Two graphical models for the analysis and comparison of cartularies

Authors: Julio Escalona, Cristina Jular Pérez-Alfaro, Anna Bellettini

Abstract

This paper presents and discusses two of a number of methods for the computer-aided analysis of cartularies that are currently under development at the Instituto de Historia – CSIC. The first one, which we call the Order/Date Model, is oriented to the integral visualization and analysis of an individual cartulary as a project. The second, which we call the Order/Order Model, is applied to pairs of cartularies that share at least part of their contents, and is aimed at revealing to what extent the most recent one made use of the oldest. Our method is based upon a relational database that stores all the information about the cartularies and a number of statistical graphs that generate a two-dimensional grid (the Order/Date or the Order/Order grids) upon which additional variables can be displayed. Our method draws on traditional codicological and palaeographical methods of analysis, but it represents a significant development, as it allows to visualize in an intuitive way very complex phenomena that are otherwise hard to grasp or difficult to analyze manually.

Keywords: Computer-aided Analysis of Cartularies, Knowledge Discovery in Databases, Graphical Models for Data Visualisation, Medieval Cartularies, Order/Date Model, Order/Order Model

DOI: http://doi.org/10.16995/dm.55

§ 1 The use of computers to facilitate and enhance the analysis of historical sources of all kinds has attracted the interest of specialists for several decades, with a range of initiatives and approaches that are consistent with the diversity of the sources themselves. More specifically, medieval manuscripts have been increasingly, and notably, part of this renewal, and as information technology is gradually being imposed as the primary means of accessing the information they contain, they will continue to be so. However, closer inspection suggests that this process experienced a significant turn at the end of the 1990s, mainly related to the expansion of the Internet and the development of text mark-up languages. Until then, the priority had been to develop desktop tools that would allow researchers to carry out an exhaustive analysis of a specific source. This changed in the late 1990s, when the spread of the Internet led to a substantial change in the amount, scope and nature of digital products. The number of major initiatives aimed at providing global users with access to large data sets increased exponentially. A constituent feature of this changing trend was the proliferation of cataloguing systems, which were designed to make it possible to recover individual items of information (an archival document, a digital image, a literary text, etc.). The relative predominance of these initiatives can be partly accounted for by their high visibility and immediate impact. This resulted in a change in funding priorities, also caused by the enormous benefits they provide to users. Nevertheless, most of these systems lacked, or had only very basic, analytical capabilities.

§ 2 The computer-assisted analysis of medieval sources and manuscripts has followed different paths, generally inspired by Big Data management and the concepts of Data Mining and Data Visualisation. This is the case, for example, in linguistic, lexicographical, diplomatic studies, as well as in some projects based on prosopography, and in some very promising advances in the field of palaeography. In parallel, the development of computer-assisted text editing has been another major trend (for a recent update with abundant references, see Tinti 2015). It is, however, very difficult to find initiatives that are focused on small-scale and high-intensity analysis of specific sources, such as those which are the subject of this paper, medieval cartularies. Research on cartularies has experienced a conceptual and methodological shift in recent decades, which has involved moving from the traditional approaches that considered cartularies as mere repositories of documentation, to now seeing them as cultural objects worthy of study in and of themselves. [1] There has also been, at the same time, a vast increase in the number of studies of specific parts or sets of
codices. These initiatives include electronic projects, usually of a cataloguing or ecdotic nature, or aimed at allowing access to a source (e.g. CartulR 2006, Becerro Galicano Digital 2013). Nonetheless, there are strikingly few examples of computer analyses of cartularies and, more significantly, they mostly date back to the early 1990s, that is, before the Internet boom (McCrank 1992). This is even more striking when considering that cartularies are an ideal type of source for a computerised approach, given that their nature as codices (susceptible of being analysed from a merely structural, formal standpoint) is linked to their content as individual pieces organised by the cartulary makers. Specialists, of course, are well aware of this and take it into account in their studies at different levels, including more or less automated quantitative and statistical analyses of document content (distribution by chronological periods, document type, type of grantors/receivers, geographical area, etc.). Nevertheless, the overall study of the cartulary as a project is usually carried out through manual and intuitive qualitative analysis.

§ 3 This is precisely the area that our study seeks to focus on. This short paper outlines some of the components of a broader methodology for the overall analysis of cartularies that we have been experimenting with at the Instituto de Historia—of the Spanish National Research Council (known by its Spanish acronym as CSIC). It is a work in progress, with some aspects being more developed, and others merely being in their exploratory phase. We want to focus on two graphical models (which we call the Order/Date Model and the Order/Order Model) for the visualisation and analysis of the cartulary as a project. These models are now mature enough to be presented, and it is our hope that, when they are tested by other researchers, their contributions will refine them beyond the capabilities of a single research group.

