DATA ANALYTICS APPLIED TO SATELLITE-DERIVED PRECIPITATION ESTIMATES AND HIGH-RESOLUTION MODEL OUTPUT

by

Taylor Alexandria Gowan

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Atmospheric Sciences

The University of Utah

May 2021

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Taylor Alexandria Gowan**

has been approved by the following supervisory committee members:

John Horel	, Chair	**2/16/2021** Date Approved
Erik Crosman	, Member	**2/16/2021** Date Approved
George Huffman	, Member	**2/16/2021** Date Approved
William J. Steenburgh	, Member	**2/16/2021** Date Approved
Courtenay Strong	, Member	**2/16/2021** Date Approved

and by **John Horel** , Chair/Dean of

the Department/College/School of **Atmospheric Sciences**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The global weather enterprise is composed of entities that specialize in a range of areas from research and operations to industry, each with a unique set of goals and challenges that impede progress. Typically, an impediment to progress in the field of atmospheric sciences lies in the lack of access to, or availability of a resource in the form of an observational dataset, model output, or computational power. This work presents a solution to two independent operational needs: 1) A reliable and representative gridded precipitation estimate dataset for fire weather applications in Alaska, and 2) A flexible, cloud-compatible file format for high volume model output for efficient operational and machine learning workflows.

The Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement Mission (IMERG) version 6B precipitation estimates are evaluated and bias-corrected for use in forecasting and situational awareness during the Alaskan fire season. Precipitation estimates from six fire seasons (1 June – 31 August 2014-2019) are aggregated to calculate empirical cumulative distributions for the four quadrants of Alaska. These distributions are utilized to evaluate the accuracy of the near real-time IMERG estimates of 24-hour precipitation and subsequently bias-correct them via regional quantile mapping. The regional quantile mapping method reduced the algorithm's wet bias, which improved the accuracy of the precipitation estimates across Alaska.

To address the second objective, an alternative file format, Zarr, is presented for the existing archive of High-Resolution Rapid Refresh (HRRR) model output. Zarr is a cloud-compatible, flexible file format which allows for N-dimensional data arrays that are chunked and compressed based on user specifications. The resulting dataset, HRRR-Zarr, has an analysis (F00) and forecast (F01-F18 or F48) file for each model run, whose arrays are subset into 96 three-dimensional chunks, with spatial dimensions of size 150 x 150 grid spaces and a time dimension determined by forecast length. This work demonstrates the utility of HRRR output in Zarr format for operational and machine learning workflows. The HRRR-Zarr archive and near real-time data are stored as objects in the Amazon Web Service Simple Storage Service and made publicly accessible via the Amazon Open Data Registry.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

A PhD is often thought of as an individual accomplishment, but that could not be further from the truth. This work and accomplishment would not be possible without the help, patience, generosity, and encouragement of many people and organizations. First and foremost, I would like to thank my advisor, John Horel. The day we met at the AMS student conference in 2015, I knew John was someone I could learn from and work with. Over the past six years, John has provided me with every opportunity possible to grow as a scientist and human being. Thank you, John, for the time you invested, knowledge you imparted on me, and laughs we shared looking at news articles from rural Alaska.

I want to thank my husband and new PhD, Tom Gowan, who still wanted to marry me after taking many graduate classes together and trouble-shooting my code more times than I want to admit. You are brilliant and I couldn't ask for a better life partner. Thank you to my wonderful, selfless parents, for your endless encouragement and love, and for always having the Weather Channel on when I was growing up. I would not be where I am without you both. I also want to thank my sister, Tara, for always making me smile, even on my worst days. To my friends and colleagues, thank you for the impact you have made on my life and for helping mold me into the person I am today.

I am very grateful for each of my committee members, Jim, George, Court, and Erik, for the time spent and effort imparted on helping me complete my degree. I want to thank Jason Knievel for hosting me as a visiting scientist at NCAR. I so appreciate the

modeling knowledge you gave me, and for all the European travel tips over lunch. A special thank you to Robert Ziel, who was my personal Alaska tour guide and expert on all things fire weather. I also want to thank Trevor Alcott, for fielding an unknown number of HRRR-related questions and always doing so graciously. I am also grateful to my fellow OU alumnus, Zac Flamig, and the Amazon Open Data Program, for providing the inspiration and support for creating the HRRR-Zarr archive as part of the Amazon Sustainability Data Initiative.

Last but not least, I am so thankful for my smiley adventure pup, Balto, whose name was inspired by my Alaska-focused research. On paper, we rescued him, but I am certain he rescued me.

This work would not be possible without a grant from the Joint Fire Science Program, (JFSP Grant L17AC00225) and a Collaborative Science, Technology, and Applied Research (CSTAR) Program Awards (55500146 and NA20NWS4680046). Creation of the HRRR-Zarr archive would not have been possible without an Amazon Sustainability Data Initiative Promotional Credits Award. The University of Utah Center for High Performance Computing (CHPC) provided computational hardware and software for this work.

CHAPTER 1

INTRODUCTION

1.1 Motivation

Each day, forecasting and decision-making agencies around the world utilize a range of data products to generate invaluable guidance on the state of the atmosphere and how it will affect the economy and society (Benjamin et al. 2018). Subsections of the weather enterprise specialize in area-specific high-impact phenomena such as hydrologic monitoring or fire weather forecasting, and thus use different resources on a daily basis whether that be numerical weather model output, satellite retrievals, or observations from ground stations. Regardless of location, specialization, or purpose, every agency that provides operational guidance faces a unique set of challenges that hinder efforts to either produce a forecast or make critical decisions. Here, we address two areas of research motivated by current operational needs: 1) Lack of availability and accuracy of precipitation estimates during Alaska fire seasons and 2) Difficulties accessing and efficiently processing numerical weather prediction model output for machine learning and forecasting applications.

1.2 IMERG Precipitation Estimates

During the summer months, the state of Alaska is prone to destructive wildfire outbreaks. The Alaska Interagency Coordination Center (AICC) is the agency tasked with forecasting weather conducive to igniting new or exacerbating existing wildfires as well as allocating firefighting resources and managing risks. The AICC determines wildfire potential using the Canadian Forest Fire Danger Rating System (CFFDRS) which relies on sensible weather parameters (temperature, relative humidity, wind, precipitation) to quantify fuel conditions and flammability (De Groot et al. 2003; Lawson and Armitage 2008; Flanigan et al. 2013; Horel et al. 2014; Field et al. 2015). The CFFDRS indices are calculated each day using observations and model analyses and made available online in gridded format as part of a program called Alaska Fire and Fuels (Fig. 1.1). These tools help fire managers and agencies like AICC communicate risk and wildfire potential to the public.

One of the major shortcomings of using the CFFDRS in Alaska is the lack of in-situ observations, especially in remote areas where wildfires are frequent. Spatially heterogenous weather parameters, like precipitation, are especially difficult to quantify during the summer months in Alaska given its predominately convective mode. Daily precipitation totals in Alaska are on average relatively small ($\leq$ 3 mm), but have a profound effect on the daily CFFDRS index calculations. The current CFFFDRS precipitation input, gridded quantitative precipitation estimates produced by the Alaska-Pacific River Forecast Center, relies heavily on observations from the sporadic National Weather Service (NWS) and RAWS stations. If one of these stations produces bad data or needs maintenance, the CFFDRS indices will not reflect the true state of the fuels

based on the meteorological information it received.

Having accurate precipitation estimates in time and space are critical for reliable representation of fuel conditions, but are not well measured by Alaska's few available ground stations. We propose an alternative precipitation source: Estimates derived using the Integrated Multi-SatellitE Retrievals for the Global Precipitation Measurement Mission (IMERG) algorithm. IMERG precipitation estimates are derived using infrared and passive microwave radiances from satellites within the Global Precipitation Measurement Mission (GPM) constellation, and are available at a 0.1° x 0.1° resolution for all latitudes (Hou et. al 2014; Fig. 1.2). The IMERG algorithm combines retrievals from all satellites in the constellation to produce a global precipitation estimate (mm hr^{-1}) every half hour (GPM ATBD 2019). For each half hour, the IMERG algorithm is run three time, with varying latencies and degree of post-processing. For the work in this dissertation, we focus on the IMERG-Early (IMERG-E) run which minimizes latency (~4h), but has less post-processing than the Late and Final runs with ~12h and ~3 months latency, respectively.

Chapter 2 of this dissertation evaluates the utility of near real-time IMERG precipitation estimates for use in operational fire weather applications in Alaska. We perform a validation of the IMERG-E 24-hour accumulated precipitation estimates across the state of Alaska, using precipitation observations from stations managed by the NWS and Remote Automated Weather Stations (RAWS) as ground truth. Since station observations are sporadic and generally unrepresentative on a large spatial scale, we split the state into 4 quadrants, and aggregate daily precipitation values over six fire seasons (June 1-August 31 2014-2019) into a cumulative distribution function. We then compare

24-hour precipitation accumulation values based on percentile ranges rather than explicit values. Following the evaluation, we build upon a commonly used bias correction method, quantile mapping, to attempt to improve the IMERG-E estimates (Erickson et al. 2012; Pierce et al. 2015; Nelson et al. 2016; Hashemi et al. 2017). We employ a regional quantile mapping approach to account for the sporadic and sometimes unreliable station data located in each quadrant. After the bias-correction, we replicate the steps from the initial validation to evaluate the quantile-mapped IMERG-E estimates.

Throughout Chapter 2 of this dissertation, we will address the following research objectives:

- How well does the IMERG-E algorithm estimate 24-hour precipitation as a function of region or fire season?
- What kind of bias (if any) does IMERG-E exhibit at each percentile range?
- What are the shortcomings of the algorithm, especially at high latitude locations?
- Does the quantile mapping bias correction method improve the IMERG-E precipitation estimates?
- Are the corrected precipitation estimates satisfactory for use in CFFDRS calculations?

1.3 The HRRR-Zarr Archive

The High-Resolution Rapid Refresh (HRRR; Benjamin et al. 2016) is a 3-km, convection-allowing numerical weather prediction model run operationally every hour by

the National Centers for Environmental Prediction's Environmental Modeling Center. Depending on the initialization time, the current HRRR model, version 4, produces hourly surface, pressure, and native grid forecasts out to 18 or 48 hours. Its domain covers the contiguous United States (CONUS) and has a separate domain for Alaska, both of which are nested within the Rapid Refresh model domain which supplies the HRRR with initial and boundary conditions (Fig. 1.3; McCorkle et al. 2018).

For decades, two-dimensional output from many numerical weather prediction models, including the HRRR model, have been stored in the highly-compressible hypercube-structured file format, GRIdded Binary Second Edition (GRIB2; Silver and Zander 2017). For a given HRRR run, 19 or 49 (F00-F18, F00-F48) GRIB2 files are produced that contain hundreds of two-dimensional fields with values for 1.9 million grid points, most of which are not needed by an end user. While GRIB2 files are effective for storing and transmitting large amounts of two-dimensional model data, they are difficult to work with given their large size when decompressed. Machine learning and other high-throughput applications might require accessing thousands of GRIB2 files, which requires an enormous amount of storage and computing power.

In Chapter 3, we propose a flexible and cloud-compatible N-dimensional file format, Zarr, as an alternative HRRR model output format intended for use in operational and machine learning workflows. Despite being a relatively new file type, the Zarr format has already been adopted for big data applications, e.g., the United Kingdom's Met Office model reanalysis datasets (McCaie 2019), Lyft's Level 5 self-driving platform (Houston et al. 2020), and for the datasets in the MalariaGEN project (Pearson et al. 2019). Files in Zarr format can easily be generated or manipulated in Python using the

Zarr Python library. The Zarr protocol is backend compatible with common libraries (e.g., xarray, iris, and dask) and has the capability to write directly to cloud platforms.

In this dissertation, we apply the Zarr format to existing HRRR surface files originally archived by the University of Utah (Blaylock et al. 2017; Blaylock et al. 2018). The resulting dataset, HRRR-Zarr, has two sets of zarr files per model run: analysis files derived from each field in the F00 GRIB2 output files, and forecast files combining all available forecast times (extending from F01 up to F48) for each field. Each GRIB2 field of size 1799 x 1059 is subdivided into 96 chunks of size 150 x 150, spanning ~450 km^2. Hence, the zarr analysis files are small squares while the forecast files are small cubes with forecast time being the third dimension (Fig. 1.4). All zarr files are compressed using the LZ4 algorithm and mimic a traditional directory structure using file names that end users are accustomed to from the field naming conventions in GRIB2 files. The archive and near-real time HRRR-Zarr files are stored as objects in the Amazon Web Service (AWS) Simple Storage Service (S3) as part of the AWS Open Data Program (Amazon 2021). These objects are located in an S3 bucket, accessible with the unique identifier “hrrrzarr”.

In Chapter 3, we discuss the current state of managing the data from numerical weather prediction model output, introduce the Zarr format, describe the workflow for creating the HRRR-Zarr dataset, and highlight potential HRRR-Zarr use cases for high stakes forecasting and machine learning applications. While some are open-ended and not completely resolved with our Zarr method, we address the following operational needs and objectives:

- How can the growing volume of model data be made more accessible to end users?
- How can end users utilize model output in a flexible manner, without having to store many GB of unneeded data or perform additional post-processing?
- What compute resources are needed and how can we convert TBs of GRIB2 HRRR output to Zarr format?
- What can be done in our HRRR-Zarr files to create a familiar environment for users of GRIB2 while being accessible to users in other fields?
- What chunk structure and compression algorithm would allow for lossless compression and promote more efficient operational and machine learning workflows?
- In what operational or research situations would the Zarr format be optimal?

The questions and objectives stated here will be addressed in Chapter 3. The final chapter of this dissertation discusses the potential applications of the HRRR-Zarr format as well as additional datasets that if in Zarr format would benefit a spectrum of users.

1.4 Research Overview

The research in this dissertation is based on the root objective of identifying gaps in resources that impede operational forecasting or the research that advances it, and posit an implementable solution. The research in Chapter 2 was spearheaded by a grant from the Joint Fire Science Consortium to study weather conditions conducive to rapid spread

of wildfires and the ability of the HRRR model to forecast them. Our previous research evaluated the HRRR-Alaska model forecasts of surface pressure (McCorkle et al. 2018), which later led to uncovering other operational challenges in the state, namely the accuracy of precipitation estimates during fire season. This work evaluated the IMERG algorithm's precipitation estimates over Alaska to determine its utility for fire weather applications. The research described in Chapter 2 was published in *Weather and Forecasting* as the first extensive validation of the IMERG precipitation estimates poleward of 60° N/S (Gowan and Horel 2020). The workflow to evaluate and bias correct the IMERG estimates was presented to researchers and fire managers at the Alaska Fall Fire Review and has the potential to be implemented operationally. In addition to addressing the utility of the algorithm for use in Alaska, the results from a high-latitude evaluation of the IMERG estimates could bring to light systematic biases that exist and could be accounted for in future algorithm iterations.

Chapter 3 presents a solution to the data bottlenecks experienced when using a high volume of model output in legacy file formats such as GRIB2 and netCDF4. We illustrate an alternate data storage method that involved transitioning from a several year effort to archive HRRR model output at the University of Utah (Blaylock et al., 2017) to providing that data in Zarr format. The HRRR-Zarr archive is now a publicly-available repository, made possible by the AWS Open Data Program and its partnership with the National Oceanic and Atmospheric Administration Big Data Program (Fig. 1.5). Chapter 3 details the workflow through which we developed the HRRR-Zarr dataset, a flexible, cloud-compatible file format that will enable efficient access of model data for machine learning and operations. This product has already garnered attention from researchers and

industry scientists alike, and has the potential to be extrapolated to output from other models or datasets. The research in Chapter 3 detailing creation of the HRRR-Zarr archive and potential use cases will be submitted for publication consideration in the *Journal of Atmospheric and Oceanic Technology*.

The dissertation concludes with Chapter 4, which summarizes the findings of the research in Chapters 2 and 3. Chapter 4 will propose extensions of the existing research and briefly discuss potential applications of these datasets in the future.

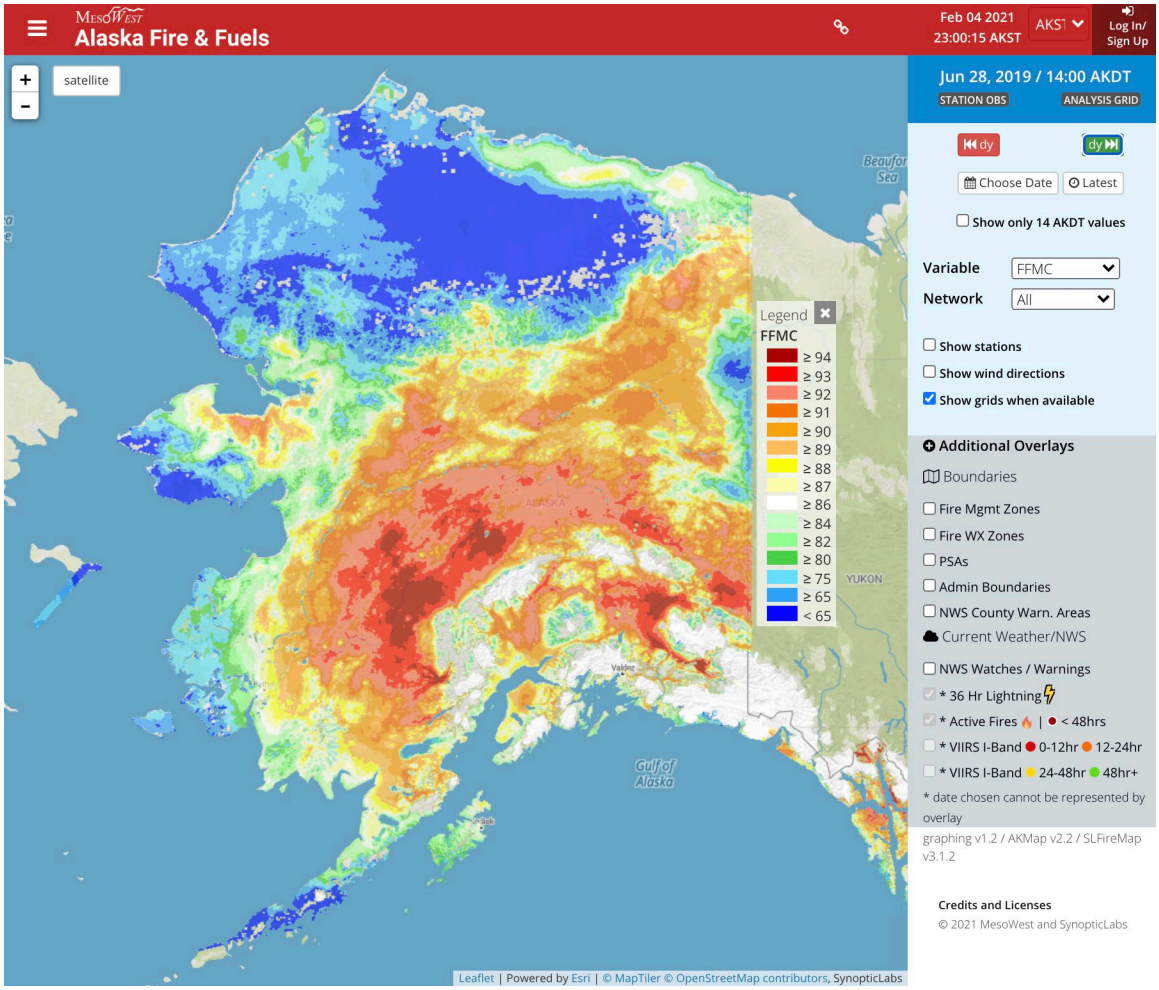

Figure 1.1: Alaska Fire and Fuels web interface displaying a CFFDRS index, Fine Fuel Moisture Code, on 28 June 2019 (https://akff.mesowest.org/).

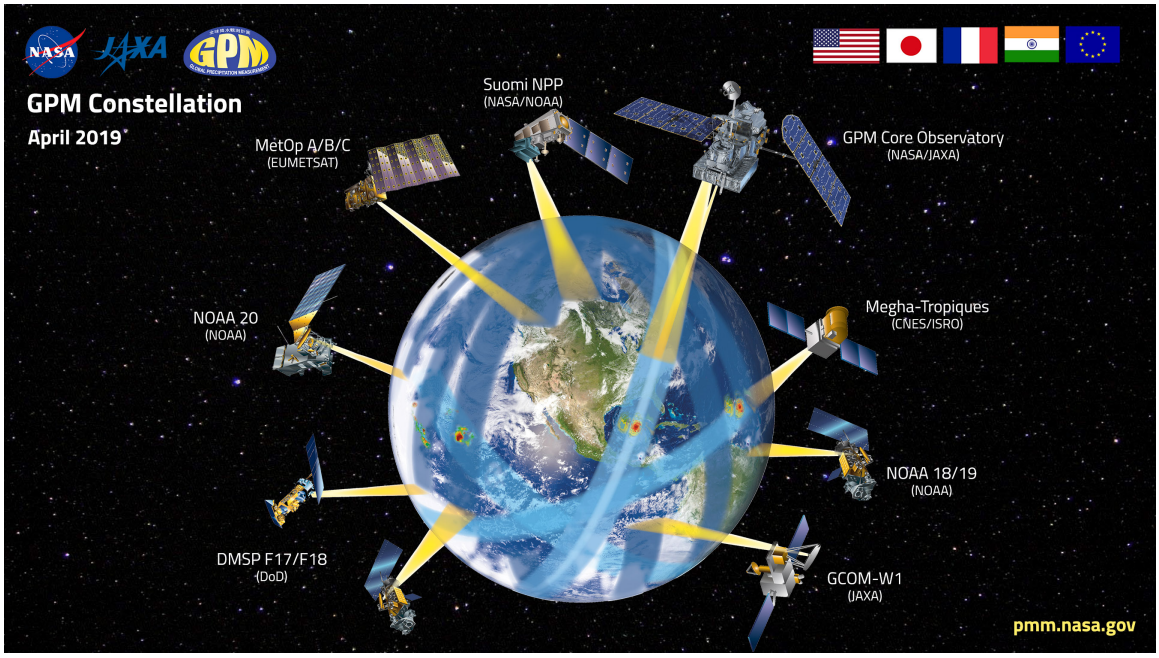

Figure 1.2: Satellites in orbit, as of April 2019, that contribute data to the Global Precipitation Measurement Mission (https://gpm.nasa.gov/).

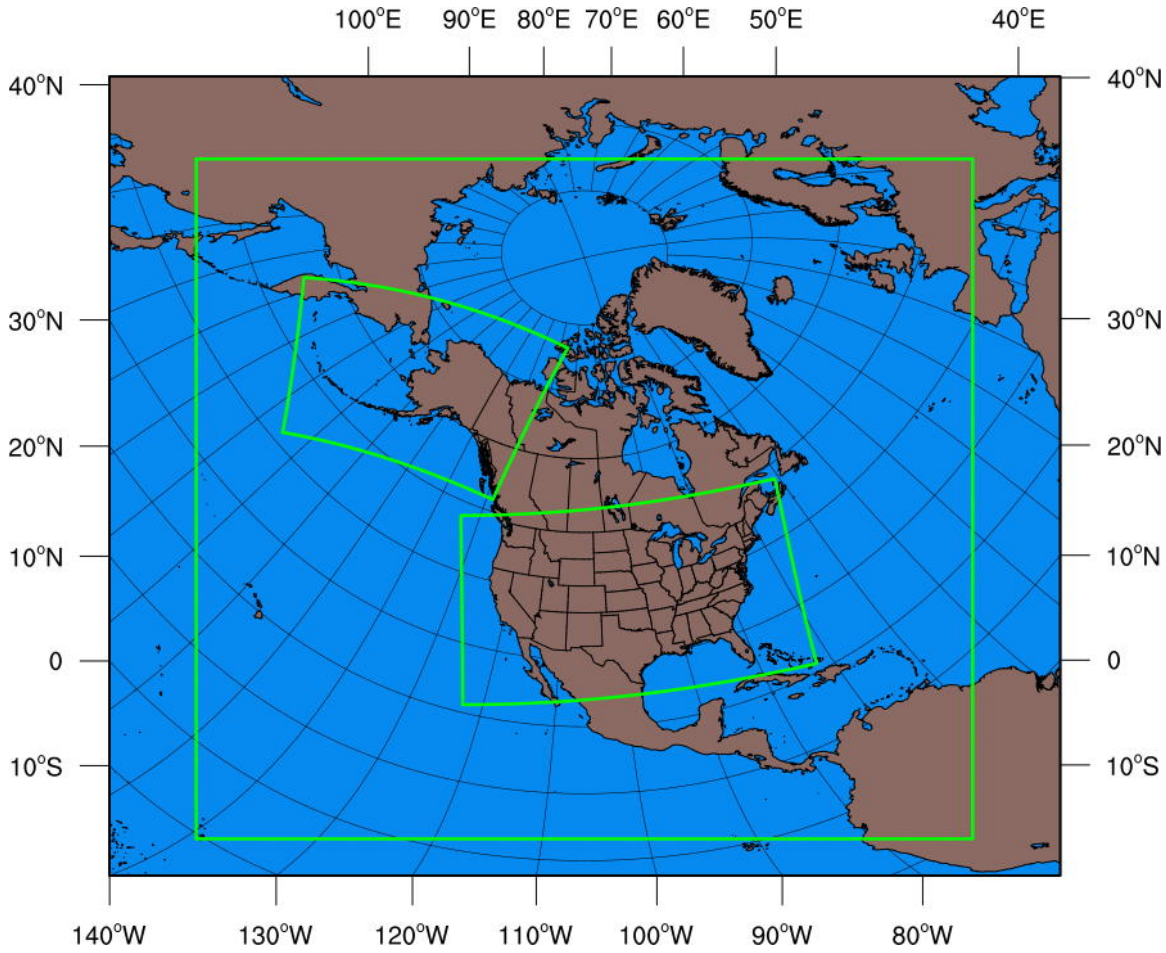

Figure 1.3: Domain boundaries for the HRRR model domains over the contiguous United States and Alaska. The HRRR domains are nested within the Rapid Refresh model domain (https://rapidrefresh.noaa.gov/).

