Geospatial modeling of pedestrian transportation networks: a case study from precolumbian Oaxaca, Mexico

Devin A. White a,*, Sarah B. Barber b

a Crow Canyon Archaeological Center, 23390 Road K Cortez, CO 81321, USA
b Department of Anthropology, University of Central Florida, Howard Phillips Hall 309, 4000 Central Florida Blvd., Orlando, FL 32816, USA

ARTICLE INFO

Article history:
Received 9 December 2011
Received in revised form 11 March 2012
Accepted 11 April 2012

Keywords:
Oaxaca
GIS
Movement
Networks
Pathfinding
High performance computing
Least cost path

ABSTRACT

Using accumulated cost surfaces and various pathfinding techniques within Geographic Information Systems (GIS) software, archaeologists and other spatial scientists have developed increasingly sophisticated models of human movement. Despite their utility, these approaches can be limited because standard GIS software cannot model movement (1) from many origins to many destinations or (2) without specific origins and destinations. Absent these capabilities, it is particularly difficult to model networks of movement over a given tract of land if you are interested in obtaining a more general sense of movement dynamics, not specific site-to-site patterns. In this paper, we present an innovative way of modeling past movement that generates both natural-looking networks and also indicates the degree of traffic that may have existed on any particular segment of those networks. The “From Everywhere to Everywhere” (FETE) model generates networks based on topography and landcover without requiring that origin and destination points be supplied in advance. We apply the FETE model to a case from the southern Mexican state of Oaxaca, a region that has extensive archaeological and ethnohistoric data sets that serve as a test of the efficacy of our technique. A comparison of the FETE output with known late precolumbian and early colonial movement corridors indicates that the method is effective and should be useful for modeling networks in other areas.

1. Introduction

For more than a decade, the use of geographic information systems (GIS) in archaeology has enabled increasingly sophisticated means of modeling human movement at various geographic scales (e.g., Bellavia, 2006; Fabrega-Alvarez, 2006; Harris, 2000; Howey, 2011; Whitley and Hicks, 2003; White and Surface-Evans, 2012). Leveraging digital data on elevation and sometimes land cover (Howey, 2007), GIS-centric movement studies in archaeology have been used primarily to create models that mathematically predict where movement was likely to have been channeled on a specific piece of terrain based on preconditions defined by the investigator (e.g., Harris, 2000). The basic methodology for most GIS movement studies draws on functions available within commercial and open source GIS software packages that are now well established within archaeology (Howey, 2011; Collischonn and Pilar, 2000; Llobera et al., 2011). First, a cost surface (also referred to as a cost-of-passage map or friction surface) is generated that numerically expresses the difficulty of moving between individual cells in a raster grid given a specific mode of transportation (Collischonn and Pilar, 2000:397; Llobera et al., 2011). Next, an accumulated cost surface is created that represents the cost of moving away from or towards a specific cell on a map based on the characteristics of the original cost surface (Howey, 2007:831; Llobera et al., 2011:844). Finally, the least costly routes of movement to or from the origin cell can be represented graphically using an algorithm appropriate to the mode of transportation (Harris, 2000; Tobler, 1993; Surface-Evans and White, 2012). Generally known as least cost paths, these mathematically generated movement models have been used in archaeology to explain site location (Carballo and Pluckhahn, 2007; Bell and Lock, 2000), to identify long-distance trade routes (e.g., Sherman, et al., 2010), to understand regional social and economic networks (e.g., Howey, 2007), and to identify possible routes for the colonization of new territories (e.g., Rademaker, et al., 2012), among others.

* Corresponding author. Tel.: +970 565 8975; fax: +970 565 4859.
E-mail addresses: devin.white@gmail.com (D.A. White), sbarber@mail.ucf.edu (S.B. Barber).

© 2012 Elsevier Ltd. All rights reserved.

doi:10.1016/j.jas.2012.04.017
Despite their clearly demonstrated value to archaeology, GIS movement studies employing these methods can be limited in several widely recognized ways—depending on the archaeological questions of interest to the researcher (Harris, 2000; Conolly and Lake, 2006:252–255). For example, accumulated cost surfaces can only be generated in reference to a specific origin point predetermined by the investigator. While this aspect of movement modeling poses no problem if the research is focused on a small number of known locations, difficulties arise if the goal of the research is to understand movement without reference to predetermined locations (Bellavia, 2006, 2002; Whitley and Hicks, 2003), if site locations are not known in advance (Barber et al., 2011), or when trying to determine if a site was positioned as a result of a pre-existing transportation route (e.g., Bell and Lock, 2000). The origin requirement poses further problems when the research question requires the generation of least cost paths because currently available software is only capable of modeling single paths between two points, or a one-to-one path relationship (Fig. 1a). This type of path relationship can be extended to include simple one-to-many and many-to-one (Fig. 1b,c) relationships by pairing the same point with many other points to create multiple single paths that share either the same origin or destination. A handful of scholars have succeeded in creating more complex one-to-many and many-to-one movement models using GIS hydrology functions (Fabrega-Alvarez, 2006; Llobera et al., 2011). In these approaches, routes to or from a specified location are calculated as the “flow” of movement over an accumulated cost surface. Areas of high flow are identified as least cost paths.