§ 4 The two models we propose are mere logical procedures, independent from any specific software, which can be applied by anyone on a conventional desktop computer. Theoretically—in simple cases—they could be applied manually, but it is more efficient to use a database management system (DBMS) capable of generating statistical tables or graphs, such as Filemaker™ or Microsoft Access™. They can also be tested, in the simplest cases, by means of any spreadsheet programme such as Microsoft Excel™, but this is not recommended, because the more sophisticated analyses almost always require a relational structure with multiple tables. The examples presented in this paper have been drawn from a relational database created in Microsoft Access™, generating visualisations using the graphics engine that comes with Access itself, although occasionally the data has been exported to Excel to take advantage of its greater flexibility in generating graphs. Graphical outputs have been designed to be deliberately simple, in order to highlight that they are conceptually operational, above and beyond their potential for visual display.

Cartularies and Diachrony

§ 5 Among medieval and modern codices, cartularies are a special theoretical case due to the role diachrony plays in both their structure and content. From a structural point of view, cartularies were often developed over a long period of time. Even when they are part of an initial single and 'closed' project, very often they go through successive stages of additions, continuations, insertions and modifications that result in “complex codices,” both from a codicological and a content point of view. Applying an archaeological metaphor, it could be said that cartularies are open to a “stratigraphic” type of analysis, that is, one capable of revealing the successive stages of the production and modification of the project, whether or not they are reflected in the materiality of the codex. This diachronic character, which is common to other types of codices (with some variations), in the case of cartularies is reinforced by the nature of their contents. Cartularies generally record—in full or abbreviated form—legal documents which did not originally form a single corpus. They do so according to one criterion or many criteria for organisation purposes, which could vary greatly depending on the plan and purpose of each project. A key attribute of the recorded texts is that they collect legal acts that occupy a specific place in time (also in space, but in this paper we shall not consider this variable). For undated documents, a chronology is assigned that depends on well-established diplomatic methods that are independent from the analysis being undertaken here. Understanding the cartulary as a project involves seeing it as a strategy. What treatment is given to individual documents? What elements are discarded and selected, and how are they organised? What is the logic behind this organisation? Is there a single project or several? Have there been changes to an initial plan? The Order/Date Model discussed in the first part of this paper exploits this diachronic dimension of the contents of the cartulary, in an attempt to visualise the project behind it and its “afterlife” (Escalona and Sirantone 2013).
The diachronic character of the cartularies can have an even more complex aspect in those cases that can be referred to as examples of “multiple cartularisation.” It is not uncommon for an institution to produce a cartulary over time whose contents overlap to a greater or lesser extent. Therefore, the same corpus of documents (or its successive development phases) could be reflected in more than one cartulary. This inevitably raises the difficult question of the successive uses given to the same documents, but also to the previous cartularies. Why is the old cartulary no longer sufficient to meet the present needs? What is it about the new one that makes it more suitable? What are the differences between them? In the second part of this paper we present the Order/Order Model, which is specifically designed to gauge the extent to which two cartularies that record (either fully or partially) the same diplomatic collection coincide or diverge in their approach as projects.

In order to present these two analytical models, some hypothetical and actual examples from cartularies have been used. The intention here is not to analyse these codices in depth, but to show the usefulness of the method and suggest possibilities for its future expansion. Therefore, many of the specific issues surrounding the individual codices used have been disregarded, and the focus instead is on the analysis of cartularies using these two models.

Visualising the structure of a cartulary: The Order/Date Model

The Order/Date Model exploits the inherently diachronic character of cartularies by working from a “null hypothesis”: that the cartulary documents are organised in chronological order. This condition is only met in very exceptional cases, such as the so-called “chronicle cartularies,” but the null hypothesis is useful to the extent that by rejecting it, other alternatives can be considered. To visualise the chronological order of a cartulary two variables are used which give the model its name:

- **Order**: the position of each document as it appears in the cartulary. This is deemed to be a perfect sequential order to be considered carefully in any real-case analysis. For the purpose of this presentation and for the sake of clarity, complexities such as the presence of pieces inserted in the margins have been disregarded.
- **Date**: the year in which the document is dated (a chronological resolution greater than one year is rarely required). Again, in this case the starting point is a theoretical homogeneous set of dated pieces, leaving aside the problems of undated documents and those to which dates are to be assigned within a certain range.

The Order/Date Model consists in graphically representing the position of each document on a coordinate axis system based on these two variables. The horizontal axis (the x-axis) shows the sequential order of the cartulary from the first to the last document. The vertical axis (the y-axis) represents the year in which the document is dated. This provides a graphical representation of each of the documents in the cartulary, which then has to be interpreted. Attention needs to be given to the graphic patterns resulting from the image.

