Assessment of Climate Change in Italy by Variants of Ordered Correspondence Analysis

Assuntina Cembalo [1]**, Rosaria Lombardo** [2],*![ID]**, Eric J. Beh** [3]![ID]**, Gianpaolo Romano** [1] **and Michele Ferrucci** [1]
and Francesca M. Pisano [1]

1   Italian Aerospace Research Center (CIRA), 81043 Capua, CE, Italy; a.cembalo@cira.it (A.C.);
    g.romano@cira.it (G.R.); m.ferrucci@cira.it (M.F.); f.pisano@cira.it (F.M.P.)
2   Department of Economics, University of Campania "L. Vanvitelli", 81043 Capua, CE, Italy
3   School of Mathematical and Physical Sciences (Statistics), University of Newcastle, Callaghan, NSW 2308,
    Australia; eric.beh@newcastle.edu.au
*   Correspondence: rosaria.lombardo@unicampania.it; Tel.: +39-0823274032

Abstract: This paper explores climate changes in Italy over the last 30 years. The data come from the European observation gridded dataset and are concerned with the temperature throughout the country. We focus our attention on two Italian regions (Lombardy in northern Italy and Campania in southern Italy) and on two particular years roughly thirty years apart (1986 and 2015). Our primary aim is to assess the most important changes in temperature in Italy using some variants of correspondence analysis for ordered categorical variables. Such variants are based on a decomposition method using orthogonal polynomials instead of singular vectors and allow one to easily classify the meteorological station observations. A simulation study, based on bootstrap sampling, is undertaken to demonstrate the reliability of the results.

1. Introduction

For over a century now, the scientific community has undertaken a huge amount of research focusing on the many aspects of climate change, which remains one of the most challenging problems that humanity shall face in the immediate future. Climate change, or global warming, refers, in part, to the rise in average surface temperatures on Earth. Scientists maintain that climate change is due primarily to the human use of fossil fuels which releases carbon dioxide and other greenhouse gases into the air. The gases trap heat within the atmosphere, which can have a range of effects on ecosystems [1–3]. Indeed, it is now well known that climate change is primarily caused by the actions of mankind and has reached a sizeable scale and affects the whole planet [4–7]. In the Fifth Assessment Report (AR5 2013–2014), the United Nations Intergovernmental Panel on Climate Change (IPCC) states that: *Warming of the climate system is unequivocal, and since the 1950s, many of the observed changes are unprecedented over decades to millennia. The atmosphere and ocean have warmed, the amounts of snow and ice have diminished, sea level has risen, and the concentrations of greenhouse gases have increased* (https://www.ipcc.ch/report/ar5/ accessed on 31 December 2020). IPCC investigations are conducted from temperature and other physical variable observations collected from the mid-19th century through to the present [8]. These observations are obtained mainly through direct measurements and from remote sensing from satellites and other platforms. In particular in Italy, the time series of temperature averages and trends, calculated from 1961 to 2004 [9,10] through three (linear, piecewise and sloped steps) trend regression models of a dataset, show a negative trend for the period 1961–1981, a more pronounced positive trend from 1981 to 2004, and an increase in the average daily temperature range for the whole period. Tomozeiu et al. [11] predict

temperature changes over different Italian agricultural areas, during the periods 2021–2050 and 2071–2100 against the base period 1961–1990, showing that a temperature increase of about 2 °C could be expected during 2021–2050, and between 2.5 °C and 4.5 °C during 2071–2100. Simolo et al. [12] investigate changes in the probability density functions and the probability of moderate extremes for maximum and minimum daily temperature anomalies in Italy from 1951 up to 2008. Their findings suggested that a simple, rigid shift in the density functions alone (without invoking any change in shape) provides evidence of changes in moderate extremes.

No studies exist that demonstrate the utility of correspondence analysis for investigating climate change in Italy. This, it seems, is primarily because the focus of analyzing such changes has involved time series modeling of temperature or precipitation for which suitable predictive models exist in the literature [12–15]. Indeed, correspondence analysis has been considered when investigating climatic change from a socio-economic perspective. For example, Ali et al. [16] used multiple correspondence analysis to analyze car users' attitudes to climate change in UK by referring to the British Social Attitude Survey (BSA) datasets of 2011 and 2014. Brunette et al. [17] analyzed the consequences of climatic changes on *forest adaptation*. However, a variant of correspondence analysis, called detrended correspondence analysis [18], has been used to summarize ecological changes through time in Mexico [19], since climatic and environmental changes *correspond to* changes in both the biotic communities and the fossil assemblages they produce.

To the best of our knowledge, our research here is unique since previous studies have focused on applying time series modeling techniques to study temperature changes in Italy for predictive and probabilistic purposes [12,15]. This paper uses correspondence analysis to highlight the *correspondence* that exists between temperature change and the month in two moments in time (1986 and 2015) and across two different regions of Italy that are characterized by different geographical conditions. Indeed, analyzing the association between temperature and the month of the year can lead to significant changes on a range of agricultural systems (such as viticulture, wheat, fruit and olive cultivation). Therefore, even if the spatial distribution of warming is projected to be quite uniform over Italy by the end of the century [11], highlighting any spatial and seasonal differences that exist over the Italian regions can be useful for addressing suitable agricultural policies, given the climatic vulnerability of a region.

In this paper some climatological investigations are carried out using two variants of ordered correspondence analysis [20,21] for providing an automatic classification of meteorological stations with respect to classes of temperature that have been observed monthly during the first and last year of reference, i.e., 1986 and 2015. We also visually investigate the association between temperature and these yearly time periods using polynomial biplots [22,23]. Finally, a simulation study based on bootstrapped sampling is proposed to show the reliability of the classification results. The paper is organized as follows. Section 2 briefly describes some variants of correspondence analysis for ordered variables suitable, for classifying the meteorological station observations and studying the association between the temperature and months, and for portraying temperature trend in a specific year. Section 3 describes the data preparation and gives the classification of the Italian meteorological stations using the polynomials and biplots of the temperature data. Section 4 show the reliability of ordered correspondence analysis through simulated data. Some final remarks are made in Section 5.

2. Methodology: Simple and Multiple Correspondence Analysis for Ordered Variables

To analyze the association between the months that span over time and classes formed from the temperature change, we consider two recent methodologies, both of which are based on orthogonal polynomials for categorical variables [24–27]. These methodologies are referred to as *ordered multiple correspondence analysis* and *doubly ordered correspondence analysis*. As discussed in Lombardo and Meulman [21], ordered multiple correspondence analysis has the advantage, when compared with the classical approach, to multiple correspondence

analysis [28,29], of automatically classifying the subjects/individuals/observations in as many clusters as the number of ordered categories; for our study the number of clusters is equivalent to the number of temperature classes. On the other hand, Beh [30] and Beh and Lombardo [22] presented doubly ordered correspondence analysis to portray the trends of ordered variables through *polynomial biplots* and, by doing so, adding new findings to the interpretation of the association between ordered categorical variables [23]. Therefore, after performing ordered multiple correspondence analysis for classifying the meteorological stations of two specific years (1986 and 2015) in different classes (highlighting the higher percentages of meteorological stations getting different classes of temperature in 2015 than 1986) we perform doubly ordered correspondence analysis to highlight those months that are mainly associated with higher temperature values. In brief, Section 2.1 discusses ordered multiple correspondence analysis and Section 2.2 provides a brief overview of doubly ordered correspondence analysis.

