Mexico's High Resolution Climate Database (MexHiResClimDB): a new daily high-resolution gridded climate dataset for Mexico covering 1951–2020

Jaime J. Carrera-Hernández[1]

[1]Instituto de Geociencias, UNAM

Correspondence: J. J. Carrera-Hernández (jaime-carrera@geociencias.unam.mx)

Abstract. This work presents Mexico's High Resolution Climate Database (MexHiResClimDB), which is a newly developed gridded, high-resolution climate dataset comprised of daily, monthly and yearly precipitation and temperature (T_{min}, T_{max}, T_{avg}). This new database provides the largest temporal coverage of the aforementioned climate variables at the highest spatial resolution (20 arc sec, or 560 m on Mexico's CCL projection) when compared to the other currently available gridded datasets for Mexico and its development has allowed to ~~analyze~~ analyse the country's climate extremes for the 1951–2020 period. By comparing the spatial distribution of precipitation from the MexHiResClimDB with other gridded data (Daymet, L15, CHIRPS and PERSIANN CDR), it was found that the precipitation provided by this new dataset ~~is the only one that~~ adequately represents the spatial variation of extreme precipitation events, in particular for the precipitation that occurred during September 15–16 of 2013, caused by the presence of Tropical storm Manuel in the Pacific Ocean and Hurricane Ingrid (Cat 1) in the Gulf of Mexico. Using data from 61 days retrieved from Automated Weather Stations located throughout Mexico – and correspoding to the two months with the largest precipitation in Mexico – it was found that precipitation data from MexHiResClimDB has a MAE=8.7 mm, followed by L15 (9.5 mm), Daymet (10.1 mm) and CHIRPS (11.7 mm). For T_{min} and T_{max}, the lowest MAE was obtained with MexHiResClimDB (1.7°C and 1.8 °C, respectively), followed by Daymet (2.0 °C for both temperatures) and L15 (2.4°C and 2.5 °C). With this new database it was possible to summarize extreme events of precipitation and temperature in Mexico for the 1951–2020 period – a summary that was not available before: the wettest year was 1958, the wettest day 1970-09-26, and September of 2013 the wettest month. It was also found that eight out of the ten days with the highest T_{min} occurred in 2020, the two months with the highest T_{min} were July and August of 2020 and that the six years with the highest T_{min} were 2015–2020. When T_{max} was ~~analyzed~~analysed, it was found that the hottest day was 1998-06-15, while June of 1998 was the hottest month and 2020 the hottest year, and that the four hottest years occurred between 2011–2020. Nationwide (and considering 1961–1990 as the baseline period), T_{min}, T_{avg} and T_{max} have increased, with their anomalies drastically increasing in recent years and reaching values above 1.0 °C in 2020. At the same time, precipitation has also decreased in recent years – which combined with the increase in temperature will have severe impacts on water availability~~that need to be analyzed in detail, for example at the watershed level~~. This new database provides a tool to quantify – in detail – the spatio-temporal variability of climate throughout Mexico.
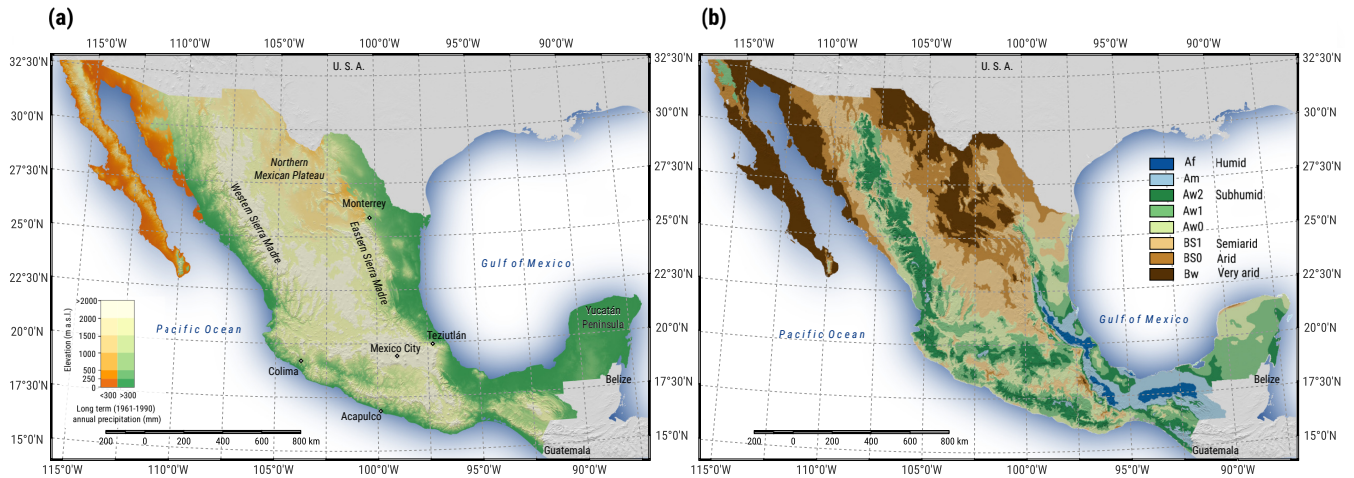
Figure 1. Mexico's (a) topography and (b) climates. The cities shown – except Mexico City, which is only shown as reference – have had catastrophic floods and landslides caused by to different Tropical Cyclones. Mexico is shown in Lambert Conformal Conical projection with shaded relief derived from the AW3D30 DSM and a cross-blended hypsometric color scale (Patterson and Jenny, 2011) to differentiate between arid and non-arid regions (the color scale varies according to the precipitation normal of 1961–1990); the distribution of climates in Mexico is adapted from García (2004)

2   Study area

Mexico – with a continental area of 1.96×10^6 km^2 – is surrounded on the west by the Pacific Ocean and on the East by the Gulf of Mexico, with an abrupt topography that varies from sea level up to ~~even~~ nearly 5,700 m a.s.l. (Fig. 1(a)). Due to the interaction between warm and cold air masses and Mexico's topography, climates that range from very arid to humid are found within it, with vegetation that varies from xerophyle shrubs and grasslands to tropical forests and even cloud forests (Fig. 1(b)). Mexico's main sierras – the Western and Eastern Sierras, which run nearly parallel to the Pacific and Gulf coasts, as shown in Fig. 1(a) – block moist air masses from the aforementioned coasts. Accordingly, the area enclosed by the ~~Western and Eastern~~ Mexico's Sierras – the Northern Mexican Plateau – has semiarid to very arid climates, while humid climates are found on the windward side of these Sierras (Fig. 1(b)). Furthermore, Tropical Cyclones that originate from both the Atlantic and the Eastern Pacific basin make landfall in Mexico (Farfán et al., 2014); in fact, on the second half of the ~~XXth~~ 20th century, a total of 65 hurricanes impacted Mexico's Pacific coast, while 27 impacted its eastern coast (Jauregui, 2003). Due to Mexico's geographic context, its precipitation is extremely variable, as on a given day during the rainy season, precipitation can vary from 0 to more than 300 ~~milimeters~~mm and the adequate representation of these extreme events is important in Mexico, in particular for those cities that depend on reservoirs that are filled by hurricanes and tropical storms or where flash-floods are a recurrent event.
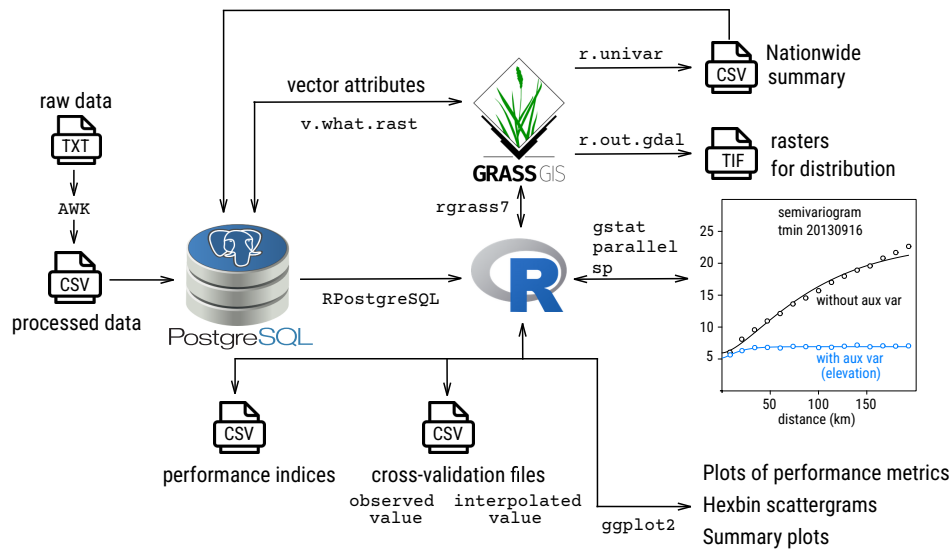
Figure 2. Flowchart describing how the relational database developed in this work can be accessed with R and the Open Source GRASS GIS. This flowchart also shows that cross-validation files were created for each variable on a daily basis in order to compute different metrics to evaluate the performance of the interpolation method applied in this work. The effect of considering elevation as an auxiliary variable to apply Kriging with External Drift is also shown on the T_{min} semivariogram for September 16th of 2013.

3 Methodology

70 The climate data used in the present study was downloaded from Mexico's meteorological service website (last date consulted: December 2024) on a station by station basis through a `bash` script, and further processed with a series of both `bash` `awk` and R scripts in order to generate a `PostgreSQL` relational database, as ~~exemplified in Carrera-Hernández and Gaskin (2008b)~~ proposed by Carrera-Hernández and Gaskin (2008b). This database can be accessed through both R and the Open Source GRASS GIS, which allows the use of data stored in GRASS to be used by different R libraries and to store the spatial data

75 created in R in GRASS's internal file structure, as detailed in Figure 2.

To develop the ~~aforementioned~~ relational database, the location of all weather stations was first verified (as some stations had wrong coordinates) and once the climate records were in PostgreSQL only those stations with more than ~~10 years~~ 80% of registered data ~~were selected; accordingly, not all available stations were used , and~~ for at least 10 continuous years were selected (wheras the L15 dataset (Livneh et al., 2015) used a 50-day threshold for stations in Mexico). Accordingly,

80 the number of stations ~~varied~~ used to undertake the interpolations across the 1951–2021 period varies, as shown in Fig. 3. ~~The aforementioned figure~~ This Figure shows that there are fewer records of temperature than precipitation (a situation also observed for North America by Tang et al. (2020)) and that the maximum number of records for both variables were registered in years 1982 and 1983 (with over 4000 daily records for precipitation and between 3000–3500 for temperature out of a total of 5467 stations from the downloaded raw data), and that years 1951 and 2021 exhibit the lowest number of records. However,
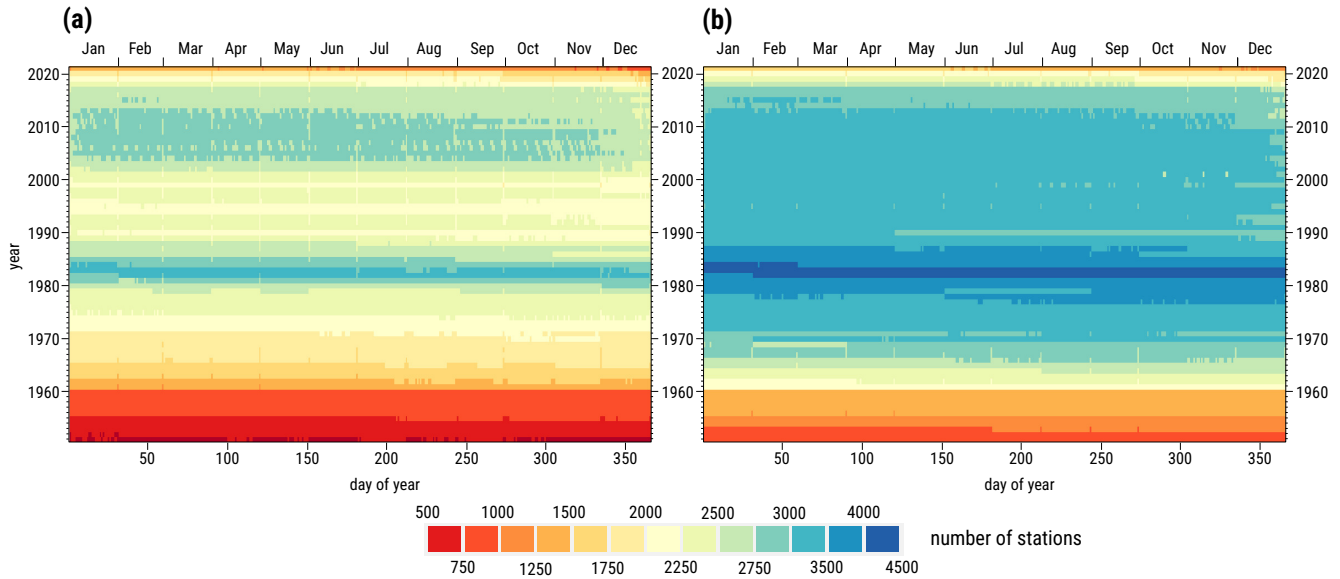
Figure 3. Number of weather stations used for daily interpolations of: (a) Temperature, (b) Precipitation. It can be seen that 1982 and 1983 were the years with the largest number of stations and that a decrease of records for both temperature and precipitation started in 2013.

85   it should be kept in mind that a larger amount of records does not represent better spatial coverage, as shown on Fig. 4, where it can be seen that by 2021 the spatial coverage of weather stations is limited outside Mexico's central region – which is why the MexHiResClimDB does not include data ~~for 2021. To avoid outliers on the dataset,~~ after year 2020. A basic Quality Control (QC) was applied to the generated database, with daily precipitation values above 600 mm, as well as temperature values below -30 °C or above 60 °C ~~were discarded~~being discarded, along with those days where $T_{min} > T_{max}$.

90   Although the removal of outliers considering neighbouring stations could have been done, this was not done due to the fact that precipitation in Mexico is highly variable within short distances due to the presence of hurricanes and these precipitation events need to be included in the gridded product. In addition, the final data selected by the previously mentioned procedure were not analysed for homogeneity and the station records were used without filling data gaps (i.e. data series reconstruction). Although gap-filling can be used to generate a complete data series for the considered time period – and thus keeping a

95   uniform number of stations for the interpolation – it was decided to avoid it in order to use the original data, because the reconstruction process is generally based on weighted averages or modeling that consists of creating a reference series formed as a weighting model of the data observed at neighbouring stations (Serrano-Notivoli and Tejedor, 2021), which is some type of interpolation (Daly, 2006). Further work can be done to address these issues and interested readers are referred to Serrano-Notivoli and Tejedor (2021) for a detailed analysis of QC on the development of gridded climate datasets.

100   The interpolations were developed using Kriging with External Drift on a local ~~neighborhood (KED_t~~neighbourhood (KED_l) using topography as an auxiliary variable because it has been found to be an adequate technique to interpolate both precipitation and temperature on a daily and yearly basis ~~(Carrera-Hernández and Gaskin, 2007; Page et al., 2022; Carrera-Hernández et al., 2024)~~
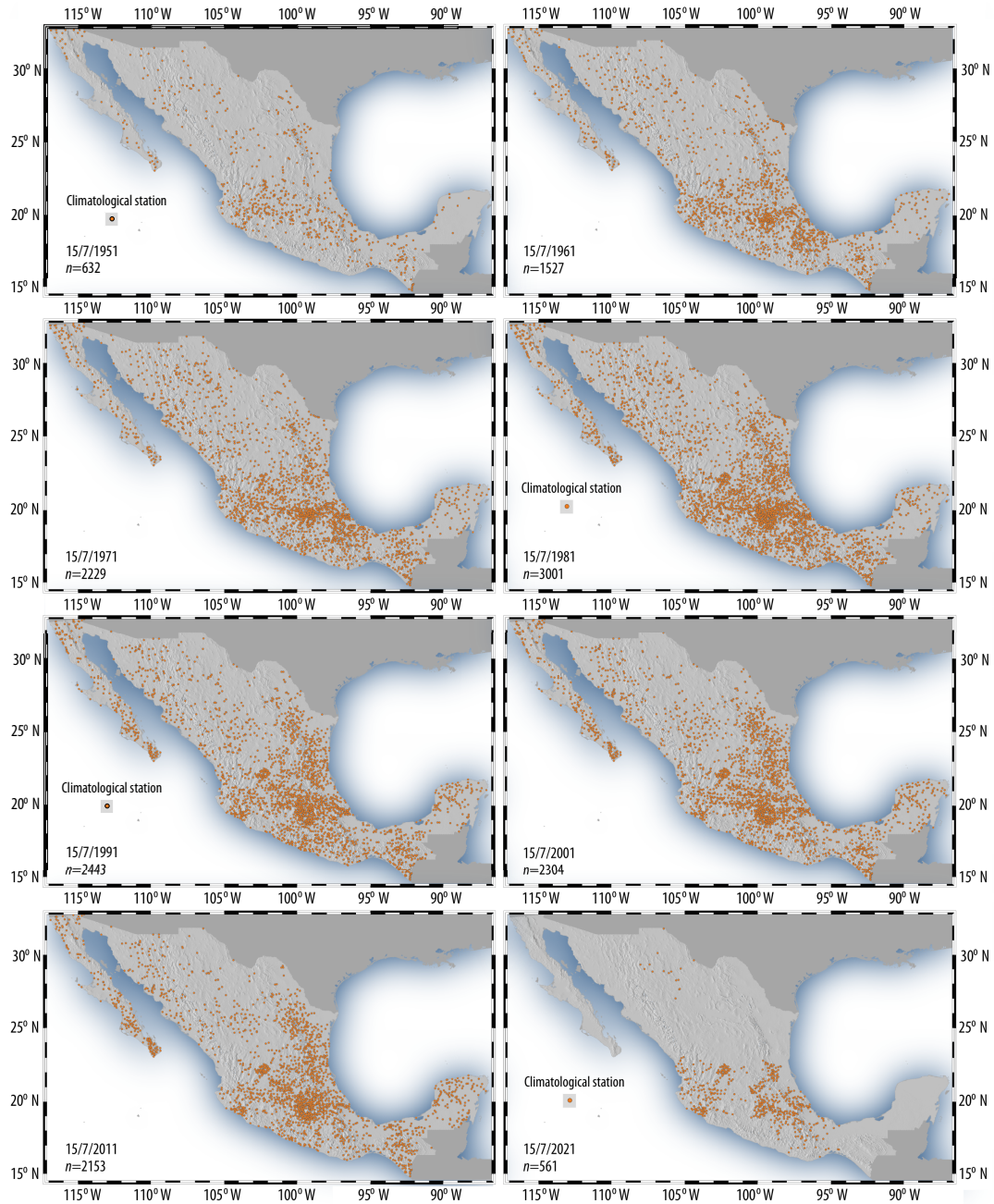
Figure 4. Spatial coverage of weather stations from 1951–2021. It can be seen that in 2021 the spatial coverage is mainly limited to Mexico's central region, which is why even though there are some climate data for 2021, the MexHiResClimDB does not include 2021.

