RIVERSIDE

global science solutions

National Weather Service
Office of Hydrologic Development

FINAL REPORT (REVISION 1)

Hydrometeorological Forcings for Hydrologic Prediction and Hydrologic Model Calibration

Prime Contract: DG133W-03-CQ-0021
Task No. 10-15

Prepared by
Riverside Technology, inc.
May 2012

This report is provided in accordance with sub-tasks 2 and 3 of Task 10-15, the deliverables for which are defined in the technical proposal as follows:

- *A report on potential and recommended data sources and methods for creating multi-decade precipitation analyses of record for hydrologic model calibration.*
- *A report on potential and recommended data sources and methods for creating multi-decade temperature analyses of record for hydrologic model calibration.*

RIVERSIDE TECHNOLOGY, INC.
2950 EAST HARMONY ROAD, SUITE 390 ■ FORT COLLINS, CO 80528
TELEPHONE (970) 484-7573 ■ FAX (970) 484-7593 ■ www.riverside.com

Summary of Recommendations

The National Weather Service Office of Hydrologic Development intends to prepare a high-resolution, long-term historical archive of gridded precipitation and temperature. This analysis of record (AOR) will be developed for a ≤1-km grid mesh and a 1-hr temporal interval for the period 1979-2010, with eventual extension back to 1948. The AOR will include the continental U.S., Alaska, Hawaii, Puerto Rico, and the portions of Canada and Mexico that drain to the U.S.

Riverside has prepared this report to summarize research into the data sources, development methods, and evaluation techniques available for developing the precipitation and temperature AOR. The following recommendations are made throughout the report:

- All reasonable historical data sources should be included for development of the AOR to satisfactorily estimate temperature and precipitation at the high spatial and temporal resolution requirements, even though doing so inevitably means that the quality of the AOR will change over time as new data sources are introduced. The limitations of the AOR datasets should be clearly described for users, particularly related to the differing data sources and approaches applied for different periods.
- All historical data must undergo quality control to ensure high quality data are being input to the AOR development process. The quality control processes should be automated to a large extent, although some manual review would be beneficial if resources permit. Users of the AOR should have access to the raw input data, the quality-controlled data, and quality control information (e.g., correction factors), in addition to the final datasets.
- A background climatology dataset must be selected to establish the long-term characteristics of the AOR datasets. The selection criteria should consider the potential impact on the hydrologic models used at the river forecast centers. Station-based climate normals and gridded climatological datasets from the PRISM system are widely employed at the river forecast centers. However, existing climatological datasets may not accurately capture the climatology in every location, and may require adjustment based on an evaluation of the AOR, for example using hydrologic simulations.
- The feasibility and benefits of a high-resolution reanalysis to support the AOR development should be evaluated. A new, high-resolution reanalysis dataset would benefit the gridding of historical point observations, spatial downscaling of gridded datasets, and temporal disaggregation of the AOR datasets to an hourly time step by representing dynamic processes that affect local weather patterns.
- A data fusion framework should be developed, or adapted from an existing system, for the AOR development. The system should be developed to produce repeatable results, to support research and development, and to apply enhanced methods to improve the AOR datasets.
- The data fusion framework should employ uncertainty estimates to produce merged, multi-sensor precipitation and temperature fields. The uncertainty estimates may be simplistic for the initial AOR development. Additional development could benefit the methods used to estimate uncertainties, as this is an active area of research.
- The data fusion process should be accomplished on a daily time step prior to disaggregation to an hourly time step, as it is possible to develop a daily dataset with

significantly less uncertainty. More historical data are available on a daily time step. In addition, a daily dataset may suffice for some applications. Both the hourly and the daily datasets should be made available to users.

- Data fusion should be accomplished using a stepwise process including bias corrections for all input data sources, gridding of point observations, spatial downscaling of gridded inputs to a common grid mesh, temporal aggregation to a daily time step, merging of the input data sources, and disaggregation to an hourly time step.
- To transform point observations to gridded estimates, climatological differences between locations must be accounted for. One recommended methodology is climatology-aided interpolation, where the interpolation method is based on the synergraphic mapping system (SYMAP) algorithm or optimal interpolation.
- To downscale coarse-resolution precipitation grids, climatological differences between locations must be accounted for. A recommended methodology for downscaling precipitation fields is spatial disaggregation with bias corrections (e.g., Wood et al., 2004). For temperature, a downscaling method that employs interpolation with terrain adjustments may suffice.
- If resources permit, it would be beneficial for the gridding and downscaling methods to perform additional development to identify and account for significant regional physiographic and atmospheric variables that are not captured in the selected background climatological dataset (e.g., Hsu et al., 2011).
- Hourly weights to disaggregate daily fields to an hourly time step should be based on a combination of observations and stochastic methods. A hierarchy needs to be defined to prioritize and merge disaggregation weights from various sources (e.g., hourly gages > radars > high-resolution reanalysis datasets > satellites > stochastic methods > coarse-resolution reanalysis datasets).
- The requirements, development methods, and eventual quality of the AOR datasets will be affected by regional differences in climate, hydrology, forecast objectives, data sources, and data quality. Prototype datasets should be developed for multiple geographic areas and limited time periods to assess the suitability of selected methods for different regions prior to full development of the AOR.
- The AOR needs to be evaluated at multiple stages in the development process using multiple methods:
    - Expert input should be solicited at the beginning of the development process to provide input on the reasonableness of the proposed approach; at the prototyping phase of the development process to determine whether preliminary results are reasonable; and at the final evaluation phase of the development process to review the final datasets. A feedback mechanism should be established to document issues identified by users and to inform additional development efforts if the AOR is regenerated in the future.
    - The AOR datasets should be compared against existing precipitation and temperature datasets to identify biases and to assess the need to re-calibrate existing hydrologic models:
        - The AOR should be compared against the legacy forcings that have been used to calibrate the hydrologic models employed at the river forecast centers.

        - The AOR should be compared against existing temperature and precipitation datasets that have been developed using long-term, stable networks or for watersheds that have been monitored intensively.
    - The AOR datasets should be evaluated using hydrologic models. The hydrologic models employed at the river forecast centers should be included for suitable watersheds. In addition, distributed models should be considered for the evaluation process.
- An important outcome from the AOR development process is the establishment of a formal evaluation procedure that identifies geographically-representative test basins and objective evaluation metrics that can be used to quantify the value of new data, inputs, and models (CBRFC, personal communication, February 22, 2012).
- The development approach for the AOR should maximize the use of available frameworks, automated processes, methodologies, and datasets to reduce the development cost and effort. Existing systems that have been developed by the NCDC, NOHRSC, NSSL, OHD, NCEP, NASA, and academic institutions should be evaluated for this purpose.
- The AOR datasets should follow established guidelines for delivery and format requirements.
- The AOR datasets should be made available to the public, ideally with tools to facilitate subsetting by geographic location and time period.
- The AOR has the potential to be an enormous advancement, exceeding the spatial extents, periods of record, and resolutions of existing gridded datasets. As such, it has potential to benefit a wide variety of stakeholders. This provides an opportunity for the NWS OHD to engage a variety of interested parties in the development and maintenance of the AOR datasets. Intended customers of the AOR should participate in defining the development approaches and evaluating the resulting datasets. For example, the river forecast centers have particular expertise in data quality control and regional issues.

Contents

1 INTRODUCTION 1

2 AOR REQUIREMENTS 2

3 DATA SOURCES 4

3.1 STATIONS 5
3.2 GROUND-BASED RADARS 9
3.3 SATELLITE 14
3.3.1 Precipitation 14
3.3.2 Temperature 17
3.4 NUMERICAL WEATHER PREDICTION MODELS AND REANALYSIS DATA 19
3.4.1 Global Reanalysis Products 20
3.4.2 Regional and High-Resolution Reanalysis Products 22
3.4.3 Issues and Benefits 22
3.5 GRIDDED CLIMATOLOGY AND OBSERVATION DATASETS 23
3.5.1 PRISM Datasets 24
3.5.2 Maurer et al. and Hamlet/Lettenmaier Datasets 25
3.5.3 NCEP Datasets 26
3.5.4 NLDAS-2 Datasets 26
3.5.5 Other Datasets 27
3.6 SUPPLEMENTAL DATA SOURCES 28
3.7 CRITERIA FOR DATA INTEGRATION INTO AOR 28

4 DATA QUALITY CONTROL 28

4.1 TYPES OF QUALITY CONTROL CHECKS 30
4.2 CONSIDERATIONS PRIOR TO REMOVING DATA 34
4.3 RECOMMENDATIONS FOR THE AOR 36

5 DATA ASSIMILATION AND DATA FUSION 36

5.1 DATA ASSIMILATION 37
5.2 DATA FUSION 38
5.2.1 Bias Adjustments 38
5.2.2 Simple Data Fusion 39
5.2.3 Optimization-based Data Fusion 39
5.2.3.1 Artificial Neural Network Framework 39
5.2.3.2 Variational Assimilation 40
5.2.4 Uncertainty-based Data Fusion 40
5.2.4.1 Kriging 40
5.2.4.2 Optimal Interpolation 40
5.2.4.3 Kalman Filter Framework 41
5.3 IMPLICATIONS FOR THE AOR 42
5.3.1 Framework 42
5.3.2 Stepwise Application of Data Fusion 43
5.3.3 Separate Daily and Hourly Data Fusion Processes 44

6 GRIDDING METHODS 44

6.1 METHODS THAT IGNORE PHYSIOGRAPHIC CHARACTERISTICS AND ATMOSPHERIC CONDITIONS 45
6.2 METHODS THAT INCORPORATE PHYSIOGRAPHIC CHARACTERISTICS 46
6.3 METHODS THAT INCORPORATE PHYSIOGRAPHIC CHARACTERISTICS AND ATMOSPHERIC CONDITIONS 47
6.3.1 Atmospheric Classification 47
6.3.2 Regression Methods 48
6.3.3 Multi-sensor Approaches 48
6.4 RECOMMENDATIONS FOR THE AOR 49
7 SPATIAL DOWNSCALING 49
7.1 INTERPOLATION METHODS 50
7.2 STATISTICAL DOWNSCALING METHODS 51
7.2.1 Spatial Disaggregation 51
7.2.2 Bias Correction and Spatial Disaggregation (BCSD) 52
7.2.3 Regression Methods 53
7.3 DYNAMICAL DOWNSCALING METHODS 55
7.4 IMPLICATIONS FOR THE AOR 58
8 TEMPORAL DISAGGREGATION 58
8.1 GRID-BASED DISAGGREGATION APPROACHES 59
8.2 STATION-BASED DISAGGREGATION APPROACHES 59
8.3 MULTI-SENSOR APPROACHES 59
8.4 STOCHASTIC APPROACHES 60
8.5 IMPLICATIONS FOR THE AOR 62
9 AOR EVALUATION 62
9.1 EXPERT INPUT 62
9.2 CROSS VALIDATION AGAINST POINT OBSERVATIONS 63
9.3 COMPARISON WITH DATASETS FROM HIGHLY INSTRUMENTED BASINS 64
9.3.1 Evaluation of Techniques for the Early Period of Record 64
9.4 COMPARISON WITH PREVIOUSLY DEVELOPED DATASETS 65
9.5 HYDROLOGIC SIMULATION 66
9.5.1 Potential Applications 67
9.5.2 Evaluation Approach 68
9.6 DISCUSSION AND IMPLICATIONS FOR THE AOR 69
10 EXISTING WEATHER DATA MANAGEMENT SYSTEMS 70
10.1 NCDC AUTOMATED DATA QUALITY CONTROL PROCEDURES 72
10.1.1 TempVal 73
10.1.2 PrecipVal 73
10.2 PRISM 74
10.3 SNODAS 75
10.4 NMQ/Q2 76
10.5 DAILYQC/MOUNTAIN MAPPER 77
10.6 MPE 78
10.7 CPC 80
10.8 NLDAS 81
11 TECHNICAL CONSIDERATIONS 82

11.1 TEMPORAL CONSIDERATIONS 82
11.2 REGIONAL CONSIDERATIONS 84
11.2.1 Alaska 84
11.2.2 Hawaii 87
11.2.3 Puerto Rico 87
11.2.4 Southeast CONUS 88
11.2.5 Intermountain West CONUS 89
11.2.6 Canada 90
11.2.7 Mexico 90
12 RFC CONSIDERATIONS 92
12.1 CONSISTENCY OVER THE HISTORICAL PERIOD 92
12.2 CONSISTENCY WITH LEGACY FORCINGS 93
12.3 SUMMARY OF RFC REQUIREMENTS 94
13 SUMMARY AND AOR RECOMMENDATIONS 94
13.1 DATA SOURCES 95
13.2 DATA QUALITY CONTROL 96
13.3 AOR FRAMEWORK 97
13.3.1 Data Assimilation/Data Fusion 99
13.3.2 Gridding Methods 100
13.3.3 Spatial Downscaling Methods 100
13.3.4 Temporal Disaggregation Methods 101
13.4 EVALUATION METHODS 102
13.5 EXISTING WEATHER DATA MANAGEMENT SYSTEMS 102
13.6 TECHNICAL CONSIDERATIONS 103
13.7 RFC CONSIDERATIONS 104
13.8 LOGISTICAL CONSIDERATIONS 104
13.9 NEXT STEPS 105
13.9.1 Parallel Development 105
13.9.2 High-Resolution NWP Model Reanalysis 106
13.9.3 Prototype Evaluations 106
14 ACKNOWLEDGEMENTS 107
15 ACRONYMS 107
16 REFERENCES 111

APPENDIX A: DATA INVENTORY

List of Tables

Table 1. Minimum Requirements for the Analysis of Record 3
Table 2. Recommended Station Networks 6
Table 3. Recommended Sources for Ground-Based Radar Precipitation 12
Table 4. Radar Reflectivity Values for Varying Snow Types (Matrosov et al., 2009) 13
Table 5. Recommended Sources for Satellite Precipitation Data 16
Table 6. Satellite Land Surface Temperature Data 18

Table 7. Recommended Sources for Reanalysis Data 19
Table 8. Gridded Data Sources 23
Table 9. Strengths and weaknesses of different data sources (reproduced from Vasiloff et al., 2007)..... 37
Table 10. Grid Meshes associated with Historical Data Archives..... 49
Table 11. Comparison of Weather Data Management System Attributes..... 72
Table 12. Historical Data Availability 83

List of Figures

Figure 1. Hourly COOP, ASOS, CRN and HADS precipitation stations in the CONUS (reproduced from Rennie, 2011a) 8
Figure 2. National Doppler Radar Sites (reproduced from NWS, 2004)..... 10
Figure 3. Terminal Doppler Weather Radar Sites (reproduced from MIT, 2011)..... 11
Figure 4. Effect of Temperature on Simulated Streamflows for the South Fork Rio Grande at South Fork, CO 29
Figure 5. Double mass analysis plot showing the effect of changes in temperature sensors at SNOTEL gages 34
Figure 6. Example of Quantile Mapping for Bias Correction (reproduced from Wood et al., 2006)..... 53
Figure 7. Geographic Regions used by GMOS for NDGD Forecast Products (reproduced from Charba and Samplatsky, 2011a)..... 55
Figure 8. Advantages and Disadvantages of Statistical and Dynamical Downscaling Techniques (reproduced from Fowler et al., 2007) 56
Figure 9. Effect of orographic blocking on precipitation-elevation relationships (reproduced from Hughes et al., 2009)..... 57
Figure 10. Determination of the diurnal temperature cycle using NWP model outputs and satellite and surface observations (reproduced from Jin and Dickinson, 1999)..... 60
Figure 11. FAA WSR-88D Radars in Alaska (ARH, 2011)..... 86
Figure 12. Mexico's Radar Network (CONAGUA, 2011)..... 91
Figure 13. Conceptual framework for the development of a data fusion system (reproduced from Esteban et al., 2005) 97
Figure 14. Potential data fusion framework for the AOR..... 98

1 Introduction

The National Weather Service (NWS) Office of Hydrologic Development (OHD) intends to prepare a high-resolution, long-term historical archive of gridded hydrometeorological forcings, including precipitation and temperature. Precipitation data are the primary meteorological inputs to the hydrologic models run operationally by the River Forecast Centers (RFCs). Temperature data are also needed to determine the form of precipitation and the rate of snowmelt. The retrospective precipitation and temperature datasets are referred to as the analysis of record (AOR).

The AOR was first envisioned as a means to verify gridded forecasts from the National Digital Forecast Database (Colman et al., 2005; Horel and Colman, 2005). A significant driver for the development of the AOR currently is that the NWS is transitioning forecasting operations from the National Weather Service River Forecast System (NWSRFS) to the Community Hydrologic Prediction System (CHPS). As part of this conversion, RFC personnel will evolve the operational processes used to develop the meteorological inputs towards high-resolution gridded inputs. Currently, many of the RFCs utilize meteorological forcings that are basin average values generated from a weighted combination of point observations. This practice dates to an earlier era when point observations were the only data available, and basin average values were sufficient for lumped hydrologic models. The relatively recent availability of gridded precipitation and temperature estimates from radars, satellites, and numerical weather prediction (NWP) models makes it desirable to retire the older method of direct computation of basin averages. Moreover, the basin average datasets are not suitable for distributed hydrologic models, which generally require gridded inputs.

The NWS intends to use the AOR to calibrate lumped and distributed hydrologic models used for forecast operations, to determine climatological statistics for forecast ensembles, and to support forecast verification (NWS, 2011a). The AOR also has broad potential to support agencies and applications outside of the NWS. In May 2011, the U.S. Army Corps of Engineers (USACE), the U.S. Geological Survey (USGS), and the National Oceanic and Atmospheric Administration (NOAA) signed a memorandum of understanding that formalizes a commitment to multiagency collaboration that will “address the goals of the Integrated Water Resources Science and Services (IWRSS) initiative and the objective of Building Strong Collaborative Relationships for a Sustainable Water Resources Future initiative to build a Federal Support Toolbox for Integrated Water Resources Management (USACE et al., 2011).” Under the IWRSS initiative, the AOR can be used by the USACE to simulate reservoir operations and by the USGS to conduct a national assessment of water availability and use through the National Water Census (USGS, 2002).

Riverside Technology, inc. (Riverside) has considerable experience with meteorological datasets and hydrologic models, and in particular with the data and models used by the RFCs for forecast operations. Riverside has developed forcing datasets, calibrated hydrologic models, and implemented forecast systems for hundreds of basins in the U.S. and around the world. As a result of this experience, Riverside is knowledgeable about the requirements for calibration datasets, including unbiased climatological characteristics and consistency with the data that are available for real-time operations. Riverside is also knowledgeable in identifying and adjusting inconsistencies in calibration datasets due to changes in data networks and sensor hardware.

Riverside is supporting the NWS OHD by researching data sources, development methods, and evaluation techniques for the AOR. Riverside collected information from a variety

of sources, including peer-reviewed journal articles, conference proceedings, webinars, presentations, websites, and expert interviews. This report presents a summary of that research, along with implications for hydrologic modeling and recommendations for the AOR. This report will be used, among other inputs, by the NWS OHD to develop an implementation plan for the AOR.

The remainder of the report is organized as follows:

- ***Section 2: AOR Requirements*** contains a summary of the requirements that have been defined for the AOR.
- ***Section 3: Data Sources*** contains information on useful sources of historical temperature and precipitation data from stations, radars, satellites, and reanalyses.
- ***Section 4: Data Quality Control*** contains information on quality control checks that should be performed on all data sources to ensure that high quality datasets are used to produce the AOR.
- ***Section 5: Data Assimilation/Data Fusion*** contains a summary of approaches that can be used to combine observations from multiple data sources.
- ***Section 6: Gridding Methods*** contains a summary of techniques that transform point observations to gridded estimates.
- ***Section 7: Spatial Downscaling*** contains a summary of techniques that generate a higher resolution grid from a lower resolution grid.
- ***Section 8: Temporal Disaggregation*** contains a summary of techniques that generate higher resolution time series from lower resolution time series.
- ***Section 9: AOR Evaluation*** contains a summary of methods for evaluating the AOR using existing temperature and precipitation datasets, as well as streamflow simulations from hydrologic models.
- ***Section 10: Existing Weather Data Management Systems*** provides an overview of several of the systems that currently collect and quality control weather data, produce gridded estimates, and/or produce multi-sensor estimates. These systems have frameworks, automated processes, methodologies, and datasets that may be beneficial for the AOR.
- ***Section 11: Technical Considerations*** discusses the temporal and regional influences on data availability and quality that will affect the development of a consistent, long-term AOR.
- ***Section 12: Summary and AOR Recommendations*** contains a summary of previous sections and recommendations for the AOR.

2 AOR Requirements

The NWS OHD has defined minimum spatial and temporal requirements for the initial AOR development, as summarized in ***Table 1***. The spatial extent of the AOR must include the continental U.S. (CONUS), Alaska (AK), Hawaii (HI), Puerto Rico (PR), and the portions of Mexico (MX) and Canada (CN) that drain to the U.S. The AOR will have a temporal interval of at least one hour and a grid mesh of less than or equal to one kilometer. The AOR will be developed initially to include precipitation and temperature, the meteorological variables required to support RFC operations. The AOR will later include variables such as

evapotranspiration, freezing level, solar radiation, longwave radiation, relative humidity, and wind speed to support the RFCs, land surface modeling, and the evaluation of meteorological forecasts (Horel and Colman, 2005; NWRFC, personal communication, February 27, 2012). The initial AOR will be developed for the period 1979-2010. Eventually, the AOR will be extended back to 1948 to support applications that would benefit from a longer archive, such as the Great Lakes pilot study being conducted as part of the National Water Census (USGS, 2011). Moving forward, the AOR will be extended using gridded estimates of temperature and precipitation that are generated in real-time at the National Water Center to support river forecast operations (OHD, personal communication, October 5, 2011).

Table 1. Minimum Requirements for the Analysis of Record

Spatial Extent	CONUS, AK, HI, PR, MX, CN
Temporal Interval	≤1 hour
Spatial Grid Mesh	≤1 kilometer
Period of Record	1979-2010
Meteorological Variables	Precipitation, temperature

In addition to the spatial and temporal requirements, additional requirements have been identified that impact the utility of the AOR for RFC forecast operations:

- The AOR needs to maintain statistical stability over the period of record, particularly with respect to long-term averages, but also with respect to variability. This requirement presents a challenge due to changes in the observing systems over time, including the addition and removal of entire networks and individual observing sites, the introduction of radars and satellites, and improvements in precipitation estimation algorithms.
- The AOR needs to be consistent with the precipitation and temperature grids that will be available in real-time to avoid introducing biases into the forecasting process. It is unlikely that the AOR will be restricted to include only those data sources that are available in real-time. On the contrary, the high spatial and temporal resolution requirements for the AOR necessitate that all reliable historical data be used to generate the best possible estimates. However, the AOR and the real-time procedures should reflect the same climatology to ensure consistency. The need for consistency implies that the methodologies for the AOR and the real-time processes should be developed in conjunction.
- The AOR needs to be evaluated against the legacy forcings that have been used to develop the hydrologic models employed at the RFCs. It is highly likely that the AOR will yield different basin average values than were used for model calibration. The NWS OHD has summarized work performed at four RFCs to evaluate the differences in operational forcings between the legacy NWSRFS and the current CHPS (NWS, 2011b). The differences in basin-average values between the AOR and the calibration datasets are expected to manifest as both random and systematic biases. The magnitude of the differences will dictate the need for adjustments to the hydrologic model parameters.

The AOR has the potential for future improvements due to continually improving data sources, data assimilation methods, and quality control techniques. The impact on the hydrologic models, including the need for model parameter adjustments and the relative improvement in simulation quality, should be considered prior to updating the AOR.

3 Data Sources

Riverside developed an inventory of data sources that are available for temperature and precipitation in the United States (including Alaska, Hawaii, and Puerto Rico) as well as the portions of Mexico and Canada that drain to the United States. The data inventory includes station, radar, satellite, and reanalysis products. While the data inventory is intended to capture major sources of historical precipitation and temperature data, Riverside acknowledges that it is likely not comprehensive, given the breadth of work being performed and the rapidity with which changes are made.

The data inventory includes a number of descriptive attributes:

- **Data type:** The assigned values include point observation, gridded station observation, gridded radar observation, gridded satellite observation, gridded multi-sensor estimate, and gridded reanalysis product.
- **Data source:** This field represents the agency responsible for generating (and often distributing) the data.
- **Spatial extent:** The data inventory indicates whether the data source is available for the CONUS, Alaska, Hawaii, Puerto Rico, Canada, and Mexico.
- **Spatial resolution/grid mesh:** This field indicates the nominal grid mesh of the gridded data sources.
- **Temporal interval:** This field indicates the temporal interval of the data source.
- **Historical period of record:** This field provides information about the period of historical data availability, which has implications for the quality and the consistency of the AOR.
- **Historical data format and coordinate system:** This information is provided to support an assessment of the effort required to include the data source in the AOR.
- **Data assimilation:** This information is provided to indicate which data sources have been adjusted using *in situ* observations. Where both raw and adjusted data sources are available, it may be preferable to use the raw data in developing the AOR so that improved quality control and data assimilation procedures can be applied.
- **Weather variables:** The data inventory indicates whether the data source provides precipitation, temperature, evaporation, freezing level, solar radiation, longwave radiation, relative humidity, and wind speed.
- **Data types:** For precipitation, this field indicates whether the historical data are accumulated or incremental values. For temperature, this field indicates whether the historical data represent minimum, maximum, average, or instantaneous temperature values.
- **Data units:** This field indicates whether the historical data are available in metric or English units.

- **Web links:** The data inventory includes uniform resource locators (URLs) to the historical data archives and metadata sources.

In addition to the attributes that describe the historical data, the inventory contains information regarding real-time data availability, format, and latency. This information may be useful to OHD for developing an implementation plan for the real-time forcings, but is also useful in considering the consistency between the AOR and the real-time forcings.

Data sources that are recommended for use in the AOR, or that are recommended for additional investigation, are described in the following sections. ***Appendix A*** contains the complete data inventory, including data sources that are not recommended for inclusion in the AOR.

3.1 Stations

Station networks are the most difficult data sources to inventory comprehensively due to the number of local and regional networks. For the purposes of the AOR development, it may be most efficient to obtain lists of stations used at each RFC for the development of calibration and real-time forcings to ensure no important stations are omitted (CBRFC, personal communication, February 22, 2012; SERFC, personal communication, December 14, 2011).

Station networks are designed for a variety of monitoring objectives, including synoptic-scale weather conditions, meso-scale weather conditions, flood warning, snowpack assessment for water supply, fire safety, climate change, soil moisture, and evapotranspiration for agricultural applications. The monitoring objectives often dictate the network density, the weather variables that are measured, and the sensor equipment that is installed. Most station networks measure both precipitation and temperature, with the exception of flood warning networks such as the Automated Local Evaluation in Real Time (ALERT) network and the Integrated Flood Observing and Warning System (IFLOWS). In addition, some of the citizen-observed weather networks are focused on precipitation, such as the Community Collaborative Rain, Hail, and Snow (CoCoRaHS) network.

Station networks can be further categorized according to manual or automated observations and agency or citizen-collected data. Increasingly, manual observations tend to be associated with citizen-collected data that are provided on a daily time step, while agency data are collected and transmitted automatically on an hourly (or even more frequent) time step. Although citizen networks tend to have shorter historical archives and uncertainties that are less well defined, the "observations have been very important for [the] understanding of small-scale variations of surface weather features" and "can be quite useful to emergency managers and local meteorologists when flooding rains or tropical cyclone conditions threaten (Ruscher, 2004)."

Table 2 presents a summary of the national and regional station networks that are recommended for inclusion in the AOR, including the spatial extent, temporal interval, and historical archive associated with each network. In terms of spatial coverage, the highest station density is found in the eastern half of the CONUS and along the west coast (***Figure 1***). The station density tends to be lower in the intermountain west and areas outside of the CONUS. In terms of temporal consistency, there are several notable dates that affect the amount and quality of the available point observation data:

- 1948: Digital records for the National Climatic Data Center (NCDC) Cooperative (COOP) observer network become available, representing a significant increase in spatial representation.
- 1979: The Natural Resources Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) network becomes available, representing a major advance in monitoring at high elevations.
- 1990s: The number of automated weather stations rapidly increases, representing a significant improvement in weather characterization on a sub-daily timescale.

Table 2. Recommended Station Networks

Station Network	**Data Source**	**Temporal Interval**	**Earliest Data Availability**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
Cooperative (COOP) Network	NCDC	Daily, hourly	1948	✓	✓	✓	✓		
Historical Climatology Network (HCN)	NCDC	Daily	1900	✓	✓				
Climate Reference Network (CRN)	NCDC	Daily, hourly	2000	✓	✓	✓			
Snowpack Telemetry (SNOTEL)	NRCS	Daily, hourly	1978	✓	✓				
Soil Climate Analysis Network (SCAN)	NRCS	Daily, hourly	1993	✓	✓	✓	✓		
Remote Automated Weather Stations (RAWS)	National Interagency Fire Center	Daily, hourly	1985	✓	✓	✓	✓		
Automated Surface Observing System (ASOS)	NOAA/ DOD/ FAA	Hourly	1991	✓	✓	✓	✓		
	FAA	Hourly	1941	✓	✓	✓	✓		
Integrated Flood Observing and Warning System (IFLOWS)	NOAA NWS	Irregular (sub-hourly)	1981	✓					
Community Collaborative Rain, Hail, and Snow (CoCoRaHS) network	CoCoRaHS	Daily	1998	✓	✓	✓			

Station Network	Data Source	Temporal Interval	Earliest Data Availability	CONUS	AK	HI	PR	CN	MX
Citizen Weather Observer Program (CWOP)	CWOP/ MADIS	5-minute	Unknown (<2003)	✓	✓	✓	✓	✓	✓
Mesonet	Various (e.g., MesoWest, MADIS)	Sub-hourly	Varies	✓	✓	✓	✓	✓	✓
Agricultural Weather Network (AgriMet)	USBR	15-minute to daily	1983	✓					
Canadian Climate Data	Environment Canada	Daily, hourly	Unknown					✓	
CLICOM Mexican Daily Surface Data	CONAGUA	Daily	1902						✓
Mexico Sistema de Información Hydrométrica (SIH) database	CONAGUA	Daily, hourly	1995						✓
Mexican automated meteorological stations	CONAGUA	10-minute	1999						✓
Mexico Agriculture Automated Weather Station Data	AGROSON	10-minute	2004						✓
Northwest Mexico NAME rain gage network; NAME supplemental rain gage Network	NCAR Earth Observing Laboratory	5-minute to daily	2002						✓

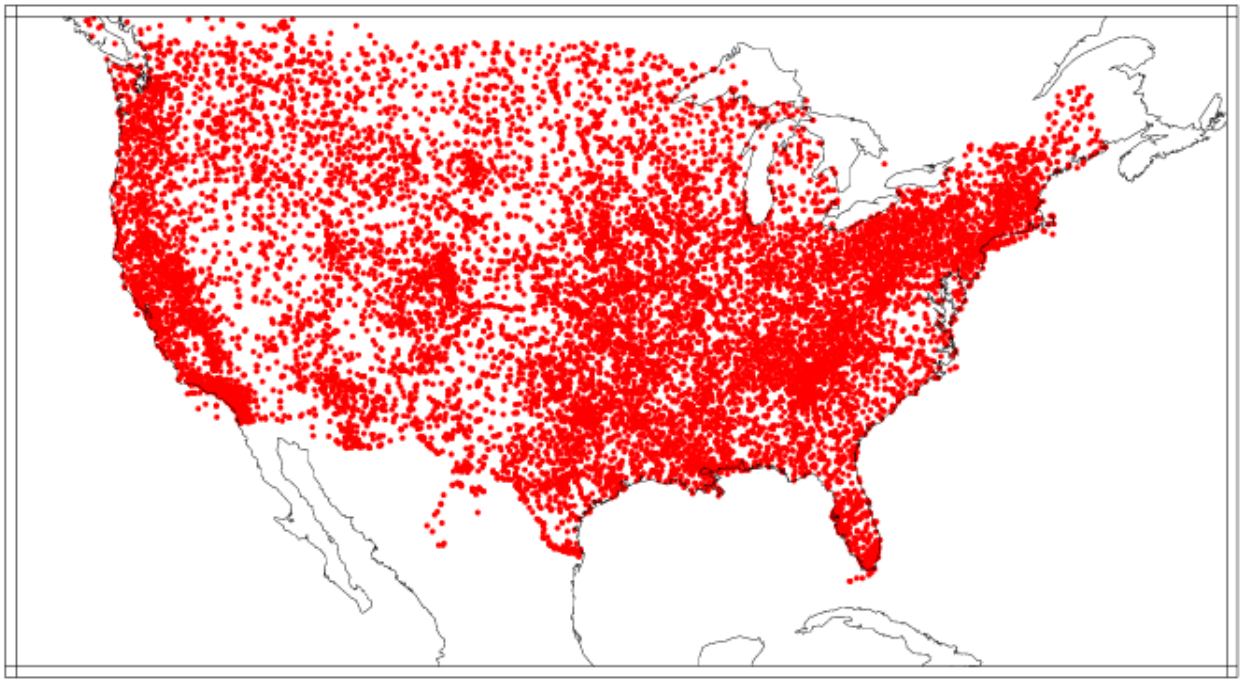

Figure 1. Hourly COOP, ASOS, CRN and HADS precipitation stations in the CONUS (reproduced from Rennie, 2011a)

Many factors affect the quality of station observations, and those factors tend to vary depending on whether the observations are manually obtained or automated. For example, automated stations are affected by mechanical and electronic malfunctions, transmission errors, decoding errors, inaccuracies in the algorithms used to convert between direct readings and temperature or precipitation, and drift over time if the instruments are not frequently recalibrated (Seo and Breidenbach, 2002; Daly et al., 2005; Kondragunta and Shrestha, 2006). In addition, automated stations more frequently suffer from gages that clog with frozen precipitation, resulting in a reduced or zero precipitation amount (West Gulf River Forecast Center [WGRFC], personal communication, November 17, 2011).

Manual stations are affected by:

- Nonstandard equipment. Relatively strict criteria exist for federal observing systems, while citizen observing networks may not have standard requirements for sensors and measurement methods (Creager, 2006). Some citizen observing networks, such as the CoCoRaHS network, do have standard equipment.
- Changing station locations. As volunteers cease participation, the station identifiers may be assigned to new volunteers at slightly different locations and elevations. Changes in station locations affect the long-term consistency of the historical dataset. This issue affects the NCDC COOP network in particular.
- Observation times. Citizen-observed data tend to be reported on a daily time step, typically in the morning at 7-8 am local time, or in the evening at 5-6 pm local time. The observation times need to be explicitly accounted for in distributing daily values to an hourly time step, as the observation times affect the period represented by the reported data.

- User errors. A common error is that citizen observers may report the depth of frozen precipitation rather than the water equivalent. Another common issue is that observers may report a precipitation value on Monday morning that represents the total precipitation accumulation over the preceding weekend rather than individual daily totals.
- Low participation rates. Citizen-observed weather data may be less reliable than automated observations. For example, a CoCoRaHS coordinator in Indiana has reported that approximately 25% of registered observers actually submit reports (Baker, 2011).
- Inadequate maintenance. Most weather instrumentation requires periodic maintenance and calibration. Volunteers may be less likely than operational agencies to perform required maintenance (Creager, 2006).

All stations are affected by:

- Catch deficiencies. The catch deficiency is the amount by which the station underreports the true amount of precipitation. The magnitude of the catch deficiency is a function of the precipitation type, gage type, and wind characteristics at the orifice of the gage (Anderson, 2002). In general, the catch deficiencies are greater for snow than for rain.
- Sensor characteristics. Some gages, such as those in the RAWS network, lack heating elements, making them unsuitable for measuring frozen precipitation (Daly et al., 2008). However, gages with heating elements are subject to evaporative losses. Changes in sensor equipment can affect the consistency of the historical record. For example, NRCS SNOTEL stations in Colorado were modified in the early 2000s to use extended range temperature sensors, producing noticeable breaks on double mass analysis plots (Riverside, 2010a).
- Inaccurate metadata. Point observations may erroneously reflect weather characteristics if the associated latitude, longitude, or elevation attributes are inaccurate.

In addition to providing high quality point observations of precipitation and temperature, station observations are often used to perform bias adjustments for remotely-sensed precipitation estimates from radars and satellites. Hourly precipitation observations can be used for distributing daily precipitation values to a sub-daily time step.

3.2 Ground-Based Radars

The radars included in the data inventory can be generally categorized according to weather radars, aviation radars, radars mounted to satellites, and distributed networks of limited-range radars. The radars differ in operating specifications such as range, frequency, and beam width. For the purposes of the AOR, the primary source of radar precipitation is the Weather Surveillance Radar-1988 Doppler (WSR-88D) radars in the Next-Generation Radar (NEXRAD) network operated by the NWS, Federal Aviation Administration (FAA), and the Department of Defense (DOD) (***Figure 2***). The observations undergo a series of data processing from the raw

analog data (Level 1) to reflectivity and precipitation (Level 2) to precipitation on an hourly time step and the Hydrologic Rainfall Analysis Project (HRAP) grid system (Level 3). The precipitation grids from individual radars are mosaicked by the NWS RFCs to cover the RFC domain and by the National Center for Environmental Prediction (NCEP) to cover the CONUS. The algorithms for quality controlling the reflectivity data and for estimating precipitation rates have improved significantly over time. At the RFCs, a series of gage-radar processing algorithms, including Stage 2, Stage 3, and RFCWide, have been replaced by the Multi-sensor Precipitation Estimation (MPE) program. MPE includes multiple gage-only and gage-radar processing options, including some that have been developed outside OHD.

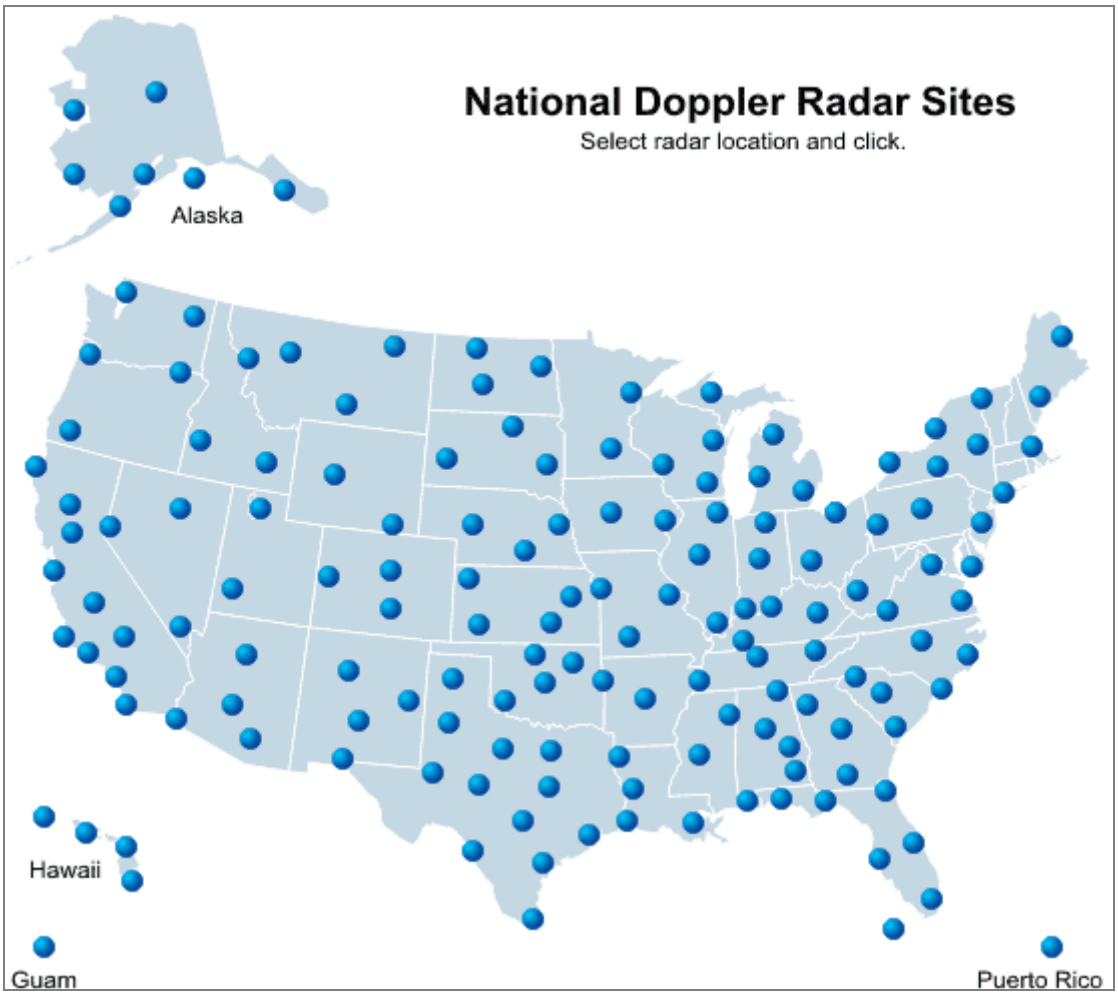

Figure 2. National Doppler Radar Sites (reproduced from NWS, 2004)

A secondary source of radar precipitation for the AOR is the Terminal Doppler Weather Radar (TDWR) network operated by the FAA (***Figure 3***). The TDWR radars are installed at or near airports to improve travel safety by detecting thunderstorms and associated wind shear and microbursts (NSSL, 2011a; MIT, 2011). The TDWR radars have a shorter range than the WSR-88D radars, but a higher operating frequency, resulting in higher resolution precipitation estimates for overlapping domains. The narrow beam and ground clutter suppression algorithms provide high quality data on boundary layer dynamics, which is important for forecasting

convective activity (MIT, 2011). However, the TDWR radar beams are significantly attenuated by heavy precipitation (Ferree, 2005). Though some correction algorithms have been developed, none have been implemented operationally for precipitation from TDWR.