Modeling more complex movement, such as a network between many origins and many destinations, has proven challenging because GIS software cannot easily calculate many-to-many path relationships (Fig. 1d). Several scholars have attempted to devise workarounds, including: conflating multiple one-to-one/one-to-many/many-to-one paths (Howey, 2007; van Leusen, 2002; Bell et al., 2002); conflating and thresholding hydrology models generated for multiple sites (Fabrega-Alvarez, 2006; Parcero-Oubina et al., in press); generating localized accessibility surfaces (Llobera et al., 2011; Llobera, 2000); creating cost-catchments around one or more locations (Surface-Evans, 2012; Ullah, 2011); using circuit modeling to calculate the probability of multiple pathways in and out of sites (Howey, 2011); full network reconstruction based on known sites (White, 2007); and “natural pathway” reconstruction using hydrological modeling that may (Whitley and Hicks, 2003) or may not (Bellavia, 2002, 2006) build on previously established path locations. Apart from the work of White, Bellavia, and Whitley and Hicks, these networks and network proxies are constrained by the analytic weight put on their origin points (or, in the case of Llobera et al., 2011; their destination points), which inherently create spoke-like networks where movement radiates out from fixed points instead of capturing the more fluid nature of real-world movement. Although more natural-looking travel networks are reconstructed in the work of White, Bellavia, and Whitley and Hicks, the significance of any given path segment, including the likelihood that it might be traversed, is unknown.

In this study, we introduce an innovative way of generating networks using accumulated path maps that avoids the problems created by the origin and one-to-one path requirements of GIS and provides insight into the level of traffic on different sections of a reconstructed network. Using an archaeological case from pre-columbian southern Mexico, we demonstrate the problems inherent in currently available techniques for modeling movement, especially when destinations are unknown. We then discuss the “From Everywhere To Everywhere” (FETE) approach for generating probabilistic travel networks where neither origins nor destinations are assumed.

2. Modeling networks in geospatial software

To exemplify the problems with currently available methods for modeling movement within most GIS software, we turn to an example from the archaeological record of Oaxaca, Mexico (Fig. 2). Covering 93,757 km², Oaxaca is a modern state located on and just to the west of the narrow Isthmus of Tehuantepec (INEGI, 2011). Landcover ranges from semi-deciduous subtropical forest in the humid Pacific and Gulf coastal lowlands to semi-arid in the highlands, where peaks can reach 3700 m above sea level (Joyce, 2010:37–42). The extremely mountainous terrain that characterizes much of Oaxaca was—and remains today—a significant impediment that channels human movement through a limited range of topographic breaks. Like the rest of North America, movement in pre-columbian Oaxaca was limited to pedestrian and canoe traffic (Hassig, 1985). The only significant navigable bodies of water in pre-columbian Oaxaca would have been a series of eight coastal lagoons and several rivers in the Atlantic watershed that are outside of the study area.

2.1. Modeling movement without destinations

Despite the challenges posed by its topography, Oaxaca was home to numerous pre-columbian complex societies, beginning with the founding of Monte Albán in the highland Valley of Oaxaca around 400 B.C. (Joyce, 2010). Until its abandonment around A.D. 800 (Markens, 2008; Blanton, 1978), Monte Albán was a demographic, economic, and political hub that would have been an important destination of human movement for nearly a millennium. An abundance of materials from distant parts of Mexico and beyond were brought to the city and neighboring sites in the Valley of Oaxaca (e.g., Joyce, 2010:155–157; Winter, 1998), including marine and estuarine shell from the Atlantic and Pacific coasts.

Fig. 1. Connectivity options for point-to-point travel. From left to right: (a) one-to-one, (b) one-to-many, (c) many-to-one, (d) many-to-many.
The coastal sources of shell found at highland sites remain poorly understood in large part due to a lack of archaeological research on much of Oaxaca’s Pacific coast (but see Joyce, 2008; Zeitlin, 1993), the region from which most of Monte Albán’s shell was likely obtained (Feinman and Nicholas, 1993:114). Identifying the coastal sources of shell in highland Oaxaca is important for resolving broader debates on long-distance political and economic relationships in western Mesoamerica during the Late Formative (400 B.C. – A.D. 250) and Classic (A.D. 250–800) periods. Several scholars have proposed that access to valued coastal resources like shell was one impetus for the development and expansion of the Monte Albán polity (e.g., Marcus and Flannery, 1996; Sherman et al., 2010; Spencer and Redmond, 2004). Furthermore, Monte Albán may have been an important node in the exchange of Pacific shell throughout the Mexican highlands, including the Central Mexican city of Teotihuacan (Feinman and Nicholas, 1993:114).

