



Spatial Interpolation of Ewert's Index of Continentality in Poland

MARIUSZ SZYMANOWSKI,¹ PIOTR BEDNARCZYK,² MACIEJ KRYZA,³ and MAREK NOWOSAD⁴

Abstract—The article presents methodological considerations on the spatial interpolation of Ewert's index of continentality for Poland. The primary objective was to perform spatial interpolation and generate maps of the index combined with selection of an optimal interpolation method and validation of the use of the decision tree proposed by Szymanowski et al. (Meteorol Z 22:577–585, 2013). The analysis involved four selected years and a multi-year average of the period 1981–2010 and was based on data from 111 meteorological stations. Three regression models: multiple linear regression (MLR), geographically weighted regression (GWR), and mixed geographically weighted regression were used in the analysis as well as extensions of two of them to the residual kriging form. The regression models were compared demonstrating a better fit of the local model and, hence, the non-stationarity of the spatial process. However, the decisive role in the selection of the interpolator was assigned to the possibility of extension of the regression model to residual kriging. A key element here is the autocorrelation of the regression residuals, which proved to be significant for MLR and irrelevant for GWR. This resulted in exclusion of geographically weighted regression kriging from further analysis. The multiple linear regression kriging was found as the optimal interpolator. This was confirmed by cross validation combined with an analysis of improvement of the model in accordance with the criterion of the mean absolute error (MAE). The results obtained facilitate modification of the scheme of selection of an optimal interpolator and development of guidelines for automation of interpolation of Ewert's index of continentality for Poland.

Key words: Ewert's index of continentality, spatial interpolation, regression kriging, geographically weighted regression, Poland.

1. Introduction

The concept of continentality in climatology refers to all characteristics of climate influenced by continents and is regarded as a notion opposed to climate oceanity. The continental impact is most frequently considered in the context of its effect on air temperature (thermal continentality) and precipitation (pluvial or hygric continentality). The continental climate is characterised by higher diurnal and annual amplitude of air temperature, hot summers, and cold winters. As the distance from the ocean increases, the cloudiness and precipitation rates decline and the annual distribution of precipitation changes and exhibits a distinct summer maximum. The oceanic climate is characterised by high air humidity and high rates of precipitation distributed evenly throughout the year, with low annual air temperature amplitudes, cool summers, and mild winters.

In a macroscale, it can be assumed that climate features of any place on the globe comprise a signal coming from the continents and a supposedly opposite signal from the oceans. To quantify these interactions, a number of indices, primarily describing the “strength” of land impacts, have been developed. Given the disproportion between the land and ocean cover (ca. 30 %—lands and 70 %—oceans), they can be regarded as a specific modification of the prevailing oceanic climate on the planet. As a rule, indicators of continentality express the relation between continental and oceanic features of climate and, hence, underline the relative nature of continentality. The continentality has been assigned particular importance in characterisation of areas with the so-called transitional climate combining continental and oceanic impacts, which exhibit high year-to-year variability depending on macro-circulation features. For Poland, such analyses have been

¹ Department of Geoinformatics and Cartography, Institute of Geography and Regional Development, University of Wrocław, pl. Uniwersytecki 1, 50-137 Wrocław, Poland. E-mail: mariusz.szymanowski@uw.edu.pl

² Laboratory of Geoinformatics, Maria Curie-Skłodowska University in Lublin, al. Kraśnicka 2cd, 20-718 Lublin, Poland.

³ Department of Climatology and Atmosphere Protection, Institute of Geography and Regional Development, University of Wrocław, ul. Kosiby 8, 51-621 Wrocław, Poland.

⁴ Department of Meteorology and Climatology, Maria Curie-Skłodowska University in Lublin, al. Kraśnicka 2cd, 20-718 Lublin, Poland.

performed by Kożuchowski and Marciniak (1986, 1992), Kożuchowski and Wibig (1988), and Kożuchowski (2003). Recently, continentality indices have been applied in environmental analyses, particularly as a variable supporting the analysis of variability and spatial distribution of, e.g., air temperature (Hogewind and Bissolli 2011), evapotranspiration (Marti and Gasque 2010) and bioclimatic changes (Torregrosa et al. 2013). The analytical usefulness of continentality indices has also been corroborated in the investigations of glacier mass balance (Holmlund and Schneider 1997), plant species range and the treeline (Caccianiga et al. 2008), and plant pollen in northern Europe (Salonen et al. 2012).

Given the recent climate changes, it is expected that continentality indices bring important signs of these changes, particularly when research-based thereon will address the dynamic year-to-year changes, multi-year trends, and predictions supplemented with detailed analysis of extreme values and circulation relationships. On the one hand, maps of continentality indices can provide data concerning a number of environmental elements, thereby being an important variable in multi-dimensional spatial modelling of these elements, as mentioned above. Methodology for accurate spatialization of continentality indices based on discrete point observations is a prerequisite for generation of such maps. This paper is focused on spatial interpolation of the thermal continentality index for Poland. It has a methodological character allowing indication of a preferable spatialization algorithm in accordance with the environmental characteristics of the study area and the nature of the input data set.

The annual temperature amplitude, i.e., the difference between the mean temperature of the warmest and coldest month, is the simplest absolute measure of thermal continentality. The drawback of this indicator lies in its dependence on seasonal changes in the quantity of incoming solar energy, which results in an increase in the annual amplitude with latitude. The effect is compensated for in the thermal continentality indices by dividing the temperature amplitude by the sine of latitude with concurrent introduction of empirical parameters facilitating rescaling of the index into the assumed

interval in accordance with theoretical assumptions (Conrad 1946).

Various continentality indices used in climatology are based on the average annual temperature amplitude scaled with the sine of latitude. However, Driscoll and Yee Fong (1992) suggest that there is no conclusive evidence for the validity of the use of this divisor with respect to amplitude changes determined by the incoming solar energy. This group comprises indicators specified by Gorczyński (1920), Johansson (1926), Raunio (1948), Conrad (Conrad and Pollak 1950), Hela (1953), Ivanov (1953), Khromov (1957), Ewert (1963), and Hogewind and Bissolli (2011). Formulas extending this type of thermal continentality indices with other environmental parameters include indicators developed by Spitaler (1922), Ringleb and Johansson (after Szreffel 1961), and Ivanov (1959). Berg (1944), Bailey (1968), and Oliver (1970) proposed formulas based on other assumptions than annual temperature amplitude. These ideas have been put forward by Driscoll and Yee Fong (1992) and Mikolášková (2009).

Classic approach to present continentality indices uses hand-drawn isolines (Swoboda 1922; Kożuchowski and Marciniak 1992), with generalisation typical of this technique, especially in small scales. In recent studies, index maps are generated in the GIS environment, taking advantage of raster maps and showing more details (Mikolášková 2009; Torregrosa et al. 2013). The spatialization approaches applied are relatively simple, one-dimensional, and disregard selection of an optimal interpolation method and accuracy of the results.

Currently, environmental research, including climatology, employs various interpolation techniques (Hengl 2007; Li and Heap 2008; Szymanowski et al. 2012), which often leads to difficulties in choosing a method that is appropriate in a given case. This problem has been addressed, e.g., in the COST 719 research project “The Use of GIS in Climatology and Meteorology” (Dobesch et al. 2007; COST Action 719 Final Report 2008), but the investigations did not yield a conclusive solution. The spectrum of methods employed for spatial interpolation of climate elements is very wide and comprises deterministic and geostatistical techniques and their combinations. The

best results are usually obtained with multivariate methods considering the role of environmental factors, in particular the elevation and coordinates (Szymanowski et al. 2012). This group includes, among others, the multiple linear regression (MLR) and residual kriging (MLRK, regression kriging) methods. Analyses of the spatial properties of climate elements, particularly non-stationarity, indicate that the local geographically weighted regression (GWR) model with residual kriging (geographically weighted regression kriging, GWRK) is more suitable for modelling the spatial variability of these elements (Szymanowski and Kryza 2011, 2012). Methodological research aiming at interpolation of the temperature field for Poland resulted in construction of a decision tree for selection of an optimal multivariate interpolation method considering the potential non-stationarity of the spatial process (Szymanowski et al. 2012, 2013).

The main objective of this study is to perform spatial interpolation and generate maps of Ewert's index of continentality (1972) for selected cases in Poland representing long-term annual mean values and years with extreme values or characteristics of the spatial distribution of this indicator. This methodological paper emphasises the choice of an optimal interpolation method and indication of guidelines for automated or semi-automated interpolation of the indicator for a large data set (multi-year series) to identify the trend in climate change. Second aim is to validate the usefulness of the scheme of selection of the optimal interpolation method developed by Szymanowski et al. (2012, 2013) for interpolation of the continentality indices.

2. Study Area

Poland (312,679 km²) is situated in the central part of Europe. The terrain elevation in the country ranges from -1.8 m in the north to 2499 m above sea level in the highest, mountainous area in the south of the country. The north-western border of Poland is delimited by the Baltic Sea, and in the south, the border extends along the main ridge of the Sudetes and Carpathians (Fig. 1).

