A new method for locating Roman transport infrastructures

Jesús María Romera[a] and Heriberto Pérez-Acebo[b]

[a]Department of Mechanical Engineering, Faculty of Engineering of Bilbao

University of the Basque Country UPV/EHU

Rafael Moreno "Pitxitxi" 2, 48013 Bilbao, Spain

e-mail: jesusmaria.romera@ehu.eus (Corresponding author)

ORCID: 0000-0002-9899-4802

[b]Department of Mechanical Engineering, Faculty of Engineering of Bilbao

University of the Basque Country UPV/EHU

Rafael Moreno "Pitxitxi" 2, 48013 Bilbao, Spain

e-mail: heriberto.perez@ehu.eus

ORCID: 0000-0003-0577-9597

ABSTRACT

Roman cities and roads, once correctly identified, can be appropriately conserved. Moreover, the correct identification of Roman transport routes will vindicate the accuracy of recent studies on the network of Roman transport infrastructure and its connectivity, functionality, and impact. With this aim in mind, a novel method is presented for computing the most likely location, from among the various proposed locations that may exist, of any Roman city that is cited both in a Roman itinerary and in Ptolemy's Geographia. In the first phase, the geographical area where the city was located is demarcated by means of the itinerary. In the second phase, Ptolemy's coordinates of well-known Roman cities from the province of the Roman Empire that is under consideration are correlated with those of the WGS 84 reference system by means of simple linear regressions. Having confirmed the normality of the regression error distribution, the bivariate normal distribution is computed, and the confidence intervals are determined. This method is implemented, to identify the most probable location of the Vaccaean city of Intercatia in Hispania, and to propose a new route for the Roman road that once passed through it.

Keywords:

Roman roads; Antonine Itinerary; Ptolemy´s *Geographia*; *Hispania*; Linear regression; Bivariate normal distribution

Word count (including bibliography): 6825 words

1. INTRODUCTION AND RESEARCH AIM

The correct correlation of archaeological sites with the cities cited in ancient sources would undoubtedly build on our historical knowledge and help conserve their remains, even though the location of Roman roads and the cities through which they passed is no easy task. However, multiple factors are involved in any economic assessment of conservation resources, such as the relevance and the significance of the material remains, their state of preservation and risk assessment, among quite a few others.

Difficulties over the identification of Roman transport infrastructure are due to imprecise knowledge of certain routes cited in the extant historic sources and, at times, due to interests that are not always coincident with gaining greater knowledge of our common historic heritage [1].

The growing interest in discovering Roman cultural heritage for modern-day society involves developing general studies with the aim of improving the "public understanding and appreciation of its historical relevance and sociocultural significance" [2]. In that regard, various projects are under development. One example is the Mercator-e project, "Quantifying the impact of transport infrastructure: network analysis diachronically applied to the study of the Iberian Peninsula (from Roman times to 19th Century)" [3]. The project covers the analysis of network connectivity, transportation system functionality, and the impact of transport networks over time in the Iberian Peninsula [4–7]. If these studies are to be accurate, precise identification of historic transport routes is, therefore, essential.

The most widely used resources for determining the alignments of Roman roads are historical texts on geography, commentaries by Roman historians, and archaeological and numismatic findings. Regarding the first-mentioned, the most important are the Antonine Itinerary [8], the *Ravennatis Anonimy Cosmographia* [9], and Ptolemy's *Geographia* [10].

The Antonine Itinerary records the stations and the distances along various roads. Traditionally ascribed to the patronage of Antoninus Pius, the 2nd century emperor, it only links the ancient place names with their intermediate distances in single itineraries. Some of the distances are known to be erroneous, making it hard to pinpoint locations where no other historical evidence on the toponym survives and distances were not always measured along straight lines. The Least Cost Path, a Geographic Information System (GIS) technique, is currently used for the reconstruction of some potential roadways [11] and to perceive the underlying logic behind their layout [12,13]. Verhagen [14] recently presented a method for assessing the distances travelled between stations on the routes of the *Tabula Peutingeriana* and the Antonine Itinerary and their relation to sites identified around the problematic landscape of the Lower Rhine *limes*. Route alignment is hardly always straight between consecutive stations, because they follow the natural levees of the meandering river. Subsequently, a new identification of sites with Roman place names was proposed up until then subject to doubt.

In Ptolemy's *Geographia*, the longitudinal and latitudinal coordinates of almost ten-thousand locations throughout Africa, Asia, and Europe are pinpointed in a list. Because of the inaccuracy of the recorded coordinates, a large number of studies have been performed to correlate Ptolemy's coordinates with present-day coordinates, in order to identify the current locations of many sites with ancient place names. Darcy and Flynn [15] proposed two equations to adjust Ptolemy's longitudes and latitudes to the modern references in Ireland. Those equations are based on the study of the radius of the Earth and the positions of the reference parallel and meridian that match the Ptolemaic coordinates of Irish settlements. Livieratos et al. [16] performed comparative analyses of geographic coordinates between various editions of the *Geographia* using geodetic techniques. Moreover, they determined the second-order polynomial transformation method that yielded the best fit with Ptolemy's coordinates for Greek sites and their modern counterparts. Tsorlini [17] determined the order of magnitude of the differences between the Ptolemaic values and the modern-day latitudes and longitudes in the Northern Mediterranean. The pattern of coordinate differences free of systematic effects up to the 2nd order is shown by means of the isolines of the latitudinal and the longitudinal differences.

Archaeological and numismatic findings are essential elements that can add weight to novel theories. Unfortunately, there is a scarcity of archaeological and numismatic sources. This problem is addressed by the Light Detection and Ranging (LiDAR) technique, widely used as a complementary archaeological resource [11].

In the present study, a new approach for the study of the Roman roads in any province of the Roman Empire is presented and a case study in *Hispania* is analyzed. Ptolemy's *Geographia* is not usually accepted as a means for siting Roman towns in Hispania. Solana [18] failed to obtain a consistent value for the Ptolemaic degree and therefore stated that the work would be unreliable for locating Roman settlements from the Iberian Peninsula. In contrast, Montero [19] proposed that Ptolemy's degree would have had a single value, but it has not been possible to derive it from the extant codices, due to the errata in successive copies of the original manuscripts. Urueña [20] sought to locate the inland cities of the Iberian Peninsula through a study on the cartographic method that Ptolemy used in the *Geographia*. Although the analysis of the cartographic method is not the aim of the present work, the concluding remarks are of interest. The author stated the impossibility of deriving the current locations of sites with ancient place names, by using transformations of Ptolemy's coordinates. In addition, it is significant that the route *ab Asturica per Cantabriam Caesaraugustam* of the Itinerary was the only one analyzed in the study, because of its unique complexity. Defaux [21] has recently presented an in-depth study on the origin of Ptolemy's coordinates corresponding to various sites of the Iberian Peninsula. Regarding other historical geographical sources, it must be indicated that the part of Peutinger's table corresponding to Hispania is under discussion, as it has been reconstructed, and was therefore not employed in this study [22]. Furthermore, the authenticity of the so-called Itinerario de Barro has recently been proven, through the application of the thermoluminescence technique to four ceramic clay tablets [23]. They can therefore be applied to hypotheses for the identification of the *mansiones*. Moreover, the comments of the historians Polibio [24], Plinio [25], and Apiano [26] on the civitates of the Iberian Peninsula could be of great utility to any verification of the proposed locations. Isaac Moreno used toponymy, epigraphic and historic sources, aerial photography, and excavations to map Roman roads in the Spanish Autonomous Region of Castile and Leon [27].

The route on the Antonine Itinerary *via ab Asturica per Cantabriam Caesaraugustam* that corresponds to *Hispania* is analyzed. This route is one of the major challenges, especially as the locations proposed by the different authors disagree. Saavedra pioneered the study of Roman roads in *Hispania* and presented a map showing the routes of the Antonine Itinerary [28]. In the first section of that Roman road, the geographic distance between the towns of *Asturica* and *Rauda* is greater than the corresponding distance of the Itinerary [29]. Additionally, the well-known Roman city of Palencia, in the vicinities of the route, was never included in the Itinerary. As a result, Saavedra added a new *mansio* in Palencia [28]. Blázquez [29] proposed another path for the same via taking into consideration the distances cited in the Itinerary. Some theories have been suggested, to resolve the inconsistent distances corresponding with some *viae* of the Antonine Itinerary. Blázquez proposed that the length of the *milia passuum* in *Hispania* differed from the same yardstick in other parts of the Roman Empire [30]. Roldán theorized that when the place name was quoted in the accusative case, the site described by the itinerary lay off the road [31].