As indicated by the recent work of Sherman et al. (2010), GIS modeling could be a valuable predictive tool to identify coastal shell sources and clarify long-distance economic and political relationships by focusing research on areas of Oaxaca’s 568 km-long Pacific coastline that were most easily accessed from Monte Albán by foot. Because standard GIS modeling can only generate one-to-one path relationships, however, specific coastal destinations are required. A lack of data on much of the coast thus substantially limits both the exploratory and explanatory power of any model that is generated. These limits are evident in GIS-based movement studies for the region (Sherman et al., 2010; Barber et al., 2011). Sherman et al. (2010) used least cost paths between Monte Albán and the Pacific coast to explain settlement patterns in several highland areas. Least cost paths were generated between Monte Albán, one known coastal site, and three arbitrary coastal points set 50 km apart (see also Carballo and Pluckhahn, 2007). While a useful first effort, the paths generated only provide information on movement between Monte Albán and a 150 km-long stretch (26%) of Oaxaca’s coast. The large sampling interval, furthermore, makes it impossible to consider other possible paths within the studied area. Using a similar approach, we expanded the geographic area to include the lengths of both the Pacific and Atlantic coasts nearest Oaxaca and narrowed the sampling interval to 30 m (Fig. 3). Our model includes 170,000 travel-time-based least cost paths between Monte Albán and the ocean. This was achieved by using Tobler’s (1993) hiking function to create an accumulated cost surface based on slope, derived from a 30 m DEM, and an automation script that leveraged both ArcGIS10 and the Interactive Data Language (IDL) to generate paths.

The paths generated by our “Monte Albán To Everywhere” (MATE) model use topography to quantify differences in the cost of movement, in travel time, between the city and the Pacific littoral (Table 1). Purely in terms of travel time, which Hassig (1985) suggests was probably the most important factor in pedestrian transportation route choice, the least costly coastal area to access from Monte Albán was thus the west-central coast via either the Sola or Ejutla valleys. This result is supported by evidence for shell-working in the highland Ejutla valley, which lies along one of the shortest routes (Feinman and Nicholas, 1993). Interestingly, the least-costly route does not appear to have been the most heavily-trafficked (Sherman et al., 2010:284) early in the city’s history, emphasizing that factors other than slope must be considered when modeling human movement across the landscape (Bellavia, 2006, 2002; White and Surface-Evans, 2012; Howey, 2007; Llobera et al., 2011; White, 2007, 2012; Pingel, 2009).

Although the MATE results are useful for understanding possible movement patterns in and out of Monte Albán, they also highlight some of the previously discussed problems with GIS-based movement modeling. First, the origin and one-to-one path requirements result in a solar pattern of paths radiating out from Monte Albán. While such a pattern may be appropriate near the city itself (following Llobera et al., 2011), at significant distances, such as in coastal regions, it seems improbable that movement would have been oriented towards a location more than a hundred kilometers away through mountainous terrain. Further, there is no reason to assume that marine shell was moved from the ocean directly to Monte Albán with no intervening transfer points (e.g., Feinman and Nicholas, 1993). A more effective model of movement in pre-columbian Oaxaca would encompass Monte Albán, sites in

![Fig. 2. The states of Mexico with Oaxaca, the study area for this project, highlighted.](image-url)
intervening areas, and those on the coast. Even if it were possible to model such a many-to-many path scenario with standard methods, sufficient site data are unavailable because many of the relevant areas have never seen systematic archaeological survey. With significant gaps in current data on precolumbian settlement, we needed a way to generate hypothetical networks without having to predetermine all possible origins and destinations.

3. The from everywhere to everywhere approach

Our solution, FETE analysis, begins by placing a regularly-spaced sampling grid on top of a selected area of terrain (Fig. 4a). Dijkstra's (1959) shortest path algorithm is then run from each sample point to every other sample point in the grid in a manner somewhat akin to the Floyd-Warshall algorithm for weighted graphs (Warshall, 1962). The algorithm produces a travel cost surface and a backlink map, which is used to calculate least cost paths from every other point on the sampling grid back to the current origin. The locations of the paths are then recorded in a shared accumulative surface. Thus if 50 paths calculated by 50 different instances of the Dijkstra algorithm all travel through the same point on the surface, that point is represented by a value of 50. The end result of the FETE process is a travel probability surface that, with some processing, resembles a dense circulatory system or road network (Fig. 4b). The FETE approach expands on the grid technique employed by Whitley and Hicks (2003), who used regularly-spaced points around the perimeter of their study area to generate a set of paths oriented north–south and east–west. FETE differs significantly from their approach in that it includes a large number of points within the study area and calculates the likelihood of travel on any segment of a network. While the extensive sampling of FETE analysis is very expensive to execute computationally, it captures complex patterns of intra-regional movement rather than just traces of people passing through. It should also be noted that while the accessibility index developed by Llobera (2000) is created using multiple shortest path calculations, the methodology behind it, and its purpose, are very different from that of FETE. The accessibility index is designed to provide a single metric that communicates how easy it is to get to a particular location from all points around it that fall within a user-specified radius. It is an extension of the standard one-to-many approach in that it moves a “Dijkstra kernel” across the supplied raster to produce a local accessibility image. That involves averaging the costs for paths from all points within the radius back to the current origin, which means that they have to be added together, but that is where the similarity to FETE ends. FETE does not average local path cost and then convey how easy it might be to traverse a landscape; it accumulates global path presence and then conveys the likelihood that any given location will be traversed.