Figure 1.4: HRRR-Zarr data chunk boundaries over the western United States. Map is filled to display terrain and land cover.

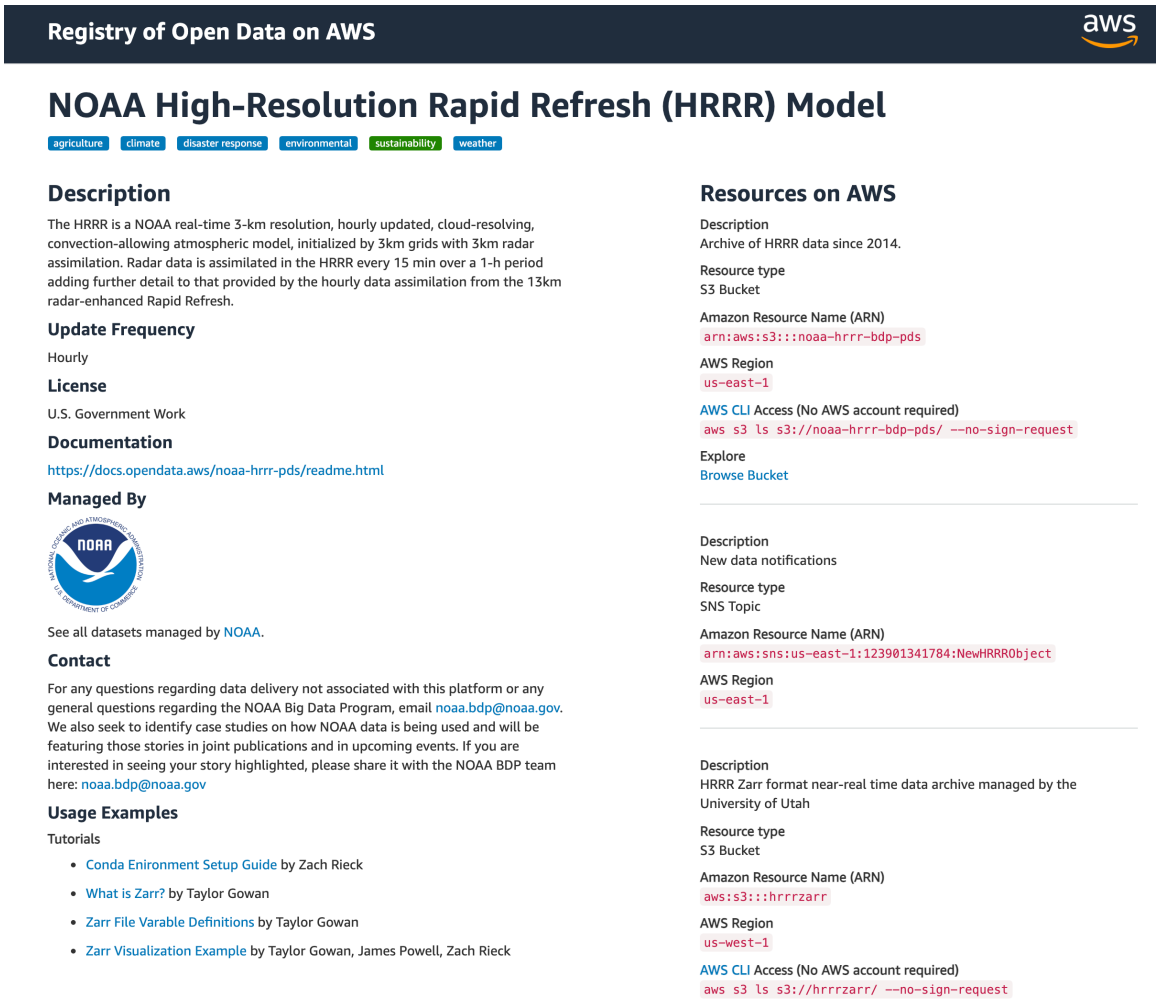

Figure 1.5: Amazon Open Data Registry landing page for the HRRR GRIB2 and Zarr repositories (https://registry.opendata.aws/noaa-hrrr-pds/)

1.5 References

Amazon, 2021: Amazon S3: Object storage built to store and retrieve any amount of data from anywhere. Accessed 2 January 2021, https://aws.amazon.com/s3/.

Benjamin, S. G., J. M. Brown, G. Brunet, P. Lynch, K. Saito, and T. W. Schlatter, 2018: *100 Years of Progress in Forecasting and NWP Applications*, *Meteor. Monogr.*, **59**, 13.1–13.67, doi:10.1175/AMSMONOGRAPHS-D-18-0020.1.

—, S. S. Weygandt, J. M. Brown, M. Hu, C. R. Alexander, T. G. Smirnova, J. B. Olson, E. P. James, D. C. Dowell, G. A. Grell, H. Lin, S. E. Peckham, T. L. Smith, W. R. Moninger, J. S. Kenyon, and G. S. Manikin, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh, 2016. *Mon. Wea. Rev.*, **144**, 1669–1694, doi:10.1175/MWR-D-15-0242.1.

Blaylock, B. K., J. Horel, E. Crosman, 2017: Impact of Lake Breezes on Summer Ozone Concentrations in the Salt Lake Valley. *J. Appl. Meteor. Clim.*,**56**, 353-370, doi:10.1175/JAMC-D-16-0216.1.

—, —, and S. T. Liston, 2017: Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output. *Computers and Geosciences*, **109**, 43–50, doi: 10.1016/j.cageo.2017.08.005.

—, —, and C. Galli, 2018: High-Resolution Rapid Refresh Model Data Analytics Derived on the Open Science Grid to Assist Wildland Fire Weather Assessment. *J. Atmos. Oceanic Technol.*, **35**, 2213–2227, doi:10.1175/JTECH-D-18-0073.1.

De Groot W. J., P. M. Bothwell, D. H. Carlsson, and K. A. Logan, 2003: Simulating the effects of future fire regimes on western Canadian boreal forests. *J. Veg. Sci.*,**14**, 355–364, doi:10.1111/j.1654-1103.2003.tb02161.1.

Erickson, M. J., B. A. Colle, and J. J. Charney, 2012: Impact of bias correction type and conditional training on Bayesian model averaging over the northeast United States. *Wea. Forecasting*, **27**, 1449–1469, doi: 10.1174/WAF-D-11-00149.1.

Field, R. D., and Coauthors, 2015: Development of a Global Fire Weather Database. *Nat. Hazards Earth Syst. Sci.*, **15**, 1407–1423, doi:10.5194/nhess-15-1407-2015.

Flannigan, M. D., A. S. Cantin, W. J. de Groot, M. Wotton, A. Newbery, and L. Gowman, 2013: Global wildland fire season severity in the 21st century. *Forest Ecol. Manage.*, **294**, 54–61, doi: 10.1016/j.foreco.2012.10.022.

Gowan, T. A., and J. Horel, 2020: Evaluation of IMERG-E Precipitation Estimates for Fire Weather Applications in Alaska. *Wea. Forecasting*, **35**, 1831–1843,

doi:10.1175/WAF-D-20-0023.1.

GPM ATBD, 2019: Global Precipitation Measurement algorithm theoretical basis documents. [Available online at: https://pps.gsfc.nasa.gov/atbd.html.]

Hashemi, H., M. Nordin, V. Lakshmi, G. J. Huffman, and R. Knight, 2017: Bias Correction of Long-Term Satellite Monthly Precipitation Product (TRMM 3B43) over the Conterminous United States. *J. Hydrometeor.*, **18**, 2491–2509, doi: 10.1175/JHM-D-17-0025.1.

Horel, J. D., R. Ziel, C. Galli, J. Pechmann, and X. Dong, 2014: An evaluation of fire danger and behaviour indices in the Great Lakes Region calculated from station and gridded weather information. *Int. J. Wildland Fire*, **23**, 202–214, doi: 10.1071/WF12186.

Hou, A. Y., R. K. Kakar, S. Neeck, A. A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, doi:10.1175/BAMS-D-13-00164.1.

Houston, J., G. Zuidhof, L. Bergamini, Y. Ye, A. Jain, S. Omari, V. Iglovikov, and P. Ondruska, 2020: One Thousand and One Hours: Self Driving Motion Prediction Dataset. Lyft, accessed 9 December 2020, https://level5.lyft.com/dataset/.

Lawson, B. D., and O. B. Armitage, 2008: Weather Guide for the Canadian Forest Fire Danger Rating System. Tech. Rep. Natural Resources Canada, Canadian Forest Service, 87 pp, ISBN: 978-1-100-11565-8.

McCaie, T., 2019: Creating a data format for high momentum datasets. Medium, accessed 20 December 2020, https://medium.com/informatics-lab/creating-a-data-format-for-high-momentum-datasets-a394fa48b671.

McCorkle, T. A., J. D. Horel, A. A. Jacques, and T. Alcott, 2018: Evaluating the Experimental High-Resolution Rapid Refresh-Alaska Modeling System Using USArray Pressure Observations. *Wea. Forecasting*, **33**, 933–953, doi: 10.1175/WAF-D-17-0155.1.

Nelson, B. R., O. P. Prat, D. Seo, and E. Habib, 2016: Assessment and Implications of NCEP Stage IV Quantitative Precipitation Estimates for Product Intercomparisons. *Wea. Forecasting*, **31**, 371–394, doi: 10.1175/WAF-D-14-00112.1.

Pearson, R. D., R. Amato, and D. P. Kwiatkowski, and MalariaGEN Plasmodium Falciparum Community Project, 2019: An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples, *BioRxiv*, doi: 10.1101/824730.

Pierce, D. W., D. R. Cayan, E. P. Maurer, J. T. Abatzoglou, and K. C. Hegewisch, 2015: Improved Bias Correction Techniques for Hydrological Simulations of Climate Change. *J. Hydrometeor.*, **16**, 2421–2442, doi: 10.1175/JHM-D-14-0236.1.

Silver, J. and C. Zender, 2017: The compression-error trade-off for large gridded data sets. *Geosci. Model Dev.*, **10**, 413–423, doi:10.5194/gmd-10-413-2017.

CHAPTER 2

EVALUATION OF IMERG-E PRECIPITATION ESTIMATES FOR FIRE WEATHER APPLICATIONS IN ALASKA

Reprinted with permission from Gowan, T. A. and J. D. Horel, 2020: Evaluation of IMERG-E Precipitation Estimates for Fire Weather Applications in Alaska. *Wea. Forecasting*, **35**, 1831–1843, doi: 10.1175/WAF-D-20-0023.1.

Evaluation of IMERG-E Precipitation Estimates for Fire Weather Applications in Alaska

TAYLOR A. GOWAN AND JOHN D. HOREL

Department of Atmospheric Sciences, University of Utah, Salt Lake City, Utah

(Manuscript received 6 February 2020, in final form 22 May 2020)

ABSTRACT

Large wildfire outbreaks in Alaska are common from June to August. The Canadian Forest Fire Danger Rating System (CFFDRS) is used operationally by Alaskan fire managers to produce statewide fire weather outlooks and forecast guidance near active wildfires. The CFFDRS estimates of fire potential and behavior rely heavily on meteorological observations (precipitation, temperature, wind speed, and relative humidity) from the relatively small number of in situ stations across Alaska with precipitation being the most critical parameter. To improve the spatial coverage of precipitation estimates across Alaska for fire weather applications, a multisatellite precipitation algorithm was evaluated during six fire seasons (1 June–31 August 2014–19). Near-real-time daily precipitation estimates from the Integrated Multisatellite Retrievals for the Global Precipitation Mission (IMERG) algorithm were verified using 322 in situ stations across four Alaskan regions. For each region, empirical cumulative distributions of daily precipitation were obtained from station observations during each summer, and compared to corresponding distributions of interpolated values from IMERG grid points ($0.1° \times 0.1°$ grid). The cumulative distributions obtained from IMERG exhibited wet biases relative to the observed distributions for all regions, precipitation amount ranges, and summers. A bias correction approach using regional quantile mapping was developed to mitigate for the IMERG wet bias. The bias-adjusted IMERG daily precipitation estimates were then evaluated and found to produce improved gridded IMERG precipitation estimates. This approach may help to improve situational awareness of wildfire potential across Alaska and be appropriate for other high-latitude regions where there are sufficient in situ precipitation observations to help correct the IMERG precipitation estimates.

1. Introduction

Wildfire outbreaks are common in Alaska's dense boreal forests during the summer months when daylight can persist for more than 20 h each day, surface heating is maximized, and convection is frequent (Shulski and Wendler 2007; Bieniek et al. 2014; Melvin et al. 2017; Pickell et al. 2017). In recent decades, wildfire outbreaks have become more numerous, especially in Alaska's interior. In total, 5 of the 10 largest fire seasons in terms of acres burned have occurred after 2000. The flammability of fuels, number of ignitions, and total area burned are largely dependent on natural, rather than anthropogenic, factors (e.g., frequency of convection and lightning and subseasonal to seasonal precipitation and temperature patterns) (Ziel et al. 2015; Pickell et al. 2017). Thus, it is critical to have accurate atmospheric data for operational products that are used for situational awareness before and during wildfire outbreaks.

The Alaska Interagency Coordination Center (AICC) is tasked with managing the resources for suppressing wildfires. The AICC staff evaluate fuel conditions and wildfire potential in part on the basis of the Canadian Forest Fire Danger Rating System (CFFDRS; Lawson and Armitage 2008). The CFFDRS depends on the assumed initial fuel state at the beginning of the fire season and continuous daily inputs, many of which are meteorological observations (rainfall, temperature, wind, and moisture), thereafter to calculate indices of fuel conditions and flammability (De Groot et al. 2003; Flannigan et al. 2013; Horel et al. 2014; Field et al. 2015). The CFFDRS indices are calculated each day and made available publicly as part of the Alaska Fire and Fuels program that is supported by the AICC (see http://akff.mesowest.org). Meteorological inputs are obtained at in situ sites from selected observational networks and

Corresponding author: Taylor A. Gowan, taylor.mccorkle@utah.edu

DOI: 10.1175/WAF-D-20-0023.1

in gridded form at 2.5-km horizontal resolution for precipitation from the Alaska–Pacific River Forecast Center (APRFC) and for temperature, wind, and relative humidity from Real Time Mesoscale Analyses (De Pondeca et al. 2011). The APRFC quantitative precipitation estimate (QPE) grids are derived by adjusting the precipitation at the widely scattered observing sites by monthly gridded climatological fields and terrain–precipitation relationships. Given the inherent range of temporal and spatial scales of summer rainfall in Alaska, the in situ data available across Alaska are likely unrepresentative of the rainfall variability within the state. The resulting potential inaccuracies in the daily precipitation estimates require fire behavior analysts to spend additional time evaluating whether the calculated CFFDRS index values reflect the actual conditions.

One of the fundamental CFFDRS indices, Fine Fuel Moisture Code (FFMC), estimates the flammability of grasses and other fine fuels and is highly sensitive to the precipitation from the previous day. High values of FFMC correspond to dry fine fuels (Lawson and Armitage 2008). Figure 1 illustrates the impact of a midsummer storm that shifted from southwest Alaska on 21 June 2018 to central Alaska the next day. The high values of FFMC on 21 June 2018 in central Alaska (Fig. 1b) computed using the APRFC QPEs dropped substantially over the next 24 h (Fig. 1d). This example obtained from the operational Alaska Fire and Fuels grids illustrates that reliable fire danger indices require accurate estimates of the amount and spatial distribution of precipitation.

Few regions of the world have sufficient in situ precipitation observations or radar coverage to monitor the complex spatial and temporal variations in precipitation around the globe (Huffman et al. 2001, 2007; Kirstetter et al. 2014; Asong et al. 2017; Kidd et al. 2017; Skofronick-Jackson et al. 2017; Lundquist et al. 2019). Satellite-derived precipitation from the joint U.S.–Japan Global Precipitation Measurement (GPM) mission (Hou et al. 2014) is an approach to monitor precipitation globally, which is particularly relevant for the types of natural disasters that occur within the 1.718 million km^2 area of Alaska (Shulski and Wendler 2007; Bieniek et al. 2012, 2014; Lader et al. 2017). Precipitation datasets are now available every half hour from the Integrated Multisatellite Retrievals for the GPM (IMERG). Since the mission's launch in 2014, extensive evaluations of IMERG precipitation estimates have been conducted in the Netherlands (Gaona et al. 2016), southeast Austria (Sungmin et al. 2017), Africa (Dezfuli et al. 2017), China (Tang et al. 2016; Tian et al. 2018), the northern highlands of Pakistan (Anjum et al. 2018), the Peruvian–Ecuadorian Amazon basin (Zubieta et al. 2017), and in Iran's regions of complex terrain (Sharafi et al. 2016). Additionally, IMERG has been used as a validation dataset for verifying global forecasts of precipitation (Kotsuki et al. 2019) as well as modeled precipitation in extratropical cyclones (Naud et al. 2020). This study aims to be the first detailed evaluation of IMERG precipitation at latitudes poleward of 60°N/S.

The objective of this research is to evaluate and improve the utility of the Early (4-h latency) IMERG QPEs, which have the greatest potential for operational applications in Alaska. The 24-h (h) accumulations (0000–0000 UTC) of IMERG precipitation are examined across Alaska during the 2014–19 fire seasons (defined here as 1 June–31 August). In situ rainfall data are used to evaluate and bias-correct the IMERG QPEs and thereby create estimates of gridded daily precipitation in addition to that presently available from the APRFC. Our approach builds upon prior precipitation evaluation and bias correction studies that have used regional or quantile-based approaches (Erickson et al. 2012; Pierce et al. 2015; Nelson et al. 2016; Hashemi et al. 2017; Gowan et al. 2018). Our verification serves as a baseline for estimating the performance of IMERG in other high-latitude regions as well as its potential for fire weather applications at short lead times.

The remainder of this manuscript is structured as follows. Section 2 describes the IMERG algorithm and verification datasets while section 3 presents the methods used for evaluation and bias correction of the IMERG precipitation estimates. The quantitative verification results and the application of the bias correction techniques to the IMERG QPEs are presented in section 4 followed by a summary of the critical results and future work in section 5.

2. Precipitation datasets

a. IMERG algorithm

The international GPM mission, a successor to the U.S.–Japan Tropical Rainfall Measurement Mission (TRMM; Liu et al. 2012), is advancing the global understanding of precipitation processes and improving the forecasting of extreme events that cause natural disasters (Hou et al. 2014). The GPM mission data products provide 0.1° × 0.1° resolution precipitation estimates for all latitudes using the IMERG algorithm (GPM 2013; Hou et al. 2014; Skofronick-Jackson et al. 2017). The current version of IMERG, 06B (v6), includes precipitation estimates for all latitudes, which is

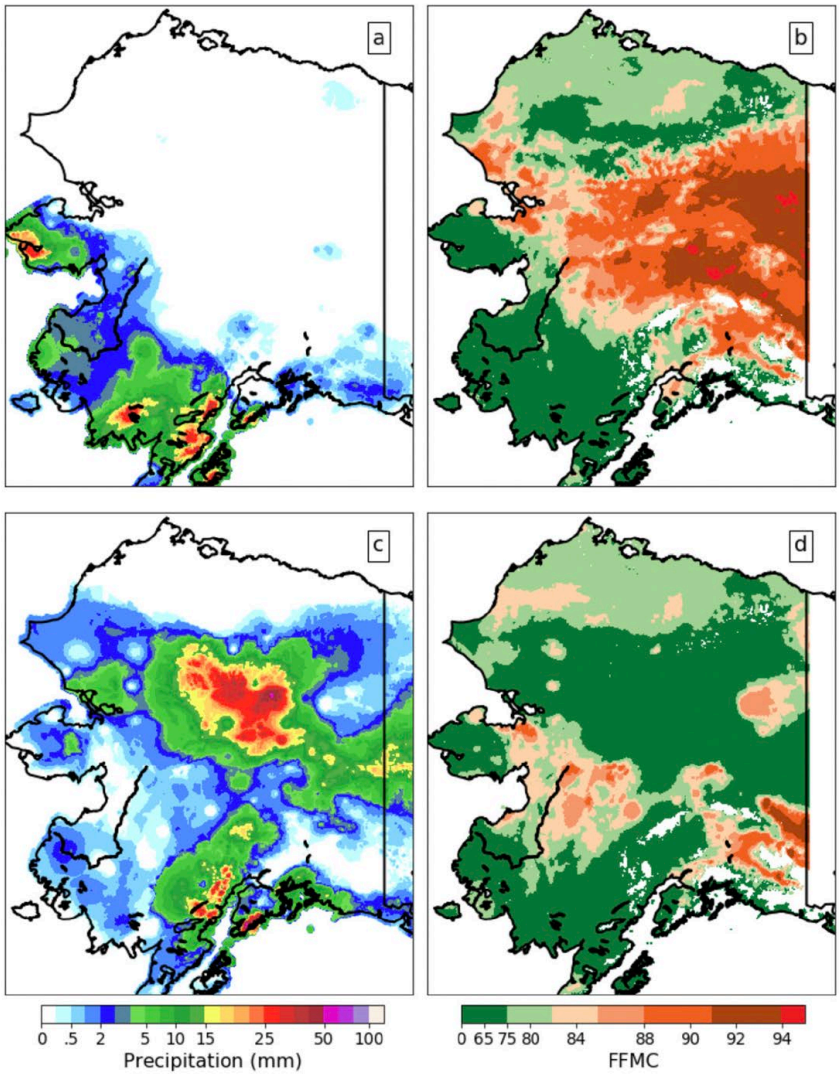

FIG. 1. (a) APRFC QPE (mm) valid for the 24-h period ending at 0000 UTC 22 Jun 2018, shaded according to the scale on the bottom left. (b) FFMC values valid at 2200 UTC 22 Jun 2018, shaded according to the scale on the bottom right. (c) As in (a), but valid for the 24 h period ending at 0000 UTC 23 Jun 2018. (d) As in (b), but valid at 2200 UTC 23 Jun 2018.

an improvement over the limited latitudinal domains of its predecessors, the TRMM Multisatellite Precipitation Analysis (40°S–40°N), and earlier versions of IMERG (60°S–60°N). With each new version of IMERG, the GPM and TRMM archives are backfilled by running the newest version of the algorithm retrospectively, allowing for consistent, comprehensive evaluation for each version.

The IMERG algorithm was developed to incorporate all microwave and infrared (IR) retrievals from sensors

on board satellites in the GPM constellation into a single global precipitation estimate (mm h^{-1}) every half hour (GPM ATBD 2019). For each time interval, the IMERG is run three times to produce independent datasets of estimated precipitation rate. These runs are referred to as Early (-E), Late (-L), and Final (-F) with latencies of 4 h, 14 h, and ~3.5 months, respectively. For this research, IMERG-E estimates of 24-h accumulated precipitation are evaluated.

For each run of the IMERG algorithm, precipitation estimates are computed from passive microwave brightness temperature retrievals using the Goddard Profiling Algorithm. This step is followed by gridding the data and performing a calibration based on the GPM Combined Instrument's sensors. Poleward of 60°, the Combined Instrument is calibrated to values from the Global Precipitation Climatology Project (GPCP; Huffman et al. 1997), due to a known low bias of GPM algorithms at high latitude (G. Huffman 2019, personal communication). Additionally, GPM IR-based estimates from the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS; Hong et al. 2004) are recalibrated based on trailing three-month matched estimates from the merged microwave field. According to the technical documents (GPM ATBD 2019), the PERSIANN-CCS IR estimates are sent along with the merged microwave field for morphing in the forward and backward direction via Lagrangian time interpolation and based on a sequence of morphing vectors. Unlike previous versions, IMERG v6 uses IR morphing vectors computed from the Modern-Era Retrospective Analysis for Research and Applications, version 2, allowing time interpolation to occur poleward of 60°N/S. Prior to this algorithm update, only passive microwave radiances over snow/ice-free surfaces were used to calculate IMERG precipitation estimates at high latitudes (GPM ATBD 2019). An important difference between IMERG runs lies in the postprocessing step, when the precipitation estimates are calibrated using in situ stations. The IMERG-E dataset is calibrated using monthly climate normals provided by stations in the GPCP, whereas the research-grade IMERG-F dataset is adjusted by monthly satellite-gauge accumulations, which explains in part the longer production latency.

b. *In situ observations*

Ground-based verification and bias correction of IMERG-E daily (0000–0000 UTC) precipitation estimates during the 2014–19 fire seasons were performed using daily precipitation accumulation observations from 155 stations managed by the NWS and 167 Remote Automated Weather Stations (RAWS). These networks complement one another for verification purposes as the NWS stations are situated near airports or in low-lying, populated areas (McCorkle et al. 2018) while the RAWS network is deployed for fire weather purposes typically in remote locations. These data were accessed and quality controlled by Synoptic Data, a Public Benefit Corporation (https://synopticdata.com), which provides access to environmental data from publicly available networks across all 50 states.