The location of the *mansio* Intercatia, essential to establish the alignment of the road, is the most controversial issue when determining the path of the first section of the route. Further information is available on that *mansio* than on other unknown *mansiones* on the road. It is cited in the narratives of ancient historians [24–26], in the Antonine Itinerary [8], in the *Ravennatis Anonimy Cosmographia* [9], and in Ptolemy's *Geographia* [10]. There is moreover archaeological evidence in the area under study [32,33]. However, *Intercatia* has been identified with numerous archaeological sites over an extensive area covering the provinces of Zamora, Palencia, and Valladolid: Medina de Rioseco [29,34], Villanueva del Campo [28], Castroverde de Campos [35], Villalpando [36], Aguilar de Campos [22], Valverde de Campos [37], Ceinos de Campos [38], Montealegre de Campos [39], and Paredes de Nava [40]. Most of them are based on the Antonine Itinerary. In the case of Montealegre de Campos, the *Ravennatis Cosmographia* has also been used. The hypothesis for its location corresponding to Paredes de Nava is based on epigraphic evidence uncovered there.

The research presented in this paper will describe a new method for computing the probability that the proposed locations are in fact a *mansio* on a route of a Roman itinerary in a province of the Roman Empire, based on the interpretation of the ancient geographical sources. The method will consist of three phases. In the first phase, the geographical domain of the most likely location of the mansion will be determined through reference to the implicit constraints of other well-known routes of the Itinerary, which will subsequently be considered. In the second phase, Ptolemy's coordinates will be correlated with those of the WGS 84 reference system by means of a statistical analysis. In the final phase, the results will be used to estimate the likelihood that the *mansio* under study was found at one particular location or another, and to assess the reliability of the suggested sites. It has been applied to the Roman city of *Intercatia* and the route *via ab Asturica per Cantabriam Caesaraugustam*. As a result, a new solution for the alignment of the route is proposed.

2. METHOD

The methodology for locating the *mansio* has three stages, which are explained in the next subsections.

2.1. Demarcation of the geographical domain

Firstly, the area of possible sites for a city on a route must be demarcated. The coastline and the provincial borders must be considered as boundaries of the geographical domains where locations either near the sea or on provincial frontiers are considered. Furthermore, a place name marked on one route and not on two adjacent routes, running more or less parallel, must be geographically bounded between those parallel routes. In other words, other routes mentioned on a Roman itinerary that run approximately parallel to the one under analysis, or that depart from the same city will help to demarcate the geographical domain for places located on the route under consideration.

2.2. Correlation between the coordinates of WGS 84 reference system and Ptolemy´s coordinates

As previously mentioned in the introduction, the coordinates cited in Ptolemy's *Geographia* are often inaccurate. We can only link coordinates from Ptolemy's gazetteer to place names that we know and then interpolate intermediate coordinates. Since the aim is to identify a place in a province of the Roman Empire, the selection of locations was restricted to those Ptolemy had included in his Geographia in the province under consideration [10], the geographic distribution of which have been fully authenticated and raise no controversy.

Having selected the locations, Ptolemy's coordinates of latitude and longitude were correlated with the WGS 84 reference system of coordinates by means of simple linear regressions.

Simple linear regression is a statistical technique employed for analyzing the relationship between a quantitative (non-qualitative) dependent variable (to be predicted), and a quantitative independent variable. The aim of this analysis is to use the independent variable, the value of which is known, to predict the dependent variable (response) [41]. A simple linear regression has the following expression [42]:

$$y = \beta_0 + \beta_1 \cdot x_1 \qquad (1)$$

where, y is the dependent variable, x_1 is an independent variable, β_0 is the intercept, and β_1 is the coefficient of the independent variable, i.e., the slope of the straight line.

The longitudinal and latitudinal coordinates of the WGS 84 reference system were respectively predicted by Ptolemy's own longitudinal and latitudinal coordinates [5].

In linear regressions, the coefficient of determination, R^2, is a descriptive measure of the global adjustment of the regression model, which quantifies on a scale from 0 to 1 the proportional variation of the dependent variable, explained by the independent variable included in the model. Moreover, the Fished–Snedecor distribution or F-statistic is used to identify confidence intervals at α% of significance level for the set of parameters, β_i, of the model; making it possible to verify the null hypothesis H_0: $\beta_1 = \beta_2 = \ldots = \beta_k = 0$. The Student-$T_i$ statistic for each of the coefficients is a test of the null hypothesis H_0: $\beta_i = 0$ for each individual coefficient. Furthermore, the linear regression model must meet the following conditions [41,42]:

- the linear relation between the dependent and the independent variable must be verifiable by means of the Pearson coefficient, R. If it is not linear, variable transformations can, in some cases, restore a degree of linearity;
- each observation must be drawn independently from the population, implying the independence of the errors between each other. This null hypothesis of no autocorrelation can be tested by means of the Durbin-Watson statistic, which ranges between 0 and 4, where a value of 2 implies total independence and where itis assumed that values between 1.5 and 2.5 indicate independent errors;
- the variance of errors must be equal across all levels, i.e., it must not be dependent on the observation. This hypothesis is called homoscedasticity. The absence of any observable pattern can be evaluated with a plot of the standardized predicted values against the standardized residuals;•the errors must follow a normal distribution, which can be verified by a Shapiro–Wilk normality test or by observing whether the errors can be plotted on a straight line in the normal probability plot.

2.3. Location probability distribution

Finally, as a result of the simple linear regressions obtained in the previous stage, the site of the city under study will follow a bivariate normal distribution [43], the center of which is located at the estimated value by means of the regression x^{city} and y^{city}:

$$f(x,y) = \frac{1}{2\pi\sigma_{LONG}\sigma_{LAT}} \exp\left\{-\frac{1}{2\sigma_{LONG}^2\sigma_{LAT}^2}\left(\sigma_{LAT}^2\left(x - x^{city}\right)^2 + \sigma_{LONG}^2\left(y - y^{city}\right)^2\right)\right\} \quad (2)$$

where x and y are the longitudes and latitudes in the WGS 84 system, σ^2_{LONG} and σ^2_{LAT} are the variances of the longitude and latitude, respectively, and $f(x,y)$ is the probability density function.

If Eq. (2) is expressed in a system (x',y') centred on ($x = x^{city}$, $y = y^{city}$), it yields,

$$(3) \qquad f(x',y') = \frac{1}{2\pi\sigma_{LONG}\sigma_{LAT}} \exp\left\{-\frac{1}{2\sigma_{LONG}^2\sigma_{LAT}^2}\left(\sigma_{LAT}^2 x'^2 + \sigma_{LONG}^2 y'^2\right)\right\}$$

The intersection lines of the density function and the planes parallel to $x'y'$ plane are ellipses centred on (0,0), the expression of which is:

$$(4) \qquad \frac{x'^2}{\sigma_{LONG}^2} + \frac{y'^2}{\sigma_{LAT}^2} = c^2$$

where, c is a parameter related to the height of the plane that intersects the surface $f(x',y')$.

Both the x' and the y' coordinates are the distances from each location to the ellipse center and are therefore random errors of that location. The probability that a Roman city can be found within the ellipse is as follows:

$$(5) \qquad P\left[\frac{x'^2}{\sigma_{LONG}^2} + \frac{y'^2}{\sigma_{LAT}^2} \leq c^2\right] = P[U \leq c^2]$$

where,

$$(6) \qquad U = \frac{x'^2}{\sigma_{LONG}^2} + \frac{y'^2}{\sigma_{LAT}^2}$$

It can be demonstrated that the variable U follows a Chi^2 distribution with two degrees of freedom:

$$(7) \qquad f(U) = \frac{1}{2}e^{-\frac{U}{2}}$$

Therefore, by combining Eq. (7) and Eq. (5), we have:

$$(8) \qquad P[U \leq c^2] = \int_0^{c^2} \frac{1}{2}e^{-\frac{U}{2}} du = 1 - e^{-\frac{c^2}{2}}$$

Eq. (8) can be used to calculate the confidence intervals (CI). The value, c^2, that corresponds to a probability, P, can be estimated, so that in consequence the region in the ellipse is the CI corresponding to a probability P.

3. RESULTS

The proposed methodology was applied to locate the city of *Intercatia*, marked on the Roman road via *ab Asturica per Cantabriam Caesaraugustam* in the Antonine Itinerary [8]. In Table 1, the first section of the itinerary under consideration is shown. The numbering of Blázquez [29] is set in parentheses next to the former name of the route. For simplicity, the same numbering used in that work will be reproduced here to name the *viae*. Wherever ancient place names are located, the corresponding current names have been added in parentheses. The distances between consecutive places are expressed in *milia passuum* (1480 m). When the values achieved from the manuscripts disagree, more than one value is given [29].

