IOP Conf. Series: Earth and Environmental Science **107** (2018) 012106 doi:10.1088/1755-1315/107/1/012106

Recovery of continuous field of ecological-climatic indices based on the meteorological stations data

A A Saveliev, A M Domnina

Institute of Environmental Sciences, Kazan Federal University, Kremlevskay str., 18, 420008, Russia

E-mail: Anatoly.Saveliev.aka.saa@gmail.com

Abstract. The possibility of using INLA (Integrated Nested Laplace Approximations) method for the recovery of continuous field of ecological-climatic indices based on the meteorological stations data has been considered and the accuracy of such recovery assessed. Meteorological data from 13 stations were used to fit INLA model, and independent data from 58 stations were used to test the model. Traditional ecological-climatic indices were modelled (annual sum of effective temperatures, mean annual temperature, mean temperature of the coldest and the warmest month within a year). All indices show good compliance of the model predictions and data test (R2 in range 0.70-0.90).

1. Introduction

Climate is one of the basic complex ecological factors determining conditions for species distribution, especially that of vegetation. The influence of the climate on vegetation is diverse and primarily related to such important conditions of occurrence as heat and humidity. Thus, ecological factor of heat can be estimated by the use of averaged daily temperatures and their derivatives, i.e. annual sum of effective temperatures, mean annual temperature, mean temperature of the coldest and the warmest month within a year. Although these indices are spatially distributed and they can be calculated only for meteorological stations, they can be interpolated for the rest of the area. Recently, Bayesian statistical approaches have been elaborated for spatial-temporal modeling based on determination of the stochastic field indices. This work is devoted to the application of INLA (Integrated Nested Laplace Approximations) method [1, 2] in spatio-temporal modeling of ecological-climatic indices.

2. Input Data

The GSOD data obtained by National Climatic Data Center (NCDC) in Asheville, the USA, in the framework of data exchange program of World Weather Watch (WWW) in accordance with the WMO Resolution 40 (Cg-XII) offering free access for scientific and non-profit institutions have been used as input data. The access was performed via FTP server (ftp://ftp.ncdc.noaa.gov/pub/data/gsod) and involved average, minimal and maximal daily temperatures and information on wind and sediments.

Three derivative ecological-climatic indices have been calculated from average daily temperature, i.e. mean annual temperature (mean.T), mean temperature of July (jul.T), mean temperature of January (jan.T), the number of days with negative temperatures (Sum.NegativeDays), sum of negative temperatures below -10 ºC (Sum.Negative10.T) and the Growing degree-day (sum of effective temperatures exceeding 10ºC, Sum.Positive10.T).

The area of Middle Volga region limited in the longitude from 43ºE to 58ºE and in the latitude from 52ºN to 58ºN was selected for the modelling. This area covers the most part of the Volga region. The area has a distinct latitude climate gradient (boreal ecocline zone). Modeling area is presented by a rectangle of 860x800 km as shown in figure 1 below. Figure 2 shows triangulation net used for the recovery of ecological-meteorological indices.

