Network Models and Archaeological Spaces

Ray Rivers, Department of Physics, Imperial College London, UK (Email: R.Rivers@imperial.ac.uk)
Carl Knappett, Department of Art, University of Toronto, Canada (Email: carl.knappett@utoronto.ca)
Tim Evans, Department of Physics, Imperial College London, UK (Email: T.Evans@imperial.ac.uk)

Contribution to “Computational Approaches to Archaeological Spaces” eds A. Bevan, M. Lake

Contents (for information only – not for journal)

I. Introduction

II. The Role of Geographic Space: First Thoughts
   • A question of ‘distance’: Fixed radius model (Ariadne VP)
   • Proximal Point Analysis (Ariadne PPA)
   • Agency: Optimisation

III. The ‘Most Likely’ Networks
    • Maximum entropy: Basic gravity models
    • Rihll and Wilson model (Ariadne RW)

IV. The ‘Most Efficient’ Networks
    • Utility functions or social potentials (Ariadne MC)
    • Organisational hierarchy
    • Is almost the best good enough?

V. How Do We Know If It Works?

VI. Summary

References

Figures

Table


I. Introduction

A fundamental ingredient of archaeological analysis is an understanding of spatial relationships. We concur with Renfrew (Renfrew 1981) that

“the very activity of examining the spatial correlates of early social structure has the useful effect of posing important problems in general and simple terms”.

Posed as generally as this, it is not altogether surprising that different authors have provided almost as many answers as there are questions and to shape the discussion which follows we have chosen to put a heavy emphasis upon the Bronze Age Aegean, the period for which our own models (Knappett et al. 2008; Evans et al. 2009) of maritime trade and exchange routes were devised. This is because the act of constructing and developing them has illuminated problems common to many approaches.

The original focus of our work was the southern Aegean in the Middle and early Late Bronze Age. This is well bounded in time, concluding with the ‘burning of the palaces’ some time after the eruption of Thera. The archaeological record poses some interesting questions. On the one hand, we wanted to understand why some sites, like Knossos on Crete, grew to be so large and influential, despite their seemingly unpromising positions (Knossos is some kilometres inland). On the other hand, we wanted to understand how the volcanic eruption of Thera seemed to have such little immediate effect on the exchange between Minoan Crete and the surrounding regions despite destroying Akrotiri, which might have been anticipated as the main gateway from Knossos into the Cyclades, Peloponnese and the Greek mainland.

Although both address spatial issues, superficially these questions seem very different. Thus, the emergence and characteristics of sites like Knossos are partially explained in local terms, with these local conditions only then facilitating exchange with other sites. The global stability of the Minoan exchange network, meanwhile, might be thought of as explicable in terms of Aegean geography and marine technology, since the appearance of the sail from about 2000 BC facilitated new levels of inter-regional interaction. However, to separate local power from global organisation too strongly is to set up a false dichotomy, since the centres of local power are also likely to be key players in global exchange. The same routes that enable these centres to maintain their influence are also conduits for the entwined cultural and economic connections between them. As a result, their growth and importance will be conditioned by the characteristics of the larger networks to which they belong, leading to what Batty terms an ‘archaeology of relations’ (Batty 2005).

In our Aegean work we have argued that network models, intrinsically encoding ‘spatial correlates’, can provide the natural framework to transcend this dualism between physical and relational space. Networks are, in essence, no more than a set of nodes or vertices connected by links or edges. Most simply, in archaeology we identify nodes with archaeological sites and the communities that inhabited these sites, and links with the exchange between them. Networks have the potential, in an almost self-evident way, of incorporating both the local attributes of sites together with their global interactions and show the reciprocal effect of the one on the other.

In this article we are primarily interested in how networks of relationships are conditioned by geographic space. The plan is as follows. In the next section we give an overview of some of the issues that arise in constructing network models for archaeology, as exemplified by overtly geographical models. This prefigures the most substantive part of the Chapter, an analysis of how ‘optimisation’ models, of the type used in urban planning and contemporary socioeconomics, might be useful in Aegean prehistory. Although such modelling has obvious limitations, in its stress on the material circumstances of human-environment interaction on social change, it allows for a ‘restricted set of spatial variables to play a role in shaping historical transformations’ (Smith, 2003). As a result, we can construct tractable quantitative
models amenable to simulation. We conclude with a comparison of these different approaches to handling space.

II. The Role of Geographic Space: First Thoughts

One attribute common to most network models is that they are construed dynamically. Links describe the flows between sites; goods, raw materials, people, ideas, both the sinews and synapses of intersite relationships. In the most simple networks, often used for null models, links are non-directional and either switched on or off, i.e. all we are concerned with is whether two sites interact or not. At a more sophisticated level, links are both directional (i.e. the relationship of A to B is not the same as B to A) and are ‘weighted’ i.e. vary from strong to weak, where the weight is a measure of the ‘flow’.

In an archaeological context nodes, thought of as sites, have two different attributes. At one level we can interpret them as junctions for flows, without needing any of their local physical and geographical properties beyond, say, position, which determines their accessibility to other sites. In this sense they are rather like traffic intersections, or airports, only acquiring their local significance from the network as a whole. This permits a variety of classifications, of which the simplest is akin to PageRank, ordering nodes by their ‘busyness’, the total flow through them. However, we should not ignore their local properties. In fact, our original interest in networks was triggered by the use of networks to explain the mismatch between local resources and exchange in the Early Bronze Age Cyclades (Broodbank 2000). In some models (our first null models) sites are treated equally but, more realistically, are ranked by their local properties (e.g. populations) as well as their properties derived from the global behaviour of the network as a whole.

As we said earlier, our main interest is to see to what extent exchange networks reflect the geographical space in which they are embedded or transcend it. Most simply, by looking at the map, could we guess the nature of the network? This is not just a question of joining the dots. Can we infer which sites are most important, both from their intrinsic properties (e.g. resources) and their importance in the network? We also need to identify the most important links since they will reflect the distribution of artefacts. We begin this discussion by presenting two null models, each of which uses nothing but physical geography, against which to test the more sophisticated hypotheses that we shall propose later.