Table 1. *Ab Asturica per Cantabriam Caesaraugustam* in the Antonine Itinerary

***Ab Asturica per Cantabriam Caesaraugustam* (Route 27)**
Asturica (Astorga)
Brigeco: 40
Intercatia: 20
Tela: 22
Pintiam: 24
Raudam (Roa): 11
Cluniam (Peñalba del Castro): 26, 16

The solutions for the alignment of Route 27 proposed by different authors are shown in Fig. 1. The white circles denote stations. The area under study in mainland Spain is colored in grey in the above map of Spain. The administrative borders of the current provinces are shown on the maps.

The study of Route 27 is complex, because of the high number of consecutive unknown stations: *Brigeco*, *Intercatia*, *Tela* and *Pintiam*. Furthermore, the geographical distance between *Asturica* and *Rauda* is 22 km greater than the distance that can be derived from the itinerary. It means that one or various stations are missing, or a notable error has been made when copying some distance in the manuscript. Consequently, the identification of *Intercatia* would be essential for a precise analysis of the alignment of this route. In Fig. 1 the black circles denote the locations suggested by other authors. The locations proposed by Saavedra [28], Wattenberg [22] and Blázquez [29] are Villanueva del Campo, Aguilar de Campos, and Medina de Rioseco, respectively. The place name that appears in the text [29] is Medina del Campo, although it is considered a misprint, because Medina de Rioseco better fits the path suggested by the author. Another key site for the determination of the alignment of the route is the ford that connects the Tierra de Campos and Cerrato regions. Saavedra affirmed that it was in the city of Palencia, mentioning a bridge of Roman origin over the Carrión River named Puentecillas. Wattenberg [22] and Blázquez [29] hypothesized that it was over the Pisuerga River in Cabezón de Pisuerga and Dueñas, respectively.

3.1. Geographical domain

As indicated in the first stage of the proposed methodology, the geographical domain of the possible location of the city must be restricted by means of adjacent routes to the one under analysis. In Table 2, the itineraries of the two adjacent routes to Route 27are shown: via *Ab Emerita Caesaraugustam* (Route 24, following the criteria of Blázquez [29]) and via *Ab Asturica Tarracone* (Route 32). The stretches of these two routes that are close to Route 27 are shown.

Table 2. Routes next to Route 27 in the Antonine Itinerary

***Ab Emerita Caesaraugustam* (Route 24)**	***Ab Asturica Tarracone* (Route 32)**
Salmatice (Salamanca)	*Asturica* (Astorga)
Sibariam: 21	*Vallata*: 16
Ocelo Duri: 21	*Interamnio*: 13, 16
Albocela: 22, 16	*Palantia*: 14
Amallobriga: 22, 27	*Viminiacio*: 31
Septimanca (Simancas): 14, 24	*Lacobriga* (Carrión de los Condes): 10
Nivaria : 12, 16, 22	*Dessobriga* (Osorno): 15
Cauca (Coca): 22	*Segisamone* (Sasamón): 15
	Deobrigula: 21, 15
	Tritium: 21
	Virovesca (Briviesca): 11

In Fig. 2, routes 24, 32 and 27 are shown. The lines corresponding to routes 24 and 32 are the borderlines of the geographical domain where Route 27 and *Intercatia* must be located. The same figure also shows that the criteria in the present study are also reflected in other studies [22,28,29].

In the case of Route 32, only the layout proposed by Saavedra is shown, because no protracted discrepancies have been reported over the distances and the whereabouts of certain intermediate stations have been clearly established.

However, in the case of the Route 24, there are several options. The stations of *Ocelum Duri*, *Albocela*, and *Amallobriga* have yet to be identified. Nevertheless, the sites suggested by the three authors [22,27,28] for the first two are next to each other. However, in the case of *Amallobriga*, the proposed locations are geographically dispersed. Saavedra [28] and Wattenberg [22] located it at Villavieja de Campos and Torrelobatón, respectively. Moreno Gallo [27] placed it at Montealegre de Campos, because a *tessera hospitalis* at that village with the Latin inscription *Amallobrigenses* was discovered in 1985. Moreover, a Roman road connected it to Toro. Gómez Fraile [39] located it at Tiedra. The path is known from Simancas (*Septimanca*), because enough intermediate stations, such as Coca

(*Cauca*) and Alcalá de Henares (*Complutum*), have been identified.

Consequently, the line corresponding to Route 32 and Saavedra's solution for Route 24 have been taken to represent the border-lines of the area in which both Route 27 and *Intercatia* are located. Saavedra's suggestion has been adopted, because it is the most conservative.

3.2. Correlation between the coordinates of the WGS 84 reference system and Ptolemy´s coordinates

Simple linear regressions were obtained between Ptolemy's coordinates of locations in the Iberian Peninsula and their corresponding coordinates in the WGS 84 reference system, in order to establish the probability distributions related to the ancient site of *Intercatia*. As the aim was to identify a place in Hispania, all of the (48) locations taken from Ptolemy's *Geographia* [10] were in the Iberian Peninsula, as listed in Table 3.

Table 3. Locations in the Iberian Peninsula employed to correlate Ptolemy's coordinates and those of WGS reference system.

***Roman name* (Present name)**	***Roman name* (Present name)**	***Roman name* (Present name)**
Corduba (Córdoba)	*Complutum* (Alcalá)	*Malaca* (Málaga)
Hispalis (Sevilla)	*Turiasso* (Tarazona)	*Ossonoba* (Faro)
Gadira (Cádiz)	*Caesaraugusta* (Zaragoza)	*Ebura* (Ébora)
Emerita (Mérida)	*Saguntum* (Sagunto)	*Valentia* (Valencia)
Salmatice (Salamanca)	*Pompaelon* (Pamplona)	*Iacca* (Jaca)
Cartago Nova (Cartagena)	*Calagurris* (Calahorra)	*Gracuris* (Alfaro)
Ilicis (Elche)	*Osca* (Huesca)	*Aeminium* (Coimbra)
Legio Germanica (León)	*Ilerda* (Lérida/Lleida)	*Barcinon* (Barcelona)
Asturica (Astorga)	*Gerunda* (Gerona/Girona)	*Segobriga* (Saelices)
Lacobriga (Carrión de los Condes)	*Tarraco* (Tarragona)	*Aqua Flavia* (Chaves)
Cauca (Coca)	*Bracaraugusta* (Braga)	*Lucentum* (Alicante)
Viroesca (Briviesca)	*Lucus Augusti* (Lugo)	*Iria Flavia* (Padrón)
Clunia (Peñalba)	*Oliosipon* (Lisboa)	*Onuba Listuria* (Huelva)
Uxama Argaela (Osma)	*Italica* (Santiponce)	*Juliobriga* (Retortillo)
Numantia (Garray)	*Pax Julia* (Beja)	*Dianium* (Denia)
Toletum (Toledo)	*Emporia* (Ampurias)	*Astigi* (Écija)

3.2.1. Correlation of the longitude coordinates

Using the above locations selected in the Iberian Peninsula and by means of a simple linear regression, a relation was established between Ptolemy's longitudinal coordinates, $Long^{PTO}$, and the longitudes in the WGS 84 system, $Long^{WGS\ 84}$, through the following equation, Eq. (9):

(9) $$Long^{WGS\,84} = -12.257 + 0.7884 \cdot Long^{PTO}$$

Both variables are expressed in decimal form, i.e. 30 minutes equal to 0.5 degrees. Eq. (9) has a coefficient of determination (R^2) of 0.9021, which means that over 90% of the variability of the values could be explained by the proposed model (Fig. 3). Moreover, as shown in Table 4, the Fisher–Snedecor (F-statistic) distribution indicated that the relationship was significant ($P < 0.001$) and the analyses of the coefficients with the Student t-test established that they were real and different from 0, with a significance of over 99%. On the other hand, the Durbin–Watson statistic was 2.067, which verified the independence of errors and homoscedasticity, as shown by Fig. 4a, confirming that the errors followed no established pattern. The significance of the Shapiro–Wilk normality test results was 0.251, over the established significance of 0.05, and the normal probability distribution plot of the standardized residuals (Fig. 4b) showed a good fit with the main diagonal, it can therefore be assumed that the errors followed a normal distribution.

Table 4. Analysis of variance of the linear regression between the longitude coordinates in the WGS 84 and Ptolemy's reference system, Eq. (9).

