	MAE (feet)	ME (meters)	ME (feet)	StdDev (meters)	StdDev (feet)
24 hours	0.07	0.24	-0.01	-0.04	0.14	0.46
48 hours	0.14	0.45	-0.05	-0.15	0.26	0.84
72 hours	0.21	0.68	-0.10	-0.33	0.37	1,22

TABLE 2. Average forecast statistics for twelve Missouri River daily forecast locations for the period 1983-2013

As would be expected the MAE increases slightly when the forecast is extended to 48 hours into the future, MAE of 0.14 meters (0.43 feet), and then for 72 hours MAE of 0.21 meters (0.68 feet) indicating more variability and uncertainty as the forecast extends further in time.

Similarly, the standard deviation of the MAE forecast errors increases when going from a 24-hour (to a 48-hour and 72-hour lead time; 0.14 meters (0.46 feet), 0.26 meters (0.84 feet), and 0.37 meters (1.22 feet) respectively. Since this database of mainstem forecast statistics contains 31 years of daily forecast data, it is felt that the statistical values shown in Table 2 form a good baseline from which to judge future improvements in forecast techniques and procedures. The Larson and Schwein study, for the period of 1980-2000, showed a 24-hour MAE of 0.09 meters (0.30 feet), a 48-hour MAE of 0.16 meters (0.53 feet), and a 72-hour MAE of 0.23 meters (0.76 feet). Thus, an improvement in MAE of about 0.02 meters (0.06 feet), 0.02 meters (0.08 feet) and 0.02 meters (0.08 feet, respectively, can be seen with the addition of the last 13 years of forecast data, implying a notable improvement in forecast accuracy during that time.

Figure 7 shows a timeline of changes related to river modelling over the years and will be discussed in more detail in Section 4. While Figures 8-10 show MAE, ME and StdDev by year for 24, 48, and 72-hour lead times. The long term trends clearly show a continued reduction in the MAE value of all three lead times. The long term trend for MAE for 24-hour forecasts has trended downward from about 0.10 meters (0.30 feet) in 1983 to about 0.05 meters (0.15 feet) in 2013. For 48-hour forecast lead time, the MAE trends downward from about 0.16 meters (0.54 feet) in 1983 to 0.10 meters (0.32 feet) in 2013. For 72-hour forecast lead time, the trend is downward from about 0.26 meters (0.81 feet) in 1983 to about 0.15 meters (0.50 feet) in 2013. Figure 8 in combination with figure 7 also clearly shows that the implementation of improved hydrologic models, quantitative precipitation forecasts (QPF), use of quantitative radar precipitation estimates (MAPX) and model time steps going from a mix of time steps to 6 hours everywhere, relates directly to improved forecasts. The plot of ME, in Figure 9, shows a continuous trend towards smaller mean errors for 24, 48 and 72-hour lead times. It should be noted that average mean errors are all negative showing that forecasts tend to underestimate corresponding observed stages. Also, beginning about 1996, QPF was added to future forecasts and the mean errors began to tend toward zero, lessening some of the negative bias in the forecasts. Similarly the StdDev (Figure 10 shows the same downward trend in error over the years for 24, 48, and 72 hour lead times.

Figure 11 shows a plot of 24, 48 and 72-hour MAEs along with the U.S. Geological Survey's (USGS) annual mean discharge (cms) for the Missouri River at Hermann, Missouri. This location was chosen because the Missouri River flow at Hermann, Missouri can be considered as the point that integrates all flows in the basin since it is the furthest downstream forecast point and near the mouth of the river that backwater from the Mississippi River does not significantly impact. For this plot the MAE values were recomputed by water year* as the annual mean discharge values provided by the USGS are computed by water year. The plot shows a clear relationship between the magnitude of the flow and the magnitude of the MAE for most years. As flows increase, the MAEs also increase. The largest MAEs occurred in 1993 during the Great Midwest Flood. MAEs peaked at about 0.15 meters (0.49 feet) for 24 hours, 0.28 meters (0.93 feet) for 48 hours and 0.45 meters (1.46 feet) for 72 hour lead times during this monumental flood. For the more recent significant flood events in water years 2010 and 2011 one again sees an increase for the year in the MAE and discharge values. For 2010 MAEs of 0.07 meters (0.23 feet) for 24 hours, 0.16 meters (0.52 feet) for 48 hours, and 0.24 meters (0.78 feet) for 72 hours, while for 2011 MAEs of 0.05 meters (0.15 feet) for 24 hours, 0.09 meters (0.30 feet) for 48 hours and 0.13 meters (0.43 feet(for 72 hour lead times. While the 2011 had major to record flooding for six of the forecast locations mostly above Kansas City, MO, for the other six forecast locations the 2011 Flood was not a major flood event.