3. Assessment of IMERG-E quantitative precipitation estimates

a. *Quantitative evaluation methods*

The baseline accuracy of the IMERG-E dataset across Alaska is evaluated by comparing daily precipitation estimates during the 2014–19 fire seasons to in situ observations. To validate IMERG-E, we compared observations at each RAWS and NWS station to precipitation estimates at nearby IMERG-E grid points using bilinear interpolation. Our approach is intended to overcome common issues for precipitation verification in Alaska and elsewhere: 1) the limited number of in situ observing sites and 2) the wide range in precipitation amounts on daily time scales across the state. We first subset the in situ stations and IMERG grid values into quadrants demarcated by the latitude and longitude lines of 63°N and 152°W. These quadrant boundaries were selected to roughly balance among the regions the available numbers of stations from the RAWS and NWS networks (Fig. 2).

As shown in Fig. 3, in situ observations of 24-h precipitation were aggregated within each region during all of the 2014–19 fire seasons to obtain climatological cumulative distribution functions (CDFs). Across Alaska, no rainfall is measured roughly 60% of the time while rainfall amounts in excess of 10 mm day^{-1} occur less than 10% of the time (Fig. 3). These climatological CDFs helped define percentile regimes and region-specific accumulation thresholds for the IMERG-E verification. In a similar manner, the IMERG-E QPEs interpolated to the observation sites during all six fire seasons were combined and transformed into climatological CDFs for the four regions (not shown).

For each region, the distribution of daily IMERG-E QPEs during the six fire seasons are evaluated with the equivalent distribution from RAWS and NWS in situ observations as a function of time, region, and event magnitude. Several error metrics are used: frequency bias, bias ratio, mean absolute error (MAE), and root-mean-square error (RMSE).

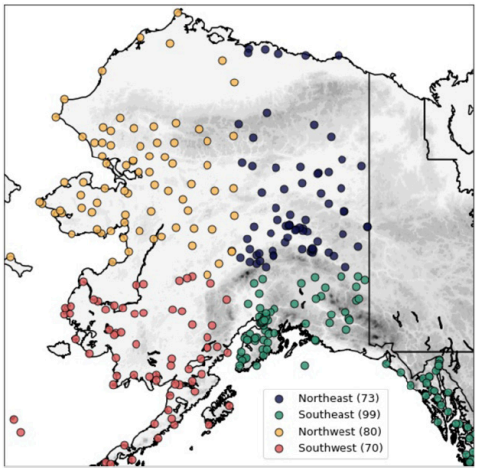

FIG. 2. In situ stations used for the IMERG-E evaluation. Station locations within the four regions are colored according to the legend with the total number of stations available in each region listed in the legend.

Frequency bias and bias ratio highlight the satellite's tendency to over or underestimate the detection and representation of an event of a given size. These metrics are defined as

$$\text{frequency bias} = \frac{N_s}{N_o}, \quad \text{bias ratio} = \frac{\sum_{i=1}^{n} S_i}{\sum_{i=1}^{n} O_i},$$

where N_s (N_o) is the number of times IMERG-E (in situ station) detected (observed) an event of a given magnitude and S_i (O_i) is the IMERG-E estimate (in situ observation) of precipitation at each grid point (station location) i, with n being the total number of locations. For both the frequency bias and the bias ratio, values near 1 imply that IMERG-E agrees well with the observations, while values less (greater) than 1 indicate an IMERG-E underestimate (overestimate).

The MAE and RMSE are additional metrics used to assess the overall agreement between the IMERG-E values and in situ observations:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|S_i - O_i|, \quad \text{RMSE} = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}(S_i - O_i)^2\right)}.$$

Both the MAE and RMSE tend to be more sensitive to large errors arising from the occasional extremely large values of rainfall.

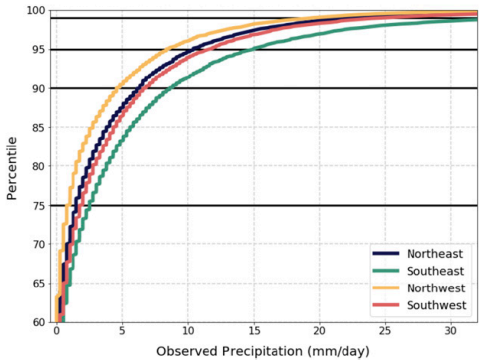

FIG. 3. Regional CDFs derived from daily precipitation observations during the six fire seasons from all stations in each region.

b. Bias correction

Globally, products that merge gauge data and satellite-based precipitation are often biased by unrepresentative and limited gauge data, particularly in mountains (Derin et al. 2016). This applies to the IMERG satellite-derived precipitation estimates, which are not expected to match in situ observations exactly since they are obtained from satellite-measured variables such as brightness temperature and other radiance quantities. Additionally, sensors have varied resolution across satellite platforms making it difficult to quantify uncertainty or even identify an error source when multiple sensors are inputs to a single algorithm, as is the case with IMERG (Bauer et al. 2011; Gebregiorgis et al. 2017). These problems appear exacerbated over regions of complex terrain or where in situ data are sparse. For these reasons, a bias correction technique is often used to mitigate for some of the uncertainties associated with satellite-derived precipitation estimate datasets.

We employ a simple bias correction method, hereafter referred to as regional quantile mapping, in order to improve the daily IMERG-E QPEs in a manner that could be used operationally. The regional aggregation of station and IMERG-E data is critical in order to avoid overfitting. The IMERG-E precipitation estimates at every grid point are bias corrected for each of the six summers independently, using the other five summers as a reference "climatology." The two steps are outlined here using a fictitious example for a specific grid point in the Northeast quadrant on 1 June 2014. We stipulate the IMERG-E estimated rainfall for that day and location was 5 mm:

1) A value of 5 mm in the Northeast quadrant in 2014 is found to correspond to the 85th percentile for that

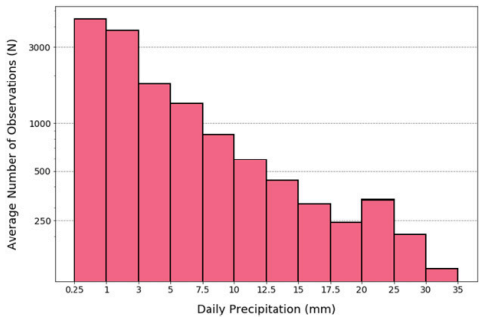

FIG. 4. Average number of daily precipitation observations across Alaska during the six fire seasons within the specified bin ranges.

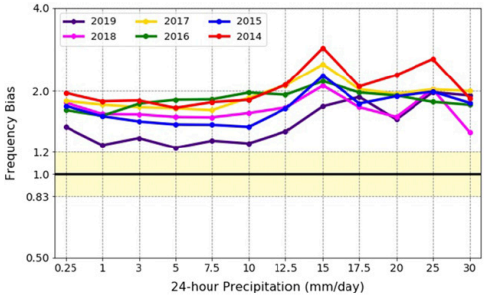

FIG. 5. IMERG-E frequency bias across Alaska as a function of daily observed precipitation (mm) during each summer season and accumulated into the specified bin ranges (e.g., the values at 0.25 represent bias values corresponding to observed precipitation in the range 0.25–1.0 mm). The shaded yellow region represents the acceptable range of frequency bias values (0.83–1.20) defined in the text.

quadrant based on all the IMERG-E values during the other 5 years (2015–19).

2) Since the 85th percentile for all in situ observations in the Northeast quadrant during the 2015–19 period is found to be 3 mm, the quantile mapped (QM) IMERG-E value for 1 June 2014 at that grid point is then set to 3 mm.

Using this technique, we bias correct the entire IMERG-E domain and evaluate those adjustments by comparing the independent observations available during each of the six summers to the gridpoint values near those locations (Hashemi et al. 2017). Limitations of this approach include the small discontinuities across the shared boundaries of the four regions arising from the CDF distributions computed for each separately. More serious limitations arise from the coarse resolution of the IMERG-E grid in mountainous regions within Alaska and when the observed conditions during the independent summer season differ substantially from those available in the dependent set of seasons used to derive the bias correction. These issues will be discussed later.

4. Results

a. *Verification of IMERG-E QPEs during the 2014–19 fire seasons*

During the Alaskan fire season, measurable precipitation occurs on approximately 40% of days (Fig. 3) with over 4000 of those observations in the 0.25–1.0 mm (0.01–0.04 in.) range, averaged over all years (Fig. 4). The frequency bias ratios within the specified precipitation ranges are shown in Fig. 5 for all stations during each of the six fire seasons. Since the regional variations in the frequency bias ratio were small, only the aggregated statewide results are presented in Fig. 5. Following other precipitation verification studies (e.g., Lewis et al. 2017; Gowan et al. 2018), we highlight reasonable frequency bias values as those between 0.83 and 1.20. The counts of IMERG-E daily amounts were more than 20% higher than those observed (frequency bias ratios larger than 1.20) during all years and within all magnitude ranges. For the ~200 observed events over 15 mm that occur on average each season (Fig. 4), the counts of IMERG-E precipitation amounts were more than 50% higher than those observed (Fig. 5).

Bias ratios calculated for each fire season, region, and specified CDF percentile ranges are shown in Fig. 6. IMERG-E consistently overestimates precipitation amounts in each region for all percentile ranges with the exception of the 2018 fire season in the Southeast region. Over the Northeast region, IMERG-E cumulative amounts were more than double those observed during all seasons for smaller rainfall amounts (60th–75th percentile range, ~0–3 mm) while the smallest bias ratios were found for large precipitation amounts in the Southeast region (95th–99th percentile range).

As should be expected, MAE and RMSE values are quite large for the highest percentile ranges (precipitation amounts) as shown in Figs. 7 and 8, respectively. During each fire season, the MAE and RMSE values are similar in magnitude, suggesting a consistent direction of the bias magnitude. The MAE and RMSE values for each fire season are consistently highest in the Northeast and Southeast regions as those are the regions with

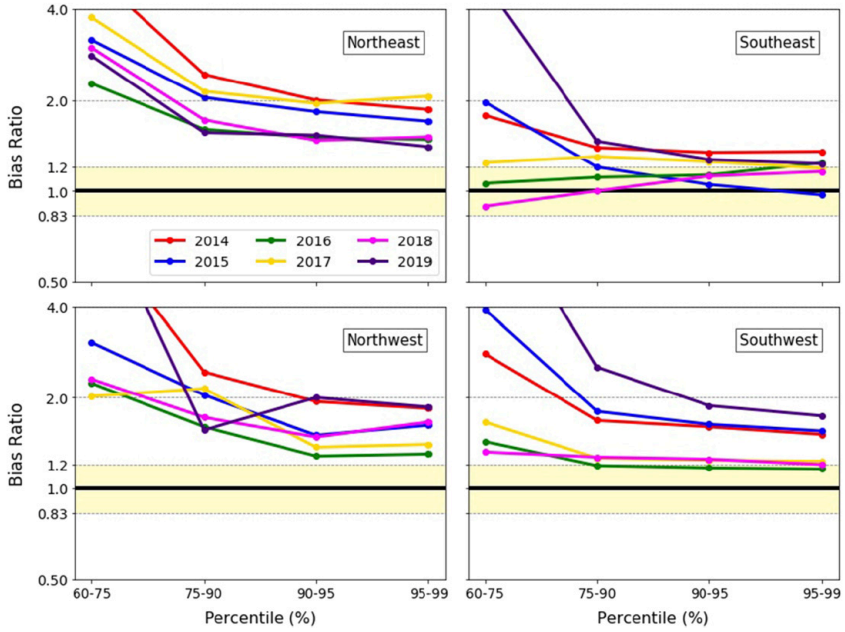

FIG. 6. Annual bias ratio as a function of region, fire season, and the observation's daily precipitation percentile (%). The shaded yellow region represents the acceptable range of frequency bias values (0.83–1.20) defined in the text.

higher climatological amounts of rainfall during the summer. For the Northwest and Southwest regions, the MAE (RMSE) values remain below 10 mm (15 mm) for all percentile ranges, with the exception of the 2019 fire season, which was marked by several anomalous precipitation events, including an atmospheric river event that produced record-breaking rain across the state.

b. *Bias correction of IMERG-E precipitation estimates*

To mitigate for the IMERG-E systematic wet bias tendency and enhance objectively the overall usability of the satellite data for fire weather applications, we employ regional quantile mapping described in section 3 to create QM IMERG-E daily QPEs at each grid point for the 2014–19 fire seasons. Observed rainfall data during the fire season for which the IMERG-E data are being corrected are not used in the quantile mapping so that they can be used to evaluate the level of improvement obtained by this approach.

The IMERG-E QPEs before (Fig. 9a) and after (Fig. 9b) the regional quantile mapping for 22 June 2018 are compared to the APRFC QPEs for the same time period (Fig. 9c, previously shown in Fig. 1c). Since the QM IMERG-E QPEs were calculated only within the state of Alaska, no precipitation is shown in Canada in Fig. 9b. The reduction of the IMERG-E wet bias as a result of the regional quantile mapping is evident by comparing Figs. 9a and 9b. The overall precipitation patterns are generally similar between the IMERG-E and APRFC estimates. However, the APRFC analysis appears to miss the north-south band of precipitation west of the Alaska–Canada border, presumably because of the limited number of observing sites in that region (Fig. 2). The IMERG-E analyses do not show the higher precipitation amounts suggested by the APRFC QPEs in the southwest quadrant, which is difficult to verify as there are no stations in that region on this day reporting large rainfall amounts. That may result from the APRFC weighting heavily climatological values in that region.

To assess the potential improvement resulting from the bias correction, we evaluate the QM IMERG-E data using the same approaches applied previously to the IMERG-E data. For example, the substantial reduction

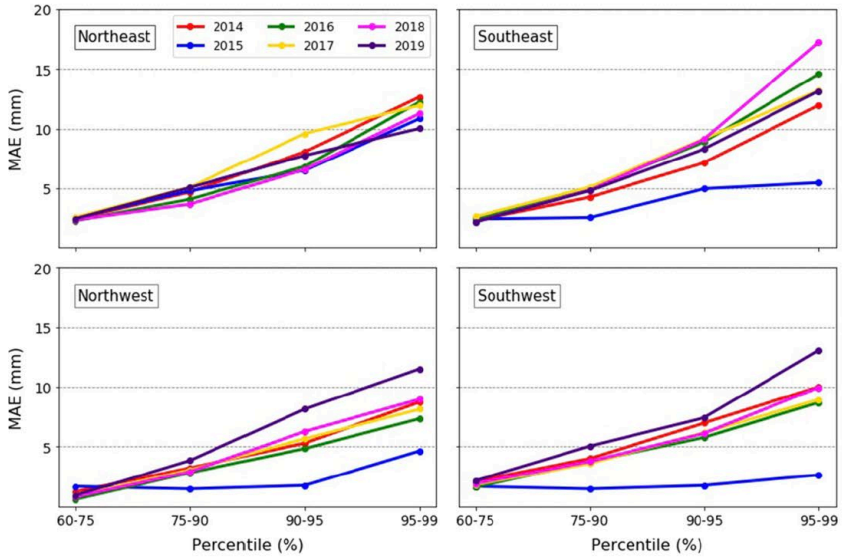

FIG. 7. MAE as a function of region, fire season, and the observation's daily precipitation accumulation percentile.

in the frequency biases for QM IMERG-E is evident for all precipitation amounts and fire seasons in Fig. 10 relative to those shown in Fig. 5 for IMERG-E. For some precipitation intervals and summer season, overestimates remain with a few cases where the QM IMERG-E now underestimates rainfall totals (e.g., for the largest precipitation amounts during 2019). Averaged over all six fire seasons, the QM IMERG-E frequency

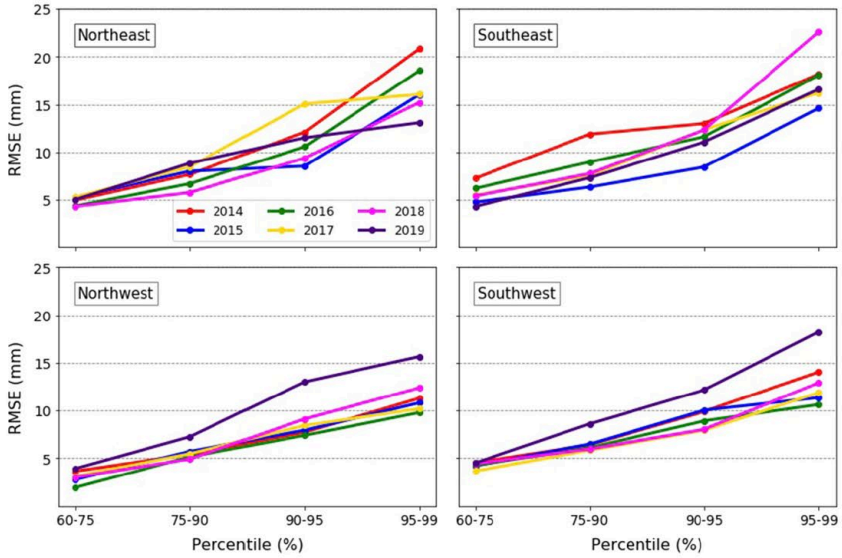

FIG. 8. As in Fig. 7, but for RMSE.

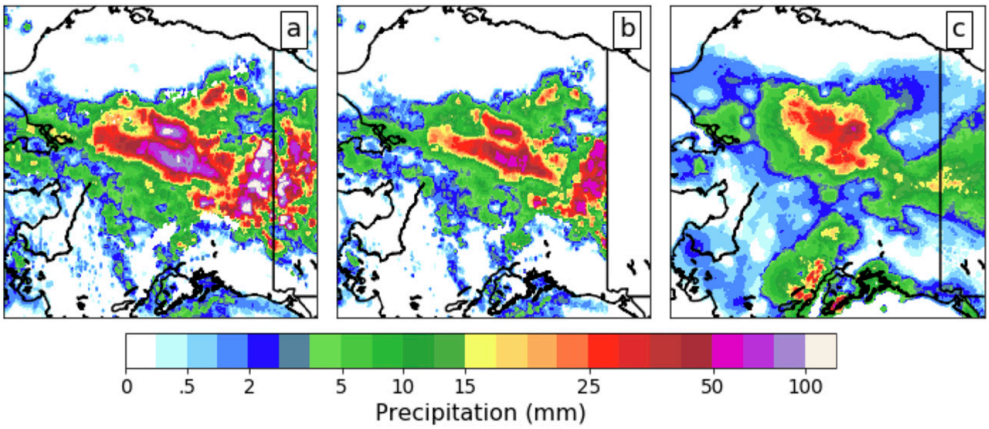

FIG. 9. 24-h precipitation estimates valid for 22 Jun 2018 shaded according to the scale below as derived by (a) IMERG-E, (b) QM IMERG-E, and (c) APRFC.

biases are improved for all regions from those obtained from IMERG-E.

Next, we discuss the QM IMERG-E bias ratios. Similar to the results from the frequency bias analysis, there was a substantial improvement in bias ratio score for all regions and percentile ranges. The improved bias ratios were similar for all regions and fire seasons (0.91–1.0), and thus are not shown graphically. The greatest percent improvement occurred for the Northeast and Northwest regions, which is encouraging from a fire weather standpoint given that many of the largest wildfires occur in these two regions. Additionally, the bias ratios for 60th–90th percentile ranges exhibit the largest percentage improvement from the baseline statistics for all regions. The average bias ratio values suggest that our bias correction method has been successful in accounting for and mitigating the wet bias that was originally present in the IMERG-E dataset.

We continue the evaluation of the QM IMERG-E precipitation estimates using the MAE (Fig. 11) and RMSE (Fig. 12) as a function of event percentile range and region. For the QM IMERG-E MAE, the most improvement occurred in the Northeast region for all percentile ranges, with the exception of the 60th–75th percentile regime. The average MAE slightly decreased in accuracy for Southeast and Southwest regions when evaluating 24-h precipitation estimates within the 90th–99th percentile ranges. For the QM IMERG-E RMSE, all regions improved for all percentile ranges, with the exception of the Southeast region showing little change. The largest RMSE improvement occurred for the Northeast region for the 60th–75th and 75th–90th percentile ranges, respectively. While not all of the results yielded a net positive change, the largest improvements with the QM IMERG-E dataset occurred for the regions that are historically highly susceptible to wildfires.

As a final illustration of both the strengths and weaknesses of bias correction approaches, time series of accumulated precipitation observations from a RAWS station near Delta Junction (OKLA2) in the Northeast region and the Anchorage International Airport (PANC) in the Southeast region during the 2019 summer season are shown in Fig. 13. IMERG-E precipitation tracks the evolution of observed rainfall at OKLA2 throughout the summer (i.e., there are no extraneous rainfall events generated by the satellite algorithm). The excessive IMERG-E precipitation estimates were largely corrected at OKLA2 by the quantile mapping

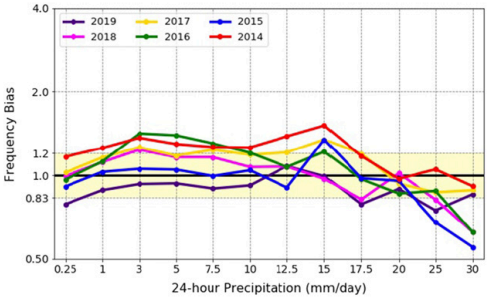

FIG. 10. As in Fig. 4, but for the QM IMERG-E dataset.

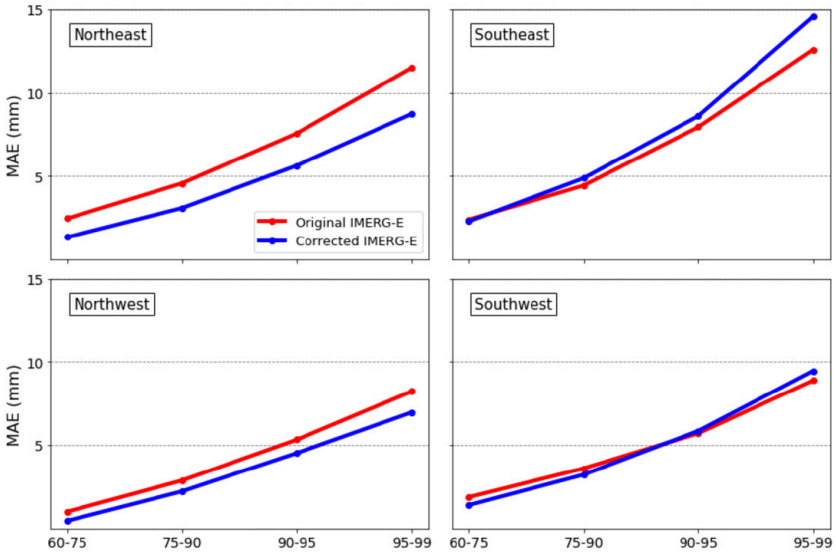

FIG. 11. Temporally averaged MAE for the IMERG-E (red line) and QM IMERG-E (blue line) as a function of region and the observation's daily precipitation accumulation percentile (%).

and similar useful corrections were observed during the other five years (not shown).

The amplitude of the estimated rainfall at PANC during the late July 2019 rain event was far in excess of that observed (Fig. 13). Interpolating the 0.1° (~10 km) resolution IMERG-E grid values to the PANC location may introduce representativeness errors, since the satellite estimates for the higher elevations in the nearby Chugach Mountains likely were larger than appropriate for PANC. The limited corrections to the IMERG-E estimates by the quantile mapping at PANC in this and the other summer seasons (not shown) are consistent with the limited changes in RMSE evident for the Southeast region during all seasons (Fig. 12).

5. Summary and future work

The IMERG-E QPEs are derived using retrievals from infrared and passive microwave sensors onboard multiple satellites. These retrievals are combined to create global snapshots of precipitation rates at a 0.1° horizontal grid spacing every half hour. This research is the first evaluation of the IMERG-E QPEs at high latitudes and is geared toward assessing the algorithm's potential for use in fire weather applications, such as for input to calculate the CFFDRS fire weather indices.

We evaluated the IMERG-E estimates of daily precipitation totals during six fire seasons relative to observations from 322 in situ stations across Alaska. Given Alaska's immense size and varied climate regimes, we sectioned the state into four regions and verified IMERG-E using the corresponding stations within those regions separately during each of the six seasons. Climatological cumulative distribution functions were used to help identify how frequent daily rainfall totals were observed and estimated by IMERG-E.

