DE GRUYTER
OPEN

Journal of Official Statistics, Vol. 34, No. 1, 2018, pp. 83–106, http://dx.doi.org/10.1515/JOS-2018-0005

Typology and Representation of Alterations in Territorial Units: A Proposal

Francisco Goerlich[1] *and Francisco Ruiz*[2]

This article proposes a typology of boundary changes in territorial units at two points in time. The different types of changes are organized in a hierarchy and represented homogeneously, independently of the number of territorial units involved and of the changes to them. Each alteration is described precisely and unambiguously, and it is codified to allow the information to be treated automatically. In addition to providing efficient storage of the information about these changes, a canonical representation facilitates the automatic detection of inconsistencies in the database. At the same time, the typology allows us to define backward and forward equivalence rules, which helps in the task of generating homogeneous time series about territorial unit characteristics, such as population or surface area, or generating the full genealogy of a territorial unit over time. We also offer an application of the proposal to inconsistencies and error detection in the database *Alterations to the Municipalities in the Population Censuses since 1842* from the Spanish National Statistical Institute (INE).

Key words: Municipal boundary changes; population census; homogeneous series; standardized representation; equivalence rules; inconsistency criteria; typology.

1. Introduction

Alterations to territorial unit boundaries occures relatively frequent, particularly in the case of smaller units – municipalities or census tracts – or when a sufficiently long time perspective is considered. Historical records of such alterations, however, tend to be literal or descriptive, with no clear standardization, and therefore very difficult to deal with on paper and practically impossible to handle in digital format. Modern advances in Geographical Information Systems (GIS) have facilitated the generation of different administrative boundary layers of territorial units – countries, regions, counties, municipalities, and so on – with different reference dates; however they do not usually provide information about the changes between two reference periods.

[1] University of Valencia and Instituto Valenciano de Investigaciones Económicas (Ivie), Ivie, Calle Guardia Civil 22 Esc. 2 1°, 46020 Valencia, Spain. Email: Francisco.J.Goerlich@uv.es.
[2] University of Castilla-La Mancha, Information Systems and Technology Institute. Dpto. Tecnologías y Sistemas de Información, Escuela Superior de Informática, Universidad de Castilla-La Mancha, Edif. Fermín Caballero (3ª planta), Paseo de la Universidad 4, 13071 Ciudad Real, Spain. Email: Francisco.RuizG@uclm.es.
Acknowledgments: Useful comments by three reviewers and an associate editor have contributed to improving the article substantially; they are not, however, responsible for any remaining errors. Francisco Goerlich acknowledges funding from the BBVA Foundation-Ivie Research Program, and project ECO2015-70632-R of the Spanish Ministry of Education and Science. Results referred to in the article are available from the authors upon request.

Various studies have compiled historical boundaries in an attempt to provide a systematic record of the alterations to them (Tir et al. 1998). More recently several authors have worked on reconstructing these boundaries using GIS techniques, which require the genealogy of current territories to be reset (Gregory 2005; Gregory and Ell 2007; Flora et al. 2015), usually with the purpose of homogenizing specific characteristics over time (Marti-Henneberg 2005; Gregory and Ell 2006).

In the process of harmonizing European regional statistics over time, considerable effort has been devoted to modeling territorial changes in the regions (NUTS 2 and 3 in European Union terms) in order to create homogenous databases on a regional scale subject to temporal evolution of the territorial hierarchy (Ben Rebah et al. 2011; Milego and Ramos 2011). Official organizations and National Statistical Institutes (NSIs) also provide compilations of changes to administrative divisions at the various territorial scales into which the state is organized. However, as far as we know, no methodological proposal has attempted to harmonize, systematize, and computerize changes in territorial units over time. The European INSPIRE directive (Directive 2007/2/EC), establishing an infrastructure for spatial information in the European Union, would probably be the most suitable framework for this harmonization process. However, the *Technical Guidelines* on *Data Specification on Statistical Units* (INSPIRE 2013) only offer a very brief guide to the temporal representation of administrative units at the object level, and tend to emphasize the modeling, for all geographical objects, considering only a life cycle defined by the attributes *beginLifespanVersion* and *endLifespanVersion*, rather than explicitly linking the different temporal versions of the same spatial object. These simple rules are clearly unsuitable for a complete characterization of a territorial hierarchy subject to spatial changes over time that can be managed in a harmonized and systematized way, taking advantage of computerization. Different NSIs have attempted to go beyond a simple compilation of territorial alterations, trying to develop spatio-temporal information systems allowing for the temporal evolution of administrative boundaries in a consistent manner (Sindoni et al. 2002; Duque 2016).

This article aims to help bridge this gap by proposing a typology and representation of alterations in administrative territorial units whose boundaries are determined by criteria of political powers or state organization. The proposal has been developed from our experience with the database of alterations to Spanish municipalities since they first appeared as such in the population censuses in the mid-nineteenth century. The proposal was therefore based on literal descriptions of changes, and our efforts have focused on developing a system to codify the alterations that is consistent and could be automated. But clearly the proposed typology and coding principles are much more general, and can be used in tracking alterations in regions, cities or urban areas, as considered, for example, in the Urban Audit pan-European project. These territories are not always consistent with administrative divisions, and the efficient monitoring of changes in borders could be very useful.

However, the principles we detail below can be applied more generally, not only at other scales such as regions or census tracts, but also based on GIS layers at two moments in time, since the geometric accuracy of the two layers will be the same in the two periods. A simple "union" GIS operation between the two layers provides all the necessary information to implement the typology proposed in this article. All that is required are the

codes for the territorial units in the two periods, together with their surface areas, and the surface area of each polygon generated by the "union" operation. Although our application below is illustrated using historical municipalities, the typology proposed in this article has been successfully applied to generate types of alterations between census tracts at two moments in time using their GIS layers, without any other additional information on alterations to them. This information is contained in the mapping itself.

The article is structured as follows. The next section presents the proposed typology in detail and the criteria for it to be computerized. We then apply the alterations in Spanish municipalities to the database going back to the first population censuses of the mid-nineteenth century. This application allows us to examine how efficient our proposal is in detecting inconsistencies. The final section offers a brief conclusion.

2. Typology of Territorial Changes: A Proposal

One initial question that must be clarified from the start is *what do we understand by a territorial unit, from the perspective of a typology?* While this may appear to be a fairly trivial question, from the point of view of a typology it must be stressed that for our purposes, the only property by which a territorial unit can be unequivocally identified is a code. A name is not usually valid information to generate a typology, since it might not be the only name, two or more territorial units may have the same name, and a name may change at any given moment. The typology can also reflect name changes, as we shall see, even though they are not territorial changes. It is true, however, that other types of geographical entities, such as census tracts, have no name and the only information of consequence is their code. NSIs and international institutions are fully aware of this need to identify territorial units using unique codes.

