information

MDPI

*Article*

# Stochastic Model of Spatial Fields of the Average Daily Wind Chill Index

**Nina Kargapolova** [1,2]

[1] Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090 Novosibirsk, Russia; nkargapolova@gmail.com

[2] Department of Mathematics and Mechanics, Novosibirsk State University, 630090 Novosibirsk, Russia

Received: 20 February 2020; Accepted: 24 March 2020; Published: 26 March 2020


**Abstract:** The objective of this paper was to construct a numerical stochastic model of the spatial field of the average daily wind chill index on an irregular grid defined by the location of the weather stations. It is shown in the paper that the field in question was heterogeneous and non-Gaussian. A stochastic model based on the real data collected at the weather stations located in West Siberia and on the method of the inverse distribution function that sufficiently well reproduce different characteristics of the real field of the average daily wind chill index is proposed in this paper. I also discussed several questions related to the simulation of the field on a regular grid. In the future, my intention is to transform the model proposed to a model of the conditional spatio-temporal field defined on a regular grid that allows one to forecast the wind chill index.

**Keywords:** stochastic simulation; spatial random field; non-Gaussian random process; wind chill index; West Siberia

## 1. Introduction

During the last decades there arose many discussions about how to determine the thermal comfort and how to grade the thermal stress. These efforts have resulted in various models attempting to describe the thermal comfort and the resultant thermal stress (one of the first models was proposed by P.A. Siple and C.F. Passel in 1945 [1]). A large number of the so-called biometeorological indices that describe the combined thermal effect of various meteorological parameters (the wind speed, the air temperature and the relative humidity, the atmospheric pressure, etc.) on human beings have been proposed. In [2–5] one may find a comprehensive review of such biometeorological indices.

The most common approach to studying the properties of the time series and spatial/spatio-temporal fields of bioclimatic indices is the statistical approach (see, for example, [6–10]). The statistical approach could be easily transformed to a stochastic one. The obtained stochastic models may be used both to study the properties of a bioclimatic index and to forecast it. For example, in [11,12] two stochastic models of the bioclimatic index of severity of climatic regime are proposed and in [13–15] the same approach was used for the time series of the heat index.

In this paper, a stochastic approach was for the first time applied to studying the spatial fields of the average daily wind chill index (WCI). Earlier, this approach was successfully applied to simulation and forecasting of the WCI time series [16,17]. The field of the WCI was simulated on an irregular grid determined by the location of the weather stations. Several questions related to the simulation of the WCI spatial field on a regular grid are discussed. Both to determine the model input parameters and to verify the model proposed, the real data collected at the weather stations located in West Siberia were used.

There is a wide range of methods for simulation of Gaussian or non-Gaussian random processes and fields [18–26]. The problem is that most of the methods could be used only for simulation of some

narrow types of processes. For instance, the method proposed in [19] allows one to generate random fields with a marginal gamma distribution only and the "repetition" method (see, for example, [18,20]) requires the same one-dimensional distributions for all the components of the process. In this paper, to simulate the heterogeneous non-Gaussian fields of the average daily WCI, the method of the inverse distribution function was used [18,26]. This method is very flexible. The correlation matrix of the simulated process could be almost arbitrary (it may be neither Toeplitz nor block-Toeplitz, the correlation coefficients may not decrease exponentially and could be both positive and negative) and the marginal distributions of the process may be time/space-dependent.

This paper is organized as follows. In Section 2 the definition of the WCI is given. Section 3 describes the real data used for determining the model input parameters and for verification of the model. The description of the stochastic model and simulation algorithms is given in the Section 4. The verification results and the conclusion are presented in Sections 5 and 6, respectively.

## 2. The Wind Chill Index

The idea of the wind chill index was introduced by P.A. Siple and C.F. Passel in 1945 [1] after the United States Antarctic Expedition. Both the impromptu experiments that were conducted during the expedition (the results of these experiments have formed the basis of the WCI concept) and the proposed formulas for calculation of the WCI were criticized by many of the researchers (see, for example, [27,28]). Nevertheless, according to the [29], the WCI has become a standard meteorological term that is used to represent the hypothetical air temperature that would, under conditions of no wind, effect the same heat loss from unclothed human skin as does the actual combination of air temperature and wind speed. In [30], one may find some research background related to innovative formulas to estimate WCI. Different formulas for calculating the WCI and the restrictions under which these formulas could be applied are presented there.

In this paper, I used the definition of the WCI presented in [2,31,32]. The wind chill index (based on the Celsius temperature scale) is a function of the air temperature $T$ °C and the wind speed $V$ $km/h$ at 10 m above the ground. The WCI $W$ °C is defined as follows:

$$W = 13.12 + 0.6215T - 11.37V^{0.16} + 0.3965TV^{0.16}, \quad (1)$$

if the air temperature is below 10 °C and the wind speed is above 5 $km/h$. In the case of the very low wind speeds, that is, less than 5 $km/h$, the WCI is either not calculated at all or is defined according to the formula [33]

$$W = T + \frac{-1.59 + 0.1345T}{5}V. \quad (2)$$

If the air temperature is above 10 °C, the WCI is not calculated. In [2], one may find a table that presents the dependence of the health risk on the values of the WCI. The lower the WCI, the higher is the risk of frostbite and hypothermia. For example, if $-27 < W < -10$, there is a low risk of frostbite and the hypothermia is possible if a person is outside for long periods without adequate protection. Meanwhile, if $-47 < W < -40$, there is a high risk of frostbite: Exposed skin can freeze in 5–10 min.

## 3. The Real Data Description and Analysis

Both to determine the model parameters and to verify the model, the real data collected at the weather stations located in the southern part of West Siberia (Russia) from 1988 to 2017 were used [34,35]. I studied the properties of the average daily WCI and constructed the proper stochastic models for the two areas. The first area (Area 1) is located between 54°54′ and 57°04′ north latitude, 82°04′ and 86°13′ east longitude. The size of the area is approximately 260 km × 240 km. There are 7 weather stations in this area: Bakchar (the weather station with the World Meteorological Organization index 29328), Pervomajskoe (29348), Tomsk (29430), Bolotnoe (29539), Tajga (29541), Ogurcovo (29638), and Kemerovo (29645). The second area (Area 2) is larger: It is situated between 53°20′ and 56°20′ north

latitude, 76°01′ and 88°18′ east longitude (the size of the area is approximately 790 km × 330 km). Weather stations Severnoe (29418), Bolotnoe (29539), Tajga (29541), Tisul′ (29557), Tatarsk (29605), Barabinsk (29612), Ogurcovo (29638), Kemerovo (29645), Kamen′-na-Obi (29822), Barnaul (29838), and Kuzedeevo (29849) are located in the Area 2. The station Ogurcovo is situated in the city of Novosibirsk. Figure 1 shows the positions of the weather stations in the areas in question.

