ORIGINAL PAPER

Exploring available landslide inventories for susceptibility analysis in Gipuzkoa province (Spain)

Txomin Bornaetxea[1] · **Juan Remondo**[2] · **Jaime Bonachea**[2] · **Pablo Valenzuela**[2]

Abstract

Similar to many mountainous regions of the world, landslides are a recurrent geological hazard in the Gipuzkoa province (northern Spain) that commonly cause damage to communication infrastructure, such as roads and railways. This geomorphological process also threatens buildings and human beings, albeit to a lesser degree. Over time, different institutions and academic research groups have individually collected crucial information on historic and ancient landslides in this region, resulting in various landslide inventories. However, these inventories have not been collectively assessed, and their suitability for landslide susceptibility modelling projects has often been assumed without comprehensive evaluation. In this study, we propose a simplified method to explore, describe, and compare the various landslide inventories in a specific study area to assess their suitability for landslide susceptibility modelling. Additionally, we present the results of an illustrative experiment that demonstrates the direct effect of using different inventories in landslide susceptibility modelling through a data-driven approach. We found that out of the five digitally available inventories in the study area, only three provide sufficient guarantees to be used as input data for susceptibility modelling. Furthermore, we observed that each individual inventory exhibited inherent biases, which directly influenced the resulting susceptibility map. We believe that our proposed methods can be easily replicated in other study areas where multiple landslide inventory sources exist, and that our work will induce other researchers to conduct preliminary assessments of their inventories as a critical step prior to any landslide susceptibility modelling project.

Keywords Landslide inventory · Landslide susceptibility model · Gipuzkoa province · Data availability

✉ Txomin Bornaetxea
txomin.bornaetxea@ehu.eus

1 Departamento de Geología, Universidad del País Vasco (UPV/EHU), Barrio Sarriena s/n, 48940 Leioa, Spain

2 Departamento de Ciencias de la Tierra y Física de la Materia Condensada, Universidad de Cantabria, Avda. de los Castros 48, 39005 Santander, Spain

1 Introduction

Landslide assessment has been a recurrent research topic since the early 1980s (Caine 1980; Evans 1982; Cruden and Varnes 1996; Cendrero and Dramis 1996; Schuster et al. 2002; Hungr 2007; Petley 2012; Froude and Petley 2018), with the majority of studies using landslide inventories as baseline data. Spatial susceptibility modelling research is currently one of the most rapidly expanding fields within the realm of landslide studies (Fell et al. 2008; Guzzetti et al. 2012; Yong et al. 2022), for which landslide inventories are essential too. In recent decades, significant advancements have been made in the field of spatial data generation and computational processing (Maguire and Longley 2005; Moraga and Baker 2022). These advancements have also extended to the domain of automatic and semi-automatic landslide feature recognition (Mondini et al. 2021). As a result, there is now a wider range of approaches available for generating landslide susceptibility maps. Recently published reviews have cited up to 3000 articles on the topic of landslide susceptibility modelling. Furthermore, there has been a huge growth in the rate of publication, with 517 articles solely dedicated to landslide susceptibility modelling released in 2020 (Reichenbach et al. 2018; Yong et al. 2022).

Due to the extensive literature available, numerous methodologies have been developed to study landslide susceptibility over time. These methodologies encompass a variety of approaches, including knowledge-based methods, statistically- or machine-learning-based techniques, and physically-based approaches. Knowledge-based methods include expert evaluation of the factors that led to slope failures, and they frequently include numerical analysis matrices like in the analytic hierarchical process (Pourghasemi et al. 2012; Kayastha et al. 2013; Jun et al. 2017). These methods are highly dependent on the researcher's expertise, and they are often poorly reproducible (Corominas et al. 2014). Statistical and machine-learning approaches—also called data-driven methods—are based on the landslides that already occurred in a study area, and whose location are already known. Using Geographical Information System (GIS) tools, the spatial features of landslide occurrence locations are coded and mathematically processed in order to find a model that best fits what the landslide inventory represents (Bovis and Jakob 1999; Tien Bui et al. 2016; Kang et al. 2017; Chmiel et al. 2021). These techniques have the advantage of being easily reproducible and applicable to large areas, although they are completely dependent on the quality and completeness of the input data (Steger et al. 2017). Physically-based approaches, in contrast, aim to simulate the comprehensive behavior of slopes through numerical modelling. These approaches rely on specific slope conditions, including factors such as ground water conditions and the presence or absence of structural reinforcements (Maleki et al. 2022; Maleki and Mir Mohammad Hosseini 2022). However, a major scientific challenge associated with physically-based models is the need for high-quality and abundant input data (Hutter et al. 1994; Gariano and Guzzetti 2016; Park et al. 2019; Palacio Cordoba et al. 2020; Marin et al. 2021).

In the case of data-driven methods, the information derived from landslide inventories, which is utilized to train and validate the susceptibility models, is frequently accepted at face value. This is probably due the limited availability of data regarding past landslide occurrences. However, landslide databases often show systematic bias related to the spatiotemporal scale, the considered data sources, and the representation criteria, affecting to their completeness and representativeness (Valenzuela et al. 2017). In fact, some authors highlight the relevant role that landslide inventories play as baselines to calibrate data-driven models (Steger et al. 2017; Jacobs et al. 2018; Bornaetxea et al. 2022), demanding

to pay more attention to the quality and accuracy of landslide databases prior to their use as input data.

Literature review revealed some good examples about landslide inventory assessments by comparing various data sources for the same study area. For instance, Brardinoni et al. (2003) compared the information provided by a photo-interpretation-based inventory and a field-work-based inventory obtained for a densely forested river basin. Their findings showcased a significant loss of information if only the photo-interpretation-based inventory had been relied upon. Bernat Gazibara et al. (2019) also analyzed the differences of three inventories based on landslide statistics, frequency-area distribution, geographical discrepancy of landslides and landslide density maps, which allowed the authors to demonstrate the unreliability of the existing historical inventories. Also, other authors conducted similar studies in larger regions of central and south Italy (Galli et al. 2008; Pellicani and Spilotro 2015). Notwithstanding, all these works only examined the mutual differences among the existing inventories, and the specific result of using each of them to model landslide susceptibility remains unknown. To date, only Pokharel et al. (2021) studied the direct effect of using different inventories to model the landslide susceptibility in the same area. For that study, the authors used five landslide inventories undertaken by different researchers after an earthquake event in central Nepal Himalayas.

1.1 Purpose of study

Landslide data collection in the Gipuzkoa province (northern Spain, Fig. 1) has been ongoing since the 1980s. Over the years, various attempts have been made to develop regional and local landslide susceptibility models (see Sect. 3). However, it is important to note that each susceptibility model created thus far has been based on different landslide inventories, and their accuracy and representativeness have not been thoroughly examined. Hence, this study aims to conduct an exploratory analysis of the available data on landslide locations in the Gipuzkoa province. The goal is to use these data in controlled statistical simulations to assess the implications of their use in landslide susceptibility mapping.
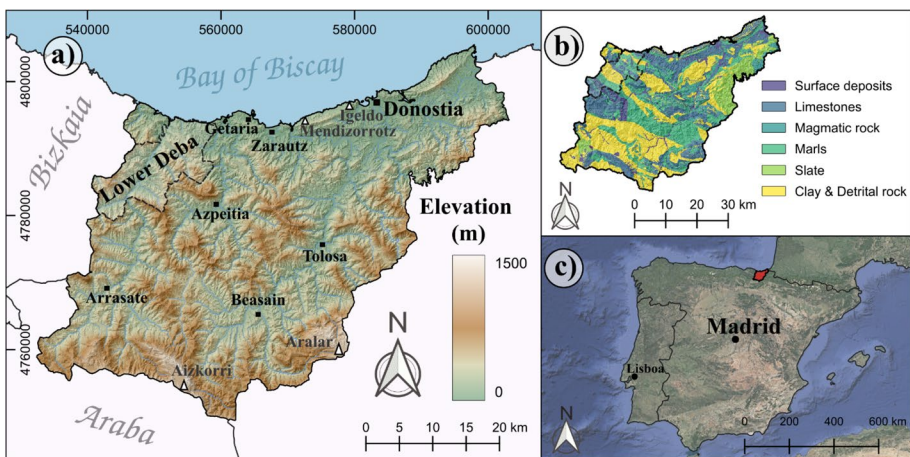

Fig. 1 Gipuzkoa province and Lower Deba area locations; **a** elevation map; **b** lithology map; **c** general location view (base map retrieved from https://www.google.com/maps/)

The specific objectives of this study are twofold: (1) to explore, describe, and compare the various landslide inventories utilized in the study area for generating susceptibility and landslide hazard maps, and (2) to visually illustrate the direct implications of employing different landslide inventories in relation to the final susceptibility map.

Unlike the approach proposed by Pokharel et al. (2021), this study presents comparison and assessment methods specifically tailored to a rainfall-induced landslide environment. These methods are designed to be applicable to various historical or geomorphological landslide data collections. To the best of our knowledge, this is the first direct comparison of landside inventories conducted in the Gipuzkoa province. The insights gained from this research will be valuable for future landslide data collection and susceptibility modelling projects in the area. Moreover, the experiments carried out in this study can be easily replicated in similar regions within Europe or worldwide. This will enable other researchers to assess the suitability of their landslide databases before using them for susceptibility modelling.

The paper is structured as follows: Sect. 2 provides a detailed description of the study area. Section 3 introduces the existing landslide inventories available in the study area. In Sect. 4, we outline the specific data employed to implement the methods presented in Sect. 5. The results obtained from the analysis are presented in Sect. 6, followed by a comprehensive discussion in Sect. 7. Finally, Sect. 8 summarizes the key conclusions drawn from this study.

2 Study area

The Gipuzkoa province (with an extension of approximately 2000 km^2) is located in the northern Iberian Peninsula, along the western end of the Pyrenees (Fig. 1), and represents the typical hilly and mountainous Atlantic landscape. This area presents generalized persistent precipitation values (average: 1597 mm/year) and two main rainy periods: November–January and April (Remondo et al. 2017). The study area comprises two main geological units: the Basque Arc, with Mesozoic and Cenozoic marine sedimentary materials, and the Paleozoic Massif of Cinco Villas, in the northeastern sector (EVE 2010). The heads of the basins are generally dominated by marls, limestone, and calcarenite, with some alternating detrital rock outcrops. In contrast, their middle and ending sectors present strata of marls, limestone, ophiolites, clays, gypsum, and thick detrital grain materials, usually covered by Quaternary deposits (Campos and García-Dueñas 1972; Campos et al. 1983). Our study area is prone to landslides due to this combination of morphological, climatic, and geological conditions.

The most common mass movements are small to medium shallow slides (cf. Hungr et al. 2014), debris flows, and rockfalls triggered by either intense rainfall or changes in the slope condition caused by humanmade infrastructures. Larger landslides have also been identified (Remondo et al. 2003; Bonachea 2006; Mavrouli et al. 2019).

3 Landslide inventories for the Gipuzkoa province

Data on landslide occurrences in Gipuzkoa (Fig. 2) have been collected since the 1980s. Tamés et al. (1986) published a work entitled *A Study of Natural Hazards of the Historical Territory of Gipuzkoa* (*Estudio de Riesgos Naturales del Territorio Histórico de Gipuzkoa*)
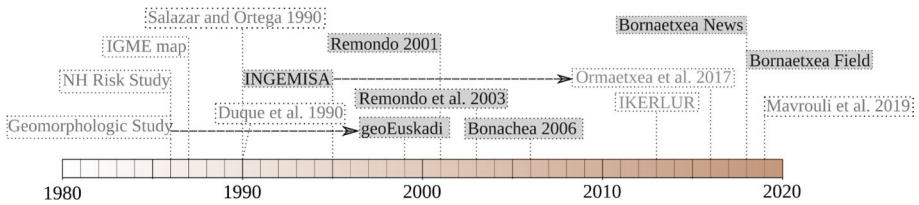
Fig. 2 Chronologic diagram of the different landslide data-collection initiatives for the Gipuzkoa province (1980–2020). Digitized databases are highlighted

in 1986 ("Natural Hazard (NH) Risk Study" in Fig. 2). The project includes written reports and paper maps of different natural hazard phenomena, including landslide processes, at a scale of 1:5000. Not all municipalities were studied for this project; as a result, the provincial map compilation shows several gaps. To the best of our knowledge, this information has not yet been digitized, and therefore can only be accessed in hard copy.

