Return Levels of Temperature Extremes in Southern Pakistan

Maida Zahid[1], Richard Blender[1] ,Valerio Lucarini[1,2] and Maria Caterina Bramati[3]

1. Meteorological Institute, University of Hamburg, Hamburg Germany
2. Department of Mathematics and Statistics, University of Reading, Reading, UK
3. Department of Statistical Science, Cornell University, New York, United States

Correspondence to: Maida Zahid (maida.zahid@uni-hamburg.de)

Abstract. Southern Pakistan (Sindh) is one of the hottest regions in the world and is highly vulnerable to temperature extremes. In order to improve rural and urban planning, it is useful to gather information about the recurrence of temperature extremes. In this work, return levels of the daily maximum temperature T_{max} are estimated, as well as the daily maximum wet-bulb temperature TW_{max} extremes. We adopt the Peaks over threshold (POT) method, which has not yet been used for similar studies in this region. Two main datasets are analyzed: temperatures observed in nine meteorological stations in southern Pakistan from 1980 to 2013, and the ERA Interim (ECMWF re-analysis) data for the nearest corresponding locations. The analysis provides the 2, 5, 10, 25, 50 and 100-year Return Levels (RLs) of temperature extremes. The 90% quantile is found to be a suitable threshold for all stations. We find that the RLs of the observed T_{max} are above 50°C in northern stations, and above 45°C in the southern stations. The RLs of the observed TW_{max} exceed 35°C in the region, which is considered as a limit of survivability. The RLs estimated from the ERA Interim data are lower by 3°C to 5°C than the RLs assessed for the nine meteorological stations. A simple bias correction applied to ERA Interim data improves the RLs remarkably, yet discrepancies are still present. The results have potential implications for the risk assessment of extreme temperatures in Sindh.

Key words

Extreme temperature, return levels, peak over threshold, Generalized Pareto Distribution, declustering.

1 Introduction

Extreme maximum temperature events have received much attention in recent years, because of the associated dangerous impact on the increased risk of mortality (IPCC, 2012). Additionally, climate change scenarios suggest that in most regions the probability of occurrence of extremely high temperature is very likely to increase in the future (Sheridan and Allen, 2015). An example of the potential impact of raising maximum temperatures is the recent heat wave in southern Pakistan (Sindh), which occurred between June 17th and June 24th 2015 and broke all the records with a death toll of 1400 people, and over 14000 people hospitalized. The temperatures in different cities of the Sindh region were in the range of 45°C - 49°C during the event (Imtiaz and Rehman, 2015). Karachi had the highest number of fatalities (1200 people approximately). The Pakistan Meteorological department issued a technical report stating a very high heat index (measuring the heat stress on humans due to high temperature and relative humidity) during this heat wave (Chaudhry et al., 2015).

In summer, Sindh becomes very hot and with the arrival of a monsoon the humidity increases in the region

44 (Chaudhry and Rasul, 2004). The extremely hot and humid conditions can have lethal effects, and can impact
45 the overall human habitability of a region (Pal and Eltahir 2015). The human body generally maintains the
46 temperature around 37°C. However, the human skin regulates at or below 35°C to release heat (Sherwood
47 and Huber, 2010). Under combined high temperatures and high levels of moisture content in the atmosphere,
48 the human body cannot maintain the skin temperature below 35°C and can develop ailments like
49 hyperthermia, heat strokes and cardiovascular problems. Hyperthermia is a condition where extremely high
50 body temperature is reached, resulting from the inability of the body to get rid of the excess heat.
51 Hyperthermia can occur even in the fittest human beings, if exposed for at least six hours to an environment
52 where wet-bulb temperature is greater than 35°C.
53
54 This study devotes special attention to Sindh (23.5° N – 28.5° N and 66.5°E - 71.1°E) because of its exposure
55 to the intense temperature extremes recently (Zahid and Rasul, 2012). It is bounded on the west by the
56 Kirthar Mountains, to the north by the Punjab plains, on the east by the Thar desert and to the south by the
57 Arabian Sea (Indian Ocean), while in the center there is a fertile land around the Indus river. Cotton, wheat,
58 sugar cane, rice, wheat and gram crops are cultivated near banks of the Indus River (Chaudhry and Rasul,
59 2004). Cotton is the cash crop of the country. High population density, limited resources, poor infrastructure
60 and high dependence of the local agriculture on climatic factors, mark this region as highly vulnerable to the
61 impacts of climate change. The Intergovernmental Panel on Climate Change (IPCC) scenarios estimates for
62 this region an increase in the temperature of the order of 4°C by the end of 2100. This may significantly
63 reduce crop yields, and cause huge economic losses to the country (Islam et al., 2009; Rasul et al., 2012;
64 IPCC, 2012, 2014). Furthermore, the risks of heat strokes, cardiac arrest, high fever, diarrhea, cholera and
65 vector borne diseases might increase.
66
67 Extreme value theory (EVT) provides the statistical basis for increasingly widespread quantitative
68 investigations of extremes in climate studies (Coles, 2001, Zhang et al., 2004; Brown et al., 2008; Faranda et
69 al., 2011; Acero et al., 2014). The peaks over threshold (POT) approach aims at describing the distribution of
70 the exceedances of the stochastic variable of interest above a threshold. Under very general conditions, the
71 exceedances are asymptotically distributed according to the Generalized Pareto Distribution (GPD). GPD has
72 remarkable properties of universality when the asymptotic behavior is considered (Lucarini et al., 2016),
73 while one can expect that the threshold level above which the asymptotic behavior is achieved depends on
74 the characteristics of the analyzed time series. In particular, when looking at spatial fields, the threshold level
75 depends on the geographical location.
76
77 In this study, we have chosen to analyze the temperature extremes in the Sindh region taking the point of
78 view of threshold exceedances associated to the GPD family of distributions, because the statistical inference
79 provided by the POT method provides a more efficient use of data and has better properties of convergence
80 when finite datasets are considered with respect to alternative methods for the analysis of extremes, such as
81 the block maxima method, which is used to fit the observed data to the generalized extreme value (GEV)
82 distribution (Lucarini et al., 2016). Additionally, we are here interested in investigating the actual tails of the
83 distributions and not the   statistics of e.g. yearly maxima, the POT approach is indeed more appropriate.
84 While the POT method has been applied for studying temperature extremes in different regions of the world

(Burgueño et al., 2002; Nogaj et al., 2006; Coelho et al., 2007; Ghil et al., 2011), to our knowledge, it has never been used to analyze the statistics of temperature extremes in Sindh. Thanks to the properties of universality of the GPD distribution (Lucarini et al. 2016), the POT approach can in principle provide reliable estimates of return periods and the return levels also for time ranges longer than what is actually observed. This information and this predictive power can be beneficial for policy makers and other stakeholders. Since, it is exactly the kind of information planners need when, e.g., designing infrastructures that are deemed to last a very long time. Note that commonly used, more empirical approaches to the study of extremes, as those more used for assessing the 'moderate extremes' (IPCC, 2012), do not have any property of universality and might have weak predictive power.

It is useful to consider two indicators of extremely hot conditions: (1) temperature extremes T_{max}, and (2) Wet-bulb temperature extremes TW_{max}. Therefore, we estimate the return levels of T_{max} and TW_{max} over different return periods during summer (May-September) in Sindh. We apply the POT method on the observational data of the nine weather stations provided by Pakistan Meteorological Department, and the ERA Interim re-analysis data of European Center for Medium range Weather Forecast (ECMWF) model for the corresponding grid points from 1980 to 2013. ERA Interim re-analysis data are generally very good at replicating also trends in temperature percentile (Cornes and Jones, 2013). Nonetheless, it is in principle not obvious that ERA Interim data can simulate well meteorological extremes, as reanalysis are constructed in such a way that typical conditions are well reproduced. This is why we look at how well ERA Interim data performs in the target area against observations. If the ERA Interim dataset characterizes well the extremes, it could be an option for the regions within Sindh where no observational data is available. Furthermore, a standard bias correction is applied on the ERA Interim data to assess whether removing the bias in the bulk of the statistics improves substantially representation of the return levels of extremes. Given the shortness of the datasets, as we will show later, it is appropriate to analyze the extremes without taking into considerations possible long-term trends (Frei and Schär, 2001); see also the discussion in Felici et al. (2007). The provision of POT-based information on stationary extremes is already quite relevant in terms of impacts for the public and private sector as it fills a big data gap in Sindh. A possibility for investigating time dependency in the temperature extremes comes for considering the centennial NCEP reanalysis (Compo et al., 2011) and using suitable bias correction procedures. Such an analysis is not performed at this stage as we focus on observational data.

The paper is organized as follows. In Section 2 we present the datasets we study and the statistical methods we use for assessing the properties of extremes. In Section 3 we show and discuss the main results. In Section 4 we make a summary of the main findings and present our conclusions and perspectives for future investigations.

2. Data and Methodology

2.1 Meteorological station data

The daily maximum temperature and relative humidity data recorded at nine meteorological stations in Sindh from 1980 to 2013 are provided by the Pakistan Meteorological Department (see Table 1). We select nine

125 stations, which contain a negligible amount of missing values after 1980, and are suitable for the POT

126 analysis (Fig. 1). An additional criterion is that only those stations are chosen where no changes occurred in

127 measuring instruments during the last 33 years (Brunetti et al., 2006). None of the station data shows gaps

128 with duration longer than two days, which are treated by replacing the missing value with the average of the

129 two previous values.

130

131 The temperature data are discretized unevenly with intervals up to 1 degree Celsius. Deidda and Puliga

132 (2006) proposed a Monte Carlo approach for addressing this issue. They showed that finite resolution in

133 precipitation data affects the convergence of parameter estimation in the extreme value analysis. They

134 suggested generating many synthetic datasets by adding numerical noise to the original data, and then

135 providing the best estimate of the parameters of the extreme value distributions by averaging over all the best

136 fits obtained in each synthetic dataset. Following their suggestion, we produce high-resolution data to

137 compensate the effect of discretization and thus to improve the convergence of the estimator. In order to

138 convert the temperature readings to higher resolution, we add a uniform random variable in the interval [-0.5,

139 0.5]. The main property of this noise is that $round(T+r) = T$, where T is the temperature with 1-degree

140 resolution and '$round$' is the numerical function, which maps the interval [T-0.5, T+0.5] to T. Thus, adding

141 the noise does not perturb the information content of the observations. This procedure is applied to all

142 temperature data, irrespective of the actual resolution, and replicated 100 times using a Monte Carlo

143 approach. For each synthetic dataset, we perform the statistical best fit described later in the paper and then

144 average the results. We check the influence of this noise parameterization and find no significant bias in the

145 return level estimates. The advantage of adding a noise is to avoid the spurious statistical effects associated

146 to the presence of discrete values assigned to the temperature readings. Using the described bootstrap method

147 we reduce such problem without biasing the data.

