**Literature review: Climate Interpolation**

**NCEAS-NASA-IPLANT ENVIRONMENTAL LAYERS PROJECTS**

**Benoit Parmentier July 22, 2012**

1. **Introduction: Background/Context**

 In this paper we review studies related to the production of spatially explicit time series of climate layers. In particular, it examines the various aspects of methods, production and validation related to the interpolation of gridded surfaces of temperature and precipitation from meteorological station data. Temperature and precipitation are major environmental factors that are major factors in determining vegetation productivity and human societies are largely dependent on the understanding and prediction of these variables for human food and fiber forecasts (Kabat et al. 2004). Currently, temperature and precipitation measurements are only available at discrete points distributed irregularly in space, the meteorological stations, at coarse resolution (on the degree scale) through climate predictions (Kabat et al. 2004) and remotely sensed products (Huffman et al. 1995, Joyce et al. 1997), or through a number of interpolated surfaces (New et al. 2001, Hijmans et al. 2005, Jolly et al. 2005). While much effort has been done to improve the resolution and quality of interpolated products using downscaling methods (von Storch et al. 1993, Wilby and Wigley 1997, Flint and Flint 2005) or remotely sensed climate products (Hay et al. 1999, Reynolds et al. 2002, Huffman et al. 2006, Neteler et al. 2010), available datasets lack the spatial or temporal resolution needed for many applications in biology (Jackson et al. 2009). Thus, there is a need for spatially explicit products at higher spatial and temporal resolution extending back multiple decades (Hijmans 2005, IPCC 2007). Specific applications range from helping vineyard production in California (Nemani et al. 2009), initializing meteorological climate process based models (Kabat et al. 2004, Nemani et al. TOPS) or predicting natural disturbances such as the occurrence of beetle infestation in North America (Logan et al. 2001, Venier et al. 1998b, Stahl et al. 2006b). Therefore, such products benefit a wide number of fields and improve modeling related to biogeochemical cycles, agriculture (Changon and Kunkel 1999), forestry (Booth and Jones 1998), climate change studies (Hulme and Jenkins 1998, Giorgi and Fransico 2000, Liu et al. 2012), species range and habitat (Venier et al. 1998a, Parra et al. 2004, Peterson et al. 2008).

 The aim of this review is to provide a synthesis of the various methods of production and validation for the available climate products. This review does not purport to be exhaustive but strives to highlight and report on some of the major challenges faced by researchers when estimating climate variables at high temporal frequencies (monthly to daily) in complex or large landscape (Price et al. 2000, REF MTCLIM??). The paper is divided in five sections: Introduction, Methods, Climate Studies and Product, Accuracy and Validation and, Conclusions. Section 2 reviews statistical methods by describing in succession: Inverse Distance Weighting (IDW), Kriging, GAM/Spline, GWR and Climatology Aided Interpolation. Section 3 examines specific products and studies, summarizes the main features and resemblances as well as the challenges of productions and, provides readers with a short review of datasets available. This section includes more detailed descriptions of five studies/products: NEW01 (New et al. 2002), PRISM (Daly et al. 2002), WORLDCLIM (Hijmans et al. 2005), DAYMET (Thornton et al. 2012), SOGS-TOPS (Jolly et al. 2005). Section 4 focuses on accuracy and validation by surveying available methods and by presenting several strategies to deal with issues arising in accuracy and error assessment. Finally, the paper ends by summarizing major points and highlighting some possible avenues of research.

1. **Methods**

 Methods to produce fine grained climate layers fall broadly in two camps: downscaling of General Circulation Models and statistical interpolation based on meteorological stations (REF). In this review, we focus on statistical interpolation methods and do not describe downscaling methods.

 Statistical interpolation methods can be divided in five overlapping broad categories: environmental correlation/gradient methods, geostatistical /moving averages methods, hybrid methods, machine learning methods and Climatology Aided Interpolation methods. Environmental correlation methods predict new values of the response variables by modeling the relationship between the response and a set of covariates (Hengl 2009) such as elevation or coastal proximity (Jarvis et al. ?). Common examples include multiple linear regressions (Goodale ?,Lennon and Turner 1995,REF), Generalized Linear Models (REF) or Generalized Additive Models (GAM) (Guan et al. 2009). Geostatistical and moving averages methods are often used in the context of spatial analysis where interpolation refers to the process of predicting unknown values of response variables (y0) at unsampled geographical locations using data observations at sampled location (yi) (Isaaks Srivista 1989, Burrough and McDonnell et al. 1998) with common examples including IDW (Willmott and Matsuura 1995) and Kriging (Phillips et al. 1992, Goevarts et al. 2000, Attore et al. 2007). Hybrid methods refer to methods using a mixture of geostatistical and environmental correlation methods (Daly et al. 1994, Daly et al. 2002, Stahl et al. 2006) i.e. where predictions are made by using both environmental factors and neighboring observations. PRISM for instance, uses linear regression (precipitation-elevation relationship) with geostatistical methods by weighting observations using covariates and Inverse Distance Weighting (Daly et al. 2002). Hybrid methods form a continuum between methods using high degrees of environmental correlation or geostatistical methods. Machine Learning (ML) methods draw inspiration from the data mining field. ML methods are based on the framework of pattern recognition where the goal is to search for meaningful patterns in large datasets (Tveito et al. 2006). In the context of climate interpolation, ML methods learn typical patterns from a training dataset to predict the response value (temperature and precipitation) using patterns in a set of features i.e. environmental covariates (latitude, longitude, elevation etc.). Typical examples include the interpolation of precipitation using Neural Network such as the Maximum Layer Perceptron (MLP), the regression trees and the Radial Function Basis Network ((Tveito et al. 2006, Attore et al. 2007, Lin et al. 2008, Snell et al. 2010).

*Table 1. Broad categories of interpolation methods for climate layers production.*

|  |  |  |
| --- | --- | --- |
| Methods | General form | Examples |
| Environmental correlation/gradients | Y= a0+a1x1+a2x2+…+an\*xn | Multiple linear regression with environmental covariates |
| Geostatistical/moving averages | y0= Σai\*yi with ai being theweights | IDW, Simple Kriging, Ordinary Kriging |
| Hybrid methods | Y= a0+a1x1+a2x2+…+an\*xnY= a0+a1x1+a2x2+… Σai\*yi ai dependent on distance and /or covariates  | Universal Kriging, PRISM,GAM-TPS. |
| Machine Learning | y= pattern(x1,x2,…,xn) | MultiLayer Perceptron (MLP), Regression Tree |
| Climatology Aided Interpolation | Y= Yav+ YdevYdev =a0+a1x1+a2x2+…+an\*xn | Two stages regression with monthly average/modeled surface and daily anomalies/deviations.  |

In the following section we focus only on a few methods cited above: IDW, Kriging, GAM and Splines and the Climatology Aided Interpolation. These are widely used in climate literature to produce gridded surfaces.

**2.1 Inverse Distance Weighting**

 IDW is a form of moving average that predicts new values based on the observed values in the vicinity of the unsampled location of interest (Sheppard 1968, Myers 1994, Waller and Gotway 2004). The predicted values are found by a weighted average sum of neighbors with weights specified by a kernel function (Myers 1994). The shape of the kernel functions typically exhibits a decreasing trend as a function of distance from the prediction location reflecting the fact that distant observations contribute less to the average prediction (Waller and Gotway 2004). Strictly speaking the term “IDW” applies to techniques that use a kernel weight function corresponding to an inverse power of distance of the form y= 1/xp . The power parameter “p” determines the speed at which the influence of observations decreases. It is typically equal to one or two but can be evaluated from the data itself through fitting (REF). The exponential decay function is also often used and many other forms are possible as long as weights sum to one. Whenever performed over large distances, the curvature of the Earth may be taken into account to improve results (Renka 1984, Wilmott 1995) by using algorithms incorporation spherical geometry with different degree of differentiability and continuity (Renka 1984).

 IDW has proven to be a popular method for interpolation because of its simplicity and relative efficiency (Willmott 1995). The method may suffer from bias however, because of the misspecification of autocorrelation and the lack of incorporation of environmental factors in the estimation (ADDREF). To alleviate these problems, the weighting function may be modified to add directional information, extended to the cosine weighting function or to include environmental gradient (Willmott et al.1995 and Willmott and Legates 1990) such as an adjustment for elevation (Thornton et al. 1997, Stahl et al. 2006).

 In the recent years, a modified version of IDW has been used in the Daymet product. Called “Truncated Gaussian Filter, “TGF”, the method predicts new values by first considering a maximum bandwidth determined locally according to the configuration of stations and second by computing new values as weighted averages using a Weighted Gaussian function (Thorton et al. 1997, Thornton et al. 2000). TGF has also been modified to adjust for the effect of elevation so that in effect it can be considered as a hybrid method that resembles in some ways to the PRISM methodology (Daly et al. 2002).

* 1. **Kriging**

 Kriging is a form of interpolation which derives predicted values using the weighted average sum of its neighbors (Krige 1966) in a manner similar to IDW. It was first introduced by Krige in 1951 in the context of mining (Krige 1951) and formalized mathematically by Matheron (1969). While it is also a form of moving average, it differs from IDW because its weighting function is derived from the spatial structure of the data itself rather than solely based on the distance to the interpolated location. Weights are determined as “best” by optimization of an objective function describing the variance of the prediction errors under the constraint that the weights must sum to 1 (Isaaks and Srivastava 1989, Waller and Gotway 2004, Burrrough and McDonnell 1998). The weight calculation uses a covariance function obtained by fitting a function to a semi-variogram plot which describes the dissimilarity of pairwise observations as a function of geographic distance and direction. When the semi-variogram plot is anisotropic its surface reduces to a one dimensional function of distance. In theory the semi-variance is equal to zero at the origin and increases until a specific distance, the “range”, at which point it becomes flat and reaches its maximum value at the “sill”. The semi-variogram shape’s describes the increasing variation in the pairwise differences as spatial autocorrelation decreases with increasing distance until spatial autocorrelation becomes zero and the variation becomes similar to the dataset ‘s variance (Isaaks and Srivastava 1989, Waller and Gotway 2004). The shape of the semivariogram at the origin reflects the continuity or smoothness of the spatial autocorrelation (Waller and Gotway 2004) with parabolic rises at the beginning indicating smooth spatial continuity with differentiability and linear rises indicating continuity without differentiability. Discontinuities are also described by nuggets which are signs of sudden jumps in neighboring values that can be due to errors in the measurements or actual discontinuities in the spatial process (Waller and Gotway 2004:276).

 Fitted semivariograms can take many forms with common models including Gaussian, exponential, spherical and linear. Spherical models display nearly linear increases at the origin while Gaussian models display a change of curvature within the range with a lesser rise near the origin. Exponential models have very smooth transitions with no formal ranges since they approach ranges asymptotically. More general models such as the Matern covariance can be defined (REF), as long as their covariance functions (or variance-covariance matrices) comply to the condition of negative definiteness or equivalently to the condition of positive-definitiveness (Waller and Gotway 2004).

 There are five types of Kriging methods commonly found in the literature: simple Kriging (SK), Ordinary Kriging (OK), Universal Kriging (UK), Kriging with External Drift (KED) and Regression-Kriging (RK) (Bivand et al. 2008, Hengl 2009). Simple Kriging is used when the mean of the response variable is constant, known or has been removed in the study area. Ordinary Kriging assumes the regional mean is constant but unknown while Universal Kriging assumes that the regional mean is not constant and must be modeled via trends analysis or regression (Bivand et al. 2008, REF). The term KED is sometime used to differentiate the surface trend calculation based on geographical coordinates only (UK) with Kriging processes where the surface trend estimation is done using both geographic coordinates and environmental covariates (REF, Hengl et al. 2009). KED, OK and UK simultaneously compute the kriged weights and coefficients of the surface trends by constructing a covariance matrix that incorporates covariates and the constant mean term. In practice, this is done by adding constraint variables (Lagrange multipliers) which translate into adding rows and columns to the covariance matrix. For instance, a column and row of 1 are added to the covariance for OK while additional columns and rows are added for each covariate in UK and KED. Regression-Kriging is a variant of Universal Kriging similar in many aspects to UK but which separates in two distinct steps the fitting of trends (mean) from the estimation of kriged weights. In other words, RK proceeds by first calculating the trend surface using linear modeling with Generalized Least Squares[[1]](#footnote-1) incorporating the spatial structure and, by second performing simple Kriging on the residuals (Hengl 2004, Hengl 2007, Odeh et al. 1995). Compared to UK and KED, RK provides additional flexibility by allowing the incorporation of non-linear relationship models in the estimation of the trend surface (e.g. GAM). All Kriging methods share commonality and Hengl 2004 and 2007 demonstrates that UK or RK reduce to OK or “pure Kriging” when the environmental covariates have no predictive power i.e. exhibit no correlation with the response in which case the mean reduces to the global mean of the response variable. Similarly, RK reduces to multiple linear regression (“pure” regression in Hengl 2007) when the semivariogram is flat and when the residuals have no autocorrelation and are characterized by “pure nugett” effects (Hengl 2009). Thus there is a continuum of cases that depends on the correlation structure between the response variable and the covariates and the spatial autocorrelation present in the response.