2.1. Ordered Multiple Correspondence Analysis

Ordered multiple correspondence analysis (OMCA) has been proposed to study the association among multiple ordered variables [21]. With respect to classic multiple correspondence analysis [28,29,31,32], OMCA is based on an alternative method of decomposition, called *hybrid decomposition*, that combines the well known singular value decomposition [33] with the less well known bivariate moment decomposition, discussed by [30]. The data are stored in an indicator (or binary) super-matrix, $\mathbf{X} = [\mathbf{X}_1|...|\mathbf{X}_p]$ of p ordered categorical variables observed on the same set of n observations. The hybrid decomposition is based on the amalgamation of singular vectors and orthogonal polynomials, the latter of which is generated by means of the recurrence formula of Emerson [26]. The orthogonal polynomials are directly linked to the moments of an ordered variable [23,34]. Let $\mathbf{\Psi}$ be the diagonal super-matrix of polynomials and $\mathbf{D}$ the diagonal super-matrix of weights, that are given by the sums of the columns of $\mathbf{X}$. The hybrid decomposition (HD) on the scaled super-indicator matrix $\mathbf{X}$ is defined as:

$$\text{HD}\left(\frac{1}{p\sqrt{n}}\mathbf{X}\mathbf{D}^{-1/2}\right) = \mathbf{\Phi}\mathbf{Z}\mathbf{\Psi}'. \tag{1}$$

where $\mathbf{\Phi}$ is the matrix of singular vectors that are associated with the rows of $\mathbf{X}$, and subject to the usual orthonormality constraint. Similarly, $\mathbf{\Psi}$ is the matrix of the orthogonal polynomials associated with the p ordered variables, subject to the orthonormality constraint with respect to the super-matrix $\mathbf{D}$. In Equation (1) the matrix $\mathbf{Z} = \frac{1}{p\sqrt{n}}\mathbf{\Phi}'\mathbf{X}\mathbf{D}^{-1/2}\mathbf{\Psi}$ represents the link between the singular vectors and the polynomials.

A set of orthogonal polynomials are generated for each variable we are studying. For variable k with j_k categories, there are $j_k - 1$ non-trivial polynomials, the first of which is linear, the second quadratic and so on. The total inertia, i.e., the variable association linked to a chi-squared statistic, can be expressed in terms of the $\mathbf{Z}$ matrix, so we can explore more about the statistical significance of the association between the variables using other features of OMCA. However, for the sake of brevity, we focus here on the main benefit of OMCA, that is, the automatic classification of the meteorological stations given in Tables 3 and 4. Indeed, one of the major advantages of ordered multiple correspondence analysis is concerned with the coordinates, or latent variables, of the individuals; the coordinates for the individuals obtained using orthogonal polynomials allow the analyst to objectively arrange individuals in distinct clusters. For each polynomial, the number of distinct scores is equal to the number of categories. For example, we can group those individuals with equal scores along the first, or linear, polynomial axis. Therefore, we can say that the first cluster includes those Italian meteorological stations with the lowest temperature class (corresponding to the lowest linear polynomial score), the second cluster concerns those station points with the second lowest temperature value (corresponding to the second lowest linear polynomial score), and so on.

After clustering the meteorological stations along the linear polynomial axis, we explore those months that are associated with the ordered classes of temperature using polynomial biplots. For this reason, we perform doubly ordered correspondence analysis.

2.2. Doubly Ordered Correspondence Analysis

When variables have ordered categories, a variant of correspondence analysis has been developed that is particularly suitable for visualizing linear, quadratic or higher order trends of ordered categories. This special variant, called *doubly ordered correspondence analysis*, uses the *bivariate moment decomposition* that makes use of Emerson's orthogonal polynomials rather than the singular value decomposition for the ordered variables. Instead of using the indicator super-matrix $\mathbf{X}$ as in ordered multiple correspondence analysis, we now consider a two-way contingency table $\mathbf{N}$ of dimension $I \times J$ that is a cross-classification of two variables consisting of I row categories and J column categories, respectively. Let the matrix of the joint relative frequencies be $\mathbf{P} = (p_{ij})$ so that $\sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} = 1$. Define the ith row marginal proportion and the jth column marginal proportion, $p_{i\bullet} = \sum_{j=1}^{J} p_{ij}$ and $p_{\bullet j} = \sum_{i=1}^{I} p_{ij}$, respectively, as general elements of the diagonal matrices $\mathbf{D}_I$ and $\mathbf{D}_J$ and of the vectors $\mathbf{r}$ and $\mathbf{c}$. To study the association between rows and columns, the Pearson chi-squared statistic can be considered and is defined as

$$X^2 = n \sum_{i=1}^{I} \sum_{j=1}^{J} p_{i\bullet} p_{\bullet j} \left(\frac{p_{ij}}{p_{i\bullet} p_{\bullet j}} - 1 \right)^2 . \tag{2}$$

In order to graphically depict the association between the rows and the columns (classes of temperature and months) in a low dimensional space, we may consider the bivariate moment decomposition of terms from Pearson's statistic that we now briefly describe. After computing the orthogonal polynomials for each ordered variable (classes of temperature and months), we perform the bivariate moment decomposition (BMD) so that

$$\text{BMD}\left[\mathbf{D}_{\mathbf{I}}^{-1} \left(\mathbf{P} - \mathbf{rc}' \right) \mathbf{D}_{\mathbf{J}}^{-1} \right] = \mathcal{A} \mathbf{Z} \mathcal{B}^T \tag{3}$$

where $\mathcal{A}$ and $\mathcal{B}$ are the row and column polynomial matrices, respectively, and defined using Emerson's recurrence formulae [26], and $\mathbf{Z}$ is the matrix of the generalized correlations [34,35]. The row and column polynomials are orthonormal with respect to the weight matrices $\mathbf{D}_I$ and $\mathbf{D}_J$, respectively. The construction of the row and column matrices of polynomials $\mathcal{A}$ and $\mathcal{B}$ requires the specification of the marginal proportions and of a priori ordered scores to reflect the ordinal structure of the row and column variables. The generalized correlation matrix, $\mathbf{Z}$, in the BMD reflects the various sources of association between the variables and are derived using Emerson's orthogonal polynomials; for further detail, see [22]. For example, when the row and column scores are defined as consecutive integers, such that $s_X(i) = i$ (for $i = 1, ..., I$) and $s_Y(j) = j$ (for $j = 1, ..., J$), then z_{11} is Pearson's product moment correlation of $\mathbf{P}$ and describes the linear-by-linear association between the row and column categories. Similarly, z_{12} describes the linear-by-quadratic association between the row and column categories and is a measure of the correlation between any change in the location of the row categories and dispersion of the column categories. As for OMCA, the advantage of this strategy is that a formal statistical test of the X^2 statistic and the sources of association (polynomial components of inertia) can be made. To test the statistical significance of the total inertia we can compare the chi-squared statistic with the χ^2 distribution with $(I-1)(J-1)$ degree of freedom and each term of $\mathbf{Z}$ can be assessed for statistical significance; see, for example, ref. [22] for further details. However, here we prefer to focus on the graphical representation of the association.

