

Response to Referee #1

Anonymous Referee #1

The authors would like to thank anonymous referee #1 for the constructive comments that helped us to improve our manuscript. The answers to the Referee 1 comments are given point by point in the following.

1. Both Generalized Extreme Value (GEV) and GPD distributions can be applied for assessing return levels and return periods of climate extreme events. The length of 33- year annual maximum values seems to be sufficient for deriving reasonable estimates using the GEV technique. Therefore, it is not so clear why the GEV distribution is not preferred here. Did authors examine the GEV-based return levels? Does GPD provide a better fit and a more reliable estimation?

Ans: We preferred GPD because in many applications, the POT approach is preferred to the BM approach for fitting time series because it provides more efficient use of data and has better properties of convergence when finite datasets are considered (Lucarini et al., 2016 ; Coles 2001 ; Holmes and Moriarity; Davison and Smith, 1990). Additionally, we are here interested in investigating the actual tails of the distributions, so the GPD point of view is more appropriate.

Revisions: Lines 11-14, p3 is added to explain the use of GPD in this study.

2. GPD approach has been widely applied for estimating the statistics of extreme rainfall and temperature (e.g. Katz et al., 2002, 2010, Cooley et al., 2005, 2007, etc.), and the techniques used in this manuscript are not nascent. In order to sufficiently support the argument that the application of GPD herein is “novel” and can provide “novel” information, highlighting any new or additional findings which can only rely upon the GPD approach is desired.

Ans: We definitely agree that the GPD approach has been applied widely for estimating the statistics of extreme rainfall and temperature. But by saying “novel” we meant that the GPD is applied for the first time to southern Pakistan region, extremes are never been estimated using the GPD here. This paper is the first one to introduce the climate extremes information in the region based on GPD. The word novel has been removed from the manuscript to avoid the misunderstanding.

Revisions: Line 23, p11 word “novel” is removed.

3. By comparing the distribution parameters and return levels derived from observations and ERA Interim data, it seems that they have a large agreement in the shape parameter estimations at some stations, while the bias in the mean and variance of model simulations is the primary factor that leads to the underestimation of the return levels. Does the agreement in the shape parameters indicate that the underlying physical process which produces extreme temperature is well represented by the climate model, though there is a bias in simulating the internal variability of extreme temperature? Please consider to extend the current discussions in this regard.

Ans: The results for the shape parameter indicate that the functional dependency of the extreme value distribution is reasonably well simulated by the ERA data. This is especially relevant for the existence of absolute maxima (Eq 3, Section 2.4) in the case of negative shape parameters. The agreement of the shape parameters in the observations and simulations means that the ERA dataset captures an important aspect of extremal behavior. This is in principle a non trivial result, as reanalysis are constructed in such a way that typical conditions are well reproduced.

Revisions: In section 3.3, lines 27-29, p8 , the information is added.

4. At stations such as JCB, MJD, RHI, the bias-corrected return levels underestimate the observed values. Which factor/parameter would be responsible for the consistent underestimation? For those locations, is there a way to conduct the bias correction for the shape value? Would a higher threshold correct such underestimation?

Ans: The disagreement of the bias corrected results indicates that the standard bias correction method based on the first two moments is not sufficient for these stations. A better agreement could be obtained by including the higher moments to improve the estimate of the extreme values. We do not think that the higher threshold can correct such underestimation because we are within the asymptotic regime.

5. In section 2.1, the purpose and benefit of adding noise to the data are not clear. By adding the noise, does the convergence of parameter estimation become more efficient? Why?

Ans: The advantage of adding a noise is to avoid the spurious statistical effects associated to the presence discrete values assigned to the temperature readings. This is discussed in detail in the cited paper of Deidda and Puliga 2006 for hydrological extremes. Using the described bootstrap method we reduce such problem without biasing the data.

Revisions: In section 2.1, lines 18-23, and 30-32, p4, the purpose and benefit advantage of noise is added

6. Please consider to rearrange the order of tables and introduce them in sequence. For instance, Table 2 is introduced ahead of Table 1 in the context, so please switch their orders.

Ans: The order of tables has been rearranged.

7. Please provide the Q-Q plots for the 9 stations, since the authors discussed the “slight deviation” revealed by examining the corresponding Q-Q plots.

Ans: Q-Q plots are provided as additional material.

References:

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

Davison A. C. and Smith R. L : 1990, Journal of the Royal Statistical Society. Series B (Methodological) Vol. 52, No. 3, 393-442

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

Holmes, J. D. and Moriarty, W. W.: 1999, Application of the Generalised Pareto Distribution to wind engineering, J. Wind Engineering and Industrial Aerodynamics 83, 1– 10.

Lucarini, V., Faranda, D., Freitas, A.C.M., Freitas, J.M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, 26 S.: Extremes and Recurrence in Dynamical Systems, John Wiley & Sons Inc, Engelska,2016.

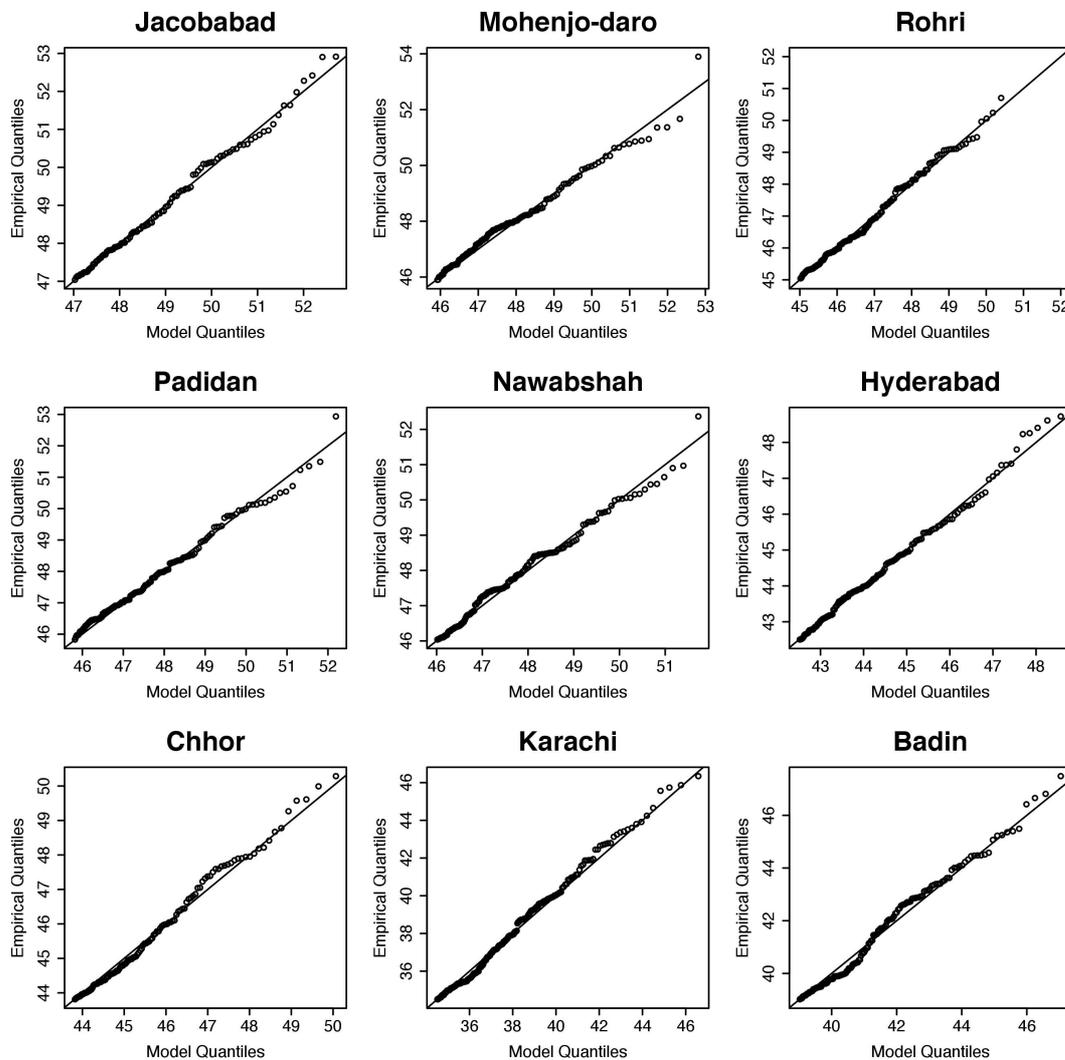


Figure 1. Quantile-Quantile plots of observed T_{max} , $u = 90\%$ for 9 stations of Southern Pakistan (Sindh).

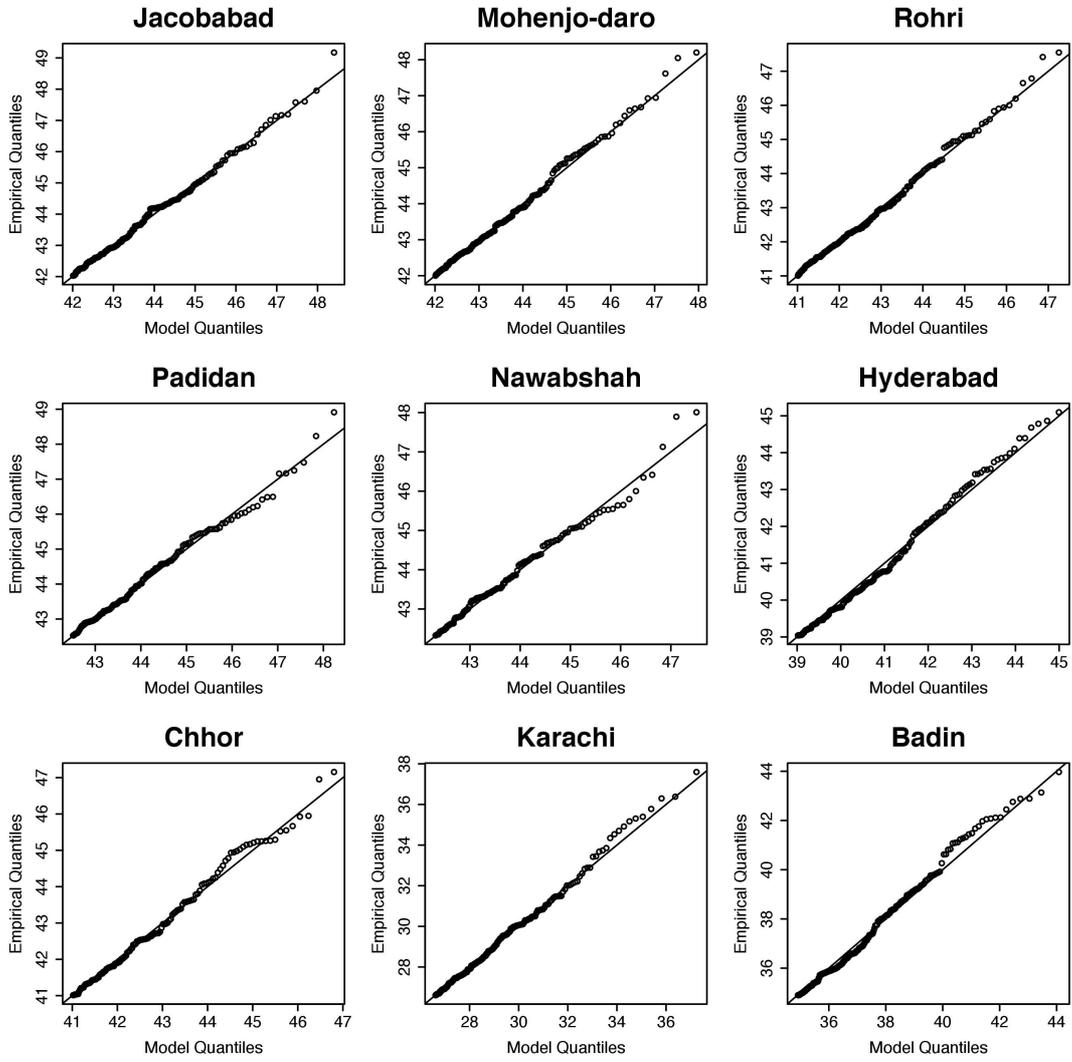


Figure 2. Quantile-Quantile plots of ERA Interim T_{\max} , $u = 90\%$ for 9 stations of Southern Pakistan (Sindh).