 Co Kriging is another Kriging method that is used in interpolation studies (Hartkamp 1999). It is a multivariate version of Kriging (Hoe and Cressie 1993, Pebesma 2004) which relies on the similarity between covariates’ spatial structure. Like UK and KED, Co-Kriging also uses covariates to improve prediction but uses a cross-variogram to include the spatial structure. Cross-variograms are similar to semi-variogram in all aspects with the exception that they represent cross-variation with respect to two different covariates rather than autocovariance (REF, Hevesi et al. 1992\*).

 In the context of the climate interpolation, Kriging has been used in many studies (Phillips et al. 1992, Goovaerts 2000) and has performed well (Dingman eta l. 1988?, Hartkamp 1999?) and several authors argue that it provides the most accurate method of interpolation (Dubrule 1983\*, Burrough and McDonnel 1998, Stein and Costein 1991, REF recent). For instance, studies indicate that Kriging and co-kriging methods might be less sensitive to irregular network than splines (REF\*). However when compared to splines, Kriging is computatively more intensive and results are highly sensitive to the accuracy of the fitted variograms. Thus, one of the main challenges of the technique is to fit a representative variogram in an automated fashion. To tackle such challenges, various methods have been used to use regression fitting approaches with RMSE (Hiemstra et al. 2008) and robust estimator techniques (Cressie and Hawkings 1980, Attore et al. 2007).

**2.3. Splines and GAM**

 First introduced by Schoenberg in 1946 in the context of mathematics, splines have found many applications in engineering, computer graphics (Bartels 1987), geometric design (Duchon 1970) and environmental sciences (Wood 2002). Over the years, splines have been used multiple times in climate studies to predict precipitation and/or temperature in studies such as NEW01 (New et al. 2002) and WorldClim (Hijmans et al. 2005), Hofierka et al. (2002) and (Tait et al. 2006). Splines form a family of methods that use a set of smooth functions to represent the relationship between a response (e.g. maximum temperature) and a set of covariates (elevation, latitude etc.) allowing a flexible representation of non-parametric relationship (Hastie and Tibshirani 1990). In smoothing splines, the relationship between response and covariates is modeled as the sum of the bases in subsections of the domain of interest. Splines may take a variety of functional forms with different properties but often used are cubic or higher degrees polynomials. Since polynomial bases can be good approximators locally but may diverge in important ways globally, the domain of prediction is divided in subintervals with each point between sections called “knots” resulting in a piecewise polynomial function. Polynomials correspond to bases functions that can be rescaled in different subsections and fitted more tightly to represent the relationship between the response and covariates. The spline is smooth in the sense that the piecewise function and its derivatives are continuous at the knots (enforced through mathematical constraints).

 The early development of splines methods took place outside the fields of statistics (Schoenberg 1946) and formal connection to linear modeling is made possible by using the statistical framework of General Additive Model (GAM) (Hastie and Tibshirani 1990, Wood 2006). Under this framework, splines may be understood as a generalization of linear modeling theory to non-parametric smooth functions and an extension of Generalized Linear Models (GLM). The term additive refers to the fact the relationship is described as a sum of terms of linear smooth functions (i.e. hence its alternative name splines). These additive terms are quite general and flexible and can include interaction among several covariates (Hastie and Tibshirani 1990). ) For instance, TPS implementation in the ANSUPLIN software (Hutchinson 2004\*) provides a mean to fit a spline in a three dimensional space with the covariates such as elevation, latitude, longitude. Thus, TPS is a form of model that allows for interactions between the three covariates. In addition to smooth terms, GAM can also include parametric terms thereby bringing further greater flexibility to the modeling framework (Figure 1).

Y= θX + f(x1) + f(x2) + f(x3,x4) +…

With the penalized least square criterion: ||y-f ||2 + λ I (f)

 where f is the smooth function to be found

                                             λ is the smooth parameter

 Y is the response variable and xn is the covariate n.

*Figure 1. The general form of a GAM model can be expressed as a sum of smooth functions f(xi) and a parametric component θX. The model is fitted using the penalized least square method with a fitting and a smoothing term.*

 One of the most difficult problems in splines is to determine the number and positions of the knots along with the polynomial coefficients. This process can be automated by using a penalized least square approach in which an objective functional is defined based on two criteria, a fitting criterion expressed as the sum of square of residuals||y-f ||2 and a roughness criterion. The solution to the optimization problem is a function that minimizes both criteria given some constraints related to smoothness (i.e. continuity at the knots). Whenever a sufficient number of knots is chosen, smoothing is independent of the knots’s position and the smoothing parameter λ controls the degree of smoothness (Wood 3003, Wood 2006). The smoothing parameter λ can be determined from the dataset by the method of Generalized Cross Validation (GCV). It represent relates to the degree of freedom of the smooth function.

 The roughness criterion, I (f) corresponds to the *wiggliness* of the smooth function. In most cases, the roughness criterion is represented using the sum of square of the second derivative or partial derivative in the multivariate case (Mitas and Mitasovas 1999). In one dimension the solution to the functional is the cubic spline while in two or more dimensions it is the Thin Plate Spline model (Dubrule 1983, Duchon 1970, Wood 3003, Wood 2006). In the multidimensional case, the roughness criterion I(f) is expressed in mathematical terms as λJmd(f) where J is a Jacobi matrix, with dimension “m” corresponding to the function space and the number of partial derivatives. The objective functional can be generalized by using a membrane component, a tension parameter and other constraints and allowing any order of derivatives in the roughness term (I(f)) in which case Regularized Spline with Tension (RST) are obtained (Mitas and Mitasova 1988,Mitas et al. 1999, Hofierka et al. 2002).

 There is also a close connection between piecewise polynomial splines, kernel functions and moving average filters (Matheron 1981, Myers 1994, Hutchinson 1994). Spectral decomposition can be used to study the local behavior or actions of the cubic spline, TPS or any spline smoothing function. This is done by performing an eigenvalue –eigenvector decomposition (Hastie and Tibshirani 1990:59) on the smoother matrix S corresponding to the smooth function. The eigenvalues describe the amount of scaling of each eigenvectors while the eigenvectors relate to the weight structure influenced by the smoothing parameters. In essence, eigendecomposition allows the recasting of the smoothing function into a transfer function that expresses the kernel action in the frequency domain. In this manner, the polynomial cubic splines can be associated with a low pass filter as evidenced by its eigenvalue spectrum (Hastie and Tibshirani 1990:59). In GAM and splines, the size of the neighborhood of the acting kernel is related to the smoothing parameter λ (Hastie and Tibshirani 1990, Woods 2006) with large and more global neighborhood having lower variance than local ones. Larger neighborhoods can however result in greater biases than smaller ones reflecting the fact that there is a tradeoff between variance and bias. When infinitely large neighborhoods are used, λ tends to infinity and the smooth function becomes linear function.

**2.4 GWR**

 Geographically weighted Regression (GWR) is a form of localized regression that can be used to model the relationship between climate response variables and environmental covariates (Daly et al. 2006, Lloyd et al. 2010). GWR provides a mean to study or account for spatial heterogeneity by fitting a regression lines at every location in the study region using weighted observations in a neighborhood (Fotheringham 2002). For every grid point, the method proceeds by setting the bandwidth which determines the size of the neighborhood to be taken into account in the regression and uses a kernel function to describe the weigh structure (Brunsdon et al. 1996). Kernel functions describe the contribution of observation as function of distance from the prediction location. Common weight functions include the exponential decay, the Gaussian function, the inverse power distance weighting with varying degree of power control the distance decay (as described earlier).

 Thus GWR is a method that attempts to integrate spatial heterogeneity in the modeling process and alleviates problems found in regression or Kriging where relationships are assumed to be stationary (Fortheringham 2002, Brundson et al. 2001). GWR can also detect local variations that do not appear in global models because effects may cancel out in the study area (Fotheringham 2002). In effect, GWR can be seen as a form of LOESS in two dimensions where the weights are dependent on the geographical coordinates (Hastie and Tibshirani 1990).

 In the climate and environmental literature, GWR has been used to model variations in the relationship between forest and rainfall (Foody 2003), Net Primary Productivity (Wang et al. 2005), changes in the lapse rate and in the precipitation-elevation slopes (Brunsdon et al. 2001). For instance, an application of the method in Great Britain revealed that the slope precipitation-elevation parameter varies from around 4.5mm in the Northwest to zero in the Southeast. Other such as Bostan and Kyurek (2007) evaluated GWR usefulness in predicting temperature and precipitation in Turkey and found that the method performed better than Kriging with environmental covariates.

**2.5 Anomalies and climatology aided interpolation**

 Climatology Aided Interpolation (CAI) and anomalies approaches are multiple steps procedures that make use of long term averages to estimate the interpolated values. The main idea is to separate the variance and its estimation in two components: a long term component and a short term variation called “anomalies” or “deviations” (Daly et al. 2006). The modeling of the long term component is done either through calculation of long term averages in which case it is equivalent to a “normal” or through some interpolation methods such as SMART interpolation (Willmott and Matsuura), TPS-GAM, Kriging (Hosfra et al. 2008) or PRISM (Hunter and Meentemeyer 2005). The short term variation is obtained by calculating the deviation from the long term base surface and can itself be modeled through Kriging or some other methods (REF). The general form of CAI can be written as:

T (estimated) = T (climatology) + T(deviation)

 The climatology approach can be traced back to the SMART interpolation by (Willmott and Matsuura 1995) and to Willmott and Robeson (1995) who introduced it formally as the Climatology Aided Interpolation. The logic behind the CAI approach is to use the high quality, higher spatial resolution surface as a base from which temporal spatially explicit layers at finer temporal time steps can be derived (Daly et al. 2006). The higher quality climatology surface can be built using more data from longer time periods which is particularly useful when sparse and intermittent network are available from shorter time period (Daly et al. 2006, Willmott and Maatsura 1995). While the average is based on a longer period, the inclusion of stations over such large time period overweighs the exclusion of them in the accuracy because of the noted improvement in spatial fidelity which often includes the resolution of finer features. For instance, Willmott and Robeson 1995 used an extensive database with long term averages for annual temperature containing some 18,000 station database (Willmott and Legate 1990) and found that this surface was able to resolve limits between deserts (e.g. Atacama) and improved substantially model predictions through time (Figure 2). Similarly when PRISM was used as the base surface corresponding to monthly long term averages and daily deviations were modeled using interpolation techniques to predict precipitation and temperature in California, Hunter and Meentemeyer (2005) found that CAI performed better with a lower MAE compared to interpolation techniques such as Kriging (Hunter and Meentemeyer 2005). Other examples of CAI and anomalies applications can be found in New et al. (2000) for the world, in Plantico et al. (2000) for USA, in Funk et al. (2002) for the African continent, in Di Luzio et al. (2008) for conterminous USA, in Perry and Hollis (2005) for the United Kingdom.

****

*Figure 2. Improvement in MAE by using CAI method as compared to the direct method (Willmott and Robeson 1995).*

**2.6 Typology of methods**

 In an attempt to summarize the different techniques used in interpolation, we provide a typology of methods by characterizing methods in terms of four characteristic: weighting scheme, use of covariates, global trend, spatial heterogeneity and degree of automation.

