A hierarchical Bayesian regression framework to analyze climate data from Central Asia region

Emerson Barili , Jorge Alberto Achcar
Ribeirão Preto Medical School University of São Paulo - Ribeirão Preto - SP

and Ricardo Puziol de Oliveira
State University of Maringá - Maringá - Republic of Brazil

A HIERARCHICAL BAYESIAN REGRESSION FRAMEWORK TO ANALYZE CLIMATE DATA FROM CENTRAL ASIA REGION

EMERSON BARILI[1], JORGE ALBERTO ACHCAR[1], AND RICARDO PUZIOL DE OLIVEIRA[2]

[1] Ribeirão Preto Medical School University of São Paulo - Ribeirão Preto - SP

[2] State University of Maringá - Maringá - Republic of Brazil

Abstract:

This study introduces a straightforward framework for analyzing climate data related to the minimum and maximum temperatures of countries in Central Asia (Kazakhstan, Kyrgyzstan, Tadjikistan, Turkmenistan, and Uzbekistan), considering annual temperature averages over a long period of time ranging from the early 1900's to the beginning of the year 2000. The data analysis used standard existing multiple linear regression models under a hierarchical Bayesian approach, assuming as covariates latitude and longitude of the climate stations, temporal factors (linear, quadratic, and cubic effects of years), and altitude of the climate station. The findings yielded highly accurate results in identifying significant factors influencing climate change, such as time (year), altitude, and spatial factors, as well as in predicting average temperatures in future years. Furthermore, the obtained results align with numerous other studies in the literature, indicating that all regions of the world are already experiencing climate change. In particular, we observed that annual average minimum temperatures in Central Asia are increasing in the five countries assumed in the study at the end of the follow-up period (close to the year 2003). We also observed similar results for the annual average maximum temperatures.

Keywords: climate data, multiple linear regression models, Bayesian inference, MCMC métodos, minimum and maximum yearly average temperatures

Introduction

Recent decades have witnessed numerous effects of climate change on the environment, including the shrinkage of glaciers, the melting of ice from rivers and lakes, the early blossoming of trees, the rapid rise in sea level, and intense heat waves, among others, underscoring the significance of studying climate change. For example, the average rainfall in the United States of America has increased since 1900, but some areas of the country have seen increases greater than the national average, and some areas have Throughout this century, we anticipate more winter and spring precipitation in the northern United States and less in the southwest.

The 2021 project uses the CMIP6 phase 6 data to examine changes in extreme temperature and precipitation events for the mid-century (2036–2065) and late-century (2070–2099) periods, comparing them to the reference period (1985–2014). It achieves this by examining various indices and socioeconomic scenarios around the world. The projected values show an increase in the intensity and frequency of hot temperatures and precipitation extremes. Valipour (2021) highlights the importance of climate data in determining global surface temperature. To do this, we look at how the mean surface temperature (ST), wind speed (WS), and albedo (AL) have changed over the last 20 years (2000–2019) from the Global Land Data Assimilation System (GLDAS) around the world and compare them to data from 1961–1990 to see if these changes have an effect on ST, WS, and AL.

As declared by the UN Intergovernmental

Panel on Climate Change (IPCC) in 2013, climate change is real, and human activities are the main cause of this change. The UN prepared the fifth assessment report, which provides information on the sea level rise in recent decades and cumulative estimates of CO2 emissions since pre-industrial time. The report found that:

From 1880 to 2012, the average global temperature increased by 0.85 °C.

The oceans warmed, the amounts of snow and ice decreased, and the sea level has risen. From 1901 to 2010, the global average sea level rose by 19 cm as the oceans expanded due to warming and melting ice.

Given the current concentrations and ongoing emissions of greenhouse gases, it is likely that by the end of this century, the global mean temperature will continue to rise above the pre-industrial level. We predict an average sea level rise of 24–30 cm by 2065 and 40–63 cm by 2100, compared to the reference period of 1986–2005. Even if we stop emissions, most aspects of climate change will persist for many centuries.

Since this topic is of significant interest, the literature presents a tremendous number of papers and reports from climate agencies related to climate change and its implications. Arnell (11999) described the effects of climate change by 2050 on hydrological regimes on a continental scale in Europe. Iglesias (2012) presented a study that consistently demonstrated the impact of climate change on arable agriculture in Europe. Guillemain (2013) investigated the primary impacts of climate change on birds, specifically ducks, in Europe. (Aguilar et al., 2005) conducted an analysis of climate change indices in Central America from 1961 to 2003. Their findings indicated a general warming trend in the region, which could potentially contribute to atypical natural events. ents. Jones and Thornton (2003) conducted a study that showed that the impacts of climate change on agriculture, such as maize production, can significantly increase the development challenges to ensure food security and reduce poverty in Africa and Latin America until 2055. Campbell et al. (2014) demonstrated that the effects of climate change in Oceania led to a rise in sea level, an increase in the frequency and severity of floods and droughts, an increase in the intensity of tropical cyclones, and a shift in the distribution of disease vectors. In 2021, Somanathan presented an association between economic production and the hottest years, estimating that workers in India experience a reduction in productivity during periods of temperature increase, thereby establishing a relationship between production and climate changes.

Arnell (1999), Bonan (2008), Costello et al. (2009), Hawkins (2017), Kabir et al. (2016), Kaczan and Orgill-Meyer (2020), Levermann et al. (2013), Li and Fang (2016), Mathews (2018), Poloczanska et al. (2013), Rahmstorf et al. (2007), Serdeczny et al. (2017), Springmann et al. (2016), Turner et al. (2020), Zhao et al. (2017).

Climate change particularly impacts Central Asia, which includes Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan. Perelet et al. (2007) noted that global climate change poses serious threats to the environment, ecological, and socioeconomic systems of the Central Asia region. This region has already experienced a decline in agricultural production, and the quantity and quality of water resources are at risk due to the severe effects of climate change. We analyzed the maximum and minimum temperatures on an annual and seasonal basis, gathering data from

108 climate stations between 1981 and 2015.

Achcar and de Oliveira (2022) used non-homogeneous Poisson processes (NHPP) to look at Kazakhstan's rainfall data from 1879 to 2002. They discovered that the average rainfall was higher 11 years before the estimated change point (year 43, which is 1921) and 46 years after that. This could be considered an indication that the average rainfall has increased since 1921. However, assuming the temperature data from Kazakhstan (1915 to 2003), they found that there were 13 years before the estimated change point (year 33, which corresponds to the year 1948) and 30 years after, with maximum temperature averages above the overall average. This suggests that the average maximum temperature has been increasing since 1948.