148

149 **2.2 ERA Interim re-analysis data**

150

151 The gridded daily maximum temperature and relative humidity data of ERA Interim re-analysis is obtained

152 from the ECMWF Public Datasets web interface (http://apps.ecmwf.int/datasets/). The ERA Interim is

153 generated by the European Center for Medium range Weather Forecast (ECMWF) model with resolution

154 0.75° × 0.75° (Dee et al., 2011). The gridded data are then extracted at the closest grid points of all stations,

155 for the period 1980-2013 (Fig. 1). The latitude and longitude of the ERA Interim stations are displayed in

156 Table 1.

157

158 The extreme temperatures analysis is restricted to the summer season (May-September) over a period of 33

159 years. We have tested the datasets by applying the Mann-Kendall test; the results show that trends are not

160 significant in such a short time interval. One of the main requirements for performing the POT analysis is

161 assuming the stationarity of the time series. Therefore, as in Bramati et al. (2014), the Augmented Dickey

162 Fuller (ADF) test of stationarity is performed on all time series (Dickey and Fuller, 1979). In all cases we

163 find no sign of long-term correlations in the data. Short-term correlations (daily time scale) typically lead to

164   clusters of extreme values and are studied by computing the extremal index θ in all time series and treated
165   using the associated standard declustering technique (see more details in Section 2.4).
166

167   **2.3 Wet-bulb temperature calculations**
168

169   The wet-bulb temperature measures the heat stress better than other existing heat indices, because it
170   establishes the clear thermodynamic limit on heat transfer that cannot be overcome by adaptations like
171   clothing, activity and acclimatization (Pal and Eltahir, 2015; Sherwood and Huber, 2010). Here, we use an
172   empirical equation developed by Stull (2011) to measure the wet-bulb temperature.
173

174    $$TW = T \text{ atan}(\alpha_1\sqrt{RH + \alpha_2}) + \text{atan}(T + RH) - \text{atan}(RH + \alpha_3) + \alpha_4(RH)^{\frac{3}{2}}\text{atan}(\alpha_5 RH) - \alpha_6$$
175
176   (1)
177
178

179   where TW is the wet-bulb temperature [°C], T is the temperature [°C], and RH is the relative humidity [%].
180   This relationship is based on an empirical fit, as in Stull (2011), where the coefficient values are α_1 =
181   0.151977, α_2 = 8.313659, α_3 = -1.676331, α_4 = 0.00391838, α_5 = 0.023101, and α_6 = 4.686035. Eq. (1)
182   covers a wide range of relative humidity and air temperatures with an accuracy of 0.3°C.
183

184   **2.4 Peaks over Threshold**
185

186   In order to determine the return levels of extreme maximum temperatures and maximum wet-bulb
187   temperatures, the peaks over threshold (POT) approach is applied to the data obtained from the
188   meteorological stations in Sindh, and from the ERA Interim archive.
189

190   Multi-occurrence is an important characteristic of extreme climatic events and is referred to as clustering.
191   Clusters are consecutive occurrences of above threshold events. It is important to post process the clustered
192   extremes in order to take into account the assumption of weak short time correlation between extreme events,
193   which is crucial for our statistical analysis. We have treated the clusters using the concept of Extremal Index
194   (EI) (see Newell, 1964, Loynes, 1965, O'Brien, 1974, Leadbetter, 1983, Smith, 1989, Davison and Smith,
195   1990). The Extremal Index θ measures the degree of clustering of extremes. It ranges between 0 and 1, (θ = 0
196   means strong clustering and dependence, θ = 1 absence of clusters and independence). Leadbetter (1983)
197   interprets 1/θ as the mean number of exceedances in a cluster.
198

199   The extremal index θ can be estimated in two different ways. Here, we apply the 'intervals estimator'
200   automatic declustering by Ferro and Segers (2003). A positive aspect of this method is that it avoids the
201   subjective choice of cluster parameters. The main ingredient is the use of an asymptotic result for the times
202   between threshold exceedances. The exceedance times are split into two types, a set of vanishing intra-
203   exceedance times within the clusters, and an exponentially distributed set of inter-exceedance times between
204   clusters. The method is iterative, starting with largest return times and stops when a limit for the inter-

205   exceedance times is reached. The standard errors of the estimated parameters is obtained by a bootstrap

206   procedure. In this study, once we select appropriate value for the threshold (see below) the extremal index

207   value is ≤ 0.5 in all the considered time series. Therefore, it is necessary to decluster the extremes by

208   choosing the largest event in each cluster, before fitting it to the GPD.

209

210   As mentioned before, we use as statistical model for the exceedances over threshold the Generalized Pareto

211   Distribution (GPD), which is characterized by two parameters, the shape ξ and the scale σ. The GPD for

212   exceedances $x - u$ of a random variable x reads as

213

$$G(x) = 1 - \left[1 + \xi\left(\frac{x - u}{\sigma}\right)\right]^{-\frac{1}{\xi}} \qquad (x > u, \xi \neq 0) \qquad (2)$$

214
215

216   where u is the threshold. The shape parameter ξ determines the tail behavior while the scale parameter σ

217   measures the variability. For a negative shape parameter, ξ < 0, the distribution is bounded (Weibull

218   distribution), for vanishing shape parameter, ξ = 0, the distribution is exponential, and for a positive shape

219   parameter, ξ > 0, the distribution has no upper bound (Pareto distribution).

220

221   In particular, for a negative shape parameters ξ <0 the GPD has the upper bound

222

$$A_{max} = u - \sigma/\xi \qquad (3)$$

223

$$G(x) = 0 \qquad (x > A_{max}, \xi < 0)$$

224
225

226   where A_{max} is an absolute maximum (Lucarini et al., 2014). In general, the best estimate for the two

227   parameters shape ξ and scale σ depend on the threshold u (Coles, 2001). The choice of the optimal threshold

228   for performing statistical inference from a time series is crucial. Choosing a very large value for u reduces

229   the number of exceedances to a few values, inflating the variance of the estimators, so that the analysis is

230   unlikely to yield any useful results. On the other hand, choosing a too small value for u would violate the

231   asymptotic nature of the model, with a possible biased estimation and wrong model selection (Coles, 2001),

232   see details later in Section 3.1. The shape ξ, the scale σ and the return levels are estimated using the

233   Maximum Likelihood Estimator (MLE) using the R software (R Development core team 2015), which also

234   provides an estimate of the standard error of the estimates.

235

236   Additionally, we wish to investigate the N - years return levels x_N, which are exceeded on the time scale of

237   N years (Coles, 2001) and can be expressed as

238
239

$$x_N = u + \frac{\sigma}{\xi}\left[(Nn_y\zeta_u)^\xi - 1\right] \qquad (4)$$

240
241

242   where N represents the return period in years, n_y is the number of observations per year , ζ_u is the probability

243   of an individual observation exceeding the threshold u, the shape parameter is ξ and the scale parameter is

244   σ.

2.5. Bias Correction Method

A simple bias correction is applied to each ERA Interim time series through a rescaling that adjust the first two moments (mean and variance) to the sample moments calculated for the corresponding observations. Therefore, the bias correction is applied to the entire time series and it is not tailored to the extreme events only. The idea is to check whether by adjusting the properties of the bulk of the statistics we improve the skill of the ERA Interim dataset considerably in describing extreme events. The bias corrected ERA Interim time series x is expressed as

$$x = \bar{z} + \frac{y_{ERA} - \bar{y}}{\sigma_y} \, \sigma_z$$

(5)

where y_{ERA} is the ERA Interim time series, $\bar{y}$ and σ_y its mean and standard deviation, whereas $\bar{z}$ and σ_z are the mean and standard deviation of the meteorological station temperatures. The properties of extremes are commonly assumed to be closely controlled by the first two moments of the underlying distribution - e.g. the IPCC (2012) relates changes in the properties of extremes to changes in the mean and in the standard deviation of the underlying distributions - EVT clarifies that, in fact, only a loose link exists between true extremes and the bulk of the events. Note that the proposed method of bias corrections has no impact on the estimates of the shape parameter, while it affects the scale and location parameters, thus impacting at any rate the return levels.

3. Results and Discussion

3.1 Threshold Selection

The threshold selection is the first step in a POT analysis. One needs to test whether the asymptotic regime is reached, i.e. whether one is choosing true extremes. It must be noted that EVT does not predict where (in terms of quantiles) one should expect the asymptotic regime to start. This can be investigated by checking whether the best fits of the shape parameter ξ and the modified scale parameter $\sigma^* = \sigma_u - \xi u$ are stable with respect to increases in the chosen value of u (Sacrrott and MacDonald, 2012).The optimal threshold u is selected as the lowest value where the two parameters are invariant in order to reach the asymptotic limit (Coles, 2001 and Furrer et al., 2010). This choice allows for having as many data as possible for performing the statistical inference, thus having lower variance for the estimators of the parameters. Fig. 2 shows the parameter stability plots of the T_{max} reading for Karachi, as an example to explain the threshold selection procedure.

In addition to diagnostic plots of the modified scale parameter σ^* and the shape parameter ξ, the mean residual life plot is used to select the appropriate threshold for the POT analysis (Davison and Smith, 1990). The idea is to select the lowest value of the threshold when the plot is approximately linear. In the case of the Karachi data for T_{max} , the plot appears to be linear and stable for u = 36°C, indicating u = 36 as the most

283 suitable threshold for the POT analysis (Fig. 3). We observe that the 90% quantile is an appropriate threshold

284 for all the station data, as well as the ERA interim datasets, and for both T_{max}, and TW_{max}.

285

3.2 GPD Fit

287

288 The goodness of fit is evaluated by Quantile-Quantile (Q-Q) plots and hypothesis testing. The Q-Q plot

289 analysis is performed for the stations observed, the ERA Interim, the bias corrected ERA Interim daily T_{max}

290 and TW_{max}. The Q-Q plots of the observed T_{max} show that the GPD fits well in most stations. However, in a

291 few stations like Jacobabad, Mohenjo-daro, Padidan and Chhor the empirical values show slight deviation

292 from the modeled values. In spite of minor deviations at some stations, still most of the exceedances are well

293 fitted by the model. The Q-Q plots of the observed TW_{max} also fits well to the model in all stations.