*Table 2. Typology of interpolation method based on four criteria:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Methods** | **Local component/****Weighted average or kernel function** | **Covariates used as predictors** | **Global component/****Trend function** | **Degree of automation** | **Studies** |
| IDW | Fixed, distance based | No | No/yes | Yes | Shepard 1968, Renka 1984, Wilmott et al. 1995, Legate and Willmott 1990, Dai et al. 1997,Thorton et al. 1997, ,Doson and Marks 1997,  |
| Kriging simple | Data/empirical and distance based | No | Null trend | No/yes | Philipts et al. 1992, Garen et al. 1994, Dingman et al. 1988, Hevesi et al. 1992, Garen et al. 1994? |
| Ordinary Kriging | Data/empirical and distance based | No | Constant trend | Yes | Jolly et al. 2005 (SOGS),  |
| Universal Kriging/Regression Kriging | Data/empirical and distance based | Yes | Trend modeled by coordinates or covariates | Yes | Hengl et al. 2007, Attore et al 2007. |
| Localized Kriging | Data/empirical and distance based | Yes | May include trends | No/yes | Add REF |
| Co-Kriging | Data/empirical and distance based | Yes | May include trends | No/yes | Philipts et al. 1992, |
| GWR | Fixed, distance based | No/yes | No trend | Yes | Llyod 1999, Brundson 1999 |
| GAM/Splines/TPS | Fixed,Solution from optimization | No | Non linear trend estimation | Yes | Wahba &Wendelberger 1980, Hutchinson et a. 1995, New et al. 199, New et al. 2002, Hijmans et al. 2005, Hong et al. 2005,Tait et al. 2006, Guan et al. 2009. |
| PRISM | Yes, empirical | Yes | No | No, Knowledge based system +statistical methods | Daly et al. 1994,Daly et al. 2002 |
| CAI/Anomaly | Dependent | Dependent | Yes | Yes | Willmott et al. Robeson 1995,Haylock et al. 2008, Hunter and Meentemeyer 2005, Perry and Hollis 2005, Di Luzo et al. 200? |
| LM and GLM | No but can be included making it GLS | Sometime | Yes | Yes | Jarvis and Stuart 2001, Bolstad et al. 1998, Xia et al.  |

 Estimating an interpolation function requires making some trade-off between fitting the data (exactness) and smoothing (filtering) noise or details (Myers 1999). In modeling terms, this may be translated into the problem of capturing two components: the general trend and the local variation. The local component is typically integrated using a kernel function (equivalent to a weight function) acting on subset of observations in the domain while the global component often takes the form of a parametric or non-parametric function with geographic coordinates and environmental covariates (Myers 1999, Mitas and Mitasovas 1999). The approach often proceeds from global to local, by first fitting a trend function T (x) to represent the global variability and by second adding local variability using a Kernel function R (x). In some cases, it is possible to estimate both T(x) and R(x) at the same time such as in Universal Kriging or in regression splines (e.g. TPS and RST). In fact, it can be shown that TPS models with latitude and longitude covariates relate to Universal Kriging models with biharmonic kernel functions corresponding to semivariogram fuctions[[2]](#footnote-2) (Dubrule et al. 1983) with pure nuggets (Myers 1990). In TPS and RST, T(x) and R(x) are formally related and form the general solution to the dual of the penalized least squares objective functionals (Wahba 1990). The dichotomy between local and global estimation may be loosely understood as reversing the role of observations and variables in the context of mathematical duality[[3]](#footnote-3). By taking higher order derivatives as roughness conditions in the penalized least squares, kernels with higher degree of smoothness are obtained and splines’s kernels can be related to Gaussian functions (Mitas and Mitasovas 1999) in which case there is infinite differentiability in interpolated surfaces. In contrast to TPS and RST, Kriging does not relate directly the trend function (T(x)) and the kernel function (R(x) so that the smoothness of the interpolated surface is not always guaranteed (Hutchinson and Gessler 1993).

 Spatial heterogeneity must also be dealt with when interpolating. When there is spatial heterogeneity, the study area may be stratified in different regions and the global trend becomes a regional trend. By increasing stratification and reducing the size of the region, we can derive a localized trend for every location obtaining a solution similar to GWR. Other methods such as Localized Kriging also account for this spatial heterogeneity by fitting semi-variograms and trends to subregions of the study area (REF).

 Covariates are also often used as a criterion to create methods’ typologies (Daly et al. 2006, Hengl 2009). Covariates are most often included in the global component but may also be incorporated in the local component in which case they influence the structure of the weight function directly such as in Splines or PRISM. In fact, in spline methods, there is no distinction between geographical covariates (latitude and longitude) and environmental factors so that covariates are always included in the calculation of distances for kernel functions resulting in non-geographical weighted distance functions. A common example found in the literature is the trivariate TPS using latitude, longitude and elevation (Hijmnans et al. 2005).

 The degree of automation is also widely used as criterion to choose a method. Methods that require manual fitting of parameters such as splines without regularization, PRISM or Kriging are at a disadvantage in large scale studies (Hutchinson) and fine temporal resolution studies where manual fitting may be impossible for every time step and subregions. Hence, TPS is often used in large scale study because of its ease of automated fit (New et al. 2001, Hijmans et al. 2005, REF) and theoretical properties (Mitas and Mitsova 1999) contrasted to Kriging even in presence of automated variogram fitting.



*Figure 3. Interpolation methods can be viewed as estimating global components (trend surfaces) and local components (kernel functions). In some case, such as Kriging, TPS and Regularized Tension Splines, the general solution may be written mathematical in separate terms as “global+ local”.*

1. **Climate studies and product**

**3.1 Overview of studies and workflow**

 Most interpolation projects can be summarized in three workflow stages: 1) database development and processing, 2) interpolation and 3) output assessment. The SOGS-TOPS system provides a prime example of such workflow in Jolly et al. (2005) (Figure 4). From the survey of studies, we have drawn a list of five issues that often occur in interpolation projects: 1) sparse and unequal station network, 2) large geographical variability and non-stationarity 3) database quality and completeness 4) Validation in a sparse data context 5) degree of automation and incorporation of expert knowledge. These issues are summarized in table 3 along with various reported strategies from reviewed papers.



*Figure 4. Interpolation Workflow from SOGS-TOPS (Jolly et al. 2005)*

*Table 3. Some of common problems and strategies for climate interpolation.*

|  |  |  |
| --- | --- | --- |
| **CLIMATE INTERPOLATION: ISSUES** | **Workflow stage** | **Strategies** |
| 1. Sparse and unequal density of station network  | Database | - Include data such as satellite information,- Evaluate accuracy from the network. - Assemble data from many alternative sources |
| 2.Large Geographical variability and non-stationary  | Interpolation | - Divide the study area in multiple sub regions- Use a model with local adjustment |
| 3.Database Quality and Completeness  | Database | - Multiple screening necessary with possibility of following WMO or NCDC like procedures. - Extend temporal period or spatial extent of study area |
| 5.Automation and incorporation of human expert knowledge  | Interpolation/Output assessment | - Reduce manual fitting of parameters.- Increase human input in validation |
| 4.Validation in a sparse data context | Output assessment | - Use cross-validation and evaluate accuracy by average gridding |

In this section, we address in briefly workflow issues stage 1 and stage 2 relating to the development of the database and the interpolation methods. Issues related to the assessment of the output are described in more details in section 4.

* 1. **Meteorological Station network and Database quality**

 While in the recent years, satellite data have also been added as an alternative source of information on climate, climate surfaces are primarily derived from observations from meteorological stations. Interpolation is therefore greatly dependent on the existence of large high quality meteorological databases (REF, Legates and Willmott\*). Before being deemed useful for interpolation, databases must undergo stringent quality controls to detect measurement errors (Durre et al. 2008) as well as spatial locations inaccuracies in coordinates: latitude, longitude and elevation (Wieczorek et al. 2004). Harmonizing databases is not an easy task and requires cooperation and standardization among many networks through international agreements and the sponsoring of international institutions such as the World Meteorological Organization (WMO) and the FAO (REF,Durre et al. 2008, Mitchel et al. 2005\*). Despite the existing international collaboration, many databases contradict or overlap each other. Studies could greatly benefit from the adoption of schemes to uniquely identify stations (Hijmans et al. 2005) for instance, using the WMO procedure so that different database can be merged together in a manner similar to the GHCN project. The majority of studies report database quality screening before the interpolation stage. For instance, Daly et al. 2002 carried out a range checking for minimum and maximum temperature and precipitation values. These were evaluated to eliminate negative and detect extreme values with application of threshold based on the country of station and climate conditions from NCDC (Daly et al. 2002). Screening is problematic because it usually reduces the number of observations. Additional loss may happen because records are incomplete over time (Hutchinson et al. 1995). In that paper, for example, averages were only included if there were calculated from 10 years of record and the time period was extended from 1960-1990 to 1950-2000 to allow for more stations to be accounted.

* 1. **Interpolation: methods selection**

 There are numerous studies that investigate multiple interpolation methods for particular study regions: Bolstad et al. (1998) in the Southern Appalachian mountains, Xia et al. (2001) in Bavaria Germany, Jarvis and Stuart (2001) in England and Wales, Attore et al. (2007) in Italy, Hosfra et al. (2008) in Europe, Stahl et al. (2006) in British Columbia, Bazgeer et al. (2012) in Fars (Iran). Results have been inconsistent, with some papers finding Kriging to be the best method (Weber and Englund 1994, Goovaerts et al. 200, Attore et al. 2007, Bazgeer et al. 2012) while others selected TPS (REF) or even IDW (Weber and Englund 1992,Willmott 1995?,Thorton et al. 1997, Lu and Wong 2008) as the best performers. For example, Attore et al. 2007 compared detrended IDW, Universal Kriging and Multilayer Percepteron and found that UK with external drift performed the best based on cross-validation test and RMSE. Similarly, Hosfra et al. (2008) compared six methods (local Kriging, global Kriging, IDW, 2D TPS, 3D TPS, natural neighbors interpolation) and found that global Kriging was the best performer for the interpolation of temperatures (tmax, tmin), precipitation and sea level pressure in Europe. In contrast, Price et al. (2000), compared a form of IDW with elevation and geographical coordinates adjustment (Gradient plus Inverse Distance Squared, GIDS) with TPS and found that IDW performance was slightly lower in terms of RMSE with TPS performing better in areas where interpolation were difficult to predict (Price et al. 2000).

 As exemplified in the contradicting reports on accuracy, much of the method selection relies on properties and specificities of the interpolation methods themselves as applied in variety of contexts (i.e. variability of the study region). Proponents of IDW highlight the fact that the method is simple and often gives similar results to more complex techniques (Thornton et al. 1997, Jolly et al. 2005). Proponents of Kriging on the other hand highlight its good performance and laud its statistical foundation based on the optimality of its best linear unbiased prediction (Hengl et al. 2007) predictions (Matheron 1981) which allows for the formal estimation of errors. In contrast, some authors highlight the challenge posed by Kriging in fitting representative variograms to datasets (Hutchinson et al. 1995) in particular when this needed multiple times in dense time series. Some authors propose a workaround by automatic fitting of variograms (Hiemstra et al. 2008). TPS/GAM methods are often presented as alternative to Kriging methods and presented as advantageous because of their high degree of automation which requires no empirical fitting of variograms to estimate kernel functions, manual tuning of parameters or weights such as in PRISM. Furthermore, evaluation of errors using Bayesian framework is also possible in TPS making it attractive for uncertainty assessment (Wood 2003, Wood 2006, Wahba 1990?). PRISM, however, clearly outperforms Kriging and other methods in North America but is not easily scalable to large scale interpolation studies due to the amount of input expert knowledge and its low degree of automation (Hijmans et al. 2005).

 Model comparison studies illustrate that in many cases methods perform similarly (REF, Attore et al. 2007, Thornton et al. 1997). For instance, Jolly et al. (2005) reported similar results in terms of MAE and RMSE in predictions of climate variables using OK, TGF and IDW in Continental USA. The interaction between the properties of the method, the size of the study area and the availability of station data and its configuration largely determines the performance of the method. Thus it appears that *there is no single best method that can be applied in all situations; accuracy is largely dependent on case by case basis with the most important factors being the density of the station network and the variability of the study area.*



*Figure 4. Summary table of accuracy.*

**3.4. Characterizing studies: short survey**

 We provide here a non-exhaustive survey of methods and studies related to climate interpolation. Since they may vary in terms of many factors such as study region, methods or temporal coverage (Hartkamp 1999, Stahl et al. 2006), we have opted to characterize studies using a set of seven dimensions: spatial extent/region, temporal extent/period, temporal resolution, spatial resolution, response variables, method and accuracy (Table 4).