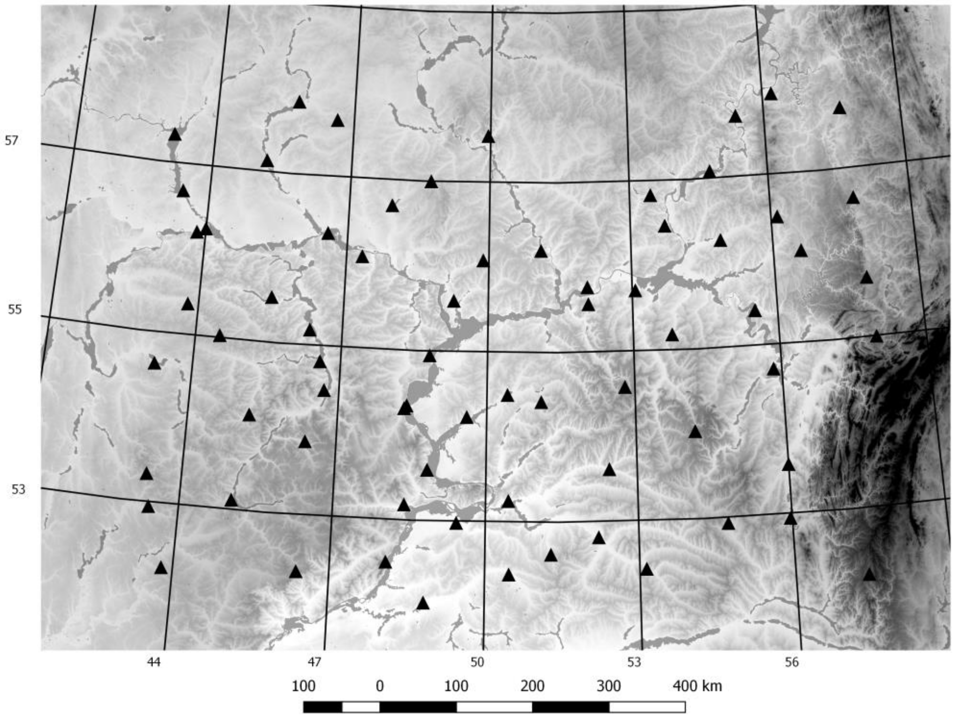


Figure 1. Modeling area and weather observing stations positions.

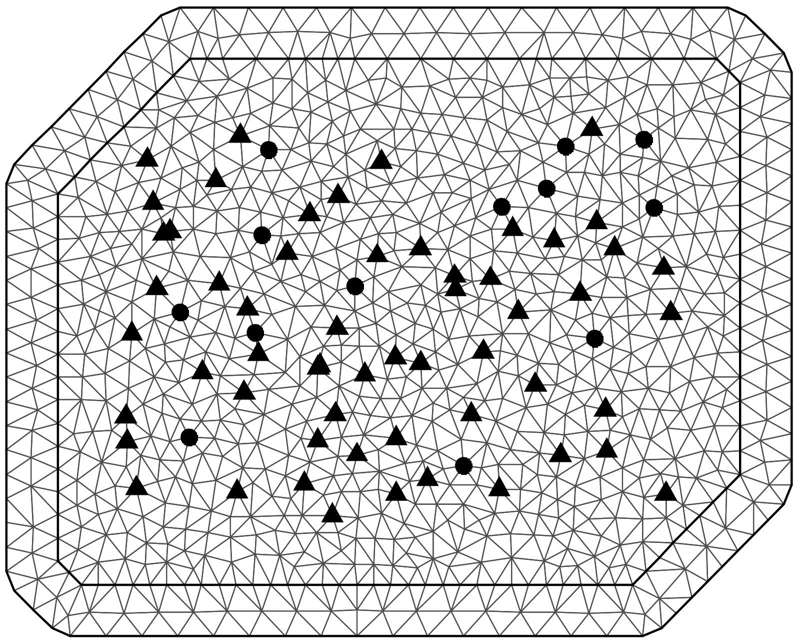


Figure 2. Triangulation net of the modeling area. Round symbols correspond to the stations with input data and triangles to the stations used for the model validation.

Modeling period covers 1980-2016 years reach with input available data. In GSOD database, 71 stations are found with annual information and used as input data source. Other 58 stations with the information covering from 3 to 36 years were applied for modeling results assessment.

3. Methods

The modeling was performed with the statistics system INLA [1] that employs the concept of random fields for the recovery of continuous spatial-temporal field of the distribution of the index to be modeled.

All the parameters modeled indexed as Y were assessed using the same type of model:

$$Y(s, year) = a_0 + a_1 \times Year + a_2 \times H + RF_{mean}(s) + RF_{year}(s, Year) \quad (1)$$

Here $Y(s, year)$ is a modeled index in the spatial location s, in a year $year$, $a_1 \times Year$ is a linear temporal trend (along the years counted from 1980), $a_2 \times H$ is a linear height trend, $RF_{mean}(s)$ is a value of mean annual random field of the index and $RF_{year}(s, Year)$ is a value of random field estimating deviation from mean average value within a $year$.

4. Results and Discussion

The following results have been obtained for linear trends (table 1), Bayesian credible intervals were used to estimate significance.

Table 1. Linear trends.

Index	Years			Height		
	Index	SD	Significance	Index	SD	Significance
mean.T	0.04	0.01	Yes	-0.010	0.003	Yes
jul.T	0.03	0.03	No	-0.011	0.001	Yes
jan.T	-0.01	0.05	No	-0.011	0.008	No
Sum.NegativeDays	-0.49	0.11	Yes	0.094	0.022	Yes
Sum.Negative10.T	0.13	1.40	No	-0.920	0.510	No
Sum.Positive10.T	7.11	1.59	Yes	-1.305	0.206	Yes

The quality of the modeling of spatial-temporal distribution of the indexes was estimated by comparison of the predicted indices with those observed at 58 test stations. For this purpose, correlation coefficient between the observed and modeled indices was calculated for the test stations within the minimal rectangle bounding 13 stations with the input data (tagged “interpolation” in text), and that for all the stations (tagged “all” in text) providing the comparison of interpolation and extrapolation result. The results obtained are given below in table 2 (correlation).