Polynomial Biplots

In the literature [22,23,30] polynomial biplots have been proposed to graphically represent trends of ordered categorical variables which can be treated as being symmetrically or asymmetrically associated. When constructing a polynomial biplot, the ordered row and column categories can be displayed in a single plot since the row and column coordinates are computed with respect to the same set of polynomial axes. For example, in a polynomial *row metric preserving* (or row isometric) biplot, the column categories are visually depicted using standard polynomial coordinates and the rows are depicted using principal polynomial. For a polynomial *column metric preserving* (or column isometric) biplot, the column categories are visually depicted using principal polynomial coordinates while the rows are visualized using standard polynomial coordinates. However, the coordinates for both the row and column categories are computed using the same set of orthogonal polynomial axes. For example, when using the first two polynomials, the standard coordinates will be ordered on the linear polynomial axis and the origin will represent the mean of the ordered categories. Along the second polynomial axis, the standard coordinates will show a parabolic trend and the associated polynomial principal category coordinates will depend on them as well as on the generalized correlations. This particular interpretation is not present in the classical forms of correspondence analysis that use singular value decomposition. Hence, the coordinates and their proximity from the origin using ordered correspondence analysis have a distinctive interpretation from those used in the classical approaches to correspondence analysis.

3. Application: Climate Change

3.1. Data Preparation

The original data were available in NetCDF format. The NetCDF file has a vectorial structure: there are three vectors of attributes (latitude and longitude to set a point in the space, plus a time vector that is a sequence of integers counting the number of days since 1 January 1950) and a very long vector of numbers that represents the specific climate variable of interest. Once the data were structured in a tabular format, the next step was to develop an Extract, Transform and Load (ETL) transformation using the Pentaho Data Integration tool (Available online: http://community.pentaho.com/projects/data-integration/ (accessed on 8 May 2012)). The main scope of the transformation was to discretize monthly mean temperatures from a set of meteorological stations (grid points) of two specific Italian regions (i.e., Lombardy and Campania) for studying the changing seasonal temperature distributions using variants of correspondence analysis [22]. The analysis of climate change in Italy is based on data from the European observation gridded dataset (E-OBS). This dataset is a free set of reference climatic data widely used for the assessment of climate modelling in Europe. The E-OBS dataset contains daily observations of precipitation, temperature and pressure at sea level across Europe from 1 January 1950 to 31 August 2016, based on European Climate Assessment & Dataset (ECA&D) information (Available online: http://www.ecad.eu/ (accessed on 31 December 2020)). Specifically, the E-OBS dataset contains gridded data for 5 climatic features: daily mean temperature, daily minimum temperature, daily maximum temperature, daily precipitation total and daily average sea level pressure (available since version 14.0 of E-OBS). In particular, we recovered values of mean daily temperature in degree Celsius for Italy in 1986 and 2015. For the sake of brevity, our analysis of the data will be confined to the regions of Lombardy and Campania. In Tables 1 and 2, we summarize the monthly average temperature in 1986 and 2015 of Lombardy and Campania, respectively; in bold are the most relevant temperature increases in 2015. In particular, Table 1 shows that, in Lombardy in 2015, the largest increase in average temperature (an increase of 2 °C from the mean temperature recorded in 1986 was statistically significant with a *p*-value = 0.111) occurred over January, February, April, July and December.

Figure 1 shows the temperature averages over 42 meteorologic stations during the 12 months of the year 1986 (left side of Figure 1-plot a) and 2015 (right side of Figure 1-

plot b). Observe that the highest temperature values are recorded during August 1986. Different from 2015, the highest temperature values are registered in July 2015.
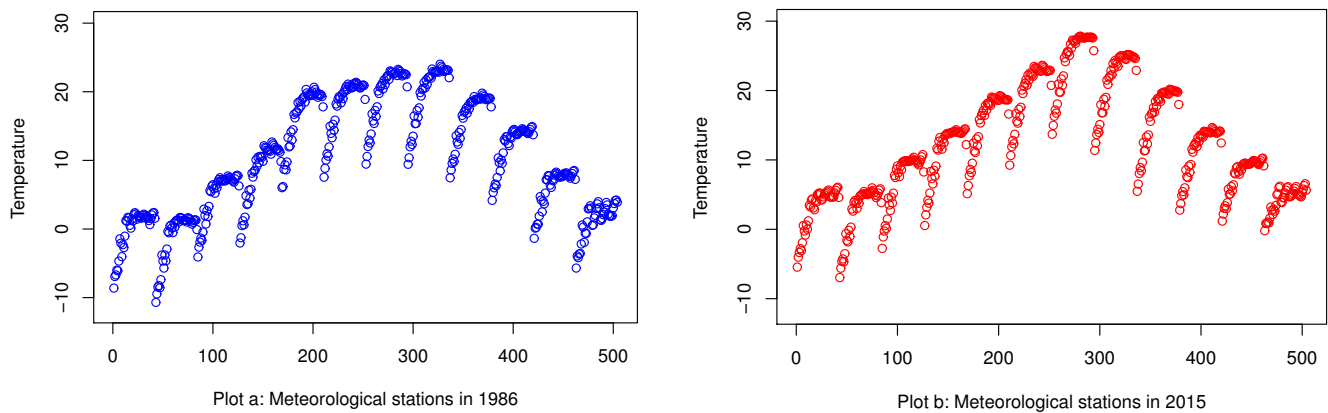
Figure 1. Average temperature in Lombardy over 12 months, recorded in 42 meteorologic stations. Plot a refers to the year 1986 and plot b to the year 2015.

Table 1. Average temperature over 42 meteorologic stations of Lombardy in 1986 and 2015. The values highlighted in bold show the most relevant temperature increases in 2015.

Month	1986	2015
January	−0.09	**2.88**
February	−1.2	**2.55**
March	4.94	7.26
April	8.36	**11.19**
May	16.56	15.73
June	17.94	20.25
July	19.81	**24.49**
August	19.94	21.63
September	16.42	16.35
October	12.19	11.46
November	6.1	7.84
December	1.29	**4.29**

Table 2 summarizes the mean monthly temperature in Campania and shows that the largest increases were observed in 2015 during June, July, and December. In 2015, an increase of 1.9 °C from the mean temperature recorded in 1986 was statistically significant with a *p*-value = 0.104.

Figure 2 shows the average temperature over 23 meteorologic stations during the twelve months of 1986 (left side of Figure 2—plot a) and of 2015 (right side of Figure 2—plot b). Observe that, also in Campania, the highest temperature values are recorded during August 1986, while in 2015, the peak of average temperature values are registered in July. Comparing Figures 1 and 2, we can observe that the distribution of average temperature values over the respective meteorologic stations has a different shape, showing more variability in Lombardy than in Campania, which is not surprising given the morphological diversity of the two regions.

Table 2. Average temperature over 23 meteorologic stations of Campania in 1986 and 2015. The values highlighted in bold show the most relevant temperature increases in 2015.