It should be noted that the WCI is not measured at a weather station, but it could be easily calculated with the above-given definition, Equations (1) and (2), using the values of the air temperature and the wind obtained at the station. At all the weather stations considered, the air temperature and the wind speed data were collected eight times per day at the synoptic hours (00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00, and 21:00 UTC). So, for each day, when the observations were made, eight values of the WCI could be calculated. These eight values of the WCI were used to determine the real average daily WCI. Hereinafter, the random field $\vec{W} = (W_1, W_2, \ldots, W_{NS})$ of the WCI (both real and simulated) will be considered, where $NS$ is the number of stations in the area in question, $W_i$ is the average daily WCI at the station number $i$ in the given day.

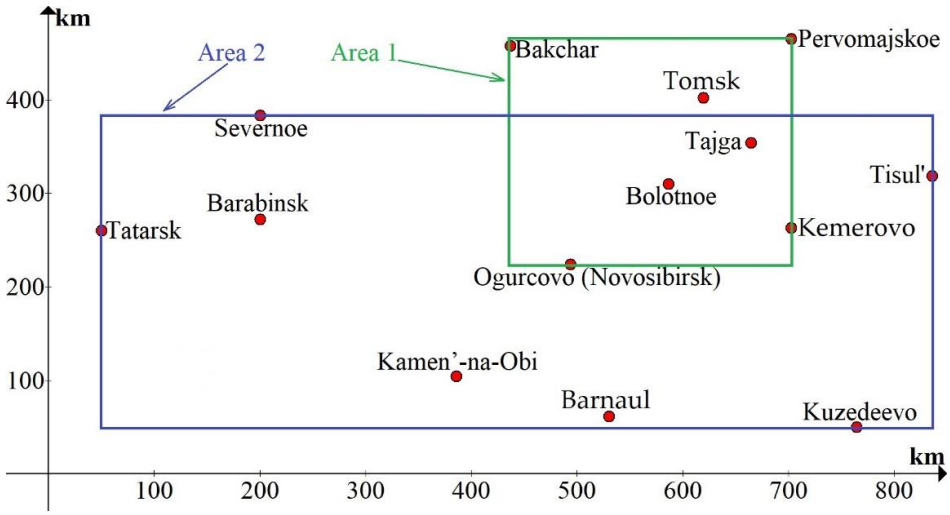



**Figure 1.** Position of the weather stations in the Area 1 and Area 2.

In order to construct a proper stochastic model, one should pay attention to the fact that the real fields $\vec{W}$ of the WCI are heterogeneous. Figure 2 shows examples of the sample mean of the average daily WCI at the weather stations in the Area 1. One can easily see that the sample mean varies from one station to another. The properties of the spatial field of the average daily WCI are time-dependent also. This should be taken into account when one constructs a model of a spatio-temporal field. Since in this paper only spatial fields were studied, this time dependence is out of range of my interests.

The spatial fields of the average daily WCI were strongly correlated. Figure 3 shows the correlation coefficients $corr\left(W_i, W_j\right)$ of the field $\vec{W}$ as a function of the distance between the stations $i$ and $j$. One can see that the correlation coefficients even between remote stations were significant. As a comparison, Figure 4 shows the correlations coefficients $corr\left(T_i, T_j\right)$ and $corr\left(V_i, V_j\right)$ of the random fields of the average daily temperature and the average daily wind speed, respectively. The correlation function of the wind speed field decreased faster than the correlation of the WCI field, and the decrease rates of the correlation functions of the temperature and WCI fields were similar. To reduce the statistical error of the estimation of all the characteristics considered over a small sample of the real data, the sample size was increased artificially. To do it, the moving averaging procedure with a symmetric two-side smoothing window with a width of $2L + 1$ days was applied. To estimate the characteristic for the given day, data collected in that day, data collected during $L$ days before the given day, and $L$ days after the day in question were used. In this paper, $L = 2$.