<table>
<thead>
<tr>
<th>Archaeological site, modern town, or geographic feature</th>
<th>Travel time from Monte Albán in MATE model (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Southwest Trunk Route</strong></td>
<td></td>
</tr>
<tr>
<td>Mouth of Río Santa Catarina, Guerrero</td>
<td>54.29</td>
</tr>
<tr>
<td>Border Oaxaca/Guerrero</td>
<td>50.55</td>
</tr>
<tr>
<td>Mouth of Corralero Lagoon</td>
<td>44.55</td>
</tr>
<tr>
<td><strong>South Trunk Route</strong></td>
<td></td>
</tr>
<tr>
<td>Collantes</td>
<td>39.69</td>
</tr>
<tr>
<td>Río Viejo</td>
<td>35.30</td>
</tr>
<tr>
<td>Río Grande</td>
<td>33.46</td>
</tr>
<tr>
<td>Manialtepec</td>
<td>30.78</td>
</tr>
<tr>
<td>Bajos de Chila</td>
<td>31.82</td>
</tr>
<tr>
<td>Mouth of Río Cozoaltepec</td>
<td>34.87</td>
</tr>
<tr>
<td>Puerto Angel</td>
<td>36.93</td>
</tr>
<tr>
<td>Huatulco Port (Crucecita)</td>
<td>36.00</td>
</tr>
<tr>
<td>Mouth of Río Zimatán</td>
<td>39.36</td>
</tr>
<tr>
<td><strong>Southeast Trunk Route</strong></td>
<td></td>
</tr>
<tr>
<td>Santiago Astatana</td>
<td>39.70</td>
</tr>
<tr>
<td>Mouth of Río Tenango</td>
<td>41.06</td>
</tr>
<tr>
<td>Salina Cruz</td>
<td>43.61</td>
</tr>
<tr>
<td>El Carrizal</td>
<td>43.82</td>
</tr>
<tr>
<td>Laguna Zope</td>
<td>46.42</td>
</tr>
</tbody>
</table>
Because the statistical distribution of paths on the FETE surface appears to follow that of a scale-free power law (Clauset et al., 2009), it can be thresholded to locate high traffic areas using only the portion of the distribution that corresponds to the well-established 80/20 Pareto Principle (Fig. 5) (Newman, 2005). The thresholded values are then color-coded using a green—yellow—red linear ramp between the minimum and maximum values, with extremely high traffic areas receiving a red color, very high traffic areas receiving a yellow color, and high traffic areas receiving a green color (Fig. 6). The colorized results are then placed on top of another reference layer (e.g., the digital elevation model) and used to guide further research. The color-coding approach has its roots in the transportation GIS research conducted by Branting (2004) at Kerkenes Dag, in which travelers are constrained to a known street grid while they move between a small number of entry/exit points to the city to build up frequency-based traffic patterns. Unlike that study, we are working with a landscape that is only constrained by topography/land cover and are using frequency distributions to find previously unidentified movement corridors.

There are two additional uses for the FETE surface that should be highlighted. It is possible to represent the FETE output in a raster form such that the frequencies tied to every traversed location, even the ones that are only touched once, are used to extrapolate more general movement corridors throughout the region. This is done by converting each location to a point that has its frequency as an attribute and then using a natural neighbor interpolation scheme (Sibson, 1981) to create a raster surface based on those points (Fig. 7). If specific origins and destinations for travel are known, FETE output can also be used as a traditional cost surface to generate “most probable” paths between the locations. This is done by subtracting each cell’s frequency from the highest one recorded to produce a new cost surface equating high frequency to low cost, which could then be used with any standard GIS software package to create least cost paths.

**Fig. 4.** An example of the regularly-spaced sampling grid used in the FETE approach (left) and sample output from the algorithm (right). Spacing between grid points is exaggerated for presentation purposes. Actual spacing used is ten pixels.

**Fig. 5.** Histogram of a FETE algorithm output raster, highlighting the frequencies used to delineate high-traffic travel regions in the dataset.