294

295 The Q-Q plots of the empirical ERA Interim T_{max} and TW_{max} data reveals substantial differences with

296 respect to the corresponding GPD fits. The empirical values of the higher quantiles are deviating from the

297 theoretical quantiles in all stations. However, if the higher quantiles are disregarded, then stations like

298 Jacobabad, Mohenjo-daro, Rohri, Padidan, Nawabshah, Chhor, and Badin fits very well with the model. The

299 Q-Q plots of the bias corrected ERA Interim T_{max}, and TW_{max} show better results than the ERA Interim. We

300 notice that the T_{max} of the ERA Interim and bias corrected ERA Interim fits better than the TW_{max} if the

301 highest quantiles are ignored, indicating the bias procedure is, as expected, unable to treat correctly the

302 statistics of the largest events.

303

304 In order to assess the goodness-of-fit, we apply the Kolmogorov-Smirnov (K-S) test and Anderson-Darling

305 (A-D) test to the data of meteorological stations, ERA Interim, bias corrected ERA Interim T_{max} and TW_{max}.

306 The p-values indicate a good performance of the fit procedure. Table 2 shows the results of the K-S and A-D

307 statistics of the T_{max} and TW_{max} in all the data sets.

308

3.3 Parameter Estimates

310

311 Here, we analyze the shape parameter ξ , the scale parameter σ, and threshold u for all considered datasets.

312 The standard errors of the shape ξ and the scale σ parameters are given in Table 3. The spatial distribution of

313 the shape parameter ξ and the scale parameter σ of the GPD in Sindh are shown in Fig. 4. The shape

314 parameters ξ are negative in all datasets at all stations. This is hardly surprising, as meteorological and

315 physical processes make sure that the temperature cannot grow locally without control. One finds a certain

316 degree of variability across stations in the estimated value of the shape parameter. In the case of the observed

317 T_{max} one obtains for ξ estimates ranging between -0.418 and -0.223, while for TW_{max} the range is between -

318 0.323 and -0.177, so that values slightly closer to zero are found, thus allowing for larger excursions towards

319 very high values with respect to the case of the extremes of the actual temperature. When looking at the bias

320 corrected ERA Interim data, the range of values for the shape parameter of T_{max} (TW_{max}) is between -0.305 to

321 -0.002 (-0.18 and -0.01). While there is a good match in the spatial patterns of the estimates for the

322 observative vs ERA Interim datasets, the presence of values much closer to zero in the second case suggests

323   the presence of some inadequacies in the representation of extremes in the reanalysis. This is not entirely

324   unexpected, as reanalysis are constructed in such a way that typical conditions are well reproduced. Note that

325   our simple bias correction procedure, while not impacting the estimates of the shape parameters, allows for

326   improving the estimates of the return levels, as discussed below.

327

328   The scale parameters σ measures the variability of the GPD distributions. The highest values of the scale

329   parameters σ of T_{max} and TW_{max} are observed at stations such as Jacobabad, Padidan, Karachi, Hyderabad

330   and Chhor in all datasets. This indicates that the variability of temperature extremes is higher at these

331   stations, and one can expect higher return values of T_{max} and TW_{max} here having similar shape parameter

332   and same threshold according to Eq. (4). The scale parameters σ of the observed T_{max} range from 2.08 to

333   2.76, and the TW_{max} are in 1.86 to 2.76. In the ERA Interim analysis, the scale parameter σ of T_{max} is

334   between 1.00 - 1.95, and TW_{max} in 0.74 -1.75. We observe a difference in the scale parameters of both the

335   observed, ERA Interim T_{max} and TW_{max}. We find that, unsurprisingly, the scale parameters of the bias

336   corrected ERA Interim data are much closer to those estimated for T_{max} and TW_{max} using the station data. In

337   the bias corrected ERA Interim T_{max} the scale parameters σ are in 1.50 - 2.75, while for TW_{max} are in a range

338   1.40 – 2.40 (Fig. 4). All the temperature scale parameters are in degree Celsius.

339

340   **3.4 Absolute Maxima**

341

342   Once the shape parameters ξ, the scale parameters σ, and the thresholds u are determined, it is possible to

343   compute the theoretical absolute maxima using Eq. (3) (Section 2.4). Theoretical absolute maxima can be

344   compared with the observed ones for each station to better understand whether our fits are in agreement with

345   the observed data. The daily maximum temperature T_{max} and the maximum wet-bulb temperature TW_{max}

346   (station data, the ERA Interim, and the bias corrected ERA Interim) have negative shape parameters ξ at all

347   stations. This means that according to Eq. (2) in section 2.4, the probability distribution function (pdf) is

348   bounded by the maximum values. These maximum values are the theoretical upper limits predicted by the

349   GPD fit. The analysis shows that the observed absolute maxima T_{max} and TW_{max} at all stations of the three

350   data sets are below the theoretical absolute maximum, as expected (Fig. 5). This gives us confidence on the

351   quality of our fit. The following piece of information can also be derived: assume that one observes in the

352   future an extreme event larger than the maximum inferred in the present dataset; this may suggest some non-

353   stationarity in the most recent portion of the dataset.

354

355   **3.5 Return Levels**

356

357   The return levels (RLs) are computed considering various return periods (2, 5, 10, 20, 50, 100-year). As

358   remarked above, using a statistical approach based on the universality of EVT, we are able to extrapolate the

359   results for time horizons longer than the one for which observations are taken. Clearly, uncertainties grow

360   when longer time horizons are considered. The return level plots of the stations observed, the ERA Interim,

361   the bias corrected ERA Interim daily maximum temperature T_{max} and daily maximum wet–bulb temperature

362   TW_{max} are displayed in Fig. 6 and Fig. 7. The values of the RLs follow the north-south gradient of the

climatic mean temperatures. The northern part of the Sindh (Jacobabad, Mohenjo-daro, Rohri, Padidan, and Nawabshah) are hotter than the southern part (Hyderabad, Chhor, Karachi, and Badin).

The 2, 5, 10, 20, 50, 100-year RLs estimated in Sindh for station observed T_{max} at time reach over 50°C in Jacobabad, Mohenjo-daro, Padidan, Nawabshah, and over 45°C in Rohri, Hyderabad, Chhor, Karachi, Badin. The corresponding ERA Interim T_{max} return levels are at least 3°C to 5°C lower in all stations, while having correct representation of the geographical variability of the field. As example, the RLs of 42°C at Badin has a 3-year return period in the observations T_{max}, but a 30-year return period in ERA Interim (Fig. 6).

The RLs of TW_{max} are above 35°C in all meteorological stations. As for the ERA Interim, the RLs of TW_{max} are greater than 30°C for all the stations except Karachi, which has RLs less than 30°C. Here, we see again that the RLs of the ERA Interim TW_{max} are lower than the RLs of station TW_{max}. Going again to the Badin stations, the 4-year return period observed for TW_{max} is 38°C, while the ERA Interim dataset show the same RL in a 15-year return period (Fig. 7).

The bias corrected ERA Interim T_{max} and TW_{max}, show some improvements in the RLs at all stations. When looking at the Nawabshah, Hyderabad, Karachi, and Badin stations, the RLs agree with those obtained from the station data in the range 5-100 years, while disagreements exist in the range 2-5 years. In the rest of the stations, the bias corrected data RLs are closer to those of the station data, yet not statistically compatible with them. When looking at the wet-bulb temperature TW_{max} analysis, the RLs of the bias corrected ERA Interim show some overlap with those derived from station observations in Mohenjo-daro, Hyderabad, Chhor, and while no overlap is found in the other stations. One understands that the proposed simple bias correction methods improves the quality of the representation of extremes by ERA Interim, but many discrepancies remain (Fig. 6 and Fig. 7).

We also plot the station and bias corrected ERA Interim T_{max}, and TW_{max} return levels spatially for the 5, 10, 25 and 50-year return periods (Fig. 8 and Fig. 9), as a detailed spatial overview of the temperature extremes in Sindh might be of interest to the policy makers. The spatial return levels of the station and bias corrected ERA Interim T$_{max}$ shows differences in temperature; the hottest stations have the highest return levels. We notice that for Jacobabad, Mohenjo-daro, Padidan, Nawabshah the return levels are between 50°C-53.6°C and for Rohri, Hyderabad, Chhor, Karachi, and Badin are between 45°C - 50°C in 5 to 50 years return period (Fig. 8). These extreme temperatures can impact the yields because crops are very sensitive to temperature variations, and even a rise of one degree Celsius can cause detrimental changes in the phenological stages of the crops (Hatfield and Preuger, 2015). Every crop has a certain limit to tolerate the temperature. When temperature exceeds this limit, the crop yield is drastically reduced. Abbas et al., (2017) notices 33% decrease in major crops of Sindh due to warmer and drier weather. Karachi and Badin are expected to decrease rice cultivation, hatching of fisheries, and mangroves forest surrounding these cities. Furthermore, temperature extremes can have serious threat to cotton, wheat, and rice yields in Rohri and Mohenjo-daro areas due to increased crop water requirements.

403

404   In summer, the temperature and humidity increase to an extent that there are high chances of a rapid pests
405   spread in the crops. Temperature extremes not just directly impact the quantity and quality of grains, but can
406   also be a reason of urban flooding affecting the agriculture lands (Luo etal ., 2015). Sindh produces cotton,
407   wheat, rice, mango, banana, and dates, so a correct estimate of temperature extremes is very important.
408

409   The spatial return levels of station and bias corrected ERA Interim TW_{max} for the 5, 10, 25 and 50-year
410   return periods show highest return level greater than 35°C at all stations (Fig. 9). This is very serious for the
411   human health due to the working day hours of population in agriculture farms, building construction, and
412   port activities. Karachi and Badin being closet to the coast are at the highest risk of temperature extremes.
413   Thus, an immediate plan for adaptations is needed in Sindh to deal with such a hazard. The high values of
414   TW_{max} also indicate high levels of humidity in the region during summer, which is also proved by Kalim
415   and Shouting, (2012), and Freychet et al. (2015).