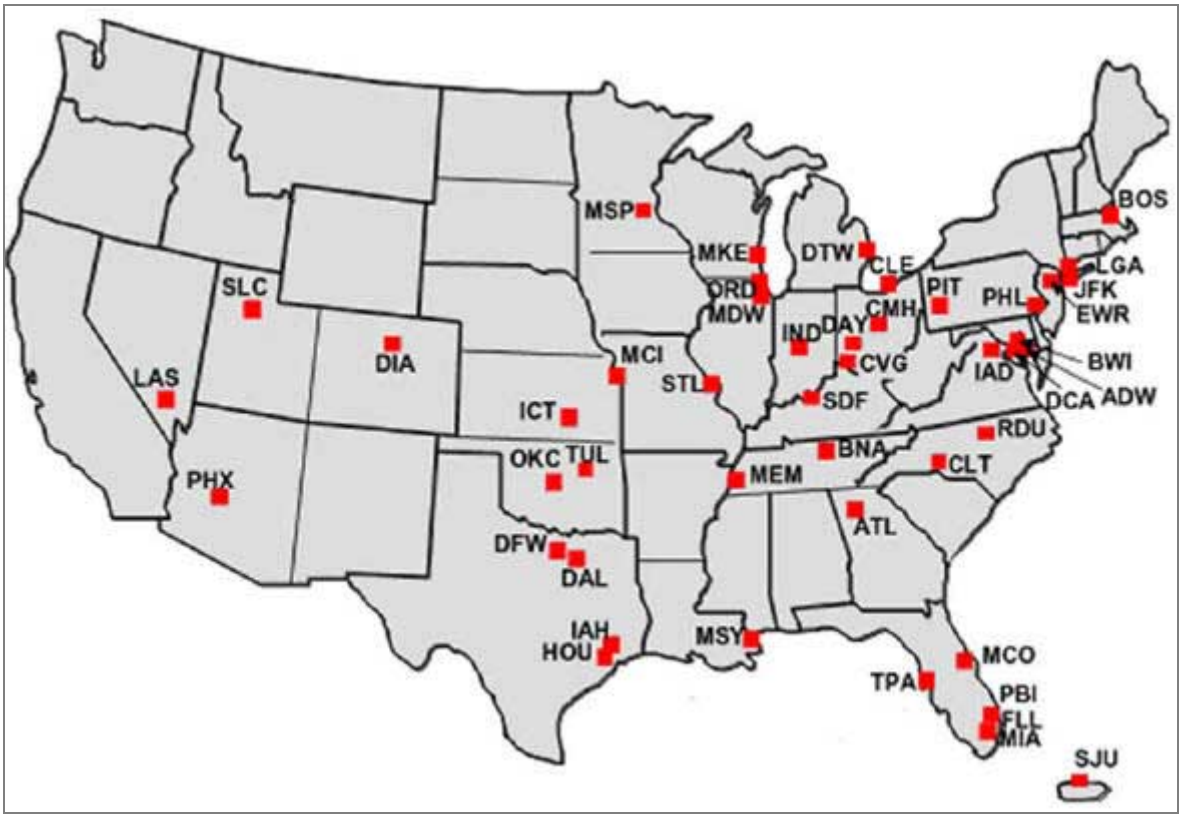

Figure 3. Terminal Doppler Weather Radar Sites (reproduced from MIT, 2011)

Table 3 presents a summary of the radar precipitation products that are recommended for consideration in the AOR. Riverside prioritized base radar precipitation products over bias-adjusted or multi-sensor products to facilitate the application of consistent quality control techniques and statistical adjustments that use the maximum amount of historical data. For gage-radar estimates prior to 2002, the NEXRAD precipitation products generated by the RFCs using Stage 3 and RFCWide will require mosaicking and statistical adjustments (OHD, personal communication, September 29, 2011). OHD is investigating the utility of hourly NCEP Stage 2 radar-only mosaics produced for the period 1996-2001.

After 2002, the national radar mosaic produced by NCEP (i.e., the NCEP Stage 2 radar-only product) can be utilized. The NWS OHD has already begun improving the NCEP Stage 2 data archive by removing temporal and spatial artifacts and applying bias adjustments (OHD, personal communication, December 2, 2011). The NCEP NEXRAD Stage 4 multi-sensor product, which benefits from quality control performed at the RFCs, may also have value for developing the AOR.

In addition to the NEXRAD and TDWR products, the National Severe Storms Laboratory (NSSL) produces a high quality national radar mosaic (NMQ) that includes the WSR-88D, TDWR, Canadian, and gap-filling radars. The NWS OHD has a NOWRAD dataset

generated by WSI that requires decoding and conversion from reflectivity to precipitation. The National Center for Atmospheric Research (NCAR) has an archive of manually-digitized radar products that may be useful for developing the AOR, but would require significant efforts to process. Canada and Mexico maintain national radar networks, but access to the full historical records is not assured.

The NSSL and the NCDC have recently developed a plan for re-creating multiple years of NMQ precipitation within the Multi-radar and Multi-sensor (MRMS) platform. Generation of the retrospective archive, including an upgrade of the spatial resolution to a 500-m grid mesh, will be implemented within two years. In the near future, the real-time MRMS will also be upgraded to higher spatial resolution (Kenneth Howard, NSSL, personal communication, 2012).

Table 3. Recommended Sources for Ground-Based Radar Precipitation

Product Name	**Data Source**	**Grid Mesh**	**Temporal Interval**	**Data Availability**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
NEXRAD Stage 3	NWS	4-km	Hourly	1993-2000	✓	✓	✓	✓		
NEXRAD RFCWide	NWS	4-km	Hourly	2000-2002	✓	✓	✓	✓		
NCEP NEXRAD Stage 2 Radar-only	NCEP	4-km	Hourly, 6-hour, Daily	1996-Present	✓	✓	✓	✓		
NCEP NEXRAD Stage 4 Multi-sensor	NCEP	4-km	Hourly, 6-hour, Daily	2002-Present	✓	✓	✓	✓		
TDWR	FAA, NWS	4-km	5-minute to Daily	1994-Present	✓			✓		
NMQ (current)	NSSL	1-km	5-minute to Daily	2008-Present	✓			✓		
NMQ (planned)	NSSL	500-m	5-minute to Daily	2012+	✓			✓		
NMQ (retrospective)	NSSL, NCDC	500-m	5-minute to Daily	2002-2010	✓					
NOWRAD	WSI	2-km	15-minute	1996-2007	✓					
Manually Digitized Radar	NWS / NCAR	40-km	Hourly	1978-1994	✓					
Canadian Weather Radar	Environment Canada	Unknown	10-minute	<2007-present	✓				✓	
Mexican Radars	CONAGUA	Unknown	Unknown	Unknown						✓

Radar precipitation estimates can be adversely affected by a number of factors, including inadequate calibration, beam attenuation, anomalous propagation, ground clutter, range effects, beam blockage, and frozen precipitation. Echoes in radar reflectivity data do not always correspond to precipitating particles. The largest errors in radar precipitation estimates appear to be caused by overshooting of precipitation by the radar beam and inadequate calibration (Hunter, 1996).

It is very difficult to accurately estimate frozen precipitation using radars. On a practical level, wet snow can accumulate on the radar antenna and degrade performance. Most of the uncertainty in accurately determining frozen precipitation rates is due to variations in the physical properties, specifically variations in the bulk density, although variations in particle shape (i.e., non-spherical particles) and size distributions contribute smaller errors (Matrosov et al., 2009). The characteristics of the frozen precipitation also affect the radar beam and returns. Beam attenuation in wet snow typically exceeds that for rain with comparable precipitation rates, while attenuation in dry snow is small and often ignored (Matrosov et al., 2009). ***Table 4*** presents a summary of radar reflectivity properties for varying types of snow.

Table 4. Radar Reflectivity Values for Varying Snow Types (Matrosov et al., 2009)

Type of Snow	**Air Temperature [°C]**	**Reflectivity [dBZ]**	**Differential Reflectivity [dB]**
Dry snow in aggregates	<0	20-35	<0.5-0.6
Wet snow	>0	>35	>0.6
Pristine single crystal snow	<-5 to -10	<20	>0.6

To correct for a wide range of errors in radar precipitation estimates, bias adjustments are applied using observed precipitation values at nearby stations. The bias adjustments are equivalent to adjusting the multiplicative factor used to compute precipitation rates from reflectivity (i.e., parameter *A* in ***Equation 1***).

$$Z = A(x, y, t) * R^b$$

Where:

Z = Rainfall rate (mm/hr)
A = Multiplicative factor (mm/hr)
R = Reflectivity (dBZ)
b = Exponent (unitless)

Equation 1. Z-R Relationship

A mean field bias adjustment applies a single multiplicative adjustment to all grid cells within a radar umbrella. This type of adjustment works well in areas with a sparse station network and a uniform bias over the radar umbrella. In order to reliably compute the mean field bias adjustment, a threshold must be established for the minimum number of grid cells in the radar umbrella that had precipitation recorded both by the radar and at nearby stations. A value of 20 has been recommended for this threshold (Seo et al., 1999). Applying a mean field bias

adjustment can exacerbate overestimation errors, particularly where radars overestimate rainfall due to brightband enhancement (Seo et al., 2002).

A local bias correction allows the multiplicative adjustment factor to vary for each grid cell within the radar umbrella. Adjustment factors are computed for each grid cell with a sufficient number of station observations within the specified radius of influence. The adjustment factors are then interpolated to the entire domain. The local bias correction is more effective than the mean field bias adjustment in areas with a dense station network and for stratiform rainfall events that occur during the cool season (Seo et al., 2002). In some cases, local bias adjustments have been shown to underestimate high rainfall amounts and to overestimate low rainfall amounts (Seo et al., 2002).

One of the major advances in radar precipitation technology is the deployment of dual polarization capability. Dual polarization aids in clutter removal and in distinguishing non-precipitating objects from precipitating particles. Precipitation estimates during the warm season due to convective activity are likely to improve significantly as a result of dual polarization (Earth System Research Laboratory [ESRL], personal communication, November 16, 2011).

In addition to precipitation detection and estimation, radar estimates are also beneficial for temporal disaggregation of precipitation estimates, spatial interpolation of point observations, and determination of the elevation separating rain from frozen precipitation.

3.3 Satellite

3.3.1 Precipitation

Precipitation rates determined from remotely-sensed data measured by satellite-mounted instruments are improving and are becoming more commonly used in hydrologic applications. Satellite-based precipitation estimates have the advantage of near-global coverage, although precipitation rates cannot currently be determined from satellite data for polar latitudes (i.e., north of approximately 50-60°N and similar in the south). Satellite precipitation estimates are valuable over oceans where alternative measurements are limited and in areas where gage and radar coverage is sparse. Although of lower utility for the AOR development in comparison to gage or radar data, satellite precipitation estimates could be incorporated for specific locations and time periods, and for spatial or temporal disaggregation. For example, the Hydroestimator geostationary infrared algorithm is used operationally for some areas of Mexico that drain to the U.S. (Scofield and Kuligowski, 2003).

Kidd and Levizzani (2011) provide a useful overview of the current state of satellite precipitation estimation, including estimation techniques associated with specific data types as well as techniques used to merge information from different satellite sources. Various sensors have been utilized for precipitation estimation, including visible, infrared, passive microwave, and active microwave (precipitation radar) sensors. The instruments are mounted on a range of satellite constellations, with a range of techniques employed to convert measured data into precipitation estimates. Visible imagery has been utilized in training algorithms to identify cloud types and subsequently to determine rainfall amounts associated with specific cloud types. Infrared (IR) measurements provide temperature data for cloud tops, which can be related to precipitation rates. However, the relationships vary widely, so although IR data are available at high spatial and temporal resolutions, precipitation estimates from IR data are less reliable than

other estimates. Passive microwave (PMW) instruments can be used to identify variations in microwave radiation emitted from the earth caused by scattering of the radiation by precipitating ice particles, and subsequently can be used to estimate precipitation more directly than with IR data. The combined PMW coverage from satellites has a lower spatial resolution than IR data and less frequent observations, with 3-6 hours between successive passes of PMW-equipped satellites over a given location. Finally, precipitation radar, such as that associated with the Tropical Rainfall Measuring Mission (TRMM), provides the most direct measurement of precipitation from space-borne instrumentation, similar to ground-based radars. However, precipitation radars have narrow swath widths and are placed on a relatively small number of satellites, with relatively infrequent passes over a given location. The data are used primarily to calibrate other instruments, but may also be useful for gap filling and temporal disaggregation of other data sources.

Various techniques have been developed to combine PMW and IR measurements and to utilize information from both sources to provide an improved satellite-based precipitation estimate. Kuligowski (2002), Behrangi et al. (2010), and Joyce and Xie (2011) describe algorithms used to combine PMW and IR information, as well as the limitations of these approaches. The techniques are categorized into four main groups:

1. Techniques that develop relationships between PMW and IR imagery to derive local mapping of IR measurements to rainfall rates (e.g., SCaMPR, Kuligowski, 2002).
2. Techniques that assign PMW-based rainfall estimates when available and IR-based rainfall rates when PMW measurements are unavailable.
3. Motion-based techniques that utilize PMW data to estimate rainfall rates and IR data to define cloud movement paths between each PMW rainfall estimate (e.g., CMORPH, Joyce et al., 2004).
4. Techniques that combine cloud movement information with cloud modeling.

Because of the wide range of satellite precipitation products that are available, the Program to Evaluate High Resolution Precipitation Products (PEHRPP) was established, from which the International Precipitation Working Group provides regular validation information on different satellite-based precipitation products over the U.S. (Ebert et al., 2007). Of the products produced on a regular basis, the motion-based techniques such as the Climate Prediction Center (CPC) Morphing Method (CMORPH; Joyce et al., 2004) and the more recent Global Satellite Mapping using a Moving Kalman Filter (GSMaP_MVK+; Ushio et al., 2009) perform better than the simpler combination techniques or single-sensor techniques (Sapiano et al., 2010; Tian et al., 2010). Recently, the CMORPH technique was modified to incorporate a Kalman filter and improve inter-calibration of PMW data, among other improvements (Joyce and Xie, 2011). The updated algorithms of the Kalman Filter CMORPH (KF-CMORPH) result in improved precipitation estimates for hours between successive PMW overpasses. The CMORPH product was recently reprocessed to incorporate better quality-controlled input data and to extend the period of record back to 1998; the CPC plans to process an equivalent period of record for the improved KF-CMORPH product before the expected launch of the first Global Precipitation Mission (GPM) satellites in 2013 (CPC, personal communication, December 20, 2011).

Recent developments in cloud modeling appear promising, improving upon motion-based precipitation estimation through inclusion of cloud classification and cloud modeling information (e.g., the Lagrangian Model [LMODEL; Bellerby et al., 2009] and Rainfall Estimation using

Forward-Adjusted advection of Microwave Estimates [REFAME; Behrangi et al., 2010] techniques), although at the time of reporting these products were not available operationally. In addition, the CPC is working with the University of California-Irvine and the National Aeronautics and Space Administration (NASA) to integrate the cloud classification system of the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CSS) algorithm (Hsu et al., 2010), the post-processing monthly bias removal of the TRMM Multisatellite Precipitation Analysis (TMPA) system (Huffman et al., 2007), and the KF-CMORPH algorithm into the Integrated Multisatellite Retrievals for the Global Precipitation Mission (IMERG), with an anticipated completion date corresponding with the GPM launch date (Huffman et al., 2012).

Table 5 presents a summary of the satellite precipitation products that are recommended for consideration in generating the AOR. Of the operational products that will be available in the near future, the KF-CMORPH product incorporates many of the improvements of the GSMaP_MVK+ technique compared with the CMORPH technique. Furthermore, the KF-CMORPH product is preferable over the GSMaP_MVK+ product since it is produced by the CPC and forms the basis for the IMERG development, which should provide further improvements to the KF-CMORPH product. Depending on the AOR development timeline, the LMODEL, REFAME, or IMERG products may also be feasible alternatives if these are generated over the complete historical record and become operational products.

Table 5. Recommended Sources for Satellite Precipitation Data

Product Name	**Data Source**	**Grid Mesh**	**Temporal Interval**	**Data Availability**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
TRMM Level 2 Rainfall Rate	NASA	4-km	3-hour	1997-present	✓		✓	✓		✓
KF-CMORPH	NCEP/CPC	8-km	30-minute	1998-present	✓		✓	✓	✓	✓
CMORPH	NCEP/CPC	8-km	30-minute	1998-present	✓		✓	✓	✓	✓
IMERG	NCEP/CPC/ NASA	Under development								
LMODEL	UC-Irvine	Under development								
REFAME	UC-Irvine	Under development								

Various data quality issues are associated with the visual, IR, PMW, and precipitation radar data. The data quality is impacted by individual sensor calibration and sensor drift issues, beam angle, and incorrect data classification due to a wide range of factors (Kidd and Levizzani, 2011; Joyce and Xie, 2011). In particular, satellite-based precipitation estimates are less accurate over snow and ice due to ice-contamination issues of PMW detections (Sapiano et al., 2008). Once precipitation estimates are derived for individual sensors, differences in precipitation estimates from different sensors need to be resolved. One of the improvements of the KF-

CMORPH algorithm compared to the original CMORPH algorithm is an improved accounting for differences between sensor precipitation estimates (Joyce and Xie, 2011). The satellite sensor constellation has changed over the available period of record, affecting the uncertainties associated with the precipitation estimates. Work presented by Joyce and Xie (2011) indicates that efforts are underway at the CPC to quantify the impact of changing satellite constellations on uncertainty estimation; the paper presents preliminary results of this investigation.

Although satellite precipitation algorithms are improving, the quality of the precipitation estimates remains relatively low compared to alternative sources of information due to the indirect nature of the precipitation estimates and the quality issues described previously. Although the satellite coverage extends to approximately 60°N, the precipitation estimates are poor at the northernmost extents due to challenges in the estimation procedures over snow and ice. In one inter-comparison study, satellite-based precipitation estimates were shown to perform better than NWP models during summer months and for convective storms (Sapiano et al., 2010). However, NWP models performed better during winter months, when snow and ice cover is more likely, and during stratiform events.

Satellite precipitation estimates suffer from relatively large biases. The TMPA 3B42 v6 product includes a post-processing step that removes bias on a monthly basis using gridded station data (Huffman et al., 2007). Xie and Xiong (2011) perform a bias adjustment for the CMORPH product based on a daily gage analysis; the CMORPH product is adjusted based on matching probability density functions between the CMORPH and the daily gage analysis over a 5° box and a moving 30-day window. The NWS MPE program also performs mean field and local bias adjustments to satellite precipitation inputs (Seo, 2003; Fulton, 2005b; NWS, 2008).

Because of their large spatial extents, satellite-based precipitation estimates are valuable data sources in data-sparse regions such as the Rio Conchos Basin in Northern Mexico, which flows to the Rio Grande River on the U.S.-Mexico border. Likewise, satellite precipitation data may have utility for precipitation detection in areas with large gaps between precipitation stations. Because satellite precipitation estimates have relatively low absolute accuracy, the estimates should be used in the AOR development by creating merged multi-sensor products, possibly using uncertainty information to assign satellite estimates lower weight relative to higher quality data sources. The hourly temporal interval of the KF-CMORPH product could be valuable for temporally disaggregating daily precipitation estimates in areas lacking hourly rain gages or ground-based radar estimates.

Although there are challenges associated with satellite precipitation estimation, satellite instrumentation and estimation algorithms continue to improve. In 2013, NASA and the Japanese Space Administration are expected to jointly launch the GPM core satellite, initiating a multi-year science mission. The new satellites will significantly augment the current satellite precipitation constellation, and are expected to result in worldwide improvements in remotely-sensed precipitation. Considering ongoing improvements in satellite-based sensors and estimation algorithms, these data will likely hold a more prominent role in future precipitation estimation over the U.S. in data sparse areas. NOAA staff are preparing proposals for eventual transition of the GPM program to operations.

3.3.2 Temperature

Temperature estimates from satellite observations have long been used to analyze the magnitude of global climate change, and more recently to assess soil moisture conditions and

crop evapotranspiration requirements. Satellite observations can be used to estimate land surface temperatures using infrared and microwave instruments (Wan and Dozier, 1996). Polar-orbiting systems equipped with Advanced Very High Resolution Radiometers (AVHRR) and the Moderate Resolution Imaging Spectroradiometers (MODIS) have been used to estimate land surface temperatures because they produce estimates at high spatial resolutions. However, polar-orbiting satellites observe a given location only twice per day, and are adversely affected by cloud cover (Pinker et al., 2009; Jin and Dickinson, 1999). Geostationary satellites have the advantage of continuously observing a fixed (but limited) location, producing estimates with high temporal resolutions (Pinker et al., 2009). Many algorithms have been developed to estimate land surface temperatures from satellite observations. Single channel methods require that surface emissivities and temperature/water vapor profiles be specified. Split-window methods require that surface emissivities be specified, but correct for atmospheric effects based on the differential absorption in adjacent infrared bands (Wan and Dozier, 1996; Pinker et al., 2009). Single channel and split-window algorithms do not work well in arid and semi-arid environments (Wan et al., 2002). Additional methods extract relative emissivities from multi-spectral thermal infrared data (Wan et al., 2002). In general, the quality of satellite-based land surface temperature estimates tends to be adversely affected by heterogeneous land surfaces, clouds, falling precipitation, and changes in satellites and instrumentation throughout the historical record. ***Table 6*** presents a summary of the historical land surface temperature data that are available from the Aqua MODIS and Terra MODIS satellites.

Table 6. Satellite Land Surface Temperature Data

Product Name	**Data Source**	**Grid Mesh**	**Temporal Interval**	**Data Availability**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
MODIS Land Surface Temperatures	NASA	1-6 km	5-min to Monthly	2000-present	✓	✓	✓	✓	✓	✓

Many studies are being conducted to develop algorithms for estimating near-surface air temperatures from satellite-based land surface temperatures (e.g., Ceccato et al., 2010; Vancutsem et al., 2010; Lin et al., 2012). These studies employ regression models to predict maximum (or minimum) air temperatures as a function of land surface temperatures, elevation, vegetation indices, and soil moisture estimates, among other variables. The estimation of daily minimum air temperatures from land surface temperatures retrieved at night tends to be more straightforward, as solar radiation does not affect the thermal infrared signal (Ceccato et al., 2010; Vancutsem et al., 2010; Lin et al., 2012). The estimation of daily maximum air temperatures from land surface temperatures retrieved during the day tends to be more variable, as the relationships vary as a function of season, ecosystem, solar radiation, and cloud cover (Vancutsem et al., 2010).

For the purposes of the AOR development, air temperatures derived from satellite observations may be less valuable in terms of absolute values, except in areas with very sparse station networks. However, satellite-based temperature estimates may be beneficial for characterizing diurnal temperature cycles (e.g., Jin and Dickinson, 1999).

3.4 Numerical Weather Prediction Models and Reanalysis Data

NWP models represent the condition of the atmosphere, and in some cases, interactions among the atmosphere, oceans, sea ice, and land surfaces. NWP models are used to forecast weather conditions based on current conditions. Multiple groups have employed NWP models to produce retrospective datasets, referred to as reanalysis products. To produce the reanalysis products, the NWP models are initialized using historical conditions prior to the start of the reanalysis period. Forecast runs are made with the NWP models. Moving forward in time, observed data such as radiance, sea surface temperature, and pressure are assimilated into the models to maintain the model states reasonably close to historical conditions. New forecast runs are produced once the observed data are assimilated into the NWP model to produce updated model states. In this manner, an entire historical record of NWP model output can be generated.

Table 7 presents a summary of the reanalysis products that are recommended for investigation for the AOR. These reanalysis products can be further categorized into global reanalysis products and regional high-resolution reanalysis products. The reanalysis products provide a variety of weather variables, including precipitation, temperature, solar radiation, longwave radiation, relative humidity, and wind speed. Of the available variables, precipitation is the most difficult to simulate well using NWP models.

Table 7. Recommended Sources for Reanalysis Data

Product Name	**Data Source**	**Grid Mesh**	**Temporal Interval**	**Data Availability**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
Climate Forecast System Reanalysis (CFSR)	NCEP	36-km	Hourly	1979-2011	✓	✓	✓	✓	✓	✓
NCEP/NCAR Reanalysis (Version 1)	NCEP/ NCAR	2.5°	Daily, 6-hour	1948-present	✓	✓	✓	✓	✓	✓
North American Regional Reanalysis (NARR)	NCEP	32-km	3-hour	1979-present	✓	✓	✓	✓	✓	✓
Arctic System Reanalysis (ASR-Interim)	Ohio State Polar Meteorology Group	10-30 km	3-hour	2000-2008	✓	✓			✓	
ERA Interim	ECMWF	0.7°	3-hour	1979-2011	✓	✓	✓	✓	✓	✓
ERA-40	ECMWF	1.125°	6-hour	1957-2002	✓	✓	✓	✓	✓	✓

3.4.1 Global Reanalysis Products

The NCEP/NCAR Reanalysis (R1; Kalnay et al., 1996) was the first major, global reanalysis run in the mid-1990s. The dataset covered the period 1948-1996, with updates extending the record through the present. The NCEP/DOE Reanalysis (R2; Kanamitsu et al., 2002) was generated a few years later to correct various errors identified in the R1 dataset. The R2 dataset was limited to begin in 1979. Subsequently, the European Center for Medium-Range Weather Forecasts (ECMWF) generated the ECMWF 15-year Reanalysis (ERA-15; Gibson et al., 1997), covering the period 1979-1993. In subsequent years, ERA-40 (Uppala et al., 2005) was generated to replace ERA-15, covering the period 1957-2002 based on updated models and improved data assimilation techniques. The Japanese Meteorological Agency (JMA) produced the Japanese 25-year Reanalysis (JRA-25; Onogi et al., 2007), using many of the same algorithms as the ERA-40 reanalysis, but with improved input data over Asia. The original JRA-25 reanalysis covers the period 1979-2004, with updates using recent data. The JMA began production of an extended reanalysis, JRA-55, in August 2010, with an expected completion date in 2013. This reanalysis will cover the period 1958-2012 using an improved version of the JMA NWP model and improved data assimilation methods. The ECMWF has generated the ERA-Interim product (Dee et al., 2011) in preparation for a longer reanalysis extending back to the early 20th century. The ERA-Interim reanalysis focused on addressing many deficiencies identified in the ERA-40 model, observations, and data assimilation system. The product was originally run for the period 1987-2006, though it continues to be extended forward in time, and was recently extended back to 1979.

In the U.S., several global reanalysis products have been produced in recent years. The Twentieth Century Reanalysis (20CR; Compo et al., 2011; Compo et al., 2012; Chang, 2012) was produced by the NOAA Cooperative Institute for Research in Environmental Sciences (CIRES). This product differs from many of the other reanalysis datasets in that the observed data that were assimilated (i.e., surface pressure observations) and the data that were used as boundary conditions (i.e., sea surface temperatures and sea ice distributions) were limited to allow a consistent reanalysis to be generated further back in time, and to provide a means to assess the impact of additional data inputs in the satellite era. The 20CR dataset has been extended to begin prior to 1900. In evaluating the 20CR dataset, Compo et al. (2012) have found that the dataset provides higher quality estimates in the northern hemisphere than in the southern hemisphere.

The NASA Goddard Global Modeling and Assimilation Office (NASA-GMAO) recently completed the Modern Era Retrospective-Analysis (MERRA; Reinecker et al., 2011) reanalysis. The NCEP recently completed the Climate Forecast System Reanalysis (CFSR; Saha et al., 2010). Both products employ similar atmospheric models. However, the MERRA reanalysis utilized an older version of the data assimilation system, and is limited to an atmospheric model reanalysis, utilizing observed data for oceanic, sea ice, and land surface boundary conditions. The CFSR product includes coupled atmospheric, oceanic, sea ice, and land surface models. The MERRA reanalysis continues to be updated with new data. The CFSR product initially covered the period 1979-2009 and was recently extended through March 2011, although no additional updates are planned. Kumar (2012) reports that the CFSR dataset depicts high-frequency (i.e., synoptic) variability on a daily time step better than low-frequency variability.

Of the global reanalysis products, the ERA-Interim, MERRA, and CFSR products represent the most recent advances in NWP models, observational datasets, and data assimilation

methods. The CFSR product is generated at the highest spatial and temporal resolutions (approximately 36-km grid mesh and 1-hr temporal interval, compared to approximately 45-km grid mesh and 1-hr temporal interval for MERRA, and 80-km grid mesh and 3-hr temporal interval for the ERA-Interim product). However, Kistler (2012) reports that these datasets are affected by changes in observing systems, in particular the introduction of the special sensor microwave/imager (SSMI) instrument in 1987 and the advanced microwave sounding unit (AMSU-A) instrument in 1998. The ERA-Interim, MERRA, and CFSR datasets were produced by assimilating data from different channels from the two instruments, resulting in inconsistencies that vary in timing and magnitude in the reanalysis datasets.

Reinecker et al. (2011) compares the ability of each of the reanalysis products to estimate precipitation on a large scale compared to the Global Precipitation Climatology Project (GPCP; Adler et al., 2003) and to the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997) observed precipitation products. The GPCP and CMAP products incorporate satellite- and ground-based observations to produce gridded precipitation estimates on a monthly basis. In Reinecker et al. (2011) and associated studies, the reanalysis products yielded comparable results. It was also demonstrated that changes in the observational network resulted in major changes in the reanalysis precipitation estimates on a monthly basis. Although the latest reanalyses show significant improvements in precipitation estimation relative to the original reanalyses, Reinecker et al. (2011) concluded that the reanalyses do not provide additional information concerning precipitation at this time compared to the GPCP or CMAP precipitation products due to the difficulties of data assimilation.

Much like satellite-based temperature estimates, the temperature fields from the global reanalyses have been used to detect and quantify climatic trends. The temperature fields have been compared to gridded observational datasets on relatively coarse spatial and temporal scales. The comparisons have employed temperature anomalies to minimize the effect of biases in the datasets and to focus on evaluating low-frequency temporal variations. For example, Simmons et al. (2004) compared anomalies in monthly mean surface air temperatures at 2-m from the ERA-40 and NCEP/NCAR reanalyses to a gage-based, gridded observational dataset (CRUTEM2v) developed by the U.K.'s Climatic Research Unit. The comparison was done for the period 1958-2001 using a 5° grid mesh, with results aggregated to hemisphere and continental scales. The authors concluded that the three datasets demonstrated similar inter-annual variability, but that the ERA-40 temperature fields tended to be closer to the observed data after 1967 in North America. Prior to 1967, the quality of the ERA-40 reanalysis suffers from a lack of synoptic-level data for assimilation. In addition, the authors reported that suspect values occurred to a limited extent in both the observational dataset (due to erroneous station values) and the reanalysis datasets. Chelliah and Ropelewski (2000) compared average monthly and seasonal temperature anomalies between the NCEP/NCAR reanalysis and gridded station observations. This comparison was performed on a 5° grid mesh for the period 1958-1996. The authors reported seasonal correlations ranging from 0.65-0.85, with higher correlations occurring in winter (December-February), and lower correlations occurring in summer (June-August) and over mountainous regions such as the Rockies. Finally, the authors report that a change in relative bias occurs in the NCEP/NCAR and ERA-15 reanalysis datasets in 1991-1997 due to a change in the satellite data inputs used in the reanalyses. Although these studies provide some insight into the potential quality of the reanalysis temperature fields, additional investigation is required at the spatial and temporal resolutions intended for the AOR.

Given the spatial and temporal resolutions, recent advances in modeling and data assimilation techniques, and comparable performance between the CFSR, MERRA, and ERA-Interim products, the CFSR is likely to be most applicable for the AOR. The R1 and ERA-40 products may also be applicable for the AOR for the period 1948-1979 if the AOR is extended to include that period, unless recent reanalysis datasets are extended to cover the earlier period.

3.4.2 Regional and High-Resolution Reanalysis Products

In addition to the global reanalysis products, numerous regional reanalyses have been completed that are relevant to the AOR. The North American Regional Reanalysis (NARR; Mesinger et al., 2006) was originally completed for the period 1979-2003 by NCEP and continues to be updated in near-real time, with the motivation to improve the precipitation fields and to increase the spatial and temporal resolution of R2 over North America. The grid mesh of the NARR is approximately 32 km, a significant improvement over R2. For the NARR, observed precipitation data were assimilated indirectly into the NWP model by adjusting the latent heat state to force precipitation to be simulated when it was observed. The assimilation of observed precipitation data was generally successful, and the NARR product better represents precipitation compared with previous reanalysis products. However, the input precipitation data were gridded point observations. Thus, the result is impacted by how the station data were interpolated to produce gridded products.

The Ohio State Polar Meteorology Group recently completed a preliminary 8-year reanalysis using the Weather Research and Forecasting (WRF) model on a 30-km grid mesh for the period 2000-2008. This preliminary reanalysis was intended as a test prior to generating the Arctic System Reanalysis (ASR) on a 10-km grid mesh for the period 2000-2010 over the Arctic region (Wang et al., 2011). Though not yet complete, the ASR product will provide a high-resolution reanalysis input for Alaska that may be valuable for the AOR.

3.4.3 Issues and Benefits

The reanalysis products have multiple potential applications for the AOR development. Although input datasets have changed over the period of record intended for the AOR, the NWP models and associated data assimilation algorithms provide a consistent means to simulate the state of the atmosphere over the period of record. For any given location and time, the associated atmospheric conditions are represented, so the unique conditions that generate rainfall and temperature conditions are represented. However, there are limitations with the reanalysis products. Most importantly, the grid mesh of the products that are available over the desired domain for the AOR is limited to 32-38 km. Within one grid element of this resolution, there are over 1,000 nested 1-km grid cells. Identifying appropriate means to downscale the reanalysis fields to a finer spatial resolution is a challenge. Furthermore, there are known limitations with the precipitation fields in the reanalysis products. Although the NARR addresses this issue to some degree through indirect data assimilation, differences between the observations and the NARR precipitation estimates still exist (Bukovsky and Karoly, 2006; West et al., 2007; Becker et al., 2009).

Although the precipitation outputs from the reanalyses may not be directly usable, the reanalysis products could be used for other purposes. Reanalysis fields besides precipitation and

temperature are necessary for the AOR, such as wind direction, wind speed, and water vapor content. These variables may be utilized to characterize atmospheric conditions, allowing for specific techniques to be applied to downscale precipitation or temperature fields. The temperature or precipitation fields may be useful for temporal disaggregation if alternative sources are not available. In places with frozen precipitation, the reanalysis products may provide the best estimate of the rain-snow elevation.

With respect to simulated precipitation, it is unclear whether the NARR or the CFSR product is preferable for use in the AOR. Given the advances in modeling and data assimilation techniques between the NARR-era reanalysis and the CFSR-era reanalysis, the CFSR likely provides a superior representation of atmospheric variables that could be used within the AOR framework. Over its available record and extent, the ASR appears to be the best reanalysis product.

3.5 Gridded Climatology and Observation Datasets

A number of gridded precipitation and temperature datasets are currently available, including both time series and multi-decade climatologies, that can be used for developing the AOR. These datasets may be useful in terms of applying the same methodologies, using the gridded products to establish the long-term statistical characteristics of the AOR, or using the gridded products as validation datasets. ***Table 8*** presents a summary of the gridded datasets that are recommended in these capacities for developing the AOR.

Table 8. Gridded Data Sources

Product Name	**Data Source**	**Grid Mesh**	**Temporal Interval**	**Period of Record**	**CONUS**	**AK**	**HI**	**PR**	**CN**	**MX**
PRISM 400-m	The Climate Source, Inc.	400-m	30-yr average	1971-2000	✓					
PRISM 2-km	The Climate Source, Inc.	2-km	30-yr average	1961-1990	✓	✓	✓	✓	✓	
PRISM 4-km	The Climate Source, Inc.	4-km	Monthly	1895-present	✓					
Atlas Climatico Digital de Mexico	Uniatmos/ CONAGUA	1-km	50-yr average	1950-2000						✓
Prototype monthly climatology datasets	NCDC	4-km	30-yr average	1978-2007	✓				✓	✓
Hamlet / Lettenmaier	University of Washington	1/8°	Daily	1915-2003	✓				✓	

Product Name	Data Source	Grid Mesh	Temporal Interval	Period of Record	CONUS	AK	HI	PR	CN	MX
Maurer et al.	Santa Clara University	1/8°	Daily, Monthly	1949-2010	✓				(1)	✓
Stage 2 National Precipitation Analysis (gage-only)	NCEP	4-km	Daily	1996-present	✓	✓	✓	✓		
US Daily Precipitation Gridded Analysis	CPC	0.25°	Daily	1979-present	✓					
Gauge-Based Analysis of Global Daily Precipitation	CPC	0.50°	Daily	1979-present		✓	✓	✓	✓	✓
Daily Gridded Temperature Analysis	CPC	0.50°	Daily	1948-present	✓	(2)	(2)	(2)	(2)	(2)
Climatology-Calibrated Precipitation Analysis (CCPA)	NCEP	4-km	6-hr	2002-2010	✓					
NLDAS-2 Forcing	NASA / NCEP	1/8°	Hourly	1979-present	✓				✓	✓
WorldClim Global Climate Data	WorldClim	1-km	50-yr average	1950-2000	✓	✓	✓	✓	✓	✓
Daymet	University of Montana	1-km	Daily	1980-1997	✓					

(1) Available in the original dataset, but not yet available for the updated dataset.
(2) A global analysis dataset for 1979-present is scheduled for completion in 2012 (CPC, personal communication, February 2012).

3.5.1 PRISM Datasets

Several gridded precipitation, maximum temperature, and minimum temperature products have been created using the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) developed at Oregon State University. These products vary in grid mesh from 400-m to 4-km and in time step from monthly grids to 30-year climatological averages. The finest

resolution products are available for the CONUS, although coarser resolution products are available for Alaska, Hawaii, Puerto Rico, and portions of Canada. The PRISM system is described in more detail in ***Section 10***.

Riverside has successfully used PRISM datasets to establish the long-term characteristics of MAP datasets developed for calibration of hydrologic models, particularly in the intermountain west. In this capacity, Riverside has employed the 2-km, 1961-1990 climatological average grids and has developed climatological average grids for varying calibration periods using the 4-km, monthly grids. To date, Riverside has not utilized the 400-m, 1971-2000 climatological grids.

3.5.2 Maurer et al. and Hamlet/Lettenmaier Datasets

Significant work has been done at the University of Washington and the Santa Clara University to develop gridded archives of precipitation, maximum temperature, and minimum temperature that have a grid mesh of 1/8° and temporal intervals of three hours, one day, and one month. These datasets cover most of the CONUS and portions of Mexico and Canada. The methodologies that have been applied to develop these datasets may be useful for developing the AOR.

The precipitation datasets that were developed by Maurer et al. (2002) used daily observations at NCDC COOP stations. Outside of the U.S., missing precipitation values were filled using the Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations data (GPCP 1DD; Huffman et al., 2001). The point observations were interpolated using the synergraphic mapping system (SYMAP) method (Shepard, 1984). The interpolated grids were adjusted using the 1961-1990 climatology grids from PRISM so that the long-term averages of the grid cells containing COOP stations are consistent with PRISM. The adjustments using PRISM were performed by developing twelve monthly adjustment factors for each grid cell. Daily precipitation values were evenly distributed to produce 3-hour values. (Maurer et al. [2002] compared uniform disaggregation to stochastic disaggregation, and concluded that the assumption of a constant diurnal cycle had minimal effect on the moisture and energy fluxes simulated by the variable infiltration capacity [VIC] model.)

To produce temperature grids, daily observations of maximum and minimum temperatures at the COOP stations were interpolated to a grid mesh of 1/8° using the SYMAP algorithm. After 1998, the dataset uses the 2-m temperature data from the NCEP/NCAR R1 reanalysis to fill missing temperature values outside of the U.S. Lapse rates were applied to account for differences in elevation. A spline interpolation technique was applied to the daily temperature values to produce the 3-hour grids.

The initial datasets developed by Maurer et al. (2002) were later improved upon by Hamlet and Lettenmaier (2005). The Hamlet and Lettenmaier dataset improves upon the Maurer et al. dataset by making an adjustment for temporally-varying biases introduced over the period of record from changing networks. The additional adjustment is made by comparing the gridded precipitation estimates using all the COOP stations to gridded precipitation estimates produced solely using HCN gages.

Both the Maurer et al. and Hamlet/Lettenmaier datasets have been used for hydrologic modeling with the VIC model (Maurer et al., 2002; Hamlet and Lettenmaier, 2005). The primary limitation of the datasets is that they rely solely upon NCDC COOP observations in the U.S. Additionally, for some days, there is a strong discontinuity at the border between Mexico and the

U.S. (Maurer, 2011). This is attributed to differences in the data sources between the two countries and the sparse station density along the border. Both the Maurer et al. and Hamlet/Lettenmaier datasets are being updated to produce longer periods of record for larger spatial extents.

3.5.3 NCEP Datasets

The CPC and the Environmental Modeling Center (EMC) within NCEP produce multiple gridded precipitation and temperature products based on station observations. These datasets have been used as validation datasets and are recommended in this capacity for the AOR (Joyce et al., 2004; Sapiano et al., 2008; Rienecker et al., 2011), in addition to having been developed using techniques that are relevant to the AOR development.

The NCEP Stage 2 National Precipitation Analysis includes both hourly and daily gage-only precipitation products. The hourly products incorporate observations from approximately 3,000 automated stations that transmit via the ASOS and the Geostationary Operational Environmental Satellites (GOES) Data Collection Platform (DCP; Katz, 2002). The daily products incorporate observations from more stations than are used in the hourly products, and also benefit from quality control that is performed at the RFCs. These precipitation products are available for the CONUS.

The CPC produces daily precipitation products for the CONUS (i.e., 0.25° grid mesh) and globally (i.e., 0.50° grid mesh). These products are developed using station observations from the World Meteorological Organization (WMO) Global Telecommunications System (GTS) and Standard Hydrologic Exchange Format (SHEF) encoded reports from the RFCs, Hydrometeorological Automated Data System (HADS), and the NRCS SNOTEL network (NWS, 2006; NCEP, 2012).

The CPC also produces daily grids of maximum temperature and minimum temperature on a 0.5° grid mesh for the CONUS. A historical dataset has been developed for the period 1948-2010 (CPC, 2012a). The historical dataset uses point observations from the COOP Summary of the Day reports produced by NCDC. The historical observations underwent both automated and manual quality control checks. The point observations were interpolated to the 0.5° grid mesh using the Cressman (1959) technique. The CPC reports that limitations in the dataset include data smoothing and limited quality in areas with sparse observations (CPC, 2012a). Real-time grids are produced using similar techniques for the CONUS (CPC, 2012b). A global temperature analysis with a 0.5° grid mesh and a daily time step for the period 1979-present is scheduled for completion in 2012 (CPC, personal communication, February 2012).

The Climatology-Calibrated Precipitation Analysis (CCPA) dataset is another dataset that has been developed by NCEP to retain the climatology from the CPC Global Daily Gauge Analysis while incorporating the high spatial and temporal resolutions associated with the NCEP Stage IV multi-sensor analysis (Hou et al., 2010).

3.5.4 NLDAS-2 Datasets

The North American Land Data Assimilation System (NLDAS) Version 2 utilizes a variety of forcing inputs, including precipitation and temperature, which have been prepared using a 1-hr temporal interval and a 1/8° grid mesh. The forcing datasets were developed for the

period 1979-present and have a spatial extent that covers central North America (NASA, 2011). Given the period of record and reasonably high resolution, the NLDAS-2 forcing datasets are relevant to the AOR in terms of data sources used and methodologies applied.

The forcing inputs for NLDAS-2, with the exception of precipitation, were derived from the NARR reanalysis dataset (Cosgrove et al., 2003). The 32-km, 3-hr fields from the NARR were interpolated to a 1/8° grid and were temporally disaggregated to a 1-hr temporal interval. The surface pressure, longwave radiation, air temperature, and specific humidity fields were adjusted vertically to account for terrain differences between the NARR and NLDAS. The vertical adjustment applied a lapse rate of 0.65°C/100 m for air temperature.