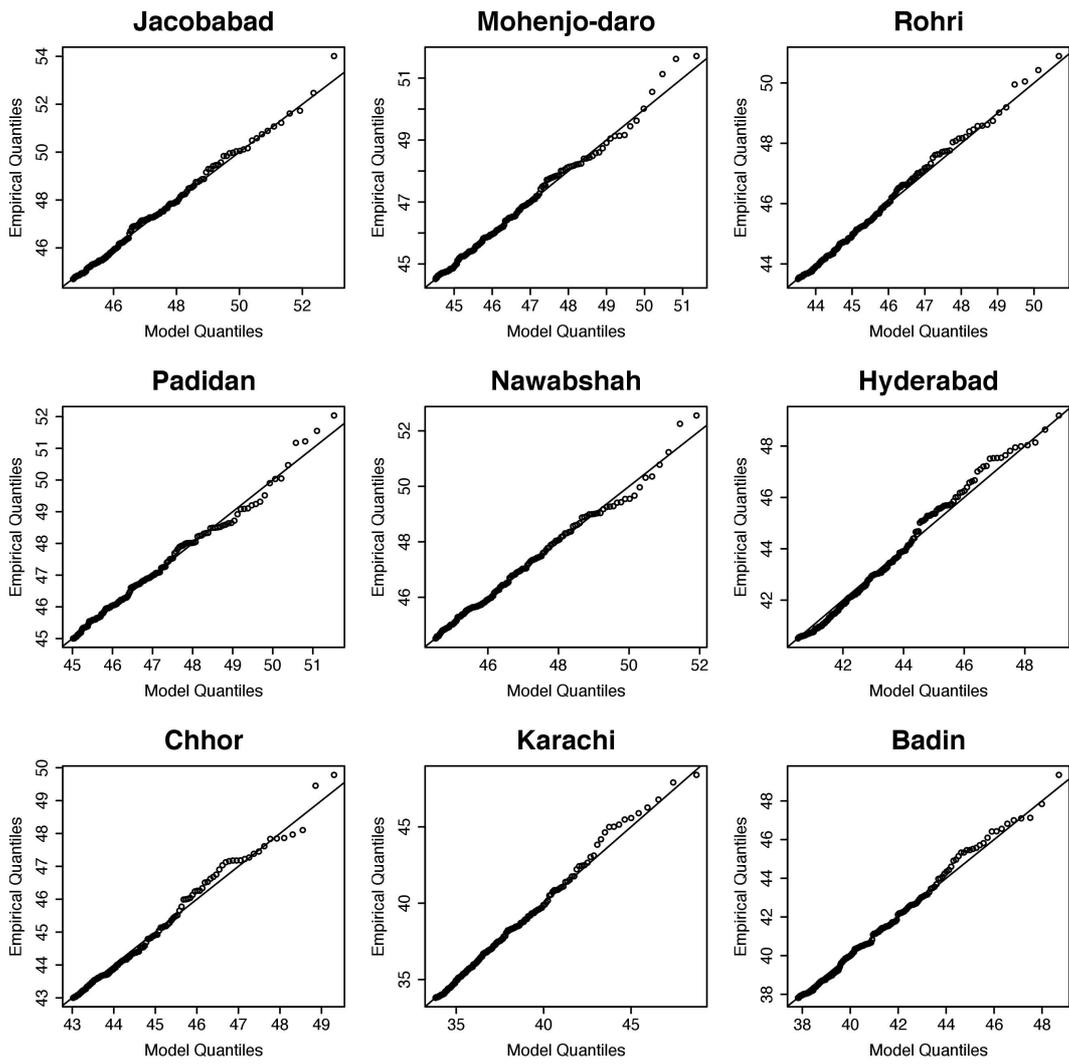


Figure 3. Quantile-Quantile plots of Bias corrected ERA Interim T_{\max} , $u = 90\%$ for 9 stations of Southern Pakistan (Sindh).

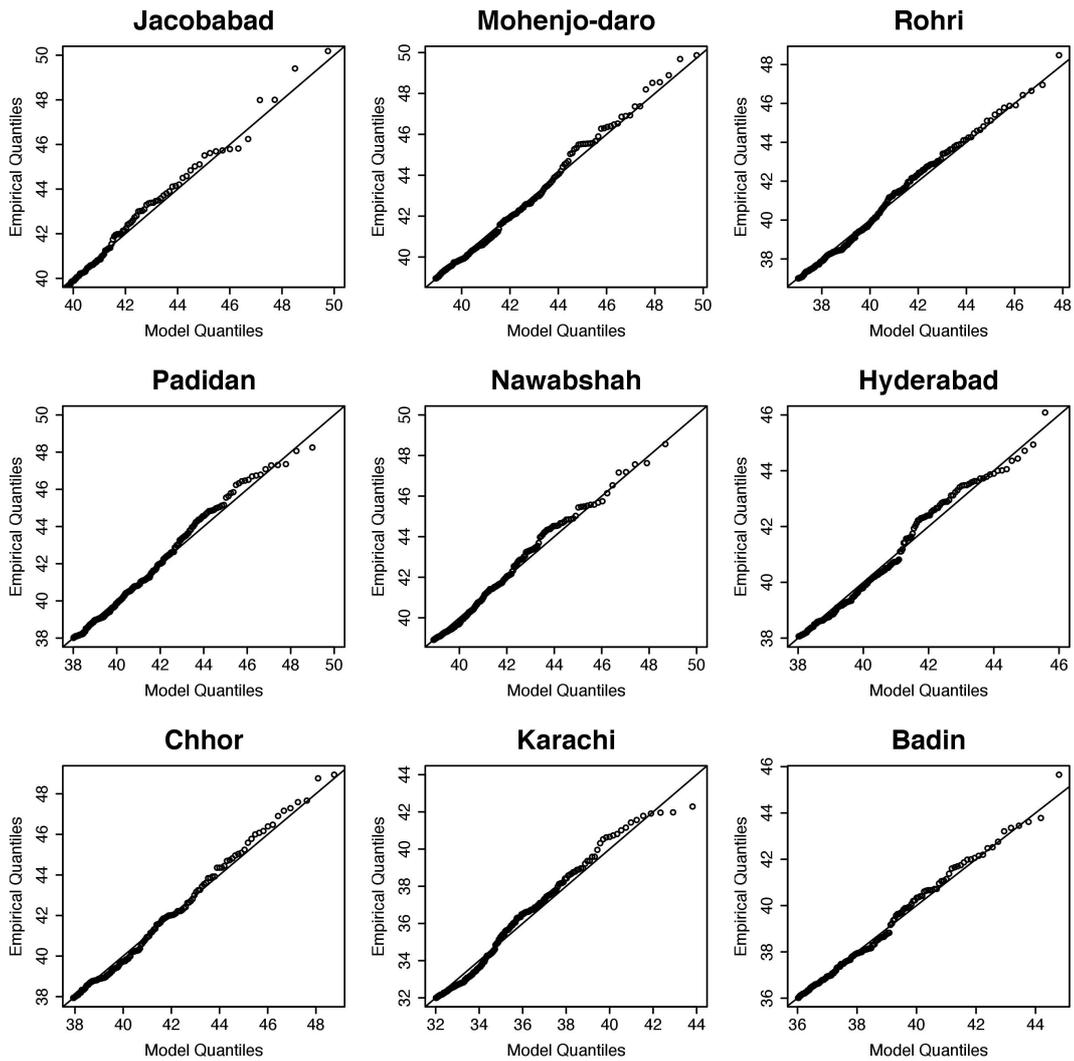


Figure 4. Quantile-Quantile plots of observed TW_{max} , $u=90\%$ for 9 stations of Southern Pakistan (Sindh).

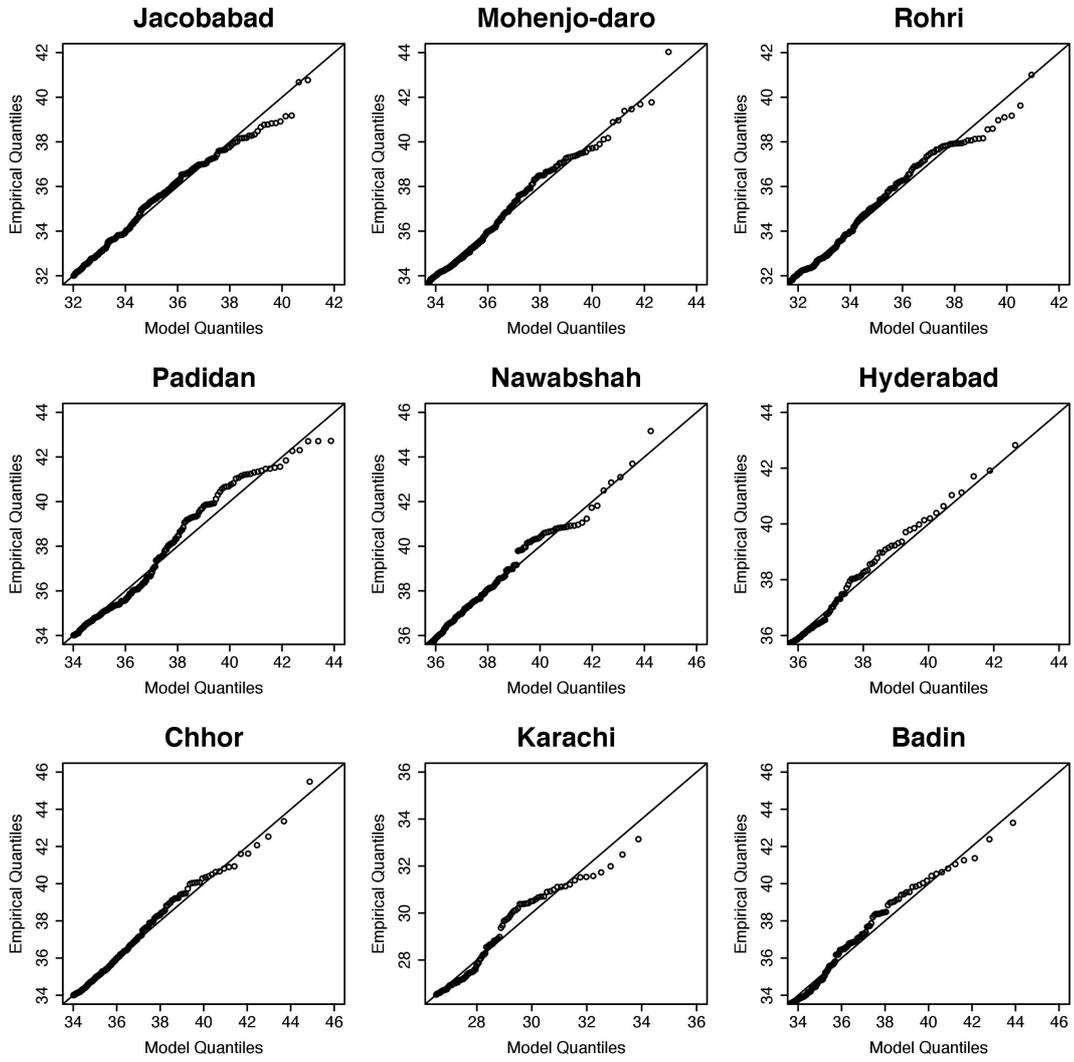


Figure 5. Quantile-Quantile plots of ERA Interim TW_{max} , $u=90\%$ for 9 stations of Southern Pakistan (Sindh).

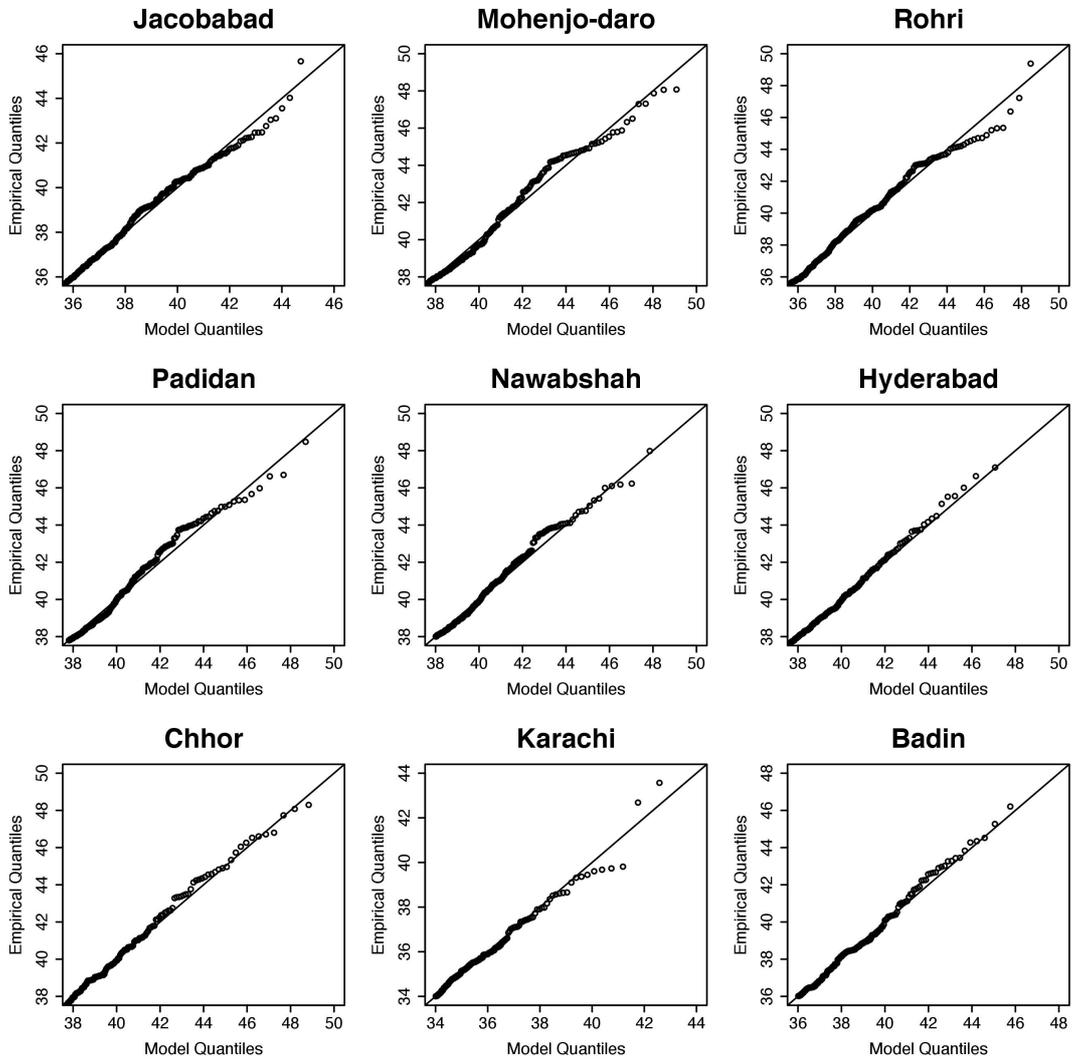


Figure 6. Quantile-Quantile plots of Bias corrected ERA Interim TW_{max} , $u = 90\%$ for 9 stations of Southern Pakistan (Sindh).

Response to Referee # 2

S. Parey (Referee)
sylvie.parey@edf.fr

The authors would like to thank Sylvie Parey (Referee 2) for the careful review and comments that helped us to improve our manuscript. Our responses to the comments are as follows.

1. Line 28 p3, it is stated “If the ERA Interim dataset characterizes well the extremes: : :”. This is very unlikely, since re-analysis has a too low spatial resolution to represent adequately local extremes, this is not the aim of reanalysis

Ans: It is in principle not obvious that ERA 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 data performs in the target area against observations. Note that Cornes and Jones, (2013) reported that the ERA-Interim reanalysis data are generally very good at replicating trends in percentile-based measures of temperature extremes. However, ERA-Interim is weak in capturing the extreme temperatures in complex terrains, but our study area has simple terrain.