This clarification is important because if a territorial unit code changes, and this is the only change made to it, from the perspective of our typology it will be treated as a 'territorial' alteration: one territorial unit disappears and another identical one is created. Fortunately, the typology allows these cases to be clearly identified.

Based on the premise that each territorial unit has its own code, and drawing on our experience of municipal alterations in a historical context, we propose a typology to classify the categories of changes to territorial units that is complete in that it incorporates all existing situations, but also open as new cases can be added to it. Because the typology was created on the basis of a specific experience – historical alterations in Spanish municipalities – certain unusual cases arise; however it is clear that the underlying philosophy can easily be adapted to other similar situations. The typology includes a codification that allows for efficient treatment of the information with computer systems, and databases in particular. We start from a two-dimensional classification and aim to establish a typology that meets the following characteristics:

i) It distinguishes the cause of the change, which may affect several territorial units at the same time, from the alteration or effect that this change produces in each of the territorial units.
ii) It considers all possible types of change. Notwithstanding, it can be extended to incorporate new types of changes or situations.

iii) It includes a textual definition for each type, as well as a precise specification that indicates the 'backward' and 'forward' rules to generate a homogenous structure of territorial units according to a given criterion, for example, homogeneous population series according to the structure of a given year, or to derive the genealogy of a territorial unit.
iv) It establishes a 'canonical form' of representing changes, namely, a common format that enables all possible situations to be dealt with by creating a database for consultation.
v) It establishes criteria to detect inconsistencies.

2.1. *Double Perspective*

When analyzing the possible types of change that can affect territorial units, and the best way of representing them, it is important to distinguish two perspectives or dimensions:

a) The cause or the type of change itself: various units merged into one new unit, one unit integrated into another, etc.
b) The alteration or specific effect that this type of change has on each of the affected territorial units: elimination, creation, modification, etc.

For example, if unit A is integrated into unit B, from the first perspective we would refer to it as an 'integration' change type, whereas from the second perspective we would say two alterations had occurred: an elimination, territorial unit A disappears as it is integrated into B, and a modification, since the territory of territorial unit B increases with the integration of A. This idea of pairs of alterations associated to types of changes is crucial to our proposal.

The following distinction must always be maintained: a change (of a certain type) is reflected (manifested) in one or more alterations. The term 'alteration' and the term 'change' must always refer to these two related ideas or concepts. From the outset the meaning of the terms used must be precisely understood so that there is no ambiguity in the way they are applied.

2.2. *Hierarchy of Basic Change Types*

The first perspective provides the base for the following hierarchy of change types that, at the first level, distinguishes between territorial and nonterritorial changes as by definition the territory is the key element of all Territorial Units (TU).

TERRITORIAL:

- *Territorial units are neither created nor eliminated:*
  - Transfer (T): one TU transfers part of its territory to another TU or other TUs.
  - Exchange (P): two TUs exchange part of their territories.
- *Territorial units are created and are not eliminated:*
  - Segregation (S): one part of a TU is separated to create a new TU.
  - Partial merger (Fp): parts of two or more TUs are combined to form a new TU.
  - Unspecified appearance (O): a new TU emerges without any specific information.

- *Territorial units are eliminated and are not created:*
  - Integration (I): a TU is fully incorporated into another TU or other TUs.
  - Distribution (R): a TU disappears when its territory is distributed among two or more pre-existing TUs.
  - Unspecified disappearance (O): a TU disappears without any specific information.
- *Territorial units are created and eliminated:*
  - Code change (C): a TU's code is changed (in practice, the old TU is eliminated and a new one is created).
  - Merger (F): two or more TUs are combined to form a new one.
  - Division (D): a TU is divided into two or more new TUs.

NONTERRITORIAL:

- Change of designation (G): a TU's name or designation is changed.
- Annotation (A*x*): other changes or information that do not affect the territory. A different letter '*x*' can be used for each situation we are interested in identifying, thus allowing the typology to be extended.

It should be noted that the cases of *Unspecified appearance* and *Unspecified disappearance* mentioned above are included because they appear in the historical lists of alterations, but they are infrequent in the present period. In any case, these situations do not arise in closed territorial systems and those with well-defined administrative divisions.

We believe the terminology is simple yet precise and each term is used unambiguously. Hence, a distribution indicates that the territory of the territorial unit that disappears results in an increase in the territory of other pre-existing territorial units, whereas a division indicates the appearance of new territorial units. In both cases the original territorial unit disappears, otherwise it would be considered as a territory transfer; but what happens to the destination territorial units, of which there must be more than one otherwise we would be dealing with an integration or a code change, depends on the specific term used.

2.3. Types of Alterations

Each type of change identified in the above hierarchy gives rise to a certain alteration or effect in each territorial unit involved in that change. There are four types of alterations and, as with the change types, we distinguish between those that have territorial effects and those that do not:

- Creation (C): the TU appears.
- Elimination (E): the TU disappears.
- Modification (M): the TU's territory changes.
- Others (O): a nonterritorial characteristic of the TU changes, such as its designation.

2.4. Codification: Canonical Representation

The alteration, and the type of change that causes it, is codified using a set of two or three letters that form a descriptive key representing the specific situation of a given territorial unit when the alteration occurs. The first letter represents the alteration in the municipality.

The remaining part of the descriptive key represents the type of change that causes the alteration. The full list of the keys used in the proposed typology is presented in Table 1, and a detailed description of all possible categories of changes, including a graphical representation, is provided in the Appendix. As can be seen from Table 1, usually two letters are sufficient for the codification, but sometimes an additional letter is convenient to denote partial splits or increasing/decreasing alterations by transfer of parts of other territoties.

We have looked for a way of representing the types of change, and their associated alterations, that can be used to represent all possible situations, and therefore generate a codification that can be represented and treated automatically. This way of structuring information corresponds to the concept of 'first normal form' used in the field of databases. In our case, this has two main consequences:

a) We represent all the changes by means of the list of the alterations they generate.
b) Each alteration is defined as a relation between two, and only two, territorial units.

Thus, the three central elements of the canonical representation are the two codes for the territorial units involved, and the key that identifies the relation between them (Table 1). The first territorial unit is the one that has the alteration, represented by the first letter of the key, while the second is related to the first by the type of change.

