The contribution of historical information to flood risk management in the Tagus estuary

Ana Rilo\textsuperscript{a,d,*}, Alexandre Tavares\textsuperscript{b}, Paula Freire\textsuperscript{a}, Pedro Pinto Santos\textsuperscript{c}, José Luis Zêzere\textsuperscript{d}

\textsuperscript{a} LNEC - National Civil Engineering Laboratory, Estuary and Coastal Zones Unit, Avenida do Brasil 101, 1700-066 Lisbon, Portugal
\textsuperscript{b} Centre for Social Studies and Earth Sciences Department, Coimbra University, Rua Silvio Lima, Pólo II UC, 3030-790 Coimbra, Portugal
\textsuperscript{c} Centre for Social Studies, Coimbra University, 3000-995 Coimbra, Portugal
\textsuperscript{d} Centre of Geographical Studies, Institute of Geography and Spatial Planning, Universidade de Lisboa, Edifício IGBT, Rua Branca Edmée Marques, 1600-276 Lisboa, Portugal

\textbf{A R T I C L E  I N F O}

\textbf{Keywords:}
Database
Flood impacts
Risk
Public policies

\textbf{A B S T R A C T}

Estuarine areas are often affected by flood episodes with significant infrastructural and human damages caused by the overlap of different triggering factors. Currently flood risk management practices are subject to increased scrutiny by the public and relevant stakeholders requiring rigorous justification by flood risk managers and careful validation of the technical options and human and financial resources allocated to the management practice. Therefore, flood risk diagnosis through historical sources might constitute an important and effective first approach to public policies validation.

In this paper is presented an estuarine flood damage database based on historical information and discussed as to the extent these types of sources can contribute to improve estuarine flood risk management in the Tagus estuary (Portugal). The paper discusses the methodological findings and limitations and highlights the usefulness of historical information integrating the results into the International Risk Governance Council (IRGC) risk management framework.

\textbf{1. Introduction}

Low lying areas and estuarine margins are often highly populated and have been subject to high risk of flooding, as demonstrated by recent events such as Hurricane Katrina in New Orleans in 2005\cite{49,65} or the storm Xynthia on the French coast in 2010\cite{15,52}, causing significant human loss\cite{48}, direct and indirect economic impacts across sectors\cite{53,64}.

The complexity of flooding process in these areas is related to the overlap of different factors (e.g. high tides, fluvial discharge, storm surge events, wind) that contribute to flood hazard. Moreover, precipitation and urban drainage systems in urbanized margins also influence flood hazard\cite{20,58,75}. The actual trend of urban expansion together with expected climate change effects that might bring a possible increase in sea level pose new challenges in estuarine flood risk assessment\cite{69}.

The importance of historical information registered in newspapers, technical reports, photographs and other published material has been highlighted from a geomorphologic point of view (e.g.\cite{79,95}) to a more applied perspective\cite{21,85} and plays an important role to better understand flood extent, damages and losses giving a valuable perspective of historical evolution about these hazardous processes.

Historical information is often organized in data bases in order to assure consistency, assist queries and analysis and maintain a record track. There is a worldwide multiplicity of disaster databases that use different inclusion criteria. In terms of global databases, the EM-DAT from the Center of Research on Epidemiology of Disasters\cite{26} is the most well-known but NatCatService from MunichRe\cite{67} or De-sinventar databases\cite{55} can also be mentioned. Concerning national databases other examples can be cited such as the Italian AVI project\cite{39}, the Spanish Catalanian\cite{56} or the Swiss flood and landslide damage database collected and maintained by the Swiss Federal Research Institute\cite{42}. More recently a database of coastal flooding in the United Kingdom from 1915 to 2016 was presented\cite{40} integrating a variety of historical sources to document the consequences of flood events around the coastal UK.

In Portugal, until 2010, the information on hydro-geomorphologic disasters at the national level was scarce. However, the work of Quaresma and Zêzere\cite{76} that built a national database of hydro-geomorphologic disasters based on newspaper records for the period of

\textsuperscript{*} Corresponding author at: LNEC - National Civil Engineering Laboratory, Estuary and Coastal Zones Unit, Avenida do Brasil 101, 1700-066 Lisbon, Portugal.
\textit{E-mail addresses:} arilo@campus.ul.pt, arilo@linc.pt (A. Rilo), atavares@ci.uc.pt (A. Tavares), pfreire@linc.pt (P. Freire), pedrosantos@ces.uc.pt (P.P. Santos), zezere@campus.ul.pt (J.L. Zêzere).

http://dx.doi.org/10.1016/j.ijdrr.2017.07.008
Received 30 March 2017; Received in revised form 22 July 2017; Accepted 23 July 2017
Available online 25 July 2017
2212-4209/ © 2017 Elsevier Ltd. All rights reserved.

The usefulness of flood damage information based on newspapers and historical records rely on the fact that these kinds of sources frequently cover more occurrences and events since they can provide news coverage on a local scale [84]. In addition, one can frequently find the same occurrence described in different newspaper sources which can contribute to a more refined and detailed description of the registered damages. Moreover, this kind of source usually covers a wide period of time and presents a continuous record, assuring consistency on the sources [55,85] and is relatively easy to access at a low-cost. For instance the UK Natural Hazards Partnership (http://www.naturalhazardspartnership.org.uk) collected, analyzed and scored impacts for eleven flood events case studies used to validate a hazard model, with information extracted from UK newspapers [27].

The information extracted from those sources is diverse and varies from quantitative information as to human loss (deaths, injuries, missing or displaced people) to the number of affected houses or infra-structure and geographic location, to more qualitative information such as the type of damage (ranging from direct economic losses to indirect losses like traffic interruption and disruption of functions). It is also common to find information regarding the weather/triggering conditions although it may be less precise [104,83].

The richness of this information associated with the widely covered period of the sources allows analysis from a social perception of risk [57] and risk governance strategies [92] as well as extraction of information on severity and probability of flood events, providing valuable information for risk management.