Revisions: An explanation regarding the use of ERA Interim is added in lines 29-35, p3.

2. Lines 31 to 35 p4: it is true that stationarity is a requirement to perform POT, but stationarity means that the distribution is invariant by translation in time, not the autocorrelation is weak. Autocorrelation is rather linked to independence, and must be handled too. Maybe here stationarity refers to the iid (independent and identically distributed) condition, then this section deals with independence, but not with identical distribution. However, identical distribution needs to be checked as well, just because of seasonality. Has seasonality been checked? Is the occurrence of the highest values restricted to the summer season or identically distributed throughout the year? If a season appears as favoring high temperature values, then the analysis should be restricted to this season, otherwise the occurrences are not distributed according to a homogenous Poisson process and the frequency of occurrence is biased. Not only seasonality can disturb the identical distribution, trends can too. There is no discussion about possible trends. Maybe it is possible to neglect the trends over the relatively short 1980-2013 period, but this could be checked.

Ans: We agree with the reviewer. We note that one can see non-stationarity and presence of trends as the presence of long-time correlation in the data. Clearly, extreme temperature and extreme heat indices are realized in summer conditions, so the analysis is restricted to summer season. Including the other seasons would make no sense. We have tested that trends are not significant in such a short time interval. Short-time correlations are studied by computing the extremal index θ in all time series and treated using the associated standard declustering technique.

Additionally, the strict stationarity means that the distribution of the random process is invariant to time shifts. Weak stationarity concerns only the invariance in time of

the first and second order moments of the random process, i.e. mean and auto-covariance function (or autocorrelations). However, if the process is Gaussian, then strict stationarity and weak stationarity are equivalent. Nevertheless, weak stationarity and ergodicity are sufficient conditions in time series which allows (when the number of time periods is quite large) to consider the observations 'as if' they were i.i.d. That is why we just check for covariance stationarity (and assume the process is ergodic, i.e. the process has not a long memory: autocovariances decay to 0 for very large time lags). Moreover, when we run the unit-root test (test for stationarity), we check for the presence of a stochastic trend. So, actually, before running the POT analysis it is definitely discussed in the paper the possible presence of a trend.

Revisions: Lines 4-5, and 8-10, p5 are added to address the seasonality, trends and correlations.

3. When computing TWmax, are we sure that Tmax and RHmax occur in the same time?

Ans: The occurrence of TWmax takes place in general in different dates than Tmax or RHmax, because one could have a day with high temperature but low humidity and vice versa. Since our study area is next to the Arabian Sea the level of Tmax and RHmax remains constantly high during summer.

4. Some considerations on independence are given again between lines 15 and 22 p6, but no indications are given on how this is used and applied in the study

Ans: Computing the extremal index θ allows for studying the degree of clustering of extremes. The inverse of θ gives the average length of a cluster. Usually $\theta = 0$ means strong clustering and dependence, $\theta = 1$ absence of clusters and independence. The extremal index value in all the time series is ≤ 0.5 referring to dependence. Therefore, it is necessary to decluster the extremes by choosing the largest event in each cluster, before fitting it to the GPD. Note that this is the practical strategy commonly adopted by practitioner as well as the rigorous prescription suggested by mathematics.

Revisions: Lines 34 - 41, p5, and 8-10, p6 are rearranged and added to clarify the use of extremal index in the study.

5. In section 3.2 concerning the GPD fits, one can read “if the higher quantiles are neglected, then the stations like : : : show that the exceedances fit very well”. But in an extreme value analysis, the higher quantiles are the targeted ones!

Ans: By saying neglected we did not mean to neglect the higher quantiles. We wanted to say if the higher quantiles are disregarded or unnoticed. We have replaced the word “neglected”, as it can be misleading for the readers.

Revisions: Lines 8 - 10, p8, the word “neglected” is removed to avoid confusion.

6. In the conclusion, it is stated that: “This paper contains novel and beneficial

information: : , which would help the local administration to prioritize the regions in terms of adaptation”. What does adaptation mean here? The estimated levels are based on observations, thus these are rare levels which could occur even if there were no additional warming. It is not clear in the paper if this distinction is made. The notion of return level is defined for stationary time series, that is without any cycles nor trends, and is devoted to the estimation of very rare levels which could happen (once every N years in average), but may not yet have been observed. Climate warming brings other difficulties in their estimation: the definition of a return level has to be changed because a past and a future period are not prone to experience the same temperature levels, and the estimation is complicated. Different papers are devoted to this problem, for example:

Cheng, L., AghaKouchak, A., Gilleland, E., & Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic change*, 127(2), 353-369
 Du, T., Xiong, L., Xu, C. Y., Gippel, C. J., Guo, S., & Liu, P. (2015). Return period and risk analysis of nonstationary low-flow series under climate change. *Journal of Hydrology*, 527, 234-250
 Obeysekera, J., & Salas, J. D. (2016). Frequency of Recurrent Extremes under Nonstationarity. *Journal of Hydrologic Engineering*, 21(5), 04016005
 Parey, S., Malek, F., Laurent, C., Dacunha-Castelle, D. (2007). Trends and climate evolutions: statistical approach for very high temperatures in France. *Clim.Change*, 81, 331–352.
 Parey, S., Hoang, R.T.H., Dacunha-Castelle, D. (2010). Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21, 698–718.

Ans: Adaptation means 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. We wish to remark that the study domain is one of the hottest region in the world as mentioned in the paper with the highest record-breaking temperature of 52°C in 2010. This region is a hub of agriculture activities and 50% of the population work outdoors. The local administrations have limited resources, therefore they want to prioritize the region for the adaptations like, early warning systems, introducing new the temperature tolerant crops, water management and providing shelters for the outdoors workers etc. Therefore, the information of return levels is good for the planning and adaptation strategies. So, a stationary analysis is already a pretty relevant contribution for the region. Moreover, we consider the stationary extreme value analysis due to short duration of the data (33 years) and to have reliable estimates with less uncertainty. Clearly, considering non-stationarity is a good idea for future work. We might consider using the centennial NCEP reanalysis [Compo et al., 2011].

Revisions: Lines 24 - 34, p11 are added to explain the meaning of adaptation and future work.

References

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, 2011: The Twentieth Century Reanalysis Project. *Quarterly J. Roy. Meteorol. Soc.*, 137, 1-28. <http://dx.doi.org/10.1002/qj.776>

Cornes, R. C., and P. D. Jones, 2013: How well does the ERAInterim reanalysis replicate trends in

Response to Referee # 3

Anonymous Referee #3

We would like to thank anonymous Referee #3 for the critical review and comments. The step-by-step responses to the main reservations of the referee are as follows.

1. Why we did not analyze the overall increasing trends in the frequency or intensity of extremes related to daily maximum temperatures?

Ans: We did not analyze the increasing trends in the frequency and intensity of extremes related to daily maximum temperatures, and considers the stationary extreme value analysis due to short duration of the data (33 years) and to have reliable estimates with less uncertainty. As described in the text, we have checked that trends are not statistically significant.

Moreover, the study domain is one of the hottest region in the world as mentioned in the paper with the highest record-breaking temperature of 52°C in 2010. This region is a hub of agriculture activities and 50% of the population work outdoors. The local administrations have limited resources, so they want to prioritize the region for the adaptations like, early warning systems, introducing new the temperature tolerant crops, water management and providing shelters for the outdoors workers etc. Therefore, the information of return levels is good for the planning and adaptation strategies. So, a stationary analysis is already a pretty relevant contribution for the region. Clearly, considering non-stationarity is a good idea for future work. We might consider using the centennial NCEP reanalysis [Compo et al., 2011].

Reference: 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, 2011: The Twentieth Century Reanalysis Project. Quarterly J. Roy. Meteorol. Soc., 137, 1-28. <http://dx.doi.org/10.1002/qj.776>

Revisions: An explanation regarding the stationary extreme value analysis is added in lines 26 - 30, p11.

2. The motivation: It does not become clear, why the re-analysis data are relevant here, except that they might be used to fill gaps with missing observations. In this respect, it appears important to correct the re-analyses for biases. That would, however, require a more advanced method as the one used here, combining local information at the stations with information on large-scale conditions.

Ans: It is common practices among meteorologists to use ERA Interim (or NCEP) data to study the local to regional to large scale climatic properties. These datasets are

also often used to assess the skill of climate models. Therefore, it seems reasonable to include them here. Additionally, the ERA Interim reanalysis data was proved to be very good at replicating trends in percentile-based measures of temperature (Cornes and Jones, 2013). However, it is still not clear that ERA data can simulate well meteorological extremes. This is why we use ERA Interim data to see how well it performs in the target area against observations. We are well aware – this is clearly explained in the text – that one could use more advanced bias correction methods. But here we want to show whether if we reduce to zero the bias in the first two moments (note that most scientists and practitioners focus only on these two statistical properties), we are still able to have a good representation of the tail of the distribution. In some stations like Nawabshah, Karachi etc, even the standard bias correction show very good agreement with observations. However, we agree with reviewer that if ERA data has to be used in the region (and elsewhere) to study extremes, a more advanced method is needed. We wish to underline the need to look into actual station data. We have added this detail in the revised version of the paper.

Revisions: Motivation for using ERA Interim data is added in lines 29-35, p3.

3. The presentation of the methodology: The presentation of the methodology fills a rather large part of the manuscript, although much of it is widely used. Therefore, this part of the manuscript could be shortened.

The presentation of the results: The presentation of the results is not very concise. Numerous numbers and maps are included in the manuscript, but often they are not properly presented.

Ans: We agree that the statistical methods used here method is widely used, but, we prefer giving some details in order to address an audience that might be not so familiar with extreme value theory. We agree that the presentation includes a lot of maps and figures, but we remind the reviewer that, as mentioned in the paper, this is the first analysis for extremes using extreme value theory in Pakistan. Therefore, we consider giving all the possible details, to provide a thorough picture of the methodology and results to the fellow researches in Pakistan and neighboring countries.

4. The discussion of the results: The results of the study are not really discussed, neither with respect to the scientific literature nor with respect to the underlying physical mechanisms and, only partly, with respect to the representativeness of the results for southern Pakistan for the rest of Pakistan or the rest of the wider region.

Ans: The results are discussed with respect to the available scientific literature.

Please see the following cited references in results and discussion.

1. Sacrott and MacDonald, 2012 (line 24, page 7)
2. Coles, 2001 (line 25, p7)
3. Furrer et al., 2010 (line 25, p7)
4. Davison and Smith, 1990 (line 31, p7)
5. Hatfield and Preuger, 2015 (line 9, p11)

6. Compo et al., 2011 (line 29, p11)

The scientific literature regarding extreme value theory and return levels is not available (mentioned between line 14 -16 page 3) for this specific region, therefore we could not discuss it in results. Understanding the meteorological mechanisms behind heat waves are obviously extremely relevant, but well beyond the scope of this paper, which is mostly statistical in nature. Much more work at this regard would be needed, and we have added a sentence in the conclusions to emphasize this. Note that we have clearly explained why the statistical properties analyzed here are relevant for human welfare and economy in the region.

5. Concluding section: The concluding section is just a repetition of the main results of the study, and no conclusions of this study are given.

Ans: The concluding section is named as “Summary and Conclusion”, therefore we have summarized the results in the beginning and conclusions are given between lines 27-34, p 10 and 1-34, p11. However, it is customary to summarize the results of the paper at the beginning of the last section, especially for a paper where statistical properties are analyzed.

Revisions: Summary and conclusion section is revised.

6. Tables: As for Table 2, it is not clear, why monthly mean values of the daily minimum and mean temperatures are presented here. As for Table 3, I am puzzled by the substantially different behaviour of the p-values according to the KS-test and the p-values according to the AD-test. As for Table 4, I am missing the units.

Ans: Table 2: shows the mean monthly climatic characteristics of the region from 1980- 2010. It is there to describe the climatology of the region. This is a useful complement to the analysis of the extremes performed in the rest of the paper.

Table 3 : shows the different p-values because KS test and AD test are two different methods, and are used here to see the goodness-of-fit at each station. For details please see line 14-17 page 8.

Table 4 : show the estimated parameters shape ξ , scale σ and standard error $\Delta\xi$ of all the data sets.

Shape parameter ξ has no unit.

Scale parameter σ has the unit “degree Celsius” like temperature.

Revisions: The units are added in figure captions.

7. Figures: Generally, the figures and/or figure captions are lacking units. Also, in many cases the use of different plotting ranges for panels, which show the same kind of estimates for different data sets or different locations make it difficult to draw firm conclusions from these figures. As for Figs. 8 and 9, it layout of the panels makes is

very hard to extract the relevant information from the map, since it the information on the magnitude is hidden in the respective column including the symbols.

Ans: The units are placed inside the figures, but now we have written them in figure captions as well. Regarding Figure 8 and 9 we think two different colors clearly distinguish between observations and bias corrected ERA Interim return levels, also different symbols are used to differentiate among the cities and return level values. In our point of view, the information on the magnitude of extremes is quite obvious here. However, suggestions to improve Figure 8 and 9 are welcome. Unfortunately it is never easy to find optimal solutions for that kind of figures. We have relied on interactions with colleagues and practitioners in multiple poster and oral presentations to gain inputs on that.

8. References: Most of the references to the scientific literature are used in the Introduction and in the methodology section, also highlighting the fact that the sections on the results and the discussion are not properly done.

Ans: Given the nature of the paper (first analysis of extremes in the region), it seems quite natural that most of the referencing goes in the introduction and in the methodology. In results and discussion, following references are cited, highlighting the fact that it is properly done.

1. Sacrott and MacDonald, 2012 (line 24, page 7)
2. Coles, 2001 (line 25, p7)
3. Furrer et al., 2010 (line 25, p7)
4. Davison and Smith, 1990 (line 31, p7)
5. Hatfield and Preuger, 2015 (line 9, p11)
6. Compo et al., 2011 (line 29, p11)

References

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

List of all changes in the manuscript

Changes according to Reviewer # 1

- Lines 11-14, p3 is added to explain the use of GPD in this study.
- Line 23, p11 word “novel” is removed.
- In section 3.3, lines 27-29, p8 the information is added.
- In section 2.1, lines 18-23, and 30-32, p4, the purpose and benefit advantage of noise is added
- The order of tables has been rearranged.
- Q-Q plots are provided as additional material.

