Budget Travel in the Mediterranean: A Methodology for Reconstructing Ancient Journeys through Least Cost Networks

FIONA CARROLL (ID)

EDWARD CARROLL (ID)

Author affiliations can be found in the back matter of this article

ABSTRACT

Least cost paths have been used extensively in the archaeological study of ancient routeways. In this paper the principal interest is less in tracing detailed paths than in modelling long-distance travel through an extensive network over land and water. We present a novel, computationally-efficient method for avoiding the direction-dependent, positive biases in least cost paths encountered in standard algorithms. A methodology for generating networks of such paths is introduced based on a trade-off between building and travel costs, minimizing the total cost. We use the Peutinger Table, an illustrated *itinerarium* of the Roman empire, to calibrate the parameter controlling network complexity. The problem of how to weight land versus sea travel costs in the network is tackled by comparing itineraries of Delphic *theoroi* of the third century BCE with solutions of the asymmetric travelling salesman problem, a classic graph theory puzzle.

1 INTRODUCTION

This study was prompted by an interest in how ancient journeys were planned and undertaken, with an emphasis on long distance travel taking days, weeks or even months. A number of questions arise, for example which itineraries travellers most likely followed, how long the journeys took, and whether it was more expedient to go by land or sea. In the early nineteenth century, antiquarians, guided by the ancient writings of the likes of Strabo and Pausanias, documented their own travels in Greece and Asia Minor, making notes on routeways, journey times, topographical features and ancient monuments. Among the better known are William Leake, William Gell and Edward Dodwell who put themselves in the shoes of their ancient guides, experiencing, as near as is possible, the paths and landscape described two millennia earlier.

Today it is possible to apply Geographic Information Systems (GIS) to model an optimal routeway and travel time between places through least cost path (LCP) analysis. However, simulating multi-stage journeys through an optimized network of paths over large distances is not easily addressed using GIS and for the following reasons we found it necessary to develop a bespoke methodology and software: (a) to allow the implementation of innovative optimization techniques; (b) to automate the process of handling, selecting and ordering large amounts of data; and (c) to speed up the intensive computation necessary to calculate the tens of thousands of LCPs over hundreds of thousands of kilometres, with over five million cost values.

The specific research aim was to create a tool to reconstruct the itineraries followed by groups of state-sponsored religious envoys travelling from Asia Minor to multiple cities across the Hellenistic world of the late third century BCE. To address this it became clear that separate optimizations must be applied at various stages and at different scales, from LCP modelling to large-scale network formulation and routing through a set of destinations on that network. In the spirit of Verhagen's exhortation (2018) that archaeologists should make the tools themselves the object of inquiry in order to free themselves from dependence on experts from different disciplines, we decided to develop new techniques for LCP generation and network construction, using ancient sources to calibrate network complexity to investigate optimal route-finding through such networks.

1.1 LEAST COST PATHS

Computational techniques to improve our understanding of the ancient world have gained much traction in the last two decades. In the field of archaeology in particular, LCP and network analysis have been used with the aim of elucidating different types of connectivity, often in the form of physical, civic, economic, cultural and social interaction (Bertoldi, Castiglia & Castrorao Barba 2019; Brughmans 2010; Carreras & de Soto 2013; Isaksen 2008). At the core of this is reconstructing the movement of people across the landscape. Movement can be investigated at the scale of single paths between adjacent settlements, such as journeys that could be undertaken in the course of a day (McHugh 2019; Parcero-Oubiña *et al.* 2019; Seifried & Gardner 2019). Where longer journeys are undertaken, an interconnected network of paths and settlements can be constructed, through which movement occurs in multiple stages (Herzog 2013; Verhagen *et al.* 2014; Verhagen, Nuninger & Groenhuijzen 2019).

Following the 'Principle of Least Effort' (Zipf 1949), Surface-Evans & White (2012) state that least cost analysis is predicated on the assumption that humans 'tend to economize many aspects of their behaviour'. In other words, humans will seek to take the easiest path, requiring the minimum cost. Finding the most efficient route from one place to another involves minimizing some cost, normally time or energy spent, of moving across a background topography often in the form of a digital elevation model (DEM). The cost is generally purely slope dependent and calculated using an empirical function, e.g. as described in Langmuir (1984), Tobler (1993), Minetti *et al.* (2002) and Márquez-Pérez, Vallejo-Villalta & Álvarez-Francoso (2017), though some authors have included additional sources of cost, such as marshy ground (Herzog 2014a; Verhagen, Nuninger & Groenhuijzen 2019). A comprehensive treatment would also require the inclusion of other factors, whether environmental, political or cultural. However, as Bevan (2013) argues, there are advantages to keeping models simple as they are easy to understand and test under controlled conditions.

The availability of tools to calculate LCPs via GIS means that they are easily accessible to the archaeologist or historian. However, incomplete understanding of how LCP applications work may lead the user to produce aesthetically-pleasing graphics despite erroneous inputs, in part due to them being treated as a 'black box'. A number of publications have addressed the pitfalls associated with LCP analysis (Bevan 2013; Herzog 2013; Herzog 2014a; Seifried & Gardner 2019; Verhagen, Nuninger & Groenhuijzen 2019). Bevan lists four such issues associated with the generation of least cost surfaces: failure to consider anisotropic costs (i.e. ones that vary according to direction of movement); the choice of appropriate hiking functions; the limitations of spreading algorithms to accumulate cost; and the absence of calibration and validation of results.

1.2 NETWORKS

Generation of LCPs is often limited to those between a small number of places due either to computational costs, or a confined geographical area of interest. For example,

McHugh (2019) creates single LCPs representing a day's journey within Attica, which she defines as up to 15 hours. However, to model a multi-stage journey, a network of individual routes must be considered. A useful way of approaching larger-scale connectivity is to represent its key components as an abstract network, consisting of nodes connected by edges. In the case of journeys, nodes may represent settlements, and edges the connecting paths weighted by some cost, such as the time or distance to travel between them. Such networks can be quite simple; some researchers of the ancient world have found them useful for conceptualising different types of connectivity and have taken to adopting the terminology of network theory without applying quantitative methods (e.g. Knappett 2011; Woolf 2016). A volume of Mediterranean Review (Malkin, Constantakopoulou & Panagopoulou 2007) was devoted to exploring where network analysis might be applied. However, while recognising the potential of quantitative approaches to networks, the volume was limited to qualitative aspects, network theory being novel for many users and the numeric data from antiquity perceived as insufficient.

On the other hand, a network stored as a matrix of values opens the door to various types of mathematical manipulation, e.g. the generation of metrics to characterize connectivity or complexity, the clustering of nodes or optimal routing through nodes. Whereas there is a clear optimization methodology for individual LCPs, the cost, or efficiency, metric for network optimization is less well defined. The construction of different least cost networks is addressed in Herzog (2013). Among these are networks of 'least cost to the user' (Waugh 2000) where every node has a connection to every other node and cost to the traveller is minimized. At the opposite end of the spectrum are networks termed 'least cost to the builder' that connect all nodes whilst minimizing the cost of construction. Techniques such as triangulation or proximal point analysis produce results which seem heuristically to be a compromise between the extremes of least cost to the user and least cost to the builder (Herzog 2014b; Groenhuijzen & Verhagen 2017), but fail to balance building and using costs formally. The justification for the term 'least cost network' seems to be that the edges are LCPs rather than the network itself being constructed through any explicit cost minimization principle, something which will be addressed here.

1.3 APPROACH

Python is used to code the elements which extract and organize the data and to call a Fortran routine which performs the computationally-intensive, heavy lifting of LCP generation, since the latter language executes much faster. The use of Python also allows the Fortran to run in parallel on multiple processors, accelerating execution significantly and extending the overarching principle of cost minimization to our computational approach.

The methodology implemented can be divided into four main steps: (a) Minimize the cost of travel between known ancient settlements avoiding the direction-dependent biases in LCPs encountered in standard algorithms; (b) generate a *minimum total cost* network using the LCPs based on a trade-off between building and travel (user) costs and a derived network complexity parameter, λ; (c) calibrate λ, and hence the network travel costs, using the ancient itinerary *Tabula Peutingeriana*, or Peutinger Table; (d) calibrate the relative cost of land and sea travel by comparing ancient itineraries of the *Delphic Theorodokoi List* with their solutions to the travelling salesman problem. The resulting tool is used to construct an extensive network spanning much of the Greek-speaking world through which optimal journeys can be calculated over land and sea.

2 LEAST COST PATH METHODOLOGY

As in GIS software, Dijkstra's algorithm (Dijkstra 1959) for finding the most efficient route through a network lies at the heart of LCP generation. Each pixel, or cell, of a digital elevation model (DEM) is treated as a grid point which effectively becomes a node of a network, the value of the line, or edge, linking adjacent nodes being the cost of progressing from one pixel to the next. For the DEM, SRTM (Shuttle Radar Topography Mission) elevations derived from a satellite-mounted radar (Yang, Meng & Zhang 2011) are used. We have masked these with ancient coastline and lake files based on the Barrington Atlas (Talbert and Bagnall 2000) downloaded from the Ancient World Mapping Center (*http://awmc.unc.edu/wordpress/*), so that over ancient water surfaces (other than rivers), elevations are forced to a null value over which land travel is not allowed. Conversely, for LCPs representing boat trips, travel is not allowed on non-null values (i.e. land). Quite significant differences with the modern coastline are seen in places, e.g. parts of Akarnania, Aitolia and Caria. The modern salt lakes of Akrotiri and Larnaca in Cyprus were anciently connected to the sea, Akrotiri formerly being an offshore island seen to the south of the mainland in ***Figure 10***. Each DEM datum is taken to be representative of a grid point, being the mean elevation of an area one arc second by one arc second, effectively 30 m in the north-south direction and, at the latitude of Greece, about 24 m in the east-west direction.

Efficiency of a path is defined in terms of minimization of the cost function with which the accumulated cost of walking along different routes is calculated. Most commonly, time elapsed is the measure of cost, the program giving the options of using the Tobler hiking function (Tobler 1993), a modified Tobler function proposed by Márquez-Pérez, Vallejo-Villalta & Álvarez-Francoso (2017) or Naismith's rule, as extended by

Langmuir (1984) to cope with downhill walking. As an alternative to time, the user may choose to find the route that minimizes energy, for which a variant of Minetti's formula for energy expenditure whilst walking is used to calculate network edges (Minetti *et al.* 2002). For each of these functions the edge values between points depend only on the distance and slope and are anisotropic, in other words they differ according to the slope in the direction of travel rather than simply depending on absolute slope. The functions all have different characteristics e.g. in the critical slope at which it becomes more efficient to adopt a zig-zagging hairpin, or switchback, path (though it is a disadvantage of the Naismith/Langmuir rule that this transition never occurs). No account is taken in this implementation of the costs of negotiating different land types, e.g. the difficulties associated with marshy ground, or of crossing rivers.

The standard approach to creating an LCP is employed initially, briefly described as follows: Working outwards from the defined start point, multiple costs are calculated for each point on the grid, representing accumulated values for different routes. Any newly-calculated value replaces a pre-existing one if it is lower. Dijkstra's algorithm makes this process efficient by greatly reducing the number of different routes which must be explored. When a lower cost to reach a given point is discovered, the direction taken in the final step to reach it is recorded in a separate, *backlink* matrix. Using this, a complete LCP can be quickly reconstituted working backwards from any point to the start once the full cost field has been determined.

A key consideration is how to define the edges connecting neighbouring points. The most common technique involves the so-called queen's move pattern, in which each of the immediately adjacent eight points on the grid is connected. This restricts changes in direction in any LCP to 45 degrees with the result that the cost of reaching a destination lying between the eight compass points is artificially increased. A map of isochrones of time elapsed walking on a flat surface shows concentric octagons rather than the circles which would result from an analytical solution (e.g. Bevan 2013: 8). One way of reducing this error is to allow edges to connect more points in the vicinity, e.g. by adding a knight's move pattern, introducing eight extra connections – this is an option in the QGIS package. The 16-sided isochrones are closer to circles, but still result in unrealistic cost penalties for destinations lying between the 16-point compass directions. In addition, it effectively doubles the processing time and allows the leap-frogging of potential barriers by jumping over adjacent pixels, therefore not recording the full cost. An improvement to this technique employed by Herzog (2013) extends the neighbourhood region out to cover 48 points by including two extra move types, allocating a cost to each by assuming a straight line route over intervening points and preventing leap-

frogging by interpolating their cost values onto the line. However, Herzog shows that even with this 48 neighbour method, 64 points in a 13 × 13 grid (38% of points) cannot be reached by an optimal, straight-line path, with deviation from such a path of up to 8%.

2.1 ROTATE AND OVERSAMPLE TECHNIQUE

The calibration of network complexity covered later in this paper requires that LCPs be made as optimal and with as little bias as possible. Therefore a method has been derived of largely eliminating such artificial, direction-dependent cost penalties at little extra computational cost. Termed the *rotate and oversample* technique, it is depicted schematically in *Figure 1* and consists of the following steps, which are novel from step 2 onwards.

