DOI:10.28974/idojaras.2022.2.1

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 2, April – June, 2022, pp. 159–184

Joint examination of climate time series based on a statistical definition of multidimensional extreme

Tamás Szentimrey [1] and **Beatrix Izsák**[2, 3,*]

[1] *Varimax Limited Partnership, Budapest, Hungary*

[2] *Hungarian Meteorological Service,*
Kitabel Pál Street 1, H-1024, Budapest, Hungary

[3] *ELTE Faculty of Science,*
Doctoral School of Earth Sciences, Budapest, Hungary

**Corresponding author E-mail: izsak.b@met.hu*

(Manuscript received in final form March 25, 2022)

Abstract— The joint examination of the climate time series may be efficient methodology for the characterization of extreme weather and climate events. In general, the main difficulties are connected with the different probability distribution of the variables and the handling of the stochastic connection between them.

The first problem can be solved by the standardization procedures, i.e., to transform the variables into standard normal ones. For example, there are the Standardized Precipitation Index (SPI) series for the precipitation sums assuming gamma distribution, or the standardization of temperature series assuming normal distribution. In case of more variables, the problem of stochastic connection can be solved on the basis of the vector norm of the transformed variables defined by their covariance matrix.

We will present the developed mathematical methodology and some examples for its meteorological applications.

Key-words: climate time series, vector variables, multidimensional extreme, transformation of vector components, vector norm by matrix, correlation matrix, SPI (Standardized Precipitation Index), STI (Standardized Temperature Index), SPTI (Standardized Precipitation and Temperature Index), hypothesis testing, extreme subsystems

1. Introduction

In the case of joint examination of several climate time series, extreme values can no longer be interpreted by the simple concepts of maximum and minimum values. However, among the element sets or vectors, there may obviously be some that are considered natural, and there may be some that seem unusual to us. The latter can also be conceivable without any extraordinariness of the individual elements, but their co-occurrence can already be considered an extreme phenomenon. Moreover, it can be assumed that the examination of such multidimensional extremes is a more effective tool for examining and characterizing climate change than dealing only with one-dimensional cases.

We began to deal with this topic and the development of the mathematical foundations more than twenty years ago (*Szentimrey*, 1999), and we have carried out a number of such examinations during this time (*Szentimrey et al.*, 2014). Now we want to present a summary of the mathematical results and give some examples for their meteorological applications.

2. Problems of the concept of multidimensional extreme

2.1 Statistical model

Let $\mathbf{X}(t) = [X_1(t), \ldots, X_N(t)]^T$ $(t = 1,2,\ldots,n)$ be a multidimensional time series, which are totally independent and identically distributed probability vector variables. The distribution functions of the components are $F_j(x)$ $(j = 1,2,\ldots,N)$, the vector of expectations is $\mathrm{E}(\mathbf{X}(\mathrm{t})) = \mathbf{E} = [E_1, \ldots, E_N]^T$ $(t = 1,2,\ldots,n)$, and the vector of standard deviations is $\mathrm{D}(\mathbf{X}(\mathrm{t})) = \mathbf{D} = [D_1, \ldots, D_N]^T$ $(t = 1,2,\ldots,n)$.

2.2 "Basic" questions, problems

- The joint examination of the vector components may be efficient for the characterization of extreme events. In general the main difficulties are connected with their different probability distribution and the handling of the stochastic connection between them.
- Which vector variable can be considered extreme?
- How can be tested the null hypothesis of the identical distribution of the vector variables on the basis of the analysis of extremes?
- How can be explained the extremity by the subsystems of the components?

2.3 Transformation of the vector components

One of the problems is that the vector components have different probability distributions and scales. This, however, can be solved by some transformation procedure.
The transformed vector variables are

$$\mathbf{Z}(t) = \mathbf{h}(\mathbf{X}(t)) = [h_1(X_1(t)), \ldots, h_N(X_N(t))]^T \qquad (t = 1,2,..,n)\,,$$

where we assume that, $\mathrm{P}\left(X_j(t) \in (a_j, b_j)\right) = 1$, and $h_j(x)$ is a strictly monotonically increasing function on interval (a_j, b_j) $(j = 1,2,\ldots,N)$.

Remark 2.3.1

If the components were examined separately, the use of the variables $X_j(t)$ or $Z_j(t) = h_j(X_j(t))$ $(t = 1,2,\ldots,n)$ would be equivalent in respect of extremity. However, at the joint examination it is important that none of the components play a dominant role, so the "similarity" of the distributions should be aimed.

2.4 Postulates to the definition of multidimensional extreme

Let $\mathbf{Z}(\mathrm{t}) = \mathbf{h}(\mathbf{X}(t))(t = 1,2,..,n)$ be transformed vector variables according to Section 2.3.

i. $\mathbf{X}(t_e)$ is extreme, i.e., extreme realization, if and only if $\mathbf{Z}(t_e) = \mathbf{h}(\mathbf{X}(t_e))$ is extreme. This postulate can be accepted because of the deterministic cause-and-effect relationship.
ii. Let us assume that, there exist $h_j(x)$ $(j = 1,2,\ldots,N)$, that fulfil $Z_j(t) = h_j(X_j(t)) \in N(0,1)$, i.e., they are standard normal variables. Then, according to the above $\boldsymbol{X}(t_e)$ is extreme, if and only if $\mathbf{Z}(t_e) = [\mathrm{Z}_1(t_e),\ldots,\mathrm{Z}_\mathrm{N}(t_e)]^\mathrm{T}$ is extreme.
iii. Let us assume further that, the joint distribution of the vector components $Z_1(t),\ldots,Z_N(t)$ is also normal, i.e., $\mathbf{Z}(t) \in N(\mathbf{0},\mathbf{R})$, where $\mathbf{R} = \mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^\mathrm{T})$ is the correlation matrix assuming the existence of the inverse matrix $\mathbf{R}^{-1}$. Then the joint density function is $g(\boldsymbol{z}) = (2\pi)^{-\frac{N}{2}}|\mathbf{R}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\ \mathbf{z}^\mathrm{T}\mathbf{R}^{-1}\mathbf{z}\right)$.

This means, the density function $g(\boldsymbol{z})$ is a strictly monotonous decreasing function of the $\boldsymbol{R}^{-1}$-norm $\|\mathbf{z}\|_{\mathbf{R}^{-1}} = (\mathbf{z}^\mathrm{T}\mathbf{R}^{-1}\mathbf{z})^{\frac{1}{2}}$. Consequently, $\boldsymbol{Z}(t_e)$ is extreme, if and only if $\|\mathbf{Z}(t_e)\|_{\mathbf{R}^{-1}} = \max_{1\le t\le n} \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}}$.

2.5 Definition of the multidimensional extreme

In summary, the multidimensional extreme can be defined as follows.

Definition 2.5.1

Let $\mathbf{X}(t) = [X_1(t), \ldots., X_N(t)]^T (t = 1,2, \ldots.., n)$ be totally independent and identically distributed probability vector variables.
Let us assume that $h_j(x)$ are strictly monotonous increasing functions on the intervals (a_j, b_j) where $\mathrm{P}\left(X_j(t) \in (a_j, b_j)\right) = 1$ $(j = 1,2, \ldots.., N)$.
Let us assume further that the components of the vector variables $\mathbf{Z}(t) = \mathbf{h}(\mathbf{X}(t)) = [h_1(X_1(t)), \ldots., h_N(X_N(t))]^T$ $(t = 1,2, \ldots.., n)$ are standard normal, i.e., $Z_j(t) = h_j(X_j(t)) \in N(0,1)$ $(j = 1,2, .., N)$, and their common correlation matrix $\mathbf{R} = \mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^{\mathrm{T}})$ has the inverse matrix $\mathbf{R}^{-1} (t = 1,2, \ldots.., n)$.
Then, $\mathbf{X}(t_e)$ is extreme, if and only if $\|\mathbf{Z}(t_e)\|_{\mathbf{R}^{-1}} = \max_{1 \le t \le n} \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}}$, where $\|\mathbf{Z}(t)\|_{\boldsymbol{R}^{-1}} = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}(t))^{\frac{1}{2}}$ is the $\mathbf{R}^{-1}$-norm of vector variable $\mathbf{Z}(t)$.

3. Examples for the transformation of climate data series and standard indexes

Various type of climate data can be transformed for standard normal distributed variable on the basis of the following well known theorem.

Theorem 3.1

Let us assume that $\mathrm{P}\left(X_j(t) \in (a_j, b_j)\right) = 1$, moreover, the distribution function $F_j(x)$ of variable $X_j(t)$ is strictly monotonous increasing and continuous on the interval (a_j, b_j) $(j = 1,2, \ldots.., N)$. Then $h_j(x) = \Phi^{-1}(F_j(x))$ is also strictly monotonous increasing and continuous function on the interval (a_j, b_j), furthermore, $Z_j(t) = h_j(X_j(t)) \in N(0,1)$ $(j = 1,2, \ldots.., N)$, where $\Phi(x)$ is the standard normal distribution function.

Proof.

$$\mathrm{P}(Z_j(t) < z) = \mathrm{P}(h_j(X_j(t)) < z) = \mathrm{P}\left(\Phi^{-1}\left(F_j(X_j(t))\right) < z\right) = \mathrm{P}\left(F_j(X_j(t)) < \Phi(z)\right) =$$
$$= \mathrm{P}\left(X_j(t) < F_j^{-1}(\Phi(z))\right) = F_j\left(F_j^{-1}(\Phi(z))\right) = \Phi(z).$$

3.1 Annual precipitation sum

$X_1(t) \in \Gamma(p,\lambda) \qquad (t = 1,2,\ldots\ldots,n)$ and
$Z_1(t) = \Phi^{-1}\big(G(X_1(t))\big) \in N(0,1) \qquad (t = 1,2,\ldots\ldots,n)$,
where $G(x)$ notates the $\Gamma(p,\lambda)$ distribution function and $\Phi^{-1}(x)$ is the inverse of the standard normal distribution function. In fact, $Z_1(t)$ is the SPI (Standardized Precipitation Index) well known in meteorology for characterization of the drought events.

3.2 Annual mean temperature

$X_2(t) \in N(E_2, D_2) \qquad (t = 1,2,\ldots\ldots,n)$ and
$Z_2(t) = \Phi^{-1}\left(\Phi\left(\frac{X_2(t)-E_2}{D_2}\right)\right) = \frac{X_2(t)-E_2}{D_2} \in N(0,1) \quad (t = 1,2,\ldots\ldots,n)$,
where $\Phi\left(\frac{x-E_2}{D_2}\right)$ is the $N(E_2, D_2)$ distribution function. We defined this index as STI (Standardized Temperature Index).