**Fig. 6.** Thresholded and colorized version of the FETE algorithm output. Moderately high traffic areas are green, very high traffic areas are yellow, and extremely high traffic areas are red. All locations represent likely candidates for prehistoric travel corridors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
3.1. Datasets used

The FETE approach is based on the research White (2007, 2012) conducted in the Western Papaguería region of the North American Southwest and requires two co-registered raster sources as input: an elevation model and a land cover model. The former provides a three-dimensional characterization of the terrain upon which people can travel and the latter provides information about the types of natural and man-made materials sitting on that terrain and how they might influence travel. The elevation model used for this study consisted of a subset of the 30 m ASTER Global DEM (GDEM, EOSDIS, 2009). ASTER elevation data were used instead of SRTM, despite known accuracy issues, because the latter are less consistent for the study area. Specifically, available SRTM data contain many voids where the active sensing system that created it (synthetic aperture radar) was unable to obtain sufficient signal due to mountainous terrain. The land cover model consisted of a subset of the 500 m MODIS Yearly Land Cover product (MCD12Q1, Strahler et al., 1999), which is based on a full year of observations from both the Terra and Aqua satellites (Fig. 8). The product contains multiple land cover classification schemes, with the primary one identifying seventeen distinct classes (Table 2). These classes were then mapped into the limited number of available terrain coefficients (Soule and Goldman, 1972) used by the metabolic rate of travel functions described below, which suggest that paved (urban) or barren terrain is easier to move across than forests or wetlands. In a more fine-grained sense, urban settings may be more difficult landscapes to

Fig. 7. Natural neighbor interpolation output for FETE, centered on Monte Albán and showing other large sites in the region. All sites fall within high traffic areas identified by the algorithm.

Fig. 8. Sample co-registered ASTER GDEM (left) and MODIS MCD12Q1 (right) raster products. For the elevation product, each pixel contains an elevation value recorded as meters above mean sea level (MSL). For the land cover product, each pixel contains a numeric value that corresponds to a broad class of materials (e.g., forest).
navigate due to the complexities of built environments, but the terrain coefficients only speak to the navigable qualities of a particular surface.

This primary scheme, developed by the International Geosphere Biosphere Programme (IGBP) (Belward and Loveland, 1995), was used in this study due to its breadth and consistency and despite the contemporaneity and coarse spatial resolution there is no other uniform land cover dataset available, past or present, for the entire study area. Currently available paleo-environmental data do indicate changes in river courses and characteristics, the size and location of coastal lagoons, and the extent of human-derived land clearance over the past 3000 years (Goman et al., 2005, 2010; Pérez Rodríguez et al., 2011). These studies, however, deal with areas far smaller than the current study area and are insufficient to generate historic land cover maps. While by no means a perfect solution, our use of modern land cover data is warranted due to the fact that topography limits agricultural development even today and because there has been considerable continuity in settlement locations since at least the Late Postclassic period (A.D. 1200–1520) (Zeitlin, 1978a; Byland and Pohl, 1994). Since the land cover data are mainly used to establish basic impedance metrics for travel, we felt that using modern data was an acceptable compromise in order to gain access to consistently measured and evaluated information for an otherwise data-poor region. Using a dataset with higher spatial resolution and direct archaeological significance (e.g., differentially weighted locations based on resource availability/desirability) would likely produce better modeling results, which is planned as part of a future extension to the existing project but is not possible at this time.

Both the elevation and land cover datasets consisted of many tiles that had to be mosaicked together in their native projections, reprojected to a meter-based system that could span such a large region (Albers Equal Area), resampled to a manageable spatial resolution (50 m), and co-registered before they could be used in the FETE algorithm.

### 3.2. Graph creation

The FETE algorithm takes advantage of Dijkstra’s (1959) well-established graph searching algorithm, which calculates the minimum cost required to travel between a given graph node (origin) and every other node on the graph. An extremely useful byproduct of this process is the ability to reconstruct the least costly path from any node back to the origin by way of a node connectivity list referred to as a backlink. For the purposes of this study, the co-registered elevation and land cover rasters are treated as one dataset consisting of two layers that represent a single graph. The center of each co-registered raster cell is treated as a node on the graph, with connections between adjacent nodes treated as edges that represent the cost of travel between the nodes. Travel between nodes is viewed as an anisotropic function, so each edge stores two costs: travel from node A to node B and travel from node B to node A. Each node is connected to its eight closest neighbors for maximum travel flexibility, generally referred to as Queen’s Case movement. The resulting connectivity relationships are stored in a master list for use by the FETE algorithm. Edge costs are generated using the co-registered digital elevation data and land cover data, along with traveler-specific information (body weight, load weight, age, height, and sex), as inputs to a set of velocity estimation and caloric expenditure estimation formulas. The resulting anisotropic costs are also stored in a master list for use by the FETE algorithm. All edges that cross water bodies are flagged as being off limits to travel. Determining whether or not edges cross water involves running a custom algorithm on Landsat ETM+ data and the regional DEM—one that takes advantage of spectral, topographic, and spatial patterns that are beyond what can be discussed here. Two of the water bodies in the sample are large modern lakes created by 20th century dam projects. We left them in place because data on the relevant ancient river courses are not available.