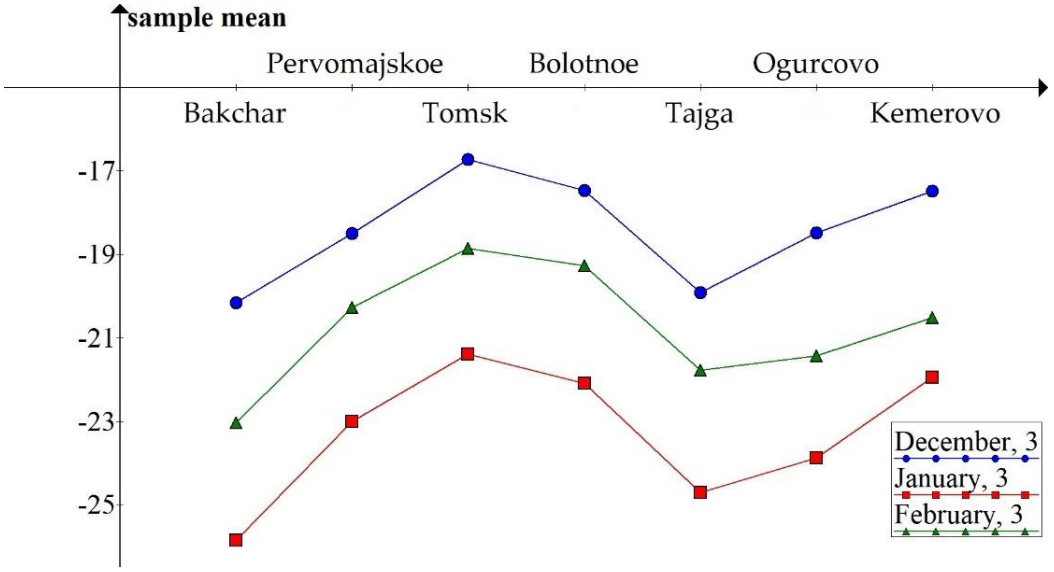



**Figure 2.** Sample mean of the average daily wind chill index (WCI), Area 1.

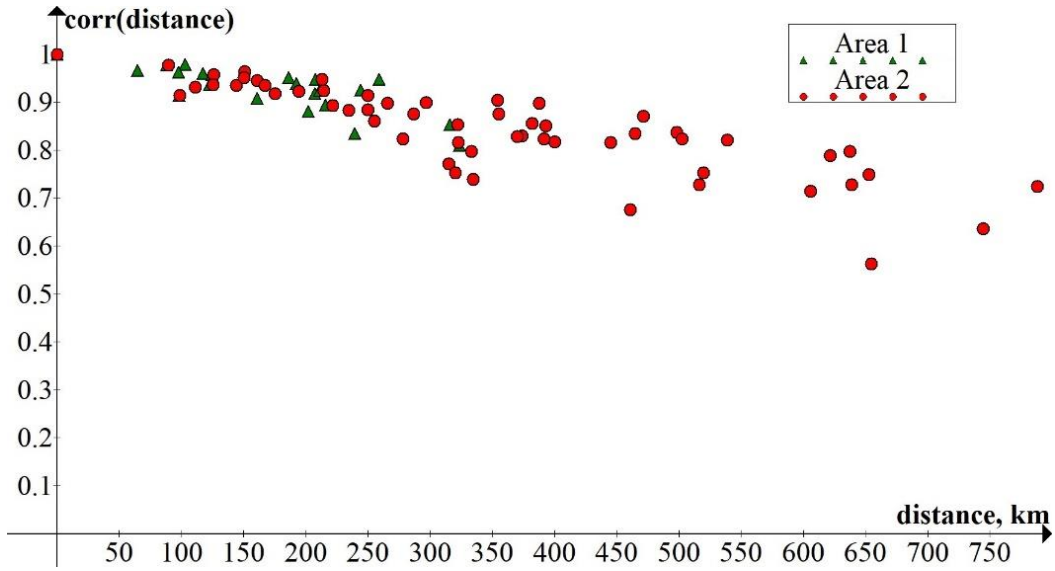
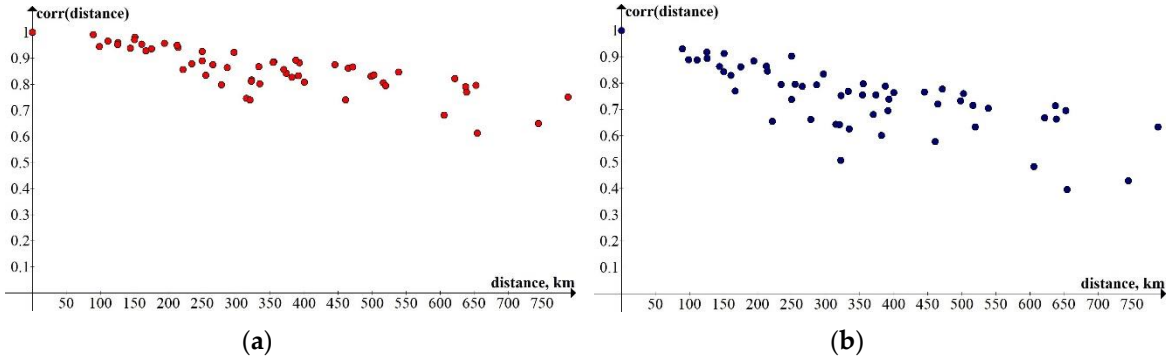



**Figure 3.** Correlation coefficients of the average daily WCI as a function of distance, 16 January.




(**a**) (**b**)

**Figure 4.** Correlation coefficients of (**a**) the average daily temperature and (**b**) the average daily wind speed as a function of distance, 16 January, Area 2.

## 4. Stochastic Model

The method of the inverse distribution function (MIDF) is used for the simulation of the random field of the average daily WCI [18,26,36]. This method allows one to simulate a non-Gaussian process with the given one-dimensional distributions and the given correlation structure. It is shown in [16] that the mixtures $g_k(x), k = \overline{1, NS}$ of the two Gaussian distributions will sufficiently approximate the sample histograms of the real average daily WCI. The correlation structure of the field $\vec{W}$ is defined by the $NS \times NS$ correlation matrix $R = \{r(i,j)\}_{i,j=1}^{NS}$ that is estimated on the real data. Let us recall the idea of the MIDF. In the framework of the MIDF, simulating of the field $\vec{W}$ with the given densities $g_k(x), k = \overline{1, NS}$ and the given correlation matrix $R$ comes to an algorithm (let us call it the Algorithm 1) with the three steps [11,16]:

---

**Algorithm 1.**

---

- Step 1. At this step, a correlation matrix $R' = \{r'(i,j)\}_{i,j=1}^{NS}$ of an auxiliary standard Gaussian process $\vec{W'} = (W'_1, W'_2, \ldots, W'_{NS})$ is calculated. The entry $r'(i,j)$ of the matrix $R'$ is the solution of the equation

$$r(i,j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i^{-1}(\Phi(x)) F_j^{-1}(\Phi(y)) \varphi(x, y, r'(i,j)) dx dy, \tag{3}$$

  where the function $\varphi(x, y, r'(i,j))$ is the distribution density of a bivariate Gaussian vector with zero mean, variance equal to 1, the correlation coefficient $r'(i,j)$, $\Phi(\cdot)$ is a cumulative distribution function (CDF) of a standard normal distribution, and $F_i(\cdot)$ and $F_j(\cdot)$ are the CDFs corresponding to the densities $g_i(x)$ and $g_j(x)$, respectively.

- Step 2. After calculating the matrix $R'$, a trajectory of the auxiliary standard Gaussian field $\vec{W'}$ with the correlation matrix $R'$ is simulated. This could be done using the Cholesky method or the spectral decomposition of the matrix $R'$ (see [18]).

- Step 3. Finally, the trajectory of the Gaussian vector $\vec{W'}$ is transformed to the trajectory of the non-Gaussian vector $\vec{W}$:

$$W_i = F_i^{-1}(\Phi(W'_i)), \quad i = \overline{1, NS}. \tag{4}$$

---

One should note that there is no analytical solution of the Equation (3). This is why one have to solve this equation numerically. Since the integral in the right part of Equation (3) is a continuous monotonically increasing function on [−1;1] (see, for example, [18]), the simplest method for solving the Equation (3) is the numerical inversion of the integral as a function of the correlation coefficient using its tabulated values. In this paper, the bisection method was used. If the matrix $R'$, obtained at the first step using the Equation (3), is not positive definite, it must be regularized. In this paper, a method of regularization based on the substitution of negative eigenvalues of the matrix $R'$ with small positive numbers was used [18]. After the regularization, the matrix $R'$ required normalization. Steps 2 and 3 were repeated as many times as trajectories were required. For the transformation Equation (4) the tabulated values of $F_i(\cdot), \quad i = \overline{1, NS}$ were used.

Let me list the main advantages of the model proposed.

- Many of the dynamic models that are used to simulate the fields of the air temperature and other meteorological parameters (using these models one may calculate the WCI) require as the input information some data related to the geographical properties of the area (whether the area is

mountainous or plain, are there any huge water bodies or not, the elevation of the weather stations above the sea level, etc.). The stochastic approach does not require this additional information.

- The most time-consuming step of the above-described algorithm is Step 1. Step 2 and Step 3 do not require a lot of time to be conducted. Once the matrix $R'$ is calculated, one may repeat Steps 2 and 3 as many times as needed to obtain the required accuracy of the estimations, and the increase in accuracy insignificantly influences the total computational time.
- The model proposed could be easily transformed to a model of a conditional random field of the average daily WCI. For the time-series of the WCI, such transformation is presented in [17]. The conditional model allows one to forecast the bioclimatic index in question.