The Spanish Geologic Survey (IGME for the Spanish acronym) published a synthesis map at national scale in 1987: *Mapa de síntesis de movimientos de terreno en España* ("IGME map" in Fig. 2; Ferrer Gijon and Ayala Carcedo 1987). Only 41 landslides mapped within our study area are available in this work.

Salazar and Ortega (1990) updated the IGME inventory in 1990 including the coastal segment between the Igeldo and Mendizorrotz mountains (Fig. 1); the data are available only in hard copy.

The Gipuzkoa Provincial Council (DFG for the Spanish acronym) conducted a mapping project, simultaneously to the NH Risk Study, that was published in 1991: *Geomorphological and Soil Study of Gipuzkoa* (*Estudio Geomorfológico y Edafológico de Gipuzkoa*; "Geomorphologic Study" in Fig. 2; Diputación Foral de Gipuzkoa DFG 1991). The DFG produced printed maps and their associated reports for several municipalities in the province at a scale of 1:25 000. In this work, the presence of slope movements is represented as morphological units within the geomorphologic maps.

The public Spatial Data Infrastructure (IDE for the Spanish acronym) digitized and published those printed maps in 1999, resulting in the geoportal geoEuskadi ("geoEuskadi" in Fig. 2; IDE geoEuskadi 2022).

Duque et al. (1990) studied the effects of intense rainfall events occurred on slopes in 1989. They exploited the collected data to assess the validity of the Natural Hazard (NH) Risk Study conducted in the early 1980s.

In 1995, the Basque Regional Council published a regional landslide inventory conducted by the private company INGEMISA (Gobierno Vasco 1995a). This inventory includes the three Basque Regional Council provinces (i.e., Gipuzkoa, Bizkaia, and Araba) and comprises three printed volumes (one per province) of individual fieldwork sheets with detailed information about each landslide and its coordinates. Ormaetxea and Sáenz de Olazagoitia (2017) contributed significantly by digitizing the landslide location points from the INGEMISA 1995 inventory.

In 2013, the Roads Management Department of the Provincial Council of Gipuzkoa conducted a geotechnical assessment of the entire road network aided by private company IKERLUR consultants (DFG 2013). The resulting PDF document contains detailed information about landslide points detected close to roads. Similarly, Mavrouli et al. (2019) designed and applied a rockfall risk assessment approach in a short segment of the road N-634 connecting Zarautz and Getaria municipalities (Fig. 1). Led by the Roads

Management Department, they generated a rockfall inventory based on reports from the emergency services spanning nearly 10 years.

Some of the most relevant contributions to the recording of landslide events in Gipuzkoa came from the Universidad de Cantabria research group. Remondo (2001) generated a landslide inventory from aerial photos taken in 1991 and field visits. This inventory includes only four municipalities located in the Lower Deba valley (western Gipuzkoa; Fig. 1). Remondo et al. (2003) also conducted a second inventory for the same area including landslides that occurred between 1991 and 1997. In addition, since 2000, annual landslide inventories have been compiled for the municipality of Deba by Bonachea (2006), Bonachea et al. (2016), and Rivas et al. (2022), all of them using photointerpretation and fieldwork.

Lastly, Bornaetxea (2018) conducted two different landslide inventories covering the entire study area. The first is a fieldwork-based inventory obtained using field surveys during 2015 and 2016 ("Bornaetxea Field" in Fig. 2; Bornaetxea et al. 2018a). The second is based on a review of newspaper articles that mention landslide events spanning from 2006 to 2015 ("Bornaetxea News" in Fig. 2; Bornaetxea et al. 2018b).

All the abovementioned independent initiatives to collect and record landslide events suggest (i) a generalized interest from several institutions and individuals in recording landslide information due to the significant problems caused by landslides in the province and (ii) that the information is largely dispersed and sometimes inaccessible. These two factors are probably the motivations behind the recently published technical report *Exposure of the Gipuzkoa Historical Territory to Landslides and Slope Instabilities under Climate Change Scenarios* (*Exposición del territorio histórico de Gipuzkoa frente a deslizamientos en masa y flujos bajo escenarios de cambio climático*; DFG 2022). This work is part of the Strategy to Prevent Climate Change 2050 promoted by the provincial government of the Gipuzkoa province. According to the authors, the reason for not using data-driven approaches was "the existing unfavorable conditions in terms of quality, representativeness and quantity of the landslide inventory".

Many of the abovementioned inventories have been used in the past to produce landslide susceptibility maps, following both statistical and heuristic (experience driven) methodologies. The INGEMISA inventory was first used to generate a heuristic landslide susceptibility map of the entire Basque Region (Gobierno Vasco 1995b); then Ormaetxea and Sáenz de Olazagoitia (2017) used the same information to generate a new landslide susceptibility map following a semiquantitative approach. Remondo (2001) took advantage of his own inventory to produce a landslide susceptibility map of the Lower Deba area using, among other statistical methods, the favorability functions approach. Subsequently, Felicísimo et al. (2013) used that same inventory to obtain enhanced statistical models (multiple logistic regression, multivariate adaptive regression splines, classification and regression trees, and maximum entropy). The geoEuskadi inventory was adopted by a private company (CGS company) to run a landslide susceptibility model (LSM) for the Gipuzkoa province, applying the linear discriminant analysis method (Diputación Foral de Gipuzkoa DFG 2007). Finally, Bornaetxea et al. (2018a) exploited the information collected from fieldwork to produce several landslide susceptibility maps for the Gipuzkoa province using logistic regression as their modelling approach.

Besides the differences in the modelling approaches, these models were all implemented at different scales or resolutions and used different explanatory variables, thus preventing their direct comparison. However, since they all refer to the susceptibility of landslide occurrence (in its general definition) in the same geographic area, we believe the degree of agreement between them should be investigated.

4 Data

Among all the landslide inventories mentioned in Sect. 3, the present work focuses on the five digitized landslide inventories that cover, totally or partially, the province of Gipuzkoa, hereinafter referred to as INGEMISA, geoEuskadi, Lower Deba, Bornaetxea Field, and Bornaetxea News (Fig. 3). Specifically, Lower Deba comprises the inventories obtained by Remondo (2001), Remondo et al. (2003), and Bonachea (2006). Landslide occurrence updates mentioned in Bonachea et al. (2016) and Rivas et al. (2022) refer to one single municipality, so we decided not to include these data in the study to maintain homogeneity within the Lower Deba area. Consequently, all landslide occurrences included in Lower Deba predate the year 2000.

Descriptive data and references regarding these inventories are presented in Table 1. The period of time in which these were generated, the method used for data collection, and even their purposes are heterogeneous, challenging their direct comparison.

One of the mayor differences between inventories is the publication date. In an optimal scenario, the most recent inventories should contain all the information collected in previous databases. Another relevant variable is the length of time during which these databases were compiled: from one year, in the case of Bornaetxea Field, to nine years, in the case of Lower Deba.

The scale, directly related to the area originally covered in each project, is a factor that can affect the level of detail of the inventories; for instance, INGEMISA and geoEuskadi were originally conducted in a total area larger than the Gipuzkoa province, while Lower Deba focused in four municipalities within this province, a much smaller area (140 km^2).

The geometrical mapping approach chosen in each project can also affect the level of detail of the inventories. Three of the databases studied in the present work are represented spatially as single points, and the other two are mapped as polygons, providing information about the shape and size of each landslide. This is directly related to the data collection method followed in each case. For the INGEMISA inventory, for example, landslide occurrence information was obtained by reviewing the previous
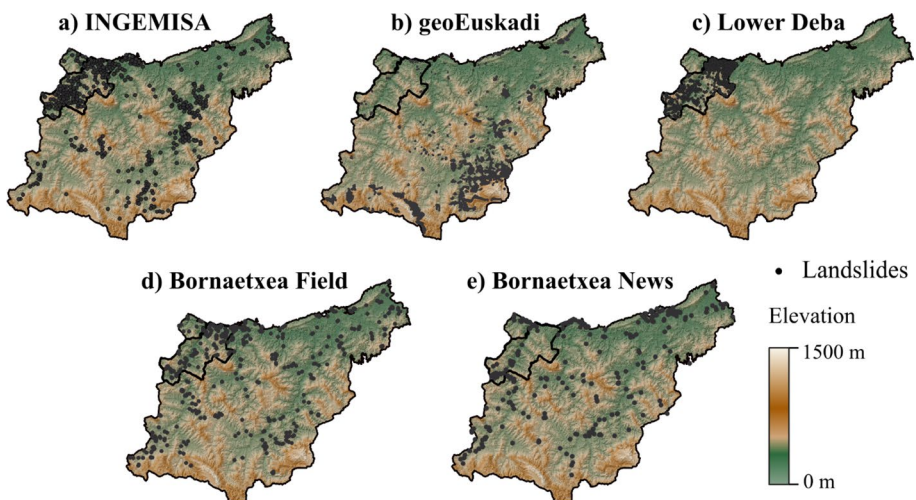

Fig. 3 Spatial distribution of the inventories analyzed

Table 1 Descriptive data and references for the five landslide inventories studied

	INGEMISA	geoEuskadi	Lower Deba	Bornaetxea Field	Bornaetxea News
Reference	Ormaetxea and Sáez de Olazagoitia (2017)	Diputación Foral de Gipuzkoa DFG (2007)	Remondo et al. (2003); Bonachea (2006)	Bornaetxea et al. (2018a)	Bornaetxea et al. (2018b)
Data collection time period	1988–1995	1984–1994	1991–2000	2015–2016	2016
Publication year	1995	1999	2003; 2006	2018	2018
Original working scale	–	1:25,000	1:25,000–1:10,000	–	–
Area covered by the full inventory	1980 km²*	1980 km²*	140 km²	1980 km²	1980 km²
Mapping geometry	Points	Polygons	Points	Polygons	Points
Data collection method	Bibliographic compilation and field visits	Unknown	Photointerpretation and field visits	Field survey	Newspapers review
Landslide types	Slides, falls, flows, topples, complex	Detachment scarp, colluvium, shallow slides, rock slides	Shallow landslides not directly affected by roads	Shallow slides	Landslides (generic)
Purpose	Risk and susceptibility analysis	Geomorphological map	Risk and susceptibility analysis	Risk and susceptibility analysis	Rainfall threshold
Number of total observations	481*	385*	2290	542	324
% Slide type	79%	22%	100%	100%	–

Values with (*) indicate inventories originally covering a wider field than the study area; only data referring to the Gipuzkoa province are considered

bibliographic references and then verifying this through field visits. In contrast, the Lower Deba inventory relies on the stereoscopic visual interpretation of aerial photographs followed by fieldwork observations. The specific approach followed for geoEuskadi remains unknown, while direct field observations were the main source of information for Bornaetxea Field. For Bornaetxea News, location and occurrence date of landslides were obtained from articles published in local newspapers.

The type of landslides mapped in each inventory also varies. INGEMISA and geoEuskadi include every type of landslides—slides, rockfalls, flows, or complex movements—while Lower Deba and Bornaetxea Field focus only in shallow slides. Moreover, in Lower Deba, the observed landslides directly related to road-side works are systematically excluded. Bornaetxea News does not specify the typology of the movements due to lack of detailed information in the newspaper reports.