416   **4. Summary and Conclusion**
417

418   The main objective of this study is the assessment of the return levels of the extreme daily maximum
419   temperatures T_{max} and wet-bulb temperatures TW_{max} in southern Pakistan (Sindh). In addition, the
420   performance of the ERA Interim TW_{max} is compared to the weather station TW_{max} to assess its ability to
421   estimate temperature extremes in Sindh. Moreover, a simple bias correction is applied to the ERA Interim
422   data to see whether correcting the first two moments of its statistics helps in improving its performance in
423   representing temperature extremes.
424

425   The POT method is applied to the daily maximum temperature (T_{max}) and wet-bulb temperature (TW_{max}) data
426   of nine stations and to the corresponding nearest ERA Interim temperature data. After testing the asymptotic
427   statistical properties, the 90% quantile is found to be appropriate threshold choice for all datasets. The Q-Q
428   plots are used to assess the GPD fit, which results to be acceptable for both T_{max} and TW_{max} station data for
429   all three datasets. However, the bias corrected ERA Interim data shows improved GPD fits than the ERA
430   Interim data. The shape parameters ξ is in general negative at all stations. The scale parameters σ show high
431   values in Jacobabad, Padidan, Karachi, Hyderabad and Chhor indicating higher variability of temperature
432   extremes in these regions. The return levels (RLs) of T_{max} and TW_{max} are estimated for the 2, 5, 10, 25, 50,
433   100-year return periods in all datasets. The RLs of T_{max} estimated using the meteorological station
434   temperatures are greater than 50°C in Jacobabad, Mohenjo-daro, Padidan, Nawabshah, and greater than 45°C
435   in Rohri, Hyderabad, Chhor, Karachi and Badin. While the RLs of TW_{max} in station data are larger than 35°C
436   in the entire Sindh, when using ERA Interim temperatures, they are estimated as greater than 45°C in
437   Northern Sindh and greater than 40°C in southern Sindh.
438

439   Our results predict extremely high values of T_{max} and TW_{max} in the region. The T_{max} extremes contribute to
440   an increase rate of evaporation, which in turn may intensify the hydrological cycle causing precipitation
441   events and flooding (Cheema et al., 2012, Luo etal., 2015). Additionally, crops variety needs to be changed

under such a hot climate to avoid the risks of temperature extremes. The extremes of daily maximum wet-bulb temperature TW_{max} are estimated as above the human survivability threshold 35°C throughout the region, so the risk of hyperthermia is very high here. The most vulnerable people are those who are involved in the everyday outdoor activities like farming, fishing, building construction, athletes, elderly and infants can have heat strokes, dehydration etc. The human habitability in such a warm region is already at risk and one can expect that these issues will be worse in future climate conditions.

We found that the RLs of station and ERA interim showed differences between 3°C and 5°C for both shorter and longer return periods due to the minor variations in the shape and scale parameters. Although the ERA Interim dataset does not capture well the magnitude of the extremes, still it provides a good representation of their spatial fields. The biases between the station and the ERA Interim data are rather relevant when one wishes to address the impact of hot climatic extremes to human life and to active crop production in the region. It would be of primary importance to understand the physical reasons behind such inconsistencies, which makes it hard to use reasonably ERA without bias correction. Clearly, they might result either from a misrepresentation of local processes dominated by near surface processes (namely, heat and water fluxes), or from an inadequacy of the re-analysis in reproducing synoptic and sub-synoptic conditions responsible for extremely hot and humid conditions. This matter is surely worth investigating but is well beyond the scope of this paper.

We applied a simple bias correction i.e. adjusting the mean and standard deviation to ERA Interim T_{max} and TW_{max} data to check the improvements in return levels. We noticed that the bias corrected ERA Interim T_{max} and TW_{max} gives the return levels closer to the meteorological stations observed ones than the original ERA Interim return levels at all stations. Although the bias corrected ERA Interim shows a good correspondence with the meteorological station data, yet statistically significant differences remain in most cases. Therefore, one must use more advanced bias correction method for analyzing extremes precisely. We propose to repeat this analysis in GCMs (CMIP5, CMIP6) and RCMs (CORDEX) to study the properties of extremes. All models use re-analysis as input, and generate information of extremes, which involves biases that if not corrected, can lead to significant errors in prediction of present and future extremes. Therefore, in order to reduce the uncertainties in impact assessment, it is necessary to improve the re-analysis before using it in GCMs and RCMs.

The results have practical implications for assessing the risk of extreme temperature events in Sindh. All the results are placed in a web-tool SindheX [www.sindhex.org] that will be freely available online soon after the publication of this paper. The maps and graphs are prepared to guide the local administrations to prioritize the regions in terms of adaptations like preparation of baseline contingency plans for dealing with strong heat waves based on the current climatology. Such measures are not yet present in the territory and lead to many casualties each year. Our results will not only contributes to the regional planning, but can also be useful for the ongoing EU projects (SUCCESS, CSCCC), World Bank project (Sindh Resilience Project) and mega construction projects like China-Pakistan Economic Corridor (CPEC).

Acknowledgements

We would like to thank Climate KIC, for funding this research. This publication is a part of a Climate KIC project "Extreme Events in Pakistan: Physical processes and impacts of changing climate", which belongs to the adaptation services platform of the Climate KIC. Thanks to Pakistan Meteorological Department (PMD) and the European Center for Medium range Weather Forecast (ECMWF) for providing the datasets. The R development core team (2015) is acknowledged for providing statistics packages. We would like to thank the DFG Cluster of Excellence CliSAP for partially supporting this research activity. We would like to thank the reviewers, whose constructive criticisms have greatly helped to improve the quality of this paper.

References

Abbas, F., Rehman, I., Adrees, M., Ibrahim, M., Saleem, F., Ali, S., Rizwan, M. and Salik, M. R.: Prevailing trends of climatic extremes across Indus-Delta of Sindh-Pakistan, Theor. Appl. Climatol., doi:10.1007/s00704-016-2028-y, 2017.

Acero, F. J., García, J. A., Gallego, M. C., Parey, S. and Dacunha-Castelle, D.: Trends in summer extreme temperatures over the Iberian Peninsula using nonurban station data, J. Geophys. Res. Atmos.,119, 39-53, doi:10.1002/2013JD020590, 2014.

Bramati, M.C., Tarragoni, C., Davoli, L., Raffi, R., Extreme Rainfall in Coastal Metropolitan Areas of Central Italy: Rome and Pescara case studies. *Geografia Fisicae Dinamica Quaternaria*, 37, 3-13, 2014.

Brunetti, M., Maugeri, M., Monti, F. and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogeneized instrumental time series, J. Climatol., 26(3), 345–381, doi:10.1002/joc.1251, 2006.

Burgueño, A., Lana, X. and Serra, C.: Significant hot and cold events at the Fabra Observatory, Barcelona (NE Spain), Theor. Appl. Climatol., 71(3), 141-156, doi:10.1007/s007040200001, 2002.

Brunetti, M., Maugeri, M., Monti, F. and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogeneized instrumental time series, J. Climatol., 26(3), 345–381, doi:10.1002/joc.1251, 2006.

Brown, S. J., Caesar, J. and Ferro, C. A. T.: Global changes in extreme daily temperature since 1950, J. Geophys. Res. Atmos., 113, D05115 doi:10.1029/2006JD008091, 2008.

Chaudhry, Q.-U.-Z. and Rasul, G.: AGRO-CLIMATIC CLASSIFICATION OF PAKISTAN, Q. Sci. Vis., 9(12), 3–4, 2004.

Chaudhry, Q. Z., Rasul, G., Kamal, A., Ahmad Mangrio, M. and Mahmood, S.: Government of Pakistan Ministry of Climate Change Technical Report on Karachi Heat wave June 2015.

Cheema S.B., Zaman Q. & Rasul G. Persistent heavy downpour in desert areas of Pakistan in South Asian Monsoon 2011. Pak J Meteorol, 9, (17), 71–84, 2012.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer London, London., 2001.

Coelho, C. A. S., Ferro, C. A. T., Stephenson, D. B. and Steinskog, D. J.: Methods for Exploring Spatial and Temporal Variability of Extreme Events in Climate Data, J. Clim., 21(10), 2072–2092, doi:10.1175/2007JCLI1781.1, 2007.

Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, B.E. Gleason, R.S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R.I. Crouthamel, A.N. Grant, P.Y. Groisman, P.D. Jones, M. Kruk, A.C. Kruger, G.J. Marshall, M. Maugeri, H.Y. Mok, Ø. Nordli, T.F. Ross, R.M. Trigo, X.L. Wang, S.D. Woodruff, and S.J. Worley: The Twentieth Century Reanalysis Project. Quarterly J. Roy. Meteorol. Soc., 137, 1-28. http://dx.doi.org/10.1002/qj.776, 2011.

Cornes, R. C., and P. D. Jones, How well does the ERAInterim reanalysis replicate trends in extremes of surface temperature across Europe? J. Geophys. Res., 118, 10 262– 10 276, doi:10.1002/jgrd.50799, 2013.

Davison, A. C. and Smith, R. L.: Models for Exceedances over High Thresholds, J. R. Stat. Soc. Ser. B,

52(3), 393–442, doi:10.2307/2345667, 1990.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137: 553–597, doi:10.1002/qj.828, 2011.

Deidda, R. and Puliga, M.: Sensitivity of goodness-of-fit statistics to rainfall data rounding off, Phys. Chem. Earth, 31, 1240–1251, doi:10.1016/j.pce.2006.04.041, 2006.

Dickey, D. A. and Fuller, W. A.: Distribution of the Estimators for Autoregressive Time Series With a Unit Root, J. Am. Stat. Assoc., 74(366), 427, doi:10.2307/2286348, 1979.

Faranda, D., Lucarini, V., Turchetti, G. and Vaienti, S.: Numerical Convergence of the Block-Maxima Approach to the Generalized Extreme Value Distribution, J. Stat. Phys., doi:10.1007/s10955-011-0234-7, 2011.

Felici, M.; Lucarini, V.; Speranza, A.; Vitolo, R. Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part II: trend detection and assessment. Journal of the Atmospheric Science, v.64, p.2159-214-75, 2007.

Ferro, C. A. T. and Segers, J.: Inference for clusters of extreme values, J. R. Stat. Soc. B, 65(2), 545–556, 2003.

Frei, C., and C. Schär, Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. J. Climate, 14, 1568–1584, 2001.

Furrer, E., Katz, R., Walter, M. and Furrer, R.: Statistical modeling of hot spells and heat waves, Clim. Res., 43(3), 191–205, doi:10.3354/cr00924, 2010.

Freychet, N., Hsu, H.-H., Chia, C., and Wu, C.-H., Asian Summer Monsoon in CMIP5 Projections : A Link between the Change in Extreme Precipitation and Monsoon Dynamics. J. Climate, pages 1477–1493, 2015.

Ghil, M., Yiou, P., Hallegatte, S., Malamud, B. D., Naveau, P., Soloviev, A., Friederichs, P., Keilis-Borok, V., Kondrashov, D., Kossobokov, V., Mestre, O., Nicolis, C., Rust, H. W., Shebalin, P., Vrac, M., Witt, A. and Zaliapin, I.: Extreme events: Dynamics, statistics and prediction, Nonlinear Process. Geophys., 18, 295-350,doi:10.5194/npg, 2011.