As an illustration, let us suppose that territorial unit A disappears because it is incorporated into territorial unit B. The 'standardized' representation of this 'integration' type change is formed by the following pair of alterations:

- (A, EI, B) $\rightarrow$ A is eliminated (E) by integration (I) into B; and
- (B, MI, A) $\rightarrow$ B is modified (M) because it integrates (I) A.

The keys that are related to the two territorial units involved are those given in Table 1, and as noted above, they use the first letter to denote the type of alteration (Elimination, E; or Modification, M), and the second and third letters to denote the type of change (Integration, I). Moreover, each key in Table 1 –first column– has a paired key –last column–, which denotes a reverse perspective, when we switch the codes of the territorial units involved in the alteration.

Nonterritorial change types only involve one territorial unit (code). For this reason, they are represented with a single alteration, in which the territorial unit code is repeated. For example, the change of name of F is represented as:

- (F, OG, F): the TU F was called . . .

Territorial changes can affect more than two territorial units at the same time, which is a fairly common case. For example, a division affects at least three territorial units: the one that is divided, which disappears, and the two or more that emerge from the division, which are created. The representation of this change must indicate all the pairs of territorial units related by an alteration caused by the change. Thus, a division in which territorial unit A disappears because it is divided into territorial units X, Y, and Z is represented with the following alterations:

- (A, ED, X) $\rightarrow$ A is eliminated (E) by division (D) into X, among others;
- (A, ED, Y) $\rightarrow$ A is eliminated (E) by division (D) into Y, among others;

Table 1. Keys representing alterations and changes.

Key	Alteration	Change	Relation (between the two municipalities)	Pair
CC	C Creation	Code Change	is created by code change of	EC
CD		Division	is created by division of	ED
CF		Merger	is created by merger, among others, of	EF
CFp		Partial merger	is created by merger of one part, among others, of	MFp
CO		Unspecified appearance	is created from territories not registered as municipalities	CO
CS		Segregation	is created by segregation of	MS
EC	E Elimination	Code change	is eliminated by code change of	CC
ED		Division	is eliminated by division, among others, into	CD
EF		Merger	is eliminated by merger into	CF
EI		Integration	is eliminated by integration into	MI
EO		Unspecified disappearance	is eliminated without further information	EO
ER		Distribution	is eliminated because it is distributed, among others, to	MR
MFp	M Modification	Partial merger	is modified because one part is merged with others to form	CFp
MI		Integration	is modified because it integrates	EI
MP		Exchange	is modified because it exchanges territories with	MP
MR		Distribution	is modified because, among others, it receives a part of the distribution of	ER
MS		Segregation	is modified because it is segregated	CS
MTc		Transfer	is modified, increasing, because it receives a transfer of a part of	MTd
MTd		Transfer	is modified, decreasing, because it transfers a part to	MTc
OA*x*	O Others (Nonterritorial)	Annotation *x*	descriptive annotation of *x*	OA*x*
OG		Name change	was called	OG

Note: Only one set of alterations in nonterritorial types of change is shown, namely *x,* simply to show that the list is open.

- (A, ED, Z) → A is eliminated (E) by division (D) into Z, among others;
- (X, CD, A) → X is created (C) by division (D) of A;
- (Y, CD, A) → Y is created (C) by division (D) of A;
- (Z, CD, A) → Z is created (C) by division (D) of A.

This standardized representation is what we refer to as 'canonical representation', which has the shortest possible label that can be used atemporally. However, when we have a sequence of alterations over time, for example municipalities in various censuses, the representation of each pair of territorial units related by an alteration must include other informative elements as well as the territorial unit codes and keys, for example, a temporal dimension of when the alteration took place, or the old and new names in the case of name changes. These extensions can easily be accommodated.

2.5. *Representation of Complex Change Types*

As well as the changes corresponding to the basic types presented above, other complex types of changes can occur where several basic types of change are combined. The proposed typology covers all the basic types needed to create, through combinations, any change, however complex it may be, by applying exactly the same ideas as for the basic types. For example, let us suppose that the change shown in Figure 1, in which territorial units X and Y are created as a result of the division of unit A, and unit B receives part of territory A. The labels that fully describe the change will be:

- (A, ED, X) → A is eliminated (E) by division (D) into X, among others;
- (X, CD, A) → X is created (C) by division (D) of A;
- (A, ED, Y) → A is eliminated (E) by division (D) into Y, among others;
- (Y, CD, A) → Y is created (C) by division (D) of A;
- (A, ER, B) → A is eliminated (E) because it is distributed (R) to B, among others;
- (B, MR, A) → B is modified (M) because, among others, it receives a part of the distribution (R) of A.

2.6. *Detecting Inconsistencies*

The 'canonical representation' has other advantages as well as allowing the homogeneous representation of all types of situations in modifications to municipalities. One particularly

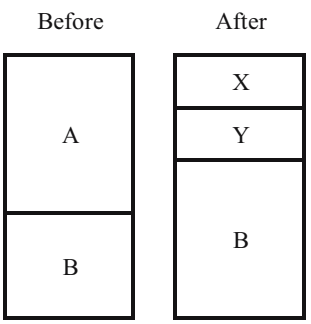

Fig. 1. Example of a complex change.

relevant advantage is that it automatically detects inconsistencies in the changes recorded. To do this, it takes advantage of the fact that alterations referring to territorial changes, those that link two or more units, must always be in pairs. Hence, in the examples given in the previous sections we can see that all type 'CD' alterations between units A and B are paired with another 'ED' type alteration between B and A. This property is due, simply, to the fact that the 'paired' alterations refer to the same information, but one is expressed actively and the other, passively. For example,

- (A, CD, B) $\rightarrow A$ is created (C) by division (D) of B;
- (B, ED, A) $\rightarrow B$ is eliminated (E) by the division (D) into A, among others.

In general, we can express a territorial alteration that affects a pair of territorial units A and B actively in terms of A (and passively in terms of B) or, conversely, actively in terms of B (and passively in terms of A). The pair of each type of alteration is shown in the last column of Table 1.

This property is very useful for detecting inconsistencies in changes to territorial units in the databases. One only has to check that the corresponding pairings match. Any unexpected pairing between two types of alterations will indicate an inconsistency in the base information that represents the changes. We apply this technique in Section 3.

2.7. Rules of Homogenization and Generating Genealogies

Another very useful application of the 'canonical representation' is the generation of homogenous series from the territorial unit structure existing at a given moment. Similarly, this allows us to generate the genealogy of changes in a territorial unit over time.