*Table 4. A non-exhaustive survey of studies and product related to climate interpolation.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Spatial extent/region** | **Temporal extent/period** | **Method** | **Temporal****Resolution/frequency** | **Spatial resolution****(grain)** | **Response Variables** | **Accuracy** | **Explanation** |
| PRISM | USA | 1980-2010 | Regression, maybe GWR | Monthly | 4km | Tmax, tmin, precipitation | MAE=1.6 C | This MAE is reported for Coastal California only. There is no full report of MAE or RMSE |
| Daymet | World | 1980-2008 | Truncated Gaussian Filter | Daily | 10minutes | Tmax, tmin, precipitation, radiation, humidity |  |  |
| SOGS-TOPS | CA,North America, World | 1982-2012? | OK, Truncated Gaussian Filter, IDW  | Daily | 1km (CA), 8km (USA), 0.5 deg (World) | Tmax, tmin, precipitation, shortwave radiation, Vapor Pressure Deficit | MAE:1.6 for Tmax OK, 1.9 Tmin OK,48 mm precip (OK) | There were three methods used and the reported accuracy is for the 2002 run from Jolly et al. 2005. OK had the lowest MAE |
| WorldClim | World | 1950-2000 | GAM/TPS | Monthly | 1km | Tmax, tmin, precipitation | 0.5-2 | The MAE in Oregon is about 0.5 for 10 degree squares |
| Willmott 1985 | World | 1881-1990 | IDW | ?? | ?? |  | 1.3-1.9C |  |
| Willmott 1995 | World | 1881-1990 | CAI:IDW | Annual | ? |  | 0.75-.1.5 |  |
| Haylock et al. 2008 | Europe | 1950-2006 | CAI: TPS+Kriging with drift | Daily | 0.1 or 25km? |  | 0.5-1C |  |
| NEW99 | World | 1961-1990 | TPS | Monthly | 0.5 deg | *Pcp,wdf,RH,sun,Trange,gff,w\_speed* |  |  |
| NEW01 | World | 1961-1990 | TPS | Monthly | 10’ | *Pcp,wdf,RH,sun,Trange,gff,w\_speed* |  |  |
| HadGHCND | World | 1946-2000 | ADW | Daily | 2.5 latBy 3.75 long | Tmax, Tmin |  |  |
| Thorton et al. 1997 | Northestern USA | 1989 | IDW: truncated GuassianWith weights dependent on elev.  | Daily | ? | *Pcp,wdf,RH,sun,Trange,gff,w\_speed* | ? |  |
| Feng et al. 2004 | China | 1951-200 |  |  |  |  |  |  |
| Groot and Orlandi 2003 | Europe | 1975-2000 | Nearest Neighbour, IDW |  | 50km | *Temperature and precipitation* | ? | ? |
| Hewitson and Crame | South Africa | 1950-2000 | Conditional Interpolation |  | 0.1 deg | *Precipitation* | ? | ? |
| Stahl et al. 2006 | British Colombia Canada | 1965-2000 | IDW, Kriging, Mulitple regression | Daily |  |  |  |  |
| McKenney et al. 2006 | Canada-USA | 1901-2000 | TPS | monthly | ? | *PRCP, tmax,tmin* | MAE: 1-1.5C, 20-40 PRCP |  |
| Leemans and Cramer 1991 | Global | ? |  |  | 0.5 deg |  |  |  |
| Hunter and Meetemeyer 2005 | California-USA | 1980-2003 | CAI, Kriging |  |  |  |  |  |

 There are some important comments that that can be drawn from the table 3 and from the review of studies (table 4). First, the spatial extent has an important influence on the methods, particularly; methods used in areas with large spatial extent must be reproducible and automatable. As mentioned earlier, PRISM appears most accurate in the USA but its methodology is not easily automatable for production at the global scale (NEW01). Second, whenever the study area is large, it is divided in multiple subregions or tiles where individual interpolations are carried out. Subregions are chosen with some overlap to allow for the calculation of smooth edges and transitions between tiles (New et al. 2001, Hijmans et al. 2005).

 From a general point of view, finer grained datasets are able to discern more variability as exemplified by the comparison between NEW01 and PRISM (New et al. 2002). In all cases however, the influence of the station network on the reconstitution of the spatial variation is paramount so that fine grain resolution may give a false sense of detail and accuracy (New et al. 2002, Daly et al. 2006). We found that there is often a relationship between the extent of the study area and the grain, with fine grained datasets such as PRISM, SOGS-TOPS and Daymet -PRISM covering smaller areas. WorlClim is an exception to the rule since it covers a large extent but is fine grained. Furthermore, the temporal resolution also plays a role in the production workflow with daily products requiring more time and effort to produce. Particularly, for such product the CAI approach is often used to model the monthly and daily variation separately (Hosfra et al. 2008, Hunter and Meentemeyer 2008?). CAI frequently outperforms direct methods at the daily time scale and simplifies some of the processing (Hunter and Meentemeyer 2008).

 Providing detailed description of every study mentioned in table 4 would be

Impractical or not fitting for this review so we chose to summarize five studies in more detail: NEW01, PRISM, WORLDCLIM, Daymet, SOGS-TPS. These studies cover either the World or North America and considered together illustrate some of the challenges and strategies used in climate interpolation projects.

**NEW01**

Originally developed for use in hydrological applications under the auspices of the International Water Management Institute, the NEW01 dataset also supports wide range applications (New et al. 2002). It is a 10’ spatial resolution dataset at a monthly time scale which covers the 1961-1990 time period. It is global and includes some eight climate response variables: temperature, temperature range, sunshine, ground frost frequency, precipitation, wet day frequency, relative humidly and windspeed. This product is an improvement on the 30’ (0.5degree) product from New et al. (1999) using a larger database with more meteorological stations records. Station data were drawn from the World Meteorological Organization (WMO) and National Meteorological Agencies (NMAs) with the NMAs databases being the largest source of data. Stations were subjected to rigorous screening using NCDC quality control protocols with tests related to internal consistency of monthly mean, minimum and maximum as well as checks on the limits/ranged of values. Climatological normals were obtained from the WMO and merged with the CRU dataset.

Prediction was performed using the Thin Plate Spline method with three covariates: latitude, longitude and elevation. For computational reasons, the world was divided in 8 overlapping subregions where individual TPS models were run using the ANUSPLIN algorithm (Hutchison 1999?). Precipitation is predicted by fitting parameters to a Gamma distribution for each month. This approach allowed the interpolation of the mean and coefficient of variation (CV). CV was interpolated using latitude, longitude and mean precipitation. After fitting the data, stations with the largest the residuals were examined to detect outliers and errors. This permitted the detection of errors in horizontal location (latitude, longitude) and elevation. Errors in time were identified by recording values that did not follow seasonal patterns. NEW01’s accuracy and uncertainty were evaluated using a form of cross-validation method, the TPS Generalized Cross Validation (GCV) measurement (REF, Wood et al. 2006). Using the square root of GCV, RTGCV, patterns of high accuracy were uncovered in dense network areas (e.g. accuracy of 10% for precipitation) and lower accuracy values in high elevation areas with sparse station network (50% accuracy).

 The NEW01 product was also compared to its older version, NEW99 and to the PRISM product using the MAD (Mean Absolute Difference) when possible. In order to do so, MADs for NEW99 and NEW99 were calculated for individual stations as well as averaged within five degree grid cells. Results indicate that NEW99 and NEW01 products resolve similar mesoscale patterns but that NEW01 provides more details in many areas because of its higher resolution. Overall improvement in MAD is greater than 25% based on aggregated grid and; in 90% of the grid cell there were more than 66% of the station-grid MAD improvement. Similarly MAD in 5 degree grid cells for mean temperature also showed greater agreement with the station data with only %7 showing increases in MAD and 75% showing decreases in MAD smaller than 0.2 C. So, results demonstrate that there is an overall improvement in prediction of temperature and precipitation as compared to NEW99 product.

***WORLDCLIM***

WorldClim is a set of monthly global climate layers at a spatial resolution of 20 arc second that includes three response variables: minimum temperature, maximum temperature and precipitation (Hijmans et al. 2005). Similarly to NEW01, the response variables were interpolated using the Thin Plate Spline method using the ANUSPLIN software (Hutchinson 1994?) with three covariates: latitude, longitude and elevation. For computational and processing purpose, the World was divided in 13 overlapping regions of 15 degrees in which individual TPS models were run. Response variables were obtained from station databases from five main sources: the GHCN database (Peterson and Vose 1997), the WMO database (WMO1996), the FAOCLIM database (FAO, 2001), the CIAT database and regional databases. The GHCN database, produced by National Climatic Data Center, constituted the most important source of data with the highest quality assessment. All datasets were assessed for quality by undergoing a rigorous screening and non-GHCN Stations were only included if there were no GHCN stations available within 5km of predicted locations. This allowed for the use of the best quality information and the removal of duplicate stations from the various databases. Even though GHCN is assembled following a rigorous quality assessment, WorldClim added additional screening that lead to the removal of errors related to elevation units (meter versus feet in India) as well as the removal of inaccurate temperature measurements.

Compared to NEW01 product, WorldClim is derived from a denser network of station for all three variables, with 63% more stations for precipitation, 48% more stations for mean temperature and 26% more stations for temperature range. The total number of stations used varies from 47,554 stations for precipitation, to 24,542 station for mean temperature and 14,835 stations for minimum and maximum temperature. WorldClim is 400 times finer grained than NEW01 but this higher resolution does not translate directly into higher accuracy because its accuracy is dependent on many factors most notably the spatial configuration of the station data that may not capture the spatial variation at one kilometer resolution and the variability of features on the ground (Hijmans et al. 2005).

The accuracy of the product was evaluated using multiple procedures including grid averaging and temporal aggregation. Maps of accuracy were produced at a 2x2 degree resolution averaged over 12 months using both data partitioning and cross-validation. Spatial patterns show that cross-validation errors for precipitation were less than 10mm per month in most locations compared to 50mm/month for some areas using data partitioning. Temperature errors are low for cross-validation, about 0.3 using RMSE, while data partitioning exhibits mean errors in the 0 and 1C range.

***PRISM***

The first PRISM model was developed at the Oregon State University with the goal of interpolating precipitation in Northern Oregon, Willamette valley (Daly et al. 1994). An improved version of the PRISM model was released in 2002 by extending the model to the interpolation of minimum and maximum temperatures over a larger study area, the USA and other parts of North America (Daly et al. 2002). The current PRISM dataset is produced at a 2 to 4 km a spatial resolution over 1950-2000 time period. The data is distributed in geographic coordinates (Lat-Long) and in the Albers Equal Area projection.

 PRISM is a hybrid method that relies on two “master” equations to describe the trend components for the interpolations of precipitation and temperature: the elevation-precipitation and elevation-temperature (lapse rate) relationships respectively. The idea behind PRISM is to tailor these fundamental and general environmental correlation relationships (Oke 1978, Barry & Chorley 1987) at every grid point using local information. This is done by controlling the slope of the master equations through weighting of input observations based on station proximity and additional environmental covariates.

 PRISM therefore distinguishes itself from other models by not including additional covariates directly in the linear regression (master equation) but by parameterizing the slopes of the relationships using a station weighting dependent on a set of covariates. There are five weighting factors included in the regression: distance to prediction location, elevation, aspect, coastal proximity and boundary inversion layer. The distance factor captures proximity using a geographical weighting function to corresponding to an exponential decay kernel. Aspect is included to capture orographic influences such as leeward and windward effects and is reclassified in 8 categories called “topographic facet” using the scheme defined in Gibson et al. (1997). Coastal proximity captures the maritime effect in temperature (reduced range near the coast) as well as in precipitation (increased precipitation near the coast).

 Observation weights are determined by a mixture of expert knowledge, distance kernel weighting and empirical weighing of covariates with boundary default values. Thus, PRISM is a form of localized regression (Daly et al. 2006) that shares modeling features with Geographically Weighted Regression (Fotheringham et al. 2002) as well as with multiple linear regressions because its slopes vary at every location and are based on both neighboring observations and covariates values.

