



# Biogeography of Crop Progenitors and Wild Plant Resources in the Terminal Pleistocene and Early Holocene of West Asia, 14.7–8.3 ka

RESEARCH PAPER

JOE ROE

AMAIA ARRANZ-OTAEGUI

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

Ju[ ubiquity press

## ABSTRACT

This paper presents the first continuous, spatially-explicit reconstructions of the palaeodistributions of 65 plant species found regularly in association with early agricultural archaeological sites in West Asia, including the progenitors of the first crops. We used machine learning to train an ecological niche model of each species based on its present-day distribution in relation to climate and environmental variables. Predictions of the potential niches of these species at key stages of the Pleistocene–Holocene transition could then be derived from these models using downsampled data from palaeoclimate simulations. Our models performed well against independent contemporary test data, but their ability to predict the occurrence of specific species at archaeological sites was much more variable, probably reflecting a tendency of the method to underestimate the species' fundamental niche. Nevertheless, the majority of species are predicted to have had more restricted geographic distributions under past climate conditions compared to today. Crop progenitors and several wild food species are modelled to have been concentrated in the Levant and, to a lesser extent, Cyprus and Western Anatolia, suggesting these regions may have served as glacial refugia. The average size of species' niche shrunk by an average of c. 25% from the terminal Pleistocene to the Early Holocene, indicating that economically significant plants were adapted to cryo-arid conditions and did not, as often assumed, initially respond positively to the 'ameliorated' climate of the Holocene.

---

## CORRESPONDING AUTHOR:

**Joe Roe**University of Copenhagen, DK  
jar@hum.ku.dk

---

## KEYWORDS:

human ecology; climate niche modeling; neolithisation; Early Neolithic agriculture

---

## TO CITE THIS ARTICLE:

Roe, J and Arranz-Otaegui A. 2026. Biogeography of Crop Progenitors and Wild Plant Resources in the Terminal Pleistocene and Early Holocene of West Asia, 14.7–8.3 ka. *Open Quaternary*, 12: 2, pp. 1–23. DOI: <https://doi.org/10.5334/oq.163>

## LAY SUMMARY

We used machine learning to reconstruct the past ranges of 65 plant species that were used by prehistoric societies in West Asia around 10,000 years ago, including the ancestors of the first crops. Our results suggest that under the cold, arid conditions of that time these were much less widespread than they are now, and that they became even scarcer just before the emergence of agriculture. The Mediterranean coast of the Levant and, to our surprise, Cyprus and Western Anatolia appear to have been refugia for many economically important plants. Our data introduces an as-yet unexplored new line of evidence on early plant domestication and has implications for how archaeologists interpret the botanical remains they recover from sites, as for example the recovery of a plant from outside its ‘native range’ is often taken as evidence that it was being managed or cultivated. It also gives insights into the broader ecosystems from which the world’s first agricultural systems emerged.

Our approach was based on ‘ecological niche modelling’, a technique widely used by ecologists to model the distribution of species now or predict where they will be in the future (for example, under climate change) but which can also be combined with palaeoclimate data to ‘retro-dict’ where they were in the past. In our case, we used palaeoclimate data from the ‘CHELSA-TraCE21k’ experiment, which took one of the global models used to predict future climate change and essentially turned it backwards. Combining this with large open datasets of where plants have been observed growing today, in relation to the same set of environmental variables, we were able to model a large number of species which we know are found at early agricultural sites in West Asia. We also used data on the species observed at archaeological sites to test the predictions of our models, though this had mixed results – something we intend to probe further in future research.

## 1. INTRODUCTION

The Pleistocene–Holocene transition in West Asia marked a turning point in global environmental history, as humans brought the first plants under cultivation and began modifying surrounding ecosystems to support their own subsistence. West Asia is part of the native range of a remarkable number of domesticable plant species, including wild relatives of wheat, barley, peas, lentils, and other crops of global importance ([Harlan and Zohary 1966](#); [Diamond 2002](#); [Zohary, Weiss, and Hopf 2012](#)). These species supported uniquely dense and complex Late Epipalaeolithic (15–11.7 ka) societies ([Bar-Yosef 1998](#); [Maher, Richter, and Stock 2012](#)) based on foraging ([Harris and Hillman 1989](#); [Colledge 2001](#);

[Weiss et al. 2004](#)) and eventually plant management and pre-domestication cultivation ([Colledge 2001](#); [Weiss, Kislev, and Hartmann 2006](#); [Harris 2007](#); [Willcox, Fornite, and Herveux 2008](#)). The first agro-ecosystems emerged as these plants were domesticated in the Pre-Pottery Neolithic period (11.7–8.5 ka) and were shaped by the broader ecosystems in which they were embedded.

Decades of research in archaeobotany and zooarchaeology have reconstructed the subsistence economies of Late Epipalaeolithic and Neolithic sites in great detail. Together with studies of other environmental archaeological records and a variety of palaeoclimate archives ([Jones et al. 2019](#)), they also tell us much about the environments surrounding these settlements. However, each of these sources of evidence is subject to the wide variety of taphonomic and recovery biases that are inherent in any direct record of the past. They are also, by definition, records of the (human) environment at particular times and places. Interpolating these snapshots to give a holistic picture of the regional ecologies is not straightforward – to date, it has tended to rely on non-explicit, inductive modelling. The majority are also filtered through human action, producing a mixed signal that makes it difficult to disentangle anthropic effects from the background of environmental change in this period of rapid climatic alteration.

In this paper we present a complementary, deductive approach based on ecological niche modelling. Rather than inferring environmental conditions from preserved physical evidence, we predict the potential niche of individual species relevant to human subsistence based on a model of their current environmental niche and simulations of past palaeoclimate. Though hypothetical, this gives us an independent line of evidence on past ecologies that is independent of the environmental archaeological and palaeoclimatic records. This means that the ancient data can be reserved for assessing the model’s ability to ‘hindcast’ past conditions. In this sense, discrepancies between the two records are perhaps the most interesting result, as they indicate processes affecting one or both records that are not fully accounted for and therefore generate new questions. Our computational approach is also readily scaled up, allowing us to model spatially-explicit palaeodistributions for a large number of species, for the whole region, under multiple reconstructed climate scenarios.

## 2. BACKGROUND

The transition to agriculture represents one of the most fundamental changes in human history. West Asia is one of the regions where this process has been studied in the most detail: decades of research have traced the gradual development of a Neolithic way of life and the changes

that occurred in the plant species and their geographical distribution as a result. Although archaeobotanical assemblages can be biased due to issues of preservation, sampling, recovery techniques, and lab procedures (Dennell 1976; Hastorf and Popper 1988)—and although they include not just food remains but plant resources that were used for other purposes or arrived at the site accidentally (Hastorf and Popper 1988)—large-scale studies have still revealed coherent patterns in the exploitation of plants over time (Colledge, Conolly, and Shennan 2004; Arranz-Otaegui et al. 2016).

The possibility of an abrupt, geographically-constrained process of plant domestication was proposed in the 1990s (Hillman and Davies 1990, 1992; Heun et al. 1997; Özkan et al. 2002) and developed as an explanatory model in the 2000s (Lev-Yadun, Gopher, and Abbo 2000a; Gopher, Abbo, and Yadun 2001; Abbo et al. 2005). As part of this model, some authors (Lev-Yadun, Gopher, and Abbo 2000a; Gopher, Abbo, and Yadun 2001; Abbo, Lev-Yadun, and Gopher 2010, 2012) argued that eight plant species, collectively referred to as ‘founder crops’ or the ‘Neolithic crop package’ (Zohary and Hopf 1988) were selected and domesticated once, without any phase of pre-domestication cultivation (Abbo, Lev-Yadun, and Gopher 2011, 177). This process could have been rapid under strong artificial selection (Hillman and Davies 1990, 1992) and may have occurred in a single region or ‘core area’ – generally located in southeast Turkey (Ladizinsky and Adler 1976; Heun et al. 1997; Özkan et al. 2002; Özkan et al. 2005; Mori 2003; Luo et al. 2007). From this single point of origin, it was supposed that domesticated or semi-domesticated plants radiated outwards to other regions (Abbo et al. 2006; Kilian et al. 2007; Özkan et al. 2011).

The ‘short gestation’ paradigm was challenged by others (Helbæk 1969; Harris 1989; Kislev 1989; Colledge 2001; Weiss et al. 2004; Willcox, Fornite, and Herveux 2008; Fuller et al. 2018). Helbæk (in Kirkbride 1966) argued that before the appearance of domesticated plants, a phase of cultivation of wild seeds must have taken place. The existence of a phase of cultivation of morphologically wild cereals or ‘pre-domestication cultivation’ was identified in the archaeological record through the study of plant domestication traits such as grain size and shattering v. non-shattering rachises. The archaeobotanical evidence shows that during the Pre-Pottery Neolithic A (PPNA) cereals exhibited sizes similar to those recorded in domestic species, but their dispersal mechanism was still the same as the one present in morphologically-wild species (i.e. shattering, see Kirkbride 1966; Kislev 1989; Hillman et al. 2001; Colledge 2001; Willcox, Fornite, and Herveux 2008). This evidence suggested that wild cereal stands could have been cultivated for as much as a thousand years before non-shattering domestic forms became prevalent in the archaeological record (Tanno and Willcox 2006, 2012; Arranz-Otaegui et al. 2016). Additional

archaeobotanical (Colledge 2001; Willcox, Fornite, and Herveux 2008; Willcox, Buxo, and Herveux 2009; Riehl, Zeidi, and Conard 2013; Arranz-Otaegui et al. 2016; Weide et al. 2018; Douché and Willcox 2018; Whitlam et al. 2018) and genetic data (Badr et al. 2000; Molina-Cano et al. 2005; Kilian et al. 2007; Özkan et al. 2011; Job and Botigué 2023) in recent years has further challenged the short-gestation model to explain the origins of plant domestication and agriculture in West Asia.

Similarly, the concept of a limited set of eight ‘founder crops’ (Zohary and Hopf 1988) that were the first species cultivated, domesticated and then spread as the basis of Neolithic agricultural systems, is not supported by the latest evidence. Our previous analyses of the composition of available archaeobotanical datasets shows that these crops were of marginal importance during the Epipalaeolithic period (Arranz-Otaegui et al. 2018) and that Neolithic subsistence did not rely either solely or primarily on the exploitation of these species (Arranz-Otaegui and Roe 2023). Instead, multiple species of grasses, legumes, fruits, nuts, and other plants were exploited over the Late Pleistocene-Early Holocene transition in southwest Asia.

## 2.1 BIOGEOGRAPHY AND AGRICULTURAL ORIGINS

The study of the natural distribution of the progenitors of domesticated crops has been a central part of discussions on the origins of agriculture and plant domestication from the beginning. von Humboldt (1807) acknowledged the importance of the natural distribution of wild species to explain the origin and domestication of crops like spelt and rye. Candolle (1886) integrated the study of plant ecology and biogeography and influenced Darwin (1859), who later reflected in detail about the geographical distribution of plants and species diversity. At that time, there was intense debate about whether there were single or multiple “centres of creation” of species. Researchers aimed to evaluate whether plant and animal species emerged in the same locations where they were currently distributed.

After Darwin, this early interest in crop origins evolved into more specific discussions about the “centres of plant domestication”. Vavilov (1926) was among the first to seek to determine the number of regions in which plants had been independently domesticated (Harris 1990). His main method was ‘differential phytogeography’: he classified the variation within a crop and established the regions of maximum diversity, to locate the geographic regions in which crops originated. Using this method, Vavilov suggested that there were at least eight centres of origin. His work was later criticised by Harlan (1971), who argued that ‘centres of origin’ and ‘centres of diversity’ had to be separated. For Harlan, a ‘centre’ was as an “area in which things originate and out of which things are dispersed” (Harlan 1971, 468), and he suggested

that three main centres of origin of domesticated crops existed. He further indicated that Vavilov's approach to the question was simplistic and that more data proxies had to be considered (e.g. archaeology, history, geology), an approach more in the tradition of Candolle (1886). Indeed, the inclusion of archaeobotany and genetics in the last decades, together with the study of wild relative distributions has been fundamental in characterising the origins of agriculture (Fuller and Colledge 2008). As a result of modern interdisciplinary studies, the number of recognised centres of plant domestication has increased considerably, from the three centres suggested by Harlan in 1971 to the six to eight centres argued for in the 1990s (Smith 1995) and up to as much as 24 potential centres reported in 2009 (Purugganan and Fuller 2009; see also Fuller 2010).

Biogeographic research into the centres of origin and/or domestication of crops has also long informed broader understanding of the process of agricultural origins. In *Man Makes Himself*, Childe (1936) correctly located the centre of origin of European agriculture in the 'Fertile Crescent' of West Asia (unlike for example Pumpelly (1908) before him). This was not based on the region's prehistoric archaeological record, which at the time had only been cursorily explored. Instead he was guided to the region by biogeographic work by Vavilov and Peake and Fleure (1927); only later was this prediction validated by archaeological work on the Epipalaeolithic and Neolithic of Palestine (Boyd 2018). In subsequent decades, the search for more precise origin zones of specific domestic plants relied on the assumption that "the locus of domestication of a wild plant would presumably be within its area of original distribution in the wild state" (Butzer 1971; paraphrasing Helbæk 1959) – and that this "natural habitat" has not changed significantly over the last 12,000 years (Butzer 1971).

Contemporary research on crop origins was pioneered by Harlan and Zohary, who compared the current distribution of the wild progenitors of domesticated plants in southwest Asia (Harlan and Zohary 1966; D. Zohary 1969; M. Zohary 1973; Zohary and Hopf 1973) to the rapidly-expanding archaeobotanical record (Harlan 1971, 1977; see also Zohary and Hopf 1988; Harlan and Zohary 1966). They both were interested in evaluating which were the wild ancestors of domesticated crops and studying their natural distribution to understand their domestication process (D. Zohary 1969; M. Zohary 1973; Zohary and Spiegel-Roy 1975). Indeed, the natural distribution of the wild relatives of domestic plant species was later used as a criterion to infer 'pre-domestic cultivation' in the archaeological record. For example, the presence of seeds of chickpea at Jericho led M. Hopf (1986) to interpret the remains as cultivars, as the natural distribution of the wild form of chickpea was located further to the north. The same rationale was applied to the einkorn remains found at several Pre-Pottery Neolithic

sites in the southern Levant (Hopf 1969; Colledge 2001), as the wild progenitors of this species was thought to be restricted to the northern Levantine area (Heun et al. 1997; Zohary, Weiss, and Hopf 2012). The same idea—presence of plants outside their natural range—has been repeated in the literature more recently by several other authors (Tanno and Willcox 2006; Willcox, Fornite, and Herveux 2008; Hillman et al. 2001).