Hatfield, J. L. and Prueger, J. H.: Temperature extremes: Effect on plant growth and development, Weather Clim. Extrem.,10, 4-10, doi:10.1016/j.wace.2015.08.001, 2015.

Imtiaz S, Rehman, ZU. 2015. June 25. Death Toll From Heat Wave in Karachi, Pakistan, Hits 1,000. *The New York Times* retrieved from http://www.nytimes.com/2015/06/26/world/asia/karachi-pakistan-heat-wave-deaths.html?_r=0

Islam, S. U., Rehman, N. and Sheikh, M. M.: Future change in the frequency of warm and cold spells over Pakistan simulated by the PRECIS regional climate model, Clim. Change, 94,35-45, doi:10.1007/s10584-009-9557-7, 2009.

Kalim, U. and Shouting, G. A. O., Moisture Transport over the Arabian Sea Associated with Summer Rainfall over Pakistan in 1994 and 2002. Advances in Atmospheric Sciences, 29(3):501–508, 2012.

Leadbetter, M. R., Extremes and local dependence in stationary sequences, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete,65, 291-306, doi:10.1007/BF00532484, 1983.

Loynes, R. M, Extreme Values in Uniformly Mixing Stationary Stochastic Processes, Ann. Math. Stat., 36(3), 993–999, doi:10.1214/aoms/1177700071,1965.

Luo, P., Apip, He, B., Duan, W., Takara, K. and Nover, D, Impact assessment of rainfall scenarios and land-use change on hydrologic response using synthetic Area IDF curves. J. Flood Risk Manage. doi:10.1111/jfr3.12164, 2015.

Luo, P., He, B., Takara, K., Xiong, Y.E., Nover, D., Duan, W.L., Fukushi, K., Historical assessment of chinese and japanese flood management policies and implications for managing future floods Environ. Sci. Policy, 48 (2015), pp. 265-277, 2015.

Lucarini, V., Faranda, D., Wouters, J. and Kuna, T.: Towards a General Theory of Extremes for Observables

of Chaotic Dynamical Systems., J. Stat. Phys., 154, 723–750, doi:10.1007/s10955-013-0914-6, 2014.

Lucarini, V., Faranda, D., .Freitas, A.C.M., Freitas, J.M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, S.: Extremes and Recurrence in Dynamical Systems, John Wiley & Sons Inc,305, ISBN: 978-1-118-63219-2

2016.

Newell, G. F.: Asymptotic Extremes for m-Dependent Random Variables, Ann. Math. Stat., 35(3), 1322–1325, doi:10.1214/aoms/1177703288, 1964.

Nogaj, M., Yiou, P., Parey, S., Malek, F. and Naveau, P.: Amplitude and frequency of temperature extremes over the North Atlantic region, 33, L10801, Geophys. Res. Lett., doi:10.1029/2005GL024251, 2006.

IPCC, 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp.

IPCC, Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.2014

Pakistan Meteorological Department, Monthly Climatic Normal of Pakistan (1980-2010), Climate Data Processing Centre (CDPC), Karachi, 2013.

Pal, J. S. and Eltahir, E. A. B.: Future temperature in southwest Asia projected to exceed a threshold for human adaptability, nature climate change, 6, 197-200, doi:10.1038/NCLIMATE2833, 2015.

Rasul, G., Mahmood, A., Sadiq, A. and Khan, S. I.: Vulnerability of the Indus Delta to Climate Change in Pakistan, Pakistan J. Meteorol., 8(16), 2012.

Rasul, G., Afzal, M., Zahid, M., Ahsan, S. and Bukhari, A.: Climate Change in Pakistan Focused on Sindh Province., Technical Report No. PMD-25, 2012.

R Development Core Team,. R, a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015.

Sheridan, S.C., Allen M. J., Changes in the Frequency and Intensity of Extreme Temperature Events and Human Health Concerns. Current Climate Change Reports 1(3): 155-162, doi:10.1007/s40641-015-0017-3, 2015.

Sherwood, S. C. and Huber, M., An adaptability limit to climate change due to heat stress, Proc. Natl. Acad. Sci. USA 107 9552–5, 2010.

Scarrott, C. and Macdonald, A.: A review of extreme valye threshold estimation and uncertainity quantification, Revstat – Stat. J., 10(1), 33–60, 2012.

Smith, R. L.: Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone, Stat. Sci., 4(4), 367–377, doi:10.1214/ss/1177012400, 1989.

Stull, R.: Wet-bulb temperature from relative humidity and air temperature, J. Appl. Meteorol. Climatol.,50, 2267-2269, doi:10.1175/JAMC-D-11-0143.1, 2011.

Tebaldi, C., Hayhoe, K., Arblaster, J. M. and Meehl, G. A.: Going to the extremes: An intercomparison of model-simulated historical and future changes in extreme events, Clim. Change, 79(3), 185–211, doi:10.1007/s10584-006-9051-4, 2006.

Zahid, M. and Rasul, G.: Rise in Summer Heat Index over Pakistan, Pakistan J. Meteorol., 6(12),85-96, 2010.

Zahid, M. and Rasul, G.: Changing trends of thermal extremes in Pakistan, Clim. Change, 113, 883-896, doi:10.1007/s10584-011-0390-4, 2012.

Zhang, X.B., Zwiers, F.W., Li, G.L., Monte Carlo experiments on the detection of trends in extreme values, *J. Clim.* 17,1945– 1952, 2004.
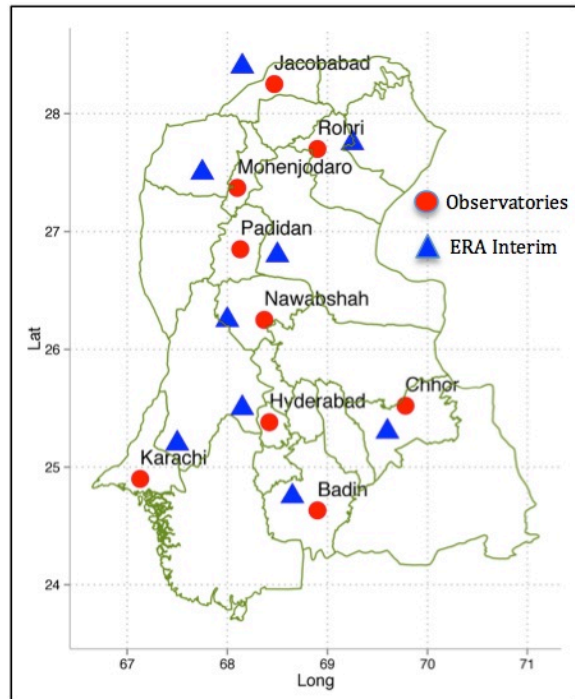
Figure 1: Study Domain (23.5 – 28.5° N , 66.5- 71.1°E)

Table 1. Code, Name, Geographic coordinates and Altitude of the stations.

Code	Name	PMD weather stations			ERA-Interim stations	
		Latitude	Longitude	Altitude (m)	Latitude	Longitude
JCB	Jacobabad	28° 18'N	68° 28'E	55	28 °4'N	68 °15'E
MJD	Mohenjo-daro	27° 22'N	68° 06'E	52.1	27°5'N	67 °75'E
RHI	Rohri	27° 40'N	68° 54'E	66	27°75'N	69 °25'E
PDN	Padidan	26° 51'N	68° 08'E	46	26°8'N	68 °5'E
NWB	Nawabshah	26° 15'N	68° 22'E	37	26°25'N	68 °0'E
HYD	Hyderabad	25° 23'N	68° 25'E	40	25°5'N	68 °15'E
CHR	Chhor	29° 31'N	69° 47' E	5	25°3'N	69 °6'E
KHI	Karachi	24° 54'N	67°08' E	21	25°2'N	67 °5'E
BDN	Badin	24° 38'N	68° 54'E	10	24 °75'N	68 °65'E

Table 2. Results of the Kolmogorov-Smirnov Goodness of fit test and Anderson-Darling test between empirical and GPD fits.

Observed T_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.947	0.340	0.996	0.139	0.941	0.385	0.928	0.306	0.666
Anderson Darling	0.553	0.978	0.654	0.857	0.157	0.649	0.233	0.869	0.145
ERA Interim T_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.169	0.125	0.553	0.456	0.322	0.187	0.419	0.456	0.332
Anderson Darling	0.355	0.263	0.165	0.587	0.615	0.398	0.266	0.687	0.425
Bias corrected ERA Interim T_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.452	0.4729	0.197	0.489	0.269	0.137	0.158	0.243	0.312
Anderson Darling	0.352	0.315	0.235	0.270	0.335	0.289	0.216	0.390	0227
Observed TW_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.981	0.111	0.341	0.226	0.457	0.545	0.441	0.385	0.211
Anderson Darling	0.623	0.745	0.587	0.884	0.199	0.123	0.789	0.669	0.473
ERA Interim TW_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.712	0.564	0.955	0.425	0.258	0.134	0.856	0.497	0.222
Anderson Darling	0.236	0.474	0.516	0.219	0.356	0.117	0.537	0.464	0.613
Bias corrected ERA Interim TW_{max}									
Test Statistics	P-value								
	JAC	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	0.268	0.688	0.127	0.372	0.268	0.229	0.591	0.582	0.478
Anderson Darling	0.373	0.484	0.278	0.432	0.306	0.283	0.365	0.445	0.483

Table 3. Estimated parameters shape ξ, scale σ and standard error Δξ, Δσ of all the data sets.

