# A Bayesian statistical model for deriving the predictive distribution of hydroclimatic variables

Hristos Tyralis and Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University, Athens Heroon Polytechneiou 5, GR-157 80 Zographou, Greece (montchrister@gmail.com)

**Abstract** Recent publications have provided evidence that hydrological processes exhibit a scaling behaviour, also known as the Hurst phenomenon. An appropriate way to model this behaviour is to use the Hurst-Kolmogorov stochastic process. The Hurst-Kolmogorov process entails high autocorrelations even for large lags, as well as high variability even at climatic scales. A problem that, thus, arises is how to incorporate the observed past hydroclimatic data in deriving the predictive distribution of hydroclimatic processes at climatic time scales. Here with the use of Bayesian techniques we create a framework to solve the aforementioned problem. We assume that there is no prior information for the parameters of the process and use a non-informative prior distribution. We apply this method with real-world data to derive the posterior distribution of the parameters and the posterior predictive distribution of various 30-year moving average climatic variables. The marginal distributions we examine are the normal and the truncated normal (for nonnegative variables). We also compare the results with two alternative models, one that assumes independence in time and one with Markovian dependence, and the results are dramatically different. The conclusion is that this framework is appropriate for the prediction of future hydroclimatic variables conditional on the observations.

**Key words** Bayesian statistics; hydroclimatic prediction; likelihood; Hurst-Kolmogorov process; hydrological statistics.

# 1 Introduction

A lot of work has been done in predicting the future of hydroclimatic processes using Bayesian statistics. Berliner et al. (2000) applied a Markov model to a low-order dynamical system of tropical Pacific SST, using a hierarchical Bayesian dynamical modelling, which led to realistic error bounds on forecasts. Duan et al. (2007) illustrated how the Bayesian model averaging (BMA) scheme can be used to generate probabilistic hydrologic predictions from several competing individual predictions. Kumar and Maity (2008) used two different Bayesian dynamic modelling approaches, namely a constant model and a dynamic regression model (DRM) to forecast the volume of the Devil's lake. Maity and Kumar (2006) used a Bayesian dynamic linear model to predict the monthly Indian summer monsoon rainfall. Bakker and Hurk (2012) used a Bayesian model to predict multi-year geostrophic winds.

On the other hand, climate models (i.e. general circulation models—GCMs) give deterministic projections of future hydroclimatic processes for some hypothesized scenarios e.g. for the increase of $CO_2$ concentration, etc. However, the uncertainty of these projections whose sources may be attributed to insufficient current understanding of climatic mechanisms, to inevitable weaknesses of numerical climatic and hydrologic models to represent processes and scales of interest, to complexity of processes and to unpredictability of causes (Koutsoyiannis et al. 2007), is not estimated by these models. Consequently, it is impossible to estimate whether any observed changes reflect the natural variability of the climatic processes or should be attributed to external forcings. Additionally, using deterministic projections and thus neglecting the uncertainty in future hydroclimatic conditions, may result in underestimation of possible range of the future hydroclimatic variation.

Koutsoyiannis et al. (2007) have done some work on the uncertainty assessment of future hydroclimatic predictions. They propose a stochastic framework for future climatic uncertainty, where climate is expressed by the 30-year time average of a natural process exhibiting a scaling behaviour, also known as the Hurst phenomenon or Hurst-Kolmogorov (HK) behaviour (Hurst 1951; Koutsoyiannis et al. 2008). To this end, they combine analytical and Monte Carlo methods to determine uncertainty limits and they apply the framework developed to temperature, rainfall and runoff data from a catchment in Greece, for which measurements are available for about a century.

In the study by Koutsoyiannis et al. (2007), the climatic variability and the influence of parameter uncertainty are studied separately. As a result, a hydroclimatic prediction needs two confidence coefficients to be defined, one referring to the uncertainty of the climatic evolution and one to the uncertainty of model parameters. In this paper we unify the study of the two uncertainties so that a climatic prediction needs only one confidence coefficient to be defined. To this end, we solve the problem of climatic predictions of natural processes using Bayesian statistics, instead of the stochastic framework developed by Koutsoyiannis et al. (2007). For physical consistency with natural processes such as rainfall and runoff, whose values are nonnegative, we also examine the case where truncation of the negative part of the distributions is applied. No prior information for the parameters of processes is assumed, so that the prior distribution is non-informative. The posterior joint distribution is derived from a mixture for the case where truncation is not applied and a Gibbs sampler for the case where truncation is applied. We derive the posterior predictive distribution (Gelman et al. 2004, p.8) of the process in closed form given the posterior distribution of the parameters. We simulate a sample from the posterior predictive distribution and use it to make inference about the future evolution of the averaged process. We apply this procedure using the same data as in Koutsoyiannis et al. (2007), and specifically runoff (Case 1 or C1), rainfall (C2) and temperature (C3) data from catchments in Greece and temperature data from Berlin (C4, C6 with the last 90 years excluded from the dataset); in addition we used temperature data from

Vienna (C5, C7 with the last 90 years excluded from the dataset). For the rainfall and runoff data we use truncated distributions.

As per the temporal dependence of the processes, three alternative assumptions are made: (a) independence in time; (b) Markovian dependence modelled by first-order autoregressive (AR(1)) process; and (c) HK dependence (see Markonis and Koutsoyiannis 2013, for a justification of the latter). In the last section we compare the results of the three models. Additional results such as the posterior distributions of the parameters and the asymptotic behaviour of the predictive distribution are also given.

While this paper uses the same case studies as those in Koutsoyiannis et al. (2007), the results of the two papers are not directly comparable to each other. Here we give posterior predictive distributions of the climatic variables, whereas Koutsoyiannis et al. (2007) give confidence limits for specified quantiles of climatic variables. The posterior predictive distribution of the variables given here is exactly what we call climatic prediction, whereas we could say that the confidence limits of the quantiles, given by Koutsoyiannis et al. (2007), are intermediate or indirect results. The Bayesian methodology applied here aims at (stochastic) prediction (Robert 2007, p.7) and is direct, while its disadvantage compared to Koutsoyiannis et al. (2007) framework is the much heavier computational burden.

## 2 Definition of AR(1) and HK process

We use the Dutch convention for notation, according to which random variables and stochastic processes are underlined (Hemelrijk 1966). We assume that $\{\underline{x}_t\}$, $t$ = 1, 2, … is a normal stationary stochastic process with mean $\mu := \mathrm{E}[\underline{x}_t]$, standard deviation $\sigma := \sqrt{\mathrm{Var}[\underline{x}_t]}$, autocovariance function $\gamma_k := \mathrm{Cov}[\underline{x}_t, \underline{x}_{t+k}]$ and autocorrelation function (ACF) $\rho_k := \mathrm{Corr}[\underline{x}_t, \underline{x}_{t+k}] = \gamma_k / \gamma_0$ ($k$ = 0, ±1, ±2, …) and that there is a record of $n$ observations $\boldsymbol{x}_n = (x_1 \ldots x_n)^\mathrm{T}$. Each observation $x_t$ represents a realization of a random variable $\underline{x}_t$, so that $\boldsymbol{x}_n$ is a realization of the vector of random variables $\underline{\mathbf{x}}_n = (\underline{x}_1 \ldots \underline{x}_n)^{\mathrm{T}}$.

We assume that $\{\underline{a}_t\}$ is a zero mean normal white noise process (WN), i.e. a sequence of independent random variables from a normal distribution with mean $\mathrm{E}[\underline{a}_t] = 0$ and variance $\mathrm{Var}[\underline{a}_t] = \sigma_a^2$. In the following discussion $\{\underline{a}_t\}$ is always referred to as WN. The following equation defines the first-order autoregressive process AR(1).

$$\underline{x}_t - \mu = \varphi_1(\underline{x}_{t-1} - \mu) + \underline{a}_t,\ |\varphi_1| < 1 \tag{1}$$

The ACF of the AR(1) is (Wei 2006, p.34)

$$\rho_k = \varphi_1^k,\ k = 0, 1, \ldots \tag{2}$$

Let $\kappa$ be a positive integer that represents a timescale larger than 1, the original time scale of the process $\underline{x}_t$. The averaged stochastic process on that timescale is denoted as

$$\underline{x}_t^{(\kappa)} := (1/\kappa) \sum_{l=(t-1)\kappa+1}^{t\kappa} \underline{x}_l \tag{3}$$

The notation implies that a superscript (1) could be omitted, i.e. $\underline{x}_t^{(1)} \equiv \underline{x}_t$. Now we consider the following equation that defines the Hurst-Kolmogorov stochastic process (HKp). (Koutsoyiannis 2003)

$$(\underline{x}_i^{(\kappa)} - \mu) \stackrel{\mathrm{d}}{=} \left(\frac{\kappa}{\lambda}\right)^{H-1} (\underline{x}_j^{(\lambda)} - \mu),\ 0 < H < 1,\ \ i, j = 1, 2, \ldots\ \text{ and } \kappa, \lambda \geq 1 \tag{4}$$

where $H$ is the Hurst parameter.

The ACF of the HKp is (Koutsoyiannis 2003)

$$\rho_k = |k+1|^{2H} / 2 + |k-1|^{2H} / 2 - |k|^{2H},\ k = 0, 1, \ldots \tag{5}$$

and does not depend on averaging time scale $\kappa$.

## 3 Posterior distribution of the parameters of a stationary normal stochastic process

The distribution of the variable $\underline{\boldsymbol{x}}_n = (\underline{x}_1 \ldots \underline{x}_n)^{\mathrm{T}}$ is

$$f(\boldsymbol{x}_n|\boldsymbol{\theta}) = (2\pi)^{-n/2}\, |\sigma^2 \boldsymbol{R}_n|^{-1/2} \exp[(-1/2\sigma^2)\, (\boldsymbol{x}_n - \mu\, \boldsymbol{e}_n)^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, (\boldsymbol{x}_n - \mu\, \boldsymbol{e}_n)] \tag{6}$$

where $\boldsymbol{R}_n$ is the autocorrelation matrix with elements $r_{ij} = \rho_{|i-j|}$, $i,j = 1,2, \ldots,n$ and $\boldsymbol{e}_n = (1\ 1 \ldots 1)^{\mathrm{T}}$ is a vector with $n$ elements. Details on the distributions used thereafter are given in Appendix 1. The autocorrelation $\rho_{|i-j|}$ is assumed to be function of a parameter (scalar or vector) $\boldsymbol{\varphi}$, so that $\boldsymbol{\theta} := (\mu, \sigma^2, \boldsymbol{\varphi})$ is the parameter vector of the process. We note that if $\underline{\boldsymbol{x}}_n$ is white noise then $\rho_0 = 1$ and $\rho_k = 0$, $k = 1, 2, \ldots$; if it is AR(1) then $\rho_k$ is given by (2) if it is HKp then $\rho_k$ is given by (5).

We assume that $\boldsymbol{\varphi}$ is uniformly distributed a priori. We set as prior distribution for $\underline{\boldsymbol{\theta}}$ the non-informative distribution (see also Robert 2007, example 3.5.6)

$$\pi(\boldsymbol{\theta}) \propto 1/\sigma^2 \tag{7}$$

(notice that we generally use the symbol $\pi$ for probability density functions of parameters).

The posterior distribution of the parameters does not have a closed form. However it can be calculated from a mixture based on conditional distributions. Specifically, it is shown (see Appendix 2) that

$$\underline{\mu}|\sigma^2, \boldsymbol{\varphi}, \boldsymbol{x}_n \sim N[(\boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)/(\boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n),\ \sigma^2/(\boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)] \tag{8}$$

$$\underline{\sigma}^2|\boldsymbol{\varphi}, \boldsymbol{x}_n \sim \text{Inv-gamma}\{(n-1)/2,\ [\boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{x}_n - (\boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)^2]/(2\, \boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)\} \tag{9}$$

$$\pi(\boldsymbol{\varphi}|\boldsymbol{x}_n) \propto |\boldsymbol{R}_n|^{-1/2}\, [\boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{x}_n - (\boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)^2]^{-(n-1)/2}\, (\boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)^{n/2-1} \tag{10}$$

As real world problems often impose upper or lower bounds on the variables $\underline{x}_t$, we assume that the distribution of $\underline{\boldsymbol{x}}_n$ is two-sided truncated by bounds $a$ and $b$, i.e.,

$$f(\boldsymbol{x}_n|\boldsymbol{\theta}) \propto \exp[(-1/2\sigma^2)\,(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)]\;\mathrm{I}_{[a,b]^n}(x_1, \ldots, x_n) \qquad (11)$$

where I denotes the indicator function, so that $\mathrm{I}_{[a,b]^n}(x_1, \ldots, x_n) = 1$ if $\boldsymbol{x}_n \in [a,b]^n$ and 0 otherwise.