Identifying graphic patterns: Chaotic pattern, coherent pattern.

Working on the null hypothesis, let us consider the (very rare) case of a cartulary organised according to the chronological principle. Fig. 1 shows this possibility using hypothetical data, which produces a consistently upward line (date values always increase or remain the same, but never decrease).
Conversely, if the cartulary is not in chronological order, the resulting image will be alternating vertical positions in which no pattern can be recognised (Fig. 2).

The disorganised appearance of this second pattern, however, can only be defined in terms of the scale of representation, as it is common for groups of documents not to have a strictly chronological order, but to be grouped within narrower time ranges than the rest of the documentation. Therefore, when defining patterns, the scale of the visualisation has to be taken into account, as illustrated eloquently below in the real case of the Arlanza cartulary.

In short, the Order/Date Model produces an image on which it is necessary to distinguish whether the sequence or parts of it fit one of two ideal patterns:

- **Chaotic pattern**: The documents are distributed without a recognisable time pattern. The initial reading would be that the order adopted is not chronological, but based on other criteria (donors, geographical distribution, legal document type) or more often, a combination of several of them.
- **Coherent pattern**: This pattern is more difficult to define, but its recognition is crucial to interpreting the model. It is a distribution of documents that has a chronological pattern, either because it follows a temporal order or because, as a whole, it keeps to a discernible temporal range, even though internally it is not organised in a strictly chronological order.

It is rare for a real case to fully fit one of the two patterns. More often some sections of the cartulary are found that fit one or the other, and identifying the turning points is the main objective in order to understand the diachronic nature of a cartulary as a historical object. In order to illustrate this more clearly it is necessary to use a real case: the Cartulary of San Pedro de Arlanza.

The simple Order/Date model: only two variables
§ 15 The first Order/Date Model was designed by Julio Escalona and Pilar Azcárate and published in 2001 as part of a strategy to analyse this important Castilian cartulary (Escalona and Azcarate Aguilar-Amat 2001, Escalona, Azcárate, and Larrañaga 2002, Azcárate Aguilar-Amat et al. 2006) which disappeared during the Spanish Civil War (1936-1939). However, data from this cartulary have been partially disseminated through an edition published about eleven years before its disappearance (Serrano 1925). This edition copied the documents in chronological order, but it hardly provided other information than the place that each document occupied in the order of the cartulary. In this case, correlating the internal data of the documents —mainly the date— with their position was the only way to explore the structure of the cartulary. The Order/Date Model produced between 1998 and 2000 —albeit not yet bearing that name— was surprising, as it revealed an eloquent pattern, as reflected in Figure 3.

![Figure 3: Order/Date Model of the cartulary of San Pedro de Arlanza. (Numbers ‘-2’ and ‘-1’ correspond to thirteenth-century additions in the first folio that have no numbering in Serrano’s edition)](image)

§ 16 What is immediately noticeable is the contrast between a chaotic pattern that is dominant at the beginning, followed, from No. 78 onwards, by a coherent set of documents confined to a narrow chronological range, between 1130 and 1156. A more detailed analysis made it possible to define not two, but up to four sets of documents (Fig. 4).

![Figure 4: Order/Date Model of the cartulary of San Pedro de Arlanza. The boxes indicate Blocks A and B](image)

- Document Set A shows a very irregular pattern, with strong variations in the dates. This is the typical pattern for a cartulary that is not organised in chronological order, with the highest dates pointing to 1119 as a date post quem for its composition.
In contrast, Document Set B represents a completely different strategy, because without having a strict chronological order, its components are restricted, as noted, to a more coherent time range, from 1130 to 1156. The pattern of Document Set B can be described as an addition made after the initial production stage of the cartulary was completed, which was merely intended to update it by incorporating documents obtained after that period.

Document Set C returns to a chaotic pattern, with the added complexity that it includes two undated pieces which, in this case, appear conventionally represented on the X axis.

Finally, Document Set D corresponds to four significantly later additions that were incorporated at the beginning and at the end of the cartulary.

Thus, the Order/Date graph enables the history of this lost cartulary to be interpreted. The initial project was perhaps conceived in the period 1120-1130, and is recognised by its chaotic chronological pattern (Document Set A). This initial cartulary was continued sometime after 1156, with the addition of a number of documents belonging to a very narrow time span (Document Set B). Later various heterogeneous parts were added (Document Set C), and finally a few additions were incorporated in the thirteenth century, at the beginning and at the end (Document Set D). In the next section it will be seen that it is possible to be more precise about the details of the Arlanza cartulary, but for now suffice it to note that the simple Order / Date Model allowed for the generation of some hypotheses about the cartularisation project, and what is important for this paper, the defining of the two basic graphical patterns, coherent and chaotic patterns.