The baseline verification revealed that the IMERG-E QPEs have a wet bias regardless of geographic region, year, or precipitation amount. More IMERG-E events were counted than observed and IMERG-E totals tended to be larger for all regions and event ranges than observed. The largest MAE and RMSE errors were found in the Northeast and Southeast regions of the state for all years and percentile regimes, which reflects in part the climatologically large amounts of summer rainfall in those areas. The results presented here may help to identify how the IMERG-E algorithm could be improved to reduce the wet bias of IMERG-E QPEs during summer.

Following the baseline verification, we bias-corrected the IMERG-E dataset using a regional quantile mapping approach, which was followed by seasonal cross

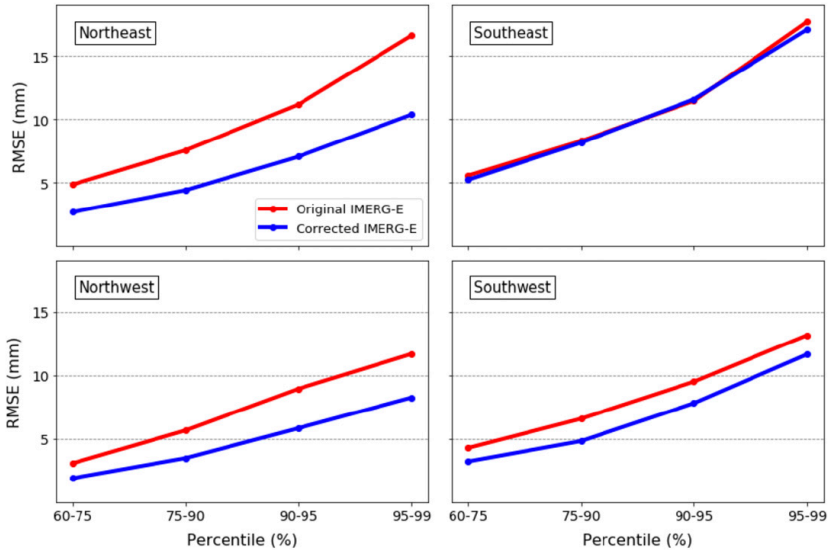

FIG. 12. As in Fig. 11, but for RMSE.

validation. The bias-corrected data were then evaluated in an identical manner to the original IMERG-E data. The bias correction yielded promising results:

1) The overall wet bias from the baseline verification was substantively reduced
2) Frequency bias and bias ratio scores fell to within acceptable ranges and approached unity for some regions and years
3) The greatest improvements in terms of MAE and RMSE were found in the Northeast and Northwest regions over all percentile ranges.

Bias correction adds value when the climatological range expected for the observed conditions is adequately sampled in space and time. As summarized by Lundquist et al. (2019), satellite-derived estimates of precipitation may not be the solution for accurately representing precipitation in complex terrain. Quantile mapping the satellite-derived estimates will also be less useful in the mountainous regions of Alaska since there are few observations available that are representative of conditions in those areas. The relatively coarse resolution of the IMERG-E grid also exacerbates such spatial representativeness concerns in regions of complex terrain. Changes over time in the types of satellite-derived products used in the IMERG-E algorithm could potentially affect the quantile mapping. However, a more serious issue is whether the observations aggregated within the region over a limited sample of years provide a sufficiently diverse climatology to reflect accurately what may happen during other summer seasons.

The IMERG-E QPEs are an invaluable resource for high-latitude locations such as Alaska where large swaths of the state have no automated precipitation observations. The quantile mapping approach was tested here by slicing Alaska into four arbitrary regions with roughly similar numbers of stations. The artificiality of the sector boundaries creates small discontinuities across the shared boundaries of the quadrants arising from the CDF distributions computed for each separately. This approach could be improved by adjusting the sectors to conform to regions with similar summer precipitation regimes, climate divisions, or the Predictive Service Areas define by the AICC.

IMERG-E QPE may be appropriate to use for computing fire potential and behavior indices in Alaska if its wet bias can be overcome. Further work is needed to assess what may be the minimum number of stations in the regions and number of summer seasons necessary to provide adequate samples for defining the observed regional CDFs on which the technique depends. Since the IMERG-E data are not available at high latitudes prior to 2014, this approach may not be practical in other high-latitude regions where there are insufficient numbers of stations available during recent years.

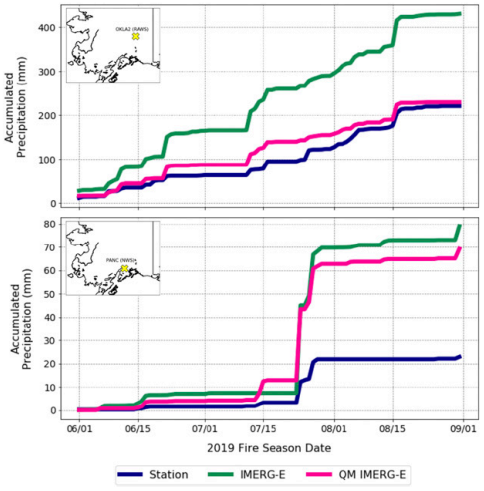

FIG. 13. Daily accumulated precipitation (mm) for the 2019 fire season as estimated by in situ stations (blue line; station code, location, and network in upper-left corner of panels) and the corresponding IMERG-E and QM IMERG-E values (green and pink lines, respectively).

Acknowledgments. The authors thank George Huffman of the NASA Goddard Space Flight Center for generously using his time and GPM IMERG expertise to help make this study possible. The authors would also like to thank Robert Ziel, an Alaskan fire behavior analyst, who went above and beyond giving us his time and sharing his first-hand knowledge and experience throughout the duration of this study. This work was supported by the Joint Fire Science Program (JFSP Grant L17AC00225) and a Collaborative Science, Technology, and Applied Research (CSTAR) Program Award (55500146). Access to the IMERG data was provided by the NASA Precipitation Processing System and the Goddard Earth Sciences Data and Information Services Center. The University of Utah Center for High Performance Computing (CHPC) provided computational hardware and software.

REFERENCES

Anjum, M. N., Y. Ding, D. Shangguan, I. Ahmad, M. W. Ijaz, H. U. Farid, Y. E. Yagoub, and M. Adnan, 2018: Performance evaluation of latest integrated multi-satellite retrievals for global precipitation measurement (IMERG) over the northern highlands of Pakistan. *Atmos. Res.*, **205**, 134–146, https://doi.org/10.1016/j.atmosres.2018.02.010.

Asong, Z. E., S. Razavi, H. S. Wheater, and J. S. Wong, 2017: Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over southern Canada against ground precipitation observations: A preliminary assessment. *J. Hydrometeor.*, **18**, 1033–1050, https://doi.org/10.1175/JHM-D-16-0187.1.

Bauer, P., and Coauthors, 2011: Satellite cloud and precipitation assimilation at operational NWP centres. *Quart. J. Roy. Meteor. Soc.*, **137**, 1934–1951, https://doi.org/10.1002/qj.905.

Bieniek, P. A., and Coauthors, 2012: Climate divisions for Alaska based on objective methods. *J. Appl. Meteor. Climatol.*, **51**, 1276–1289, https://doi.org/10.1175/JAMC-D-11-0168.1.

——, J. E. Walsh, R. L. Thoman, and U. S. Bhatt, 2014: Using climate divisions to analyze variations and trends in Alaska temperature and precipitation. *J. Climate*, **27**, 2800–2818, https://doi.org/10.1175/JCLI-D-13-00342.1.

De Groot, W. J., P. M. Bothwell, D. H. Carlsson, and K. A. Logan, 2003: Simulating the effects of future fire regimes on western Canadian boreal forests. *J. Veg. Sci.*, **14**, 355–364, https://doi.org/10.1111/j.1654-1103.2003.tb02161.x.

De Pondeca, M. S., and Coauthors, 2011: The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: Current status and development. *Wea. Forecasting*, **26**, 593–612, https://doi.org/10.1175/WAF-D-10-05037.1.

Derin, Y., and Coauthors, 2016: Multiregional satellite precipitation products evaluations over complex terrain. *J. Hydrometeor.*, **17**, 1817–1836, https://doi.org/10.1175/JHM-D-15-0197.1.

Dezfuli, A. K., C. M. Ichoku, G. J. Huffman, K. I. Mohr, J. S. Selker, N. van de Giesen, R. Hochreutener, and F. O. Annor, 2017: Validation of IMERG precipitation in Africa. *J. Hydrometeor.*, **18**, 2817–2825, https://doi.org/10.1175/JHM-D-17-0139.1.

Erickson, M. J., B. A. Colle, and J. J. Charney, 2012: Impact of bias correction type and conditional training on Bayesian model averaging over the northeast United States. *Wea. Forecasting*, **27**, 1449–1469, https://doi.org/10.1175/WAF-D-11-00149.1.

Field, R. D., and Coauthors, 2015: Development of a global fire weather database. *Nat. Hazards Earth Syst. Sci.*, **15**, 1407–1423, https://doi.org/10.5194/nhess-15-1407-2015.

Flannigan, M. D., A. S. Cantin, W. J. de Groot, M. Wotton, A. Newbery, and L. Gowman, 2013: Global wildland fire season severity in the 21st century. *For. Ecol. Manage.*, **294**, 54–61, https://doi.org/10.1016/j.foreco.2012.10.022.

Gaona, M. F. R., A. Overeem, H. Leijnse, and R. Uilenhoet, 2016: First-year evaluation of GPM rainfall over the Netherlands: IMERG day 1 final run (V03D). *J. Hydrometeor.*, **17**, 2799–2814, https://doi.org/10.1175/JHM-D-16-0087.1.

Gebregiorgis, A. S., P. Kirstetter, Y. E. Hong, N. J. Carr, J. J. Gourley, W. Petersen, and Y. Zheng, 2017: Understanding overland multisensor satellite precipitation error in TMPA-RT products. *J. Hydrometeor.*, **18**, 285–306, https://doi.org/10.1175/JHM-D-15-0207.1.

Gowan, T. M., W. J. Steenburgh, and C. S. Schwartz, 2018: Validation of mountain precipitation forecasts from the convection-permitting NCAR ensemble and operational forecast systems over the western United States. *Wea. Forecasting*, **33**, 739–765, https://doi.org/10.1175/WAF-D-17-0144.1.

GPM, 2013: Precipitation measurement missions. Global Precipitation Measurement Program, NASA, accessed 15 July 2019, http://gpm.nasa.gov/.

GPM ATBD, 2019: Global precipitation measurement algorithm theoretical basis documents. NASA, accessed 22 July 2019, https://pps.gsfc.nasa.gov/atbd.html.

Hashemi, H., M. Nordin, V. Lakshmi, G. J. Huffman, and R. Knight, 2017: Bias correction of long-term satellite monthly precipitation product (TRMM 3B43) over the conterminous United States. *J. Hydrometeor.*, **18**, 2491–2509, https://doi.org/10.1175/JHM-D-17-0025.1.

Hong, Y., K. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial

neural network cloud classification system. *J. Appl. Meteor.*, **43**, 1834–1853, https://doi.org/10.1175/JAM2173.1.

Horel, J. D., R. Ziel, C. Galli, J. Pechmann, and X. Dong, 2014: An evaluation of fire danger and behaviour indices in the Great Lakes Region calculated from station and gridded weather information. *Int. J. Wildland Fire*, **23**, 202–214, https://doi.org/10.1071/WF12186.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20, https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

——, R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.*, **2**, 36–50, https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

——, and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum, 2017: So, how much of the Earth's surface is covered by rain gauges? *Bull. Amer. Meteor. Soc.*, **98**, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1.

Kirstetter, P. E., Y. Hong, J. J. Gourley, Q. Cao, and M. Schwaller, and W. Petersen, 2014: Research framework to bridge from the global precipitation measurement mission core satellite to the constellation sensors using ground-radar-based national mosaic QPE. *Remote Sensing of the Terrestrial Water Cycle, Geophys. Monogr.*, Vol. 206, Amer. Geophys. Union, 61–79.

Kotsuki, S., K. Kurosawa, S. Otsuka, K. Terasaki, and T. Miyoshi, 2019: Global precipitation forecasts by merging extrapolation-based nowcast and numerical weather prediction with locally optimized weights. *Wea. Forecasting*, **34**, 701–714, https://doi.org/10.1175/WAF-D-18-0164.1.

Lader, R., J. E. Walsh, U. S. Bhatt, and P. A. Bieniek, 2017: Projections of twenty-first-century climate extremes for Alaska via dynamical downscaling and quantile mapping. *J. Appl. Meteor. Climatol.*, **56**, 2393–2409, https://doi.org/10.1175/JAMC-D-16-0415.1.

Lawson, B. D., and O. B. Armitage, 2008: Weather guide for the Canadian forest fire danger rating system. Tech. Rep. Natural Resources Canada, Canadian Forest Service, 87 pp.

Lewis, W. R., W. J. Steenburgh, T. I. Alcott, and J. J. Rutz, 2017: GEFS precipitation forecasts and the implications of statistical downscaling over the western United States. *Wea. Forecasting*, **32**, 1007–1028, https://doi.org/10.1175/WAF-D-16-0179.1.

Liu, Z., D. Ostrenga, W. Teng, and S. Kempler, 2012: Tropical Rainfall Measuring Mission (TRMM) precipitation data and services for research and applications. *Bull. Amer. Meteor. Soc.*, **93**, 1317–1325, https://doi.org/10.1175/BAMS-D-11-00152.1.

Lundquist, J., M. Hughes, E. Gutmann, and S. Kapnick, 2019: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. *Bull. Amer. Meteor. Soc.*, **100**, 2473–2490, https://doi.org/10.1175/BAMS-D-19-0001.1.

McCorkle, T. A., J. D. Horel, A. A. Jacques, and T. Alcott, 2018: Evaluating the experimental high-resolution Rapid Refresh–Alaska modeling system using USArray pressure observations. *Wea. Forecasting*, **33**, 933–953, https://doi.org/10.1175/WAF-D-17-0155.1.

Melvin, A. M., J. Murray, B. Boehlert, J. A. Martinich, L. Rennels, and T. S. Rupp, 2017: Estimating wildfire response costs in Alaska's changing climate. *Climatic Change*, **141**, 783–795, https://doi.org/10.1007/s10584-017-1923-2.

Naud, C. M., J. Jeyaratnam, J. F. Booth, M. Zhao, and A. Gettelman, 2020: Evaluation of modeled precipitation in oceanic extratropical cyclones using IMERG. *J. Climate*, **33**, 95–113, https://doi.org/10.1175/JCLI-D-19-0369.1.

Nelson, B. R., O. P. Prat, D. Seo, and E. Habib, 2016: Assessment and implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Pickell, P. D., N. C. Coops, C. J. Ferster, C. W. Bater, K. D. Blouin, M. D. Flannigan, and J. Zhang, 2017: An early warning system to forecast the close of the spring burning window from satellite-observed greenness. *Sci. Rep.*, **7**, 14190, https://doi.org/10.1038/s41598-017-14730-0.

Pierce, D. W., D. R. Cayan, E. P. Maurer, J. T. Abatzoglou, and K. C. Hegewisch, 2015: Improved bias correction techniques for hydrological simulations of climate change. *J. Hydrometeor.*, **16**, 2421–2442, https://doi.org/10.1175/JHM-D-14-0236.1.

Sharafi, E., R. Steinacker, and B. Saghafian, 2016: Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.*, **8**, 135, https://doi.org/10.3390/rs8020135.

Shulski, M., and G. Wendler, 2007: *The Climate of Alaska*. University of Alaska Press, 216 pp.

Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Amer. Meteor. Soc.*, **98**, 1679–1695, https://doi.org/10.1175/BAMS-D-15-00306.1.

Sungmin, O., U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, and W. A. Petersen, 2017: Evaluation of GPM IMERG early, late, and final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.*, **21**, 6559–6572, https://doi.org/10.5194/hess-21-6559-2017.

Tang, G., Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016: Evaluation of GPM day-1 IMERG and TMPA version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167, https://doi.org/10.1016/j.jhydrol.2015.12.008.

Tian, F., S. Hou, L. Yang, H. Hu, and A. Hou, 2018: How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? *J. Hydrometeor.*, **19**, 339–349, https://doi.org/10.1175/JHM-D-17-0161.1.

Ziel, R. H., J. Wolken, T. St. Clair, and M. Henderson, 2015: Modeling fire growth potential by emphasizing significant growth events: Characterizing a climatology of fire growth days in Alaska's boreal forest. *11th Symp. on Fire and Forest Meteorology*, Minneapolis, MN, Amer. Meteor. Soc.,1.2, https://ams.confex.com/ams/11FIRE/webprogram/Paper272864.html.

Zubieta, R., A. Getirana, J. C. Espinoza, W. Lavado-Casimiro, and L. Aragon, 2017: Hydrological modeling of the Peruvian-Ecuadorian Amazon Basin using GPM-IMERG satellite-based precipitation dataset. *Hydrol. Earth Syst. Sci.*, **21**, 3543–3555, https://doi.org/10.5194/hess-21-3543-2017.

CHAPTER 3

ARCHIVAL AND ANALYSIS OF HIGH-RESOLUTION RAPID REFRESH MODEL OUTPUT USING ZARR FILES IN THE CLOUD

3.1 Abstract

Output from operational numerical weather prediction models is relied upon for diverse applications. The dominant format (GRIdded Binary Second Edition, GRIB2) used by operational centers to disseminate files within a model run relies on highly efficient compression of two-dimensional grids and results in O(100 MB) file sizes. Processing time and storage needs are high if large numbers of GRIB2 files are required. We illustrate one approach to overcome such bottlenecks by reformatting GRIB2 model output from the High-Resolution Rapid Refresh (HRRR) model of the National Centers for Environmental Prediction to an N-dimensional, cloud-compatible file type, Zarr. The resulting data archive (HRRR-Zarr) is stored using the Amazon Web Service Simple Storage Service and available publicly through the Amazon Open Data program.

Currently, 576 HRRR files are generated every day for each of three sets of grids (surface, pressure, and native). The surface set contains 173 grids representing a mix of variables at levels in the vertical of highest interest for many applications. Each grid contains 1.9 million grid points for the contiguous United States region that are

subdivided into 96 chunks containing 150x150 grid points. Chunks for all forecast times within a model run are combined into O(1 MB) files for each field type (parameter, level, etc.). Over 17,000 small file objects are created from each model run in order to provide flexibility to access the subdomains and variables of interest. The HRRR-Zarr approach is illustrated for use cases that benefit from the HRRR surface files, such as real-time alerts for high-impact situations and efficient access to output from hundreds of model runs relevant to machine-learning applications of sensible weather parameters.

3.2 Introduction

The global weather enterprise relies on millions of large, two-dimensional data fields created each day by operational numerical weather prediction (NWP) models (Benjamin et al. 2018). The perceptions, uses, and values for that vast amount of information depends in part on how it is disseminated and accessible to end users (Lazo et al. 2009). Advances in computing processing power and storage have allowed operational centers to run models at finer spatial scales and higher temporal frequency (Benjamin et al. 2018), yet only a small fraction of the information available from the models are typically available to end users. Pragmatic decisions are made by operational forecast centers in order to disseminate global and regional model output for dozens of parameters that are restricted to specific ranges and frequencies of valid times and horizontal and vertical grid spacings. Those decisions have been heavily influenced by internal and external limitations on storing and accessing the hundreds of gigabytes (GB) of model output generated by each model run. These challenges are not unique to the weather sector; many disciplines are struggling to overcome the “Volume, Variety, and

Velocity" of data cubes (datasets in space and time) available from earth observation systems (Giuliani et al. 2020; Yao et al. 2020). Improved data cube cyber-infrastructures are recognized to be needed for environmental data sets to allow the ingestion, storage, access, analysis, and use of data elements ordered by geolocation and other shared attributes (Nativi et al. 2017).

The National Oceanic and Atmospheric Administration (NOAA) Big Data Program (BDP) was established in 2015 to address agency-wide issues to access the tens of terabytes of observations and model output created each day within the agency (NOAA 2020; Ansari et al. 2018). This initiative involves public-private partnerships with several infrastructure-as-a-service (IaaS) providers for data storage for over 130 data streams, including such high-demand data sets as current and historical Next Generation Weather Radar (NEXRAD) products from 160 sites in the United States (Ansari et al. 2018). IaaS providers (e.g., Amazon Web Services- AWS; Google Cloud Platform; IBM; and Microsoft Azure) have the capacity to store enormous datasets and provide public access and computing resources for end users to post-process these data streams within their IaaS environment to reduce the time and cost to download the information to local compute resources (Molthan et al. 2015; Siuta et al. 2016).

The Google Cloud Platform and AWS Simple Storage Service (S3) began providing public access during 2020 to output from the High-Resolution Rapid Refresh (HRRR) model of the National Centers for Environmental Prediction (NCEP). The HRRR is a convection-allowing model that was developed by the Earth Systems Research Lab (ESRL) and is run operationally every hour by the NCEP's Environmental Modeling Center. Its output is available for dozens of surface and upper-atmospheric

variables at 3-km grid spacing for a 1.9 million grid-point domain that covers the contiguous United States (CONUS; Benjamin et al. 2016; Blaylock et al. 2017). The NOAA partnership with IaaS providers creates storage redundancy as well as facilitating public access to data, such as HRRR model output and many other datasets, that are crucial for research and development.

From 2016-2020, over 1000 operational and research users relied on the only publicly-accessible archive of HRRR model output that was managed by researchers at the University of Utah using S3-type storage provided by the Center for High Performance Computing (Blaylock et al. 2017; Blaylock et al. 2018). As the dataset grew (Fig. 3.1), it became clear that the storage capacity to continue to maintain and expand the HRRR archive at the University of Utah would no longer be feasible with the introduction of HRRRv4. Research began to explore approaches for storing HRRR model output appropriate for machine learning applications.

International standards were established by the World Meteorological Organization (WMO) to efficiently store and disseminate numerical model output in hypercube-structured file formats with built-in compression algorithms. The GRIdded Binary Second Edition (GRIB2) format has been in use during the past several decades to archive two-dimensional files that are efficiently compressed using a method similar to JPEG image compression (Silver and Zender 2017). While GRIB2 files effectively help store and transmit large amounts of meteorological data as two-dimensional slices, they can be cumbersome to work with and rely on WMO-defined tables that are unfamiliar to users in other disciplines (Wang 2014). Many users rely on software tools to transform GRIB2 files into other self-describing formats such as netCDF (Silver and Zender 2017).

Decoding the two-dimensional slices in GRIB2 format leads to expanded file sizes that contribute to inefficiencies when an end user may, for example, be interested in certain parameters for all forecast times available from a specific model run within a local or regional subdomain. However, it is possible to access individual variables within GRIB2 files by selecting their byte range or specifying a bounding box for domain subsets, but doing so requires loading each file into memory and performing additional post-processing (Blaylock et al. 2017).

An alternative WMO format used for vertical slices of model output at specific locations is the Binary Universal Form for the Representation of meteorological data (BUFR). However, the external-table descriptors for the BUFR format contributes to impediments to decoding for many applications, leading to the need for tools to convert to other formats such as netCDF.

Researchers generally use high-level programming environments that rely on Matlab, IDL or Python to examine, post-process, and visualize operational model data. For open-source languages such as Python, few libraries exist that read GRIB2 files efficiently and the hundreds of encoded variables that they contain. Data science and machine learning (ML) techniques applied to operational model output typically require multivariate training datasets with long periods of record for which alternative model data structures beyond GRIB2 or netCDF are necessary (Vannitsem et al. 2020). As summarized by McGovern et al. (2017), these big data and ML methods have been used to improve forecasts of high-impact weather parameters such as storm duration (Cintineo et al. 2014), severe wind (Lagerquist 2016), large hail (Adams-Selin and Ziegler 2016), precipitation type (Reeves et al. 2014; Elmore et al. 2015) and aviation turbulence

(Sharman 2016). To continue applying these ML and artificial intelligence techniques to the ever-growing model output repositories, it will be critical to have data in structures that allow for flexible dissection in space, time, across many runs or ensemble members (McGovern et al. 2017).

An alternative file format, Zarr, is applied in this study to HRRR files originally archived by the University of Utah (Blaylock et al. 2017; Blaylock et al. 2018). Zarr is a relatively new file format, developed in 2016 for use in a Malaria genome project, which chunks and compresses N-dimensional datasets for storage in memory, disk, or cloud platforms (Vance et al. 2019; Miles et al. 2020). Zarr is a hypercube file format that provides more flexibility for structure and storage options. The Zarr format is already being used for promising ML and big data applications in other disciplines, e.g., Lyft Level 5 self-driving dataset (Houston et al. 2020), the MalariaGEN project (Pearson et al. 2019), and the Pangeo project (Signell and Pothina 2019; Eynards-Bontemps et al. 2019). In the weather enterprise, the United Kingdom's Met Office has adopted Zarr as its file storage format of choice for the 200+ terabytes (TB) of data produced by high-resolution NWP models each day (McCaie 2019).