We assume that the truncation set of $\mu$ is $[a,b]$, $a,b \in \mathrm{R}\cup\{-\infty,\infty\}$. The following Gibbs sampler is used to obtain a posterior sample from $\underline{\boldsymbol{\theta}} = (\underline{\mu}, \underline{\sigma}^2, \underline{\boldsymbol{\varphi}})$ (see Appendix 2).

$$\pi(\mu|\sigma^2, \boldsymbol{\varphi}, \boldsymbol{x}_n) \propto \exp\{-[\mu - (\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,\boldsymbol{e}_n)/(\boldsymbol{e}_n^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,\boldsymbol{e}_n)]^2/(2\sigma^2/\boldsymbol{e}_n^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,\boldsymbol{e}_n)\}\;\mathrm{I}_{[a,b]}(\mu) \qquad (12)$$

$$\underline{\sigma}^2|\mu, \boldsymbol{\varphi}, \boldsymbol{x}_n \sim \text{Inv-gamma}\{n/2, (\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)/2\} \qquad (13)$$

$$\pi(\boldsymbol{\varphi}|\,\mu, \sigma^2, \boldsymbol{x}_n) \propto |\boldsymbol{R}_n|^{-1/2}\exp[-(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)/2\sigma^2] \qquad (14)$$

## 4 Posterior predictive distributions

As we stated in the Introduction, we seek to make an inference about the future evolution of a process given observations of its past. To this end, in this section we derive the posterior predictive distributions of $\underline{\boldsymbol{x}}_{n+1,n+m}\,|\boldsymbol{x}_n$ for the cases of the white noise, the AR(1) and the HKp, where $\underline{\boldsymbol{x}}_{n+1,n+m} := (\underline{x}_{n+1},\ldots,\underline{x}_{n+m})^{\mathrm{T}}$.

### 4.1 White noise

We assume that $\underline{x}_t$, $t = 1, 2, \ldots$ is white noise, with $f(x_t|\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp[-(x_t-\mu)^2/(2\sigma^2)]$. A non-informative prior distribution for $\underline{\boldsymbol{\theta}} = (\underline{\mu},\underline{\sigma}^2)$ is $\pi(\boldsymbol{\theta}) \propto 1/\sigma^2$. The posterior distributions of the parameters are given by (Gelman et al. 2004, p.75-77)

$$\underline{\mu}|\boldsymbol{x}_n \sim \mathrm{t}_{n-1}(\bar{x}_n, s_n^2/(n-1)) \qquad (15)$$

$$\underline{\sigma}^2|\boldsymbol{x}_n \sim \text{Inv-gamma}((n-1)/2, n\,s_n^2/2) \qquad (16)$$

Notice that (15) and (16) are derived from (8),(9),(10) for $\boldsymbol{R}_n = \boldsymbol{I}_n$ (the former after integrating out $\sigma^2$). The posterior predictive distribution is

$$\underline{x}_t|\boldsymbol{x}_n \sim t_{n-1}(\overline{x}_n,((n+1)/(n-1))s_n^2),\ t = n+1, n+2, \ldots \tag{17}$$

where $\underline{x}_{n+1}, \underline{x}_{n+2},\ldots$ are mutually independent,

$$\overline{x}_n := \sum_{i=1}^{n} x_i/n \tag{18}$$

$$s_n^2 := \sum_{i=1}^{n} (x_i - \overline{x}_n)^2/n \tag{19}$$

are the maximum likelihood estimates of $\mu$ and $\sigma^2$ respectively and $t_v(\mu,\sigma^2)$ is the Student's distribution with $v$ degrees of freedom.

**4.2 AR(1) and HKp**

When there is dependence among the elements of $\underline{\boldsymbol{x}}_{n+m}$, the posterior predictive distribution of $\underline{\boldsymbol{x}}_{n+1,n+m}$ given $\boldsymbol{\theta}$ and $\boldsymbol{x}_n$ is (Eaton 1983, p.116,117)

$$f(\boldsymbol{x}_{n+1,n+m}|\boldsymbol{\theta},\boldsymbol{x}_n) = (2\pi\sigma^2)^{-m/2}\,|\boldsymbol{R}_{m|n}|^{-1/2}\exp[(-1/2\sigma^2)\,(\boldsymbol{x}_{n+1,n+m} - \boldsymbol{\mu}_{m|n})^{\mathrm{T}}\,\boldsymbol{R}_{m|n}^{-1}\,(\boldsymbol{x}_{n+1,n+m} - \boldsymbol{\mu}_{m|n})] \tag{20}$$

where $\boldsymbol{\mu}_{m|n}$ and $\boldsymbol{R}_{m|n}$ are given by:

$$\boldsymbol{\mu}_{m|n} = \mu\boldsymbol{e}_m + \boldsymbol{R}_{[(n+1):(n+m)]\,[1:n]}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n - \mu\boldsymbol{e}_n) \tag{21}$$

$$\boldsymbol{R}_{m|n} = \boldsymbol{R}_{[(n+1):(n+m)]\,[(n+1):(n+m)]} - \boldsymbol{R}^{\mathrm{T}}_{[1:n]\,[(n+1):(n+m)]}\,\boldsymbol{R}_n^{-1}\,\boldsymbol{R}_{[1:n]\,[(n+1):(n+m)]} \tag{22}$$

where $\boldsymbol{R}_{[k:l]\,[m:n]}$ is the submatrix of $\boldsymbol{R}$ which contains the elements $r_{ij}$, $k \le i \le l$, $m \le j \le n$, whereas the notation $\boldsymbol{R}_{[1:n]\,[1:n]}$ with identical subscripts $[1:n]$ can be simplified to $\boldsymbol{R}_n$ as defined above. The elements of the correlation matrices $\boldsymbol{R}_n$ and $\boldsymbol{R}_{m+n}$ are obtained from (2) for the case of the AR(1) and from (5) for the case of HKp. In the implementation of the AR(1) model we assume that all three parameters $\mu$, $\sigma$, $\varphi_1$ are unknown. For the HKp we examine two cases: (a) all three parameters $\mu$, $\sigma$, $H$, are unknown, and (b) $\mu$, $\sigma$, are unknown but $H$ is considered to be known and equal to its maximum likelihood estimate (Tyralis and Koutsoyiannis 2011).

In the case that all three parameters of the AR(1) or HKp are unknown, we obtain a simulated sample of $\underline{\boldsymbol{\theta}}$ from (8),(9),(10) and use this sample to simulate $\underline{\boldsymbol{\mu}}_{m|n}$ and $\underline{\boldsymbol{R}}_{m|n}$ from (21) and (22) and generate a sample of $\underline{\boldsymbol{x}}_{n+1,n+m}$ from (20). In the case where $H$ is considered as known, we obtain a simulated sample of $\underline{\boldsymbol{\theta}} = (\underline{\mu}, \underline{\sigma}^2)$ from (8),(9) and use this sample to simulate $\underline{\boldsymbol{\mu}}_{m|n}$ and $\underline{\boldsymbol{R}}_{m|n}$ from (21) and (22) and generate a sample of $\underline{\boldsymbol{x}}_{n+1,n+m}$ from (20).

### 4.3 Asymptotic behaviour of AR(1) and HKp

In most applications, it is useful to know the ultimate confidence regions as prediction horizon tends to infinity. This is expressed by the distribution of $\underline{\boldsymbol{x}}_{n+m+1,n+m+l} := (\underline{x}_{n+m+1},\ldots,\underline{x}_{n+m+l})$ as $m \to \infty$, conditional on $\boldsymbol{x}_n$. For given $\boldsymbol{\theta}$ this distribution is:

$$f(\boldsymbol{x}_{n+m+1,n+m+l}|\boldsymbol{\theta},\boldsymbol{x}_n) = (2\pi\sigma^2)^{-l/2}|\boldsymbol{R}_{l|n}|^{-1/2}\exp[(-1/2\sigma^2)(\boldsymbol{x}_{n+m+1,n+m+l} - \boldsymbol{\mu}_{l|n})^{\mathrm{T}}\boldsymbol{R}_{l|n}^{-1}(\boldsymbol{x}_{n+m+1,n+m+l} - \boldsymbol{\mu}_{l|n})] \quad (23)$$

where $\boldsymbol{\mu}_{l|n}$ and $\boldsymbol{R}_{l|n}$ are given by:

$$\boldsymbol{\mu}_{l|n} = \mu\boldsymbol{e}_l + \boldsymbol{R}_{[(n+m+1):(n+m+l)]\,[1:n]}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n - \mu\boldsymbol{e}_n) \quad (24)$$

$$\boldsymbol{R}_{l|n} = \boldsymbol{R}_{[(n+m+1):(n+m+l)]\,[(n+m+1):(n+m+l)]} - \boldsymbol{R}^{\mathrm{T}}_{[1:n]\,[(n+m+1):(n+m+l)]}\,\boldsymbol{R}_n^{-1}\,\boldsymbol{R}_{[1:n]\,[(n+m+1):(n+m+l)]} \quad (25)$$

We observe that, as $m \to \infty$, $\boldsymbol{R}_{[1:n]\,[(n+m+1):(n+m+l)]}$ and $\boldsymbol{R}_{[(n+m+1):(n+m+l)]\,[1:n]}$ become zero matrices and $\boldsymbol{R}_{[(n+m+1):(n+m+l)]\,[(n+m+1):(n+m+l)]} = \boldsymbol{R}_l$. This implies that:

$$\boldsymbol{\mu}_{l|n} = \mu\boldsymbol{e}_l \quad (26)$$

$$\boldsymbol{R}_{l|n} = \boldsymbol{R}_l \quad (27)$$

where $\boldsymbol{R}_l$ is again obtained from (2) for the case of the AR(1) and from (5) for the case of HKp.

Accordingly, the application can proceed as follows. We obtain a simulated sample of $\underline{\boldsymbol{\theta}}$

from (8),(9),(10) and use this sample to simulate $\underline{\boldsymbol{\mu}}_{l|n}$ and $\underline{\boldsymbol{R}}_{l|n}$ from (26) and (27) and generate a sample of $\underline{\boldsymbol{x}}_{n+m+1,n+m+l}$ from (23) for a large $m$.

### 4.4 Truncated white noise, AR(1) and HKp

To examine real world problems which often impose upper or lower bounds on the variables $\underline{x}_t$, we assume that the distribution of $\underline{\boldsymbol{x}}_n$ is two-sided truncated, and is given by (11). We obtain a posterior sample of $\underline{\boldsymbol{\theta}}$ using the Gibbs sampler defined by (12), (13), (14). When $\boldsymbol{\varphi}$ is known, we obtain a posterior sample of $(\underline{\mu},\underline{\sigma}^2)$ using the Gibbs sampler defined by (12) and (13). Then $\underline{\boldsymbol{x}}_m|\boldsymbol{\theta}$ follows a truncated normal multivariate distribution and according to Horrace (2005) the conditional multivariate distributions of $\underline{\boldsymbol{x}}_{n+1,n+m}|\boldsymbol{\theta},\boldsymbol{x}_n$ are again truncated normal. As a result (20) still holds after slight modifications and (21), (22) are valid. The posterior predictive distribution of $\underline{\boldsymbol{x}}_{n+1,n+m}|\boldsymbol{\theta},\boldsymbol{x}_n$ is then a multivariate truncated normal distribution:

$$f(\boldsymbol{x}_{n+1,n+m}|\boldsymbol{\theta},\boldsymbol{x}_n) \propto \exp[(-1/2\sigma^2)\,(\boldsymbol{x}_{n+1,n+m}-\boldsymbol{\mu}_{m|n})^{\mathrm{T}}\,\boldsymbol{R}_{m|n}^{-1}\,(\boldsymbol{x}_{n+1,n+m}-\boldsymbol{\mu}_{m|n})]\;\mathrm{I}_{[a,b]^m}(\boldsymbol{x}_{n+1,n+m}) \quad (28)$$

Now for the case of white noise, (15), (16) and (17) are not valid. But from (21), (22) and for $\rho_0 = 1$ and $\rho_k = 0$, $k = 1, 2, \ldots$, we obtain that $\boldsymbol{\mu}_{m|n} = \mu\boldsymbol{e}_m$ and $\boldsymbol{R}_{m|n} = \boldsymbol{R}_m$.

