















# Stochastic models support rapid peopling of Late Pleistocene Sahul

Corey J. A. Bradshaw <sup>1,2✉</sup>, Kasih Norman<sup>2,3</sup>, Sean Ulm <sup>2,4</sup>, Alan N. Williams<sup>2,5,6</sup>, Chris Clarkson <sup>2,7,8,9</sup>, Joël Chadœuf<sup>10</sup>, Sam C. Lin <sup>2,3</sup>, Zenobia Jacobs <sup>2,3</sup>, Richard G. Roberts <sup>2,3</sup>, Michael I. Bird <sup>2,11</sup>, Laura S. Weyrich<sup>2,12</sup>, Simon G. Haberle <sup>2,13</sup>, Sue O'Connor <sup>2,13</sup>, Bastien Llamas <sup>2,14,15</sup>, Tim J. Cohen<sup>2,3</sup>, Tobias Friedrich <sup>16</sup>, Peter Veth <sup>2,17</sup>, Matthew Leavesley <sup>2,4,18</sup> & Frédérik Saltré <sup>1,2</sup>

The peopling of Sahul (the combined continent of Australia and New Guinea) represents the earliest continental migration and settlement event of solely anatomically modern humans, but its patterns and ecological drivers remain largely conceptual in the current literature. We present an advanced stochastic-ecological model to test the relative support for scenarios describing where and when the first humans entered Sahul, and their most probable routes of early settlement. The model supports a dominant entry via the northwest Sahul Shelf first, potentially followed by a second entry through New Guinea, with initial entry most consistent with 50,000 or 75,000 years ago based on comparison with bias-corrected archaeological map layers. The model's emergent properties predict that peopling of the entire continent occurred rapidly across all ecological environments within 156–208 human generations (4368–5599 years) and at a plausible rate of 0.71–0.92 km year<sup>-1</sup>. More broadly, our methods and approaches can readily inform other global migration debates, with results supporting an exit of anatomically modern humans from Africa 63,000–90,000 years ago, and the peopling of Eurasia in as little as 12,000–15,000 years via inland routes.

<sup>1</sup> Global Ecology, College of Science and Engineering, Flinders University, Adelaide, SA, Australia. <sup>2</sup> ARC Centre of Excellence for Australian Biodiversity and Heritage, Wollongong, NSW, Australia. <sup>3</sup> Centre for Archaeological Science, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, NSW, Australia. <sup>4</sup> College of Arts, Society and Education, James Cook University, Cairns, QLD, Australia. <sup>5</sup> Climate Change Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia. <sup>6</sup> EMM Consulting, St Leonards, NSW, Australia. <sup>7</sup> School of Social Science, University of Queensland, Brisbane, QLD, Australia. <sup>8</sup> Max Planck Institute for the Science of Human History, Jena, Germany. <sup>9</sup> Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, NSW, Australia. <sup>10</sup> UR 1052, French National Institute for Agricultural Research (INRA), Montfavet, France. <sup>11</sup> College of Science and Engineering, James Cook University, Cairns, QLD, Australia. <sup>12</sup> Department of Anthropology, Pennsylvania State University, University Park, PA, USA. <sup>13</sup> Department of Archaeology and Natural History, School of Culture, History and Language, Australian National University, Canberra, ACT, Australia. <sup>14</sup> School of Biological Sciences, Environment Institute, University of Adelaide, Adelaide, SA, Australia. <sup>15</sup> National Centre for Indigenous Genomics, Australian National University, Canberra, ACT, Australia. <sup>16</sup> Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawai'i at Manoa, Honolulu, Hawai'i, USA. <sup>17</sup> Archaeology and the Centre for Rock Art Research and Management M257, School of Social Sciences, University of Western Australia, Crawley, WA, Australia. <sup>18</sup> Department of Anthropology and Sociology, University of Papua New Guinea, Port Moresby, Papua New Guinea. ✉email: [corey.bradshaw@flinders.edu.au](mailto:corey.bradshaw@flinders.edu.au)

Around 65,000–50,000 years ago (or earlier<sup>1</sup>), large<sup>2</sup> and apparently well-organized groups<sup>3</sup> of anatomically modern humans first entered the continent of Sahul—the landmass connecting New Guinea, some small islands of modern-day Indonesia, mainland Australia and Tasmania at lower sea levels than today. However, the exact timing of initial peopling remains controversial<sup>4,5</sup> because of arid Sahul's poor preservation environment and the relatively few reliable archaeological sequences compared to other regions of the world. Nonetheless, models are converging on the idea that first entry into New Guinea through Sulawesi was more likely<sup>3,6,7</sup>, although a southern route through Bali, Timor and onto the now-drowned Sahul Shelf in the Arafura Sea north of the modern-day Kimberley region in Western Australia is also plausible (or possibly both)<sup>3,8–10</sup>.

To date, the most likely migration routes and overall continental pathways of initial peopling have only been speculated from distant point estimates of genetic distance<sup>11–13</sup>, inferred biogeographic connectivity<sup>14,15</sup> or sparse archaeological data<sup>16–20</sup>. Further, the movements of people through New Guinea have been largely ignored, even though the region was directly connected to northern Australia throughout the hypothesized period of initial peopling<sup>2,3</sup>. Recent genetic evidence from Aboriginal Australians and Papuans<sup>11,12,21–23</sup> suggests that these peoples originated from a common ancestral population<sup>11,13,21</sup>, eventually diverging into their respective genetic clusters after an initial bottleneck<sup>11</sup>, with little evidence for subsequent gene flow<sup>12,13,22,23</sup>. While genomic evidence proposes that the peopling of Sahul was the result of a single migration event<sup>11</sup>, mitochondrial DNA evidence concludes that settlement occurred via both the southern and northern routes<sup>13,23</sup>. Subsequent gene flow apparently occurred predominantly along the coasts of Australia<sup>11,12</sup>, with the arid interior acting as a barrier to movement<sup>11</sup>.

Regardless of the true timing(s) and point(s) of initial entry into Sahul, there is little ecological insight into how these first wave(s) of humans moved across and established populations in the rest of the continent. There is evidence that hunter-gatherer use of space is structured by ecological factors<sup>24–27</sup>. For instance, the average distance of residential mobility is a direct function of the available energy in the local environment<sup>28–30</sup>. These relationships reflect fundamental features in human land use, regardless of the physical or cultural environment, and are thus useful for constructing generalized predictions of past human dispersal at broad spatial scales to test against archaeological data.

We constructed a fully stochastic, cellular-automaton model at a resolution of  $0.5^\circ \times 0.5^\circ$  to quantify how the patterns of peopling of Sahul could have occurred based on measured ecological mechanisms. Stochastic, complex-system models like the one we present here are useful not only for forecasting phenomena like global and regional climate patterns<sup>31</sup>, biodiversity extinctions<sup>32</sup> and land use<sup>33</sup>, but also for hindcasting past phenomena based on hypothesized, but hitherto untested mechanisms<sup>34,35</sup>. Our model and analyses incorporate data-based human-population dynamics to project cell-specific population densities relative to hindcasted<sup>35</sup>, climate-dependent<sup>36</sup> predictions of maximum carrying capacity derived from established relationships between hunter-gatherer densities and environmental conditions<sup>37</sup>, data-based rules of migration between neighbouring cells, measured long-distance dispersal<sup>38–41</sup> contingent on water availability<sup>15</sup> and landscape ruggedness<sup>42</sup> (a proxy for accessibility), as well as spatially clustered catastrophic mortality events operating at a generational scale based on data from across vertebrate taxa<sup>43</sup>.

We ran a series of 120 scenarios of the model modifying the dates and places of initial entry into Sahul from the islands of Wallacea<sup>3</sup>, as well as the hypothesized relationship between hindcasted net primary production (available biomass to consumers in an ecosystem) and maximum human carrying capacity. To rank the

relative realism<sup>44</sup> of which timings of initial entry, point(s) of initial entry and carrying capacity scenarios were best supported by the available archaeological data, we designed an analysis of the oldest-available, quality-rated archaeological dates to generate a map of initial arrival across the continent corrected for spatial bias and temporal uncertainty. We calculated the correspondence between this bias-corrected map and each of the 120 modelled scenarios to rank their relative likelihood given the archaeological data.

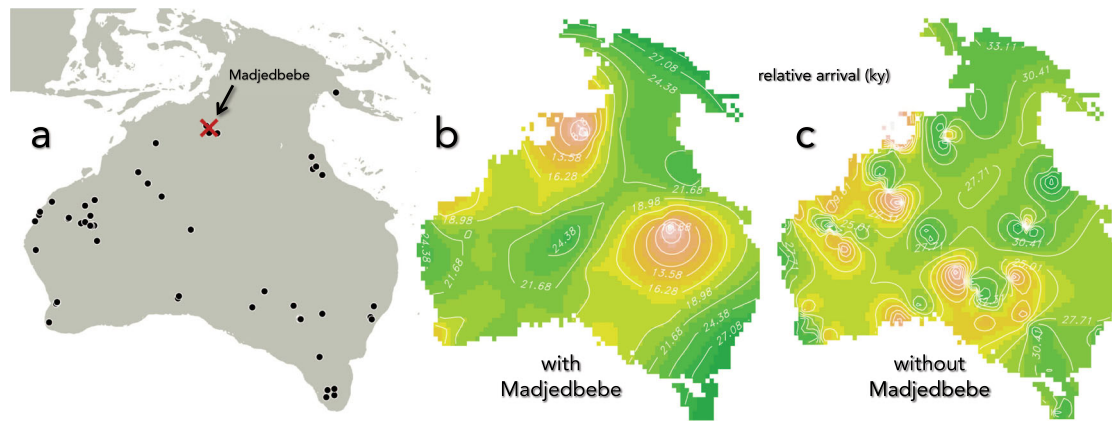
Our model provides a baseline approximation for the relative timing and spread of human dispersal in Sahul on the basis of ecological parameters alone. The reliance here on ecological principles explicitly does not mean that human dispersal in Sahul did not involve deliberate social, cultural and/or economic decisions. Rather, the movement parameters we used were derived from a large global dataset of contemporary and recent hunter-gatherer populations, whose behaviours and decision-making processes were embedded with cultural knowledge and practices. The stochastic nature of the model allows the effect of cultural adaptation on human movement, as captured in the ethnographic dataset, to be expressed probabilistically as part of the modelled human–environment relationship.