3.3 Annual precipitation sum and mean temperature together

$\boldsymbol{X}(t) = [X_1(t), X_2(t)]^T \quad (t = 1,2,..,n)$ and the
SPI, STI indexes together are $\boldsymbol{Z}(t) = [Z_1(t), Z_2(t)]^T \quad (t = 1,2,..,n)$.

Definition 3.3.1

The SPTI (Standardized Precipitation and Temperature Index) can be defined as

$$\|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}} = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}(t))^{\frac{1}{2}} \quad (t = 1,2,..,n) ,$$

where $\mathbf{R}$ is the common correlation matrix of the vector variables $\mathbf{Z}(t)$.

4. Hypothesis testing, statistical test based on norm

4.1 Basic properties of the norms of vector variables

Theorem 4.1.1

Let us assume about the vector variables $\mathbf{Z}(t) = [\mathrm{Z}_1(t),\ldots,\mathrm{Z}_\mathrm{N}(t)]^\mathrm{T}$ $(t = 1,2,\ldots\ldots,n)$, that the vector of expectations is $\mathrm{E}(\mathbf{Z}(t)) = \mathbf{0}$, the vector of standard deviations is $\mathrm{D}(\mathbf{Z}(t)) = \mathbf{1}$, and the correlation matrix is $\mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^\mathrm{T}) = \mathbf{R}$.

Then the following properties are true:

i. Let $\mathbf{R}^{\frac{1}{2}} = \left(\mathbf{R}^{\frac{1}{2}}\right)^{\mathrm{T}}, \mathbf{R}^{\frac{1}{2}}\mathbf{R}^{\frac{1}{2}} = \mathbf{R}$. Then the covariance matrix of $\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)$ is the identity matrix $\mathbf{I}$.

ii. If matrix $\mathbf{P}$ satisfies that the covariance matrix of $\mathbf{PZ}(t)$ is $\mathbf{I}$, then $\|\mathbf{PZ}(t)\| = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}(t))^{\frac{1}{2}} = \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}}$, where $\|\mathbf{PZ}(t)\|$ is the Euclidean norm.

iii. If $\mathbf{Z}(t) \in N(\mathbf{0}, \mathbf{R})$, that means that the joint distribution of the components is normal, then $\|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}} \in \chi_N$, i.e., this $\mathbf{R}^{-1}$-norm is chi distributed with degrees of freedom *N*.

Proof.

i. If $\mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^{\mathrm{T}}) = \mathbf{R}$, then $\mathrm{E}\left(\left(\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)\right)\left(\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)\right)^{\mathrm{T}}\right) =$

$$= \mathrm{E}\left(\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-\frac{1}{2}}\right) = \mathbf{R}^{-\frac{1}{2}}\mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^{\mathrm{T}})\mathbf{R}^{-\frac{1}{2}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{R}\mathbf{R}^{-\frac{1}{2}} = \mathbf{I}.$$

ii. If $\mathrm{E}\left((\mathbf{PZ}(t))(\mathbf{PZ}(t))^{\mathrm{T}}\right) = \mathbf{I}$, then $\mathbf{PRP}^{\mathrm{T}} = \mathbf{I}$, since $\mathbf{PRP}^{\mathrm{T}} = \mathrm{E}\left((\mathbf{PZ}(t))(\mathbf{PZ}(t))^{\mathrm{T}}\right)$. Therefore, $\mathbf{R} = \mathbf{P}^{-1}(\mathbf{P}^{\mathrm{T}})^{-1}$, and so $\mathbf{R}^{-1} = \mathbf{P}^{\mathrm{T}}\mathbf{P}$. Consequently, $\|\mathbf{PZ}(t)\| = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{PZ}(t))^{\frac{1}{2}} =$

$$= (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}(t))^{\frac{1}{2}} = \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}}.$$

iii. If $\mathbf{Z}(t) \in N(\mathbf{0}, \mathbf{R})$ then $\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t) \in N(\mathbf{0}, \mathbf{I})$ by item (i) and therefore, $\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)\right\| \in \chi_N$ according to the definition of chi distribution. Furthermore, $\|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}} = \left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Z}(t)\right\|$ as a consequence of item (ii).

4.2 Statistical test

Using the concept of multidimensional extreme, the hypothesis test for the identical distribution of vector variables can be implemented as follows.
Let us assume that the vector variables $\mathbf{X}(t)$ $(t = 1,2,\ldots\ldots,n)$ are totally independent. Then the null hypothesis for their identical distribution can be accepted if and only if

$$\|\mathbf{Z}(t_e)\|_{\mathbf{R}^{-1}} = \max_{1 \le t \le n} \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}} < cr,$$

where cr is a critical value on a given significance level. This critical value can be calculated on the basis of the chi distribution.

Remark 4.2.1

The question may arise as to whether the application of $\mathbf{R}^{-1}$-norm is optimal, as other $\mathbf{A}$-norm $\|\mathbf{Z}(t)\|_{\mathbf{A}} = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{A}\mathbf{Z}(t))^{\frac{1}{2}}$ could be used, where $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ is a positive definite square matrix. According to the following theorem, if a different $\mathbf{A}$-norm were used, the efficiency of the test would be expected to decrease.

Theorem 4.2.1 ("minimal" acceptance region)

Let us assume about the vector variable $\mathbf{Z} = [Z_1, \ldots, Z_N]^T$, that the vector of expectations is $\mathrm{E}(\mathbf{Z}) = \mathbf{0}$, the vector of standard deviations is $\mathrm{D}(\mathbf{Z}) = \mathbf{1}$, and the correlation matrix is $\mathrm{E}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}) = \mathbf{R}$. Let us assume further, that the positive definite matrix $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$satisfies the following criterion for the expected value of the norm square: $\mathrm{E}(\mathbf{Z}^{\mathrm{T}}\mathbf{A}\mathbf{Z}) = \mathrm{E}(\mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}) = N$.
Then the inequality $|\mathbf{R}| \le |\mathbf{A}^{-1}|$ is true for the determinants, consequently

$$\int_{\mathbf{z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}<\alpha} 1 \ \mathrm{d}\mathbf{z} = |\mathbf{R}|^{\frac{1}{2}} \int_{\mathbf{y}^{\mathrm{T}}\mathbf{y}<\alpha} 1 \ \mathrm{d}\mathbf{y} \le |\mathbf{A}^{-1}|^{\frac{1}{2}} \int_{\mathbf{y}^{\mathrm{T}}\mathbf{y}<\alpha} 1 \ \mathrm{d}\mathbf{y} = \int_{\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}<\alpha} 1 \ \mathrm{d}\mathbf{z}.$$

Proof.

First it needs to be seen that the determinants fulfil the inequality $|\mathbf{R}| \le |\mathbf{A}^{-1}|$.
According to our assumption:

$$\mathrm{E}(\mathbf{Z}^{\mathrm{T}}\mathbf{A}\mathbf{Z}) = \mathrm{E}\left(\left(\mathbf{A}^{\frac{1}{2}}\mathbf{Z}\right)^{\mathrm{T}}\left(\mathbf{A}^{\frac{1}{2}}\mathbf{Z}\right)\right) = N.$$

Consequently, the sum of the diagonal elements of matrix $\mathbf{A}^{\frac{1}{2}}\mathbf{R}\mathbf{A}^{\frac{1}{2}}$ is also equals to N, since

$$\mathbf{A}^{\frac{1}{2}}\mathbf{R}\mathbf{A}^{\frac{1}{2}} = \mathrm{E}\left(\left(\mathbf{A}^{\frac{1}{2}}\mathbf{Z}\right)\left(\mathbf{A}^{\frac{1}{2}}\mathbf{Z}\right)^{\mathrm{T}}\right) = \mathrm{E}\left(\mathbf{A}^{\frac{1}{2}}\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{A}^{\frac{1}{2}}\right).$$

Therefore, the arithmetic mean of the eigenvalues of matrix $\mathbf{A}^{\frac{1}{2}}\mathbf{R}\mathbf{A}^{\frac{1}{2}}$ equals to 1, consequently their geometric mean and their product are less or equal to 1. Using that the product of the eigenvalues equals to the determinant of the matrix we obtain,

$$\left|\mathbf{A}^{\frac{1}{2}}\mathbf{R}\mathbf{A}^{\frac{1}{2}}\right| = |\mathbf{R}||\mathbf{A}| \le 1, \text{ and so } |\mathbf{R}| \le |\mathbf{A}^{-1}|.$$

Then applying the substitutions $\mathbf{y} = \mathbf{R}^{-\frac{1}{2}}\mathbf{z}$ and $\mathbf{y} = \mathbf{A}^{\frac{1}{2}}\mathbf{z}$, respectively, we have proved the following relation,

$$\int_{\mathbf{z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}<\alpha} 1 \ \mathrm{d}\mathbf{z} = |\mathbf{R}|^{\frac{1}{2}} \int_{\mathbf{y}^{\mathrm{T}}\mathbf{y}<\alpha} 1 \ \mathrm{d}\mathbf{y} \le |\mathbf{A}^{-1}|^{\frac{1}{2}} \int_{\mathbf{y}^{\mathrm{T}}\mathbf{y}<\alpha} 1 \ \mathrm{d}\mathbf{y} = \int_{\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}<\alpha} 1 \ \mathrm{d}\mathbf{z}.$$

Remarks 4.2.2

1. According to Theorem 4.2.1, applying a different $\mathbf{A}$-norm for which the probability of the first type error would be similar to that of the $\mathbf{R}^{-1}$-norm would increase the probability of the second type error, therefore, it would decrease the efficiency of the test. In addition, the test is expected to be more effective if the determinant is smaller, that means a stronger linear connection between the components of the vector variable $\mathbf{Z}$.
2. Consequently, in respect of the second type error, the following series of statistics are optimal for testing: $ST(t) = \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}} = (\mathbf{Z}(t)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z}(t))^{\frac{1}{2}}$ $(t = 1,2,\ldots\ldots,n)$.
3. In order to determine and select the extremes optimally, it is expedient also to examine the subsystems of the components.