When looking for the asymptotic behaviour of the process, (23) still holds after slight modifications, according to Horrace (2005). As a result, the distribution of $\underline{\boldsymbol{x}}_{n+m+1,n+m+l}|\boldsymbol{\theta},\boldsymbol{x}_n$ is truncated multivariate normal, while (26) and (27) remain valid:

$$f(\boldsymbol{x}_{n+m+1,n+m+l}|\boldsymbol{\theta},\boldsymbol{x}_n) \propto$$

$$\propto \exp[(-1/2\sigma^2)(\boldsymbol{x}_{n+m+1,n+m+l}-\boldsymbol{\mu}_{l|n})^{\mathrm{T}}\boldsymbol{R}_{l|n}^{-1}(\boldsymbol{x}_{n+m+1,n+m+l}-\boldsymbol{\mu}_{l|n})]\;\mathrm{I}_{[a,b]^l}(\boldsymbol{x}_{n+m+1,n+m+l}) \quad (29)$$

### 4.5 Asymptotic convergence of MCMC

To simulate from (10) we use a random walk Metropolis-Hastings algorithm with a normal instrumental (or proposal) distribution (Robert and Casella 2004, p.271). We implement the algorithm using the function MCMCmetrop1R of the R package ‘MCMCpack’ (Martin et al., 2011). The variable 'burnin' in this package is given the value 0, whereas the other variables keep their default values.

There are a lot of methods to decide whether convergence can be assumed to hold for the generated sample (see Gamerman and Lopes 2006, p.157-169; Robert and Casella 2004, p.272-276). We use the methods of Heidelberger and Welch (1983) and Raftery and Lewis (1992). These methods are described by Smith (2007), whose notation we use here. We use the R package 'coda' (Plummer et al. 2011) to implement these methods. We assume that we have obtained a sample $\psi_1$, $\psi_2$,... of a scalar variable $\varphi$ using the MCMC algorithm.

The diagnostic of Heidelberger's method provides an estimate of the number of samples that should be discarded as a burn-in sequence and a formal test for non-convergence. The null hypothesis of convergence to a stationary chain is based on Brownian bridge theory and uses the Cramer-von-Mises test statistic $\int_0^1 B_n(t)^2 dt$, where

$$B_n(t) = (T_{\lfloor nt \rfloor} - \lfloor nt \rfloor \, \bar{\psi})/\sqrt{nS(0)} \tag{30}$$

$$T_k = \sum_{j=1}^{k} \psi_j,\ k = 1, 2, \ldots \text{ and } T_0 = 0 \tag{31}$$

where $\lfloor x \rfloor$ denotes the floor of $x$ (the greatest integer not greater than $x$) and $S(0)$ is the spectral density evaluated at frequency zero. In calculating the test statistic, the spectral density is estimated from the second half of the original chain. If the null hypothesis is rejected, then the first 0.1$n$ of the samples are discarded and the test is reapplied to the resulting chain. This process is repeated until the test is either non-significant or 50% of the samples have been discarded, at which point the chain is declared to be non-stationary. For more details see Smith (2007).

The methods of Raftery and Lewis are designed to estimate the number of MCMC samples needed when quantiles are the posterior summaries of interest. Their diagnostic is applicable for the univariate analysis of a single parameter and chain. For instance, let us consider the estimation of the following posterior probability of a model parameter $\theta$:

$$P(f(\theta) < a \mid \boldsymbol{x}) = q \tag{32}$$

where $\boldsymbol{x}$ denotes the observed data. Raftery and Lewis sought to determine the number of MCMC samples to generate and the number of samples to discard in order to estimate $q$ to within $\pm r$ with probability $s$. In practice, users specify the values of $q$, $r$ and $s$ to be used in applying the diagnostic (For more details see Smith, 2007).

To simulate from (14) we use an accept-reject algorithm (Robert and Casella 2004, p.51-53) with a uniform instrumental density. Simulation from (12) and (13) is trivial. We assess the convergence of the chain simulated from (12), (13), (14) using the method of Gelman and Rubin (1992; see also Gelman 1996; Gamerman and Lopes 2006, p.166-168). An indicator of convergence is formed by the estimator of a potential scale reduction (PSR) that is always larger than 1. Convergence can be evaluated by the proximity of PSR to 1. Gelman (1996) suggested accepting convergence when the value of PSR is below 1.2.

# 5 Case studies

In this section we apply the methodology developed in the previous sections to five historical datasets; three of them obtained from the Boeoticos Kephisos River basin, one from Berlin and one from Vienna. The choice of these datasets was dictated by the fact that they have been also studied in other works with similar objectives, i.e. Koutsoyiannis et al. (2007) and Koutsoyiannis (2011), so that the interested reader can make some comparisons. We present the results of the application of the methodology to the aforementioned datasets.

## 5.1 Historical datasets

The first case study is performed on an important catchment in Greece, which is part of the water supply system of Athens and has a history, as regards hydraulic infrastructure and management that extends backward at least 3500 years. This is the closed (i.e. without outlet to the sea) basin of the Boeoticos Kephisos River (Figure 1), with an area of 1955.6 $km^2$, mostly formed over a karstic subsurface. Owing to its importance for irrigation and water supply, data availability for the catchment extends for about 100 years (the longest dataset in Greece) and modelling attempts with good performance have already been carried out on the hydrosystem (Rozos et al. 2004).

The long-term dataset for the basin extends from 1908 to 2003 and comprises a flow record at the river outlet at the Karditsa station (C1), rainfall observations in the raingage Aliartos (C2) and a temperature record at the same station (C3); the station locations are shown in Figure 1. Further details on the construction of these datasets are given by Koutsoyiannis et al. (2007). The relatively long records have already made it possible to identify the scaling behaviour of rainfall and runoff in this basin (Koutsoyiannis 2003), and make the catchment ideal for a case study of uncertainty assessment.

The two other datasets which we use are the mean annual temperature record of Berlin/Templehof and Vienna, two of the longest series of instrumental meteorological observations. For further details on the Berlin mean annual temperature dataset see Koutsoyiannis et al. (2007) and for the Vienna mean annual temperature dataset see Koutsoyiannis (2011). We examine two cases. In the first case we assume that the update of the prior information is done (C4, C5), using the whole dataset. In the second case the update is done excluding the last 90 years of the datasets (C6, C7).

### 5.2 Application of the method

We classified the data into three classes, the first containing the data from the Boeoticos Kephisos River basin (C1-C3), the second containing the data from Berlin and Vienna (C4, C6) and the third containing again the data from Berlin and Vienna (C5, C7) but excluding the last 90 years. In the third case the posterior results were compared to the actual 90 last years.

First we calculated the maximum likelihood estimates of the parameters for all the examined cases (WN, AR(1), HKp). The results are given in Tables 1a and 1b. Truncated models were used for C1 and C2 datasets due to the relatively high estimated $\sigma$ which otherwise would result in negative values. Instead, when we examined the temperature datasets (C3-C7), simulated values near the absolute zero never appeared, indicating a good behaviour of the non-truncated model.

The procedure for the temperature datasets is described below. We used (15) and (16) to generate a posterior sample from $\underline{\mu}$ and $\underline{\sigma}^2$ for the WN case. To simulate from (10) for the $\underline{\varphi}_1$ and $\underline{H}$ posterior distribution of the AR(1) and HK cases correspondingly, we used a random walk Metropolis-Hastings algorithm. We simulated a single chain with 3 000 000 MCMC samples. The Metropolis acceptance rates are given in Table 2. To decide whether convergence has been achieved, we used the Heidelberger and Welch method (1983). We tested four cases, the first case containing all the 3 000 000 samples, the second containing the last 2 000 000 samples and so forth. The results are presented in Tables 3a and 3b, from where we conclude that stationary chain hypothesis holds in every case. We also used the methods of Raftery and Lewis (1992), to estimate the number of MCMC samples needed when quantiles are the posterior summaries of interest. The minimum number of samples and the burn-in period for the simulation is given in Tables 4a and 4b, where $q$ = 0.025, 0.500, 0.975 are the quantiles to be estimated, $r$ = 0.005 is the desired margin of error of the estimate and $s$ = 0.95 is the probability of obtaining an estimate in the interval ($q-r$, $q+r$). We decided to use the last 2 000 000 samples of the chains, to obtain the histograms of the posterior distributions of the parameters $\underline{\varphi}_1$ and $\underline{H}$. The simulation of $\underline{\mu}$, $\underline{\sigma}^2$ from (8) and (9) is then trivial. Summarized results for the parameters of the AR(1) and HK cases respectively are shown in Tables 5a and 5b.

From the simulated samples we obtained the posterior probability plots of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the AR(1) and HK cases (Figures 2 and 3). The last 100 000 simulated samples of the parameters, described in the previous paragraph were used to obtain samples from the required posterior predictive probabilities. The samples from the posterior predictive probability of $\underline{x}_t|\boldsymbol{x}_n$, $t = n+1, n+2, ..., n+90$ were used to obtain samples for the variable of interest $\underline{x}_t^{(30)}$ given by (33).

$$\underline{x}_t^{(30)} := (1/30)\Big(\sum_{l=t-29}^{n} x_l + \sum_{l=n+1}^{t} \underline{x}_l\Big),\ t = n+1, \ldots, n+29 \text{ and } \underline{x}_t^{(30)} := (1/30)\sum_{l=t-29}^{t} \underline{x}_l,\ t = n+30, n+31, \ldots \quad (33)$$

We examined the cases of WN, AR(1), asymptotic behaviour of AR(1), HK where $H$ is considered to be known and has the value of the maximum likelihood estimate, HK when $H$ is not known, and its asymptotic behaviour. Figures 4, 5a, 5b show the 0.025, 0.500 and 0.975 quantiles of the posterior predictive distributions of $\underline{x}_t^{(30)}|\boldsymbol{x}_n$, $t = n+1, n+2, ..., n+90$.

The procedure for C1 and C2 is described below. We simulated from (12), (13) and (14) to obtain a posterior sample from $\underline{\mu}$, $\underline{\sigma}^2$ and $\underline{\varphi}$ for all cases. We simulated 10 chains with each one having 300 000 MCMC samples. To decide whether convergence has been achieved, we used the Gelman and Rubin (1992) rule. In all cases PSR ≈ 1 which shows that the chains converged to the target distribution. We decided to use the last 200 000 samples of each chain, to obtain the histograms of the posterior distributions of the parameters $\underline{\varphi}_1$ and $\underline{H}$. Summarized results for the parameters of the AR(1) and HK cases respectively are shown in Table 5a.

From the simulated samples we obtained the posterior probability plots of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the AR(1) and HK cases (Figure 2a, 2b). The last 10 000 simulated samples of the parameters of each chain, described in the previous paragraph are used to obtain samples from the required posterior predictive probabilities. The samples from the posterior predictive probability of $\underline{x}_t|\mathbf{x}_n$, $t = n+1, n+2, ..., n+90$ are used to obtain samples for the variable of interest $\underline{x}_t^{(30)}$ given by (33). We examined the cases of WN, AR(1), asymptotic behaviour of AR(1), HK where $H$ is considered to be known and has the value of the maximum likelihood estimate, HK with unknown $H$ and its asymptotic behaviour. Figure 4 shows the 0.025, 0.500 and 0.975 quantiles of the posterior predictive distributions of $\underline{x}_t^{(30)}|\mathbf{x}_n$, $t = n+1, n+2, ..., n+90$.

