

Functional Time Series Analysis of Land Surface Temperature

Karim Azumah¹, Ananda Omutokoh Kube², Bashiru Imoro Ibn Saeed³

¹ Pan African University Institute for Basic Sciences, Technology and Innovation, Juja, Kenya

² Department of Mathematics, Kenyatta University, Kenya

³ Department of Mathematics and Statistics, Kumasi Technical University, Kumasi, Ghana

Correspondence: Karim Azumah, Pan African University Institute for Basic Sciences, Technology and Innovation, Juja, Kenya

Received: May 24, 2020 Accepted: August 17, 2020 Online Published: August 24, 2020

doi:10.5539/ijsp.v9n5p61 URL: <https://doi.org/10.5539/ijsp.v9n5p61>

Abstract

Parametric modeling imposes rigid assumption on abstraction of physical characteristics of a phenomenon, which in case of model misspecification could give erroneous results. To address the drawbacks, efforts have been channeled on semi-parametric and nonparametric modeling and inference. This study focuses on constructing an estimator and consequently modeling a meteorological temperature time series first by constructing a penalized spline estimator based on cubic splines. The penalized spline estimator proposed, which are known to impose very minimal restrictions on estimation process, provides good fits to observed data with very attractive properties namely consistent as observed in values of the Mean Squared Error from the analysis. The results of our simulations compared favorably with the empirical analysis on average monthly meteorological temperature data obtained from Climate Knowledge Portal World Bank Organization on Ghana for periods 1901-2016.

Keywords: splines, smoothing parameter, cubic spline functions, temperature, roughness penalty

1. Introduction

A major fundamental problem of all types of meteorological time series variables is their nonlinear features see e.g. (Wang et al., 2013; Bradley and Kantz, 2015; Hyrkkänen et al., 2016; Krishnamurthy, 2019; Benrhmach et al., 2020; Md. Karimuzzaman & Md.Hossain, 2020). Approaching from the data generating processes and its inherent physical conditions. Nonlinearity can cause severe methodological inaccuracies in modeling meteorological times series variables if not taken into account. As a corrective or remedy, the numerical technique of Spline functions (B-spline interpolation) has been modified to the field of statistical time series modeling, see e.g. (Eilers and Marx, 2010; Grimstad and Sandnes, 2016; Jauch et al., 2017; Singh et al., 2020; Yeh et al., 2020; Michel and Zidna, 2020). An unlimited advantage of this method is the ability to correct nonlinearity in the meteorological time series variable. As cited in (Tamsir et al., 2016; Bluemm et al., 2010; Remontet et al., 2019) with a simple extension of this technique, the elimination of systematic influences, such as the effect of nonstationarity coming from the seasonality and cyclical, is it possible to model temperature variable to obtain a high degree efficient temperature model. The study proposed to derive a penalized spline estimator and study its mean squared error measure in relations to selection of smoothing parameter with illustration to real meteorological temperature variable.

A study by Mitchell (2001), indicated that, global climate change has already had observable effects on the environment such as flooding, earth quick, drought, extreme rainfall etc. The effects that scientists had predicted in the past would result from global climate change are now occurring. Thus, loss of sea ice, accelerated sea level rise and longer, more intense heat waves. Scientists have predicted that global temperatures will continue to rise for decades to come, as a result of greenhouse gases produced by human activities. Again, the extent of climate change effects on individual regions will vary over time and therefore, societies needs conscious effort to mitigate or adapt to changes, (IPCC, 2001).

A key to effective managerial decision-making will be forecasts of climate change that are accurate and cost effective. An important indicator of climate change, although by no means the only one, is global temperature change, (Romilly, 2005). As Wangler and Backlund (2005), suggested, systems are defined by a set of entities, named as components, mutually connected where those connection are responsible for defining the various relationships and dependencies among all components. The relationships and the dependencies are often unknown. Therefore, the knowledge of the components and understanding their connections according to (Robert H. Shumway, 2006), one can model the system in order to improve decision making, forecasting, controlling as well as detecting outliers or failure in the system.

More so, the experimental data study area aims at designing methods to formulate or model and thus understand system behavior, (Robert H. Shumway, 2006). Also, global temperature is a popular metric for describing the state of global

climate. However, its effects are felt locally, but the global distribution of climate response to many global climate changes is reasonably congruent in climate models, suggesting that the global metric measure is useful, (Hansen et al., 2006). By IPCC (2007) the Intergovernmental Panel on Climate Change (IPCC), has forecasted a temperature rise of 2.5 to 10 degrees Fahrenheit over the next century.

According to Godfrey et al. (2012) climate change is one of the extreme significant problems on the worldwide political and economic agendas. The complexities, confusing, and at times contested scientific information resulted in a slow public and political response to the climate crisis.

In the African context, climate change is far from intangible, it is presently defining the course of people's lives. Africa experienced extreme weather events and more irregularity in weather patterns, leading to serious outcomes for the people, who depend on land and some water bodies to survive. As a result, Africa's engagement with the issue is evolving rapidly, presenting an opportunity to leapfrog the slow evolution of western public opinion and political action, (Godfrey et al., 2012).

According to Adedeji et al. (2014), climate change is one of the major challenges of our time and adds considerable stress to our societies and to the environment. This is seen from shifting weather patterns that threaten the survival of the likelihood of mankind from food production, rising sea levels increasing the risk of catastrophic flooding, this impacts of climate change are global in scope and unprecedented in scale. This requires a drastic action today, adapting to these impacts in the present and future will be more difficult and costly.

A report by DeGhetto et al. (2016) on African Agenda 2063 of the Africa we want, point out that, whilst Africa at present contributes less than 5% of global carbon emissions, it bears the brunt of the impact of climate change. Africa shall address the global challenge of climate change by prioritizing adaptation in all our actions, drawing upon skills of diverse disciplines to ensure implementation of actions for the survival of the most vulnerable populations, including islands states, and for sustainable development and shared prosperity.

Africa will participate in global efforts for climate change mitigation that support and broaden the policy space for sustainable development on the continent. Africa shall continue to speak with one voice and unity of purpose in advancing its position and interests on climate change, (DeGhetto et al., 2016).

Another study by, Amegah et al. (2016) on sub-Saharan Africa found that, sub-Saharan Africa augments very little to global climate change and nonetheless it is projected to endure the utmost problem of climate change, with 34% of the global Disability-Adjusted Life Years (DALYs) attributable to the effects of climate change found in sub-Saharan Africa.

As a result, Ye et al. (2013); Mudelsee (2018) modeled the absolute temperature using the traditional parametric methods to model the relationships and using the "trial and error" method to determine the order of the polynomial function for estimation process. These concerns are to be address by using appropriate spline functions.

The study sought to achieved the following objectives: to derive a penalized spline estimator and to study its mean squared error measure in relations to selection of smoothing parameter with illustration to real meteorological temperature variable; to study the gains of the models obtained here over other existing models; to assess the accuracy as well as the consequence of the results achieved.

2. Methods

2.1 Spline Models

A spline model is a piece-wise well-defined function with the separable pieces joined together employing continuity and smoothness constraints. They are worthwhile in explaining the relationship between a response variable and one or more independent variables when the relationship involves a curve or flexible model. The segments of a spline function are usually low order polynomials of up to third degree and the polynomial segments connect at a set of finite points known as knots. There are numerous forms of spline functions and in order to offer a definition, the ensuing representation is presented. Let Y and X be jointly distributed random variables such that

$$E(Y|X) = l(x) \quad (2.1)$$

where Y is the response variable and X is the explanatory variable. The goal is to estimate the function l based on observed data, (y_i, x_i) , where $i = 1, \dots, n$. Given that the observed data exhibits a relationship such that linear and quadratic functions would not be a good fit to the data and using a high-order polynomial does not essentially offer an improved fit to the data and pose challenges in understanding the coefficients of the large number of polynomial terms. By way of an alternative methodology, one can subdivide the domain $[a, b]$ of the function l into $k + 1$ equally spaced segment.

Definition 2.1. Let $a < t_1 < \dots < t_k < b$ be fixed points called knots. Let $t_0 = a$ and $t_{k+1} = b$. Generally, splines functions

are piecewise polynomials joined together smoothly at the knots. Formally, a *spline function* of order m , is a real-valued function on the closed interval $[a, b]$, such that

- (i) l is piecewise polynomial of order m on the $[t_i, t_{i+1}), i = 0, 1, \dots, k$
- (ii) l has $m - 2$ continuous derivatives and the $(m - 1)$ st derivative is a step function with jumps at the knots. For orders represented by $m = 2r$, the function l is a *natural spline function* of order $2r$ if, in addition to (i) and (ii), it satisfies the natural boundary conditions
- (iii) $l^{(j)}(a) = l^{(j)}(b) = 0, j = r, \dots, 2r - 1$.

The natural boundary conditions ensure that l is spline function of order r on the two outside intervals $[a, t_1]$ and $[t_k, b]$. Represent the function space of the natural spline function of order $2r$ with knots t_1, \dots, t_k as $NS^{2r}(t_1, \dots, t_k)$.