5. Analysis of the subsystems of the components

5.1 The subsystems of the components and their properties

According to the former notations, let $\mathbf{Z}(t) = [Z_1(t),\ldots,Z_N(t)]^T$ $(t = 1,2,\ldots\ldots,n)$ be vector variables, assuming that the vector of expectations is $\mathrm{E}(\mathbf{Z}(t)) = \mathbf{0}$, the vector of standard deviations is $\mathrm{D}(\mathbf{Z}(t)) = \mathbf{1}$, and the correlation matrix is $\mathrm{E}(\mathbf{Z}(t)\mathbf{Z}(t)^{\mathrm{T}}) = \mathbf{R}$.

The subsystem of the components can be defined by the subsets of indexes $J = \{j_1,\ldots,j_L\} \subseteq \{1,\ldots,N\}$, and the appropriate subsystem vector variables are, $\mathbf{Z}_J(t) = \left[\mathrm{Z}_{j1}(t),\ldots,\mathrm{Z}_{jL}(t)\right]^{\mathrm{T}}$ $\qquad (t = 1,2,\ldots\ldots,n)$.

The correlation matrix is $\mathrm{E}\left(\mathbf{Z}_J(t)\mathbf{Z}_J(t)^{\mathrm{T}}\right) = \mathbf{R}_J \qquad (t = 1,2,\ldots\ldots,n)$.

The series of the statistics characterizing the extremity of the subsystem vector variables defined by the index subset J are as follows:

$$ST(t,J) = \left\|\mathbf{Z}_J(t)\right\|_{\mathbf{R}_J^{-1}} = \left(\mathbf{Z}_J(t)^T\mathbf{R}_J^{-1}\mathbf{Z}_J(t)\right)^{\frac{1}{2}} \qquad (t = 1,2,\ldots\ldots,n)\,.$$

Theorem 5.1.1

If $J_1 = \{j_1,\ldots,j_{L_1}\} \subseteq J_2 = \{j_1,\ldots,j_{L_1+1},\ldots,j_{L_2}\} \subseteq \{1,\ldots,N\}$, then with probability 1,

$$ST(t,J_1) = \left\|\mathbf{Z}_{J_1}(t)\right\|_{\mathbf{R}_{J_1}^{-1}} \leq \left\|\mathbf{Z}_{J_2}(t)\right\|_{\mathbf{R}_{J_2}^{-1}} = ST(t,J_2) \qquad (t = 1,2,\ldots\ldots,n)$$

Proof.

According to our notations, $\mathbf{Z}_{J_1}(t) = \left[\mathrm{Z}_{j1}(t),\ldots,\mathrm{Z}_{jL_1}(t)\right]^T$,

$$\mathbf{Z}_{J_2}(t) = \left[\mathrm{Z}_{j1}(t),..,\mathrm{Z}_{jL_1}(t),..,\mathrm{Z}_{jL_2}(t)\right]^{\mathrm{T}}.$$

Let us see the linear Euclidean space generated by the components $\mathrm{Z}_{j1}(t),..,\mathrm{Z}_{jL_1}(t),..,\mathrm{Z}_{jL_2}(t)$ of the vector variable $\mathbf{Z}_{J_2}(t)$, where the scalar product is the covariance.

In this space, the components of the vector variable $\mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)$ form an orthonormal system by item (i) of Theorem 4.1.1. Then according to the orthonormalization procedure, there exists such orthonormal basis in the space, whose first L_1 elements are exactly these components.

Formalized, there is a matrix $\mathbf{P} = \left[\mathbf{P}_1^{\mathrm{T}}, \mathbf{P}_2^{\mathrm{T}}\right]^{\mathrm{T}}$ that the covariance matrix of the vector variable $\mathbf{PZ}_{J_2}(t) = \left[\left(\mathbf{P}_1\mathbf{Z}_{J_2}(t)\right)^{\mathrm{T}}, \left(\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right)^{\mathrm{T}}\right]^{\mathrm{T}}$ is the identity matrix **I**, and $\mathbf{P}_1\boldsymbol{Z}_{J_2}(t) = \mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)$.

Consequently, according to item (ii) of Theorem 4.1.1.:

$\left\|\mathbf{Z}_{J_2}(t)\right\|_{\mathbf{R}_{J_2}^{-1}} = \left\|\mathbf{PZ}_{J_2}(t)\right\|$ and $\left\|\mathbf{Z}_{J_1}(t)\right\|_{\mathbf{R}_{J_1}^{-1}} = \left\|\mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)\right\|$, therefore

$$\left\|\mathbf{Z}_{J_2}(t)\right\|_{\boldsymbol{R}_{J_2}^{-1}}^2 = \left\|\mathbf{PZ}_{J_2}(t)\right\|^2 =$$

$$= \left[\left(\mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)\right)^{\mathrm{T}}, \left(\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right)^{\mathrm{T}}\right]\left[\left(\mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)\right)^{\mathrm{T}}, \left(\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right)^{\mathrm{T}}\right]^{\mathrm{T}} =$$

$$= \left(\mathbf{R}_{J_1}^{-\frac{1}{2}}\boldsymbol{Z}_{J_1}(t)\right)^{\mathrm{T}}\left(\mathbf{R}_{J_1}^{-\frac{1}{2}}\mathbf{Z}_{J_1}(t)\right) + \left(\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right)^{\mathrm{T}}\left(\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right) =$$

$$= \left\|\mathbf{Z}_{J_1}(t)\right\|_{\mathbf{R}_{J_1}^{-1}}^2 + \left\|\mathbf{P}_2\mathbf{Z}_{J_2}(t)\right\|^2.$$

Thus, $\left\|\mathbf{Z}_{J_2}(t)\right\|_{\mathbf{R}_{J_2}^{-1}}^2 \geq \left\|\mathbf{Z}_{J_1}(t)\right\|_{\mathbf{R}_{J_1}^{-1}}^2$.

5.2 Extreme subsystems

As a consequence of Theorem 5.1.1, if a subsystem of the components is "extreme", then it is presumably visible throughout the system.

A further consequence is that this theorem allows a meaningful, consistent definition of the series of the L-element $(L = 1,2,.....,N)$ extreme subsystems.

Let $JL(t) \subseteq \{1,...,N\}$ $(\#JL(t) = L)$, where # denotes cardinality be the index set of the L-element extreme subsystem at time t $(L = 1,2,......,N;\ \ t = 1,2,...,n)$, if

$$ST(t,JL(t)) = STL(t) = \max(ST(t,J)|\ J \subseteq \{1,...,N\},\ \#J = L)\ .$$

According to the notations, the L-element extreme subsystem at time *t* is $\mathbf{Z}_{JL(t)}(t)$. Furthermore, as a consequence of Theorem 5.1.1, it is true for the series of statistics $STL(t)$ $(t = 1,2,\ldots\ldots,n)$ belonging to the L-element extreme subsystems that

$$\max_{1\leq j\leq N}\left|Z_j(t)\right| = ST1(t) \leq ST2(t) \leq \cdots \ldots \leq STN(t) = \|\mathbf{Z}(t)\|_{\mathbf{R}^{-1}}$$
$$(t = 1,2,\ldots\ldots,n).$$

5.3 Methodological basis for the analysis procedure

1. Examination of the series of the statistics $STL(t)$ $(t = 1,\ldots,n)$ showing which series $(L = 1,..,N)$ majorize each other, followed by the looking for the extremes and testing them on the basis of the critical values.
2. Calculation of the critical values on the given significance level assuming the chi distribution.
3. Examination of the subsystems $JL(t)$ belonging to the statistics
   $STL(t) = ST(t, JL(t))$ $(t = 1,\ldots,n;\quad L = 1,..,N)$.
4. Comparison of the statistics $STL(t)$ $(t = 1,\ldots,n;\quad L = 1,..,N)$ for different L can be done on the basis of the following probabilities: $p(STL(t)) = 1 - H_L(STL(t))$ $(t = 1,\ldots,n;\quad L = 1,..,N)$, where $H_L(x)$ is the chi distribution function with degrees of freedom L. The less probability $p(STL(t))$ is the more extreme value.
5. Examination of the climate indexes SPI, STI, and SPTI is also recommended.

6. Meteorological applications

6.1 Data

For our study, we used daily data from the last 151 years. For temperature, this means 11 stations from January 1, 1870, 33 from January 1, 1901, 55 from January 1, 1951, and 114 from January 1, 1975 to December 31, 2020. For precipitation, we used data from 11 precipitation stations from January 1, 1870, 131 from January 1, 1901, and 461 from January 1, 1951 to December 31, 2020 (*Figs. 1* and *2*).

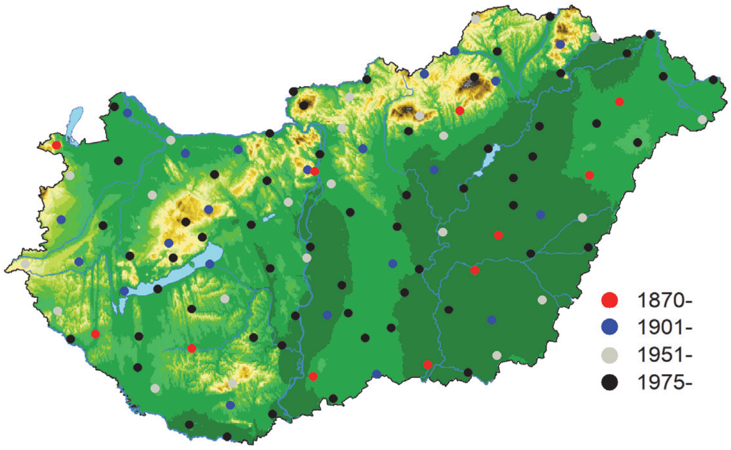

Fig. 1. Location of the stations in case of temperature.

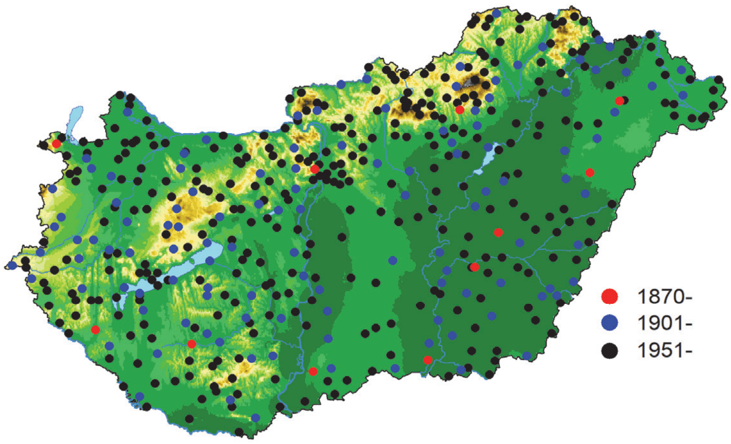

Fig. 2. Location of the stations in the case of precipitation.