The HRRR model output in Zarr format developed in this study (hereafter HRRR-Zarr) is one approach to extract and disseminate model output intended for typical ML workflows that may require from one model run a subset of variables within regional model domains for all times in its forecast period. HRRR-Zarr makes it more practical to then aggregate that relatively small fraction of data originally in GRIB2 format from thousands of model runs. The capability to do so is possible since HRRR-Zarr formatted

files are being stored in the AWS S3 environment and publicly accessible as part of the AWS Open Data Program to complement the HRRR GRIB2 model archive available.

The remainder of this manuscript will be organized in the following manner. We first detail the HRRR model specifications, Zarr capabilities and limitations, and the AWS HRRR-Zarr archive structure. Next, we explore potential use cases for the HRRR-Zarr dataset, for both research or operational applications. We will detail the benefits of the HRRR-Zarr format in a general sense, as well as demonstrate its utility in analyzing a high-impact meteorological event from September 2020 that included record-breaking downslope windstorms in two states, devastating wildfire spread, and an early season snowstorm. Finally, the conclusions and future work are presented.

3.3 Data and Methods

3.3.1 The High-Resolution Rapid Refresh Model

The High-Resolution Rapid Refresh (HRRR) is a 3-km, convection-allowing model that is run operationally by NCEP's Environmental Modeling Center (Benjamin et al. 2016). It was developed by the Earth Systems Research Laboratory and was first run operationally September 2014. The latest version of the HRRR model (version 4, deployed 2 December 2020), is initialized each hour, with hourly forecasts out to either 18 or 48 hours depending on the initialization time (Table 3.1). The operational HRRR domain covers the entire CONUS, with a separate domain for the state of Alaska (McCorkle et al. 2018). The HRRR is nested within the larger domain of the Rapid Refresh model (RAP), from which it receives its initial and boundary conditions for each model run (Fig. 3.2). The RAP employs identical parameterization schemes, with the

exception of a convection parameterization, and assimilates data using the NOAA Gridpoint Statistical Interpolation system, which was modified to include hourly radar data, boundary layer observations, and other cloud processes (Kleist et al. 2009). As discussed by McCorkle et al. (2018), the HRRR model is initialized 1 hour prior to its analysis time, known as a pre-forecast, in order to assimilate 3-dimensional radar reflectivity data, which impacts latent heating estimations and thus the HRRR's ability to forecast convection (James and Benjamin 2017).

Staff in the NOAA Big Data Program (BDP) manage the distribution of HRRR model output to IaaS providers Google and AWS. We rely on the archive and real-time HRRR GRIB2 files available publicly as part of the AWS Open Data Program (https://registry.opendata.aws/noaa-hrrr-pds/). The HRRR GRIB2 files are publicly accessible via the AWS S3 using the unique identifier "noaa-hrrr-bdp-pds".

Currently, 576 HRRR GRIB2 files are generated every day for each of three sets of grids (surface, pressure, and native). We focus our description on the surface set that contains currently 173 grids representing a mix of variables at levels in the vertical of highest interest for many applications (Table 3.1). The pressure and native files contain meteorological variables at fixed pressure or terrain-following levels, respectively that are most relevant for users who need the HRRR output for initial and boundary conditions to initialize high resolution forecasts or research simulations (e.g., Crosman and Horel 2017; Foster et al. 2017). Output from HRRR model runs initialized at 00, 06, 12, and 18 UTC are available from the analysis time (F00) and from hourly forecasts through 48 h (F48). The HRRR model runs initialized at the other hours of the days are available from F00-F18.

3.3.2 Zarr

Zarr is a flexible file format for storing N-dimensional data arrays that are chunked and compressed based on the user specifications. In terms of its functionality, the Zarr protocol is similar to the popular data format, Hierarchical Data Format version 5 (HDF5; Delaunay et al. 2019). Zarr files are read and written with the Zarr Python library, which depends on the widely-used NumPy library (Harris et al. 2020). The Zarr format is becoming a more desirable file structure choice among data scientists and researchers because of its seamless ability to read and write to cloud platforms. Other benefits include its library of compression options, multithreading and multiprocessing capabilities, and its backend compatibility with format-agnostic, array-manipulation Python libraries (e.g., xarray, iris, and dask).

A Zarr file is initialized by first defining the file store, which can be in memory, as a directory on local disk, in distributed or cloud storage, or as a zip file. Next, Zarr arrays (hereafter, zarrays) are created and filled in a similar manner to NumPy arrays by defining a dtype and shape, and then assigning values and defining zarray attributes (zattrs) that will serve as the key references for that zarray. These zarrays can be chunked along any specified dimension and in any shape, which allows a dataset to be manipulated and stored efficiently for use in specific applications. All chunks in a zarray are uniform in shape and stored as individual objects that are identified by their integer index location in the array (e.g., row and column). The process of defining an optimal chunk structure for the HRRR-Zarr dataset is outlined in the next subsection.

Before sending the chunked zarrays to the Zarr store, they first must be encoded and compressed. The encoding instructions are located in the metadata of a zarray and define the data type's byte order (little endian or big endian), character code (integer, floating point, Boolean, etc.), and the number of bytes. All data types in the NumPy array protocol are acceptable for zarray encoding. Once the encoding parameters have been defined, the zarrays can be compressed using a number of compression algorithms and data filters. The Numcodecs library was designed specifically for data storage applications like Zarr, and serves as an interface to other compressor libraries such as Blosc, Zstandard, LZ4, Zlib, and LZMA. This allows the user to choose the primary compressor, the compression algorithm, and the compression level that will perform best based on the applications of the dataset. When choosing compression codec and level, the user takes into consideration the potential compression ratio (Eqn. 1) and decoding speed.

$$Compression\ Ratio = \frac{Uncompressed\ Data\ Volume}{Compressed\ Data\ Volume} \tag{3.1}$$

Fortunately, there exists a plethora of literature that details compression algorithm performance and benchmark test results that can aid choosing appropriate compression schemes for a particular use case (Donoho 1993; Alted 2010; Almeida et al. 2014; Wang et al. 2015; Kuhn et al. 2016). In addition to choosing a compression scheme, the Numcodecs library also offers a number of data filters that can be implemented. The filter sorts the data and transforms it in a way that would streamline compression, such as shuffling bytes and bits when adjacent values in an array are correlated. The compatibility

of the Zarr protocol and Numcodecs libraries allows for the configuration of an external filter for use with any chosen compressor, even if the filter is not an option by default. A detailed overview of the encoding and compression schemes used to create the HRRR-Zarr archive is presented in the next subsection

3.3.3 The HRRR-Zarr Archive

The HRRR-Zarr archive available publicly as part of the AWS Open Data Program with the unique bucket identifier "hrrrzarr" was designed to be relevant for users less familiar with environmental data set formats while supporting a familiar environment for users who have been exposed to model output in netCDF or GRIB2 format. This subsection details the conversion workflow and hierarchical structure of the data cubes in each HRRR-Zarr file, followed by a summary of the encoding and compression choices, and concludes with a description of the structure of the hrrrzarr AWS S3 bucket.

The HRRR-Zarr conversion workflow follows that of the United Kingdom's Met Office Informatics Lab, where they are actively using Zarr to store large datasets (Donkers 2020). Due to the challenges that surround manipulating data cubes in various file formats, the Met Office developed the Iris Python library, a format-agnostic library for processing datasets and converting between file formats (Iris 2020). Unlike other Python libraries, Iris and its companion package, Iris-grib, were built to read data cube formats such as GRIB2 and recognize the Climate and Forecast (CF; Eaton et al. 2020) metadata conventions used in model data. For this reason, the Iris libraries were the optimal choice for transitioning the HRRR archive from GRIB2 to Zarr format without compromising metadata. The list of Python libraries required for file creation and retrieval use can be found in Table 3.2.

The HRRR-Zarr archive files are built using the Iris and Iris-grib libraries. HRRR-Zarr data files rely on the same self-describing metadata (keywords and CF naming scheme) as the corresponding GRIB2 files obtained from their associated index (.idx) files. As an example, consider the workflow required to process the 49 GRIB2 surface files containing 173 grids from the HRRR model runs initialized at 00 UTC. (The CF names for all 173 HRRR variables are listed in the Appendix as they are not readily available online.) All of the grids from the 49 hourly forecast files are read into memory and then organized into unique Iris data cubes containing data and metadata. The iris data cubes are then converted to zarrays that are subdivided (chunked), encoded, and output into separate files identified by the parameter's CF name and atmospheric level or layer (e.g., 2-m, 500 mb). We maintain separate directory structures for (1) analysis files only and (2) files containing all possible times (F00-F48).

As shown in Fig. 3.3, the 1799 x 1059 grid is subdivided into 96 chunks of size 150 x 150 based on recommendations for optimal data compression. (The topmost 12 chunks contain data in only 9 rows.) We chose the LZ4 compression codec, which is a lossless compression algorithm with the ability to quickly and efficiently compress large amounts of data (Collett 2020). The zarrays are encoded as 16-bit little-endian floats, with the exception of the surface pressure parameter, which requires 32-bit little-endian floats. We access the LZ4 algorithm using the Numcodecs library class, Blosc, which is a meta-compressor that imitates the utility of the built-in Python library zlib. When using the LZ4 compression algorithm, additional modifications can be made to tailor the scheme for a particular use. We chose byte shuffling and a compression level of 9 within a range of 1-12 where levels 1 and 12 provide the fastest compression speed and highest

compression ratio, respectively. While other compression codecs have been shown to produce higher data compression ratios, their decompression speeds are much slower (Collett 2020).

It is widely recognized that optimal use of cloud resources requires having data processing and analysis within the same compute environment as the data archive. The HRRR-Zarr files are created shortly after the entire model run is accessible as objects in the AWS GRIB2 S3 archive using AWS Elastic Cloud Compute resources in the same region (West-1) as our hrrrzarr bucket. Our Python code uses the libraries mentioned in the previous subsection along with the Amazon Boto3 Python library.

Within the S3:hrrrzarr bucket, all files (or objects) are contained in a flat structure where the concept of folder or directory structure is provided by using shared name prefixes or suffixes for objects that mimic traditional directory structure. The zarr files derived from the surface and pressure sets of HRRR GRIB2 files are stored with the prefixes “sfc/” and “prs/”, respectively (Fig. 3.4). Model runs are accessible by date using suffixes for analysis (anl.zarr) and forecast (fcst.zarr) files for each model run, e.g., files with the prefix sfc/20200907_12z_fcst.zarr/ were generated from the F00-F48 HRRR GRIB2 surface files initialized at 1200 UTC 7 Sept 2020. Prefixes follow then based on level (e.g., 700 mb/ or 10m_above_ground/) and CF naming conventions for variables a (e.g., TMP/ or UGRD/) with the final part of the file name being the chunk identifier. A full list of variables (abbreviation and full name) available in the HRRR v4 output and HRRR-Zarr files is located in Appendix B of this manuscript. Users can download the specific files of interest to them by accessing them by full name using web tools or from within programs in Python or other languages.

3.4 HRRR-Zarr Applications and Discussion

The HRRR-Zarr archive was developed with the intention of expanding its utility for machine learning and other applications that require high velocity file throughput. While demonstrating a full machine learning scenario is outside the scope here, this section illustrates examples of situations where the Zarr file format may be optimal in terms of efficiency and ease of use. We will use a high-impact meteorological event from September 2020 to showcase the utility of model data in Zarr format for not only research applications, but operational decision making and forecasting use cases as well. This section of the paper will be comprised of subsections that detail the event we are analyzing, followed by example use cases for HRRR model output in Zarr format.

3.4.1 Overview of the Labor Day Weather Event

(7 – 9 September 2020)

In the days leading up to the historic 2020 Labor Day weather event, forecasters in the western half of the United States were on high alert as Typhoon Julian began recurving poleward towards the midlatitudes, a phenomenon called extratropical transition. When extratropical transition occurs, a tropical cyclone frequently interacts with the midlatitude flow, which can amplify the ridge-wave patterns and produce high-impact weather downstream (Bosart and Carr 1978; Cordeira et al. 2013; Feser et al. 2015; Keller et al. 2019). In this case, Typhoon Julian did modify the midlatitude wave pattern by amplifying both the anticyclone over the Gulf of Alaska and the midlatitude cyclone situated over Western Canada. The rapid intensification of this ridge-trough pair ultimately produced far-reaching effects from historic windstorms and unrelenting

wildfire spread (Fig. 3.5) in the Pacific Northwest, to strong downslope winds in Utah and a snowstorm in Colorado. We focus here on the data and model forecasts pertaining to the events that occurred in Oregon, west of the Cascade Mountains.

The Labor Day weather event was synoptically-driven and thus was well-forecasted several days in advance. Prior to the trough arrival and onset of the downslope windstorm, the Pacific Northwest was experiencing extreme fire danger due to warm and dry conditions, with several fires already burning in Washington and Oregon. By 12 UTC on September 7, a thermal trough was situated over coastal Oregon with a tightening pressure gradient orthogonal to it. These conditions are indicative of impending strong northeast and easterly winds in western Oregon. As forecasted, strong easterly winds arrived on the western side of the Oregon Cascades by 00 UTC on September 8. In a near worst case scenario, wind gusts along the western slopes of the Oregon Cascades ignited new fires (Riverside Fire) and significantly intensified existing wildfires (Beechie Creek). For nearly a week after the onset of the downslope winds, persistent easterly flow propagated wildfire smoke west, resulting in historic PM2.5 measurements in excess of 500 micrograms per cubic meter in Portland, Salem, and Eugene, OR (Green 2020). Suppression efforts were minimal given the steep terrain surrounding the wildfires, making it too dangerous for fire crews to extinguish them safely. Ultimately, the Riverside and Beechie Creek fires burned over 332,000 acres of land. The following subsections will use the data from this high-stakes weather event to illustrate use cases for future machine learning applications for research, forecasting, and beyond.

3.4.2 Forecast Time Series for a Point

Time series are one of the most straightforward and widely understood visualizations used to show how a given parameter evolves over a period. Scientists and consumers alike are exposed to time series every day when looking at stock market trends, weather forecasts, and health tracking applications. Despite their inherent simplicity, requiring only time and a dependent variable as input, they can be time consuming and challenging to create when starting from data files that represent a single time in space for millions of locations, as is the case with NWP model output in GRIB2 format. As discussed earlier, a HRRR GRIB2 file of size O(100 MB) contains hundreds of two-dimensional forecast fields for a single valid time. Retrieving, storing, and unpacking 18, 36, or 48 such files up to 24 times a day is beyond what many users want to deal with in terms of compute power and storage.

Efficient access to model output in the form of a time series was a cornerstone in determining the structure and organization of the HRRR-Zarr format. While identical time series can be constructed from both GRIB2 and Zarr file formats, the process and requirements are quite different. As discussed in the Data and Methods section, the tiny two-dimensional analysis HRRR-Zarr files can be easily accessed to estimate prior conditions at a location while the three-dimensional forecast HRRR-Zarr files contain all forecast hours from a model run to assess how future conditions at that location may unfold.

To illustrate the utility of the Zarr format for this use case, we plot time series of forecast wind gusts from the 00, 06, 12, and 18 UTC HRRR model runs for a single point from 12 UTC 6 Sept-18 UTC 8 Sept 2020 (Fig. 3.6). For this case, we chose the HRRR

grid point nearest to the Horse Creek (Station ID: HSFO3) Remote Automated Weather Station (RAWS), located downwind of the Beechie Creek Fire. To create this visualization of forecasted wind gust, 10 small chunks of data (one from each model run) totaling ~ 10 MB were retrieved from the hrrrzarr bucket to obtain all necessary model output. In contrast, 360 GRIB2 files totaling ~54 GB might be downloaded to replicate this process or else values within byte ranges in each of those files would need to be determined and accessed. The single access point to all forecast hours in a model run can reduce processing time and optimize workflows for applications such as creating a training dataset for a ML model.

Plotting sequential model runs creates a time-lagged ensemble (TLE), showing an estimate of the uncertainty over a span of forecasts for a given valid time. Ensemble forecast systems more generally are comprised of multiple model runs (ensemble members) that differ in their initial conditions, parameterizations, or dynamical cores, and when considered simultaneously, can offer probabilistic forecast guidance. While not an ensemble system in the classical sense, a TLE from HRRR output can provide useful diagnostics for evaluating the uncertainty or spread in values among recent forecasts from a model that would otherwise provide only deterministic guidance (Xu et al. 2019). A TLE can be constructed using a set of sequential HRRR forecasts, with each model run treated as an ensemble member. In this case, we use forecasts from all model runs initialized from 12 UTC 5 September – 06 UTC 9 September to calculate statistics at valid times 00 UTC 7 September – 12 UTC 9 September. Diagnostic values such as mean, minimum, and maximum forecasted wind gust provide a simple evaluation of the model's uncertainty leading up to an event (Fig. 3.7). In this specific case, the high wind

gusts forecasted by the HRRR were observed by only one observation at the RAWS site, in part due to its location within a heavily forested region (not shown). These same statistical evaluations would be similarly powerful for comparison to output from other modeling systems.

3.4.3 Spatial Analysis of Forecast Data

Many operational or research applications of HRRR model output require only a fraction of the 1.9 million grid points in the HRRR CONUS domain, requiring the user to implement methods to subset areas of interest from the complete grids. The 450 x 450 km area within each HRRR-Zarr chunk is an approach to simplify that process for many local- to regional-scale applications. Adjacent chunks can be stitched together to evaluate conditions within larger contiguous areas or chunks from different portions of the HRRR grid can be accessed and processed as needed.

To demonstrate the utility of the HRRR-Zarr chunk structure, we illustrate a potential use case for the HRRR-Zarr analysis data. Model analyses are often used as proxies for observations, especially in areas of complex terrain where in-situ measurements may not be available or not necessarily representative of prevailing conditions in all situations (e.g., Fig. 3.7). In this case, we use the HRRR-Zarr analysis files to determine the onset time of wind gusts exceeding 10 m/s for every point within the chunk (Fig. 3.8). This wind gust threshold was chosen based on criteria commonly used for red flag warnings issued by the National Weather Service. The filled contours in Figure 8 depict the approximate onset time of the downslope windstorm event across western Oregon, with the event beginning along the highest reaches of the Cascade

Range and then progressing westward later. Such diagnostics can then be related to available wind observations and damage reports to help evaluate the ability of the HRRR model to forecast the temporal evolution of the event.

Building on the TLE concept available from consecutive HRRR forecasts, we calculate the probability of a HRRR wind gust forecast exceeding 10 m/s at a given time during the downslope windstorm for all grid points within the Western Oregon chunk. The 21 model runs (F01-F18, F24, F30, and F36) available from HRRRv3 forecasts valid at 00 and 06 UTC 8 September 2020 are used to calculate the fraction of wind gusts forecasts exceeding that threshold in this subregion (Figure 9). Using such probabilistic guidance as the event developed, forecasters might have higher confidence that the HRRR model forecasts issued earlier are being confirmed by more recent forecasts as the downslope winds continued. For a single valid time, this probability metric utilized wind gust from 20 HRRR-Zarr files, which required less than 20 MB of storage capacity, an amount easily manageable in computer memory. Actual forecast applications might limit the TLE members to those available at least 12 hours in advance, e.g., forecasts with lead times from F12-F18 and those available every 6 h out to 48 h from the HRRRv4 model output now available.

3.4.4 Empirical Cumulative Distributions

Empirical cumulative distributions of model data and observations are often utilized to better understand the range of possible values for a given parameter as a function of time and/or location and can be used to correct for model biases (Blaylock et al. 2018, Gowan and Horel 2020). If enough data are available over an adequate period of

time, these cumulative distributions can be thought of as a climatology and used for comparison to a parameter at an equivalent time or location in order to recognize conditions that are likely anomalous. Creating distributions from observations or model output typically requires data from thousands of input times and files for the information to be considered useful. This can be a daunting and time-consuming task since a large amount of storage and compute power are needed to efficiently process thousands of GRIB2 data files.

Blaylock et al. (2018) presented an approach to compute empirical cumulative distributions of HRRR model output at all 1.9 million gridpoints that required harnessing the power of the Open Science Grid (OSG). The OSG allows users to send jobs that are repetitive in nature (e.g., statistical calculations using large datasets, data mining, etc.) to unused or idle computing resources at hundreds of locations within the OSG consortium, reducing the overall processing time for a given workflow. The OSG method enables large amounts of data to be simultaneously processed, but its complexities can be a drawback for most users without a thorough understanding of the system. Continually updating cumulative distributions using this approach is also difficult to sustain.

A quick and efficient method is illustrated here to generate empirical cumulative distributions of atmospheric parameters from the HRRR-Zarr archive. To assess the anomalous nature of the downslope wind event in September 2020 in northern Oregon, we generated cumulative distributions of wind gust data for each grid point in that region by accessing all HRRR hourly analyses during the month of September during the preceding years 2016-2019. Each grid point's cumulative distribution is derived then from 2,880 wind gust values, one from every hourly analysis during the four calendar

months of HRRR model output. A range of percentiles can be derived from the empirical distributions to estimate what are normal and above normal wind gusts in this area during the month of September.

As expected, the highest wind gusts evident from the 95th percentile values during this month tend to occur over the Cascade Range and offshore (Fig. 3.10). Using this four-year distribution, we then compare the 95th percentile values to the analysis and F06, F12, and F18 forecasts valid at 06 UTC on 8 September 2020 (Fig. 3.11). To emphasize the severity of the event across the region, the excess magnitude of wind gusts values above the 95th percentile are shown. Comparing these forecasts and analysis to the cumulative distribution is a simple way to show how anomalous this event was, with wind gusts exceeding the 95th percentile values by 15-30 m/s over the Cascade and portions of the Coast Ranges and extending into sections of the Willamette Valley.

The empirical distributions computed using 4 months of data for a single variable and chunk required less than a minute on a typical workstation. We compare this to the method used by Blaylock et al. (2018), which calculated empirical cumulative distributions for all HRRR model grid points. These distributions were then used to output wind speed values at 19 percentiles at all HRRR grid points for each day of the year. As previously stated, this was a rigorous and time-intensive endeavor that required an enormous amount of model output. Ultimately, calculating these distributions resulted in the need to store 6,935 additional files containing the percentiles at each of the 1.9 million HRRR grid points.

Calculating empirical cumulative distributions, as well as other large-scale statistical metrics, with data in Zarr format gives the end user the ability to continually

update their statistics as new information is received. This method especially benefits users who are interested in time-sensitive datasets, like those from numerical weather prediction models. Using the HRRR-Zarr method, a user will be able to efficiently compute statistics that are tailored to a specific application or workflow, without dealing with the overhead of many GB of excess data.

3.5 Conclusions

The global weather enterprise relies on the vast amount of output produced by numerical weather prediction models each day for applications ranging from operational forecasting to research and machine learning. As advancements in technology allow for higher-resolution and more frequent weather model output, users are struggling to keep up, and may only have access to a small fraction of the data produced. Much of this model output is in available in GRIB2 format, a file type that efficiently stores hundreds of two-dimensional variable fields for a single valid time. Despite the highly compressible nature of GRIB2 files, they are often on the order of several hundred MB each, making high-volume input/output applications challenging due to the memory and compute resources needed to parse these files.

We present a solution to the GRIB2 bottleneck problem by illustrating the conversion to the N-dimensional, cloud compatible Zarr file format of the current AWS S3 storage of HRRR GRIB2 files made possible by the Amazon Sustainability Data Initiative and the NOAA Big Data Program. Our supplementary S3 bucket, hrrrzarr, is publicly accessible as part of the Amazon Initiative and is composed of sets for each

model run of analysis (F00) and forecast files for 173 surface fields sectioned into 96 small chunks.

The structure of the HRRR-Zarr files was designed to allow users the flexibility to access only the data they need through selecting subdomains and parameters of interest without the overhead that comes from accessing numerous GRIB2 files. Users may retrieve the analysis files needed to diagnose prior conditions or retrieve the forecast files in combination with the analysis files to evaluate future conditions or validate prior forecasts.

Using a high-impact weather event from September 2020, we present workflow examples for analyzing large amounts of sensible weather parameters from the HRRR-Zarr data archive. Applications of the HRRR-Zarr data format are illustrated utilizing a few examples: assembling time series for a specific grid point of forecast conditions over a range of model runs; examining similarities and differences among samples of model forecasts for the same valid times from successive model runs; calculating empirical cumulative distributions over multiyear periods; and detecting forecasts of extreme conditions relative to conditions during other recent years. The small, compressed chunks of data are ideal for high-throughput workflows where minimizing processing time or accessing files corresponding to many different valid times is critical. However, GRIB2 HRRR files remain the best option for applications, such as initializing high resolution model simulations, that require many variables at multiple levels for a limited set of valid times.

The HRRR-Zarr archive is a supplementary resource for users who need data found in the HRRR GRIB2 files. Future dataset conversions of HRRR and other model

data to Zarr are likely given its structural flexibility and compatibility with cloud platforms. For example, GRIB2 files from the NOAA Global Ensemble Forecast System (GEFS) Forecasts and Re-Forecasts are available as part of the NOAA Open Data Program and the Amazon Sustainability Data Initiative (https://registry.opendata.aws/noaa-gefs/; https://registry.opendata.aws/noaa-gefs-reforecast/). Four-dimensional zarr data files for sensible weather parameters of high interest would likely be of high interest where the additional dimension corresponds to ensemble member from the GEFS. Establishing conventions to aid creation and exchange of Zarr files will likely be a useful future step as the variety of options available to create the files might lead to unnecessary complexity for end users. The Zarr format may help expand utilization of numerical weather prediction model output for diverse applications of data scientists in many fields.