Despite the importance of biogeography in the development and validation of hypotheses regarding the origins of agriculture, there have been few studies of the wild range of specific crop progenitors or other relevant plant species (cf. for domestic animals, e.g. Yeomans, Martin, and Richter 2017). Observations regarding translocation or range expansion must therefore rely on a relatively rough and ahistoric notion of a species' 'natural distribution' – that is, one based primarily on contemporary or recent-historic occurrences. Yet we know there has been considerable climatic and environmental change in West Asia since the terminal Pleistocene (Jones et al. 2019), so it is very unlikely that these ranges were in fact static. Reconstructions of broader environments have attempted to trace their fluctuations through time, either for the entire region (e.g. van Zeist and Bottema 1991; Hillman in Moore, Hillman, and Legge 2000) or parts of it (e.g. Cordova 2007), but these are at the level of the vegetation zone rather than individual species. They also invariably rely on what might be called 'expert interpolation' (where the author composes a map based on his or her own knowledge of the relevant data) rather than an explicit modelling process. This makes it difficult, if not impossible, for users of such reconstructions to understand exactly how they were derived or what could explain, for example, the significant discrepancies between the predictions of different experts.

## 2.2 ECOLOGICAL NICHE MODELLING IN ARCHAEOLOGY

Ecological niche modelling or species distribution modelling is widely used by ecologists to mathematically predict the geographic distribution of a species from a sample of occurrences (Franklin and Miller 2009; Sillero et al. 2021). Essentially, it involves combining records of where an organism has been observed with environmental data (climate, topography, etc.) for those locations to model the range of environmental values at which that species – its environmental niche. This model can then be used to predict the potential niche of the organism in question either in the same or a different environment. Townsend Peterson and Soberón (2012) suggests reserving the term 'species distribution modelling' for when the method is used to recover the verifiable range of a species in a real and existing environment, and using 'ecological niche modelling' as the broader term covering hypothetical or predictive applications – a convention we follow here when referring to predictive or 'hindcast'

models of past niches. Within this overarching framework, ecological niche modelling encompasses a wide range of applications and a variety of potential environmental predictors, modelling approaches, and methodologies, which we will not attempt to review here.

Ecological niche modelling has long been of interest to archaeologists as both a means of exploring the biological niche of humans and for reconstructing the past environments they inhabited ([Polly and Eronen 2011](#); [Franklin et al. 2015](#)). In the first sense, it has been used most extensively to model the niche of humans and other hominin species (e.g. [Benito et al. 2017](#); [Yousefi et al. 2020](#); [Banks et al. 2021](#); [Yaworsky, Hussain, and Riede 2024](#); [Yaworsky, Nielsen, and Nielsen 2024](#); [Guran et al. 2024](#)), especially in the Palaeolithic. This overlaps with what archaeologists usually call generically ‘predictive modelling’ ([Verhagen and Whitley 2020](#))—or more precisely ‘site distribution modelling’—which is essentially the same approach as (and often borrows methodologies from) ecological niche modelling but applied to the occurrence of archaeological sites. Here what is modelled is not strictly a biological niche alone, but also aspects of human geography, taphonomy, and archaeological visibility. These applications can be distinguished from ‘palaeoecological niche modelling’, where the object of model remains, as in ecology, a non-human biological niche.

[Franklin et al. \(2015\)](#) review palaeoecological niche modelling and advocate for its greater adoption in environmental archaeology. In an early application to West Asia, [Conolly et al. \(2012\)](#) used the occurrence of wild and domestic *Bos* remains at prehistoric archaeological sites to map the evolving niche of cattle over the Pleistocene–Holocene transition. It has been used to model the availability of fauna exploited by humans at wider scales (e.g. [de Andrés-Herrero, Becker, and Weniger 2018](#); [Yaworsky, Hussain, and Riede 2023](#)) and, in a West Asian context, of foraged plant resources in the landscape around the Neolithic sites on the Konya Plain ([Collins et al. 2018](#)). Modelling the spread of crops has been another significant archaeological application (e.g. [Krzyszanska et al. 2022](#); [Krzyszanska 2023](#)), though not as yet applied to West Asia.

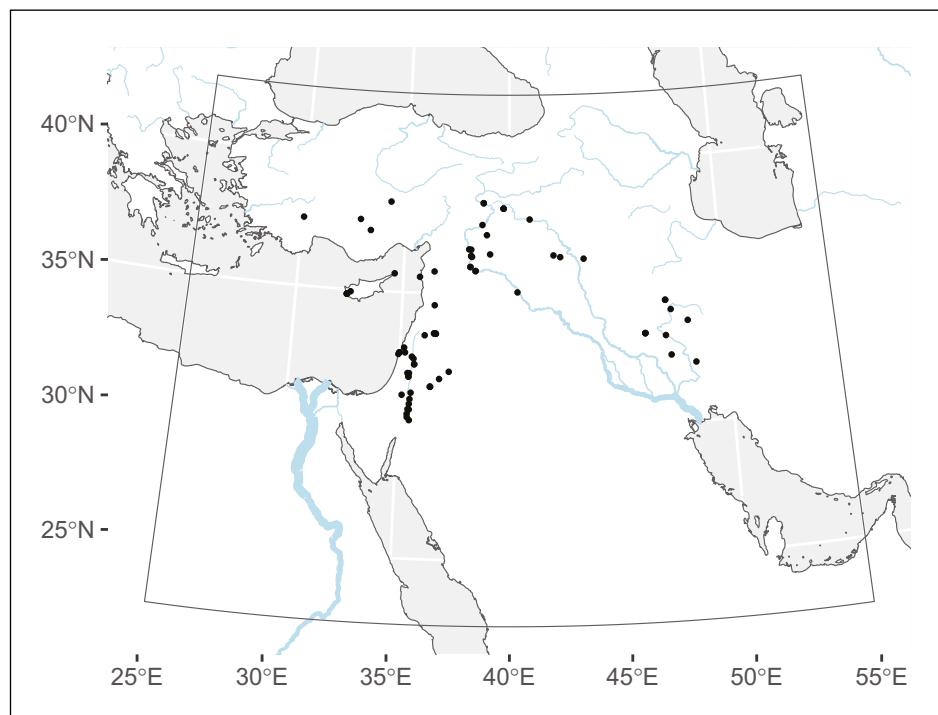
In the majority of studies to date (palaeo)ecological niche modelling has been applied to archaeological data in an ‘inductive’ fashion, i.e. faunal and botanical remains from ancient sites are used as the occurrence dataset for training a model using either past or present environmental data. However, both the zooarchaeological and archaeobotanical records are sparse and subject to a complex array of depositional, taphonomic and recovery biases, many of which are not fully understood and/or cannot be corrected for. This means that while the archaeological attestation of the presence of a species might generally be relied upon, it is highly unlikely that its absence is representative of true past distributions.

The alternative approach is to train the model using contemporary occurrence and environmental data and then use palaeoenvironmental data to ‘hindcast’ its predictions backwards in time. Like [Franklin et al. \(2015\)](#), we view the hindcasting approach as more promising, because training datasets for both occurrences and environment are far more readily available, complete and reliable for the present than the past. There is some scepticism in the ecological niche modelling literature about the ability of such models to make accurate predictions in unknown environments (like the past, [Franklin et al. 2015](#)), but here the hindcasting approach also presents an opportunity: it reserves archaeological occurrence data as an independent dataset that can be used to assess the retrodictive performance of the model. This possibility was suggested by [Franklin et al. \(2015\)](#) but to our knowledge our study represents the first attempt to actually do so.

The major practical limitation of the hindcasting approach is that it relies on spatially explicit, high resolution palaeoenvironmental surfaces with continuous coverage of the region and periods of interest. Until recently, this has not been widely available for most applications, which is perhaps why only a minority of studies use it (cf. [Krzyszanska et al. 2022](#); [Yaworsky, Hussain, and Riede 2023](#)). In this study, we are able to take advantage of the increasing availability of high resolution, global palaeoclimate data derived from simulation experiments with general circulation models of climate ([Brown et al. 2018](#); [Brown et al. 2020](#); [Karger et al. 2023](#)).

### 3. DATA AND MODEL

The aim of our study was to model the biogeography of species relevant to human subsistence economies in West Asia (excluding the Southern Arabian peninsula, see [Figure 1](#)) during the archaeological Late Epipalaeolithic (15–11.7 ka) and Pre-Pottery Neolithic (11.7–8.3 ka) periods. Based on current understandings, we assume that plant-based subsistence during this period was broad, geographically- and temporally-varied, and reflects a gradual, geographically decentred, and nonlinear transition to greater reliance on cultivars (i.e. agriculture, see Section 2). Our starting point was a list of 68 taxa ([Table 1](#)) comprising the identified species observed at at least 3 Late Epipalaeolithic/Pre-Pottery Neolithic sites, according to our previous study of the regional archaeobotanical data ([Arranz-Otaegui and Roe 2023](#)). This was based on dataset collated from three previously published regional archaeobotanical databases: ADEMNES ([Riehl and Kümmel 2005](#)), ORIGINS ([Wallace, Livarda, et al. 2018](#)), and COMPAG ([Lucas and Fuller 2018; Fuller et al. 2018](#); based on [Colledge, Conolly, and Shennan 2004](#); [Shennan and Conolly 2007](#)). We did not attempt to distinguish between the



**Figure 1** Map of the study region (West Asia, grey box) with locations of Late Epipalaeolithic and Pre-Pottery Neolithic archaeobotanical assemblages.

TAXON	OCCURRENCES		MODEL			
	ARCH.	CUR.	ACC.	ROC-AUC	SENS. CUR.	SENS. ARCH.
<i>Aegilops crassa</i>	4	223	0.98	0.98	0.26	0.00
<i>Aegilops speltoides</i>	0	448	0.99	0.99	0.77	—
<i>Aegilops tauschii</i>	0	1257	0.97	0.99	0.84	—
<i>Aizoanthemopsis hispanica</i> <sup>1</sup>	4	762	0.99	1.00	0.87	0.50
<i>Ammi majus</i>	4	10337	0.97	0.99	0.98	0.00
<i>Androsace maxima</i>	17	4946	0.94	0.98	0.90	0.00
<i>Arenaria serpyllifolia</i>	3	101617	0.98	0.99	1.00	0.00
<i>Arnebia decumbens</i>	21	617	0.97	0.97	0.59	0.05
<i>Arnebia linearifolia</i>	13	239	1.00	1.00	0.83	0.00
<i>Atriplex prostrata</i>	4	80186	0.99	0.99	0.99	0.00
<i>Avena sterilis</i>	5	15906	0.96	0.99	0.98	0.40
<i>Bassia arabica</i>	4	262	1.00	0.99	0.88	0.00
<i>Bolboschoenus glaucus</i>	5	418	0.98	0.97	0.63	0.00
<i>Bolboschoenus maritimus</i> <sup>2</sup>	31	53813	0.98	1.00	0.99	0.00
<i>Brachypodium distachyon</i>	5	13800	0.98	0.99	0.99	0.20
<i>Bromus sterilis</i>	3	89764	0.99	0.99	1.00	0.00
<i>Buglossoides arvensis</i>	23	23004	0.95	0.98	0.98	0.09
<i>Buglossoides tenuiflora</i>	26	399	0.99	0.98	0.79	0.04
<i>Capparis spinosa</i>	4	8617	0.96	0.99	0.95	0.00
<i>Carex divisa</i>	9	7400	0.96	0.99	0.97	0.00
<i>Chenopodium album</i>	5	184412	0.99	0.99	1.00	0.00

(Contd.)

TAXON	OCCURRENCES		MODEL			
	ARCH.	CUR.	ACC.	ROC-AUC	SENS. CUR.	SENS. ARCH.
<i>Cicer reticulatum</i> <sup>3</sup>	3	52	1.00	1.00	1.00	0.00
<i>Citrullus colocynthis</i>	4	1823	0.94	0.96	0.71	0.00
<i>Crithopsis delileana</i>	3	430	1.00	1.00	0.93	0.00
<i>Euclidium syriacum</i>	3	338	0.98	0.96	0.31	0.00
<i>Ficus carica</i>	8	64440	0.98	0.99	0.99	0.38
<i>Fumaria densiflora</i>	4	3704	0.96	0.99	0.95	0.00
<i>Gypsophila elegans</i>	3	1613	0.96	0.99	0.83	0.00
<i>Gypsophila pilosa</i>	5	348	0.99	0.99	0.69	0.00
<i>Gypsophila vaccaria</i> <sup>4</sup>	8	1177	0.95	0.96	0.69	0.00
<i>Henrardia pubescens</i>	3	1 <sup>5</sup>	—	—	—	—
<i>Hordeum bulbosum</i>	4	4134	0.97	0.99	0.94	0.00
<i>Hordeum murinum</i>	5	74770	0.98	0.99	1.00	0.00
<i>Hordeum spontaneum</i> <sup>6</sup>	76	3098	0.99	1.00	0.98	0.21
<i>Lathyrus aphaca</i>	4	18652	0.96	0.99	0.98	0.00
<i>Lathyrus oleraceus</i> <sup>7</sup>	8	819	0.98	0.99	0.78	0.00
<i>Lathyrus sativus</i>	4	2120	0.93	0.97	0.77	0.00
<i>Lepidium perfoliatum</i>	3	2764	0.94	0.98	0.84	0.00
<i>Linum bienne</i> <sup>8</sup>	14	12824	0.98	0.99	0.99	0.07
<i>Lolium rigidum</i>	5	10620	0.97	0.99	0.98	0.00
<i>Lolium temulentum</i>	3	4463	0.94	0.98	0.94	0.00
<i>Medicago radiata</i>	20	834	0.98	1.00	0.90	0.30
<i>Phalaris paradoxa</i>	3	3287	0.96	0.99	0.94	0.00
<i>Phragmites australis</i>	4	373066	0.99	0.99	1.00	0.00
<i>Pistacia atlantica</i>	6	2419	0.98	1.00	0.93	0.33
<i>Poa bulbosa</i>	5	30769	0.96	0.98	0.99	0.20
<i>Polygonum arenarium</i> <sup>9</sup>	7	25 <sup>5</sup>	—	—	—	—
<i>Polygonum corrigioloides</i>	6	12 <sup>5</sup>	—	—	—	—
<i>Prosopis farcta</i>	5	2512	0.99	1.00	0.96	0.00
<i>Rumex pulcher</i>	6	17606	0.97	0.99	0.99	0.00
<i>Salsola kali</i>	6	14701	0.98	1.00	0.99	0.00
<i>Salvia absconditiflora</i> <sup>10</sup>	3	128	1.00	1.00	0.81	0.00
<i>Secale cereale</i>	4	14813	0.95	0.98	0.98	0.00
<i>Secale strictum</i> <sup>11</sup>	3	162	0.99	0.98	0.62	0.00
<i>Suaeda fruticosa</i>	3	577	0.99	0.99	0.76	0.00
<i>Taeniatherum caput-medusae</i>	4	1969	0.97	0.99	0.90	0.00
<i>Triticum aestivum</i> <sup>12</sup>	4	217	0.99	0.96	0.50	0.00
<i>Triticum durum</i>	3	65	0.99	0.95	0.11	0.00
<i>Triticum monococcum</i> <sup>13</sup>	47	870	0.97	0.99	0.80	0.04
<i>Triticum turgidum</i> <sup>14</sup>	53	345	0.98	0.91	0.25	0.00
<i>Triticum urartu</i>	0	420	1.00	1.00	0.95	—

(Contd.)

