Climate and population: risk exposure to precipitation concentration in mainland Spain (1950–2010)

Clima y población: exposición al riesgo por concentración de la precipitación en la España peninsular (1950–2010)

Roberto Serrano Notivoli
rsnotivoli@gmail.com
Estación Experimental de Aula Dei
Consejo Superior de Investigaciones Científicas (EEAD–CSIC) (Spain)

María Zúñiga-Antón
mz@unizar.es
Departamento de Geografía y Ordenación del Territorio
University of Zaragoza (Spain)

Ángel Pueyo
apueyo@unizar.es
Departamento de Geografía y Ordenación del Territorio
University of Zaragoza (Spain)


Received: 05.04.2020  Accepted: 24.07.2020  Published: 03.09.2020

Published under the terms and conditions of an Attribution-NonCommercial 4.0 International license.
Abstract

The term risk exposure combines the notions of natural hazard and the human exposure to that hazard. Spatial and temporal variations in risk exposure, therefore, can be caused by changes in hazard, in exposure, or both. In this work a novel methodology for computing and representing risk exposure and its temporal changes are presented, and applied to the analysis of risk exposure to extreme rainfall in mainland Spain between 1950 and 2010. For that, two complementary high-resolution gridded datasets, one of population potentials and another one of precipitation concentration, are combined. Despite a great spatial variability over time, the highest exposure was consistently found in the surrounding areas of the largest cities and along the Mediterranean coast. The relative influence of population changes and precipitation concentration evolution is analyzed. Results suggest that hazard (precipitation concentration) changes led most of the observed changes in risk exposure, except in those decades where population movements were massive and widespread.

Key words: risk; population; climate; precipitation; Spain.

Resumen

El término exposición al riesgo combina las nociones de peligro natural y exposición humana a ese peligro. Las variaciones espaciales y temporales en esa exposición al riesgo, por tanto,
pueden ser causadas por cambios en el peligro, la exposición, o en ambos. En este trabajo, se presenta una novedosa metodología para el cálculo y representación de la exposición al riesgo y sus cambios a lo largo del tiempo, aplicada al análisis de la exposición al riesgo de la precipitación extrema en la España peninsular entre 1950 y 2010. Para ello, se combinan dos bases de datos complementarias de alta resolución, una de potenciales de población y otra de concentración de la precipitación. A pesar de la gran variabilidad espacial a lo largo del tiempo, la exposición más alta se encontró de manera consistente en el entorno de las grandes ciudades y en la costa mediterránea. Se analizó la influencia relativa, en el tiempo, de los cambios en la población y de la concentración de la precipitación. Los resultados sugieren que los cambios en el peligro (concentración de la precipitación) son responsables de la mayoría de los cambios observados en la exposición al riesgo, excepto en aquellas décadas donde los movimientos de la población fueron masivos y generalizados.

**Palabras clave:** riesgo; población; clima; precipitación; España.

1 Introduction

While most of the attention in the recent decades has focused on the evolution of climatic extremes under the influence of climate change, the socioeconomic exposure to these extremes has also experienced widespread and even more dramatic changes. The present population of 7.8 billion people around the world is projected to increase to 9.7 billion by 2050 (United Nations, 2019), and urban areas are expected to absorb almost all this growth, since 68% of the population will live in cities (United Nations, 2018). This situation creates diverse urban-rural scenarios where highly-populated and densely urbanized areas will become more vulnerable to climate change consequences, increasing the exposure to the consequences of climate change through increasingly high exposure to extreme events while densely populated rural areas would experience a comparatively lower change in risk exposure (Zhang et al., 2018). In any case, the natural climatic variability includes the occurrence of extreme events that might be more frequent due to climate change depending on the territory, and populated areas are obligated to face with these scenarios and, based on them, learn about better ways to manage and plan urban areas development. A detailed knowledge about the recent behavior of climate and its spatial and temporal variability is of paramount importance from a social and economic outlook and from an applied perspective of natural systems management. In light of this, the thorough characterization of how the frequency and intensity of the extreme climatic events can affect the territorial decision-making, is a priority action.
Risk is usually defined as a combination of three components: hazard, exposure and vulnerability. Hazard refers to the occurrence of natural events in a magnitude, intensity or frequency out of the regular or expected values, while exposure is the total amount of elements (people, infrastructures, assets, etc.) in contact with that hazard. Vulnerability is a more complex term that involves the damaged caused to the exposed elements and its cost and it has been treated with different approaches (e.g.: Downing et al., 2001; Kumpulainen, 2006; Fuchs et al., 2012; Birkman et al., 2013; Olcina et al., 2016), but all of them require data or information about the economic and/or social losses related to specific disasters that are not always available. Therefore, the term risk represents the consequences of the events (hazard) over the human settlements (exposure) that are compounded in those areas with higher vulnerability. Here we use the term risk exposure to refer to the combination of hazard and exposure, that is the raw combination of the natural hazard and the human exposure, without considering the potential damages or its costs.

The Western Mediterranean region is expected to be one of the areas where the effects of climate change will be more pronounced and will result in will become in high-intensity impacts over socioeconomic activities (Ruelland et al., 2015; Nguyen et al., 2016). Although Spain holds a wide range of climatic settings, a high temporal concentration and spatial variability of precipitation is a general characteristic, especially in the eastern half of the country (Serrano-Notivoli et al., 2017a). This, together with a large great seasonal variability and a high frequency of extreme precipitation events (Merino et al., 2015), produce potential high-risk areas. In addition, in the second half of the twentieth century we have witnessed an unprecedented urbanization process has taken place in Spain, inducing an increase of risk exposure, which is translated into higher vulnerability. The Mediterranean coast of Spain is the most representative area of this situation by not only because recent urbanization tended in many cases to occupy flood-prone lands but also due to alterations of the hydrological cycles through soil sealing (Ribas et al., 2020). Despite the well-known effects of extreme rainfall events, there is a lack of an analysis at the national level identifying the most exposed areas and their changes in time.