#### 3.2.1. Edge cost generation

Given graph node A, surrounded by nodes B–I, and the desire to travel between nodes A and D (Fig. 9), the process outlined below is used to calculate bidirectional edge costs in the form of kilocalories expended. First, elevation values and land cover values for each node are retrieved from the DEM. The horizontal \( d_h \) and vertical \( d_v \) distances traveled, in meters, are determined and used to calculate the actual three-dimensional distance traveled \( d_t \), fractional slope \( s_f \), and percentage slope \( s_p \):

\[
d_t = \sqrt{d_h^2 + d_v^2}
\]

\[
s_f = \frac{d_c}{d_h}
\]

\[
s_p = \left( \tan^{-1}(s_f) \right) \frac{2}{\pi} \times 100,
\]

where \( s_f \) is the fraction of a quadrant-averaged arctangent function (e.g., atan2).

The percentage slope calculation described above, also referred to as grade, returns values between −100.00% and 100.00%, where the bounds represent −90° and 90°, respectively. Because arctangent functions, as implemented in standard programming languages, return angular values in radians, additional conversion is necessary to obtain percentages.

Second, travel velocity along the edge \( v_e \), in meters per second, is calculated by supplying the fractional slope value to Tobler’s (1993) hiking function and multiplying the result \( \text{km/h} \) by a conversion factor:

\[
v_e = \left( 6 \cdot e^{-3.5 \cdot |s_f + 0.05|} \right) \cdot \left( \frac{1000}{3600} \right)
\]

Supplying angular (degrees between 0 and 90) or percentage (values between 0 and 100) slope to Tobler’s function is a common mistake, as those are the only types of slope that can be easily generated in raster form by commercial and open source GIS software packages.
Third, the metabolic rate of travel (MRT) along the edge, in watts, is calculated using travel velocity, weight in kilograms (w), load weight in kilograms (l), percentage slope, and landcover type ($\eta$). For level and positive (uphill) slopes, the standard Pandolf equation is used (Pandolf et al., 1977; Soule and Goldman, 1972):

$$\text{MRT} = 1.5 \cdot w + 2 \left( \frac{l}{w} \right)^2 + \eta \cdot (w + l) \cdot \left( 1.5 \cdot v_e^2 + 0.35 \cdot v_e \cdot sp \right)$$

For negative (downhill) slopes, a modified version of the Pandolf equation is used, based on the work of Santee (Santee et al., 2001; Wood and Wood, 2006):

$$\text{MRT} = 1.5 \cdot w + 2 \left( \frac{l}{w} \right)^2 + \eta \cdot (w + l) \cdot \left( 1.5 \cdot v_e^2 + 0.35 \cdot v_e \cdot sp \right) - \frac{\eta \cdot \left( v_e \cdot sp \cdot (w + l) - (w + l) \cdot (sp + 6)^2 \right)}{3.5 \cdot w} + \left( 25 - v_e^2 \right)$$

Since the traveled edge spans two raster cells with potentially different types of land cover, the MRT is calculated for both land cover values ($\eta_1$, $\eta_2$).

In the fourth step, MRTs are combined, changed from watts (Joules per second) to kilocalories per second via multiplication by 0.000239 kcal/s, and used in combination with the travel velocity and edge distance (the length is split between both calculated MRTs) to calculate the number of kilocalories expended during travel:

$$\text{kcal}_t = \frac{0.000239 \cdot (\text{MRT}_{\eta_1} + \text{MRT}_{\eta_2}) \cdot (0.5 \cdot d_t)}{v_e}$$

Fifth, using age in years ($a$), weight in kilograms ($w$), height in centimeters ($h$), and sex, standing metabolic rate in watts is calculated via the Harris-Benedict Basal Metabolic Rate equations and a scale factor (Wood and Wood, 2006):

$$\text{SMR}_{\text{male}} = (663 + (13.7 \cdot w) + (5 \cdot h) - (6.8 \cdot a)) \cdot 1.2$$

$$\text{SMR}_{\text{female}} = (655 + (9.6 \cdot w) + (1.8 \cdot h) - (4.7 \cdot a)) \cdot 1.2$$

The standing metabolic rate is then used to calculate an alternative estimate of caloric expenditure that can be substituted for kcal$_t$ at slow speeds on downhill slopes, which sometimes underpredicts actual expenditure (Wood and Wood, 2006):

$$\text{kcal}_a = \frac{0.000239 \cdot \text{SMR} \cdot d_t}{v_e}$$

### 3.3. Implementation of the FETE model

The FETE model involves a complex graph building approach and is so computationally intensive that it cannot be efficiently or correctly executed inside any currently available commercial or open source desktop GIS software package. It does not help that as the size of the terrain being investigated increases, the computational challenge also increases, but at a much faster rate. Tsui and Wood (2009) encountered this issue when they developed a way to run Dijkstra’s algorithm on an elevation model for the entire United States in an attempt to find historic trails. They ended up having to employ both search constraint mechanisms and a dual-threaded algorithm in order to solve for a single path. The approach used here sidesteps the need for search constraints because it takes advantage of an extremely efficient custom version of Dijkstra’s algorithm that was written in C++ and can solve for multiple paths simultaneously, in parallel, by taking advantage of today’s multicore desktop computers and servers.