According to Feng et al. (2018), there was an increase in all extreme temperature values on an annual scale, with the maximum temperature increasing faster than the minimum temperature. Different statistical models reveal an increase of 0.37°C/decade in temperature projections for the period 2021-2060 compared to the period 1965-2004, with higher latitudes and mountainous areas experiencing accelerated warming (Luo et al., 2019). Some research that looked at changes in average temperature and rainfall for the years 2011–2040, 2041–2070, and 2071-2100 in Central Asia used climate models to show that the rise in temperature and fall in rainfall compared to 1971–2000 can impact the social and economic systems of this area, which is mostly dry or semi-dry (Ozturk et al., 2017).

The goal of this work is to present a multiple linear regression model that uses a Bayesian hierarchical framework. This model can be very useful for getting very accurate fits to see if factors like longitudes and latitudes have a big impact on the average highest and lowest temperatures that happen each year. Some studies have looked at the relationship between the highest and lowest temperatures and the effects of global warming. These studies usually use simple linear regression models and the classic approach of looking for linear trends over time (Brown, 2008; Donat et al., 2013; Stocker, 2014). In particular, Feng et al. (2018) introduced a study that also examined maximum and minimum temperatures, demonstrating the increasing linear climate trends and their spatial and temporal effects in the Central Asia region. Other studies considering extreme temperatures in other parts of the world were also introduced in the literature (Alexander et al., 2006; Donat et al., 2013; Stocker, 2014; Tao et al., 2014; Cassano et al., 2015; Fang et al., 2015; Horton et al., 2015; Zhong et al., 2017), but as pointed out by Feng et al. (2018), few studies have been performed in arid and semi-arid regions (Klein et al., 2006; Wang et al., 2013; Tonkaz et al., 2007) as considered in the present work.

We want to make a multiple linear regression model that looks at the average lowest and highest temperatures over the course of a year, includes a latent factor (a variable that can't be measured), and includes a number of independent variables, such as time, latitude, longitude, and elevation of the climate stations, along with their linear, quadratic, and cubic effects. The proposed approach does not require numerical methods that strongly depend on precise initial values to reach convergence. With the suggested framework and common MCMC (Markov Chain Monte Carlo) techniques like Gibbs Sampling, Metropolis-Hastings, and Metropolis-within-Gibbs (MwG), it is possible to use a fully Bayesian approach to guess and come to the right conclusions. We chose to use the MwG algorithm (Gilks et al., 1995) in this paper to get pseudo-random samples from the roughly emphasized posterior distribution of model parameters. In summary, this study aims to achieve the following primary objectives:

We aim to present a statistical examination of climate data in Central Asian countries such as Kazakhstan, Kyrgyzstan, Tadjikistan, Turkmenistan, and Uzbekistan, spanning from the early 1900s to the start of 2000. We focus on the annual average minimum and maximum temperatures, taking into account a latent factor (an unobserved variable) and certain independent variables influenced by time, latitude, longitude, and elevation of the climate stations.

We used multiple linear regression models with a hierarchical Bayesian approach to look at the data. We accounted for the latitude and longitude of the climate stations, temporal factors (linear, quadratic, and cubic effects of years), and altitude as covariates.

We want to draw accurate conclusions and predictions using standard MCMC (Markov Chain Monte Carlo) methods, assuming prior distributions for the parameters of the proposed models that are not very informative. This is part of a fully Bayesian approach.

This paper is organized as follows. In Section 2, we present the dataset and the fundamental concepts regarding formulation and estimation of the proposed model. In Section 3, we analyze and discuss the results obtained using the proposed methodology for modeling the climate data in Central Asia. We also present model comparisons between linear and polynomial approximations. Section 4 addresses general comments and concluding remarks.

Materials and Methods

2.1 Climate Data

The study considers a climate data set introduced by Williams and Konovalov (2008) related to the annual average minimum and maximum temperature rather than the usual annual average temperatures in many studies introduced in the literature for five countries in Central Asia: 211 Kazakhstan, 213 Kyrgyzstan, 227- Tajikistan, 229 Turkmenistan, and 231 Uzbekistan (Figure1). Although this data set has some limitations since the time series end in the year 2003, it is possible to verify the behavior of the climate series for a long period. Analysis of the extreme minimum and maximum temperature means can give a more realistic view of climate change.

Figure 1: Map of Central Asia.

The original data set consists of monthly average temperatures (minimum and maximum) reported in different climate stations for many years. We only took into account the annual averages of complete data because there were a large number of missing observations in the data set, particularly in the early years of the follow-up periods. This refers to the annual averages of the maximum and minimum temperatures for the years when there was data available for a full year. It is important to note that the average monthly temperatures in each of the four seasons (summer, winter, spring, and autumn) exhibit seasonality. Therefore, calculating annual average temperatures based on only a few months of the year rather than the complete 12-month data for each year would be meaningless.

The annual minimum temperature data set consists of n = 3467 annual averages (average of the 12-month averages in each year ranging from the year 1883 to the year 2003) reported in different climate stations (Appendix 1) for each country (Figure 2). The original data set has deleted some observations due to missing data, and each sample observation represents the average of the monthly minimum temperature averages for each year. On the other hand, various climate stations (Appendix 1) report the annual average maximum temperatures, which comprise n = 3959 annual averages (the mean of the 12-month means in each year from 1894 to 2003) for each country (Figure 2). Moreover, in Appendix 2 at the end of the manuscript, we have the latitude, longitude, altitude, and study period for each station.

2.2 Statistical Analysis

In recent years, the literature has extensively explored the topic of climate change. However, new studies using various databases and statistical models can provide valuable insights and aid public authorities in making sometimes drastic decisions to mitigate the impact of climate change on future generations. The latitude and longitude factors provide precise information about how neighboring regions influence the temperature measurements at each location. This information can serve as a viable substitute for the existing Bayesian spatial-temporal models, such as the conditional autoregressive (CAR) and the simultaneously autoregressive (SAR) models introduced in the literature (Besag, 1974; Wall, 2004; Cressie and Chan, 1989). Under this model approach, Tawn et al. (2018) considered applications with environmental extreme events. On the other hand, Zarei et al. (2021) assumed multiple Bayesian linear regression models to investigate the impact of minimum

