

TOWARDS NEAR-REAL-TIME SPATIAL FORECASTING OF RAINFALL TRIGGERED LANDSLIDES

XUETONG WANG
August, 2022

SUPERVISORS:
dr.L.Lombardo
dr.H.Tanyas



TOWARDS NEAR-REAL-TIME SPATIAL FORECASTING OF RAINFALL TRIGGERED LANDSLIDES

XUETONG WANG

Enschede, The Netherlands, August, 2022

Thesis submitted to the Faculty of Geo-information Science and Earth
Observation of the University of Twente in partial fulfilment of the requirements
for the degree of Master of Science in Geo-information Science and Earth
Observation.

Specialization: M-GEO

SUPERVISORS:

dr.L.Lombardo

dr.H.Tanyaş

THESIS ASSESSMENT BOARD:

dr.C.J. van Westen (chair)

dr.A.C (Harry) Seijmonsbergen (External examiner)

Disclaimer

This document describes work undertaken as part of a programme of study at the Faculty of Geo-information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

With the increase of frequency and intensity of heavy precipitation in the future, rainfall triggered landslides (RTL) can be one of the major threat to human life and property security. Early warning systems of natural hazards are one of the most effective measure for reducing disaster losses and risks. However, the forecast of RTL in near-real-time (NRT) is extremely difficult since the quality of NRT precipitation data is relatively poor. Quantile regression forest, a state-of-the-art statistical postprocessing method, has been proved to reduce the difference existing between NRT satellite precipitation estimates and ground-based rainfall data.

However, this field has still a long way to go before becoming operational. In fact, the level of bias-reduction sought in this work did not reach what was expected. This implies that the reliability of these postprocessing practices still requires improvement. This been said, when the bias-corrected rainfall maps are put side by side with raw satellite product, the pattern of the first matches much more closely the locations where landslide events have been mapped in a test site in North-Eastern Turkey. This still leave an optimistic perspective on the application of post-processing techniques in the field of weather science and in general for natural hazard assessment. Ideally, by correcting the continuous information in space and time provided by satellite rainfall estimates, one could create a new operational tool for landslide early warning system, not bound to the financial and deployment requirement typical of rain gauge and terrestrial radar stations.

Keywords

Statistical Post-Processing, Quantile Regression Forest, Rainfall Triggered Landslides, IMERG-E, Black Sea Region

ACKNOWLEDGEMENTS

Several days ago, I sat in front of my laptop in despair, feeling as if I would never be able to finish my master thesis. I felt I can never write so much content. Two years ago, I was at a low point in my life and had no idea what to do in the future. Eight years ago, I had just finished my high school studies and thought I would never enter a scientific field at that time. However, at this moment, I have just finished the writing of my master thesis. I know exactly what I want to do in the future and I am proud and enjoyed in what I have learned. Without people who helped me along the way, I don't think I would have made it this far.

First of all, I want to express my most sincere gratitude to my supervisors and advisor: Dr. Luigi Lombardo, Dr. Hakan Tanyaş and Ashok Dahal. My supervisors are always extremely patient to me in our weekly meeting. They keep encouraging me even when I am inefficient at work or feeling low. Their passion for academics motivating me all time. Ashok helps me a lot when I encounter technical problems. I have gained a deeper understanding of machine learning with his help.

Secondly, I want to thank my supervisors for my internship at Royal Netherlands Meteorological Institute (KNMI): Eva van der Kooij, Dr. Kirien Whan and Dr. Maurice Schmeits. They taught me a lot during my six-month internship, and what I have learned in the internship inspired the most important part for my thesis.

I also want to thank to Dr. Andrew Bush, Dr. Patrick Getty and all other professors from the department of earth sciences at University of Connecticut. They led me into this amazing field I would never studied this major without them.

I would like to express my respect to all professors I have met at ITC Faculty Geo-Information Science and Earth Observation (ITC). I have learned a lot during the past two years, and I also got better idea about what I want to do in the future.

Certainly, I am also grateful to ITC for providing such an special international community. I have made friends from all over the world, and I hope those precious friendships can stay with me for the rest of my life.

Meanwhile, I want to thank my parents. They are best parents in the world. They always support me no matter what choice I have made. Being your child is the luckiest thing in my life.

The last part is for my boyfriend who has been with me all the time. He has cooked and encouraged for me during my busiest and hardest time. Your love and support mean a lot to me.

TABLE OF CONTENTS

Abstract	i
Acknowledgements	ii
1 Introduction	1
1.1 Background	1
1.2 Literature Review	3
1.3 Research problem	4
1.4 Research Objectives and questions	5
1.4.1 Objective	5
1.4.2 Sub-Objectives	5
1.4.3 Research Questions	5
2 Study area and datasets	7
2.1 Study Area	7
2.2 Datasets	11
2.2.1 Ground Observations	11
2.2.2 Landslide Inventories	11
2.2.3 NRT Satellite Precipitation Data	12
2.2.4 Covariates	12
3 Methodology	14
3.1 Interpolation	14
3.1.1 Ordinary Cokriging	14
3.1.2 Inverse Distance Weighting	14

3.2	Statistical Postprocessing	15
3.2.1	Ensemble Model Output Statistics	15
3.2.2	QRF	16
3.3	Model Justification	17
3.4	Verification	18
3.4.1	Model Comparison	18
3.4.2	Model Assessment	19
3.4.3	General overview	20
4	Results	22
4.1	Rain Gauge Assessment	22
4.2	Interpolation	23
4.3	IMERG Assessment	23
4.4	Model Validation	29
4.5	Model Testing	30
4.5.1	Model Comparison	30
4.5.2	Model Assessment	33
4.5.3	Predictor Importance	35
4.6	Attempt to explain landslide occurrences through rainfall patterns	36
5	Discussion	40
5.1	Observational precipitation interpolation	40
5.2	Statistical Postprocessing	40
6	Conclusion and recommendation	43
	List of References	44

LIST OF FIGURES

1.1	Statistical Postprocessing (Li et al., 2017)	3
2.1	Distribution of fatal landslide in major geographical provinces of Turkey. a: The number of fatal landslides and b: fatalities in major geographical regions and their percentages (Görüm and Fidan, 2021)	8
2.2	Distribution of triggers for 389 fatal landslide events. a: Spatial distribution of events and b: distribution of fatal landslides, and c: fatalities according to the reported triggers for each event. “N/A” (not available) indicates 45 events with unknown triggers, and “Others” specifies some exceptional landslides (n = 3) triggered during archaeological excavations (Görüm and Fidan, 2021)	9
2.3	Combined maps of a: slope gradient and annual rainfall and b: slope gradient, annual rainfall, and population density. Black dots show the location of fatal landslide events (n = 389) (Görüm and Fidan, 2021)	10
2.4	Elevation of study area. Black dots represent ground stations	11
3.1	PDF and CDF for ZAGA distribution (Stasinopoulos et al., 2017)	15
3.2	An example of random forests (Dimitriadis et al., 2018)	17
3.3	Examples of reliability diagram (Wilks, 2011)	20
3.4	Flowchart	21
4.1	Ground Stations Distribution. Each blue point represents a ground station	22
4.2	Interpolation comparison for landslide Artvin Hopa (units: mm/day)	24
4.3	Interpolation comparison for landslide Ordu-Perşembe (units:mm/day)	25
4.4	Interpolation comparison for landslide Rize Kaptanpaşa (units: mm/day)	26
4.5	IMERG comparison for landslide Artvin Hopa (units: mm/day)	27
4.6	IMERG comparison for landslide Ordu-Perşembe (units: mm/day)	28
4.7	IMERG comparison for landslide Rize Kaptanpaşa (units: mm/day)	29
4.8	RMSE comparison (units: mm/day)	31
4.9	PCC comparison	32

4.10 ME comparison (units: mm/day)	32
4.11 Brier skill score for three models with different thresholds (blue is for Nov-Feb, red is for Mar-Jun, green is for Jul-Oct, the units of thresholds are mm/day) . . .	33
4.12 Reliability Diagram for multiple daily rainfall thresholds (Nov-Feb)	34
4.13 Reliability Diagram for multiple daily rainfall thresholds (Mar-Jun)	34
4.14 Reliability Diagram for multiple daily rainfall thresholds (Jul-Oct)	35
4.15 Prediction for precipitation/landslide: Artvin Hopa	37
4.16 Prediction for precipitation/landslide: Ordu-Perşembe	38
4.17 Prediction for precipitation/landslide: Rize Kaptanpaşa	39

LIST OF TABLES

2.1	Landslide Inventories	12
3.1	Predictors Selection	18
3.2	Training, validation and testing groups	18
4.1	RMSE comparison for validation folds (mm/day)	30
4.2	RMSE and PCC comparison for the testing group (units of RMSE: mm/day) . .	31
4.3	The five most important predictors	36

Chapter 1

Introduction

1.1 BACKGROUND

It has been mostly agreed that both frequency and intensity of heavy precipitation have increased significantly during past decades due to the climate change at the global scale (Alexander, 2016; Berg et al., 2013; Fischer and Knutti, 2016; Kharin et al., 2013; Myhre et al., 2019; Sillmann et al., 2013). Myhre et al. (2019) have shown that the global frequency of extreme precipitation events, which can occur with an average of twice per decade nowadays, will increase in frequency by 1-2 events per decade following an increase of just one degree Celsius (C) in the future. The increased rate of precipitation intensity is similar to the increased rate of vapor pressure, which is 6-7 percent for each degree increase (Allan et al., 2014; Fischer and Knutti, 2016; Myhre et al., 2019; O’Gorman, 2015). Worryingly, precipitation pattern is closely related to landslides because precipitation is a major trigger (Petley, 2012). The increase of frequency and intensity of heavy precipitation can influence the magnitude, frequency, and temporal-spatial distribution of landslides and ultimately raise the disaster risks (Gallina et al., 2016). In this scenario, landslide could also be an even more severe threat in the coming years (Stoffel et al., 2014), especially in tropical countries and mountainous regions.

Landslide is a type of mass wasting process that occurs in all continents, and in all terrestrial environments both with natural and engineered slopes (Froude and Petley, 2018; Gariano and Guzzetti, 2016). Landslides can be threats to life and property security in inhabited area, transport facilities including multiple types of roads, and industries (Gariano and Guzzetti, 2016). When considering disaster risk assessments for a specific event, we usually need to consider hazard, vulnerability, and exposure. The “hazard” is usually expressed as the probability of an event occurring within a given time frame, so it needs to consider both spatial and temporal probabilities (van Westen et al., 2006). Since temporal probabilities of landslide hazards are usually difficult to calculate because of the lack of historical data (van Westen et al., 2006), spatial probability information becomes particularly significant for landslides assessment. Landslide susceptibility represents the relative probability of landsliding on a given hillslope (Guzzetti et al., 2006). Landslide susceptibility can be modelled both via physical-based and statistical-based methods. Much research focusing on landslides susceptibility by using physical methods has greatly contributed to understanding the mechanisms of landslides (Marin and Mattos, 2020; Medina et al., 2021; Ray et al., 2018). However, the scale at which physics-based analysis can be applied is generally localized, involving single slopes or catchments (N. Wang et al., 2021). Conversely, statistically-based methods can be extended over larger areas with more diverse geomorphological situations.

In both cases, information on the rainfall spatio-temporal patterns are of fundamental im-

portance. However, detail rainfall data is often not available. This research aims to improve the near-real-time (NRT) predictability of statistical modeling for rainfall triggered landslides (RTL) by improving the qualities of NRT precipitation products. Geo-environmental variables required to assess landslide susceptibility can be classified as predisposing (e.g., morphology, geology, land use, etc.) and triggering factors (precipitation, earthquake and anthropogenic factors)(Reichenbach et al., 2018). Compared to other variables, rainfall information in the predictive models developed for RTL is of utmost importance as it controls the landslide temporal onset as well as influencing the landslide scenario in space (Chaithong and Komori, 2020; Fan et al., 2020). In recent years, the prediction RTL susceptibility has developed towards space-time models and thus rainfall records and their forecasts could become part of the modeling procedures (Fathi et al., 2021). With the application of multiple satellite precipitation products, for instance, based on Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Measurement (GPM), the NRT spatiotemporal pattern of precipitation has also become available in four hours (Hsu et al., 2021; Qi et al., 2021), and they have also been adopted to improve the predictability for RTL susceptibility (Hong et al., 2007; Michaud et al., 2021). However, because of the data source and algorithm, NRT precipitation products with short latency are (still) biased and include large errors (H. Chen et al., 2020). Statistical postprocessing methods can be applied to improve the NRT products' quality (Y. Zhang et al., 2022).

The common approaches for measuring precipitation are ground stations with rain gauges, local radar systems and satellite precipitation estimates products (Gilewski and Nawalany, 2018). Each of them has some advantages and disadvantages. Rain gauges usually can represent "true" measurements of rainfall at a specific point, but the typical problem is the lack of spatial density of the rain gauge network, especially in mountainous regions or sparsely populated areas. Moreover, the data quality from rain gauges in large area might be degraded due to limited numbers of gauges and interpolation errors (Tang et al., 2018). Ground based radars can provide better spatial-temporal resolution of rainfall information but construction and maintenance of radar systems can be expensive, so they are not available in many countries, especially in mountainous and landslide prone regions (Nguyen et al., 2018). In addition, ground based radar systems also need long time for signal calibration and corrections, and those can take up to months for them to become actively part of operational systems. On the other hand, many paper indicated that satellite precipitation estimates are promising and demonstrate a noticeable improvement because of the emergence of Global Precipitation Measurement (GPM) based products (Gilewski and Nawalany, 2018; Nguyen et al., 2018). GPM based products can provide precipitation information for almost the whole globe, and cover many areas that couldn't be measured by rain gauges and local radar systems. In addition, GPM based NRT products are available in four hours, and this is a big advantage of operational flooding/landslides monitoring and forecasting systems. However, the accuracy and resolution of GPM based NRT products still need to be largely improved (Gentilucci et al., 2021; Gilewski and Nawalany, 2018). Improving the accuracy of satellite based precipitation products has been a concern in many studies (Dong et al., 2020; Gumindoga et al., 2019; Yu et al., 2021), and with the development of statistical postprocessing in weather forecast, methods such as ensemble model output statistics (EMOS) and quantile regression forests (QRF) have also been applied to correct satellite-based biases (Yang, 2020; T. Zhang et al., 2022; Y. Zhang et al., 2022).

Statistical postprocessing is a key technique for many national meteorological services to adjust the current forecast through correction of systematic errors. These errors become evident when comparing past raw forecasts to observational data (Scheuerer and Hamill, 2015; Vannitsem et al., 2021), thus opening up to statistical routines aimed at leveling these differences. In other words, the relationship between forecasts and observations from historical data can be built into

statistical models. The new forecasts can be adjusted based on the joint probability distribution from those models and output a predictive cumulative distribution function (CDF). The raw forecasts can be transferred to a calibrated deterministic or ensemble forecasts with less errors, which in turn implies a better rainfall information for end users. An example of a common statistical postprocessing procedure is shown in the Figure 1.1.

Recently, researchers have begun to apply statistical postprocessing (which was mainly used in weather forecasting) to satellite estimates (Y. Zhang et al., 2022). If such protocols would prove to be efficient in correcting the errors between observed and estimated rainfall, this would open up a number or research lines towards compound hazard modeling. For instance, postprocessed NRT satellite estimates could be used to explain landslides/floods patterns. And, if the postprocessed results from NRT satellite estimates can be shown to better predict landslides or floods, then statistical postprocessing methods could be even become part of NRT landslide/flood susceptibility modeling in the future. Unlike for other weather variables, statistical postprocessing of precipitation is more challenging because of the nature of precipitation, because precipitation values are non-negative but include many zero values and extreme values. Thus, modeling them together in one distribution model is relatively difficult (Scheuerer, 2014; van Straaten et al., 2018).

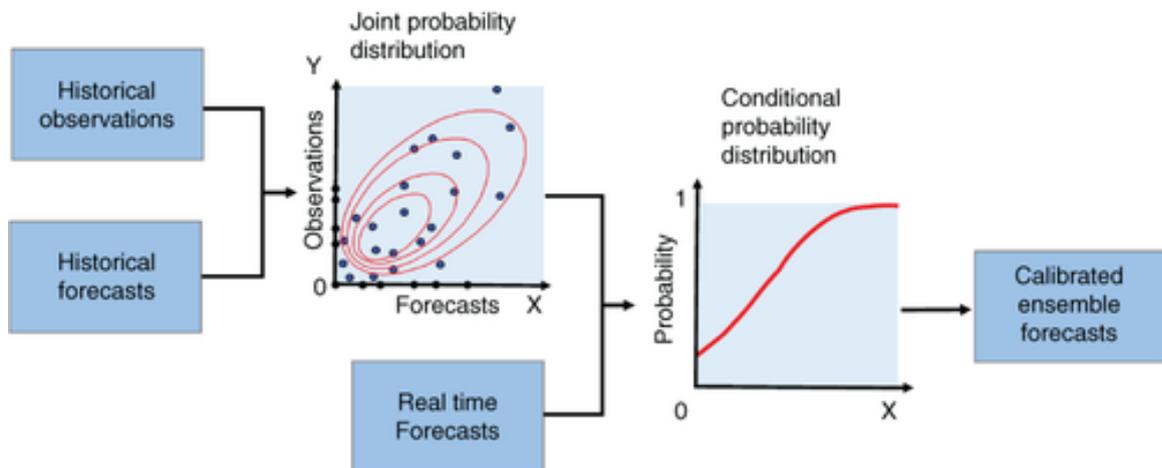


Figure 1.1: Statistical Postprocessing (Li et al., 2017)

1.2 LITERATURE REVIEW

In 2005, Gneiting et al. proposed ensemble model output statistics (EMOS), a variant of multiple linear regression, where various parametric distributions can be used to find the relationship between predictors and response variables (Gneiting et al., 2005). Since then, EMOS has been tried with many different statistical distributions. A generalized extreme value distribution (GEV) was applied to the COSMO-DE ensemble prediction system operated by the Germany Meteorological Service (Scheuerer, 2014). A censored and shifted gamma distribution (CSG/CSGD) was proposed and tested for precipitation postprocessing with two ensemble forecasting systems in the United States (Baran and Nemoda, 2016; Scheuerer and Hamill, 2015). In 2018, a zero-adjusted gamma distribution (ZAGA) was applied to KNMI's high-resolution, non-hydrostatic numerical weather prediction (NWP) model HARMONIE-AROME (HA) (van Straaten et al., 2018; Whan

and Schmeits, 2018). CSG and ZAGA are both variants of gamma distributions, and they can model precipitation better because of their heavy tails (Whan and Schmeits, 2018). They have also shown better performance compared to GEV (Baran and Nemoda, 2016) and lognormal or inverse Gaussian distributions (Bentzien and Friederichs, 2012).