 PRISM prediction relies on a database of meteorological station assembled from very diverse set of sources originating in10 different databases but with most of the data items coming from COOP dataset. All meteorological stations underwent careful quality assessment. Stations with hourly measurements were aggregated to the daily time scale by taking the highest and lowest values for temperature or by summing the amount of precipitation for rainfall. Monthly values were only used if they were derived from a minimum of 85% valid observation following the completeness criterion from NDCD (NOAA-NCDC, 2003) and the World Meteorological Organization (WMO, 1989).

***Daymet***

Daymet is a climate dataset produced by the Numerical Terradynamic Group (NTSG) at University of Montana, School of Forestry. Initially developed as an input layers for plant growth models it is now used in a wide variety of application (Thornton et al. 1997 and Thornton et al. 2000). The product covers a large part of North America including continental USA, parts of Canada and Mexico. The dataset has a 1 km spatial resolution and is produced at daily and weekly time scales over 1980-2008 time period (Thornton et al. 2012). There are six response variables currently interpolated at the daily time scale: minimum and maximum temperature, precipitation, humidity/vapor pressure, shortwave radiation and snow water equivalent.

Daymet relies on the methodology developed by Thornton et al. (2000) for the production of radiation and Thornton et al. 1(997) for the interpolation of temperature and precipitation. Daymet uses a form of distance weighting with a Gaussian decay function modified to take into account the influence of elevation. The radius of the function varies depending on the density of the data i.e. it increases when there are fewer stations. The method, named the “Truncated Gaussian Function” (TGF), was first introduced in Thornton et al. (1997) for the interpolation of Northwestern USA region (an area of about 400,000km) over one year at daily time scale.

The TGF method performed well with MAE in the 0.7 and 1.2 range for annual maximum and minimum temperature and; with biases in the 0,1 and -0.1 respectively. MAE for annual precipitation was 19.3% in 134 mm. The meteorological station data used came from the Cooperative Day Summary from the NCDC and the Snowpack and Telemetry (SNOTEL) network from Natural Resources Conservation Service (NRCS). The Cooperative Day Summary was recently incorporated in the GHCN initiative (REF, Peterson et al. 1997?, Menne et al. 2012). Data from outside the USA were made available indirectly through GHCND and directly through Environment Canada and the Mexican Meteorological National Service (Servicio Meteorológico Nacional).

***SOGS-TOPS***

The Surface Observations Gridding System (SOGS) is a framework and methodology developed by Jolly et al. (2005) currently in use in the Terrestrial Observation and Prediction System (TOPS) from NASA (Nemani et al. 2009). TOPS products are part of a data and modeling software approach that integrates satellite, ground and air observations to derive operational products relevant to ecological monitoring and forecasting .TOPS performs Nowcasting and Forecasting of more than 30 variables such as gross primary productivity (GPP), evapotranspiration (Nemani and Running 2009). Within this framework, SOGS produces “Nowcast” of five daily meteorological variables using stations data: minimum temperature, maximum temperature, precipitation, vapor pressure deficit and radiation. SOGS-TOPS datasets are produced at three different levels of spatial resolution, one-kilometer resolution for the state of California, 8 kilometer resolution for the USA and at one degree resolution for the world.

SOGS utilizes three methods of interpolation: Truncated Gaussian Filter (TGF), Ordinary Kriging (OK) and IDW. TGF is a moving average method introduced by Thornton et al. (1997) which predicts new values by using a Gaussian weighting function truncated at a specified distance. In effect, TGF is form of IDW which uses a Gaussian weighting scheme rather than an inverse distance function to specify the weights. In addition to TGF prediction, IDW and OK predictions are performed. Rather than fitting a variogram for every daily time step, OK uses a fixed variogram for all days so that, according to the authors, it can be compared to the IDW and TGF methods (Jolly et al. 2005).

Temperature data are first reduced to the 1000 mb pressure surface (Barry and Chorley, 2008) with surface pressures estimated using elevation (Iribane and Godson, 1981). This phase is equivalent to detrending before interpolation. The final predicted surface is obtained by adding back the detrended surface to the interpolated surface at the end of the prediction process. Precipitation was also interpolated in two steps. First, the occurrence and non-occurrence of rain was modeled using a Boolean variable for all three methods. Second, the amount of rain was modeled using latitude, longitude and elevation as inputs.

Accuracy and validation was carried out for all three methods using cross-validation. Results for the 2004 prediction in California indicate that all three methods performed similarly with an MAE ranging from 1.6 to 1.9 and a ME (bias) ranging from 0.01 to 0.11 for temperatures. Precipitation predictions have an MAE ranging from 47 mm to 49 with a bias ranging from 27mm to 35mm.

1. **Accuracy and validation**

We summarize accuracy methods in 9 non-exclusive procedures that are often used in some combination in studies: reporting fit metric, data partitioning and holdout, cross-validation, grid aggregation, error uncertainty, error regression study, visualization/mapping of errors and residuals, product comparison, temporal aggregation.

*Table 5. Common procedures for validation*

|  |  |
| --- | --- |
| **Procedures** | **Studies** |
| 1.Reporting fit metric  | Everywhere: Jolly et al. 2005, Willmott and Matsuura 1995, New 2001, Attore et al. 2007, Daly et al. 2002 etc. |
| 3.Cross-validation | Jolly et al. 2005, Willmott and Matsuura 1995, New 1999 etc. |
| 3.Data partitioning/hold out | Price et al. 2000, Vicente-Serrano et al. 2003, Hijmans et al. 2005, Attore et al. 2007, McKenney et al. 2006. |
| 4.Grid aggregation | Hijmans et al. 2005, Hosfra et al. 2008, Haylock et al. 2008 |
| 5.Error uncertainty  | Hijmans et al. 2005, Daly et al. 2002,  |
| 6.Error regression study | Thornton et al. 1997, Price et al. 2000, Stahl et al. 2006. |
| 7.Visualization /mapping of errors/residuals | Hijmans et al. 2005, Jarvis and Stuart 2001 |
| 8. Product comparison | Hijmans et al. 2005, Daly et al. 2002, New et al. 2002,… |
| 9. Temporal aggregation | Hijmans et al. 2005 |

**4.1. Reporting fit metric**

 We found that this is the most commonly reported form of accuracy assessment with nearly all reviewed studies reporting some type of goodness of fit metrics (Willmott and Matsuura 1995, Hijmans et al. 2005, Hong et al. 2005, Daly et al. 2006, Bazgeer et al. 2012) such as the coefficient of determination (R2), the root means square error (RMSE) or the Akaike Information Criterion (AIC). The most widely used measured metrics are RMSE, mean absolute error (MAE) and mean error (ME/BIAS) (New et al. 2002, Hijmans et al. 2005, Hosfra et al. 2008, Fleming et al. 2008, and Bazgeer et al. 2012). ME is a measure of central tendency which correspond to the average error or bias from the true value while MAE and RMSE are measures of spread of errors. Since RMSE is affected by outliers and tends to weigh heavily extreme residuals, MAE is often used in a supplement or as an alternative (Willmott and Matsuura 2005, Willmott and Matsuura 2006). Additional method specific metrics are often reported such as the square root of the generalized cross-validation (RTGCV) for the TPS methods (New et al. 2002) or AIC for regressions.

**4.2. Cross-validation**

 Cross validation is performed by removing one station (the “validation” station) at a time from the station network and fitting the interpolation surface. The “error” or difference between the fitted value and the observed value at the validation station can be calculated every time a new prediction is obtained. When this procedure is repeated iteratively for each station in the form of the jackknife method, it allows the calculation of the overall errors or accuracy metrics such as MAE and ME (Legates and McCabe 1999, Daly et al. 2006, Willmott and Matsuura 2006). Cross-validation is one of the most ubiquitous validation procedures found in many interpolation studies (Willmott and Matsuura 1995, New et al. 1999, Gyalistras 2003, Jolly et al. 2005, Stahl et al. 2006, Daly et al. 2006, Hosfra et al. 2008, Bazgeer et al. 2012). There are several disadvantages to CV, first the influence of the removal of one station can be minimum because results are smoothed and local details reduced (Daly et al. 2006). Second, CV underestimates errors when stations are close together i.e. clustered spatially. Therefore, CV accuracy metric provides an indication of both the method performance and the adequacy of the network (Robeson 1993 and Willmott 1995) because cross-validation is influenced by the quality of the network (i.e. its density and spatial configuration). When errors are compared among different methods, the influence of the network can however, be considered somewhat constant, thus providing a mean to validate the methods (Willmott and Robeson 1995).

**4.3. Data partitioning and hold out**

 Data partitioning consists in dividing the observations, meteorological stations, into training and testing (validation) data sets by selecting a certain proportion of hold-out (REF). The training dataset is used to fit the model while the testing dataset allows the assessment of the model using independent observations that were not used in the model fitting (Price et al. 2000, Vicente-Serrano et al. 2003, Hijmans et al. 2005 and Attore et al. 2007). The method of selection of hold-out varies from fixed numbers of stations (Price et al. 2000, McKinney et al. 2006) to fixed proportions (Hijmans et al. 2005, Attore et al. 2007) using random or stratified sampling with co-variates variables such as elevation (McKenney et al. 2006). For instance, Hutchinson 2004 SPLINA method includes options to select the validation dataset using a form of stratified random sampling with elevation. Other refinements include multiple partitioning to average sampling effects. An example is found Attore et al. (2007) where multiple sampling is used by repeating the selection of training and validation sets 20 times and by calculating the validation metric on each validation set. Since the density of the network is the most important factor in the accuracy, holding out data result in a decrease in accuracy (New et al. 2001, Stahl et al. 2006, Hutchinson et al. 1995) so that in some instances no partitioning is performed to validate the model.

 **4.4. Grid-averaging**

 Accuracy assessments reported at stations are point-based and do not allow easy assessment of the spatial patterns and large scale performance of models. In consequence, grid or box-averaging is often used as way to display the spatial accuracy in the study area. This is usually done by aggregating data to lower resolutions into cells to study the RMSE and MAE metrics derived from cross validation or data partitioning (Hijmans et al. 2005, Hosfra et al. 2008). Similarly, aggregation at lower resolution is often used in cross-products comparisons. For example, NEW99 and NEW01 were evaluated by aggregation of both products at 5 and 0.5 degrees and through the calculation of MAEs at individual stations as well as averages in each grid cell. Using such procedure, the authors demonstrated that NEW99 and NEW01 were comparable in their ability to resolve similar mesoscale patterns but that NEW01 showed better resolution in many areas because of its higher spatial resolution (New et al. 2001).

**4.5. Error uncertainty**

 Uncertainty relates to the interval of precision of the predicted values for temperatures, precipitations or other climate variables at each location in the study area. It is provided in the form of standard error estimates or interval of confidence in linear regression models and can be obtained for many other interpolation methods such as Kriging in the form of the variance of prediction (Waller and Gotway 2004). For the GAM and TPS methods, standard errors estimates are provided for every prediction using the Bayesian framework (Woods et al. 2006, McKenney et al. 2006). While reporting of uncertainty is possible many instances, it is not always done because it may be computationally burdensome (McKenney et al. 2006) or may increase substantially the amount of data to produce and distribute (REF).

**4.6. Error regression study**

 Prediction errors are closely related to many modeling factors namely: the complexity of elevation and its variability of in the area (Price et al. 2000, Willmott and Robeson, 1995, Thornton et al. 1997), the spatial density (Willmott et al. 1995) and the configuration of network (Willmott et al 1991). For instance, Hijmans et al (2005) found low accuracy in areas where the station network is spatially sparse and where there are large variations in elevation. This problem is further compounded by the topographic bias present in the station network (Briggs et al. 1996). For instance, Hijmans et al. (2005) reported that most stations were found in low lying areas in high latitude regions and at high elevation in the Tropics and low latitudes regions. This relationship can be portrayed using a measure of error such as MAE (Thornton et al. 1997:230) but there exists no solutions to solve for such bias.