If one simulates the field of the average daily WCI only at the weather stations, the above-described approach could be used (let us call it the Model 1). In the next section, the results of the verification of the Model 1 are given. However, if it is necessary to simulate the field not only at the weather stations but also in the nodes of some regular or irregular grid, one must determine the distribution density of the WCI in each node and determine the $(NS + NN) \times (NS + NN)$ correlation matrix of the field, where $NN$ is a number of nodes in the grid. There are a lot of methods (both deterministic and stochastic) for determining the distribution parameters in a node of the grid. For example, one may calculate the parameters using the interpolation from the nearest neighbor, triangulation, different types of kriging, conditional interpolation technique, etc. (see, for example, [37,38]). The most common approach to determining the $(NS + NN) \times (NS + NN)$ correlation matrix of the field is approximating the sample correlation coefficients (that form the matrix $R$) with an analytical function of the continuous argument and calculating the correlation coefficients between each pair of nodes as a value of this function in the corresponding points [39,40]. For approximating of the correlation coefficients of the meteorological time series and fields functions

$$corr_a\left(x_i, x_j, y_i, y_j\right) = corr(\rho) = \exp\left(-\alpha\rho^{\beta}\right), \quad \alpha, \beta > 0 \tag{5}$$

$$corr_a\left(x_i, x_j, y_i, y_j\right) = corr(\rho) = (1 + \alpha\rho)\exp(-\alpha\rho), \quad \alpha > 0 \tag{6}$$

$$corr_a\left(x_i, x_j, y_i, y_j\right) = corr(\rho) = \frac{\sin\beta\rho}{\beta\rho}\exp(-\alpha\rho), \quad \alpha, \beta > 0 \tag{7}$$

$$corr_a\left(x_i, x_j, y_i, y_j\right) = corr(\rho) = \cos(\beta\rho)\exp(-\alpha\rho), \quad \alpha, \beta > 0 \tag{8}$$

$$corr_a\left(x_i, x_j, y_i, y_j\right) = \exp\left(-\alpha\left(a\left(x_i - x_j\right)^2 + b\left(x_i - x_j\right)\left(y_i - y_j\right) + c\left(y_i - y_j\right)^2\right)^{\beta}\right), \quad \alpha, a, c > 0, \ 0 < \beta < 4, \ b \in \mathbb{R} \tag{9}$$

are often used. Here $\rho = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ is a distance between two points with the coordinates $(x_i, y_i)$ and $\left(x_j, y_j\right)$. The correlation functions defined by Equations (5)–(8) depend only on the distance between points and do not depend on the exact position of the points $(x_i, y_i)$ and $\left(x_j, y_j\right)$ in the considered area or on the direction of the vector $\left(x_i - x_j, y_i - y_j\right)$. The function Equation (9) depends both on distance and on direction and does not depend on the parallel shift of the vector $\left(x_i - x_j, y_i - y_j\right)$. The numerical experiments show that the sample correlation coefficients of the field of the average daily WCI were approximated sufficiently well with the function Equation (5) when $\beta = 2$ and with the function Equation (9). To choose the parameters of the approximating functions, the functional

$$\sum_{i=1}^{NS}\sum_{j=i+1}^{NS}\left(r(i,j) - corr\left(x_i, x_j, y_i, y_j\right)\right)^2 \tag{10}$$

was minimized on the condition that the function $corr_a\left(x_i, x_j, y_i, y_j\right)$ was positive definite.

Before one starts to simulate the field of the WCI on a grid with the correlation functions Equations (5) or (9), it is necessary to study how such an approximation influences the quality of the simulation of the field at the weather stations. Let the Model 2 and Model 3 be the models of the field of the WCI based on the MIDF, in which the matrix $R$ was substituted with the matrices calculated with the Equations (5) and (9) with the parameters that minimize the functional Equation (10). In the next section, the results of the verification of the Model 2 and the Model 3 are given.

## 5. Verification of the Model

To verify a model, it was necessary to compare estimations of various characteristics that were based on the simulated and real data. Only such characteristics must be considered that, on the one hand, were reliably estimated by means of real data and, on the other hand, were not input parameters of the model. In this section, several examples of such characteristics used for verification of the Models 1–3 are given.

It should be noted that it is possible to simulate as many trajectories as needed to provide a required accuracy of the estimation when the simulated data are used. In this paper, for all estimations based on the simulated data, I attained the accuracy above 0.0001. Thus, in all the tables presented, the estimations based on the simulated trajectories are given with significant digits only.

From now on, $\sigma$ is a standard deviation of the characteristic under consideration when estimating with the real data. Since for the Models 1–3 only one-dimensional distributions of the field and its correlation structure were defined, it was not possible to write down theoretical formulas for the variance (and, therefore, for $\sigma$) of the characteristics of the field related to its multi-dimensional distributions. This is why the values $\sigma$, presented in tables below, were numerically estimated.

The first characteristic considered was the average number $AN(l)$ of stations in the area where the average daily WCI was above the given level $l$. The estimations of the $AN(l)$ on the real and simulated data are presented in the Tables 1 and 2. The number $AN(l)$ did not depend on the correlation structure of the field, this is why the estimations of this characteristic on the trajectories obtained with the Models 1–3 were the same. The values of $AN(l)$ was defined by the one-dimensional distributions in each point of the field. For both Area 1 and Area 2 and for all the considered days, the estimations of the $AN(l)$ based on the simulated data belonged to the confidence intervals $(AN(l) - \sigma, AN(l) + \sigma)$ corresponding to the real data. This fact confirmed that the mixtures of the two Gaussian distributions, used for approximation of the sample histograms, described sufficiently well the one-dimensional distributions of the real field.

**Table 1.** The average number, $AN(l)$, of stations in the area where the average daily WCI was above the given level $l$, Area 1, 15 January.

| Level $l$ °C | Real Data, $AN(l) \pm \sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| −21 | 3.660 ± 0.248 | 3.586 | 3.586 | 3.586 |
| −27 | 5.227 ± 0.203 | 5.167 | 5.167 | 5.168 |
| −33 | 6.307 ± 0.126 | 6.368 | 6.368 | 6.368 |
| −39 | 6.867 ± 0.056 | 6.871 | 6.871 | 6.871 |
| −45 | 6.987 ± 0.016 | 6.986 | 6.986 | 6.986 |
| −51 | 7.000 ± 0.003 | 6.999 | 6.999 | 6.999 |
| −57 | 7.000 ± 0.001 | 7.000 | 7.000 | 7.000 |

The next characteristic used for the verification of the models was the probability $p(l) = P(W_1 < l, W_2 < l, \ldots, W_{NS} < l)$. This characteristic was closely related to the $NS-$ dimensional distribution of the average daily WCI. Tables 3 and 4 show the examples of the corresponding estimations, obtained with the real and simulated data. When I considered the relatively high levels $l$, where the estimating of $p(l)$ was reliable (lines 1–5 in Tables 3 and 4), in approximately 95% of numerical experiments performed the absolute difference of $p(l)$ assessed with real data and data simulated with the Model 1 and Model 3 was less than $\sigma$, and this difference never exceeded $2\sigma$. The deviations of

the simulated with the Model 2 data estimations of $p(l)$ from the corresponding estimations based on the real data did not exceed $\sigma$ and $2\sigma$ in 86% and 94% of the experiments conducted, respectively. This means that all three models reproduced well the characteristic in question. In the last four lines (lines 6–9) in the Tables 3 and 4 the estimations of $p(l)$ for the extremely low values of $l$ are given. The estimations with the real data were statistically unreliable and, even more, if we consider the real data, the event "$W_1 < l, W_2 < l, \ldots, W_{NS} < l$ when $l < -37$" never happened. One could see that the Models 1–3 gave realistic results. This results were close to each other because the only difference between these models was the type of the correlation function (that characterizes the two-dimensional distributions of the field), and $p(l)$ was related to the $NS-$ dimensional distribution of the WCI field.