Poland is located in the temperate transitional climate zone with clear continental and oceanic impacts. Westerly winds and polar air masses are predominant. The spatial distribution of the lowest values of the mean annual air temperature is determined by elevation, with the top parts of the mountains (-0.7 °C), and continentality, with the north-eastern part of the country being the second coolest region (<7 °C). The mean annual temperature increases from the north-east to the south-west, where it exceeds 8.5 °C (Woś 2010). July, with temperatures ranging from 17.3 to 18.8 °C (except for the mountains), is the warmest month, and January, with temperatures in the range between -3.4 and -1.3 °C, is the coldest (Kozuchowski 2011). In January, longitudinal distribution of isotherms dominates, with temperatures decreasing eastwards. In July, the course of isotherms exhibits latitudinal distribution with temperatures decreasing from the central regions northwards, towards the Baltic Sea, and southwards along the increasing elevation in the mountains. The mean annual temperature amplitude varies between ca. 15 °C in high mountain areas and 22 °C in the eastern part of the country. The longitudinal course of isoamplitudes typical for the east of the country is deformed in the west. The longitudinal distribution is only observed there in the central part, whereas in the north and south, the arrangement has a latitudinal direction similar to the course of the coast and mountain ranges (Woś 2010). A similar distribution is characteristic for the multi-year continentality indices. The areal average value of Ewert's index of continentality is 44.3 % (Kozuchowski 2011). In decades dominated by oceanic influences, it ranges from below 38 % at the Baltic Sea to over 48 % in the east of the country. In decades dominated by continental influences, the values of Ewert's index range from below 44 % and over 56 %, respectively (Kozuchowski and Marciniak 1992).

3. Data and Methods

3.1. Ewert's Index of Continentality

One of the thermal continentality indices presented in the Introduction section, i.e., Ewert's index

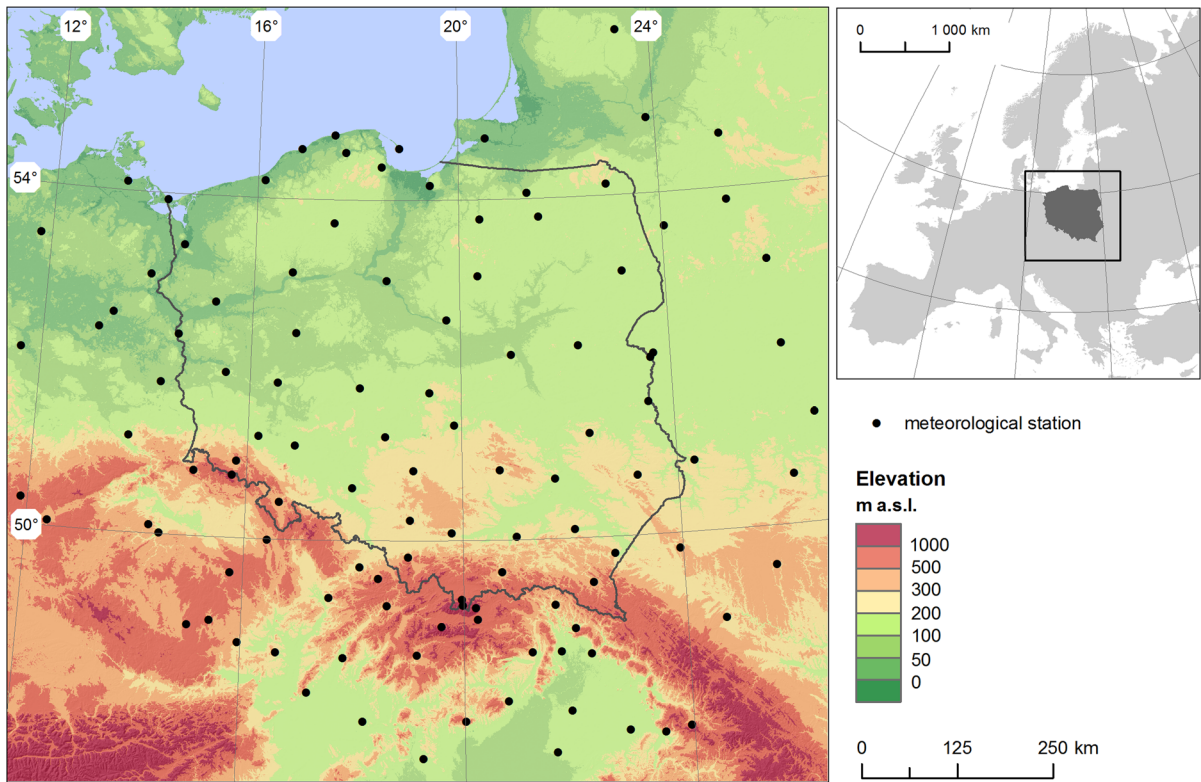


Figure 1
Study area and location of meteorological stations used in the study

(1972), was selected for the analysis. Compared with other indicators involving compensation of the influence of latitude, Ewert's index includes additional parameterization considering the impact of land cover on the value of the annual air temperature amplitude in the individual latitude zones. These relationships were estimated particularly carefully for Central Europe, where the best results are achieved (Ewert 1972). Additional important assumptions for the development of the index include the possibility of applying it to all latitudes and any—annual or multi-year—period. Ewert's index of continentality K [%] is expressed by the formula:

$$K = \frac{A - (3.81 \sin \varphi + 0.1)}{38.39 \sin \varphi + 7.47} \times 100, \quad (1)$$

where A is the annual temperature amplitude and φ is latitude. The index is not limited to the range of 0–100 %, but it has the values from –1.5 % for Turk Island to 141.5 % for Yakutsk. Areas with negative

index values are characterised by “hyperoceanity” and those with the values >100 % by “hypercontinentality” (Ewert 1972). Minimal “hyperoceanity” has been identified at several Pacific stations and “hypercontinentality” in a large part of eastern Siberia (Ewert 1972).

To achieve one of the objectives of this study, i.e., development of a method and general guidelines for spatial interpolation of Ewert's index of continentality for Poland, the analyses were based on five data sets representing mean and extreme values of the index from 1981 to 2010 (Table 1):

- the mean value of the index ($K_{1981-2010}$);
- the years 1989 and 2006 with the lowest and highest areal means for Poland, respectively (K_{1989} , K_{2006}); and
- the years 1990 and 2002 with the smallest and largest index ranges, respectively (K_{1990} , K_{2002}).

The study was based on measurement data provided by 111 synoptic meteorological stations,

Table 1

Summary of statistics of Ewert's index of continentality (K) in Poland for the four selected years and the 1981–2010 average

Statistics	K_{1989}	K_{1990}	K_{2002}	K_{2006}	$K_{1981-2010}$
Mean	36.19	38.52	58.71	67.52	50.26
Minimum	22.79	31.26	35.36	49.25	39.54
Maximum	40.88	42.78	70.10	75.92	55.52
Range	18.09	11.51	34.74	26.67	15.97
Standard deviation	3.43	2.72	7.01	6.03	4.18

including 53 Polish stations and 58 stations located outside Polish borders (Fig. 1). The localization of each station was carefully verified by comparing station metadata with orthophotomaps and digital elevation model. The values of the annual amplitude of air temperature were calculated based on information contained in global summary of the day (GSOD) database and provided by German weather service—Deutscher Wetterdienst (Klimadaten für Deutschland—online—frei; <http://www.dwd.de>).

3.2. Environmental Variables

The findings of the COST 719 project (COST Action 719 Final Report 2008) and the previous experience related to interpolation of meteorological elements for Poland (Ustrnul and Czekierda 2005, 2009; Szymanowski et al. 2012) indicate a significant value of multivariate interpolation algorithms constituted for large range of climatological applications. The methods consider the deterministic relationships between the modelled element and environmental variables—predictors. The spatial distribution of the predictors is represented in the interpolation procedure by raster layers with a fixed spatial resolution. Variables prepared previously for interpolation of temperature (Szymanowski et al. 2012, 2013), which could have a significant impact on the distribution of the continentality index, were used in the study. Three groups of variables were included in the set of potential predictors:

- terrain elevation represented by the digital elevation model (DEM) and derivative layers: the concavity/convexity index (CCI), the foehn index (FI), and insolation (IT); and
- land cover and its derivatives: the percentage of natural surfaces (NS) and artificial surfaces (AS) in the neighbourhood of the data point.

Variables X and Y denote coordinate values in the local coordinate system PUWG-92. The influence of the Baltic Sea is illustrated by the SDI index. To consider the decreasing impact of the Baltic Sea along with distance, the index was constructed as a square root of the smallest Euclidean distance (expressed by the number of cells) of a raster cell from the coastline. SRTM-3 elevation data (<http://www2.jpl.nasa.gov/srtm>) were used as a digital elevation model (DEM). To achieve the goal of this study, the data were transformed into the PUWG-92 system and resampled into 250-m resolution. This resolution is a compromise between the computational costs and details of information introduced to the models and expressed on the maps. High resolution is of special importance when the interpolated variable is potentially strongly dependant on local factors (including changes in altitude). The choice of spatial resolution applied is supported by earlier studies of Szymanowski et al. (2012, 2013) and Szymanowski and Kryza (2015). Based on the 250-m DEM, the concavity/convexity index (CCI) expressing the cool air accumulation effects of concave terrain forms and the foehn index (FI) illustrating the thermal impact of the foehn wind were calculated (Szymanowski et al. 2007). Insolation (IT), expressed by sums of energy of potential total radiation incoming to the terrain surface, is a variable

- variables of the overall spatial trend comprising layers of raster cell coordinates (X , Y) and the sea distance index (SDI);

illustrating the role of energy factors. The calculations were performed with the use of the *r.sun* program implemented in the GIS—GRASS software (GRASS Development Team 2011). The *r.sun* program is a functionally best-developed module for calculation of radiation in GIS, which can be successfully applied to large areas (Šuri and Hofierka 2004) and works with high-resolution terrain models (Kryza et al. 2010). Variables describing the percentage of the surface area arbitrarily called “artificial” (AS) and “natural” (NS) in the surroundings with a radius of 2500 m around each point were prepared on the basis of CLC2000—CORINE Land Cover 2000 database (2004) for European Union countries and the USGS Land Cover database (2011) for Ukraine, Belarus, and Russia. The values of all environmental variables were extracted from raster layers at the coordinates of localization of each meteorological station and were used to specify regression models as described below.