To do this we have identified backward and forward equivalence rules for each type of change presented in Table 1. These rules allow us to automate the knowledge we have about the different parts that a territorial unit, existing in a given moment, has in the territorial units at other moments in time, whether past or future, and regardless of whether or not these territorial units exist in the reference period. Thus, we establish a systematic way of creating homogeneous series for groups of territories, subject to territorial changes over time, based on applying a type of equivalence rule between the territorial units existing before the change and those existing afterwards. Hence, for each territorial unit A existing before a change, the forward rules allow us to establish the unit, units and/or parts of territorial units that are 'equivalent' to A after the change. Similarly, for each territorial unit B existing after the change, the backward rules allow us to establish the unit, units and/or parts of territorial units that are 'equivalent' to B before the change.

For example, let us suppose a change took place in year T that caused territorial unit A to disappear because it was divided into two new ones, B and C, as illustrated in Figure 2. Before year T only A existed, whereas after year T, B, and C exist, but A does not. The only forward rule will be $A \rightarrow B + C$. Its application to the case of the population will imply, for example, that we must compare the population of A before year T with the sum of the populations of B and C after year T. There are two backward rules for the same example, one for each territorial unit existing after year T: $B \rightarrow A(b)$ and $C \rightarrow A(c)$, where $A(x)$

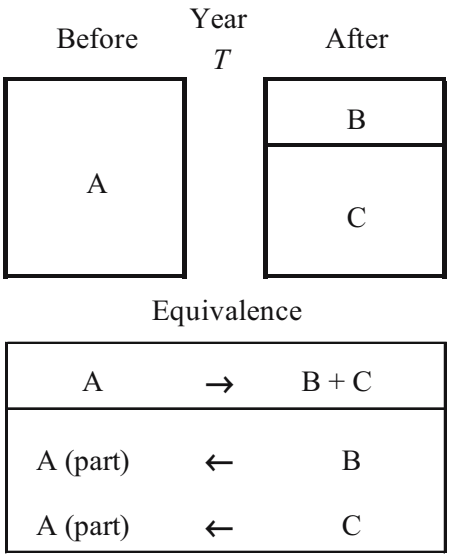

Fig. 2. Example of equivalence rules.

represents the part of territorial unit *A* that was transferred to territorial unit *X*. If we are interested in populations, applying these rules will mean that the population of *B* after year *T* should be compared with the population of a certain part of territorial unit A before year *T*; and the same for territorial unit *C*.

As a real example of the above rules, we may consider the case of *Santurce* (*A* in Figure 2), a municipality that disappeared at the beginning of the twentieth century by splitting its territory into two new municipalities, *Santurce Antiguo* and *Santurce Ortuella* (*B* and *C* in Figure 2). Application of the forward rule is very simple, because it entails comparing the particular characteristic of *Santurce*, that is population or surface, with the sum of the characteristic of the two new municipalities created by division, *Santurce Antiguo* and *Santurce Ortuella*. The application of the backward rules is not so straightforward, since it entails knowing the characteristic of the part of *Santurce* assigned to each of the new municipalities. For example, if we are interested in the homogeneous population of *Santurce Ortuella*, we must know or estimate the part of the population of *Santurce* that is in the actual territory of *Santurce Ortuella*. These backward and forward equivalence rules for each type of change included in Table 1 are fully described in the Appendix.

To automatically apply the equivalence rules it must be kept in mind that each change is represented in the form of one or various alterations. For this reason an algorithm has also been designed that allows the equivalence rules to be applied for a set of territorial units subject to a set of alterations. The algorithm is based on applying the substitutions shown in Table 2, where each row displays what substitutes a territorial unit *X* affected by an alteration (X, $<key>$, Y) by making a forward or backward homogenization. With this algorithm we can generate complete genealogies for a given territorial unit. Naturally, the rules are only applied to territorial type changes as they are not necessary for nonterritorial type changes. Some NSIs have been interested in keeping track of these forward and

Table 2. Substitutions for the automatic application of the equivalence rules.

Key	Relation (X, <key>, Y)	X is substituted by	
		Forward	Backward
CC	is created by code change of	none	Y
CD	is created by division of	none	Y
CF	is created by merger of, among others,	none	Y(part)
CFp	is created by merger of a part of, among others,	none	Y(part)
CO	is created from territories not registered as municipalities	none	none
CS	is created by segregation of	none	Y
EC	is eliminated by code change of	Y	none
ED	is eliminated by division into, among others,	Y(part)	none
EF	is eliminated by merger into	Y	none
EI	is eliminated by integration into	Y	none
EO	is eliminated without additional information about the circumstances	none	none
ER	is eliminated because it is distributed to, among others,	Y(part)	none
MFp	is modified because one part is merged with others to form	X(part) + Y(part)	X
MI	is modified because it integrates	X	X(part) + Y(part)
MP	is modified because it exchanges territories with	X(part) + Y(part)	X(part) + Y(part)
MR	is modified because, among others, it receives part of the distribution of	X	X(part) + Y(part)
MS	is modified because it is segregated	X(part) + Y(part)	X
MTc	is modified, increasing, because it receives a transfer of a part of	X	X(part) + Y(part)
MTd	is modified, decreasing, because it transfers a part to	X(part) + Y(part)	X

backward changes for developing spatio-temporal information systems (Sindoni et al. 2002; Duque 2016). Our typology, and the derived homogeneous rules, provides a complete and neat solution to this problem. An application of this technique to generate homogeneous population series at the municipal level for all the twentieth century censuses based on the municipality structure in the 2011 Census can be found in Goerlich et al. (2015).

3. Application to Detect Inconsistencies in the INE Database *Alterations to the Municipalities in the Population Censuses Since 1842*

This section provides an illustration of how the above typology was used to refine inconsistencies in the database of alterations to Spanish municipalities since their first appearance in censuses in the mid-nineteenth century (INE 2005).

The first attempts to compile the alterations to Spanish municipalities were made by the INE during the work carried out for the 1981 population census (INE 1981a, 1981b). The

aim was to discover which municipalities had disappeared since 1900, without offering any type of systematization, and to find out which municipalities had absorbed them and detect any name changes. Following these initial endeavors, various authors have studied alterations to municipalities, essentially with the aim of constructing homogeneous population series according to the structure of a given census (García 1985, 1994; Goerlich et al. 2006).