Risk management models articulate process analysis, risk assessment and judgment of territorial impacts, along with the implementation of mitigation and decision actions [103,9]. There is an increasing need for a comprehensive and strategic view of mitigation processes and actions, as well as for the institutional settings and decision mechanisms [50,73]. As Van Asselt and Renn [98] and Renn and Klinke [81] pointed out, the aim is to improve decision-making processes based on social dimensions and institutional capacity, framed by routine/linearity, complexity, uncertainty and ambiguity. In particular, this allows for the dimensioning of institutional and financial resources for the decision-making process, as well as the assessment of available technical and human resources [87,99]. The design of risk response or mitigation measures, the definition of the level of monitoring, together with the knowledge of the costs and benefits associated with each option contributes to improving systems’ responsiveness based on a precautionary approach [9], as well as supporting organizations in resource sharing, local involvement and effectiveness of adaptation measures [18].

Risk matrices are a tool used to manage risk built on the intersection of probability and damage. They can be drawn up from database analysis and are a widely used instrument in risk management, present in different national and international standards and guidelines (e.g. [80,5,47]). They are semi-quantitative tools, belonging to probability vs. consequence diagrams that prioritize risk and present it graphically in order to provide a simplified version of complex concepts to a wider audience [1]. Nevertheless, it is worth mentioning the debate on the advantages and disadvantages of risk matrices with some authors arguing for a more careful use of this instrument [19,23]. Despite the ongoing debate, this tool is widely used from corporate governance to natural hazards management playing an important role in the risk management process, especially for risk information purposes [63].

Flood risk management [51] is a continuous and developing process usually conducted by public authorities that involves a decision making process with options’ identification, assessment and implementation measures along with feedback from the flood risk management practice, including risk communication [101,25], local knowledge [86] and resilience [10]. Currently, citizen and stakeholder scrutiny is growing and becoming more important having a huge influence on the way institutions communicate with the public [17,88]. On the other hand, flood risk managers need to justify and validate options concerning flood risk public policy, including institutional means and human, financial and technical resources allocated to flood risk management practice [89]. Thus, historical information, namely long-term database records are an important starting point for flood risk management diagnosis, allowing risk assessment through risk matrices and being an effective first approach to public policies validation (Fig. 1)

The present work uses a database of historical records (based on a systematic collection of printed newspapers) as a starting point to present an innovative contribution to improve estuarine flood risk management. This is implemented through the definition of impact profiles based on information collected in the above mentioned database and the integration of those profiles into the IRGC risk management framework.

Overall, this paper aims to discuss to what extent the analysis of historical sources can contribute to flood risk management in the Tagus estuary (Portugal). To address this objective two conceptual questions were formulated: (1) is it relevant to know the spatial and temporal evolution of past estuarine flood occurrences? (2) Is historical information able to inform impact profile definition that can be integrated into a risk management framework?

The paper was structured as follows: first, the overall database exploration is presented, including the spatial and temporal evolution of estuarine flood occurrences, the main triggering factors found in the historical sources and the temporal and spatial characterization of damages associated with past estuarine floods. The temporal distribution of flood occurrences allowed the definition of three time periods; each was analyzed and the human losses typology distribution is presented. Second two different tools for risk management were applied: risk matrices (two different standards were used) along with the definition of impact profiles based on cluster analysis of the damage data. Finally the hypotheses are discussed and analyzed as to what extent historical information can contribute to improve flood risk management, inserting the results into the IRGC risk management escalator.

1.1. Study area

1.1.1. Geophysical context of the Tagus estuary

The Tagus estuary is located in the western part of the Iberian Atlantic coast (Fig. 2a) which extends from the Strait of Gibraltar, in the south, to the Spanish-French border in the north. According to Fortunato et al. [32], on the Iberian Atlantic coast, tides are semi-diurnal.

![Fig. 1. Graphic synthesis on the importance of historical information to public policy.](image-url)
and storm surges increase from south to north. Furthermore, the western and northern parts of the coast are affected by a very energetic wave regime that is explained by its direct exposure to the waves coming from the Atlantic Ocean [32].

Water levels in estuaries depend on the relative importance of different triggering factors. In the case of the Tagus estuary, water levels are influenced by the particular estuarine morphology that promotes tidal range amplification by resonance [37]. The estuary has a deep and narrow inlet channel and an extensive, shallow inner domain that corresponds to 43% of the total estuarine surface [71]. The existence of a narrow inlet channel constrains the propagation of ocean waves into the estuary, but the geometry of the inner domain favors the local generation of waves [29].

Storm surges can also contribute to water level increase and have been estimated at 46 and 58 cm in Cascais (estuary mouth) for return periods of 5 and 100 years, respectively [3]. Recent studies in the Tagus estuary [30] estimated water level variations for the 100-year return period of 4.5 m (CD-chart datum:2.08 m below mean sea level (MSL)) in Cascais, at the mouth, to 5.1 m (CD) at its head in Vila Franca de Xira, 50 km upstream from Cascais.

Additionally, fluvial discharge influence on water level is restricted to the upstream areas of the estuary [100], the Tagus and Sorraia rivers being the main sources of fresh water (Fig. 2b). The Tagus estuary is drained by an extensive hydrographic basin of about 80,629 km² [12]. The main Tagus estuary tributaries are the Tagus river and the Sorraia river (Fig. 2b). Over the last century, the Tagus and Sorraia rivers discharges have been largely modified by the construction of a large number of dams (more than 140 including Portuguese and Spanish territories) mainly between 1951 and 1957 [16,6]. Since the beginning of the 20th century, with the 1912 Treaty between Portugal and Spain, the two countries have agreed on the industrial exploitation of transboundary rivers among other issues, but only in 1998 with the Albufeira Convention [12] did the two countries agree to the shared management of the water discharges of dams, especially those located upstream in Spanish territory.

1.1.2. Regional land use and social vulnerability

The Lisbon metropolitan area comprises twelve waterfront municipalities, five in the north margin and seven in the south margin (Fig. 2a). The northern margin is mostly occupied by urban areas (Fig. 2a), whereas natural and semi-natural spaces are predominant in the southern part, along with extensive agricultural areas located between the Tagus and the Sorraia rivers and across the Benavente municipality (Fig. 2a). Additionally, the area is classified and legally protected as a Natural Reserve and a Special Protection Area for birds playing a crucial role in the ecological functions of the Tagus estuary.