To obtain a good approximation of the natural spline function, we define smoothing spline approach where the number and location of knots is worth considering. We begin, by defining a well-defined model space for the function and introduce a penalty to account for overfitting.

Suppose l is "smooth". Specifically, assume that $l \in W_2^r[a, b]$ where the *Sobolev space*

$$W_2^r[a, b] = \left\{ l : l, l', \dots, l^{r-1} \text{ are absolutely continuous } \int_a^b (l^{(r)})^2 dx < \infty \right\} \tag{2.2}$$

Let $L(t)$ be a natural cubic interpolation spline. On each interval, $[t_i, t_{i+1}], i = 1, \dots, m - 1$, where, $l_i(t)$ is given by a different cubic polynomial, l_i , define as

$$L(t) = \sum_{i=1}^{m-1} 1_{[t_i, t_{i+1})}(t) l_i(t) \tag{2.3}$$

since $L(t)$ is natural cubic spline, the following must hold,

$$l_i(t) = a_i(t - t_i)^3 + b_i(t - t_i)^2 + c_i(t - t_i) + d_i \quad 1 \leq i \leq m - 1 \tag{2.4}$$

Using Truncated Power Basis to model natural cubic spline function as given in equation (2.3) is

$$L(t) = \sum_{j=0}^3 \beta_j t^j + \sum_{k=1}^K \theta_k (t - \eta_k)_+^3 \tag{2.5}$$

where η_k is the knots and K is number of knots of the data-set The truncated power basis has $r + l - 1$ basis function. Hence the truncated power basis function is indeed a basis of the vector space of the splines function. In general the truncated power basis of a function with degree r and K knots can be written as

$$L(t) = \sum_{j=0}^r \beta_j t^j + \sum_{k=1}^K \theta_k (t - \eta_k)_+^3 \tag{2.6}$$

where

$$(t - \eta_k)_+^3 = \begin{cases} (t - \eta_k)_+^3, & \text{if } (t - \eta_k)_+^3 > 0 \\ 0, & \text{if } (t - \eta_k)_+^3 \leq 0 \end{cases} \tag{2.7}$$

where K and η_k are the number of knots and knots position of the data set. The data represents average monthly meteorological temperature data collected from (<https://climateknowledgeportal.worldbank.org/download-data>) for the period between January,1901-December,2016 on Ghana.

Based on the data pairs $(Y_t, y_{t-1}), y_{t-1} \in [a, b], t = 1, \dots, n$ with a true relationship

$$Y_t = l(y_{t-1}) + \epsilon_i \tag{2.8}$$

we aim to estimate the unknown smooth function $l(\cdot) \in C^{p+1}([a, b])$, a $p + 1$ times continuously differentiable function, with penalized splines. The residuals ϵ_i are assumed to be uncorrelated with zero mean and variance $\sigma^2 > 0$.

Suppose that, $n \geq r = 2$ and $a < t_1 < t_2 < \dots < t_n \leq b$. Then, for fixed $0 < \lambda < \infty$ (2.9) has a unique minimizer \hat{l} and $\hat{l} \in NS^2(t_1, \dots, t_n)$, Eubank (1999). By this result indicates that even though we begun with the infinite dimensional space $W_2^r[a, b]$ as the model space for l , the solution is the minimizer of the equation (2.9) belonging to a finite dimensional space. Specifically, the solution is a natural spline function with knots at distinct design points. One approach to computing the polynomial spline function estimate is to denote \hat{l} as a linear combination of a basis of $NS^2(t_1, \dots, t_n)$.

2.2 The Penalized Least Squares Estimation Method

The function $l(t)$ is to be estimated using penalized cubic spline least squares method where generalized cross-validation (GCV) will be used to select smoothing parameter. Given any twice-differentiable function $l(t)$ defined on $[a, b]$, and a smoothing parameter $\lambda > 0$, define the penalized sum of squares as

$$S^2(\epsilon_t) = \sum_{i=1}^n \{Y(t_i) - l(t_i)\}^2 + \lambda \int_a^b \{l''(t)\}^2 dx \tag{2.9}$$

where $l(t_i)$ is defined in equation (2.4). The addition of the roughness penalty term $\lambda \int l''(t)^2$ in equation (2.9) ensures that the cost $S^2(\epsilon_t)$ of a particular curve is determined not only by its goodness-of-fit to the data as quantified by the residual sum of squares $\sum_{i=1}^n \{Y(t_i) - l(t_i)\}^2$ but also by its roughness $\int l''(t)^2$. The smoothing parameter λ represents the "rate of exchange" between residual error and local variation and gives the amount in terms of summed square residual error that corresponds to one unit of integrated squared second derivative.

In general, we want the resulting function to exhibit some degree of smoothness. The general approach to a formal generalization of this is to introduce the roughness measure,

$$P(l) = \int_{t_1}^{t_m} \{l''(t)\}^2 dt \tag{2.10}$$

which clearly measure the total curvature of the smoothing function. A fundamental results in spline theory is that the natural cubic spline (the cubic spline imposing, $l'(t_i) = l'(t_m) = 0$ ensures the smoothest fit by minimizing equation (2.10) among all C^2 .

Theorem 2.2. *With l denoting the natural cubic interpolation spline, we have for any interpolation function $f \in C^2$ such that $P(f) \geq P(l)$, where $P(l)$ is defined in equation (2.10), with equality if and only if $f = l$.*

Proof. Since $l \in C^2$ is a cubic spline, we have that,

$$l'''(t) = k_i, \forall t \in (t_i, t_{i+1}) \text{ for some } k_i \in \mathbb{R}, i = 1, \dots, m - 1 \tag{2.11}$$

$P(f) = 0$ if and only if $f''(t) = 0$, for all t , that is, if and only if f is a first order polynomial. For all, $x, y \in \mathbb{R}$, $x^2 - y^2 = (x - y)^2 + 2(x - y)y$, and let $\tau = f - l$. This implies that,

$$\begin{aligned} P(f) - P(l) &= \int_{t_1}^{t_m} [f''(t)]^2 dt - \int_{t_1}^{t_m} [l''(t)]^2 dt \\ &= \int_{t_1}^{t_m} [f''(t) - l''(t)]^2 + 2 \int_{t_1}^{t_m} [f''(t) - l''(t)] l''(t) dt \\ &= \int_{t_1}^{t_m} [\tau''(t)]^2 dt + 2 \int_{t_1}^{t_m} \tau''(t) l''(t) dt \\ &= P(\tau) + 2 \int_{t_1}^{t_m} \tau''(t) l''(t) dt \end{aligned}$$

Now,

$$P(\tau) = 0, \Leftrightarrow \tau'' = 0 \Leftrightarrow f'' = l''$$

so we have $P(\tau) \geq 0$, with equality if and only if $f'' = l''$. Considering second term, we see that,

$$\begin{aligned} 2 \int_{t_1}^{t_m} \tau''(t) l''(t) dt &= 2 \left([\tau'(t) l''(t)]_{t_1}^{t_m} - 2 \sum_{i=1}^{m-1} \int_{t_i}^{t_{i+1}} \tau'(t) l'''(t) dt \right) \\ &= 2 [\tau'(t) l''(t)]_{t_1}^{t_m} - 2 \sum_{i=1}^{m-1} k_i (\tau(t_{i+1}) - \tau(t_i)) \end{aligned}$$

using partial integration, the continuity of the derivatives up to and including second order, and (2.11). Recall the definition of τ ,

$$\tau(t_{i+1}) - \tau(t_i) = f(t_{i+1}) - l(t_{i+1}) - f(t_i) + l(t_i) = y_{i+1} - y_{i+1} - y_i + y_i = 0 \tag{2.12}$$

making the second term zero. Remaining with,

$$2 [\tau'(t)l''(t)]_{t_1}^{t_m} = 2 (\tau'(t_m)l''(t_m) - \tau'(t_1)l''(t_1)) \tag{2.13}$$

We note that by letting,

$$l''(t_1) = l''(t_m) = 0 \tag{2.14}$$

(or with $l' = f'$, in which case $l = f$), we have,

$$2 \int_{t_1}^{t_m} \tau''(t)l''(t)dt = 0 \tag{2.15}$$

This immediately allows us to conclude,

$$P(f) - P(l) = P(\tau) > 0 \tag{2.16}$$

with equality if and only if $l = f$ □

In other words, the function l is expressed as a linear combination of the basis functions, $l(t) = \sum_{j=1}^n N_j(t)\theta_j$, where

$$N_1(t) = 1,$$

$$N_2(t) = t,$$

N_1, \dots, N_n are the natural spline basis functions:

$$N_{k+2}(t) = \frac{(t - t_k)_+^3 - (t - t_n)_+^3}{t_n - t_k} - \frac{(t - t_{n-1})_+^3 - (t - t_n)_+^3}{t_n - t_{n-1}}, \quad k = 3, 4, \dots, n - 2.$$

