# Supplementary Information

The supplementary material to Hoebe et al 2023 below refers to excel and R files provided in the folder SupplMat The supplementary data can be used to check the methodology. Unfortunately a portion of the dates is not yet available for re-use, as it is under embargo until a later date. This material can be provided upon request. These supplements will be made available on the first author’s github: https://github.com/pirhoebe.

# Source overview

See SupplMat/SI-1 Source Overview.xlsx

Source overview gives metadata for the source databases and references used.

# Dataset

See SupplMat/SI-2 Supplementary dataset.xlsx

# Scripts

Scripts are available in R files in the SupplMat folder. Below the scripts are listed with their input, used subscripts, data output and plots.

1 Dataset preparation.R

Data input:

* Supplementary dataset.xlsx

Scripts used:

* read libraries – *script used to facilitate installing and reading packages*
* calibrate data.R – *script used to facilitate calibration*

Data output:

* data\_full (caldates, bins and chronozones added)
* data\_full.cal
* data\_vet (caldates, bins and chronozones added)
* data\_vet.cal
* chronozones (timing)
* data in timeframe.xlsx
* data count overview.xlsx

2 Aoristic weight.R

Data input:

* data\_full
* data\_vet
* chronozones

Scripts used:

* function aoristic().R – *custom function to calculate aoristic count per bin*
* phaserect.R – *script used to plot cold phase rectangles*

Data output:

* chronozones – *aoristic weights for dates and phases in the full and vetted dataset*

SI-5: chronozones aoristic distribution.xlsx

Plots:

* aoristic weight dates per chronozone
* aoristic weight phases per chronozone
* Fig. 2: comparison aoristic weight dates and phases per chronozone
* SI-5: comparison aoristic weight full and vetted dates per chronozone

3 Vetting impact.R

Data input:

* data\_full
* data\_vet

Scripts used:

* SPD aggregation.R – *script used to facilitate using the spd function*

Data output:

* pivot – *pivot table of dates per country in the full and vetted dataset*

SI 5: *vetting comparison by country.xlsx*

* data\_full.spd
* data\_vet.spd
* spd\_vetcompare

Plots:

* SI-5: vetting comparison plotting the full and vetted spds, and an spd of the difference in density

4 Frequency distribution.R

Data input:

* data\_full

Scripts used:

* calibrate data.R – *script used to facilitate calibration*
* SPD aggregation.R – *script that facilitates using the spd() function*
* KDE analysis.R – *scripts that facilitates using the ckde() function*
* phaserect.R – *script used to plot cold phase rectangles*

Data output:

* Intcal20
* NGRIPdata
* event – timing of cold phases
* data\_full\_n.cal
* data\_full\_n.spd

Plots:

* Multiplot of:

1. NGRIP d18O and climate chronozones
2. Binsense and bin medians
3. Non-normalised and normalised SPD with Intcal inset
4. KDE

5 Analysis.R

Data input:

* data\_full
* data\_full.cal
* data\_full.spd
* NGRIPdata

Scripts used:

* SPD aggregation.R – *script that facilitates using the spd() function*
* Modeltest.R– *script that facilitates using the spd() function*
* Modeltest climate.R – *script that facilitates using the modelTest() function*
  + *Output: supporting SI-7 figures for the climate modeltest*
  + *Figure 4: alignment of modes in NGRIP and SPD*
* phaserect.R – *script used to plot cold phase rectangles*

Data output:

* data\_full.spd\_m
* data\_full.expmod
* data\_full.climod
* perm.elevation
* perm.country
* perm.material

Plots:

* Fig. 5: exponential and climate model test
* Fig. 7: landscape zone permutation test
* Fig. 8: country permutation test
* Fig. 9: sample material permutation test

6 Calculations in the text.R

Data input:

* data\_full
* perm.elevation
* perm.country

Scripts used:

* calibrate data.R – *script used to facilitate calibration*
* SPD aggregation.R – *script that facilitates using the spd() function*

Data output:

* SI-8: Various calculations
* corylus, corylus.cal, corylus.spd
* pits , pits.cal, pits.spd

Plots:

* SI-8: Distribution of hazelnut and pit hearth dates

# Distribution of dates across source and country

See SupplMat/SI-4 SourcexCountry.xlsx

Below: number of dates and sites per country per source. This includes duplicates between sources and serves to illustrate only the distribution of data within each source. src= source, fin = final dataset, dup = duplicates