3.3. Statistical and Spatial Analysis Methods

In this study, a scheme of selection of an optimal interpolation method developed with an example of air temperature is employed (Szymanowski et al. 2012, 2013). The scheme is based on Hengl’s decision tree (2007), and was extended with methods dedicated to non-stationary spatial processes. It is generally assumed that the analysed continentality index is characterised by a statistically significant correlation with at least one environmental variable; therefore, the deterministic component can be modelled with the regression method. In the absence of the correlation, spatial interpolation would have to be performed using one of the one-dimensional methods, e.g., ordinary kriging (OK) or inverse distance weighting (IDW), depending on the degree of spatial autocorrelation of the analysed variable (Hengl 2007). In the scheme discussed, the modelled indicator is regarded as a regionalised variable (Matheron 1971), i.e., a random variable comprising structural, local, and random components. This cumulative model was termed a universal model of spatial variability by Matheron. The structural—deterministic ($\hat{m}(s_0)$) and local—stochastic (deterministic model residuals— $\hat{e}(s_0)$) components of variability can be

modelled separately and the estimated value of variable z in position s_0 — $\hat{z}(s_0)$ is a sum of these components (Eq. 2):

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0). \quad (2)$$

If the deterministic component is modelled with the regression method, and kriging is used for modelling regression residuals, such a model is called residual kriging (Hengl 2007). The decision scheme is based on Matheron’s model and offers a possibility of application of one of the four interpolation methods at a general assumption of existence of a correlation between the modelled variable and environmental variables. The choice of an optimal method follows decisions that are taken in two steps. Initially, based on the goodness-of-fit of the model to the observations, the stationarity of the spatial process is assessed. The choice is made between the global, multiple linear regression (MLR) model dedicated to stationary processes, and the local, geographically weighted regression (GWR) model employed for non-stationary processes.

The global multiple linear regression MLR model can be expressed as the following equation:

$$\hat{m}(s_0) = \sum_{k=0}^p \hat{\beta}_k q_k(s_0), \quad (3)$$

where $\hat{\beta}_k$ are estimated regression coefficients ($\hat{\beta}_0$ —estimated intercept and $q_k(s_0)$ —explanatory variables). The explanatory variables were selected for the model with the stepwise regression method. The prerequisite for inclusion of a variable was its statistically significant ($p < 0.05$) correlation with the continentality index and the absence of collinearity with other explanatory variables determined by the variance inflation factor, VIF (VIF < 10).

The strongest reservations concerning the MLR model in the context of spatial interpolation are that the possible local variations and non-constant spatial relation between the predictor and the estimated variable are ignored both during model calibration and in the prediction stage. These factors are considered in the geographically weighted regression model (Fotheringham et al. 2002). The validity of the application of this model for air temperature in Poland has been demonstrated by Szymanowski et al. (2012). When the GWR model is employed, the

deterministic component of the spatial variability model will be expressed as the following equation:

$$\hat{m}(s_0) = \sum_{k=0}^p \hat{\beta}_k(s_0) q_k(s_0), \quad (4)$$

where $\hat{\beta}_k(s_0)$ are estimated regression coefficients in location s_0 and $q_k(s_0)$ —explanatory variables. Estimation of regression parameters is performed locally using the weighted least-squares method. The schemes of weighing in GWR moving windows are developed in such a way that the weight would decrease with the distance between the estimation point and the data point. In this study, we used the bi-square function in adaptive kernels, whose value was selected on the basis of minimisation of the Akaike information criterion (AIC; Fotheringham et al. 2002). Exactly, the same auxiliary variables as in the MLR were used in the GWR model.

In practice, the decision concerning the choice of the regression model for approximation of the deterministic component of the universal model can be taken only on the basis of the degree of the goodness-of-fit of the model to the observation data. It was based on such measures of the fit as the adjusted coefficient of determination (R_{adj}^2), standard error of estimation (STE), corrected Akaike information criterion (AICc), and analysis of variance (ANOVA) for regression model residuals. ANOVA was used to check whether the improvement of the model fit, expressed as a decrease in the sum of squared residuals of the model, was statistically significant (Szymanowski and Kryza 2012).

The justification for the use of the GWR model instead of MLR was tested using two tests of the spatial variability of local geographically weighted regression coefficients, i.e., Monte Carlo implemented in the GWR3.0 software (Fotheringham et al. 2002) and the geographical variability (GV) test from the GWR4.0 software (Nakaya 2016). In the case of stationarity of any of the explanatory variables, it would be justified to apply a mixed GWR (MGWR) model instead of a fully local approach (Nakaya et al. 2005). The mixed models allow to mix in one model the explanatory variables which are spatially non-stationary (like for ordinary GWR) and the stationary predictors (like used in

MLR). Calibration of the semiparametric GWR model was performed for comparison using the GWR4.0 software.

The other decision in the presented scheme of selection of an optimal interpolation method concerns the possibility of extending the model with a stochastic component, i.e., interpolation of regression model residuals with ordinary kriging. This procedure is employed when there is a significant spatial autocorrelation of regression residuals, which is a basis for modelling a variogram that is different from the pure nugget effect. If the pure nugget effect was the only variogram possible to fit to the experimental variogram of regression residuals, the deterministic model would be corrected at each studied point by an average of the residuals, which equals zero, in accordance with the assumptions of MLR as a best linear unbiased predictor (BLUP) model. In the GWR model, which does not meet the criterion of unbiasedness, the average of the residuals varies, although it is sufficiently close to zero (Fotheringham et al. 2002) to assume that the modification of prediction by residual kriging is negligible in the absence of autocorrelation (Fotheringham et al. 2002). Therefore, the absence of autocorrelation of the residuals excludes the extension of the MLR or GWR models into MLRK or GWRK, respectively (Szymanowski et al. 2012).

Spatial autocorrelation is a mathematical expression of spatial relationships described by Tobler's first law of geography (Tobler 1970), i.e., decreasing similarity of features of geographical objects along the increasing distance between them. It describes the degree of correlation of the variable value in one location with the value of the same variable in a different location, which implies that the values of the analysed variable determine and, concurrently, are determined by realisation of the variable in different locations. These relationships result in spatial clustering of similar values, which is referred to as positive autocorrelation. The value of the autocorrelation and its statistical significance was determined by calculating Moran's I statistics (Moran 1950).

The deterministic component residuals were modelled with ordinary kriging (Eq. 5):

$$\hat{e}(s_0) = \sum_{i=1}^n \lambda_i e(s_i), \quad (5)$$

where λ_i are kriging weights determined from the spatial dependence of the deterministic component residuals and $e(s_i)$ —residual in position s_i . Hence, the full model of residual kriging based on the global regression MLRK model (Eqs. 2, 3, 5) can be expressed as the following equation:

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0) = \sum_{k=0}^p \hat{\beta}_k q_k(s_0) + \sum_{i=1}^n \lambda_i e(s_i), \quad (6)$$

whereas residual kriging based on the geographically weighted regression (GWR) model (Eqs. 2, 4, 5) can be expressed as the following equation:

$$\begin{aligned} \hat{z}(s_0) &= \hat{m}(s_0) + \hat{e}(s_0) \\ &= \sum_{k=0}^p \hat{\beta}_k(s_0) q_k(s_0) + \sum_{i=1}^n \lambda_i e(s_i). \end{aligned} \quad (7)$$

Modelling of variograms of the regression residuals was carried out using automated fitting procedure in ArcGIS Geostatistical Analyst. This approach is based on the Levenberg–Marquardt method (Press et al. 1988) of non-linear least-squares approximation. The choice between candidate variogram models (spherical, circular, and exponential) was done basing on the fit quality, using the criterion of the lowest root-mean-squared error (RMSE). To avoid inconsistencies on the edges of local neighbours' search area and subsequent artificial tearing of continuity of interpolated variable, global settings were used to specify kriging models, i.e., all data points were considered in determination of the weights at any interpolation point.

Due to the limited 111-element input data set, the quantification of the modelling results was performed using the leave-one-out cross-validation (CV) technique. It yielded a 111-element set of CV errors, which were used in the validation of the model in two ways. The values of the summary diagnostic measures were calculated and analysis of the spatial distribution of the CV errors was carried out, particularly in terms of systematic local and regional trends.

Three synthetic measures, i.e., mean error (ME), mean absolute error (MAE), and RMSE, were used in the analysis of the CV errors. The relationship between the sizes of the aforementioned errors can be defined as $ME \leq MAE \leq RMSE$, with the two latter measures having only non-negative values with the predicted zero value. The use of the square of the CV errors in the RMSE index makes the measure substantially biased even by an inconsiderable number of large errors, although the other errors may be small and acceptable. According to some researchers, MAE is regarded as the most natural diagnostic measure (Willmott and Matsuura 1995).