As a first step, representative time series had to be produced from the raw measurements. Climate studies, in particular those related to climate change, require long, high-quality, controlled data sets which are both spatially and temporally representative. Changing the context in which the measurements were taken, for example relocating the station, or a change in the frequency of measurements or in the instruments used may result in an unduly fractured time series (*Izsák* and *Szentimrey*, 2020). Data errors and inhomogeneities are eliminated and data gaps are filled in using the MASH (Multiple Analysis of Series for Homogenization; *Szentimrey*, 2017) homogenization procedure. The homogenized station data series were interpolated to a regular grid covering the whole area of Hungary using the MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey* and *Bihari*, 2014) method, thus obtaining a high quality, representative data set. The method was used for both temperature and precipitation amount series, thus we obtained a grid point time series system with daily data at a resolution of 0.1 × 0.1 degrees. The grid point system is shown in *Fig. 3*. In order to use the longest measurement data from as many stations as possible for our dataset, we harmonized four MASH systems during homogenization and used the grid point data set interpolated using the most stations at each time point during interpolation (*Izsák et al*., 2022).

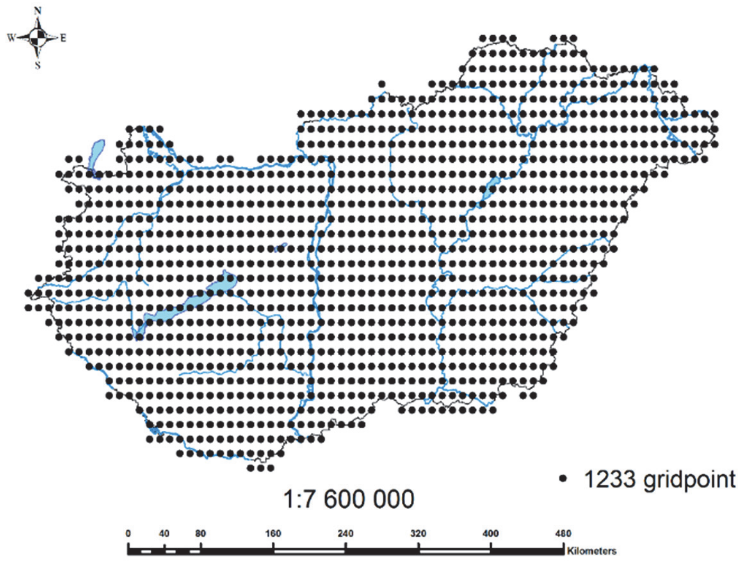

Fig. 3. Spatial location of the gridpoints.

In the next step, the *SPI* (Section 3.1) and *STI* (Section 3.2) variable series are calculated from the grid point data series (*WMO,* 2012; *Szentimrey et al.,* 2014). For annual studies, we will have *SPI12*, i.e., calculated for 12-month precipitation accumulation periods, similarly *STI12* for temperature as our transformed data series. For seasonal studies, we calculate *SPI3* and *STI3* for all 1233 grid points, correspondingly to a resolution of 0.1 degrees. If we perform our analyses in space, we compute the grid point *SPTI* series (Definition 3.3.1) from the corresponding *SPI* and *STI* series. This is shown for two-dimensional studies in Section 6.3. and for eight-dimensional studies in Section 6.4.

If we are studying a regional average, we compute the spatial average of the *STI* and *SPI* series, and then standardize these, since the average of standard normal series will no longer necessarily have a standard normal distribution. Then these standardized spatial average *SPI* and *STI* series will be the components of the multivariate spatial average vector variable series. Consequently, the spatial average norm and *SPTI* series are computed also from the standardized spatial averages of *SPI* and *STI* series. Specifically, for the eight-dimensional studies, the standardized spatial average *SPI3* and *STI3* for the seasons are the eight components of the vector variable. From this we compute the eight-dimensional *ST8* norms (Definition 2.5.1) and the extreme subsystems *ST1*, *ST2*,..., *ST8* (Section 5.2). The corresponding values are presented in Section 6.4. Calibration period for both the two- and eight-dimensional statistics presented in this paper is the period 1871–1900.

6.2 Statistical tests

As indicated in Sections 4.2 and 5.3, to accept the null hypothesis of identical distribution of vector variables, the critical value for a given significance level must be specified. In the present case, certain critical values have been defined at a significance level of 0.1. Three different statistical tests are used to see whether climate change can be detected by analyzing different meteorological elements together. In the following subsections, the procedure for determining the tests and their corresponding critical values is described.

6.2.1 Test 1

In this test, we seek to answer the question: are extreme events more frequent than would be expected for the identical distribution?

The first critical value is defined as follows:

- For any *STN(t)* (t = 1, 2, ..., n) statistic, there is a probability of 0.1 that the critical value *Cr1* is reached, i.e., it is expected to occur in 10% of the statistics.

Cr1 values are given in *Table 1* determined assuming joint normality of the transformed components, therefore, the chi distribution was applied for the statistics *STN(t)* according to item (iii) of Theorem 4.1.1.

The test procedure is as follows:

- Calculate the *STN(t)* norms and then to determine the frequency of norms exceeding the *Cr1* value for the total period. This frequency is denoted by v.
- If the null hypothesis is true then $v \in B(n,p)$, where $B(n,p)$ denotes the Binomial distribution with parameters n and p, specifically n=150 and p=0.1.
- Consequently, according to the central limit theorem, the standardized value *TS1* of the frequency ν converges to the standard normal distribution.
- This gives the critical value *Cr3*=1.65 for the standardized value *TS1* at the significance level 0.1. If *TS1*≥ Cr3, the null hypothesis for the identical distribution is not accepted, if *TS1*<Cr3, it is accepted.

6.2.2 Test 2

In the second test, we seek to answer the question whether there are extreme cases in the data set that would have a low probability of occurring if the distribution were identical.

This test is detailed in Section 4.2, and the critical value *Cr2* is defined as follows:

- The maximum of the statistics *STN(t)* ($t = 1, 2, ..., n$) with probability at most 0.1 can reach the critical value *Cr2*.

Table 1 shows these *Cr2* values as a function of the different degrees of freedom. For the present study, the *Cr2* critical values refer to the n=150 years data set. In determining the critical values, the chi distribution was assumed for the statistics *STN(t)* according to item (iii) of Theorem 4.1.1.

Table 1. Critical values

Degrees of freedom	2	3	4	6	8	12	24
Cr1	2.15	2.50	2.79	3.26	3.66	4.31	5.76
Cr2	3.81	4.12	4.39	4.83	5.20	5.82	7.24

6.2.3 Test 3

Trend analysis is used to answer the question: is there a one-way change in extremality over time that would occur with low probability given the same distribution?

We fit an exponential trend to *STN(t)* statistics. We test the significance of the resulting trend coefficient. If the one-way change is significant, we reject the null hypothesis at the given significance level. The significance level in the present case is 0.1.

6.3 Two-dimensional application: the SPTI index

In the analysis, we consider the *SPTI* (Definition 3.3.1) series defined above, both for the spatial average and for grid points. In the two-dimensional case, *Cr1*=2.15 and *Cr2*=3.81 for n=150 (*Table 1*). Based on Section 6.2, statistical tests were carried out to see if there is a change in climate when the two elements are considered together.

Fig. 4 shows the spatial average *SPI*, *STI* and *SPTI* values based on annual precipitation sum and mean temperature. It can be clearly seen that there are values above both *Cr1* and *Cr2,* suggesting that there is climate change when looking at spatial average annual mean temperature and precipitation sum series together. The change in the *SPTI* index over 150 years was determined using exponential trend estimation. On spatial average, this index increased by 100% over the whole period. This change is significant at the 0.1 significance level. This is indicated by the green line in *Fig. 4*. Based on Section 6.2.1, the value of *TS1* was determined (*Table 3*). The null hypothesis of an identical distribution for the annual values, a spatial average, cannot be accepted on the basis of this test either, since a *TS1* value well above the critical value was obtained.

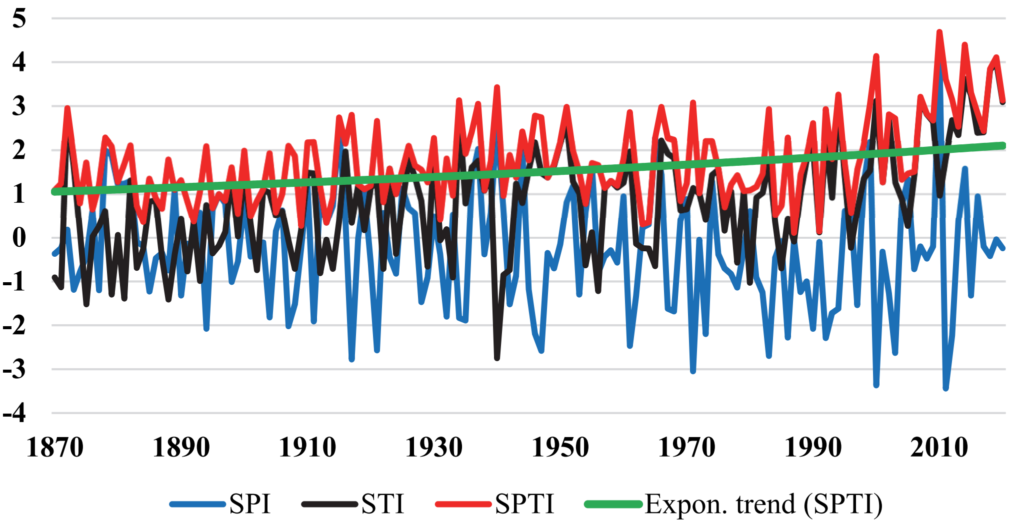

Fig. 4. Spatial average of SPI, STI and SPTI values, based on annual precipitation sum and mean temperature.

If the *SPTI* values are calculated as grid points instead of spatial averages, we can get an idea of which areas of Hungary are experiencing climate change at the 0.1 significance level, when the precipitation and temperature series are taken into account. In *Fig. 5* we show the areas where the maximum of the SPTI values exceeds *Cr2*, these are the red colored areas, while in the green colored areas, these statistics exceed only *Cr1*.