---

* Water year is defined as the 12-month period October 1, for any given year through September 30, of the following year. The water year is designated by the calendar year in which it ends and which includes 9 of the 12 months.

In 1993 the Missouri River contributed about 75 percent the flow in the Mississippi River at St. Louis, Missouri when it crested on August 1, 1993. In 2011 the peak discharge for the Missouri River at Hermann, MO occurred on May 27, 2011 and at that time the Missouri River was contributing about 45 percent of the flow in the Mississippi River at St. Louis, Missouri.

Figure 12 is a plot of MAE versus Hermann, Missouri annual mean discharges for 24 hours lead time forecasts. This plot shows how the MAE increases as discharge increases. Looking at the trend line, one can see the average MAE is about 0.03 meters (0.09 feet) for discharges of around 1700 cms (60,000 cfs) and increases to 0.15 meters (0.48 feet) for discharges around 4250 cms (150,000 cfs). Figures 13 and 14 show a similar trend of increasing MAE with increasing discharge. The three plots also show that as forecast lead time increases, so do MAEs. Figure 13 shows the MAEs versus the flow at Hermann for 48 hours lead time. The MAE for discharges of around 1700 cms (60,000 cfs) is 0.07 meters (0.23 feet) and increases to about 0.25 meters (0.83 feet) for discharges of 4250 cms (150,000 cfs). Figure 14 shows MAEs versus the annual mean discharge at Hermann for 72 hours lead time. The MAE for discharges of around 1700 cms (60,000 cfs) is 0.10 meters (0.32 feet) and rises to 0.38 meters (1.26 feet) for discharges of around 4250 cms (150,000 cfs). Table 3, below, summarizes these results and shows the MAE increase with lead time and with flow (values are approximate and are from Figures 12-14).