For the precipitation input to NLDAS-2, the CPC daily precipitation product was used as the starting point. The daily estimates were temporally disaggregated using the NCEP Stage 2 radar product, CMORPH satellite estimates, the CPC hourly precipitation dataset, or the NARR precipitation fields (in that order of priority) depending on data availability for specific periods of time.

Since daily and hourly precipitation stations are sparse over Canada, precipitation from the NARR was used over Canada within the NLDAS domain. For a 1° swath along the U.S.-Canada border, the precipitation input to the NLDAS is a weighted combination of the CPC-derived value and the NARR value.

3.5.5 Other Datasets

A gridded climatological dataset has been produced for Mexico using a similar methodology to the PRISM system. The Atlas Climatico Digital de Mexico includes 50-year climatological averages for the period 1950-2000 for precipitation, maximum temperature, and minimum temperature (UNIATMOS, 2011). Riverside has successfully utilized this dataset to establish long-term characteristics for calibration MAPs and MATs developed for the Rio Sabinal Basin.

The WorldClim datasets have been developed on a 1-km grid mesh to represent climatological averages for the period 1950-2000. Climatological datasets were developed for precipitation and minimum, maximum, and average temperature. Average monthly climate data from stations were interpolated using thin plate smoothing splines, with latitude, longitude, and elevation serving as independent variables (Hijmans et al., 2005).

The Numerical Terradynamic Simulation Group at the University of Montana developed the Daymet dataset to support plant growth models. Datasets of precipitation, temperature, humidity, and radiation were developed on a daily time step for the period 1980-1997 using point observations and a digital elevation model (Daymet, 2012).

An experimental 4-km gridded climatology for monthly average precipitation, minimum temperature, and maximum temperature has been provided to OHD staff by NCDC (Russell Vose, NCDC, personal communication, 2011). The grids cover the conterminous U.S., southern Canada, and northern Mexico. Documentation for the dataset is being developed but is not yet available.

3.6 Supplemental Data Sources

In addition to the temperature and precipitation data sources included in the data inventory, additional data sources may prove beneficial for the AOR for the purposes of validation and improved quality control. For example, soil moisture data from stations, satellites, and land surface models can be used to validate remotely sensed observations of precipitation from radars and satellites. Lightning data can be useful for delineating areas of convective and stratiform events. Areas experiencing convective precipitation are likely to show lower spatial correlations, which has implications for data quality control, interpolating point observations, estimating missing precipitation values, and making bias adjustments to radar and satellite observations.

3.7 Criteria for Data Integration into AOR

In developing the data inventory and determining the potential utility of each data source for the AOR, Riverside applied the following criteria:

- The data source should have a sufficient record in the historical archive to aid in establishing long-term precipitation and temperature characteristics. The required length of the historical record varies. For example, a shorter period of observed data at a station may be acceptable if a nearby station is highly correlated and has a long historical record that can be used to estimate missing values.
- Data sources that have known quality limitations should be considered for use in areas with sparse data coverage.
- Input data sources produced using a single sensor are favored over multi-sensor estimates, if the previously applied data fusion method is not fully transparent. This principle is based on the assumption that the AOR will benefit from increased data availability and improved quality control methods and multi-sensor merging techniques than have been applied historically.
- The general effort required to collect, reformat, and quality control the data were weighed against the potential value of the data source in improving the AOR.
- In some cases, it may be more efficient to obtain quality-controlled or processed data rather than the raw data from source agencies.

Data sources that are recommended for use or for further investigation may require spatial or temporal restrictions. For example, certain rain gage networks that lack heating elements, or radar data, may not be suitable for estimating frozen precipitation. Regional and temporal issues are further explored in ***Section 11***.

4 Data Quality Control

The data sources that are used to produce the AOR must undergo quality control to remove erroneous data and to correct for inconsistencies that occur over the historical record. All observations are subject to errors. Inconsistencies in a data record may occur due to changes in

locations, sensors, or algorithms, or drift if the instrument is not recalibrated sufficiently frequently. It is preferable to perform data quality control on the input time series data from individual sensors rather than the final output grids (Gibson et al., 2002).

Data errors and temporal inconsistencies can have a significant adverse impact on hydrologic simulations. Data errors can affect soil moisture and snowpack states long past the time steps in which the erroneous data occurred. Temporal inconsistencies can produce varying biases in the simulated streamflows for different periods of time. While the hydrologic models used by the RFCs contain parameters to address long-term biases (i.e., precipitation and evapotranspiration multipliers, temperature lapse rates), the parameters cannot be used to make seasonal adjustments or to rectify inconsistencies that vary throughout the historical record (Anderson, 2002). In addition, it is critical to ensure that biases are minimized between the historical inputs used for model calibration and the real-time inputs used for operational modeling.

The SNOW-17 snow accumulation and ablation model and the SAC-SMA soil moisture accounting model are sensitive to biased inputs. The effects from biased inputs are nonlinear; a given bias in precipitation may cause a larger bias in the simulated streamflows. Small temperature errors that are difficult to detect during quality control (e.g., ±2°F) have a significant impact on hydrologic model simulations during the snowmelt runoff period (***Figure 4***).

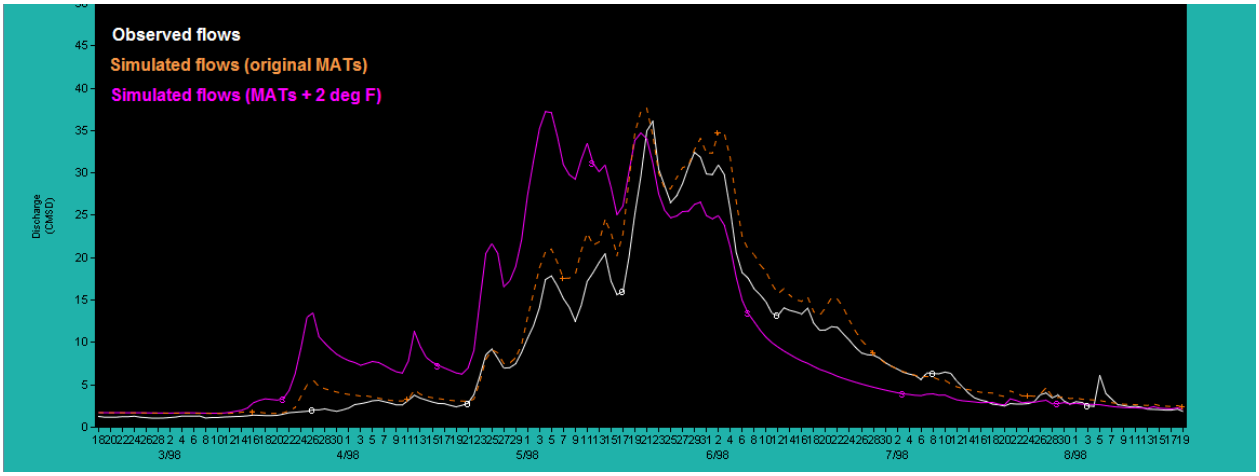

Figure 4. Effect of Temperature on Simulated Streamflows for the South Fork Rio Grande at South Fork, CO

Manual review of data observations is quite effective; the human eye can spot outlying values, periods with reduced or increased variability, or repeating values in a time series quickly when plotting multiple time series together. Spatially, anomalous values can be easy to spot as a "bull's-eye," or a value that is inconsistent with values at nearby locations. Given the intended spatial and temporal extent of the AOR, it is desirable to implement a data quality control system that is automated to the extent possible. However, because rainfall is highly variable in space and time, it is difficult to develop fully automated quality control tools (Kondragunta and Shrestha, 2006). Typically, automated checks are developed to flag suspicious values, which are then automatically removed if the values are clearly erroneous, or manually reviewed if additional judgment is required (Angel et al., 2002; Gibson et al., 2002; Fulton, 2005b; Riverside, 2010b; Riverside, 2011).

In locations where precipitation and temperature data vary dramatically over short distances due to elevation changes or proximity to coastlines, it can be difficult to perform data quality control. It is always a risk that naturally occurring extreme values may be removed, resulting in reduced variance in the final dataset (Seo and Breidenbach, 2002). To minimize this outcome, Riverside adheres to the following guidance when developing calibration MAPs and MATs: "The underlying rule when making [data] adjustments should be that if there is any doubt that the correction should be made, then don't make the adjustment (Anderson, 2002)."

4.1 Types of Quality Control Checks

Data sources can introduce errors into estimated precipitation and temperature in a number of ways:

- ***Erroneous metadata.*** Erroneous metadata, including incorrect identifiers, latitude, longitude, elevation, sensor equipment, and dates that sensors were reset or recalibrated, can cause misrepresentation of weather characteristics for a location. Metadata errors are more impactful in areas with steep topographic relief because of the variation in weather characteristics over short distances. Incorrect identifiers can mask duplicate datasets. Identifiers can also mask the fact that a station was moved, or obscure the fact that records associated with multiple identifiers should be combined. While station histories contain extremely valuable information regarding location changes and changes to sensor equipment or algorithms, it may be too laborious to review all metadata in detail prior to developing the AOR. Riverside recommends that metadata checks be automated to the extent possible, for example by intersecting station locations and a digital elevation model (DEM) to identify large differences between the computed and the assigned elevation attributes. Additionally, plotting long-term monthly, seasonal, and annual characteristics against elevation can help to identify metadata errors.
- ***Unreliability.*** Data sources with large amounts of missing data prevent the reliable computation of long-term characteristics. The PRISM system requires a minimum of 85% non-missing data for a station to be included in an analysis for a given month (Daly et al., 2008). In the NCDC temperature validation (TempVal) system, stations that are missing nine or more days in a given month, or five consecutive days, are excluded to avoid skewing estimated monthly characteristics (Angel et al., 2002). In developing calibration MAPs and MATs, Riverside has required a minimum of 5-10 years of data within the calibration period to estimate long-term climate characteristics (Riverside, 2010b; Riverside, 2011). Additionally, Riverside may eliminate data sources with significant periods of missing data if a nearby data source situated at a similar elevation has a high quality dataset.

  Similarly, hourly observations at SNOTEL stations tend to show high levels of variability and inconsistency with daily values. The daily observed values are significantly more reliable (NWRFC, personal communication, February 27, 2012). The reliability of the observations is improved by the periodic revisions performed by the NRCS, typically once or twice per month. Riverside typically uses only the daily historical observations from SNOTEL stations in developing calibration MAPs and MATs.

- ***Seasonal limitations.*** Several data sources (e.g., stations with no heating elements, radars, satellites) have issues measuring frozen precipitation accurately, typically resulting in underestimation of the actual precipitation that occurred. For these data sources, it may be necessary to set all data values to missing when air temperatures are below freezing. Another common issue is that sensors may fail to register frozen precipitation until the precipitation melts, causing the precipitation to be observed during the incorrect time interval (CNRFC, personal communication, February 27, 2012). This type of issue is difficult to detect using automated quality control procedures, and will likely result in erroneously removing the precipitation values.

Erroneous data values can be identified using a variety of quality control checks:

- ***Range checks.*** The most basic quality control check evaluates each observation against a reasonable range of values to identify obvious errors. For example, all precipitation values should be greater than or equal to zero. Maximum temperature values should be greater than minimum temperature values for the same time interval. Additional range checks can be established for combinations of variables, for example the ratio of water equivalent to precipitation depth, or the ratio of frozen precipitation to total precipitation.
- ***Climatological range checks.*** In addition to basic range checks, data values can be evaluated for reasonableness using climatological range checks, or thresholds computed from historical climate data. Reasonable ranges will vary by location and time of year. In applying climatological range checks, it is assumed that the historical climate is a good indicator of future weather characteristics. The climatological range checks can be run on multiple time steps, for example both daily and monthly values, to reduce the variability in the observed data and to provide confirmation that the flagged data are erroneous.

  The PRISM system flags potentially erroneous data using climatological range checks. Daily precipitation values that are greater than 115% of the state's record daily precipitation are flagged, while minimum (maximum) temperature values are flagged if they are 3°C less (greater) than the state's record daily values (Daly et al., 2008). The NCDC TempVal system flags temperature values if they differ significantly from climate extremes or climate normals that are computed for stations (Angel et al., 2002).
- ***Internal consistency checks.*** Data can be evaluated using internal consistency checks, for example checking that daily and hourly precipitation characteristics are consistent for a given data source (Riverside, 2011). Maximum and minimum temperature characteristics can be plotted against elevation to check that the relationships reflect reasonable lapse rates and inversion conditions, and that the lapse rates reflect physically-realistic seasonal variations (Anderson, 2002).
- ***Spatial consistency checks.*** Spatial consistency checks are effective for flagging anomalous data values, particularly for variables that exhibit high spatial correlations, such as temperature and stratiform precipitation. In developing calibration MAPs, Riverside has produced automated scripts that flag daily and monthly precipitation values that differ significantly from values recorded at surrounding stations (Riverside, 2011). For example, daily and monthly precipitation values may be

flagged for review if they are a factor of two larger (or smaller) than the precipitation values observed at nearby stations. The thresholds are established based on the variability in the weather characteristics for the area being modeled.

Spatial consistency checks are less effective for quality controlling convective precipitation, which exhibits low spatial correlations. The NWS OHD has developed a real-time technique to retain flagged data values if a lightning strike occurred within one hour and 10 km of the station (Kondragunta and Shrestha, 2006). The NSSL NMQ/Q2 system utilizes lightning data to delineate areas of convective and stratiform precipitation (Vasiloff et al., 2007; NSSL, 2011b).

- ***Multi-sensor checks.*** The development of automated multi-sensor quality control checks is an active area of research and development. Multi-sensor quality control techniques have primarily been applied for developing precipitation estimates using station observations and radar data. However, satellite data and NWP model outputs are also being integrated. For each data source, statistical characteristics are computed such as the probability of rainfall detection, correlations between data sources, and the ratios and differences between paired observations. Thresholds are established for the computed statistics to identify and remove questionable data sources and outlying data values (Seo and Breidenbach, 2002). Multi-sensor quality control checks are often performed separately for the warm and cool seasons.

  In the real-time NWS MPE system, gages are assumed to be malfunctioning when radar detects precipitation, but a co-located gage reports no precipitation (Fulton, 2005b; Kondragunta and Shrestha, 2006). The NCDC precipitation validation (PrecipVal) system compares precipitation estimates from multiple station networks, WSR-88D radars, GOES satellites, and the Rapid Update Cycle (RUC) NWP model to identify data outliers (Urzen et al., 2004).
- ***Checks between predicted and observed values.*** The PRISM system uses spatial interpolation and multivariate regression techniques to develop expected values to which station observations are compared (Daly, 2006). The process developed for the PRISM system (ASSAY_QC) applies cross validation to identify potential data errors on a monthly time step. The differences between the predicted and observed values are normalized to account for the fact that higher variances are expected for higher precipitation values (Gibson et al., 2002). The quality control process is used to set the observed values to missing, to replace the observations with the predicted values, or to linearly combine the observed and predicted values (Daly et al., 2005).
- ***Temporal consistency checks.*** Repeating data values and temporal inconsistencies in the historical record can be identified using temporal consistency checks. While weather variables (particularly temperature) are expected to be temporally correlated, exact values that repeat for an extended period of time typically indicate quality issues. For maximum and minimum temperatures, values that repeat for fewer than five days are common and likely naturally occurring, values that repeat for more than ten days are typically caused by erroneous data that should be set to missing, and values that repeat for 5-10 days are difficult to assess (Daly et al., 2005). The PRISM system flags temperature values if the temperature differences are less than 0.1°C over a 10-day period. In developing calibration MAPs and MATs, Riverside has produced automated scripts that flag values for review if the same station value repeats for more than five days or for more than ten hours (Riverside, 2010b).

Riverside typically does not remove repeating values from the dataset if the repeating value is the minimum reporting level for the data source.

Double mass analysis can be used to identify inconsistencies in the historical record caused by changes in location, sensor equipment, and the surrounding environment. Noticeable shifts in the double mass analysis plots can indicate natural variability, outlying data, or temporal shifts. Therefore, it is important to review the history of the data source before making changes. For example, NRCS SNOTEL stations in Colorado and New Mexico underwent changes in temperature sensors in the 2000s. This change in instrumentation manifests as a change in slope on the double mass analysis plot (***Figure 5***).

The results from the double mass analysis should be used to correct inconsistencies prior to inputting the data into a hydrologic model, to avoid having different biases between one portion of the historical record and another (Anderson, 2002). The data are typically corrected to make the earlier period consistent with the recent period (i.e., the current location and instrumentation), particularly if the data source is still in operation. Riverside has established ranges for reasonable correction factors applied to individual stations. For example, Riverside has typically avoided making precipitation changes larger than a factor of two or temperature adjustments larger than ±5°F (Riverside, 2010b; Riverside, 2011). If the double mass analysis conducted for individual stations indicates that larger adjustments are required, Riverside often chooses to eliminate the station due to suspect data quality prior to computing calibration MAPs and MATs.

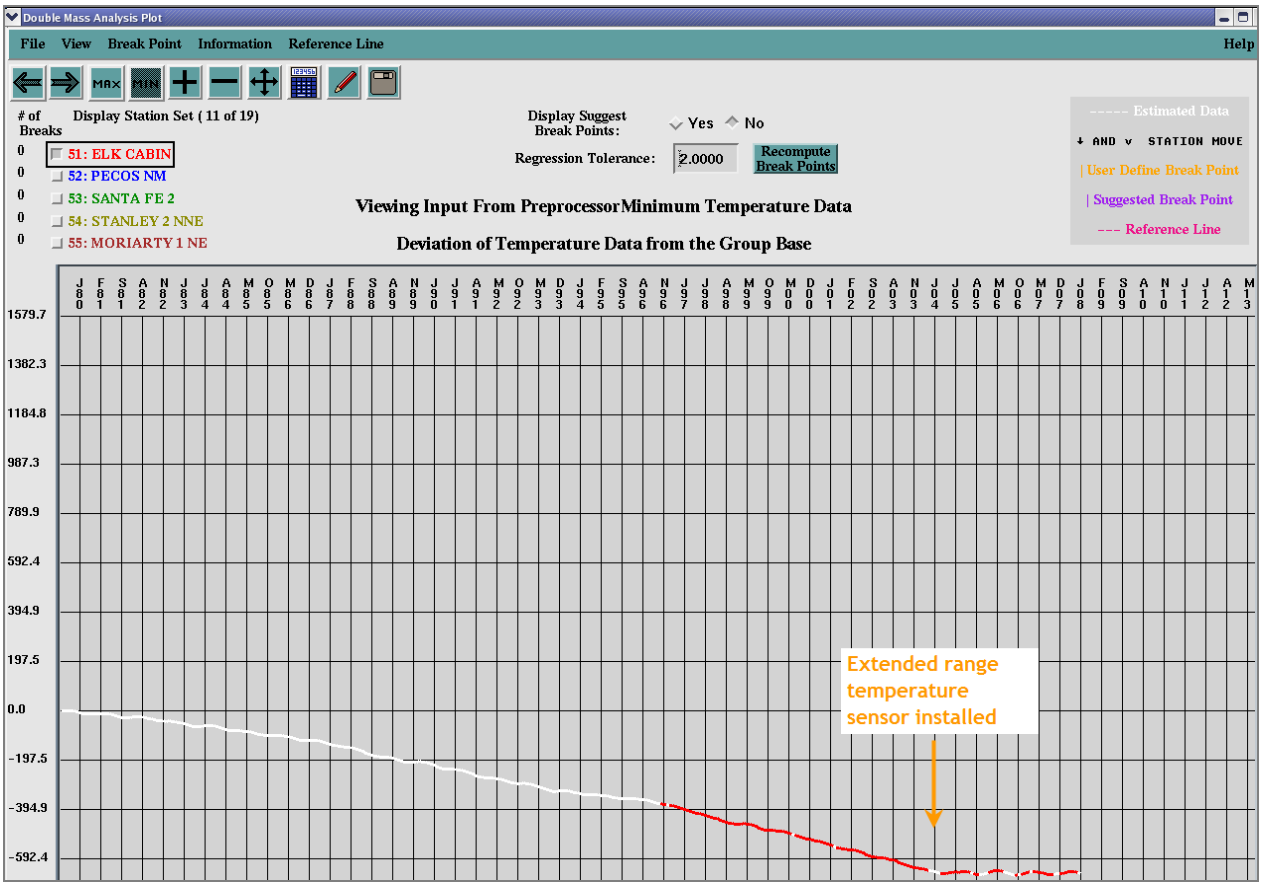

Figure 5. Double mass analysis plot showing the effect of changes in temperature sensors at SNOTEL gages

The best quality control approach requires manual review and expert judgment to assess the reasonableness of outlying values and to incorporate knowledge about the location, data source, and weather event. For example, the PRISM system categorizes flagged data values into clear outliers that are automatically removed and questionable values that require manual review to determine the appropriate action (Gibson et al., 2002). For developing calibration MAPs and MATs, Riverside applies a similar approach, where data values that are flagged using basic range checks are automatically set to missing, while data values flagged using other techniques are manually reviewed to determine the appropriate action (Riverside, 2010b; Riverside, 2011). The extent to which manual reviews are feasible for developing the AOR depends upon available resources and the development timeline.

4.2 Considerations Prior to Removing Data

Before removing data values from a dataset that have been flagged during the quality control process, the following factors should be considered:

- It is difficult to perform quality control on hourly precipitation data due to weather pattern movement and high spatial and temporal variability. It may be more effective to perform quality control on a daily or monthly time step.
- Data values that have been flagged by comparing predicted and observed values, or using spatial consistency checks, should be considered for retention if large differences in distance or elevation exist between the point being evaluated and the comparison points, or if the precipitation amounts are small (Gibson et al., 2002).
- Data values that have been flagged in a data sparse region should be considered for retention. It is difficult to provide a high level of quality control in these areas due to a lack of corroborating information (Daly et al., 2005).
- Observation times and reporting times must be accounted for to ensure the data values being compared represent the same time interval. Data that are reported from manual observing stations often have timestamps in local time observing daylight savings. Automated data tend to be reported in standard time with no adjustments for daylight savings. It is best to adjust all data to a common time zone, or if this is not possible, to perform the quality control using a longer time step (e.g., monthly).

  At times, the data values that are reported with a particular timestamp actually apply to the previous day, and may need to be shifted for consistency with other data sources. For example, daily maximum and minimum temperature values reported on the NRCS website for SNOTEL stations actually apply to the previous day (NRCS, 2012).

  In the legacy NWSRFS system, the calibration MAP and MAT preprocessors account explicitly for station observation times (NWS, 2012). For example, for stations with morning observation times, the MAT preprocessor assumes that the maximum temperature occurred on the previous day while the minimum temperature occurred on the current day. For stations with afternoon observation times, the MAT preprocessor assumes that both the maximum and minimum temperatures occurred on the current day. The MAP preprocessor distributes precipitation values to the 24 hours ending at the observation time. The AOR will need to similarly explicitly account for observation times.
- Special data codes that indicate missing data or quality issues should be reviewed prior to performing data quality control. It is best to convert all input data to the same format and to apply the same data codes. For example, the historical data files for NCDC COOP stations use a value of -999 to indicate missing values, while a code of -998 indicates that the subsequent precipitation value represents an accumulated total over the time intervals indicated with the -998 code.

  In addition, the data codes should be evaluated for potential errors. The NWS OHD has found cases where NCDC digital precipitation reports confuse multi-day and single-day accumulations due to miscoding of missing report indicators (OHD, personal communication, February 12, 2012). These types of issues might require manual effort to identify and correct the errors.
- The weather conditions should be considered when evaluating flagged data values. It is difficult to accurately quality control data when convective activity, inversions, warm air advection, strong frontal passages, or wedging along coastlines occur, because these phenomena create strong gradients over short horizontal distances

(Angel et al., 2002). Spatial consistency checks should be limited under these conditions.

4.3 Recommendations for the AOR

For the development of the AOR, it will be important to include as much meaningful historical data as possible to satisfactorily estimate temperature and precipitation at high spatial and temporal resolutions. Data values that are removed or replaced during the quality control process should be archived along with notes about why the data values were removed or altered. Data values that were flagged but retained should include notes about the quality issues to convey lower confidence in the final estimated values. The raw and the quality-controlled data should be made available to users, along with supplemental quality control information such as climatological ranges that were applied and time series plots of the raw and final time series (Daly et al., 2005).

Depending on time constraints, it will be potentially difficult to perform significant manual review of all inputs. The quality control processes should be automated to the extent possible to minimize the need for manual review of flagged values, while acknowledging that this approach may permit some bad data values to enter into the AOR. Many organizations, including the NWS RFCs and Weather Forecast Offices (WFOs), the National Operational Hydrologic Remote Sensing Center (NOHRSC), NCDC, NCEP, NASA, and NSSL have devoted significant resources to collecting and quality controlling weather data. To the extent possible, the OHD should capitalize on work that has already been performed by collecting historical data, metadata, and quality control methods and processes from these organizations. In particular, it would be informative to collect information from the RFCs regarding common quality control issues that are currently performed manually but could be automated with additional development efforts (CBRFC, personal communication, February 22, 2012).

As automated quality control systems continue to evolve in the future, it is possible that the AOR may need to be re-generated to incorporate updates and to maximize consistency with the real-time forcings process (Daly et al., 2005).

5 Data Assimilation and Data Fusion

The requirements for the AOR specify high spatial and temporal resolutions over an extended period of record. In order to meet the high resolution requirements, the AOR development process should incorporate as much data as are available and reasonable to capitalize upon the relative strengths and weaknesses of the different data types (***Table 9***). Each data source has varying potential to inform precipitation detection and quantification, temperature estimation, and spatial and temporal downscaling processes.

Table 9. Strengths and weaknesses of different data sources (reproduced from Vasiloff et al., 2007)

Sensor	Strengths	Weaknesses	Time–space scales	Operational applications
Weather radar	• High spatial and temporal resolution • Good areal coverage	• Range effects • Coverage in complex terrain • Z_e–R and Z_e–S uncertainties • Nonmeteorological target contamination	• 5–10 min • 1 km	• Precipitation nowcasting • Flash flood forecasting • River forecasting (after bias correction) • Land surface modeling (after bias correction)
Geostationary satellite	• Continuous spatial coverage	• Indirect measurements of precipitation • Difficulty with nonprecipitation clouds	• 15 min • 4 km	• Nowcasting • Flash flood forecasting • River forecasting (after bias correction) • Land surface modeling (after bias correction)
Polar-orbiting satellite (passive microwave)	• Continuous spatial coverage	• Poor spatial/temporal resolution • Indirect measurements of precipitation • Difficulty with non-ice-bearing clouds	• 3–6 h (6+ satellite constellation) • 15 km	• Tropical Rainfall Potential (TraP) • Adjustment of GOES precipitation estimates
Precipitation gauge	• Direct measurement of precipitation	• Nonuniform spatial distribution • Latency in real-time data transfer • Problems with quality of measurements • Frozen hydrometeors • Wind effects • Uncalibrated (tipping-bucket type in high rain rate)	• 10 min–1 day	• Flash flood forecasting • River forecasting • Land surface modeling • Water supply forecasting • Correction of radar and satellite QPE • Hydroclimatological studies • Verification

This section summarizes various approaches that may be used to combine observations from multiple data sources. Data assimilation focuses on the use of a NWP model to perform this combination (e.g., reanalysis). Observed data are ingested into a data assimilation system to adjust model states and produce model outputs that more closely represent the true value over a grid cell. In contrast to data assimilation, data fusion defines a simplified model to merge multiple data sources to produce an improved estimate of a variable. The approaches are not strictly distinct, as data fusion may utilize outputs from NWP models, which in turn may benefit from data assimilation to improve the quality of the NWP model output.

5.1 Data Assimilation

The data assimilation process incorporates information from multiple sensors into a combined product by assimilating station, radar, and/or satellite estimates into a NWP model. The NWP model represents the underlying physics that control weather and reflects atmospheric

characteristics that determine when and where precipitation occurs and how temperature varies over a landscape.

Extensive work has been applied to develop methods that assimilate observations into NWP models. The NWP models employed for the different reanalysis products utilize 3- or 4-dimensional variational data assimilation procedures to incorporate a range of observed values, including radiance, sea surface temperatures, and mean sea level pressure. In variational assimilation, a least-squares problem is formulated as the minimization of an objective function with weights based on error covariances of model variables and observations. The 4-dimensional variational data assimilation is in effect a smoothing algorithm, because it operates over a finite number of model time steps.

Unlike the aforementioned variables, precipitation observations cannot currently be assimilated directly into NWP models because precipitation is a diagnostic model output resulting from a combination of model states that represent atmospheric conditions at particular locations. Instead, precipitation observations are indirectly assimilated by adjusting other model states, such as the latent heat of the atmosphere, in an attempt to force the model to produce precipitation at the location and time represented by the observation. Although this form of data assimilation can improve the precipitation fields generated by NWP models compared with earlier reanalysis products (as was accomplished with the NARR), the precipitation fields are not as accurate as the observed fields in locations with gage or radar observations (Bukovsky and Karoly, 2006; West et al., 2007; Becker et al., 2009). Research efforts continue to improve the indirect assimilation of precipitation observations, but at this point in time, alternative merging approaches are required to fully utilize observed data (Colorado State University [CSU], personal communication, December 21, 2011). The precipitation outputs from NWP models still represent one estimate of the true precipitation, with associated uncertainties that can be used in data fusion. If additional high-resolution weather (e.g., WRF) modeling is pursued for the AOR, precipitation data may be indirectly assimilated as a part of the modeling to improve the outputs. Current versions of the WRF model can indirectly assimilate gage and radar precipitation observations (ESRL, personal communication, November 16, 2011).

5.2 Data Fusion

Data fusion combines inputs from multiple sources into a single product, often in a stepwise fashion. Typically, the steps include interpolating station observations to produce a grid (***Section 6***) and downscaling grids to equivalent resolutions (***Section 7***). The merged product can be disaggregated to produce grids on shorter time intervals (***Section 8***). Data fusion can involve one or more of the approaches described in the following sections to merge multiple estimates into a single product that more accurately reflects the true field.

5.2.1 Bias Adjustments

A basic data fusion technique involves using one dataset (e.g., rain gages) to remove the bias associated with a second dataset (e.g., radar precipitation fields). The mean field bias adjustments and local bias adjustments for radar estimates based on gage observations are described in ***Section 3.2***. Bias adjustment methods applied for satellite precipitation are

described in ***Section 3.3***. Similar approaches could be employed to perform bias corrections for precipitation estimates from reanalysis datasets.

5.2.2 Simple Data Fusion

The MPE program is used by the RFCs to apply bias adjustments to radar and satellite precipitation products (***Section 10.6***). The program also includes tools to quality control and to adjust the final precipitation field. Within the program, various input datasets can be specified for subsequent use. Through the use of polygon replacement tools, different input datasets can be utilized for regions when the default field is clearly incorrect, resulting in a manually-merged precipitation estimate.

A second example of simple data fusion includes developing weighted averages of gridded inputs from different sources, often employing a transformation of the inputs prior to merging.

5.2.3 Optimization-based Data Fusion

Although simple data fusion techniques have value, the value of the data fusion process can be increased through optimization of weighting parameters and transformation functions. Various data fusion methods can be classified as optimization-based methods. Thin plate splines, regression-based approaches, and artificial neural networks fall under this broad category of data fusion. Equations and parameters relating each dataset are determined through some form of optimization to minimize errors in the output dataset. These methods do not require *a priori* specification of error covariances. One limitation of the optimization methods is that they require an observed dataset for parameterization purposes. For the AOR, an observed gridded dataset is not directly available, though a high-quality input could be treated as the known quantity to optimize the merging of other inputs (e.g., using radar data to assess data fusion of satellite and NWP model outputs).

5.2.3.1 Artificial Neural Network Framework

Artificial neural networks (ANNs) can be used for various purposes in data fusion, including the classification of information from input fields and the optimization of merging parameters. ANNs can effectively identify complex relationships and patterns between input variables that can be difficult to identify using alternative approaches. ANNs define classifications and relationships between variables through a training phase where the inputs are compared with a known output. Once the training phase is complete, and the hidden transformation layer and weights are optimized, a recognition phase is employed to utilize the defined transformations and weights for the remainder of the record. This approach has limitations because the true field is unknown, yet the approach could be used to define relationships for certain aspects of the final merging technique.

For example, Chiang et al. (2007) combined MAP estimates derived from gages and from satellite data through a simple linear combination of the two inputs. The relative weights associated with the MAP estimates were adjusted to optimize the quality of a hydrologic simulation within an ANN framework.

The PERSIANN-CCS satellite precipitation algorithm employs an ANN to classify different cloud types and to calibrate infrared-based precipitation estimates using available passive microwave data (Hsu et al., 2010).

Turlapaty et al. (2010) utilize an ANN in combination with a preliminary vector space transformation to merge multiple satellite-based precipitation products developed using different estimation techniques. The merged product represented an improvement over the input products based on a comparison with the MPE estimates developed at the Arkansas Basin River Forecast Center (ABRFC).

5.2.3.2 Variational Assimilation

Variational assimilation can be viewed as an optimization-based data fusion method. The major difference from the previously described methods is that variational assimilation requires estimates of the model and observation error covariances. Although variational assimilation is primarily associated with real-time data assimilation for NWP models, it can be applied as a smoothing process as part of a reanalysis. Variational assimilation has been applied for other purposes as well. For instance, Seo et al. (2009) apply variational assimilation to adjust soil moisture states in the SAC-SMA model through assimilation of precipitation, evapotranspiration, and streamflow data.

5.2.4 Uncertainty-based Data Fusion

In contrast to optimization-based approaches, uncertainty-based approaches (e.g., kriging, optimal interpolation, and Kalman filtering) utilize explicit definition of error characteristics to merge multiple data sources. If the error characteristics can be accurately specified, each of the analyses yields an estimate that minimizes the sum of the squared errors. Kriging and optimal interpolation are optimal linear estimators that are commonly used for generating gridded estimates of gage data. Weights are computed by minimizing the squared error term based on an assumed knowledge of the spatial covariance relationships typically estimated from sample data.

5.2.4.1 Kriging

Kriging is a geostatistical interpolation technique that is often applied to the interpolation of hydrometeorological data. Several forms of kriging have been developed. Ordinary kriging is the most common method. It assumes stationarity in the mean as well as intrinsic stationarity. The variogram is used to define the spatial variability of the random field. Universal kriging allows the specification of a trend surface for the mean. Cokriging provides the framework for a multi-variable approach, and has been used to merge radar and gage data (Krajewski, 1987; Seo, 1998a; Seo, 1998b; Goovaerts, 2000).

5.2.4.2 Optimal Interpolation

The optimal interpolation framework developed by Gandin (1965) develops a final estimate by modifying a first-guess field by the weighted difference between observations and the first guess. The weights are defined based on the background and observation error

covariance relationships. Optimal interpolation is extensively used for data assimilation into NWP models.

Xie et al. (2007) apply optimal interpolation for the gridding of gage observations as described in ***Section*** 6. Xie and Xiong (2011) use optimal interpolation to merge satellite precipitation estimates with a gage-based precipitation grid. First, the bias in the satellite precipitation estimates is removed based on gage data. Then, the bias-corrected satellite dataset is combined with the daily gage analysis using optimal interpolation. The process requires estimates of errors and the spatial correlation of errors for each input field. The error in the gage analysis is specified empirically as a function of station density. The error in the bias-corrected satellite data is defined empirically to vary with precipitation intensity. The spatial correlation of errors is defined empirically as a correlation-distance function. Luo et al. (2011) are extending the work of Xie and Xiong (2011) to an hourly time step over eastern China. Luo et al. (2011) are also including radar data in the data merging process as follows: the radar and satellite estimates are merged to produce a remotely-sensed data field, which is then bias-adjusted and merged with an hourly gage analysis using optimal interpolation.

Sapiano et al. (2008) applied optimal interpolation to merge satellite precipitation estimates with the ERA-40 NWP model reanalysis. Errors associated with each dataset were determined empirically through comparison with a third global precipitation dataset. The errors were defined as a function of latitude to weight the satellite precipitation estimates more heavily in the tropics and the reanalysis estimates more heavily in the polar regions to capitalize on the strengths of each input.

Kummerow and Sapiano are working to improve upon the merging of satellite and NWP model precipitation estimates through improved quantification of errors and error correlations, and by recognizing that atmospheric conditions affect the errors and error correlations for models and observations (CSU, personal communication, December 21, 2011).

5.2.4.3 Kalman Filter Framework

The Kalman filter provides a recursive framework for combining observations with model state estimates of a linear system through the computation of a Kalman gain assuming Gaussian error distributions. Kalman filtering involves a forecast and an analysis step. In the forecast step, a linear state-space model is utilized to compute the state of a variable and the associated error variance based on the previous conditions. In the analysis step, observations and associated error statistics are used to refine the forecast. Although computationally intensive, the Kalman filter provides a direct estimate of the error covariance of the forecasted state.

The Kalman filter methodology has been applied to non-linear systems using the extended Kalman filter approach. The extended Kalman filter uses a local linearization of the state-space equations to propagate the state error covariance. More recently, the ensemble Kalman filter has been used to estimate the state error covariance for complex non-linear modeling systems.

The mean field bias and local field bias corrections employed in the MPE program represent a simplified form of Kalman filtering (Seo et al., 1999; Seo and Breidenbach, 2002; Seo, 2003).

The KF-CMORPH algorithm utilizes a Kalman filter to combine PMW and IR precipitation estimates, accounting for differences in error characteristics associated with different satellite sensors and the length of time between PMW overpasses (Joyce and Xie,

2011). The KF-CMORPH algorithm is a step towards a final goal of developing a pole-to-pole product that merges satellite precipitation estimates from PWM and IR sources with NWP model outputs, and potentially other sources (Joyce and Xie, 2011; CPC, personal communication, December 20, 2011).

The Multiscale Kalman Smoother algorithm can be used to fuse inputs of different spatial or temporal resolutions. Wang et al. (2011) evaluate the use of a Multiscale Kalman Smoother-based framework to fuse gridded precipitation inputs of different spatial resolutions, and investigate the impact of white noise and bias errors through hypothetical experiments. In their investigations, Wang et al. found that the data fusion process increases the accuracy of the resulting grid both when the output is a coarse-resolution grid as well as when the output is a fine-resolution grid. Wang et al. (2011) includes references to multiple applications of the Multiscale Kalman Smoother for precipitation data fusion (e.g., Gorenburg et al., 2001; Gupta et al., 2006).

Ebtehaj and Foufoula-Georiou (2011) recognize that precipitation does not have a Gaussian distribution in either normal or lognormal space. In order to account for the unique characteristics of precipitation, Ebtehaj and Foufoula-Georiou first transform the input data into the wavelet domain prior to applying a Kalman filter to merge satellite and radar precipitation fields.

Bowler et al. (2006) describe the Short-Term Ensemble Prediction System, an ensemble-based probabilistic forecasting system designed to merge advected radar observations with forecast outputs from NWP models to produce a precipitation nowcast. The ensemble process is used to account for uncertainty in the radar precipitation advection and the NWP model spatial downscaling. The radar fields are advected using motion vectors that include a random component to characterize the uncertainty. The coarse-resolution NWP model output is downscaled using a technique that incorporates stochastically-generated small-scale features to represent uncertainty. An ensemble of noise is generated to represent the loss of forecast skill moving forward in time. The data fusion process utilizes the ensembles of advected radar fields, downscaled NWP model outputs, and noise to produce a forecast ensemble that can be used to derive uncertainty information for the merged fields.

5.3 Implications for the AOR

5.3.1 Framework

Riverside recommends that an uncertainty-based approach using optimal interpolation, cokriging, or Kalman filtering form the basis for a data fusion framework for the AOR. ANNs or other optimization techniques may be employed for specific development aspects, such as quantification of uncertainties or error correlations that may be difficult to determine otherwise.

Uncertainty-based data fusion methods require the quantification of uncertainties. Many factors make this quantification difficult. The spatial and serial correlation of weather variables varies across space and time depending upon the physiographic characteristics of a region and the particular atmospheric conditions occurring at any point in time. Furthermore, the spatial and serial characteristics of weather vary depending upon the spatial and temporal scales under consideration, with higher variability associated with smaller spatial scales and shorter time steps (Daly, 2006). All of these factors affect the spatial and temporal correlations of errors associated

with particular sensors. For instance, precipitation recorded at a station in relatively flat terrain under stratiform conditions will likely reflect high spatial correlations that decay slowly with distance. However, under convective conditions, the spatial correlations for the same station will likely show a much steeper degradation with distance. In complex terrain, the spatial correlations will vary directionally. Serial correlations are high on an hourly time step, but relatively low on a daily time step. Spatial and serial correlations will be stronger in certain directions depending upon atmospheric conditions such as the predominant wind direction.

Multiple factors affect the errors associated with each of the major data sources (***Section*** 3). For gages, the sensor and observer type, frequency of maintenance, site-specific characteristics, and form of precipitation all impact the gage accuracy and representativeness. For radars, the radar type, surrounding terrain, atmospheric conditions, precipitation intensity, and form of precipitation affect the accuracy of the radar precipitation estimates. For satellites, the sensor type, timing and location of PMW overpasses, snow/ice contamination, and other factors affect the error associated with the precipitation estimates. For NWP models and reanalysis products, the accuracy varies for different atmospheric conditions and terrain characteristics, as well as the range of data available for model assimilation.

Currently, the quantification of uncertainties associated with precipitation observations is not well developed and is a major area of ongoing research. For example, Dr. Chris Kummerow is working on quantifying observational uncertainties and defining uncertainties as a function of atmospheric states (CSU, personal communication, December 21, 2011). However, Sapiano et al. (2008) suggests that even with simplistic, empirical uncertainty estimates, fused analyses yielded improved precipitation estimates compared with the input datasets. With the ongoing efforts to improve error quantification, it seems appropriate to construct the AOR under the assumption that information from each of the major data sources will be merged into a single product. Although the uncertainty information used to perform this merger may be preliminary for the initial AOR development, the framework will permit incorporation of higher-quality uncertainty estimates as they become available.

5.3.2 Stepwise Application of Data Fusion

Data fusion is often completed in a stepwise fashion to simplify the computations and inputs needed at each step of the process. For instance, rather than directly merging point observations with gridded estimates, station data are often interpolated to a grid that is then combined with other gridded data sources (e.g., Xie and Xiong, 2011; Luo et al., 2011). The order of merging is dependent upon the characteristics of the input datasets and the fields being generated.

Riverside recommends that a stepwise data fusion process be applied for the AOR, rather than attempting to merge all inputs in a single processing step:

- Downscale the input radar, satellite, and reanalysis grids to the target grid mesh.
- Perform bias corrections and merge the downscaled radar, satellite, and reanalysis grids.
- Grid the point observations to produce the target grid mesh.
- Merge the gridded gage analysis with the gridded radar/satellite/reanalysis field.

