



NUREG/CR-7271
ORNL/SPR-2019/1323

Application of Point Precipitation Frequency Estimates to Watersheds

AVAILABILITY OF REFERENCE MATERIALS IN NRC PUBLICATIONS

NRC Reference Material

As of November 1999, you may electronically access NUREG-series publications and other NRC records at the NRC's Public Electronic Reading Room at <http://www.nrc.gov/reading-rm.html>. Publicly released records include, to name a few, NUREG-series publications; *Federal Register* notices; applicant, licensee, and vendor documents and correspondence; NRC correspondence and internal memoranda; bulletins and information notices; inspection and investigative reports; licensee event reports; and Commission papers and their attachments.

NRC publications in the NUREG series, NRC regulations, and Title 10, "Energy," in the *Code of Federal Regulations* may also be purchased from one of these two sources.

1. The Superintendent of Documents

U.S. Government Publishing Office
Washington, DC 20402-0001
Internet: <http://bookstore.gpo.gov>
Telephone: 1-866-512-1800
Fax: (202) 512-2104

2. The National Technical Information Service

5301 Shawnee Road
Alexandria, VA 22161-0002
<http://www.ntis.gov>
1-800-553-6847 or, locally, (703) 605-6000

A single copy of each NRC draft report for comment is available free, to the extent of supply, upon written request as follows:

U.S. Nuclear Regulatory Commission

Office of Administration
Multimedia, Graphics and Storage & Distribution Branch
Washington, DC 20555-0001
E-mail: distribution.resource@nrc.gov
Facsimile: (301) 415-2289

Some publications in the NUREG series that are posted at the NRC's Web site address <http://www.nrc.gov/reading-rm/doc-collections/nuregs> are updated periodically and may differ from the last printed version. Although references to material found on a Web site bear the date the material was accessed, the material available on the date cited may subsequently be removed from the site.

Non-NRC Reference Material

Documents available from public and special technical libraries include all open literature items, such as books, journal articles, transactions, *Federal Register* notices, Federal and State legislation, and congressional reports. Such documents as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings may be purchased from their sponsoring organization.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at—

The NRC Technical Library

Two White Flint North
11545 Rockville Pike
Rockville, MD 20852-2738

These standards are available in the library for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from—

American National Standards Institute

11 West 42nd Street
New York, NY 10036-8002
<http://www.ansi.org>
(212) 642-4900

Legally binding regulatory requirements are stated only in laws; NRC regulations; licenses, including technical specifications; or orders, not in NUREG-series publications. The views expressed in contractor-prepared publications in this series are not necessarily those of the NRC.

The NUREG series comprises (1) technical and administrative reports and books prepared by the staff (NUREG-XXXX) or agency contractors (NUREG/CR-XXXX), (2) proceedings of conferences (NUREG/CP-XXXX), (3) reports resulting from international agreements (NUREG/IA-XXXX), (4) brochures (NUREG/BR-XXXX), and (5) compilations of legal decisions and orders of the Commission and Atomic and Safety Licensing Boards and of Directors' decisions under Section 2.206 of NRC's regulations (NUREG-0750).

DISCLAIMER: This report was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any employee, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product, or process disclosed in this publication, or represents that its use by such third party would not infringe privately owned rights.



NUREG/CR-7271
ORNL/SPR-2019/1323

Application of Point Precipitation Frequency Estimates to Watersheds

Manuscript Completed: May 2020
Date Published: February 2021

Prepared by:
Shih-Chieh Kao*
Scott T. DeNeale*

Elena Yegorova, NRC Project Manager

Oak Ridge National Laboratory*
P.O. Box 2008
Oak Ridge, TN 37831

Office of Nuclear Regulatory Research

ABSTRACT

This report documents work sponsored by the U.S. Nuclear Regulatory Commission (NRC) at the Oak Ridge National Laboratory (ORNL) as part of the RES project, "Application of Point Precipitation Frequency Estimates to Watersheds." This project was implemented as part of the Probabilistic Flood Hazard Assessment (PFHA) Research Program. The objective of the PFHA Research Program is to develop tools and guidance on the use of PFHA methods to risk-inform NRC's licensing of new facilities as well as licensing and oversight of currently operating facilities as they relate to flooding hazards.

Many nuclear power plants (NPPs) are located on or near rivers so riverine flooding hazards need to be considered in their design and operation. Probabilistic riverine flood models are important tools for realistic assessment of flooding risks. However, these models require areal estimates of the depth, duration, and frequency of rainfall distributed over the watershed, which are not often available. Point precipitation frequency estimates are more widely available. For example, the National Oceanic and Atmospheric Administration (NOAA) has published NOAA Atlas 14, which provides point precipitation frequency estimates for 5-minute through 60-day durations at average recurrence intervals of 1-year through 1,000-year. The research documented in this report addresses areal reduction factors (ARFs), which can be used to convert the widely available point precipitation frequency estimates, to estimates of areal precipitation frequency over a watershed.

The most widely used ARF source is Technical Paper 29 (TP-29) published by the then U.S. Weather Bureau in 1958. However, both the methods and the underlying precipitation data used to produce TP-29 are seriously out of date. For example, due to the small gauge network available at the time of TP-29's compilation, ARF estimates developed are only for watersheds smaller than about 400 square miles. Due to the relatively short record lengths of precipitation data available, frequency considerations could not be accurately determined. Other factors such as regional climate and seasonality were not addressed.

Several newer methods have been published since TP-29 was developed and both the type and quantity of precipitation data have increased significantly, along with computational resources and analytical tools such as geographic information systems. This report reviewed and assessed the available precipitation products and methods for conducting ARF analysis. The work applied up-to-date precipitation data products and analysis methods with a novel watershed-based approach to investigate how ARF estimates vary across different methods, data sources, geographical locations, return periods, and seasons.

The overall findings reported here regarding basic ARF trends are in line with other recent studies showing that ARFs decrease with increasing area, increase with increasing duration, and decrease with increasing return period. This study found significant differences among the available ARF methods. This work also found a strong geographical variability across different US hydrologic regions, suggesting that the ARF are specific to regional climate patterns and geographical characteristics and should not be applied arbitrarily to other locations. The results also reveal the importance of data record length, especially for high return level ARFs.

The work reported in NUREG/CR-7271 will assist NRC staff in assessing different classes of ARF methods in conjunction with available rainfall data sets. It will also support the development of guidance for application of point precipitation data in PFHAs. It should be noted that the ARF values presented in this report for any location or region were developed for the purposes of

comparing methods and investigating the factors that influence ARFs. They should not be considered official and should not be used in leu of a site-specific analysis.

FOREWORD

This report (NUREG/CR-7271) documents work sponsored by the U.S. Nuclear Regulatory Commission (NRC) as part of the RES project “Application of Point Precipitation Frequency Estimates to Watersheds”. The research conducted supports the NRC’s Probabilistic Flood Hazard Assessment (PFHA) program. The objective of the project was to assist NRC in assessing different classes of fixed-area precipitation areal reduction factors (ARF) methods in conjunction with available rainfall data sets to support the development of guidance for application of PFHA. Given the limitations of available ARF products being used today (e.g., TP-29) and with the advance of recent observational precipitation products and computational capabilities, more comprehensive ARF evaluations can be made to understand how ARF estimates vary across different methods, data sources, geographical locations, return periods, and seasons. These topics are explored in this study.

This research is part of the NRC’s PFHA Research Program and is to assist NRC in assessing different classes of fixed-area ARF methods in conjunction with available rainfall datasets to support development of guidance for application of NPP-PFHA. The work will aid the development of guidance on the use of PFHA methods and support risk-informing NRC’s licensing framework (flood hazard design standards at proposed new facilities as well as significance determination tools for evaluating potential deficiencies related to flood protection at operating facilities). The tools and guidance developed will support and enhance NRC’s capacity to perform thorough and efficient reviews of license applications and license amendment requests. They will also support risk-informed significance determination of inspection findings, unusual events and other oversight activities.

NUREG/CR-7271 summarizes available precipitation products and methods for conducting ARF analysis. Using this information, a series of use case studies are developed for both regional ARF assessment and for contiguous U.S. (CONUS)-scale assessment. The use case findings produce ARF relationships which are in line with available literature; they also demonstrate the importance of precipitation data source and ARF fitting method which both contribute to ARF uncertainty. In particular, the importance of available data length is highlighted given ARFs are often sought for long return periods. The study demonstrates the need to improve ARFs with new data and methods for more reliable areal extreme precipitation estimates to support PFHA applications.

The main objective of this study is to assist NRC in assessing different classes of fixed-area ARF methods in conjunction with available rainfall data sets to support the development of guidance for application of PFHA. The results of this study are for demonstration purposes only and are not intended to be used for ARF application. Additional research and development efforts, with thorough quality assurance and control performed, should be performed to develop a reliable national ARF product suitable for PFHA application.

The discussion of specific references, methods, software, or tools in this NUREG/CR does not constitute an endorsement or approval for any specific use by the U.S. Nuclear Regulatory Commission or Oak Ridge National Laboratory. The case study results presented herein are the result of research efforts only, do not incorporate uncertainty quantification, and should not be directly incorporated for application. They are intended to demonstrate some of the primary factors affecting areal reduction factor estimation.

TABLE OF CONTENTS

ABSTRACT	iii
FOREWORD.....	v
LIST OF FIGURES.....	ix
LIST OF TABLES	xiii
EXECUTIVE SUMMARY	xv
ACKNOWLEDGMENTS	xvii
ABBREVIATIONS AND ACRONYMS	xix
1 INTRODUCTION	1-1
1.1 Background and Motivation	1-1
1.2 Nuclear Regulatory Context.....	1-2
1.3 Areal Reduction Factors	1-2
1.4 Fixed-Area versus Storm-Centered ARFs	1-4
1.5 Research Objective	1-5
2 PRECIPITATION PRODUCTS FOR ARF ANALYSIS	2-1
2.1 Overview of Precipitation Products	2-1
2.1.1 Categories of Precipitation Products.....	2-1
2.1.2 Key Metrics for Consideration.....	2-2
2.2 Summary of Available Precipitation Products	2-2
2.2.1 Gauge-only Precipitation Data Sets.....	2-3
2.2.2 Gauge-driven Precipitation Products	2-4
2.2.3 Radar-driven Precipitation Products	2-6
2.2.4 Satellite-driven Precipitation Products	2-8
2.2.5 Reanalysis-driven Precipitation Products.....	2-9
2.3 Considerations for ARF	2-10
3 METHODS FOR ARF ANALYSIS	3-1
3.1 Factors Affecting ARF	3-1
3.1.1 Storm Characteristics	3-1
3.1.2 Geographic Features.....	3-3
3.1.3 ARF Methodology.....	3-5
3.1.4 Data.....	3-6
3.2 Critical Review of ARF Methods	3-8
3.2.1 Qualitative Assessment Considerations.....	3-8
3.2.2 Available ARF Methods	3-8
3.3 Selection of Potential ARF Methods for PFHA.....	3-24
4 ARF USE CASE STUDIES.....	4-1
4.1 Study Domain.....	4-2
4.1.1 Regional Assessments	4-2
4.1.2 CONUS Assessment	4-5

4.2	Selected Precipitation Products	4-6
4.3	ARF Assessment.....	4-9
4.3.1	AMS Identification.....	4-9
4.3.2	Sample ARF Calculation.....	4-13
4.3.3	ARF Model Fitting.....	4-14
4.4	Use Case Results.....	4-19
4.4.1	Ohio (Region 05)	4-19
4.4.2	Mid-Atlantic (Region 02)	4-28
4.4.3	South Atlantic-Gulf (Region 03)	4-35
4.4.4	CONUS Reconnaissance-level Assessment.....	4-42
5	SUMMARY	5-1
5.1	ARF Characteristics in the United States.....	5-1
5.2	Challenges for PFHA Applications.....	5-3
6	REFERENCES	6-1
APPENDIX A	SUMMARY OF AVAILABLE PRECIPITATION PRODUCTS.....	A-1
APPENDIX B	QUALITATIVE ASSESSMENT RESULTS FOR ARF METHODS....	B-1
APPENDIX C	OHIO (REGION 05) DETAILED RESULTS	C-1
APPENDIX D	MID-ATLANTIC (REGION 02) DETAILED RESULTS.....	D-1
APPENDIX E	SOUTH ATLANTIC-GULF (REGION 03) DETAILED RESULTS	E-11
APPENDIX F	CONUS ASSESSMENT DETAILED RESULTS	F-1

LIST OF FIGURES

Figure 1-1	TP-29 ARF Curves.....	1-3
Figure 3-1	TP-29 ARF Curves.....	3-11
Figure 3-2	UK Flood Studies Report ARF Curves Based on Koutsoyiannis and Xanthopoulos Fitted Model	3-13
Figure 3-3	TR-24 ARF Calculation Process	3-15
Figure 4-1	Ohio (Region 05) 4-digit, 6-digit, and 8-digit Hydrologic Units (From Top Left to Bottom)	4-3
Figure 4-2	Mid-Atlantic (Region 02) 4-digit, 6-digit, and 8-digit Hydrologic Units (From Top Left to Bottom).....	4-4
Figure 4-3	South Atlantic–Gulf (Region 03) 4-digit, 6-digit, and 8-digit Hydrologic Units (From Top Left to Bottom).....	4-5
Figure 4-4	HUC02 Hydrologic Regions in CONUS.....	4-6
Figure 4-5	NCEI Hourly Rainfall Stations with 30+ Years of Record in Ohio (Region 05)	4-9
Figure 4-6	Spatial and Temporal Aggregation Diagram Used for Gridded Precipitation Products	4-10
Figure 4-7	Example of AMS Searing in the Ohio River Basin.....	4-11
Figure 4-8	Example Spatial Aggregation of Ohio Region HUC08s to form a HUCac Upstream of HUC 08 05090203	4-12
Figure 4-9	Example of M1 ARF Model Fitting.....	4-15
Figure 4-10	Example of M2 ARF Model Fitting.....	4-15
Figure 4-11	Example of M3 ARF Model Fitting.....	4-16
Figure 4-12	Example of M4 ARF Model Fitting.....	4-17
Figure 4-13	Example of M5 ARF Model Fitting.....	4-17
Figure 4-14	Example of M6 ARF Model Fitting.....	4-18
Figure 4-15	Comparison of 1-day, 10-year Ohio ARF Fitting Using PRISM Precipitation Across Different ARF Models.....	4-20
Figure 4-16	Comparison of 1-day, 10-year Ohio M5 ARF Fitted by Different Precipitation Products	4-22
Figure 4-17	Comparison of PRISM-based 1-day Ohio M5 ARF across Different Frequency Levels.....	4-24
Figure 4-18	Comparison of 10-year Ohio M5 ARF Using PRISM Precipitation across Different Durations	4-25
Figure 4-19	Comparison of 10-year Ohio M5 ARF Using DSI-3240 Precipitation across Different Durations.....	4-26
Figure 4-20	Comparison of 1-day, 10-year Ohio M5 ARF Using PRISM Precipitation across Different Seasons	4-27

Figure 4-21	Comparison of 1-day, 10-year Mid-Atlantic ARF Fitting Using PRISM Precipitation across Different ARF Models	4-29
Figure 4-22	Comparison of 1-day, 10-year Mid-Atlantic M5 ARF Fitted by Different Precipitation Products	4-30
Figure 4-23	Comparison of PRISM-based 1-day Mid-Atlantic M5 ARF across Different Frequency Levels.....	4-32
Figure 4-24	Comparison of 10-year Mid-Atlantic M5 ARF Using PRISM Precipitation across Different Durations.....	4-33
Figure 4-25	Comparison of 1-day, 10-year Mid-Atlantic M5 ARF Using PRISM Precipitation across Different Seasons.....	4-34
Figure 4-26	Comparison of 1-day, 10-year South Atlantic-Gulf ARF Fitting Using PRISM Precipitation across Different ARF Models.....	4-36
Figure 4-27	Comparison of 1-day, 10-year South Atlantic-Gulf M5 ARF Fitted by Different Precipitation Products.....	4-37
Figure 4-28	Comparison of PRISM-based 1-day South Atlantic-Gulf M5 ARF across Different Frequency Levels	4-39
Figure 4-29	Comparison of 10-year South Atlantic-Gulf M5 ARF Using PRISM Precipitation across Different Durations	4-40
Figure 4-30	Comparison of 1-day, 10-year South Atlantic-Gulf M5 ARF Using PRISM Precipitation across Different Seasons.....	4-41
Figure 4-31	CONUS ARF Assessment for 1-day Duration and 100-year Return Period Using PRISM-Daily Data and M5 Model Fitting.....	4-43
Figure 4-32	CONUS ARF Assessment for 1-day Duration, 10,000-mi ² Area, and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting.....	4-43
Figure 4-33	Map of CONUS 1-day ARFs Using PRISM-daily Data and M5 Model Fitting across Different Areas and Return Periods.....	4-44
Figure 4-34	Map of CONUS 100-year ARFs Using PRISM-daily Data and M5 Model Fitting across Different Durations and Areas	4-46
Figure C-1	Calculated Ohio ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models <i>Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.</i>	C-2
Figure C-3	Calculated Ohio ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models <i>Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.</i>	C-4
Figure C-4	Calculated Ohio ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-5
Figure C-5	Calculated Ohio ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-6
Figure C-6	Calculated Ohio ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-7
Figure C-7	Calculated Ohio ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-8

Figure C-8	Calculated Ohio ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-9
Figure C-9	Calculated Ohio ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-10
Figure D-1	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models	D-2
Figure D-2	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models	D-3
Figure D-3	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models	D-4
Figure D-4	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-5
Figure D-5	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-6
Figure D-6	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-7
Figure D-7	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-8
Figure D-8	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-9
Figure D-9	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-10
Figure E-1	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-12
Figure E-2	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-13
Figure E-3	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-14
Figure E-4	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-15
Figure E-5	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-16
Figure E-6	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-17
Figure E-7	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-18
Figure E-8	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-19
Figure E-9	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-20

Figure F-1	CONUS ARF Assessment Maps for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes.....	F-2
Figure F-2	CONUS ARF Assessment Maps for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-3
Figure F-3	CONUS ARF Assessment Maps for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-4
Figure F-4	CONUS ARF Assessment Maps for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-5
Figure F-5	CONUS ARF Assessment Maps for 3-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-6
Figure F-6	CONUS ARF Assessment Plot for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting	F-7
Figure F-7	CONUS ARF Assessment Plot for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-8
Figure F-8	CONUS ARF Assessment Plot for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-9
Figure F-9	CONUS ARF Assessment Plot for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-10
Figure F-10	CONUS ARF Assessment Plot for 3-day duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-11

LIST OF TABLES

Table 3-1	Summary of Factors Affecting Precipitation ARF Calculations	3-1
Table 3-2	Summary of Available ARF Calculation Methods	3-9
Table 4-1	Summary of the Overall Design of the Use Case Study	4-1
Table 4-2	Selected Precipitation Products and Features Used in Demonstrations	4-8
Table 4-3	Summary of fitting model complexity	4-19
Table 4-4	Comparison of 1-day Ohio ARF Fitting Using PRISM Precipitation Across Different ARF Models.....	4-21
Table 4-5	Comparison of 1-day Ohio M5 ARF Fitting across Different Data Sources.....	4-22
Table 4-6	Comparison of Ohio M5 ARF fitting Using PRISM Precipitation across Different Durations	4-25
Table 4-7	Comparison of 1-day Ohio M5 ARF Fitting Using PRISM Precipitation across Different Seasons	4-28
Table 4-8	Comparison of 1-day Mid-Atlantic ARF Fitting Using PRISM Precipitation across Different ARF Models	4-29
Table 4-9	Comparison of 1-day Mid-Atlantic M5 ARF Fitting across Different Data Sources	4-31
Table 4-10	Comparison of Mid-Atlantic M5 ARF Fitting Using PRISM Precipitation across Different Durations.....	4-33
Table 4-11	Comparison of 1-day Mid-Atlantic M5 ARF Fitting Using PRISM Precipitation across Different Seasons.....	4-35
Table 4-12	Comparison of 1-day South Atlantic-Gulf ARF Fitting Using PRISM Precipitation across Different ARF Models	4-36
Table 4-13	Comparison of 1-day South Atlantic-Gulf M5 ARF Fitting across Different Data Sources	4-38
Table 4-14	Comparison of South Atlantic-Gulf M5 ARF Fitting Using PRISM Precipitation across Different Durations	4-40
Table 4-15	Comparison of 1-day South Atlantic-Gulf M5 ARF Fitting Using PRISM Precipitation across Different Seasons.....	4-42
Table 4-16	Comparison of 1-day CONUS Regional M5 ARF Fitting Using PRISM Precipitation across Different Return Periods.....	4-45
Table 4-17	Comparison of 100-year CONUS Regional M5 ARF Fitting Using PRISM Precipitation across Different Durations	4-46
Table B-1	ARF Method Assessment Matrix	B-2

EXECUTIVE SUMMARY

This research is part of the NRC's PFHA Research Program and is to assist NRC in assessing different classes of fixed-area ARF methods in conjunction with available rainfall datasets to support development of guidance for application of NPP-PFHA. The work will aid the development of guidance on the use of PFHA methods and support risk-informing NRC's licensing framework (flood hazard design standards at proposed new facilities as well as significance determination tools for evaluating potential deficiencies related to flood protection at operating facilities). The tools and guidance developed will support and enhance NRC's capacity to perform thorough and efficient reviews of license applications and license amendment requests. They will also support risk-informed significance determination of inspection findings, unusual events and other oversight activities.

To support PFHA of nuclear power plants (NPPs), probabilistic estimates of extreme rainfall depth across various watershed sizes are required. Nevertheless, most existing precipitation frequency analysis (PFA) products (such as NOAA Atlas 14) provide frequency estimates of "point" precipitation that can only be representative for a small domain and are not appropriate for large-scale watershed modeling applications. The ARF examined in this study, which is the ratio of areal extreme rainfall depth to point-based extreme rainfall depth, is one commonly-used approach to derive areal extreme rainfall estimates from conventional point-based PFA products.

An ARF can be generally defined as the ratio of areal extreme rainfall depth (i.e., total observed rainfall volume across an area divided by the area of interest) to point-based extreme rainfall depth (i.e., observed rainfall depth at a point location or for a representatively small area). The use of ARF is necessary because networks of rain gauges with long periods of record, which are needed for accurate areal rainfall frequency estimation, are generally sparse and do not allow for an appropriate characterization of the associated spatial rainfall patterns. If the ARF relationship is known, point-based PFA products and ARF can be used to approximate areal extreme rainfall for a watershed.

Compared to modern PFA products, the progress of ARF development in the U.S. is relatively slow, and the TP-29 ARFs published in the 1950s are still used in practice. These TP-29 estimates suffer from major limitations, including the use of very limited rain gauge data, the application to only small area sizes, and the lack of variation across geographic location, return period, and seasonality. Given these limitations and with the advance of recent observational precipitation products and computational capabilities, more comprehensive ARF evaluations can be made to understand how ARF estimates vary across different methods, data sources, geographical locations, return periods, and seasons. These topics are explored in this study.

To improve the understanding of ARF variability, this study conducts a comprehensive review of recent ARF research, summarizes potential precipitation products for ARF applications, and provides use case studies to demonstrate the derivation of ARF in several selected hydrologic regions in the U.S. The survey of available precipitation products covers several major categories, including gauge-only, gauge-driven, radar-driven, satellite-driven, and reanalysis-driven products. Each precipitation product is evaluated based on its availability, accuracy, spatiotemporal resolution, latency, and suitability for ARF assessment. Several precipitation products are subsequently selected for comparison in the use case studies.

A review of ARF research identifies the key factors affecting ARF estimation, including storm characteristics, geographic features, ARF methodology, and data. A critical review is conducted to summarize available ARF methods across five major types: empirical methods, spatial correlation methods, statistical crossing properties methods, spatial and temporal scaling methods, and extreme value theory methods. Each method is evaluated by its data dependency, required assumptions, analytical complexity, spatiotemporal scale, and whether it has been independently evaluated in other studies. Several suitable ARF methods are then selected and compared in the use case studies.

The use case studies implement a novel watershed-based annual maximum precipitation searching approach to identify ARF samples across different watershed sizes for further ARF model fitting. Through these use cases, a quantitative comparison of major factors affecting ARFs is provided. The use cases include (1) regional assessments of ARFs for three selected hydrologic regions focusing on different precipitation products, and fitting models to demonstrate major factors affecting ARFs and (2) a CONUS assessment of ARFs across all hydrologic regions, focusing on the use of one precipitation product and one fitting model to demonstrate geographic variation in ARFs.

The use case study results are generally in line with available literature which suggest ARFs decrease with increasing area, increase with increasing duration, and decrease with increasing return period. The results also demonstrate the importance of precipitation data source and ARF fitting method which both contribute to ARF uncertainty. In particular, the importance of available data length is highlighted given ARFs are often sought for long return periods. The study demonstrates the need to improve ARFs with new data and methods for more reliable areal extreme precipitation estimates to support PFHA applications. Based on the results, ARF characteristics and PFHA application challenges are also summarized in this report. The results of this study are for demonstration purposes only and are not intended to be used for ARF application.

The discussion of specific references, methods, software, or tools in this NUREG/CR does not constitute an endorsement or approval for any specific use by the U.S. Nuclear Regulatory Commission or Oak Ridge National Laboratory. The case study results presented herein are the result of research efforts only, do not incorporate uncertainty quantification, and should not be directly incorporated for application. They are intended to demonstrate some of the primary factors affecting areal reduction factor estimation.

ACKNOWLEDGMENTS

US Nuclear Regulatory Commission

The authors would like to acknowledge and express their appreciation to the U.S. Nuclear Regulatory Commission (NRC) for overseeing and funding this research opportunity to evaluate point-to-area precipitation conversion approaches and summarize key information in this NUREG/CR. The following NRC staff were heavily involved in supporting this research effort.

- Elena Yegorova, Project Manager
- Joseph Kanney
- Meredith Carr

Oak Ridge National Laboratory

In addition, the following individuals from Oak Ridge National Laboratory provided technical review and support for this report and associated review.

- Deborah Counce
- Kathy Jones
- Kevin Stewart

ABBREVIATIONS AND ACRONYMS

AMS	annual maximum series
ARF	areal reduction factor
ARR	Australian Rainfall & Runoff
ASOS	Automated Surface Observing System
AWOS	Automated Weather and Observing System
CFR	Code of Federal Regulations
CFSR	Climate Forecast System Reanalysis
CoCoRaHS	Community Collaborative Rain, Hail, and Snow Network
CONUS	contiguous United States
COOP	Cooperative Observer Program
CPC	Climate Prediction Center
DAD	depth-area-duration
DPA	Digital Precipitation Array
ESP	early site permit
FAA	Federal Aviation Administration
GDC	General Design Criterion/Criteria
GEV	generalized extreme value
GHCN	Global Historical Climatology Network
GPM	Global Precipitation Measurement
H&H	hydrologic and hydraulic
HUC	hydrologic unit code
HUCac	accumulated HUC
JAXA	Japan Aerospace Exploration Agency
JRA-55	Japanese 55-year Reanalysis
MERRA-2	Modern-Era Retrospective analysis for Research and Application, Volume 2
MPE	multisensor precipitation estimator
MRMS	Multi-Radar/Multi-Sensor System
NARR	North American Regional Reanalysis
NASA	National Aeronautical and Space Administration
NCAR	National Center for Atmospheric Research
NCEI	National Centers for Environmental Information
NCEP	National Centers for Environmental Prediction
NERC	(UK) Natural Environment Research Council
NEXRAD	Next-Generation Radar
NOAA	National Oceanic and Atmospheric Administration
NPP	nuclear power plant
NRC	Nuclear Regulatory Commission
NSE	Nash-Sutcliffe efficiency
NWS	National Weather Service

ORNL	Oak Ridge National Laboratory
PFA	precipitation frequency analysis
PFHA	Probabilistic Flood Hazard Assessment
PMF	probable maximum flood
PMP	probable maximum precipitation
PRISM	Parameter-elevation Regressions on Independent Slopes Model
QC	quality controlled
QPE	Quantitative Precipitation Estimates
RFC	River Forecasting Center
RMSE	root-mean-square error
TMPA	TRMM Multi-Satellite Precipitation Analysis
TP-29	Technical Paper No. 29
TP-40	Technical Paper No. 40
TR-24	Technical Report NWS 24
TRMM	Tropical Rainfall Measuring Mission
UK	United Kingdom
U.S.	United States
VIC	Variable Infiltration Capacity

1 INTRODUCTION

1.1 Background and Motivation

Extreme precipitation and subsequent flooding play major roles in infrastructure design and engineered systems operation. For probabilistic flood hazard assessment (PFHA), probabilistic estimates of extreme rainfall depth across various durations (e.g., T -year d -hour rainfall depth) are key inputs for hydrologic and hydraulic (H&H) modeling. Such estimates are quantified through precipitation frequency analysis (PFA) based on long-term rain gauge observations. To avoid going through the entire chain of PFA (including rain gauge data collection and processing, annual or partial duration maxima searching, probabilistic density function selection and fitting, goodness-of-fit test, and regionalization), H&H engineers have often opted to look up pre-calculated T -year rainfall depths from existing PFA products such as the U.S. Weather Bureau Technical Paper No. 40 (TP-40; Hershfield, 1961) or its successor, the National Oceanic and Atmospheric Administration (NOAA) Atlas 14 (Bonnin et al., 2004 and other volumes). Spatiotemporal distribution and area adjustments are then applied to form rainfall hyetographs for H&H modeling application.

However, one key feature that is easily overlooked is that most of the PFA products (including NOAA Atlas 14) provide frequency estimates of “point” precipitation, meaning that the results are representative only for a small domain and are not directly appropriate for large-scale watershed modeling applications. This happens because the annual (or partial duration) maxima of each rain gauge, which are the main inputs for PFA, are usually identified independently in time. Therefore, although they capture the local maximum precipitation, they do not jointly represent the areal maximum for a watershed. For instance, for watersheds governed by small-scale convective storm systems, the annual maximum precipitation captured by each gauge may be individually high but may result from different local thunderstorm events occurring at different times. In directly using the gauge-based extreme rainfall estimates to build design rainfall hyetographs for the entire watershed (without further adjustment), one may end up creating an unrealistically large rainfall depth that is not supported by historic observations.

To properly identify areal extreme rainfall estimates for watershed-scale applications, one needs to either (1) perform watershed-specific PFA that is based on precipitation maxima searched for the entire watershed (rather than for individual rain gauges) or (2) use a suitable precipitation areal reduction factor (ARF) to perform conversion (reduction) from point-based extreme rainfall estimates to areal-based extreme rainfall estimates. Although the watershed-specific PFA approach may theoretically be more precise, it involves much larger labor and resource efforts that are not always feasible. On the other hand, the ARF approach can provide quick estimation of areal extreme rainfall and hence has been a popular alternative in many H&H applications.

From a methodological perspective, the PFA methods (for either point-based or watershed-specific assessments) have been more extensively studied. Since the publication of TP-40 in 1961, widely-recognized frequency analysis methods have been developed, including the L-moment approach (Hosking and Wallis, 1992) that was used in the development of NOAA Atlas 14. For ARF, while new methods have been developed in recent decades (see Section 3.2.2), research efforts in the United States (U.S.) are lagging, and there has yet to be a national study of ARF comparable to NOAA Atlas 14. Some new ARF methods have been developed to make use of modern spatiotemporal rainfall data (such as weather radar), while some still rely upon rain gauge networks. However, none have found widespread acceptance and use in H&H engineering practice.

With H&H engineering applications and PFHA efforts requiring watershed-scale probabilistic precipitation estimates, there is a need to better understand how ARF methods can be applied to more effectively leverage existing point-based PFA products (such as NOAA Atlas 14). These considerations motivate the study described in this report.

1.2 Nuclear Regulatory Context

The U.S. Nuclear Regulatory Commission (NRC) has developed regulations regarding the siting and design of nuclear power plants (NPPs) which consider various natural hazards, including flooding. *Code of Federal Regulations* (CFR) Title 10, Part 50 and Part 52, address design criteria for NPPs with respect to natural hazards, whereas 10 CFR Part 100 addresses siting criteria. Title 10 CFR Part 50, Appendix A, General Design Criterion (GDC) 2, “Design bases for protection against natural phenomena,” states that structures, systems, and components important to safety shall be designed to withstand the effects of natural phenomena that have been historically reported for the site and surrounding area, with sufficient margin for the limited accuracy, quantity, and period of time in which the historical data have been accumulated. The regulation also states that the design bases shall reflect appropriate combinations of the effects of normal and accident conditions with the effects of the natural phenomena.

Title 10 CFR Part 52, more specifically 10 CFR Part 52.17(a)(1)(vi), for early site permits (ESPs) and 10 CFR Part 52.79 (a)(1)(iii) for combined licenses provide the requirements for new reactor applications as they relate to the hydrologic characteristics of the proposed site. These regulations require consideration of the most severe of the natural phenomena that have been historically reported for the site and surrounding area and with sufficient margin for the limited accuracy, quantity, and period of time in which the historical data have been accumulated. The requirements to consider physical site characteristics (including hydrologic features) in site evaluations are specified in 10 CFR Part 100.10(c) for applications before January 10, 1997, and 10 CFR Part 100.20(c) for applications on or after January 10, 1997.

NRC regulatory guidance for flood hazard assessments currently focuses on using deterministically derived, conservative estimates of key flood-causing mechanisms (e.g., probable maximum precipitation [PMP], probable maximum flood [PMF]) to provide the “sufficient margin” called for in the regulations. The magnitude of the provided margin is not explicitly quantified in either a physical or risk perspective. Probabilistic treatment of flood hazard phenomena can provide quantitative estimates of the flood safety margin and thus contribute to the risk-informed assessment of flooding hazards, but regulatory guidance on the use of probabilistic methods for riverine flood hazard assessment at NPP sites is lacking.

1.3 Areal Reduction Factors

An ARF can be generally defined as the ratio of areal extreme rainfall depth (i.e., total observed rainfall volume across an area divided by the area of interest) to point-based extreme rainfall depth (i.e., observed rainfall depth at a point location or for a representatively small area). For instance, for a watershed A of interest, given that P_W is the areal average extreme rainfall depth estimate for the entire watershed (can be derived from watershed-specific PFA) and P_{pt} is the representative point-based extreme rainfall depth within the watershed (can be determined from existing PFA products), the ARF can be formulated as

$$ARF = \frac{P_W}{P_{pt}}, \quad (1)$$

Since the value of P_W should be always smaller than that of P_{pt} , the value of ARF should be less than 1. Also, when the watershed area is very small, it is expected that the value of P_W will approach P_{pt} and hence ARF will have an upper bound of 1 when a watershed area approaches 0.

The use of ARF is necessary because networks of rain gauges with long periods of record, which are needed for accurate areal rainfall frequency estimation, are generally sparse and do not allow for an appropriate characterization of the associated spatial rainfall patterns. If the ARF relationship is known, point-based PFA products and ARF can be used to approximate areal extreme rainfall for a watershed. This approach has been used in many hydrologic applications, including stormwater management (e.g., MGNDCT, 2012; CWCB, 2006) and dam safety assessment (e.g., USBR, 2004). In the U.S., the most commonly used ARF chart was published by the U.S. Weather Bureau in the five-part Technical Paper No. 29 (TP-29; U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960). The same ARF chart (Figure 1-1) was provided for all TP-29 regions with durations ranging from 30-minute to 24-hour durations and watershed sizes up to 1,036 km² (400 mi²).

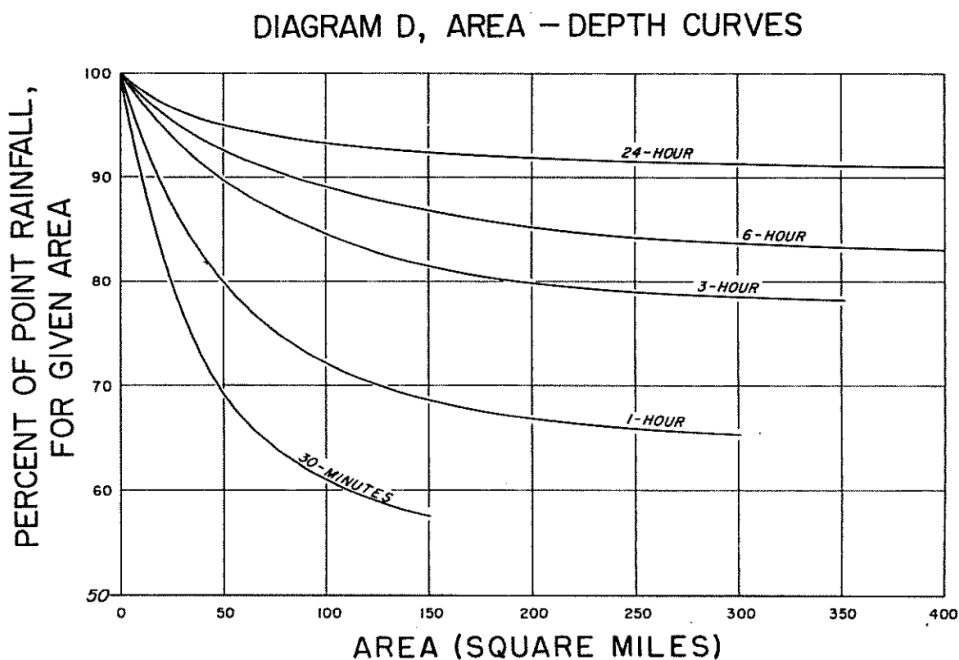


Figure 1-1 TP-29 ARF Curves (Source: U.S. Weather Bureau, 1957)

Although the TP-29 ARF chart was published decades ago, it is still used in many engineering applications. Considering current data, methodology, and numerical standards, this practice is problematic from several perspectives:

- The TP-29 ARFs are based on very limited data (approximately 20 rain gauges from a dense network with an average record length of about 11 years). Newer, high-resolution spatial rainfall observations have not been incorporated to update TP-29.
- The TP-29 ARFs are available only for relatively small areas (less than 1,036 km² [400 mi²]). Many NPP watersheds are far larger than this threshold.

- The TP-29 ARFs do not vary with geographic location (i.e., the same ARFs are used across the five regions of TP-29). This condition implies that the same ARF values can be applied regardless of local climate conditions, which is inconsistent with findings from other follow-up ARF studies.
- The TP-29 ARFs do not vary with return period, which is inconsistent with findings from many follow-up ARF studies.
- The TP-29 ARFs do not vary with season, which is questionable since the controlling extreme rainfall processes in a region can be quite different across warm and cool seasons.

Clearly, there is a need to update ARFs based on improved data and methods. With the advance of recent observational precipitation products and computational capabilities, more comprehensive ARF evaluations can be made to understand how ARF estimates vary across different methods, data sources, geographical locations, return periods, and seasons. These topics are explored in this study.

1.4 Fixed-Area versus Storm-Centered ARFs

In the context of PFHA, the goal of ARF is to bridge point-based and areal-based probabilistic extreme rainfall estimates. Therefore, the annual (or partial duration) maximum precipitation samples used in ARF analysis should also be searched following a similar approach to the maxima for PFA (i.e., precipitation maxima searched with a given spatial-filter and d -hour temporal window at a specific geographic location). Under this approach, the identified maximum samples represent the highest precipitation depth within a d -hour window, in which it may be a subset of a long-term rainfall event or composed of multiple sequential short-term rainfall events. In any case, these maxima represent the heaviest precipitation observed within a d -hour window at a selected location of interest, which is important for forming the most critical design events for H&H modeling application. In other words, ARFs for use with precipitation frequency estimates are normally referred to as “fixed-area” (or “geographically-fixed-area”) methods. Further formulation and discussion of maximum precipitation searching through the “fixed-area” approach are provided in Section 4.3.1 of this report.

To avoid potential confusion, it should be clarified that the “fixed-area” ARFs are distinct from “storm-centered” ARFs that are developed based on the analysis of individual storm events. The storm-centered ARFs are usually seen in storm-based, deterministic PMP analysis. Through the analysis of major historic storms, the maximum average rainfall depths across various durations and storm areas are identified; then they are presented in the form of depth-area-duration (DAD) tables or curves. By calculating the depth ratios between different cells in the DAD table, the storm-centered ARFs can be formed.

While both types of ARFs are used in flood hazard assessments, because of their different methodological features, their areas of application are distinct; thus, they should not be mixed. Because of a lack of formal PFA considerations, it is challenging to assign frequency and risk estimates to storm-centered ARFs. Numerically speaking, the values of storm-centered ARFs are smaller than those of fixed-area ARFs (Sivapalan and Blöschl, 1998; Svensson and Jones 2010); hence, applying storm-centered ARFs in PFHA (without further adjustment or treatment) will lead to underestimation of areal extreme precipitation estimates.

Given the specific objective to explore issues associated with ARFs for future PFHA applications, this research effort focuses only on fixed-area ARFs.

1.5 Research Objective

The objective of this study is to assist NRC in assessing different classes of fixed-area ARF methods in conjunction with available rainfall data sets to support the development of guidance for application of NPP PFHA. As mentioned in Section 1.2, NRC regulatory guidance for flood hazard assessments is currently deterministic in nature and does not enable explicit quantification of flood safety margin. In contrast, probabilistic approaches to flood hazard assessment could contribute to the risk-informed assessment of flooding hazards. However, regulatory guidance on the use of probabilistic methods for riverine flood hazard assessment at NPP sites is lacking.

This research project is part of the NRC's PFHA Research plan.¹ The work will aid the development of guidance on the use of PFHA methods and support informing NRC's licensing framework regarding risk (flood hazard design standards at proposed new facilities, as well as significance determination tools for evaluating potential deficiencies related to flood protection at operating facilities). The tools and guidance developed will support and enhance NRC's capacity to perform thorough and efficient reviews of license applications and license amendment requests. They will also support risk-informed significance determination of inspection findings, unusual events, and other oversight activities.

This report summarizes the overall findings from this ARF project. The report is organized into the following sections:

- Section 2 provides a summary of available precipitation data products that can potentially be used for ARF assessment.
- Section 3 provides a critical review of available ARF methods with a view to addressing the deficiencies in the commonly used methods for PFHA.
- Section 4 demonstrates the use of the promising method/data set combinations for ARF estimation using case studies for selected river basins.
- Section 5 provides the overall conclusion and recommendations for the development and application of ARF in future PFHA studies.

It should be emphasized that the main purpose of this research project is to improve the overall understanding of ARFs to support broader NPP PFHA efforts. However, the goal is not to provide national ARF estimates for direct applications. In the authors' view, the application of ARF for flood hazard assessment requires further site- or watershed-specific considerations, and it is not within the scope of this project to provide application-ready ARFs for a wide variety of potential usages. The discussion of specific references, methods, software, or tools in this report also does not constitute an endorsement or approval for any specific use by the NRC or Oak Ridge National Laboratory (ORNL). The case study results presented herein are the result of research efforts only, do not incorporate uncertainty quantification, and should not be directly

¹ The NRC PFHA Research Plan (Version 2014-10-23) is available online at <https://www.nrc.gov/docs/ML1429/ML14296A442.pdf>

incorporated for application. They are intended to demonstrate some of the primary factors affecting areal reduction factor estimation.

2 PRECIPITATION PRODUCTS FOR ARF ANALYSIS

In this section, various existing precipitation products that can potentially be used for ARF analysis are reviewed and discussed. Each precipitation product is evaluated based on its availability, accuracy, spatiotemporal resolution, latency, and suitability for ARF assessment. Several precipitation products are subsequently selected for comparison in the use case studies as described in Section 4.

2.1 Overview of Precipitation Products

The U.S. government began taking organized weather observations through the establishment of the Army-operated Weather Bureau in 1870. Following the deadly rain-induced South Fork Dam failure near Johnstown, Pennsylvania, Congress established the weather service in 1890 as a civilian agency by transferring the meteorological duties of the Army Signal Service to the newly created Weather Bureau in the Department of Agriculture. Soon thereafter, the nation's largest and oldest weather network, the Cooperative Observer Program (COOP), was established in 1891. At that time, more than 2,000 weather stations were recording observations by volunteers; today more than 11,000 COOP volunteers record weather observations across the country. The current National Weather Service (NWS) was established in 1970 as part of the NOAA within the Commerce Department (National Ocean Service, 2018; National Weather Service, 2017a; National Weather Service, 2018).

Over the years, various technology advancements have enabled increased spatial and temporal precipitation observation coverage in the U.S. Standard rain gauges have been installed across all U.S. states and territories. During World War II, military radar operators who had noticed returned echoes from precipitation began developing operational weather radars. Further development of radar in the 1940s provided a way to measure weather-related phenomena across wide areas using a single device or network of devices, and the first radar designed specifically for meteorological use was unveiled in 1954. In the late 1950s, weather satellites were developed and launched, enabling long-distance meteorological observations made from orbit. In the 1990s, meteorological reanalysis data sets were first produced to derive meteorological estimates based on incomplete observational data and simulated atmospheric processes; such reanalysis products have been used primarily for climate assessments. Additional information and detailed descriptions of gauge, radar, satellite, and meteorological reanalysis products are provided in Section 2.2.

2.1.1 Categories of Precipitation Products

Regarding the use of the term “data,” while the precipitation products described herein are developed based on various information types and methodologies, many products do not provide direct precipitation measurement. Gauge observations provide direct precipitation measurement; however, radar, satellite, and reanalysis products include interpretative precipitation estimates computed through various methods. Although data sets are developed for each precipitation product, it is important to note the difference between direct precipitation measurements and indirect estimates.

In this report, precipitation products are organized into five different categories:

- Gauge-only: direct rain gauge observations (for more information, see Section 2.2.1)

- Gauge-driven: gridded estimates produced from a series of rain gauges (for more information, see Section 2.2.2)
- Radar-driven: estimates produced from a radar network, often merged with gauge estimates (for more information, see Section 2.2.3)
- Satellite-driven: estimates produced from satellites, often merged with gauge estimates (for more information, see Section 2.2.4)
- Reanalysis-driven: reconstructed historical weather using global or regional weather forecasting models (for more information, see Section 2.2.5)

2.1.2 Key Metrics for Consideration

To develop areal precipitation frequency estimates and/or ARF, various metrics may be useful in considering the quality of available precipitation products. Among the most important metrics are the following:

- Accuracy/precision: How reliable are the precipitation estimates available from the product, and what sources of error and uncertainty exist? (See the product-specific sections later in this report for more information on product error and uncertainty.)
- Temporal coverage: For what time period are the precipitation estimates available, and are there any gaps in temporal coverage?
- Data latency: How regularly are the precipitation estimates uploaded online?
- Spatial coverage: For what regions are the precipitation estimates available?
- Temporal resolution: How frequently are precipitation estimates provided?
- Spatial resolution: For what horizontal spacing or area size are individual precipitation estimates available?

Ideally, to derive reliable ARF and/or areal extreme rainfall estimates through PFA, one would use precipitation products with the lowest measurement uncertainties (such as gauges), longest records (preferably more than 30 years of records), highest spatial density and largest spatial coverage (to capture extreme rainfall depth across various storm areas), and hourly or subhourly temporal resolution (to capture local intense precipitation). However, such criteria are unlikely to be satisfied universally across all parts of the U.S. Depending on data availability, controlling extreme rainfall types, and the selected ARF method in a watershed of interest, a preferred precipitation product or a mix of various products may be required.

2.2 Summary of Available Precipitation Products

For a detailed comparative summary of the available precipitation products available in the U.S., please see Appendix A.

Each of the following subsections provides a general overview, metric-based description, and error and uncertainty discussion of the available precipitation products.

2.2.1 Gauge-only Precipitation Data Sets

Gauge-based precipitation observations are collected from a variety of sources across the U.S. Most measurements are made using either standard 8-inch (20-cm) diameter non-recording rain gauges or Fischer-Porter recording rain gauges. Standard rain gauges include a funnel emptying into a graduated cylinder positioned inside a larger container for overflow. Fischer-Porter rain gauges mechanically convert the water weight into precipitation depth and record every 15 minutes. Tipping buckets are also used to automatically record precipitation accumulations. Some gauges may include heating capabilities to melt frozen precipitation for accurate measurement (National Weather Service, 2017b; Kuligowski, 1997). Wind shields are sometimes installed in locations where wind effects that reduce gauge catch cannot be reduced by site selection.

2.2.1.1 Product Description

For all U.S. states and territories, NOAA National Centers for Environmental Information (NCEI)² provide hourly and 15-minute precipitation gauge data (items 1 and 2 in Table A-1). At the time of this report's preparation,³ observations since 2013 had not yet been publicly released for both data sets. The hourly data set (DSI-3240) includes measurements from more than 7,000 NWS, Federal Aviation Administration (FAA), and COOP stations in the U.S. and U.S. territories, with data available since 1940 for some stations. The 15-minute data set (DSI-3260) includes measurements from more than 3,600 NWS, FAA, and COOP stations in the U.S. and U.S. territories, with data available since 1970 for some stations. The hourly and 15-minute stations mostly consist of Automated Weather and Observing System (AWOS) units, primarily managed by the FAA, and Automated Surface Observing System (ASOS) units, managed cooperatively by the NWS, FAA, and Department of Defense.

NOAA NCEI also provides a Global Historical Climatology Network (GHCN) database consisting of more than 100,000 stations reporting daily precipitation worldwide (item 3 in Table A-1). Data are available since 1870 for some stations, and the station data are updated every month. Data in the U.S. have been collected from FAA; COOP; the Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS); and other daily data sources.

2.2.1.2 Error and Uncertainty

In general, gauged data are considered “ground truth” for precipitation estimates. The three NOAA NCEI data sets represent actual observed measurements that may offer the most accurate precipitation estimates at point locations. Gauge observations have historically been used as the main input in PFA products (including NOAA Atlas 14). However, since the data collected are point measurements, they may not be representative of a region and are thus commonly used to produce gridded products or to calibrate radar products.

Aside from tipping error, wetting loss, and potential mechanical failure, the most apparent error with gauged data is undercatch, which may occur during measurement, especially during windy or snowy conditions (Sieck et al., 2007; Mekonnen et al., 2015). A comprehensive review of possible sources of rainfall observation uncertainty is presented by McMillan et al. (2012). Overall, existing literature suggests undercatch of gauges mounted at a 1 m height could have

² Formerly the National Climatic Data Center (NCDC) before dissolving in 2015.

³ Data accessed in February 2019.

5–16% error on average and 0–75% error per storm. Wind field deformation may cause 2–10% error for rain and 10–50% error for snow, wetting loss can cause 2–15% error in summer and 1–8% error in winter, and tipping error per 1 mm of rain can be up to 10%, depending on gauge type and rain rate (see Table 1 of McMillan et al. [2012] for further information).

Unlike the DSI-3240 and DSI-3260 data sets, the GHCN data set pulls from CoCoRaHS. This network of volunteer observers represents a significant source of precipitation observations but is not subject to a strict quality control protocol. While volunteers are encouraged to complete training courses, it is not required. Limited automated, web-based checks and personnel checks are conducted to identify errors, but the nature of a volunteer network introduces increased error potential compared with a federally managed and quality-controlled network.

The distribution of gauged data in both time and space is inconsistent, making precipitation estimates for certain locations difficult. For example, gauge coverage over some of the mountains, deserts, and plains of the western U.S. and Great Plains is sparse and could increase uncertainty in spatially interpolating precipitation estimates. In addition, with the diversity of gauge types available, differences in instrument performance (e.g., undercatch) can affect results, and correction may be needed in merging data.

2.2.2 Gauge-driven Precipitation Products

Given the temporal and spatial variability among observing gauge locations, gauge-driven precipitation products provide a useful way to assimilate individual gauge data and provide gridded estimates.

2.2.2.1 Product Description

Gauge-driven precipitation products are gridded products that (primarily) process daily gauge-based data to form gridded estimates (items 4–9 in Table A-1). During the gridding process, topographical, orographical, or statistical adjustments are made in many of the products. Therefore, gridded precipitation products are not simple spatial interpolations across available gauge observations.

Daymet, maintained by ORNL, provides daily gauge-based gridded precipitation estimates for 1980–2017 throughout all of North America at a 1-km horizontal resolution (Thornton et al., 1997 and 2017; item 4 in Table A-1). Daymet currently offers the highest spatial resolution among all publicly available gauge-based gridded precipitation data sets in the U.S. Daymet estimates include topographical adjustment based on elevation. However, since Daymet did not fully resolve rain shadow barrier effects, it may overestimate precipitation in the mountainous regions (e.g., in many parts of the western U.S.). The Daymet data set is currently updated annually.

Another widely used, gridded precipitation product is the Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 1994) data set, produced by the Oregon State University. With a daily temporal resolution, PRISM has offered gridded precipitation estimates since 1981 for the lower 48 U.S. states (i.e., the Contiguous U.S. [CONUS]) at a 1/24° (~4 km) horizontal resolution (item 5 in Table A-1). PRISM brings a combination of climatological and statistical concepts to the analysis of orographic precipitation. Given PRISM's ability to account for topographical effects and some other orographic adjustment factors (Daly et al. 2002), it is one of the best available grid-based meteorological data sets. NOAA Atlas 14 also used the PRISM spatial interpolation algorithm in its assessment. The daily PRISM is a relatively recent

product of the widely used monthly PRISM data set, which has offered monthly gridded precipitation estimates since 1900 at the same 4-km horizontal spacing. The gridded daily PRISM data are updated automatically and are partially assimilated with radar data in its post-2002 daily precipitation output. An 800-m horizontal resolution data set is also available for a fee from the PRISM website.

The Livneh Daily CONUS Near-surface Gridded Meteorological Data produced by the University of Colorado offers daily gauge-based gridded precipitation estimates for 1950–2013 for the CONUS, Mexico, and the part of Canada south of 53°N at a 1/16° (~6 km) horizontal resolution (Livneh et al., 2015; item 6 in Table A-1). The Livneh product builds upon (and essentially replaces) the previous Maurer Gridded Meteorological Data product from Santa Clara University, which provides daily gauge-based gridded precipitation estimates for 1949–2010 for the CONUS and British Columbia in Canada at a 1/8° (~12 km) horizontal resolution (Maurer et al., 2002; item 7 in Table A-1). Both the Livneh and Maurer data sets include elevation-based topographical adjustments. They have been widely used as the meteorological input to drive the Variable Infiltration Capacity hydrologic model (VIC; Liang et al. 1994 and 1996) in many hydroclimatic studies (e.g., Bennett et al., 2018; Gutmann et al., 2014; Mizukami et al., 2017; Sheffield et al., 2006; Wood et al., 2004).

The NOAA Climate Prediction Center (CPC) produces the CPC Unified Gauge-based Analysis of Daily Precipitation over CONUS, which provides daily gauge-based gridded precipitation estimates for 1948–2006 for the CONUS at a 1/4° (~25 km) horizontal resolution (item 8 in Table A-1). Given its coarser spatial resolution, topographic and orographic adjustments were not applied. This can lead to underestimated precipitation in mountainous regions.

Compared with the previous gridded precipitation products (items 4–8 in Table A-1), the National Center for Atmospheric Research (NCAR) Gridded Ensemble Precipitation and Temperature Estimates over the Contiguous United States (Newman et al., 2015; item 9 in Table A-1)—available from 1980–2012 for the CONUS, northern Mexico and southern Canada at a 1/8° (~12 km) horizontal resolution—is a conceptually different product. Newman et al. (2015) expanded the concept of probabilistic interpolation by Clark and Slater (2006). Instead of providing only one set of most likely values (as is provided by other gridded products), Newman et al. (2015) provided a 100-member ensemble of historic precipitation in which each realization is embedded with spatially-correlated random signals to account for the uncertainties from measurements and other sources. Although each ensemble member is different, their collective ensemble mean is similar to those of other gridded precipitation data sets. This data set can be easily used for ensemble hydrologic simulation to understand how uncertainties may propagate through different steps of H&H modeling.

2.2.2.2 Error and Uncertainty

For gridded gauge-driven precipitation products, any error or uncertainty in the gauge data propagates to the gridded applications. Spatial smoothing may help reduce some of the errors or uncertainties presented from a single station, but such benefits have not been quantified.

One of the major differences among the various gridded products (also a source of uncertainty) is how precipitation is spatially distributed over complex terrain. PRISM spatially distributes precipitation using precipitation/elevation regressions. Daymet spatially distributes precipitation through an iterative station density algorithm. Livneh spatially distributes precipitation over complex terrain using a satellite-based estimate of peak snow water equivalent. Maurer spatially distributes precipitation using PRISM. An intercomparison of gridded precipitation data sets

covering PRISM, Daymet, Livneh, and Newman in the western U.S. by Henn et al. (2017) suggested that the greatest absolute differences in annual total precipitation occurred in maritime mountain ranges and high-elevation areas of the Western U.S. (200 mm/year or greater on average, around 5–60%). Oubeidillah et al. (2014) compared the mean and 95% percentiles of PRISM, Daymet, and Maurer and suggested that higher-spatial-resolution data sets (i.e., PRISM and Daymet) performed better than the coarser-resolution data sets, particularly in capturing precipitation extremes.

2.2.3 Radar-driven Precipitation Products

While early weather radar systems (e.g., WSR-57 and WSR-74) enabled detection of precipitation in the atmosphere, a major breakthrough came through the development and deployment of Doppler radar. The Next-Generation Radar (NEXRAD), WSR-88D Doppler radar, was developed in the 1970s and 1980s and achieved first operational use in 1992. Beyond detecting the position of precipitation, Doppler radar also captures movement toward or away from the radar by sending horizontal radio waves and interpreting the shift in response (NCEI, 2018; Rinehart, 1997). Another major breakthrough came with the development of dual-polarization Doppler radar, in which both horizontal and vertical waves are sent. The addition of vertical wave transmittal provides information on the vertical motion of particles and helps distinguish differences in precipitation type (e.g., rain, hail, or snow). Dual polarization was added to existing WSR-88D radars starting in 2010, and by 2013 all NEXRAD radars were equipped with dual polarization. For a more detailed description of dual-polarization radar, see Cifelli and Chandrasekar (2013). The NEXRAD system currently comprises 160 sites throughout the United States and select overseas locations (NCEI, 2018).

2.2.3.1 Product Description

A series of data post-processing steps are required to convert the measured radar reflectivity into estimated rainfall depth. Multiple radar-driven precipitation products are available and include varying levels of post-processing and/or bias correction. These products are introduced below and presented in order from lowest to highest complexity.