Month	1986	2015
January	6.44	8.11
February	5.36	7.27
March	9.63	9.98
April	12.57	13.03
May	18.31	18.30
June	18.95	**21.76**
July	22.55	**26.75**
August	24.54	25.37
September	20.26	21.75
October	16.67	16.57
November	11.28	13.32
December	6.42	**10.57**

Plot a: Meteorological stations in 1986

Plot b: Meteorological stations in 2015
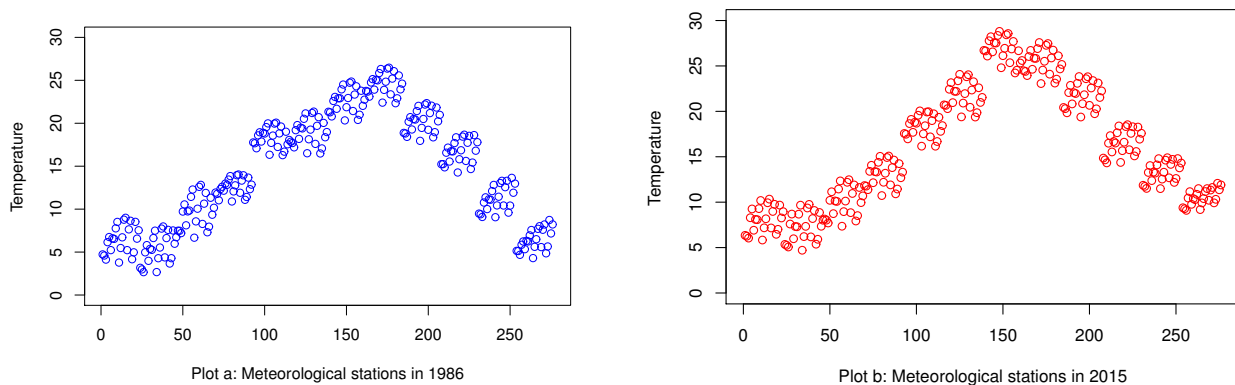
Figure 2. Average temperature in Campania over 12 months recorded in 23 meteorologic stations. Plot a: referred to the year 1986 and plot b to the year 2015

To further investigate the issues concerned with climate change in these two Italian regions, we considering classes of temperature that have been observed monthly in 1986 and 2015, and perform some variants of ordered correspondence analysis.

Note that the extremes of temperature classes for both regions was determined using the quantiles of the temperature distribution (where the quantiles are cut points dividing the range of a probability distribution into continuous intervals with equal probabilities). A two-way contingency table was then constructed for each Italian region, which cross-tabulates the temperature classes and the months, for two specific years (1986 and 2015). From this contingency table a super-indicator table was created to perform ordered multiple correspondence analysis (OMCA).

By first performing OMCA, we are able to classify the meteorologic stations according to the temperature classes. In doing so, we can highlight the increased number of meteorological stations in 2015 with respect to 1986 for a particular range of temperature values; see Tables 3 and 4. Secondly, we perform doubly ordered correspondence analysis (DOCA) for identifying those months that observed, for a particular temperature class, an increased number of meteorological stations in 2015 when compared with their number in 1986.

The temperature distribution is slightly different for the two regions due to differences in the specification of the data. The two correspondence analysis variants, OMCA and DOCA, were implemented using the R packages `MCAvariants` and `CAvariants` [36], respectively; these packages are available from the Comprehensive R Archive Network (CRAN) at https://CRAN.R-project.org/package (accessed on 31 December 2020)) .

3.2. Lombardy Region

The Lombardy region, owing to its peculiar geography (extended mountain areas, steep valleys and high-flow rivers) and to its socio-economic features (employment, health, agriculture, industry), is subjected to a high vulnerability to the impacts of climate change. Past trends and future projections suggest a marked rise in mean temperature, as highlighted in Table 3.

Table 3. Percentage of meteorological stations in temperature classes for Lombardy in 1986 and 2015. Those percentages of meteorological stations that are higher than 1% in 2015 with respect to 1986 are shown in bold.

Temperature	1986	2015
(−10.9, −1.6]	7.5	6.2
(−1.6, 1]	7.4	6.6
(1, 2.3]	10.2	5.5
(2.3, 3.5]	5.7	6.6
(3.5, 4.5]	5.6	**6.9**
(4.5, 5.7]	4.8	**12**
(5.7, 6.7]	5.8	3.5
(6.7, 8.2]	9	4.4
(8.2, 9.5]	5.8	**8.3**
(9.5, 10.9]	6.2	**8.1**
(10.9, 12.2]	6.5	6.1
(12.2, 13.4]	5.8	**6.8**
(13.4, 14.6]	3.2	**5**
(14.6, 16.4]	1.3	1.4
(16.4, 18.1]	1.9	2.7
(18.1, 19.5]	3.9	3
(19.5, 20.9]	3.7	2.7
(20.9, 22.6]	2.7	1.6
(22.6, 24.1]	3	2.6

The Lombardy region includes 42 meteorological monitoring stations so that in every 12 month period we have a total of 504 observations. For this, the temperature is divided into the 19 classes, detailed in Table 3. The column variable, *Month*, comprises 12 months, from January to December. By using OMCA, we first calculate the indicator super-matrix (with $p = 2$ number of ordered variables) according to the ordered temperature classes and the month of reference. As a result, 504 meteorological stations were objectively classified in 19 clusters.

In Table 3, we highlight in bold the highest percentages of meteorological stations recorded during 2015. It is clear that there was an increase (higher than 1%) in the number of meteorological stations in 2015 compared with 1986 for only some temperature classes. In particular there was an increase in the number of stations that recorded temperatures that lie within the following *critical* classes (3.5, 4.5], (4.5, 5.7], (8.2, 9.5], (9.5, 10.9], (12.2, 13.4], (13.4, 14.6] .

In order to determine those months that were mainly characterized by these temperature changes, we perform doubly ordered correspondence analysis (DOCA) using the two-way contingency tables for 1986 and 2015.

The total inertia, or strength of association between the temperature and month is formally tested using the X^2 statistic. Its value is 1589.29 with 198 degrees of freedom, and thus, there exists a statistically significant association between the two variables. While we have not shown it here, all polynomials for both variables are statistically significant for modeling the association.

To portray the association between the temperature classes and the months, we construct polynomial biplots. On the left and right side of Figure 3, plot a and plot b