Changes according to Reviewer # 2

- An explanation regarding the use of ERA Interim is added in lines 29-35, p3.
- Lines 4-5, and 8-10, p5 are added to address the seasonality, trends and correlations.
- Lines 34 - 41, p5, and 8-10, p6 are rearranged and added to clarify the use of extremal index in the study.
- Lines 8 - 10, p8, the word “neglected” is removed to avoid confusion.
- Lines 24 - 34, p11 are added to explain the meaning of adaptation and future work.

Changes according to Reviewer # 3

- An explanation regarding the stationary extreme value analysis is added in lines 26 - 30, p11.
- Motivation for using ERA Interim data is added in lines 29-35, p3.
- Summary and conclusion section is revised.
- The units are added in figure captions.

Return Levels of Temperature Extremes in Southern Pakistan

Maida Zahid¹, Richard Blender¹, Valerio Lucarini^{1,2} and Maria Caterina Bramati³

1. Meteorological Institute, University of Hamburg, Hamburg Germany
- 5 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)

10 **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 have not yet been used for similar studies in this region](#). Two main datasets are
15 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
20 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.

25 Key words

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

1 Introduction

30 Extreme maximum temperature events have received much attention in recent years, because of the associated [dangerous impact on the increased](#) risk of mortality. [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
35 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).

40

In summer, Sindh becomes very hot and with the arrival of a monsoon the humidity increases in the region (Chaudhry and Rasul, 2004). [The extremely hot and humid conditions can have lethal effects, and can impact the human habitability of a region](#) (Pal and Eltahir 2015). The human body generally maintains the temperature

around 37°C. However, the human skin regulates at or below 35°C to release heat (Sherwood and Huber, 2010). Under high levels of the moisture content in the atmosphere, the human body cannot maintain the skin temperature below 35°C and can develop ailments like hyperthermia, heat strokes and cardiovascular problems. Hyperthermia is a condition where extremely high body temperature is reached, resulting from the inability of the body to get rid of the excess heat. It occurs mostly when temperature and relative humidity levels are extremely high at the same time. [Hyperthermia can occur even in the fittest human beings, if exposed to an environment where wet-bulb temperature is greater than 35°C for at least six hours.](#)

This study devotes special attention to Sindh because of its exposure to the frequent and intense temperature extremes in the past (Zahid and Rasul, 2012). This region is considered as one of the most vulnerable regions in Pakistan. Sindh stretches from 23.5° N – 28.5° N and 66.5°E - 71.1°E, and is bounded on the west by the Kirthar Mountains, to the north by the Punjab plains, [on the east by the Thar desert](#) and to the south by the Arabian Sea (Indian Ocean), [while in the center there is a fertile land](#) around [the](#) Indus river. The Indus river is the source of water for the agricultural lands. Cotton, wheat and sugar cane are grown on the left bank of the Indus and rice, wheat and gram on the right bank (Chaudhry and Rasul, 2004). Cotton is the cash crop of the country.

The climate in Sindh is arid and subtropical with less than 250 mm annual rainfall. The temperature frequently exceeds 45°C in summer (May-September) and the minimum average temperature recorded during winter (December- January) is 2°C. Table 1 shows the mean monthly climatic characteristics of the region from 1980-2010. Figure 1 shows the spatial distribution of all nine weather stations of Pakistan meteorological department, and the ERA Interim grid points close to the corresponding locations. High population density, limited resources, poor infrastructure and high dependence of the local agriculture on climatic factors, mark this region as highly vulnerable to the impacts of climate change.

The Intergovernmental Panel on Climate Change (IPCC) scenarios estimates for this region an increase in the surface temperature of the order of 4°C by the end of 2100. This may significantly reduce crop yields, and cause huge economic losses to the country (Islam et al., 2009; Rasul et al., 2012; IPCC, 2012, 2014). Furthermore, the risks of heat strokes, cardiac arrest, high fever, diarrhea, cholera and vector borne diseases might increase. Heat waves became more frequent and intense during 90's in Southern Pakistan. Zahid and Rasul (2010) reports a significant rise in the heat index and heat wave events longer than ten days in Sindh. The enhanced mortality rate related to the heat waves is a serious problem, and two obvious examples are the 1991 and the previously mentioned 2015 heat waves (Imtiaz and Rehman, 2015).

The analysis of extreme climatic events is a very active area of research in geosciences (Christidis et al., 2005, 2010; Tebaldi et al., 2006; Zwiers et al., 2011; Morak et al., 2011, 2013). In order to facilitate and standardize the analysis of extremes, the World Meteorological Organization (WMO) has suggested 27 specific climate indices, like the number of hot days, cold days, wet days, dry days, etc. (Tank et al., 2006; 2009, Frisch et al., 2002; Choi et al., 2009; Lustenberger et al., 2014). The investigation and analysis of such climate indices has now reached a high level of popularity.

Extreme value theory (EVT) [provides the statistical basis for](#) increasingly widespread [quantitative investigation of extremes](#) in climate studies (Coles, 2001, Zhang et al., 2004; Brown et al., 2008; Faranda et al., 2011; Acero et al., 2014). The peaks over threshold (POT) approach [aims at describing](#) the distribution of the exceedances [of the stochastic variable](#) of interest above a threshold. [Under very general conditions](#), the exceedances are

5 asymptotically distributed according to the Generalized Pareto Distribution (GPD). GPD has remarkable properties of universality when the asymptotic behavior is considered (Lucarini et al., 2016), while one can expect that the threshold level above which the asymptotic behavior is achieved depends on the characteristics of the analyzed time series. In particular, when looking at spatial fields, the threshold level depends on the geographical location.

10

[In this study, we have chosen to use the POT method to assess the temperature extremes in the Sindh region, because it provides a more efficient use of data and has better properties of convergence when finite datasets are considered \(Lucarini et al., 2016\). Additionally, we are here interested in investigating the actual tails of the distributions, so the POT approach is more appropriate.](#) It is applied for studying temperature extremes in

15 different regions of the world (Burgueño et al., 2002; Nogaj et al., 2006; Coelho et al., 2007; Ghil et al., 2011). However, to our knowledge, the POT method was never used to analyze the risk of temperature extremes in Sindh. The POT approach provides estimates of return periods and the return levels also for time ranges even longer than what is currently observed. This information and this predictive power can be beneficial for policy makers and other stakeholders. Note that this is exactly the kind of information planners need when, e.g.,

20 designing infrastructures that are deemed to last a very long time.

25

It is useful to consider two indicators of [extremely hot conditions](#): (1) temperature extremes T_{max} , and (2) Wet-bulb temperature extremes TW_{max} . [Up to now, there has been no investigation using EVT of the temperature extremes in southern Pakistan \(Sindh\).](#) Thus, considering the need and relevance of the information such a study is necessary and timely. Therefore, [we estimate the return levels of extreme daily maximum temperatures \$T_{max}\$ and daily maximum wet-bulb temperatures \$TW_{max}\$ over the 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-](#)

30 [analysis data are generally very good at replicating trends in percentile-based measures of temperature extremes \(Cornes and Jones, 2013\). But it is in principle not obvious that ERA 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 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](#)

35 [available.](#) Furthermore, a standard bias correction is applied on the ERA Interim data to assess whether the return levels of extremes are better predicted after the rescaling. As described below, given the shortness of the datasets, it is appropriate at this stage to analyze the extremes without taking into considerations possible long-term trends.

40

The paper is organized as follows. In Section 2, the statistical modeling of extremes using [the POT](#) method is briefly [described](#) along with a description of the data used. Section 3 presents the main results of the POT analysis of the meteorological station observations, ERA Interim, and bias corrected ERA Interim daily

maximum temperature T_{max} and wet-bulb temperature TW_{max} data at nine locations, viz. Jacobabad, Mohenjodaro, Rohri, Padidan, Nawabshah, Hyderabad, Chhor, Karachi, and Badin. The performance of the ERA Interim and bias corrected ERA Interim in comparison to observations is also described in Section 3. All computations and graphics in this work are done using the R free open source statistical software, using the packages ismev and extRemes (see www.R-project.org and R Development core team 2015). Section 4 summarizes the major findings of the study and concludes our work.

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 2). We select nine stations, which contain a negligible amount of missing values after 1980, and are suitable for the POT analysis. An additional criterion is that only those stations are chosen where no changes occurred in measuring instruments during the last 33 years (Brunetti et al., 2006). None of the station data shows gaps with a duration longer than two days, which are treated by replacing the missing value with the average of the two previous values.

The temperature data are discretized unevenly with intervals up to 1 degree Celsius. Deidda and Puliga (2006) [proposed a Monte Carlo approach for addressing this issue. They showed that finite resolution in precipitation data affects the convergence of parameter estimation in the extreme value analysis. They suggested generating many synthetic datasets by adding numerical noise to the original data, and then providing the best estimate of the parameters of the extreme value distributions by averaging over all the best fits obtained in each synthetic dataset. Following their suggestion, we produce high resolution data to compensate the effect of discretization and thus to improve the convergence of the estimator. In order to convert the temperature readings to higher resolution, we add a uniform random variable in the interval \[-0.5, 0.5\]. The main property of this noise is that \$\text{round}\(T+r\) = T\$, where \$T\$ is the temperature with 1-degree resolution and 'round' is the numerical function, which maps the interval \$\[T-0.5, T+0.5\]\$ to \$T\$. Thus, adding the noise does not perturb the information content of the observations. This procedure is applied to all temperature data, irrespective of the actual resolution, and replicated 100 times using a Monte Carlo approach. For each synthetic dataset, we perform the statistical best fit described later in the paper and then average the results.](#) We check the influence of this noise parameterization and find no significant bias in the return level estimates. [The advantage of adding a noise is to avoid the spurious statistical effects associated to the presence discrete values assigned to the temperature readings. Using the described bootstrap method we reduce such problem without biasing the data.](#)

2.2 ERA Interim re-analysis data

The gridded daily maximum temperature and relative humidity data of ERA Interim re-analysis is obtained from the ECMWF Public Datasets web interface (<http://apps.ecmwf.int/datasets/>). The ERA Interim is generated by the European Center for Medium range Weather Forecast (ECMWF) model with resolution $0.75^\circ \times 0.75^\circ$ (Dee et

al., 2011). The gridded data are then extracted at the closest grid points of all stations, for the period 1980-2013. The latitude and longitude of the ERA Interim stations are displayed in Table 2.

[The extreme temperatures analysis is restricted to the summer season \(May-September\) over a period of 33 years. We have tested that trends are not significant in such a short time interval.](#) One of the main requirements for performing the POT analysis is assuming the stationarity of the time series. Therefore, as in Bramati et al. (2014), the [Augmented Dickey Fuller](#) (ADF) test of stationarity is performed on all time series (Dickey and Fuller, 1979). [In all cases we find no sign of long-term correlations in the data. Short-term correlations \(daily time scale\) typically lead to clusters of extreme values and are studied by computing the extremal index \$\theta\$ in all time series and treated using the associated standard declustering technique \(see more details in Section 2.4\).](#)

2.3 Wet-bulb temperature calculations

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

$$TW = T \operatorname{atan}(\alpha_1 \sqrt{RH + \alpha_2}) + \operatorname{atan}(T + RH) - \operatorname{atan}(RH + \alpha_3) + \alpha_4 (RH)^{\frac{3}{2}} \operatorname{atan}(\alpha_5 RH) - \alpha_6 \quad (1)$$

where TW is the wet-bulb temperature [°C], T is the temperature [°C], and RH is the relative humidity [%]. This relationship is based on an empirical fit, as in Stull (2011), where the coefficient values are $\alpha_1 = 0.151977$, $\alpha_2 = 8.313659$, $\alpha_3 = -1.676331$, $\alpha_4 = 0.00391838$, $\alpha_5 = 0.023101$, and $\alpha_6 = 4.686035$. Equation (1) covers a wide range of relative humidity and air temperatures with an accuracy of 0.3°C.

2.4 Peaks over Threshold

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

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

The extremal index θ can be estimated in two separate ways. Here, we apply the ‘intervals estimator’ automatic declustering by Ferro and Segers (2003). A positive aspect of this method is that it avoids the subjective choice of cluster parameters. The main ingredient is the use of an asymptotic result for the times between threshold exceedances. The exceedance times are split into two types, a set of vanishing intra-exceedance times within the clusters, and an exponentially distributed set of inter-exceedance times between clusters. The method is iterative, starting with largest return times and stops when a limit for the inter-exceedance times is reached. The standard errors of the estimated parameters is obtained by a bootstrap procedure. [In this study, once we select appropriate value for the threshold \(see below\) the extremal index value is \$\leq 0.5\$ in all the considered time series. Therefore, it is necessary to decluster the extremes by choosing the largest event in each cluster, before fitting it to the GPD.](#)

[As mentioned before, we use as statistical model for the exceedances](#) over threshold [the](#) Generalized Pareto Distribution (GPD), which is characterized by two parameters, the shape ξ and the scale σ . The GPD for exceedances $x - u$ of a random variable x reads as

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

where u is the threshold. The shape parameter ξ determines the tail behavior while the scale parameter σ measures the variability. For a negative shape parameter, $\xi < 0$, the distribution is bounded ([Weibull](#) distribution), for vanishing shape parameter, $\xi = 0$, the distribution is exponential, and for a positive shape parameter, $\xi > 0$, the distribution has no upper bound (Pareto distribution).