Alternatively, uncertainty and correlation characteristics could be defined for individual stations so that point observations could be directly merged with the remotely-sensed data, eliminating the need to grid the point observations.

5.3.3 Separate Daily and Hourly Data Fusion Processes

Serial correlations can be neglected, or treated in a simplified fashion, on a daily time step compared to an hourly time step. The spatial distribution of daily precipitation is much smoother than hourly precipitation, with stronger spatial correlations that persist for greater distances. Hourly precipitation can be serially correlated across many time steps, but daily data typically do not exhibit strong serial correlations.

In addition to reduced serial correlations and increased spatial correlations, the daily gage network is denser than the hourly network. For many users, it will be more important to have accurate estimates of daily precipitation, maximum temperature, and minimum temperature than hourly estimates.

For these reasons, Riverside recommends that a daily analysis be completed as a preliminary processing step, aggregating all sub-daily data sources to a daily time step. In a subsequent step, the daily analysis can be disaggregated to an hourly time step using hourly data sources or stochastic techniques. Methods for temporal disaggregation are addressed in ***Section*** 8.

6 Gridding Methods

Few of the existing data sources are available at the intended spatial and temporal resolutions of the AOR. Although not a required step, point observations are often transformed to gridded estimates prior to merging the point data with other data sources. There are a number of factors that need to be considered when developing gridded values from point observations:

- Point observations are affected by sampling errors, most notably catch deficiencies, that prevent point observations from being perfectly representative of even the 4-8” diameter area being measured.
- Point measurements can differ significantly from the areal average value of the grid cell containing the station. Topographic features cause small-scale variability in weather patterns that affect how representative a point value is of the surrounding environment. In data-sparse areas, if gages are installed in locations that are not representative, the grid estimates created from the point observations may not represent the variation in weather characteristics over the larger area.
- Spatial correlations, serial correlations, and weather patterns are very different for different temporal and spatial scales. For example, at monthly time steps, precipitation values show relatively smooth patterns of spatial variability that exhibit strong relationships with physiographic characteristics. At daily or hourly time intervals, precipitation characteristics exhibit high spatial variability that is primarily controlled by atmospheric conditions associated with the specific weather event. Likewise, higher variability is associated with finer spatial resolutions. Temperature values typically

exhibit less variability than precipitation at short time intervals, but can still exhibit high variability at fine spatial resolutions.

This section presents an overview of methods to develop gridded estimates from point observations.

6.1 Methods that Ignore Physiographic Characteristics and Atmospheric Conditions

Gridding methods that do not account for physiographic characteristics and atmospheric conditions can yield reasonable results in locations with high station densities and minimal climatic gradients. However, these techniques do not account for physiographic characteristics (e.g., orographic effects, coastal proximity) or storm-specific atmospheric conditions (e.g., moisture content, wind direction) that affect weather processes. Some of these methods retain the exact values at the station locations (e.g., thin plate splines), while others do not (e.g., kriging, optimal interpolation).

The most basic type of gridding method estimates cell values from point observations considering only gage locations. This category of methods includes inverse distance weighting (IDW), which has multiple possible formulations for the distance weighting function, and often includes a maximum radius of influence beyond which stations are not weighted. The natural neighbor method determines weights based on representative areas, similar to Thiessen polygons (Sibson, 1981). Several methods apply techniques to reduce weights on stations in close proximity, to avoid over-representing a particular area in developing the grid cell value. For example, Shepard (1968) incorporates a directional correction to weight isolated gages more heavily than closely-spaced gages. Recent applications of this approach include the SYMAP method (Shepard, 1984; Maurer et al., 2002; Hamlet and Lettenmaier, 2005) and the Angular Distance Weighting method (Hofstra and New, 2009). The techniques employed by Cressman (1959) and Barnes (1964) refine the distance-weighted interpolations using multiple passes, with each refinement incorporating finer-resolution information in areas with high gage densities (e.g., Chen et al., 2002). Thin plate splines are used to fit a surface through observed point data in a manner that prevents sharp breaks at observed data points (Hutchinson, 1993). Parameters controlling the smoothness of the resulting surface can be adjusted through optimization to minimize cross validation errors.

A more advanced category of gridding methods, including kriging and optimal interpolation, explicitly accounts for the spatial covariance of point observations (Cressie, 1993; Gandin, 1965; Daly, 2006; Hofstra et al., 2009). An advantage of these methods is that they generate error estimates associated with the interpolated grids. However, the error estimates should be used with caution, as the quality of the estimates depends on the validity of the assumptions associated with the interpolation approach (Daly, 2006).

Numerous studies have demonstrated that interpolation methods that account for the spatial covariance of observations perform better than methods that consider only the locations of gages (e.g., Chen et al., 2002; Hofstra et al., 2008). However, in these studies, approaches based on Shepard (1968) were also shown to perform well.

6.2 Methods that Incorporate Physiographic Characteristics

Physiographic characteristics influence the distribution of precipitation and temperature over an area. Variables such as elevation, aspect, latitude, longitude, and distance from major waterbodies have been included in regression-based gridding techniques to capture these influences. The relationships between physiographic variables and precipitation or temperature can be developed across an entire domain (e.g., generalized operator equations) or within limited areas (e.g., regionalized operators) (Glahn et al., 2009; Charba and Samplatsky, 2011a; Charba and Samplatsky, 2011b). The challenges of developing regression methods that account for physiographic influences include determining predominant controls on weather patterns and identifying appropriate transitions between regions with different controlling variables.

The PRISM system is an excellent example of a method that interpolates point observations using regression models that incorporate physiographic effects (Daly et al., 2002). The PRISM system accounts for the effects of elevation, aspect, coastal proximity, and station clustering, among other variables. The PRISM system has been applied primarily to develop monthly precipitation and temperature grids, as well as long-term climatological averages, but has also been applied on a daily time step for limited regions (Daly et al., 2007).

Some gridding methods make use of datasets that have already been developed to account for physiographic effects, rather than developing independent relationships. An example method is climatology-aided interpolation (CAI), which interpolates point observations using a climatology dataset, for example based on historical station data or developed using the PRISM system. The general procedure for CAI is as follows:

- For each station, a ratio (or difference) is computed between the observation and the climatology value at the station location. Typically, ratios are computed for precipitation, while differences are computed for temperature. The ratio (or difference) fields are referred to as anomalies.
- The anomalies are interpolated using basic gridding methods (i.e., those that ignore physiographic characteristics and atmospheric conditions).
- The interpolated anomaly grid is multiplied by (or added to) the climatology grid to generate the final interpolated precipitation (or temperature) grid.

Any of the interpolation methods described in ***Section 6.1*** can be applied for CAI. Two of the major assumptions of the CAI method are that the anomalies are more spatially homogeneous than the point observations and that applying the interpolated anomalies to the climatology dataset results in a final interpolated grid that benefits from the information accounted for in the climatology grid, such as the effects of physiographic variables. Multiple studies have shown that interpolating precipitation anomalies yields improved gridded fields compared to interpolating precipitation observations (e.g., Chen et al., 2002; Hofstra and New, 2009). Additionally, CAI has the benefit of producing interpolated grids that reflect differences in average characteristics (as represented by the climatology), even at locations where no observations are available.

CAI can be applied using annual, seasonal, or monthly climatology grids. Variability in regional and seasonal weather patterns suggests that monthly climatology grids, such as those available from the PRISM system, should be used. However, using monthly climatology grids can produce sharp discontinuities at the transitions between months. A possible solution was

developed by Xie et al. (2007), who computed climatology grids for each day of the year from smoothed average daily precipitation distributions adjusted to the PRISM monthly climatology grids. This approach provided a smooth transition across months while maintaining the long-term monthly characteristics reflected in the PRISM climatology grids.

The CAI method assumes that the spatial variability on shorter time intervals is the same as reflected in the climatology grid. Therefore, the CAI method can be used to reproduce long-term weather patterns over an area, but this technique does not produce realistic weather patterns for individual weather events in the absence of a dense station network.

Most of the gridding methods that are used to interpolate temperature observations, particularly in areas of complex terrain, incorporate lapse rates to account for elevation differences between stations and prediction locations. The specified lapse rates can be constant values (e.g., 6-6.5°C/km), constant values that vary by month, or dynamic values that are computed using observed station data for each time interval. Dynamic lapse rates are advantageous in that they capture atmospheric conditions affecting temperature patterns and transient conditions such as temperature inversions, but they do not perform well in areas with sparse station networks and insufficient high-elevation observations. For example, Stahl et al. (2006) found that dynamically computed, local lapse rates outperformed other methods at high elevations in recent years, but performed poorly prior to 1965 due to a lack of high elevation data. In some applications, the lapse rates are applied to adjust all temperature observations to a common elevation prior to interpolation, and are then used to adjust the interpolated values back to the local elevations.

6.3 Methods that Incorporate Physiographic Characteristics and Atmospheric Conditions

Gridding methods that incorporate physiographic characteristics successfully capture typical effects created by relatively static features such as topography. However, dynamic atmospheric conditions also affect weather patterns. For example, storm directions influence where precipitation occurs. As a result, weather characteristics differ under differing atmospheric conditions.

Numerous applications have incorporated the effects of atmospheric conditions into interpolation schemes. One approach is atmospheric classification, which involves identifying significant atmospheric classes, categorizing historical observations according to the defined classes, and developing interpolation parameters for each classification (e.g., Daly and Taylor, 2006). A second approach is to develop regression models that incorporate independent variables that act as surrogates for the controlling atmospheric conditions (e.g., Glahn et al., 2009; Guan et al., 2005).

6.3.1 Atmospheric Classification

The PRISM system has been used to develop “pseudo-climatology” grids, which are defined as long-term weather characteristics conditioned on an atmospheric classification (Daly and Taylor, 2006). Historical precipitation observations were classified based on the predominant wind direction at a specified level over the study area. In a similar fashion, various authors have investigated adjusting interpolation parameters (e.g., correlation decay functions)

for different atmospheric classifications (Hofstra and New, 2009; Hewitson and Crane, 2005; Tveito, 2002; Courault and Monestiez, 1999).

Theoretically, pseudo-climatology grids or classification-specific interpolation parameters can be utilized in the CAI method to incorporate information on atmospheric conditions and to improve the final interpolated grids. However, most studies have found that applying atmospheric classifications to gridding techniques resulted in minimal improvements in interpolation skill (Hofstra and New, 2009; Courault and Monestiez, 1999). This may be due to the selected atmospheric classifications, and demonstrates the difficulty in successfully identifying the atmospheric conditions that dominate weather patterns. It also may be a reflection of the limitations of CAI. If atmospheric classification approaches are applied to the AOR, they will need to be developed regionally, as the relevant classification schemes will vary by region.

6.3.2 Regression Methods

The Auto-Searched Orographic and Atmospheric Effects Detrended Kriging (ASOADeK) method is a geostatistical model that has been applied for long-term monthly precipitation characteristics in northern New Mexico (Guan et al., 2005). The ASOADeK method incorporates independent variables that represent both physiographic and atmospheric factors, including elevation, the relationship between moisture flux direction and aspect, and the spatial gradient of moisture distribution.

The gridded model output statistics (GMOS) method is used to develop gridded weather forecasts for NWS forecasters (Glahn et al., 2009; Gilbert et al., 2009). Regression relationships are developed at point locations using weather variables (e.g., temperature, precipitation) as the dependent variables and physiographic variables (e.g., elevation, latitude) and atmospheric variables from the Global Forecast system (GFS) NWP model (e.g., wind components, divergence of the moisture flux) as the independent variables. The point forecasts are interpolated to produce gridded forecast products.

Personnel at ESRL are working to improve upon the PRISM system by incorporating atmospheric variables into the regression models (Hsu et al., 2011). This work includes developing a nonlinear, hierarchical model on a daily time step to identify additional controlling variables beyond those incorporated into the PRISM system. For example, atmospheric and oceanic variables such as vertical motion, sea surface temperature, Froude number, and water vapor are being evaluated to determine if they improve upon the general PRISM formulation.

Clark and Slater (2006) present a locally-weighted regression approach to generate daily precipitation grids from daily station values in mountainous terrain. Regression models are used at each time step to estimate the conditional distribution of precipitation. Ensembles of precipitation grids are generated using stochastically-generated, correlated random fields. Clark and Slater suggest that the technique could be extended to merge the ensemble grids with precipitation estimates from radar and satellites.

6.3.3 Multi-sensor Approaches

Background fields to guide the gridding of station data, similar to the CAI method, can be generated from other gridded data sets, such as radar products, satellite products, or high-

resolution NWP model outputs. This type of multi-sensor approach could be achieved through direct application of point observations in the data fusion process.

6.4 Recommendations for the AOR

A wide range of techniques are available to spatially interpolate point observations. If gridding is employed as part of the AOR development, at a minimum the gridding method should utilize CAI, with an interpolation method based on the SYMAP algorithm or optimal interpolation. Such an approach would account for typical physiographic controls on the distribution of precipitation and temperature.

The SYMAP algorithm based on Shepard (1964) requires no characterization of correlation characteristics, yet has yielded comparable results to more complex approaches.

Optimal interpolation allows differences in correlation characteristics between stations to be represented. Since correlation characteristics can be defined independently for each station, optimal interpolation provides greater flexibility across the domain compared to kriging. In addition, optimal interpolation would yield error statistics related to the interpolation of the anomalies that may be useful for developing error statistics for the final interpolated grid.

The hierarchical regression-based approach of Hsu et al. (2011) also merits further investigation for the AOR due to its ability to incorporate non-linear atmospheric effects into the interpolation process.

7 Spatial Downscaling

Spatial downscaling is herein defined as the process of creating a higher resolution grid from a lower resolution grid. Because most of the gridded precipitation and temperature data sources are available at a coarser resolution than the grid mesh desired for the AOR (***Table 10***), it is likely that spatial downscaling techniques will need to be applied during AOR development.

Table 10. Grid Meshes associated with Historical Data Archives

Data Type	**Nominal Grid Mesh [km]**
Radar (Level III Digital Precipitation Array)	4
Radar (NMQ)	0.5-1
Satellite	4-28
Reanalysis	10-124

Several comprehensive reviews of downscaling techniques have been published in the literature (e.g., Wilby et al., 2004; Fowler et al., 2007). These reviews summarize the downscaling methods that are available, along with their associated benefits and limitations. Many of these articles have been written with a focus on climate change and the need to downscale outputs from general circulation models and global climate models. In addition to these published references, personnel at the NOHRSC and the ESRL have performed surveys of downscaling techniques and have documented their results (Johnson, 2008; Johnson and Schneider, 2009; NOHRSC, personal communication, September 8, 2011).

In this section, Riverside does not attempt to repeat the extensive work that has been done to describe and compare spatial downscaling techniques. Rather, this section is intended to provide a brief summary of techniques that are likely to be useful for the AOR development, either because they are relatively straightforward to implement, because they are in use operationally, or because they have the highest potential to capture weather variability at the grid mesh intended for the AOR.

Downscaling methods and categories are not strictly distinct. Many applications apply multiple methods or combinations of methods (Wood et al., 2004; Johnson, 2008). Although the published literature varies in how the downscaling techniques are categorized, almost all of the references differentiate between statistical and dynamical downscaling techniques. In addition, some sources separate interpolation techniques into a third category; Riverside adapts this convention rather than including interpolation techniques within the statistical downscaling category.

7.1 Interpolation Methods

Basic interpolation methods, such as bilinear interpolation and majority assignment, add the least amount of value to downscaled grids because they do not account for topographic and atmospheric factors that influence climate (Wood et al., 2004). The value of interpolation methods for the AOR depends on the starting resolution of the dataset; the finer the starting resolution, the more accurate the downscaled results could be using interpolation methods. Likewise, interpolation methods are more suitable when weather gradients are small, as variables with less spatial variability can be more accurately downscaled using basic interpolation methods.

Many tools have been developed that include basic interpolation functions for grids:

- ***ArcGIS:*** Interpolation options include nearest neighbor assignment, majority assignment, bilinear interpolation, and cubic interpolation.
- ***NCEP CopyGB utility:*** Interpolation options include bilinear interpolation, bicubic interpolation, nearest neighbor assignment, nearest neighbor averaging, and spectral interpolation (Iredell, 2000).
- ***Grid Analysis and Display System (GrADS):*** GrADS includes a function to perform bilinear interpolation.
- ***NCAR Command Language (NCL):*** NCL includes multiple functions to perform bilinear interpolation.
- ***Geographic Resources Analysis Support System (GRASS):*** Interpolation options include nearest neighbor assignment, bilinear interpolation, and bicubic interpolation.
- ***R:*** Interpolation options include nearest neighbor assignment, linear interpolation from nearest neighbors, piece-wise cubic hermite interpolating polynomial, cubic interpolation, and cubic spline interpolation.

Operational systems that employ interpolation methods typically apply a vertical adjustment to account for elevation differences between the coarse resolution grid being downscaled and the high-resolution downscaled grid. The vertical adjustments help to introduce

variability into the downscaled grid, since basic interpolation techniques tend to produce downscaled grids that are relatively smooth.

There are multiple operational applications within NOAA that apply downscaling methods based on interpolation. For example, the NCEP Smart Initializations (SmartInits) system downscales NWP model outputs to a nominal mesh of 2.5-5.0 km for the National Digital Forecast Database (NDFD). Bilinear interpolation is applied to downscale all weather variables from the coarser resolutions used by the NWP models to the higher resolution used by the NDFD. In addition, the temperature fields are adjusted using lapse rates and a high-resolution DEM to account for elevation differences between the coarse NWP grids and the high-resolution NDFD grids (Johnson, 2008).

Precipitation fields from the RUC NWP model are downscaled from 13 km to 1 km using bilinear interpolation for input to the NOHRSC Snow Data Assimilation System (SNODAS; NOHRSC, personal communication, July 20, 2011). The temperature inputs to SNODAS also originate from the RUC model. The temperature fields are downscaled from 13 km to 1 km using two-dimensional bilinear interpolation, and are then adjusted for terrain differences between the RUC and SNODAS grids. Lapse rates are computed in real-time based on level 1 and level 4 temperature fields from the RUC model. The lapse rates are limited to the adiabatic lapse rate. Lapse rates are not applied if temperature inversions are occurring. Additionally, the downscaled temperature grids are adjusted using station observations prior to input to the SNODAS. Anomalies are computed between station observations and the downscaled, terrain-adjusted temperature field. The anomalies are interpolated using inverse distance weighting, although prior station weights are incorporated in areas with complex terrain. The interpolated anomalies are applied to adjust the downscaled temperature field. The assimilation of observations helps to capture temperature inversions in the final temperature field that is input to the SNODAS.

7.2 Statistical Downscaling Methods

Statistical downscaling methods rely on developing statistical relationships between climate variables on a fine scale and climate, physiographic, and atmospheric variables on a coarse scale. These methods add value to the downscaled product by incorporating spatial variability from higher resolution fields that capture physiographic and/or atmospheric influences. Statistical downscaling approaches are particularly useful in areas where climatic variables exhibit strong relationships with synoptic-scale factors (Wilby et al., 2004).

One of the major limitations of statistical downscaling methods is that they tend to underestimate variance and to poorly represent extreme weather events. Therefore, techniques are sometimes employed to add variability to the downscaled weather variables (Fowler et al., 2007).

7.2.1 Spatial Disaggregation

Spatial disaggregation techniques are also referred to as anomaly methods, perturbation methods, and delta-change approaches. They are analogous to the CAI techniques described for gridding station observations in ***Section*** 6. Spatial disaggregation methods require two datasets: a low-resolution gridded dataset that will be downscaled and a high-resolution gridded dataset with more realistic spatial variability. This type of method maintains the long-term characteristics of

the coarse grid being downscaled, while introducing spatial variability using a higher resolution dataset. For example, spatially disaggregating satellite precipitation grids (i.e., 4-28 km) using PRISM climatology grids (i.e., 400 m) will retain the long-term climatology of the original satellite precipitation dataset, while introducing spatial variability into the downscaled product that reflects the topographic and coastal effects accounted for in developing the PRISM dataset. A general description of the spatial disaggregation methodology follows (Wood et al., 2004):

- Use the high-resolution dataset to develop an anomaly field.
    - Compute gridded climatological averages from the high-resolution dataset. The climatology can be developed on a monthly, seasonal, or annual basis. This step is not necessary if the starting high-resolution dataset represents long-term climatological averages (e.g., PRISM).
    - Aggregate the climatological grids to the coarse resolution of the dataset being downscaled, and then interpolate back to the original high resolution.
    - Compute an anomaly field between the original high-resolution climatology and the interpolated high-resolution climatology. Typically, the anomalies are expressed as ratios for precipitation and as differences for temperature.
- Downscale the low-resolution dataset using the anomaly field.
    - Interpolate the low-resolution dataset to the high resolution using the same interpolation procedure that was applied to the climatology dataset.
    - Apply the anomaly fields to the interpolated low-resolution dataset using multiplication for precipitation, or addition for temperature, as appropriate.

7.2.2 Bias Correction and Spatial Disaggregation (BCSD)

The bias correction and spatial disaggregation (BCSD) technique improves upon the spatial disaggregation technique by incorporating a bias adjustment to the long-term characteristics of the coarse resolution dataset. Without the adjustment, biases in the original dataset being downscaled will be retained in the downscaled dataset. Though the downscaled data may have realistic spatial variability, they can be unsuitable for hydrologic simulations, depending on the characteristics of the bias (Wood et al., 2004).

A technique for correction of real-time gage-radar hourly analyses using monthly total gage-only analyses was demonstrated by Zhang et al. (2011). The approach succeeded in correcting long-term, time-dependent biases due to radar underestimation over the eastern U.S.

Bias correction techniques such as mean field bias adjustments and local bias adjustments have been discussed in ***Section 3.2***. Additional bias correction techniques exist, such as quantile mapping, that are intended to better capture the probability distribution function of the weather variable so that extreme events, which have been difficult to capture using statistical downscaling, are better represented in the downscaled dataset (Wood et al., 2004; Wood et al., 2006; Hashino et al., 2006; Gutmann et al., 2011). As with the spatial disaggregation technique, the bias correction requires two datasets: a low-resolution gridded dataset that will be downscaled, and a high-resolution gridded dataset that will be used to compute the bias corrections. A general description of a bias correction technique based on quantile mapping is as follows (Wood et al., 2004; Hashino et al., 2006):

- Develop the cumulative distribution functions for the coarse-resolution dataset being downscaled and the high-resolution dataset using gridded climatological averages from both datasets.
- Perform quantile mapping at the coarse resolution, so that the adjustments vary spatially by month for each grid cell at the coarse resolution. The quantile mapping approach replaces the climatological average for a grid cell in the coarse dataset with the value from the high-resolution dataset that has the same non-exceedance probability. Mathematical formulations can be used to fit and relate the cumulative distribution functions, or a one-to-one mapping between the cumulative distribution functions can be applied (Hashino et al., 2006).
- Apply the bias adjustments to the coarse-resolution dataset being downscaled, using multiplication for precipitation, or addition for temperature, as appropriate. ***Figure 6*** presents an example of applying a one-to-one mapping to the cumulative distribution functions, and the effect of the bias adjustments on the time series of gridded values.
- Once the bias correction has been applied, perform spatial disaggregation as described above to introduce spatial variability from the high-resolution dataset.

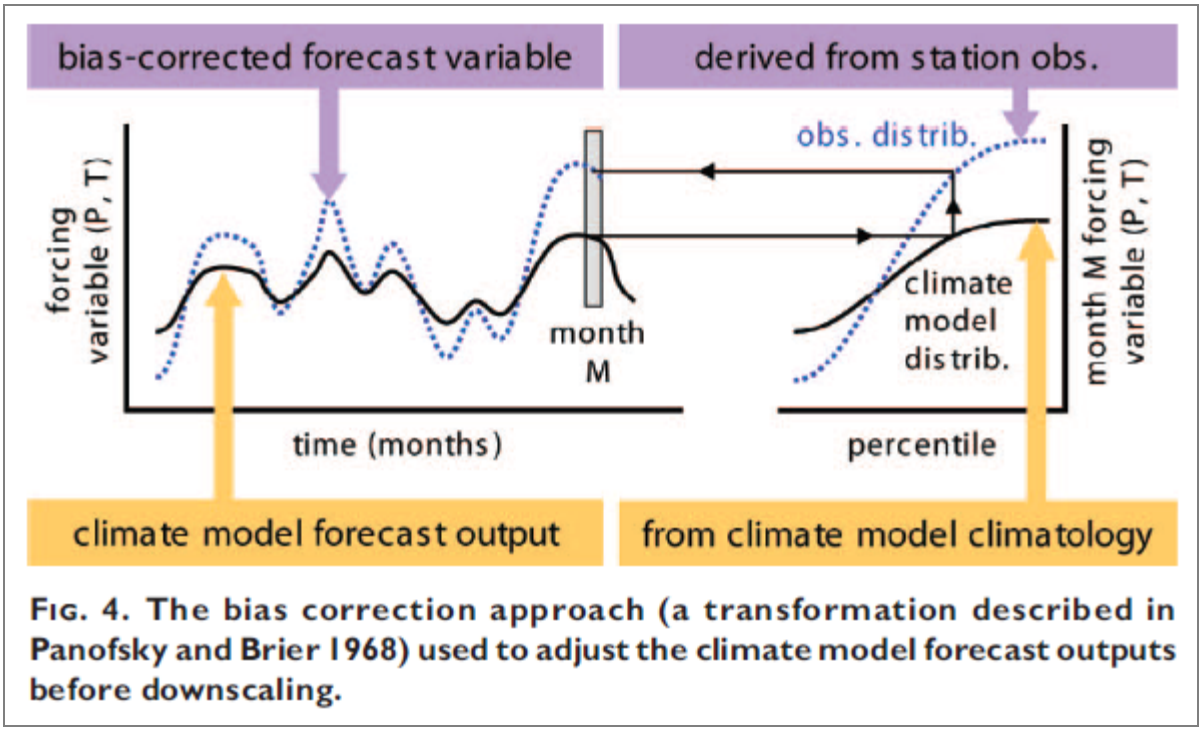

Figure 6. Example of Quantile Mapping for Bias Correction (reproduced from Wood et al., 2006)

7.2.3 Regression Methods

Statistical regression models (also known as transfer functions) can be used to downscale temperature and precipitation variables as a function of location, topography, climatological averages, and atmospheric variables. The challenge in developing regression models is

accurately identifying the variables that control the weather patterns for the area being downscaled. The benefit in developing regression models is that additional variables can be introduced into the downscaling process that may not be captured in existing datasets.

A wide variety of statistical techniques can be employed, including multiple linear regression, principal components analysis, canonical correlation analysis, and artificial neural networks, among others. Common predictor variables include elevation, aspect, latitude, longitude, air temperature, humidity, sea level pressure, wind speed, wind direction, dew point, dew point temperature depression, geopotential height, moisture flux direction, and the spatial gradient of moisture distribution. As the regression models become more complex, incorporating more variables as surrogates for physical processes and allowing interactions between variables to capture nonlinear processes, the results from the statistical downscaling can be very similar to those achieved using dynamical downscaling approaches (Fowler et al., 2007).

When spatial downscaling is applied to grids, the regression equations are developed for grid cells, as opposed to the gridding techniques described in ***Section 6***, which develop regression equations for point locations. The regression models can be developed globally, regionally, or for individual grid cells. Typically, the statistical power of the relationship improves as the development area becomes smaller, but the potential for discontinuities between regions increases (Glahn, 2006; Charba and Samplatsky, 2011a). Regional relationships are superior to global relationships when the predictors exhibit regional variations in controlling weather patterns, or the datasets have regionally-varying biases (Charba and Samplatsky, 2011a).

Multiple linear regression models are being developed using a research version of the GMOS methodology to directly produce gridded forecast products for the National Digital Guidance Database (NDGD), including quantitative precipitation forecasts and the probability of precipitation (Charba and Samplatsky, 2011a; Charba and Samplatsky, 2011b). The regression models incorporate static predictors such as location, elevation, and climatological averages, in addition to dynamic predictors based on GFS NWP model outputs. Terms representing the multiplicative product between selected static and dynamic predictors are included in the regression models to capture nonlinear interactions.

The regression models are being developed for the geographic regions shown in ***Figure 7***, which were determined based on geographical variations in precipitation forcing mechanisms, regional biases in the GFS fields, assessments of the model performance, and the costs required to develop the models. To address discontinuities between the regions, Charba and Samplatsky developed a technique to weight the predicted values in overlapping areas, wherein the weights are a function of the distance to the regional boundaries (Charba and Samplatsky, 2011a).

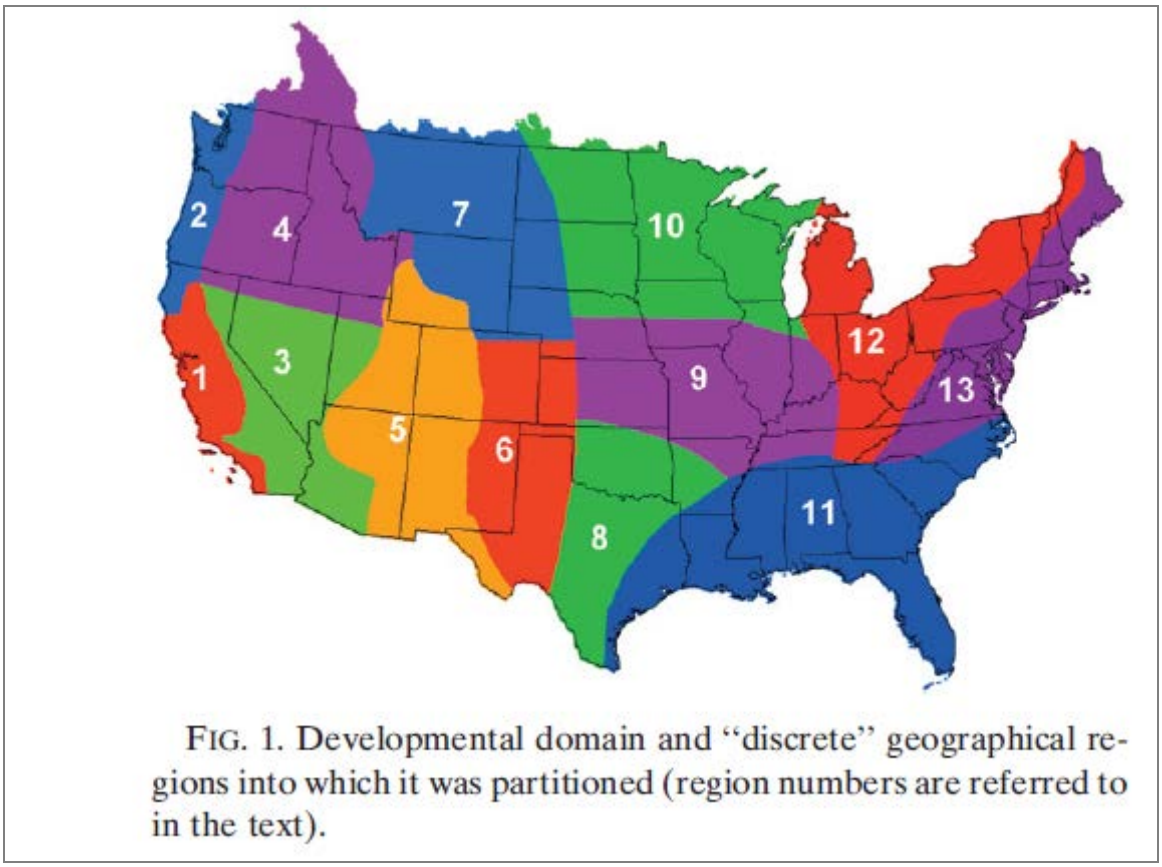

Figure 7. Geographic Regions used by GMOS for NDGD Forecast Products (reproduced from Charba and Samplatsky, 2011a)

7.3 Dynamical Downscaling Methods

Dynamical downscaling methods rely on high-resolution weather models that can simulate the physical processes that control weather. The weather models can account for factors such as local frontal systems, topographic channeling of flow, and the interplay of atmospheric dynamics with hydrometeor microphysics that are difficult to incorporate into statistical downscaling methods (Gutmann et al., 2011). The weather models are run using initial and boundary conditions from coarse-resolution NWP model outputs. The outputs from the weather model are then incorporated into a statistical downscaling framework.

Dynamical downscaling methods that employ high-resolution weather model outputs are typically able to represent regional features, such as orographic precipitation, extreme climate events, regional-scale climate anomalies, and non-linear effects (e.g., the El Nino Southern Oscillation), better than statistical downscaling techniques (Fowler et al., 2007). The relative advantages and disadvantages of statistical and dynamical downscaling methods are summarized in ***Figure 8***.

Table I. Comparative summary of the relative merits of statistical and dynamical downscaling techniques (adapted from Wilby and Wigley, 1997).

	Statistical downscaling	Dynamical downscaling
Advantages	• Comparatively cheap and computationally efficient • Can provide point-scale climatic variables from GCM-scale output • Can be used to derive variables not available from RCMs • Easily transferable to other regions • Based on standard and accepted statistical procedures • Able to directly incorporate observations into method	• Produces responses based on physically consistent processes • Produces finer resolution information from GCM-scale output that can resolve atmospheric processes on a smaller scale
Disadvantages	• Require long and reliable observed historical data series for calibration • Dependent upon choice of predictors • Non-stationarity in the predictor-predictand relationship • Climate system feedbacks not included • Dependent on GCM boundary forcing; affected by biases in underlying GCM • Domain size, climatic region and season affects downscaling skill	• Computationally intensive • Limited number of scenario ensembles available • Strongly dependent on GCM boundary forcing

Figure 8. Advantages and Disadvantages of Statistical and Dynamical Downscaling Techniques (reproduced from Fowler et al., 2007)

The potential advantages of dynamical downscaling techniques for the AOR development are directly proportional to the extent that weather models can help to identify important physical processes that control weather variability. In some cases, variables representing these controlling processes can be incorporated to improve long-term climatological datasets such as those developed using the PRISM system. In other cases, the variability for an individual weather event cannot be represented well using a long-term climatological dataset.

Several groups have used high-resolution regional weather models to simulate domains with complex terrain in the intermountain west (Hughes et al., 2009; Ikeda et al., 2010; Gutmann et al., 2011; Rasmussen et al., 2011; Silverman and Maneta, 2011). These studies have identified dynamic atmospheric conditions that significantly affect weather patterns in the west.

Researchers at NCAR applied the WRF model over Western Colorado for the period 2001-2008 (Ikeda et al., 2010; Gutmann et al., 2011; Rasmussen et al., 2011). The WRF model was run using grid meshes ranging from 2 to 36 km using initial and boundary conditions from the NARR reanalysis dataset. The outputs from WRF were used to develop climatological averages, which were then compared to the 800-m climatological averages developed using the PRISM system for the period 1971-2000. (Although the climatological averages were developed for different periods, the researchers felt a comparison was valid because the period 2001-2008 included wet, dry, and average years.) The WRF and PRISM climatologies represented similar patterns of precipitation over Western Colorado, although localized differences existed (Gutmann et al., 2011). For example, the WRF and PRISM climatologies had significant differences in the maximum precipitation estimated in the San Juan Mountains; recent observations seem to support the WRF estimate. Additionally, the researchers found that synoptic features such as prevailing wind directions and “up-slope” storms significantly impact precipitation patterns (Gutmann et al., 2011). These factors may not be fully represented by the PRISM system, although predominant wind directions that vary by season should be captured in the monthly climatological grids. In theory, a high-resolution regional weather model should

provide better estimates in areas that lack observations than the PRISM system (Gutmann et al., 2011).

Researchers at UCLA employed the NCAR Mesoscale Model (MM5) over the intermountain west for the period 1995-2006, using a nested domain with grid meshes that ranged from 6 km over California (the location of interest) to 54 km (Hughes et al., 2009). The MM5 model was run using initial and boundary conditions from a NCEP 40-km Eta model reanalysis dataset. The major finding of this work is that orographic blocking substantially affects the relationship between precipitation and elevation. Orographic blocking occurs when the kinetic energy of airflow approaching a mountain is less than the energy needed to pass over the mountain, causing the airflow to decelerate and produce increased precipitation upstream of the mountain. The researchers found that the Froude number (i.e., the ratio of the kinetic energy of the ambient airflow to the energy required to pass over the mountain) can be used to classify the strength of orographic blocking. For example, the results in ***Figure 9*** demonstrate differing relationships between precipitation and elevation for two ranges of Froude numbers. When blocking is weak (i.e., high Froude number), precipitation increases significantly with elevation. When blocking occurs (i.e., low Froude number), precipitation shows little increase with elevation. This work is one example of how weather models can be used to identify dynamical processes that need to be accounted for when downscaling is applied. Variables such as the Froude number that represent the atmospheric conditions could be incorporated into statistical downscaling methods that employ regression methods to improve the downscaled outputs. Otherwise, regional weather models are needed to capture the physical processes.

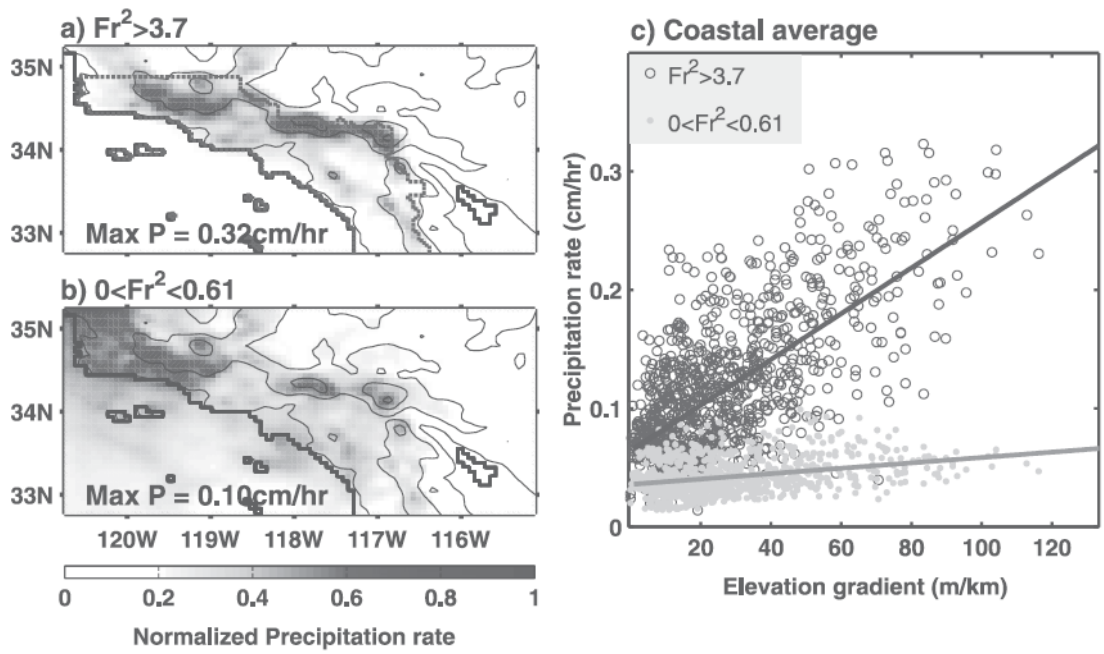

FIG. 4. Average precipitation rates, normalized by the domain maximum precipitation rate (Max P), for hours with (a) $Fr^2 > 3.7$ and (b) $0 < Fr^2 < 0.61$. (c) Average precipitation rates for coastal land points vs. elevation gradient for hours with $0 < Fr^2 < 0.61$ (light gray dots) and $Fr^2 > 3.7$ (darker gray circles). Gray dotted line in (a) shows grid points used in (c). Terrain is shown as in Fig. 1. Light and dark gray lines show corresponding regression lines.

Figure 9. Effect of orographic blocking on precipitation-elevation relationships (reproduced from Hughes et al., 2009)

7.4 Implications for the AOR

A wide range of spatial downscaling techniques is available. Some of the downscaling techniques are simple and easily implemented (i.e., interpolation methods), while others are complex and require further investigation and development prior to applying to the AOR development (i.e., dynamical methods).

The interpolation downscaling methods are readily available and easily implemented, with several operational applications already taking advantage of these techniques. For the AOR, it may be an acceptable minimum to utilize interpolation with terrain adjustments to downscale temperature fields, but interpolation methods are not recommended for downscaling precipitation fields.

The spatial disaggregation methods are also reasonably developed and straightforward to implement. For the purposes of the AOR, it may be an acceptable minimum to downscale precipitation fields using spatial disaggregation with bias corrections.

Spatial downscaling techniques that employ regression methods are desirable for incorporating dynamic atmospheric processes that control weather patterns. However, additional development is required to identify controlling variables and to develop suitable relationships for different geographic regions. If time and resources permit further investigation for the AOR, the GMOS method using reanalysis products and other inputs would be a reasonable starting point to develop these relationships. The predictor variables can be used both for gridding point observations and downscaling gridded datasets.

Similarly, dynamical downscaling methods that employ outputs from high-resolution regional weather models are desirable for incorporating atmospheric influences. If a high-resolution weather model can be run to generate a reanalysis dataset for the AOR, it would be extremely beneficial in providing a consistent, high-resolution dataset over the full historical period and spatial extent. Additional investigation is required to determine if it is feasible to complete a WRF reanalysis for the AOR, and to weigh the relative effort and potential benefits.

8 Temporal Disaggregation

Temporal disaggregation methods are needed for the AOR development to downscale daily values to an hourly time step. Because weather variables exhibit higher spatial correlations and lower serial correlations on a daily time step, and because more historical data are available on a daily time step, Riverside recommends developing daily fields before temporally disaggregating to an hourly time step.

Hourly weather characteristics can be difficult to infer depending upon the location and period of time. Hourly gages are available over the full period intended for the AOR, although the density of gages varies in time and space. Radar data are available in 5-min to 1-hr increments from approximately 1996 to present, but the spatial coverage is incomplete and the data quality is limited for frozen precipitation. Likewise, satellite estimates are available in 15-min to 3-hr increments for the recent period, but the uncertainties are relatively high, particularly for precipitation estimation over areas covered by snow and ice. The reanalysis products are available for the full period intended for the AOR with time steps of 1-6 hr. However, the grid meshes of the reanalysis products are relatively coarse, and the uncertainties associated with the reanalysis data are also high.

8.1 Grid-based Disaggregation Approaches

The radar, satellite, and reanalysis data sources are gridded products that can be used for temporal disaggregation. The primary issue in using these data sources for temporal disaggregation is that they require spatial downscaling to the target grid mesh (***Section 7***). The downscaled radar, satellite, or reanalysis datasets could be used to temporally disaggregate daily grids using the hourly distributions defined by the hourly gridded datasets. For example, Wüest et al. (2009) first developed a daily precipitation field from station observations, and then disaggregated the daily product using radar data. This approach is limited to the period with radar (or satellite or reanalysis) estimates.

