



TECHNICAL UNIVERSITY OF CRETE

TITLE

Evaluation of gridded precipitation datasets over Greece

BY

Kalliopi – Mikaela Papa

A thesis submitted in fulfilment of the partial requirements for the degree of

BSc Environmental Engineering
School of Chemical and Environmental Engineering
Technical University of Crete

Under the supervision of
Associate Professor Aristeidis Koutroulis

COMMITTEE

Aristeidis Koutroulis, Associate Professor, TUC
Apostolos Voulgarakis, Professor, TUC
Emmanouil Grillakis, Laboratory Teaching Staff, TUC

Chania, Greece

April 2024

Copyright ©TUC, 2024

Με επιφύλαξη παντός δικαιώματος

Απαγορεύεται η αντιγραφή, αποθήκευση σε αρχείο πληροφοριών, διανομή, αναπαραγωγή, μετάφραση ή μετάδοση της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό, υπό οποιαδήποτε μορφή και με οποιοδήποτε μέσο επικοινωνίας, ηλεκτρονικό ή μηχανικό, χωρίς την προηγούμενη έγγραφη άδεια του συγγραφέα και των επιβλεπόντων καθηγητών. Επιτρέπεται η αναπαραγωγή, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν στη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Η έγκριση της διπλωματικής εργασίας από τη Σχολή Χημικών Μηχανικών και Μηχανικών Περιβάλλοντος του Πολυτεχνείου Κρήτης δεν υποδηλώνει αποδοχή των απόψεών του (Ν. 5343/1932, Άρθρο 202).

Copyright ©TUC, 2024

All Rights Reserved

Neither the whole nor any part of this diploma thesis may be copied, stored in a retrieval system, distributed, reproduced, translated, or transmitted for commercial purposes, in any form or by any means now or hereafter known, electronic or mechanical, without the written permission either from the author or the supervisor(s). Reproducing, storing, and distributing this thesis for non-profitable, educational, or research purposes is allowed, without prejudice to reference to its source and inclusion of the present text. Any queries concerning the use of the present thesis for commercial purposes must be addressed to its author.

Approval of this diploma thesis by the School of Chemical and Environmental Engineering of the Technical University of Crete (TUC) does not constitute in any way an acceptance of the views of the author contained herein by the said academic organization (L. 5343/1932, art. 202).

Περίληψη

Τα χωροχρονικά μοτίβα βροχόπτωσης στην Ελλάδα επηρεάζονται από διάφορους παράγοντες, συμπεριλαμβανομένης της σύνθετης τοπογραφίας και των πολυποίκιλων κλιματικών συστημάτων. Τα βροχόμετρα, αν και αποτελούν τη βάση για την αξιόπιστη ποσοτικοποίηση της βροχόπτωσης, παρουσιάζουν ελλείψεις καταγραφές, είναι σποραδικά και συχνά συντηρούνται πλημμελώς. Σε αυτές τις περιπτώσεις, τα πλεγματικά δεδομένα μπορούν να αποτελέσουν μια λύση, παρέχοντας χωρικά και χρονικά συνεχή δεδομένα βροχόπτωσης. Τα δεδομένα αυτά, ωστόσο, παρουσιάζουν περιορισμούς όσον αφορά τη ρεαλιστική αποτύπωση της βροχόπτωσης, οι οποίοι οφείλονται κυρίως στις εγγενείς ατέλειες των υποκείμενων μεθόδων που χρησιμοποιούν. Στην παρούσα εργασία γίνεται αξιολόγηση για εννέα από τα πιο χωροχρονικά λεπτομερή μοντέλα, και συγκεκριμένα των ERA5-Land (ERA5L), AgERA5, CHELSA-W5E5 v1.1 (CHELSA), MSWEP V2.8, CHIRPS05, IMERG V06, PERSIANN-CCS-CDR (PCCSCDR) και E-OBS, σε αντιπαραβολή με παρατηρήσεις πεδίου που λήφθηκαν από 304 βροχομετρικούς σταθμούς σε όλη την Ελλάδα. Η αξιολόγηση πραγματοποιείται σε ημερήσια και μηνιαία κλίμακα, για μια περίοδο 32 ετών (1984 - 2016), αξιολογώντας την απόδοση των πλεγματικών δεδομένων, εξετάζοντας τη συνέπεια τους τόσο σε επίπεδο χώρας όσο και κατά περιοχικές συστάδες παρόμοιων χαρακτηριστικών. Η ικανότητα των πλεγματικών δεδομένων να αποτυπώνουν σωστά τις ακραίες τιμές και τα μοτίβα βροχόπτωσης εξετάζεται με στατιστικούς δείκτες, ενώ περαιτέρω συμπεράσματα εξάγονται με στατιστική ανάλυση τυποποιημένων κλιματικών δεικτών και διαχρονικών τάσεων. Τα μοντέλα CHELSA, CERRAL και AgERA5 παράγουν συστηματικά συνεπέστερα αποτελέσματα για όλους τους δείκτες έναντι άλλων, όπως για παράδειγμα το PCCSCDR, το οποίο παρουσιάζει υποδεέστερες επιδόσεις και στις δύο χρονικές κλίμακες. Η στατιστική ανάλυση αποκαλύπτει διακριτά μοτίβα εντονότερων βροχοπτώσεων στις βόρειες και δυτικές περιοχές της χώρας, με έντονη εποχιακή διακύμανση στα δυτικά και νότια. Προσδιορίζεται επίσης μια πιθανή μέση αύξηση άνω των 110 mm στη μέση ετήσια βροχόπτωση και 30 mm στην ακραία βροχόπτωση, ανά δεκαετία, κατά την περίοδο αξιολόγησης. Συνολικά, τα πλεγματικά δεδομένα αδυνατούν να αποδώσουν με ακρίβεια τις ακραίες βροχοπτώσεις, αλλά το CHELSA και το CERRAL ξεχωρίζουν ως πιο αξιόπιστες επιλογές για την περιγραφή της δυναμικής των βροχοπτώσεων στην Ελλάδα.

Abstract

The spatiotemporal precipitation patterns in Greece are accentuated by several factors, including the complex topography and the multifaceted climatic regimes of the country. Rain gauges, albeit a reliable tool for the accurate quantification of rainfall, are scarce, sporadic, and not properly maintained. In these instances, gridded datasets may provide a solution by administering spatially and temporally continuous precipitation data. The products, however, reveal limitations in the realistic simulation of precipitation, primarily caused by the intrinsic flaws of the underlying methods used. The assessment of nine of the most spatially and temporally detailed precipitation datasets, namely ERA5-Land (ERA5L), AgERA5, CHELSA-W5E5 v1.1 (CHELSA), MSWEP V2.8, CHIRPS05, IMERG V06, PERSIANN-CCS-CDR (PCCSCDR), and E-OBS, compared against field observations acquired from 304 gauging stations across Greece has not been previously attempted. The evaluation is conducted on a daily and a monthly timescale, over a 32-year period (1984 – 2016), assessing the performance of the gridded products by considering both the country as a whole and its individual regions. The ability of the datasets to correctly portray the occurrence of extreme events and precipitation patterns is examined by statistical metrics and further insights are provided by the application and statistical analysis of climate indices on ground observations. CHELSA, CERRAL and AgERA5 consistently yield acceptable results across statistical metrics, outperforming others like PCCSCDR, which exhibits inferior performance in both temporal scales. The statistical analysis reveals distinct patterns of heavier precipitation in northern and western regions, with strong seasonal variability in the West and South and a possible average decennial increase of over 110 mm in mean annual rainfall and 30 mm in extreme rainfall, over the assessment period. Overall, the datasets fail to accurately depict extreme precipitation, but CHELSA and CERRAL stand out as more reliable options for describing the precipitation dynamics in Greece.

Acknowledgments

Upon finalising the present thesis, I would like to express my sincerest gratitude towards my supervisor, Associate Professor Aristeidis Koutroulis, whose input and guidance were invaluable throughout the whole process. I want to thank him for assigning me the topic of the study and introducing me to new concepts of academic research. His devotion to his work and willingness to help and provide answers to the questions posed on him, is deeply admirable and inspiring.

I would like to praise Dr Emmanouil Grillakis for his extensive body of knowledge and constructive criticism, which were imperative for the realisation of the study. The conversations we engaged in were intriguing and thought-provoking.

I also want to thank Professor Apostolos Voulgarakis for agreeing to act as a member of the thesis committee, devoting time to evaluate the thesis, and for welcoming me to participate many times in his weekly lab meetings.

Highly important is the contribution of Konstantinos Seiradakis and Athanasios Tsilimigkras. Their readiness to advise and help me, even in the most mundane of tasks, was extremely beneficial.

Finally, I want to commend with great respect, my family, and friends on the patience they have demonstrated, and the much-needed support and encouragement during this time.

Table of Contents

1. Introduction.....	9
1.1 Field of Research.....	9
1.2 Thesis Structure	10
2. Background.....	11
2.1 Study Area.....	11
2.2 Precipitation Measurement	11
2.2.1 Gauge-based observations	11
2.2.2 Satellite-based observations.....	12
2.2.3 Reanalysis datasets.....	13
2.3 Relevant research, gaps, and study aim	14
3. Data and Methods.....	16
3.1 In Situ Observations.....	16
3.1.1 Data selection	16
3.1.2 Data processing.....	17
3.2 Gridded Precipitation Datasets.....	18
3.2.1 Data selection	18
3.2.2 Data processing.....	21
3.2.3 Data pre-processing.....	22
3.3 Evaluation	22
3.3.1 Performance metrics	22
3.3.2 Precipitation Indices.....	24
4. Results	25
4.1 Precipitation patterns based on ground observations	25
4.2 Precipitation patterns in gridded datasets	33
5. Conclusions	51
6. Literature	53
7. Appendix	60

List of Figures

Figure 3.1. Location and elevation of rainfall stations and fourteen water districts of Greece.	16
Figure 3.2. Number of stations with daily rainfall data over time for the full dataset (blue) and for the two different data providers, YPEN (green) and HNMS (yellow). The flat grey line corresponds to the monthly filled and homogenised dataset (251 stations). The shaded period (1984 - 2016) highlights the period of assessment.	18
Figure 4.1. Mean annual precipitation of 251 homogenised timeseries for 1962-2020.....	26
Figure 4.2. a) Mean annual precipitation at the level of the fourteen water districts is illustrated in the central map and their corresponding seasonal distribution in the surrounding bar-graphs. (b) illustrates the mean seasonal rainfall of Greece and (c) the variation of mean annual precipitation with altitude.....	28
Figure 4.3. Spatial distribution of seasonal variability of mean accumulated precipitation across stations.....	29
Figure 4.4. a.) Visualisation of eight station clusters and their location. b) Analysis fields (parameters) and c) number of stations in each cluster.....	31
Figure 4.5. Spatial distribution and median values of ETCCDI climate indices for R95p, rx1, rx5, SDII, R10mm and R20mm.....	32
Figure 4.6. Trends of PRCPTOT and R95p indices for 1985 – 2016.....	33
Figure 4.7. Mean annual rainfall distribution resulting from the nine gridded datasets and point observations of the 1985 – 2016 period. *For IMERG: mean annual rainfall corresponds to the 2000 – 2021 period.	34
Figure 4.8. NSE, KGE, R score values for 1985 – 2016, derived from 304 daily time series, and the median value of the nine gridded precipitation datasets.	35
Figure 4.9. NSE values of 304 daily time series and median values of the nine gridded datasets.	36
Figure 4.10. Median values of KGE for the nine datasets.....	37
Figure 4.11. R values of 304 daily time series and median values of the nine gridded datasets.	38
Figure 4.12. MAE, RMSE, PBIAS score values for 1985 – 2016, derived from 304 daily time series, and the median value of the nine gridded precipitation datasets.....	39

Figure 4.13. MAE values of 304 daily time series and median values of the nine gridded datasets.	40
Figure 4.14. RMSE values of 304 daily time series and median values of the nine gridded datasets.	41
Figure 4.15. PBIAS values of 304 daily time series and median values of the nine gridded datasets.	42
Figure 4.16. NSE, KGE, R score values for 1985 – 2016, derived from 251 monthly time series, and the median value of the nine gridded precipitation datasets.....	45
Figure 4.17. MAE, RMSE, PBIAS score values for 1985 – 2016, derived from 251 monthly time series, and the median value of the nine gridded precipitation datasets. 45	
Figure 4.18. Scatter plots of the jointly statistically significant slope values (red points) and the remaining slope values (black points), regardless of significance, based on the statistical analysis of the PRCPTOT index. The vertical axis corresponds to the observational values, whereas the horizontal axis to the values of the datasets. Note that the range and size of the vertical axis is not identical among all datasets.	48
Figure 4.19. Scatter plots of the jointly statistically significant slope values (red points) and the remaining slope values (black points), regardless of significance, based on the statistical analysis of the R95p extreme index. Note that the range and size of the vertical axis is not identical among all datasets.	50

List of Tables

Table 3.1: Summary information of fill rate, operation period, and available data.....	17
Table 3.2. Summary information of gridded precipitation datasets.....	21
Table 4.1. Number of stations and station coverage in the fourteen water districts.....	27
Table 4.2. Number of stations and median values of features in the clusters.	31
Table 4.3. Median values of the statistical metrics of the nine datasets and number of stations, at three different altitude ranges.....	43
Table 4.4. Median correlation coefficient (R) and percent bias of extreme indices for the nine datasets, between 1985 – 2016. Disclosed in brackets are the corresponding values of each index, as derived from the observations. *The extreme indices of IMERG were based on the period between 2000 – 2020.....	46
Table 4.5. Values of mean slope, and its coefficient of variation (CV) and R metric, fraction of joint significance, as well as CV, R, and mean values of jointly significant stations, based on the statistical analysis of the PRCPTOT index. Joint significance fraction (JSF) refers to the fraction of statistically significant stations of the datasets, where both the gridded datasets and the station observations (OBS) present statistical significance. CV is an indicator of the average deviation of the values from the mean value, relative to the mean of the specific series. The data provided by IMERG were proven statistically insignificant and the corresponding estimations were not calculated. **Single value.....	47
Table 4.6. Values of mean slope, and its coefficient of variation (CV) and R metric, fraction of joint significance, as well as CV, R, and mean values of jointly significant stations, based on the statistical analysis of the R95p extreme index. Joint significance fraction (JSF) refers to the fraction of statistically significant stations of the datasets, where both the datasets and the observations (OBS) present statistical significance. CV is an indicator of the average deviation of the values from the mean value, relative to the mean of the specific series. The estimations marked in bold italic were derived from two values.	49

1. Introduction

1.1 Field of Research

Precipitation plays a pivotal role in the hydrological cycle and atmospheric regulation, so accurate measurement is significant for a multitude of applications, such as water resources management, risk assessment and mitigation in extreme weather conditions (i.e. floods, droughts), and climate research. Achieving a precise description of rainfall can be challenging across different time scales, since it is a highly unpredictable physical phenomenon that lacks consistent spatial distribution and as the time intervals of measurement or analysis decrease, the complexity of describing it becomes more pronounced.

In hydrological modelling, the accurate measurement and understanding of precipitation patterns is essential in simulating the movement and availability of water resources within a region. The simulated behaviour of rivers, lakes, and groundwater systems improve upon decisions related to water management, such as reservoir operations and drought mitigation strategies. Agriculture also relies heavily on precipitation data for crop planning and irrigation regimes, as it directly influences soil moisture levels and plant growth. Timely and accurate precipitation data help farmers adapt to weather variations and make informed decisions that contribute to food security. In flood prediction, a comprehensive understanding of precipitation patterns is fundamental for anticipating and responding to potential flood events, allowing for early warnings and effective disaster planning.

In the last few decades, there has been an upsurge in the development and use of gridded datasets, bringing new prospects in precipitation monitoring. The datasets are products of reanalysis, satellite observations, spatial interpolation of point-scale observations, data from ground-based meteorological radars, or a combination of them. Information derived from such products is extremely valuable, considering it benefits the study of the spatio-temporal continuity of the rainfall distribution and supports hydrological applications, especially where gauge observations are scarce. Nevertheless, there are limitations in the realistic simulation of precipitation, primarily caused by the intrinsic flaws of the underlying method of the dataset. Therefore, validating the consistency of the datasets using rain-gauge observations is essential.

The assessment of nine of the most spatially and temporally detailed precipitation datasets, compared against field observations deriving from approximately 300 gauging stations across Greece, has not been previously attempted. The main objective of this study is to identify the most reliable product for Greece, considering both the country as a whole and its individual regions, which would prove invaluable for hydrological studies of local water catchments and limited drainage basins. The assessment is carried out on a daily and a monthly timescale, covering a wide range of hydrological processes, like the occurrence of extreme events and precipitation pattern analysis.

1.2 Thesis Structure

The present thesis comprises the following chapters,

- **Chapter 2:**
Introduction to the study area and a basic theoretical background regarding the means of acquiring precipitation measurements and the limitations that arise with each method. This chapter also offers an overview of relevant studies conducted on Greece and the shortcomings in research the present study aims to abridge.
- **Chapter 3:**
Presentation of the chosen gridded precipitation datasets and ground observations and the processes followed to ensure the quality of the data for the assessment. We describe the range and scope of the analysis, in addition to the statistical indicators and indices used for the evaluation.
- **Chapter 4:**
This chapter contains the results of the analysis of the datasets, accompanied by corresponding figures and tables, which further describe and highlight the findings.
- **Chapter 5:**
A summary of the results and outcomes of the study, and recommendations for future research.

2. Background

2.1 Study Area

Greece is located in the south-eastern part of the European Continent, between the latitudes 35°N – 42°N and longitudes 19°E – 28°E, occupying a total area of 131,957 km², 80% of which is considered mountainous. It has an intricate coastline of 13,676 km and two major mountain ranges, the Pindus and Rhodope Mountain Ranges, that extend across the largest portion of northern and western continental Greece. Therefore, a diverse climate can be observed with varying patterns of rainfall distribution (**Figure 4.1**), both temporally and spatially, as is identified in mountainous and coastal regions (Ballantyne, 1983; Barros & Lettenmaier, 1994; Basist et al, 1994; Marra et al., 2021). The climate of the country is characterised by mild and wet winters over the central and southern part, as well as cold winters accompanied by heavy snowfall mainly in the mountainous areas of northern and central Greece. During summer, precipitation is sparse across the entire country. According to the updated Köppen-Geiger climate classification system, Greece falls primarily under the Mediterranean climate type, but other types can be observed to a minimal extent (Beck et al., 2020). Greece is divided into fourteen water districts (**Figure 4.2**), of areas that share similar morphological characteristics, available water resources, and water demands (Baltas, 2008).

2.2 Precipitation Measurement

2.2.1 Gauge-based observations

Ground-based precipitation measurements are collected using a variety of instruments, including rain gauges, snow gauges and disdrometers. Rain gauges record point-specific rainfall accumulation and are divided into non-automatic and automatic. Non-automatic types, like the Helmann rain gauge, rely on human observation by collecting the rainwater in a funnel. The funnel then, directs the accumulated rain in a measuring container with calibrated markings (in inches or millimetres), which the observer must physically read and record at regular time intervals (hourly or daily). Automatic rain gauges (e.g. the Tipping Bucket Rain Gauge, Weighing Precipitation Gauge, Doppler Radar) are devices equipped with sensors and data recording systems designed to automatically and consistently measure and document rainfall. These devices offer precise detailed information on the timing and intensity of precipitation events, as well as remote sensing of inaccessible locations. Data from automatic rain gauges are digitally recorded, eliminating the need for manual measurements, therefore reducing the risk of human errors associated with manual data collection. Non-automatic rain gauges are less suited for real-time or continuous monitoring, compared to automatic rain gauges which can provide more frequent and immediate information.

Consequently, there are certain disadvantages that stem solely from the placement of rain gauges, related to location and altitude. Since precipitation magnitude is closely linked to elevation, the orographic characteristics of a given region can produce measured precipitation that varies eminently between two points of different heights (Briggs & Cogley, 1996). This is considered an effect of the orographic lift, which occurs when elevated terrain

causes humid air masses to ascend and quickly become colder. The ascending air eventually reaches its dew point temperature, becoming more saturated and the water vapor within it more condensed forming water droplets. Influenced by factors such as wind, humidity, and temperature, the water droplets can coalesce and produce various types of clouds. A process which often results in precipitation, as the condensed moisture falls to the ground. The loss of a significant amount of moisture leads to less cloud formations and reduced precipitation on the leeward side, thus creating a rain shadow or a drier area.

Dispersed gauge networks do not incorporate enough stations, especially at high altitudes, to correctly describe the variations in precipitation over regions with complex topography. The interpolation of data for unknown points is necessary in these cases but can lead to a misrepresentation of the actual occurring rainfall amount (Tapiador et al., 2012).

2.2.2 Satellite-based observations

Satellite-based observations are a critical component of meteorological and hydrological monitoring systems. These observations involve the use of low-orbiting satellites, equipped with advanced sensors to remotely detect, and quantify precipitation patterns over vast geographic areas (Levizzani et al., 2007). By continuously scanning the surface of the Earth, the satellites produce homogeneous information about the distribution, intensity, and movement of precipitation.

One of the main advantages of satellite-based precipitation observations is the global coverage, which allows for the monitoring of rainfall in remote and inaccessible regions where ground-based data may be scarce or non-existent. Satellite sensors use a variety of techniques, including passive microwave sensors, which detect microwave radiation emitted or scattered by precipitation particles, and radar sensors, which use active beams of electromagnetic waves to measure the intensity of precipitation (Levizzani et al., 2020). These sensors provide data at different spatial and temporal resolutions, making the analysis of precipitation events possible on a variety of scales, from large weather systems to local storms.