**5.3 Results**

A first important result of the proposed framework is that it provides good estimates of the model parameters without introducing any assumptions (i.e., using non-informative priors). While common statistical methods give point estimates of parameters, the Bayesian framework provides also interval estimates based on their posterior distributions. The estimated values of $\mu$ are given in Table 6. It turns out that irrespective of the method used (MLE or posterior medians) they are almost equal. When examining temperatures, HKp resulted in the largest $\hat{\mu}$ and AR(1) in the second largest. In C4 and C6, $\hat{\mu}$ was larger than in C5 and C7 respectively. From the density diagrams of the posterior distributions (Figures 2-3) it seems that the posterior distribution of $\underline{\mu}$ is wider when HKp is used. The posterior distribution of $\underline{\sigma}$ is also wider on the right (see the values of the 0.975 quantiles in Tables 5a,5b) for the HKp. However the estimated values of $\sigma$ are almost equal for the three used models (Tables 1a and 1b). The estimated $\varphi_1$ and $H$ are given in Tables 1a and 1b. Their estimated values for C5 are considerably higher compared to C7, but their posterior distributions are narrower (Table 5b), probably because of the bigger sample size in the former case. Their posterior distributions are also narrower for C4 compared to C6.

The second result of the framework is the predictive distribution of the future evolution of the process of interest. The posterior predictive 0.95-confidence regions for the 30-year moving averages are given in Figures 4, 5a and 5b. For C1 the confidence region is not symmetric with respect to the estimated mean, owing to the lower truncation bound alongside with the relatively big $\hat{\sigma}$. In contrast, there is a symmetry for C2 owing to the relatively small $\hat{\sigma}$, which justifies our decision to use models without truncation in those cases where $\hat{\sigma}$ is even smaller (compared to mean). For all cases, the widest confidence regions correspond to the HKp (due to the existence of persistence), followed by the AR(1), while the narrowest confidence regions appear for the WN. Of course the confidence regions for unknown $H$ are wider than in the case where $H$ was considered to be known and equal to its maximum likelihood estimate. In C5 and C7 the HKp seems to be the best model, because it captures better than the others the observed values of the climate variable for the last 90 years based on the observed values of the previous years. In C7 it seems that the HKp did not capture the increase of temperature in last decades. But when we examine the full dataset (C5), the behaviour in last 90 years does not appear extraordinary. For the asymptotic values in the HKp, the 0.95-confidence region ranges at intervals of the order of 150 mm (C1), 220 mm (C2), 1.6°C (C3), 1.9°C (C4), 1.4°C (C5) for the 30-year moving average. The corresponding values for the case of the WN of the order of 50 mm (C1), 75 mm (C2), 0.5°C (C3), 0.6°C (C4), 0.6°C (C5) are considerably smaller compared to the case of the HKp.

## 6 Summary

We developed a Bayesian statistical methodology to make hydroclimatic prognosis in terms of estimating future confidence regions on the basis of a stationary normal stochastic process. We applied this methodology to five cases, namely the runoff (C1), the rainfall (C2) and the temperature (C3) at Boeoticos Kephisos river basin in Greece, as well as the temperature at Berlin (C4, C6) and the temperature at Vienna (C5, C7). The Bayesian statistical model consisted of a stationary normal process (or truncated stationary normal process for the runoff and rainfall cases) with a non-informative prior distribution. Three kinds of stationary normal processes were examined, namely WN, AR(1) and HKp. We derived the posterior distributions of the parameters of the models, the posterior predictive distributions of the variables of the process and the posterior predictive distribution of the 30-year moving average which was the climatic variable of interest. The methodology can also be applied to other structures of the ACF.

A first important conclusion is that for all the examined cases and for all the examined processes their estimated means are almost equal as expected. However the posterior distributions of the means are wider when using the HKp, due to the persistence of the process, and even wider when all parameters of the process are assumed to be unknown. This results in wider confidence regions for future climatic variables of the processes. Moreover the confidence regions of truncated future variables are asymmetric. This asymmetry depends on the variance of the examined process. However the posterior distributions of the means of all processes were less asymmetric.

Another important conclusion is that the use of short-range dependence stochastic processes is not suitable to model geophysical processes, because they underestimate uncertainty. However stationary persistent stochastic processes are suitable to achieve this purpose. In the examined cases they performed well and were able to explain the fluctuations of the process.

One may claim that, when climate is to be predicted, an assumption of stationarity is not an appropriate one as currently several climate models project a changing future climate. Nonetheless, an assessment of future climate variability and uncertainty based on the stationarity hypothesis is a necessary step in establishing a stochastic method, whose generalization at a second step would enable incorporating nonstationary components. In addition, without knowing the variability under stationary conditions, it would not be possible to quantify the credibility of climate models and even their usefulness. Work on the generalization of the methodology to incorporate deterministic predictions by climate models is under way and its results will be reported in due course.

## Appendix 1: Standard probability distributions

For easy reference, the details of the distribution functions used in this paper are summarized in Table 7.

**Table 7**. Distributions used in the Bayesian framework

| Distribution | Notation | Parameters | Density function |
|---|---|---|---|
| Normal | $\underline{x} \sim \mathrm{N}(\mu,\sigma^2)$ | location $\mu$<br>scale $\sigma > 0$ | $f_{\mathrm{N}}(x \mid \mu,\sigma^2) = (2\pi\sigma^2)^{-1/2} \exp[(-1/2\sigma^2)\,(x-\mu)^2]$ |
| Truncated normal | $\underline{x} \sim \mathrm{TN}(\mu,\sigma^2,a,b)$ | location $\mu$<br>scale $\sigma > 0$<br>$a$ minimum value<br>$b$ maximum value | $f_{\mathrm{TN}}(x \mid \mu,\sigma^2,a,b) = [f_{\mathrm{N}}((b-\mu)/\sigma) - f_{\mathrm{N}}((a-\mu)/\sigma)]^{-1}(1/\sigma)\, f_{\mathrm{N}}((x-\mu)/\sigma)$<br>$x \in [a,b]$, $f_{\mathrm{N}}(x) := f_{\mathrm{N}}(x \mid 0,1^2)$ |
| Multivariate normal | $\underline{\boldsymbol{x}} \sim \mathrm{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$<br>(implicit dimension $n$) | location $\boldsymbol{\mu}$<br>symmetric, pos. definite<br>$n$ x $n$ variance matrix $\boldsymbol{\Sigma}$ | $f_{\mathrm{MN}}(\boldsymbol{x} \mid \boldsymbol{\mu},\boldsymbol{\Sigma}) = (2\pi)^{-n/2}\, \lvert\boldsymbol{\Sigma}\rvert^{-1/2} \exp[(-1/2)\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\, \boldsymbol{\Sigma}^{-1}\, (\boldsymbol{x}-\boldsymbol{\mu})]$ |
| Inverse-gamma | $\underline{x} \sim \text{Inv-gamma}(\alpha,\beta)$ | shape $\alpha > 0$<br>scale $\beta > 0$ | $f_{\mathrm{IG}}(x \mid \alpha,\beta) = \beta^{\alpha}\, [\Gamma(\alpha)]^{-1}\, x^{-(\alpha+1)} \exp(-\beta/x)$, $x > 0$ |
| Student-t | $\underline{x} \sim t_n(\mu,\sigma^2)$ | degrees of freedom $n$<br>location $\mu$<br>scale $\sigma > 0$ | Not needed in the manuscript |

## Appendix 2: Mathematical proofs

In Appendix 2 the proofs of (8),(9),(10),(12),(13),(14) are given. It is easily shown that

$$(\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, (\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n) = \boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \mu^2 - 2\, \boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \mu + \boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{x}_n \qquad (34)$$

After completing the squares the above expression becomes:

$$\boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \mu^2 - 2\, \boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \mu + \boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{x}_n = \boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, [\mu - (\boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n) / (\boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)]^2 + [\boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n\, \boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{x}_n - (\boldsymbol{x}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n)^2] / (\boldsymbol{e}_n^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, \boldsymbol{e}_n) \qquad (35)$$

From (6) and (7) we obtain the following:

$$\pi(\boldsymbol{\theta})\, f(\boldsymbol{x}_n \mid \boldsymbol{\theta}) \propto \sigma^{-(n+2)}\, \lvert\boldsymbol{R}_n\rvert^{-1/2} \exp[(-1/2\sigma^2)\, (\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)^{\mathrm{T}}\, \boldsymbol{R}_n^{-1}\, (\boldsymbol{x}_n - \mu\,\boldsymbol{e}_n)] \qquad (36)$$

From (34),(35) and (36) we obtain (8). After integration of (36) we obtain (37) which proves (9):

$$\pi(\sigma^2|\boldsymbol{\varphi},\boldsymbol{x}_n) \propto (\sigma^2)^{-(n+1)/2}\,|\boldsymbol{R}_n|^{-1/2}\,\exp[(-1/2\sigma^2)[\boldsymbol{e}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{e}_n\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{x}_n-(\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{e}_n)^2]/(\boldsymbol{e}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{e}_n)] \quad (37)$$

After integration of (36) we obtain (38), which proves (10) after integration:

$$\pi(\boldsymbol{\varphi}|\boldsymbol{x}_n) \propto \iint \sigma^{-(n+2)}\,|\boldsymbol{R}_n|^{-1/2}\,\exp[(-1/2\sigma^2)\,(\boldsymbol{x}_n-\mu\,\boldsymbol{e}_n)^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n-\mu\,\boldsymbol{e}_n)]\,\mathrm{d}\mu\,\mathrm{d}\sigma^2 \quad (38)$$

See also Falconer and Fernadez (2007) for some results.

Now for the case where truncation is applied we obtain from (7) and (11):

$$\pi(\boldsymbol{\theta})\,f(\boldsymbol{x}_n|\boldsymbol{\theta}) \propto \sigma^{-(n+2)}\,|\boldsymbol{R}_n|^{-1/2}\,\exp[(-1/2\sigma^2)\,(\boldsymbol{x}_n-\mu\,\boldsymbol{e}_n)^{\mathrm{T}}\,\boldsymbol{R}_n^{-1}\,(\boldsymbol{x}_n-\mu\,\boldsymbol{e}_n)]\,\mathrm{I}_{[a,b]^n}(x_1,\ldots,x_n) \quad (39)$$

Conditional on $\mu \in [a,b]$, $a,b \in \mathrm{R}\cup\{-\infty,\infty\}$the derivation of (12), (13) and (14) from (39) is then trivial.

**Acknowledgements**: The authors wish to thank the eponymous reviewer Dr. Federico Lombardo and an anonymous reviewer for their encouraging and constructive comments which helped to improve the quality of the manuscript significantly.