Here we show that full peopling of the entire continent of Sahul could have occurred within as few as 150–200 human generations (<5000 years), with up to 6.5 million people potentially living on the continent at saturation (New Guinea: ~1 million and Australia: ~5.5 million). Compared to the mapped archaeological layer, we find most relative support for an initial entry from Wallacea via the southern route onto the now-drowned Sahul Shelf, with a possible second entry later in western New Guinea. The results support both entry times of 50 or 75 ka (1000 years ago = ka), and a non-linear (rotated parabolic) relationship between human carrying capacity and net primary production as most likely of the three variants we test.

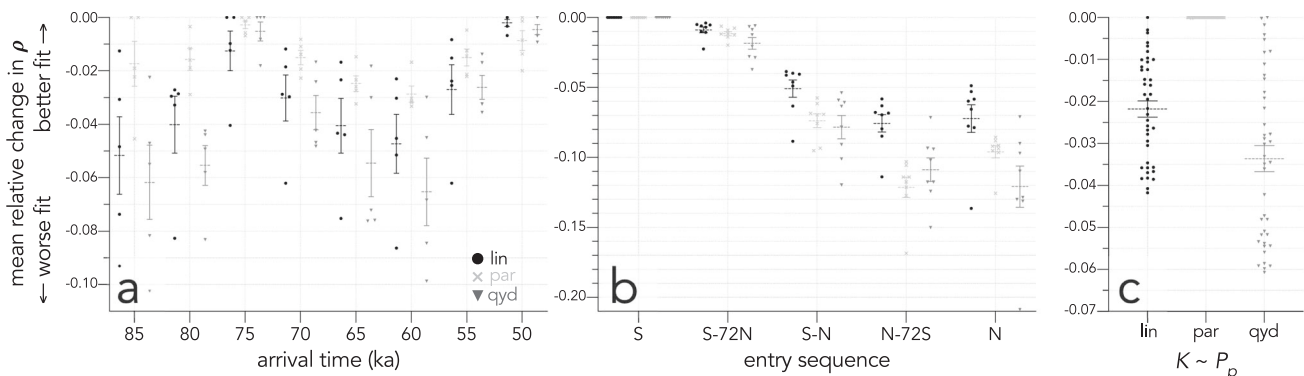
## Results

To estimate the most likely migration routes and overall continental pathways of initial peopling of Sahul, we ran 120 different scenarios (see details in 'Methods') where we varied entry time in 5000-year increments between the plausible limits of 85 and 50 ka, number of entries (single/dual), entry location (following a dominant northern and/or southern route) and sequence (northern–southern or southern–northern), and the relationship between human carrying capacity and net primary production (linear, parabolic or quadratic yield density; see details in 'Methods'). By comparing our simulated predictions to the spatial bias- and Signor–Lipps-corrected map layer of timing of first arrival derived from the most-reliable (quality-rated) and currently available archaeological data (Fig. 1; see also Supplementary Information and Supplementary Data 1), the following characteristics emerged:

First, of the eight potential arrival times we considered, there was near-equal support (mean relative change in Spearman's  $\rho$  compared to the top-ranked model =  $-0.005$ , Fig. 2a) for an entry time at 50 ka compared to 75 ka (mean relative change in Spearman's  $\rho$  compared to the top-ranked model =  $-0.007$ ; Fig. 2a). For the procedure to determine the correlation between modelled and archaeological estimates of initial arrival, see 'Spatial correlation between modelled and archaeologically derived dates of initial arrival' in the Supplementary Information. The fits derived from these two entry times arise because the older projected dates for the northwest of Sahul near the now-drowned Sahul Shelf entry corresponded closer to the 75-ka date, whereas the distribution of dates in the southeast of Sahul (modern-day Victoria and New South Wales) best aligned with entry at 50 ka (Fig. 1).



**Fig. 1 Spatial reconstructions of oldest archaeological dates.** **a** Distribution of reliable (A\* and A quality-rated) archaeological dates older than 30,000 years used to generate **b** a Signor-Lipps- and spatial-bias-corrected map layer of first arrival (see details in ‘Methods’); in **c** we removed the disputed<sup>1,4,5,45–48,95</sup> Madjedbebe dates (position indicated in **a** by a red cross) and reconstructed the layer. All date isohyets in **b** and **c** are expressed in 1000-year (ky) increments relative to arrival time and are based on 100 simulated replicates of the spatial-inference algorithm. Colours indicate progressively earlier (pink) to later (green) archaeologically derived dates. See ‘Methods’ (Cellular-automaton framework) for the source of the map extent for Sahul.



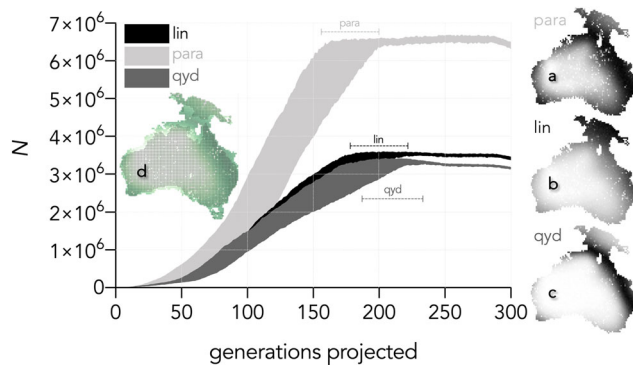
**Fig. 2 Correlation differences between first-arrival scenarios and archaeological dates.** Relative reduction in Spearman’s  $\rho$  rank correlation between the predicted time of first arrival over all grid cells for each scenario and the Signor-Lipps- and spatial-bias-corrected archaeological map layer. Holding all other input parameters equal, the mean relative change in average  $\rho$  is expressed as a function of modifying **a** arrival time (from 85 to 50 ka, in 5000-year increments; five scenarios for each increment), **b** entry-point sequence (S = southern route through the Sahul Shelf; N = northern route through Bird’s Head of New Guinea, and combinations of these with lags expressed in terms of human generations: 72 generations ~2000 years; 8 scenarios each), and **c** assumed relationship between human carrying capacity ( $K$ ) and hindcasted net primary production ( $P_p$ ) (lin = linear; par = rotated parabolic; qyd = reciprocal quadratic yield density—see details in ‘Methods’; three scenarios each). The horizontal bars represent the mean relative change in  $\rho$  and the error bars represent standard errors of the mean.

In terms of entry sequence, there was clearly more support (mean relative change in Spearman’s  $\rho$  compared to the top-ranked model = 0.000; Fig. 2b) for a single entry point at the Sahul Shelf offshore from the modern-day Kimberley region of Western Australia (southern entry), with slightly less support (mean relative change in Spearman’s  $\rho$  compared to the top-ranked model = -0.013, Fig. 2b) for a subsequent entry at western Bird’s Head (Vogelkop Peninsula) in modern-day western New Guinea (i.e., a single southern entry is most supported, followed by a dual entry first in the south and a second entry in the north). In contrast, all scenarios where the northern entry preceded a southern entry, or when there was only a northern entry, always provided worse fits to the primary archaeological map layer (Fig. 2b, Supplementary Fig. 1 and Supplementary Data 1).

Holding all parameters constant while only modifying the assumed relationship between human carrying capacity ( $K$ ) and net primary production ( $P_p$ ), we found that the rotated parabolic relationship (Fig. 2c and Supplementary Fig. 2c, f) gave the best fit between predicted settlement patterns and the primary

archaeological map layer, followed by the linear (Fig. 2c and Supplementary Fig. 2b, e) and the quadratic yield density relationships (Fig. 2c and Supplementary Fig. 2d, g). However, both the linear and the rotated parabolic relationships appeared to overemphasize extreme values of  $K$  across the landscape (Supplementary Fig. 2b, c, e, f). When we repeated the analysis excluding the Madjedbebe site<sup>1,4,5,45–48</sup> from the derived archaeological map layer (Fig. 1c), the overall fits were much weaker (Supplementary Figs. 3 and 4). While the results using this layer were inconclusive with respect to entry time (Supplementary Fig. 3a), there remained more support for a dominant southern entry (Supplementary Fig. 3b) and the rotated parabolic relationship between  $K$  and  $P_p$  (Supplementary Fig. 3c).

Based on the projection with a 50-ka southern entry, but varying the underlying relationship between carrying capacity and net primary production, we find that the best-supported rotated parabolic relationship projected a total population of 6.3–6.5 million (0.55–0.57 people  $\text{km}^{-2}$ ) within 300 generations (New Guinea: ~0.93–0.96 million; Australia: 5.4–5.5 million), compared to ~3.1–3.5 million inhabitants (0.27–0.30 people  $\text{km}^{-2}$ ) using the



**Fig. 3 Projected population size according to three assumed relationships between human carrying capacity ( $K$ ) and net primary production ( $P_p$ ).** In all three curves, we applied top-ranked scenario settings (50-ka entry at the southern entry-point), but altered the underlying carrying capacity–net primary production relationship. Shaded areas represent 95% confidence intervals based on 100 stochastic simulations. The best-supported relationship was for the rotated parabolic (a. para; see also Supplementary Fig. 2), followed by the linear (b. lin) and then the quadratic yield density (c. qyd). The rotated parabolic resulted in a final population size of 6.31–6.51 million people at 300 generations, or an average population density of 0.55–0.57 people  $\text{km}^{-2}$ . The linear model projected a total continental population of 3.37–3.46 million inhabitants at 300 generations (an average population density of 0.29–0.30 people  $\text{km}^{-2}$  over the entire continent), and the quadratic yield density projected a population of 3.13–3.19 million (density = 0.27–0.28 people  $\text{km}^{-2}$ ). These values are considered high compared with existing archaeological data (see ‘Discussion’), but provide an indication and trends of the populations during the initial peopling of the continent. Inset d shows the extent of Sahul overlaid with the modern grid for Australian and New Guinea. This reveals that 27.5% of the population estimate is attributed to now-drowned parts of the continental shelf. Further, the horizontal bars indicate the range of generations required to reach saturation of Sahul for each  $K$ – $P_p$  relationship. The maps on the right show the relative population distribution at 300 generations for a single iteration following the three different  $K$ – $P_p$  relationships. See ‘Methods’ (cellular-automaton framework) for the source of the map extent for Sahul.

other (linear or quadratic yield density)  $K$ – $P_p$  relationships (Fig. 3). These values are considerably greater than previous estimates<sup>49</sup>, and we provide several caveats regarding the interpretation of this result in ‘Discussion’. The rotated parabolic model (Supplementary Fig. 2f) predicted the highest densities in a general band from northern Australia, east through what is today the Gulf of Carpentaria, up into southeast New Guinea, and down the eastern coastal/inland region of Australia to modern-day New South Wales. Smaller pockets of predicted high carrying capacity occur along the western coast in the Kimberley, and farther south in south-western Western Australia, as well as in Tasmania. These are all areas with ample archaeological evidence of use by humans as refuges before and during the Last Glacial Maximum at ~21 ka<sup>50</sup>.