Analysis of variance								
Source	**Sum of squares**	**Degrees of freedom**	**Mean square**	**F value**	**p-value**	**Durbin-Watson**	**Shapiro-Wilk**	**p-value**
Model	460.752	1	460.752	423.829	< 0.001	2.067	0.970	0.251
Error	50.007	46	1.0871		**Root mean square error**		**R**	**R^2**
Corrected total	510.759	47			1.0426		0.950	0.902
Parameter estimates								
Variable	**Coefficient (βi)**	**Standard error**	**t value**	**p-value**	**95% confidence limits**			
					Lower bound		**Upper bound**	
Intercept	-12.257	0.446	-27.497	< 0.001	-13.154		-11.360	
Lat^{PLO}	0.788	0.038	20.587	< 0.001	0.711		0.865	

3.2.2. Correlation of the latitude coordinates

Similarly, Ptolemy's latitudinal coordinates, Lat^{PTO}, and those of the WGS 84 reference system, $Lat^{WGS\ 84}$, both expressed in decimal form, when correlated by means of a simple linear regression, generated the following Eq. (10):

(10) $$Lat^{WGS\,84} = 5.1308 + 0.8606 \cdot Lat^{PTO}$$

Eq. (10) has a R^2 value of 0.907, highlighting a very good linear relationship between both variables (Fig. 5). The analysis of variance of the model is shown in Table 5. The Fisher-Snedecor distribution indicated that it was a linear relationship and that all the coefficients could not simultaneously be 0 ($P < 0.001$). The Student t-test results showed that the coefficients were different from 0 at a confidence level of over 99% ($P < 0.01$). The value of the Durbin–Watson statistic (1.705) confirmed the independence of errors. Fig. 6a verified the homoscedasticity of the model, as no pattern was observed. The normal distribution of errors was verified by the P-value of the Shapiro–Wilk statistic, 0.051, over the established significance level (0.05), and by the adjustment to the main diagonal of the standardized errors in the normal probability plot (Fig. 6b).

Table 5. Analysis of variance of the linear regression between the latitude coordinates in the WGS 84 and Ptolemy's reference systems, Eq. (10).

Analysis of variance								
Source	**Sum of squares**	**Degrees of freedom**	**Mean square**	**F value**	**p-value**	**Durbin-Watson**	**Shapiro-Wilk**	**p-value**
Model	177.321	1	177.321	448.689	< 0.001	1.705	0.953	0.051
Error	18.179	46	0.395		**Root mean square error**		**R**	**R^2**
Corrected total	195.500	47			0.62865		0.952	0.907
Parameter estimates								
Variable	**Coefficient (βi)**	**Standard error**	**t value**	**p-value**	**95% confidence limits**			
					Lower limit		**Upper limit**	
Intercept	5.131	1.666	3.080	0.003	1.777		8.484	
Lat^{PLO}	0.861	0.041	21.182	< 0.001	0.779		0.942	

3.3. Probability distribution of *Intercatia*

Finally, once the relations between the longitudinal references both of Ptolemy and of the WGS 84 reference system coordinates had been established, the coordinates for *Intercatia* could be estimated in the WGS 84 system: $LONG^{WGS} = -4.18$ and $LAT^{WGS} = 42.50$ by substituting the Ptolemaic coordinates [10]: $LONG^{PTO} = 10.25$ and $LAT^{PTO} = 43.42$ in Eq. (9) and Eq. (10). These coordinates corresponded to the village of Tagarrosa in the province of Burgos.

The longitudinal and latitudinal values of the WGS 84 system follow normal distributions, the average values of which are those corresponding to the regression lines and the standard deviations are $\sigma_{LONG} = 1.0426$ and $\sigma_{LAT} = 0.62865$, respectively, as can be seen from Tables 4 and 5.

Therefore, the site of *Intercatia* follows a bivariate normal distribution the center of which is located at the estimated value bymeans of the regression $x^{inter} = -4.18$ and $y^{inter} = 42.50$.

Eq. (8) was used to calculate the confidence intervals (CI), which in this study were CI of 50, 70, and 95%. In Fig. 7 the ellipses of those intervals and the borderlines of the geographical domain obtained in Section 3 are shown. Moreover, a triangle represents the center of the ellipses and the grey circles represent the start and finish of the first section of the route. The black circles correspond to the locations proposed by the other authors [22,28,29].

The regions in the ellipses were the 50, 70, and 95% confidence intervals, related to high, medium and low probabilities, respectively. In Fig. 8 the intersections between the geographical domain and the ellipses corresponding to the confidence intervals are shown. Additionally, the locations proposed by other authors, which are cited in Section 1, are marked in white circles.

Paredes de Nava and Montealegre de Campos are the only sites in the 50% interval. Ceinos de Campos is next to although slightly outside the ellipse of 50%. According to one tradition [36], Villalpando is the site of *Intercatia*, but it is the least likely option, because it is outside the 70% CI. Among the others within the 70% CI, Ceinos de Campos is perhaps the most likely. From among the proposed sites that could be considered towns, Paredes de Nava would be the most likely candidate. The probability value must be lowered by as much as 23% for Paredes de Nava to fall outside the confidence region. The difference in terms of probability between Paredes de Nava and the other proposals was significant (between 27% and50%).

4. DISCUSSION

As can be seen in Fig. 8, Paredes de Nava and Montealegre de Campos are the only towns within the 50% CI. It is worth noting that an inscription with a demonym corresponding to *Amallobriga*, which as mentioned was a village on Route 24, was discovered in the *tessera* of Montealegre de Campos. Moreover, the demonym *intercatiensi* is written on another *tessera* at Paredes de Nava. Despite the discovery, in 1869, of the latter demonym [32], the suggestion of Paredes de Nava was not taken into account, because it appeared not to fit on Route 27. A second *tessera* from Paredes de Nava with the same demonym, published in 1999, is recorded on the epigraphic database of the University of Heidelberg [33]. According to recent archaeological postulations [44], based on the texts of both *tesserae*, Paredes de Nava might have been the site of *Intercatia*. The results of the statistical model agreed with the archaeological evidence.

The authors are of the opinion that Tierra de Campos and the Comarca del Cerrato regions were connected by a ford over the Carrión River, at the city of Palencia. This hypothesis matches Saavedra's proposal. In fact, no great adaptation of Saavedra's layout is necessary for Route 27 to pass through *Intercatia*. In the alignment suggested by the authors, which passes through Paredes de Nava, the route would border the Campos Sea by the north, as can be seen in Fig. 9. In this figure, the flooding area of this ancient lake [45] has been represented, on the basis of the plans of the Canal of Castile, drawn up in 1806 [46].

5. CONCLUSIONS

The identification of Roman roads and their corresponding towns and cities will greatly assist their conservation. Furthermore, the correct identification of the routes is essential to obtain accurate results in future studies on Roman transport networks their connectivity, functionality and impact over time [3].

The alignments of Roman roads that remain unknown to this day can be determined by pinpointing some intermediate stations. A new method has been presented for assessing the probability that various suggested sites were stations on the Roman roads cited in the *Cosmographia*. The procedure could feasibly be applied to any Roman Empire province.

The main contribution of the present procedure with regard to previous proposals [14–17] is its use of probability theory. Firstly, the mathematical expressions that relate Ptolemy's coordinates and their current counterparts were obtained by means of linear regressions, while these relations are described in other studies on the basis of the radius of the Earth and the position of the reference parallel and meridian. Secondly, a bivariate normal distribution isused to determine the probabilities corresponding to the regions under analysis. The areas corresponding to the confidence intervals are limited by using geographical and political constraints, as well as those imposed by other routes from the Roman itineraries.

The proposed methodology is interesting for problematic cases when other alternative techniques [14] cannot be used, because the number of unknown consecutive stations is considerable. Moreover, it is appropriate when the proposed locations are geographically dispersed, because the probability values would otherwise be very similar, due to the accuracy level of the *Cosmographia*. The accuracy level is determined by the standard latitudinal and longitudinal deviations of the province under study.

The method has been applied to locate *Intercatia* marked on Route 27 of the itinerary. The first stretch of that route was the most controversial in *Hispania*, because of the various unknown consecutive stations and the wide variation between the proposals from different authors. The alignment of the route depends significantly on the location of this *mansio*, as can be seen from the proposed sites from the literature review.

Since Paredes de Nava is the most likely placement and two *tesserae* with the inscription *intercatiensi* have been discovered there, its identification with *Intercatia* has been assumed. Consequently, a new suggestion for the alignment of Route 27 has been proposed by redrawing the layout that Saavedra proposed, so that it passes through Paredes de Nava, bordering the Campos Sea (former wetlands) to the north.