A question of ‘distance’: Fixed radius model

The map that we choose to study is of the S. Aegean as given in Fig.1

Insert Fig.1

On it we have labelled 39 of the most significant sites in the Middle Bronze Age. Names and details of the sites are given in the Appendix in Table I. In general we do not need them, since we are mainly using the resulting networks for comparative purposes, and will not address the record in any detail. In the first instance, what we see is four regional groupings; Crete, the Cyclades, the Peloponnese and the Dodecanese. One of the tests for the importance of geography lies in the extent to which the exchange networks reflect these regional groupings.

The next step is to decide what is meant by the distance $d_{ij}$ between sites $i$ and $j$ (where $i$ and $j$ take values from 1 to 39). For the period in question, the MBA, communication is by sailing ships which do not remain at sea at night. Our distances are not line-of-sight distances but correspond to the best sea routes, negotiating headlands where necessary. Even that is not sufficient, since land travel may produce routes which are shorter and some sites are inland, notably Knossos and Mycenae. Given the slow and
laborious nature of land travel in this era, we introduce a ‘frictional coefficient’ of three to one for land travel over sea travel. This makes many coastal island sites, as on N. Crete, behave rather like islands themselves given the mountainous terrain. Our distance estimates were obtained by analysing the physical geography by hand at a scale of around a few km. We expect our estimates may be wrong by up to 5%. However, we argue that this is an appropriate level of error given the other uncertainties in our knowledge of the physical and social geography of the time. Our effective distances thus reflect actual time-of-travel and we use the same estimates in all the models discussed here. We could do more and make distance (or time-of-travel) directional to take into account the prevailing winds and currents. However we will leave these for later studies.

Once we have encoded the effect of geography and technology in our effective distances, the simplest way to impose geography on the network is to connect sites only if they are less than a certain distance $D$ (e.g. 100km) apart, which may be encoded as a dendrogram such as the one shown in Fig.2. A horizontal line at some fixed distance scale $D$ cuts the tree into a number of smaller branches lying below the line, each defining a separate cluster of sites. Only sites within the same cluster are connected to each other by paths in which each section is less than the distance $D$ apart.

**Insert Fig.2**

Of course, this approach is too precise. If we set $D=100km$ then two sites which are 99km apart are connected, while a pair 101km apart are not. To take account of the uncertainties in distance qualitatively, it is better to imagine drawing a band of width 10km rather than a line, such that the branches defined by the top and bottom of the bands are almost identical. That is we look for regions in the dendrogram with few horizontal lines. Looking at our dendrogram of Fig.2, we see that distances between 110km and 130km (the effective distance between Knossos and Akrotiri is about 130km, once land travel is taken into account) are the transitional distances. At 110km the sites are split into the four isolated zones we would expect; Crete, the Cyclades, the Dodecanese and the Peloponnesse (itself fractured into adjacent parts). On the other hand, almost all the sites are connected by routes involving steps of 130km or less. This structure becomes very clear if we show the network for different distance scales, as in Fig.3. The dark links correspond to distances up to the specified distance.

**Insert Fig.3**

Simple as this model is, one thing we take from it is the very striking fact that the distance scale $D=120km$ at which we move from ‘under’ connected to ‘over’ connected networks is more or less the maximum distance scale we would expect from MBA sailing vessels for a single journey. The arrival of MBA marine technology makes the Aegean a ‘Goldilocks’ sea in the sense that it is just right for the appearance of an active trading network with a handful of key sites. Too small a journey distance and the network is overwhelmed by the prevailing geography, as is the case for the EBA, for which canoes have a distance scale more like 30km. requiring the Cyclades and other regions to be relatively self-contained [Broodbank 2000], as seen in the leftmost diagram in Figs.2. Too large a distance scale, as happens with ferries going at 20 knots, the whole region becomes overconnected. Thus our MBA networks live in the transition between social organisation being strongly influenced by geography and it being much less relevant.

**Proximal Point Analysis (PPA)**

A very different null model, but still implicitly geographical, is provided by Proximal Point Analysis (PPA), also simple enough to be done with ruler and compass. Rather than look for clustering, the aim in PPA is, primarily, to decide which sites have a more prominent role in regional interactions. This is determined by
‘connectivity’, the number of links that a site has with its network neighbours. There are several variations on this theme but, most simply, sites are unweighted and links are non-directional. If the premise of our previous model is that communities need to interact but it is just too difficult to travel far, the premise of PPA is that communities need to interact, but it is just too difficult to sustain more than a few important interactions (a Dunbar number for communities, rather than individuals: Dunbar 1992). To perform the analysis links are drawn outward from each site to a specified number \( k \) of nearest neighbours (typically three or four). By ‘nearest’ is meant geographically closest once headlands and land travel are taken into account as before. If, when this is completed, directional arrows are removed from the links, some sites will emerge as being more connected than others, with five, six, or more, links to other sites. These sites possess greater ‘centrality’ in the network, and are anticipated to be more dominant in regional interactions. In fact, our interest in archaeological networks was fired by the application of PPA to the Early Bronze Age Cyclades by Broodbank (Broodbank 2000). Of course, as we have already implied, networks do not just form for the sake of it, but arise because of a mixture of imperatives. Broodbank suggests that the EBA Cyclades were agriculturally marginal, requiring social storage networks.

It is no surprise that island archipelagos lend themselves to PPA network analysis. With a dominant means of interaction (sea travel) and, typically, a dominant sailing technology (canoe or sailing vessels) they allow for simple networks. In particular, their early application was to anthropological networks in Oceania (Terrell 1977; Irwin 1983; Hage & Harary 1991; 1996). Coastal sites are equally amenable to a PPA analysis (Terrell 2010). However, some PPA models are for land-based networks (Collar 2007) for sites that are discrete. Networks built on PPA have a strong emphasis on local geography because most neighbours are relatively close. However, there is also the contrary effect that even isolated sites will interact with the same number of neighbours, however far they are removed from them. The end result is that interactions are inclined to lie in ‘strings’ of sites, including remote sites on the periphery, rather than the clustering of the previous model. We have shown this for our sites in Fig. 5, in which we join each site to its three or four nearest neighbours.