The first satellite, created exclusively for the research of precipitation over the tropics, was the Tropical Rainfall Measuring Mission (TRMM) by the National Aeronautics and Space Administration (NASA), launched in 1997. The TRMM used multiple satellite and radar sources as input to measure the microwave energy emitted by the Earth and assess the presence of water and the intensity of rainfall in the atmosphere (Huffman et al., 2007). Several datasets were later introduced that utilise satellite imaging to effectively estimate precipitation rates and adequately capture the spatial and temporal fluctuations in precipitation in great detail, across most regions of the world (Mahmoud et al., 2018).

Present satellite sensors use integrated methods involving microwave and infrared radiation to detect variations in the temperature of the cloud tops, based on the understanding that colder temperatures are indicative of larger vertical cloud expansion, thereby signifying a higher probability of precipitation (Levizzani et al., 2007). But considering the intricate layers of cloud formations, this method might not identify the precipitating cloud or correctly calculate the amount of rainfall, as the temperature of the cloud tops and precipitation intensity are not clearly correlated. To address this, datasets introduce cloud classification

systems which categorise clouds by type and distinguish their unique features determined by their infrared brightness temperature (So & Shin, 2018). An example being the Precipitation Estimation from Remote Sensed Information using Artificial Neural Networks (PERSIANN) Cloud Classification System (CCS) (Hong et al., 2004), which uses infrared sensors to detect rain occurrences. Nevertheless, infrared satellites can mistake high temperature clouds as non-precipitating instances and disregard rain above a certain temperature threshold. Additionally, the cloud classification system incorporated in this product, requires manual characterisation of cloud types, which reduces the variations of precipitating clouds to their most evident features (Sadeghi et al., 2021). Cases such as these, highlight the inherent challenges and limitations of rainfall estimation across varying elevations and point out the need for improved methodologies that focus on data integration to enhance accuracy in complex terrains.

Satellite products evaluated in this study include the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System-Climate Data Record (PERSIANN-CCS-CDR) and the Integrated Multi-satellitE Retrievals for GPM (IMERG) datasets. The products provide 3-hourly and half-hourly precipitation estimations, at 0.04° and 0.2° horizontal resolutions, respectively, with more than 20 years of global data available.

2.2.3 Reanalysis datasets

Reanalysis datasets are generated by assimilating various observational data sources, such as satellite observations, ground-based measurements, and model simulations into a unified numerical framework (Donat et al., 2014). The aim is to produce a temporally and spatially continuous distribution of precipitation by interpolating and extrapolating information from areas where direct measurements are unavailable.

The spatial continuity of the datasets is achieved by estimating the unknown values between the known data points, using common interpolation methods like the nearest neighbour remapping, inverse distance weighting, bi-linear interpolation, and geostatistical methods (e.g., the kriging interpolation). Although some methods are better suited for specific applications, no explicit method is known to overall accommodate every aspect of reanalysis. For instance, bi-linear interpolation tends to produce smoother estimates over large areas but underestimates the total maximum rainfall and overestimates the total minimum, whilst nearest neighbour does not affect the total maxima of rainfall but can be positively biased due to the spreading of precipitation (Accadia et al., 2003). Kriging interpolation is prone to less errors (S. Vicente-Serrano et al., 2003), as it is a stochastic method that incorporates non-biased weights into the estimations and considers both the variability and spatial correlation present in the data (Tabios & Salas, 1985), but can be a complex and computationally heavy process.

The choice of assimilation techniques and model parameterisation can introduce biases and errors, particularly in earlier periods with sparse data coverage. So, the accuracy of the datasets depends greatly on the quality and coverage of past observations, as well as the processing of the models used (Bengtsson et al., 2004). Nonetheless, their ability to provide a coherent and homogeneous time series of precipitation estimates, even in regions with

sparse or irregular gauge networks, makes them valuable for studying large-scale climate phenomena and for assessing changes in precipitation patterns over decades.

Commonly used reanalysis precipitation datasets include products like CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) (Funk et al., 2015), ERA5 (Fifth Generation of the European Centre for Medium-Range Weather Forecasts Reanalysis) (Hersbach et al., 2020), and MSWEP (Multi-Source Weighted-Ensemble Precipitation) (Beck et al., 2019). Each dataset has its own strengths and limitations, and the choice depends on the specific requirements of the research or application. The field of atmospheric reanalysis is continuously evolving, with advancements in computational capabilities, improvements in observational technologies, and refinement in modelling techniques. Future directions include enhancing the resolution, extending reanalysis to include more components of the Earth's system, like ocean and land processes, and improving the representation of uncertainties.

2.3 Relevant research, gaps, and study aim

Preceding studies of similar interest, conducted on Greece, are mainly directed towards the spatial and temporal variability of rainfall, or focus on the analysis of the varying trends of precipitation, largely combined with the inspection of other climatic factors. A few of them evaluate gridded precipitation datasets, inasmuch as the availability of the data permits, but are confined by the limited extent of the time series used and the number of gridded datasets under evaluation. A common conclusion of these studies is the acknowledgment of precipitation as a highly variable physical phenomenon, recognising that complex topography, such as mountainous terrain, hinders its description to a sufficient degree. The usefulness of reliable gridded datasets is emphasised in those instances, where in situ data are unattainable or non-existent.

A range of studies have evaluated gridded precipitation over Greece, with varying results. Nastos et al. (2016) found that the TRMM 3B43 product generally underestimates high precipitation in high altitude areas, while Mavromatis & Voulanas (2021) identified E-OBS and Agri4Cast as the most reliable products, with E-OBS performing better in capturing variability and extreme precipitation. Feidas (2010) in contrast, highlighted the excellent performance of the TRMM 3B43 product, particularly at different spatial scales. Kazamias et al. (2022) further investigated the performance of GPM-IMERG rainfall estimates over a 5-year period and found that IMERG faces limitations in capturing orographic precipitation over western Greece and underestimates rainfall in islands and coastal areas. In a recent study, Alexandridis et al. (2023) evaluated the ERA5 and ERA5-Land datasets and concluded that ERA5-Land describes rainfall seasonality more accurately than ERA5, but both datasets underestimate precipitation during spring and summer.

However, these studies rely either on the use of a limited number of gauging stations and have examined a limited array of datasets or span over short periods of time. The present study evaluates nine datasets against 304 stations, considering both daily and monthly temporal scales. Some of these datasets are relatively new products, but are equipped with promising features, like the highest spatial resolution to date and a substantially long temporal coverage. Our analysis incorporates metrics for mean, seasonal, and extreme

precipitation, along with their respective trends, spanning from 1984 to 2016 and identifies the best performing datasets, inclusive of specific factors such as the update frequency and suitability for various types of applications.

Over the past several decades, there has been a significant increase in both the development and use of gridded precipitation datasets, offering new opportunities for the extensive monitoring of precipitation. These datasets are products of integrated reanalysis, satellite observations, terrestrial data, or a combination. In regions like Greece, where traditional gauge-based observations may be sparse or unevenly distributed, gridded datasets prove a solution by providing spatially continuous precipitation information, crucial for hydrological and climatic studies (Olmo & Bettolli, 2021; Najmi et al., 2023). Despite their utility, gridded datasets are not without limitations. Technical demands, intrinsic to the design of each dataset, can lead to inaccuracies in precipitation simulation, amplified in cases of elaborate climatic conditions, such as those of the Eastern Mediterranean. Validating these datasets against ground-based observations is therefore essential to ensure their reliability for regional applications (Bouizrou et al., 2023; Viviroli et al., 2011).

The present study aims to rigorously evaluate nine gridded precipitation datasets: ERA5-Land, AgERA5, CERRA-Land, CHELSA-W5E5 v1.1, MSWEP V2.8, CHIRPS05, IMERG V06, PERSIANN-CCS-CDR, and E-OBS, against field observations from around 300 gauging stations across Greece. This effort is unprecedented in its scope and depth for the study area, focusing on identifying the most reliable dataset for various Greek regions, therefore addressing a critical gap in regional hydrometeorological research. The extensive range of the study, which includes analysis on both daily and monthly timescales, facilitates an in-depth examination of hydrological processes, including the analysis of extreme events, as well as seasonal and spatial precipitation patterns. The methodology involves a comparative evaluation of these datasets against ground-truth, assessing their performance across different hydrological districts within Greece. This approach is pertinent given the complex climate dynamics of the Mediterranean, which distinctly challenge precipitation modelling and remote sensing. The outcomes of the study are expected to guide dataset developers in refining their methodologies and offer invaluable insights to practitioners selecting appropriate precipitation datasets for hydrological and agricultural applications in the Mediterranean region.

3. Data and Methods

3.1 In Situ Observations

3.1.1 Data selection

Rainfall data were collected from 307 gauging stations, 264 managed by the Hellenic Ministry of Environment and Energy (YPEN) and 43 by the Hellenic National Meteorological Service (HNMS). The data comprises rainfall values at a daily timestep which spans several decades, predominantly starting from 1945 and ending in 2020. The stations are distributed across the entirety of Greece, with the majority being in the water division of the West Central Greece and the Peloponnese peninsula, at various altitudes (**Figure 3.1**).

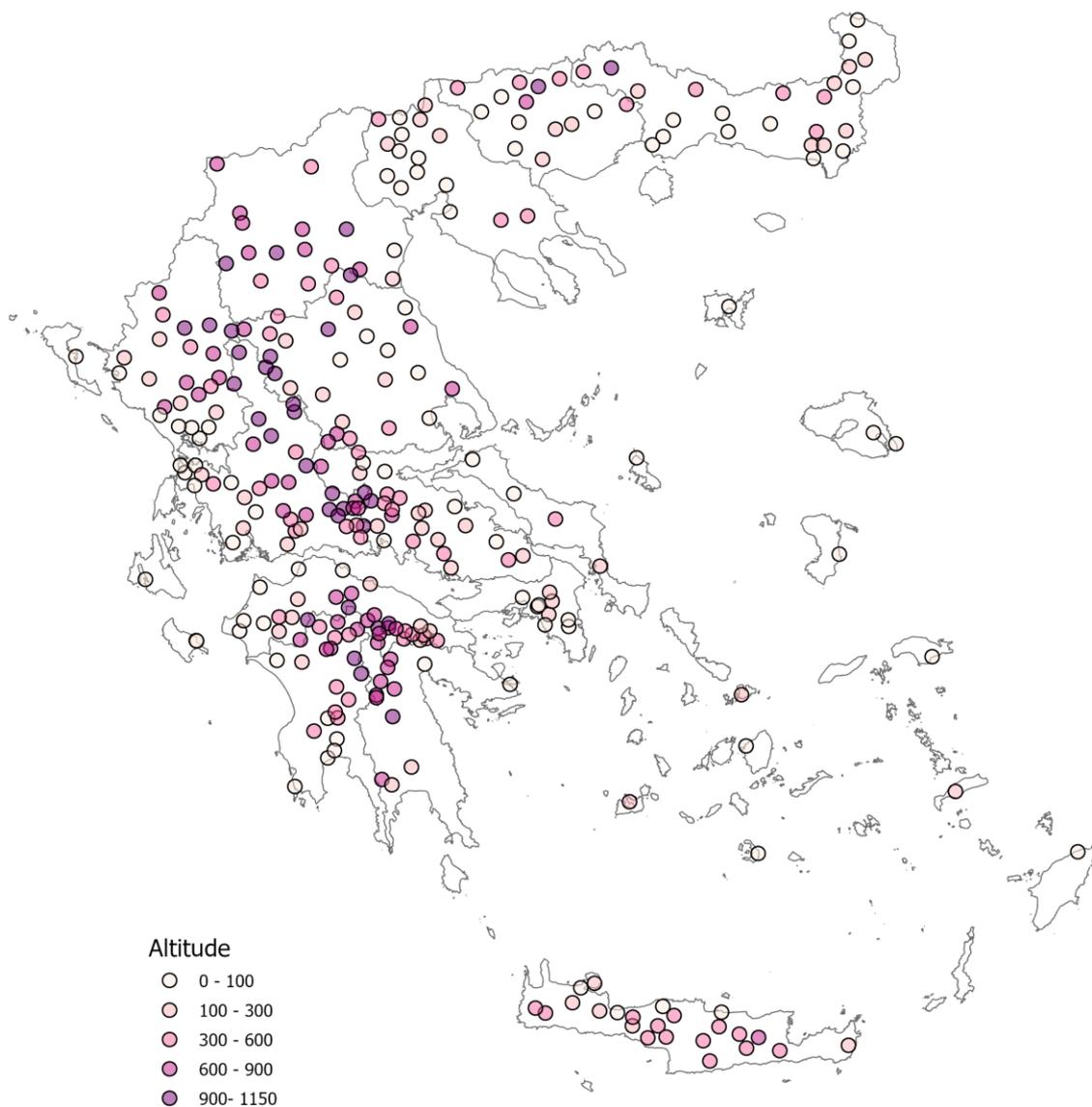


Figure 3.1. Location and elevation of rainfall stations and fourteen water districts of Greece.

3.1.2 Data processing

An in-depth evaluation of rainfall time series was conducted to identify and address potential errors related to precipitation frequency and abnormal values. The datasets obtained both from YPEN and HNMS, were based on raw daily data and had not been subjected to any process prior to this study, so a basic quality control has been implemented as an essential step in ensuring the integrity and usefulness of the precipitation data in our analysis.

Specifically, for values greater than 300 mm nearby stations were examined, and contingent upon inconsistency, values were excluded. Similarly, for values between the range of 200 mm and 300 mm, a thorough examination of neighbouring observations was performed to verify their validity, and in the case of spurious values the data were removed.

The gauging stations were operational for an average period of 57.9 years, with the longest duration being 77.3 years for station YPEN150, and the shortest duration being 9.5 years for station 200331. The period of available records, after excluding missing values, ranged from a minimum of 8.4 years (for station 200331) to a maximum of 73.1 years (for station YPEN150), with an average period of 55.1 years. To assess the applicability of the stations, we calculated a fill rate based on available data, ranging from 31.75% to 100%, with an average fill rate of 95.16%.

For 10 stations, up to 20 – 30% of the data were missing, and for 30 stations up to 10 – 20%, whereas 258 stations had less than 10% of missing data. Out of the 307 stations, 6 had a fill rate of 100%. Finally, we dismissed 3 stations with a fill rate of less than 70% (ID numbers: YPEN067, YPEN127, YPEN233), resulting in a total of 304 stations for further assessment. **Table 3.1** provides an overview of the fill rate and period of operation of the stations.

Table 3.1: Summary information of fill rate, operation period, and available data.

Fill Rate (%)			Operating period (years)			Available Data (years)		
Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum
31.75	95.16	100.00	9.50	57.88	77.33	8.38	55.08	73.07

In addition to the assessment of the daily time series, our analysis was extended to monthly time series since several applications are based on a monthly time scale. Thus, daily data were aggregated to monthly using a homogenisation and data filling method developed by Vicente-Serrano et al. (2009), which was applied to the overall dataset. The method aims at reconstructing incomplete precipitation series by combining the shortest time series from nearby gauging stations into a single dataset. Any discontinuities in the time series are corrected using the nearest neighbour interpolation method and then checked for errors. The homogeneity of the datasets is inspected through parameters relating to frequency and intensity of precipitation, and periods that do not meet the criteria are removed. The final data are composed of long-term continuous series, containing values that preserve the distribution of rainfall probability.

The application of the method resulted in a complete dataset of 251 out of 304 stations over the period 1961 – 2020. **Figure 3.2** compares the number of stations containing daily records, before and after implementing the homogenisation procedure.

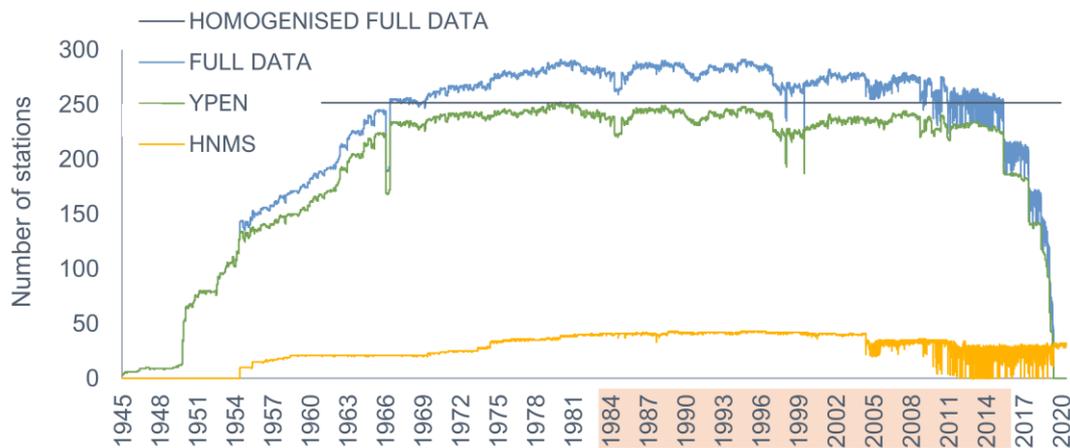


Figure 3.2. Number of stations with daily rainfall data over time for the full dataset (blue) and for the two different data providers, YPEN (green) and HNMS (yellow). The flat grey line corresponds to the monthly filled and homogenised dataset (251 stations). The shaded period (1984 – 2016) highlights the period of assessment.

3.2 Gridded Precipitation Datasets

3.2.1 Data selection

The evaluation encompasses nine distinct gridded precipitation datasets, classified under three broad categories: reanalysis, satellite, and gauge-based products. The following datasets were considered:

Reanalysis datasets:

ERA5-Land (Muñoz-Sabater et al., 2021)

The ERA5-Land dataset is a product of reanalysis on the ERA5 atmospheric model, which incorporates data collected from rain gauges into the estimated data. By emphasizing the ground-based components and considering the atmospheric variables of ERA5, the ERA5-Land provides a temporally consistent global dataset, at $0.1^\circ \times 0.1^\circ$ horizontal resolution, which covers more than seventy years (from 02/01/1950 to 31/12/2021) at hourly timesteps.

AgERA5 (Boogaard et al., 2020)

In the same manner, AgERA5 derives from reanalysis on the ERA5 model, by processing the hourly timesteps and integrating them into a dataset with a daily temporal resolution. The horizontal resolution of $0.1^\circ \times 0.1^\circ$ was achieved by interpolating ERA5 into a 0.1° grid, prior to applying specific regression equations used on the ECMWF’s high-resolution atmospheric model (HRES) and readjusted at a 0.1° resolution. The AgERA5 dataset consists of global surface meteorological data from 01/01/1950 to 31/12/2020 and is best suited for research pertaining to agriculture, due to it being adapted to accommodate a more detailed landscape.

CERRA-Land (Verrelle et al., 2022)

The CERRA-Land dataset is produced using the CERRA-Land system, which combines the SURFEX V8.1 land surface model, the CERRA atmospheric reanalysis model, and a daily accumulated precipitation analysis system (MESCAN). The system operates by transferring data from the CERRA reanalysis model into SURFEX, which in turn generates forecasts by consolidating the input data. The observations of the MESCAN system are used for the reanalysis process and affect the output indirectly through other variables (atmospheric forcings) included in the CERRA model. Succeeding the CERRA-Land system, the dataset covers the European continent at a 5.5 km x 5.5 km spatial resolution and contains daily precipitation estimates from 01/09/1984 to 31/12/2020.

CHELSA-W5E5 v1.0 (Karger et al., 2022)

The daily climate dataset provided by ISIMIP (CHELSA-W5E5 v1.0), is created by downscaling the W5E5 v1.0 dataset to a 1 km horizontal resolution. The model produces precipitation estimates from 01/01/1979 to 31/12/2016 and uses wind components from the ERA5 model to allow for better representation of precipitation in intricate topography. The CHELSA v2.0 downscaling algorithm implemented to achieve the 1 km spatial resolution, considers the contour of the terrain, and introduces readjustments influenced by orographic variables (e.g., altitude) but omits corrected estimations over the sea, where the readjustment is not applied.

MSWEP V2.8 (Beck et al., 2019)

MSWEP V2.8 combines satellite and gauge derived information to create global precipitation datasets at a 0.1° spatial and 3-hourly temporal resolution (from 01/02/1979 to 30/12/2020) and uses reanalysis to minimise temporal inaccuracies between the data. The dataset is suitable for areas with high convection, frontal precipitation, and dense rain gauge networks due to the three different means of administering data.

CHIRPS05 (Funk et al., 2015)

The CHIRPS05 dataset is a mixture of five different data sources: CHPClim, thermal infrared NOAA satellites, the TRMM 3B42 product (developed by NASA), the NOAA CFSv2 atmospheric model, and in situ observations; all aggregated into 5 – day intervals. Satellite observations are expressed as a percentage of the time where the temperature of the cloud tops falls below 235° K and are converted into millimetres of precipitation using local regression with the TRMM 3B42 precipitation data. The 5-day satellite estimates are reconveyed as Percent of Normal Precipitation indices, which are then multiplied by the corresponding CHPClim intervals to generate new unbiased estimates (CHIRP). The final dataset is available at a daily temporal (01/01/1981 – 31/12/2020) and 0.05° horizontal resolution by readjusting CHIRP values with data obtained from gauging stations.

Satellite datasets:

IMERG V06 (Huffman et al., 2020)

The IMERG V06 dataset is created by combining data from multiple satellite sensors to produce half-hourly and monthly precipitation estimates. The produced data are filtered through an algorithm that considers the errors of each sensor and provides a weighted combination of these estimates to obtain a more accurate and reliable estimation. The dataset contains global daily rainfall values from 19/06/2000 to 31/12/2020, at a 0.2° horizontal resolution.

PERSIANN-CCS-CDR (Sadeghi et al., 2021)

PERSIANN-CCS-CDR provides global precipitation estimates at 0.04° spatial and 3-hourly temporal resolution and has a temporal coverage of 37 years (01/01/1983 – 28/02/2021). The dataset is produced by merging the algorithms used in PERSIANN-CCS and PERSIANN-CDR, where the former is applied on the GridSat-B1 and NOAA CPC-4km satellite data, and the latter introduces Artificial Neural Networks to acquire estimations. The merged algorithm employs the Cloud Classification System of PERSIANN-CCS to accommodate cloud – rain distribution curves and bias adjustment is achieved using the Global Precipitation Climatology Project (GPCP) dataset.