**4.7. Visualization and mapping of uncertainty and residuals**

 Visualization or mapping of residuals and uncertainty provide a very effective method to detect patterns and to uncover potential causes. For instance, using visualization, the WorlClim team reported that predictions and residuals in Poland and UK were unusually high and low respectively. This lead to the uncovering of erroneous recordings in temperatures related to the inclusion of extremes in the mean temperature. Similarly, in cases where residuals are obtained in many temporal steps, Jarvis and Stuart (2000) used mapping of average residuals to identify recurrent patterns over one year.

**4.8. Product comparison**

 Product comparison provides a useful way to compare relative accuracy and spatial differences in prediction can lead to further understanding of prediction errors. Data product comparison may also be problematic however, because 1) Products may not cover the same extent so a common area must be chosen and the choice may not reflect the overall accuracy, 2) products may not have been produced at the same spatial resolution which means that coarsening may be necessary 3) Products may not have been produced at the same temporal resolution. Aggregation in time may be necessary. Regardless much can be gained by comparing across products. For instance, Hijmans et al. (2005) found that the largest differences between the WorldClim and NEW01 products were in areas where the meteorological station network density was low thereby highlighting again the primary importance of the station network in the determination product quality.

**4.9. Temporal aggregation and assessment**

Datasets are generally produced as time series and variation in temporal accuracy can be assessed for every time step. Accuracy at every time step can be visualized through time series of errors (Willmott and Robeson 1995) or an averaged through space (Hijmans et al. 2005). Analysis of temporal patterns in accuracy for PRISM revealed that annual errors were lower than monthly prediction errors. The PRISM group suggests that the lower performance at the monthly time scale may be due to higher temporal variability and weaker relationships in the precipitation-elevation master equation. In addition, the authors found that there was a temporal pattern in the cross-validation with higher MAE values during the summer time period. The authors relate this pattern to the smaller rainfall amounts and the high and complex spatial variability during summer months (Daly et al. 2004:155). Temporal aggregation is also required in cases where there is a need to compare different temporal resolutions such daily and monthly time predictions. Monthly predictions can be obtained by averaging of daily predictions or by prediction at monthly time scale using the same methodology.

**4.10. Validation overview, issues and possible solutions**

 There is no single accuracy method that can provide an overall accuracy assessment without suffering from some disadvantage or drawback (Daly et al. 2006). For instance, data partitioning data holdout may result in a decrease in accuracy (New et al. 2001, Stahl et al. 2006, Hutchinson et al. 1995) because the density of the network is the most important factor in determining the accuracy of the predictions. Spatial autocorrelation in the dataset will also reduce the effective number of observations and influence accuracy measures (e.g. MAE). Accuracy may also depend on the coverage of validation sample with areas lacking validation stations incorrectly portrayed in terms of accuracy. This entails that random sampling may not be appropriate in many cases because of the sparsity and clustering of stations (Attore et al. 2007, Hutchinson et al. 1995). Designing a truly representative sample is however difficult or impossible because the network itself is often biased (Briggs et al. 1996, Hengl et al. 2009). Cross-validation suffers from similar problems with accuracy values influenced by both the station network configuration and the prediction technique.

In the next sections, we suggest and comment on a few possible avenues to deal with accuracy issues highlighted previously.

**Issue 1: Data partitioning and hold out**

 While data partitioning is fraught with problems, an independent dataset is one of the only ways to avoid overfitting and to provide information on the predictive capacity of the model at unknown locations (Daly et al. 2006). Thus, possible solutions to the problems highlighted include evaluating the effect of partitioning by varying hold out proportion and by providing an average accuracy (Figure 7). This assessment should include multiple hold out for each proportion (Attore et al. 2007) so that the effect of varying spatial configuration of the network is somewhat averaged out in the accuracy. The effect of permuting the training and testing datasets may be also estimated by plotting the variance along with the average accuracy metric. Other options to assess the effect of holding out can include providing a graph of the difference between accuracy of the validation and testing data set. The reasoning is that the accuracy may be over-evaluated at unknown locations due to overtraining/overfitting (Figure 7).

 

*Figure 7. Multiple partitioning and accuracy assessment*

**Issue 2: Spatial autocorrelation and accuracy**

 Given that observations are intrinsically spatial, there will frequently be spatial autocorrelations in validation and training datasets. Possible solution to deal with such problem may include stratified or systematic sampling using range based limit i.e. taking into account only observations beyond the range of spatial autocorrelation as determined from the variogram dataset. This may be difficult in practice and may fail to recognize the fact that most methods exploit autocorrelation itself to predict new values. Thus, it may be more appropriate to provide a report of how uncertainty and accuracy metrics vary in terms of the spatial configuration (i.e. map clusters) and in terms of distance to station point (Figure 8).



*Figure 8. Autocorrelation and accuracy*

**Issue 3: Comparison between products**

This approach is very common in the literature. It raises important issues because similarities and/or dissimilarities among product do not imply accurate results. Furthermore, visual comparison is often carried out without a metric. This practice can be misleading in some cases because visual poor tool of map comparison (Pontius et al.?). Nevertheless, comparison between earlier products is useful for users and may provide insight in possible improvement avenues. Thus a possible solution to inter map comparison may be to generate and compare products to an additional more “neutral” reference surface derived from the station network (Figure 9). This is the approach taken in New et al. 2002 where common error grids are derived by calculating the mean of error metrics within each grid box from the station network. In this manner, decreases in average grid box errors are detected between NEW01 and NEW99 products indicating improvements in the interpolation.



*Figure 9. Product comparison using the station network as reference.*

1. **Conclusion**

 This paper provides a review of interpolation methods, products and accuracy assessment procedures for the prediction of climate variables such as temperature and precipitation on gridded surfaces. In examining the literature, we have found five major points:

1. There is no single best method that can be applied in all situations and accuracy is largely dependent on case by case basis with the most important factors being the density of the station network and the variability of the meteorological surfaces across the study area.
2. Covariates are helpful at the scale of 100km to estimate fine grained variation in temperature and precipitation. Elevation is the most often used covariates but other covariates are often included such as coastal proximity and aspect.
3. There are many overlaps between methods in a field where researchers may come from climatology, ecology or geography. Many methods are related formally, for instance, TPS with latitude and longitude is a form of Universal Kriging with geographical coordinates and; GAM relates to splines but also to multiple regression. Methods can often be understood as a hybrid of geostatistical and multiple regression methods. Methods can also be understood as attempting to predict values by focusing on the estimation of the global variance, local variance or both.
4. Accuracy assessment is difficult because of bias in the observation, autocorrelation and scarcity of information. Since there is no best practice, multiple ways of reporting accuracy metrics through maps, graphic and tables is generally done. When comparing between products a neutral grid surface comparison can be built from average accuracy metrics at a common grid resolution and accuracy and uncertainty can be described in terms of the various input covariates.
5. Cleaning of station network or QC is often needed even when it comes from trusted source such as GHCN. Such cleaning is paramount and may improve results because the station network is the most important factor in the accuracy of the product.

 The production of gridded surface is a complex enterprise that requires tackling many challenges related to dealing with scarce data, assembling databases from multiple resources, automation of workflows in large computatively intensive production and to understanding of statistical techniques and climate and weather related phenomena. Predictions of climate variables are crucial for a host of studies involved in a wide range of research related to the global environment and human societies (Kabat et al. 2004). Practical applications include understanding of crop yields (Kabat et al. 2004), modeling of species ranges (Peterson et al. 2008) and climate change studies (IPCC2007, REF).

1. **References**

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res., 111, D05109, doi:10.1029/2005JD006290.

Apaydin, H., F. K Sonmez, and Y. E Yildirim. “Spatial Interpolation Techniques for Climate Data in the GAP Region in Turkey.” *Climate Research* 28, no. 1 (2004): 31–40.

Attorre, F., et al. (2007), Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale, Int. J. Climatol., 27,1825–1843, doi:10.1002/joc.1495.

Bartels, R., Beatty, J. and B. Barsky. 1987. An Introduction to Splines for Use in Computer Graphics and Geometric Modeling, M. Kaufmann Publishers, pp.476.

Bazgeer S., Oskuee E.A., Hagigat M., Darban Astane A.R. (2012). Assessing the performance of spatial interpolation methods for mapping precipitation data: a case study in Fars Province, Iran. Trends in applied Sciences Research, 7(6),467-475.

Boer E.P.J. de Beurs K.M, Hartkamp A.D. 2001. Kriging and thin plate splines for mapping climate variables. JAG, Volume 3, issue 2.

Bolstad, P.V., L. Swift, F. Collins, and J. Regniere, 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains, Agriculture and Forest Meteorology. 91:167-176.

Booth TH, Jones PG (1998) Identifying climatically suitable areas for growing particular trees in Latin America. Forest

Ecology and Management 108:167–173

Bostan P. A., Akyurek Z. 2007. Exploring the mean annual precipitation and temperature values over Turkey by using environmental variables. Proceedings of the ISPRS conference, vol XXXVI.4/W45.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, p. 333.

Briggs PR, Cogley JG (1996) Topographic bias in mesoscale precipitation networks. J Climate 9:205–218

Brunsdon, C., Fotheringham, A. S., Charlton, M. E. 1996. Geographically weighted regression: a method for exploring spatial non-stationarity. Geographical Analysis, 28: 281-298.

Brunsdon, C., McClatchey J., Unwin D.J. 2001. Spatial variations in the average rainfall-altitude relationship in great Britain: an approach using geographically weighted regression. International Journal of Climatology. 21:455-466. DOI: 10.1002/joc.614

Caesar, J., et al. (2006), Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded dataset, J. Geophys. Res., 111, D05101, doi:10.1029/2005JD006280.

Changnon SA, Kunkel KE (1999) Rapidly expanding uses of climate data and information in agriculture and water resources: causes and characteristics of new applications.Bull Amer Meteor Soc 80:821–830

Choi, J., Chung, U., Yun, J.I., 2003. Urban effect correction to improve accuracy of spatially interpolated temperature estimates in Korea.J. Appl. Meteorol. 42 (12), 1711–1719.

Christidis, N., Stot,P. A., Brown, S.J. The role of human activiey in the recent warming of extrememely Warm daytime temperatures.

Chung, U., Yun, J.I., 2004. Solar irradiance—corrected spatial inter-polation of hourly air temperature in complex terrain. Agric.Forest Meteorol. 126, 129–139.

Cooley, Daniel, Douglas Nychka, and Philippe Naveau. “Bayesian Spatial Modeling of Extreme Precipitation Return Levels.” *Journal of the American Statistical Association* 102 (September 2007): 824–840.

Coulibaly M., Becker S. 2009. Spatial interpolation of annual precipitation in South Africa-comparison and evaluation of methods. Water International, 32:3,494-502.

Couralt, D., Monestiez, P., 1999. Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France. Int. J. Climatol. 19, 365–378.

Cramer W, Fischer A (1996) Data requirements for global terrestrial ecosystem modelling. In: Walker B, Steffen W (eds)

Global change and terrestrial ecosystems. Cambridge University Press, Cambridge, MA, p 530–565

Cressie, N.A., 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., New York, p. 900.

Cressman, George P. “An Operational Objective Analysis System.” *Monthly Weather Review* 87, no. 10 (October 1959): 367–374.

Cressie N.and Hawkings D.M. (1980). Robust estimation of variogram, I: Math. Geology, 12, 2, 115-125.

Dai A, Fung IY, Del Genio AD. 1997. Surface observed global land precipitation variations during 1900–88. *Journal of Climate* **10**: 2943–2962.

Daly C, Johnson GL. 1999. PRISM spatial climate layers: their development and use. short course on topics in applied climatology. *79thAnnual Meeting of the American Meteorological Society*, American Meteorological Society: Dallas, TX, 10–15 January, available at:http://www.ocs.orst.edu/prism/prisguid.pdf.

Daly, C., et al. (2002), A knowledge-based approach to the statistical mapping of climate, Clim. Res., 22, 99–113.

Daly, C. (2006), Guidelines for assessing the suitability of spatial climate data sets, Int. J. Climatol., 26, 707– 721, doi:10.1002/joc.1322.

Daly, C., M. Halbleib, J. I Smith, W. P Gibson, M. K Doggett, G. H Taylor, J. Curtis, and P. P Pasteris. 2008. “Physiographically Sensitive Mapping of Climatological Temperature and Precipitation Across the Conterminous United States.” *International Journal of Climatology* 28 (15): 2031–2064.

Diggle, P. J, and P. J Ribeiro. *Model-based Geostatistics*. Vol. 846. Springer New York, 2007.