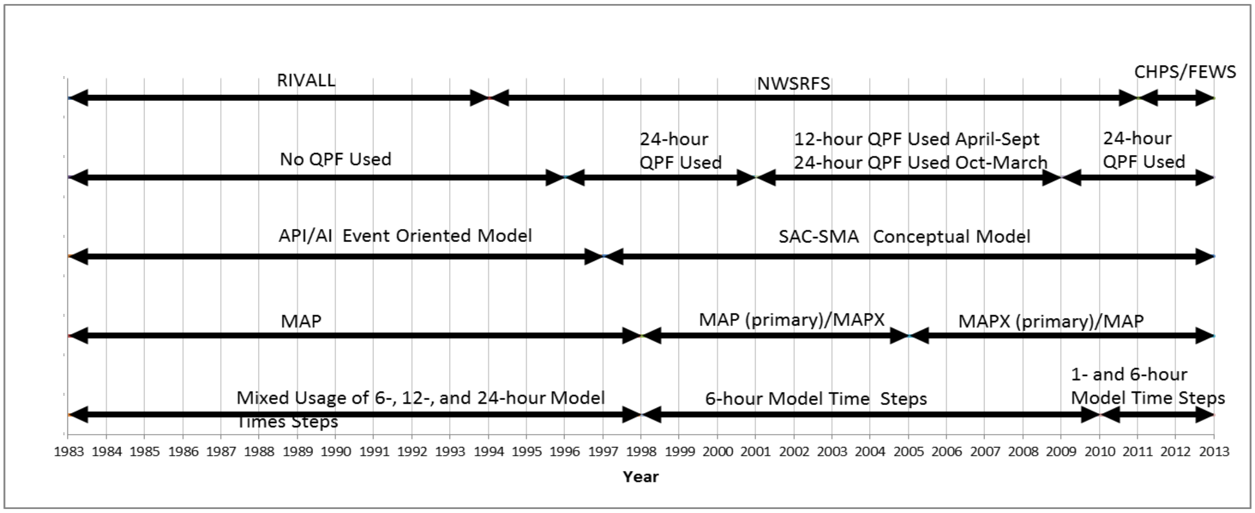

FIG. 7. Timeline of Changes

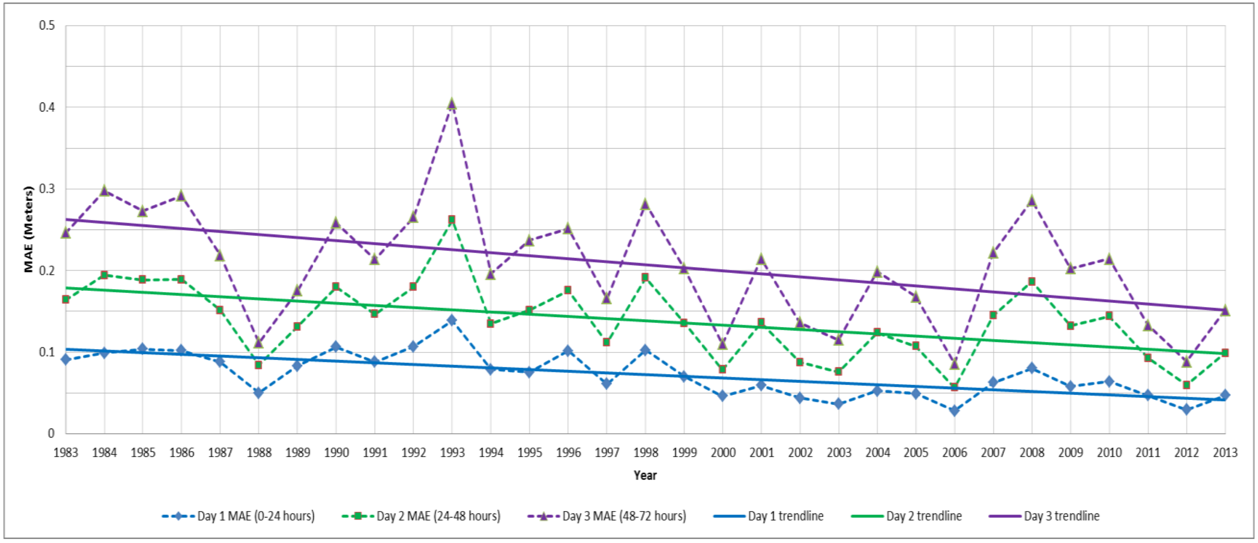

FIG. 8. Mean Absolute Error by Lead Time and by Year