Eventually, INE released a database of the original population censuses at municipal level with literal descriptions of all inter-census changes: "*Alterations to the Municipalities in the Population Censuses since 1842*" (INE 2005). The current municipal codification system at a national level dates from the 1970 Census and comprises five digits in the form *PPNNN*, where *PP* is the provincial code and *NNN* a serial number for the municipality within each province. The INE, aware of the importance of a codification system that would allow municipalities to be accurately traced over time, extended the current codification to the municipalities that had disappeared before the 1970 Census. In addition, this process involved tracing the names of municipalities through the censuses, and creating a gazetteer of names, which is the database we use.

After the INE made its database public, the Ministry of Public Administrations prepared a database of municipal alterations since 1842 in which, first, the type of alteration was identified: creation (C), extinction (E), or modification (M) of the municipality in question; and second, the cause of the alteration was identified according to a series of keys (MAP 2008). Unfortunately this database does not have codes, which means it is practically unworkable for our purposes, and it focuses on name changes; however, it is this idea of double entry –alteration *versus* cause – that underlies our typology proposal.

3.1. Detecting Inconsistencies

The first step was to annotate the literal descriptors in the INE (2005) database with the classification keys in Table 1. This preliminary stage was peformed using standard Microsoft tools for extracting, loading, and transforming data – Access and Excel-Power Query – and required a fairly large amount of work, which shows again the benefits of a systematic treatment of the information. Once all the alterations are expressed in the 'canonical form', we apply the technique to detect inconsistencies based on the pairing of alterations presented in Subsection 2.6. The typology defined establishes that the alterations corresponding to territorial changes must always appear in pairs. This information is displayed in Table 3, generated from Table 1.

This information is used directly to detect inconsistencies in the alterations. It consists of identifying the situations in which for an alteration (row or entry in the table) between the territorial units A and B of type P in year T, (A, P, B, T), there is no paired alteration (B, Q, A, T), where Q is the pair key that corresponds to P according to Table 3. The result was that of the 13,424 alterations included in the table in the canonical form with the original INE data, 175 errors or inaccuracies were found, affecting a total of 334 alterations.

Once an inconsistency has been discovered the typology considerably reduces the effort involved in finding the problem, as the origin of the error can be pinpointed to within a few

Table 3. Pairing of types of alterations for each type of territorial change.

Type of change	*Key* 1	*Key* 2
Code change	CC	EC
Division	CD	ED
Merger	CF	EF
Partial merger	CFp	MFp
Integration	MI	EI
Exchange	MP	MP
Distribution	MR	ER
Segregation	CS	MS
Transfer	MTc	MTd

lines, from more than one hundred thousand in the original data. All the inconsistencies detected were investigated and corrected, resulting in a table of 13,415 alterations, of which 8,935 correspond to territorial changes. These figures are reported in Table 4, and the pairing rules are now satisfied in all cases.

It is worth mentioning that the statistics for territorial changes after the corrections displayed in Table 4 does not include all the errors found in the INE database (2005), but only those detected automatically by means of the pairing technique. Goerlich et al. (2015, Sec. 2.2 and 3.3) provide all the errors detected, together with the final statistics for the database of alterations to Spanish municipalities once it had been refined.

The full list of errors and inaccuracies can be consulted in Ruiz and Goerlich (2014). The corrections involved modifications to:

- The type of alteration (263 occasions).
- The census year (27 occasions).

Table 4. Statistics of territorial changes after the corrections.

Change	Alteration 1	Alteration 2	No. cases
Code change	CC	EC	60
Division	CD	ED	32
Merger	CF	EF	446
Partial merger	CFp	MFp	3
Segregation	CS	MS	452
Integration	EI	MI	3,380
Distribution	ER	MR	61
Exchange*	MP	MP	1
Transfer	MTc	MTd	16
Total pairs			**4,451**
Creation from others	CO	–	9
Unspecified disappearance	EO	–	24
Total			**8,935**

*Each exchange also involves two alterations, both MP type, but with the pair of municipalities the opposite way round.

Note: 4,451 pairs = 8,902 alterations.

- The code of the second municipality (22 occasions).
- Eliminating the alteration (14 occasions).
- Adding a new alteration (5 occasions).
- The type of alteration, together with the code of the second municipality (2 occasions).
- The type of alteration, together with the census year (1 occasion).

A comparative review of these errors, their corrections and the content of the entries on the INE website uncovered the reasons that most probably caused these errors in the original INE data. The following are highlighted:

a) Errata or isolated errors, such as a mistake in writing down the code for a municipality in an entry (for example, 1717007 instead of 17007) or in a date (for example, 1960 instead of 1860). Another frequent case involved a municipality being assigned the code that came either before or after it alphabetically in the province list. This seems to suggest that the entries recorded on the INE website were introduced manually one by one and, therefore, without any effective possibility of checking for consistency.
b) Inaccurate use of terms. The same verb was used for different situations, that is, different types of changes (for example, the verbs 'to group' for mergers or partial mergers, and 'to integrate' for distributions, segregations or divisions). It is very difficult to discover these situations through a manual review of the entries; however, they are quickly identified by applying the classification, codification and pairing techniques presented here.
c) Contradictory situations noted in two or more municipalities associated with the same change type. This usually occurs in municipalities where various types of alterations coincided (for example, a municipality grows in size because it incorporates others, and at the same time, because other municipalities transfer part of their territory to it). The system of representation used in our study was an essential tool to avoid these problems as it allowed us to 'visualize' the whole picture of all the alterations associated with each change.

3.2. Detecting Codification Errors

The pairing technique allowed us to uncover situations where the entries on the INE website were inconsistent. However, this does not rule out the possibility of other errors that are not reflected in inconsistencies. For example, if there is a mistake in the code for a municipality in the province of Madrid, 28979 instead of the correct one, 28079, and this mistake appears in every reference to this municipality, there is no way of knowing that this is the wrong code without comparing it with another external data source. To ensure that the codes the INE assigned in its alterations database are the correct ones, we checked that for each code, the municipality corresponds to that indicated in the official 1970 Census, which is when the municipal codes were first established, based on the alphabetical order generated in that census.

This verification revealed that the INE had wrongly assigned codes in the four municipalities reported in Table 5, all of which belong to the province of Teruel. On

Table 5. Errors in assigning codes in the INE alterations.

Code in the INE alterations website	Municipality	Correct code
44138	Luco de Bordón	44139
44139	Luco de Jiloca	44140
44140	Lledó	44141
44141	Loscos	44138

observing their names and the connection between the wrong codes and the correct codes, it seems that the four mistakes were due to the same human error on entering the data for the INE website, which arose because the alphabetical order created by computers differs from the traditional Spanish alphabetical order, which considered 'LL' as a separate letter following all other entries beginning with 'L'.