Diggle, P. J, and P. J Ribeiro Jr. “Bayesian Inference in Gaussian Model-based Geostatistics.” *Geographical and Environmental Modelling* 6, no. 2 (2002): 129–146.

Di Luzio 2008. Constructing Retrsopective Griddedd Daily Precipitation and Temperature datasets for the conterminous United States. Journal of Applied Meteorology and Climatology, 47, 2, 457-497.

Dinku, T., P. Ceccato, E. Grover‐Kopec, M. Lemma, S. J. Connor, and C. F. Ropelewski. “Validation of Satellite Rainfall Products over East Africa’s Complex Topography.” *International Journal of Remote Sensing* 28, no. 7 (2007): 1503–1526.

Dingman SL, Seely-Reynolds DM, Reynolds RC III (1988) Application of kriging to estimate mean annual precipita- tion in a region of orographic influence. Wat Resour Bull 24:29 – 339

Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim Res 8:1–20

Dubrule, O. (1983) Two methods with different objectives: Splines and Kriging. *Mathematical* *Geology*, 15, pp. 245-257.

Dubrule O. (1984). Comparing splines and Kriging. Computers & Geosciences, 10,2-3: 327-338.

Duchon. (1975). Fonctions splines du type plaque mince en dimension 2. Seminaire d’analyse numerique. Universite de Grendoble, no. 231, 13p.

Durre, Imke, Matthew J. Menne, and Russell S. Vose. “Strategies for Evaluating Quality Assurance Procedures.” *Journal of Applied Meteorology and Climatology* 47, no. 6 (June 2008): 1785–1791.

FAO. (2002) Food and Agriculture Organization of the United Nations - FAOCLIM Ver.2 database- Environment and Natural Resources Service SDRN, Publication No. 9 in the SDRN workingpaper

Feng, S., et al. (2004), Quality control of daily meteorological data in China, 1951 – 2001: A new dataset, Int. J. Climatol., 24, 853 – 870,doi:10.1002/joc.1047.

Fleming M.D., Chapin III F.S, Cramer W., Hufford G.L., Serreze M.C. Geographic patterns an dynamics of Alaskan climate interpolated from a sparse station record. Global Change Biology, 6, 49-58.

Flint L.E., Flint A.L. 2012. Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. Ecological Processes. 1:1.

Foody GM (2003) Geographical weighting as a further refinement to regression modeling: an example focused on the NDVI-rainfall relationship. Remote Sens Environ 88:283–293

Foody GM (2004) Spatial nonstationarity and scale-depen- dency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. Global Ecol Biogeogr 13:315–320

Fotheringham AS, Brunsdon C, Charlton M (2002) Geographical weighted regression: the analysis of spatially relationships. Wiley, Chichester.

Funk C, Michaelsen J, Verdin J, Artan G, Husak G, Senay G, Gadain H, Magadazire T. 2002. The collaborative historical African rainfall model: description and evaluation. *International Journal of Climatology* **1**: 47–66.

Garen DC, Johnson GL, Hanson CL (1994) Mean areal pre- cipitation for daily hydrologic modeling in mountainous terrain. Wat Resour Bull 30:481–491

Gerstner, E.-M., and G. Heinemann (2008), Real-time areal precipitation determination from radar by means of statistical objective analysis,J. Hydrol., 352, 296– 308, doi:10.1016/j.jhydrol.2008.01.016.

Goodale C, Aber J, Ollinger S. 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Research* **10**: 35–49.

Goovaerts, P. (2000), Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113– 129.

Groot, E. v. d., and S. Orlandi (2003), Technical description of interpolation and processing of meteorological data in CGMS, 16 pp, Internal Report,Space Applications Institute, Joint Research Centre of the EU, ISPRA,Italy.

Guan, B. T, H. W Hsu, T. H Wey, and L. S Tsao. (2009). “Modeling Monthly Mean Temperatures for the Mountain Regions of Taiwan by Generalized Additive Models.” *Agricultural and Forest Meteorology* 149, 2, 281–290.

Gyalistras D. (2003). Development and validation of a high-resolution monthly gridded temperature and precipitation data set for

Switzerland (1951–2000). *Climate Research* **25**: 55–83.

Hamlet AF, Lettenmaier DP. 2005. Production of temporally consistent gridded precipitation and temperature fields for the continental U.S. *Journal of Hydrometeorology* **6**: 330–336.

Handcock, M. S., et al. (1994), Kriging and splines—An empirical-comparison of their predictive performance in some applications—Comment,J. Am. Stat. Assoc., 89, 401–403.

Hart, Q J, M. Brugnach, B. Temesgen, C. Rueda, S L Ustin, and K. Frame. “Daily Reference Evapotranspiration for California Using Satellite Imagery and Weather Station Measurement Interpolation.” *Civil Engineering and Environmental Systems* 26, no. 1 (2009): 19–33.

Hartkamp, A.D., De Beurs, K., Stein, A., White, J.W., 1999. Interpolation Techniques for Climate Variables. CIMMYT, Mexico, D.F.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. 1st ed., 335pp . Chapman & Hall/CRC.

Hay S.I., Lennon J.J. (1999). Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. Trop Med Int Health; 4(1): 58–71.

Haylock, M., et al. (2008), A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950-2006. J. Geophys. Res.,113, D20119, doi:10.1029/2008JD010201.

Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables. JRC Scientific and Technichal Reports. Office for Official Publication of the European Communities, Luxembourg.

Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-

Hevesi JA, Istok JD, Flint AL. 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. 1. Structural

analysis. *Journal of Applied Meteorology* **31**: 661–676...

Hewitson, B. C., and R. G. Crane (2002), Self-organizing maps: Applications to synoptic climatology, Clim. Res., 22, 13– 26.

Hewitson, B. C., and R. G. Crane (2005), Gridded area-averaged daily precipitation via conditional interpolation, J. Clim., 18, 41–57.

Hiemstra P.H., Pebesma E.J., Twenhofel C.J.W., Heuveling G.B.M. (2008). Automatic real-time interpolation of radiation hazards: prototype and system architecture considerations.Intertional Journal of Spatial Data Infrastructures Research,3,58-72.

Hijmans, R J, S E Cameron, J L Parra, P G Jones, A. Jarvis, and Others. “Very High Resolution Interpolated Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 25, no. 15 (2005): 1965–1978.

Hofierka J., Parajka J., Mitasova H., Mitas L (2002). Multivariate interpolation of precipitation using regularized spline with tension.Transactions in GIS, 2002, 6(2): 135±150

Hong, Y., Nix, H.A., Hutchinson, M.F., Booth, T.H. (2005). Spatial Interpolation of monthly mean climate data for China. International Journal of Climatology. 25: 1369-1379.

Huffman, G. J., et al. (1995), Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information, J. Clim., 8, 1284–1295.

Huang, B., B. Wu, and M. Barry, *Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices.* International Journal of GeographicalInformation Science, 2010. **24**(3): p. 383 - 401.

Hunter R.D. and Meentemeyer R.K. (2005). Climatology aided mapping of daily interpolation and temperature. Journal of Applied Meteorology, 44,10, 1501-1510.

Hutchinson, M.F., Gessler, P.E., 1994. Splines — more than just a smooth interpolator. Geoderma 62, 45–67.

Hutchinson, M. F. (1993), On thin plate splines and kriging, in Computing and Science in Statistics, edited by M. E. Tarter and M. D. Lock, pp. 55 – 62, Interface Foundation of North America, University of California, Berkeley, Calif.

Hutchinson MF. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems* **9**: 385–403.

Hutchinson, M. F. (2004). Anusplin Version v4.3. Canberra, Australia, Centre for Resource and Environmental Studies. The Australian National University.

Jackson, S. T., J. L. Betancourt, et al. (2009). "Ecology and the ratchet of events: Climate variability, niche dimensions, and species distributions." Proceedings of the National Academy of Sciences **106**: 19685-19692.

Jarvis CH, Stuart N. 2001. A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II:the interaction between the number of guiding variables and the type of interpolation method. *Journal of Applied Meteorology* **40**:1075–1084.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, p. 561.

Johansson, M.A., and G.E. Glass. “High-resolution Spatiotemporal Weather Models for Climate Studies.” *International Journal of Health Geographics* 7, no. 1 (2008): 52.

Johnson L., Pierce L., Michaelis A., Scholasch T., Nemani R. (2006). Remote Sensning and Water balance modeling in California drip-irrigated Vineyards. Proceedins, ASCE World Environmental & Water Resources Congress, Omaha NE, 21-25 May 2006.

Joubert, Sarah Joan. “High-resolution Climate Variable Generation for the Western Cape.” Thesis, 2007. <http://scholar.sun.ac.za/handle/10019.1/2452>.

Kabat P., Claussen M, Dirmeyer PA, Gash JH, Deguenni LB, Myebeck M., Pielke RA, Vorosmarty CJ, Hutjes R WA,

Lutkemeier S. 2004. Vegetation, Water, Humans and the climate. A new perspective on an Interactive System. Global Change. IGBP series. Springer Verlag.

Klok, L., and A. M. G. Klein Tank (2008), Updated and extended European dataset of daily observations, Int. J. Climatol., in press.

Krige, D. G. (1951), A statistical approach to some basic mine valuation problems on the Witwatersrand, J. Chem. Metall. Min. Soc. S. Afr., 52, 119 – 139.

Krige, D. G. (1966), Two-dimensional weighted moving average trend surfaces for ore evaluation, J. S. Afr. Inst. Min. Metall., 66, 13–38.

Leemans R, Cramer WP. 1991. *The IIASA Database for mean Monthly Values of Temperature, Precipitation and Cloudiness of a GlobalTerrestrial Grid* . Research Report RR-91-18 November 1991. International Institute of Applied Systems Analyses: Laxenburg, Austria.

Legates DR, Willmott CJ. 1990. Mean seasonal and spatial variability of global surface air temperature. *Theoretical and Applied*

*Climatology* **41**: 11–21.

Legates DR, McCabe GJ. 1999. Evaluating the use of “goodness of fit” measures in hydrologic and hydroclimatic model validation.*Water Resources Research* **35**: 233–241.

Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673.

Lennon J.J., Turner J.R.G. (1995). Predicting the spatial distribution of climate: temperature in Great Britain. Journal of Animal Ecology, 64, 3:370-392.

Lin, G., Chen, L., 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. Journal of Hydrology 288,288–298.

Lloyd, C. D. (2010) Nonstationary models for exploring and mapping monthly precipitation in the United Kingdom International Journal of Climatology.30: 390–405.doi: 10.1002/joc.1892

Logan, J.A., Powell, J.A., 2001. Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). Am. Entomol.47, 160–173.

Lu, G.Y. and D.W. Wong. 2008. An adaptive inverse-distance weighting spatial interpolation technique. Comp. Geosci., 34: 1044-1055.

Matheron G (1969) Le krigeage universel. Centre de Morphologie, Fontainebleau.

McKenney D.W., Pedlar J.H., Papadopol P. and M.F. Hutchinson 2006. The development of 1901-2000 historical monthly climate models for Canada and the United States. Agricultural and Forest Meteorology 138, 69-81.

Menne, M., I. Durre, R. Vose, B. Gleason, and T. Houston, 2012: An Overview of the Global Historical Climatology Network-Daily Database. J. Atmos. Oceanic Technol. doi:10.1175/JTECH-D-11-00103.1.

Mitas, L. and Mitasova, H. (1999). Spatial Interpolation. In: Longley, P. Goodchild, M.F., Maguire, D. and Rhind, D. (eds.) Geographical Information Systems. 2nd Edition. Vol. 1: Principles and Technical Issues: 481-492.

Mitchell TD, Jones PD. 2005. An improved method of constructing a database of monthly climate observations and associated high resolution grids. *International Journal of Climatology* **25**: 693–712.

Mokhov, II, and MG Akperov. “Tropospheric Lapse Rate and Its Relation to Surface Temperature from Reanalysis Data.” *Izvestiya Atmospheric and Oceanic Physics* 42, no. 4 (2006): 430–438.

Moral, F.J. (2010). Comparison of different geostatistical approaches to map climate variables: application to precipitation. International Journal of Climatology. 30: 620–631.

Murphy, A. H., and E. S. Epstein (1989), Skill scores and correlation coefficients in model verification, Mon. Weather Rev., 117, 572–581.