(1) Derive the LCP using the queen's move.
(2) Break this path into sections – most experiments used a section length of 1 kilometre (A to B in *Figure 1*). Rotate each of these in turn so that the end is due east of the start.
(3) Create a separate high-resolution grid which encompasses each rotated section. This has a grid spacing *n* times smaller than the DEM. Most experiments conducted used *n* = 6, giving a sub grid spacing of about 5 × 4 metres depending on the angle of rotation.
(4) Populate each high-resolution grid with linearly-interpolated elevation values mapped from the main DEM grid.
(5) Re-apply Dijkstra's algorithm to derive a cost field, thence a high-resolution LCP from the start to end point on the high-resolution sub grid. The original LCP route section is discarded, having served its purpose in giving start and end points as well as defining the grid dimensions.
(6) Construct a new path by taking every *n*th point of the high-resolution LCP, rotated back to the coordinate space of the original grid. If the terrain slope exceeds the critical value at which time or energy-saving zig-zagging is expected, retain all points to avoid smoothing over such detail. This critical slope is calculated for all cost functions other than Naismith's, for which it does not exist.
(7) Following this new path, recompute the cost from point to point, for the full journey. If the cost function minimized with Dijkstra is energy expended, there is the opportunity here of computing the time taken to follow this path using one of the time-based cost functions.

The rotation step solves the standard queen's move distance bias for travel on flat surfaces, leading to straight paths and isochrones in concentric circles rather than polygons. *Figure 2* compares both methodologies for paths to close destinations (less than a kilometre
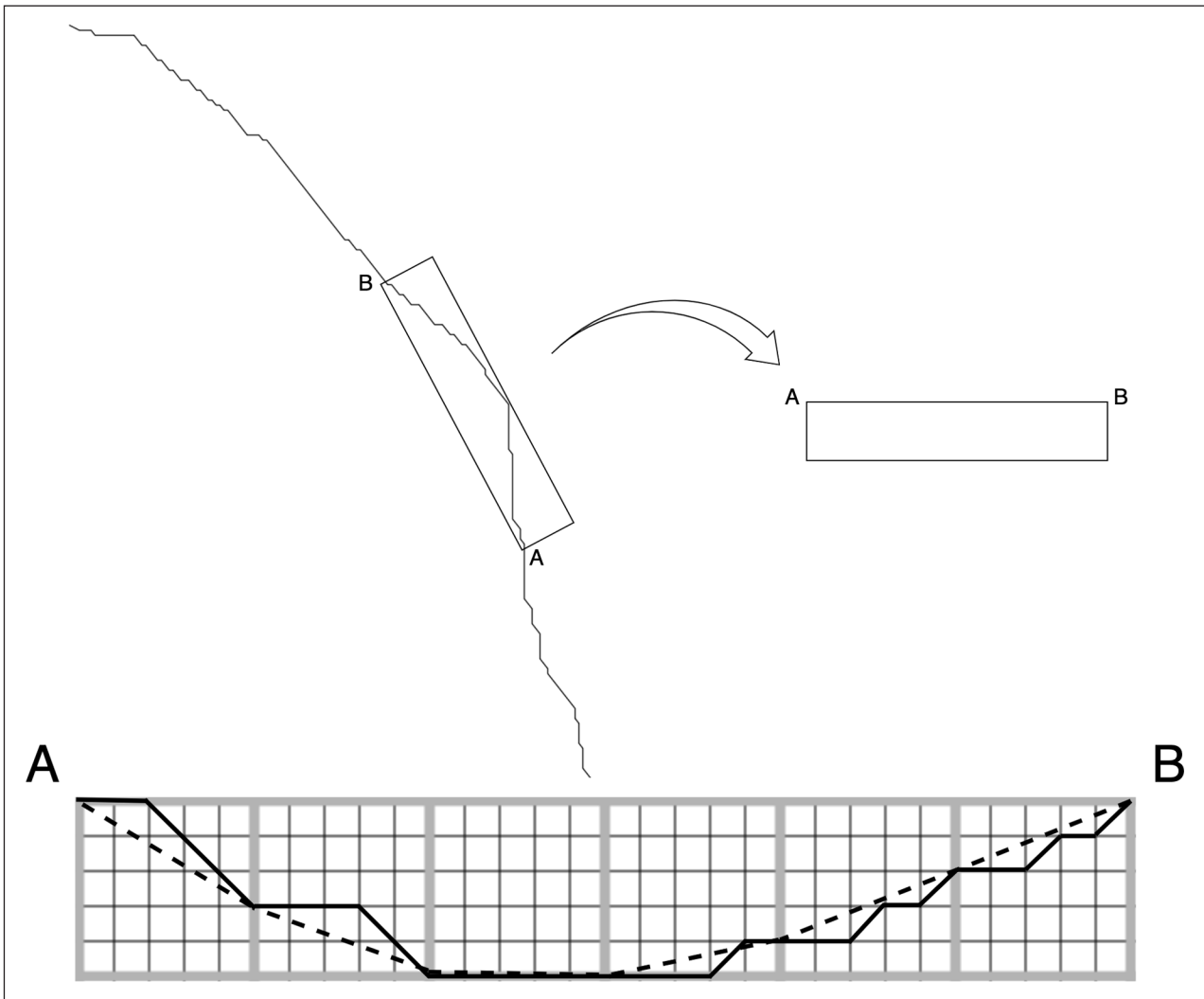
Figure 1 Schematic illustration of *rotate and oversample* method. Here n = 5, the solid line resulting from step 5, dashed from step 6.
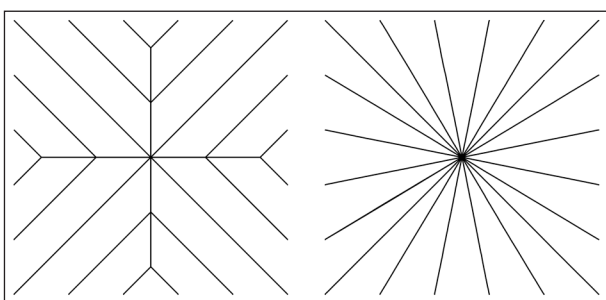

Figure 2 Least costs paths on a flat surface radiating from a central point. Standard queen's move (left), rotate methodology (right).

distant) arranged at equal intervals around the edge of a square from a central point on a flat surface.

However, the deviations from straight-line paths still lead to errors, which are significantly lessened by the oversample step. The end result is a path which is constrained to follow the general route of the originally-derived LCP, but in which directional changes of about 45/n degrees (e.g. 7.5 degrees with n = 6) are allowed, giving shorter paths and, more importantly, largely

eliminating distance biases depending on the direction of destination relative to start. Unlike Herzog's 48 or higher neighbour method, all these slight directional changes can be accomplished over sections as short as one DEM grid length, e.g. in a winding valley path (*Figure 3*). With n = 6, all points within a 13 × 13 square are reachable from the centre by straight-line paths with no deviation from a straight-line path, compared to 62% of points with the 48 neighbour method with up to 8% deviation (Herzog 2013). The final step of selecting only every nth point effectively brings the resolution of the path back into equivalence with that of the DEM though allows points to be defined intermediate to the main grid positions. Because the sub-grids on which the second Dijkstra step is made are very small compared to the main DEM, it only results in a modest increase in computational time (generally less than 10% relative to standard queen's move), whereas the 48 neighbour method would approximately quadruple the processing time for deriving the cost surface relative to queen's move. In a comparison of over 10,000 km of LCPs in the Peloponnese, the technique generated paths which were

on average 3.5% shorter than those produced by the standard queen's move. More importantly, there is no bias in the computed distance associated with average path direction.

There is a special problem associated with paths in domains in which there is perfectly uniform (or no) terrain slope. In this situation, with no variation in forcing from topography, there are many equally long queen's move solutions to the shortest path and, as Herzog (2013) shows, the calculated path can deviate up to 20% of its length from the optimal path. However, with real data there is always some slight unevenness in terrain, noisiness in the SRTM data artificially adding to this. For the purposes of travel over the sea the expedient of adding a small (less than one percent) random component to speed gets round this problem.

It should be noted that an LCP between two settlements varies according to which is nominated as the start point due to the anisotropic treatment of cost on slopes, with the journey in one direction occasionally differing substantially in its route from that in the opposite direction. In order to generate a plausible road network, we follow Herzog (2014b) in using a cost function which represents walking in both directions at each stage, so that the path generated is that which minimizes the cost of a return journey on the same route. Whilst the path is isotropic inasmuch as one line on a map represents the path taken in both directions, the cost of walking it in each direction still varies, and is calculated and stored separately so that the network will retain its anisotropic character. The path is stored as a series of latitude and longitudes, costs are stored in an $N \times N$ matrix where N is the number of locations, and COST(i,j) represents the cost of going from place i to place j. Anisotropy is reflected in the fact that COST(i, j) ≠ COST(j, i) unless the surface is perfectly flat, which is the case with travel over water.

It is possible to reduce computational time considerably by applying what will be called an *undersampling* method to the first step. This involves altering the first queen's move Dijkstra call so that rather than searching immediately-neighbouring points on the DEM, it establishes direct links to points at some number of grid lengths away, known as the undersampling factor. The initial LCP is then necessarily more approximate with

likely underestimated temporal costs, since gradients will generally be artificially reduced. However, having divided it into rotated sections at the oversample stage, we go back to the usual neighbourhood point definition for the fine-scale LCP, so gradients are ultimately fully resolved. The full process can then be termed *undersample, rotate and oversample*. This expedient reduces processing time by around the square of the undersampling factor, adding further to the time savings achieved by parallel processing. As an example, it allows the network of the Peloponnese (*Figure 9*) containing over 10,000 km of LCPs, to be created in 41 seconds on the eight processing threads of a 2013 MacBook Pro. This shortcut is used as standard when generating boat journeys where gradient calculation does not apply and it is simply resolution of the position of the coastline which is at issue. An undersampling factor of four is used, a precision of 120 metres being considered sufficient. It was also used as standard when creating a master, least cost to the builder network to supply first-guess costs prior to full-resolution network generation (see Appendix 1).

3 MINIMUM TOTAL COST NETWORK METHODOLOGY

One may question the degree to which real routes achieved the degree of optimization represented in LCPs. Even today, someone travelling by road from one town to another rarely takes an optimal route, instead being constrained to move through a network of roads which generally forces an approach to the destination which is indirect to some degree. The complexity of the network is a strong determinant of how direct longer routes or itineraries are, so to recreate the degree of connectedness of settlements across a wide region in the ancient world, a suitable network must be created. Networks lying between the two extremes of 'least cost to the builder' and 'least cost to the user' (Waugh 2000) can be produced by various techniques (e.g. Evans & Rivers 2017; Groenhuijzen & Verhagen 2017; Manière, Crépy & Redon 2021; Prignano *et al*. 2019). Some have LCP-derived edges, though others only allow geodesic (effectively straight-line) paths and none appears
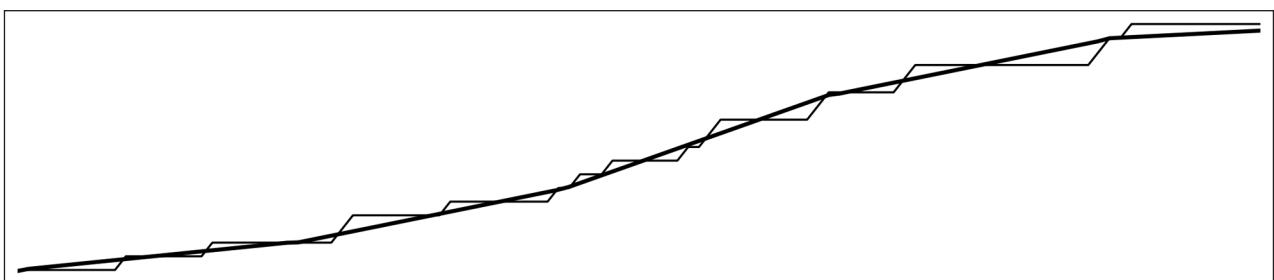

Figure 3 Rotate and oversample LCP (thick) compared with standard queen's move LCP (thin).

to offer any minimization of combined cost of travel and cost of road building. Prignano *et al.* (2019) have some representation of building cost in the form of a fixed, upper limit to the total length of network edges permitted.

Here we introduce a methodology that addresses these notional costs, which we call *minimum total cost*. It generalizes the concepts of competing costs into a system in which least cost to the builder and least cost to the user fall out as the extremes of one technique, and allows all gradations in between.