show the polynomial biplots for the year 1986 and 2015, respectively. Plot a of Figure 3 highlights the strong association between the temperature class (19.5, 20.9] and the months of July and August, while, plot b of Figure 3 shows that, in 2015, the association between the temperature classes and months has changed. In particular, plot b of Figure 3 shows that the temperature class (19.5, 20.9] is strongly associated with the months of June and August. Furthermore, the cold months of January and February are associated with the warmer temperature class (2.3, 3.5] (instead of (−1.6, 1] in 1986). With respect to the *critical* temperature classes, plot b of Figure 3 shows a very strong association between the temperature class (3.5, 4.5] and December, and between (12.2, 13.4] and April. In agreement with the findings discussed by [11], the average temperature increase recorded in Lombardy of about 2 °C in April, and of about 2.2 °C in December shows statistically significant results.
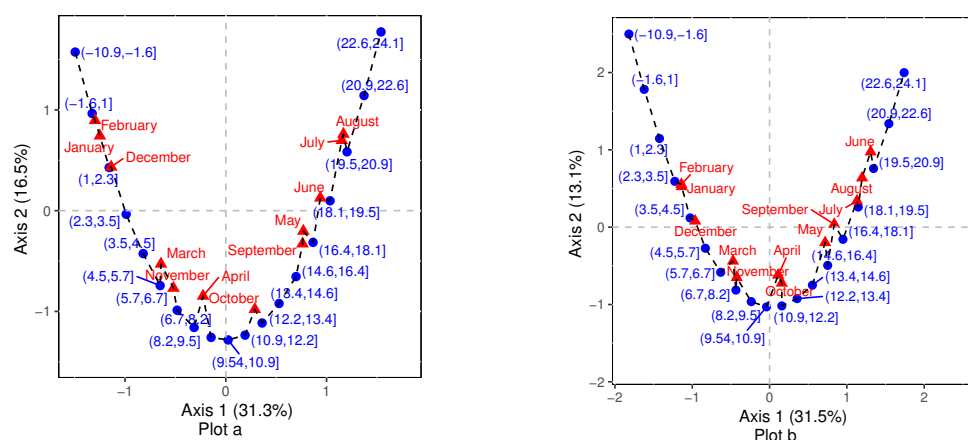
Figure 3. DOCA polynomial biplot. Temperature and month association in Lombardy. Plot a refers to the year 1986 and plot b to the year 2015.

Understanding those *critical* months that experienced a climatic change can allow for an assessment of the vulnerability of the socio-economic and natural systems. Doing so paves the way for the implementation of a regional adaptation strategy.

3.3. Campania Region

The region of Campania, which lies in the southern part of the Italian peninsula, has an area of about 13,500 km^2, and a coastline along the Tyrrhenian Sea. The landscape is dominated by the Apennine mountain ranges, reaching altitudes of 1000–2000 m, and accounts for 32% of the land area. The region of Campania has a Mediterranean climate, affected by the Azores, Siberian and South African anticyclones and the Aleutian and Icelandic lows, with hot, dry summers and moderately cool rainy winters. For Campania from January to December, temperature values are divided into the 20 classes detailed in Table 4. The Campania region includes 23 meteorological stations collecting data during the 12 months of the year for a total of 276 observations in a year. We consider again 1986 and 2015 as our reference years. By performing OMCA, with $p = 2$ ordered categorical variables (the temperature and the months of the year) we have 276 observations that were automatically classified. Of course, this classification does not give the same result for both years, so that a comparison of the distribution differences are of particular interest for defining good climate adaptation practices. In Table 4, the percentages of meteorological stations higher than 1% in 2015 with respect to 1986 are highlighted in bold. Additionally, in Campania, the increase in the number (percentages) of observations in 2015 with respect to 1986 concerns only some classes of temperature; those being (8.7, 9.6], (9.6, 10.4], (11.2, 12.2], (14.5, 16] and (25.1, 28.8], of course, these *critical* classes of temperature of Campania are not the same of Lombardy.

To identify what months were strongly associated with these *critical* temperature classes that recorded an increase in the number of meteorological stations over the period, we perform DOCA.

The total inertia is 1092.1 for the two-way contingency table formed by cross-classifying temperature class and month for Campania in 2015 with 209 degrees of freedom, there exists a strong association between the temperature class values and the months. To visually represent the association between the two variables in 1986 and 2015, we consider the two plots of Figure 4. Plot a of Figure 4 shows the visual summary of the association for 1986. It shows that there exists a strong association between the temperature class (21.8, 22.9] and the month of July, while plot b of Figure 4 shows, for 2015, the association between July and the warmer temperature class (25.1, 28.8]. These conclusions are in agreement with the findings discussed by [11], the average test of temperature recorded in Campania shows an increase of about 4 °C in July and December that is statistically significant. To analyze the correspondence between those months that were mainly characterized by these temperature changes, we perform doubly ordered correspondence analysis (DOCA) using two-way contingency tables for 1986 and 2015. Furthermore, we observe the association between the *critical* temperature classes in Campania (which recorded an increased number of observations in 2015) with the months; plot b of Figure 4 shows a strong association between (8.7, 9.6] and January, (9.6, 10.4] and March, (11.2, 12.2] and April, (14.5, 16] and October and (25.1, 28.8] and July.

Note also that, in 2015 (plot b of Figure 4), the variability of the warmer temperature values (see class (25.1, 28.8]) is lower when compared with 1986 (plot a of Figure 4) being the polynomial coordinate along the second axis lower. Conversely, the variability in the colder temperature values (see class (1.7, 5.8]) has increased relative to the temperatures recorded in 1986. These results highlight a shift in the temperature distribution and is in agreement with previous studies on temperature extremes in Italy [12].

Table 4. Percentage of meteorological stations in temperature classes for Campania in 1986 and 2015. Those percentages of meteorological stations higher than 1% in 2015 with respect to 1986 are in bold.

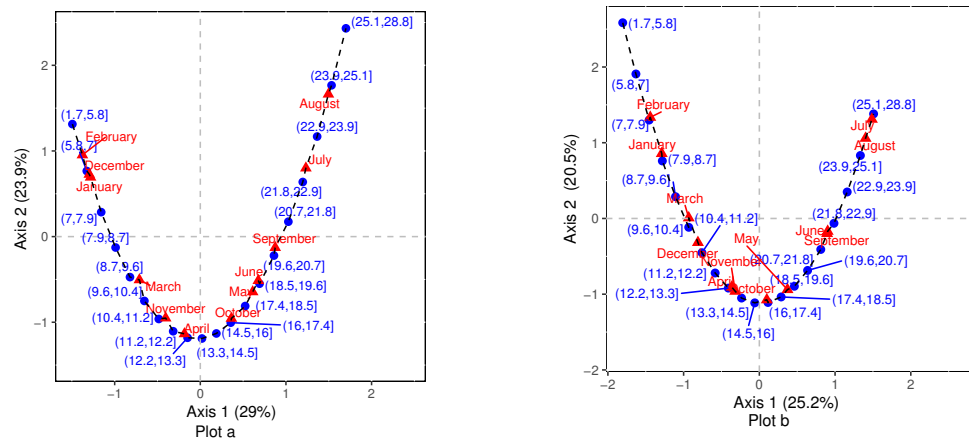| Temperature	1986	2015
(1.7, 5.8]	9.8	5.1
(5.8, 7]	7.4	6
(7, 7.9]	6.9	6.2
(7.9, 8.7]	6.2	6.3
(8.7, 9.6]	5.6	**6.7**
(9.6, 10.4]	5.4	**7.4**
(10.4, 11.2]	6.2	6.2
(11.2, 12.2]	6.5	**7.8**
(12.2, 13.3]	6.7	6.7
(13.3, 14.5]	6.2	7.1
(14.5, 16]	5.6	**7.1**
(16, 17.4]	6.9	6.2
(17.4, 18.5]	3.6	2.5
(18.5, 19.6]	3.8	1.8
(19.6, 20.7]	2.9	2.7
(20.7, 21.8]	2.5	1.4
(21.8, 22.9]	2	2.3
(22.9, 23.9]	2.7	2.5
(23.9, 25.1]	1.8	2
(25.1, 28.8]	1.3	**6**

Figure 4. DOCA polynomial biplot. Temperature and month association in Campania. Plot a is reflects the year 1986 and plot b the year 2015.