Regarding the purpose of the inventories, INGEMISA, Lower Deba, and Bornaetxea Field were produced specifically to investigate the landslide susceptibility of the territory; geoEuskadi was also used for this purpose even though its original aim was to describe the general geomorphological characteristics of the province (Table 1). The aim of Bornaetxea News inventory was more specific: to study the triggering effect of rainfall in landslide occurrences.

Number of total observations varies significantly among datasets. The largest inventory is Lower Deba, with 2290 observations, despite covering a relatively small area compared with those covered in the other inventories. The second largest inventory, Bornaetxea Field, shows 542 observations, followed by INGEMISA, geoEuskadi, and Bornaetxea News with 481, 385, and 324 observations, respectively. Three of the inventories studied herein show a significant abundance of slide-type movements, whereas the other two show either a relatively small amount of slide typology (geoEuskadi; 22%) or no landslide type classification whatsoever (Bornaetxea News). The two inventories mapped as polygons indicate a large difference in size. In geoEuskadi, the mapped features range between 1144 m^2 and 6 772 478 m^2, whereas in Bornaetxea Field, the polygons range between 7.27 m^2 and 11 620 m^2. Therefore, landslides in the latter are more than three orders of magnitude smaller than those in the former. In fact, most mapped objects in geoEuskadi seem to represent sliding areas in general terms instead of the specific location of each single landslide (Fig. 3), therefore covering large extensions.

5 Methodology

Several examples of statistical and geospatial comparative analyses of different landslide inventories for a particular area were found in literature (Carrara et al. 1993; Galli et al. 2008; Pokharel et al. 2021; Bornaetxea et al. 2022). However, similar experiments have never been undertaken in the Gipuzkoa province. Furthermore, the diversity of the available landslide inventories in our study area forced us to adapt some of the experiments proposed in literature to the specifications of the data available. Hence, the methods described in this section were designed in order to highlight the discrepancies between multi-source inventories and to show their predictive capability when used to produce LSMs.

We performed pairwise comparisons that could be summarized in four successive analyses:

5.1 Computation of the pairwise spatial matching index

The study area was divided in slope units (SUs) following the approach suggested by Alvioli et al. (2016). As inputs, we used a digital elevation model with 5×5 m cell resolution, 150 000 m^2 as the minimum area for SU subdivision, a value of 0.2 as circular variance threshold, and a reduction factor of 5. We used a threshold of 25 000 m^2 to clean the SU map. Then, we counted the number of SUs containing at least one landslide. We decided to express the amount of SUs with landslides in relative form—i.e., in relation to the total number of SUs in the area covered by each inventory—given that the extension of the area covered by the Lower Deba inventory was smaller than that of those covered in the other databases. The use of SUs allowed the comparison between inventories with different geometry, such as points and polygons, and reduced the positional errors due to scale differences or spatial inaccuracies during digitation.

Next, we adapted the conventional error index proposed by Carrara et al. (1993) and used by other authors (Alvioli et al. 2018; Fiorucci et al. 2018; Pokharel et al. 2021). We calculated the spatial matching index (SMI) between two inventories with the following expression:

$$\mathrm{SMI} = \frac{A \cap B}{A + B - (A \cap B)} \times 100 \tag{1}$$

where A and B represent the number of SUs with landslides in each inventory and $A \cap B$ represents the number of SUs with landslides in inventory A that are also present in inventory B. The unity value (SMI = 100) meant that 100% of the SUs with landslides in inventory A coincided with those in inventory B, and *vice versa*. The goal of this test was to assess the degree of correspondence in displaying the same landslide locations between different inventories and to detect whether some inventories report landslide occurrences that others do not.

5.2 Assessment of the spatial homogeneity using the visibility class test

We followed the approach proposed in Bornaetxea et al. (2022) to assess the spatial homogeneity of the studied inventories. We simulated the visibility of a landslide of ~20 m^2 (the fifth percentile of the total landslide sizes among the two polygon-based inventories) from the main road network in Gipuzkoa, counted the number of landslides in each of the five visibility classes, and plotted them into normalized landslide count plots. The rationale behind the visibility class test was that a spatially homogeneous inventory should show a nearly regular distribution of landslides among the visibility classes (Bornaetxea et al. 2022).

5.3 Exploration of the distribution of the explanatory variables

We selected slope, lithology, and land use as independent variables to perform distribution analyses. With the exception of Bornaetxea News, all inventories have been used in the past for landslide susceptibility modelling. Table 2 presents the spatial variables selected as the best predictors by the different authors who conducted LSMs in the past in Gipuzkoa. For this test, we selected only the explanatory variables present in all the

Table 2 Summary of the explanatory variables used in previous landslide susceptibility models (LSMs) in Gipuzkoa

	INGEMISA	geoEuskadi	Lower Deba	Bornaetxea Field
Elevation			*	
Slope	*	*	*	*
Slope_m		*		
Slope_s				*
Surface area ratio				*
Land use/vegetation	*	*	*	*
Lithology	*	*	*	*
Fracturation		*		
Permeability	*			*
Regolith thickness	*	*	*	*
Aspect			*	*
twi				*

The asterisk indicates that the explanatory variable has been used in previous landslide susceptibility modelling works while using the inventory specified in the corresponding column header

Slope_m: mean slope of the contributing basin; slope_s: sinusoidal slope; twi: topographic wetness index

precedent works, i.e., slope, lithology, and land use; these are the three most commonly used variables in literature (Reichenbach et al. 2018). We downloaded the thematic maps from the public spatial data infrastructure of the Basque Country (IDE geoEuskadi 2022) and rasterized them into a grid file with a resolution of 5×5 m. Since the original lithology and land use maps contained a large number of classes, these were reclassified following the expert criteria as specified in Bornaetxea (2018). More detailed description and the maps of the explanatory variables can be found in Appendix I.

For land use and lithology, we divided the landslide frequency in each class by the total frequency of the same class in the entire study area to represent the distribution of these two variables in a relative manner. For slope, we sampled the slope values and then represented them in a frequency histogram with bins of five degrees each. To facilitate comparison between datasets, we decided to convert the two polygon-based inventories, i.e., geoEuskadi and Bornaetxea Field, into point location databases.

5.4 Landslide susceptibility modelling and model comparison

We performed LSMs using the different landslide inventories and following the logistic regression (LR) modelling approach. LR is one of the most commonly used statistical approaches for landslide susceptibility in literature (Reichenbach et al. 2018; Shano et al. 2020) and was useful and reliable in several studies (Goyes-Peñafiel and Hernandez-Rojas 2021; Xing et al. 2021; Rossi et al. 2022), and also for the study area of Gipuzkoa (Felicísimo et al. 2013; Bornaetxea 2018). The mathematical relationship between the dependent dichotomous variable (presence or absence of a landslide in the

mapping unit) and the independent variables (i.e., slope, lithology, land use; $X_1, \ldots, X_n$) can be expressed as

$$p = \frac{1}{1 + e^{-z}} \tag{2}$$

where p represents the probability of landslide occurrence and z is a linear combination of independent variables. The linear combination in Eq. (1) reads as follows:

$$z = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \cdots + b_n \cdot x_n \tag{3}$$

where b_0 is the intercept of the linear model, $b_i (i = 0, 1, 2, \ldots, n)$ represents the coefficient of the regression model, and $x_i (i = 0, 1, 2, \ldots, n)$ represents the independent variable.

A different landslide inventory was used as dependent variable for each LSM; slope, land use, and lithology were used as independent variables.

To ensure robust validation of our models, we employed the LAND-SUITE software (Rossi et al. 2022) for all landslide susceptibility assessments. Following the approach outlined in Rossi et al. (2022), we implemented a rigorous cross-validation strategy for each of the four experiments. Notably, the training and validation subsets were independently divided for every experiment.

In the training phase, we randomly selected 70% of the inventory to train the models, ensuring a representative dataset for learning. Subsequently, the remaining 30% of the dataset was used as an independent validation set to evaluate the models' performance against real observations.

To compare the models' performance, we calculated the area under the receiver operating characteristic (ROC) curve (AUC). This metric allowed us to objectively compare and assess the predictive capabilities of each model. Additionally, we examined the spatial similarity among the resulting four susceptibility maps using pairwise scatter plots and the Pearson correlation index.

Finally, the resulting susceptibility maps were classified in five susceptibility classes—i.e., Very Low (0–0.2), Low (0.2–0.4), Medium (0.4–0.6), High (0.6–0.8), and Very High (0.8–1)—to display easily comprehensible maps that show the spatial patterns of the proneness to landslide occurrence.

Table 3 Relative presence of landslides with respect to the total number of SUs in the area covered in each inventory

	INGEMISA	geoEuskadi	Lower Deba	Bornaetxea Field	Bornaetxea News
Total SUs	7930	7930	538	7930	7930
SUs with at least one landslide	415	1009	322	353	239
Relative presence%	5.23	12.72	59.85	4.45	3.01

6 Results

6.1 Analysis 1

Table 3 presents the results of the landslide count. The values indicate highly variable presence of landslides between the five inventories analyzed. A clear difference can be found between Lower Deba inventory, where 59.85% of the slopes presented at least one landslide occurrence, and the rest of the databases. The second most abundant inventory was geoEuskadi, with 12.72% of SUs showing a landslide presence. The rest of the datasets showed values of 5% or lower.

The pairwise comparison of the SMI is shown in Fig. 4. We observe a moderate spatial coincidence between INGEMISA and Lower Deba as these presented the same information in 32% of SUs. A slightly smaller matching level was found between Bornaetxea Field and Lower Deba (20% of spatial coincidence). The third larger SMI corresponds to the combination of INGEMISA and Bornaetxea Field (10% of spatial coincidence). Both INGEMISA and geoEuskadi showed low reciprocal SMI (6.5%), even though these two inventories were conducted almost in the same period of time covering the same area. Bornaetxea News shows very low values of spatial coincidence compared with those of the rest of the inventories, and the mismatch is the highest with respect to the geoEuskadi inventory (2.4% of coincidence).

Therefore, the group composed by INGEMISA, Lower Deba, and Bornaetxea Field showed moderate spatial matching between them, whereas geoEuskadi and Bornaetxea News showed poor coincidence with any of the other inventories.
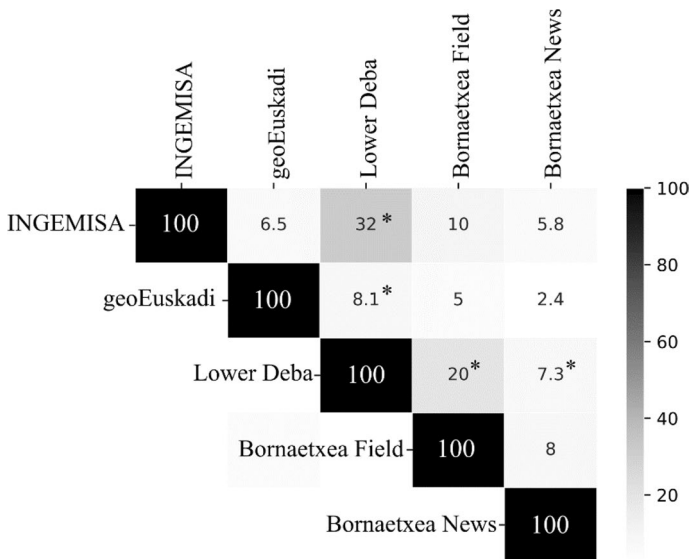