ACKNOWLEDGEMENTS

The first author wants to thank María del Mar Torrado and Javier Romera because of their support and confidence during the preparation of this work.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

DECLARATIONS OF INTERESTS: None

REFERENCES

[1] Moreno Gallo, I., (2018). Vías romanas: tecnología romana del transporte terrestre (y II). *Cimbra*, 413, 36-49

[2] Sánchez, A., (2016). La puesta en valor de las calzadas romanas. Estudio, excavación y recuperación social del patrimonio arqueológico inmueble. *Cuadernos de Arqueología de la Universidad de Navarra*, 24, 215-231. doi:10.15581/012.24.003

[3] *Mercator-e Project*, http://fabricadesites.fcsh.unl.pt/mercator-e/

[4] De Soto, P. (2016). Mercator-e. A new diachronic analysis on historical transport infrastructures in the Iberian Peninsula. *Bridging social and geographical space through networks. Topoi – Freie Universität.* Berlin

[5] De Soto, P. (2017). The Mercator-e project. Analysing the Iberian Peninsula seaports as ancient main penetration vertices through the inland territories. *North meets East. Recent research on ancient ports IV.* Hamburg

[6] De Soto, P. (2018). The evolution of the Iberian historical transportation networks through very detailed digitised models. *Digital Approaches to Cartographic Heritage. 13th International Conference of the International Cartographic Association Commission on Cartographic Heritage into the Digital*. Madrid

[7] De Soto, P. Network Analysis to Model and Analyse Roman Transport and Movility, in: Verhagen, P., Joyce, J., and Groenhuijzen, M.R. (Eds.). *Finding the Limits of the Limes*, Springer, Amsterdam, 2019, pp. 271-289

[8] Pinder, M. and Parthey, G. (1848). *Itinerarium Antonini Augusti et Hierosolymitanum*. Berolini: inpemsis Friderici Nicolai, https://archive.org/details/bub_gb_uAVYqu_XP7IC/page/n3

[9] Ravenna, G., Pinder, M. and Parthey, G. (1860). *Ravennatis Anonymi Cosmographia et Gvidonis Geographica*. Berolini: in aedibvs Friderici Nicolai, https://archive.org/details/ravennatisanony02ravegoog/page/n8

[10] Ptolomeo, C. *Cosmographia*, http://bdh-rd.bne.es/viewer.vm?id=0000061306&page=1

[11] Verbrugghe, G., De Clercq, W. and Van Eetvelde, V. (2017). Routes across the Civitas Menapiorum: using least cost paths and GIS to locate the Roman Roads of Sandy Flanders. *J. Hist. Geogr.*, 57, 76-88

[12] Fonte, J., Parcero-Oubiña, C. and Costa-García, J. M. (2017). A GIS-based analysis of the rationale behind

roman roads. The case of the so-called via XVII (NW Iberian Peninsula). *Mediterranean Archaeology and Archaeometry*, 17(3), 163-189

[13] Güimil-Fariña, A. and Parcero-Oubiña, C. (2015). "Dotting the joins": a non-reconstructive use of Least Cost Paths to approach ancient roads. The case of the Roman roads in the NW Iberian Peninsula. *J. Archaeol. Sci.,* 54, 31-44

[14] Verhagen, J. G. M. (2014). Using distances to identify Roman places in *Itineraria* – A case study on the Lower Rhine *Limes. Archäologisches Korrespondenzblatt,* 44(4), 543-562

[15] Darcy, R. and Flynn, W. (2008). Ptolemy´s map of Ireland: a modern decoding. *Irish Geography*, 41(1), 49-69

[16] Livieratos, E., Tsorlini, A., Boutoura, C. and Manoledakis, M. (2008). Ptolemy´s *Geographia* in digits. *e-Perimetron,* 3(1), 22-39

[17] Tsorlini, A. (2009). Spatial distribution of Ptolemy´s *Geographia* coordinate differences in North Mediterranean eliminating systematic effects. *e-Perimetron,* 4(4), 247-266

[18] Solana, J.M. (1972). Ensayos sobre el valor del grado Ptolemaico. *Hispania Antiqua*, 2, 69-76

[19] Montero, J. (1991). *Carpetanos y vettones en la Hispania de Ptolomeo. Ciudades y vías romanas en Carpetania y Vettonia en época altoimperial.* Universidad Complutense de Madrid

[20] Urueña, J. (2014). El método cartográfico de Ptolomeo: análisis del sistema de localización utilizado en la *Geographia* para la ubicación de las poblaciones del interior de la Península Ibérica. *Palaeohispanica*, 14, 153-185

[21] Defaux, O. (2017). *The Iberian Peninsula in Ptolemy's Geography. Origins of the coordinates and textual history*. Berlin: Edition Topoi

[22] Wattenberg, F. (1959). *La región vaccea. Celtiberismo y romanización en la cuenca media del Duero*. Madrid: Bibliotheca Præhistorica Hispana

[23] Fernández, C., Morillo, A. and Gil, F. (2012). El itinerario de barro. Cuestiones de autenticidad y lectura. *Zephyrus*, LXX, 151-179.

[24] Polibio and Balasch, M. (1981). *Historias: libros V-XV*. Madrid: Gredos

[25] Plinio, C., Hernández, F. and de la Huerta, J. (1999). *Historia natural de Cayo Plinio Segundo*. Madrid: Visor Libros, S.L.

[26] Apiano and Sancho, A. (1995). *Sobre Iberia. Sobre Africa*. Barcelona: Planeta Deagostini

[27] Moreno Gallo, I. (2017). *Vías Romanas en Castilla y León,* http://www.viasromanas.net/

[28] Saavedra, E. (1862). *Mapa itinerario de la España romana con sus divisiones territoriales,* http://bibliotecadigital.rah.es/dgbrah/es/consulta/registro.cmd?id=28968

[29] Blázquez, A. (1892). Nuevo estudio sobre el "Itinerario" de Antonino. *Boletín de la Real Academia de la Historia*, 21, 54-128

[30] Blázquez, A. (1899). La milla romana. *Boletín de la Real Academia de la Historia*, 34, 25-52

[31] Roldán, J.M. (1966). Sobre los acusativos con "ad" en el Itinerario de Antonino. *Zephyrus: Revista de*

prehistoria y arqueología, 17, 109-120

[32] Hübner, E.W.E. (1869). *Corpus Inscriptionum Latinarum. Inscriptionum Hispaniae Latinarum Supplementum*. Berolini apud Georgium Reimerum, https://archive.org/details/bub_gb_Cm0_AQAAMAAJ/page/n3

[33] Heidelberg Academy of Sciences and Humanities. *Epigraphic Database Heidelberg,* https://edh-www.adw.uni-heidelberg.de/home

[34] Ceán, J.A. (1832). *Sumario de las antigüedades romanas que hay en España, en especial las pertenecientes á las Bellas Artes*. Madrid: publicase de Real Orden

[35] Gómez-Moreno, M. (1927). *Catálogo monumental de España. Provincia de Zamora*. Madrid: Ministerio de Instrucción Pública y Bellas Artes

[36] Negrete, J. (2018). *La conquista romana de Hispania*. Madrid: La esfera de los libros

[37] Martino, D. (2004). *Las ciudades romanas de la meseta norte de la Península Ibérica*: identificación, estatuto jurídico y oligarquías (SS. I-III D.C.). Universidad Complutense de Madrid

[38] Brezmes, M.A. (2014). *Intercatia*, el estado de la cuestión. *Investigaciones arqueológicas en el valle del Duero*, 2, 257-265

[39] Gómez Fraile, J.M. (1998). Intercatia y Amallobriga en los itinerarios antiguos de la Península Ibérica. Trabajo presentado en el *IV Congreso Internacional de Caminería Hispánica*, Guadalajara

[40] Hernández Guerra, L. (2010). Consideraciones sobre la ubicación de dos ciudades vacceas, *Intercatia* y *Pallantia*. *Dialéctica histórica y compromiso social*, 2, 961-980

[41] Pérez López. C. (2014). Técnicas estadísticas predictivas con IBM SPSS. Ibergaceta Publicaciones, S.L., Madrid, Spain.

[42] Montgomery, D.C., Peck, E.A., Vining, G.G. (2012). Introduction to Linear Regression Analysis (5th Ed). John Wiley & Sons, Hoboken, NJ, USA

[43] Blasco, A. and Pérez-Díaz, S. (2015). *Modelos aleatorios en Ingeniería*. Madrid: Ediciones Paraninfo

[44] Abarquero Moras, F.J., Gutiérrez Pérez, J. and Pérez Rodríguez, F.J. El yacimiento vacceo-romano de La Ciudad, Paredes de Nava, Palencia, in: Sanz Míguez, C. and Blanco García, J.F. (Eds.). *Novedades arqueológicas en cuatro ciudades vacceas: Dessobriga, Intercatia, Pintia y Cauca*, Centro de Estudios Vacceos "Federico Wattenberg", Valladolid, 2018, pp. 75-92

[45] Ruiz Monge, J. J. (2018). *El mar que hubo en Tierra de Campos. La desecación de la laguna de la Nava.* Palencia: José Javier Ruiz Monge (Ed.)