On a spatial average, 2010 was the extreme year with the highest precipitation but average temperature (*Table 2*) If we look at individual areas, *Fig. 6* shows that the year 2014 is still the one with the largest area. This year is markedly extreme in the sense that it is recorded in the yearbooks as a hot and wet year. The extreme drought and heat of the year 2000 makes it appear in even larger areas. Generally speaking, recent years appear as extreme. The largest areas are dominated by warm years with precipitation, but dry hot years also appear on the list. In the case of annual studies, the pairs of dry and cool years are not listed as extreme, but will only appear in seasonal studies.

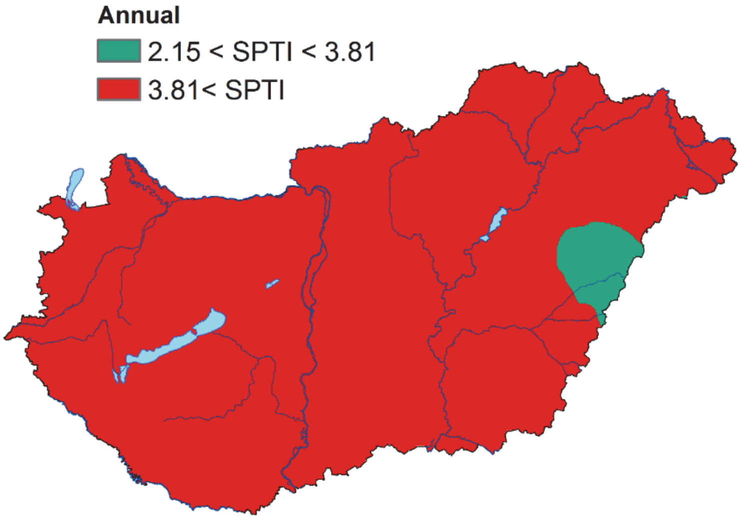

Fig. 5. Maximum of the SPTI values, based on annual precipitation sum and mean temperature, 1870–2020.

Table 2. Extreme years and seasons by SPTI, spatial average, 1871–2020

Annual		Winter		Spring		Summer		Autumn	
2010	4.69	**2006/07**	2.7	**1934**	3.8	**1984**	4.13	**1986**	4.08
2014	4.4	**1909/10**	2.62	**1875**	3.25	**2019**	4.07	**1920**	3.95
2000	4.14	**1935/36**	2.56	**2003**	3.02	**2003**	3.81	**1908**	3.69
2019	4.11	**1989/90**	2.49	**1946**	2.9	**1976**	3.57	**2011**	3.62
2018	3.85	**1950/51**	2.47	**1920**	2.85	**1978**	3.5	**2006**	3.31
2011	3.61	**1976/77**	2.44	**2018**	2.79	**2018**	3.49	**1978**	3.01
1940	3.43	**1879/80**	2.41	**1872**	2.78	**1962**	3.46	**1947**	2.93
2015	3.31	**1962/63**	2.4	**1937**	2.7	**2012**	3.46	**1959**	2.84
1994	3.26	**1881/82**	2.35	**2007**	2.66	**1919**	3.45	**1953**	2.76
2007	3.21	**1997/98**	2.25	**2010**	2.65	**1923**	3.43	**1912**	2.75

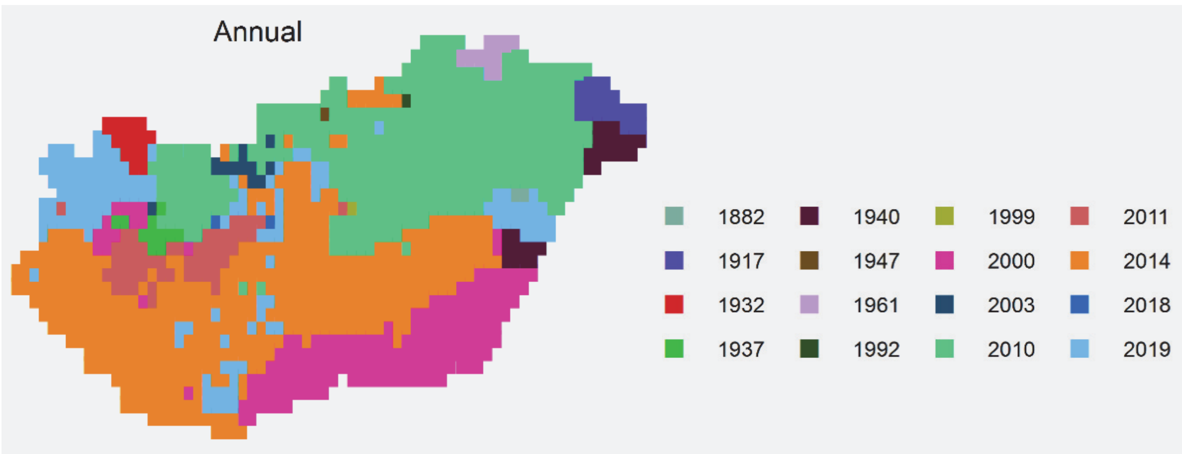

Fig. 6. Year of maximum SPTI values, based on annual precipitation sum and mean temperature, 1870–2020.

Similar to annual studies, seasons and months can be considered together in terms of precipitation sum and mean temperature.

If only the seasonal precipitation sums and mean temperatures are considered, the winter norm values (SPTI) are the lowest and exceed only the mildest *Cr1* critical value (*Fig. 7*). If we want to decide on the basis of the first test (Section 6.2.1), we can see the *TS1* values in *Table 3*. Only the winter value remains below the critical value. The other two tests (Sections 6.2.2 and 6.2.3) were also applied to the seasonal *SPTI* values. None of the spring norm values reaches *Cr2*, but here relatively many values exceed *Cr1*. The summer norm values are the highest and there are values exceeding *Cr2*. The autumn values also include some norms above *Cr2*, but on average they are lower than in the summer study. When an exponential trend is fitted to the *SPTI* values, significant changes are observed only in spring and summer. When the temperature and precipitation time series are considered together, significant increases are observed in both cases. Summer norm values increased by 90% and spring values by 55%.

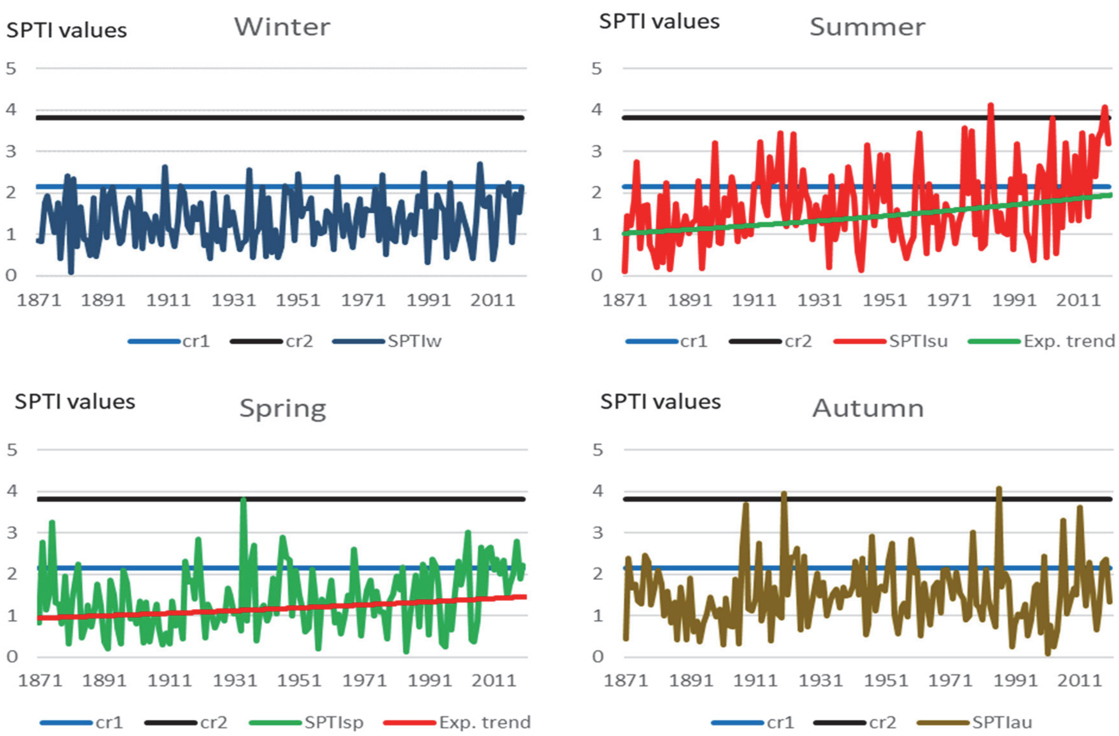

Fig 7. SPTI values by season, standardised spatial average.

Table 3. *TS1* values for Test1, the critical value is Cr3=1.65

1871-2020	**Annual**	**Summer**	**Autumn**	**Winter**	**Spring**
TS1	**9.25**	**8.16**	**2.72**	0.27	**2.99**

Of course, the seasonal analyses also examined the areas where the maximum *SPTI* values exceeded each critical value. The areas exceeding *Cr1* and *Cr2* are shown in *Fig. 8*. In all cases, the milder *Cr1* critical value is exceeded everywhere by the maximum of the normal values, these are the areas marked in green. Where the *SPTI* maxima exceed the more severe *Cr2* critical value, the area is red colored.

It can be clearly seen in *Fig. 8A* that the winter values do not reach *Cr2* anywhere, while for the autumn (*Fig. 8D*) and summer (*Fig. 8B*) maxima we get the largest areas with maximum *SPTI* values above the more stringent critical value.