Gauge-based datasets:

E-OBS (Cornes et al., 2018)

E-OBS is a daily (24-hour timestep) observational dataset that uses station data from the European Climate Assessment & Dataset (ECA&D) project, assisting in model validation, monitoring the European climate, and the study of daily extremes. The data are procured mainly from the European National Meteorological and Hydrological Services (NMHSs), without adjusting the time differences of data collection from each region. It offers a high spatial resolution of 0.1° and covers data from 01/01/1950 to 31/12/2021.

Succinctly, the AgERA5, ERA5-Land, and CERRA-Land datasets use the process of reanalysis, CHELSA-W5E5 uses gauge corrected reanalysis, IMERG (V6) uses gauge corrected information from satellites, and E-OBS is gauge-based. Two of the datasets, CHIRPS and MSWEP (V2.8), use both gauge and satellite correction for the reanalysis. Distinctively, PERSIANN-CCS-CDR was created by combining Artificial Neural Networks and satellite information. The datasets offer a high spatial resolution of 0.04°, 0.05°, 0.1°, 0.2°, 1 km, 5.5 km, and are also available at the same daily temporal resolution. **Table 3.2** summarises information on the precipitation datasets evaluated in this study, i.e., the dataset type, the spatial and temporal resolution, the temporal coverage, and the corresponding sources.

Table 3.2. Summary information of gridded precipitation datasets.

Name	Developer	Type	Spatial Resolution	Temporal Resolution	Temporal Coverage	Reference
ERA5-Land	ECMWF	Reanalysis	0.1°	Hourly	01/02/1950 – present	DOI:10.2438 1/CDS.E216 1BAC
AgERA5	ECMWF	Reanalysis	0.1°	Daily	01/01/1979 – present	DOI:10.2438 1/CDS.6C68 C9BB
CERRA-Land	ECMWF	Reanalysis	5.5 km	Daily	01/09/1984 – 30/04/2021	DOI:10.2438 1/CDS.A7F3 CD0B
CHELSA-W5E5 v1.0	ISIMIP	Gauge Corrected Reanalysis	1km	Daily	01/01/1979 – 31/12/2016	DOI:10.4836 4/ISIMIP.836 809.2
MSWEP V2.8	GloH2O	Reanalysis (Gauge + Satellite)	0.1°	3-Hourly	01/02/1979 – present	DOI:10.1175 /BAMS-D- 17-0138.1
CHIRPS05	CHG	Reanalysis (Gauge + Satellite)	0.05°	Daily	01/01/1981 – present	DOI:10.1038 /sdata.2015. 66
IMERG V06	NASA	Gauge Corrected Satellites	0.1°	Half-Hourly	19/06/2000 – present	DOI:10.5067 /GPM/IMER GDL/DAY/06
PERSIANN-CCS-CDR	CHRS	Gauge Corrected Satellites + ANN	0.04°	3-Hourly	01/01/1983 – present	DOI:10.1157 2/P24W2F
E-OBS	ECMWF	Gauge-Based	0.1°	Daily	01/01/1950 – present	DOI:10.2438 1/CDS.151D 3EC6

3.2.2 Data processing

Point scale time series of precipitation values were extracted from each dataset using the nearest neighbour interpolation method. Missing values or trace precipitation, similar to the in-situ data, were represented with values unlikely to occur. Those values were discarded, and trace precipitation values (between -0.1 mm and 0.1 mm) were substituted with values of zero.

In reanalysis and model data, trace precipitation refers to very small amounts of precipitation that are present but fall below a certain measurable threshold. In many cases, these small amounts are not captured accurately or are considered insignificant for specific applications. The threshold of 0.1 mm is commonly used to distinguish trace precipitation from measurable precipitation (Ma et al., 2009; Barrett et al., 2020). In the present study, precipitation events with values below 0.1 mm were considered as trace amounts and were approximated to 0.

3.2.3 Data pre-processing

The Climate Data Operators (CDO) software by the Max-Planck Institute for Meteorology (Schulzweida, 2023) was selected for its robust handling of gridded precipitation datasets, stored in netCDF files, and to produce manageable file formats. To extract the point scale timeseries we used the *-remapnn* operator provided by CDO, which performs the nearest neighbour interpolation method to transform the input square grids into isometric triangular tiles, thus distributing the values among finer and more precise points. This method was chosen because it preserves the original data values of the source points, suitable for variables like precipitation of daily temporal resolution, where maintaining peak values within the dataset is crucial. Following the interpolation, Python was utilised to conduct further data processing and calculate key evaluation metrics. Finally, the processed data and evaluation results were imported into ArcGIS Pro software by ESRI to create the corresponding maps.

3.3 Evaluation

For the assessment, a period of approximately 32 years (from 01/09/1984 to 31/12/2016) was selected, where daily precipitation measurements are available on most of the datasets and field data. The IMERG dataset was evaluated separately since the product of this dataset corresponds to the period 2000 – 2021, therefore restricting the assessment period to 2000 – 2016 and is denoted with an asterisk as IMERG* in the respective figures and tables.

3.3.1 Performance metrics

To evaluate the performance of the datasets in simulating precipitation, six indicators were calculated, two of which are used widely in hydrology.

Nash – Sutcliffe Efficiency [NSE] (Nash & Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum(E - O)^2}{\sum(O - \bar{O})^2} \quad (3.1)$$

where, E the estimated values, O the observed values, and \bar{O} the mean of the observed values.

NSE takes a maximum value of one, which states that the estimated values are comparable to the observed. A value of zero expresses an estimation ability on a par with that of the mean observed values and negative values indicate a poor prognostic performance of the datasets. Generally, a dataset with NSE equal or greater than 0.5 is considered to have acceptable performance (Duc & Sawada, 2023).

Kling – Gupta Efficiency [KGE] (Gupta et al., 1999):

$$KGE = 1 - \sqrt{(R - 1)^2 + \left(\frac{\sigma_E}{\sigma_O} - 1\right)^2 + \left(\frac{\bar{E}}{\bar{O}} - 1\right)^2} \quad (3.2)$$

where, \bar{E} the mean of the estimated values, σ_E the standard deviation of the estimated values, σ_O the standard deviation of the observed values, and R the Pearson correlation coefficient.

The values of the KGE metric can be compared to those of the NSE regarding model performance, with the exception being that KGE considers the linear correlation between the estimated and observed values, the variance of the values, and the bias error. Similar to NSE, positive values of KGE are associated with improved model performance and negative values with poor performance, although a specific threshold is not placed (Knoben et al., 2019).

Percent Bias [PBIAS]:

$$PBIAS = \left[\frac{\sum(O - E)}{\sum O} \right] \cdot 100\% \quad (3.3)$$

Percent Bias expresses the average deviation of the estimated values relative to the observed. Negative values of PBIAS indicate an underestimation tendency of the model and positive values an overestimation tendency. A value of zero is considered optimal, and near-zero values state accurate model estimation.

Root Mean Square Error (RMSE) [mm]:

$$RMSE = \sqrt{\frac{\sum(E - O)^2}{N}} \quad (3.4)$$

The RMSE metric measures the level of accuracy of a specific dataset, by calculating the square root of the mean squared errors between the estimated values and the observed data. Larger errors have a greater effect on the metric, and data points that have a significant difference (outliers) can produce inordinate results (Chai & Draxler, 2014). RMSE gives positive values and lower values indicate better model performance, while a value of zero is considered a perfect fit.

Mean Absolute Error (MAE) [mm]:

$$MAE = \frac{1}{N} \sum |O - E| \quad (3.5)$$

MAE calculates the average of the absolute errors between the observed and the estimated data. The produced errors represent the average absolute distance between the datasets and the $Y = X$ line.

Pearson Correlation Coefficient (R) (Pearson, 1895):

$$R = \frac{\sum(O - \bar{O})(E - \bar{E})}{\sqrt{\sum(O - \bar{O})^2} \sqrt{\sum(E - \bar{E})^2}} \quad (3.6)$$

Pearson's R assesses the linear correlation between the observed and the estimated data. The values range from -1 to 1 , with 1 (optimal value) signifying a positive relation and 0 a non-existent relation. A value of -1 indicates a negative correlation between the datasets, meaning that as one increases the other tends to decrease.

3.3.2 Precipitation Indices

Further to the typical evaluation metrics we calculated the following statistical measures, referred to as climate indices, which give insight on trends in mean and extreme precipitation patterns. The ETCCDI (Expert Team on Climate Change Detection and Indices), with the aim to develop a standardised set of climate indices for extreme events, created a flexible approach that has broad applicability and limits background noise originating from the input data. These indices derive from daily temperature and precipitation data, offering advantages over restrictions when managing daily datasets (Karl et al., 1999; Schär et al., 2016).

Annual total precipitation on wet days (PRCPTOT):

$$\text{PRCPTOT} = \sum \text{RR} \quad (3.7)$$

where, RR is the daily amount of precipitation above or equal to 1.0 mm.

Annual total 95th percentile precipitation events (R95p):

$$\text{P95p} = \sum \text{RR}_{95} \quad (3.8)$$

R95p expresses the sum of the daily precipitation on wet days ($\text{RR} \geq 1.0$ mm), where the amount of precipitation surpasses the 95th percentile of the given period.

Annual count of days when $RR \geq 10$ mm (R10mm):

$$R10mm = \sum I_{10} \quad (3.9)$$

where, I_{10} represents a day with a precipitation amount equal or greater than 10 mm.

Annual count of days when $RR \geq 20$ mm (R20mm):

$$R20mm = \sum I_{20} \quad (3.10)$$

where, I_{20} represents a day with a precipitation amount equal or greater than 20 mm.

Simple Precipitation Intensity Index (SDII):

$$SDII = \frac{\sum RR}{N} \quad (3.11)$$

where, RR dictates the amount of daily precipitation on wet days (equal or greater than 1.0 mm), and N equals to the number of wet days for that period.

4. Results

4.1 Precipitation patterns based on ground observations

To obtain a preliminary overview of the rainfall distribution across Greece, mean annual precipitation values were extracted from the homogenised time series for the hydrological years (from 1st of September to 31st of August) of 1962 – 2020. The spatial distribution of precipitation across Greece, as shown in **Figure 4.1**, highlights the substantial regional variations in rainfall patterns, which are intrinsically linked to the complex topography. The mean annual precipitation across the examined stations is 749.2 mm, but with large spatial disparities. The western mountainous part of the study region receives the highest precipitation accumulations, up to more than 2000 mm/yr, due to orographic uplift, while the lowest values (249.5 mm/yr) are recorded over the Aegean Islands (Santorini), as a result of their geographical location and prevailing wind patterns, which further contribute to their semi-arid conditioning. This variability poses a significant challenge for gridded precipitation products to realistically capture the complex precipitation patterns that are affected by the topography, the microclimates, and the interplay of atmospheric circulation patterns.

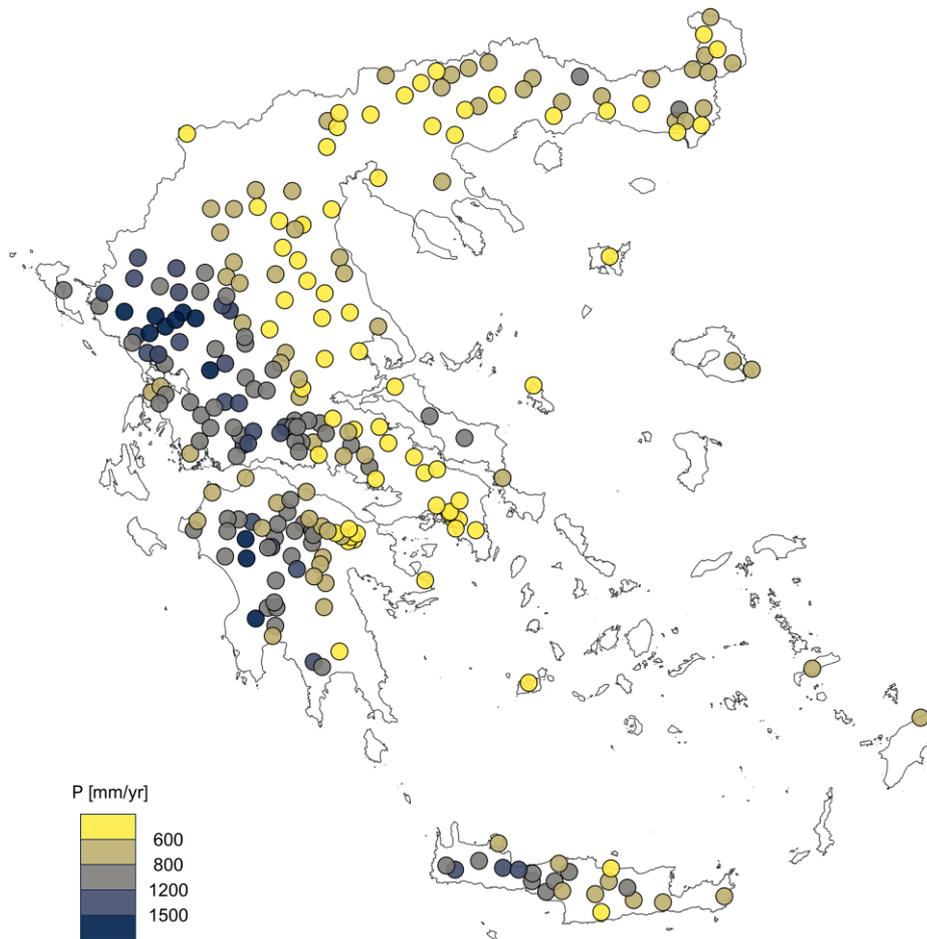


Figure 4.1. Mean annual precipitation of 251 homogenised timeseries for 1962-2020.

Regarding the homogenised time series, the largest mean annual precipitation amount was recorded on station YPEN084 with a total of 2331.8mm, and the smallest on station YPEN006 with a total of 281.6mm. The stations are situated in the water districts of West Central Greece and West Macedonia, at 941m and 808.2m, respectively. Whereas the maximum mean annual precipitation of the daily time series is seen on station YPEN084 with a value of 2157.3mm and the minimum on station WMO16744 with 250.6mm (located in Santorini of the Aegean Islands water district, at 19m). Attica and East Central Greece receive the least amount of rain, as rainfall decreases from West to East and South to North.

The network of the available stations included in this study is considered sufficiently dispersed among the water districts, therefore eliminating larger errors that derive from the congregation of stations to a particular area. **Table 4.1** contains the number of stations in the daily and monthly datasets, as well as the water district in which they are located, the total area of each district, and the area that each station covers.

Table 4.1. Number of stations and station coverage in the fourteen water districts.

Water District	Area (km ²)	Number of stations		Station Coverage (km ² /station)	
		Daily	Monthly	Daily	Monthly
GR01	7234	22	18	329	402
GR02	7396	33	24	224	308
GR03	8442	14	11	603	767
GR04	10496	42	36	250	292
GR05	9980	24	21	416	475
GR06	3186	9	8	354	398
GR07	12290	29	26	424	473
GR08	13141	26	26	505	505
GR09	13619	16	11	851	1238
GR10	10164	16	7	635	1452
GR11	7320	14	13	523	563
GR12	11242	24	22	468	511
GR13	8344	23	21	363	397
GR14	9141	12	7	762	1306

Figure 4.2 illustrates the spatial and seasonal precipitation variability across the fourteen water districts of the country. Greece experiences the heaviest precipitation during winter, predominantly in the western part of the mainland, the Peloponnese peninsula, and western Crete, averaging at 294.9 mm. The maximum mean DJF (December-January-February) rainfall amounts to 840.8 mm, observed in the water district of West Central Greece. These seasonal heterogeneities are less pronounced across the northern region of the country, as the mean JJA (June-July-August) rainfall varies between 8.2 mm and 220.9 mm and the mean DJF rainfall does not exceed 546.5 mm. Overall, the mean annual precipitation increases with altitude, at a rate of 39.3 mm/100 m across the stations. **Figure 4.3** offers supplemental information on the spatial distribution of mean seasonal precipitation, as derived from the observations.

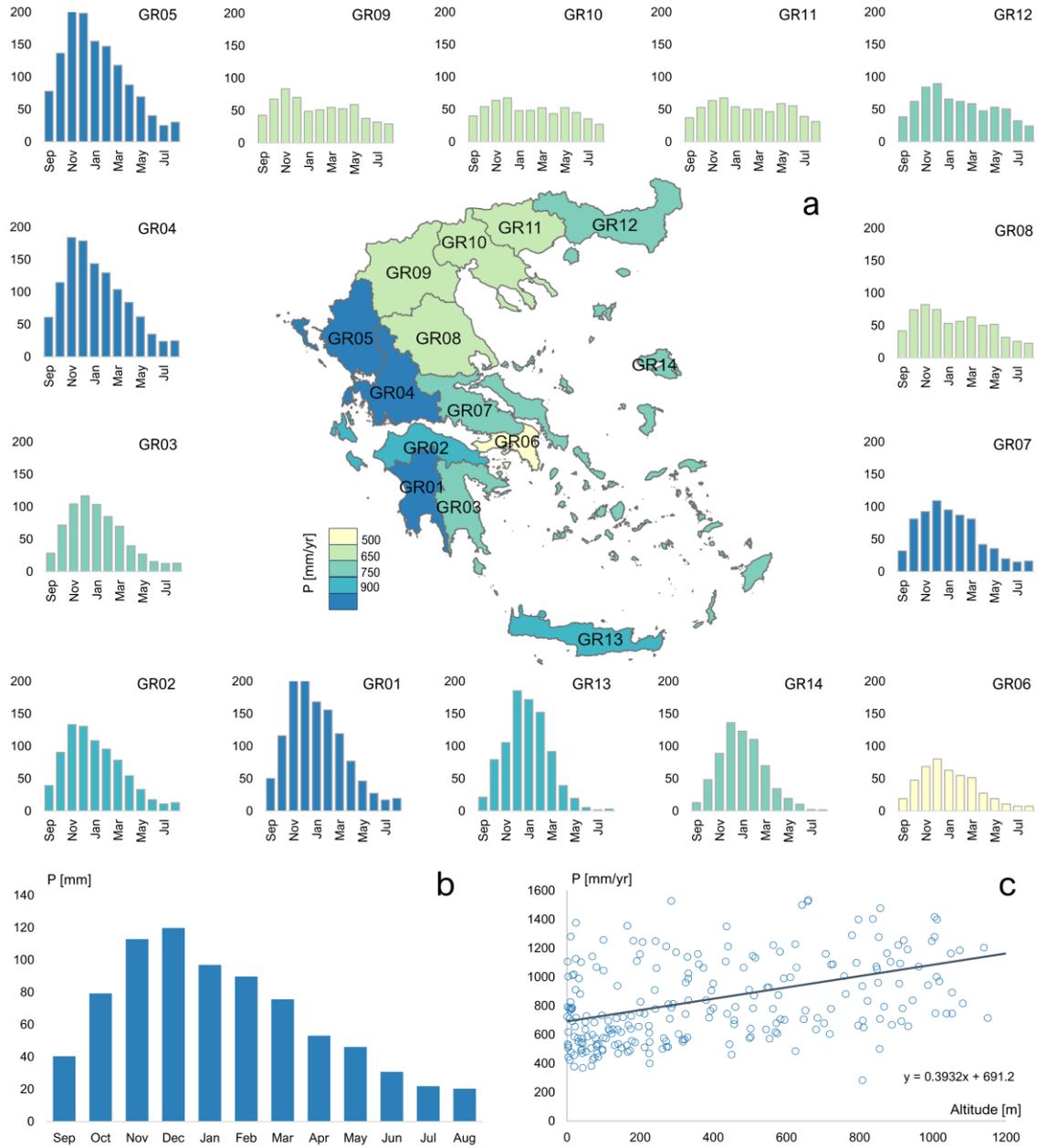


Figure 4.2. a) Mean annual precipitation at the level of the fourteen water districts is illustrated in the central map and their corresponding seasonal distribution in the surrounding bar-graphs. (b) illustrates the mean seasonal rainfall of Greece and (c) the variation of mean annual precipitation with altitude.