## References

Berliner LM, Wikle CK, Cressie N (2000) Long-Lead Prediction of Pacific SSTs via Bayesian Dynamic Modeling. J. Climate 13:3953-3968

Bakker A, Hurk B (2012) Estimation of persistence and trends in geostrophic wind speed for the assessment of wind energy yields in Northwest Europe. Climate Dynamics 39(3-4):767-782. doi: 10.1007/s00382-011-1248-1

Duan Q, Ajami NK, Gao X, Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resources 30(5):1371-1386

Eaton ML (1983) Multivariate Statistics: a Vector Space Approach. Institute of Mathematical Statistics, Beachwood, Ohio

Falconer K, Fernadez C (2007) Inference on fractal processes using multiresolution approximation. Biometrica 94(2):313-334

Gamerman D, Lopes H (2006) Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference, second edition, Chapman & Hall/CRC, London

Gelman A (1996) Inference and monitoring convergence. In: Markov Chain Monte Carlo in Practice (ed Gilks WR, S. Richardson S, Spiegelhalter DJ). Chapman & Hall, New York, 131-143

Gelman A, Carlin J, Stern H, Rubin D (2004) Bayesian Data Analysis. second edition, Chapman & Hall/CRC, Boca Raton, FL

Gelman A, Rubin DR (1992) A single series from the Gibbs sampler provides a false sense of security. In: Bayesian Statistics 4. (ed Bernardo JM, Berger JO, Dawid AP, Smith AFM). Oxford University Press, Oxford, 625-632

Heidelberger P, Welch PD (1983) Simulation run length control in the presence of an initial transient. Operations Research 31(6):1109-1144. doi:10.1287/opre.31.6.1109

Hemelrijk J (1966) Underlining random variables. Statistica Neerlandica 20:1-7. doi:10.1111/j.1467-9574.1966.tb00488.x

Horrace W (2005) Some results on the multivariate truncated normal distribution. Journal of Multivariate Analysis 94(1):209-221

Hurst HE (1951) Long term storage capacities of reservoirs. Transactions of the American Society of Civil Engineers 116:776-808 (published in 1950 as Proceedings Separate no.11)

Koutsoyiannis D (2003) Climate change, the Hurst phenomenon, and hydrological statistics. Hydrological Sciences Journal 48(1):3–24. doi:10.1623/hysj.48.1.3.43481

Koutsoyiannis D (2011) Hurst-Kolmogorov dynamics as a result of extremal entropy production. Physica A: Statistical Mechanics and its Applications 390(8):1424–1432

Koutsoyiannis D, Efstratiadis A, Georgakakos KP (2007) Uncertainty Assessment of Future Hydroclimatic Predictions: A Comparison of Probabilistic and Scenario-Based

Approaches. Journal of Hydrometeorology 8(3):261-281. doi:http://dx.doi.org/10.1175/JHM576.1

Koutsoyiannis D, Efstratiadis A, Mamassis N, Christofides A (2008) On the credibility of climate predictions. Hydrological Sciences Journal 53(4):671-684. doi:10.1623/hysj.53.4.671

Kumar DN, Maity R (2008) Bayesian dynamic modeling for nonstationary hydroclimatic time series forecasting along with uncertainty quantification. Hydrological Processes 22(17):3488-3499. doi:10.1002/hyp.6951

Maity R, Kumar DN (2006) Bayesian dynamic modeling for monthly Indian summer monsoon using El Nino-Southern Oscillation (ENSO) and Equatorial Indian Ocean Oscillation (EQUINOO). Journal of Geophysical Research 111 D07104. doi:10.1029/2005JD006539

Markonis Y, Koutsoyiannis D (2013) Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst-Kolmogorov dynamics. Surveys in Geophysics 34(2):181–207

Martin A, Quinn K, Park JH (2011) MCMCpack: Markov Chain Monte Carlo (MCMC). R package version 1.2-1, URL http://cran.r-project.org/web/packages/MCMCpack/index.html

Plummer M, Best N, Cowles K, Vines K (2011) coda: Output analysis and diagnostics for MCMC. R package version 0.14-6, URL http://cran.r-project.org/web/packages/coda/index.html

Raftery AL, Lewis S (1992) How many iterations in the Gibbs sampler? In: Bayesian Statistics 4. (ed Bernardo JM, Berger JO, Dawid AP, Smith AFM). Oxford University Press, Oxford, 763-774

Robert C (2007) The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation. Springer, New York

Robert C, Casella G (2004) Monte Carlo Statistical Methods. second edition, Springer-Verlag New York, Inc., Secaucus, NJ

Rozos E, Efstratiadis A, Nalbantis I, Koutsoyiannis D (2004) Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows. Hydrological Sciences Journal 49(5):819-842

Smith B (2007) boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. Journal of Statistical Software 21(11):1-37

Tyralis H, Koutsoyiannis D (2011) Simultaneous estimation of the parameters of the Hurst-Kolmogorov stochastic process. Stochastic Environmental Research & Risk Assessment 25(1):21-33. doi:10.1007/s00477-010-0408-x

Wei WWS (2006) Time Series Analysis, Univariate and Multivariate Methods. second edition, Pearson Addison Wesley, Chichester

**Table 1a**. Summarized results and maximum likelihood estimates for the cases of WN, AR(1) and HKp at Boeoticos Kephisos River basin.

| | Boeoticos basin | | |
|---|---|---|---|
| | Runoff (mm) | Rainfall (mm) | Temperature (°C) |
| Start year | 1908 | 1908 | 1898 |
| End year | 2003 | 2003 | 2003 |
| Size, n | 96 | 96 | 106 |
| WN | | | |
| $\hat{\mu}$ | 197.63 | 658.36 | 16.96 |
| $\hat{\sigma}$ | 81.25 | 155.82 | 0.69 |
| AR(1) | | | |
| $\hat{\mu}$ | 197.65 | 658.22 | 16.96 |
| $\hat{\sigma}$ | 81.22 | 155.81 | 0.69 |
| $\hat{\varphi}_1$ | 0.34 | 0.10 | 0.31 |
| HK | | | |
| $\hat{\mu}$ | 195.11 | 657.38 | 16.97 |
| $\hat{\sigma}$ | 80.47 | 155.00 | 0.70 |
| $\hat{H}$ | 0.71 | 0.60 | 0.71 |

**Table 1b**. Summarized results and maximum likelihood estimates for the cases of WN, AR(1) and HKp at Berlin and Vienna.

| | Berlin Temperature (°C) | Vienna Temperature (°C) | Berlin Temperature (°C) | Vienna Temperature (°C) |
|---|---|---|---|---|
| Start year | 1756 | 1775 | 1756 | 1775 |
| End year | 2009 | 2009 | 1919 | 1919 |
| Size, n | 254 | 235 | 164 | 145 |
| WN | | | | |
| $\hat{\mu}$ | 9.17 | 9.58 | 9.04 | 9.36 |
| $\hat{\sigma}$ | 0.91 | 0.87 | 0.92 | 0.84 |
| AR(1) | | | | |
| $\hat{\mu}$ | 9.18 | 9.58 | 9.05 | 9.36 |
| $\hat{\sigma}$ | 0.92 | 0.87 | 0.92 | 0.84 |
| $\hat{\varphi}_1$ | 0.37 | 0.30 | 0.30 | 0.11 |
| HK | | | | |
| $\hat{\mu}$ | 9.27 | 9.64 | 9.10 | 9.37 |
| $\hat{\sigma}$ | 0.91 | 0.86 | 0.92 | 0.84 |
| $\hat{H}$ | 0.73 | 0.70 | 0.70 | 0.59 |

**Table 2**. Metropolis acceptance rate for the MCMC simulation of $\underline{\varphi}_1$ and $\underline{H}$, respectively, at Boeoticos Kephisos River basin.

| | Aliartos temperature | Berlin temperature (1756-2009) | Vienna temperature (1775-2009) | Berlin temperature (1756-1919) | Vienna temperature (1775-1919) |
|---|---|---|---|---|---|
| $\varphi_1$ | 0.70731 | 0.70603 | 0.70612 | 0.70649 | 0.70654 |
| $H$ | 0.706037 | 0.70551 | 0.70599 | 0.70601 | 0.70638 |

**Table 3a**. Heidelberger and Welch test, for significance level 0.05, at Boeoticos Kephisos River basin.

| | Aliartos temperature | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Parameter | $\varphi_1$ | | | | *H* | | | |
| Stationarity test | passed | passed | passed | passed | passed | passed | passed | passed |
| Start iteration | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *p*-value | 0.427 | 0.745 | 0.46 | 0242 | 0.869 | 0.567 | 0.338 | 0.618 |

**Table 3b**. Heidelberger and Welch test, for significance level 0.05, at Berlin and Vienna.

| | Berlin temperature (1756-2009) | | | | Vienna temperature (1775-2009) | | | |
|---|---|---|---|---|---|---|---|---|
| Data start | 1 | 1000000 | 2000000 | 2900000 | 1 | 1000000 | 2000000 | 2900000 |
| Parameter | $\varphi_1$ | | | | $\varphi_1$ | | | |
| Stationarity test | passed | passed | passed | passed | passed | passed | passed | passed |
| Start iteration | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *p*-value | 0.943 | 0.738 | 0.342 | 0.448 | 0.928 | 0.696 | 0.366 | 0.0761 |
| Parameter | *H* | | | | *H* | | | |
| Stationarity test | passed | passed | passed | passed | passed | passed | passed | passed |
| Start iteration | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *p*-value | 0.837 | 0.466 | 0.279 | 0.691 | 0.789 | 0.501 | 0.296 | 0.84 |
| | Berlin temperature (1756-1919) | | | | Vienna temperature (1775-1919) | | | |
| Parameter | $\varphi_1$ | | | | $\varphi_1$ | | | |
| Stationarity test | passed | passed | passed | passed | passed | passed | passed | passed |
| Start iteration | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *p*-value | 0.94 | 0.589 | 0.376 | 0.425 | 0.777 | 0.55 | 0.308 | 0.592 |
| Parameter | *H* | | | | *H* | | | |
| Stationarity test | passed | passed | passed | passed | passed | passed | passed | passed |
| Start iteration | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *p*-value | 0.833 | 0.606 | 0.339 | 0.923 | 0.885 | 0.83 | 0.373 | 0.323 |

**Table 4a**. Raftery and Lewis test for the case of Boeoticos Kephisos River basin.

| | | Aliartos temperature | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *q* | Burn-in | Total | Lower bound | Dependence factor | | Burn-in | Total | Lower bound | Dependence factor |
| $\varphi_1$ | 0.025 | 21 | 31794 | 3746 | 8.49 | *H* | 18 | 35784 | 4899 | 7.3 |
| | 0.500 | 24 | 356752 | 38415 | 9.29 | | 24 | 464024 | 50239 | 9.24 |
| | 0.975 | 28 | 32298 | 3746 | 8.62 | | 28 | 42161 | 4899 | 8.61 |

Note: $q$ is the quantile to be estimated, $r$ = 0.005 is the desired margin of error of the estimate, $s$ = 0.95 the probability of obtaining an estimate in the interval ($q$−$r$, $q$+$r$), eps = 0.001 is the precision required for estimating time to convergence.

**Table 4b**. Raftery and Lewis test for the cases of Berlin and Vienna.

| | $q$ | Berlin temperature (1756-2009) Burn-in | Total | Lower bound | Dependence factor | Vienna temperature (1775-2009) Burn-in | Total | Lower bound | Dependence factor |
|---|---|---|---|---|---|---|---|---|---|
| $\varphi_1$ | 0.025 | 21 | 31416 | 3746 | 8.39 | 21 | 31612 | 3746 | 8.44 |
| | 0.500 | 24 | 356512 | 38415 | 9.28 | 21 | 322441 | 38415 | 8.39 |
| | 0.975 | 21 | 31731 | 3746 | 8.47 | 21 | 31745 | 3746 | 8.47 |
| $H$ | 0.025 | 18 | 27288 | 3746 | 7.28 | 18 | 35670 | 4899 | 7.28 |
| | 0.500 | 21 | 322777 | 38415 | 8.4 | 21 | 422975 | 50239 | 8.42 |
| | 0.975 | 28 | 32732 | 3746 | 8.74 | 28 | 42882 | 4899 | 8.75 |
| | | Berlin temperature (1756-1919) | | | | Vienna temperature (1775-1919) | | | |
| $\varphi_1$ | 0.025 | 21 | 31780 | 3746 | 8.48 | 21 | 31780 | 3746 | 8.48 |
| | 0.500 | 24 | 356656 | 38415 | 9.28 | 21 | 323631 | 38415 | 8.42 |
| | 0.975 | 21 | 32193 | 3746 | 8.59 | 21 | 32137 | 3746 | 8.58 |
| $H$ | 0.025 | 18 | 27330 | 3746 | 7.3 | 18 | 27072 | 3746 | 7.23 |
| | 0.500 | 21 | 323330 | 38415 | 8.42 | 21 | 324177 | 38415 | 8.44 |
| | 0.975 | 18 | 32991 | 3746 | 8.81 | 27 | 39690 | 3746 | 10.6 |

Note: $q$ is the quantile to be estimated, $r$ = 0.005 is the desired margin of error of the estimate, $s$ = 0.95 the probability of obtaining an estimate in the interval ($q$−$r$, $q$+$r$), eps = 0.001 is the precision required for estimating time to convergence.