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Britain** | | | **Belgium** | | | **Netherlands** | | | **Germany** | | | **Denmark** | | | **Total** | | |
|  | **src** | **fin** | **dup** | **src** | **fin** | **dup** | **src** | **fin** | **dup** | **src** | **fin** | **dup** | **src** | **fin** | **dup** | **src** | **fin** | **dup** |
| **Databases** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ADS | 1135 | 669 | 466 |  |  |  |  |  |  |  |  |  |  |  |  | **1135** | **669** | **466** |
| CalPal 2020 | 881 | 167 | 714 | 419 | 165 | 254 | 589 | 421 | 168 | 1113 | 684 | 429 | 244 | 162 | 82 | **3246** | **1599** | **1647** |
| CIO Groningen | 167 | 35 | 132 | 234 | 76 | 158 | 2654 | 1170 | 1484 | 346 | 81 | 265 | 13 | 6 | 7 | **3414** | **1368** | **2046** |
| EuroEvol | 418 | 0 | 418 | 156 | 0 | 156 | 500 | 1 | 499 | 368 | 7 | 361 | 52 | 2 | 50 | **1494** | **10** | **1484** |
| KIK-IRPA | 5 | 5 | 0 | 681 | 530 | 151 | 22 | 20 | 2 |  |  |  |  |  |  | **708** | **555** | **153** |
| Niekus & Peeters Meso |  |  |  |  |  |  | 795 | 302 | 493 |  |  |  |  |  |  | **795** | **302** | **493** |
| ORAU | 320 | 36 | 284 | 82 | 1 | 81 | 8 | 0 | 8 | 120 | 2 | 118 | 5 | 0 | 5 | **535** | **39** | **496** |
| PACEA | 333 | 18 | 315 | 98 | 5 | 93 | 58 | 1 | 57 | 487 | 19 | 468 | 44 | 1 | 43 | **1020** | **44** | **976** |
| Palaeolithic DB v26 | 330 | 0 | 330 | 86 | 1 | 85 | 49 | 0 | 49 | 519 | 6 | 513 | 35 | 1 | 34 | **1019** | **8** | **1011** |
| RADON | 583 | 1 | 582 | 173 | 5 | 168 | 517 | 0 | 517 | 484 | 8 | 476 | 112 | 5 | 107 | **1869** | **19** | **1850** |
| **Literature** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Bevan 2017 | 595 | 378 | 217 |  |  |  |  |  |  |  |  |  |  |  |  | **595** | **378** | **217** |
| Cziesla 2015 | 8 | 0 | 8 | 26 | 1 | 25 | 35 | 0 | 35 | 193 | 33 | 160 | 2 | 0 | 2 | **264** | **34** | **230** |
| Gehlen EM2 Appendix 2 |  |  |  |  |  |  |  |  |  | 117 | 37 | 80 |  |  |  | **117** | **37** | **80** |
| Gehlen et al. 2020: CCA |  |  |  | 10 | 2 | 8 | 2 | 0 | 2 | 68 | 21 | 47 |  |  |  | **80** | **23** | **57** |
| Gehlen et al. 2020: pits |  |  |  |  |  |  |  |  |  | 131 | 56 | 75 |  |  |  | **131** | **56** | **75** |
| Griffiths & Robinson 2018 | 62 | 8 | 54 | 10 | 1 | 9 | 1 | 0 | 1 |  |  |  | 88 | 49 | 39 | **161** | **58** | **103** |
| Grimm & Weber 2008 |  |  |  |  |  |  | 16 | 0 | 16 | 43 | 2 | 41 | 11 | 0 | 11 | **70** | **2** | **68** |
| Jensen et al. 2020 |  |  |  |  |  |  |  |  |  |  |  |  | 158 | 85 | 73 | **158** | **85** | **73** |
| Maier 2015 |  |  |  | 38 | 7 | 31 |  |  |  | 206 | 23 | 183 |  |  |  | **244** | **30** | **214** |
| Steele & Shennan 2000 | 47 | 0 | 47 | 48 | 0 | 48 | 18 | 0 | 18 | 116 | 1 | 115 | 33 | 1 | 32 | **262** | **2** | **260** |
| Street et al. 2019 |  |  |  |  |  |  |  |  |  | 48 | 43 | 5 |  |  |  | **48** | **43** | **5** |
| Waddington & Wicks 2017 | 123 | 32 | 91 |  |  |  |  |  |  |  |  |  |  |  |  | **123** | **32** | **91** |
| Miscellaneous literature | 8 | 1 | 7 | 2 | 0 | 2 | 9 | 0 | 9 | 85 | 14 | 71 | 20 | 7 | 13 | **124** | **22** | **102** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **17612** | **5415** | **12197** |