With the fast development of machine learning in recent years, tree-based nonparametric techniques have also been widely used in precipitation postprocessing. Unlike parametric methods that need to set a specific distribution in advance, nonparametric methods can be more flexible while representing nonlinear relationships better (Vannitsem et al., 2021; Whan and Schmeits, 2018). A quantile regression forest (QRF) was applied to surface temperature and wind speed with Météo-France 35-member ensemble forecast (PEARP) and it performed better than EMOS (Taillardat et al., 2016). Based on this work, QRF was also applied to the HA model and Grand Limited Area Model Ensemble Prediction System (GLAMEPS), and was found to be more skillful than EMOS with a ZAGA distribution, especially for higher thresholds (van Straaten et al., 2018; Whan and Schmeits, 2018).

Applying statistical postprocessing techniques on NRT precipitation satellite images is still in the early stages of research. In 2022, a QRF-based postprocessor (QRF4P-NRT) has been proposed with a study case at the Yalong River basin in China (Y. Zhang et al., 2022). Researchers found that the results from QRF4P-NRT has improved the raw NRT GPM images significantly. This finding can be meaningful for rainfall related natural hazards such as flooding and landslides since NRT satellite products are quite important for forecasting those hazards (Khan et al., 2021; Soo et al., 2022).

1.3 RESEARCH PROBLEM

There has been a great number of research focusing on statistical postprocessing over the years (Vannitsem et al., 2021), and the development of such technology is also rapidly growing (Harris et al., 2022). More and more promising machine/deep learning based methods have been proposed and their performances also significantly improved with respect to traditional techniques. In the meantime, weather forecasting systems of many countries have also improved thanks to the technological progress. Much of the research in this field has been at the meteorological level only, but in reality they can also be helpful to many fields such as disaster risk reduction and natural resource management.

As mentioned in the background, though the emergence of NRT satellite products makes it is possible to predict RTL both temporally and spatially, there are still large errors in NRT satellite products (Gentilucci et al., 2021; Gilewski and Nawalany, 2018). This research aims to improve the accuracy of NRT satellite estimates for explaining the occurrence of landslides in the Black Sea Region of Turkey. This will be attempted by using statistical postprocessing techniques aimed at decreasing the error between rain gauge and satellite precipitation measurement. And, any improvement with respect to the raingauge baseline could mean an extention of the same protocol in other mountainous regions all over the world. Specifically, if the accuracy of NRT satellite precipitation estimates can be significantly improved, then NRT flooding/landslides susceptibility system can also be implemented in the future.

Statistical postprocessing on satellite precipitation estimates are supposed to be more complex and unstable compared to ensemble forecasts with short forecast lead time. Ensemble fore-

cast systems usually also include multiple weather elements such as wind direction and convective available potential energy (CAPE) which are related to precipitation information (Mendoza et al., 2015), and they can be potential predictors in the training procedure. A variable selection procedure could then even select the most important predictors and remove redundant information in space and time. On the other hand, NRT satellite estimates only provide precipitation information with large errors, and potential predictors should be collected from other sources. In this research, some potential predictors are collected from NRT satellite estimates and others express the potential landscape influence onto the spatio-temporal pattern of the precipitation.

1.4 RESEARCH OBJECTIVES AND QUESTIONS

1.4.1 Objective

The main objective of this research is to:

1. explore and implement methods for statistically postprocessing on near-real-time GPM satellite precipitation data in eastern Black Sea Region of Turkey.
2. test whether the improved rainfall estimates can explain the distribution of landslides in the area.

1.4.2 Sub-Objectives

1. To explore methods for interpolating local rain gauges data.
2. To evaluate the performance of GPM NRT product.
3. To evaluate the performance of different statistical postprocessing techniques on GPM NRT data.
4. To explore precipitation features in different seasons.
5. To improve the accuracy of landslide prediction by upgrading GPM NRT data.

1.4.3 Research Questions

The following research questions are designed to achieve the above sub-objectives:

1. Which considerations are needed for interpolating rain gauges in mountainous regions? What covariates can be included in the interpolation procedure? Which method is outperformed for precipitation interpolating in this study case?
2. Whether GPM NRT product is reliable for providing precipitation information?

3. Which predictors should be selected on statistical postprocessing models for NRT satellite estimates? Whether statistical postprocessing techniques can be applied on NRT satellite products? Which method is outperformed and why?
4. Which predictors are more important in the model for different seasons? How is the model perform for different precipitation thresholds?
5. Can statistical postprocessing applied on NRT satellite estimates be helpful for landslides susceptibility analysis? What are further direction for NRT landslides forecasting?

Chapter 2

Study area and datasets

2.1 STUDY AREA

Given that hillslopes make up of a significant portion of the world's ice-free terrestrial regions (Huggett, 2016), landslides can be one of the most significant threats to human life (Görüm and Fidan, 2021). Like other landslide-prone alpine countries, Turkey also suffers from landslides. Landslides in Turkey are responsible for the second highest number of death caused by natural hazards, just after earthquakes. Between 1995-2004, Turkey accounts for a quarter of the total number of landslide deaths in Europe, and it represents the European country with most victims caused by slope failures (Haque et al., 2016).

In a study of 90 years (1929-2019) of fatal landslide data from Turkey (Görüm and Fidan, 2021), around 37.8% of fatal landslides occurred in the Black Sea Region, and more than half of the fatalities happened in this region (Figure 2.1). The reasons that cause this region to hold such unfortunate record might be related to the high precipitation, combined to high slope gradients and population density. As shown in the Figure 2.2, most fatal landslides happen in the eastern Black Sea Region, these being triggered by precipitation. Thus, improving the precipitation forecast for this region can help local authorities to reduce the losses and risks from landslides. If NRT precipitation forecasts can be improved or landslides early warning system can be built for this eastern Black Sea Region in the future, the number of fatalities could be reduced.

Figure 2.3 shows two combined maps of Turkey. Figure 2.3a is a combined map of slope gradient and annual rainfall. Obviously, eastern Black Sea Region is marked to be much more prone to RTL as compared to other Turkish sectors. Moreover, Figure 2.3b is the combined map of slope gradient, annual rainfall and population density, which indicates that Elements at Risks (EaR) in this region are also particularly high.

Meanwhile, the decision of the study area for this research also depends on the data availability. Landslide inventories and ground truth of precipitation data collected by local radars or rain gauges are relatively difficult to get. In north eastern Turkey though, being the situation particularly appalling, local institutions have deployed a dense network of rain gauges and have also mapped landslides in space and time. Specifically, precipitation daily data from local ground stations over five and a half years have been accessed, together with three event-based landslide inventories spread over the same period.

For all of the above reasons, the eastern Black Sea Region has been selected as the study area for this research. Since the quality and quantity of the observational data are important for the performance of statistical models (Caldwell et al., 2013), only the region has available ground sta-

tions data is analyzed in this study (Figure 2.4). The elevation information is also shown in Figure 2.4, where the largest topographic change appears more to be more prominent in the eastern part of this region.

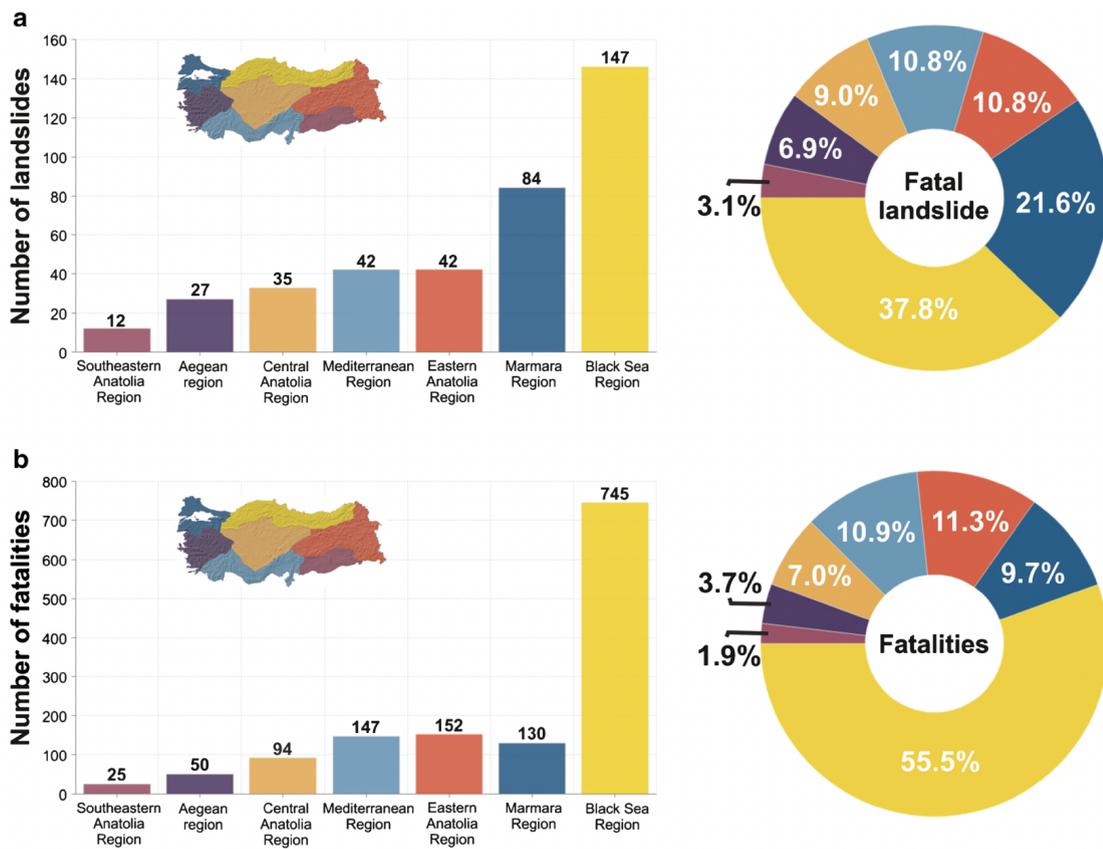


Figure 2.1: Distribution of fatal landslide in major geographical provinces of Turkey. a: The number of fatal landslides and b: fatalities in major geographical regions and their percentages (Görüm and Fidan, 2021)

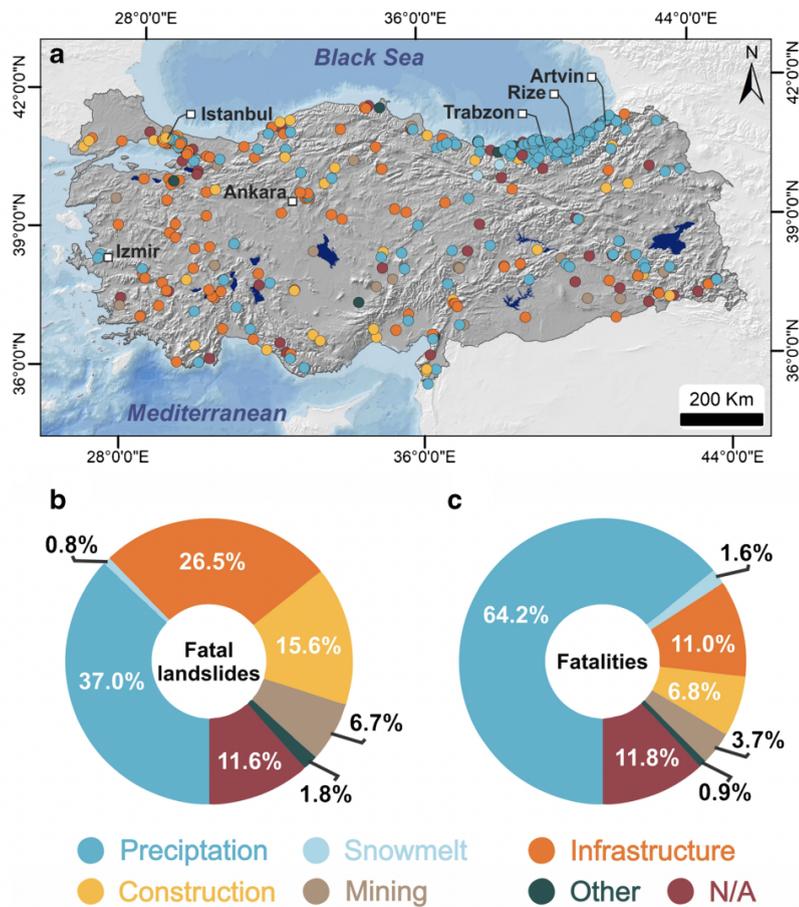


Figure 2.2: Distribution of triggers for 389 fatal landslide events. a: Spatial distribution of events and b: distribution of fatal landslides, and c: fatalities according to the reported triggers for each event. “N/A” (not available) indicates 45 events with unknown triggers, and “Others” specifies some exceptional landslides (n = 3) triggered during archaeological excavations (Görüm and Fidan, 2021)

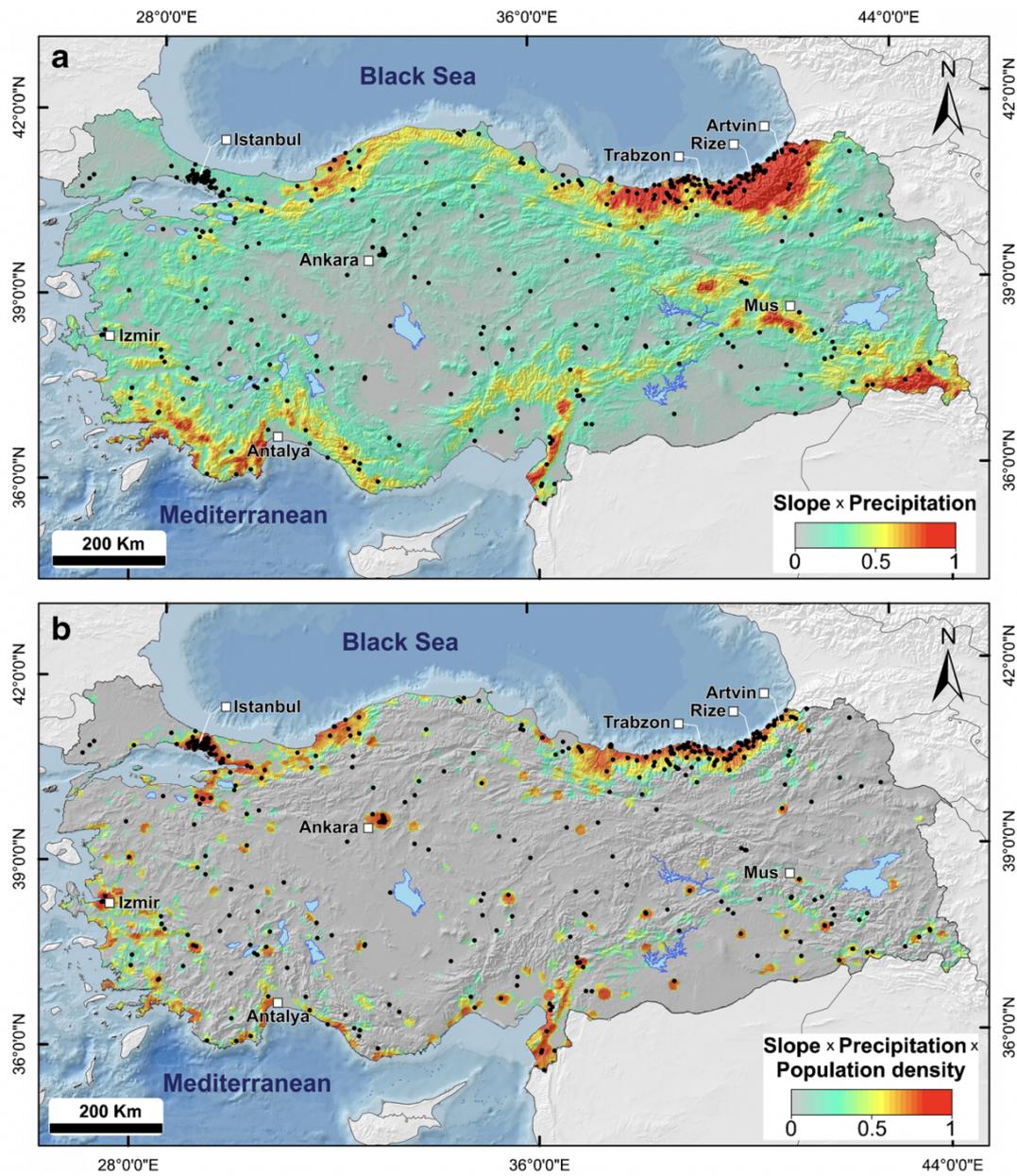


Figure 2.3: Combined maps of a: slope gradient and annual rainfall and b: slope gradient, annual rainfall, and population density. Black dots show the location of fatal landslide events ($n = 389$) (Görüm and Fidan, 2021)