**Table 5a**. Summary results for the parameters of the AR(1) and HK cases at Boeoticos Kephisos River basin.

| Case | Mean | Standard Deviation | Quantiles 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|
| Boeoticos runoff | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 197.7 | 12.69 | 172.5 | 189.4 | 197.7 | 205.9 | 222.8 |
| $\sigma$ | 83.93 | 7.41 | 71.50 | 78.78 | 83.23 | 88.29 | 100.45 |
| $\varphi_1$ | 0.35 | 0.10 | 0.16 | 0.28 | 0.35 | 0.42 | 0.55 |
| HK | | | | | | | |
| $\mu$ | 194.85 | 31.30 | 132 | 178.1 | 195 | 211.6 | 256.1 |
| $\sigma$ | 86.51 | 12.35 | 71.19 | 79.15 | 84.40 | 91.06 | 114.22 |
| $H$ | 0.74 | 0.07 | 0.62 | 0.69 | 0.74 | 0.78 | 0.88 |
| Aliartos rainfall | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 658.18 | 18.57 | 621.5 | 646 | 658.2 | 670.4 | 694.7 |
| $\sigma$ | 159.9 | 12.24 | 138.3 | 151.3 | 159.1 | 167.5 | 186.2 |
| $\varphi_1$ | 0.11 | 0.10 | −0.09 | 0.04 | 0.11 | 0.18 | 0.32 |
| HK | | | | | | | |
| $\mu$ | 657.09 | 31.98 | 592.5 | 638.4 | 657.3 | 676.1 | 720.4 |
| $\sigma$ | 160.7 | 13.45 | 137.9 | 151.4 | 159.5 | 168.6 | 190.3 |
| $H$ | 0.62 | 0.06 | 0.51 | 0.58 | 0.62 | 0.66 | 0.75 |
| Aliartos temperature | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 16.96 | 0.10 | 16.76 | 16.89 | 16.96 | 17.02 | 17.15 |
| $\sigma$ | 0.71 | 0.06 | 0.61 | 0.67 | 0.70 | 0.74 | 0.84 |
| $\varphi_1$ | 0.33 | 0.10 | 0.14 | 0.26 | 0.33 | 0.39 | 0.52 |
| HK | | | | | | | |
| $\mu$ | 16.97 | 0.29 | 16.44 | 16.83 | 16.97 | 17.11 | 17.52 |
| $\sigma$ | 0.75 | 0.13 | 0.62 | 0.68 | 0.73 | 0.79 | 0.99 |
| $H$ | 0.74 | 0.07 | 0.61 | 0.69 | 0.74 | 0.79 | 0.88 |

**Table 5b**. Summary results for the parameters of the AR(1) and HK cases respectively at Berlin and Vienna.

| Case | Mean | Standard Deviation | Quantiles 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|
| Berlin temperature (1756-2009) | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 9.18 | 0.09 | 9.01 | 9.12 | 9.18 | 9.24 | 9.35 |
| $\sigma$ | 0.93 | 0.05 | 0.84 | 0.89 | 0.92 | 0.96 | 1.03 |
| $\varphi_1$ | 0.38 | 0.06 | 0.26 | 0.34 | 0.38 | 0.42 | 0.49 |
| HK | | | | | | | |
| $\mu$ | 9.28 | 0.25 | 8.80 | 9.13 | 9.27 | 9.43 | 9.79 |
| $\sigma$ | 0.94 | 0.06 | 0.83 | 0.89 | 0.93 | 0.97 | 1.08 |
| $H$ | 0.75 | 0.03 | 0.67 | 0.72 | 0.75 | 0.77 | 0.83 |
| Vienna temperature (1775-2009) | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 9.58 | 0.08 | 9.42 | 9.53 | 9.58 | 9.63 | 9.74 |
| $\sigma$ | 0.88 | 0.05 | 0.80 | 0.85 | 0.88 | 0.91 | 0.98 |
| $\varphi_1$ | 0.31 | 0.06 | 0.19 | 0.27 | 0.31 | 0.35 | 0.43 |
| HK | | | | | | | |
| $\mu$ | 9.64 | 0.19 | 9.27 | 9.52 | 9.64 | 9.76 | 10.03 |
| $\sigma$ | 0.88 | 0.05 | 0.79 | 0.84 | 0.87 | 0.91 | 0.99 |
| $H$ | 0.71 | 0.04 | 0.64 | 0.68 | 0.71 | 0.73 | 0.79 |
| Berlin temperature (1756-1919) | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 9.05 | 0.10 | 8.85 | 8.98 | 9.05 | 9.12 | 9.25 |
| $\sigma$ | 0.94 | 0.06 | 0.83 | 0.89 | 0.93 | 0.97 | 1.06 |
| $\varphi_1$ | 0.31 | 0.08 | 0.16 | 0.26 | 0.31 | 0.37 | 0.46 |
| HK | | | | | | | |
| $\mu$ | 9.11 | 0.26 | 8.60 | 8.95 | 9.10 | 9.26 | 9.64 |
| $\sigma$ | 0.96 | 0.08 | 0.83 | 0.90 | 0.95 | 1.00 | 1.14 |
| $H$ | 0.72 | 0.05 | 0.63 | 0.69 | 0.72 | 0.76 | 0.83 |
| Vienna temperature (1775-1919) | | | | | | | |
| AR(1) | | | | | | | |
| $\mu$ | 9.36 | 0.08 | 9.20 | 9.31 | 9.36 | 9.42 | 9.52 |
| $\sigma$ | 0.86 | 0.05 | 0.76 | 0.82 | 0.85 | 0.89 | 0.97 |
| $\varphi_1$ | 0.12 | 0.08 | −0.04 | 0.06 | 0.12 | 0.18 | 0.29 |
| HK | | | | | | | |
| $\mu$ | 9.37 | 0.13 | 9.10 | 9.29 | 9.37 | 9.45 | 9.63 |
| $\sigma$ | 0.86 | 0.06 | 0.76 | 0.82 | 0.86 | 0.89 | 0.98 |
| $H$ | 0.61 | 0.05 | 0.51 | 0.57 | 0.61 | 0.64 | 0.72 |

**Table 6**. Estimates of $\mu$ using various methods.

| | Maximum likelihood estimate | | | 50% quantile | |
|---|---|---|---|---|---|
| Examined case | WN | AR(1) | HKp | AR(1) | HK |
| Boeoticos runoff | 197.63 | 197.65 | 195.11 | 197.7 | 195 |
| Aliartos rainfall | 658.36 | 658.22 | 657.38 | 658.2 | 657.3 |
| Aliartos temperature | 16.96 | 16.96 | 16.97 | 16.96 | 16.97 |
| Berlin temperature (1756-2009) | 9.17 | 9.18 | 9.27 | 9.18 | 9.28 |
| Vienna temperature (1775-2009) | 9.58 | 9.58 | 9.64 | 9.58 | 9.64 |
| Berlin temperature (1756-1919) | 9.04 | 9.05 | 9.10 | 9.05 | 9.11 |
| Vienna temperature (1775-1919) | 9.36 | 9.36 | 9.37 | 9.36 | 9.37 |


**Figure 1**. The Boeoticos Kephisos River basin.

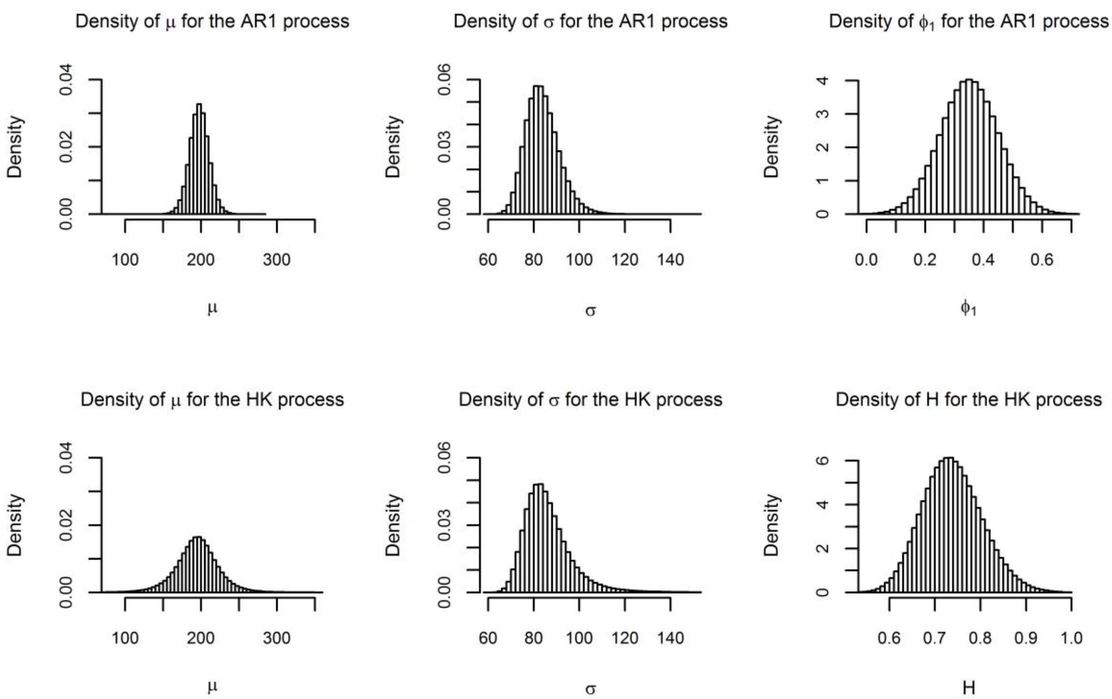


**Figure 2a**. Posterior probability distributions of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the cases of AR(1) and HK processes, for the runoff of Boeoticos Kephisos.

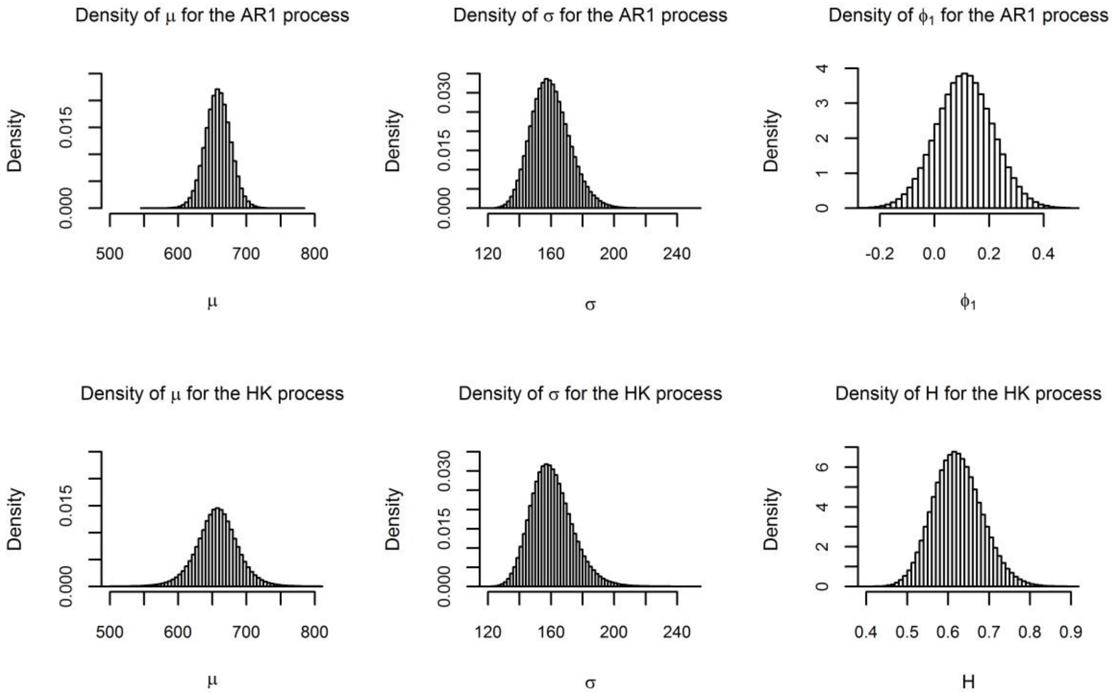


**Figure 2b**. Posterior probability distributions of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the cases of AR(1) and HK processes, for the rainfall at Aliartos.