4. Simulation: Climate Change

A simulation study has been performed to assess the reliability and stability of the clusters for Lombardy and Campania regions. A total of 1000 bootstrap samples have been generated from the original temperature data in order that 1000 contingency tables were generated by fixing the table marginals and hypothesizing the multinomial distribution of the categorical variables. What is interesting is that the average number of observations for each class of temperature is equal to the original average in each class for both regions.

Table 5 includes the average and standard deviations of meteorological stations for all classes of temperature in the Lombardy region in 1986 and 2015.

Table 5. Data Simulated: Mean and standard deviation of 1000 meteorological bootstrapped observations of Lombardy for temperature classes in 1986 and 2015. The mean values highlighted in italics are slightly different from the original table. The standard deviations (sd) highlighted in bold are those of the *critical* temperature classes.

	Temperature	1986-Mean	1986-sd	2015-Mean	2015-sd
1	$(-10.9, -1.6]$	7.5	0.8	6.2	0.7
2	$(-1.6, 1]$	*7.5*	0.7	6.6	0.7
3	$(1, 2.3]$	10.2	0.7	5.5	0.6
4	$(2.3, 3.5]$	5.7	0.6	6.6	0.6
5	$(3.5, 4.5]$	*5.7*	**0.6**	6.9	**0.6**
6	$(4.5, 5.7]$	*5.0*	**0.6**	*11.9*	**0.8**
7	$(5.7, 6.7]$	5.8	0.6	3.5	0.5
8	$(6.7, 8.2]$	9.0	0.7	4.4	0.5
9	$(8.2, 9.5]$	*5.6*	**0.6**	*8.4*	**0.7**
10	$(9.5, 10.9]$	*6.4*	**0.6**	*8.0*	**0.7**
11	$(10.9, 12.2]$	*6.3*	0.6	*6.3*	0.6
12	$(12.2, 13.4]$	5.8	**0.6**	6.8	**0.6**
13	$(13.4, 14.6]$	3.2	0.5	5.0	0.5
14	$(14.6, 16.4]$	1.3	0.3	1.4	0.3
15	$(16.4, 18.1]$	*2.1*	0.4	*2.5*	0.4
16	$(18.1, 19.5]$	3.9	0.5	2.9	0.4
17	$(19.5, 20.9]$	3.7	0.5	2.7	0.4
18	$(20.9, 22.6]$	2.7	0.4	1.6	0.3
19	$(22.6, 24.1]$	3	0.4	2.6	0.4

Given the findings of the bootstrapped data in Table 5, we can confirm the reliability of the percentage of meteorological stations in each temperature class for Lombardy and for the both reference years. Only small differences between the bootstrapped and the original data table for some temperature classes are highlighted in italics.

Similarly, Table 6 shows the average values and standard deviations of the number (percentage) of meteorological stations in Campania for all classes of temperature in the two reference years. The average of 1000 percentages for each temperature class (columns 1 and 3) is identical to the original percentage of each temperture class given in Table 4. As a result of the summaries in Table 6, there are no mean values marked in italics suggesting that the simulated results compare very well with the observed temperature recordings. Furthermore, the standard deviation for each temperature class is very low, confirming the reliability of the percentages of meteorological stations for each temperature class.

Table 6. Data Simulated: Mean and standard deviation of 1000 meteorological bootstrapped observations of Campania for temperature classes in 1986 and 2015. The standard deviations (sd) highlighted in bold are those of the *critical* temperature classes.

	Temperature	1986-mean	1986-sd	2015-mean	2015-sd
1	(1.7, 5.8]	9.8	0.8	5.1	0.6
2	(5.8, 7]	7.4	0.8	6	0.7
3	(7, 7.9]	6.9	0.7	6.2	0.6
4	(7.9, 8.7]	6.2	0.6	*6.3*	0.6
5	(8.7, 9.6]	5.6	**0.6**	6.7	**0.6**
6	(9.6, 10.4]	5.4	**0.6**	7.4	**0.7**
7	(10.4, 11.2]	6.2	0.6	6.2	0.6
8	(11.2, 12.2]	6.5	**0.6**	7.8	**0.7**
9	(12.2, 13.3]	6.7	0.6	6.7	0.6
10	(13.3, 14.5]	6.2	0.6	7.1	0.6
11	(14.5, 16]	5.6	**0.6**	7.1	**0.7**
12	(16, 17.4]	6.9	0.6	6.2	0.6
13	(17.4, 18.5]	3.6	0.5	2.5	0.4
14	(18.5, 19.6]	3.8	0.5	1.8	0.3
15	(19.6, 20.7]	2.9	0.4	2.7	0.4
16	(20.7, 21.8]	2.5	0.4	1.4	0.3
17	(21.8, 22.9]	2	0.4	2.3	0.4
18	(22.9, 23.9]	2.7	0.4	2.5	0.4
19	(23.9, 25.1]	1.8	0.3	2	0.4
20	(25.1, 28.8]	1.3	**0.3**	6	**0.6**

From the simulation study, Figures 5 and 6 illustrate the 19 box-plots concerning the 19 classes of average temperature for the Lombardy region in 1986 and 2015. These figures show the presence of many outliers as well as the different distribution of the meteorological stations between the two years of reference. For example, the median of the observations in class 6 is greatly increased from 1986 to 2015.