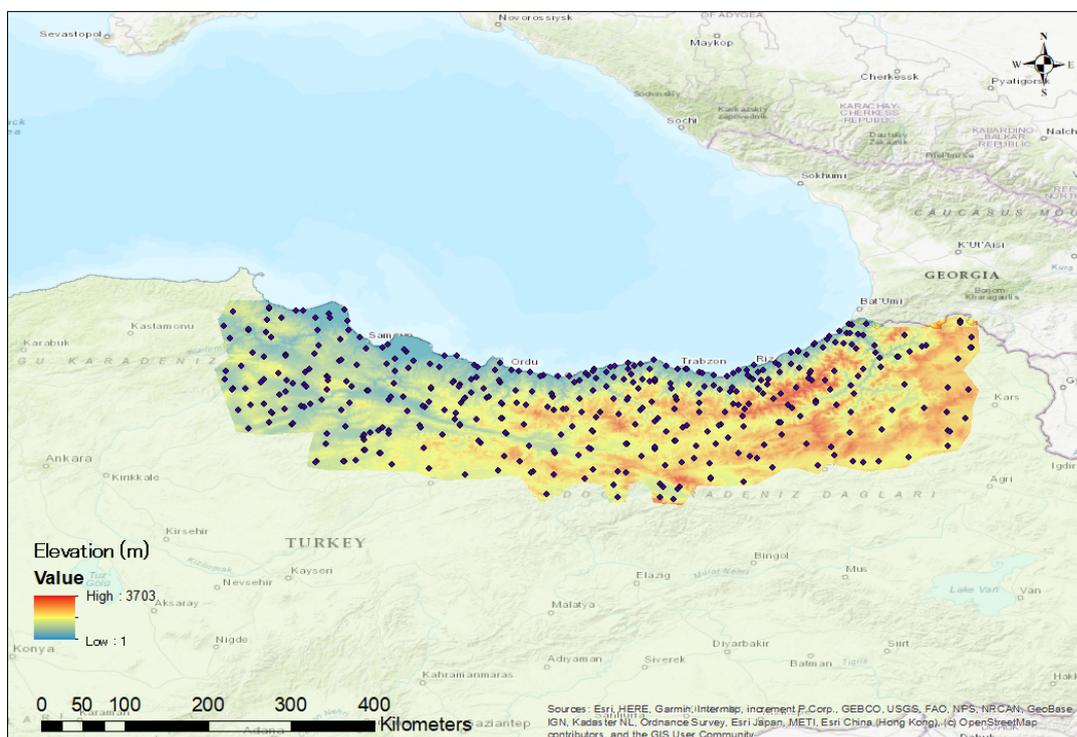


Figure 2.4: Elevation of study area. Black dots represent ground stations

2.2 DATASETS

2.2.1 Ground Observations

Daily precipitation data from ground stations are available from 01 Jan 1960 to 31 May 2020 for this study. Considering the launch date of GPM, data quality and computation time, only ground observations from 01 Jan, 2015 to 31 May, 2020 have been used in this study. There are total of 411 ground stations in the eastern Black Sea Region, but some of them have stopped running and some of them present a relevant part of the record with missing values. On average, each day has around 200 ground stations effectively running. Elevation, latitude and longitude data for those ground stations are also available. In order to make better correction on the GPM satellite estimates, observational data has been interpolated to the resolution same as GPM data (More details can be found in Methodology and Results chapters)

2.2.2 Landslide Inventories

Three landslides inventories are available for this study in the period between 01 Jan 2015 to 31 May 2020. Table 2.1 includes some simple information for those landslide inventories. In the later sections of this study, the results based on dates of those three landslides events are displayed.

Table 2.1 Landslide Inventories

Location	Date	Counts
Rize Kaptanpaşa	28-30 Sep 2017	1058
Ordu-Perşembe	05 Jul 2016	638
Artvin Hopa	24-25 Aug 2015	1367

2.2.3 NRT Satellite Precipitation Data

The Tropical Rainfall Measuring Mission (TRMM) was a joint space mission initiated in 1997 by National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), and it has carried the first satellite based precipitation radar. In 2014, Global Precipitation Measurement (GPM), the successor of TRMM, has been launched by NASA and JAXA. Compared to TRMM, GPM provides more details on the precipitation, a better spatial resolution and a larger global coverage. The spatial resolution for TRMM was 0.25 (around 25km) degree, whereas the spatial resolution of GPM is 0.01 (around 10 km) degree. Two major precipitation products are based on GPM: Integrated Multi-satellitE Retrievals for GPM (IMERG) developed by NASA and Global Satellite Mapping of Precipitation (GSMaP) developed by JAXA (H. Wang and Yong, 2020). Both IMERG and GSMaP provide NRT and gauge-calibrated precipitation data, and there are many research have compared them (Aslami et al., 2019; Nepal et al., 2021; H. Wang and Yong, 2020). In a number of studies, IMERG have been found having closer precipitation data to the rain gauges than GSMaP, this being the case for Iran (Aslami et al., 2019). And, IMERG has also been found having better performance for monitoring daily extreme precipitation data (Nepal et al., 2021). Therefore, IMERG was decided to be used in this study.

When this study started, IMERG V06B was the latest algorithm, so the data based on IMERG V06B has been used in this study. There are three runs of IMERG data: Early Run, Late Run and Final Run. The processing time for each runs are 4 hours, 14 hours and 3.5 months separately (Sakib et al., 2021). Considering the purpose that we want the improve the NRT prediction on precipitation and landslides, Final Run is not applicable in this study. There are many study have compared the performance between IMERG Early Run and Late Run, and most of them find that the difference between them is relatively small (Kawo et al., 2021; Zhou et al., 2021).

The local time in Turkey is UTC+03:00, and the ground observation precipitation is daily based. So the daily IMERG data has been accumulated by using 30 minute IMERG data plus a three-hour time gap. For example, the daily precipitation data from IMERG on 01 Jan, 2015 has been derived by accumulating 30 minute IMERG data from 31 Dec, 2014, 9:00PM to 1 Jan, 2015, 8:59PM. An accuracy assessment for this procedure has been done, and the result is in Figures 4.5d, 4.6d, and 4.7d.

2.2.4 Covariates

Digital Elevation Model (DEM) data (Hastings et al., 1999) and Normalized Difference Vegetation Index (NDVI) data (Vermote et al., 2014) can be helpful as reference data since they are related with precipitation pattern (Ranjbar et al., 2020; Tennant et al., 2017). Moreover, North-

ness and Eastness have also been calculated based on the DEM data for each grid. The resolution of reference data have been rescaled to the same resolution of GPM data.

Chapter 3

Methodology

3.1 INTERPOLATION

Many paper have compared several precipitation interpolation methods (Antal et al., 2021; Katipoğlu, 2022; Usowicz et al., 2021). However, the aim of this thesis was not to interpolate rainfall data from one rain gauge to another. For this reason, the two most common spatial interpolators have been chosen, compared and only one of the two has later been adopted to regionalize the discrete precipitation information. Specifically, Ordinary CoKriging (OCK) and Inverse Distance Weighting (IDW) have been tested in this study.

3.1.1 Ordinary Cokriging

Ordinary Kriging (OK) is one of the commonly used geostatistical approach has been widely used on climate variables interpolation (D. Chen et al., 2010). A spatial model (called variogram) contains the spatial structure that has been retrieved from known points, and a unique variance value is then assigned for each unknown point (Katipoğlu, 2022). Ordinary Cokriging (OCK) is an extension of OK, and it adopts the information from auxiliary variables for interpolating the target one. The value of unknown point can be calculated based on a semivariogram that also exploits the spatial correlation between target variable and auxiliary variables (Antal et al., 2021).

3.1.2 Inverse Distance Weighting

Inverse Distance Weighting (IDW) is a interpolation method for estimating unknown points entirely based on the distance between observation points. IDW assumes that points next to each other have more similar values, and unknown points can be calculated as a function of the distance and values between nearby points (Antal et al., 2021; Katipoğlu, 2022). As a result, values of observation points have greater weights on nearby region, and this weights are proportional to the inverse of the distance. The exponent decrease rate of weights is the only parameter needed to be set for this method, and it has been set to 2 in this study.

On the basis of the literature review, it is still difficult to say which method is more effective. Some research found that OCK outperformed IDW (Katipoğlu, 2022) since covariates bring additional information to the OCK interpolation routine. However, some research also mentioned that covariates do not often bring meaningful information and can even act as noise, thus leading to performance of OCK being worse than IDW (Antal et al., 2021).

3.2 STATISTICAL POSTPROCESSING

3.2.1 Ensemble Model Output Statistics

As mentioned in the literature review section, Ensemble Model Output Statistics (EMOS) uses various parametric distribution for modeling the relationship between predictors (estimates) and predictands (observations). And, zero-adjusted gamma distribution (ZAGA) has been proved successful for modeling precipitation relying on a variant of the gamma distribution. Figure 3.1 shows the probability density function (PDF) and cumulative distribution function (CDF) for the ZAGA distribution.

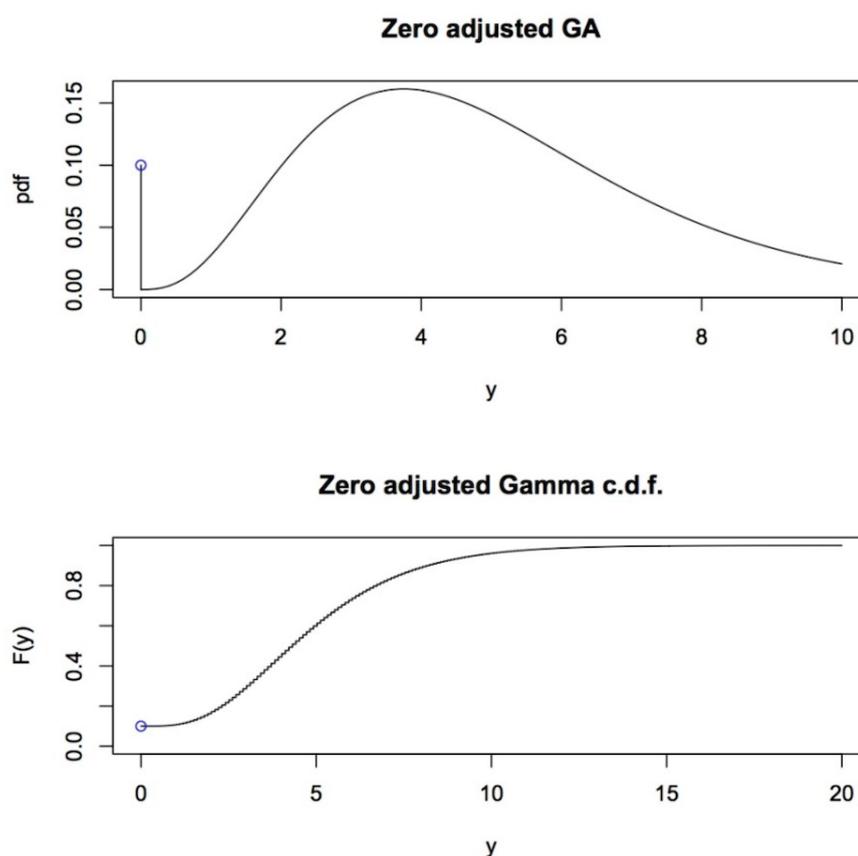


Figure 3.1: PDF and CDF for ZAGA distribution (Stasinopoulos et al., 2017)

According to the upper graph of Figure 3.1, the ZAGA distribution can be understood as a mixture of a gamma distribution and a separate probability at the value 0. There are three parameters for modeling a ZAGA distribution: μ is the location parameter, and it is usually related to the “center” of the distribution; σ is the scale parameter, and it represents the “spread” of the distribution. Larger σ indicates that the distribution is more dispersive. ν is an extra parameter specific of ZAGA, and it is equal to the probability at the value 0. When the value is larger than 0, the probability can be calculated by a gamma distribution with probability $(1-\nu)$ (Rigby et al., 2019). As figure 3.2.1 shows, the ZAGA distribution matches the precipitation pattern: many days have no rainfall (value 0), and extreme precipitation is much less likely to happen (with small

probabilities).

The probability density function for ZAGA is defined as follows:

$$f_Y(y|\mu, \sigma, \nu) = \begin{cases} \nu, & \text{if } y = 0 \\ (1 - \nu) \left[\frac{1}{(\sigma^2\mu)^{1/\sigma^2}} \cdot \frac{y^{(1/\sigma^2)-1} \cdot e^{-y/(\sigma^2\mu)}}{\Gamma(1/\sigma^2)} \right], & \text{if } y > 0 \end{cases} \quad (3.1)$$

Predictors for parameters μ , σ and ν are selected by Generalized Akaike Information Criterion (GAIC), which has been commonly used in model selection by calculating the information loss (Wagenmakers and Farrell, 2004). Strategy A of the GAIC has been applied in this research, and it starts with a forward stepwise approach for calculating the first predictor for the parameter μ . The predictor with the smallest AIC will be selected, and the number of predictors of each parameter will be discussed in the validating procedure. After the model for μ has been built, the model for σ will be built on the basis of the model for μ . The model for ν will be built later. After all predictors for all three parameters have been decided, a backward elimination procedure will be applied.

3.2.2 QRF

Random forests (RF) is a non-parametric machine learning method that does not need to assume a specific function to represent the relationship between predictors and response variables in advance (Carella et al., 2020). RF is a tree-based algorithm that consists of many decision trees together. The decision tree technique splits the observation data into two homogeneous groups for each node by some threshold of the predictor variables. The split procedure will stop when the minimum number of observations for each node has reached, and all observation data will be distributed to a set of terminal nodes (van Straaten et al., 2018). Modeling with only one decision tree can induce large variances, so multiple trees are grown together on bootstrapped samples of the original data. The average of different trees has been calculated (also called bagging). To make trees more independent, RF was proposed in 2001. In the RF, each split of each tree can be built on a random subset of the predictors (Breiman, 2001). In conclusion, RF is a relatively flexible model which can be set with a different number of trees (ntree), minimum number of the observation for each node (nodesize), and number of variables randomly sampled at each split (mtry). QRF is an extension of random forest that can be used to perform a quantile regression and generate a probabilistic forecast (Schulz and Lerch, 2022). Unlike RF, which only estimates the response variable's mean at each terminal node, QRF delivers the full CDF. The final forecast using QRF is calculated by averaging the CDF from all nodes of all trees. More details about how QRF can be applied to statistical postprocessing can be found in Taillardat's paper (Taillardat et al., 2016).

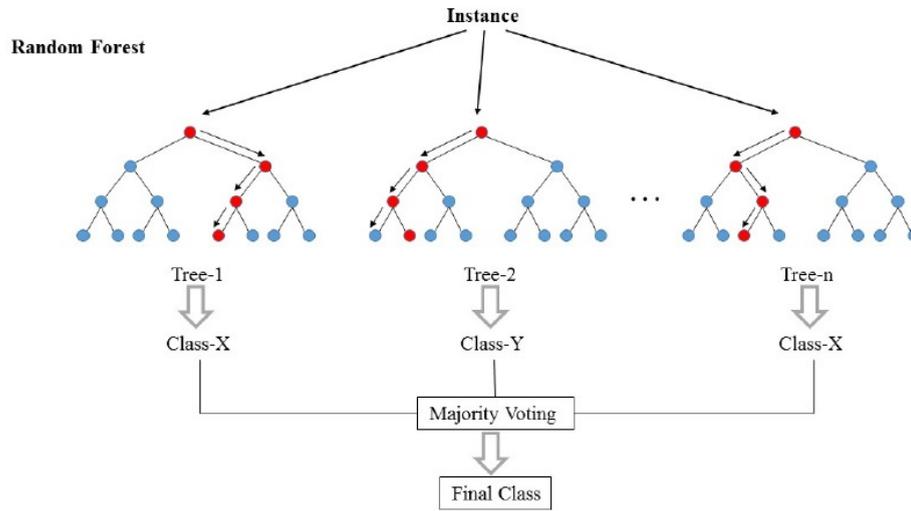


Figure 3.2: An example of random forests (Dimitriadis et al., 2018)

3.3 MODEL JUSTIFICATION

The predictors' selection in this study is based on a previous study (Y. Zhang et al., 2022), to which additional covariates have been added to bring further information on the landscape characteristics, hence on the orographic effect on rainfall patterns. Table 3.1 shows all selected predictors for training the model. Overall, six covariates have been selected (Lon, Lat, Elevation, Eastness, Northness, NDVI), and precipitation data on the 20 closest grids of the target grid of the target day and the previous have also been selected. The reason for taking precipitation data from such many grids is that the difference between IMERG and interpolated data is large, and more features can be captured by the model with more predictors.

The reason behind such predictor choice is briefly summarized as follows: elevation can be responsible to create an orographic barrier to the incoming clouds. In the specific case of this test site, the clouds come from the north, as a result of evaporated and re-condensed water from the Black Sea. Therefore, Eastness and Northness can partially explain local rainfall as for instance, North facing slopes may receive more precipitation than South facing ones at higher altitudes. As for the inclusion of the NDVI signal 10 days before each examined day, this predictor was selected in the hope of capturing the evapotranspiration effect of the vegetation, under the assumption that denser forests would release more moisture in the air, which can be converted into rain. Finally, the choice of the 20 closest grids is meant to bring the information of the neighboring areas, adding more spatial coherence to the model.

Considering there are five and a half years ground stations data are available (2015.1.1-2020.5.31), four years of data (2015.1.1-2018.12.31) have been used to perform a four-fold cross validation. After the validation procedure, the final model used four years of data (2015.1.1-2018.12.31) for training and one and a half year of data (2019.1.1-2020.5.31) for testing (Table 3.2). Since precipitation in different seasons usually show different features, and statistical postprocessing on climate variables need to consider the seasonal effects (Ratri et al., 2021; van Straaten et al., 2018). All data have been divided into three groups: November - February (rainfall season), July-October (land-

slide prone season) and March-June (others). Models for each season have been trained and tested independently.