The most fundamental radar-driven rainfall products are NEXRAD Level-II and Level-III. The NCEI provides NEXRAD Level-II and Level-III data for most of the United States (item 10 in Table A-1) and U.S. territories. Measurements are updated automatically and are made every 4 to 10 minutes. The WSR-88D radars originally provided radar reflectivity at 1.0 degrees azimuthal by 1 km range gate resolution to a range of 460 km, and Doppler velocity and spectrum width at 1.0 degree azimuthal by 250 m range gate resolution to a range of 230 km. In 2008, the WSR-88D radars were upgraded with increased spatial resolution (referred to as “super resolution”) to provide radar reflectivity at 0.5 degrees azimuthal by 250 m range gate resolution to a range of 460 km, and Doppler velocity and spectrum width at 0.5 degree azimuthal by 250 m range gate resolution to a range of 300 km. Level-II data contain initially processed (base) data at high resolution and are essentially raw data. These data include reflectivity, radial velocity, spectrum width for single-polarization radar and the addition of differential reflectivity, co-polar correlation coefficient, and differential phase for dual-polarization radar. Level-III data include further processed data at a lower resolution deemed more appropriate for common use. Over 75 Level-III products are routinely provided by NCEI, and the precipitation-specific data include 1-hour, 3-hour, and total storm precipitation. Both Level-II and Level-III products are derived solely from NEXRAD radar observations (NCEI, 2018).

To support operational river forecasting, the NOAA/NWS National Centers for Environmental Prediction (NCEP) and River Forecasting Centers (RFCs) further conduct a series of data processing efforts to increase the accuracy of radar-driven rainfall estimates. These efforts were originally structured into four consecutive stages (Stage I to Stage IV), although the structure has been revised with the advancement of procedures and technologies. Some terminologies are still in use now but with adjusted meanings (NOAA, 2018):

- **Stage I** referred to the radar-only digital precipitation arrays (DPAs) that use radar reflectivities to estimate rainfall depth on 4 km, 131×131 polarstereographic grids centered on individual radar sites. Note that one main difference between Level-II/III and Stage I products is the different spatial grids.
- **Stage II** referred to the merged data of Stage I and automatic rainfall gauge observations (available since 1996 in the CONUS; item 11 in Table A-1). In the past, this process was carried out at each of the 12 CONUS RFCs, as well as at the NCEP, and was used as input for Stage III. With the change of process from Stage III to quantitative precipitation estimates (QPE) at RFC, the meaning of Stage II has changed. Currently, Stage II refers to the multisensor product that has not undergone quality control (QC), which is generated directly from the radar and gauge data at NCEP (i.e., NCEP Stage II).
- **Stage III** referred to the 1-hour and 6-hour analyses conducted by RFCs based on RFC-mosaicked Stage II with manual QC (available since 2002 in all 14 RFCs; item 12 in Table A-1). Currently most RFCs have transitioned from Stage III to multisensor precipitation estimator. Western RFCs have transitioned to Mountain Mapper, and Arkansas-Red Basin RFC is using a local bias adjustment algorithm. The regional 1-hour/6-hour estimates provided by RFCs to NCEP are collectively called QPE.
- **Stage IV** referred to the final radar-driven rainfall product produced at NCEP (available since 2002 in the CONUS, excluding the California-Nevada and Northwest RFCs; item 13 in Table A-1). It is based on the QPEs (after manual QC) provided by the RFCs. NCEP merges all RFC QPEs into a national NCEP Stage IV product.

The Multi-Radar/Multi-Sensor System (MRMS) produced by the NOAA National Severe Storms Laboratory provides automatic 2-minute resolution precipitation estimates based on a series of algorithms integrating radar and gauge data with lightning detection systems and forecast models (item 14 in Table A-1). The information produced is mostly used for extreme weather forecasting and other purposes, and historical simulations are not publicly archived. The CONUS and adjacent portions of Canada and Mexico are covered by MRMS.

2.2.3.2 Error and Uncertainty

Although radar can capture the spatial distribution of extreme storms that cannot be measured by conventional gauges, the accuracy of radar-based precipitation products can be limited by nonlinear reflectivity-rainfall relationships, variations in vertical reflectivity, blockages, and spatial and temporal sampling (AghaKouchak et al., 2010). Because of some coverage gaps and limitations of radar penetration in mountainous terrain, radar data are not available in some parts of the U.S. Krajewski et al. (2010) compared radar rainfall estimates and rain gauge observations for 20 selected storms and reported around 15–91% average differences. From a statistical sampling perspective, unlike rain gauges that continuously measure cumulative precipitation at a fixed location, radar rainfall is based on measurements of instantaneous

reflectivity during each scan (currently every 4–10 minutes). The intrinsic assumptions are that the scan interval is sufficiently short, and each instantaneous scan can be representative for the entire interval. Nevertheless, given the high variability of extreme storms, such assumptions involve high uncertainties, and hence further adjustments and assimilation using gauge observation are required. An evaluation performed by Gourley et al. (2010) suggests that the Stage IV product has the highest correlation coefficient to gauge observations among various gridded rainfall products. Cunha et al. (2013) find that dual-polarization radars generally provide lower error than single-polarization radars, but that the error is fundamentally dependent on range sampling; radar rainfall error is found to decrease for larger temporal and spatial scales. Seo et al. (2015) also find dual-polarization to provide higher accuracy than single-polarization.

2.2.4 Satellite-driven Precipitation Products

Satellites have been used to observe weather phenomena for decades, with temperature measurements first recorded in the late 1960s and precipitation measurements becoming available in the 1990s. Precipitation detection is accomplished through the use of radar, microwave imaging, and lightning sensors.

2.2.4.1 Product Description

In 1997, a joint effort between the National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) launched the Tropical Rainfall Measuring Mission (TRMM) satellite equipped with the first orbiting precipitation radar. Before ending its mission in 2015, the TRMM orbited at an elevation of approximately 400 km and completed an orbit every 92.5 minutes. The TRMM Multi-Satellite Precipitation Analysis (TMPA) merges the data collected from TRMM with data from other satellites, gauges, and other sources to produce 3-hourly precipitation estimates at a 1/4 deg (~25 km) horizontal resolution (item 15 in Table A-1). Global coverage is provided from 60 degrees S to 60 degrees N. Although the TRMM went offline in 2015, TMPA products were still produced through early 2018.

NASA and JAXA jointly launched TRMM's replacement, the Global Precipitation Measurement (GPM) international satellite mission, in 2014 (item 16 in Table A-1). GPM provides global coverage from 60 degrees S to 60 degrees N. Compared with TMPA, GPM provides higher spatial and temporal resolution, offering 30-minute precipitation accumulations at a resolution of 0.1 deg (~10 km).

2.2.4.2 Error and Uncertainty

Given satellites' broad spatial coverage, TRMM is a valuable resource for regions without a dense gauge network or weather monitoring radars. However, the use of satellite-based precipitation products is not advantageous in many parts of the U.S. given the existence of various gauge- or radar-based precipitation products.

From a statistical sampling perspective, satellite-based precipitation products also suffer from a similar but more severe measurement issue compared with weather radar. Though using a multi-satellite approach can increase measuring frequency, significant inconsistencies and gaps in temporal coverage exist which may decrease accuracy compared with radar and gauge estimates. Several studies have reported that systematic error in satellite measurements increases as the rain rate increases (AghaKouchak et al., 2012). Petkovic and Kummerow (2017) evaluate the source of error in passive TRMM microwave imager sensors over certain regions. Tian and Peters-Lidard (2010) present a global map of measurement uncertainty in

satellite-based precipitation estimates, finding higher uncertainty nearer the poles. An evaluation performed by Gourley et al. (2010) suggests that Stage IV provides more accurate precipitation measurements than TRMM. The use of satellite-based precipitation products is most beneficial for regions with a minimum number of rain gauges and without reliable radar measurement, such as parts of the mountainous western United States.

2.2.5 Reanalysis-driven Precipitation Products

Meteorological reanalysis products have been widely used by the meteorological, climatological, and hydrological communities in understanding historic weather patterns. Meteorological reanalysis provides reconstructed historic weather simulations using global weather forecasting models with observations collected from various instruments (but not rainfall depth from rain gauges until recently; see Reichle et al., 2017). The simulations provide comprehensive snapshots of 3-dimensional meteorological conditions (e.g., wind, pressure, total precipitable water) at regular intervals over long time periods, often years or decades (Parker, 2016). There have been multiple international efforts to generate reanalysis data sets to support various missions.

2.2.5.1 Product Description

NOAA NCEP has produced several precipitation reanalysis data sets, including NCEP/NCAR Reanalysis 1 (item 17 in Table A-1), NCEP/Department of Energy Reanalysis 2 (item 18 in Table A-1), North American Regional Reanalysis (NARR; item 19 in Table A-1), and Climate Forecast System Reanalysis (CFSR; item 20 in Table A-1). Reanalysis 1 offers 6-hourly global precipitation estimates at a 1.875° horizontal resolution for the period since 1948 and is updated daily. Reanalysis 2 provides similar information, is available from 1979, and is updated monthly. CFSR provides similar global information and coverage as Reanalysis 2 but includes an enhanced 0.313° (~35 km) horizontal resolution and has started to incorporate gauge rainfall observations in its data assimilation. NARR is a specific use case of the high-resolution NCEP Eta Model to provide 3-hourly precipitation estimates over North America at a 32 km resolution; data are available from 1979 to 2014, and no update is currently scheduled.

NASA produces the Modern-Era Retrospective analysis for Research and Application, Version 2 (MERRA-2; item 21 in Table A-1), which offers global hourly precipitation estimates at a 0.5° latitude by 0.625° longitude resolution and has been available since 1980. Compared with NCEP reanalysis, NASA reanalysis uses more remotely sensed information courtesy of NASA satellites. MERRA-2 has also started to incorporate gauge rainfall observations in its data assimilation system.

The Japan Meteorological Agency produces the Japanese 55-year Reanalysis (JRA-55; item 22 in Table A-1), which offers global 3-hourly precipitation estimates at a 55-km horizontal resolution and is available since 1958.

The European Centre for Medium-Range Weather Forecasts produces the ERA-Interim reanalysis product, which offers global 6-hourly precipitation estimates at a 0.7° (~78 km) horizontal resolution and is available from 1979 (item 23 in Table A-1). ERA-Interim will be replaced in 2019 by ERA5, which will offer 31-km global resolution (compared with 79-km in ERA-Interim), hourly data (compared with 6-hourly in ERA-Interim) and improved atmospheric parameterization.

2.2.5.2 Error and Uncertainty

Among the various types of precipitation products summarized in this report, reanalysis-driven precipitation is most different from rain gauge observations, mainly because two factors: (1) reanalysis data sets typically have very coarse spatial resolution and (2) reanalysis data sets typically do not assimilate observations from rain gauges. Recent reanalysis products such as CFSR and MERRA-2 start to incorporate gauge observations to correct model generated precipitation (Reichle et al., 2017). Despite the relatively poor utility of reanalysis-based precipitation, other reanalysis variables such as total precipitable water and 3-dimensional wind and pressure fields are considered reliable and sometimes have been treated as surrogates of observations. A comprehensive global-scale evaluation of 23 precipitation data sets (satellite- and reanalysis-based) by Beck et al. (2017) reported large differences in estimation accuracy among the data sets and highlighted the benefits of careful data merging across gauge-, satellite- and reanalysis-based precipitation estimates.

2.3 Considerations for ARF

In this section, a total of 23 publicly available precipitation products based on gauge, radar, satellite, and reanalysis data are described. Data metrics (including data set type, temporal and spatial coverage, and latency) for these products are tabulated in APPENDIX A for further ARF analysis consideration.

To reliably derive areal extreme rainfall and ARF estimates through PFA, precipitation products with low measuring uncertainty, long records, high spatial density, large spatial coverage, and fine temporal resolution are desired. In general, the gauge-only precipitation data sets should have the highest (point) measurement accuracy. In addition, since the gauge-only precipitation data sets are also the common input to most point-based PFA products (including NOAA Atlas 14), they may provide the best data consistency for ARF derivation. However, spatial disaggregation of gauge precipitation measurements into watersheds is not a trivial task since one needs to consider various topographic, orographic, and local adjustments (especially in complex terrain). The processing of gauge observations (e.g., format transform, quality control, treatment of missing values) is also more cumbersome than other approaches. As a result, the use of gauge-only precipitation data sets likely demands more labor and resources compared with other alternative precipitation data products.

Compared with gauge-only precipitation data sets, several gauge-driven precipitation products are easier to use and quite popular in many hydrologic studies. Despite some differences in methodologies, these products all consider potential topographic effects on precipitation. However, although they are all (mostly) based on gauge observations, different data processing and quality control procedures (that are not fully open to the public) are used in each product. As a result, it is challenging to conduct further review and understand their inter-product differences. Most gauge-driven precipitation products are also limited to daily temporal resolution and hence cannot support subdaily ARF analysis.

Given the advance of monitoring technologies, both radar-driven and satellite-driven precipitation products provide new opportunities to better understand the spatiotemporal structures of extreme precipitation. While both types of products can better capture the instantaneous and relative magnitude of extreme storms, they cannot provide temporally continuous and accurate measurement of precipitation depth at a given location. As a result, assimilation and bias-correction of raw radar-driven and satellite-driven precipitation products by gauge observations are necessary. When both types of products are available, it is also

recommended to first consider radar-driven rather than satellite-driven precipitation products. In addition, given their relatively shorter periods of record, radar- and satellite-based products may not be used to credibly estimate long-return-period extreme rainfall (e.g., it is questionable to estimate 100-year rainfall depth from 16 years of data). The limited period of data collection is of particular concern for the purpose of PFHA.

Among all types of precipitation products, reanalysis-driven precipitation estimates can be the most limiting. Various studies have reported large differences between reanalysis-driven and observed precipitation. Although the reanalysis products can provide a suite of best-available meteorological conditions (e.g., wind, pressure, total precipitable water), they do not seem to have the required accuracy to support PFA and ARF derivation. Therefore, use of the reanalysis-driven precipitation in PFA is not recommended (unless all other alternative products are unavailable).

Since no existing precipitation product can meet all desired data criteria consistently across all of the U.S., region-specific considerations are needed. To explore and demonstrate the sensitivity of ARF associated with precipitation products, several precipitation products are selected for comparison in the use case studies documented in Section 4.

3 METHODS FOR ARF ANALYSIS

This section reviews and discusses various methods that can potentially be used for ARF analysis. Each method is evaluated by its data dependency, required assumptions, analytical complexity, spatiotemporal scale, and whether it has been independently evaluated in other studies. Several suitable ARF methods are then selected and compared in the use case studies in Section 4.

3.1 Factors Affecting ARF

We start by summarizing the main factors that affect ARF in Table 3-1. These factors are classified by storm characteristics (Section 3.1.1), geographic features (Section 3.1.2), ARF methodology (Section 3.1.3), and source of rainfall data (Section 3.1.4). Recent review articles (e.g., Svensson and Jones, 2010; Pietersen et al., 2015; Pavlovic et al., 2016) provide insights into how these factors could qualitatively or quantitatively affect ARF. Discussion, along with comments on the relevance to PFHA at NPP sites (included as indented, bulletized text throughout Section 3.1), are provided in the remainder of this section.

Table 3-1 Summary of Factors Affecting Precipitation ARF Calculations

Storm Characteristics (Section 3.1.1)	Geographic Features (Section 3.1.2)	ARF Methodology (Section 3.1.3)	Sources of Rainfall Data (Section 3.1.4)
<ul style="list-style-type: none"> • Storm duration • Storm type • Seasonality • Return period 	<ul style="list-style-type: none"> • Geographic location • Regional climate pattern • Watershed geometry • Watershed topography 	<ul style="list-style-type: none"> • Empirical versus analytical • Storm-centered versus fixed-area • Spatial averaging • Probabilistic distributions 	<ul style="list-style-type: none"> • Types of rainfall products • Period of record • Data integration • Measurement uncertainty

3.1.1 Storm Characteristics

Storm characteristics are perhaps the most sensitive factors affecting ARF. Pietersen et al. (2015) identified predominant weather types, storm durations, seasonal factors, and recurrence intervals as primary contributors to inconsistencies across multiple ARF findings.

Storm Duration: In addition to storm area, storm duration is one of the most-commonly reported factors affecting ARF. Many conventionally used ARFs (e.g., TP-29) reported higher ARF values (closer to 1) at longer durations (1 day or above) and lower ARF values at shorter durations (subdaily). However, some studies (e.g., Huff, 1995; Clark and Rakhecha, 2002; Ramos et al., 2005) reported minimal impact of storm duration on ARF, although such studies analyzed only limited storm durations.

- For the purpose of PFHA, longer-duration ARF is likely to be the focus for NPPs located in large watersheds. Other NPPs located in smaller watersheds or in watersheds with short lag times may require shorter-duration ARF. Depending on the specific modeling

needs (e.g., for watershed-scale flooding or for local intense precipitation), ARF with the most suitable duration should be derived to support PFHA applications.

Storm Type: Extreme precipitation events are usually classified into larger-scale tropical and synoptic (e.g., frontal) systems and smaller-scale convective (e.g., thunderstorm) systems that have distinct spatial rainfall patterns. Skaugen (1997) identified spatial correlation structures for small- and large-scale precipitation using statistical pattern recognition of daily rainfall parameters and found that precipitation intensity for small-scale convective events decreases more rapidly than for large-scale frontal systems (i.e., the ARF would decay more rapidly for small-scale events). However, since storm type can be addressed only through the less conservative storm-centered approach, it is typically not called out in most fixed-area assessments. Storm type can instead be indirectly represented through storm area and duration, with larger area/duration ARF primarily covering tropical/synoptic systems and smaller area/duration ARF primarily covering convective systems. In addition, certain storm types occur more commonly and would be more highly represented in the historical record in certain geographic locations. For instance, convective systems are more frequent in the Midwest, while tropical and extra-tropical storm remnants are more common in the Southeast and Northeast, respectively. Since intense, small-scale convective events typically occur during warm seasons (e.g., May–October), seasonal assessment can be another way to indirectly assess the effects of storm type on ARF.

- Existing literature does not offer an effective way to quantify the effects of storm type on ARF when a fixed-area assessment approach is used. Employing storm type classification on historic events is a tedious and somewhat subjective process which may not yield specific benefits for ARF derivation. Given that the effects of storm type can be indirectly incorporated through storm area and duration and vary by region, for the purpose of PFHA on NPP sites, assessment of effects of storm type on ARF may not be a priority. If needed, a hybrid approach can be developed to specifically address the effects of storm type on ARF.

Seasonality: Huff and Shipp (1969) and Allen and DeGaetano (2005b) both report smaller ARFs associated with the warm season compared with the cool season. Svensson and Jones (2010) reasons that this finding likely results from increased convective activities present in the warm season. Various literature suggests that increased convection (which is associated with concentrated, high-intensity precipitation over relatively small areas) could explain why warm season ARFs are smaller than cool season ARFs. For regions with significant seasonal rainfall variability, the effects of seasonality on ARF can be expected.

- In considering the impacts of all season precipitation events versus cool season rain-on-snow events on an NPP site of interest, calculation of season-specific ARFs may be needed for hydrologic applications. The (default) all season ARF is likely to be smaller than cool season ARF and would lead to an underestimation of ARF during cool season applications. Thus, a dedicated cool season ARF is likely needed.

Return Period: Strong evidence exists in the literature to suggest that ARF varies with return period. Bell (1976) reported that ARFs decrease more rapidly for increasing return periods, an observation which is also found in Stewart (1989), Skaugen (1997), Allen and DeGaetano (2005a), and Asquith and Famiglietti (2000). However, perhaps because of data limitation, both U.S. Weather Bureau TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960) and United Kingdom (UK) Flood Studies Report (NERC, 1975) ARFs disregard the potential effects due to return period, while Grebner and Roesch (1997) evaluate return period and find ARFs to

be independent of return period in Switzerland. Using more recent radar-based rainfall product, Pavlovic et al. (2016) show that ARF derived for the State of Oklahoma are dependent upon return periods.

- Although return period was not represented in the conventional TP-29 ARF, its potential influences should not be neglected and are certainly relevant to PFHA at NPP sites. However, the main challenge arises in estimating long-return-period (e.g., greater than 100-year level) ARF estimates based on limited historic observations (i.e., decades of data). To overcome the data limitation, a possible approach in assisting the development of long-return-period (i.e., low frequency) ARF is through multi-ensemble simulation using numerical weather simulation models. Modern high-performance-computing capabilities have gradually matured enough to enable examining the areal-point extreme precipitation relationship through computationally intensive, process-based modeling. Until such progress has been made to improve the reliability of low-frequency extreme precipitation depth, the ARF estimates at long return periods solely based on observations should be treated with extreme caution.

3.1.2 Geographic Features

Geographic features such as location, regional climate pattern, watershed geometry, and topography affect both point and areal precipitation features as well as their ARF relationship. However, the U.S. Weather Bureau TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960) and United Kingdom Flood Studies Report (NERC, 1975) ARFs were usually applied across various locations without considering the potential geographical differences. This is an area that can be improved in PFHA for NPP sites.

Geographic Location: Comparisons among ARFs in different geographic locations have supported the concept that ARFs vary with location. For example, Omolayo (1993) found higher 1-day ARFs in the U.S. than in Australia; Asquith and Famiglietti (2000) found higher ARFs in the eastern U.S. than in Texas; and Zehr and Myers (1984) found more rapid ARF decline in the Southwest U.S. than the rest of the country. Bell (1976) and Stewart (1989) both found very weak correlations between ARF and latitude, indicating that ARFs may be higher at more northerly latitudes. In addition, the frequency and intensity of certain storm types varies across locations. As Skaugen (1997) points out, extreme point precipitation tends to occur more frequently inland, whereas large, synoptic-scale precipitation events tend to occur closer to the coast.

- When local data are sufficient, the most defensible approach will be to derive site-specific ARF that can best reflect the geographical influences for a site of interest. If the preference is to use ARF from other published studies, the selection should consider if geographical features are sufficiently similar to avoid erroneous ARF values for applications.

Regional Climate Pattern: Since precipitation patterns are fundamentally tied to regional climate, ARF studies have commonly reported (mostly qualitatively) the impact of regional climate on ARFs. As mentioned earlier, the results of Omolayo (1993), Asquith and Famiglietti (2000), and Zehr and Myers (1984) all indicate variations in ARF with geographic location; however, these findings also indicate lower ARFs in drier climates (e.g., Australia, Texas, Southwest U.S.) compared with wetter climates (U.S., eastern U.S.). Stewart (1989) shows that ARFs were correlated with long-term average annual rainfall, indicating a connection between climate and ARF.

Microclimates from urban rainfall effects may also affect ARFs. Huff (1995) evaluates ARFs for storms occurring in urban and rural areas, finding that precipitation may decrease more slowly in urban areas than in surrounding rural areas for a 500-km² (193-mi²) area and may decrease more rapidly in urban areas than in surrounding rural areas for larger areas. The study, however, was based on only 8 urban storms in Chicago and 67 rural storms in the surrounding rural region, and Huff (1995) reported that the anomaly could result from natural variation.

- Although it is clear that regional climate pattern has a direct linkage to ARF, there has not been an effort to include climate variables (e.g., annual precipitation) in the numerical ARF representation. This can be worth exploring in future research efforts, and the results should help in deriving a more generalized ARF model for regions without sufficient data to derive site-specific ARF.

Watershed Geometry: In computing areal average rainfall (and hence ARF), the geometry of the watershed can play a role. For example, if typical rainfall patterns (e.g., shape and movement) vary greatly from the watershed geometry, observed areal rainfall characteristics may vary. This effect could be most pronounced for irregular or elongated catchment shapes. However, Veneziano and Langousis (2005) found that the effect of watershed geometry is generally small and that highly elongated watersheds are rare.

- Although watershed shape and geometry may affect ARF, the influence should be relatively minor and can be reasonably neglected for most NPP sites. The potential impact may still need to be evaluated if the shape of a contributing watershed is highly irregular or elongated.
- Note that in considering ARFs for a large watershed, ARFs specific to multiple sub-watersheds may be sought for hydrologic modeling application. In such cases, it is important to ensure that the total mass balance of rainfall estimated for the full area of interest is reflected in the sub-watershed ARF estimates. To be more specific, the overall watershed-wide ARF should be governed by the total area of the watershed. If the sub-watershed ARFs are evaluated for each of the hydrologic modeling units (based on the individual area of each sub-watershed), these sub-watershed ARFs should be further rescaled to match the overall watershed-scale ARF. In other words, since the aggregation of ARFs applied within multiple smaller sub-watersheds would increase the total precipitation volume compared with application of the overall watershed ARF, adjustments would be needed to ensure conservation of total rainfall mass.

Watershed Topography: Topography plays an important role in precipitation processes through various orographic effects, and the frequency of particular storm patterns may change with topography. For example, Stewart (1989) reports a relationship between ARF and long-term average annual rainfall; however, a strong correlation between long-term average annual rainfall and elevation also was found. Thus, topography could have contributed to the relationship found between ARF and long-term average annual rainfall. Gauged precipitation observations are typically sparser in high-elevation locations, partially because of lower population densities. From a calculation standpoint, using low-density, high-elevation observations to calculate areal precipitation could prove unrepresentative of the region being considered; however, Allen and DeGaetano (2005b) find little variation in ARF based on gauge density.

- To ensure that the potential influence of topography is properly accounted for in ARF, it is important to assess whether the selected precipitation product has been adjusted for topography (e.g., PRISM and Daymet). The conventional Thiessen polygon spatial

averaging approach does not account for the potential influence due to topography and can lead to biased areal precipitation estimates in topographically complex regions. As with regional climate pattern, the quantitative influence of topography on ARF can be worth exploring in future research efforts, and the results should help in deriving a more generalized ARF model for regions without sufficient data to derive site-specific ARF.

3.1.3 ARF Methodology

While multiple ARF methods have been developed since TP-29, there is no consensus guidance on which newer ARF methods may be superior (or preferable). Most studies were proposed and conducted for specific regions. Intercomparison studies such as Pietersen et al. (2015) and Pavlovic et al. (2016) are needed for more locations and to cover more potential methods. Further research could guide ARF model selection by providing quantifiable metrics for comparison among alternative ARF methods. The general classification and considerations are discussed below with further method-specific review provided in Section 4.

Empirical versus Analytical: Empirical ARF refers to the regionally smooth ARF relationship based on a large number of samples (derived from pairs of gauges or from smaller sub-areas) without an underlying analytical theory. Most ARFs used in practice (e.g., U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960; NERC, 1975; Bell, 1976) fall into the empirical category. Several newer analytical approaches have been developed using various methodologies, including correlation analysis, crossing properties, scaling relationships, and storm movement.

- To select an appropriate ARF model for application, both empirical and analytical ARF calculation approaches should be considered and compared for the purpose of PFHA at NPP sites. A hybrid method (i.e., fitting parameters of an analytical model by site-specific empirical samples) offers another promising approach.

Fixed-area versus Storm-centered: The more commonly used empirical ARF approaches (e.g., U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960; NERC, 1975; Bell, 1976) follow a fixed-area approach in which the ARF calculation domain is geographically fixed. In contrast, some other studies calculate ARF using a storm-centered approach in which the area analyzed changes for each storm, while the point precipitation for ARF calculation is based on the maximum rainfall observed. Asquith and Famiglietti (2000) point out that one challenge associated with storm-centered ARF approaches is difficulty with handling multi-centered storms. Omolayo (1993) report that storm-centered ARFs are incorrect for estimating areal precipitation frequency from point precipitation observations. However, Omolayo (1993) and Svensson and Jones (2010) acknowledge that they can be used for PMP studies since there is no frequency associated with PMP. From a conservativeness standpoint, storm-centered ARFs have been shown to be less conservative than fixed-area ARFs (Sivapalan and Blöschl, 1998; Svensson and Jones 2010).

- Although storm-centered ARF relationships have been used in PMP studies (in the form of DAD curves), those are in a different context and should not be used to derive areal rainfall frequency estimate. Furthermore, given concerns with the appropriateness of storm-centered ARF approaches (Omolayo, 1993) and application-related challenges (e.g., multi-center storms), storm-centered ARF is not recommended for PFHA at NPP sites. As highlighted in Section 1.4, this study focuses only on the fixed-area ARF approach.

Spatial Averaging: In spatially aggregating gauge rainfall observations to estimate areal precipitation, various methods may be used. Most commonly, simple unweighted areal averages, Thiessen polygons, and inverse distance weighting are used to compute areal precipitation. In topographically complex regions, additional adjustments based on topography need to be considered during spatial averaging. The case studies by Allen and DeGaetano (2005a) found small differences in ARFs derived by using simple unweighted areal averages, Thiessen polygons, and inverse distance weighting, but also acknowledge that these simpler interpolation methods do not account for topography. Pavlovic et al. (2016) suggest that the spatial averaging method has a significant influence on the ARF; hence, they eventually use radar-based data instead of gauge rainfall in their Oklahoma study.

- The potential effect of spatial averaging on ARF is likely affected by the homogeneity of precipitation extremes within a region. As discussed previously with respect to watershed topography, spatial averaging methods that cannot address topographical influences (e.g., Thiessen polygons) would likely result in biased areal rainfall estimates in topographically complex regions. This issue can be further challenged by storm type considerations and gauge density. In regions with very few gauge stations, small-scale convective storms may not be fully captured, leading to underestimation of extreme storms. Therefore, while simple unweighted areal averages can be adequate for spatially homogeneous regions, various weighting methods or meteorological data assimilation approaches should be compared and evaluated in highly heterogeneous regions.

Probabilistic Distributions: Some ARF methods rely on various types of probabilistic distributions (e.g., Rodriguez-Iturbe and Mejía, 1974; Bacchi and Ranzi, 1996; Grebner et al., 1998; De Michele et al., 2001). These distributions are used to model point extreme rainfall processes, areal dependence structure, or both. For instance, Sivapalan and Blöschl (1998) assume that point rainfall would follow an exponential distribution, whereas the areal average rainfall would follow a gamma distribution. Overeem et al. (2009) and (2010) proposed a regional generalized extreme value (GEV) distribution that captures both point and areal probability distribution. Consequently, typical issues associated with frequency analysis—such as annual maximum series (AMS) processing, fitting techniques (e.g., maximum likelihood), and goodness-of-fit tests—would affect the derivation of ARF.

- Since the main purpose of ARF within the PFHA framework is to estimate areal PFA estimates from point-based PFA products that have been derived through rigorous frequency analysis, the specific requirements and consideration of probability distribution fitting and testing should be quite familiar to the users. For a targeted NPP watershed, one may follow the concept of conventional PFA to test and select a most suitable probabilistic distribution to model areal extreme precipitation (i.e., a specific distribution such as GEV is not necessary). However, while it is important to evaluate the type of distribution to be used, it is also critical to assess whether there are sufficient statistical samples to support parameter estimation.

3.1.4 Data

The final category relates to the rainfall data used to estimate ARF. Factors include types of rainfall products, period of record, data integration, and measurement uncertainty.

Types of Precipitation Products: The conventional ARFs (e.g., U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960; NERC, 1975; Bell, 1976) were developed using gauge rainfall observations; however, the increased availability of other rainfall products (e.g., radar-based

rainfall estimates) have enabled additional accessibility for ARF calculation. Durrans et al. (2002) report radar-based 100-year return period estimates to be 20–35% lower than gauge-based estimates found in TP-40 (Hershfield, 1961) and HYDRO-35 (Frederick et al., 1977). Short radar records, heterogeneous data treatment, natural climate variability, and radar data calibration have all been identified as potential contributors to the differences in radar-based and gauge-based extreme rainfall estimates (Svensson and Jones, 2010). In calculating ARF, however, these differences are generally assumed to cancel out, since point and areal estimates are subject to the same types of biases. Durrans et al. (2002) report that while radar-based ARFs are largely consistent with earlier gauge-based ARFs, they do not decrease with area as rapidly as do gauge-based ARFs; however, Allen and DeGaetano (2005b) report the opposite conclusion. Other rainfall products, such as those from satellite observations or reanalysis, could also be evaluated but have not been commonly applied to ARF calculation. For example, Kok et al. (2017) evaluate the feasibility of using satellite observations to compute ARFs in Malaysia, while Fouchier et al. (2015) evaluate probabilistic ARF in France using a reanalysis product. Pavlovic et al. (2016) state that the daily ARF curves obtained using the radar-based Stage IV and gauge-based PRISM match each other closely.

- Given the availability of a variety of precipitation products (as discussed in Section 2), the required effort for intensive gauge data processing is largely reduced. However, this also raises a new concern regarding how the differences among existing precipitation products may affect ARF. When a situation allows, both gauge and radar-based precipitation products can be evaluated and compared to check how their differences may affect the derived ARF values. The satellite- and reanalysis-driven precipitation products, on the other hand, would likely involve larger uncertainty. ARF computed using those data may be less accurate and should be applied with caution.

Period of Record: Since radar-based precipitation products are relatively new, the short period of record represents a major limitation for computing radar-based ARFs (Durrans et al., 2002). Satellite-based precipitation products also suffer from a relatively short period of record. Although gauge observations benefit from a longer period of record, the periods among gauges may be inconsistent and are more difficult to process and analyze.

- The limited period of record would prohibit reliable ARF estimation for long return periods. If long-return-period ARF estimates are sought, the maximum number of data should be collected for analysis.

Data Integration: Asquith and Famiglietti (2000) evaluate the aggregation of multiple overlapping rain gauge networks near Houston, TX, and conclude that “differing precipitation-monitoring networks cannot be indiscriminately combined.” In terms of station density, Allen and DeGaetano (2005a) find that differences in station density and interpolation method had minimal effect on ARF calculation in New Jersey and North Carolina. Moving forward, the integration among different precipitation products will likely receive increasing attention, although such an approach has yet to be used for ARF study.

- For data-limited regions, aggregation of multiple precipitation products will likely be one of the few viable approaches. However, potential issues resulting from data integration (specifically for ARF) are not clearly understood. ARF products created using integrated data sets should be carefully reviewed.

Measurement Uncertainty: The various precipitation products exhibit different measurement error and uncertainty, as described in Section 2. When applicable, preference should be given

to those precipitation products with relatively smaller measurement uncertainty. Nevertheless, since the ARF represents a depth ratio, it is usually assumed that the systematic measuring biases can be largely canceled out. Thus, the measurement uncertainty may not be as critical to other factors that affect ARF.

- Conventionally, hydrologic and hydraulic engineers have treated gauge observation (and increasingly, radar-based precipitation data) as ground truth values, often without considering its uncertainty. With the improved understanding of measurement uncertainty and availability of multiple precipitation products, measurement uncertainty can be quantitatively addressed in PFHA for NPP sites.

3.2 Critical Review of ARF Methods

3.2.1 Qualitative Assessment Considerations

In evaluating the suitability of an ARF method for PFHA, several key qualities should be considered, including the following:

- Spatiotemporal Scale and Resolution—Was the method developed at a comparable spatial scale (e.g., over 10,000 km²) to the targeted NPP watershed size and does it offer the desired temporal resolution (e.g., 3-day or 1-hour)?
- Data Sufficiency and Dependency—Was the method developed based on a sufficiently large and complete data set? Does the method heavily rely on the quality and quantity of the input data, and are those data readily available, accessible, and reliable? If data availability is an issue for a specific NPP site, does the method entail a sufficient process-based foundation to support reasonable implementation with minimum data input?
- Required Assumptions—Does the method require various assumptions that may have been oversimplified or impractical?
- Analytical Complexity—Is the method difficult to follow or to reproduce (e.g., requires excessive computational resources or cannot be easily applied using existing software packages or tool sets) so that it is challenging to implement or review?
- Independent Evaluation—Has the method been tested and shown to perform as expected through independent evaluation by a third party?
- Site-specific Transferability—Is the method applicable to the NPP site or watershed of interest or is its application beyond the original study limited by geographic or climatologic features?

These considerations are discussed and commented for various ARF methods reviewed in this section.

3.2.2 Available ARF Methods

Some of the currently available ARF methods are summarized in this section, including

- Empirical methods (Section 3.2.2.1)

- Spatial correlation methods (Section 3.2.2.2)
- Statistical crossing properties methods (Section 3.2.2.3)
- Spatial and temporal scaling methods (Section 3.2.2.4)
- Extreme value theory methods (Section 3.2.2.5)

A comprehensive summary of available ARF methods is also provided in Pietersen et al. (2015); additional information on ARF methods can be found in Svensson and Jones (2010), while an intercomparison among a few selected methods is provided in Pavlovic et al. (2016). The reviewed methods are summarized in Table 3-2 and discussed in the sections below.

Table 3-2 Summary of Available ARF Calculation Methods

Empirical Methods
US Weather Bureau (TP-29; 1957, 1958a, 1958b, 1959, 1960) Leclerc and Schaake (1972)
United Kingdom Approach (NERC, 1975; Bell, 1976) Koutsoyiannis and Xanthopoulos (1999)
National Weather Service (TR-24; Myers and Zehr, 1980)
Annual-maxima-centered (Asquith and Famiglietti, 2000)
Swiss Approach (Grebner et al., 1998)
Australian Rainfall and Runoff Approach (Nathan and Weinmann, 2016)
Spatial Correlation Methods
Rodriguez-Iturbe and Mejía (1974)
Sivapalan and Blöschl (1998)
Omolayo (1989)
Statistical Crossing Properties Methods
Bacchi and Ranzi (1996)
Spatial and Temporal Scaling Methods
De Michele et al. (2001)
Veneziano and Langousis (2005)
Extreme Value Theory Methods
Durrans et al. (2002)
Allen and DeGaetano (2005a)
Lombardo et al. (2006)
Overeem et al. (2010)

3.2.2.1 Empirical Methods

Empirical ARF methods originated from TP-29 by the U.S. Weather Bureau (1957, 1958a, 1958b, 1959, 1960). The following text describes TP-29 and various empirical ARF calculation methods that have since been developed.

U.S. Weather Bureau TP-29 Approach (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960)

The U.S. Weather Bureau TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960) represents the first assessment of ARFs and provides a series of curves to estimate areal rainfall based on point rainfall. ARFs were developed for 0.5-, 1-, 3-, 6-, and 24-hour durations for area sizes up to 1,036-km² (400-mi²; see Figure 3-1) and were assumed generalizable regardless of geographic location or return period. The TP-29 estimates were derived by analyzing six rain gauge networks in the U.S. with individual gauges nearly uniformly spaced and gauge records ranging from 7 to 15 years. These networks are located between 80 and 90 degrees W longitude and 35–40 degrees N latitude.

To compute ARF, the TP-29 approach relates the mean annual maximum areal rainfall to the mean annual maximum point rainfall for all stations and all years. The ARF is calculated for a specific duration as follows (formula modified from Pietersen et al., 2015):

$$ARF = \frac{N \sum_{i=1}^N \sum_{j=1}^n w_i \overline{P_{ij}}}{\sum_{i=1}^N \sum_{j=1}^n P_{ij}}, \quad (2)$$

where

N = number of stations within the catchment area

n = record length (years)

$\overline{P_{ij}}$ = point rainfall of station i coincident with the annual maximum areal rainfall in year j (mm)

P_{ij} = annual maximum point rainfall of station i in year j (mm)

w_i = Thiessen weighting factor for station i

During year j , P_{ij} represents the annual maximum point rainfall of station i which may occur at different timing across the year. On the other hand, $\overline{P_{ij}}$ refers to the single largest storm event during year j . A Thiessen weighting factor w_i is used to assign the importance of each station in which $\sum_{i=1}^N w_i = 1$. Since P_{ij} should always be greater than $\overline{P_{ij}}$, ARF calculated by Equation (2) has an upper bound at 1. Given the nearly uniform spatial distribution of the gauge stations, TP-29 assumes equal weighting among all gauge observations. For nonuniform rain gauge networks, Equation (2) can be revised to include a weighted factor for each point rainfall (see Pietersen et al. 2015).

DIAGRAM D, AREA – DEPTH CURVES

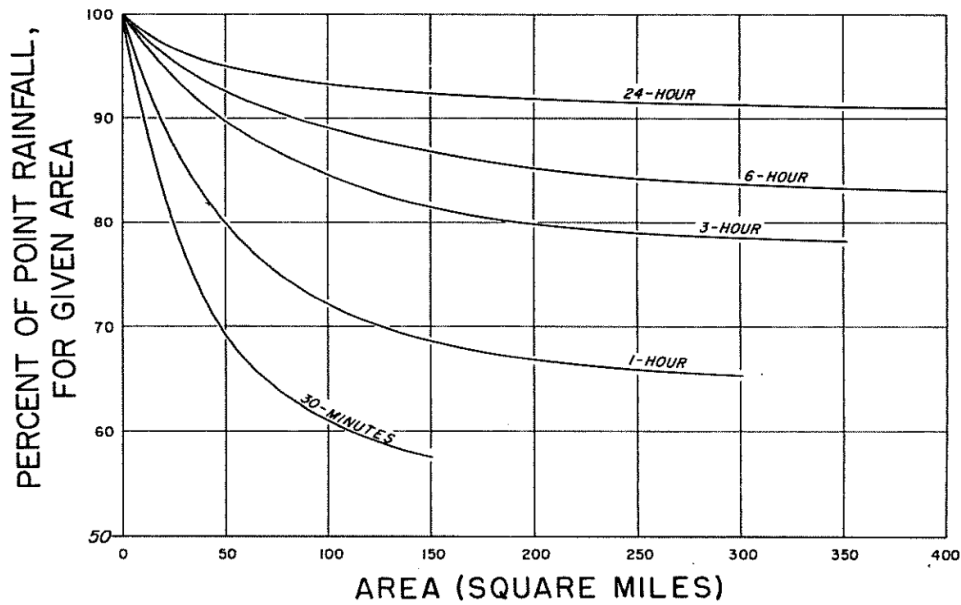


Figure 3-1 TP-29 ARF Curves (Source: U.S. Weather Bureau, 1957)

Although newer methods have been developed, TP-29 remains the most commonly used ARF calculation method applied in the U.S. Leclerc and Schaake (1972) fit the TP-29 ARF results using the following equation:

$$ARF(A, D) = 1 - e^{aD^b} + e^{(aD^b - cA)}, \quad (3)$$

where

A = area (km²)

D = duration (hours)

a, b, c = fitted parameters

United Kingdom Approach (NERC, 1975; Bell, 1976)

Essentially following the same approach as used in TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960), the United Kingdom Flood Studies Report (NERC, 1975) calculates ARF using a computationally convenient assumption that ARF can be calculated as the average of the area-to-point ratios computed from annual maximum events. The ARF is calculated as (formula from Pietersen et al., 2015)

$$ARF = \frac{1}{nN} \sum_{i=1}^N \sum_{j=1}^n \left(\frac{\overline{P_{ij}}}{P_{ij}} \right), \quad (4)$$

where

N = number of stations within the catchment area

n = record length (years)

$\overline{P_{ij}}$ = point rainfall for station i coincident with the annual maximum areal rainfall in year j (mm)

P_{ij} = annual maximum point rainfall of station i in year j (mm)

Compared with TP-29, the UK ARF covered a wider range of durations (1 minute to 25 days) and areas (1 km² to 30,000 km²).

An empirical equation representing the UK Flood Studies Report ARF curves was established by Koutsoyiannis and Xanthopoulos (1999). The ARF curve is calculated as

$$ARF(A, D) = 1 - \frac{aA^{(b-c \ln A)}}{D^d} \geq 0.25, \quad (5)$$

where

A = area (km²)

D = duration (hours)

a, b, c, d = fitted parameters

A graphical representation of the Koutsoyiannis and Xanthopoulos (1999) fitted model is shown in Figure 3-2. The UK ARF curves were plotted in semi-log scale to better illustrate small and large area ARFs simultaneously on the same figure. This plot formatting technique is different from the linear technique used in Figure 1-1 for the TP-29 curves and results in curves which are concave down rather than concave up.

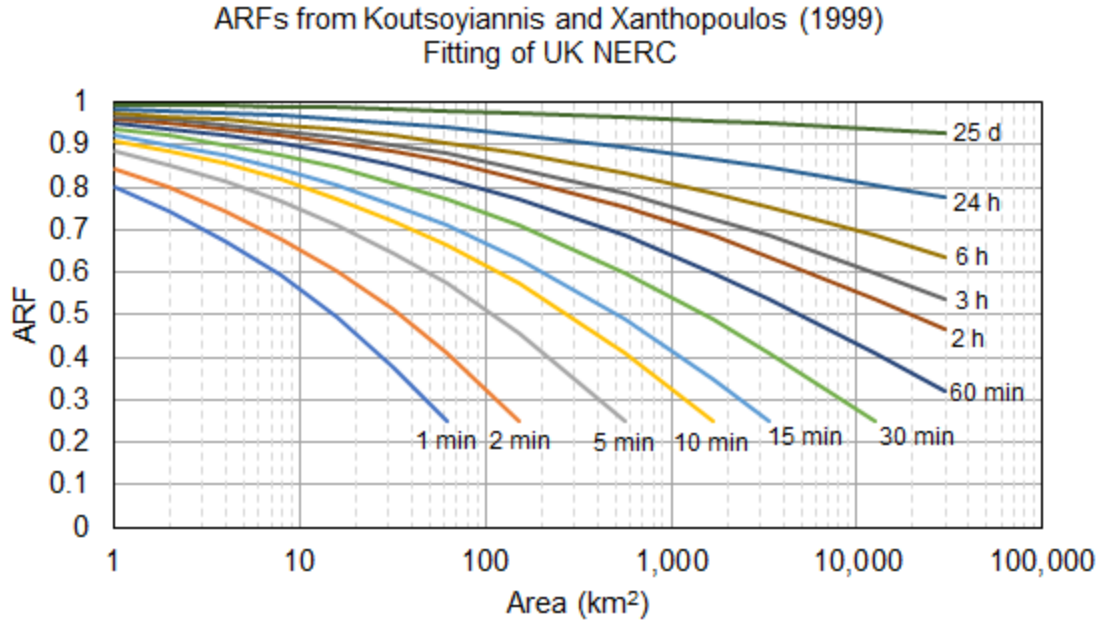


Figure 3-2 UK Flood Studies Report ARF Curves Based on Koutsoyiannis and Xanthopoulos Fitted Model

While the Flood Studies Report (NERC, 1975) assumed return period had little effect on ARF calculation, Bell (1976) reexamined the approach by fitting exponential distributions to the areal and point annual maximum rainfall series and computing ARF for areal and point rainfall estimates of the same return period. Circular areas of 1,000 km² (386 mi²) were evaluated with ARFs computed for up to a 24 h duration. Given spatial variability in gauge observation locations, Thiessen weighting was used to fit the annual maximum areal rainfall series. To better represent the area of consideration, the point rainfall frequency curve was fitted based on the Thiessen-weighted means of annual maximum point rainfall. Fitting to an exponential distribution using the method of maximum likelihood was performed using a partial duration series of the 20 highest events; this was performed separately for each sample area and for the point rainfall estimates. The ARF for the same return period was then calculated for return periods of 2–20 years. The ARF is calculated as follows:

$$ARF_T = \frac{\sum_{i=1}^N (w_i \overline{P_{ij}})_T}{\sum_{i=1}^N (w_i P_{ij})_T}, \quad (6)$$

where

ARF_T = areal reduction factor at return period T

T = Return period (year)

$\overline{P_{ij}}$ = point rainfall for station i coincident with the annual maximum areal rainfall in year j (mm)

P_{ij} = annual maximum point rainfall of station i in year j (mm)

w_i = Thiessen weighted factor for station i

Overall, the Bell (1976) results provide reasonable agreement with the NERC (1975) results, with specific additional findings including these:

- A slight tendency for 24-h ARFs to increase for higher latitude (less than 3% bias)
- A statistically significant tendency for ARFs to decrease for longer return periods (2–5% bias for 24-h duration and 5–15% bias for 1-h and 2-h durations).

National Weather Service TR-24 Approach (Myers and Zehr, 1980; Zehr and Myers, 1984)

NOAA Technical Report NWS 24 (TR-24; Meyers and Zehr, 1980) represents the current U.S. approach to ARF estimation, yet practitioners often still use TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960). TR-24 uses annual maximum rainfall across station pairs to perform frequency analysis and was developed using data from the Chicago area but deemed applicable nationwide. The frequency analysis is estimated following Chow (1951, 1964) and based on a Gumbel fitting of a Fisher-Tippet type I distribution. Svensson and Jones (2010) states that “it is questionable whether the complicated methodology [used in TR-24] is justified as precipitation observations become more plentiful with time.” Figure 3-3 shows TR-24’s complex calculation process. The ARF is calculated as follows (formula from Pietersen et al., 2015):

$$ARF = \frac{\overline{P}_A(f, \Delta t, A)}{\overline{P}_P(f, \Delta t, 0)}, \quad (7)$$

where

\overline{P}_A = average areal rainfall for a specific frequency (f), duration (Δt), and area (A) (mm)

\overline{P}_P = point rainfall for a specific frequency (f), duration (Δt), and area (A) (mm)

The subsequent publication of NOAA Technical Memorandum NWS HYDRO-40 (Zehr and Meyers, 1984) applied some modifications (by using a mixed statistical model) to account for the non-dense network of gauges in Arizona and western New Mexico. Compared with the previous national study, NWS HYDRO-40 reports larger decreases in ARFs for larger area sizes in the southwest U.S.

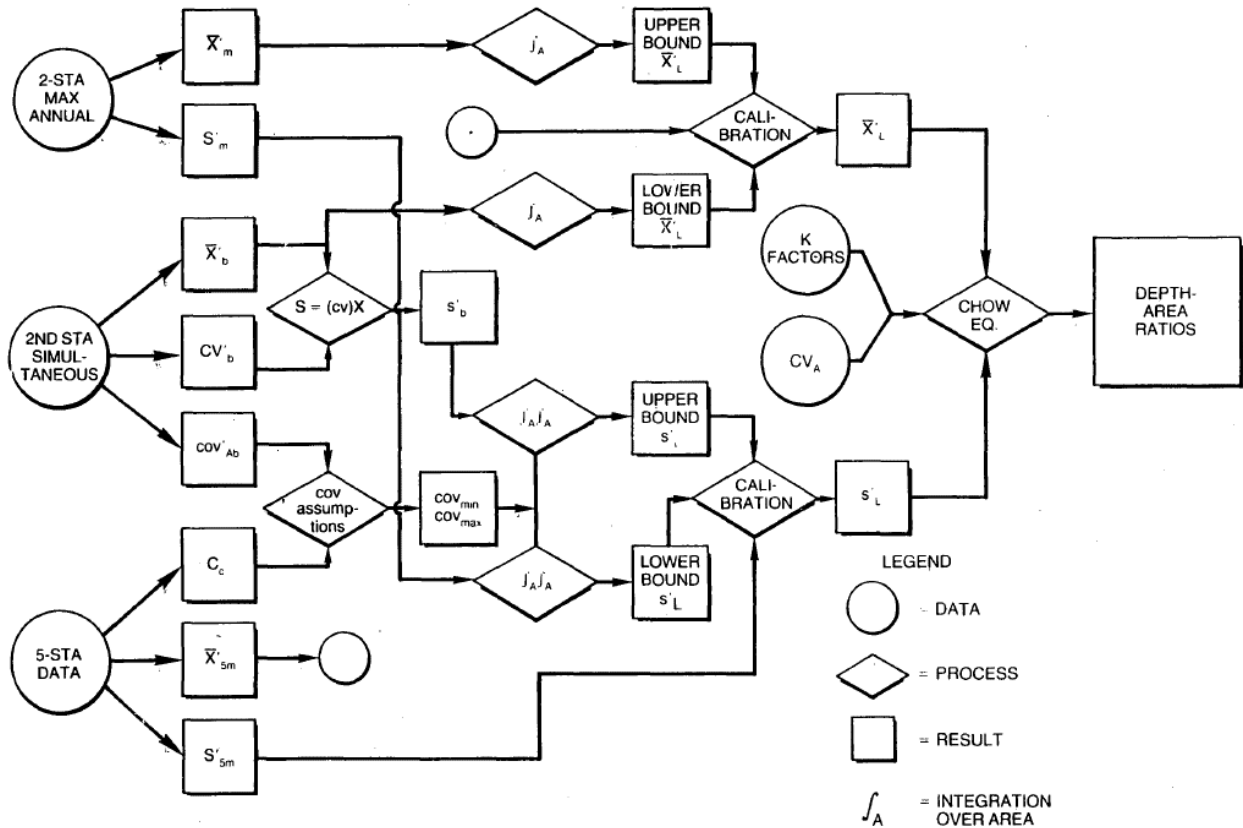


Figure 3-3 TR-24 ARF Calculation Process (Myers and Zehr, 1980)

Annual-maxima-centered Approach (Asquith and Famiglietti, 2000)

Asquith and Famiglietti (2000) follows a slightly different ARF calculation approach in which both the areal and point components of the ARF are calculated based on annual maximum point events; in this way, the approach focuses on specific events and represents a storm-centered approach. Once an annual maximum point precipitation is found, a pair-wise series of calculations are performed to find the ratio between each surrounding point precipitation and the target annual maximum point precipitation. These ratios, shown as $S_T(r)$ in Equation (8), are then plotted against the distance between each point pairing, and fitting is performed. The resulting areal estimate is thus calculated as a circular area. The ARF is calculated as follows (formula from Pietersen et al., 2015):

$$ARF = \frac{\int_0^R 2rS_T(r)\Delta r}{R^2}, \quad (8)$$

where

R = maximum radius of circular catchment or integration limit (km); the total circular area = πR^2

r = radius of concentric circle within the catchment (km)

$S_T(r)$ = ratio between rainfall depth at a specific location, distance r from the point of the design storm and annual maximum rainfall

This process is conducted separately for each return period. The approach also enables simple customized evaluations, as it does not require a fixed area and can be spatially integrated once the paired ratios are determined. This approach requires dense gauge networks and results in more rapidly decreasing ARFs than does TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960). As mentioned in Svensson and Jones (2010), since the approach centers around the annual maximum point events and therefore does not consider the annual maximum areal events, the areal estimate likely underestimates the ARF.

Swiss Approach (Grebner et al., 1998)

The Hydrological Atlas of Switzerland (Grebner et al., 1998) provides ARFs across eight geographic zones for durations of 3–72 hours and areas of up to 5,000 km² (1,931 mi²). Using hourly precipitation gauge data from 1981 to 1993, the 26 most intensive precipitation events per geographic zone and per duration class were summarized; thus, this approach is storm-centered. The ARF is then statistically fit to the following formula:

$$ARF = \frac{a_0}{(A+a_2)^{a_1}} + a_3e^{-a_4A}, \quad (9)$$

where

A = rainfall storm areas (km²)

a_0, a_1, a_2, a_3, a_4 = fitted parameters

When $A = 0$, ARF should be 1. Therefore, an additional equation can be obtained:

$$a_3 = 1 - \frac{a_0}{a_2^{a_1}}, \quad (10)$$

As do other studies, Grebner et al. (1998) find ARFs for longer durations are flatter (i.e., decrease less rapidly) than ARFs for shorter durations. ARFs are also reported to be independent of return period.

Australian Rainfall and Runoff Approach (Nathan and Weinmann, 2016)

The Australian Government's Australian Rainfall and Runoff (ARR) Guidelines (Nathan and Weinmann, 2016) provides a modified version of Bell's ARF calculation method. For areas with sufficient data, hypothetical circular catchments are defined, and areal rainfall series are developed using areal weighting (e.g., Theissen weighting). The ARF is then calculated from the

ratio of the areal weighted rainfall quantile divided by the weighted point rainfall quantile. After ARFs are calculated across various catchment areas, durations, and return periods for as many locations as possible, the ARFs are averaged across the catchments and fitted to provide a prediction model for the region of interest. Nathan and Weinmann (2016) provide detailed application guidelines for computing ARFs for catchments up to 30,000 km² (11,583 mi²) and for durations of up to 7 days. For application to small catchments, the authors recommend linearly interpolating ARFs between the equation-based 10-km² ARF and an ARF of 1.0 for 0 km².

A generalized equation for catchment areas of 10–1,000 km² and durations of 1–7 days (that may be suitable for PFHA at many NPP sites) is formulated as

$$\begin{aligned}
 ARF(A, D, AEP) = & 1 - a(A^b - c \log_{10} D)D^{-d} \\
 & + eA^f D^g (0.3 + \log_{10} AEP) \\
 & + h10^{iAD} (0.3 + \log_{10} AEP)
 \end{aligned} \tag{11}$$

where

A = area (km²)

D = duration (hours)

AEP = annual exceedance probability

a through i = fitted parameters

When $A = 0$, ARF should be 1. Therefore, an additional equation can be obtained to eliminate h as a fitted parameter:

$$h = - \frac{ac(\log_{10} D)D^{-d}}{0.3 + \log_{10} AEP} , \tag{12}$$

Overall, although empirically based methods require minimum analytical assumptions, they heavily rely on the amount and quality of the underlying data. The computational challenges are mostly for data processing and QC, but not so much for the implementation of a numerical ARF model. Also, given the empirical nature, empirically based ARF products may not be regionally transferable (nevertheless, many applications still use them in various distinct regions). These various empirically based ARFs are commonly used in practice across various countries.

3.2.2.2 Spatial Correlation Methods

Various ARF calculation methods are based on spatial correlation of rainfall fields and rely on assumptions of isotropy (i.e., spatial correlation does not vary significantly along a specific horizontal direction/orientation) and particular statistical distributions of the rainfall process.

Original Spatial Correlation Approach (Rodriguez-Iturbe and Mejía, 1974)

Rodriguez-Iturbe and Mejía (1974) represents the first assessment of ARFs using spatial correlation parameters. The ARF in this case is simplified to depend only on the correlation coefficient of rainfall data between any two point rainfall locations chosen at random from the study area, as follows:

$$ARF = \sqrt{E(\rho(d))}, \quad (13)$$

where

$E(\rho(d))$ = expected correlation coefficient for the characteristic correlation distance

The method uses all precipitation data rather than only considering extreme events; thus, the randomly chosen locations are likely more evenly distributed than locations associated only with more extreme events. The correlation distance measures the mean separation between two point locations selected randomly from within the area of interest. A spatial correlation structure, either an exponentially decaying function or a Bessel-type correlation structure, is assumed to fit the rainfall data, with additional assumptions that the point precipitation is isotropic (i.e., any location within the area of interest exhibits the same probability law) and Gaussian. Svensson and Jones (2010) point out that this distribution is not typical of extreme, short-duration precipitation events and that a non-Gaussian distribution will result in inexact correspondence between the point and areal precipitation frequencies (i.e., a theoretically correct ARF will not be represented).