[46] Helguera, J. (1992). El *Canal de Castila. Cartografía de un proyecto ilustrado*. Madrid: Centro de Estudios y Experimentación de Obras Públicas

FIGURE CAPTIONS

Figure 1. Solutions for the Route 27 between *Asturica* (Astorga) and *Cluniam* (Peñalba del Castro).

Figure 2. Solutions for routes 24, 27 y 32.

Figure 3. Longitude coordinates in Ptolemy's and WGS 84 reference systems of the 48 selected locations in the Iberian Peninsula

Figure 4. Error analysis in Eq. (2): a) scatter plot of the observed standardised residuals vs. predicted standardised residuals, b) adjustment of errors to standardised normal distribution

Figure 5. Latitude coordinates in Ptolemy's and WGS 84 reference systems of the 48 selected locations in the Iberian Peninsula

Figure 6. Error analysis in Eq. (3): a) scatter plot of the standardised predicted residuals vs. standardised residuals, b) adjustment of errors to standardised normal distribution

Figure 7. Geographical domain and error ellipses for locating *Intercatia*.

Figure 8. Confidence intervals and suggested locations for *Intercatia*.

Figure 9. Suggested layouts for the Route 27 between *Asturica* and *Rauda*.

TABLE CAPTIONS

Table 1. *Ab Asturica per Cantabriam Caesaraugustam* in the Antonine Itinerary

Table 2. Routes next to Route 27 in the Antonine Itinerary

Table 3. Locations in the Iberian Peninsula employed to correlate Ptolemy's coordinates and those of WGS reference system.

Table 4. Analysis of variance of the linear regression between the longitude coordinates in the WGS 84 and Ptolemy's reference system, Eq. (9).

Table 5. Analysis of variance of the linear regression between the latitude coordinates in the WGS 84 and Ptolemy's reference systems, Eq. (10).

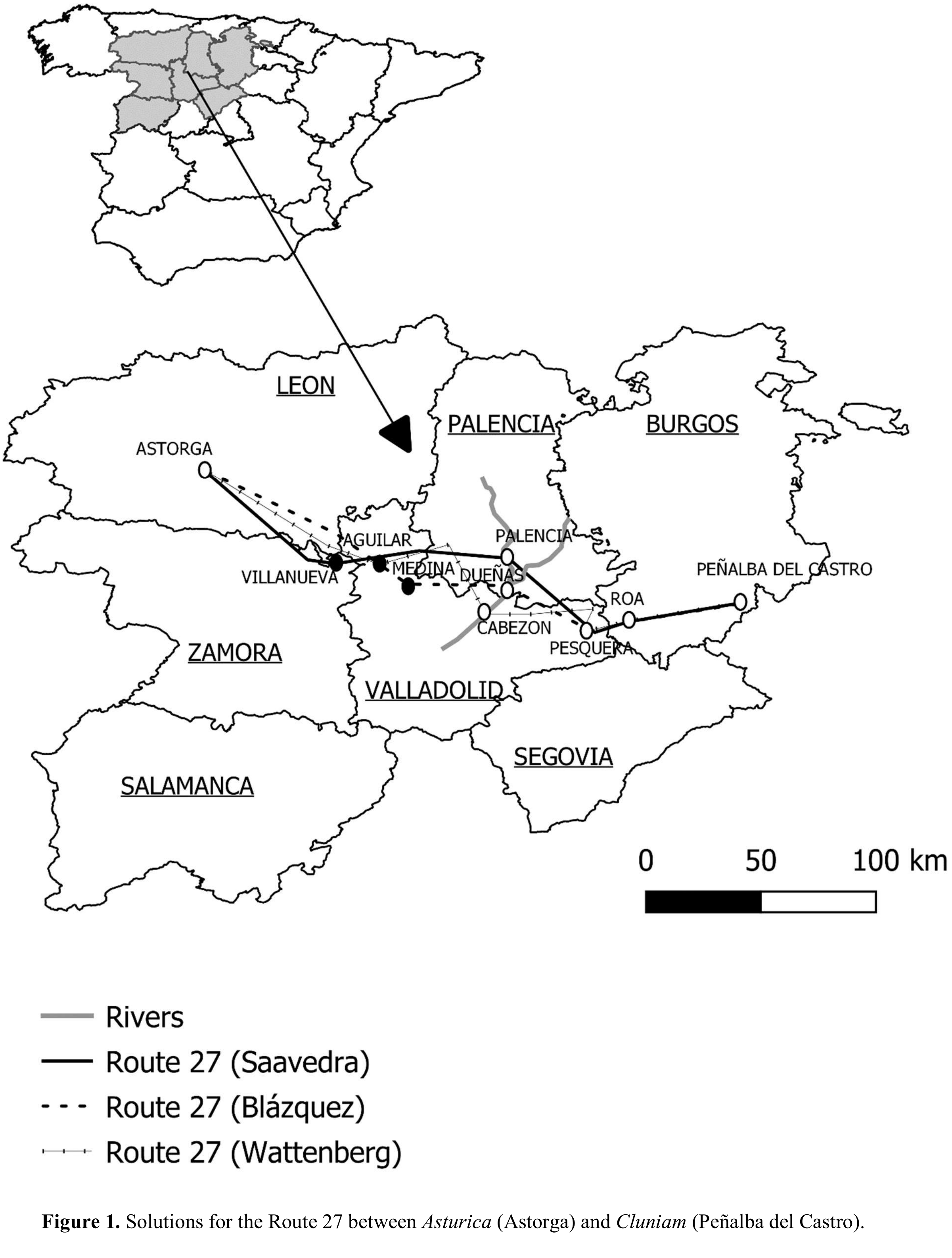

Figure 1. Solutions for the Route 27 between *Asturica* (Astorga) and *Cluniam* (Peñalba del Castro).

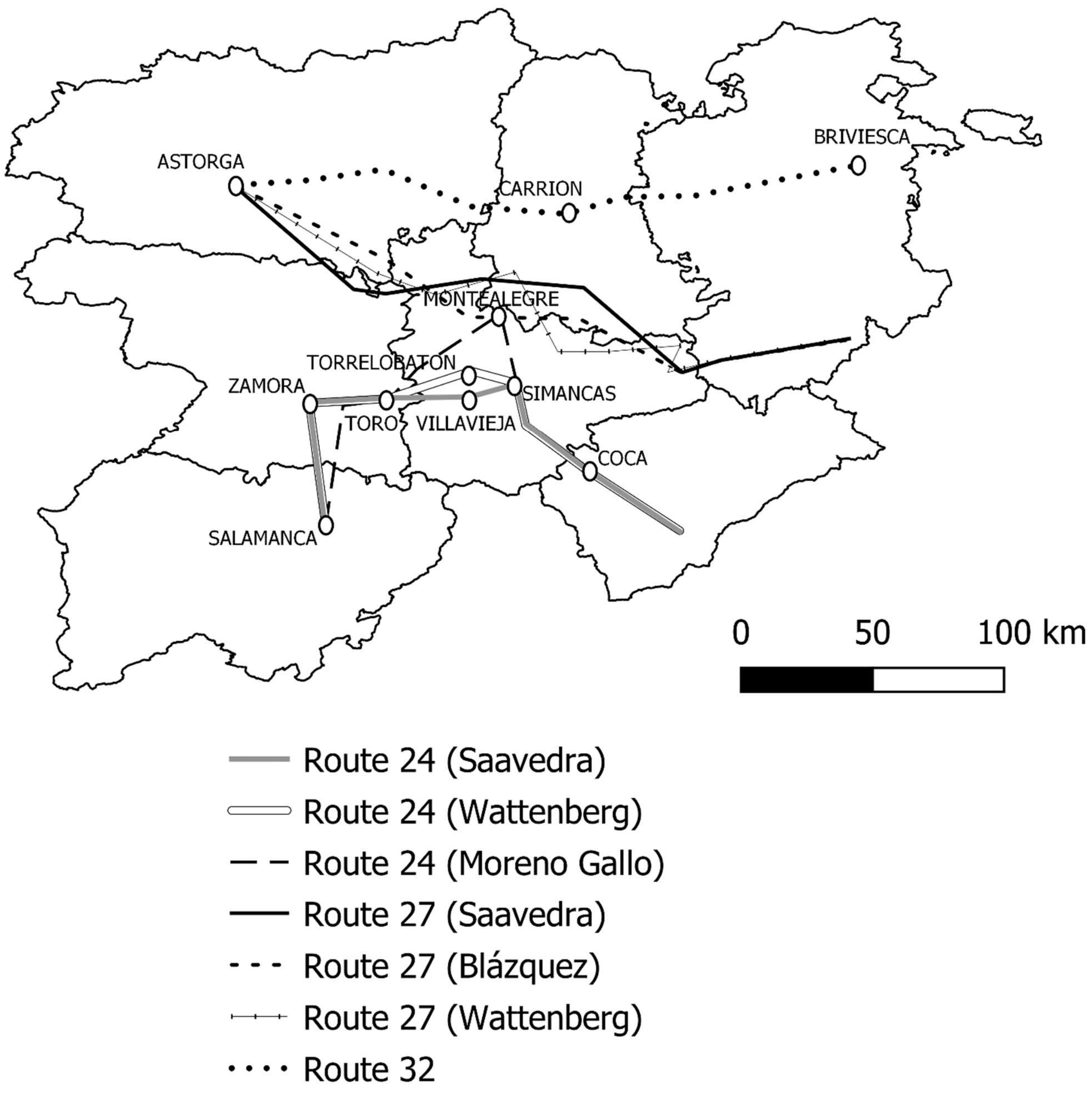

Figure 2. Solutions for routes 24, 27 y 32.