According to the top-ranked parameters for the different scenarios we considered, full occupation of Sahul occurred between 156 and 200 generations (4368–5599 years) after initial (southern) entry at 50 ka, or between 169 and 208 generations (4732–5796 years) after initial (southern) entry at 75 ka followed by a second northern entry at 73 ka (both using a rotated parabolic relationship between  $K$  and  $P_p$ ). If we take a direct line from the Sahul Shelf entry-point to the south-eastern coast of mainland Australia north of (but excluding) Tasmania (50-ka scenario), or to the south-eastern tip of Tasmania (75-ka scenario),

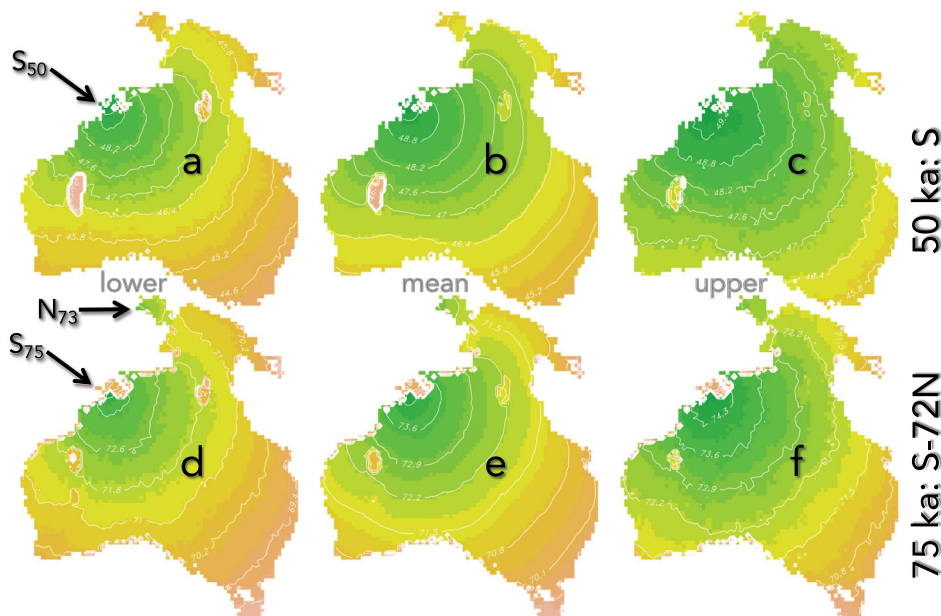
this implies a maximum rate of progression for regional occupation of 0.71–0.92  $\text{km year}^{-1}$ .

The general pattern of spread in these scenarios followed an eastward expansion from the southern entry-point from the Sahul Shelf (adjacent to the modern-day Kimberley coast), and then down both the west coast and south-east through east-central Australia (Fig. 4). The eastward expansion across inland Sahul follows the rotated-parabolic model’s predicted regions of highest density, suggesting early inland expansion through the Gulf Country and onto the eastern side of the continent. In both the single, southern entry- and dual-entry scenarios, early movement occurred from the Sahul Shelf northeast into New Guinea (Fig. 4). In the dual-entry scenario (75 ka at the Sahul Shelf followed by 73 ka at Bird’s Head), dispersal northeast from the Sahul Shelf and east from Bird’s Head reached the main body of New Guinea at approximately the same time.

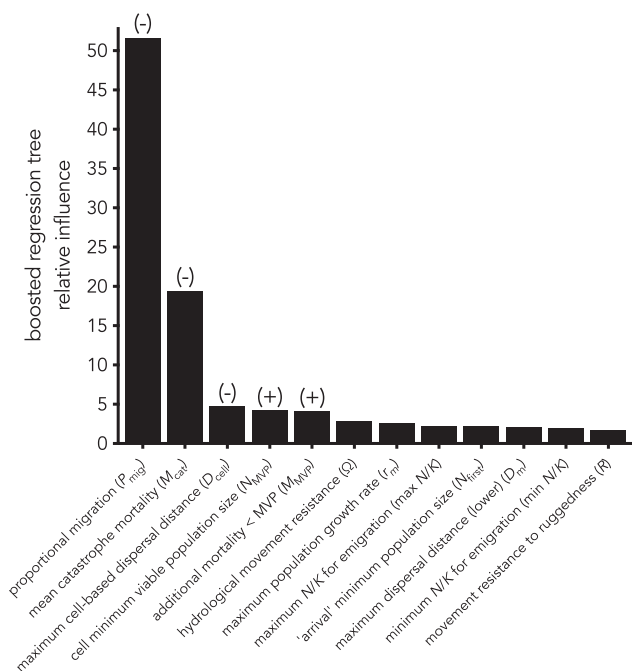
There was a notable gap in the peopling of the arid interior of Australia where the Gibson and Great Victoria Deserts are found today, which is imposed by the limitation of long-distance dispersal and by the water-availability dispersal and occupation-limitation functions (see details in ‘Methods’). This is consistent with the genetic evidence concluding the same phenomenon of extensive arid landscapes restricting settlement<sup>11</sup>. The presence of the inland deserts resulted in population divergence to either side of the arid interior. Most movement was eastward (as stated above), with a smaller dispersal south from the Kimberley, which followed a narrower dispersal path between the coastline and the arid zone. The spread continued southeast toward Tasmania in the 75-ka scenario (Fig. 4d–f). It is also notable that due to submergence of the land bridge between Tasmania and mainland Australia between 62 and 46 ka<sup>51</sup>, scenarios with an initial entry at 50 ka did not result in the peopling of Tasmania during this initial phase (Fig. 4a–c). This phenomenon arises because we assumed only terrestrial progression of the migration wave, whereas sufficiently advanced watercraft technology<sup>3</sup> could have brought the first people to Tasmania across the intermittently flooded Bass Strait<sup>51</sup>. We note, however, that there is no evidence of voyaging between Tasmania and the mainland since the flooding of Bass Strait ~12 ka<sup>52,53</sup>.

Testing the model sensitivity to variation in its parameters<sup>54</sup>, the boosted-regression tree emulator (explaining 97.8% of the deviance) for the twelve-dimension, Latin hypercube-sampled parameter space indicated that only two parameters had reasonable influence on the time to saturation of the continent (Fig. 5): (i) variation in the proportion of the population migrating between cells when an inter-cell movement occurred ( $P_{\text{mig}}$ )—a 50% decrease in this proportion resulted in a 21% reduction in the time to continental saturation (Fig. 5 and Supplementary Fig. 5a), and (ii) the negative effect of the mean mortality occurring during a catastrophe event ( $M_{\text{cat}}$ )—a 50% decrease in this value resulted in a 17% slowing of the time to continental saturation (Supplementary Fig. 5b). Three other variables had a weak effect on time to saturation (Fig. 5): (iii) the cell-based maximum dispersal distance modifier ( $D_{\text{cell}}$ ), (iv) cell minimum viable population size ( $N_{\text{MVP}}$ ) and (v) additional mortality <MVP ( $M_{\text{MVP}}$ ). Increasing the cell-based maximum dispersal distance ( $D_{\text{cell}}$ ) by five times resulted in a 5% faster time to continental saturation (Supplementary Fig. 5c), whereas a 50% increase in  $N_{\text{MVP}}$  resulted in a 2% slowing of the time to saturation (Supplementary Information Fig. 5d). Likewise, a 50% increase in the additional mortality below MVP ( $M_{\text{MVP}}$ ) slowed saturation time by ~4% (Supplementary Fig. 5e). All other parameters considered had substantially weaker influence on the rate of progression of the peopling wave (Fig. 5).

As another example of the relative insensitivity of the time to continental saturation by even large changes in maximum



**Fig. 4 Pattern of first arrival for the two top-ranked modelled scenarios.** Here, entry is via the southern route (entry points denoted by arrow in **a**) at 50 ka ( $S_{50}$ ; top row, **a-c**; '50 ka: S'), and an initial entry via the southern route at 75 ka ( $S_{75}$ ) followed by a second entry via the northern route 72 generations (~2000 years) later ( $N_{73}$ ; bottom row, **d-f**; '75 ka: S-72N'). Shown are the lower- and upper-percentile (95% confidence interval; **a, c** and **d, f**) and mean maps (**b** and **e**) for each scenario based on 100 runs of each model. Note in the 50-ka scenario (top row, **a-c**), Tasmania was not peopled within the modelled time frame (300 generations) because Bass Strait was flooded during this period (position of Tasmania shown in bottom row). Colours indicate progressively older (green) to younger (pink) dates. All date isohyets are expressed as ky relative to arrival time. See 'Methods' (cellular-automaton framework) for the source of the map extent for Sahul.



**Fig. 5 Global sensitivity analysis results.** Shown are relative inference scores (summing to 100 across all parameters considered) from a boosted-regression tree<sup>93</sup> of the relative importance of the model parameter ranges on the time taken to reach continental saturation. See main text for parameter descriptions and ranges tested. The most influential parameters (five top-ranked) are also given with the direction of their influence on the timing of continental saturation: (-) = negative, (+) = positive. Settings for the boosting regression tree were: error distribution = Gaussian, bag fraction = 0.75, learning rate = 0.008, tolerance = 0.0001, maximum number of trees = 10,000, and tree complexity = 2. See also Supplementary Fig. 5.

dispersal distance, we re-ran 100 iterations of the top-ranked scenario (southern entry only at 50 ka) with  $D_{cell}$  increased to five times the value in the original model. This only changed the estimated time to saturation to 153–185 generations (4285–5152 years) compared to 156–200 generations predicted from the original model.