This article presents, as a potential application for territorial planning, an analysis of changes in risk exposure to extreme precipitation over mainland Spain. As a secondary objective, we assess the relative importance of changes in extreme precipitation and population density in the observed evolution of exposure risk. We present the results in a well-readable cartographic way in order for it to be useful for a wide audience. The article is organized as follows: Section 2
describes the datasets and how we combined them; Section 3 briefly explains the results that are discussed in Section 4. The research is summarized in Section 5.

2 Data and methods

High-resolution in the spatial dimension helps to detect territorial processes that are not able to be perceived at coarser scales. However, the creation of this kind of datasets requires huge amounts of data. In the recent decades, the automation of data gathering through a wide and diverse set of techniques has greatly improved the availability of information, especially in climatic variables but also in demographic ones. In addition to the official data providers, new initiatives as citizen science are presented as helpful solutions but still need to be explored as scientific data (Theobald et al., 2015; Kullenberg & Kasperowski, 2016; Burgess et al., 2017). In any case, first processes of quality control and filtering are always necessary to avoid important data loss and wrong data that could mask confusing territorial processes. Ideally, homogenous information of different variables over the same spatial configurations at national (or lower) scales would represent a great benefit to perform joint analyses of different processes in territory but, unfortunately, it is not yet available. This, in fact, is not a problem of lack of data but a lack of will of organizing the available information around a common, precise and comparable spatial structure.

In this work, we used two high-resolution datasets (5x5 km) over the same regular matrix configuration, which greatly facilitated the analysis. Although the sources of data are different, the precipitation dataset was initially conceived over the base of the population spatial basis, thinking about future potential analyses. We focused the study on the confluence of two indices (population potentials and precipitation concentration), but both datasets include raw data with unlimited analysis possibilities. Furthermore, both of them are scalable, providing dynamical analysis capabilities that even allow for predictive analyses of risk patterns that need to be explored.

2.1 The SPREAD dataset

In Spain, the Climatic Data National Bank (BNDC) of the Spanish Meteorological Agency (AEMET) and other datasets from official national and regional organisms, contain discontinued information about the daily precipitation in more than 10,000 observatories, from which only 0.14% of them are complete in more than 99% of the days in 1950–2012 period. This reality, that is not very different from other countries, reveals an incomplete dataset that is usable in two situations: 1)
rejecting the shorter series; or 2) using all the available data to reconstruct as much as possible of the precipitation evolution. While the first option is commonly adopted to avoid further consistency problems, the second one implies a statistical approach based on the final reliability of the data series. Almost all the available precipitation reconstructions in Spain are at a monthly scale and they allow neither the understanding of the extreme events dynamics nor their related risks. In addition, these reconstructions are based on a limited number of observatories with long series as those used to the development of the Climate Research Unit (CRU) dataset (35 observatories at a monthly scale, 1900–2015); European Climate Assessment and Dataset (ECA&D) (199 observatories at a daily scale, 1896–2019); Spain02 (Herrera et al., 2012, 2756 observatories at a daily scale, 1950–2003), MOPREDAS (González-Hidalgo et al. 2011, 2670 observatories at a monthly scale, 1946–2005); González-Rouco et al. (2001) (100 approximately observatories at a monthly scale, 1899–1989); Ninyerola et al. (2007) (1999 observatories at a monthly scale, 1950–1999), amongst others.

Serrano-Notivoli (2017) developed a novel method for daily precipitation reconstruction maximizing the use of the available information and applying a daily climatic reconstruction at a proper spatial resolution to be useful to study the local variations of precipitation. This methodological challenge allows for a detailed description of the spatial behavior of the climate as well as a comprehensive and precise analysis of the climatic indices at a daily scale. Both spatial and temporal high resolutions are essential to address the study of the extreme events and the evaluation of their associated risks.

a) Methodological approach

The methodological process provides a workflow to reconstruct daily precipitation by filtering and completing original data series and create new continuous series for any location of the territory. The method aims to preserve the high-frequency variability of precipitation considering both temporal and spatial dimensions, which also allows using all the available information regardless of the length of the original series. The approach is simple, effective and clear in its development, being able to: i) independently estimate precipitation values for each day and location without assuming stationary relationships of precipitations in space or time; ii) capture the high spatial and temporal variability of daily precipitation, providing an uncertainty value to each one of the predicted values, reporting the reliability of each predicted value; and ii) allow that a researcher of any discipline is able to create an individual daily precipitation series for a single location.
The reconstruction is based on the calculation, independently for each day and location, of reference values (RVs) relying on two predicted values: a binomial prediction (BP) depicting the probability of occurrence of a wet ($P(X > 0)$) or dry ($P(X = 0)$) day and a magnitude prediction (MP) estimating the amount of precipitation ($P(X = x)$). Multiple logistic regressions (MLR) are used to compute the RVs using the precipitation data (occurrence and magnitude) of the 10 nearest neighbours as the dependent variable, while the geographic data the observatories (latitude, longitude and altitude) are used as the independent variables. Since the availability of observations vary from day to day, the models are accordingly different, reflecting the daily influence of the geographical parameters. For each day and location, an individual RV is computed based on the BP and the MP. The first one uses MLR based on the observations codified as 0 or 1 to predict the probability (from 0 to 1) of precipitation occurrence. The final RV is determined by combining MP and BP using a threshold of BP $\geq 0.5$ to set a wet day (otherwise RV is 0). The most important difference between this method and others used in climatic reconstruction is that RVs are independently computed instead of the use of reference series, that require a minimum temporal length of nearest observations. RVs include the estimation of the standard error of the model, which can be used to evaluate the uncertainty of predictions related to each moment and location and to propagate the uncertainty to further calculations such as monthly or annual aggregates. Guidelines for RVs computing and treatment have been widely described in Serrano-Notivoli et al. (2017b).

The calculation of the RVs allows for: 1) applying an independent quality control process (QC) for each precipitation observation; 2) estimating new precipitation values in days without record (or removed by the QC); 3) create new data series at not-gauged locations; and 4) replicate this new-series creation at regularly-separated locations to build daily precipitation gridded datasets.