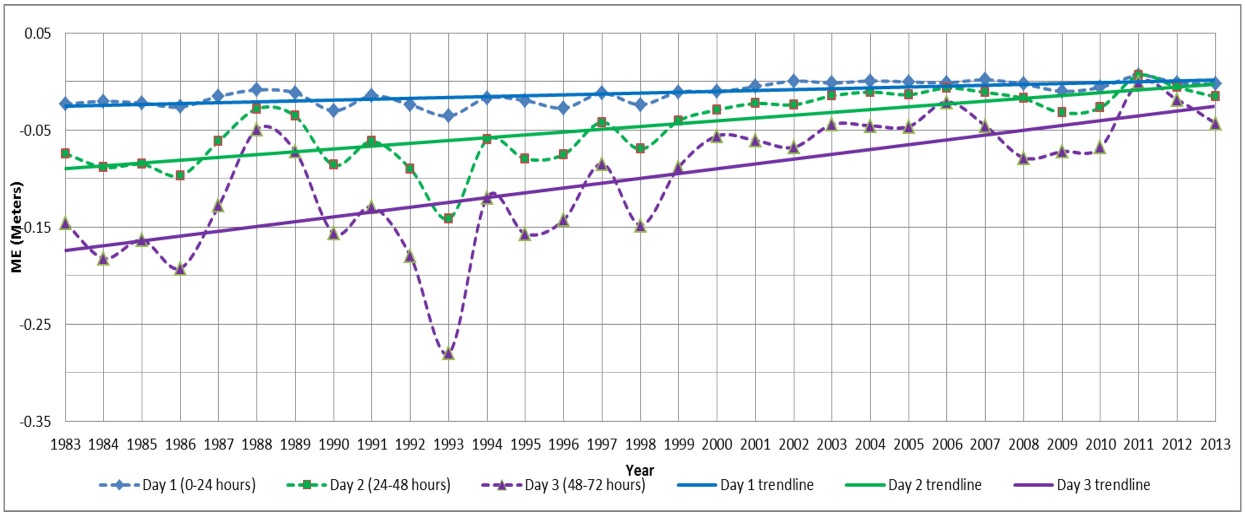

FIG. 9. Mean Error by Lead Time by Year

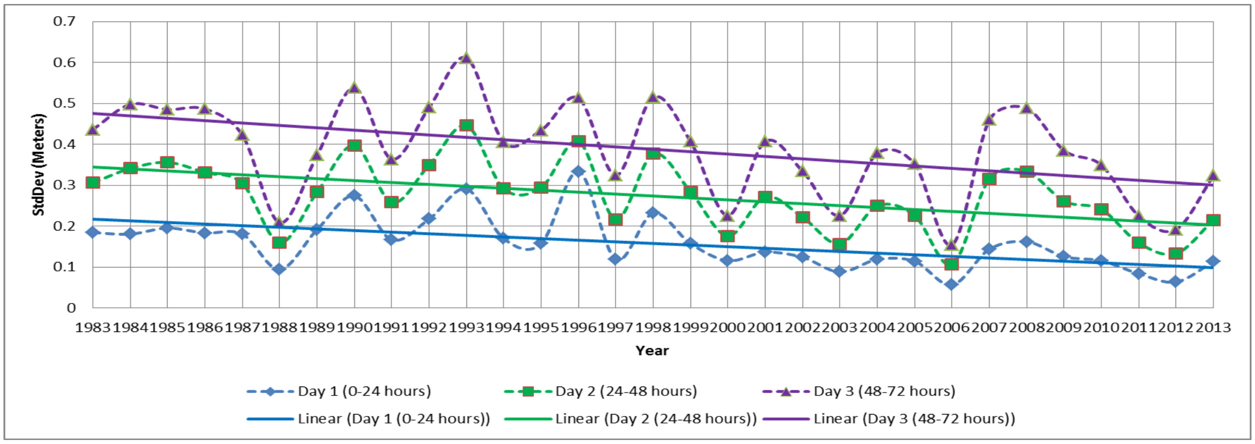

FIG. 10. Forecast Error Standard