

This is the author-created version of the following work:

Crabtree, Stefani A., White, Devin A., Bradshaw, Corey J.A., Saltré, Frédérik, Williams, Alan N., Beaman, Robin J., Bird, Michael I., and Ulm, Sean (2021)
Landscape rules predict optimal superhighways for the first peopling of Sahul.
Nature Human Behaviour, 5 pp. 1303-1313.

Access to this file is available from:

<https://researchonline.jcu.edu.au/68058/>

© National Technology and Engineering Solutions of Sandia, LLC under exclusive license to Springer Nature Limited 2021.

Please refer to the original source for the final version of this work:

<https://doi.org/10.1038/s41562-021-01106-8>

Landscape rules predict optimal super-highways for the first peopling of Sahul

Stefani A. Crabtree^{1,2,3,4†*}, Devin A. White^{5,6†}, Corey J. A. Bradshaw^{7,8}, Frédéric Saltré^{7,8}, Alan N. Williams^{9,10,11}, Robin J. Beaman¹², Michael I. Bird^{3,12}, Sean Ulm^{3,13,♦}

¹Department of Environment and Society, Utah State University, 5200 Old Main Hill, Logan, Utah 84322, USA

²The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

³ARC Centre of Excellence for Australian Biodiversity and Heritage, James Cook University, PO Box 6811, Cairns, Queensland 4870, Australia

⁴Université de Paris, INSERM U1284, Center for Research and Interdisciplinarity (CRI), F-75006 Paris, France

⁵Autonomous Sensing and Perception, Sandia National Laboratories, Albuquerque, New Mexico 87123, USA

⁶Department of Anthropology, University of Tennessee, Knoxville, 502 Strong Hall, Knoxville, Tennessee 37996, USA

⁷ARC Centre of Excellence for Australian Biodiversity and Heritage, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia

⁸Global Ecology, College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia

⁹Climate Change Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales 2052, Australia

¹⁰ARC Centre of Excellence for Australian Biodiversity and Heritage, University of New South Wales 2052, Australia

¹¹EMM Consulting Pty Ltd, Suite 01, 20 Chandos Street, St Leonards, New South Wales 2065, Australia

¹²College of Science and Engineering, James Cook University, PO Box 6811, Cairns, Queensland 4870, Australia

¹³College of Arts, Society and Education, James Cook University, PO Box 6811, Cairns, Queensland 4870, Australia

†These authors contributed equally

*corresponding author: stefani.crabtree@usu.edu

ABSTRACT

Archaeological data and demographic modelling suggest that the peopling of Sahul required substantial populations, occurred rapidly within a few thousand years, and encompassed environments ranging from hyper-arid deserts to temperate uplands and tropical rainforests. How this migration occurred, and how humans responded to the physical environments they encountered have, however, remained largely speculative. By constructing a high-resolution Digital Elevation Model for Sahul and coupling it with fine-scale viewshed analysis of landscape prominence, least-cost pedestrian travel modelling, and high-performance computing, we create over 125 billion potential migratory pathways, whereby the most parsimonious routes traversed emerge. Our analysis revealed several major pathways—super-highways—transecting the continent, that we evaluated using archaeological data. These results suggest that the earliest Australian ancestors adopted a set of fundamental rules shaped by physiological capacity, attraction to visually prominent landscape features, and freshwater distribution to maximise survival, even without previous experience of the landscapes they encountered.

Introduction

The ability of humans to expand our range rapidly and efficiently exploits a capacity to modify the environment for our benefit¹, and to occupy a wide ecological niche. Humans are capable not only of switching food sources when required²⁻⁴, but also of migrating to more favourable environments quickly as the need arises⁵, and adapting to changing environmental conditions⁶. This ecological versatility⁷ has enabled our rapid global expansion without requiring expensive evolutionary adaptation to micro-environments or even extensive technological modifications.

But have modern humans always had the capacity to exploit novel environments rapidly and efficiently? One way to answer this question is to examine the ecological limits of migrations of the groups of modern humans who exited Africa as early as 120,000 years ago to people the rest of the world. Foundational studies of these migratory events can be traced to the 19th Century^{8,9}, and they continue to be refined as researchers investigate the first Out-of-Africa events¹⁰, the peopling of the Americas¹¹, the exploration of the South Pacific¹², and the peopling of Australia¹³. Yet our understanding of such movements is often only conceptual, the result of interpolation or inference among points across landscapes where sparse archaeological investigations have occurred. Genomic research has shed some light on

movements of people across Eurasia and elsewhere¹⁴; however, even these analyses are constrained by the availability of underlying data and the temporal resolution of the genetic process underlying the patterns. While increasingly complex analytical approaches are being adopted to fill the gaps among rare spatial data, such as spatio-temporal analysis of radiocarbon dates¹⁵ and agent-based/cellular-automaton models^{5,16–19}, such models frequently have to make assumptions about demographics and speed of dispersal.

We overcome these limitations here by developing an approach that objectively generates parsimonious corridors of movement over an entire continent without assuming any particular human demographic rates or any other variables for which few data currently exist; the *From Everywhere To Everywhere* approach²⁰ (see Methods). This allows an agnostic identification of the most-probable routes by which people would have first populated an entire continent.

The super-continent of Australia and New Guinea that was connected during much of the late Quaternary—known as the continent of Sahul (Fig. 1)—provides a case study to test and examine human migration into novel landscapes using this approach. We apply this method to generate corridors of movement over Sahul at the broadest spatial scales yet considered. We couple fine-resolution palaeo-geographic and hydrological data with physiological characteristics of hunter-gatherers²¹ to identify optimal routes across the continent—super-highways—that we then evaluate against the available archaeological data. In addition to providing an improved quantification of likely routes traversed during the initial peopling of this landscape, our methodology can be applied to any of the current spatio-temporal debates on human migration events since the late Pleistocene.

Studying migratory events into Sahul 50,000–70,000 years ago is potentially illuminating, providing the challenge of a novel landscape, novel faunal types, and a landscape devoid of other hominins. While our understanding of the possible oceanic-voyaging pathways through the islands of modern eastern Indonesia into Sahul is relatively more advanced—derived from studies of least-cost paths²², palaeogeography^{23–25}, inter-visibility^{24–26}, and coupled demographic-drift voyaging models²⁷—explaining how people likely traversed the large and diverse continent of Sahul once they arrived so far remains largely speculative²⁸.

There have been even fewer attempts to quantify *how* this movement of people occurred or where likely impacts were concentrated. Birdsell²⁹ speculated that people moved rapidly in all directions from their arrival point, whereas Bowdler³⁰ hypothesised that people initially followed the coasts before occupying the interior. Horton³¹ and Tindale³² further argued that people initially moved through the relatively wetter northern and eastern inland woodlands

along riverine corridors and then the coast. More recent studies have reworked these various hypotheses to consider a range of biogeographic^{33,34}, ecological/climatic^{35,36}, and sociological/technological^{37,38} determinants. For example, O'Connell and Allen^{28,39} drew on the concept of optimal foraging theory to propose a conceptual model of human movement throughout the continental interior driven by resource availability, with the major river basins representing the most-attractive environments to human foragers. Most recently, Bird et al.⁴⁰ hypothesised pathways for movement based on the distance to water and a travel-cost surface, leading the way toward identifying the mechanisms underlying how these movements might have occurred.

Of course, several challenges arise with examining the movement of people 70,000 years ago, including attempting to capture the demographics of the travellers, the ways that geography and ecology provided constraints to, or opportunities for, movement, and the choices people made to explore the open space. Yet, previous research has established that extant hunter-gatherers use space in similar ways to other animals^{41–43}. The models we present here build on this literature and show how human decision-making in novel landscapes can lead to distinct patterns of movement that are detectable via advanced computational methods.

By exposing the underlying geological attributes—or baseline landscape structure—of the continent while incorporating models that encapsulate decision-making via visual landform prominence⁴⁴, water distribution, and anisotropic (directionally dependent) movement, we identified a network of optimal movement corridors transecting one of the driest inhabited continents on Earth³¹. Comparing the emergent corridors to the distribution of known archaeological sites dating to $\geq 35,000$ years ago, we demonstrate that the most parsimonious paths correspond closely to regions of earliest occupation. Our results emphasise the extraordinary cultural plasticity of humans to adapt to novel and challenging environments rapidly and efficiently. Our findings are also critical for modern policy and land-management issues; since our results suggest fundamental rules for migration, highlighting the important routes and pathways that populations used to survive arid conditions and climatic disruption in the past, these same rules could be used to highlight important routes and pathways that populations may employ in the future in other regions, according to certain climate scenarios^{45,46}.

Summary of Methods

To examine how people entered Sahul >50,000 years ago, we first created a Sahul digital elevation model that links modern New Guinea, Australia, and Tasmania into one supercontinent (Supplementary Figures 1-2). Then, to examine how people traversed Sahul, we applied various weighting layers, including (i) a layer that calculates the visibility of all landforms across Sahul based on the topography from our Sahul digital elevation model (500 m resolution; Supplementary Figure 4), and (ii) a layer that calculate streams with Strahler Order 9 (representing the most probable permanent bodies of water) extended to the coastline (Supplementary Figure 3). These layers are collated for our simulations to create Sahul's topography, visual prominence, and a layer of permanent streams and rivers.

To model the travellers, we applied ethnographic data on the movement characteristics of a typical 25-year-old woman, suggesting that, while male travellers might make longer forays, female travellers would likely be leading families into new territories and holding groups together. Using the physical characteristics of these women, we then simulate various scenarios of travel across Sahul using the *From Everywhere To Everywhere* modelling platform. We parameterise the travel function assuming the least anisotropically 'costly' movement (in terms of caloric expenditure) across the landscape, incorporating the weighting layers outlined above. Travellers in our simulations attempt to move from their origin point to their input destination point while minimising caloric costs. We ran these models enough times to create over 125 billion pathways, tracking every time a traveller passes through a pixel (500 m grid cells), thus creating pathways that transect Sahul. The number of times a pixel is crossed indicates the 'attractiveness' of that cell for movement. We then applied crossing thresholds to the pathways to indicate those in the 1%, 5%, and 10% attractiveness considered across simulations.

To estimate which modelled paths were most likely migration routes, we developed statistical methods to compare the network of pathways against our archaeological data. We compiled a database of sites of $\geq 35,000$ years old to create a concordance dataset. Next, we developed several goodness-of-fit metrics, including approximation of the Bayesian information criterion weight ($wBIC_a$) that assesses portions of paths, stitching together the paths that provide the highest likelihood of performance against our concordance archaeological dataset. This then gives relative probabilities of likelihood that paths were traversed. This taken together creates the paths we report below. We describe all methods in greater detail in the Methods, provide code in the Supplementary Data, and all model outputs are available freely on our GitHub repository.