Modified Spatial Correlation Approaches (Omolayo, 1989; Sivapalan and Blöschl, 1998)

Omolayo (1989) documented an ARF approach in which rainfall depths are assumed to be log-normally distributed in space. The calculated method produces ARFs that vary directly with spatial correlation coefficient and inversely with return period, standard deviation, and number of gauges. The full calculation is shown in Eq. (14), while Eq. (15) shows a reduced form when a normal distribution is assumed and Eq. (16) shows a further reduced form for when n is large (formulas from Pietersen et al., 2015). Equation (16) provides a form resembling that derived by Rodriguez-Iturbe and Mejía (1974); however, the Omolayo (1989) expression includes a correlation coefficient averaged over the rain gauges rather than being expressed for a particular separation distance. The following formulas are taken from Pietersen et al. (2015):

LN distributed rainfall:

$$ARF_1 = Exp \left\{ K_T \sigma \left[\sqrt{\frac{1 + (N - 1)\rho}{N}} - 1 \right] \right\} \quad (14)$$

Normal distributed rainfall:

$$ARF_2 = \sqrt{\frac{1 + (N - 1)\rho}{N}} \quad (15)$$

Normal distributed rainfall (large number of rainfall stations)

$$ARF_3 = \sqrt{\rho} \quad (16)$$

where

K_T = frequency factor corresponding to return period

N = number of rainfall stations

T = return period (years)

σ = standard deviation of rainfall depth in the log domain (mm)

ρ = Average spatial correlation coefficient

Sivapalan and Blöschl (1998) propose a modification to the Rodriguez-Iturbe and Mejía (1974) approach, noting that the latter approach considers mean areal average rainfall, which does not change with the averaging area. To address the concern, Sivapalan and Blöschl (1998) consider extreme value distributions rather than parent distributions only and assume an exponential distribution for the point rainfall intensity. The final Sivapalan and Blöschl (1998) ARF expression using this method is complex but is dependent on the catchment area, spatial correlation length, duration, and return period; for very large return periods, Sivapalan and Blöschl (1998) find the ARF to vary with catchment area and correlation structure only. While ARFs are found to be loosely associated with duration, the authors note that the spatial correlation length (the most critical parameter) is closely related to duration and to storm type. The ARF is calculated as follows (formula from Pietersen et al., 2015):

$$ARF \left[k^2 \left(\frac{A}{\lambda^2} \right), T_d, T \right] = \frac{b(T_d)c(T_d)k^2 F_2(k^{-2}) - \frac{k^2}{F_1(k^{-2})} \ln \left[\ln \left[\frac{T}{T-1} \right] \right]}{b(T_d)c(T_d) - \ln \left[\ln \left[\frac{T}{T-1} \right] \right]} \quad (17)$$

where

A = catchment area (km²)

b = function of duration, where $b(T_d) = -0.05 + 0.25T_d^{0.49}$

c = function of duration, where $c(T_d) = 0.2 + 20T_d^{-0.7}$

$F_1(k^{-2}) = 1 - 0.17 \ln(k^{-2})$

$F_2(k^{-2}) = 0.39 + 0.61(k^{-2})^{0.8}$

k^2 = rainfall correlation structure

T = return period (years)

T_d = storm duration (hours)

λ = spatial correlation length (km)

Although the development of the spatial correlation method can be traced back to Rodriguez-Iturbe and Mejía (1974), it is unclear whether this type of method has been applied in practice. The method is, in theory, less dependent upon data (compared with the empirically based approach). However, retuning of the parameters based on local data may still be needed. Computational complexity will be mainly from modeling implementation and parameter fitting. It will also require further efforts to understand the underlying theory for proper applications, which may not be appealing to many users of ARF.

3.2.2.3 Statistical Crossing Properties Methods

Bacchi and Ranzi (1996) present a method based on the statistical crossing properties of rainfall. Following the method, rainfall properties that cross a high threshold of rainfall intensity are analyzed and integrated over space and time, assuming the crossings follow a Poisson distribution. The ARF is calculated as follows:

$$\begin{aligned}
ARF(A, D, AEP) &= F_{A,D}^{-1}(1 - AEP)/F_D^{-1}(1 - AEP) \\
F_D(b_D) &= Prob[\max X_D(x, y, t) \leq b_D, (x, y) \in A, t \in D] \\
F_{A,D}(b_{A,D}) &= Prob[\max X_{A,D}(x, y, t) \leq b_{A,D}, (x, y) \in A, t \in D]
\end{aligned} \tag{18}$$

where

A = area under consideration (km²)

D = duration (hours)

AEP = annual exceedance probability

$F_D(b_D)$ = probability that the maximum value of the point rainfall intensity X_D does not exceed b_D over the period D and the spatial domain A

$F_{A,D}(b_{A,D})$ = probability that the maximum value of the areal rainfall intensity $X_{A,D}$ does not exceed $b_{A,D}$ over the period D and the spatial domain A

The authors note that the method may be useful for design storm applications in small urban catchments for short durations. This method may not be particularly useful for PFHA at NPP sites, since longer-duration and larger-area ARFs should be needed in most cases.

3.2.2.4 Spatial and Temporal Scaling Methods

Scaling Properties of Annual Maxima Approach (De Michele et al., 2001)

De Michele et al. (2001) present an ARF method that reflects the scaling properties of rainfall in space in time, using the concepts of dynamic scaling and statistical self-affinity. The authors develop a model calibrated to empirically derived ARFs from 8 years of data gathered near Milan, Italy. The fitted results show reasonable agreement for durations of 1 and 3 h, but less agreement for durations of 20 min and 6 h. The results also show less agreement for increasing area sizes. This method was later tested by Pavlovic et al. (2016) in the Oklahoma study based on radar-driven observation. The ARF is calculated as follows (formula modified from Pietersen et al., 2015):

$$ARF(A, D) = \left[1 + \omega \left[\frac{A^z}{D} \right]^b \right]^{-\frac{v}{b}}, \tag{19}$$

where

A = area (km²)

D = duration (hours)

b, v, ω, z = fitted parameters

T = return period (years)

Multifractal Scaling Properties of Annual Maxima Approach (Veneziano and Langousis, 2005)

Veneziano and Langousis (2005) analyze ARF under the condition that space-time rainfall displays approximate multifractal scale invariance and found the multifractal approach to explain many features of empirical ARF charts. The study found that ARF varies with rainfall advection, basin shape, and return period and that such dependencies are difficult to quantify empirically.

Although the concepts of dynamic scaling and statistical self-affinity may not be familiar to many H&H modelers, the final ARF form (Eq. [19]) is in fact reasonably easy to fit and use. This spatial and temporal scaling approach also did not require the assumption of a specific type of distribution. Further testing can be performed to understand its transferability across different regions.

3.2.2.5 Extreme Value Theory Methods

Given the wide use of extreme value theory and GEV distribution in point rainfall frequency analysis, many methods extend such a concept into the estimation of ARF. It can either include a direct extension of point rainfall frequency analysis to areal rainfall observation at various areas (e.g., radar-driven rainfall aggregated at different grid resolutions) or fitting of a single regional GEV distribution across various areas. Many of these new methods use the radar-driven observation (as opposed to the conventional gauge observation) to support the derivation of ARF.

Extension of Point Rainfall Frequency Analysis (Durrans et al., 2002; DeGaetano, 2005a)

Durrans et al. (2002) evaluated 4-km (2.49-mi) resolution radar-driven rainfall data over a rectangular area of the central U.S. for 1-, 2-, and 4-hour durations over a 7.5-year period. They used Gumbel distribution (a special case of GEV) for modeling the annual maximum series of both grid cell and areally averaged precipitation depths, which is a direct extension from the conventional point rainfall analysis approach. Durrans et al. (2002) found unexpectedly high ARFs above 1 for some averaging areas, noted issues with the period of record available and data processing heterogeneity, and reported that edge effects of the spatial smoothing algorithm can affect the results and produce ARFs greater than 1. Durrans et al. (2002) also reported 100-year estimates that were 20–35% lower than gauge-based estimates from previous literature. The study also reports that radar-based ARFs are reasonably consistent with the TP-29 (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960) estimates but that they do not decrease with area as rapidly as gauge-based estimates; it also reports that the radar-based ARF curves are similar across return periods but that ARFs decrease more rapidly for larger return periods.

A study by Allen and DeGaetano (2005a) evaluated ARFs in New Jersey and North Carolina using 5 years' worth of daily radar data with 2-km (1.24-mi) resolution. In contrast to findings from Durrans et al. (2002), Allen and DeGaetano (2005a) found that radar-based ARFs decay more rapidly for increasing area than gauge-based ARFs. For a 20,000 km² (7,722 mi²) basin, the study found radar-based and gauge-based ARFs to differ by 11–32%. The study concluded that higher return periods are associated with lower ARFs, warm season (April–September) ARFs decay faster than cold season (October–March) ARFs, geographic variation in ARF is minimal but may vary depending on different primary precipitation mechanisms, Thiessen-weighted or inverse distance weighted averaging provides improvements over unweighted averages, and station density has little effect on ARF for the densities tested.

A study by Lombardo et al. (2006) used 1-km² (0.386-mi²) resolution radar data to estimate rainfall for areas of up to 900 km² (347 mi²). The study found that ARFs for 200–900 km² (77.2–347 mi²), 25–50 year return periods, and 1–2 hour durations were much lower than empirical values from literature, ranging from 0.1–0.3 compared with 0.4–0.8 from literature. The authors suggest that the methodology should be applied in other locations to generalize the results. The ARF is calculated as (formula from Pietersen et al., 2015)

$$ARF_{(T_d, T)} = \frac{i_A(T_d, T)}{i_{A=1}(T_d, T)}, \quad (20)$$

where

A = area under consideration (km²)

i = rainfall intensity (mm/h)

T = return period (years)

T_d = storm duration (hours)

Regional Generalized Extreme Value Distribution Approach (Overeem et al., 2010)

Another promising approach is to use the concept of covariance to generalize the point-based GEV into regional GEV distribution. Using 11-year high-quality radar rainfall data, Overeem et al. (2010) develop a generalized regional GEV distribution to estimate ARF for durations of from 15 min to 24 h and area sizes of 6 to 1700 km² for the Netherlands. This approach tries to fit a more generalized GEV that incorporates area and duration as co-variates in the GEV parameters. The ARFs are then calculated by using the generalized GEV. The approach is formulated as

$$\begin{aligned} ARF(A, D, AEP) &= P(A, D, AEP) / P(A^*, D, AEP) \\ P(A, D, AEP) &= GEV^{-1}(1 - AEP | \mu(A, D), \gamma(A, D), \kappa(A)) \\ \mu(A, D) &= aD^b + (c + d \ln D)A^e \\ \gamma(A, D) &= f \ln A + g \ln D + h \\ \kappa(A) &= i \ln A + j \end{aligned} \quad (21)$$

where

A = area (km²)

D = duration (hours)

AEP = annual exceedance probability

a through j = Overeem et al. (2010) parameters

A^* = additional parameter representing selected minimum area

Although parameter estimation is a bit more challenging than the conventional frequency analysis, this approach can fit a single distribution across various durations and areas and help derive ARF naturally. Nevertheless, note that depending on the form of the variates, the regional GEV may not have a bound when area approaches 0. In such cases, the values of ARF will also be affected by the selection of minimum area A^* in the calculation. This factor will be tested in a follow-up case study.

Extreme value theory methods represent a natural extension to the conventional point rainfall frequency analysis. Therefore, the data dependency and required assumptions are generally similar to the point rainfall frequency analysis. Computational complexity would inevitably increase given the expanded scope at the areal dimension. The co-variate-based GEV approach is a promising method. Nevertheless, further independent analysis and validation should be performed to better understand its capabilities and limitations.

3.3 Selection of Potential ARF Methods for PFHA

To help narrow down the list of available ARF methods and identify several potentially suitable options for further evaluation in the use case study (Section 4), qualitative assessment considerations from Section 3.2.1 are used to guide the overall selection. Each qualitative assessment category (except for Site-specific Transferability) is assigned a 0 to 1 score; higher scores indicate the method could be more suitable for the PFHA need. The detailed qualitative assessment results are provided in Appendix B, Table B-1. General scoring rationales are described as follows:

- Spatiotemporal Scale and Resolution—A score close to 1 indicates the original method was developed for a wide range of areas and durations; a score close to 0 indicates the method was developed only for limited areas and/or durations. In other words, considering the varying sizes of NPP watersheds in the U.S., a method with a wider range of spatiotemporal scales is preferred.
- Data Sufficiency and Dependency—This includes two aspects related to data. Regarding data sufficiency, a higher score suggests the original method was developed based on a large set of observational data set (and vice versa). Regarding data dependency, a higher score suggests the data used for calculation are easily accessible and generally reliable and various data sources could be used; a lower score suggests the required data are highly limited and prevents general application.
- Required Assumptions—A score close to 1 indicates few assumptions are required and the required assumptions are clear and supported by available literature; a score close to 0 indicates the assumptions are overly simplified or impractical.
- Analytical Complexity—A score close to 1 indicates the method can be implemented with little or moderate effort; a score close to 0 indicates the method is computationally complex and/or requires significant effort to implement.
- Independent Evaluation—A score close to 1 indicates the method has been independently evaluated with highly positive findings or has been applied widely in practice; a score close to 0 indicates the method either has not been independently reviewed in literature, has been found to have major flaws, or would not be practical for case study application.
- Site-specific Transferability—This consideration is associated with a specific NPP site and is not considered now. If a specific NPP site is being considered, a higher transferability would suggest more similar geographical settings, local climate patterns, and/or site-specific features; a lower transferability would suggest dissimilar background and inappropriateness for NPP-PFHA application.

For Empirical Methods, overall, a lower score is assigned to “Data Sufficiency and Dependency” and a higher score to “Required Assumption.” This reflects the nature of empirical methods, which are highly data driven. Among them, the size of study area mainly affects the evaluation of “Spatiotemporal Scale and Resolution.” While the U.S. Weather Bureau Approach receives a lower score (given its development for limited area sizes), a higher score is given to the United Kingdom Approach given its coverage of a large range of area and durations. It is worth noting that the fitted formulae were developed and provided by multiple empirical methods (e.g., Grebner et al., 1998). For the purpose of PFHA, one may easily refit these formulae using local

extreme precipitation information. Therefore, their “Analytical Complexity” receives a higher score than other empirical ARF products that provide only graphical charts. Note that while both the Asquith and Famiglietti (2000) and Grebner et al. (1998) approaches are storm-centered, the empirical model developed for the Swiss approach can be easily used to fit precipitation data for comparative curve fitting.

Compared with the Empirical Methods, other methods developed with supporting theories generally received higher scores in “Data Sufficiency and Dependency” and lower scores in “Required Assumptions.” Again, “Spatiotemporal Scale and Resolution” was mainly evaluated based on the sizes of study areas and the range of durations demonstrated in the original papers or reports. The “Analytical Complexity” of these methods varies but in general should be more involved than the empirical methods. Further insight can be obtained from testing various methods in the case study in Section 4.

It should be emphasized that the scores listed in Appendix B merely represent a qualitative understanding obtained during literature review. Depending on the intended PFHA applications, these selection criteria should be revised based on more specific needs. Overall, the following methods are selected for further evaluation and comparison in Section 4:

- Empirical Methods
 - M1: Leclerc and Schaake (1972)— fitted formula of TP-29 ARF (Eq.[3])
 - M2: Koutsoyiannis and Xanthopoulos (1999)—fitted formula of UK ARF (Eq. [5])
 - M3: Swiss Approach—fitted formula (Eq. [6]) by Grebner et al. (1998)
 - M4: Australian Rainfall and Runoff Approach—fitted formula (Eq. [11]) by Nathan and Weinmann (2016)
- Spatial and Temporal Scaling Method
 - M5: Dynamic Scaling Model (De Michele et al., 2001), Eq. ([19])
- Extreme Value Theory Method
 - M6: Regional GEV Model (Overeem et al., 2010), Eq. ([21])

4 ARF USE CASE STUDIES

Section 4 summarizes the study domain, data, methods, and results for several use cases designed to demonstrate ARF estimation in three selected regions and across the CONUS. A watershed-based annual maximum precipitation searching approach is used to identify ARF samples across different watershed sizes for further ARF model fitting. The use case study only considers fixed-area ARF; storm-centered ARFs are considered less conservative for the purpose of PFHA and not evaluated in this study (as discussed in Section 1.4). Through these use cases, a quantitative comparison of major factors affecting ARFs is provided. The overall design of the use case study is summarized in Table 4-1.

Table 4-1 Summary of the Overall Design of the Use Case Study

Factors	Considerations in the Use Case Study
Study domains (geographical variability)	<ul style="list-style-type: none"> • Regional assessments (with multiple precipitation products and ARF models) <ul style="list-style-type: none"> ○ Ohio (Hydrologic Region 05) ○ South Atlantic-Gulf (Hydrologic Region 03) ○ Mid-Atlantic (Hydrologic Region 02) • CONUS assessment (with one precipitation product and one ARF model) <ul style="list-style-type: none"> ○ 18 national hydrologic regions
Precipitation products	<ul style="list-style-type: none"> • DSI-3240: 1950–2013 NCEI hourly gauge observations • PRISM: 1981–2017 daily gridded precipitation (Daly et al., 1994) • Daymet: 1980–2017 daily gridded precipitation (Thornton et al., 1997) • Livneh: 1950–2013 daily gridded precipitation (Livneh et al., 2015) • ST4: 2002–2017 NCEP Stage-IV hourly radar-driven precipitation
Maxima searching	<ul style="list-style-type: none"> • Watershed-based annual maximum precipitation searching approach
Area	<ul style="list-style-type: none"> • From point location to different watersheds sizes
Duration	<ul style="list-style-type: none"> • Daily and above: 1-, 2-, and 3-day • Subdaily: 1-, 2-, 3-, 6-, 12-, and 18-hour (depending on data sets)
Seasonality	<ul style="list-style-type: none"> • All season: AMS searched during each calendar year • Warm season: AMS searched during May through October • Cool season: AMS searched during November through April
Frequency (return period)	<ul style="list-style-type: none"> • Average annual maximum precipitation (around 2-year return level) • 10- and 100-year precipitation (fitted by GEV)
ARF models	<ul style="list-style-type: none"> • M1: Leclerc and Schaake (1972) TP-29 Model • M2: Koutsoyiannis and Xanthopoulos (1999) United Kingdom–Natural Environment Research Council (UK-NERC) Model • M3: Hydrological Atlas of Switzerland Model (Grebner et al., 1998) • M4: Australian ARR Model (Nathan and Weinmann, 2016) • M5: De Michele (2001) Dynamic Scaling Model • M6: Regional GEV Model (Overeem et al., 2010)
Fitting statistics	<ul style="list-style-type: none"> • NSE: Nash-Sutcliffe model efficiency coefficient • RMSE: Root mean square error

The authors note that additional return periods beyond 100 years were also analyzed, but that due to the poor fitting model performance, the results of such analyses are not summarized in this report. The issue largely related to the limited observed data record length available from the precipitation products leading to high uncertainty in the model results.

In addition, given the significant effort required to separate historical precipitation data based on storm type, the effect of storm type is not explicitly explored in this use case study. Instead, seasonality is used as an appropriate proxy parameter to demonstrate the potential influence of storm type on ARF.

4.1 Study Domain

The use cases include (1) regional assessments of ARFs for three selected hydrologic regions focusing on different precipitation products, and fitting models to demonstrate major factors affecting ARFs and (2) a CONUS assessment of ARFs across all hydrologic regions, focusing on the use of one precipitation product and one fitting model to demonstrate geographic variation in ARFs.

4.1.1 Regional Assessments

Three hydrologic regions (also known as the 2-digit hydrologic units, HUC02) are assessed to develop regional ARF models across different areas, durations, and return periods using different precipitation products and fitting models. These three HUC02 regions include Ohio (Region 05), Mid-Atlantic (Region 02), and South Atlantic-Gulf (Region 03). Maps of each region with its finer hydrologic units (i.e., HUC04 subregions, HUC06 basins, and HUC08 subbasins) are provided in Figure 4-1, Figure 4-2, and Figure 4-3, respectively. These regions were selected in this use case study because they represent regions with differing precipitation drivers for which data availability is good, and because of their relevance to existing and proposed NPPs.

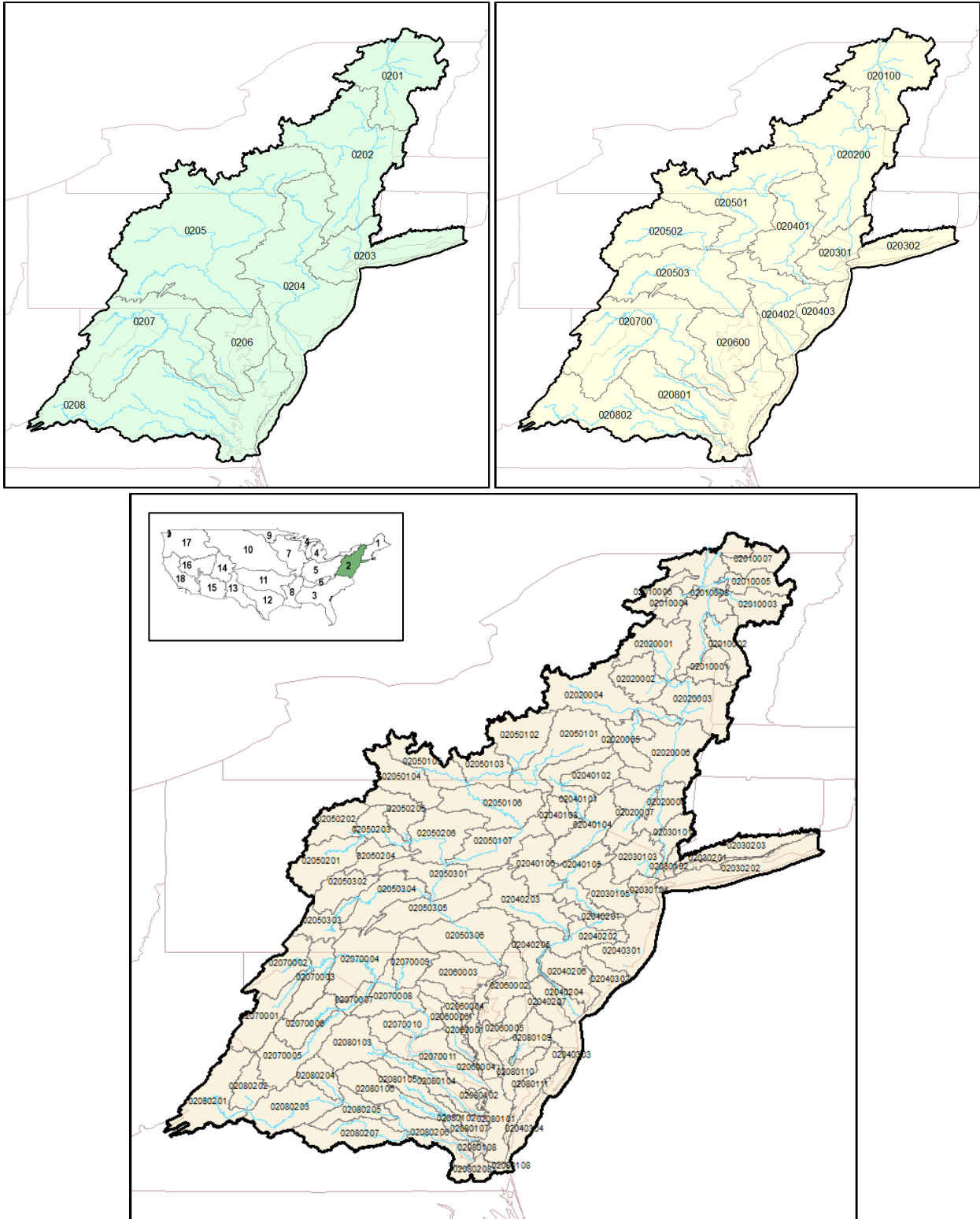


Figure 4-2 Mid-Atlantic (Region 02) 4-digit, 6-digit, and 8-digit Hydrologic Units (From Top Left to Bottom)

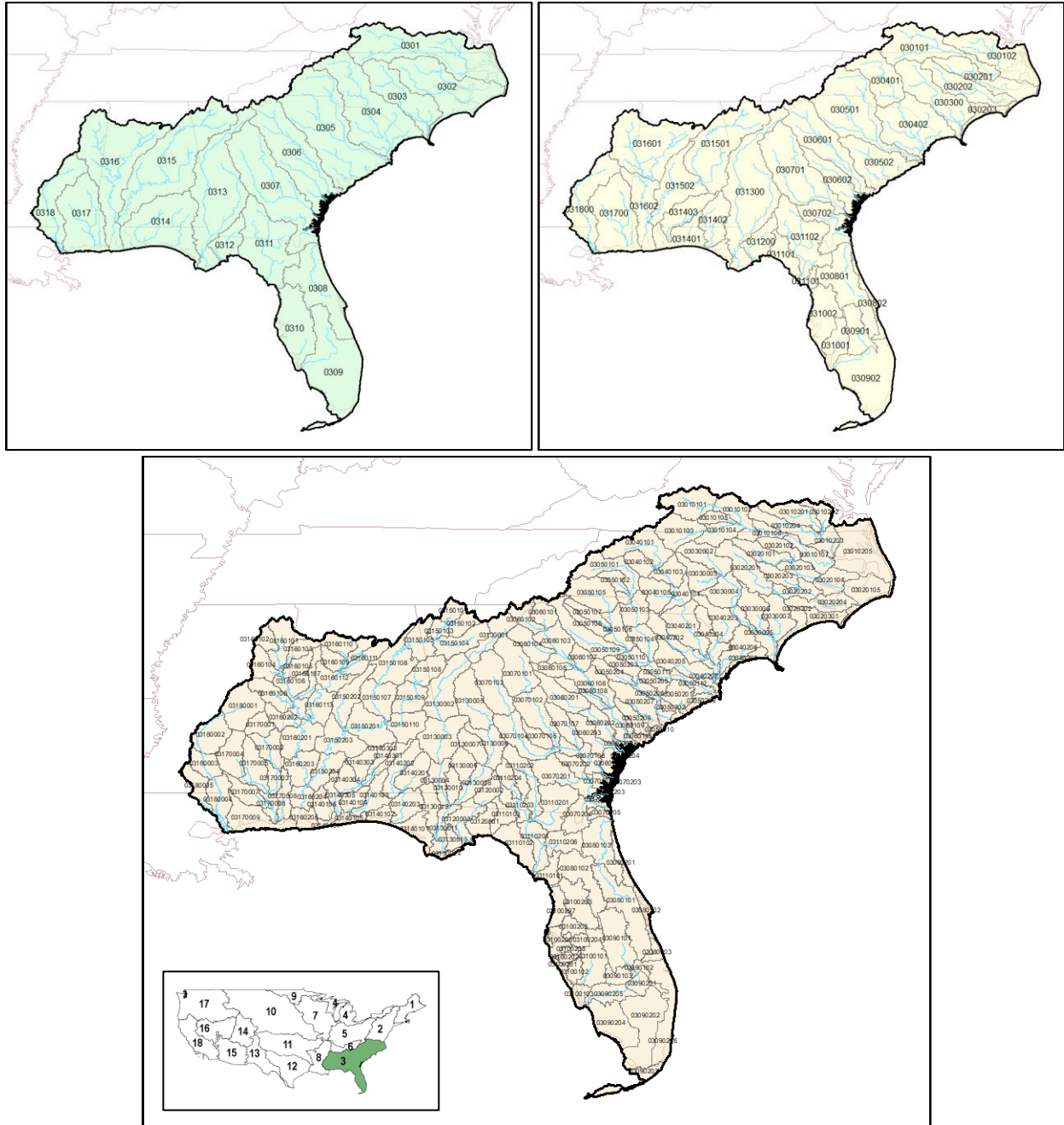


Figure 4-3 South Atlantic–Gulf (Region 03) 4-digit, 6-digit, and 8-digit Hydrologic Units (From Top Left to Bottom)

4.1.2 CONUS Assessment

The CONUS assessment uses one precipitation product (PRISM) and one fitting model (M5: De Michele Dynamic Scaling Model) to assess the geographic variations in ARFs across different areas, durations, and return periods. Each HUC02 region in the CONUS (Figure 4-4) is evaluated independently using the PRISM data and M5 fitting model. The decision to use PRISM data was largely motivated by its used in NOAA Atlas 14, complete CONUS coverage,

and good spatial resolution. The decision to use the M5 fitting model was largely motivated by its simplicity and generally good performance.



Figure 4-4 HUC02 Hydrologic Regions in CONUS

4.2 Selected Precipitation Products

The data collected for use case application are from five different precipitation products, including hourly gauge observations (DSI-3240), daily PRISM, daily Daymet, daily Livneh, and hourly NCEP Stage-IV (ST4). A summary of the five selected precipitation products is provided in Table 4-2. DSI-3240 is selected for its high temporal resolution (hourly) and because such gauge observation is generally considered to be “ground truth” for observed precipitation. The radar-driven ST4 is selected because of its high temporal resolution (hourly) and because of radar’s unique strength in detecting spatial storm structures that cannot be measured by sparse rain gauges. The three gridded daily precipitation data products (PRISM, Daymet, Livneh) are selected for their wide applications in various hydrologic studies and because they can be easily analyzed using a similar approach. It is worth mentioning that the daily PRISM data also incorporate radar information in their post-2002 gridded precipitation. Given the influence of radar information, ST4 and post-2002 PRISM generally show more spatial variability than Daymet and Livneh, especially near boundaries of radar data availability.

Among the gridded precipitation products, Daymet has the finest spatial resolution (1 km), followed by PRISM and ST4 (4 km), and then Livneh (6 km). Despite Livneh’s coarser spatial resolution, it has the longest record (64 years from 1950–2013), followed by Daymet (38 years from 1980–2017), PRISM (37 years from 1981–2017), and ST4 (16 years from 2002–2017). Since both spatial resolution and data length are important features affecting the accuracy of ARF, trade-offs exist in selecting the most appropriate precipitation product to support ARF calculation.

Additional treatment (e.g., spatial interpolation or Thiessen polygon approach) is needed to process the DSI-3240 gauge observations for ARF calculation. Since DSI-3240 data collection, processing, quality control, and analysis require more effort, hourly gauge observations are analyzed only in the Ohio (Region 05) regional assessment. As an example, over 300 NCEI DSI-3240 stations with more than 30 years of record in Ohio (with density ~1400 km² per station) are shown in Figure 4-5. The PRISM, Daymet, Livneh, and NCEP ST4 data are used across the three regional assessments. The CONUS assessment uses only PRISM data. Further discussion regarding the background and features of each selected data set can be found in Section 2.2.

Table 4-2 Selected Precipitation Products and Features Used in Demonstrations

N	Precipitation Data Set	Provider	Data Set Type	Start	End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution
1	Hourly Precipitation Data (DSI-3240)	NOAA National Centers for Environmental Information (NCEI)	Gauge observation	1950	2013	Data since 2014 have not been released (checked 10/17/2017)	U.S. (including AK, HI, PR)	Hourly	Gauge
2	Daymet	Oak Ridge National Laboratory (ORNL)	Gridded from gauge observation	1980	2017	Annual update	North America	Daily	1 km * 1 km
3	PRISM (daily)	Oregon State University	Gridded from gauge observation (and partially with radar)	1981	2017	Operational (updated automatically)	U.S. (48 states)	Daily	1/24 deg * 1/24 deg (~ 4 km * 4 km)
4	Livneh Daily CONUS Near-surface Gridded Meteorological Data	University of Colorado, Boulder	Gridded from gauge observation	1950	2013	No scheduled update (checked 10/17/2017)	U.S. (48 states), Mexico, and Canada (south of 53N)	Daily	1/16 deg * 1/16 deg (~ 6 km * 6 km)
5	NCEP National Stage IV Analyses	NOAA National Centers for Environmental Prediction (NCEP)	Merged radar and gauges (with QC)	2002	2017	Operational (updated automatically)	U.S. (48 states), excluding California-Nevada and Northwest RFCs	Hourly	4 km * 4 km

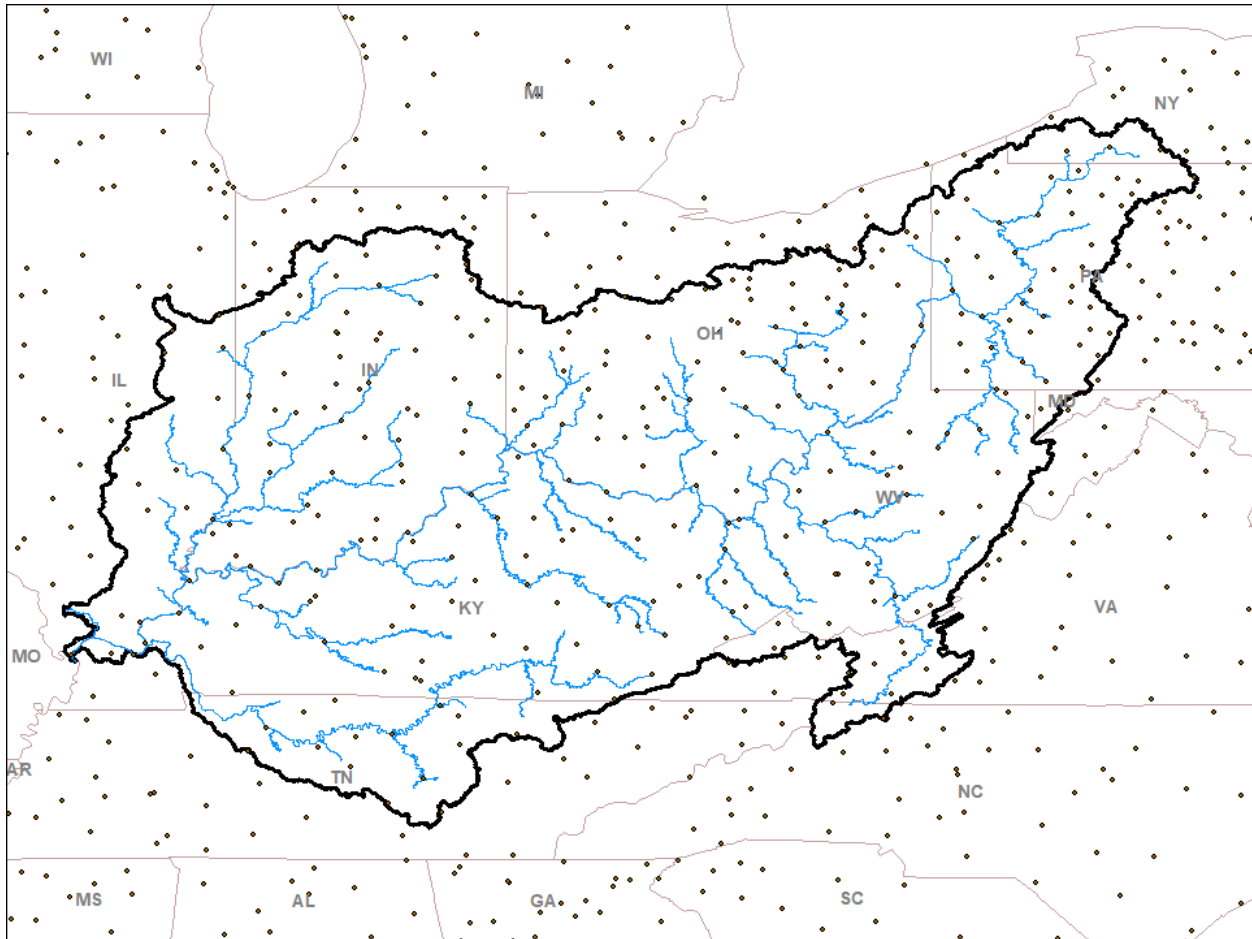


Figure 4-5 NCEI Hourly Rainfall Stations with 30+ Years of Record in Ohio (Region 05)

4.3 ARF Assessment

The calculation of ARF involves three main steps: (1) AMS searching, (2) sample ARF calculation, and (3) ARF model fitting. Conceptually speaking, ARF derivation involves initial steps which are similar to the procedures of point-based frequency analysis, but extends the AMS searching and probabilistic density function (PDF) fitting from each point location to the watershed scale. The use case demonstrations follow a data-driven approach in which some level of quality assurance and control are performed, although the raw data are largely unaltered. This section provides a summary of calculation procedures employed in this case study.

4.3.1 AMS Identification

Similar to the procedures of PFA, the first step of ARF calculation is to identify maximum precipitation from either the AMS or partial duration series (PDS) approaches (i.e., PDS is an alternative approach that identifies all maximum precipitation samples above a defined threshold). To be consistent with the prominent literature (including NOAA Atlas 14), the AMS approach is used in this use case study. Also, for consistency (with NOAA Atlas 14), AMS are searched for each calendar year (January to December) instead of each water year (October to

September). While some AMS may be different when searched by water year, the overall influence on the final PFA estimates should be very limited.

Since the main purpose of ARF is to suggest how the values of extreme precipitation change across different spatial scales, the AMS identification should be conducted at various aggregated watershed or catchment scales. Using Figure 4-6 as an example, consider $R_{grid}(d, g)$ as a daily gridded precipitation field in which d represents a certain day and g represents a certain grid location. Given that extreme precipitation (e.g., mesoscale convective systems, tropical storms, hurricanes) may occur at different scales with large spatial variability, the annual maximum precipitation across all grids may occur at different times. With regard to understanding the frequency of extreme precipitation for a given catchment in this field, the average of grid-based AMS will likely overestimate the magnitude of extreme precipitation. Therefore, one should use the catchment shape as a spatial filter to first spatially aggregate daily precipitation from $R_{grid}(d, g)$ to $R_{Area}(d)$, and then identify AMS from this aggregated precipitation time series for analysis.

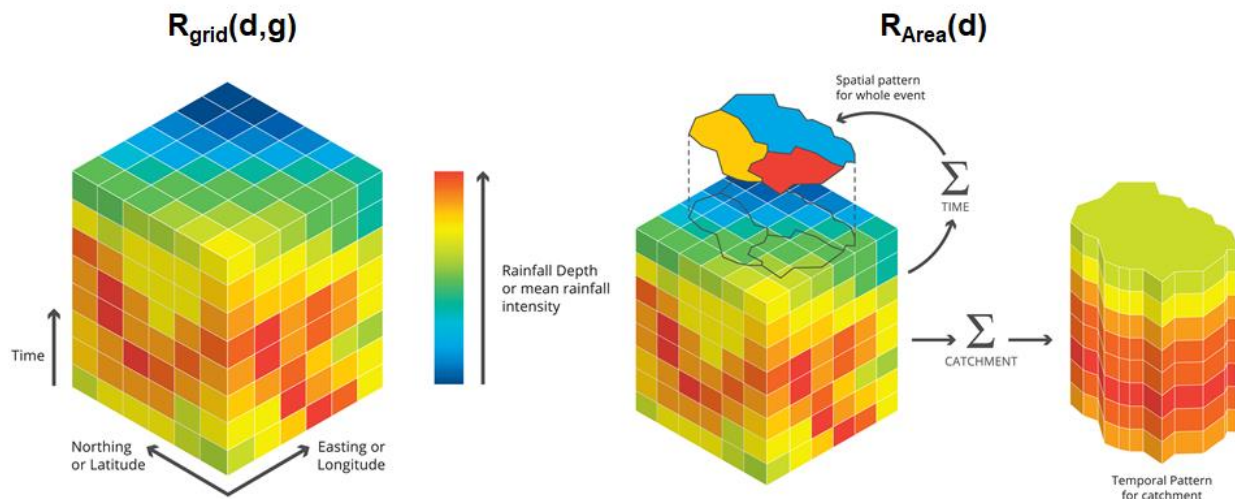


Figure 4-6 Spatial and Temporal Aggregation Diagram Used for Gridded Precipitation Products Source: ©Commonwealth of Australia (Geoscience Australia) 2019. This Product is Released Under the Creative Commons Attribution 4.0 International Licence. <http://creativecommons.org/licenses/by/4.0/legalcode>

Following this concept, and taking the calendar year 2002 as an example, in searching the maximum 1-day precipitation at each PRISM grid within the entire Ohio River Basin, it is found that the timing of grid-based AMS spreads across different seasons and belongs to different events (Figure 4-7a). On the other hand, in searching the basin-wide maximum average precipitation of the entire Ohio River Basin, the 1-day AMS event is found to be on September 27, 2002. In comparing grid-based and basin-wide AMS (Figure 4-7b and Figure 4-7c), it can be seen that while nearly half of the grid-based AMS are the same as basin-wide AMS, grid-based AMS reports high precipitation across the watershed. Therefore, using grid-based AMS to conduct point-based PFA will lead to higher extreme rainfall estimates than using areal PFA based on basin-wide AMS. It is intuitive that, when the catchment size increases, the AMS identified for the watershed will also become smaller compared with the average of AMS across all corresponding grids. This concept forms the basis of ARF as a reduction factor needed to

convert point-based precipitation frequency estimates to areal estimates for watershed application.

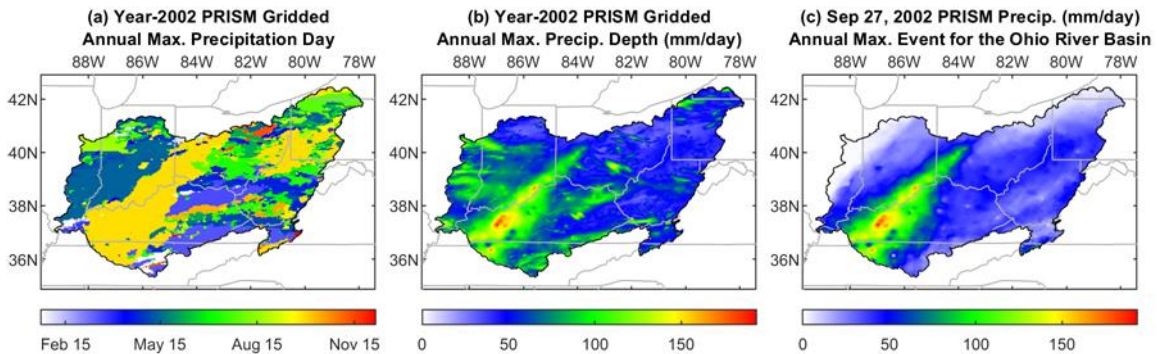


Figure 4-7 Example of AMS Searing in the Ohio River Basin

Although previous areal AMS approaches have used different spatial aggregation techniques to analyze areal extreme precipitation (e.g., square n -by- n grids or circular windows), the U.S hydrologic unit code (HUC) watersheds are used as a spatial filter in this study. The HUC is a hierarchical labeling structure to organize U.S. watersheds across different sizes (e.g., Figure 4-1, Figure 4-2, and Figure 4-3). For each hydrologic HUC02 region, spatially aggregated AMS are searched for each HUC04 subregion, HUC06 basin, and HUC08 subbasin. The overall AMS at the HUC02 region scale is not searched since in many cases HUC02 itself is not a single watershed (e.g., Region 02 Mid-Atlantic and Region 03 South Atlantic–Gulf). Since these HUC units are defined by watershed boundaries, this HUC-based spatial aggregation approach lends itself well to hydrologic applications, including the PFHA for NPPs.

In testing this HUC-based AMS identification approach, a larger data gap for large-area samples was observed (i.e., there are fewer HUC04s than HUC06s and HUC08s). To address this issue and increase the AMS samples to cover a wider range of watershed sizes, HUC accumulation technique is applied in this use case study for developing accumulated HUC units (HUCac). The concept of HUCac is illustrated in Figure 4-8. In the left panel of Figure 4-8, all HUC08s in the Ohio Region are shown with their upstream and downstream subbasin connectivity. Using this HUC08 connectivity information, taking HUC08 05090203 as an example (marked in blue in the right panel of Figure 4-8), all upstream contributing HUC08s to 05090203 are labeled and hence identify the entire upstream contributing area as HUCac (to HUC08 05090203). Following this approach, HUCac is identified for each of the HUC08s based on their upstream connectivity, and then search AMS for each HUCac. When an HUCac is identical to an existing HUC04, HUC06, and HUC08, it is neglected, to avoid double-sampling.

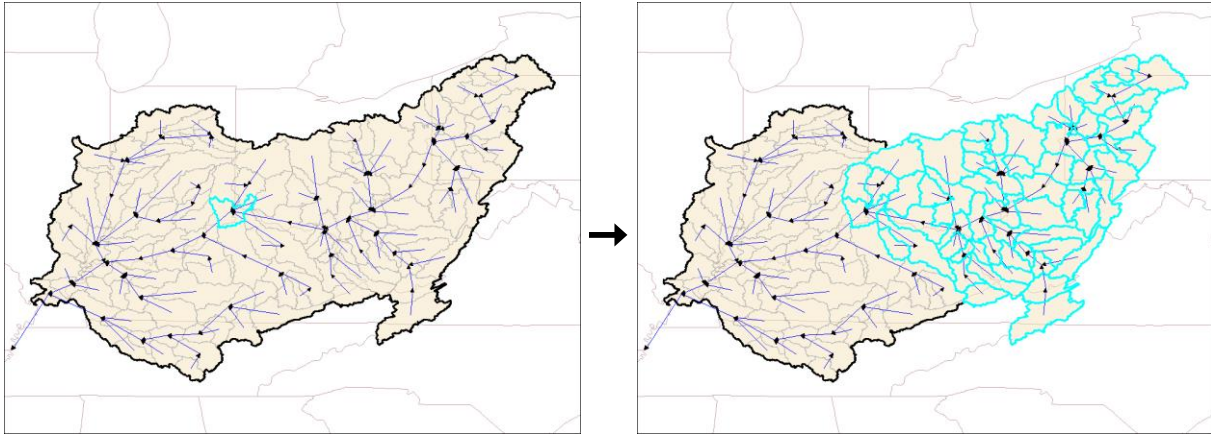


Figure 4-8 Example Spatial Aggregation of Ohio Region HUC08s to form a HUCac Upstream of HUC 08 05090203

Following the HUCac implementation, the final Ohio (Region 05) AMS sampling consists of the following HUC-based spatial units, where the largest HUCac in this case is the entire Ohio River Basin (420,000 km²):

- 120 HUC08: 290 – 840 km²
- 21 HUC06: 4,400 – 54,000 km²
- 7 HUC04: 15,000 – 85,000 km²
- 46 HUCac: 4,600 – 420,000 km²

To effectively summarize gridded precipitation (PRISM, Daymet, Livneh, and ST4), a conversion table is established indicating what grid points should be included in a specific HUC unit. This conversion table is then used to spatially average all hourly and daily gridded precipitation into different HUC-based precipitation for AMS identification.

Since the DSI-3240 hourly gauge precipitation requires great analysis effort, it is only analyzed in Region 05 Ohio to assess the sensitivity of using different data. After performing processing and QC of all DSI-3240 hourly gauge data in and surrounding Region 05 Ohio, bilinear interpolation is used to spatially interpolate DSI-3240 values at each PRISM grid location. The HUC-based AMS identification approach is then followed to determine the AMS of DSI-3240. While the gauge data are ground-based and should ideally include topographic adjustments (e.g., elevation lapse rate), these considerations are not accounted for during spatial interpolation. This is acceptable in most of Region 05 Ohio, given its relatively flatter terrain, but it can present a greater impact for certain watersheds that contain significant topographic variation. In such cases, proper topography-informed adjustment approaches (such as those used in PRISM, Daymet, and Livneh) should be used to analyze spatial DSI-3240 precipitation.

Overall, AMS are searched for different

- Data
 - 1950–2013 DSI-3240 hourly gauge precipitation, spatially interpolated to the ~4 km resolution PRISM grids (Ohio only)
 - 1981–2017 PRISM daily gridded precipitation, ~4 km resolution
 - 1980–2017 Daymet daily gridded precipitation, 1 km resolution

- 2002–2017 ST4 hourly gridded precipitation, 4 km resolution
- 1950–2013 Livneh daily gridded precipitation, ~6 km resolution
- Durations
 - Daily and above: 1-, 2-, and 3-day
 - Subdaily for ST4 and DSI-3240 (Ohio only): 1-, 2-, 3-, 6-, 12-, and 18-hour
- Spatial units
 - Grid (P_{grid}): annual at each grid
 - Areal (P_{HUC}): annual at each HUC08, HUC06, HUC04, and HUCac
- Seasons
 - All seasons (January through December)
 - Warm season (May through October)
 - Cool season (January through April and November through December)
- Geographic coverage
 - Regional for Region 05 Ohio, Region 03 South Atlantic–Gulf, and Region 02 Mid-Atlantic
 - National for each HUC02 hydrologic region (PRISM only)

These calculated AMS values form the basis of the ARF calculation in this use case study.

4.3.2 Sample ARF Calculation

After AMS has been comprehensively searched for all HUCs and each grid point, the next step is to calculate the sample ARF at each HUC. These sample ARFs would represent the best-available, watershed-specific ARF estimates at each HUC (assuming that there are sufficient historic observations to support the ARF estimate). The ARF samples across all HUCs can then be grouped and jointly fitted into an ARF model for more generalized representation (discussed in the following section).

Let $P_{grid}(y, g)$ represent the annual maximum precipitation at year y and grid g , and $P_{HUC}(y)$ represent the annual maximum precipitation at year y for a particular HUC. Considering all grids in the HUC unit, the first type of sample ARF, ARF_{avg} , can be defined as

$$ARF_{avg} = \frac{P_{HUC,avg}}{P_{grid,avg}}, \quad (22)$$

$$P_{HUC,avg} = \frac{\sum_{y=1}^{N_y} P_{HUC}(y)}{N_y}, \quad (23)$$

$$P_{grid,avg} = \frac{\sum_{y=1}^{N_y} \sum_{g=1}^{N_g} P_{grid}(y,g)}{N_y N_g}, \quad (24)$$

where N_y is the total number of years, and N_g is the total number of grid points in the HUC unit. Following the definition, ARF_{avg} presents the ratio of average areal AMS to the average grid-based AMS. Approximately speaking, ARF_{avg} defined in this fashion would have a return period around 2 years. In rare cases, when calculating individual ARF values, it is possible to compute an ARF above 1.00. In these cases, the authors set the value to an upper ARF limit of 1.00.

Another type of sample ARF involves the selection of a suitable PDF and PDF fitting so that frequency (return period) can be introduced into the calculation of ARF. Assuming that the GEV distribution is a suitable PDF in this case, the T -year ARF_{Tyr} can be defined as

$$ARF_{T_{yr}} = \frac{P_{HUC,T_{yr}}}{P_{grid,T_{yr},avg}} \quad , \quad (25)$$

$$P_{HUC,T_{yr}} = GEV^{-1} \left(1 - \frac{1}{T} \left| \mu_{HUC}, \gamma_{HUC}, \kappa_{HUC} \right. \right) \quad , \quad (26)$$

$$P_{grid,T_{yr},avg} = \frac{\sum_{g=1}^{N_g} GEV^{-1} \left(1 - \frac{1}{T} \left| \mu_g, \gamma_g, \kappa_g \right. \right)}{N_g} \quad , \quad (27)$$

where GEV^{-1} represents the inverse of GEV; $\mu_{HUC}, \gamma_{HUC}, \kappa_{HUC}$ represent the GEV parameters estimated at the selected HUC unit using the HUC-based AMS; and $\mu_g, \gamma_g, \kappa_g$ represent the GEV parameters estimated at each grid point using the grid-based AMS. Clearly, other suitable PDFs may also be used to replace GEV to provide proper frequency estimates.

In this use case study, the maximum likelihood approach is used to estimate the GEV parameters. After parameter fitting, the Kolmogorov–Smirnov (KS) test is used to examine the goodness-of-fit at a 5% significance level. If a specific case fails to pass the KS test, the data point is disregarded for further ARF model fitting. Overall, the average ARF from Eq. (21) and the 10- and 100-year estimates from Eq. (24) are used in the following section.

Note that frequency can also be introduced into ARF from other approaches (e.g., Overeem et al., 2010). However, regardless of which approach is used, the limited data record would likely be the biggest hurdle to estimating long-return-level precipitation and ARF. Even with around 60 years of long-term records (from Livneh and DSI-3240 data sets), there may not be sufficient data to support the estimate of precipitation and ARF with return periods greater than 100 years, as the uncertainty associated with extrapolating return period estimates well beyond the record length continues to rise. In addition, nonstationarity in a changing environment would add further complication to the frequency analysis. These more involved issues are noted but not examined in this use case study.

4.3.3 ARF Model Fitting

The final step of the regional ARF model development is to fit all ARF samples in a hydrologic region with a generalized ARF model. Six ARF models are used in this study. They include the Leclerc and Schaake TP-29 Model (M1), the Koutsoyiannis and Xanthopoulos UK-NERC Model (M2), the Hydrological Atlas of Switzerland Model (M3), the Australian ARR Model (M4), the De Michele Dynamic Scaling Model (M5), and the Regional GEV Model (M6). These ARF models are discussed in Section 3.

M1: Leclerc and Schaake TP-29 Model

The Leclerc and Schaake (1972) Model is an empirical equation representing the U.S. Weather Bureau TP-29 ARF curves (U.S. Weather Bureau, 1957, 1958a, 1958b, 1959, 1960). The M1 model (Eq. [3] in Section 3.2.2.1) is a function of area A (km^2 or mi^2) and duration D (hours) with three parameters a , b , and c .

To fit the three M1 parameters, the root mean square error (RMSE) between ARF samples and ARF model is minimized using the nonlinear programming solver (from Matlab). The performance of fitting is further evaluated by the Nash–Sutcliffe efficiency (NSE; McCuen et al.,

2006) coefficient⁴. Given that M1 is not a function of return period, fitting is performed separately for each frequency level of interest (i.e., average AMS, 10-, and 100-year). An example of a typical M1 fitting is provided in Figure 4-9 (in semi-log scale to better illustrate ARF across a wide range of areas).

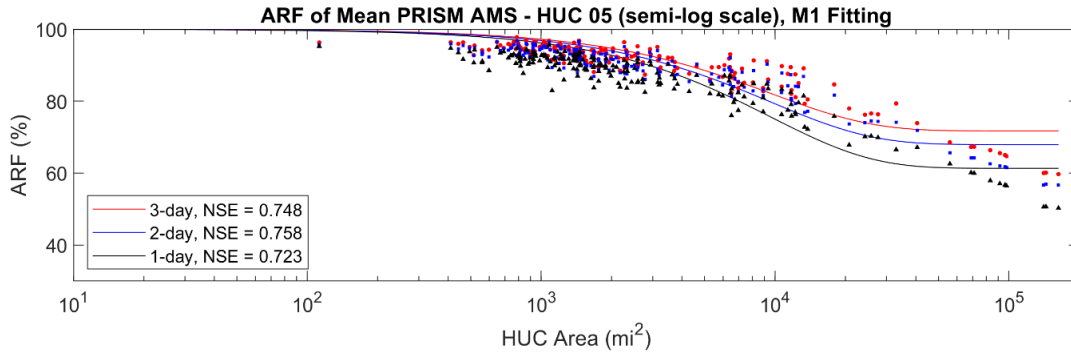


Figure 4-9 Example of M1 ARF Model Fitting

M2: Koutsoyiannis and Xanthopoulos UK-NERC Model

An empirical equation representing the UK Flood Studies Report ARF curves was established by Koutsoyiannis and Xanthopoulos (1999). The M2 model (Eq. [5] in Section 3.2.2.1) is a function of area A (km^2 or mi^2) and duration D (hours) with four parameters a , b , c , and d .

The fitting procedure for the four M2 parameters is similar to that for M1. Also, given that M2 is not a function of return period, fitting is performed separately for each frequency level of interest (i.e., average AMS, 10-, and 100-year). An example of a typical M2 fitting is provided in Figure 4-10.

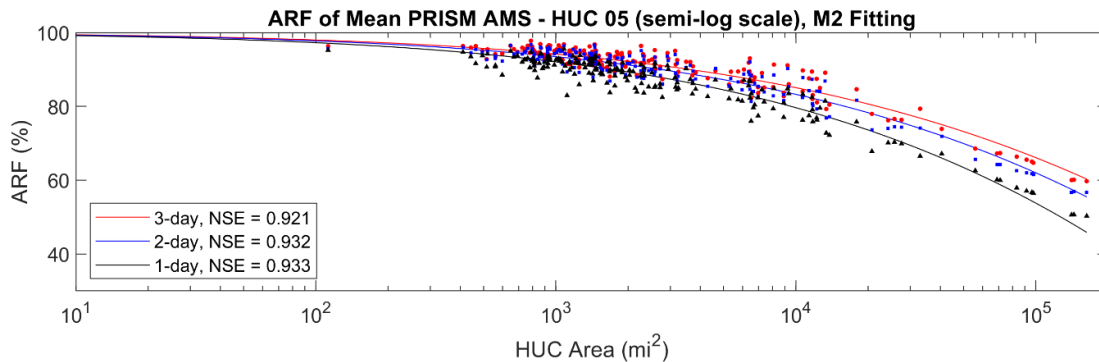


Figure 4-10 Example of M2 ARF Model Fitting

⁴ NSE coefficient values range from $-\infty$ to 1 with an efficiency of 0 indicating the model predictions are as accurate as the mean of the observed data. Higher values correspond to better model predictions, and values of 0.5 or above are generally considered to indicate good performance.

M3: Hydrological Atlas of Switzerland Model

The Hydrological Atlas of Switzerland (Grebner et al., 1998) provides a generalized equation for ARFs across different geographic zones of Switzerland. The M3 model (Eq. [9] in Section 3.2.2.1) is a function of area A (km^2 or mi^2) with five parameters a_0 , a_1 , a_2 , a_3 , and a_4 . Since ARF should be 1 when $A = 0$, an additional equation (Eq. [10] in Section 3.2.2.1) can be used to reduce one parameter.

We use a similar procedure to fit the remaining four M3 parameters. Unlike M1 and M2, M3 ARF is only a function of area (not duration). Therefore, fitting is performed separately for each duration and frequency level. An example of a typical M3 fitting is provided in Figure 4-11. Because of the reduced sample size, the fitting of M3 is generally more challenging than fitting of M1 and M2 and may be more sensitive to outliers in the fitting samples, especially at the higher return period level.