**Discussion**

One of the most revealing insights provided by our stochastic model is that the rapid peopling of the entire continent of Sahul occurred potentially within as few as 4370–5600 years, across a wide range of environments, including rainforests, savannas, deserts, alpine regions, grasslands and temperate forests. This result was largely insensitive to assumptions regarding landscape carrying capacity (minimum and maximum ratios of  $N/K$  per cell invoking an emigration, and maximum population growth) (Fig. 5). Based on the ecological and anthropological mechanisms we applied, the emergent pattern of spread across Sahul was at least partially congruent with the overall pattern of first arrival derived from the bias-corrected primary archaeological map layer. Rapid expansion across Sahul has been hypothesized previously<sup>12,19</sup>, whereas our model provides the first robust test that such a swift saturation (at least from the perspective of the time people first entered Sahul) was biologically and ecologically feasible. Relying solely on archaeological or genetic data would be unlikely to identify, or be able to test, such rapid rates of human migration given the scarcity of reliable data at the continental scale during the period of initial peopling. The complexity of this fully stochastic model also ensures that no single component—measured or otherwise assumed—can dominate the outcomes of any scenario run. The emergent property of the many interacting and stochastically sampled functions and parameters provide the net spatial patterns of peopling and the plausible pace of this process. No single function within the model could achieve the same outcomes.

The emergent pattern of occupation revealed by our stochastic model is also largely consistent with the available genetic and archaeological evidence, especially with respect to the dominant pathways suggested for Sahul's western and eastern regions<sup>55</sup>, rather than through the dry interior. Indeed, our hydrologically limited dispersal function effectively supported the genetic view that the dry interior was at least partially a barrier to movement and permanent settlement for the earliest human occupants<sup>11,12</sup>, even though the realized value of the parameter itself had little effect on the rate of progression of the continental migration wave (Fig. 5). Where our model diverges, however, is that the movements southward were not necessarily coastal exclusively, especially within the central and eastern margins of the continent. Indeed, the hindcasted carrying capacity model suggests that expansive areas east of the western deserts were entirely amenable to migration and settlement according to the data-based functions applied in the model. Conversely, the extensive distribution of the arid lands westward (beyond even the modern-day extents of the western deserts) during most of the plausible period of initial settlement pushed the western movement corridor close to the west coast of Sahul (Fig. 4). It is important to be aware that our model predicts migration in terms of settlement patterns as opposed to temporary forays that do not result in local settlement. Second, while a genetic study<sup>11</sup> found evidence only for southward movement from New Guinea to Australia, our scenarios predict rapid, northeast movement onto the Sahul Shelf into New Guinea due to the high carrying capacity predicted for the Gulf region.

Our model has also provided the capacity to test the relative likelihood of some alternative entry points, as well as the demographic context for the initial peopling of Sahul from Wallacea. First, the highest-ranked scenarios all point toward a prevailing entry via the southern route (via the Sahul Shelf). Any scenario where a northern entry occurred before a southern entry led to a lower relative ranking. This suggests that although the northern route was likely an easier pathway for entry<sup>3</sup>, the dominance of the southern route is more consistent with the available archaeological evidence (recognising that archaeological investigation of New Guinea remains limited to date). Given that our highest-ranking scenarios provided the strongest support for a southern entry, with a potential secondary northern entry, our treatment of temporally restricted (<700 years) initial peopling events of >1000 individuals<sup>2</sup> is realistic. This also potentially implies that populations in Wallacea required sufficient time to grow large enough before widespread migration events took place<sup>2</sup>. Further, the mounting genetic evidence<sup>13,23</sup> concurs with our conclusions regarding timing, support for a dual entry and Late Pleistocene population expansion. Unlike our model, however, genetics cannot yet shed light on precedence of particular or multiple entry points.

The highest-ranked relationship between carrying capacity and net primary production predicted a maximum continental-wide population of 6.3–6.5 million (0.55–0.57 people km<sup>-2</sup>) based on the carrying capacity predicted from the LOVECLIM Earth-systems model and the particular form of the best-supported relationship between human-population density and net primary production we considered. One of the main determinants of this range—which is substantially larger than previous reconstructions of population size for Late Pleistocene Sahul<sup>49</sup>—is the relatively higher support for the non-linear relationship (rotated parabolic) between human carrying capacity and net primary production compared to the other two forms. Although a non-linear relationship is consistent with evidence from elsewhere that ancient people selected habitats that were structurally intermediate (ecotones) between open plains and dense forests (e.g., savannas)<sup>56</sup>, thus providing vegetation mosaics that were more ecologically diverse and favourable to human visitation and

settlement than maximum-productivity environments<sup>57</sup>, we did not have empirical data to support the exact form of that relationship. Testing the relative support among three different types of relationships does not mean that the rotated parabolic is in fact the correct one; instead, it merely gives more support for the hypothesis of non-linearity versus linearity. The combination of the spatial pattern of net primary production across Sahul (Supplementary Fig. 2) and the latter's relationship with human density drives the magnitude of the total population estimate (Fig. 3). Without a specific, empirically supported relationship, it is entirely plausible that the total population size could have been substantially smaller, especially if other, unknown conditions prevented populations from achieving local carrying capacity. The true magnitude of the population estimate is therefore not central to our approach nor the ensuing results.

Nonetheless, the large estimate and the resulting population-density range are not at odds with post-contact estimates of hunter-gatherer densities across temperate coastal regions of Australia (mean ± SD: 0.34 ± 0.24 people km<sup>-2</sup>, with an upper limit of 0.81 people km<sup>-2</sup> assuming a Gaussian distribution) and densities of shifting agriculturalists in New Guinea (4.37 ± 4.22 people km<sup>-2</sup>)<sup>58</sup>. In addition, the only existing estimates of population size in Sahul<sup>49</sup> did not account for arrival before 50 ka, and instead focussed on periods much later than the arrival window we considered. Neither did those estimates include New Guinea, which represents ~15% of the Sahul landmass during the period of interest (and contains areas of higher ecological productivity than much of the remainder of Sahul), nor did they include the sunken coastal-shelf (~1.43 million km<sup>2</sup>, or ~19% more land area than Australia today). Overlaying the modern extent of Australia and New Guinea on the predicted population grid at 300 generations from initial entry indicates that ~27.5% of the continent's total population is attributed to cells that are now under water (Fig. 3). In the case of the 50-ka single-entry scenario shown in Fig. 3, the above-water component of the Australian part of Sahul would give a total population of 3.1 million people at saturation.

The large assemblage of artefacts in the earliest phase of human occupation at Madjedbebe is also potentially consistent with a large population size<sup>1</sup>. For example, a 2 × 2-m area of the lower (Phase 2) assemblage at Madjedbebe contains tens of thousands of stone artefacts, >1000 pieces of ground ochre, >400 grindstones and ground fragments, >30 ground edge axes/fragments and includes distinctive technologies not recovered from the overlying layers (e.g., discoidal cores, thinning flakes, convergent flakes, points)<sup>1</sup>. This density of artefacts is greater than that seen for the Holocene layers at the site and the richness of cultural materials exceeds even that of many dense Holocene sites in Australia and New Guinea<sup>5,59</sup>. These data support the idea that Madjedbebe was intensively occupied, supportive of the early high population densities predicted by our model. In addition, recent evidence suggests that northern Sahul was relative wet during the period of human arrival as the intertropical convergence zone shifted south<sup>60</sup>, potentially increasing productivity in this region. Improving and expanding the archaeological record in places of initial arrival and where population densities were highest, including now-drowned but previously productive coastal-shelf sites<sup>61</sup>, must therefore form a research priority to improve model predictions of the development of human societies during the Late Pleistocene.

At a global scale, our conclusions have important implications for the migration of anatomically modern humans in the Late Pleistocene, and especially for those moving out of Africa. While specific to Sahul, our model does support the notion that during this period humans were able to migrate and occupy a wide range of environments and climatic zones, with only the most extreme

arid environments forming barriers to movement in this region. Our models do not indicate a preference for coastal environments, with inland grasslands and rainforests settled first; these results lend support to the hypothesis of inland dispersal<sup>62</sup>. Our estimated migration rates of 0.71–0.92 km year<sup>-1</sup> are consistent with previous estimates of rapid spread throughout Sahul from genetic evidence<sup>12</sup>. Our rates also provide a realistic model for maximum movement rates of humans across Eurasia, where as few as 12,000–15,500 years would be required to reach Sahul from Africa (assuming a minimum global-arc distance of 11,000 km; or 16,300–21,100 years assuming a minimum-distance coastal route of ~15,000 km). Assuming initial entry to Sahul between 75 and 50 ka, this lends support to arguments and genetic evidence<sup>63</sup> for permanent movement out of Africa at the end of Marine Isotope Stage 5 or beginning of Stage 4, and perhaps movement into Sahul<sup>64</sup> during both Stages 4 and 3.

Along with other recent models of the demographic conditions needed for successful entry to Sahul<sup>2</sup>, our model provides further evidence for the adaptability and technologically advanced state of anatomically modern humans during the Late Pleistocene. Our model also supports the notion of large populations being present across Southeast Asia and Sahul during this period<sup>2,3</sup>, which contrasts with past views of dispersed, small bands of hunter-gatherers. However, the lack of expansive archaeological records means that disparities persist, especially regarding the general absence of early sites in the vicinity of Bird's Head and the highlands of New Guinea. There is a marked disparity also between the high maximum population sizes proposed here and the comparative paucity of documented archaeological sites across Sahul prior to the Last Glacial Maximum.

## Methods

**Cellular-automaton framework.** We constructed the cellular-automaton model in the open-access R statistical computing environment (cran.rproject.org). We provide all code, data and instructions to repeat the analysis<sup>65</sup>, which can be run on any desktop computer. Our spatial model is based on a 0.5° × 0.5° raster grid of Sahul from 0.5 to 43.0° S latitude, and 110.5 to 153.5° E longitude (86 rows and 87 columns). The land area of Sahul changes with fluctuating sea levels, so we estimated exposed land in 1000-year time slices to follow our available hindcasts of carrying capacity (see 'Carrying capacity' below) based on a digital elevation model and estimated sea-level change over the period of interest (from 85 to 40 ka; see Scenarios). We used the ETOPO1 global relief model of Earth's surface<sup>66</sup> to estimate the exposed landmass of Sahul through time. To reconstruct the landmass changes of Sahul every 1000 years, we applied sea-level variability outputs<sup>67</sup> to the ETOPO1 model. We also included fluctuations in Lake Carpentaria that could potentially act as a natural barrier for human movement over time. We modified the contour of the lake based on modelled sea-level changes<sup>68</sup> applied to the digital elevation model.

From an initially peopled cell (see Scenarios), the new population can grow following a Ricker population-dynamics model, and emigrate to adjacent cells following stochastically resampled rules of dispersal; likewise, each cell can receive immigrants from adjacent cells following similar dispersal rules (see 'Emigration and immigration', and 'Long-distance dispersal').