Since the differences in the MAE of CV errors may be negligible in the compared methods, the decision as to whether any of the methods yielded significant improvement, i.e., has a significantly lower MAE, was made by comparison of $MAE \pm \hat{\sigma}_{MAE}$ intervals, where $\hat{\sigma}_{MAE}$ was a MAE calculation error. Assuming that the mean μ of CV errors is $\mu_{CV} \cong 0$, and $\hat{\sigma}_{MAE}$ for an n element set with standard deviation σ_{CV} can be calculated with the formula (Kalarus et al. 2010; Eq. 8):

$$\hat{\sigma}_{MAE} = \frac{\sigma_{CV}}{\sqrt{\pi c}}, \quad c = \frac{\pi}{\pi - 2}. \quad (8)$$

Therefore, if the $MAE \pm \hat{\sigma}_{MAE}$ intervals for the two methods are disjoint, the method with a smaller MAE can be considered better. If, however, the analysed intervals have a certain common range, this statement is not justified.

The quantitative analysis of the model was also accompanied by visual evaluation, which by definition serves identification of features of the model that cannot be shown by methods based on the actual values of the interpolated variable. In particular, it facilitates detection of such little realistic effects as spatial discontinuity, unusually large or small values of the modelled variable, strong directional or regional trends, and various artefacts illustrating the characteristics of the interpolation algorithm rather than those of the interpolated variable. Although it is based on the expert knowledge of the modelled element, such validation is subjective and only complementary to quantitative evaluation.

A number of various computer programs were used in the study, both commercial packages and free software. The spatial analyses primarily concerning the structure of the layers of the explanatory variables and spatial interpolation as well as the final maps were generated in the ArcGIS and GIS GRASS software. The analysis of stepwise regression and the global model was carried out in the STATISTICA program, and the complementary analyses of geographically weighted regression (ANOVA, non-stationarity tests, calibration of semiparametric GWR model) were performed using the GWR3.0 (Fotheringham et al. 2002) and GWR4.0 (Nakaya 2016) packages. For cross validation, the R scripts were developed, using gstat and spgwr packages.

4. Results and Discussion

4.1. Ewert's Index Spatial Predictors

The specification of the multiple linear regression (MLR) model using the stepwise method showed that only four environmental variables in the five analysed cases significantly determined the value of Ewert's index: coordinates (X and Y), elevation (DEM), and the distance from the sea (SDI; Table 2). These variables indicate the global nature of the determinants of the continentality index distribution. It should be stressed that variables combining global and local factors, such as insolation (IT), regional factors, such as the foehn index (FI), or local factors reflecting the effect of the terrain relief (CCI) or land cover (AS, NS), were not introduced into the model in any of the analysed cases. These variables were, therefore, not correlated significantly with the continentality index based on air temperature amplitude, although they often proved to be significantly correlated with air temperature alone (Szymanowski et al. 2012, 2013). This may have been caused by the fact that the input data for the analysis presented in this paper were only obtained from synoptic stations that, in accordance with the WMO guidelines, had been located in a way to minimise local impacts; no data from climatological stations were available in the study.

Table 2

Explanatory variables and standardized regression coefficients in multiple linear regression (MLR) models of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average (descriptions of variables in the text)

Model parameter (in the order of significance)	Standardized coefficients	P value
K_{1989}		
SDI	0.753	0.000
DEM	-0.612	0.000
X	0.121	0.050
K_{1990}		
Y	-0.956	0.000
DEM	-0.359	0.000
K_{2002}		
X	0.725	0.000
DEM	-0.569	0.000
SDI	0.101	0.014
K_{2006}		
DEM	-0.525	0.000
X	0.323	0.000
$K_{1981-2010}$		
SDI	0.754	0.000
DEM	-0.576	0.000
X	0.363	0.000

DEM elevation, *SDI* sea distance index, *X* easting coordinate, *Y* northing coordinate

Each time, a maximum of three variables was introduced into one model, and only elevation was included in all the five models. This is related to the expected decline in the temperature amplitude together with elevation. Variable X describing the rate of change in the east–west orientation was introduced four times, except for the index calculated for 1990 (Table 2). Similarly, one of the variables characterising the zonal distribution of Ewert's index, Y , or SDI was introduced in four cases. Importantly, in none of the models, do these variables appear simultaneously. In the geographical conditions of Poland with the Baltic Sea in the north of the country, and given arrangement of meteorological stations, variable Y (northing) and the sea distance index SDI are correlated, and hence, only one of them was introduced into the regression model even if both were significantly correlated with the continentality index (SDI—3 times, Y —1 time; Table 2). There was no significant correlation with one of the zonal variables Y or SDI only in 2006.

Table 3

Selected statistics of multiple linear regression (MLR), geographically weighted regression (GWR), and mixed geographically weighted regression (MGWR) models of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

Regression model	Auxiliary variables (in the order of significance)	Bandwidth size	Adjusted R^2	Corrected Akaike information criterion (AICc)	Standard error of estimation (STE)
K_{1989}					
MLR	SDI, DEM, X	–	0.66	586.36	3.33
GWR		65	0.77	550.19	2.74
MGWR		65	0.77	549.32	2.74
K_{1990}					
MLR	Y, DEM	–	0.76	510.86	2.38
GWR		48	0.84	472.13	1.92
K_{2002}					
MLR	X, DEM, SDI	–	0.87	566.53	3.05
GWR		65	0.91	538.4	2.59
MGWR		65	0.91	537.07	2.59
K_{2006}					
MLR	DEM, X	–	0.37	684.13	5.20
GWR		32	0.75	596.74	3.27
$K_{1981-2010}$					
MLR	SDI, DEM, X	–	0.90	427.45	1.63
GWR		65	0.91	538.4	2.59

DEM elevation, SDI sea distance index, X easting coordinate, Y northing coordinate

The values of standardized regression coefficients indicate a varying impact of the environmental variables on the individual cases, which concurrently imply significant differences in the distribution of the continentality index. The sign of the coefficients identifies the direction of the relationships: in general, Ewert's index of continentality decreases together with the terrain elevation and northwards and increases along the distance from the sea and eastwards (Table 2).

4.2. Regression Models

Global regression model (MLR) explains over 66 % of the Ewert's index variability in four analysed cases (Table 3). The coefficient of determination was as low as 0.37 only for 2006. The model for the 1981–2010 average was the best fitted one ($R^2_{adj.} = 0.90$) with a 1.63 % standard error of estimation. In the poorest fitted model for 2006, STE was found to be 5.20 %.

Local geographically weighted regression models (GWR) for all the cases were characterised by a better fit to the observations. The highest increase in the coefficient of determination was found for 2006:

from 0.37 for MLR to 0.75 for GWR. In the other cases, there was an increase in the explanation of the variance of Ewert's index from +2 % for the multi-year average to +11 % for 1989. Increase in explained variance was accompanied by a simultaneous decrease in the STE value and a significant (>3) reduction in the Akaike information criterion (AICc, Table 3). The ANOVA analysis indicated a statistically significant decrease in the residual sum of squares of the GWR models, compared with MLR (Table 4).

The Monte Carlo and geographical variability (GV) tests for spatial non-stationarity indicated possible spatial stationarity of the DEM variable GWR coefficients in two cases: for 1989 (GV test only) and 2002 (both tests, but the result of the GV test was not significant at 0.05; Table 5).

The mixed MGWR models for the two cases considered were characterised by a slightly lower residual sum of squares than the GWR models (Table 4). However, the coefficients of determination and estimation errors were similar to the GWR model, and AICc did not indicate a significant improvement in the fit to the observations (Table 3).

Table 4

ANOVA of multiple linear regression (MLR), geographically weighted regression (GWR), and mixed geographically weighted regression (MGWR) models of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

Source	K_{1989} Sum of squares	K_{1990}	K_{2002}	K_{2006}	$K_{1981-2010}$
MLR residuals	1186.5	612.8	992.3	2918.9	283.5
GWR residuals	733.3	353.1	659.6	945.9	195.0
GWR improvement	453.2	259.7	332.7	1972.9	88.5
MGWR residuals	748.4	–	670.2	–	–
MGWR improvement	438.1	–	322.1	–	–

The analysis presented in this section shows a better fit of the local GWR regression model than that of the global MLR model. This suggests the non-stationarity of the spatial process in each analysed case of the spatial distribution of Ewert's index of continentality. No significant difference between the GWR and mixed MGWR models was found either. Thus, accordingly to Occam's razor principle, where simpler models are preferable to more complex ones, because they are better testable and falsifiable, simpler GWR model was used for further analysis.

4.3. Residual Kriging Models

Extension of the regression models to the form of residual kriging was preceded by an analysis of the spatial autocorrelation of the regression residuals. Moran's I statistics showed a statistically significant positive autocorrelation, i.e., clustering of similar values of residuals, for all the MLR models (Table 6). For the GWR residuals, a tendency towards clustering of similar values was observed in four cases, but the autocorrelation was not significant at the 0.05 level. In the case of year 2002, a random distribution of the GWR residuals, i.e., the absence of autocorrelation, was observed (Table 6).