The Lisbon metropolitan area plays a strategic role in the economic framework of the country, since the vast majority of governmental agencies are located here along with the largest public and private Portuguese companies and critical infrastructures.

The social and economic exchanges between the two margins are a crucial feature for understanding the metropolitan dynamics. In fact, 26% of Portugal’s total population is concentrated here particularly in the northern margin, namely in the Lisbon, Loures and Oeiras municipalities. The Lisbon municipality alone was responsible for about 26% of the total turnover in Portugal in 2011 [44]. The last census data from 2011 demonstrated there is a high proportion of residents that work or study outside their municipality of residence, justifying the intense...

![Fig. 2. (a) Tagus estuary in the context of the Iberian Atlantic coast; (b) Tagus estuary marginal municipalities and land use map from 2007, level 1 of disaggregation [22].](image-url)
commuting movements between the two margins particularly towards the Lisbon municipality, the national capital, but also between other municipalities [28]. The education level of northern margin stands out with a higher number of the resident population holding a university degree (22.1% compared to 11.9% in the south margin). In addition, the percentage of the resident population with no formal educational level is slightly higher in the southern margin (17.5% compared with 15.8%) [93].

Important commuting movements between the two margins of the estuary are a strong feature of the study area [28]. About 135,000 vehicles per day [45] cross the two existing bridges (25 de Abril and Vasco da Gama) along with more than 40,000 passengers by train [28]. The river is also crossed by 65,000 boat passengers each day [45]. Additionally, commodities transport and tourism are growing sectors. In fact, the Lisbon Port Authority verified an increase of 20% in cruise ship dockings between 2009 and 2013 along with a 13% increase in commodities commerce between 2012 and 2013 [93].

Recently, Tavares et al. [93] studied the territorial vulnerability to flooding of the Tagus estuary margins applying statistical methods to census data and identified the urban context, the family structure and the socio-economic conditions as the main vulnerability drivers. The authors acknowledge that housing characteristics (namely the age of the buildings), education of the individuals (namely age and level) and mobility (namely the proximity to fluvial transport interfaces) are the most important dimensions to explain the territorial flood vulnerability in the Tagus estuary.

1.1.3. Historical evolution of estuarine margins

During the second half of the 19th century, the city of Lisbon was the main center of social and economic activities and the population doubled in less than fifty years, along with extension of urbanized areas driven by the late industrialization period that occurred in Portugal [94]. Also during that period Oeiras, Cascais and Vila Franca de Xira in the northern margin and the Almada and Barreiro regions in the southern margin followed this growth trend [72] which led to unplanned urban areas with deficient sanitary conditions. During this period, the first widely circulated Portuguese newspaper, “Diário de Noticias” was launched and it still exists [90]. At the beginning of the 20th century (Fig. 3), the first harbor infrastructures were built on landfill terrains along the Lisbon waterfront [24].

From the middle of the 20th century onwards, almost all the municipalities grew rapidly in terms of population bringing into existence a continuous urban area between Lisbon and the surrounding regions. This was especially prevalent in the northern margin [72]. The 1950s and 1960s were marked by intense industrialization, with construction of transport infrastructures accompanied by a growing awareness of the need for urban planning guidelines (Fig. 3).

The Tagus estuary now frames the largest metropolitan area in Portugal, containing about 1.6 million inhabitants mainly concentrated in the northern margin. The Lisbon world exhibition in 1998 along with the construction of a 3rd bridge (called Vasco da Gama), linking Lisbon and Alcochete in the southern margin, were responsible for the main territorial developments.

More recently in 2008, an extensive territorial development plan was launched in the southern margin called “Lisbon South Bay” [60] which sought to reclaim about 912 ha of the southern margin waterfront, formerly occupied by one of the largest ship repair companies, Lissave; an important steel company, Siderurgia Nacional, and an industrial chemical company, Quimiparque. The target territory located on the Almada, Seixal and Barreiro waterfronts was environmentally degraded with unoccupied areas. Therefore, the project goals were environmental reclamation with an ecological corridor linking the three municipalities, along with new urban infrastructures dedicated to education, services, health, and sport facilities.

Furthermore, the government’s recent intention to expand the capabilities of Lisbon airport might lead to new airport facilities in Montijo’s estuarine margins, along with a new metropolitan line on Vasco da Gama Bridge which would generate new and intense commuting between the two margins [36].

2. Methods

The conceptual methodology followed in this study is summarized in Fig. 4. The outlined scheme was organized in two central steps that are described in detail below.

Previous studies conducted in Portugal [104] have provided an unprecedented and extensive compilation of information regarding hydro-geomorphologic (i.e., floods - except coastal floods - and
landsides) disasters for the whole country, based on a broad and systematic collection of newspapers covering the period 1865–2010. Learning from this experience, the project DISASTER (http://riskam.ul.pt/disaster/) that gathered all the newspaper articles provided the vast majority of the printed sources used in the scope of this study. For completion, other institutional sources were collected (ANPC- National Authority for Civil Protection; APL - Lisbon Port Authority).

Due to the societal and economic national importance of the Lisbon Metropolitan Area, only national newspapers and national institutional sources were used as data sources (Table 1). From all the consulted printed media sources, a total of 147 newspapers and magazines specimens were used, one geodatabase of flood occurrences provided by the National Authority for Civil Protection (ANPC) and 21 photographs provided by the Lisbon Port Authority (APL) historical archives (Table 1).

The consulted sources (Table 1) were used to build a geodatabase of Tagus flood occurrences considering losses, damages and flood triggering factors, comprising 235 flood occurrences. An occurrence is defined as a geographically defined place described in the consulted sources as affected by estuarine flooding, independent of its severity [85]. An event is defined as a group of occurrences with the same date or identified in the sources’ descriptions as being part of the same episode.