Myers, D.E. (1994). Spatial interpolation: an overview. Geoderman, 62, 17-28.

Myers, D.E. (1999). Smoothing and Interpolation with radial basis functions. In Boundary Element Technology XIII, C.S. Chen, C.A. Brebbia and D.W. Pepper (eds), pp: 365-376, WIT Presss, Southhampton.

Nemani R., Hashimoto H., Votava P., Melton F., Wang W., Michaelis A., Mutch L., Milesi C., Hiatt S., White M. (2009). Monitoring and forecasting ecosystem dynamics using the Terrestrial Observation and Prediction System (TOPS). Remote Sensing of the Environment 113, 1497-1509.

Neteler M. 2010. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST Data. Remote Sensing, 2, 333-351.

New, M. (1999), Uncertainty in representing observed climate, in Representing Uncertainty in Climate Change Scenarios and Impact Studies, edited by T. Carter et al., pp. 59–66, Climate Research Unit, Norwich, UK.

New, M., et al. (2000), Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate, J. Clim., 13, 2217–2238.

New, M., et al. (2001), Precipitation measurements and trends in the twentieth century, Int. J. Climatol., 21, 1899–1922, doi:10.1002/joc.680.

New M, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* **21**: 1–15.

Newlands, Nathaniel K, Andrew Davidson, Allan Howard, and Harvey Hill. “Validation and Inter‐comparison of Three Methodologies for Interpolating Daily Precipitation and Temperature Across Canada.” *Environmetrics* 22, no. 2 (March 1, 2011): 205–223.

Pandey GR, Cayan D, Dettinger MD, Georgakakos KP. 1999. A hybrid orographic plus statistical model for downscaling daily

precipitation in northern California. *Journal of Hydrometeorology* **1**: 491–506.

Parra J.L., Graham CC, Freile JF. 2004. Evaluating alternative data sets for ecological niche models of birds in the Andes. *Ecography* **27**: 350–360.

Perry, M., Hollis D., (2005) “The generation of monthly gridded datasets for a range of climatic variables over the UK.” *International Journal of Climatology* 25, 1041–1054.

Peterson C.T., Basist A.N., Williams C.N., Grody N.C. 2000. A blended satellite-In Situ Near-Global Surface Temperature Dataset. Bulletin of the American Meteorological Society.

Peterson, AT, and Y. Nakazawa. “Environmental Data Sets Matter in Ecological Niche Modelling: An Example with Solenopsis Invicta and Solenopsis Richteri.” *Global Ecology and Biogeography* 17, no. 1 (2008): 135.

Phillips DL, Dolph J, Marks D (1992) A comparison of geostatistical procedures for spatial analysis of precipitation inmountainous terrain. Agric For Meteorol 58:119–141

Plantico MS, Goss LA, Daly C, Taylor G. 2000. A new U.S. climate atlas. *Proceedings of 12th AMS Conference on Applied Climatology*.American Meteorological Society: Asheville, NC; 247–248, May 8–11, http://www.ncdc.noaa.gov/oa/about/cdrom/climatls2/info/atla-sad.html.

Price, D. T., et al. (2000), A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data, Agric. For. Meteorol., 10, 81–94.

Regniere, J. and P. Bolstad, 1994. Statistical simulation of daily air temperature patterns in eastern North America to forecast events in insect pest management, Environmental Entomology, 23: 1368-1380.

Renka RJ. 1984. Interpolation of data on the surface of a sphere. *ACM Transactions on Mathematical Software* **10**: 417–436.

Riccio, A. “A Bayesian Approach for the Spatiotemporal Interpolation of Environmental Data.” *Monthly Weather Review* 133, no. 2 (2005): 430–440.

Riccio, A., G. Barone, E. Chianese, and G. Giunta. “A Hierarchical Bayesian Approach to the Spatio-temporal Modeling of Air Quality Data.” *Atmospheric Environment* 40, no. 3 (2006): 554–566.

Schuurmans, J. M., M. F. P. Bierkens, E. J. Pebesma, and R. Uijlenhoet. “Automatic Prediction of High-Resolution Daily Rainfall Fields for Multiple Extents: The Potential of Operational Radar.” *Journal of Hydrometeorology* 8, no. 6 (December 2007): 1204–1224.

Shepard D (1968) A two-dimensional interpolation function for irregularly spaced data. In: Proc 23rd National Conf

ACM. American Meteorological Society, p 517–523

Simolo, C., M. Brunetti, M. Maugeri, and T. Nanni. “Improving Estimation of Missing Values in Daily Precipitation Series by a Probability Density Function-preserving Approach.” *International Journal of Climatology* 9999, no. 9999 (2009).

Sluiter R. 2009. Interpolation Methods for climate data. Intern rapport.

Snell S.E., Gopal S., Kaufmann R.K. (2000). Spatial interpolation of surface temperatures using artificial neural networks: evaluating their use for downscaling GCMs. Journal of Climate, 13,886-895.

Spencer, R. W. and Christy, J. R. **1990.** ‘Precise monitoring of global temperature trends from satellites’, ***Science,* 247, 1558-1562.**

Stahl, K., Moore, R.D., Floyer, J.A., Asplin, M.G., McKendry, I.G., 2006a. Comparison ofapproaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agricultural and Forest Meteorology 139, 224–236.

Stahl, K., Moore, R.D., McKendry, I.G., 2006b. Climatology of winter cold spells in relation to mountain pine beetle mortality in British Columbia, Canada. Climate Research 32, 13–23.

Stein M.L. (1991). A kernel approximation to the Kriging predictor of a spatial process. Ann. Inst. Statist. Math., 45, 1, 61-75.

Tait, Andrew, Roddy Henderson, Richard Turner, and Xiaogu Zheng. “Thin Plate Smoothing Spline Interpolation of Daily Rainfall for New Zealand Using a Climatological Rainfall Surface.” *International Journal of Climatology* 26, no. 14 (November 30, 2006): 2097–2115.

Thornton, P.E., Running, S.W., White, M.A., (1997). Generating surfaces of daily meteorological variables over large regions of

complex terrain. J. Hydrol. 190, 214±251.

Thornton, P.E., Hasenauer H., White M.A. (2000). Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. Agricultural and Forest Meteorology, 104, 255-271.

Thornton, PE, MM Thornton, BW Mayer, N Wilhelmi, Y Wei, RB Cook (2012). Daymet: Daily surface weather on a 1 km grid for North America,1980 - 2008. Acquired online (http://daymet.ornl.gov/) on [insert acquisition date here: DD/MM/YYYY] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi:10.3334/ORNLDAAC/Daymet\_V2.

Trnka, M., J. E Olesen, K. C Kersebaum, a.O. Skjelv\aa g, J. Eitzinger, B. Seguin, P. Peltonen-Sainio, et al. “Agroclimatic Conditions in Europe Under Climate Change.” *Global Change Biology* 17, no. 7 (July 2011): 2298–2318.

Tveito, O.E., Wegehenkel M., Van der Wel F., Dobesch H. (2006). The use of Geographic Information Systems in climatology and meteorology.- Final Report COST Action 719.

Venier, L.A., McKenney, D.W., Wang, Y., McKee, J., 1998a. Models of large-scale breeding-bird distribution as a function

of macroclimate in Ontario, Canada. J. Biogeography 26, 315–328.

Venier, L.A., Hopkin, A.A., McKenney, D.W., Wang, Y., 1998b.A spatial climate-determined risk rating for Scleroderris

disease of pines in Ontario. Can. J. For. Res. 28, 1398–1404.

Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Climate Research 24, 161–180.

Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science 41, 473–490.

von Storch, H., Zorita, E. and Cubasch, U. 1993: Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. Journal of Climate 6, 1161-71.

Wahba G., and Wendelberger J. 1980. Some new mathematical methods for variational objective analysis using splines and cross validation.

Wahba, G. 1990. Spline models for observational data. CBMS-NSF. Regional Conference Series in Mathematics, 59. SIAM, Philadelphia. pp169.

Waller L.A. and Gotway C.A. (2004). Applied spatial statistics for public health data. Wiley, Hoboken, New Jersey. pp 494.

Wang Q., Ni J. and Tenhunen J. (2005) Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. Global Ecology and Biogeography. 14, 379-393.

Weber, D., Englund, E., 1992. Evaluation and comparison of spatial interpolators. Mathematical Geology 24, 381–391.

Weber, D. and Englund, E. 1994. Evaluation and comparison of spatial interpolation. Mathematical Geology, 24: 381-391.

Wilby, R.L. and Wigley, T.M.L., (1997) Downscaling general circulation model output: a review of methods and limitations, Progress in Physical Geography, 21, 530–548.

Wilks, D. S. (2006), Statistical Methods in the Atmospheric Sciences, 2nd ed., 627 pp., Elsevier, New York.

Willmott, C.J., 1981. On the validation of models. Physical Geography 2, 184–194.

Willmott CJ, Rowe CM, Philpot WD. 1985a. Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *American Cartographer* **18**: 2249–2251.

Willmott, C. J., et al. (1991), Influence of spatially variable instrument networks on climatic averages, Geophys. Res. Lett., 18(12), 2249–2251.

Willmott, C. J., et al. (1994), Estimating continental and terrestrial precipitation averages from rain-gauge networks, Int. J. Climatol., 14,403 – 414.

Willmott, C. J., and K. Matsuura (1995), Smart interpolation of annually averaged air temperature in the United States, J. Appl. Meteorol., 34,2577 – 2586.

Willmott, C. J., and S. M. Robeson (1995), Climatologically Aided Inter-polation (CAI) of terrestrial air temperature, Int. J. Climatol., 15, 221–229.

Willmott CJ, Johnson ML. 2005. Resolution errors associated with gridded precipitation fields. *International Journal of Climatology* **25**:1957–1963; (Published Online: 7 Oct 2005, DOI: 10.1002/joc.1235).

Willmot C.J., and Maatsura K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30, 79-82.

Willmott, C. J., and K. Matsuura (2006), On the use of dimensioned measures of error to evaluate the performance of spatial interpolators, Int. J.Geogr. Inf. Sci., 20, 89–102.

Wimberly M.C., Yasblley M.J., Baer A.D., Dugan V.G., Davidson W.R. 2008. Spatial heterogeneity of climate and land-cover constraints on distributions of tick-borne pathogens. Global Ecology and Biogeography, 17:189-202.

Wieczorek JR, Guo Q, Hijmans RJ. 2004. The point-radius method for georeferencing point localities and calculating associated

uncertainty. *International Journal of Geographical Information Science* **18**: 745–767.

WMO. 1996. *Climatological Normals (CLINO) for the period 1961–1990*. World Meteorological Organization: Geneva, Switzerland.

Wood S. N., Augustin N.H. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modeling. Ecological Modeling 157: 157-177.

Wood, S.N. (2003) Thin plate regression splines. J.R.Statist.Soc.B 65(1):95-114.

Wood,S.N. (2006). Generalized Additive Models. An introduction with R. 1st ed., 410pp. Chapman and Hall/CRC.

Wood, S.N. (2008) Fast stable direct fitting and smoothness selection for generalized additive models. J.R.Statist. Soc. B 70(3):495-518

Wood, S. N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models." Journal of the Royal Statistical Society: Series B (Statistical Methodology) **73**(1): 3-36.

Wood, S. (2012). R Package “mgcv” Version 1.7-17, Mixed Gam Computation Vehicle with GCV/AIC/REML smoothness estimation.

1. The Generalized Least Square incorporates the spatial structure of observations in the estimation. While GLS estimates are theoretically needed, Hengl 2007 reports that OLS estimates may be used in practice since results are similar. [↑](#footnote-ref-1)
2. To be precise, the “inverse” of the semi-variogram function since kernel functions may relate to auto-correlogram and covariance rather than semi-variance. [↑](#footnote-ref-2)
3. Duality allows the recasting of the solution of the objective functional (penalized least square regression and regularized least square with tensions) in terms of a trend component T(x) corresponding to a sum of monomials and a local component (R(x)) corresponding to a sum of radial basis function (R(x)). This recasting necessitates expressing the smooth functions in a Hilbert space along with an associated reproducing Kernel Hilbert space for the Kernel radial basis function (Wahba 1990, REF). [↑](#footnote-ref-3)