The QC was developed to detect and remove suspect data by comparing daily observations at each observatory with the estimates (RVs) computed from their ten nearest records. Five criteria are defined to identify and remove suspect data from the original dataset: (QC1) suspect data (isolated wet conditions): observed value is over zero and the precipitation at the 10 nearest observations are zero; (QC2) suspect zero (isolated dry conditions): observed value is zero and all its 10 nearest observations are over zero; (QC3) suspect outlier: the magnitude of the observed value is 10 times higher or lower than that predicted by its 10 nearest records; (QC4) suspect dry day: observed value is zero, wet probability is over 99 %, and predicted magnitude is over 5 mm; and (QC5) suspect wet day: observed value is over 5 mm, dry probability is over 99 %, and predicted magnitude is under 0.1 mm. After this process, new RVs are computed.
using the filtered dataset to fill the gaps in the original data series. Since the observed values do not participate in the calculation of the RVs for those observations, they can be directly compared, representing a leave-one-out cross-validation process, which is useful as a validation of the quality of the reconstruction. The RVs at days and locations without observations fill the missing observations in the filtered dataset, constituting serially-continuous data series. Once the QC and gap filling processes are completed, the reconstructed series can be used to estimate new RVs at single and not-gauged locations. Also, it is useful for computing gridded datasets by applying the same approach to a set of regularly distributed pairs of coordinates.

To facilitate the use of the described processes, an R language package of functions called reddPrec was developed. This package is freely available at the official R repository (http://cran.r-project.org/web/packages/reddPrec) and the details of its operation are fully explained in Serrano-Notivoli et al. (2017c).

b) A high-resolution dataset for Spain and derived products

The methodology was applied, through the reddPrec package, to the complete daily precipitation network in Spain including 12,898 observatories in the 1950–2012 period. Most of the data sourced from the AEMET, but data from regional hydrological and meteorological services and from the national agronomic network were also used. The QC process flagged and removed an annual average of 2.4% of data in peninsular Spain with no major differences between years. Suspect data and suspect zeros were the two main reasons for data removing and the rate of outlier detection was relatively low in all years. Suspect dry and wet days were more frequent in the first 20 years of the series. The validation of the methodological protocol showed a good agreement between observations and estimates as well as in the monthly, seasonal and annual aggregates, with high correlation values both in daily means and by stations.

The missing values in the original (filtered) dataset were filled with the RVs and those reconstructed series with more than 10 years of original data (7604 series) were used to create 5x5 km spatial resolution gridded dataset over peninsular Spain and the Balearic and Canary Islands. This high-resolution daily gridded dataset was called SPREAD (Spanish PREcipitation At Daily scale) and further details of its configuration can be found in Serrano-Notivoli et al. (2017d).
The dataset is freely available for any user in raw format\textsuperscript{2} at and as a climatic service (data on demand).\textsuperscript{3}

Despite the daily precipitation values were estimated independently for each day and location, the reconstruction method showed spatially-coherent results. The application to the Spanish network confirmed that the methodology properly adapts to very different climatic situations or rainfall regimes, being able to use all the available information, avoiding the assumption of pre-constructed relationships between data series and maintaining the local variability of daily precipitation.

Based on the daily gridded dataset, a suite of traditional climatologies (monthly, seasonal and annual aggregates), mean typical values of daily precipitation (mean intensity, number of wet days, mean length of wet and dry spells), and 9 extreme precipitation indices based on those proposed by the WMO Expert Team on Climate Change Detection and Indices (ETCCDI). The spatial distribution of their average values, uncertainty and trends, were analyzed and widely described in several publications (Serrano-Notivoli, 2017 and Serrano-Notivoli et al., 2017a; 2017d; 2018). These comprehensive analyses provided new insights about the regional differences in the spatial distribution of extreme precipitation indices in Spain, mainly due to the increase of the spatial resolution in regard of the previous works. The use of all the available precipitation observations and the calculation of the relative error of predictions for the whole dataset, were very useful to show that the uncertainty (representing the difficult to predict daily precipitation) was inversely proportional to the amount of precipitation. In regard of the rainfall regimes, the high-resolution helped to reveal the Central Range (more specifically Sierra de Gredos) as the area with highest intensity in rainy days across the country. In this sense, the study of trends (Serrano-Notivoli et al., 2018) showed a general decrease of precipitation intensity in Spain both in average and extreme terms with a slightly increase of the duration of the precipitation episodes and in the frequency of low precipitation events. Conversely, the Mediterranean coast showed the inverse behavior, a positive significant trend in the number of rainy days, but also a clear negative trend in the intensity of precipitation events.

\begin{itemize}
\item \textsuperscript{2} Available at http://dx.doi.org/10.20350/digitalCSIC/7393
\item \textsuperscript{3} Available at http://spread.csic.es
\end{itemize}
2.2 Precipitation concentration in Spain

In order to detect and assess these zones, a detailed and reliable precipitation dataset is essential. As we have already discussed, Serrano-Notivoli (2017) developed a novel method for daily precipitation reconstruction which is able to translate the raw information from a dense network of individual observatories into a quality-controlled and serially-complete data series at grid points of high spatial resolution. The actual importance of its capabilities lies in the potential of creating reliable climatic data series for any location and temporal resolution (from daily to coarser time steps), which can be combined with any other available variable at the same locations. This methodological process was coded and published as a freely available R package (Serrano-Notivoli et al., 2017b) and applied to the whole Spanish precipitation monitoring network (more than 12,000 stations), to create a 5x5 km gridded dataset of daily precipitation (SPREAD, Serrano-Notivoli et al. (2017c), over the same regular set of coordinates as the population potentials used in this study.

The CI (Concentration Index) (Martín-Vide, 2004) is a useful index for exploring the risks related to extreme precipitation events (Serrano-Notivoli et al., 2017a) since it relates the magnitude to the time period in which they occur. It is a measure of rainfall intensity over specific locations and it is comparable when applied in different climatic areas (e.g.: De Luis et al., 1996; Cortesi et al., 2012; Monjo & Martín-Vide, 2016; Nuñez-Gonzalez, 2019). The CI calculation, very similar to the Gini index, begins by sorting the daily precipitation values in ascending order and then, the cumulative percentage of days where precipitation is higher than zero is plotted against the corresponding cumulative percentage of rainfall amounts. These curves are of the kind:

\[ Y = aX \exp(bX) \]  

(1)

where \( a \) and \( b \) are constants. They can be calculated by means of the least-squares method. The definite integral of the exponential curve between 0 and 100 is the area (\( A \)) under the curve, and 5000-A is the area compressed by the curve, \( Y = X \), and \( X = 100 \), lets name \( S \). Then CI is defined as:

\[ CI = S / 5000 \]  

(2)

The CI measures the relative separation of the exponential curve from the \( Y = X \) line.