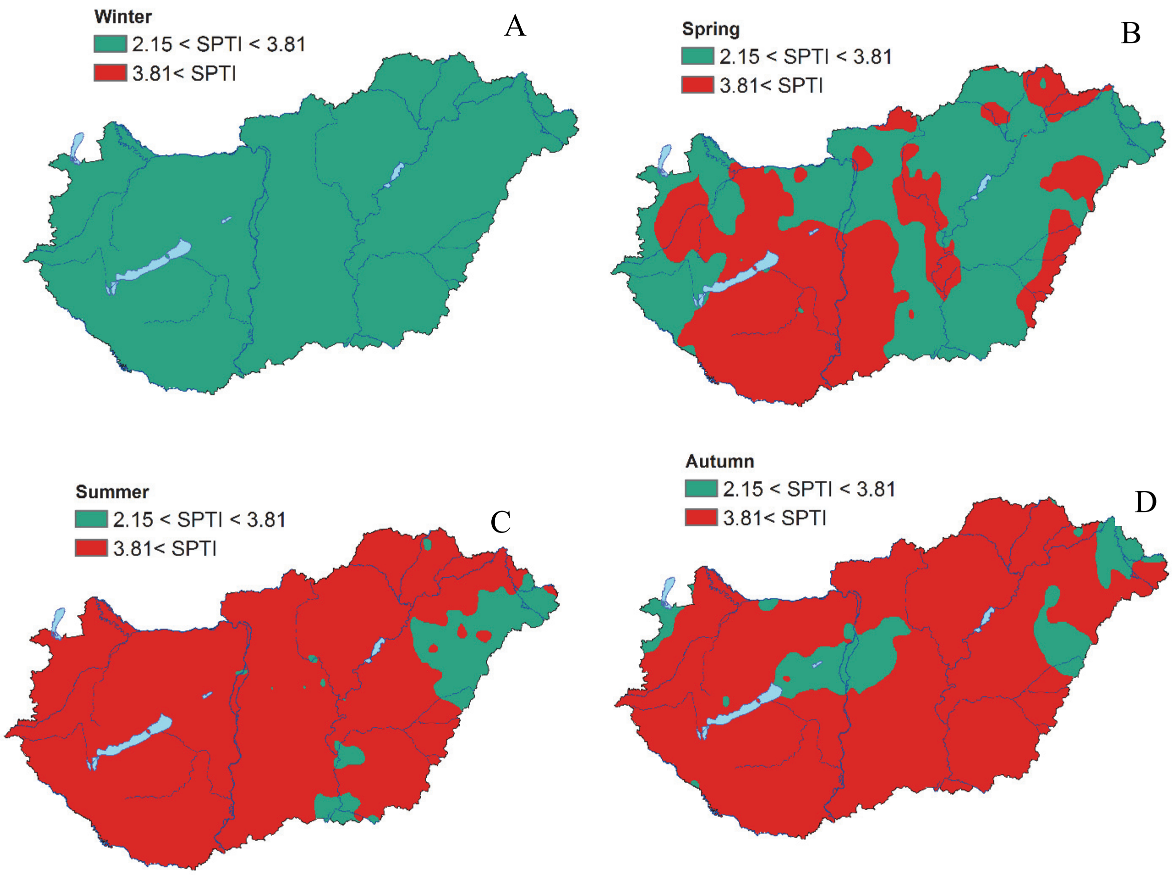

Fig. 8. Maximum SPTI values by season: winter (A), spring (B), summer (C), and autumn (D), based on seasonal precipitation sum and mean temperature, 1871–2020.

Let us look at the spring extremes for grid point data (*Fig. 9*). The hot, dry spring of 1934 is the most unusual spatial average, and the largest area for this year is shown on the map. Dry, hot springs occur over an even larger area in 2003 and 2012. A wet, warm spring is 2010, while the dry, cool feature that was missing in the annual surveys appears here: the springs of 1956 and 1875, for example. There is also an example of a wet, cool spring: 1919 was just such a spring, but it is only shown in one small area on the map (*Fig. 9*).

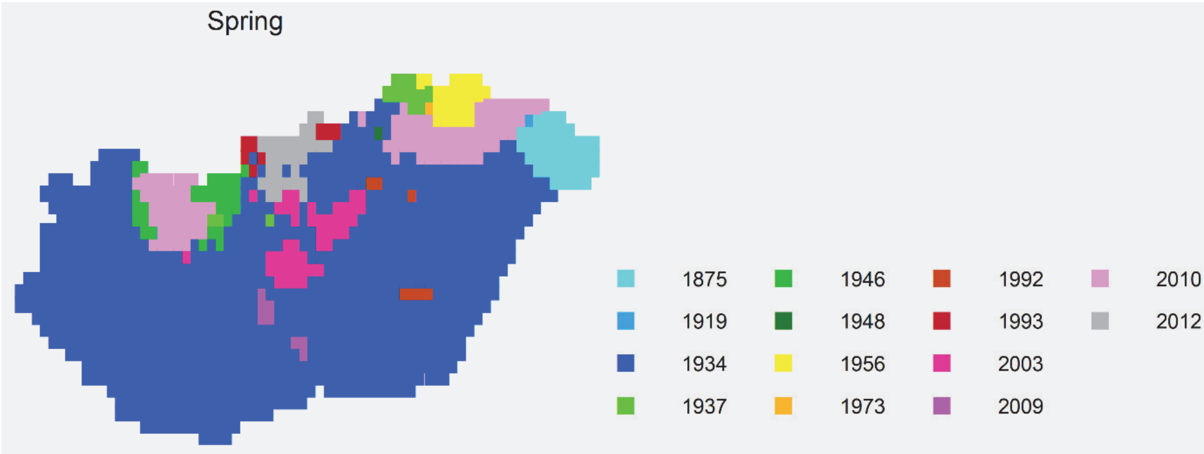

Fig. 9. Spring: year of maximum SPTI values.

The palette of summers is the most colorful (Fig. 10). On spatial average, the dry, cold summer of 1984 is the extreme. Although a large area of *Fig. 10* shows the summer of that year, it is the hot, dry summer of 2003 that is the most extreme over the largest area, precisely in areas where wet, less hot summers are otherwise common, namely the southern and western parts of the Transdanubian region. In general, dry, cool summers dominate the extreme lists, with a large number of dry, hot summers still on the map. Wet, cold summers include 1913 and wet, hot summers include, for example, 1999.

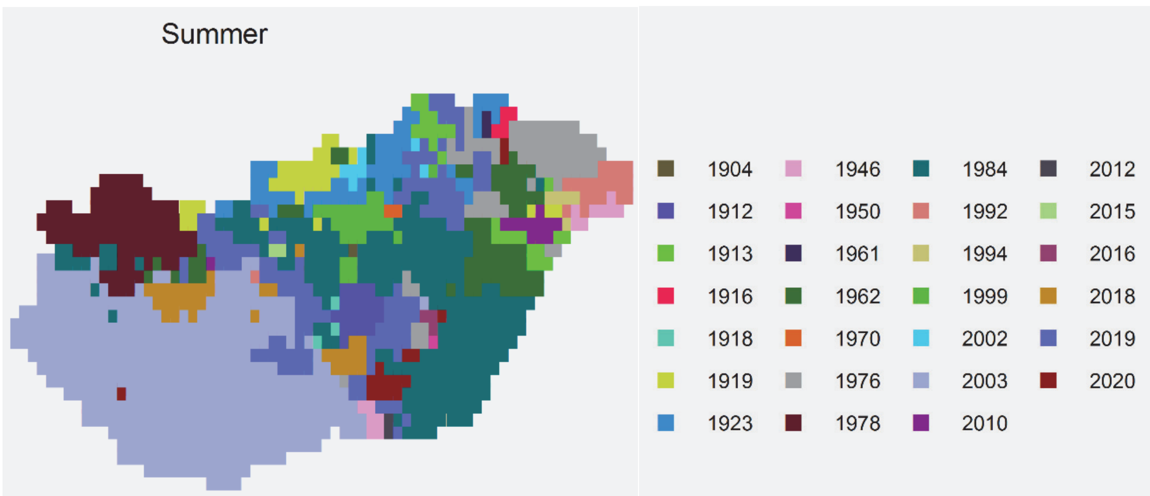

Fig. 10. Summer: year of maximum SPTI values.

The autumn extremes are also shown in *Table 2*. *Fig. 11* shows that the extreme dry and slightly warmer than average autumn of 1986 is the one with the largest area. It can also be said of the years shown in *Fig. 11*, that if we split the country in two parts imaginatively, the western half of the country has the most unusual cold, dry autumns, while the eastern half of the country has the most unusual dry, warm autumns.

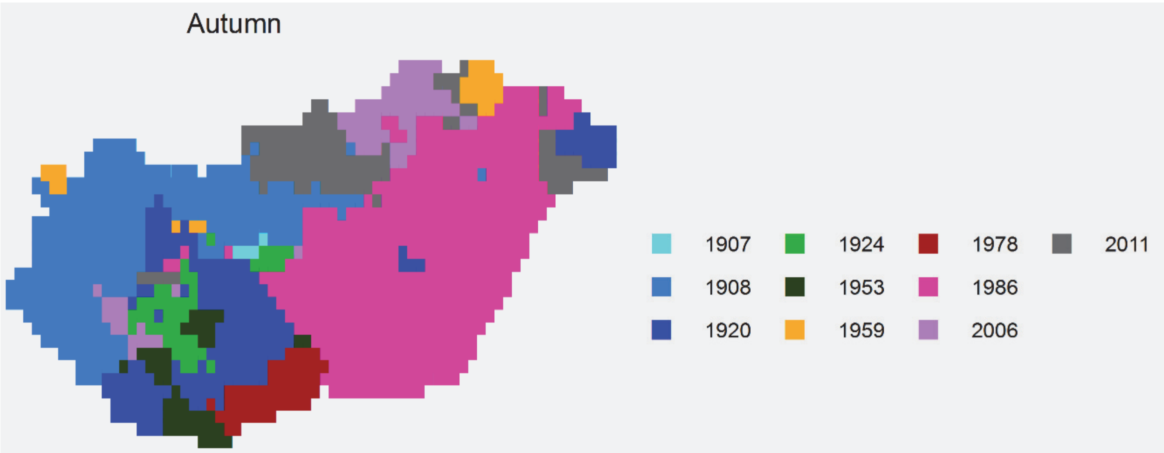

Fig. 11. Autumn: year of maximum SPTI values.

Finally, let us consider the combined extremes of winter (*Fig. 12*). Whether averaged by area or grid point, the *SPTI* indices show that the warmest winter with average precipitation in 2006/2007 is the most unusual of the 150-year-long temperature and precipitation series. Similarly to the summer values, all possible combinations are shown in *Fig. 12*. Warm, wet extremes include the winters of 1909/1910 and 1935/1936, while wet cool extremes include the winter of 1890/1891, but there are also examples of dry, cold winters, such as the winter of 1879/1880.