(Carrera-Hernández and Gaskin, 2007; Page et al., 2022; Carrera-Hernández et al., 2025). The ALOS AW3D30 DEM was used as auxiliary variable because it has been shown to better represent Mexico's topography (Carrera-Hernández, 2021), and the original 1 arc-sec AW3D30 DEM was resampled to a 20 arc-sec resolution using the r.neighbors command of the GRASS GIS. There are currently several books and articles that describe the equations and assumptions behind Kriging and its different variants (Isaaks and Srivastava, 1989; Goovaerts, 1997, 2000; Carrera-Hernández and Gaskin, 2007; Pebesma, 2014); accordingly, they are not discussed in this work, ~~but interested readers are referred to some of them (Isaaks and Srivastava, 1989; Goovaerts, while guided~~. Interested readers can find examples on the implementation of Kriging with R and gstat ~~are given~~ in Pebesma and Benedikt (2023) and Pebesma and Bivand (2023).

The interpolations were carried out using the GRASS GIS and R with the libraries RPostgreSQL, parallel, gstat, sp and rgrass7 (Conway et al., 2023; Pebesma and Bivand, 2023). The climate records (i.e. time series data) were stored in PostgreSQL while all the generated raster files were stored in the file structure of GRASS (Fig. 2). A Gaussian semivariogram was fitted to the experimental semivariogram of each variable for each day using GSTAT's automated fitting procedure – which uses an iterative reweighted least squared estimation (Pebesma, 2014). However, some semivariograms had to be manually fitted when the automated procedure could not fit the semivariogram parameters. Based on the author's previous work on nation-wide interpolation of yearly precipitation in Mexico, a cut-off distance of ~~140~~ 180 km and a local ~~neighborhood~~ neighbourhood of 30 stations was used, because KED_{zl} with these parameters adequately interpolates precipitation even when this process is anisotropic ~~(Carrera-Hernández and Gaskin, 2007; Carrera-Hernández et al., 2024)~~(Carrera-Hernández and Gaskin, 2007; Carrera-Hernán. These parameters were recommended by Carrera-Hernández et al. (2025) after a detailed comparison of different Kriging variants at the national level using both global and stratified domains that used different auxiliary variables with a combination of different cut-off distances and local neighbourhoods. This comparison showed that KED_l using elevation as a secondary variable with a cut-off distance of 180 km and a local neighbourhood of 30 stations provided the best representation of yearly precipitation in Mexico (a code snippet is provided in Carrera-Hernández (2025p)).

Each ~~interpolation required~~ daily interpolation took 26 GB of RAM and a processing time of approximately 16 minutes; ~~the interpolations~~ and they were carried out on three workstations with multi-core processors and 256 GB of RAM located at the Hydrogeomatics Laboratory of the Geosciences Institute, UNAM. As previously mentioned, the interpolations were carried out using R's parallel library and it was found that the use of five cores for each interpolation provided the fastest interpolation time ($\approx$16 minutes, because parallelizing the process also requires time). For each variable 25,564 rasters were interpolated, thus yielding a total of 76,692 rasters, which required a computation time of ~~1.227×10^6 minutes~~ 28 months (although this time was larger because some days required a manual adjustment of the semivariograms).

4 ~~Validation~~

In order to estimate the errors in spatial climate datasets, a combination of approaches should be used, involving data that are as independent from those used to generate the datasets along with common sense in the interpretation of results (Daly, 2006). Accordingly, the MexHiResClimDB was validated using : 1) leave-one-out cross validation – from which different performance

metrics were obtained, 2) visual comparison of extreme precipitation events, 3) use of independent data acquired at different Automatic Weather Stations (AWSs) located throughout Mexico, and 4) comparison of the spatial distribution of accumulated monthly precipitation at both national and watershed scales.

~~To validate the undertaken interpolations,~~

4 Validation

140

The first validation step consisted of leave-one-out cross validation (which ~~basically~~ computes an interpolated value at the location of each station used, without using its value for interpolation) ~~was applied to~~ of each variable for each day~~. With these data~~, which were written as `.csv` files for further analyses. Using these files, four different measures of error were used to validate the interpolations: Coefficient of Determination (R^2), Coefficient of Efficiency (COE), Mean Absolute Error (MAE),

145 and Index of Agreement (IOA), which are explained in detail by Legates and McCabe (1999). The ~~Coefficient of Efficiency (COE), Mean Absolute Error (MAE) and the Index of Agreement (IOA)~~ COE, MAE and IOA are reported in this work because average-error and agreement measures based on sums of error magnitudes are – in general – superior to comparable measures based on sums of squared errors ~~(Willmott et al., 2015)~~(i.e. RMSE, Willmott et al. (2015)). Furthermore, the RMSE is not reported in this work because it is a function of the MAE, the variablity within the distribution of errors and the area

150 represented by the interpolated variable (Willmott and Matsuura, 2006).

In addition to the aforementioned indices, the coefficient of determination (R^2) is also reported in this work due to the ease of its interpretation, as it describes the proportion of the total variance in the observed data that can be explained by the model (i.e. if R^2=0.80, then the model ~~explaines~~ explains 80% of the variability in the observed data) and is given by:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(Tm_i - \overline{Tm})(T_i - \overline{T})}{\sum_{i=1}^{n}\left(Tm_i - \overline{Tm}\right)^{0.5}\sum_{i=1}^{n}\left(T_i - \overline{T}\right)^{0.5}} \right)^2 \tag{1}$$

155 where T_i refers to the ith interpolated temperature value, Tm_i refers to the ith measured temperature value, n is the number of measurements, while $\overline{Tm}$ and $\overline{T}$ represent the mean for the entire dataset of the observed and interpolated values, respectively. The coefficient of determination (R^2) varies from 0.0–1.0; however, ~~it~~ the information it provides is limited because it standardizes the differences between the observed and simulated means and variances because it only evaluates linear relationships between the variables, thus it is insensitive to the additive and proportional differences between the observed and

160 interpolated values (Legates and McCabe, 1999).

The Coefficient of Efficiency (COE) is an improvement over R^2 because it is sensitive to differences in the observed and interpolated means and variables (Legates and McCabe, 1999), ranging from $-\infty$ to 1.0, and obtained by:

$$COE = 1.0 - \frac{\sum_{i=1}^{n}|Tm_i - T_i|}{\sum_{i=1}^{n}|Tm_i - \overline{Tm}|} \tag{2}$$

where a value of COE=1.0 represents a perfect model, and a value of COE=0.0 indicates that the model is not better to predict the observed values than the observed mean, while negative values indicate that the model is less effective than the observed mean in predicting the variation in observations.

The Mean Absolute Error (MAE) is also ~~used~~ reported in this work because it is an unambiguous and more natural measure of average error than the Root Mean Square Error (RMSE, Willmott and Matsuura (2005)) due to the bias of RMSE when large outliers are present (Legates and McCabe, 1999) and because the RMSE does not describe average error alone and its use has been discouraged (Willmott and Matsuura, 2005). The MAE is determined by:

$$MAE = \frac{\sum_{i=1}^{n} |Tm_i - T_i|}{n} \tag{3}$$

The modified Index of Agreement (IOA, Legates and McCabe (1999)) has the advantage that errors and differences are not inflated by their squared values and is computed by:

$$IOA = 1.0 - \frac{\sum_{i=1}^{n} |Tm_i - T_i|}{\sum_{i=1}^{n} \left(|T_i - \overline{Tm}| + |Tm_i - \overline{Tm}| \right)} \tag{4}$$

Another advantage of the IOA is that it is related to the Mean Average Error (MAE) and the Mean Absolute Deviation (MAD) as follows:

$$IOA = 1.0 - \frac{MAE}{MAD} \tag{5}$$

The above metrics were computed daily for each of the three interpolated climate values through the use of the `OpenAir` library for R (Carslaw and Ropkins, 2012). Due to the large variability of precipitation in Mexico, the Mean Absolute Error (MAE) was only determined for T_{min} and T_{max} as shown in Fig. 5, where it can be seen that the MAE values are lower for T_{min} than for T_{max}, and that for both temperatures the MAE is lower for the summer months. For T_{min}, the MAE varies between 1.5–2.0 °C from June through October and between 2.0–2.5 °C for the remainder months, except for some days in January and December, where it reaches 2.5–3.0 °C. A similar behavior is observed for the MAE values of T_{max}, but with MAE values 0.5°C higher.

The three dimensionless performance indices (R^2, COE, and IOA) obtained for the three climate variables are shown in Fig. 6, where it can be seen that T_{min} exhibits the highest values for the three performance indices; in fact, for T_{min}, $R^2 > 0.7$ on most days (with $0.6 < R^2 < 0.7$ only for some days). For the same climate variable (T_{min}), the Index of Agreement (IOA) is > 0.8 for several days – in a similar (but more persistent) temporal distribution as MAE (Fig. 5) – with the remaining days having values of $0.7 < IOA < 0.8$. Of interest is the fact that the values of the three performance indices are ~~better during summerand particularly better~~ higher during summer, in particular for the 1960–1990 period (as can be easily observed in Fig. 5(b)), with precipitation exhibiting a similar – but more subtle – ~~behavior~~behaviour. The values of R^2, COE, and IOA for precipitation are lower due to its heterogeneity (which is why the MAE is not shown for this variable in this first validation step).
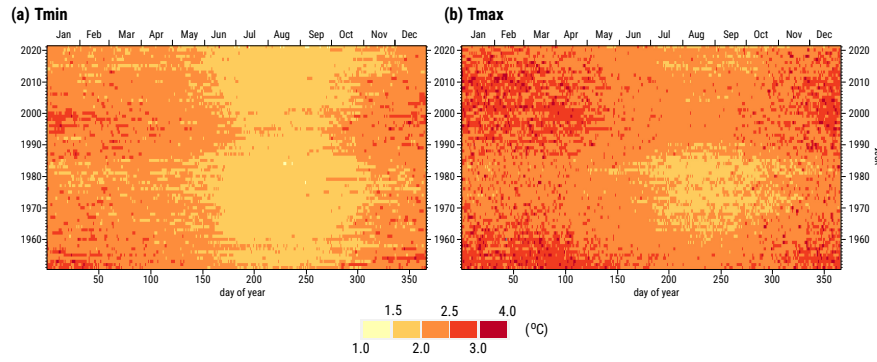
Figure 5. Mean Absolute Error of daily interpolations for: (a) T_{min}, (b) T_{max}.
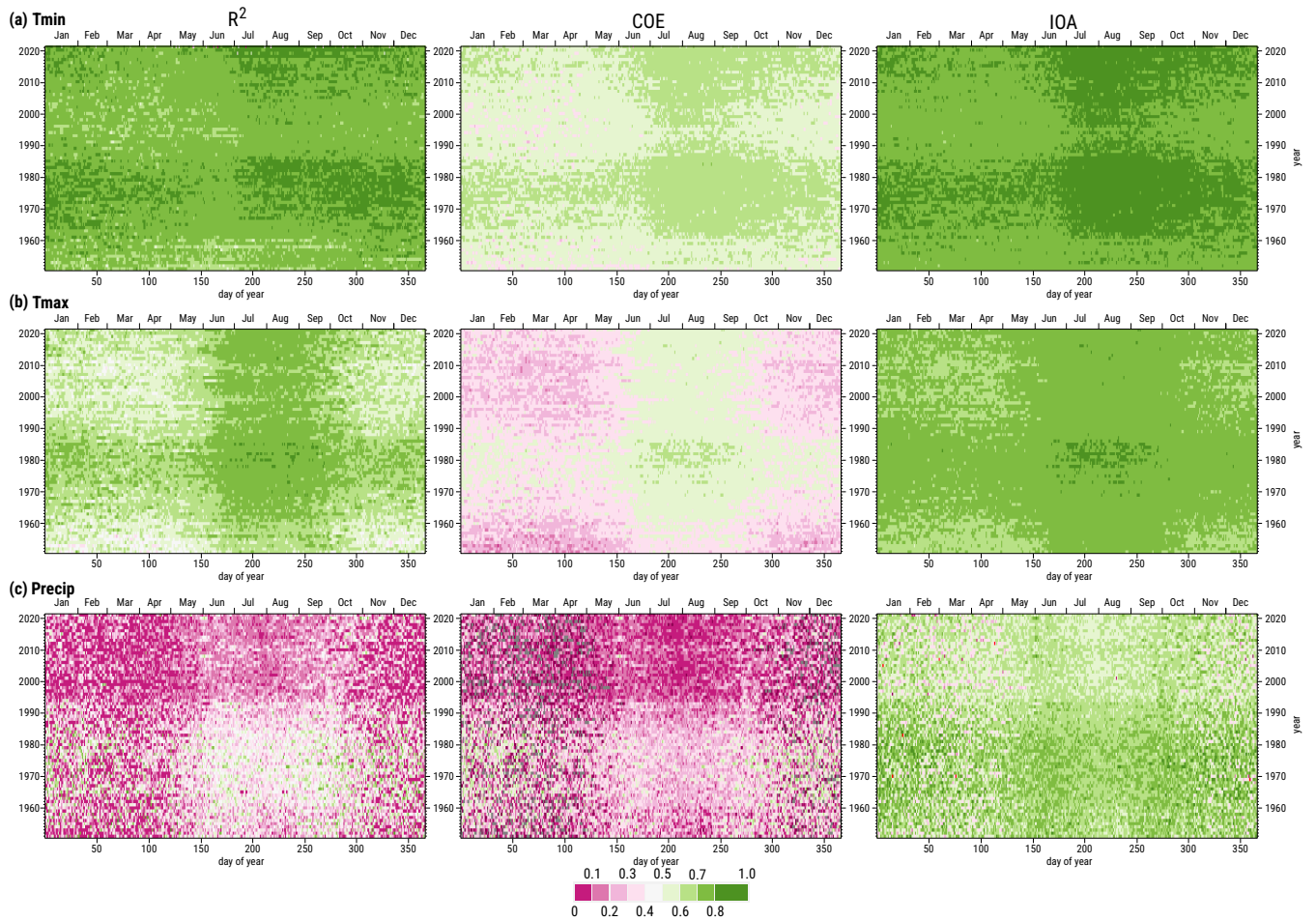
Figure 6. Performance indices – coefficient of determination (R^2), Coefficient of Efficiency (COE) and Index of Agreement (IOA) – for the daily interpolations developed in this work: (a) Minimum Temperature, (b) Maximum Temperature, and (c) Precipitation.

5 Comparison with other datasets and applications

Different gridded climate datasets that cover Mexico are currently available, as summarized in Table 4. These datasets have been used in Mexico for different purposes: to ~~analyze~~ analyse the duration and intensity of Mexico's midsummer drought during the 1981–2010 period (Perdigón-Morales et al., 2018), to study climate trends in both the North American Monsoon (NAM) and Mid-Summer drought (MSD) regions (Cavazos et al., 2020), to ~~analyze~~ analyse trends of daily rainfall indices (Colorado-Ruiz and Cavazos, 2021), or to validate the accuracy of their precipitation values on Mexico's northwestern region (Esquivel-Arriaga et al., 2024; de la Fraga et al., 2024).

Some of these datasets were developed through interpolation of observed climate values at meteorological stations (Daymet, L15) while others were derived from satellite observations (PERSIANN-CDR) or a mixture of these two methodologies (CHIRPS). Due to Mexico's abrupt topography – and its effect on both temperature and precipitation – it is imperative for any method used to estimate the aforementioned climate variables in the country to consider the impact of topography on their spatial distribution. In order to validate the daily precipitation values of the HiResMexClimDB, different datasets obtained through either interpolation, remote sensing or a mixture of both are compared for five events of extreme precipitation. As can be seen in Table 4, the MexHiResClimDB has the finest spatial resolution and the longest temporal coverage of the five datasets that were compared, which are briefly described in the following paragraphs.

~~Spatial resolution and temporal coverage of the MexHiResClimDB and the other four gridded climate datasets used for comparison. Dataset Source Climate variable Spatial resolution Temporal Area MexHiResClimDB (Carrera-Hernández, 2025a) Weather station T_{min}, T_{max}, Precip 20" ≈ 0.6 km 1951 − 2020 Mexico Daymet (Thornton et al., 2021)Weather station T_{min}, T_{max}, Precip + 4 more* 35" ≈ 1.0 km 1980 − 2023 North AmericaL15 (Livneh et al., 2015) Weather station T_{min}, T_{max}, Precip 3'20" ≈ 6.0 km 1950 − 2013 North AmericaCHIRPS (Funk et al., 2015) Satellite and weather station Precip 3'00" ≈ 5.4 km 1981 − 2023 Global PERSIANN CDR (Ashouri et al., 2015) Satellite Precip15'00" ≈ 25.0 km 1983 − 2024 Global~~

5.1 Daymet

The latest release of the Daymet database (version 4, Thornton et al. (2021)) provides daily T_{min} and T_{max} along with Precip and other variables for North America – Mexico, the Conterminous U. S. (CONUS) and Canada – at a resolution of 1 km for the 1980–2023 period. The estimation of T_{min}, T_{max} and Precip in Daymet is achieved through a truncated Guassian filter that uses inputs from multiple weather stations and weights that reflect the spatio-temporal relationships between each cell and the surrounding stations. For each grid cell the station list and associated weights are calculated on a yearly basis for T_{min}, T_{max} and Precip; additionaly, this dataset also includes other secondary variables such as daylight average shortwave radiation, daily average water vapor pressure, daylength and an estimate of accumulated snowpack. The daylength estimate is based on geographic location and time of year, while the remainder secondary variables are derived from T_{min} and T_{max} and Precip based on atmospheric theory and empirical relationships, as detailed in Thornton et al. (2021).

5.2 L15

The L15 dataset (Livneh et al., 2015) is a daily gridded dataset with a resolution of 1/16° (3'20" or ≈6 km) derived from observed Precip, T_{min} and T_{max} for North America (Mexico, the CONUS, and regions of Canada south of 53 N°) for the 1950–2013 period. This dataset was created by applying the SYMAP interpolation algorithm, which uses Inverse Distance Weighting (IDW, Shepard (1984)). To develop this dataset, weather stations from Mexico, CONUS and Canada were used; however, the selection criteria on which stations to use were different according to their location: a minimum of >20 years of data were required for weather stations located in CONUS or Canada, while only >50 days of data were required for the weather stations located in Mexico.

In order to consider the effect of topography on both temperature and precipitation, a lapse rate of 6.5°C/km was applied to T_{min} and T_{max}, while precipitation was scaled based on existing estimates of monthly precipitation that were developed by considering topography into account. For the CONUS, the Parameter-elevation Regressions on Independent Slopes Model (PRISM, Daly et al. (1997)) dataset was used to scale precipitation, while for both Canada and Mexico the gridded climate dataset developed by Wehner et al. (2011) – which was obtained by using trivariate thin plate smoothing splines that employed latitude, longitude and elevation as predictors – was used to incorporate the effect of topography on precipitation.

5.3 PERSIANN-CDR

The global dataset Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR, Ashouri et al. (2015)) provides world wide daily precipitation at a resolution of 0.25° (≈25 km at the equator) since 1983. The PERSIANN dataset is estimated through an Artificial Neural Network (ANN) model that extracts cold-cloud pixels and ~~neighboring~~ neighbouring features from GEO infrared images that associates variations in each pixel's brightness temperature to estimate the pixel's rainfall rate and uses monthly data from the Global Precipitation Climatology Project (GPCP) to reduce biases in the estimation of precipitation (Ashouri et al., 2015).