Table 3.1: Selected Characteristics of HRRR CONUS Versions from 2014-Present

HRRR CONUS		Forecast Length for Initialization Times		Number of Output Files		
v.	First Date	0, 6, 12, 18 UTC	Other Hours	Surface	Pressure	Native
1	9/30/2014	15	15	102	659	778
2	8/23/2016	18	18	135	687	1110
3	7/12/2018	36	18	151	701	1126
4	12/2/2020	48	18	173	711	1136

Table 3.2: Python libraries recommended for HRRR-Zarr file creation and retrieval.

Python Library	HRRR-Zarr Creation	HRRR-Zarr Retrieval
Python 3	X	X
Zarr	X	X
Eccodes	X	X
Numpy		X
Xarray	X	X
Requests	X	
Iris	X	
Iris-Grib	X	
Boto3	X	X

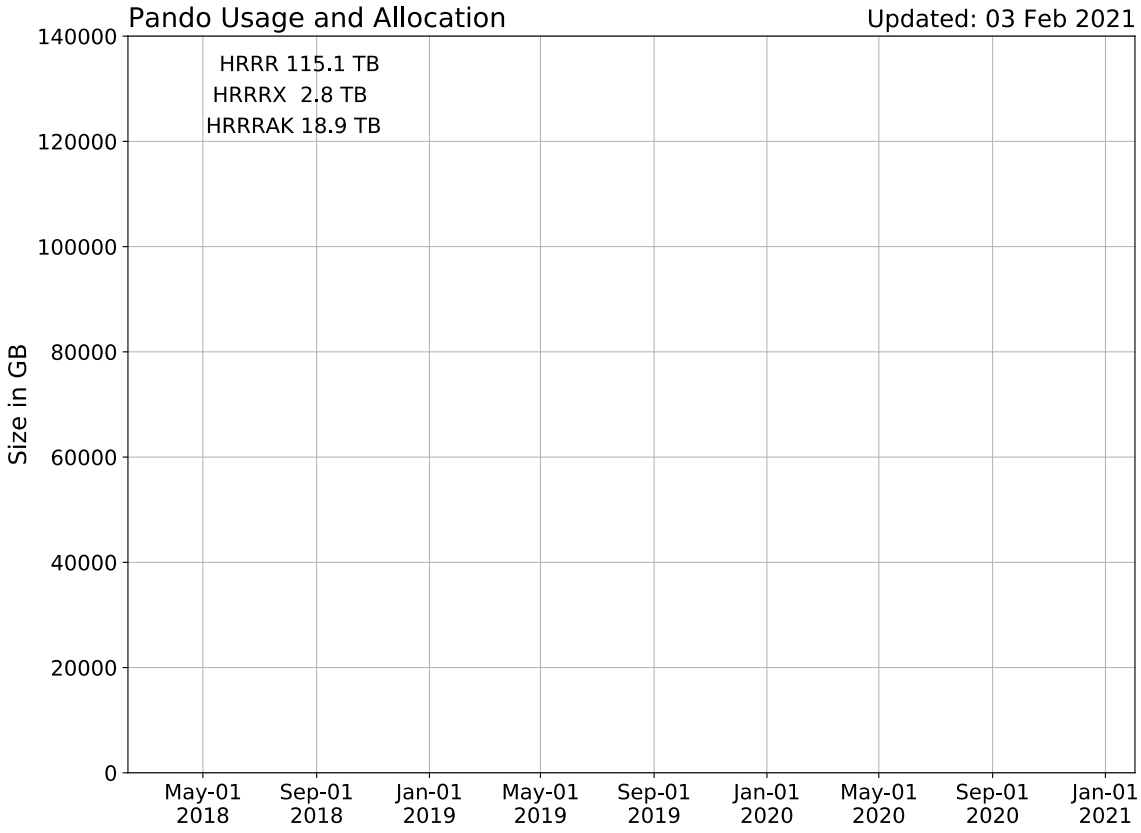

Figure 3.1: Storage growth on University of Utah Pando system during the past three years of: operational CONUS HRRR (Red), operational HRRR Alaska (Blue), experimental HRRR analyses (Gold).

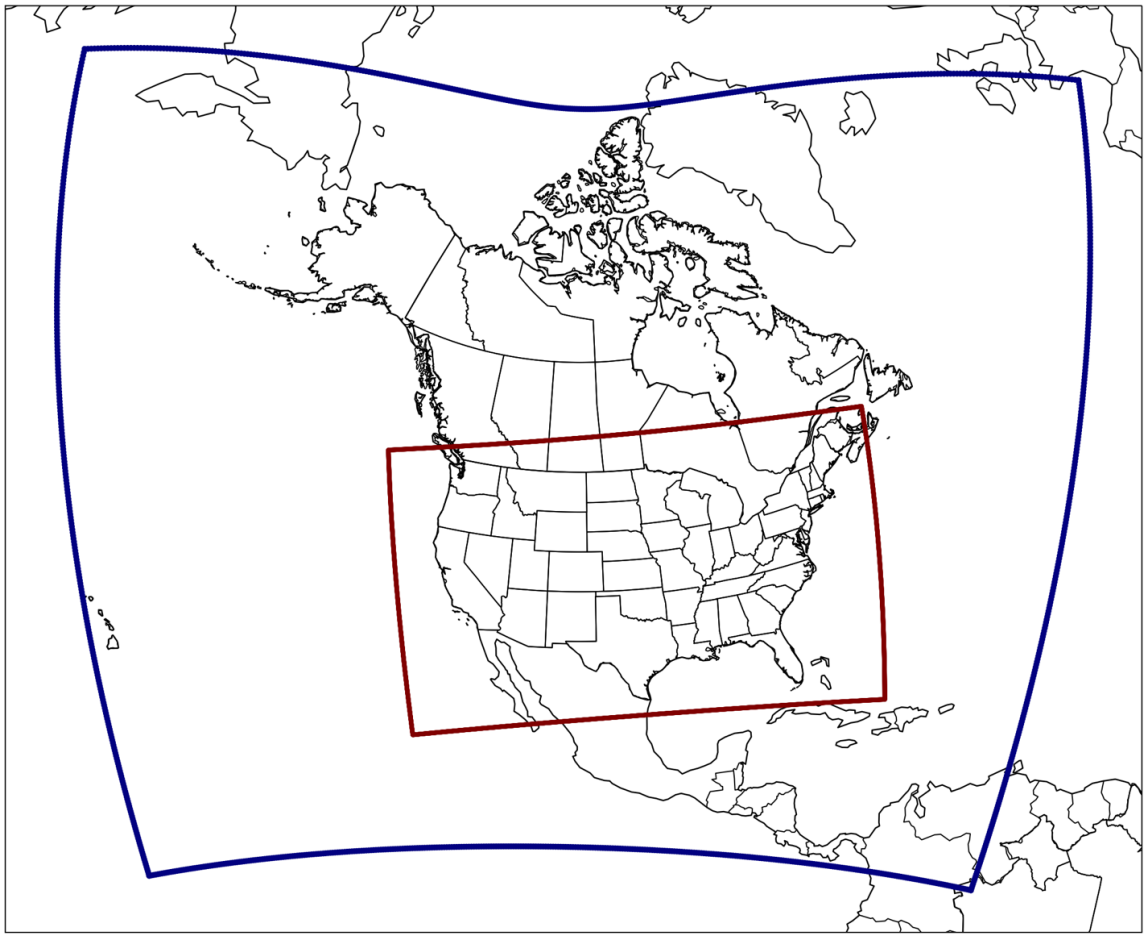

Figure 3.2: CONUS HRRR domain (red) within the Rapid Refresh domain (blue) that supplies initial and boundary conditions to the HRRR

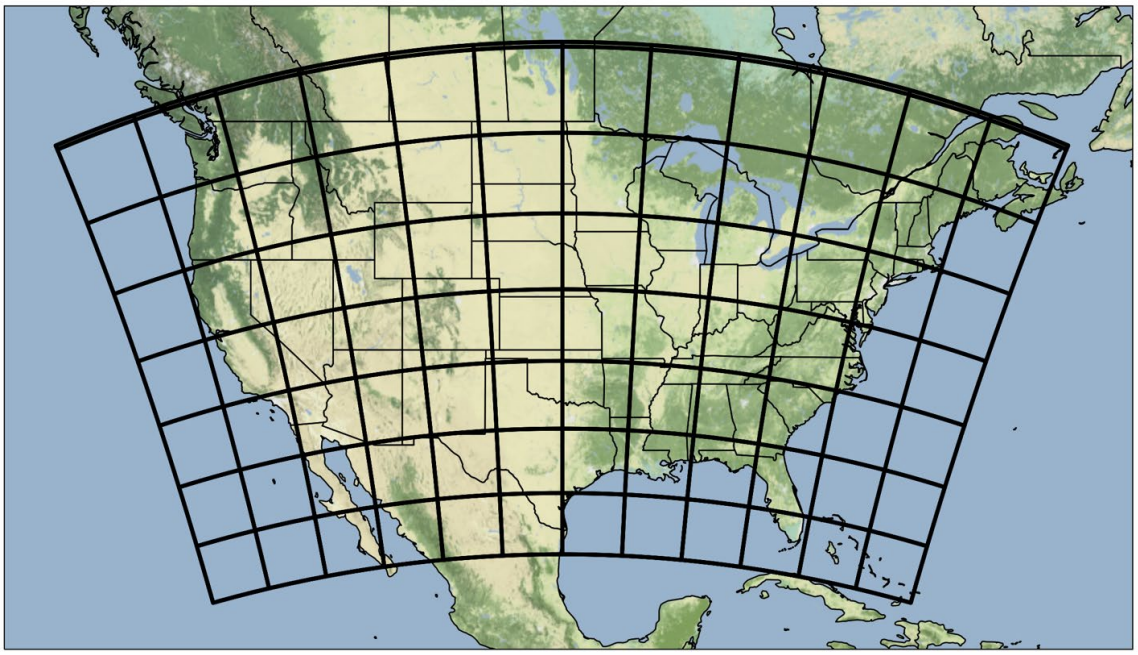

Figure 3.3: HRRR Domain divided into 96 equal chunks of size 150 x 150 grid points. The 12 chunks on the top row contain only 9 rows of non-NaN data.

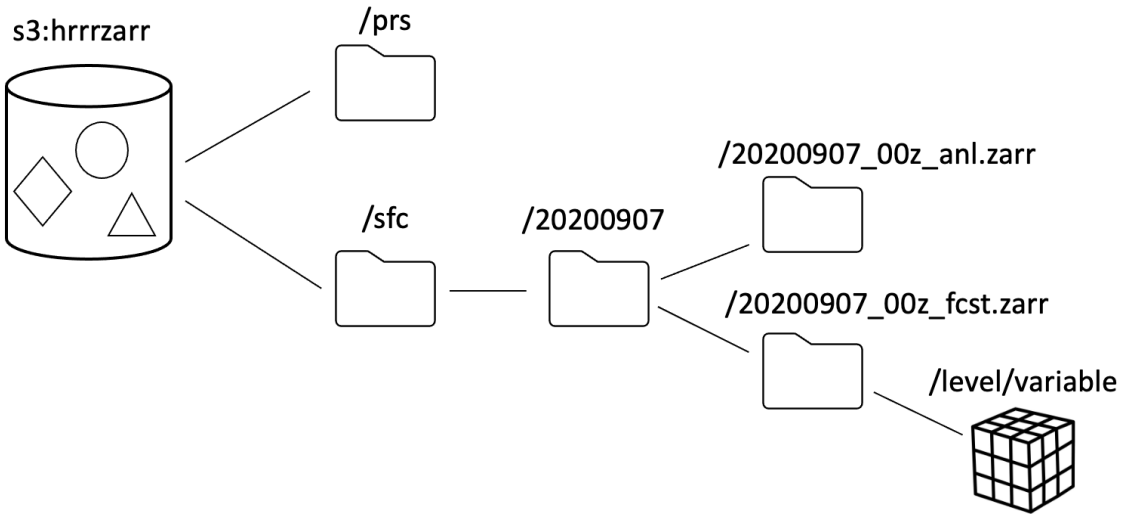

Figure 3.4: Files within the AWS S3 bucket hrrrzarr are named to emulate a hierarchical data structure.

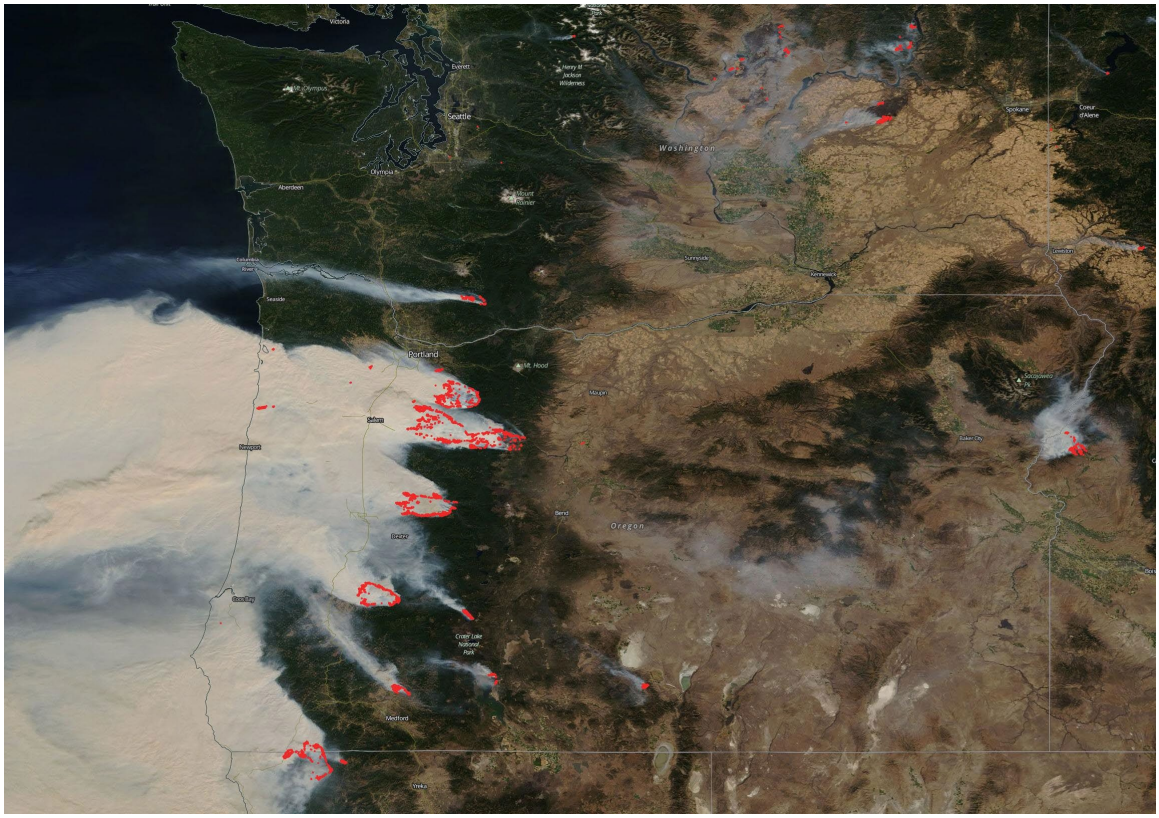

Figure 3.5: Boundaries of active fires (red outlines), estimated using VIIRS 375 m thermal anomalies, and smoke from wildfires in the Pacific Northwest on 9 September 2020 (source: https://worldview.earthdata.nasa.gov/).

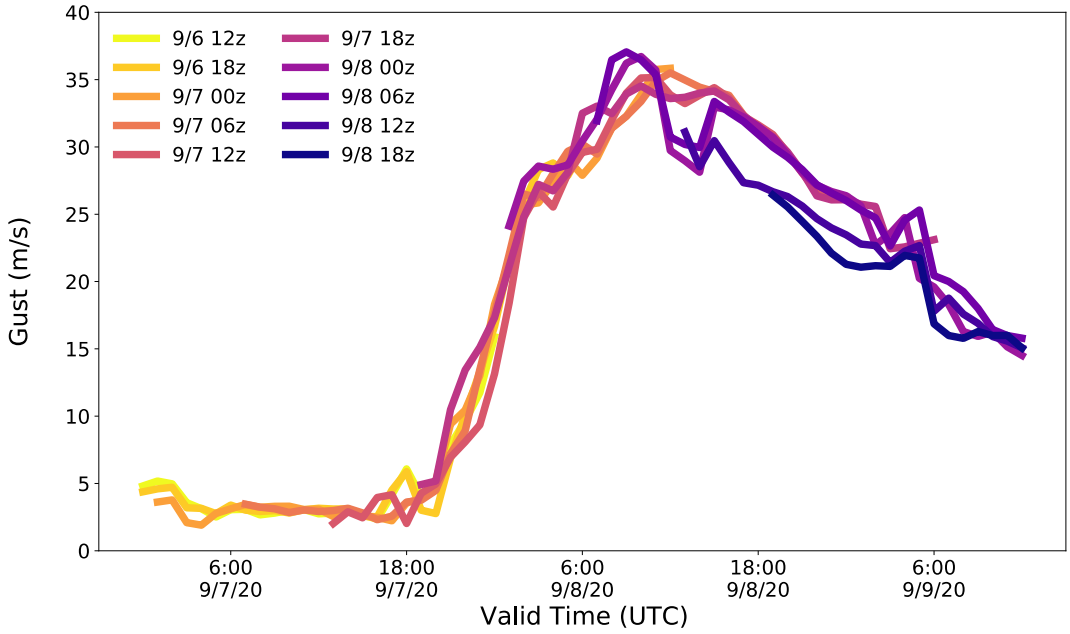

Figure 3.6: HRRR wind gust (m s^{-1}) forecasts near the Beechie Creek fire and the Horse Creek (HSFO3) RAWS. Lines are colored corresponding to model initialization time.

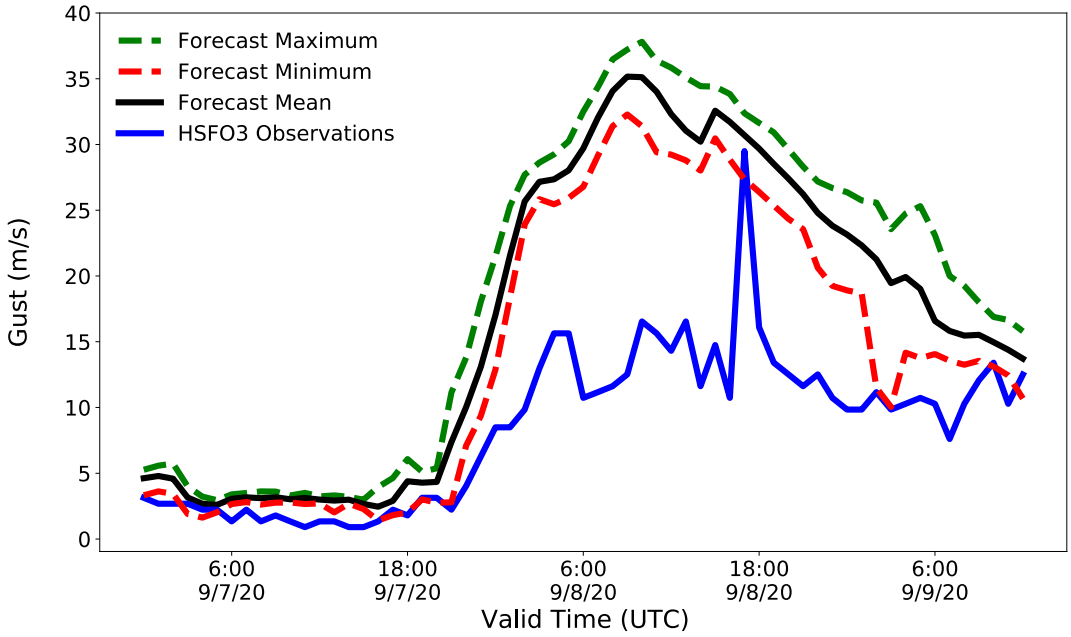

Figure 3.7: Minimum, mean, and maximum HRRR wind gust (m s^{-1}) forecasts near the Beechie Creek fire and the Horse Creek (HSFO3) RAWS observations.

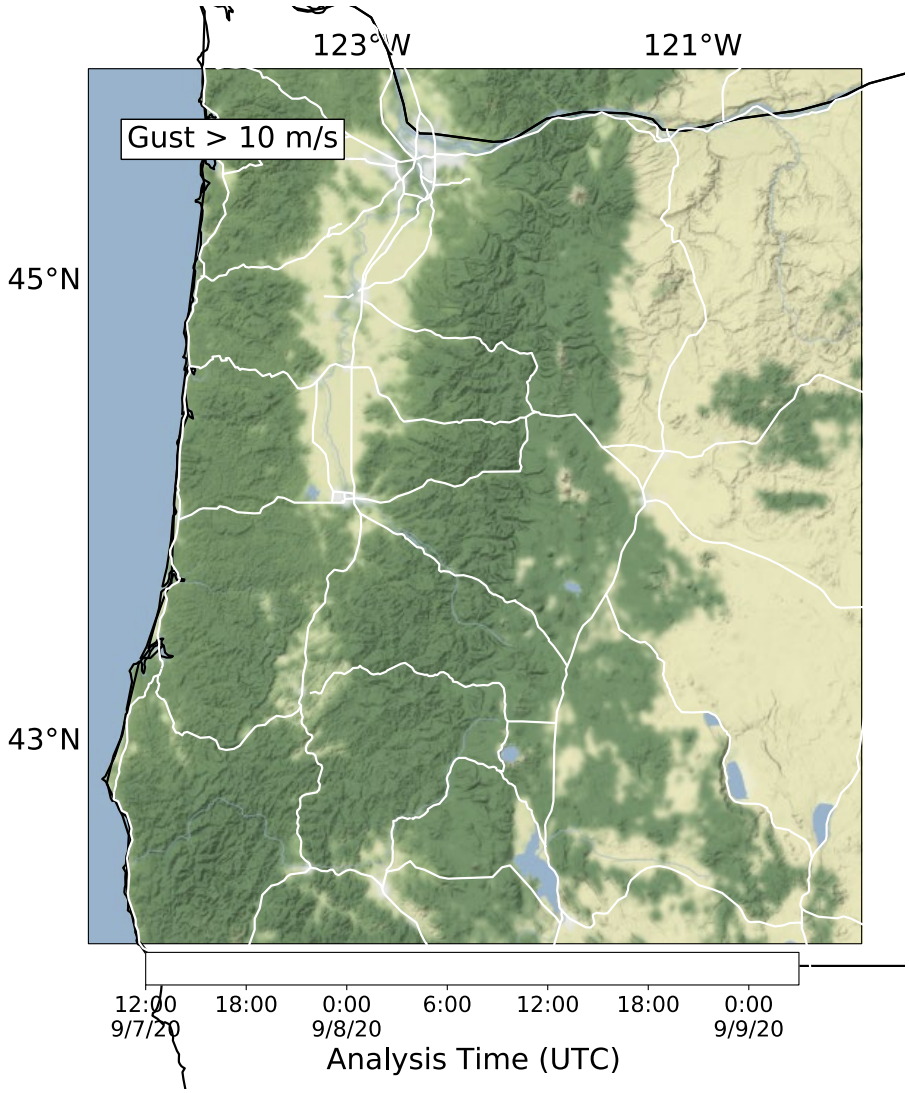

Figure 3.8: Time of first HRRR analysis (F00) with a wind gust exceeding 10 m/s at each grid point. Contours are shaded according to scale and consider model runs initialized between 12 UTC 7 September – 03 UTC 9 September.