If fields such as those available from the NARR, which are available on a 3-hr time step, are used for temporal disaggregation of daily grids, an additional step is required to disaggregate from a 3-hr to a 1-hr time step. A simple alternative is to assume a uniform distribution of values for every 3-hr interval, although more sophisticated techniques could be developed. Also,1-hr radar or satellite precipitation could be used, as is done for the NLDAS dataset (Cosgrove et al., 2003).

8.2 Station-based Disaggregation Approaches

Temporal disaggregation based on station data could be completed using a variety of approaches. Hourly gage observations could be directly interpolated using one of the approaches described in ***Section 6*** (e.g., Luo et al., 2011). The temporal distribution from the hourly grids could be used to disaggregate daily grids to a 1-hr time step.

If the hourly station network is relatively dense, station-based disaggregation approaches can successfully represent the spatial distribution of events, including the movement of storms across a basin. However, spatial distributions will not be well-represented in areas with sparse hourly station networks. In these areas, the spatial correlations would be overestimated, resulting in overly smoothed datasets (Debele et al., 2007). An alternative approach that better accounts for spatial and temporal correlations in the dataset would be preferable.

8.3 Multi-sensor Approaches

Data fusion could be used to integrate information from all available sources into the temporal disaggregation process. This could involve developing hourly datasets through the data fusion process to incorporate information from all sources of hourly information (i.e., gages, radars, satellites, and NWP model outputs), and then scaling the hourly grids so that they reproduce the daily values developed in a previous data fusion step.

Similarly, a multi-sensor approach could be used to characterize the diurnal distribution of temperatures. Jin and Dickinson (1999) developed a multi-sensor approach that combines satellite and surface observations with NWP model outputs to characterize the diurnal cycle of land surface skin temperatures (***Figure 10***). The authors formulated the diurnal cycle for a particular day as a combination of a daily average, a daily periodic component, and a random component. The daily average component was characterized using satellite and gage observations (depending on cloud cover conditions). The periodic component was characterized

using simulated climatologies from the CCM3/BATS model, where the model simulations were used to determine typical diurnal patterns for different latitudes, seasons, and vegetation types.

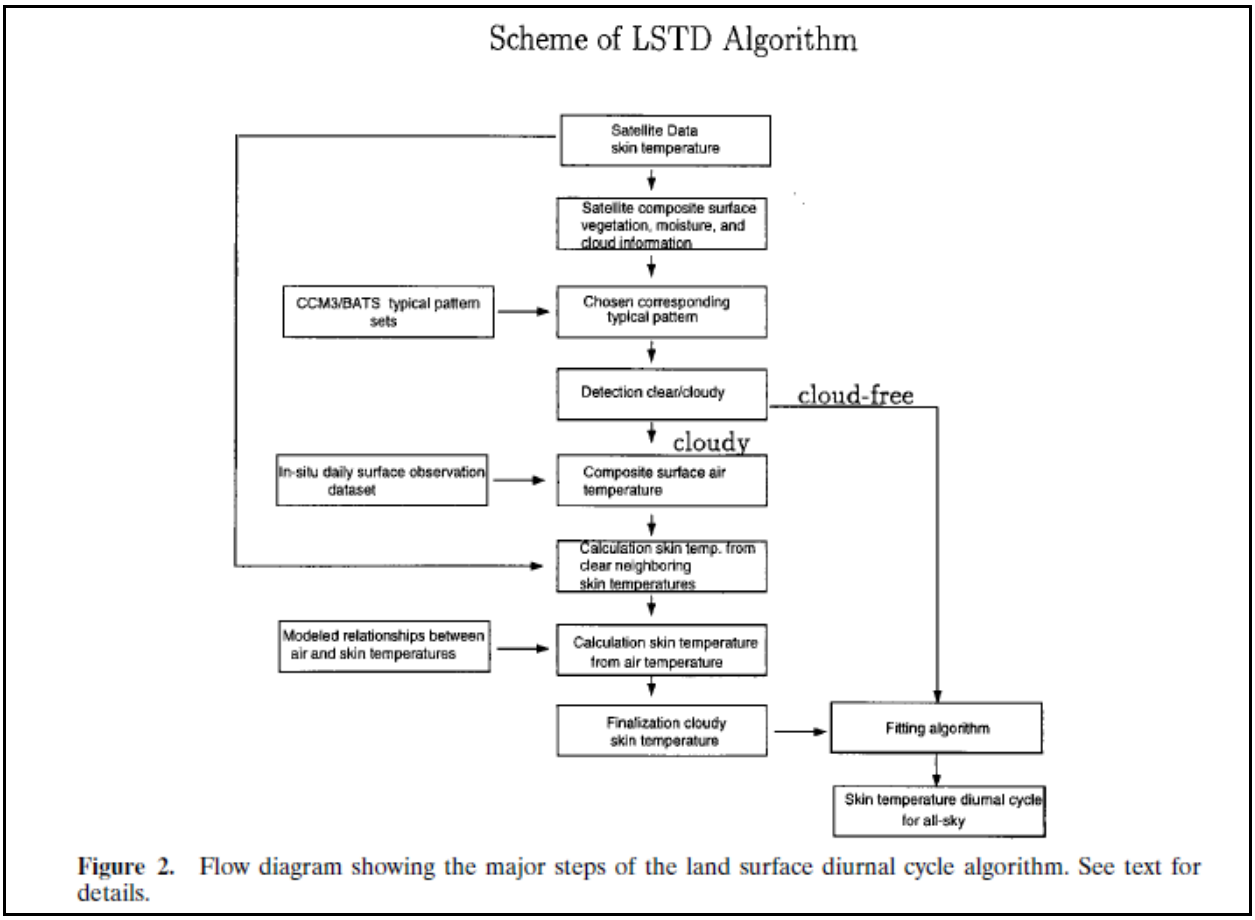

Figure 10. Determination of the diurnal temperature cycle using NWP model outputs and satellite and surface observations (reproduced from Jin and Dickinson, 1999)

8.4 Stochastic Approaches

Stochastic approaches could be employed to temporally disaggregate daily grids. Stochastic approaches would not reflect hourly observations for a particular day, but instead would attempt to maintain the statistical characteristics of the daily fields.

Many studies have investigated disaggregation techniques for individual stations, or the extension of hourly stations, to maintain statistical characteristics at the station (e.g., Debele et al., 2007; Wheater et al., 2005). For a relatively small basin, Wheater et al. (2005) applied the disaggregation approach of Koutsoyiannis and Onof (2001) to temporally distribute daily station observations to an hourly time step. Subsequently, the researchers used this distribution to disaggregate stochastically-generated daily data across the basin. This disaggregation approach, particularly on larger scales and in regions with complex terrain, over-estimates the cross correlation between locations (Mezghani and Hingray, 2009). Alternatively, individual stations

could be stochastically disaggregated using these approaches prior to spatial interpolation; however, this would ignore cross correlation of events between stations.

Weather generation models have been applied to maintain spatial and temporal correlations for multiple-site disaggregation and two-dimensional storm generation. Wheater et al. (2005) present a generalized linear regression model that represents the development and movement of storms and storm cells based on stochastic characteristics of precipitation fields. The parameters defining the precipitation field characteristics are determined based on analysis of radar precipitation estimates. Although such an approach appears promising, there are major limitations. The model is developed on a daily basis because of the added complexities and differences in spatial and temporal correlations at finer time intervals. Non-homogeneous precipitation characteristics such as orographic effects are not represented in the model. The weather generation models have been developed over limited domains. Finally, no mechanism is available to adjust the models to observations.

K-nearest neighbor (K-nn) resampling approaches have been used in multiple applications to preserve the statistical characteristics of time series and gridded datasets. With K-nn approaches, a database of historical events is developed and similarity criteria are defined. Data from the current time step are compared with the historical database to identify appropriate historical analogs. An analog is randomly selected from the K number of nearest analog events that are weighted based on their similarity to the conditions at the current time step (Buishand and Brandsma, 2001). A wide variety of information (e.g., station observations, atmospheric states from NWP models) can be combined to define the similarity criteria, as long as the information is available for the historical record as well as for the current time step. Brandsma and Können (2006) applied K-nn resampling to disaggregate hourly temperature data from limited observations collected over a day. Gangopadhyay et al. (2005) used K-nn resampling to spatially and temporally downscale precipitation and temperature from the R1 reanalysis. In a two-step process, Mezghani and Hingray (2009) statistically downscaled daily precipitation and temperature grids from the R1 reanalysis, and then applied K-nn resampling to temporally disaggregate the daily grids to an hourly time step.

Mezghani and Hingray (2009) and Gangopadhyay et al. (2005) summarize the strengths and limitations of K-nn resampling. The K-nn approach does not make assumptions regarding the underlying probability distributions or joint distributions of the functions. However, the cross-correlations of data across space can be represented in the method. The approach may also be utilized to select an ensemble of possibilities and to define confidence information based on these ensembles (Gangopadhyay et al., 2005). K-nn resampling techniques rely on a representative database of historical events that reflect the range of possibilities, and as such events cannot be generated that exceed the limits of the historical record. Temporal correlations of weather variables may be underestimated with K-nn approaches, although this can be overcome in part by including lagged variables in the predictor similarity functions. A K-nn resampling approach would ignore hourly observations from gages or other sources. However, these data might be used to identify analog events. Although the K-nn approach can produce hourly data that are statistically consistent, the hourly distribution on any given day could have large errors.

8.5 Implications for the AOR

For the AOR development, temporal disaggregation methods based on observations or stochastic methods can be employed to develop datasets on a 1-hr time step. The final approach will ultimately depend upon available resources, but the following options could be explored:

- Employing data fusion to develop hourly fields using all available hourly data (i.e., gages, radars, satellites, and NWP model outputs), and then adjusting the hourly fields to be consistent with daily fields developed in previous analysis steps. Observation-based distributions are likely to be better in the recent period and in locations with hourly gage and radar observations.
- Developing a K-nn resampling approach using the recent period to populate the historical database and to characterize the statistical properties. This method is likely to be most applicable prior to 1996.
- Developing an approach that combines observation-based distributions and stochastic distributions based on location, period, data availability, and confidence in the available data.

9 AOR Evaluation

An important aspect of the AOR development process is evaluating the quality of the precipitation and temperature datasets. An evaluation process is necessary during development to compare and adjust potential methodologies. Although inter-comparison studies have been done to evaluate specific techniques that may be applied to the AOR, the results of these studies are highly dependent upon the input datasets and the climatological characteristics of the area being studied. Once the AOR is complete, an evaluation process is necessary to understand the quality and uncertainties associated with the datasets.

This section focuses on evaluation methods that can be used during the AOR development process, although the same techniques can be applied to the final datasets. A single evaluation approach will not provide definitive conclusions regarding estimation accuracy across the entire country for the entire dataset. Therefore, Riverside recommends using multiple evaluation techniques to assess the benefits and limitations of different approaches. The recommended evaluation methods have been categorized as follows:

- Expert input (i.e., peer reviews)
- Cross validation against point observations
- Comparisons with previously developed datasets
- Comparisons with high-quality datasets from heavily instrumented basins
- Assessments of simulation quality using hydrologic models

9.1 Expert Input

The AOR development process touches on many fields, including precipitation and temperature estimation from gage, radar, satellite, and NWP model sources; automated data

quality control; gridding; downscaling; temporal disaggregation; data assimilation; data fusion; and uncertainty estimation. The AOR development process will benefit from review and feedback by experts in each of these fields. Although expert input can lead to growing or conflicting requirements and potential changes in methodology, it can provide invaluable guidance to avoid pursuing approaches with limited value (Daly, 2006). In soliciting expert input, it will be helpful to manage expectations by conveying an accurate understanding of the needs, limitations, and timeline for the AOR development.

Riverside recommends that expert input be solicited at various points in the development process. For this report, Riverside conducted expert interviews in different areas to better understand research that has been completed and to solicit recommendations regarding feasible approaches. As OHD refines the development plan, Riverside recommends soliciting another round of expert input, particularly to help assess the benefits and limitations of each alternative method. Once the AOR has been developed, expert review would also be beneficial in assessing whether the final dataset reflects reasonable weather patterns.

9.2 Cross Validation against Point Observations

One commonly applied evaluation method is cross validation of the dataset using station observations. In this approach, individual stations are withheld from the development process, the estimation is completed without the excluded station, and the estimated value at the grid cell containing the excluded station is compared to the observed values. This process is repeated for all stations. Summary statistics are computed to assess the quality of the applied methodology. Cross validation can also be completed by removing groups of stations (e.g., high elevation SNOTEL stations) to assess predictive skill under various conditions.

Cross validation is appealing because it provides a statistical assessment that can be used to compare different techniques. Willmott and Matsuura (2006) emphasize the use of the mean absolute error and mean bias error in comparing spatial interpolation techniques. Li and Heap (2011) provide a summary of the most commonly used statistics to assess the performance of spatial interpolation methods. Hofstra et al. (2008) utilize various statistics that emphasize differing aspects of the estimation quality, and present the results of the analysis in different formats to visualize and interpret the variation in estimation quality across a large area.

Although cross validation can provide insights into the strength of a chosen methodology, Daly (2006) summarized various limitations of cross validation analysis. Station observations are not necessarily equal to the areal average associated with a grid cell due to local site characteristics, surrounding topographic features, and other factors. Accepting that a station observation may be the best estimate for the corresponding grid cell, cross validation only provides an assessment of estimation quality where station observations are available. In many cases, stations are located at low elevations, so the results of the cross validation analysis may not represent the ability of a technique to reproduce values at higher elevations. Importantly, a cross validation analysis can only compare different methodologies if equivalent input data sources are included in each method. If cross validation statistics are computed with differing inputs, the results could be misleading. For example, the addition of stations in topographically complex areas will generally improve the precipitation or temperature estimation in these regions. However, a cross validation analysis including these stations would tend to yield poorer statistical results than an analysis that included only lower elevation stations, since it is more difficult to estimate values at high elevations. Thus, it would be inappropriate to use cross

validation to assess the value of including or excluding specific station networks. Finally, cross validation statistics tend to favor techniques that yield smoother interpolated surfaces that may not represent the true variability of the precipitation or temperature field.

9.3 Comparison with Datasets from Highly Instrumented Basins

There are several watersheds around the nation where dense measurements of precipitation and temperature have been collected over relatively limited periods of time for a variety of purposes. In such cases, the spatial and temporal distributions of precipitation and temperature are understood with greater confidence than is possible under typical monitoring conditions. The high-resolution precipitation and temperature fields from these locations provide a unique benchmarking opportunity for evaluation of the AOR methodologies. Various precipitation or temperature inputs could be withheld to represent typical data inputs that would be available, similar to the cross validation analysis. Precipitation and temperature fields could be generated based on the limited data inputs and compared against the high-resolution precipitation or temperature fields derived using all available data inputs.

The advantage of such an analysis is that the baseline dataset is known to a greater degree than in the case of a cross validation analysis based on point observations. The inclusion or exclusion of different inputs or the use of different methodologies could be assessed more completely. However, there are limitations with this approach as well. The highly instrumented basins typically have limited records, so the range of precipitation and temperature conditions, storm types, extreme events, and other conditions for the watershed may not be represented in the available historical record. Furthermore, although the evaluation may indicate that a particular methodology yields superior results for a specific watershed, this may be due to region-specific characteristics that may not be representative of a larger area. However, the goal for the AOR development would be to identify multiple highly instrumented watersheds located in geographically and climatologically distinct regions. In this way, the combined analysis would provide insight into the skill under different conditions.

The following locations have been identified for consideration for this type of analysis:

- ESRL Hydrometeorology Testbed (HMT) locations (e.g., American River, Russian River, and HMT Southeast)
- Distributed Model Inter-comparison Project (DMIP) locations (e.g., Oklahoma)
- Colorado Basin River Forecast Center (CBRFC) Snow Modeling and Data Assimilation (SMADA) Testbed

9.3.1 Evaluation of Techniques for the Early Period of Record

A similar technique could be employed to evaluate different estimation techniques for the early period of record before radar data became available (i.e., approximately 1980-1996). Gridded precipitation and temperature datasets could be generated for the recent, more highly instrumented period of record using the techniques that are determined to be most effective based on all available data. This dataset could then be used for comparison of methodologies that will be employed for precipitation and temperature estimation for the early period. The data sources available only in recent years (e.g., radar data) could be withheld, and precipitation and

temperature grids could be generated without those inputs after 1996. The results could be compared against the datasets generated utilizing all available data to provide insight into the estimation quality during the early period.

9.4 Comparison with Previously Developed Datasets

It may also be insightful to compare the AOR against other gridded or basin average datasets that have been produced. This type of comparison would be beneficial for indicating the degree of consistency between methodologies, for identifying errors in the computational procedures, and for identifying uncertainties in the estimates.

Multiple datasets have been developed that could provide insights into the quality of the precipitation and temperature datasets developed for the AOR:

- ***PRISM climatology and monthly grids:*** The monthly climatological datasets developed with the PRISM system are considered the best representation across the nation of long-term average precipitation, maximum temperatures, and minimum temperatures. These datasets provide an excellent source of information for comparison of the AOR at a fine resolution, as the climatology grids are available on a 400-m grid mesh across the CONUS. The final methodology for the AOR development will likely include adjustments to maintain the long-term means of a baseline climatology, which may be the PRISM datasets. However, in some cases, Riverside has identified discrepancies between the basin average characteristics from PRISM and the precipitation or temperature required to simulate streamflows at NWS forecast locations; this indicates that the PRISM values are subject to errors. Recognizing that adjustments to the PRISM datasets may be necessary in some locations, the PRISM climatology grids should be used as a validation dataset.

  The PRISM system has also been used to generate gridded monthly time series on a 4-km grid mesh from 1895 to present. These datasets could be used for validation by accumulating the AOR grids to a monthly time step and comparing the results to the PRISM grids. This type of analysis could help to identify shifts in the AOR grids due to changing inputs or computational procedures.
- ***Calibration MAPs and MATs:*** The NWS RFCs have developed 6-hr MAP and MAT time series for the sub-basins in their forecast areas for input into lumped hydrologic models. The MAP and MAT time series typically employ observations from major national networks (e.g., COOP, SNOTEL, and RAWS). The time series developed using station observations typically begin in 1948 or 1978 and extend to the mid 2000s. Some RFCs have developed historical MAP time series using radar precipitation estimates; these time series typically begin in 1996 and extend to the mid or late 2000s. The MAP and MAT time series are adjusted during calibration as needed to reduce biases in the resulting streamflow simulations at the basin outlets. Because of this evaluation, the MAP and MAT time series are likely to reflect the long-term basin characteristics reasonably well. Furthermore, the RFCs often complete more rigorous quality control on the input data than may be feasible for the AOR development.

  A comparison between basin-average time series computed from the AOR and the historical MAP and MAT time series would provide valuable information. If

significant biases are identified, the AOR procedures may need to be adjusted. Furthermore, the degree of difference between the two time series would indicate the relative importance of reviewing the hydrologic model calibration prior to employing gridded inputs. It is likely that in many cases, recalibration of the hydrologic models must be carried out with the new gridded inputs.

- ***Long-term stable datasets (e.g., HCN, CRN):*** Some networks such as the HCN and the CRN have been designed and maintained to produce high quality, continuous, datasets with long historical records. The characteristics of these stations are well known. The long-term precipitation and temperature time series recorded at these stations could be used to identify artificial shifts in characteristics in the AOR dataset for specific points, aggregations of points, or areas (e.g., Hamlet and Lettenmaier, 2005).
- ***Gridded datasets:*** Although PRISM is generally accepted as the best gridded precipitation and temperature climatology, other datasets ***(***described in ***Section 3.5)*** have been developed that cover the region of interest and could provide alternative comparison information. In addition, several RFCs are currently working to develop high resolution, gridded datasets. For example, the APRFC is working to develop a 10-yr precipitation dataset in Hawaii using gage observations (OHD, personal communication, January 31, 2012). The CBRFC is working to develop 30-yr gridded precipitation and temperature datasets that are consistent both with their lumped calibration forcings as well as their real-time forcings, and documentation on this effort is forthcoming (CBRFC, personal communication, February 19, 2012). The CNRFC is currently documenting their process for developing gridded forcings for input to CHPS; that documentation should be available in April 2012 (CNRFC, personal communication, February 27, 2012).

  The available gridded datasets could also serve a secondary purpose related to uncertainty estimation. Since each dataset was developed using different methodologies and input datasets, the grids have the potential to yield different precipitation and temperature characteristics. In areas where there is little difference between the PRISM climatology grids and alternative datasets, one would expect the values to have a higher level of certainty. On the other hand, if the available climatology grids yield widely varying precipitation and temperature estimates for a given location, this indicates greater uncertainty. Sapiano et al. (2008) utilized differences in precipitation estimates from different sources to quantify the uncertainty in NWP model and satellite precipitation estimates. The uncertainties were then used to merge the data sources into a single product.

  A similar approach could be used for the AOR development to quantify the uncertainty in the available precipitation and temperature climatology grids at different locations. This quantification of uncertainty in the long-term averages could be used to determine how much emphasis to place on the selected climatology versus the long-term averages from the AOR for determining long-term bias adjustments.

9.5 Hydrologic Simulation

One of the difficulties with evaluating the quality of precipitation and temperature fields is the lack of independent meteorological data for the evaluation. Often, the data that could be

used for evaluation are also used in the generation of the dataset. Streamflow measurements provide an independent means of evaluating the precipitation and temperature fields through hydrologic simulations. While other factors can affect observed streamflows, such as regulation and groundwater interactions, in many watersheds there is a clear relationship between precipitation and temperature and the resulting streamflows that can be represented well in a hydrologic model. Various lumped and distributed hydrologic models have been developed and calibrated for watersheds across the nation, most notably the hydrologic models employed at the NWS RFCs. The VIC distributed hydrologic model has also been widely implemented. Since the NWS RFCs are the primary intended recipients of the AOR dataset, and the RFCs have lumped hydrologic models calibrated across most of the U.S., the following discussion focuses on evaluation using these hydrologic models. Many of the same principles apply using other lumped or distributed hydrologic models for evaluation of the AOR.

9.5.1 Potential Applications

An evaluation based on the quality of hydrologic simulations would serve multiple purposes. First, the hydrologic models could be used to compare simulation quality between the historical MAPs and MATs developed for model calibration and basin-average values computed from the AOR. This assessment may necessitate parameter adjustments to tune the models for each input. This assessment would help to identify the impact of changing the model inputs, as well as to identify potential biases in the AOR datasets.

A second application is to utilize the existing hydrologic models to test the quality of precipitation and temperature inputs developed using different algorithms. Model parameter adjustments may be required to provide a fair comparison. This application would provide a means to evaluate alternative algorithms before finalizing the methodology for the AOR development.

A third application is to assess the impact of changes in the AOR over the period of record. There are distinct periods when the input datasets to the AOR will change significantly, such as the incorporation of radar data in the late 1990s. The hydrologic simulations could be reviewed across the period of record to assess the impact of changing inputs on the simulation quality and biases.

A fourth application is to use the hydrologic simulations to identify long-term biases in the precipitation and temperature datasets, and to use this information to make adjustments to the underlying methods or data inputs. Acceptable ranges have been developed for the hydrologic model parameters. If the models require parameter adjustments beyond these ranges, that finding would suggest the input time series are biased. The biases identified based on the hydrologic simulations could be used to make adjustments to the inputs for the AOR, for example adjusting the background climatology that was used to determine the long-term characteristics of the AOR datasets. The specific adjustments should be made based on the confidence associated with the hydrologic simulations, and should be based on the hydrologic simulations for larger groups of basins.

9.5.2 Evaluation Approach

Evaluation of the AOR using hydrologic simulations will provide valuable feedback regarding the input and output datasets, resulting in improved methodologies and outputs that better meet the needs of the hydrologic modeling community, in particular the RFCs. In the past, various RFCs have conducted inter-comparisons of different precipitation and temperature inputs and their impact on hydrologic simulation quality, such as comparisons of gage-only and gage-radar precipitation inputs. However, such analyses have been limited and have varied in approach and detail between RFCs. The AOR could potentially result in a significant change in the approach used for historical precipitation and temperature estimation. There is an opportunity with the AOR development to apply a systematic evaluation approach based on hydrologic simulations.

First, the RFCs could identify basins in their jurisdiction that meet the following criteria:

- ***No (or minimal) streamflow regulation.*** Ideally, the basins will be undisturbed to prevent anthropogenic influences from impacting the results. However, some basins with urban development should be included in the evaluation.
- ***Consistent simulation results using historical MAPs and MATs.*** Basins should not be selected if shifts in runoff characteristics have occurred during the period intended for the AOR. For example, changes in baseflow characteristics due to groundwater pumping can interfere with the evaluation.
- ***Identifiable local runoff contribution.*** Both headwaters and local areas should be selected if the local runoff response can be identified.
- ***Full range of hydrologic responses.*** The basins that are selected should represent the range of hydrologic conditions, for example both rainfall-runoff basins and snowmelt-runoff basins should be selected, as well as basins with a wide range of times of concentration.
- ***Full range of simulation quality.*** Basins should be selected where the hydrologic models perform well, as well as problematic basins where the simulation quality can be improved from better precipitation and temperature inputs.

While the number of basins included in the analysis will depend upon available resources, a sufficient number should be selected so that the evaluation results are representative. If only a few basins are considered, the validity of any conclusions may be limited, since various external factors and basin-specific issues could impact the results. A larger number of basins would require a more automated, optimization-based approach to the evaluation process. In this case, it may be more difficult to evaluate the results and to draw common, clear conclusions. The best approach may be to include a sufficiently large sample that represents the range of hydrologic conditions and reduces the impact of basin-specific issues, yet allows a detailed analysis of the results at each location.

The number of basins could also vary based on the stage of development at which the evaluation is completed. During initial development, the purpose of the evaluation is to provide guidance for selecting appropriate methodologies. In this case, a smaller sample may suffice. Once the final methodology is selected, a larger sample may be used, for example to identify bias adjustments to the inputs.

Once the basins have been identified, the following steps could be used to evaluate the simulation quality:

1. Set up the hydrologic models with plots and statistical outputs, for example using the NWSRFS calibration files and the Interactive Calibration Program (ICP).
2. Establish baseline simulations using the hydrologic models and the historical MAP and MAT time series.
3. Change the inputs to the hydrologic models to be MAP and MAT time series computed from the AOR datasets.
4. Assess the initial modeling results, which have been produced using identical model parameters.
    a. This initial assessment provides an indication of the potential biases in the input time series, and the degree to which the hydrologic simulations are impacted by the change in inputs.
    b. The statistics used to assess the simulation quality will vary by region. For instance, seasonal volumes may be most important in snowmelt-runoff areas, whereas peak flow statistics may be more critical in rainfall-runoff areas.
5. Adjust a limited number of model parameters to improve the simulation quality.
    a. The parameters that need to be changed would vary based on the hydrologic regime and the simulation sensitivities for a region.
    b. The parameter changes could be done using manual calibration or an optimizer tool such as the NWSRFS OPT3 program. The use of an optimization tool may be less labor intensive, but can yield results that are more difficult to interpret. Automated optimization should employ multi-objective criteria for the best results (CBRFC, personal communication, February 22, 2012). If parameter changes are limited to major model parameters, manual calibration could also be completed relatively quickly.
6. Assess the adjusted modeling results using the same statistics as were employed in Step 4. This assessment provides an indication of the potential benefit from the AOR datasets.

It may be possible to use an automated simulation tool, such as the Adjoint Baseline Optimizer (AB-OPT; Seo et al., 2009), to execute multiple lumped simulations to assess the overall quality of the forcings input.

Once the AOR datasets are generated for the full spatial extent and historical period, an evaluation needs to be conducted across each RFC to identify basins where the changes in precipitation and temperature result in significant deterioration in simulation quality, perhaps necessitating re-calibration.

9.6 Discussion and Implications for the AOR

As different methods and input datasets are explored for developing the AOR, natural questions arise regarding how to evaluate the results of the AOR, to select the most appropriate methodologies, and to ensure the results are spatially and temporally consistent. The answers to those questions are not simple or straightforward. Even in well-instrumented areas, it is difficult to know the average precipitation and temperature values with a high degree of certainty over a

1-km area on a 1-hr time step. In poorly instrumented areas, the problem is compounded. Although various methods are proposed to evaluate the AOR, each has strengths and limitations. A combination of the proposed methods would yield the most comprehensive evaluation of the AOR and would help to ensure the best estimate of the true precipitation and temperature values over the historical record.

Several issues require additional exploration to further clarify the evaluation approach. The proposed evaluation approach using hydrologic simulations focuses on the lumped hydrologic models run at the RFCs because the RFCs will be the primary users of the AOR datasets, because the RFC models cover most of the nation, and because significant resources have been devoted to calibrate the models. However, given that the AOR datasets will be gridded, it may be more appropriate to perform the evaluation using distributed hydrologic models to identify problems on a finer resolution. Distributed models may be more sensitive to discrepancies in different meteorological inputs. Similarly, the lumped hydrologic models developed for the RFCs typically represent relatively large watersheds. Completing an evaluation on smaller watersheds may yield different results.

Another issue concerns the incremental value of including additional data for the AOR development. Additional datasets (e.g., citizen-observed networks) increase the spatial and temporal resolution of available information. However, the additional data sources may be less reliable. At some point, the additional data may increase the noise and uncertainty in the resulting datasets. One challenge will involve determining an appropriate point to stop adding data due to the need for increasingly complex quality control measures and relatively small gains in the resulting accuracy.

In selecting methodologies, input datasets, and the evaluation approach, the NWS will ultimately need to grapple with the larger purpose of the AOR datasets. Various constraints, such as available resources, the timeline for completion of the datasets, and the immediate versus long-term needs of the datasets, may limit the effort that can be put into developing the datasets. However, if the AOR datasets are intended to replace the gridded precipitation and temperature datasets that have already been developed, a detailed evaluation is necessary. The evaluation would involve significant resources to complete a thorough analysis, however the quality of the resulting datasets would be increased, users could be made aware of limitations of the datasets, and there would be more confidence in the results. An important outcome from the AOR development process is the establishment of a formal evaluation procedure that identifies geographically-representative test basins and objective evaluation metrics that can be used to quantify the value of new data, inputs, and models (CBRFC, personal communication, February 22, 2012).

10 Existing Weather Data Management Systems

Significant resources have been devoted by personnel at federal and academic institutions to develop systems that collect, quality control, and process weather data. The development approach for the AOR should maximize the use of available frameworks, automated processes, methodologies, and datasets to reduce the development cost and effort. Many of the relevant weather data management systems have been developed within NOAA by the NCDC and the NWS (i.e., NOHRSC, NSSL, OHD, and NCEP). Additional development work has been performed jointly by the NCEP and NASA to integrate satellite data and NWP model outputs

with land surface modeling. Additionally, the PRISM Climate Group at Oregon State University has developed one of the most widely used climatological datasets.

The purpose of this section is to summarize the system features that have potential utility for the AOR. The data in ***Table 11*** provide a comparison across the systems based on criteria that affect how readily the desired features could be applied to the AOR. The attributes in ***Table 11*** describe whether the existing systems:

- ***Are being run operationally.*** If an existing system is being run operationally, it likely has automated pieces that can be adapted for the AOR.
- ***Are integrated into the RFC operational systems.*** Because the AOR is intended to support river forecast operations, a system or data source that has been integrated into the RFC operations is advantageous.
- ***Cover the spatial extent intended for the AOR.*** If the system has been applied to a limited spatial extent, additional development will be required to apply it for the AOR development.
- ***Meet (or exceed) the spatial and temporal resolution requirements that have been defined for the AOR.*** If the system has been applied using coarser spatial and temporal resolutions, additional development will be required to apply it for the AOR development.
- ***Generates both gridded temperature (T) and precipitation (P) products.*** Some of the systems presented have been developed only for precipitation, not temperature. Some of the systems work with maximum and minimum temperature grids, rather than average or instantaneous temperature grids.
- ***Have the capability to produce multi-sensor temperature (T) and precipitation (P) estimates.*** The AOR is intended to develop multi-sensor precipitation and temperature grids so existing data fusion capabilities are advantageous.

Table 11. Comparison of Weather Data Management System Attributes

System	Is run operationally	System is integrated with the RFCs	Data are integrated with the RFCs	Covers the CONUS	Includes AK, HI, PR	Includes MX, CN	Grid mesh ≤ 1 km	Temporal interval ≤ 1 hr	Produces gridded T and P	Produces multi-sensor T and P
NCDC	✓		✓	✓	P	P		P	✓ (1)	P
PRISM			✓	✓	✓	P	P		✓ (1)	
SNODAS	✓		✓	✓		P	✓	✓	✓	P
NMQ/Q2	✓	P	✓	✓		P	✓	✓	P (2)	P
Daily QC/ Mountain Mapper	✓	✓	✓	✓	✓	✓		P	✓	
MPE	✓	✓	✓	✓	✓	✓		✓	P (2)	P
CPC	P			✓	P	P		P	✓ (1)	P
NLDAS	✓			✓		P		✓	✓	

✓ System meets this criterion.
P System partially meets this criterion.
1 System provides daily maximum and minimum temperature grids, not instantaneous or average temperature grids.
2 System provides precipitation grids, but not temperature grids.

10.1 NCDC Automated Data Quality Control Procedures

The NCDC is the official federal repository for historical weather data. As such, the NCDC has long had responsibility for quality controlling station observations. Historically, the quality control process included extensive manual reviews that delayed the publication of data (Rennie et al., 2011b). To reduce the amount of effort required for quality control, and to eliminate inconsistencies that can result from manual data review, the NCDC has been developing automated data quality control systems for point observations that could be useful for the AOR.

In addition to the operational systems that have been developed to quality control observations at COOP stations (described below), the NCDC is also developing automated procedures to perform quality control on hourly precipitation data in the DSI-3240 dataset. This system includes automated checks to identify data spikes, global extremes, gaps, and climatological outliers (Rennie et al., 2011b). Once the development work is completed, these quality control procedures will replace the current process of manual review and editing for the DSI-3240 hourly precipitation dataset.

10.1.1 TempVal

An automated temperature validation system (TempVal) has been integrated with the operational COOP temperature validation system. The TempVal system uses hourly and daily observations at ASOS and AWOS stations to quality control daily observations of maximum and minimum temperature at COOP stations (Angel et al., 2002).

First, a benchmarking dataset is developed using the ASOS and AWOS observations. Gridded anomaly fields are developed by computing departures between the maximum (or minimum) temperature observations and the station's climatological value. The anomalies are interpolated to a 0.5° grid mesh using inverse distance squared weighted interpolation. The gridded anomalies undergo an automated "bull's-eye" test to compare each grid cell value to the average value of surrounding grid points.

The daily COOP data are also used to develop gridded anomaly fields. The daily maximum (or minimum) temperature values at the COOP stations are used to compute anomalies from the station's climatological value, which are then interpolated using the same technique applied to the ASOS/AWOS anomalies. If the COOP anomaly for a grid cell differs from the ASOS/AWOS anomaly by more than a specified threshold (currently 7°F plus the gradient of the ASOS field computed from surrounding cells), the COOP observation is replaced by a value that is computed by applying the ASOS/AWOS anomaly to the COOP station's climatological value.

The TempVal system was run in parallel with the previous operational system to assess the performance. NCDC personnel concluded that the automated quality control procedures performed better in 90% of the cases where large temperature differences existed (Angel et al., 2002). However, the TempVal system does not perform well when strong temperature gradients exist due to inversions, warm air advection episodes, and strong frontal passages. The NCDC has developed a list of stations that are affected by these weather phenomena and therefore undergo limited spatial review. Due to a lack of benchmarking data, the TempVal system is not applied in Puerto Rico, the Virgin Islands, or other U.S. territories.

10.1.2 PrecipVal

Similar to TempVal, the NCDC has developed an automated quality control process for precipitation observations. While the precipitation validation system (PrecipVal) can be applied to quality control any data source, it has been applied primarily to the COOP network (Urzen et al., 2004).

The PrecipVal system develops a gridded benchmarking dataset on a 1-hr temporal interval and the nominal 4-km grid mesh associated with the HRAP grid. The benchmarking dataset is developed using information from multiple data sources, including point observations from the ASOS and CRN networks and gridded estimates from the RUC NWP model, GOES satellite, and NEXRAD radar network (Urzen et al., 2004). All precipitation layers are compared to identify outlying values (i.e., an inter-layer comparison test). For each grid cell, the final precipitation estimate is an average value computed from the precipitation estimates that survived the inter-layer comparison test. Statistics, including the number of input data sources, the standard deviation of the input values, and the inner quartile range of the input values, are computed for each grid cell to assign a confidence level to the estimated precipitation value.

Point observations are compared to the gridded benchmarking dataset to identify outlying data values. Thresholds are applied to determine if the point observations are in agreement with the benchmarking dataset. In addition, the statistics computed during the development of the benchmarking dataset are used to convey the confidence associated with the validation process.

10.2 PRISM

The PRISM system developed at Oregon State University employs point observations and multiple linear regression techniques to develop gridded estimates of precipitation, maximum temperature, and minimum temperature (Daly et al., 2002; Gibson et al., 2002). The PRISM system includes automated data quality control procedures and gridding techniques that may be beneficial for the AOR. In addition, the gridded climatological datasets that have been developed using the PRISM system may be useful for gridding point observations using CAI, or spatially disaggregating coarse-resolution datasets.

The PRISM system includes point observations from numerous networks, including the COOP, SNOTEL, RAWS, Agrimet, Weather Bureau Army Navy (WBAN), California Data Exchange Center (CDEC), and Environment Canada networks. In a given month, a station must have at least 85% non-missing data values to be included in subsequent analyses. Due to a lack of heating elements, observations from RAWS stations are incorporated into the system from May-September only.

The data quality control features in the PRISM system have been developed to identify errors, not catch deficiencies or inconsistencies due to changes in station locations or sensors. The PRISM system flags potentially erroneous data using climatological range checks. Daily precipitation values that are greater than 115% of the state's record daily precipitation are flagged, while minimum (maximum) temperature values are flagged if they are 3°C less (greater) than the state's record daily values (Daly et al., 2008). The PRISM system also flags temperature values if the temperature differences are less than 0.1°C over a ten-day period.

The PRISM system uses spatial interpolation and multivariate regression techniques to develop expected values to which station observations are compared (Daly, 2006). The process developed for the PRISM system (ASSAY_QC) applies cross validation to identify potential data errors on a monthly time step. The differences between the predicted and observed values are normalized to account for the fact that higher variances are expected for higher precipitation values (Gibson et al., 2002). The quality control process is used to set the observed values to missing, to replace the observations with the predicted values, or to linearly combine the observed and predicted values (Daly et al., 2005).

The PRISM system was developed assuming that elevation is the most important factor in distributing temperature and precipitation values in a localized region. Station weights are used to control the effects of variables other than precipitation and temperature. In the general PRISM formulation, the combined weight of a station is a function of distance, elevation, station clustering, coastal proximity, topographic facets, vertical layer effects, and terrain effects. A station is weighted less when it is far from the cell being estimated, when it is at a different elevation, when it is near other stations, or when it has a different topographic facet.

Both the significant benefits and some of the drawbacks associated with the outputs from the PRISM system stem from the fact that elevation is the primary factor controlling weather patterns. Studies have shown that elevation is not always a significant predictor of precipitation (e.g., Guan and Wilson, 2005), or that even if elevation is a significant predictor, it may only

explain 25% of the variation in winter precipitation (Peck and Schaake, 1990). In addition, the quality of the PRISM grids suffers in areas with sparse station networks. In order to maintain a water balance on a watershed scale, Riverside has sometimes adjusted the long-term characteristics derived from PRISM that were used to develop calibration MAPs, especially in small watersheds and in areas with sparse station networks. The CBRFC has developed software that can be used to adjust the 800-m climatological grids from the PRISM system based on the calibration climatology. For example, the CBRFC has adjusted the temperature climatology grids from the PRISM system, with adjustments ranging from 0-4°F, to match the climatology developed for their synthetic temperature stations (CBRFC, personal communication, February 19, 2012). The adjusted climatology grids were then used in the DailyQC program to develop historical temperature grids that were input to the Research Distributed Hydrologic Model (RDHM). The hydrologic simulations were significantly improved with the adjusted climatological grids (CBRFC, personal communication, February 19, 2012).

Efforts continue by the PRISM Climate Group at Oregon State University and other groups to evolve the PRISM system. For example, some work has been done to develop a daily time series of PRISM grids for the period 1960-2001 on a 4-km grid mesh over the CONUS (DiLuzio et al., 2008). The monthly PRISM grids were the starting point for this application. For each station, a daily fraction of the monthly precipitation total was computed. The fractions were interpolated using inverse distance weighting, and then were multiplied by the monthly PRISM grids to develop daily grids. One limitation of this work is it used only the COOP station network.

Additionally, the PRISM system was used to develop daily grids of minimum temperature, maximum temperature, total precipitation, rainfall, snowfall, solar radiation, and radiation-adjusted maximum temperature on a 50-m grid mesh for a small watershed in the Cascade Mountains in Oregon for the year 2003 (Daly et al., 2007). The quality of the grids, particularly precipitation and solar radiation, was affected by data scarcity and data quality issues.

Personnel at ESRL are working to improve upon the PRISM system by incorporating atmospheric variables into the regression models (Hsu et al., 2011). This work includes developing a nonlinear, hierarchical model on a daily time step to identify additional controlling variables beyond those already incorporated into the PRISM system. For example, atmospheric and oceanic variables such as vertical motion, sea surface temperature, Froude number, and water vapor are being evaluated to determine if they improve upon the general PRISM formulation.

The NCDC recently released climatological averages for the period 1981-2010. It is likely that the PRISM system will be used to develop gridded climatological datasets for this period as well, though the schedule of this work is unknown (Colorado Climate Center, personal communication, October 13, 2011). Because the period 1981-2010 more closely aligns with the desired period of record for the AOR, it would be advantageous to have these datasets available for the AOR development.

10.3 SNODAS

The SNODAS developed by the NOHRSC employs NWP model outputs, point observations, and satellite data to drive a spatially distributed snow model “developed to provide the best possible estimates of snow cover and associated variables to support hydrologic

modeling and analysis (Barrett, 2003)." The SNODAS is currently run on a 1-km spatial scale and a 1-hr temporal scale for the CONUS, although the system will be expanded to cover Alaska (NOHRSC, personal communication, July 20, 2011). The SNODAS includes automated data quality control procedures, spatial downscaling methods, and data assimilation techniques that may be useful for the AOR development.

The meteorological inputs to the SNODAS include barometric pressure, solar radiation, incident longwave radiation, air temperature, wind speed, relative humidity, precipitation type and quantity, and cloud coverage. The radiation fields are obtained from the National Environmental Satellite, Data, and Information Service (NESDIS) Continental Scale International Project (GCIP). The remaining meteorological inputs are obtained from the RUC NWP model. The RUC outputs are downscaled from a nominal 13-km grid to the 1-km grid used by the SNODAS. The downscaling process is specific to the variable. Precipitation fields from the RUC NWP model are downscaled using bilinear interpolation (NOHRSC, personal communication, July 20, 2011). The temperature fields are downscaled using bilinear interpolation, and are then adjusted for terrain differences. Additionally, the downscaled temperature grids are adjusted using station observations prior to input to the SNODAS (NOHRSC, personal communication, July 20, 2011).