Insert Fig.4

For \( k = 3 \) nearest neighbours we see a Western ‘string’ encompassing Crete and the Peloponnese (with a separated northern part), The Cyclades and Dodecanese, just simply connected to each other, have no connection to N.Crete. If we were to join each site to its four nearest neighbours, the Cyclades becomes connected to the East. Unfortunately, this conclusion is not stable to our initial choice of sites, our only input. This is primarily because we have chosen so few (39) sites. If we were to double the number of sites, the Cyclades would become disconnected again for \( k = 4 \). However, we think of the number of connections that a site will make as something intrinsic to the nature of society, a Dunbar number, whereas the number of sites on the map is purely artificial, perhaps a reflection of our knowledge. We need to distinguish between conclusions that follow from the principles of the model and conclusions that are a consequence of our ignorance about the record, an issue that will recur throughout this discussion. This is not to damn PPA as a qualitative guide for those cases, in its applications such as contemporary and recent anthropology (see Terrell 2010), for which the record is very good.

Agency: Optimisation

Our null models assumed that sites would like to interact directly with other sites for the purposes of exchange but are limited in doing so, either because it is difficult to travel too far, or difficult to sustain too many exchanges. This need to establish a balance has an element of truth and suggests that social networks are ‘optimal’ in some sense. We need to be careful that we do not create too simplistic a Panglossian ‘best of all possible worlds’ but recent years have seen a revival of the idea that social networks do evolve to some form of ‘best’ behaviour, particularly in economic theory (Jackson 2008). While not always correct, the simplicity of the approach gives us a more sophisticated set of null models than those we have considered so far.
There are two different approaches. The first looks to identify the ‘best’ network with the ‘most likely’ network. The second of these is more social utilitarian, and identifies ‘best’ with ‘most efficient’.

III. The ‘Most Likely’ Networks

The first optimisation approach looks to identify the ‘best’ network with the ‘most likely’ network, all other things being equal, within the constraints of our knowledge. There is a long literature to this approach, which maximises ‘information’, understood as Shannon entropy. To choose less likely networks would, in some sense, correspond to assuming information that we did not have (Batty 2010). In particular, models of this type have been used extensively in modelling transport flows (Erlander 1990) and we shall just pick out those aspects that might be relevant to the ‘flows’ of Aegean maritime networks.

The discussion that follows is fairly technical, probably not enough so for readers really wanting to understand the models, and too much for readers uncomfortable with algebra. For the former more details about the approach can be found elsewhere in our work (Knappett et al. 2008; Evans et al. 2009).

Maximum entropy: Basic models

One of the best known families of models used for flows are known as Gravity Models (e.g. see Jensen-Butler 1972 as one of many summaries). Define $F_{ij}$ as the ‘flow’ from site $i$ to site $j$ where, for our Aegean network of Fig.1, $i$ and $j$ take values from 1 to 39, as before. We can think of $F_{ij}$ as, say, the number of vessels per year travelling from $i$ to $j$, or some equivalent measure of exchange.

In the `doubly constrained’ gravity model, the total outflow from a site $i$, $O_i$ (e.g. the total number of vessels leaving $i$ in a year), and the total inflow to a site $j$, $I_j$ (e.g. number of vessels arriving) are inputs to be specified, so

$$O_i = \sum_j F_{ij}, \quad I_j = \sum_i F_{ij}$$  \hspace{1cm} (1)

Likewise the ‘cost’ of sustaining each link has to be specified, say $c_{ij}$, and then the total ‘cost’ of maintaining a given network is fixed to be some parameter $C$

$$C = \sum_{ij} F_{ij}c_{ij}$$

The idea is that the most likely pattern of flows is that which maximises the entropy $S$

$$S = -\sum_{ij} F_{ij}(\ln(F_{ij}) - 1)$$

(e.g. see Ball (2004), Batty (2010)) subject to the constraints on flows and costs given above. That is, one finds the most likely flows if trips between sites are themselves equally likely, provided the total pattern conforms to the specified constraints. This may seem an unreasonable assumption to make but it has found wide applicability in a variety of fields. In fact one can prove that the optimal network of flows is of the form (Batty 2010)

$$F_{ij} = A_i O_i B_j I_j \exp(-\beta c_{ij})$$  \hspace{1cm} (2)

Here the parameters $A_i$ and $B_j$ are found from the constraints on the input and output flows (1)

$$\frac{1}{A_i} = \sum_j B_j I_j \exp(-\beta c_{ij}) , \quad \frac{1}{B_j} = \sum_i A_i O_i \exp(-\beta c_{ij})$$

while choosing the parameter $\beta$ is equivalent to setting the total cost $C$. It has to be said that this is more sophisticated than the earlier approach to gravity models of just specifying the form of the flow to be that of equation (2) by fiat, where we define $A_i$, $O_i$, and $B_j$, $I_j$ to be the populations of $i$ and $j$ respectively, with no internal consistency. Nonetheless, many applications of this approach exist. For example, for applications to imperial trade see Mitchener and Weidenmier (2008), and to Toltec networks see Alden (1979).

In archaeology we rarely have a direct handle on the flows in and out of a site. It is more likely that we have some way of estimating the size of a site, be it through the area occupied, size of key monument
(Renfrew Malta XTent – Renfrew and Level 1979) or analysis or available local resources (Broodbank 2000). We might therefore consider the (over)simple case in which the inflow \( I_j \) and outflow \( O_j \) are always equal to each other and are set equal to the size of the site \( S_j \), so that \( O_j = I_j = S_j \) for all \( j \). For our MBA Aegean example we have classified our sites as large, medium or small (see Appendix for details) and taken our \( S_j \) to have values 1, \( \frac{1}{2} \) or \( \frac{1}{3} \) respectively (the model depends only on the relative values of flow).