4. Conclusions

In this article we have presented a proposal for a typology that defines and classifies different types of change that may occur in different territorial units between two moments in time. The types of change were classified into two categories: territorial and nonterritorial. The first category includes various groups, depending on whether they create, eliminate or modify territorial units. The second group includes types of change that do not involve territorial modifications or appearances or disappearances in the list of territorial units, such as for example, name changes.

Each type of change is represented in what we term 'canonical form', which consists of expressing its effect (alterations) between pairs of territorial units. This allows every possible situation to be represented, however complex it may be, in a common format. Additionally, in this type of representation territorial alterations are presented in pairs, such that each type P alteration, between territories A and B, corresponds to another type Q alteration between B and A. In this way, errors due to inconsistencies in the original data sources can be detected automatically. This technique was applied to the data on municipal alterations available on the INE website (2005), which covers all alterations occurring between the censuses of 1842 and 2001. As a result of this procedure, 175 inconsistencies and inaccuracies in the database, as well as four wrongly assigned codes, were detected and corrected.

Appendix

Detailed Description of Basic Types of Change

This Appendix describes in detail the types of changes listed in the hierarchy presented in the study. A schema is used for each type of change that provides the specific name of each change type, a textual description, a graphical illustration of the state before and after the change, canonical representation – the list of alterations involved, using the keys from Table 1, forward and backward equivalence rules to generate homogeneous constructions over time, and an example taken from the INE database *Alterations to the Municipalities in the Population Censuses since 1842*. For this reason, territorial unit is equivalent to municipality in the annex. Additional comments are provided in some cases.

1. Territorial, without creation or elimination of municipalities

This includes types of changes that involve a territorial change, but neither the creation of new municipalities nor the elimination of pre-existing ones.

Type of change:	**TRANSFER**
Description:	One municipality transfers part of its territory to another or other pre-existing municipalities.
States:	Municipality *B* transfers part of its territory to municipality *A*. Before: **A** / **B** After: **A** / **B**
Representation:	(*A*, MTc, *B*): *A* is modified, increasing, because it receives a part of *B*. (*B*, MTd, *A*): *B* is modified, decreasing, because it transfers a part to *A*.
Forward rules:	$A \rightarrow A(a)$ $B \rightarrow B+A(b)$
Backward rules:	$A \rightarrow A+B(a)$ $B \rightarrow B(b)$
Example:	**09323 Rojas**: Between the 1897 Census and the previous one, the municipal area grew because it incorporated 095101 (Quintana Urria) which was previously an entity of 09071 (Carcedo de Bureba). => (09323, MTc, 09071) (09071, MTd, 09323)

Type of change:	**EXCHANGE**
Description:	Two municipalities exchange parts of their territories.
States:	Municipalities *A* and *B* exchange parts of their territories. Before: **A** / **B** After: **A** | **B**
Representation:	(*A*, MP, *B*): *A* is modified because it exchanges territories with *B*. (*B*, MP, *A*): *B* is modified because it exchanges territories with *A*.
Forward rules:	$A \rightarrow A(a)+B(a)$ $B \rightarrow A(b)+B(b)$
Backward rules:	$A \rightarrow A(a)+B(a)$ $B \rightarrow A(b)+B(b)$
Example:	**25138 Montgai**: Between the 1877 Census and the previous one, the municipal area altered because it transferred 25092 (Floresta) to 25153 (Omellons) and received 255079 (Butsenit) from 25153 (Omellons). => (25138, MP, 25153) (25153, MP, 25138)
Comments:	This is the same as two transfers in the opposite direction, one from *A* to *B* and another from *B* to *A*.

2. Territorial, where municipalities are created but not eliminated

This includes types of territorial changes in which a municipality is created but no municipality is eliminated.

Type of change:	**SEGREGATION**
Description:	A new municipality is created by the segregation of part of another municipality's territory.
States:	The new municipality *A* is created by the segregation of part of municipality *B*. Before: **B** After: **A** / **B**
Representation:	(*A*, CS, *B*): *A* is created by the segregation of *B*. (*B*, MS, *A*): *B* is modified because *A* is segregated from it.
Forward rules:	$B \rightarrow A+B$
Backward rules:	$A \rightarrow B(a)$ $B \rightarrow B(b)$
Example:	**02004 Albatana**: This municipality appeared between the 1920 Census and the previous one because it was segregated from municipality 02056 (Ontur). => (02004, CS, 02056) (02056, MS, 02004)

Type of change:	**PARTIAL MERGER**
Description:	Two or more municipalities cede parts of their territories to create a new municipality.
States:	Municipalities *B* and *C* transfer part of their territories to create the new municipality *A*. Before: **B** / **C** After: **B** / **C** / **A**
Representation:	(*A*, CFp, *B*): *A* is created by the merger of onepart, among others, of *B*. (*A*, CFp, *C*): *A* is created by the merger of onepart, among others, of *C*. (*B*, MFp, *A*): *B* is modified because one part is merged with others to create *A*. (*C*, MFp, *A*): *C* is modified because one part is merged with others to create *A*.
Forward rules:	$B \rightarrow B+A(b)$ $C \rightarrow C+A(c)$
Backward rules:	$A \rightarrow B(a)+C(a)$ $B \rightarrow B(b)$ $C \rightarrow C(c)$
Example:	**30902 Alcázares, Los**: This municipality appeared in the 1991 Census as a result of the merger of two parts, one from 30035 (San Javier) and the other from 30037 (Torre Pacheco). => (30902, CFp, 30035) (30902, CFp, 30037) (30035, MFp, 30902) (30037, MFp, 30902)

Type of change:	**UNSPECIFIED APPEARANCE OR CREATED FROM OTHER TERRITORIES**
Description:	A new municipality is created from unspecified territories or from territories that were not municipalities.
States:	Municipality *A* is created from territories that are not municipalities. Before: (empty) After: **A**
Representation:	(*A*, CO, *A*): *A* is created from territories that are not registered as municipalities.
Forward rules:	not applicable
Backward rules:	not applicable
Example:	**52001 Melilla**: This municipality appeared between the 1877 Census and the previous one; previously it was defined as a '*plaza de soberanía*' or sovereign stronghold. => (52001, CO, 52001)
Comments:	Used for special circumstances such as the change of category of certain sovereign strongholds in North Africa to municipalities. This alteration is not noted as such in the INE database, where Melilla appears with its present code, 52001, in the 1877 Census with no additional comment.