Consider three sites, A, B and C as in the schematic at *Figure 4*. The lengths of the lines are proportional to the cost (temporal or energetic) of travelling an LCP road, and their straightness is not to be interpreted as representing straight routes. We use the notation AB, BC and AC to represent the costs of travelling individual legs and ABC the cost of travelling from A to C via B. The first principle is that each site should have a road connecting it to its nearest (in cost terms) neighbour, so roads connect A with B and B with C. The next principle is that a road be built connecting A with C only if there is a net positive benefit, i.e. that cost savings to travellers between A and C are greater than the cost of building and maintaining a road from A to C. Let *b* represent building and maintenance costs per year, and *t* represent travelling costs per year, which are proportional to the length of the journey and number of people and/or transported goods making the journey. The total benefit (negative cost) of instituting a road can be given as

$$benefit = (ABC - AC) \times t - AC \times b$$
or
$$benefit = t(ABC - AC(1 + b/t))$$

For the creation of a road to be cost-effective the total benefit should be positive, i.e.

$$AC(1 + b/t) < ABC$$

We will use the term 'build cost' to include the initial construction cost as well as the ongoing maintenance and repair cost. Whilst construction might be thought to involve a one-off, capital cost, we assume that this is spread over an extended period of time, so that like travel costs, build costs are deemed to be continuous. The build cost of a road might at first be thought of as proportional to the geometric length of the road rather than time taken or energy expended to walk it. However, it seems logical that, like travel, road building is more challenging on difficult terrain, so for our purposes build cost per unit length of road has been treated as being proportional to time or energetic cost of travelling that length. The
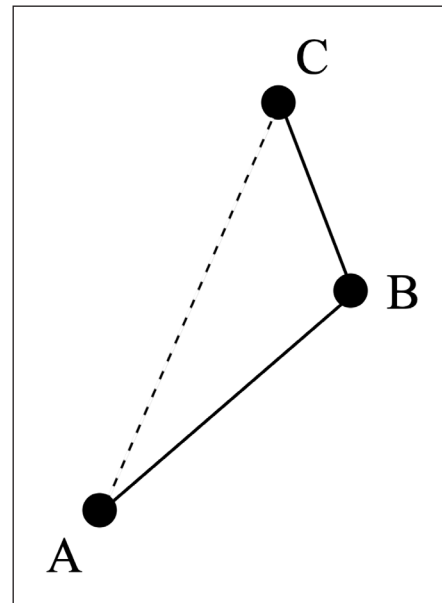

Figure 4 Schematic of travel costs between three points.
Lines AB and BC represent the costs of travelling established roads. Dotted line AC represents a proposed direct road from A to C.

term *distance* will be loosely used when discussing road cost, being a catch-all for spatial, temporal or energetic distance.

The formula could be made more elaborate by factoring in the traffic volumes on the road, which could be made proportional to the population of the sites A and C. However, in the absence of population data and to keep things as simple as possible, we reason that not only cost of travel, but also of road building is likely to be in some degree proportional to traffic levels, so that this factor would tend to cancel out in *b/t*. For a route which only has two travellers per day a simple path might do, whereas a major highway with wheeled traffic would require a minimum width, foundations, paving and periodic repairs.

Denoting the ratio *t/b* by the dimensionless parameter *λ*, the criterion for establishing a direct road from A to C rests on whether AC times $1 + 1/\lambda$ is less than ABC, and the break-even cost of building the road can be given by *t/λ*. Take the example of a situation where *λ* = 1.0, corresponding to travel cost equalling build cost. New roads will be established where the indirect route takes more than twice as long as the proposed direct route. If road building is expensive and *λ* = 0.5, the indirect route would have to be three times as long as the direct one before a road is built; but if it is cheaper, or travel time more valuable, e.g. with *λ* = 5.0, the indirect route only has to be 1.2 times as long before a new road is created. Following the rule that a road only be instituted if savings in travel exceed the break-even build cost, the total cost is minimized.

3.1 NETWORK COST MINIMIZATION IMPLEMENTATION

In order to make computation manageable direct links are limited to those which can be made within some defined cost termed the maximum temporal radius, usually set to 36 hours. This is tunable, and at first 12 hours was specified on the basis that this was the approximate limit of a day's travel, and that travellers would not want to overnight between settlements. However, where settlements are sparsely distributed it left holes in some networks, and it was reasoned that there would likely be stations or small settlements of which we have no record. Two methods for implementing the cost minimization technique were coded, *local outward search* (LOS) and *global pairwise optimization* (GPO) – for details of algorithmic implementation see Appendix 2. Both require a master network to be created first, effectively a least cost to user network in which every node is connected to every other node within the maximum temporal radius. Since this simply provides a lookup table to inform decisions about which nodes to link in the network, its precision is not critical and the computational burden of generating the master network is significantly lessened by employing the undersample, rotate and oversample methodology.

3.1.1 Local Outward Search methodology

The LOS method requires that $\lambda \geq 1$ and replicates a decision-making process which takes place at the level of individual settlements. Nodes are placed in order of temporal distance by direct LCP from the central settlement under consideration. A route is established to the closest node. Thereafter, each node is examined in turn to weigh the net benefit (dependent on λ) of building a direct road against using a stepping stone to which roads, direct or indirect, have already been established in this outward search. The route which yields the greatest benefit is chosen. We could imagine point A in *Figure 4* as the central settlement, B as an intermediate node which has already been connected, directly or indirectly, and C as a settlement to which a direct road is built according to whether $AC(1 + 1/\lambda)$ < ABC. There is a stricture that AB must be less than AC, in other words that the staging post be closer to the destination than the origin. If a direct route is not established, the temporal distance AC is updated with the indirect value ABC. A new central location is then chosen and the process repeated until all nodes have been treated as a central location.

3.1.2 Global Pairwise Optimization methodology

The GPO method replicates a central authority in decision making. It involves looking across the whole network and finding the road, the establishment of which would confer most benefit per unit distance to the two nodes at either end, using benefit = t (ABC – AC(1 + 1/λ)). Every edge in the network is given an arbitrarily large starting value, so the smallest roads are created first. Every time a road is established, Dijkstra's routine is called to update the indirect temporal distances till there are at least indirect routes connecting all nodes. The process continues until the maximum benefit of new road establishment falls to zero or less.

3.2 NETWORK EXAMPLES FOR CYPRUS

By specifying varying values of λ, LCP networks of different complexity can be constructed across a region. *Figure 5* shows examples of LCP networks across Cyprus utilizing nodes for the late third century BCE for values of λ ranging from 0.1 to ∞. Here λ = 0.1 is effectively a least-cost to the builder network and λ = ∞ is least cost to the user. λ = 0.1 produces elements of Herzog's (2013) 'main routeway with later subsidiary pathway' model, with some settlements lying at the end of spurs leading off a main route. The Troödos mountains in the interior of the western half of Cyprus force LCPs into valleys, with the result that even with large λ, routeways remain well defined and separate. However, in the flatter, eastern half of the island the roads spread out across the plain in an unrealistic sprawl. Quite apart from considerations of travel and build costs, the cost in loss of agricultural land is likely to exert a strong constraint against such proliferation.

A weakness of the technique is the failure to re-use sections of existing roads in establishing new routes by defining junctions, as is done e.g. in the Steiner tree approach (Verhagen, Polla & Frommer 2014). However, these networks will certainly exhibit junction-like points, e.g. where routes are forced together in complex terrain, as can be seen in the λ = 8 network in *Figure 5* around the Troödos mountains. An extension to the technique which may be worth pursuing is the assignment of junction nodes based on the master network and their use in the network optimization process. Alternatively, new nodes could be created e.g. every kilometre of LCP, though this would result in prohibitive costs for anything other than small networks.

4 NETWORK CALIBRATION

Ideally, known ancient roads and networks of roads would be put to this purpose, but their fragmentary nature makes this unviable. Instead we can turn to the ancient *itineraria*, which were used to aid travellers on their journeys, since distances between locations depend on network complexity, with simpler networks of lower λ making for longer distances between places on average.
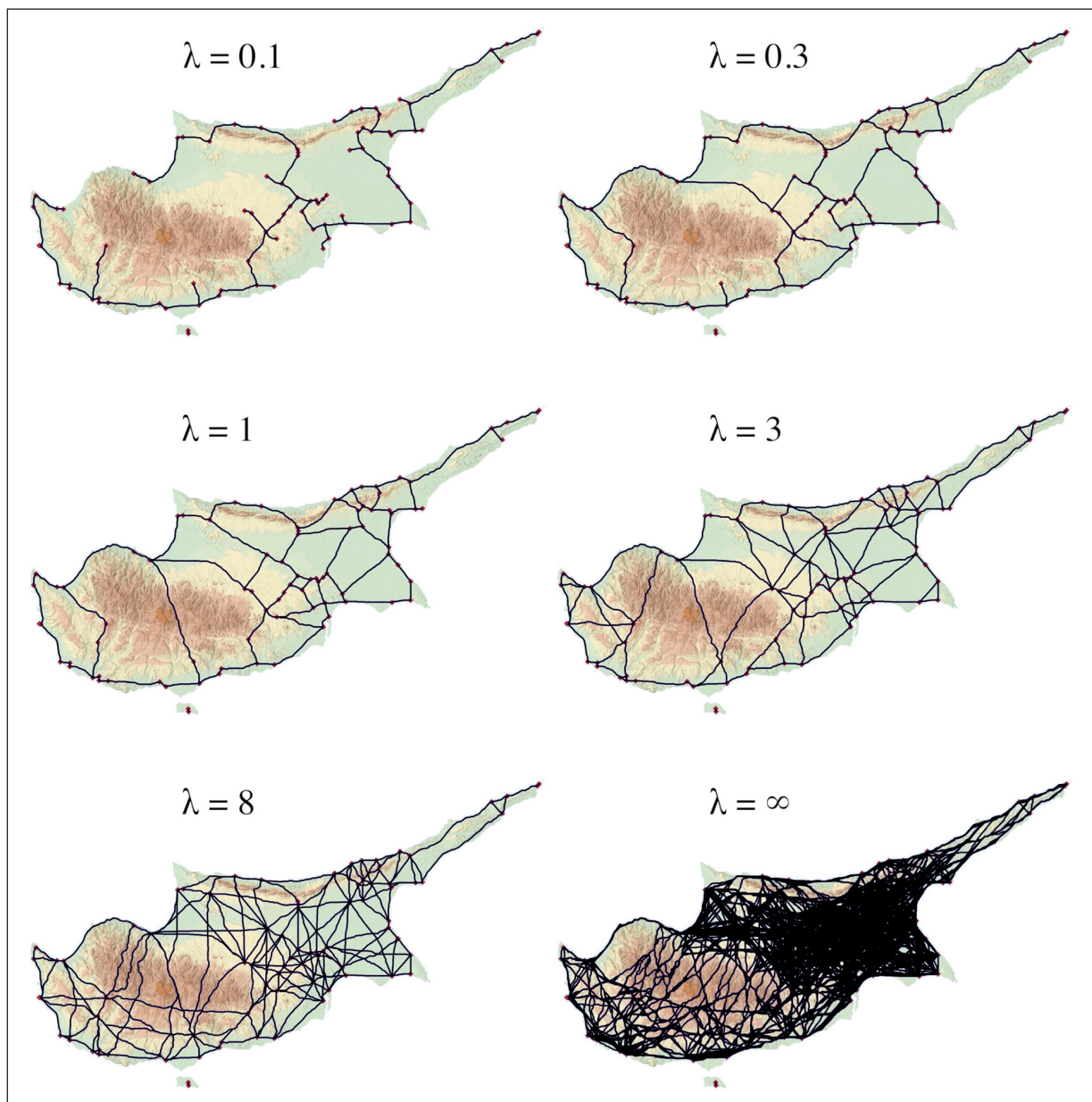
Figure 5 **LCP networks of increasing complexity generated for late third century BCE Cyprus.** GPO method used for λ < 1, others using LOS.

4.1 TABULA PEUTINGERIANA (TP)

Itineraria in their simplest form were lists of place names with the estimated distances between them. No such lists over land have survived from before the Roman period, although sea itineraries (*periploi*) exist from the Hellenistic and classical periods. The most famous and extensive ancient itinerary is the *Tabula Peutingeriana* (TP), or Peutinger Table, which is in the form of a semi-pictorial map, and is thought to have been compiled from multiple smaller itineraries designed for practical use (Salway 2001: 47). A section of the TP will be used to calibrate simulated network complexity, and in this context an understanding of its chronology is necessary by way of introduction. The TP is a medieval copy of an earlier map of the Roman imperial period, which itself went through a number of iterations, but excepting a few cases (e.g. Constantinople rather than Byzantium and the

marking of St Peter's outside Rome), the place-names are generally those current before the era of Diocletian and the Tetrarchy (293–305 CE), and some elements appear to belong to a considerably earlier period (Salway 2005: 120). Its inclusion of Pompeii, Herculaneum and Oplontis, all destroyed in 79 CE, suggests its descent from an earlier prototype (Elsner 2000: 185), while Rathmann (2016: 342) argues for the TP archetype pre-dating the Roman era and belonging to the Hellenistic period. Because it cannot be viewed as a snapshot of the ancient world, a decision has to be made as to which date to use in the selection of known ancient settlements to provide network nodes. Given previous considerations, 280 CE was chosen.