Serrano-Notivoli et al. (2017a) applied it to the SPREAD dataset providing, for the first time, a comprehensive description of the mean values, variability, extremes, and trends of CI in Spain. Given the reliable representation in SPREAD of the local precipitation values in terms of
frequency of wet and dry days and extreme values, it represents a proper approach to assess the precipitation concentration evolution as described by the CI.

The CI is presented as a great index to evaluate daily precipitation risks, especially in Spain, a diverse territory in regard of its precipitation spatial and temporal distribution, regimes, trends and frequency of extreme events, amongst others. A change in precipitation concentration from the mid-20th century also means a change in a varied range of environmental risks such as erosivity, floods and extreme events. These changes in the frequency and intensity of the hazard can be complemented with changes in exposure (e.g. changes in amounts and concentration of population) and vulnerability (e.g. occupation of flood-prone areas), leading to an increase of the risk for society. The results obtained from the application of CI to Spain represent a useful tool for regional planning and for decision-making policies related to these environmental risks.

The global spatial distribution showed that highest values are present along the coasts with mountain ranges that are parallel to them. When this mountain ranges are disrupted, for instance in the Ebro or Guadalquivir valleys (no physical obstacles exist between the coast and inland areas), the CI progresses inland. In this regard, the Central Range, was an exception due to it is perpendicularly orientated to the Atlantic coast. Its high values respond to the absence of high mountain barriers in the western Iberian Peninsula (IP), Atlantic wet air masses break into the IP without obstructions. Through a comprehensive analysis, four different geographical regions can be noted in the spatial distribution of CI in Spain:

1. The Mediterranean coast, showing the highest values of the index and a high inter-annual variability, with very extreme values in some years and moderate rates as average. Torrential rainfall events are relatively frequent producing that, in combination with a high concentration of population in the summer season, it is configured as a high-risk area;

2. the southwestern IP, obtained low to moderate CI and also a high inter-annual variability. Very low precipitation concentration values (CI < 0.4) were found in some years and very high in some others (CI > 0.7). This great variability depicts a highly unstable scenario where the average situation is not representative of a high environmental risk, but this can change unexpectedly;

3. the Pyrenees and the Ebro Valley. This represents the transition sector from a) the Mediterranean area, with a high frequency of convective events in summer-to-fall time and its torrential subsequent precipitation amounts, contributing to high CI values, to b) the Atlantic influence, with a lower variability due to the influence of the Atlantic fronts and cut-off lows in
winter. The whole area constitutes a corridor between the two zones for the Mediterranean precipitation. Main consequences of the precipitation concentration in this northern sector are the risks related to a water scarcity scenario in the Ebro valley and flood-related events in the Pyrenees (resulting from short-duration and high-intensity precipitation events) that drain into the low lands of the Ebro Valley, leading to great economic losses every year;

4. the northern and southern plateaus, which are the inner and elevated areas of the IP, flanked by mountain barriers avoiding the entry of high-precipitation events, are characterized by the lowest mean precipitation concentration and a low variability, meaning that CI values differ over a short range. The lower frequency of extreme precipitation events configures the area as a low level of risk occurrence.

The comparison of the reached values with those in the rest of the world in relative terms showed that the maximum concentration in Spain (in the eastern IP) are comparable with the maximums in the Middle West, southern South America, or eastern continental Asia.

The study of trends showed a general significant increase of CI, that was more intense in those areas with a higher inter-annual variability. Also, the presence of mountain barriers near to the coast were important to explain the spatial distribution of the highest concentration values. CI trends over the 1950-2012 period indicated a global increase in precipitation concentration in Spain. By regions, the Mediterranean coast showed a generalized positive trend, meaning an increase in concentration and an increase in associated risks. The southern IP were also positive and very high, showing a pattern matching with the CI distribution, following the direction of the Guadalquivir River. This area, where the arrival of the Atlantic fronts produces (not very high) CI values, shows the need to be monitored in the following years due to CI has been increasing since mid-20th century. The Ebro Valley and the Pyrenees showed also positive trends but slightly lower, probably due to the already high values of CI and the more regular inter-annual variability of precipitation. The northern plateau showed only a small part with significant and low magnitude positive trends, even a few isolated areas showed negative trends. Southern plateau showed a general positive trend, especially in its western part, engaging the Central Range and Guadalquivir Valley high positive trends.

The CI computed over the SPREAD dataset helped to better understand the spatial and temporal behavior of this hazard, and resulted in a grid that perfectly fits with the population potentials developed by Pueyo et al. (2013), thereby facilitating a joint analysis to assess the potential risk exposure of population to precipitation hazards.
2.3 Population potentials

Instead of using a static representation of traditional demographic variables by showing population distribution maps in absolute numbers or densities, we used a representation through a potential model, representing the intensity of the possible interaction between social and economic groups at different locations. More specifically, the population potential is a measure of the influence of people at a distance, or a simplification about how close to each other are the population centers. Population potential is also useful to summarize the spatial distribution of the population over a territory (Craig, 1987). The concept of potential allows for a geographic explanation (1) of the spatial interactions between the demographic settlements and (2) of the weight of certain areas regardless of the structure of the settlements. This formulation offers a representation of the influence of the population in the territory, as well as of its spatial structures, which are of great interest to the study of the environmental alterations in the last century both in local and regional or general areas. Population potentials are based on the Newton’s law of Universal Gravitation to measure the intensity of interaction between elements of a geographical nature. This was applied from mid-18th century to the study of emotions, people and cities (Berkeley, 1713; Algarotti, 1737; Lagrange, 1773; Carlyle, 1837). Carey (1858) introduced the concept of social gravitation, applying this law to the population analysis with the consideration of that the influence of a population is proportional to the number of inhabitants and that influence decreases with distance. Reilly (1937), Stewart (1947), Converse (1949) and Zipf (1949) developed and improved the concept and formulation to analyze the interactions in a wide range of social phenomena. Gravitational models assume that the greater the population centers and the lower distance between them, the greater are the reciprocal inferences of potential.