Table 3.1 Predictors Selection

Variable Name	Abbreviation	Unit
Longitude	Lon	/
Latitude	Lat	/
Elevation	Elevation	m
Aspect (Eastness)	Eastness	/
Aspect (Northness)	Northness	/
Normalized Difference Vegetation Index	NDVI	/
Precipitation of the grid with the minimal distance (target grid) on the target day	min1	mm/day
Precipitation of the grid with the minimal distance (target grid) on the previous day	minP1	mm/day
Precipitation of the grid with the 2nd minimal distance to the target grid on the target day	min2	mm/day
Precipitation of the grid with the 2nd minimal distance to the target grid on the previous day	minP2	mm/day
...
Precipitation of the grid with the 20th minimal distance to the target grid on the target day	min20	mm/day
Precipitation of the grid with the 20th minimal distance to the target grid on the previous day	minP20	mm/day

Table 3.2 Training, validation and testing groups

	2015	2016	2017	2018	2019	2020
Fold1	Train	Train	Train	Validation	/	/
Fold2	Train	Train	Validation	Train	/	/
Fold3	Train	Validation	Train	Train	/	/
Fold4	Validation	Train	Train	Train	/	/
Final	Train	Train	Train	Train	Test	Test

3.4 VERIFICATION

3.4.1 Model Comparison

Three commonly used model comparison methods have been carried out in this study: mean error (ME), root mean square error (RMSE) and Pearson correlation coefficient (PCC). (3.2) to (3.4) are equations for calculating those three indicators. Smaller RMSE and ME indicate the estimate/predict values are much closer to the observational data. PCC is for measuring the relationship between estimate/predict values and observational data. When PCC equals to 0, there are no relationship between two values. When PCC equals to 1 or -1, two values has a perfectly positive or negative linear relationship. Value of PCC between 0 to 0.3 (0 to -0.3) represents a weak linear relationship, between 0.3 to 0.7 (-0.3 to -0.7) represents a moderate linear relationship, and between 0.7 to 1 (-0.7 to -1) represents a strong linear relationship (Ratner, 2009).

$$ME = \frac{1}{n} \sum (y_i - x_i) \quad (3.2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - x_i)^2} \quad (3.3)$$

$$PCC = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (3.4)$$

3.4.2 Model Assessment

Brier Score (BS) and Reliability Diagram are two widely used methods for assessing accuracy, reliability and sharpness of statistical postprocessing models. Different from model comparison methods, BS and reliability diagrams are more focused the performance of the models themselves.

Brier Score

The BS is a strictly proper score that can measure the accuracy of the model, and the formula of the BS is shown in (3.5). The brier score calculates the mean square error (MSE) of different forecast probabilities for binary events o_i with a threshold ($o_1 = 1$ if the event occurs and $o_2 = 0$ if the event does not occur). n indicates the verification set size. The brier score is negative-oriented, and it can be converted to the brier skill score (BSS) for showing the forecast skills according to the reference climatology (3.6).

$$BS = \frac{1}{n} \sum_{i=1}^n (p_i - o_i)^2 \quad (3.5)$$

$$BSS = 1 - \frac{BS_{mod}}{BS_{ref}} \quad (3.6)$$

Reliability diagram

BS can give a simple initial impression, and a reliability diagram can show the full joint distribution of NRT satellite data and observations. The reliability diagram is a well-designed graphical device that provides reliability, resolution, and sharpness information for the forecast. More introduction about reliability diagram can be found in (Wilks, 2011), and some typical examples of reliability diagrams are shown in the Figure 3.3a. The observed relative frequency is plotted against the forecast estimate probability in the reliability diagram. The center panel in Figure 3.3a is the most well-calibrated forecast among the five diagrams since the observed relative frequency

of each data group is relatively matching with the forecast probability from the model. The top panel in Figure 3.3a is an example of overestimation (or overforecasting), in which the observed relative frequency is relatively lower than the forecast probability. The bottom panel in the figure is the opposite of the top one, and it corresponds to underforecasting. The left and right panels in the Figure 3.3a are related to the panels in the Figure 3.3b. The left one is underconfident with good resolution, and the observed outcomes change obviously with the forecast change. On the other hand, the right panel is overconfident with poor resolution, and the observed outcomes will not change obviously applying a forecast change. The reliability improvement can be achieved by adjusting the extreme probabilities to be less extreme. The equation (3.5) can be expressed in the equation (3.7), including reliability, resolution, and uncertainty. Since the sign for the part of the resolution is negative, the poor resolution can increase the BS and decrease the BSS. The sharpness diagram is also presented with the reliability diagram simultaneously, indicating the tendency to forecast the extreme probabilities rather than the mean probabilities.

$$BS = \frac{1}{n} \sum_{i=1}^I N_i (y_i - \bar{o}_i)^2 - \frac{1}{n} \sum_{i=1}^I N_i (\bar{o}_i - \bar{o})^2 + \bar{o}(1 - \bar{o}) \quad (3.7)$$

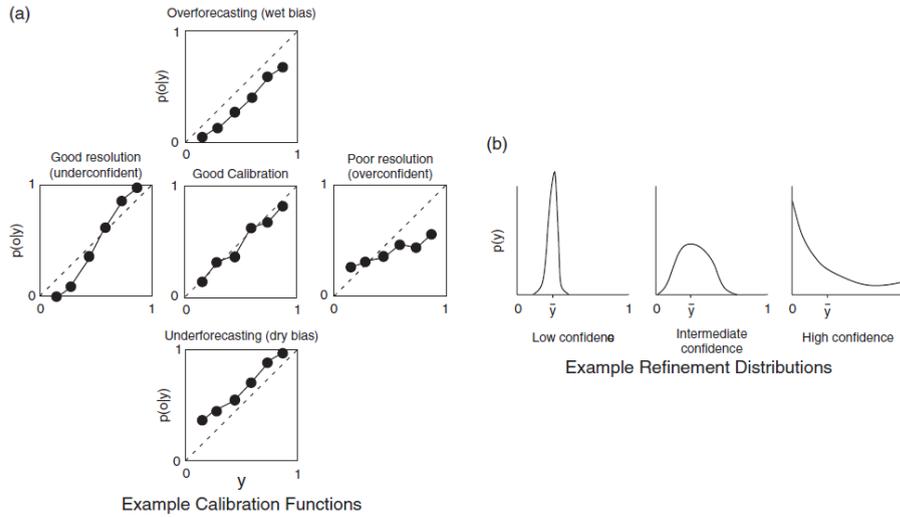


Figure 3.3: Examples of reliability diagram (Wilks, 2011)

The Brier score and reliability diagram have been widely used in the statistical postprocessing studies for weather forecasting (Whan and Schmeits, 2018). The basic principles of statistical postprocessing on either weather forecasting or NRT satellite estimates is similar, since the NRT satellite estimate couldn't reflect the actual rainfall situation accurately at this stage.

3.4.3 General overview

A summary of the methods described above and how they are interconnected into a single analytical protocol are shown in figure the Figure 3.4.

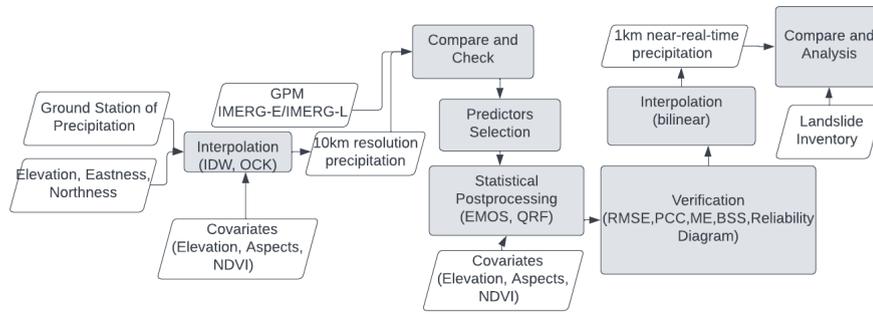


Figure 3.4: Flowchart

Chapter 4

Results

4.1 RAIN GAUGE ASSESSMENT

In the study area, 177 ground stations have run since 01 Jan, 2015 and up to 282 ground stations have run until 31 May, 2020. Figure 4.1 shows the distribution of ground stations on 01 Jan, 2015 and 31 May, 2020, separately. Precipitation interpolation in mountainous region is more complex than it in plain area (Katipoğlu, 2022) since the intensity of precipitation changes with the change of altitude (Dimri et al., 2022). After more ground stations have been built in this area, a better accuracy of the “true” precipitation data is expected, as the investment from the Turkish administrations deployed new stations also on highly elevated slopes.

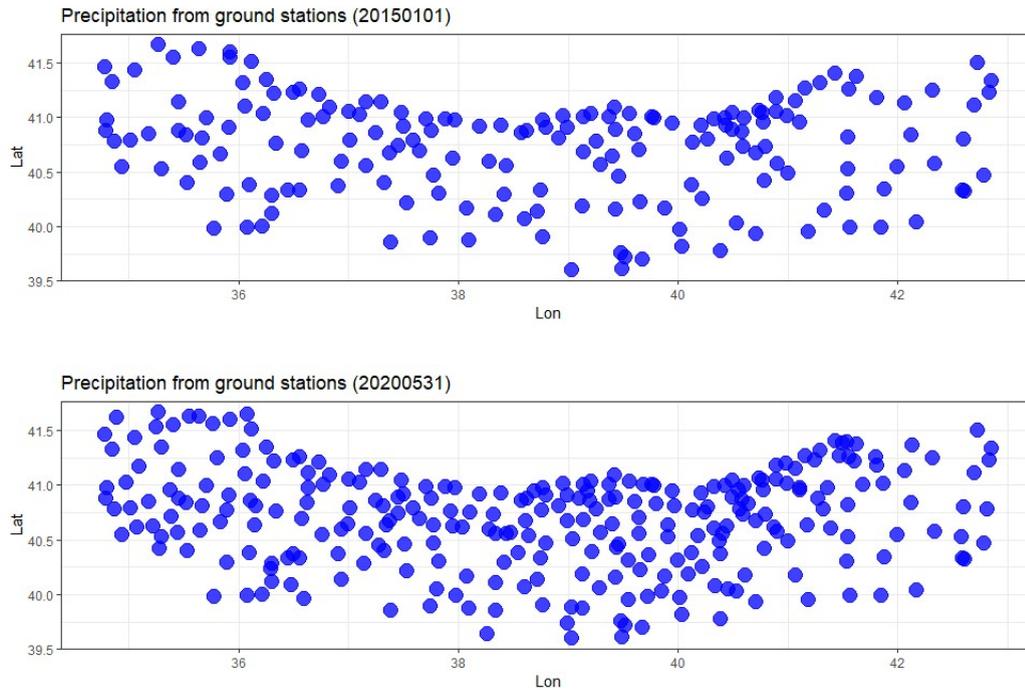


Figure 4.1: Ground Stations Distribution. Each blue point represents a ground station

The subsequent section will present an overview of how the discrete information collected at the rain gauges has been regionalized in a continuous manner over the study area.

4.2 INTERPOLATION

Two commonly used spatial interpolation methods have been used in this study: Inverse Distance Weighting (IDW) and Ordinary CoKriging (OCK). Daily rainfall data have been interpolated by both methods to the resolution same as GPM-IMERG data (0.1 degree), and here shows three sets of comparison results (Figure 4.2-4.4) based on the date of a selected landslide inventory (24 Aug 2015, 05 July 2016 and 28 September 2017). According to those three figures, both interpolation methods are able to show the precipitation pattern from the ground stations. However, figure 4.3 implies that OCK may miss some extreme precipitation features (the maximum precipitation from ground stations is larger than 120mm, but the maximum precipitation showed in the OCK interpolation map is just 27.78mm). Extreme precipitation is vital to explain landslide occurrences. Therefore, we opted to use the results from IDW interpolation for the later analysis. Specifically, the interpolated rainfall in space and time will be used as the baseline to which GPM estimates will be projected to, on the basis of a machine learning based, bias correction procedure.

4.3 IMERG ASSESSMENT

Though IMERG has been applied in several RTL susceptibility analysis (LaJoie et al., 2021; Titti et al., 2021), there are still large error involved with IMERG NRT data (Gilewski and Nawalany, 2018; Nguyen et al., 2018). Daily data on the dates of three landslides occurred from early run (IMERG-E) and late run (IMERG-L) have been compared (Figure 4.5a/b, 4.6a/b, 4.7a/b). It can be seen that the difference between IMERG-E and IMERG-L are relatively small. To buy more time for landslide prediction, IMERG-E has been selected as the target for our bias-reduction attempt. As mentioned in section 2.2.3, there are three hours difference between UTC and Turkey Time (TRT). Therefore, the daily IMERG-E data has been recalculated by using 30 minutes data on the basis of TRT (see Figure 4.5d, 4.6d, 4.7d). Compared to the observational data from ground stations (4.5c, 4.6c, 4.7c), errors between ground stations and IMERG-E are relatively large. Considering the difference between observational data and IMERG-E, the statistical post-processing routine we implemented was framed on the same resolution of IMERG-E (around 0.1 degree). In other words, the IDW interpolation described in the previous section was constrained to match the same resolution of IMERG-E and the following bias-reduction will act at this scale.

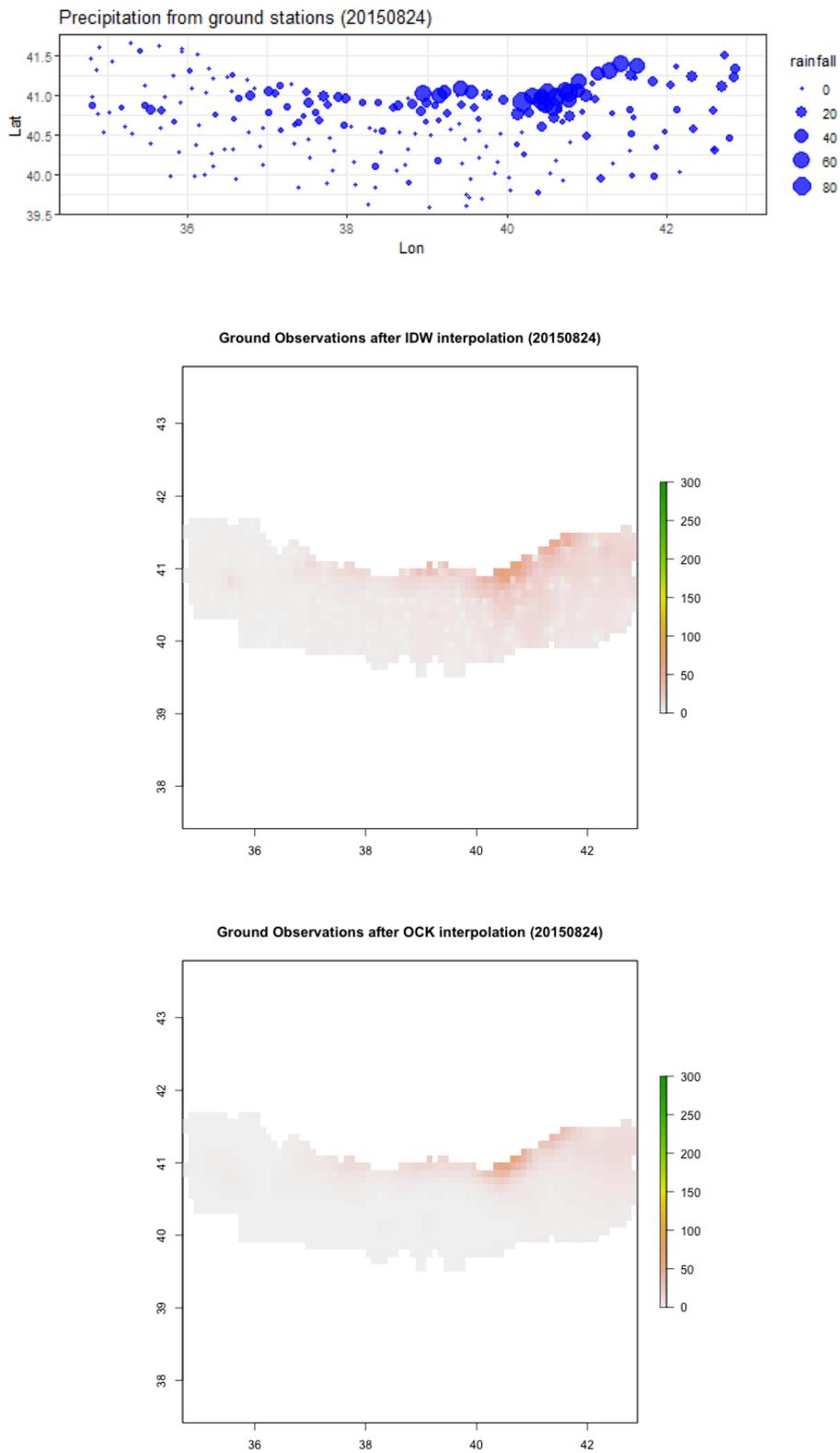


Figure 4.2: Interpolation comparison for landslide Artvin Hopa (units: mm/day)