The internal organization of the geodatabase is provided in Fig. 5 and comprises four groups of information.

To assure that the geodatabase would only capture estuarine flood occurrences, the geographic incidence was constrained to the area where estuarine processes prevail, between Oeiras and Vila Franca de Xira, corresponding to the upstream limit of the salt intrusion (Fig. 1), and between the highest astronomical tide line – the upper limit of intertidal domain [82] – and 20 m. above mean sea level but sufficiently near the margin to be clearly associated with estuarine processes. The topographic data was obtained from Portuguese Topographic Maps, series 888, at 1:25,000 scale by the Portuguese Military Center for Spatial Information (IgeoE).

Table 1
Sources used to construct the Tagus estuarine flood geodatabase.

<table>
<thead>
<tr>
<th>Source</th>
<th>Source Typology</th>
<th>Coverage period (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Newspapers and magazines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diário de Notícias</td>
<td>Dail (p)</td>
<td>1864-2013</td>
</tr>
<tr>
<td>Jornal de Notícias</td>
<td>Dail (p)</td>
<td>1888-2013</td>
</tr>
<tr>
<td>O Século</td>
<td>Dail (p)</td>
<td>1880-1978</td>
</tr>
<tr>
<td>Público</td>
<td>Dail (d)</td>
<td>1990-2013</td>
</tr>
<tr>
<td>Correio da Manhã</td>
<td>Dail (d)</td>
<td>1979-2013</td>
</tr>
<tr>
<td>24 Horas</td>
<td>Dail (d)</td>
<td>1998-2010</td>
</tr>
<tr>
<td>Ilustração</td>
<td>Fortnightly (d)</td>
<td>1926-1939</td>
</tr>
<tr>
<td>O SéculoIlustrado</td>
<td>Weekly (d)</td>
<td>1933-1989</td>
</tr>
<tr>
<td><strong>Institutional Sources</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANPC geodatabase</td>
<td></td>
<td>2006-2013</td>
</tr>
<tr>
<td>APL photographs</td>
<td></td>
<td>1941-2010</td>
</tr>
</tbody>
</table>

(p) - printed source; (d) - digital sources consulted individually for validating and cross-checking purposes.

ANPC - National Authority for Civil Protection; APL - Lisbon Port Authority.

The methodology for content analysis and extraction of information from the newspapers was based on Berelson [11], GAO [34] and Krippendorf [54], methods that focus on the conceptual and relational elements present in the documents.

The approach followed to introduce newspaper information into the geodatabase started with extensive reading of each individual newspaper, then extraction and compiling the information in a spreadsheet based on geographic coordinates, for each occurrence and later introduced in a geodatabase using ArcGIS 10.1 software. Moreover for each occurrence the geodatabase fields (Fig. 5) were filled with the information retrieved from the consulted sources (Table 1).

The geographical locations extracted from the information given by the sources were geo-referenced in ETRS89 PT TM06 geographic coordinate systems, using a toponymical database supplied by ANPC and a geographical database of municipalities and civil parishes by Statistics Portugal (INE). In the cases of old street names and locations, a series of online data sources were accessed (e.g. [41,59]). When convenient a collection of orthophotographs from 2007 provided by the Directorate-General for the Territory (DGT), with 0.50 m. of spatial resolution and RGB spectral resolution were used for validation purposes.

To extract the relevant occurrences from the two institutional sources consulted (Table 1) the ANPC geodatabase was filtered by location and by triggering factors using the ANPC internal manual for occurrence classification. Furthermore, the information extracted from the photographs provided by APL was incorporated through cross-checking with other online sources (e.g. [41,59]).

To assure data integration consistency a set of rules were followed, namely, that the database would be filled in by one person only, who
consulted all the sources and filled in the geodatabase fields (Table 1 and Fig. 5). Information was gathered and compared to acquire more refined details in the cases where the same event was recorded in more than one data source. When the source was considered vague or imprecise in terms of damages and flood processes the occurrence was not incorporated (for instance when is referred a flood in a neighborhood but no other information is given).

Concerning database exploration a simple spatial and temporal analysis was done, filtering the 235 occurrences by location (municipality) and representing the geographic distribution of estuarine flood occurrences as well as their distribution over time. The classification of occurrences into events allowed a simple probabilistic analysis using the Poisson distribution function since each event was considered independent [85].

The temporal distribution of flood occurrences was graphically represented and allowed the empirical visualization of three distinct time periods between 1865 and 2013. The exact time intervals were achieved through the cumulative curve obtained using the relative frequencies for 10 year periods between1865 and 2013, where time intervals correspond to the points that marked a slope change of the curve. These three time intervals were used to perform a temporal risk analysis through the application of risk matrices as risk assessment tools. Two distinct standards were followed namely the AS/NZS 4360:2004 Risk Management Guidelines [8] and the HB205-2004 OHS Risk Management Handbook [7].

The first standard provides a level of risk through the combination of likelihood and consequence (or severity). Occurrence likelihood was calculated for each time period and classified according to the standard classification. The AS/NZS 4360:2004 Risk Management Guidelines consider seven levels of likelihood: (A-almost certain; B- likely; C-possible; D-unlikely; E-rare; F-very rare; G-almost incredible) and the respective indicative frequency (A-once a year or more frequently; B-once every three years; C-once every ten years; D-once every thirty years; E-once every 100 years; F-once in 1000 years; G-once in 10,000 years). The severity level for each time period was classified using the health and safety type of consequences established by the standard, considering the different types of human injuries reported in the Tagus estuary database. The standard establishes five levels of severity (I-no medical treatment required; II-objective but reversible disability requiring hospitalization; III-moderate irreversible disability or impairment (< 30%) to one or more persons; IV-single fatality and/or severe irreversible disability (> 30%) to one or more persons; V-multiple fatalities, or significant irreversible effects to > 50 persons). The overall risk level is obtained using the risk classification matrix of the standard.