TAXON	OCCURRENCES		MODEL			
	ARCH.	CUR.	ACC.	ROC-AUC	SENS. CUR.	SENS. ARCH.
<i>Verbena officinalis</i>	3	89655	0.99	0.99	1.00	0.00
<i>Vicia ervilia</i>	26	1924	0.95	0.98	0.84	0.27
<i>Vicia faba</i>	7	42644	0.97	0.99	0.99	0.29
<i>Vicia narbonensis</i> <sup>15</sup>	3	2826	0.95	0.99	0.90	0.33
<i>Vicia orientalis</i> <sup>16</sup>	16	41	1.00	0.97	0.14	0.00
<i>Vitis sylvestris</i>	3	63	1.00	1.00	0.50	0.00
<i>Zygophyllum fabago</i>	3	3554	0.97	0.99	0.94	0.00

**Table 1** Summary of species modelled.<sup>1</sup>Including *Aizoon hispanicum*.<sup>2</sup>Including *Scirpus maritimus*.<sup>3</sup>Including *Cicer arietinum*.<sup>4</sup>Including *Vaccaria pyramidata*.<sup>5</sup>Excluded from modelling due to sample size.<sup>6</sup>Including *Hordeum vulgare*.<sup>7</sup>Including *Pisum sativum* and *Pisum elatius*.<sup>8</sup>Including *Linum usitatissimum*.<sup>9</sup>Including *Polygonum venantianum*.<sup>10</sup>Including *Salvia cryptantha*.<sup>11</sup>Including *Secale montanum*.<sup>12</sup>Including *Triticum spelta* and *Triticum aestivocompactum*.<sup>13</sup>Including *Triticum boeticum*.<sup>14</sup>Including *Triticum aestivum*, *Triticum dicoccum*, and *Triticum dicoccoides*.<sup>15</sup>Including *Vicia narbonense*.<sup>16</sup>Including *Lens culinaris* and *Lens orientalis*.

source of the remains (cf. [Wallace, Jones, et al. 2018](#)); archaeobotanical assemblages are subject to a variety of preservational and recovery biases, so by no means were all the species on our list consumed or even deliberately collected by people. However, we assume that their presence at a site of human settlement at least implies that they were part of the wider ecosystem that supported habitation there.

The taxonomic identifications of archaeobotanical material given in our source databases were previously controlled to ensure consistency between sources and to remove taxa that cannot be reliably distinguished (for details see [Arranz-Otaegui and Roe 2023](#)). Taxonomic names were then matched to the canonical form specified in the GBIF Backbone Taxonomy ([GBIF Secretariat 2023](#)) so they could be related to modern occurrences. Archaeologically-attested domestic species meeting our inclusion criteria were substituted for their wild progenitors (where different) when gathering occurrence data, since the domestic forms are now widespread and presumably uninformative of the species's original niche.

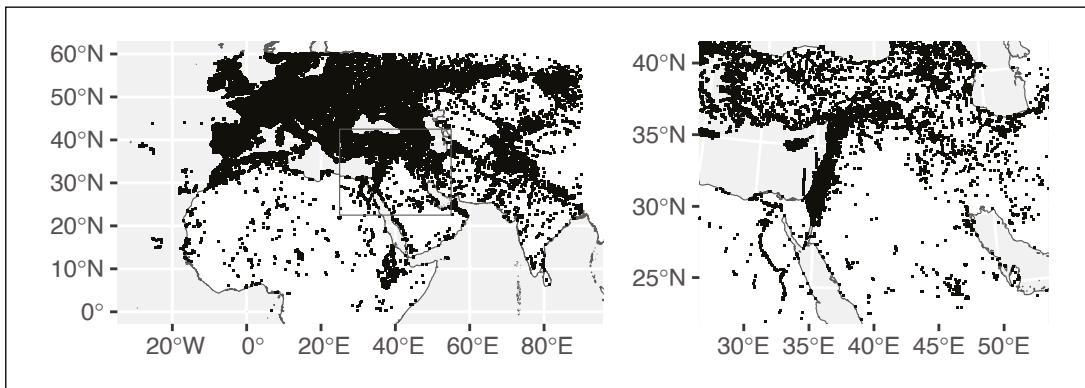
### 3.1 OCCURRENCE DATA

Georeferenced occurrence data for West Eurasia between 0 and 60° of latitude was obtained from the Global

Biodiversity Information Facility (GBIF) using via and the R package 'rgbif' ([Chamberlain and Boettiger 2017](#); [Chamberlain et al. 2024](#)). The GBIF dataset ([GBIF 2025a](#)) excluded fossil occurrences, recorded absences, and records with missing or dubious coordinates. Although niche models have reasonable predictive power even with small training samples ([Stockwell and Peterson 2002](#); [Hernandez et al. 2006](#); [Wisz et al. 2008](#)), we did not attempt to model three taxa with less than 40 usable occurrences, following recommendations for niche models generally and Random Forest-based models specifically ([Stockwell and Peterson 2002](#); [Luan et al. 2020](#)). Multiple records of the same taxon at the same coordinate were discarded because they do not impart information to the model. The resulting cleaned dataset used to train our niche models comprises 3,392,920 occurrences from 4769 constituent datasets.

GBIF is currently the best available general-purpose occurrence dataset for the West Asia region, its coverage is uneven both geographically and from species to species. The Southern Levant, and Israel specifically, is significantly more densely sampled than other parts of West Asia ([Figure 2](#)).

Random Forest is a presence-absence approach to niche modelling and therefore requires not just data



**Figure 2** Georeferenced occurrence records from West Eurasia used to train models ( $N = 1412083$ ). Inset, right: prediction region (West Asia).

on where a species is present, but where it is definitely not present. However ‘absence’ data is rarely available because it requires exhaustive survey. In practice, most applications of niche modelling are ‘presence-only’ and, where absence data is required (as for Random Forest), it is supplied as a random background sample of points.<sup>1</sup> The purpose of this sample is to inform the model about the nature of the underlying environment. The stochastic generation process means that some of these points will overlap or fall close to presences, so ensuring the model is not overly influenced by background samples is critical to its predictive importance (Valavi et al. 2022). Here we follow the advice of Barbet-Massin et al. (2012) for regression-based species distribution models and use a large ( $\approx 10000$ ) uniform sample of points from across the land area of the study region. These points are then weighted equally against the presences in the regression to produce a ‘balanced Random Forest’ (Valavi et al. 2022).

### 3.2 PREDICTOR DATA

We modelled the occurrence of species as a function of 24 geospatial predictor variables. These included:

- Sixteen ‘bioclimatic’ variables derived from monthly temperature and precipitation values, following standard practice for species distribution models (Hijmans et al. 2005). Contemporary bioclimatic predictor data for West Asia was extracted from the global CHELSA dataset (Karger et al. 2017), which predicts temperature and precipitation from downscaled general circulation model output at 1 km resolution.
- Terrain aspect and slope, which at high resolution perform well as proxies for solar radiation when modelling plant occurrence (Austin and Van Niel 2011; Leempoel et al. 2015); and the topographic wetness index (TWI), which serves as a proxy for soil moisture and is particularly important in modelling arid environments (Kopecký and Čížková 2010; Campos et al. 2016; Di Virgilio et al. 2018). All three

were derived from the SRTM30+ digital elevation model using algorithms from WhiteboxTools (Lindsay 2016).

- Edaphic data from SoilGrids (Hengl et al. 2014, 2017), which improves model performance for plants (Dubuis et al. 2013; Mod et al. 2016; Velazco et al. 2017). Based on a recent assessment of the reliability of SoilGrids data for species distribution modelling (Miller, Blackwood, and Case 2024), we used a subset of four variables relating to soil texture (clay, silt, sand) and pH at the surface (0–5 cm depth).

For hindcasting, we used reconstructed bioclimatic data for three key climatologies generated from downscaled paleoclimate simulations from the HadCM3 general circulation model (Fordham et al. 2017; Brown et al. 2018): the Bølling–Allerød (c. 14.7–12.9 ka), the Younger Dryas (c. 12.9–11.7 ka), and the Early Holocene (11.7–8.3 ka). Terrain and soil predictors were held constant, since reconstructions of these variables in the past are not available at sufficient scale. It is unlikely that either macroscale topography or soil characteristics have altered significantly over the period of time considered here, so we assume that this does not degrade model performance, and may in fact benefit it by providing ‘anchoring’ predictors that are independent of climate change.

For training, test predictions, and archaeological predictions predictor data was left in its native projection and resolution. For hindcast palaeodistributions, it was transformed to common equal-area projection and resolution of 5 km.

### 3.3 RANDOM FOREST

Ecological niche modelling is a classification problem that can be approached with a wide range of statistical methods. A substantial literature exists on the relatively performance of these approaches and their respective parameterisations (reviewed in Valavi et al. 2022). Random Forest, a widely-used machine learning algorithm, is amongst the best performing methods for

presence-only species distribution models, providing it is appropriately parameterised to account for the class imbalance between presence and background samples ([Valavi et al. 2021, 2022](#)). For our application, it also has the advantage of requiring little to no manual parameter tuning to achieve good predictive results, which makes it easier to model a larger numbers of taxa.

For each taxon we trained a classification model to predict occurrence (presence or absence/background) based on up to 24 predictor variables (Section 3.2). Highly correlated (Pearson's  $r > 0.7$ ) predictors were removed on a taxon-by-taxon basis, to mitigate issues of overfitting due to collinearity ([Dormann et al. 2013](#)), as were redundant predictors with zero variance. We used the Random Forest algorithm implemented in the R package 'ranger' ([Wright and Ziegler 2017](#)) and the 'tidymodels' ([Kuhn and Silge 2022](#)) framework for data preprocessing and model selection. To avoid overfitting, we follow Valavi et al. ([2021](#)) in their recommended hyperparameters and use of down-sampling to balance presence and background samples. Models for each taxon were fit independently, with redundant zero-variance predictors excluded, and assessed based on balanced training (3%) and test (1%) partitions.

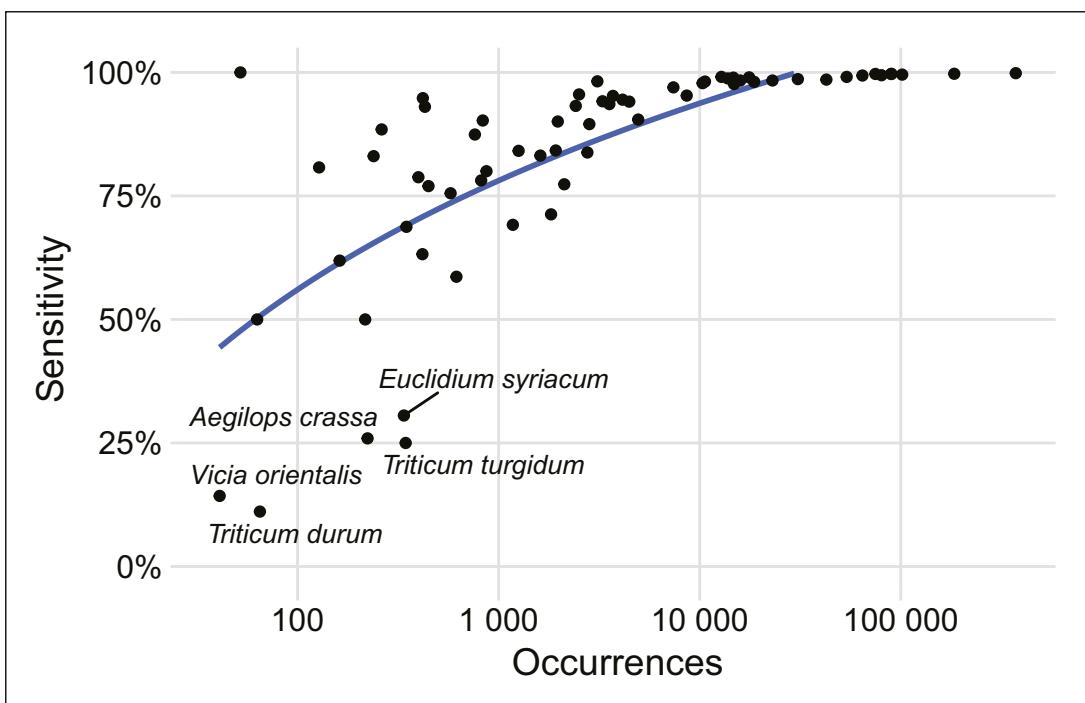
#### 4. MODEL ASSESSMENT

We trained Random Forest models for 65 taxa using contemporary occurrence data from GBIF, a random sample of background points, and the predictor variables described in Section 3.2. Substituting the "current"

climate predictors for those derived from palaeoclimatic simulations ([Brown et al. 2018](#)), we could then generate hindcast predictions for reconstructed past environments in 4 key climate periods – a total of 260 modelled palaeodistributions. Predicted distributions for individual taxa are presented in the appendix and accompanying material.

We assessed the predictive performance of the fitted niche models in the contemporary environment based on the reserved test partition ([Table 1](#)). Model accuracy (proportion of correctly classified presence and background samples) ranged between 93% and 100%, with an average of 98%. Sensitivity (proportion of correctly classified presence samples) ranged between 11% and 100%, with an average of 83%. The area under the models' receiver operating characteristic curves (ROC-AUC) was on average  $0.986 \pm 0.017$ . Model sensitivity is loosely correlated with the number of occurrences available for training ([Figure 3](#)), with the worst-performing models all having less than 1000 recorded occurrences: *Triticum durum*, *Euclidium syriacum*, *Triticum turgidum*, *Aegilops crassa*, and *Vicia orientalis*. Test metrics and ROC curves for the individual models are included in the appendix.

The ability of the hindcast models to predict the occurrence of specific species at archaeological sites is worse, with just 8% of presences in archaeobotanical assemblages successfully predicted at a threshold of 0.5. Model sensitivity (proportion of correctly predicted presences) in relation to the archaeological data is on average  $0.06 \pm 0.12$  ([Table 1](#)). A full assessment of the hindcasting performance of the individual models can be found in the appendix.