We took 117 route distances in Roman miles from the TP for a region extending from Sicily to Cyprus, this region chosen to correspond with the simulated

network for application to itineraries of the earlier Greek world. Network nodes were extracted from the Pleiades gazetteer of ancient places (Barker *et al.* 2016), filtered by area and date. Note that we do not use the network itself depicted in the TP, which is assumed to represent a small portion of the total road network, rather simply the given distances.

4.2 OPTIMIZATION OF λ BY STATISTICAL ANALYSIS OF NETWORK DISTANCES AGAINST TP

Different networks of least cost paths (LCPs) were derived across the region based on the *total cost minimization* principle applied using both the LOS and GPO methods. Varying values of λ were specified using costs based on Minetti's energy function (Minetti *et al.* 2002) and the modified Tobler hiking function (Márquez-Pérez, Vallejo-Villalta & Álvarez-Francoso 2017). Dijkstra's algorithm (Dijkstra 1959) was applied to the networks in order to find an optimal path from node to node through the network and derive the 117 route distances corresponding to the legs taken from the TP. Distances obtained in this way are denoted by LCPnet. The best fit between TP and LCPnet values was produced with the modified Tobler hiking function in a network of $\lambda = 7$ generated by the LOS method. This both maximized the number of distances for which LCPnet was within 5% of TP (33) and generated the median normalized difference closest to zero, whereas the best fit obtained using the GPO method had 29 instances of LCPnet within 5% of TP.

Ultimately, consideration of the quality of the TP data should be made in light of the fact that the comparisons here are not made against truth, which cannot be known, but against the LCP network, which doubtless has its own bias and noise. However, given the wide range of ratios and evident errors in TP values, with e.g. around 25% of the sample being shorter than geodesic (effectively straight-line) distances, it is considered that the large majority of the variance comes from them rather than the LCPnet values. The large spread in ratios and the presence of gross errors in the TP may seem to

militate against extracting value from them; however, it is a fundamental tenet of statistics that poor quality of data can be compensated by quantity, and we contend that judicious use of diagnostics can yield a useful signal sufficient for our purpose. At the same time something about the nature of the errors in the TP can be inferred.

Figure 6 shows the frequency plots of normalized difference between LCPnet and TP for three different levels of network complexity defined by $\lambda = 4$, 7 and 10 at 0.1 bin intervals. $\lambda = 7$ scored highest independently both for symmetry and number of differences of route length of less than 5% (indicated by the bin interval –0.05 to +0.05).

Errors in the TP values could have many different causes, such as transcription mistakes (Bekker-Nielsen 2004), inaccuracies in the sources from which the original table was constructed, original surveying errors, or modern misidentification of the positions of the ancient sites referred to. On the other hand, errors in LCPnet values could arise due to inadequacies in the assumptions underlying the technique, such as the very principle of cost minimization and the applicability of the hiking formula employed. According to the Central Limit Theorem, one would expect random errors in a population arising for many independent reasons to approach a statistically normal, or Gaussian, distribution. In all the frequency plots in *Figure 6*, but especially so for $\lambda = 7$, a narrow peak is in evidence, superimposed on an approximately Gaussian distribution, suggesting the amalgamation of two distinct populations.

To test this, an iterative process was adopted in which incremental changes were made independently to N (number), σ (standard deviation) and μ (mean) of two idealized subpopulations, and those which increase R^2 are retained, causing the composite modelled population to converge towards the actual population until no significant improvements are made. The distribution was found to be fitted best to a composite distribution consisting of the sum of two unbiased ($\mu = 0.0$) Gaussian populations, one larger ($N = 78.2$) and noisier ($\sigma = 0.37$), the other smaller ($N = 34.1$) and less noisy ($\sigma = 0.04$), the
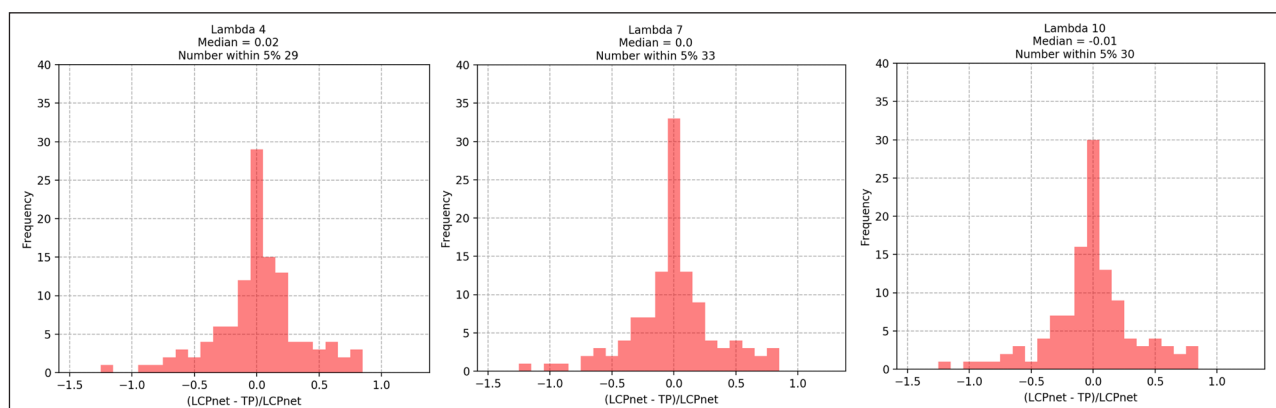

Figure 6 Frequency plots of normalized difference between LCPnet and TP. For $\lambda = 4$, 7 and 10 at 0.1 bin intervals.

results shown in *Figure 7*. The high correlation between this composite and the observed distribution (R^2 = 0.981) gives confidence that the modelled distribution approximates the observed one well, indicating that 98.1% of the observed frequency within bins can be explained by the idealized distributions. Two of the values not accounted for in the Gaussian populations lie outside of the plotted range shown, with errors up to a factor of 5.

4.3 COMPARISON OF INFERRED NETWORK COMPLEXITY WITH ARCHAEOLOGICAL AND HISTORICAL EVIDENCE

The network generated by λ = 7 (shown for the Peloponnese in *Figure 8*) may seem intricate compared with many reconstructed historical road networks. One measure of complexity is the β index, the ratio of the number of edges to nodes (Kansky 1963). In our simulated network for Cyprus, β = 2.60 compared with a value of 1.68 calculated by Bekker-Nielsen (2004: 225). He based his network on fieldwork throughout Cyprus
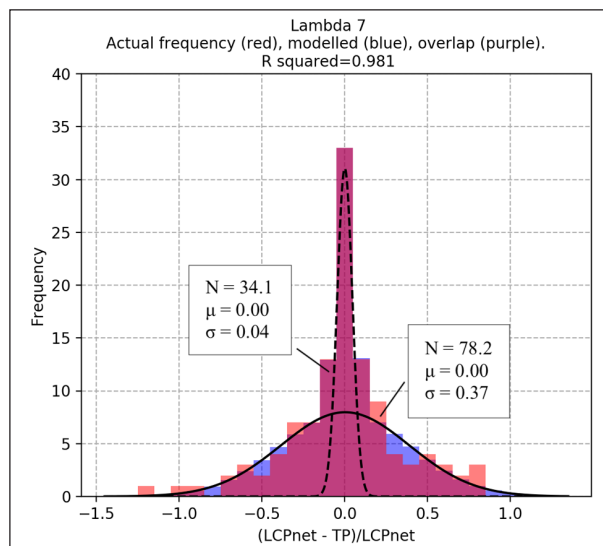

Figure 7 Frequency plots of actual and modelled normalized difference. Red+purple covers actual frequency distribution (as in *Figure 6*), whilst blue+purple give that modelled by the sum of the two Gaussian functions shown, defined by population sizes (*N*), standard deviation (σ) and mean (μ).
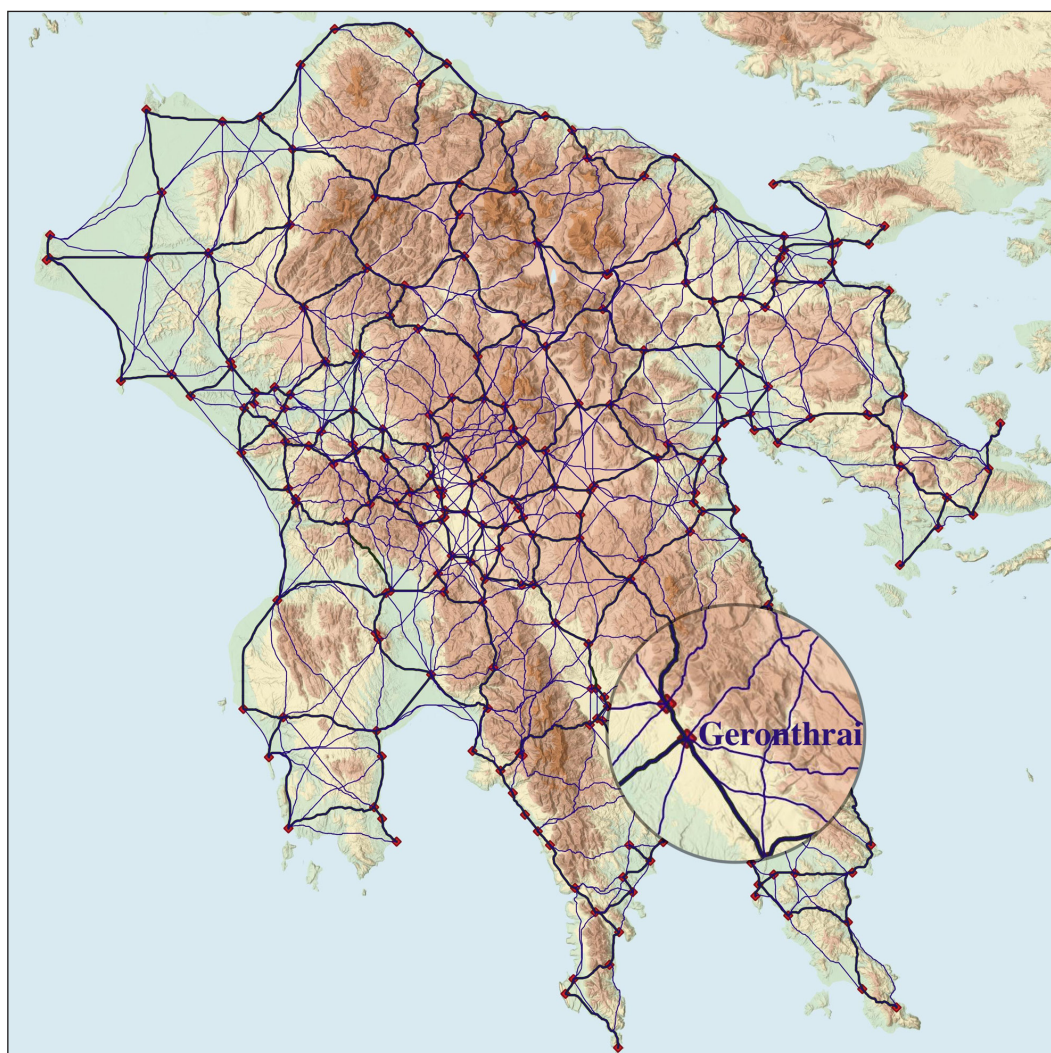

Figure 8 Modelled networks in late third century BCE Peloponnese. Major (thick black) and minor (thin blue) roads from a putative network created with λ = 0.65 (β = 1.31) and λ = 7.0 (β = 3.0). All roads in the simpler network are also in the more complex network. The area around Geronthrai is magnified.

and use of literary sources, though he acknowledges that in his study many local roads remain unidentified, and that even main roads may be missing. He also quotes (p. 225) β = 1.31 to 1.35 for reconstructed Roman road networks in Gaul, northern Italy and Greece, and 'more than two' for the area around Carthage. His value for the Peloponnese of β = 1.31 is close to the 1.33 given by Sanders and Whitbread (1990) which they calculated from a map of 1822. Justification for the use of early nineteenth century evidence to make inferences about travel in the ancient world is similar to that advanced by authors who quote accounts of travellers such as Leake, Dodwell and Gell (e.g. Seifried & Gardner 2019; McHugh 2019). Both are significantly lower than the β = 3.00 for our λ = 7 network for the Peloponnese.

However, Pikoulas (2012) in his exhaustive study of the ancient cart-road network of Lakonike in the Peloponnese found strong archaeological evidence for a much more complex network than had previously been imagined. He documented, for instance, eight roads leading to Geronthrai, one of the *poleis* of Lakonike, which compares well with the seven roads to Geronthrai from surrounding nodes generated by setting λ = 7. Bekker-Nielsen's value of β = 1.31 for the Peloponnese can be replicated by using λ = 0.65, the resulting network having only three roads into Geronthrai (*Figure 8*), so Pikoulas's archaeological findings better support the order of complexity exhibited by our network. If we take it to include minor routes which, in the hierarchy of roads, lie below the main *viae publicae* or equivalent, it is possible to imagine a combined network of routes of such complexity. These could include the likes of the *viae vicinales*, the local, village roads of the Roman system, which may have been no more than unpaved tracks or well-trodden footpaths, leaving little or no trace today.