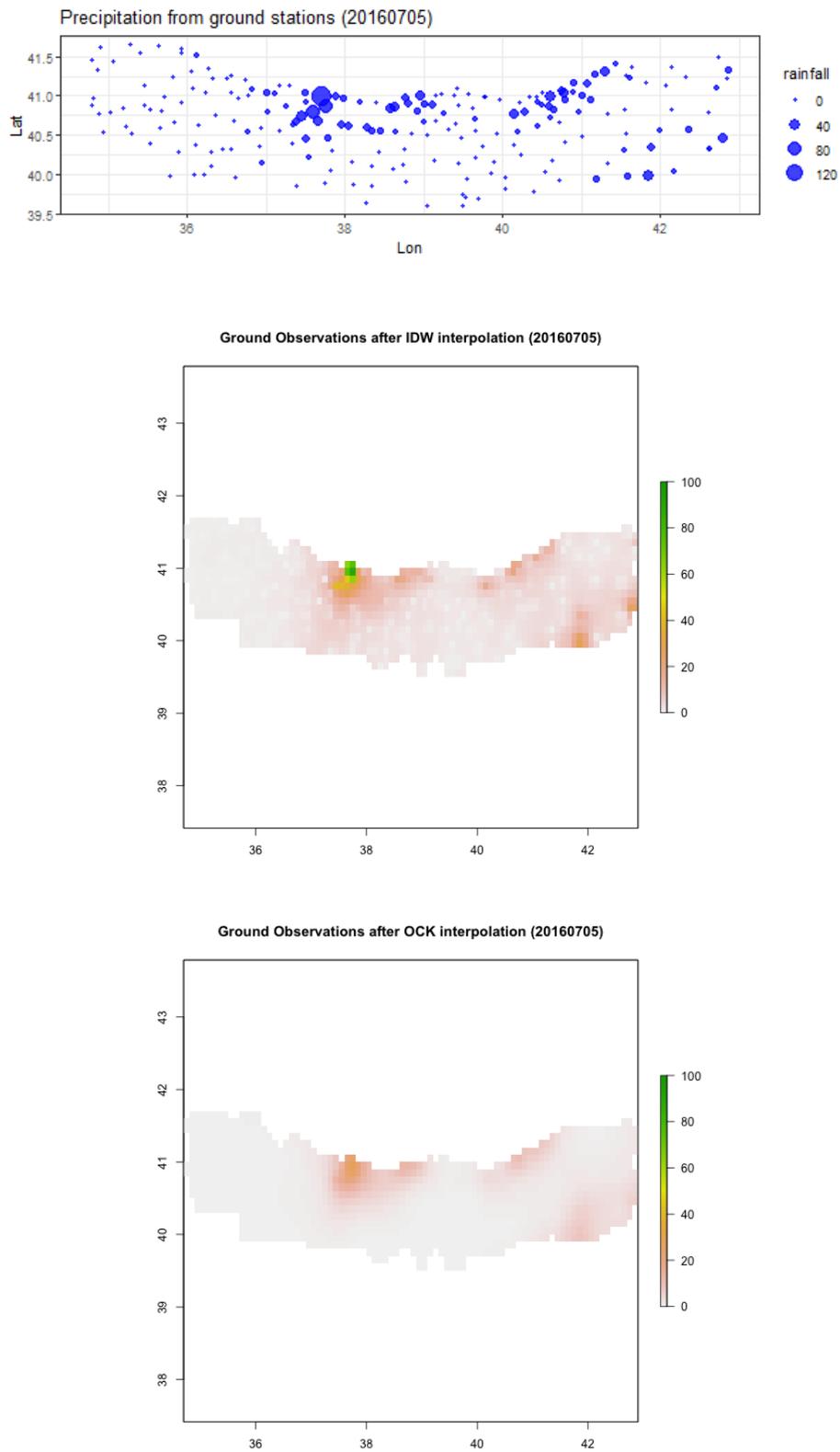


Figure 4.3: Interpolation comparison for landslide Ordu-Perşembe (units:mm/day)

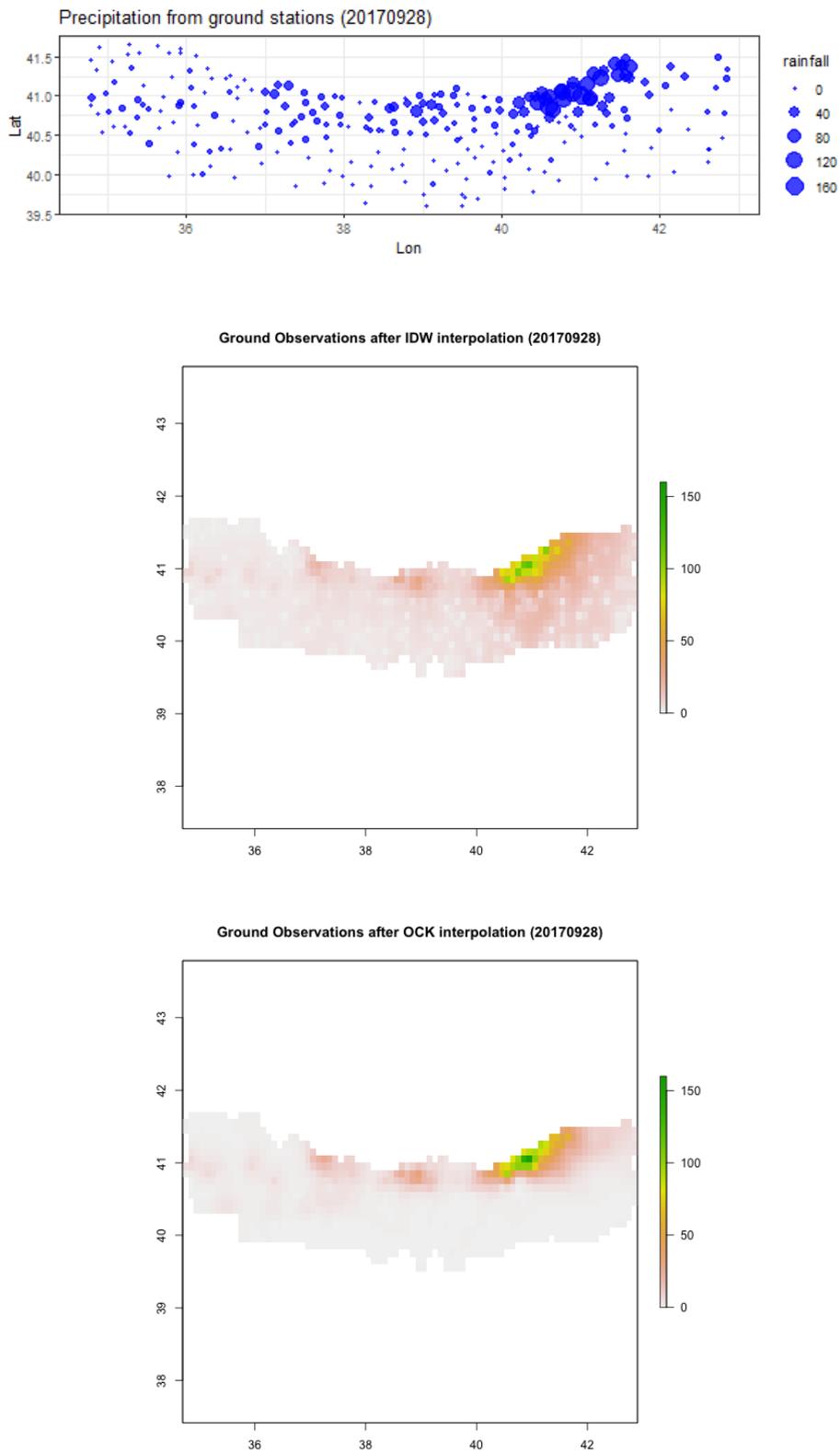


Figure 4.4: Interpolation comparison for landslide Rize Kaptanpaşa (units: mm/day)

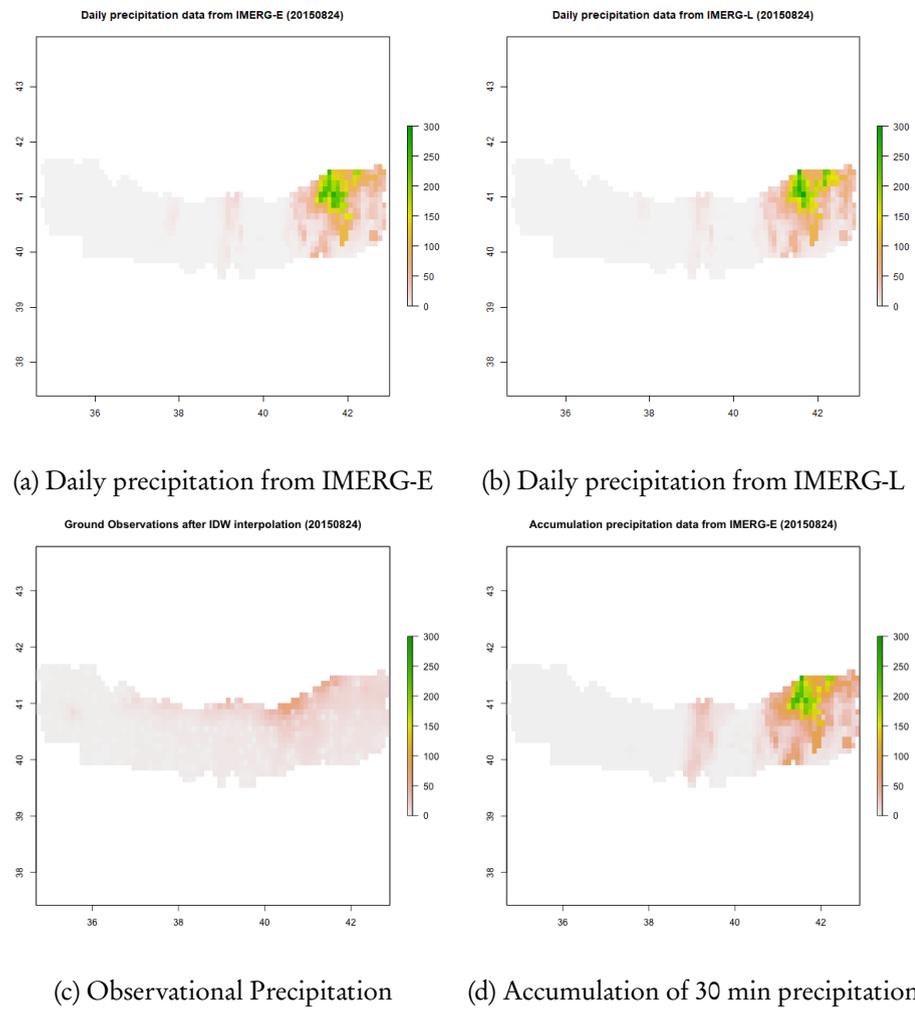


Figure 4.5: IMERG comparison for landslide Artvin Hopa (units: mm/day)

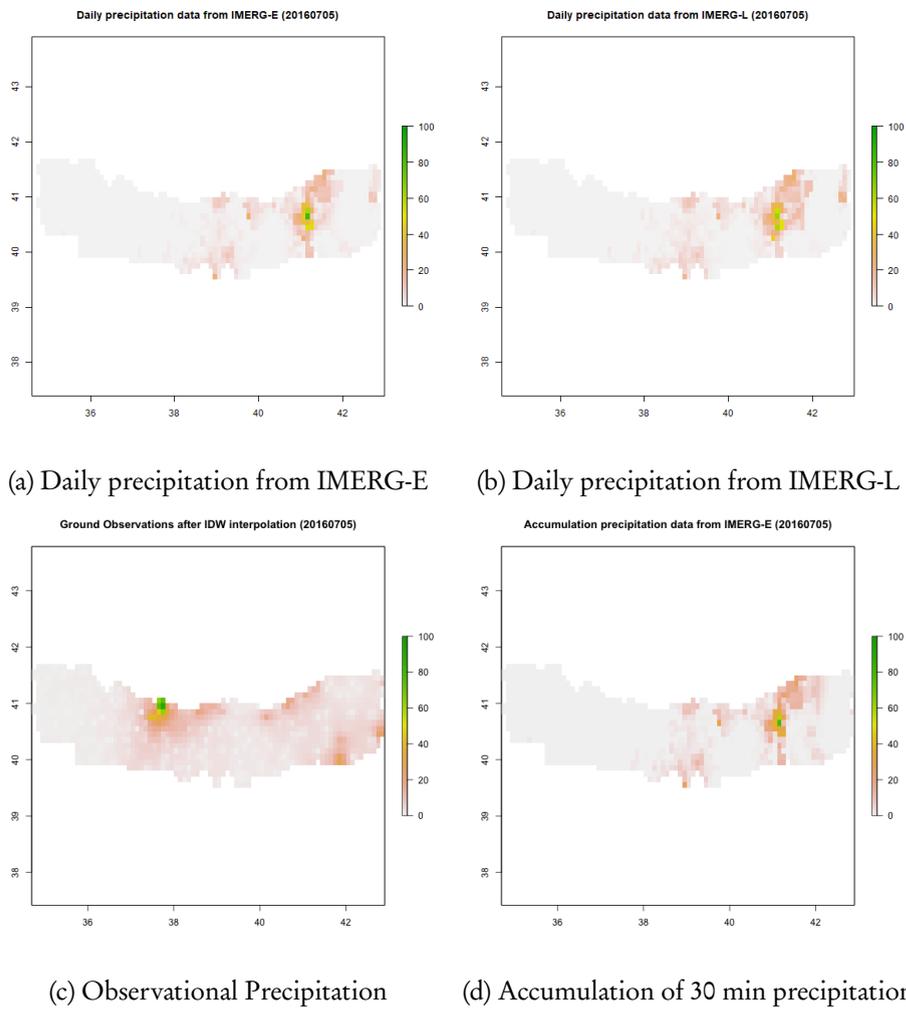


Figure 4.6: IMERG comparison for landslide Ordu-Perşembe (units: mm/day)

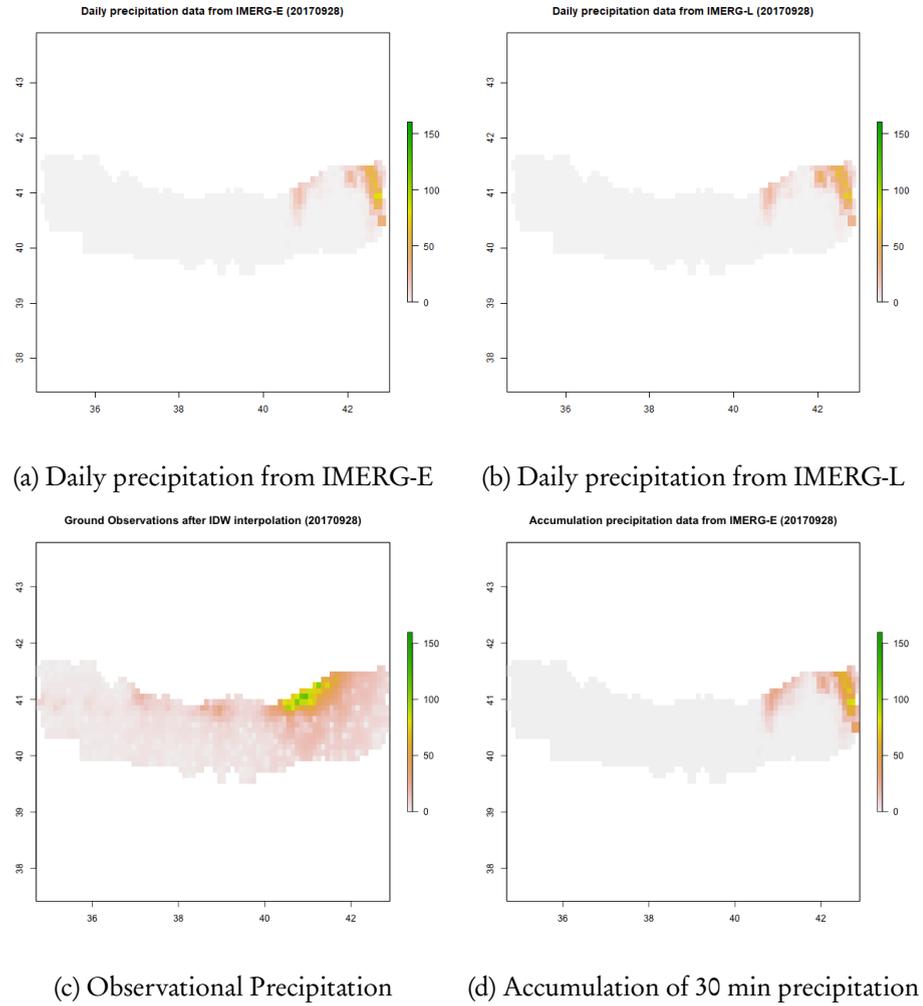


Figure 4.7: IMERG comparison for landslide Rize Kaptanpaşa (units: mm/day)

4.4 MODEL VALIDATION

The comparison between RMSE of observational data and IMERG-E, QRF prediction and ZAGA prediction have been shown in the table 4.1. For each validation fold, QRF has significantly improved the results for all folds in all seasons comparing to the raw IMERG-E values. However, ZAGA couldn't improve the estimates from IMERG-E and have larger RMSE for all seasons. By checking the predict precipitation values from the model ZAGA, some output values are extreme large and they are unrealistic. More discussion about the behaviours difference between QRF and ZAGA can be found in the discussion chapter. Considering the results in the validation procedure, only the model QRF has been decided to use in the testing procedure.

Table 4.1 RMSE comparison for validation folds (mm/day)

	IMERG-E	QRF	ZAGA
Nov-Feb			
Fold1	7.942166	4.321199	1351.912
Fold2	6.896297	4.166793	/
Fold3	7.561765	5.323488	/
Fold4	6.018691	4.40065	/
Mar-Jun			
Fold1	7.778902	3.566019	22.1255
Fold2	6.529668	4.014457	/
Fold3	7.977819	3.997963	/
Fold4	6.009583	3.541323	/
Jul-Oct			
Fold1	7.685115	4.191181	3161.067
Fold2	6.489428	3.551965	/
Fold3	9.053709	4.029116	/
Fold4	7.795569	3.749676	/

4.5 MODEL TESTING

4.5.1 Model Comparison

During the testing procedure, four years of data have been used for training (2015-2018), and the rest of data have been used for testing (2019-2020). Since the testing groups are kept independent from other data in the validation procedure, the results here are of particular relevance because they can inform on what one can expect when predicting future rainfall patterns. RMSE and PCC have been calculated for all data on different seasons (table 4.2), and the results here are similar to the validation procedure. QRF have significantly improved raw IMERG-E for all seasons according to the RMSE values. As for PCC, the raw IMERG-E for all seasons have relatively poor performance, and the relationship between the observational data and the IMERG-E estimates are weak before the postprocessing. After the postprocessing, by using QRF, the relationship between observational and predict data change to moderate according to the rules of PCC (Ratner, 2009).

The distribution of RMSE, PCC, and ME comparison of observational data and raw IMERG-E value or the predictions from QRF model have been shown in Figure 4.8-4.10. Figure 4.9 indicates that the predict values and observational data have stronger relationship in the eastern part of the study area rather than the western part.