Here, as in the previous studies (Szymanowski et al. 2012, 2013), an exclusion criterion was used, i.e., the absence of a significant (<0.05) positive autocorrelation prevented correct fitting of the theoretical variogram different from the pure nugget effect. Thus, according to none significant autocorrelation of all GWR residuals (Table 6), variogram modelling was performed only for MLR residuals (Table 7). In all the cases, the spherical model was

the best fitted (lowest or at least the same RMSE) while comparing to exponential and circular ones (Table 7).

4.4. Cross-Validation Results

The diagnostic measures of CV errors of the continentality index for Polish synoptic stations are presented in Table 8. Summary statistics were calculated for the two regression models (MLR and GWR) and one residual kriging model (MLRK). The choice of the model that would be optimal considering cross-validation results was made based on ME, MAE, and RMSE measures and extreme errors (Table 8). MAE was regarded as the main measure in this case, since it facilitated assessment of the statistically significant difference between the error sizes and, consequently, identification of significant improvement provided by the model with lower MAE values. This is particularly important, as the differences in the diagnostic measures, including MAE, are often inconsiderable, particularly when similar categories of models, e.g., regression or residual kriging, are compared (Table 8).

The better fit of the GWR models, comparing with MLR, was also confirmed by cross validation. In general, the GWR model was characterised by better CV summary statistics in each analysed case. This did not imply that the model was better in terms of all the five measures employed. Sometimes, the ME or one of the extreme errors was closer to zero in MLR than in GWR. However, the MAE and RMSE for GWR were always lower than for MLR. Importantly, GWR was characterised by a significantly lower MAE only in 2006 (Table 8).

Table 5

Local parameter variability tests for geographically weighted regression (GWR) models of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

GWR model parameter	Monte Carlo test (GWR3.0 software) p value	Geographical variability test (GWR4 software) difference of criterion
K_{1989}		
Intercept	0.000	−6.234
DEM	0.020	0.874
SDI	0.000	−29.289
X	0.000	−10.570
K_{1990}		
Intercept	0.000	−15.714
DEM	0.000	−0.843
Y	0.000	−24.766
K_{2002}		
Intercept	0.000	−6.920
DEM	0.140*	1.363*
SDI	0.000	−28.086
X	0.000	−16.900
K_{2006}		
Intercept	0.000	−0.305
DEM	0.000	−5.312
X	0.000	−22.551
$K_{1981-2010}$		
Intercept	0.000	−12.387
DEM	0.000	−0.286
SDI	0.000	−14.413
X	0.000	−11.086

* Not significant at 0.05; bold numbers suggest no spatial variability

The absence of autocorrelation of GWR residuals limits the model of spatial variation only to explanation of the deterministic component. Thus, even if the fit of the GWR model is significantly better than that of MLR (as in the case of 2006; Table 3), GWRK is not effective when regression residuals are randomly distributed. Paradoxically, this may imply that a better fitted deterministic model does not necessarily guarantee better spatial interpolation. Consequently, the scheme of the optimal interpolator selection applied so far (Szymanowski et al. 2012, 2013) requires revision, so that it will consider the above-mentioned possibility of the absence of autocorrelation of GWR residuals accompanied by a significant positive autocorrelation of MLR residuals (Fig. 2). Therefore, a question arises whether a reverse situation, i.e., the absence of autocorrelation of MLR residuals

Table 6

Spatial autocorrelation of multiple linear regression (MLR) and geographically weighted regression (GWR) residuals of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

Model	Moran's I statistics $E(I) = -0.009$	p value
K_{1989}		
MLR	0.380	0.000
GWR	0.083	0.145*
K_{1990}		
MLR	0.357	0.000
GWR	0.049	0.357*
K_{2002}		
MLR	0.244	0.000
GWR	−0.019	0.870*
K_{2006}		
MLR	0.585	0.000
GWR	0.098	0.095*
$K_{1981-2010}$		
MLR	0.313	0.000
GWR	0.078	0.165*

* Not significant at 0.05

accompanied by a significant positive autocorrelation of GWR residuals, should be considered in such scheme. Such situation is highly unlikely, since it would have to result from a better fit of the MLR than the GWR model. In fact, MLR can be regarded as a special case of GWR with the constant weight along the distance and all observation data points used for model calibration. This implies that the GWR model may work better or be comparable to MLR, but the reverse situation should be rather excluded. Thus, such option is not included in the proposed selection tree (Fig. 2).

In all the analysed cases, the MLRK models have smaller errors than both the regression models. Moreover, due to MAE criterion, MLRK was significantly different from MLR and GWR, so it can be assumed the optimal spatial interpolation algorithm. To confirm the choice, an additional visual assessment of maps of the continentality index was carried out.

4.5. Maps of Ewert's Index of Continentality

To present the characteristics and differences between the maps generated with the three spatial models analysed, two cases were selected, i.e., the

Table 7

Variogram fitting evaluation and parameters of spherical variograms of multiple linear regression (MLR) residuals of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

Variogram model fitting—RMSE			Spherical model parameters		
Spherical	Circular	Exponential	Partial sill	Range	Nugget
K_{1989}					
2.61	2.62	2.66	5.3279	382,579	5.9659
K_{1990}					
2.34	2.34	2.34	2.3679	870,121	4.0273
K_{2002}					
2.69	2.70	2.72	3.0954	709,466	6.7452
K_{2006}					
2.83	2.83	2.96	23.618	380,159	5.1770
$K_{1981-2010}$					
1.43	1.44	1.44	0.29,026	500,120	2.3367

years 2002 and 2006. Due to the large range and considerable variability of the index (Table 1), these were the most demanding cases in terms of spatial interpolation of the analysed data set. This was also reflected in the largest cross-validation errors (Table 8).

The largest differences between the fit of the MLR and GWR models were found for 2006 (Table 3). This was clearly reflected on maps generated with these two methods (Fig. 3a, b). The DEM and X variables included in the MLR model are responsible for the overall decrease in the index with elevation and an increase eastwards. The process, however, is highly non-stationary (Table 5) and the model significantly overestimates the values obtained on the Baltic Sea coast and underestimates the values from the rest of the country (Fig. 3a). The considerably better fitted GWR model substantially increases the value of the index in the central part of the country and slightly decreases it on the coast, contributing to the reduction of the CV errors size. The general tendency towards overestimation is still visible in the northern part of Poland (Fig. 3b). The map generated using the GWR method clearly emphasises the role of terrain relief, particularly in the northern part of the country. This is related to the small kernel bandwidth (32 points), i.e., the smallest of all the analysed cases (Table 3), which can locally emphasise the effect of the explanatory variable. The correction introduced by the geostatistical component is significant in the MLRK model

(Fig. 3c) which can be regarded as optimal in this case.

In the case of 2002, although the improvement of the fit introduced by the GWR model compared with MLR is statistically significant (Table 4), the maps of Ewert's index do not show significant differences, the spatial distribution and size of the CV errors are similar (Fig. 4a, b), and the MAE criterion does not indicate a significant difference between the methods. All three models for 2002 exhibit a poor local fit to the observations in the central-northern part of the country (Fig. 4). Despite this limitation, MLRK, which is statistically significantly better than MLR and GWR in terms of the MAE criterion, can be regarded as an optimal model for 2002 (Table 8).

In the other analysed cases, MLRK were also regarded as optimal models. This supported the conclusions drawn from the analysis of the cross-validation errors also confirming the necessity of introduction of a modified scheme of selection of an optimal predictor (Fig. 2).

The differences in the size, range, and variation of Ewert's index indicated in the observations (Table 1) were also reflected on the maps (Fig. 5). The spatial distribution of the 1981–2010 multi-year average is characterised by a distinct upward trend from NW to SE with a clearly decreased value of the index in the top parts of the highest mountains. This is particularly visible in the western part of the Sudetes, SW Poland (Fig. 5a).

Table 8

Summary statistics of cross-validation (CV) errors for multiple linear regression (MLR), geographically weighted regression (GWR), and multiple linear regression—kriging (MLRK) models of Ewert's index of continentality (K) for the four selected years and the 1981–2010 average

Statistics	MLR	GWR	MLRK
K_{1989}			
ME	1.080	1.068	0.408
MAE	2.180 (1.991 ÷ 2.369)	2.065 (1.893 ÷ 2.237)	1.109 (1.000 ÷ 1.218)
RMSE	2.526	2.385	1.400
MIN	−4.054	−3.254	−2.520
MAX	5.504	5.117	3.690
K_{1990}			
ME	1.010	0.862	0.434
MAE	1.797 (1.652 ÷ 1.942)	1.739 (1.599 ÷ 1.879)	1.399 (1.275 ÷ 1.523)
RMSE	2.111	2.018	1.660
MIN	−3.452	−2.977	−2.631
MAX	4.341	4.135	3.649
K_{2002}			
ME	−0.479	−0.518	−0.111
MAE	2.158 (1.938 ÷ 2.378)	2.071 (1.866 ÷ 2.276)	1.819 (1.621 ÷ 2.017)
RMSE	2.982	2.848	2.690
MIN	−5.582	−6.134	−7.992
MAX	9.654	8.956	8.496
K_{2006}			
ME	−2.989	−2.171	−0.442
MAE	5.078 (4.727 ÷ 5.429)	4.317 (3.998 ÷ 4.636)	1.645 (1.500 ÷ 1.790)
RMSE	5.748	5.044	2.107
MIN	−9.414	−7.551	−5.536
MAX	10.631	11.024	6.357
$K_{1981-2010}$			
ME	0.341	0.323	0.243
MAE	1.331 (1.223 ÷ 1.439)	1.285 (1.182 ÷ 1.388)	1.118 (1.025 ÷ 1.211)
RMSE	1.608	1.565	1.398
MIN	−3.233	−2.975	−2.557
MAX	3.353	3.598	3.664

ME mean error, MAE mean absolute error, RMSE root-mean-square error, MIN minimum, MAX maximum; in brackets—MAE error range (description in the text)

Although discussion of the reasons for this situation lies beyond the scope of this paper, it should be noted that the differences in the distribution between the individual years in relation to the averaged field are significant. In 1990, there was a pronounced zonal trend in the distribution of the continentality index; in 2002, the distribution was longitudinal, and in 2006, high values of the continentality index covered large areas of the central Poland (Fig. 5). Such strong variability of the index distribution may indicate its analytical usefulness as a complementary indicator of macro- and mesoscale atmospheric processes.