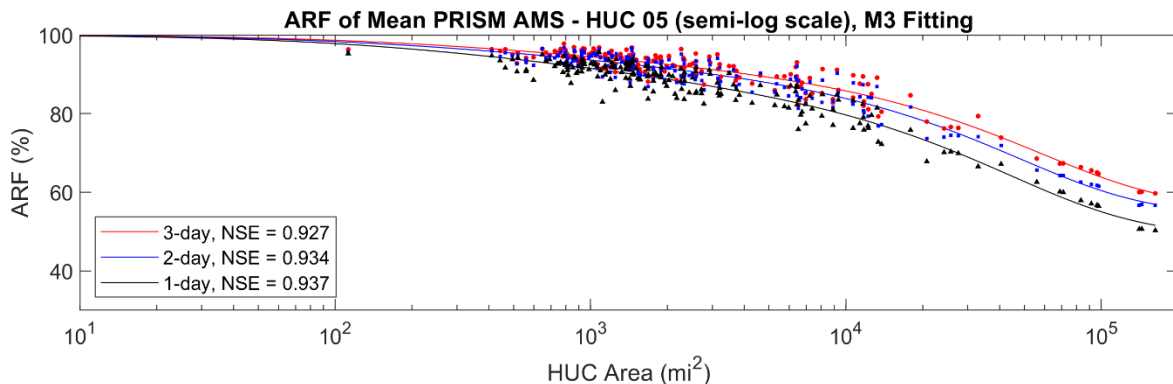


Figure 4-11 Example of M3 ARF Model Fitting

M4: Australian ARR Model

The ARR Guidelines (Nathan and Weinmann, 2016) provide a series of equations for ARF calculation, including a generalizable equation used for catchment areas of 10–1,000 km^2 , and durations of 1–7 days. This equation is deemed reasonable to use for demonstration purpose. The M4 model (Eq. [11] in Section 3.2.2.1) is a function of area A (km^2 or mi^2), duration D (hours), and annual exceedance probability (AEP) with nine parameters (a through i). Similar to M3, an additional equation (Eq. [12] in Section 3.2.2.1) can be used to reduce one parameter for fitting.

We use a similar procedure to fit the remaining eight M4 parameters. Unlike M1 through M3 models, M4 ARF is a function of area, duration, and return period. Therefore, fitting is needed only once across all frequency levels. While fitting eight model parameters through nonlinear programming solver is more challenging, the M4 model has an extra return period dimension and can be more flexible for application. Nevertheless, note that unless credible higher-return-level ARF samples (e.g., over 200 years) are included as a part of fitting, the fitted M4 model should not be used to derive ARF at a long-return-period level. An example of a typical M4 fitting is provided in Figure 4-12.

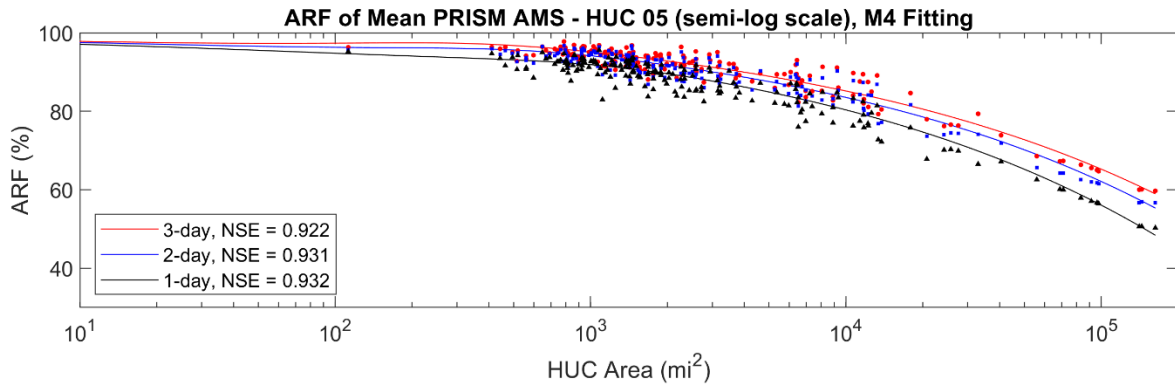


Figure 4-12 Example of M4 ARF Model Fitting

M5: De Michele Dynamic Scaling Model

The previous four methods employed are empirical methods. The fifth model, De Michele (De Michele et al., 2001), is selected because of its good underlying theory based on spatial and temporal rainfall scaling. The M5 model (Eq. [19] in Section 3.2.2.4) is a function of area A (km^2 or mi^2) and duration D (hours) with four parameters b , v , w , and z .

A similar fitting procedure is used to estimate the four M5 parameters. Similar to M1 and M2, the M5 model is also a function of area and duration. Therefore, fitting is performed separately at each frequency level (i.e., average AMS, 10-, and 100-year). An example of a typical M5 fitting is provided in Figure 4-13.

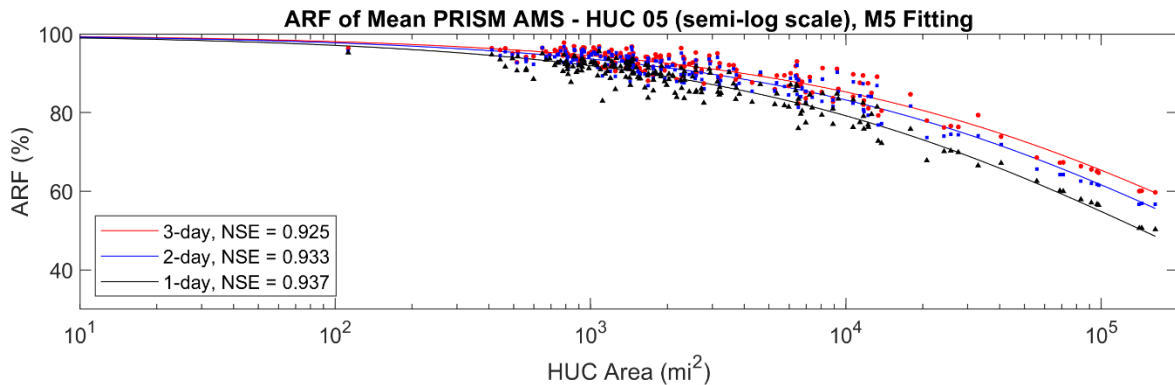


Figure 4-13 Example of M5 ARF Model Fitting

M6: Regional GEV Model

A final ARF model based on the regional GEV distribution approach is developed by Overeem et al. (2010). Instead of using HUC-specific ARF samples for parameter estimation and regional ARF model development (such as M1 through M5), this approach tries to fit a more generalized GEV that incorporates area and duration as co-variables in the GEV parameters. The ARFs are then calculated using the generalized GEV. The M6 model (Eq. [21] in Section 3.2.2.5) is a function of area A (km^2 or mi^2), duration D (hours), and AEP with ten parameters (a through j).

Note that, theoretically, A^* in the denominator should be 0 (so that ARF represents the ratio of precipitation with area A divided by precipitation with area 0). However, the current formulation, $P(A^*, D, AEP)$, will approach infinity when A approaches 0.⁵ Therefore, A^* needs to be treated as an additional parameter for estimation. Parameters a through j are estimated by the maximum likelihood approach through the same Matlab-based nonlinear programming solver used for M1–M5. Parameter A^* is estimated by minimizing the RMSE between the ARF samples and the ARF model. Similar to the case of M4 model, fitting is needed only once across all frequency levels. An example of a typical M6 fitting is provided in Figure 4-14.

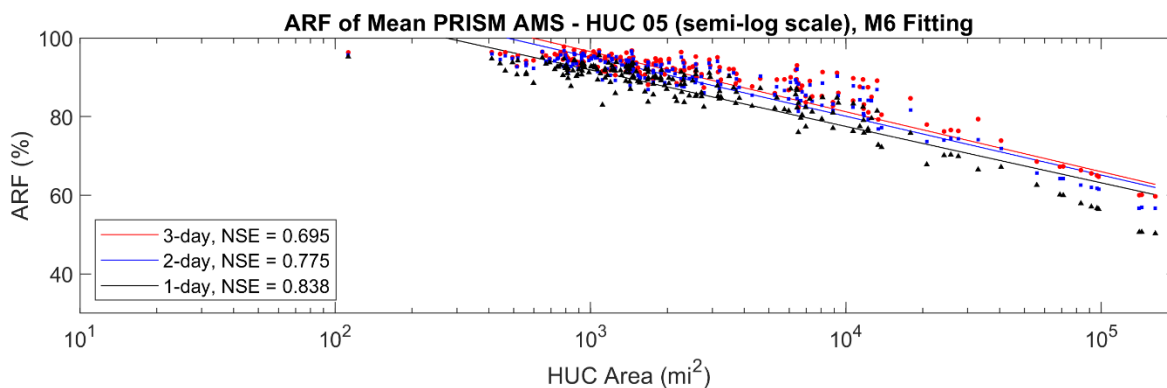


Figure 4-14 Example of M6 ARF Model Fitting

These six ARF models are fitted with different data sources in the three regions. In the national assessment, only the M5 De Michele model is fitted across all hydrologic regions for inter-regional comparison.

It is important to note that when comparing the fitted model results that the complexity of each model can impact model performance and uncertainty. Table 3-1 summarizes the fitting model complexity for M1 through M6. The six selected models each include area as an independent variable, with some including duration and AEP. They also include varying numbers of fitted parameters, ranging from 4 to 10. Since the case studies provide fitted model results as a function of area, duration, and AEP, the number of models fitted for each case study region is a

⁵ Overeem (2010) documents its approach as applicable for areas of 6 km^2 and larger.

function of the number of fitted parameters and the number of independent variable selections for which a unique model is needed.⁶

It is noted that while a model form that includes area, duration, and AEP as independent variables (e.g., M4 and M6) helps reduce the number of models needed, it may yield lower performance since the variation in multiple variables is being explained by a single model. On the other hand, a model form that only includes area (e.g., M3) may yield good performance but requires a larger number of models to assess impacts of duration and AEP on ARF.

In addition, a model which introduces many fitted parameters can increase performance but may not be theoretically justifiable given its complexity and uncertainty.

Table 4-3 Summary of fitting model complexity

Fitting Model	Model Form	Fitted Parameters	# Models Fitted
M1	$ARF = f(A, D)$	3	(3 fitted parameters) × (# AEPs)
M2	$ARF = f(A, D)$	4	(4 fitted parameters) × (# AEPs)
M3	$ARF = f(A)$	4	(4 fitted parameters) × (# durations) × (# AEPs)
M4	$ARF = f(A, D, AEP)$	8	(8 fitted parameters)
M5	$ARF = f(A, D)$	4	(4 fitted parameters) × (# AEPs)
M6	$ARF = f(A, D, AEP)$	10	(10 fitted parameters)

4.4 Use Case Results

Using the data and methods described in Sections 4.2 and 4.3, respectively, the regional assessment results are summarized for Ohio (Section 4.4.1), Mid-Atlantic (Section 4.4.2), and South-Atlantic-Gulf (Section 4.4.3) regions and for the entire CONUS (Section 4.4.4).

4.4.1 Ohio (Region 05)

As described in Section 4.1.1, the regional assessment includes ARF estimates across the Ohio region using different fitting models, data sources, return periods, durations, and seasons. Results demonstrating the effects of these different features are provided below. Detailed results for Region 05 Ohio are provided in APPENDIX C.

4.4.1.1 Effect of Fitting Model

Figure 4-15 provides a comparison of six ARF models (M1–M6) and their NSE fitting statistics in Ohio at 1-day duration and 10-year return period using PRISM. The figure shows variability in the individual, site-specific HUC unit ARFs (black dots) with fitted model curves (colored lines) which generally follow the traditional ARF curve shape. The M1 and M6 models reveal different shapes with worse overall performance. The M2–M5 models produce similar curves, with the full-HUC02 ARF (~422,000 km²; 163,000 mi²) only varying from 0.40 to 0.47. The summary of NSE across different ARF models and frequency levels are further shown in Table 4-4. As noted

⁶ For example, for M1, a unique model is needed for each AEP analyzed since the ARF model is a function of area and duration, but not AEP.

in the Section 4 introduction below Table 4-1, additional return periods beyond 100 years were also analyzed. However, due to the poor fitting model performance, the results of such analyses are not summarized in this report. The issue largely related to the limited observed data record length available from the precipitation products leading to high uncertainty in the model results.

These findings indicate the importance of model selection and establish M1 and M6 as less-preferred options. For M1, the Leclerc and Schaake (1972) equation was originally developed for the TP-29 results, which only provided ARF values or areas less than 1,036 km² (400 mi²). When applying M1 for larger area ARF (e.g., 422,000 km² [163,000 mi²] in Figure 4-15), it appears that M1 would reach a lower bound and cannot provide suitable fitting at the right tail. For M6, the ARF curve would reduce to a linear line at the semi-log scale (lower panel in Figure 4-15) and cannot provide good fit compared to other models.

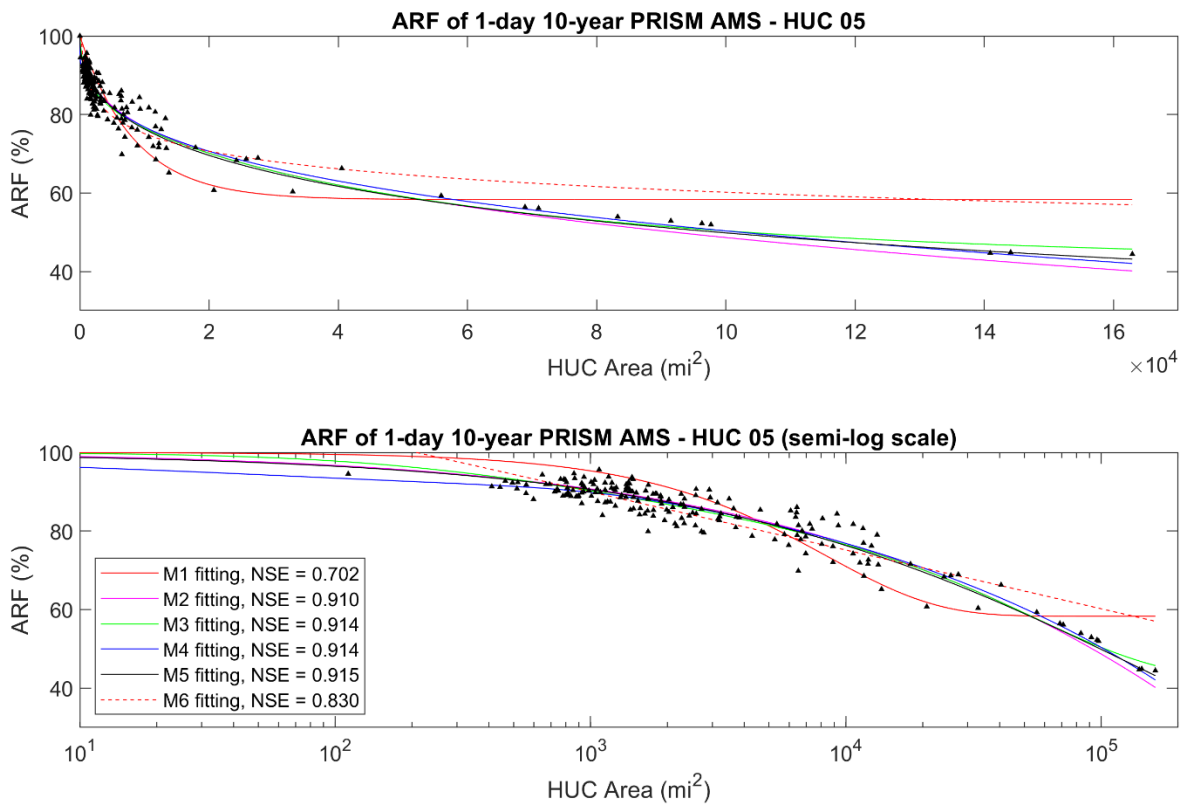


Figure 4-15 Comparison of 1-day, 10-year Ohio ARF Fitting Using PRISM Precipitation Across Different ARF Models *Black dots represent calculated ARFs across HUC units; solid curves represent fitted models for M1–M6. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.*

Table 4-4 Comparison of 1-day Ohio ARF Fitting Using PRISM Precipitation Across Different ARF Models *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE					
	M1	M2	M3	M4	M5	M6
Average AMS (~ 2-year)	0.72	0.93	0.94	0.93	0.94	0.84
10-year	0.70	0.91	0.91	0.91	0.91	0.83
100-year	0.49	0.67	0.69	0.68	0.68	0.62

4.4.1.2 Effect of Data Source

Figure 4-16 provides a comparison of five data sources (PRISM, Daymet, ST4, Livneh, and DSI3240) and their NSE fitting statistics in Ohio at 1-day duration and 10-year return period. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. The ST4 case reveals overall lower ARFs than other cases, with worse overall performance. The models using Daymet, Livneh, PRISM, and DSI-3240 data produce more similar results. The summary of NSE across different precipitation products and frequency levels are further shown in Table 4-5. To better illustrate the effects of data source (and data length) on ARF fitting performance, analysis is extended to 200-, 500-, and 1000-year in Table 4-5.

While the effect of data source is generally smaller than the effect of model, the differences across data sources are still large enough to be non-negligible. In particular, with the increase of return period, the radar-driven ST4 data can lead to significantly different ARFs than when using the gauge-only (DSI-3240) or gauge-driven (Daymet, Livneh, PRISM) precipitation products. This may be a result of shorter ST4 data record length. While ST4 is in hourly time step and can better represent spatial variability (through radar), it only has 16 years of record as compared to 37–64 years of record for other precipitation products.

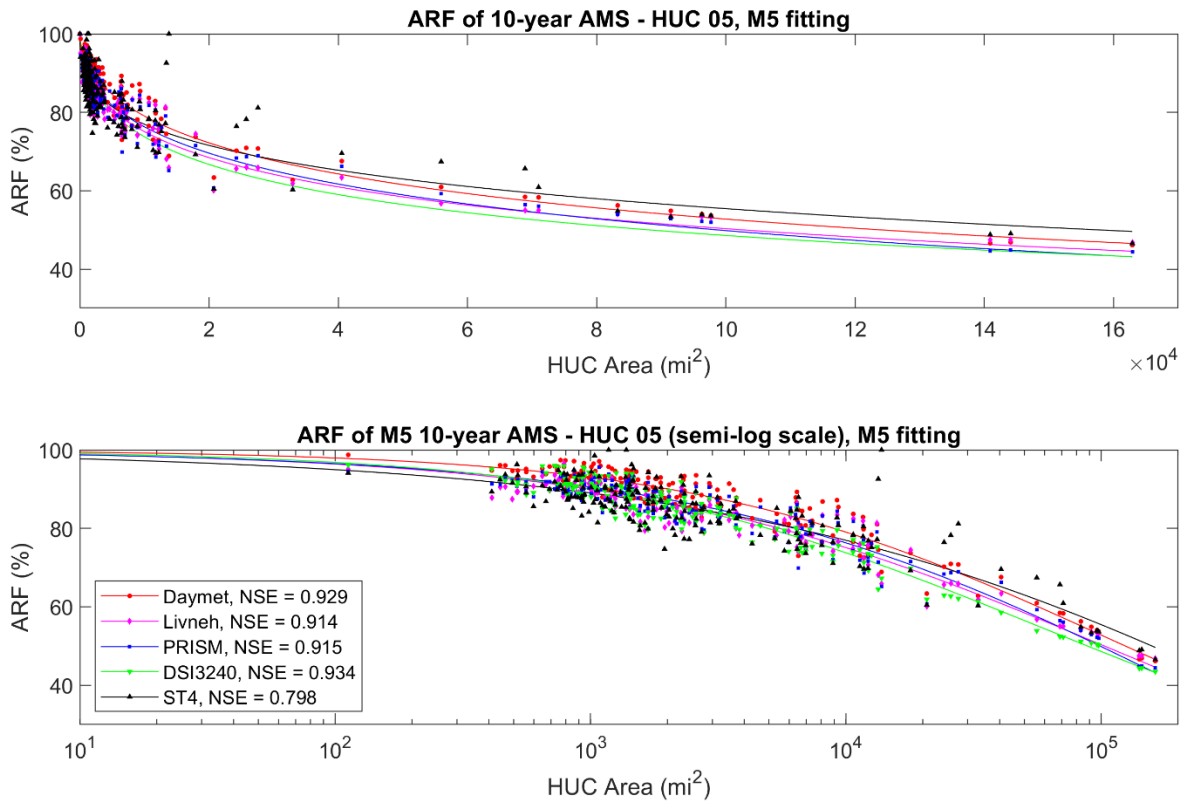


Figure 4-16 Comparison of 1-day, 10-year Ohio M5 ARF Fitted by Different Precipitation Products Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-5 Comparison of 1-day Ohio M5 ARF Fitting across Different Data Sources Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.

Return Period	NSE				
	PRISM (1981–2017)	Daymet (1980–2017)	ST4 (2002–2017)	Livneh (1950–2013)	DSI3240 (1950–2013)
Average AMS (~ 2-year)	0.94	0.95	0.91	0.92	0.95
10-year	0.91	0.93	0.80	0.91	0.93
100-year	0.68	0.74	0.09	0.80	0.85
200-year	0.60	0.66	-0.01	0.73	0.79
500-year	0.58	0.56	0.02	0.63	0.74
1000-year	0.56	0.48	0.08	0.62	0.67

4.4.1.3 *Effect of Return Period*

Figure 4-17 provides the 1-day Ohio M5-fitted ARF and associated performance statistics using PRISM precipitation product across different return periods. Additional fitting statistics across different return periods for other ARF models and precipitation products are shown in Table 4-4, and Table 4-5. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 100-year curve reveals overall lower ARFs than the lower return periods, with worse overall performance.

The model performance results demonstrate the importance of data record length. With the PRISM-based analysis using 37 years of data, estimation of longer return periods (e.g., 100-year and above) becomes more challenging and results in worse performance.

While one can numerically calculate higher return period ARFs (e.g., 1000-year) through fitted GEV distribution for regional ARF model development, it was noticed that the spread across HUCs will become too large to reliably estimate ARF. Considering the limited observation record, it is likely that one may not credibly estimate long return period precipitation and ARF only through data driven approach. Further research that can utilize numerical weather model to assist the development of higher return ARFs should be considered.

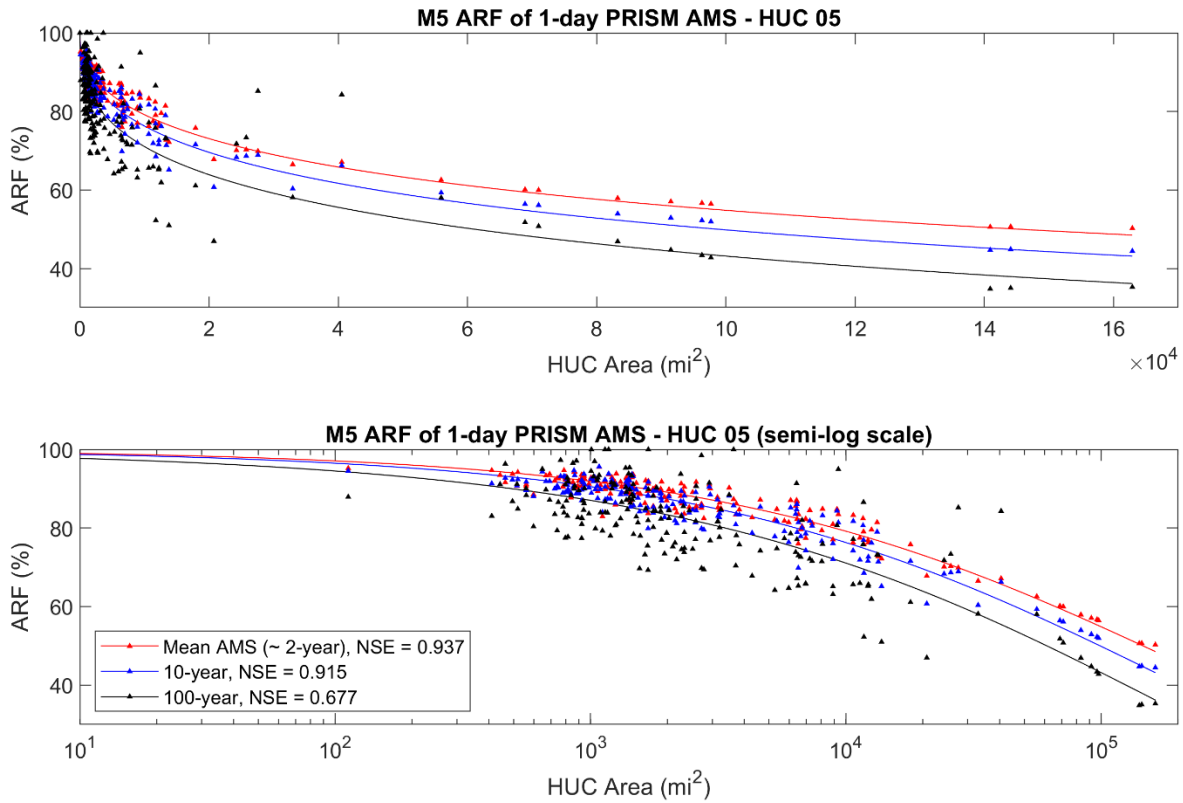


Figure 4-17 Comparison of PRISM-based 1-day Ohio M5 ARF across Different Frequency Levels *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.*

4.4.1.4 Effect of Duration

Figure 4-18 provides the 10-year Ohio M5-fitted ARF and associated performance statistics using PRISM precipitation product across different durations. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 3-day curve reveals overall higher ARFs than the shorter durations, which matches the trend found in literature. The performance also slightly decreases for longer duration. The summary of NSE across different durations and frequency levels are further shown in Table 4-6.

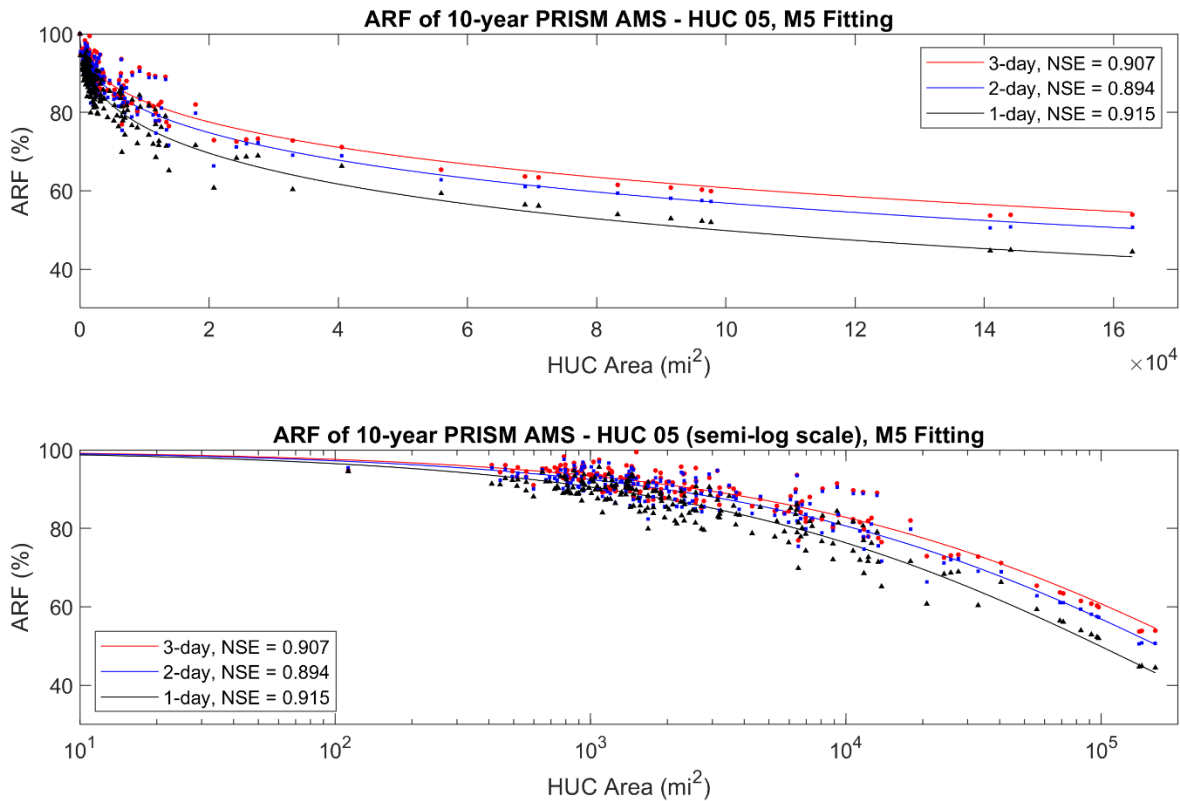


Figure 4-18 Comparison of 10-year Ohio M5 ARF Using PRISM Precipitation across Different Durations Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.

Table 4-6 Comparison of Ohio M5 ARF fitting Using PRISM Precipitation across Different Durations Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.

Return Period	NSE		
	1-day	2-day	3-day
Average AMS (~ 2-year)	0.94	0.93	0.93
10-year	0.91	0.89	0.91
100-year	0.68	0.70	0.80

To gain further insights into the variability of subdaily ARF, in Figure 4-19 provides a comparison of 10-year Ohio M5-fitted ARF using DSI-3240 hourly precipitation across 9 different durations. DSI-3240 is selected for this comparison given its hourly resolution and relatively long record. Similar to the findings reported by NERC (1975), the subdaily ARF decreases with shortening durations. Such analysis can also be performed using the radar-

driven hourly ST4 precipitation, but the limited ST4 period of record would lead to more noisy results especially at higher return levels (not shown).

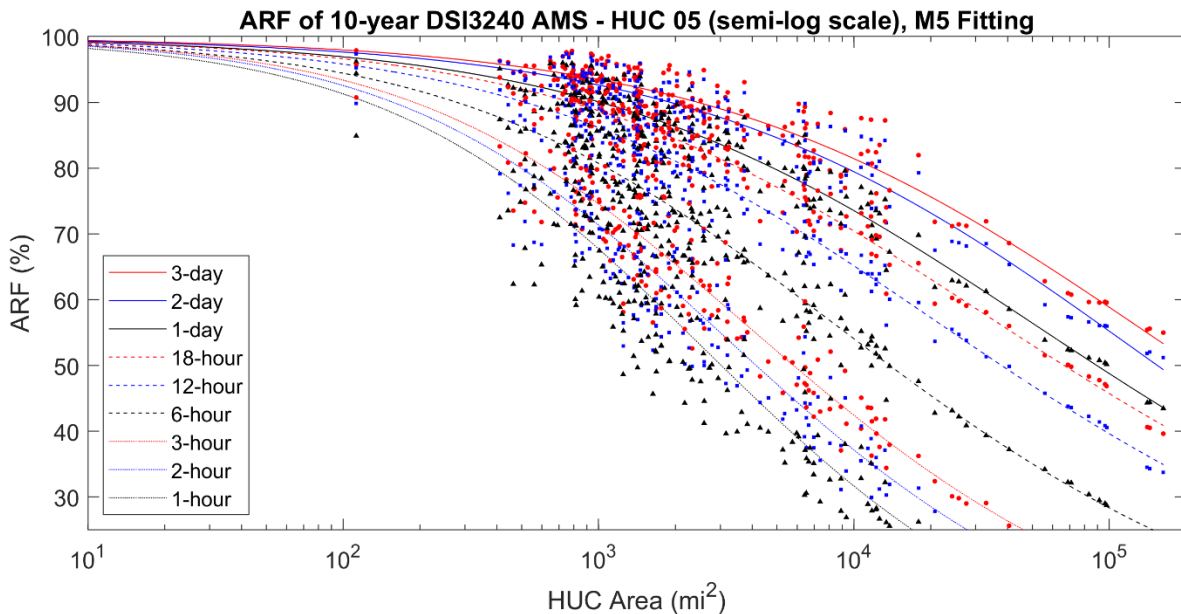


Figure 4-19 Comparison of 10-year Ohio M5 ARF Using DSI-3240 Precipitation across Different Durations Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using DSI-3240.

When performing fitting of M5 model, a further complication was encountered when trying to fit a single set of ARF parameters across both longer (1-, 2-, and 3-day) and shorter (less than 1-day) duration ARF samples. With additional subdaily samples, the performance of fitting would start to decline (not shown). Therefore, instead of fitting a single set of ARF parameters across all durations, in Figure 4-19 each duration is fit separately. The implication is that, for the derivation of a generalized ARF model, one may need to develop separate models for long, medium and short durations.

To provide a consistent comparison in this use case study, 1-, 2-, and 3-day ARF results are analyzed and presented since they are the common durations across different precipitation products. As discussed in Section 2.3, there are fewer subdaily resolution precipitation products to support subdaily ARF analysis. DSI-3240 has longer period of records and the highest point measurement accuracy, but it involves a much larger effort for data processing and spatial interpolation (in particular, considering the effects of topographic during interpolation). The novel radar-driven ST4 precipitation product can better capture the storm structures and has already been in grid format (i.e., easier to process). However, ST4 only has 16 years of records that significantly limits its applicability for higher return period ARF. Moving forward, alternative data approach such as merging of different precipitation products or incorporation of numerically simulated extreme events can be potential areas for further exploration.

4.4.1.5 Effect of Seasonality

Figure 4-20 provides the 1-day 10-year Ohio M5-fitted ARF and associated performance statistics using PRISM precipitation product across different seasons. The figure shows

variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. The summary of NSE across different seasons and frequency levels are further shown in Table 4-7.

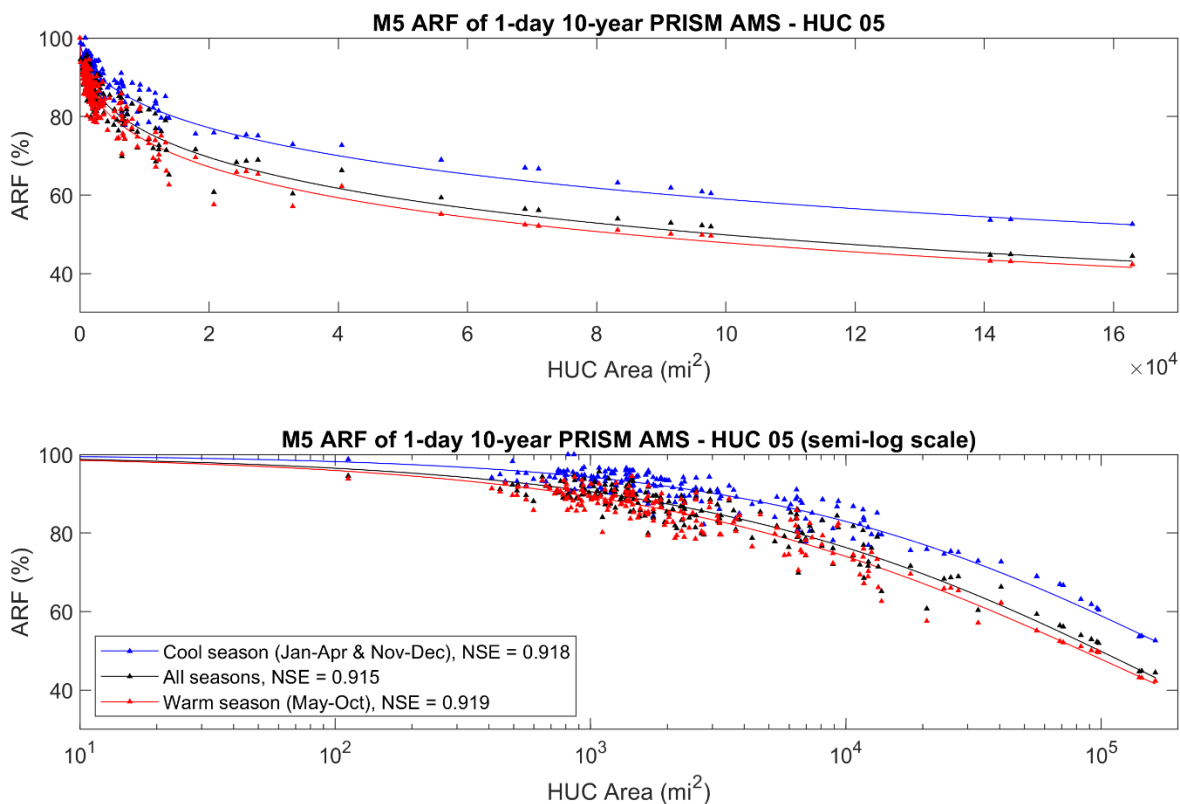


Figure 4-20 Comparison of 1-day, 10-year Ohio M5 ARF Using PRISM Precipitation across Different Seasons Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

The different seasons refer to when the AMS is looked up annually (e.g., warm season AMS is searched during May through October every year). The results suggest that the warm season ARF is close to all season ARF, while cool season ARF has a much higher value. The closeness between warm and all seasons indicates that the annual extreme precipitation in Ohio mainly occurs during warm season. The differences in warm and cool seasons ARF can be explained by their respective controlling extreme precipitation processes. In the Ohio Region, the major extreme precipitation events during warm season are meso-scale convective storms that area generally smaller in size and have larger spatial variability (leading to smaller ARF). On the other hand, the major extreme precipitation events in the Ohio Region during cool season are mostly large-scale frontal systems with relatively smaller spatial variability (as compared to warm season convective storms) that leads to larger ARF. For H&H applications such as simulation of rain-on-snow during cool season, the results suggest that a specific cool season ARF will be needed. The fitting performance of ARF across different seasons is largely similar with cool season having slightly smaller NSE.

Table 4-7 Comparison of 1-day Ohio M5 ARF Fitting Using PRISM Precipitation across Different Seasons *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE		
	All seasons	Warm season (May–Oct)	Cool season (Jan–Apr & Nov–Dec)
Average AMS (~ 2-year)	0.94	0.95	0.94
10-year	0.91	0.92	0.92
100-year	0.68	0.67	0.61

4.4.2 Mid-Atlantic (Region 02)

As described in Section 4.1.1, the regional assessment includes ARF estimates across the Mid-Atlantic region using different fitting models, data sources, return periods, durations, and seasons. Results demonstrating the effects of these different features are provided below. Detailed demonstration results for the Mid-Atlantic Region are provided in APPENDIX D.

4.4.2.1 Effect of Fitting Model

Figure 4-21 provides a comparison of six ARF models (M1–M6) and their NSE fitting statistics in Mid-Atlantic at 1-day duration and 10-year return period using PRISM. The figure shows variability in the individual, site-specific HUC unit ARFs (black dots) with fitted model curves (colored lines) which generally follow the traditional ARF curve shape. The summary of NSE across different ARF models and frequency levels is further shown in Table 4-8.

The models reveal similarities, yet larger differences, compared to the better performance found in the Ohio region. Comparatively speaking, the M1 and M6 models (identified as less-preferred in Ohio) are also some of the worst performers in the Mid-Atlantic Region, while M4 also produces relatively low performance results, though no model produces particularly good results. These findings again indicate the importance of model selection. The relatively poorer performance in Region 02 (than Region 05) can be explained by the wider variability in the calculated HUC unit ARFs. Unlike the single watershed in Region 05 Ohio, Region 02 Mid-Atlantic contains multiple independent HUC04-level watersheds and hence involves larger variabilities. Additionally, the terrain in Region 02 is more complicated than Region 05, and hence involves more types of extreme precipitation processes (e.g., from coastal hurricanes to topographically enhanced precipitation). Despite the weaker fitting, the overall findings across different models are still consistent with general understanding and expectation.

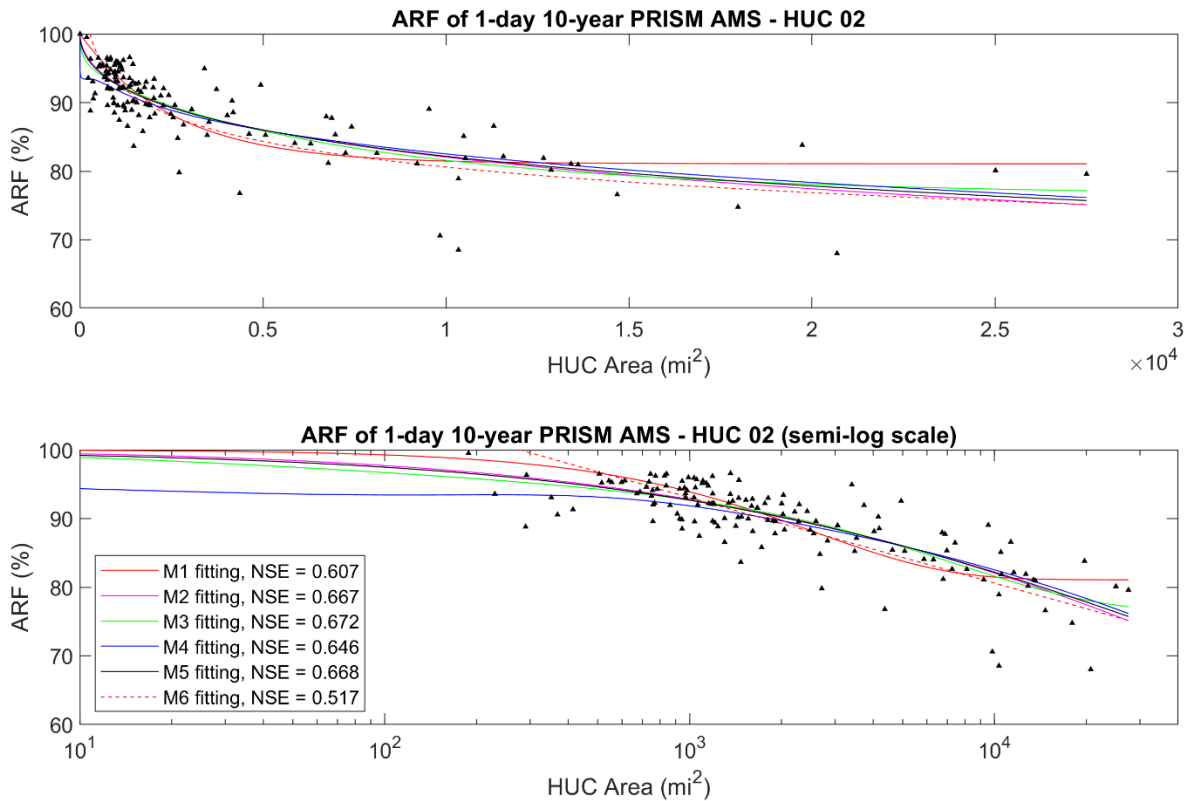


Figure 4-21 Comparison of 1-day, 10-year Mid-Atlantic ARF Fitting Using PRISM Precipitation across Different ARF Models *Black dots represent calculated ARFs across HUC units; solid curves represent fitted models for M1–M6. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.*

Table 4-8 Comparison of 1-day Mid-Atlantic ARF Fitting Using PRISM Precipitation across Different ARF Models *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE					
	M1	M2	M3	M4	M5	M6
Average AMS (~ 2-year)	0.72	0.80	0.81	0.80	0.80	0.65
10-year	0.61	0.67	0.67	0.65	0.67	0.52
100-year	0.12	0.15	0.15	0.11	0.15	0.02

4.4.2.2 Effect of Data Source

Figure 4-22 provides a comparison of four data sources (PRISM, Daymet, ST4, and Livneh) and their NSE fitting statistics in Mid-Atlantic at 1-day duration and 10-year return period (unlike

Ohio, the Mid-Atlantic region did not include DSI-3240). The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. The summary of NSE across different precipitation products and frequency levels are further shown in Table 4-9. To better illustrate the effects of data source (and data length) on ARF fitting performance, the analysis is extended to 200-, 500-, and 1000-year in Table 4-9.

As with the Ohio region results, the ST4 case reveals overall lower ARFs than other cases, with worse overall fitting performance. The models using Daymet, Livneh, PRISM, and DSI-3240 data produce similar but not identical results. This again highlights the need to examine different datasets, as well as the importance of data record length (especially for higher return period estimates). As with the findings noted in Section 4.4.2.1, the overall performance in the Mid-Atlantic region is relatively poorer, owing to variability in calculated HUC unit ARFs. ST4 again provides worse fitting for long return period ARF, due to its relatively shorter period of records.

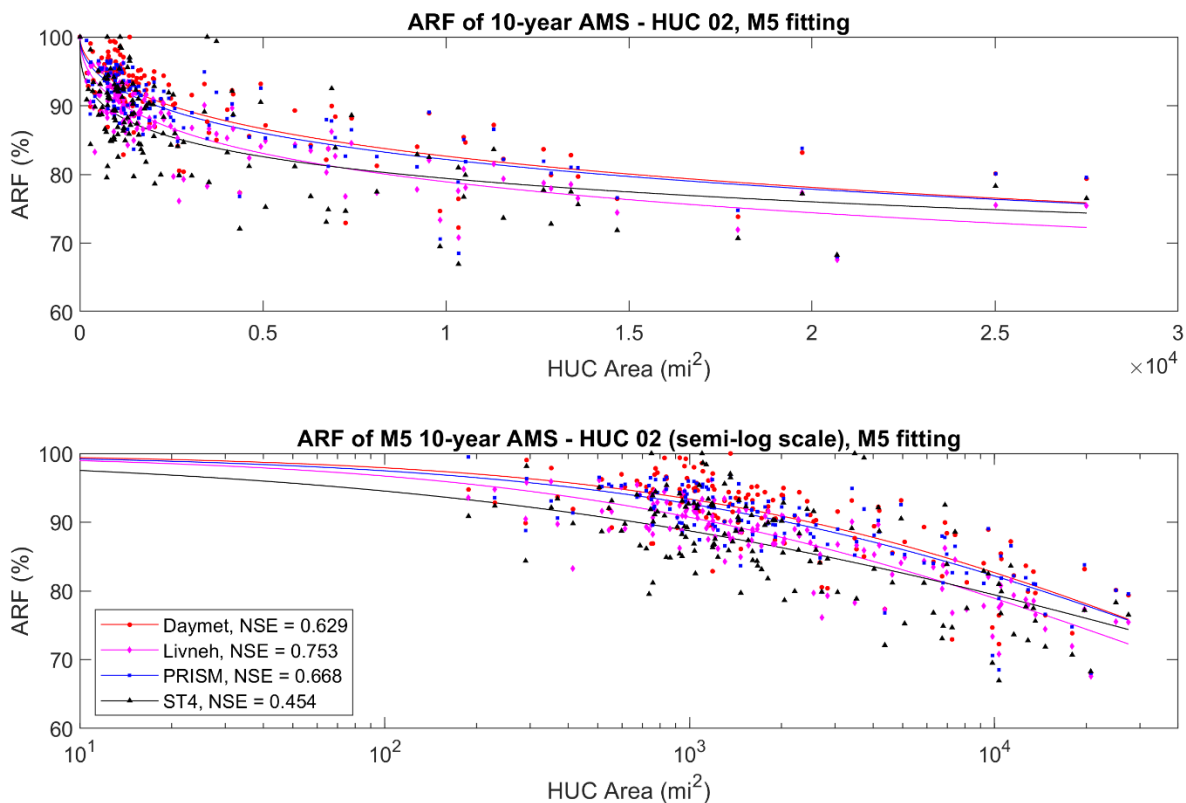


Figure 4-22 Comparison of 1-day, 10-year Mid-Atlantic M5 ARF Fitted by Different Precipitation Products Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-9 Comparison of 1-day Mid-Atlantic M5 ARF Fitting across Different Data Sources *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE				
	PRISM (1981–2017)	Daymet (1980–2017)	ST4 (2002–2017)	Livneh (1950–2013)	DSI3240 (1950–2013)
Average AMS (~ 2-year)	0.80	0.71	0.74	0.83	--
10-year	0.67	0.63	0.45	0.75	--
100-year	0.15	0.17	0.29	0.28	--
200-year	0.12	0.11	0.27	0.18	--
500-year	0.12	0.09	0.25	0.17	--
1000-year	0.09	0.08	0.24	0.14	--

4.4.2.3 Effect of Return Period

Figure 4-23 provides the 1-day Mid-Atlantic M5-fitted ARF and associated performance statistics using PRISM precipitation product across different return periods. Additional fitting statistics across different return periods for other ARF models and precipitation products are shown in Table 4-8, and Table 4-9. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 100-year curve reveals overall lower ARFs than the lower return periods, with worse overall performance. Notably, this worse performance also results in a different shape to the curve, resulting in the 100-year curve and ARF values being higher than the other two curves at high area size. This is a result of poor statistical modeling given limited data.

The model performance results demonstrate the importance of data record length. With the PRISM-based analysis using 37 years of data, estimation of longer return periods (e.g., 100-year and above) becomes more challenging and results in worse performance. The challenges associated the higher return period ARF are similar to the observations in Region 05 Ohio.

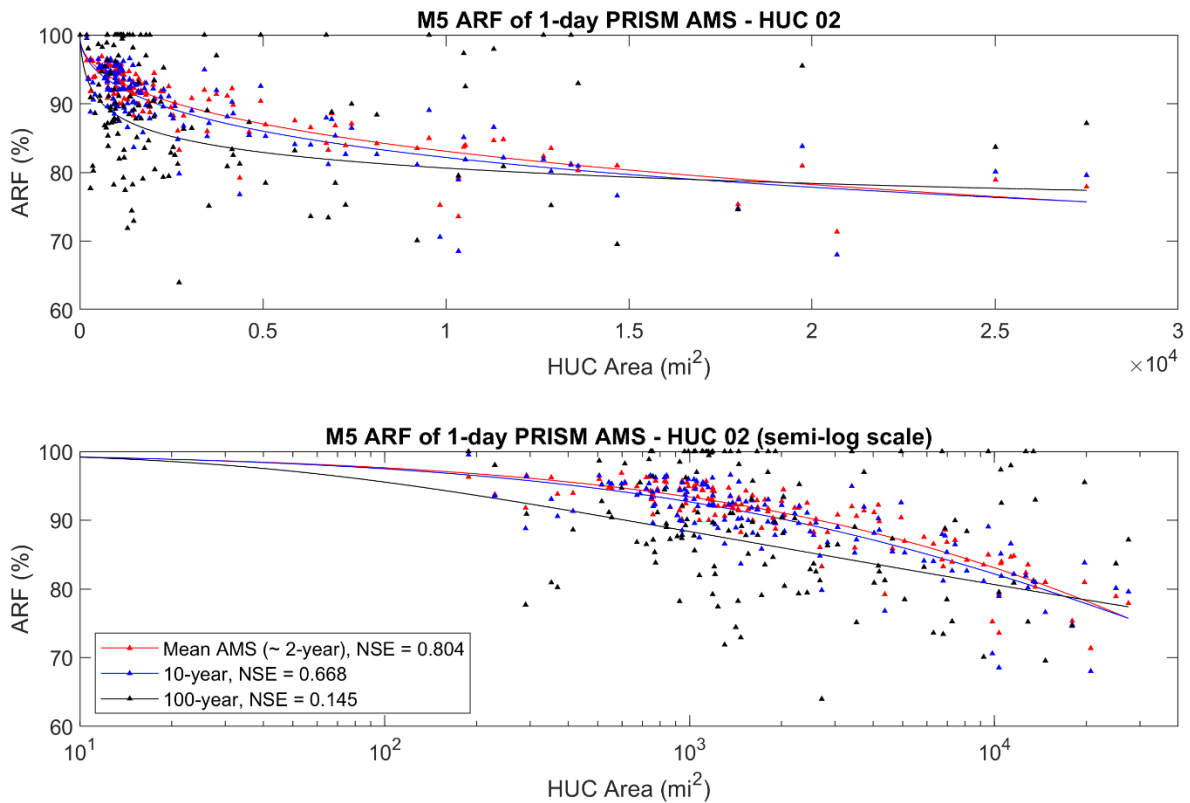


Figure 4-23 Comparison of PRISM-based 1-day Mid-Atlantic M5 ARF across Different Frequency Levels Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.

4.4.2.4 Effect of Duration

Figure 4-24 provides the 10-year Mid-Atlantic M5-fitted ARF and associated performance statistics using PRISM precipitation product across different durations. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 3-day curve reveals overall higher ARFs than the shorter durations, which matches the trend found in literature. Unlike in the Ohio region, the Mid-Atlantic performance slightly increases for longer duration, but the increase is fairly small that may not be meaningful. However, the fitting performance in the Mid-Atlantic is lower than in the Ohio region. The summary of NSE across different durations and frequency levels are further shown in Table 4-10.

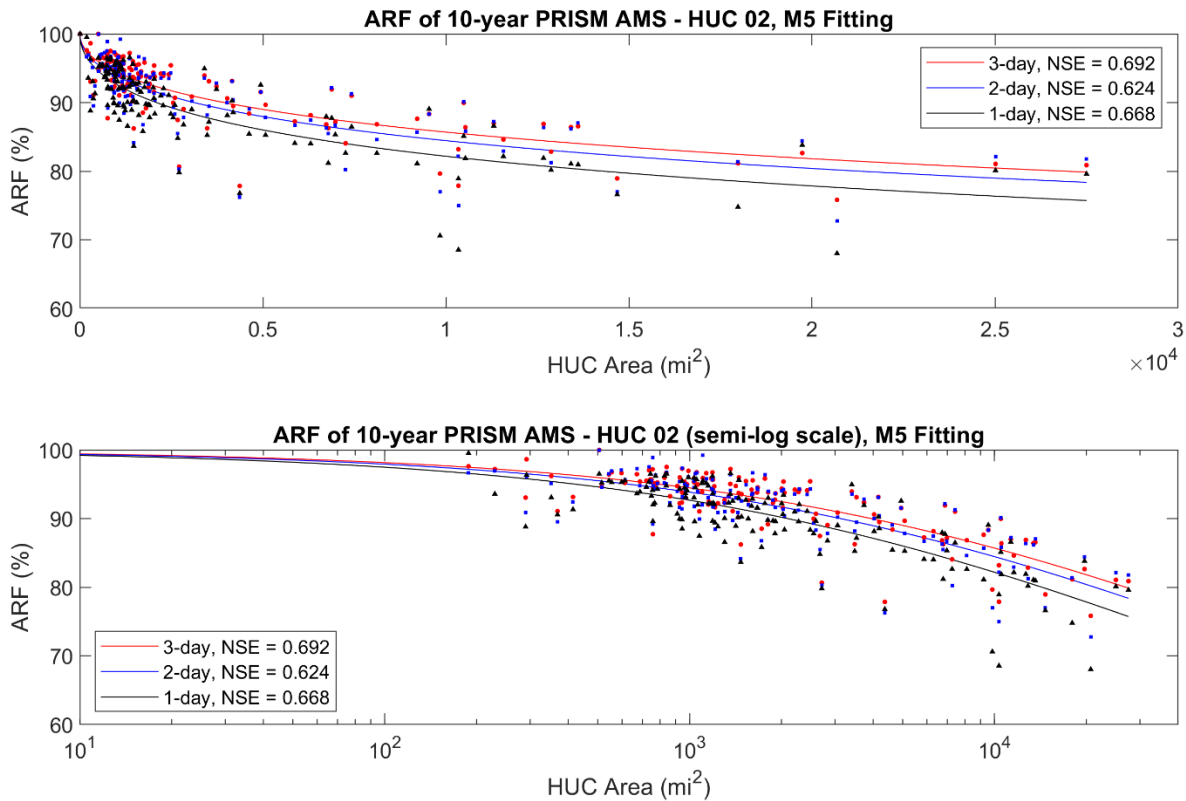


Figure 4-24 Comparison of 10-year Mid-Atlantic M5 ARF Using PRISM Precipitation across Different Durations Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.

Table 4-10 Comparison of Mid-Atlantic M5 ARF Fitting Using PRISM Precipitation across Different Durations Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.

Return Period	NSE		
	1-day	2-day	3-day
Average AMS (~ 2-year)	0.80	0.78	0.79
10-year	0.67	0.62	0.69
100-year	0.15	0.17	0.22

4.4.2.5 Effect of Seasonality

Figure 4-25 provides the 1-day 10-year Mid-Atlantic M5-fitted ARF and associated performance statistics using PRISM precipitation product across different seasons. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves

(colored lines) which follow the traditional ARF curve shape. The summary of NSE across different seasons and frequency levels are further shown in Table 4-11.

The different seasons refer to when the AMS is looked up annually (e.g., warm season AMS is searched during May through October every year). Similar to Ohio, the warm season ARF is close to all season ARF. Cool season ARF is higher than both warm and all seasons, but in Mid-Atlantic the difference is relatively smaller. The closeness between warm and all seasons indicates that the annual extreme precipitation in Mid-Atlantic also mostly occurs during warm season. The difference between warm and cool season ARF can also be explained by the different controlling extreme precipitation processes (e.g., meso-scale versus frontal systems) across seasons. For H&H applications such as simulation of rain-on-snow during cool season, the results also suggest that a specific cool season ARF will be needed in the Mid-Atlantic region. Regarding fitting performance, both warm and all seasons provide worse fitting, particularly at higher return period.

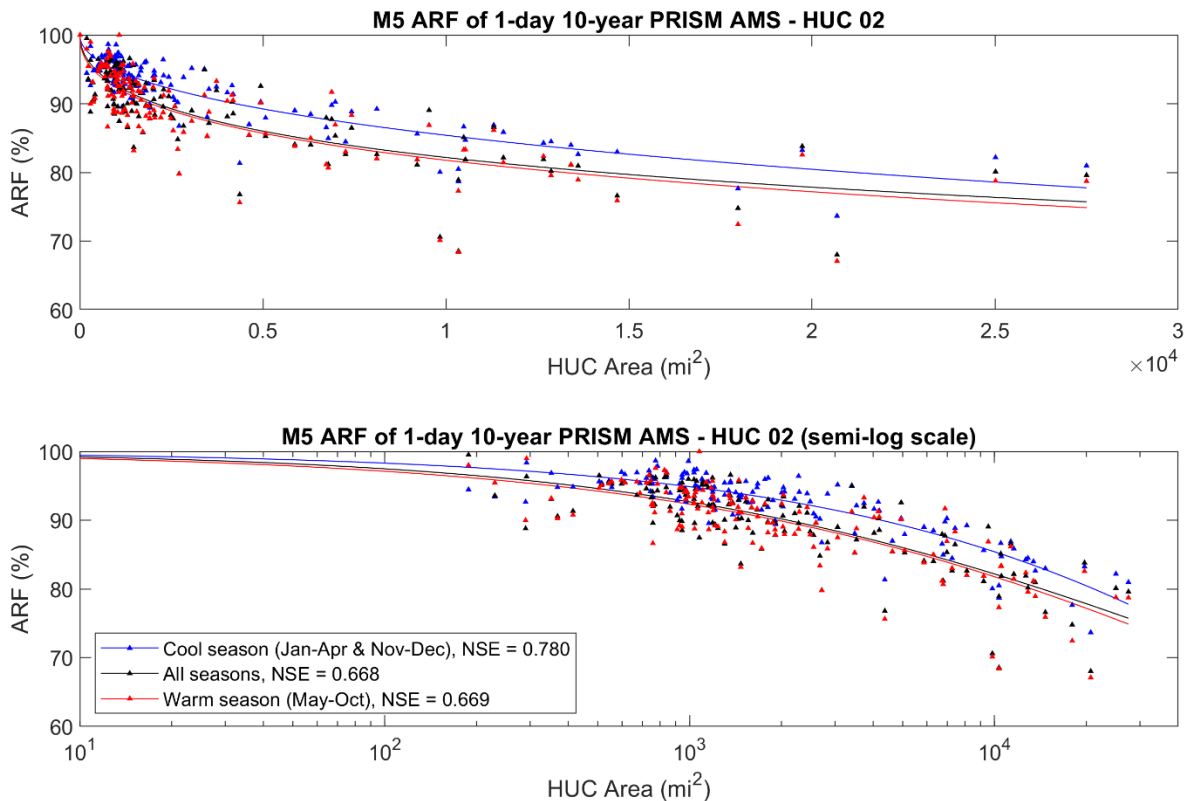


Figure 4-25 Comparison of 1-day, 10-year Mid-Atlantic M5 ARF Using PRISM Precipitation across Different Seasons Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-11 Comparison of 1-day Mid-Atlantic M5 ARF Fitting Using PRISM Precipitation across Different Seasons *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE		
	All seasons	Warm season (May–Oct)	Cool season (Jan–Apr & Nov–Dec)
Average AMS (~ 2-year)	0.80	0.81	0.80
10-year	0.67	0.67	0.78
100-year	0.15	0.12	0.53

4.4.3 South Atlantic-Gulf (Region 03)

As described in Section 4.1.1, the regional assessment includes ARF estimates across the South Atlantic-Gulf region using different fitting models, data sources, return periods, durations, and seasons. Results demonstrating the effects of these different features are provided below. Detailed demonstration results for the South Atlantic-Gulf Region are provided in APPENDIX E.