Writing $\mathbf{N} \in \mathbb{R}^{n \times n}$ with $N_{ij} = N_j(t_i)$, and $\mathbf{\Omega}_N \in \mathbb{R}^{n \times n}$ with $\Omega_{ij} = \int N_i''(t)N_j''(t)dt$, the minimization problem over l now becomes a minimization problem over $\theta = (\theta_1, \dots, \theta_n)^T \in \mathbb{R}^n$:

$\hat{\theta}_\lambda = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} (y - \mathbf{N}\theta)^T (y - \mathbf{N}\theta) + \lambda \theta^T \mathbf{\Omega}_N \theta$. We can solve for θ , thus,

$$\begin{aligned} & (y^T - \theta^T \mathbf{N}^T) (y - \mathbf{N}\theta) + \lambda \theta^T \mathbf{\Omega}_N \theta \\ & y^T y - y^T \mathbf{N}\theta - \theta^T \mathbf{N}^T y + \theta^T \mathbf{N}^T \mathbf{N}\theta + \lambda \theta^T \mathbf{\Omega}_N \theta \\ & y^T y - 2\theta^T \mathbf{N}^T y + \theta^T \mathbf{N}^T \mathbf{N}\theta + \lambda \theta^T \mathbf{\Omega}_N \theta \\ & \frac{\partial}{\partial \theta} (y^T y - 2\theta^T \mathbf{N}^T y + \theta^T \mathbf{N}^T \mathbf{N}\theta + \lambda \theta^T \mathbf{\Omega}_N \theta) = 0 \\ & -2\mathbf{N}^T y + 2\mathbf{N}^T \mathbf{N}\theta + 2\lambda \mathbf{\Omega}_N \theta = 0 \\ & 2\mathbf{N}^T \mathbf{N}\theta + 2\lambda \mathbf{\Omega}_N \theta = 2\mathbf{N}^T y \\ & \mathbf{N}^T \mathbf{N}\theta + \lambda \mathbf{\Omega}_N \theta = \mathbf{N}^T y \\ & \theta (\mathbf{N}^T \mathbf{N} + \lambda \mathbf{\Omega}_N) = \mathbf{N}^T y \\ & \hat{\theta} = (\mathbf{N}^T \mathbf{N} + \lambda \mathbf{\Omega}_N)^{-1} \mathbf{N}^T y \end{aligned} \tag{2.17}$$

Since $l(t) = \sum_{j=1}^n N_j(t)\theta_j$, from equation (2.17)

$$\hat{l}(\lambda) = \mathbf{N} (\mathbf{N}^T \mathbf{N} + \lambda \mathbf{\Omega}_N)^{-1} \mathbf{N}^T y \tag{2.18}$$

From equation (2.18), we represent the smoother matrix $S_\lambda = \mathbf{N} (\mathbf{N}^T \mathbf{N} + \lambda \mathbf{\Omega}_N)^{-1} \mathbf{N}^T$. Hence $\hat{l}(\lambda) = S_\lambda y$. By the Reinsch (1967) form, for a smoothing spline is by manipulating the singular value decomposition (SVD) of $\mathbf{N} = \mathbf{U}\mathbf{Z}\mathbf{V}^T$, we can rewrite S_λ as

$$\begin{aligned} S_\lambda &= \mathbf{N} (\mathbf{N}^T \mathbf{N} + \lambda \mathbf{\Omega}_N)^{-1} \mathbf{N}^T \\ &= \mathbf{U}\mathbf{Z}\mathbf{V}^T \left((\mathbf{U}\mathbf{Z}\mathbf{V}^T)^T \mathbf{U}\mathbf{Z}\mathbf{V}^T + \lambda \mathbf{\Omega}_N \right)^{-1} (\mathbf{U}\mathbf{Z}\mathbf{V}^T)^T \\ &= \mathbf{U}\mathbf{Z}\mathbf{V}^T \left((\mathbf{V}\mathbf{Z}\mathbf{U}^T) \mathbf{U}\mathbf{Z}\mathbf{V}^T + \lambda \mathbf{\Omega}_N \right)^{-1} (\mathbf{V}\mathbf{Z}\mathbf{U}^T) \\ &= \mathbf{U}\mathbf{Z}\mathbf{V}^T (\mathbf{V}\mathbf{Z}\mathbf{U}^T \mathbf{U}\mathbf{Z}\mathbf{V}^T + \lambda \mathbf{\Omega}_N)^{-1} \mathbf{V}\mathbf{Z}\mathbf{U}^T \\ &= (\mathbf{U}\mathbf{Z}^{-1} \mathbf{V}^T \mathbf{V}\mathbf{Z}\mathbf{U}^T \mathbf{U}\mathbf{Z}\mathbf{V}^T \mathbf{V}\mathbf{Z}^{-1} \mathbf{U}^T + \lambda \mathbf{U}\mathbf{Z}^{-1} \mathbf{V}^T \mathbf{\Omega}_N \mathbf{V}\mathbf{Z}^{-1} \mathbf{U}^T)^{-1} \\ &= (\mathbf{I} + \lambda \mathbf{K})^{-1} \end{aligned} \tag{2.19}$$

where $\mathbf{K} = \mathbf{U}^T \mathbf{Z}^{-1} \mathbf{V}^T \mathbf{\Omega}_N \mathbf{V}\mathbf{Z}^{-1} \mathbf{U}$ and depends only on the t'_i s and nothing else.

2.3 Selecting the Smoothing Parameter and Model Diagnostics

In this section, we discuss various data driven criteria for selecting λ in fitting criterion (2) conditional on the value of K : Akaike’s information criterion corrected (AICc), Bayesian information criterion (BIC), Cross-Validation(CV) and Generalized Cross-Validation (GCV) criterion. Each of these data driven criteria provides an approach to select the value of λ conditional on the number of knots, and each criterion is a function of λ . For the AICc, BIC, CV, and GCV criteria, the value of λ that gives the minimum value of the criteria is taken to be a good value for the smoothing parameter conditional on the value of K . Each of these methods is dependent on the sum of squares error,

$$SSE(\lambda) = \sum_{i=1}^n (Y_i - \hat{l}(t_i, \lambda))^2 \tag{2.20}$$

Ruppert et al. (2003) The procedure explained by Ruppert et al. (2003) using a Demmler-Reinsch Orthogonalization to compute our fitted values for each smoothing parameter value is adopted in this study. This is accomplished using the identity

$$SSE(\lambda) = y^T y - 2y^T \hat{l}_\lambda + \tilde{l}^T \hat{l}_\lambda \tag{2.21}$$

The Demmler-Reinsch Orthogonalization is explained below Ruppert et al. (2003). Assumed that

$$\hat{l}_\lambda = N(N^T N + \lambda \Omega_N)^{-1} N^T y \tag{2.22}$$

- The Procedure

1. Obtain matrix R using the Cholesky decomposition: $N^T N = R^T R$.
2. Use Singular Value Decomposition(SVD) on the symmetric matrix $R^{-T} \Omega_N R^{-1}$ to obtain $U \text{diag}(s) U^T$.
3. Compute A and b as follows: $A \equiv N R^{-1} U$ and $b = A^T y$.
4. The fitted values, \hat{l}_λ , are $\hat{l}_\lambda = A \left(\frac{b}{\mathbf{1} + \lambda s} \right)$, where s is from the singular value decomposition.
5. The degrees of freedom, $df_{fit}(\lambda)$, is, $df_{fit}(\lambda) = \mathbf{1}^T \left(\frac{\mathbf{1}}{\mathbf{1} + \lambda s} \right)$

The cross-validation criterion is

$$\min_{\lambda} CV(\lambda) = n^{-1} \sum_{i=1}^n \{Y_i - \tilde{l}^{(-i)}(t_i; \lambda)\}^2 \tag{2.23}$$

The CV method is also known as the leave- one-out method. The value of λ that gives the minimum CV score is taken to be a good choice for the smoothing parameter. If a smoother matrix exists, then the CV formula Silverman (1985) may be expressed as

$$CV(\lambda) = \sum_{i=1}^n \left(\frac{Y - \hat{l}(t_i)}{1 - S_{ii}} \right)^2 \tag{2.24}$$

Where S_{ii} denotes the i^{th} diagonal element of S_λ , defined as

$$S_\lambda = N(N' N + \lambda \Omega_N)^{-1} N' \tag{2.25}$$

and $\hat{l}(t_i)$ is the spline smoother with λ . Akaike’s information criterion corrected, $AICc$, was introduced by Hurvich et al. (1998) because the commonly used Akaike’s information criterion Eubank (1999) may have a tendency to over-fit the curve estimate for small samples. The $AICc$ criterion may be expressed as

$$AICc(\lambda) = \ln(n^{-1} SSE(\lambda)) + \frac{2(df(\lambda) + 1)}{n - df(\lambda) - 2} \tag{2.26}$$

The BIC criterion Schwarz et al. (1978) may be defined as,

$$BIC(\lambda) = \ln(n^{-1} SSE(\lambda)) + \frac{df(\log n)df(\lambda)}{n} \tag{2.27}$$

where $df(\lambda)=tr(S_\lambda)$ which is similar to $AICc$ but penalizes a model fit with a larger $df(\lambda)$ more strongly than the $AICc$ for large n .