**Figure 3** Model sensitivity by number of training occurrences.

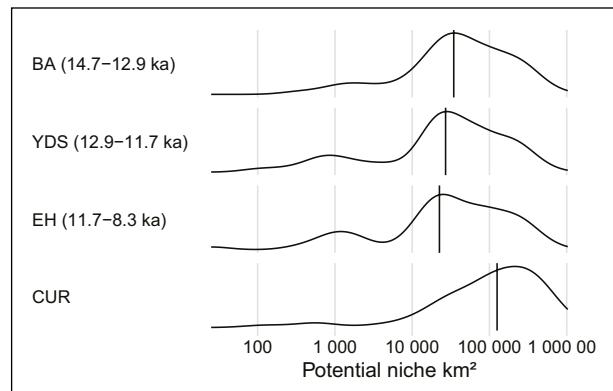
In summary, the majority of our models perform poorly as simple binary classifiers with a threshold of  $P(\text{present}) > 0.5$ . However, this is an arbitrary threshold; lowering it can improve sensitivity (proportion of occurrences correctly predicted) at the cost of specificity (proportion of non-occurrences incorrectly predicted). There are several approaches to selecting informative thresholds for binary classification in the ENM literature (Liu et al. 2005; Liu, White, and Newell 2013; Liu, Newell, and White 2016), with most aiming for an optimum compromise between sensitivity and specificity. We were unable to find a thresholding technique that significantly improved the hindcast models' performance as binary classifiers. By definition, binarizing model output also discards information from the underlying probabilistic prediction, so we have opted to present results probabilistically as far as possible. In principle the archaeological occurrences themselves could be used to 'calibrate' the hindcast models and select an optimum threshold, but there aren't enough archaeological attestations of most of the modelled species for this to be practicable at this time.

## 5. DISCUSSION

### 5.1 REDUCTION IN POTENTIAL NICHE SIZES OVER THE PLEISTOCENE/HOLOCENE BOUNDARY

Our reconstructed palaeodistributions (shown in full in the appendix) indicate that the majority of species had significantly different potential geographic niches in the Late Pleistocene/Early Holocene compared today. 55 of 65 species are predicted reduced niches in the past; 52 of more than 10% or more. Though the magnitude of the change in potential niche size from prehistory to the present likely reflects a degree of overfitting in the model (discussed further in Section 5.3), fluctuations in modelled niche size between the Bølling-Allerød (14.7–12.9 ka), Younger Dryas (12.9–11.7 ka), and Early Holocene (11.7–8.3 ka) are more directly comparable (Figure 4). The average potential niche size of modelled species was 35% smaller in the Early Holocene compared to the Bølling-Allerød, and 22% smaller during the Younger Dryas. This perhaps indicates that although this period is considered one of climatic amelioration globally (Jones et al. 2019), the colder conditions of the Pleistocene may have supported more extensive plant-based economies in West Asia specifically.

Many taxa that occur (or are predicted to occur) across the 'hilly flanks' today—including most crop progenitors—are reconstructed to have had a significantly more restricted distribution in the terminal Pleistocene/Early Holocene (Figure 5). These include *Ficus carica* (fig); *Hordeum* spp. (wild barleys); *Lathyrus aphaca* and *L. sativus* (both marginally edible legumes); *Triticum*



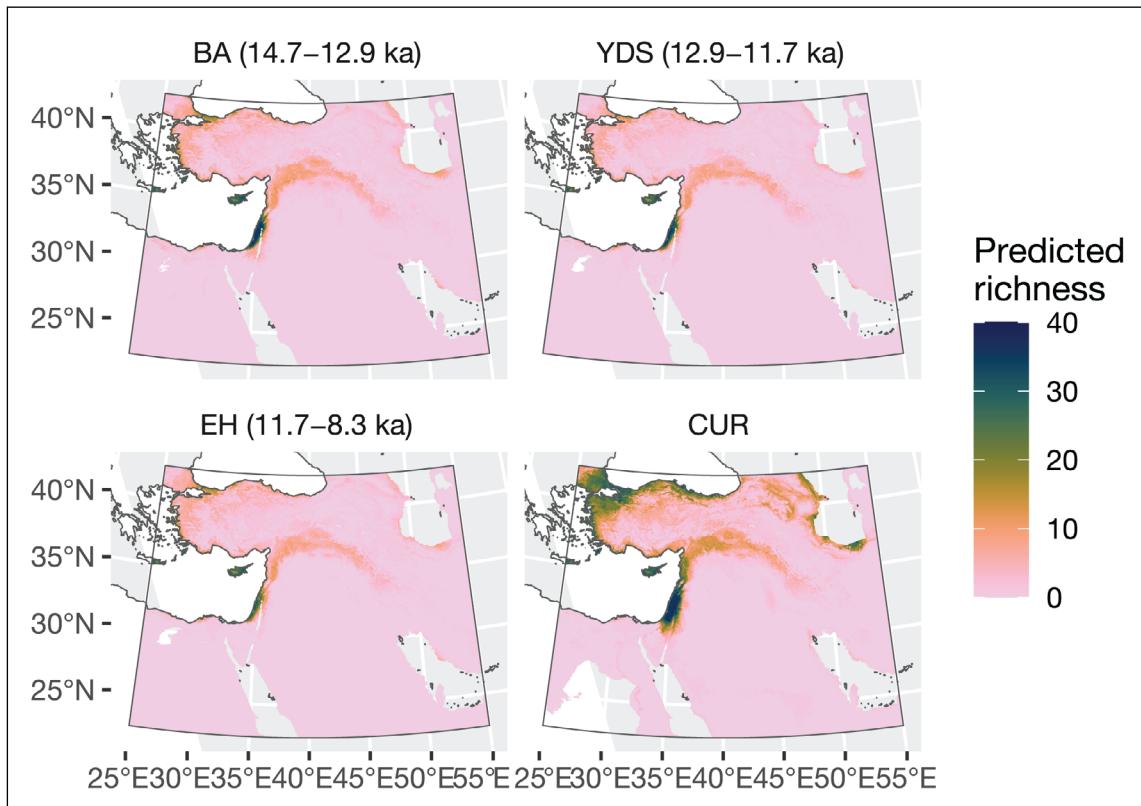
**Figure 4** Distribution of predicted species niche size by period (logarithmic scale). Dashed lines indicate the median niche size.

*aestivum compactum* (in the N. Levant), *T. monococcum aegilopoides*, *T. durum*, and *Triticum urartu* (but not the other wheat progenitor, *T. turgidum dicoccum* – see Section 5.2); *Aegilops speltoides*, but not *Aegilops tauschii* (goatgrasses); and *Vicia* spp. (vetches), including *Vicia faba* (broad beans). Most of Anatolia, Northern Mesopotamia, and the Zagros Mountains in particular disappear from the predicted niches of these species, leaving the Levant and to a lesser extent the Aegean and Cyprus as refugia.

Our results for the Levant are consistent with the current understanding of this region as developing early intensive foraging economies (the Natufian culture, Bar-Yosef 1998) and as a centre of origin of agriculture (Zeder 2011). Within the Levant, many species show moderate reductions in potential niche size over the Pleistocene/Holocene boundary, retreating from the Badia/Transjordan region (e.g. barley, Figure 8).

Loss of the Northern Mesopotamia-Anatolia region from the predicted potential niches of crop progenitors is interesting in light of the 'golden triangle' hypothesis (Lev-Yadun, Gopher, and Abbo 2000b; Kozłowski and Aurenche 2005; Abbo, Lev-Yadun, and Gopher 2010), which puts this region at the centre of the development of agriculture and plant domestication. Multiple lines of archaeological evidence have emerged that point away from this hypothesis and towards a more geographically diverse origin (Asouti 2006; Fuller, Willcox, and Allaby 2011; Arranz-Otaegui et al. 2016), and our reconstructions are also consistent with the late arrival of intensive plant-based foraging economies in this region (cf. the Natufian of the Levant).

The near-absence of the Zagros in any predicted niches is also surprising, given mounting evidence that animal domestication took place just as early in the eastern Mashriq as it did in the west (Zeder 2024). We consider that the most likely explanation for this is that our flora does not include the species that were most important to plant subsistence in the east. Archaeobotanical data on Neolithic sites in the Zagros is limited (compared to the Levant in particular) due to a hiatus in field research



**Figure 5** Predicted species richness (sum of predicted ranges) by period.

there from the 1980s to early 2010s ([Zeder 2024](#)). Recent research ([Riehl, Zeidi, and Conard 2013; Weide et al. 2017, 2018; Whitlam et al. 2018; González Carretero et al. 2023](#)) indicates that plant subsistence in this region was based on a distinct set of species than that of the Levant and Anatolia.

Cyprus and the Aegean are not conventionally considered part of the primary zone of domestication but rather amongst the first regions that acquired agriculture from West Asia. Our analysis complicates this picture, as it indicates that the wild distribution of many crop progenitors included these regions. Early examples of several domesticates are recorded at sites in Cyprus, Western Anatolia and Greece ([Arranz-Otaegui and Roe 2023](#)), and the Aegean region was probably connected to West Asia by a land bridge via Anatolia until the Early Holocene ([Aksu and Hiscott 2022](#)). Were these areas part of the same broader ‘interaction sphere’ that produced Neolithic agriculture in West Asia?

Exceptions to the dominant trend of niche size reduction include *Cicer reticulatum* (wild chickpea), which has a relatively stable potential niche centered on Northern Mesopotamia; and *Triticum turgidum dicoccum* (wild emmer wheat), which is predicted to occur in two limited zones centered around the Black Sea Coast of Anatolia and the Palmyra basin. In the latter case, neither of these areas are part of the predicted (by our model) modern distribution of wild emmer, which is centered around the Caucasus and Northern Mesopotamia. But it would be consistent with archaeological evidence for

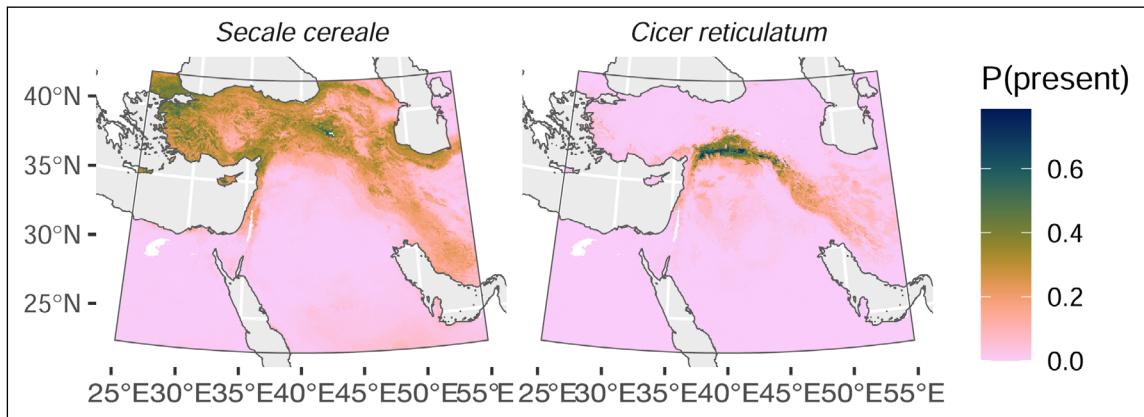
early cultivation at sites in the Upper Euphrates ([Willcox 2024](#)).

## 5.2 BIOGEOGRAPHY OF CROP PROGENITORS

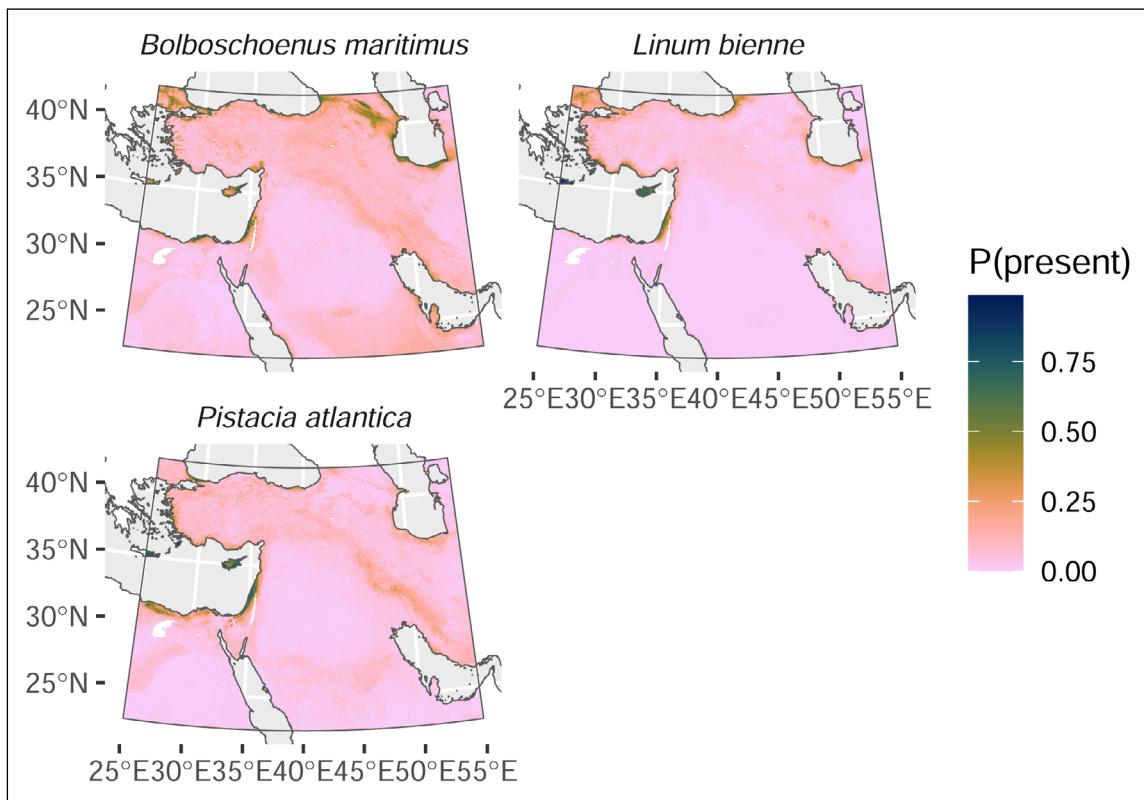
Almost all the cereal and legume crop progenitors we modelled are predicted to have only been found in the Levant during the terminal Pleistocene and Early Holocene (see appendix). Part of this may be due to the fact that both our initial flora and training occurrence dataset have a strong bias towards the southern Levant, but the modelled current potential niche of these plants do tend to include Anatolia and the Zagros, so this cannot be the only factor. It has previously been proposed that the Southern Levant was an “important glacial refuge area for wild cereals” ([Roberts et al. 2018](#)), which accords with these results.

One notable exception is the wild ancestor of chickpea (*Cicer reticulatum*), which is predicted to have a distribution centered on Northern Mesopotamia but encompassing much of the ‘hilly flanks’ (except the southern Levant, [Figure 6](#)). Another is rye (*Secale cereale*), which is inferred to be primarily Anatolian ([Figure 6](#)). This is perhaps relevant to rye’s unusual domestication history, as a crop of West Asian origin that was intensively exploited ([Hillman et al. 2001; Douché and Willcox 2018](#)) but apparently not first cultivated until much later than the ‘founder crops’, in Europe ([Schreiber et al. 2021](#)).

Flax (*Linum bienne*) is predicted to have had a highly concentrated distribution in Cyprus and along the Mediterranean coast of the southern Levant ([Figure 7](#)).