4.4.3.1 Effect of Fitting Model

Figure 4-26 provides a comparison of six ARF models (M1–M6) and their NSE fitting statistics in South Atlantic-Gulf at 1-day duration and 10-year return period using PRISM. The figure shows variability in the individual, site-specific HUC unit ARFs (black dots) with fitted model curves (colored lines) which generally follow the traditional ARF curve shape. The summary of NSE across different ARF models and frequency levels is further shown in Table 4-12.

The overall model performance in the South Atlantic-Gulf region is better than the Mid-Atlantic region but worse than the Ohio region. Comparatively speaking, the M1 and M6 models (identified as less-preferred in other two regions) are also some of the worst performers in the South Atlantic-Gulf Region. These findings again indicate the variabilities introduced by different models. The relatively weaker performance in Region 03 (than Region 05) can be explained by the wider variability in the calculated HUC unit ARFs. South Atlantic-Gulf is the 2nd largest hydrologic region in the U.S. that contains 18 independent HUC04-level watersheds along the southeastern coastline. Given its wider geographical coverage, it involves multiple types of extreme precipitation processes (e.g., from coastal hurricanes to topographic enhance precipitation near the Smokies) that lead to larger sample ARF variability. Nevertheless, the overall findings across different models are consistent with general understanding and expectation.

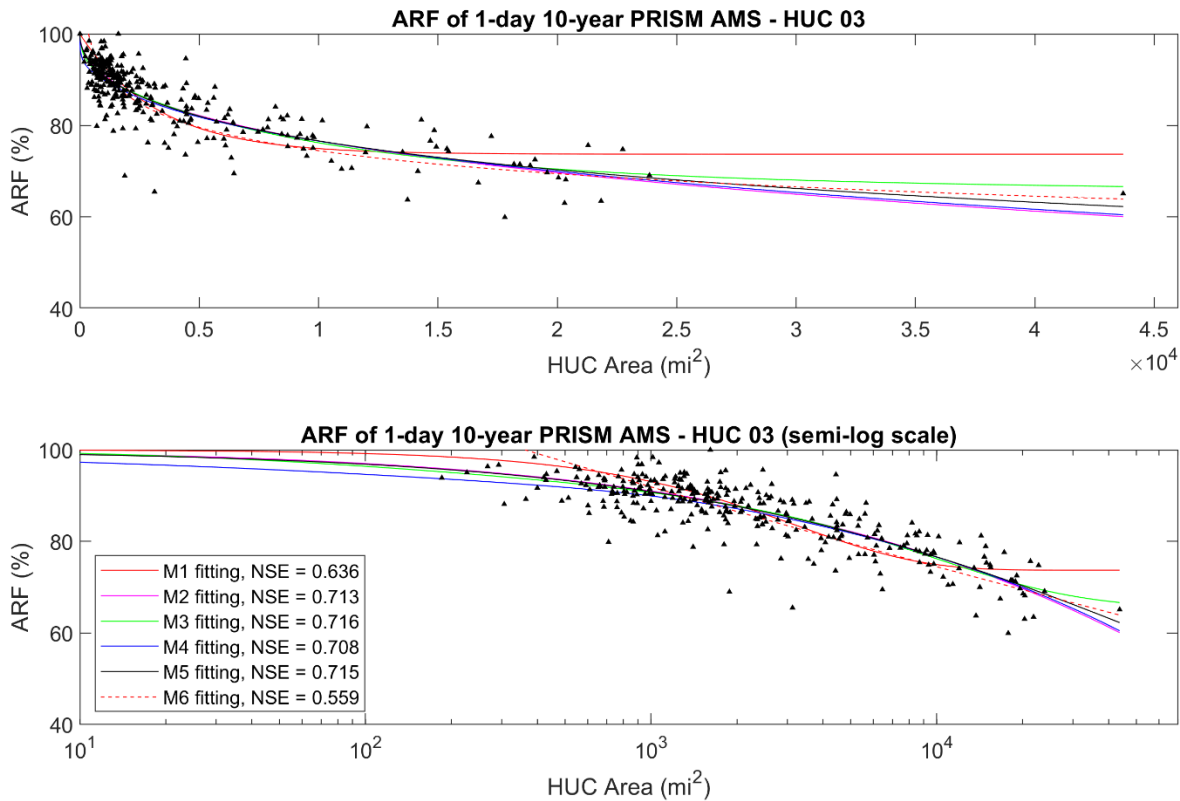


Figure 4-26 Comparison of 1-day, 10-year South Atlantic-Gulf ARF Fitting Using PRISM Precipitation across Different ARF Models Black dots represent calculated ARFs across HUC units; solid curves represent fitted models for M1–M6. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-12 Comparison of 1-day South Atlantic-Gulf ARF Fitting Using PRISM Precipitation across Different ARF Models Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.

Return Period	NSE					
	M1	M2	M3	M4	M5	M6
Average AMS (~ 2-year)	0.64	0.72	0.72	0.71	0.72	0.65
10-year	0.64	0.71	0.72	0.71	0.72	0.56
100-year	0.35	0.43	0.44	0.43	0.44	0.24

4.4.3.2 Effect of Data Source

Figure 4-27 provides a comparison of four data sources (PRISM, Daymet, ST4, and Livneh) and their NSE fitting statistics in South Atlantic-Gulf at 1-day duration and 10-year return period

(unlike Ohio, the Mid-Atlantic region did not include DSI-3240). The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. The summary of NSE across different precipitation products and frequency levels are further shown in Table 4-13. To better illustrate the effects of data source (and data length) on ARF fitting performance, the analysis is extended to 200-, 500-, and 1000-year in Table 4-13.

As with the Ohio and Mid-Atlantic results, the ST4 case reveals overall lower ARFs than other cases, with worse overall fitting performance. The models using Daymet, Livneh, PRISM, and DSI-3240 data produce similar but not identical results. This again highlights the need to examine different datasets, as well as the importance of data record length (especially for higher return period estimates). As with the findings noted in Section 4.4.2.1, the overall performance in the South Atlantic-Gulf region is relatively poorer, owing to the large spatial coverage. ST4 also provides worse fitting for long return period ARF, due to its relatively shorter period of records.

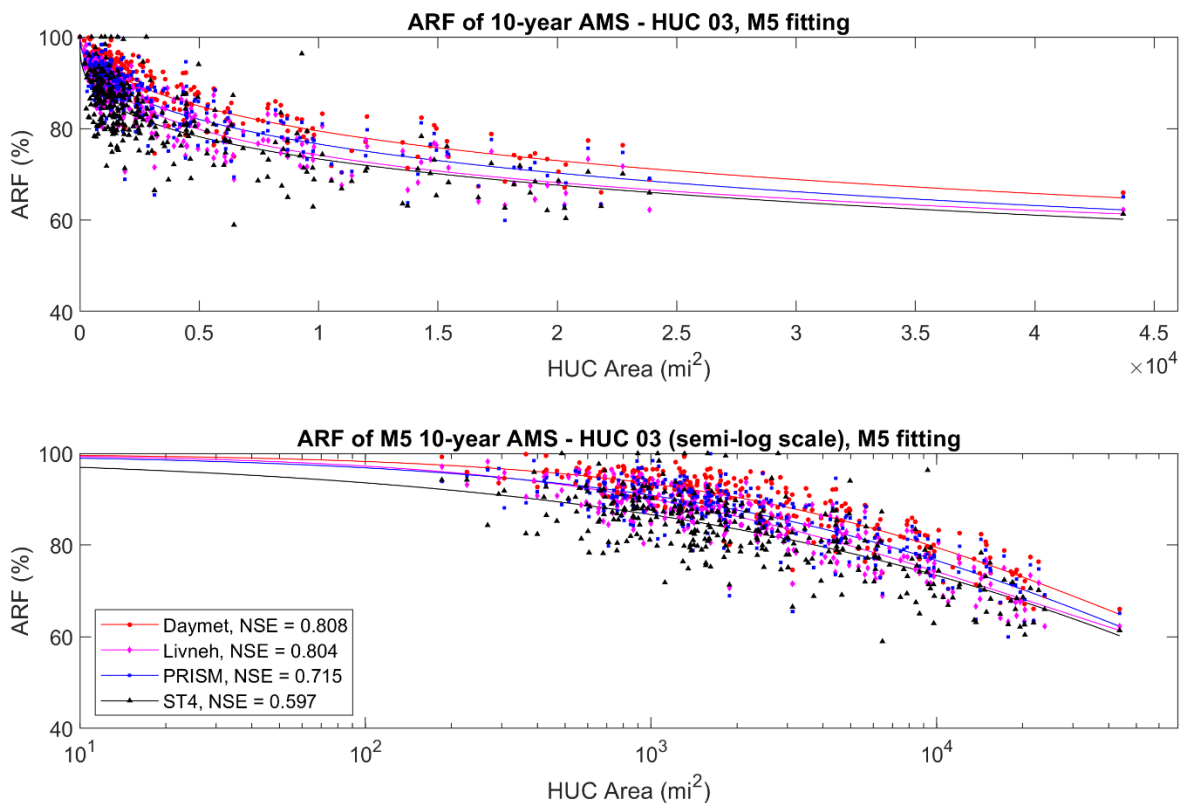


Figure 4-27 Comparison of 1-day, 10-year South Atlantic-Gulf M5 ARF Fitted by Different Precipitation Products Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-13 Comparison of 1-day South Atlantic-Gulf M5 ARF Fitting across Different Data Sources *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE				
	PRISM (1981–2017)	Daymet (1980–2017)	ST4 (2002–2017)	Livneh (1950–2013)	DSI3240 (1950–2013)
Average AMS (~ 2-year)	0.72	0.83	0.72	0.80	--
10-year	0.72	0.81	0.60	0.80	--
100-year	0.44	0.48	0.19	0.62	--
200-year	0.36	0.42	0.15	0.57	--
500-year	0.31	0.38	0.15	0.50	--
1000-year	0.27	0.34	0.14	0.44	--

4.4.3.3 Effect of Return Period

Figure 4-28 provides the 1-day South Atlantic-Gulf M5-fitted ARF and associated performance statistics using PRISM precipitation product across different return periods. Additional fitting statistics across different return periods for other ARF models and precipitation products are shown in Table 4-12, and Table 4-13. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 100-year curve reveals overall lower ARFs than the lower return periods, with worse overall performance.

The model performance results demonstrate the importance of data record length. With the PRISM-based analysis using 37 years of data, estimation of longer return periods (e.g., 100-year and above) becomes more challenging and results in worse performance. The challenges associated the higher return period ARF are similar to the observations in the Ohio and Mid-Atlantic regions.

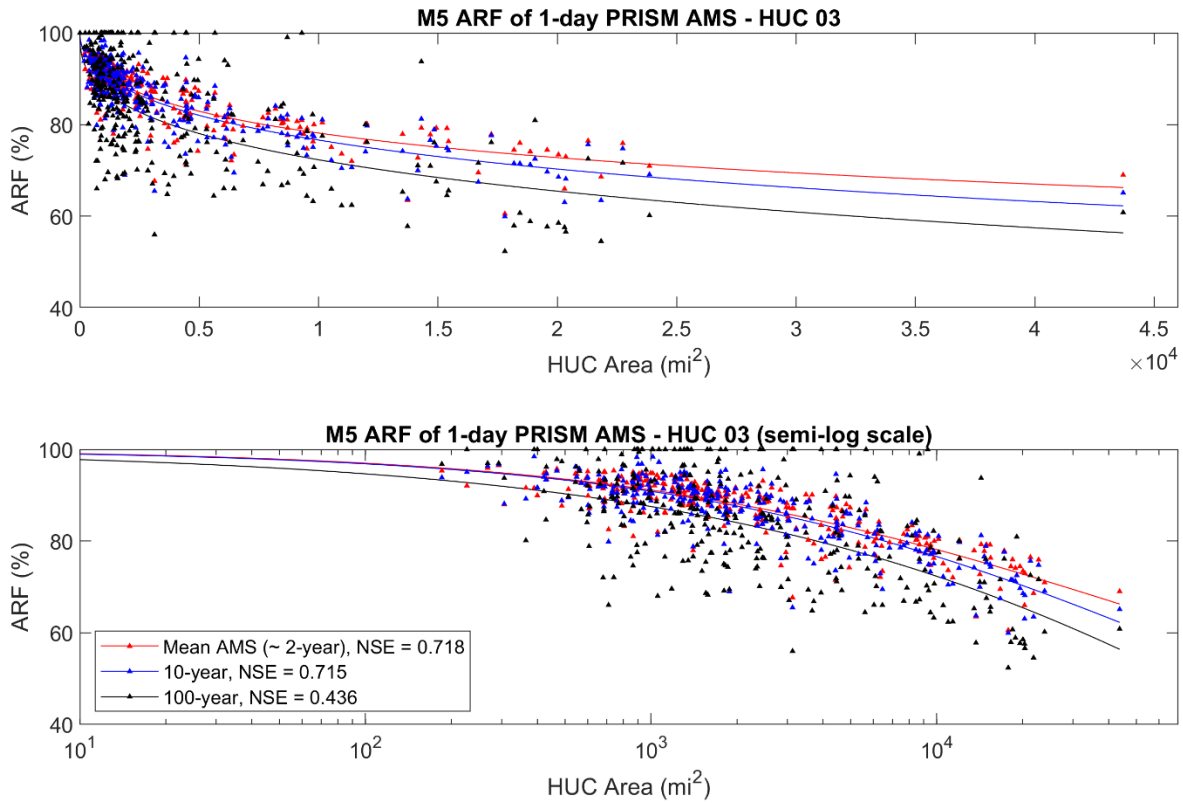


Figure 4-28 Comparison of PRISM-based 1-day South Atlantic-Gulf M5 ARF across Different Frequency Levels Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.

4.4.3.4 Effect of Duration

Figure 4-29 provides the 10-year South Atlantic-Gulf M5-fitted ARF and associated performance statistics using PRISM precipitation product across different durations. The figure shows variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. As expected, the 3-day curve reveals overall higher ARFs than the shorter durations, which matches the trend found in literature. Unlike in the Ohio region, the South Atlantic-Gulf performance slightly increases for longer duration. The summary of NSE across different durations and frequency levels are further shown in Table 4-14.

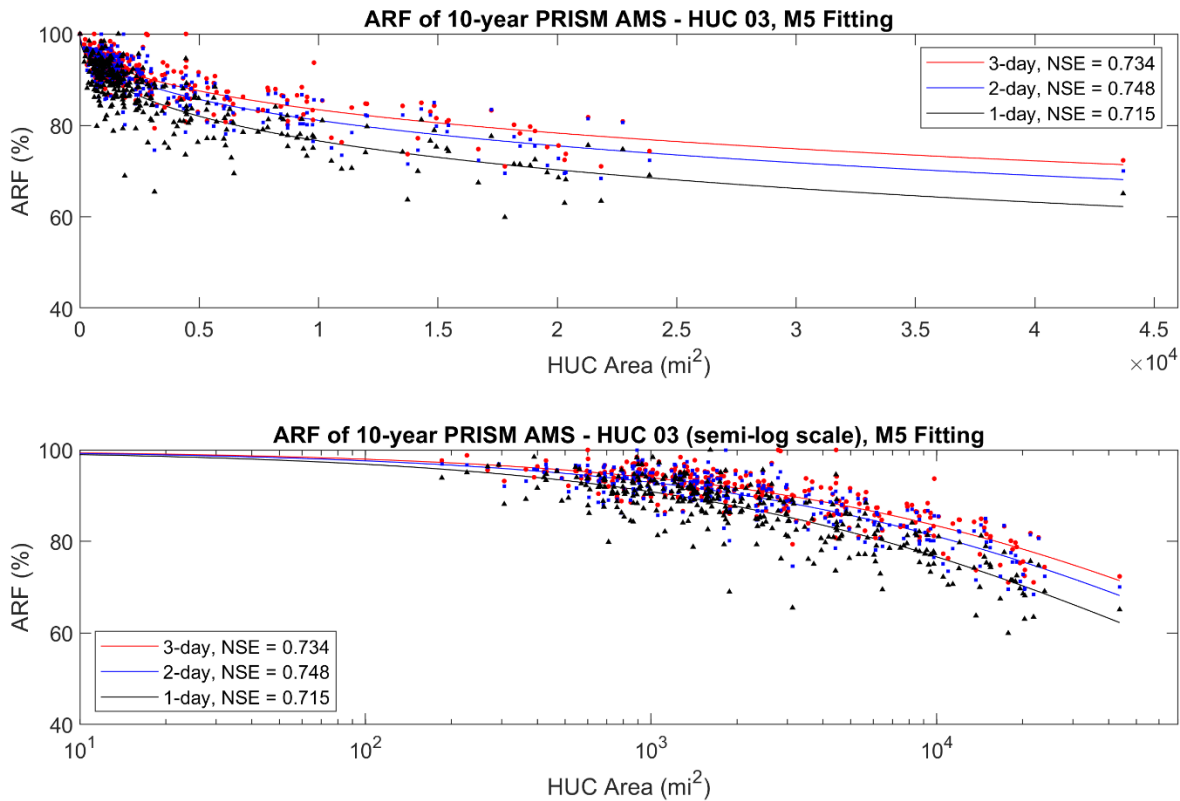


Figure 4-29 Comparison of 10-year South Atlantic-Gulf M5 ARF Using PRISM Precipitation across Different Durations Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using PRISM.

Table 4-14 Comparison of South Atlantic-Gulf M5 ARF Fitting Using PRISM Precipitation across Different Durations Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.

Return Period	NSE		
	1-day	2-day	3-day
Average AMS (~ 2-year)	0.72	0.75	0.75
10-year	0.72	0.75	0.73
100-year	0.44	0.46	0.47

4.4.3.5 Effect of Seasonality

Figure 4-30 provides the 1-day 10-year Mid-Atlantic M5-fitted ARF and associated performance statistics using PRISM precipitation product across different seasons. The figure shows

variability in the individual, site-specific HUC unit ARFs (colored dots) with fitted model curves (colored lines) which follow the traditional ARF curve shape. The summary of NSE across different seasons and frequency levels are further shown in Table 4-15.

The different seasons refer to when the AMS is looked up annually (e.g., warm season AMS is searched during May through October every year). Similar to Ohio and Mid-Atlantic, the warm season ARF is close to all season ARF, and the cool season ARF is higher than others. The closeness between warm and all seasons indicates that the annual extreme precipitation in South Atlantic-Gulf also mostly occurs during warm season. However, the difference between warm and cool season ARF is the smallest among all regions, suggesting that the effect of seasonality is less significant in the South Atlantic-Gulf region. This finding is consistent with the regional climate pattern of South Atlantic-Gulf that it's the warmest among all three regions. Given that there is much less annual snowpack in this region, and the smaller differences between warm and cool season ARF, there may not be a significant need to develop season-specific ARF for H&H applications.

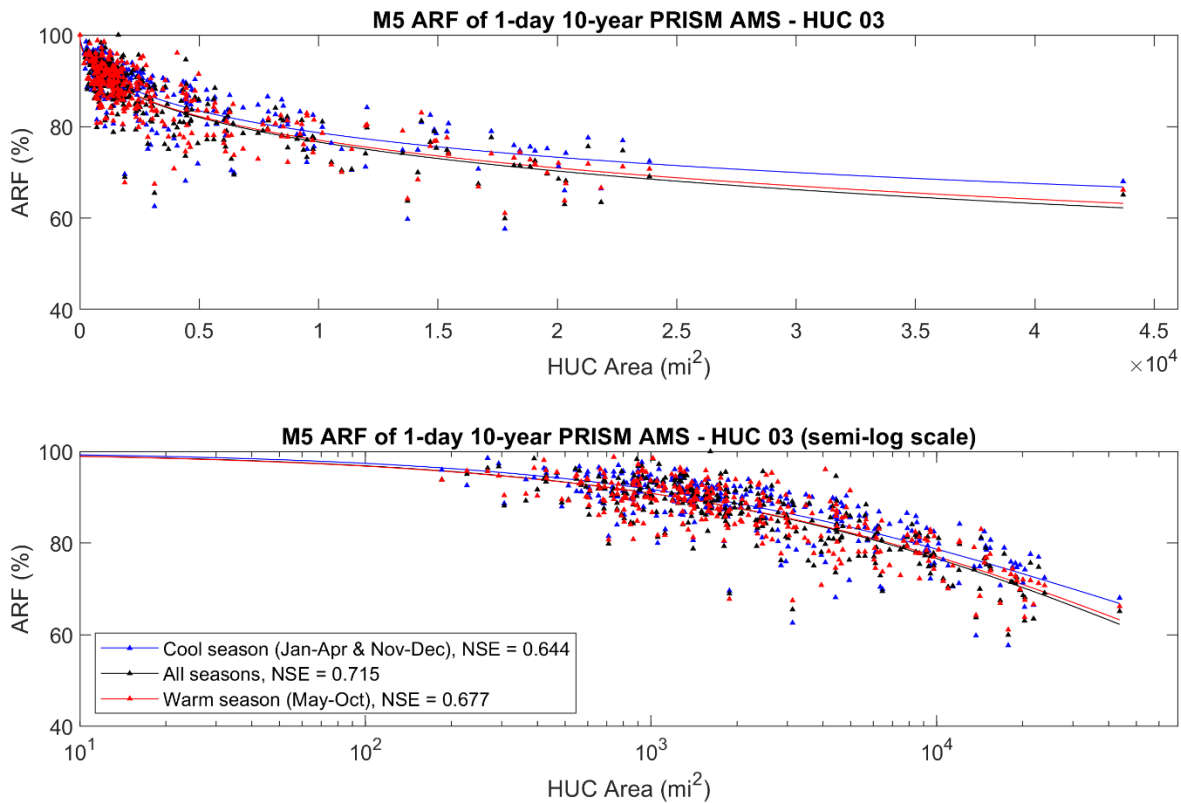


Figure 4-30 Comparison of 1-day, 10-year South Atlantic-Gulf M5 ARF Using PRISM Precipitation across Different Seasons Colored dots represent calculated ARFs across HUC units; solid curves represent fitted M5 models using different precipitation products. Top and bottom panels include the same information in linear (top) and semi-log (bottom) scale.

Table 4-15 Comparison of 1-day South Atlantic-Gulf M5 ARF Fitting Using PRISM Precipitation across Different Seasons *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance.*

Return Period	NSE		
	All seasons	Warm season (May–Oct)	Cool season (Jan–Apr & Nov–Dec)
Average AMS (~ 2-year)	0.72	0.75	0.67
10-year	0.72	0.68	0.64
100-year	0.44	0.35	0.41

4.4.4 CONUS Reconnaissance-level Assessment

As described in Section 4.1.2, the national assessment includes ARF estimates at the HUC02 level across all CONUS regions. These estimates are produced using daily PRISM precipitation data and the M5 De Michele fitting model. Additional demonstration results for the national assessment are provided in APPENDIX F.

4.4.4.1 Effect of Geographic Location

Figure 4-31 shows example ARF curves across the 18 CONUS HUC02 regions for 1-day duration and 100-year return period. The maximum area plotted for each region represents the largest HUC unit analyzed in the region. The results demonstrate that while ARF values are somewhat consistent for small area sizes (e.g., ARFs are within about 6% for areas below 259 km² [100 mi²]), the variation increases for larger areas. For example, at 25,900 km² (10,000 mi²), the fitted ARFs range from approximately 0.85 in the New England region (Region 01) to 0.56 in the Souris-Red-Rainy (Region 09) and Texas-Gulf (Region 12) regions. This regional variation is illustrated in Figure 4-32, which generally shows that lower ARFs are found in the central U.S., with higher ARFs found in the eastern and western U.S. Several factors could contribute to this pattern, but it is expected that regional climate, coastal proximity, topographical influences, and frequency of certain storm types/intensities could influence AMS values across regions and influence ARF estimates. It is worth noting that the results in Figure 4-31 and Figure 4-32 are for a 100-year return period based on 37 years of data; thus, the fitting performance (summarized in Table 4-16 and Table 4-17) is relatively poorer than for lower return periods.

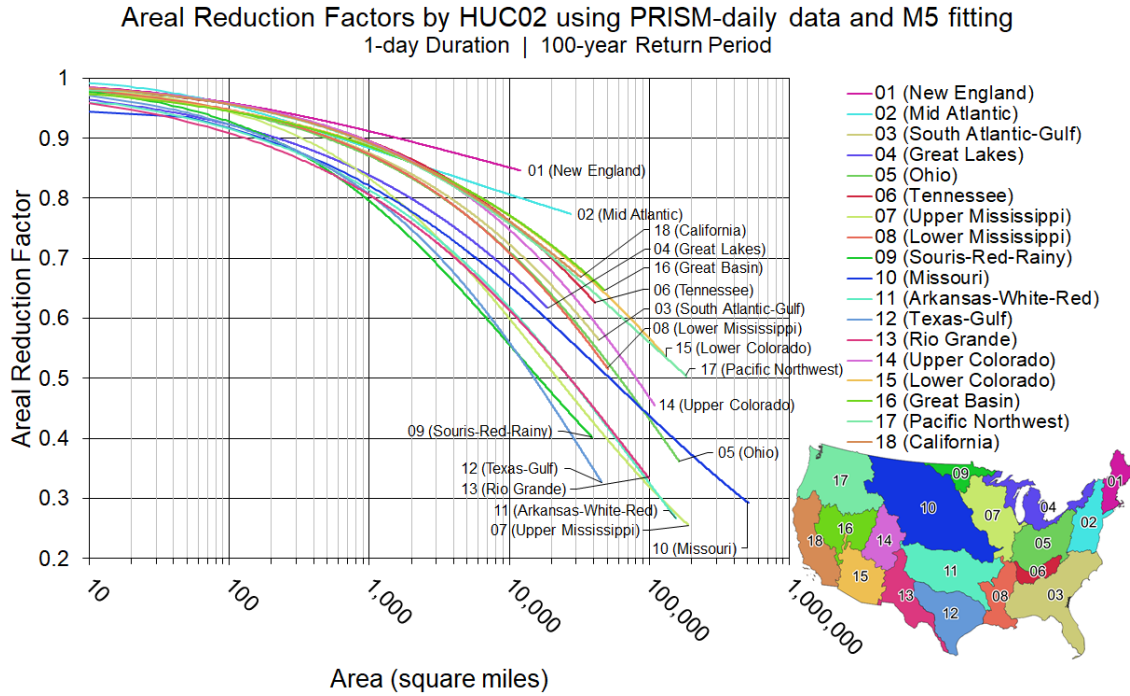


Figure 4-31 CONUS ARF Assessment for 1-day Duration and 100-year Return Period Using PRISM-Daily Data and M5 Model Fitting

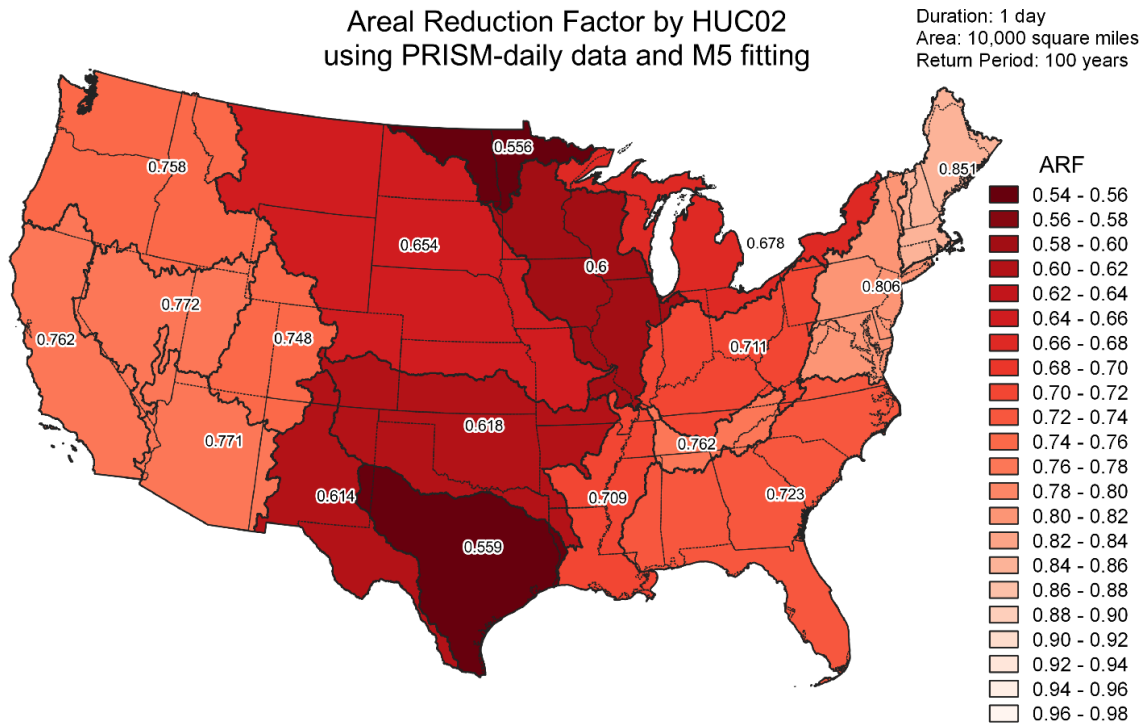


Figure 4-32 CONUS ARF Assessment for 1-day Duration, 10,000-mi² Area, and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting

4.4.4.2 Effect of Return Period

Figure 4-33 shows various maps of HUC02-level ARFs for 1-day duration. The maps are organized to demonstrate the influence of increasing return period when moving from left to right. Example plots are provided for different area sizes of 1,036 km² (400 mi²; top row), 5,180 km² (2,000 mi²; middle row), and 25,900 km² (10,000 mi²; bottom row). The results show a tendency for ARFs to decrease with increasing return period and to decrease with increasing area. The same general geographic pattern as discussed in Section 4.4.4.1 is also seen in these maps. Table 4-16 summarizes the regional NSE model performance statistics associated with the 1-day fittings. As expected, the performance degrades for higher return periods. The NSE values also vary regionally, with the lowest values found in regions 01 (New England), 02 (Mid-Atlantic), 04 (Great Lakes), and 16 (Great Basin). While regions 01, 02, and 04 are all in the northeastern U.S., a clear explanation for the performance results is not apparent and would require further investigation.

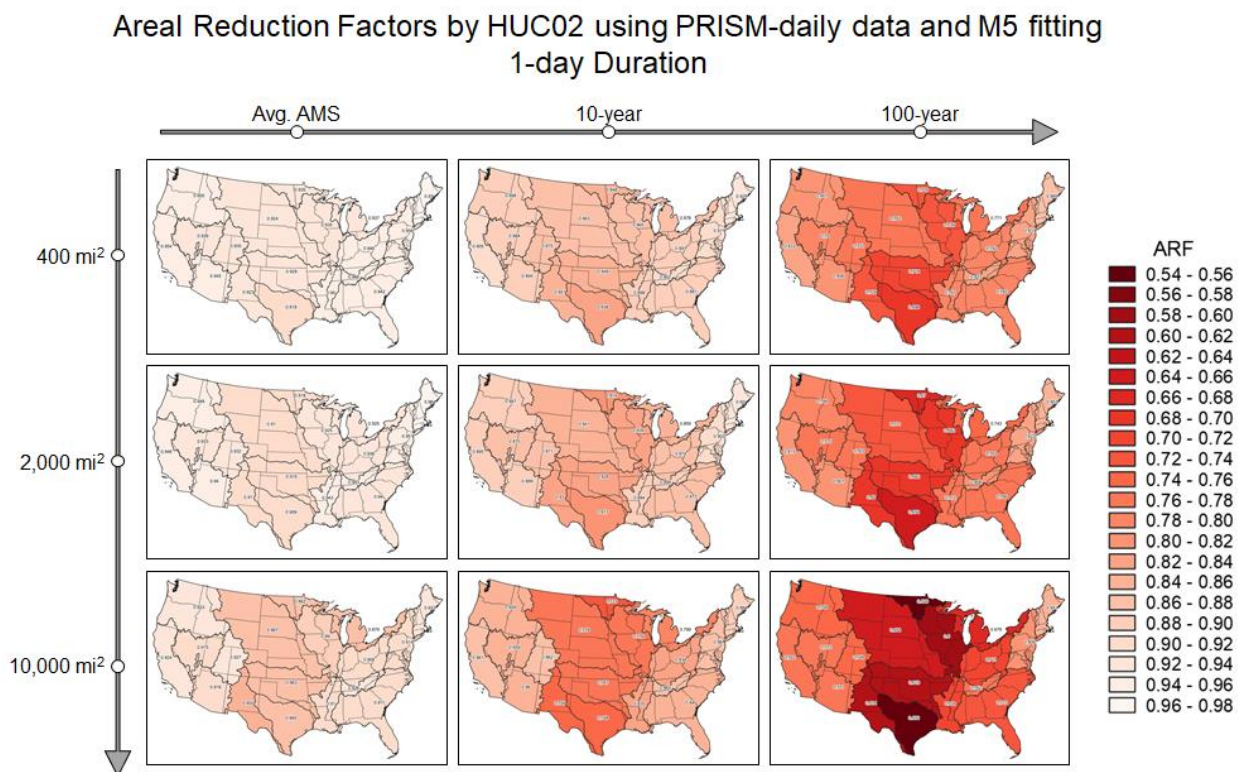


Figure 4-33 Map of CONUS 1-day ARFs Using PRISM-daily Data and M5 Model Fitting across Different Areas and Return Periods

Table 4-16 Comparison of 1-day CONUS Regional M5 ARF Fitting Using PRISM Precipitation across Different Return Periods *Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance. Note that the region numbers correspond to Figure 4-4, with lower numbers generally in the eastern U.S. and higher numbers generally in the western U.S.*

Return Period	NSE																	
	Region Number																	
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
Avg. AMS	0.68	0.80	0.72	0.69	0.94	0.91	0.93	0.87	0.88	0.85	0.87	0.88	0.92	0.83	0.84	0.81	0.85	0.72
GEV 10-yr	0.66	0.67	0.72	0.58	0.91	0.89	0.90	0.83	0.85	0.78	0.81	0.89	0.90	0.81	0.79	0.77	0.84	0.74
GEV 100-yr	0.20	0.15	0.44	0.31	0.68	0.46	0.72	0.59	0.73	0.57	0.59	0.70	0.72	0.65	0.51	0.37	0.70	0.63

4.4.4.3 Effect of Area and Duration

Figure 4-34 shows various maps of HUC02-level ARFs for 100-year return period. The maps are organized to demonstrate the influence of increasing duration when moving from left to right. Example plots are provided for different area sizes of 1,036 km² (400 mi²; top row), 5,180 km² (2,000 mi²; middle row), and 25,900 km² (10,000 mi²; bottom row). The results show a tendency for ARFs to increase with increasing duration and to decrease with increasing area. The same general geographic pattern as discussed in Section 4.4.4.1 is also seen in these maps. Table 4-17 summarizes the regional NSE model performance statistics associated with the 100-year fittings. The results demonstrate relatively low differences in performance across different durations. However, the NSE values vary regionally, with the lowest values found in regions 01 (New England), 02 (Mid-Atlantic), 03 (South Atlantic-Gulf), 04 (Great Lakes), and 16 (Great Basin). While regions 01, 02, and 04 are all in the northeastern U.S., a clear explanation for the performance results is not apparent and would require further investigation. The relatively poorer performance in certain regions can be explained by the wider variability in the calculated HUC unit ARFs. One possible explanation for the wider variability is that these regions contain multiple independent HUC04-level watersheds and hence involves larger variabilities. It is also possible that terrain effects could play a role and that a mixture of more extreme precipitation processes (e.g., from coastal hurricanes to topographically enhanced precipitation).

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting
100-y Return Period

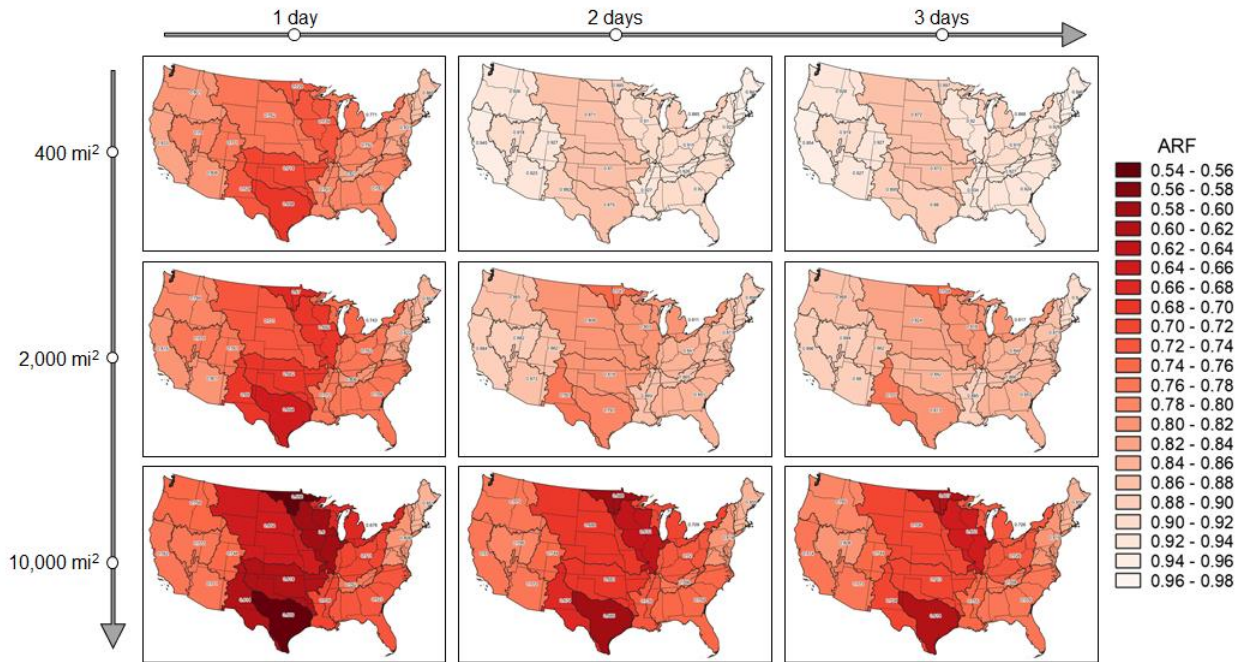


Figure 4-34 Map of CONUS 100-year ARFs Using PRISM-daily Data and M5 Model Fitting across Different Durations and Areas

Table 4-17 Comparison of 100-year CONUS Regional M5 ARF Fitting Using PRISM Precipitation across Different Durations Cell coloration indicates relative performance, with darker red colors indicating worse performance and white colors indicating better performance. Note that the region numbers correspond to Figure 4-4, with lower numbers generally in the eastern U.S. and higher numbers generally in the western U.S.

Duration	NSE																	
	Region Number																	
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
1-day	0.20	0.15	0.44	0.31	0.68	0.46	0.72	0.59	0.73	0.57	0.59	0.70	0.72	0.65	0.51	0.37	0.70	0.63
2-day	0.29	0.17	0.46	0.33	0.70	0.67	0.70	0.58	0.70	0.59	0.62	0.71	0.73	0.62	0.57	0.50	0.65	0.60
3-day	0.28	0.22	0.47	0.30	0.80	0.65	0.72	0.67	0.72	0.60	0.68	0.66	0.78	0.70	0.50	0.43	0.58	0.58

5 SUMMARY

To support PFHA of NPPs, probabilistic estimates of extreme rainfall depth across various watershed sizes are required. Nevertheless, most existing PFA products (such as NOAA Atlas 14) provide frequency estimates of “point” precipitation that can only be representative for a small domain and are not appropriate for large-scale watershed modeling applications. The ARF examined in this study, which is the ratio of areal extreme rainfall depth to point-based extreme rainfall depth, is one commonly-used approach to derive areal extreme rainfall estimates from conventional point-based PFA products. The use of ARF is necessary because high spatiotemporal resolution precipitation observations with long period of records, which are needed for accurate areal rainfall frequency estimation, are generally lacking and do not allow for an appropriate characterization of the associated spatial rainfall patterns.

However, compared to modern PFA products, the progress of ARF development in the U.S. is relatively slow, and the TP-29 ARFs published in the 1950s are still used in practice. To improve understanding of ARF variabilities across different precipitation products, ARF models, return periods, geographical locations, and seasons, this study conducts a comprehensive review of recent ARF research, summarizes potential precipitation products for ARF applications, and provides use case studies to demonstrate the derivation of ARF in several selected hydrologic regions in the U.S.

This research is part of the NRC’s PFHA Research Program and is to assist NRC in assessing different classes of fixed-area ARF methods in conjunction with available rainfall datasets to support development of guidance for application of NPP-PFHA. The work will aid the development of guidance on the use of PFHA methods and support risk-informing NRC’s licensing framework (flood hazard design standards at proposed new facilities as well as significance determination tools for evaluating potential deficiencies related to flood protection at operating facilities). The tools and guidance developed will support and enhance NRC’s capacity to perform thorough and efficient reviews of license applications and license amendment requests. They will also support risk-informed significance determination of inspection findings, unusual events and other oversight activities.

The discussion of specific references, methods, software, or tools in this NUREG/CR does not constitute an endorsement or approval for any specific use by the U.S. Nuclear Regulatory Commission or Oak Ridge National Laboratory. The case study results presented herein are the result of research efforts only, do not incorporate uncertainty quantification, and should not be directly incorporated for application. They are intended to demonstrate some of the primary factors affecting areal reduction factor estimation.

The main findings and recommendations are summarized in this section.

5.1 ARF Characteristics in the United States

Overall, the most general observations that can be made across all methods, precipitation products, and regions are:

- ARF decreases with decreasing duration
- ARF decreases with increasing area
- ARF decreases with increasing return period (Figure 4-33)

In particular, the use case study results clearly suggest that ARF decreases with increasing return period (which was not provided in the conventional TP-29 ARF curves). This supports the need to develop frequency- (return period) specific ARF for H&H application. However, the results also suggest that the performance of long return period ARF is strongly controlled by the length of historic observation (e.g., the poor performance of 100-year ST4 ARF fitting given only 16 years of observation). Therefore, when using purely observation-based approach (such as the one used in this study), it is questionable if one has sufficient ability and confidence to derive long-return period areal extreme rainfall and ARF estimates to support H&H application. This limitation should be clearly acknowledged when selecting and designing PFHA for NPP. Unless the data sufficiency issue is addressed, long return period ARFs should be estimated using more conservative (and reliable) short return period ARFs. Process-based numerical weather simulation can be a promising alternative to resolve this issue.

Regarding ARF methods, the difference across different ARF models is found to be one factor affecting the estimation of regional ARF. In general, M2 (Koutsoyiannis and Xanthopoulos UK-NERC Model), M3 (Hydrological Atlas of Switzerland Model), M4 (Australian ARR Model), and M5 (De Michele Dynamic Scaling Model) provide better fitting. While M3 (Hydrological Atlas of Switzerland Model) can fit well, it does not include duration as a variable and hence can be more sensitive to sample size and data quality. M4 (Australian ARR Model) is more difficult to fit (8 parameters), but it includes frequency levels in the model and can be overall more robust. M5 (De Michele Dynamic Scaling Model) can fit well and has a good underlying theory, and hence is selected in the CONUS reconnaissance level assessment. Given that M1 (Leclerc & Schaake TP-29 Model) was developed for smaller area TP-29 values, it cannot provide good fitting when considering larger area ARF samples (which are likely be the case for many U.S. NPPs). While M6 (Regional GEV Model) has a good underlying theory, it's more challenging for the ARF application. Further ad hoc adjustment is likely needed for M6. It's difficult to fit one set of parameters for both longer and shorter durations. For the derivation of a generalized ARF model, one should develop separate models for long, medium and short durations. It is also acknowledged that other ARF models, which not selected in this use case study, can also be suitable choices as long as a good fitting performance can be demonstrated. Overall, only fixed-area ARF (not storm-centered ARF) is evaluated since the fixed-area ARF is a more suitable choice for PFHA. If the goal is to identify ARF for deterministic applications (e.g., PMP and PMF), the storm-centered ARF would be a more suitable approach instead.

Regarding data sources, while the effect of data sources is relatively smaller than the effect of ARF model, non-negligible differences are still found. The gridded precipitation products (e.g., Daymet, PRISM, Livneh) are easy to use, but given their limited temporal resolution, they cannot be used to derive subdaily ARF. The radar-driven precipitation product (ST4) can better capture high-resolution spatiotemporal storm structure. However, given its limited period of record, it provides the worst long return period ARF model fitting across all precipitation products. While gauge data (DSI-3240) is harder to process, it leads to one of the best ARF model fitting in Region 05 Ohio. With further consideration of topography-informed spatial interpolation, there is value to consider gauge data-based approach in future site-specific ARF studies.

Regarding seasonality, in the three hydrologic regions examined in this study, the results suggest that the warm season ARF is close to all season ARF, while cool season ARF has a higher value. The closeness between warm and all seasons indicates that the annual extreme precipitation in these regions mainly occurs during warm season. The differences in warm and cool seasons ARF can be explained by their respective controlling extreme precipitation processes. In most regions, the major extreme precipitation events during warm season are

meso-scale convective storms that area generally smaller in size and have larger spatial variability (leading to smaller ARF). On the other hand, the major extreme precipitation events in most regions during cool season are mostly large-scale frontal systems with relatively smaller spatial variability (as compared to warm season convective storms) that leads to larger ARF. For the South Atlantic-Gulf region, the difference between warm and cool season ARF is the smallest, suggesting that the effect of seasonality is less significant in the overall warmer South Atlantic-Gulf region. For H&H applications such as simulation of rain-on-snow during cool season in Ohio and Mid-Atlantic regions, the results suggest that a specific cool season ARF will be needed.

Regarding inter-regional differences, it is found that ARFs are lower in the central U.S., and higher in eastern and western U.S. Texas-Gulf (R12) & Souris-Red-Rainy (R09) are generally the lowest among all regions. The results clearly suggest a strong geographical variability associated with ARF. Therefore, ARF values produced from previous studies (e.g., TP-29) may not be indiscriminately used at different geographical locations. There is a need to derive region- or watershed-specific ARF for robust PFHA of national NPPs.

5.2 Challenges for PFHA Applications

While the results of this study suggest that, with the advance of precipitation products and more advanced ARF methods, more defensible ARF estimates can be derived for H&H applications, several major challenges are also identified. For the purpose of PFHA for NPPs, these challenges will need to be addressed in future studies.

- **Long return period areal extreme rainfall and ARF derivation:** One major finding from this study is the challenge associated with long return period areal extreme rainfall and ARF derivation. “Long” return period in this case refers to when the desired return period is much longer than the period of records of the supporting precipitation data (e.g., deriving 1000-year extreme rainfall estimate based on 66 years of data). In particular, when compared to other general H&H applications, PFHA-NPP needs to evaluate the risks from highly extreme precipitation events (i.e., greater than 1000-year return levels), which may not be fully supported only by historic observations. In addition, the potentially changing climatic conditions may further complicate this challenge. From this perspective, this is a need to explore the derivation of long return period areal extreme rainfall and ARF through the assistance of process-based, numerical weather simulation.
- **Uncertainty quantification:** While the variability of ARF across different factors (e.g., methods, data sources, geographical locations, and seasonality) were examined in this study, the statistical uncertainty of these ARF estimates were not evaluated. Existing statistical methods such as bootstrapping can be used to analyze the uncertainty and develop confidence intervals of the selected ARF. For the purpose of PFHA, the uncertainty of ARF should be considered as a part of the H&H application.
- **High spatiotemporal resolution dataset:** Among the gridded precipitation products considered in this study (PRISM, Daymet, Livneh, and ST4), the spatial resolution (from 1-km Daymet to ~6km Livneh) appears to be sufficient for the development of watershed scale ARF. In contrast, temporal data resolution may not be sufficient for some purposes. While DSI-3240 and ST4 provide hourly data, PRISM, Daymet, and Livneh only provide daily data. In addition, reanalysis-driven products provide subdaily data but are of higher spatial resolution. For hydrologic modeling purposes, subdaily data with

high spatial resolution may be needed, leaving DSI-3240 and ST4 as the potential options. However, these datasets each present challenges. While data from ST4 is of limited historical record (since 2002), data from DSI-3240 may extend across several decades; however, DSI-3240 requires greater processing and interpolation effort. As with the challenge of estimating long return period ARF, challenges associated with high spatiotemporal resolution could be overcome by using numerical weather simulation in future research efforts.

- **Subwatershed application:** The use case studies documented in this report provide modeled results for HUC02 (hydrologic region) level assessments. For some regions, particularly where precipitation-producing processes (e.g., in coastal or mountainous locations) vary widely within the region, or regions that are composed of multiple independent watersheds (e.g., South-Atlantic Gulf region), some level of heterogeneity may exist in the ARF results and yield poorer fitting performance. Site-specific subwatershed application could yield more reliable results with better performance.
- **Need for a national ARF product:** Perhaps most importantly, there is currently no new national ARF product (i.e., similar to NOAA Atlas 14) that provides geographically variable ARFs. Existing ARF products also do not consider the effects of return period or seasonality and are applicable for only small area sizes. While this study examined how ARFs may vary based on these different considerations, the results presented herein are for demonstration purposes only and are not intended to be used for ARF application without further site-specific evaluation. Additional research and development efforts, with thorough quality assurance and control performed, would be required to develop a reliable national ARF product suitable for PFHA application. Such efforts may be suitable for a federal agency with specific mission and objectives related to weather prediction and precipitation monitoring, such as NOAA.

6 REFERENCES

- AghaKouchak, A., A. Mehran, H. Norouzi, and A. Behrangi (2012), Systematic and random error components in satellite precipitation data sets, *Geophys. Res. Lett.*, 39, L09406, doi:10.1029/2012GL051592.
- AghaKouchak, A., E. Habib, and A. Bárdossy (2010), Modeling radar rainfall uncertainties: Random error model, *J. Hydrol. Eng.*, 4, 265–274, doi:10.1061/(ASCE)HE.1943-5584.0000185.
- Allen, R. J., and A. T. DeGaetano (2005a), Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts, *J. Hydrol.*, 315(1–4), 203–219, doi:10.1016/j.jhydrol.2005.03.028.
- Allen, R. J., and A. T. DeGaetano (2005b), Areal Reduction Factors for Two Eastern United States Regions with High Rain-Gauge Density, *J. Hydrol. Eng.*, 10(4), 327–335, doi:10.1061/(ASCE)1084-0699(2005)10:4(327).
- Asquith, W. H., J. S. Famiglietti (2000), Precipitation areal-reduction factor estimation using an annual-maxima centered approach, *J. Hydrol.*, 230(1–2), 55–69, doi:10.1016/S0022-1694(00)00170-0.
- Bacchi B. & Ranzi R. (1996), On the derivation of the areal reduction factor of storms. *Atmos Res*, 42, 123–135, doi:10.1016/0169-8095(95)00058-5.
- Beck, H. E., N. Vergopolan, M. Pan, V. Levizzani, A. I. J. M. van Dijk, G. Weedon, L. Brocca, F. Pappenberger, G. J. Huffman, and E. F. Wood (2017), Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling, *Hydrol. Earth Syst. Sci.*, in press, doi:10.5194/hess-2017-508.
- Bell, F. C. (1976), *The areal reduction factor in rainfall frequency estimation*, Institute of Hydrology Report No 35, Natural Environment Research Council, UK.
- Bennett, K. E., T. J. Bohn, K. Solander, N. G. McDowell, C. Xu, E. Vivoni, and R. S. Middleton (2018), Climate-driven disturbances in the San Juan River sub-basin of the Colorado River, *Hydrol. Earth Syst. Sci.*, 22, 709–725, doi: 10.5194/hess-22-709-2018.
- Bonnin, G. M., D. Martin, B. Lin, T. Parzybok, M. Yekta, and D. Riley (2004), *Precipitation-Frequency Atlas of the United States, NOAA Atlas 14, Volume 2*, U. S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD.
- Chow V.T. (1951), A general formula for hydrologic frequency analysis, *Trans Am Geophys Union*, 32, 231–237.
- Chow V.T. (1964), Statistical and probability analysis of hydrologic data. In: V.T. Chow, ed. *Handbook of hydrology*. New York: McGraw-Hill, 8.1–8.42.
- Cifelli, R. and V. Chandrasekar (2013). Dual-Polarization Radar Rainfall Estimation. Rainfall: State of the Science, American Geophysical Union: 105-125.

- Clark, C. and P.R. Rakhecha (2002), Areal PMP distribution of one-day to three-day duration over India, *Meteorol. Appl.*, 9, 399–406, doi:10.1017/S1350482702004024.
- Clark, M. P., and A. G. Slater (2006), Probabilistic Quantitative Precipitation Estimation in Complex Terrain, *J Hydrometeorol.*, 7(1), 3–22, doi:10.1175/JHM474.1.
- Cunha, L. K., J. A. Smith, M. L. Baeck, and W. F. Krajewski (2013). An early performance evaluation of the NEXRAD dual-polarization radar rainfall estimates for urban flood applications, *Weather and Forecasting*, 28, 1478–1497, doi: 10.1175/WAF-D-13-00046.1.
- CWCB (Colorado Water Conservation Board) (2006), “Hydrologic Analysis: Rainfall”, In *Colorado Floodplain and Stormwater Criteria Manual*, Colorado Water Conservation Board.
- Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteor.*, 33, 140—158, [https://doi.org/10.1175/1520-0450\(1994\)033<0140:ASTMFM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2).
- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris (2002), A knowledge-based approach to the statistical mapping of climate, *Climate Res.*, 22, 99–113, doi:10.3354/cr022099.
- De Michele, C., N. T. Kottegoda, and R. Rosso (2001), The derivation of areal reduction factor of storm rainfall from its scaling properties, *Water Resour. Res.*, 37(12), 3247–3252, doi:10.1029/2001WR000346.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. De Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, F (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828.
- Durrans, S. R., L. T. Julian, and M. Yekta (2002), Estimation of Depth-Area Relationships using Radar-Rainfall Data, *J. Hydrol. Eng.*, 7(5), 356–367, doi:10.1061/(ASCE)1084-0699(2002)7:5(356).
- Dyrrdal, A.V., T. Skaugen, F. Stordal, and E.J. Førland (2016), Estimating extreme areal precipitation in Norway from a gridded dataset, *Hydrological Sciences Journal*, 61, (3), 483–494, doi:10.1080/02626667.2014.947289.
- Fouchier, C., A. Maire, P. Arnaud, and P. Cantet (2015), Assessment of probabilistic areal reduction factors of precipitations for the whole French territory with gridded rainfall data, Proc. 10th International Workshop on Precipitation in Urban Areas, Dec 2015, Pontresina, Switzerland. Rainfall in Urban and Natural Systems, 5 p.

- Frederick R.H., V.A. Myers and E.P. Auciello (1977), *Five- to 60-minute precipitation frequency for the eastern and central United States*. NOAA Technical Memorandum NWS HYDRO-35, Office of Hydrology, National Weather Service, Silver Spring, MD.
- Gourley, J. J., Y. Hong, Z. L. Flamig, L. Li and J. Wang (2010), Intercomparison of rainfall estimates from radar, satellite, gauge, and combinations for a season of record rainfall, *J. Appl. Meteor. Climatol.*, 49, 437–452, doi:10.1175/2009JAMC2302.1.
- Grebner, D. and T. Roesch (1997), Regional dependence and application of DAD relationships, Proc. FRIEND'97 – Regional hydrology: Concepts and models for sustainable water resource management.
- Grebner, D., T. Roesch, and M. Schwarb (1998), Extreme regional precipitation of varying duration and return period 1981-1993, *Hydrological Atlas of Switzerland*, Plate 2.5.
- Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen (2014), An intercomparison of statistical downscaling methods used for water resource assessments in the United States, *Water Resour. Res.*, 50, 7167–7186, doi:10.1002/2014WR015559.
- Henn, B., A. J. Newman, B. Livneh, C. Daly, and J. D. Lundquist (2017), An assessment of differences in gridded precipitation datasets in complex terrain, *J. Hydrol.*, in press, doi:10.1016/j.jhydrol.2017.03.008.
- Hershfield, D. M. (1961), Rainfall frequency atlas of the U.S. for duration from 30 minutes to 24 hours and return periods from 1–100 years, U.S. Weather Bureau Technical Paper No. 40, Washington, D.C.
- Hosking, J. R. M., and Wallis, J. R. (1997), *Regional Frequency Analysis*, 1st Ed., Cambridge University Press, New York.
- Huff, F.A. (1995), Characteristics and contributing causes of an abnormal frequency of flood-producing rainstorms at Chicago, *Water Resour. Bull.*, 31, (4), 703–714, doi:10.1111/j.1752-1688.1995.tb03395.x.
- Huff, F.A. and W.L. Shipp (1969), Spatial correlations of storm, monthly and seasonal precipitation, *J. Appl. Meteorol.*, 8, (4), 542–550, doi:10.1175/1520-0450(1969)008<0542:SCOSMA>2.0.CO;2.
- Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi (2015), The JRA-55 Reanalysis: General Specifications and Basic Characteristics, *J. Meteorol. Soc. JPN*, 93(1), 5–48, doi:10.2151/jmsj.2015-001.
- Kok, K., W. Yu, J. Kim, L. Mohd Sidek, and K. Jung (2017), Feasibility study of deriving areal reduction factor for storm design application in Malaysia using satellite rainfall products, *J. Korean Soc. Hazard Mitig.*, 17, (6), 399–409, doi:10.9798/KOSHAM.2017.17.6.399.
- Koutsoyiannis, T. and D. Xanthopoulos (1999), *Engineering Hydrology*, Edition 3, National Technical University of Athens, Athens, Greece.