Developed by Craven and Wahba (1979), Generalized cross-validation (GCV) may be defined by

$$\min_{\lambda} GCV(\lambda) = n^{-1} \frac{\sum_{i=1}^n (Y_i - \hat{l}(t_i))^2}{\{1 - n^{-1}trS(\lambda)\}^2} \tag{2.28}$$

To compare the proposed model, the mean squared error, predictive squared error, Mallows’s C_p measures will be employed. Given that our model is defined as

$$y = \mu + \epsilon \tag{2.29}$$

we assume that $E(y) = \mu$. The vector $\mu = (\mu_1, \dots, \mu_n)$ is the spline function and evaluated at the design model $\mu_i = l(t_i)$ and the components of the error vector ϵ are normally with mean zero and $\sigma^2 < \infty$. Since $\hat{l}(\lambda) = S_\lambda y$, then its expectation is

$$E(\hat{l}(\lambda)) = E(S_\lambda y) = S_\lambda E(y) = S_\lambda \mu \tag{2.30}$$

2.4 Bias of the Estimator

Let b_λ be bias of the estimator, then

$$b_\lambda = E(\hat{l}(\lambda)) - \mu = S_\lambda E(y) - \mu = S_\lambda \mu - \mu = (S_\lambda - I)\mu \tag{2.31}$$

2.5 Variance of the Estimator

The variance of y is the same as ϵ , that is,

$$Var(y) = \sigma^2 \tag{2.32}$$

$$Var(\hat{l}(\lambda)) = Var(S_\lambda(y)) = S_\lambda^T Var(y) S_\lambda = \sigma^2 S_\lambda S_\lambda^T \tag{2.33}$$

Note that the variance of the residual vector is given as

$$\begin{aligned} Var(y - \hat{l}(\lambda)) &= Var(y - S_\lambda y) \\ &= Var[(I - S_\lambda)y] \\ &= \sigma^2 (I - S_\lambda)(I - S_\lambda)^T \\ &= \sigma^2 (I - S_\lambda - S_\lambda^T + S_\lambda S_\lambda^T) \end{aligned} \tag{2.34}$$

2.6 The Variance Estimate, $\hat{\sigma}^2$

To estimate the variance, σ^2 , in case of regression, we divide the residual sum of squares $\|y - \hat{l}(\lambda)\|^2 = \sum_{i=1}^n (y_i - \hat{l}(\lambda)_i)^2$ by the degrees of freedom for the error to obtain an unbiased estimator of σ^2 . We need the expectation of the residual sum of squares. For any random variable, W ,

$$Var(W) = E(W^2) - (E(W))^2$$

Thus

$$E(W^2) = Var(W) + (E(W))^2$$

Using this, with $W = y_i - S_\lambda y_i$

$$E(W) = b_{\lambda,i}, \text{ is it bias}$$

$$Var(W) = \sigma^2 (I - S_\lambda - S_\lambda^T + S_\lambda S_\lambda^T)$$

The sum of the diagonal elements of a matrix is called its trace. A trace of a matrix $\mathbf{A} = tr(\mathbf{A})$

$$\begin{aligned} E\{\|y - \hat{l}(\lambda)\|^2\} &= \sigma^2 tr(I - S_\lambda - S_\lambda^T + S_\lambda S_\lambda^T) + b_\lambda^T b_\lambda \\ &= \sigma^2 (n - 2tr(S_\lambda) + tr(S_\lambda S_\lambda^T)) + b_\lambda^T b_\lambda \end{aligned}$$

The estimate of σ^2 is to divide the MSE by the degrees of freedom given s

$$df_{err} = n - 2tr(S_\lambda) + tr(S_\lambda S_\lambda^T)$$

Then

$$\hat{\sigma}^2 = \frac{\|y - \hat{l}(\lambda)\|^2}{df_{err}}$$

is our biased estimate of σ^2 . Generally, $\hat{\sigma}^2$ over estimates σ^2 due to fact that its bias, $b_\lambda^T b_\lambda / df_{err}$ is necessarily non-negative.

3. Performance Criteria

3.1 Mean Squared Error

The mean Squared Error of the estimator, $\hat{l}_\lambda(t)$ is

$$MSE(\hat{l}_\lambda(t)) = E\left\{(\hat{l}_\lambda(t) - l(t))^2\right\} \tag{3.1}$$

Remark 1. MSE is variance plus the biased squared, but the bias is as

$$b_\lambda(t) = E(\hat{l}_\lambda(t)) - l(t)$$

Then the Mean squared error

$$\begin{aligned} MSE(\hat{l}_\lambda(t)) &= E\left\{(\hat{l}_\lambda(t) - l(t))^2\right\} \\ &= E\left\{[\hat{l}_\lambda(t) - E(\hat{l}_\lambda(t)) + b_\lambda(t)]^2\right\} \\ &= E\left\{[(\hat{l}_\lambda(t) - l(t) - b_\lambda(t)) + b_\lambda(t)]^2\right\} \\ &= E\left\{[(\hat{l}_\lambda(t) - l(t) - b_\lambda(t))^2 + 2b_\lambda(t)(\hat{l}_\lambda(t) - l(t) - b_\lambda(t)) + (b_\lambda(t))^2]\right\} \\ &= E\left\{(\hat{l}_\lambda(t) - l(t) - b_\lambda(t))^2\right\} + 2b_\lambda(t)E(\hat{l}_\lambda(t) - l(t) - b_\lambda(t)) + E(b_\lambda(t))^2 \end{aligned}$$

From Remark (1)

$$b_\lambda(t) + l(t) = E(\hat{l}_\lambda(t))$$

$$\begin{aligned} MSE(\hat{l}_\lambda(t)) &= E\left\{[\hat{l}_\lambda(t) - E(\hat{l}_\lambda(t))]^2\right\} + b_\lambda(t)^2 \\ &= Var(\hat{l}_\lambda(t)) + b_\lambda(t)^2 \end{aligned}$$

3.2 The Predictive Squared Error

The predictive Squared Error of the estimator, $\hat{l}_\lambda(t)$ is defined as

$$PSE(\lambda, t) = E\left\{(y^* - \hat{l}_\lambda(t))^2\right\}$$

where y^* is a new response value associated with the predictor, ("new" here means not of one of the "old" values used in the estimation of the $\hat{l}_\lambda(t)$). There is a direct relationship between the *PSE* and *MSE* of the estimator.

$$\begin{aligned} PSE(\lambda, t) &= E\left\{(y^* - \hat{l}_\lambda(t))^2\right\} \\ &= E\left\{(y^* - l(t) + l(t) - \hat{l}_\lambda(t))^2\right\} \\ &= E\left\{[(y^* - l(t)) + (l(t) - \hat{l}_\lambda(t))]^2\right\} \\ &= E\left\{(y^* - l(t))^2 + 2(y^* - l(t))(l(t) - \hat{l}_\lambda(t)) + (l(t) - \hat{l}_\lambda(t))^2\right\} \\ &= E\left\{(y^* - l(t))^2\right\} + 2E(y^* - l(t))(l(t) - \hat{l}_\lambda(t)) + E(l(t) - \hat{l}_\lambda(t))^2 \\ &= E\left\{(y^* - l(t))^2\right\} + 0 + E(l(t) - \hat{l}_\lambda(t))^2 \\ &= E\left\{(y^* - l(t))^2\right\} + E(l(t) - \hat{l}_\lambda(t))^2 \\ &= \sigma^2 + MSE(\lambda, t) \end{aligned}$$

To obtain a uniform criterion of performance, we averaged the *MSE* and *PSE* over the observed observations as

$$\begin{aligned}
 MSE(\lambda) &= \frac{1}{n} \sum_{i=1}^n MSE(\lambda, t_i) \\
 &= \frac{1}{n} \sum_{i=1}^n Var(\hat{l}_\lambda(t_i)) + \frac{1}{n} \sum_{i=1}^n b_\lambda(t_i)^2 \\
 &= \frac{tr(Var(\hat{l}_\lambda(t_i)))}{n} + \frac{b_\lambda^T b_\lambda}{n} \\
 &= \frac{\sigma^2 tr(S_\lambda S_\lambda^T)}{n} + \frac{b_\lambda^T b_\lambda}{n}
 \end{aligned}$$

and

$$\begin{aligned}
 PSE(\lambda) &= \sigma^2 + \frac{\sigma^2 tr(S_\lambda S_\lambda^T)}{n} + \frac{b_\lambda^T b_\lambda}{n} \\
 &= \sigma^2 \left(1 + \frac{tr(S_\lambda S_\lambda^T)}{n} \right) + \frac{b_\lambda^T b_\lambda}{n}
 \end{aligned}$$

3.3 Mallow's C_p

A very bad approximation of the *PSE* (λ) is the average squared residual

$$ASR(\lambda) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{l}(\lambda))^2 = \frac{\|y - \hat{y}_\lambda\|^2}{n}$$

the expectation is

$$\begin{aligned}
 E(ASR(\lambda)) &= \frac{\sigma^2 (n - 2tr(S_\lambda) + tr(S_\lambda S_\lambda^T)) + b_\lambda^T b_\lambda}{n} \\
 &= \sigma^2 \left(1 - \frac{2tr(S_\lambda)}{n} + \frac{tr(S_\lambda S_\lambda^T)}{n} \right) + \frac{b_\lambda^T b_\lambda}{n} \\
 &= PSE(\lambda) - \sigma^2
 \end{aligned} \tag{3.2}$$

Hence

$$C_p(\lambda) = ASE(\lambda) + \hat{\sigma}^2 \frac{2tr(S_\lambda)}{n} \tag{3.3}$$

is a sensible estimate of the predictive squared error.