The second standard provides a level of risk, called risk rating, based on the overlap of three descriptors – consequences (C), exposition (E) and likelihood (L). The HB205-2004 OHS Risk Management Handbook considers six levels of consequences rated between 1 and 100: 1-notable (with small cuts and bruises; small environmental effects in public health and non-permanent effects in the environment); 2-infrequent (from once a month to once a year); 3-occasional (once a week to once a month); 6-frequent (almost daily); 10-continuous (various occurrences in one day).

Finally, in what concerns likelihood (L) the standard stipulates six levels ranging from 0.1 to 10: 0.1-almost impossible (unknown occurrence) with a residual incidence (1/10000); 0.5-conceivable (very low incidence (1/1000)); 1-remote (low incidence (1/100)); 3-possible but not common; 6-probable (common occurrence with a probability of 50%); 10-almost certain.

For the different periods of time each descriptor was assessed and the final risk score was obtained by equation 1.

\[
\text{Risk score} = C \times E \times L
\]

The level of risk (risk rating) for each result of equation 1 is achieved through reference to the standard classification table which establishes four risk ratings accordingly: for a risk score higher than 600 the risk level is classified as Very High; for a risk score between 300 and 599 the risk level is classified as High; for a risk score between 90 and 299 the risk level is classified as Moderate and for a risk score lower than 90 the risk level is classified as Low.

To evaluate the relational degree between the different damage typologies the data was processed using cluster analysis. This is an exploratory statistical technique of multivariate analysis that allows grouping variables into homogeneous groups based on common characteristics of those variables [62]. The geodatabase group of information b) “type of impacts” was used to do a statistical exploration using cluster analysis techniques to assess the degree of relation between the different types of impacts. For the classification a hierarchical cluster analysis was used considering 3 clusters using the Ward method and phi coefficient as a similarity measure. The procedure was done using SPSS Statistics® software.

3. Database exploration

3.1. Spatial and temporal occurrences

The acquired data comprise 235 occurrences corresponding to 44 events over a period of 148 years, between 1865 and 2013 that affected 10 municipalities in the Lisbon metropolitan area between Oeiras (downstream) and Vila Franca de Xira (upstream) (Fig. 6). Of the 235 occurrences, 54% were registered in the Lisbon municipality followed by 18% in Vila Franca de Xira and 13% in the Oeiras municipality. Fig. 6 demonstrates the contrast between the two margins regarding occurrence distribution, with the Montijo municipality registering only one occurrence.

The temporal distribution of flood occurrences and events is represented in Fig. 7 and reveal that the period from 1900 to almost 1970 is characterized by an increase in flood occurrences and events. The decades of the 1930’s and 1940’s showed the highest number of occurrences registered during the study period.

3.2. Triggering factors

The relative percentage of each forcing factor by municipality is presented in Fig. 8 and reveals that in the majority (exceptions are Alcochete and Montijo) of the Tagus estuary municipalities estuarine flood processes are associated with the simultaneous action of multiple triggering factors. However, these assumptions should be regarded carefully since in the cases of the Alcochete, Montijo and Seixal municipalities, only a limited number of occurrences were registered.

The figure indicates that rainfall has been cited by the historical sources as the most relevant triggering factor, in all estuarine municipalities along with the tide influence, although the latter has less importance when compared to the rainfall. Also, other anthropogenic factors are worth noting (destruction of flood protection walls and dykes) that are particularly relevant for the upstream municipality of Vila Franca de Xira.

3.3. Spatial and temporal impact

In order to reveal which impacts are relevant in Tagus estuary floods, an impact analysis was made through the exploration of direct and tangible damages (infrastructures impacts and reported economic
losses), intangible damages (human losses and environmental deterioration) and indirect and tangible records (traffic interruption and function disruption).

Fig. 9 shows the percentage of each type of impact by municipality. Herein, the percentages have to be interpreted considering the total number of occurrences with registered impacts by each municipality. The analysis revealed that out of the 235 occurrences registered in the database only 15 do not have any associated impact. The majority of impact typologies are widely distributed among Tagus municipalities with exception of the Montijo municipality registering only infrastructure-related impacts (Fig. 9).

Infrastructure impacts (physical damages on built infrastructures) are widely distributed among all the Tagus estuary municipalities (Fig. 9). Economic losses (only direct economic losses were considered) were found in five of the ten municipalities (Alcochete, Lisboa, Moita, Oeiras and Vila Franca de Xira) and human losses were present in seven of ten municipalities (Almada, Barreiro, Lisboa, Loures, Moita, Oeiras and Vila Franca de Xira). Environmental degradation was registered in eight municipalities (Alcochete, Almada, Barreiro, Lisboa, Loures, Moita, Oeiras and Vila Franca de Xira) and institutional involvement, traffic interruption and functions disruption were registered in nine of the ten municipalities (Fig. 9). At a glance, Fig. 9 reveals that historically, Tagus estuary flood damages are diverse and were registered at the vast majority of the municipalities with the exception of Montijo where only one typology of damages is associated.

Human losses were described by typology (dead, injured, evacuated, missing and displaced persons) and by municipality (Fig. 10). None of the ten municipalities registered any record of missing people and Alcochete, Seixal and Montijo did not register any record of human losses. Lisbon and Vila Franca de Xira are the municipalities that stand out with the largest number of recorded human losses. The other municipalities have a residual number of records with the evacuated and displaced being the most common features. Database records also revealed that the number of occurrences with registered evacuees are relevant in Lisbon (17 records), Vila Franca de Xira (15 records) and Oeiras (4 records) along with the death records that account for 2 in...
Lisbon, 3 in Oeiras, and 8 in Vila Franca de Xira.

Overall, Fig. 10 reveals the historical relevance of evacuees as the preponderant human loss typology in the Tagus estuary, since they occurred in five of the ten municipalities and have the highest number of records.

Fig. 11 illustrates the historical overview of the total number of occurrences per year and the number of occurrences with registered human casualties per year between 1865 and 2013. Between the early and the middle part of the 20th century there were a higher number of occurrences registered which led to the definition of three distinct time periods (Table 2).