**Population-dynamics model.** Each cell within the grid acts as a particular sub-population unit within the overall dynamics of Sahul, and the summary information provided at the end of a simulation is an overall expression of all cells. The change in human abundance ( $N$ ) within each cell is governed by the following phenomenological (Ricker) equation of population dynamics:

$$N_{i,j,t+1} = N_{i,j,t} e^{r_m \left(1 - \frac{N_{i,j,t}}{K_{i,j,t}}\right)} - (E_{i,j,t} - I_{i,j,t}) \quad (1)$$

where  $i$  is the cell row number in the 0.5° × 0.5° latitude lattice,  $j$  is the cell column number,  $t$  is the time interval in units of human generations (1 g = 27.9 years)<sup>2</sup>,  $N_{i,j,t+1}$  is the number of individuals in cell  $i, j$  at the next time interval ( $t+1$ ),  $N_{i,j,t}$  is the number of individuals in cell  $i, j$  at time interval  $t$ ,  $r_m$  is the maximum rate of population increase when resources are not limiting,  $K_{i,j,t}$  is the cell-specific carrying capacity (see 'Carrying capacity' below), and the  $E_{i,j,t}$  and  $I_{i,j,t}$  parameters represent the number of individuals emigrating from and immigrating into the focal cell  $i, j$  per time interval  $t$ , respectively (see 'Emigration and immigration'). As an estimate of  $r_m$ , we set the age-structured Leslie matrix for Aboriginal hunter-gatherers<sup>2</sup> to have a survival probability (subdiagonal matrix entries) all equal to 1 (complete survival in every age class), and then took the log<sub>e</sub> of that matrix's dominant eigenvalue to the power of  $g$  multiplied by

2 as the generationally scaled  $r_m$  estimate required for Eq. (1). Finally, we imposed a beta-resampled additional mortality parameter  $M_{MVP} = 0.2$  for cells with a population size <100 individuals. This threshold of  $N_{MVP} = 100$  is derived from the ecological concept of minimum viable population size<sup>69</sup> where ~100 effective individuals are required to avoid inbreeding depression<sup>70</sup>. This assumes that individual fitness declines when populations decline below the minimum size due to inbreeding depression from non-random mating and other Allee effects related to social disruption.

**Carrying capacity.** We constructed a theoretical carrying capacity ( $K$ ) for each cell from a hindcasted estimate of net primary production based on the LOVECLIM climate reconstruction<sup>2,3,35</sup>. In this case, we define carrying capacity as the upper limit to total abundance<sup>71</sup>, which is expressed mathematically as the point at which net population growth becomes stable ( $r = 0$ )<sup>72,73</sup>. LOVECLIM is a three-dimensional Earth-system model of intermediate complexity<sup>35</sup> that includes representations of the atmosphere, ocean and sea ice, land surface (including vegetation), ice sheets and the carbon cycle. LOVECLIM produces climates over the past 800 ka in 1000-year time-averaged increments (layers) that we downscaled (using a bilinear interpolation)<sup>74,75</sup> from a spatial resolution of 5.625° × 5.625° to 0.5° × 0.5°. For each grid cell and each 1000-year layer, we extracted net primary production<sup>76</sup> (kg C m<sup>-2</sup> year<sup>-1</sup>) as the comprehensive indicator of relative carrying capacity through time<sup>2</sup>. To translate net primary production into a carrying capacity expressed in units of humans the landscape was capable of supporting, we used the predicted relationship between net primary production and human density for hunter-gatherer societies<sup>37</sup>. Here, we rescaled the net primary production values over all grid cells to occur with the minimum (0.018 km<sup>-2</sup>) and maximum (1.152 km<sup>-2</sup>) human densities, multiplied by cell area (3080.25 km<sup>2</sup>) to give per-cell  $K$ . We then linearly interpolated these values between the 1000-year layers for each cell.

This approach assumes a linear relationship between  $K$  and net primary production (Supplementary Fig. 2). However, others have suggested a non-linear relationship where maximum  $K$  occurs at mid-range net primary production, because past humans possibly struck a compromise between high productivity and ease of passage and/or visibility to hunt prey by tending towards ecotones of mid-range productivity<sup>56,57</sup>; this type of non-linear relationship has also been identified between herbivore biomass and mean annual rainfall across sub-Saharan Africa<sup>77</sup>. We therefore also applied a 180°-rotated parabola model between  $K$  and net primary production ( $P_p$ ) of the form:

$$K = a(P_p - h)^2 + K_{\max} \quad (2)$$

where we set  $K_{\max}$  for the rotated parabolic relationship at  $0.5K_{\max}$  for the linear relationship,  $a = -3$ ,  $h =$  the median of  $P_p$ , and where we scaled the rotated parabolic  $\Sigma K_{\max}$  across all cells and temporal layers to equal  $\Sigma K_{\max}$  of the linear relationship (because of a higher frequency of cells with mid-range compared to high net primary production) (Supplementary Fig. 2). This approach assumes a logarithmic growth to and decline from a peak, so we also considered a third non-linear relationship—a reciprocal quadratic yield density—between  $K$  and  $P_p$  of the form:

$$K = \frac{P_p}{a + bP_p + cP_p^2} \quad (3)$$

where  $a = 200$ ,  $b = 0.6$  and  $c = 0.2$ , which instead assumes a logistic growth to and decline from a more pronounced peak (Supplementary Fig. 2). Whether we invoked the linear, rotated parabolic or reciprocal quadratic yield relationship (see Scenarios), we Poisson-resampled the resultant  $K$  for each cell per generational time step to simulate spatio-temporal uncertainty in  $K$ .

**Emigration and immigration.** We reasoned that emigration out of a focal cell  $i, j$  to its eight immediately neighbouring cells would be a function of the gradient in  $K_t$  between two cells<sup>44</sup>, as well as how close the abundance  $N_{i,j,t}$  of cell  $i, j$  was to  $K_{i,j,t}$ . If the Poisson-resampled  $K_t$  ratio between two adjacent cells ( $K_{\text{rel}} = K_{i,j,t} / K_{i+y,j+x,t}$ , where  $x$  and  $y$  are integers ranging from  $-1$  to  $1$  to define the immediate one-cell neighbourhood of the focal cell  $i, j$ ) was  $<1$ , and  $N_{i,j,t} / K_{i,j,t}$  was  $>$  a random uniform proportion between 0.3 and 0.7, then emigration ensued. The values of 0.3 and 0.7 originate from Birdsell<sup>78</sup> who stated that group 'budding off' or 'fissioning' occurred when a population reached 30–70% of carrying capacity. We further imposed an exponential decay function to describe the declining probability of emigration ( $\text{Pr}(E)$ ) as the ratio of  $K_{i,j,t} / K_{i+y,j+x,t} = K_{\text{rel}}$  increased towards 1 (Supplementary Fig. 6), where:

$$\text{Pr}(E) = e^{-3.2K_{\text{rel}}} \quad (4)$$

For the size of the emigrating group  $N_{\text{mig}}$ , we assumed that this represented a beta-resampled proportion  $P_{\text{mig}} = 1/3$ , such that  $N_{\text{mig}}$  is centred on  $N_{i,j,t} \times P_{\text{mig}}$  (with  $\text{SD} = 0.05N_{i,j,t} \times P_{\text{mig}}$ ) based on the observation that when coastal populations of Aboriginal Australians fission, the population tends to separate into one larger and one smaller group, with the latter emigrating<sup>79</sup>. We applied this calculation for each of the eight cell comparisons, removing those already emigrated from each cell's subsequent estimate (moving in sequence through directions NW, N, NE, W, E, SW, S and SE of focal cell  $i, j$ ). Likewise, when  $K_{i,j,t} / K_{i+y,j+x,t} > 1$ , immigration into cell  $i, j$  occurred following the same movement rules as for emigration.

**Long-distance dispersal.** We used the allometric relationship of natal dispersal for omnivorous and herbivorous mammals<sup>40</sup> to predict a dispersal probability for humans. Assuming a mean adult mass of  $M = 50$  kg, maximum natal dispersal distance  $D_m$  is estimated as  $aM^b$ , where  $a = 3.31 \pm 1.17$  and  $b = 0.65 \pm 0.05$  for omnivores and herbivores combined<sup>40</sup>. This produced an estimated maximum dispersal distance  $D_m$  ranging from 22.4 to 69.3 km. As a maximum dispersal range, this compares well to the average mobility of African hunter-gatherers of 1400–3900 km<sup>2</sup>/generation<sup>39</sup> (equivalent to a radius of 21.1–35.2 km assuming a perfect circle), and the 0.4–1.1 km yr<sup>-1</sup> (11.2–30.7 km/generation) estimates for Palaeolithic human expansions in northern Europe<sup>41</sup>. Also, Gould<sup>38</sup> reported journeys by Aboriginal Australians of 400 to 560 km as ‘not unusual’ and perhaps the greatest mobility ever recorded, moving as many as nine times in three months, and covering an area of ~2600 km<sup>2</sup> (radius = 28.8 km).

Next, we used the estimated probability of dispersal ( $\Pr(d_{\max})$ ) of  $d$  exceeding multiples (1 to 10) of one cell width ( $0.5 \times 111.12 = 55.6$  km) as  $\Pr(d_{\max}) = e^{-d/aM^b}$  (Supplementary Fig. 7). However, there is evidence globally that the territory size of hunter-gatherer groups is strongly related to local productivity, with a greater need to expand foraging areas as productivity declines<sup>19,80</sup>. Using territory size and rainfall data from Hiscock<sup>80</sup>, we assumed the same relative change in rainfall applied to net primary productivity, but shifted the power-law relationship upwards to match the slope of the upper limit of maximum dispersal distance (Supplementary Fig. 7b). Thus, for every tenfold decrease in relative net primary production, maximum dispersal distance increases by 12.7 times (Supplementary Fig. 7b). Once a long-distance dispersal event occurred, we Poisson-resampled the maximum dispersal distance to provide a  $\delta x$  and a  $\delta y$  to move from the focal cell in cell units (including a random direction: east-west for  $\delta x$ , and north-south for  $\delta y$ ). The size of the long-distance-dispersing population followed the same rules as for neighbouring-cell emigration.