Results

Our results show several optimal pathways across Sahul according to the underlying datasets we applied: (1) a harmonised bathymetry/topography layer at 500 m resolution covering Sahul ~ 50,000 years ago combined with, (2) a hydrological network, propagated to the Sahul palaeocoastline and based on the topographic layer (Fig. 1), and (3) visual prominence, also developed from the topographic layer (Fig. 2).

Several highly ranked ‘super-highways’ are clearly visible (Fig. 3, Supplementary Figures 5-7, Table 1; see Methods), both around the continental margin (probability of choosing path $p = 0.9947$) and, perhaps more unexpectedly, through the continental interior ($p = 0.9729$). Table 1 reports the probability that Besag’s L , a multi-distance spatial cluster analysis (see Methods), exceeds the upper confidence limit of 10,000 random permutations. We discuss the ensuing probabilities in terms of equality (‘=’) to indicate a precisely estimated value, and the inequality (‘>’) to indicate that the estimated value is larger than what is stated, which is only used when the probability is close to 1.0 but exceeds the precision used for reporting (e.g., $p = 0.99997$ reported as $p > 0.9999$).

Many ‘feeder’ routes extend from the hypothesised ‘southern route’ landfall^{26,27} in the northwest ($p = 0.8496$), to highly ranked corridors connecting known, old archaeological sites in the Kimberley, Pilbara, and Arnhem Land regions (Fig. 3; Supplementary Figure 13). The lesser-travelled fork of the highway that meanders along the edge of the Great Sandy Desert ($p = 0.6489$) aligns well with archaeological data including Parnkupirta (Lake Gregory) and Puritjarra (Cleland Hills), two of the earliest archaeological sites in the arid centre. This corridor, along with several others, converge to a single super-highway linking the northwest of Sahul through the centre ($p = 0.9861$), to Lake Eyre (**LE**, Fig. 3), and on to the southeast of the continent. A corridor through southern Sahul also passes close (20-50 km, a long day of walking) to known $\geq 35,000$ year-old archaeological sites, including Devils Lair (Margaret River), Koonalda Cave (Nullarbor) and Warraty rock shelter (Flinders Range), linking the southwest and southeast of the continent along the southern coast ($p = 0.9709$), close to the modern coastline. This super-highway runs east-west toward the south-eastern highlands through the northern Murray-Darling River system, close to well-documented early sites near Menindee Lakes (**ML**, Fig. 3). South-eastern Sahul is linked to Tasmania ($p > 0.9999$) via a corridor east of Lake Bass (**LB**, Fig. 3) (see Supplementary Data and Code for individual route probabilities).

Table 1. 0.005 DEGREE RESOLUTION Rank of models used to produce optimal path network. Origin = origin starting points; destination = destination points; thresh (%) = percentile of most frequent paths; c = proportion of Sahul land cells with a path generated at that percentile threshold for all Sahul; BIC_a = approximate Bayesian information criterion; $wBIC_a$ = relative BIC_a weight of model (across all models, $\Sigma wBIC_a = 1$); p = probability that Besag's L (multi-distance spatial cluster analysis) exceeds the upper confidence limit after 10,000 random permutations (see Methods). Two parameters held constant in these tests: k number of model parameters, set to 4; n points = total number of points (archaeological sites) used to calculate statistics, set to 40 archaeological sites older than 35,000 years. Models that do not perform highly ($p < 0.9$) are reported in Supplemental Tables 2-7.

| rank | origin | destination | weighting | thresh (%) | c | BIC_a | $wBIC_a$ | p |
|------|-----------|-------------|------------|------------|--------|---------|----------|--------|
| 1 | coastline | grid | prominence | 1 | 0.0012 | 4.2330 | 0.2644 | 0.9913 |
| 2 | grid | grid | prominence | 1 | 0.0034 | 4.6730 | 0.2122 | 0.9602 |
| 3 | water | grid | prominence | 1 | 0.0026 | 6.2100 | 0.0984 | 0.9835 |
| 4 | coastline | coastline | prominence | 5 | 0.0019 | 6.9861 | 0.0667 | 0.9918 |
| 5 | coastline | grid | prominence | 5 | 0.0058 | 7.4484 | 0.0530 | 0.9931 |
| 6 | water | water | prominence | 1 | 0.0020 | 7.8145 | 0.0441 | 0.9946 |
| 7 | grid | grid | prominence | 5 | 0.0172 | 7.9753 | 0.0407 | 0.9209 |
| 8 | coastline | coastline | prominence | 10 | 0.0037 | 8.3627 | 0.0335 | 0.9966 |
| 9 | coastline | water | prominence | 5 | 0.0036 | 8.6438 | 0.0291 | 0.9985 |
| 10 | coastline | grid | prominence | 10 | 0.0116 | 8.9469 | 0.0250 | 0.9388 |
| 11 | grid | grid | prominence | 10 | 0.0343 | 9.3189 | 0.0208 | 0.9408 |
| 12 | water | grid | prominence | 5 | 0.0129 | 9.4852 | 0.0191 | 0.9562 |
| 13 | coastline | coastline | prominence | 20 | 0.0075 | 9.7473 | 0.0168 | 0.9974 |
| 14 | coastline | water | prominence | 10 | 0.0072 | 10.0304 | 0.0146 | 0.9981 |
| 15 | water | grid | prominence | 10 | 0.0258 | 10.8737 | 0.0096 | 0.9552 |
| 16 | water | water | prominence | 5 | 0.0099 | 11.0241 | 0.0089 | 0.9985 |
| 17 | water | water | rivers | 5 | 0.0108 | 11.2172 | 0.0080 | 0.9909 |
| 18 | coastline | water | prominence | 20 | 0.0144 | 11.4226 | 0.0073 | 0.9952 |
| 19 | coastline | water | rivers | 20 | 0.0175 | 11.8649 | 0.0058 | 0.9696 |
| 20 | water | water | none | 5 | 0.0160 | 12.0708 | 0.0053 | 0.9570 |
| 21 | water | water | prominence | 10 | 0.0198 | 12.4105 | 0.0044 | 0.9984 |
| 22 | water | water | rivers | 10 | 0.0216 | 12.5940 | 0.0040 | 0.9955 |
| 23 | water | water | none | 10 | 0.0320 | 13.4515 | 0.0026 | 0.9597 |
| 24 | water | water | prominence | 20 | 0.0396 | 13.8029 | 0.0022 | 0.9954 |
| 25 | water | water | rivers | 20 | 0.0433 | 13.9809 | 0.0020 | 0.9953 |
| 26 | water | water | none | 20 | 0.0640 | 14.8309 | 0.0013 | 0.9630 |

From the ‘northern route’ entry^{27,47} (Fig. 3) into the modern western tip of New Guinea, the main movement corridor is along the southern slopes of the New Guinea highlands ($p = 0.9112$) linking to southern Sahul via the then-exposed Torres Strait region ($p = 0.9171$). From the southern Torres Strait region this corridor bifurcates, one corridor linking to Arnhem Land ($p > 0.9999$) around the northern shore of Lake Carpentaria, the other linking to south-eastern Sahul offshore of the modern eastern coastline ($p = 0.9827$). A more diffuse network of lower-ranked corridors crosses north-eastern Sahul ($p = 0.5061$)—aligning with ecological refuges centred on Lawn Hill and Gregory River systems—and south-western ($p = 0.5274$) Sahul and its northernmost coast ($p = 0.5484$).

No old archaeological sites are currently known within the large areas that are distant from identified movement corridors, apart from the Willandra Lakes system (**WL**, Fig. 3). In the case of Willandra Lakes, river avulsion after the Last Glacial Maximum led to the irreversible drying of what were permanent waterbodies during the period of early human arrival, which accounts for the lack of pathways revealed by our model in this area^{48–50}. As our modelled freshwater comes from the Water Observations from Space dataset⁵¹, the lack of pathways approaching the Willandra Lakes system reflects the lack of forcing the model to conform to expectations, and rather allowing the paths to emerge parsimoniously from empirical data.