This type of thematic cartography considers population as a continuous phenomenon in the space, which have, at least, two benefits in terms of patterns analysis: 1) facilitates the exploration of spatial trends and clustering; and 2) allows for a direct comparison with any other continuous variable over the same spatial domain.

Population potential maps take into consideration the distance from each grid point to the population centers and assess their impact over the surrounding space. This consideration is not new since Carey (1858) already stated that relationship. From the mathematical point of view, this type of gravitational models allows for the transformation from a discrete distribution of point-based clusters over space into a topological continuous surface, which is useful for spatial interactions and for the human resources management in territorial planning. The concept...
potential enables the engagement between population and distance, providing a geographical explanation to the spatial interactions between population centers. Population potentials help to understand the distribution and organization of the population, the location and delimitation of influence areas, and the inclusion in the same map of different spatial scales. It is a logical system of population modelling since the pressure on the territory is carried out in the first place by the resident population itself, but it also depends on the possibilities of access that the rest of the citizens have to it. This is a consequence of distance from the rest of the population centers. Although the sense of proximity is different in each period of time and spatial situation, which would require the use of specific adjustments, in this case a homogeneous model has been used for the whole of the Iberian Peninsula (Zúñiga, 2009). All of this contributes to a better delimitation of metropolitan, suburban, rural or empty areas, independently from the administrative units to which they belong (López Escolano, 2017).

Compared, for instance, to the traditional population density cartographic representation, it only shows the resident population, while potential maps show areas that, due to their location in relation to the others, are able to explain territorial movements. This is helpful to understand mobility, especially around the main cities, related to rental prices, development of economic activities, infrastructures and services, amongst others.

a) Methodological approach

The 5x5 km spatial resolution grid was chosen based on the size and density variation between municipalities in Spain (more than 8,000 municipalities). Each of the grid cells assumed the total population of all the population centers they covered. As grids spanned 25 km², the maximum distance between two municipal seats in the same grid area was 7 km, which gives the assumption of that municipal centers 7 km far from each other are considered as the same demographic unit (Calvo et al., 2008).

This methodology for population potentials calculation in Spain uses a variation of the gravitational model that has been used during three decades by researchers from the University of Zaragoza (Spain) in several demographic studies (Calvo & Pueyo, 1989; 1991; Calvo et al., 1992; 2007; 2008; 2009; Pueyo et al., 2012; 2013; 2016). The model applies the already mentioned gravitational model by putting the information in a matrix-structured dataset (raster format) over the Spanish territory. Each of the grid cells are mapped according to their population amounts in each census, plus the sum of values inferred from the rest of the values of the system (the rest of
the territory), resulting from the population and distance to each of the remaining grid cells (Figure 1).

**Figure 1. Population potentials calculation process**

![Population potentials calculation process](image)

Source: elaborated by the authors

Obviously, a grid cell has only influence when it contains a main population center, according to Equation 3, which represents the mathematical adaptation of a gravity model. Thus, the calculation of population potentials for each grid cell ($POT_i$) is expressed as

$$POT_i = \sum_{j=1}^{n} \left( \frac{P_j}{d_{ij}} \right) + P_i$$

(3)

where $P_j$ is the number of inhabitants at all the grid cells in the spatial domain; $P_i$ the number of inhabitants in grid cell $i$ and $d_{ij}$ the distance between grid cells $i$ and $j$.

Population potentials in Spain are available from 1950 to 2010 based on the national census data developed by the Spanish Statistical Office for the following years (1950, 1960, 1970, 1981, 1991, 2001 and 2010). It was necessary to homogenize the population series to calculate the
population potentials from the beginning of the series. From 1897 to the 1991 census, the concepts of Population in law and Population in fact were recognized (Vinuesa, 2005). However, in the 2001 census round, the concept of Population in law was replaced by Resident population and the concept of Population in fact was replaced by Connected population understood as the group of persons with habitual residence in Spain who have some type of habitual link with the municipality: they work, study or have a second home there. This methodological change is better adjusted to the territorial reality and eliminates the concept of transient, counting only the population that is registered and has a legal right to receive services in the municipality (Vinuesa, 2005).

2.4 Precipitation concentration risk exposure calculation

We used the CI annual values computed and analyzed in Serrano-Notivoli et al. (2017a) (Figure 2B) as a measure of Hazard (H). Only years with POT data were considered.

Following Vörösmarty et al. (2017), exposure (E) was assumed as the population potential of people (POT) in contact with a particular hazard level during a time step.

\[ E = POT|H \]  

(4)

In a first step, both POT and CI values at each grid point \( x_i \) were normalized from 0 to 1 to make them comparable:

\[ x_i = (x_i - \text{min}(x))/(\text{max}(x) - \text{min}(x)) \]  

(5)

where minimum (min) and maximum (max) values of POT and CI were computed taking into consideration the complete dataset (data from all years) in each case. Then, Exposure \( E_i \) was computed as the natural logarithm of the ratio between normalized POT and CI at each grid point:

\[ E_i = \log(POT_i/CI_i) \]  

(6)

2.5 Mapping risk exposure changes

The challenge of represent the combination of such different variables as precipitation concentration and population were solved using efficient ways of cartographic representation through the use of the appropriate spatial scale and communicative color schemes. A thematic map allows the reader to visualize and understand the spatial problems. To this end, the use of different representations to facilitate the reflection, consideration and construction of knowledge about the represented variable, are of crucial interest (Pueyo et al., 2016).
The use of the 5x5 km spatial resolution grid cells was of key importance in the representation of the risk exposure. From the population point of view, the 25 km² area is a value lower than the mean and median area of the Spanish municipalities (62 and 35 km² in 2006, respectively). Calvo et al. (2008) verified that the population axes (spatial directional clusters) are better represented in potential maps than in population density maps, absolute population maps, or demographic evolution maps. In addition, most of them present serious difficulties to make comparisons between population centers, while the use of homogeneous grid cells in which the potential inferred from the population center is projected, the importance for spatial planning is emphasized. Indeed, this can be accentuated with the surrounding grid cells to draw the population axes of urban expansion. From the precipitation concentration side, the spatial resolution was detailed enough to account for the local spatial variations of precipitation. However, those zones where the spatial variation of precipitation is higher, especially at mountain areas, a higher spatial resolution would provide better results in terms of precise representation of reality as done, for example, in the Pyrenees (Serrano-Notivoli et al., 2019).