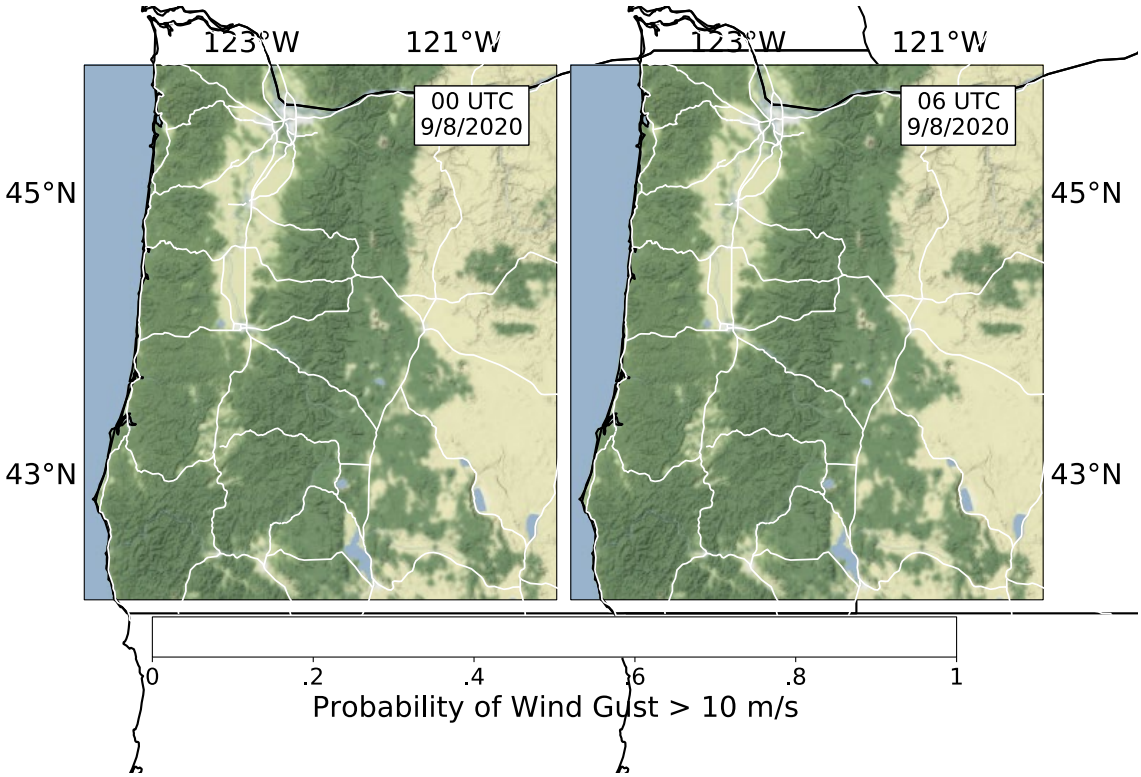

Figure 3.9: Fraction of HRRR wind gust forecasts exceeding 10 m/s at valid times 00 UTC (left) and 06 UTC (right) on 8 September 2020. Contours correspond to probability values (0-1) and are shaded according to the scale below the figure.

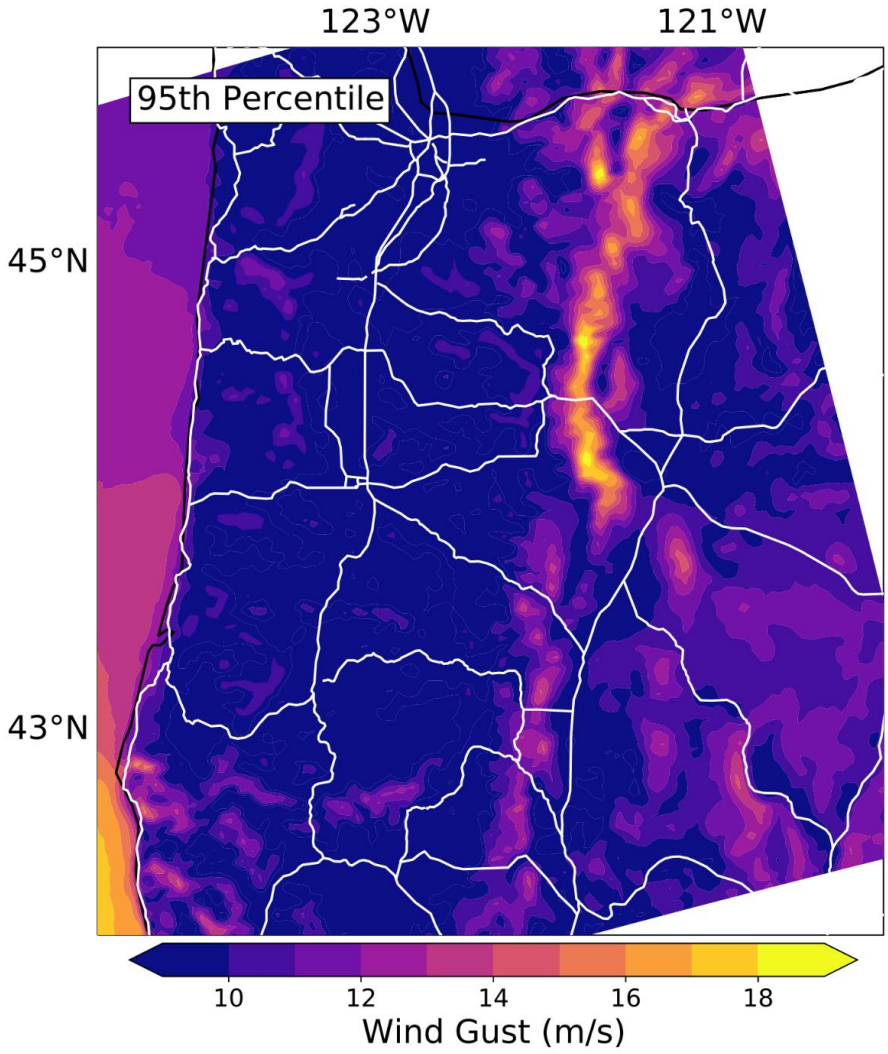

Figure 3.10: 95th percentile wind gust values (shaded according to scale below figure) calculated at each grid point from empirical cumulative distributions derived from HRRR analyses valid during the month of September during the four preceding years, 2016-2019.

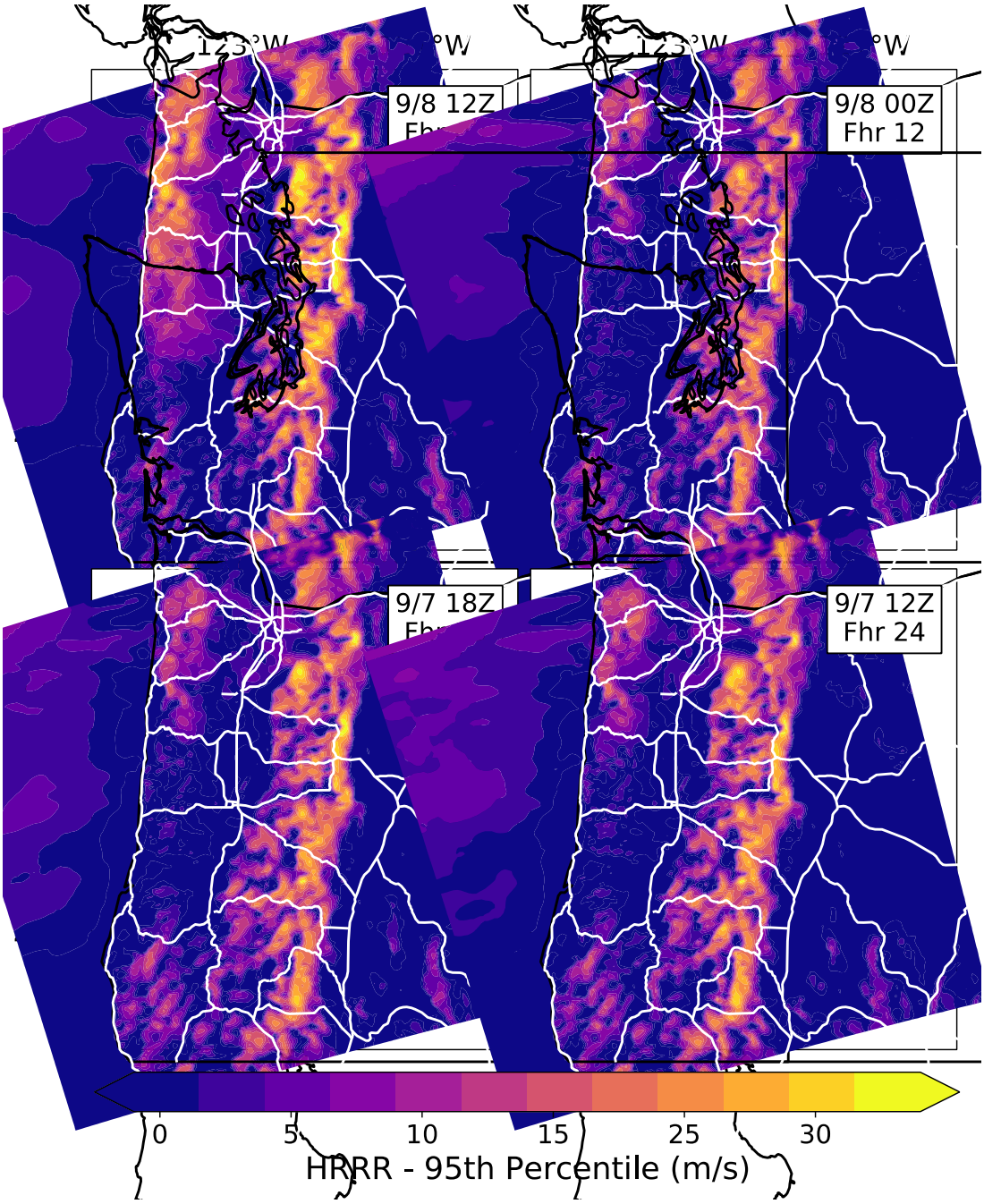

Figure 3.11: Wind speed (m/s; shaded according to scale) exceeding 95th percentile wind gust values at 12 UTC 8 September 2020. Subplots correspond to the verifying analysis (upper left) and F12, F18, and F24 forecasts valid at that time.

3.6 Appendix: HRRRv4 Variable List

#	Parameter Long Name and Units	Vertical Level	Parameter Short Name
1	Composite Reflectivity (dB)	Entire Atmosphere	REFC
2	Echo Top (m)	Cloud Top	ETOP
3	Radar-based Vertically Integrated Liquid (kg/m^2)	Entire Atmosphere	RADARVIL
4	Vertically Integrated Liquid (kg/m^2)	Entire Atmosphere	VIL
5	Visibility (m)	Surface	VIS
6	Reflectivity (dB)	1000-m Above Ground	REFD
7	Reflectivity (dB)	4000-m Above Ground	REFD
8	Reflectivity (dB)	263 K Level	REFD
9	Wind Gust ($m\ s^{-1}$)	Surface	GUST
10	U-Component of Wind ($m\ s^{-1}$)	250 mb	UGRD
11	V-Component of Wind ($m\ s^{-1}$)	250 mb	VGRD
12	U-Component of Wind ($m\ s^{-1}$)	300 mb	UGRD

13	V-Component of Wind ($m\ s^{-1}$)	300 mb	VGRD
14	Geopotential Height (gpm)	500 mb	HGT
15	Temperature (K)	500 mb	TMP
16	Dew Point Temperature (K)	500 mb	DPT
17	U-Component of Wind ($m\ s^{-1}$)	500 mb	UGRD
18	V-Component of Wind ($m\ s^{-1}$)	500 mb	VGRD
19	Geopotential Height (gpm)	700 mb	HGT
20	Temperature (K)	700 mb	TMP
21	Dew Point Temperature (K)	700 mb	DPT
22	Vertical Velocity ($m\ s^{-1}$)	700 mb	DZDT
23	U-Component of Wind ($m\ s^{-1}$)	700 mb	UGRD
24	V-Component of Wind ($m\ s^{-1}$)	700 mb	VGRD
25	Geopotential Height (gpm)	850 mb	HGT
26	Temperature (K)	850 mb	TMP

27	Dew Point Temperature (K)	850 mb	DPT
28	U-Component of Wind (m s^{-1})	850 mb	UGRD
29	V-Component of Wind (m s^{-1})	850 mb	VGRD
30	Temperature (K)	925 mb	TMP
31	Dew Point Temperature (K)	925 mb	DPT
32	U-Component of Wind (m s^{-1})	925 mb	UGRD
33	V-Component of Wind (m s^{-1})	925 mb	VGRD
34	Temperature (K)	1000 mb	TMP
35	Dew Point Temperature (K)	1000 mb	DPT
36	U-Component of Wind (m s^{-1})	1000 mb	UGRD
37	V-Component of Wind (m s^{-1})	1000 mb	VGRD
38	Hourly Maximum of Upward Vertical Velocity (m s^{-1})	100-1000 mb Above Ground	MAXUVV
39	Hourly Maximum of Downward Vertical Velocity (m s^{-1})	100-1000 mb Above Ground	MAXDVV
40	Vertical Velocity (m s^{-1})	0.5-0.8 Sigma Layer	DZDT

41	Mean Sea Level Pressure (Pa)	Mean Sea Level	MSLP
42	Geopotential Height (gpm)	1000 mb	HGT
43	Hourly Maximum of Simulated Reflectivity at 1 km AGL (dB)	1000 m Above Ground	MAXREF
44	Reflectivity (dB)	263 K Level	REFD
45	Hourly Maximum of Updraft Helicity ($m^2 s^{-2}$)	5000-2000 m Above Ground	MXUPHL
46	Hourly Minimum of Updraft Helicity ($m^2 s^{-2}$)	5000-2000 m Above Ground	MNUPHL
47	Hourly Maximum of Updraft Helicity ($m^2 s^{-2}$)	2000-0 m Above Ground	MXUPHL
48	Hourly Minimum of Updraft Helicity ($m^2 s^{-2}$)	2000-0 m Above Ground	MNUPHL
49	Hourly Maximum of Updraft Helicity ($m^2 s^{-2}$)	3000-0 m Above Ground	MXUPHL
50	Hourly Minimum of Updraft Helicity ($m^2 s^{-2}$)	3000-0 m Above Ground	MNUPHL
51	Relative Vorticity (s^{-1})	2000-0 m Above Ground	RELV
52	Relative Vorticity (s^{-1})	1000-0 m Above Ground	RELV
53	Hail (m)	Entire Atmosphere	HAIL
54	Hail (m)	0.1 Sigma Layer	HAIL

55	Hail (m)	Surface	HAIL
56	Total Column Integrated Graupel (kg m-^{2})	Entire Atmosphere Single Layer	TCOLG
57	Lightning Potential Index (J kg^{-1})	1-m Above Ground	LTPINX
58	Lightning Potential Index (J kg^{-1})	2-m Above Ground	LTPINX
59	Lightning (Non-Dim)	Entire Atmosphere	LTNG
60	U-Component of Wind (m s^{-1})	80-m Above Ground	UGRD
61	V-Component of Wind (m s^{-1})	80-m Above Ground	VGRD
62	Pressure (Pa)	Surface	PRES
63	Geopotential Height (gpm)	Surface	HGT
64	Temperature (K)	Surface	TMP
65	Total Snowfall (m)	Surface	ASNOW
66	Moisture Availability (%)	0-m Underground	MSTAV
67	Plant Canopy Surface Water (kg m-^{2})	Surface	CNWAT
68	Water Equivalent of Accumulated Snow Depth (kg m-^{2})	Surface	WEASD

69	Snow Cover (%)	Surface	SNOWC
70	Snow Depth (%)	Surface	SNOD
71	Temperature (K)	2-m Above Ground	TMP
72	Potential Temperature (K)	2-m Above Ground	POT
73	Specific Humidity (kg/kg)	2-m Above Ground	SPFH
74	Dew Point Temperature (K)	2-m Above Ground	DPT
75	Relative Humidity (kg/kg)	2-m Above Ground	RH
76	Mass Density (Concentration; ($kg\ m^{-3}$))	8-m Above Ground	MASSDEN
77	U-Component of Wind ($m\ s^{-1}$)	10-m Above Ground	UGRD
78	V-Component of Wind ($m\ s^{-1}$)	10-m Above Ground	VGRD
79	Wind Speed ($m\ s^{-1}$)	10-m Above Ground	WIND
80	U-Component of Hourly Maximum Wind ($m\ s^{-1}$)	10-m Above Ground	MAXUW
81	V-Component of Hourly Maximum Wind ($m\ s^{-1}$)	10-m Above Ground	MAXVW
82	Percent Frozen Precipitation (%)	Surface	CPOFP

83	Precipitation Rate (kg $m^{-2}s^{-1}$)	Surface	PRATE
84	Total Precipitation (kg m^{-2})	Surface (UTC day)	APCP
85	Water Equivalent of Accumulated Snow Depth (kg m^{-2})	Surface (UTC day)	WEASD
86	Frozen Rain (kg m^{-2})	Surface (UTC day)	FROZR
87	Freezing Rain (kg m^{-2})	Surface (UTC day)	FRZR
88	Storm Surface Runoff (kg m^{-2})	Surface	SSRUN
89	Baseflow-Groundwater (kg m^{-2})	Surface	BGRUN
90	Total Precipitation (kg m^{-2})	Surface	APCP
91	Water Equivalent of Accumulated Snow Depth (kg m^{-2})	Surface	WEASD
92	Frozen Rain (kg m^{-2})	Surface	FROZR
93	Categorical Snow	Surface	CSNOW
94	Categorical Ice Pellets	Surface	CICEP
95	Categorical Freezing Rain	Surface	CFRZR
96	Categorical Rain	Surface	CRAIN

97	Surface Roughness (m)	Surface	SFCR
98	Frictional Velocity (m s^{-1})	Surface	FRICV
99	Sensible Heat Net Flux (W m^{-2})	Surface	SHTFL
100	Latent Heat Net Flux (W m^{-2})	Surface	LHTFL
101	Vegetation (%)	Surface	VEG
102	Annual Maximum Vegetation Fraction	Surface	MXVEG
103	Annual Minimum Vegetation Fraction	Surface	MNVEG
104	Leaf Area Index	Surface	LAI
105	Ground Heat Flux (W m^{-2})	Surface	GFLUX
106	Vegetation Type	Surface	VGTP
107	Surface Lifted Index (K)	500-1000 mb	LFTX
108	Convective Available Potential Energy (J kg^{-1})	Surface	CAPE
109	Convective Inhibition (J kg^{-1})	Surface	CIN
110	Precipitable Water (kg m^{-2})	Entire Atmosphere Single Layer	PWAT

111	Aeorosol Optical Thickness	Entire Atmosphere Single Layer	AOTK
112	Total Column-Integrated Mass Density (kg m^{-2})	Entire Atmosphere Single Layer	COLMD
113	Total Column-Integrated Cloud Ice (kg m^{-2})	Entire Atmosphere Single Layer	TCOLWold
114	Total Column-Integrated Mass Density (kg m^{-2})	Entire Atmosphere Single Layer	COLMDold
115	Total Cloud Cover (%)	Boundary Layer Cloud Cover	TCDC
116	Low Cloud Cover (%)	Low Cloud Layer	LCDC
117	Middle Cloud Cover (%)	Middle Cloud Layer	MCDC
118	High Cloud Cover (%)	High Cloud Layer	HCDC
119	Total Cloud Cover (%)	Entire Atmosphere	TCDC
120	Geopotential Height (gpm)	Cloud Ceiling	HGT
121	Geopotential Height (gpm)	Cloud Base	HGT
122	Pressure (Pa)	Cloud Base	PRES
123	Pressure (Pa)	Cloud Top	PRES
124	Geopotential Height (gpm)	Cloud Top	HGT

125	Upward Long-Wave Radiation Flux ($W\ m^{-2}$)	Top of Atmosphere	ULWRF
126	Downward Short-Wave Radiation Flux ($W\ m^{-2}$)	Surface	DSWRF
127	Downward Long-Wave Radiation Flux ($W\ m^{-2}$)	Surface	DLWRF
128	Upward Short-Wave Radiation Flux ($W\ m^{-2}$)	Surface	USWRF
129	Upward Long-Wave Radiation Flux ($W\ m^{-2}$)	Surface	ULWRF
130	Cloud Forcing Net Solar Flux ($W\ m^{-2}$)	Surface	CFNSF
131	Visible Beam Downward Solar Flux ($W\ m^{-2}$)	Surface	VBDSF
132	Visible Diffuse Downward Solar Flux ($W\ m^{-2}$)	Surface	VDDSF
133	Upward Short-Wave Radiation Flux ($W\ m^{-2}$)	Surface	USWRF
134	Storm Relative Helicity ($m^2\ s^{-2}$)	3000-0 m Above Ground	HLCY
135	Storm Relative Helicity ($m^2\ s^{-2}$)	1000-0 m Above Ground	HLCY
136	U-Component Storm Motion ($m\ s^{-1}$)	0-6000 m Above Ground	USTM
137	V-Component Storm Motion ($m\ s^{-1}$)	0-6000 m Above Ground	VSTM
138	Vertical U-Component Shear ($m\ s^{-1}$)	0-1000 m Above Ground	VUCSH

139	Vertical V-Component Shear (m s^{-1})	0-1000 m Above Ground	VVCSH
140	Vertical U-Component Shear (m s^{-1})	0-6000 m Above Ground	VUCSH
141	Vertical V-Component Shear (m s^{-1})	0-6000 m Above Ground	VVCSH
142	Geopotential Height (gpm)	0 Celsius Isotherm	HGT
143	Relative Humidity (%)	0 Celsius Isotherm	RH
144	Pressure (Pa)	0 Celsius Isotherm	PRES
145	Geopotential Height (gpm)	Highest Tropospheric Freezing Level	HGT
146	Relative Humidity (%)	Highest Tropospheric Freezing Level	RH
147	Pressure (Pa)	Highest Tropospheric Freezing Level	PRES
148	Geopotential Height (gpm)	263 K Level	HGT
149	Geopotential Height (gpm)	253 K Level	HGT
150	Best (4 layer) Lifted Index (K)	180-0 mb Above Ground	4LFTX
151	Convective Available Potential Energy (J kg^{-1})	180-0 mb Above Ground	CAPE
152	Convective Inhibition (J kg^{-1})	180-0 mb Above Ground	CIN

153	Planetary Boundary Layer Height	Surface	HPBL
154	Geopotential Height (gpm)	Level of Adiabatic Condensation	HGT
155	Convective Available Potential Energy ($J\ kg^{-1}$)	90-0 mb Above Ground	CAPE
156	Convective Inhibition ($J\ kg^{-1}$)	90-0 mb Above Ground	CIN
157	Convective Available Potential Energy ($J\ kg^{-1}$)	255-0 mb Above Ground	CAPE
158	Convective Inhibition ($J\ kg^{-1}$)	255-0 mb Above Ground	CIN
159	Geopotential Height (gpm)	Equilibrium Level	HGT
160	Pressure of level from which parcel was lifted (Pa)	255-0 mb Above Ground	PLPL
161	Convective Available Potential Energy ($J\ kg^{-1}$)	0-3000 m Above Ground	CAPE
162	Geopotential Height (gpm)	No Level	HGT
163	Effective Storm Relative Helicity ($m^2\ s^{-2}$)	Surface	EFHL
164	Critical Angle (degee)	0-5000 m Above Ground	CANGLE
165	Layer Thickness (m)	261 K – 256 K Level	LAYTH
166	Enhanced Stretching Potential	0-3000 m Above Ground	ESP

167	Relative Humidity with Respect to Precipitable Water (%)	Entire Atmosphere	RHPW
168	Land Cover	Surface	LAND
169	Ice Cover Proportion	Surface	ICEC
170	Simulated Brightness Temperature GOES 12, Channel 3 (K)	Top of Atmosphere	SBT123
171	Simulated Brightness Temperature GOES 12, Channel 4 (K)	Top of Atmosphere	SBT124
172	Simulated Brightness Temperature GOES 11, Channel 3 (K)	Top of Atmosphere	SBT113
173	Simulated Brightness Temperature GOES 11, Channel 4 (K)	Top of Atmosphere	SBT114

3.7 References

Adams-Selin, R. D., and C. L. Ziegler, 2016: Forecasting hail using a one-dimensional hail growth model within WRF. *Mon. Wea. Rev.*, **144**, 4919–4939, doi:10.1175/MWR-D-16-0027.1.

Almeida, S., V. Oliveira, A. Pina, and M. Melle-Franco, 2014: Two High-Performance Alternatives to ZLIB Scientific-Data Compression. *Proc. of 14th International Conference on Computational Science and Its Applications (ICCSA 2014)*, Guimarães, Portugal, 623–638, doi:10.1007/978-3-319-09147-1_45.

Alted, F., 2014: Why Modern CPUs Are Starving and What Can Be Done About It. *Comput. Sci. Eng.*, **12**, 68–71, doi:10.1109/MCSE.2010.51.

Amazon, 2021: Amazon S3: Object storage built to store and retrieve any amount of data from anywhere. Accessed 2 January 2021, https://aws.amazon.com/s3/.

Ansari S., S. Del Greco, E. Kearns, O. Brown, S. Wilkins, M. Ramamurthy, J. Weber, R. May, J. Sundwall, J. Layton, A. Gold, A. Pasch, and V. Lakshmanan, 2018: Unlocking the potential of NEXRAD data through NOAA's Big Data Partnership. *Bull. Amer. Meteor. Soc.*, **99**, 189–204, doi:10.1175/BAMS-D-16-0021.1.

Benjamin, S. G., J. M. Brown, G. Brunet, P. Lynch, K. Saito, and T. W. Schlatter, 2018: *100 Years of Progress in Forecasting and NWP Applications*, *Meteor. Monogr.*, **59**, 13.1–13.67, doi:10.1175/AMSMONOGRAPHS-D-18-0020.1.

—, S. S. Weygandt, J. M. Brown, M. Hu, C. R. Alexander, T. G. Smirnova, J. B. Olson, E. P. James, D. C. Dowell, G. A. Grell, H. Lin, S. E. Peckham, T. L. Smith, W. R. Moninger, J. S. Kenyon, and G. S. Manikin, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh, 2016. *Mon. Wea. Rev.*, **144**, 1669–1694, doi:10.1175/MWR-D-15-0242.1.