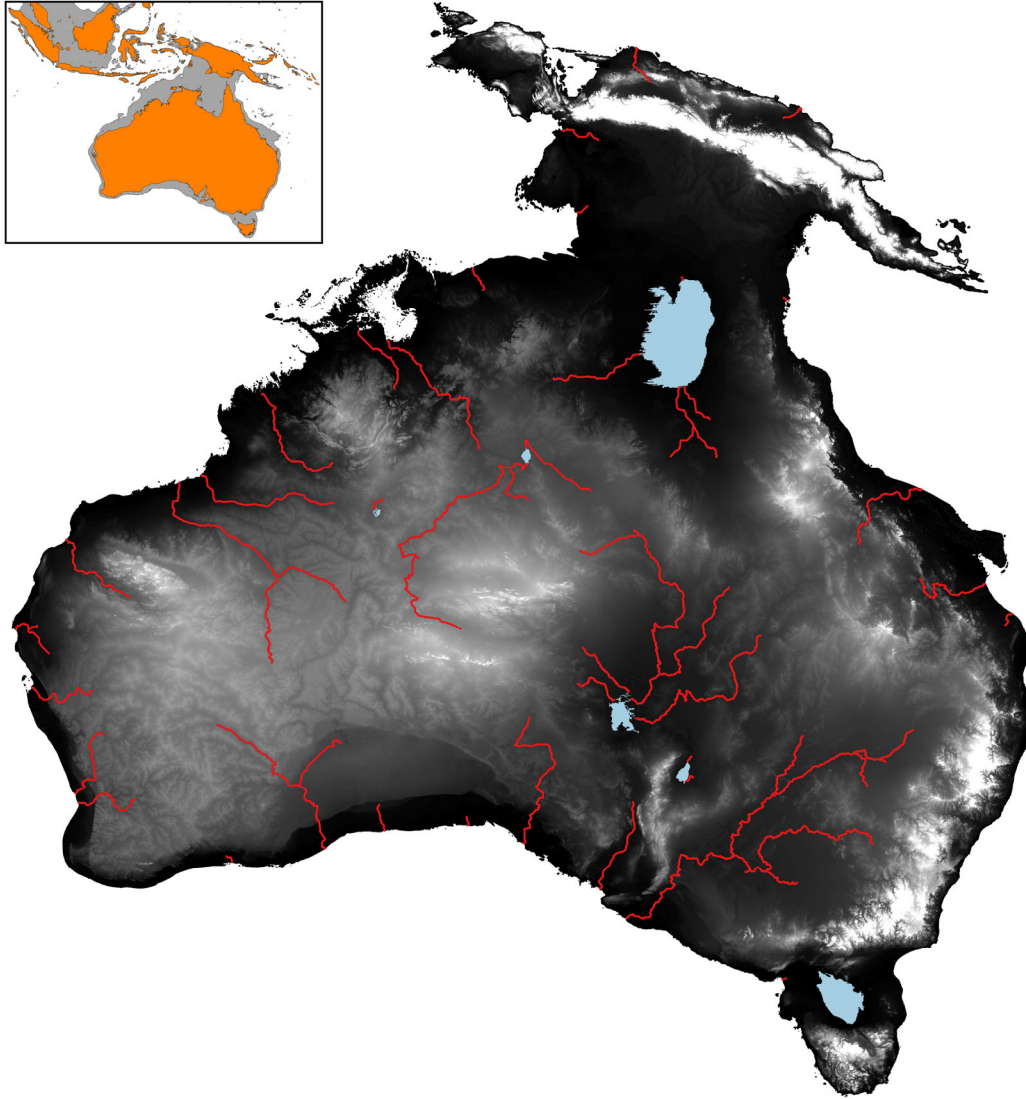
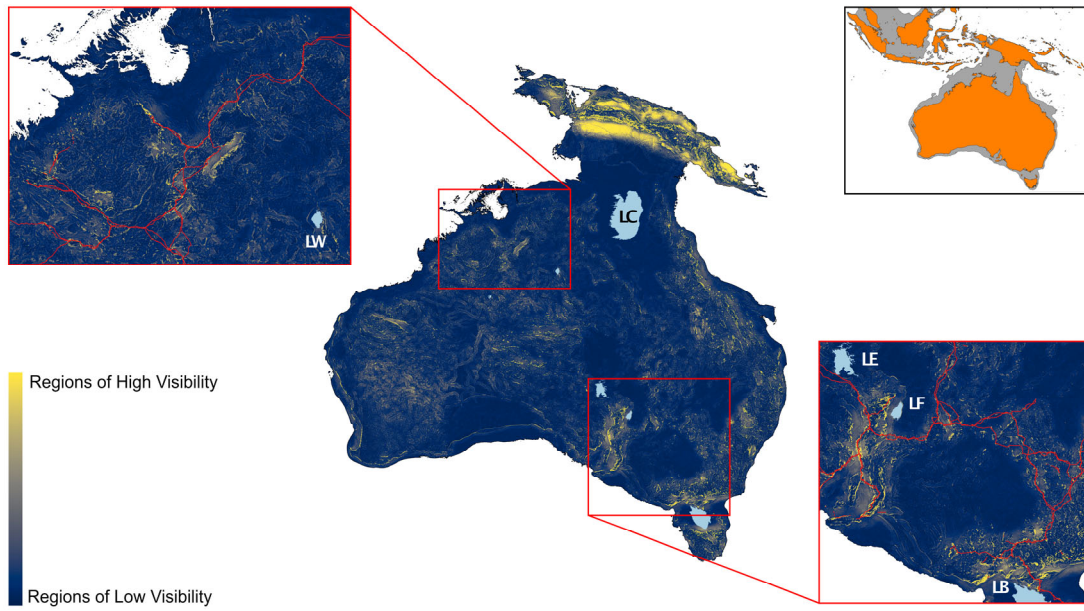


Fig. 1. Data layers used to create optimal pathways. **Legend:** Combined data layers used to create optimal pathways. Inset map: modern country boundaries superimposed on the coastlines of Sahul c. 50,000 years ago (during the approximate time of first human arrival). Main: Combined contemporary elevation data and bathymetric data. We defined the coastline as the -85 m isobath²⁷ demonstrating the extent of the Sahul landmass toward the continental shelf. The approximated extent of large inland lakes (in blue) is delineated by palaeoshoreline elevation. A hydrologically-modelled stream network (Strahler order 9 and higher, in red) indicates the likely locations and extents of major rivers and their courses to the ancient coast during the time of initial peopling of the continent.



Visual-prominence aggregate viewshed for the entire Sahul landmass. **Legend:** Visual-prominence aggregate viewshed for the entire Sahul landmass. Yellow = visually prominent features; dark blue = low visual prominence. Each cell (grid resolution of 0.004 degrees) contains a normalised value indicating how visible it is compared to every other cell. Red lines in the insets show sample output from a FETE run that used prominence weighting, which illustrates how much influence prominence has on the resulting network in that scenario. To create a layer of visual prominence for human travellers, the most visible locations are determined by beginning with the mean visibility after log-transforming the aggregate viewshed and keeping cells with values ≥ 3 standard deviations above the mean (i.e., only the most prominent locations remain). The prominence of several important mountain ranges, rocky ranges, and landforms are identifiable at this scale. The raw data underlying Figures 1 and 2 provide the foundation for our analyses. The modern coastline is visually prominent at the lower sea level around much of the continent, likely due to the modern coastline creating strong erosional features. **LC** = Lake Carpentaria; **LE** = Lake Eyre; **LF** = Lake Frome; **LB** = Lake Bass; **LW** = Lake Woods.

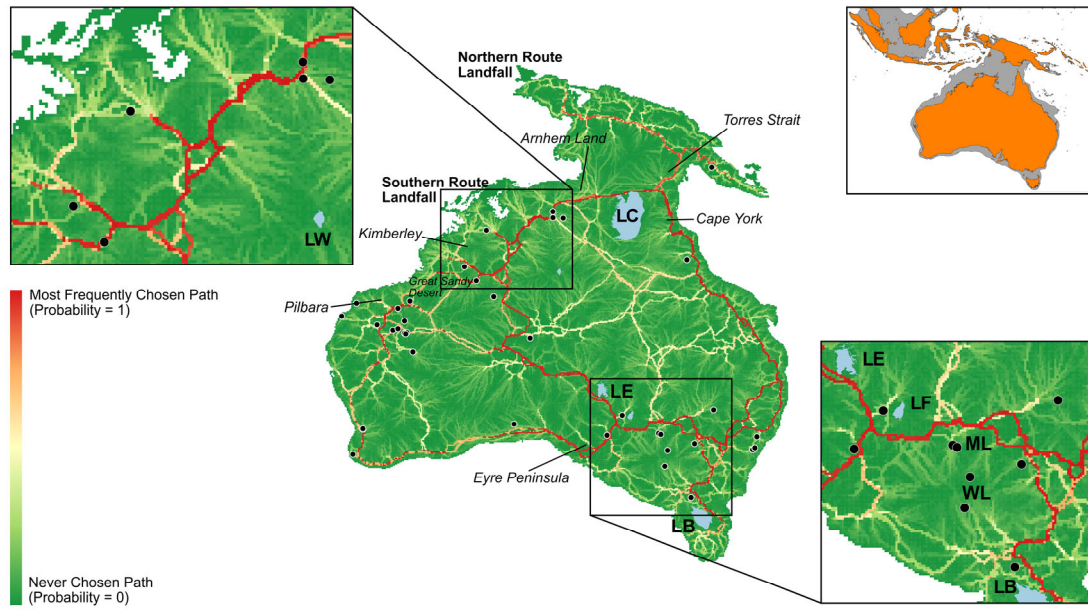


Fig. 3. Model-averaged pathway probabilities. **Legend:** Model-averaged pathway probability calculated from ~ 125 billion paths across the continent of Sahul. Areas that are marked green have paths that were never or rarely chosen, while paths of bright red have higher probabilities of optimal paths present in a grid cell across all models (grid resolution of 0.1 degrees); paths of yellow-tan have a 50% probability. Black dots indicate locations of archaeological sites older than 35,000 years. **ML** = Menindee Lakes; **WL** = Willandra Lakes; **LC** = Lake Carpentaria; **LE** = Lake Eyre; **LF** = Lake Frome; **LB** = Lake Bass; **LW** = Lake Woods. Inset map: modern country boundaries superimposed on the coastlines of Sahul <50,000 years ago (during the approximate time of first human arrival).

Table 2: Total number of paths generated for each set of origins and destinations for six travel scenarios. Scenario values are multiplied by 3 to account for the three weighting conditions used (no weighting, prominence, and rivers). Combined, we generated more than 125 billion paths to examine peopling of Sahul.

| Origins | Destinations | Scenario | Total paths generated for each scenario | Total paths generated under all weighting conditions |
|----------------|---------------------|--|--|---|
| grid | grid | generalized landscape movement and exploration across entire continent | 1,473,024,400 | 4,419,073,200 |
| coastline | grid | generalized movement from periphery to anywhere on continent | 3,620,462,160 | 10,861,386,480 |
| coastline | water | movement from periphery to reliable water sources | 10,639,328,952 | 31,917,986,856 |
| coastline | coastline | continent traversal | 8,898,526,224 | 26,695,578,672 |
| water | grid | movement from reliable water sources to anywhere on continent | 4,328,726,680 | 12,986,180,040 |
| water | water | movement between reliable water sources | 12,720,681,796 | 38,162,045,388 |

Discussion

The debate on the patterns and drivers of the peopling of Sahul began in the 1950s^{52–55}, and continues to be a subject of research and controversy. Our results move this debate beyond conceptual models to identify empirically the optimal corridors of movement throughout the continent. While underlying geological attributes generally structure movement pathways⁵⁶, ignoring constraints like the need for freshwater (Fig. 1) and the attraction to visible environmental cues (Fig. 2) always led to poor-fitting models (Supp. Tables 2-7; Supp. Figures 8-12); when ranked by explanatory power, the top 16 models are driven by visual prominence (Table 1). Two models account for almost half (0.47) of the explanatory power ($wBIC_a$). The first model ($wBIC_a=0.2644$) corresponds to a scenario of moving from the coasts inland, while the second model ($wBIC_a=0.2122$) focuses on transecting the interior. Including the model with the third-most explanatory power ($wBIC_a=0.0984$), which has routes connecting known water sources and brings total explanatory power up to 0.575, our models suggest a strategy of first moving inland, then discovering water sources in the interior, then travelling among them, all while navigating by prominent landscape features. The super-highways that emerge from our models (Fig. 3) are both the least-calorically costly and most ‘attractive’ paths, while also corresponding most closely to known early archaeological sites. This concordance demonstrates that while there were a range of routes available to facilitate the rapid and efficient movement of people throughout the continent, people likely chose the most optimal visually, calorically, and hydrologically most frequently. These suggest fundamental rules guiding human movement—people orient themselves using visual cues, they search for freshwater, and they minimise caloric expenditure as much as possible as they enter new regions.