Deviation by Lead Time and by Year

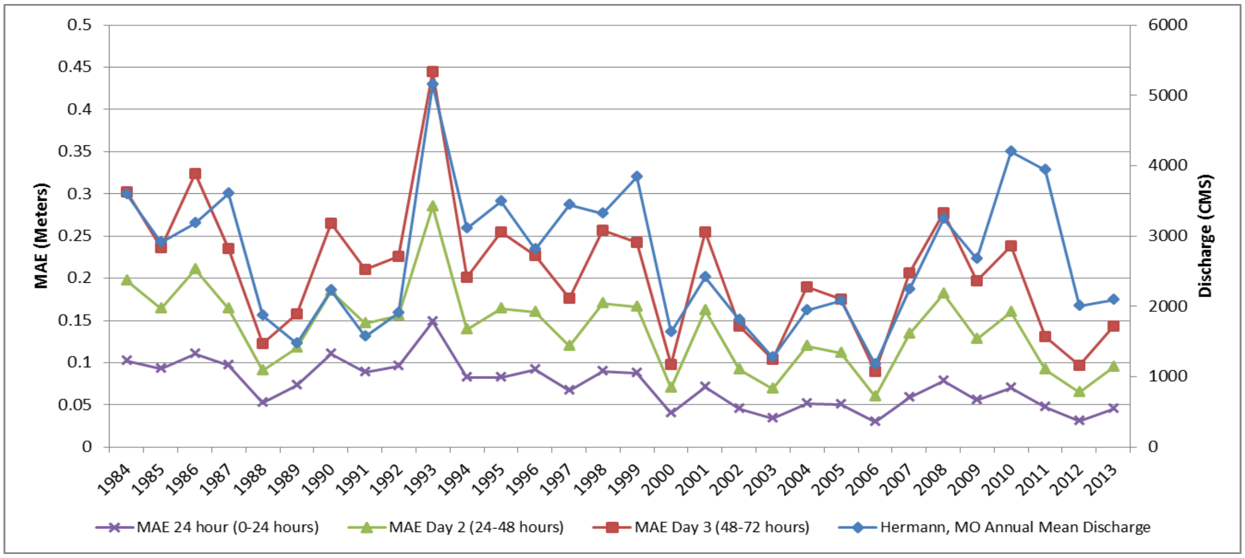

FIG. 11. Mean Absolute Error and Hermann, Missouri Annual Mean Discharge by Water Year