In many problems, including archaeological ones, the ‘costs’ are not known. For this reason it is common to choose the costs such that the longer the distance between two sites, the higher the cost. In our work we have chosen our costs such that \( \exp(-\beta c_{ij}) = V(d_{ij}/D) \) where the \( d_{ij} \) are the effective distances we introduced earlier, and \( D \) is a distance scale to be specified (it plays a similar role to \( \beta \)). For the moment we can think of \( V(x) \) as a single journey ‘likelihood function’ or an ‘ease-of-travel’ function that quantifies the fuzziness that we argued for in the fixed radius model.’ We choose \( V \) to have the form shown in Fig.5. This shape means that all short journeys, \( d_{ij}<D \), are easy (\( V \) close to 1) and have a cost close to the minimum possible. All long journeys, \( d_{ij}>D \), are difficult (\( V \) close to 0) and are of a similar high cost. For intermediate distances, \( d_{ij}\sim D \), the function falls smoothly but rapidly.

**Insert Fig.5**

By way of comparison, the networks of Fig.4 are effectively using a simple cut-off function in which \( V(x) = 1 \) for \( 0<x<1 \) and \( V(x) = 0, x>1 \) where \( x=d_{ij}/D \). Our function avoids making such a crude distinction between trips which are and are not possible, and is better suited to our imprecise knowledge of what was possible (e.g. \( D=100 \text{km} \) or \( D=110 \text{km} \) a better representation of the maximum distance travelled in a day), the errors in our estimates of ancient travelling times, and the inevitable variations from day to day of what was feasible in practice.

**Insert Fig.6**

In Fig.6 we have shown the maximum entropy gravity model for \( D=70 \text{ km} \). We see that, despite the (intentionally) small 70km distance scale, the model enforces connectivity across the S. Aegean, which we might not have expected. This is because, with both site outflows and inflows determined by (non-zero) site size, they have to go somewhere, even for a site far removed from others. The outcome for \( D = 70 \text{ km} \) is something rather like PPA, with a strong Western link and strong connections around the rim of the network, but with an infill of weak links. As \( D \) increases to 150km, so does the number of links but, since the total flows remain fixed, links become weaker. This is not how trade networks should behave.

**Maximum entropy: An enhanced model**

One of the limitations of all the models considered so far is that differences in the site sizes or their local resources play no role in our fixed radius model and PPA models, or our fixed inputs, as in the doubly constrained gravity model just considered. However, a variant of the gravity model does produce a variable measure of site ‘attractiveness’. This model, originally designed for urban planning (Wilson 1967; 1970) was applied by Rihll and Wilson (Rihll and Wilson 1987; 1991) to an archaeology context, that of Iron Age Mainland Greek city states. In this variant the flows are given by

\[
F_{ij} = A_i O_j W_j^\beta V(d_{ij}/D)
\]

(2)

Rihll and Wilson chose \( V(d_{ij}/D) = \exp(-d_{ij}/D) \) using direct distances and a typical scale of 20km to represent travel by foot. As before the output of each site is an input parameter (chosen to be 1.0 for all sites in Rihll and Wilson) and \( A_i \) is set self-consistently as above. The parameter \( W_j \) is the attractiveness of site \( j \). This is an output of the model since it is set equal to the total flow into site, \( W_j = \sum_i F_{ij} \), which is no longer a fixed input parameter \( I_j \). The new input parameter \( \beta \) is usually chosen to be a little above 1.0.
As we take larger and larger values for $\beta$, we find fewer and fewer sites attract larger and larger fractions of the total flow into their sites which must therefore come from sites further and further away. Thus this parameter controls the nature of this ‘rich get richer’ phenomena. A smaller and smaller $D$ counteracts this effect, especially for the sharp exponential fall off used by Rihll and Wilson.

We have applied the RW model to our context, setting $O$ equal to our input site sizes $O = S$, and using our $V(x)$ from Fig 5 which has a less severe cutoff of large distances than the exponential. The results are shown in Fig.7. The networks produced by this model tends to have star-like structures as a small number of sites (‘terminals’ in the language of Rihll and Wilson 1987; 1991) suck in most of the flow from a neighborhood and their output is directed between themselves. Conversely, most sites have no input and all their output is focused on the nearest terminal site. This is not the type of global pattern produced by a network of exchange and we do not think this is a useful model for our purposes. While dominant trading centres will emerge, and trade will be unbalanced at many sites, we would still expect to see some level of exchange to occur at all levels. We would also expect to see more significant coupling between attractiveness/total input and the output, yet the latter remains fixed. Indeed it does not matter how remote a site is, or how high its interaction costs are; it will always maintain a flow to the nearest terminal site. Where this model may have a role in archaeology is in power networks as a star is a sensible representation of the domination of a local region by a single site, so perhaps it is better considered alongside models such as the XTent model (Renfrew and Level 1979; Bevan 2010).

### IV. The ‘Most Efficient’ Networks

An alternative approach to optimisation is to adopt a social utilitarian principle and interpret ‘best’ as ‘most beneficial’, or ‘most efficient’ (although these may not be synonymous). These models are common in socioeconomics, or any system of exchange in which there are identifiable costs and benefits, for which ‘most beneficial’ or ‘most efficient’ means achieving the greatest benefits relative to the costs. Even simple networks suggest a spectrum of benefits that encompass social storage, exogamy, acquisition of raw materials, distribution of prestige goods, cultural exchange and trade and it is not unreasonable to assume that networks evolve to accrue higher benefits.

#### Utility functions or social potentials

To each network we associate what in economics we would call a ‘cost/benefit’ or ‘utility’ function (Jackson 2008) and in sociology a ‘social potential’ (e.g. Bejan and Merkx 2007). With two of us as physicists, we call it a ‘Hamiltonian’ $H$. Whatever it is called, it describes the ‘costs’ minus the ‘benefits’ of the network. If we assume that the network adjusts so as to increase the surplus of benefits over costs or increase utility, we are seeking to find the networks that minimise $H$. This notion is referred to as ‘strong efficiency’. See Jackson (2008) for discussions of this and other definitions of efficiency.