3.4 Asymptotic Properties of Penalized Spline Estimator

In studying the theoretical properties of the penalized spline estimator as the minimizer of the (2.9), we considered the average mean squared error (AMSE) and the asymptotic bias and variance of the model. In addition, we also deliberate on the optimum choice of the smoothing parameter. Asymptotic properties of the penalized spline estimator are discussed under subsection 3.4, subsection 3.5 Average mean squared error of the estimator and Asymptotic of the variance and the bias in subsection 3.6.

See Zhou et al. (1998) for asymptotic properties of the regression spline estimator and we adopt the following assumptions.

Assumption 1. let $\delta = \max_{0 \leq j \leq K} (\kappa_{j+1} - \kappa_j)$. There exist a constant $M > 0$, such that $\delta / \min_{0 \leq j \leq K} (\kappa_{j+1} - \kappa_j) \leq M$ and $\delta = o(K^{-1})$.

Assumption 2. For deterministic designs points $y_t \in [a, b]$, $t = 1, \dots, n$, assume that there exists a distribution function Q with corresponding positive continuous design density ρ such that, with Q_n the empirical distribution of y_1, \dots, y_n $\sup_{y \in [a, b]} |Q_n(y) - Q(y)| = o(K^{-1})$.

Assumption 3. The number of knots $K = o(n)$

The asymptotic bias and variance for the regression splines are obtainable from Zhou et al. (1998) as

$$E[\hat{l}_{reg}(y_t)] - l(y_t) = b_a(y_t; p + 1) + o(\delta^{p+1})$$

$$Var[\hat{l}_{reg}(y_t)] = \frac{\sigma^2}{n} N(y_t) G^{-1} N^T(y_t) + o\{(n\delta)^{-1}\}$$

where

$$b_a(y_t; p + 1) = -\frac{l^{(p+1)}(y_t)}{(p + 1)!} \sum_{j=0}^K I_{[\kappa_j, \kappa_{j+1})}(y_t) (\kappa_{j+1} - \kappa_j)^{p+1} B_{p+1}\left(\frac{y_t - \kappa_j}{\kappa_{j+1} - \kappa_j}\right) \tag{3.4}$$

and

$$G = \int_a^b N^T(y_t) N(y_t) \rho_{y_t} dx \tag{3.5}$$

given that $B_{p+1}(\cdot)$ is the $(p + 1)$ th Bernoulli polynomial see Abramowitz and Stegun (1972)

3.5 Average Mean Squared Error of the Estimator

Agreeing to Demmler and Reinsch (1975), it is possible to express the average bias and variance in terms of the eigenvalues having been obtained from the singular value decomposition.

$$(N^t N)^{-1/2} D_q (N^t N)^{-1/2} = U \text{diag}(s) U^T \tag{3.6}$$

where U is the matrix of vectors and s is the vector of eigenvalues s_j . Representing $A = N(N^t N)^{-1/2} U$. Forming a semi-orthogonal matrix with $A^t A = [N(N^t N)^{-1/2} U]^t [N(N^t N)^{-1/2} U]$, simplified as $A^t A = I_{K+p+1}$ and $AA^t = N(N^t N)^{-1} N^t$. Where AA^t is express as $[N(N^t N)^{-1/2} U][N(N^t N)^{-1/2} U]^t$ Rewriting the penalized spline estimator, $\hat{l} = N(N^t N + \alpha D_q)^{-1} N^t Y$ as

$$\hat{l} = A \{I_n + \alpha \text{diag}(s)\}^{-1} A^t Y \tag{3.7}$$

$$= \{I_n + A \alpha \text{diag}(s) A^t\}^{-1} AA^t Y$$

$$= \{I_n + A \alpha \text{diag}(s) A^t\}^{-1} \hat{l}_{reg} \tag{3.8}$$

We obtain the *AMSE* as

$$AMSE(\hat{l}) = \frac{1}{2} E\{\hat{l} - l\}^t \{\hat{l} - l\}$$

$$= \frac{\sigma^2}{n} \sum_{j=1}^{K+p+1} \frac{1}{(1 + \alpha s_j)^2} + \frac{\alpha^2}{n} \sum_{j=1}^{K+p+1} \frac{s_j^2 b_j^2}{(1 + \alpha s_j)^2} + \frac{1}{n} l^t (I_n - AA^t) l \tag{3.9}$$

where $AA^t l = E(\hat{l}_{reg})$, $l = \{l(t_1) \dots l(t_n)\}^t$, $b = A^t l$ with components b_j and AA^t an Idempotent matrix, obtaining

$$AMSE(\hat{l}) = \frac{\sigma^2}{n} \sum_{j=1}^{K+p+1} \frac{1}{(1 + \alpha s_j)^2} + \frac{\alpha^2}{n} \sum_{j=1}^{K+p+1} \frac{s_j^2 b_j^2}{(1 + \alpha s_j)^2} + \frac{1}{n} \sum_{j=1}^n [E\{\hat{l}_{reg}(t_j)\} - l(t_j)]^2 \tag{3.10}$$

From *AMSE* equation (3.10) the first term is the average variance while the second term is the average squared bias (shrinkage) as a result of penalization and the third term is the average squared approximation bias.

Let define

$$K_q = (K + p + 1 - q)(\alpha c_1)^{1/(2q)} n^{-1/(2q)} \tag{3.11}$$

Theorem 3.1. Under Assumptions 1-3, we consider two asymptotic cases: (i) if $K_q < 1$ and $l(\cdot) \in C^{p+1}[a, b]$, then

$$AMSE(\hat{l}) = O\left(\frac{K}{n}\right) + O\left(\frac{\alpha^2}{n^2} K^{2q}\right) + O(K^{-2(p+1)})$$

where $K \sim C_1 n^{1/(2p+3)}$, with C_1 a constant, $\alpha = O(n^\gamma)$ with $\gamma \leq (p+2-q)/(2p+3)$, the penalized spline estimator attains the optimal rate of convergence for $l \in C^{p+1}[a, b]$ with $AMSE(\hat{l}) = O(n^{-(2p+2)/(2p+3)})$. (ii) if $K_q \geq 1$ and $l \in W^q[a, b]$

$$AMSE(\hat{l}) = O\left(\frac{n^{1/(2q-1)}}{\alpha^{1/(2q)}}\right) + O\left(\frac{\alpha}{n}\right) + O(K^{-2(q)})$$

and for $\alpha = O(n^{1/(2q+1)})$, such that $\alpha n^{2q-1} \rightarrow \infty$ and $K \sim C_2 n^v$ with $v \geq 1/(2q+1)$ and C_2 a constant, the penalized spline estimator attains the optimal rate of convergence for $l \in W^q[a, b]$ with $AMSE(\hat{l}) = O(n^{-2q/(1+2q)})$

Considering case (i) with $K_q < 1$ the result we obtained is similar to the regression splines. In this case, the AMSE is determined by the squared approximation bias and the mean asymptotic variance. Thus the smaller the smoothing parameter, α , the insignificant the shrinkage bias is, that is $\gamma < (p+2-q)/(2p+3)$ with the asymptotically optimal number of knots having same order as the regression splines, that is $K \sim C_1 n^{1/(2p+3)}$. We also notice in the case(ii) with $K_q \geq 1$ that this result obtained is similar to the smoothing spline. Where AMSE in this case is dominated by the squared shrinkage bias and average asymptotic variance. The mean squared approximation bias has the same order as that of the shrinkage bias for $K_q = 1$ and is negligible order for $K_q > 1$. The results suggest that convergence rate for the penalized spline estimator is much faster in the case (i) where $K_q < 1$ relying on the assumption that $q \leq p$.