5.4 CHIRPS

The Climate Hazards group Infrared Precipitation with Stations (CHIRPS) was developed using precipitation estimates based on infrared cold cloud duration observations calibrated through the Tropical Rainfall Measuring Mission Multi-satellite PRecipitation Analyis version 7 and a moving window regression that used latitude, longitude, elevation and slope – as detailed in Funk et al. (2015). The CHIRPS dataset provides world wide daily precipitation since 1981 with a spatial resolution of 3' (≈5.4 km at the equator). Although the main advantage of precipitation estimates derived from satellite observations is their aereal coverage, it should be kept in mind that the main drawback of estimates based on IR techniques (PERSIANN-CDR and CHIRPS) is that they are based on the assumption that colder clouds produce more intense rainfall and may miss heavy precipitation from shallow clouds (Behrangi et al., 2016).

Table 1. Spatial resolution and temporal coverage of the MexHiResClimDB and the other four gridded climate datasets used for comparison.

Dataset	Source	Climate variable	Spatial resolution	Coverage Temporal	Area
MexHiResClimDB (Carrera-Hernández, 2025a)	Weather station	T_{min}, T_{max}, Precip	20" ≈ 0.6 km	1951 – 2020	Mexico
Daymet (Thornton et al., 2021)	Weather station	T_{min}, T_{max}, Precip + 4 more*	35" ≈ 1.0 km	1980 – 2023	North America
L15 (Livneh et al., 2015)	Weather station	T_{min}, T_{max}, Precip	3'20" ≈ 6.0 km	1950 – 2013	North America
CHIRPS (Funk et al., 2015)	Satellite and weather station	Precip	3'00" ≈ 5.4 km	1981 – 2023	Global
PERSIANN CDR (Ashouri et al., 2015)	Satellite	Precip	15'00" ≈ 25.0 km	1983 – 2024	Global

*The other three variables included in Daymet are shortwave radiation, vapor pressure, snow water equivalent and day length.

5.5 Validation and comparison of Precipitation

The spatial variation of daily precipitation from the MexHiResClimDB dataset is compared with the previously mentioned datasets using five different events of extreme precipitation in Mexico (shown in Fig. 7 and caused by hurricanes or tropical storms). These events and their impact in Mexico are briefly described in the following paragraphs:

1. On September 16th of 1988, Hurricane Gilbert – once labelled the "storm of the century" because of the meteorological records it set (Meyer-Arendt, 1991) – caused torrential rains, which led to floods that caused fatalities and destruction in Monterrey, Mexico's third largest city (Meyer-Arendt, 1991). The first landfall of this hurricane was in Cozumel on September 14th as a Force 5 Hurricane, and due to the floods it caused, 225 people died in the Monterrey Metropolitan Area (Aguilar-Barajas et al., 2019).

2. The large rainfall events of October 4–6 of 1999 caused by Tropical Depression 11, that triggered several landslides in the northern Sierra of Puebla, affecting different areas (Capra et al., 2003a, b; Alcántara-Ayala, 2004; Borja-Baeza and Alcántara-Ayala, 2004; Alcántara-Ayala et al., 2006). A total of nearly 3000 landslides that ranged from soil slides to debris flows and avalanches occured in this area and on October 5, in the town of *Teziutlán*, a single landslide caused approximately 150 casualties in this town – with a total of 263 casualties – affecting 1.5 million people.

3. In July 2010 Hurrican Alex affected the Monterrey Metropolitan Area. Hurricane Alex was a Category 2 hurricane before landfall in Tamaulipas (Cázares-Rodríguez et al., 2017) and flash floods triggered by it caused fifteen fatalities in the Monterrey Metropolitan Area (Aguilar-Barajas et al., 2019).

4. The simultaneous ocurrance of the tropical storm Manuel on the Pacific coast and of hurricane Ingrid (Cat 1) on the Gulf of Mexico in September 2013 caused flooding and substantial damage in several states of Mexico. The rainfall from tropical storm Manuel triggered a landslide at the *La Pintada* village in Guerrero, causing 78 fatalities – with an additional eight missing people (Alcántara-Ayala et al., 2017). The floods caused by Manuel destroyed several highways
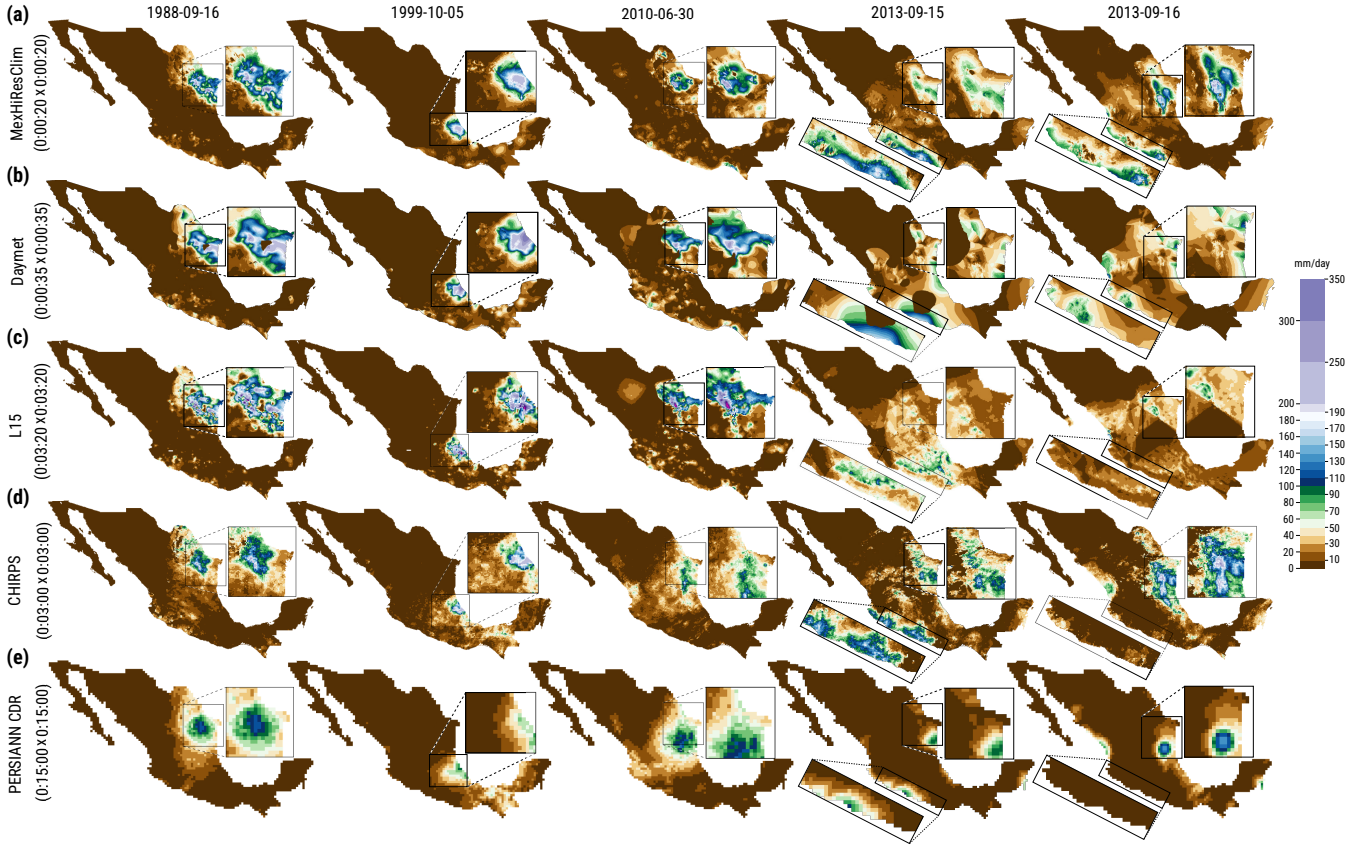
Figure 7. Spatial distribution of daily precipitation on five different days according to five different datasets: (a) MexHiResClim, (b) Daymet, (c) L15, (d) CHIRPS, and (e) PERSIANN CDR. These days were selected due to the large precipitation registered on them: September 16th, 1988, when Hurricane Gilbert caused large precipitation events – and flooding – in the Monterrey Metropolitan Area (MMA); October 5th 1999, when Tropical Depression 11 caused several landslides in the state of Puebla; June 30th 2010, when Hurricane Alex caused flooding in the MMA; September 15–16 of 2013, when Tropical storm Manuel in the Pacific and Hurricane Ingrid (Cat 1) in the Gulf of Mexico occurred simultaneously, causing floods – and triggering landslides – in different parts of Mexico. Only the MexHiResClim dataset adequately represents the precipitation caused in both the Pacific and the Gulf of Mexico regions for September 15–16 of 2013.

275 and bridges in the state of Guerrero, leaving 40,000 tourists stranded in the Acapulco bay because its main highway could not be used for a week, and its airport was flooded (CENAPRED, 2014).

The spatial distribution of precipitation for each of the aforementioned events is shown in Fig. 7, where it can be seen that PERSIANN-CDR exhibits lower precipitation values than the other four datasets for the five events considered, and that the areal coverage of precipitation obtained with both CHIRPS and PERSIANN CDR is smaller than the one obtained with the
280 other three datasets (which can be clearly seen for the events of 1988, 1999 and 2010). Of particular interest are the events of 2013: for September 15th, the MexHiResClimDB shows a large area with precipitation along the Pacific Coast caused by the Tropical Storm Manuel and another area with precipitation on the Northeastern region of Mexico, near the Gulf of Mexico's coast, caused by Hurricane Ingrid. This precipitation pattern is similar to that reported in Pedrozo-Acuña et al. (2014) and Rosengaus-Moshinsky et al. (2016), but not present on the L15 and Daymet datasets (as shown in Fig. 7(b) and (c)). In fact,
285 for this day (2013-09-15), the L15 dataset shows larger precipitation on the leeward side of the mountain range near Acapulco (Fig. 1(a)) – which is opposite to what actually occured, and represented on the other four datasets, although with some caveats, as is the case of Daymet (i.e. interpolation artifacts).

To provide further insight into how well these five datasets represent the registered precipitation of these five events, the Coefficient of Determination (R^2), the Coefficient of Efficiency (COE) and the Index of Agreement (IOA) were determined
290 for each event and dataset by leave-one-out cross-validation in the case of the MexHiResClimDB and through raster sampling of the remainder datasets using the `v.what.rast` command of the GRASS GIS (which updates the table attributes of a vector map, which in this case it consisted of the weather stations used to undertake KED$_1$ for each day). These performance statistics are shown in Fig. 8 along with their respective scattergrams of differences; however, it should be kept in mind that the performance metrics shown in the aforementioned figure for the MexHiResClimDB can not be directly compared with
295 the metrics of Daymet, L15 or CHIRPS, because for the latter three cases some of the weather stations used to compute the metrics were used to develop the datasets – in summarythus, the performance metrics obtained through cross-validation for the MexHiResClimDB are expected to be lower. As However, as can be seen on Fig. 8, the PERSIANN CDR is the dataset with the lowest metrics for four of the five events considered, followed by CHIRPS. For the precipitation events considered, L15 showed the largest performance values for only one event (1988-09-16), while Daymet for two (1999-10-05, 2010-06-
300 05) and MexHiResClimDB for the remainder two events (2013-09-15, 2013-09-16). However, the The differences between the performance statistics of L15 or Daymet compared to MexHiResClimDB for the first three events are not drastic: if the IOA is considered, for 1988-09-16, IOA$_{L15}$ = 0.863, while IOA$_{Daymet}$ = 0.811 and IOA$_{MexHiResClimDB}$ = 0.751. For the last two days – for which the performance indices of the MexHiResClimDB are better – these differences are larger, as for 2013-09-15, IOA$_{MexHiResClimDB}$ = 0.703, IOA$_{Daymet}$ = 0.525 and IOA$_{L15}$ = 0.334. However, if the Coefficient of Efficiency
305 (COE) is considered, the performance difference for the 2013 events is even larger, as COE$_{L15}$ = -0.331, COE$_{Daymet}$ = 0.051 and COE$_{MexHiResClimDB}$ = 0.405. From this analysis, it can be concluded that the precipitation patterns obtained with the MexHiResClimDB represent precipitation in a better way than the other four datasets considered in this work.
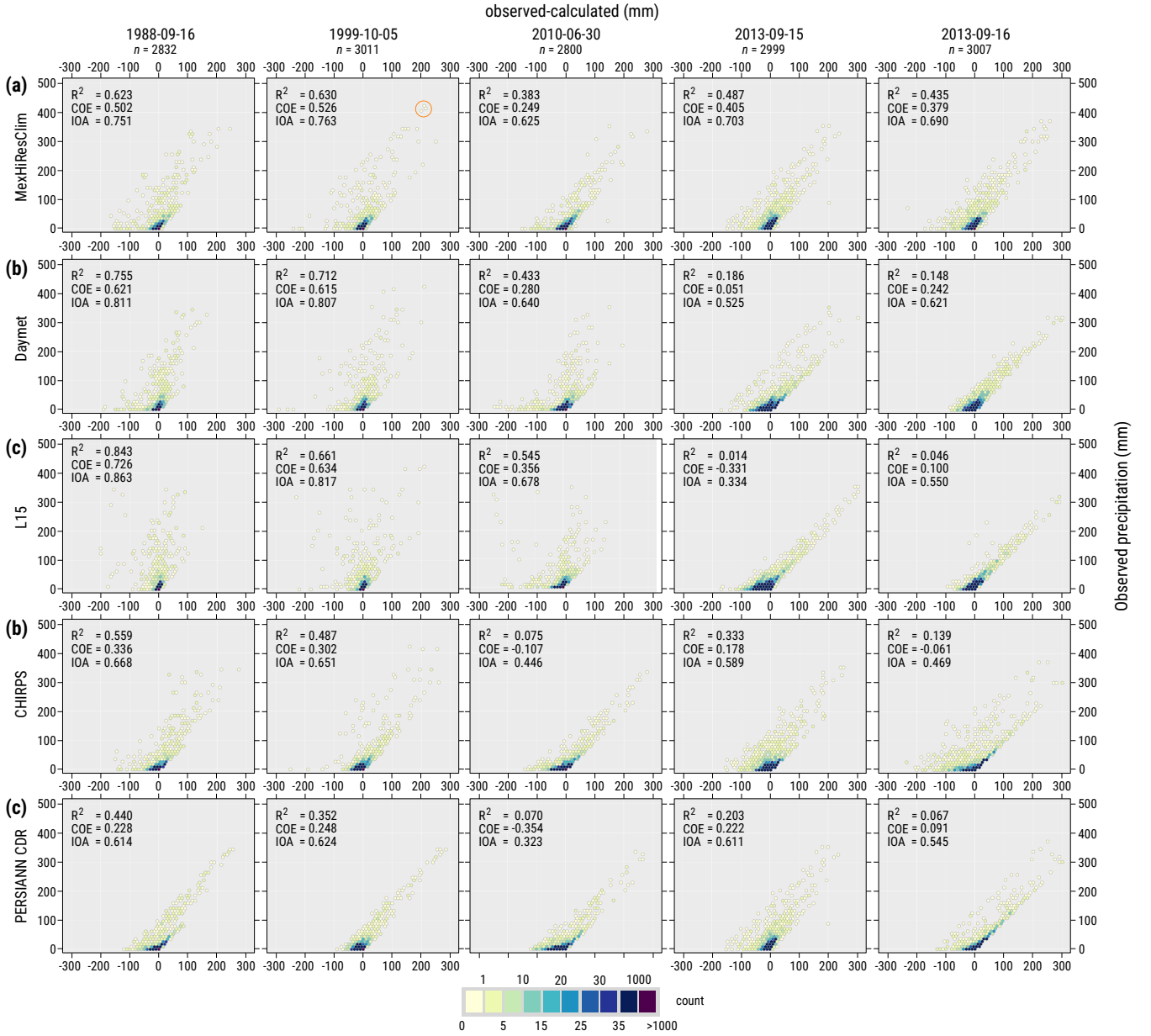
Figure 8. Validation scatterplots and error metrics for the precipitation events shown in Figure 7 for (a) MexHiRexClim, (b) Daymet, (c) L15, (d) CHIRPS, and (e) PERSIANN CDR. It is important to keep in mind that the values shown for Daymet and L15 do not correspond to cross-validation, but rather to sampling of the rasters provided by each dataset; accordingly, cross-validation of both Daymet and L15 would exhibit lower values for the performance indices used. It can be seen that MexHiResClim is the only dataset that adequately represents the precipitation events of September 15–16 caused by the presence of Tropical Storm Manuel and Hurricane Ingrid.

Table 2. Countrywide (a) Maximum and (b) Minimum values of daily, monthly and yearly accumulated values for precipitation.

Maximum						Minimum					
daily		monthly		yearly		daily		monthly		yearly	
date	precip (10^9 m^3)	date	precip (10^9 m^3)	date	precip (10^9 m^3)	date	precip (10^9 m^3)	date	precip (10^9 m^3)	date	precip (10^9 m^3)
1970-09-26	28.878	2013-09	387.716	1958	1,826.167	1951-01-17	0.000	1953-01	6.129	1953	1,106.309
1967-09-22	27.725	2010-07	379.970	2013	1,646.250	1953-02-05	0.000	1984-04	8.882	1957	1,158.624
1973-06-22	26.975	1958-09	355.837	1981	1,591.904	2007-03-01	0.001	1955-04	8.910	1994	1,161.331
1974-09-21	26.763	1955-07	354.232	1984	1,552.483	1956-03-05	0.001	1998-04	9.381	1951	1,164.111
1974-09-22	26.333	1955-09	351.492	1978	1,532.719	1977-03-13	0.001	1975-04	9.427	2011	1,164.234
2013-09-16	25.952	1973-08	349.515	1992	1,524.137	1953-01-05	0.001	1960-03	9.699	1962	1,176.275
2013-09-15	24.925	1978-09	339.886	2015	1,522.222	1953-01-10	0.002	1955-03	10.246	1987	1,179.864
1974-09-20	23.987	2014-09	338.307	1976	1,517.643	1956-03-04	0.002	1962-02	10.497	1982	1,191.108
1978-09-22	23.954	1976-07	334.457	2010	1,517.590	1952-02-13	0.003	1984-03	10.934	2009	1,192.888
2010-02-03	23.818	1969-08	333.554	1955	1,509.553	1958-02-27	0.003	1970-04	11.447	1956	1,201.528

5.5.1 Extreme records of precipitation

The ~~MexHiResCLimDB~~ MexHiResClimDB provides not only daily precipitation, but monthly and yearly aggregated values as
310 well. One application of this new database is the generation of a summary of precipitation extremes (i.e. wettest and driest) at
the aforementioned aggregation times (which will provide information that is not available for Mexico at this time). In order to
generate this information, the volume of precipitation was computed for each day and aggregated at the national level (the use
of volume was selected instead of mm due to the variability of precipitation in Mexico, as shown in Fig. 7). In order to create
this information, the r.univar command of the GRASS GIS – which computes univariate statistics (such as minimum,
315 maximum or total sum) from the non-null cells of a raster map – was used to compute the nation-wide precipitation volume
at the required aggretion time and stored in a relational database (PostgreSQL) for efficient handling ~~–~~(which is required for
monthly and daily data). With this procedure, the ten wettest and driest days, months and years were obtained and summarized
in Table 2, where it can be seen that the wettest day was 1970-09-26 (28.878×10^9 m^3) – which surprisingly was not caused
by a hurricane – while September of 1974 had the largest precipitation events for two and three consecutive days (21–22 and
320 20–22 respectively)~~which were~~ caused by the Fifi-Orlene Hurricane (which entered as Tropical Storm on the eastern side
of Mexico and moved westwards over the country to regain energy once entering the Pacific Ocean to turn into Hurricane
Orlene). Of interest is the 10th wettest day (2010-02-03), which occurred in February (outside Mexico's rainy season). The
month with the largest precipitation was September of 2013 (387.716×10^9 m^3, ~~due to the precipitation~~ caused by both Manuel
and Ingrid~~, as described in the previous section~~), while the wettest year was 1958 ($1,826.167 \times 10^9$ m^3). The driest year was
325 1953 while January of the same year was the driest month. An interesting finding was that the precipitation event of 2013-09-16
(25.952×10^9 m^3) is ranked as the 6th wettest day and that no wettnes or dryness tendency is easily seen on the values shown
in Table 2.