The complex Order/Date Model: adding more variables

The core of the Arlanza cartulary features a chaotic chronological pattern. This makes it possible to distinguish it from the later stage, but it is not in itself a very powerful analysis tool. However, the fact that the pattern is chaotic does not mean its analysis has to be discarded. It is possible to set up and test alternative hypotheses by adding supplementary variables to the two basic ones (order and date), and to explore whether there are conceptual groupings independent from the chronological order. In the case of Arlanza, if a third variable is added, such as the grantor of the documents, the underlying strategies can be interpreted further. In Figure 5, the Order/Date Model for Document Set A and Document Set B (as defined in Fig. 4) distinguishes the documents issued by the Kings of León and the Counts of Castile from the rest.

It is immediately apparent that Document Set B shows a very high proportion of royal documents, suggesting that this not only consisted of the addition of pieces from the 1130 to 1156 range, but that the documents had been selected in such a way as to give preference to royal charters. Document Set A also indicates an internal division. The charters granted by royalty and counts do not form a continuum at the beginning of the codex, but they do appear in a much higher proportion in the first thirty texts than in the remainder. This suggests that the original cartulary project was organised by blocks of ownership, which often originated from donations by counts or royalty; hence these pieces appear concentrated, but without
continuity. The rest of Document Set A contains more varied documentation, with a high proportion of private documents.

§ 20 So far we have used this case as an illustration because it is a fairly simple one, but the true potential of the method lies in applying it to complex cartularies, as they are difficult to visualise in a comprehensive manner. As far as complexity goes, one could hardly find a more complex case in the Iberian Peninsula than the so-called Becerro Gótico de Valpuesta. This is a factitious codex that brings together various fragments created at different points in time over the eleventh and twelfth centuries, each with its own page layout, calligraphy, content, etc., which never made up a single project, but were bound together some time before 1236. For the purposes of this paper, it has the major advantage of having recently undergone a splendid edition that includes a facsimile reproduction of the codex and a comprehensive study of its codicological, palaeographic and diplomatic aspects (Ruiz Asencio, Ruiz Albi, and Herrero Jiménez 2010). Therefore, without providing any further details about the codex, it shall be used as an example to show how, the richer the information available, the more analytically powerful the Order/Date Model is. Figs. 6 and 7 display the Order/Date Model of the Becerro Gótico de Valpuesta, first in its simplest form, using order and date as sole variables, and then with the addition of a third variable.

![Figure 6: Order/Date Model of the Becerro Gótico de Valpuesta](image)

§ 21 Its heterogeneous appearance is immediately striking, with an alternation between very contrasting sections: some show the pattern we have called chaotic, while others, as in Document Set B of Arlanza, have a coherent pattern: the documents—not necessarily in chronological order—cover a short time range. But in the case of Valpuesta, the available information allows the incorporation of more variables and further analysis. For example, the graph can be coded on the basis of the codicological structure, with a different colour marking each quire (Fig. 7).
§ 22 By doing this it becomes obvious that there is a high correlation between the changes in the pattern indicated above and the codicological structure, which, given that it is a factitious codex, leads to the interpretation that there were separate production stages or subsequent alterations to the codicological structure.

§ 23 It is still possible to generate a more complex image by considering that the quires are a sequential variable, which can therefore be represented on the x-axis, leaving space for the addition of one more variable, in this case, script type (Fig. 8).

Figure 7: Order/Date Model of the Benserro Gótico de Valpuesta symbolized by quires

Figure 8: Order/Date Model of the Benserro Gótico de Valpuesta symbolized by writing type (C = Caroline; V = Visigothic)