Fig. 4 Spatial mismatch index (SMI) values between the inventories analyzed. Values with (*) indicate that the index was calculated with respect to the Lower Deba area only
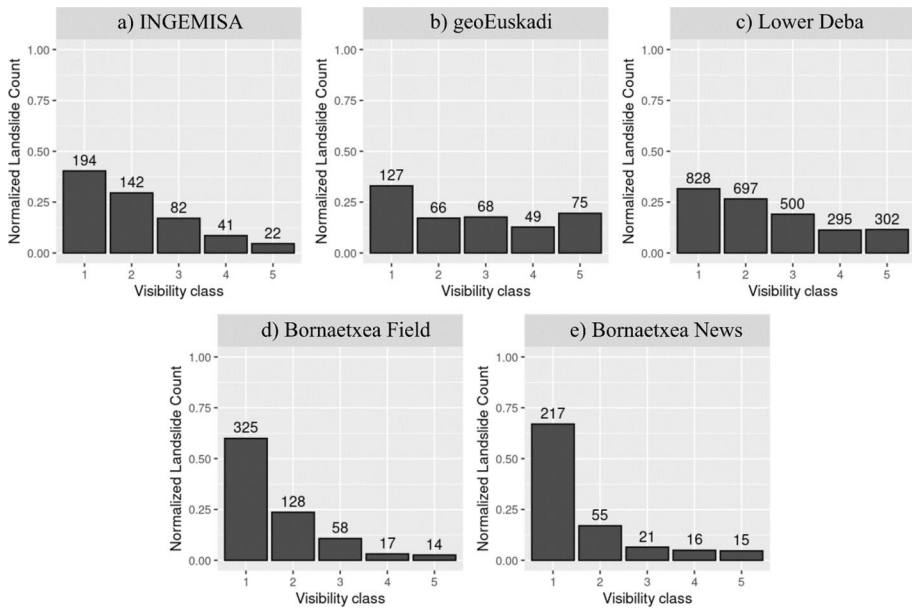
Fig. 5 Distribution of landslides among the visibility classes following the approach described in Bornaetxea et al. (2022) for the five inventories analyzed. The values on top of the bars represent the total number of landslides present within each visibility class

6.2 Analysis 2

The results from the spatial homogeneity assessment are shown in Fig. 5. INGEMISA (40% in class 1), Bornaetxea Field (60% in class 1), and Bornaetxea News (67% in class 1) show a high accumulation of landslides in the areas most visible from the roads (class 1) and a pronounced monotonic decreasing trend over the visibility classes.

The case of Lower Deba is more complex. In Fig. 5c we can observe a slight reduction of the collected information related to the visibility classes. However, this behavior is much less pronounced compared with the three abovementioned inventories. Indeed, the difference between the first class (32% of the landslides) and the rest of the classes is not as large as in Fig. 5a, d, and e. Moreover, the amount of data within the less visible classes (23% in class 4 and 5) is much more numerous.

In opposition, the geoEuskadi inventory shows a completely different distribution (Fig. 5b). Landslides are evenly distributed among visibility classes 2–5, although in the first class, a relevant abundance of observations can be found, similarly to the other inventories.

6.3 Analysis 3

Figure 6 presents the relative distributions of land use and lithology, along with the frequency distribution of slope, for each of the inventories. We decided to convert the two polygon-based inventories, i.e., geoEuskadi and Bornaetxea Field, into point location databases to facilitate the comparisons. Based in the chosen methodology and the size of the
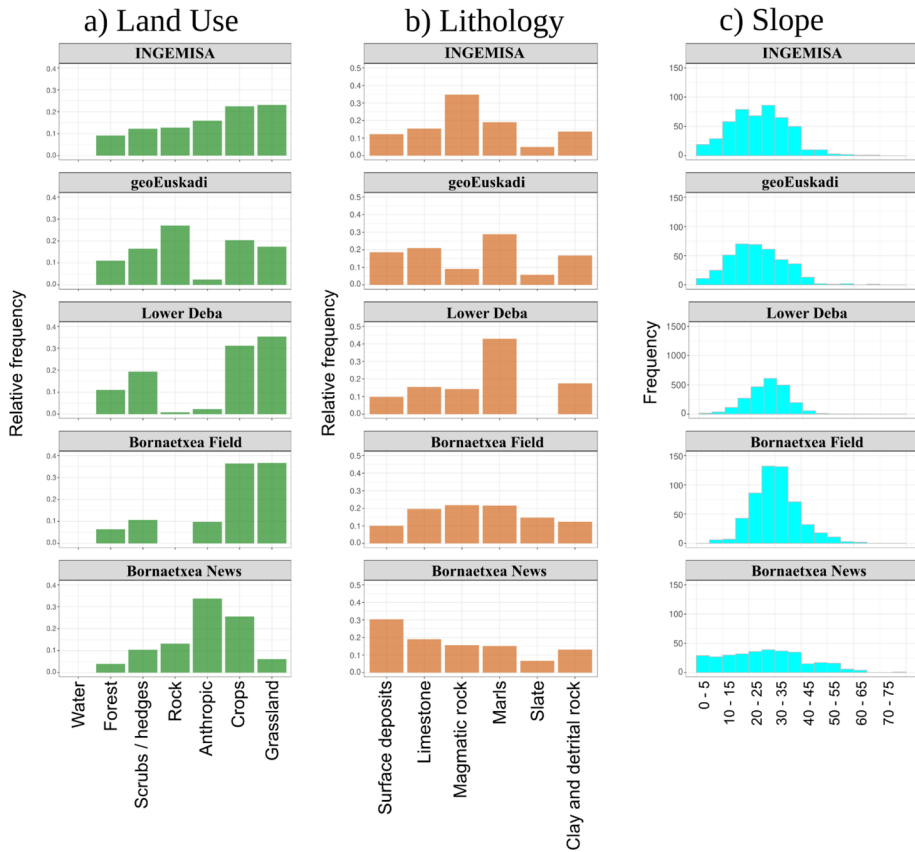
Fig. 6 Relative distribution of Land Use and Lithology and frequency distribution of Slope among the five inventories analyzed

mapped landslides, we assumed Bornaetxea Field to fairly represent the boundaries of the mass movements; therefore, the points were sampled at the higher position of the polygon to simulate the location of the primary displacement scar. For geoEuskadi, polygon centroids were sampled, since the large areas covered by many of the features suggest an inexact individualization of the mapped landslides.

The relative distribution of land use (Fig. 6a) shows some similarities between INGEMISA, Lower Deba, and Bornaetxea Field. For these three inventories, the most important land use classes are crops and grassland. However, anthropic land use and rock outcrops are the third and fourth most relevant classes, respectively, for INGEMISA, while these are almost irrelevant for the Lower Deba inventory. For the latter, scrubs or hedges are more significant, whereas forest land use shows a moderate presence in the landslide locations. In Bornaetxea Field, scrubs and hedges are also the third class, closely followed by anthropic land use. Forest is the least present land use, and rock outcrop is completely absent.

In geoEuskadi and Bornaetxea News, the distribution of land use classes shows substantial differences with respect to the other three inventories. The former presents rock outcrops as the most relevant class, while the latter locates the most part of the landslides

in anthropic land use. Moreover, the less important land use class for Bornaetxea News is grassland, one of the most frequent land use classes in the other inventories.

Relative distributions among the lithology classes are more heterogeneous (Fig. 6b). For the INGEMISA inventory, the most relevant lithology is the magmatic rock, followed by marls and limestone. Clay and detrital rock, together with surface deposits, show a moderate relative frequency, while slate is clearly the less important lithology class. For geoEuskadi, marls are the most frequent lithology, followed by limestone, surface deposits, and clay and detrital rock. In this case, slate is also the least relevant class, while magmatic rock, the most important for INGEMISA, is the second least important lithology. For Lower Deba, the abundance of landslides in marls is remarkable over the rest of the classes; clay and detrital rock, limestone, and magmatic rock maintain a moderate relative frequency, while surface deposits are less relevant. The complete absence of slate responds to the real absence of this lithology in the area analyzed for Lower Deba. In contrast, differences between lithology classes are much less pronounced in Bornaetxea Field. Magmatic rocks, marls, and limestone show similar relative frequencies, and slate is more relevant than for any of the other inventories. Surface deposits remains the least important lithology. For Bornaetxea News, however, the majority of landslides fall in the surface deposits lithological class. Limestone is the second most frequent class, followed by magmatic rock, marls, and clay and detrital rock. Again, the least abundant lithology is slate.

Slope frequency distributions are shown in Fig. 6c, where INGEMISA, Lower Deba, and Bornaetxea Field present again some similitude. In these three cases, frequency follows a normal distribution, with a peak slope angle in the range of 30–35º, although a higher presence of landslides can be found with lower slope values in INGEMISA than in Lower Deba and Bornaetxea Field. The geoEuskadi inventory also follows a Gaussian distribution, but the peak can be found in much lower slope values, i.e., in the range of 15–25°. Once again, the inventory that differs the most is Bornaetxea News, for which a flattish distribution can be observed, suggesting that landslides can be observed at almost any slope inclination from 0 to 40º.

6.4 Analysis 4

Figure 7 shows the LSMs calibrated with INGEMISA, geoEuskadi, Lower Deba, and Bornaetxea Field. In view of the evident differences resulting from the previous tests, we decided not to include Bornaetxea News inventory in this experiment. Considering the spatial patterns shown by the LSMs, we highlight some remarkable differences between inventories. We observe distinctive zones of very low susceptibility in the east sector of both INGEMISA and geoEuskadi maps (Fig. 7a and b); a susceptibility class that does not appear in the Bornaetxea Field map (Fig. 7d). This area corresponds to the only slate outcrop in the province (see Fig. 10b), so the maps agree well with the relative distribution of the lithology observed in Fig. 6b, where slate shows the lowest relative frequency for INGEMISA and geoEuskadi, but not for Bornaetxea Field. In Fig. 7a, the very low susceptibility class is scarce, as is the very high class. Thus, the majority of the province territory is classified between low and high susceptibility. Conversely, Fig. 7d shows very polarized susceptibility values, with the very low class covering a large sector of the territory, and some very high-class zones distributed throughout the study area. The smooth spatial distribution of the susceptibility obtained with INGEMISA and Bornaetxea Field contrasts with the evident concentration visible in geoEuskadi map and, in a less pronounce manner, in the Lower Deba map (Fig. 7b and c). In both these maps, the valley bottom areas are
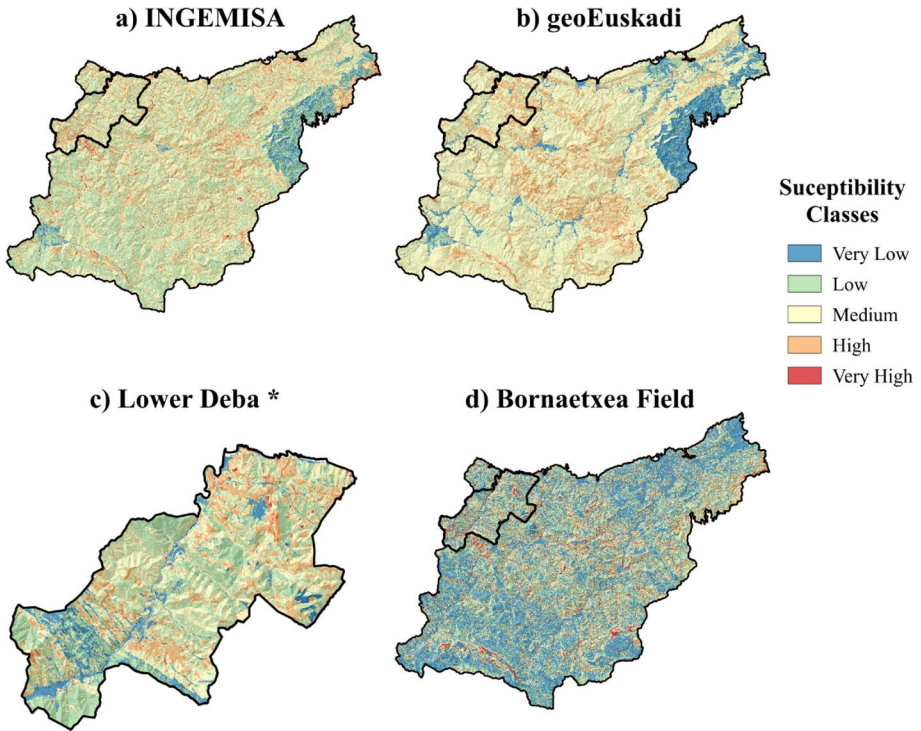
Fig. 7 Landslide susceptibility maps (LSMs) obtained using LR modelling and calibrated with four different inventories
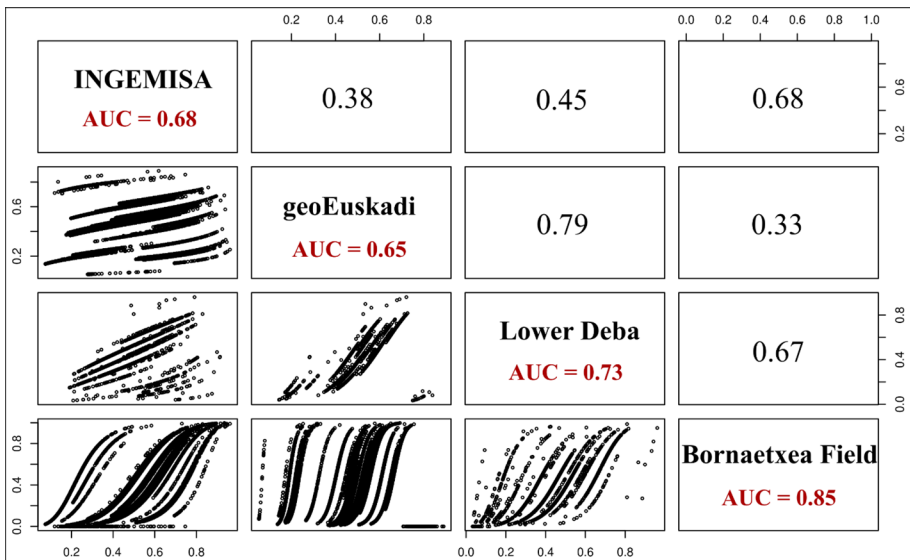