**Figure 6** Predicted palaeodistribution of wild chickpea and rye in the Early Holocene (11.7–8.2 ka).

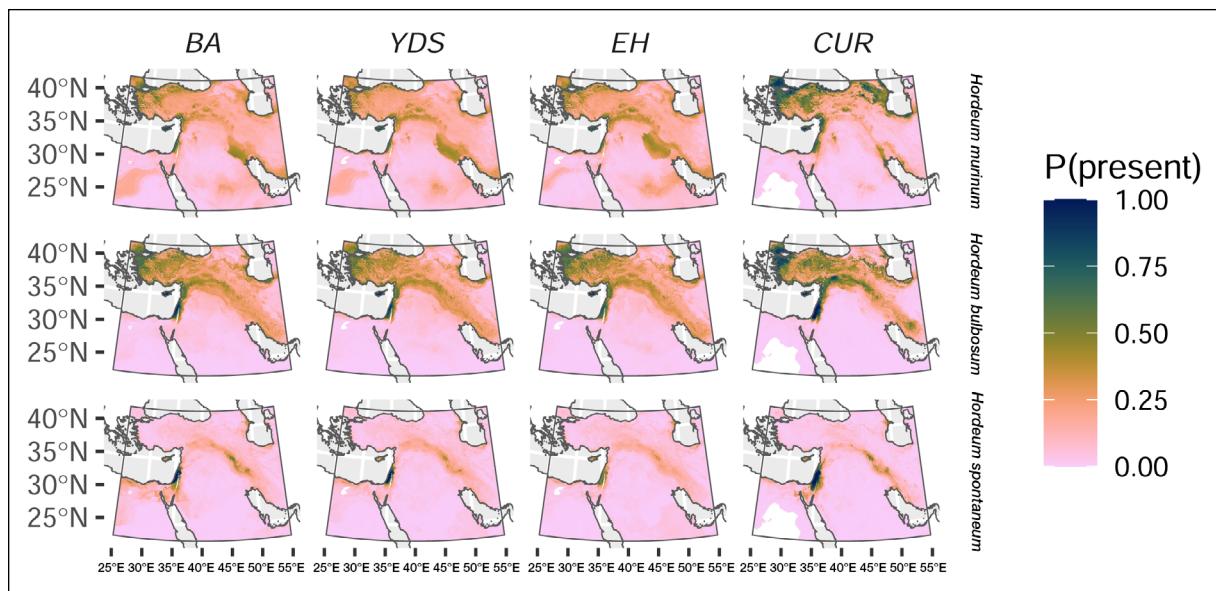


**Figure 7** Predicted palaeodistribution of flax, pistachio and clubrush in the Early Holocene (11.7–8.2 ka).

This is consistent with its low ubiquity in archaeobotanical assemblages (Arranz-Otaegui and Roe 2023), despite conventionally being considered a ‘founder crop’, and presumably implies that its domestication was similarly geographically constrained. It is the only unambiguous crop progenitor with such a restricted potential niche, though pistachio (*Pistacia atlantica*) and clubrush (*Bolboschoenus maritimus*) are similarly constrained to the Mediterranean coast (and North Africa, in the case of pistachio) (Figure 7). This is despite the fact that they are well-attested in the archaeobotanical record from across West Asia.

Wild barley (*Hordeum spontaneum*) and its relatives show a contraction of their predicted potential niche from the Pleistocene to the Holocene, concurrent with

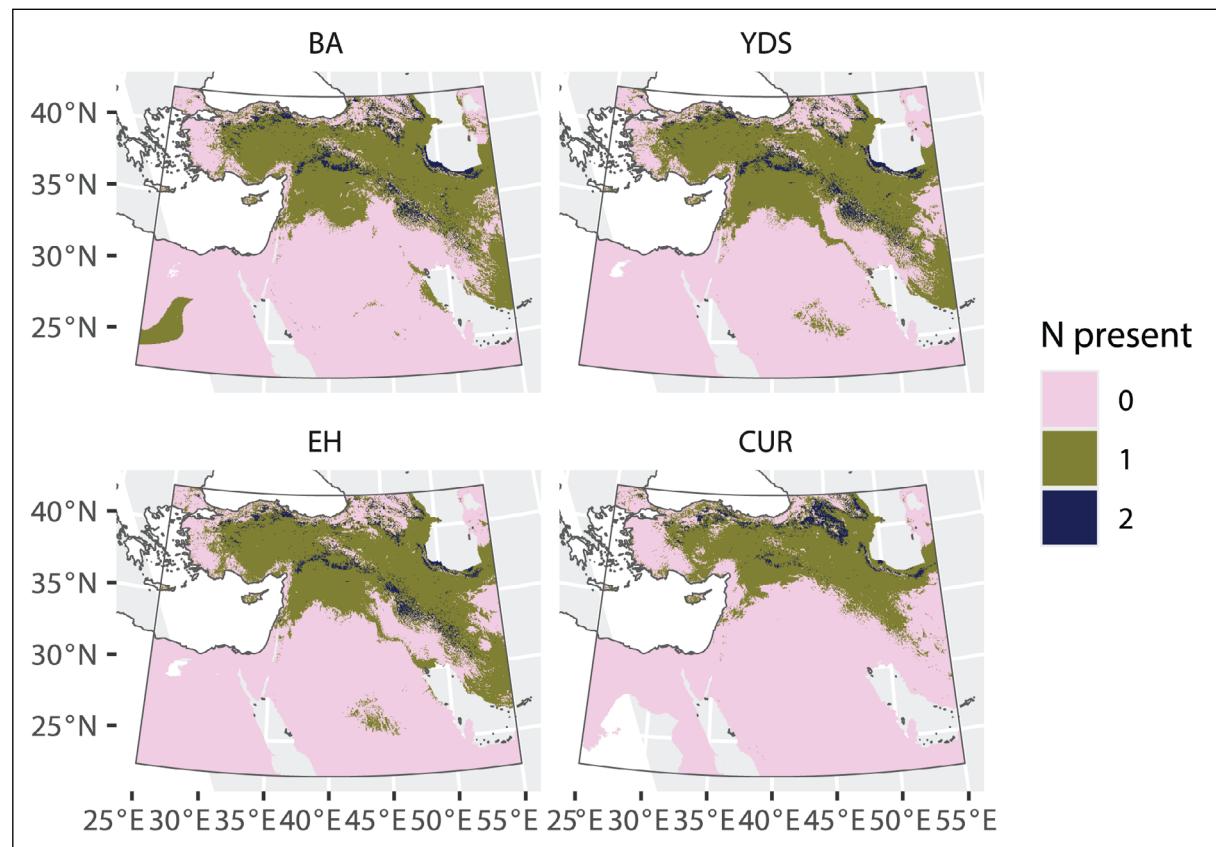
it being brought into cultivation (Figure 8). It also sees a marked decline in the archaeobotanical record from the Early PPNA/Early PPNB (where it was amongst the most common taxa) to the Late PPNB and Late Neolithic (Arranz-Otaegui and Roe 2023). Pistachio (*Pistacia atlantica*) shows similar trends, but it is less certain that this species was managed in the Neolithic. Conversely, the various wild wheat species native to West Asia show almost no response to Pleistocene/Holocene climate change, even within the Levant, and in the archaeobotanical record wheat displays the opposite trend to barley and pistachio – becoming gradually more abundant through the course of the Neolithic and dominant by its end (Arranz-Otaegui and Roe 2023).



**Figure 8** Predicted palaeodistribution of wild barleys.

Bread wheat (*Triticum aestivum*), the most common wheat cultivar today, has a complex ancestry that involves two recent hybridisation events (Levy and Feldman 2022): most recently between domestic emmer (*Triticum turgidum dicoccum*) and a goatgrass (*Aegilops tauschii*) c. 9 ka, and before that, in emmer, between wild red einkorn (*T. urartu*) and another goatgrass (*Aegilops speltoides*). The predicted potential niches of the two recent progenitors—*T. turgidum dicoccum* and *A. tauschii*—have relatively limited areas of overlap.

These include the Syrian Jazira (Figure 9), close to the only two sites in our archaeobotanical database where bread wheat (*T. aestivum*) co-occurs with both of the recent progenitors (*A. tauschii* and *T. turgidum dicoccum*): Tell Abu Hureyra and El Kowm II (Arranz-Otaegui and Roe 2023).<sup>2</sup> Taken together this suggests an origin of domesticated bread wheat, otherwise only loosely geographically constrained to the Levant–Upper Euphrates corridor (Levy and Feldman 2022), in this vicinity.



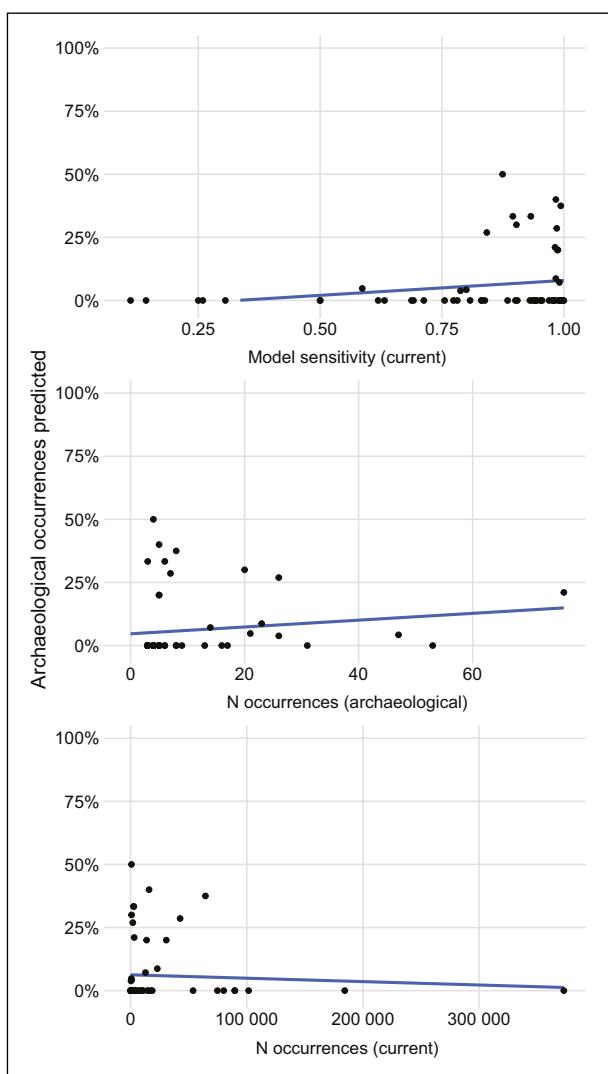
**Figure 9** Combined predicted palaeodistributions of bread wheat progenitors.

### 5.3 DISAGREEMENT BETWEEN HINDCAST MODELS AND ARCHAEOBOTANICAL COMPOSITION

In general there is middling to low agreement between our hindcast predictions of species niches and observed archaeological occurrences of the same species (Table 1).

The hindcasting performance of models is not strongly correlated with either the number of occurrences in the training dataset, the performance of the model on test data, or the number of attested archaeological occurrences (Figure 10). This indicates against a single, simple explanation for the discrepancy, i.e. that one or the other signal is ‘wrong’.

The core issue is that the two sets of evidence discussed here—modelled palaeodistributions and archaeologically-attested palaeooccurrences—are signals of related but distinct phenomena. The hindcast niche models are projections of the species’ present realised niche into a reconstructed past climatology. The archaeobotanical record is an incomplete sample of



**Figure 10** Summary of hindcasting performance of models in relation to model performance and available archaeological test data.

a past realised niche passed through various anthropic and taphonomic filters. Neither can be assumed to be fully accurate representations of the ultimate aim of this study, which is the *fundamental* niche of the species in relation to reconstructable climate and environmental factors – though both carry information about it.

For each specific species there are a variety of potential explanations for why the signals differ, but without further lines of evidence we are not in a position to distinguish them. In general, our models probably underestimate the fundamental niche or potential distribution of the target taxon for a variety of reasons:

- The inherent tendency of machine learning models to overfit to the training scenario, despite our efforts to mitigate this;
- Uneven spatial sampling intensity of the GBIF occurrence dataset (see Section 3.1), meaning that some parts of the species’ niche are probably either over- and under-represented in the training data;
- Use of fairly coarse (2000 year) time-averaged climate slices, which capture a general trend but are probably unrepresentative of the environment around sites at the specific time at which they were occupied.

At the same time, there are several reasons to believe that the archaeobotanical data (at least as it has been compiled for this study) overestimates the fundamental niche:

- Archaeological deposits are also time-averaged signals, in some cases over quite large spans of time;
- The chronology attached to our archaeobotanical data is imprecise, e.g. primarily based on point estimates of radiocarbon dates ([Michałczyński 2007](#)), producing a further averaging effect;
- Archaeological occurrences could be the result of human transplanting, range expansion and/or transport, especially considering this is the period where agriculture began to be practiced;
- Most archaeobotanical samples in our archaeological test dataset were small, so individual errors and misidentifications have a correspondingly large impact on quantification of model performance.

At the same time, we cannot rule out more substantive reasons for the discrepancy between predicted and observed archaeological occurrences. The niches of the modelled species could have changed since the Early Holocene, which would not be captured in a model trained purely on modern specimens. Human economic choices—mobility, foraging strategies, cultivation, etc.—could also produce archaeobotanical assemblages whose composition depart significantly from that of the surrounding local flora.

What are the implications of this assessment? In the immediate term, the modelled palaeodistributions presented here should be taken as conservative estimates – a minimal likely potential niche of the species. Closer inspection of discrepancies between the modelled and attested distributions of particular species in relation to specific environmental, taphonomic and cultural factors could yield further insights. Future applications of hindcast palaeoecological niche models in archaeology could refine the methodology in several ways: using more finely resolved palaeoclimate sequences (e.g. Karger et al. 2023); building more refined archaeological chronologies; performing hyperparameter tuning to further mitigate overfitting;<sup>3</sup> and compiling expanded archaeological datasets that could be used to calibrate binary classification.

## 6. CONCLUSION

We present the first continuous, spatially explicit models of the palaeodistributions of 65 plant species found regularly in association with Late Epipalaeolithic and Early Neolithic sites in West Asia. This deductive approach—modelling the niche of a species based on its occurrence in relation to environmental factors today, and using this together with palaeoclimate simulations to infer its past distribution—represents a new line of evidence on the archaeoecology of the world's first agricultural societies. It provides a complementary picture to that gleaned from environmental archaeology and climatological archives because it is independent of the taphonomic, anthropic, and recovery-related processes that affect these records.