Table 2. Correlation between observed and modelled indexes.

Index	Correlation (Interpolation)	Correlation (All)	R^2 (Interpolation)
mean.T	0.86	0.87	0.75
jul.T	0.92	0.90	0.81
jan.T	0.89	0.87	0.75
Sum.NegativeDays	0.89	0.88	0.78
Sum.Negative10.T	0.76	0.71	0.51
Sum.Positive10.T	0.88	0.87	0.76

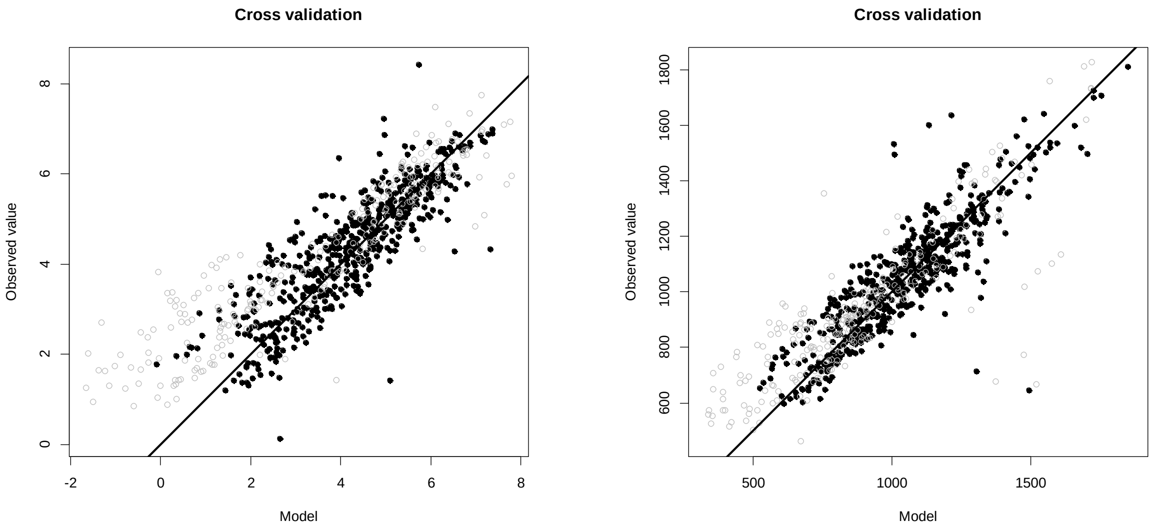
Linear regression results of indices on the corresponding models are given in table 3 (bias, slope and standard deviation).

IOP Conf. Series: Earth and Environmental Science **107** (2018) 012106 doi:10.1088/1755-1315/107/1/012106

Table 3. Linear regression coefficients for models of observed and modelled indices.

Index	Interpolation			All		
	bias	slope	SD	bias	slope	SD
mean.T	0.91	0.81	0.68	1.54	0.69	0.72
jul.T	1.56	0.93	0.82	3.00	0.86	0.95
jan.T	-0.20	0.95	1.64	-0.81	0.87	1.83
Sum.NegativeDays	3.14	0.97	7.15	21.24	0.83	7.88
Sum.Negative10.T	-82.6	0.72	81.4	-123.8	0.54	91.7
Sum.Positive10.T	155.6	0.85	103.1	235.2	0.78	115.0

Figures 3-5 illustrate some results.



mean.T Sum.Positive10.T

Figure 3. The comparison of model and real indices corresponding to test stations. Filled symbols correspond to interpolation data and open symbols to extrapolation results obtained for the stations outside the initial input data area.

The model obtained makes it possible to estimate continuous distribution of the index along the area used for modeling. Thus, model distribution of effective temperatures for 1980 and 2010 is presented below in figure 3 (obtained with mean height value to provide more smooth contours).

The consideration of the results shows that the applied INLA method is a reliable tool for recovery of continuous field of ecological-climatic indices even if a limited set of data from meteorological stations is available. The method allows operating incomplete temporal data. This makes it convenient for imputation of missed stations data. The spatial distribution of the ecological-climatic indices obtained can be used for territory zoning, especially in agriculture and geobotany. Besides, modeling made it possible to discover statistically significant temporal trend of mean temperature and of the sum of Growing degree-days on the area considered. The applied method (INLA) provides not only the indices of expectation estimate, but also the probability distribution in the form of p.d.f. at each location, as well as uncertainty estimate.