§ 24 The x-axis now shows a double grouping: on the one hand, by the order of documents and by quires, and on the other, by the two main script types in the codex. The editors of the cartulary identified a total of thirty-two different types of handwriting, the graphical representation of which allows for further details to be identified. However, for the sake of clarity, in this graph we simply distinguish between Visigothic and Caroline scripts. The graph eloquently displays how the blocks of Caroline script tend to form compact units.
embedded among others in Visigothic script, and how this usually coincides with codicological units (quires 5-7, quire 11), confirming their status of independent blocks bound in between older fragments. Moving further, the Order/Date Model also reveals that two of the fragments in Visigothic script (quires 1-4, quire 13) adopt the chaotic pattern, typical of a non-chronological order and would indicate that it is a “typical” cartulary. This does not happen in quire 8, which, with the exception of a piece at the beginning, follows a coherent pattern (narrow chronological range), indicating an addition to the set formed by quires 1-4. This is not the case for quire 14 either. It is a monographic dossier on the monastery of San Pedro de Buezo which constitutes a special case, as it only contains documents from 950, with a folio inserted in between containing several texts in Caroline script (Ruiz Asencio, Ruiz Albi, and Herrero Jiménez 2010, p. 54). In contrast, the blocks in Caroline script clearly adopt the coherent pattern. This indicates a significant divergence from the plan. The two main blocks in Visigothic script correspond to “typical” cartularies (in fact it has been suggested, and with good reason, that the Becerro Gótico de Valpuesta contains fragments of two originally separate cartularies in Visigothic script: (Ruiz Asencio, Ruiz Albi, and Herrero Jiménez 2010, p. 58-60), while the Caroline script blocks formed either loose dossiers, or later additions to one or other of the blocks in Visigothic script.

§ 25 The cases presented so far have been selected to gradually illustrate the basic possibilities of application of the Order/Date Model. The potential (and the analytical difficulties) increase significantly with more complex cartularies, as in the example of Valpuesta. We have only shown a few basic features in order to demonstrate how the method can produce very intuitive visualisations, in this case from a study which involved a very thorough analysis previously carried out by the editors. It also enables further analysis by adding variables and producing ad hoc graphs to examine specific aspects. Nevertheless, this is not the purpose of the study being discussed here, and it would make it excessively long. It would be appropriate, however, to provide some details about the potential application of the Order/Date Model to the analysis of complex cartularies such as the Becerro Gótico de Cardeña, the main Castilian cartulary prior to the twelfth century (Serrano 1910, Martínez Díez 1998, Fernández Flórez 2014). It is a cartulary of great formal quality and size, containing 373 texts distributed in 99 folios, with an arrangement that, in general, favours clusters by units owned by the monastery. The Order/Date Model allows this complexity to be comprehensively visualised (Fig. 9), but also the analysis of specific components to be shown (Fig. 10), and such complexity to be deconstructed in a multi-scalar way, zooming in on specific sections (Fig. 11).

Figure 9: Order/Date Model of the Becerro Gótico de Cardeña

§ 26 The sequence, as visualized through the simple graph version (Fig. 9) is really hard to interpret because of information density. However, several patterns are immediately apparent, and in general, it is easy to perceive that the bulk of documents in the Cardeña collection were issued either in the mid-tenth or in the second third of the eleventh centuries, with a much thinner representation in between. This, however, is not much to say. The overall visual impression is too chaotic to handle.
Figure 10: Order/Date Model of the Becerro Gótico de Cardeña symbolized by subject. The number of identified subjects (based upon the rubrications that accompany each entry) is 98. The resulting pattern is very hard to interpret integrally, but the graph is suitable to orientate further analysis (see Fig. 11).

§ 27 In the case of Cardeña, we can count on the additional aid of document titles, since almost every piece included in the cartulary has one that identifies the main theme by which they were classified. By adding that information as our third variable, we can display the location of documents with the same tag across the Cardeña cartulary. Fig. 10 yields a very eloquent—even if at first sight puzzling—picture. There are evident clusters of documents with the same tag (sharing the same symbol). Vertical or almost vertical lines linking adjacent dots (as in the first two clusters at the top left) mean that the documents refer to the same topic (or place) while their dates may vary notably. However, in most cases there are also outliers (represented by long coloured horizontal or diagonal lines). This is the case of the second cluster on the left (brown colour), which is placed at the beginning of the book, but has one outlier in the middle and two more at the end. Thus, Cardeña's apparently unreadable pattern does have a meaning: the Cardeña cartulary makers might have identified all pieces with a tag, but their sorting is anything but watertight, which may be interpreted in a number of ways, including the hypothetical re-working of a previous cartulary.

Figure 11: Order/Date Model of the Becerro Gótico de Cardeña showing one specific subject against all others.

§ 28 The efficiency of the Order-Date model in studying a complex cartulary like Cardeña's can be further exploited by filtering the data to represent single cases (or small groups thereof). Fig. 11 shows the general graph with the superimposition of charters under the tag “Burgos.” The image clearly illustrates the aforesaid
overall impression: pieces tend to cluster in one specific section, but have outliers, especially at the final leaves of the codex. A case-by-case analysis may confirm whether this apparent inconsistency repeats itself throughout the collection or is limited to specific tags or periods.