Table 4.2 RMSE and PCC comparison for the testing group (units of RMSE: mm/day)

Nov-Feb		
	IMERG	QRF
RMSE	7.579431	4.021487
PCC	0.1667672	0.3856719
Mar-Jun		
RMSE	9.078529	3.603529
PCC	0.1436403	0.3278707
Jul-Oct		
RMSE	8.725887	3.820046
PCC	0.09079039	0.3075987

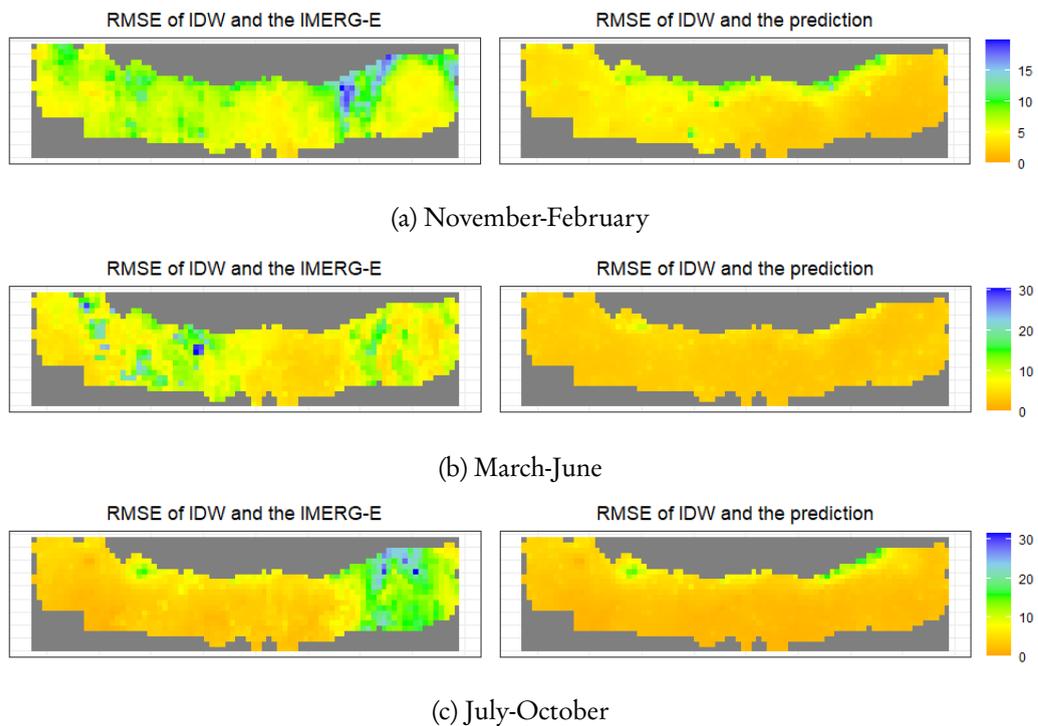


Figure 4.8: RMSE comparison (units: mm/day)

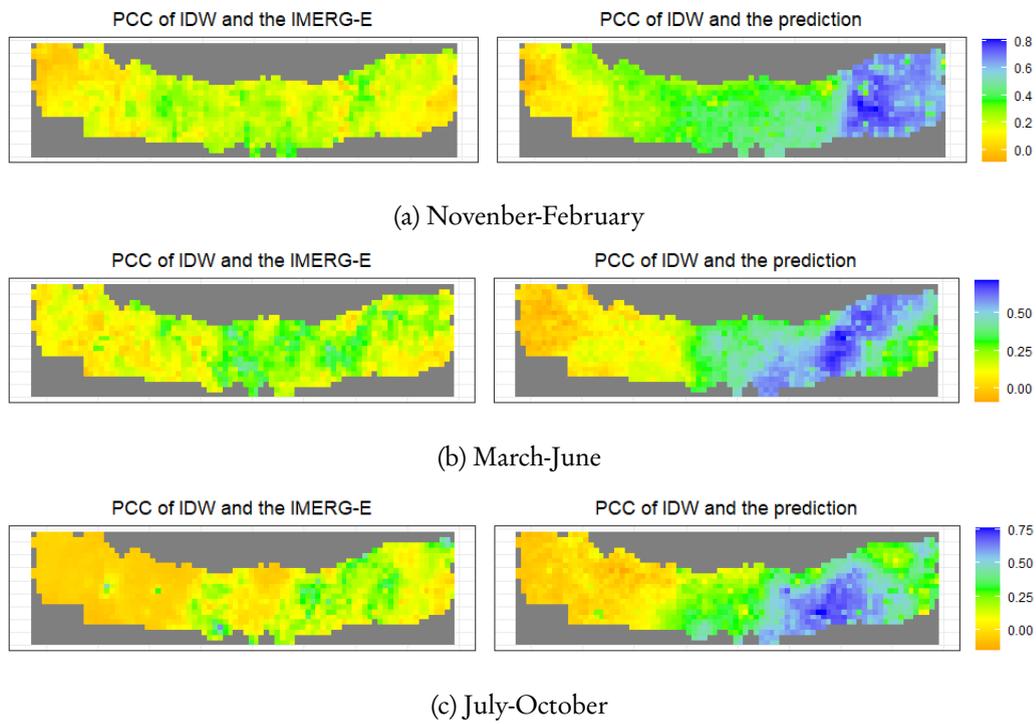


Figure 4.9: PCC comparison

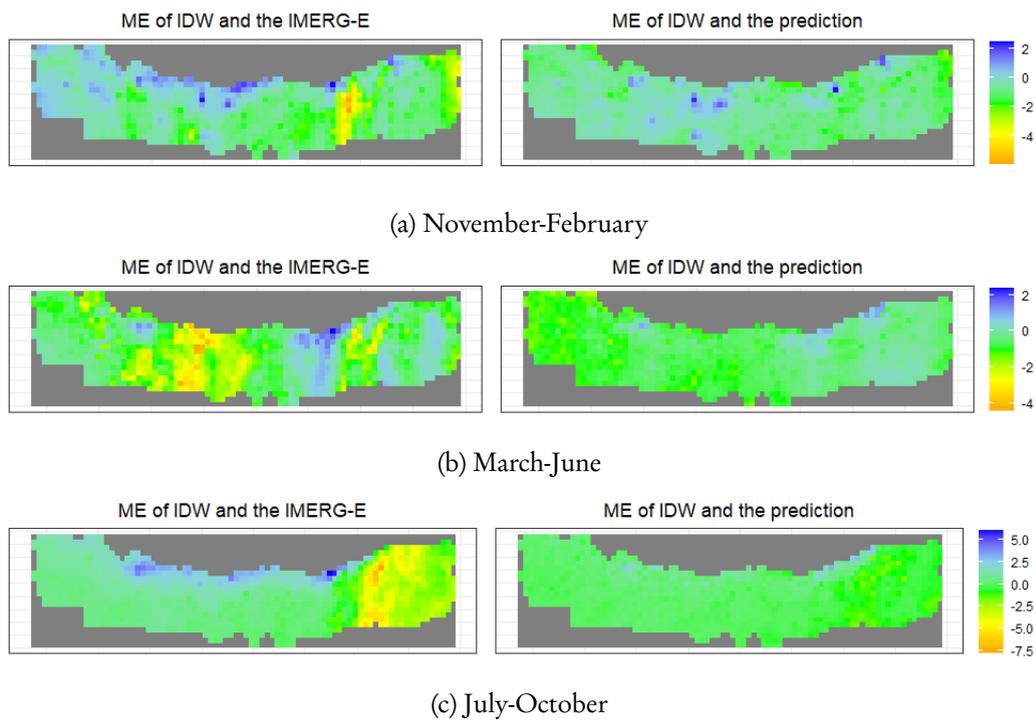


Figure 4.10: ME comparison (units: mm/day)

4.5.2 Model Assessment

After demonstrating that the predicted results have better performance than the raw IMERG-E data, information about the models themselves have also been checked. Figure 4.11 reports BSS for three models with different thresholds. All of them drop abruptly before reaching 20mm, and the BSS calculated for all of them are relatively low. Notably, higher BSS value indicate better performance. This implies that even though the statistical postprocessing models has improved the raw IMERG-E estimates, the predicted values are still not close to the observational data.

More information about models can be collected from reliability diagrams (4.12-4.14). These graphs can be read by assuming that a good performance would align data across a theoretical 45 degree line. Such lines reports on the y-axis the observed frequency of a given rainfall event discharging a total daily rainfall depending on the plot itself. Then the abscissa reports the forecast probability, which should ideally follow the observed frequency. At low thresholds (0.3mm & 1mm), models of all seasons show good calibration. However, the performance slightly worsen reaching a medium threshold (10mm). And, all models even stops issuing higher probabilities for higher thresholds (50mm), which in turn mean that the prediction is not capable of reflecting extreme rainfall events. A sharpness diagram has also been showed at the corner of each reliability diagram. Sharpness diagram represents the relative frequency of forecasts falling into different probability bin (Liu et al., 2017). From a medium threshold closer to 10mm, most of forecasts are falling into the low forecast probability, so the forecast systems only have little sharpness.

BSS for different thresholds

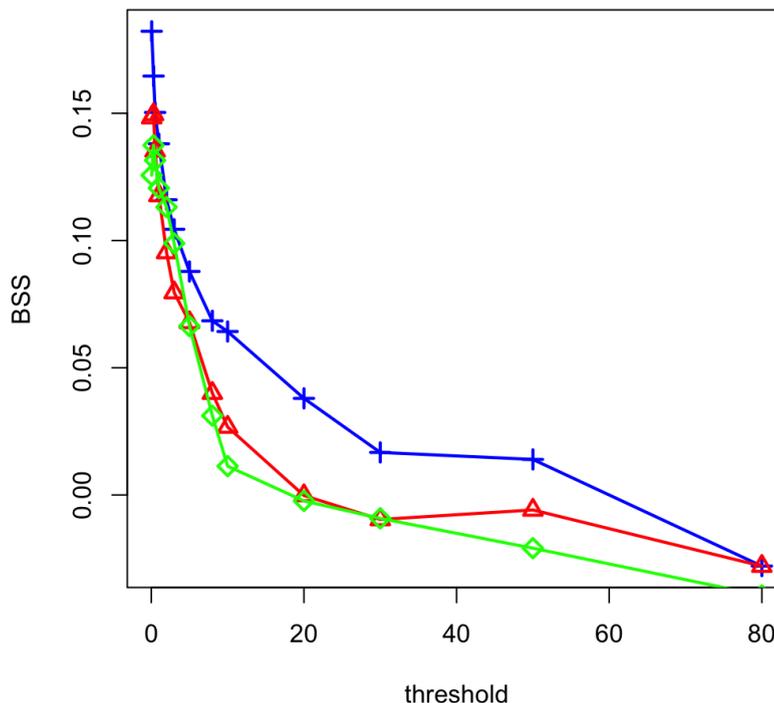


Figure 4.11: Brier skill score for three models with different thresholds (blue is for Nov-Feb, red is for Mar-Jun, green is for Jul-Oct, the units of thresholds are mm/day)

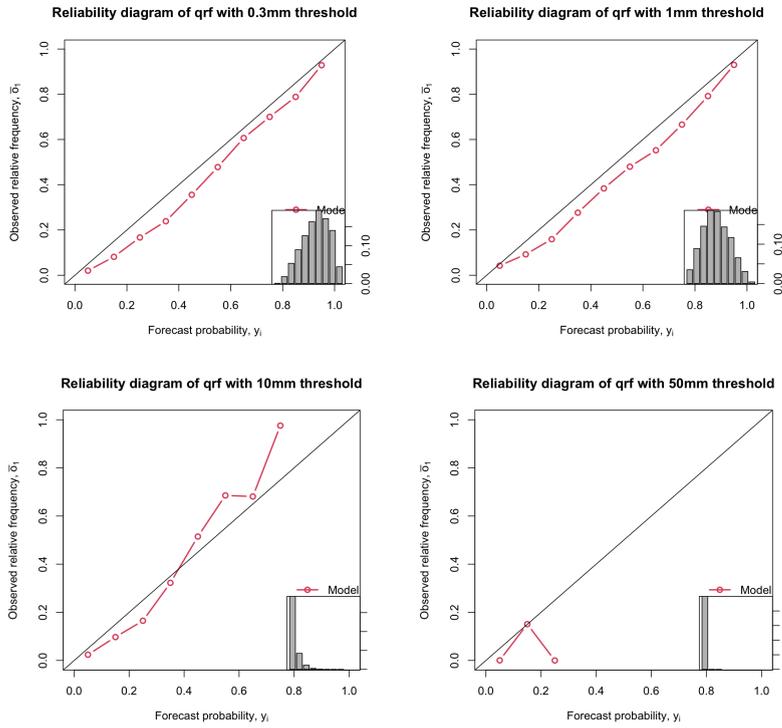


Figure 4.12: Reliability Diagram for multiple daily rainfall thresholds (Nov-Feb)

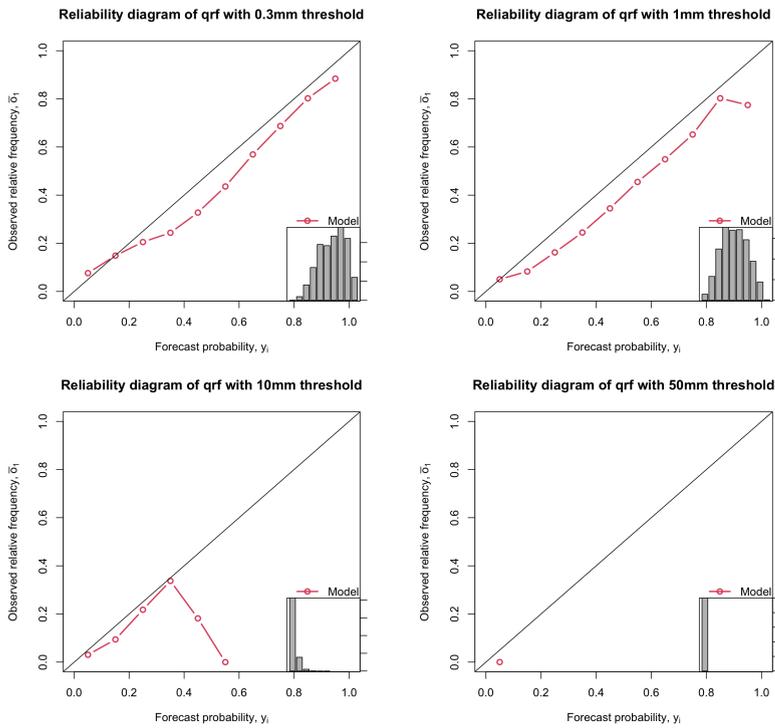


Figure 4.13: Reliability Diagram for multiple daily rainfall thresholds (Mar-Jun)

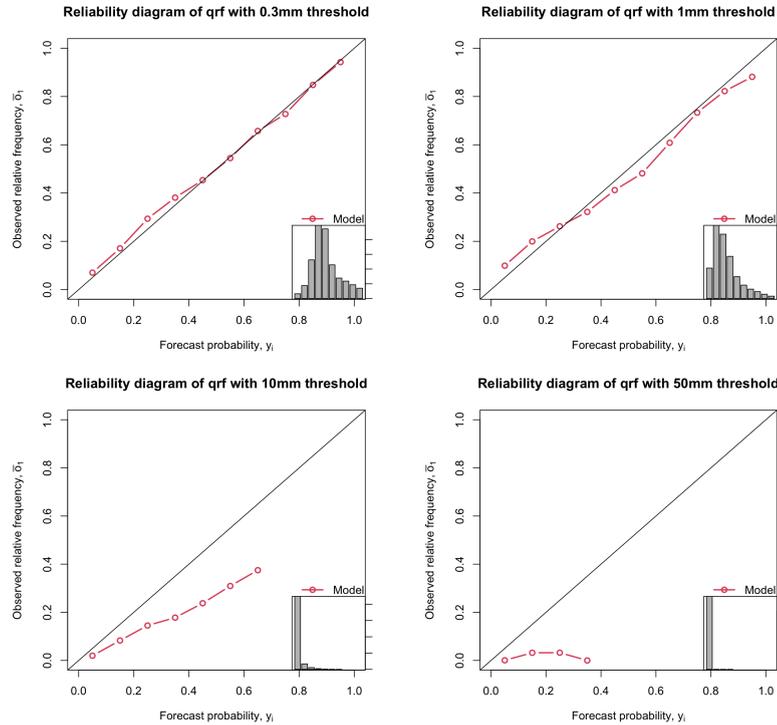


Figure 4.14: Reliability Diagram for multiple daily rainfall thresholds (Jul-Oct)

4.5.3 Predictor Importance

The importance of all predictors is also helpful to understand the model and how certain predictions are produced. When a predictor has higher importance, it has more weights and plays as a more important role in the model. Table 4.3 lists the five most important predictors for each model. NDVI seems to be one of the most important predictor for models in all seasons. Elevation data are more important than the location information during July to October (landslide prone season), and this matches the precipitation types in this region (Türkoglu et al., 2003). This is extremely relevant because it supports the validity of the model itself. In fact, during the summer, prolonged warm days induce high evaporation from the Black Sea surface and whenever the wind pushes the humid air toward the study area to the south, a strong orographic effect is expected to control the rainfall pattern and amounts (Türkoglu et al., 2003).

The precipitation data from the previous days also bring some interesting information. In fact, it seems more important than the precipitation data from the target day, and this finding is consistent with others (Y. Zhang et al., 2022). In addition, the grids are closer to the target grid do not seem more important than farther grids, and this is probably due to the bad performance of IMERG-E and a residual spatial effect that the model then captures from the neighboring structure.