In particular, for a negative shape parameters $\xi < 0$ the GPD has the upper bound

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

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

where A_{max} is an absolute maximum (Lucarini et al., 2014). [In general, the best estimate for the two parameters shape \$\xi\$ and scale \$\sigma\$ depend on the threshold \$u\$ \(Coles, 2001\).](#) The choice of the optimal threshold for performing statistical inference from a time series is crucial. [Choosing a very large value for \$u\$ reduces the number of exceedances to a few values, inflating the variance of the estimators, so that the analysis is unlikely to yield any useful results.](#) On the other hand, [choosing a too small value for \$u\$ would violate the asymptotic nature of the model, with a possible biased estimation and wrong model selection \(Coles, 2001\), see details later in Section 3.1.](#) The shape ξ , the scale σ and the return levels are estimated using the Maximum Likelihood Estimator (MLE) using the R software (R Development core team 2015), which also provides [an estimate of the standard error of the estimates.](#)

[Additionally, we wish to investigate the \$N\$ - years return levels \$x_N\$, which are exceeded on the time scale of \$N\$ years \(Coles, 2001\) and can be expressed as](#)

$$x_N = u + \frac{\sigma}{\xi} \left[(N n_y \zeta_u)^\xi - 1 \right], \quad (4)$$

where N represents the return period in years, n_y is the number of observations per year, ζ_u is the probability of an individual observation exceeding the threshold u , the shape parameter is ξ and the scale parameter is σ .

5 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 on 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 considerably the skill of the ERA Interim dataset in describing extreme events.](#) The bias corrected ERA Interim time series x is expressed as

$$x = \bar{z} + \frac{y_{ERA} - \bar{y}}{\sigma_y} \sigma_z \quad (5)$$

15 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.

3. Results and Discussion

3.1 Threshold Selection

20 The threshold selection is the first step in a POT analysis. [One needs to test whether the asymptotic regime is reached, i.e. whether we are choosing true extremes. This can be investigated by checking whether the best fits of the shape parameter \$\xi\$ and the modified scale parameter \$\sigma^* = \sigma u - \xi u\$ are stable with respect to increases in the chosen value of \$u\$ \(Sacrott and MacDonald, 2012\).](#) 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.](#) Figure 2 shows the parameter stability plots of the T_{max} [reading for Karachi](#), as an example to explain the threshold selection procedure.

30 [In addition to diagnostic plots of the modified scale parameter \$\sigma^*\$ and the shape parameter \$\xi\$, 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^\circ\text{C}\$, indicating \$u = 36\$ as the most suitable threshold for the POT analysis \(Figure 3\). We observe that the 90% quantile is an appropriate threshold for all the station data, as well as the ERA interim datasets, and for both \$T_{max}\$ and \$TW_{max}\$.](#)

35 3.2 GPD Fit

The goodness of fit is evaluated by Quantile-Quantile (Q-Q) plots and hypothesis testing. The Q-Q plot analysis is performed for the stations observed, the ERA Interim, the bias corrected ERA Interim daily T_{max} and TW_{max} .

The Q-Q plots of the observed T_{max} show that the GPD fits well in most stations. However, in a few stations like Jacobabad, Mohenjo-daro, Padidan and Chhor the empirical values show slight deviation from the modeled values. In spite of minor deviations at some stations, still most of the exceedances are well fitted by the model. The Q-Q plots of the observed TW_{max} also fits well to the model in all stations.

5

The Q-Q plots of the ERA Interim T_{max} and TW_{max} reveals substantial differences with respect to the GPD. The empirical values of the higher quantiles are deviating from the theoretical quantiles in all stations. However, [if the higher quantiles are disregarded, then stations like Jacobabad, Mohenjo-daro, Rohri, Padidan, Nawabshah, Chhor, and Badin fits very well with the model. The Q-Q plots of the bias corrected ERA Interim \$T_{max}\$ and \$TW_{max}\$ show better results than the ERA Interim.](#) We notice that the T_{max} of the ERA Interim and bias corrected ERA Interim fits better than the TW_{max} if the higher quantiles are ignored, [indicating the bias procedure is, as expected, unable to treat correctly the statistics of the largest events.](#)

10

In order to assess the goodness-of-fit, we apply the Kolmogorov-Smirnov (K-S) test and Anderson-Darling (A-D) test to the data of meteorological stations, ERA Interim, bias corrected ERA Interim T_{max} and TW_{max} . The p-values indicate a good performance of the fit procedure. Table 3 [shows](#) the results of the K-S and A-D statistics of the T_{max} and TW_{max} in all the data sets.

15

3.3 Parameter Estimates

Here, we analyze the shape parameter ξ , the scale parameter σ , and threshold u for all considered datasets. The standard errors of the shape ξ and the scale σ parameters are given in Table 4. The spatial distribution of the shape parameter ξ and the scale parameter σ of the GPD in Sindh are shown in Figure 4. The shape parameters ξ are negative in all datasets at all stations. This is hardly surprising, as meteorological and physical processes make sure that the temperature cannot grow locally without control. Figure 4 displays the bias corrected ERA Interim results only. The observed T_{max} shape parameters ξ are between -0.418 to -0.223, and for TW_{max} within -0.323 to -0.177. The bias corrected ERA Interim T_{max} shape parameters ξ range from -0.305 to -0.002, and TW_{max} are between -0.18 to -0.01. [The agreement in the values of the shape parameters in the observations and simulations means that the ERA dataset captures an important aspect of extremal behavior. This is in principle a non-trivial result, as reanalysis are constructed in such a way that typical conditions are well reproduced.](#)

20

25

30

The scale parameters σ of the observed T_{max} range from 2.08 to 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 between 1.00 - 1.95, and TW_{max} in 0.74 -1.75. We observe a difference in the scale parameters of both the observed, ERA Interim T_{max} and TW_{max} . We find that, [unsurprisingly,](#) the scale parameters of the bias corrected ERA Interim data are much closer to those estimated for T_{max} and TW_{max} using the station data. In 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 1.40 – 2.40 (Figure 4). All the temperature scale parameters are in degree Celsius.

35

3.4 Absolute Maxima

40

Once the shape parameters ξ , the scale parameters σ , and the thresholds u are [determined](#), it is possible to compute the theoretical absolute maxima using Eq. (3) (Section 2.4). Theoretical absolute maxima can be compared with the observed ones for each station to better [understand whether our fits are in agreement with the observed data](#). The daily maximum temperature T_{max} and the maximum wet-bulb temperature TW_{max} (station data, the ERA Interim, and the bias corrected ERA Interim) have negative shape parameters ξ at all stations. This means that according to Eq. (2) in section 2.4, the probability distribution function (pdf) is bounded by the maximum values. These maximum values are the theoretical upper limits predicted by the GPD fit. The analysis shows that the observed absolute maxima T_{max} and TW_{max} at all stations of the three data sets are below the theoretical absolute maximum, as expected (Figure 5). This gives us confidence on the quality of our fit. The following piece of information can also be derived: [assume](#) that one observes in the future an extreme event larger than the maximum inferred in the present dataset; this may suggest some non-stationarity in the most recent portion of the dataset.

3.5 Return Levels

The return levels (RLs) are computed considering various return periods (2, 5, 10, 20, 50, 100-year). The return level plots of the stations observed, the ERA Interim, the bias corrected ERA Interim daily maximum temperature T_{max} and daily maximum wet-bulb temperature TW_{max} are displayed in Figures 6 and 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](#) (Figure 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](#) (Figure 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 the 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](#) (Figures 6 and 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 (Figures 8 and 9), as a detailed spatial overview of the temperature extremes in Sindh might be of interest to the policy makers.

4. Summary and Conclusion

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

The peaks over threshold method is applied to the daily T_{max} and TW_{max} data of nine [stations](#) and to the corresponding nearest ERA Interim temperature data. Standard declustering technique is applied to all time series to achieve the independence assumption of extremes. [After testing the asymptotic statistical properties](#), the 90% quantile is [found to be](#) appropriate threshold choice for [all the](#) weather stations, the ERA Interim and the bias corrected ERA Interim maximum temperature and wet-bulb temperature. A Generalized Pareto Distribution (GPD) is fit to both T_{max} and TW_{max} for all three datasets. The results conclude that the shape parameter ξ is negative at all stations. The scale parameter σ estimated on weather station temperatures is much closer to the bias corrected ERA Interim estimates than the original ERA Interim data ones. The theoretical absolute maxima of the time series are higher than the observed absolute maxima in all stations. The Q-Q plots are used to assess the GPD fit, which results to be acceptable for both T_{max} and TW_{max} station data as compared to the ERA Interim data. However, the bias corrected ERA Interim [data](#) shows improved GPD fits than ERA Interim [data](#).

[The](#) return levels (RLs) of T_{max} and TW_{max} are estimated for the 2, 5, 10, 25, 50, 100-year return periods in all datasets. The RLs of T_{max} estimated using the meteorological station temperatures are greater than 50°C in Jacobabad, Mohenjo-daro, Padidan, Nawabshah, and greater than 45°C in Rohri, Hyderabad, Chhor, Karachi and Badin. While the RLs of TW_{max} in station data are larger than 35°C in the entire Sindh, when using ERA Interim temperatures, they are estimated as greater than 45°C in Northern Sindh and greater than 40°C in southern Sindh. The differences in the RLs using the two datasets are 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, but it provides a good representation of their spatial fields.

The bias corrected ERA Interim T_{max} and TW_{max} gives 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, [statistically](#) differences remain [in most cases](#). [Therefore, one must use more advanced bias correction method for analyzing extremes precisely.](#)

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 involve 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.

The 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. In summer, the temperature and humidity increase to an extent that there are high chances of a rapid pests spread in the crops. Sindh produces cotton, wheat, rice, mango, banana, and dates, so a correct estimate of temperature extremes is very important.

This clarifies that 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.

This paper contains beneficial information regarding the assessment of the temperature extremes in Sindh, which could help 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. While the stationary analysis presented here has already relevance in terms of impacts for the public and private sector as it fills a research gap, and is statistically motivated by the short duration of the observational dataset (33 years). Indeed, it seems relevant to investigate time dependency in the temperature extremes. We might consider using the centennial NCEP reanalysis (Compo et al., 2011) and using suitable bias correction procedures. 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.

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.

References

- 5 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.
- 10 Bramati, M.C., Tarragoni, C., Davoli, L., Raffi, R., Extreme Rainfall in Coastal Metropolitan Areas of Central Italy: Rome and Pescara case studies. *Geografia Fisica e Dinamica Quaternaria*, 37, 3-13, 2014.
- 15 Brunetti, M., Maugeri, M., Monti, F. and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogenized instrumental time series, *J. Climatol.*, 26(3), 345–381, doi:10.1002/joc.1251, 2006.
- 20 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 homogenized instrumental time series, *J. Climatol.*, 26(3), 345–381, doi:10.1002/joc.1251, 2006.
- 25 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.
- 30 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.
- 35 Choi, G., Collins, D., Ren, G., Trewin, B., Baldi, M., Fukuda, Y., Afzaal, M., Pianmana, T., Gomboluudev, P., Huong, P. T. T., Lias, N., Kwon, W.-T., Boo, K.-O., Cha, Y.-M. and Zhou, Y.: Changes in means and extreme events of temperature and precipitation in the Asia-Pacific Network region, 1955-2007, *Int. J. Climatol.*, 29(13), 1906–1925, doi:10.1002/joc.1979, 2009.
- 40 Christidis, N., Stott, P. A., Brown, S., Hegerl, G. C. and Caesar, J.: Detection of changes in temperature extremes during the second half of the 20th century, *Geophys. Res. Lett.*, 32, L20716, doi:10.1029/2005GL023885, 2005.
- Christidis, N., Stott, P. A., Zwiers, F. W., Shiogama, H. and Nozawa, T.: Probabilistic estimates of recent changes in temperature: A multi-scale attribution analysis, *Clim. Dyn.*, 34, 1139–1156, doi:10.1007/s00382-009-0615-7, 2010.
- 45 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.
- 50 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.
- 55 Cornes, R. C., and P. D. Jones, 2013: 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.
- 60 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.
- 65 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.,

- 5 Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavalato, 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.
- 10 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.
- 15 Ferro, C. A. T. and Segers, J.: Inference for clusters of extreme values, *J. R. Stat. Soc. B*, 65(2), 545–556, 2003.
- Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Tank Klein, A. M. G. and Peterson, T.: Observed coherent changes in climatic extremes during the second half of the twentieth century, *Clim. Res.*, 19,193-212, doi:10.3354/cr019193, 2002.
- 20 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.
- 25 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.
- 30 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.
- 35 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
- IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, edited by C. B. Field, V. Barros, T. F. Stocker, and Q. Dahe, Cambridge University Press, Cambridge.,582, 2012.
- 40 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.
- 45 Klein Tank, A. M. G., Peterson, T. C., Quadir, D. A., Dorji, S., Zou, X., Tang, H., Santhosh, K., Joshi, U. R., Jaswal, A. K. and Kolli, R. K.: Changes in daily temperature and precipitation extremes in central and south Asia, *J. Geophys. Res.*, 111, D16105, doi:10.1029/2005JD006316, 2006.
- Klein Tank, A.M.G., Zwiers, F.W., Zhang, X.: Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation., 2009.
- 50 Leadbetter, M. R.: Extremes and local dependence in stationary sequences, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*,65, 291-306, doi:10.1007/BF00532484, 1983.
- 55 Loynes, R. M.: Extreme Values in Uniformly Mixing Stationary Stochastic Processes, *Ann. Math. Stat.*, 36(3), 993–999, doi:10.1214/aoms/1177700071, 1965.
- 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.
- 60 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.
- 65 Lustenberger, A., Knutti, R. and Fischer, E. M.: The potential of pattern scaling for projecting temperature-related extreme indices, *Int. J. Climatol.*, 34(1), 18–26, doi:10.1002/joc.3659, 2014.
- Morak, S., Hegerl, G. C. and Kenyon, J.: Detectable regional changes in the number of warm nights, *Geophys. Res. Lett.*, 38, L17703, doi:10.1029/2011GL048531, 2011.
- 70 Morak, S., Hegerl, G. C. and Christidis, N.: Detectable changes in the frequency of temperature extremes, *J. Clim.*, doi: 26, 1561-1574, 10.1175/JCLI-D-11-00678.1, 2013.
- Newell, G. F.: Asymptotic Extremes for β -Dependent Random Variables, *Ann. Math. Stat.*, 35(3), 1322–1325,

doi:10.1214/aoms/1177703288, 1964.