Fig. 8 Pairwise scatter plots and Pearson correlation coefficients between the four LSMs. AUC: area under the ROC curve

classified as very low susceptibility and the highest values are defined by marls lithology type, suggesting that susceptibility is highly controlled by lithology in these two models. However, Fig. 7b shows complete absence of very high susceptibility values.

6.4.1 Validation of the models

Each susceptibility model presented above are calibrated with a random sample of 70% among the respective inventories. Computed susceptibility values corresponding to the remaining 30% of the observations are extracted to obtain ROC curves for each. The resulting areas under the curves (AUC) are shown highlighted in red in Fig. 8. The figures obtained suggest that the model calibrated with the geoEuskadi inventory performs the worst (AUC = 0.65); the model calibrated with the INGEMISA model performs slightly better (AUC = 0.68); and the models calibrated with Lower Deba and Bornaetxea Field perform the best (AUC = 0.73 and AUC = 0.85, respectively). Therefore, our results indicate that Lower Deba and Bornaetxea Field fit the data used for calibration satisfactorily, whereas INGEMISA and geoEuskadi show a lower predictive capacity.

Clearly different susceptibility maps are obtained by the visual interpretation of map evidence in each of the four inventories. However, the pairwise scatter plots and their numerical correlations observed in Fig. 8 offer a more objective way to compare these results. The first map, INGEMISA, shows a very low coincidence with geoEuskadi, with which it shares a correlation coefficient value of 0.38 of; the correlation with the Lower Deba map increases to 0.45; while the largest similarity, a correlation coefficient of 0.68, is shown with the Bornaetxea Field map. The geoEuskadi inventory shows the largest correlation coefficient with Lower Deba (0.79); however, its results agree with neither INGEMISA nor Bornaetxea Field (with which the correlation is the lowest: 0.33). Lower Deba and Bornaetxea Field show a significant correlation of 0.67, which indicates that these two maps are moderately in agreement, at least within the Lower Deba area boundaries.

7 Discussion

The purpose of this study was to investigate, describe, and compare different landslide inventories used for creating susceptibility and hazard maps in the Gipuzkoa province (northern Spain). Additionally, we aimed to visually demonstrate how the use of different landslide inventories directly affects the final susceptibility map. By comparing and analyzing multiple landslide inventories, we sought to advance the understanding of landslide susceptibility modelling and improve the accuracy of the resulting maps. We stress that the objective of this paper is not the finding of the best landslide susceptibility model possible for this region. To do so, apart from the need of a previously assessed and reliable inventory, we would also need to experiment with multiple modelling approaches and an extended list of explanatory variables (Zeng et al. 2023). Instead, in this paper we want to propose some specific methods that can be used to compare the existing inventories in an area, and to highlight the direct effect of using those different inventories on the final susceptibility maps. In this section, we discuss the results of our research and their implications, considering their alignment with existing literature, while addressing their limitations.

According to the results showed in Sect. 6.1, Lower Deba present a very large number of SUs with landslides. This responds to the number of records in the database due to the detailed scale in which this inventory was conducted. However, the second larger inventory, in terms of data quantity, is Bornaetxea Field, which shows the second lowest percentage in Table 3. The use of direct field observations to develop this inventory implies that a relevant part of the study area remained unvisited. As a result, observations are only concentrated in some areas, so more than one landslide record may be located in the same SU. In contrast, Lower Deba was developed using photointerpretation at a higher resolution combined with fieldwork, which allows a more homogeneous treatment of the study area. Surprisingly, the geoEuskadi inventory covers a relevant part of the SUs (1009 SUs with landslide), despite the small number of records in this dataset (385 landslides; Table 1), indicating that many landslide polygons cover more than one SU. This is likely the effect of large and very large polygons that represent sliding zones—rather than specific locations of single landslides—that cross SU borders covering various units at the same time. This presumably overestimates the real extension of individual landslides.

The relevant spatial mismatch observed in Fig. 4 cannot be attributed to the spatial inaccuracy of the mapped landslides because we used SUs for the comparison of landslide locations. Thus, the absence of landslides in some inventories should be attributed to other factors. A possible reason is the large spatial incompleteness of information in some cases, which restricted data collection to specific areas, leaving other parts of the study area unsurveyed. As mentioned above, this issue is often linked to the methodology and the type of data source considered (Steger et al. 2017). The interpretation of aerial photographs allows a large percentage of the study area to be covered, while fieldwork is often limited by accessibility, time constrains, or economic reasons (Brardinoni et al. 2003). When press archives are used, the majority of the records are referred to landslide causing some damage, which implies a bias towards the infrastructure network and the most populated areas (Stanley and Kirschbaum 2017). The chosen methodology also determines the type of landslide inventoried, since a recent database may ignore certain types of landslides previously registered in older inventories because different objectives were pursued.

So, these tests confirm that the five inventories analyzed provide a considerably different data, which was something to be expected. Firstly, because differences in the working spatial scale inevitably affect the amount of collected data (Malamud et al. 2004; Galli et al. 2008; van Westen et al. 2013); but also because of the different duration of the data collection time periods. While the Lower Deba database has been almost continuously updated for a period of 9 years, data in Bornaetxea Field was compiled between 2015 and 2016. This may be relevant since the geomorphic features of landslides in steep terrain swiftly disappear due to the action of weathering agents, growth of vegetation, or changes in land use. Thus, a database covering a period of several decades that is frequently updated is more likely to register a larger number of landslides. In contrast, a short period of data collection or the irregular updates of an inventory can result in the loss of information between the successive landslide mapping campaigns (Yang et al. 2018; Tanyas et al. 2021).

Regarding Sect. 6.2, Bornaetxea et al. (2022) highlighted that, inventories created using field surveys and/or historic legacy data may suffer from compromised uniformity in landslide mapping. This is often reflected in the high concentration of landslides in the areas most visible from the roads (class 1 in Fig. 5), and in a pronounced monotonic decreasing trend, indicating significant heterogeneity in completeness. In fact, INGEMISA, Bornaetxea Field, and Bornaetxea News (Fig. 5a, d and e) exhibited the expected behaviour, suggesting that these inventories may be considered complete only within specific regions, such as those proximal to roads and urban areas, but not across the entire study area.

In the case of data collection for Lower Deba, a different methodology was employed, although a slight, but still monotonic, descending pattern can also be observed in Fig. 5c. This can be attributed to the inclusion of landslides identified during fieldwork near the dense road network prevalent in the study area. Furthermore, the presence of roads could be a contributing factor to the higher abundance of inventoried landslides in the more visible classes (Bruschi et al. 2013), potentially influenced by unstable cut slopes, altered water runoff patterns due to drainage works, and other factors (Taylor et al. 2020). Notably, a relevant percentage of the road network within the study area runs through the hillsides of the valleys, where the presence of steeper slopes is a clear conditioning factor for the occurrence of landslides. Thus, we cannot reject the hypothesis that a greater number of landslides actually occur in close proximity to roads in Gipuzkoa. Indeed, class 1 is the most prevalent column in Fig. 5 for all cases. Notwithstanding, the decreasing presence of landslides in classes 2–5 indicates that all inventories, except geoEuskadi, exhibit some bias toward areas visible from roads. Among them, Lower Deba displays a relatively lower impact of this bias, as evidenced by the considerable abundance of observations in less visible zones (classes 4 and 5). Conversely, the nonmonotonic distribution observed in the geoEuskadi inventory (Fig. 5b) is likely attributable to inaccurately located landslides rather than the inventory's completeness.

Concerning the distributions of explanatory variables among the inventories (see Sect. 6.3), we observe that INGEMISA, Lower Deba, and Bornaetxea Field show similar behavior regarding land use and slope. In the three inventories, crops and grassland types are dominant, and forests are among the less represented classes. This is likely due to the generalized underestimation of landslide occurrences in forested areas, which is most common in inventories obtained by means of traditional approaches (Brardinoni et al. 2003; Schmaltz et al. 2017). The comparable slope distributions also indicate that landslides can be expected in similar slope conditions, according to these databases. Furthermore, considering that Lower Deba and Bornaetxea Field collected exclusively shallow slide types of movements, they are consistent with the almost complete absence of rock outcrop, since these mass movements require, by definition, the presence of well-formed soil or altered regolith over the base rock (Cruden and Varnes 1996; Hungr et al. 2014).

Conversely, geoEuskadi and Bornaetxea News show relevant differences with respect to other inventories. The former presents rock outcrops as the most relevant class, while the latter locates most landslides in relation to anthropic land use. This behavior can be directly related to the data collection methods applied in each inventory. The large polygons mapped in geoEuskadi likely include large unstable bedrock walls outcropping within them. On the other hand, the information obtained through the newspaper chronicles for Bornaetxea News inevitably concentrates data around urbanized areas (Pereira et al. 2014; Taylor et al. 2015). Moreover, in densely urbanized areas and along roadsides, the spatial resolution of the digital elevation model utilized in this research may not be accurate enough to correctly consider the abrupt slope changes. This is probably the reason behind differences also observed in slope distributions that must necessarily result in significantly different susceptibility map. In addition, the generalized mismatch between inventories with respect to lithology indicates these databases fail to display, in fully representative manner, landslide occurrence in the Gipuzkoa province.

Taking into consideration the factors mentioned above, we made the decision to exclude Bornaetxea News from the susceptibility modelling experiment (Sect. 6.4). More specifically, we found three key reasons that influenced this choice: (i) based on the findings presented in Sects. 6.1–6.3, the accuracy of landslide locations in this inventory is not guaranteed, and there is a notable bias towards urbanized areas; (ii) the absence of information

about landslide types in this inventory introduces uncertainty in the interpretation of the results; and (iii) the original purpose of this inventory was not specifically geared towards susceptibility modelling. In contrast, while there are valid arguments to also exclude geoEuskadi from the experiment, we decided to include this inventory due to its previous use in a landslide susceptibility modelling project (Diputación Foral de Gipuzkoa DFG 2007). It is nevertheless interesting to observe the susceptibility map that can result using this dataset, and to compare it with those derived from the other inventories, in a controlled experiment.