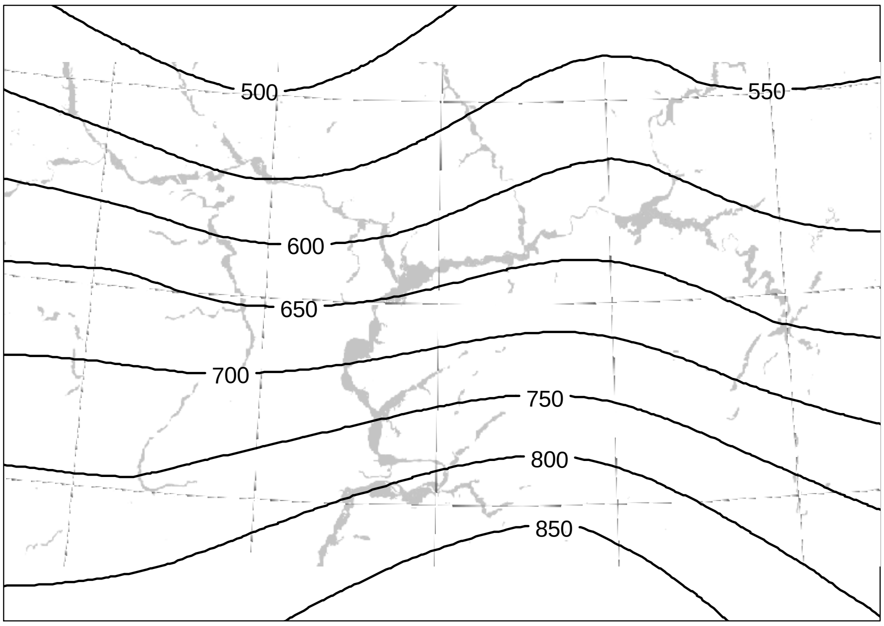


Figure 4. Model distribution of the Growing degree-day (Sum.Positive10.T) values in 1980.

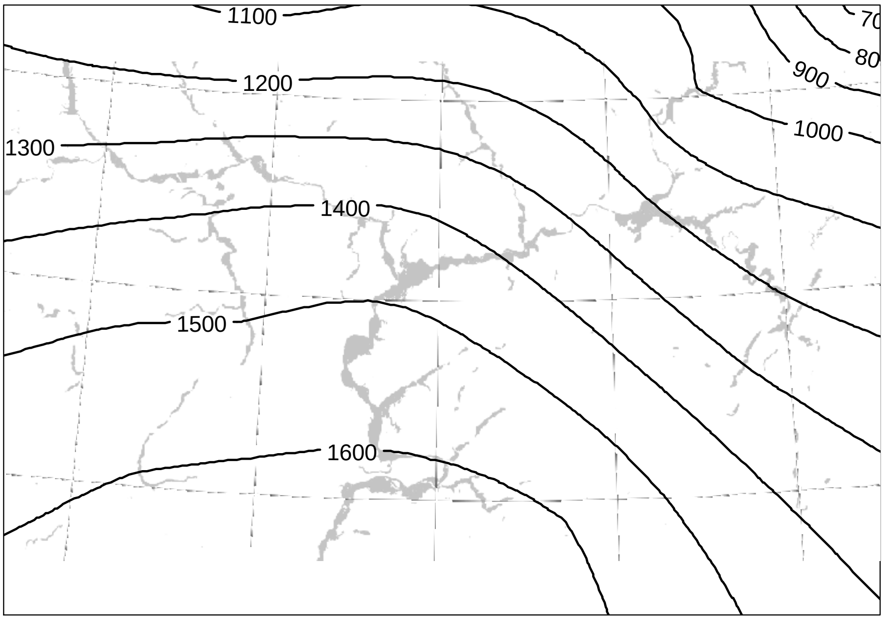


Figure 5. Model distribution of the Growing degree-day (Sum.Positive10.T) values in 2010.

References

[1] Rue H, Martino S and Chopin N 2009 Approximate Bayesian Inference for latent Gaussian models using Integrated Nested Laplace Approximations. *JRSS-series B.* **71**(2) 319-392

[2] Blangiardo M, Cameletti M 2015 *Spatial and Spatio-temporal Bayesian Models with R – INLA.*Wiley, 320 p