**Distance-to-water limitation.** While territory size, and hence, maximum dispersal distances increase with increasing aridity according to the relationships described above, there is evidence that human dispersal is ultimately limited by water availability<sup>15</sup>. This is likely to be even more relevant in Australia, the driest inhabited continent on Earth—indeed, estimated routes of gene flow among Aboriginal Australians suggest that the arid interior acted as a barrier to migration<sup>11</sup>. We therefore invoked an additional limitation on dispersal by calculating a probability of realizing a long-distance dispersal event ( $P_i$ ) according to the following equation previously designed to limit modelled species migrations<sup>81</sup>:

$$P_i = 1 - \left( \frac{D_i}{D_{H_2O}} \right)^\Omega \quad (5)$$

where  $D_i$  is the realized maximum dispersal distance generated from the algorithm described above,  $D_{H_2O}$  is the distance to water in units of map cells derived from the Australian Water Observations from Space dataset<sup>15</sup>, and  $\Omega$  is the hydrological resistance parameter set arbitrarily to a value of 3 to invoke landscape-scale resistance to movement only in the driest areas of Sahul per generational time step.

**Ruggedness.** We hypothesized that high landscape ruggedness (elevational gradient) might at least partially impede the progress of human expansion across the landscape<sup>42</sup>, so we tested this using data available in an ethnographic and environmental dataset compiled by Binford<sup>42</sup>. Available in the `binford` library<sup>82</sup> in R, the dataset includes >200 variables measuring aspects of hunter-gatherer subsistence, mobility and social organization for 339 ethnographically documented groups. Given the evidence that mobility is a function of productivity<sup>36,80</sup>, we constructed a simple linear model of annual movement varying with annual rainfall and the difference between maximum and minimum elevation within a 25- (40.2 km) mile radius of the group’s centroid (equivalent to an elevational gradient; i.e., ruggedness). Taking the cube root of annual movement and the difference in maximum and minimum elevation to comply with the assumption of Gaussian error distributions, the expected relationship between movement and rainfall prevailed, and there was a weak effect of elevational difference—a maximum of 1% reduction in annual movement (Supplementary Fig. 8). Expressed on the linear scale and standardizing annual movement and elevational difference to the range of 0–1 (assuming a constant median annual rainfall value), an exponential decay function of the form:

$$M_{\text{red}} = a + b\sqrt[3]{G_{\text{rel}}} \quad (6)$$

where  $M_{\text{red}}$  is the proportion of expected total annual movement,  $a = 1.001116$ ,  $b = -0.0104453$  and  $G_{\text{rel}}$  is the standardized ruggedness from 0 to 1, described the reduction in annual movement rates up to a maximum of 1% (Supplementary Fig. 8). For all instances of emigration, immigration and long-distance dispersal, we assigned this function to the total number of people migrating for each cell based on its standardized ruggedness. We computed the topographic ruggedness index<sup>83</sup> as the difference in elevation between a given cell and its eight neighbouring central cells, based on our digital elevation model. For a given cell, we then squared each of the eight elevation difference values (to render them positive), and calculated the square root of the averages of the squares. We updated the spatial resolution of our results to  $0.5^\circ \times 0.5^\circ$  to match the other environmental layers.

**Catastrophic mortality events.** Palaeo-demographic investigations of past human populations suggest that long-term population growth rates were just slightly higher than zero as a result of episodes of catastrophic mortality arising from pandemics, natural disasters and violent conflicts occurring every few generations<sup>84</sup>. This also agrees well with estimates of the probability of mass mortality events scaling to generation time for vertebrates ( $\Pr(\text{catastrophe}) = 0.14$  per generation)<sup>43</sup>. We thus sampled binomially at  $\Pr = 0.14$  for whether a catastrophe occurred in each focal cell, and then beta-sampled the severity of the event centred on  $M_{\text{cat}} = 0.5$  (SD = 0.5/10) to emulate a stochastic catastrophe event of 50% mortality, on average, for that cell<sup>43</sup>.

However, we reasoned that a random allocation of catastrophes among cells across the entirety of Sahul was not realistic, for the reason that mortality events arising from natural disasters, warfare or disease outbreaks would likely be spatially aggregated. We therefore imposed a Thomas cluster process using the `rThomas` function from the `spatstat` R library<sup>85</sup>, setting the intensity of the Poisson process of cluster centres  $\kappa$  to a linear relationship between the number of cells occupied per iteration and a vector ranging from 0.3 to 1.2, the standard deviation of random displacement along each coordinate axis of the grid of a given cell away from the cluster centre  $\sigma_{\text{scale}} = 0.015$ , and the mean number of cells per cluster  $\mu = 0.6 \times$  the mean dimension of the occupied grid per iteration. This combination of parameters led to a reasonable degree of spatial clustering while maintaining a random spread of cells around a catastrophe focal point, as well as maintaining the overall proportion of cells across the landscape experiencing a catastrophic mortality event ~0.14 per generational iteration.

**Scenarios.** We ran 120 scenarios (8 entry times,  $\times 5$  entry sequences,  $\times 3$  relationships between carrying capacity and net primary production) where we modified three main components of the stochastic simulations: (i) the timing of first entry to Australia (from 85 to 50 ka, in 5000-year increments), (ii) the place of entry (northern, southern, simultaneous northern and southern, northern followed by southern 2000 years later, or southern followed by northern 2000 years later) and (iii) the form of the relationship between hindcasted net primary productivity and human carrying capacity (linear, rotated parabolic or reciprocal quadratic yield density). We repeated each scenario 100 times to generate a per-cell confidence interval of time of first arrival. Here, we deemed a cell to have been populated for the first time once it received  $\geq 100$  individuals ( $N_{\text{first}}$ ), which is considered the minimum viable effective population size to avoid inbreeding depression<sup>70</sup>.

**Comparison layers.** To test the resultant outputs against real archaeological data, we compiled a conservative list of ages older than 30 ka obtained from across Sahul (see ‘Compiling reference archaeological dates’ in the Supplementary Information and Supplementary Data 1). However, the spatial coverage of these ages is highly uneven (Fig. 1a), so we applied a maximum-likelihood method to correct for the Signor–Lipps effect first developed by Solow<sup>86</sup> and adapted for spatial inference of both first-arrival and extinction patterns<sup>87</sup>. While described in more detailed elsewhere<sup>87</sup>, we briefly summarize the approach here.

To correct for the inherent spatial bias of dates in a landscape, let  $x_1, \dots, x_n$  be the spatial locations of  $n$  dated specimens in an area  $W$  and  $a_1, \dots, a_n$  their respective ages. The estimated average age  $\hat{M}(x)$  of a putative date at a given location  $x$  is based on a standard kriging procedure<sup>88</sup> derived from the spatial covariance between the age of two dated specimens as a function of their respective pairwise distance, so that:

$$\hat{M}(x) = \sum_{i \leq n} w_i(x) a_i \quad (7)$$

where  $w_1(x), \dots, w_n(x)$  follows  $\sum_{i \leq n} w_i(x) = 1$  and minimizes

$$\sum_{i \leq n} w_i(x) \gamma(x_i - x_j) + \mu = \gamma(x - x_j) \quad (8)$$

for  $j \leq n$ , with  $\mu$  being a Lagrange multiplier so that  $\mu = \sum_{i \leq n} \gamma(x_i - x)$  and  $\gamma$  is the variogram:

$$\gamma(u) = \frac{1}{2} E(a(z) - a(z+u))^2 = \sigma^2 - c(u) \quad (9)$$

where  $a(z)$  is the age  $a$  of a specimen found at a given location  $z$  (with  $z \in W$ ),  $\sigma^2$  is the variance of  $a(z)$  and  $c(u)$  is the covariance between  $a(z)$  and  $a(z+u)$ , with any two locations in  $W$  separated by distance  $u$ .

We then modified Solow’s method<sup>89</sup> to correct for taphonomic bias, which assumes initially that the distribution of ages through time is uniform between a given age  $A_0$  when individuals are assumed to be present, and the date of arrival  $A$ . For  $n$  ages of a given time series at a given location, the estimated terminal age  $\hat{A}$  is therefore:

$$\hat{A} = A_0 + \frac{n+1}{n} \max_i(a_i - A_0) \quad (10)$$

To integrate this method into a spatial context, we estimated a preliminary age  $A_p$  across space assuming  $\hat{M}(x)$  follows a stationary random field:

$$\hat{A}_p(x) = 2\hat{M}(x) - A_0 \quad (11)$$

But this generates a spatial bias  $\hat{A}_p(x) - A(x)$ , in every  $\hat{A}_p(x)$ , so we applied a simulation-based, spatial-bias-correction procedure<sup>90</sup> to estimate the bias



generated by Eq. (11) at each  $x$  across  $W$ . The first step assumes that  $\hat{A}_p(x)$  is the ‘true’ date of the terminal event in  $x$ . Based on these  $\hat{A}_p(x)$ , we generated  $k$  age samples  $a^{(k)} = (a_1^{(k)}, \dots, a_n^{(k)})$  at the same locations  $x_1, \dots, x_n$  following the same spatial pattern and characteristics as the dated record and sampled independently from a uniform distribution on  $[A_0, \hat{A}_p(x_i)]$ . We then inferred  $\hat{A}^{(k)}(x)$ , the timing of the terminal event for the  $k$  new simulated time series and calculated an estimated total bias  $\hat{B}(x)$  across all  $k$  ages:

$$\hat{B}(x) = \frac{1}{k} \sum_k \hat{A}^{(k)}(x) - \hat{A}_p(x) \quad (12)$$

The final estimate of the timing of the terminal event of interest  $\hat{A}(x)$  is the distribution of the preliminary dates  $\hat{A}_p(x)$  for every location  $x$  corrected by  $\hat{B}(x)$ , such that:

$$\hat{A}(x) = \hat{A}_p(x) - \hat{B}(x) \quad (13)$$

Because archaeological age estimates  $a_i$  are always associated with an inherent dating uncertainty  $\sigma_i$ , we assumed that age uncertainties are Gaussian and independent<sup>91</sup> so that the probability density of the estimated age of the terminal event  $A(x)$  follows:

$$\int_{\epsilon_1, \dots, \epsilon_n} A_{\{a_1 + \epsilon_1, \dots, a_n + \epsilon_n\}}(x) \prod_i g_i(\epsilon_i) d\epsilon_1 \dots d\epsilon_n \quad (14)$$

where  $g_i$  is the density of the Gaussian random variable with mean 0 and variance  $\sigma_i^2$ , and  $A_{\{a_1 + \epsilon_1, \dots, a_n + \epsilon_n\}}(x)$  is the final estimate at a given location  $x$  for a time series of age  $a_1 + \epsilon_1, \dots, a_n + \epsilon_n$  located at  $x_1, \dots, x_n$ , respectively. We applied the same Cook and Stefanski bias-correction procedure so that the  $k$  ages are independently sampled from a uniform distribution on  $[A_0, \hat{A}_p(x_i)]$  at the same locations  $x_1, \dots, x_n$  following the same spatial pattern and characteristics as the dated record. This gives  $a^{(k)} = (a_1 + \epsilon_1^{(k)}, \dots, a_n + \epsilon_n^{(k)})$  with  $\epsilon_i^{(k)}$  independently sampled as a function of the probability density described in Eq. (14) to account for the dating uncertainty associated with each age. We then use the terminal ages  $\hat{A}^{(k)}(x) = A_{\{a_1^{(k)}, \dots, a_n^{(k)}\}}(x)$  to estimate the bias in Eq. (12) and apply this to provide a corrected timing of  $\hat{A}(x)$  for every  $x$  following Eq. (13).