Our resulting network explicitly identifies vectors of connectivity among regions that have so far only been qualitatively inferred from linguistic^{57,58}, genetic^{59,60}, or archaeological^{24,27,28} data (although see recent work by Pedro et al.⁶¹ that suggested vectors of connectivity); our findings are complimentary to these studies. Recent genetic mapping of O, S, P and M haplotypes suggest that people moved down the eastern and western coastlines before joining in the Australian Bight and moving into the interior^{13,61}. Conversely, linguistic mapping—acknowledging that this is constrained to the Holocene—suggests a general transmission from the northeast to the southwest. Our results broadly support these findings, while also adding additional nuance and detail across Sahul during this time. Few of the previously published studies identified the importance of a major corridor through the arid centre, nor

for many of the lesser, but still probable, routes running through the semi-arid and channel country of inner south-western Queensland.

Several of the major routes our models identified that traverse central Australia are echoed by well-documented ethnographic trade routes, including Cape York to South Australia via Birdsville in the trade of pituri native tobacco (*Duboisia hopwoodii*), greenstone axes, Flinders Ranges greenstones, and baler shell^{62,63}, and Kimberley to Eyre Peninsula via the trade of pearl and baler shell^{64,65}. While these ethnographic trade routes are considerably younger than the period we modelled, the main edaphic features dictating the network of optimal pathways would have remained largely intact even over tens of millennia, suggesting extraordinary persistence of these routes as primary corridors of movement. As people expanded across the continent and distinct cultural and linguistic groups formed, the deeply established connections across these physical networks would have facilitated the movement of people, objects, and stories that led to the rich and highly interconnected Aboriginal, Torres Strait Islander, and Melanesian cultures encountered at European contact.

A limitation of this study is that our methods intentionally avoid considering many types of potential biases, including concerns of demographic composition or traveller velocity, the types of ecosystems encountered, and perhaps most controversially, cultural influence. However, it is telling that the similar pathways between the period of first peopling of the continent and late Holocene trade networks suggest the underlying environmental variables structuring pathway persistence likely influenced socio-economic and ritual practices. The facilitation of movement that these super-highways provided would have enabled efficient spatial use of the continent.

Our models also provide understanding of the underlying baseline landscape structure for the most plausible migratory pathways across the entire continent of Sahul. In turn, our results link with previous work predicting the most-likely pathways of entry into Sahul from Wallacea^{22,24,27}. Basic behavioural and landscape features enable the calculation of ‘cost’ of movement, which provide a foundation for understanding how different demographic processes or dispersal capabilities would have influenced these early humans’ relationship with the landscape. For example, we can now use these pathways in agent-based or other dynamic models to examine how groups might have progressed through the landscape with demographic and other environmental constraints⁵. Our methods also are applicable to other migratory events beyond Sahul, such as the peopling of the Americas or the waves of Out-of-Africa migration. For example, combining maritime path models⁶⁶ with land-based models developed using the *From Everywhere To Everywhere* approach (see Methods) could identify

links from Alaska⁶⁷ to South America^{68,69}. While these models would not definitively solve the *Clovis First* controversy⁷⁰, the ability to detect pathways would be an important step in this regard.

The routes we identified also suggest potential target areas for archaeological prospection by pinpointing the regions that humans would have frequently traversed and where archaeological material would have likely accumulated⁷¹. Extant terrestrial locations along these identified corridors therefore should become priority locations to search for new archaeological data. We recommend that future archaeological prospection focus on where modern shores intersect major ancient corridors, and along the identified super-highways transecting the continent. Many of the predicted corridors are now located on the flooded continental shelf, making prospection of these routes challenging, although not impossible with the use of modern technologies^{66,72,73}. Combining our approach with maritime path models^{22,74}, we might be able to predict where migrations likely occurred, helping focus exploration into those regions.

The ways that people migrate have not changed in the hundreds of generations since the first people entered Sahul. The cognitive capacity of these early explorers and their behavioural plasticity enabled the rapid expansion across a diverse range of arid, tropical and temperate environments. The consistency of these abilities and behaviours have implications for predicting future migrations. As climate disruption impacts the liveability of many parts of the globe⁴⁶, we will be confronted with large-scale migration events as people escape drowning coastlines, wildfires, violence, or drought. Predicting where populations are likely to move will be crucial for international planning and coordination. The methods we employ can be used to examine these issues, both within Australia and other landmasses across the globe.

References (www.mendeley.com/community/fete-peopling-of-sahul)

1. Laland, K. N. & O'Brien, M. J. Niche construction theory and archaeology. *J. Archaeol. Method Theory* **17**, 303–322 (2010).
2. Dunne, J. A. *et al.* The roles and impacts of human hunter-gatherers in North Pacific marine food webs. *Sci. Rep.* **6**, 21179 (2016).
3. Crabtree, S. A., Bird, D. W. & Bird, R. B. Subsistence transitions and the simplification of ecological networks. *Nat. Hum. Behav.* 165–177 (2019).
4. Crabtree, S. A., Vaughn, L. J. S. & Crabtree, N. T. Reconstructing Ancestral Pueblo food webs in the southwestern United States. *J. Archaeol. Sci.* **81**, 116–127 (2017).
5. Romanowska, I., Gamble, C., Bullock, S. & Sturt, F. Dispersal and the Movius Line: Testing the effect of dispersal on population density through simulation. *Quat. Int.* **431**, 53–63 (2017).
6. Grothmann, T. & Patt, A. Adaptive capacity and human cognition: The process of individual adaptation to climate change. *Glob. Environ. Chang.* **15**, 199–213 (2005).
7. Potts, R. Variability selection in hominid evolution. *Evol. Anthropol. Issues, News, Rev.* **7**, 81–96 (1998).
8. Thomsen, C. J. *Ledetraad Til Nordisk Oldkyndighed. Kjöbenhavn.* (S.L. Møllers bogtr, 1836).
9. Montelius, O. *Der Orient und Europa.* (Königliche Akademie der schönen Wissenschaften, Geschichte und Altertumskunde, 1899).
10. Hershkovitz, I. *et al.* The earliest modern humans outside Africa. *Science (80-)*. **359**, 456 LP – 459 (2018).
11. Braje, T. J., Dillehay, T. D., Erlandson, J. M., Klein, R. G. & Rick, T. C. Finding the first Americans. *Science (80-)*. **358**, 592 LP – 594 (2017).
12. Skoglund, P. *et al.* Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510–513 (2016).
13. Tobler, R. *et al.* Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544**, 180–184 (2017).
14. Nielsen, R. *et al.* Tracing the peopling of the world through genomics. *Nature* **541**, 302–310 (2017).
15. Crema, E. R., Habu, J., Kobayashi, K. & Madella, M. Summed Probability Distribution of 14C Dates Suggests Regional Divergences in the Population Dynamics of the Jomon Period in Eastern Japan. *PLoS One* **11**, 1–18 (2016).
16. Antón, S. C., Leonard, W. R. & Robertson, M. L. An ecomorphological model of the initial hominid dispersal from Africa. *J. Hum. Evol.* **43**, 773–785 (2002).
17. Breeze, P. S. *et al.* Palaeohydrological corridors for hominin dispersals in the Middle East ~250–70,000 years ago. *Quat. Sci. Rev.* **144**, 155–185 (2016).
18. Hughes, J. K., Haywood, A., Mithen, S. J., Sellwood, B. W. & Valdes, P. J. Investigating early hominin dispersal patterns: developing a framework for climate data integration. *J. Hum. Evol.* **53**, 465–474 (2007).
19. R. Vahdati, A., Weissmann, J. D., Timmermann, A., Ponce de León, M. S. & Zollikofer, C. P. E. Drivers of Late Pleistocene human survival and dispersal: an agent-based modeling and machine learning approach. *Quat. Sci. Rev.* **221**, 105867 (2019).
20. White, D. A. & Barber, S. B. Geospatial modeling of pedestrian transportation networks: a case study from precolumbian Oaxaca, Mexico. *J. Archaeol. Sci.* **39**, 2684–2696 (2012).
21. Binford, L. R. *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data.*

- (University of California Press, 2001).
22. Kealy, S., Louys, J. & O'Connor, S. Least-cost pathway models indicate northern human dispersal from Sunda to Sahul. *J. Hum. Evol.* **125**, 59–70 (2018).
 23. Kealy, S., Louys, J. & O'Connor, S. Islands under the sea: a review of early modern human dispersal routes and migration hypotheses through Wallacea. *J. Isl. Coast. Archaeol.* **11**, 364–384 (2016).
 24. Bird, M. I. *et al.* Palaeogeography and voyage modeling indicates early human colonization of Australia was likely from Timor-Roti. *Quat. Sci. Rev.* **191**, 431–439 (2018).
 25. Kealy, S., Louys, J. & O'Connor, S. Reconstructing palaeogeography and inter-island visibility in the Wallacean Archipelago during the likely period of Sahul colonization, 65–45 000 years ago. *Archaeol. Prospect.* **24**, 259–272 (2017).
 26. Norman, K. *et al.* An early colonisation pathway into northwest Australia 70–60,000 years ago. *Quat. Sci. Rev.* **180**, 229–239 (2018).
 27. Bird, M. I. *et al.* Early human settlement of Sahul was not an accident. *Sci. Rep.* **9**, 8220 (2019).
 28. O'Connell, J. F. & Allen, J. The restaurant at the end of the universe: modelling the colonisation of Sahul. *Aust. Archaeol.* **74**, 5–31 (2012).
 29. Birdsell, J. B. Some population problems involving Pleistocene man. in *Cold Spring Harbor Symposia on Quantitative Biology* 47–69 (Cold Spring Harbor Lab Press, 1957). doi:10.1101/sqb.1957.022.01.008
 30. Bowdler, S. The coastal colonisation of Australia. in *Sunda and Sahul: Prehistoric Studies in Southeast Asia Melanesia and Australia* (eds. Allen, J., Golson, J. & Jones, R.) 205–246 (Academic Press, 1977).
 31. Horton, D. R. Water and woodland: the peopling of Australia. *Aust. Inst. Aborig. Stud. Newsl.* **16**, 21–27 (1981).
 32. Tindale, N. B. Prehistory of the Aborigines: some interesting considerations. in *Ecological Biogeography of Australia* (ed. Keast, A.) 1761–1797 (Dr W Junk Publishers, 1981). doi:10.1007/978-94-009-8629-9_63
 33. Veth, P. Islands in the interior: a model for the colonization of Australia's arid zone. *Archaeol. Ocean.* **24**, 81–92 (1989).
 34. Veth, P. *Islands in the interior: the dynamics of prehistoric adaptations within the arid zone of Australia*. (International Monographs in Prehistory, 1993).
 35. Hiscock, P. & Wallis, L. A. Pleistocene settlement of deserts from an Australian perspective. in *Desert Peoples* (eds. Veth, P., Smith, M. & Hiscock, P.) 34–57 (Blackwell Publishing Ltd, 2008). doi:10.1002/9780470774632.ch3
 36. Veth, P., O'Connor, S. & Wallis, L. A. Perspectives on ecological approaches in Australian archaeology. *Aust. Archaeol.* **50**, 54–66 (2000).
 37. Hallam, S. J. The relevance of Old World archaeology to the first entry of man into new worlds: colonization seen from the Antipodes. *Quat. Res.* **8**, 128–148 (1977).
 38. Lourandos, H. & Ross, A. The great 'intensification debate': its history and place In Australian archaeology. *Aust. Archaeol.* **39**, 54–63 (1994).
 39. O'Connell, J. F. & Allen, J. The process, biotic impact, and global implications of the human colonization of Sahul about 47,000 years ago. *J. Archaeol. Sci.* **56**, 73–84 (2015).
 40. Bird, M. I., O'Grady, D. & Ulm, S. Humans, water, and the colonization of Australia. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11477–11482 (2016).
 41. Brown, C. T., Liebovitch, L. S. & Glendon, R. Lévy Flights in Dobe Ju/'hoansi Foraging Patterns. *Hum. Ecol.* **35**, 129–138 (2007).
 42. Hamilton, M. J., Milne, B. T., Walker, R. S. & Brown, J. H. Nonlinear scaling of space