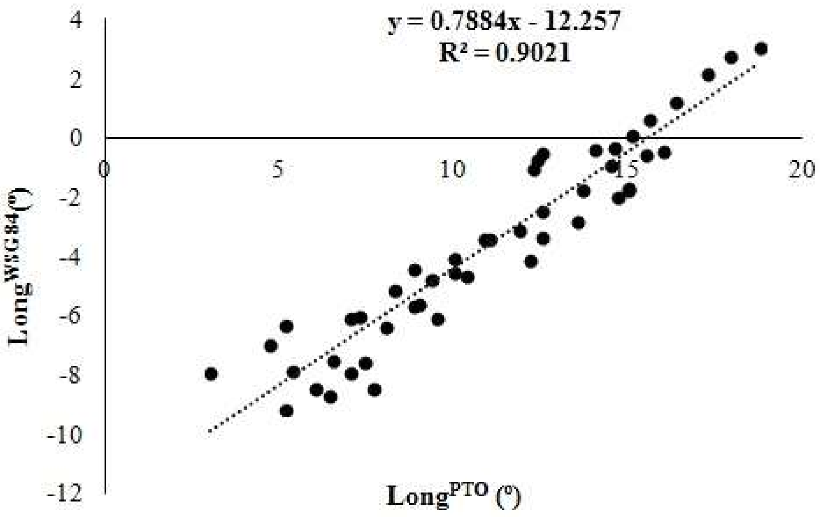

Figure 3. Longitude coordinates in Ptolemy's and WGS 84 reference systems of the 48 selected locations in the Iberian Peninsula

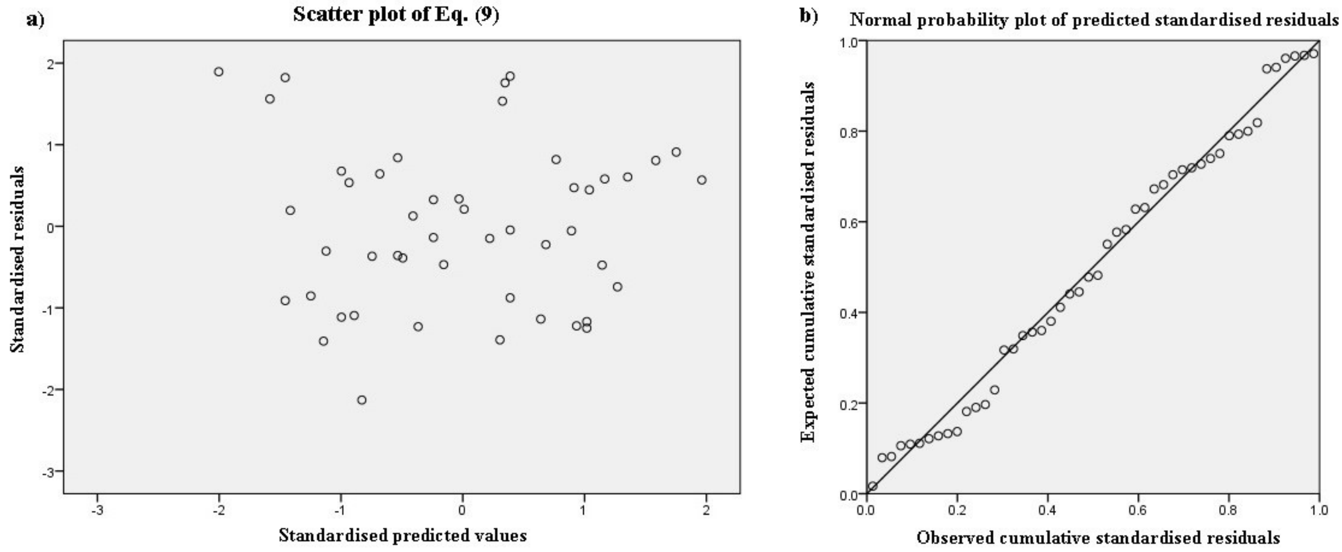

Figure 4. Error analysis in Eq. (9): a) scatter plot of the standardised predicted residuals vs. standardised residuals, b) adjustment of errors to standardised normal distribution

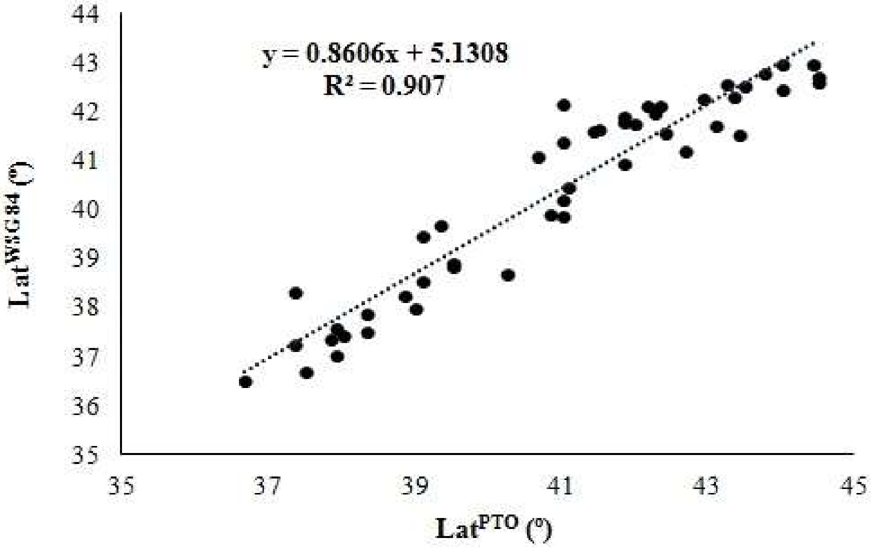

Figure 5. Latitude coordinates in Ptolemy's and WGS 84 reference systems of the 48 selected locations in the Iberian Peninsula

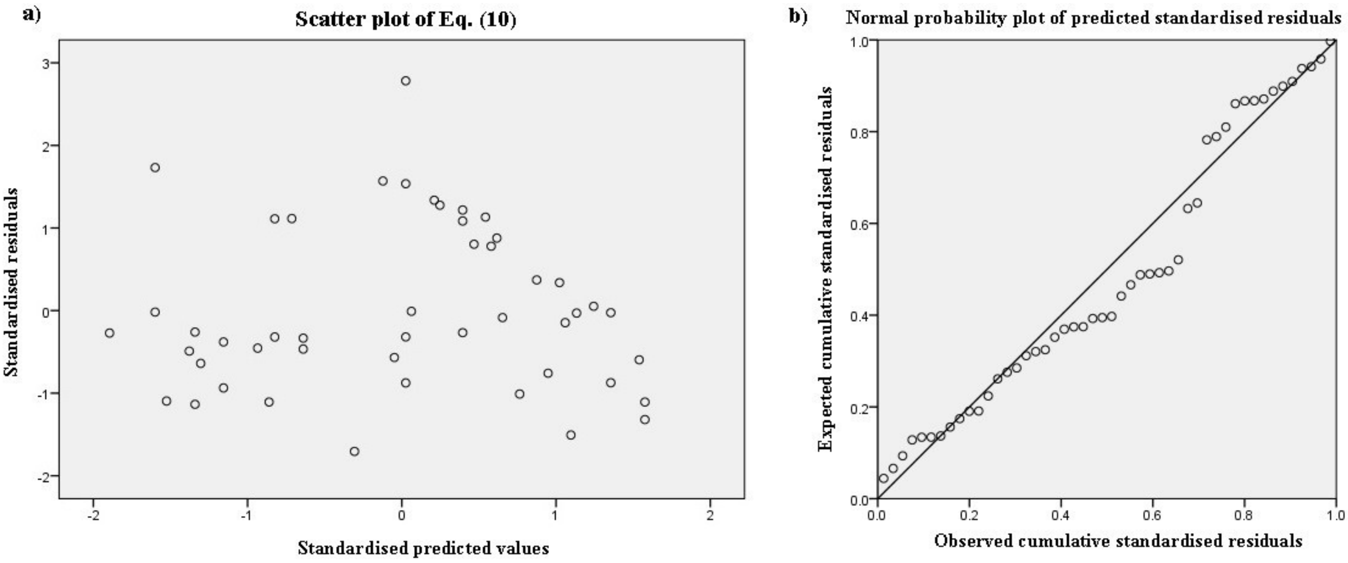

Figure 6. Error analysis in Eq. (10): a) scatter plot of the standardised predicted residuals vs. standardised residuals, b) adjustment of errors to standardised normal distribution