3. Territorial, where municipalities are eliminated but not created

This includes types of territorial changes involving the elimination of a pre-existing municipality without the creation of any new municipalities.

Type of change:	**INTEGRATION**
Description:	A pre-existing municipality is fully incorporated into one or more other municipalities.
States:	Municipality *B* disappears because it is integrated into municipality *A*. Before → After **A** / **B** → **A**
Representation:	(*A*, MI, *B*): *A* is modified because it is integrated into *B*. (*B*, EI, *A*): *B* is eliminated because it is integrated into *A*.
Forward rules:	$A \rightarrow A(a)$ $B \rightarrow A(b)$
Backward rules:	$A \rightarrow A+B$
Example:	**01016 Bernedo**: The surface area of this municipality increased between the 1981 Census and the previous one because it incorporated 01005 (Arlucea Marquínez). => (01016, MI, 01005) (01005, EI, 01016)

Type of change:	**DISTRIBUTION**
Description:	A municipality disappears because all its territory is distributed among two or more pre-existing municipalities.
States:	Municipality *C* disappears because it is distributed between municipalities *A* and *B*. Before → After **C** (**A** / **B**) → **A** / **B**
Representation:	(*A*, MR, *C*): *A* is modified because it receives, among others, one part from the distribution of *C*. (*B*, MR, *C*): *A* is modified because it receives, among others, one part from the distribution of *C*. (*C*, ER, *A*): *C* is eliminated because it is distributed to *A*, among others. (*C*, ER, *B*): *C* is eliminated because it is distributed to *B*, among others.
Forward rules:	$A \rightarrow A(a)$ $B \rightarrow B(b)$ $C \rightarrow A(c)+B(c)$
Backward rules:	$A \rightarrow A+C(a)$ $B \rightarrow B+C(b)$
Example:	**49001 Abelón**: This municipality disappeared between the 1981 Census and the previous one because it was integrated, in parts, into municipalities 49023 (Bermillo de Sayago) and 49152 (Pereruela). => (49001, ER, 49023) (49001, ER, 49152) (49023, MR, 49001) (49152, MR, 49001)

Type of change:	**UNSPECIFIED DISAPPEARANCE**
Description:	A municipality disappears and no information is available on the reason why or where its territory goes.
States:	Municipality *A* disappears with no further information about the circumstances. Before: **A** — After:
Representation:	(*A*, EO, *A*): *A* is eliminated with no further information about the circumstances.
Forward rules:	not applicable
Backward rules:	not applicable
Example:	**255122 Cícera**: This municipality, which appears in the 1842 Madoz Census, is not found in the 1857 Census (it belonged to Partido de Cervera). => (255122, EO, 255122)
Comments:	In practice, this circumstance was only found in the 1857 Census with regard to municipalities recorded in the previous census of 1842 (the first to provide a list of municipalities). Presumably this is because the 1842 Census was not totally accurate in the way it distinguished between municipalities and submunicipal entities, since the concept of the municipality had not yet been fully clarified.

4. Territorial, where municipalities are created and eliminated

This includes the types of territorial changes involving both the creation of a new municipality and the elimination of a pre-existing municipality.

Type of change:	**CODE CHANGE**
Description:	The code for a municipality changes; that is, it disappears and another one appears in its place, although it is really the same one (with the same territory).
States:	Municipality with code *B* changes to *A*. Before: **B** — After: **A**
Representation:	(*A*, CC, *B*): *A* is created due to the change of the code for *B*. (*B*, EC, *A*): B is eliminated due to the change of the code for *A*.
Forward rules:	$B \rightarrow A$
Backward rules:	$A \rightarrow B$
Example:	**27901 Baralla**: This municipality appeared between the 1981 Census and the previous one because its name changed, and municipality 27036 (Neira de Jusá) disappeared. => (27901, CC, 27036) (27036, EC, 27901)
Comments:	This is not, strictly speaking, a territorial change, but because what identifies the municipalities is the code, in practice it is the same as eliminating one municipality and creating a new one, which has the same territory and all its characteristics. Code changes are often due to an error when a simple name change occurs. A special case is when the code change is the result of a municipality transferring to a different province.

Type of change:	**MERGER**
Description:	Two or more municipalities disappear because they combine to form a new municipality.
States:	Municipalities *B* and *C* disappear because they are merged into a new municipality *A*. Before: **B**, **C** After: **A**
Representation:	(*A*, CF, *B*): *A* is created from the merger, among others, of *B*. (*A*, CF, *C*): *A* is created from the merger, among others, of *C*. (*B*, EF, *A*): *B* is eliminated by merger into *A*. (*C*, EF, *A*): *C* is eliminated by merger into *A*.
Forward rules:	$B \rightarrow A(b)$ $C \rightarrow A(c)$
Backward rules:	$A \rightarrow B+C$
Example:	**17902 Forallac**: This municipality appeared between the 1981 Census and the previous one following the merger of municipalities 17072 (Fonteta), 17131 (Peratallada) and 17235 (Vulpellach). => (17902, CF, 17072) (17902, CF, 17131) (17902, CF, 17235) (17072, EF, 17902) (17131, EF, 17902) (17235, EF, 17902)

Type of change:	**DIVISION**
Description:	A municipality disappears because it is divided into two or more new municipalities.
States:	Municipality *C* disappears because it is divided into municipalities *A* and *B*. Before: **C** After: **A**, **B**
Representation:	(*A*, CD, *C*): *A* iscreated by the division of *C*. (*B*, CD, *C*): *B* iscreated by the division of *C*. (*C*, ED, *A*): *C* is eliminated by the division into *A*, among others. (*C*, ED, *B*): *C* is eliminated by the division into *B*, among others.
Forward rules:	$C \rightarrow A+B$
Backward rules:	$A \rightarrow C(a)$ $B \rightarrow C(b)$
Example:	**48531 Santurce**: This municipality disappeared between the 1910 Census and the previous one, because it was integrated, in parts, into municipalities 48082 (Santurce Antiguo) and 48083 (Santurce Ortuella). => (48082, CD, 48531) (48083, CD, 48531) (48531, ED, 48082) (48531, ED, 48083)

5. Nonterritorial

This includes the types of change that involve no alterations to the municipalities' territory or to the list of existing municipalities.