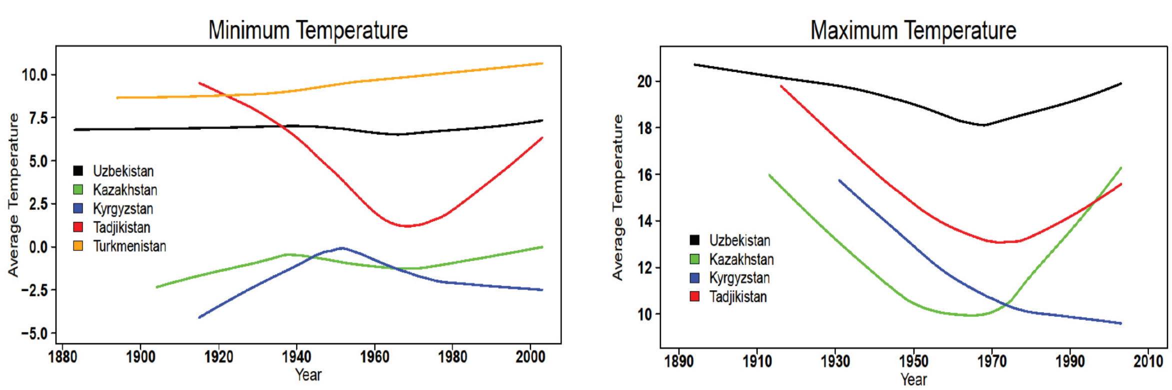

Figure 2: Annual average minimum and maximum temperatures in Central Asia (Kazakhstan, Kyrgyzstan, Tadjikistan, Turkmenistan and Uzbekistan).

and maximum temperatures on plantations in Iran during the period 1968–2017. In another paper, Singh et al. (2020) assumed a Bayesian approach to implement a simulation study for the association between climate and the impact of rice productivity in the state of Louisiana, USA, from 1960 to 2015.

The model also includes a non-observed latent factor, representing the effect of other covariates not included in the model, that could affect the variability of the response variable (Mcculloch and Searle, 2004; Laird and Ware, 1982; Verbeke, 1997). Assuming linear regression statistical models, we want to look at time factors like years (linear, quadratic, and cubic effects), space factors like longitudes (denoted as long) and latitudes (denoted as lat) of each climate station, and other factors like the altitudes (denoted as alt) of each climate station.

2.2.1 Linear Regression Model

Let us consider the data presented in Section 2.1 related to the annual average minimum and maximum temperatures. In the statistical analysis of the data set, we have adopted the following linear regression model for the annual average minimum temperature,

$$Y_{1i}=\beta_0+b_i+\beta_1(lat_i\text{-}avg.lat)+\beta_2(long_i\text{-}avg.long)+\beta_3(alz_i\text{-}avg.alt)+\beta_4x_i+\beta_5x_i^2+\beta_6xi^3+\epsilon_i, \quad (i=1,\dots,3467) \qquad (1)$$

where Y_{1i} is the annual average minimum temperature; x_i = year-1883; b_i is a latent or random factor assumed as a random variable with a normal distribution $N(0,\sigma_b^2)$ and ϵ_i is a error term assumed to be independent identically distributed with a normal distribution $N(0,\sigma^2)$. The equation (1) defines a latente variable model, that is, a statistical model that contains latent, or unobserved variables (effects not measured by the covariate effects of latitude, longitude, altitude and the temporal effects of years). Equation (1) connects the observed and latent variables using a linear model where the outcome (observed) dependent variable is assumed to be a continuous random variable. On other hand, for the annual average maximum temperatures, we have adopted the following linear regression model,

$$Y_{2i}=\beta_0+b_i+\beta_1(lat_i\text{-}avg.lat)+\beta_2(long_i\text{-}avg.long)+\beta_3(alt_i\text{-}avg.alt)+\beta_4x_i+\beta_5x_i^2+\beta_6x_i^3+\epsilon_i, \quad (i=1,\dots,3959) \qquad (2)$$

where Y_{2i} is the annual average maximum temperature. In both models, we included a polynomial terms to capture possible effects of linearity, quadratic and cubic effects of year. A quadratic term or cubic term transforms a linear regression model into a curve. Since the regression model has the squared and cubic year as covariates, and not transformations on the regression coefficient, the model remains a linear regression model (Draper and Smith, 1998; Seber, 2015). The presence of a quadratic term in the model creates a U-shaped curve or an inverted U, as seen in the graphs presented in Figure 2. A cubic term has two distinct parts: one facing up and one facing down, that is, the curve go down, back up and back again.

In both hierarchical Bayesian models (model (1) and model (2)), assuming a normal probability distribution for the random effects, we have a conjugated prior distribution providing a posterior distribution in its closed form. The use of conjugated prior distributions also is usually assumed for other hierarchical Bayesian models in presence of latent non-observed factors considering other distributions, such as, binomial data, where the conjugated prior distribution for the random effects is a Beta distribution, thus obtaining the Beta-Binomial model. In the same way, if the data are results from a Poisson distribution, the prior conjugate distribution for the random effects is a Gamma distribution, obtaining the Gamma-Poisson model. Furthermore, the literature presents many studies analyzing the impact of other probability distributions for the random effects (Bell et.al, 2019), showing that the results (parameter estimates and standard errors) are nearly identical when assuming a normal distribution. In addition, we have great computational advantage having a closed posterior distribution when we use the normal distribution for both random effects, b_i and ϵ_i in (1) and (2). In our study assuming the modeling structure based on random effects with normal distribution, we obtained a good fit of the model for the data as observed in Figure 4 presented in the application section. The assumption of normality for the random errors usually is verified from residual plots.

2.2.2 Polynomial Regression Model

As second approach, we have adopted only the covariate x=year (114 years from 1883 to 2003). In this case, we assume the following polynomial regression model for the annual average minimum temperature,

$$Y_{1i}=\beta_0+b_i+\beta_1 x_i+\beta_2 x_i^2+\beta_3 x_i^3+\epsilon_i, \quad (i=1,\ldots,114) \qquad (3)$$

and, for annual average maximum temperatures (covariate x = year with 108 years from 1894 to 2003), the following polynomial regression model,

$$Y_{2i}=\beta_0+bi+\beta_1 x_i+\beta_2 x_i^2+\beta_3 x_i^3+\epsilon_i, \quad (i=1,\ldots,108) \qquad (4)$$

2.2.3 Approximate Hierarchical Bayesian Inference

In this subsection, we address the problem of estimating and making inferences from proposed models under a fully Bayesian perspective. In this way, assuming a Hierarchical Bayesian framework, we have adopted weakly-informative Normal prior distributions for the vector $\boldsymbol{\beta}=(\beta_0,\ldots,\beta_p)$ (p is the number of covariates) of regression parameters, that is

$$\beta \sim N_k\,(0,1_k)$$

where $\mathbf{1}_k$ is identity matrix of size k. As for parameters $\zeta=1/\sigma^2$ and $\zeta_b=1/\sigma_b^2$, we have adopted a Gamma prior distribution with both hyperparameters equal to 0.01. We further assume prior independence among all parameters. From the Bayesian point of view, inferences for the elements of $\boldsymbol{\beta}$ can be derived from their marginal posterior distribution. Here, we have opted to use a suitable iterative procedure to draw pseudo-random samples from the approximate posterior density in order to make inferences for $\boldsymbol{\beta}$. Thus, in order to generate N pseudo-random values for each element of $\boldsymbol{\beta}$, we have adopted the MWG algorithm in which a total of N=110,000 pseudo-random values from the approximate *posterior* distribution of $\boldsymbol{\beta}$ were obtained. After generating the values, the first 10,000 samples were discarded (burn-in period). Then one out of every 100 generated values was kept, resulting in sequences of size B=1,000 for each element of $\boldsymbol{\beta}$. Finally, trace plots were used to assess the stationarity of the obtained chains.