5.6 ~~Validation and comparison of temperature~~

~~The MexHiResClimDB also includes temperature data~~

5.5.1 Validation with independent data from Automatic Weather Stations

The comparison undertaken in the previous section used cross-validation for the MexHiResClimDB and sampling at the location of conventional weather stations for the other gridded datasets. To improve the validation – ~~as well as the~~ and intercomparison – of the gridded climate datasets, independent data from different Automatic Weather Stations (AWSs) operated by Mexico's National Weather Service (SMN) are used in this section. The AWSs acquire data every 10 minutes and these data are distributed by the SMN after being requested, with the data being provided in one spreadsheet file per month, which had to be proccessed in a similar way to that shown in Fig. 2. For the validation reported in this section, the two wettest months (September 2013 and June 2010, Table 2) were selected in order to have the maximum number of days with precipitation. Once the data from the AWSs were imported into PostgreSQL, stations with full data (registered every ten minutes) had to found; the original idea was to use the same stations for both months, but this proved to be unfeasible due to the lack of continuous data. The stations that were selected for this validation are shown in Fig. 9, where it can be seen that they are located at different elevations (with station 3102 found at an elevation of 4 meters in the Yucatán Península, while station 1504 is found at 4077 meters) and throughout Mexico (except Baja California, due to the lack of data on the AWSs located there). It can be noted that stations 501, 502 and 810 are very close to the border between Mexico and the U.S., and are useful to explore the impact of interpolation near Mexico's northern limit, because the MexHiResClimDB was developed by only using data from Mexico, whereas Daymet and L15 ~~datasets (Table 4). To validate and compare the precipitation of~~ used weather stations from Mexico, the U.S. and Canada, or globally (CHIRPS).

The ten-minute precipitation data recorded at each of the AWS was aggregated at a daily level through the use of the `subdaily2daily` function of the `hydroTSM` library for R (Zambrano-Bigiarini, 2024) for September 2013 and July 2010 (61 days) and their corresponding mass curve was also computed in order to compare it with the precipitation mass curve obtained from the different gridded precipitation datasets, as shown in Figures 10(a) and (b); for this analyses the PERSIANN dataset was not considered due to its low spatial resolution. The acquisition time of data from Mexico's conventional weather stations is 8:00 a.m., while the timestamp (i.e. date and time) from AWS data is in UTC time; this required to modify the acquisition time of the AWS data to match the acquisition time of the conventional stations (which was done in PostgreSQL).

It was decided to use the precipitation mass curve in order to show the difference in accumulated precipitation throughout and at the end of the month for each case. Of interest is the fact that the majority of the mass curves shown in Figures 10(a) and (b) show a similar behaviour to that of the AWS. In addition, the location of the AWS used also showed the limitations of the spatial resolution used in both L15 and Daymet for September 2013, as they do not provide data for the AWS 2307 (Fig. 10(a) – located near the border to Belize, Fig. 9) or for July 2010 in the case of L15, which could not provide data for the AWS 3102 (Fig. 10(b) – located at the north of the Yucatán Peninsula, Fig. 9).
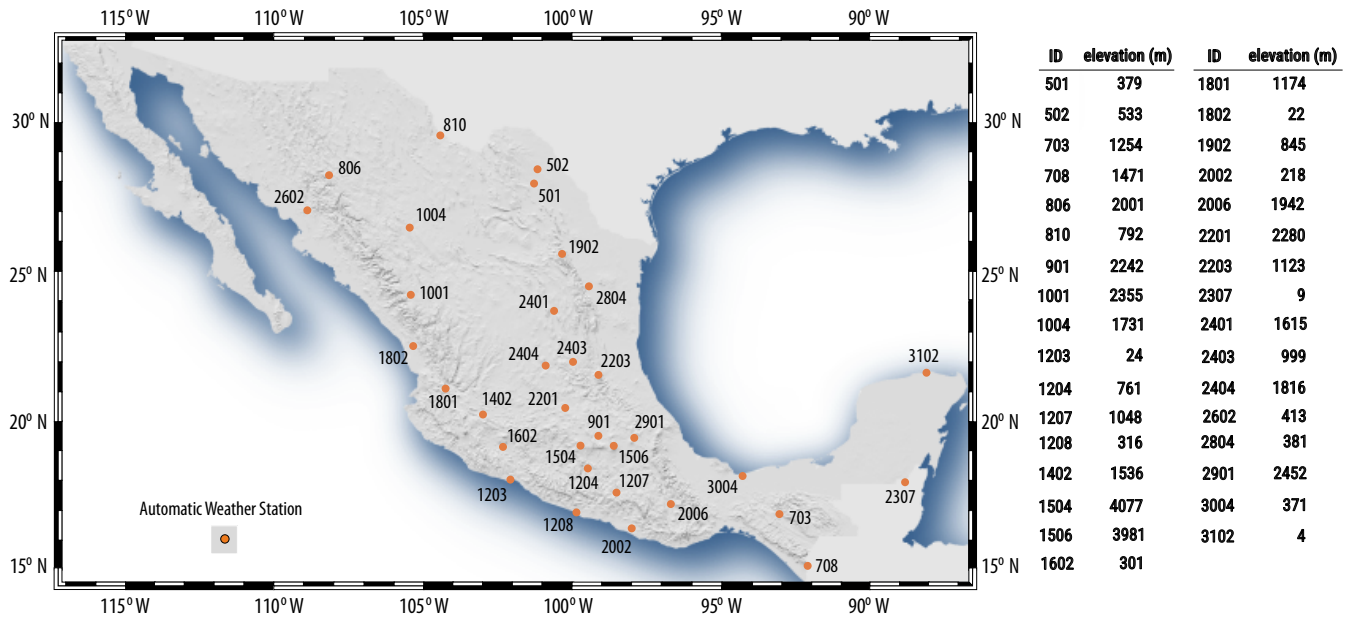
Figure 9. Location of Automatic Weather Stations used for further validation.

For the case of September 2013, the MexHiResClimDB provided a lower MAE for 17 of the 20 cases shown in Fig. 10(a) and for 7 of the 20 cases shown for July 2010 (Fig. 10(b)). Of interest is the precipitation registered at AWS 1208 in September 2013, as the monthly precipitation registered by this station was around 1100 mm, while the monthly value obtained with data from MexHiREsClimDB was around 600 mm (and 200 mm in the case of L15). For this AWS, the mass curves for all gridded datasets show a clear underestimation of precipitation for the events of September 14 and 15. This underestimation could have been caused by the fact that these events were caused by Tropical Storm Manuel and that the data registered at the conventional station (used on all the gridded datasets being compared) were wrong (i.e. lack of accesibility to the weather station or damage caused by strong winds). This situation contrasts with that of AWS 2203, which registered a total of 1025 mm, but which was represented similarly – although overestimated (1150 mm) – by MexHiResClimDB, while the monthly total by CHIRPS at this station was around 550 mm, and of only 50 mm by L15 Fig. 10(a).

Of particular interest are the mass curves of AWS 501 for Septmeber 2013, because this station is very close to the Mexico-U.S.A. border. At this station, the precipitation mass curve with the lowest MAE was the one obtained with precipitation from the MexHiResClimDB (MAE = 28.2 mm) followed by CHIRPS (MAE = 29.6 mm), L15 (MAE = 60.1 mm) and Daymet (MAE = 76.2 mm). Unfortunately, there was no data for station 502 on this date, but both stations have data for July 2010, when Hurricane Alex caused large precipitation events – and flooding – in northeastern Mexico. As can be seen on Fig. 10(b), the large precipitation events registered at station 501 and 502 for July 2010 were underestimated by all gridded datasets, in particular for station 501, which registered nearly 500 mm in 9 days, while the precipitation from MexHiresClimDB was nearly 300 mm (MAE = 181 mm), followed by Daymet (160 mm, MAE = 288 mm), CHIRPS (120 mm, MAE = 329 mm) and L15

(100 mm, MAE = 349 mm). For station 502 (which is closest to the border between Mexico and the U.S.A), data obtained from the MexHiResClimDB represented in a better way the precipitation registered at the AWS (with a MAE = 40.4 mm while
380   the MAE obtained for the remainder gridded datasets was above 200 mm).

While Figure 10 compares the monthly precipitation mass curves between the registered values at the AWSs and the different gridded datasets for September 2013 and July 2010, the daily comparison between these datasets is shown in Figure 11 along with their corresponding performance metrics. For the 61 days sampled at the location of the AWSs listed in Fig. 10, the lowest MAE (8.7 mm) was obtained with precipitation data from the MexHiResClimDB, followed by L15 (9.5 mm), Daymet
385   (10.1 mm) and CHIRPS (11.7 mm). As can be seen on Fig. 11, all gridded datasets tend to underestimate daily precipitation above 50 mm, which can also be observed on Fig. 8 when daily data from the traditional weather stations were used for the dataset comparison. Of note is how both Daymet and L15 could not be sampled at some stations due to their spatial extent and resolution and because L15 has an areal data gap where Lake Chapala – Mexico's largest lake (Carrera-Hernandez, 2018) – is located. This is why the hex-bin scatterplots of these datasets exhibit a lower number of points (1160 and 1098 for Daymet and
390   L15, respectively).

5.5.2   Monthly precipitation values at national and watershed levels

To conclude the comparison between the gridded precipitation datasets, this section compares the nation-wide accumulated monthly precipitation for both September 2013 and July 2010, while computing the corresponding monthly value for basins where one of the AWSs used in the precipitation mass curve analysis was found. The watersheds were obtained from the
395   recently developed Mexico's Watershed Database (MexWatDB, Carrera-Hernandez (2025)), which consists of a nested classification sytem for Mexico based on the Pfafstter organization scheme. These watersheds were selected in order to show the spatial variability of climate variables inherent to the spatial resolution of each dataset.

The accumulated monthly precipitation (at both national and basin levels) for September 2013 and July 2010 are shown in Figures 12 and 13 where it can be seen that Daymet and L15 show interpolation artifacts in large areas of northern Mexico – in particular for September 2013. At the national level, the MexHiResClimDB provides the largest volume of precipitation $(387.16 \times 10^9$ m^3), while with L15 the computed volume is of 309.789 $\times 10^9$ m^3. At the watershed level, for the basin where AWS 708 is found (with an area of 330 km^2), the volume obtained with MexHiResClimDB is 203.9 $\times 10^6$ m^3, which is four times the volume obtained with L15 (49.4 $\times 10^6$ m^3); this basin corresponds to $the MexHiResClimDB with other datasets, five differe$ 10^6 m^3 respectively).

The spatial distribution of precipitation for July 2010 (Fig. 13) near the Mexico-U.S.A. border – where AWS 502 is located – shows the large difference in precipitation obtained with the four different datasets considered. As can be seen on Fig. 10(b), the AWSs 501 and 502 registered large precipitation events for the first nine days of July 2010, due to Hurricane Alex, and flooding the city of *Piedras Negras*, near the outlet of the *Escondido* river and whose watershed corresponds to the one where AWS 502 is located. As previously mentioned, the MexHiResClimDB was the dataset that better represented the precipitation mass curve at this location, and the precipitation volumes obtained for the *Escondido* river watershed vary drastically between the gridded datasets (737.9 $\times 10^6$ m^3 for the MexHiResClimDB, and less than 375 $\times 10^6$ m^3 for the remainder datasets).
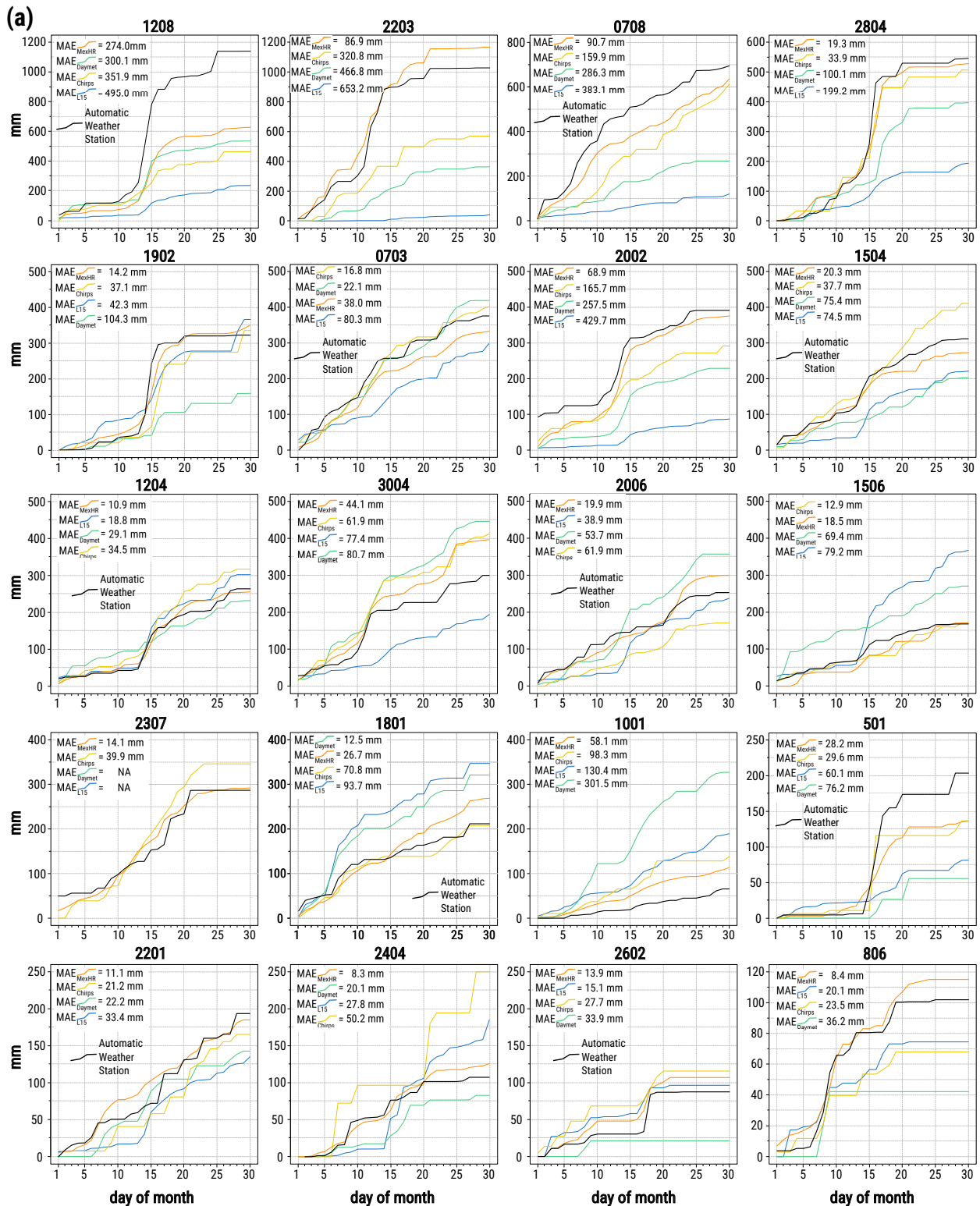
Figure 10. Comparison of precipitation mass curves obtained with data from Weather Stations and MexHiResClimDB, Daymet, L15 and CHIRPS for (a) September 2013, and (b) July 2010. Note the change in vertical scale for all plots, as it was modified in order to show the differences between the precipitation mass curves. For July 2010, the L15 could not be sampled because this dataset has an areal data gap where Lake Chapala – Mexico's largest lake (Carrera-Hernandez, 2018) – is located.

(b)
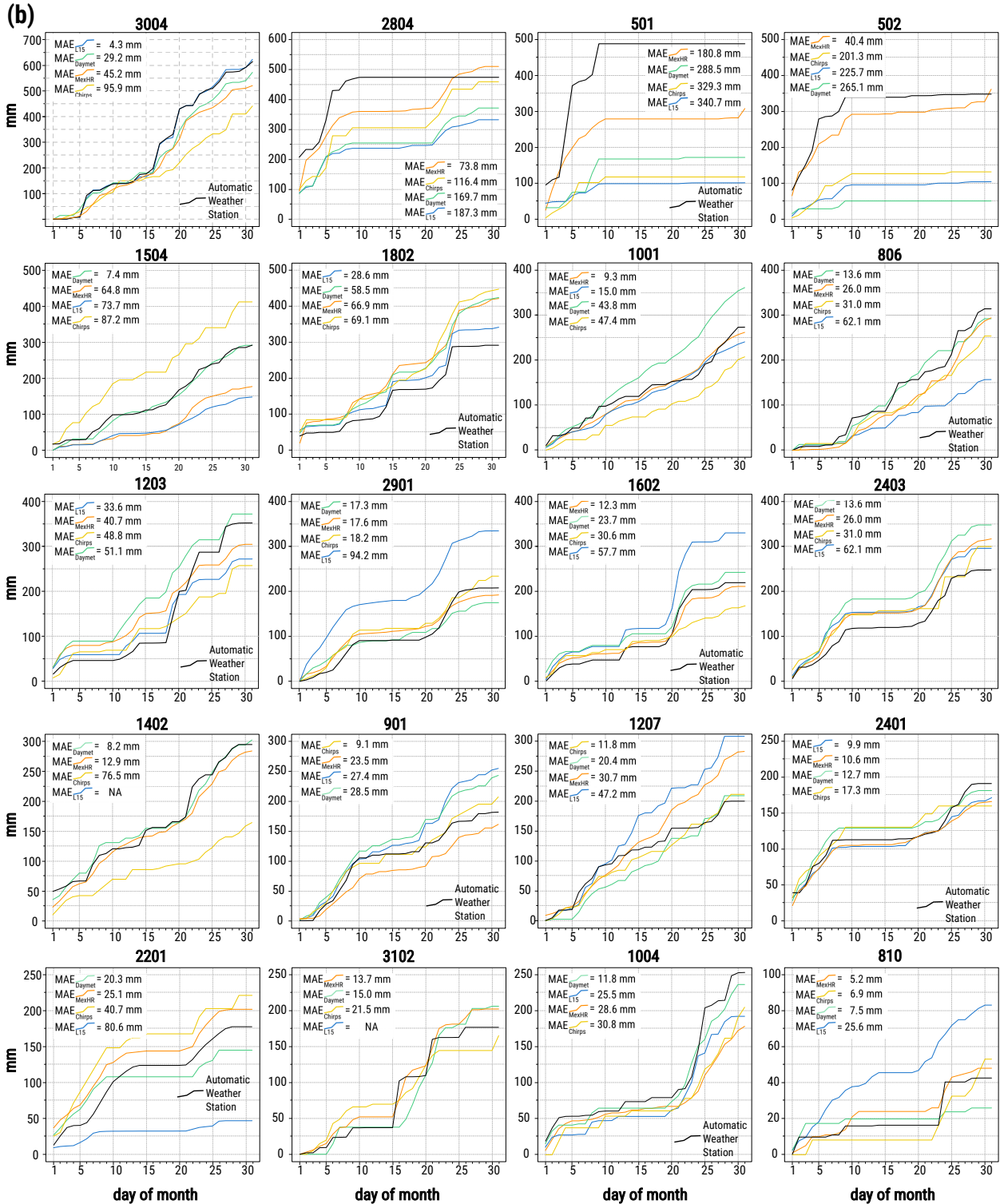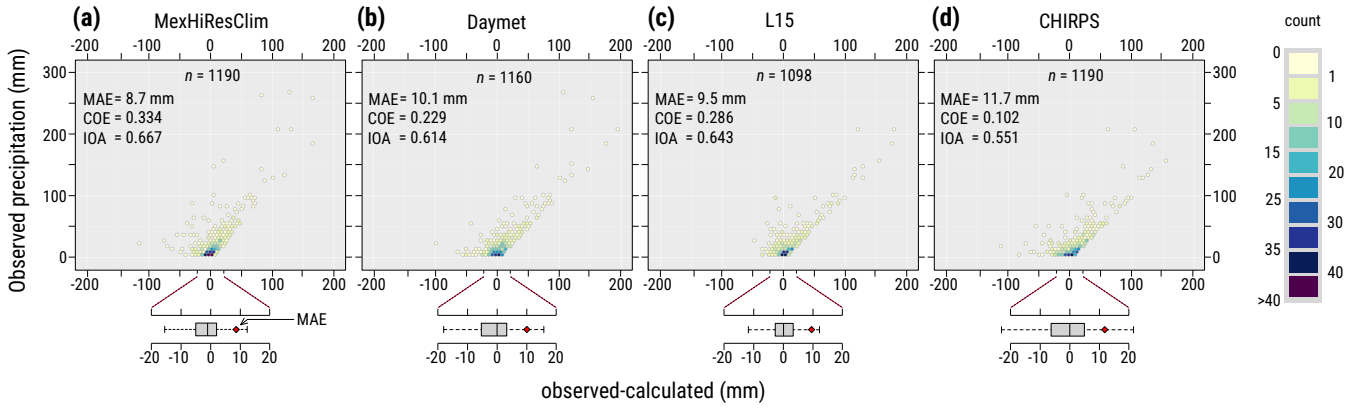
Figure 11. Comparison between daily values observed at the Automatic Weather Stations and gridded daily precipitation values for (a) MexHiResClimDB, (b) Daymet, (c) L15, (d) CHIRPS. The number of samples is lower for Daymet and L15 because they could not be sampled at the location of AWSs 2307 and 3102 due to their spatial resolution.