**Global sensitivity analysis.** We designed a global sensitivity analysis to provide robust sensitivity measures of the probability of the time to saturation of the entire Sahul continent to variation in the underlying parameters of our stochastic model<sup>54,92</sup>; this analysis does not repeat the scenario-testing parameters (i.e., time of entry, point(s) of entry,  $K$ – $P_p$  relationship). For this global sensitivity analysis, we used the initial scenario parameters of a 50-ka entry at the southern route, the rotated parabolic relationship between  $K$  and  $P_p$ , and assuming a founding population size stochastically sampled between 1300 and 1500 people for the entry point<sup>2</sup>.

Here, we ran the cellular-automaton spatial model 1000 times, randomly sampling 12 of its parameters uniformly for each iteration based on a Latin hypercube-sampling protocol<sup>54</sup>. We set the 12 parameters to be sampled with  $\pm 50\%$  variation on the median value used in the model (except for  $N_{cat}$  and max  $N/K$  with a maximum upper bound of 0.99, and for maximum  $D_{cell}$ —see below); these 12 parameters were: (i) the maximum generational rate of population increase  $r_m$  used to parameterize the phenomenological population-dynamics model per cell (range: 0.10–0.31), (ii) the minimum maximal dispersal distances  $D_m$  estimated from the allometric prediction (11–34 km), (iii) the cell-based maximum dispersal distance modifier (max  $D_{cell}$ ), ranging from  $1 \times$  to  $5 \times$  the value set in the original model (10 cells), (iv) the cell-based minimum viable population size  $N_{MVP}$  (50–150 individuals) below which we set (v) an additional mortality parameter  $M_{MVP}$  (0.1–0.3), (vi) the hydrological resistance parameter  $\Omega$  invoking landscape-scale resistance to movement only in the driest areas of Sahul per generational time step (1.5–4.5), (vii) the beta-resampled mean mortality of a cell during a catastrophe event  $M_{cat}$  (0.38–0.99), (viii) the beta-resampled proportion of people moving between cells when a migration event occurs  $P_{mig}$  (0.17–0.50), the beta-resampled (ix) minimum and (x) maximum ratios of  $N/K$  per cell invoking an emigration event (0.15–0.45 and 0.35–0.99, respectively), (xi) a resistance modifier  $R$  that modified the relationship between landscape ruggedness and maximum dispersal probability (0.5–1.5) and (xii) the population threshold  $N_{first}$  above which we determined a cell to be occupied for the calculation of the date of first arrival in a cell (50–150 individuals).

We chose to summarise the output of each of these 1000 parameter-sampled runs of the spatial model as the time taken to achieve continental saturation (i.e., the number of years taken from initial entry to occupy every cell in Sahul). In a separate analysis, we then tested the influence of the per-model run parameter values (predictors) on the time to continental saturation (response) using a boosted-regression tree<sup>93</sup> emulator with the function `gbm.step` in the `dismo` R library<sup>94</sup>. Here, we set the error distribution family as Gaussian, the bag fraction to 0.75, the learning rate to 0.008, the tolerance to 0.0001, the maximum number of trees to 10,000 and the tree complexity to 2 (first-order interactions only). To assess the relative contribution of each of the 12 randomly sampled parameters to the

time to spatial saturation, we calculated the boosted-regression tree metrics of relative influence<sup>54</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All data are available for download at [github.com/cjabradshaw/SahulHumanSpread](https://github.com/cjabradshaw/SahulHumanSpread) (<https://doi.org/10.5281/zenodo.4453767>).

## Code availability

All R code is available for download at [github.com/cjabradshaw/SahulHumanSpread](https://github.com/cjabradshaw/SahulHumanSpread) (<https://doi.org/10.5281/zenodo.4453767>).

Received: 20 April 2020; Accepted: 2 February 2021;

Published online: 29 April 2021

## References

- Clarkson, C. et al. Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).
- Bradshaw, C. J. A. et al. Minimum founding populations for the first peopling of Sahul. *Nat. Ecol. Evol.* **3**, 1057–1063 (2019).
- Bird, M. I. et al. Early human settlement of Sahul was not an accident. *Sci. Rep.* **9**, 8220 (2019).
- O’Connell, J. F. et al. When did *Homo sapiens* first reach Southeast Asia and Sahul? *Proc. Natl Acad. Sci. USA* **115**, 8482 (2018).
- Clarkson, C. et al. Reply to comments on Clarkson et al. (2017) ‘Human occupation of northern Australia by 65,000 years ago’. *Aust. Archaeol.* **84**, 84–89 (2018).
- Kealy, S., Louys, J. & O’Connor, S. Reconstructing palaeogeography and inter-island visibility in the Wallacean Archipelago during the likely period of Sahul colonization, 65–45000 years ago. *Archaeol. Prospect.* **24**, 259–272 (2017).
- 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).
- Norman, K. et al. An early colonisation pathway into northwest Australia 70–60,000 years ago. *Quat. Sci. Rev.* **180**, 229–239 (2018).
- 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).
- Yuen, L. K. W. et al. Tracing ancient human migrations into Sahul using hepatitis B virus genomes. *Mol. Biol. Evol.* **36**, 942–954 (2019).
- Malaspinas, A.-S. et al. A genomic history of Aboriginal Australia. *Nature* **538**, 207–214 (2016).
- Tobler, R. et al. Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544**, 180–184 (2017).
- Pedro, N. et al. Papuan mitochondrial genomes and the settlement of Sahul. *J. Hum. Genet.* **65**, 875–887 (2020).
- Veth, P. Islands in the interior: a model for the colonization of Australia’s arid zone. *Archaeol. Ocean.* **24**, 81–92 (1989).
- Bird, M. I., O’Grady, D. & Ulm, S. Humans, water, and the colonization of Australia. *Proc. Natl Acad. Sci. USA* **113**, 11477–11482 (2016).
- Birdsell, J. B. Some population problems involving Pleistocene man. *Cold Spring Harb. Symp. Quant. Biol.* **22**, 47–69 (1957).
- Bowdler, S. *Sunda and Sahul: Prehistoric Studies in Southeast Asia Melanesia and Australia* (eds Allen J., Golson J. & Jones R.) 205–246 (Academic Publishing, 1977).
- Tindale, N. B. *Ecological Biogeography of Australia* (ed Keast A.) 1761–1796 (Dr W Junk Publishers, 1981).
- O’Connell, J. F. & Allen, J. The restaurant at the end of the universe: modelling the colonisation of Sahul. *Aust. Archaeol.* **74**, 5–17 (2012).
- Allen, J. & O’Connell, J. Both half right: updating the evidence for dating first human arrivals in Sahul. *Aust. Archaeol.* **79**, 86–108 (2014).
- Roberts-Thomson, J. M. et al. An ancient common origin of Aboriginal Australians and New Guinea highlanders is supported by alpha-globin haplotype analysis. *Am. J. Hum. Genet.* **58**, 1017–1024 (1996).
- Bergström, A. et al. Deep roots for Aboriginal Australian Y chromosomes. *Curr. Biol.* **26**, 809–813 (2016).
- Nagle, N. et al. Mitochondrial DNA diversity of present-day Aboriginal Australians and implications for human evolution in Oceania. *J. Hum. Genet.* **62**, 343–353 (2017).
- Brown, C. T., Liebovitch, L. S. & Glendon, R. Lévy flights in Dobe Ju/’hoansi foraging patterns. *Hum. Ecol.* **35**, 129–138 (2007).