Below: number of dates and sites per country by source, filtered for duplicates. Filtering for duplicates was done in a certain order to ensure data quality. Data from databases with a lot of information were retained. The difference in total dates and sites with the table above illustrates the overlap between sources. d= dates, p= phases (bins), s= sites

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Britain** | | | **Belgium** | | | **Netherlands** | | | **Germany** | | | **Denmark** | | | **Total** | | |
|  | d | p | s | d | p | s | d | p | s | d | p | s | d | p | s | **d** | **p** | **s** |
| **Databases** | **931** | **755** | **377** | **783** | **603** | **299** | **1915** | **1346** | **783** | **807** | **570** | **301** | **177** | **129** | **73** | **4613** | **3403** | **1833** |
| ADS | 669 | 542 | 251 |  |  |  |  |  |  |  |  |  |  |  |  | **669** | **542** | **251** |
| CalPal 2020 | 167 | 133 | 79 | 165 | 133 | 87 | 421 | 292 | 156 | 684 | 453 | 208 | 162 | 115 | 62 | **1599** | **1126** | **592** |
| CIO Groningen | 35 | 31 | 20 | 76 | 75 | 61 | 1170 | 930 | 589 | 81 | 77 | 64 | 6 | 6 | 3 | **1368** | **1119** | **737** |
| EuroEvol |  |  |  |  |  |  | 1 | 1 | 1 | 7 | 7 | 5 | 2 | 2 | 2 | **10** | **10** | **8** |
| KIK-IRPA | 5 | 3 | 1 | 530 | 383 | 142 | 20 | 15 | 4 |  |  |  |  |  |  | **555** | **401** | **147** |
| Niekus & Peeters Meso |  |  |  |  |  |  | 302 | 107 | 32 |  |  |  |  |  |  | **302** | **107** | **32** |
| ORAU | 36 | 28 | 13 | 1 | 1 | 1 |  |  |  | 2 | 2 | 2 |  |  |  | **39** | **31** | **16** |
| PACEA | 18 | 17 | 12 | 5 | 5 | 3 | 1 | 1 | 1 | 19 | 18 | 11 | 1 | 1 | 1 | **44** | **42** | **28** |
| Palaeolithic DB v26 |  |  |  | 1 | 1 | 1 |  |  |  | 6 | 6 | 5 | 1 | 1 | 1 | **8** | **8** | **7** |
| RADON | 1 | 1 | 1 | 5 | 5 | 4 |  |  |  | 8 | 7 | 6 | 5 | 4 | 4 | **19** | **17** | **15** |
| **Literature** | **419** | **271** | **162** | **11** | **9** | **8** |  |  |  | **230** | **162** | **91** | **142** | **85** | **64** | **802** | **527** | **325** |
| Bevan 2017 | 378 | 243 | 147 |  |  |  |  |  |  |  |  |  |  |  |  | **378** | **243** | **147** |
| Cziesla 2015 |  |  |  | 1 | 1 | 1 |  |  |  | 33 | 24 | 11 |  |  |  | **34** | **25** | **12** |
| Gehlen EM2 Appendix 2 |  |  |  |  |  |  |  |  |  | 37 | 34 | 19 |  |  |  | **37** | **34** | **19** |
| Gehlen et al. 2020: CCA |  |  |  | 2 | 2 | 1 |  |  |  | 21 | 11 | 6 |  |  |  | **23** | **13** | **7** |
| Gehlen et al. 2020: pits |  |  |  |  |  |  |  |  |  | 56 | 45 | 24 |  |  |  | **56** | **45** | **24** |
| Griffiths & Robinson 2018 | 8 | 7 | 4 | 1 | 1 | 1 |  |  |  |  |  |  | 49 | 11 | 2 | **58** | **19** | **7** |
| Grimm & Weber 2008 |  |  |  |  |  |  |  |  |  | 2 | 2 | 2 |  |  |  | **2** | **2** | **2** |
| Jensen et al. 2020 |  |  |  |  |  |  |  |  |  |  |  |  | 85 | 70 | 60 | **85** | **70** | **60** |
| Maier 2015 |  |  |  | 7 | 5 | 5 |  |  |  | 23 | 19 | 14 |  |  |  | **30** | **24** | **19** |
| Steele & Shennan 2000 |  |  |  |  |  |  |  |  |  | 1 | 1 | 1 | 1 | 1 | 1 | **2** | **2** | **2** |
| Street et al. 2019 |  |  |  |  |  |  |  |  |  | 43 | 17 | 6 |  |  |  | **43** | **17** | **6** |
| Waddington & Wicks 2017 | 32 | 20 | 10 |  |  |  |  |  |  |  |  |  |  |  |  | **32** | **20** | **10** |
| Miscellaneous literature | 1 | 1 | 1 |  |  |  |  |  |  | 14 | 9 | 8 | 7 | 3 | 1 | **22** | **13** | **10** |
| **Grand Total** | **1350** | **1026** | **539** | **794** | **612** | **307** | **1915** | **1346** | **783** | **1037** | **732** | **392** | **319** | **214** | **137** | **5415** | **3930** | **2158** |
| **Total unique records** | **1350** | **974** | **478** | **794** | **596** | **291** | **1915** | **1334** | **764** | **1037** | **691** | **349** | **319** | **207** | **130** | **5415** | **3802** | **2012** |