By observing Fig. 7, it becomes evident that in statistical landslide susceptibility modelling projects, if we only change the input landside inventory, while keeping all other parameters constant, the resulting maps exhibit significantly different spatial patterns; even if these inventories correspond to the same study area and were obtained for the same purpose. This matter highlights the extraordinary capacity of the statistical and machine learning approaches (such as the Logistic Regression) to find very tailored patterns that fit the input data (Goetz et al. 2015; Kavzoglu et al. 2019; Akinci and Zeybek 2021; Mehrabi 2022). We believe that this is a great opportunity for landslide susceptibility modelling in cases where there is a fully representative and reliable landslide inventory available. However, when the origin of the available data and its completeness are not verified, this may become a real drawback, as we have seen in this illustrative experiment.

As an example, in INGEMISA, Lower Deba, and Bornaetxea Field, data collection methods are well defined, and field verifications were performed in all cases. Therefore, we assume that location accuracy has been assessed, and thus the statistics obtained through these datasets present low spatial uncertainty. However, one of the reasons for the low performance of the model calibrated with INGEMISA (AUC = 0.68) could be the heterogeneity of the landslide typologies contained in the inventory, where 21% of the points represent landslides different to slide type (Table 1). Instead, in Lower Deba (AUC = 0.73) and Bornaetxea Field (AUC = 0.85), 100% of the dataset refers to the same landslide type; this specificity in the event typology probably facilitates the finding of common spatial features. In fact, correlation coefficients in Fig. 8 show a moderate agreement between these three maps, and their model fitting performance resulted in satisfactory AUC metrics, according to the conventions in literature (Zhao and Chen 2020). On the other hand, the susceptibility maps ultimately obtained reveal that the models calibrated with these inventories are heavily influenced by their inherent biases, such as a bias towards roads and urban areas (see Sect. 6.2). This results in relevant differences in the spatial distribution of the areas with extreme levels of landslide susceptibility, both very low and very high. Indeed, the highly polarized susceptibility map obtained with Bornaetxea Field inventory (Fig. 7d) probably indicates an overfitting situation, challenging the accuracy of this model if validated with future landslide occurrences.

The Lower Deba inventory represents the opposite scenario. The combination of photointerpretation and periodic field visits, together with the long observational time-span resulted in an abundant and accurate dataset, which ensures the detection of landslides terrain with limited ground visibility or in hardly accessible zones, so reducing the bias toward roads and urban area observed in the other inventories. The performance of the susceptibility model is also good (Fig. 8). Unfortunately, the small study area in which this inventory is applied is not fully representative of the entire Gipuzkoa province, and hence Lower Deba, by itself, is not applicable to our entire study area.

Although a larger variety of landslide types (only 22% of slide type; Table 1) is contained in geoEuskadi, we ignore the data collection methodology chosen for this inventory. Therefore, we cannot assume that the accuracy of the data location has been assessed by

field verifications, for example, unlike the other three inventories. In addition, the location accuracy error obtained by considering the large polygons in this inventory as single centroid points represents another relevant issue that makes this database unsuitable for generating LSMs. We thus believe that too much uncertainty is implicit in the results obtained by the geoEuskadi inventory, even though its susceptibility map shows moderate to low prediction capacity (AUC = 0.65).

The data verification tests proposed in this research have been demonstrated to adapt correctly to any kind of landside inventory type, no matter the geometry or the collection methodology, and they allowed us to compare not only event-based inventories, but also historical and geomorphological inventories. These tests lead us to conclude that currently the only valid inventories for statistical susceptibility modelling in Gipuzkoa are INGEMISA, Lower Deba, and Bornaetxea Field. Furthermore, we believe that the
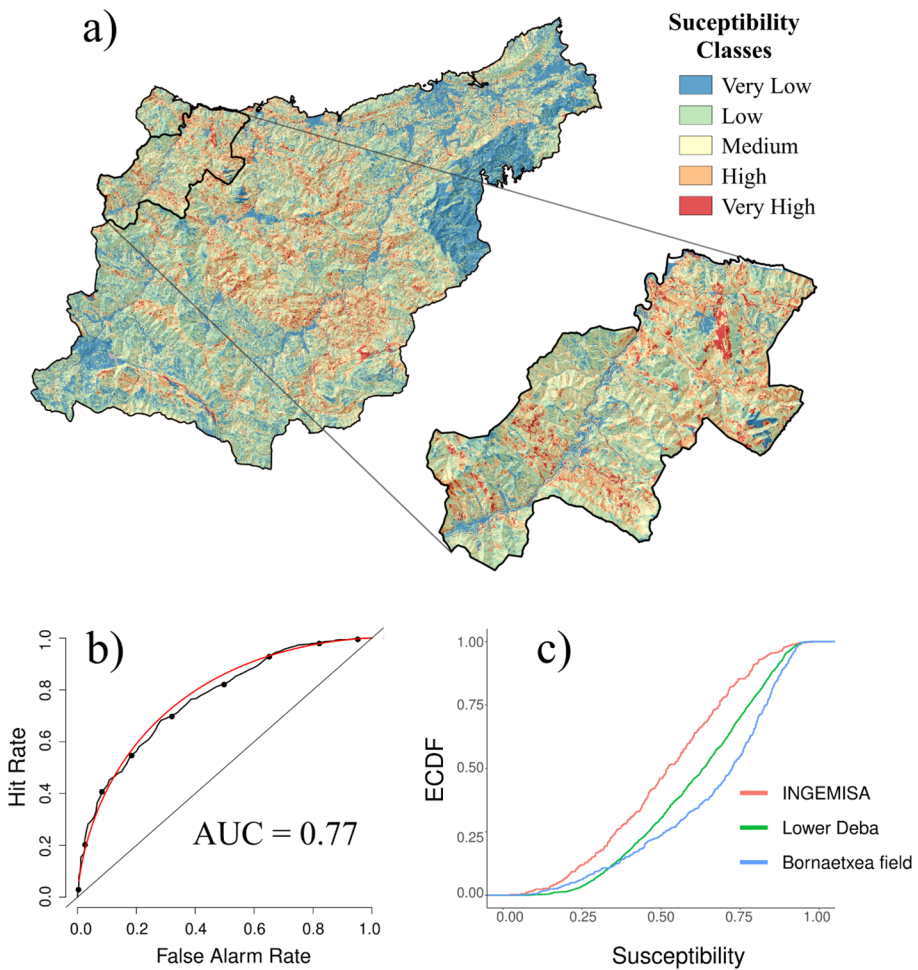
Fig. 9 **a** LSMs obtained using LR modelling and calibrated with the combination of INGEMISA, Lower Deba, and Bornaetxea Field inventories. **b** ROC curve computed with the 30% random sample reserved for validation. **c** Empirical cumulative density function (ECDF) curve of the modelled susceptibility values for each inventory

combined use of these three inventories could compensate for the inherent bias of each, resulting in a more homogeneous and representative susceptibility model. To verify this hypothesis, we conducted one last susceptibility model.

We extracted only the points corresponding to slide mass movements from INGEMISA and combined them with the Lowe Deba and Bornaetxea Field inventories. Subsequently, we conducted another logistic regression (LR) model following the exact settings applied in all previous experiments. The resulting combined landslide susceptibility map presented in Fig. 9, exhibited a prediction performance of 0.77 AUC, which can be considered good, albeit not optimal. Comparatively, the model calibrated with only the Bornaetxea Field inventory demonstrated better performance (Fig. 8). However, as noted in the literature (Reichenbach et al. 2018; Steger et al. 2021), high AUC values do not necessarily imply geomorphic causation in data-driven LSMs, and models with lower AUC may be more reliable and valuable from a geomorphological perspective. Moreover, the three inventories (individually) show a good agreement with the susceptibility map. The empirical cumulative density function (ECDF) curves displayed in Fig. 9c indicate that approximately 55% of the landslides present in each of the three inventories show susceptibility values greater than 0.5. Furthermore, this value increases up to approximately 70% when we reduce the threshold to a 0.4 of susceptibility. These results support the idea that the combined use of different inventories (once their suitability has been confirmed) can contribute to the generation of more representative susceptibility models.

8 Conclusions

The present work highlights the exceptional efforts made by different institutions and private individuals to compile sixteen landslide inventories (eleven in hard copy and five in electronic form) with different purposes within the Gipuzkoa province. The statistical analyses of five digitized landslide inventories—four of them used in the past to generate susceptibility maps—revealed remarkable differences between them. Such differences are mainly related to the characteristics of the single/multiple data sources used to obtain the information, in addition to the objectives of each database which condition the size of the study area, the temporal period covered, and the type of landslides studied, among other aspects. Notwithstanding, despite their methodological disparities, we observe a moderate spatial coincidence between some inventories and a great variability between others.

The results from our comparison tests (Sects. 6.1–6.3) indicate that the Bornaetxea News inventory does not offer a spatially representative picture of landslide occurrence in the study area. The spatial coincidence with respect to the other inventories analyzed is minimal, as shown in Figs. 5e and 6, and data are highly concentrated in urbanized areas.

Our results indicate the geoEuskadi inventory has minimal spatial accuracy, with absence of documentation detailing the precise methods followed to generate this dataset hindering our conclusions. A visual analysis of the inventory map (Fig. 3) and the results obtained from comparisons with other inventories suggest that geoEuskadi offers an extremely generalized landslide location pattern, and therefore it is hardly suitable for specific landslide analysis projects, such as susceptibility, hazard, or risk assessment.

Hence, out of the five landslide inventories digitally available in our study area, we conclude that only INGEMISA, Lower Deba, and Bornaetxea Field present enough guarantees to perform susceptibility analysis. However, we must acknowledge that specific limitations of each of these inventories are directly propagated into the final products, as shown in

Fig. 7. Therefore, the combined use of all three databases offers more plausible and generalizable results according to the AUC and ECDF tests, and to our own judgement (Fig. 9).

The methods presented in this work allowed the comparison of heterogenous landslide databases and can be easily reproduced in other study areas where multiple landslide inventory sources exist. The preliminary assessment of such a relevant input should become a critical step in any landslide susceptibility modelling project.

In addition, we believe that the continued effort of landslide data collection is key to advance on the research of susceptibility, hazard, and risk models; and a cost–benefit approach to this laborious activity should always be a priority, especially for large regions such as the Gipuzkoa province. In Appendix II, we propose some general recommendations for the compilation of landslide inventories that will hopefully be useful to motivate future landslide mapping projects in either the Gipuzkoa province, or any other national or foreign territory.

Appendix I

Description of the explanatory variables used in the experiments mentioned in Sects. 5.3 and 5.4. See Fig. 10.

Land use

It is the corresponding part to the study area of the Spanish National Forest Inventory of the 2010 (for its acronym in Spanish, IFN version (4), which is an update of IFN version (3), carried out in 2005. The update was conducted by aerial photo-interpretation of images with 25 cm of spatial resolution obtained in 2009. The original 32 classes were re-classified by expert criteria as it is shown in Bornaetxea (2018). The source layer can be found under the code INV_FORESTAL_2010_10000_ETRS89 in the geoportal geoEuskadi (IDE geoEuskadi 2022).
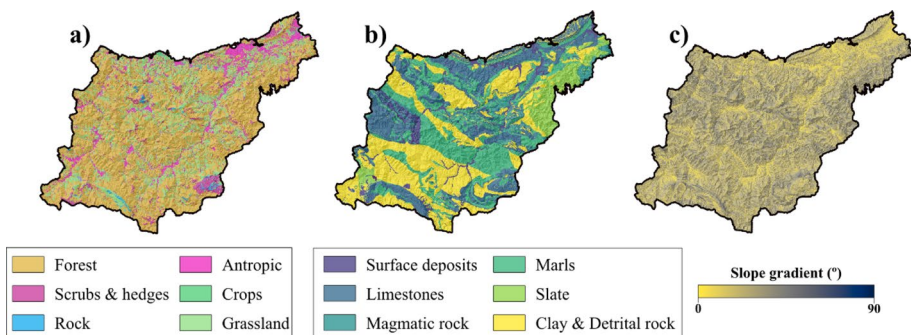

Fig. 10 **a** Land use map; **b** Lithology map; **c** Slope map

Lithology

Lithology, or rock typology, constitutes the geological substrate of the territory classified according to its composition and its physical–chemical behavior. The original layer, available in the geoportal geoEuskadi (IDE geoEuskadi 2022) and named CT_LITOLOG-ICO_25000_ETRS89, was undertaken in 1999 based on the contributions of the Geological Map of Euskadi developed by the EVE (Ente Vasco de Energía). According to this map, 21 different rock typologies can be found in study area. Knowing that the resistance against the shear tension, as well as the water infiltration capacity could be considered similar in some of those lithologies, the original 21 classes were re-classified into 6 classes following the expert criteria specified in Bornaetxea (2018).