Blaylock, B., J. Horel, E. Crosman, 2017: Impact of Lake Breezes on Summer Ozone Concentrations in the Salt Lake Valley. J. Appl. Meteor. Clim.56, 353-370. http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-16-0216.1

Blaylock, B. K., J. D. Horel, and S. T. Liston, 2017: Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output. *Computers and Geosciences*, **109**, 43–50, doi: 10.1016/j.cageo.2017.08.005.

—, —, and C. Galli, 2018: High-Resolution Rapid Refresh Model Data Analytics Derived on the Open Science Grid to Assist Wildland Fire Weather Assessment. *J. Atmos. Oceanic Technol.*, **35**, 2213–2227, doi:10.1175/JTECH-D-18-0073.1.

Bosart, L. F., and F. H. Carr, 1978: A Case Study of Excessive Rainfall Centered Around Wellsville, New York, 20–21 June 1972. *Mon. Wea. Rev.*, **106**,

doi:10.1175/1520-0493(1978)106<0348:ACSOER>2.0.CO;2.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29**, 639–653, doi:10.1175/WAF-D-13-00113.1.

Collett, Y., 2020: LZ4 – Extremely fast compression, version 1.6.2. Accessed 1 January 2021, https://lz4.github.io/lz4/.

Cordeira, J. M., F. M. Ralph, and B. J. Moore, 2013: The development and evolution of two atmospheric rivers in proximity to western North Pacific tropical cyclones in October 2010. *Mon. Wea. Rev.*, **141**, 4234–4255, doi:10.1175/2010MWR2888.1.

Crosman, E., J. Horel, 2017: Large-eddy simulations of a Salt Lake Valley cold-air pool. Atmospheric Research. 193, 10–25, doi:10.1016/j.atmosres.2017.04.010

Delaunay, X., A. Courtois, and F. Gouillon, 2019: Evaluation of lossless and lossy algorithms for the compression of scientific datasets in netCDF-4 or HDF5 files. *Geosci. Model Dev.*, **12**, 4099–4113, doi:10.5194/gmd-12-4099-2019.

Donkers, K., 2020: To the cloud and back again. Medium, accessed 28 December 2020, https://medium.com/informatics-lab/create-zarr-from-pp-files-ffa6b7972d6f.

Donoho, D. L., 1993: Unconditional Bases Are Optimal Bases for Data Compression and for Statistical Estimation. *Appl. Comput. Harmonic Anal.*, **1**, 100–115, doi:10.1006/acha.1993.1008.

Eaton, B., J. Gregory, B. Drach, K. Taylor, S. Hankin, J. Blower, J. Caron, R. Signell, P. Bentley, G. Rappa, H. Höck, A. Pamment, M. Juckes, M. Raspaud, R. Horne, T. Whiteaker, D. Blodgett, C. Zender, and D. Lee, 2020: NetCDF Climate and Forecast (CF) Metadata Conventions version 1.8. Accessed 27 December 2020, https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.pdf.

Elmore, K. L., Z. L. Flamig, V. Lakshmanan, B. T. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz, 2015: Verifying forecast precipitation type with mPING. *Wea. Forecasting*, **30**, 656–667, doi:10.1175/WAF-D-14-00068.1.

Eynard-Bontemps, G., R. Abernathey, J. Hamman, A. Ponte, and W. Rath, 2019: The Pangeo Big Data Ecosystem and its use at CNES. *Proc. of the 2019 conference on Big Data from Space (BiDS' 2019),* Munich, Germany, 49–52, doi:10.2760/848593.

Feser, F., M. Schubert-Frisus, H. von Storch, M. Zahn, M. Barcikowska, S. Haeseler, C. Lefebvre, and M. Stendel, 2015: Hurrican Gonzalo and its extratropical transition to a strong European storm. *Bull. Amer. Meteor. Soc.*, **96**, S51–S55,

doi:10.1175/BAMS-D-15-00122.1.

Foster, C., E. Crosman, J. Horel, 2017: Simulations of a Cold-Air Pool in Utah's Salt Lake Valley: Sensitivity to Land Use and Snow Cover. Boundary Layer Meteorology. 164, 63-87, doi:10.1007/s10546-017-0240-7

Giuliani, G., B. Chatenoux, T. Piller, F. Moser, and P. Lacroix, 2020: Data Cube on Demand (DcoD): Generating an earth observation Data Cube anywhere in the world. *Int. J. Appl. Earth. Obs. Geoinformation*, **87**, 1–6, doi:10.1016/j.jag.2019.102035.

Gowan, T. A., and J. Horel, 2020: Evaluation of IMERG-E Precipitation Estimates for Fire Weather Applications in Alaska. Wea. Forecasting, 35, 1831–1843, doi:10.1175/WAF-D-20-0023.1.

Green, A., 2020: Portland's air quality is off the charts Sunday, and parts of Oregon are just as bad due to wildfires. The Oregonian/Oregon Live. Accessed 15 January 2021, https://www.oregonlive.com/news/2020/09/portlands-air-quality-is-off-the-charts-on-sunday-and-much-of-oregon-is-just-as-bad-due-to-wildfires.html.

Harris, C. R. and Coauthors, 2020: Array programming with NumPy. *Nature*, **585**, 357–362, doi:10.1038/s41586-020-2649-2.

Houston, J., G. Zuidhof, L. Bergamini, Y. Ye, A. Jain, S. Omari, V. Iglovikov, and P. Ondruska, 2020: One Thousand and One Hours: Self Driving Motion Prediction Dataset. Lyft, accessed 9 December 2020, https://level5.lyft.com/dataset/.

Iris, 2020: Iris: A Python library for analysing and visualising meteorological and oceanographic datasets, version 2.4.0. Met Office, accessed 22 December 2020, https://scitools.org.uk/iris.

James, E. P., and S. G. Benjamin, 2017: Observation System Experiments with the Hourly Updating Rapid Refresh Model Using GSI Hybrid Ensemble–Variational Data Assimilation. *Mon. Wea. Rev.*, **145**, 2897–2918, doi:10.1175/MWR-D-16-0398.1.

Keller, J. H., C. M. Grams, M. Reimer, H. M. Archambault, L. Bosart, J. D. Doyle, J. L. Evans, T. J. Galarneau Jr., K. Griffin, P. A. Harr, N. Kitabatake, R. McTaggart-Cowan, F. Pantillon, J. F. Quinting, C. A. Reynolds, E. A. Ritchie, R. D. Torn, and F. Zhang, 2019: The Extratropical Transition of Tropical Cyclones. Part II: Interaction with the Midlatitude Flow, Downstream Impacts, and Implications for Predictability. *Mon. Wea. Rev.*, **147**, 1077–1106, doi:10.1175/MWR-D-17-0329.1.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea.*

Forecasting, **24**, 1691–1705, doi:10.1175/2009WAF2222201.1.

Kuhn, M., J. M. Kunkel, and T. Ludwig, 2016: Data compression for climate data. *Supercomput. Front. Innovations*, **3**, 75–94, doi:10.14529/jsfi160105.

Lagerquist, R., 2016: Using machine learning to predict damaging straight-line convective winds. M.S. thesis, School of Meteorology, University of Oklahoma, 251 pp. [Available online at http://hdl.handle.net/11244/44921.]

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bull. Amer. Meteor. Soc.*, **90**, 785–798, doi:10.1175/2008BAMS2604.1.

McCaie, T., 2019: Creating a data format for high momentum datasets. Medium, accessed 20 December 2020, https://medium.com/informatics-lab/creating-a-data-format-for-high-momentum-datasets-a394fa48b671.

McCorkle, T. A., J. D. Horel, A. A. Jacques, and T. Alcott, 2018: Evaluating the Experimental High-Resolution Rapid Refresh – Alaska Modeling System using USArray Pressure Observations, *Wea. Forecasting*, **33**, 933–953, doi: 10.1175/WAF-D-17-0155.1.

McGovern, A., K. L. Elmore, D. J. Gagne, S. E. Haupt, C. D. Karstens, R. Lagerquist, T. Smith, and J. K. Williams, 2017: Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bull. Amer. Meteor. Soc.*, **98**, 2073–2090, doi:10.1175/BAMS-D-16-0123.1.

Miles, A., and Coauthors, 2020: zarr-developers/zarr-python: v2.5.0. https://zenodo.org/record/4069231.

Molthan, A. L., J. L. Case, J. Venner, R. Schroeder, M. R. Checchi, B. T. Zavodsky, A. Limaye, and R. G. O'Brien, 2015: Clouds in the Cloud: Weather Forecasts and Applications within Cloud Computing Environments. *Bull. Amer. Meteor. Soc.*, **96**, 1369–1379, doi:10.1175/BAMS-D-14-00013.1.

National Oceanic and Atmospheric Administration (NOAA), 2020: Big Data Program. Accessed 8 December 2020, https://www.noaa.gov/organization/information-technology/big-data-program.

Nativi, S., P. Mazzetti, and M. Craglia, 2017: A view-based model of data-cube to support big earth data systems interoperability. *Big Earth Data*, **1**, 75–99, doi:10.1080/20964471.2017.1404232.

Pearson, R. D., R. Amato, and D. P. Kwiatkowski, and MalariaGEN Plasmodium Falciparum Community Project, 2019: An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples, *BioRxiv*, doi:

10.1101/824730.

Reeves, H. D., K. L. Elmore, A. Ryzhkov, T. Schuur, and J. Krause, 2014: Sources of uncertainty in precipitation-type forecasting. *Wea. Forecasting*, **29**, 936–953, doi:10.1175/WAF-D-14-00007.1.

Sharman, R., 2016: Nature of aviation turbulence. *Aviation Turbulence: Processes, Detection, Prediction*, R. Sharman and T. Lane, Eds., Springer, 3–30, doi:10.1007/978-3-319-23630-8_1.

Signell, R. P., and D. Pothina, 2019: Analysis and Visualization of Coastal Ocean Model Data in the Cloud. *J. Mar. Sci. Eng.*, **7**, 1–12, doi: 10.3390/jmse7040110.

Silver, J. and C. Zender, 2017: The compression-error trade-off for large gridded data sets. *Geosci. Model Dev.*, **10**, 413–423, doi:10.5194/gmd-10-413-2017.

Siuta D., G. West, H. Modzelewski, R. Schigas, and R. Stull, 2016: Viability of Cloud Computing for Real-Time Numerical Prediction. *Wea. Forecasting*, **31**, 1985–1996, doi:10.1175/WAF-D-16-0075.1.

Vance, T. C., M. Wengren, E. Burger, D. Hernandez, T. Kearns, E. Medina-Lopez, N. Merati, K. O'Brien, J. O'Neil, J. T. Potemra, R. P. Signell, and K. Wilcox, 2019: From the Oceans to the Cloud: Opportunities and Challenges for Data, Models, Computation and Workflows. *Front. Mar. Sci.*, **6**, 1–18, doi:10.3389/fmars.2019.00211.

Vannitsem, S., J. B. Bremnes, J. Demaeyer, G. R. Evans, J. Flowerdew, S. Hemri, S. Lerch, N. Roberts, S. Theis, A. Atencia, Z. Ben Bouallègue, J. Bhend, M. Dabernig, L. De Cruz, L., Hieta, O. Mestre, L. Moret, I. O. Plenković, M. Schmeits, M., Taillardat, J. Van den Bergh, B. Van Schaeybroeck, K. Whan, and J. Ylhaisi, 2020: Statistical Postprocessing for Weather Forecasts – Review, Challenges and Avenues in a Big Data World. *Bull. Amer. Meteor. Soc.*, Early Online Release,1–44, doi:10.1175/BAMS-D-19-0308.1.

Wang, N., J. Bao, J. Lee, F. Moeng, and C. Matsumoto, 2015: Wavelet Compression Technique for High-Resolution Global Model Data on an Icosahedral Grid. *J. Atmos. Oceanic Technol.*, **32**, 1650–1667, doi:10.1175/JTECH-D-14-00217.1.

Wang, Y. Q., 2014: MeteoInfo: GIS Software for meteorological data visualization and analysis. *Meteor. Appl.*, **21**, 360–368, doi:10.1002/met.1345.

Xu, M., G. Thompson, D. R. Adriaansen, and S. D. Landolt, 2019: On the value of time-lag-ensemble averaging to improve numerical model predictions of aircraft icing conditions. Wea. Forecasting, 34, 507–519, doi:10.1175/WAF-D-18-0087.1.

Yao, X., G. Li, J. Xia, J. Ben, Q. Cao, L. Zhao, Y. Ma, L. Zhang, and D. Zhu, 2020: Enabling the Big Earth Observation Data via Cloud Computing and DGGS: Opportunities and Challenges. *Remote Sens.*, **12**, 1–15, doi:10.3390/rs12010062.

CHAPTER 4

CONCLUSION

4.1 Research Summary

The impacts of the global weather enterprise are far reaching, directly or indirectly affecting decisions made in all aspects of society and the economy. Accurate, accessible weather data are critical for not only the average person, but for the researchers, developers, and forecasters who are tasked with disseminating and communicating it. The primary focus of this work was to identify obstructions to progress or operational needs in high-stakes weather forecasting, research, or industry applications. In addition, this work addresses needs to post-process output from existing models to perform data analytics and optimize workflows. In this dissertation, we identified and addressed two distinct operational challenges: 1) Availability and accuracy of estimated precipitation during fire seasons in Alaska, and 2) Accessibility of high-resolution numerical weather prediction model output for use in operational forecasting and machine learning workflows.

4.1.1 IMERG-E Precipitation Estimates in Alaska

This work began with the availability of IMERG version 06B, which for the first time, made IMERG precipitation estimates available poleward of 60° N/S. At high-

latitudes, model output is coarse and in-situ observations are sporadic, making the spatial distribution of sensible weather parameters difficult to determine with any degree of accuracy (Huffman et al. 1997 Huffman et al. 2001; Kidd et al. 2017). The poor representation of atmospheric conditions is only made more challenging for vast landscapes with varying terrain and the potential for high-impact weather, as in the 1.718 million km^2 of land within the borders of Alaska.

In Alaska, wildfire outbreaks are common in the summer months (June – August), especially in sparsely-populated areas of the interior where firefighting resources are limited (Partain et al. 2016). Forecasting and quantifying the wildfire potential is an incredible challenge across this enormous state with few observational resources, albeit a critical one, as the safety of firefighters depends on it. Forecasters and decision makers depend on metrics calculated using the CFFDRS, which relies on sensible weather parameters (temperature, moisture, wind, precipitation) as input. The most challenging of these parameters, daily precipitation, is arguably the most critical for accurate representation of fuel conditions and moisture content (Ziel et al. 2015; Pickell et al. 2017). This work presented an alternative source of gridded precipitation input, global precipitation estimates derived from the near real-time IMERG-Early (IMERG-E) algorithm.

The baseline accuracy of the IMERG-E algorithm was determined by comparing 24-hour accumulated precipitation estimates to observations from in situ stations over six fire seasons (1 June-31 August 2014-2019). To account for the limited number of ground stations and varying spatial distribution of precipitation, we subset the stations and IMERG-E grid into quadrants demarcated by the 63ºN and 152ºW latitude and longitude

lines. For each region, the 24-hour IMERG-E precipitation estimates and station observations were aggregated to calculate comparable cumulative distributions functions and identify pertinent event magnitude and the corresponding percentile ranges. The IMERG-E estimates were validated as a function of season, geographical region and event magnitude percentile. We quantified the error in the IMERG-E estimates using the following statistical metrics: frequency bias, bias ratio, mean absolute error (MAE), and root-mean-square-error (RMSE).

Upon evaluation, it was found that the IMERG-E algorithm consistently overestimated precipitation amounts for all percentile ranges and regions, with the exception of the Southeast region during the 2018 fire season. MAE and RMSE values were large for the highest percentile ranges (precipitation amounts), with the largest error values found in the eastern half of the state. The RMSE and MAE were found to be similar in magnitude during each fire season, which in this scenario, suggests the existence of a systematic wet bias in the algorithm at high-latitudes.

Using precipitation estimates that have a known wet bias can result in an underestimation of fire potential, inaccurate communication of risk, or misallocation of critical resources. For this reason, we employed a bias correction method, regional quantile mapping, to address this bias in the IMERG-E precipitation estimates. The regional quantile mapping method allowed for correction of these estimates not only at the station locations, but for the entire grid. Following the bias correction, a seasonal cross-validation was performed to calculate the same error metrics as in the baseline verification. The employed bias correction method yielded promising results:

1. The systematic wet bias identified in the baseline IMERG-E validation was

considerably reduced.

2. Frequency bias and bias ratio scores were improved, and in some regions and fire seasons, approached unity.
3. The largest improvements in MAE and RMSE values were seen in the Northwest and Northeast geographical regions over all percentile ranges and event magnitudes.

Overall, the corrected IMERG-E estimates proved to be valuable as a daily precipitation resource in Alaska. Studies have noted that satellite-derived precipitation estimates may not be adequate for representing in complex terrain (Lundquist et al. 2019), and given the lack of measurements in mountainous areas, a quantile mapping correction would likely be ineffective in those areas. The IMERG-E precipitation estimates could be a viable option for computing indices describing fire potential and fuel conditions in Alaska, if the underlying wet bias can be reduced.

Future work to evaluate precipitation estimates derived from later iterations of the IMERG algorithm would be beneficial to bolster the case for potential use in operational fire weather applications. An expansion of the regional quantile mapping method using less arbitrary regional boundaries would be of considerable interest and could be done by grouping stations with similar average seasonal precipitation. IMERG estimates have now been reprocessed using retrievals from the Tropical Rainfall Measurement Mission era, extending the potential period of record for calculating distributions for bias correction methodologies.

4.1.2 The HRRR-Zarr Archive

The second half of this work addressed the issues concerning accessibility of high-resolution numerical weather prediction model output for use in operational forecasting and machine learning workflows. Recent advances in computing power have enabled numerical weather prediction models to be run more frequently and at higher spatial resolutions, rapidly increasing the velocity and volume of available data, while end users struggle to keep up (Benjamin et al. 2018; Giuliani et al. 2020; Yao et al. 2020). Output from numerical weather prediction models are typically stored in file formats such as GRIB2 and netCDF which are efficient for transporting and compressing large amounts of data, but require considerable computational resources to store, process, and analyze. These challenges have forced operational forecasters and researchers to make pragmatic decisions regarding which models, parameters, and periods of record are feasible to store and disseminate for operational applications.

The NOAA Big Data Program (BDP) was established in 2015 to address the issues related to the accessibility of the increasing volume of operational model output produced each day in that agency. This program involved partnerships with several IaaS providers (e.g., AWS, Google Cloud Platform, IBM, and Microsoft Azure) that have the capacity to store massive datasets and provide compute resources for users to post-process them without leaving the IaaS environment (Molthan et al. 2015; Siuta et al. 2016). In an effort to address this operational need and build upon the NOAA BDP initiative, we present a workflow for converting the HRRR model archive originally managed by the University of Utah (Blaylock et al. 2017; Blaylock et al. 2018), to a

flexible, cloud-compatible file format, Zarr. In addition to the NOAA BDP, the conversion of the HRRR GRIB2 dataset was made possible by the Amazon Sustainability Data Initiative.

The resulting dataset, HRRR-Zarr, is a solution to the data bottleneck problem that occurs as a result of repeated access of GRIB2 files for high-throughput applications such as machine learning and operational forecasting. The HRRR-Zarr archive is stored in an S3 bucket with the unique identifier “hrrrzarr”, and for each model run contains a set of analysis (F00) and forecast files (F01 up to F48) that are composed of 173 surface parameter arrays divided into 96 regional chunks. Each HRRR-Zarr file contains nearly 17,000 tiny objects that can be accessed individually or simultaneously.

Using HRRR-Zarr data files, we demonstrated the following workflow examples for analyzing large amounts of sensible weather parameters:

1. Assembling location-specific time series of forecast data over a range of model runs,
2. Identifying consistencies and differences among many model forecasts for the same valid time,
3. Calculating empirical cumulative distributions for sensible weather parameter over multiple years,
4. Identifying model forecasts that predict extreme conditions relative to typical conditions from recent years.

In addition to illustrating optimized workflows for big data applications, we have contributed to making the HRRR-Zarr archive accessible to end users across disciplines regardless of computational or storage resources through the Amazon Open Data

Registry program.

The GRIB2 to Zarr conversion of the HRRR model archive is a first step to make model data more accessible to end users. As technology and computational power continues to advance, the inevitable path of numerical weather prediction models is towards probabilistic guidance from ensemble forecasting systems (Frogner et al. 2019; Schwartz et al. 2019). Ensemble models introduce an additional data dimension (number of members), which compounds the volume of data produced by each model run. Plans are being developed for GRIB2 to Zarr conversion of model output from an ensemble forecast system such as the NOAA Global Ensemble Forecast System (GEFS), which are also available through the NOAA BDP and the Amazon Sustainability Initiative. Additional work to streamline a Zarr conversion process in the cloud environment would benefit users who want the flexibility to customize a Zarr file structure for their own purposes. Utilizing the Zarr format as an alternative file structure for the vast amount of numerical weather prediction output may help expand its already wide reach to data scientists in other disciplines, while optimizing workflows for end users throughout the weather enterprise.

4.2 References

Benjamin, S. G., J. M. Brown, G. Brunet, P. Lynch, K. Saito, and T. W. Schlatter, 2018: *100 Years of Progress in Forecasting and NWP Applications*, *Meteor. Monogr.*, **59**, 13.1–13.67, doi:10.1175/AMSMONOGRAPHS-D-18-0020.1.

Blaylock, B. K., J. D. Horel, and S. T. Liston, 2017: Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output. *Computers and Geosciences*, **109**, 43–50, doi: 10.1016/j.cageo.2017.08.005.

—, —, and C. Galli, 2018: High-Resolution Rapid Refresh Model Data Analytics Derived on the Open Science Grid to Assist Wildland Fire Weather Assessment. *J. Atmos. Oceanic Technol.*, **35**, 2213–2227, doi:10.1175/JTECH-D-18-0073.1.

Frogner, I-L, A. Singleton, M. Køltzow, U. Andrae, 2019: Convection-permitting ensembles: Challenges related to their design and use. *Q. J. R. Meteorol Soc.*, **2019**, 1–17. doi:10.1002/qj.3525.

Giuliani, G., B. Chatenoux, T. Piller, F. Moser, and P. Lacroix, 2020: Data Cube on Demand (DcoD): Generating an earth observation Data Cube anywhere in the world. *Int. J. Appl. Earth. Obs. Geoinformation*, **87**, 1–6, doi:10.1016/j.jag.2019.102035.

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bull. Amer. Meteor.* Soc., **78**, 5–20, doi: 10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

——, ——, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *J. Hydrometeor.*, **2**, 36–50, doi: 10.1175/1525-7541(2001)002<0036:GPAODD>2.0.C0;2.

Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, D. B. Kirschbaum, 2017: So, how much of the Earth's Surface is covered by rain gauges? *Bull. Amer. Meteor. Soc.*, **98**, 69–78, doi:10.1175/BAMS-D-14-00283.1.

Lundquist, J., M. Hughes, E. Gutmann, and S. Kapnick, 2019: Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. *Bull. Amer. Meteor. Soc.*, **100**, 2473–2490, doi: 10.1175/BAMS-D-19-001.1.

Molthan, A. L., J. L. Case, J. Venner, R. Schroeder, M. R. Checchi, B. T. Zavodsky, A. Limaye, and R. G. O'Brien, 2015: Clouds in the Cloud: Weather Forecasts and Applications within Cloud Computing Environments. *Bull. Amer. Meteor. Soc.*, **96**, 1369–1379, doi:10.1175/BAMS-D-14-00013.1.

Pickell, P. D., N. C. Coops, C. J. Ferster, C. W. Bater, K. D. Blouin, M. D. Flannigan, and J. Zhang, 2017: An early warning system to forecast the close of the spring burning window from satellite-observed greenness. *Sci. Rep.*, **7**, 14190, doi: 10.1038/s41598-017-14730-0.

Schwartz, C.S., G.S. Romine, R.A. Sobash, K.R. Fossell, and M.L. Weisman, 2019: NCAR's Real-Time Convection-Allowing Ensemble Project. *Bull. Amer. Meteor. Soc.*, **100**, 321–343, doi:10.1175/BAMS-D-17-0297.1.

Siuta D., G. West, H. Modzelewski, R. Schigas, and R. Stull, 2016: Viability of Cloud Computing for Real-Time Numerical Prediction. *Wea. Forecasting*, **31**, 1985–1996, doi:10.1175/WAF-D-16-0075.1

Yao, X., G. Li, J. Xia, J. Ben, Q. Cao, L. Zhao, Y. Ma, L. Zhang, and D. Zhu, 2020: Enabling the Big Earth Observation Data via Cloud Computing and DGGS: Opportunities and Challenges. *Remote Sens.*, **12**, 1–15, doi:10.3390/rs12010062.

Ziel, R. H., and Coauthors; 2015: Modeling fire growth potential by emphasizing significant growth events: characterizing climatology of fire growth days in Alaska's boreal forest. *11th Symp. on Fire and Forest Meteorology*, Minneapolis MN, Amer. Meteor. Soc.,1.2. [Available online https://ams.confex.com/ams/11FIRE/webprogram/Paper272864.html.].