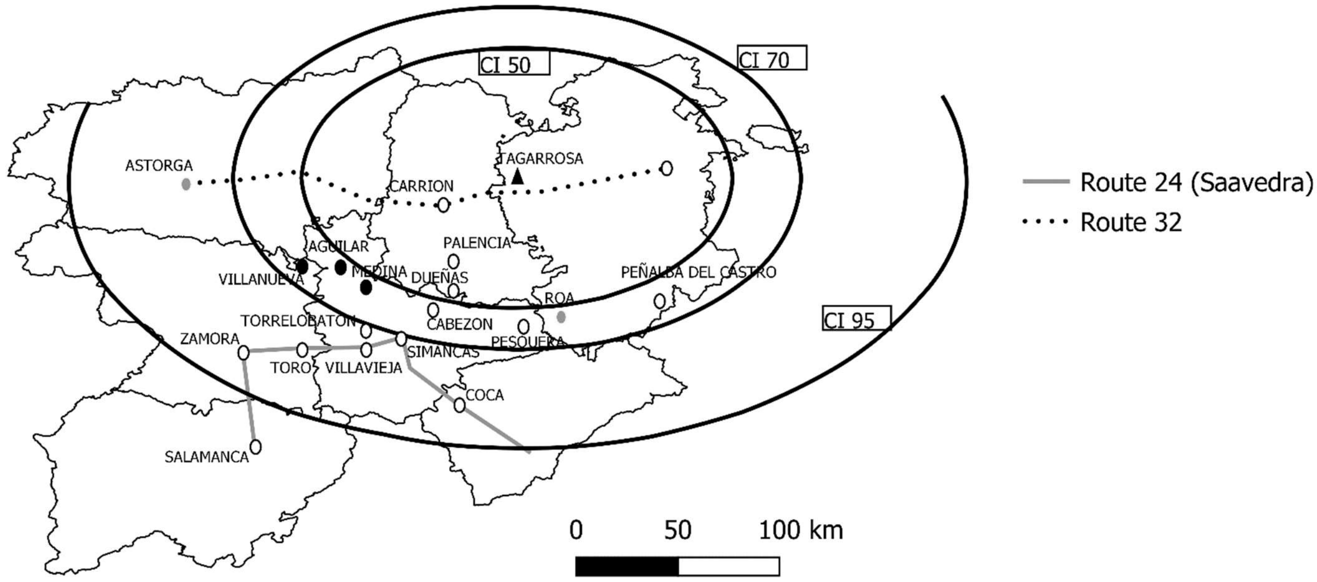

Figure 7. Geographical domain and error ellipses for locating *Intercatia*.

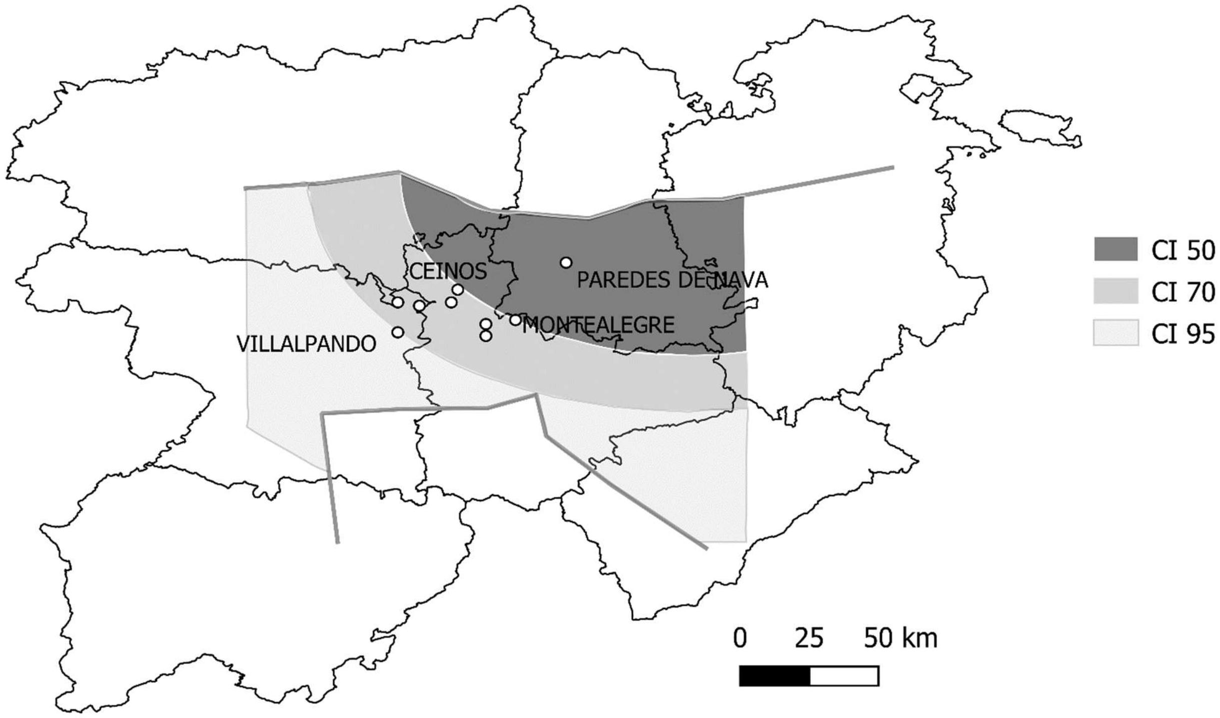

Figure 8. Confidence intervals and suggested locations for *Intercatia*.

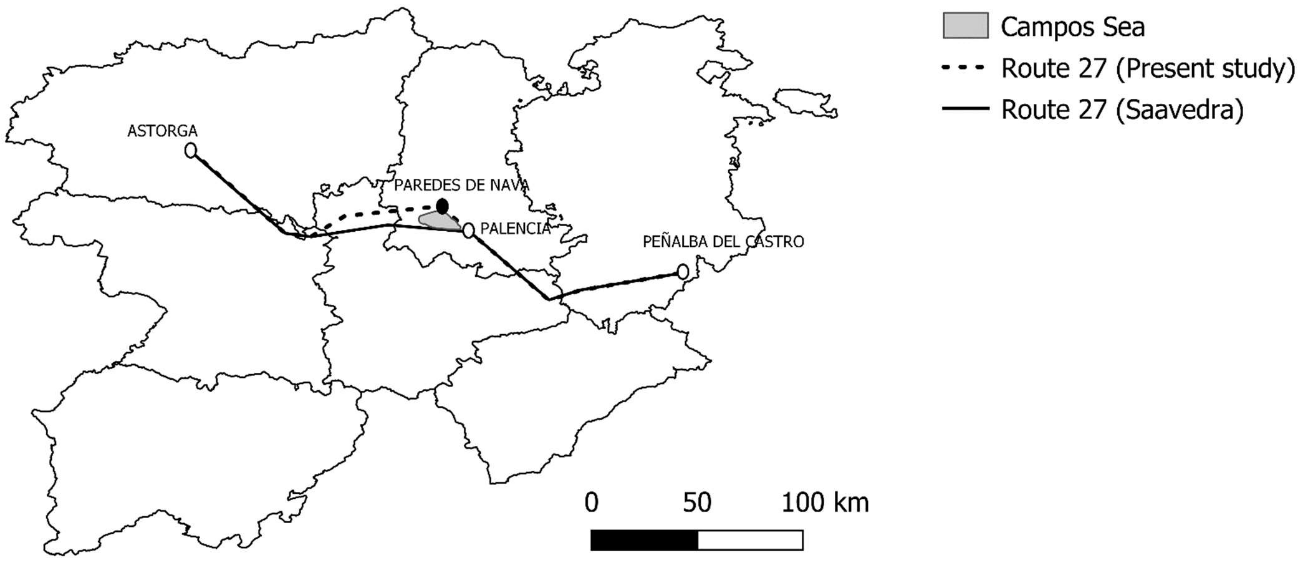

Figure 9. Suggested layouts for the Route 27 between *Asturica* and *Rauda*.