The modelled palaeodistributions of the species presented here (see Appendix) represent plausible minimum potential niche under the average climatic conditions of the Bølling-Allerød, Younger Dryas, and Early Holocene periods. The models' generally high performance as assessed against independent contemporary test datasets lends confidence to these predictions. Their application as a binary predictor of archaeological presence or absence is significantly less promising, which likely reflects a combination of methodological limitations of our modelling approach, the incomplete and coarsely temporally-resolved nature of the archaeobotanical test dataset, and genuine discrepancies between the species' fundamental niche and its occurrence in the archaeological record. Whether on the broad scale (e.g. the restricted geographic range of most species compared to their attestation in the archaeological record), or relating to specific species (i.e. false positives and false negatives), these discrepancies suggest several avenues for future investigation.

Modelling a large number of species using machine learning, the substantial occurrence datasets available for the present day, and a hindcasting approach to past distributions also represents a significant advance in the methodology of palaeoecological niche modelling. This approach is enabled by the availability of high quality, global open datasets in ecology (GBIF 2025b; GBIF Secretariat 2023), earth science (Farr et al. 2007), and climatology (Karger et al. 2017; Brown et al. 2018). However, several areas of methodological improvement are evident. Notably, the state of open data availability in archaeology lags conspicuously behind the fields mentioned above. Though we have benefited from the relatively long tradition of compiling archaeobotanical data in our region of study (Colledge, Conolly, and Shennan 2004; Shennan and Conolly 2007; Riehl and Kümmel 2005; Lucas and Fuller 2018; Fuller et al. 2018; Wallace, Jones, et al. 2018; Wallace, Livarda, et al. 2018), further development of open, comprehensive and up-to-date 'backbone' datasets on site locations and chronologies is needed to advance archaeoecological modelling to the same level.

## DATA ACCESSIBILITY STATEMENT

The data and R code used to produce this study is archived with Zenodo at <https://doi.org/10.5281/zenodo.14629984>. Modelled palaeodistributions in raster format (.TIF) can be found in the same repository.

## NOTES

- 1 We use a background sample that is representative of the entire study area, as opposed to attempting to construct a 'pseudo-absence' sample. See Sillero et al. (2021) for the distinction.
- 2 We are grateful to Benjamin Nowak for pointing this out to us.
- 3 We did not perform hyperparameter tuning in this study due to computational limitations.

## ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendix:** The ecological niche models of 65 plant species relevant to the subsistence of Late Epipalaeolithic (15–11.7 ka) and Neolithic (11.7–8.2 ka) societies in West Asia. DOI: <https://doi.org/10.5334/oq.163.s1>

## ACKNOWLEDGEMENTS

We are grateful to Alex Weide and the second, anonymous reviewer of our manuscript, whose suggestions have greatly improved this study.

## FUNDING INFORMATION

This work was partially funded by the European Research Council (ERC-2021-STG project 101039060, *PalaeOrigins: Tracing the Epipalaeolithic origins of plant management in southwest Asia*).

## COMPETING INTERESTS

The authors have no competing interests to declare.

## AUTHOR AFFILIATIONS

**Joe Roe**  [orcid.org/0000-0002-1011-1244](https://orcid.org/0000-0002-1011-1244)

University of Copenhagen, DK

**Amaia Arranz-Otaegui**  [orcid.org/0000-0002-5091-6426](https://orcid.org/0000-0002-5091-6426)

University of the Basque Country, ES

## REFERENCES

- Abbo, S, Gopher, A, Peleg, Z, Saranga, Y, Fahima, T, Salamini, F and Lev-Yadun, S.** 2006. ‘The ripples of “The Big (agricultural) Bang”: The spread of early wheat cultivation’. *Genome*, 49(8): 861–863. DOI: <https://doi.org/10.1139/g06-049>
- Abbo, S, Gopher, A, Rubin, B and Lev-Yadun, S.** 2005. ‘On the Origin of Near Eastern Founder Crops and the ‘Dump-heap Hypothesis’’. *Genetic Resources and Crop Evolution* 52(5): 491–495. DOI: <https://doi.org/10.1007/s10722-004-7069-x>
- Abbo, S, Lev-Yadun, S and Gopher, A.** 2010. ‘Agricultural Origins: Centers and Noncenters; A Near Eastern Reappraisal’. *Critical Reviews in Plant Sciences*, 29(5): 317–328. DOI: <https://doi.org/10.1080/07352689.2010.502823>
- Abbo, S, Lev-Yadun, S and Gopher, A.** 2011. ‘Origin of Near Eastern plant domestication: Homage to Claude Levi-Strauss and ‘La Pensée Sauvage’’. *Genetic Resources and Crop Evolution*, 58(2): 175–179. DOI: <https://doi.org/10.1007/s10722-010-9630-0>
- Abbo, S, Lev-Yadun, S and Gopher, A.** 2012. ‘Plant Domestication and Crop Evolution in the Near East: On Events and Processes’. *Critical Reviews in Plant Sciences*, 31(3): 241–257. DOI: <https://doi.org/10.1080/07352689.2011.645428>
- Aksu, AE and Hiscott, RN.** 2022. ‘Persistent Holocene outflow from the Black Sea to the eastern Mediterranean Sea still contradicts the Noah’s Flood Hypothesis: A review of 1997–2021 evidence and a regional paleoceanographic synthesis for the latest Pleistocene–Holocene’. *Earth-Science Reviews*, 227: 103960. DOI: <https://doi.org/10.1016/j.earscirev.2022.103960>
- Arranz-Otaegui, A, Colledge, S, Zapata, L, Teira-Mayolini, LC and Ibáñez, JJ.** 2016. ‘Regional diversity on the timing for the initial appearance of cereal cultivation and domestication in southwest Asia’. *Proceedings of the National Academy of Sciences*, 113(49): 14001–14006. DOI: <https://doi.org/10.1073/pnas.1612797113>
- Arranz-Otaegui, A, González Carretero, L, Roe, J and Richter, T.** 2018. ‘‘Founder crops’ v. Wild plants: Assessing the plant-based diet of the last hunter-gatherers in southwest Asia’. *Quat. Sci. Rev.*, 186: 263–283. DOI: <https://doi.org/10.1016/j.quascirev.2018.02.011>
- Arranz-Otaegui, A and Roe, J.** 2023. ‘Revisiting the concept of the ‘Neolithic Founder Crops’ in southwest Asia’. *Vegetation History and Archaeobotany*. DOI: <https://doi.org/10.1007/s00334-023-00917-1>
- Asouti, E.** 2006. ‘Beyond the Pre-Pottery Neolithic B interaction sphere’. *Journal of World Prehistory*, 20(2): 87–126. DOI: <https://doi.org/10.1007/s10963-007-9008-1>
- Austin, MP and Van Niel, KP.** 2011. ‘Improving species distribution models for climate change studies: Variable selection and scale’. *Journal of Biogeography*, 38(1): 1–8. DOI: <https://doi.org/10.1111/j.1365-2699.2010.02416.x>
- Badr, A, Müller, K, Sch, R, Rabey, HE, Effgen, S, Ibrahim, HH, Pozzi, C, Rohde, W and Salamini, F.** 2000. ‘On the Origin and Domestication History of Barley (*Hordeum vulgare*)’. *Molecular Biology and Evolution*, 17(4): 499–510. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a026330>
- Banks, WE, Moncel, M-H, Raynal, J-P, Cobos, ME, Romero-Alvarez, D, Woillez, M-N, Faivre, J-P, Gravina, B, d'Errico, F, Locht, J-L and Santos, F.** 2021. ‘An ecological niche shift for Neanderthal populations in Western Europe 70,000 years ago’. *Scientific Reports*, 11(1): 5346. DOI: <https://doi.org/10.1038/s41598-021-84805-6>
- Barbet-Massin, M, Jiguet, F, Albert, CH and Thuiller, W.** 2012. ‘Selecting pseudo-absences for species distribution models: How, where and how many?: How to use pseudo-absences in niche modelling?’ *Methods Ecol. Evol.*, 3(2): 327–338. DOI: <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Bar-Yosef, O.** 1998. ‘The Natufian culture in the Levant, threshold to the origins of agriculture’. *Evolutionary Anthropology*, 6(5): 159–177. DOI: [https://doi.org/10.1002/\(SICI\)1520-6505\(1998\)6:5<159::AID-EVAN4>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<159::AID-EVAN4>3.0.CO;2-7)
- Benito, BM, Svenning, J-C, Kellberg-Nielsen, T, Riede, F, Gil-Romera, G, Mailund, T, Kjaergaard, PC and Sandel, BS.** 2017. ‘The ecological niche and distribution of Neanderthals during the Last Interglacial’. *J. Biogeogr.*, 44(1): 51–61. DOI: <https://doi.org/10.1111/jbi.12845>
- Boyd, B.** 2018. ‘Settled? Recent debates in the archaeology of the Epipalaeolithic and Pre-Pottery Neolithic of Southwest Asia’. *Asian Archaeology*, 1(1): 63–73. DOI: <https://doi.org/10.1007/s41826-018-0006-3>
- Brown, JL, Hill, DJ, Dolan, AM, Carnaval, AC and Haywood, AM.** 2018. ‘PaleoClim, high spatial resolution paleoclimate surfaces for global land areas’. *Sci Data*, 5: 180254. DOI: <https://doi.org/10.1038/sdata.2018.254>

- Brown, SC, Wigley, TML, Otto-Bliesner, BL and Fordham, DA.** 2020. ‘StableClim, continuous projections of climate stability from 21000 BP to 2100 CE at multiple spatial scales’. *Scientific Data*, 7(1): 335. DOI: <https://doi.org/10.1038/s41597-020-00663-3>
- Butzer, KW.** 1971. ‘Agricultural origins in the Near East as a geographical problem’. In: Streuver, S (ed.) *Prehistoric Agriculture*. American Museum Sourcebooks in Anthropology.
- Campos, VE, Cappa, FM, Viviana, FM and Giannoni, SM.** 2016. ‘Using remotely sensed data to model suitable habitats for tree species in a desert environment’. *Journal of Vegetation Science*, 27(1): 200–210. DOI: <https://doi.org/10.1111/jvs.12328>
- Chamberlain, S, Barve, V, Mcglinn, D, Oldoni, D, Desmet, P, Geffert, L and Ram, K.** 2024. *Rgbif: Interface to the global biodiversity information facility API*.
- Chamberlain, S and Boettiger, C.** 2017. ‘R python, and ruby clients for GBIF species occurrence data’. *PeerJ PrePrints*. DOI: <https://doi.org/10.7287/peerj.preprints.3304v1>
- Childe, VG.** 1936. *Man Makes Himself*. London: Watts & Co.
- Colledge, S.** 2001. *Plant exploitation on Epipalaeolithic and early Neolithic sites in the Levant*. BAR International series. Oxford: Oxbow. DOI: <https://doi.org/10.30861/9781841711904>
- Colledge, S, Conolly, J and Shennan, S.** 2004. ‘Archaeobotanical Evidence for the Spread of Farming in the Eastern Mediterranean’. *Current Anthropology*, 45(S4): S35–S58. DOI: <https://doi.org/10.1086/422086>
- Collins, C, Asouti, E, Grove, M, Kabukcu, C, Bradley, L and Chiverrell, R.** 2018. ‘Understanding resource choice at the transition from foraging to farming: An application of palaeodistribution modelling to the Neolithic of the Konya Plain, south-central Anatolia, Turkey’. *Archaeol. Sci.*, 96: 57–72. DOI: <https://doi.org/10.1016/j.jas.2018.02.003>
- Conolly, J, Manning, K, Colledge, S, Dobney, K and Shennan, S.** 2012. ‘Species distribution modelling of ancient cattle from early Neolithic sites in SW Asia and Europe’. *Holocene*, 22(9): 997–1010. DOI: <https://doi.org/10.1177/0959683612437871>
- Cordova, CE.** 2007. *Millennial Landscape Change in Jordan: Geoarchaeology and Cultural Ecology*. Tucson, AZ: University of Arizona Press. DOI: <https://doi.org/10.1353/book.114867>
- Darwin, C.** 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray. DOI: <https://doi.org/10.5962/bhl.title.68064>
- David Polly, P and Eronen, JT.** 2011. ‘Mammal Associations in the Pleistocene of Britain: Implications of Ecological Niche Modelling and a Method for Reconstructing Palaeoclimate’. In: Ashton, N, Lewis, SG, and Stringer, C (eds.) *Developments in Quaternary Sciences*. Elsevier. pp. 279–304. DOI: <https://doi.org/10.1016/B978-0-444-53597-9.00015-7>
- de Andrés-Herrero, M, Becker, D and Weniger, G-C.** 2018. ‘Reconstruction of LGM faunal patterns using Species Distribution Modelling. The archaeological record of the Solutrean in Iberia’. *Quat. Int.*, 485: 199–208. DOI: <https://doi.org/10.1016/j.quaint.2017.10.042>
- de Candolle, A.** 1886. *The Origin of Cultivated Plants*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139107365>
- Dennell, RW.** 1976. ‘The economic importance of plant resources represented on archaeological sites’. *Journal of Archaeological Science*, 3(3): 229–247. DOI: [https://doi.org/10.1016/0305-4403\(76\)90057-1](https://doi.org/10.1016/0305-4403(76)90057-1)
- Di Virgilio, G, Wardell-Johnson, GW, Robinson, TP, Temple-Smith, D and Hesford, J.** 2018. ‘Characterising fine-scale variation in plant species richness and endemism across topographically complex, semi-arid landscapes’. *Journal of Arid Environments*, 156: 59–68. DOI: <https://doi.org/10.1016/j.jaridenv.2018.04.005>
- Diamond, J.** 2002. ‘Evolution, consequences and future of plant and animal domestication’. *Nature*, 418(6898): 700–707. DOI: <https://doi.org/10.1038/nature01019>
- Dormann, CF, Elith, J, Bacher, S, Buchmann, C, Carl, G, Carré, G, Marquéz, JRG, Gruber, B, Lafourcade, B, Leitão, PJ, Münkemüller, T, McClean, C, Osborne, PE, Reineking, B, Schröder, B, Skidmore, AK, Zurell, D and Lautenbach, S.** 2013. ‘Collinearity: A review of methods to deal with it and a simulation study evaluating their performance’. *Ecography*, 36(1): 27–46. DOI: <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Douché, C and Willcox, G.** 2018. ‘New archaeobotanical data from the Early Neolithic sites of Dja’de el-Mughara and Tell Aswad (Syria): A comparison between the Northern and the Southern Levant’. *Paléorient*, 44(2): 45–58.
- Dubuis, A, Giovanettina, S, Pellissier, L, Pottier, J, Vittoz, P and Guisan, A.** 2013. ‘Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables’. *Journal of Vegetation Science*, 24(4): 593–606. DOI: <https://doi.org/10.1111/jvs.12002>
- Farr, TG, Rosen, PA, Caro, E, Crippen, R, Duren, R, Hensley, S, Kobrick, M, Paller, M, Rodriguez, E, Roth, L, Seal, D, Shaffer, S, Shimada, J, Umland, J, Werner, M, Oskin, M, Burbank, D and Alsdorf, D.** 2007. ‘The Shuttle Radar Topography Mission’. *Rev. Geophys.*, 45(2): RG2004. DOI: <https://doi.org/10.1029/2005RG000183>
- Fordham, DA, Saltré, F, Haythorne, S, Wigley, TML, Otto-Bliesner, BL, Chan, KC and Brook, BW.** 2017. ‘PaleoView: A tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales’. *Ecography*, 40(11): 1348–1358. DOI: <https://doi.org/10.1111/ecog.03031>
- Franklin, J and Miller, JA.** 2009. *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press.