- Krajewski, W. F., G. Villarini, and J. A. Smith (2010), RADAR-Rainfall Uncertainties: Where are We after Thirty Years of Effort?, *B. Am. Meteorol. Soc.*, 91(1), 87–94, doi:10.1175/2009BAMS2747.1.
- Kuligowski, R. J. (1997), *An Overview of National Weather Service Quantitative Precipitation Estimates*, TDL Office Note 97-4, National Weather Service.
- Leclerc, G. and J.C. Schaake (1972) Derivation of hydrologic frequency curves. Report 142, Mass. Inst. of Technol., Cambridge, 151 pp.
- Liang, X., D. P. Lettenmaier, and E. F. Wood (1996), Surface soil moisture parameterization of the VIC-2L model: evaluation and modification, *Global Planet. Change*, 13, 195–206, doi:10.1016/0921-8181(95)00046-1.
- Liang, X., D. P. Lettenmaier, S. J. Burges, and E. F. Wood (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, 99, 14415–14428, doi:10.1029/94JD00483.
- Livneh, B., T. J. Bohn, D. W. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. R. Cayan, and L. Brekke (2015), A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and southern Canada 1950—2013, *Nature Scientific Data*, 5, 150042, doi:10.1038/sdata.2015.42.
- Lombardo, F., F. Napolitano, and F. Russo (2006), On the use of radar reflectivity for estimation of the areal reduction factor, *Nat. Hazards Earth Syst. Sci.*, 6, 377–386, doi:10.5194/nhess-6-377-2006.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, *J. Climate*, 15, 3237—3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.
- McCuen, R. H., Z. Knight, and A. G. Cutter (2006), Evaluation of the Nash-Sutcliffe efficiency index. *J. Hydrol. Eng.*, 11 (6).
- McMillan, H., T. Krueger, and J. Freer (2012), Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality, *Hydrol. Process.*, 26, 4078–4111, doi:10.1002/hyp.9384.
- Mekonnen, G. B., S. Matula, F. Doležal, and J. Fišák (2015), Adjustment to rainfall measurement undercatch with a tipping-bucket rain gauge using ground-level manual gauges, *Meteorol. Atmos. Phys.*, 127, 241—256, doi: 10.1007/s00703-014-0355-z.
- MGNDCT (2012), *Stormwater Management Manual*, Metropolitan Government of Nashville and Davidson County, Tennessee.
- Mizukami, N., M. P. Clark, A. J. Newman, A. W. Wood, E. D. Gutmann, B. Nijssen, O. Rakovec, and L. Samaniego (2017), Towards seamless large-domain parameter estimation for hydrologic models, *Water Resour. Res.*, 53, 8020–8040, doi:10.1002/2017WR020401.

- Myers, V.A. and R.M. Zehr (1980), A methodology for point-to-area rainfall frequency ratios, NOAA Technical Report NWS 24. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Washington DC.
- Nathan, R and E. Weinmann (2016), Rainfall Estimation, Book 2 in Australian Rainfall and Runoff - A Guide to Flood Estimation, Commonwealth of Australia.
- National Ocean Service (2018), A History of Observing the Weather, https://celebrating200years.noaa.gov/foundations/weather_obs/, accessed Mar. 14, 2018.
- National Weather Service (2017a), History of the National Weather Service, <https://www.weather.gov/timeline>, accessed Nov. 6, 2017.
- National Weather Service (2017b), National Weather Service Manual 10-1315: Cooperative Station Observations and Maintenance, accessed Oct. 9, 2017.
- National Weather Service (2018), Cooperative Observer Program, <https://www.weather.gov/pah/coop>, accessed Mar. 14, 2018.
- NCEI (2018), NEXRAD, National Centers for Environmental Information, <https://www.ncdc.noaa.gov/data-access/radar-data/nexrad>, accessed Mar. 14, 2018.
- NERC (1975), *Flood studies report, Vol. II*, Natural Environment Research Council, UK.
- Newman, A. J., M. P. Clark, J. Craiget, B. Nijssen, A. Wood, E. Gutmann, N. Mizukami, L. Brekke, and J. R. Arnold (2015), Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States, *J. Hydrometeorol.*, 16, 2481–2500, doi:10.1175/JHM-D-15-0026.1.
- NOAA (National Oceanic and Atmospheric Administration) (2018), Q&A about NCEP Stage II/Stage IV, <http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/QandA/>, accessed Apr. 5, 2018.
- Omolayo A.S. (1989), Relationships between point and areal rainfall for flood and drought assessments. PhD Thesis, University of New South Wales, Australia.
- Omolayo A.S. (1993), On the transposition of areal reduction factors for rainfall frequency estimation. *J Hydrol*, 145, (1-2), 191-205, doi:10.1016/0022-1694(93)90227-Z.
- Oubeidillah, A. A., S.-C. Kao, M. Ashfaq, B. S. Naz, and G. Tootle (2014), A Large-Scale, High-Resolution Hydrological Model Parameter Data Set for Climate Change Impact Assessment for the Conterminous US, *Hydrol. Earth Syst. Sci.*, 18, 67–84, doi:10.5194/hess-18-67-2014.
- Overeem, A., T. A. Buishand, and I. Holleman (2010), Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar, *Water Resour. Res.*, 45, W10424, doi: 10.1029/2009WR007869.

- Overeem, A., T. A. Buishand, I. Holleman, and R. Uijlenhoet (2010), Extreme value modeling of areal rainfall from weather radar, *Water Resour. Res.*, *46*, W09514, doi:10.1029/2009WR008517.
- Parker, W.S. (2016), Reanalyses and Observations: What's the Difference?, *B. Am. Meteorol. Soc.*, *97*(9), 1565–1572, doi:10.1175/BAMS-D-14-00226.1.
- Pavlovic, S., S. Perica, M. St Laurent, and A. Mejía (2016), Intercomparison of selected fixed-area areal reduction factor methods, *J. Hydrol.*, *537*, 419-430, doi:10.1016/j.jhydrol.2016.03.027.
- Petković, V., and C. D. Kummerow (2017), Understanding the sources of satellite passive microwave rainfall retrieval systematic errors over land, *J. Appl. Meteor. Climatol.*, *56*, 597–614, doi: 10.1175/JAMC-D-16-0174.1.
- Pietersen, J.P.J., O.J. Gericke, J.C. Smithers, and Y.E. Woyessa (2015), Review of current methods for estimating areal reduction factors applied to South African design point rainfall and preliminary identification of new methods, *J. South African Institution of Civil Eng.*, *57*, (1), 16–30, doi:10.17159/2309-8775/2015/v57n1a2.
- Ramos, M.H., J.-D. Creutin, and E. Leblois (2005), Visualization of storm severity, *J. Hydrol.*, *315*, 295–307, doi:10.1016/j.jhydrol.2005.04.007.
- Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. Mahanama, and G. S. Partyka (2017), Land Surface Precipitation in MERRA-2, *J. Climate*, *30*, 1643–1664, doi: 10.1175/JCLI-D-16-0570.1.
- Rinehart, R. E. 1997. *Radar for Meteorologists*. Grand Forks, ND: Rinehart Publishing.
- Rodriguez-Iturbe I. & Mejía J.M. (1974), On the transformation of point rainfall to areal rainfall. *Water Resour Res*, *10*, (4), 729–735, doi:10.1029/WR010i004p00729.
- Seo, B.-C., B. Dolan, W. F. Krajewski, S. A. Rutledge, and W. Petersen (2015), Comparison of single- and dual-polarization-based rainfall estimates using NEXRAD data for the NASA Iowa flood studies project, *J. Hydromet.*, *16*, 1658–1675, doi: 10.1175/JHM-D-14-0169.1.
- Sheffield, J., G. Goteti, and E.F. Wood (2006), Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, *J. Climate*, *19*, 3088–3111, doi:10.1175/JCLI3790.1.
- Sieck, L. C., S. J. Burges, and M. Steiner (2007), Challenges in obtaining reliable measurements of point rainfall, *Water Resour. Res.*, *43*, W01420, doi: 10.1029/2005WR004519.
- Sivapalan, M., and G. Blöschl (1998), Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves, *J. Hydrol.*, *204*(1–4), 150–167, doi:10.1016/S0022-1694(97)00117-0.
- Skaugen, T. (1997), Classification of rainfall into small- and large-scale events by statistical pattern recognition, *J. Hydrol.*, *200*, 40-57, doi:10.1016/S0022-1694(97)00003-6.

- Stewart, E.J. (1989), Areal reduction factors for design storm construction: Joint use of raingauge and radar data. In: New directions for surface water modeling, Proc. Baltimore Symposium, May, IAHS Publ., 181, 31–40.
- Svensson, C., and D. A. Jones (2010), Review of methods for deriving areal reduction factors, *J. Flood Risk Manag.*, 3, 232–245, doi:10.1111/j.1753-318X.2010.01075.x.
- Thornton, P. E., M. M. Thornton, B. W. Mayer, Y. Wei, R. Devarakonda, R. S. Vose, and R. B. Cook (2017), Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3, ORNL DAAC, Oak Ridge, Tennessee, USA, <https://doi.org/10.3334/ORNLDAAAC/1328>.
- Thornton, P. E., S. W. Running, and M. A. White (1997), Generating surfaces of daily meteorology variables over large regions of complex terrain, *J. Hydrol.*, 190, 214–251, doi:10.1016/S0022-1694(96)03128-9
- Tian, Y., and C. D. Peters-Lidard (2010), A global map of uncertainties in satellite-based precipitation estimates, *Geophys. Res. Lett.*, 37, L24407, doi:10.1029/2010GL046008.
- U.S. Weather Bureau (1957), *Rainfall Intensity-Frequency Regime: Part 1 – The Ohio Valley*, Technical Paper No. 29, U.S. Weather Bureau, Washington, D.C.
- U.S. Weather Bureau (1958a), *Rainfall Intensity-Frequency Regime: Part 2 – Southeastern United States*, Technical Paper No. 29, U.S. Weather Bureau, Washington, D.C.
- U.S. Weather Bureau (1958b), *Rainfall Intensity-Frequency Regime: Part 3 – The Middle Atlantic Region*, Technical Paper No. 29, U.S. Weather Bureau, Washington, D.C.
- U.S. Weather Bureau (1959), *Rainfall Intensity-Frequency Regime: Part 4 – Northeastern United States*, Technical Paper No. 29, U.S. Weather Bureau, Washington, D.C.
- U.S. Weather Bureau (1960), *Rainfall Intensity-Frequency Regime: Part 5 – Great Lakes Region*, Technical Paper No. 29, U.S. Weather Bureau, Washington, D.C.
- USBR (2004), *Causey Dam, Utah 1,000-Year Rainfall-Runoff Report*, Report No. DSO-03-01, U.S. Bureau of Reclamation.
- Veneziano, D., and A. Langousis (2005), The areal reduction factor: A multifractal analysis, *Water Resour. Res.*, 41, W07008, doi:10.1029/2004WR003765.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, *Climatic Change*, 62(1–3), 189–216, doi: 10.1023/B:CLIM.0000013685.99609.9e.
- Zehr, R. M. and V. A. Myers (1984), Depth-Area Ratios in the Semi-Arid Southwest United States. NOAA Tech. Memo. *HYDRO-40*, Washington, DC, 1984.

APPENDIX A

SUMMARY OF AVAILABLE PRECIPITATION PRODUCTS

Table A-1 summarizes the precipitation products available for ARF analysis and various features of these products. They are organized by product type, including gauge-only, gauge-driven, radar-driven, satellite-driven, and reanalysis-driven precipitation products.

Descriptions of the precipitation products are provided in Section 2.2.

Table A-1 Summary of Available Precipitation Products

N	Precipitation Dataset	Provider	Dataset Type	Coverage Start	Coverage End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution	Project Main Page	Citation	Data FTP Site
Gauge-only Datasets												
1	Hourly Precipitation Data (DSI-3240)	NOAA National Centers for Environmental Information (NCEI)	Gauge observation	1940	2013	Data since 2014 have not been released (checked 10/17/2017)	U.S. (including AK, HI, PR)	Hourly	Gauge	https://data.nodc.noaa.gov/cgi-bin/iso?id=go.v.noaa.ncdc:C00313		ftp://ftp.ncdc.noaa.gov/pub/data/hourly_precip-3240/
2	15 Minute Precipitation Data (DSI-3260)	NOAA National Centers for Environmental Information (NCEI)	Gauge observation	1970	2013	Data since 2014 have not been released (checked 10/17/2017)	U.S. (including AK, HI, PR)	15-minute	Gauge	https://data.nodc.noaa.gov/cgi-bin/iso?id=go.v.noaa.ncdc:C00505		ftp://ftp.ncdc.noaa.gov/pub/data/15min_precip-3260/
3	Global Historical Climatology Network - Daily (GHCND)	NOAA National Centers for Environmental Information (NCEI)	Gauge observation	1870	present	Database updated every month	Global	Daily	Gauge	https://data.nodc.noaa.gov/cgi-bin/iso?id=go.v.noaa.ncdc:C00861	Menne et al. (2012), An Overview of the Global Historical Climatology Network-Daily Database, J. Atmos. Oceanic Technol., 29, 897-910. doi:10.1175/JTECH-D-11-00103.1.	ftp://ftp.ncdc.noaa.gov/pub/data/ghecndaily/
Gauge-driven Products												
4	Daymet	Oak Ridge National Laboratory (ORNL)	Gridded from gauge observation	1980	2017	Annual update	North America	Daily	1 km * 1 km (21,533,332 grid points)	https://daymet.ornl.gov/	Thornton et al. (2017), Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3, ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1328	https://daac.ornl.gov/cgi-bin/dataset_viewer.pl?p=32
5	PRISM (daily)	Oregon State University	Gridded from gauge observation (and	1981	present	Operational (updated automatically)	U.S. (48 states)	Daily	1/24 deg * 1/24 deg (~ 4 km * 4 km, 481,631 grid points)	http://www.prism.oregonstate.edu/	Daly et al. (1994), A statistical-topographic model for mapping climatological precipitation over	ftp://prism.nas.nasa.gov/

Table A-1 Summary of Available Precipitation Products (cont.)

N	Precipitation Dataset	Provider	Dataset Type	Coverage Start	Coverage End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution	Project Main Page	Citation	Data FTP Site
			partially with radar)								mountainous terrain, J. Appl. Meteor., 33, 140-158.	
6	Livneh Daily CONUS Near-surface Gridded Meteorological Data	University of Colorado, Boulder	Gridded from gauge observation	1950	2013	No scheduled update (checked 10/17/2017)	U.S. (48 states), Mexico, and Canada (south of 53N)	Daily	1/16 deg * 1/16 deg (~ 6 km * 6 km, 333,579 grid points)	https://data.nodc.noaa.gov/cgi-bin/iso?id=go.v.noaa.nodc:0129374.view.html	Livneh et al. (2015), A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and southern Canada 1950-2013, Nature Scientific Data, 5:150042, doi:10.1038/sdata.2015.42.	ftp://ftp.nodc.noaa.gov/pub/data.nodc/no_dc/archive/data/0129374/
7	Maurer Gridded Meteorological Data	Santa Clara University	Gridded from gauge observation	1949	2010	No scheduled update (checked 10/17/2017)	U.S. (48 states) and Canada (British Columbia)	Daily	1/8 deg * 1/8 deg (~ 12 km * 12 km)	http://www.earthsystem.nasa.gov/ftp/ncar/ensemble/precip/analysis-of-daily-precipitation-over-conus	Maurer et al. (2002), A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, J. Climate 15, 3237-3251.	http://hydro.earthsystem.nasa.gov/ftp/ncar/ensemble/precip/analysis-of-daily-precipitation-over-conus
8	CPC Unified Gauge-based Analysis of Daily Precipitation over CONUS	NOAA Climate Prediction Center (CPC)	Gridded from gauge observation	1948	2006	No scheduled update (checked 10/17/2017)	U.S. (48 states)	Daily	1/4 deg * 1/4 deg	https://data.nodc.noaa.gov/datasets/cpc-us-p-recip/	Newman et al. (2015), Gridded Ensemble Precipitation and Temperature Estimates for the Conterminous United States, J. Hydrometeorol., 16, 2481-2500.	https://www.earthsystem.nasa.gov/ftp/ncar/ensemble/precip/analysis-of-daily-precipitation-over-conus
9	Gridded Ensemble Precipitation and Temperature Estimates over the Conterminous United States	National Center for Atmospheric Research (NCAR)	Probabilistic interpolation using gauge observation	1980	2012	No scheduled update (checked 10/17/2017)	U.S. (48 states)	Daily	1/8 deg * 1/8 deg (~ 12 km * 12 km)	https://www.earthsystem.nasa.gov/ftp/ncar/ensemble/precip/analysis-of-daily-precipitation-over-conus	Newman et al. (2015), Gridded Ensemble Precipitation and Temperature Estimates for the Conterminous United States, J. Hydrometeorol., 16, 2481-2500.	https://www.earthsystem.nasa.gov/ftp/ncar/ensemble/precip/analysis-of-daily-precipitation-over-conus
Radar-driven Products												

Table A-1 Summary of Available Precipitation Products (cont.)

N	Precipitation Dataset	Provider	Dataset Type	Coverage Start	Coverage End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution	Project Main Page	Citation	Data FTP Site
10	Next Generation Weather Radar (NEXRAD) Level II and Level III Products	NOAA National Centers for Environmental Prediction (NCEP)	Radar	1992	present	Operational (updated automatically)	U.S. (including AK, HI, PR)	4 to 10 minutes	1 km x 1 deg azimuth (original); 250 m x 0.5 deg azimuth (super resolution)	https://www.ngdc.noaa.gov/metadata/view.php?url=NOAA/NESDIS/NCDC/Geoportal/iso/xml/C00708.xml&view=getDataView&header=none		https://www.ngdc.noaa.gov/nexradinv/
11	NCEP National Stage II Analyses	NOAA National Centers for Environmental Prediction (NCEP)	Merged radar and gauges (no QC)	1996	present	Operational (updated automatically)	U.S. (48 states)	Hourly	4 km * 4 km	http://www.nc.ncep.noaa.gov/mmb/yli/n/pcpam/stag/e2/	Katz (2006), GCIP/EOP Surface: Precipitation NCEP/EMC 4 KM Gridded Dataset. Version 1.0, UCAR/NCAR - Earth Observing Laboratory, https://data.eol.ucar.edu/dataset/21.006 .	https://data.eol.ucar.edu/dataset/21.006
12	River Forecasting Center (RFC) Stage III / Quantitative Precipitation Estimate (QPE) Analyses	NOAA River Forecasting Centers (RFCs)	Merged radar and gauges (with QC)	2002	present	Operational (updated automatically)	All 14 RFCs	Hourly	4 km * 4 km	https://dipper.nws.noaa.gov/hdsb/data/nexrad.html		
13	NCEP National Stage IV Analyses	NOAA National Centers for Environmental Prediction (NCEP)	Merged radar and gauges (with QC)	2002	present	Operational (updated automatically)	U.S. (48 states), excluding California-Nevada and Northwest RFCs; AK available since 2017	Hourly	4 km * 4 km	http://www.nc.ncep.noaa.gov/mmb/yli/n/pcpam/stag/e4/	Lin (2011), GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data, Version 1.0, UCAR/NCAR - Earth Observing Laboratory, https://doi.org/10.5065/D6FG1QDD	https://data.eol.ucar.edu/dataset/21.093
14	Multi-Radar Multi-Sensor System (MRMS)	NOAA National Severe Storms Laboratory (NSSL)	Merged radar, gauges, model and other sensors (no manual QC)	N/A	present	Operational (updated automatically), historical data not available	U.S. (48 states) and neighboring regions in Canada and Mexico	2-minute	1 km * 1 km	http://www.nssl.noaa.gov/projects/mrms/		http://mrms.ncep.noaa.gov/data/

Table A-1 Summary of Available Precipitation Products (cont.)

N	Precipitation Dataset	Provider	Dataset Type	Coverage Start	Coverage End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution	Project Main Page	Citation	Data FTP Site
Satellite-driven Products												
15	TRMM-based Multi-satellite Precipitation Analysis (TMPA/3B42)	NASA, Japan Aerospace Exploration Agency	Multi-satellite	2000	present	Operational (updated automatically)	Global 60S-60N	3-hourly	0.25 deg	https://trmm.gsfc.nasa.gov/3b42.html	Huffman et al. (2007), The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale, J. Hydrometeor., 8(1), 38-55.	https://pmm.nasa.gov/data-access/downloads/trmm
16	Global Precipitation Measurement (GPM)	NASA, Japan Aerospace Exploration Agency	Multi-satellite	2014	present	Operational (updated automatically)	Global 60S-60N	30-minute	0.1 deg * 0.1 deg	https://pmm.nasa.gov/GPM		https://pmm.nasa.gov/data-access/downloads/gpm
Reanalysis-driven Products												
17	NCEP/NCAR Reanalysis 1	NOAA National Centers for Environmental Prediction (NCEP)	Meteorological Reanalysis	1948	present	Operational (updated every day)	Global	6-hourly	1.875 deg * 1.875 deg (192 * 94 grid points)	https://www.eol.ucar.edu/data/ncar	Kalnay et al. (1996), The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 437-470.	https://www.eol.ucar.edu/data/ncar
18	NCEP/DOE Reanalysis 2	NOAA National Centers for Environmental Prediction (NCEP)	Meteorological Reanalysis	1979	present	Database updated every month	Global	6-hourly	1.875 deg * 1.875 deg (192 * 94 grid points)	https://www.eol.ucar.edu/data/ncar	Kanamitsu et al. (2000), NCEP-DOE AMIP-II Reanalysis (R-2), Bulletin of the American Meteorological Society, 1631-1643.	https://www.eol.ucar.edu/data/ncar
19	North American Regional Reanalysis (NARR)	NOAA National Centers for Environmental Prediction (NCEP)	Meteorological Reanalysis	1979	2014	No scheduled update (checked 10/17/2017)	North America	3-hourly	32 km	https://www.ncep.noaa.gov/data-access/model-data/narr	Mesinger et al. (2006), North American Regional Reanalysis, Bulletin of the American Meteorological Society, 87, 343-360, doi:10.1175/BAMS-87-3-343	ftp://nomads.ncep.noaa.gov/NARR/

Table A-1 Summary of Available Precipitation Products (cont.)

N	Precipitation Dataset	Provider	Dataset Type	Coverage Start	Coverage End	Data Latency	Spatial Coverage	Temporal Resolution	Spatial Resolution	Project Main Page	Citation	Data FTP Site
20	Climate Forecast System Reanalysis (CFSR)	NOAA National Centers for Environmental Prediction (NCEP)	Meteorological Reanalysis	1979	present	Operational (updated every day)	Global	6-hourly	0.313 deg * 0.313 deg	https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2	Saha et al. (2010), The NCEP Climate Forecast System Reanalysis, Bulletin of the American Meteorological Society, 91, 1015–1057, doi:10.1175/2010BA MS3001.1	https://nomad.srncdc.noaa.gov/modeldata/cmd_fixf/
21	Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2)	NASA	Meteorological Reanalysis	1980	present	Operational (updated every day)	Global	Hourly	0.5 deg * 0.625 deg	https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/	Ronald Gelaro et al. (2017), The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Clim., doi:10.1175/JCLI-D-16-0758.1.	https://disc.sci.gsfc.nasa.gov/datasets/pa_ge=1&keywo_rds=MERRA-2
22	Japanese 55-year Reanalysis (JRA-55)	Japan Meteorological Agency	Meteorological Reanalysis	1958	present	Database updated every month	Global	3-hourly	55 km	https://rda.ucar.edu/dataset/s/ds628.0/#access	Kobayashi et al. (2015), The JRA-55 Reanalysis: General Specifications and Basic Characteristics, J. Meteorol. Soc. JPN, 93(1), 5–48, doi:10.2151/jmsj.2015-001.	https://rda.ucar.edu/dataset/s/ds628.0/#access
23	ERA-Interim	European Centre for Medium-Range Weather Forecasts	Meteorological Reanalysis	1979	present	Database updated every month. Will be fully replaced by ERA5 in 2019.	Global	6-hourly	0.7 deg * 0.7 deg	https://www.ecmwf.int/en/forecasts/datasets/era-interim	Dee et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828.	http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/

APPENDIX B

QUALITATIVE ASSESSMENT RESULTS FOR ARF METHODS

Table B-1 documents the qualitative assessment results for ARF methods as described in in Section 3.3. The results were used to inform which methods would be used in the use case studies.

Table B-1 ARF Method Assessment Matrix

	Qualitative Assessment Consideration						Total
	Spatiotemporal Scale and Resolution	Data Sufficiency and Dependency	Required Assumption	Analytical Complexity	Independent Evaluation		
Empirical Methods							
US Weather Bureau Approach (TP-29; 1957, 1958a, 1958b, 1959, 1960)	0	0	1	0.5	1	2.5	
Leclerc and Schaake (1972) – curve fitting of TP-29	0	0	1	1	1	3	
United Kingdom Approach (NERC, 1975)	1	0.5	1	0.5	1	4	
Koutsoyiannis and Xanthopoulos (1999) – curve fitting of NERC (1975)	1	0.5	1	1	1	4.5	
Bell (1976) – extension of United Kingdom ARF	1	0.5	1	0.5	1	4	
National Weather Service (TR-24; Myers and Zehr, 1980)	0	0	0.5	0	0	0.5	
Annual-maxima-centered (Asquith and Famiglietti, 2000)	1	0.5	0.5	0.5	0	2.5	
Swiss Approach (Grebner et al., 1998) – formula-based model available	1	0.5	0.5	1	1	4	
Australian Rainfall & Runoff Approach (Nathan and Weimann, 2016)	1	0.5	1	0.5	1	4	
Spatial Correlation Methods							
Rodriguez-Iturbe and Mejia (1974)	0	0.5	0	0.5	0	1	
Sivapalan and Blöschl (1998)	0.5	1	0.5	0.5	1	3.5	
Omolayo (1989)	0.5	0.5	0.5	0.5	0	2	
Statistical Crossing Properties Method							
Bacchi and Ranzi (1996)	0.5	0.5	0	0.5	0.5	2	
Spatial and Temporal Scaling Methods							
De Michele et al. (2001)	0.5	1	0.5	1	1	4	
Veneziano and Langousis (2005)	1	1	0.5	0.5	0	3	
Extreme Value Theory Methods							
Durrans et al. (2002)	0.5	0.5	0.5	0.5	0.5	2.5	
Allen and DeGaetano (2005a)	0.5	0.5	0.5	0.5	0.5	2.5	
Lombardo et al. (2006)	0.5	1	0.5	0.5	0.5	3	
Overeem et al. (2010)	1	1	0.5	0.5	1	4	

Scores range from 0 to 1:

- 0: consideration significantly limits method's applicability
- 0.5: consideration limits method's applicability
- 1: consideration is satisfied and method is applicable

APPENDIX C OHIO (REGION 05) DETAILED RESULTS

Additional information on the Ohio regional assessment results is provided in the figures below. Example information and discussion is provided in the main report. The plots include:

Figure C-1	Calculated Ohio ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.	C-2
Figure C-3	Calculated Ohio ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models <i>Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.</i>	C-4
Figure C-4	Calculated Ohio ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-5
Figure C-5	Calculated Ohio ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-6
Figure C-6	Calculated Ohio ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	C-7
Figure C-7	Calculated Ohio ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-8
Figure C-8	Calculated Ohio ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-9
Figure C-9	Calculated Ohio ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models	C-10

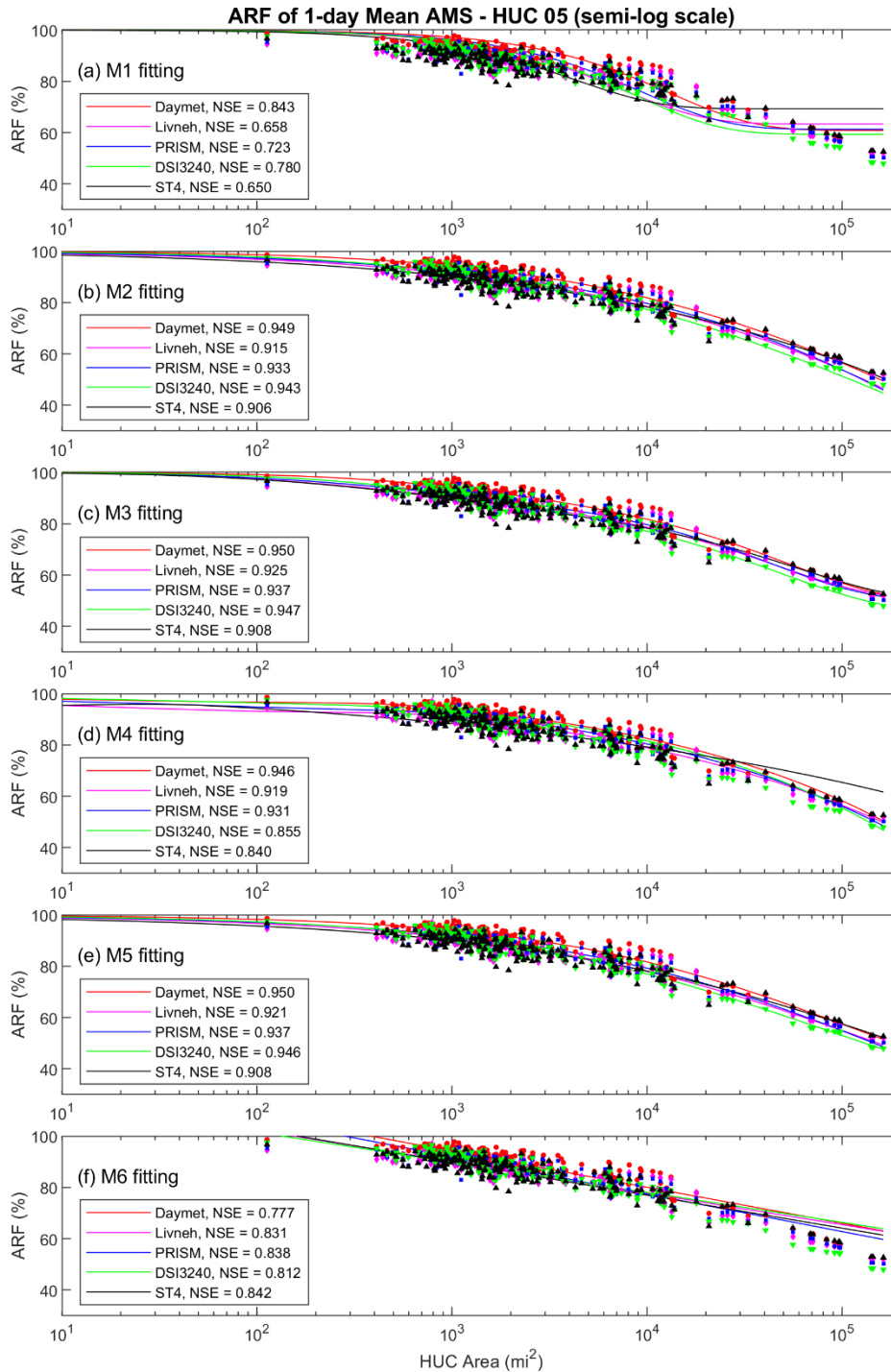


Figure C-1 Calculated Ohio ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

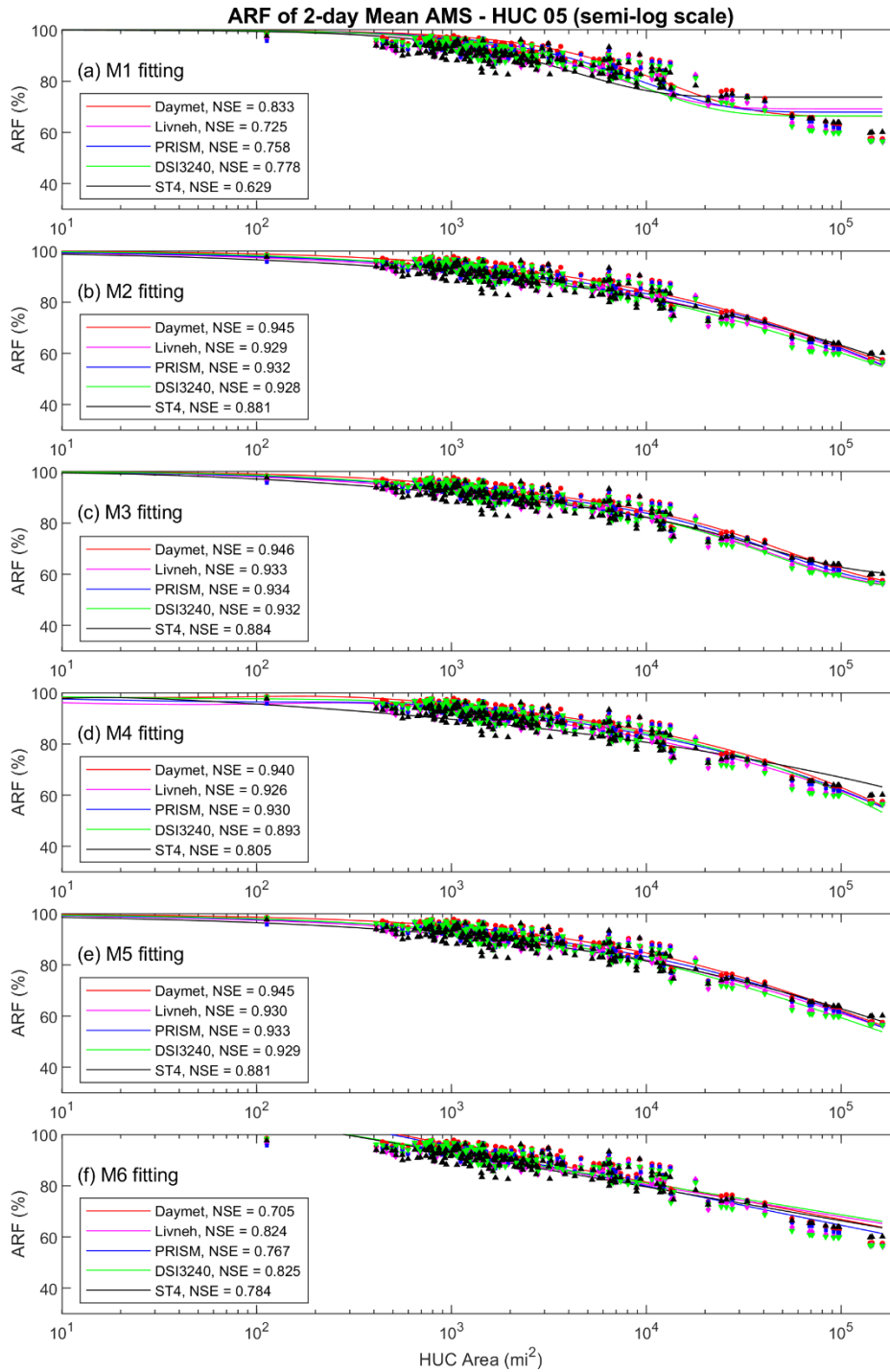


Figure C-2 Calculated Ohio ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

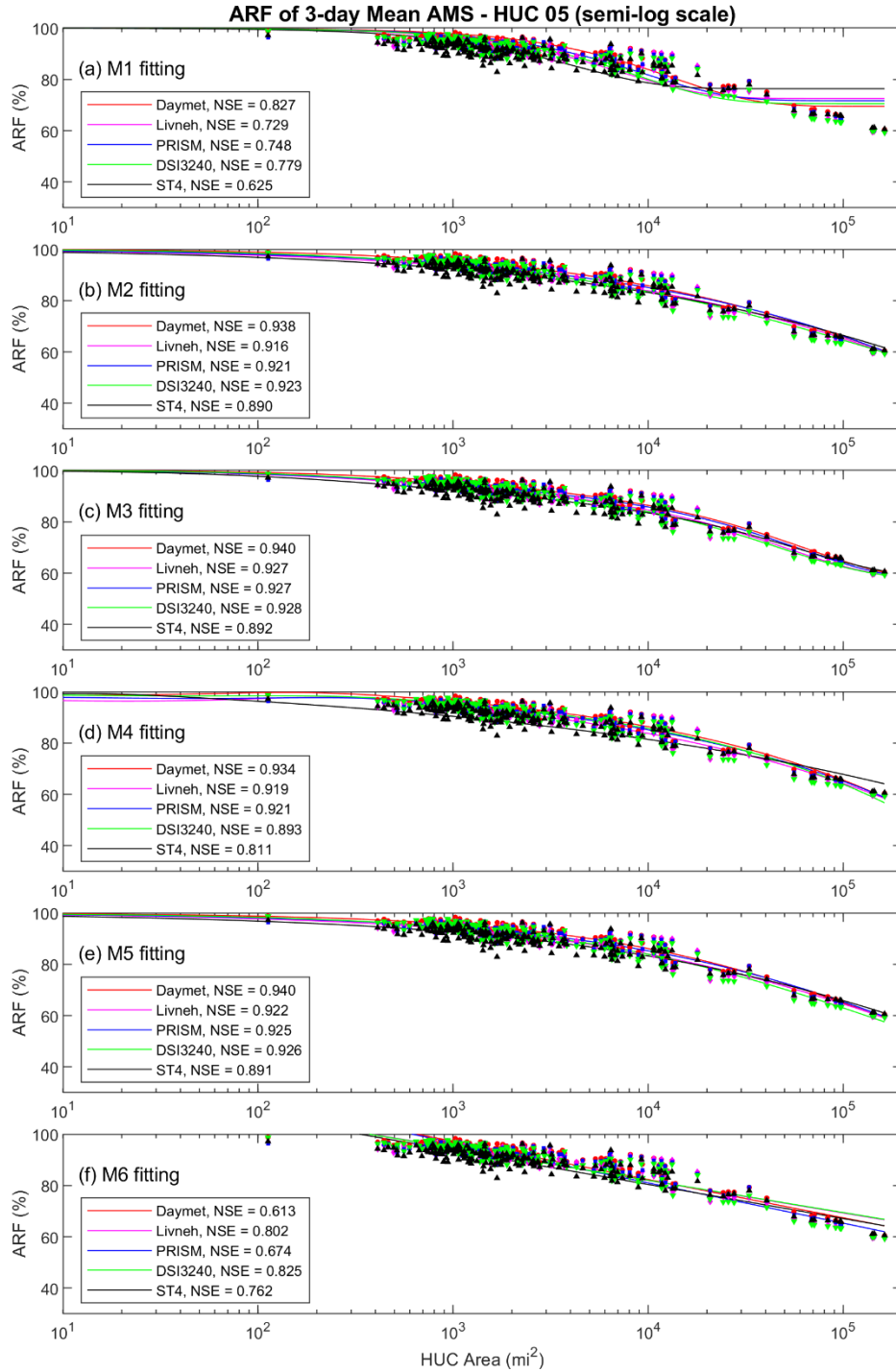


Figure C-3 Calculated Ohio ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

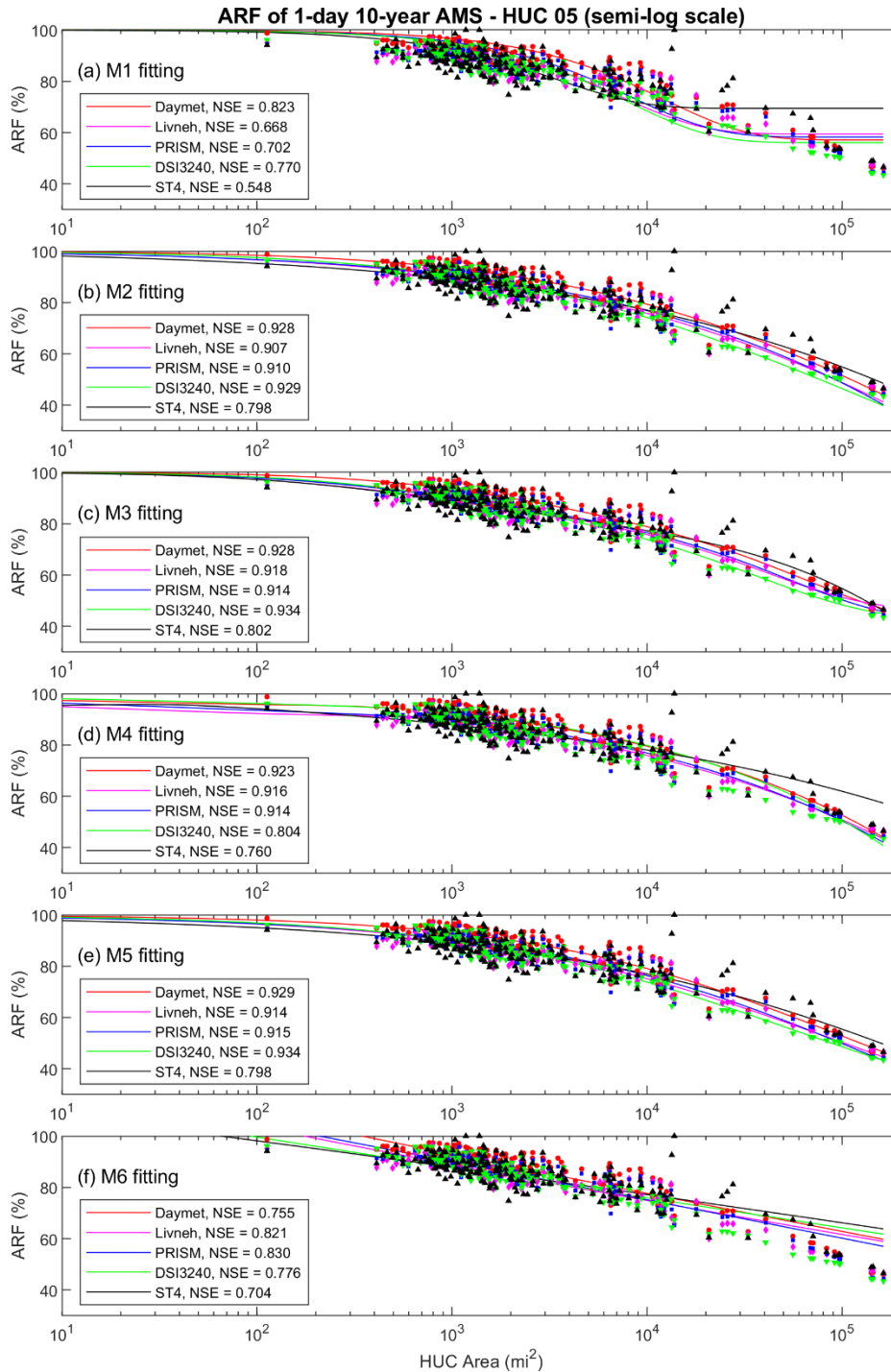


Figure C-4 Calculated Ohio ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

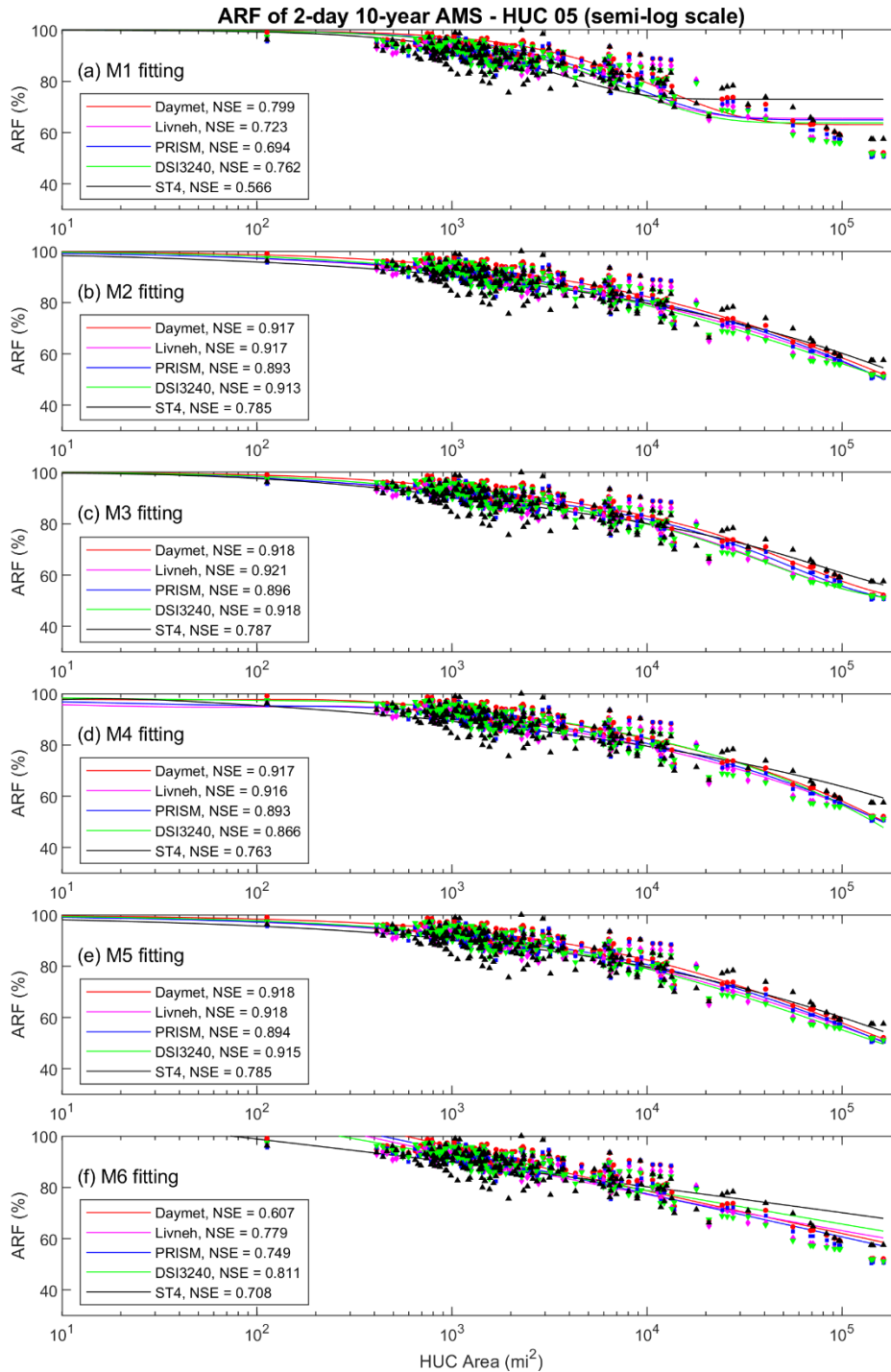


Figure C-5 Calculated Ohio ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

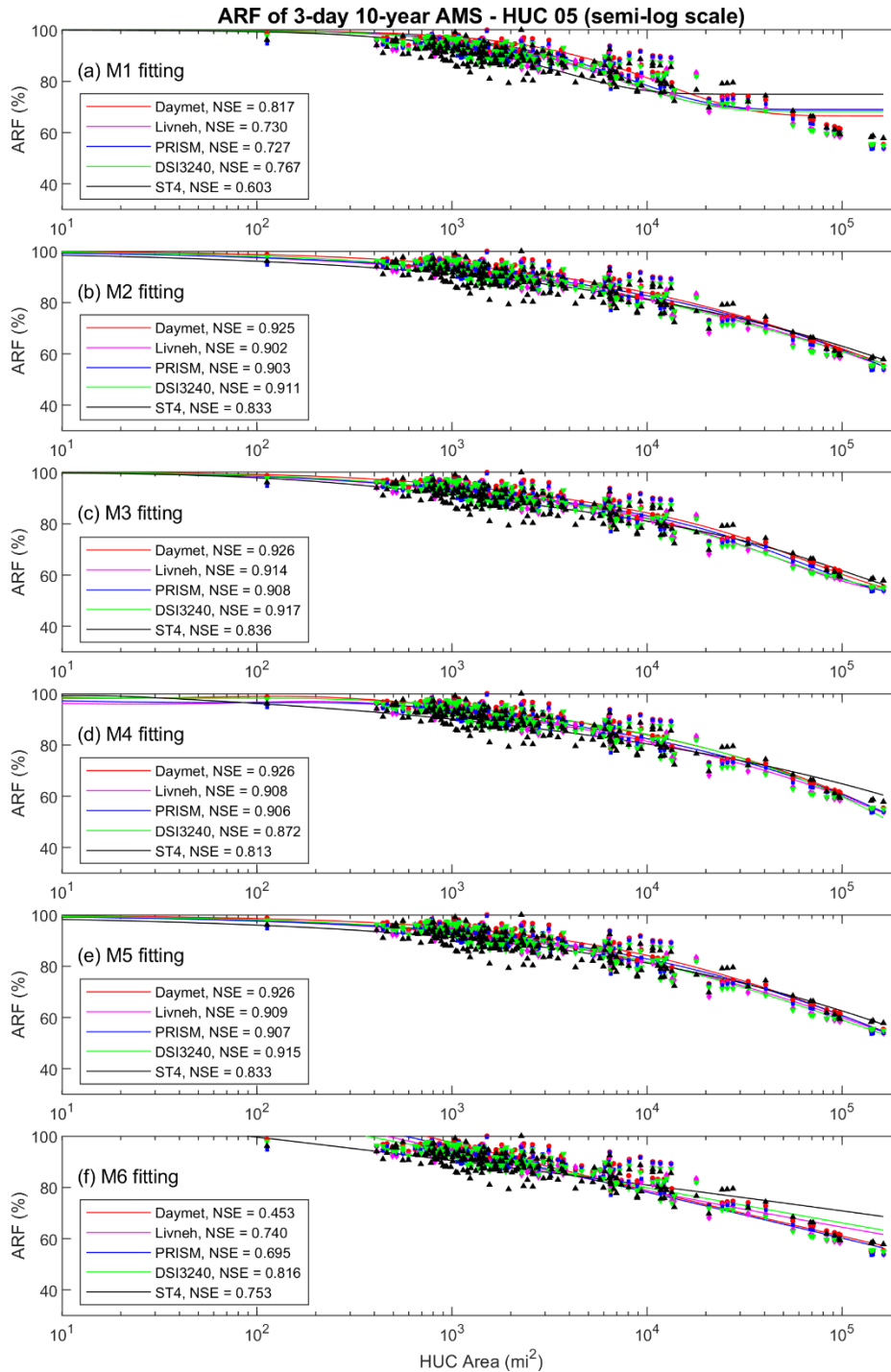


Figure C-6 Calculated Ohio ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

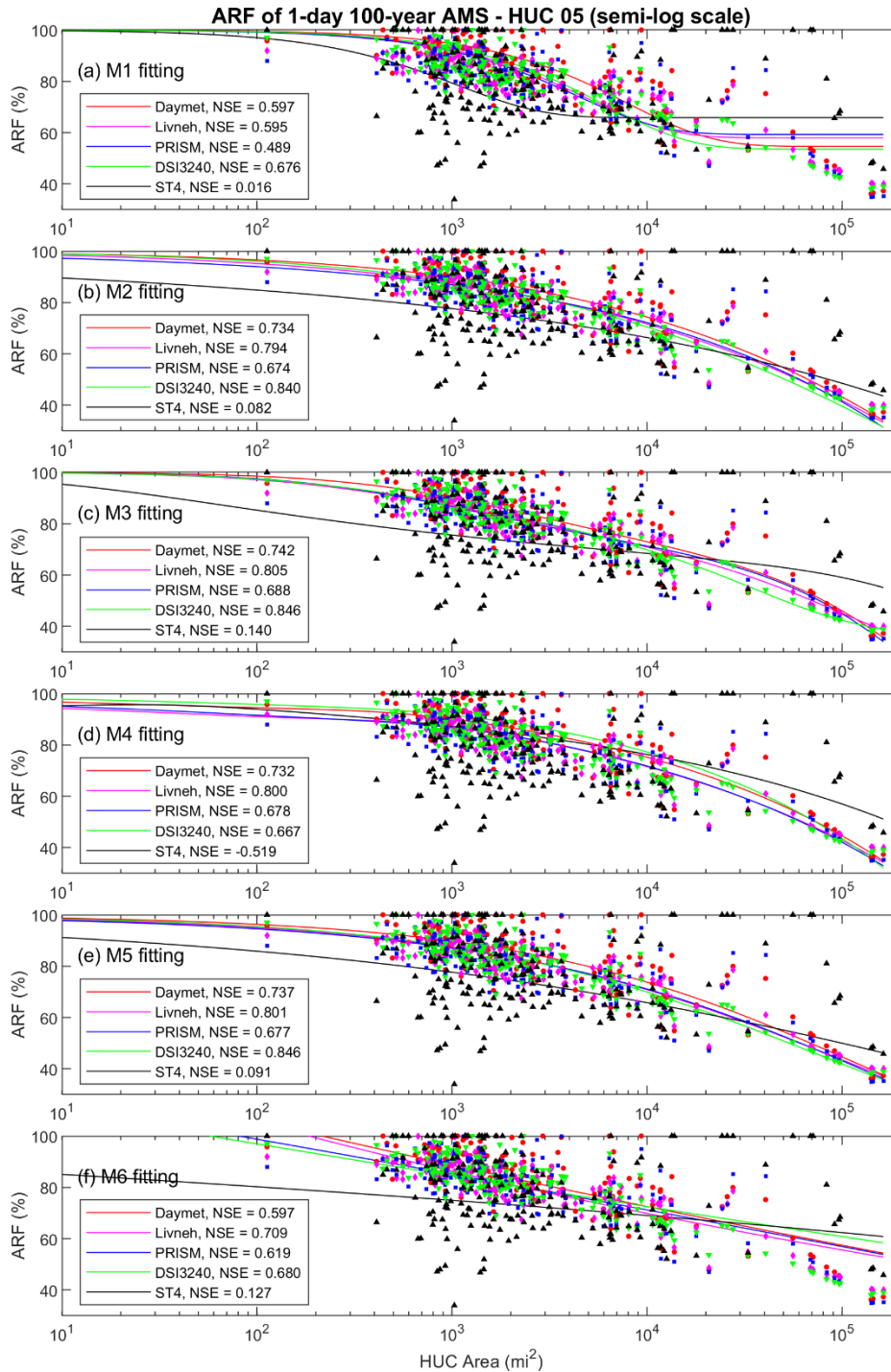


Figure C-7 Calculated Ohio ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

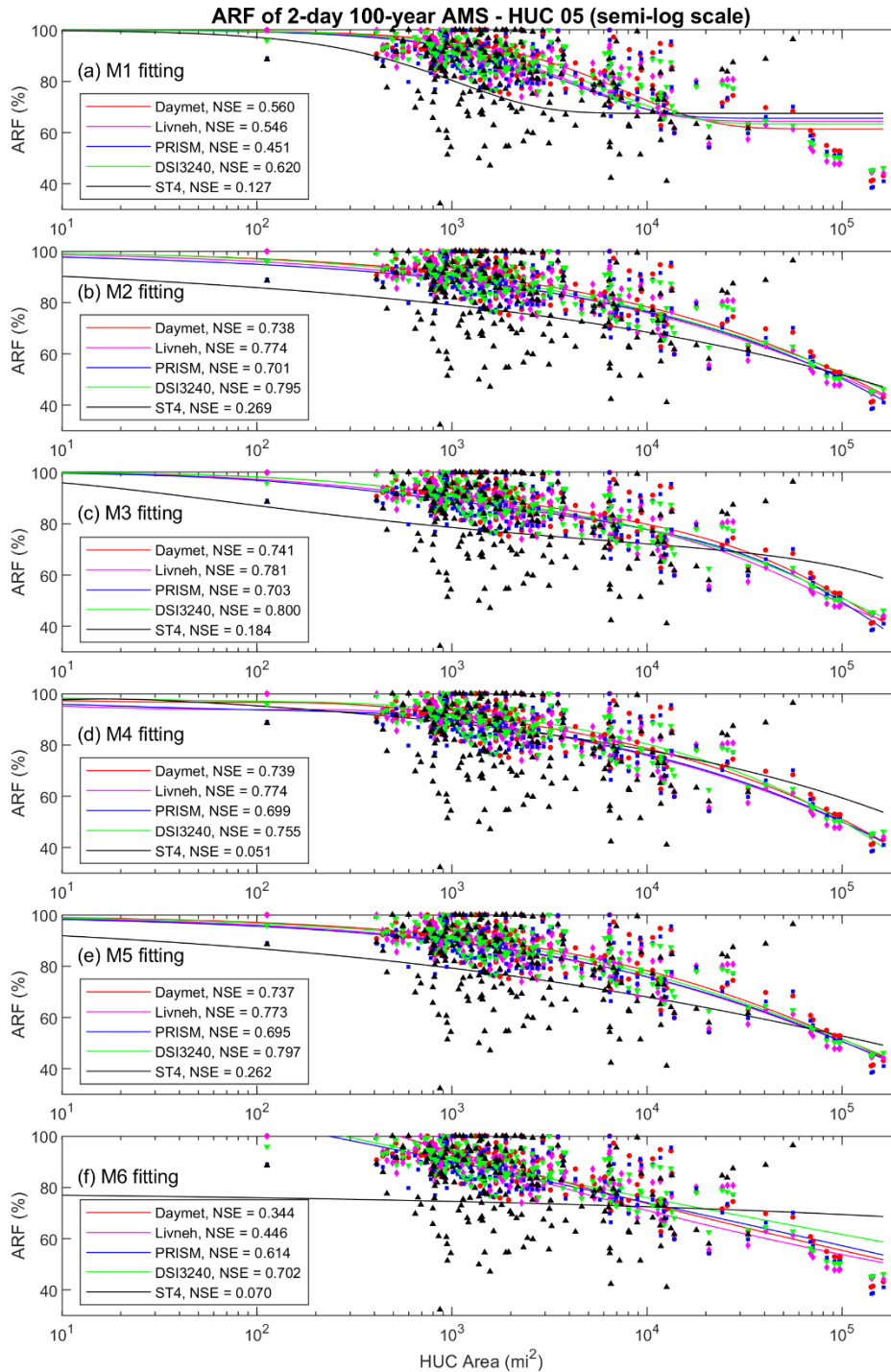


Figure C-8 Calculated Ohio ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

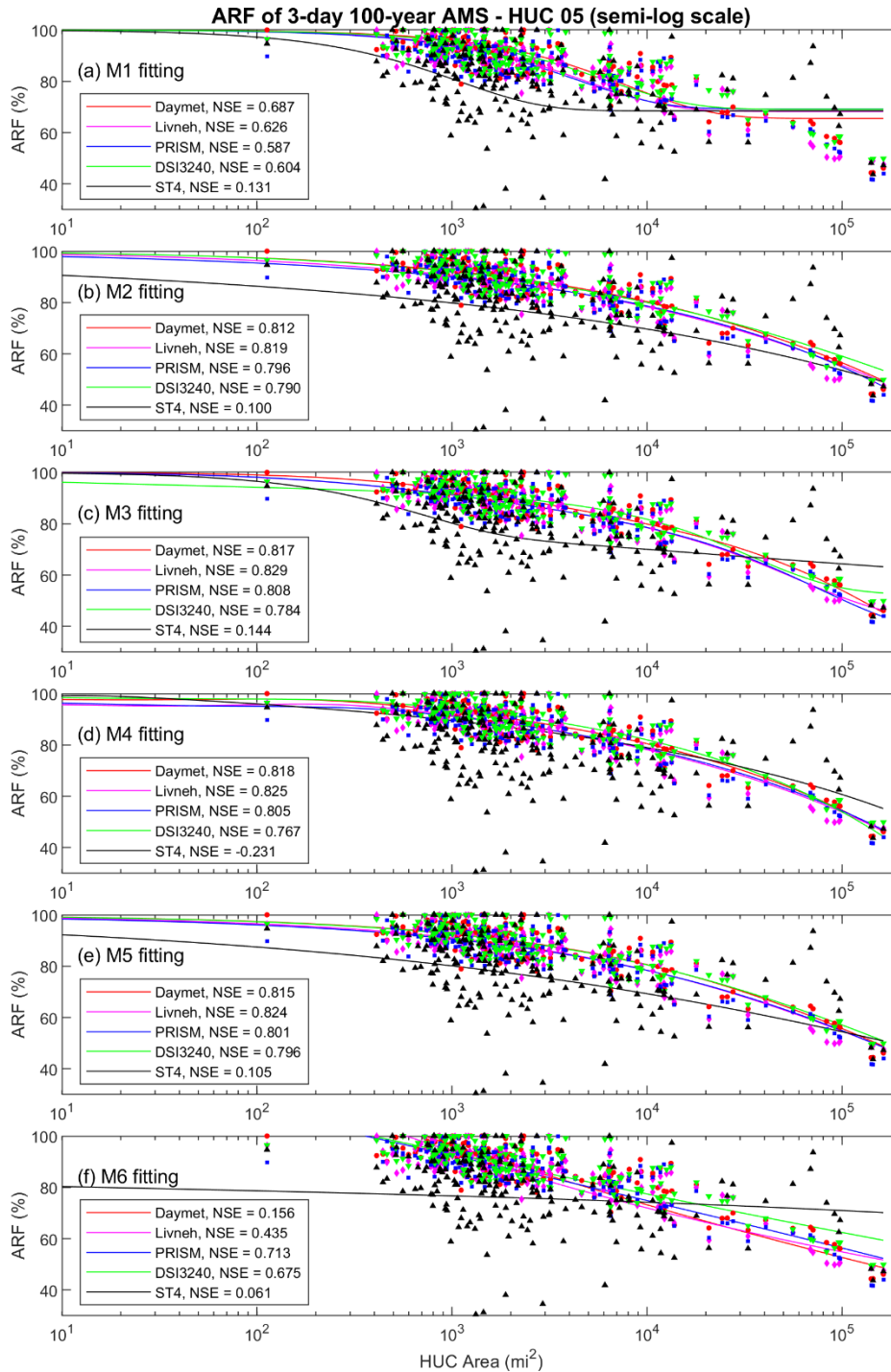


Figure C-9 Calculated Ohio ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

APPENDIX D MID-ATLANTIC (REGION 02) DETAILED RESULTS

Additional information on the Mid-Atlantic regional assessment results is provided in the figures below. Example information and discussion is provided in the main report. The plots include:

Figure D-1	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models	D-2
Figure D-2	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models	D-3
Figure D-3	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models	D-4
Figure D-4	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-5
Figure D-5	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-6
Figure D-6	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models	D-7
Figure D-7	Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-8
Figure D-8	Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-9
Figure D-9	Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	D-10

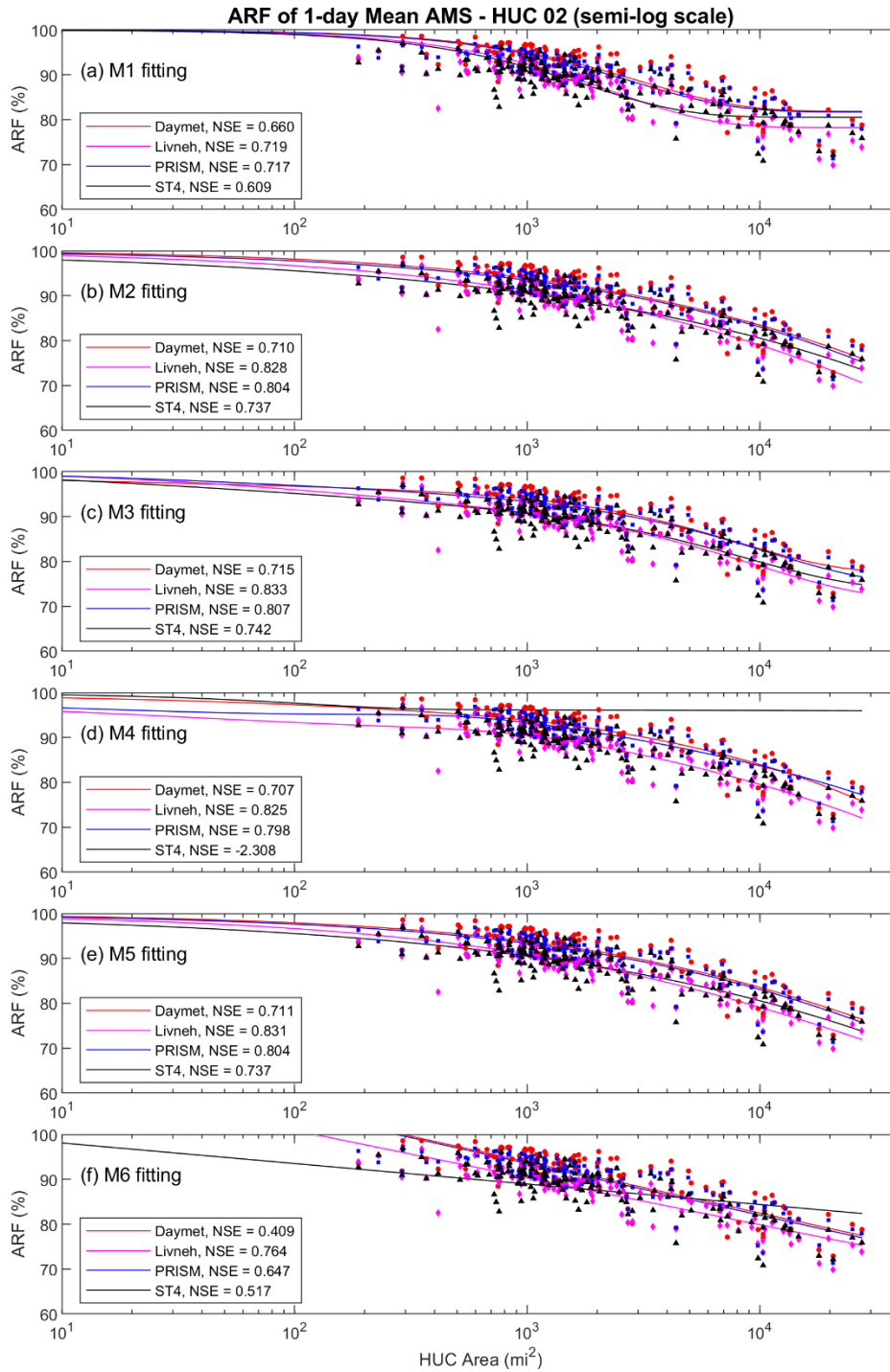


Figure D-1 Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

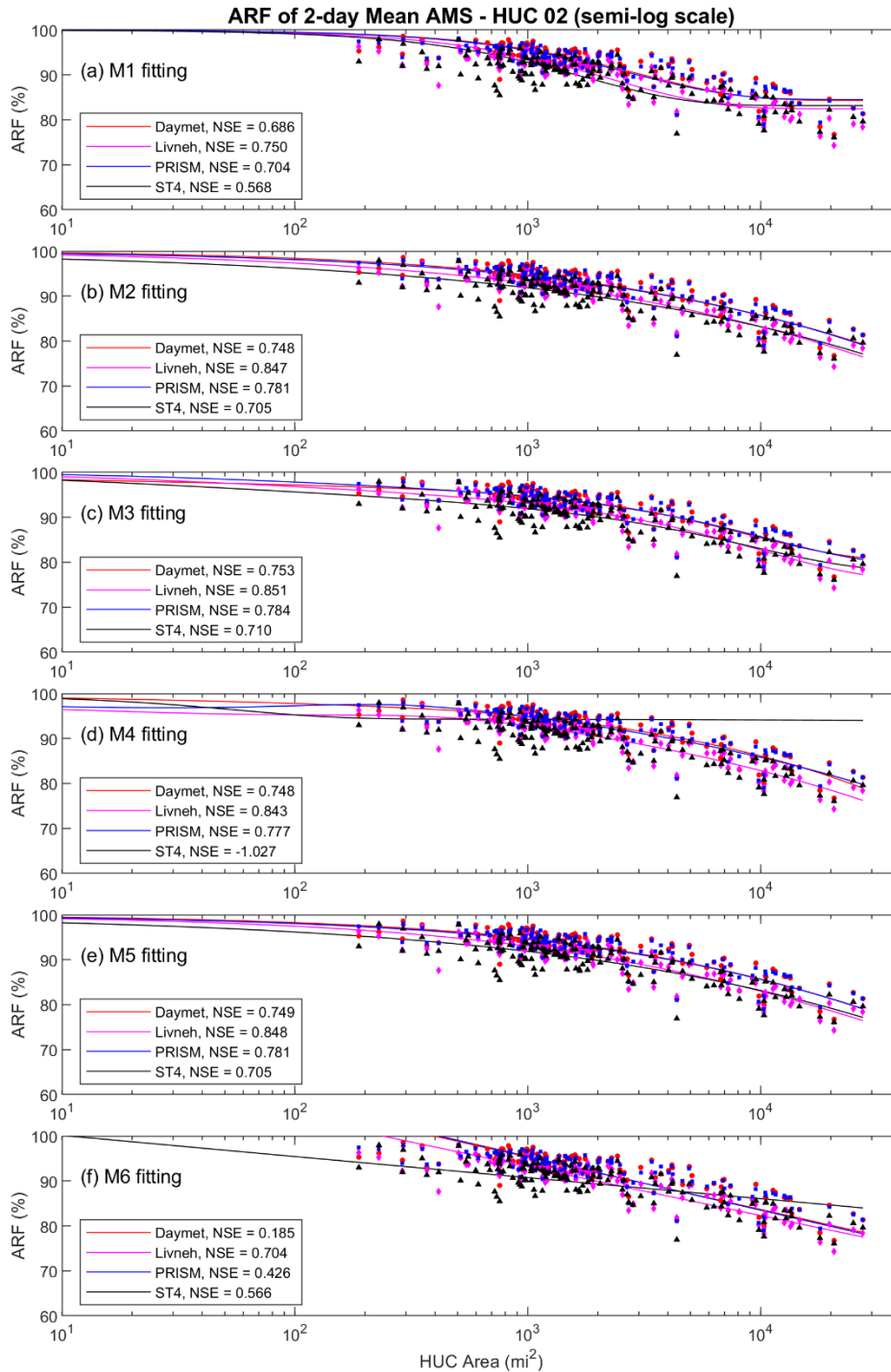


Figure D-2 Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

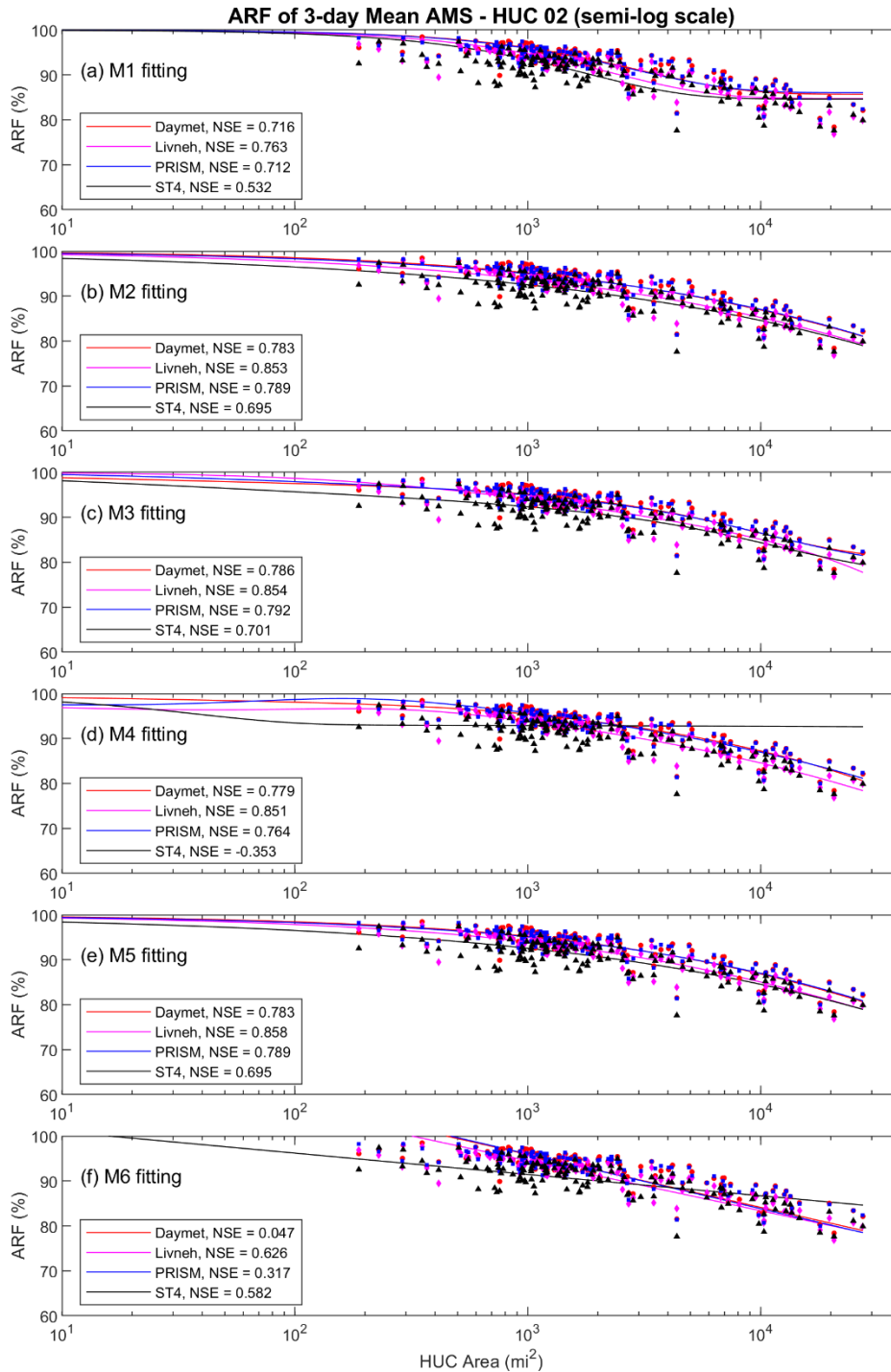


Figure D-3 Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

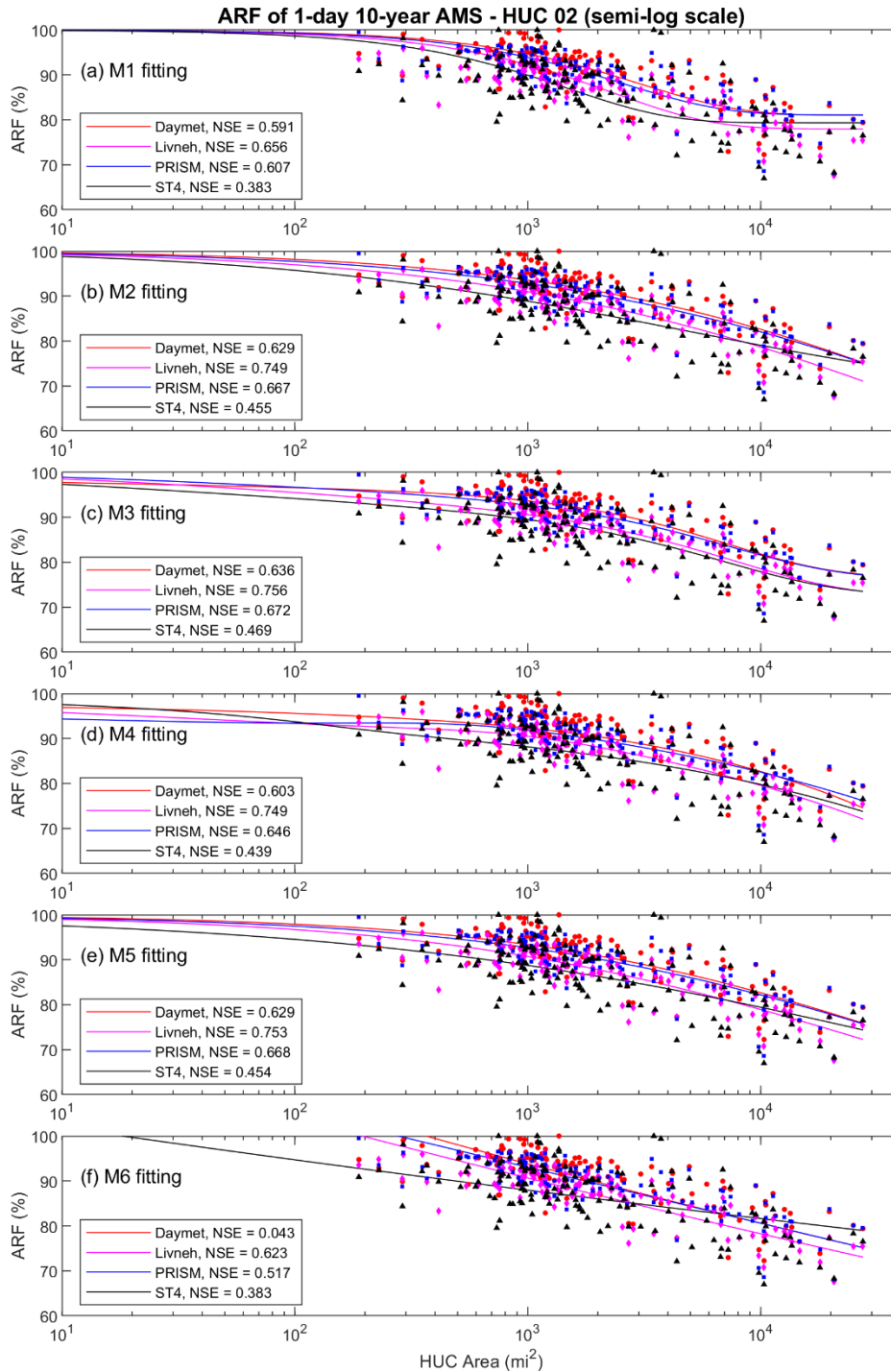


Figure D-4 Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

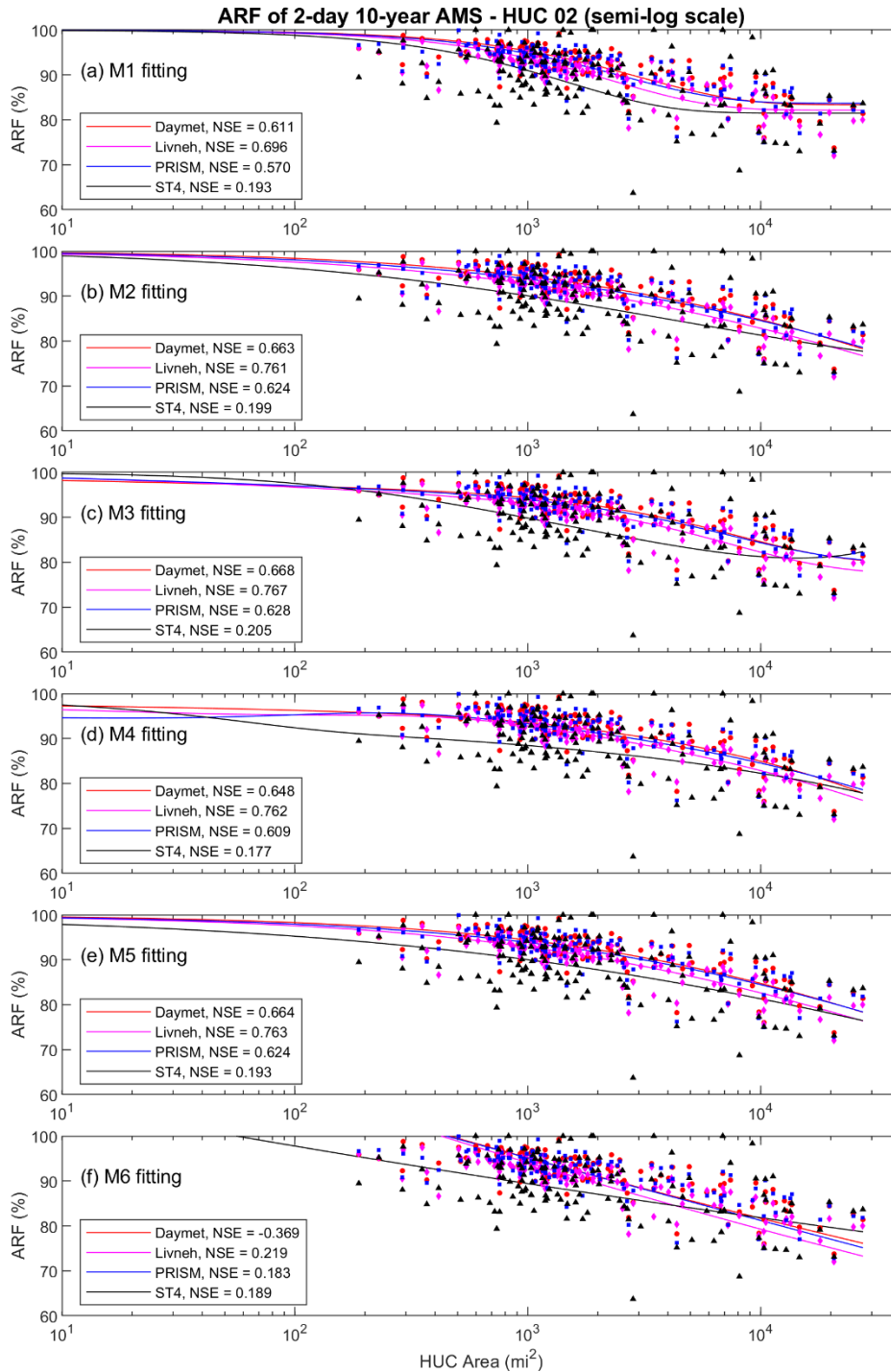


Figure D-5 Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

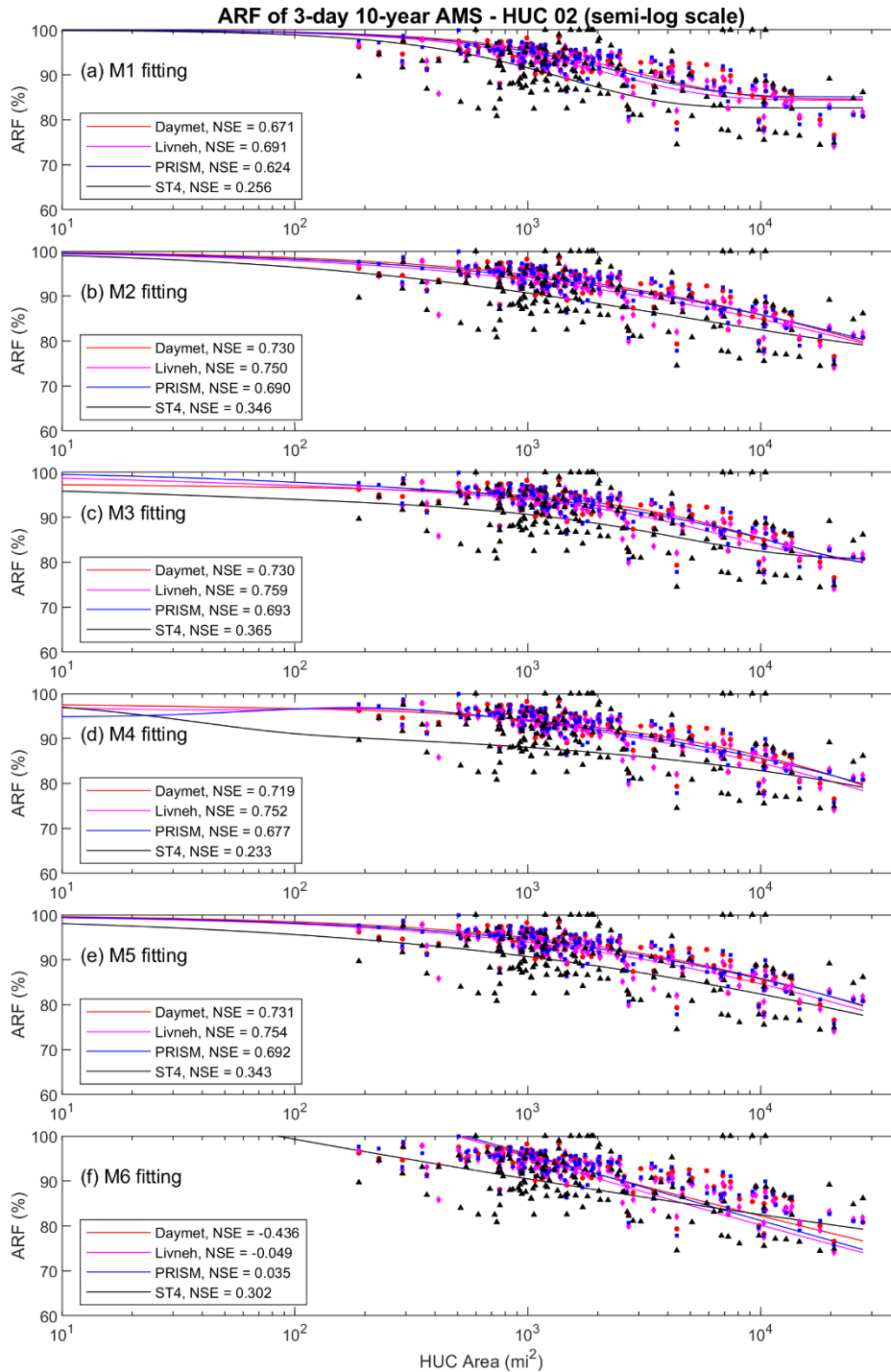


Figure D-6 Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

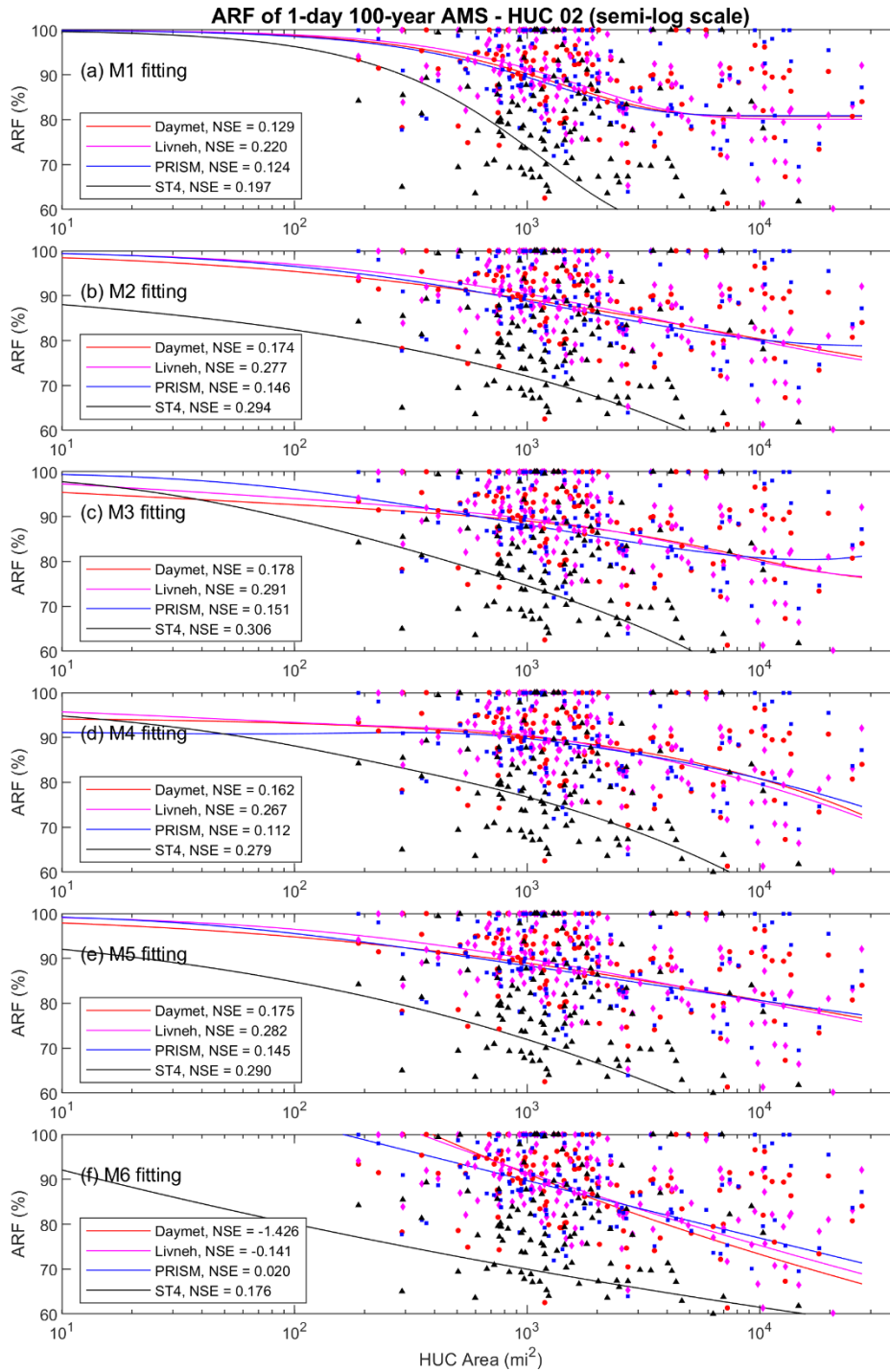


Figure D-7 Calculated Mid-Atlantic ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

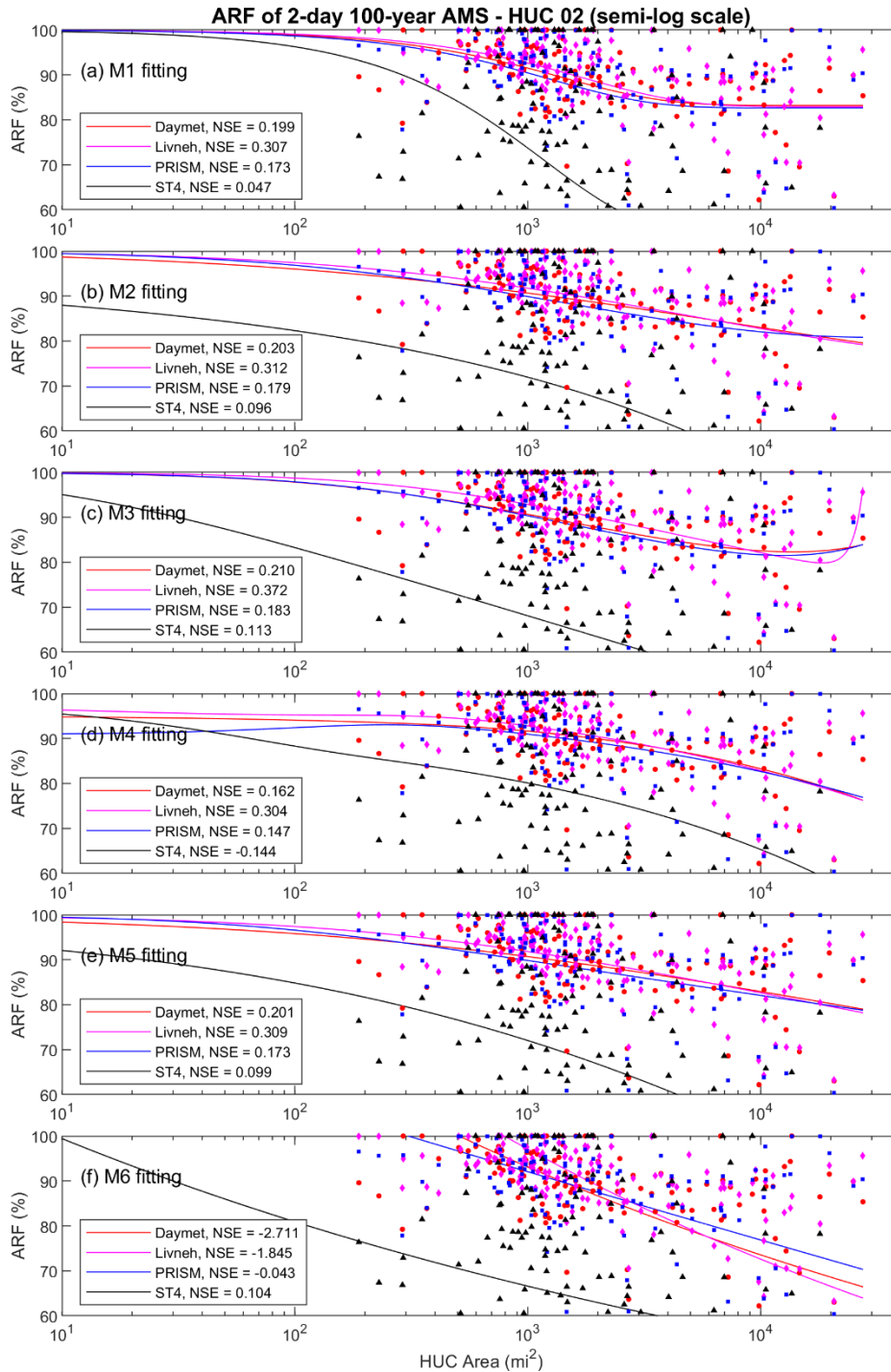


Figure D-8 Calculated Mid-Atlantic ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

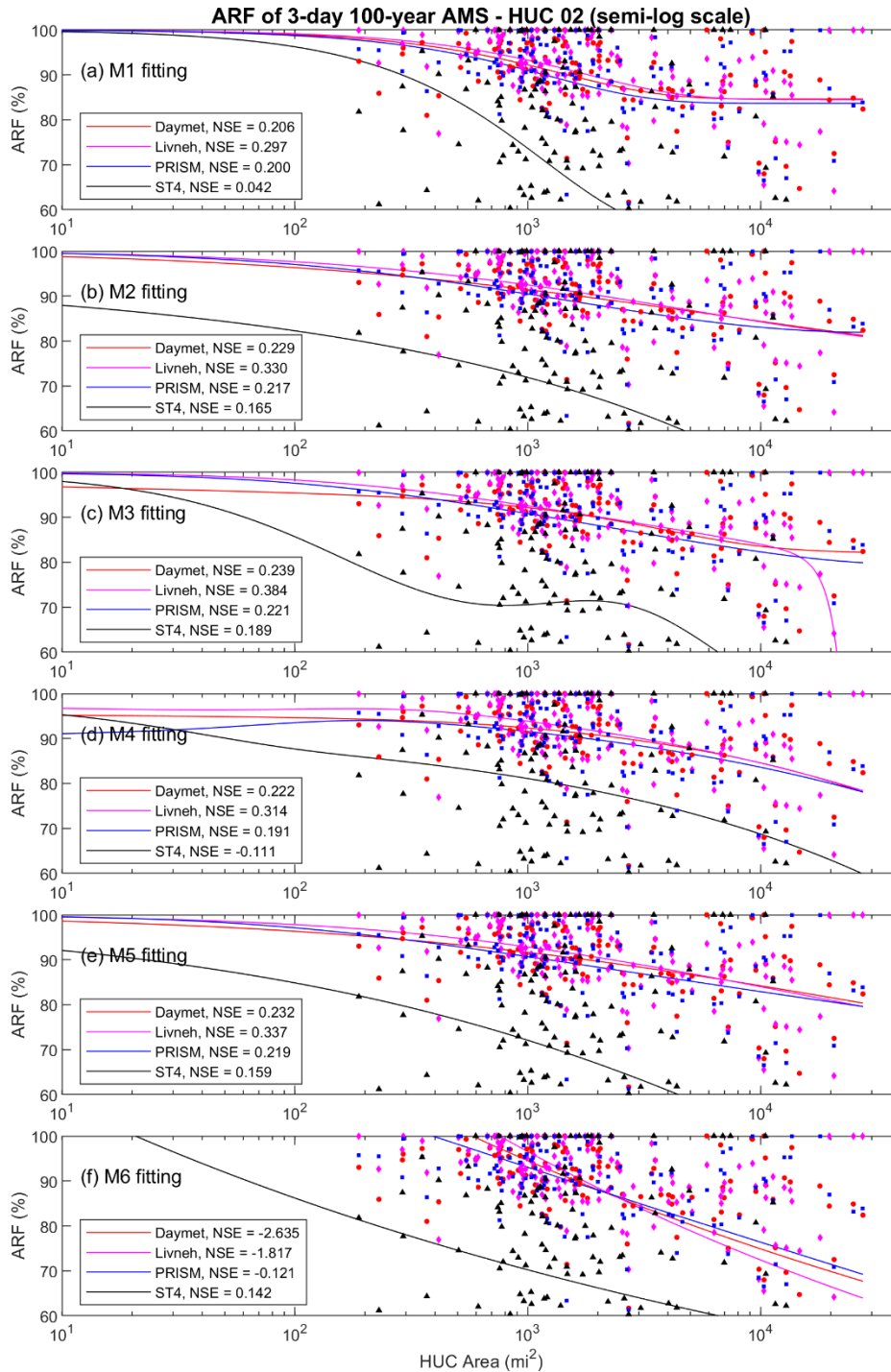


Figure D-9 Calculated Mid-Atlantic ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

APPENDIX E

SOUTH ATLANTIC-GULF (REGION 03) DETAILED RESULTS

Additional information on the South Atlantic-Gulf regional assessment results is provided in the figures below. Example information and discussion is provided in the main report. The plots include:

Figure E-1	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-12
Figure E-2	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-13
Figure E-3	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models.....	E-14
Figure E-4	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-15
Figure E-5	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-16
Figure E-6	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models.....	E-17
Figure E-7	Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-18
Figure E-8	Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-19
Figure E-9	Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models.....	E-20

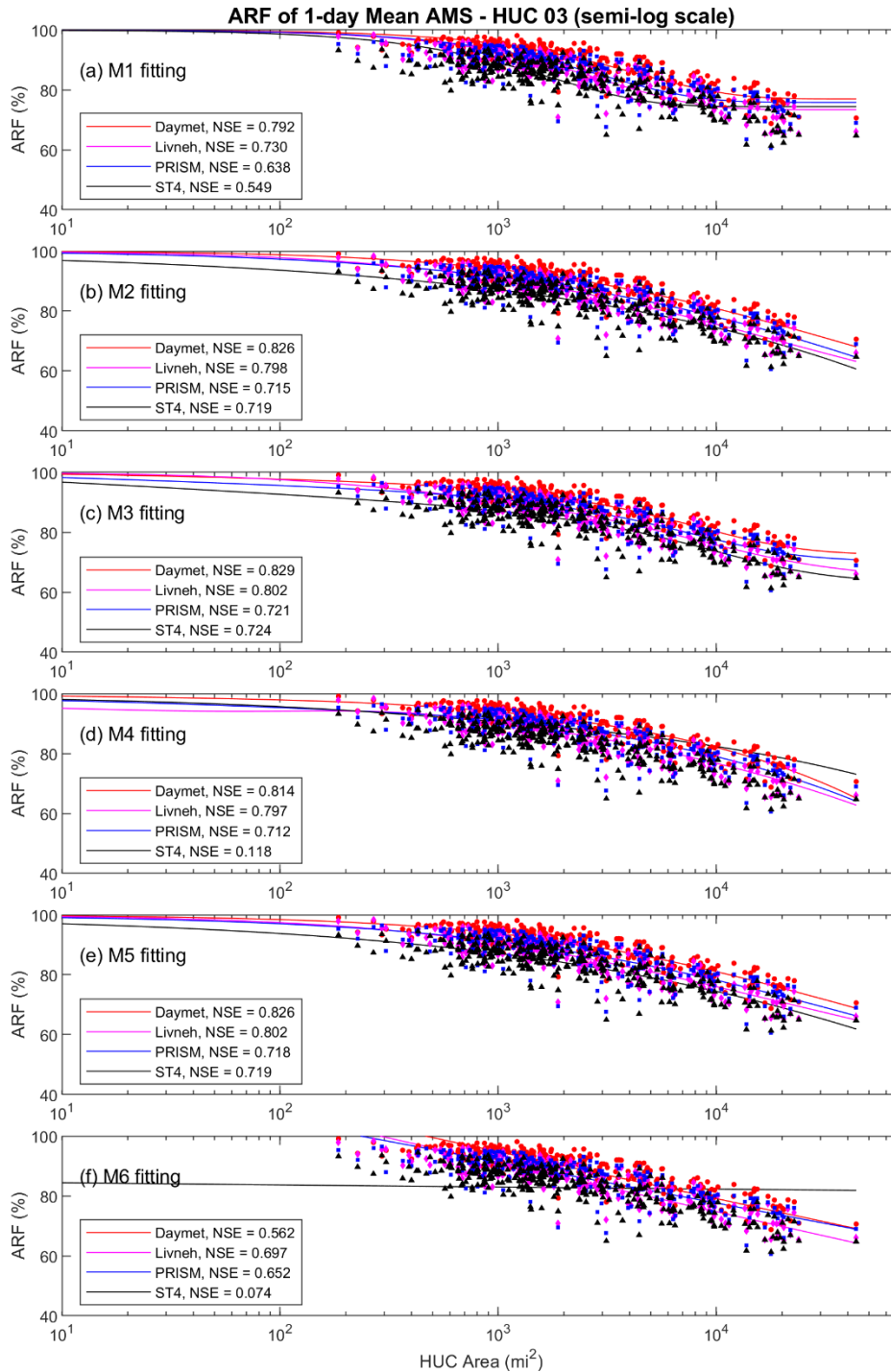


Figure E-1 Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

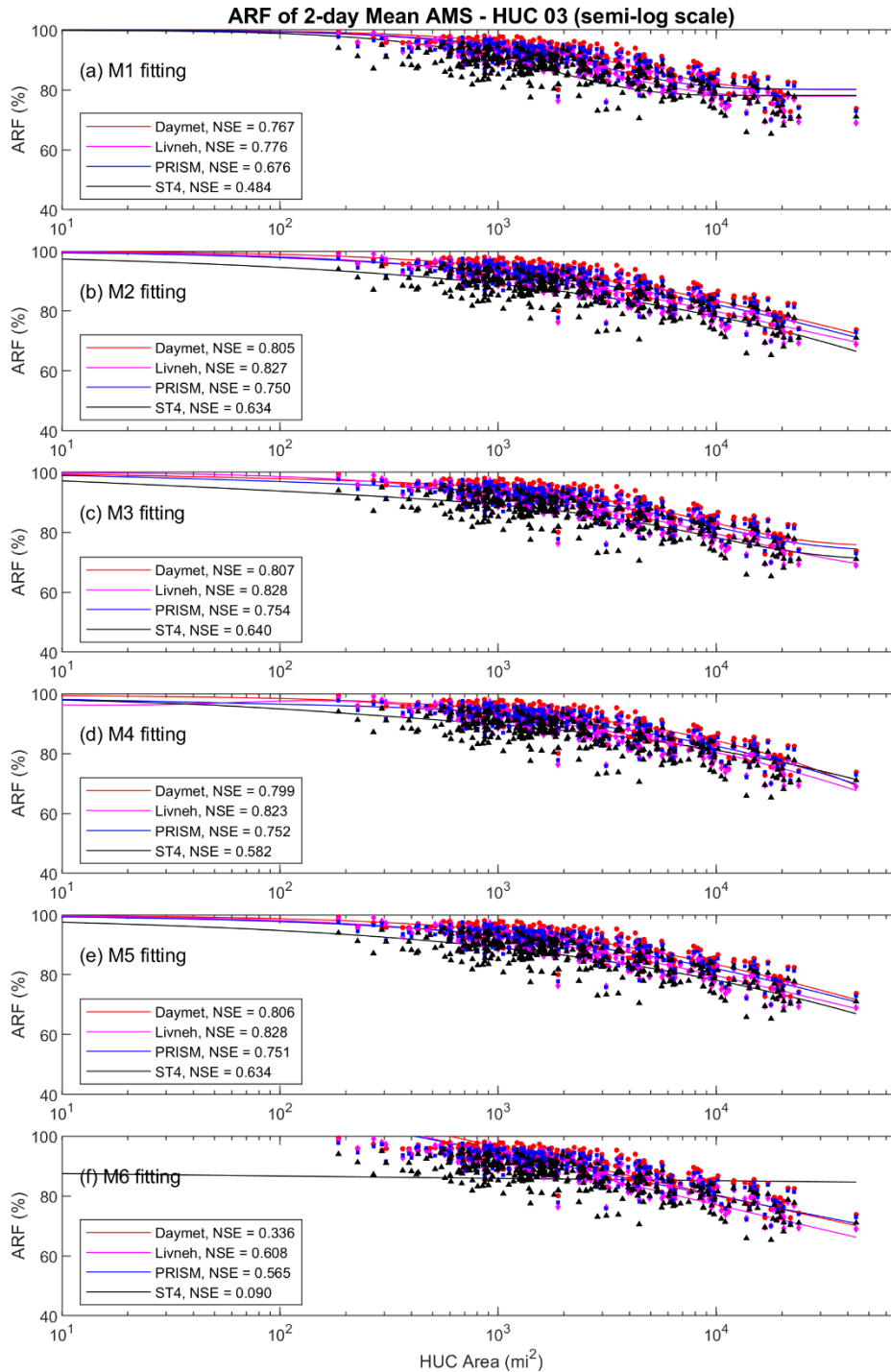


Figure E-2 Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

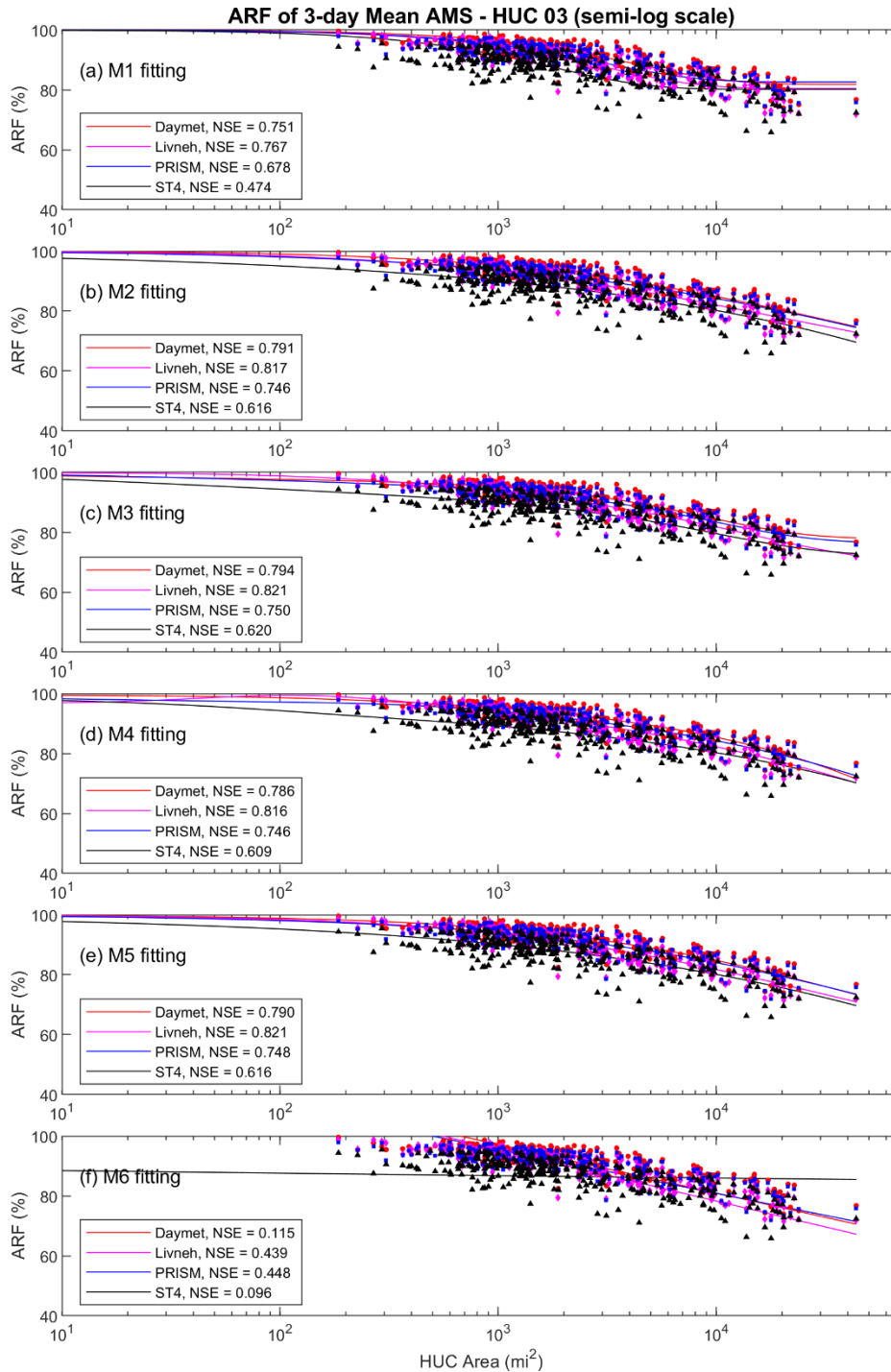


Figure E-3 Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and Average AMS Using Different Datasets and Fitted Models *Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.*

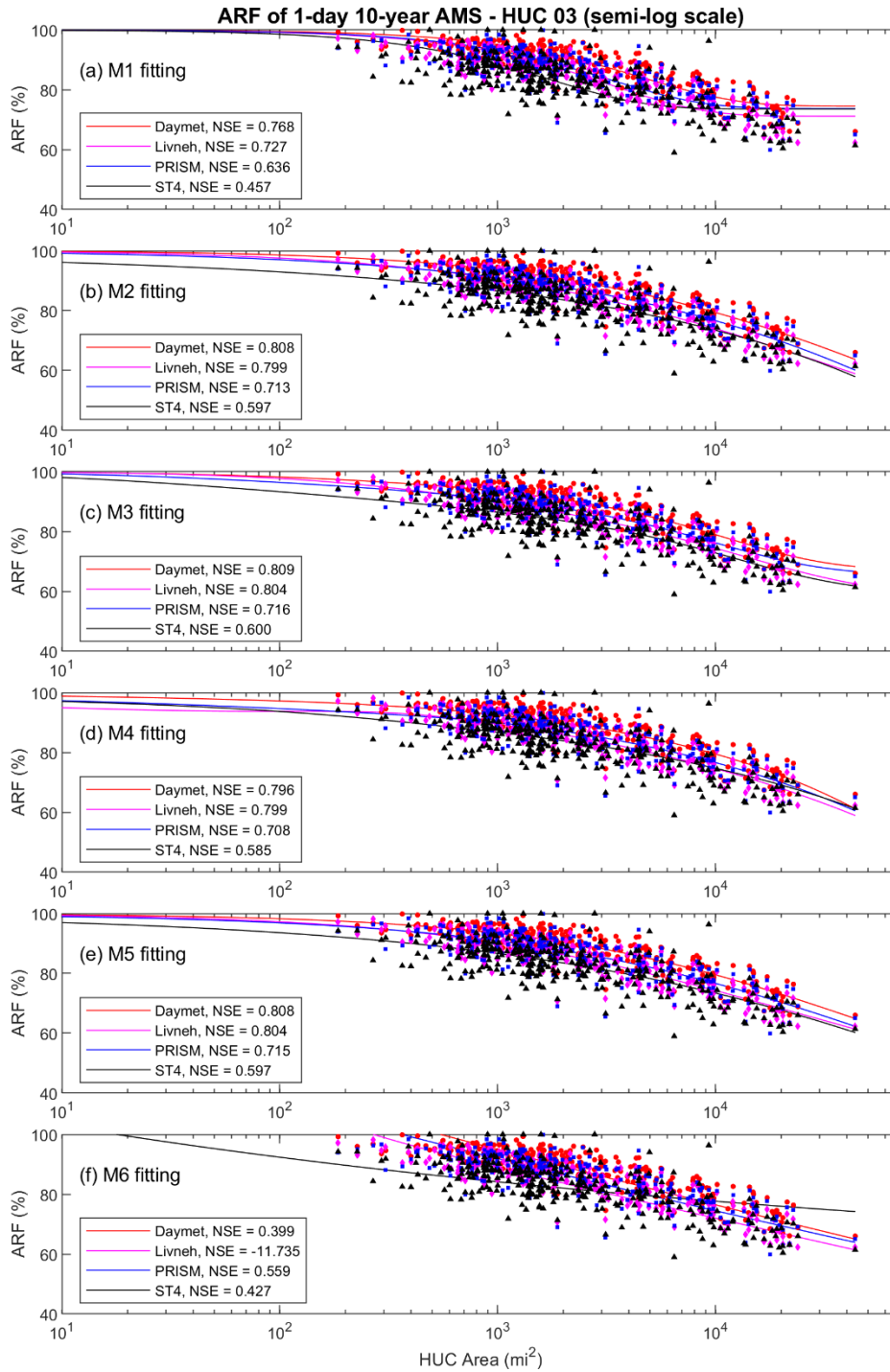


Figure E-4 Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 10-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

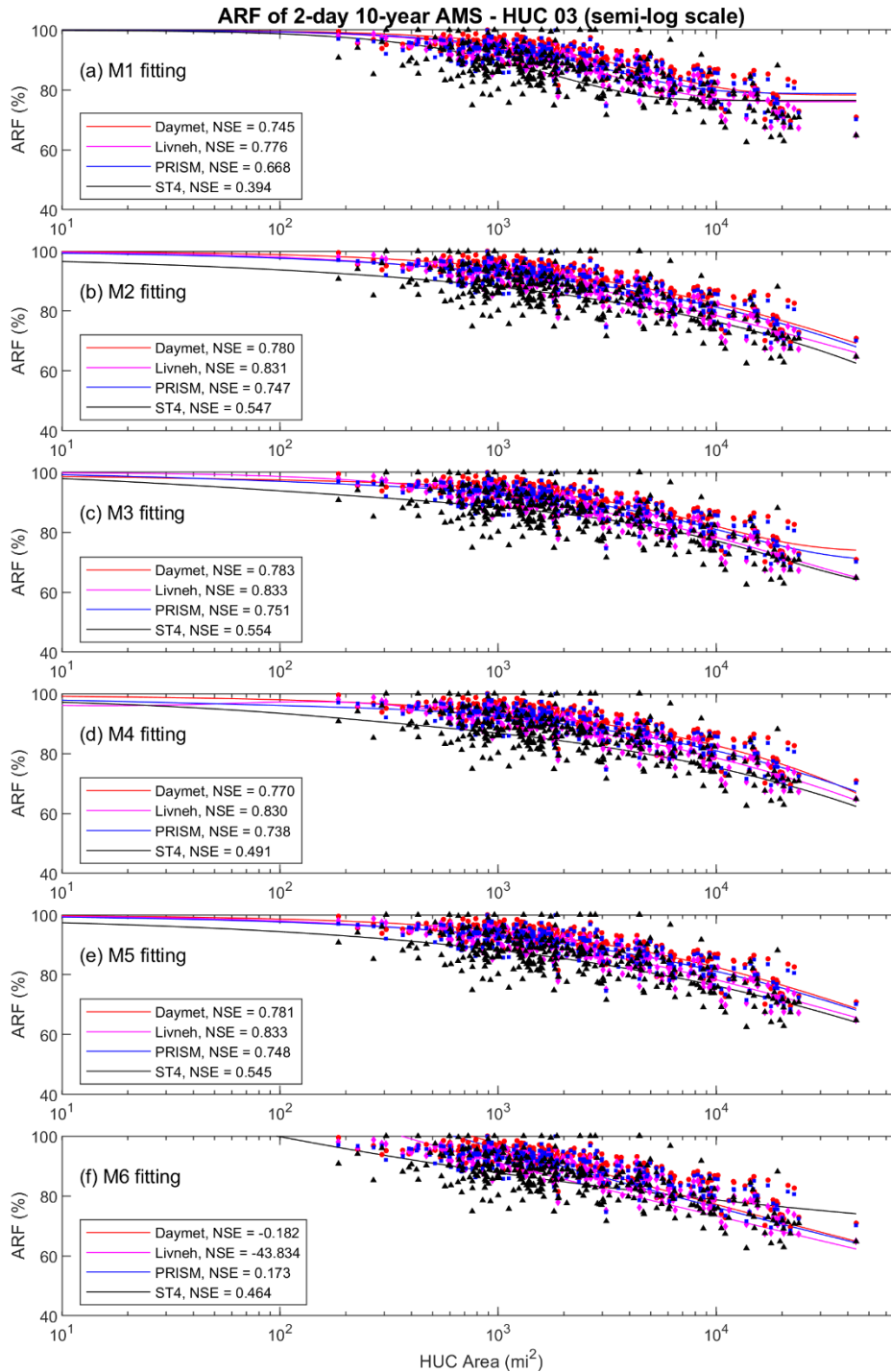


Figure E-5 Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 10-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