3 Results and Discussion

3.1 Results

In this subsection, we address the results of the proposed methodology for the annual average minimum temperatures and maximum temperature data for each country (211 Kazakhstan, 213 Kyrgyzstan, 227- Tajikistan, 229 Turkmenistan, and 231 Uzbekistan). As a preliminary analysis, we considered an ANOVA model and verified that the means for both responses are statistically different for each country (p-value < 0.05). Figure 3 shows the 95% confidence intervals for the minimum and maximum temperature values in each region.

Table 1 presents the posterior parameter estimates and the 95% Credible Intervals (CIs) for annual average minimum temperatures based on the fitted models. From the displayed results, we can notice that the CI of parameters $\beta_j, j=2,\ldots,6$ of model 1 do not contain the value zero, which constitute longitude, altidude, x (linear effect), x^2 (quadratic effect) and x^3 (cubic effect) as relevant covariates to explain part of the response' variability. Also, since altitude has a negative Monte Carlo Bayesian estimator for the regression parameter β_3 (-0.1131), we conclude that the annual average minimum temperature decreases with altitude, an expected result. Moreover, for model given by (3), we can observe that the factor year shows significant effect on the response average minimum temperature in terms of quadratic and cubic effects since zero is not included in the 95% credible intervals for the regression parameters β_2 and β_3.

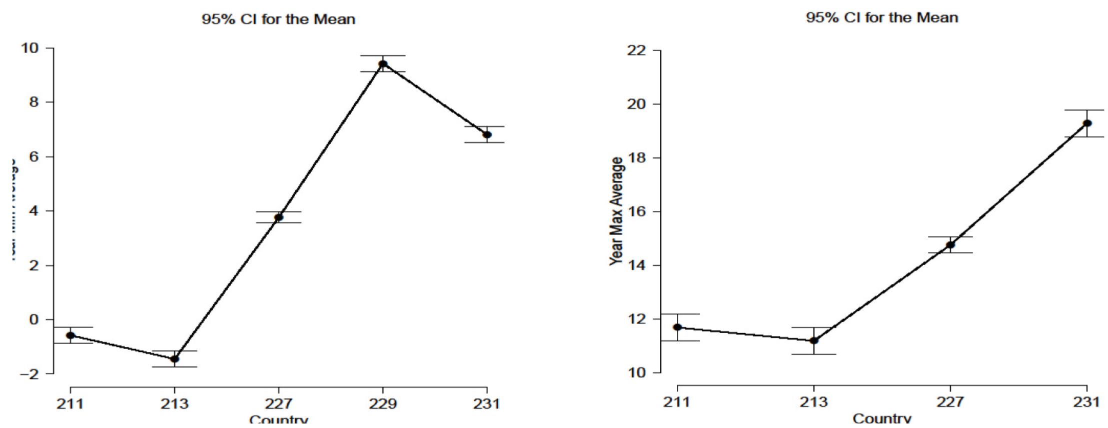

Figure 3: 95% confidence intervals for the temperature means (211-Kazakhstan, 213-Kyrgyzstan, 227-Tadjikistan, 229-Turkmenistan and 231-Uzbekistan).

Table 1: Posterior parameter estimates and 95% credible intervals – Annual average minimum temperatures

Model	Parameter	Mean	Std. Dev.	95% CI	
				Lower	Upper
Eq.(1)	β_0	45.400	0.997	43.660	47.350
	β_1	1.107	0.683	-0.316	2.717
	β_2	-0.675	0.398	-1.536	0.011
	β_3	-0.113	0.002	-0.117	0.107
	β_4	-85.080	0.043	-85.140	84.980
	β_5	2.025	0.001	2.022	2.028
	β_6	-0.011	0.0002	-0.011	0.011
	ζ	0.859	0.908	0.036	3.332
	ζ_b	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Eq.(3)	β_0	8.123	0.728	6.689	9.511
	β_1	0.010	0.046	-0.080	0.105
	β_2	-0.002	0.001	-0.0044	-0.001
	β_3	< 0.0001	< 0.0001	< 0.0001	< 0.0001
	ζ	0.599	0.349	0.323	1.416
	ζ_b	0.912	0.303	0.427	1.600

Table 2 presents the posterior parameter estimates and the 95% Credible Intervals (CIs) for annual average maximum temperatures based on the fitted models. From the displayed results, we can notice using model (2), except for β_2, the significant effects are the same as the average minimum temperatures to explain part of the response' variability. In addition, for model given by (4), we can observe that the linear and cubic effects are significant since zero is not included in the 95% credible intervals for the regression parameters β_1 and β_3.

Figure 4 shows the minimum and maximum mean temperatures and the estimated means versus years fitted by models (3) and (4)

Table 2: Posterior parameter estimates and 95% credible intervals – Annual average maximum temperatures

Model	Parameter	Mean	Std. Dev.	95% CI	
				Lower	Upper
Eq.(2)	β_0	65.180	0.916	63.190	66.850
	β_1	1.164	0.689	-0.285	2.245
	β_2	-0.213	0.327	-1.005	0.251
	β_3	-0.077	0.001	-0.080	-0.074
	β_4	-93.200	0.050	-93.290	-93.110
	β_5	2.340	0.000	2.339	2.341
	β_6	-0.014	0.0001	-0.014	-0.014
	ζ	1.185	1.002	0.195	3.990
	ζ_b	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Eq.(4)	β_0	28.650	2.302	24.180	33.050
	β_1	-0.348	0.134	-0.593	-0.078
	β_2	0.001	0.002	-0.002	0.005
	β_3	< 0.0001	< 0.0001	< 0.0001	< 0.0001
	ζ	0.334	0.172	0.164	0.688
	ζ_b	0.941	0.329	0.403	1.686