Comparing cartularies: The Order/Order Model

§ 29 The second method we propose is aimed to facilitate comparison between two cartularies which contain at least a significant block of documentation in common.[2] This mainly occurs when an institution produces more than one cartulary throughout its history, a situation whose individual causes and characteristics are subject to much variability. There are “second generation” cartularies which merely served to update and enhance an older one that was considered to be obsolete; but new cartularies could also be produced for use in specific operations or with a selective theme (using, for example, only royal, or papal documents). When two cartularies exist that, either totally or partially, include the same documentation, they can be compared by asking a basic question: to what extent does the second cartulary replicate the structure of the first? In this case, the null hypothesis to be verified would be whether the second cartulary closely follows in the footsteps of the first. To do so, we propose a method of graphical representation we call Order/Order Model. The reason for this name is that it compares the order of the documents in each of the codices on two axes.

§ 30 When building the Order/Order Model, the starting point is a matrix organised on the basis of the documentary pieces copied in each of the two cartularies to be compared. To that first column should be added two others that record the position of each text in cartulary A and cartulary B. These are sets of three values in which the document number identifies the individual piece, while the other two (position in A, position in B) function as a pair of coordinates for graphical representation. From this matrix a graph is generated in which the x-axis represents the order of positions in cartulary A and the y-axis represents the order of positions in cartulary B. When a document appears in only one of the two cartularies, its position value in the other one will be 0, and therefore it will be represented on the relevant axis.

§ 31 How is an Order/Order Model read? The Model starts from a simple principle: a cartulary is based on the structure of an older one, to the extent that its documents are copied in the same order that they had in the older one. The purpose is to assess to what extent there is a correlation in the order of the documents in both cartularies. A correlation exists when an increase of one unit in the position of the documents in cartulary A results in an equal increase in cartulary B. When this happens, a straight line is generated whose slope is 1 (45 degrees). Fig. 12 shows an (improbable) ideal case of perfect correlation: the documents occupy exactly the same position in A and B, and hence the second is an exact copy of the first (in structure, not necessarily in the details of the texts).

<table>
<thead>
<tr>
<th>Id-doc</th>
<th>Pos. A</th>
<th>Pos. B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 12: Order/Order Model of a hypothetical case of total correlation between two cartularies (cartulary B is a perfect copy of cartulary A)

§ 32 But in practice correlations are not perfect, and to interpret the model it is important to realise that what matters are not the absolute positions, but the relative increases in the order of the documents. In another ideal, but less homogeneous example (Fig. 13), a more complex sequence is shown: cartulary B omits the first four documents that appear in A (shown as 0 on the x-axis), and then matches documents 5-16 in the
same order and position as in cartulary A, with the exception of no. 15, which it omits (again as 0). Finally, the documents that occupy positions 17-20 in cartulary A appear at the beginning in cartulary B.

§ 33 As in the previous section, the Order/Order Model can be applied to real cases. To demonstrate this the two cartularies known as Becerro Gótico (“Gót””) and Becerro Galicano de Santa María de Valpuesta (“Gal”) (Fig. 14) will be used. The older one (Gót) is a very heterogeneous factitious codex composed between the eleventh and twelfth centuries (Ruiz Asencio, Ruiz Albi, and Herrero Jiménez 2010). The more modern one (Gal) is a very homogeneous codex, dating from 1236, whose colophon explicitly states that it is a copy of the Gótico, which it aimed to “renew” (Ruiz de Loizaga 1995).
On a first approach the graph shows a very high correlation, marked by the dominant diagonal line. The anomalies consist mostly of documents contained in only one of the two and in some elements that appear in a different position. The Gót documents omitted in Gal are represented as 0 on the x-axis. The only three cases of documents added to Gal that were not in Gót are represented as 0 on the y-axis. The documents that are located at different points in the two cartularies are shown as points that do not follow the general trend of the straight line with slope 1. Without going into specifics (a question remaining for another study), the Order/Order Model clearly reflects how Gal closely follows the organisation of Gót, as well as making it possible to visualise the points at which there is no correlation, in order to focus the analysis on them by incorporating other issues such as codicology or text content.

The Premonstratensian monastery of Treviño has two cartularies from very distant dates: the twelfth to the thirteenth centuries for the first, to be referred to as Cartulary, and 1676 for the second, to be referred to as Becerro. From a total of 231 identified documents, both codices have 122 (53%) in common, which is a sufficient overlap for analysis. As in the case of Valpuesta, it is not possible in this paper to make a detailed presentation of these codices, the problems entailed in their study, and their vast formal richness and content. The scope of this section will be limited to a comparison of their respective structures by using the Order/Order Model. Treviño's case is particularly eloquent in this regard because in the prologue of the 1676 Becerro there is a statement which reveals its internal logic, and that logic is not the same as that in the
medieval Cartulary. To probe this organising principle different symbols have been used to represent document types. As these are different cartulary projects, one would expect there to be a complete mismatch between the two codices, but the Order/Order Model shows that this is not the case (Fig. 15).