5 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, *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

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

15 Pal, J. S. and Eltahir, E. A. B.: Future temperature in southwest Asia projected to exceed a threshold for human adaptability, *nature cliamte 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.

20 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, Vienn, Austria, 2015.

25 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.

30 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 uncertainty quantification*, *Revstat – Stat. J.*, 10(1), 33–60, 2012.

35 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.

40 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.

45 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.

50 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.

Zwiers, F. W., Zhang, X. and Feng, Y.: *Anthropogenic influence on long return period daily temperature extremes at regional scales*, *J. Clim.*, 24, 881-892, doi:10.1175/2010JCLI3908.1, 2011.

55

60

65

70

[Table 1. Monthly](#) mean climatic characteristics of all nine stations from 1980-2010.

Stations	Mean Temperature (°C)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Jacobabad	15.2	18.2	24	30.5	35.6	37	34.8	33	31.4	27.8	22.3	16.7	27
Mohenjo-daro	13.9	16.7	23	29.1	34.1	35	33.9	32.9	30.9	26.7	21.1	15.9	25.9
Rohri	15.6	18.2	23.6	29.8	34.5	35.6	33.9	32.3	31.2	27.6	22.1	16.9	26.4
Padidan	14.8	17.7	23.5	29.9	34.4	35.5	33.7	32.1	31	27.5	22.4	16.4	26.5
Nawabshah	15.4	18	24	29.8	34.5	35.6	34	32.3	31.5	28	22.4	16.9	26.7
Hyderabad	18	21	26.2	30.9	33.3	34	32.4	31.1	31	29.6	24.8	19.6	27.6
Chhor	16.5	19.5	25	30.1	33.5	33.7	31.6	30.1	30.1	28.2	22.6	17.9	26.3
Karachi	18.6	21.2	25.4	28.9	31.1	31.9	30.5	29.2	29.5	28.9	24.6	20.4	26.4
Badin	17.5	20.5	25.8	30.1	32.6	32.8	31	29.6	29.6	28.7	24	19	26.6

Stations	Minimum Temperature (°C)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Jacobabad	7.9	10.9	16.6	22.4	27.4	29.8	29.3	28.4	26.3	20.5	14.3	8.9	19.9
Mohenjo-daro	4.7	7.9	13.3	18.9	24	27.4	27.9	27	24.7	18.2	11.8	7.3	17.3
Rohri	8.3	10.8	15.9	21.7	26.1	27.7	27.1	26	24.4	19.9	14.2	9.6	18.7
Padidan	6.5	8.9	14.5	20.2	24.7	27	26.9	25.8	23.7	18.3	12.4	7.6	17.8
Nawabshah	6.3	8.7	14.2	19.4	24.6	27.3	27.2	25.9	23.8	18.4	12.4	7.8	17.9
Hyderabad	11.4	13.9	18.8	22.8	26.1	27.9	27.6	26.5	25.4	22.5	17.4	13	21.1
Chhor	5.9	8.9	14.8	20.3	24.8	26.9	26.5	25.3	23.9	18.7	11.8	7	17.6
Karachi	11.5	14	18.6	23	26.6	28.3	27.6	26.3	25.6	21.9	16.8	12.7	20.7
Badin	9.9	12.6	17.9	22.3	25.7	27.6	27.1	26	25	22.1	16.5	11.4	20.2

Stations	Maximum Temperature (°C)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Jacobabad	22.6	25.6	31.4	38.6	43.9	44.4	40.2	37.6	36.8	35.1	30.3	24.4	34.1
Mohenjo-daro	23.1	26.2	32.1	38.7	43.8	44.2	40.9	38.7	37.5	35.2	30.5	24.8	34.5
Rohri	22.6	25.6	31.2	38.1	43	43.5	40.5	38.3	37.8	35.2	30	24.3	34
Padidan	23.1	26.4	32.2	39.4	43.9	44.1	40.6	38.4	38.3	36.3	31.1	25.3	34.8
Nawabshah	24.5	27.9	33.8	40.2	44.2	43.9	40.7	38.8	39	37.7	32.3	26.1	35.5
Hyderabad	24.7	28.1	33.7	38.8	41.3	40	37.2	35.6	36.3	36.7	31.9	26.2	34.1
Chhor	26.9	29.9	35.2	40	42	40.6	36.8	34.9	36.3	37.6	33.5	28.7	35
Karachi	26.3	28.4	32.2	34.7	35.5	35.4	33.3	32.1	33.2	35.5	32.5	28.2	32
Badin	25.2	28.3	33.7	37.8	39.4	37.9	34.9	33.2	34.2	35.2	31.4	26.5	32.9

5

[Table 2. 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 3. Results of the Kolmogorov-Smirnov Goodness of fit test and Anderson-Darling test between empirical and GPD fits.

Observed T_{max}										
Test Statistics	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.947	0.340	0.996	0.139	0.941	0.385	0.928	0.306	0.666
Anderson Darling	Equality of probability distribution	0.553	0.978	0.654	0.857	0.157	0.649	0.233	0.869	0.145
ERA Interim T_{max}										
Test Statistics	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.169	0.125	0.553	0.456	0.322	0.187	0.419	0.456	0.332
Anderson Darling	Equality of probability distribution	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	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.452	0.4729	0.197	0.489	0.269	0.137	0.158	0.243	0.312
Anderson Darling	Equality of probability distribution	0.352	0.315	0.235	0.270	0.335	0.289	0.216	0.390	0.227
Observed TW_{max}										
Test Statistics	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.981	0.111	0.341	0.226	0.457	0.545	0.441	0.385	0.211
Anderson Darling	Equality of probability distribution	0.623	0.745	0.587	0.884	0.199	0.123	0.789	0.669	0.473
ERA Interim TW_{max}										
Test Statistics	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.712	0.564	0.955	0.425	0.258	0.134	0.856	0.497	0.222
Anderson Darling	Equality of probability distribution	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	Null Hypothesis	P-value								
		JAC	MJD	RHI	PDN	NWS	HYD	CHR	KHI	BDN
Kolmogorov Smirnov	Equality of probability distribution	0.268	0.688	0.127	0.372	0.268	0.229	0.591	0.582	0.478
Anderson Darling	Equality of probability distribution	0.373	0.484	0.278	0.432	0.306	0.283	0.365	0.445	0.483

5

10

Table 4. Estimated parameters shape ξ , scale σ and standard error $\Delta\xi$ of all the data sets.

Station observed T_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.3875	-0.2550	-0.4182	-0.3261	-0.3323	-0.3292	-0.3108	-0.2225	-0.3292
Standard Error $\Delta\xi$	0.0317	0.0226	0.0226	0.0218	0.0208	0.0312	0.0371	0.0341	0.0312
Scale σ	2.7540	2.0819	2.3510	2.2144	2.1391	2.2286	2.5629	2.5685	2.2286
Standard Error $\Delta\sigma$	0.1421	0.1040	0.1075	0.1076	0.1031	0.1166	0.1462	0.1444	0.1166
ERA Interim T_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.1959	-0.1788	-0.2076	-0.2185	-0.2135	-0.3380	-0.2850	-0.0376	-0.2514
Standard Error $\Delta\xi$	0.0320	0.0348	0.0343	0.0287	0.0265	0.0316	0.0337	0.0508	0.0371
Scale σ	1.4643	1.3230	1.3440	1.5045	1.5630	2.0656	1.8497	1.3303	2.0410
Standard Error $\Delta\sigma$	0.0798	0.0739	0.0741	0.0788	0.0788	0.1082	0.0949	0.0908	0.1153
Bias Corrected ERA Interim T_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.1959	-0.1788	-0.2076	-0.2185	-0.2135	-0.3380	-0.2850	-0.0376	-0.2514
Standard Error $\Delta\xi$	0.0320	0.0348	0.0343	0.0287	0.0265	0.0316	0.0337	0.0508	0.0371
Scale σ	1.9834	1.7918	1.8205	2.0382	2.1164	2.7980	2.3081	1.8016	2.7636
Standard Error $\Delta\sigma$	0.1081	0.1001	0.1004	0.1068	0.1068	0.1467	0.1233	0.1229	0.1562
Station observed TW_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.1769	-0.1860	-0.2150	-0.2157	-0.2164	-0.3231	-0.2423	-0.2190	-0.1867
Standard Error $\Delta\xi$	0.0383	0.0354	0.0347	0.0442	0.0266	0.0269	0.0347	0.0368	0.0322
Scale σ	2.7590	2.0454	1.9600	2.0780	1.8572	2.3724	2.5126	2.3375	1.9032
Standard Error $\Delta\sigma$	0.1596	0.1146	0.1084	0.1289	0.0938	0.1191	0.1380	0.1328	0.1055
ERA Interim TW_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.0896	-0.0946	-0.0687	-0.1257	-0.1583	-0.1771	-0.0902	-0.0194	-0.1733
Standard Error $\Delta\xi$	0.0379	0.0293	0.0327	0.0342	0.0313	0.0377	0.0357	0.0359	0.0378
Scale σ	1.2879	1.2437	1.2311	1.4408	1.6104	1.6499	1.3423	0.6801	1.7886
Standard Error $\Delta\sigma$	0.0748	0.0660	0.0676	0.0804	0.0875	0.0959	0.0760	0.0398	0.1028
Bias Corrected ERA Interim TW_{max}									
Estimates	JCB	MJD	RHI	PDN	NWB	HYD	CHR	KHI	BDN
Shape ξ	-0.08961	-0.0946	-0.06870	-0.12570	-0.15831	-0.17711	-0.09017	-0.01942	-0.17332
Standard Error $\Delta\xi$	0.03786	0.02931	0.03275	0.03424	0.03134	0.03767	0.03571	0.03593	0.03782
Scale σ	1.35674	1.64650	1.75852	1.49477	1.52013	2.05281	2.14609	1.39943	2.15299
Standard Error $\Delta\sigma$	0.07878	0.08736	0.09651	0.08347	0.08254	0.11924	0.12145	0.08193	0.12370

5

10

15

20

25

30

35

40

45

50

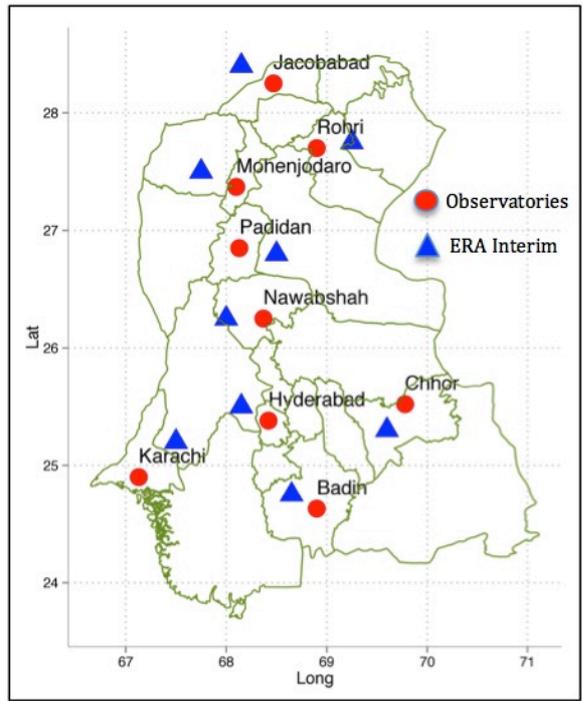


Figure 1: Study Domain (23.5 – 28.5° N , 66.5- 71.1°E)

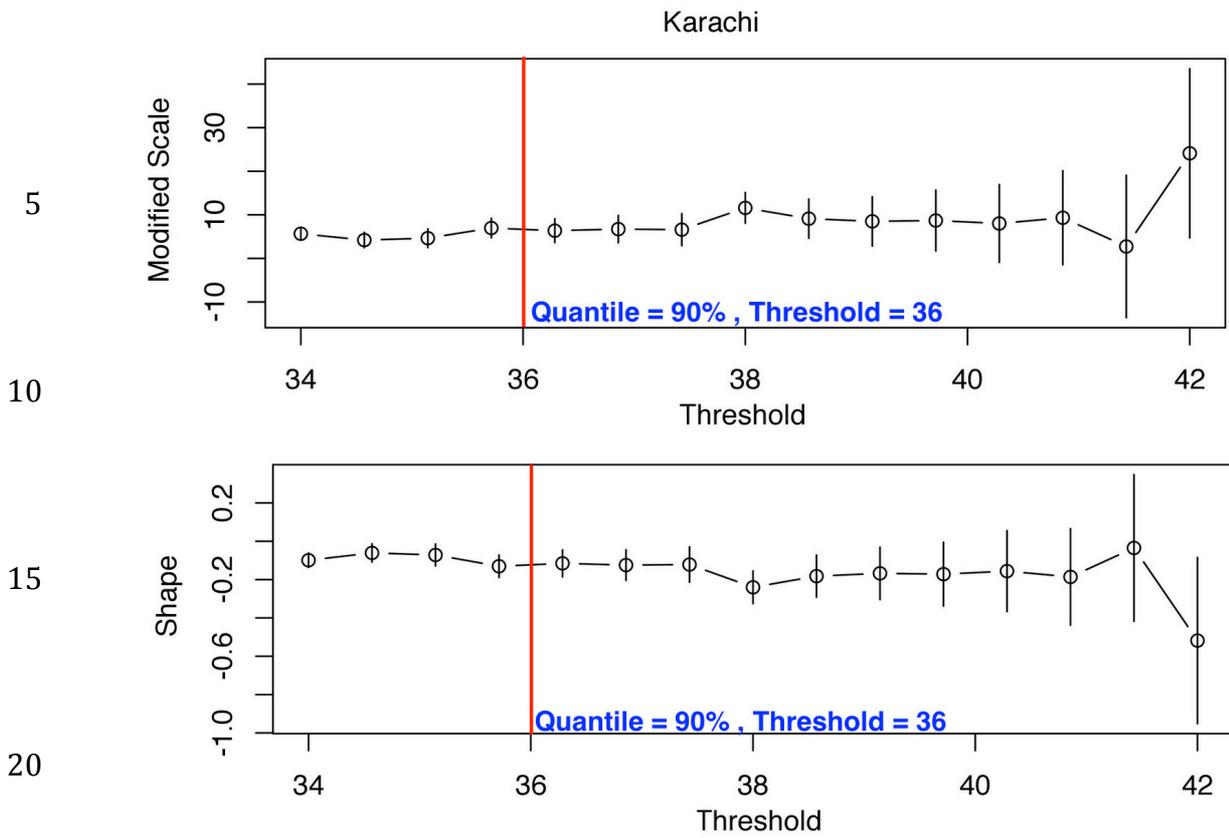


Figure 2. Modified scale (σ^*) and shape parameter (ξ) of the observed T_{max} ($^{\circ}\text{C}$) Karachi. The red vertical lines represent the selected threshold according to the station quantiles.