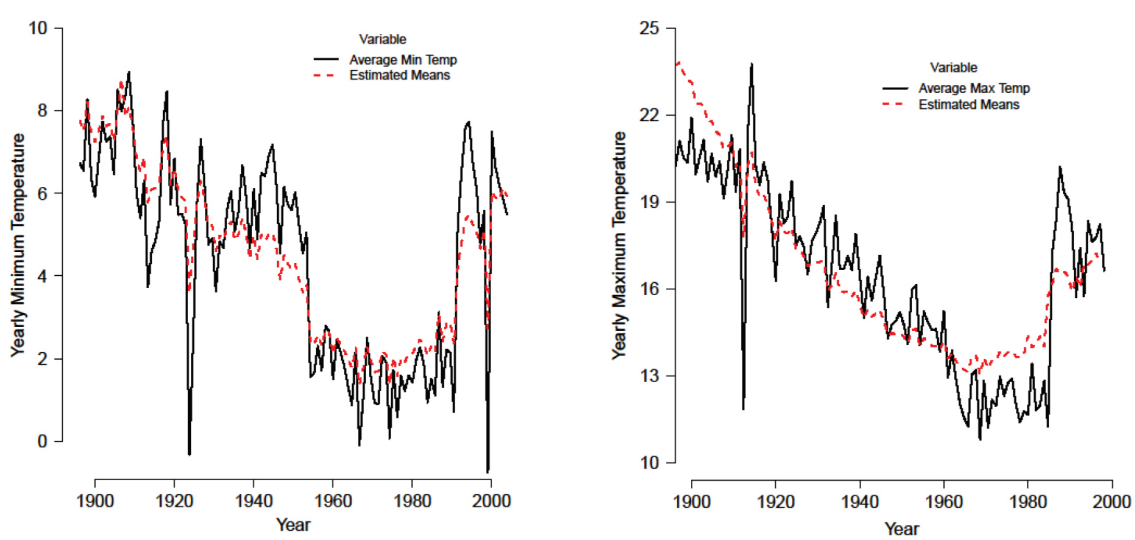

Figure 4: Average minimum and maximum temperatures and estimated means versus years.

Figure 5 illustrates the simulated envelope for the residuals assuming the polynomial regression models (Equations (3) and (4)) for each response which suggests that the model fit is adequate since all the estimated residuals are lying within the simulated envelope, which also indicates that there is no severe violation of model assumptions. Additionally, one can notice that the proposed models fits considerably well since the estimated residuals are close to the dashed median line.

The needed assumptions (normality, constant variance and non-correlated) for the errors in the proposed regression models were verified by residual plots. Figure 6 shows the residual plots considering models (1) and (2).

One of the findings from this study is that from the year close to 1970 there is a consistent increasing of the annual average temperatures (minimum and maximum) in the Central Asia region. Hu et al. (2014) corroborate this result, pointing out that despite the lack of accurate climate data, multiple data sets show regional temperature increases from 1979 to 2011. They also discovered that the rate is higher in recent years than in the early years of the follow-up period, with a greater increase in surface temperature in the spring season and accelerated warming in Central Asia compared to other regions of the world (Brohan et al., 2006; Smith and Reynolds, 2005).

Another finding from this study is that the high temperature variability may cause changes in vegetation and agriculture in the Central Asia region. Propastin et al. (2008) looked at annual

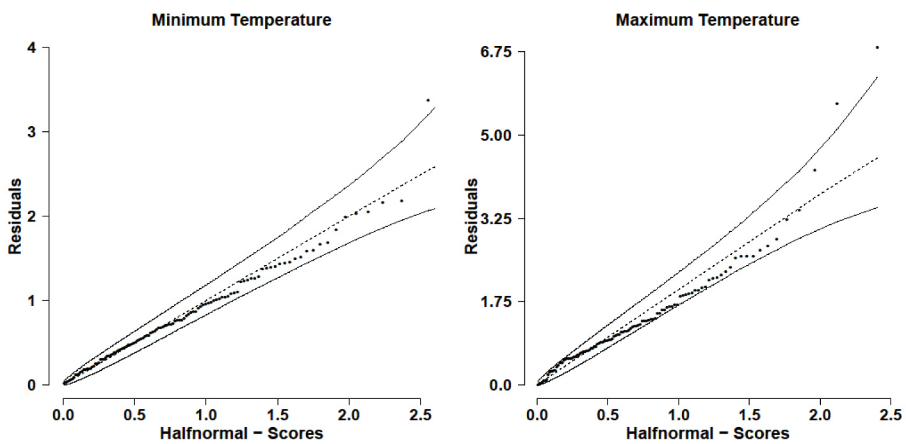

Figure 5: Normal plot with simulated envelope for the residuals for annual average minimum and maximum temperatures (left-panel: Model (3); right-panel: Model (4).

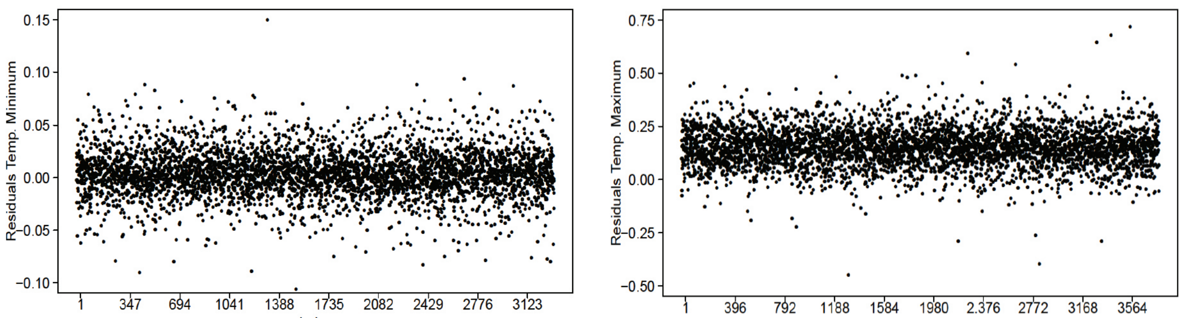

Figure 6: Residual plots (minimum and maximum temperatures) for the proposed models (left-panel: Model (1); right-panel: Model (2)).

and seasonal plant activities in Central Asia using time series models for the Normalized Difference Vegetation Index (NDVI) from the NOAA/AAVHRR data set. They found that temperature change is the only thing that affects the trend in NDVI in the spring, since temperature is a measure of the energy that plants can use to grow. Yin et al. (2016) demonstrated a significant increasing trend in the annual normalized difference vegetation index (NDVI) between 1982 and 1994, followed by a decreasing trend since 1994. The study's findings, which show significant fluctuations in minimum and maximum temperatures as well as an upward trend towards the end of the follow-up period, validate these results and have implications for the region's vegetation. In parallel, Chen and Guo (2012) fitted a statistical model to quantify the impact of temperature variations on pasture productivity during 1982–2015, concluding that climate change had a severe influence on pasture productivity in Central Asia since the year 1980. According to Reyer et al. (2017), climate changes have a negative impact on the availability of water in the Central Asia region, which could potentially lead to a decrease in crop yields and pasture productivity.