At the very least, $H$ contains two types of term; those describing the benefits of exchange and those describing the cost of maintaining the resulting network. In our models, the Hamiltonian utility function $H$, which characterises each configuration of the system, separates into four terms

$$H = -\kappa S - \lambda E + (j P + \mu F).$$

(5)

With all coefficients positive, $E$ represents the benefits of exchange and $F$ the cost of maintaining the network. In addition, the first term $S$ represents the benefits of local resources and $P$ the cost of maintaining the population. The parameters $\kappa, \lambda, j, \mu$ which control $H$ are measures of site self-sufficiency, constraints on population size, etc. All other things being equal, increasing $\lambda$ enhances the importance of
inter-site interaction, whereas increasing $\kappa$ augments the importance of single site behaviour. On the other hand, increasing $\mu$ effectively corresponds to reducing population, and increasing $\mu$ reduces exchange. However, unlike for the constrained entropy models, these constraints are averaged over the network and do not determine specific site behaviour.

As with the entropy models, beyond these generalised constraints, the inputs are the sites’ fixed carrying capacities $S$, and the intersite ‘potentials’ $V(d_i/D)$ of Fig.5, again understood as a measure of the difficulty to travel from site $i$ to site $j$ in a single journey. The direct outputs are again the flows $F_{ij}$ between sites $i$ and $j$. However there is an important difference from the treatment in Gravity models. For us the total output from a site is not fixed but is allowed to vary. We represent this via an output variable occupation fraction $v_i$ for each site $i$ from which we construct the ‘site weight’ (actual site size or population) $P_i=Sv_i$, which can be bigger or smaller than the fixed site carrying capacity $S$. In another departure from the gravity model approach we also allow the total output flow from a site to be less or equal to its site weight, $P_i=Sv_i$. We represent this by assigning a variable $e_i$ to each link, describing the fraction of the possible output from site $i$ which is flowing along the link to site $j$. Put another way, $e_{ij}$ represents the likelihood of an individual (or vessel) at site $i$ travelling to site $j$. Thus $F_{ij} = S_i v_i e_{ij}$ but $0 \leq \sum_j F_{ij} \leq S_i v_i = P_i$. The advantage over the gravity model description is that for some parameters in our model, when exchange between sites is very expensive, no interaction need occur and each site can exist on its own resources and be of a reasonable size. As we have noted before, in a constrained gravity model, every site will always have an output, regardless of how inefficient trade may be e.g. for a remote site.

These outputs relate only to the local properties of the sites and links. We have also constructed outputs which reflect the effect of the network on individual sites. The most important of these is site rank (essentially the same as PageRank). This is a measure of the global flow of people/trade passing through a site, an attribute of how the network functions as a whole. As a first approximation this is the flow from nearest neighbours, used by Rihll and Wilson to indicate attractiveness. However in general our site rank depends on the structure of the whole network. Sites with high ranking in comparison to their site weights have high impact and are, literally, punching above their weight. In fact, we use the ratio of rank/weight to define ‘impact.’ See our work elsewhere (Knappett et al. 2008; in press).

In detail, $S = \sum_i S_i v_i (1 - v_i)$ describes the benefits of local resources. As such, it is a sum of terms, one for each site, which describes the exploitation of the site as a function of its ‘population’. The detail is not crucial, as long as over-exploitation of resources incurs an increasingly non-linear cost, whereas under-exploitation permits growth.

$E = \sum_{i,j} V(d_{ij}/D)(S_i v_i e_{ij} (S_j v_j)$ denotes the benefits from exchange. It is a sum of terms for every pair of sites. Thus, direct long distance interactions give virtually no benefit and are unlikely to appear in our simulations; if not impossible they are deemed to carry prohibitively high overheads. It is much better to island-hop. With $S_i v_i = P_i$ the population at site $i$, the form of $E$ is ‘gravitational’, based on the premise that it is advantageous, in cultural exchange or trade, if both a site and its exchange partner, have large resources.

The final terms (in brackets) enable us to impose constraints on population size $P = \sum_i S_i v_i$ and on total trading links (and/or journeys made) $F = \sum_{i,j} F_{ij} = \sum_{i,j} S_i v_i e_{ij}$

Organisational hierarchy

Before going further, we explore the nature of our ‘gravitational’ exchange benefits in $E$ in more detail, since it is an important part of our model-making. Our earlier gravity models concerned trade flows, and we did not discuss social organisation, beyond aggregating communities into sites characterised by total population, their gravitational ‘centre of mass’. In a social context it was suggested as far back as 1713 (attributed to Georges Berkeley, but quoted in The Guardian (Guardian 1713)) that
‘There is a principle of attraction in ... the minds of men’, which draws people together into ‘communities, clubs, families, friendships, and all various aspects of society’ Further, ‘the attraction is strongest between those which are placed nearest to each other’

Benefits and trade are not the same, but should the quotation above be better understood as a statement about the former, rather than a consequence for the latter? This requires us to think about the levels of social organisation, which lead to different scales of interaction within the network. For island archipelagos we would define the relevant units of organisation as

- the local community (microscopic)
- the island (mesoscopic)
- the exchange network (macroscopic)

Similar hierarchies can be constructed for other social systems. For the moment we ignore how individual households might be subsumed into the local communities. We shall return to this later.

We now have a practical issue and a conceptual issue. The practical issue is one of where to put the sites, particularly when the record is poor. Ideally, we would have liked the model to tell us where the most important sites should be, but we don’t have models sufficiently sophisticated to be able to start with a blank map and tell us where to put the pins. Rather, we take the sites as given, and then use the model to tell us which are important and which are not. The conceptual problem is will our model give us a different answer if the effect of an excavation (typically a rescue excavation) gives us a new site? If it does in any significant way, we are beholden to the whims of the positioning of the next airport or supermarket. Again, this is particularly important when the record is poor, as is the case for the MBA Aegean.