**Table 2.** The average number $AN(l)$ of stations in the area where the average daily WCI was above the given level $l$, Area 2, 15 January.

| Level $l$ °C | Real Data, $AN(l) \pm \sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| −21 | 5.813 ± 0.369 | 5.684 | 5.684 | 5.685 |
| −27 | 8.407 ± 0.308 | 8.350 | 8.347 | 8.350 |
| −33 | 10.093 ± 0.182 | 10.179 | 10.178 | 10.179 |
| −39 | 10.860 ± 0.060 | 10.870 | 10.870 | 10.870 |
| −45 | 10.987 ± 0.015 | 10.989 | 10.989 | 10.989 |
| −51 | 11.000 ± 0.001 | 11.000 | 11.000 | 11.000 |

**Table 3.** Probability $p(l)$, Area 1, 5 December.

| Line | Level $l$ °C | Real Data, $p(l) \pm \sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|---|
| 1 | −3 | 0.980 ± 0.013 | 0.977 | 0.978 | 0.976 |
| 2 | −6 | 0.927 ± 0.023 | 0.913 | 0.916 | 0.910 |
| 3 | −9 | 0.707 ± 0.035 | 0.765 | 0.772 | 0.762 |
| 4 | −12 | 0.527 ± 0.039 | 0.566 | 0.576 | 0.563 |
| 5 | −15 | 0.367 ± 0.046 | 0.396 | 0.405 | 0.392 |
| 6 | −33 | 0.040 ± 0.436 | 0.027 | 0.028 | 0.026 |
| 7 | −36 | 0.013 ± 0.498 | 0.012 | 0.013 | 0.011 |
| 8 | −39 | 0.000 ± 0.572 | 0.004 | 0.004 | 0.004 |
| 9 | −42 | 0.000 ± 0.837 | 0.000 | 0.000 | 0.000 |

**Table 4.** Probability $p(l)$, Area 2, 5 December.

| Line | Level $l$ °C | Real Data, $p(l) \pm \sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|---|
| 1 | −3 | 0.967 ± 0.017 | 0.958 | 0.962 | 0.957 |
| 2 | −6 | 0.873 ± 0.027 | 0.866 | 0.877 | 0.863 |
| 3 | −9 | 0.633 ± 0.039 | 0.684 | 0.702 | 0.680 |
| 4 | −12 | 0.440 ± 0.041 | 0.463 | 0.483 | 0.459 |
| 5 | −15 | 0.287 ± 0.057 | 0.288 | 0.308 | 0.282 |
| 6 | −30 | 0.027 ± 0.541 | 0.013 | 0.015 | 0.013 |
| 7 | −33 | 0.000 ± 0.602 | 0.005 | 0.005 | 0.005 |
| 8 | −36 | 0.000 ± 0.699 | 0.001 | 0.001 | 0.001 |
| 9 | −39 | 0.000 ± 0.781 | 0.000 | 0.000 | 0.000 |

Another characteristic used for the comparison of the real and the simulated fields of the average daily WCI was the probability $pr(k,l)$ that at least at $k$ stations the average daily WCI did not exceed the level $l$. Figure 5 shows the estimations of $pr(k,l)$ for several values of $k$ and $l$. One can clearly see that the simulation data based on estimations of $pr(k,l)$ rapidly decreased both as functions of $k$ and as functions of $l$. Nonmonotonicity of the real data based on estimations is explained with the statistical error of the estimation. For both areas in question, for all the considered days, values of $k$, and levels $l$, the estimations of the $pr(k,l)$ based on the data simulated with the Models 1–3 belonged to the confidence intervals $(pr(k,l) - 3\sigma, pr(k,l) + 3\sigma)$, corresponding to the real data.

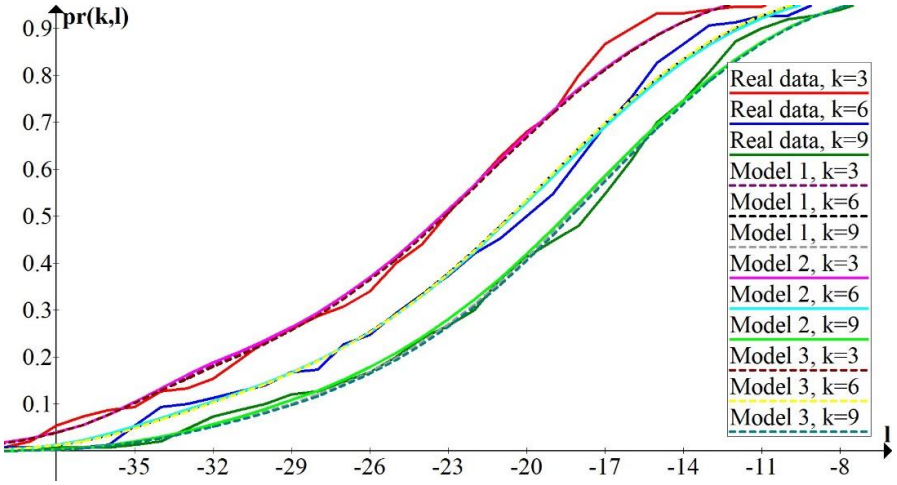



**Figure 5.** Probability $pr(k,l)$ that at least at $k$ stations the average daily WCI did not exceed the level $l$, Area 2, 1 February.