Station observed T$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.387	-0.255	-0.418	-0.326	-0.332	-0.329	-0.310	-0.222	-0.329
Standard Error Δξ	0.031	0.022	0.022	0.021	0.020	0.031	0.037	0.034	0.031
Scale σ	2.754	2.081	2.351	2.214	2.139	2.228	2.562	2.568	2.228
Standard Error Δσ	0.142	0.104	0.107	0.107	0.103	0.116	0.146	0.144	0.116
ERA Interim T$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.195	-0.178	-0.207	-0.218	-0.213	-0.338	-0.285	-0.037	-0.251
Standard Error Δξ	0.032	0.034	0.034	0.028	0.026	0.031	0.033	0.050	0.037
Scale σ	1.464	1.323	1.344	1.504	1.563	2.065	1.849	1.330	2.041
Standard Error Δσ	0.079	0.073	0.074	0.078	0.078	0.108	0.094	0.090	0.115
Bias Corrected ERA Interim T$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.195	-0.178	-0.207	-0.218	-0.213	-0.338	-0.285	-0.037	-0.251
Standard Error Δξ	0.032	0.034	0.034	0.028	0.026	0.031	0.033	0.050	0.037
Scale σ	1.983	1.791	1.820	2.038	2.116	2.798	2.308	1.801	2.763
Standard Error Δσ	0.108	0.100	0.100	0.106	0.106	0.146	0.123	0.122	0.156
Station observed TW$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.176	-0.186	-0.215	-0.215	-0.216	-0.323	-0.242	-0.219	-0.186
Standard Error Δξ	0.038	0.035	0.034	0.044	0.026	0.026	0.034	0.036	0.032
Scale σ	2.759	2.045	1.960	2.078	1.857	2.372	2.512	2.337	1.903
Standard Error Δσ	0.159	0.114	0.108	0.128	0.093	0.119	0.138	0.132	0.105