For the Campania region, Figures 7 and 8 show the 20 box-plots concerning the observation numbers and the 20 classes of temperature in 1986 and 2015. Additionally, for this region, the climate change between the two reference years for the different temperature classes is very evident. For example, the median of the observations in class 20 (highest temperature) is greatly increased from 1986 to 2015.
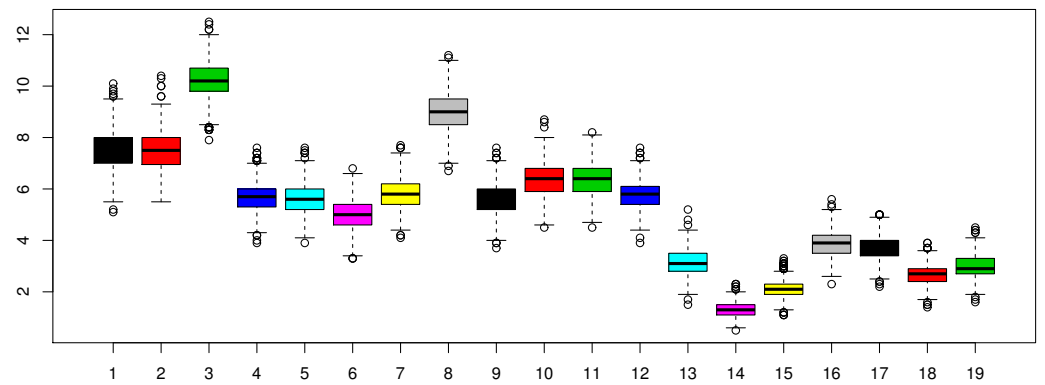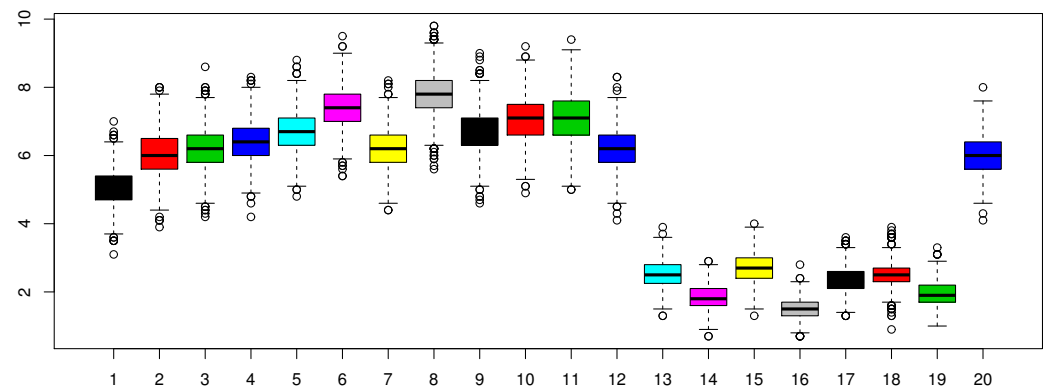
Figure 5. Lombardy, bootstrapped observations referring to 1986. Distribution of the number of meteorological stations in each of the 19 temperature classes.
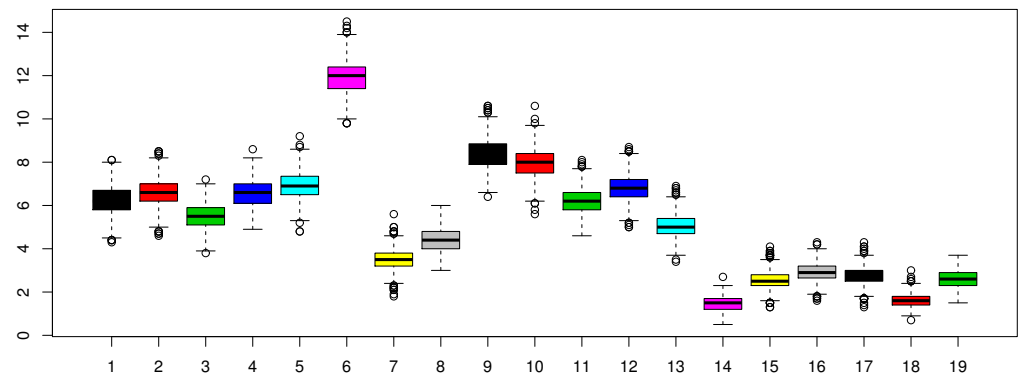
Figure 6. Lombardy, bootstrapped observations referring to 2015. Distribution of the number of meteorological stations in each of the 19 temperature classes.
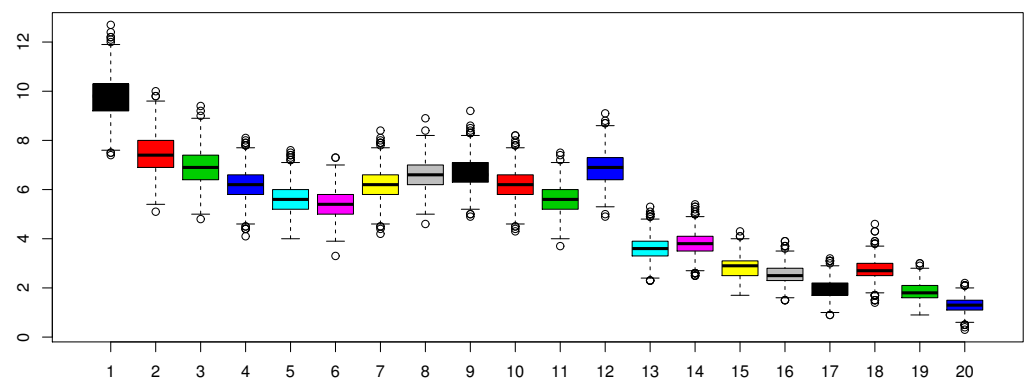
Figure 7. Campania, bootstrapped observations referring to 1986. Distribution of the number of meteorological stations in each of the 20 temperature classes.

Figure 8. Campania, bootstrapped observations referring to 2015. Distribution of the number of meteorological stations in each of the 20 temperature classes.

5. Conclusions

The challenge of climate change requires a clear response strategy: apart from the adoption of measures to reduce greenhouse gas emissions (i.e., mitigation policies), there should be more actions taken to reduce the vulnerability of the local natural and socio-economic systems. Doing so will increase their resilience towards the inevitable impacts of a changing climate (i.e., adaptation measures). Climate change and its impact on growth and development in recent years is of central importance in the overall context of spatial planning, given the territorial vulnerability arising from risks associated with flooding and desertification. While analyzing data concerned with climate change we demonstrate the utility of some variants of ordered correspondence analysis for objectively classifying meteorological station observations [21] and visually exploring the total association using polynomial biplots. Indeed, the orthogonal polynomials [24–26] allow the analyst to portray the complex dependence associations that exist in the ordered data. Polynomial biplots are visual displays that can be constructed as part of an investigation of the data using ordered correspondence analysis [23,30]. Such plots visually portray the association between the classes of temperature and months of a specific year of study. As a result, the visualization of the changes in temperature for particular months allows for one to assess the most important differences across Italy and over time. The reliability of the ordered correspondence analysis results have been assessed through bootstrapped samples. Here, the aim was to describe the main results of the statistical analysis of temperature changes in two morphologically different regions of Italy (Lombardy and Campania) using variants of correspondence analysis. This analysis has shown that the average temperature increase recorded in Campania and Lombardy is in agreement with the findings discussed by Tomozeiu et al. [11]. By observing plot b of Figure 4, it shows that the variability of the warmer temperature values (class (25.1, 28.8]) is lower in 2015 than in 1986 (plot a of Figure 4) being the polynomial coordinate along the second axis lower. On the other hand, plot b of Figure 3 shows that the variability of the colder temperature values (class (−10.9, −1.6]) is higher in 2015 than in 1986 (plot a of Figure 3). For Campania, these findings show a shift in the temperature distribution in agreement with previous studies on temperature extremes in Italy [12].

The application of correspondence analysis to the data examined in this paper suggests that a strong association exists between temperature class and the month of the year and gaining a deeper understanding of this association can be useful to policy makers for addressing suitable climate and agricultural policies. Indeed, by better understanding of the increasing climatic changes being observed around the world, not just in Italy, using visual strategies such as correspondence analysis (as we did above) can better prepare the agricultural industry. For example, such strategies can help support farmers with the adaptation of new farming practices (such as providing ad hoc insurance mechanisms or tools for fighting plant diseases).

Author Contributions: A.C. with G.R., M.F. and F.M.P. prepared the data and contributed to the introduction section and to illustrate the results of the applications. R.L. and E.J.B. wrote the statistical methodological and applicative sections and made the simulation study. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: No applicable.

Informed Consent Statement: The authors consent to the publication of the findings of their research.

Data Availability Statement: The authors agree to make available the data of this research.

Conflicts of Interest: The authors declares no conflict of interest.