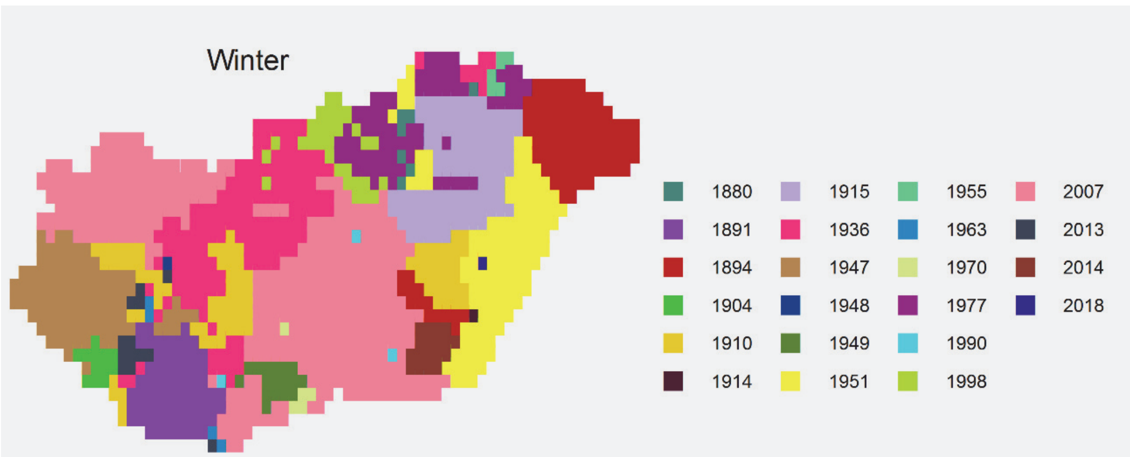

Fig. 12. Winter: year of maximum SPTI values.

6.4 Eight-dimensional application

As indicated in Section 6.1, *SPI3* and *STI3* values were derived from seasonal precipitation sum and mean temperature values for all four seasons. The standardized spatial averages become the components of the eight-dimensional vector variable. To determine the eight-dimensional extreme, the *ST8* statistics were calculated (Sections 2.5, 5.3), and these statistics were applied to the full dataset to decide on the presence of climate change in eight dimensions. Performing the first test (Section 6.2.1), the *TS1* was found to be 6.25, while the critical value is only 1.65, so the null hypothesis of an identical distribution in the eight-dimension cannot be accepted. *Fig. 13* shows that *Cr1*=3.66 is exceeded in a relatively large number of years, but the more stringent *Cr2*=5.20 is only reached in 2018 and 2019. Therefore, we can say that there is climate change with respect to the spatial average, when the mean temperature and precipitation series for the four seasons are considered together. Using an exponential trend estimation for the eight-dimensional norm series, a significant change can be seen, with these values increasing by 37%, shown in green in *Fig. 13*. 2018 was the second warmest and drier than average year, while 2019 was the warmest and slightly wetter than average year. What is striking from *Fig. 13* is that the last decade can be considered an extreme decade. If we plot the probability (Section 5.3) associated with the *ST8* eight-dimensional norm values (*Fig. 14*), this extreme period is more distinct, showing how low probability events have occurred in the past period. In terms of the eight-dimensional norms, 1947 has the third highest norm: it was essentially a dry, hot year. Only winter was wetter and cooler than average, the other three seasons were dry and warm in 1947. The most unusual of the components for 1947 is the autumn precipitation, which is listed as the 5th driest autumn in the 150-year-long time series. The year 1920 and 2020 are not considered extreme on average, but when the eight variables are analyzed together, they can be considered extreme years. In order, they are the fourth and fifth highest *ST8* values. In the year 1920, two seasons were cooler than average and two warmer, while in terms of precipitation, two seasons had above average precipitation and two were extremely dry. In 2020, we also had two very dry seasons and two seasons with above average precipitation, but temperatures were above average in all four seasons.

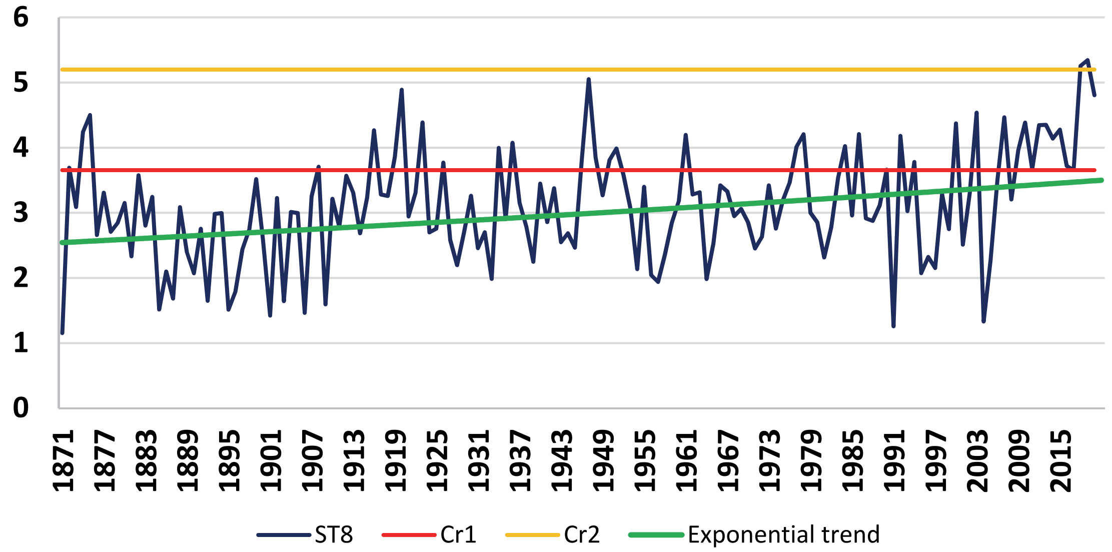

Fig. 13. ST8 values, spatial average.

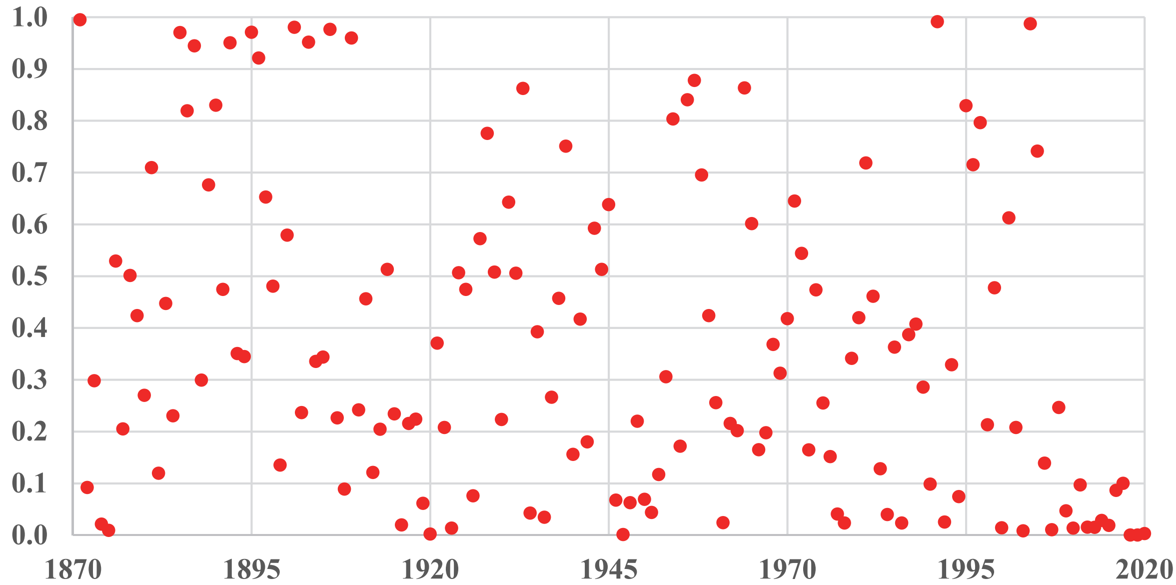

Fig. 14. The probability of ST8 values.

It is also interesting to look at the norm values for the extreme sub-systems (Section 5.2) in *Fig. 15*. For 2019, the two-dimensional extreme sub-system already exceeds the critical value *Cr1* for the eight-dimensional norm values. If we want to know which subsystem is the most unusual, we consider the probabilities in *Fig. 16B*. For the total period and all eight elements, the lowest probability is for the temperature of summer 2003, while the most extreme two-dimensional subsystem, with the addition of the spring drought of the same year, is the most extreme two-dimensional subsystem, with all other extreme subsystems belongs to 2019.

The lowest probability event belongs to the year 2019, of which this value also belongs to the four-element extreme subsystem: warm, dry summer, warm autumn, and wet spring. (*Fig. 16A*)

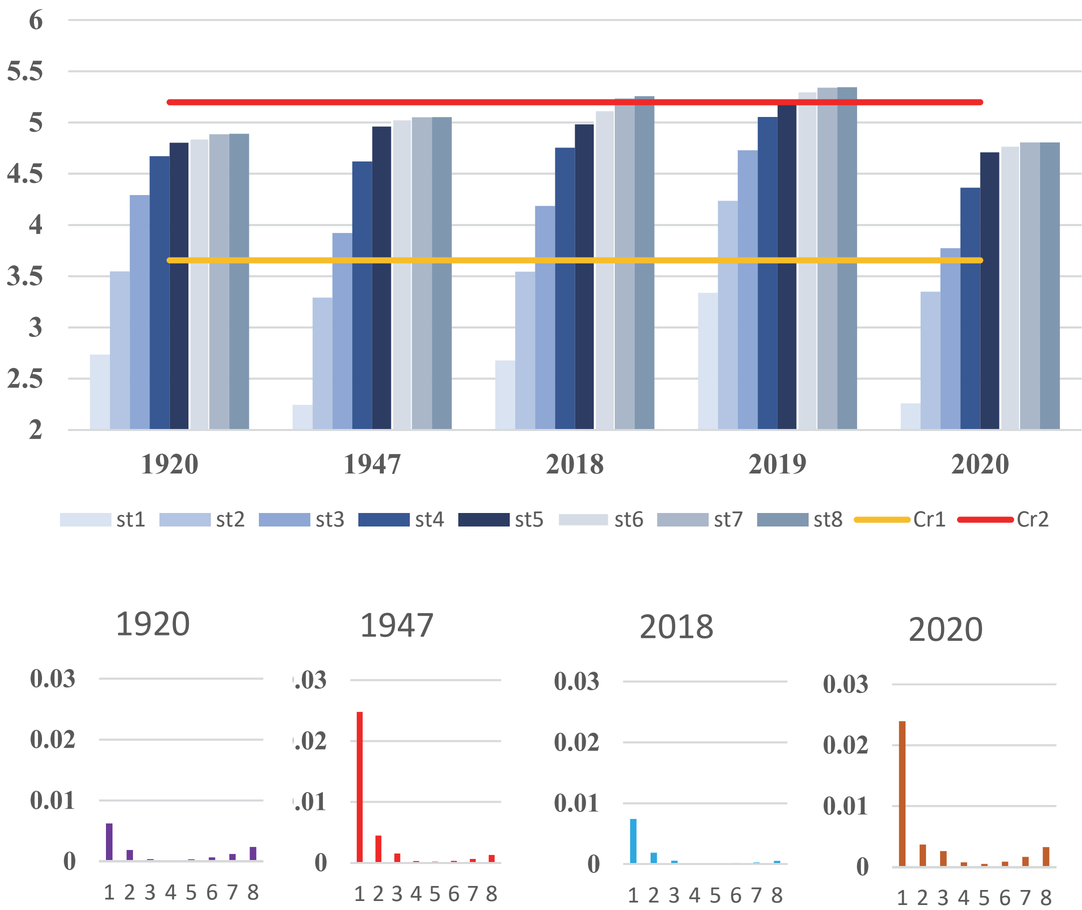

Fig. 15. ST1-8 values for extreme subsystems (top) and their probabilities per year (bottom).