Figure 15: Complex Order/Order Model of the Cartulary and the Becerro de San Miguel de Treviño symbolized by content type

§ 36 Firstly, the above-mentioned 53% overlap is shown in the large number of documents represented as 0 on the two axes: those on the horizontal axis are the texts that are not included in the medieval Cartulary, but are included in the seventeenth-century Becerro; those on the vertical axis are pieces from the Cartulary that were not included in the Becerro. Secondly, in the part where the two coincide, five distinct blocks can be seen:

1. By reading from left to right, there is indication that for documents 2 to 20 the Becerro followed its own plan: when documents were copied that were already in the Cartulary, the result is a totally chaotic pattern in which a clear trend cannot be established; when moving to the right, the documents are sometimes higher and others lower on the y-axis. Initially, and without excluding ulterior corrections, this image could be interpreted as an indication that the author of the Becerro may have elaborated this part from the documentation of the archive, or may have used the Cartulary, but in either case, ignored its organisation. It is important to note here that if the whole of the work had been independent, the entire sequence should be equally chaotic, but, as we shall see, this is not the case.

2. A different pattern can be seen between documents 20 to 40: all documents that are copies of those in the Becerro appear in the same sequential order as in the Cartulary, whether or not they are consecutive. From left to right, each document is always displayed with a higher value of y than before, although there are intermittent values. This is extremely revealing, because it indicates that while in Block A, the author of the Becerro was keen to present his own plan, in this new Block (the chapter on Donations from Benefactors) some pieces were extracted from the Cartulary that matched the desired typology (donations) while the rest were disregarded, but always in a forward progression.

3. From documents numbers 40 to 75 there is a new change to the plan. Now the Becerro copies later documentation, obtained by the monastery after composing the Cartulary, which therefore had not been incorporated into it. This is why they appear on the x-axis, with the exception of some sporadic documents that were in the Cartulary, of which the author of the Becerro probably was not aware and therefore he included them on the basis of the original records. Similarly, while the donation and confirmation types predominate, three documents of different types are interspersed within this block, perhaps because the author felt that this typology could be homogenised, as is also the case for the sales and purchases.

4. From document 75 onwards the Becerro follows the Cartulary closely. It ignores some documents, but the upward trend of the line leaves no room for doubt: it is based on the Cartulary. In this case, when using codes on the graph based on document types, the decisions made by the author of the Becerro's
can be better understood. The sale and purchase and documents were not reorganised, but taken directly from the *Cartulary*.

5. The last block corresponds to later typologies (property rights and censuses) that did not exist at the time when the *Cartulary* was composed and, therefore, appear on the *x*-axis.

§ 37 An additional element to consider is that the author of the Becerro is **not interested** in certain types of documents, such as surveys, which occupy an important part of the *Cartulary*, but are systematically excluded from the *Becerro*, and so they appear on the *y*-axis of the graph. Ultimately, the Order/Order Model shows that the author of the 1676 *Becerro* made a complex use of the medieval *Cartulary*. He disregarded its content to articulate his own discourse at the start but followed it closely when a rearrangement of the original material was not necessary. The capability for the Order/Order Model to combine the relative positions of the documents in both codices with the symbols of their qualitative elements is revealed as a powerful tool for visualising relationships that are otherwise difficult to identify and to test.

**Conclusions**

§ 38 In this paper two graphical methods of analysing cartularies have been presented. The first, the Order/Date Model, allows for the whole structure of a cartulary to be visualised and analysed; in addition, it makes it possible to focus on some sections of a cartulary, based on a two-dimensional representation (order/date of the documents) to which other variables can be added for a more sophisticated analysis. Some avenues have also been shown for the interpretation of the results, based on the identification of two basic patterns: a *chaotic* pattern, typical of cartularies that are not organised chronologically, and a *coherent* pattern, which shows either a chronological order, or a concentration of items within a limited time range. The examples discussed illustrate both situations and the various interpretations that may arise when they are identified in real cases, such as those of *Arlanza* and *Valpuesta*.

§ 39 Another method has also been presented, the Order/Order Model, to compare cartularies with sufficiently overlapping contents. As in the previous case, it is based on a two-dimensional graph generated from the position values of each document in the two cartularies being compared, to which a third variable could be added for further analysis. This model indicates to what extent one cartulary reproduces the order of a previous one, and facilitates the visualisation of the points of alignment/misalignment. The examples used illustrated how powerful the method can be. The cartularies of *Valpuesta* followed a very coherent pattern, while those of Treviño showed a much richer and complex scenario where the different strategies used by the author of the seventeenth *Becerro* (when handling the medieval *Cartulary*) can be seen.