5. Summary and Conclusions

The primary aim of the paper was to perform spatial interpolation and to generate maps of Ewert's index of continentality in Poland for the selected years and an average from the period 1981–2010. The main emphasis was placed on the methodological side of spatial interpolation to develop guidelines for automation of the interpolation process.

Additional objective was to test, on the example of Ewert's index, the validity of the scheme of selection of an optimal interpolation method developed by Szymanowski et al. (2012, 2013) for

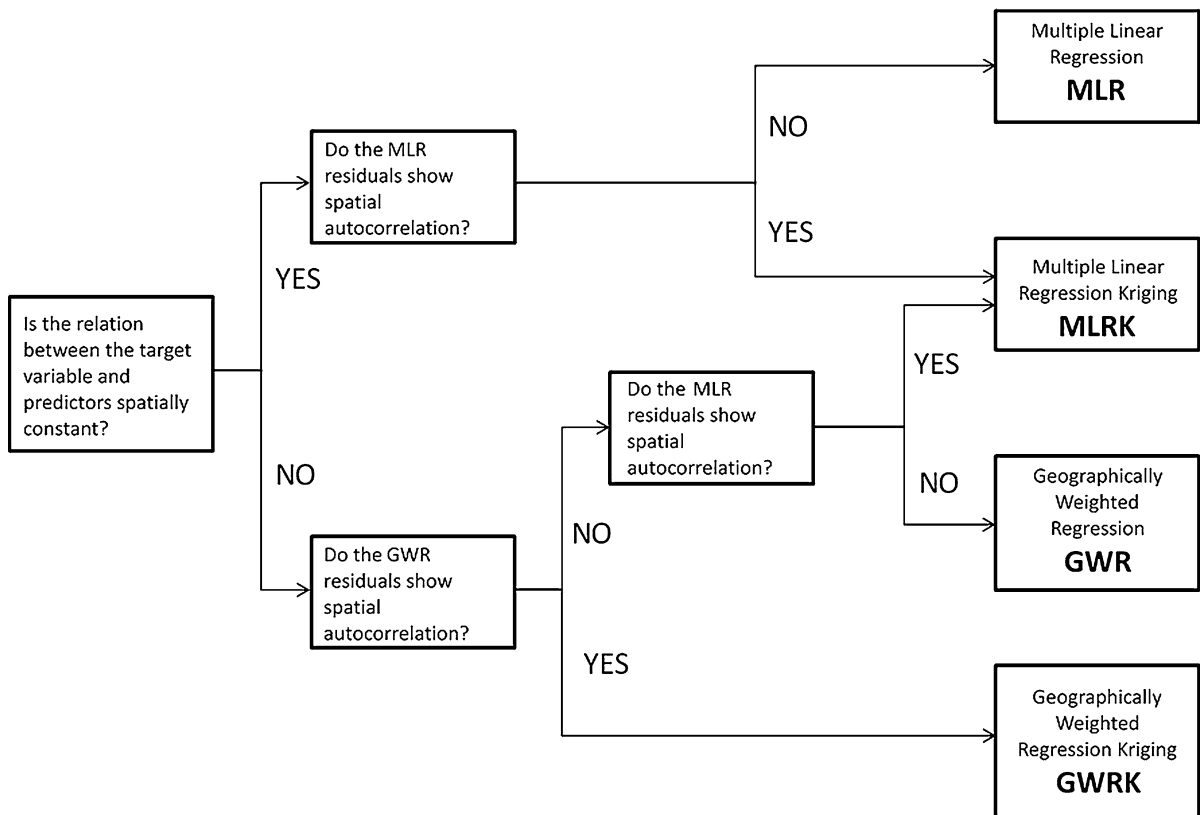


Figure 2

Decision tree for selecting a suitable spatial prediction model under the assumption of existing environmental correlations (after Szymanowski et al. 2012, 2013, extended)

spatialization of air temperature. An optimal method was yielding the smallest errors and simultaneously allowing generation of an acceptable map on the basis of expert knowledge about the spatial characteristics of the interpolated variable.

The continentality index for the period 1981–2010 was calculated based on the daily data from 111 meteorological stations. Four years (1989, 1990, 2002, and 2006) characterised by extreme values of the areal average and index range were selected for the analysis. The set of analysed cases was complemented with the mean value of Ewert's index for the multi-year period 1981–2010.

The set of potential predictors of the continentality index comprised nine environmental variables previously used for spatial interpolation of air temperature in Poland (Szymanowski et al. 2012, 2013). However, the stepwise regression analysis

demonstrated that only four variables in the analysed cases exhibited a significant correlation with Ewert's index, i.e., coordinates X and Y , elevation, and the distance from the sea. Regional or local variables were not significantly correlated with the index, which implies that the features of the continentality index field are mainly determined by macroscale and regional factors (distance from the Baltic Sea) with modification depending on terrain elevation. Noteworthy, the analysis was performed on the basis of the index calculated for the synoptic stations; the location of which, by definition, minimises the local impacts on climate elements. Data from lower-rank, climatological stations, where the impact may be more significant, were not included.

The correlation analysis performed to eliminate collinearity in the regression model additionally showed that variables Y (northing) and SDI were

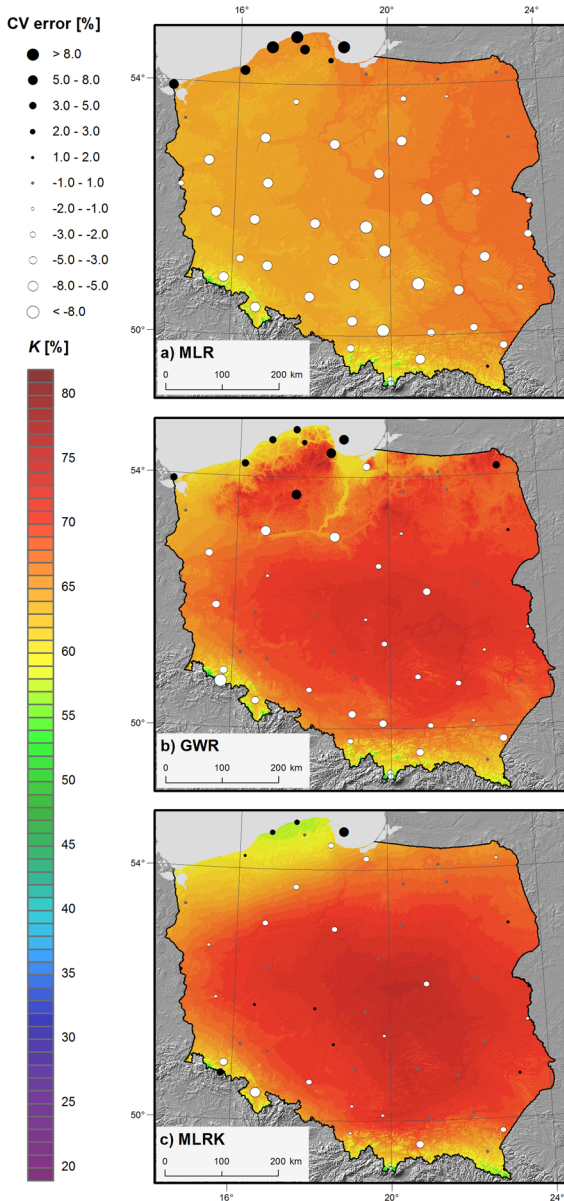


Figure 3

Ewert's index of continentality (K) in Poland in 2006 spatialized using selected interpolation methods

correlated in the geographical conditions of Poland and considering given set of meteorological station. Therefore, only one of them was included in the multiple linear regression MLR model (SDI—3 times, $Y-1$ time). The elevation variable was included in the models of each of the five cases, although this does not mean that its role was dominant. In fact, the role of the explanatory variables in