Slope

The slope was digitally calculated as a derived product of the DEM. By means of Q-GIS software, the maximum rate of change in value from each cell to its neighbours was calculated, assuming that the maximum change in elevation over the distance between the cell and its eight neighbours identified the steepest downhill descent from the cell. For this reason, the calculations were carried out over the original DEM, and then the resulting layer was clipped in order to fit the study area. This way, border errors were avoided. This variable represents the existing angle between the terrain surface and the horizontal plane in degrees, and it shows a direct relation with the tangent and normal cutting stress of the surface formation. In addition, the slope influences very importantly the water flow velocity and distribution.

Appendix II

Based on the characteristics of the inventories analyzed in the present work and in our own personal experience, we enumerate some general recommendations for the compilation of landslide inventories. We hope these can motivate eventual future landslide mapping projects in either the Gipuzkoa province or any other national or foreign territory:

1. Inventories that cover longer time periods and are frequently updated are expected to register a large number of landslides, minimizing the loss of information due to fast natural processes of denudation of the landslide signs.
2. Landslides should always be mapped individually, since the representation of large areas where instability processes are or could be frequent will strongly distort the calibration and validation processes.
3. Polygons are preferred over points for representation purposes.
4. Among the classical data collection approaches, photointerpretation of aerial or satellite images is considered the most suitable method, as this ensures a homogenous data collection capacity over the entire study area. Regular field visits are also highly recommended to assess the spatial accuracy of the mapped features.
5. Use of different data sources to collect information can compensate the inherent bias of each database, resulting in a more homogeneous and representative inventory. Population surveys, for example, can be useful; however, the data collection method corre-

sponding to each observation must be properly registered to facilitate the decision of its use by future potential users.

6. The concentration of records in the vicinity of the road network is attributed to both natural and anthropogenic causes, but this can also be magnified depending on the type of data source used or the subsequent methodology. Both overrepresentation and underrepresentation of landslide records in these areas can affect to the calibration of landslide susceptibility models. The initial exclusion of presumed human-induced landslides from data collection is discouraged; their use can be discarded in subsequent phases of the modelling process.

Declarations

References

Akinci H, Zeybek M (2021) Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. Nat Hazards 108:1515–1543. https://doi.org/10.1007/s11069-021-04743-4

Alvioli M, Marchesini I, Reichenbach P, Rossi M, Ardizzone F, Fiorucci F, Guzzetti F (2016) Automatic delineation of geomorphological slope units with r. slope units v1. 0 and their optimization for landslide susceptibility modeling. Geosci Model Dev 9:3975–3991. https://doi.org/10.5194/gmd-9-3975-2016

Alvioli M, Mondini AC, Fiorucci F, Cardinali M, Marchesini I (2018) Topography-driven satellite imagery analysis for landslide mapping. Geomat Nat Hazards Risk 9:544–567. https://doi.org/10.1080/19475705.2018.1458050

Bernat Gazibara S, Krka M, Mihali Arbanas S (2019) Verification of historical landslide inventory maps for the Podsljeme area in the City of Zagreb using LiDAR-based landslide inventory. Min Geol Pet Eng Bull 34:45–58. https://doi.org/10.17794/rgn.2019.1.5

Bonachea J, Remondo J, Rivas V, Sánchez Espeso J, Bruschi VM, Cendrero A, Díaz de Terán JR, Fernández Maroto G, Gómez Arozamena J, González-Díez AA, Sainz C (2016) Desarrollo de escenarios de peligrosidad y riesgo por deslizamientos (proyecto Espérides). In: Durán Valsero JJ et al. (ed) Comprendiendo el relieve: del pasado al futuro: actas de la XIV Reunión Nacional de Geomorfología Málaga. Instituto Geológico y Minero de España, pp 205–212

Bonachea J (2006) Desarrollo, aplicación y validación de procedimientos y modelos para la evaluación de amenazas, vulnerabilidad y riesgo debidos a procesos geomorfológicos. PhD tesis, Universidad de Cantabria. http://hdl.handle.net/10803/10610

Bornaetxea T, Rossi M, Marchesini I, Alvioli M (2018a) Effective surveyed area and its role in statistical landslide susceptibility assessments. Nat Hazards Earth Syst Sci 18:2455–2469. https://doi.org/10.5194/nhess-18-2455-2018

Bornaetxea T, Marchesini I, Kumar S, Karmakar R, Mondini A (2022) Terrain visibility impact on the preparation of landslide inventories: a practical example in Darjeeling district (India). Nat Hazards Earth Syst Sci 22:2929–2941. https://doi.org/10.5194/nhess-22-2929-2022

Bornaetxea T, Ormaetxea O, Antigüedad I, Melillo M (2018b) Landslide and rainfalls: press inventory, conditioning factors characterization and precipitation thresholds for Gipuzkoa province (Basque country). In: Lourenço L, Vieira A, (ed) Metodologia de Análise de Riscos através de Estudos de Casos. RISCOS, pp 201–222

Bornaetxea T (2018) Methodological approach for landslide analysis in a regional scale. Data collection, susceptibility models and precipitation thresholds. Application in Gipuzkoa province (Basque Country). PhD thesis, Universidad del País Vasco-Euskal Herriko Unibertsitatea. https://addi.ehu.eus/handle/10810/32947

Bovis MJ, Jakob M (1999) The role of debris supply conditions in predicting debris flow activity. Earth Surf Process Landf 24:1039–1054. https://doi.org/10.1002/(SICI)1096-9837(199910)24:11%3c1039::AID-ESP29%3e3.0.CO;2-U

Brardinoni F, Slaymaker O, Hassan MA (2003) Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data. Geomorphology 54:179–196. https://doi.org/10.1016/S0169-555X(02)00355-0

Bruschi VM, Bonachea J, Remondo J, Gómez-Arozamena J, Rivas V, Barbieri M, Cendrero A (2013) Land management versus natural factors in land instability: some examples in northern Spain. Environ Manage 52:398–416. https://doi.org/10.1007/s00267-013-0108-7

Caine N (1980) The rainfall intensity-duration control of shallow landslides and debris flows. Geogr Ann A Phys Geogr 62:23–27. https://doi.org/10.1080/04353676.1980.11879996

Campos J, García-Dueñas V (1972) Mapa Geológico de España escala 1:50.000. 2ª Serie (MAGNA), Hoja de San Sebastián. IGME. Servicio de Publicaciones Ministerio de Industria

Campos J, Olivé A, Ramírez JI, Solé J, Villalobos L (1983) Mapa Geológico de España escala 1:50.000. 2ª Serie (MAGNA), Hoja de Tolosa. IGME. Servicio de Publicaciones Ministerio de Industria

Carrara A (1993) Uncertainty in evaluating landslide hazard and risk. In: Nemec J et al (eds) Prediction and perception of Natural hazards. Springer, Berlin, pp 101–109

Cendrero A, Dramis F (1996) The contribution of landslides to landscape evolution in Europe. Geomorphology 15:191–211. https://doi.org/10.1016/0169-555X(95)00070-L

Chmiel M, Walter F, Wenner M, Zhang Z, McArdell BW, Hibert C (2021) Machine learning improves debris flow warning. Geophys Res Lett. https://doi.org/10.1029/2020GL090874

Corominas J, van Westen C, Frattini P, Cascini L, Malet JP, Fotopoulou S, Smith JT (2014) Recommendations for the quantitative analysis of landslide risk. Bull Eng Geol Environ 73:209–263. https://doi.org/10.1007/s10064-013-0538-8

Cruden DM, Varnes DJ (1996) Landslides Types and Processes. In: Transportation research board (ed) Landslides: Investigation and Mitigation. National Research Council, pp 247

Diputación Foral de Gipuzkoa DFG (1991) Geomorfología y edafología de Gipuzkoa. Diputación Foral de Gipuzkoa

Diputación Foral de Gipuzkoa DFG (2007) Elaboración de modelos de predicción de riesgos naturales de incendios, deslizamientos e inundaciones en el Territorio Histórico de Gipuzkoa. Dirección General de Medio Ambiente. Technical Report

Diputación Foral de Gipuzkoa DFG (2013) Evaluación y gestión integrada de riesgos geotécnicos en la red de carreteras de la Diputación Foral de Gipuzkoa. Mugikortasun eta Bide Azpiegituren Saila. Technical Report

Diputación Foral de Gipuzkoa DFG (2022) Exposición del Territorio Histórico de Gipuzkoa frente a deslizamientos en masa y flujos bajo escenarios de cambio climático. Departamento de Medio Ambiente y Obras Hidráulicas. Technical Report

Duque A, Echeverría G, Fernández E, Kerejeta A, Cendrero A, Tamés P (1990) Ensayo de un modelo empírico de carácter predictivo para la evaluación de la inestabilidad de laderas. In: IV Reunión Nacional de Geología Ambiental y Ordenación del Territorio. Instituto Geológico Nacional, pp 105–113

Evans SG (1982) Landslides and surficial deposits in urban areas of British Columbia: a review. Can Geotech J 19:269–288. https://doi.org/10.1139/t82-034

EVE (2010) Mapa Geológico del País Vasco 1:100.000. Ente Vasco de la Energía

Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2013) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10:175–189. https://doi.org/10.1007/s10346-012-0320-1

Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ (2008) Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. J Eng Geol 102:85–98. https://doi.org/10.1016/j.enggeo.2008.03.022

Ferrer Gijon M, Ayala Carcedo FJ (1987) Mapa de movimientos del terreno de España a escala 1:1.000.000. IGME, Technical Report

Fiorucci F, Giordan D, Santangelo M, Dutto F, Rossi M, Guzzetti F (2018) Criteria for the optimal selection of remote sensing optical images to map event landslides. Nat Hazards Earth Syst Sci 18:405–417. https://doi.org/10.5194/nhess-18-405-2018

Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. Nat Hazards Earth Syst Sci 18:2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Galli M, Ardizzone F, Cardinali M, Guzzetti F, Reichenbach P (2008) Comparing landslide inventory maps. Geomorphology 94:268–289. https://doi.org/10.1016/j.geomorph.2006.09.023

Gariano SL, Guzzetti F (2016) Landslides in a changing climate. Earth Sci Rev 162:227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Goetz JN, Brenning A, Petschko H, Leopold P (2015) Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput Geosci 81:1–11. https://doi.org/10.1016/j.cageo.2015.04.007

Goyes-Peñafiel P, Hernandez-Rojas A (2021) Landslide susceptibility index based on the integration of logistic regression and weights of evidence: A case study in Popayan. Colombia. J Eng Geol 280:105958. https://doi.org/10.1016/j.enggeo.2020.105958

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang KT (2012) Landslide inventory maps: New tools for an old problem. Earth Sci Rev 112:42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Hungr O (2007) Dynamics of rapid landslides. In: Sassa K, Fukuoka H, Wang F, Wang G (eds) Progress in landslide science. Springer, Berlin, pp 47–57

Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. Landslides 11:167–194. https://doi.org/10.1007/s10346-013-0436-y