Let us consider one more characteristic used for verification of the stochastic models. Let $pd(station_1, station_2, \Delta)$ be a probability of the event "the average daily WCI at the $station_1$ and $station_2$ differs by more than $\Delta$ °C", i.e.,

$$pd(station_1, station_2, \Delta) = P\left(\left|W_{station\ 1} - W_{station\ 2}\right| > \Delta\right). \quad (11)$$

Tables 5 and 6 show the examples of the corresponding estimations obtained with the real and simulated data. The results in these tables are given for the pairs of stations Bolotnoe and Ogurcovo and Bolotnoe and Kemerovo when the simulation was conducted for the Area 1. The numerical experiments showed that all of the three models in question more or less precisely reproduced this characteristic of the real field. The estimations of the probability Equation (11), calculated with the trajectories obtained with the Model 3, approximately 1.2–1.3 times more often belonged to the confidence interval $(pd(station_1, station_2, \Delta) - 2\sigma, pd(station_1, station_2, \Delta) + 2\sigma)$ than the estimations based on the Model 2. The possible explanation is that the characteristic in question was closely related to the correlation structure of the field, and the correlation function Equation (9) used in the Model 3 was closer (in the sense of the functional Equation (10)) to the sample correlation coefficients of the real field of the average daily WCI. Figure 6 presents the estimations of $pd(station_1, station_2, \Delta)$ for the pairs of stations Bakchar and Tomsk and Pervomajskoe and Bolotnoe. One can see that, despite the fact that all four stations are situated in the same climatic zone and the distance between Bakchar and Tomsk is almost equal to the distance between Pervomajskoe and Bolotnoe, there is a significant distinction between the decreased rates of $pd(\cdot, \cdot, \Delta)$ as a function of $\Delta$. The most probable explanation of this fact is that the correlation coefficient between the average daily WCI at the weather stations Pervomajskoe and Bolotnoe was higher than the correlation between the average daily WCI in Bakchar and Tomsk, and, therefore, the probability of the sufficient difference in values of the WCI at the first pair of stations was less than the corresponding probability at the second pair of stations.

**Table 5.** Probability $pd(Bolotnoe, ogurcovo, \Delta)$, Area 1, 15 January.

| Difference Δ °C | Real Data, $pd(Bolotnoe,Ogurcovo,\Delta) \pm 2\sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| 1 | 0.693 ± 0.074 | 0.691 | 0.685 | 0.699 |
| 2 | 0.467 ± 0.079 | 0.444 | 0.436 | 0.457 |
| 3 | 0.320 ± 0.073 | 0.149 | 0.261 | 0.283 |
| 4 | 0.153 ± 0.058 | 0.270 | 0.140 | 0.162 |
| 5 | 0.060 ± 0.044 | 0.072 | 0.065 | 0.084 |
| 6 | 0.020 ± 0.029 | 0.030 | 0.026 | 0.038 |

**Table 6.** Probability $pd(Bolotnoe, Kemerovo, \Delta)$, Area 1, 15 January.

| Difference Δ °C | Real Data, $pd(Bolotnoe,Kemerovo,\Delta) \pm 2\sigma$ | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|
| 1 | 0.687 ± 0.075 | 0.713 | 0.623 | 0.664 |
| 2 | 0.420 ± 0.084 | 0.471 | 0.335 | 0.395 |
| 3 | 0.260 ± 0.076 | 0.292 | 0.159 | 0.214 |
| 4 | 0.173 ± 0.061 | 0.171 | 0.067 | 0.107 |
| 5 | 0.113 ± 0.047 | 0.094 | 0.025 | 0.048 |
| 6 | 0.053 ± 0.034 | 0.048 | 0.008 | 0.019 |

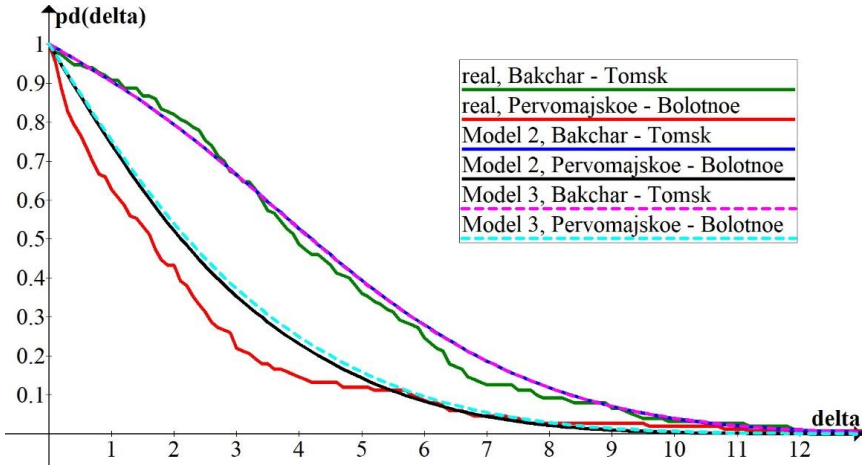



**Figure 6.** Probability $pd(station_1, station_2, \Delta)$, Area 1, 15 January.

In this section, four characteristics used for the verification of the models were presented. For each of the characteristics only several values of the arguments were considered. For instance, the probability $pd(\cdot, \cdot, \Delta)$ was presented only for four pairs of stations (Bolotnoe-Ogurcovo, Bolotnoe-Kemerovo, Bakchar-Tomsk, and Pervomajskoe-Bolotnoe), though for each of the areas there were $NS(NS-1)/2$ pairs of stations. In fact, for both Area 1 and Area 2 I considered 15 different characteristics. For each of the characteristic in question, I compared the simulated and the real data based on estimations for all possible values of the arguments (if the number of possible values was finite) or, if an argument was continuous (like $l$ in $AN(l)$), a fine grid with a small step in this argument was used for the verification. The results of this detailed numerical analysis showed that the trajectories obtained with the models proposed were close in their statistical properties to the real fields of the average daily wind chill index. Therefore, these models could be used for studying of those properties of the field that are unreliably estimated by means of real data. Since the Model 3 reproduced the characteristics of the real field a little more accurately than the Model 2, the Model 3 will be used for simulating of the spatial and spatio-temporal fields of the average daily wind chill index on a regular grid.

## 6. Conclusions

For the first time ever, the stochastic approach was used for studying and simulation of the random fields of the average daily wind chill index. It was shown in this paper that the spatial fields in question, defined as two areas located in West Siberia, were heterogeneous and non-Gaussian. I proposed a stochastic model that fairly well reproduces the statistical properties of the real spatial fields. This model could be used for estimating of those characteristics of the random field that could not be estimated from the real data in view of the small sample size of the real data.

In the future, our intention is to transform the proposed model to a model of the conditional spatio-temporal field defined on a regular grid that allows one to forecast the wind chill index. To achieve this goal, it is necessary to solve the next few problems. The first one is to determine such interpolation technique that allows one to interpolate values of the WCI from the weather stations to the grid nodes with the smallest error of the interpolation. The interpolation should

let one take into account both spatial and temporal variability of the WCI. The next problem we have to solve is simulation of the nonstationary heterogeneous non-Gaussian field on a dense grid. Simulation of such a process on a grid with a large number of nodes is a time-consuming procedure and it requires also a lot of computer memory. It is quite probable that it will be necessary to substitute the Algorithm 1 with some approximate algorithm with the lesser computational complexity (see, for example, [23,24]). Finally, it would be necessary to transform the non-conditional model to a conditional one. This transformation is rather simple if the conditions are point conditions, when several values (either consecutive or inconsecutive) of the WCI are given. Since there is no efficient algorithm to simulate the conditional Gaussian time series with the interval conditions, the transformation of the non-conditional model to a conditional one is much trickier and requires further research.