5.6 Validation and comparison of temperature

The MexHiResClimDB also includes temperature data, and the analysis of extreme values of temperature is useful in climate change analysis; however, there is currently no information available on the coldest or hottest days in Mexico for the 1951–2020 period. To provide this information, the temperature rasters generated in the MexHiResClimDB were processed in a similar way as was done with precipitation in order to summarize the ten hottest and coldest days, months and years for T_{min}, T_{max} and also T_{avg} (which is a ~~secondary~~ derived product of the interpolated temperature values), as shown in Table 3. In contrast to the precipitation values shown on Table 2 (which showed no particular trend), the highest temperature values (Table 3(a)) show a clear warming trend (in particular for minimum temperature): eight out of the ten days with the highest T_{min} occurred in 2020, the two months with the highest T_{min} were July and August of 2020 and the six years with the highest T_{min} were 2015–2020. The hottest day was 1998-06-15, while June of 1998 was the hottest month and 2020 the hottest year; of interest is to note that the four hottest years occurred between 2011–2020. The values of Table 3(b) show that the coldest day for the 1951–2020 period was January 12th of 1962, while the coldest month was January of 1987, with 1987 being the coldest year – as can be seen, the 1970–1979 decade was the coldest for the 1951–2020 period – and surprisingly, 2010 was the third coldest year.

To validate the interpolated temperature maps, the maximum and minimum values of T_{max} (1998-6-15, 1967-1-10) and T_{min} (2020-8-31, 1962-1-12) were selected to report their validation in detail. The results obtained with the leave-one-out cross-validation are shown on Fig. 14, where it can be seen that the performance statistics are better for T_{min} than for T_{max} – a fact that was previously pinpointed in the previous section. ~~For temperature, the minimum value for the Index of Agreement (IOA) is 0.714 (T_{max} for 1998-06-15), while its maximum value is IOA=0.822 (T_{min} for 2020-8-31).~~ The scattergrams of Fig. 14 only show the differences obtained through cross-validation for the MexHiResClimDB and not the other datasets because it is not possible to obtain cross-validation values for L15 or Daymet – and using the sampled values would yield values that are not
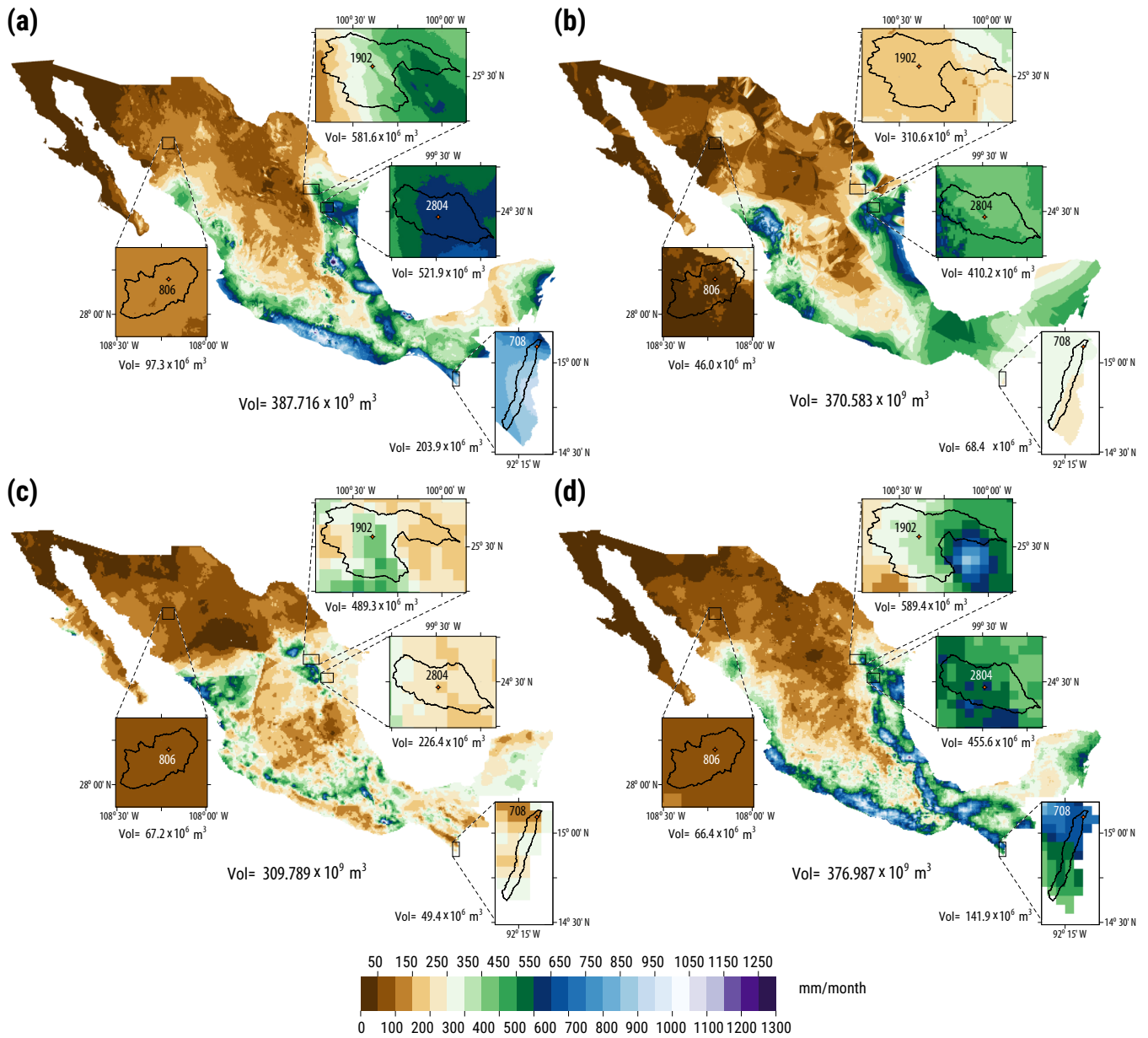
Figure 12. Monthly precipitation for September 2013 obtained by aggregating the daily gridded precipitation values of (a) MexHiResClimDB, (b) Daymet, (c) L15 and (d) CHIRPS. The precipitation volumes shown on the insets correspond to the monthly volume for each basin. Of note is the underestimation of precipitation for this month on Mexico's southwestern coast by both the Daymet and L15 datasets.
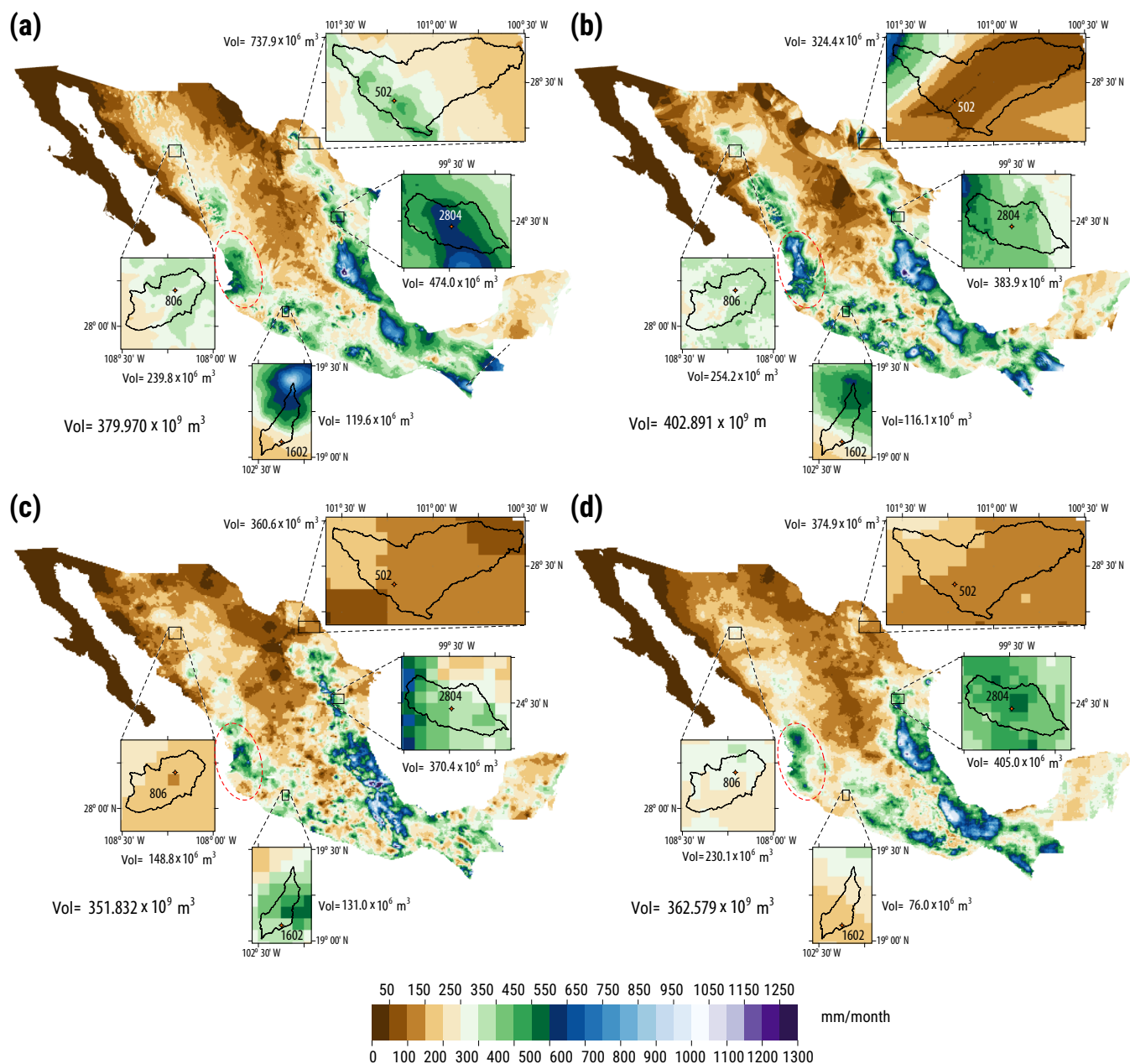
Figure 13. Monthly precipitation for July 2010 obtained by aggregating the daily gridded precipitation values of (a) MexHiResClimDB, (b) Daymet, (c) L15 and (d) CHIRPS. The precipitation volumes shown on the insets correspond to the monthly volume for each basin. Of note is the overestimation of precipitation by Daymet on the area highlighted by the red ellipse.

comparable to those shown on Fig. 14. ~~However, a visual comparison of the spatial distribution of both~~ To overcome this issue, the Automatic Weather Stations (AWSs) used in the validation of precipitation data were also used for temperature; however,

420 not all of the AWSs shown on Fig. 10 had a full record of temperature for September 2013 or July 2010 and other stations had to be selected. As done with precipitation data, the 10-minute registered acquisition time had to be modified in order to match the registration time of the conventional weather stations, after which the daily T_{min} and T_{max} ~~for the dates with minimum and maximum values of the aforementioned temperature values is done in Fig. ??, where it can be seen that the MexHiResClimDB isthe database with the longest temporal coverage and that neither~~ values were selected and compared with those obtained

425 from MexHiResClimDB, Daymet and L15~~nor Daymet were capable of showing the temperature extremes that were obtained through the MexHiResClimDB. For the hottest day of the 1951–2020 period (1998-06-15),~~, as shown in Figure 15(a) and (b) for September 2013 and July 2010 respectively, along with the daily MAE determined for each dataset.

The daily variation of both T_{min} and T_{max} is, in general, well represented by all gridded datasets and the largest MAE was obtained at the AWS 806 for the two months analysed: for September 2013 Daymet had a MAE = 7.7 and 7.1 °C for T_{min}

430 and T_{max}, while L15 ~~and Daymet show colder areas in the northern region of the Baja California Peninsula; a situation that is also apparent for the other three days of the comparison (although Daymet does not have data for the coldest day or the day when T_{max} is minimum~~ had a MAE of 9.1 and 4.8 °C for the considered temperatures. Finally, the plots for AWS 810 – ~~1962-01-12 and 1967-01-10, respectively– and L15 does not have data for~~ located at the Mexico-U.S. border, in the ~~day when the minimum value of T_{max} was obtained: 2020-08-31)~~. ~~Of note is the patchiness observed on the temperature maps provided by~~ city of *Ojinaga* (Fig. 9) – show that the three datsets provide similar temperature values at that location, although the data

435 from the MexHiResClimDB provided the lowest T_{max} MAE for September 2013 and for T_{min} of July 2010 (1.1 and 1.6 °C respectively).

To summarize the results of the 32 comparisons shown in Figure 15(a) and (b), for both T_{min} and T_{max}, the MAE obtained with the MexHiResCimDB was the lowest in 15 and 14 cases (T_{min} and T_{max} respectively) followed by Daymet (11 and 14

440 cases), while L15 provided the lowest MAE for 6 and 4 cases.

The daily differences between the gridded temperature datasets and the AWS data are summarized in Fig. 16, where it can be seen that temperature data from MexHiResClimDB exhibit the lower MAE of all datasets (1.7 and 1.8 °C for T_{min} and T_{max} respectively), followed by Daymet (2.0 °C for both temperatures) and L15 (~~Fig. ??~~2.4 and 2.5 °C).

5.7 Climate summary and ~~trends~~anomalies

445 The MexHiResClimDB (Carrera-Hernández, 2025a) provides daily, monthly and yearly data for T_{min}, T_{max}, Precip and also for T_{avg} (Carrera-Hernández, 2025b, c, d, e, f, g, h, i, j) – which is a derived product of the interpolated temperature variables – along with monthly and yearly climate normals of the four previously mentioned variables (Carrera-Hernández, 2025k, l, m, n, o). This new database has been used to determine climate extremes (both minimum and maximum, as shown in Tables 2 and 3). To show how the climate variables vary through the 1951–2020 period according to the different aggregation times on which

450 they are distributed, Fig. 17 was created, where it can be seen how temperature has increased in recent years.

Table 3. Country wide (a) Maximum and (b) Minimum values of daily, monthly and yearly ~~accumulated~~ averaged values for T_{min}, T_{avg} and T_{max}.