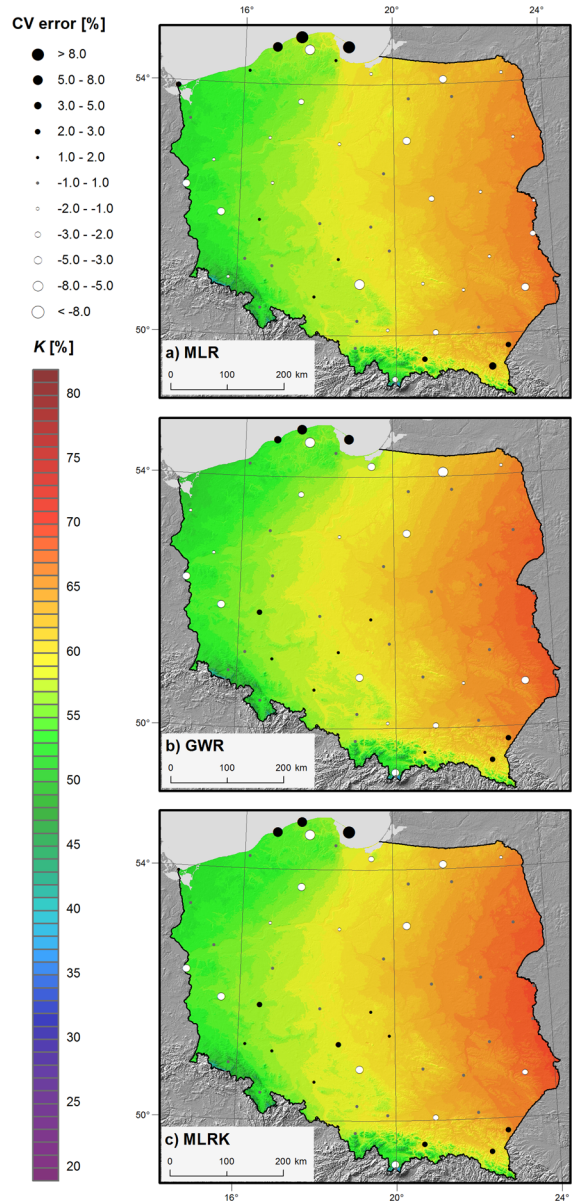


Figure 4

Ewert's index of continentality (K) in Poland in 2002 spatialized using selected interpolation methods

the regression models varied between the cases, which indicates significant changes in the determinants and, hence, variable features of the Ewert's index distribution. The general regularities of the distribution allow a conclusion that the continentality index for Poland decreases with elevation and northwards and increases along the distance from the sea and eastwards, which is also supported by

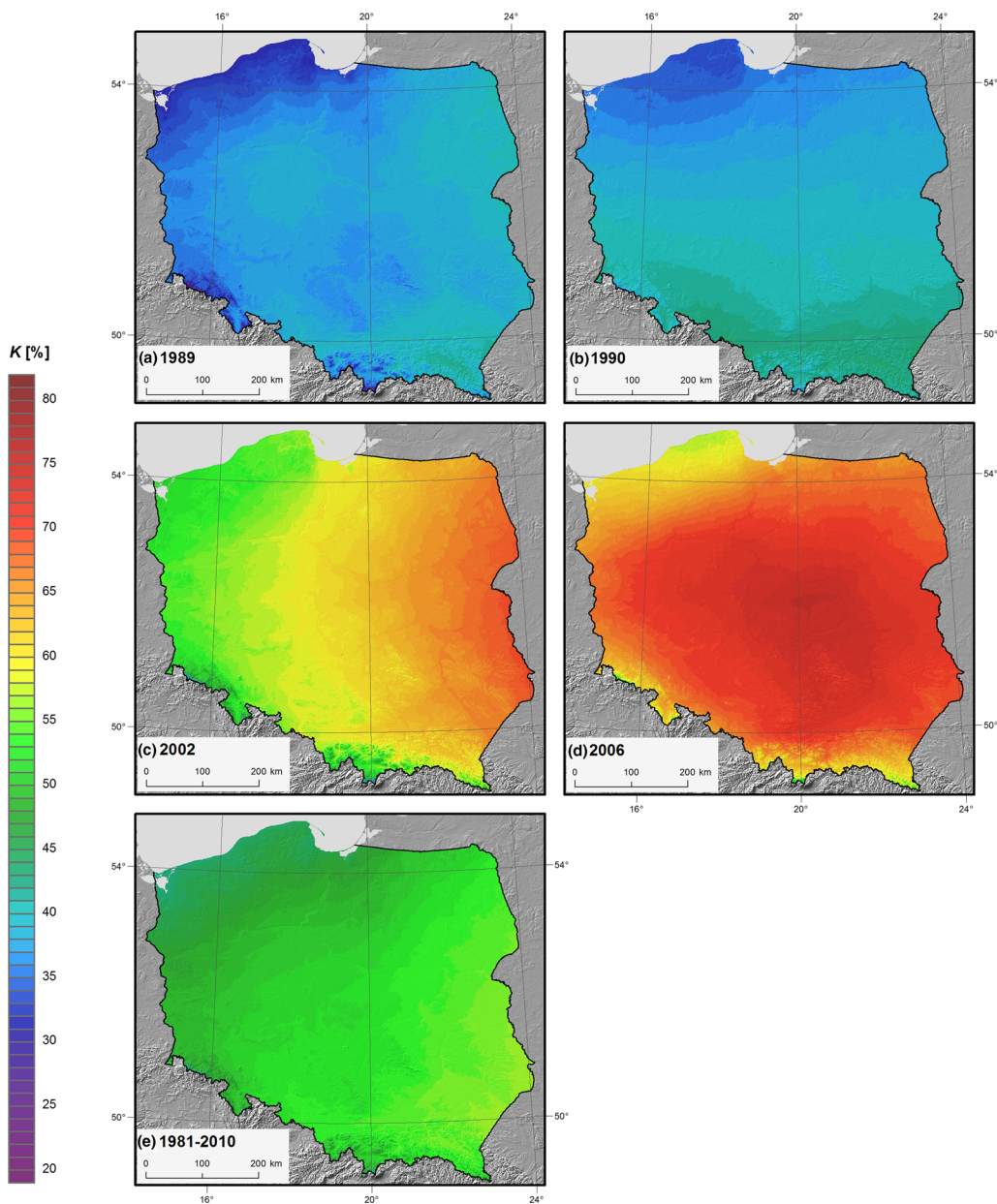


Figure 5

Maps of Ewert's index of continentality in Poland for the four selected years and the 1981–2010 average

the previous studies (Kozuchowski and Marciniak 1992).

The environmental correlation for each of the analysed cases was statistically significant, although for 2006, the MLR model explained solely 37 % of the Ewert's index variation. In each case, the local GWR model was better fitted to the

observations and it significantly improved the MLR regression results. The tests of the spatial variability of the GWR regression coefficients indicated their stationarity. However, compared with the GWR, the use of mixed local–global regression models (MGWR) did not significantly improve the fit.

The residuals of the MLR models were characterised by a positive spatial autocorrelation, which clearly justified extension of the model to the form of residual kriging, MLRK. In contrast, no statistically significant autocorrelation was found in any case of GWR, so the GWRK model was excluded from further considerations.

The cross validation was carried out for three models of each case of Ewert's index. The local GWR model was characterised by smaller errors. In terms of the MAE criterion, MLR was found to be statistically significantly worse than GWR only for 2006. CV carried out for residual kriging models clearly indicated that the MLRK models produced smaller errors than both regression models. However, the specification of GWRK was not justified and gave grounds for modification of the scheme of optimal interpolator selection in the presence of the environmental correlation. The modified scheme of Szymanowski et al. (2013) now allows a situation where the absence of an autocorrelation of GWR residuals is accompanied by an autocorrelation of MLR residuals, which indicates that MLRK is an optimal method, even if GWR is better fitted to observations than MLR.

The visual assessment of the maps confirmed the results of the cross validation and the MAE criterion. The MLRK model is recommended to spatialize the Ewert's index for Poland.

Based on the results obtained, it can be postulated that the following elements should be considered in the procedure of automation of the interpolation of the continentality index for Poland:

- Regression models should be specified using four potential predictors: coordinates (X , Y), elevation (DEM), and distance from the sea (SDI).
- The decision concerning selection of the interpolator should in each case be based on the proposed scheme (Fig. 2), including comparison of the fit of the regression models to observations by, e.g., comparing the coefficient of determination and autocorrelation regression residuals using Moran's I statistics.
- For autocorrelated residuals, the automatic fitting of the variogram can be done using the spherical model.

In the previous papers (Szymanowski and Kryza 2012; Szymanowski et al. 2012, 2013), it was assumed that the choice of an optimal interpolation method was determined by the stationarity or non-stationarity of the spatial process, which could be inferred from the better fit of one of the regression models, i.e., global or local. The results obtained in this study demonstrated greater importance of the geostatistical component in the universal model of spatial variation (Hengl 2007). Residual kriging is highly efficient if regression residues exhibit a strong positive autocorrelation contributing to the well fit of the theoretical variogram to the experimental one. This does not imply, however, a possibility of exclusion of the deterministic component if the environmental correlation is significant. Although this would lead to maintenance of the good fit of the model to observation in measurement points, the spatial distribution of the index would simultaneously "diverge" from environmental features beyond these points.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

- Bailey, H. (1968). Hourly temperatures and annual range. *Yearb. Assoc. Pac. Coast Geogr.*, 32, 25–40.
- Berg, H. (1944). Zum Begriff der Kontinentalität. *Meteorologische Zeitschrift*, 61, 283–284.
- Caccianiga, M., Andreis, C., Armiraglio, S., Leonelli, G., Pelfini, M., & Sala, D. (2008). Climate continentality and treeline species distribution in the Alps. *Plant Biosyst.*, 142, 66–78.
- Conrad, V. (1946). Usual formulas of continentality and their limits of validity, Eos T. *American Geophysical Union*, 27, 663–664.
- Conrad, V., & Pollak, L. W. (1950). *Methods in climatology* (2nd ed.). Cambridge: Harvard University Press.
- CORINE Land Cover 2000 in Poland. (2004). Final Report, Warsaw. http://www.igik.edu.pl/images/stories/sip/clc_final_report_pl.pdf. Accessed Aug 2012.
- COST Action 719 Final Report. (2008). In O. E. Tveito, M. Wegehenkel, F. van der Wel, & H. Dobesch (Eds.), *The use of geographic information systems in climatology and meteorology*. Luxembourg: Office for Official Publications of the European Communities.