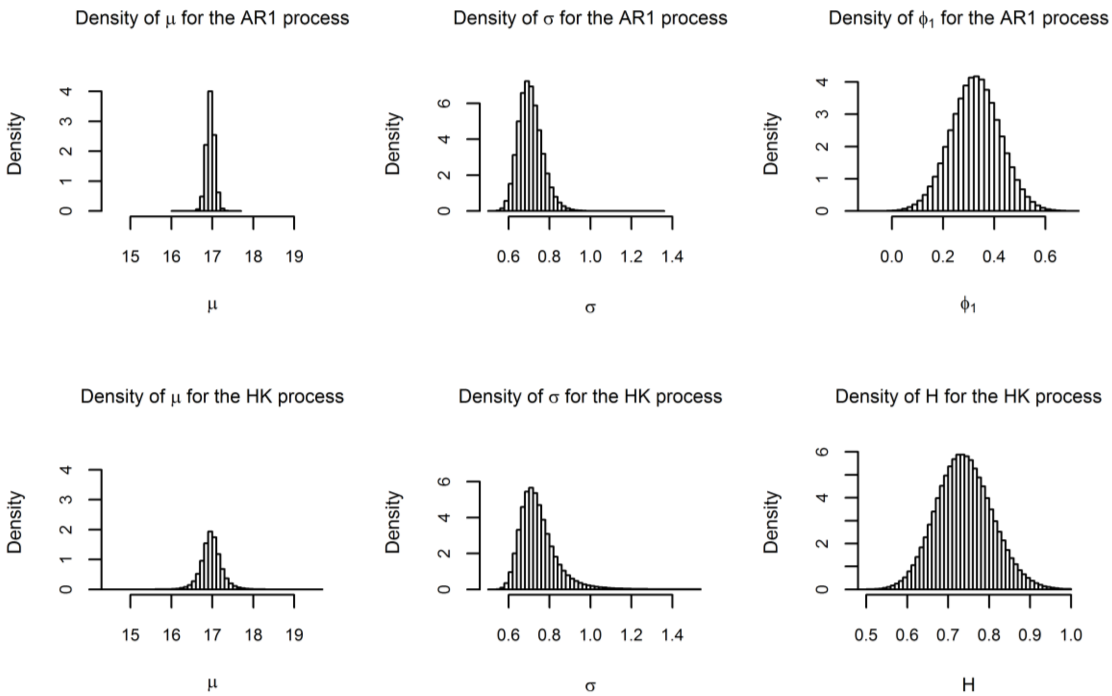


**Figure 2c**. Posterior probability distributions of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the cases of AR(1) and HK processes, for the temperature at Aliartos.

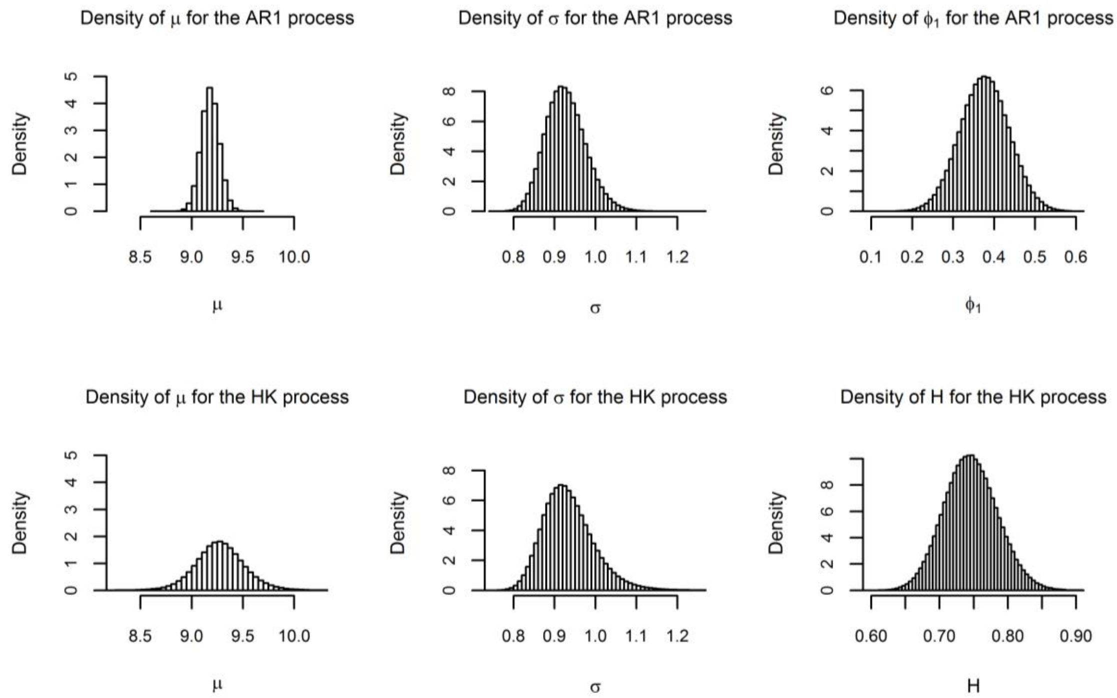


**Figure 3a**. Posterior probability distributions of $\mu$, $\sigma$, $H$, $\varphi_1$ for the cases of AR(1) and HK processes, for the temperature at Berlin/Tempelhof. In this case the parameters are estimated from years 1756-2009.

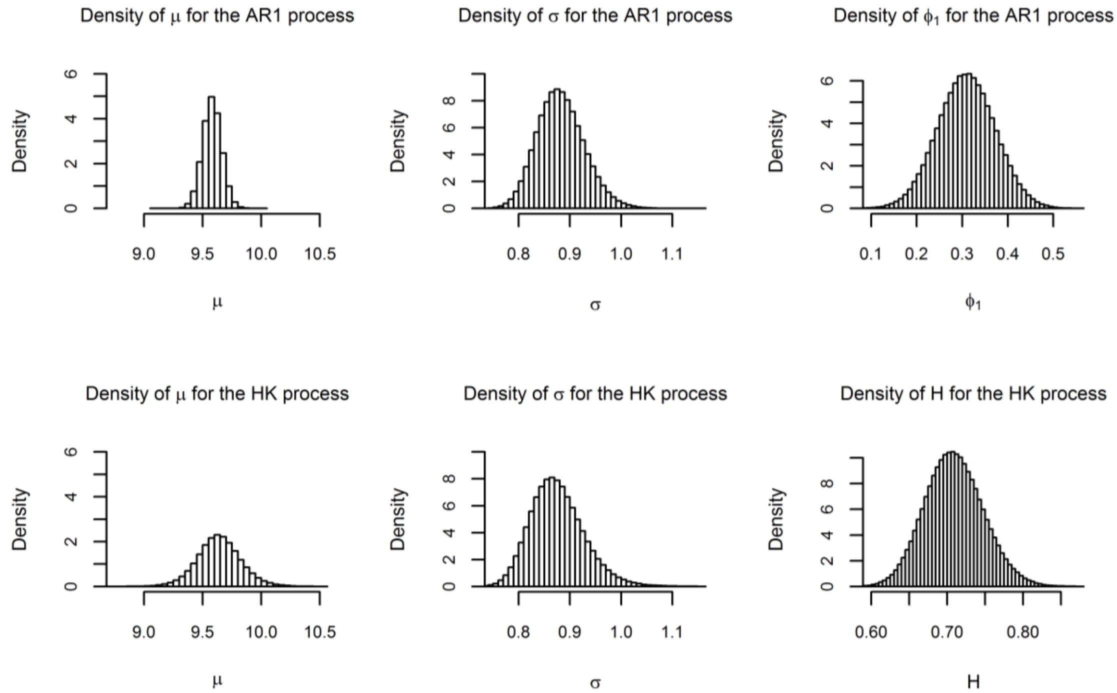


**Figure 3b**. Posterior probability distributions of $\mu$, $\sigma$, $H$, $\varphi_1$ for the cases of AR(1) and HK processes, for the temperature at Vienna. In this case the parameters are estimated from years 1775-2009.

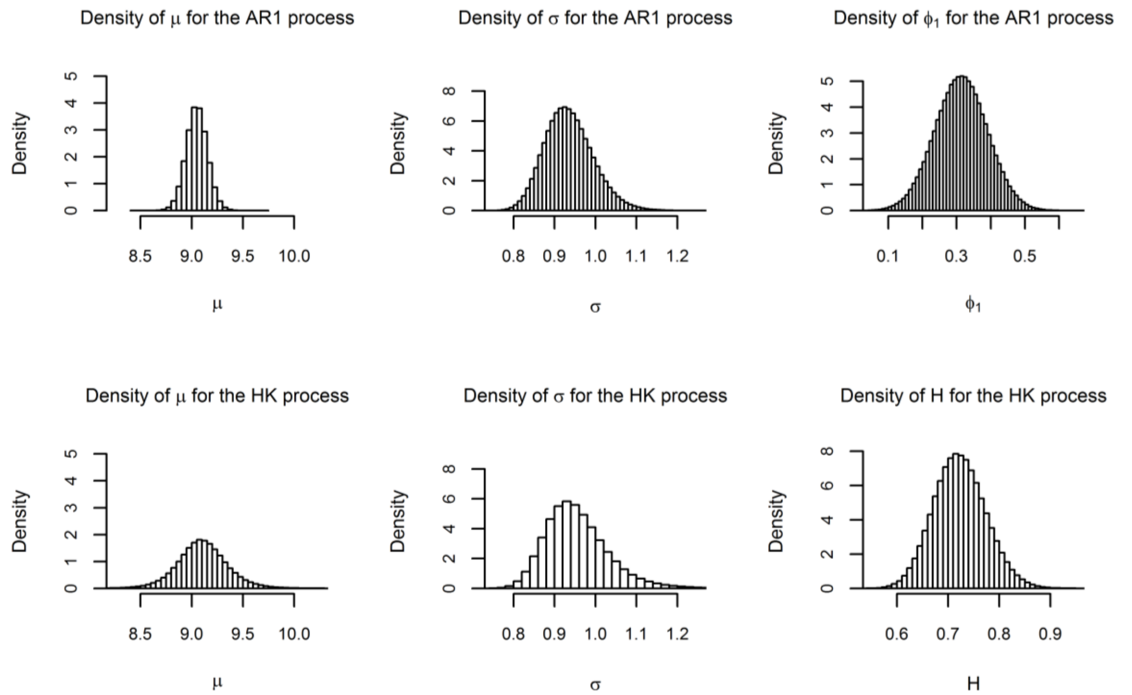


**Figure 3c**. Posterior probability distributions of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the cases of AR(1) and HK processes, for the temperature at Berlin/Tempelhof. In this case the parameters are estimated from years 1756-1919.

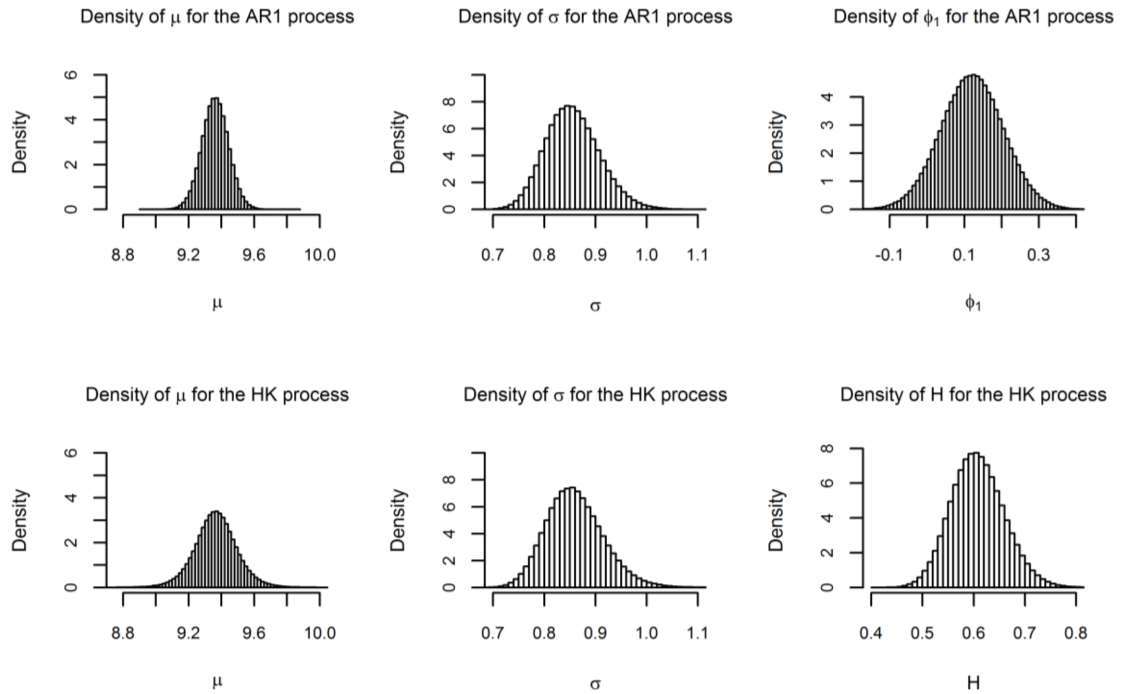


**Figure 3d**. Posterior probability distributions of $\underline{\mu}$, $\underline{\sigma}$, $\underline{H}$, $\underline{\varphi}_1$ for the cases of AR(1) and HK processes, for the temperature at Vienna. In this case the parameters are estimated from years 1775-1919.