3.6 Asymptotic of the Variance and the Bias

Theorem 3.2. Under Assumptions 1-3, we derive the Asymptotic of the Variance and the Bias under following two cases:
 (i) If $K_q < 1$ and $l(\cdot) \in C^{p+1}[a, b]$, then

$$Var\{\hat{l}(t)\} = \frac{\sigma^2}{n} N(t)(G + \alpha D_q/n)^{-1} G(G + \alpha D_q/n)^{-1} N^t(t) + o\{(n\delta)^{-1}\}$$

Giving that,

$$E\{\hat{l}(t)\} - l(t) = b_a(t; p+1) + b_\alpha(t) + o(\delta^{p+1}) + o(\alpha n^{-1} \delta^{-q})$$

(ii) If $K_q \geq 1$ and $l \in W^q[a, b]$

$$Var\{\hat{l}(t)\} = \frac{\sigma^2}{n} N(t)(G + \alpha D_q/n)^{-1} G(G + \alpha D_q/n)^{-1} N^t(t) + o\{(n^{-1}(\delta/n)^{-1/2q})\}$$

Where,

$$E\{\hat{l}(t)\} - l(t) = b_a(t; q) + b_\alpha(t) + o(\delta^q) + o\{(\alpha/n)^{1/2}\}$$

The shrinkage bias b_α is defined as $b_\alpha(t) = -\alpha n^{-1} N(t)(G + \alpha D_q/n)^{-1} D_q \beta$, where G and β are define as in Zhou et al. (1998).

Proof. See proof of of Theorem (3.1) and Theorem (3.2) in Claeskens et al. (2009) □

4. Simulation, Analysis, Results and Discussion

4.1 Simulation Studies

A climate data (indicator) was generated using simulation method to measure strength of our proposed modeling adequacy. We anticipate the climate or whether signal may have composed of different components, namely, trend, periodicity as well as stochastic patterns.

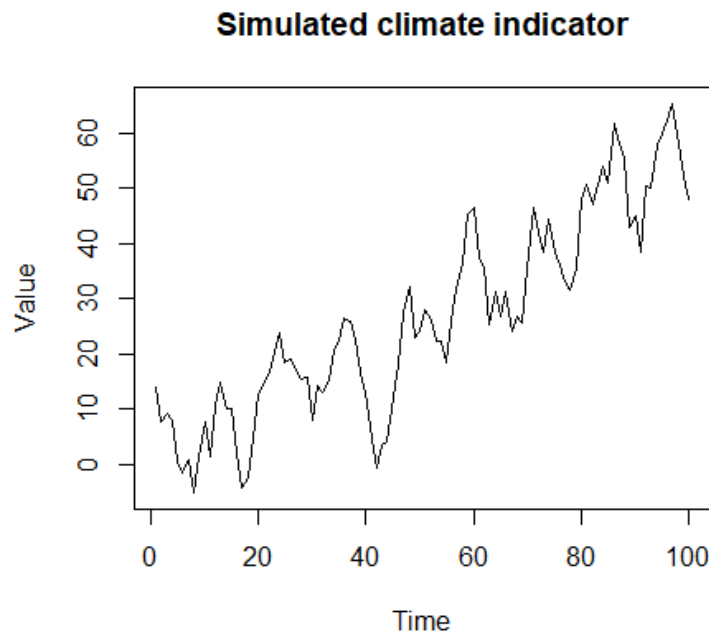
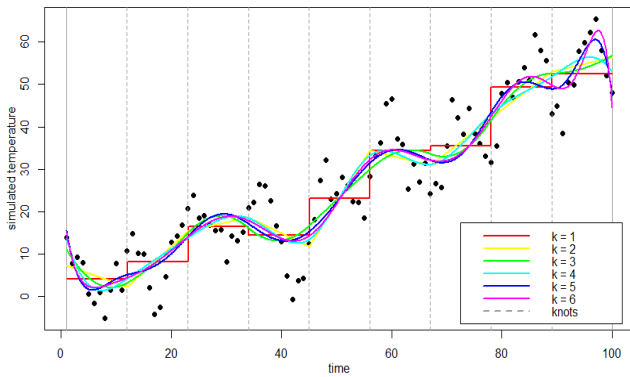


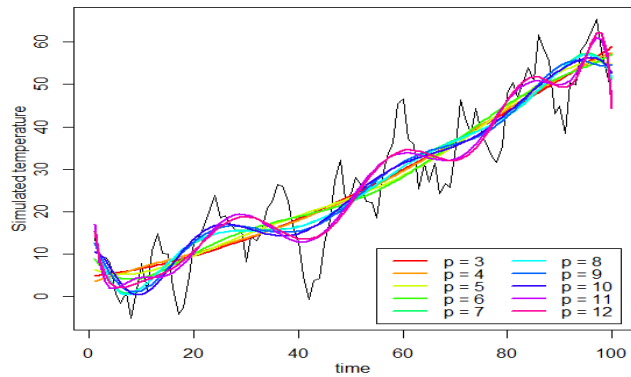
Figure 1. Time plot of the components of the climate indicator

The Figure1 show the time plot exhibiting an increasing trend and seasonal components. The highly oscillatory nature of simulated signal requires the use of nonlinear functions such as the splines functions to model it, since the traditional model might fail with mis-specified function form of the model.

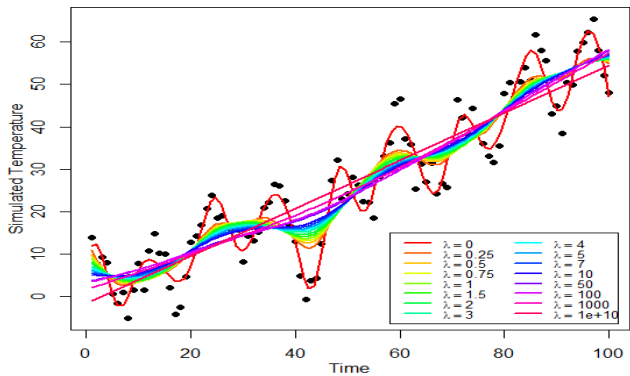
In Figure 2a, using a spline function to estimate the function for varying order k with a fixed number of $m = 8$ inner equidistant knot. For $k = 1$ the function consists of horizontal straight lines with jumps at the knots and for $k > 2$ indicates a smooth function but the orders cannot well be distinguished visually anymore. In contrast, Figure 2b shows the curves for polynomials of degree $p = 3, \dots, 12$. The curves are characterized with degrees of fluctuations and for higher values of p , the polynomial curves exhibit an oscillatory form. Also, in Figure 2c, the curve is estimated using Penalized splines with $m = 20$ equidistant inner knots and varying smoothing parameter values. Figure 2d shows the plot of selected smoothing parameter for simulated data for the penalized spline function. While Figure 2e represents the smoothing with $k = 4$ and varying values of λ . Figure 2f shows the plot of selected smoothing parameter for simulated data for the smoothing spline function.



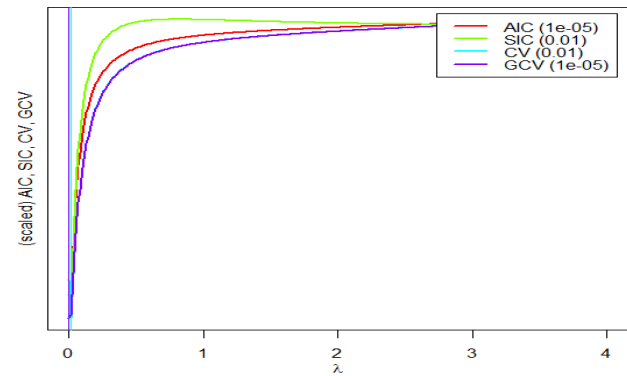
(a) Spline functions with 8 equidistant inner knots and different orders $k = 1, \dots, 6$



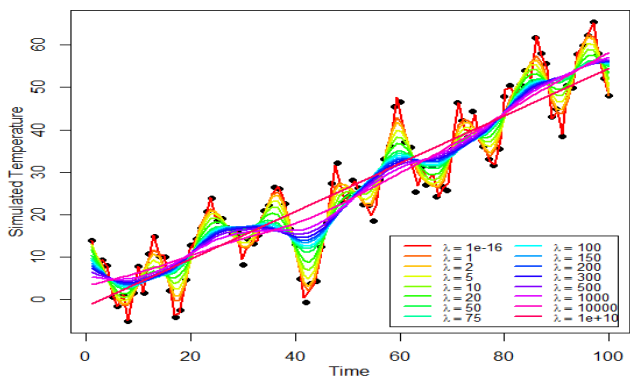
(b) A polynomial function fitting the simulated temperature with different orders $p = 3, \dots, 12$.



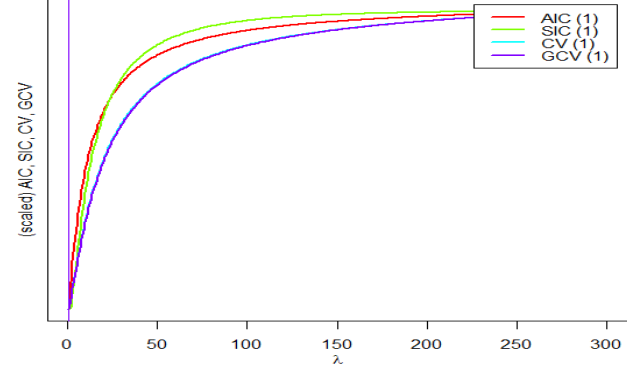
(c) Penalize spline functions of order $k = 4$ with 20 equidistant inner knots and varying values for λ .