Table 4.3 The five most important predictors

	1st	2nd	3rd	4th	5th
Nov-Feb					
Fold1	NDVI	Lon	Elevation	Lat	Northness
Fold2	Lon	NDVI	Elevation	Lat	minP18
Fold3	Lon	NDVI	Elevation	Lat	minP15
Fold4	NDVI	Lon	Elevation	Lat	Eastness
Testing	Lon	NDVI	Elevation	Lat	Northnessz
Mar-Jun					
Fold1	NDVI	Lon	Elevation	Lat	Eastness
Fold2	NDVI	Lon	Elevation	Eastness	minP15
Fold3	NDVI	Lon	Elevation	minP15	Eastness
Fold4	NDVI	Lon	Elevation	minP10	minP15
Testing	NDVI	Lon	Elevation	minP20	minP16
Jul-Oct					
Fold1	NDVI	Elevation	Lon	Lat	minP13
Fold2	NDVI	Elevation	Lon	minP15	Lat
Fold3	NDVI	Elevation	Lon	minP20	minP16
Fold4	NDVI	Elevation	Lon	minP20	minP8
Testing	NDVI	Elevation	Lon	minP20	minP16

4.6 ATTEMPT TO EXPLAIN LANDSLIDE OCCURRENCES THROUGH RAINFALL PATTERNS

The last and the most important step for this study is to map observational precipitation, IMERG-E estimates and the predicted (bias-corrected) precipitation in map form. In order to match the size of landslide better, the resolution of precipitation from all maps have been down-scaled to 1km by a “bilinear” interpolator. Comparative maps of the dates when the three landslide events happened have been selected and shown in Figures 4.15-4.17. Black polygons in maps are landslide inventories. Precipitation information for all maps have been classified into three groups based on the natural breaks on ArcGIS: red represents high precipitation, orange represents medium precipitation and yellow represents low precipitation. The three maps show that QRF models are able to correct most of the precipitation errors from IMERG-E estimates and the output of the model can offer useful information for landslide susceptibility prediction in near-real-time. The result in the figure 4.16 is especially exciting. Raw IMERG-E map shows that there is no precipitation in the landslide region, although this is clearly wrong according to the observational precipitation. The predicted results from QRF model has the capacity of correcting this error to a certain extent, which is something promising to extent the use of bias corrected IMERG-E data in near-real-time.

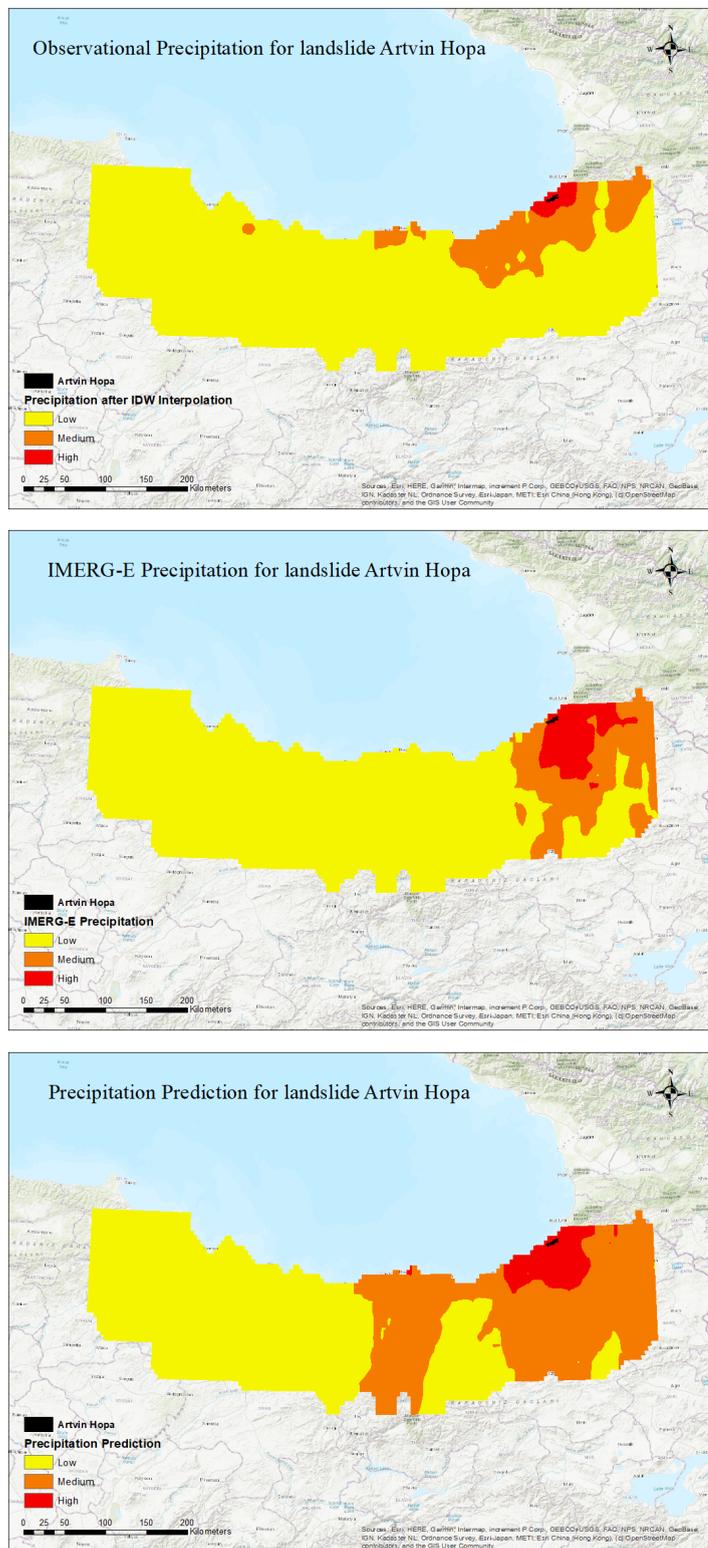


Figure 4.15: Prediction for precipitation/landslide: Artvin Hopa

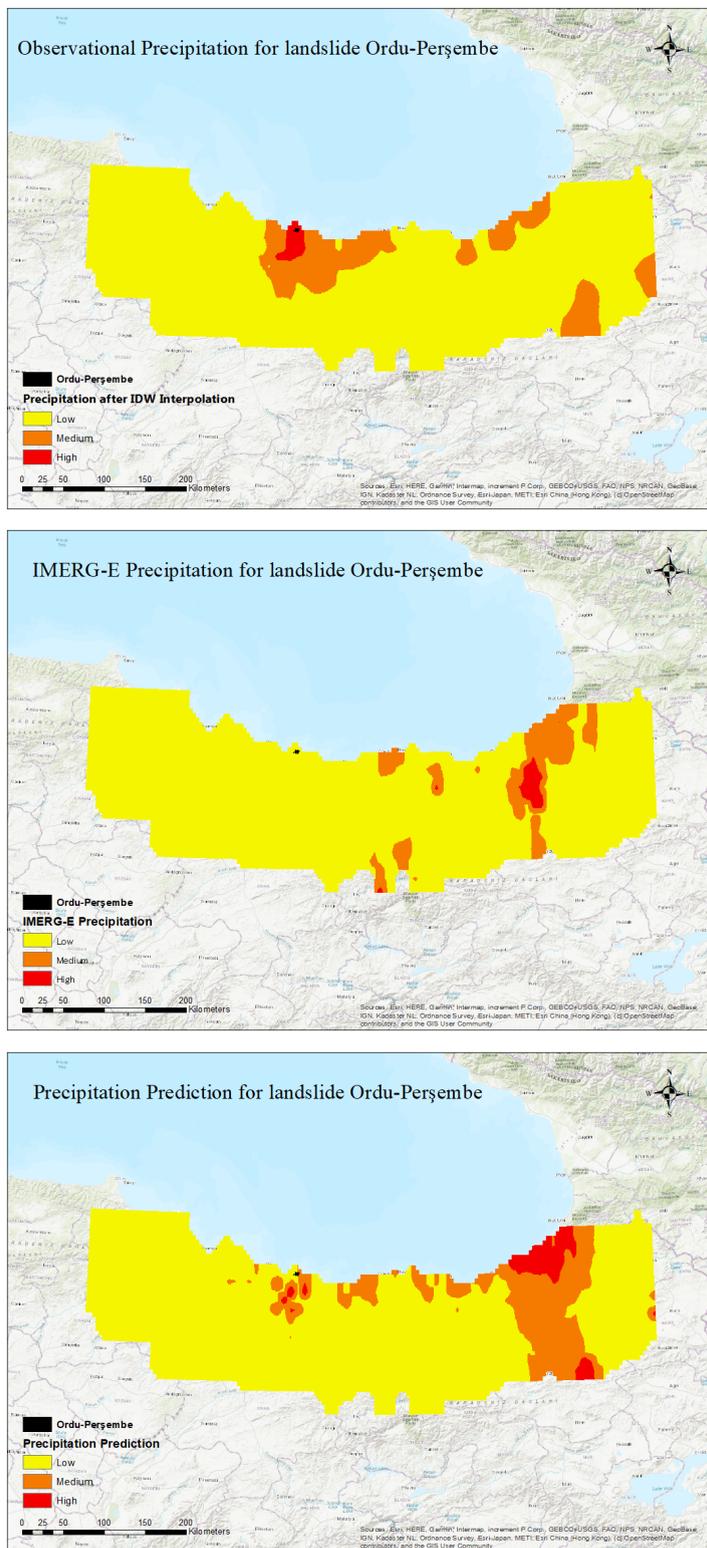


Figure 4.16: Prediction for precipitation/landslide: Ordu-Perşembe

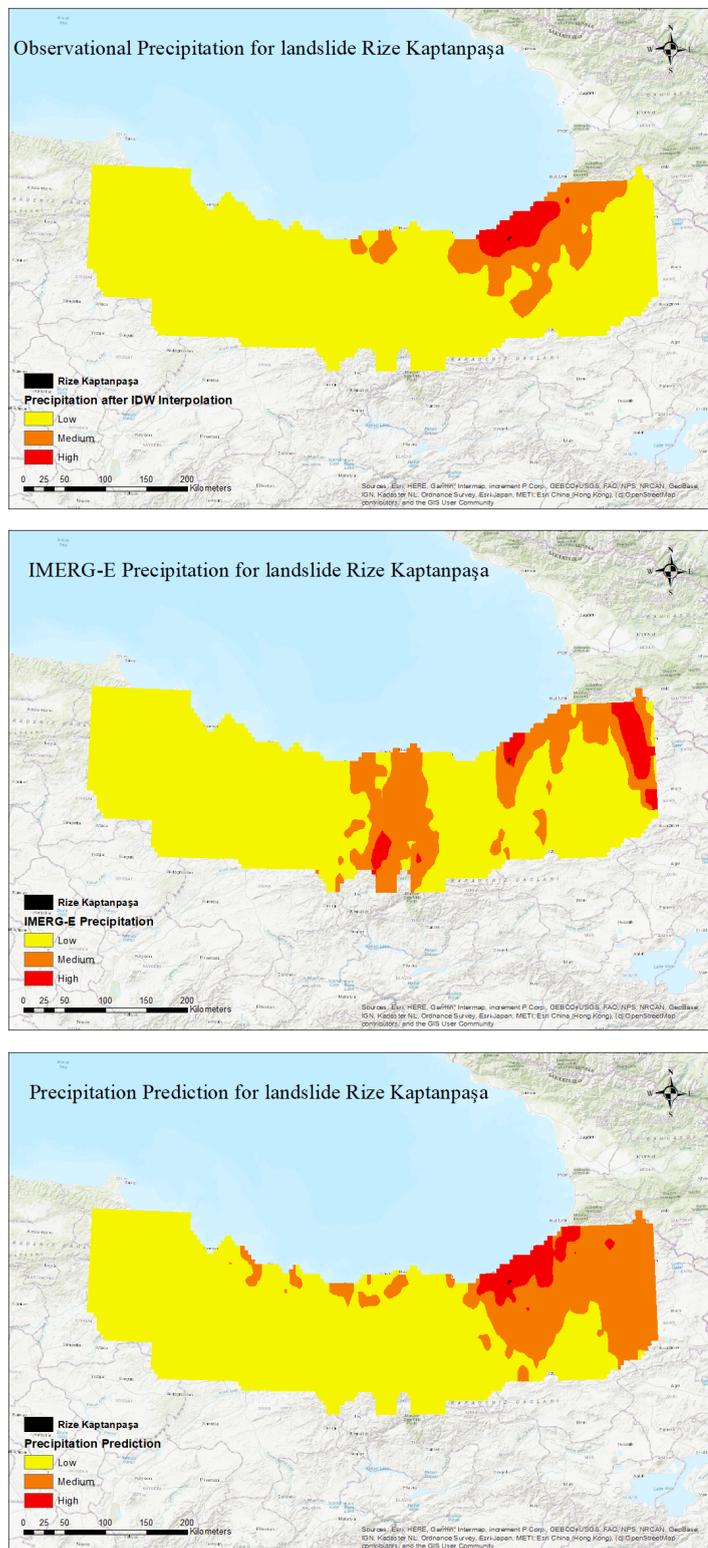


Figure 4.17: Prediction for precipitation/landslide: Rize Kaptanpaşa

Chapter 5

Discussion

Statistical postprocessing models have been proven to be powerful tools in the weather forecast field. However, not much has been done applying those methods on NRT satellite estimates and hardly no applications have been tested in the landslide context. In other words, it is important to stress that this is a field at its infancy stage.

Looking at results provided in the few articles published on this theme, the expectation was to largely reduce the difference between observed and satellite based rainfall measurements. For this reason, a full downscaling experiment was planned in the early stage of this research. However, the quality of the NRT IMERG product proved to be quite poor, forcing a different direction more towards bias reduction routines rather than spatio-temporal data resolution enhancement. Therefore, the sections below will expand the discussion for the bias correction aspects. Nevertheless, especially to help defining better landslide prediction strategies, rainfall downscaling in NRT should still be a field where scientific efforts should be invested.

5.1 OBSERVATIONAL PRECIPITATION INTERPOLATION

Even though OCK did not perform well in this study, other examples in the literature have pointed out at its good performance. Therefore, a potential explanation for such results can be potentially due to the local landscape. Even this finding matches with another study (Antal et al., 2021), OCK had also been proved outperformed than IDW in another study for mountainous regions in Turkey (Katipoğlu, 2022). The difference between this study and the above one is they were using monthly data from ground stations rather than daily data. The monthly precipitation data may have a more uniformly distribution, so the spatial structures are easily to built by using kriging.

Notably, the use of IDW was also the optimal case because similar landscape covariates have entered the modeling procedure for the bias removal. Therefore, using them as part of interpolating the rain gauge data and for the bias-removal would have inflated the performance artificially, thus being a grave mistake.

5.2 STATISTICAL POSTPROCESSING

Predictors' selection can also be improved in the future. The selection of predictors in this study was based on other's experience (Lombardo et al., 2018; Y. Zhang et al., 2022). However,

the process of deciding how many grids to choose within a neighborhood as predictors can for instance be improved by adding a sensitivity analysis. In such case, one could examine the inclusion of smaller or larger neighborhoods and improve the prediction as a consequence. In this research, the main trade off upon which the choice of neighborhood was justified was based on computational costs. In unreported results tests were made for smaller neighborhood and at each grid number increase, the model performance was recorded. However, at a neighborhood comprising 20 grid the procedure was interrupted because the computational time was increasing to the point of making the procedure unfeasible in the time frame allocated for this MSc research.

According to the table 4.5, precipitation from raw IMERG-E were not that important in the model so maybe more reference data should be considered rather than IMERG-E data. In the future, another NRT precipitation data or forecast data could be explored and bias-reduced using statistical postprocessing. The finding from this study also tells that IMERG-E may not be a good choice for modeling NRT RTL in eastern Black Sea Region in Turkey.

The difference between ZAGA and QRF is that ZAGA is a parametric method while QRF is non-parametric one. ZAGA can only select a certain number predictors for each parameter, and the distribution of the data also need to follow a specific shape. On the other hand, QRF is more flexible because it belongs to a class of machine-learning approaches. This makes QRF more suited to solve prediction tasks and it is most likely the reason behind the large performance gap with respect to the ZAGA approach. The observational data and satellite estimates in this study also had a relatively large gap. Traditional parametric methods may be difficult to built in this case. Currently, more and more deep learning based methods have been proposed in the field of statistical postprocessing (Harris et al., 2022; Schulz and Lerch, 2022), and those more flexible methods appear quite promising to boost the use of NRT satellite estimates in the future.

The BSS and reliability diagram in this research showed relatively poor performance compared to traditional statistical postprocessing studies on weather forecast (Whan and Schmeits, 2018). This can be explained by the big difference between observational data and satellite estimates (similar as forecasts). BSS and reliability diagram are more focused on the output values comparing to continuous metrics (RMSE, PCC and ME). Even though models can correct some errors in map, but the output values from the model are more difficult to postprocess. In fact, improving the accuracy of predict values can be a direction for future research. According to the figure 4.15-4.17, relative high risks area can be discovered by the model, but the output values from the model still can be improved.

It should be noted that the study site is relatively small compared to other areas where such postprocessing tools have been reported in the literature. For this reason, it is possible that the scale at which bias-corrections are applied can also play a large role. For instance, over regional, national or even continental extents, the match between observed and remotely-sensed rainfall can substantially improve whereas for very localized studies, satellite products may still be unsuited.

The predictor importance table (table 4.5) is also an important element for statistical postprocessing. Predictors were more important should be kept in the future analysis since they had played a positive role for the modeling. Moreover, according to the table, precipitation between July to October are more likely to be orographic precipitation while in March to June are more likely to be convective precipitation. Similar to the result from others (Y. Zhang et al., 2022), the precipitation data from the previous day is more important than the precipitation from the target day, and this is probably due to the infrared images are able to reflect signal earlier. Further im-

provements though can still be envisioned. For instance, temperature and wind data also largely control the spatio-temporal pattern of rainfall events. Therefore, further research efforts could be made to find the optimal predictor set to support any bias-reduction procedure in the context of weather science.

Chapter 6

Conclusion and recommendation

For predicting RTL in near-real-time, IMERG-E data involved large errors and it could not be used directly. This study has found that QRF is able to correct some major errors of IMERG NRT data in eastern Black Sea Region of Turkey. Several evaluation methods have also been carried out in this study for models comparison and model assessment. Two widely used statistical postprocessing methods for weather forecast have been tested in this study. The results found that nonparametric methods may perform better than parametric ones for NRT satellite estimates corrections. By comparing RMSE/PCC of observational data and IMERG-E to RMSE/PCC of observational data and predict values, the raw IMERG-E has been significantly improved by QRF, and the relationship between predict values and observational data is also stronger than the relationship between raw IMERG-E data and observational data. This finding is an important step for near-real-time forecasting rainfall forecast at an operational scale typical of other surface phenomena, to which class landslide belong too. Future work can be focused on finding better predictors and more advanced (better performance) techniques, but also to test whether other satellite products equipped with a forecast service can be used instead of IMERG.