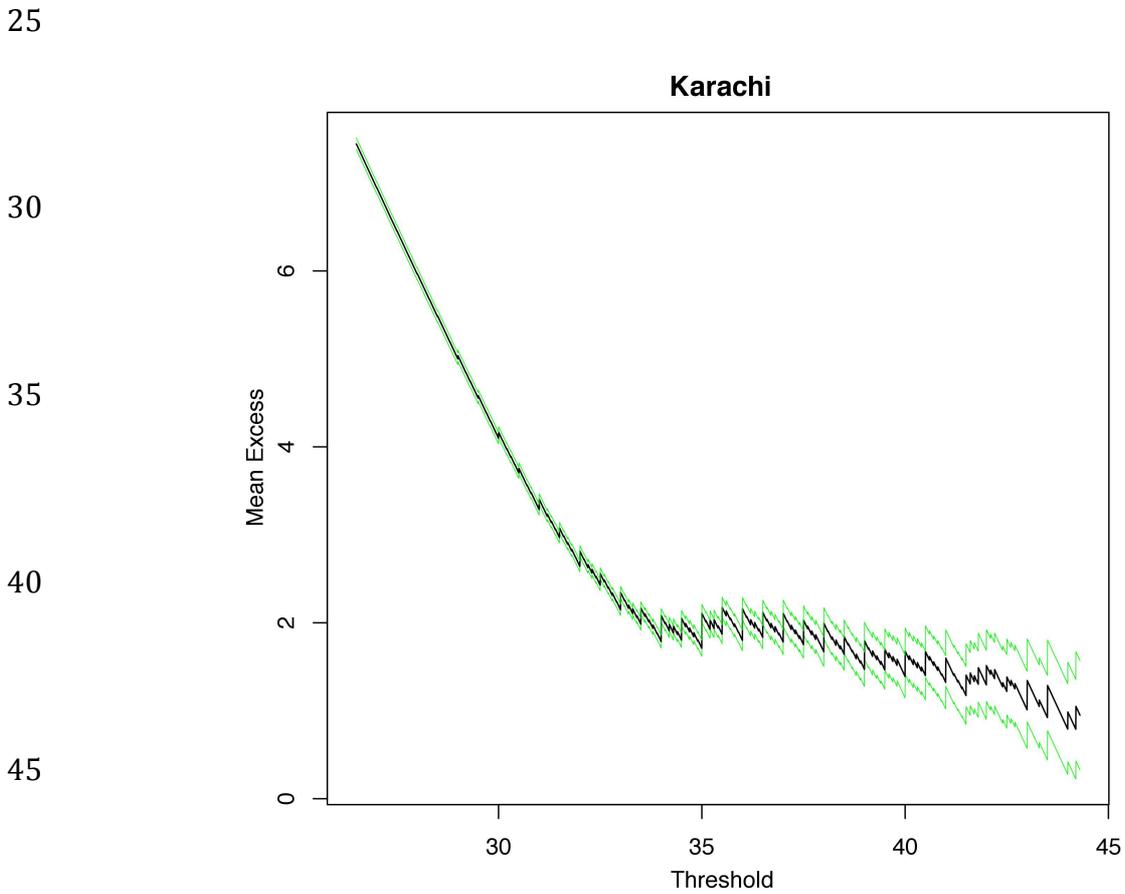


Figure 3. Mean residual life plot of the station observed T_{max} ($^{\circ}\text{C}$) Karachi.

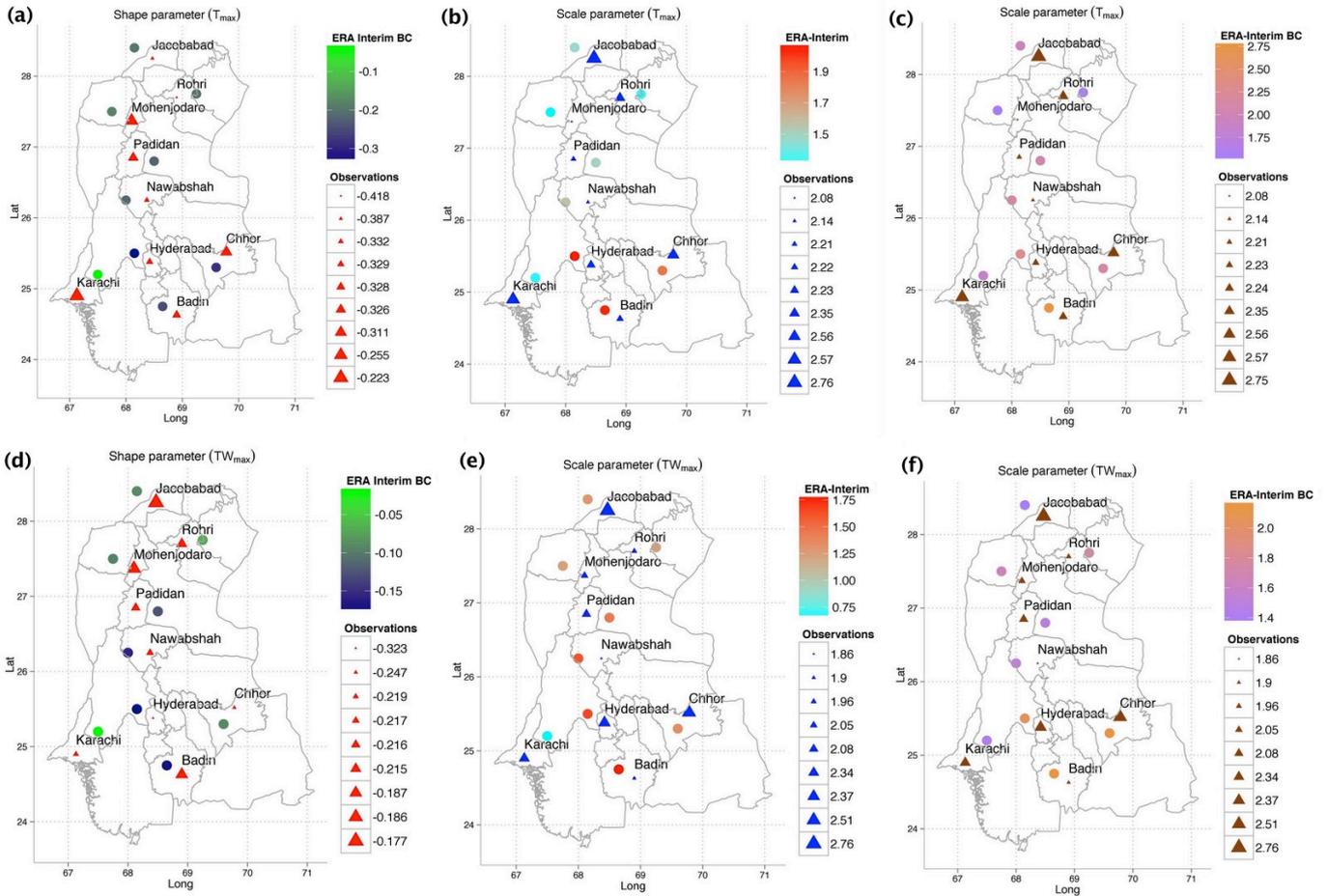


Figure 4. Spatial distribution of the shape parameters ξ and scale parameters σ of the station observed, ERA Interim, and bias corrected ERA Interim T_{max} (upper panel) and TW_{max} (lower panel) in [degree Celsius](#).

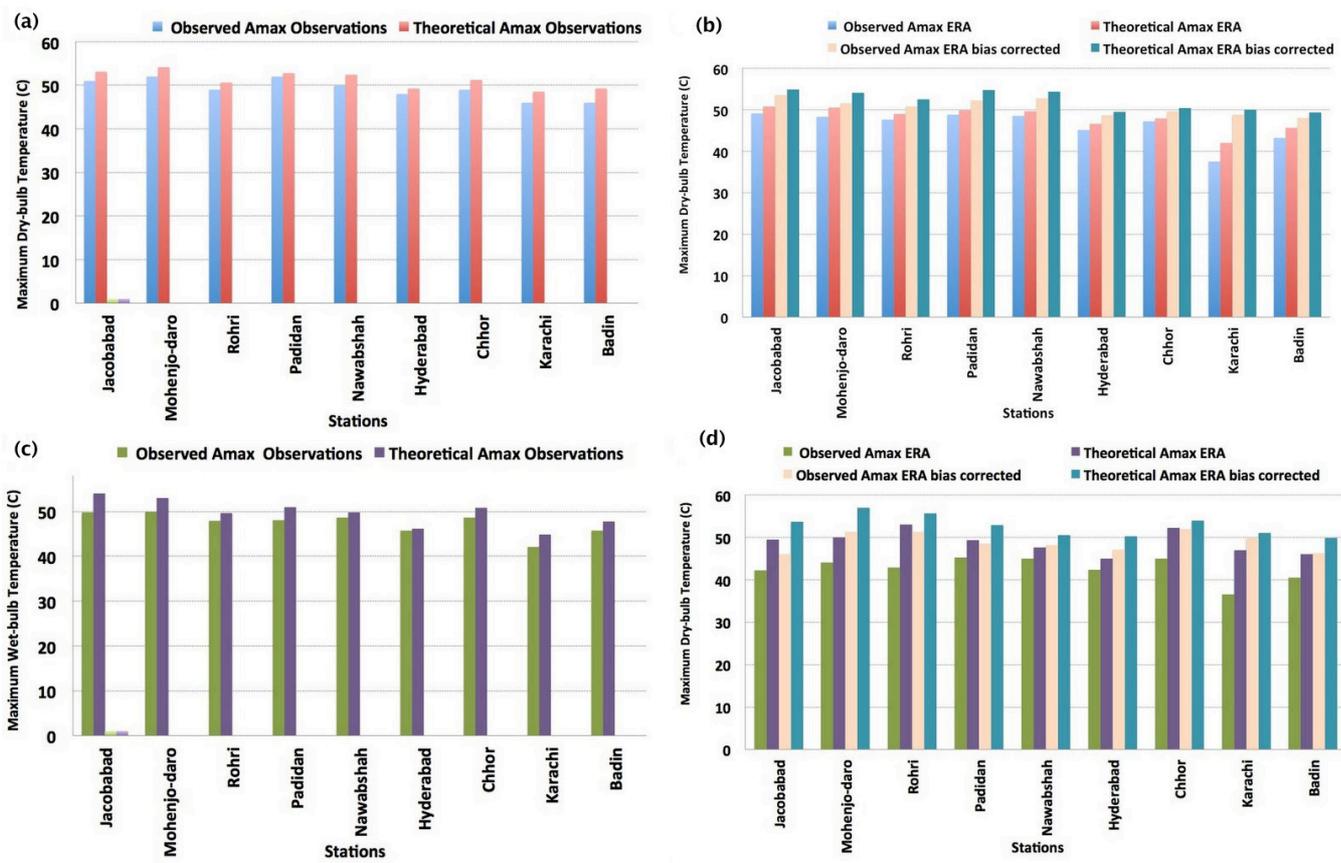


Figure 5. Absolute maxima A_{max} in degree Celsius (a) station observed T_{max} (b) ERA Interim and bias corrected ERA Interim T_{max} (c) station observed TW_{max} (d) ERA Interim and bias corrected ERA Interim TW_{max} .

5

10

15

20

25

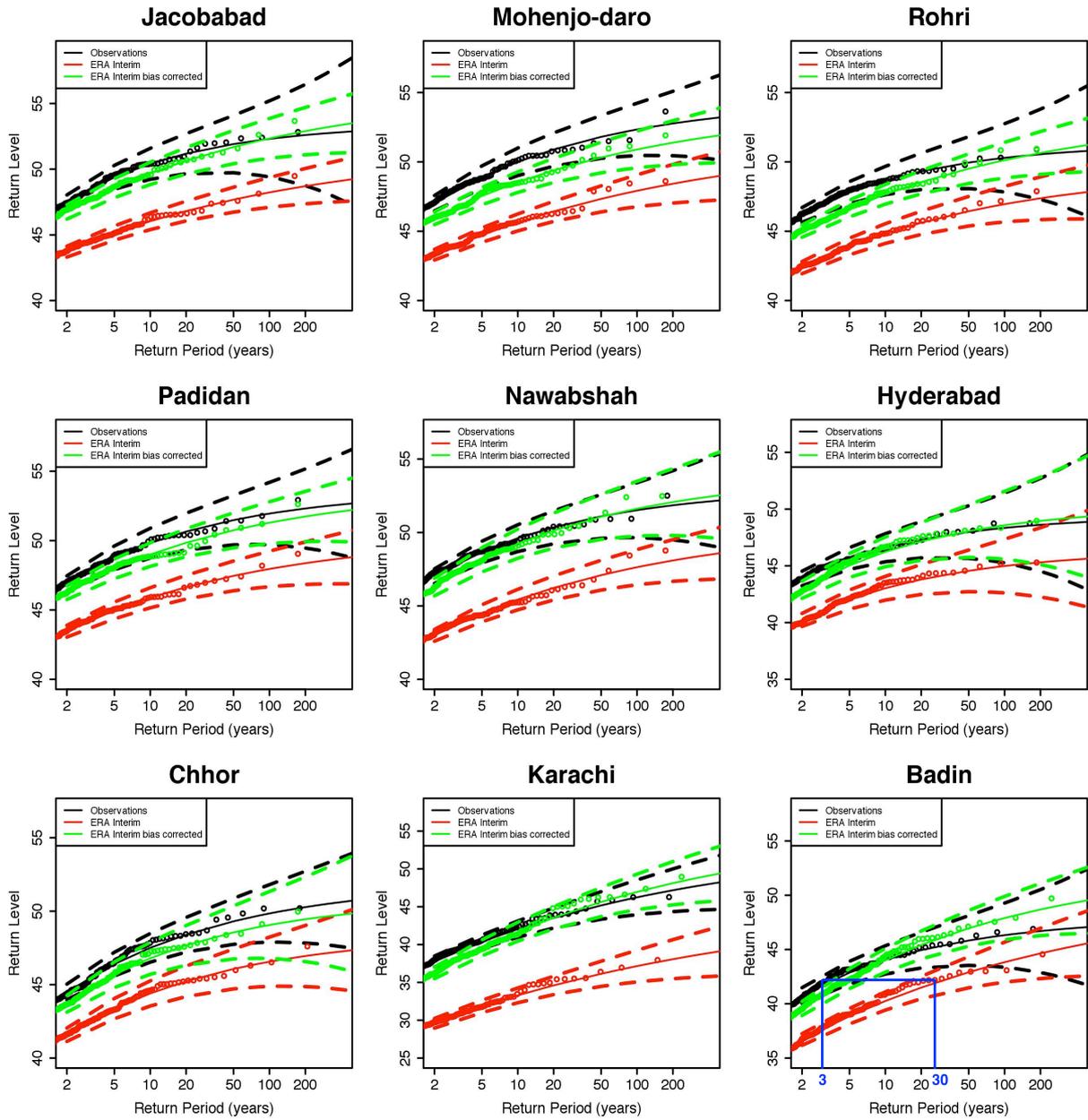


Figure 6. Return level plots of the station observed T_{max} (black) , ERA Interim T_{max} (red) and bias corrected ERA Interim T_{max} (green) in degree Celsius. The blue line is to show a difference in the observed and ERA Interim RLs.