The experiment described below was conducted with a 5000 x 4000 pixel terrain dataset using a 100-pixel sampling grid with Threading Building Blocks (TBB, Intel Corporation, 2012) running on a twelve-core Intel Core i7 computer with 24 GB of RAM. The TBB process monopolized all twelve cores and consumed close to 50% of the RAM for 2.5 h, producing 2000 cost surfaces and 4,000,000 least cost paths. The FETE software was developed to be scalable so that, if necessary, it can run on hundreds or even tens of thousands of cores (i.e., a high-performance computing cluster). As the size of the region of interest increases beyond what can be examined by using a desktop computer, there are still viable options that will not require a significant reworking of the underlying code.

It is also important to highlight the fact that FETE can operate in five different modes that are designed to explore multiple movement scenarios. The most basic mode connects all points on a regularly-spaced grid to each other, as described above and used below. The second mode connects all points in a user-supplied set to each other. The third mode connects a user-supplied set of origin points to a user-supplied set of destination points (and vice versa). The fourth mode connects user-supplied origin and destination points to a regularly-spaced grid of waypoints that spans the distance between them.
fifth mode substitutes a user-supplied set of waypoints for the regularly-spaced ones.

4. A proposed network for ancient Oaxaca

4.1. Model parameters

Three experiments were conducted to test the efficacy of the FETE approach to identifying movement corridors within the Oaxaca study region, mainly to confirm that the high traffic paths being generated appeared consistently even when the parameters of the algorithm were altered significantly. The first used a baseline set of parameters derived from White’s (2007, 2012) work in the Western Papaguería portion of the North American Southwest. This experiment used a fairly dense grid spacing of 100 pixels; a 25-year-old, 1.7 m tall male traveler weighing 75.0 kilos; and a minimum additional load of 7.0 kilos. Once the baseline was established, two additional experiments were run to see how much the generated networks differed: (1) varying the sampling grid from 100 to 500 pixels while holding everything else constant and (2) varying the additional load based on values documented in a recent study of Nepalese porters (Malville, 2001) while holding everything else constant (Table 3).

We found that the most frequently traveled routes changed little when the sampling interval changed or the weight of the load was adjusted. For instance, in the Valley of Oaxaca, the main northwest—southeast routes persisted even when the weight of the porter’s load was adjusted from 23 kilos, to 71.6 kilos, and then to 111.0 kilos. What did change was the level of traffic on particular routes. We found that (1) as sampling decreased, fewer routes were generated, but the same high traffic ones were consistently identified; and (2) as load increased, the number of high-traffic routes decreased significantly while the number of lower-trafficked but very direct routes increased. Models run with the maximum load also generated a number of unnaturally direct routes, particularly over level terrain but also in mountain zones (Fig. 10). The increase in direct routes mirrors Hassig’s (1985:32) expectation that pedestrian routes will be direct even if the such paths entail surmounting a hill rather than going around it. In all cases, the mountains served to funnel movement into a limited number of corridors.

4.2. FETE in Oaxaca

Oaxaca’s extensive ethnohistoric record provides independent corroboration that the movement corridors generated with the FETE approach likely indicate the approximate location of routes used for precolumbian movement (Fig. 11). The most important historically documented routes passed through the Valley of Oaxaca to the Tehuacan Valley to the north and the southern Isthmus of Tehuantepec to the southeast (Redmond, 1983; Zeitlin, 1978b). Our results successfully identified both routes as very high traffic. An ethnographically-documented route between Mitla and the southern Isthmus was also identified as a high-traffic route (Parsons, 1936:Map III). A second route into the Valley of Oaxaca passed through the sierra town of Sosola, the location of an Aztec garrison (Ball and Brockington, 1978:111). Our results identify this high-traffic corridor and demonstrate how it may have connected with important trade centers farther north like Tamazulapan (see below). The model postulates a high-traffic route along the Rio Santo Domingo, a tributary of the Papaloapan that connects the Cuicatlán Cañada with the Gulf Coast that has archaeological and ethnohistoric documentation (Ball and Brockington, 1978; Rees, 1975). The FETE output also shows a very high traffic path between the Valley of Oaxaca and Tuxtepec on the Gulf Coast that bypasses the Santo Domingo. Based on colonial documents, Rees (1975:Fig. 1; see also Gutierrez Mendoza and Van Rossum, 2006) has proposed that a similar route was used by the Aztec. Finally, the model identifies a series of high traffic routes extending northwest-southeast between Ejutla and the coast near Huatulco. Borah (1954:27) sketched a similar route using early Colonial records. We argue that the correspondence between our results and documented transportation routes supports the use of the FETE approach to identify movement corridors even in cases where written documentation is not available.