(a) Maximum values

T_{min}						T_{avg}						T_{max}					
daily		monthly		yearly		daily		monthly		yearly		daily		monthly		yearly	
date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)
2020-08-31	19.7	2020-07	18.9	2020	13.8	2018-07-24	26.8	1998-06	25.8	2020	21.1	1998-06-15	35.4	1998-06	34.2	2020	29.2
2020-07-11	19.6	2020-08	18.8	2015	13.6	2017-06-23	26.6	2019-08	25.7	2017	21.0	2011-05-28	35.2	1980-06	33.6	2017	29.0
2020-07-09	19.4	2019-08	18.7	2019	13.6	2018-07-25	26.6	2020-07	25.5	2019	20.9	1951-06-19	35.2	2011-06	33.5	2011	28.8
2020-08-30	19.4	2016-07	18.6	2018	13.5	2020-07-13	26.6	2020-08	25.5	2016	20.8	2011-05-27	35.1	2003-05	33.4	2019	28.8
2020-07-08	19.4	1957-07	18.5	2017	13.5	2020-07-12	26.5	1980-06	25.4	2018	20.7	1951-06-18	35.1	2005-06	33.4	1998	28.8
2020-08-29	19.3	1960-07	18.5	2016	13.5	2020-07-11	26.5	1980-07	25.3	2015	20.6	2018-05-31	35.0	1960-06	33.2	1995	28.7
2020-08-13	19.3	1969-07	18.5	1958	13.3	1998-07-13	26.5	2016-07	25.2	1995	20.6	2003-05-17	35.0	1998-05	33.2	1996	28.7
2020-07-13	19.3	1980-07	18.4	2014	13.2	1951-06-19	26.5	1960-06	25.2	1994	20.6	1998-06-19	35.0	1996-05	33.2	1999	28.6
1998-07-13	19.3	1953-07	18.4	1957	13.2	1998-07-14	26.5	1998-07	25.2	1954	20.5	2003-05-18	34.9	1953-06	33.2	2009	28.6
1998-07-14	19.3	1998-07	18.4	1994	13.1	2018-07-23	26.4	1953-06	25.2	2014	20.5	2018-06-02	34.9	2019-08	33.1	1953	28.6

(b) Minimum values

T_{min}						T_{avg}						T_{max}					
daily		monthly		yearly		daily		monthly		yearly		daily		monthly		yearly	
date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)	date	temp (°C)
1962-01-12	1.3	1987-01	5.7	1987	12.1	1962-01-11	9.0	1985-01	13.1	1976	19.4	1967-01-10	15.2	1992-01	20.0	1976	27.1
2011-02-04	1.8	1967-01	5.7	1975	12.1	2011-02-04	9.1	1966-01	13.2	1968	19.6	1967-01-09	16.0	1985-01	20.6	1966	27.2
1962-01-11	2.3	1951-01	5.9	2010	12.2	1967-01-10	9.2	1958-01	13.3	1966	19.6	1967-01-11	16.1	1966-01	20.6	1968	27.2
1951-02-03	2.5	1964-01	5.9	1999	12.2	1967-01-09	9.6	1964-01	13.4	1987	19.6	1962-01-11	16.1	1958-01	20.9	1992	27.4
1951-02-02	2.6	1973-12	5.9	1976	12.2	1962-01-12	9.7	1987-01	13.6	1975	19.7	1992-01-16	16.6	1981-01	21.1	1984	27.4
1997-12-15	2.6	1976-01	6.0	1970	12.3	1967-01-11	9.9	1973-01	13.7	1984	19.7	1992-01-17	16.7	2007-01	21.2	1958	27.5
1962-01-13	2.7	1999-12	6.0	1979	12.3	2011-02-03	9.9	1979-01	13.7	1973	19.7	1981-01-18	16.7	1976-12	21.3	1985	27.5
1973-12-21	2.7	2010-12	6.1	1973	12.3	1997-12-13	10.0	1967-01	13.8	2010	19.7	2011-02-04	16.8	1984-01	21.4	1987	27.6
1997-12-14	2.8	1960-02	6.1	1971	12.4	1971-01-07	10.3	1981-01	13.8	1964	19.8	2011-02-03	17.0	1964-01	21.4	1973	27.6
2011-02-05	2.9	1985-01	6.1	1974	12.4	1964-01-14	10.3	1992-01	13.8	1985	19.8	1985-01-13	17.1	1979-01	21.5	1964	27.6
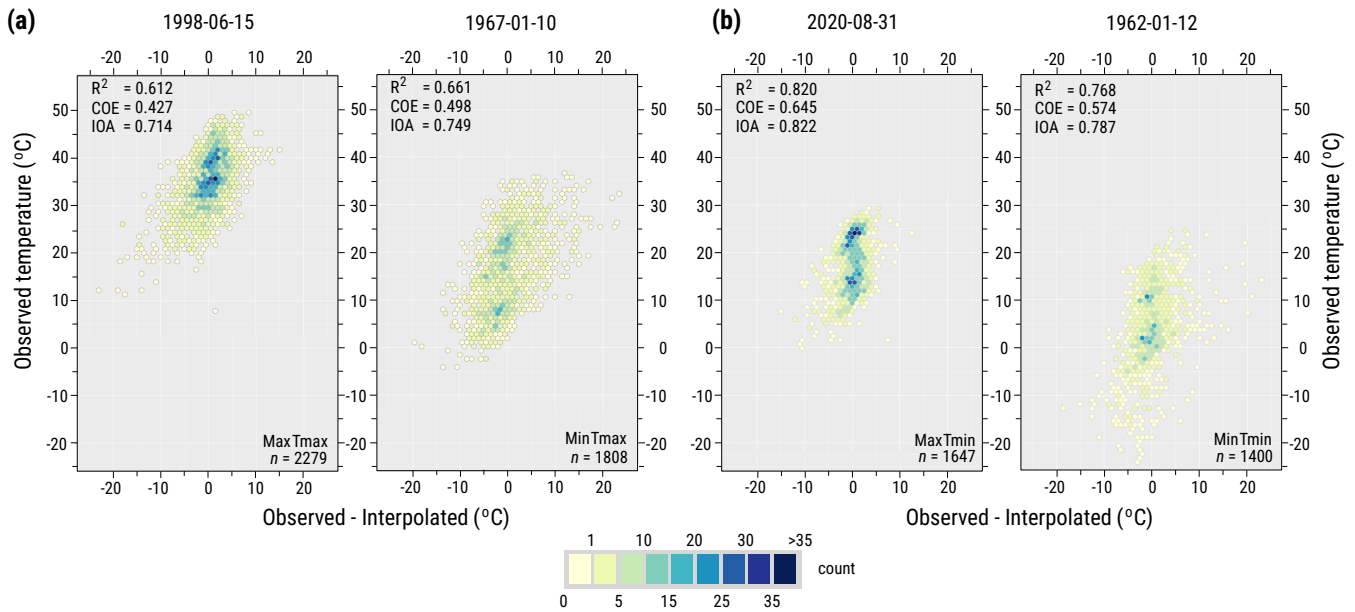
Figure 14. Scatterplots for minimum and maximum temperatures of (a) T_{max} and (b) T_{min} obtained through cross-validation.

It is interesting to note that for daily temperature, the distribution of warmer days for the three temperatures shown in Fig. 17 has an hourglass shape, thus showing colder temperatures for the 1961–1990 period – in particular for those years between 1970–1979. However, the bottom of this hourglass (for the three temperatures reported) is narrower than it is at its top, which clearly shows ~~a warming trend~~ an increase in temperature – ~~a trend that~~ which is easier to note when T_{min} (Fig. 17(c)) is
455   observed, as the days with T_{min} between 17.5–20.0 °C have increased since the summer of 2013. The ~~daily trends~~ increase in daily temperature observed in this Figure ~~are~~ is consistent with the values of Table 3, where it can be seen that the eight days with the highest T_{min} occurred in 2020. On a monthly basis, ~~the warming trend can also~~ this warming can also be seen on Fig. 17 because 2019 was the year where three months (Jun-Aug) exceeded 32.5°C, while on 2020 four months (May-Aug) also exceeded this temperature (Fig. 17(a)); similar warmer blocks are also observable for both T_{avg} and T_{min}. By looking at
460   the five-year moving average of monthly temperature, it can be seen that T_{min} started to increase in 2013 and by 2020 most months show T_{min} anomalies above 1°C, with July showing the largest increase for both T_{min} and T_{max}.

The summary plots of precipitation (Fig. 17(d)) clearly show that the rainy season in Mexico is from May to October, with the wettest months being July, August and September. The monthly plot of precipitation clearly shows that January of 1992 was an exceptionally wet January, while the yearly plot shows that 1958 was the wettest year for the period covered by the
465   MexHiResClimDB. However, although the aforementioned extremes can be easily seen on Fig. 17(d), ~~a trend~~ an increase or decrease in precipitation is difficult to observe, as is the case for yearly temperatures (due to the color scale selected in order to represent the temperature range between T_{min} and T_{max}). To provide further insight into this situation, the yearly anomalies of the four climate variables available on the MexHiResClimDB were computed for the ~~1951-2020~~ 1951–2020 period as shown
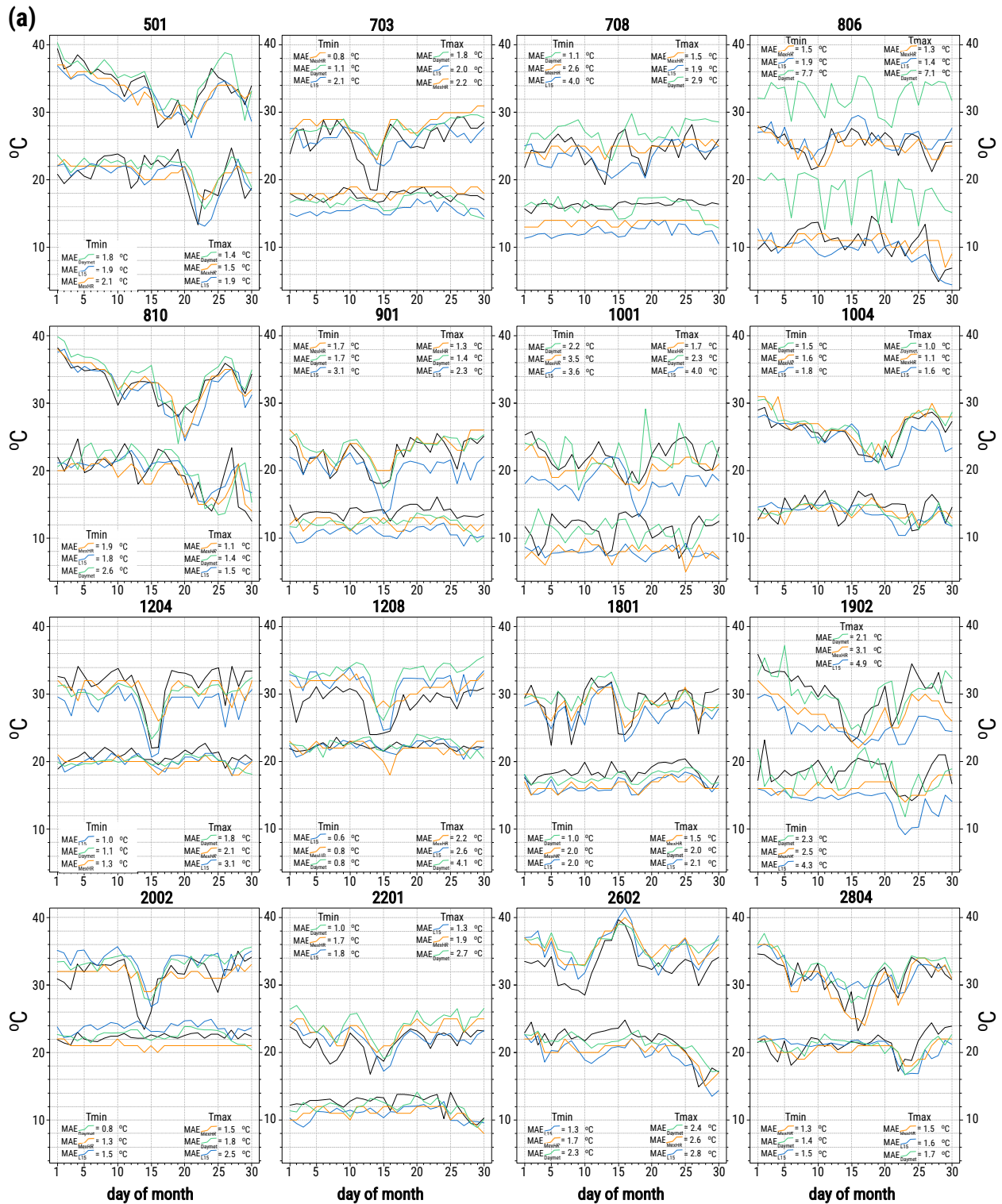
Figure 15. ~~Spatial distribution~~ Comparison of daily temperature obtained with data from Weather Stations, MexHiResClimDB, Daymet and L15 for (a) ~~Tmax~~ September 2013, and (b) ~~Tmin for their maximum and minimum events from 1951–2020. The temperature values shown on each map represent~~ July 2010. This figure shows the ~~average~~ limitation of storing temperature ~~for that day considering data in~~ the ~~entire country. The~~ MexHiResClimDB ~~is the only dataset that covers the entire 1951–2020 period~~ as integers (Tmin) for AWSs 708 and 2002.

(b)
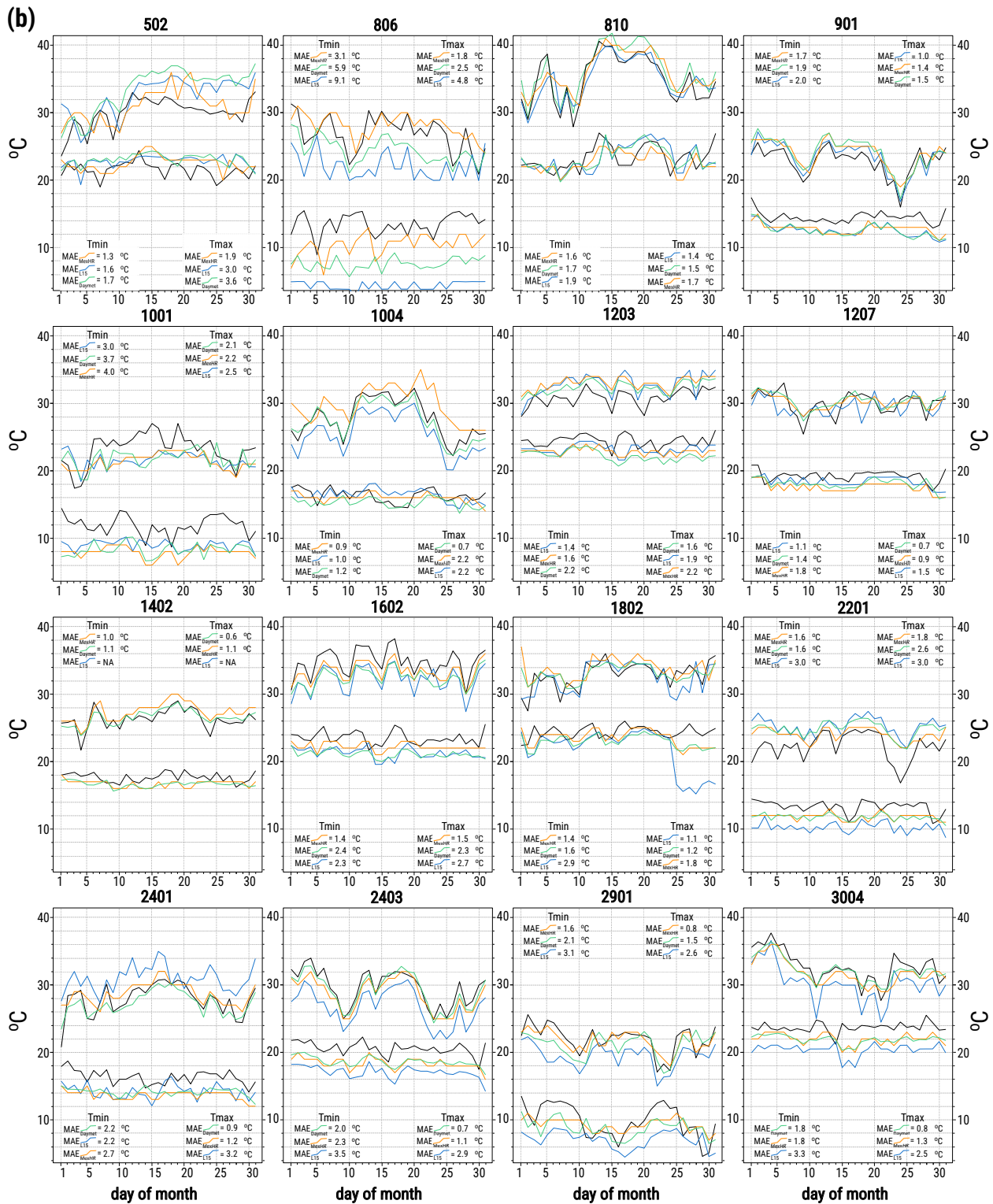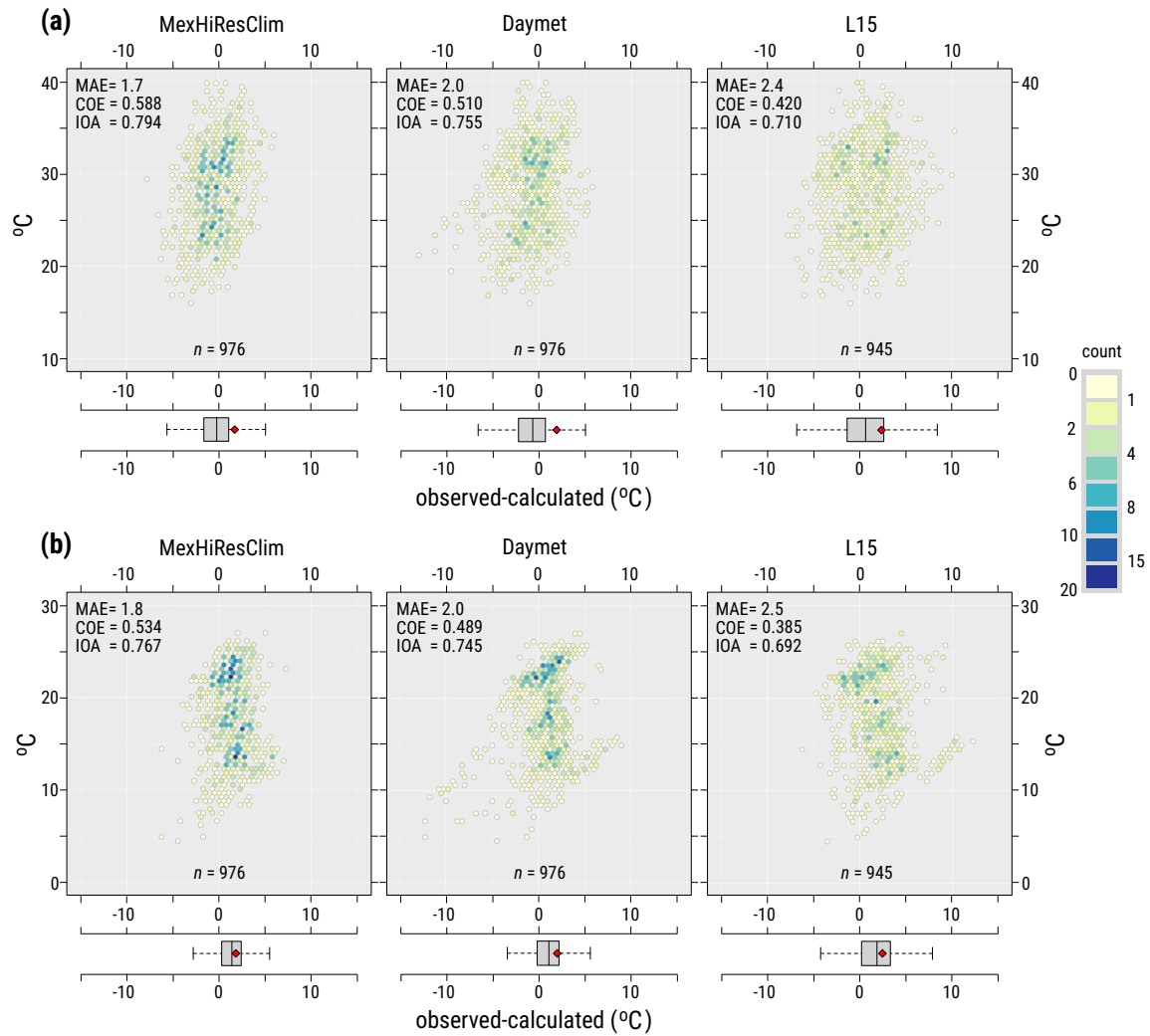
Figure 16. Validation of gridded temperature using data from the Automatic Weather Stations and 61 days sampled at 32 locations for (a) Tmax and (b) Tmin. It should be noted that the L15 dataset has a data gap on the western region of Central Mexico, which is why both the T_{min} and T_{max} hex-bin scattergrams were developed using 945 points instead of the 976 points used for MexHiResClimDB and Daymet.
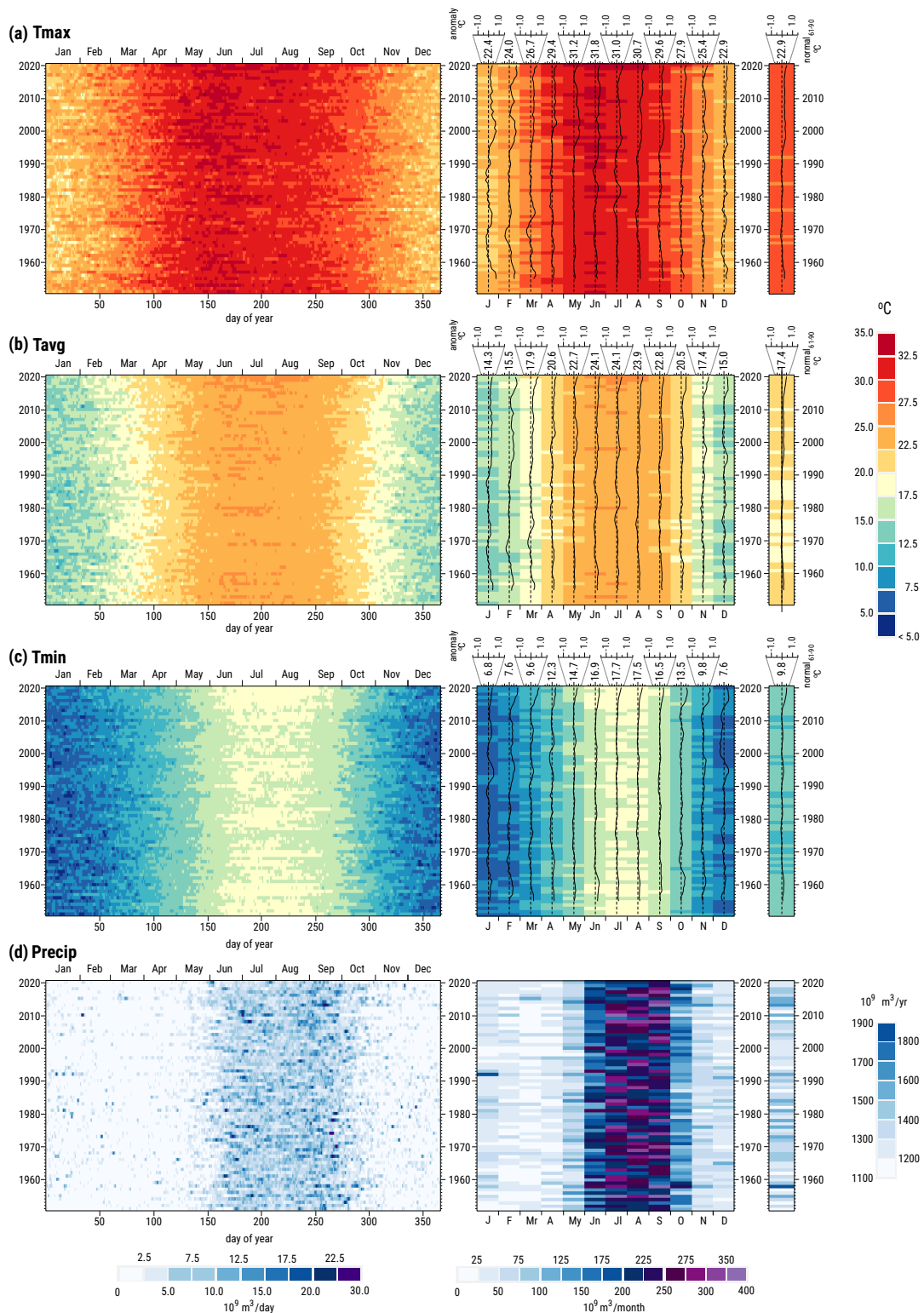
Figure 17. Countrywide daily, monthly and yearly values for the climate variables derived in this work. Precipitation values are given in volume (10^9 m^3) and it can clearly be seen that Mexico's rainy season is from May to October, that 1958 was the year with the largest precipitation for the 1951–2020 period, and that January 1992 was an exceptionally wet month. The ~~daily and monthly~~ variability of the five-year moving average of temperature ~~averages show a warming trend – particularly during~~ with respect to the ~~summer months~~ 1961-1990
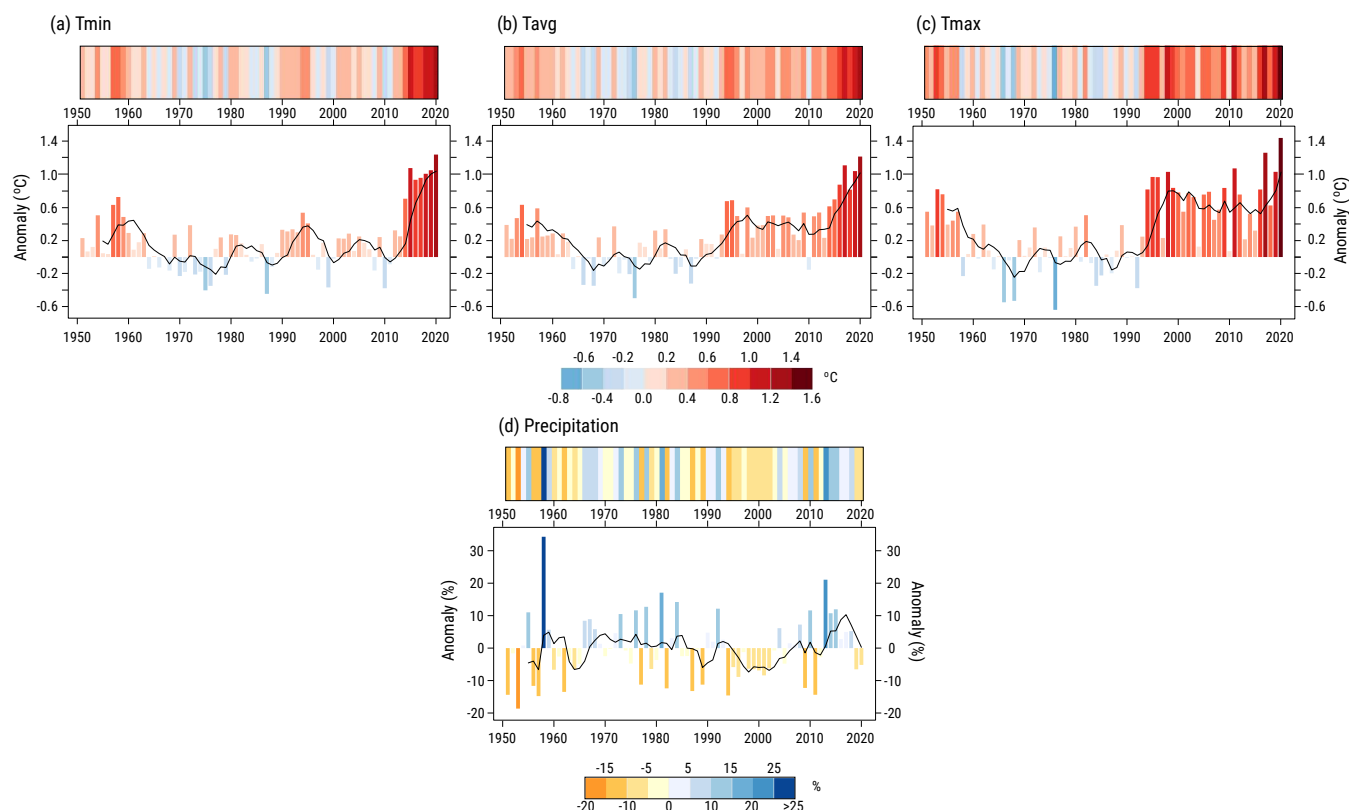
Figure 18. Yearly anomalies with respect to the 1961–1990 normals of (a) T_{min}, (b) T_{avg}, (c) T_{max} and (d) Precipitation