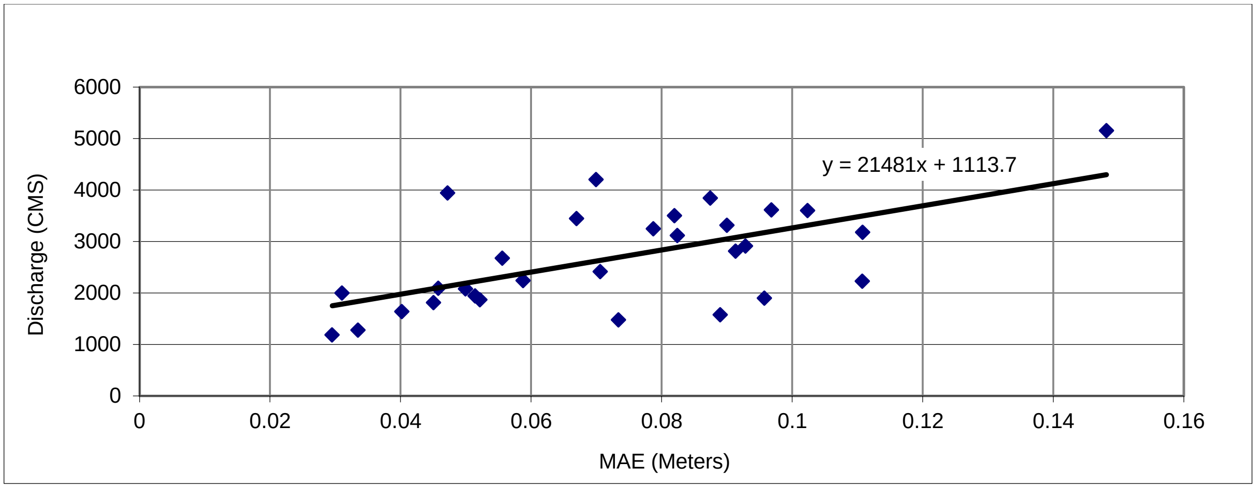

FIG. 12. Plot of 24 hours Lead time MAE versus Hermann Annual Mean Discharge

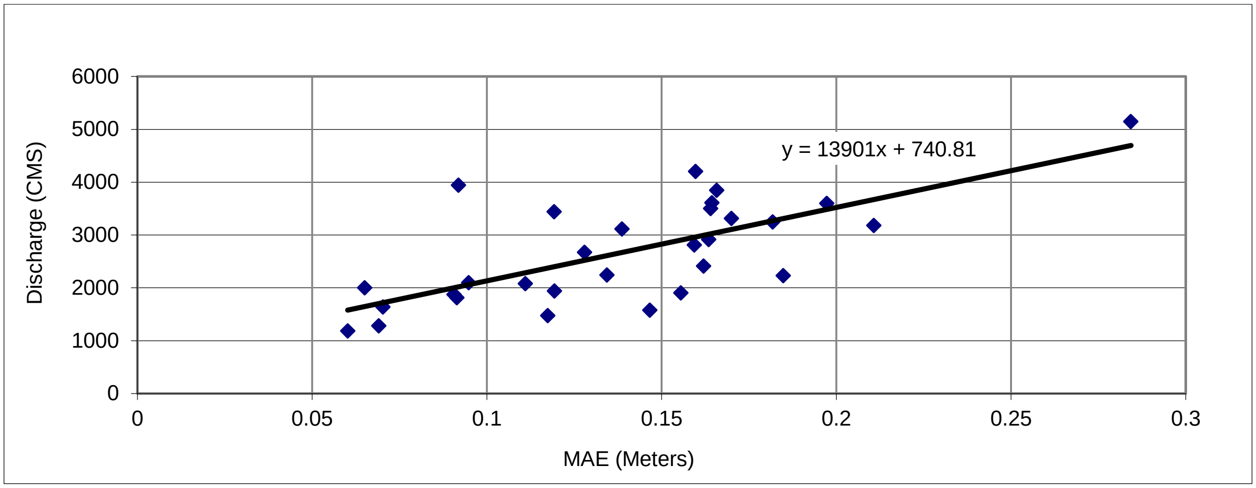

FIG.13. Plot of 48 hours Lead time MAE versus Herman Annual Mean Discharge

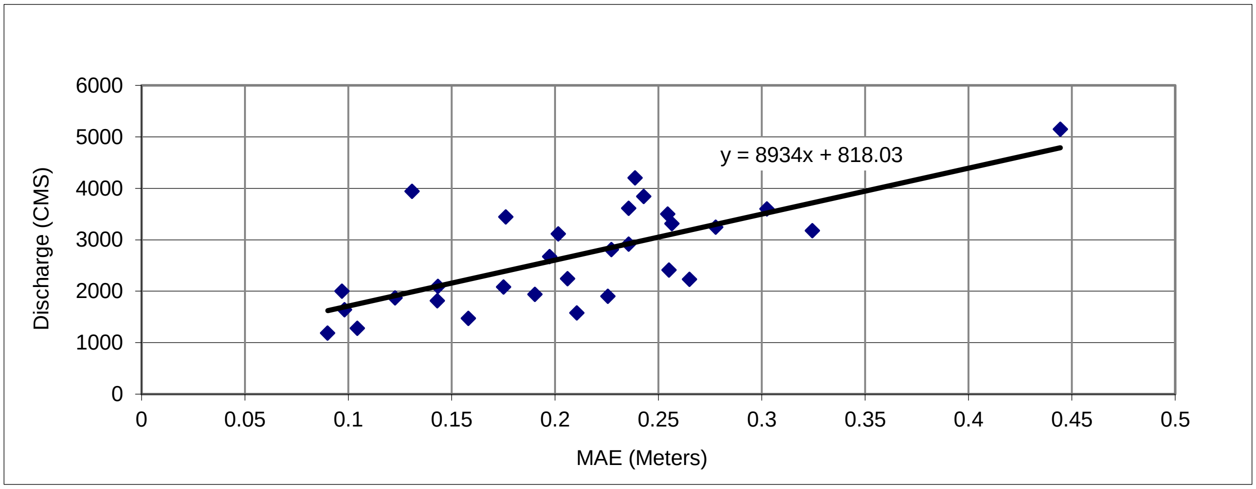

FIG.14. Plot of 72 hours Lead time MAE versus Hermann Annual Mean Discharge

MAE at Specific Discharges

Lead Time	Discharge			
	1700 cms	60,000 cfs	4250 cms	150,000 cfs
24 hours	0.03 meters	0.09 feet	0.15 meters	0.48 feet
48 hours	0.07 meters	0.23 feet	0.25 meters	0.83 feet
72 hours	0.10 meters	0.32 feet	0.38 meters	1.26 feet