Indeed, the results of our analysis suggest that there were several high-traffic routes through Oaxaca that are not as widely discussed in the archaeological literature (see Fig. 11). Ball and Brockington (Ball and Brockington, 1978:110) mention Tamazulapan and Coixtlahuaca in the Mixteca Alta region as part of a Late Postclassic period (A.D. 1200–1521) Aztec trade route between Tenochtitlan and the Valley of Oaxaca. Borah (1954) discusses a more southerly route through Yanhuitlán that our model indicates may have extended south to Sosola and north to Tamazulapan. The FETE results suggest that both were very high traffic routes and that the Sosola to Tamazulapan route had equivalent traffic to the more widely discussed route through the Cuicatlán Cañada. The model also predicts a high-traffic west-corridor that passed from the southern Valley of Oaxaca into the Sola Valley with less heavily-trafficked corridors extending south to the coast. These paths in aggregate are similar to one generated by Sherman et al. (2010:Fig. 13) using standard GIS methods. Because the FETE model renders networks, the results highlight the multi-directional nature of movement through this region, particularly the likelihood of high traffic heading west out of the Sola Valley rather than just south to the coast. Finally, the model predicts a high-traffic route along the coast for which there is little documentation but that archaeologists have generally assumed existed (e.g., Ball and Brockington, 1978).

In terms of understanding the long-distance exchange of Pacific coastal resources, especially shell, in precolumbian Oaxaca, the FETE results provide further support to existing models and suggest new possibilities. The FETE results support current archaeological evidence that the highest traffic between Monte Albán and the Pacific would have been through the Sola and Ejutla valleys to the west and central coasts. Because the FETE model does not focus specifically on Monte Albán, however, the results also indicate several possible high-traffic routes between the coast and the Mixtec highlands to the west of the Valley of Oaxaca. These latter routes could have been used in long-distance exchange between the Valley of Mexico to the northwest and the Pacific coast of Oaxaca while bypassing Monte Albán. The number of high-traffic routes traversing the Mixtec highlands more generally supports archaeological observations that this region of

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Details of the eight FETE runs used for this study. Sampling grid size and load carriage values were varied and compared against an established baseline.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FETE run</strong></td>
<td><strong>Grid spacing</strong></td>
</tr>
<tr>
<td>Baseline</td>
<td>100</td>
</tr>
<tr>
<td>Increase Grid Spacing × 2</td>
<td>200</td>
</tr>
<tr>
<td>Increase Grid Spacing × 3</td>
<td>300</td>
</tr>
<tr>
<td>Increase Grid Spacing × 4</td>
<td>400</td>
</tr>
<tr>
<td>Increase Grid Spacing × 5</td>
<td>500</td>
</tr>
<tr>
<td>Minimum Load</td>
<td>100</td>
</tr>
<tr>
<td>Average Load</td>
<td>100</td>
</tr>
<tr>
<td>Maximum Load</td>
<td>100</td>
</tr>
</tbody>
</table>
Oaxaca had diverse long-distance connections throughout the precolumbian era (e.g., Blomster and Glascock, 2011). Finally, the FETE model predicts a series of east–west trending routes south of Monte Albán, suggesting that an extensive trade in coastal goods may have occurred in many areas of Oaxaca without reference to Monte Albán.

Given the parity between modeled and documented movement corridors, future research can employ FETE results both as a predictive tool and as a means of analyzing relationships between fixed points and probable loci of movement. For instance, since the output of the model is based entirely on ease of movement without reference to specific settlements, it can be used to generate hypotheses regarding whether a site developed or flourished as a result of proximity to a pre-existing trade route or junction while controlling for the possibility that the trade routes developed as a result of the presence of the site. There are obviously a wide range of factors that influence the growth of ancient towns and cities, however a number of Oaxaca's most important precolumbian towns and cities were located within 5 km of high-traffic path intersections generated by the model (Fig. 12; see Surface-Evans,

**Fig. 10.** Baseline FETE output near Monte Albán (top) compared to output when maximum load carriage is used (bottom).
The most obvious example is Monte Albán, which sat atop the junction of the most heavily-trafficked routes in the state.

The model we have presented here, and the computational technique we used to generate it, combines Western assumptions of rational economic action with principles of human physiognomy and modern physical geography. As many scholars have observed, human movement was influenced by many other factors, such as political conditions, ritual obligations, and historical relationships. Indeed, the Oaxacan ethnohistoric record demonstrates how movement, particularly walking, embodied specific histories, social values, and spatial perceptions (e.g., Zeitlin, 2005:14). We thus offer our results not as a finished map of ancient pathways but rather as a tool for considering how people in the past navigated the landscape.

**Fig. 11.** Documented precolombian travel corridors in Oaxaca overlaid on interpolated FETE results, showing good agreement with hypothesized high traffic routes.

**Fig. 12.** Proximity of important precolombian towns and cities in Oaxaca to intersections of FETE-generated high traffic routes. Each circle has a 5 km radius.
Acknowledgements

We would like to thank John Walker for helpful comments on this paper and the audiences at the 2011 SAA meetings and VAacmulate 2011 for asking insightful questions after the presentation of our initial results. Jack Gibson provided valuable assistance in generating GIS datasets for this study and in copy-editing an earlier draft.

References