Hutter K, Svendsen B, Rickenmann D (1994) Debris flow modeling: A review. Contin Mech Thermodyn 8:1–35. https://doi.org/10.1007/BF01175749

IDE geoEuskadi (2022) Infraestructura de datos espaciales de Euskadi, URL, www.geo.euskadi.eus

Jacobs L, Dewitte O, Poesen J, Sekajugo J, Nobile A, Rossi M, Thiery W, Kervyn M (2018) Field-based landslide susceptibility assessment in a data-scarce environment: the populated areas of the Rwenzori Mountains. Nat Hazards Earth Syst Sci 18:105–124. https://doi.org/10.5194/nhess-18-105-2018

Jun H, Min DH, Yoon HK (2017) Determination of monitoring systems and installation location to prevent debris flow through web-based database and AHP. Mar Georesources Geotechnol 35:1049–1057. https://doi.org/10.1080/1064119X.2017.1280716

Kang S, Lee SR, Vasu NN, Park JY, Lee DH (2017) Development of an initiation criterion for debris flows based on local topographic properties and applicability assessment at a regional scale. J Eng Geol 230:64–76. https://doi.org/10.1016/j.enggeo.2017.09.017

Kavzoglu T, Colkesen I, Sahin EK (2019) Machine learning techniques in landslide susceptibility mapping: a survey and a case study. In: Pradhan SP, Vishal V, Singh TN (eds) Landslides: theory, practice and modelling. Springer, Cham, pp 283–301

Kayastha P, Dhital MR, De Smedt F (2013) Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. Comput Geosci 52:398–408. https://doi.org/10.1016/j.cageo.2012.11.003

Maguire DJ, Longley PA (2005) The emergence of geoportals and their role in spatial data infrastructures. Comput Environ Urban Syst 29:3–14. https://doi.org/10.1016/j.compenvurbsys.2004.05.012

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Landslide inventories and their statistical properties. Earth Surf Process Landf 29:687–711. https://doi.org/10.1002/esp.1064

Maleki M, Mir Mohammad Hosseini SM (2022) Assessment of the Pseudo-static seismic behavior in the soil nail walls using numerical analysis. Innov Infrastruct Solut 7:262. https://doi.org/10.1007/s41062-022-00861-5

Maleki M, Khezri A, Nosrati M (2022) Seismic amplification factor and dynamic response of soil-nailed walls. Model Earth Syst Environ 9:1181–1198. https://doi.org/10.1007/s40808-022-01543-y

Marin RJ, Velásquez MF, Sánchez O (2021) Applicability and performance of deterministic and probabilistic physically based landslide modeling in a data-scarce environment of the Colombian Andes. J South Am Earth Sci 108:103175. https://doi.org/10.1016/j.jsames.2021.103175

Mavrouli O, Corominas J, Ibarbia I, Alonso N, Jugo I, Ruiz J, Luzuriaga S, Navarro JA (2019) Integrated risk assessment due to slope instabilities in the roadway network of Gipuzkoa, Basque Country. Nat Hazards Earth Syst Sci 19:399–419. https://doi.org/10.5194/nhess-19-399-2019

Mehrabi M (2022) Landslide susceptibility zonation using statistical and machine learning approaches in Northern Lecco, Italy. Nat Hazards 111:901–937. https://doi.org/10.1007/s11069-021-05083-z

Mondini AC, Guzzetti F, Chang KT, Monserrat O, Martha TR, Manconi A (2021) Landslide failures detection and mapping using Synthetic Aperture Radar: Past, present and future. Earth-Sci Rev 216:103574. https://doi.org/10.1016/j.earscirev.2021.103574

Moraga Paula, Baker Laurie (2022) rspatialdata: a collection of data sources and tutorials on downloading and visualising spatial data using R. F1000Res 11:770

Ormaetxea O, Sáez de Olazagoitia A (2017) Análisis y caracterización de los factores que intervienen en los movimientos de ladera y aproximación de la susceptibilidad en el País Vasco. Lurralde: investigación y espacio 40:81–109

Palacio Cordoba J, Mergili M, Aristizábal E (2020) Probabilistic landslide susceptibility analysis in tropical mountainous terrain using the physically based r.slope.stability model. Nat Hazards Earth Syst Sci 20:815–829. https://doi.org/10.5194/nhess-20-815-2020

Park HJ, Jang JY, Lee JH (2019) Assessment of rainfall-induced landslide susceptibility at the regional scale using a physically based model and fuzzy-based Monte Carlo simulation. Landslides 16:695–713. https://doi.org/10.1007/s10346-018-01125-z

Pellicani R, Spilotro G (2015) Evaluating the quality of landslide inventory maps: comparison between archive and surveyed inventories for the Daunia region (Apulia, Southern Italy). Bull Eng Geol Environ 74:357–367. https://doi.org/10.1007/s10064-014-0639-z

Pereira S, Zêzere JL, Quaresma ID, Bateira C (2014) Landslide incidence in the North of Portugal: analysis of a historical landslide database based on press releases and technical reports. Geomorphology 214:514–525. https://doi.org/10.1016/j.geomorph.2014.02.032

Petley D (2012) Global patterns of loss of life from landslides. Geology 40:927–930. https://doi.org/10.1130/G33217.1

Pokharel B, Alvioli M, Lim S (2021) Assessment of earthquake-induced landslide inventories and susceptibility maps using slope unit-based logistic regression and geospatial statistics. Sci Rep 11:1–15. https://doi.org/10.1038/s41598-021-00780-y

Pourghasemi HR, Pradhan B, Gokceoglu C (2012) Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed. Iran Nat Hazards 63:965–996. https://doi.org/10.1007/s11069-012-0217-2

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth Sci Rev 180:60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Remondo J (2001) Elaboración y validación de mapas de susceptibilidad de deslizamientos mediante técnicas de análisis espacial. PhD thesis, Universidad de Oviedo

Remondo J, González-Díez A, Díaz De Terán JR, Cendrero A (2003) Landslide susceptibility models utilising spatial data analysis techniques. A case study from the lower Deba Valley, Guipúzcoa (Spain). Nat Hazards 30:267–279. https://doi.org/10.1023/B:NHAZ.0000007202.12543.3a

Remondo J, Bonachea J, Rivas V, Sánchez-Espeso J, Bruschi V, Cendrero A, Díaz De Terán JR, Fernández-Maroto G, Gómez-Arozamena J, González-Díez A, Sainz C (2017) Landslide hazard scenarios based on both past landslides and precipitation. In: Sassa K et al (eds) 4th World landslide forum. Springer, pp 981–988

Rivas V, Remondo J, Bonachea J, Sánchez-Espeso J (2022) Rainfall and weather conditions inducing intense landslide activity in northern Spain (Deba, Guipúzcoa). Phys Geogr 43:419–439. https://doi.org/10.1080/02723646.2020.1866790

Rossi M, Bornaetxea T, Reichenbach P (2022) LAND-SUITE V1. 0: a suite of tools for statistically based landslide susceptibility zonation. Geosci Model Dev 15:5651–5666. https://doi.org/10.5194/gmd-15-5651-2022

Salazar A, Ortega LI (1990) Los deslizamientos de la alineación costera guipuzcoana (sector Mendizorrotz-Igueldo). In: Gutiérrez Elorza M (ed) Actas de la 1ª Reunión Nacional de Geomorfología. Diputación Provincial de Teruel, pp 551–560

Schmaltz EM, Steger S, Glade T (2017) The influence of forest cover on landslide occurrence explored with spatio-temporal information. Geomorphology 290:250–264. https://doi.org/10.1016/j.geomorph.2017.04.024

Schuster RL, Salcedo DA, Valenzuela L (2002) Overview of catastrophic landslides of South America in the twentieth century. Rev Eng Geol 15:1–34. https://doi.org/10.1130/REG15-p1

Shano L, Raghuvanshi TK, Meten M (2020) Landslide susceptibility evaluation and hazard zonation techniques–a review. Geoenviron Disasters 7:1–19. https://doi.org/10.1186/s40677-020-00152-0

Stanley TA, Kirschbaum DB (2017) Effects of inventory bias on landslide susceptibility calculations. In: Proceedings of the 3rd North American symposium on landslides. AEG Publishing, California

Steger S, Brenning A, Bell R, Glade T (2017) The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides 14:1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Steger S, Mair V, Kofler C, Pittore M, Zebisch M, Schneiderbauer S (2021) Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling–Benefits of exploring landslide data collection effects. Sci Total Environ 776:145935. https://doi.org/10.1016/j.scitotenv.2021.145935

Tamés P, Elosegui JM, Salazar A, Almoguera J, Albert J, Solesio J, Fernández Areces JL, López Vilchez L (1986) Estudio de riesgos naturales del territorio de Guipúzcoa (2ª fase). Diputación Foral Guipúzcoa, Open file report

Tanyaş H, Kirschbaum D, Görüm T, van Westen CJ, Tang C, Lombardo L (2021) A closer look at factors governing landslide recovery time in post-seismic periods. Geomorphology 391:107912. https://doi.org/10.1016/j.geomorph.2021.107912

Taylor FE, Malamud BD, Freeborough K, Demeritt D (2015) Enriching Great Britain's national landslide database by searching newspaper archives. Geomorphology 249:52–68. https://doi.org/10.1016/j.geomorph.2015.05.019

Taylor FE, Tarolli P, Malamud BD (2020) Preface: Landslide–transport network interactions. Nat Hazards Earth Syst Sci 20:2585–2590. https://doi.org/10.5194/nhess-20-2585-2020

Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13:361–378. https://doi.org/10.1007/s10346-015-0557-6

Valenzuela P, Domínguez-Cuesta MJ, Mora García MA, Jiménez-Sánchez M (2017) A spatio-temporal-landslide inventory for the NW of Spain: BAPA database. Geomorphology 293:11–23. https://doi.org/10.1016/j.geomorph.2017.05.010

van Westen CJ, Ghosh S, Jaiswal P, Martha TR, Kuriakose SL (2013) From landslide inventories to landslide risk assessment; an attempt to support methodological development in India. In: Margottini Claudio, Canuti Paolo, Sassa Kyoji (eds) Landslide science and practice: volume 1: landslide inventory and susceptibility and hazard zoning. Springer, Berlin, pp 3–20. https://doi.org/10.1007/978-3-642-31325-7_1

Gobierno Vasco (1995a) Inventario y análisis de las áreas sometidas al riesgo de inestabilidades del terreno en la CAPV. Anexos I-VII. Dirección de Atención de Emergencias. Technical Report

Gobierno Vasco (1995b) Inventario y análisis de las áreas sometidas al riesgo de inestabilidades del terreno en la CAPV. Mapa de Peligrosidad. Departamento de Interior

Xing X, Wu C, Li J, Li X, Zhang L, He R (2021) Susceptibility assessment for rainfall-induced landslides using a revised logistic regression method. Nat Hazards 106:97–117. https://doi.org/10.1007/s11069-020-04452-4

Yang W, Qi W, Zhou J (2018) Effects of precipitation and topography on vegetation recovery at landslide sites after the 2008 Wenchuan earthquake. Land Degrad Dev 29:3355–3365. https://doi.org/10.1002/ldr.3098

Yong C, Jinlong D, Fei G, Bin T, Tao Z, Hao F, Li W, Qinghua Z (2022) Review of landslide susceptibility assessment based on knowledge mapping. Stoch Environ Res Risk Assess 36:2399–2417. https://doi.org/10.1007/s00477-021-02165-z

Zeng T, Wu L, Peduto D, Glade T, Hayakawa YS, Yin K (2023) Ensemble learning framework for landslide susceptibility mapping: different basic classifier and ensemble strategy. Geosci Front 14:101645. https://doi.org/10.1016/j.gsf.2023.101645

Zhao X, Chen W (2020) Optimization of computational intelligence models for landslide susceptibility evaluation. Remote Sens 12:2180. https://doi.org/10.3390/rs12142180