Our findings also showed that the annual average minimum temperatures in Central Asia are increasing in all countries (Figure 4) at the end of the follow-up period (close to the year 2003). Despite significant data variability, we observe this trend. The same happens for the annual average maximum temperatures (Figure 4). It is important to point out that climate data from different climate stations, especially in third-world countries, usually present many missing data or outliers in the reported data sets, sometimes due to errors. This typically poses a challenge in obtaining accurate results from climate data analysis. Moreover, the factors longitude (long) and altitude (alt) have significant effects on the annual average minimum temperatures; that is, the annual average minimum temperatures depend on the spatial location of each climate station. The spatial factors longitude (long) and altitude (alt) do not show significant effects on the annual average maximum temperatures; that is, the annual average maximum temperature does not depend on the spatial location of each climate station. Also, considering the combined data of the five countries, we observe that from the year close to 1970 there is a consistent increasing of the annual average temperatures (minimum and maximum), despite the substantial heterogeneity between the annual average temperatures of the five countries in Central Asia (Figure 4).

Overall, this study confirms the findings of numerous recently published papers in the literature: the warming effects of climate change observed in recent decades are the primary cause of the large variations in average annual maximum temperatures in Central Asia. Asia is a large continent, including parts near the polar regions and parts near the equator, leading to large differences in climate. Other potential factors, such as the proximity of some parts of the Asia continent to seas and oceans, as well as the presence of various landforms, also contribute to this climatic behavior. This phenomenon is evident in Central Asia, where the impact of climate change is particularly severe. Other regions of the world, such as the arid region of northeastern South America, particularly in Brazil, are also experiencing significant effects of climate change due to human activities such as deforestation, the establishment of large farms for cattle raising, and agricultural production. These activities have resulted in increasing annual maximum and minimum temperatures in recent decades, leading to catastrophic economic consequences (Soares et al., 2021).

4 Conclusion

Polynomial and linear regression approaches typically involve choosing a model among many existing linear and non-linear formulations, which can be a burden in many applications. In some situations, adding a random factor could be more precise to obtain the inferences of interest; however, most numerical iterative methods for model fitting strongly depend on choosing precise initial values. In this sense, we aimed to introduce a simple regression model that incorporates a latent variable, a linear, a quadratic, and a cubic effect as an alternative to many existing linear time trends techniques in the literature. We obtained approximate posterior inferences for the model parameters using a fully hierarchical Bayesian approach based on the MwG and weakly informative priors as inference methods. The findings demonstrated great accuracy in identifying important factors affecting climate change (such as time (year), altitude, and spatial factors) and predicting average temperatures in the future. Furthermore, the obtained results align with numerous other studies in the literature and hold potential for application in other arid regions of the world, particularly the northeast of Brazil, which has experienced catastrophic climate change effects in recent years. In conclusion, the results obtained from this study may be of great interest for authorities in the areas of environment, agriculture, sanitation, and irrigation in the countries of the Central Asia region to plan public policies for agricultural plantations and the construction of water deposits, among many challenges that the planet will face in the coming decades.

References

*Achcar, J. A. and de Oliveira, R. P. (2022) Climate change: use of non-homogeneous poisson processes for climate data in presence of a change-point. Environmental Modeling & Assessment, 27(2):385–398.

*Aguilar E, Peterson T, Obando PR, et. al. (2005) Changes in precipitation and temperature extremes in central america and northern south america, 1961–2003. Journal of Geophysical Research: Atmospheres 110(D23).

*Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., et al. (2006) Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research: Atmospheres, 111(D5).

*Almazroui, M., Saeed, F., Saeed, S., Ismail, M., Ehsan, M. A., Islam, M. N., Abid, M. A., O'Brien, E., Kamil, S., Rashid, I. U., et al. (2021) Projected changes in climate extremes using CMIP6 simulations over SREX regions. Earth Systems and Environment, 5(3):481–497.

* Arnell, N. W. and Lloyd-Hughes, B. (1999) The global-scale impacts of climate change on water resources and -oodingunder new climate and socio-economic scenarios. Climatic change, 122(1):127–140.

* Bell A, Fairbrother M, Jones K. (2019) Fixed and random effects models: making an informed choice. Quality and quantity 53:1051–1074.

*Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society: Series B (Methodological), 36(2):192–225.

*Bonan, G. B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefts of forests. Science, 320(5882):1444-1449.

* Brohan, P., Kennedy, J. J., Harris, I., Tett, S.

F., and Jones, P. D. (2006) Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. Journal of Geophysical Research: Atmospheres, 111(D12).

* Brown, S. J., Caesar, J., and Ferro, C. A. 2008 Global changes in extreme daily temperature since 1950, Journal of Geophysical Research: Atmospheres, 113(D5).

* Campbell J, Bedford R, Bedford R. (2014) Migration and climate change in oceania. People on the move in a changing climate: The regional impact of environmental change on migration pp 177–204.

* Casella, G. and George, E. I. (1992) Explaining the Gibbs sampler. The American Statistician, 46(3):167–174.

* Cassano, E. N., Glisan, J. M., Cassano, J. J., Gutowski Jr, W. J., and Seefeldt, M. W. (2015) Self-organizing map analysis of widespread temperature extremes in Alaska and Canadá. Climate Research, 62(3):199–218.

* Chen, H., Xu, C.-Y., and Guo, S. (2012) Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. Journal of hydrology, 434:36–45.

* Chib, S. and Greenberg, E. (1995) Understanding the Metropolis-Hastings algorithm. The american statistician, 49(4):327–335.

* Costello, A., Abbas, M., Allen, A., Ball, S., Bell, S., Bellamy, R., Friel, S., Groce, N., Johnson, A., Kett, M., et al. (2009) Managing the health effects of climate change: Lancet and University College London Institute for Global Health Commission.

* Cressie, N. and Chan, N. H. (1989) Spatial modeling of regional variables. Journal of the American Statistical Association, 84(406):393–401.

* Donat, M., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J., Willett, K. M., Aguilar, E., Brunet, M., Caesar, J., et al. (2013) Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, Journal of Geophysical Research: Atmospheres, 118(5):2098–2118.

** Draper, N. R. and Smith, H. (1998) Applied regression analysis, volume 326. JohnWiley and Sons.

* Fang, S., Qi, Y., Han, G., and Zhou, G. (2015) Changing trends and abrupt features of extreme temperature in mainland China during 1960 to 2010. Earth System Dynamics Discussions, 6(1):979–1000.