(d) Smoothing parameter λ vs. resulting selection criteria for Penalized spline function of order $k = 4$ with 20 equidistant inner knots.



(e) Smoothing spline functions of order $k = 4$ with varying values for λ .



(f) Smoothing parameter λ vs. resulting selection criteria for smoothing spline function of order $k = 4$

Figure 2. Simulation of Temperature data

4.2 Application to Real data

This study used a secondary data on average monthly temperature from the World Bank Climate Change Knowledge Portal (CCKP), an online tool that provides access to comprehensive global and country data information related to climate change and development on temperature of Ghana for period 1901 to 2016. The open source R.3.5.2 version is employed in analyzing the data.

4.2.1 Results and Discussion

The Figure3 below comprises panels 3a-3h representing different spline function of different orders, number of knots and smoothing parameter λ . From Figure 3a shows the time series plots of the average monthly temperature exhibiting irregular oscillations and increasing trend. It is also, observed in Figure 3b that, using a spline function to estimate the function for varying order k with a fixed number of $m = 8$ inner knots in an equidistant knot sequence. For $k = 1$ the function consists of horizontal straight lines with jumps at the knots and for $k = 2$ the curve is constructed from straight lines with non-zero slopes which are continuously connected. For $k = 3$ and $k = 4$ the quadratic and cubic pieces can be observed but for $k > 5$ the orders cannot well be distinguished visually anymore. In contrast, Figure 3c shows the curves for polynomials of degree $p = 3, \dots, 12$. The curves is characterized with degree of fluctuations and for higher values of p , the polynomial curves exhibit an undesirable behavior by oscillating more and more especially at the end point. Also, in Figure 3d, the curve is estimated using Penalized splines with $m = 20$ equidistant inner knots and varying smoothing parameter values. Figure 3e shows the plot of selected smoothing parameter for data under considerations. With the selected smoothing parameters in Figure 3g, the B-spline, the penalized spline and smoothing splines were used to fit data in Figure 3h.

The performance criteria of fitted models are presented below based on MSE, PSE, and ASR taken into account the influence of the smoothing parameter, position and the number of knots as in the theoretical studies shown in section 3

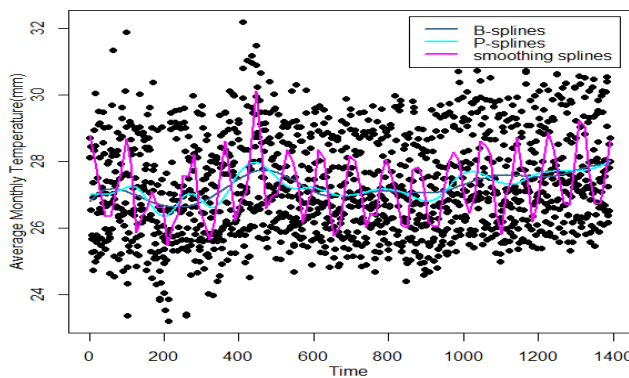
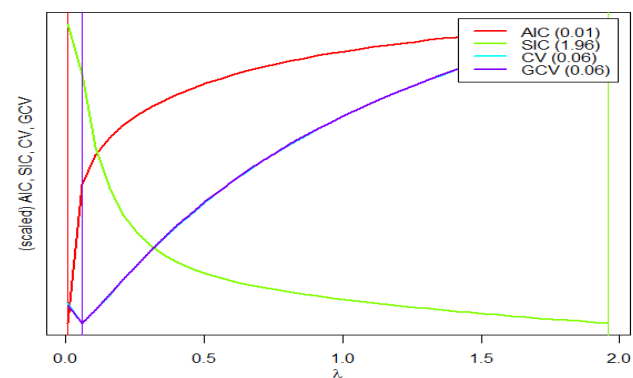
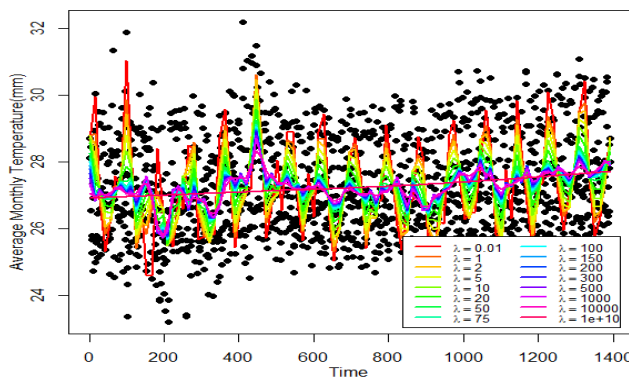
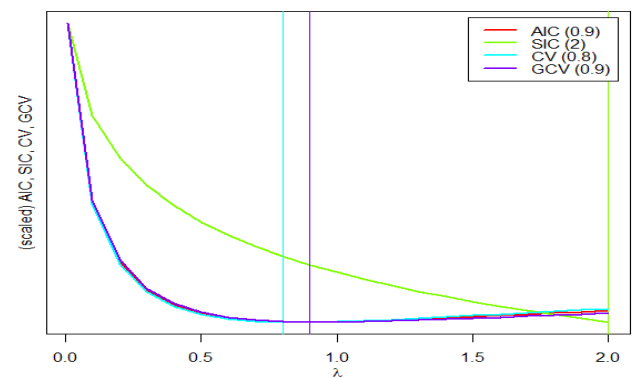
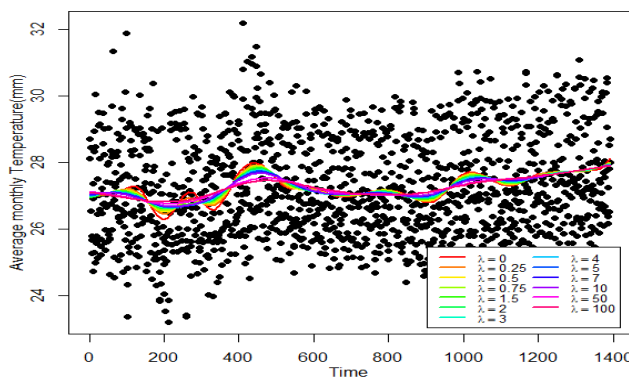
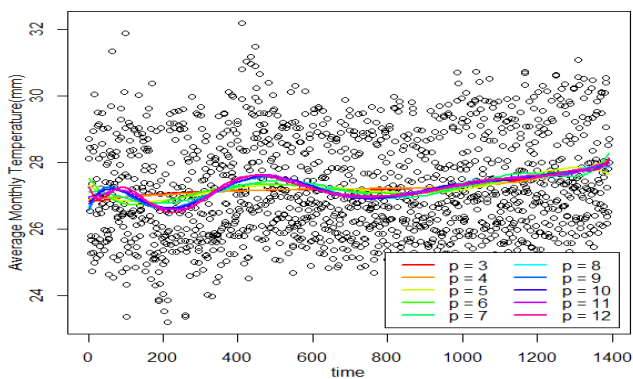
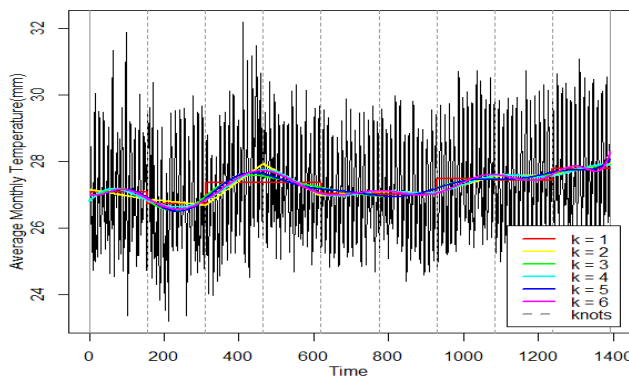
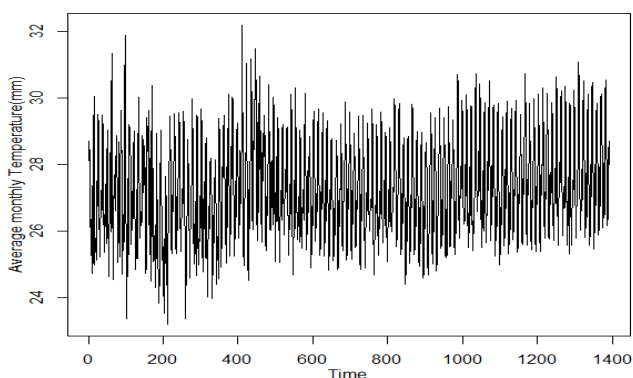


Figure 3. Fitting of the Average Monthly Temperature data

Table 1. Model performance measures

Function	Variance	MSE	PSE	ASR	Cp
B-spline ($k = 4, m = 8$, equidistant knots)	13.43152	2.254234	4.588284	2.234756	2.256641
P-spline ($k = 4, m = 20$, equidistant knots, $\lambda = 0.06$)	47.89018	2.215432	4.549482	2.180061	2.249155
Smoothing spline ($k = 4, \lambda = 6$)	544.045	1.454607	3.788657	0.712444	1.139812

From Table(1), the model performance measures for B-spline, Penalized spline and the Smoothing spline functions fitted to temperature data. The three spline functions perform fairly comparable, particularly the B-spline function and Penalized spline curves with suitably selected values for λ using the CV and GCV criteria. Largely, the specifications for the three curves seem all well chosen to represent the average monthly temperature for Ghana.