TABLE 3. MAE by Lead Times for Specific River Discharges

4. SIGNIFICANT ADVANCEMENTS

The science of hydrology and hydrologic forecasting has made significant advancements over the past 60 plus years. The history of River Forecast Centers goes back to the late 1940s. At that time, all river and flood forecasting was done manually with hydrologists calculating rainfall, runoff and routing water downstream using manual and graphical techniques. Unit hydrographs and antecedent precipitation index (API) models were the general tools in vogue. Over the years, these techniques were transferred to early computers and then later to scientific workstations. In addition, dynamic routing methodologies and soil moisture accounting (SMA) models were added. The science of meteorology also advanced the ability to better forecast where and how much precipitation would fall, which enabled the RFCs to add this parameter to the hydrologic modeling mix. Better estimates of observed precipitation which are incorporated into river models have also improved with the implementation of the operational radar network of WSR-88D systems across the United States. This has allowed for better quantitative precipitation estimates (QPE) across areas where observation networks are stretched thin. Tables 4-8 identify some of these technological changes along with a rough time frame of implementation. It should be noted that Tables 4-8 refers only to the Missouri Basin RFC in Pleasant Hill, Missouri. However, the 12 other River Forecast Centers in the National Weather Service have gone through similar changes.

Another change since the earlier study is the implementation of a standard verification program at the River Forecast Centers in 2001. Prior to 2001 the Missouri Basin RFC had its own local verification program that was developed and implemented back in 1980. This local program was limited to daily forecast locations, which at that time were only available on the Missouri River. This standard verification program has allowed MBRFC to provide the river forecasters and others verification statistics for every river forecast location.

River Models Used

Year	***Changes***
Fall 1994 thru early 1997	NWSRFS (API/AI - Event Oriented Hydrological Model)
Early 1997 thru late 2011	NWSRFS (SAC-SMA Continuous Accounting Hydrologic Model)
2002	Begin implementing Ensemble Streamflow Probabilistic (ESP) Forecasting; about 3/4 complete in early 2014
Late 2011 to present	CHPS/FEWS (SAC-SMA - Continuous Accounting Hydrologic Model)
Mid-2013	CHPS/HEC-RAS Hydraulic Model implemented on Missouri River from Nebraska City, NE downstream to confluence with Mississippi River; work is in progress to model from Missouri River from below Gavin's Point Dam downstream to Nebraska City, NE

TABLE 4. Advances in River Modeling

River Model Time Steps	
Year	***Changes***
Late 1982 thru 1997	Mix of 6-, 12-, and 24-hour time steps
1998 thru mid-2010	All model segments at a 6-hour time step
Mid 2010 to present	Several locations in Lower Missouri Tributaries forecast group switched to 1-hour time step; all other locations remain on a 6-hour time step

TABLE 5. Changes in River Model Time Steps

Quantitative Precipitation Forecast (QPF) Usage	
Year	***Changes***
Mid-1996 thru mid-2001	24-hour QPF included in the river model
Mid-2001 thru mid-2009	April-Sept. 12-hour QPF and Oct.-March 24-hour QPF included in the river model
Mid-2009 to present	24-hour QPF included in the river model

TABLE 6. Changes in Quantitative Precipitation Forecast (QPF) Usage

Computer Technology	
Year	***Changes***
1981	AFOS DG-S230 installed at RFC (communications); teletypewriters phased out over the next few years
1981	DATACOL (DG-S140) installed at RFC (database/data collection system); over 600 river gages dialed 4 times per day
1985	Prime mini-computer installed at RFC; computer models now executed locally
Fall 1994	Operations moved to pre-AWIPS government development platform (gdp) system; Prime phased out
Fall 1996	AWIPS computer system installed
1993	DATACOL (DG-S140) replaced with PC HYDROMET
1999	Operations moved from pre-AWIPS to AWIPS; AWIPS ties to AFOS "cut" and all forecasts now sent via WAN/SBN; PC HYDROMET phased out and replaced by AWIPS/LDAD system

TABLE 7. Changes in Computer Technology

Data Availability	
Year	***Changes***
Fall 1946 thru early 1980s	Mostly manual readings for precipitation and temperature data, generally once per day (24-hr precipitation report at 7am, snow depth and snow water equivalent at 7am, max and min temperature for previous day); data relayed to RFC by weather office via teletypewriters.
Late 1960s to early 1980s	Many river gages have telemetry (telemark, DBT, DARDC, LARC), still quite a few manual river gages
Early 1980s to present	Exponential growth of all data due to new technology and computer software, such as satellite data collection platforms (DCP), meteor burst data, ALERT, IFLOWS, ASOS, AWOS, ROSA, WXCODER

TABLE 8. Advancements in Data Availability

5. CONCLUSIONS

It is clear that over the last 31 years continued improvements have been made in the accuracy of MBRFC hydrologic forecasts. These improvements have been due to more sophisticated techniques, data availability and handling, and modeling capabilities. It is also clear that MAEs are highly dependent on time frame and flows within the river system.