- use in human hunter–gatherers. *Proc. Natl. Acad. Sci.* **104**, 4765 LP – 4769 (2007).
43. Hamilton, M. J., Davidson, A. D., Sibly, R. M. & Brown, J. H. Universal scaling of production rates across mammalian lineages. *Proc. R. Soc. B Biol. Sci.* **278**, 560–566 (2011).
 44. Lock, G. & Pouncett, J. Walking the Ridgeway Revisited: The Methodological and Theoretical Implications of Scale Dependency for the Derivation of Slope and the Calculation of Least-Cost Pathways. in *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference* (eds. Frischer, B., Webb Crawford, J. & Koller, D.) 192–203 (Archaeopress, 2010).
 45. Thomas, Z. A. *et al.* Tipping elements and amplified polar warming during the Last Interglacial. *Quat. Sci. Rev.* **233**, 106222 (2020).
 46. Xu, C., Kohler, T. A., Lenton, T. M., Svenning, J.-C. & Scheffer, M. Future of the Human Climate Niche. *Proc. Natl. Acad. Sci.* 1–6 (2020). doi:10.1073/pnas.1910114117
 47. Bradshaw, C. J. A. *et al.* Minimum founding populations for the first peopling of Sahul. *Nat. Ecol. Evol.* **3**, 1057–1063 (2019).
 48. Bowler, J. M. *et al.* New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature* **421**, 837–840 (2003).
 49. Fitzsimmons, K. E., Stern, N. & Murray-Wallace, C. V. Depositional history and archaeology of the central Lake Mungo lunette, Willandra Lakes, southeast Australia. *J. Archaeol. Sci.* **41**, 349–364 (2014).
 50. Kemp, J., Pietsch, T., Gontz, A. & Olley, J. Lacustrine-fluvial interactions in Australia’s Riverine Plains. *Quat. Sci. Rev.* **166**, 352–362 (2017).
 51. Mueller, N. *et al.* Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.* **174**, 341–352 (2016).
 52. Tindale, N. B. Ecology of Primitive Aboriginal Man in Australia BT - Biogeography and Ecology in Australia. in (eds. Keast, A., Crocker, R. L. & Christian, C. S.) 36–51 (Springer Netherlands, 1959). doi:10.1007/978-94-017-6295-3_3
 53. Bowler, J. M., Jones, R., Allen, H. & Thorne, A. G. Pleistocene human remains from Australia: A living site and human cremation from Lake Mungo, western New South Wales. *World Archaeol.* **2**, 39–60 (1970).
 54. Jones, R. The Geographical Background to the Arrival of Man in Australia and Tasmania. *Archaeol. Phys. Anthropol. Ocean.* **3**, 186–215 (1968).
 55. Reynolds, H. The land, the explorers and the aborigines. *Hist. Stud.* **19**, 213–226 (1980).
 56. Bellavia, G. Extracting “Natural Pathways” from a Digital Elevation Model. in *Archaeological Informatics: Pushing The Envelope. Proceedings of the CAA 2001* (eds. Burenhult, G. & Arvidsson, J.) 5–12 (Archaeopress, 2002).
 57. Clendon, M. Reassessing Australia’s Linguistic Prehistory. *Curr. Anthropol.* **47**, 39–61 (2006).
 58. McConvell, P. The Linguistic Prehistory Of Australia: Opportunities For Dialogue With Archaeology. *Aust. Archaeol.* **31**, 3–27 (1990).
 59. Redd, A. J. & Stoneking, M. Peopling of Sahul: mtDNA Variation in Aboriginal Australian and Papua New Guinean Populations. *Am. J. Hum. Genet.* **65**, 808–828 (1999).
 60. van Holst Pellekaan, S. Genetic evidence for the colonization of Australia. *Quat. Int.* **285**, 44–56 (2013).
 61. Pedro, N. *et al.* Papuan mitochondrial genomes and the settlement of Sahul. *J. Hum. Genet.* (2020). doi:10.1038/s10038-020-0781-3

62. McBryde, I. Goods from Another Country: Exchange Networks and the People of the Lake Eyre Basin. in *Australians to 1788* (eds. Mulvaney, D. J. & White, J. P.) 252–273 (Academy of the Social Sciences in Australia, 1987).
63. McAdam, L. & Davidson, I. Beads and boundaries. in *The Archaeology of Portable Art Southeast Asian, Pacific, and Australian Perspectives* (eds. Langley, M., Litster, M., Wright, D. & May, S. K.) (2016).
64. Smith, M. A. & Veth, P. M. Radiocarbon Dates For Baler Shell In The Great Sandy Desert. *Aust. Archaeol.* **58**, 37–38 (2004).
65. Robertson, S. Significance Of Baler Shell(Melo) At Olympic Dam, South Australia. *J. Anthropol. Soc. South Aust.* **36**, (2012).
66. Maloney, J. *et al.* Refining search areas for submerged archaeological resources using subbottom data and applied geomorphology. in *Ocean Sciences Meeting* (2020).
67. Flegontov, P. *et al.* Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. *Nature* **570**, 236–240 (2019).
68. Vialou, D., Benabdelhadi, M., Feathers, J., Fontugne, M. & Vialou, A. V. Peopling South America's centre: the late Pleistocene site of Santa Elina. *Antiquity* **91**, 865–884 (2017).
69. Huguin, R., Franco, N. V & Flegenheimer, N. Humans, Technology, and Environment in the Early Peopling of South America. *PaleoAmerica* **5**, 307–308 (2019).
70. Klein, H. S. The First Americans: The Current Debate. *J. Interdiscip. Hist.* **46**, 543–561 (2016).
71. Binford, L. R. The archaeology of place. *J. Anthropol. Archaeol.* **1**, 5–31 (1982).
72. Benjamin, J. *et al.* Submerged Aboriginal artefacts on the continental shelf reveal ancient drowned cultural landscapes in northwest Australia. *PLoS One* **15**, (2020).
73. King, R. *et al.* Controlled-Source Electromagnetic Methods to Investigate the Submerged Cultural Landscapes of the Pacific Continental Shelf. in *Ocean Sciences Meeting* (2020).
74. Gustas, R. & Supernant, K. Least cost path analysis of early maritime movement on the Pacific Northwest Coast. *J. Archaeol. Sci.* **78**, 40–56 (2017).
75. Whiteway, T. G. *Australian Bathymetry and Topography Grid, Report No. Geoscience Australia Record 2009/21*, 46. (2009).
76. Beaman, R. J. *3DGBR: A high-resolution depth model for the Great Barrier Reef and Coral Sea. Report No. Marine and Tropical Sciences Research Facility (MTSRF) Project 2.5i.1a Final Report, 13 plus Appendix 11.* (2010).
77. Beaman, R. J. *High-resolution depth model for Northern Australia - 100 m.* (2018).
78. Group, G. C. *GEBCO Grid.* (2019).
79. Daniell, J. J. Development of a bathymetric grid for the Gulf of Papua and adjacent areas: A note describing its development. *J. Geophys. Res. Earth Surf.* **113**, 1–8 (2008).
80. Tickle, P. *Digital Elevation Models User Guide: 1 second DSM, DEM & DEM-S; 3 second DSM, DEM & DEM-S. Report No. Version 1.0.3*, 84. (Geoscience Australia, 2010).
81. Jarvis, A., Guevara, E., Reuter, H. I. & Nelson, A. D. Hole-filled SRTM for the globe : version 4 : data grid . (2008).
82. Whiteway, T. G. *Australian Bathymetry and Topography Grid. Record 2009/21* (Geosci- Geoscience Australia, 2009).
83. Beaman, R. J. *A high-resolution depth model for the Great Barrier Reef and Coral Sea. Report No. Marine and Tropical Sciences Research Facility (MTSRF) Project 2.5i.1a Final Report, 13 plus Appendix 11.* (Reef and Rainforest Research Centre (RRRC), 2010).