25. 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. USA* **104**, 4765 (2007).
26. Raichlen, D. A. et al. Evidence of Lévy walk foraging patterns in human hunter-gatherers. *Proc. Natl Acad. Sci. USA* **111**, 728 (2014).
27. Hamilton, M. J. & Walker, R. S. A stochastic density-dependent model of long-term population dynamics in hunter-gatherer populations. *Evol. Ecol. Res.* **19**, 85–102 (2018).
28. Grove, M. Hunter-gatherer movement patterns: causes and constraints. *J. Anthropol. Archaeol.* **28**, 222–233 (2009).
29. Johnson, A. L. Exploring adaptive variation among hunter-gatherers with Binford's frames of reference. *J. Archaeol. Res.* **22**, 1–42 (2014).
30. Hamilton, M. J., Lobo, J., Rupley, E., Youn, H. & West, G. B. The ecological and evolutionary energetics of hunter-gatherer residential mobility. *Evol. Anthropol. Issues News Rev.* **25**, 124–132 (2016).
31. Rummukainen, M. State-of-the-art with regional climate models. *WIREs Clim. Change* **1**, 82–96 (2010).
32. Strona, G. & Bradshaw, C. J. A. Co-extinctions annihilate planetary life during extreme environmental change. *Sci. Rep.* **8**, 16724 (2018).
33. Dietzel, C. & Clarke, K. The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. *Computers, Environ. Urban Syst.* **30**, 78–101 (2006).
34. Caley, T., Roche, D. M. & Renssen, H. Orbital Asian summer monsoon dynamics revealed using an isotope-enabled global climate model. *Nat. Comm.* **5**, 5371 (2014).
35. Goosse, H. et al. Description of the Earth system model of intermediate complexity LOVECLIM version 1.2. *Geosci. Mod. Dev.* **3**, 603–633 (2010).
36. Eriksson, A. et al. Late Pleistocene climate change and the global expansion of anatomically modern humans. *Proc. Natl Acad. Sci. USA* **109**, 16089–16094 (2012).
37. Tallavaara, M., Eronen, J. T. & Luoto, M. Productivity, biodiversity, and pathogens influence the global hunter-gatherer population density. *Proc. Natl Acad. Sci. USA* **115**, 1232–1237 (2018).
38. Gould, R. A. Puntutjarpa Rockshelter and the Australian desert culture. *Anthropological papers of the American Museum of Natural History*, 54, 1–187 (1977).
39. Hewlett, B. S., van de Koppel, I. M. H. & Cavalli-Sforza, L. L. *African Pygmies* (ed Cavalli-Sforza L. L.) 65–79 (Academic Press, 1986).
40. Sutherland, G. D., Harestad, A. S., Price, K. & Lertzman, K. P. Scaling of natal dispersal distances in terrestrial birds and mammals. *Conserv. Ecol.* **4**, 1–16 (2000).
41. Fort, J., Pujol, T. & Cavalli-Sforza, L. L. Palaeolithic populations and waves of advance. *Camb. Archaeol. J.* **14**, 53–61 (2004).
42. Binford, L. R. *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data Sets*. (University of California Press, 2001).
43. Reed, D. H., O'Grady, J. J., Ballou, J. D. & Frankham, R. The frequency and severity of catastrophic die-offs in vertebrates. *Anim. Conserv.* **6**, 109–114 (2003).
44. Vahdati, A. R., 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).
45. Veth, P. Breaking through the radiocarbon barrier: Madjedbebe and the new chronology for Aboriginal occupation of Australia. *Aust. Archaeol.* **83**, 165–167 (2017).
46. Hiscock, P. Discovery curves, colonisation and Madjedbebe. *Aust. Archaeol.* **83**, 168–171 (2017).
47. Wood, R. Comments on the chronology of Madjedbebe. *Aust. Archaeol.* **83**, 172–174 (2017).
48. Dortch, J. & Malaspinas, A.-S. Madjedbebe and genomic histories of Aboriginal Australia. *Aust. Archaeol.* **83**, 174–177 (2017).
49. Williams, A. N. A new population curve for prehistoric Australia. *Proc. R. Soc. Lond. B* **280**, 20130486 (2013).
50. Williams, A. N., Ulm, S., Cook, A. R., Langley, M. C. & Collard, M. Human refugia in Australia during the Last Glacial Maximum and Terminal Pleistocene: a geospatial analysis of the 25–12 ka Australian archaeological record. *J. Archaeol. Sci.* **40**, 4612–4625 (2013).
51. Lambeck, K. & Chappell, J. Sea level change through the last glacial cycle. *Science* **292**, 679–686 (2001).
52. Cosgrove, R. Forty-two degrees south: the archaeology of Late Pleistocene Tasmania. *J. World Prehist.* **13**, 357–402 (1999).
53. Davidson, I. & Roberts, D. *Turning Points in Australian History* (eds Roberts D. A. & Crotty M.) 18–31 (UNSW Press, 2008).
54. Proust, T. A. A. et al. An efficient protocol for the global sensitivity analysis of stochastic ecological models. *Ecosphere* **7**, e01238 (2016).
55. Crabtree, S. A. et al. Landscape rules predict optimal super-highways for the first peopling of Sahul. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-021-01106-8> (2021).
56. Finlayson, C. et al. The *Homo* habitat niche: using the avian fossil record to depict ecological characteristics of Palaeolithic Eurasian hominins. *Quat. Sci. Rev.* **30**, 1525–1532 (2011).
57. Hope, J. & Hope, G. *The Origin of the Australians* (eds Kirk R. L. & Thorne A. G.) 29–54 (Institute of Aboriginal Studies, 1976).
58. Lourandos, H. *Continent of Hunter-Gatherers: New Perspectives in Australian Prehistory*. (Cambridge University Press, 1997).
59. Clarkson, C. et al. The archaeology, chronology and stratigraphy of Madjedbebe (Malakunanja II): a site in northern Australia with early occupation. *J. Hum. Evol.* **83**, 46–64 (2015).
60. De Deckker, P. et al. Climatic evolution in the Australian region over the last 94 ka—spanning human occupancy—, and unveiling the Last Glacial Maximum. *Quat. Sci. Rev.* **249**, 106593 (2020).
61. Benjamin, J. et al. Aboriginal artefacts on the continental shelf reveal ancient drowned cultural landscapes in northwest Australia. *PLoS ONE* **15**, e0233912 (2020).
62. Bae, C. J., Douka, K. & Petraglia, M. D. On the origin of modern humans: Asian perspectives. *Science* **358**, eaai9067 (2017).
63. Haber, M. et al. A rare deep-rooting D0 African Y-chromosomal haplogroup and its implications for the expansion of modern humans out of Africa. *Genetics* **212**, 1421–1428 (2019).
64. Pagani, L. et al. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature* **538**, 238–242 (2016).
65. Bradshaw, C. J. A. et al. Stochastic models support rapid peopling of Late Pleistocene Sahul. *GitHub::SahulHumanSpread*, <https://doi.org/10.5281/zenodo.4453767> (2021).
66. Amante, C. & Eakins, B. W. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. 19 (2009).
67. Grant, K. M. et al. Sea-level variability over five glacial cycles. *Nat. Comm.* **5**, 5076 (2014).
68. Reeves, J. M. et al. The sedimentary record of palaeoenvironments and sea-level change in the Gulf of Carpentaria, Australia, through the last glacial cycle. *Quat. Int.* **183**, 3–22 (2008).
69. Traill, L. W., Bradshaw, C. J. A. & Brook, B. W. Minimum viable population size: a meta-analysis of 30 years of published estimates. *Biol. Conserv.* **139**, 159–166 (2007).
70. Frankham, R., Bradshaw, C. J. A. & Brook, B. W. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Conserv.* **170**, 56–63 (2014).
71. Bradshaw, C. J. A. & Brook, B. W. *Conservation Biology for All* (eds Sodhi N. S. & Ehrlich P. R.) 313–334 (Oxford University Press, 2010).
72. Bell, T. *The Princeton Guide to Ecology*. 239–246 (Princeton University Press, 2009).
73. Turchin, P. *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. (Princeton University Press, 2003).
74. Wilby, R. L. & Wigley, T. M. L. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geogr. Earth Environ.* **21**, 530–548 (1997).
75. Lorenz, D. J., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C. & Williams, J. W. Downscaled and debiased climate simulations for North America from 21,000 years ago to 2100 AD. *Sci. Data* **3**, 160048 (2016).
76. Timmermann, A. & Friedrich, T. Late Pleistocene climate drivers of early human migration. *Nature* **538**, 92–95 (2016).
77. Hempton, G. P., Archibald, S. & Bond, W. J. A continent-wide assessment of the form and intensity of large mammal herbivory in Africa. *Science* **350**, 1056–1061 (2015).
78. Birdsell, J. B. Some environmental and cultural factors influencing the structuring of Australian Aboriginal populations. *Am. Nat.* **87**, 171–207 (1953).
79. McNiven, I. J. *Australian Coastal Archaeology* (eds Hall H. J. & McNiven I. J.) 157–168 (ANH Publications, 1999).
80. Hiscock, P. *Archaeology of Ancient Australia*. (Routledge, 2008).
81. Saltré, F., Chuine, I., Brewer, S. & Gaucherel, C. A phenomenological model without dispersal kernel to model species migration. *Ecol. Model.* **220**, 3546–3554 (2009).
82. Marwick, B., Johnson, A., White, D. & Eff, E. A. *binford: Binford's Hunter-Gatherer Data*, R package version 0.1.0 CRAN.R-project.org/package=binford (2016).
83. Riley, S. J., DeGloria, S. D. & Elliot, R. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermount. J. Sci.* **5**, 23–27 (1999).
84. Chamberlain, A. Archaeological demography. *Hum. Biol.* **81**, 275–287 (2009).
85. Baddeley, A., Rubak, E. & Turner, R. *Spatial Point Patterns: Methodology and Applications with R*. (Chapman and Hall/CRC Press, 2015).
86. Solow, A. R., Roberts, D. L. & Robbirt, K. M. On the Pleistocene extinctions of Alaskan mammoths and horses. *Proc. Natl Acad. Sci. USA* **103**, 7351–7353 (2006).

87. Saltré, F. et al. Climate-human interaction associated with southeast Australian megafauna extinction patterns. *Nat. Comm.* **10**, 5311 (2019).
88. Wackernagel, H. *Multivariate Geostatistics: An Introduction with Applications*. (Springer, 2003).
89. Solow, A. R. Inferring extinction from sighting data. *Ecology* **74**, 962–964 (1993).
90. Cook, J. R. & Stefanski, L. A. Simulation-extrapolation estimation in parametric measurement error models. *J. Am. Stat. Assoc.* **89**, 1314–1328 (1994).
91. Bradshaw, C. J. A., Cooper, A., Turney, C. S. M. & Brook, B. W. Robust estimates of extinction time in the geological record. *Quat. Sci. Rev.* **33**, 14–19 (2012).
92. Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q. & Birkholzer, J. T. Making sense of global sensitivity analyses. *Comput. Geosci.* **65**, 84–94 (2014).
93. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
94. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species distribution modeling. R package version 1.1-4. [CRAN.R-project.org/package=dismo](http://CRAN.R-project.org/package=dismo). (2017).
95. Bowdler, S. 'Human occupation of northern Australia by 65,000 years ago' (Clarkson et al. 2017): a discussion. *Aust. Archaeol.* **83**, 162–163 (2017).

## Acknowledgements

This study was supported by the Australian Research Council through a Centre of Excellence Grant (CE170100015) to R.G.R., S.U., M.I.B., Z.J., C.J.A.B., S.O'C., S.G.H. and L.S.W., fellowships to S.U. (FT120100656), M.I.B. (FL140100044), R.G.R. (FL130100116), Z.J. (FT150100138), C.C. (FT160100242), L.S.W. (FT180100407) and B. L. (FT170100448) and an Australian Government Research Training Program Award to K.N. We acknowledge the Indigenous Traditional Owners of the land on which Flinders University is built—the Kaurna people of the Adelaide Plains.

## Author contributions

C.J.A.B., K.N. and F.S. conceptualized the paper, and C.J.A.B. and F.S. designed the research. C.J.A.B. did the analysis and sourced the data. S.U., A.N.W., K.N. and M.L. sourced the archaeological data for the comparison layer. F.S. and J.C. generated the

archaeological map layers. C.J.A.B., F.S., S.U., A.N.W. and K.N. wrote the paper. All other co-authors contributed substantially to developing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21551-3>.

**Correspondence** and requests for materials should be addressed to C.J.A.B.

**Peer review information** *Nature Communications* thanks Michael Petraglia, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021