5

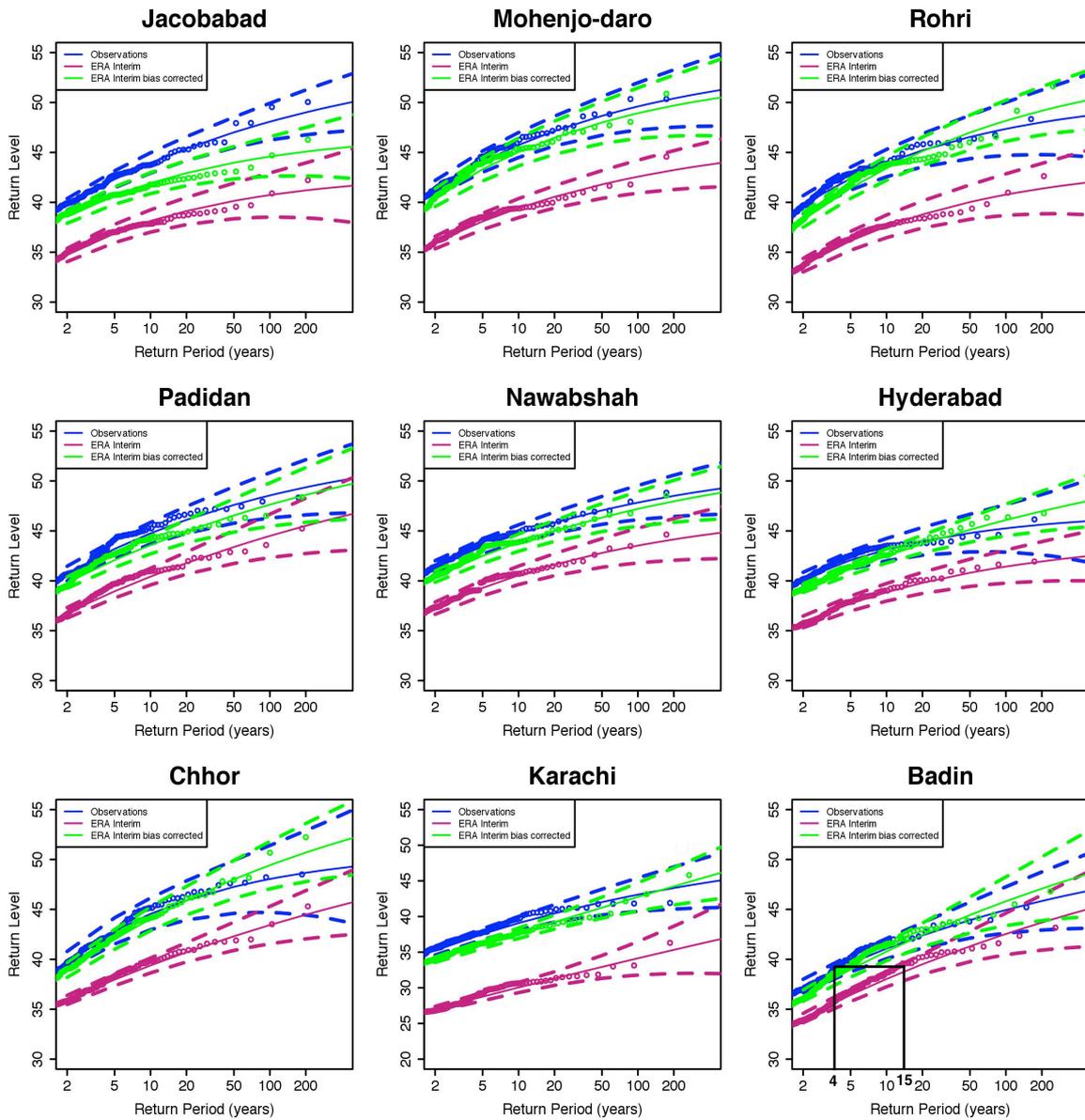


Figure 7. Return level plots of the station observed TW_{max} (blue), ERA Interim T_{max} (pink) [in degree Celsius](#), and bias corrected ERA Interim T_{max} (green). The black line is to show a difference in the observed and ERA Interim RLs.

5

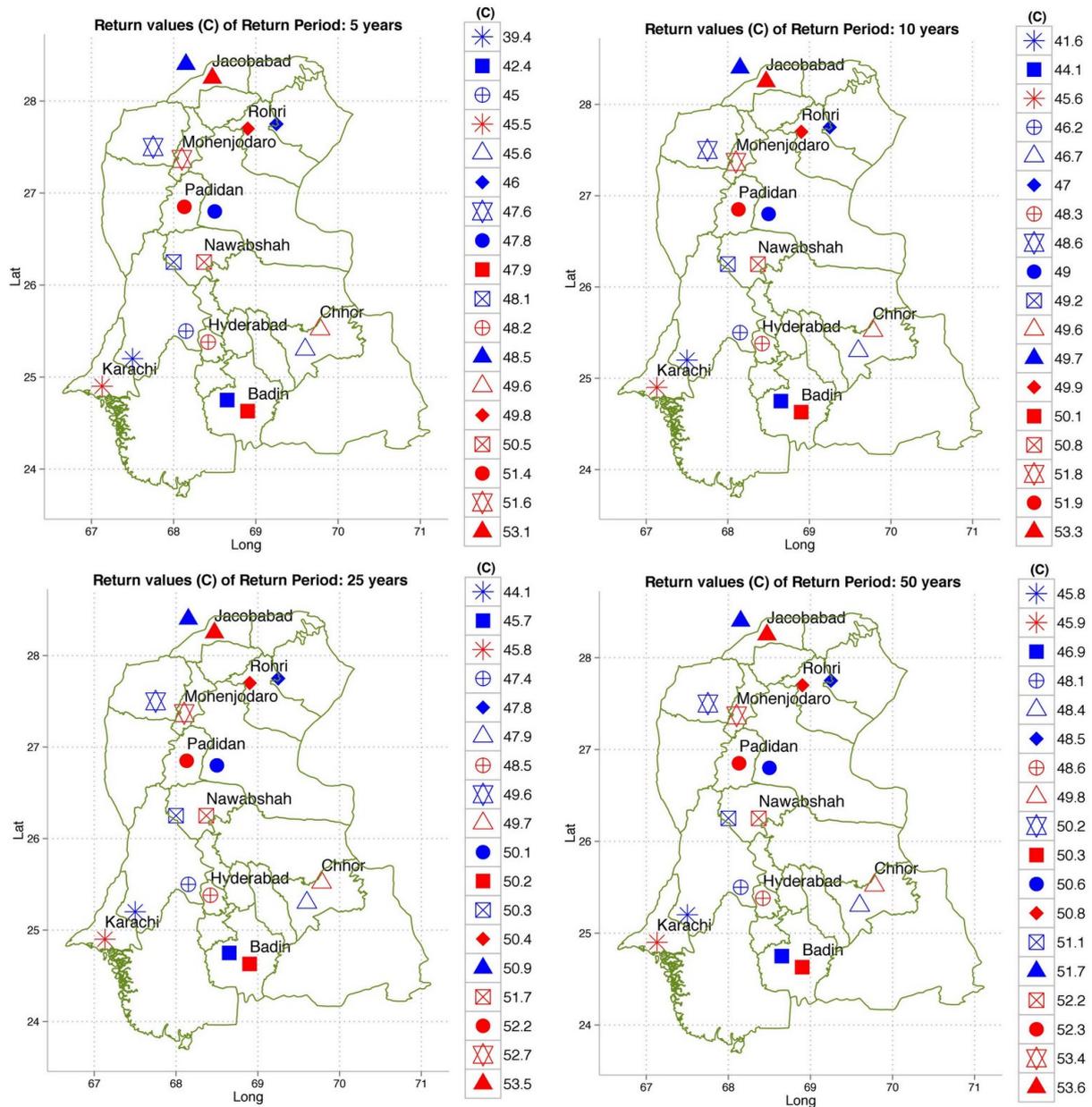


Figure 8. Spatial distribution of the station observed T_{max} (red) and bias corrected ERA Interim T_{max} (blue) return levels in degree Celsius corresponding to return periods of 5, 10, 25 and 50 years in southern Pakistan.

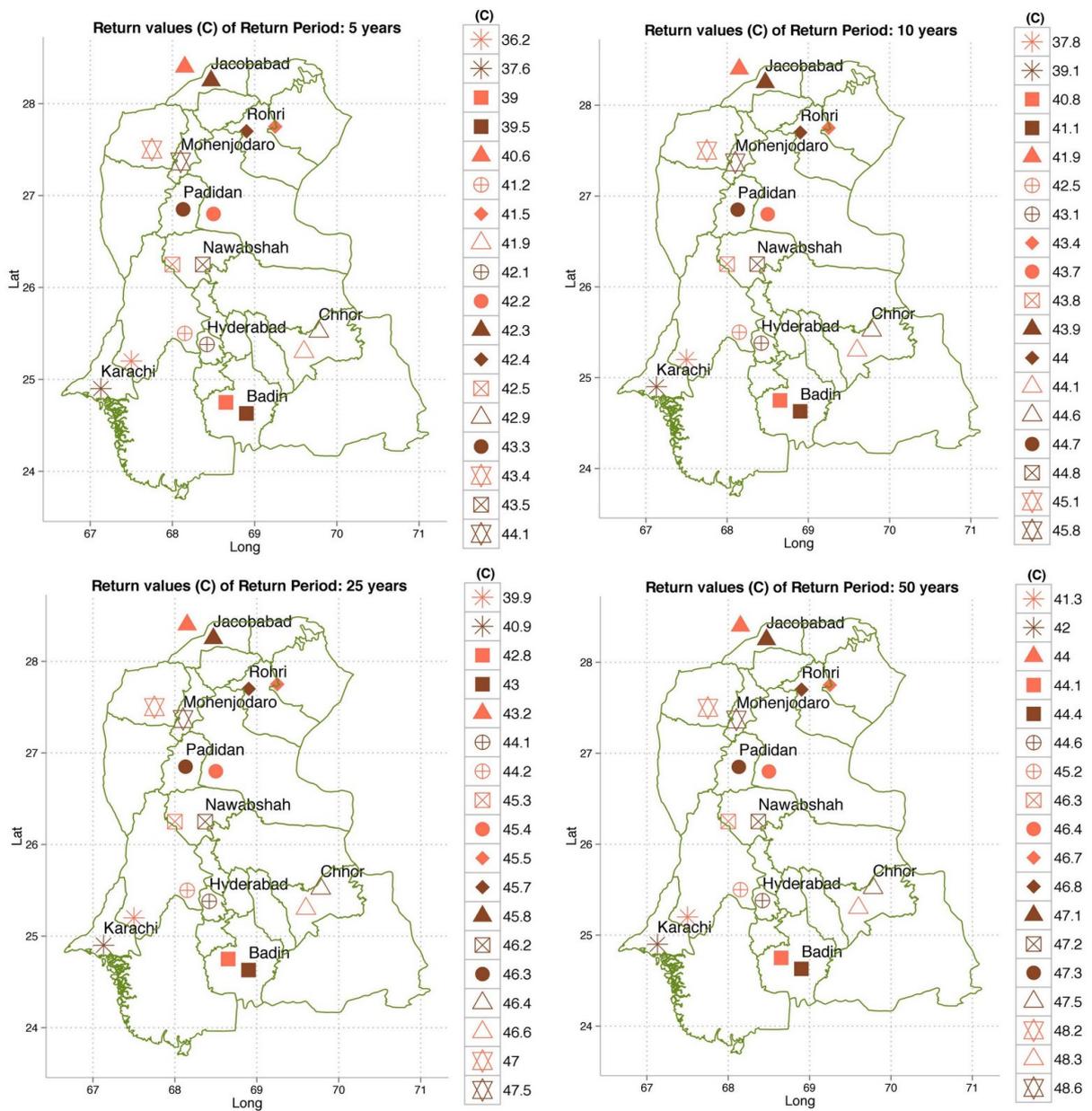


Figure 9. Spatial distribution of the station observed TW_{max} (brown) and bias corrected ERA Interim TW_{max} (orange) return levels in degree Celsius corresponding to return periods of 5, 10, 25 and 50 years in southern Pakistan. 5

10

15