**Funding:** This work was partly financially supported by the Russian Foundation for Basic Research (grant No. 18-01-00149-a), by the Russian Foundation for Basic Research and the Government of the Novosibirsk region according to the research project No. 19-41-543001-r_mol_a.

**Conflicts of Interest:** The author declares no conflict of interest.

## Nomenclature

| | |
|---|---|
| Algorithm 1 | An algorithm for simulation of the heterogeneous non-Gaussian random field based on the MIDF |
| Area 1 | An area located between 54°54′ and 57°04′ north latitude, 82°04′ and 86°13′ east longitude. The size of the area is approximately 260 km × 240 km. There are 7 weather stations in this area. |
| Area 2 | An area located between 53°20′ and 56°20′ north latitude, 76°01′ and 88°18′ east longitude. The size of the area is approximately 790 km × 330 km. There are 11 weather stations in this area. |
| CDF | The cumulative distribution function |
| MIDF | The method of the inverse distribution function |
| Model 1 | The model of the field of the WCI based on the MIDF with the correlation matrix $R$ and distribution densities $g_i(x)$, $i = \overline{1, NS}$ |
| Model 2 | The model of the field of the WCI based on the MIDF, in which the matrix $R$ is substituted with the matrices calculated with the Formula (5) |
| Model 3 | The model of the field of the WCI based on the MIDF, in which the matrix $R$ is substituted with the matrices calculated with the Formula (9) |
| UTC | Coordinated Universal Time |
| WCI | Wind chill index |
| $\alpha, \beta$ | Parameters of the approximating functions (5)–(9) |
| $\Phi(\cdot)$ | CDF of a standard normal distribution |
| $\varphi(x, y, r'(i, j))$ | Distribution density of a bivariate Gaussian vector with zero mean, variance equal to 1 and the correlation coefficient $r'(i, j)$ |
| $\rho$ | Distance between points with the coordinates $(x_i, y_i)$ and $(x_j, y_j)$. |
| $\sigma$ | Standard deviation of the characteristic under consideration when estimating with the real data. |
| $AN(l)$ | Average number of stations in the area where the average daily WCI is above the given level $l$ |
| $a,\ b,\ c$ | Parameters of the approximating function (9) |
| $corr(T_i, T_j)$, $corr(V_i, V_j)$, $corr(W_i, W_j)$ | Correlation coefficients between air temperature, wind speed and WCI at the stations $i$ and $j$, respectively |
| $corr_a(x_i, x_j, y_i, y_j)$ | Approximating correlation function |
| $F_i(\cdot)$ | CDF corresponding to the density $g_i(x)$ |
| $g_i(x)$ | Distribution density |
| $i,\ j,\ k$ | Integer numbers used for enumeration of the weather stations |
| $L$ | Parameter that determines the width of a symmetric two-side smoothing window |
| $NN$ | Number of a grid nods |
| $NS$ | Number of stations in the area |
| $p(l)$ | Probability $P(W_1 < l, W_2 < l, \ldots, W_{NS} < l)$ |
| $pd(station_1, station_2, \Delta)$ | Probability of the event "the average daily WCI at the $station_1$ and $station_2$ differs by more than $\Delta$ °C" |
| $pr(k, l)$ | Probability that at least at $k$ stations the average daily WCI does not exceed the level |
| $R = \{r(i, j)\}_{i,j=1}^{NS}$ | Correlation matrix of the random field that is estimated on the real data |
| $R' = \{r'(i, j)\}_{i,j=1}^{NS}$ | Correlation matrix of an auxiliary standard Gaussian process |
| $T$ °C, $V$ km/h, $W$ °C | Air temperature, wind speed, WCI |
| $\vec{W} = (W_1, W_2, \ldots, W_{NS})$ | Random field of the WCI |
| $\vec{W}' = (W'_1, W'_2, \ldots, W'_{NS})$ | Auxiliary standard Gaussian process |
| $(x_i, y_i)$ | Coordinates of a point $i$ |

## References


1. Siple, P.A.; Passel, C.F. Measurements of dry atmospheric cooling in sub-freezing temperatures. *Proc. Am. Philos. Soc.* **1945**, *89*, 177–199.
2. Blazejczyk, K.; Epstein, Y.; Jendritzky, G.; Staiger, H.; Tinz, B. Comparison of UTCI to selected thermal indices. *Int. J. Biometeorol.* **2012**, *56*, 515–535. [CrossRef] [PubMed]
3. Kobisheva, N.V.; Stadnik, V.V.; Klueva, M.V.; Pigoltsina, G.B.; Akentieva, E.M.; Galuk, L.P.; Razova, E.N.; Semenov, U.A. *Guidance on Specialized Climatological Service of the Economy*, 1st ed.; Asterion: Saint Petersburg, Russia, 2008. (In Russian)
4. McGregor, G.R. (Ed.) *Heatwaves and Health: Guidance on Warning-System Development*, 1st ed.; WMO: Geneva, Switzerland, 2015.
5. Zare, S.; Hasheminejad, N.; Shirvan, H.E.; Hemmatjo, R.; Sarebanzadeh, K.; Ahmadi, S. Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year. *Weather Clim. Extremes* **2018**, *19*, 49–57. [CrossRef]
6. Shartova, N.; Shaposhnikov, D.; Konstantinov, P.; Revich, B. Cardiovascular mortality during heat waves in temperate climate: An association with bioclimatic indices. *Int. J. Environ. Health Res.* **2018**, *28*, 522–534. [CrossRef]
7. Wang, F.; Duan, K.; Zou, L. Urbanization Effects on Human-Perceived Temperature Changes in the North China Plain. *Sustainability* **2019**, *11*, 3413. [CrossRef]
8. Pichugina, N.Y.; Voronina, L.V. Novosibirsk region bioclimatic conditions estimation. In Proceedings of the Interexpo Geo-Siberia, Novosibirsk, Russia, 19–29 April 2010; Volume 4, pp. 124–128. (In Russian).
9. Wang, Y.; Chen, L.; Song, Z.; Huang, Z.; Ge, E.; Luo, M. Human-perceived temperature changes over South China: Long-term trends and urbanization effects. *Atmos. Res.* **2019**, *215*, 116–127. [CrossRef]
10. Founda, D.; Pierros, F.; Katavoutas, G.; Keramitsoglou, I. Observed Trends in Thermal Stress at European Cities with Different Background Climates. *Atmosphere* **2019**, *10*, 436. [CrossRef]
11. Kargapolova, N.A. Stochastic model of the time series of the average daily bioclimatic index of severity of climatic regime. In Proceedings of the 33rd European Simulation and Modelling Conference, Palma de Mallorca, Spain, 28–30 October 2019; pp. 185–189.
12. Akenteva, M.S.; Kargapolova, N.A.; Ogorodnikov, V.A. Numerical study of the bioclimatic index of severity of climatic regime based on a stochastic model of the joint meteorological time series. In Proceedings of the 5th International Workshop «Applied Methods of Statistical Analysis. Statistical Computation and Simulation», Novosibirsk, Russia, 18–20 September 2019; pp. 311–319.
13. Kargapolova, N.; Khlebnikova, E.; Ogorodnikov, V. Numerical study of properties of air heat content indicators based on stochastic models of the joint meteorological series. *Russ. J. Num. Anal. Math. Model.* **2019**, *34*, 95–104. [CrossRef]
14. Kargapolova, N. Stochastic Models of Non-stationary Time Series of the Average Daily Heat Index. In Proceedings of the 9th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, Praha, Czech Republic, 29–31 July 2019; Volume 1, pp. 209–215. [CrossRef]
15. Kargapolova, N.A. Numerical Study of the Conditional Time Series of the Average Daily Heat Index. In Proceedings of the International Conference on Time Series and Forecasting, Granada, Spain, 25–27 September 2019; Volume 1, pp. 226–234.
16. Kargapolova, N. Numerical Stochastic Model of Non-stationary Time Series of the Wind Chill Index. *Methodol. Comput. Appl. Probab.* **2020**. [CrossRef]
17. Kargapolova, N. Conditional Models of Non-Stationary Time Series of the Wind Chill Index in West Siberia. In Proceedings of the 10th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, Lieusaint, France, 8–10 July 2020.
18. Ogorodnikov, V.A.; Prigarin, S.M. *Numerical Modelling of Random Processes and Fields: Algorithms and Applications*, 1st ed.; VSP: Utrecht, The Netherlands, 1996.
19. Liu, Y.; Shields, M.D. A direct simulation method and lower-bound estimation for a class of gamma random fields with applications in modelling material properties. *Probab. Eng. Mech.* **2017**, *47*, 16–25. [CrossRef]
20. Mikhailov, G.A. On the "repetition" method for modelling of random vectors and processes (randomization of correlation matrices). *Theory Probab. Its Appl.* **1974**, *19*, 873–878.