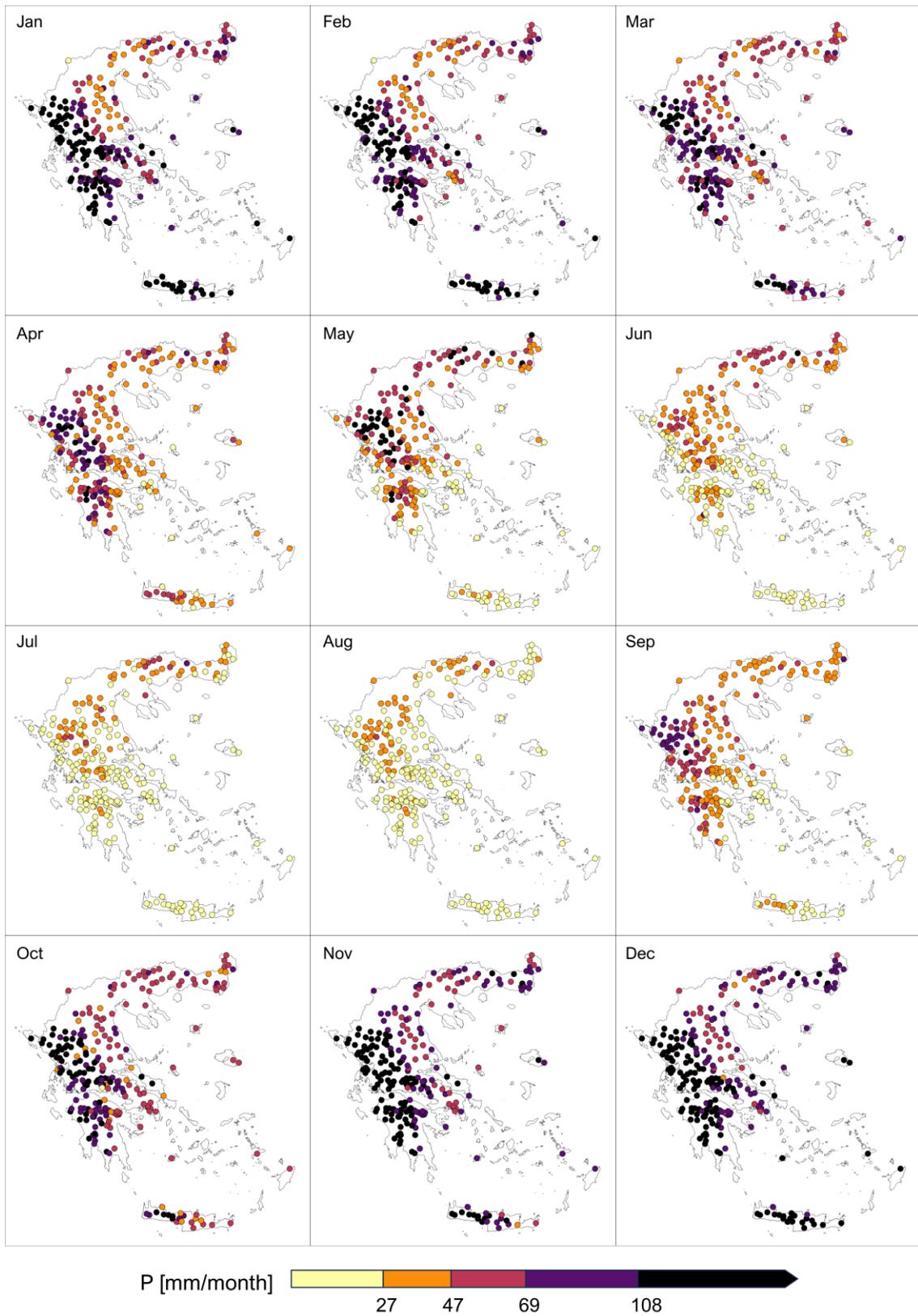


Figure 4.3. Spatial distribution of seasonal variability of mean accumulated precipitation across stations.

The spatial variability of rainfall, and whether a relation between the stations exists, was further investigated with the k-means clustering method using nine descriptive parameters: the latitude, the longitude, the 95th percentile threshold, the mean annual precipitation, the mean seasonal December-January-February (DJF), June-July-August (JJA), March-April-May (MAM), September-October-November (SON) precipitation, and the mean annual precipitation above 20 mm of every station. The application resulted in eight different subdivisions of stations, as shown in **Figure 4.4**, grouped by their similar characteristics.

Cluster 3 presents the highest values of each feature, with a mean annual precipitation range of 1241.8 mm to 2087.6 mm and a median value of 1380.9 mm (**Table 4.2**). The stations included in Cluster 3 are mostly located in the water district of Epirus and West Central Greece. Cluster 2 displaying dissimilar mean annual accumulations to Cluster 3, despite being geographically near, can be attributed to the Pindus Mountain Range casting a rain shadow effect on its leeward side (Sindosi et al., 2015). All stations receive the heaviest rainfall during winter (DJF) and autumn (SON), which in Cluster 8 accounts for up to 78% of the mean annual total.

A research by Markonis et al. (2017), employing, among others, the k-means clustering method, reached a similar conclusion. Their analysis was conducted on 136 monthly and annual time series, which largely encompasses a period between 1940 and 2012, with the objective of providing new information on the spatiotemporal variability of precipitation across Greece. They decidedly created eight regional subgroups, depending on the climatic aspects of each area. Larger emphasis was placed on the longitude of the station rather than the altitude, stating that accumulation of stations above a certain elevation (1000 m) could produce significant bias and skewed results. The stations incorporated in this research are situated between 1 m to 1200 m, above sea level, and have an adequately consistent coverage across different altitude ranges, therefore an upper elevation boundary is not applied.

Based on their study, the West presents the highest precipitation values, while the lowest are observed over the North, the North-East, and Eastern Central Greece. Similarly, the highest rainfall accumulations can be seen in Cluster 3 and Cluster 6 and the lowest in Clusters 2, 5, and 1 of this study, which mostly correspond to the same general vicinities and hydrological districts. Also, in accordance with our findings, they detected a strong seasonal variability in the South, as well as clusters of comparable characteristics in the western midlands and western Peloponnese.

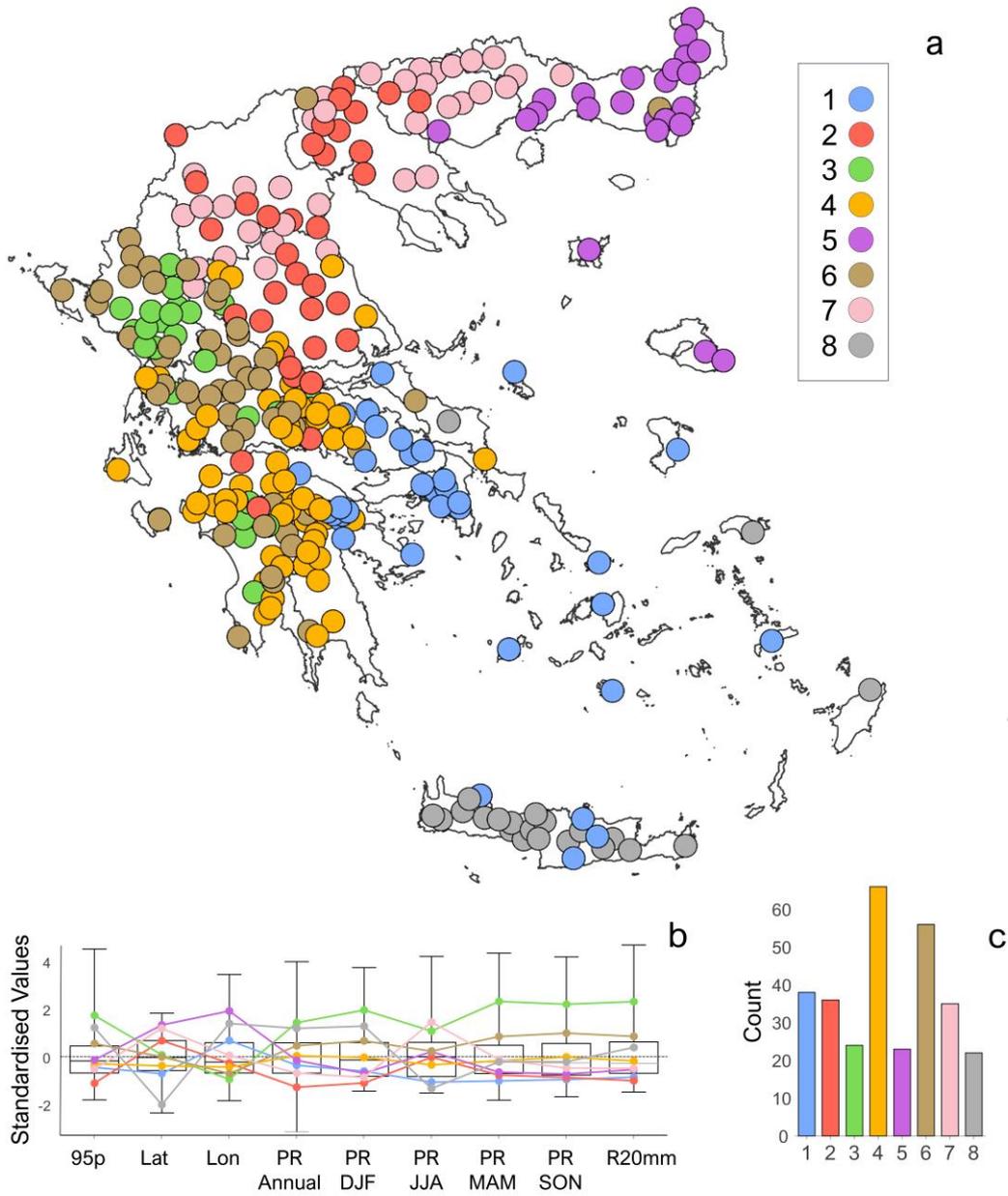


Figure 4.4. a.) Visualisation of eight station clusters and their location. b.) Analysis fields (parameters) and c) number of stations in each cluster.

Table 4.2. Number of stations and median values of features in the clusters.

Cluster ID	Number of stations	Annual (mm)	DJF (mm)	MAM (mm)	JJA (mm)	SON (mm)	95p (mm)	R20mm (mm)
1	38	468.7	201.2	102.2	21.2	130.3	34.3	6.6
2	36	450.9	128.7	119.5	60.8	133.7	25.9	5.0
3	24	1380.9	560.8	309.3	104.8	419.2	57.6	26.6
4	66	715.7	290.9	156.6	46.4	223.7	36.0	10.7
5	23	520.7	179.7	121.1	71.4	150.5	38.7	8.1
6	56	1015.6	389.5	223.0	68.8	323.2	45.9	17.9
7	35	612.4	163.5	157.5	109.4	159.3	33.0	8.4
8	22	877.0	473.2	163.0	6.4	211.3	52.4	14.9

Figure 4.5 provides an overview of the results based on the ETCCDI climate indices, with a focus on extreme precipitation. The median R95p index (accumulated precipitation of wet days exceeding the 95th percentile) across all stations is 152.2 mm/yr, with values reaching up to 478.3 mm/yr, indicating the existence of areas with precipitation extremity in the western and central parts of Greece. The medians of maximum 1-day (rx1) and 5-day (rx5) precipitation are 132.4 mm/d and 225.5 mm/d, respectively, but vary substantially among the inspected stations, while the precipitation intensity of wet days (SDII) is 12.4 mm/d by median. The median values of the R10mm and R20mm indices, showing the annual frequency of moderate and heavy rainfall days, respectively, are 22.5 d/yr and 9.9 d/yr. The patterns of the examined indices suggest that while the study region experiences a considerable degree of variability in extreme precipitation events, there are consistent patterns that emerge, particularly in the occurrence of heavy rainfall.

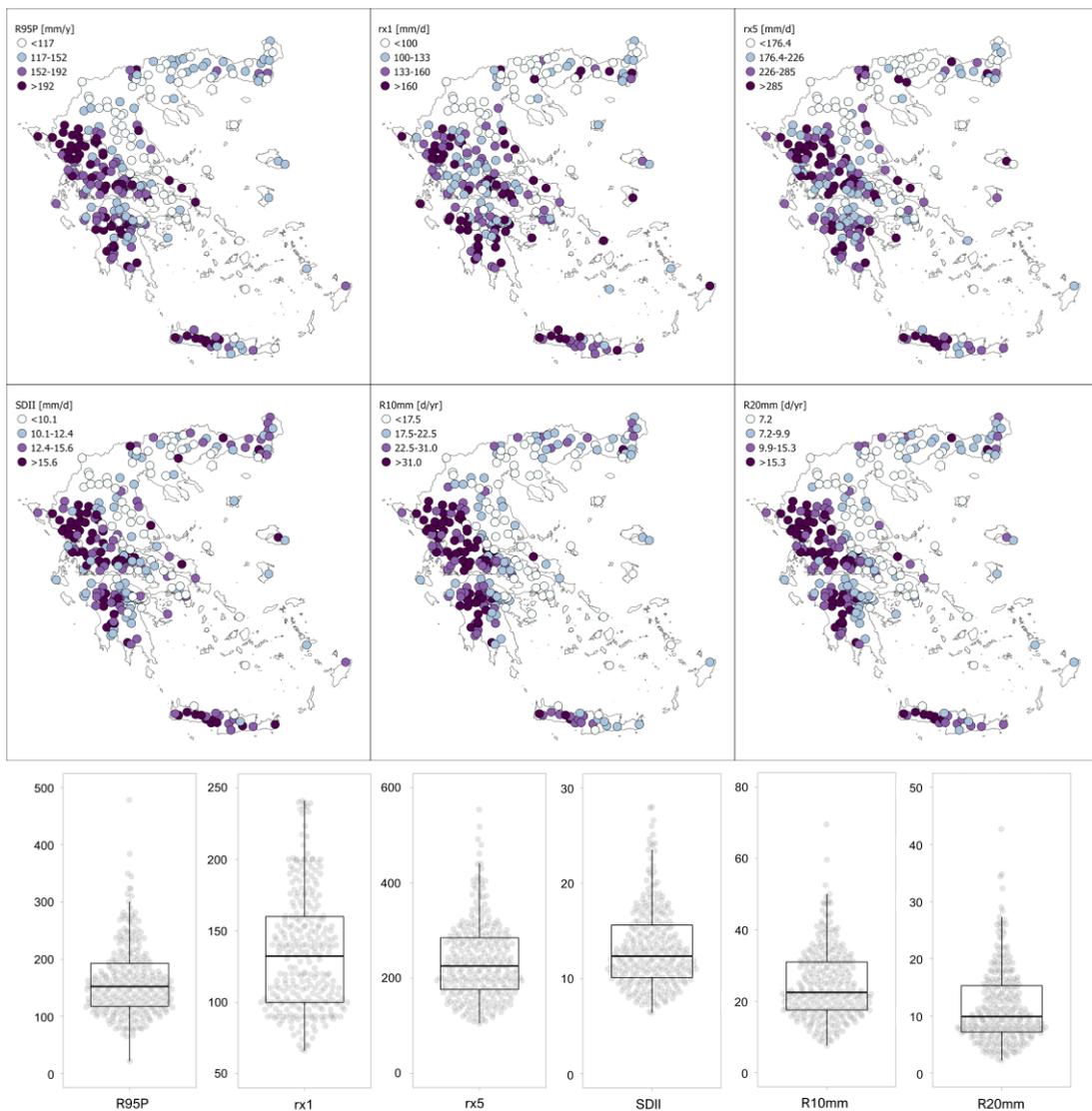


Figure 4.5. Spatial distribution and median values of ETCCDI climate indices for R95p, rx1, rx5, SDII, R10mm and R20mm.

A further scope of this study is to broadly identify trends in extreme climate indices and their spatial variability across the study area and assess the consistency of the derived trends with the corresponding estimates from the gridded datasets. For the description of the trends found in ETCCDI climate indices, we used the non-parametric Theil-Sen (Sen, 1968) approach, as well as the Mann-Kendall statistical test (Mann, 1945; Kendall, 1948), both widely used methods in hydrological studies (Hamed, 2008; Y. Trambly et al., 2013; Westra et al., 2013; Tzani et al., 2019). The tests were conducted on a 32-year period (from September of 1984 to August of 2016), at a 95% confidence level. The outcome of the statistical analysis may serve as a benchmark in the examination of the gridded datasets and their capacity of supporting realistic trend analyses.

The trends of the R95p index of 66 stations (21.71% of the total) are estimated as statistically significant, increasing with an average decennial rate of 33.1 mm. In comparison, 103 stations (37.5% of the total) exhibit a statistically significant trend of the PRCPTOT index, with an average decadal increase of 114.4 mm. Approximately, a third of those stations are found in northern Greece, specifically in western, northern, eastern Macedonia, and Thrace. The distribution of the PRCPTOT and R95p trends suggests that the two occurrences do not necessarily coincide with each other. Meaning that one area could receive less precipitation through high-intensity, short-term events, while another could display an overall increase in precipitation accumulation, which is more spread out and long-term. **Figure 4.6** showcases the spatial variability of the PRCPTOT and R95p trends and their statistical significance, across Greece.

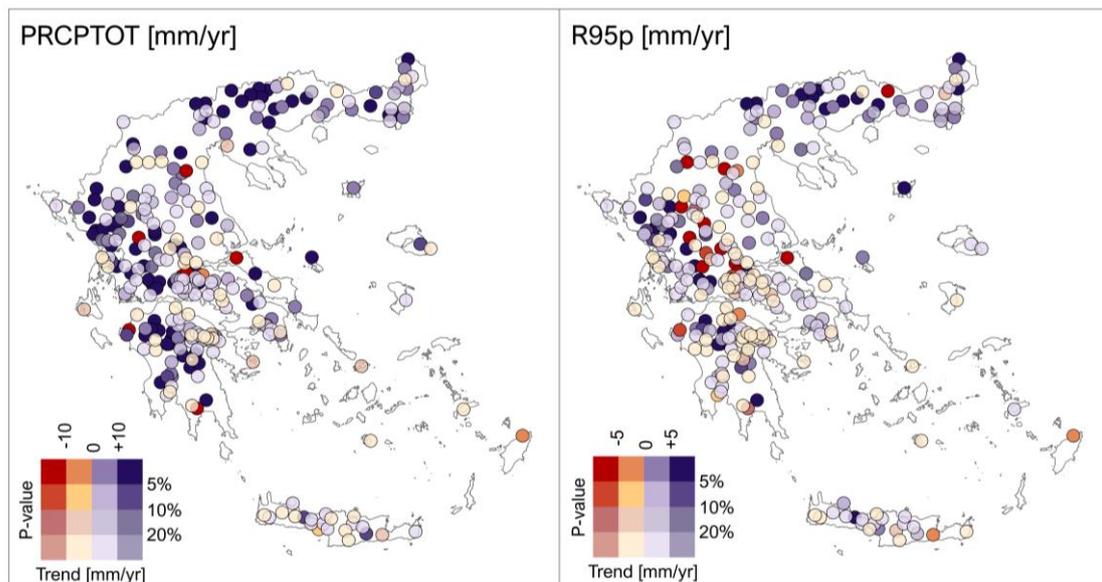


Figure 4.6. Trends of PRCPTOT and R95p indices for 1985 – 2016.

4.2 Precipitation patterns in gridded datasets

All datasets deviate in the representation of the mean annual rainfall; most evident when compared visually against the recorded data (**Figure 4.7**). The degree of misrepresentation varies for each dataset and might depend on the availability of data records, the composition

of the model, or the method used to achieve the resolution of interest (downscaling). Ostensibly, CERRAL captures the spatial variability of precipitation more accurately, while maintaining a resolution of 5.5 km, but slightly overestimates the rainfall accumulation in northern regions. The ERA5L, CHIRPS, AgERA5, and IMERG datasets seem to competently interpret the spatial variability of heavier precipitation, although the small-scale variability between neighbouring grid cells is reduced. Both CHELSA and MSWEP produce similar results, with MSWEP displaying a greater over-smoothing effect, whereas E-OBS vastly underestimates the mean annual accumulation across the entire country. The largest disparities are observed in the water districts of East Macedonia, Thrace, Crete, as well as the Aegean Islands, but can be partially attributed to lack of sufficient station data until recent years.

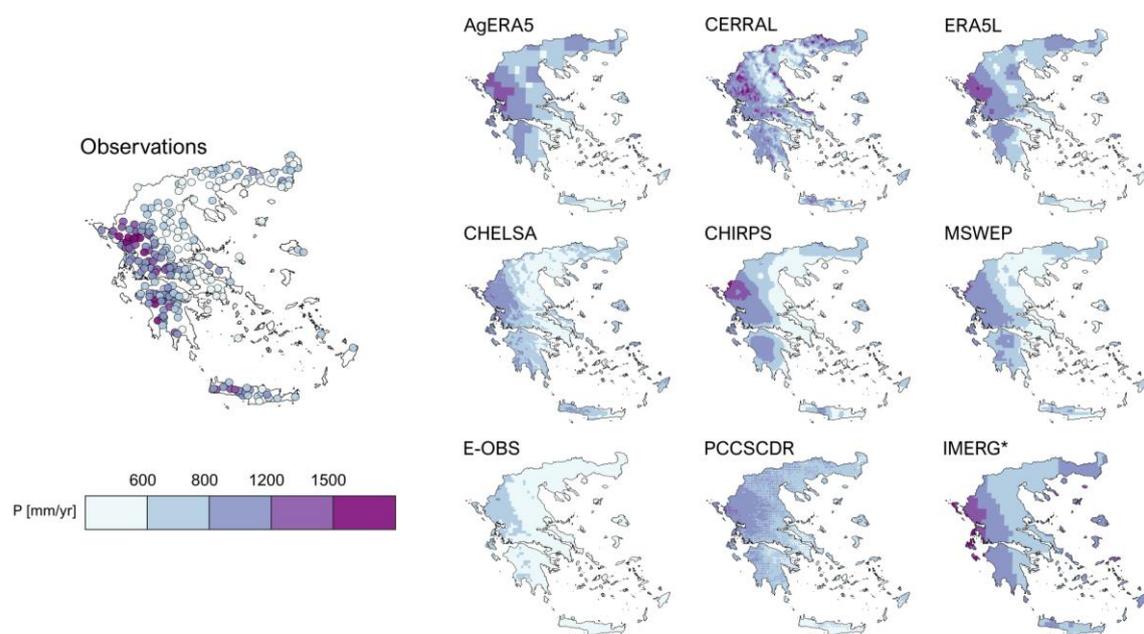


Figure 4.7. Mean annual rainfall distribution resulting from the nine gridded datasets and point observations of the 1985 – 2016 period. *For IMERG: mean annual rainfall corresponds to the 2000 – 2021 period.

CHELSA and AgERA5 performed relatively better in terms of NSE, with median values of 0.23 and 0.16, respectively. However, none of the models demonstrate acceptable performance based on this metric, as the range of NSE values varies from -1.11 (PCCSCDR) to 0.23 (CHELSA). Regarding KGE, MSWEP achieved the highest median score of 0.36, while PCCSCDR had the lowest at 0.06. All datasets suggest a slight positive correlation between the observed and estimated precipitation, out of which CHELSA presents the highest correlation coefficient (R), equal to 0.50, and PCCSCDR the lowest, equal to 0.17. **Figure 4.8** shows the performance of the gridded datasets in 304 stations, concerning NSE, KGE, and R , along with the respective median values of each metric.