Two major components within the SNODAS are a spatially distributed snow model that estimates snowpack characteristics and a data assimilation algorithm that adjusts snowpack characteristics based on observations. The NOHRSC snow model is a hybrid of SNTHERM and the Utah Energy Balance Model (NOHRSC, personal communication, September 2007; USACE, 2007). The snow model simulates snowpack states and fluxes for each grid cell based on energy- and mass-balance principles. During the data assimilation process, the snow water equivalent (SWE) states are adjusted using observations of snow cover and SWE in a process referred to as "nudging." The estimates include satellite-derived, airborne, and ground-based observations.

Eight data products are available from the SNODAS: liquid and frozen precipitation, SWE, snow depth, snow pack temperature, snowmelt, snow pack sublimation, and blowing snow sublimation. Daily grids are produced for all eight products. Hourly grids are produced for all data products except snow pack temperature and blowing snow sublimation.

The primary limitation with the SNODAS as currently implemented is that the spatial downscaling procedure applied to the RUC precipitation field does not make bias corrections or account for topographic features that affect precipitation patterns. For example, in a previous study, Riverside found that the downscaled RUC precipitation fields showed little variation with elevation, tended to underestimate winter precipitation, and tended to overestimate summer precipitation (Riverside, 2007). This analysis was performed for selected watersheds in the Upper Colorado River Basin and may not be representative of results in other parts of the country.

10.4 NMQ/Q2

The NMQ/Q2 system developed by the NSSL produces a national mosaic of radar precipitation estimates, in addition to quantitative precipitation estimates and short-term quantitative precipitation forecasts. The NMQ/Q2 system is being run in real-time by the NSSL with high reliability, but the system will not be considered fully operational until a redundant version of the system is implemented at the NCEP. In addition, a separate research version of the NMQ/Q2 system is being run at the ESRL. The NMQ/Q2 system is currently being run on a 1-

km grid mesh and a 5-min temporal interval over the CONUS and portions of Canada and Mexico. Future developments will include extending the NMQ/Q2 system to cover Puerto Rico, and reducing the grid mesh to 500 m. The NMQ/Q2 system includes automated data quality control procedures (particularly convective screening); radar processing, mosaicking, and bias correction algorithms; and a framework for data assimilation that may be beneficial for the AOR development.

The NMQ/Q2 system currently ingests precipitation estimates from NEXRAD radars, a limited number of TDWR radars, and Canadian radars; GOES satellite data; precipitation observations from stations; outputs from the RUC NWP model; and lightning data (NSSL, 2011b; Zhang et al., 2005). The radar data form the backbone of the NMQ/Q2 system. The radar reflectivity data undergo a series of quality control checks to eliminate artifacts (e.g., noise, sunbeams, bright bands; Zhang et al., 2005) and to perform corrections based on the vertical profiles of reflectivity. Lightning data are used to delineate areas of convective and stratiform precipitation. The precipitation estimates from individual radars are used to develop a national mosaic of radar precipitation (i.e., NMQ). Station observations are used to develop mean field bias-adjusted and local bias-adjusted radar precipitation products, in addition to gage-radar multi-sensor products.

The NMQ/Q2 system has excellent radar quality control and mosaicking methods, producing a high quality national radar mosaic. The quality control routines are excellent at removing ground clutter, although they sometimes result in the removal of light precipitation that is actually occurring (Southeast River Forecast Center [SERFC], personal communication, December 14, 2011; Story, 2011). The NMQ/Q2 products suffer from limited radar coverage and inaccurate estimation of frozen precipitation.

Future areas of enhancement include improving the vertical profile of reflectivity corrections that were developed in the plains region, to provide corrections that will perform better in complex mountainous terrain (ESRL, personal communication, November 16, 2011). Although the use of gage and radar data has been good, improvements can be made in the application of satellite and NWP model outputs.

The NCDC and the NSSL might soon generate a 10-yr reanalysis dataset using the NMQ/Q2 system, which would improve the potential benefit for the AOR development by providing a longer historical record.

10.5 DailyQC/Mountain Mapper

The DailyQC program developed by the CBRFC is used to quality control observations of temperature, precipitation, and freezing level on 6-hr and 24-hr time steps (CBRFC, 2009; NWS, 2008; NWS, 2009). The DailyQC program was developed to perform quality control in areas with complex mountainous terrain, but is also applied in areas with smaller elevation gradients. The DailyQC program has been integrated into the MPE and the Mountain Mapper programs. The automated data quality control procedures in DailyQC may be useful for the AOR development.

The DailyQC program performs automated spatial and temporal consistency checks to flag questionable values. Observations are compared to surrounding observations at nearby stations to assess their spatial consistency. A gage whose value is more than a specified number of standard deviations different than surrounding values is flagged as questionable (NWS, 2008). Temporal consistency checks are used to ensure that the 6-hr and the 24-hr data are consistent.

Data values that are flagged by the automated quality control procedures are denoted using a quality control flag that is set to "questionable." These values require subsequent manual review.

The DailyQC program includes a spatial interface to display the observed values, station elevation, station network, sensor type, spatial and temporal consistency metrics, and quality control flags (e.g., verified, questionable, estimated, bad, missing) to facilitate the manual review process. Additional functions are available through the interface to apply range checks, remove precipitation values from stations with tipping bucket gages if the air temperature is less than the freezing level, and to review relationships between the observations and the station elevations.

In the version of DailyQC that has been integrated into the MPE program, the 6-hr precipitation values can be disaggregated to a 1-hr time step for use in developing gridded precipitation estimates (NWS, 2008). This temporal disaggregation process may be useful for the AOR development. In addition, an automated version of the DailyQC program (i.e., auto_daily_qc) can be used to generate gridded datasets in a batch, automated fashion (CBRFC, personal communication, February 22, 2012).

The Mountain Mapper system developed by the CBRFC generates gridded precipitation totals and instantaneous temperature values in areas of complex mountainous terrain (CBRFC, 2009; Schaake et al., 2004). The gridding techniques in Mountain Mapper may be beneficial for the AOR development. The Mountain Mapper system uses the CAI technique (***Section 6***) to transform point observations into gridded estimates using monthly climatology datasets from the PRISM system or monthly station climatologies. The program weights the closest stations the highest, unlike the predetermined station weights employed at most western RFCs using the NWSRFS calibration MAP preprocesser (CBRFC, personal communication, February 22, 2012). The instantaneous temperature grids need to be adjusted to represent average temperature values before being input to the hydrologic models run at the river forecast centers (CBRFC, personal communication, February 22, 2012).

The DailyQC/Mountain Mapper system is in widespread use at the western RFCs, and likely will represent a benchmarking system against which the AOR will be compared. For example, the NWRFC is using Mountain Mapper to estimate missing precipitation values at stations before computing MAPs using predetermined station weights in CHPS (NWRFC, personal communication, February 27, 2012). The CBRFC may be redeveloping the DailyQC program to use a grid mesh that is smaller than the standard HRAP grid, for example an 800-m grid mesh (CBRFC, personal communication, February 22, 2012).

10.6 MPE

The MPE program developed by the NWS OHD develops gridded precipitation estimates on a 1-hr temporal interval and a nominal 4-km grid mesh for an RFC service area (Seo, 2003; Fulton, 2005a; Fulton, 2005b; NWS, 2008; NWS, 2009). The MPE program is run at the RFCs as part of their river forecast operations. The MPE program generates multiple precipitation grids, from which the RFC personnel choose the best representation to input into their operational hydrologic models. In addition, the DailyQC and Mountain Mapper programs (described above) have been integrated into the MPE program to facilitate quality control of point observations and grid generation using climatological datasets. However, the MPE program does not generate gridded temperature fields. The MPE program includes automated data quality control, bias correction, and multi-sensor precipitation features that that may be beneficial for the AOR.

The MPE program ingests the Digital Precipitation Array (DPA) products from the WSR-88D radars, satellite precipitation estimates, and station observations. These data are used by the MPE program to generate the following precipitation grids:

- ***A mosaic of radar precipitation (RMOSAIC).*** This grid is generated by mosaicking the DPA rainfall estimates from the radars in the service area. By default, in areas of overlapping radar coverage, the rainfall value is selected from the radar that provides coverage at the lowest height above sea level that is free of significant beam blockage (NWS, 2008; Seo, 2003). Options are available in the MPE program to set the grid cell value to the average or the maximum precipitation value from the radars that cover the grid cell. If a grid cell is not covered by a radar, then it is assigned a missing value. Climatological values developed for each radar are used to quality control the precipitation estimates; if a precipitation value for a grid cell is questionable; the cell value is set to missing.
- ***A mosaic of mean field bias-adjusted radar precipitation (BMOSAIC).*** The rainfall estimates from each radar are adjusted using a mean field bias correction prior to mosaicking the radar estimates for the service area (Seo et al., 1999). The mean field bias correction value is computed for each radar using stations within the effective radar coverage. The mean field bias values are computed for multiple time intervals (i.e., sub-hourly, hourly, daily, weekly, monthly, seasonal, and period of record). The MPE program selects the best bias correction value among the time intervals that have a minimum number of gage-radar pairs. The threshold value is configurable, though a minimum value of 10 is the default.
- ***A mosaic of local bias-adjusted radar precipitation (LMOSAIC).*** The radar precipitation mosaic (RMOSIAC) is adjusted using a local bias correction (Seo and Breidenbach, 2002). A radius of influence (e.g., 40-km) is used to select gage-radar pairs for each grid cell. Similar to the mean field bias correction, the local bias adjustments are computed for multiple time intervals (i.e., sub-hourly, hourly, daily, weekly, monthly, seasonal, and period of record). The MPE program selects the bias correction value using the shortest time interval that has a minimum number of gage-radar pairs. The threshold value is configurable, though a minimum value of 10 is the default. The local bias adjustment factors grid can be viewed in MPE. The resulting grid of local bias values is applied to the raw radar mosaic to produce the local bias corrected radar mosaic.
- ***A gridded gage-only precipitation field (GMOSAIC***). A precipitation grid based on point observations is generated using the Single Optimal Estimator approach (SOE; Seo, 1998a; Fulton, 2005b). A radius of influence parameter is used (e.g., 60 km) to estimate the precipitation value for a grid cell; if no stations are located within 60 km of the grid cell, it is assigned a missing value. This also means that isolated gages can produce “bull’s-eyes” in the final grid. If the grid cell is within 60 km of multiple gages, the grid cell value is determined by weighting the gage values based on their distance from the cell and making a climatological adjustment by comparison of the grid cell climatology to the station climatologies. The default climatology datasets originate from the PRISM system, although climatological data from other sources may be used instead (NWS, 2008).

- ***Gridded gage-radar analyses (MMOSAIC, MLMOSAIC).*** The default gage-radar precipitation grid (MMOSAIC) is generated by linearly weighting the mean field bias-adjusted radar mosaic (BMOSAIC) and station observations and then making a climatological adjustment (Seo, 1998b). For a grid cell, the final precipitation estimate is a function of the BMOSAIC value at the grid cell and observations at nearby stations that are weighted as a function of distance from the cell. If the grid cell contains a station, it is assigned its observation value; if the grid cell is near several stations, their values are heavily weighted compared to the radar precipitation value; if the cell is far from stations, the radar precipitation value is heavily weighted. If a grid cell has no radar coverage or nearby stations, the estimated precipitation value is filled using climatological differences based on the PRISM dataset. The MMOSAIC layer is the default layer for the best precipitation estimate unless the user changes the settings in the MPE program.

  A gage-radar analysis can also be performed using the local bias-adjusted radar precipitation grid (LMOSAIC) rather than the mean field bias-adjusted radar mosaic (BMOSAIC). In this case, the multi-sensor precipitation estimate is referred to as the local-bias adjusted multi-sensor mosaic (MLMOSAIC).

Satellite precipitation values can be used to generate gridded precipitation estimates using the same options described above for radar precipitation. The satellite products are intended primarily to provide precipitation estimates in areas that lack radar and station coverage.

The user interface developed for the MPE program permits interactive quality control for the station observations and the precipitation grids. For example, once the RFC has selected a precipitation layer as the best estimate of precipitation, the layer can be altered using multiple techniques such as substituting precipitation values from other layers for a polygon specified by the user.

Some of the MPE features described above are designed for real-time use. For example there is no fully-automated QC of gage or radar input within the package. However, a more flexible, offline version of MPE is under development within OHD, which can ingest other inputs.

Some of the limitations in the MPE program relate to the radar precipitation processing. Removing ground clutter from the DPA products can be time-consuming (SERFC, personal communication, December 14, 2011). Neither the NEXRAD precipitation processing system nor the MPE program include some of the quality control processes for the radar data that are included in the NMQ/Q2 system, such as the corrections based on vertical profiles of reflectivity.

Ongoing improvements to the MPE system are focused on the data quality control procedures, including developing additional automated procedures, applying multi-sensor techniques, and incorporating lightning and NWP model outputs into the quality control procedures.

10.7 CPC

Researchers at the CPC are developing methods that may have relevance for the AOR in the following areas (CPC, personal communication, December 21, 2011):

- Improving satellite precipitation estimation through development work on the IMERG algorithm.
- Developing historical and real-time precipitation grids on a 1-hr time step for the CONUS using gage observations.
- Developing global historical and real-time temperature grids on a 1-day time step using gage observations.
- Merging satellite estimates and NWP model outputs.
- Merging satellite estimates, NWP model outputs, gage observations, and radar precipitation into a single multi-sensor estimate.

The hourly gage-only precipitation analysis under development will be constructed by first interpolating hourly station data to the HRAP grid using optimal interpolation, and then adjusting the hourly analysis to match a daily gage analysis (Xie et al., 2007). The dataset will initially cover the period 1995-present, with plans to extend the analysis back to 1948. The CPC plans to have a prototype of this analysis completed by the end of 2012, and is currently working to finalize parameters associated with the interpolation algorithm.

The pole-to-pole CMORPH project is focused on merging the KF-CMORPH with NWP model outputs on a global basis. The CPC has investigated using the CFSR, MERRA, and ERA-Interim reanalysis datasets for this purpose, and will likely select the CFSR due to the finer spatial and temporal resolutions of this product. The initial framework for this merging process has been developed; ongoing work is focused on developing appropriate bias adjustments and uncertainty estimation techniques.

An optimal interpolation-based algorithm forms the basis of an ongoing project to develop a multi-sensor, merged hourly precipitation analysis (Xie and Xiong, 2011). The product will incorporate precipitation data from stations, radars, satellites, and NWP models using the HRAP grid over the CONUS for the period 1995-present, and is anticipated to be completed in 2-3 years.

The CPC produces daily grids of maximum temperature and minimum temperature on a 0.5° grid mesh for the CONUS. The real-time analysis uses point observations from the RFCs, HADS, and the Climate Anomaly Database (CPC, 2012b). If observations are available at fewer than 350 stations, the temperature analysis is not performed for the day. The point observations are interpolated to a 0.5° grid mesh using a modified Cressman technique. In 2012, the CPC intends to produce a global temperature analysis with a 0.5° grid mesh and a daily time step for the period 1979-present (CPC, personal communication, February 2012).

10.8 NLDAS

The NLDAS developed by NASA and NCEP is intended to “construct quality-controlled, and spatially and temporally consistent, land-surface model datasets from the best available observations and model output to support modeling activities (NASA, 2011).” The NLDAS develops forcings datasets to input into multiple land surface models that simulate surface fluxes, soil moisture, and snow cover. The NLDAS has automated quality control, gridding, and downscaling procedures that may be relevant for the AOR development.

The forcing inputs to NLDAS, including precipitation and temperature, are generated for a 1-hr temporal interval and a 1/8° grid mesh. The forcing datasets have a spatial extent that covers central North America (NASA, 2011a). The forcing inputs (except for precipitation) are

derived from the NCEP NARR reanalysis (Cosgrove et al., 2003). The 32-km, 3-hr fields from NARR are interpolated to a 1/8° grid and are temporally disaggregated to a 1-hr temporal interval. The surface pressure, longwave radiation, air temperature, and specific humidity fields are adjusted vertically to account for terrain differences between the NARR and the NLDAS. The vertical adjustment applies a lapse rate of 6.5 K/km for air temperature.

For the precipitation input to NLDAS-2, the CPC daily precipitation product (which includes an adjustment based on the PRISM climatological grids) is used as the starting point. The gage-only daily precipitation analyses are processed to fill missing values. The daily estimates are temporally disaggregated using the NCEP Stage 2 radar product, CMORPH satellite estimates, the CPC hourly precipitation dataset, or the NARR precipitation fields (in that order) depending on data availability for specific periods of time.

Since daily gage and hourly precipitation data are sparse over Canada, the NARR precipitation field was used over Canada within the NLDAS domain. For a 1° swath along the U.S.-Canada border, the NLDAS precipitation forcing value is a weighted combination of the CPC-derived and the NARR precipitation values.

11 Technical Considerations

This section discusses the temporal and regional influences on data availability and quality that will affect the development of a consistent, long-term AOR.

11.1 Temporal Considerations

The AOR is intended to be developed for the period 1979-2010, with eventual extension back to 1948. Within this historical period, there are several notable dates that affect the amount and quality of the available data:

- ***1948:*** Digital records for the NCDC COOP network become available, representing a significant increase in spatial representation. The NCEP R1 reanalysis also begins in 1948, providing coarse-resolution information on atmospheric conditions.
- ***1978:*** The major reanalysis products (e.g., CFSR, NARR, and ERA-Interim) become available, providing higher quality, higher resolution information on atmospheric conditions. These reanalyses use NWP models and data assimilation techniques that are consistent over the reanalysis period, although the assimilation inputs may have varied over the period.
- ***1979:*** The NRCS SNOTEL network becomes available, representing a major advance in monitoring at high elevations. Western RFCs have experienced difficulty estimating MAP and MAT values at high elevations prior to the introduction of the SNOTEL network (CBRFC, personal communication, February 19, 2012).
- ***1990s:*** The number of automated weather stations rapidly increases, representing a significant improvement in weather characterization on a sub-daily timescale.
- ***1996:*** The NCEP NEXRAD Stage 2 hourly radar archive begins, representing a significant improvement in the spatial and temporal characterization of precipitation.
- ***1998:*** The TRMM satellite network was launched, providing significant advancements in satellite-based precipitation estimation.

- ***Mid 2000s:*** Several mesonets and citizen networks begin operation, increasing the spatial density of point observations.

Based on these dates, the quality of the AOR is expected to vary among three major periods due to differing amounts of historical data (***Table 12***).

Table 12. Historical Data Availability

Period	**Available data**
1948-1977	• Daily station data • Limited hourly station data • R1 reanalysis
1978-1995	• Daily station data • Limited hourly station data • SNOTEL high-elevation station data • CFSR, NARR, and ERA-Interim reanalyses
1996-2010	• Daily station data • Expanded hourly station data • SNOTEL high-elevation station data • CFSR, NARR, and ERA-Interim reanalyses • Radar data • Satellite data

One of the goals of the AOR is to produce a dataset that has consistent characteristics over the period of record. The consistency of the AOR can be characterized using multiple metrics:

- ***Long-term climatology.*** Any shift in long-term averages should reflect natural or anthropogenic changes in climate, and not changes caused by the data networks, datasets, or methodologies being applied to the AOR.
- ***Daily variability.*** The variability captured in the AOR should be consistent on a daily basis across the period of record.
- ***Hourly distributions.*** The AOR should reflect realistic distributions of precipitation and temperature on an hourly time step. This criterion may necessitate the use of stochastic methods for the early period when limited hourly data are available to adequately characterize weather characteristics at less than a daily time step. The AOR should also realistically capture extreme weather events across the period of record.

Temporal consistency for the period 1948-2010 is a significant, if not impossible, challenge given the major differences in data availability. Multiple approaches are available to help maintain consistency across the period:

- To correct for changes in long-term climatological averages:

    - After developing the AOR for the period 1996-2010 using all available data sources, a second analysis could be completed that uses only data sources that are available for the full period of 1979-2010. The differences in characteristics between the datasets could be applied to adjust the characteristics of the AOR for the period 1979-1996.
    - The AOR could be aggregated to a monthly time step for comparison and correction against the long-term climatological datasets developed using the PRISM system.
- To maintain consistent spatial and temporal variability throughout the period of record:
    - A reanalysis from a high-resolution regional weather model (e.g., WRF) could be used for spatial and temporal disaggregation during the early period to compensate for a lack of radar and satellite estimates.
    - A K-nn resampling approach could be applied to temporally disaggregate daily precipitation or temperature grids to an hourly time step for the period 1979-1996, where the AOR developed for the 1996-2010 period is used as the historical development sample.

Following development of the AOR, the limitations of the dataset should be clearly described for users, particularly related to the differing data sources and approaches applied for each historical period.

11.2 Regional Considerations

Given the regional differences in climate, hydrology, forecast objectives, data sources, and data quality, the AOR development process will have to address regional requirements and issues.

11.2.1 Alaska

River forecasts for Alaska are produced by the Alaska-Pacific River Forecast Center (APRFC). The forecasts are mostly concerned with flooding from multiple sources that varies in duration, although APRFC personnel run ESP and have a small water supply component to support users, for example related to hydropower (APRFC, personal communication, December 8, 2011). In addition, low flows are a concern due to barge operations. Alaska experiences significant flow contributions from glacier melt, particularly in the month of August. Therefore, the APRFC hydrologic models include the glacier model and require MAP and MAT estimates for all glacier areas.

APRFC has historically relied upon seasonal predetermined station weights for both calibration and forecast operations. In 2011, APRFC began computing MAPs using the DailyQC/Mountain Mapper program for operations and for comparison against the MAPs computed using predetermined station weights. Station-based MAP estimates are used everywhere except a small area around the city of Fairbanks that employs MAP estimates based on the mean field bias-corrected radar product from the MPE program.

The major issues in Alaska that will affect the AOR development are data scarcity and problems accurately measuring frozen precipitation. Prior to the 1970s, almost no historical weather data exist for Alaska. Most of the federal networks, including the NRCS SNOTEL network, have some stations in Alaska, but the station density is low. The largest source of hourly point observations is the RAWS network, which has a short period of record. The operational network is much smaller than the historical network used to compute calibration MAPs because many of the stations are located in remote regions and provide aggregated reports on an infrequent basis. APRFC estimates approximately 50% of the stations in the historical network are in the operational network. None of the satellite products cover Alaska except for the small portion south of 60°N. Alaska has limited radar coverage, including 7 WSR-88D radars in the south that are operated by the FAA (***Figure 11***). The radar precipitation estimates are adversely affected by terrain blockage and frozen precipitation. Historically, the radar precipitation estimates have not been archived in the NCEP Stage 4 product; the historical data are solely available from the NCDC archive for individual radar umbrellas (APRFC, personal communication, December 8, 2011). Operationally, the APRFC has been using the temperature field from the Real-Time Mesoscale Analysis (RTMA). For calibration MAPs, the APRFC has utilized the total precipitation field from NARR for remote portions of Alaska and Yukon Canada that otherwise have no data. The APRFC has not evaluated reanalysis products other than the RTMA and NARR. It is possible that the ASR-Interim product could add value to the AOR in this region.

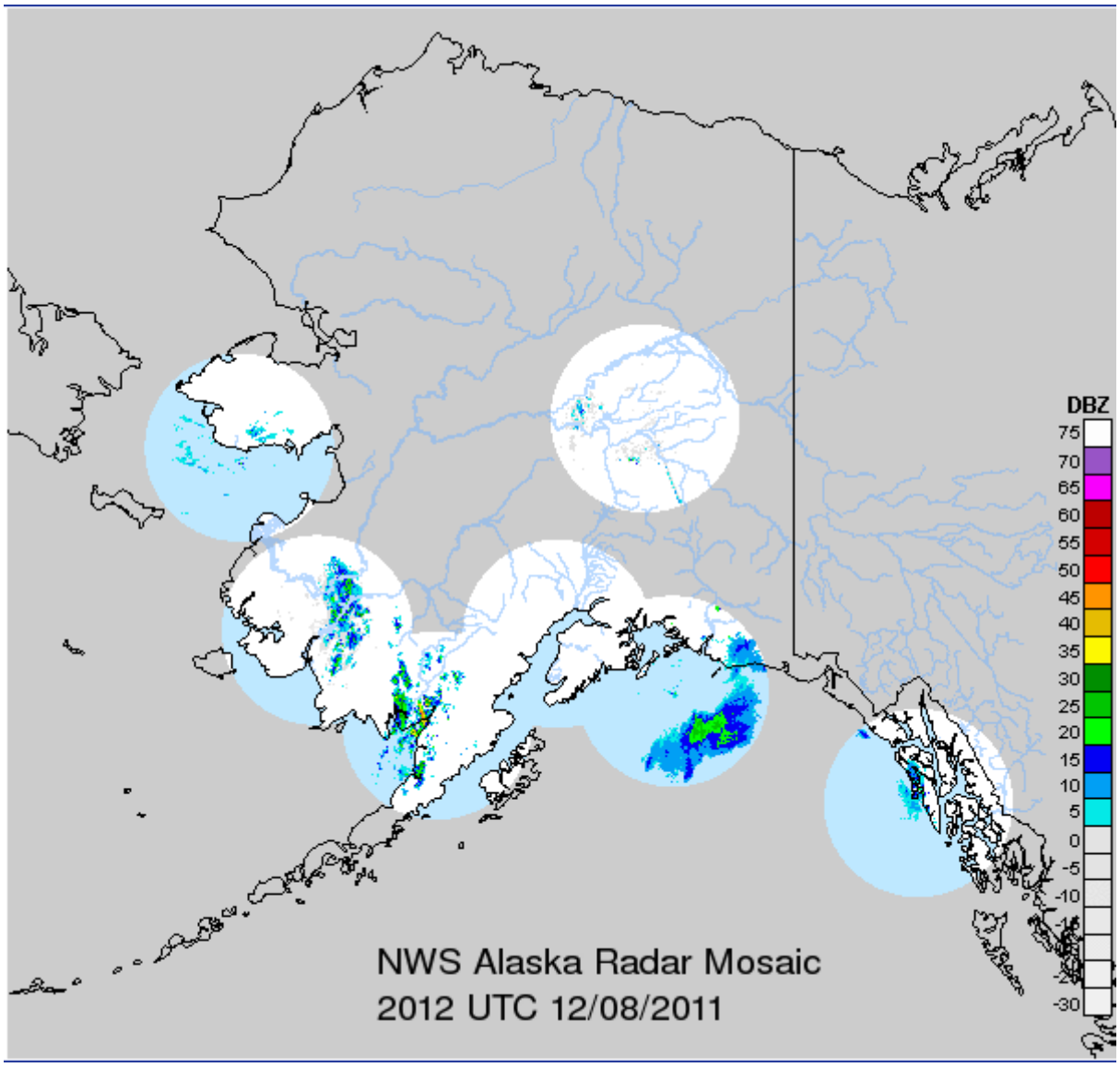

Figure 11. FAA WSR-88D Radars in Alaska (ARH, 2011)

Given the data scarcity and data quality issues, reliable gridded climate normals are difficult to develop. APRFC has historically relied on station-based climatology (APRFC, personal communication, December 8, 2011). APRFC has done some preliminary investigations of the long-term MAP characteristics compared to the PRISM climatology. One apparent issue that has been identified in the PRISM climatology grids for Alaska is that PRISM-based precipitation estimates appear to monotonically increase with elevation to a larger degree than may actually occur, and certainly to a larger degree than the MAPs indicate.

It will be difficult to develop a high-resolution AOR on a 1-km grid mesh and 1-hr temporal interval. Due to a lack of data, the AOR in Alaska may be the most consistent over the historical 30-yr period, but also may have high uncertainty due to the sparse data network. From a hydrologic perspective, the APRFC would benefit from an AOR that has at a minimum a 4-km (HRAP) grid mesh and a 6-hr temporal interval, although there are a few flashy rivers that would

benefit from hourly precipitation estimates. It may be reasonable to consider developing the AOR with limited spatial and temporal resolutions given these constraints.

11.2.2 Hawaii

The watersheds in Hawaii are small and very quick-responding. As a result, no river forecasts are issued. The NWS Honolulu WFO solely issues flash flood guidance. The WFO assumes most of the forest responsibility, though the RFC is supporting the WFO in implementing the RDHM Threshold Frequency (TF) to enhance the flash flood program (APRFC, personal communication, December 8, 2011).

In terms of data coverage, most of the federal networks include some stations in Hawaii. There are four WSR-88D radars located in Hawaii that are operated by the FAA. One radar is located on Kauai, one radar is located on Molokai, and the remaining two radars are located on the Big Island of Hawaii. The radars are adversely affected by terrain blockage. The radar precipitation estimates have not been relied upon in the past. The satellite products cover Hawaii, including the TRMM precipitation radar, though the temporal interval (i.e., 3 hr) and grid mesh (i.e., 4 km) of that product are too coarse to be useful for flash flooding purposes.

In addition to the complexities of estimating temperature and precipitation for Hawaii in areas affected by coastal effects and steep topographic relief, Hawaii also experiences temperature inversions associated with the trade winds.

Given the flash flood focus and a lack of deterministic and probabilistic river forecasts, a retrospective AOR would be used for computing the statistics required to implement the RDHM-TF, requiring at least a 10-yr historical period. Any gridded precipitation estimates developed for Hawaii must have a grid mesh of at least 1 km and a temporal interval of at least 1 hr. Finer resolution estimates would be beneficial.

11.2.3 Puerto Rico

There are 13 forecast locations in Puerto Rico (i.e., 11 headwater locations and 2 local area locations). No deterministic river forecasts are issued by the SERFC. The SERFC computes real-time precipitation inputs using the MPE program. The San Juan WFO provides quantitative precipitation forecasts (QPF) to the SERFC (SERFC, personal communication, December 14, 2011). In addition, the SERFC utilizes an experimental 6-hr QPF product from the Hydrometeorological Prediction Center (HPC) at NCEP in a qualitative capacity. The SERFC creates 1-hr, 3-hr, and 6-hr distributions of QPF, inputs them to SAC-SMA, UNIT-HG, and channel routing models, and then publishes the crest heights and times associated with each QPF scenario (SERFC, personal communication, December 14, 2011). The SAC-SMA model is applied on a 1-hr time step because the watersheds in Puerto Rico are quick-responding; the maximum crest travel time is approximately eight hours. Given these hydrologic characteristics, the forecasts are more sensitive to precipitation intensity, times of concentration, and channel travel times than the initial soil moisture states.

The data coverage over Puerto Rico is reasonably good, but suffers from quality issues. Because of the quick hydrologic response times and the one-hour simulation time step employed in the SAC-SMA models, the SERFC relies upon hourly observations from stations in the HADS network. The quality of the observations is related to station maintenance, because the tropical

environment in Puerto Rico rapidly produces vegetation that interferes with the monitoring equipment. The FAA operates one WSR-88D radar and one TDWR radar in San Juan. Despite the steep terrain relief in Puerto Rico, the radars are apparently unaffected by terrain blockage. However, the WSR-88D radar is considered unreliable by SERFC staff. The SERFC receives reflectivity data from the TDWR radar, and is working to get products representing quantitative precipitation estimates (QPE). When no radar precipitation estimates are available, the SERFC transitions to precipitation estimates that are computed solely based on station observations.

The SERFC has not evaluated satellite precipitation estimates in Puerto Rico, though they are available through the MPE program.

The NSSL is working to produce the NMQ/Q2 suite of products in Puerto Rico, but the status of this work is unclear. The SERFC is not currently receiving NMQ/Q2 products for Puerto Rico.

Climatological grids have been developed for Puerto Rico using the PRISM system. The SERFC has performed preliminary comparisons between the MAPs and basin-average values computed from the PRISM climatology datasets, but the PRISM climatology is not used in real-time and is not used to adjust the MAPs.

The primary technical challenge in computing high-resolution grids in Puerto Rico for the AOR is likely to be the steep gradients that result from the elevation changes on the island. The SERFC is likely to use the AOR once available to update the SAC-SMA model parameters (SERFC, personal communication, December 14, 2011).

11.2.4 Southeast CONUS

The southeast and central plains regions of the CONUS are relatively flat, receive little frozen precipitation, and thus have benefited greatly from radar-based precipitation estimates. In the southeast of the CONUS, the river forecast objectives are focused on potential flooding. For developing calibration MAPs, the SERFC has relied upon station observations that are processed into basin averages using the grid point weighting option available in the calibration MAP preprocessor in NWSRFS (SERFC, personal communication, December 14, 2011). Because the weather characteristics show lower spatial variability than occurs in other areas with significant topographic relief, the RFCs, including the SERFC, have not needed to adjust the calibration MAPs to an independent data source such as the PRISM climatology grids. Precipitation adjustments (via the PXADJ parameter in the SAC-SMA model) are rarely required. Because precipitation gradients are smaller, and climatological adjustments have not been applied, the RFCs may not have quantified the differences between the calibration and operational MAPs.

For real-time precipitation estimation, the SERFC switched to the MPE program about three years ago. The SERFC staff typically uses the multi-sensor product, but may choose different layers depending upon the gage density for a given area. Sometimes the SERFC uses the NMQ/Q2 product to replace a polygon in the MPE multi-sensor product, particularly when there is too much ground clutter in the MPE product and no rainfall occurred. The SERFC sometimes uses the P3 and the gage-only products from MPE in winter in the mountains of NC. When using the radar products, the SERFC almost exclusively relies upon the mean field bias-corrected product. Rarely is the local bias-corrected product used. Historically, the time to generate the local bias correction product was prohibitive; however, this restriction may no longer be an issue. Overall, the SERFC personnel are pleased with the radar precipitation estimates in their area and they rarely use gage-only data.

The southeast has very strong data coverage from stations and radars; the satellite and reanalysis products are rarely employed or required. Some of the technical issues that complicate the development of high-resolution temperature and precipitation grids in the southeast include coastal effects, terrain blocking both from the mountains and the coastlines, and accurate estimation of frozen precipitation. Along the coast, precipitation estimation is complicated by the small, quick-responding coastal watersheds and the sparse station networks. The radar products are affected by the sparse station coverage along the coast, as fewer stations are available under the radar umbrellas to compute reliable bias adjustments. Moving forward, the deployment of dual polarization capability on radars in the southeast will affect radar precipitation estimation and the consistency of real-time values with historical values.

The most significant benefit from the AOR for areas that already utilize gridded precipitation estimates from remotely-sensed sources is maximizing consistency with the gridded inputs that will be generated in real-time at the National Water Center. In addition, the SERFC intends to develop the SNOW-17 model for the forecast locations in the mountains of VA, NC, and SC. The intent of these snow models is to take advantage of precipitation typing, and to avoid issuing flood forecasts when heavy precipitation occurs as snow rather than rain. The snow models are not required to accurately accumulate and deplete snowpacks. However, the implementation of these snow models indicates that the temperature grids in the AOR may be beneficial for more areas than have employed temperature data in the past.

11.2.5 Intermountain West CONUS

The intermountain west of the CONUS is defined to include the mountainous, snowmelt portions of the West Gulf, Missouri Basin, Arkansas-Red Basin, Colorado Basin, California Nevada, and the Northwest RFCs. In these areas, long-term seasonal water supply forecasts are an important potential application for the AOR. In addition, peak flow forecasts are very important in many areas. The primary difficulty in developing high-resolution climate grids is the steep gradients that occur over short distances in these areas of sharp elevation differences. Gutmann et al. (2011) state the problem very well: "For hydrologic simulations, the inability of a 10-km forcing dataset to adequately resolve important topographic variations in the Colorado Rockies presents serious limitations for snowpack modeling." Fowler et al. (2007) cited a study that found that a grid mesh of 0.125° is required for hydrologic simulations of monthly flow in mountainous catchments.

The major issue in developing high-resolution grids for the AOR is orographic effects by which moisture is forced by the topography to rise, generally resulting in increasing precipitation with elevation and more precipitation on the windward side of topographic features and less precipitation on the leeward side. However, orographic effects depend on elevation, topographic features, and the moisture flux direction (Guan et al., 2005). Due to complex climatic processes such as channeling and convergence, it does not always hold that precipitation increases with elevation (Guan et al., 2005; Hughes et al., 2009).

The primary source for both historical and real-time precipitation estimates has been station observations, particularly valuable high-elevation observations recorded at the NRCS SNOTEL stations. In some cases, western RFCs (e.g., CBRFC, WGRFC) have limited the period of historical forcings due to poor estimation of MAPs and MATs at high elevations prior to the introduction of the SNOTEL network (CBRFC, personal communication, February 19, 2012). Radar precipitation estimates have not been widely used due to issues with sparse coverage,

difficulties in accurately estimating frozen precipitation, and a tendency for the radar beams to overshoot precipitation (CNRFC, personal communication, February 27, 2012). Because it is more difficult to compute precipitation and temperature characteristics in complex terrain, the western RFCs typically have done more work comparing the long-term characteristics of the calibration MAPs and MATs to climate normals, typically derived from the PRISM system, and also tying the real-time estimates to the same climatology, ensuring higher consistency between the historical and real-time estimates.

The largest potential for benefit from the AOR for the intermountain west is development of gridded temperature and precipitation estimates, which have not been in widespread use to date.

11.2.6 Canada

The Northwest River Forecast Center (NWRFC) and APRFC have developed hydrologic models for the portions of Canada that flow into the U.S. The primary forecast objective for the NWRFC is water supply, while the APRFC is more concerned with the potential for flooding. For calibration MAPs, the APRFC has utilized the total precipitation field from the NARR reanalysis dataset for remote portions of Alaska and Yukon Canada that otherwise have no data (APRFC, personal communication, December 8, 2011).

Environment Canada maintains a network of surface stations as well as a network of weather radars. The historical station data can be obtained through an on-line tool, or can be ordered through the climate service. Most of the station data have undergone some quality control. The historical precipitation values may only be available on a daily time step, whereas the temperature data are available on hourly and daily time steps. The weather radar network includes 31 radar sites that cover 98% of Canada's populated areas (Environment Canada, 2011). The radar precipitation data appear to be available for 10-min increments back to 2007. At least some of the station and radar data have been made available through the Advanced Weather Interactive Processing System (AWIPS). Many of the satellite products cover most of Canada (up to a latitude of 60°N), but may not be high quality due to inaccuracies in determining frozen precipitation and location errors due to parallax effects.

The technical issues in developing the AOR over Canada are likely to be similar to those enumerated for Alaska, including data scarcity, poor data quality, difficulty estimating frozen precipitation, and difficulty estimating weather in areas of complex terrain. In addition, it may be more difficult to obtain the required historical data. The APRFC and NWRFC may have some of the historical data archived.

11.2.7 Mexico

The WGRFC is concerned with the potential for flooding on the Rio Conchos and the Rio Bravo (i.e., Rio Grande), which impacts downstream forecast locations on the Rio Grande in the U.S. in Texas.

The CONAGUA agency in Mexico maintains multiple station databases, including the CLICOM Daily Surface Data Set, the Hydrologic Information System (SIH), and automated meteorological stations. The CLICOM Daily Surface Data Set provides a long-term historical dataset of point precipitation and temperature values as far back as 1902. There are about 3,300

stations, of which approximately 900 report daily. However, weekend reporting is poor when observers are away from their posts, and missing data are often reported as zero values. The SIH is an operational database that extends back to 1995. It is used by CONAGUA for the collection of precipitation, temperature, and streamflow data at instantaneous and daily stations. The automated meteorological station network includes 136 automated meteorological stations and 44 automated synoptic meteorological stations. The archived data extend back to 1999 and include 10-min precipitation and temperature values. The Mexico Agriculture Automated Weather Station Data network maintained by AGROSON provides a full suite of meteorological variables in real-time; historical data begin in 2004. The NCAR has a historical dataset from the North American Monsoon Experiment (NAME) for the period 2002-2005. CONAGUA also maintains a radar network that includes 13 sites, but only small portions of the Rio Conchos and Rio Bravo watersheds appear to be covered (***Figure 12***).

Figure 12. Mexico's Radar Network (CONAGUA, 2011)

The biggest challenges in developing a high-resolution AOR over Mexico will include a scarcity of data sources and availability of historical weather data. The WGRFC may have some of these data archived, but data delivery has been intermittent in the past. The historical data can be obtained from various sources, namely by contacting CONAGUA, from the NCAR Earth Observing Laboratory, or from websites.

Mexico has developed climatological grids using a process similar to the PRISM system (i.e., the Atlas Climatico Digital de Mexico). Riverside has successfully used the climatological datasets to develop calibration time series for the Rio Sabinal.

12 RFC Considerations

The AOR is intended for use by the RFCs in their river forecast operations. As such, there are implications for the hydrologic modeling that influence a suitable development approach for the AOR. For example, the AOR needs to be consistent over the historical period to support both model calibration and the Hydrologic Ensemble Forecast System (HEFS). Ensemble streamflow prediction (ESP) is one component of HEFS. To the extent that the AOR differs from the legacy MAP and MAT forcings used to calibrate the existing models, the model inputs or parameters may need to be adjusted during a re-calibration process. Additionally, the AOR needs to be consistent with the forcings that will be generated in real-time at the National Water Center to avoid introducing biases into the forecasting process.

12.1 Consistency over the Historical Period

The AOR dataset will be developed to be consistent with accepted reference observations over the entire historical period to the extent possible. However, because the available data sources vary over the historical period, the quality of the AOR will also vary over time if all data sources are used. Any change in quality over the historical period affects the utility of the AOR for hydrologic model calibration and HEFS.

The minimum period that is required for model calibration varies by location and is a function of the climate variability for that location. The calibration period that is selected should include average, wet, and dry years. In regions where the lumped hydrologic models perform satisfactorily, 10 years is typically sufficient to adequately capture the variability in conditions that affects parameter estimation (Anderson, 2002). In semi-arid regions with infrequent events, a 10-yr period may not be adequate for model calibration. Beyond 25 years of record, there is little incremental benefit for improving model parameter estimation. Additionally, the longer the period, the more likely that the characteristics of the watershed being modeled have changed (e.g., urbanization, changes in consumptive use), which can adversely affect the calibration results.

For HEFS, an extended historical record is desirable to capture the range of possible future weather conditions, and additional years are needed to estimate the extremes of the conditional forecast distribution. The RFCs have typically used a minimum of 30 years of historical record for this purpose, although some RFCs are using 60 or more years of historical record for producing probabilistic forecasts. Although the AOR may have significant uncertainty in the period 1979-1996 in terms of representing the actual temporal and spatial distributions at an hourly time step, the dataset should have less uncertainty at six hour and daily time steps, which are temporal resolutions of more importance for ensemble prediction.