in Fig. 18 in both the well known "warming stripes" format and as bar plots (considering 1961–1990 as the baseline period,
470 according to the recommendation of the World Meteorological Organization, (WMO, 2017)).

The temperature anomalies shown in Fig. 18 show a warm period before 1964, and a fluctuation between colder and warmer periods between 1964–1990, although the anomalies for T_{min} (Fig. 18(a)) still fluctuated until 2010. However, in contrast to Fig. 17, the anomalies shown in Fig. 18 clearly ~~show a warming trend~~increase, in particular for T_{max}, because its ~~latest warming trend~~ positive anomalies (which started in 1992) reached 1.0 °C above the 1961–1990 normal in 1998 (thus yielding a T_{max}
475 increase rate of 0.13 °C/yr for the 1992–1998 period) and an anomaly of 1.4 °C in 2020, with a five-year moving average ~~trend~~ of around 0.6 °C between 2000–2016 that reached 1.0 °C in 2020 (Fig. 18(c)). For the 2016–2020 period the T_{max} increase rate was similar to that of 1992.

As previously mentioned, the T_{min} anomalies fluctuated between cold and warm periods until 2012, when the T_{min} anomaly started to increase from two cold years (2010 and 2011) to 0.2 °C in 2012 and to 1.2 °C in 2020 (Fig. 18(a)), which represents
480 a warming ~~trend~~ rate of 0.15 °C per year. Finally, the T_{avg} anomalies show a warming ~~trend~~ tendency that started in 1990 that drastically increased in 2014, with a maximum value of 1.2 °C in 2020.

The precipitation anomalies shift between dry and wet spells, although the 1995–2003 were dry years which were folowed by a couple of wet years before the 2009 and 2011 dry years (in fact, the five-year moving average shows a dry spell of 12 years between 1995–2007). This tendency in precipitation is important for water management and water supply, because even though the 2013–2020 five-year moving average shows a seven year wet period, 2019 and 2020 were dry years. When this tendency is compared with that of temperature, it can be concluded that higher temperatures will increase evapotranspiration and less water will be available. However, it should be kept in mind that even though the plots of Fig. 18 show ~~an undeniable warming trend in Mexico~~ that Mexico is warming, further studies are needed in order to pinpoint the areas where climate change is having a profound impact – for example, at the watershed level – a task that can be now done using the MexHiResClimDB.

6 Conclusions

This work presents Mexico's High Resolution Climate Database (MexHiResCLimDB), which is a new gridded, high-resolution ($\approx$600 m) climate dataset comprised of daily, monthly and yearly precipitation and temperature (T_{min}, T_{max}, T_{avg}). The monthly and yearly values were derived from daily interpolations obtained by using Kriging with External Drift on a local ~~neighborhood~~ neighbourhood through `gstat` within `R` and `GRASS` along a relational database developed in `PostgreSQL`.

Although different gridded climate datasets that cover Mexico are currently available, the MexHiResClimDB improves the spatio-temporal representation of climate variables over the country, as it is now the gridded climate database with the largest time coverage (1951–2020) and the highest spatial resolution (with 20" or $\approx$ 600 m). Furthermore, the daily precipitation data provided by this database ~~is the only one that~~ adequately represents the spatial variation of the extreme precipitation events ~~, in particular for the precipitation~~ that occurred during September 15–16 of 2013, caused by the presence of Tropical storm Manuel in the Pacific Ocean and Hurricane Ingrid (Cat 1) in the Gulf of Mexico. Using data from 61 days retrieved from Automated Weather Stations located throughout Mexico, it was found that precipitation data from MexHiResClimDB has a MAE=8.7 mm, followed by L15 (9.5 mm), Daymet (10.1 mm) and CHIRPS (11.7 mm). For T_{min} and T_{max}, the lowest MAE was obtained with MexHiResClimDB (1.7°C and 1.8 °C, respectively), followed by Daymet (2.0 °C for both temperatures) and L15 (2.4°C and 2.5 °C).

With this new database it was possible to summarize extreme events of precipitation and temperature in Mexico for the 1951–2020 period – a summary that was not available before. With this summary, it was found that the wettest year was 1958, the wettest day 1970-09-26 (which surprisingly was not caused by a hurricane) and that September of 1974 had the largest precipitation events for two and three consecutive days (21–22 and 20–22 respectively) – with September of 2013 being the wettest month. Regarding temperature extremes, it was found that eight out of the ten days with the highest T_{min} occurred in 2020, the two months with the highest T_{min} were July and August of 2020 and that the six years with the highest T_{min} were 2015–2020. When T_{max} was ~~analyzed~~analysed, it was found that the hottest day was 1998-06-15, while June of 1998 was the hottest month and 2020 the hottest year, and that the four hottest years occurred between 2011–2020.

Table 4. Datasets included in Mexico's High Resolution Climate Database (MexHiResClimDB, Carrera-Hernández (2025a)). The daily datasets are comprised of data for all days of the 1951–2020 period, monthly data for all months (i.e., 12 months × 70 years) and yearly data for all years (70).

Citation	Data	Digital Object Identifier
Carrera-Hernández (2025b)	Daily T_{min} for 1951–2020	DOI:10.6084/m9.figshare.28462808
Carrera-Hernández (2025c)	Daily T_{avg} for 1951–2020	DOI:10.6084/m9.figshare.28462835
Carrera-Hernández (2025d)	Daily T_{max} for 1951–2020	DOI:10.6084/m9.figshare.28462820
Carrera-Hernández (2025e)	Daily Precip for 1951–2020	DOI:10.6084/m9.figshare.28462796
Carrera-Hernández (2025f)	Monthly T_{min} for 1951–2020	DOI:10.6084/m9.figshare.28124789
Carrera-Hernández (2025g)	Monthly T_{avg} for 1951–2020	DOI:10.6084/m9.figshare.28462769
Carrera-Hernández (2025h)	Monthly T_{max} for 1951–2020	DOI:10.6084/m9.figshare.28462679
Carrera-Hernández (2025i)	Monthly Precip for 1951–2020	DOI:10.6084/m9.figshare.28462787
Carrera-Hernández (2025j)	Yearly data for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.28074998
Carrera-Hernández (2025k)	Monthly and yearly normals (1951–1980) for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.28464398
Carrera-Hernández (2025l)	Monthly and yearly normals (1961–1990) for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.28464458
Carrera-Hernández (2025m)	Monthly and yearly normals (1971–2000) for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.28464461
Carrera-Hernández (2025n)	Monthly and yearly normals (1981–2010) for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.28464488
Carrera-Hernández (2025o)	Monthly and yearly normals (1991–2020) for T_{min}, T_{avg}, T_{max} and Precip.	DOI: 10.6084/m9.figshare.~~28074998~~ 28464494

The anomalies obtained with this dataset show an undeniable ~~warming trend~~ increase of temperature in Mexico; however, further studies are needed in order to pinpoint the areas where climate change is having a profound impact – ~~for example, at~~
515   ~~the watershed level –~~ a task that can be now done using the MexHiResClimDB due to its spatio-temporal resolution.

7   Data availability

The datasets included in Mexico's High Resolution Climate Database (MexHiResClimDB, Carrera-Hernández (2025a)) are distributed as GeoTiffs (and not in NetCDF format in order to provide data for the 366 days of leap years); in addition to Precip, T_{min}, and T_{max}, the derived T_{avg} from the aforementioned temperature values is also distributed. The available data
520   are distributed at different aggregation (precip) or average (temperature) time steps (daily, monthly and yearly) as well as monthly and yearly normals for five different periods (1951–1980, 1961–1990, 1971–2000, 1981–2010 and 1991–2020). The datasets are provided on geographic coordinates referred to the WGS84 ellipsoid and are available on Figshare through the links provided in Table 4.

8   Code availability

525   A code snippet showing how Kriging can be parallelized in R using data stored in both PostgreSQL and the GRASS GIS is provided in Carrera-Hernández (2025p).

Competing interests

The author declares no conflicts of interest.

Acknowledgements

530 Funding for this project was provided by UNAM through project grant PAPIIT-IN101524. ALOS AW3D30 Digital Elevation Model provided by JAXA. The processing and generation of data, as well as the figures and this manuscript were all created with Open Source Software (PostgreSQL, R, GRASS GIS, QGIS, Inkscape and LaTeX) on Linux.

References

Aalto, J., Pirinen, P., and Jylhä, K.: New gridded daily climatology of Finland: Permutation-based uncertainty estimates and temporal trends in climate, Journal of Geophysical Research, 121, 3807–3823, https://doi.org/10.1002/2015JD024651, 2016.

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015, Scientific Data, 5, 1–12, https://doi.org/10.1038/sdata.2017.191, 2018.

Aguilar-Barajas, I., Sisto, N. P., Ramirez, A. I., and Magaña-Rueda, V.: Building urban resilience and knowledge co-production in the face of weather hazards: flash floods in the Monterrey Metropolitan Area (Mexico), Environmental Science and Policy, 99, 37–47, https://doi.org/10.1016/j.envsci.2019.05.021, 2019.

Alcántara-Ayala, I.: Hazard assessment of rainfall-induced landsliding in Mexico, Geomorphology, 61, 19–40, https://doi.org/10.1016/j.geomorph.2003.11.004, 2004.

Alcántara-Ayala, I., Esteban-Chávez, O., and Parrot, J. F.: Landsliding related to land-cover change: A diachronic analysis of hillslope instability distribution in the Sierra Norte, Puebla, Mexico, Catena, 65, 152–165, https://doi.org/10.1016/j.catena.2005.11.006, 2006.

Alcántara-Ayala, I., Garnica-Peña, R. J., Domínguez-Morales, L., González-Huesca, A. E., and Calderón-Vega, A.: The La Pintada landslide, Guerrero, Mexico: hints from the Pre-Classic to the disasters of modern times, Landslides, 14, 1195–1205, https://doi.org/10.1007/s10346-017-0808-9, 2017.

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies, Bulletin of the American Meteorological Society, 96, 69–83, https://doi.org/10.1175/BAMS-D-13-00068.1, 2015.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth System Science Data, 5, 71–99, https://doi.org/10.5194/essd-5-71-2013, 2013.

Beguería, S., Peña-Angulo, D., Trullenque-Blanco, V., and González-Hidalgo, C.: MOPREDAScentury: a long-term monthly precipitation grid for the Spanish mainland, Earth Syst. Sci. Data, 15, https://doi.org/10.5194/essd-15-2547-2023, 2023.

Behrangi, A., Guan, B., Neiman, P. J., Schreier, M., and Lambrigtsen, B.: On the quantification of atmospheric rivers precipitation from space: Composite assessments and case studies over the eastern north pacific ocean and the Western United States, Journal of Hydrometeorology, 17, 369–382, https://doi.org/10.1175/JHM-D-15-0061.1, 2016.

Borja-Baeza, R. C. and Alcántara-Ayala, I.: Procesos de remocion en masa y riesgos asociados en zacapoaxtla, puebla, Investigaciones Geograficas, 53, 7–26, 2004.

Capra, L., Lugo-Hubp, J., and Borselli, L.: Mass movements in tropical volcanic terrains: The case of Teziutlán (México), Engineering Geology, 69, 359–379, https://doi.org/10.1016/S0013-7952(03)00071-1, 2003a.

Capra, L., Lugo-Hubp, J., and Dávila-Hernández, N.: Fenómenos de remoción en masa en el poblado de Zapotitlán de Méndez, Puebla: Relación entre litología y tipo de movimiento, Revista Mexicana de Ciencias Geologicas, 20, 95–106, 2003b.

Carrera-Hernandez, J. J.: A tale of Mexico's most exploited-and connected-watersheds: the Basin of Mexico and the Lerma-Chapala Basin, Wiley Interdisciplinary Reviews: Water, 5, e1247, https://doi.org/10.1002/wat2.1247, 2018.

Carrera-Hernandez, J. J.: The MexHiResClimDB and the MexWatDB: Two new country-wide databases in Mexico to improve hydrological modelling from regional to local scales., https://doi.org/10.5194/egusphere-egu25-4925, 2025.