- Franklin, J, Potts, AJ, Fisher, EC, Cowling, RM and Marean, CW.** 2015. 'Paleodistribution modeling in archaeology and paleoanthropology'. *Quat. Sci. Rev.*, 110: 1–14. DOI: <https://doi.org/10.1016/j.quascirev.2014.12.015>
- Fuller, DQ.** 2010. 'An emerging paradigm shift in the origins of agriculture'. *General Anthropology*, 17(2): 1–12. DOI: <https://doi.org/10.1111/j.1939-3466.2010.00010.x>
- Fuller, DQ and Colledge, S.** 2008. 'Recent lessons from Near Eastern archaeobotany: Wild cereal use, pre-domestication cultivation and tracing multiple origins and dispersals'. *Pragdvara*, 18: 105–134.
- Fuller, DQ, Lucas, L, Carretero, LG and Stevens, C.** 2018. 'From intermediate economies to agriculture: Trends in wild food use, domestication and cultivation among early villages in Southwest Asia'. *Paléorient*, 44(2): 59–74.
- Fuller, DQ, Willcox, G and Allaby, RG.** 2011. 'Cultivation and domestication had multiple origins: Arguments against the core area hypothesis for the origins of agriculture in the Near East'. *World Archaeol*, 43(4): 628–652. DOI: <https://doi.org/10.1080/00438243.2011.624747>
- GBIF.** 2025a. GBIF Occurrence Download (14 August 2025). DOI: <https://doi.org/10.15468/dl.t8sqqm>
- GBIF.** 2025b. What is GBIF? <https://www.gbif.org/what-is-gbif> [Last accessed 23 January 2026].
- GBIF Secretariat.** 2023. GBIF Backbone Taxonomy. DOI: <https://doi.org/10.15468/39OMEI>
- González Carretero, L, Lucas, L, Stevens, C and Fuller, DQ.** 2023. 'Investigating early agriculture, plant use and culinary practices at Neolithic Jarmo (Iraqi Kurdistan)'. *Journal of Archaeological Science: Reports*, 52: 104264. DOI: <https://doi.org/10.1016/j.jasrep.2023.104264>
- Gopher, A, Abbo, S and Yadun, SL.** 2001. 'The 'when', the 'where' and the 'why' of the Neolithic revolution in the Levant'. *Documenta Praehistorica*, 28: 49–62. DOI: <https://doi.org/10.4312/dp.28.3>
- Guran, SH, Yousefi, M, Kafash, A and Ghasidian, E.** 2024. 'Reconstructing contact and a potential interbreeding geographical zone between Neanderthals and anatomically modern humans'. *Scientific Reports*, 14(1): 20475. DOI: <https://doi.org/10.1038/s41598-024-70206-y>
- Harlan, JR.** 1971. 'Agricultural Origins: Centers and Noncenters'. *Science*, 174(4008): 468–474. DOI: <https://doi.org/10.1126/science.174.4008.468>
- Harlan, JR.** 1977. 'The Origins of Cereal Agriculture in the Old World'. In: Reed, CA (ed.) *Origins of Agriculture*. De Gruyter Mouton. pp. 357–384. DOI: <https://doi.org/10.1515/9783110813487.357>
- Harlan, JR and Zohary, D.** 1966. 'Distribution of Wild Wheats and Barley'. *Science*, 153(3740): 1074–1080. DOI: <https://doi.org/10.1126/science.153.3740.1074>
- Harris, DR.** 1989. 'An evolutionary continuum of people-plant interaction'. In: *The Emergence of Agriculture*. Routledge.
- Harris, DR.** 1990. 'Vavilov's concept of centres of origin of cultivated plants: Its genesis and its influence on the study of agricultural origins'. *Biological Journal of the Linnean Society*, 39(1): 7–16. DOI: <https://doi.org/10.1111/j.1095-8312.1990.tb01608.x>
- Harris, DR.** 2007. 'Agriculture, Cultivation and Domestication: Exploring the Conceptual Framework of Early Food Production'. In: *Rethinking Agriculture*. Routledge.
- Harris, DR and Hillman, GC.** 1989. 'Introduction'. In: *Foraging and Farming: The Evolution of Plant Exploitation*. Routledge.
- Hastorf, CA and Popper, VS.** 1988. *Current Paleoethnobotany: Analytical Methods and Cultural Interpretations of Archaeological Plant Remains*. University of Chicago Press.
- Helbæk, H.** 1959. 'How farming began in the Old World'. *Archaeology*, 12(3): 183–189.
- Helbæk, H.** 1969. *Plant collecting, dry-farming and irrigation agriculture in prehistoric Deh Luran*. University of Michigan Ann Arbor.
- Hengl, T, de Jesus, JM, MacMillan, RA, Batjes, NH, Heuvelink, GBM, Ribeiro, E, Samuel-Rosa, A, Kempen, B, Leenaars, JGB, Walsh, MG and Gonzalez, MR.** 2014. 'SoilGrids1km—global soil information based on automated mapping'. *PLoS One*, 9(8): e105992. DOI: <https://doi.org/10.1371/journal.pone.0105992>
- Hengl, T, Mendes de Jesus, J, Heuvelink, GBM, Ruiperez Gonzalez, M, Kilibarda, M, Blagotić, A, Shangguan, W, Wright, MN, Geng, X, Bauer-Marschallinger, B, Guevara, MA, Vargas, R, MacMillan, RA, Batjes, NH, Leenaars, JGB, Ribeiro, E, Wheeler, I, Mantel, S and Kempen, B.** 2017. 'SoilGrids250m: Global gridded soil information based on machine learning'. *PLoS One*, 12(2): e0169748. DOI: <https://doi.org/10.1371/journal.pone.0169748>
- Hernandez, PA, Graham, CH, Master, LL and Albert, DL.** 2006. 'The effect of sample size and species characteristics on performance of different species distribution modeling methods'. *Ecography*, 29(5): 773–785. DOI: <https://doi.org/10.1111/j.0906-7590.2006.04700.x>
- Heun, M, Schäfer-Pregl, R, Klawan, D, Castagna, R, Accerbi, M, Borghi, B and Salamini, F.** 1997. 'Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting'. *Science*, 278(5341): 1312–1314. DOI: <https://doi.org/10.1126/science.278.5341.1312>
- Hijmans, RJ, Cameron, SE, Parra, JL, Jones, PG and Jarvis, A.** 2005. 'Very high resolution interpolated climate surfaces for global land areas'. *Int. J. Climatol.*, 25(15): 1965–1978. DOI: <https://doi.org/10.1002/joc.1276>
- Hillman, GC and Davies, MS.** 1990. 'Measured domestication rates in wild wheats and barley under primitive cultivation, and their archaeological implications'. *Journal of World Prehistory*, 4(2): 157–222. DOI: <https://doi.org/10.1007/BF00974763>
- Hillman, GC and Davies, MS.** 1992. 'Domestication Rate in Wild Wheats and Barley under Primitive Cultivation: Preliminary Results and Archaeological Implications of Field Measurements of Selection Coefficient'. *Monographie du CRA*, (6): 113–158.

- Hillman, G, Hedges, R, Moore, A, Colledge, S and Pettitt, P.** 2001. 'New evidence of Lateglacial cereal cultivation at Abu Hureyra on the Euphrates'. *The Holocene*, 11(4): 383–393. DOI: <https://doi.org/10.1191/095968301678302823>
- Hopf, M.** 1969. 'Plant remains and early farming in Jericho 1'. In: *The Domestication and Exploitation of Plants and Animals*. Chicago: Aldine.
- Hopf, M.** 1986. 'Archaeological Evidence of the Spread and Use of Some Members of the Leguminosae Family'. In: Barigozzi, C (ed.) *Developments in Agricultural and Managed Forest Ecology*. The origin and domestication of cultivated plants. Elsevier. pp. 35–60. DOI: <https://doi.org/10.1016/B978-0-444-42703-8.50008-7>
- Iob, A and Botigué, L.** 2023. 'Genomic analysis of emmer wheat shows a complex history with two distinct domestic groups and evidence of differential hybridization with wild emmer from the western Fertile Crescent'. *Vegetation History and Archaeobotany*, 32(5): 545–558. DOI: <https://doi.org/10.1007/s00334-022-00898-7>
- Jones, MD, Abu-Jaber, N, AlShdaifat, A, Baird, D, Cook, BI, Cuthbert, MO, Dean, JR, Djamali, M, Eastwood, W, Fleitmann, D, Haywood, A, Kwiecien, O, Larsen, J, Maher, LA, Metcalfe, SE, Parker, A, Petrie, CA, Primmer, N, Richter, T, Roberts, N, Roe, J, Tindall, JC, Ünal-İmer, E and Weeks, L.** 2019. '20,000 years of societal vulnerability and adaptation to climate change in southwest Asia'. *WIREs Water*, 6(2): e1330. DOI: <https://doi.org/10.1002/wat2.1330>
- Karger, DN, Conrad, O, Böhner, J, Kawohl, T, Kreft, H, Soria-Auza, RW, Zimmermann, NE, Linder, HP and Kessler, M.** 2017. 'Climatologies at high resolution for the earth's land surface areas'. *Sci Data*, 4: 170122. DOI: <https://doi.org/10.1038/sdata.2017.122>
- Karger, DN, Nobis, MP, Normand, S, Graham, CH and Zimmermann, NE.** 2023. 'CHELSA-TraCE21k – high-resolution (1&thinsp;km) downscaled transient temperature and precipitation data since the Last Glacial Maximum'. *Climate of the Past*, 19(2): 439–456. DOI: <https://doi.org/10.5194/cp-19-439-2023>
- Kilian, B, Özkan, H, Walther, A, Kohl, J, Dagan, T, Salamini, F and Martin, W.** 2007. 'Molecular Diversity at 18 Loci in 321 Wild and 92 Domesticate Lines Reveal No Reduction of Nucleotide Diversity during *Triticum monococcum* (Einkorn) Domestication: Implications for the Origin of Agriculture'. *Molecular Biology and Evolution*, 24(12): 2657–2668. DOI: <https://doi.org/10.1093/molbev/msm192>
- Kirkbride, D.** 1966. 'Five Seasons at the Pre-Pottery Neolithic Village of Beidha In Jordan'. *Palestine Exploration Quarterly*. DOI: <https://doi.org/10.1179/peq.1966.98.1.8>
- Kislev, ME.** 1989. 'Origins of the cultivation of *lathyrus sativus* and *L. Cicera* (fabaceae)'. *Economic Botany*, 43(2): 262–270. DOI: <https://doi.org/10.1007/BF02859868>
- Kopecký, M and Čížková, Š.** 2010. 'Using topographic wetness index in vegetation ecology: Does the algorithm matter?' *Appl. Veg. Sci.*, 13(4): 450–459. DOI: <https://doi.org/10.1111/j.1654-109X.2010.01083.x>
- Kozłowski, SK and Aurenche, O.** 2005. *Territories, boundaries and cultures in the Neolithic Near East*. Archaeopress. DOI: <https://doi.org/10.30861/9781841718071>
- Krzyżanska, M.** 2023. *Modelling spatio-temporal changes in the ecological niches of major domesticated crops in China: Application of species distribution modelling*. PhD thesis. University of Cambridge.
- Krzyżanska, M, Hunt, HV, Crema, ER and Jones, MK.** 2022. 'Modelling the potential ecological niche of domesticated buckwheat in China: Archaeological evidence, environmental constraints and climate change'. *Vegetation History and Archaeobotany*, 31(4): 331–345. DOI: <https://doi.org/10.1007/s00334-021-00856-9>
- Kuhn, M and Silge, J.** 2022. *Tidy Modeling with R: A Framework for Modeling in the Tidyverse*. O'Reilly.
- Ladizinsky, G and Adler, A.** 1976. 'The origin of chickpea *Cicer arietinum* L.'. *Euphytica*, 25(1): 211–217. DOI: <https://doi.org/10.1007/BF00041547>
- Leempoel, K, Parisod, C, Geiser, C, Daprà, L, Vittoz, P and Joost, S.** 2015. 'Very high-resolution digital elevation models: Are multi-scale derived variables ecologically relevant?' *Methods in Ecology and Evolution*, 6(12): 1373–1383. DOI: <https://doi.org/10.1111/2041-210X.12427>
- Lev-Yadun, S, Gopher, A and Abbo, S.** 2000a. 'The Cradle of Agriculture'. *Science*, 288(5471): 1602–1603. DOI: <https://doi.org/10.1126/science.288.5471.1602>
- Lev-Yadun, S, Gopher, A and Abbo, S.** 2000b. 'The Cradle of Agriculture'. *Science*, 288(5471): 1602–1603. DOI: <https://doi.org/10.1126/science.288.5471.1602>
- Levy, AA and Feldman, M.** 2022. 'Evolution and origin of bread wheat'. *The Plant Cell*, 34(7): 2549–2567. DOI: <https://doi.org/10.1093/plcell/koac130>
- Lindsay, JB.** 2016. 'Whitebox GAT: A case study in geomorphometric analysis'. *Comput. Geosci.*, 95: 75–84. DOI: <https://doi.org/10.1016/j.cageo.2016.07.003>
- Liu, C, Berry, PM, Dawson, TP and Pearson, RG.** 2005. 'Selecting thresholds of occurrence in the prediction of species distributions'. *Ecography*, 28(3): 385–393. DOI: <https://doi.org/10.1111/j.0906-7590.2005.03957.x>
- Liu, C, Newell, G and White, M.** 2016. 'On the selection of thresholds for predicting species occurrence with presence-only data'. *Ecology and Evolution*, 6(1): 337–348. DOI: <https://doi.org/10.1002/ece3.1878>
- Liu, C, White, M and Newell, G.** 2013. 'Selecting thresholds for the prediction of species occurrence with presence-only data'. *Journal of Biogeography*, 40(4): 778–789. DOI: <https://doi.org/10.1111/jbi.12058>
- Luan, J, Zhang, C, Xu, B, Xue, Y and Ren, Y.** 2020. 'The predictive performances of random forest models with limited sample size and different species traits'. *Fisheries Research*, 227: 105534. DOI: <https://doi.org/10.1016/j.fishres.2020.105534>
- Lucas, L and Fuller, D.** 2018. *Dataset: From intermediate economies to agriculture: Trends in wild food use, domestication and cultivation among early villages in southwest Asia*.