It is possible that a shorter period within the AOR may be recommended to the RFCs for determining model parameters (e.g., 1996-2010), while the full period is suitable for HEFS. This recommendation may vary by region, depending upon the changes in available data inputs and the hydrologic characteristics of the region. Alternatively, the RFCs have indicated a desire to have access to multiple datasets: one that represents the “best available” dataset and employs all available input data sources, and a second “consistent” dataset that employs only the data inputs that are available for the full historical period (CBRFC, personal communication, February 22,

2012; CNRFC, personal communication, February 27, 2012; NWRFC, personal communication, February 27, 2012).

12.2 Consistency with Legacy Forcings

The AOR has the potential to differ from the legacy forcings developed at the RFCs in terms of inputs, methods, and spatial and temporal intervals. As a result, the RFCs will need to evaluate and understand the differences between the AOR and the forcing datasets that were used for model calibration, and the impact of using the AOR on the quality of hydrologic simulations in their service areas. Depending on the results of this evaluation, the RFCs may need to make adjustments to the model inputs or parameters.

In terms of input data sources, much of the hydrologic model calibration performed to date at the RFCs has been done with gage-based MAP and MAT time series. Some model calibration has been done using gage-radar estimates. Satellite-based precipitation inputs and NWP model outputs have been used on a limited basis in data sparse areas (such as Northern Mexico and Alaska) where alternatives are very limited. Because the AOR may incorporate different data sources than have been used historically, it can be expected to have both random and systematic biases compared to legacy forcings used to calibrate the hydrologic models.

Additionally, many of the hydrologic models run at the RFCs utilize a 6-hr simulation time step. The RFCs could choose to aggregate the AOR datasets to a 6-hr time series for use with their existing models. Alternatively, the RFCs could choose to move to a 1-hr simulation time step. Because the parameter values are dependent on the defined spatial and temporal scales, a change from a 6-hr time step to a 1-hr time step would necessitate adjustments to the model parameters (Smith et al., 1996).

The NWS OHD could provide several features with the AOR datasets to provide the RFCs with maximum flexibility and to minimize the potential disruption to operations. Firstly, the RFCs could be provided with a tool to adjust the AOR datasets to an existing climatology (CBRFC, personal communication, February 22, 2012). Such a tool would allow the RFCs to utilize information from their existing calibrations and to transition to using the AOR datasets even if re-calibration of existing models is not immediately feasible. Secondly, the AOR datasets could be provided using multiple time intervals, for example ranging from one hr to one day, to support different applications (CNRFC, personal communication, February 27, 2012; NWRFC, personal communication, February 27, 2012).

The level of effort to conduct a comprehensive evaluation of the AOR will vary by RFC, but could require significant resources. Based on a limited number of discussions with RFC personnel, this is a necessary and worthwhile process to address inconsistencies between the AOR and legacy datasets, and to ensure consistency between the calibration and real-time forcings. Ideally, a mechanism should be established to allow the RFCs to submit feedback to the NWS OHD, particularly during the initial AOR development and evaluation process. The RFCs could identify deficiencies and suggest changes to inform both the AOR development process and the real-time forcing process.

12.3 Summary of RFC Requirements

Riverside conducted interviews with the APRFC, SERFC, CBRFC, CNRFC, and NWRFC to solicit input on RFC requirements for the AOR. The following notable points were identified through those discussions:

- The AOR datasets must be consistent with the real-time forcings and forecast grids provided to the RFCs. Consistency includes climatological characteristics, spatial resolution, and temporal resolution. For example, the forecast grids from the NDFD are currently provided on a 2.5-5.0 km grid mesh; the forecast grids should be provided on the same grid mesh intended for the AOR.
- The RFCs expressed interest in having access to multiple versions of the AOR: a "best" dataset that uses all available data, as well as a "consistent" version that uses limited data.
- The RFCs expressed interest in having access to the AOR datasets at multiple temporal intervals that range from one hr to one day. Some locations would benefit from an even finer temporal interval, such as 15 minutes, but sub-hourly resolution is not practical for the initial AOR development. Regional differences and intended usage affect the appropriate time interval.
- The RFCs would benefit from having historical gridded datasets that expand beyond temperature and precipitation to include freezing level, wind speed, relative humidity, etc.
- The RFCs exhibit regional preferences in using gage and radar data. The framework selected to develop the AOR should avoid favoring one data type. This requirement is consistent with an uncertainty-based data fusion framework.
- To strengthen the likelihood that the AOR will be adopted within the RFCs (superceding their own forcing datasets), it will be essential to survey the RFCs regarding quality control procedures and data networks that are currently incorporated into their forcing analyses, and to explore ways that these can be incorporated into the AOR development (CBRFC, personal communication, March 7, 2012).
- The RFCs anticipate that the AOR will evolve as data sources and quality control techniques improve. The AOR should be regenerated as new information becomes available and should be extended over time to keep the datasets current.
- The RFCs would benefit from tools that enhance the flexibility of the datasets to meet RFC needs, for example a tool that would allow the RFCs to adjust the AOR to a desired climatology.
- Overall, the RFC personnel are enthusiastic about the potential benefit of the AOR datasets. The RFC personnel who were interviewed expressed a willingness to work with the OHD to provide regional expertise, to provide available data, to evaluate the AOR, and to evaluate the need to re-calibrate the existing models.

13 Summary and AOR Recommendations

The NWS OHD intends to prepare a high-resolution, long-term historical archive of gridded precipitation and temperature. This AOR will be developed for a ≤1-km grid mesh and a 1-hr temporal interval for the period 1979-2010, with eventual extension back to 1948. The AOR

will include the CONUS, Alaska, Hawaii, Puerto Rico, and the portions of Canada and Mexico that drain to the U.S.

The primary purpose of the AOR is to support the development and calibration of lumped and distributed hydrologic models for river forecast operations at the RFCs using CHPS. The AOR is also intended to support meteorological and hydrologic applications for other federal agencies under the IWRSS initiative. To support these intended uses, the AOR must maintain statistical stability over the period of record, as represented by the long-term characteristics, daily spatial variability, and hourly temporal variability. The AOR needs to be consistent with the real-time forcings process that is being established at the National Water Center to provide temperature and precipitation grids to the RFCs in real-time.

13.1 Data Sources

Riverside developed an inventory of data sources that provide historical temperature and precipitation data in the CONUS, Alaska, Hawaii, Puerto Rico, Mexico, and Canada (***Appendix A***). The data inventory includes gage, radar, satellite, and reanalysis products. Historical precipitation data are available from all of these sources, while historical temperature data are limited to gage, satellite, and reanalysis products. Supplemental data sources such as lightning observations and additional NWP model outputs can provide useful information on storm direction, the presence of convective activity, and other factors that influence the precipitation and temperature distributions.

Historical gage observations are available for the full period of record intended for the AOR. However, the amount and quality of available data changes significantly with the introduction of digital records (1948), the introduction of high-elevation stations (1978), and improved spatial and temporal characterization due to the introduction of automated stations and citizen-observed weather stations (1990s-present). Station observations require a significant amount of quality control to identify and remove data errors and to ensure temporal consistency throughout the historical record. For use in the AOR, daily observations will need to be temporally disaggregated, and point observations will need to be transformed to gridded values.

Historical radar precipitation products became available in the mid 1990s. The RFCs have traditionally relied on precipitation estimates from WSR-88D radars in the NEXRAD network, although some use has been made of precipitation estimates from the TDWR radars operated by the FAA. Most of the historical radar products were produced using the HRAP grid coordinate system, which has a nominal grid mesh of 4 km. Therefore, the radar products will need to be spatially downscaled, or reanalyzed, to the target resolution intended for the AOR. Radar processing algorithms, including quality control, precipitation estimation, and bias corrections, have improved over time. Therefore, the historical record of radar precipitation needs to be reprocessed to apply common procedures, a task that the NWS OHD has already begun. Radar precipitation estimates have not been used widely in the intermountain west of the CONUS or in Alaska due to terrain blocking and limitations in accurately determining frozen precipitation. The challenges in assessing frozen precipitation rates using radar data remains an issue for the AOR. As noted above, a high-resolution reprocessed dataset for the period 2002-2010 using the NMQ algorithm might soon be available. OHD staff are also working to generate a quality-controlled and bias-adjusted radar dataset for the period 1996-2001.

Historical satellite data became available in the late 1990s. The precipitation estimates have been developed using multiple sensor types, including visible, infrared, passive microwave,

and active microwave (precipitation radar) sensors. Like radar precipitation, algorithms for quality controlling satellite data and estimating precipitation have improved substantially throughout the historical record. The satellite products typically extend to 50-60°N, meaning that coverage is limited over Alaska and Canada. However, satellite estimates would have limited utility in these areas as they tend to be inaccurate for frozen precipitation. Although satellite precipitation algorithms are improving, the quality of the precipitation estimates remains relatively low compared to other sources due to the indirect nature of the precipitation estimates and other issues. Historically, precipitation estimates from satellite data have been used by the RFCs only in areas with very sparse station networks. Of the available historical satellite products, the CMORPH dataset appears to have the highest potential for the AOR (including the KF-CMORPH dataset that will soon be reprocessed). The 30-min temporal interval makes the CMORPH dataset useful for temporal disaggregation of other data sources. However, the CMORPH data are available on an 8-km grid that will require spatial downscaling to the target grid mesh intended for the AOR. The TRMM precipitation radar dataset may also be beneficial.

NWP models have been used to develop reanalysis datasets. Most of the reanalysis datasets exist for the period 1979-2010, although two datasets begin prior to 1979. Thus, the reanalysis datasets are the only data sources besides point observations that exist in the early period intended for the AOR. The reanalysis datasets have been produced on relatively coarse grid meshes ranging from 10-124 km, indicating they will require spatial downscaling to the AOR target mesh. Temporal disaggregation may or may not be required for these datasets, as they have been produced using time steps ranging from 1-3 hrs. The challenges in applying the reanalysis datasets to the AOR include developing suitable bias correction techniques and applying appropriate spatial downscaling methods so that physiographic features influencing local weather patterns are captured in the downscaled product. The CFSR reanalysis dataset or the NARR reanalysis dataset appear to have the most value for the AOR given the spatial and temporal resolutions, recent advances in modeling and data assimilation techniques, and an assessment of the dataset performance. In addition, the ASR-Interim dataset may have value for developing the AOR in Alaska.

Existing gridded precipitation and temperature datasets may be useful for the AOR development in terms of transferring existing methodologies, using the gridded products to establish the long-term characteristics of the AOR, or using the gridded products as validation datasets. These datasets include products that have been developed using the PRISM system, the Hamlet and Lettenmaier dataset developed at the University of Washington, and the operational products produced by the NCEP, among others.

13.2 Data Quality Control

The data sources used to produce the AOR must undergo quality control procedures to remove erroneous data and to correct for temporal inconsistencies that can have a significant adverse impact on hydrologic simulations. While manual review of data observations is quite effective, it is preferable to implement an automated data quality control system given the intended spatial and temporal extents of the AOR.

Automated processes can be applied to identify metadata errors that may cause weather characteristics for a location to be misrepresented, to assess the amount of missing data for a given data source, and to remove data values that are questionable under specific circumstances (e.g., frozen precipitation). Existing systems include automated data quality checks that can be

categorized as range checks, climatological range checks, internal consistency checks, spatial consistency checks, multi-sensor checks, checks between predicted and observed values, and temporal consistency checks.

Automated techniques in all of these categories should be adapted for the AOR. However, the quality control processes for the AOR need to be adaptive to avoid applying inappropriate techniques (e.g., spatial consistency checks during convective activity) and to avoid removing data values in data sparse regions that cannot be reliably validated.

13.3 AOR Framework

Information from each input data source needs to be merged to produce a combined output for the AOR. In data fusion, a model is used to merge multiple data sources to produce an improved estimate. This approach has been utilized in a wide range of applications. ***Figure 13*** presents a generalized framework for the development of a data fusion system:

- In the ***identification*** step, the input data sources are characterized (e.g., long-term characteristics, biases, uncertainties) and appropriate methods to combine the data are defined. The identification step can incorporate information from ***data mining*** techniques (e.g., ANNs, principal components analysis [PCA]).
- The ***estimation*** step applies the methods to merge the input data sources. The data fusion process can be applied to data from individual sources or datasets that have been processed.
- Once the data fusion process is complete, a ***validation*** step employs ***benchmarking*** datasets and a ***performance assessment*** to identify weaknesses in the processing.
- The ***administration*** feedback loop incorporates changes into the identification and estimation steps to address the weaknesses.

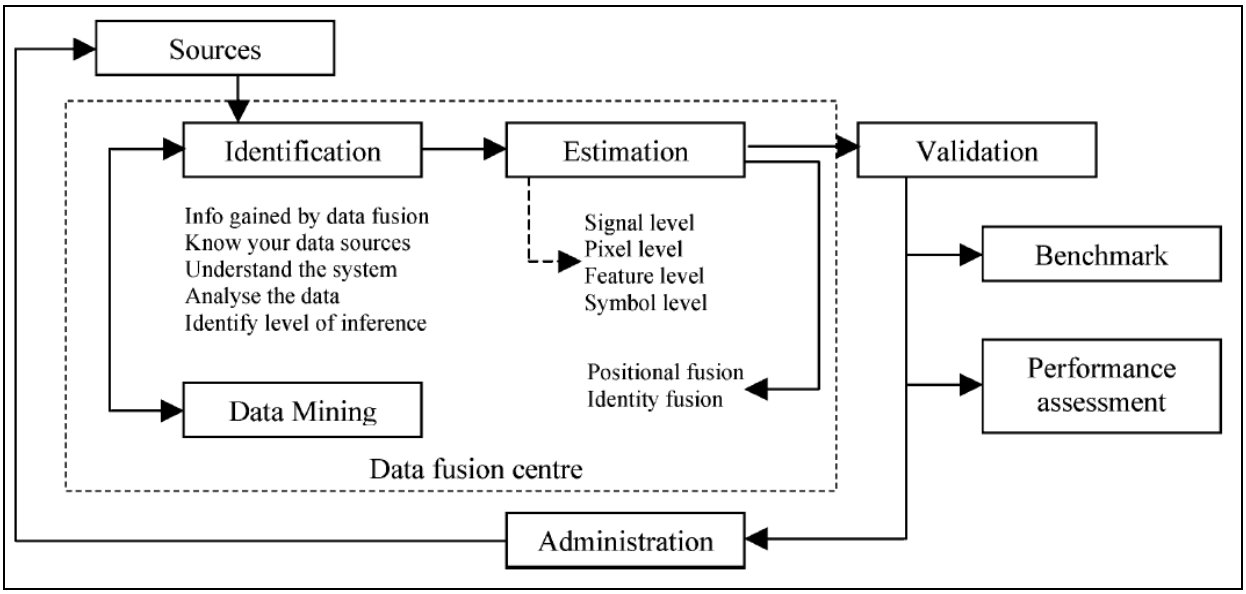

Figure 13. Conceptual framework for the development of a data fusion system (reproduced from Esteban et al., 2005)

Figure 14 presents one possible data fusion framework that could be applied to the AOR.

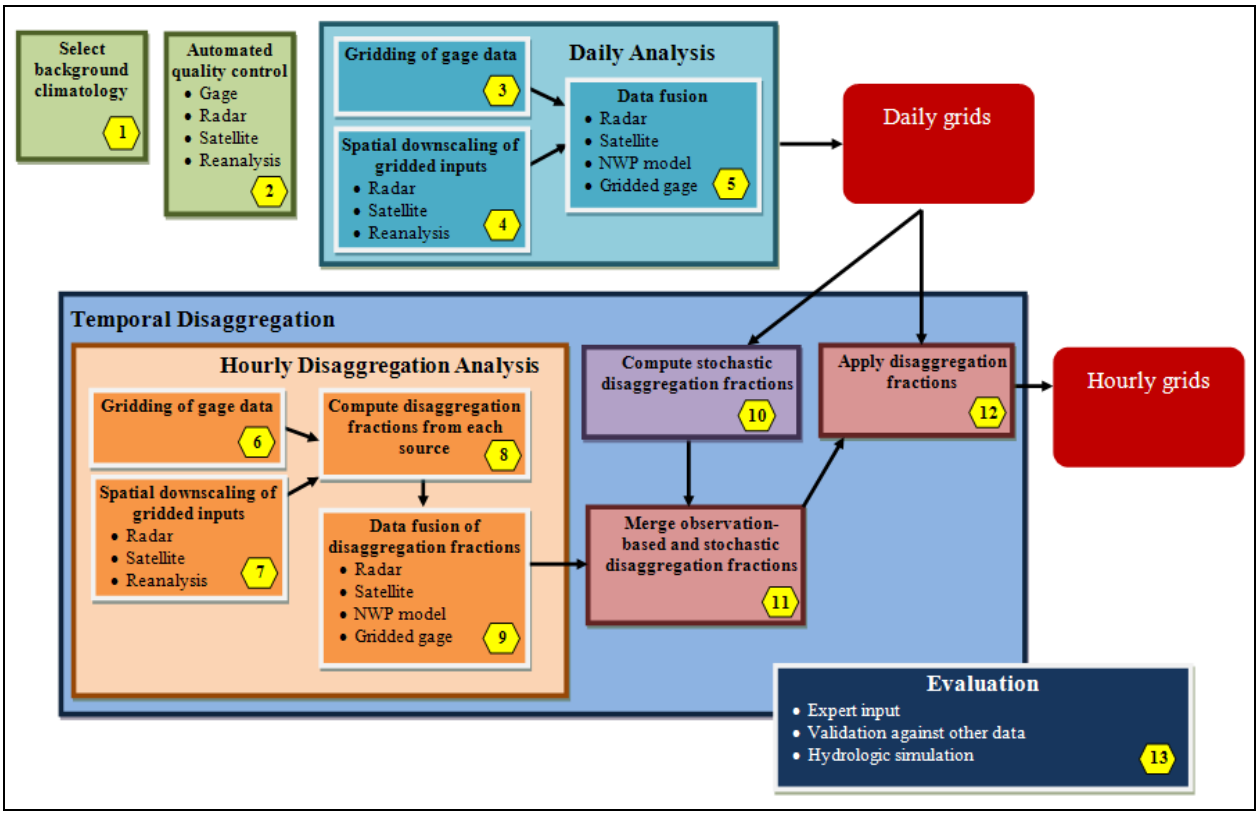

Figure 14. Potential data fusion framework for the AOR

- In Step 1, a ***background climatology*** dataset is selected, recognizing that the climatology dataset may be adjusted in an iterative fashion upon evaluating the AOR.
- In Step 2, the input data sources undergo ***automated data quality control***.
- In Steps 3-5, all data sources are processed to a daily time step and the target grid mesh to generate a merged daily product.
  - In Step 3, the daily point observations are transformed to a gridded field using ***gridding methods***.
  - In Step 4, all gridded datasets are downscaled to the target grid mesh using ***spatial downscaling methods***.
  - In Step 5, the gridded gage fields and the downscaled grids are merged using uncertainty characteristics in a ***data fusion*** process.
- In Steps 6-12, temporal disaggregation weights are developed and applied to produce grids on the target time interval using ***temporal downscaling*** methods.
  - In Step 6, grid sets available on an hourly time step are downscaled to the target grid mesh using ***spatial downscaling*** methods.
  - In Step 7, the hourly point observations are transformed to a gridded field using ***gridding methods***.

- In Step 8, the hourly grids are used to compute hourly temporal disaggregation weights.
- In Step 9, the gridded disaggregation weights from each source are merged using uncertainty characteristics in a ***data fusion*** process.
- In Step 10, a stochastic process is used to produce hourly disaggregation weights that maintain a defined set of statistical characteristics that are determined based on the specific approach applied.
- In Step 11, the disaggregation fractions determined in steps 9 and 10 are merged using a ***data fusion*** process. The disaggregation fractions computed using observations are applied in regions that have dense hourly gage networks or radar coverage, while the disaggregation fractions computed using a stochastic process are applied in regions with high uncertainties in the observed hourly distributions.
- In Step 12, the hourly disaggregation fractions are applied to the daily analysis to produce the final hourly gridded field.

- In Step 13, the final hourly datasets undergo an ***evaluation*** process, which may include comparisons (and adjustments) based on long-term climatological datasets, cross validation, hydrologic simulation quality, and expert input. If the evaluation process reveals weaknesses in the development approach, the process may be refined.

The framework in ***Figure 14*** shows how the data quality control processes, gridding methods, spatial downscaling methods, temporal disaggregation methods, data fusion techniques, and evaluation methods fit together to produce the AOR. However, the framework in ***Figure 14*** does not represent the only solution. For example, a separate daily gage-only analysis may not be required; gage data could be directly fused with the gridded data sources.

13.3.1 Data Assimilation/Data Fusion

Data assimilation and data fusion methods are used to merge multiple data sources into a single product. Data assimilation methods incorporate observations into a NWP model to improve the model outputs. Precipitation observations cannot currently be assimilated directly into NWP models; precipitation observations are indirectly assimilated by adjusting other model states, such as the latent heating of the atmosphere, in an attempt to force the model to produce precipitation at the location and time represented by the observation.

Data fusion uses a statistical model to combine multiple data inputs into a merged product. Data fusion techniques can be categorized as multi-sensor bias adjustments, simple data fusion techniques, optimization-based techniques, and uncertainty-based techniques, recognizing that there is overlap between the categories.

Quantifying uncertainties for application to data fusion is an active area of research and development. Riverside recommends that an uncertainty-based data fusion approach be developed for the AOR. Although the uncertainty information used to perform data fusion may be rudimentary in the early development stages, simplifying assumptions can be made that yield improved results relative to the individual data inputs, and that reflect regional quality differences in available data sources. As uncertainty estimates improve, the framework developed for the AOR can incorporate this information. It may be possible to utilize an optimization-based technique such as ANNs to classify uncertainty based on atmospheric states

or otherwise improve the data fusion process. Furthermore, an ensemble-based approach can be used for some systems to approximate uncertainties and represent non-linear uncertainty information.

Riverside recommends that a stepwise data fusion process be developed for the AOR that includes development of daily merged datasets (i.e., daily precipitation, maximum temperature, minimum temperature) using all available data sources prior to producing an hourly dataset. This recommendation is based on the differences in weather characteristics at hourly and daily time steps, the fact that more historical data are available on a daily time step, and the fact that accurate daily estimates are very important for the intended hydrologic applications.

Some issues merit further investigation through prototyping prior to finalizing the AOR development process. For example, given the many available options, the most appropriate sequence of data fusion should be tested using prototyping. Existing data fusion frameworks should be evaluated prior to finalizing the data fusion sequence.

13.3.2 Gridding Methods

Gridding methods that interpolate point observations to produce gridded estimates can be categorized into three groups:

- Methods that ignore physiographic characteristics and atmospheric states.
- Methods that consider physiographic influences.
- Methods that consider both physiographic characteristics and atmospheric states.

Most of the approaches in the first group can be modified to account for physiographic characteristics through a climatology-aided interpolation approach, where each station observation is converted to an anomaly from the climatological normal, the anomalies are interpolated, and the interpolated anomalies are multiplied by the background climatology grid. Various approaches incorporate atmospheric classifications to modify the interpolation parameters and improve the interpolation results.

Optimal interpolation and kriging incorporate information regarding covariances into the gridding techniques. These techniques, applied in a climatology-aided interpolation framework, outperform other spatial interpolation techniques. The hierarchical non-linear regression technique of Hsu et al. (2011) also seems promising if time and resources permit the development of the technique and associated regression relationships.

13.3.3 Spatial Downscaling Methods

The methods available for spatial downscaling can be categorized into interpolation methods, statistical downscaling methods, and dynamical downscaling methods.

The interpolation downscaling methods are readily available and easily implemented, with several operational applications already taking advantage of these techniques. For the AOR, it may be acceptable to utilize interpolation with terrain adjustments to downscale temperature fields, but interpolation methods are not recommended for downscaling precipitation fields. The spatial disaggregation methods are also reasonably well-developed and straightforward to implement.

Spatial downscaling techniques that employ regression methods are desirable for incorporating dynamical atmospheric processes that control weather patterns. However, additional development is required to identify controlling variables and to develop suitable relationships for different geographic regions. If time and resources permit further investigation for the AOR, the GMOS method would be a reasonable starting point to develop these relationships. The predictor variables can be used to develop methods both for transforming point observations to gridded estimates as well as for spatial downscaling of gridded datasets.

Similarly, dynamical downscaling methods that employ outputs from high-resolution regional weather models are desirable for incorporating dynamical atmospheric influences. If a high-resolution weather model can be run to generate a reanalysis dataset for the AOR, it would be extremely beneficial in providing a consistent, high-resolution dataset over the full historical period and spatial extent. Additional investigation is required to determine if it is feasible to complete a reanalysis for the AOR, and to evaluate the relative levels of effort and potential benefits between regression-based techniques and dynamical downscaling techniques.

13.3.4 Temporal Disaggregation Methods

Temporal disaggregation may be completed using an observation-based approach or a stochastic approach.

Observation-based approaches may use downscaled gridded products (i.e., radar, satellite, or NWP model outputs), hourly gage data, or a combination of sources developed through a data fusion process.

Major limitations were identified for most of the stochastic approaches. A form of K-nearest neighbor resampling may be feasible, where the hourly distribution applied for a grid cell is selected from a historical database of hourly distributions based on the similarity with the day being disaggregated. A distance function would be defined using characteristics of the daily precipitation field, and possibly atmospheric information from NWP models (e.g., moisture content, wind direction). The benefit of a stochastic approach is that it would more realistically represent the statistical characteristics of the true hourly precipitation distribution when insufficient data sources are available on an hourly time step. The limitations of stochastic approaches are that a unique solution is not obtained from the analysis procedure (particularly relevant if the AOR is reprocessed in the future), and the distributions would not reflect hourly observations where available.

Riverside recommends that a combined approach be taken to develop hourly disaggregation factors. In locations and time periods with a dense hourly gage network or radar observations, the hourly distribution will be known with higher confidence. In locations isolated from hourly gage or radar observations, the uncertainties associated with the observation-based hourly distributions will be very high. In these cases, a stochastic approach may be more appropriate. A combined approach would utilize stochastic estimates of the hourly distributions when uncertainties are high, otherwise observation-based estimates would be used. It may be most appropriate to select a particular disaggregation methodology for each region. Prototype applications could help identify the feasibility and limitations of the selected disaggregation approaches.

13.4 Evaluation Methods

Multiple approaches are available to evaluate the AOR for the purposes of evaluating competing methodologies, parameterizing algorithms, and assessing the final datasets. The evaluation methods include expert review, comparisons with existing datasets, and evaluations based on hydrologic models. Each method has benefits and limitations:

- By collecting expert input at different points of the development process, strengths, limitations, and pitfalls of different approaches will be identified, helping to guide the development process towards an appropriate balance of effort and benefit.
- Cross validation against point observations can be beneficial if the results are interpreted carefully.
- Multiple watershed studies have incorporated a dense network of gages and high-quality radars, allowing for enhanced estimation of precipitation and temperature on a fine resolution for a limited spatial extent and time period. These datasets provide valuable information for comparison and validation purposes.
- Many gridded datasets exist, for example climatological datasets developed using the PRISM system. These datasets can be used for assessing bias characteristics, temporal consistencies, and uncertainties in the AOR.
- Hydrologic simulations provide important validation information since observed streamflows are independent and indirect measures of the impact of precipitation and temperature on runoff in the contributing watershed. Since the RFCs are intended as the primary beneficiaries of the AOR development effort, it would be appropriate to utilize the calibrated models from the RFCs for the evaluation process. In addition to these lumped hydrologic models, the NWS OHD may wish to include basins modeled with a distributed model to evaluate the AOR using a model that requires gridded inputs.

13.5 Existing Weather Data Management Systems

Significant resources have been devoted by personnel at federal and academic institutions to develop systems that collect, quality control, and process weather data. The development approach for the AOR should maximize the use of available frameworks, automated processes, methodologies, and datasets to reduce the development cost and effort. The following systems and features were identified as having potential benefit for the AOR development:

- The NCDC PrecipVal and TempVal systems may provide useful automated data quality control procedures for both temperature and precipitation.
- The PRISM system developed at Oregon State University includes automated data quality control procedures and gridding techniques that may be beneficial for the AOR. In addition, the gridded climatological datasets that have been developed using the PRISM system may be useful for gridding point observations using climatology-aided interpolation, or spatially disaggregating coarse-resolution datasets.
- The SNODAS developed by the NOHRSC includes automated data quality control procedures, spatial downscaling methods, and data assimilation techniques that may be useful for the AOR development.

- The NMQ/Q2 system developed by the NSSL includes automated data quality control procedures (particularly convective screening); radar processing, mosaicking, and bias correction algorithms; and a framework for data fusion that may be beneficial for the AOR development. Additionally, the NCDC and the NSSL plan to generate a 10-yr reanalysis dataset using the NMQ/Q2 system, which would improve the potential benefit for the AOR development by providing a longer historical record.
- The DailyQC program, which has been integrated into the MPE and the Mountain Mapper programs, contains automated data quality control procedures and gridding techniques that may be useful for the AOR development.
- The Mountain Mapper system developed by the CBRFC uses climatology-aided interpolation to transform point observations into gridded estimates. This gridding technique may be beneficial for the AOR development. In addition, the Mountain Mapper system is in widespread use at the western RFCs, and likely will represent a benchmarking dataset against which the AOR will be compared.
- The MPE program run by the RFCs includes automated data quality control, bias correction, and multi-sensor features that that may be beneficial for the AOR.
- Researchers at the CPC are developing methods that may have relevance for the AOR in the following four areas: improving satellite precipitation estimation; developing historical and real-time grids on a 1-hr time step for the CONUS using only gage observations; merging satellite estimates and NWP model outputs; and merging satellite estimates, NWP model outputs, gage observations, and radar precipitation into a single multi-sensor estimate.
- The NLDAS developed by NASA and NCEP has automated quality control, gridding, and downscaling procedures that may be relevant for the AOR development.

13.6 Technical Considerations

The AOR is intended to be developed for the period 1979-2010, with eventual extension back to 1948. Within this historical period, there are several notable periods that will be affected by the amount and quality of the available historical data (i.e., 1948-1977, 1978-1995, and 1996-2010). Consistency among these periods is desired and will be assessed in terms of long-term climatology, daily spatial variability, and hourly distributions.

The AOR will be utilized to evaluate hydrologic models developed at the RFCs, potentially requiring calibration adjustments to improve streamflow simulations. The consistency of the AOR over the period of record affects its utility for calibration purposes. It is possible that a shorter period within the AOR will be recommended for model calibration, whereas the full period will be suitable for climatological forcing of ensemble forecasts. The quality of the hydrologic simulations will provide valuable information on the success of the AOR development process that could subsequently inform ongoing improvements.

The AOR is intended to be developed for the CONUS, Alaska, Hawaii, Puerto Rico, and the portions of Canada and Mexico that drain to the U.S. Given the regional differences in climate factors, hydrology, forecast objectives, data sources, and data quality, the AOR development process will have to address regional requirements and issues. The NWS OHD must evaluate whether one development approach can address these issues, or whether regional development approaches are required.

13.7 RFC Considerations

The following notable points were identified through discussions with selected RFC personnel:

- The AOR datasets must be consistent with the real-time forcings and forecast grids provided to the RFCs. Consistency includes climatological characteristics, spatial resolution, and temporal resolution. For example, the forecast grids from the NDFD are currently provided on a 2.5-5.0 km grid mesh; the forecast grids should be provided on the same grid mesh intended for the AOR.
- The RFCs expressed interest in having access to multiple versions of the AOR: a "best" dataset that uses all available data, as well as a "consistent" version that uses limited data.
- The RFCs expressed interest in having access to the AOR datasets at multiple temporal intervals that range from one hour to one day. Some locations would benefit from an even finer temporal interval, such as 15 minutes, but sub-hourly resolution is not practical for the initial AOR development. Regional differences and intended usage affect the appropriate time interval.
- The RFCs would benefit from having historical gridded datasets that expand beyond temperature and precipitation to include freezing level, wind speed, relative humidity, etc.
- The RFCs exhibit regional preferences in using gage and radar data. The framework selected to develop the AOR should avoid favoring one data type. This requirement is consistent with an uncertainty-based data fusion framework.
- To strengthen the likelihood that the AOR will be adopted within the RFCs (superceding their own forcing datasets), it will be essential to survey the RFCs regarding quality control procedures and data networks that are currently incorporated into their forcing analyses, and to explore ways that these can be incorporated into the AOR development (CBRFC, personal communication, March 7, 2012).
- The RFCs anticipate that the AOR will evolve as data sources and quality control techniques improve. The AOR should be regenerated as new information becomes available and should be extended over time to keep the datasets current.
- The RFCs would benefit from tools that enhance the flexibility of the datasets to meet RFC needs, for example a tool that would allow the RFCs to adjust the AOR to a desired climatology.
- Overall, the RFC personnel are enthusiastic about the potential benefit of the AOR datasets. The RFC personnel who were interviewed expressed a willingness to work with the OHD to provide regional expertise, to provide available data, to evaluate the AOR, and to evaluate the need to re-calibrate the existing models.

13.8 Logistical Considerations

NOAA has established best practices for data archives and delivery and format requirements for climate data records (Nelson and Semunegus, 2011). These guidelines for documentation, data formats, and metadata should be followed for the AOR.

For the RFCs, the AOR will be delivered using existing NWS data transmission processes that have been established for CHPS. However, the data also should be made available

to external users, for example on a public website. The website should be developed to include subsetting tools based on date and geographic extent so that users can easily access data for their areas and periods of interest, rather than downloading the entire AOR, which is likely to be a large dataset. The NWS OHD should also consider providing access to the input datasets (raw and quality-controlled) along with quality control information.

Moving forward in time, the NWS OHD intends to append to the AOR using the real-time process established at the National Water Center. Riverside recommends that data grids developed using the real-time process be assigned a preliminary status. After a suitable delay, the grids can be updated using all available historical data sources and updated to reflect published values. For example, the Mesoscale Analysis Committee (MAC) convened in 2004 envisioned that real-time forcings could be available within 30 minutes, with subsequent updates issued with one day and one month delays to augment the available data sources (Colman et al., 2005). In having preliminary datasets (developed using the real-time process) and published datasets (developed using the AOR process) the NWS OHD can evaluate the consistency between the methods used and the resulting datasets, and can continue to refine both processes as needed. This process would be similar to plans to release a preliminary version of the IMERG satellite precipitation dataset, followed by a gage-adjusted dataset with a latency of approximately two months (Huffman et al., 2012).

Additionally, the AOR may need to be periodically regenerated as significant advances are made to data inputs (e.g., reanalysis or climatological datasets) or methods (e.g., automated quality control). However, the potential improvements in the AOR should be weighed against the efforts required at the RFCs to adjust the hydrologic models.

13.9 Next Steps

13.9.1 Parallel Development

The AOR development process will involve a review of efforts underway at various agencies, the collection and quality control of large datasets, and coordination among many organizations. Although some aspects of the development effort cannot be started without selecting a framework, numerous preparatory activities can be initiated:

- Identify primary data sources that will be utilized and begin collecting data from those sources. Given the wide range of data sources, particularly for station data, the data collection may require a large effort.
- Begin performing quality control of station data.
- Continue efforts to quality control and adjust radar data. The NWS OHD has initiated this activity, recognizing its importance for the AOR.
- Identify high-resolution datasets that will be utilized for evaluation purposes and begin assembling these data.
- Identify primary components of existing systems (e.g., software) that may be integrated into the processing system.
- Identify a starting development platform and data fusion framework. Multiple existing systems could provide this framework.

13.9.2 High-Resolution NWP Model Reanalysis

Currently, there are multiple high-resolution (i.e., 4-5 km grid mesh) NWP models that are run in real-time for forecasting purposes (e.g., WRF-NMM, WRF-ARW, and NAM-Nest). However, besides limited applications of high-resolution NWP models over limited domains and time periods, there are no equivalent high-resolution reanalysis datasets available for input into the AOR development process. The reanalysis products available for the period of record intended for the AOR are much coarser (i.e., 32 and 38-km grid meshes for the NARR and CFSR, respectively), and they employ different algorithms than what is available in real-time. These differences constitute one of the more significant inconsistencies between the inputs for the AOR and the real-time forcings process. If the AOR is intended to be used to evaluate the impact of the real-time forcings on hydrologic simulations, a high-resolution reanalysis would be necessary to more completely achieve this requirement.

A high-resolution reanalysis would provide valuable input into the AOR development for gridding and spatial downscaling. A high-resolution reanalysis would provide more realistic fields of precipitation and temperature than could be achieved through statistical downscaling of the current reanalysis products. In preliminary discussions with the ESRL, it appears that such an analysis would be feasible given current computational capabilities (ESRL, personal communication, November 16, 2011). If such a reanalysis could be completed for the full period of record, it would constitute the only consistent, high-resolution gridded input in the development process.

A high-resolution reanalysis would provide the greatest benefit in areas with complex terrain and limited radar coverage. If a complete reanalysis of the entire spatial domain is not feasible, it would still be advantageous to perform the reanalysis for a limited domain. Similarly, if the reanalysis could not be completed for the full period, it would still be advantageous to perform the reanalysis for a limited period. A reanalysis over the later period (e.g., 1996-2010) would improve consistency with the real-time forcings process and would provide a better sample for stochastic approaches employed for temporal disaggregation in the early period. A reanalysis over the earlier period (e.g., 1980-1996) would improve the available information and alter the methodology to prioritize observation-based disaggregation weights computed using the reanalysis data over stochastic methods.

Given the widespread use of the NARR reanalysis dataset in research applications, an updated, high-resolution reanalysis over North America would be valuable to researchers in many disciplines. However, the costs associated with such an effort may be prohibitive.

13.9.3 Prototype Evaluations

Riverside recommends that the NWS OHD conduct prototype evaluations to evaluate and test available data sources and methods for the AOR. For example, Riverside recommends developing prototypes of the AOR for several regions (e.g., the southeast CONUS and the intermountain west) to allow development, testing, refinement, and automation of a proposed method. The prototype should include an evaluation phase where the prototype datasets are input to existing hydrologic models. Additionally, the quality-controlled individual data sources could be input to the hydrologic models as basin averages to assess their utility for the AOR development process.

14 Acknowledgements

Riverside would like to thank the following individuals for providing guidance, information, and valuable insights that were incorporated into this report. In acknowledging these contributions, Riverside does not imply endorsement of the work performed or the conclusions made herein.

- Dave Kitzmiller – Office of Hydrologic Development
- Carroll Campbell – Earth System Research Laboratory
- Rob Cifelli – Earth System Research Laboratory
- Greg Fall – National Operational Hydrologic Remote Sensing Center
- Rob Hartman – California Nevada River Forecast Center
- Chengmin Hsu – Earth System Research Laboratory
- Mimi Hughes – Earth System Research Laboratory
- Joe Intermill – Northwest River Forecast Center
- Lynn Johnson – Earth System Research Laboratory
- Chris Kummerow – Colorado State University
- Sanghun Lim – Earth System Research Laboratory
- Harold Opitz – Northwest River Forecast Center
- Craig Peterson – Colorado Basin River Forecast Center
- Robin Radlein – Alaska-Pacific River Forecast Center
- Andrea Ray – Earth System Research Laboratory
- John Schaake – Independent Consultant
- John Schmidt – Southeast River Forecast Center
- Tim Schneider – Office of Hydrologic Development
- Wei Shi – Climate Prediction Center
- Greg Story – West Gulf River Forecast Center
- Dave Streubel – Alaska-Pacific River Forecast Center
- Andy Wood – Colorado Basin River Forecast Center
- Wanru Wu – Office of Hydrologic Development
- Pingping Xie – Climate Prediction Center
- Yu Zhang – Office of Hydrologic Development
- Ziya Zhang – Office of Hydrologic Development

15 Acronyms

ABRFC - Arkansas Basin River Forecast Center
AgriMet - Agricultural Weather Network
AGROSON – Red Estatal de Estaciones Meteorologicas de Sonoro
AK - Alaska
ALERT – Automated Local Evaluation in Real Time
ANN - Artificial Neural Network
AOR – Analysis of Record
APRFC - Alaska-Pacific River Forecast Center

ASOADeK - Auto-Searched Orographic and Atmospheric Effects Detrended Kriging
ASOS - Automated Surface Observing System
ASR - Arctic System Reanalysis
AVHRR - Advanced Very High Resolution Radiometer
AWIPS - Advanced Weather Interactive Processing System
AWOS - Automated Weather Observing System
BCSD - Bias Correction and Spatial Disaggregation
CAI - Climatology-Aided Interpolation
CBRFC - Colorado Basin River Forecast Center
CCPA - Climatology-Calibrated Precipitation Analysis
CDEC - California Data Exchange Center
CFSR - Climate Forecast System Reanalysis
CHPS – Community Hydrologic Prediction System
CIRES - Cooperative Institute for Research in Environmental Sciences
CLICOM – Climate Computing Project
CMAP - Climate Prediction Center Merged Analysis of Precipitation
CMORPH - Climate Prediction Center Morphing Method
CN – Canada
CoCoRaHS – Community Collaborative Rain, Hail, and Snow
CONAGUA - Comisión Nacional del Agua
CONUS – Continental United States
COOP – Cooperative Network
CPC – Climate Prediction Center
CRN - Climate Reference Network
CSU – Colorado State University
CWOP - Citizen Weather Observer Program
DCP - Data Collection Platform
DEM - Digital Elevation Model
DMIP - Distributed Model Inter-comparison Project
DOD – Department of Defense
DOE – Department of Energy
DPA - Digital Precipitation Array
ECMWF - European Centre for Medium-Range Weather Forecasts
EMC - Environmental Modeling Center
ESP – Ensemble Streamflow Prediction
ESRL - Earth System Research Laboratory
FAA – Federal Aviation Administration
GCIP - Continental Scale International Project
GFS – Global Forecast System
GMOS - Gridded Model Output Statistics
GOES - Geostationary Operational Environmental Satellites
GPCP - Global Precipitation Climatology Project
GPCP 1DD - Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations
GPM - Global Precipitation Mission
GrADS - Grid Analysis and Display System

GRASS - Geographic Resources Analysis Support System
GSMaP_MVK+ - Global Satellite Mapping using a Moving Kalman Filter
GTS - Global Telecommunications System
HADS - Hydrometeorological Automated Data System
HCN - Historical Climatology Network
HEFS - Hydrologic Ensemble Forecast System
HI - Hawaii
HMT – Hydrometeorology Testbed
HPC - Hydrometeorological Prediction Center
HRAP - Hydrologic Rainfall Analysis Project
ICP - Interactive Calibration Program
IDW - Inverse Distance Weighting
IFLOWS – Integrated Flood Observing and Warning System
IMERG - Integrated Multisatellite Retrievals for the Global Precipitation Mission
IR - Infrared
IWRSS – Integrated Water Resources Science and Solutions
JMA - Japanese Meteorological Agency
KF-CMORPH - Kalman Filter Climate Prediction Center Morphing Method
K-nn - K-Nearest Neighbor
LMODEL - Lagrangian Model
MAC - Mesoscale Analysis Committee
MADIS - Meteorological Assimilation Data Ingest System
MAP – Mean Areal Precipitation
MAT – Mean Areal Temperature
MCP3 - Manual Calibration Program
MERRA - Modern Era Retrospective-Analysis
MIT - Massachusetts Institute of Technology
MM5 - Mesoscale Model
MODIS - Moderate Resolution Imaging Spectroradiometer
MOS - Model Output Statistics
MPE - Multi-sensor Precipitation Estimation
MRMS - Multi-radar and Multi-sensor
MX - Mexico
NAME – North American Monsoon Experiment
NARR - North American Regional Reanalysis
NASA - National Aeronautics and Space Administration
NASA-GMAO - NASA Goddard Global Modeling and Assimilation Office
NCAR - National Center for Atmospheric Research
NCDC – National Climatic Data Center
NCEP - National Center for Environmental Prediction
NCL - National Center for Atmospheric Research Command Language
NDFD - National Digital Forecast Database
NDGD - National Digital Guidance Database
NESDIS - National Environmental Satellite, Data, and Information Service
NEXRAD - Next-Generation Radar
NLDAS - North-American Land Data Assimilation Systems

NMQ - National Radar Mosaic
NMQ/Q2 - National Radar Mosaic/Next Generation Multi-Sensor Quantitative Precipitation Estimation
NOAA – National Oceanic and Atmospheric Administration
NOHRSC - National Operational Hydrologic Remote Sensing Center
NRCS – Natural Resources Conservation Service
NSSL – National Severe Storms Laboratory
NWP – Numerical Weather Prediction
NWRFC - Northwest River Forecast Center
NWS – National Weather Service
NWSRFS – National Weather Service River Forecast System
OHD – Office of Hydrologic Development
OPT3 – Calibration System Automatic Parameter Optimization Program
PCA – Principal Components Analysis
PEHRPP - Program to Evaluate High Resolution Precipitation Products
PERSIANN-CSS - Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System
PMW - Passive Microwave
PR – Puerto Rico
PRISM - Parameter-Elevation Regressions on Independent Slopes Model
PXADJ – Precipitation Adjustment Factor
QPE – Quantitative Precipitation Estimate
QPF – Quantitative Precipitation Forecasts
RAWS - Remote Automated Weather Stations
RDHM - Research Distributed Hydrologic Model
RDHM-TF - Research Distributed Hydrologic Model Threshold Frequency
REFAME - Rainfall Estimation using Forward-Adjusted advection of Microwave Estimates
RFC – River Forecast Center
RTMA - Real-Time Mesoscale Analysis
RUC - Rapid Update Cycle
SAC-SMA – Sacramento Soil Moisture Accounting Model
SCAN - Soil Climate Analysis Network
SERFC - Southeast River Forecast Center
SHEF - Standard Hydrologic Exchange Format
SIH - Sistema de Información Hydrométrica
SmartInits - Smart Initializations
SNODAS - Snow Data Assimilation System
SNOTEL – Snowpack Telemetry
SOE - Single Optimal Estimator
STEPS - Short-Term Ensemble Prediction System
SWE – Snow Water Equivalent
SYMAP - Synergraphic Mapping System
TDWR - Terminal Doppler Weather Radar
TMPA - Tropical Rainfall Measuring Mission Multisatellite Precipitation Analysis
TRMM - Tropical Rainfall Measuring Mission
URL – Uniform Resource Locator

US – United States
USACE – United States Army Corps of Engineers
USBR – United States Bureau of Reclamation
USGS – United States Geological Survey
VIC - Variable Infiltration Capacity
WBAN - Weather Bureau Army Navy
WFO – Weather Forecast Office
WGRFC – West Gulf River Forecast Center
WMO - World Meteorological Organization
WRF - Weather Research and Forecasting
WSR-88D - Weather Surveillance Radar-1988 Doppler

16 References

Adler, R. F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P-P., Janowiak, J., Rudolf, B., et al. (2003). "The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present)." *Journal of Hydrometeorology* 4 (December): 1147-1167.