Gravitational models help resolve both of these problems. Newtonian gravity is remarkable in that breaking up the masses into smaller parts does not change the result as long as the position of the centres of mass is unchanged. All that matters is the total masses and their positions. We can imagine an island counterpart to this, in which the contribution of the communities on an island to the benefits of exchange only depends on its total population, say. Then how this population is distributed throughout the island may not matter much and how resources are distributed through the island may not matter much either to the network as a whole. In one sense the whole is the sum of the parts. That is, we can subsume the microscopic level into the mesoscopic (Evans et al. 2009). A little caution is required, in that our ability to think of the whole as just the sum of its parts in Newtonian physics depends on the particular form of the gravitational potential, requiring an inverse square law for forces. We would not expect the benefits of exchange to be prescribed so exactly. To achieve this we exclude all interactions between sites closer than some small distance; we use 10km in our work. This is equivalent to modifying the potential $V$ of Fig.5 so that its is large and negative below this value. However, for fast enough falloff, there will be an approximate independence of how we distribute resources and people between sites, good enough for the impoverished data sets that we are considering. It is in this regard that that these models minimise the effects of our ignorance of the record. We stress again that our gravitational modelling lies in the exchange benefits as given in the term $E$. Our outputs are not gravitational in the sense of the entropy models (or the original gravitational models), although qualitatively there are similarities; the flow between large sites is large.

However, before discussing how realistic our model might be, we digress as to how we find the best network.

Is almost the best good enough?

The act of finding the most ‘efficient network’ looks highly deterministic, but the reality is more subtle. We are familiar in our personal lives with the experience of wanting to make the best choice, but finding very
little to distinguish between several of the choices available, and perhaps making a final choice with the toss of a coin. Our models reflect this. We can think of $H$ as describing a ‘landscape’, both for our model (and interpret $H$ as minus the constrained entropy in maximum entropy models). Each network that we can write down is a point on that landscape. What optimisation does is to look for the lowest part of the landscape, its *global* minimum, the network describing that point being the ‘best’ network. In practice, the landscape has many dimensions and is full of dips and bumps. As a result, there are several *local* minima offering significant improvements on our starting point but otherwise being comparably good.

As always, there are several ways to proceed. Consider the analogous optimal problem of putting a ball on our model ‘landscape’ of hills, mountains and plains and wanting to find the lowest point as it rolls under gravity. One possibility is to put the ball in some given initial position, and then ‘shake’ the landscape, giving the ball every incentive to roll downhill. After a while it gets trapped in some local minimum, which network we identify. We then repeat the process, either beginning from the same initial state or from a different one. Since we are trying to get as far downhill as possible, it shouldn’t really matter where we begin. Final outcomes will vary, as we find networks that are comparably efficient but, if the initial shaking has been good enough, they will be commensurate with the undiscovered best. However, locally they may differ significantly and we have to interpret them statistically. Technically we adopt the standard procedure of using a Boltzmann distribution to assign to each network $G$ a probability $p(G) \propto \exp (-\beta H(G))$ where $\beta$ is a large constant, the inverse volatility. The implicit statistical fluctuations in the networks we construct reflect the normal variations in a real world system.

Two examples are given in Figs.8 in which we have taken $D = 120$ km and begin from the same initial conditions. We have chosen a set of parameter values that best seems to mimic what we see archaeologically, setting the costs of trade low with respect to its benefits, with a sufficient exploitation of resources. This stimulates connectivity and encourages a number of ‘weak’ ties between clusters, this number increasing as the benefits of local resources increases. The outputs are the ‘site weights’ (populations) $S_{r}$ and the ‘link weights’ $F_{r}$. In Fig.8, the sizes of the nodes are proportional to the former and the thicknesses of the lines to the latter. We will not display the site ranks. We merely observe that, for these networks, Akrotiri and the states of N. Crete have among the highest impacts, but slightly differently in the two cases.

**Insert Fig.8**

Entropy models have a comparable problem for computing the optimal networks. An alternative approach (Wilson 2008; Wilson and Dearden 2010) is to have a deterministic algorithm for finding the nearest local minimum from any point on the landscape. The same initial condition now gives the same outcome, unlike for the stochastic approach above. However, different initial conditions will give different outcomes and we can think about these statistically. The work of Rihll and Wilson (Rihll and Wilson 1987; 1991) is closer to this way of thinking.

From a statistical mechanics viewpoint, $j$ and $\mu$ are ‘chemical potentials’ and, in our cost/benefit analysis we are working with a Grand Canonical ensemble. On the other hand, maximum entropy models correspond to working with microcanonical ensembles. The advantage of canonical ensembles over microcanonical ensembles it that, for the former, only the average behaviours are fixed and we do not have to constrain inflows or outflows on a site by site basis. Thus we do not find ourselves in the peculiar position of having to enforce long single journeys over unreasonable distances to balance flows. If we were to set $D=70$ km in our model for the same parameter values as in Fig.8 it becomes almost indistinguishable from the $D=70$ km figure of our original geographic model of Fig.3, and totally unlike the contortions required by our constrained entropy models.

However, whichever we choose, this ensemble approach leads to a different approach to how we consider the smallest scale of the social hierarchy, the household, or individual. At this fundamental level of the individual there is a whole world of Agent Based Modelling (ABM) in the literature, which shows
how a limited collective ‘social’ behaviour emerges from very limited inter-personal interactions. A problem with ABM is that it cannot encompass the range of social hierarchies present in networks as complex as Aegean trade or Greek city states. However, we can think of our statistical ensemble as describing interacting agents in a way complementary to ABM. This complementarity is familiar in describing physical systems like gases, say, where we have the option to work with the individual gas atoms in a statistical way (statistical mechanics) or through the behaviour of bulk properties of the system, like pressure and free energy (thermodynamics). In this language our approach is thermodynamical, and motivated the Boltzmann distribution quoted earlier.