4.4 DISCUSSION

Whilst the TP depicted a system of roads on the grand scale, it is easy to imagine that some of the individual edges grew out of minor routes serving local networks, but which came to achieve a greater significance when repurposed and upgraded as arteries in an empire-wide network. The better fit provided by the LOS network generation method over GPO gives some support to this idea, GPO having 29 results within 5% of parity between TP and LCPnet, versus 33 for LOS, marginally suggesting that the network evolved principally from local rather than large-scale organization. This contrasts with the findings of Prignano *et al.* (2019) in relation to road networks of Iron Age Etruria, though the difference could be explained by the fact that here we are considering a much larger area.

One might speculate that the Roman roads on which the TP distances are based could have been on average shorter than LCPs since Romans roads are known for following more direct courses than their predecessors,

their surveyors realizing impressively rectilinear constructions, sometimes for tens of kilometres at a time (Hucker 2009), achieving efficiencies in building materials and garnering prestige. We took all LCPs of 10 Roman miles or less to minimize the risk of indirect routes through intermediate nodes and of copyist omission errors, and they matched their TP equivalents to less than half a percent. Too much should not be read into this because of the small sample size (5), though it certainly gives no evidence of systematic divergence between direct LCP distances and TP values. It also lends some support to the idea that in tuning the LCP network to the TP values, we are mainly matching network complexity rather than correcting biases in LCPs. However, for our purposes the distinction between the two sources of error is not of great importance since our main interest is in generating realistic edge costs between nodes regardless of how that is achieved. In the same way, the actual paths taken between nodes are of secondary interest, in contrast to many archaeological studies.

Salway (2001: 58) maintains that the variegated nature of the TP reflects the fact that it was a compilation of data from publicly-displayed lists of stages and their distances from different locations and times. We speculate that the smaller, more accurate subset of TP figures which correlate well with LCPnet (*Figure 7*) may have come disproportionately from roads which were built and measured directly by the Romans, whose practice it was to survey carefully and mark out prior to building (Hucker 2009). The larger, less accurate subset may have contained essentially unaltered roads whose pre-existing, recorded measurements were used by the Romans, which seems to be the case with the distances given on the monumental *stadiasmus provinciae Lyciae* at Patara (Salway 2001: 58). Here 64 route distances between some 50 locations were recorded by the Romans shortly after they had conquered Lycia in Asia Minor (Rousset 2013: 65). A case-by-case examination would be necessary to cast further light on this, but at the very least, Salway's contention would support the idea of distinct subpopulations contained in the TP as strongly suggested by the statistical analysis.

Verhagen, Polla & Frommer (2014) point out that the cost functions commonly used to derive LCPs are not necessarily applicable to a road designed for wheeled transport, particularly in respect of the differences in coping with slopes, and estimate that a mule-drawn cart with a load of 500 kg would be unable to get in motion on a slope of greater than 9%, based on an equation by Raepsaet (2002). They therefore tried a cost function which avoided all slopes of > 9% (though Hucker (2009) reports gradients of 1 in 6, i.e. about 17%, as the critical slope for onset of zig-zagging of Roman roads). The modified Tobler function has a critical slope for cost-saving zig-zagging of 27%, as defined by Llobera & Sluckin's (2007) equation 20. As an experiment, we modified it to

force the critical slope to 9% by heavily penalizing the cost of travelling on slopes steeper than this. For the 117 TP legs it resulted in slightly longer average road sections (by about 2%), partly by lengthening the routes through switchbacks and partly by going around areas of high ground. However, we decided not to implement this as standard in part because of some uncertainty about how the resulting increase in frequency of a switchback road configuration would have been measured. Bekker-Nielsen (2004: 200) states that the Romans did not generally include the extra distance resulting from switchbacks in their surveyed distances, at the same time noting that one TP distance in Cyprus evidently *does* include the contribution from switchbacks, perhaps, he speculates, because its recorded mileage predated Roman surveyors.

The applicability to earlier periods of a value of λ which has been derived for a date of 280 CE is open to question. However, there is some evidence that the Roman road network in territories of the earlier Hellenistic world did not differ significantly from its antecedents. Verbrugge (1976), in his study of the road network of Sicily, found that all bar two of the major routes of the island were in place before the advent of the Romans, and that the routes used by travellers and the imperial post show no signs of Roman engineering or realignment. He concludes that where they found an adequate road network they made few attempts to improve or extend it. Pikoulas (2007: 85) makes a similar point in relation to the road network of the Peloponnese and the rest of southern Greece, where there are no known exclusively Roman road constructions. Taken with Rathmann's contention (2016) that the TP has Hellenistic origins these observations support the use of the derived value of $\lambda = 7$ for this earlier period, so long as the network is reconstructed with contemporary nodes.

5 NETWORKS OVER WATER

A network of water routes (over sea and lakes) was constructed using $\lambda = \infty$, on the assumption that boats are unconstrained in the directions they take, other than not being allowed to go over land points. An upper limit was placed on the amount of time allowed for a single journey over water of 36 hours which means that whilst routes are allowed across the Aegean Sea (aided by island hopping), journeys across the Ionian Sea were restricted to the narrower part between Epirus and Italy, consistent with a practice of cabotage, or not straying too far from the coast (Constantakopoulou 2007: 176; Rougé 1966: 173).

No account of prevailing winds or currents was taken, as is done e.g. in the ORBIS tool (*https://history.stanford. edu/publications/orbis-stanford-geospatial-network-model-roman-world*), mainly for the sake of simplicity, but also on the basis that the majority of edges represent journeys

close to land. Since these journeys are more commonly undertaken in summer, sea breeze effects are frequently likely to generate winds at odds with the climatological, open-sea values, giving a fairly reliable onshore wind component for much of the day regardless of the large-scale prevailing wind. According to Heikell (2018), such winds in the Mediterranean set in during the morning and last till evening, blowing onshore and extending as much as 50 miles out to sea. Not only is this cycle favoured by a hotter land, but is also reinforced by the inertial effects of the earth's rotation in much of the Mediterranean due to its proximity to 30 degrees latitude (Hsu 1988). At night and into the first part of the morning a weaker land breeze, blowing offshore, predominates, easing an early-morning exit from a harbour, the transition to the onshore sea breeze often being highly predictable. Ships of the time could make headway perpendicular to the wind, even somewhat against the wind (Gal, Saaroni & Cvikel 2021; Whitewright 2011), suggesting that the sea-land breeze cycle would have been much utilized by ancient mariners whose courses parallel to the coast would lie approximately perpendicular to the wind. Capturing the cycle in a quantitative way, however, is very challenging because the effect tends to cancel out over any averaging period of 24 hours or more and the numerical models used in generating such climatologies are generally too coarse to represent such small-scale effects. Gal, Saaroni & Cvikel (2021) have significantly improved the sophistication of estimates of travel under sail with both high resolution (27 km) meteorological data and sailing software which models ship performance in different conditions. However, they concede that this works principally for direct runs over open sea and that a separate method is required for modelling sea-breeze assisted coastal sailing.

In view of these difficulties, a constant, adjustable speed was specified for the network, which could be tuned up or down when combined with the land network values to form a super network. A small random perturbation of between plus and minus 1% is added to the distances between grid points resulting in trajectories which are not quite straight. Largely cancelling in aggregate, these perturbations change journey times by much less than 1%. Clearly, tacking courses adopted when sailing into the wind are not represented, the velocity being interpreted as an effective velocity, or 'velocity made good' as defined by Whitewright (2011). Any settlement within two kilometres of the sea or a lake is deemed to be able to access travel by boat, so no account is taken of e.g. the presence or absence of suitable deep water harbours, it being assumed that boats could load or unload passengers and cargo alike whilst lying at anchor (Votruba 2017). There is no representation of rivers either as conduits for travel by boat or as barriers to land travel, mainly because of the extra layers of complexity in implementation that this would have entailed.
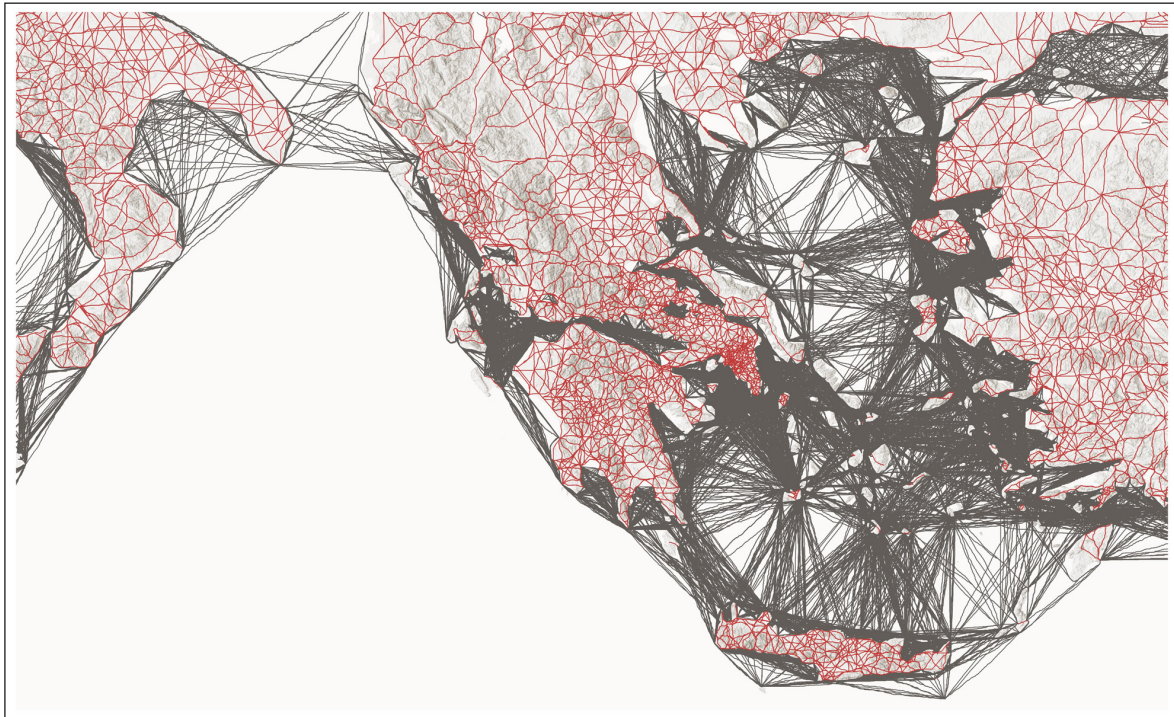
Figure 9 Section of combined land and sea network. Based on Pleiades locations extant in the late third century BCE.

Figure 9 shows part of the network of land and sea routes.

6 THE TRAVELLING SALESMAN PROBLEM

In attempting to reconstruct the routes taken by ancient travellers, the relative cost of land and sea travel becomes an important consideration affecting decisions ranging from whether an individual leg of a journey is more likely to have been made by land or by sea, to the whole order of an itinerary.

In order to calibrate the relative costs of land and sea travel, we turn to the Delphic *theorodokoi* list (DTL) dated to around 220–210 BCE (Perlman 1995; Plassart 1921). *Theorodokoi* acted as hosts for religious delegates (*theoroi*) visiting from afar, and cities which sent out these delegates kept lists inscribed in stone for public display containing the names of the *theorodokoi* they would call upon, alongside the names of their cities. The DTL originally contained some 300 cities across the Greek world to which the *theoroi* were sent to announce Delphic festivals. Such lists were usually organized geographically and within them some itineraries can be inferred (Perlman 2000; Rutherford 2013). Two itineraries from the DTL are used here to calibrate the sea speed, one for Cyprus and the other Crete, island itineraries being specifically chosen as they are more likely to allow both land and sea route options. To do this we extend the least cost path problem to finding the optimal route through a list of destinations which are nodes in a network.

6.1 TSP SOLUTION METHODOLOGY

This can be tackled as a variant of the classic travelling salesman problem (TSP), in which the most efficient (least cost) order in which to visit a set of locations is established, starting and finishing at the same point. Specifically, this is an asymmetric travelling salesman problem because costs between nodes vary with direction. The TSP can be solved for a small number of destinations (less than about 14) by using a brute force method whereby the cost of every ordering permutation is compared and the lowest selected. Since computation time is proportional to the factorial of the number of destinations, this method soon becomes impractical as the number increases. For example, going through permutations on a computer we were able to solve the TSP by comparing all possible routes for the nine locations in Cyprus (*Figure 10*) in about 0.006 seconds. Increasing this to 18 locations would have taken over three years using the same program, whilst the Crete example (*Figure 11*) with its 29 locations would have required a period many times in excess of the age of the universe. This being so, we use a technique which directs the search by allowing whole groups of permutations to be ruled out *a priori* via a *branch and bound* process. Specifically, direct use is made of a Fortran routine documented by Carpaneto, Dell'Amico & Toth (1995), which returns the optimal itinerary order on input of an anisotropic matrix of times taken between each pair of destinations.