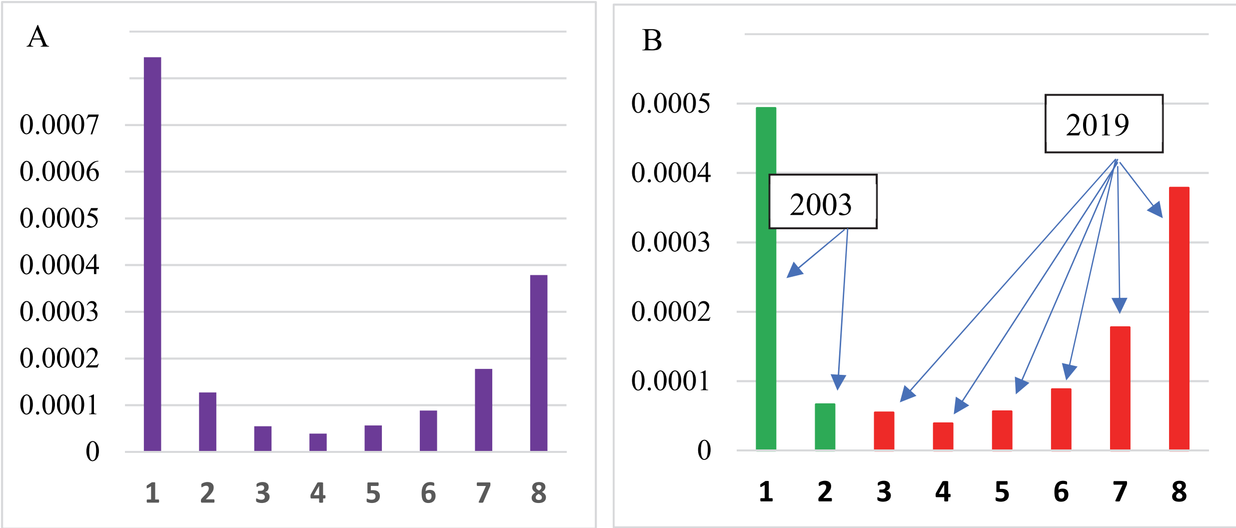

Fig. 16. A: Probability of extreme subsystems in 2019, B: Probability of extreme subsystems of the complete dataset.

If we want to know which seasons and which elements are responsible for the extremes of a given year, we need to analyze the extreme subsystems (Section 5.2), as we did above for the five extreme years (*Figs. 15, 16A*). For most of these, both elements deviated significantly from the average in all four seasons, resulting in high *ST8* norm values. On the other hand, there are years with *ST8* values above the critical value of *Cr1*, but the extremes are caused by only one or two elements. A good example of this is the year 1934 (*Fig 17*), when the probability of the one-dimensional subsystem was the lowest. This is the extreme high mean temperature in spring. Also, *Cr1* is above the norm for the year 1936, but here the two-dimensional extreme subsystem has the lowest probability, and this is the two-dimensional extreme subsystem of a wet winter and a warm spring (*Fig 17*).

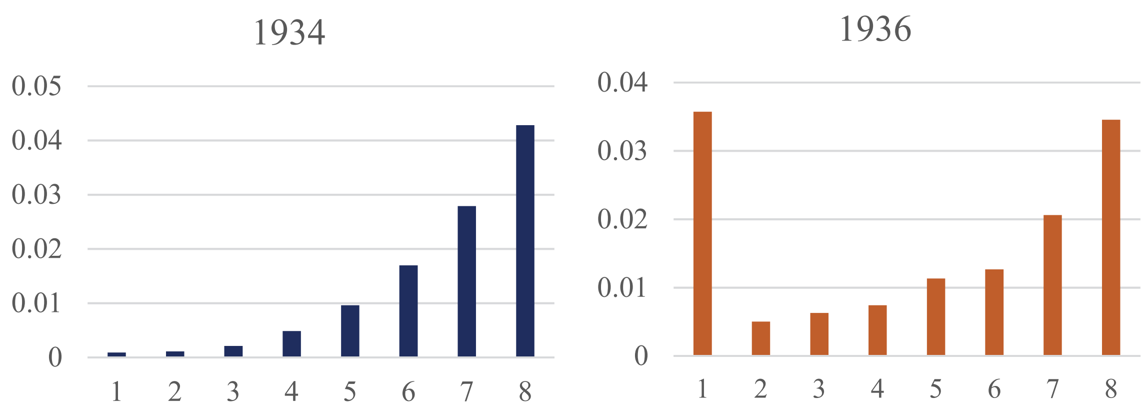

Fig. 17. Probability of ST1-8 norms for the years 1934 and 1936.

7. Summary

The primary objective of this paper was to present a vector norm methodology based on the probability distribution of multidimensional climate time series, which can be used to complement our knowledge of the extremes of climate elements. We have shown why this method was developed in order to study climate elements together and how it works in practice to determine multidimensional extremes.

The results of the statistical tests described in Section 6.2 are summarized in *Table 4*. For the first test (T1 in Section 6.2.1), the change in the distribution of the probability vector variables is not proven only for the winter temperature and precipitation values. For the second test (T2 in Section 6.2.2), the spring values do not exceed the *Cr2* values in addition to the winter values. In the third test (T3 in Section 6.2.3), no unidirectional change is observed for winter and autumn values. In summary, when looking at the two-dimensional annual and summer mean temperature and precipitation values together, all three tests result in rejecting the null hypothesis that there is climate change. The same is true for the four seasonal mean temperature and precipitation values, so that climate change can be detected in eight-dimensions with all three tests. In the case of the spring tests, the T1 and T3 tests were used to detect change, while for the autumn values the T1 and T2 tests were used. Only the winter tests can not demonstrate the change in the probability distribution of the vector variables over time.

Table 4. Summary table of which test and statistics have changes in probability distribution, 1871–2020

	Annual ST2	Winter ST2	Spring ST2	Summer ST2	Autumn ST2	Annual ST8
Test1	✔	X	✔	✔	✔	✔
Test2	✔	X	X	✔	✔	✔
Test3	✔	X	✔	✔	X	✔

In general, we can say that in Hungary cooler years tend to have more precipitation, while drier years tend to be warmer. This makes it unusual to have a very hot, wet year and a dry, cool year. It is also clear that a very dry, very hot year is also unusual, while a very cold, very wet year can be considered extreme. Looking at the temperature and precipitation time series together, we could see examples of each of these cases. Multidimensional extremes included years that could be considered extreme in one dimension, e.g., 2019, 2010, and 1940, but also included years that were not particularly unusual in one dimension, e.g., 1920, 1947, 2020.

Considering the spatial averages, years with hot, rainy weather are the least likely, followed by a series of dry, very hot years and only one case of a cool, rainy year. The top ten did not include a pair of dry, cool years. In the case of winters, the top ten is dominated by warm, wet winters, with dry, warm and cool, wet winters also found in *Table 2*. Most of the spring extremes are associated with dry, warm springs. There is one case of dry, cool spring and two cases of wet, warm spring in our list. For summers, the extremes are the dry, cool and dry, warm seasons. For autumn, these two combinations are also on the list. The exception is the 10th most extreme autumn, which has a very cold period with a lot of rainfall as an extreme. It is now accepted (*Harangi*, 2017), that the eruption of the Katmai-Novarupta volcano in Alaska (June 6, 1912) is responsible for this extreme autumn.

Of course, other elements and dimensions can be investigated further. Two-dimensional studies can be complemented by six-dimensional studies. For example, summers have the largest statistics, in which case it is worth looking at the mean temperature and precipitation totals for the three summer months together. But it is also possible to look at the precipitation sum and temperature of a single month together, or to determine the SPTI index for the summer and winter semesters. After the mathematical description, we have presented only some meteorological applications, the aim of which was to illustrate, through examples, how the theoretical method works in practice.

References

Harangi, Sz., 2017: Vulkánkitörések klímaváltoztató hatása: a kicsi is számít! *Magyar Tudomány 188*, 664–674. (In Hungarian)

Izsák, B. and *Szentimrey, T.,* 2020: To what extent does the detection of climate change in Hungary depend on the choice of statistical methods? *Int. J. Geomath. 11*, 17 https://doi.org/10.1007/s13137-020-00159-7

Izsák, B., Szentimrey, T., Lakatos, M., Pongrácz, R., and *Szentes, O.*: Creation of a representative climatological database for Hungary from 1870 to 2020. *Időjárás 126*, 1–26. https://doi.org.10.28974/idojaras.2022.1.1

Szentimrey, T., 1999: Többdimenziós éghajlati idősorok "extrémumainak" vizsgálata. Időjárási és éghajlati szélsőségek. Meteorológiai Tudományos Napok 1999. OMSZ, Budapest. 77–88. (in Hungarian)

Szentimrey, T., Lakatos, M., and *Bihari, Z.*, 2014: Joint examination of climate variables, Standardized Precipitation and Temperature Index (SPTI)). European Conference on Applied Climatology (ECAC), Prague, Czech Republic, 5-10 October, 2014.

Szentimrey, T. and *Bihari, Z.,* 2014: Manual of interpolation software MISHv1.03. OMSZ, Budapest.

Szentimrey, T., 2017: Manual of homogenization software MASHv3.03. OMSZ, Budapest.

WMO, 2012: Standardized Precipitation Index User Guide. WMO-No. 1090. https://library.wmo.int/index.php?lvl=notice_display&id=13682#.X4lJJ-28qUk