5. Conclusion

Considering the average monthly temperature data characterized by nonlinear the trend, seasonal and periodic components, the study proposed a penalized spline model. The study compared the B-spline, Penalized spline and the smoothing spline which in all cases with the appropriate determination of the smoothing parameter adequately fit the average monthly temperature for Ghana.

Acknowledgment

The authors express their appreciation to Pan African University Institute for Basic Sciences, Technology and Innovation and African Union for Supporting this research. Data information related to climate change and development on temperature of Ghana for period 1901 to 2016 was used and available on (<https://climateknowledgeportal.worldbank.org/downloaddata>)

Conflicts of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- Adedeji, O., Reuben, O., & Olatoye, O. (2014). Global climate change. *Journal of Geoscience and Environment Protection*, 2(02), 114. <https://doi.org/10.4236/gep.2014.22016>
- Amegah, A. K., Rezza, G., & Jaakkola, J. J. (2016). Temperature-related morbidity and mortality in sub-saharan africa: A systematic review of the empirical evidence. *Environment international*, 91, 133–149. <https://doi.org/10.1289/isee.2016.3018>
- Benrhmach, G., Namir, K., Namir, A., & Bouyaghroumni, J. (2020). Nonlinear autoregressive neural network and extended kalman filters for prediction of financial time series. *Journal of Applied Mathematics*, 2020. <https://doi.org/10.1155/2020/5057801>
- Bluemm, C., Weiss, R., Weigel, R., & Brenk, D. (2010). Correcting nonlinearity and temperature influence of sensors through b-spline modeling. In *2010 IEEE International Symposium on Industrial Electronics*, 3356–3361. IEEE. <https://doi.org/10.1109/ISIE.2010.5637991>
- Bradley, E., & Kantz, H. (2015). Nonlinear time-series analysis revisited. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(9), 097610. <https://doi.org/10.1063/1.4917289>
- Claeskens, G., Krivobokova, T., & Opsomer, J. D. (2009). Asymptotic properties of penalized spline estimators. *Biometrika*, 96(3), 529–544.
- Craven, P., & Wahba, G. (1979). Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31, 377–403.
- DeGhetto, K., Gray, J. R., & Kiggundu, M. N. (2016). The african union's agenda 2063: Aspirations, challenges, and opportunities for management research. *Africa Journal of Management*, 2(1), 93–116.
- Demmler, A., & Reinsch, C. (1975). Oscillation matrices with spline smoothing. *Numerische Mathematik*, 24(5), 375–382. <https://doi.org/10.1007/BF01437406>
- Eilers, P. H., & Marx, B. D. (2010). Splines, knots, and penalties. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(6), 637–653. <https://doi.org/10.1002/wics.125>

- Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC press. <https://doi.org/10.1201/9781482273144>
- Godfrey, A., Burton, M., & LeRoux-Rutledge, E. (2012). "africa talks climate". comparing audience understandings of climate change in ten african countries. *The handbook of global media research*, pages 504–520. <https://doi.org/10.1002/9781118255278.ch29>
- Grimstad, B., & Sandnes, A. (2016). Global optimization with spline constraints: a new branch-and-bound method based on b-splines. *Journal of Global Optimization*, 65(3), 401–439. <https://doi.org/10.1007/s10898-015-0358-4>
- Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D. W., & Medina-Elizade, M. (2006). Global temperature change. *Proceedings of the National Academy of Sciences*, 103(39), 14288–14293. <https://doi.org/10.1073/pnas.0606291103>
- Hurvich, C. M., Simonoff, J. S., & Tsai, C.-L. (1998). Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2), 271–293.
- Hyrkkänen, J., Kilpinen, J., Nurmi, P., Kaurola, J., & Brockmann, M. (2016). Error characteristics of temperature forecast in finland for the period 1979–2011 in relation to various weather patterns. *Meteorological Applications*, 23(2), 244–253. <https://doi.org/10.1002/met.1550>
- IPCC (2001). Climate change 2001: The scientific basis. contribution of working group i to the third assessment report of the intergovernmental panel on climate change [houghton, j.t., y. ding, d.j. griggs, m. noguer, p.j. van der linden, x. dai, k. maskell, and c.a. johnson (eds.)]. cambridge university press, cambridge, united kingdom and new york, ny, usa, 881pp. Technical report.
- IPCC (2007). Technical summary. climate change 2007: The physical science basis. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change, s. solomon et al., eds., Cambridge University Press, Cambridge, 74pp.
- Jauch, J., Bleimund, F., Rhode, S., & Gauterin, F. (2017). Recursive b-spline approximation using the kalman filter. *Engineering Science and Technology, an International Journal*, 20(1), 28–34. <https://doi.org/10.1016/j.jestch.2016.09.015>
- Krishnamurthy, V. (2019). Predictability of weather and climate. *Earth and Space Science*, 6(7), 1043–1056. <https://doi.org/10.1029/2019EA000586>
- Md. Karimuzzaman, & Md. Hossain, M. (2020). Forecasting performance of nonlinear time-series models: an application to weather variable. *Modeling Earth Systems and Environment*. <https://doi.org/10.1007/s40808-020-00826-6>
- Michel, D., & Zidna, A. (2020). A new deterministic heuristic knots placement for b-spline approximation. *Mathematics and Computers in Simulation*. <https://doi.org/10.1016/j.matcom.2020.07.021>
- Mitchell, J. F. (2001). Detection of climate change and attribution of causes. *Climate Change 2001: The Scientific Basis*.
- Mudelsee, M. (2018). Trend analysis of climate time series: A review of methods. *Earth-science reviews*. <https://doi.org/10.1016/j.earscirev.2018.12.005>
- Reinsch, C. H. (1967). Smoothing by spline functions. *Numerische mathematik*, 10(3), 177–183. <https://doi.org/10.1007/BF02162161>
- Remontet, L., Uhry, Z., Bossard, N., Iwaz, J., Belot, A., Danieli, C., ... Group, C. W. S. (2019). Flexible and structured survival model for a simultaneous estimation of non-linear and non-proportional effects and complex interactions between continuous variables: Performance of this multidimensional penalized spline approach in net survival trend analysis. *Statistical methods in medical research*, 28(8), 2368–2384.
- Robert, H., & Shumway, D. S. S. (2006). *Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics)*. Springer.
- Romilly, P. (2005). Time series modelling of global mean temperature for managerial decision-making. *Journal of environmental management*, 76(1), 61–70. <https://doi.org/10.1016/j.jenvman.2005.01.008>
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press. <https://doi.org/10.1017/CBO9780511755453>

- Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(1), 1–21. <https://doi.org/10.1111/j.2517-6161.1985.tb01327.x>
- Singh, A., Dahiya, S., & Singh, S. (2020). A fourth-order b-spline collocation method for nonlinear burgers–fisher equation. *Mathematical Sciences*, 14(1), 75–85. <https://doi.org/10.1007/s40096-019-00317-5>
- Tamsir, M., Dhiman, N., & Srivastava, V. K. (2016). Extended modified cubic b-spline algorithm for non-linear burgers’ equation. *Beni-Suef University journal of basic and applied sciences*, 5(3), 244–254. <https://doi.org/10.1016/j.bjbas.2016.09.001>
- Wang, S., Feng, J., & Liu, G. (2013). Application of seasonal time series model in the precipitation forecast. *Mathematical and Computer modelling*, 58(3-4), 677–683. <https://doi.org/10.1016/j.mcm.2011.10.034>
- Wangler, B., & Backlund, A. (2005). Information systems engineering: What is it? In *CAiSE Workshops (2)*, pages 427–437.
- Ye, L., Yang, G., Van Ranst, E., & Tang, H. (2013). Time-series modeling and prediction of global monthly absolute temperature for environmental decision making. *Advances in Atmospheric Sciences*, 30(2), 382–396.
- Yeh, R., Nashed, Y. S., Peterka, T., & Tricoche, X. (2020). Fast automatic knot placement method for accurate b-spline curve fitting. *Computer-Aided Design*, page 102905. <https://doi.org/10.1016/j.cad.2020.102905>
- Zhou, S., Shen, X., & Wolfe, D., et al. (1998). Local asymptotics for regression splines and confidence regions. *The annals of statistics*, 26(5), 1760–1782. <https://doi.org/10.1214/aos/1024691356>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).