84. Beaman, R. J. *High-resolution depth model for Northern Australia - 100 m.* (Geosci-Geoscience Australia, 2018).
85. Daniell, J. J. Development of a bathymetric grid for the Gulf of Papua and adjacent areas: A note describing its development. *J. Geophys. Res. Earth Surf.* **113**, (2008).
86. Dungan, K. A., White, D., Déderix, S., Mills, B. J. & Safi, K. A total viewshed approach to local visibility in the Chaco World. *Antiquity* **92**, 905–921 (2018).
87. Bliege Bird, R. *et al.* Fire mosaics and habitat choice in nomadic foragers. *Proc. Natl. Acad. Sci.* **117**, 12904 LP – 12914 (2020).
88. White, D. A. & Barber, S. B. Geospatial modeling of pedestrian transportation networks: a case study from precolumbian Oaxaca, Mexico. *J. Archaeol. Sci.* **39**, 2684–2696 (2012).
89. Pandolf, K. B., Givoni, B. & Goldman, R. F. Predicting energy expenditure with loads while standing or walking very slowly. *J. Appl. Physiol.* **43**, 577–581 (1977).
90. Looney, D. P. *et al.* Metabolic Costs of Military Load Carriage over Complex Terrain. *Mil. Med.* **183**, e357–e362 (2018).
91. Soule, R. G. & Goldman, R. F. Terrain coefficients for energy cost prediction. *J. Appl. Physiol.* **32**, 706–708 (1972).
92. Santee, W., Blanchard, L., Speckman, K., Gonzalez, J. & Wallace, R. Load Carriage Model Development and Testing with Field Data. Technical Note. *Natick, MA, U.S. Army Res. Inst. Environ. Med.* (2003).
93. Wood, B. M. & Wood, Z. J. Energetically optimal travel across terrain: visualizations and a new metric of geographic distance with anthropological applications. in *Proc.SPIE* **6060**, (2006).
94. Tobler, W. Three Presentations on Geographical. *Natl. Cent. Geogr. Inf. Anal.* (1993).
95. Mifflin, M. D. *et al.* A new predictive equation for resting energy expenditure in healthy individuals. *Am. J. Clin. Nutr.* **51**, 241–247 (1990).
96. Frankenfield, D., Roth-Yousey, L. & Compher, C. Comparison of Predictive Equations for Resting Metabolic Rate in Healthy Nonobese and Obese Adults: A Systematic Review. *J. Am. Diet. Assoc.* **105**, 775–789 (2005).
97. Williams, A. N., Ulm, S., Smith, M. & Reid, J. AustArch: A database of 14C and non-14C ages from archaeological sites in Australia - Composition, compilation and review. *Internet Archaeol.* **36**, (2014).
98. Rodríguez-Rey, M. *et al.* Criteria for assessing the quality of Middle Pleistocene to Holocene vertebrate fossil ages. *Quat. Geochronol.* **30**, 69–79 (2015).
99. Bird, M. I. *et al.* The efficiency of charcoal decontamination for radiocarbon dating by three pre-treatments – ABOX, ABA and hypy. *Quat. Geochronol.* **22**, 25–32 (2014).
100. Alex, B. *et al.* Radiocarbon chronology of Manot Cave, Israel and Upper Paleolithic dispersals. *Sci. Adv.* **3**, e1701450 (2017).
101. David, B. *et al.* 45,610–52,160 years of site and landscape occupation at Nawarla Gabarnmang, Arnhem Land plateau (northern Australia). *Quat. Sci. Rev.* **215**, 64–85 (2019).
102. Wood, R. *et al.* Towards an accurate and precise chronology for the colonization of Australia: The example of Riwi, Kimberley, Western Australia. *PLoS One* **11**, e0160123 (2016).
103. Kiskowski, M. A., Hancock, J. F. & Kenworthy, A. K. On the use of Ripley’s K-function and its derivatives to analyze domain size. *Biophys. J.* **97**, 1095–1103 (2009).
104. Kolmogorov, A. Sulla determinazione empirica di una legge di distribuzione, 1st. *Ital. Attuari. G.* **4**, 1–11 (1933).
105. Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *Ann. Math. Stat.* **19**, 279–281 (1948).

106. Besag, J. Efficiency of pseudolikelihood estimation for simple Gaussian fields. *Biometrika* **64**, 616–618 (1977).

Acknowledgements

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This research was done by the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage (CE170100015) and the ASU-SFI Center for Biosocial Complex Systems. M.I.B. is the recipient of an Australian Research Council Laureate Fellowship (FL140100044). We thank the Santa Fe Institute for support in beginning this research. For assistance in developing the basis of the paper, we thank K. Norman, F. Petchey, M. Price, S. Slater, C. McGuire and R. Wood. We thank B. Roberts for comments on the paper, and three anonymous reviewers.

Author Contributions S.A.C., S.U., M.I.B. and C.J.A.B. obtained funding from the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage and the ASU-SFI Center for Biosocial Complex Systems. D.A.W. and S.A.C. developed the *From Everywhere To Everywhere* models. D.A.W., S.A.C., and F.S. did statistical evaluations. R.J.B. compiled seamless terrestrial and bathymetric digital elevation model data. S.A.C., D.A.W., S.U., A.N.W, C.J.A.B., F.S. and M.I.B. wrote the main text with specialist contributions from other authors. S.A.C. and D.A.W. contributed equally; S.U. and M.I.B. contributed equally; C.J.A.B. and F.S. contributed equally. All authors provided comments and revisions.

Competing Interests. The authors declare no competing interests.

Methods

Developing a digital elevation model

Our Sahul digital elevation model is an amalgamation of existing regional- and global-scale digital elevation models. The primary bathymetry data used the *ausbathytopo* grid⁷⁵ for Australia, with a pixel size of 0.0025° (~ 250 m) and clipped to $110\text{--}156^\circ$ E, $8\text{--}46^\circ$ S. Additional bathymetric data for the Australian region included the *gbr100* grid⁷⁶ with a pixel size of 0.001° (~ 100 m), clipped to $142\text{--}156^\circ$ E, $10\text{--}29^\circ$ S, and the *nthaus100* grid⁷⁷ with a pixel size of 0.001° (~ 100 m) and an area of $121\text{--}133^\circ$ E, $8\text{--}18^\circ$ S. Bathymetric data for areas north of Australia toward New Guinea used the latest GEBCO grid⁷⁸ with a pixel size of 15 arc-seconds (~ 416 m) and clipped to $110\text{--}156^\circ$ E, $5^\circ\text{--}8.2^\circ$ S. A small area of Gulf of Papua bathymetry used the *gop100* grid⁷⁹, with a pixel size of 0.001 (~ 100 m) and clipped to $142.6\text{--}146.7^\circ$ E, $7\text{--}8.2^\circ$ S. The primary land-elevation data used the 3 arc-second (~ 83 m) Shuttle Radar Topographic Mission (SRTM) digital surface model for Australia⁸⁰. We defined the coastline as the -85 m isobath²⁷. Additional land elevation north of Australia used the 3 arc-second (~ 83 m) SRTM data sourced from the Consortium of International Agricultural Research Centres–Consortium for Spatial Information⁸¹ and clipped to $110\text{--}155^\circ$ E, $5\text{--}15^\circ$ S. We removed anomalous land elevation < 0 m height in the SRTM datasets using the ArcGIS *SetNull* tool. We created an ArcGIS mosaic dataset with all bathymetry and elevation source grids added and the *ZOrder* changed to reflect the priority viewing of datasets: Australia SRTM⁸², other SRTM⁸³, *gbr100*⁸⁴, *nthaus100* (GEBCO Compilation Group, 2019), *ausbathytopo*⁸⁵, *gop100*⁸⁰, GEBCO⁸¹. We then applied the ArcGIS *Resample* tool to generate a compiled digital elevation model with a pixel size of 0.0025° (~ 250 m), which corresponds to the coarsest resolution of the above datasets, and clipped to the final grid area of $111\text{--}155^\circ$ E, $3\text{--}45^\circ$ S. This provides the best resolution DEM possible that can then be used for future products.

Model inputs

We based our construction of theoretical pedestrian-transportation networks on the Sahul digital elevation model described above (see *Developing a digital elevation model*) merged with a reprocessed version of the Shuttle Radar Topography Mission digital surface model data that was one of the sources used during its creation. The reason for doing so was to extract the most accurate topographic information possible across the entire region, which required processing modern land surfaces and their bathymetric counterparts separately due

to how the best available spatial resolution for each was quite different (30 m per cell for land, 250 m per cell for bathymetric). This was especially important for identifying and preserving small, but visually prominent, landscape features that are missing in the Sahul model, which used the SRTM model in a very different way. As part of the processing, both models were independently resampled twice to a common coarser resolution of 500 m, where cells in the first represented the average height of all contributing data from the original products, and the cells in the second represented the absolute maximum height. We selected a cell size of 500 m for analysis because it was computationally tractable, i.e., the original elevation model (at 250 m) was simply too large to process in a reasonable amount of time—even on a supercomputer. The resampling techniques used preserved critical information from the original model that we discuss in more detail below. We then merged those resampled surfaces to create two new mean- and maximum-height surfaces for Sahul, applying a 3×3 grid cell smoothing kernel along the merge boundaries to mitigate artefacts such as prominent landscape features created by coastline erosion that might negatively impact subsequent processing. We used the mean- and maximum-height surfaces as direct inputs to the transportation and aggregate viewshed-modelling applications discussed below, respectively. We also masked out ancient lakes in the mean-height dataset to prevent travellers from walking over them, and to treat those regions as filled with water in the maximum-height dataset to reflect actual visibility. Leveraging accepted elevation values for known lakes, we created a continuous contour for each one to either create impassable (mean) or flat (maximum) regions in the surfaces.

We used both height surfaces to create what we call a ‘visual prominence’ weighting layer that represents how close each elevation cell is to the most noticeable landscape features across the entire continent, where closeness is based on three-dimensional walking distance. We first processed the maximum-height surface, which preserves the most contrast in vertical information from the higher-resolution source elevation model and facilitates identifying prominent points, using an aggregate viewshed application that is accelerated by a graphical processing unit to shorten runtime, resulting in a visibility-frequency map whose cell-value distribution closely resembles a power law; this is the first time such an application has been applied at a continental scale⁸⁶. We then log-transformed the data to approximate a normal distribution, and selected cells whose values were ≥ 3 standard deviations (σ) from the mean as the most visible locations. We selected the 3σ threshold after generating location maps at multiple thresholds (from 1σ to 6σ) and empirically evaluating the results. Fewer multiples of σ led to a large number of locations, and more multiples of σ resulted in few locations. Both

would have provided little-to-no value because they would produce roughly equal weights everywhere, because a traveller would either never be far away from one of them or almost never close to one of them.