Type of change:	**CHANGE OF DESIGNATION**
Description:	A municipality's name changes.
States:	The name of municipality A changes. Before: **A** → After: **A**
Representation:	(*A*, OG, *A*): *A* was called...
Forward rules:	not applicable
Backward rules:	not applicable
Example:	**01001 Alegría-Dulantzi**: In censuses 1842 to 1981 this municipality was called Alegría. => (01001, OG, 01001)
Comments:	The old and new names are recorded separately in additional columns.

Type of change:	**ANNOTATION**
Description:	Another type of nonterritorial change, or complementary information about changes in a municipality.
States:	Another change or other information about changes in municipality *A*. Before: **A** → After: **A**
Representation:	(*A*, OA*x*, *A*): where lower case letter *x* can mean (this list may be extended): - *a*: municipality is known to have changed province, but previous code is not available. - *b*: complementary information about the changes. - *c*: no population data in a given census. - *d*: also appears with a second alternative name. - *e*: change of capital entity ('cap' key in MAP 2008) - *f*: the capital entity name changes ('dca' key in MAP 2008)
Forward rules:	*not applicable*
Backward rules:	*not applicable*
Example:	**13012 Almadenejos**: In the 1842 Census it is mistakenly referred to as "Almagro". => (13012, OA*b*, 13012)
Comments:	The descriptive text must be recorded separately in an additional column.

5. References

Ben Rebah, M., C. Plumejeaud, R. Ysebaert, and D. Peeters. 2011. *Modeling Territorial Changes and Time Series Database Building Process: Empirical Approach and Applications*. ESPON Technical Report 2.1. ESPON 2013 Database. March 2011. Available at: https://www.espon.eu/sites/default/files/attachments/2.1_TR_time_series.pdf (accessed 5 September 2017).

Duque, I. 2016. "Improving Quality and Added Value of Geo-Statistical units." European Forum for Geography and Statistics (EFGS) Conference. Paris, November 15–17, 2016. Available at: http://www.efgs.info/wp-content/uploads/conferences/efgs/2016/S4-1_presentationV8_IgnacioDuque_EFGS2016.pdf (accessed 8 November 2017).

Flora, P., F. Kraus, R. Walton, D. Caramani, J. Marti-Henneberg, and J. Schweikart. 2015. *European Regions: The Territorial Structure of Europe since 1870*. Series: Societies of Europe. Gordonsville: Palgrave Macmillan.

García Fernández, P. 1985. *Población de los actuales términos municipales 1900–1981*. Madrid: Instituto Nacional de Estadística.

García Fernández, P. 1994. *Población de hecho de los municipios de España según la configuración geográfica del censo de 1991. Cifras homogéneas de los censos 1900-1991*. Madrid: Obra Social de la Confederación Española de Cajas de Ahorro.

Goerlich, F.J., M. Mas, J. Azagra, and P. Chorén. 2006. *La localización de la población española sobre el territorio. Un siglo de cambios: un estudio basado en series homogéneas (1900–2001)*. Bilbao: Fundación BBVA.

Goerlich, F.J., F. Ruiz, P. Chorén, and C. Albert. 2015. *Cambios en la estructura y localización de la población: Una visión de largo plazo (1842–2011)*. Bilbao: Fundación BBVA.

Gregory, I.N. 2005. "The Great Britain Historical GIS." *Historical Geography* 33: 132–134. Available at: https://ejournals.unm.edu/index.php/historicalgeography/article/view/2933 (accessed 8 November 2017).

Gregory, I.N. and P.S. Ell. 2006. "Error-Sensitive Historical GIS: Identifying Areal Interpolation Errors in Time-Series Data." *International Journal of Geographical Information Science* 20(2): 135–152. Doi: http://dx.doi.org/10.1080/13658810500399589.

Gregory, I.N. and P.S. Ell. 2007. *Historical GIS: Technologies, Methodologies and Scholarship*. Cambridge: Cambridge University Press.

INSPIRE. 2013. "D2.8.III.1 INSPIRE Data Specification on Statistical Units – Technical Guidelines" Available at: http://inspire.ec.europa.eu/id/document/tg/su (accessed 8 November 2013).

Instituto Nacional de Estadística. 1981a. *Relación de municipios desaparecidos desde principio de siglo*. Madrid: INE.

Instituto Nacional de Estadística. 1981b. *Relación de municipios y códigos al 31 de diciembre de 1980*. Madrid: INE.

Instituto Nacional de Estadística. 2005. Alteraciones de los municipios en los Censos de Población desde 1842. Database distributed by Instituto Nacional de Estadística. Available at: http://www.ine.es/intercensal/ (accessed 8 November 2013).

Ministerio de Administraciones Públicas (MAP). 2008. *Variaciones de los municipios de España desde 1842*. Madrid: Secretaría General Técnica, MAP. Available at: http://www.sefp.minhafp.gob.es/dam/es/web/publicaciones/centro_de_publicaciones_de_la_sgt/Monografias0/parrafo/011113/text_es_files/Variaciones-INTERNET.pdf (accessed 8 November 2017).

Marti-Henneberg, J. 2005. "Empirical Evidence of Regional Population Concentration in Europe, 1870–2000." *Population, Space and Place* 11: 269–281. Doi: http://dx.doi.org/10.1002/psp.37.

Milego, R. and M.J. Ramos. 2011. *Disaggregation of Socioeconomic Data into a Regular Grid: Results of the Methodology Testing Phase*. ESPON Technical Report 2.2. ESPON 2013 Database. March 2011. Available at: https://www.espon.eu/sites/default/files/attachments/M4D_FR_TechnicalReports.zip (accessed 8 November 2017).

Ruiz, F. and F.J. Goerlich. 2014. "Taxonomía y representación de los cambios en los municipios españoles." Working Paper No. WP-EC 2014-01. Valencia: Instituto Valenciano de Investigaciones Económicas. Doi: http://dx.medra.org/10.12842/WPEC_201401.

Sindoni, G., S. De Francisci, M. Paolucci, and L. Tininini. 2002. "Experiences in Developing a Spatio-Temporal Information System." *Research in Official Statistics* 1: 45–57. Available at: http://ec.europa.eu/eurostat/documents/3217494/5644045/KS-CS-02-001-EN.PDF/69d90c27-c788-45bc-87c9-ac3ffa09b06f?version=1.0 (accessed 5 September 2017).

Tir, J., P. Shafer, P. Diehl, and G. Goertz. 1998. "Territorial Changes, 1816–1996: Procedures and Data." *Conflict Management and Peace Science* 16(1)(Spring): 89–97. Doi: https://doi.org/10.1177/073889429801600105.

Received March 2015
Revised September 2017
Accepted November 2017