ERA Interim TW$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.089	-0.094	-0.068	-0.125	-0.158	-0.177	-0.090	-0.019	-0.173
Standard Error Δξ	0.037	0.029	0.032	0.034	0.031	0.037	0.035	0.035	0.037
Scale σ	1.287	1.243	1.231	1.440	1.610	1.649	1.3423	0.680	1.788
Standard Error Δσ	0.074	0.066	0.067	0.080	0.087	0.095	0.076	0.039	0.102
Bias Corrected ERA Interim TW$_{max}$									
Estimates	**JCB**	**MJD**	**RHI**	**PDN**	**NWB**	**HYD**	**CHR**	**KHI**	**BDN**
Shape ξ	-0.089	-0.094	-0.068	-0.125	-0.158	-0.177	-0.090	-0.019	-0.173
Standard Error Δξ	0.037	0.029	0.032	0.034	0.031	0.037	0.035	0.035	0.037
Scale σ	1.356	1.646	1.758	1.494	1.520	2.052	2.146	1.399	2.152
Standard Error Δσ	0.078	0.087	0.096	0.083	0.082	0.119	0.121	0.081	0.123
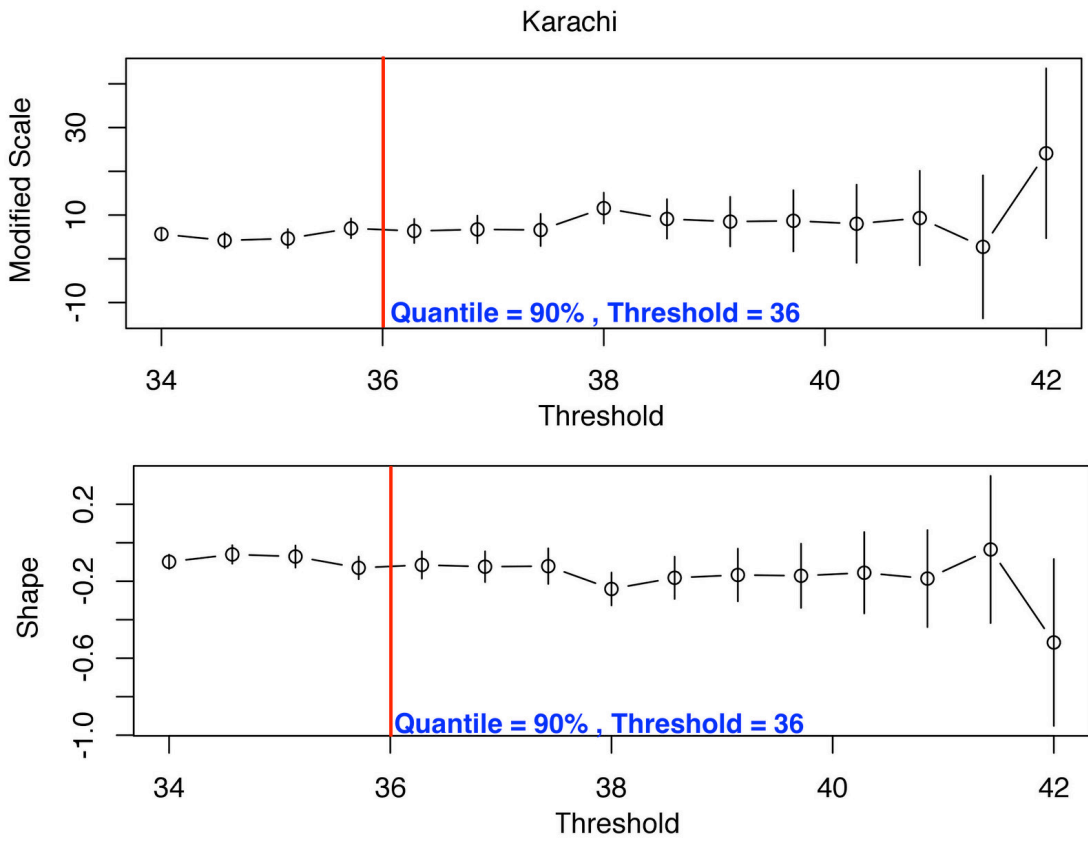
Figure 2. Modified scale (σ*) and shape parameter (ξ) of the observed T$_{max}$ (°C) Karachi. The red vertical lines represent the selected threshold according to the station quantiles.
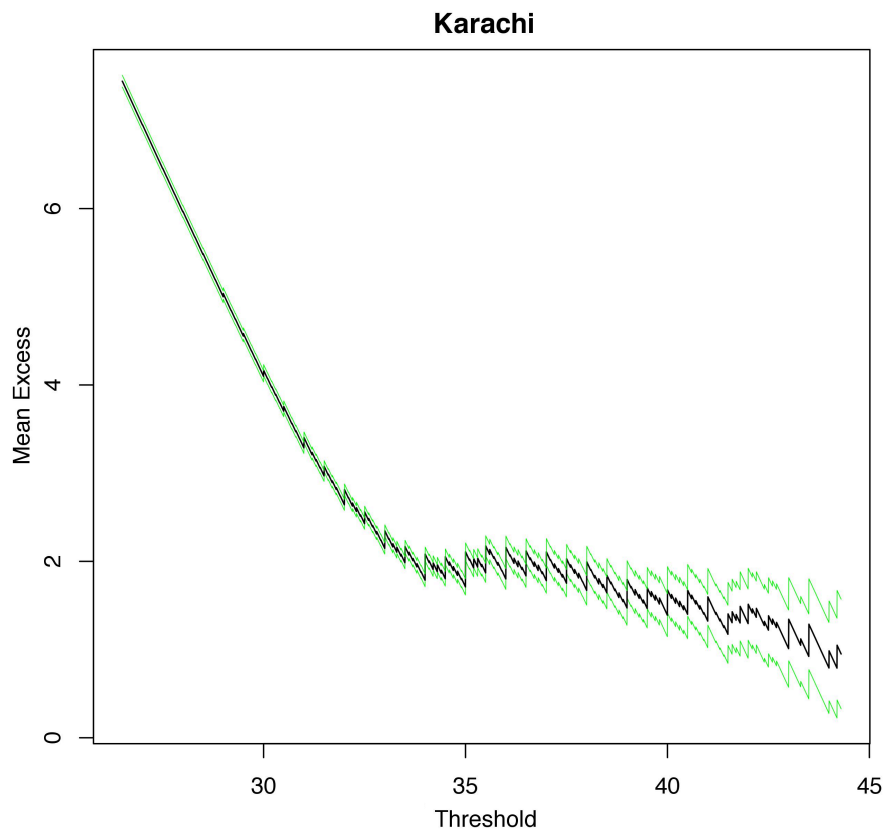
Figure 3. Mean residual life plot of the station observed T$_{max}$ (°C) Karachi.
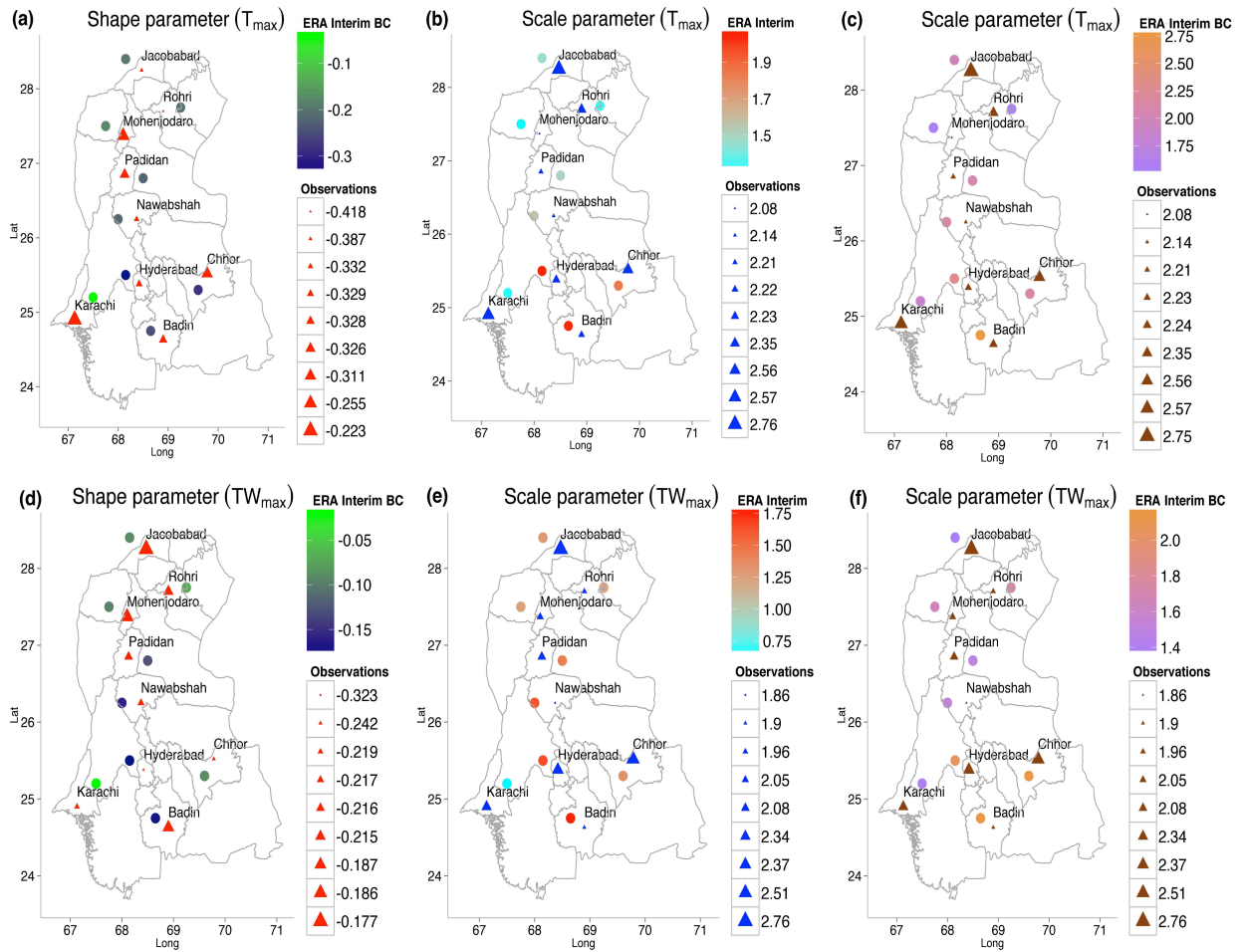
Figure 4. Spatial distribution of the shape parameters ξ and scale parameters σ of the station observed, ERA Interim, and bias corrected ERA Interim T$_{max}$ (upper panel) and TW$_{max}$ (lower panel) degree Celsius.