For our purposes we depart from the pure TSP in two respects: in allowing the route to go through the same place twice if this produces a shorter overall journey, and

in allowing itineraries in which start and end locations are not the same. The solution algorithm is presented with an anisotropic matrix of costs between the nodes representing the destinations. The algorithm orders these in the most efficient way, then the network backlink matrix is used to find the intermediate nodes from the full network and the journey can be reconstructed. These intermediate nodes themselves may coincide with destinations, in which case the route is allowed to pass through that destination more than once. If it is required to specify which is the last place visited, the cost matrix is altered so that the cost from the nominated last destination to the start point is very small (e.g. 0.1 seconds), whilst those to all the other destinations from the nominated last place are very large (e.g. 1000 years).

Just as one may question the extent to which real paths between two nodes of a network really were optimal enough to be considered LCPs, it is a matter of conjecture whether ancient travellers would even have the data to compare the efficiency of different itineraries, let alone be able to solve what can be a computationally challenging task. However, the DTL itineraries are considered more useful than an account of a one-off journey since they result from the experience of regular journeys over multiple years during which there will have been ample opportunity for fine tuning.

6.2 DELPHIC *THEOROI* ITINERARY IN CYPRUS

Figure 10 (top panel) shows the order (1 – 9) in which *theoroi* visited destinations in Cyprus, as inferred from the Delphic *theorodokoi* list. Routes between destinations are derived by taking the quickest paths from a land and sea network with $\lambda = 7$ over land. The middle map shows a travelling salesman solution for the fastest route starting at Salamis and finishing at Arsinoe using a standard ship speed of 3.9 knots, close to Whitewright's (2011) 'favourable conditions' velocity made good for a ship
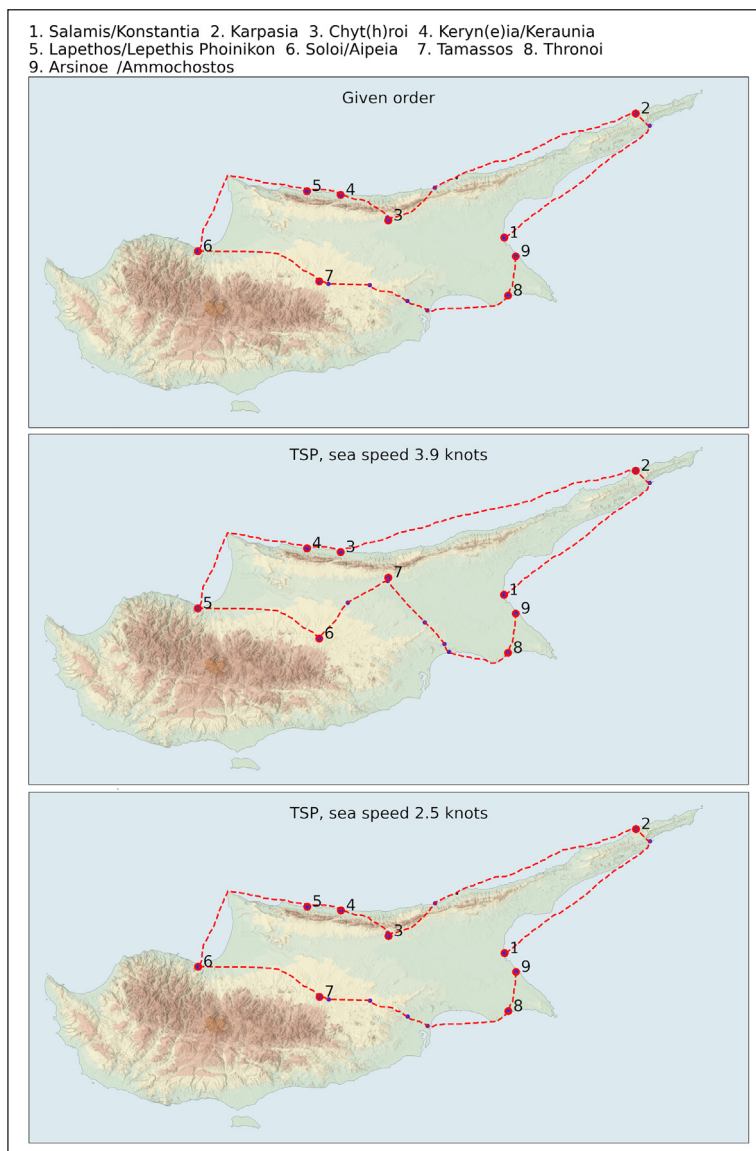

Figure 10 Delphic *theoroi* itinerary for Cyprus. Top: Actual itinerary from *theorodokoi* list with routes between numbered destinations inferred from network defined by $\lambda = 7$. Unnumbered are intermediate nodes in network derived from backlink matrix. Middle: TSP solution, sea speed 3.9 knots. Bottom: TSP solution, sea speed 2.5 knots.

of the first century BCE. This produces a different route, taking in all the *poleis* on the north coast sequentially, rather than breaking off to visit Chythroi between Karpasia and Keryneia as the actual itinerary indicates. Chythroi lies on the other side of the ridge of the Kyrenian mountains, so taking it in the given *theorodokoi* list order requires crossing the high ground twice, and the first TSP solution (middle panel) in which Chythroi is visited after Tamassos seems to make sense on the basis that it avoids this. Gradually reducing sea speed leads to a TSP solution (bottom panel) which matches the given order of the actual itinerary; this first happens at 2.5 knots, close to midway between Whitewright's sailing ship speeds for favourable and unfavourable conditions. When the sea speed is reduced to 1.6 knots, around Whitewright's

unfavourable conditions speed, an itinerary with no sea legs is produced (not shown) though the order in which the destinations are taken does not alter.

6.3 DELPHIC *THEOROI* ITINERARY IN CRETE

Figure 11 gives the itinerary of Delphic *theoroi* visiting Crete. They start with Kythera, an island off the southeast of the Peloponnese (not shown on map), and after visiting 27 settlements on Crete itself, it is inferred from the list that they went on to visit a number of *poleis* in Libya (Rutherford 2013: 75), starting with Cyrene. In view of this, the island of Kaudos is used to finish off the itinerary since it lies in the direction of Cyrene. A TSP problem was set with Kythera and Kaudos forced as start and end points respectively, and with a 3.9 knot sea speed (middle
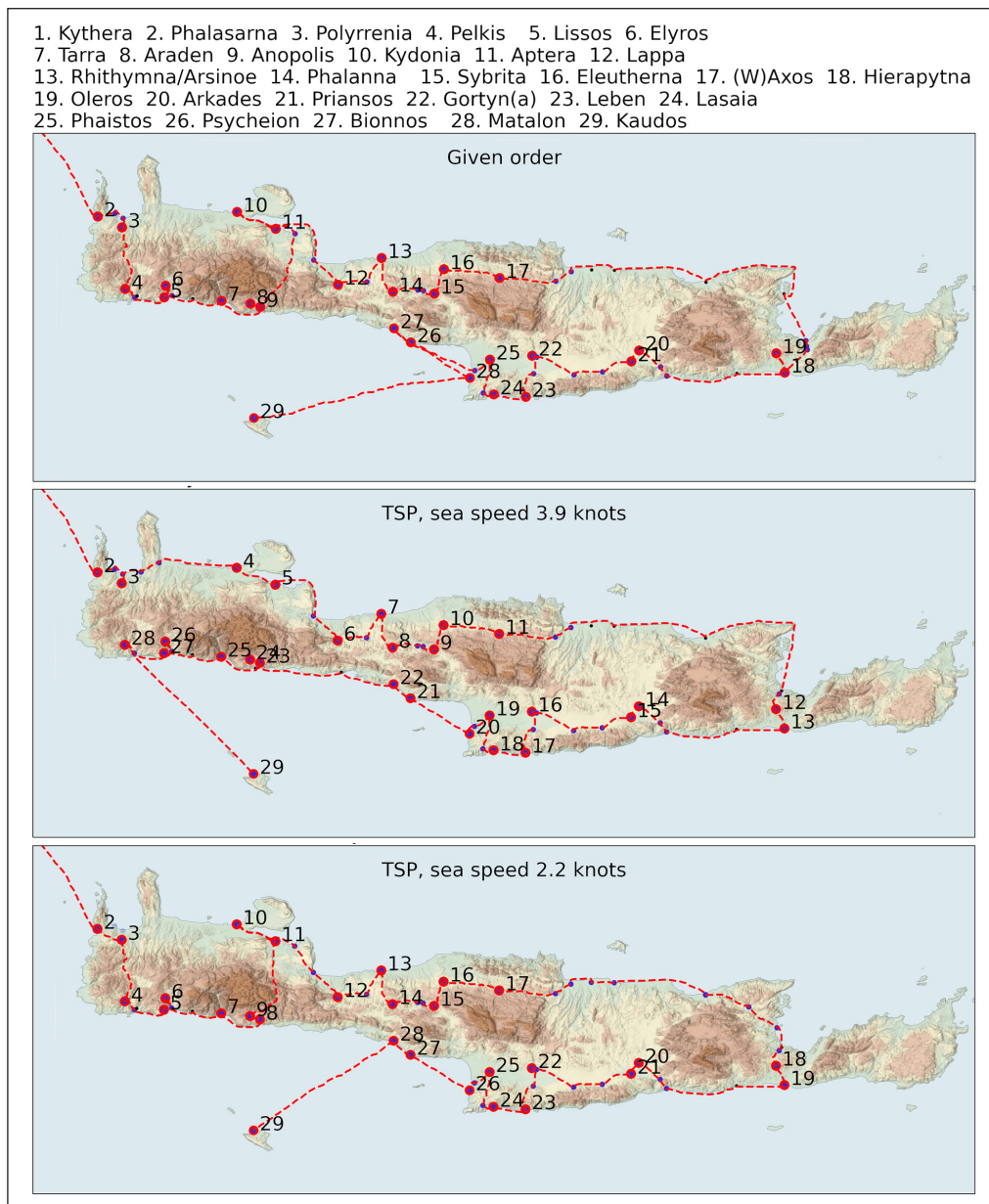

1. Kythera 2. Phalasarna 3. Polyrrenia 4. Pelkis   5. Lissos 6. Elyros
7. Tarra 8. Araden 9. Anopolis 10. Kydonia 11. Aptera 12. Lappa
13. Rhithymna/Arsinoe 14. Phalanna   15. Sybrita 16. Eleutherna 17. (W)Axos 18. Hierapytna
19. Oleros 20. Arkades 21. Priansos 22. Gortyn(a) 23. Leben 24. Lasaia
25. Phaistos 26. Psycheion 27. Bionnos   28. Matalon 29. Kaudos

Figure 11 Delphic *theoroi* itinerary for Crete. Top: Actual itinerary from *theorodokoi* list with routes between numbered destinations inferred from network defined by $\lambda = 7$. Unnumbered are intermediate nodes in network derived from backlink matrix. Middle: TSP solution, sea speed 3.9 knots. Bottom: TSP solution, sea speed 2.2 knots. Kaudos is inserted as an end point since it seems that the itinerary then continued with Cyrene in Libya.

panel *Figure 11*). This produces an itinerary in which the destinations are taken clockwise around the island with liberal use of boat trips between coastal settlements and the group of six in the southwest visited last rather than amongst the first, as in the actual itinerary. Again, gradually reducing the sea speed eventually leads to a sudden switch to a TSP solution which is much more similar to the actual itinerary – the threshold speed being 2.2 knots. It differs from the actual itinerary in the order in which locations in the last section, Phaistos to Matalon, are taken, which, in the given order involves some doubling back – perhaps because it was easier to take a boat to Libya from the port of Matalon.

Mount Pachnes (2453 m) lies between the Araden/Anopolis area (numbers 8 and 9 in top map) and Kydonia (10), forcing an indirect route which goes through Aptera (11), resulting in a double visit here. The TSP solution based on a more complex network of λ = 10 (not shown) allows a more direct route between nodes 9 and 10 bypassing 11, though the overall order in which named destinations are visited is not changed. Reducing the sea speed to 1.2 knots flips the TSP solution to a different configuration which in part reverts to the clockwise solution seen at speeds above 2.2 knots. All three itineraries depicted go through northeast Crete despite the absence there of supplied destinations. It is stated by Perlman (1995: 130) that about a dozen names lost from the *theorodokoi* inscription's fourth column, between Oaxos and Hierapytna, are probably from northeast Crete.