We first mapped the annual exposure to detect year-to-year variations as well as the linear trend using the Theil-Sen estimator (Sen, 1968) to identify homogeneous spatial patterns. To this end, only significant trends ($p < 0.05$), computed through Mann-Kendall test (Mann, 1945) were represented in color, while non-significant areas were left white. A diverging color scheme was used (blue to red) to show the trends, getting a quick identification of the spaces with a tendency to be considered.

The combination of color hue and lightness to map the annual risk exposure required to complete a sequential color scheme to distinguish the different values within each year and make them comparable between years. The legend uses a suite of cold colors with a transition from blues representing higher values of risk exposure to a lightless yellow for lowest values. This selection involves a semiotic conception of the cartography, which means that the document expounds the reality in an objective way (Zuñiga et al., 2012).

Risk exposure maps computed in the proposed way resulted in a range of values with a difficult interpretation, thus, we opted for a representation of changes between timesteps. For each grid point, and year (except for 1950, the first one), the relative change in risk exposure was computed, and expressed as a percentage with respect to the first year. This resulted in six maps at the original spatial resolution of 5x5 km showing the changes within a range of values from -10 to +10%. Since the variable is continuous, we did not cut legend values in classes, taking
advantage of the smooth transitions in a diverging color scheme. The legend is organized around the value of zero (no changes). Positive values range a transition of low-to-high saturations of red, while negative ones use the same scheme with blue as base color. Following Zuñiga et al. (2012), the positive and negative schemes were designed to: 1) be symmetrical; 2) emphasize the extremes; and 3) use the same percentage of each color. CMYK system was used to design the color legends of all maps.

The legends on the two main cartographic documents have been designed to be applicable in all timesteps (1950–2010 and 1950/1960–2001/2010). Only one legend is incorporated in each document, which is shared by all the intermediate maps. Using this cartographic model, the capacity of representation for each specific date is lower. However, this model has several advantages that are considered fundamental to this study. On the one hand, it makes the final maps comparable with each other, which is one of the key elements that allows 1) to identify shared spatial patterns and 2) to compare the intensity of trends. In addition, it makes the maps simpler and allows to focus the attention on the mapped area.

**Figure 2.** A: Population potentials (POT) density in 2010. B: Mean annual precipitation concentration (CI) in 1950-2010 period

![Image of maps](image_url)

Source: Pueyo et al. (2013) (A) and Serrano-Notivoli et al. (2017a) (B)

3 Results

3.1 Spatial and temporal distribution of exposure

Potential population maps help to precisely delimitate the metropolitan, suburban, rural and empty areas regardless of the administrative units where they are referred. They are also useful to
detect and assess the impacts received by the high pressure of human presence, territorial mobility and activity, as well as the implications of the climatic exposures as the extreme precipitation risks. The maps make visible the distribution and influence of population and the final result of the concentration that occurred in the IP from the mid-20th century (Calvo et al., 2008). The evolution in the 6 studied decades is accompanied by productive changes and higher interdependence between activities, making evident the higher exposure to any extreme natural risk. The late 1950s’ industrialization and economic reorganization intensified the urbanization and the abandonment of rural areas started in the end of 19th century, consolidating the metropolitan areas or urban conurbations that constitute the current Spanish urban framework with the dramatic rural emptying. The demographic growth has been concentrated in the main urban centers and in a few main district areas. This is the case of great cities or metropolitan areas (Madrid, Barcelona, Bilbao, Sevilla, Valencia, Málaga, Zaragoza), other major settlements (e.g.: Palma de Mallorca, Santa Cruz de Tenerife, Las Palmas, Alicante, Logroño, San Sebastián, Burgos, Valladolid, Oviedo, Gijón Vigo, A Coruña, Huelva, Cádiz, Córdoba, Granada, Almería, Cartagena, etc.) and municipalities of metropolitan areas from the Basque Country, Barcelona, coast of Catalonia, Asturian crossroad, Valencia, Alicante, Murcia or surrounding areas of Zaragoza. The complexity of the activities and infrastructures change these areas into high-sensitive zones to extreme events risks, with high costs of recovery.

When comparing POT and CI, they showed little changes in time in absolute figures. Despite the increase of population in global terms from 1950 to 2010 (+60%), the variation by decades is almost negligible when the complete range of values is shown (Figure 3A). This is due to the great spatial diversity of POT per grid cell (standard deviation varied from 16,000 in 1950 to 31,000 in 2010, while medians were 1,273 and 1,417, respectively). Similarly, the CI showed a slightly higher variation (Figure 3B), especially in the decade of 1960 with a lower value and then, it remained in a short range between 0.54 and 0.56 until 2010. When looking at the variations of CI depending on different population categories (Figure 3C), the general pattern of CI drives the temporal evolution with internal variations. In 1950, the CI decreased as the POT increased, a pattern that gradually changed in the following decades, being the CI higher in areas with highest POT until 2010, when areas with POT>50,000 stayed with CI values lower than areas of the immediately lower class.
Figure 3. POT (A) and CI (B) variation per year and CI variation per year and by population categories (C).

Red dashed lines indicate the 1st and 99th percentile of POT and CI of all years.