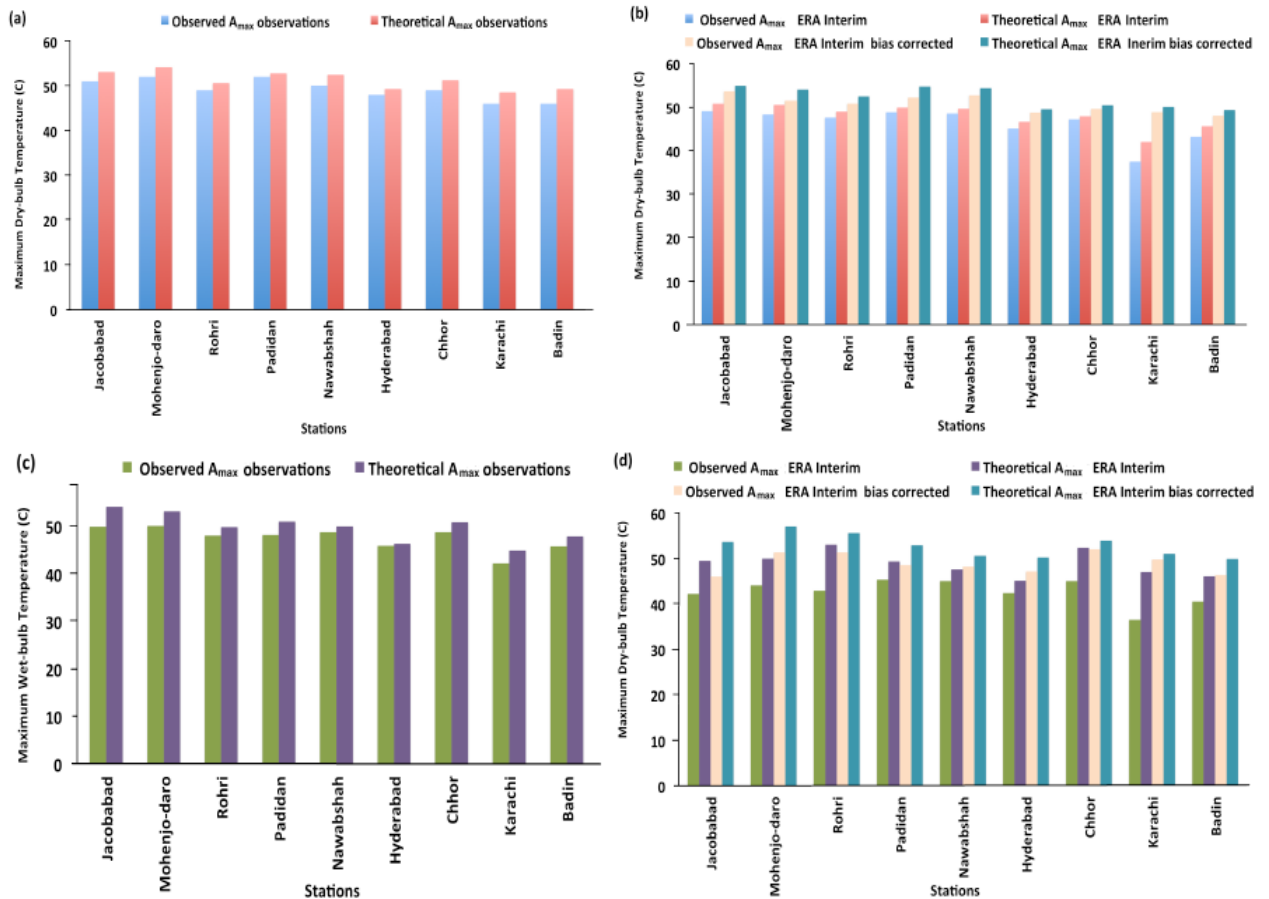
Figure 5. Absolute maxima A_{max} in degree Celsius (a) station observed T_{max} (b) ERA Interim and bias corrected ERA Interim T_{max} (c) station observed TW_{max} (d) ERA Interim and bias corrected ERA Interim TW_{max}.
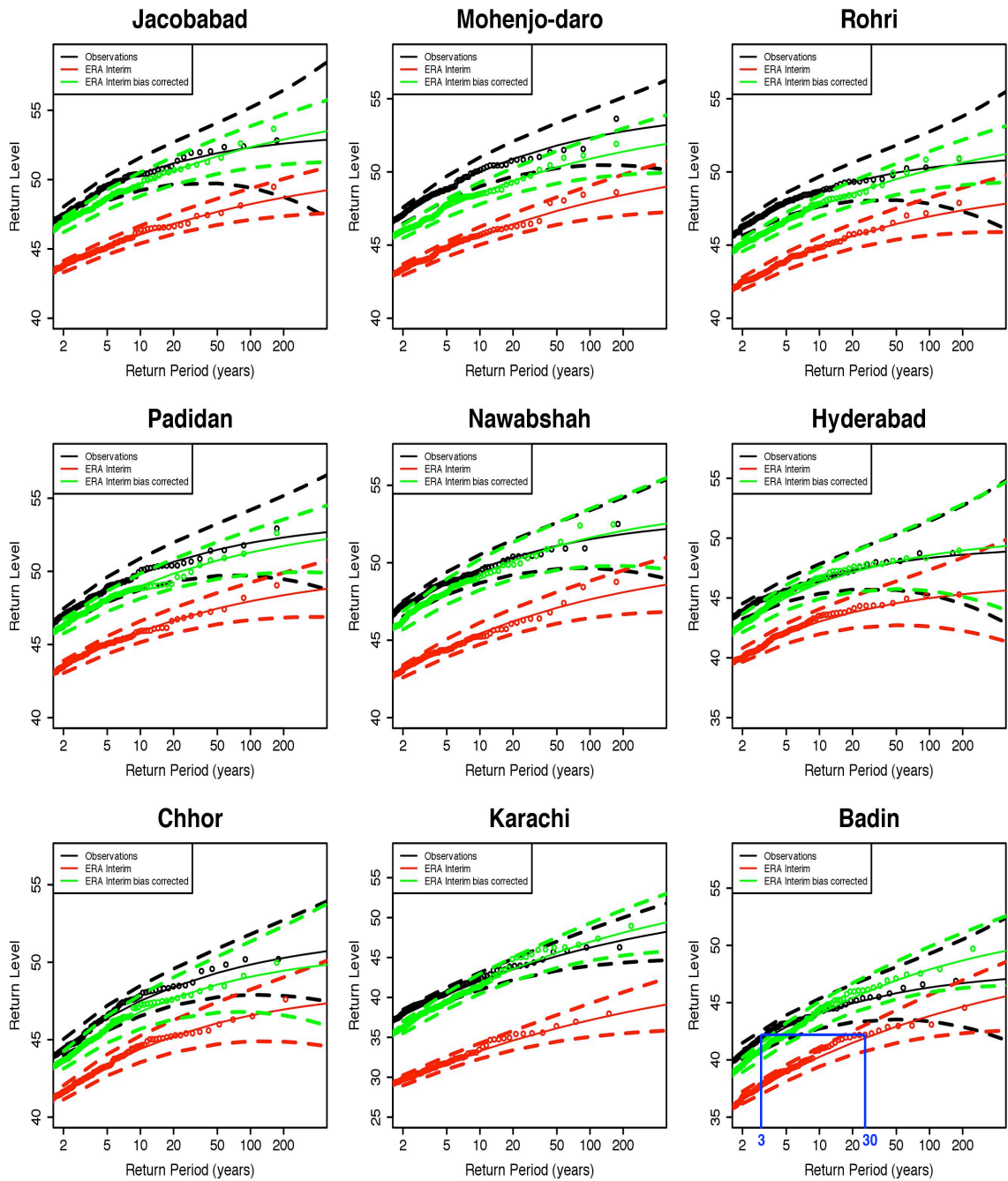
Figure 6. Return level plots of the station observed T$_{max}$ (black) , ERA Interim T$_{max}$ (red), and bias corrected ERA Interim T$_{max}$ (green) in degree Celsius. The blue line is to show a difference in the observed and ERA Interim RLs.
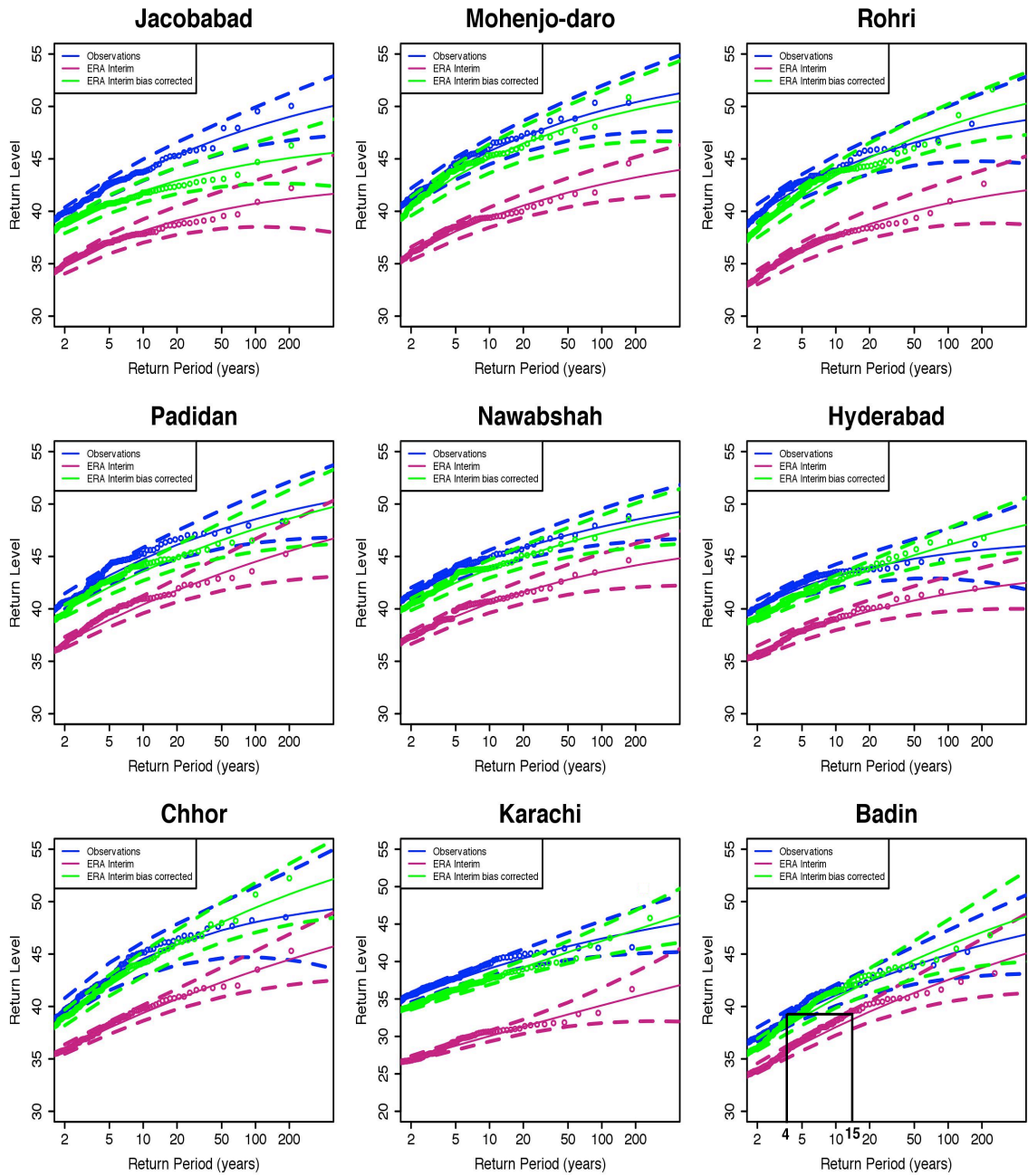
Figure 7. Return level plots of the station observed TW$_{max}$ (blue), ERA Interim T$_{max}$ (pink), and bias corrected ERA Interim T$_{max}$ (green) in degree Celsius. The black line is to show a difference in the observed and ERA Interim RLs.
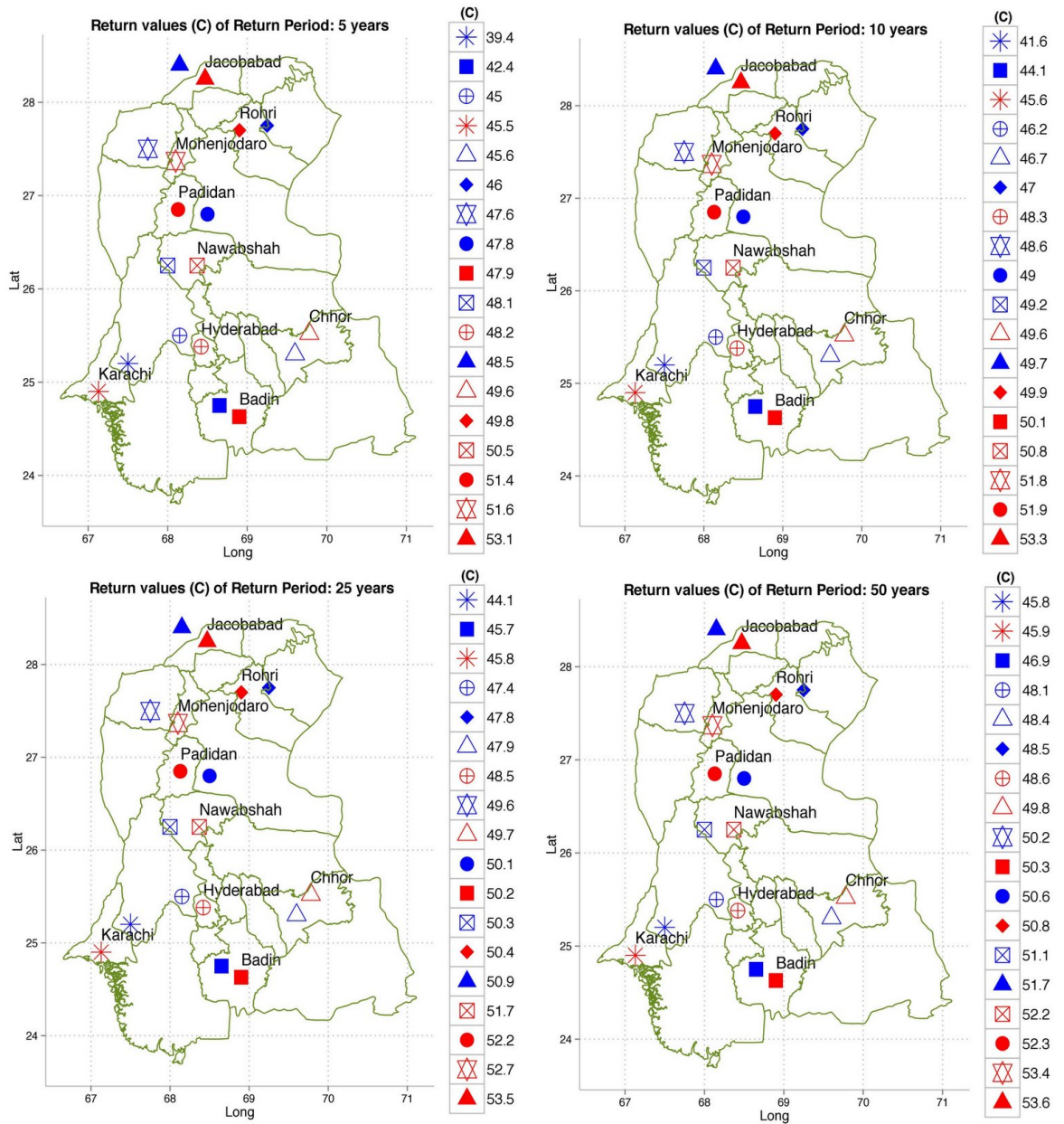
Figure 8. Spatial distribution of the station observed T$_{max}$ (red) and bias corrected ERA Interim T$_{max}$ (blue) return levels in degree Celsius corresponding to return periods of 5, 10, 25 and 50 years in southern Pakistan.
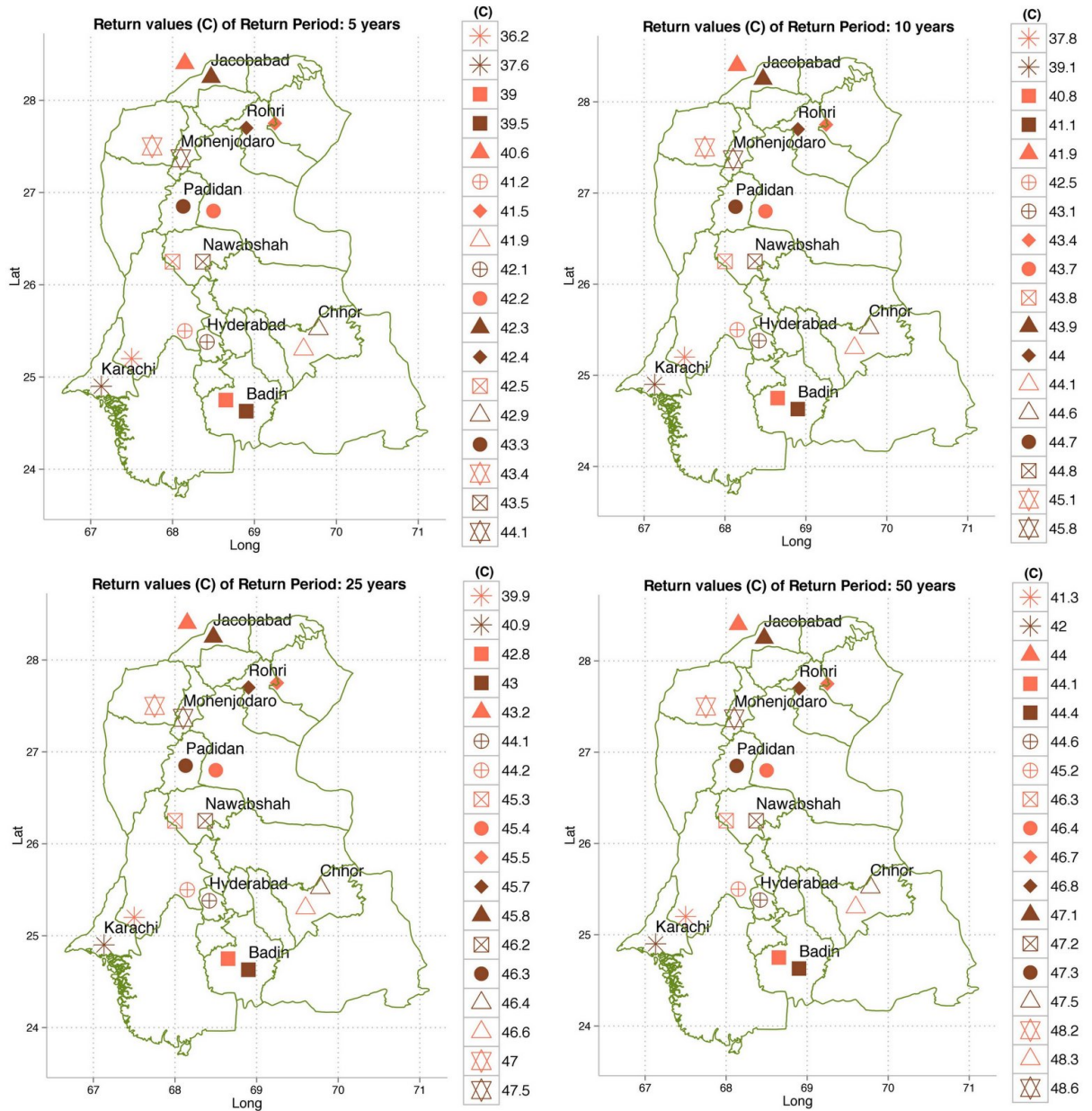
Figure 9. Spatial distribution of the station observed TW$_{max}$ (brown) and bias corrected ERA Interim TW$_{max}$ (orange) return levels in degree Celsius corresponding to return periods of 5, 10, 25 and 50 years in southern Pakistan.