6.4 OTHER EXPERIMENTS

It is documented that there was an overland *theoroi* route from Athens to Delphi. Going through a similar exercise (not shown) for a TSP return journey from Athens to Delphi gives a threshold sea speed of 2.1 knots to force a land journey. Since this route was taken as a symbolic procession (Rutherford 2013: 183), the overland itinerary may well have been followed for considerations which go beyond those of travel cost, though at the same time could have evolved originally for such reasons.

In order to test the sensitivity to networks of differing complexity, the experiment was repeated with networks of λ = 4 and 10, with the results given in *Table 1*. The simpler Cyprus itinerary showed no sensitivity over this range, whilst that for Crete showed an increase in threshold sea speed as the land network became more complex. It is notable that the ability of the TSP algorithm to reproduce similar routes does not seem too sensitive to defined network complexity.

6.5 DISCUSSION

Overall the TSP experiments provide some data points which suggest that a sea speed in the range of around 2.0 to 2.5 knots should be specified in a combined land/sea network based on time costs. In addition to fitting centrally in Whitewright's (2011) range of speeds for ships of the first century BCE, it is consistent with de Soto (2019: 281) who estimates an average speed of 2.3 knots for coasting journeys. The results also suggest that plausible reconstructions of journeys can be made by applying a sea speed in this range to groups of destinations for which we have no itineraries.

Note that it is the ratio of land to sea speed which is important in this context, rather than absolute sea speed. Tobler's hiking function (Tobler 1993) is about 23% faster than the Modified Tobler function when walking on the flat, and using it would suggest sea speeds faster in proportion (about 2.5 to 3.0 knots). However, given the evidence that the Modified Tobler function is more accurate (Márquez-Pérez, Vallejo-Villalta & Álvarez-Francoso 2017; Seifried & Gardner 2019) the lower range is preferred. It might be more realistic to define the time costs of sea travel not merely by a speed, but also by a fixed cost which could represent the time taken to wait for a boat to set sail. It should be borne in mind that the time costs represented by these speeds can also be taken as proxies for a number of other considerations such as monetary cost, physical effort, organizational difficulties and risks associated with the different modes of travel, and that the speeds of land and sea travel would better be considered as virtual rather than actual speeds, a term which is expanded upon in Appendix 2. Another consideration is that time measured in elapsed days needs to be calculated separately since this depends on the spacing of nodes in the network, with routes which go through nodes conveniently separated by a day's walk over land making more efficient use of travel time.

7 CONCLUSION

This study has brought to bear three different quantitative optimization procedures on the study of ancient travel. The well-known least cost path technique has been extended by a *rotate and oversample* method which improves its error characteristics and directional biases at little extra computational expense. The problem of network construction has been tackled by framing it as one of minimization of travel and building/maintenance costs for which a *minimum total cost* principle has been formulated. The λ parameter (ratio of travel to build cost) required to achieve this has been tuned by use

	λ = 4	λ = 7	λ = 10
Cyprus	2.5	2.5	2.5
Crete	1.9	2.2	2.3

Table 1 Threshold sea speeds (knots) at which TSP routes become similar to Delphic *theoroi* routes for networks of different complexity (λ).

of a subset of the Peutinger Table, whose inferred error characteristics have been explored. An appropriate sea speed has been estimated for use in a combined sea and land network by treating known itineraries of Delphic *theoroi* in Crete and Cyprus as solutions to travelling salesman problems. The good agreement between optimized and actual itineraries suggests that, given a list of destinations whose order is unknown, other journeys could be reconstructed using the combined sea and land network.

The same principles outlined in this paper could be applied to other geographic areas, such as the wider Roman empire, as well as to later or earlier periods. A source of data on node locations, such as Pleiades, is necessary along with a suitable DEM and some means of estimating an appropriate network complexity metric, *λ*. This could be done, as here, by using a dataset of inter-node road distances, though it could also be derived through other means such as tuning to an estimated *β* value. Use of ancient lakes and coastlines to modify the DEM is not necessarily a prerequisite of successful application of the technique, whilst, conversely, it may be appropriate to include in the hiking model the effects of other terrain costs where terrain height is not considered to be the main determinant in the choice of routes. The resulting network could also be turned to other uses such as clustering studies or the evaluation of network diagnostics, e.g. betweenness centrality, in order to evaluate the importance of individual nodes.

The authors are willing to make available the software used to generate networks of least cost paths upon reasonable request.

APPENDIX 1. ALGORITHMIC DETAILS OF NETWORK CREATION USING THE MINIMUM TOTAL COST METHOD

The technique relies firstly on establishing the $N \times N$ first-guess cost matrix, FGCOST, where N is the number of locations in the network, and FGCOST (i, j) represents the cost of going directly along an LCP from place i to place j. Multiple LCPs are derived linking all places to all others within a specified temporal radius, MAXRAD (36 hours for experiments here), in what is effectively a 'least cost to the user' network. For maximum convenience at the expense of significant loss of accuracy, geodesic (straight-line) distances could have been used. For full accuracy, LCPs may be derived at full resolution using the *rotate and oversample* method (*Figure 1*). The compromise adopted for this study to reduce computation time was to use the *undersample, rotate and oversample* method. It is important to note that these LCP costs are simply used to inform the decision-making about which connections to make in the network, and that LCPs are regenerated to

the required precision afterwards.

Two alternative algorithms have been coded for network construction. The first is termed *global pairwise optimization* (GPO) and at each stage chooses the road, the building of which confers the greatest benefit per unit distance between any pair of locations within MAXRAD hours of each other. It consists of the following steps:

(1) Create a new $N \times N$ cost matrix COST, each element initially populated by an arbitrarily large number. 10^{10} was used in the experiments.
(2) Calculate the matrix BENEFIT = COST/FGCOST – (1 + 1/$λ$). This gives the net benefit (negative total cost) of creating a road between each pair of places per unit temporal distance of that road. Take the indices of the largest element of BENEFIT, *imax* and *jmax*. If BENEFIT(*imax,jmax*) is greater than 0, establish the LCP between locations *imax* and *jmax*, populating COST(*imax,jmax*) and COST(*jmax, imax*) with the resulting values. If BENEFIT(*imax, jmax*) is less than 0 the network has been completed and the algorithm terminates.
(3) Run the local elements of the COST matrix (within MAXRAD hours of the nodes at either end of the route just established) through Dijkstra's algorithm to update network distances in the light of the latest route added. Go back to step (2).

The GPO method has some similarities to the EE (*equitable efficiency*) model of Prignano *et al.* (2019), though they choose the next path on the basis of that which has the lowest FGCOST/COST and do not weigh the benefit and building cost of each path. The GPO method is efficient from the point of view of coding and replicates a network evolution whereby decisions about where to build the next road are made collectively at a network-wide level on the basis of which pair of settlements across the network will benefit most. The most computationally expensive part of the process, that of running Dijkstra's algorithm, is reduced by running it only on a subset of the cost matrix. The technique could be made more comprehensively global if the establishment of each link were tested on the basis of the benefit it conferred not just on the pair of settlements it links but on all settlement linkages simultaneously. However, this would involve running Dijkstra's algorithm on the full network many times, so would be prohibitively expensive.

A second approach named *local outward search* (LOS) has also been developed. Taking each node in turn, and working outwards to other settlements, decisions on which direct roads to establish are made on the basis of the existence or otherwise of potential staging-posts, or stepping stones.

(1) Set i = 1.
(2) Examine values of FGCOST(i, j = 1..N). Establish a list of potential destinations by rejecting those which are greater than MAXRAD hours. Of the remaining destination list, put them in ascending order of FGCOST.
(3) Establish an LCP to the closest j and populate values of COST(i,j) and COST(j, i).
(4) For each subsequent destination j find amongst the closer places a potential staging-post k which is closer to j, i.e. FGCOST(k, j) < FGCOST(i, j) and which also satisfies the condition that FGCOST(i, k) + FGCOST(k, j) < (1 + 1/λ)FGCOST(i, j). If there is more than one, take that with the lowest FGCOST(i, k) + FGCOST(k, j). If a suitable k is found, reject j as a direct destination, increase its distance to reflect the new indirect route to it, and move onto next until all have been explored. Otherwise, unless such a link already exists, establish LCP between i and j and populate values of COST(i, j) and COST(j, i).
(5) Increment i. If i = N + 1 terminate algorithm. Otherwise go to (2).

The LOS method differs from the GPO method in placing the decision-making for road building at the level of the individual settlement, with regard only to optimization of routes from each settlement separately. The order in which the start-point settlements is taken makes no difference to the final network. It has the important limitation of only allowing networks for which $\lambda > 1$ due to the specification that the distances from start to staging-post and staging-post to destination each be less than the distance from start to destination. It is possible to lift this constraint in a variant of the scheme.

APPENDIX 2. VIRTUAL SPEED

When calculating the cost surfaces for the purposes of establishing LCPs, the most easily quantified costs are time or energetic output. In reality, a whole host of costs can simultaneously be attributed to travel including monetary expenditure, organizational effort, and, along with a direct energetic cost, the mental and physical stress of a journey. A second class of cost is the risk of mishap, such as falling prey to pirates or brigands, or of shipwreck. Whilst these eventualities will not materialize on most journeys, even a low probability that they might can be treated as a continuous cost.

It is not easy to represent these costs without introducing complexity, much less to quantify them accurately. However, we can do so approximately and with minimal complexity by introducing the concept of *virtual speed*.

Taking the example of time (t) as a measurement of cost, the distance (D) covered per unit cost is speed (S), as in

$$S = D/t$$

Generalizing to include other costs, we define a *virtual speed* as the distance per unit total cost incurred, given by:

$$S_v = D/(t + M + Ph + ...)$$

where S_v is virtual speed, M is money spent, P is the probability of encountering a hazard, h is the cost of that hazard, e.g. loss of life or limb, then a series of other potential cost factors. Clearly, the different costs on the denominator have different units, though could all be converted to the same units given a suitable conversion factor in the same way that in modern-day economics a life is assigned a monetary cost when making decisions on spending on transport safety measures (Hauer 1994). Since the non-time costs can be taken as more or less proportional to time spent travelling, we could rewrite the expression as

$$S_v = D/t(1 + m + \text{ph} + ...) = S/(1 + m + \text{ph} + ...)$$

where m is the money expended per unit time, p is the probability per unit time.

When ancient travellers made journeys, it seems likely that their choice of route and mode of travel would be made not just on the basis of time or effort, but on an amalgam of costs as outlined above. It might be the case that a leg of a journey could be made more quickly by sea than by land, but that the traveller would opt to go by land because it was safer. Alternatively, it might be quicker to go by land than by sea but the traveller might chose the sea route because it involved less energy expenditure.

When using the Delphic *theorodokoi* lists to weight the relative speeds of land and sea travel, it is really a virtual speed which we are ascertaining since these routes are likely to have been established over many years with a view to minimizing all costs, not simply that of time. The ratio of the virtual speed over land to that over sea is the quantity which determines the route taken. The actual travelling time may have been different from that derived by using the virtual speed as an actual speed. Of course, such treatments are necessarily approximate since there will be geographically and temporally varying costs, e.g. of piracy and shipwreck, which would require a higher order of complexity to represent. Applying a uniform value is equivalent to the approximation entailed in using a set sea speed without regard for the geographical and seasonal variation in winds and currents.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Fiona Carroll orcid.org/0000-0002-5346-7400
Classics and Ancient History, University of Exeter, GB
Edward Carroll orcid.org/0000-0001-5384-7633
Independent researcher, GB

REFERENCES

Barker, E, Simon, R, Isaksen, L and **de Soto Canamares, P.** 2016. The Pleiades gazetteer and the Pelagios project. In: Berman, ML, Mostern, R and Southall, H (eds.), *Placing Names: Enriching and Integrating Gazetteers,* 97–109. Bloomington: Indiana University Press. DOI: *https://doi.org/10.2307/j.ctt2005zq7.12*

Bekker-Nielsen, T. 2004. *The roads of ancient Cyprus.* Copenhagen: Museum Tusculanum Press.

Bertoldi, S, Castiglia, G and **Castrorao Barba, A.** 2019. A Multi-scalar Approach to Long-Term Dynamics, Spatial Relations and Economic Networks of Roman Secondary Settlements in Italy and the Ombrone Valley System (Southern Tuscany): Towards a Model? In: Verhagen, P, Joyce, J and Groenhuijzen, MR (eds.), *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*, 191–214. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-030-04576-0_10*

Bevan, AH. 2013. Travel and interaction in the Greek and Roman world: a review of some computational modelling approaches. In: Dunn, S and Mahoney, S (eds.), *The Digital Classicist 2013,* 3–24. London: University of London Press.