V. How Do We Know If It Works?

In the first instance, it is hard to improve upon the statement in Rihll and Wilson (Rihll and Wilson 1987; 1991) that

“The purpose of a good model is to formulate simple concepts and hypotheses concerning them, and to demonstrate that, despite their simplicity, they give approximate accounts of otherwise complex behaviour of phenomena. If a model ‘works’ (faithfully represents the known evidence) then it shows that the assumptions and hypotheses built into the model contribute to an explanation of the phenomena”

Null models are useful in this, in helping to provide simple benchmarks. Although the ‘evidence’, the archaeological record, may be good for some eras, for the prehistoric era of our models it is sufficiently patchy in both space and time as to be ambiguous. A priori, this leads to an immediate problem. For our model of (4), there are four variable parameters ($D$ and three of $\kappa, \lambda, j, \mu$), provided we take the functional forms of the terms in $H$ as given. If the data set were good, whether the model could be made to ‘work’ would become clear quite quickly, and would strongly constrain the parameters. Unfortunately, with the data as poor as it is, we might expect to find so many choices of parameters (i.e. a large part of ‘parameter space’) to be commensurate with the data that it would not be clear how to proceed.

This is where non-linear optimisation comes to our rescue. A non-linear system is one in which, on prodding it, the response is not proportional to the strength of the prod. In our case the non-linearities are the conventional ones that arise when the disproportionate benefits that large sites accrue from interacting with other large sites (the ‘gravitational’ benefit) are swamped by the disproportionate costs of shortages due to high population. In the Rihll and Wilson models the non-linearities lie in the constraints upon the entropy, manifest in the non-linear benefits of site ‘attraction’.

Let us return to Fig.8. Our model turns out to be very sensitive to parameter values (having fixed $D$), only giving a picture of healthy networks for very limited choices of them as we tread a fine line between ‘boom’ and ‘bust’. This is commensurate with the observation by Broodbank et al. (2005):

“For the southern Aegean islands in the late Second and Third Palace periods, an age of intensifying trans-Mediterranean linkage and expanding political units, there may often have been precariously little middle ground to hold between the two poles of (i) high profile connectivity, wealth and population, or (ii) an obscurity and relative poverty in terms of population and access to wealth that did not carry with it even the compensation of safety from external groups”.

Empirically, this is a first sign that our model ‘works’. Not primarily because we are looking for societal collapse, although that seems to happen all too easily, but because models incorporating instability tend to lead to settlements with a wide variety of population. If we change their form to make them structurally more stable it becomes increasingly difficult to generate a wide enough range of site sizes (‘populations’) to match the record.
What seems to keep the network together is the web of ‘weak’ links. It was proposed by Granovetter (Granovetter 1973; 1983) many years ago in a seminal paper that weak links play an important role in social networks for facilitating the exchange of information and facilitating innovation. This proposition has been continually examined since and borne its weight. More generally, it has been argued (Csermely 2004) that “weak links stabilise all complex systems”, of which networks are but one type. There are different definitions of stability, which we shall hint at below, but there are certainly many examples where the presence of many weak links does indeed aid stability (Csermely 2004), including our models. Something similar seems to be happening in constrained entropy models of the Rihill and Wilson type (see Fig.7) but we have not analysed this in any detail.

This narrow path between ‘boom and bust’ plays an important role in the evolution of exchange networks. As to how the networks evolve in time, systems evolve for a variety of reasons. The models we have considered here are not subtle enough to bootstrap themselves and we have to enforce change exogenously, usually by varying the parameters smoothly, distorting the landscape. Thus, for example, as populations grow or total trade volume increases, the optimal network (lowest energy configuration) changes. It is not surprising that the model shows ‘tipping points’ as ‘valley bottoms’ rise and new valleys are formed. ‘Boom’ and ‘bust’ is a feature of non-linear systems. Another way of thinking about this is in the language of statistical mechanics or thermodynamics, where the notion of a phase transition is familiar and these collapses constitute just such a change (Wilson 2008; Wilson and Dearden 2010).

If, for example, beginning from a typical network of Fig.9, we increase the costs of trade we find the network puts its eggs in fewer baskets. The weak links are unrewarding to maintain and trading becomes confined to fewer and fewer networks until it collapses.

Insert Fig.9

We should not assume that instability is inevitable as the networks evolve. There is still enough room in the parameter space to stay in an area of stability and evolve with gentle growth or decline but the opportunity both for collapse and rapid growth (equally unstable) is always there. It is in circumstances like this that social space and geographical space come together before going violently apart. The benign networks of Fig.8 in some sense go beyond the geography while being strongly conditioned by it, in giving details that we could not have predicted (albeit stochastically) with some outliers punching above their weight. However, regional geography reinstates itself in primitive form as we move towards instability, almost along the lines of our null models, with only a few strong links between collapsing regional clusters.

VI. Summary

We began this article with a general question; does the S. Aegean exchange network reflect the geographical space in which it is embedded? We have argued how conditional this is on MBA marine technology, which makes the S. Aegean a Goldilocks sea for the appearance of trading networks, with single journey distance correlated to the scale separating the main regional clusterings.

The bulk of our analysis has been devoted to a discussion of ‘best’ networks, with ‘best’ meaning ‘most likely’ or ‘most efficient’.

The former, used in town planning and transport networks, adopts a microcanonical approach in arguing for a maximisation of constrained network entropy, essentially information. We have not had time to pursue this too far but it seems, in general, that the constraints in terms of inflows and outflows on a site by site basis that are conventionally imposed are too strong to permit networks that could match the record. The reason why they will fail for some maritime networks is that imposing constraints on a site by site basis requires sites to connect willy-nilly, even to sites that are inaccessible by simple sea-travel.
Even for networks like our MBA networks, where travelling distances match typical regional separations, some direct connections look highly implausible. In this regard the models have something in common with PPA, which argues for a maximum number of connections, rather a minimum separation. This needs further work. However, unlike PPA, which is very sensitive to our knowledge (or ignorance) of the record, entropy maximisation leads to gravitational models which are relatively insensitive to our ignorance.

‘Most efficient’ or cost/benefit models are used widely in socioeconomics and avoid constraints on a site by site basis by adopting a canonical ensemble approach. As a result, we get networks that, in general, are strongly conditioned by geography. However, a characteristic of the optimal models we have discussed here is that they are prone to instabilities, and the links to geographic space make themselves starkly evident before they break down if we begin to walk off the delicate line between ‘boom’ and ‘bust’ that they embody.