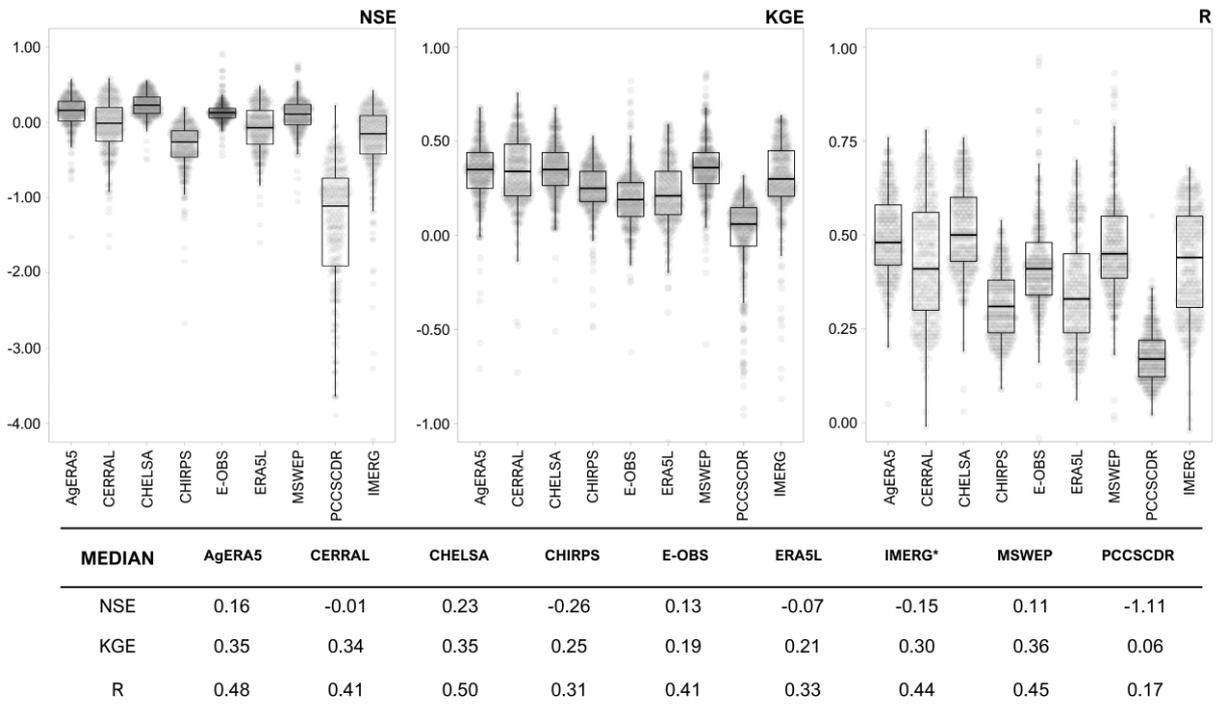


Figure 4.8. NSE, KGE, R score values for 1985 – 2016, derived from 304 daily time series, and the median value of the nine gridded precipitation datasets.

Figure 4.9, Figure 4.10, and Figure 4.11 provide the spatial distribution of NSE, KGE, and R, respectively, in 304 stations across the country, and the median values of the metrics for the nine datasets.

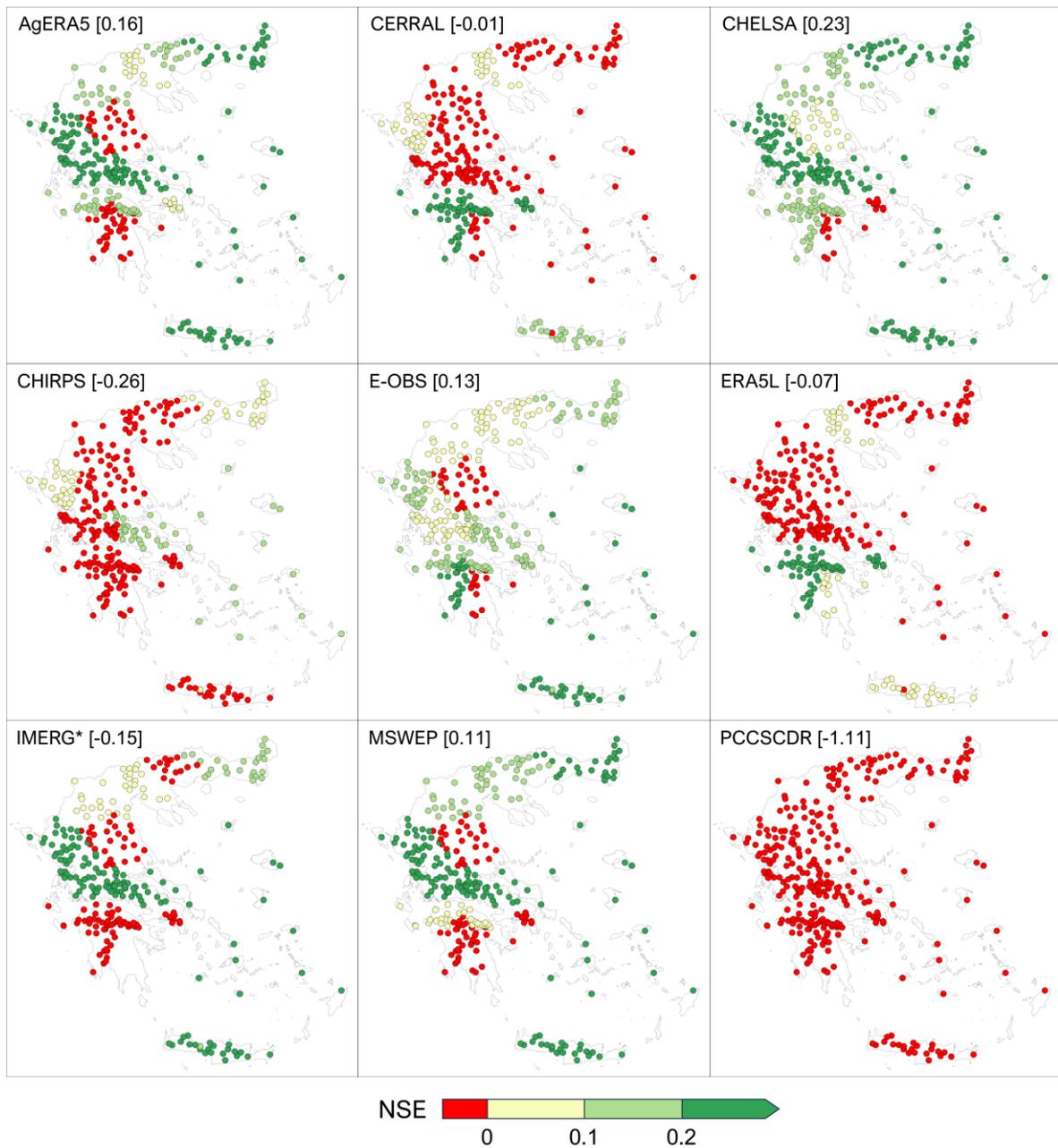


Figure 4.9. NSE values of 304 daily time series and median values of the nine gridded datasets.

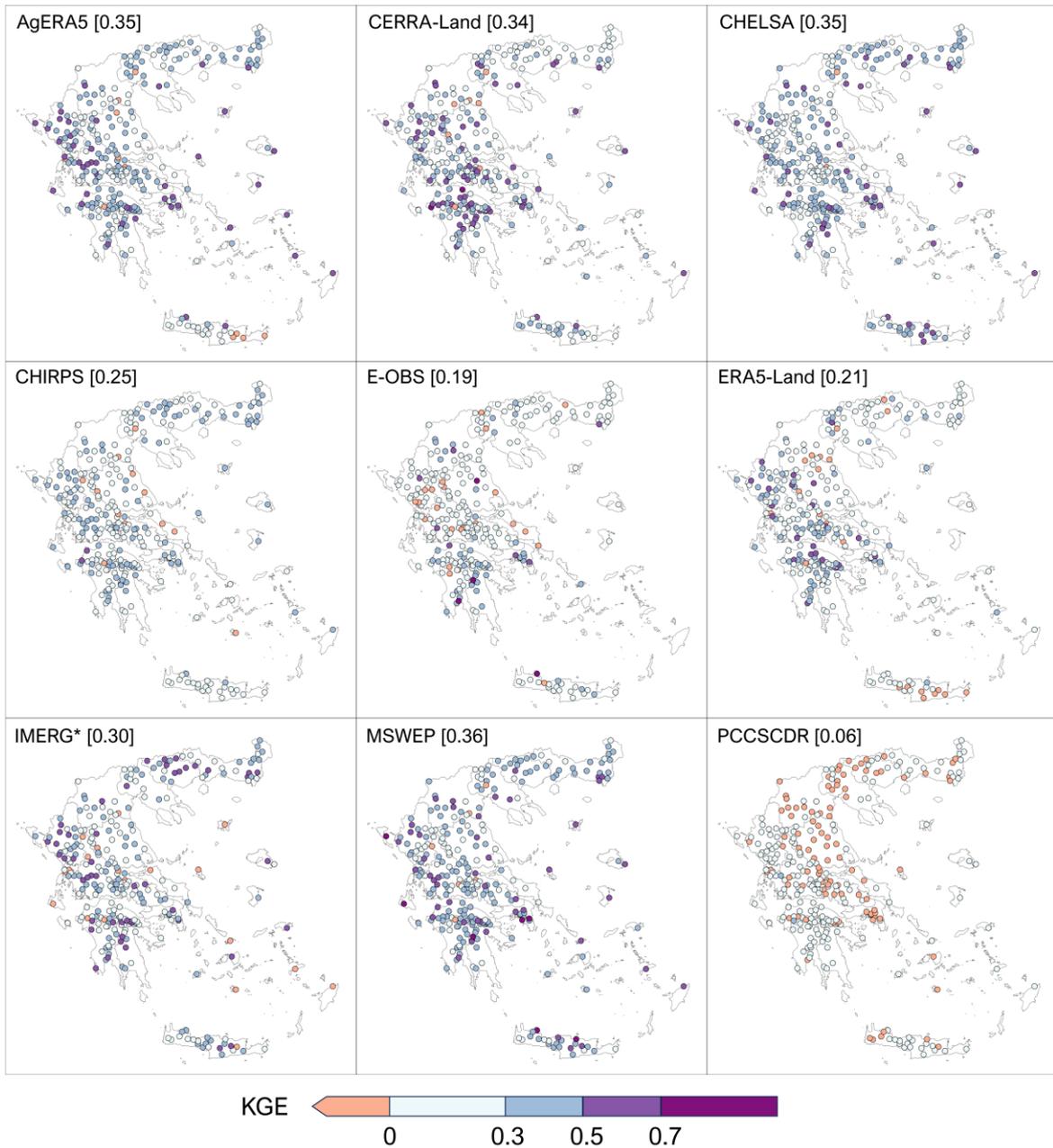


Figure 4.10. Median values of KGE for the nine datasets.



Figure 4.11. R values of 304 daily time series and median values of the nine gridded datasets.

IMERG, ERA5L, AgERA5, CERRAL, and PCCSCDR tend to overestimate daily precipitation by a maximum of 8.17% (IMERG), while E-OBS, CHELSA, MSWEP, and CHIRPS underestimate it by a maximum of 30.93% (E-OBS). Specifically, PCCSCDR presents the lowest overestimation, reflected in a PBIAS value of -0.81% . The datasets exhibit similar MAE scores, ranging from 2.14 mm (E-OBS) to 3.18 mm (PCCSCDR) and produce minimal fluctuations in the root-mean-square error (RMSE), as the values range from 6.30 mm for CHELSA to 10.95 mm for PCCSCDR. The performance of the gridded datasets in 304 stations, concerning MAE, RMSE, and PBIAS, along with the respective median values of each metric are shown in **Figure 4.12**.

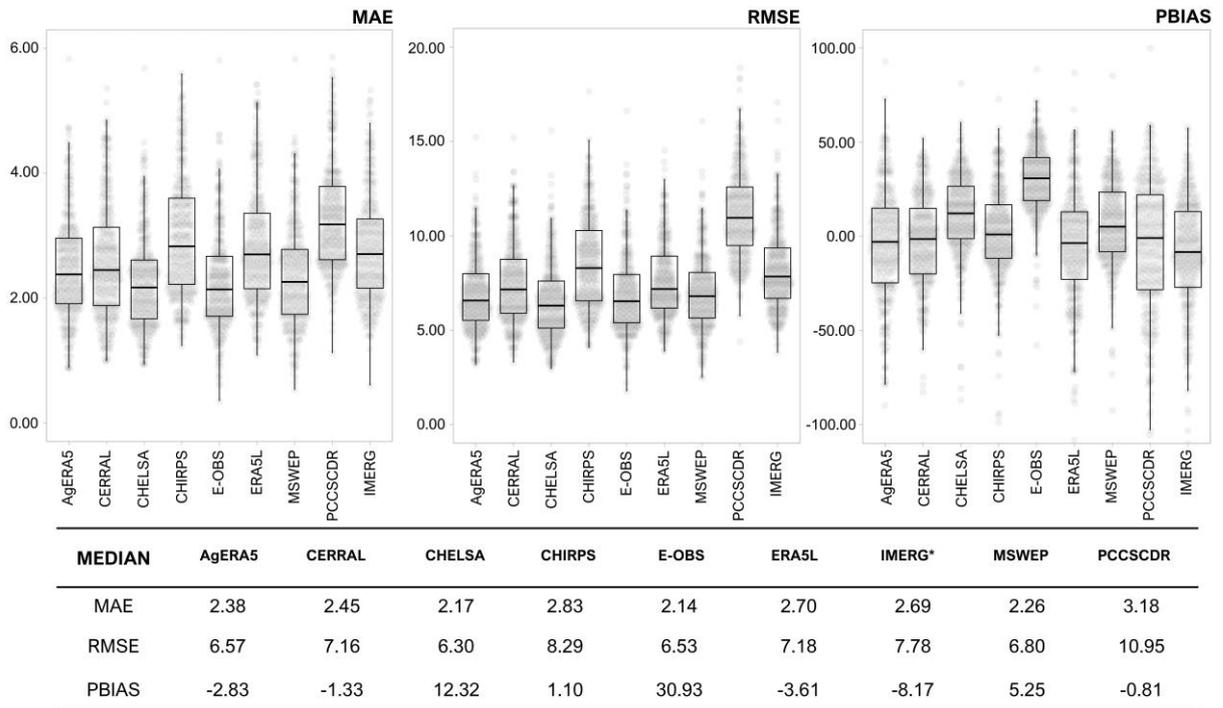


Figure 4.12. MAE, RMSE, PBIAS score values for 1985 – 2016, derived from 304 daily time series, and the median value of the nine gridded precipitation datasets.

Figure 4.13, Figure 4.14, and Figure 4.15 provide the spatial distribution of MAE, RMSE, and PBIAS, respectively, in 304 stations across the country, and the median values of the metrics for the nine datasets.



Figure 4.13. MAE values of 304 daily time series and median values of the nine gridded datasets.



Figure 4.14. RMSE values of 304 daily time series and median values of the nine gridded datasets.

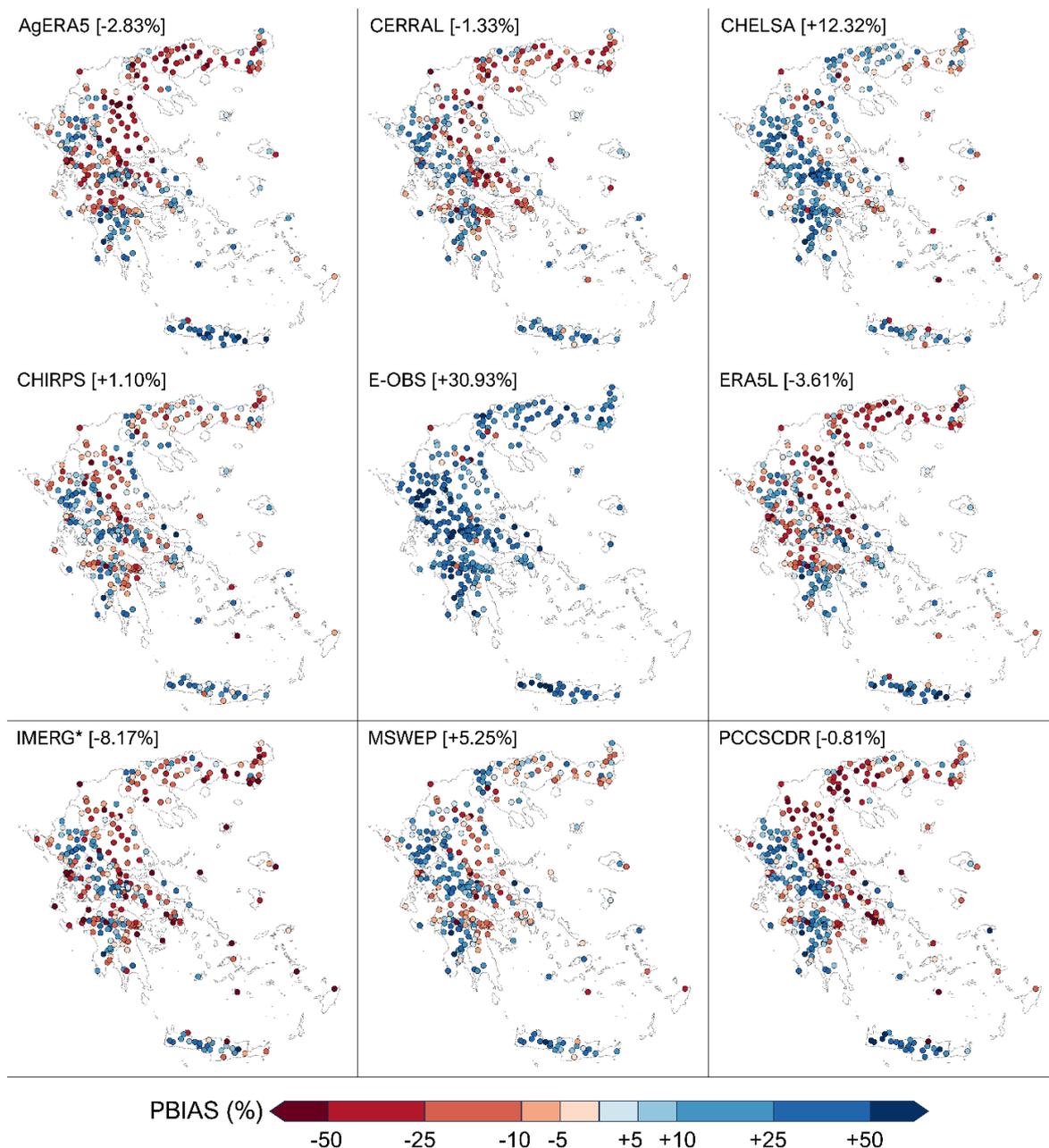


Figure 4.15. PBIAS values of 304 daily time series and median values of the nine gridded datasets.

On a point scale, AgERA5, CERRAL, ERA5L, IMERG, and PCCSCDR overestimated rainfall at over 50% of the stations, whereas CHELSA, CHIRPS, E-OBS, and MSWEP underestimated rainfall, on average, at 70.7% of the stations. The datasets, seemingly, perform worse at higher altitudes, apart from IMERG, PCCSCDR, MSWEP, and ERA5L, where NSE displays a minor improvement within the 700 – 1200 m range (Table 4.3). The KGE scores remain constant for both ERA5L and PCCSCDR.

The percentage of bias reveals an increasing tendency towards underestimation at higher elevation, likely due to the limitations of the datasets in accurately capturing diverse types of hydrometeors and the smoothing effect of the chosen interpolation method (Accadia et al., 2003). The input data for E-OBS are procured mainly from the European National

Meteorological and Hydrological Services (NMHSs) and are limited to the availability and density of the station network of each country. In higher altitudes, where rain gauges might be insufficient, the interpolation methods used can generate significant inaccuracies and bias errors (Hofstra et al., 2009). In that respect, E-OBS produces the highest absolute value of PBIAS, equal to 33.63%.

Many techniques use only temperature boundaries or cloud classification to calculate orographic precipitation and do not consider the upward change in atmospheric pressure. A mechanism that ultimately dictates the formation of clouds and by extension precipitation (Houze, 2012). For example, PCCSCDR utilises the predated PERSIANN-CCS (PCCS) dataset as input, which uses infrared sensors to detect rain occurrences. Infrared satellites can mistake high temperature clouds as non-precipitating instances and disregard rain above a certain temperature threshold. In addition, the cloud classification system incorporated in PCCS, requires manual characterisation of cloud types, which reduces the variations of precipitating clouds to their most evident features (Sadeghi et al., 2021). These cases highlight the inherent challenges and limitations of rainfall estimation across varying elevations and point out the need for improved methodologies that focus on data integration to enhance accuracy in complex terrains.

Table 4.3. Median values of the statistical metrics of the nine datasets and number of stations, at three different altitude ranges.