The determination of hydrologic forecasting errors is difficult and has been the subject of numerous studies within the National Weather Service. In 2001, a national program of hydrologic forecasting verification was implemented. Such a system is desirable for many reasons. It is necessary to determine the current level of forecasting skills so the effects of future innovations can be determined and documented. This will allow the river forecaster to better understand all the ramifications which computer systems, communications, data collection, data archiving and analysis, hydrologic model improvement, and forecaster training have on forecast accuracy. The verification statistics described here for twelve locations on the mainstem Missouri River is a first step for this type of analysis. This system of MAEs for the forecast points along the river gives us a good estimate of the current level of hydrologic forecasting skills. This verification system has also provided, at least in a limited fashion, a benchmark against which future improvements in all phases of forecasting skills can be measured (Welles, et al, 2007).

While some conclusions can be drawn from these data, there are many issues that can impact the quality of the hydrologic forecasts in addition to those discussed here. An area not discussed is the impact of seasonal precipitation on the overall MAE for any given year. An above average year of overall precipitation will result in an increased MAE as seen in the 1993 MAEs of this study. The comparison of MAEs to annual mean discharge at Hermann is a good start. An extension of this study would be to relate detailed precipitation records to river forecast MAEs for each year. A study relating to radar precipitation estimates might be of interest. It would also be useful to relate, in detail, the implementation of large scale NWS technological upgrades to the MAEs.

The U.S. Army Corps of Engineers Hydrologic Engineering Center’s River Analysis System (HEC-RAS) hydraulic model was recently implemented and should help in improving forecasts during flood events when levee breaches and overtopping can occur, the backwater impacts on the lower end of tributaries, and in the development of better routings for use in the hydrologic model. Verification of these advancements should be documented.

Unexplored at Missouri Basin RFC at this time is the accuracy of the ensemble streamflow prediction (ESP) forecasts which are nearing implementation completion. ESP forecasts are based on historical mean areal precipitation (MAP) and temperature (MAT) data for several years. The model uses current conditions as the starting point and then runs each year’s MAPs and MATs to create an ensemble of possible outcomes for the next several months. It is from this ensemble of possible outcomes on which probabilities are based. Again, verification of these advancements should be documented.

The Missouri Basin RFC has begun implementation (Fall 2014) of the Hydrologic Ensemble Forecast Service (HEFS) suite of software. The HEFS will allow the RFC to issue hydrologic forecasts that are “uncertainty aware”, i.e. they will provide information about forecast uncertainty. Unlike the deterministic forecasts discussed in this paper, these ensemble forecasts will provide a set of possible values. How HEFS will aid in achieving more forecast accuracy is unexplored at this time.

Again, verification should be considered and documented.

Finally, additional forecast accuracy will require greater attention to developing more sophisticated hydrologic models, the development and maintenance of additional and more accurate sources of input data which drive the models, additional systems capabilities for the RFCs, and additional training for RFC forecasters.

6. ACKNOWLEDGEMENTS

The authors wish to thank Lisa Holts (MBRFC) and Scott Dummer (MBRFC) for their significant input and contributions to this manuscript.

REFERENCES

Larson, L.W. and Schwein, N., *Temperature, Precipitation and Streamflow Trends in the Missouri Basin, 1895 to 2001*, 84th Annual AMS Meeting, Seattle, Jan. 2004.

Larson, L.W. and Schwein, N., *A Statistical Evaluation of Mainstem Forecasting Errors for the Missouri and Mississippi Rivers*, AMS Meeting, Orlando, Jan. 2002.

Natural Disaster Survey Report, *The Great Midwest Flood of '93*, National Weather Service, Central Region Hqs., Kansas City, MO., 1993.

USGS, Surface Water Statistics, http;//waterdata.usgs.gov/,2013

Welles, E., Sorooshian, S., Carter, G., and Olsen, B., *Hydrologic Verification: A Call for Action and Collaboration*, BAMS, April. 2007