Next, we supplied those locations as a binary raster to a least-cost surface-generator application that accumulates three-dimensional travel distance, along with the mean height surface, which is smoother than the maximum-height surface and facilitates more realistic estimates of travel distance, to produce a cost-distance surface that we then normalised to create the final weighting layer. We treated each cell with a value of 1 in the binary raster as an ‘origin’ location for travel, and ignored cells with 0, which is a standard way to provide origins to least-cost surface-generator applications. While we used a custom application (separate from *FETE*) due to its computational efficiency, both commercial and open-source GIS software packages contain similar capabilities, e.g., *Path Distance* in *ArcMap* and *r.walk* in *GRASS*. Once normalised, each cell communicates how attractive it is with respect to being physically proximal to noticeable features, meaning that they act as potential magnets for travellers depending on their values. A value closer to zero mean that a cell is within a short walking distance of a noticeable feature and a value closer to one is much farther away, but this is with respect to the entire continent, so “close” is a relative term.

We processed a binary raster representing the locations of all streams at Strahler Stream Order 9 or higher in a similar fashion as the visibility layer to create the second normalised weighting layer. Briefly, we used standard GIS hydrological flow-modelling tools in conjunction with the original Sahul digital elevation model to create the stream network, which represents the likely locations and extents of major rivers and their courses to the ancient coast during the time of initial peopling of the continent. We selected Order 9 as the threshold after empirically evaluating the results of selecting higher and lower values. Much like with visual prominence, values $<$ and $>$ 9 produced either far too many or far too few locations, respectively, to be useful. In this case, Order 9 represents all major streams and rivers potentially present on the continent during the period of interest. Using empirically ranked streams also removes researcher bias in deciding which streams are ‘important’, allowing for the most parsimonious model to emerge.

Modelling travel

We created theoretical pedestrian-transportation networks using the *From Everywhere To Everywhere* (FETE) modelling application to generate many candidate least-cost routes.

FETE leverages elevation data, optional land cover (not employed here due to annual changes in land cover⁸⁷), an optional weighting layer (developed for use in this study), physical traveller characteristics (age, height, weight, sex, additional load carried) to identify the least calorically costly walking paths, and sets of origin and destination points⁸⁸. We needed both general and specific sets of origins and destinations for travel to test several hypotheses. For the general sets of points, we created a regularly spaced grid with spacing selected as one point every 50 surface cells in each direction, which was empirically determined to represent a reasonable balance between point density and computational cost. We did test runs with spacings that ranged from 10 to 100, with the resulting normalised outputs not containing enough new information at densities > 50 to justify the substantial computational resources required to create them (38,380 points total). For the specific sets of points, we converted the vector shoreline associated with the mean-height surface into a series of individual points (one per relevant surface cell; 94,332 points total), and also converted the *Water Observations from Space*^{40,51} dataset, which maps the presence/absence of surface water, into a series of individual points (variable point density based on location; 112,786 points total). This allowed for several hundred thousand potential starting and ending points for travel across all modelled scenarios, which used the individual sets as origins or destinations (e.g., grid-to-grid, grid-to-coastline, grid-to-water), resulting in billions of pathways. Given that the New Guinea portion of the Sahul landmass is substantially smaller than the Australia portion, there are many more points in the latter that can be used as origins and destinations for travel. This means that there is a higher likelihood of creating the kinds of ‘super-highways’ in that region that are the focus of this study. While that algorithmic behaviour aligns well with our aims to examine movement into Australia, it also highlights an opportunity to look at New Guinea more closely on its own in a future study.

While the full methodology is described in White and Barber⁸⁸, we present the main elements of *FETE*’s per-cell movement cost-estimation process here. *FETE* ingests either user-supplied sets of origin and destination points to explore specific connectivity questions, or can automatically generate regularly spaced grids of origin and destination points at user-specific densities to explore more general movement trends. The locations of calculated routes are mapped onto a regularly spaced raster grid and accumulated to produce a frequency-based surface where each cell indicates how likely it is to be used for travel with respect to the supplied origins and destinations. We then applied a threshold frequency to create networks of routes representing corridors for movement. *FETE*’s anisotropic cost-minimisation function relies on well-established caloric expenditure^{89–93} and travel time⁹⁴

equations for pedestrians (Equations 1–7 below). Traditional GIS software can be used to create similar outputs, but least-cost route generation—especially for large areas, many origin points, and many travel scenarios—is computationally expensive. *FETE* is specifically designed to address those challenges efficiently, enabling the calculation of tens of billions of routes for the entire continent of Sahul each time it is run, compared to tens or perhaps hundreds that are generally created for smaller areas (Supplemental Table 9).

When estimating the energetic cost of travel from one grid cell to another, *FETE* first calculates a generalized estimate of walking speed for the traveller in m hr^{-1} , v_e , using an appropriately scaled version of Tobler’s hiking function⁹⁴:

$$v_e = 1000 \cdot (6e^{-3.5|s_f+0.05|}) \quad \text{[Equation 1]}$$

where s_f is fractional slope (rise over run) of the terrain between the two grid cells. Since the traveller is potentially walking on two different surface types as she moves from one grid cell to another, a metabolic rate of travel (MRT) is calculated for each surface type using one of two equations:

$$MRT = 1.5w + 2\left(\frac{l}{w}\right)^2 + \eta(w+l)(1.5v_e^2 + 0.35v_e s_p) \quad \text{[Equation 2]}$$

$$MRT = 1.5w + 2\left(\frac{l}{w}\right)^2 + \eta(w+l)(1.5v_e^2 + 0.35v_e s_p) - \eta\left(\frac{(v_e s_p (w+l))}{3.5} - \frac{(w+l)(s_p+6)^2}{w} + (25 - v_e^2)\right) \quad \text{[Equation 3]}$$

where w is the weight of the traveller in kg, l is the optional load carried by the traveller in kg, s_p is % slope (converted from s_f), and η is a unitless terrain coefficient associated with surface type⁹¹. *FETE*’s selection of Equation 2 or Equation 3 is driven by the slope present between the two cells. For uphill or flat surfaces (fractional slope ≥ 0), Equation 2 is used⁸⁹. For downhill surfaces (fractional slope < 0), Equation 3 is used^{92,93}. In this way, *From Everywhere to Everywhere*’s cost estimation is anisotropic, or direction-dependent, because the cost of travelling uphill from point A to point B (positive slope or no slope) is different than the cost of travelling downhill from Point B to Point A (negative slope). This is clearly seen in the asymmetric cost output from Tobler’s hiking function on its own, but is also reflected in Equation 3, which includes an empirically derived nonlinear corrective term that is associated only with downhill movement.

The two calculated MRTs, MRT for surface type in the origin cell (η_1) and MRT for surface type in the destination cell (η_2), are then combined, converted from watts (J s^{-1}) to kcal s^{-1} via multiplication by a conversion factor ($1 \text{ W} = 0.000239 \text{ kcal s}^{-1}$), and used in combination with the travel velocity and travel distance (the length is split between both calculated MRTs) to calculate the energy expended during travel, kcal_t :

$$\text{kcal}_t = \frac{0.000239 \cdot (\text{MRT}_{\eta_1} + \text{MRT}_{\eta_2}) \cdot (0.5d_t)}{v_e} \quad [\text{Equation 4}]$$

where d_t is the three-dimensional travel distance in m. Estimates of kcal_t at slow speeds on downhill slopes sometimes under-predict actual expenditure⁹³, so a sex-specific alternative estimate, kcal_a , is calculated using the Mifflin St. Jeor basal metabolic rate equation^{95,96} and a scale factor:

$$\text{SMR}_{\text{male}} = 1.2 \cdot (5.0 + 10.0w + 6.5h - 5.0a) \quad [\text{Equation 5}]$$

$$\text{SMR}_{\text{female}} = 1.2 \cdot (-161.0 + 10.0w + 6.5h - 5.0a) \quad [\text{Equation 6}]$$

$$\text{kcal}_a = \frac{0.000239 \cdot \text{SMR} \cdot d_t}{v_e} \quad [\text{Equation 7}]$$

where h = the height of the traveller in cm and a = the age of the traveller in years. *FETE* selects the larger of kcal_t or kcal_a as the final energetic cost estimate, kcal_f . If a normalised weighting layer is supplied, the final cost estimate is scaled by the weight value present in the destination cell prior to use:

$$\text{kcal}_{fw} = \text{weight} \cdot \text{kcal}_f \quad [\text{Equation 8}]$$

A normalised weighting value of 1.0 for a destination cell results in no change to the travel cost associated with reaching that cell (a smaller value will reduce the cost and thus make travel to the cell more attractive). With respect to the two weighting layers that we used, smaller weights were assigned to cells in closer spatial proximity to rivers and visually prominent locations, and larger weights were assigned to more distant ones. We treated each weighting layer equally and used them independently to support the scenarios of visual prominence and river-travel.

The traveller characteristics we selected are an average of data extracted from Binford²¹ combined with our own ethnographic interviews among Martu Aboriginal foragers³: female, 23 years old, 155.8 cm height, 45 kg weight. From our own ethnographic work, we determined that travellers walking long distances carry loads of an average of 10 kg,

including water- and food-procurement tools. We assigned all our travellers as female and mid-way through their reproductive years, suggesting that, while male travellers might make longer forays, female travellers would likely be leading families into new territories and holding groups together. Since female travellers will travel with different efficiencies than male travellers, this downscales our models to those that would accommodate women and families, on average. From our ethnographic work we determined that foraging decisions would be highly local and provide fewer constraints to a continent-wide model than search for fresh water would. However, incorporating an optimal foraging theory framework to future models that examine highly localised conditions for movement would be an opportunity for future study.

We ran *FETE* a total of 18 times to explore six different feasible origin/destination travel scenarios for people moving into a new region, with and without the influence of visually prominent landscape features and Strahler Stream Order 9 streams (Table 2). This added up to a total of 125,042,250,636 paths.

Outputs

We put the frequency surface output for each *FETE* run through a multi-step process with the goal of producing layers that can be evaluated statistically for evidence of non-random distribution. The first step was to use four different thresholds of the frequency surface representing the top 20%, 10%, 5%, and 1% of all cells with respect to the number of paths that crossed through them. This produced four binary masks for each run that show networks with potential explanatory value. We inputted each binary mask into the three-dimensional least-cost travel distance application discussed above, along with the mean-height surface, to produce a new surface where the value in each cell represents the shortest travel distance from that cell to the threshold-based network.