The first period (1865–1916) registered the lowest number of occurrences and only two typologies of human losses are found (deaths and evacuated). In the second period (1917–1962) there is a significant increase in the number of flood occurrences along with human losses typology diversification. More recently (1963–2013) the number of occurrences and related human losses typologies decreased significantly. Generally, the Tagus estuary flood history reveals a distinct prevalence of flood occurrences and related human consequences over the considered time periods.

4. Tools for risk management

4.1. Risk matrices

Risk matrices provide a broad view of risk assessment synthesizing the evolution of estuarine flood risk over time. Table 3 presents the risk classification obtained according to the described methodologies ([8,7]) for each considered time period. Despite the different indicators and assumptions considered by each standard the results are similar and indicate that the period 1917–1962 registered the highest risk within the entire geodatabase time frame. The oldest and the more recent time periods have similar risk classifications.

4.2. Impact profiles

Damage types were assessed through pre-defined groups of damage typologies that were considered in the database. The typologies range from physical damages on built structures to human losses or traffic disruption. We evaluated the degree of relation between the different typologies using cluster analysis. Fig. 12 represents the operational steps used to establish the three clusters.

These clusters express the relationship between the different types of damages. Considering that the information is obtained from newspaper analysis, it is possible to assume that the results reflect a social and political context that those types of sources are frequently embedded in. Therefore, each cluster is understood as a unique and individual impact profile that might be interpreted in the context of Tagus estuary flood risk evaluation (Table 4).

Impact profile A (operational impacts) is related to episodes with small physical damages and no significant human or economic losses. Typically, this corresponds to episodes where there is traffic or public transport interruption with associated functions’ disruption. Profile B concerns infrastructure impacts and is associated with episodes where the destruction of built structures is the relevant feature, sometimes with economic damages mentioned in newspapers but with no significant human losses. Finally, profile C is associated with major public impacts and is linked to episodes where significant human and economic losses are the most relevant feature, but where other types of damages also appear due to the importance of the episode. Hence, damage profiles constitute a snapshot of estuarine flood damage typologies giving valuable insights on flood risk management policies.
5. Discussion

5.1. Historical information appraisal

In agreement with other studies (e.g., [78,104]85), the present work shows that databases built upon newspaper reports have the potential to give valuable information on historical flood occurrences. However, the results have some limitations since it is recognized that newspaper reports carry a clear bias (e.g., [91]). As already recognized by other authors [14,38] these types of sources usually give excessive media coverage to events with significant societal consequences (namely with human causalities or significant social disruption) which might lead to an underrepresentation of the real number of events, compromising the completeness of the database. Another acknowledged pitfall is a frequent lack of scientific accuracy in the reported information [43]. Relevant to the case of the Tagus estuary, it was also recognized that newspapers due to their intrinsic characteristics, reflect journalistic criteria and the socio-political importance relative to certain municipalities when compared to others which might explain the larger percentage of occurrences registered in the Lisbon municipality.

Another source of bias is related to the methodology used to perform content analysis and information extraction in media sources. There are two main types of content analysis: the one performed by human coders executed by people, and computer aided text analysis (also designated as CATA), which involves the use of software to execute analyses [70]. Although it is recognized that the latter type favors reliability and comparability, the two types are equally effective when compared [66]. Furthermore, despite all efforts in algorithm development it is acknowledged that CATA is not able to interpret subjective messages as human coders can [54]. For instance, during this work on flood extent was frequently encountered like this example: "só se viam os dois últimos degraus da escadaria do Teatro Nacional" – only the last two steps of the stairs were visible in the front of the National Theatre (Jornal de Noticias 25/10/1954). This type of sentence needs to be interpreted in order to be translated into useful information, only possible if interpreted by human coders. Following the best practice [54] a set of previously established rules of content analysis were followed to improve accuracy and reliability.

In order to extract the maximum amount of information possible from the sources, a field was added in the database structure to capture the information about triggering factors (Fig. 5). Although efforts were made to incorporate accuracy, the limitations of the data sources in describing meteorological and hydrodynamic conditions have to be acknowledged. In fact, this has already been noticed by others [43]. The lack of accuracy and detail in this particular type of information reveals that media sources (specifically newspapers) are not the most adequate base for study and analysis of triggering factors of natural hazards, especially if the hazard is a result of compound triggers such as estuarine floods. Nevertheless, the data presented in Fig. 8 was able to capture the estuarine system complexity, demonstrating that flood events are associated with the simultaneous presence of multiple forcing factors with rainfall being particularly relevant.

Table 2

<table>
<thead>
<tr>
<th>Time periods</th>
<th>Number of years</th>
<th>Number of occurrences</th>
<th>Human losses typology</th>
<th>Deaths</th>
<th>Missing</th>
<th>Injured</th>
<th>Evacuated</th>
<th>Displaced</th>
</tr>
</thead>
<tbody>
<tr>
<td>1865 – 1916</td>
<td>51</td>
<td>31</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1917 – 1962</td>
<td>45</td>
<td>141</td>
<td>33</td>
<td>0</td>
<td>24</td>
<td>1021</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>1963 – 2013</td>
<td>50</td>
<td>63</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>408</td>
<td>131</td>
<td></td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C E L Risk score</td>
<td>Likelihood label</td>
</tr>
<tr>
<td>1865 – 1916</td>
<td>25 6 3 450  High</td>
<td>B   III  High</td>
</tr>
<tr>
<td>1917 – 1962</td>
<td>50 6 3 900  Very high</td>
<td>A       V  Very high</td>
</tr>
<tr>
<td>1963 – 2013</td>
<td>25 6 3 450  High</td>
<td>A       III  High</td>
</tr>
</tbody>
</table>

Fig. 11. Temporal distribution of total number of occurrences per year along with the number of occurrences with registered human casualties.
Cluster analysis and impact profiles definition.

<table>
<thead>
<tr>
<th>Impact profile</th>
<th>Types of impacts</th>
<th>Cluster name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Traffic interruption</td>
<td>Operational impacts</td>
</tr>
<tr>
<td></td>
<td>Functions disruption</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Physical impacts on built infrastructures</td>
<td>Infrastructures impacts</td>
</tr>
<tr>
<td></td>
<td>Economic losses</td>
<td>Major public impacts</td>
</tr>
<tr>
<td></td>
<td>Human losses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Environmental degradation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Institutional involvement</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 12. SPSS cluster analysis dendogram displaying three clusters association.