- Luo, M-C, Yang, Z-L, You, FM, Kawahara, T, Waines, JG and Dvorak, J.** 2007. 'The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication'. *Theoretical and Applied Genetics*, 114(6): 947–959. DOI: <https://doi.org/10.1007/s00122-006-0474-0>
- Maher, LA, Richter, T and Stock, JT.** 2012. 'The Pre-Natufian Epipaleolithic: Long-term Behavioral Trends in the Levant'. *Evolutionary Anthropology: Issues, News, and Reviews*, 21(2): 69–81. DOI: <https://doi.org/10.1002/evan.21307>
- Michczyński, A.** 2007. 'Is it Possible to Find a Good Point Estimate of a Calibrated Radiocarbon Date?' *Radiocarbon*, 49(2): 393–401. DOI: <https://doi.org/10.1017/S003382200042326>
- Miller, T, Blackwood, CB and Case, AL.** 2024. 'Assessing the utility of SoilGrids250 for biogeographic inference of plant populations'. *Ecology and Evolution*, 14(3): e10986. DOI: <https://doi.org/10.1002/ece3.10986>
- Mod, HK, Scherrer, D, Luoto, M and Guisan, A.** 2016. 'What we use is not what we know: Environmental predictors in plant distribution models'. *Journal of Vegetation Science*, 27(6): 1308–1322. DOI: <https://doi.org/10.1111/jvs.12444>
- Molina-Cano, J-L, Russell, JR, Moralejo, MA, Escacena, JL, Arias, G and Powell, W.** 2005. 'Chloroplast DNA microsatellite analysis supports a polyphyletic origin for barley'. *Theoretical and Applied Genetics*, 110(4): 613–619. DOI: <https://doi.org/10.1007/s00122-004-1878-3>
- Moore, AMT, Hillman, GC and Legge, AJ.** 2000. *Village on the Euphrates: From Foraging to Farming at Abu Hureyra*. Oxford: Oxford University Press.
- Mori, N.** 2003. 'Origins of domesticated emmer and common wheat inferred from chloroplast DNA fingerprinting'. In: *Proc. 10th. Int. Wheat Genetics Symp., Paestum, Italy*, 2003. 2003. pp. 25–28.
- Ozkan, H, Brandolini, A, Pozzi, C, Effgen, S, Wunder, J and Salamini, F.** 2005. 'A reconsideration of the domestication geography of tetraploid wheats'. *Theoretical and Applied Genetics*, 110(6): 1052–1060. DOI: <https://doi.org/10.1007/s00122-005-1925-8>
- Özkan, H, Brandolini, A, Schäfer-Pregl, R and Salamini, F.** 2002. 'AFLP Analysis of a Collection of Tetraploid Wheats Indicates the Origin of Emmer and Hard Wheat Domestication in Southeast Turkey'. *Molecular Biology and Evolution*, 19(10): 1797–1801. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a004002>
- Özkan, H, Willcox, G, Graner, A, Salamini, F and Kilian, B.** 2011. 'Geographic distribution and domestication of wild emmer wheat (*Triticum dicoccoides*)'. *Genetic Resources and Crop Evolution*, 58(1): 11–53. DOI: <https://doi.org/10.1007/s10722-010-9581-5>
- Peake, H and Fleure, HJ.** 1927. *Corridors of Time III: Peasants and Potters*. Oxford: Clarendon Press.
- Pumpelly, R.** 1908. *Explorations in Turkestan*. Carnegie Institution Publication 73. Washington, DC: Carnegie Institution.
- Purugganan, MD and Fuller, DQ.** 2009. 'The nature of selection during plant domestication'. *Nature*, 457(7231): 843–848. DOI: <https://doi.org/10.1038/nature07895>
- Riehl, S and Kümmel, C.** 2005. *Archaeobotanical Database Of Eastern Mediterranean And Near Eastern Sites (ADEMNEs)*. <https://www.ademnes.de/>.
- Riehl, S, Zeidi, M and Conard, NJ.** 2013. 'Emergence of Agriculture in the Foothills of the Zagros Mountains of Iran'. *Science*, 341(6141): 65–67. DOI: <https://doi.org/10.1126/science.1236743>
- Roberts, N, Woodbridge, J, Bevan, A, Palmisano, A, Shennan, S and Asouti, E.** 2018. 'Human responses and non-responses to climatic variations during the last Glacial-Interglacial transition in the eastern Mediterranean'. *Quaternary Science Reviews*, 184: 47–67. DOI: <https://doi.org/10.1016/j.quascirev.2017.09.011>
- Schreiber, M, Özkan, H, Komatsuda, T and Mascher, M.** 2021. 'Evolution and Domestication of Rye'. In: Rabanus-Wallace, MT and Stein, N (eds.) *The Rye Genome*. Cham: Springer International Publishing. pp. 85–100. DOI: [https://doi.org/10.1007/978-3-030-83383-1\\_6](https://doi.org/10.1007/978-3-030-83383-1_6)
- Shennan, SJ and Conolly, J.** 2007. *Dataset: The Origin and Spread of Neolithic Plant Economies in the Near East and Europe*. DOI: <https://doi.org/10.5284/1000093>
- Sillero, N, Arenas-Castro, S, Enriquez-Urzelai, U, Vale, CG, Sousa-Guedes, D, Martínez-Freiría, F, Real, R and Barbosa, AM.** 2021. 'Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling'. *Ecological Modelling*, 456: 109671. DOI: <https://doi.org/10.1016/j.ecolmodel.2021.109671>
- Smith, BD.** 1995. *The emergence of agriculture*. New York, NY: Scientific American Library.
- Stockwell, DRB and Peterson, AT.** 2002. 'Effects of sample size on accuracy of species distribution models'. *Ecological Modelling*, 148(1): 1–13. DOI: [https://doi.org/10.1016/S0304-3800\(01\)00388-X](https://doi.org/10.1016/S0304-3800(01)00388-X)
- Tanno, K and Willcox, G.** 2006. 'How Fast Was Wild Wheat Domesticated?' *Science*, 311(5769): 1886–1886. DOI: <https://doi.org/10.1126/science.1124635>
- Tanno, K and Willcox, G.** 2012. 'Distinguishing wild and domestic wheat and barley spikelets from early Holocene sites in the Near East'. *Vegetation History and Archaeobotany*, 21(2): 107–115. DOI: <https://doi.org/10.1007/s00334-011-0316-0>
- Townsend Peterson, A and Soberón, J.** 2012. 'Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right'. *NatCon*, 10(2): 102–107. DOI: <https://doi.org/10.4322/natcon.2012.019>
- Valavi, R, Elith, J, Lahoz-Monfort, JJ and Guillera-Arroita, G.** 2021. 'Modelling species presence-only data with random forests'. *Ecography*, 44(12): 1731–1742. DOI: <https://doi.org/10.1111/ecog.05615>
- Valavi, R, Guillera-Arroita, G, Lahoz-Monfort, JJ and Elith, J.** 2022. 'Predictive performance of presence-only species distribution models: A benchmark study with reproducible code'. *Ecological Monographs*, 92(1): e01486. DOI: <https://doi.org/10.1002/ecm.1486>

- van Zeist, W and Bottema, S.** 1991. *Late Quaternary vegetation of the Near East*. Tübinger Atlas des Vorderen Orients, Reihe A (Naturwissenschaften) 18. Wiesbaden: Reichert Verlag.
- Vavilov, N.** 1926. *Studies on the origin of cultivated plants*. Leningrad: Institute of Applied Botany and Plant Breeding.
- Velazco, SJE, Galvão, F, Villalobos, F and De Marco Júnior, P.** 2017. ‘Using worldwide edaphic data to model plant species niches: An assessment at a continental extent’. *PLoS One*, 12(10): e0186025. DOI: <https://doi.org/10.1371/journal.pone.0186025>
- Verhagen, P and Whitley, TG.** 2020. ‘Predictive Spatial Modelling’. In: *Archaeological Spatial Analysis*. Routledge. DOI: <https://doi.org/10.4324/9781351243858-13>
- von Humboldt, A.** 1807. *Essai sur la géographie des plantes*. Paris.
- Wallace, M, Jones, G, Charles, M, Forster, E, Stillman, E, Bonhomme, V, Livarda, A, Osborne, CP, Rees, M, French, G and Preece, C.** 2018. ‘Re-analysis of archaeobotanical remains from pre- and early agricultural sites provides no evidence for a narrowing of the wild plant food spectrum during the origins of agriculture in southwest Asia’. *Veg. Hist. Archaeobot.* DOI: <https://doi.org/10.1007/s00334-018-0702-y>
- Wallace, M, Livarda, A, Charles, M and Jones, G.** 2018. *Origins of agriculture: Archaeobotanical database*. DOI: <https://doi.org/10.5284/1046750>
- Weide, A, Riehl, S, Zeidi, M and Conard, NJ.** 2017. ‘Reconstructing subsistence practices: Taphonomic constraints and the interpretation of wild plant remains at aceramic Neolithic Chogha Golan, Iran’. *Vegetation History and Archaeobotany*, 26(5): 487–504. DOI: <https://doi.org/10.1007/s00334-017-0607-1>
- Weide, A, Riehl, S, Zeidi, M and Conard, NJ.** 2018. ‘A systematic review of wild grass exploitation in relation to emerging cereal cultivation throughout the Epipalaeolithic and aceramic Neolithic of the Fertile Crescent’. *PLOS ONE*, 13(1): e0189811. DOI: <https://doi.org/10.1371/journal.pone.0189811>
- Weiss, E, Kislev, ME and Hartmann, A.** 2006. ‘Autonomous Cultivation Before Domestication’. *Science*, 312(5780): 1608–1610. DOI: <https://doi.org/10.1126/science.1127235>
- Weiss, E, Wetterstrom, W, Nadel, D and Bar-Yosef, O.** 2004. ‘The broad spectrum revisited: Evidence from plant remains’. *Proceedings of the National Academy of Sciences*, 101(26): 9551–9555. DOI: <https://doi.org/10.1073/pnas.0402362101>
- Whitlam, J, Bogaard, A, Matthews, R, Matthews, W, Mohammadifar, Y, Ilkhani, H and Charles, M.** 2018. ‘Pre-agricultural plant management in the uplands of the central Zagros: The archaeobotanical evidence from Sheikh-e Abad’. *Vegetation History and Archaeobotany*, 27(6): 817–831. DOI: <https://doi.org/10.1007/s00334-018-0675-x>
- Willcox, G.** 2024. ‘Sowing, harvesting and tilling at the end of the Pleistocene/beginning of the Holocene in northern Syria: A reassessment of cereal and pulse exploitation’. *Vegetation History and Archaeobotany*. DOI: <https://doi.org/10.1007/s00334-023-00984-4>
- Willcox, G, Buxo, R and Herveux, L.** 2009. ‘Late Pleistocene and early Holocene climate and the beginnings of cultivation in northern Syria’. *The Holocene*, 19(1): 151–158. DOI: <https://doi.org/10.1177/0959683608098961>
- Willcox, G, Fornite, S and Herveux, L.** 2008. ‘Early Holocene cultivation before domestication in northern Syria’. *Vegetation History and Archaeobotany*, 17(3): 313–325. DOI: <https://doi.org/10.1007/s00334-007-0121-y>
- Wisz, MS, Hijmans, RJ, Li, J, Peterson, AT, Graham, CH, Guisan, A and NCEAS Predicting Species Distributions Working Group.** 2008. ‘Effects of sample size on the performance of species distribution models’. *Diversity and Distributions*, 14(5): 763–773. DOI: <https://doi.org/10.1111/j.1472-4642.2008.00482.x>
- Wright, MN and Ziegler, A.** 2017. ‘Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R’. *Journal of Statistical Software*, 77: 1–17. DOI: <https://doi.org/10.18637/jss.v077.i01>
- Yaworsky, PM, Hussain, ST and Riede, F.** 2023. ‘Climate-driven habitat shifts of high-ranked prey species structure Late Upper Paleolithic hunting’. *Scientific Reports*, 13(1): 4238. DOI: <https://doi.org/10.1038/s41598-023-31085-x>
- Yaworsky, PM, Hussain, ST and Riede, F.** 2024. ‘The effects of climate and population on human land use patterns in Europe from 22ka to 9ka ago’. *Quaternary Science Reviews*, 344: 108956. DOI: <https://doi.org/10.1016/j.quascirev.2024.108956>
- Yaworsky, PM, Nielsen, ES and Nielsen, TK.** 2024. ‘The Neanderthal niche space of Western Eurasia 145 ka to 30 ka ago’. *Scientific Reports*, 14(1): 7788. DOI: <https://doi.org/10.1038/s41598-024-57490-4>
- Yeomans, L, Martin, L and Richter, T.** 2017. ‘Expansion of the known distribution of Asiatic mouflon (*Ovis orientalis*) in the Late Pleistocene of the Southern Levant’. *Royal Society Open Science*, 4(8): 170409. DOI: <https://doi.org/10.1098/rsos.170409>
- Yousefi, M, Heydari-Guran, S, Kafash, A and Ghasidian, E.** 2020. ‘Species distribution models advance our knowledge of the Neanderthals’ paleoecology on the Iranian Plateau’. *Scientific Reports*, 10(1): 14248. DOI: <https://doi.org/10.1038/s41598-020-71166-9>
- Zeder, MA.** 2011. ‘The Origins of Agriculture in the Near East’. *Curr. Anthropol.*, 52(S4): S221–S235. DOI: <https://doi.org/10.1086/659307>
- Zeder, MA.** 2024. ‘Out of the Shadows: Reestablishing the Eastern Fertile Crescent as a Center of Agricultural Origins: Part 1’. *Journal of Archaeological Research*. DOI: <https://doi.org/10.1007/s10814-024-09195-5>
- Zohary, D.** 1969. ‘The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the Old World’. In: *The Domestication and Exploitation of Plants and Animals*. Routledge.

- Zohary, D and Hopf, M.** 1973. 'Domestication of Pulses in the Old World'. *Science*, 182(4115): 887–894. DOI: <https://doi.org/10.1126/science.182.4115.887>
- Zohary, D and Hopf, M.** 1988. *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. 1st ed. Oxford: Oxford University Press.
- Zohary, D and Spiegel-Roy, P.** 1975. 'Beginnings of Fruit Growing in the Old World'. *Science*, 187(4174): 319–327. DOI: <https://doi.org/10.1126/science.187.4174.319>
- Zohary, D, Weiss, E and Hopf, M.** 2012. *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. 4th ed. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:osobl/9780199549061.001.0001>
- Zohary, M.** 1973. *Geobotanical foundations of the Middle East*. Stuttgart: Gustav Eischer.

---

**TO CITE THIS ARTICLE:**

Roe, J and Arranz-Otaegui A. 2026. Biogeography of Crop Progenitors and Wild Plant Resources in the Terminal Pleistocene and Early Holocene of West Asia, 14.7–8.3 ka. *Open Quaternary*, 12: 2, pp.1–23. DOI: <https://doi.org/10.5334/oq.163>

**Submitted:** 01 April 2025    **Accepted:** 27 August 2025    **Published:** 03 February 2026

**COPYRIGHT:**

© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Open Quaternary* is a peer-reviewed open access journal published by Ubiquity Press.