21. Liu, Y.; Lee, F.-H.; Quek, S.-T.; Beer, M. Modified linear estimation method for generating multi-dimensional multi-variate Gaussian field in modelling material properties. *Probab. Eng. Mech.* **2014**, *38*, 42–53. [CrossRef]
22. Ammon, D. Approximation and Generation of Gaussian and Non-Gaussian Stationary Processes. *Struct. Saf.* **1990**, *8*, 153–160. [CrossRef]
23. Svanidze, G.G. *Mathematical Modeling of Hydrologic Series for Hydroelectric and Water Resources Computations*, 1st ed.; Water Resources Publications: Littleton, CO, USA, 1980.
24. Marchenko, A.S.; Seomochkin, A.G. *FΦΦF*-method for the time series simulation by observed realizations. In *Numerical Methods of Statistical Simulation*, 1st ed.; Mikhailov, G.A., Ed.; Computational Center SB AS USSR: Novosibirsk, Russia, 1987; pp. 6–14. (In Russian)
25. Yamazaki, F.; Shinozuka, M. Digital Generation of Non-Gaussian Stochastic Fields. *J. Eng. Mech.* **1988**, *114*, 1183–1197. [CrossRef]
26. Piranashvili, Z.A. Some problems of statistical probabilistic modelling of random processes. *Probl. Oper. Res.* **1966**, *1*, 53–91. (In Russian)
27. Molnar, G. An evaluation of wind chill. In *Sixth Conference on Cold Injury*; Josiah Macy Foundation: New York, NY, USA, 1960; pp. 175–221.
28. Steadman, R.G. Indices of Windchill of Clothed Persons. *J. Appl. Meteorol.* **1971**, *10*, 674–683. [CrossRef]
29. Bluestein, M. An Evaluation of the Wind Chill Factor: Its Development and Applicability. *J. Biomech. Eng.* **1998**, *120*, 255–258. [CrossRef]
30. Roshan, G.; Mirkatouli, G.; Shakoor, A.; Mohammad-Nejad, V. Studying wind chill index as a climatic index effective on the health of athletes and tourists interested in winter sports. *Asian J. Sports Med.* **2010**, *1*, 108–116. [CrossRef]
31. *Report on Wind Chill Temperature and Extreme Heat Indices: Evaluation and Improvement Projects*; FCM-R19-2003; Office of the Federal Coordinator for Meteorological Services and Supporting Research: Washington, DC, USA, 2003.
32. Osczevski, R.; Bluestein, M. The New Wind Chill Equivalent Temperature Chart. *Bull. Am. Meteorol. Soc.* **2005**, *86*, 1453–1458. [CrossRef]
33. Mekis, É.; Vincent, L.A.; Shephard, M.W.; Zhang, X. Observed Trends in Severe Weather Conditions Based on Humidex, Wind Chill, and Heavy Rainfall Events in Canada for 1953–2012. *Atmos. Ocean* **2015**, *53*, 383–397. [CrossRef]
34. Russian Institute of Hydrometeorological Information—World Data Center. Available online: http://meteo.ru/data/163-basic-parameters (accessed on 31 January 2020).
35. Buligina, O.N.; Veselov, V.M.; Razuvaev, V.N.; Aleksandrova, T.M. The Description of the Data of the Main Meteorological Parameters at Weather Stations Situated in Russia. 2014. Available online: http://meteo.ru/data/163-basic-parameters#%D0%BE%D0%BF%D0%B8%D1%81%D0%B0%D0%BD%D0%B8%D0%B5-%D0%BC%D0%B0%D1%81%D1%81%D0%B8%D0%B2%D0%B0-%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85 (accessed on 31 January 2020).
36. Cario, M.C.; Nelson, B.L. *Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix*; Working Paper; Department of Industrial Engineering and Management Sciences, Northwestern University: Evanston, IL, USA, 1997.
37. Hartkamp, A.D.; de Beurs, K.; Stein, A.; White, J. *Interpolation Techniques for Climate Variables*; Geographic Information Systems Series 99-01; International Maize and Wheat Improvement Center (CIMMYT): Mexico, Mexica, 1999.
38. Sluiter, R. *Interpolation Methods for Climate Data (Literature Review)*, 1st ed.; KNMI: De Bilt, The Netherlands, 2009.
39. Gandin, L.S.; Kagan, R.L. *Statistical Methods of Interpretation of the Meteorological Data*, 1st ed.; Gidrometeoizdat: Leningrad, Russia, 1976. (In Russian)
40. Ogorodnikov, V.; Sereseva, O. Approximate numerical modelling of inhomogeneous stochastic fields of daily sums of liquid precipitation. *Russ. J. Num. Anal. Math. Model.* **2014**, *29*, 375–382. [CrossRef]