Brughmans, T. 2010. Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, 29(3): 277–303. DOI: *https://doi.org/10.1111/j.1468-0092.2010.00349.x*

Carpaneto, G, Dell'Amico, M and **Toth, P.** 1995. Exact solution of large-scale, asymmetric traveling salesman problems. *ACM Transactions on Mathematical Software*, 21(4): 394–409. DOI: *https://doi.org/10.1145/212066.212081*

Carreras, C and **de Soto, P.** 2013. The Roman transport network: a precedent for the integration of the European mobility. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(3): 117–133. DOI: *https://doi.org/10.1080/01615440.2013.803403*

Constantakopoulou, C. 2007. *The dance of the islands: insularity, networks, the Athenian empire, and the Aegean world*. Oxford: Oxford University Press. DOI: *https://doi.org/10.1017/S0009840X08002485*

De Soto, P. 2019. Network Analysis to Model and Analyse Roman Transport and Mobility. In: Verhagen, P, Joyce, J and Groenhuijzen, MR (eds.), *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*, 271–289. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-030-04576-0_13*

Dijkstra, EW. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1): 269–271. DOI: *https://doi.org/10.1007/BF01386390*

Elsner, J. 2000. The Itinerarium Burdigalense: Politics and salvation in the geography of Constantine's empire. *The Journal of Roman Studies*, 90: 181–195. DOI: *https://doi.org/10.2307/300206*

Evans, TS and **Rivers, RJ.** 2017. Was Thebes Necessary? Contingency in Spatial Modeling. *Frontiers in Digital Humanities*, 4: 8. DOI: *https://doi.org/10.3389/fdigh.2017.00008*

Gal, D, Saaroni, H and **Cvikel, D.** 2021. A new method for examining maritime mobility of direct crossings with contrary prevailing winds in the Mediterranean during antiquity. *Journal of Archaeological Science*, 129: 105369. DOI: *https://doi.org/10.1016/j.jas.2021.105369*

Groenhuijzen, MR and **Verhagen, P.** 2017. Comparing network construction techniques in the context of local transport networks in the Dutch part of the Roman limes. *Journal of Archaeological Science: Reports*, 15: 235–251. DOI: *https://doi.org/10.1016/j.jasrep.2017.07.024*

Hauer, E. 1994. Can one estimate the value of life or is it better to be dead than stuck in traffic? *Transportation research part A: policy and practice*, 28(2): 109–118. DOI: *https://doi.org/10.1016/0965-8564(94)90032-9*

Heikell, R. 2018. *The Adlard Coles Book of Mediterranean Cruising* (4th Edition). Camden: Bloomsbury Publishing.

Herzog, I. 2013. The potential and limits of optimal path analysis. In: Bevan, A and Lake, M (eds.), *Computational Approaches to Archaeological Spaces*, 179–211. Walnut Creek, CA: Left Coast Press. DOI: *https://doi.org/10.1080/00293652.2013.773368*

Herzog, I. 2014a. Least-cost paths – some methodological issues. *Internet Archaeology*, 36. DOI: *https://doi.org/10.11141/ia.36.5*

Herzog, I. 2014b. Least-cost networks. In: Earl, G, et al. (eds.), *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA),* Southampton, 26–29 March 2012, 237–248. Amsterdam: Amsterdam University Press. DOI: *https://doi.org/10.1515/9789048519590*

Hsu, SA. 1988. *Coastal meteorology*. New York: Academic Press.

Hucker, RA. 2009. How did the Romans achieve straight roads? Paper Presented at *International Federation of Surveyors (FIG) Working Week 2009: Surveyors' Key Role in Accelerated Development,* Eilat 3–8 May.

Isaksen, L. 2008. The application of network analysis to ancient transport geography: A case study of Roman Baetica. *Digital Medievalist*, 4. DOI: *https://doi.org/10.16995/dm.20*

Kansky, KJ. 1963. *Structure of transportation networks: relationships between network geometry and regional characteristics*. Chicago: University of Chicago.

Knappett, C. 2011. *An archaeology of interaction: network perspectives on material culture and society.* Oxford: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:osobl/9780199215454.001.0001*

Langmuir, E. 1984. *Mountaincraft and leadership: a handbook for mountaineers and hillwalking leaders in the British Isles.* Edinburgh: Scottish Sports Council.

Llobera, M and **Sluckin, TJ.** 2007. Zigzagging: Theoretical insights on climbing strategies. *Journal of Theoretical Biology*, 249(2): 206–217. DOI: *https://doi.org/10.1016/j.jtbi.2007.07.020*

Malkin, I, Constantakopoulou, C and **Panagopoulou, K.** 2007. Preface: Networks in the ancient Mediterranean. *Mediterranean Historical Review*, 22(1): 1–9. DOI: *https://doi.org/10.1080/09518960701539414*

Manière, L, Crépy, M and **Redon, B.** 2021. Building a Model to Reconstruct the Hellenistic and Roman Road Networks of the Eastern Desert of Egypt, a Semi-Empirical Approach Based on Modern Travelers' Itineraries. *Journal of Computer Applications in Archaeology*, 4(1): 20–46. DOI: *https://doi.org/10.5334/jcaa.67*

Márquez-Pérez, J, Vallejo-Villalta, I and **Álvarez-Francoso, JI.** 2017. Estimated travel time for walking trails in natural areas. *Geografisk Tidsskrift-Danish Journal of Geography*, 117(1): 53–62. DOI: *https://doi.org/10.1080/00167223.2017.1316212*

McHugh, M. 2019. Going the extra mile: Travel, time and distance in classical Attica. *The Annual of the British School at Athens*, 114: 207–240. DOI: *https://doi.org/10.1017/S0068245419000017*

Minetti, AE, Moia, C, Roi, GS, Susta, D and **Ferretti, G.** 2002. Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, 93(3): 1039–46. DOI: *https://doi.org/10.1152/japplphysiol.01177.2001*

Parcero-Oubiña, C, Güimil-Fariña, A, Fonte, J and **Costa-García, JM.** 2019. Footprints and cartwheels on a pixel road: On the applicability of GIS for the Modelling of Ancient (Roman) Routes. In: Verhagen, P, Joyce, J and Groenhuijzen, MR (eds.), *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*, 291–311. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-030-04576-0_14*

Perlman, PJ. 1995. Theorodokountes en tais polesin. Panhellenic Epangelia and Political Status. In: Hansen, MH (ed.), *Sources for the Ancient Greek City-State. Symposium* August, 24–27 1994. *Acts of the Copenhagen Polis Centre Vol. 2*, 113–164. Copenhagen: Munksgaard.

Perlman, PJ. 2000. *City and sanctuary in ancient Greece: the theorodokia in the Peloponnese.* Göttingen: Vandenhoeck & Ruprecht. DOI: *https://doi.org/10.13109/9783666252181*

Pikoulas, Y. 2007. Travelling by land in ancient Greece. In: Adams, R (ed.), *Travel, geography and culture in ancient Greece, Egypt and the Near East.* Oxford: Oxbow Books.

Pikoulas, Y. 2012. *Το οδικό δίκτυο της Λακωνικής (The Road-network of Lakonikē).* Horos: Athens.

Plassart, A. 1921. Inscriptions de Delphes, la liste des Théorodoques. *Bulletin de correspondance hellénique*, 45(1): 1–85. DOI: *https://doi.org/10.3406/bch.1921.3046*

Prignano, L, Morer, I, Fulminante, F and **Lozano, S.** 2019. Modelling terrestrial route networks to understand inter-polity interactions (southern Etruria, 950-500 BC). *Journal of Archaeological Science*, 105: 46–58. DOI: *https://doi.org/10.1016/j.jas.2019.02.007*

Raepsaet, G. 2002. *Attelages et techniques de transport dans le monde gréco-romain.* Brussels: Le Livre Timperman.

Rathmann, M. 2016. 19 The Tabula Peutingeriana and Antique Cartography. In: Bianchetti, S, Cataudella, M and Gehrke, H-J (eds.), *Brill's Companion to Ancient Geography,* 335–362. Leiden: Brill. DOI: *https://doi.org/10.1163/9789004284715_020*

Rougé, J. 1966. *Recherches sur l'organisation du commerce maritime en Méditerranée sous l'Empire romain.* Paris: Imprimerie Nationale. DOI: *https://doi.org/10.1086/365340*

Rousset, D. 2013. Le stadiasme de Patara et la géographie historique de la Lycie: itinéraires et routes, localités et cités. In: Brun, P, et al. (eds.), *Euploia. La Lycie et la Carie antiques. Actes du colloque de Bordeaux*, 63–76. Bordeaux: Ausonius Éditions. DOI: *https://doi.org/10.1017/S1047759418002209*

Rutherford, I. 2013. *State-Pilgrims and Sacred Observers in Ancient Greece: A Study of Theoria and Theoroi.* Cambridge: Cambridge University Press. DOI: *https://doi.org/10.1017/CBO9781139814676*

Salway, B. 2001. Travel, itineraria and tabellaria. In: Adams, C and Laurence, R. (eds.) *Travel and geography in the Roman Empire,* 32–76. London/New York: Routledge.

Salway, B. 2005. The nature and genesis of the Peutinger Map. *Imago Mundi*, 57(2): 119–135. DOI: *https://doi.org/10.1080/03085690500094867*

Sanders, GD and **Whitbread, IK.** 1990. Central places and major roads in the Peloponnese. *Annual of the British School at Athens*, 85: 333–361. DOI: *https://doi.org/10.1017/S0068245400015719*

Seifried, RM and **Gardner, CAM.** 2019. Reconstructing historical journeys with least-cost analysis: Colonel William Leake in the Mani Peninsula, Greece. *Journal of Archaeological Science: Reports*, 24: 391–411. DOI: *https://doi.org/10.1016/j.jasrep.2019.01.014*

Surface-Evans, SL and **White, DA.** 2012. An introduction to the least cost analysis of social landscapes. In: White, DA and Surface-Evans, SL (eds.), *Least Cost Analyses of Social Landscapes: Archaeological Case Studies*, 1–10. Salt Lake City: University of Utah Press. DOI: *https://doi.org/10.11141/ia.34.7*

Talbert, RJA and **Bagnall, RS.** 2000. *Barrington Atlas of the Greek and Roman World.* Princeton: Princeton University Press.

Tobler, W. 1993. *Three presentations on geographical analysis and modeling. Non–Isotropic Modeling, Speculations on the Geometry of Geography, Global Geographical Analysis (93–1).* Santa Barbara, CA: National Center for Geographic

Information and Analysis, UC Santa Barbara. DOI: *https://doi.org/10.29268/stsc.2014.2.4.1*

Verbrugge, G. 1976. *Sicilia (Itinera Romana, 2)*. Bern: Kimmerly and Frey.

Verhagen, P. 2018. Spatial analysis in archaeology: moving into new territories. In: Siart, C, Forbriger, M and Bubenzer, O (eds.), *Digital Geoarchaeology*, 11–25. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-319-25316-9_2*

Verhagen, P, Brughmans, T, Nuninger, L and **Bertoncello, F.** 2014. The long and winding road: combining least cost paths and network analysis techniques for settlement location analysis and predictive modelling. In: Earl, G, et al. (eds.), *Archaeology in the digital era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA),* Southampton, 26–29 March 2012, 357–366. Amsterdam: Amsterdam University Press. DOI: *https://doi.org/10.1515/9789048519590-039*

Verhagen, P, Nuninger, L and **Groenhuijzen, MR.** 2019. Modelling of pathways and movement networks in archaeology: an overview of current approaches. In: Verhagen, P, Joyce, J and Groenhuijzen, MR (eds.), *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*, 217–249. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-030-04576-0_11*

Verhagen, P, Polla, S and **Frommer, I.** 2014. Finding Byzantine Junctions with Steiner Trees. In: Polla, S and Verhagen, P (eds.), *Computational approaches to movement in archaeology: Theory, Practice and Interpretation of Factors and Effects of Long Term Landscape Formation and Transformation*, 73–98. Berlin/Boston: De Gruyter. DOI: *https://doi.org/10.1515/9783110288384.73*

Votruba, GF. 2017. Did vessels beach in the ancient Mediterranean? An assessment of the textual and visual evidence. *The Mariner's Mirror*, 103(1): 7–29. DOI: *https://doi.org/10.1080/00253359.2017.1274138*

Waugh, D. 2000. *Geography: An integrated approach*. Cheltenham: Nelson Thornes.

Whitewright, J. 2011. The potential performance of ancient Mediterranean sailing rigs. *International Journal of Nautical Archaeology*, 40(1): 2–17. DOI: *https://doi.org/10.1111/j.1095-9270.2010.00276.x*

Woolf, G. 2016. Only connect? Network analysis and religious change in the Roman World. *Hélade. Revista de História Antiga*, 2(2): 43–58. DOI: *https://doi.org/10.22409/rh.v2i2.10568*

Yang, L, Meng, X and **Zhang, X.** 2011. SRTM DEM and its application advances. *International Journal of Remote Sensing*, 32(14): 3875–3896. DOI: *https://doi.org/10.1080/01431161003786016*

Zipf, GK. 1949. *Human behavior and the principle of least effort: an introduction to human ecology.* Cambridge, MA: Addison-Wesley. DOI: *https://doi.org/10.1126/science.110.2868.669*