§ 40 The methods presented here are only two of several ways to approach the study of cartularies by the generation of two-dimensional graphs. Our group has experimented with other options to effectively probe specific aspects of cartularies, such as iconography, codicological structure, format and page layout, additions in different hand-writing, and the analysis of insertions and alterations. These questions will form the subject matter of future papers. One of the most promising areas is the ability to analyse cartularies whose original codicological structure has been altered in order to generate reconstruction hypotheses. Similarly, the integration of the data into a Geographic Information System allows for the application of spatial analysis techniques to explore the organisation of a cartulary by geographical area. All these aspects are already being used and studied by our research group.

§ 41 Since the principles of the method rest on logical processes independent from any particular system, in theory they could be implemented manually. However, implementation on a computer system is essential to optimise efforts and achieve a sufficiently sophisticated analysis. A database management system associated with a graphics engine are the minimum tools necessary for this to work successfully. However, our work is aimed at integrating these procedures in a specific piece of software to analyse cartularies, which will require not only more developed visualisation capabilities than those presented here, but also routines for the automatic recognition of patterns that to date have only been recognised visually. The same applies to the identification of point discontinuities or changing trends that may seem obvious in the simpler graphs, but are much less clear in large and complex codices. It must be stressed that the potential ability to visualise the entire codex in a single image is a much more powerful tool than one based on tables or text descriptions.
Finally, we would like to encourage the scientific community to make their contributions to test and refine these methods, not only for research purposes but also for teaching and public outreach purposes. It is our hope that some of them, such as the Order/Date Model in the near future, may lead to standards generally applicable to the study and edition of this type of historical sources.

Notes

[1]. The classical reference for this conceptual shift is the 1991 Paris Conference Les Cartulaires (Guyotjeannin, Morelle, and Parisse 1993), which triggered a sustained wave of studies in most European countries, up to the present. French research remains the most active in this field, especially through the works of Pierre Chastang, who has proposed the term ‘cartularisation’ to define the emergence of cartularies as a documentary genre in the central Middle Ages (Chastang 2001, 2006). Major landmarks in the process are a number of national and international conferences and collective volumes including the Colloquium of the Commission Internationale de Diplomatique in Princeton and New York in 1999 (Kosto and Winroth 2002) and the 2002 Béziers Conference on the cartularies from southern France (Le Blévec 2006). The most recent major international meeting is the Lisboa 2015 Conference Cartularies in Medieval Europe (http://medievalcartularies.letras.ulisboa.pt/). For a more detailed contextual discussion, see (Escalona and Sirantoine 2013).

[2]. The cases of multiple cartularisation have been insufficiently researched to date. Their study is a very promising avenue of research for the immediate future. In the Iberian Peninsula steps have recently been taken in this direction, as shown in the comparative study of three cases presented at the international conference entitled “Cartularies in Medieval Europe: Texts and Contexts” held in Lisbon in June 2015 (see Tinti, Peterson and Agúndez 2015).


[4]. The Becerro plan announced that there would be eight chapters: 1) Foundation of the monastery; 2) Donations by the founders; 3) Donations from benefactors; 4) Papal and Royal Confirmations; 5) Sales and Purchases; 6) Property rights; 7) Censuses; 8) List of Abbots (Abadologio) (which probably eventually gave rise to a separate volume). AHN, Cod. 1375, fols 1v-2r.

Acknowledgements

This study was conducted with the support of the research projects of the Spanish National R & D Plan ref. HAR2013-47889-C3-2-P, led by Julio Escalona and ref. HAR2012-34756, led by Cristina Jular Pérez-Alfaro, respectively. Some preliminary aspects of this research were presented at the international conference entitled “Cartularies in Medieval Europe: Texts and Contexts,” held in Lisbon, Portugal, on 11-12 June, 2015 and the PraME Séminaire (Séminaire Pratiques médiévales de l’écrit) at the University of Namur, Belgium, on 3 December, 2015. The authors wish to acknowledge the criticisms and comments received from those present at these events. The draft was read by Francesca Tinti and we particularly appreciate the suggestions received from her. In the early stages of research, we benefitted from the comments and insights provided by Carmen Escalona Noguero, whom we heartedly thank. Finally we want to acknowledge the very helpful criticism received from Digital Medievalist's two anonymous readers, who have helped us greatly in defining the final version of the text. Any errors, however, are the sole responsibility of the authors.

Works cited

http://www.ehu.eus/galicano.