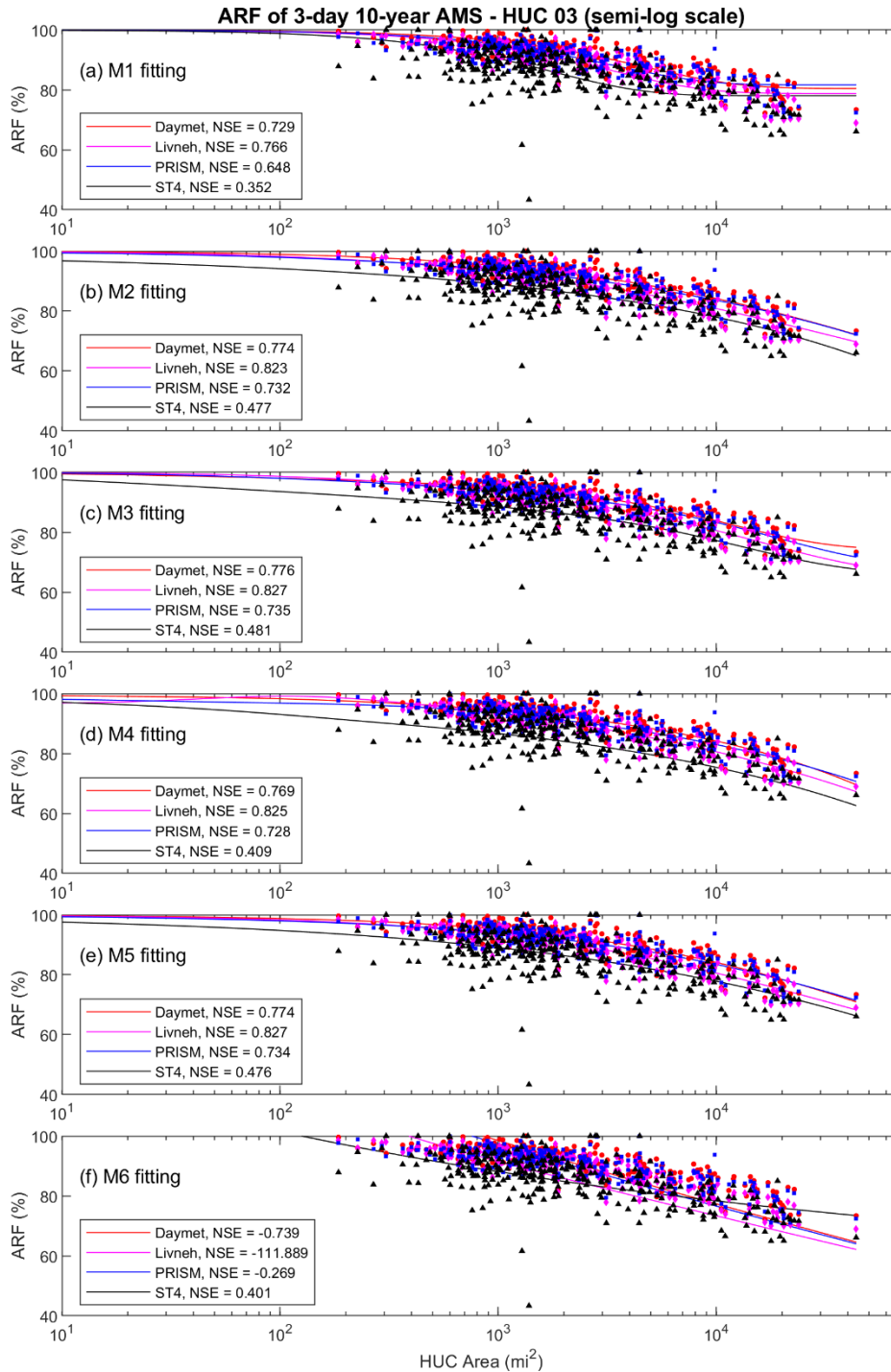


Figure E-6 Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 10-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

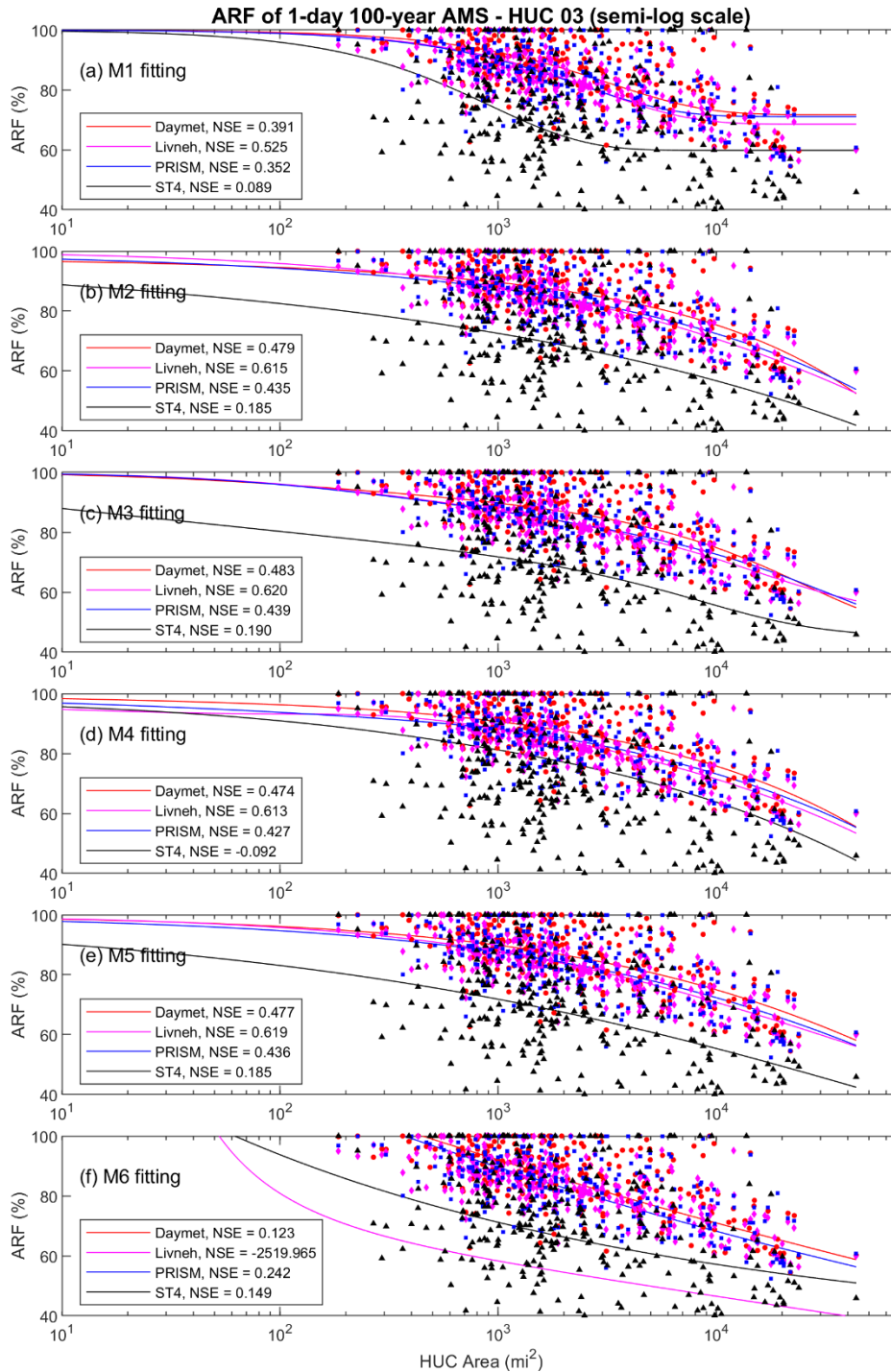


Figure E-7 Calculated South Atlantic-Gulf ARFs and Fitted Models for 1-day Duration and 100-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

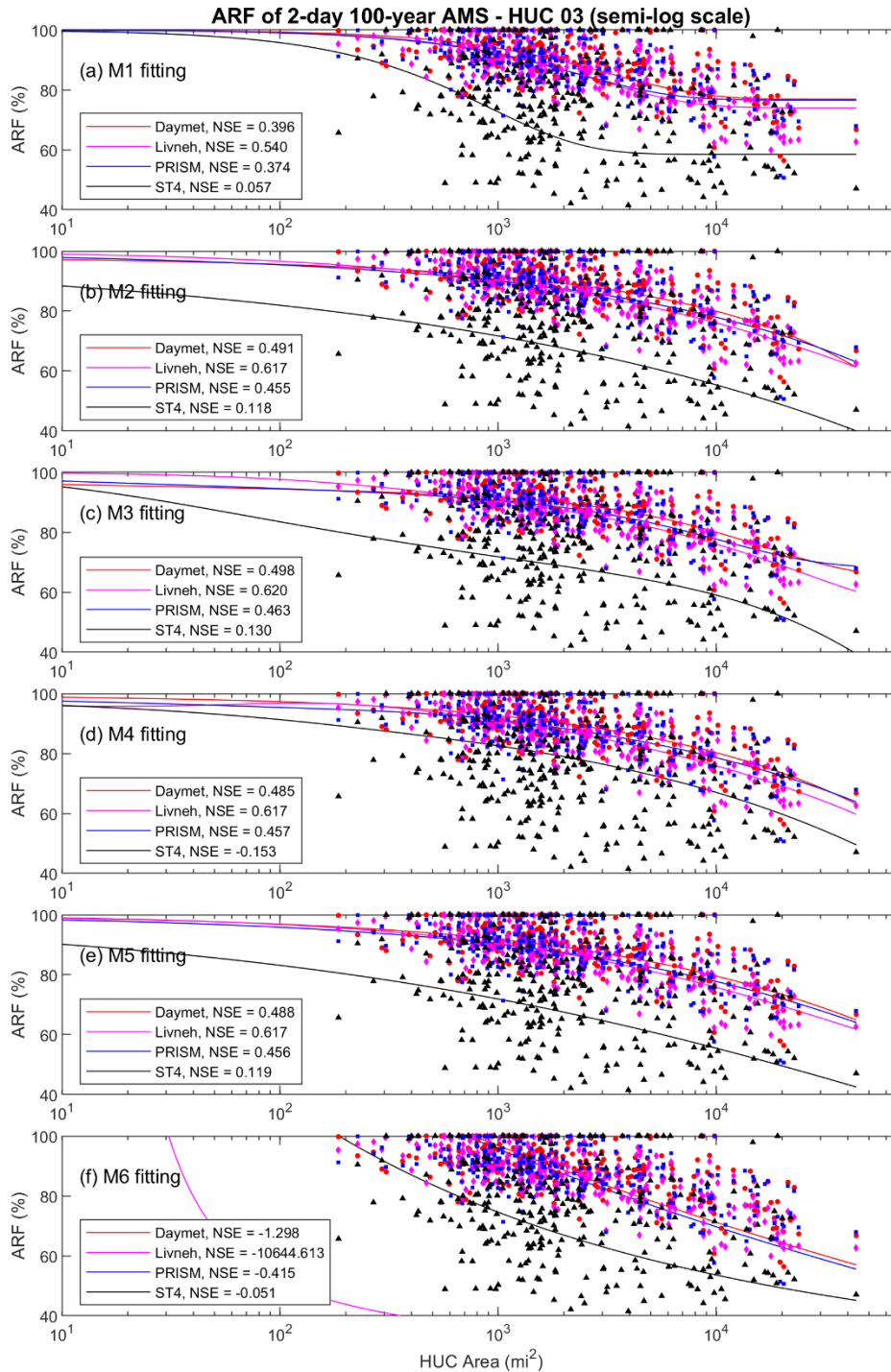


Figure E-8 Calculated South Atlantic-Gulf ARFs and Fitted Models for 2-day Duration and 100-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

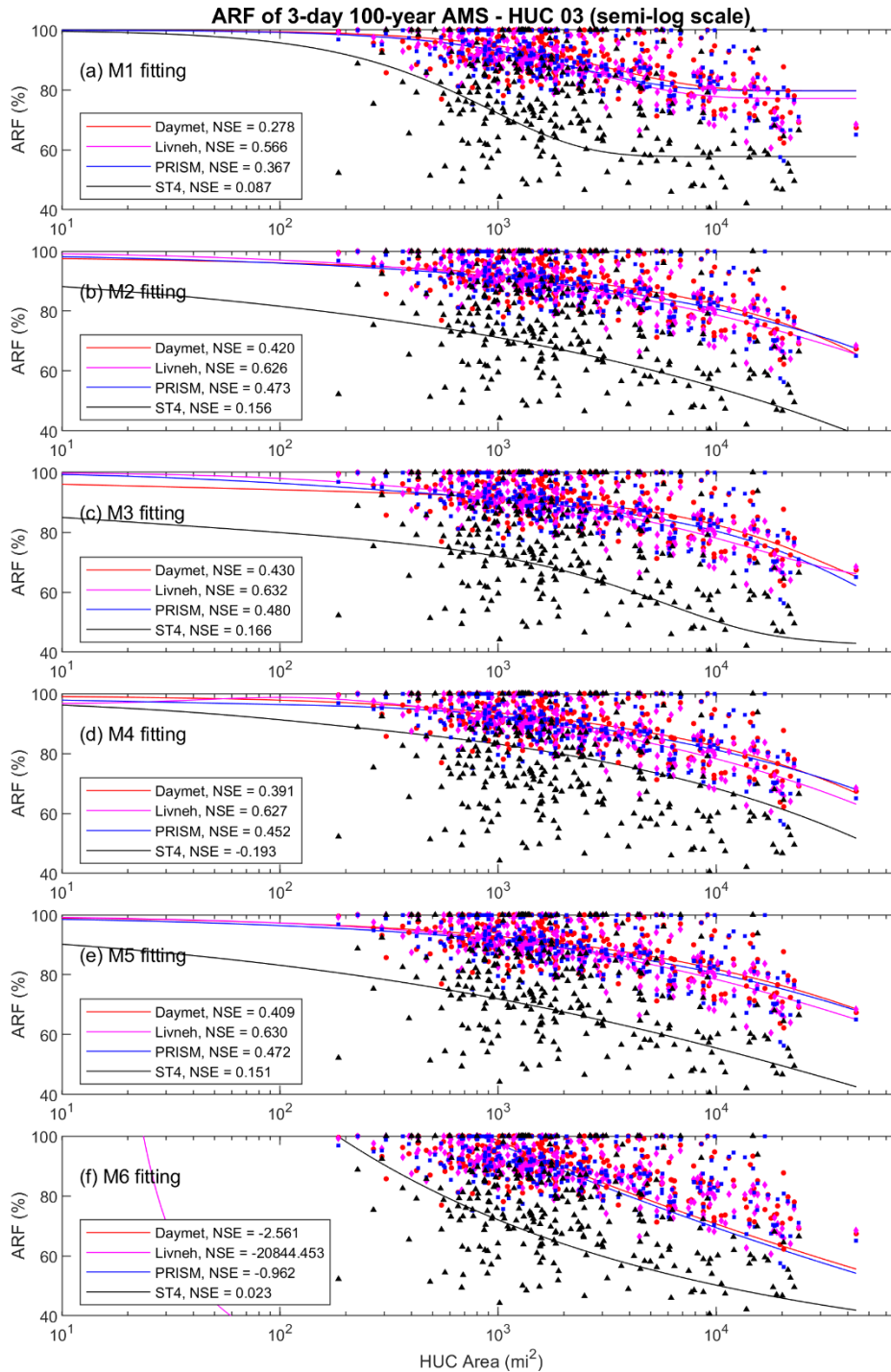


Figure E-9 Calculated South Atlantic-Gulf ARFs and Fitted Models for 3-day Duration and 100-year Return Period Using Different Datasets and Fitted Models
Colored dots represent calculated ARFs across HUC units; solid curves represent fitted models.

APPENDIX F

CONUS ASSESSMENT DETAILED RESULTS

Additional information on the CONUS assessment results is provided in the figures below. Example information and discussion is provided in Section 4.4 of the main report. These results stem from ARF analysis using daily PRISM data and M5 fitting. The plots include:

Figure F-1	CONUS ARF Assessment Maps for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes.....	F-2
Figure F-2	CONUS ARF Assessment Maps for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-3
Figure F-3	CONUS ARF Assessment Maps for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-4
Figure F-4	CONUS ARF Assessment Maps for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-5
Figure F-5	CONUS ARF Assessment Maps for 3-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes	F-6
Figure F-6	CONUS ARF Assessment Plot for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting	F-7
Figure F-7	CONUS ARF Assessment Plot for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-8
Figure F-8	CONUS ARF Assessment Plot for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-9
Figure F-9	CONUS ARF Assessment Plot for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-10
Figure F-10	CONUS ARF Assessment Plot for 3-day duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting	F-11

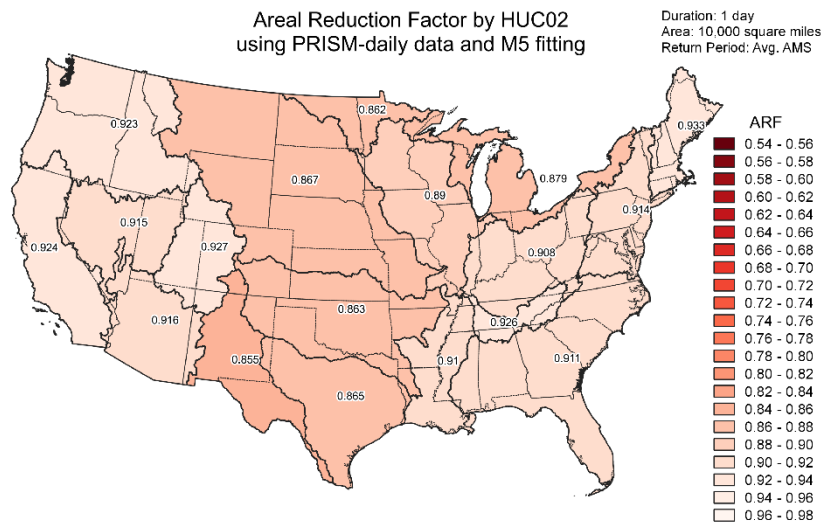
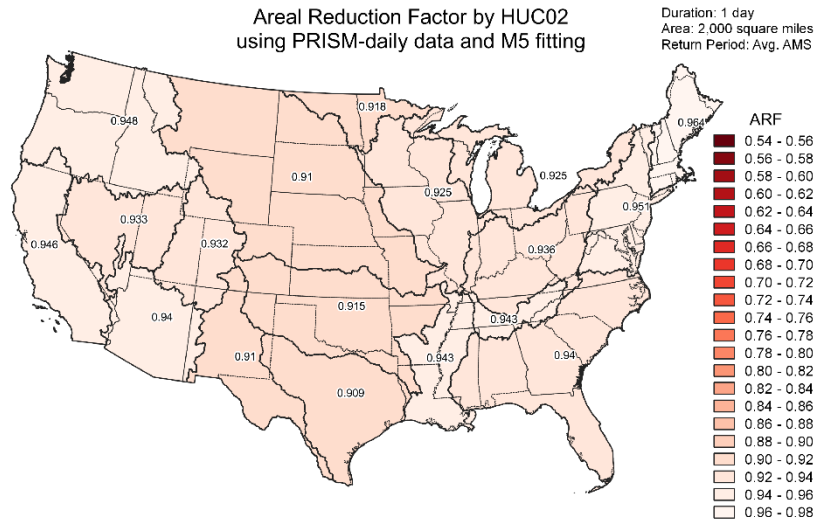
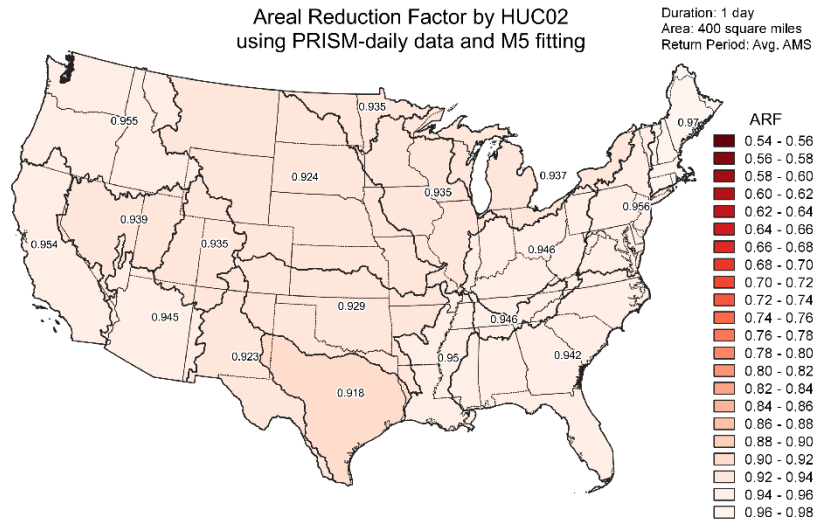


Figure F-1 CONUS ARF Assessment Maps for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes

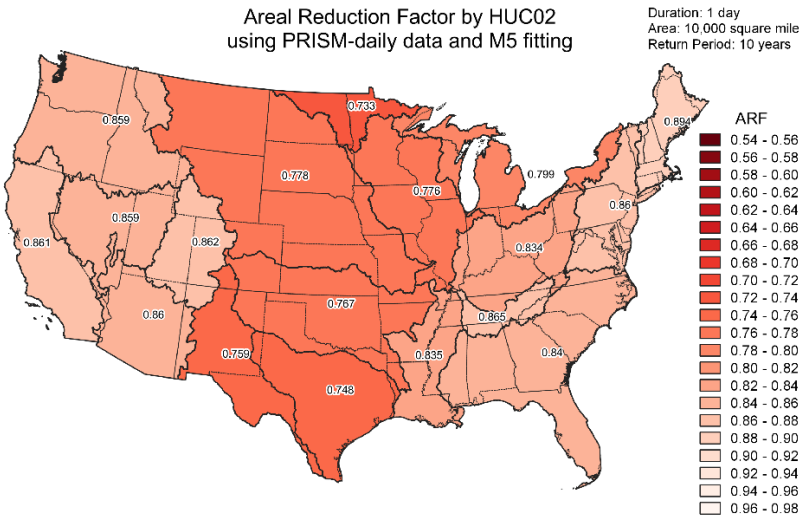
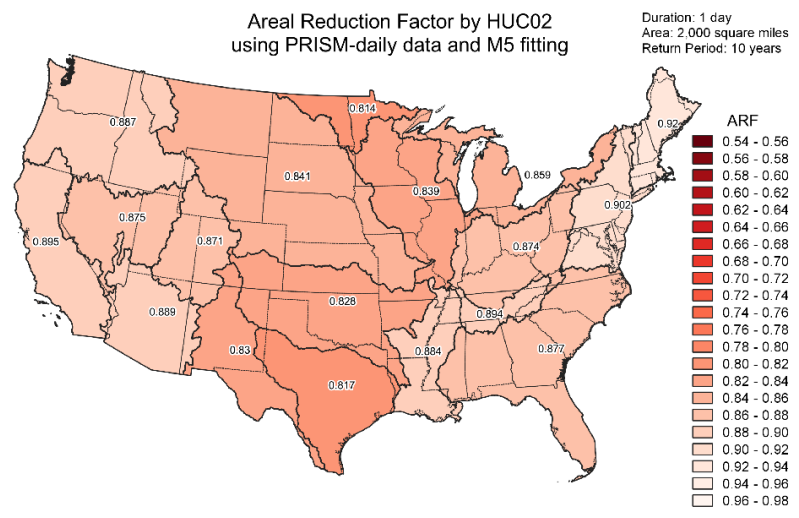
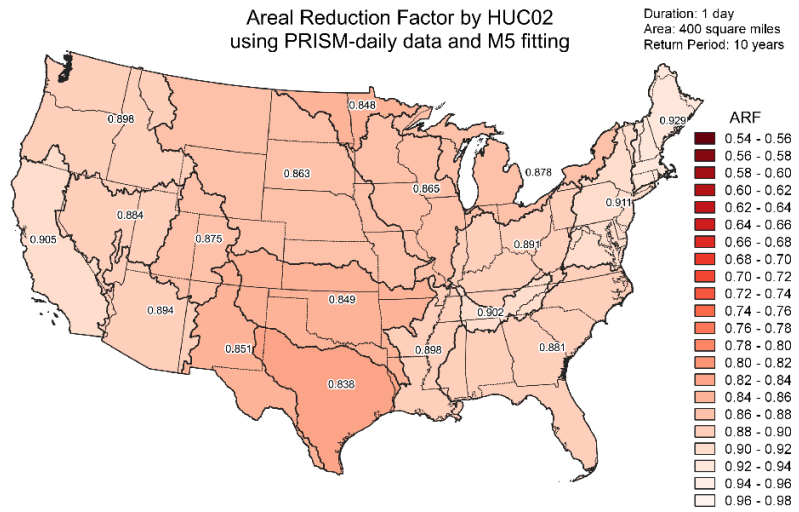


Figure F-2 CONUS ARF Assessment Maps for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes

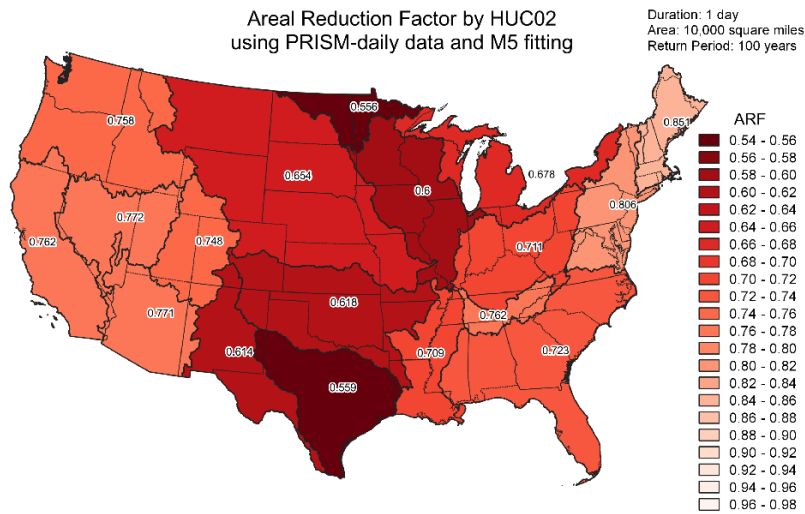
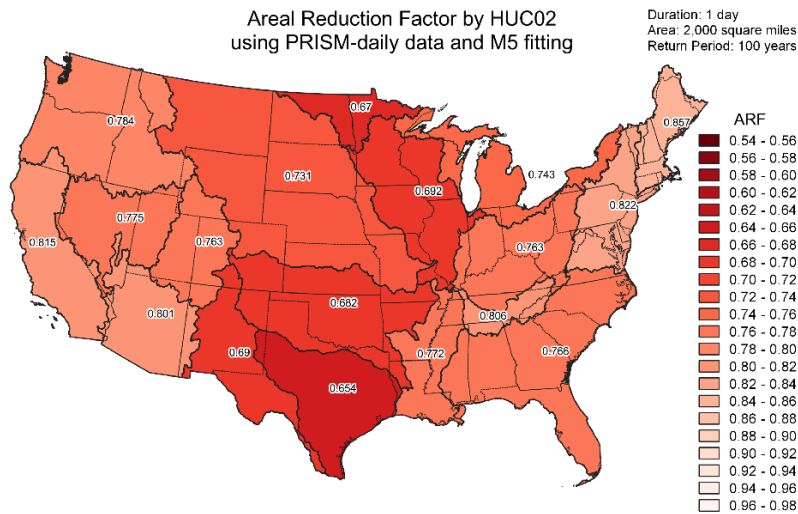
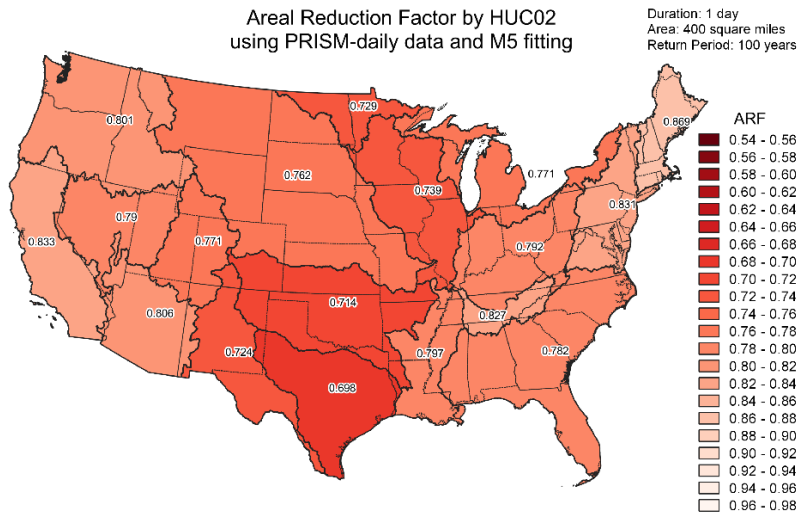


Figure F-3 CONUS ARF Assessment Maps for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes

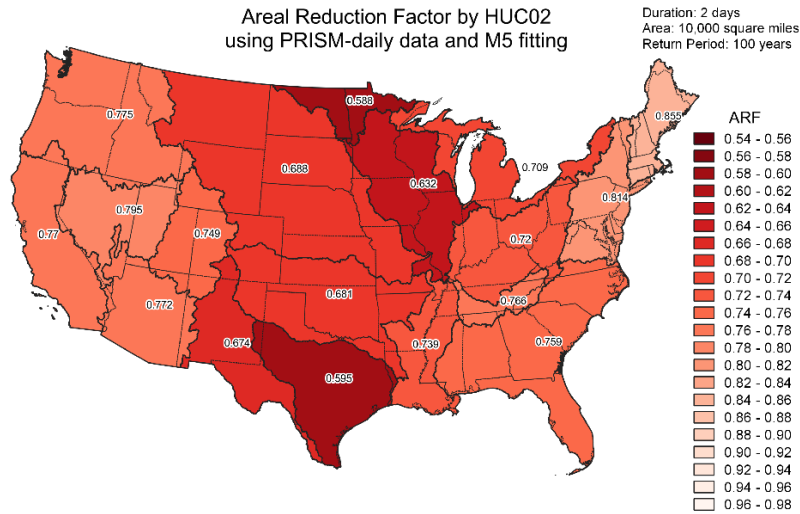
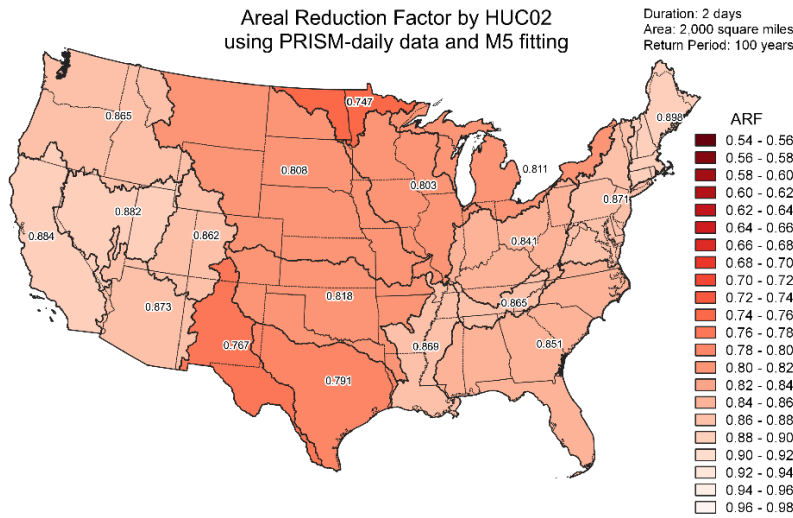
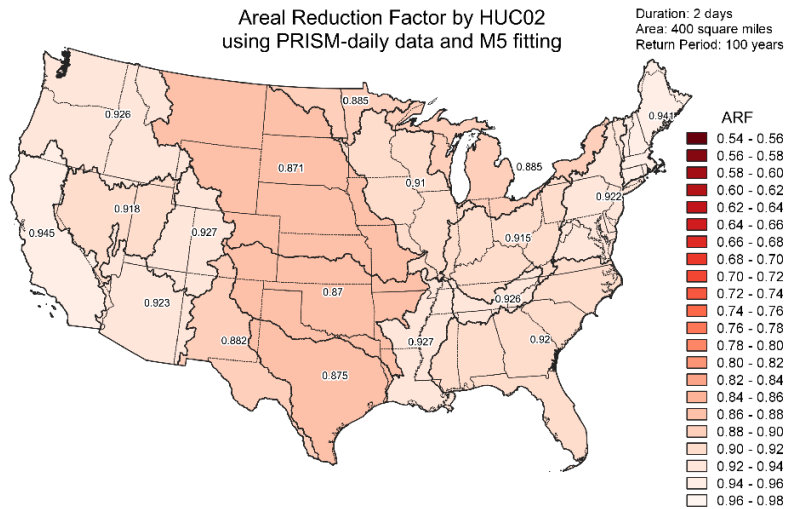


Figure F-4 CONUS ARF Assessment Maps for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes

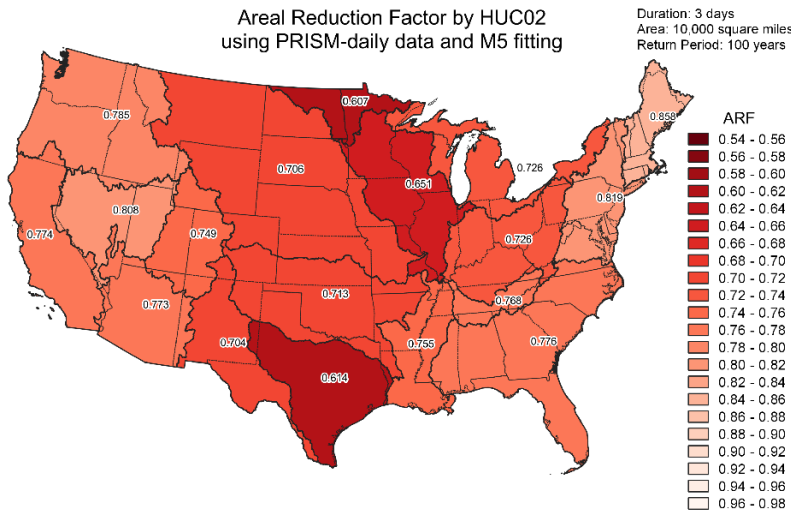
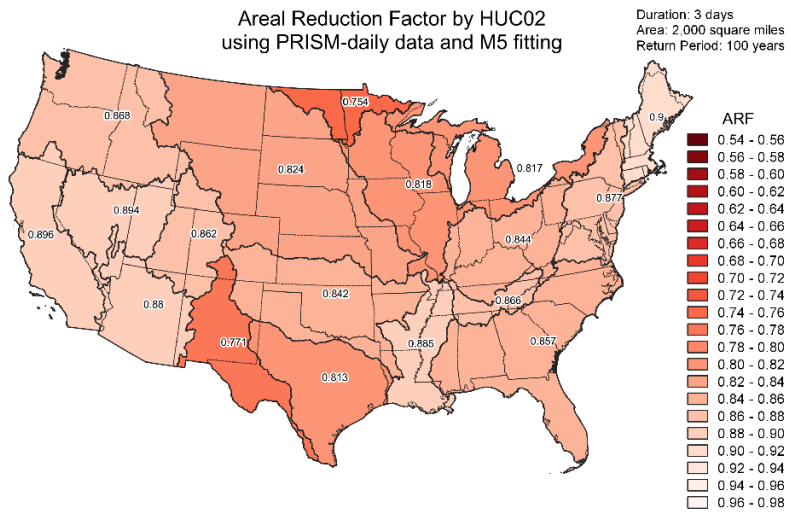
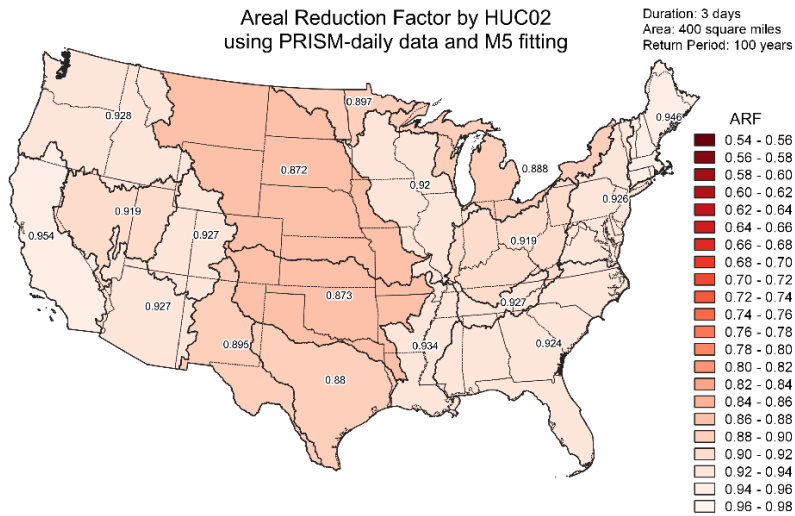


Figure F-5 CONUS ARF Assessment Maps for 3-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting across Different Area Sizes

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting

1-day Duration | Average AMS (~2-year return period)

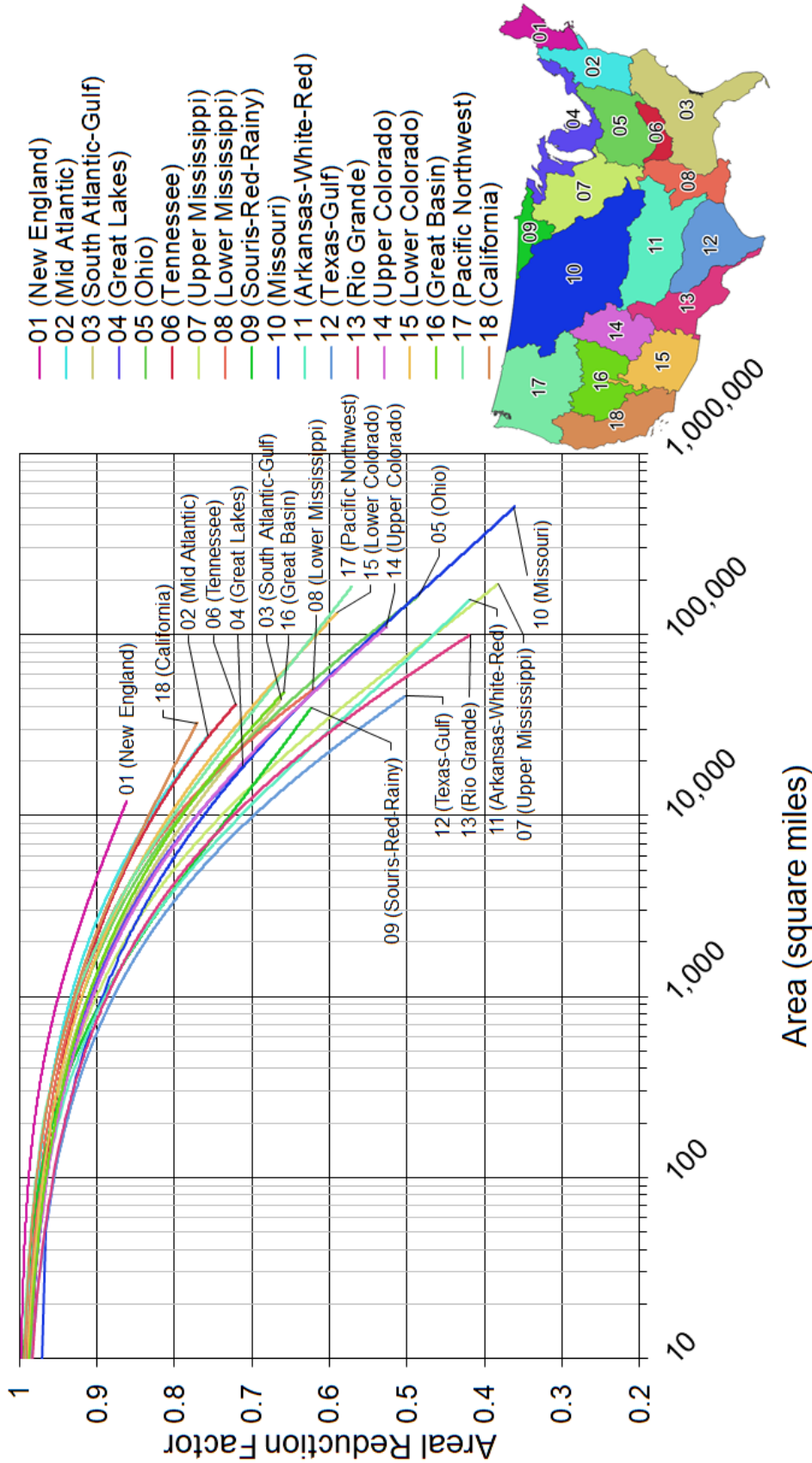


Figure F-6 CONUS ARF Assessment Plot for 1-day Duration and Average AMS Using PRISM-daily Data and M5 Model Fitting

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting 1-day Duration | 10-y Return Period

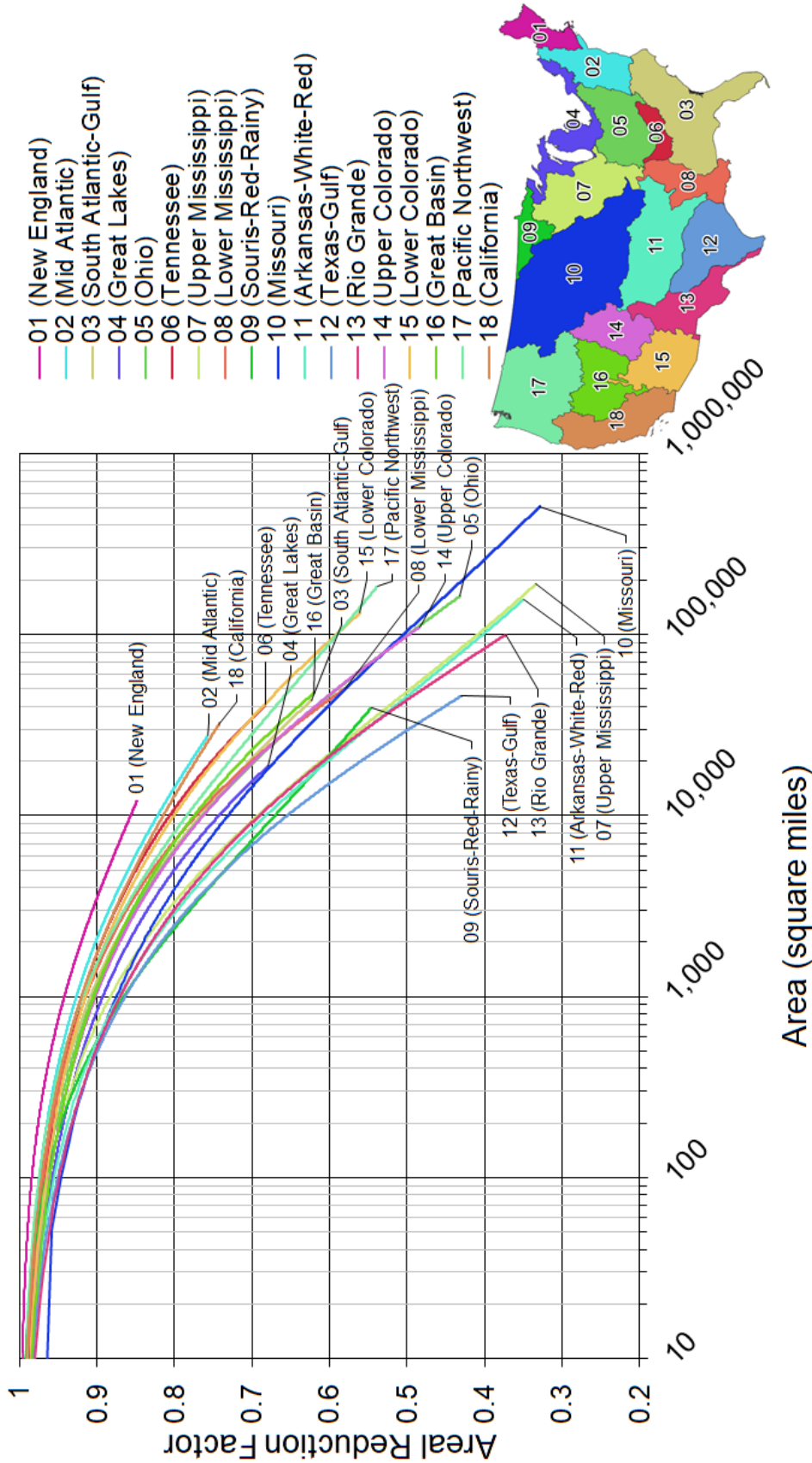


Figure F-7 CONUS ARF Assessment Plot for 1-day Duration and 10-year Return Period Using PRISM-daily Data and M5 Model Fitting

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting 1-day Duration | 100-year Return Period

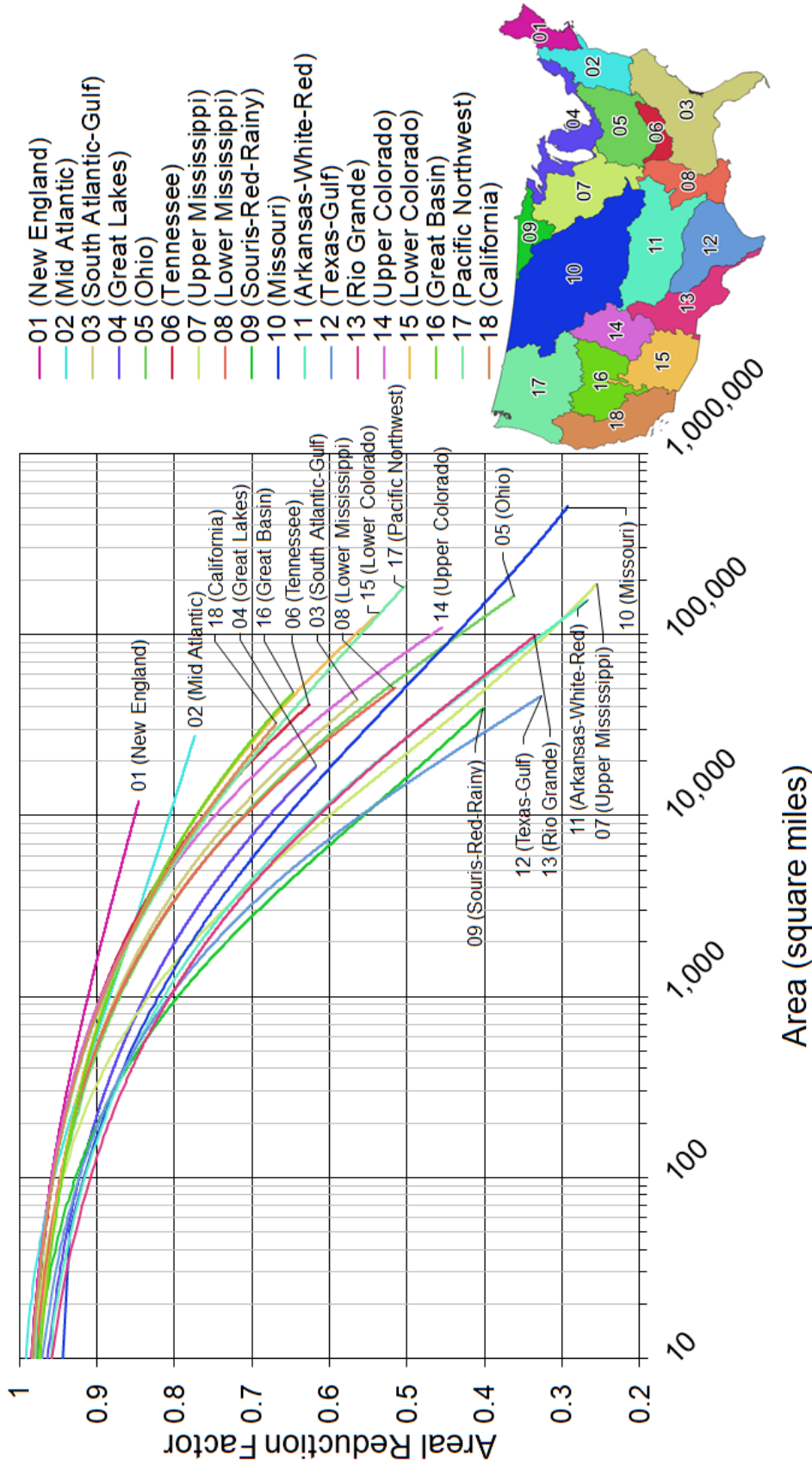


Figure F-8 CONUS ARF Assessment Plot for 1-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting 2-day Duration | 100-y Return Period

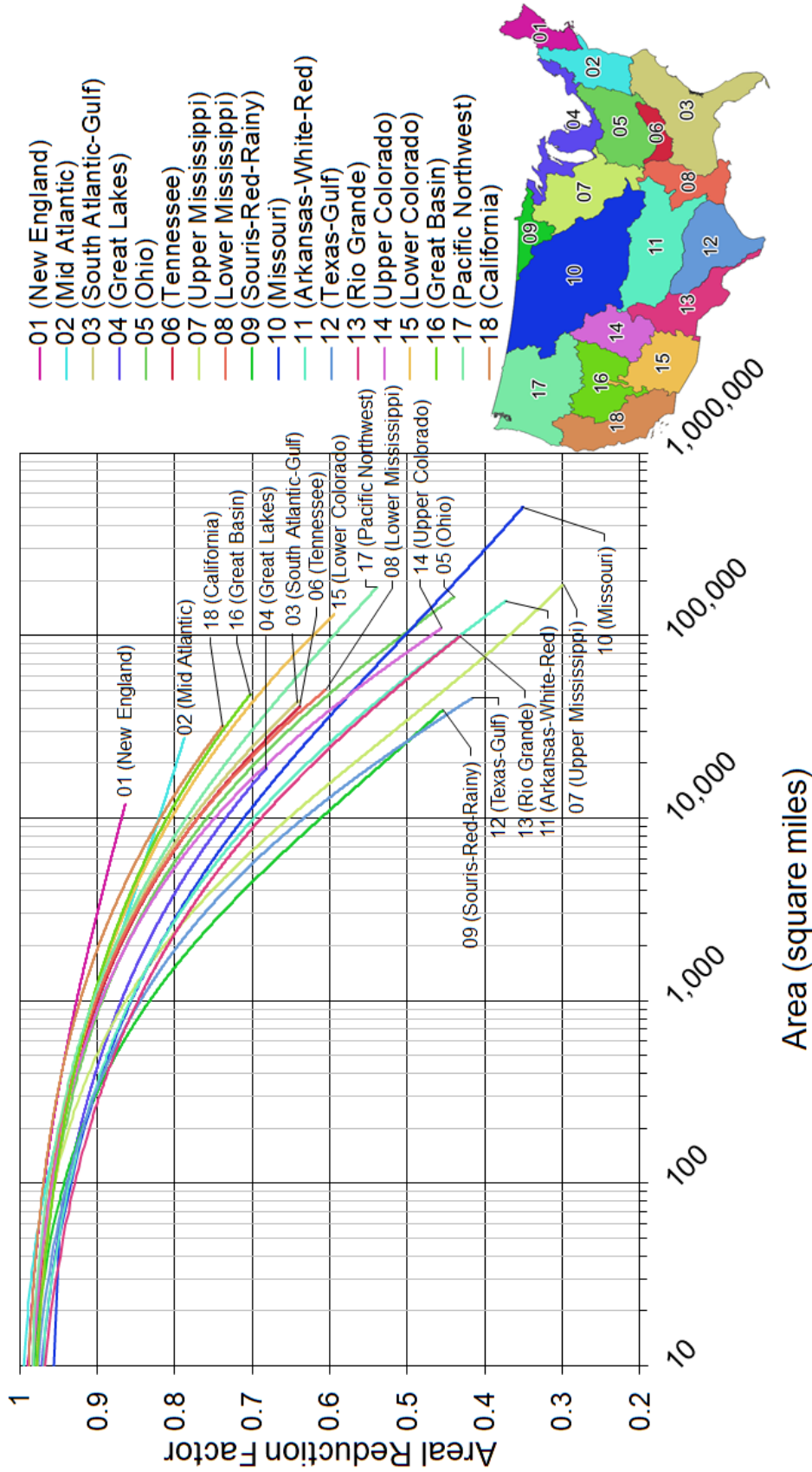


Figure F-9 CONUS ARF Assessment Plot for 2-day Duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting

Areal Reduction Factors by HUC02 using PRISM-daily data and M5 fitting 3-day Duration | 100-y Return Period

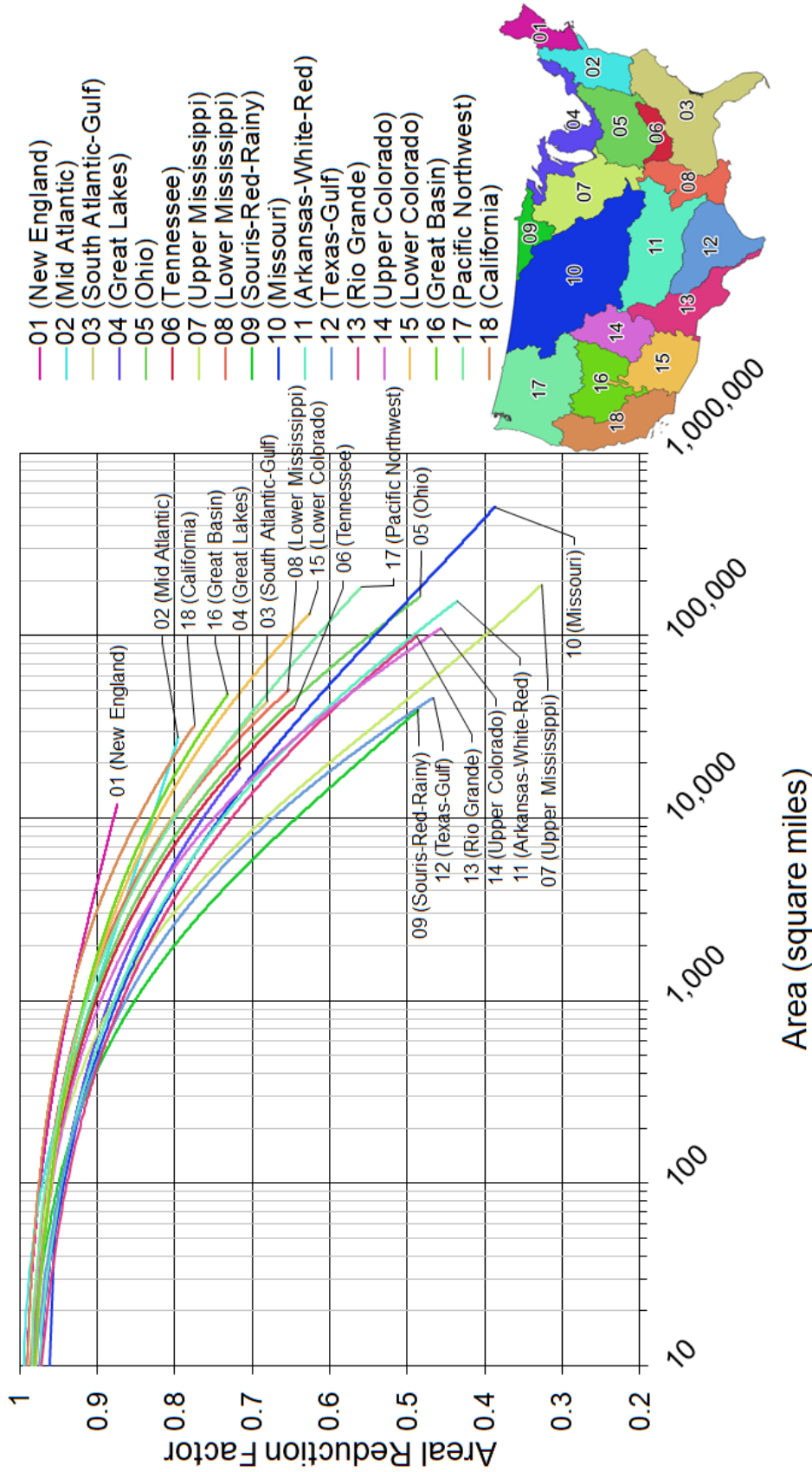


Figure F-10 CONUS ARF Assessment Plot for 3-day duration and 100-year Return Period Using PRISM-daily Data and M5 Model Fitting

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

NUREG/CR-7271

2. TITLE AND SUBTITLE

Application of Point Precipitation Frequency Estimates to Watersheds

3. DATE REPORT PUBLISHED

MONTH	YEAR
February	2021

4. FIN OR GRANT NUMBER

5. AUTHOR(S)

Shih-Chieh Kao*
Scott T. DeNeale*

6. TYPE OF REPORT

Technical

7. PERIOD COVERED (Inclusive Dates)

8. PERFORMING ORGANIZATION - NAME AND ADDRESS (If NRC, provide Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Oak Ridge National Laboratory*
P.O. Box 2008
Oak Ridge, TN 37831

9. SPONSORING ORGANIZATION - NAME AND ADDRESS (If NRC, type "Same as above", if contractor, provide NRC Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address.)

Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555-0001

10. SUPPLEMENTARY NOTES

11. ABSTRACT (200 words or less)

This report documents work sponsored by the U.S. Nuclear Regulatory Commission (NRC) at the Oak Ridge National Laboratory (ORNL) as part of the RES project, "Application of Point Precipitation Frequency Estimates to Watersheds." This project was implemented as part of the Probabilistic Flood Hazard Assessment (PFHA) Research Program. The objective of the PFHA Research Program is to develop tools and guidance on the use of PFHA methods to risk-inform NRC's licensing of new facilities as well as licensing and oversight of currently operating facilities as they relate to flooding hazards. Many nuclear power plants (NPPs) are located on or near rivers so riverine flooding hazards need to be considered in their design and operation. Probabilistic riverine flood models are important tools for realistic assessment of flooding risks. However, these models require areal estimates of the depth, duration, and frequency of rainfall distributed over the watershed, which are not often available. The research documented in this report addresses areal reduction factors (ARFs), which can be used to convert the widely available point precipitation frequency estimates, to estimates of areal precipitation frequency over a watershed.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

ARF, areal reduction factor, point precipitation frequency, hydrology, PFHA

13. AVAILABILITY STATEMENT

unlimited

14. SECURITY CLASSIFICATION

(This Page)

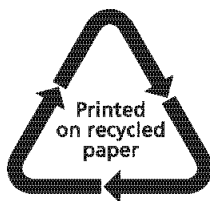
unclassified

(This Report)

unclassified

15. NUMBER OF PAGES

16. PRICE



Federal Recycling Program



UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, DC 20555-0001

OFFICIAL BUSINESS



@NRCgov

NUREG/CR-7271

Application of Point Precipitation Frequency Estimates to Watersheds

February 2021