As for our original concerns regarding Knossos, the importance of the N. Cretan sites is seen clearly in Fig.8. Statistically Knossos is capable of being dominant. As for Thera, the resilience of the network under the removal of one of its key nodes is, crudely, sustained by its weak links, as discussed in (Knappett et al. in press). However, intrinsic network instability along the lines indicated above could be the cause for the ultimate collapse of the network.
References


Knappett, Carl, Evans, Tim, and Ray Rivers. in press. “Modelling Maritime Interaction in the Aegean Bronze Age, II. The Eruption of Thera and the Burning of the Palaces.” *Antiquity*.


Figures

Fig. 1

Important sites, for the MBA Aegean, including Knossos (1) and Thera (10). The sea journey from the N. Cretan coast to Thera is a little over 100km.
Fig. 2

Dendrogram for the sites of Fig.1, using an effective shortest distance between sites in which sea travel is counted in physical kilometres but land travel is penalised by a `friction' factor of 3.0. A horizontal line cutting the vertical axis at distance scale $D$ cuts the dendrogram into a number of disconnected branches below the horizontal line. Each of these separate branches defines a cluster (or community) of sites within which it is possible to move between any pair of sites via a sequence of sites each of which is separated by an effective distance of $D$ or less. It is not possible to move between sites in different branches without following a link effective length greater than $D$. 

![Dendrogram image](image-url)
Fixed Radius Networks for the sites of Fig.1. Sites are placed at their geographical location. Distances between sites are judged via shortest sensible routes with land travel penalised by a factor of 3.0 but without taking currents, winds etc in to account. Two sites are linked (dark links) if the effective distance is less than $D$, where $D = 70\text{km}$, $100\text{km}$, $130\text{km}$ and $150\text{km}$. Light grey links are edges which are longer than $D$ but by no more than 20%. Sites of same colour are connected by routes via black links (i.e. all hops are less than $D$ km).
Fig. 4:

PPA for the S.Aegean sites of Fig.1, joining each site to either three or four nearest neighbours (thick black links). Thin grey links indicate links to fifth and sixth nearest neighbours. Sites of same colour are connected by routes via thick black links.
Interaction potentials as a function of distance $d$, where $x=d/D$ and by construction $V(1) = 0.5$. The solid red line is $V(x)=1/(1+x^4)$ as used in this work. The dashed blue line is $V(x)=\exp(-\beta x)=2^x$ with $\beta=ln(2)/D$ as used in many Gravity Models.
The doubly constrained gravity model, a maximum entropy model, for D=70km, 100km, 130km and 150km (reading from left to right then top to bottom). For each site $i$ the outflow $O_i$ and inflow $D_i$ is fixed equal to the fixed site size $S_i$ (taking values $1/3, 1/2,$ or $1.0$, see table 1). Only edges with flow over $1/39 \approx 0.0256$ are included. Vertices of same colour are connected by the edges shown.
Fig. 7

Rihll and Wilson model for values $D = 100\text{km}$ on the top row with $\beta = 1.04$ (on the left) and $\beta = 1.30$ (on the right). $D = 150\text{km}$ on the second row, for $\beta = 1.04$. Note how the number of weak links decreases as $\beta$ increases. For each site $i$ only the outflow $O_i$ is fixed equal to the site size $S_i$ (taking values $1/3, 1/2$, or $1.0$, see table 1). Only edges with flow over $1/39 \approx 0.0256$ are shown.
Fig. 8

Two runs of our model for values $D = 100$km with identical input parameters ($\lambda =3$, $\kappa=1, \mu=0.1, j=-2.0$). Akrotiri and Knossos are important in each. The size of vertex is proportional to the site weight $S_{iv}$, with an edge shown if the flow along that edge the flow along an edge. $S_{iv}e_{ij}$, is greater than 0.1.
Fig. 9

The effect of increasing the cost of an edge in our model. From left to right, then top to bottom, the parameter $\mu$ is raised from 0.1, to 0.3, 0.5, 1.0 and finally 1.5. For $D = 120$km ($\lambda = 3$, $\kappa = 1$, $j = -2.0$).
Table 1:
The sites enumerated in Fig.1 and the size of their local resource base, with (S), (M), (L) denoting 'small', medium' or 'large' respectively in terms of their resource base (input). This is to be distinguished from their 'populations', which are outputs.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knossos (L)</td>
<td>14</td>
<td>Kea (M)</td>
</tr>
<tr>
<td>2</td>
<td>Malia (L)</td>
<td>15</td>
<td>Karpathos (S)</td>
</tr>
<tr>
<td>3</td>
<td>Phaistos (L)</td>
<td>16</td>
<td>Rhodes (L)</td>
</tr>
<tr>
<td>4</td>
<td>Kommos (M)</td>
<td>17</td>
<td>Kos (M)</td>
</tr>
<tr>
<td>5</td>
<td>Ayia Triadha (L)</td>
<td>18</td>
<td>Miletus (L)</td>
</tr>
<tr>
<td>6</td>
<td>Palaikastro (L)</td>
<td>19</td>
<td>Iasos (M)</td>
</tr>
<tr>
<td>7</td>
<td>Zakros (M)</td>
<td>20</td>
<td>Samos (M)</td>
</tr>
<tr>
<td>8</td>
<td>Gournia (L)</td>
<td>21</td>
<td>Petras (L)</td>
</tr>
<tr>
<td>9</td>
<td>Chania (L)</td>
<td>22</td>
<td>Rethymnon (L)</td>
</tr>
<tr>
<td>10</td>
<td>Thera (M)</td>
<td>23</td>
<td>Paroikia (M)</td>
</tr>
<tr>
<td>11</td>
<td>Phylakopi (M)</td>
<td>24</td>
<td>Amorgos (S)</td>
</tr>
<tr>
<td>12</td>
<td>Kastri (M)</td>
<td>25</td>
<td>Ios (S)</td>
</tr>
<tr>
<td>13</td>
<td>Naxos (L)</td>
<td>26</td>
<td>Aegina (M)</td>
</tr>
</tbody>
</table>