The initial idea behind this work still stands and it should be mentioned here before closing. If and only if, rainfall forecasts would be proven to be correctable and downscalable, then a number of applications of fundamental importance could be enabled. The initial assumption for this thesis was to build a model capable of correcting differences between observational and remotely-sensed rainfall, and to downscale it in the end. Such a model could then be used to convert the rainfall pattern from forecasted estimated thus creating an opening towards landslide forecasting services. This may be of fundamental relevance for communities living in mountainous landscapes and could be the base for alternative early-warning systems, as most of the current ones are predominantly based on rain gauge and terrestrial radar data.

List of References

- Alexander, L. V. (2016). Global observed long-term changes in temperature and precipitation extremes: A review of progress and limitations in ipcc assessments and beyond. *Weather and Climate Extremes*, 11, 4–16.
- Allan, R. P., Liu, C., Zahn, M., Lavers, D. A., Koukouvagias, E., & Bodas-Salcedo, A. (2014). Physically consistent responses of the global atmospheric hydrological cycle in models and observations. *Surveys in Geophysics*, 35(3), 533–552.
- Antal, A., Guerreiro, P. M., & Cheval, S. (2021). Comparison of spatial interpolation methods for estimating the precipitation distribution in portugal. *Theoretical and Applied Climatology*, 145(3), 1193–1206.
- Aslami, F., Ghorbani, A., Sobhani, B., & Esmali, A. (2019). Comprehensive comparison of daily imerg and gsmap satellite precipitation products in ardabil province, iran. *International journal of remote sensing*, 40(8), 3139–3153.
- Baran, S., & Nemoda, D. (2016). Censored and shifted gamma distribution based emos model for probabilistic quantitative precipitation forecasting. *Environmetrics*, 27(5), 280–292.
- Bentzien, S., & Friederichs, P. (2012). Generating and calibrating probabilistic quantitative precipitation forecasts from the high-resolution nwp model cosmo-de. *Weather and Forecasting*, 27(4), 988–1002.
- Berg, P., Moseley, C., & Haerter, J. O. (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, 6(3), 181–185.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Caldwell, R., Gangopadhyay, S., Bountry, J., Lai, Y., & Elsner, M. (2013). Statistical modeling of daily and subdaily stream temperatures: Application to the methow river basin, washington. *Water Resources Research*, 49(7), 4346–4361.
- Carella, G., Vrac, M., Brogniez, H., Yiou, P., & Chepfer, H. (2020). Statistical downscaling of water vapour satellite measurements from profiles of tropical ice clouds. *Earth System Science Data*, 12(1), 1–20.
- Chaithong, T., & Komori, D. (2020). Application of satellite precipitation data to model the extreme rainfall-induced landslide event. *22nd Congress of the International Association for Hydro-Environment Engineering and Research-Asia Pacific Division: Creating Resilience to Water-Related Challenges, IAHR-APD 2020*.
- Chen, D., Ou, T., Gong, L., Xu, C.-Y., Li, W., Ho, C.-H., & Qian, W. (2010). Spatial interpolation of daily precipitation in china: 1951–2005. *Advances in Atmospheric Sciences*, 27(6), 1221–1232.
- Chen, H., Yong, B., Shen, Y., Liu, J., Hong, Y., & Zhang, J. (2020). Comparison analysis of six purely satellite-derived global precipitation estimates. *Journal of Hydrology*, 581, 124376.
- Dimitriadis, S. I., Liparas, D., Initiative, A. D. N., et al. (2018). How random is the random forest? random forest algorithm on the service of structural imaging biomarkers for alzheimer’s disease: From alzheimer’s disease neuroimaging initiative (adni) database. *Neural regeneration research*, 13(6), 962.

- Dimri, A., Palazzi, E., & Daloz, A. (2022). Elevation dependent precipitation and temperature changes over indian himalayan region. *Climate Dynamics*, 1–21.
- Dong, J., Crow, W. T., & Reichle, R. (2020). Improving rain/no-rain detection skill by merging precipitation estimates from different sources. *Journal of Hydrometeorology*, 21(10), 2419–2429.
- Fan, L., Lehmann, P., Zheng, C., & Or, D. (2020). Rainfall intensity temporal patterns affect shallow landslide triggering and hazard evolution. *Geophysical Research Letters*, 47(1), e2019GL085994.
- Fathi, M., Haghi Kashani, M., Jameii, S. M., & Mahdipour, E. (2021). Big data analytics in weather forecasting: A systematic review. *Archives of Computational Methods in Engineering*, 1–29.
- Fischer, E. M., & Knutti, R. (2016). Observed heavy precipitation increase confirms theory and early models. *Nature Climate Change*, 6(11), 986–991.
- Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181.
- Gallina, V., Torresan, S., Critto, A., Sperotto, A., Glade, T., & Marcomini, A. (2016). A review of multi-risk methodologies for natural hazards: Consequences and challenges for a climate change impact assessment. *Journal of environmental management*, 168, 123–132.
- Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252.
- Gentilucci, M., Bufalini, M., D’Aprile, F., Materazzi, M., & Pambianchi, G. (2021). Comparison of data from rain gauges and the imerg product to analyse precipitation in mountain areas of central italy. *ISPRS International Journal of Geo-Information*, 10(12), 795.
- Gilewski, P., & Nawalany, M. (2018). Inter-comparison of rain-gauge, radar, and satellite (imerg gpm) precipitation estimates performance for rainfall-runoff modeling in a mountainous catchment in poland. *Water*, 10(11), 1665.
- Gneiting, T., Raftery, A. E., Westveld III, A. H., & Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133(5), 1098–1118.
- Görüm, T., & Fidan, S. (2021). Spatiotemporal variations of fatal landslides in turkey. *Landslides*, 18(5), 1691–1705.
- Gumindoga, W., Rientjes, T. H., Haile, A. T., Makurira, H., & Reggiani, P. (2019). Performance of bias-correction schemes for cmorph rainfall estimates in the zambezi river basin. *Hydrology and earth system sciences*, 23(7), 2915–2938.
- Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., & Galli, M. (2006). Estimating the quality of landslide susceptibility models. *Geomorphology*, 81(1-2), 166–184.
- Haque, U., Blum, P., Da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., Malet, J.-P., Auflič, M. J., Andres, N., Poyiadji, E., et al. (2016). Fatal landslides in europe. *Landslides*, 13(6), 1545–1554.
- Harris, L., McRae, A. T., Chantry, M., Dueben, P. D., & Palmer, T. N. (2022). A generative deep learning approach to stochastic downscaling of precipitation forecasts. *arXiv preprint arXiv:2204.02028*.
- Hastings, D. A., Dunbar, P. K., Elphinstone, G. M., Bootz, M., Murakami, H., Maruyama, H., Masaharu, H., Holland, P., Payne, J., Bryant, N. A., et al. (1999). The global land one-kilometer base elevation (globe) digital elevation model, version 1.0. *National Oceanic and Atmospheric Administration, National Geophysical Data Center*, 325, 80305–3328.
- Hong, Y., Adler, R. F., Negri, A., & Huffman, G. J. (2007). Flood and landslide applications of near real-time satellite rainfall products. *Natural Hazards*, 43(2), 285–294.

- Hsu, J., Huang, W.-R., & Liu, P.-Y. (2021). Performance assessment of gpm-based near-real-time satellite products in depicting diurnal precipitation variation over taiwan. *Journal of Hydrology: Regional Studies*, 38, 100957.
- Huggett, R. J. (2016). *Fundamentals of geomorphology*. Routledge.
- Katipoğlu, O. M. (2022). Spatial analysis of seasonal precipitation using various interpolation methods in the euphrates basin, turkey. *Acta Geophysica*, 70(2), 859–878.
- Kawo, N. S., Hordofa, A. T., & Karuppannan, S. (2021). Performance evaluation of gpm-imerg early and late rainfall estimates over lake hawassa catchment, rift valley basin, ethiopia. *Arabian Journal of Geosciences*, 14(4), 1–14.
- Khan, S., Kirschbaum, D., & Stanley, T. (2021). Investigating the potential of a global precipitation forecast to inform landslide prediction. *Weather and Climate Extremes*, 33, 100364.
- Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the cmip5 ensemble. *Climatic change*, 119(2), 345–357.
- LaJoie, P., Cronin, E., Perrotti, J., Shives, E., & Webster, S. (2021). Cincinnati & covington urban development ii: Assessing flooding and landslide susceptibility along the ohio-kentucky border.
- Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., & Di, Z. (2017). A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdisciplinary Reviews: Water*, 4(6), e1246.
- Liu, L., Gao, C., Xuan, W., & Xu, Y.-P. (2017). Evaluation of medium-range ensemble flood forecasting based on calibration strategies and ensemble methods in lanjiang basin, southeast china. *Journal of Hydrology*, 554, 233–250.
- Lombardo, L., Opitz, T., & Huser, R. (2018). Point process-based modeling of multiple debris flow landslides using inla: An application to the 2009 messina disaster. *Stochastic environmental research and risk assessment*, 32(7), 2179–2198.
- Marin, R. J., & Mattos, Á. J. (2020). Physically-based landslide susceptibility analysis using monte carlo simulation in a tropical mountain basin. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 14(3), 192–205.
- Medina, V., Hürlimann, M., Guo, Z., Lloret, A., & Vaunat, J. (2021). Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale. *Catena*, 201, 105213.
- Mendoza, P. A., Rajagopalan, B., Clark, M. P., Ikeda, K., & Rasmussen, R. M. (2015). Statistical postprocessing of high-resolution regional climate model output. *Monthly Weather Review*, 143(5), 1533–1553.
- Michaud, C., Malet, J.-P., Oppikofer, T., Emberson, R., Kirschbaum, D., Pacini, F., Horton, P., Puissant, A., Mazzanti, P., Poteau, M., et al. (2021). Landslide information system for disaster risk financing: Earth observation and modelling products for near-real-time assessment. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 1812–1814.
- Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M., et al. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. *Scientific reports*, 9(1), 1–10.
- Nepal, B., Shrestha, D., Sharma, S., Shrestha, M. S., Aryal, D., & Shrestha, N. (2021). Assessment of gpm-era satellite products’(imerg and gsmap) ability to detect precipitation extremes over mountainous country nepal. *Atmosphere*, 12(2), 254.
- Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., & Thorstensen, A. R. (2018). The persiann family of global satellite precipitation data:

- A review and evaluation of products. *Hydrology and Earth System Sciences*, 22(11), 5801–5816.
- O’Gorman, P. A. (2015). Precipitation extremes under climate change. *Current climate change reports*, 1(2), 49–59.
- Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, 40(10), 927–930.
- Qi, W., Yong, B., & Gourley, J. J. (2021). Monitoring the super typhoon lekima by gpm-based near-real-time satellite precipitation estimates. *Journal of Hydrology*, 603, 126968.
- Ranjbar, A., Vali, A., Mokarram, M., & Tarippanah, F. (2020). Investigating variations of vegetation: Climatic, geological substrate, and topographic factors—a case study of kharestan area, fars province, iran. *Arabian Journal of Geosciences*, 13(14), 1–18.
- Ratner, B. (2009). The correlation coefficient: Its values range between + 1/- 1, or do they? *Journal of targeting, measurement and analysis for marketing*, 17(2), 139–142.
- Ratri, D. N., Whan, K., & Schmeits, M. (2021). Calibration of ecmwf seasonal ensemble precipitation reforecasts in java (indonesia) using bias-corrected precipitation and climate indices. *Weather and Forecasting*, 36(4), 1375–1386.
- Ray, R. L., Jacobs, J. M., & Douglas, E. M. (2018). Modeling regional landslide susceptibility using dynamic soil moisture profiles. *Journal of Mountain Science*, 15(8), 1807–1824.
- Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-science reviews*, 180, 60–91.
- Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., & De Bastiani, F. (2019). *Distributions for modeling location, scale, and shape: Using gamlss in r*. CRC press.
- Sakib, S., Ghebreyesus, D., & Sharif, H. O. (2021). Performance evaluation of imerg gpm products during tropical storm imelda. *Atmosphere*, 12(6), 687.
- Scheuerer, M. (2014). Probabilistic quantitative precipitation forecasting using ensemble model output statistics. *Quarterly Journal of the Royal Meteorological Society*, 140(680), 1086–1096.
- Scheuerer, M., & Hamill, T. M. (2015). Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. *Monthly Weather Review*, 143(11), 4578–4596.
- Schulz, B., & Lerch, S. (2022). Machine learning methods for postprocessing ensemble forecasts of wind gusts: A systematic comparison. *Monthly Weather Review*, 150(1), 235–257.
- Sillmann, J., Kharin, V. V., Zwiers, F., Zhang, X., & Bronaugh, D. (2013). Climate extremes indices in the cmip5 multimodel ensemble: Part 2. future climate projections. *Journal of geophysical research: atmospheres*, 118(6), 2473–2493.
- Soo, E. Z. X., Wan Jaafar, W. Z., Lai, S. H., Othman, F., & Elshafie, A. (2022). Enhancement of satellite precipitation estimations with bias correction and data-merging schemes for flood forecasting. *Journal of Hydrologic Engineering*, 27(9), 05022009.
- Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., & De Bastiani, F. (2017). *Flexible regression and smoothing: Using gamlss in r*. CRC Press.
- Stoffel, M., Tiranti, D., & Huggel, C. (2014). Climate change impacts on mass movements—case studies from the european alps. *Science of the Total Environment*, 493, 1255–1266.
- Taillardat, M., Mestre, O., Zamo, M., & Naveau, P. (2016). Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Monthly Weather Review*, 144(6), 2375–2393.
- Tang, G., Behrangi, A., Long, D., Li, C., & Hong, Y. (2018). Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *Journal of hydrology*, 559, 294–306.
- Tennant, C. J., Harpold, A. A., Lohse, K. A., Godsey, S. E., Crosby, B. T., Larsen, L. G., Brooks, P. D., Van Kirk, R. W., & Glenn, N. F. (2017). Regional sensitivities of seasonal snow-

- pack to elevation, aspect, and vegetation cover in western north america. *Water Resources Research*, 53(8), 6908–6926.
- Titti, G., Borgatti, L., Zou, Q., Cui, P., & Pasuto, A. (2021). Landslide susceptibility in the belt and road countries: Continental step of a multi-scale approach. *Environmental Earth Sciences*, 80(18), 1–18.
- Türkoğlu, N., Cicek, I., Gurgun, G., & Gürer, Ý. (2003). Recent flood disasters at eastern black sea region of turkey. *EGS-AGU-EUG Joint Assembly*, 14568.
- Usovich, B., Lipiec, J., Łukowski, M., & Słomiński, J. (2021). Improvement of spatial interpolation of precipitation distribution using cokriging incorporating rain-gauge and satellite (smos) soil moisture data. *Remote Sensing*, 13(5), 1039.
- Vannitsem, S., Bremnes, J. B., Demaeyer, J., Evans, G. R., Flowerdew, J., Hemri, S., Lerch, S., Roberts, N., Theis, S., Atencia, A., et al. (2021). Statistical postprocessing for weather forecasts: Review, challenges, and avenues in a big data world. *Bulletin of the American Meteorological Society*, 102(3), E681–E699.
- van Straaten, C., Whan, K., & Schmeits, M. (2018). Statistical postprocessing and multivariate structuring of high-resolution ensemble precipitation forecasts. *Journal of Hydrometeorology*, 19(11), 1815–1833.
- van Westen, C., van Asch, T. W., & Soeters, R. (2006). Landslide hazard and risk zonation—why is it still so difficult? *Bulletin of Engineering geology and the Environment*, 65(2), 167–184.
- Vermote, E., Justice, C., Csiszar, I., Eidenshink, J., Myneni, R., Baret, F., Masuoka, E., Wolfe, R., & Claverie, M. (2014). Noaa climate data record (cdr) of normalized difference vegetation index (ndvi), version 4. *NOAA Natl. Clim. Data Cent.*
- Wagenmakers, E.-J., & Farrell, S. (2004). Aic model selection using akaike weights. *Psychonomic bulletin & review*, 11(1), 192–196.
- Wang, H., & Yong, B. (2020). Quasi-global evaluation of imerg and gsmap precipitation products over land using gauge observations. *Water*, 12(1), 243.
- Wang, N., Cheng, W., Lombardo, L., Xiong, J., & Guo, L. (2021). Statistical spatiotemporal analysis of hydro-morphological processes in china during 1950–2015. *Stochastic environmental research and risk assessment*, 1–21.
- Whan, K., & Schmeits, M. (2018). Comparing area probability forecasts of (extreme) local precipitation using parametric and machine learning statistical postprocessing methods. *Monthly Weather Review*, 146(11), 3651–3673.
- Wilks, D. S. (2011). Forecast verification. *International geophysics* (pp. 301–394). Elsevier.
- Yang, D. (2020). Ensemble model output statistics as a probabilistic site-adaptation tool for satellite-derived and reanalysis solar irradiance. *Journal of Renewable and Sustainable Energy*, 12(1), 016102.
- Yu, C., Zheng, J., Hu, D., Di, Y., Zhang, X., & Liu, M. (2021). Evaluation and correction of imerg late run precipitation product in rainstorm over the southern basin of china. *Water*, 13(2), 231.
- Zhang, T., Liang, Z., Wang, H., Wang, J., Hu, Y., & Li, B. (2022). Merging multisatellite precipitation products using stacking and emos-csg methods.
- Zhang, Y., Ye, A., Nguyen, P., Analui, B., Sorooshian, S., & Hsu, K. (2022). Qrf4p-nrt probabilistic post-processing of near-real-time satellite precipitation estimates using quantile regression forests. *Water Resources Research*, e2022WR032117.
- Zhou, C., Gao, W., Hu, J., Du, L., & Du, L. (2021). Capability of imerg v6 early, late, and final precipitation products for monitoring extreme precipitation events. *Remote Sensing*, 13(4), 689.