In previous work, Ramos and Reis [77] and Trigo et al. [96] associated important flood episodes in the Tagus basin with two different meteorological situations, a) relatively long lasting rainfall periods in winter months, and b) the existence of a strong, regional convection activity and a low pressure system that triggers intense and short term rainfall episodes. In the estuary, particularly in urban areas, the coincidence of heavy rainfall episodes with high tide, along with the limited capacity of the sewer system can cause severe urban floods [13] as observed along the estuary margins [97]. The results of database analysis (Fig. 8) indicate the relevance of the fluvial discharges in the flood episodes in the upstream areas, especially in the Vila Franca de Xira municipality, which is in agreement with the results of a previous study that analyzed the influence of Tagus discharge on the estuarine water level [100].

Fig. 8 also reveals the importance of other anthropogenic factors in the Vila Franca de Xira municipality. The newspapers indicated the destruction of flood protection walls in the agriculture lands namely on the three islands formed by alluvial deposition (Fig. 2b) and in the extensive agricultural areas located between the Tagus and the Sorraia rivers (called Lezírias) as a triggering factor in the above mentioned municipality. This is consistent with area development since 1836 in establishment of a large farmstead company (called Companhia das Lezírias) that built and maintained flood protection walls and dykes along the margins to prevent salt water from entering these fertile lands. In spite of this, other studies revealed that this area has been subject to flood events during the last century [2,33,61].

As already acknowledge by Grantham and Vieira [35] newspapers and content analysis of them have been a primary source of environmental information playing an important role in risk communication and political decisions. The present work brings to light the fact that in spite of the various sources of bias, the extended analysis of newspaper stories made possible the acquisition of valuable knowledge regarding the dimension of damages/consequences of flood risk. In fact, the richness of detail those newspapers provide on damages quantification and description is significant as demonstrated in the Figs. 9, 11 and Table 2.

5.2. Spatial and temporal evolution of past estuarine flood occurrences

The historical perspective given by this study revealed the relevance of geographical context since Lisbon, followed by the Oeiras and Vila Franca de Xira municipalities account for the vast majority of the registered occurrences. Furthermore, the contrast between the two margins is remarkable, the northern margin being the most affected by estuarine floods (Fig. 6). This circumstance reflects the socio-political importance, replicated in the data sources that Lisbon and the northern margin municipalities had over time when compared to the southern margin municipalities, as previously described in Section 1.1.3.

The temporal distribution of flood occurrences allowed the definition of three time periods that differ in number of occurrences and simultaneously on the features of human losses. The first period (1865–1916) presented the lowest number of occurrences and associated human losses. This period corresponded to the beginning of the industrialization and transport infrastructure planning in the Lisbon metropolitan area and the rest of the country. The population was thus connected to Lisbon and in relatively small urban areas, poorly connected with the capital.

The second period (1917–1963) presented the highest number of flood occurrences along with the diversification of the human losses typology. This period corresponds to the consolidation of infrastructure development and the acceleration of urban expansion, along with population growth on both margins of the estuary. In the Lisbon metropolitan area, the construction of landfills to install the Lisbon harbor facilities was especially relevant, changing dramatically the waterfront landscape [24] along with marginal roads, tram and railway lines and the increase of fluvial connections between the two margins of the Tagus estuary. These historical occurrences might help to explain the higher number of human losses, when compared to the other two periods. Regarding the exceptional number of flood occurrences registered, Zézere et al. [104] computed the mean annual precipitation per decade in Lisbon during the 20th century and concluded that the decades between 1930 and 1960 presented an increased trend in mean
annual precipitation. This fact, associated with a larger urban area with limited drainage capacity might also help to explain the highest number of flood occurrences.

More recently (1964–2013), a decrease in flood occurrences and associated human losses was registered. The end of the 20th century and the beginning of the 21st century corresponded to a relative stagnation of new infrastructure development and to a decrease in urban expansion projects. The exceptions were the World Exhibition in 1998 that transformed a large degraded area on the northern margin into a new urban area, and more recently, the so-called “Lisbon South Bay” project in the southern margin. Concerning triggering factors, Zêzere et al. [104] verified a decrease of mean annual precipitation between the end of the 1960s and the 1980s followed by an increase until the decade following 2000. This recent decreasing trend in flood occurrences and associated impacts might also be associated with the reinforcement of flood protection measures in the upstream area of the estuary (Lezírias) and is consistent with the past knowledge of the system [30].

Over time and regardless of the number of occurrences, there is a background record of human casualties, as illustrated in Fig. 11. Also worth noting is the importance of evacuees and displaced people when compared to the number of dead, injured and missing (Fig. 10). The results illustrate that typology of human losses is dependent on the land use, and thus urban and populated municipalities registered a higher number of human losses (Fig. 10). This fact might be a crucial consideration in flood risk management strategies for this area, especially with recent trends towards reconversion of degraded industrial areas into new urban centers along the Tagus estuarine margins. Despite the information constraints discussed in the previous section, the overall knowledge on past estuarine flood occurrences and related impacts provides a strong base to inform flood risk management plans that usually use historical data to validate options and models.

5.3. The use of historical information to define impact profiles

In the framework of this study two risk management tools based on historical information were applied: risk matrices and impact profiles. The application of risk matrices to three time periods revealed that risk level is high in the overall period of analysis increasing to very high between 1917 and 1963. This might be explained not only by a consolidation of infrastructure development and population increase, but also because during this period the Tagus estuary was hit by two major events recorded in the database. Namely the storm of January 27th, 1937 [74] and the February 15th, 1941 windstorm [31,33,68] that contributed to 31 of the 33 deaths registered during that period, along with 14 of the 24 registered injured and 293 of the 1021 registered evacuees (Table 2). The November 25th, 1967 event [97] is acknowledged as the more recent and significant flood event represented in the third period of analysis (1963–2013) with an overall of 12 occurrences in the database. The event was responsible for the 7 deaths registered during that period along with 155 of the total 408 evacuees (Table 2).