Altitude (m)	Name	NSE	KGE	PBIAS	RMSE	MAE	R
0 – 200	AgERA5	0.18	0.38	-12.85	5.91	1.18	0.51
No. of stations	CERRAL	-0.03	0.35	-5.13	6.61	1.21	0.44
123	CHELSA	0.26	0.39	3.19	5.36	1.06	0.54
	CHIRPS	-0.17	0.29	-2.98	6.83	1.40	0.35
	E-OBS	0.13	0.24	25.47	5.63	1.07	0.41
	ERA5L	-0.16	0.22	-14.73	6.77	1.43	0.33
	IMERG*	-0.20	0.31	-25.39	7.28	1.31	0.50
	MSWEP	0.15	0.40	-4.13	5.93	1.09	0.49
	PCCSCDR	-1.37	0.06	-19.86	10.16	1.65	0.20
Altitude (m)	Name	NSE	KGE	PBIAS	RMSE	MAE	R
200 – 700	AgERA5	0.17	0.35	3.89	6.85	1.14	0.48
No. of stations	CERRAL	0.03	0.31	-0.34	7.22	1.20	0.39
113	CHELSA	0.23	0.34	16.46	6.46	1.06	0.50
	CHIRPS	-0.32	0.24	3.98	8.52	1.39	0.29
	E-OBS	0.14	0.17	32.71	6.63	1.06	0.42
	ERA5L	-0.02	0.20	4.73	7.25	1.26	0.33
	IMERG*	-0.10	0.32	-1.29	8.05	1.22	0.43
	MSWEP	0.11	0.36	8.47	6.97	1.13	0.44
	PCCSCDR	-1.02	0.06	6.48	11.24	1.50	0.18
Altitude (m)	Name	NSE	KGE	PBIAS	RMSE	MAE	R
700 – 1200	AgERA5	0.13	0.32	5.91	7.15	1.15	0.44

No. of stations							
68	CERRAL	0.03	0.33	3.54	7.75	1.17	0.41
	CHELSA	0.18	0.30	18.26	6.88	1.07	0.46
	CHIRPS	-0.36	0.23	4.07	9.37	1.39	0.26
	E-OBS	0.11	0.16	33.63	7.11	1.06	0.39
	ERA5L	-0.01	0.22	4.79	7.51	1.23	0.34
	IMERG*	-0.14	0.28	4.88	8.38	1.22	0.34
	MSWEP	0.08	0.33	12.73	7.28	1.13	0.42
	PCCSCDR	-0.99	0.06	15.24	11.39	1.47	0.14

Figure A1 (in the Appendix) presents the metric scores of CHELSA in three altitude ranges, as an indication of the performance of the datasets when considering the change in elevation.

Regionally, there are discernible variations in the performance of the datasets. AgERA5 performs well within the central part of the country, resulting in the best values of the statistical metrics in Cluster 1, and displays acceptable performance throughout the West and South (i.e. Clusters 3, 6, 8). CHELSA estimates rainfall with greater precision than the rest of the products in northern and northcentral regions but possesses a larger underestimation bias in western parts. Overall, MSWEP and E-OBS are unable to interpret the distribution of precipitation, as they deliver the poorest scores across the entirety of the clusters. **Table A1**, of the Appendix, contains the daily metric scores of the datasets, regarding the 8 Clusters.

The efficiency of the datasets improves on a monthly timescale, as demonstrated by the statistical metrics depicted in **Figure 4.16** and **Figure 4.17**, which exhibit greater uniformity and minimal discrepancies. PCCSCDR displays the poorest performance in five out of the six metrics (NSE = -0.03, KGE = 0.38, RMSE = 68.32 mm, MAE = 42.06 mm, R = 0.52) but produces a satisfactory PBIAS with a value of 1.21%, revealing possible inaccuracies in the temporal representation of precipitation or significant limitations of reproduction of short-term precipitation events. The CHELSA dataset shows promising results with a NSE and KGE equal to 0.61, RMSE of 37.02 mm, MAE of 23.55 mm, R of 0.83, and PBIAS of 13.02%, but is limited to a data record of 38 years, from 1979 to 2016 (at the time of this study), making it most suitable for research pertaining to that assessment period. Evidently, when assessed on a monthly timescale, CERRAL and CHELSA achieve the best median scores out of the nine products, while PCCSCDR exhibits the poorest performance across both temporal scales.

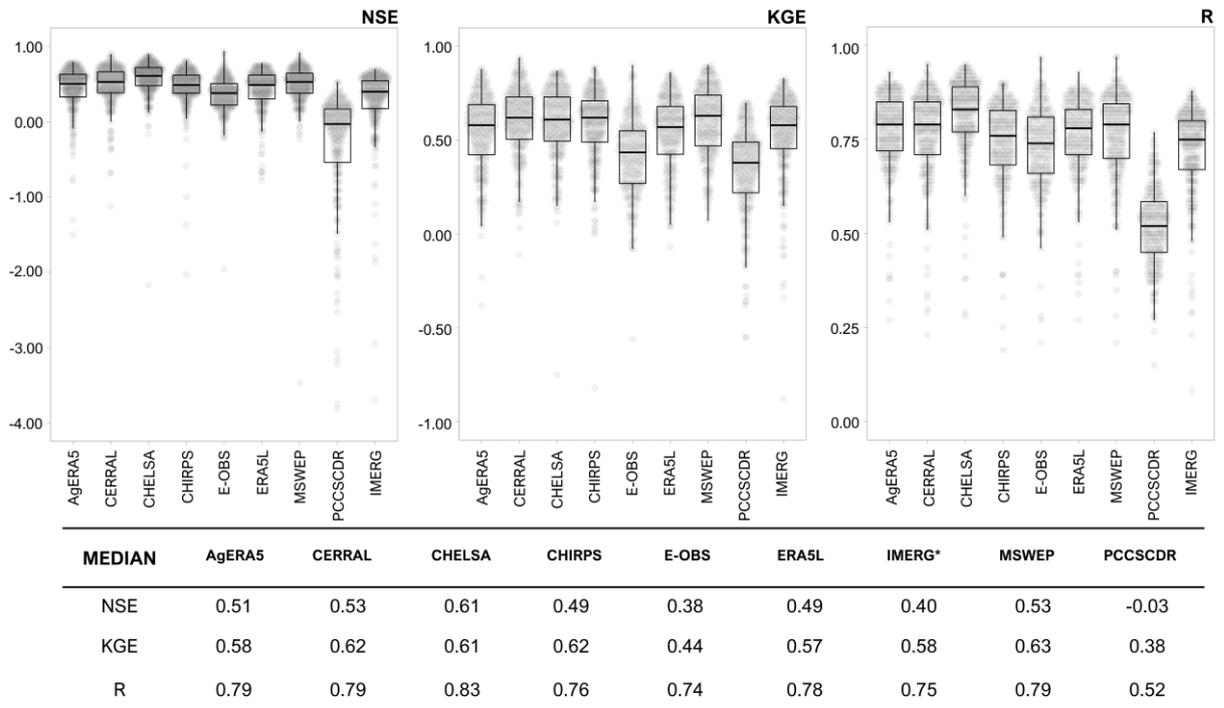


Figure 4.16. NSE, KGE, R score values for 1985 – 2016, derived from 251 monthly time series, and the median value of the nine gridded precipitation datasets.

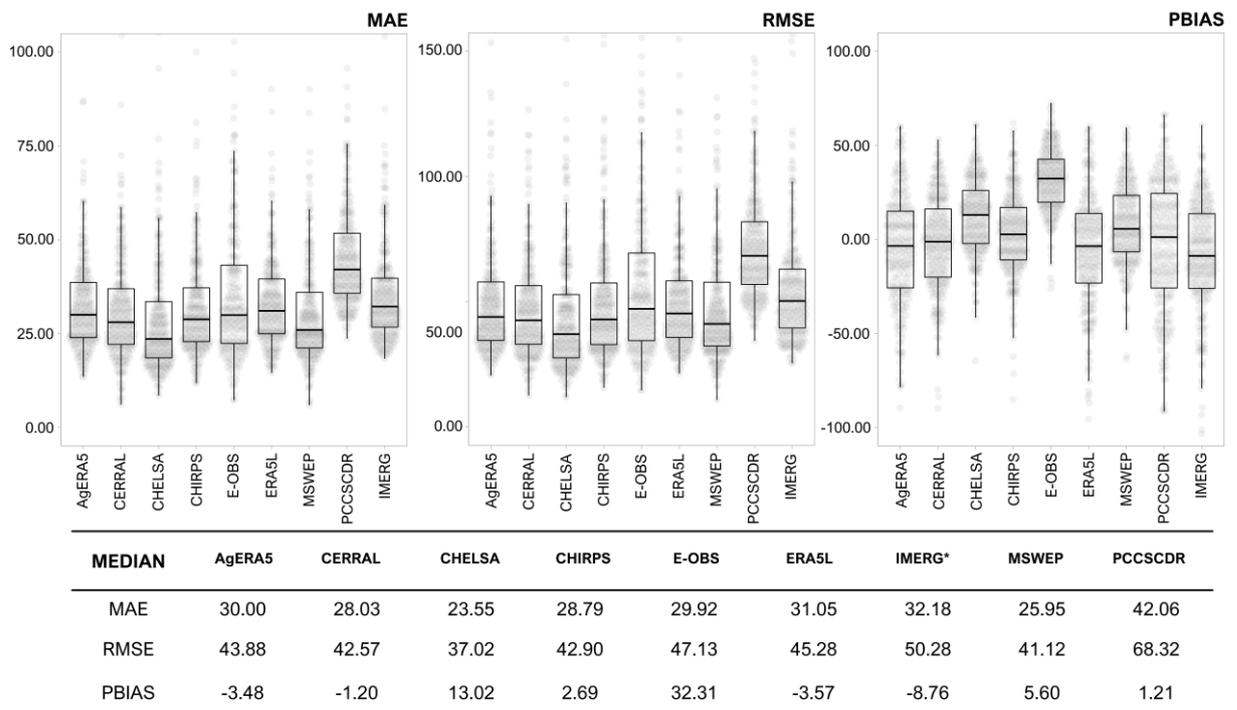


Figure 4.17. MAE, RMSE, PBIAS score values for 1985 – 2016, derived from 251 monthly time series, and the median value of the nine gridded precipitation datasets.

The competence of the datasets in estimating extreme precipitation events, was also evaluated using the ETCCDI indices and contrasted with those obtained from station data. The 95th percentile threshold, in most datasets, is overestimated by a maximum of –56.96% (IMERG) and underestimated by a maximum of 54.87% (E-OBS). PCCSCDR vastly overestimates the daily maxima of rainfall, scoring a PBIAS value of –162.86%, but performs relatively well in moderate and heavy precipitation (PBIAS of 4.78% and 6.63%, respectively), which further confirms the temporal obscurities of this specific product. IMERG produces a PBIAS value of –56.96% when estimating the 95th percentile threshold but displays a minor improvement in less acute conditions. E-OBS underestimates all aspects of extreme precipitation by equable amounts.

The strength of the correlation between the datasets and the observations is low or medium in most extreme indices, except for CERRAL, CHELSA, CHIRPS, and MSWEP where they produce values that range from 0.14 to 0.75 and a mean R value equal to 0.55, among the four datasets. CERRAL and CHELSA outperform the other products by producing mean R scores, of 0.56 and 0.60 respectively, that suggest moderate correlation. **Table 4.4** includes the median values of R and PBIAS of the datasets in relation to extreme precipitation, as well as the median values of the ETCCDI indices, as estimated by the observations. The median scores of R and PBIAS within the 8 Clusters, with respect to the extreme indices of the datasets, are presented in **Table A2**, which is located in the Appendix.

Table 4.4. Median correlation coefficient (R) and percent bias of extreme indices for the nine datasets, between 1985 – 2016. Disclosed in brackets are the corresponding values of each index, as derived from the observations. *The extreme indices of IMERG were based on the period between 2000 – 2020.

		AgERA5	CERRAL	CHELSA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [152.2 mm]	R	0.52	0.64	0.70	0.62	0.51	0.50	0.47	0.62	0.40
	PBIAS	-19.33	-28.68	7.00	6.13	54.87	15.64	-56.96	-11.85	-54.39
rx1 [132.4 mm]	R	0.21	0.23	0.43	0.33	0.21	0.18	0.11	0.14	0.13
	PBIAS	27.03	-8.40	40.91	10.46	50.70	30.32	-18.37	-0.17	-162.86
rx5 [225.5 mm]	R	0.29	0.49	0.52	0.47	0.24	0.27	0.25	0.20	0.20
	PBIAS	22.65	0.63	33.85	14.65	56.17	27.18	-4.32	13.04	-73.45
SDII [12.4 mm/d]	R	0.38	0.56	0.55	0.54	0.50	0.37	0.31	0.42	0.38
	PBIAS	39.00	27.82	41.40	-20.23	46.99	41.75	23.67	36.31	-4.82
R10mm [22.5 days]	R	0.65	0.73	0.72	0.74	0.76	0.67	0.67	0.75	0.47
	PBIAS	-4.28	-5.58	13.48	-15.65	44.88	-1.41	-8.84	7.20	4.78
R20mm [9.9 days]	R	0.55	0.72	0.68	0.73	0.63	0.57	0.55	0.67	0.50
	PBIAS	25.98	13.75	45.60	-8.30	79.34	28.51	-2.24	32.41	6.63

The trends of the PRCPTOT index produced by the datasets seen in **Figure 4.18**, are notably lower in value and exhibit less fluctuation compared to the trends based on observations. As revealed by the coefficient of variation (CV) in **Table 4.5**, the datasets score lower in that metric against the observations, meaning that the slope estimations deviate only slightly from the mean slope value and indicating that the datasets are incapable of replicating the minute regional characteristics of precipitation. Not true for datasets that demonstrate inaccuracies on account of deficient data (i.e. E-OBS and IMERG).

Statistically significant trends are detected predominantly in larger values of slope, both for the datasets and the observations. The variations of the slope with respect to the mean slope value, also in the case of the statistically significant trends, display minimal contrast in all datasets and do not coincide with the value assumed by the observations. Correspondingly, the linear correlation (R) between the observational and estimated indices is low and veers on non-existent in the statistically significant constrained subset of stations. Similar indications emerge from the trend analysis of the R95p index (**Table 4.6** and **Figure 4.19**).

Table 4.5. Values of mean slope, and its coefficient of variation (CV) and R metric, fraction of joint significance, as well as CV, R, and mean values of jointly significant stations, based on the statistical analysis of the PRCPTOT index. Joint significance fraction (JSF) refers to the fraction of statistically significant stations of the datasets, where both the gridded datasets and the station observations (OBS) present statistical significance. CV is an indicator of the average deviation of the values from the mean value, relative to the mean of the specific series. The data provided by IMERG were proven statistically insignificant and the corresponding estimations were not calculated. **Single value.

PRCPTOT	OBS	AgERA5	CERRAL	CHELSA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
Mean Slope	6.01	6.02	6.25	6.43	4.65	-3.98	6.15	-2.16	4.09	6.99
CV	1.62	0.71	0.75	0.52	0.62	-2.71	0.71	-3.04	0.85	0.93
R	-	0.30	0.33	0.31	0.24	0.12	0.31	0.01	0.32	0.13
JSF	-	24%	23%	29%	19%	25%	24%	0%	16%	21%
Mean Slope of JS	13.61	9.44	10.03	8.26	7.08	-1.10	9.51	-18.89**	8.03	12.28
CV of JS	0.81	0.36	0.36	0.34	0.38	-11.99	0.40	-	0.31	0.27
R of JS	-	0.11	0.31	0.27	0.33	0.01	0.17	-	0.18	0.02

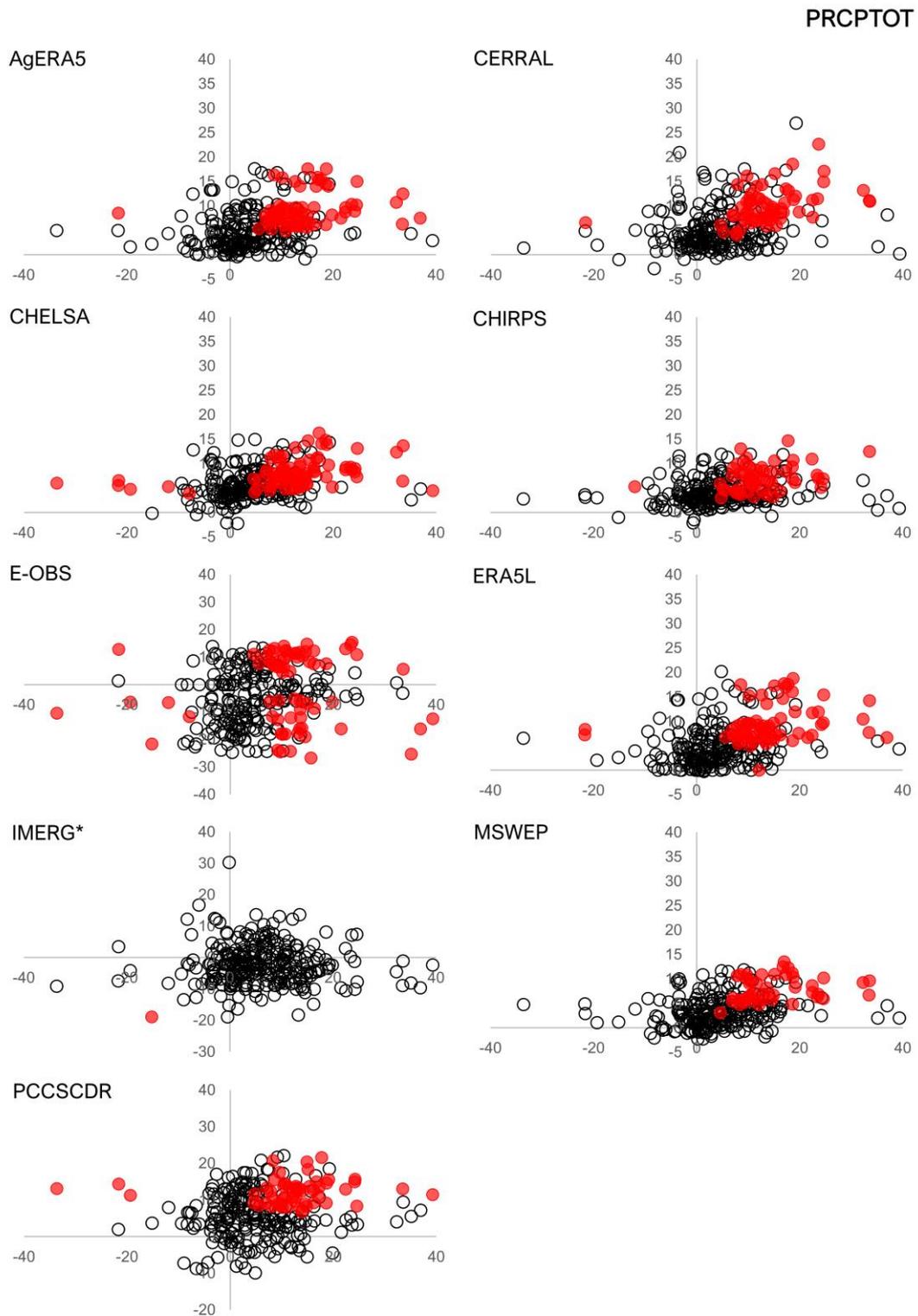


Figure 4.18. Scatter plots of the jointly statistically significant slope values (red points) and the remaining slope values (black points), regardless of significance, based on the statistical analysis of the PRCPTOT index. The vertical axis corresponds to the observational values, whereas the horizontal axis to the values of the datasets. Note that the range and size of the vertical axis is not identical among all datasets.

Table 4.6. Values of mean slope, and its coefficient of variation (CV) and R metric, fraction of joint significance, as well as CV, R, and mean values of jointly significant stations, based on the statistical analysis of the R95p extreme index. Joint significance fraction (JSF) refers to the fraction of statistically significant stations of the datasets, where both the datasets and the observations (OBS) present statistical significance. CV is an indicator of the average deviation of the values from the mean value, relative to the mean of the specific series. The estimations marked in bold italic were derived from two values.

R95p	OBS	AgERA5	CERRAL	CHELSA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
Mean Slope	0.33	1.67	1.32	1.80	-1.11	1.98	1.68	-2.32	1.47	-5.77
CV	2.95	1.36	2.11	1.10	-1.93	1.10	1.32	-1.84	1.59	-0.78
R	-	0.31	0.30	0.31	0.19	0.14	0.32	-0.06	0.23	0.00
JSF	-	9%	6%	13%	5%	16%	10%	1%	11%	12%
Mean Slope of JS	4.46	4.54	6.19	3.66	-0.86	3.25	3.92	<i>11.64</i>	4.19	-9.09
CV of JS	0.96	0.23	0.49	0.28	-4.37	0.55	0.43	<i>0.08</i>	0.42	-0.25
R of JS	-	0.24	0.23	-0.01	-0.08	-0.06	0.38	<i>1.00</i>	0.09	0.21

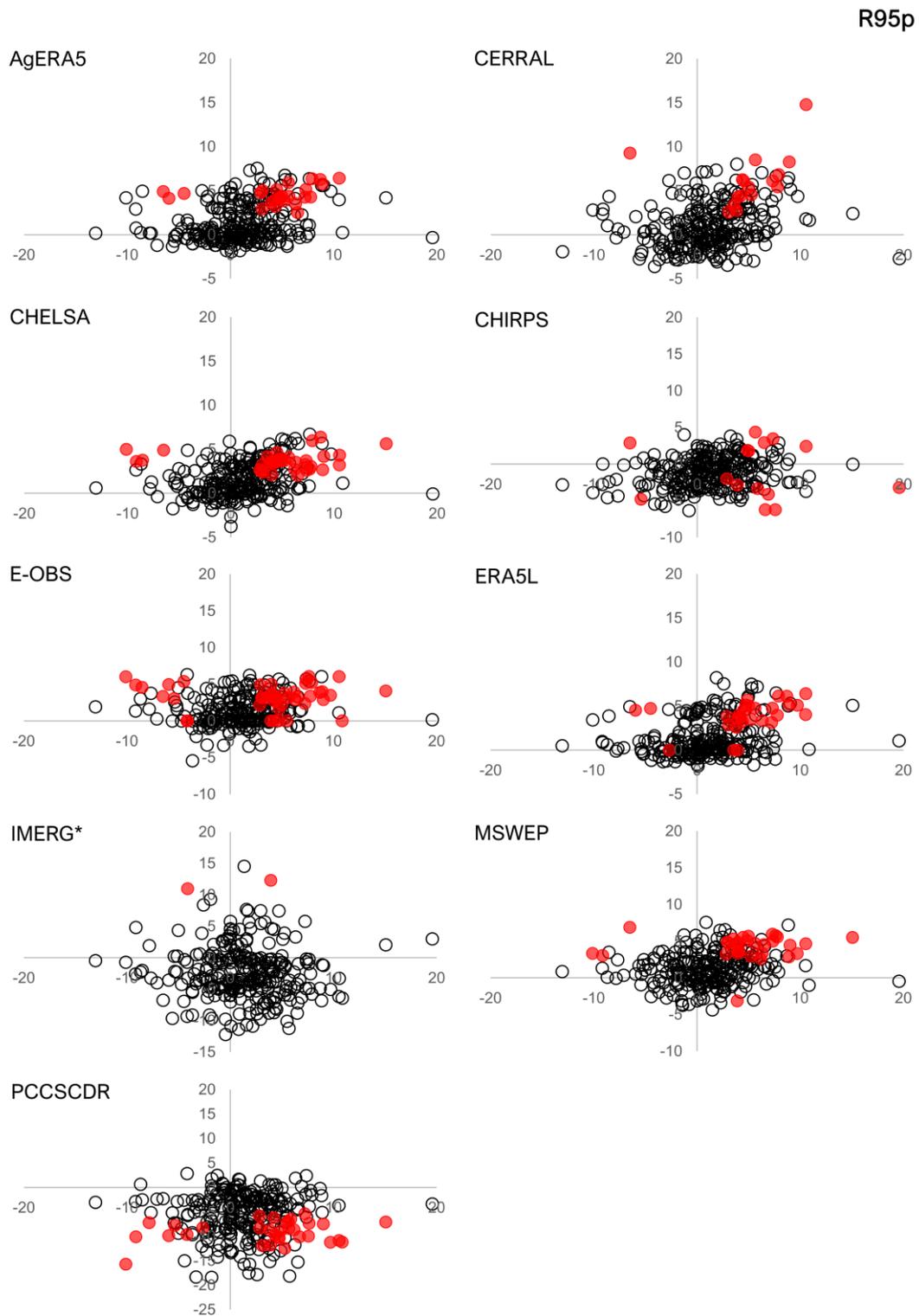


Figure 4.19. Scatter plots of the jointly statistically significant slope values (red points) and the remaining slope values (black points), regardless of significance, based on the statistical analysis of the R95p extreme index. Note that the range and size of the vertical axis is not identical among all datasets.