The spatial distribution of the risk exposure values also does not show dramatic changes between years (Figure 4). However, a gradual shift of increasing values around the initial highest ones can be clearly seen from the beginning to the end of the study period. In regard of this, two different spatial patterns can be inferred depending on high values of POT or CI:

- **POT dependence:** Most of the high-risk exposure values are related to highly populated areas, such as Madrid, Barcelona, Sevilla, Valencia and Bilbao. Very high POT value always results in high risk exposure, which makes obvious the increase of exposure in surrounding areas of the largest cities over the time, responding to the population increase. This premise is also valid for the southern Mediterranean coast where, while cities are not extremely large, the urbanization has been greatly increased in the last decades in relation to the tourism development.

- **Precipitation concentration:** The highest values of CI are concentrated along the Mediterranean coast, in the Ebro Valley and in the Cantabric coast, resulting in high risk exposure values.
Besides, Serrano-Notivoli et al. (2017a) stressed a positive CI trend in the Guadalquivir Valley and southeastern PI, which is also noted in the temporal evolution of risk exposure in these areas in the last years (mostly since 1991) of the series.

We emphasize that the methodology computes the POT exposed to different levels of CI in individual years and it is not intended to be interpreted as a prediction of extreme precipitation occurrence. Still, linear trend was computed as a measure of the average change (expressed in percentage) per decade all over the study period. The spatial distribution of this trend (Figure 4) summarizes the main changes, which are shown as an increase of exposure (positive trend) mainly around the cities of Madrid and Barcelona, and also in southeast, southwest and central northern part of the IP. Several other smaller areas, mostly coinciding with cities of medium size, had positive values. The negative trends are spatially disseminated all over the territory with a barely perceivable concentration on northwestern part, coinciding with northern plateau. Most of the latter areas were revealed as with negative trend of CI (Serrano-Notivoli et al., 2017a).

**Figure 4. Spatial distribution of risk exposure (E) over the census years (1950–2010) and linear trend expressed as relative change per decade in percentage**

Source: elaborated by the authors

### 3.2 Relative changes of exposure

The relative differences in risk exposure between decades is shown in Figure 5. In the first period (1950–1960), risk exposure heavily decreased (near to -10%) in most of the inner eastern Spain, while the highest increases were along the southern and northern Mediterranean coast. In
the period 1960–1970, the highest changes were found surrounding Madrid and all over the Mediterranean coast. The pattern was similar in the 1970–1980 period, but more intense in the latter and almost negligible around Madrid. The 1980–1990 period changed the behavior in the Mediterranean coast to a heavy decrease in its central part while an increase occurred in the Guadalquivir valley, a pattern that was extended to the next decade (less intense in the coast). During the last period of analysis, a general increase in risk exposure was found, especially in the Mediterranean coast and around Madrid, but also in northwest Spain and along the Cantabric coast, for the first time in the whole period.

These relative changes in risk exposure mostly coincide with the urbanization and specialization stages of the different regions as suggested by Calvo et al. (2008). Until the decade of 1970, the urbanization process was related to the rural-urban migrations, while since the decade of 1980 the touristic specialization led the urbanization process in the (southern) Mediterranean coast, along with the strengthening of ancillary industries and the reinforcement of the agri-food sector, transforming this space in the market garden of Europe.

Figure 5. Relative change (expressed in percentage) of exposure between decades

![Relative change (expressed in percentage) of exposure between decades](image)

Source: elaborated by the authors
4 Discussion

Despite the wide availability of scientific literature about climatic hazard, recurrence of extreme events and forensic analysis of extreme episodes, the attention received on how these climatic factors can impact on population has been significantly lower. We are living a changing-climate scenario where these impacts could have a great influence on social and economic activities. However, several previous works made great efforts to analyze how climatic risks are considered in the decision-making processes. For instance, Jurgilevivh et al. (2017) studied 42 sub-national climate risk and vulnerability plans all over the world, finding that biophysical factors, at the expense of the socio-economic ones, still prevail in all future-oriented assessments. They stated that the usual consideration of exposure is tackled from two perspectives: 1) as a manifestation of a hazard, just mapping the physical event (e.g.: extreme rainfall, flood, etc.); and 2) as a geographical location, through the representation of the layer in which hazard impacts (e.g.: population, land use, etc.). Our approach goes beyond by merging both into a single risk exposure index as made before at different spatial scales (e.g.: Vörösmarty et al., 2017; Mysiak et al., 2018). Sherbinin et al. (2019) reviewed 84 plans mapping social vulnerability to climate impacts and they found serious lacks such as map validation or engagement with policy audience, deficiencies that we tackled by using reliable sources of information and well-readable cartographic representations, respectively. In this regard, Viner et al. (2020) emphasized the importance of facing risk as a dynamic concept, spanning the temporal context as much as possible, showing the evolution of the risk over time and, in essence, avoiding the static perspective of the risk. The use of high-resolution maps also helps in this process of developing decision-making materials tailored to individual locations, which is the final objective in applied regional (and local) planning.

We performed a first attempt to analysing the risk exposure to precipitation concentration that has a clear reading related with the spatio-temporal variability of population: The evolution of CI based on the population (Figure 3C) showed higher values in less populated areas in 1950 due to that was the usual demographic configuration: higher number of towns of small and very small size (<1,000) in a mostly rural-based society. This dynamic began to change the next year of the series (1960) with a massive rural depopulation in favor of largest cities where the new industrial clusters and tourist areas absorbed the population of small towns. In 1970 and 1980 rural population clearly decreases and this can be noted in a lower exposure in inner IP, while the census of year 1991 showed that areas with more than 5,000 inhabitants begin their growth, strengthening the new suburban areas. In 2001 this trend is not so intense and some degree of
similarity between mid-sized populated areas can be observed. In 2010, the highly-populated areas remained stable while the actual increase is in their surrounding areas. This decade (2001-2010) corresponds with the highest foreign migration to Spain (about 4 million people). Through a different way than the large cities, the Mediterranean coast, especially the southern part, also increased its risk exposure related to population. The drastic change in its territorial model related to the mass tourism trend, beginning in the decade of 1960, produced large urbanized areas and massive seasonal population migrations. These factors partly explain why extreme climatic situations became important in these areas over the last half of the study period.