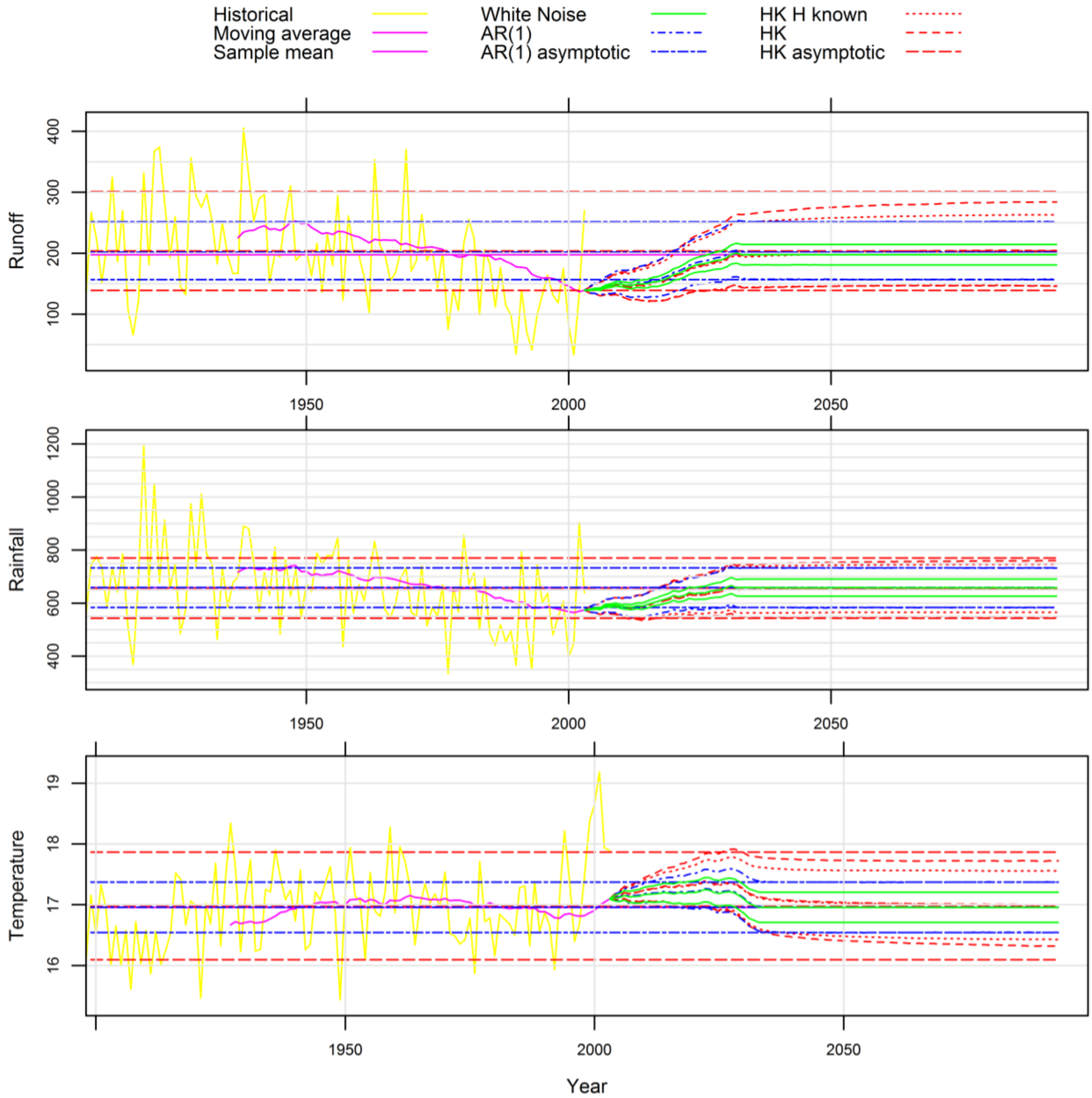


**Figure 4**. Historical climate and confidence regions of future climate (for $1 - a = 0.95$ and climatic time scale of 30 years) for (upper) runoff of Boeoticos Kephisos, (middle) rainfall at Aliartos, and (lower) temperature at Aliartos.

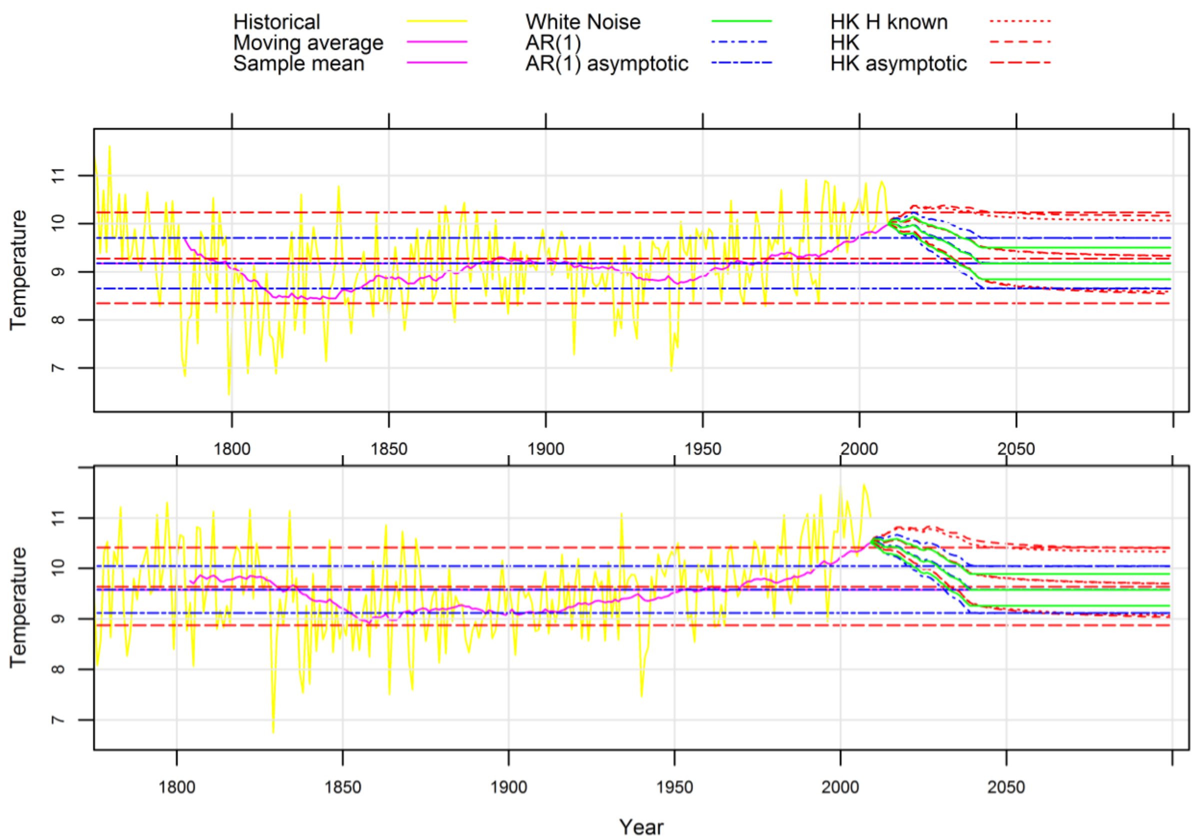


**Figure 5a**. Historical climate and confidence regions of future climate (for $1 - a = 0.95$ and climatic time scale of 30 years) for (upper) temperature at Berlin, and (lower) temperature at Vienna.

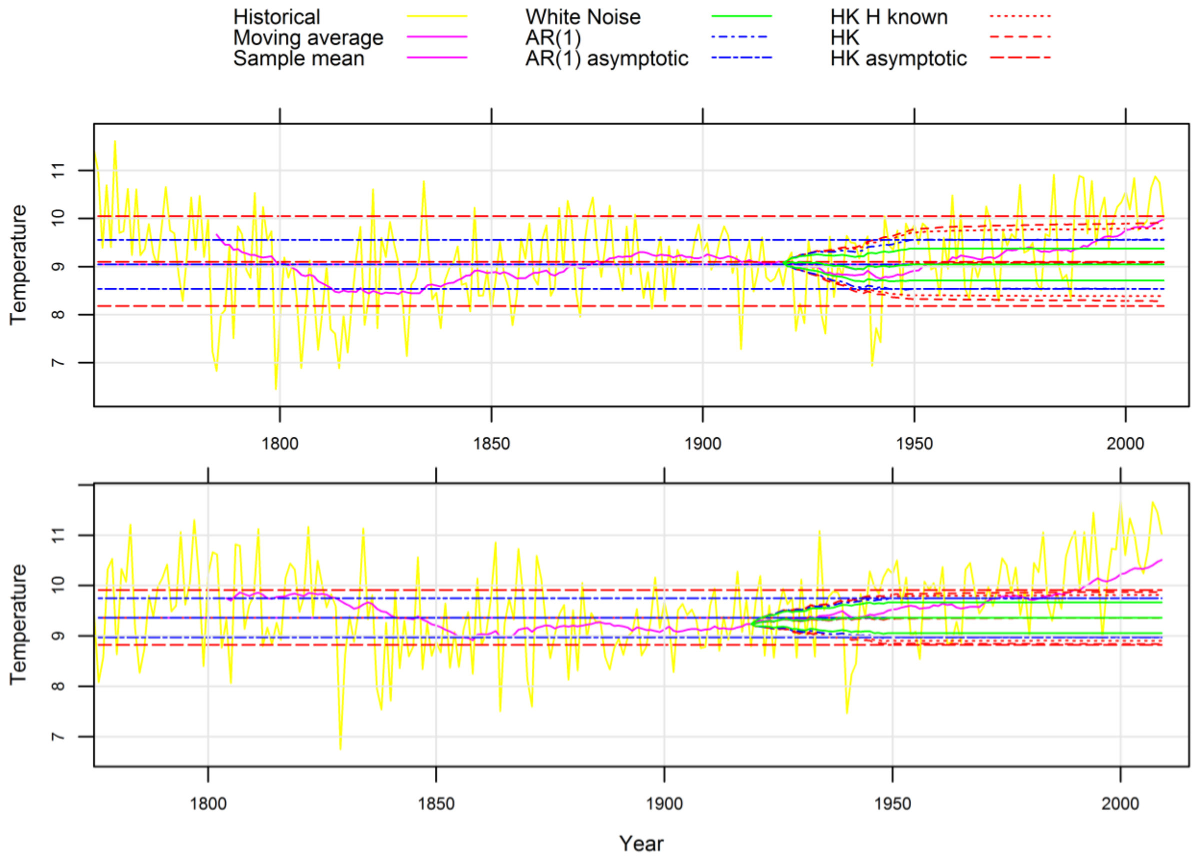


**Figure 5b**. Historical climate and confidence regions of climate (for $1 - a = 0.95$ and climatic time scale of 30 years) for (upper) temperature at Berlin/Tempelhof after the year 1920 and (lower) temperature at Vienna after the year 1920.