References

1. Dong, B.; Sutton, R. Dominant role of greenhouse-gas forcing in the recovery of Sahel rainfall. *Nat. Clim. Chang.* **2015**, *5*, 757–760. [CrossRef]
2. Kisrchvink, J.; Kopp, R. Palaeoproterozoic ice houses and the evolution of oxygen mediating enzymes: The case for a late origin of photosystem II. *Philos. Trans. R. Soc. B* **2008** *363*, 2755–2765.
3. IPPC14. *Intergovernmental Panel on Climate Change. Fifth Assessment Report: Climate Change 2007*; Cambridge University Press: Cambridge, UK, 2014. Available online http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml (accessed on 31 December 2020).
4. Christidis, N.; Donaldson, G.C.; Stott, P.A. Causes for the recent changes in cold- and heat-related mortality in England and Wales. *Clim. Chang.* **2010**, *102*, 539–553. [CrossRef]
5. Hansen, G.; Stone, D. Assessing the observed impact of anthropogenic climate change. *Nat. Clim. Chang.* **2016**, *6*, 532–537. [CrossRef]
6. Stone, D.; Auffhammer, M.; Carey, M.; Hansen, G.; Huggel, C.; Cramer, C.; Yohe, G. The challenge to detect and attribute effects of climate change on human and natural systems. *Change* **2013**, *121*, 381–395. [CrossRef]
7. Rosenzweig, C.; Karoly, D.; Vicarelli, M.; Neofotis, P.; Wu, Q.; Casassa, G.; Imeson, A. Attributing physical and biological impacts to anthropogenic climate change. *Nature* **2008**, *453*, 353–357. [CrossRef]
8. Bates, B.C.; Kundzewicz, Z.W.; Wu, S.; Palutikof, J.P. *IPCC Technical Paper VI*; IPCC Secretariat: Geneva, Switzerland, 2008.
9. Toreti, A.; Desiato, F. Temperature trend over Italy from 1961 to 2004. *Theor. Appl. Climatol.* **2008**, *91*, 51–58. [CrossRef]
10. Toreti, A.; Desiato, F. Changes in temperature extremes over Italy in the last 44 years. *Int. J. Climatol.* **2008**, *28*, 733–745. [CrossRef]
11. Tomozeiu, R.; Pasqui, M.; Quaresima, S. Future changes of air temperature over Italian agricultural areas: A statistical downscaling technique applied to 2021–2050 and 2071–2100 periods. *Meteorol. Atmos. Phys.* **2018**, *130*, 543–563. [CrossRef]
12. Simolo, C.; Brunetti, M.; Maugeri, M.; Nanni, T.; Speranza, A. Understanding climate change–induced variations in daily temperature distributions over Italy. *J. Geophisical Res.* **2010**, *115*, 1–12. [CrossRef]
13. Brunetti, M.; Maugeri, M.; Monti, F.; Nanni, T. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. *Int. J. Climatol.* **2006**, *26*, 345–381. [CrossRef]
14. Brunetti, M.; Maugeri, M.; Nanni, T.; Simolo, C.; Spinoni, J. High-resolution temperature climatology for Italy: Interpolation method intercomparison. *Int. J. Climatol.* **2014**, *34*, 1278–1296. [CrossRef]
15. Mudelsee, M. Trend analysis of climate time series: A review of methods. *Earth-Sci. Rev.* **2019**, *190*, 310–322. [CrossRef]
16. Ali, F.; Dissanayake, D.; Bell, M.; Farrow, M. Investigating car users' attitudes to climate change using multiple correspondence analysis. *J. Transp. Geogr.* **2018**, *72* 237–247. [CrossRef]
17. Brunette, M.; Bourke, R.; Hanewinkel, M.; Yousefpour, R. Adaptation to climate change in forestry: A Multiple Correspondence Analysis (MCA). *Forests* **2018**, *9*, 1–14. [CrossRef]
18. Hill, M.O.; Gauch, H.G. Detrended correspondence analysis: An improved ordination technique. *Vegetatio* **1980**, *42*, 47–58. [CrossRef]
19. Correa-Metrio, A.; Dechnik, Y.; Lozano-García, S.; Caballero, M. Detrended correspondence analysis: A useful tool to quantify ecological changes from fossil data sets. *Boletín De La Soc. Geológica Mex.* **2014**, *66*, 135–143. [CrossRef]
20. Lombardo, R.; Beh, E.J. Simple and multiple correspondence analysis using orthogonal polynomials. *J. Appl. Stat.* **2010**, *37*, 2101–2116. [CrossRef]
21. Lombardo, R.; Meulman, J.J. Multiple correspondence analysis via polynomial transformations of ordered categorical variables. *J. Classif.* **2010**, *27*, 191– 216. [CrossRef]
22. Beh, E.J.; Lombardo, R. *Correspondence Analysis, Theory, Practice and New Strategies*; Wiley: Chichester, UK, 2014.
23. Lombardo, R.; Beh, E.J.; Kroonenberg, P.M. Modelling trends in ordered correspondence analysis using orthogonal polynomials. *Psychometrika* **2016**, *81*, 325–349. [CrossRef] [PubMed]
24. Chihara, T.S. *An Introduction to Orthogonal Polynomials, Mathematics and Its Applications*; Gordon and Breach Science Publishers: New York, NY, USA, 1978.
25. Favard, J. Sur les polynomes de Tchebicheff. *Comptes Rendus De L'Académie Des Sci.* **1935**, *200*, 2052–2053.

26. Emerson, P. L. Numerical construction of orthogonal polynomials from general recurrence formula. *Biometrics* **1968**, *24*, 696–701. [CrossRef]
27. Rayner, J.C.W.; Thas, O.; De Boeck, B. A Generalized Emerson Recurrence Relation. *Aust. N. Z. J. Stat.* **2008**, *50*, 235–240. [CrossRef]
28. Greenacre, M.J. *Theory and Application of Correspondence Analysis*; Academic Press: London, UK, 1984.
29. Lebart, L.; Morineau, A.; Warwick, K.M. *Multivariate Descriptive Statistical Analysis*; Wiley: New York, NY, USA, 1984.
30. Beh, E.J. Simple Corresp. Anal. Ordinal Cross-Classif. Using Orthogonal Polynomials. *Biom. J.* **1997**, *39*, 1–25. [CrossRef]
31. Benzecri, J.P. *L'Analyse des Correspondances*; Dunod: Paris, France, 1973.
32. Gifi, A. *Non-Linear Multivariate Analysis*; Wiley: Chichester, UK, 1990.
33. Eckart, C.; Young, G. The approximation of one matrix by another of lower rank. *Psychometrika* **1936**, *1*, 211–218. [CrossRef]
34. Rayner, J.C.W.; Best, D.J. Smooth extensions of Pearson's product moment correlation and Spearman's rho. *Stat. Probab. Lett.* **1996**, *30*, 171–177. [CrossRef]
35. Rayner, J.C.W.; Beh, E.J. Towards a better understanding of correlation. *Stat. Neerl.* **2009**, *63*, 324–333. [CrossRef]
36. Lombardo, R.; Beh, E.J. Variants of Correspondence Analysis. *R J.* **2016**, *8*, 167–184. [CrossRef]