* Feng, R., Yu, R., Zheng, H., and Gan, M. (2018) Spatial and temporal variations in extreme temperature in Central Asia. International Journal of Climatology, 38:e388–e400.

* Gelfand, A. E. and Smith, A. F. (1990) Sampling-based approaches to calculating marginal densities. Journal of the American statistical association, 85(410):398–409.

* Gilks, W. R., Best, N. G., and Tan, K. K. (1995) Adaptive rejection Metropolis sampling within Gibbs sampling. Journal of the Royal Statistical Society: Series C (Applied Statistics), 44(4):455–472.

* Guillemain M, P$\ddot{o}$ys$\ddot{a}$ H, Fox AD, et al. (2013) Effects of climate change on european ducks: what do we know and what do we need to know? Wildlife Biology 19(4):404–419.

Hawkins, E., Ortega, P., Suckling, E., Schurer, A., Hegerl, G., Jones, P., Joshi, M., Osborn, T. J., Masson-Delmotte, V., Mignot, J., et al. (2017) Estimating changes in global temperature since the preindustrial period. Bulletin of the

American Meteorological Society, 98(9):1841-1856.

* Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S. (2015) Contribution of changes in atmospheric circulation patterns to extreme temperature trends. Nature, 522(7557):465–469.

* Hu, Z., Zhang, C., Hu,Q., and Tian, H. (2014) Temperature changes in Central Asia from 1979 to 2011 based on multiple datasets. Journal of Climate, 27(3):1143–1167.

* Iglesias A, Garrote L, Quiroga S, et al. (2012) A regional comparison of the effects of climate change on agricultural crops in europe. Climatic change 112:29–46.

* Jones P G, Thornton P K. (2003) The potential impacts of climate change on maize production in africa and latin america in 2055. Global environmental change 13(1):51–59.

* Kabir, R., Khan, H. T., Ball, E., and Caldwell, K. (2016) Climate change impact: the experience of the coastal areas of Bangladesh affected by cyclones Sidr and Aila. Journal of environmental and public health.

* Kaczan, D. J. and Orgill-Meyer, J. (2020) The impact of climate change on migration: a synthesis of recent empirical insights. Climatic Change, 158(3):281-300.

* Klein Tank, A. M., Peterson, T., Quadir, D., Dorji, S., Zou, X., Tang, H., Santhosh, K., Joshi, U., Jaswal, A., Kolli, R., et al. (2006) Changes in daily temperature and precipitation extremes in central and south Asia. Journal of Geophysical Research: Atmospheres, 111(D16).

* Laird, N. M. and Ware, J. H. (1982) Random-effects models for longitudinal data. Biometrics, pages 963–974.

* Levermann, A., Clark, P. U., Marzeion, B., Milne, G. A., Pollard, D., Radic, V., and Robinson, A. (2013) The multimillennial sea-level commitment of global warming. Proceedings of the National Academy of Sciences , 110(34):13745-13750.

* Li, Z. and Fang, H. (2016) Impacts of climate change on water erosion: A review. Earth-Science Reviews, 163:94-117.

* Luo, M., Liu, T., Meng, F., Duan, Y., Bao, A., Frankl, A., and De Maeyer, P. (2019) Spatiotemporal characteristics of future changes in precipitation and temperature in Central Asia. International Journal of Climatology, 39(3):1571–1588.

* Matthews, T. (2018) Humid heat and climate change. Progress in Physical Geography: Earth and Environment, 42(3):391-405.

* McCulloch, C. E. and Searle, S. R. (2004) Generalized, linear, and mixed models. JohnWiley & Sons.

* Ozturk, T., Turp, M. T., Türke³, M., and Kurnaz, M. L. (2017) Projected changes in temperature and precipitation climatology of Central Asia CORDEX Region 8 by using RegCM4 v3.5, Atmospheric Research, 183:296–307.

* Perelet, R. et al. (2007) Central Asia: Background paper on climate change. Fighting climate change: Human solidarity in a divided world, UNDP Human Development Report.

* Poloczanska, E. S., Brown, C. J., Sydeman, W. J., Kiessling, W., Schoeman, D. S., Moore, P. J., Brander, K., Bruno, J. F.,Buckley, L. B., Burrows, M. T., et al. (2013) Global imprint of climate change on marine life. Nature Climate Change, 3(10):919-925.

* Propastin, P., Kappas, M., and Muratova, N. (2008) Inter-annual changes in vegetation activities and their relationship to temperature and precipitation in Central Asia from 1982 to 2003. Journal of Environmental Informatics, 12(2).

* Rahmstorf, S., Cazenave, A., Church, J. A.,

Hansen, J. E., Keeling, R. F., Parker, D. E., and Somerville, R. C. (2007) Recent climate observations compared to projections. Science, 316(5825):709-709.

* Reyer, C. P., Otto, I. M., Adams, S., Albrecht, T., Baarsch, F., Cartsburg, M., Coumou, D., Eden, A., Ludi, E., Marcus, R., et al. (2017) Climate change impacts in Central Asia and their implications for development. Regional Environmental Change, 17(6):1639–1650.

* Seber, G. A. (2015) Nonlinear regression models. In The linear model and hypothesis, pages 117–128. Springer.

* Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, M., and Reinhardt, J. (2017) Climate change impacts in Sub-Saharan Africa: From physical changes to their social repercussions. Regional Environmental Change, 17(6):1585-1600.

* Singh, N. K., Bhattacharya, R., and Borrok, D. M. (2020) A Bayesian framework to unravel food, groundwater, and climate linkages: A case study from Louisiana. PloS one, 15(7):e0236757.

* Smith, T. M. and Reynolds, R. W. (2005) A global merged land-air-sea surface temperature reconstruction based on historical observations (1880–1997). Journal of climate, 18(12):2021–2036.

* Soares, M. d. O., Campos, C. C., Carneiro, P., Barroso, H., Marins, R. V., Teixeira, C. E. P., Menezes, M. O. B., Pinheiro, L. d. S., Viana, M. B., Feitosa, C. V., et al. (2021) Challenges and perspectives for the Brazilian semi-arid coast under global environmental changes. Perspectives in Ecology and Conservation, 19(3):267–278.

* Somanathan E, Somanathan R, Sudarshan A, et al. (2021) The impact of temperature on productivity and labor supply: Evidence from indian manufacturing. Journal of Political Economy 129(6):1797–1827, 2021.

* Springmann, M., Mason-D'Croz, D., Robinson, S., Garnett, T., Godfray, H. C. J., Gollin, D., Rayner, M., Ballon, P., and Scarborough, P. (2016) Global and regional health effects of future food production under climate change: A modelling study. The Lancet, 387(10031):1937-1946.