5. Conclusions

All datasets present inaccuracies when estimating daily rainfall and are not completely efficient in reproducing the intricacies of regional precipitation in Greece, both spatially and temporally. CHELSA and AgERA5 outperform the other datasets by systematically achieving acceptable scores throughout the six statistical metrics, although CHELSA underestimates daily rainfall by a median of 12.3%. PCCSCDR produces the poorest results in five out of six metrics (NSE = -1.11, KGE = 0.06, RMSE = 10.95 mm, MAE = 3.18 mm, and R = 0.17), but only slightly overestimates precipitation by -0.81%, therefore indicating errors of temporal representation of precipitation events. In higher elevation (700 – 1200 m) the competence of the datasets decreases, however CHELSA, CERRAL, and MSWEP capture orographic precipitation more precisely. The performance expectedly improves on a monthly timestep, but slight underestimation tendencies are still detected in most datasets. PCCSCDR underperforms in both temporal analyses.

Based on station observations and the implementation of the k-means clustering, there are distinct patterns of heavier precipitation, introduced mostly in northern and western regions of Greece. On average, Greece receives 749 mm of total rainfall. The West receives approximately 40% of the mean annual precipitation, whereas 30% is observed over the general northern Greek area. There is a strong seasonal variability in the West and South, while rainfall is minimal during summer months, especially in the central and southern parts of the country. The performance of the datasets, regarding specific regions, is limited, but AgERA5 provides better estimations for the hydrological district of Central Greece and CHELSA for West, Central Macedonia, and Thrace.

The analysis of the ETCCDI indices, suggests an average increase of 114.4 mm in mean annual rainfall and 33.1 mm in extreme rainfall per decade, during the 32 years of the assessment, as well as subareas of precipitation extremity in the western and central regions. The products fail to accurately describe the occurrence of extreme precipitation by producing poor correlation between the measured and estimated datasets, as well as pronounced biases.

Many studies suggest that there has been a noticeable shift across the Mediterranean countries regarding the increase of temperature and reduction of rainfall, stating that the Mediterranean region is undergoing a drier transformation, accompanied by prolonged droughts and severe storms (Stockhecke et al., 2016; Drobinski et al., 2018; Hertig & Trambly, 2017; Caloiero et al., 2018; Trambly et al., 2020). However, recent research mentions of an increase in the total amount of annual precipitation in the Northern Mediterranean, through short-term, high intensity rainfalls (Trambly & Somot, 2018; Ribes et al., 2019). This rate of atmospheric water exchange is considered a flood triggering parameter, does not contribute to soil water renewal, and has been proven disastrous in numerous occasions for Greece (Diakakis, 2017).

Therefore, a gridded precipitation dataset of a sub-daily or daily resolution, which allows for specific applications pertaining to risk mitigation and management, such as rapid analysis of extreme events or drought evolution monitoring, is beneficial to these regions. The suitable application of each product depends on its capabilities and the frequency with

which it is renewed. Six out of the nine datasets under evaluation use the method of reanalysis, which may incorporate several sources of data as input (e.g. gauges and/or satellite imaging), that undergo a process of refinement before the final product. This particular approach is used for the ERA5L and MSWEP reanalysis datasets, which, in addition to the timely (near real-time) data outputs, renders them appropriate for operational climate analysis, systematic monitoring, assessment, and forecasting of weather and climatic conditions. Regions with insufficient gauging stations, may benefit from satellite products, such as IMERG or PCCSCDR, which do not require ground-based data as input and can provide information on remote areas, at a high spatial resolution and half-hourly and 3-hourly timesteps, respectively. CHELSA has proven suitable for a diversity of applications, but the confined temporal coverage of the dataset is restrictive for the potential temporal extent of the analyses.

The evaluation and assessment of the gridded products, using an updated dataset of rainfall observations, may provide greater insights into the behaviour of precipitation and further contribute to the interpretation of long-term climate trends and their variability. A suggestion is the utilisation of the latest NOA Automatic Network (NOAAN) gauge-data collection, provided and managed by the National Observatory of Athens (NOA), which incorporates contemporary, real-time, measurements of climatic parameters, but is of limited temporal coverage (less than 20 years) (Lagouvardos et al., 2017).

6. Literature

- Accadia, C., Mariani, S., Casaioli, M., Lavagnini, A., & Speranza, A. (2003). Sensitivity of Precipitation Forecast Skill Scores to Bilinear Interpolation and a Simple Nearest-Neighbor Average Method on High-Resolution Verification Grids. *Weather and Forecasting*, *18*(5), 918–932. doi: 10.1175/1520-0434(2003)018<0918:SOPFSS>2.0.CO;2
- Alexandridis, V., Stefanidis, S., & Dafis, S. (2023). Evaluation of ERA5 and ERA5-Land Reanalysis Precipitation Data with Rain Gauge Observations in Greece. *16th International Conference on Meteorology, Climatology and Atmospheric Physics—COMECAP 2023*, 104. MDPI. doi: 10.3390/environsciproc2023026104
- Ballantyne, C. K. (1983). PRECIPITATION GRADIENTS IN WESTER ROSS, NORTH-WEST SCOTLAND. *Weather*, *38*(12), 379–387. doi: 10.1002/j.1477-8696.1983.tb04832.x
- Baltas, E. A. (2008). Climatic Conditions and Availability of Water Resources in Greece. *International Journal of Water Resources Development*, *24*(4), 635–649. doi: 10.1080/07900620802230129
- Barrett, A. P., Stroeve, J. C., & Serreze, M. C. (2020). Arctic Ocean Precipitation From Atmospheric Reanalyses and Comparisons With North Pole Drifting Station Records. *Journal of Geophysical Research: Oceans*, *125*(1), e2019JC015415. doi: 10.1029/2019JC015415
- Barros, A. P., & Lettenmaier, D. P. (1994). Dynamic modeling of orographically induced precipitation. *Reviews of Geophysics*, *32*(3), 265–284. doi: 10.1029/94RG00625
- Basist, A., Bell, G. D., & Meentemeyer, V. (1994). Statistical Relationships between Topography and Precipitation Patterns. *Journal of Climate*, *7*(9), 1305–1315. doi: 10.1175/1520-0442(1994)007<1305:SRBTAP>2.0.CO;2
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., ... Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. doi: 10.1175/BAMS-D-17-0138.1
- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2020). Publisher Correction: Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *7*(1), 274. doi: 10.1038/s41597-020-00616-w

- Bengtsson, L., Hagemann, S., & Hodges, K. I. (2004). Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research: Atmospheres*, *109*(D11), 2004JD004536. doi: 10.1029/2004JD004536
- Boogaard, H., Schubert, J., De Wit, A., Lazebnik, J., Hutjes, R., & Van der Grijn, G. (2020). *Agrometeorological indicators from 1979 up to 2019 derived from reanalysis* [Data set]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS): ECMWF. doi: 10.24381/CDS.6C68C9BB
- Bouizrou, I., Bouadila, A., Aqnouy, M., & Gourfi, A. (2023). Assessment of remotely sensed precipitation products for climatic and hydrological studies in arid to semi-arid data-scarce region, central-western Morocco. *Remote Sensing Applications: Society and Environment*, *30*, 100976. doi: 10.1016/j.rsase.2023.100976
- Briggs, P. R., & Cogley, J. G. (1996). Topographic Bias in Mesoscale Precipitation Networks. *Journal of Climate*, *9*(1), 205–218. doi: 10.1175/1520-0442(1996)009<0205:TBIMPN>2.0.CO;2
- Caloiero, T., Veltri, S., Caloiero, P., & Frustaci, F. (2018). Drought Analysis in Europe and in the Mediterranean Basin Using the Standardized Precipitation Index. *Water*, *10*(8), 1043. doi: 10.3390/w10081043
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. doi: 10.5194/gmd-7-1247-2014
- Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J. M., & Jones, P. D. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, *123*(17), 9391–9409. doi: 10.1029/2017JD028200
- Diakakis, M. (2017). Flood seasonality in Greece and its comparison to seasonal distribution of flooding in selected areas across southern Europe. *Journal of Flood Risk Management*, *10*(1), 30–41. doi: 10.1111/jfr3.12139
- Donat, M. G., Sillmann, J., Wild, S., Alexander, L. V., Lippmann, T., & Zwiers, F. W. (2014). Consistency of Temperature and Precipitation Extremes across Various Global Gridded In Situ and Reanalysis Datasets. *Journal of Climate*, *27*(13), 5019–5035. doi: 10.1175/JCLI-D-13-00405.1
- Drobinski, P., Silva, N. D., Panthou, G., Bastin, S., Muller, C., Ahrens, B., ... Torma, C. Z. (2018). Scaling precipitation extremes with temperature in the Mediterranean: Past climate assessment and projection in anthropogenic scenarios. *Climate Dynamics*, *51*(3), 1237–1257. doi: 10.1007/s00382-016-3083-x

- Duc, L., & Sawada, Y. (2023). A signal-processing-based interpretation of the Nash–Sutcliffe efficiency. *Hydrology and Earth System Sciences*, *27*(9), 1827–1839. doi: 10.5194/hess-27-1827-2023
- Feidas, H. (2010). Validation of satellite rainfall products over Greece. *Theoretical and Applied Climatology*, *99*(1–2), 193–216. doi: 10.1007/s00704-009-0135-8
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. doi: 10.1038/sdata.2015.66
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, *4*(2), 135–143. doi: 10.1061/(ASCE)1084-0699(1999)4:2(135)
- Hamed, K. H. (2008). Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, *349*(3–4), 350–363. doi: 10.1016/j.jhydrol.2007.11.009
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803
- Hertig, E., & Trambly, Y. (2017). Regional downscaling of Mediterranean droughts under past and future climatic conditions. *Global and Planetary Change*, *151*, 36–48. doi: 10.1016/j.gloplacha.2016.10.015
- Hofstra, N., Haylock, M., New, M., & Jones, P. D. (2009). Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *Journal of Geophysical Research: Atmospheres*, *114*(D21), 2009JD011799. doi: 10.1029/2009JD011799
- Hong, Y., Hsu, K.-L., Sorooshian, S., & Gao, X. (2004). Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *Journal of Applied Meteorology*, *43*(12), 1834–1853. doi: 10.1175/JAM2173.1
- Houze, R. A. (2012). Orographic effects on precipitating clouds. *Reviews of Geophysics*, *50*(1), 2011RG000365. doi: 10.1029/2011RG000365
- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., ... Xie, P. (2020). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow,

- K. Nakamura, & F. J. Turk (Eds.), *Satellite Precipitation Measurement* (pp. 343–353). Cham: Springer International Publishing. doi: 10.1007/978-3-030-24568-9_19
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., ... Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, 8(1), 38–55. doi: 10.1175/JHM560.1
- Karger, D. N., Lange, S., Hari, C., Reyer, C. P. O., & Zimmermann, N. E. (2022). *CHELSA-W5E5 v1.0: W5E5 v1.0 downscaled with CHELSA v2.0* (Version 1.0.2) [Data set]. ISIMIP Repository. doi: 10.48364/ISIMIP.836809.2
- Karl, T. R., Nicholls, N., & Ghazi, A. (1999). CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary. In T. R. Karl, N. Nicholls, & A. Ghazi (Eds.), *Weather and Climate Extremes* (pp. 3–7). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-015-9265-9_2
- Kazamias, A.-P., Sapountzis, M., & Lagouvardos, K. (2022). Evaluation of GPM-IMERG rainfall estimates at multiple temporal and spatial scales over Greece. *Atmospheric Research*, 269, 106014. doi: 10.1016/j.atmosres.2021.106014
- Kendall, M. G. (1948). *Rank correlation methods*. Griffin.
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). *Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores* [Preprint]. Catchment hydrology/Modelling approaches. doi: 10.5194/hess-2019-327
- Lagouvardos, K., Kotroni, V., Bezes, A., Koletsis, I., Kopania, T., Lykoudis, S., ... Vougioukas, S. (2017). The automatic weather stations NOANN network of the National Observatory of Athens: Operation and database. *Geoscience Data Journal*, 4(1), 4–16. doi: 10.1002/gdj3.44
- Levizzani, V., Bauer, P., & Turk, F. J. (Eds.). (2007). *Measuring Precipitation From Space: EURAINSAT and the Future*. Dordrecht: Springer Netherlands. doi: 10.1007/978-1-4020-5835-6
- Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., & Turk, F. J. (Eds.). (2020). *Satellite Precipitation Measurement: Volume 2*. Cham: Springer International Publishing. doi: 10.1007/978-3-030-35798-6
- Ma, L., Zhang, T., Frauenfeld, O. W., Ye, B., Yang, D., & Qin, D. (2009). Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *Journal of Geophysical Research*, 114(D9), D09105. doi: 10.1029/2008JD011178

- Mahmoud, M. T., Al-Zahrani, M. A., & Sharif, H. O. (2018). Assessment of global precipitation measurement satellite products over Saudi Arabia. *Journal of Hydrology*, *559*, 1–12. doi: 10.1016/j.jhydrol.2018.02.015
- Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245. doi: 10.2307/1907187
- Markonis, Y., Batelis, S. C., Dimakos, Y., Moschou, E., & Koutsoyiannis, D. (2017). Temporal and spatial variability of rainfall over Greece. *Theoretical and Applied Climatology*, *130*(1–2), 217–232. doi: 10.1007/s00704-016-1878-7
- Marra, F., Armon, M., Borga, M., & Morin, E. (2021). Orographic Effect on Extreme Precipitation Statistics Peaks at Hourly Time Scales. *Geophysical Research Letters*, *48*(5), e2020GL091498. doi: 10.1029/2020GL091498
- Mavromatis, T., & Voulanas, D. (2021). Evaluating ERA-Interim, Agri4Cast, and E-OBS gridded products in reproducing spatiotemporal characteristics of precipitation and drought over a data poor region: The Case of Greece. *International Journal of Climatology*, *41*(3), 2118–2136. doi: 10.1002/joc.6950
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., ... Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. doi: 10.5194/essd-13-4349-2021
- Najmi, A., Igmoullan, B., Namous, M., El Bouazzaoui, I., Brahim, Y. A., El Khalki, E. M., & Saidi, M. E. M. (2023). Evaluation of PERSIANN-CCS-CDR, ERA5, and SM2RAIN-ASCAT rainfall products for rainfall and drought assessment in a semi-arid watershed, Morocco. *Journal of Water and Climate Change*, *14*(5), 1569–1584. doi: 10.2166/wcc.2023.461
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. doi: 10.1016/0022-1694(70)90255-6
- Nastos, P. T., Kapsomenakis, J., & Philandras, K. M. (2016). Evaluation of the TRMM 3B43 gridded precipitation estimates over Greece. *Atmospheric Research*, *169*, 497–514. doi: 10.1016/j.atmosres.2015.08.008
- Olmo, M. E., & Bettolli, M. L. (2021). Extreme daily precipitation in southern South America: Statistical characterization and circulation types using observational datasets and regional climate models. *Climate Dynamics*, *57*(3–4), 895–916. doi: 10.1007/s00382-021-05748-2

- Pearson, K. (1895). Note on Regression and Inheritance in the Case of Two Parents. *Proceedings of the Royal Society of London Series I*, 58, 240–242.
- Ribes, A., Thao, S., Vautard, R., Dubuisson, B., Somot, S., Colin, J., ... Soubeyroux, J.-M. (2019). Observed increase in extreme daily rainfall in the French Mediterranean. *Climate Dynamics*, 52(1–2), 1095–1114. doi: 10.1007/s00382-018-4179-2
- Sadeghi, M., Nguyen, P., Naeini, M. R., Hsu, K., Braithwaite, D., & Sorooshian, S. (2021). PERSIANN-CCS-CDR, a 3-hourly 0.04° global precipitation climate data record for heavy precipitation studies. *Scientific Data*, 8(1), 157. doi: 10.1038/s41597-021-00940-9
- Schär, C., Ban, N., Fischer, E. M., Rajczak, J., Schmidli, J., Frei, C., ... Zwiers, F. W. (2016). Percentile indices for assessing changes in heavy precipitation events. *Climatic Change*, 137(1–2), 201–216. doi: 10.1007/s10584-016-1669-2
- Schulzweida, U. (2023). *CDO User Guide*. doi: 10.5281/ZENODO.10020800
- Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. doi: 10.1080/01621459.1968.10480934
- Sindosi, O. A., Bartzokas, A., Kotroni, V., & Lagouvardos, K. (2015). Influence of orography on precipitation amount and distribution in NW Greece; A case study. *Atmospheric Research*, 152, 105–122. doi: 10.1016/j.atmosres.2014.06.013
- So, D., & Shin, D. (2018). Classification of precipitating clouds using satellite infrared observations and its implications for rainfall estimation. *Quarterly Journal of the Royal Meteorological Society*, 144(S1), 133–144. doi: 10.1002/qj.3288
- Stockhecke, M., Timmermann, A., Kipfer, R., Haug, G. H., Kwiecien, O., Friedrich, T., ... Anselmetti, F. S. (2016). Millennial to orbital-scale variations of drought intensity in the Eastern Mediterranean. *Quaternary Science Reviews*, 133, 77–95. doi: 10.1016/j.quascirev.2015.12.016
- Tabios, G. Q., & Salas, J. D. (1985). A COMPARATIVE ANALYSIS OF TECHNIQUES FOR SPATIAL INTERPOLATION OF PRECIPITATION ¹. *JAWRA Journal of the American Water Resources Association*, 21(3), 365–380. doi: 10.1111/j.1752-1688.1985.tb00147.x
- Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., ... De Castro, M. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, 104–105, 70–97. doi: 10.1016/j.atmosres.2011.10.021

- Tramblay, Y., El Adlouni, S., & Servat, E. (2013). Trends and variability in extreme precipitation indices over Maghreb countries. *Natural Hazards and Earth System Sciences*, *13*(12), 3235–3248. doi: 10.5194/nhess-13-3235-2013
- Tramblay, Yves, Koutroulis, A., Samaniego, L., Vicente-Serrano, S. M., Volaire, F., Boone, A., ... Polcher, J. (2020). Challenges for drought assessment in the Mediterranean region under future climate scenarios. *Earth-Science Reviews*, *210*, 103348. doi: 10.1016/j.earscirev.2020.103348
- Tramblay, Yves, & Somot, S. (2018). Future evolution of extreme precipitation in the Mediterranean. *Climatic Change*, *151*(2), 289–302. doi: 10.1007/s10584-018-2300-5
- Tzanis, C. G., Koutsogiannis, I., Philippopoulos, K., & Deligiorgi, D. (2019). Recent climate trends over Greece. *Atmospheric Research*, *230*, 104623. doi: 10.1016/j.atmosres.2019.104623
- Verrelle, A., Glinton, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Ridal, M., ... Soci, C. (2022). *CERRA-Land sub-daily regional reanalysis data for Europe from 1984 to present* [Data set]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS): ECMWF. doi: 10.24381/CDS.A7F3CD0B
- Vicente-Serrano, S. M., Beguería, S., López-Moreno, J. I., García-Vera, M. A., & Stepanek, P. (2009). A complete daily precipitation database for northeast Spain: Reconstruction, quality control, and homogeneity. *International Journal of Climatology*, *30*(8), 1146–1163. doi: 10.1002/joc.1850
- Vicente-Serrano, S., Saz-Sánchez, M., & Cuadrat, J. (2003). Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): Application to annual precipitation and temperature. *Climate Research*, *24*, 161–180. doi: 10.3354/cr024161
- Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., ... Woods, R. (2011). Climate change and mountain water resources: Overview and recommendations for research, management and policy. *Hydrology and Earth System Sciences*, *15*(2), 471–504. doi: 10.5194/hess-15-471-2011
- Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global Increasing Trends in Annual Maximum Daily Precipitation. *Journal of Climate*, *26*(11), 3904–3918. doi: 10.1175/JCLI-D-12-00502.1

7. Appendix

Table A1. Median values of the statistical metrics of the nine datasets and number of stations included in the 8 clusters.

Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
1	AgERA5	0.27	0.43	1.86	5.07	1.49	0.55
No. of stations	CERRAL	-0.09	0.34	-16.42	6.21	1.82	0.40
38	CHELSA	0.15	0.37	-1.97	5.27	1.63	0.46
	CHIRPS	-0.24	0.26	-8.15	6.56	2.17	0.29
	E-OBS	0.15	0.28	18.75	5.40	1.58	0.43
	ERA5L	-0.05	0.25	3.32	6.07	1.74	0.35
	IMERG*	0.00	0.39	-9.10	5.95	1.73	0.41
	MSWEP	-1.59	0.00	-28.77	9.55	2.56	0.15
	PCCSCDR	-0.23	0.24	-20.96	6.84	2.08	0.39
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
2	AgERA5	-0.07	0.33	-41.12	5.19	1.89	0.45
No. of stations	CERRAL	-0.32	0.31	-20.25	5.52	1.88	0.38
36	CHELSA	0.13	0.39	-3.86	4.61	1.59	0.47
	CHIRPS	-0.39	0.20	-16.94	5.99	2.08	0.29
	E-OBS	0.01	0.26	13.02	4.89	1.64	0.35
	ERA5L	-0.47	0.17	-41.34	5.92	2.17	0.30
	IMERG*	-0.01	0.36	-14.77	4.96	1.71	0.43
	MSWEP	-2.78	-0.37	-53.99	9.49	2.72	0.14
	PCCSCDR	-0.57	0.23	-29.84	6.14	2.06	0.37
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
3	AgERA5	0.23	0.29	26.41	10.67	4.18	0.51
No. of stations	CERRAL	0.10	0.25	28.01	11.47	4.44	0.40
24	CHELSA	0.25	0.21	38.81	10.39	3.89	0.52
	CHIRPS	-0.10	0.20	31.14	12.74	4.86	0.33
	E-OBS	0.12	-0.03	55.51	10.97	3.92	0.41
	ERA5L	0.06	0.21	21.24	11.67	4.73	0.35
	IMERG*	0.17	0.18	37.10	10.82	4.11	0.45
	MSWEP	-0.50	0.11	35.50	14.87	5.07	0.19
	PCCSCDR	0.09	0.28	27.21	11.47	4.50	0.42
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
4	AgERA5	0.09	0.35	-5.39	6.86	2.58	0.43
No. of stations	CERRAL	0.13	0.47	-1.82	6.81	2.32	0.52
66	CHELSA	0.16	0.33	15.61	6.59	2.34	0.45
	CHIRPS	-0.44	0.24	-2.98	8.54	3.02	0.27
	E-OBS	0.14	0.27	25.11	6.59	2.19	0.43
	ERA5L	0.09	0.34	-5.46	6.86	2.62	0.42
	IMERG*	0.05	0.38	-0.52	7.05	2.48	0.42

	MSWEP	-1.26	0.08	1.15	10.99	3.29	0.16
	PCCSCDR	-0.29	0.26	-4.96	8.08	2.85	0.32
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
5	AgERA5	0.17	0.41	-25.47	6.02	1.98	0.51
No. of stations	CERRAL	-0.23	0.21	-18.22	6.79	2.43	0.30
23	CHELSA	0.30	0.45	-5.18	5.39	1.77	0.57
	CHIRPS	-0.08	0.33	-5.39	6.68	2.20	0.37
	E-OBS	0.12	0.20	26.53	5.96	1.80	0.38
	ERA5L	-0.28	0.16	-27.63	7.08	2.56	0.20
	IMERG*	0.06	0.41	-6.50	6.22	1.80	0.45
	MSWEP	-0.94	0.15	-21.02	9.34	2.59	0.22
	PCCSCDR	-0.11	0.37	-35.45	7.09	2.22	0.51
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
6	AgERA5	0.23	0.39	4.91	7.99	3.09	0.53
No. of stations	CERRAL	0.00	0.33	10.58	8.91	3.40	0.40
56	CHELSA	0.29	0.32	28.28	7.68	2.83	0.55
	CHIRPS	-0.25	0.29	11.22	10.24	3.75	0.35
	E-OBS	0.14	0.14	40.83	8.21	2.87	0.42
	ERA5L	-0.07	0.24	2.92	9.18	3.64	0.33
	IMERG*	0.20	0.36	20.75	8.12	3.00	0.48
	MSWEP	-0.86	0.11	16.83	12.61	4.01	0.18
	PCCSCDR	0.00	0.39	12.17	9.07	3.23	0.50
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
7	AgERA5	0.07	0.33	-19.69	5.96	2.35	0.44
No. of stations	CERRAL	-0.34	0.24	-14.45	6.51	2.55	0.29
35	CHELSA	0.14	0.34	10.58	5.47	1.98	0.47
	CHIRPS	-0.24	0.26	-2.91	6.76	2.50	0.29
	E-OBS	0.06	0.16	25.47	5.76	2.00	0.36
	ERA5L	-0.30	0.14	-17.44	6.54	2.80	0.23
	IMERG*	0.00	0.33	3.50	6.07	2.10	0.40
	MSWEP	-1.47	0.05	-17.22	9.92	3.02	0.17
	PCCSCDR	-0.17	0.32	-6.77	7.08	2.42	0.37
Cluster ID	Name	NSE	KGE	PBIAS	RMSE	MAE	R
8	AgERA5	0.28	0.16	39.01	8.01	2.25	0.56
No. of stations	CERRAL	0.13	0.26	25.66	8.57	2.52	0.43
22	CHELSA	0.32	0.36	22.65	7.82	2.38	0.58
	CHIRPS	-0.33	0.18	26.51	11.05	3.23	0.26
	E-OBS	0.26	0.17	46.63	7.75	2.28	0.52
	ERA5L	0.08	0.05	42.73	8.56	2.57	0.34
	IMERG*	0.21	0.34	31.56	8.21	2.39	0.52
	MSWEP	-0.64	0.06	37.19	11.86	3.31	0.15
	PCCSCDR	0.09	0.28	15.21	9.53	2.88	0.45

Table A2. Median correlation coefficient and percent bias of extreme indices of the nine datasets, in the 8 clusters, between 1985 – 2016. Disclosed in brackets are the corresponding values of each index, as calculated from the observations. The asterisks (**) denote insufficient data due to the pre-processing of E-OBS and implementation of the 70% fill rate of daily values.

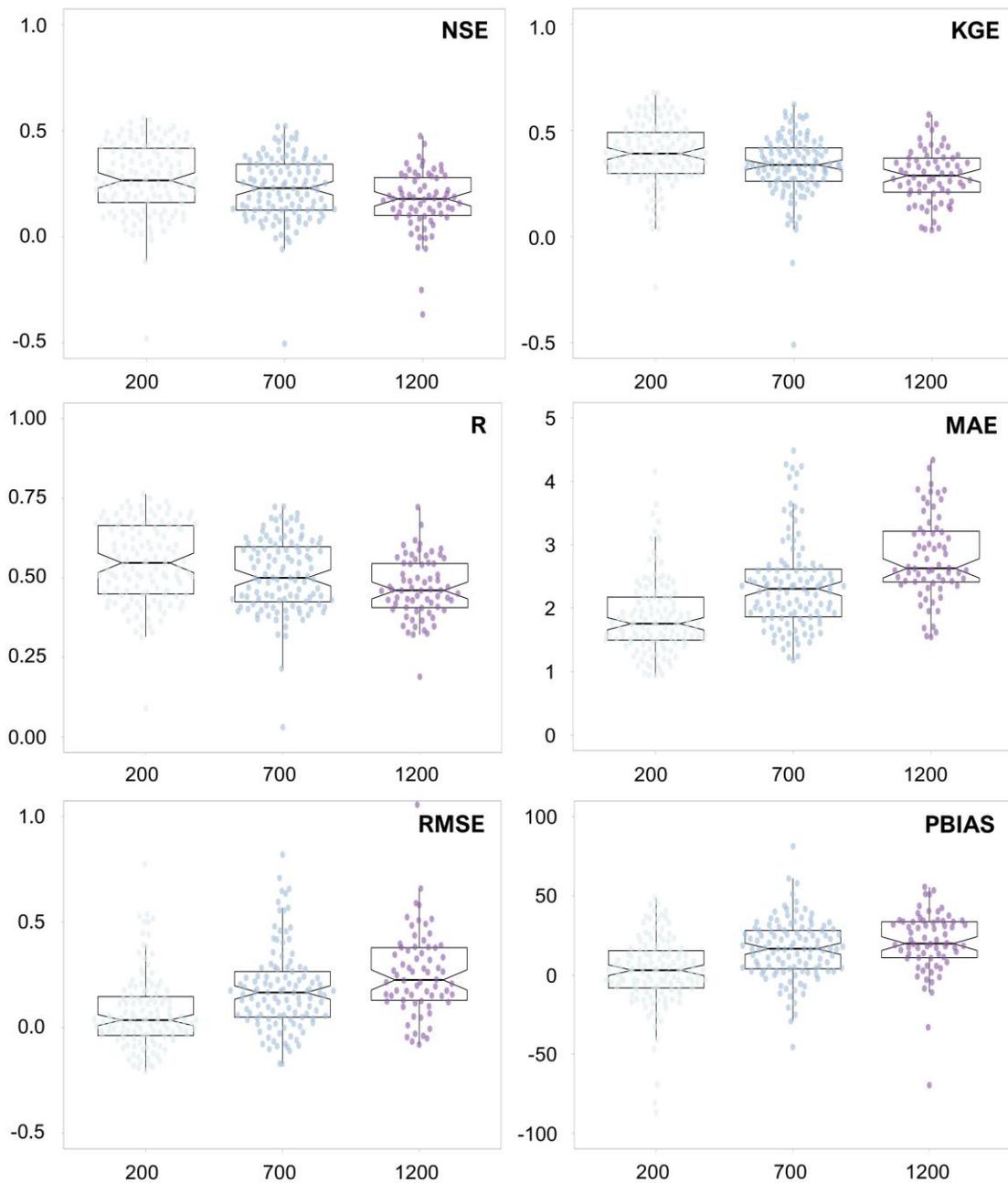
1		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [111.4 mm]	R	0.38	0.41	0.43	0.67	0.03	0.48	0.09	0.71	0.02
	PBIAS	-3.46	-31.17	-11.11	3.24	65.33	7.47	-79.83	-24.77	-80.84
rx1 [113.2 mm]	R	0.24	0.44	0.33	0.25	0.11	0.25	-0.20	0.23	-0.05
	PBIAS	20.71	-9.21	37.80	11.64	42.15	23.66	-32.70	-11.54	-181.97
rx5 [178.7 mm]	R	0.15	0.27	0.19	0.18	0.06	0.08	0.11	-0.08	-0.21
	PBIAS	23.17	6.58	35.24	17.70	50.23	28.38	-5.14	14.66	-67.10
SDII [9.9 mm/d]	R	0.43	0.48	0.39	0.45	0.38	0.40	0.32	0.32	0.36
	PBIAS	36.17	27.37	42.02	-18.70	38.78	39.57	22.76	36.83	-6.13
R10mm [15.2 days]	R	0.68	0.63	0.72	0.75	0.72	0.69	0.65	0.71	0.70
	PBIAS	-5.81	-5.02	17.10	-16.45	39.08	-2.34	-8.20	8.45	2.76
R20mm [6.5 days]	R	0.69	0.72	0.67	0.75	0.67	0.69	0.57	0.69	0.69
	PBIAS	22.64	15.75	49.75	-4.56	77.81	24.98	-1.71	35.55	5.13
2		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [99.3 mm]	R	0.30	0.27	0.16	0.26	-0.13	0.27	0.11	0.25	0.31
	PBIAS	-110.88	-100.99	-34.23	-38.17	8.74	105.44	-139.04	-71.62	-186.32
rx1 [95.2 mm]	R	0.16	0.19	0.38	0.31	0.13	0.04	0.13	0.00	0.04
	PBIAS	22.49	-21.43	34.71	0.40	36.75	24.07	-25.68	-8.82	-173.38
rx5 [162.0 mm]	R	0.54	0.62	0.65	0.66	0.26	0.48	0.35	-0.08	0.09
	PBIAS	18.71	-14.70	25.88	-0.77	45.50	21.69	-11.86	4.55	-88.34
SDII [8.1 mm/d]	R	0.35	0.53	0.60	0.61	0.65	0.48	0.33	0.43	0.50
	PBIAS	34.82	22.55	38.06	-32.68	35.37	36.72	19.23	33.16	-12.80
R10mm [14.3 days]	R	0.58	0.61	0.56	0.73	0.66	0.63	0.58	0.67	0.47
	PBIAS	-10.86	-15.99	11.02	-26.20	33.80	-9.58	-11.31	2.78	-1.36
R20mm [4.8 days]	R	0.34	0.60	0.55	0.74	0.60	0.47	0.47	0.57	0.52
	PBIAS	23.29	1.87	43.46	-21.75	75.00	24.39	-3.51	29.16	0.54
3		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR

R95p [273.6 mm]	R	0.44	0.45	0.40	0.18	0.25	0.43	0.46	0.24	0.20
	PBIAS	7.19	-2.46	31.02	29.39	60.82	4.37	-7.25	25.18	-3.26
rx1 [164.7 mm]	R	0.11	-0.09	0.39	0.34	0.30	-0.02	-0.14	0.34	0.12
	PBIAS	13.11	-19.85	40.96	13.83	37.36	19.94	-11.73	-1.43	-157.73
rx5 [345.2 mm]	R	0.25	0.75	0.66	0.54	0.26	0.31	0.13	0.51	0.26
	PBIAS	17.76	-3.33	36.01	17.32	46.81	23.05	3.81	15.41	-66.57
SDII [19.3 mm/d]	R	0.60	0.80	0.69	0.56	0.43	0.53	0.48	0.47	0.24
	PBIAS	39.84	28.78	46.19	-10.34	35.99	41.78	29.95	41.14	-2.03
R10mm [44.4 days]	R	0.75	0.89	0.90	0.86	0.85	0.80	0.79	0.82	0.61
	PBIAS	-16.92	-20.11	15.59	-23.84	31.76	-13.82	-10.72	2.14	-5.20
R20mm [25.7 days]	R	0.82	0.90	0.88	0.85	0.73	0.77	0.78	0.76	0.58
	PBIAS	17.89	-3.44	48.45	-7.54	75.77	18.39	0.04	31.34	-2.14
4		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [160.5 mm]	R	-0.07	-0.12	0.04	-0.13	-0.04	-0.13	-0.10	-0.22	0.03
	PBIAS	-18.22	-24.11	10.21	6.12	58.02	-13.59	-50.24	-15.18	-50.28
rx1 [149.0 mm]	R	0.34	0.22	0.39	0.23	0.24	0.29	0.20	0.14	0.27
	PBIAS	26.03	-3.58	41.23	10.86	43.21	26.09	-7.79	0.43	-166.66
rx5 [231.1 mm]	R	0.35	0.42	0.48	0.43	0.32	0.42	0.25	0.37	0.31
	PBIAS	20.98	1.31	33.09	14.61	50.01	21.37	5.84	11.73	-73.55
SDII [11.8 mm/d]	R	0.45	0.54	0.60	0.43	0.43	0.45	0.45	0.48	0.41
	PBIAS	41.96	32.45	45.26	-15.99	41.86	42.15	30.14	40.99	1.40
R10mm [24.2 days]	R	0.52	0.68	0.73	0.61	0.75	0.58	0.69	0.74	0.25
	PBIAS	-0.86	-2.74	19.91	-3.76	40.05	-3.00	-1.11	10.74	10.03
R20mm [10.2 days]	R	0.47	0.66	0.70	0.60	0.65	0.51	0.57	0.71	0.37
	PBIAS	31.17	18.17	50.32	-6.72	78.45	30.16	8.86	36.78	12.29
5		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [123.7 mm]	R	0.38	0.22	0.29	0.17	0.05	0.41	-0.05	0.34	0.45
	PBIAS	-27.77	-40.56	1.12	14.15	48.10	-19.15	-90.59	-21.74	-61.22
rx1 [110.0 mm]	R	0.34	0.20	0.42	0.32	0.42	0.11	0.21	-0.04	0.50
	PBIAS	36.85	8.47	49.08	23.74	55.61	39.07	-11.37	-1.86	-129.77

rx5 [182.5 mm]	R	0.57	0.26	0.69	0.70	0.52	0.54	0.53	0.06	0.43
	PBIAS	34.29	14.63	41.39	20.29	61.86	38.15	9.31	22.27	-43.09
SDII [12.4 mm/d]	R	0.60	0.69	0.66	0.66	0.75	0.54	0.73	0.58	0.48
	PBIAS	48.76	41.44	50.96	-1.13	55.73	51.53	35.07	46.76	14.63
R10mm [18.7 days]	R	0.83	0.71	0.82	0.84	0.84	0.84	0.84	0.88	0.70
	PBIAS	5.27	6.92	23.13	-3.04	50.40	9.15	4.18	18.15	17.44
R20mm [8.0 days]	R	0.70	0.73	0.76	0.78	0.76	0.72	0.82	0.78	0.68
	PBIAS	34.25	28.14	50.98	7.42	82.22	37.48	14.15	41.88	24.13
6		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [203.7 mm]	R	-0.27	0.24	-0.09	-0.22	-0.36	-0.34	-0.30	-0.17	-0.15
	PBIAS	-18.18	-20.13	12.99	7.25	52.59	-13.15	-49.00	-4.50	-38.39
rx1 [132.5 mm]	R	0.29	0.37	0.46	0.32	0.17	0.32	0.34	-0.01	0.07
	PBIAS	28.48	-8.70	43.65	12.24	44.58	27.95	-15.05	11.51	-159.73
rx5 [276.0 mm]	R	0.37	0.61	0.45	0.45	0.21	0.39	0.50	0.12	0.22
	PBIAS	25.66	7.48	39.13	19.69	52.85	25.56	-0.74	23.68	-61.01
SDII [15.4 mm/d]	R	0.58	0.76	0.71	0.63	0.62	0.63	0.48	0.60	0.50
	PBIAS	40.27	32.29	44.81	-15.92	40.97	39.62	25.80	40.01	-1.42
R10mm [33.9 days]	R	0.78	0.86	0.86	0.81	0.81	0.79	0.81	0.84	0.52
	PBIAS	-0.81	3.46	23.49	-8.36	42.42	-1.46	-0.18	16.03	9.51
R20mm [16.5 days]	R	0.66	0.84	0.79	0.80	0.67	0.70	0.71	0.75	0.55
	PBIAS	26.84	22.26	53.00	-1.01	77.85	26.77	5.98	39.66	10.25
7		AgERA5	CERRAL	CHELSEA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [125.8 mm]	R	0.16	0.20	0.00	-0.06	-0.16	0.16	0.00	-0.19	0.07
	PBIAS	-42.27	-50.00	2.40	3.19	36.38	-43.69	-70.16	-17.18	-83.85
rx1 [91.0 mm]	R	0.19	0.25	0.65	0.49	0.66	0.09	-0.01	0.32	0.06
	PBIAS	44.65	9.10	46.42	14.35	83.19	47.61	-2.25	13.73	-143.81
rx5 [158.8 mm]	R	0.07	0.62	0.74	0.49	0.53	-0.12	-0.02	0.57	0.24
	PBIAS	33.92	11.13	38.62	24.42	82.73	43.18	-0.48	20.64	-76.26
SDII [10.4 mm/d]	R	-0.23	0.41	0.61	0.65	0.87	-0.24	-0.11	0.48	0.30
	PBIAS	44.81	29.19	39.36	-15.71	79.02	50.04	28.72	34.53	0.58

R10mm [19.8 days]	R	0.26	0.70	0.81	0.42	0.91	0.25	0.32	0.51	-0.11
	PBIAS	10.51	5.31	3.34	0.53	76.20	18.81	-7.75	6.91	14.19
R20mm [7.9 days]	R	0.11	0.70	0.76	0.63	0.74	0.07	0.19	0.47	-0.06
	PBIAS	35.23	21.29	37.05	1.17	89.40	44.67	-3.13	27.71	15.55
8		AgERA5	CERRAL	CHELSA	CHIRPS	E-OBS	ERA5L	IMERG*	MSWEP	PCCSCDR
R95p [176.9 mm]	R	0.74	0.70	0.65	0.34	**	0.63	0.33	0.59	0.45
	PBIAS	29.59	-1.00	12.96	9.90	**	30.71	-36.54	-1.04	-9.92
rx1 [158.5 mm]	R	0.02	0.22	0.11	0.37	-0.55	0.27	-0.10	0.15	0.40
	PBIAS	8.99	-53.64	27.19	-12.22	67.56	34.19	-72.29	-32.51	-191.09
rx5 [288.8 mm]	R	0.15	0.39	0.35	0.44	-0.47	0.40	0.02	0.26	0.08
	PBIAS	-19.79	-59.55	2.21	-19.04	65.36	9.21	-83.55	-49.73	-160.57
SDII [16.9 mm/d]	R	0.88	0.91	0.76	0.63	0.71	0.96	0.72	0.77	0.91
	PBIAS	-0.22	-30.27	-3.99	-95.41	53.94	22.90	-47.48	-17.25	-78.17
R10mm [25.7 days]	R	0.87	0.88	0.70	0.81	0.92	0.87	0.84	0.90	0.68
	PBIAS	-65.87	-75.91	-75.55	-90.73	55.45	-28.27	-123.66	-66.20	-62.88
R20mm [14.1 days]	R	0.88	0.87	0.68	0.81	0.62	0.87	0.69	0.85	0.77
	PBIAS	-39.83	-66.34	-30.84	-106.82	77.79	-11.80	-173.28	-50.05	-85.80

CHELSEA



MEDIAN	NSE	KGE	R	MAE (mm)	RMSE (mm)	PBIAS
0 – 200m	0.27	0.23	0.18	1.75	5.36	2.99
200m – 700m	0.39	0.34	0.29	2.31	6.66	16.56
700m – 1200m	0.55	0.50	0.46	2.63	7.26	19.82

Figure A1. Distribution of statistical metrics of the CHELSA dataset in three altitude ranges and the corresponding median scores.