- USGS Land Cover (2011). <http://landcover.usgs.gov/usgslandcover.php>. Accessed Mar 2011.
- Dobesch, H., Dumolard, P., & Dyras, I. (Eds.). (2007). *Spatial interpolation for climate data: The use of GIS in climatology and meteorology*. London: ISTE Ltd.
- Driscoll, D. M., & Yee Fong, J. M. (1992). Continentality: A basic parameter re-examined. *International Journal of Climatology*, 12, 185–192.
- Ewert, A. (1963). Kontynentalizm termiczny klimatu. *Przegl. Geofiz.* XVI, 3, 143–150.
- Ewert, A. (1972). O obliczaniu kontynentalizmu termicznego klimatu. *Przegl. Geogr.* XLIV, 2, 273–286.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester: Wiley.
- Gorczyński, W. (1920). Sur le Calcul du Degré du Continentalisme et son Application dans la Climatologie. *Geografiska Annaler*, 2, 324–331.
- Johansson, O. V. (1926). Über die Asymmetrie der Meteorologische Schwankungen, *Societas Scientiarum Fennica Communications in Mathematical Physics*, 3 I, Helsingfors, 124.
- Matheron, G. (1971). The theory of regionalised variables and its applications (Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau, 5. Ecole Nationale Supérieure des Mines de Paris).
- GRASS Development Team. (2011). Geographic resources analysis support system (GRASS) Software, Version 6.4.0. Open Source Geospatial Foundation. <http://grass.osgeo.org>.
- Hela, I. (1953). Regional distribution of the continentality in the climate of the oceans. *Geophysica*, 4, 41–47.
- Hengl, T. (2007). *A practical guide to geostatistical mapping of environmental variables*. Luxembourg: Office for Official Publications of the European Communities.
- Hogewind, F., & Bissolli, P. (2011). Operational maps of monthly mean temperature for WMO Region VI (Europe and Middle East). *Idojaras*, 115, 31–49.
- Holmlund, P., & Schneider, T. (1997). The effect of continentality on glacier response and mass balance. *Annals of Glaciology*, 24, 272–276.
- Ivanov, N. (1953). Ob opredeleni welichyny kontinentalnosti klimata, *Izv. Vses. Geogr. Obshch.* 85.
- Ivanov, N. (1959). Belts of continentality on the globe. *Izv. Vses. Geogr. Obshch.* 91, 410–423.
- Kalarus, M., Schuh, H., Kosek, W., Akyilmaz, O., Bizouard, C., Gambis, D., et al. (2010). Achievements of the Earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, 84, 587–596.
- Khromov, S. P. (1957). K voprosu o kontinentalnosti klimata, *Izv. Vses. Geogr. Obshch.* 89.
- Koźuchowski, K. (2003). Cyrkulacyjne czynniki klimatu Polski. *Czasopismo Geograficzne*, LXXIV, 1–2, 93–105.
- Koźuchowski, K. (2011). *Klimat Polski. Nowe Spojrzenie*. Warszawa: Wydawnictwo Naukowe PWN.
- Koźuchowski, K., & Marciniak, K. (1986). Fluktuacje kontynentalizmu klimatu Polski na tle warunków cyrkulacyjnych i solarnych (1881–1980). *Przegląd Geofizyczny*, 31(39), 139–152.
- Koźuchowski, K., & Marciniak, K. (1992). Kontynentalizm termiczny klimatu na obszarze Polski, *Wiad. IMGW XV(XXXVI)*, 4, 89–93.
- Koźuchowski, K., & Wibig, J. (1988). Kontynentalizm pluwiálny w Polsce: Zróżnicowanie geograficzne i zmiany wieloletnie. *Acta Geographica Lodziensia*, 55, 102.
- Kryza, M., Szymanowski, M., Mięła, K., & Pietras, M. (2010). Spatial information on total solar radiation: Application and evaluation of the r.sun model for the Wedel Jarlsberg Land, Svalbard, Pol. *Polar Research*, 31(1), 17–32.
- Li, J., & Heap, A. D. (2008). *A review of spatial interpolation methods for environmental scientists*. Canberra: Geoscience Australia.
- Marti, P., & Gasque, M. (2010). Ancillary data supply strategies for improvement of temperature-based ET₀ ANN models. *Agricultural Water Management*, 97, 939–955.
- Mikolášková, K. (2009). A regression evaluation of thermal continentality, *Geografie—Sborník České Geografické Společnosti*, 114(4), 350–362.
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1), 17–23.
- Nakaya, T. (2016). GWR4 User Manual. GWR 4 Windows Application for Geographically Weighted Regression Modelling, GWR4 Development Team. https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf.
- Nakaya, T., Fotheringham, A. S., Brunson, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease associative mapping. *Statistics in Medicine*, 24, 2695–2717.
- Oliver, J. (1970). An air mass evaluation of the concept of continentality. *The Professional Geographer*, 22, 83–87.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1988). *Numerical recipes in C, the art of scientific computing*. New York: Cambridge University Press.
- Raunio, N. (1948). The effect of local factors on meteorological observations at Tórshavn. *Geophysica*, 3, 173–179.
- Salonen, J. S., Seppä, H., Luoto, M., Bjune, A. E., & Birks, H. J. B. (2012). A North European pollen—climate calibration set: Analysing the climatic responses of a biological proxy using novel regression tree methods. *Quaternary Science Reviews*, 45, 95–110.
- Spitaler, R. (1922). Klimatische Kontinentalität und Ozeanität. *Petermann's Geographische Mitteilungen*, 68, 113.
- Šuri, M., & Hofierka, J. (2004). A new GIS-based solar radiation model and its application to photovoltaic assessments. *Transactions in GIS*, 8, 175–190.
- Swoboda, G. (1922). Linien gleicher Kontinentalität und Ozeanität 1. Weltkarte 2. Europa, Justus Perthes/Petermanns Mitteilungen, Gotha.
- Szrefel, C. (1961). Przegląd ważniejszych sposobów charakterystyki stopnia kontynentalizmu. *Przegląd Geograficzny* VI, 3, 191–199.
- Szymanowski, M., & Kryza, M. (2011). Application of geographically weighted regression for modelling the spatial structure of urban heat island in the city of Wrocław (SW Poland). *Procedia Environmental Sciences*, 3, 87–92.
- Szymanowski, M., & Kryza, M. (2012). Local regression models for spatial interpolation of urban heat island—an example from Wrocław, SW Poland. *Theoretical and Applied Climatology*, 108, 53–71.
- Szymanowski, M., & Kryza, M. (2015). The role of auxiliary variables in deterministic and deterministic-stochastic spatial models of air temperature in Poland. *Pure and Applied Geophysics*, doi:10.1007/s00024-015-1199-2.

- Szymanowski, M., Kryza, M., Smaza, M. (2007). A GIS approach to spatialize selected climatological parameters for wine-growing in Lower Silesia, Poland. In: Štrelcová, K., Škvarenina, J., Blaženc, M (Ed.) Bioclimatology and natural hazards, International Conference, Poľana nad Detvou, Slovakia, 17–20 September 2007, CD-ROM, ISBN 978-80-228-17-60-8.
- Szymanowski, M., Kryza, M., & Spallek, W. (2012). *Air temperature atlas for Poland: The methodical approach*. Wrocław: Uniwersytet Wrocławski. (in Polish, English summary).
- Szymanowski, M., Kryza, M., & Spallek, W. (2013). Regression-based air temperature spatial prediction models: An example from Poland. *Meteorologische Zeitschrift*, 22, 577–585.
- Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.*, 46, 234–240.
- Torregrosa, A., Taylor, M. D., Flint, L. E., & Flint, A. L. (2013). Present, future, and novel bioclimates of the San Francisco, California Region. *PLoS ONE*, 8(3), e58450. doi:10.1371/journal.pone.0058450.
- Ustrnul, Z., & Czekierda, D. (2005). Application of GIS for the development of climatological air temperature maps: An example from Poland. *Meteorological Application*, 12, 43–50.
- Ustrnul, Z., & Czekierda, D. (2009). *Atlas of extreme meteorological phenomena and synoptic situations in Poland*. Warszawa: IMGW.
- Willmott, C. J., & Matsuura, K. (1995). Smart interpolation of annually averaged air temperature in the United States. *Journal of Applied Meteorology*, 34, 2577–2586.
- Woś, A. (2010). *Klimat Polski w Drugiej Połowie XX Wieku*. Poznań: Wydawnictwo Naukowe UAM.

(Received September 22, 2014, revised September 22, 2016, accepted October 5, 2016, Published online October 19, 2016)