The analysis of historical information allowed a comprehensive view of past Tagus estuary flood episodes, giving at the same time a snapshot of flood risk assessment. As Van Asselt and Renn [98] or Renn and Klinke [81] point out, the enhancement of decision making processes based on the correct dimension of technical, human, institutional and financial resources available for those decisions is crucial to advance flood risk mitigation. Therefore, the impact profiles outlined in Table 4 represent a conceptual tool for Tagus estuary flood risk management, since they present three different risk problems whose analysis might give valuable insights on future resource allocation to face estuarine flood episodes. The three outlined profiles can be situated in the IRGC risk management framework [46] (Fig. 13), giving a comprehensive view on the type of risk problems faced in the Tagus estuary along with the set of actors involved and type of resources needed.

Profile A (Fig. 13) exemplifies damages related to traffic interruption and functions disruption caused by flood episodes. These types of damages are mainly operational and require minimal institutional staff intervention (only municipal and national civil protection agents) who configure a simple risk problem with a straightforward answer. Typically, this profile stands for routine interventions with a small amount of resources allocated by public authorities. Nevertheless, given the intense commuting verified between the two Tagus estuary margins, it is reasonable to expect that public and private transport companies might be the most affected.

Profile B (Fig. 13) typifies essentially infrastructure damages but with no human or economic losses described in the sources. Usually, infrastructure damages signify damages to small business like water-front cafes, public leisure facilities, harbor facilities or even private houses. Considering the IRGC risk management framework these types of damages configure a complex risk problem requiring scientific simulation, and therefore, civil protection agents are insufficient to address the challenge. Thus, external experts from national institutions from different sectors should be called to share their insight.

Profile C (Fig. 13) illustrates major public impacts due to significant and relevant flood episodes with human and economic losses, environmental degradation and consequently, institutional involvement. The damages are associated with human losses (dead, injured, displaced and evacuees) along with economic losses whose relevance is reflected by the fact that, in the reports, the money estimated lost due to the flood is frequently recorded. Environmental degradation usually corresponds to large amounts of waste and wreckage left by the flood. Institutional involvement is thus a transversal feature often mentioned in the sources specially when there are human and economic losses. Sometimes traffic interruption can be deduced from the sources but is not described and thus does not appear associated with the profile.

This impact profile embodies an uncertainty induced risk problem and therefore requires a cost-benefit analysis with the crucial involvement of other relevant stakeholders, like the state-owned agro-forestry farmstead that owns the northeast Lisbon agriculture areas, civil parish councils, residents’ associations or public transport users’ associations, besides civil protection agents and external experts. The February 1941 storm or, more recently, the November 1969 floods can be considered examples of episodes fitting into this profile.

A noteworthy outcome of this study is the recognition that past records do not reveal an ambiguous risk problem. In fact, although some relevant and destructive flood episodes took place in the Tagus estuary, none of the profiles show a severity of records that configure disruption lacking a debate of conflicting views about flood risk or even flood risk protection measures which is in line with what is known about past flood records in the area.

6. Conclusions

The study discussed the usefulness and relevance of information from historic sources particularly concerning damage assessment along with an appraisal of constrains related to the use of these types of sources and methods. An analytical overview of the spatial and temporal distribution of Tagus estuary flood occurrences was done and demonstrated that in spite of historical information constrains, this knowledge is useful to inform future flood risk management plans, since those instruments usually validate options based on the historical record. Finally, based on historical damage information, impact profiles were outlined that were placed into the IRGC risk management escalator, giving a complete portrait of Tagus estuary flood risk management options.

The appraisal of historical information made in this study indicated that in spite of the intrinsic constrains of the sources and methods used newspapers, are still useful to extract damage data that can be a very valuable asset for improved flood risk management strategies. Moreover the easy, almost free access to a systematic source like this with such a wide cover period should be acknowledged. However, it is
worth noting that there is a lack of a systematic standard for damage data extraction from this type of source as well as from other types of sources that register losses and damages and without a common ontology for the terms used to describe damage types. These are issues that call for further research.

However, given the absence of spatial impact profiles and taking into account the discussed constraints on the sources and methods used, these results are not suitable for direct use by operational actors or to inform future flood risk assessments considering climate change scenarios. Nevertheless, the conceptual development made in the framework of this study contributes an overview of the different risk problems, including the relevant actors and the type of answers that inform the relevant stakeholders about the necessary resources for emergency and mitigation planning, thus contributing to avoidance of over-sizing or downsizing of measures and resources.

From a more theoretical point of view the study contributes to policy learning practices as presented by Voss and Wagner [102], who recognize that small scale events create valuable opportunities to put into practice learning processes that will allow coping with larger damaging events in the future, further demonstrating the richness of historical sources on flood damages magnitude that provide a valuable set of information to flood risk managers. The necessity of justification and validation of flood risk public policy options, namely the institutional means and resources allocated to the risk management practice is inevitability growing. Thus, historical information based on long-term database records can establish an important basis for diagnosis and management of flood risk.

Acknowledgements

This work was supported by the projects FORLAND – Hydrogeomorphologic risk in Portugal: driving forces and application for land use planning [PTDC/ATPGEO/1660/2014] and MOLINES - Modeling floods in estuaries. From the hazard to the critical management [PTDC/AAG-MAA/2811/2012] both funded by the Portuguese Foundation for Science and Technology (FCT), Portugal. The work was also supported by European Union’s H2020 Programme (BINGO, Grant Agreement; 641739). The first author benefits from a doctoral grant funded by the Portuguese Foundation for Science and Technology (FCT) – (SFRH/BD/111166/2015).

The authors thank the data provided by the project DISASTER (PTDC/CS-GEO/103231/2008) also funded by FCT and the following institutions: Administração do Porto de Lisboa (APL), Autoridade Nacional de Proteção Civil (ANPC).

The authors would also like to thank the three anonymous reviewers for their valuable comments which helped improve the quality of this paper.

References