* Stocker, T. (2014) Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge university press.

* Stocker, T. (2014) Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge university press.

* Tao, H., Fraedrich, K., Menz, C., and Zhai, J. (2014) Trends in extreme temperature indices in the Poyang Lake Basin, china. Stochastic Environmental Research and Risk Assessment, 28(6):1543–1553.

* Tawn, J., Shooter, R., Towe, R., and Lamb, R.(2018) Modelling spatial extreme events with environmental applications. Spatial Statistics, 28:39–58.

* Tonkaz, T. and Çetin, M. (2007) Effects of urbanization and land-use type on monthly extreme temperatures in a developing semi-arid region, Turkey. Journal of arid environments, 68(1):143–158.

* Turner, M. G., Calder, W. J., Cumming, G. S., Hughes, T. P., Jentsch, A., LaDeau, S. L., Lenton, T. M., Shuman, B. N., Turetsky, M. R., Ratajczak, Z., et al. (2020) Climate change, ecosystems and abrupt change: Science priorities. Philosophical Transactions of the Royal Society B, 375(1794):20190105.

* Valipour, M., Bateni, S. M., and Jun, C. (2021) Global surface temperature: A new insight. Climate, 9(5), 2021.

* Verbeke, G. (1997) Linear mixed models for longitudinal data. In Linear mixed models in practice, pages 63–153. Springer.

* Wall, M. M.(2004) A close look at the spatial structure implied by the CAR and SAR models. Journal of statistical planning and inference, 121(2):311–324, 2004.

* Wang, H., Chen, Y., Chen, Z., and Li,W. (2013) Changes in annual and seasonal temperature extremes in the arid region of China, 1960–2010. Natural hazards, 65(3):1913–1930.

* Williams, M. and Konovalov, V. (2008) Central Asia temperature and precipitation data, 1879–2003. Boulder, Colorado: USA National Snow and Ice Data Center, 2008.

* Yin, G., Hu, Z., Chen, X., and Tiyip, T. (2016) Vegetation dynamics and its response to climate change in Central Asia. Journal of Arid Land, 8(3):375–388.

* Zarei, A. R., Mahmoudi, M. R., and Shabani, A. (2021) Investigating of the climatic parameters effectiveness rate on barley water requirement using the random forest algorithm, Bayesian multiple linear regression and cross-correlation function. Paddy andWater Environment, 19(1):137–148.

* Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., et al. (2017) Temperature increase reduces global yields of major crops in four independent estimates. Proceedings of the National Academy of Sciences, 114(35):9326-9331.

* Zhong, K., Zheng, F.,Wu, H., Qin, C., and Xu, X. (2017) Dynamic changes in temperature extremes and their association with atmospheric circulation patterns in the Songhua River Basin, China. Atmospheric Research, 190:77–88.

Appendix 1

Climate Stations in each country (minimum temperatures) in Central Asia (211-Kazakhstan, 213-Kyrgyzstan, 227-Tadjikistan, 229-Turkmenistan and 231-Uzbekistan)

Country	Clima Station
Kazakhstan	Almaty, Balhash, Big Almaty Lake, Karaganda, Kegen, Kugaly, Mynzhilky, Narynkol, Podgornoe, Sarydzhas, Semipalatinsk, Verhniy, Gorelnik.
Kyrgyzstan	Aktash, Angren, Arpa, Atbashi, Baityk, Bishkek, Chatirkul, Daraut-Kurgan, Dolon, Dzhergetal, Gulcha, Haidarkan, Isfana, Kochkorka, Naryn, Przhevalsk, Rybach'e, Sarytash, Susamyr, Tien-Shan, Tyuya-ashu, Southern, Ustie r.Ters.
Tadjikistan	Anzobsky pereval.Bulunkul, Bustonabad, Dehavz, Dushanbe, Dzhavshangoz, Fedchenko Glacier, Garm, Gushari, Haburabad, Haramkul, Horog, Hovaling, Humrogi, Irht, Ishkashim, Iskanderkul, Kalai-Khumb, Karakul, Kulyab, Kurgan-tyube, Leninabad, Lyairun, Lyakhsh, Madrushkent, Muminabad, Murgab, Rushan, Sangiston, Sanglok, Shahristanskiy Pereval, Shaimak, Tavildara, Uratyube.
Turkmenistan	Ashgabat, Bairam-ali, Chardzhou, Gasan-kuli, Kizyl-arvat, Krasnovodsk, Kushka, Serahs.
Uzbekistan	Akrabat, Fergana, Kizilcha, Minchukur, Oigaing, Pskem, Samarkand , Sanzar, Severtsova Glacier, Tamdy, Tashkent, Termez.

Appendix 2

Climate Stations in each country (maximum temperatures) in Central Asia (211 Kazakhstan, 213-Kyrgyzstan, 227-Tadjikistan, and 231-Uzbekistan)

Country	Clima Station
Kazakhstan	Almaty, Big Almaty Lake, Chilik, Issyk, Kegen', Kugaly, Mynzhilky, Narynkol, Podgornoe, Sarydzhas, Ust'-Gorelnik, Verhniy Gorelnik.
Kyrgyzstan	Aktash, Angren, Arpa, Atbashi, Baityk, Bishkek, Chatirkul, Daraut-Kurgan, Dolon, Dzhergetal, Gulcha, Haidarkan, Isfana, Kochkorka, Naryn, Przhevalsk, Rybach'e, Sarytash, Susamyr, Tien-Shan, Tyuya-ashu Southern, Ustie r.Ters.
Tadjikistan	Anzobsky pereval, Ayvadzh, Bulunkul, Bustonabad, Dushanbe, Dzhavshangoz, Faizabad, Fedchenko Glacier, Garm, Gushari, Haburabad, Haramkul, Horog, Hovaling, Humrogi, Irht, Isfara, Ishkashim, Iskanderkul, Kalai-Khumb, Karakul, Komsomolabad, Kulyab, Kurgan-tyube, Leninabad, Lyairun, Lyakhsh, Madrushkent, Muminabad, Murgab, Obigarm, Parhar, Pendjikent, Pyandzh(Kirovobad), Rushan, Sangiston, Sanglok, Shahrinau, Shahristanskiy Pereval, Shaimak, Tavildara, Uratyube.
Uzbekistan	Ablyk, Akrabat, Amankutan, Baisun, Bogarnoe, Charvak, Dehkanabad, Denau, Dukant, Dzhizak, Fedchenko, Fergana, Gallyaaral, Hiva, Kassansai, Kizilcha, Kokand, Minchukur, Namangan, Naugarzan, Nurata, Oigaing, Pskem, Samarkand, Sanzar, Severtsova, Glacier, Tamdy, Tashkent, Termez, Urgench.