Alaska Region Headquarters (ARH). 2011. Alaska Radar. Accessed January 2012 at: http://www.arh.noaa.gov/radar.php.

Anderson, E. (2002). *Calibration of Conceptual Models for Use in River Forecasting*. August. http://www.nws.noaa.gov/oh/hrl/calb/calibration1102/main.htm.

Angel, W. E., Urzen, M. L., Del Greco, S. A., and Bodosky, M. W. (2002). "Automated Validation for Summary of the Day Temperature Data." *American Meteorological Society, 19th Conference on IIPS, 83rd Annual Meeting*, Long Beach, CA. http://ams.confex.com/ams/annual2003/techprogram/paper_57274.htm.

Baker, Adam. (2011). "CoCoRaHS: At the Core of Climate Outreach." *Climate Services Workshop*; June 23, 2011. Accessed December 2011 at: http://iclimate.org/workshop-c/Baker_CoCoRaHS_climate_outreach_final.pdf

Barnes, S. L. (1964). "A technique for maximizing details in numerical weather map analysis." *Journal of Applied Meteorology* 3: 396-409.

Barrett, A. (2003). "National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS) Products at NSIDC." *NSIDC Special Report 11*, National Snow and Ice Data Center, Boulder, CO. Digital media.

Becker, E. J., Berbery, E. H., and Higgins, R. W. (2009). "Understanding the Characteristics of Daily Precipitation over the United States Using the North American Regional Reanalysis." *Journal of Climate* 22 (23) (December): 6268-6286. doi:10.1175/2009JCLI2838.1.

Behrangi, A., Imam, B., Hsu, K., Sorooshian, S., Bellerby, T. J., and Huffman, G. J. (2010). "REFAME: Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates." *Journal of Hydrometeorology* 11 (6) (December): 1305-1321.

Bellerby, T., Hsu, K-L., and Sorooshian, S. (2009). “LMODEL: A Satellite Precipitation Methodology Using Cloud Development Modeling. Part I: Algorithm Construction and Calibration.” *Journal of Hydrometeorology* 10 (5) (October): 1081-1095.

Bowler, N.E., C.E. Pierce, and A.W. Seed. (2006). “STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP.” Quarterly Journal of the Royal Meteorological Society 132, no. 620: 2127–2155.

Brandsma, T., and Können, G.P. (2006). “Application of nearest-neighbor resampling for homogenizing temperature records on a daily to sub-daily level.” *International Journal of Climatology* 26 (January): 75-89. doi:10.1002/joc.1236.

Buishand, T.A., and Brandsma, T. (2001). “Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling.” *Water Resources Research* 37 (11) (November): 2761-2776.

Bukovsky, M. S., and Karoly, D. J. (2007). “A Brief Evaluation of Precipitation from the North American Regional Reanalysis.” *Journal of Hydrometeorology* 8 (4) (August): 837-846. doi:10.1175/JHM595.1.

Ceccato, P., C. Vancutsem, and M. Temimi. (2010). “Monitoring Air and Land Surface Temperatures from Remotely Sensed Data for Climate-human Health Applications.” In *Geoscience and Remote Sensing Symposium (IGARSS)*, 2010 IEEE International, 178–180. IEEE. doi:10.1109/IGARSS.2010.5649810.

Chang, Edmund. (2012). “Using 20th Century Reanalysis Data to Examine Northern Hemisphere Storm Track Trend in the 20th Century” presented at the NOAA CPO MAPP Webinar, February 14.

Charba, J. P., and Samplatsky, F. G. (2011a). “Regionalization in Fine-Grid GFS MOS 6-h Quantitative Precipitation Forecasts.” *Monthly Weather Review* 139 (1) (January): 24-38. doi:10.1175/2010MWR2926.1.

Charba, J. P., and Samplatsky, F. G. (2011b). “High-Resolution GFS-Based MOS Quantitative Precipitation Forecasts on a 4-km Grid.” *Monthly Weather Review* 139 (1) (January): 39-68. doi:10.1175/2010MWR3224.1.

Chelliah, Muthuvel, and C. F. Ropelewski. (2000). “Reanalyses-Based Tropospheric Temperature Estimates: Uncertainties in the Context of Global Climate Change Detection.” *Journal of Climate* 13 (17) (September): 3187–3205. doi:10.1175/1520-0442(2000)013<3187:RBTTEU>2.0.CO;2.

Chen, M., Xie, P., Janowiak, J.E., and Arkin, P.A. (2002). “Global land precipitation: A 50-yr monthly analysis based on gauge observations.” *Journal of Hydrometeorology* 3 (3): 249–266.

Chiang, K. Hsu, F. Chang, Y. Hong, and S. Sorooshian. (2007). “Merging multiple precipitation sources for flash flood forecasting.” *Journal of Hydrology* 340 (3-4): 183-196. doi:10.1016/j.jhydrol.2007.04.007.

Clark, M. P., and Slater, A. G. (2006). “Probabilistic Quantitative Precipitation Estimation in Complex Terrain.” *Journal of Hydrometeorology* 7 (February): 3-22.

Climate Prediction Center (CPC). (2012a). *Gridded U.S. Daily Temperature (Maxima and Mini[m]a).* Accessed March 2012 at: ftp://ftp.cpc.ncep.noaa.gov/precip/daily_grids/DOCUMENTATION.

Climate Prediction Center (CPC). (2012b). *Metadata for the Daily Gridded Temperature Analyses (C).* Accessed March 2012 at: http://www.cpc.ncep.noaa.gov/products/GIS/GIS_DATA/us_temp/metadata/.

Colman, Brad, John Horel, and Lee Anderson. (2005). "Progress Toward a Mesoscale Analysis of Record" August 8. http://www.nws.noaa.gov/ost/ifps_sst/presentations/AOR_Update_Presentation_080805_FINAL.ppt.

Colorado Basin River Forecast Center (CBRFC). (2009). *daily_qc(2.5) Users Guide*. July 20.

Comision Nacional del Agua (CONAGUA). (2011). *Radares*. Accessed January 2012 at: http://smn.cna.gob.mx/index.php?option=com_content&view=article&id=14&Itemid=84.

Compo, Gilbert P., Jeffrey S. Whitaker, Prashant D. Sardeshmukh, and Benjamin Giese. (2012). "Evaluating the 20th Century Reanalysis" presented at the NOAA CPO MAPP Webinar, February 14.

Compo, G. P, Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., et al. (2011). "The twentieth century reanalysis project." *Quarterly Journal of the Royal Meteorological Society* 137 (654): 1–28.

Cosgrove, Brian A., Dag Lohmann, Kenneth E. Mitchell, Paul R. Houser, Eric F. Wood, John C. Schaake, Alan Robock, et al. (2003). "Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project." *Journal of Geophysical Research* 108 (October 31): 12 PP. doi:200310.1029/2002JD003118.

Courault, D., and P. Monestiez. (1999). "Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France." *International Journal of Climatology* 19, no. 4: 365–378.

Creager, G. (2006). "Densifying the in-situ surface observing network: The Citizen Weather Observer Program, its role in observation and verification, and avenues for enhancement." *22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, Atlanta, GA.

Cressie, N. (1993). *Statistics for Spatial Data*. New York: John Wiley and Sons.

Cressman, G. P. (1959). "An operational ojective analysis system." *Monthly Weather Review* 87: 367-374.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). "A knowledge-based approach to the statistical mapping of climate." *Climate Research* 22 (2): 99–113.

Daly, C., Redmond, K., Gibson, W., Doggett, M., Smith, J., Taylor, G., Pasteris, P., and Johnson, G. (2005). "Opportunities for improvements in the quality control of climate observations." *15th AMS Conf. on Applied Climatology, American Meteorological Society*, Savannah, GA.

Daly, C. (2006). "Guidelines for assessing the suitability of spatial climate data sets." *International Journal of Climatology* 26 (6) (May): 707-721. doi:10.1002/joc.1322.

Daly, C., and Taylor, G. H. (2006). Smart PRISM draft final report.

Daly, C., Smith, J. W., Smith, J. I., and McKane, R. B. (2007). "High-Resolution Spatial Modeling of Daily Weather Elements for a Catchment in the Oregon Cascade Mountains, United States." *Journal of Applied Meteorology and Climatology* 46 (10) (October): 1565-1586. doi:10.1175/JAM2548.1.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P. (2008). "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology* 28 (15) (December): 2031-2064. doi:10.1002/joc.1688.

Daymet. (2012). *Daymet Daily Surface Weather and Climatological Summaries*. Accessed February 2012 at: http://www.daymet.org/.

Debele, B., R. Srinivasan, and J. Yves Parlange. (2007). "Accuracy evaluation of weather data generation and disaggregation methods at finer timescales." *Advances in Water Resources* 30, no. 5: 1286–1300.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., et al. (2011). "The ERA-Interim reanalysis: configuration and performance of the data assimilation system." *Quarterly Journal of the Royal Meteorological Society* 137 (April): 553-597. doi:10.1002/qj.828.

Di Luzio, Mauro, Gregory L. Johnson, Christopher Daly, Jon K. Eischeid, and Jeffrey G. Arnold. (2008). "Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States." *Journal of Applied Meteorology and Climatology* 47 (2) (February): 475-497. doi:10.1175/2007JAMC1356.1.

Ebert, E. E., Janowiak, J. E., and Kidd, C. (2007). "Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models." *Bulletin of the American Meteorological Society* 88 (January): 47-64.

Ebtehaj, Ardeshir Mohammad, and Efi Foufoula-Georgiou. (2011). "Adaptive fusion of multisensor precipitation using Gaussian-scale mixtures in the wavelet domain." *Journal of Geophysical Research* 116, no. 22 (November 19). http://www.agu.org/pubs/crossref/2011/2011JD016219.shtml.

Environment Canada. (2011). Canadian Weather Radar. Accessed January 2012 at: http://www.weatheroffice.gc.ca/radar/index_e.html.

Esteban, Jaime, Andrew Starr, Robert Willetts, Paul Hannah, and Peter Bryanston-Cross. (2005). "A Review of data fusion models and architectures: towards engineering guidelines." *Neural Computing and Applications* 14, no. 4 (June 21): 273-281.

Ferree, J. (2005). "TDWR Data Into AWIPS: OB5 Training." *NOAA Warning Decision Training Branch*. Accessed December 2011 at: http://www.wdtb.noaa.gov/buildTraining/TDWR/TDWR.pdf.

Fowler, H. J, Blenkinsop, S., and Tebaldi, C. (2007). "Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling." *International Journal of Climatology* 27 (12) (October 1): 1547-1578. doi:10.1002/joc.1556.

Fulton, R. (2005a). “Quantitative Precipitation Estimation in the National Weather Service.” *MPE Workshop*. Silver Spring, Maryland.

Fulton, R. (2005b). “Multisensor Precipitation Estimator (MPE) Workshop.” *MPE Workshop*. Silver Spring, Maryland.

Gandin. (1965). Objective analysis of meteorological fields. Jerusalem: Israel Program for Scientific Translations.

Gangopadhyay, S., Clark, M. P., and Rajagopalan, B. (2005). “Statistical downscaling using K-nearest neighbors.” *Water Resources Research* 41 (2). doi:10.1029/2004WR003444.

Gibson, J.K., Kallberg, P., Uppala, S., Hernandez, A., Nomura, A., and Serrano, E. (1997). “ERA Description.” *ERA-15 Project Report Series No. 1*, 89 pp. July.

Gibson, W., Daly, C., Kittel, T., Nychka, D., Johns, C., Rosenbloom, N., McNab, A., and Taylor, G. (2002). “Development of a 103-year high-resolution climate data set for the conterminous United States.” *Proceedings of the 13th AMS Conference on Applied Climatology, American Meteorological Society*, 181-183. Portland, OR.

Gilbert, K.K., Glahn, B., Cosgrove, R. L., Sheets, K. L., and Wagner, G. A. (2009). “Gridded Model Output Statistics: Improving and Expanding.” *23rd Conference Weather Analysis and Forecasting/19th Conference Numerical Weather Prediction, American Meteorology Society*, Omaha, NE, Accessed January 2012 at: http://www.nws.noaa.gov/mdl/synop/amspapers/Gilbert_et_al_2009.pdf.

Glahn, B. (2006). “The Role of Model Output Statistics (MOS) in Downscaling NWP Model Output.” *AMS Short Course Methods and Problems of Downscaling Weather and Climate Variables*, Atlanta, GA.

Glahn, B., Gilbert, K. K., Cosgrove, R. L., Ruth, D. P., and Sheets, K. L. (2009). “The Gridding of MOS.” *Weather and Forecasting* 24 (2): 520-529. doi:10.1175/2008WAF2007080.1.

Goovaerts, P. (2000). “Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall.” *J Hydrology* 228, 113-129.

Gorenburg, I.P., D. McLaughlin, and D. Entekhabi. (2001). “Scale-recursive assimilation of precipitation data.” *Advances in Water Resources* 24, no. 9-10: 941-953.

Guan, H., Wilson, J. L., and Makhnin, O. (2005). “Geostatistical Mapping of Mountain Precipitation Incorporating Autosearched Effects of Terrain and Climatic Characteristics.” *Journal of Hydrometeorology* 6 (December): 1018-1031. doi:10.1175/JHM448.1.

Gupta, Rohit, V. Venugopal, and Efi Foufoula-Georgiou. (2006) . “A methodology for merging multisensor precipitation estimates based on expectation-maximization and scale-recursive estimation.” *Journal of Geophysical Research* 111, no. 2. http://www.agu.org/pubs/crossref/2006/2004JD005568.shtml.

Gutmann, E. D., Rasmussen, R. M., Liu, C., Ikeda, K., Gochis, D. J., Clark, M. P., Dudhia, J., and Thompson, G. (2011). “A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation Over Complex Terrain.” *Journal of Climate* (July 19): 110719122723006. doi:10.1175/2011JCLI4109.1.

Hamlet, A. F., and Lettenmaier, D. P. (2005). “Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States.” *Journal of Hydrometeorology* 6 (3) (June): 330-336. doi:10.1175/JHM420.1.

Hashino, T., Bradley, A. A., and Schwartz, S. S. (2006). “Evaluation of bias-correction methods for ensemble streamflow volume forecasts.” *Hydrology and Earth System Sciences Discussions* 3 (April): 561-594.

Hewitson, B. C, and R. G Crane. (2005). “Gridded area-averaged daily precipitation via conditional interpolation.” *Journal of Climate* 18, no. 1: 41–57.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis. (2005). “Very high resolution interpolated climate surfaces for global land areas.” *International Journal of Climatology* 25: 1965-1978.

Hofstra, N., Haylock, M., New, M., Jones, P., and Frei, C. (2008). “Comparison of six methods for the interpolation of daily, European climate data.” *Journal of Geophysical Research* 113. doi:10.1029/2008JD010100.

Hofstra, N., and New, M. (2009). “Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe.” *International Journal of Climatology* 29 (12) (October): 1872-1880. doi:10.1002/joc.1819.

Horel, John, and Brad Colman. (2005). “Real-Time and Retrospective Mesoscale Objective Analyses.” *Bulletin of the American Meteorological Society* 86 (10) (October): 1477–1480. doi:10.1175/BAMS-86-10-1477.

Hou, Dingchen, Mike Charles, Yan Luo, Zoltan Toth, Yuejian Zhu, Roman Krzysztofowicz, Ying Lin, Pingping Xie, Dong-Jun Seo, Malaquias Pena, and Bo Cui. (2010). *Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of Stage IV towards CPC Gauge-Based Analysis*. December 16. Submitted for publication. Accessed February 2012 at: http://www.emc.ncep.noaa.gov/gmb/yzhu/gif/pub/Manuscript_CCPA_20101216-1.pdf.

Hsu, K., Behrangi, A., Imam, B., and Sorooshian, S. (2010). “Extreme Precipitation Estimation Using Satellite-Based PERSIANN-CCS Algorithm.” *Satellite Rainfall Applications for Surface Hydrology*: 3–22.

Hsu, C., Cifelli, R., Zamora, R. J., and Johnson, L. E. (2011). “Toward the Estimation of High-Resolution Daily Precipitation in Complex Regions – The Study of Intertwined Physiographic, Vegetative, and Climatologic Factors for PRISM Enhancement.” *American Geophysical Union Fall Meeting*, San Francisco, CA.

Huffman, G.J., Adler, R.F., Morrissey, M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., Susskind, J. (2001). “Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations.” *Journal of Hydrometeorology*, 2, 36-50.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). “The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales.” *Journal of Hydrometeorology* 8 (1) (February): 38-55. doi:10.1175/JHM560.1.

Huffman, G. J., Bolvin, D. T., Braithwaite, K. L., Hsu, K-L., Joyce, R. J., Kidd, C., Sorooshian, S., Xie, P., and Yoo, S. H. (2012). "Design and Early Tests of the Day-1 GPM Combined Precipitation Algorithm: IMERG." *92nd American Meteorological Society Meeting*. New Orleans, LA: January 22. http://ams.confex.com/ams/92Annual/webprogram/Paper197814.html.

Hughes, M., Hall, A., and Fovell, R. G. (2009). "Blocking in Areas of Complex Topography, and Its Influence on Rainfall Distribution." *Journal of the Atmospheric Sciences* 66 (2) (February): 508-518. doi:10.1175/2008JAS2689.1.

Hunter, S. M. (1996). "WSR-88D Radar Rainfall Estimation: Capabilities, Limitations and Potential Improvements." *National Weather Service Weather Forecast Office Morristown, TN*. Accessed December 2011 at: http://www.srh.noaa.gov/mrx/research/precip/precip.php.

Hutchinson, M.F. (1993). "On thin plate splines and kriging." In Computing and Science in Statistics, 55-62. Edited by M. E. Tarter and M. D. Lock. University of California, Berkeley, CA: Interface Foundation of North America.

Ikeda, K., Rasmussen, R., Liu, C., Gochis, D., Yates, D., Chen, F., Tewari, M., et al. (2010). "Simulation of seasonal snowfall over Colorado." *Atmospheric Research* 97 (4) (September): 462-477. doi:10.1016/j.atmosres.2010.04.010.

Iredell, M. (2000). CopyGB utility documentation. *NCEP*. January 19.

Jin, Menglin, and Robert E. Dickinson. (1999). "Interpolation of Surface Radiative Temperature Measured from Polar Orbiting Satellites to a Diurnal Cycle 1. Without Clouds." *Journal of Geophysical Research* 104 (D2): PP. 2105–2116. doi:199910.1029/1998JD200005.

Johnson, L. (2008). *Survey of NOAA Downscaling Approaches and Applications*. November 14.

Johnson, L., and Schneider, T. (2009). "Survey of NOAA Downscaling Approaches and Applications." *AWRA Spring Specialty Conference: Managing Water Resources Development in a Changing Climate*, Anchorage, AK.

Joyce, R. J, Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution." *Journal of Hydrometeorology* 5 (3) (June): 487–503.

Joyce, R. J., and Xie, P. (2011). "Kalman Filter Based CMORPH." *Journal of Hydrometeorology* (June): 110601110645038. doi:10.1175/JHM-D-11-022.1.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., et al. (1996). "The NCEP/NCAR 40-Year Reanalysis Project." *Bulletin of the American Meteorological Society* 77 (March): 437-471.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S-K., Hnilo, J. J., Fiorino, M., and Potter, G. L. (2002). "NCEP–DOE AMIP-II Reanalysis (R-2)." *Bulletin of the American Meteorological Society* 83 (November): 1631-1643.

Katz, S. (2002). NCEP Stage II Data Read Me File. Accessed January 2012 at: http://data.eol.ucar.edu/codiac_data/gcip_eop/docs/katz_stageII_readme.txt.

Kidd, C., and Levizzani, V. (2011). "Status of satellite precipitation retrievals." *Hydrology and Earth System Sciences* 15: 1109–1116.

Kistler, Robert. (2012). “Evaluating Recently Developed Reanalysis Projects” presented at the NOAA CPO MAPP Webinar, February 14.

Kondragunta, C., and Shrestha, K. (2006). “Automated Real-Time Operational Rain Gauge Quality-Control Tools in NWS Hydrologic Operations.” *20th Conference on Hydrology, American Meteorological Society*, Atlanta, GA. http://ams.confex.com/ams/Annual2006/techprogram/paper_102834.htm.

Koutsoyiannis, D., and C. Onof. (2001). “Rainfall disaggregation using adjusting procedures on a Poisson cluster model.” *Journal of Hydrology* 246, no. 1-4: 109–122.

Krajewski, W.F. (1987). Cokriging Radar-Rainfall and Rain Gage Data. *Journal of Geophysical Research*, 92, 9571-9580.

Kuligowski, R. J. (2002). “A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates.” *J. Hydrometeor.* 3, 112-130.

Kumar, Arun. (2012). “The NCEP Climate Forecast System Reanalysis (CFSR)” presented at the NOAA CPO MAPP Webinar, February 14.

Li, J., and Heap, A. D. (2010). “A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors.” *Ecological Informatics*.

Lin, Shengpan, Nathan J. Moore, Joseph P. Messina, Mark H. DeVisser, and Jiaping Wu. (2012). “Evaluation of Estimating Daily Maximum and Minimum Air Temperature with MODIS Data in East Africa.” *International Journal of Applied Earth Observation and Geoinformation* 18 (0) (August): 128–140. doi:10.1016/j.jag.2012.01.004.

Luo, Y., F. Zhao, Q. Shan, Z. Yang, and P. Xie. (2011). “A Gauge-Radar-Satellite Merged Analysis of High-Resolution Regional Precipitation over Eastern China: Preliminary Investigation”. San Francisco, CA: American Geophysical Union Fall Meeting.

Massachusetts Institute of Technology (MIT). (2011). FAA Weather Systems. Accessed December 2011 at: http://www.ll.mit.edu/mission/aviation/faawxsystems/mpar.html.

Matrosov, S. Y., Campbell, C., Kingsmill, D., and Sukovich, E. (2009). “Assessing Snowfall Rates from X-Band Radar Reflectivity Measurements.” *Journal of Atmospheric and Oceanic Technology* 26 (11): 2324-2339. doi:10.1175/2009JTECHA1238.1.

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B. (2002). “A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States.” *Journal of Climate* 15 (22): 3237–3251.

Maurer, E. (2011). “Gridded Meteorological Data: 1949-2010.” Accessed December 2011 at: http://www.engr.scu.edu/~emaurer/gridded_obs/index_gridded_obs.html.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jovi, D., et al. (2006). “North American regional reanalysis.” *Bulletin of the American Meteorological Society* 87 (3): 343–360.

Mezghani, A., and Hingra, B. (2009). “A combined downscaling-disaggregation weather generator for stochastic generation of multisite hourly weather variables over complex terrain: Development and multi-scale validation for the Upper Rhone River basin.” *Journal of Hydrology* 377 (3-4): 245–260.

National Air and Space Administration (NASA). (2011a). "NLDAS-2 Forcing Dataset Information." Accessed December 2011 at: http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php.

National Air and Space Administration (NASA). (2011b). "LDAS Land Data Assimilation Systems." Accessed January 2012 at: http://ldas.gsfc.nasa.gov/nldas/NLDASgoals.php.

National Center for Environmental Prediction (NCEP). (2012). "CPC Gauge-Based Analysis of Global Daily Precipitation." Accessed January 2012 at: ftp://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB/DOCU/PRCP_CU_GAUGE_V1.0GLB_0.50deg_README.txt.

Natural Resources Conservation Service (NRCS). (2012). "SNOTEL Data & Products." Accessed January 2012 at: http://ftp.wcc.nrcs.usda.gov/snow/.

National Severe Storm Laboratory (NSSL). (2011a). "WSR-88D/TDWR Operational Product Development and Improvement." Accessed December 2011 at: http://www.nssl.noaa.gov/divisions/warning/swat/tdwr.php.

National Severe Storms Laboratory (NSSL). (2011b). "NMQ Verification System (QVS) Tutorial." Accessed January 2012 at: http://www.nssl.noaa.gov/projects/q2/tutorial/3dderived.php.

National Weather Service (NWS). (2004). "National Doppler Radar Sites." October 6. Accessed December 2011 at: http://www.srh.noaa.gov/radar/.

National Weather Service (NWS). (2006). "US Daily Precipitation Gridded Analysis." *Climate Prediction Center*. Accessed January 2012 at: http://www.cpc.ncep.noaa.gov/products/precip/realtime/GIS/USMEX/analysis.shtml.

National Weather Service (NWS). (2008). "MPE Editor User's Guide." February 5. Accessed January 2012 at: https://ocwws.weather.gov/intranet/whfs/.

National Weather Service (NWS). (2009). "MPE Editor DailyQC Features System Guide version OB9.2." December 17. http://www.nws.noaa.gov/os/water/RFC_support/resources/DailyQC_SystemGuideOB9x.doc.

National Weather Service (NWS). (2011a). "Draft HOSIP Research Project Plan for Gridded Hydrometeorological Forcings for Community Hydrologic Prediction System (CHPS) – [Calibration Analysis of Record]." Prepared by the *Office of Hydrologic Development*, Silver Spring, MD.

National Weather Service (NWS). (2011b). "Summary of Results of Meteorological Forcings Comparisons from CNRFC, ABRFC, NERFC, NWRFC." Prepared by the *Office of Hydrologic Development*. Silver Spring, MD.

National Weather Service (NWS). (2012). "NWSRFS User Manual Documentation" – Adobe Acrobat Format. Accessed January 2012 at: http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/htm/xrfsdocpdf.php.

Nelson, Brian, and Hilawe Semunegus. (2011). "Best Practices for Data Archive, Format, and Delivery" presented at the *2nd NOAA User Workshop on the Global Precipitation Mission*, November 29, College Park, MD.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., et al. (2007). "The JRA-25 Reanalysis." *Journal of the Meteorological Society of Japan* 85: 369-432. doi:10.2151/jmsj.85.369.

Peck, Eugene L, and John C Schaake. (1990). "Network Design for Water Supply Forecasting in the West1." *JAWRA Journal of the American Water Resources Association* 26 (1) (February 1): 87-99. doi:10.1111/j.1752-1688.1990.tb01354.x.

Pinker, Rachel T., Donglian Sun, Meng-Pai Hung, Chuan Li, and Jeffrey B. Basara. (2009). "Evaluation of Satellite Estimates of Land Surface Temperature from GOES over the United States." *Journal of Applied Meteorology and Climatology* 48 (1) (January): 167–180. doi:10.1175/2008JAMC1781.1.

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., Tewari, M., et al. (2011). "High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate." *Journal of Climate* 24 (June): 3015-3048. doi:10.1175/2010JCLI3985.1.

Rennie, J. (2011a). "Maintenance and Streamlining of the Hourly Precipitation Data (HPD) Network Dataset." *Scientific Report*, National Climatic Data Center Cooperative Institute for Climate and Satellites. Accessed December 2011 at: http://www.cicsnc.org/people/jared-rennie/progress-update/.

Rennie, J., Wilson, A., Lawrimore, J. H., Ray, R., and Menne, M. J. (2011b). "Implementing New Quality Control and Processing Systems for Hourly Precipitation Data." *2011 AGU Fall Meeting*, San Francisco, CA.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., et al. (2011). "MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications." *Journal of Climate* 24 (July): 3624-3648. doi:10.1175/JCLI-D-11-00015.1.

Riverside Technology, inc. (Riverside). (2007). Task Completion Report: Hydrologic Model Adaptation for Snowpack Runoff. August 2007.

Riverside Technology, inc. (Riverside). (2010a). Final Task Report Task 9-0009 Phase II WGRFC Rio Grande Calibration. August 2010.

Riverside Technology, inc. (Riverside). (2010b). Temperature quality control and MAT Development Standard Procedure. December 16. Internal standard operating procedure.

Riverside Technology, inc. (Riverside). (2011). Precipitation quality control and MAP Development Procedure. January 25. Internal standard operating procedure.

Ruscher, Paul H. (2004). "The Florida GLOBE Partnership Program." *Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone*; January 12-15, 2004; Seattle, WA. Accessed at: http://ams.confex.com/ams/84Annual/techprogram/paper_70010.htm.

Saha, S., Moorthi, S., Pan, H-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2010). "The NCEP Climate Forecast System Reanalysis." *Bulletin of the American Meteorological Society* 91 (8) (August): 1015-1057. doi:10.1175/2010BAMS3001.1.

Sapiano, M. R. P., Smith, T. M., and Arkin, P. A. (2008). "A new merged analysis of precipitation utilizing satellite and reanalysis data." *Journal of Geophysical Research* 113

(D22) (November). doi:10.1029/2008JD010310. http://www.agu.org/pubs/crossref/2008/2008JD010310.shtml.

Sapiano, M. R.P, Janowiak, J. E., Shi, W., Higgins, R. W., and Silva, V. B. S. (2010). "Regional Evaluation Through Independent Precipitation Measurements: USA." *Satellite Rainfall Applications for Surface Hydrology*, 169–191. New York: Springer.

Schaake, J, Henkel, A., and Cong, S. (2004). "Application of PRISM climatologies for hydrologic modeling and forecasting in the western U.S." *84th Annual Meeting of the American Meteorological Society*, Seattle, WA.

Scofield, R. A. and R. J. Kuligowski (2003). "Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events." *Mon. Wea. Rev.*, 18, 1037-1051.

Seo, D-J. (1998a). "Real-time estimation of rainfall fields using radar rainfall and rain gage data." *Journal of Hydrology* 208 (1-2) (July 10): 37-52.

Seo, D-J. (1998b). "Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions." *Journal of Hydrology* 208 (1-2) (July 10): 25-36.

Seo, D-J, Breidenbach, J.P., and Johnson, E.R.. (1999). "Real-time estimation of mean field bias in radar rainfall data." *Journal of Hydrology* 223 (3-4): 131-147. doi:10.1016/S0022-1694(99)00106-7.

Seo, D-J., and Breidenbach, J. P. (2002). "Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements." *Journal of Hydrometeorology* 3 (2) (April): 93-111. doi:10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2.

Seo, D-J. (2003). "Multi-Sensor Precipitation Estimation." *NWSRFS International Workshop* October 21, Kansas City.

Seo, Dong-Jun, Lee Cajina, Robert Corby, and Tracy Howieson. (2009). "Automatic state updating for operational streamflow forecasting via variational data assimilation." *Journal of Hydrology* 367 (3-4): 255-275.

Shepard, D. (1968). "A two dimensional interpolation function for regularly spaced data." *Proceedings: 23rd National Conference of the Association for Computing Machinery*, 517-524. Pinceton, NJ: ACM.

Shepard, D. S. (1984). "Computer mapping: The SYMAP interpolation algorithm." *Spatial Statistics and Models*, G. L. Gaile and C. J. Willmott, Eds., D. Reidel, 133–145.

Sibson, R. (1981). "A brief description of natural neighbour interpolation." *Interpreting Multivariate Data*, 21-36. Edited by V. Barnett. Chichester, UK: John Wiley and Sons.

Silverman, N, and Maneta, M. P. (2011). "Evaluation of a WRF dynamic downscaling simulation over Western Montana." *AGU Fall Meeting 2011*. San Francisco, CA.

Simmons, A. J., P. D. Jones, V. da Costa Bechtold, A. C. M. Beljaars, P. W. Kållberg, S. Saarinen, S. M. Uppala, P. Viterbo, and N. Wedi. (2004). "Comparison of Trends and Low-frequency Variability in CRU, ERA-40, and NCEP/NCAR Analyses of Surface Air Temperature." *Journal of Geophysical Research* 109 (December 31): 18 PP. doi:200410.1029/2004JD005306.

Smith, M.B., Seo D.J., Finnerty, B.D., and V. Koren. (1996). "Distributed Parameter Hydrologic Modeling and NEXRAD for River Forecasting: Scale Issues Facing the National Weather Service." *1996 ASCE Conference*, June 24-28, 1996. Accessed 2012 at: http://www.nws.noaa.gov/oh/hrl/papers/distmodel/asceb.htm.

Stahl, K, R D Moore, J A Floyer, M G Asplin, and I G Mckendry. (2006). "Comparison of Approaches for Spatial Interpolation of Daily Air Temperature in a Large Region with Complex Topography and Highly Variable Station Density." *Agricultural and Forest Meteorology* 139 (3-4): 224–236. doi:10.1016/j.agrformet.2006.07.004.

Story, Greg. (2011). "Recent Changes Made in the Hydrometeorological Analysis and Support Function at the NWS West Gulf RFC." *2011 National Hydrologic Warning Council Training Conference*; May 9-12, 2011; San Diego, CA. Session 30. http://documents.clubexpress.com/documents.ashx?key=X3T8Wc5cElH8MTweMhXQMYeG34f6Pxjf

Tian, Y., Peters-Lidard, C. D., Adler, R. F., Kubota, T., and Ushio, T. (2010). "Evaluation of GSMaP Precipitation Estimates over the Contiguous United States." *Journal of Hydrometeorology* 11 (2) (April): 566-574.

Turlapaty, Anish C., Valentine G. Anantharaj, Nicolas H. Younan, and F. Joseph Turk. (2010). "Precipitation data fusion using vector space transformation and artificial neural networks." *Pattern Recogn. Lett.* 31 (10) (July): 1184–1200. doi:10.1016/j.patrec.2009.12.033.

Tveito, O. (2002). "Spatial distribution of winter temperatures in Norway related to topography and large scale atmospheric circulation." In *AHS Workshop*.

Uppala, S. M., KÅllberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., et al. (2005). "The ERA-40 re-analysis." *Quarterly Journal of the Royal Meteorological Society* 131 (October): 2961-3012.

Unidad de Informatica para las Ciencias Atmosfericas y Ambientales (UNIATMOS). (2011). Atlas Climatico Digital de Mexico. Accessed December 2011 at: http://atlasclimatico.unam.mx/atlas/uniatmos.html.

Urzen, M. L., Ansari, S., and Del Greco, S. A. (2004). "Automated Spatial Precipitation Estimator (PrecipVal)." *20th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Seattle, WA. http://ams.confex.com/ams/84Annual/techprogram/paper_70681.htm.

U.S. Army Corps of Engineers (USACE). (2007). "SNTHERM-Snow Cover Temperature Model." Accessed July 2007 from: http://snow.usace.army.mil/model_info/sntherm.html.

U.S. Army Corps of Engineers (USACE), U.S. Geological Survey (USGS), and National Oceanic and Atmospheric Administration (NOAA). (2011). "Memorandum of Understanding: Collaborative Science, Services, and Tools to Support Integrated and Adaptive Water Resources Management." Accessed November 2011 at: http://www.building-collaboration-for-water.org/Documents/FederalToolboxMOU_051111.pdf.

U.S. Geological Survey (USGS). (2002). "Report to Congress: Concepts for National Assessment of Water Availability and Use." *Circular 1223*. Reston, VA. Accessed November 2011 at: http://pubs.usgs.gov/circ/circ1223/.

U.S. Geological Survey (USGS). (2011). "Great Lakes Basin Pilot." Last Updated February 2011. Accessed December 2011 at: http://water.usgs.gov/wateravailability/greatlakes/index.html.

Ushio, T., Sasashige, K., Kubota, K., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., et al. (2009). "A Kalman Filter Approach to the Global Satellite Mapping of Precipitation (GSMaP) from Combined Passive Microwave and Infrared Radiometric Data." *Journal of the Meteorological Society of Japan. Ser. II* 87A: 137-151.

Vancutsem, Christelle, Pietro Ceccato, Tufa Dinku, and Stephen J. Connor. (2010). "Evaluation of MODIS Land Surface Temperature Data to Estimate Air Temperature in Different Ecosystems over Africa." *Remote Sensing of Environment* 114 (2) (February 15): 449–465. doi:10.1016/j.rse.2009.10.002.

Vasiloff, S., Seo, D-J., Howard, K. W., Zhang, J., Kitzmiller, D. H., Mullusky, M. G., Krajewski, W. F., et al. (2007). "Improving QPE and Very Short Term QPF." *Bulletin of the American Meteorological Society* 88: 1899-1911.

Wan, Zhengming. (2008). "New Refinements and Validation of the MODIS Land-Surface Temperature/Emissivity Products." *Remote Sensing of Environment* 112 (1) (January 15): 59–74. doi:10.1016/j.rse.2006.06.026.

Wan, Zhengming and J. Dozier. (1996). "A Generalized Split-window Algorithm for Retrieving Land-surface Temperature from Space." *IEEE Transactions on Geoscience and Remote Sensing* 34 (4) (July): 892–905. doi:10.1109/36.508406.

Wang, S. H., Bromwich, D. H., and Bai, L. (2011). "Arctic System Reanalysis Interim (ASR-interim) output - A 30 km resolution dataset." *11th Conference on Polar Meteorology and Oceanography*, Boston, MA.

West, G. L., Steenburgh, W. J., and Cheng, W. Y. Y. (2007). "Spurious Grid-Scale Precipitation in the North American Regional Reanalysis." *Monthly Weather Review* 135 (6) (June): 2168-2184. doi:10.1175/MWR3375.1.

Wheater, H. S., R. E. Chandler, C. J. Onof, V. S. Isham, E. Bellone, C. Yang, D. Lekkas, G. Lourmas, and M.-L. Segond. (2005). "Spatial-temporal rainfall modelling for flood risk estimation." *Stochastic Environmental Research and Risk Assessment* 19 (October 14): 403-416.

Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., and Mearns, L. O. (2004). "Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods." *Intergovernmental Panel on Climate Change*, August.

Willmott, C. J., and Matsuura, K. (2006). "On the use of dimensioned measures of error to evaluate the performance of spatial interpolators." *International Journal of Geographical Information Science* 20 (January): 89-102. doi:10.1080/13658810500286976.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P. (2004). "Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs." *Climatic Change* 62 (1-3) (January): 189-216. doi:10.1023/B:CLIM.0000013685.99609.9e.

Wood, A. W., and Lettenmaier, D. P. (2006). “A Test Bed for New Seasonal Hydrologic Forecasting Approaches in the Western United States.” *Bulletin of the American Meteorological Society* 87 (12) (December): 1699-1712. doi:10.1175/BAMS-87-12-1699.

Wüest, Marc, Christoph Frei, Adrian Altenhoff, Martin Hagen, Michael Litschi, and Christoph Schär. (2009). “A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation.” *International Journal of Climatology.*

Xie, Pingping, and An-Yuan Xiong. (2011). “A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses.” *Journal of Geophysical Research* 116 (November 8): 14 PP. doi:201110.1029/2011JD016118.

Xie, P., and Arkin, P. A. (1997). “Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs.” *Bulletin of the American Meteorological Society* 78 (November): 2539-2558.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., and Liu, C. (2007). “A Gauge-Based Analysis of Daily Precipitation over East Asia.” *Journal of Hydrometeorology* 8 (3) (June): 607-626. doi:10.1175/JHM583.1.

Zhang, J., Howard, K., and Vasiloff, S. (2005). “3-D Radar Mosaic and Initial Q2 Development Plans.” *Q2 Workshop*, Norman, OK.

Zhang, Yu, Seann Reed, and David Kitzmiller. (2011). “Effects of Retrospective Gauge-Based Readjustment of Multisensor Precipitation Estimates on Hydrologic Simulations.” *Journal of Hydrometeorology* 12 (3) (June): 429–443. doi:10.1175/2010JHM1200.1.