Carrera-Hernández, J.: Not all DEMs are equal: An evaluation of six globally available 30 m resolution DEMs with geodetic benchmarks and LiDAR in Mexico, Remote Sensing of Environment, 261, 19, https://doi.org/10.1016/j.rse.2021.112474, 2021.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): daily, monthly, yearly and 30 year normals of precipitation and temperature (minimum, average and maximum) for the 1951-2020 period at a resolution of 20 arc sec. https://doi.org/10.6084/m9.figshare.c.7689428.v2, 2025a.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): daily minimum temperature for Mexico https://doi.org/10.6084/m9.figshare.28462808, 2025b.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): daily average temperature for Mexico https://doi.org/10.6084/m9.figshare.28462835, 2025c.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): daily maximum temperature for Mexico https://doi.org/10.6084/m9.figshare.28462820, 2025d.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): daily precipitation for Mexico https://doi.org/10.6084/m9.figshare.28462796, 2025e.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): monthly minimum temperature for Mexico https://doi.org/10.6084/m9.figshare.28124789, 2025f.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): monthly average temperature for Mexico https://doi.org/10.6084/m9.figshare.28462769, 2025g.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): monthly maximum temperature for Mexico https://doi.org/10.6084/m9.figshare.28462679, 2025h.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): monthly precipitation for Mexico https://doi.org/10.6084/m9.figshare.28462787, 2025i.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): yearly data of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28074998, 2025j

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): Monthly and yearly normals (1951–1980) of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28464398, 2025k.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): Monthly and yearly normals (1961–1990) of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28464458, 2025l.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): Monthly and yearly normals (1971–2000) of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28464461, 2025m.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): Monthly and yearly normals (1981–2010) of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28464488, 2025n.

Carrera-Hernández, J. J. Mexico's High Resolution Climate Database (MexHiResClimDB): Monthly and yearly normals (1991–2020) of T_{min}, T_{avg}, T_{max} and Precipitation for Mexico https://doi.org/10.6084/m9.figshare.28464494, 2025o.

Carrera-Hernández, J. J. Parallel Kriging in R using PostgreSQL and GRASS; code snippet. https://doi.org/10.6084/m9.figshare.29840129, 2025p.

Carrera-Hernández, J. and Gaskin, S.: Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico, Journal of Hydrology, 336, 231–249, https://doi.org/10.1016/j.jhydrol.2006.12.021, 2007.

Carrera-Hernández, J. and Gaskin, S.: Spatio-temporal analysis of potential aquifer recharge: Application to the Basin of Mexico, Journal of Hydrology, 353, 228–246, https://doi.org/10.1016/j.jhydrol.2008.02.012, 2008a.

Carrera-Hernández, J. and Gaskin, S.: The Basin of Mexico Hydrogeological Database (BMHDB): Implementation, queries and interaction with open source software, Environmental Modelling & Software, 23, 1271–1279, https://doi.org/10.1016/j.envsoft.2008.02.012, 2008b.

Carrera-Hernández, J., Smerdon, B., and Mendoza, C.: Estimating groundwater recharge through unsaturated flow modelling: Sensitivity to boundary conditions and vertical discretization, Journal of Hydrology, 452, 90–101, http://www.sciencedirect.com/science/article/pii/S0022169412004283, 2012.

Carrera-Hernández, J. J., Carreón-Freyre, D., Cerca-Martínez, M., and Levresse, G.: Groundwater flow in a transboundary fault-dominated aquifer and the importance of regional modeling: the case of the city of Querétaro, Mexico, Hydrogeology Journal, 24, 373–393, https://doi.org/10.1007/s10040-015-1363-x, 2016.

Carrera-Hernández, J. J., Levresse, G. P., and Hernández-Espriú, J. A.: Geostatistical Analysis of Yearly Precipitation at the National Level: Stratification and Anisotropy Considerations in Mexico, , 2024. Hydrol. Sc. J., https://doi.org/https://doi.org/10.1080/02626667.2025.2534192, 2025.

Carslaw, D. C. and Ropkins, K.: Openair - An r package for air quality data analysis, Environmental Modelling and Software, 27-28, 52–61, https://doi.org/10.1016/j.envsoft.2011.09.008, 2012.

Castañeda, S., Botello, F., Sánchez-Cordero, V., and Sarkar, S.: Spatio-Temporal Distribution of Monarch Butterflies Along Their Migratory Route, Frontiers in Ecology and Evolution, 7, https://doi.org/10.3389/fevo.2019.00400, 2019.

Cavazos, T., Luna-Niño, R., Cerezo-Mota, R., Fuentes-Franco, R., Méndez, M., Martínez, L. F. P., and Valenzuela, E.: Climatic trends and regional climate models intercomparison over the CORDEX-CAM (Central America, Caribbean, and Mexico) domain, International Journal of Climatology, 40, 1396–1420, https://doi.org/10.1002/joc.6276, 2020.

CENAPRED: Socioeconomical impact of the main disasters that occurred in Mexico during 2013, Tech. rep., https://www.cenapred.unam.mx/es/Publicaciones/archivos/324-NO.15-IMPACTOSOCIOECONMICODELOSPRINCIPALESDESASTRESOCURRIDOSENMXICOENELAO20 PDF, 2014.

Colorado-Ruiz, G. and Cavazos, T.: Trends of daily extreme and non-extreme rainfall indices and intercomparison with different gridded data sets over Mexico and the southern United States, International Journal of Climatology, 41, 5406–5430, https://doi.org/10.1002/joc.7225, 2021.

Conway, J., Eddelbuettel, D., Nishiyama, T., Kumar, S., and Tiffin, N.: RPostgreSQL: R Interface to the 'PostgreSQL' Database System. R package version 0.7-5, https://cran.r-project.org/package=RPostgreSQL, 2023.

Cuervo-Robayo, A. P., Téllez-Valdés, O., Gómez-Albores, M. A., Venegas-Barrera, C. S., Manjarrez, J., and Martínez-Meyer, E.: An update of high-resolution monthly climate surfaces for Mexico, International Journal of Climatology, 34, 2427–2437, https://doi.org/10.1002/joc.3848, 2014.

Cuervo-Robayo, A. P., Ureta, C., Gómez-Albores, M. A., Meneses-Mosquera, A. K., Téllez-Valdés, O., and Martínez-Meyer, E.: One hundred years of climate change in Mexico, PLoS ONE, 15, 1–19, https://doi.org/10.1371/journal.pone.0209808, 2020.

Cázares-Rodríguez, J. E., Vivoni, E. R., and Mascaro, G.: Comparison of two watershed models for addressing stakeholder flood mitigation strategies: Case study of Hurricane Alex in Monterrey, México, Journal of Hydrologic Engineering, 22, 1–16, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001560, 2017.

Daly, C.: Guidelines for assessing the suitability of spatial climate data sets, International Journal of Climatology, 26, 707–721, https://doi.org/10.1002/joc.1322, 2006.

Daly, C., Taylor, G., and Gibson, W.: The Prism Approach to Mapping Precipitation and Temperature, in: Proc., 10th AMS Conf. on Applied Climatology, 1997.

de la Fraga, P., Del-Toro-Guerrero, F. J., Vivoni, E. R., Cavazos, T., and Kretzschmar, T.: Evaluation of gridded precipitation datasets in mountainous terrains of Northwestern Mexico, Journal of Hydrology: Regional Studies, 56, https://doi.org/10.1016/j.ejrh.2024.102019, 2024.

Durre, I., Arguez, A., Schreck, C. J., Squires, M. F., and Vose, R. S.: Daily High-Resolution Temperature and Precipitation Fields for the Contiguous United States from 1951 to Present, Journal of Atmospheric and Oceanic Technology, 39, 1837–1855, https://doi.org/10.1175/JTECH-D-22-0024.1, 2022.

Englehart, P. J. and Douglas, A. V.: Characterizing regional-scale variations in monthly and seasonal surface air temperature over Mexico, International Journal of Climatology, 24, 1897–1909, https://doi.org/10.1002/joc.1117, 2004.

Esperon-Rodriguez, M., Correa-Metrio, A., Beaumont, L. J., Baumgartner, J. B., and Lenoir, J.: Evaluating the impact of protected areas in lowering extinction risks in a biodiversity hotspot, Biological Conservation, 297, 110 728, https://doi.org/10.1016/j.biocon.2024.110728, 2024.

Esquivel-Arriaga, G., Huber-Sannwald, E., Reyes-Gómez, V. M., Bravo-Peña, L. C., Dávila-Ortiz, R., Martínez-Tagüeña, N., and Velázquez-Zapata, J. A.: Performance Evaluation of Global Precipitation Datasets in Northern Mexico Drylands, Journal of Applied Meteorology and Climatology, 63, 1545–1558, https://doi.org/10.1175/JAMC-D-23-0227.1, 2024.

Farfán, L. M., D'Sa, E. J., biu Liu, K., and Rivera-Monroy, V. H.: Tropical Cyclone Impacts on Coastal Regions: the Case of the Yucatán and the Baja California Peninsulas, Mexico, Estuaries and Coasts, 37, 1388–1402, https://doi.org/10.1007/s12237-014-9797-2, 2014.

Feeley, K. J., Bravo-Avila, C., Fadrique, B., Perez, T. M., and Zuleta, D.: Climate-driven changes in the composition of New World plant communities, Nature Climate Change, 10, 965–970, https://doi.org/10.1038/s41558-020-0873-2, 2020.

Fernández-Eguiarte, A. R., Romero-Cenento, R., and Zavala-Hidalgo, J.: Atlas climático de México y áreas adyacentes, vol. 1, Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, https://atlasclimatico.unam.mx/ACM/, 2012.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International Journal of Climatology, 37, 4302–4315, https://doi.org/10.1002/joc.5086, 2017.

Frick, C., Steiner, H., Mazurkiewicz, A., Riediger, U., Rauthe, M., Reich, T., and Gratzki, A.: Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity, Meteorologische Zeitschrift, 23, 15–32, https://doi.org/10.1127/0941-2948/2014/0560, 2014.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes, Scientific Data, 2, 1–21, https://doi.org/10.1038/sdata.2015.66, 2015.

García, E.: Modification to Koppen's climate classification system, Institute of Geography, UNAM, 5 edn., http://www.igeograf.unam.mx/sigg/utilidades/docs/pdfs/publicaciones/geo_siglo21/serie_lib/modific_al_sis.pdf, 2004.

Goovaerts, P.: Geostatistics for Natural Resources and Evaluation, Oxford University Press, 1997.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. of Hydrology, 228, 113–129, 2000.

Herrera, S., Cardoso, R. M., Soares, P. M., Espírito-Santo, F., Viterbo, P., and Gutiérrez, J. M.: Iberia01: A new gridded dataset of daily precipitation and temperatures over Iberia, Earth System Science Data, 11, 1947–1956, https://doi.org/10.5194/ESSD-11-1947-2019, 2019.

685 Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, International Journal of Climatology, 25, 1965–1978, https://doi.org/10.1002/joc.1276, 2005.

Hollis, D., McCarthy, M., Kendon, M., Legg, T., and Simpson, I.: HadUK-Grid—A new UK dataset of gridded climate observations, Geoscience Data Journal, 6, 151–159, https://doi.org/10.1002/gdj3.78, 2019.

Hopkinson, R. F., McKenney, D. W., Milewska, E. J., Hutchinson, M. F., Papadopol, P., and Vincent, L. A.: Impact of Aligning Climatological Day on Gridding Daily Maximum–Minimum Temperature and Precipitation over Canada, Journal of Applied Meteorology 690 and Climatology, 50, 1654–1665, https://doi.org/10.1175/2011JAMC2684.1, 2011.

Huntington, T. G., Weiskel, P. K., Wolock, D. M., and McCabe, G. J.: A new indicator framework for quantifying the intensity of the terrestrial water cycle, Journal of Hydrology, 559, 361–372, https://doi.org/10.1016/j.jhydrol.2018.02.048, 2018.

Hutchinson, M.: ANUSPLIN version 4.3, Tech. rep., Australian National University, 2007.

Isaaks, E. H. and Srivastava, M.: Applied geostatistics, Oxford University Press, 1989.

695 Jauregui, E.: Climatology of landfalling hurricanes and tropical storms in Mexico, Atmósfera, 16, 193–204, http://revistas.unam.mx/index.php/atm/article/view/8516/7986, 2003.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, 233–241, https://doi.org/10.1029/1998WR900018, 1999.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D. R., and Brekke, L.: A spatially 700 comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013, Scientific Data, 2, 150 042, https://doi.org/10.1038/sdata.2015.42, 2015.

Mankin, K. R., Mehan, S., Green, T. R., and Barnard, D. M.: Review of gridded climate products and their use in hydrological analyses reveals overlaps, gaps, and the need for a more objective approach to selecting model forcing datasets, https://doi.org/10.5194/hess-29-85-2025, 2025.

705 Meyer-Arendt, K. J.: Hurricane Gilbert: The storm of the century, GeoJournal, 23, 323–325, https://doi.org/10.1007/BF00193605, 1991.

Page, T., Beven, K. J., Hankin, B., and Chappell, N. A.: Interpolation of rainfall observations during extreme rainfall events in complex mountainous terrain, Hydrological Processes, 36, 1–21, https://doi.org/10.1002/hyp.14758, 2022.

Patterson, T. and Jenny, B.: The Development and Rationale of Cross-blended Hypsometric Tints, Cartographic Perspectives, pp. 31–46, 2011.

710 Pebesma, E.: GSTAT user's manual, http://www.gstat.org/gstat.pdf, 2014.

Pebesma, E. and Benedikt, G.: GSTAT: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation., https://github.com/r-spatial/gstat/, 2023.

Pebesma, E. and Bivand, R.: Spatial Data Science, Chapman and Hall/CRC, https://doi.org/10.1201/9780429459016, 2023.

Pedrozo-Acuña, A., Breña-Naranjo, J. A., and Domínguez-Mora, R.: The hydrological setting of the 2013 floods in Mexico, Weather, 69, 715 295–302, https://doi.org/10.1002/wea.2355, 2014.

Perdigón-Morales, J., Romero-Centeno, R., Pérez, P. O., and Barrett, B. S.: The midsummer drought in Mexico: perspectives on duration and intensity from the CHIRPS precipitation database, International Journal of Climatology, 38, 2174–2186, https://doi.org/10.1002/joc.5322, 2018.

Perez-Navarro, M. A., Broennimann, O., Esteve, M. A., Moya-Perez, J. M., Carreño, M. F., Guisan, A., and Lloret, F.: Temporal variability is key to modelling the climatic niche, Diversity and Distributions, 27, 473–484, https://doi.org/10.1111/ddi.13207, 2021.

Ramírez-Arce, D., Ochoa-Ochoa, L., Lira-Noriega, A., and Martorell, C.: Reptile Diversity Patterns Under Climate and Land Use Change Scenarios in a Subtropical Montane Landscape in Mexico, Journal of Biogeography, https://doi.org/10.1111/jbi.15017, 2024.

Razafimaharo, C., Krähenmann, S., Höpp, S., Rauthe, M., and Deutschländer, T.: New high-resolution gridded dataset of daily mean, minimum, and maximum temperature and relative humidity for Central Europe (HYRAS), Theoretical and Applied Climatology, 142, 1531–1553, https://doi.org/10.1007/s00704-020-03388-w, 2020.

Rosengaus-Moshinsky, M., Arreguín-Cortés, F. I., Korenfeld-Federman, D., and Rubio-Gutiérrez, H.: Visión panorámica de las precipitaciones pluviales combinadas por los efectos de los ciclones tropicales Ingrid-Manuel, Tecnologia y Ciencias del Agua, 7, 73–92, 2016.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theoretical and Applied Climatology, 115, 15–40, https://doi.org/10.1007/s00704-013-0860-x, 2014.

Sekulić, A., Kilibarda, M., Protić, D., and Bajat, B.: A high-resolution daily gridded meteorological dataset for Serbia made by Random Forest Spatial Interpolation, Scientific Data, 8, 1–12, https://doi.org/10.1038/s41597-021-00901-2, 2021.

Serrano-Notivoli, R. and Tejedor, E.: From rain to data: A review of the creation of monthly and daily station-based gridded precipitation datasets, https://doi.org/10.1002/wat2.1555, 2021.

Serrano-Notivoli, R., Beguería, S., Ángel Saz, M., Longares, L. A., and de Luis, M.: SPREAD: a high-resolution daily gridded precipitation dataset for Spain – an extreme events frequency and intensity overview, Earth System Science Data, 9, 721–738, https://doi.org/10.5194/essd-9-721-2017, 2017.

Shepard, D. S.: Computer mapping: the SYMAP interpolation algorithm, pp. 133–45, D. Reidel Publishing Company, 1984.

Sáenz-Romero, C., Rehfeldt, G. E., Crookston, N. L., Duval, P., St-Amant, R., Beaulieu, J., Richardson, B. A., and Bryce, B.: Spline models of contemporary , 2030 , 2060 and 2090 climates for Mexico and their use in understanding climate-change impacts on the vegetation, pp. 595–623, https://doi.org/10.1007/s10584-009-9753-5, 2010.

Tadić, M. P.: Gridded Croatian climatology for 1961-1990, Theoretical and Applied Climatology, 102, 87–103, https://doi.org/10.1007/s00704-009-0237-3, 2010.

Tang, G., Clark, M. P., Newman, A. J., Wood, A. W., Papalexiou, S. M., Vionnet, V., and Whitfield, P. H.: SCDNA: A serially complete precipitation and temperature dataset for North America from 1979 to 2018, https://doi.org/10.5194/essd-12-2381-2020, 2020.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., and Wilson, B. E.: Gridded daily weather data for North America with comprehensive uncertainty quantification, Scientific Data, 8, 190, https://doi.org/10.1038/s41597-021-00973-0, 2021.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., and Bairlein, F.: Ecological responses to recent climate change, Nature, 416, 389–395, https://doi.org/10.1038/416389a, 2002.

Wehner, M., Easterling, D. R., Lawrimore, J. H., Heim, R. R., Vose, R. S., and Santer, B. D.: Projections of Future Drought in the Continental United States and Mexico, Journal of Hydrometeorology, 12, 1359–1377, https://doi.org/10.1175/2011jhm1351.1, 2011.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Research, 30, 79–82, https://doi.org/10.3354/cr030079, 2005.

Willmott, C. J. and Matsuura, K.: On the use of dimensioned measures of error to evaluate the performance of spatial interpolators, International Journal of Geographical Information Science, 20, 89–102, https://doi.org/10.1080/13658810500286976, 2006.

Willmott, C. J., Robeson, S. M., Matsuura, K., and Ficklin, D. L.: Assessment of three dimensionless measures of model performance, Environmental Modelling and Software, 73, 167–174, https://doi.org/10.1016/j.envsoft.2015.08.012, 2015.

WMO: WMO Guidelines on the Calculation of Climate Normals, WMO-No. 1203, p. 29, https://library.wmo.int/doc_num.php?explnum_id=4166, 2017.

Xavier, A. C., Scanlon, B. R., King, C. W., and Alves, A. I.: New improved Brazilian daily weather gridded data (1961–2020), International Journal of Climatology, 42, 8390–8404, https://doi.org/10.1002/joc.7731, 2022.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., and Kitoh, A.: APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Bulletin of the American Meteorological Society, 93, 1401–1415, https://doi.org/10.1175/BAMS-D-11-00122.1, 2012.

Zambrano-Bigiarini, M.: {hydroTSM}: Time Series Management, Analysis and Interpolation for Hydrological Modelling, https://doi.org/10.5281/zenodo.839565, 2024.