Despite the fact that a great part of the changes in risk exposure can be explained by the evolution of population over time, it cannot be completely attributed to it. Serrano-Notivoli et al. (2017a) addressed that the CI also had a clear trend in Spain during the whole considered period, with some spatial variability. With the aim of ascertain that the global spatial variations do not depend only in population variations, we correlated the decadal changes of exposure (Figure 5) with the decadal changes of both POT and CI, considering all grid points. The results (Figure 6) shows that the highest correlations with POT were of 0.51 and 0.48 in the decades of 1960 and 1970, which means that the changes in risk exposure in these decades are largely due to changes in population amounts. This is consistent with the last years of rural depopulation and the beginning of urbanization in the Mediterranean coast. However, the rest of correlations are too low to be considered as a reliable explanation, always in general terms (specific patterns in smaller areas can be explained mostly with population variations). The correlation with CI, however, was higher (ranging from 0.62 to 0.83), which explains that changes in exposure in all the considered years are aligned with changes in hazard (CI). This agreement is stronger since the decade of 1980 (0.83) until 2010, coinciding with a decreasing correlation with population in the same period. These results suggest that CI variations mostly explain the changes in exposure, except in those moments when the population variations were most intense (decades of 1960 and 1970).

These variations, depending on different and unpredictable factors, make difficult a potential estimation of risk exposure in the near future, considering different scenarios in a climate change context. On one side, the small variation of annual CI values and the demand of daily values of precipitation for its calculation, makes difficult obtaining reliable values of future CI. On the other side, future population estimations (apart from the predicted natural movements) rely on the knowledge about critical socioeconomic changes, which have been shown as the main drivers of risk exposure when they take place.
Although we found a variable evolution of exposure over mainland Spain, with a mostly-increasing change in the last decade of the period (2001 to 2010), we did not take into consideration the vulnerability since it was not available as a quantifiable variable for each grid point as POT or CI. However, previous studies in Spain (Ribas et al., 2020), Europe (Paprotny et al., 2018) and global (Formetta & Feyen, 2019) indicate that changes in vulnerability are not necessarily linear with respect to changes in hazard or exposure, indicating that further research in this respect is needed. Building on this, hazard and exposure will be closely related in the near future to i) the design and territorial planning of urban areas; ii) the management models of social and economic activities; and iii) the community resilience to face extreme or unforeseen situations. The current context exemplifies the lack of foresight and the arrogance in the belief that, as society, all the natural processes are under control.

The analysis we made is of great importance to regional planning and decision-making processes about territorial development. Most of the actions focused on risk assessment in environmental plans in Spain have been dedicated to hazard mapping (e.g.: precipitation or flood probability/return periods, etc.) but only a few were focused on exposure and vulnerability involving, for instance, authorities and stakeholders (Martín et al., 2017). The latter should be analyzed from a common perspective but, at least in Spain, the information required to that aim is not centralized yet, which makes difficult a common approach. Our analysis provides an example with CI, but the methodology can be applied to any other natural hazard. The most important issue in this kind of analyses is that is necessary a correct definition of the hazard is needed, and how it can affect to the rest of the system (Hagenlocher et al., 2019). In addition, a key factor is the availability of public access to data at the proper spatial resolution for risk evaluation analyses. A better understanding of exposure and vulnerability involves the use of multiple types of
information such as activities, infrastructures and facilities exposed to specific natural hazards. Also, the disaggregation of socio-demographic data by age groups, vulnerable communities, etc. allows for a better focus of the assessment models.

5 Conclusions

We combined two high-resolution datasets (5x5 km spatial resolution) of population potentials (POT) and precipitation concentration (CI) over mainland Spain to obtain a risk exposure index. The spatial variations of this index over time in the 1950–2010 period were assessed through its relative change per decade. Despite the high spatial variability, our results suggest that variations in population explained the risk exposure variations in the decades of 1970 and 1980, while CI variations were largely responsible of risk exposure changes in the decade of 1950, and from 1981 to 2010. These changes also explain the important increase of damages caused by extreme climatic situations in the last three decades in Spain. This analysis is the first attempt to evaluate the exposure of population climatic risks and the source information of CI and population potentials will be updated to assess the risk exposure in temporal periods comparable to the existing ones (decades). We did not consider the potential social and/or economic losses due to the absence of a complete and continuous dataset over time and space in Spain. Therefore, we did not include the vulnerability in the equation of risk assessment. However, the work serves as an essay of what can be done with a reliable and consistent baseline information in regard of decision-making processes in regional planning.

Regarding the population potentials, some of the results do not take into consideration the discontinuities linked to tunnel areas caused by commutes such as flights, high speed rail network, or the new flexidimensional conceptualizations of the geographical space (López et al., 2015). Although this is not the aim of this work, a new approach using gravitational models including roads network, airports and high speed rail network could be used in future research.

Acknowledgments: R.S.N. is funded by a “Juan de la Cierva” postdoctoral grant FJCI-2017-31595. R.S.N., M.A.S and M.D.L. are supported by the Government of Aragón through the “Programme of research groups” (group H38, “Clima, Agua, Cambio Global y Sistemas Naturales”) and thank to the project CGL2015-69985-R. R.S.N. and S.B. thank to the CLICES Project (CGL2017-83866-C3-2-R).

Authorship statement: The authors declare no conflict of interest. The participation of the authors in the article is as follows. R. Serrano-Notivoli: Conceptualization, design, data collection and
execution of this study; analysis; paper writing. M. Zúñiga-Anton: Discussion and advice; paper writing, cartographic analysis. A. Pueyo: Literature review, methodology and results. S. Beguería: Data analysis; discussion and advice. M.A. Saz: Discussion and advice; Conclusions. M. de Luis: Conceptualization, discussion and conclusions.
References


Boletín de la Asociación de Geógrafos Españoles, (86)


Zhang, W., Zhou, T., Zou, L., Zhang, L., & Chen, X. (2018). Reduced exposure to extreme precipitation from 0.5 °C less warming in global land monsoon regions. *Nature Communications, 9*, 3153. [https://doi.org/10.1038/s41467-018-05633-3](https://doi.org/10.1038/s41467-018-05633-3)

Appendix S1. Spatial distribution of CI over the census years (1950-2010)

Source: prepared by the authors

Appendix S2. Relative change (expressed in percentage) of CI between decades

Source: prepared by the authors