Reference archaeological data

We constructed the reference archaeological site dataset for Sahul by building on the *AustArch* compilation⁹⁷; for more information on age reliability we recommend reviewing that publication, although we detail our process below. We removed sites younger than 35,000 years old and sites with ages more than 35,000 years old that are documented as not directly related to human occupation (e.g., ages taken on samples below archaeological deposits). We selected 35,000 years ago as a threshold to encompass ages representing the earliest phases of human occupation of the continent, and before significant societal change

from the Last Glacial Maximum. We systematically compiled Australian sites published since the completion of *AustArch* in 2014 from the literature, ages missed in the original *AustArch* compilation, and extended the dataset to include New Guinea to encompass all of Sahul. We did not include archaeological sites on islands that have never formed part of the Sahul landmass (e.g., the Bismarck and Solomon Archipelagos) in the dataset.

We consulted the original sources for all sites, where available, to compile details about the dating technique, with a particular focus on sample pretreatment, provenance, and cultural associations. Missing sample pretreatment information was, where available, provided by the University of Waikato Radiocarbon Dating Laboratory and the Australian National University Radiocarbon Laboratory, or in some cases we obtained this information directly from the original researchers.

We quality-rated ages largely following the protocol developed by Rodríguez-Rey *et al.*⁹⁸ for evaluation of ages associated with Sahul megafauna deposits. Quality rating comprised a two-step process resulting in allocation of ages to one of four categories of reliability. The first step evaluated the dating technique itself resulting in assignment of the age to one of four categories (best to worst: m*, m, B, C). The second step evaluates the strength of association of ages rated m* and m with the dating target, in this case the association of the sample with cultural deposits, resulting in a final reliability rating (best to worst: A*, A, B and C). Where a sample is reported as directly associated with a stratigraphic unit containing cultural deposits, we have accepted it as ‘associated’.

Quality ratings for other dating techniques such as optically stimulated luminescence follow the criteria outlined in detail in Rodríguez-Rey *et al.*⁹⁸. We reduced the dating quality of any infinite age or any age without a reported error to C. We reduced the overall quality rating of any age > 75 ka to a C, because this is beyond the accepted age of initial peopling of Sahul. We reduced the overall quality rating of any age to a C where the authors reporting the age state that it is unreliable, or where the authors note that it is not associated with cultural material.

We deviated from the Rodríguez-Rey *et al.*⁹⁸ approach in accepting charcoal ages pretreated with acid-base oxidation (ABOx) as ‘m*’ (rather than ‘m’) and acid-base-acid (ABA) or acid-alkali-acid (AAA) as ‘m’ (rather than ‘B’). We also rated hydrogen pyrolysis (hypy) as ‘m’, which was not included in Rodríguez-Rey *et al.*⁹⁸. For marine shells, we rated samples subject to either x-ray diffraction or Feigl staining as ‘m*’ rather than ‘m’. The original protocol only accepted charcoal ages pretreated with ABOx. Although ABOx remains the standard for pretreatment of older charcoal samples (> 30 ka)⁹⁹, recent studies

have shown that ABA- and hypy-treated pairs produce comparable results to each other and to ABOx-pretreated samples^{100,101}. Only a few samples in Sahul have been subjected to ABOx pretreatment, so this permits a much broader consideration of the available data. ABOx pretreatment is also unsuitable for many samples because it results in losses of sample material¹⁰².

To avoid spatial bias, we collapsed two landscapes with closely spaced clusters of sites into single sites. Here, we collapsed four sites from the Ivane Valley (AAXD Airport Mound, AAXE South Kov, AAXF Vilakuav, AER Kosipe Mission Trench 2) into a single site designation ('Ivane Valley'); these sites are all within 2 km of one another. Likewise, we collapsed fifteen sites from Willandra Lakes (Lake Arumpo (Top Hut Site 3), Outer Arumpo (Top Hut Site 3), Lake Arumpo (Top Hut Site 1), Leaghur Peninsula, Mungo (WOC-4), Mungo 76E, Mungo B Trench, Mungo Hearth F12, Mungo Hearth F7, Mungo Hearth F8, Mungo I Residual, Mungo I Transect, Mungo III Transect, Mungo WLH3, Mungo WLH4) into a single site ('Willandra Lakes'); these fifteen sites were within 25 km of one another.

Using the processes above, we ultimately included 40 sites with ages $\geq 35,000$ ka and quality-rated as A*, A or B in the reference archaeological site dataset used in the modelling process.

Model performance

To examine the correspondence of each threshold-based network for each *From Everywhere to Everywhere* run to the locations of archaeological sites during this early time period, we first created a way to treat the network as a single point, which enabled the use of traditional point-based statistical tests for complete spatial randomness. These tests normally compute the distances between all possible pairs of points in two similarly sized sets, one of which is held constant while the other is randomly generated from a Monte Carlo process. We instead substituted the estimates of shortest travel distance to the network associated with the points of interest and the same number of points generated 10,000 times randomly. Here we are in effect creating a new coordinate system where the origin (single point) represents the entire network and the points of interest (and their randomly generated counterparts) are arranged around it based on how far away they are from the network. This establishes a common frame of reference for statistically evaluating the spatial distributions of points of interest and similar sets of randomly generated points: one that is focused on the most salient metric for this study — the proximity of sites to the network. We are interested in understanding how sites are distributed with respect to the network *as a whole*, not with respect to any one part

of it, and this new coordinate system enables that kind of analysis. We began with the most simple and straightforward Kolmogorov-Smirnov complete spatial randomness test to check for basic randomness before proceeding to the more sophisticated Besag's L test (itself a normalisation of Ripley's K -function¹⁰³), which tests for clustering at user-specified distance thresholds. Given the continental scale at which this analysis occurs, we applied 20 evenly spaced thresholds, ranging from 5 to 100 km. This allowed for the possibility of local clustering, regional clustering, or both. We pooled the results of the threshold tests for a specific network prior to doing a final test to calculate overall statistical significance and pooled probability estimates. We include all code in the Supplementary Data and Code.

Model comparisons and averaging

We tested the ability of a given network to reproduce the spatial pattern of archaeological sites by using a multi-staged process. We first calculated the shortest, three-dimensional travel distance from each archaeological site to the network, depending on surface elevation, using a standard anisotropic approach for generating a least-cost surface. Each cell in the resulting grid (here corresponding to our 500-m resolution cells) represents the shortest travel distance from that location to the network, and we averaged the values for a 1-km neighbourhood of cells around each site to produce more comprehensive distance estimates. This creates a new coordinate system where the entire network is represented by the origin (0, 0) as a single point. In other words, the frame of reference, and per-site distances from that point, can be used as a metric in two standard statistical clustering tests for complete spatial randomness—Kolmogorov-Smirnov^{104,105} and Besag's L -function¹⁰⁶—to determine how well the network explains site locations. The stronger the clustering, the stronger the relationship.

Both clustering tests require the comparison of site distances to those obtained from the same number of randomly generated points—a process that is generally repeated many times to estimate an empirical value; for this analysis we selected 10,000 trials. The Kolmogorov-Smirnov test is straightforward and indicates whether two samples are drawn from the same distribution. If a large enough number of the comparisons between site distances and those for random points pass the test, which is based on a standard uniform distribution probability, complete spatial randomness cannot be ruled out and the network in question is declared uninformative.

If enough simulated estimates do not pass the Kolmogorov-Smirnov test, indicating that site locations are not random with respect to the network, Besag's L -function evaluates how

resilient the relationship is by examining it at multiple spatial scales. The greater the number of scales where it passes, the stronger the explanatory power. Instead of comparing site distances to each set of randomly generated points, we use all sets of randomly generated points to create a single empirical estimate of randomness against which the set of sites can be compared at a series of distances (5 to 100 km, in 5-km increments). We then pooled results across all distances to create a single probability estimate for the network. The L -function can test for either clustering or dispersion with respect to a given frame of reference. For this analysis, we considered clustering.

To assess the information-theoretic weight of evidence (the parsimony-weighted index of maximum likelihood) of each network, we calculated an approximation of the Bayesian information criterion (BIC_a). We defined BIC_a for a given network (i) as:

$$BIC_{a_i} = k \log_e n - 2 \log_e \left(p_i \left(\frac{\max(C)}{c_i} \right) \left(\frac{\max(OD)}{OD_i} \right) + \varepsilon \right) \quad [\text{Equation 9}]$$

where k = the number of model parameters (4; origin type, destination type, weighting layer/no weighting layer as one option, and surface elevation), n = the number of archaeological sites (40), p = the pooled probability of Besag's L -function, C = the number of land cells used, OD = the number of origin and destination pairs supplied (equal to the number of paths generated), and $\varepsilon = 1.0 \times 10^{-16}$ (small error term to ensure 0 is not used in logarithm calculations). This version of the equation ranks the most efficient and parsimonious models most highly by rewarding high-probability networks that are simultaneously compact and for which a smaller number of origin and destination pairs were supplied. We used pooled probability instead of residual sum of squares because that goodness-of-fit estimate is already included in Besag's L -function. We examined 20 spatial scales from 5 to 100 km, resulting in a more robust estimation of network performance.

We assumed that no single network can explain all site locations, and that each network i has some value, so we calculated the network weights as the relative weight of its BIC_a ($wBIC_a$):

$$wBIC_{a_i} = \frac{e^{\Delta BIC_{a_i}/-2}}{\sum_{j=1}^M e^{\Delta BIC_{a_j}/-2}} \quad [\text{Equation 10}]$$

where M is the total number of networks,

$$\sum_{i=1}^M wBIC_{a_i} = 1, \quad [\text{Equation 11}]$$

and

$$\Delta\text{BIC}_{a_i} = \text{BIC}_{a_i} - \min(\text{BIC}_a) \quad [\text{Equation 12}]$$

We applied the per-network $w\text{BIC}_{a_i}$ weights to reduced-resolution (0.1 degrees) binary representations of the networks and summed the results across all modelled networks to create a single, model-averaged, composite grid. These provide the probabilities of paths chosen reported above in the Results.

Data availability statement

All thresholded binary masks created from all *FETE* runs, the model-averaged composite grid, lakes, and archaeological sites are made publicly available in standard geospatial data formats as part of our Supplementary Data and Code. Due to large file size, we provide the two digital elevation models and aggregate viewshed in standard geospatial data formats via GitHub: github.com/dawwhite/sfa.

Code availability statement

Due to how the development of *FETE* was funded, it is not currently possible to make the source code available to the public. As an alternative to public release of the code, the full methodology of the baseline version of *FETE* is described in White and Barber⁸⁸ and we describe modifications made to support our study here. Researchers who are interested in replicating the functionality of *FETE* can do so either by using an open-source or commercial GIS software package in combination with basic scripting in a high-level programming language like Python and following the methodology described herein. We provide all code used for analyses via GitHub: github.com/dawwhite/sfa.