# Restructuring and data transformation procedure

Datasets were restructured using R, renaming, combining and removing columns to arrive at a unified structure. In the restructuring process, dates were removed if they did not belong to the spatiotemporal framework (Britain, Belgium, Netherlands, Germany, Denmark; 16000 – 7500 cal BP). Unfortunately the original datasets and restructuring scripts cannot be supplied here at this time.

After restructuring the dataset data in some columns were cleaned up, removing incongruities e.g. in the country column and in the labcode column (esp. issues with capitalisation and hyphens).

Similar incongruities had to be dealt with in the sitename column. This required manually going through the unique values and combining variations under a single identifier.

A subset of the dataset did not have coordinates, which was necessary for assigning sites to correct geographical regions. A large portion of these dates belonged to sites for which coordinates were available but not yet assigned. For others coordinates were estimated using the primary literature in which the date was published. Final published coordinates were rounded to the km. Additionally the dataset was appended with elevation data from Copernicus.org [[1]](#footnote-1)

Information on the sampling material or species was often but not always clearly present as a separate category in the different datasets. The process to find this information in unstructured text involved combining information from several columns (e.g. sample descriptions, comments, etc.) into the ‘sample’ column, and searching it for relevant keywords, using string detection (str\_detect())*.*

Example:

|  |
| --- |
| tooth\_ivory <- c("tooth", "molar", "ivory", "tusk")  for (CatNr in 1:length(tooth\_ivory)) { xMat <- tooth\_ivory[CatNr]  message(paste( "Material:", CatNr, xMat ))  data$MatCat[str\_detect(data$material, fixed(xMat, ignore\_case = TRUE))] <- "tooth / ivory"} |

# Vetting procedure and data density

To assess the impact of vetting, less reliable sample materials and animal species were removed, for instance because of the possible impact of old-wood effect or reservoir effect. The data is divided amongst the following material categories: Shell\*, sediment\*, wood, other\*, pottery, botanical remains, charcoal, bone, antler, bone / antler, tooth/ivory, amino acid, adhesives, same sample\*, bulk\*, bone collagen. Categories highlighted with an asterisk were deselected. Dates on material from animal species that can produce unreliable dates due to their diet (reservoir effects), notably humans, carnivores and omnivores as well as aquatic species and reptiles. The vetting procedure as presented in the script ‘SupplMat/scripts/1 Dataset preparation.R’ is repeated here:

|  |
| --- |
| data <- read\_excel("data/Supplementary dataset final.xlsx")  # 1 Vetting procedure ####  data$vet <- 1 # adding vetting column  # remove dates on unreliable materials  data$vet[data$material %in% c("bulk","other", "same sample",  "sediment", "shell")] <- 0  # remove dates on aquatic species, reptiles, omnivores and carnivores.  data$vet[data$species %in% c("Homo","Ursus", "Vipera",  "Lynx", "Panthera", "Canis lupus",  "Canis familiaris", "Frogs",  "Esox", "Anguis", "Toads", "Turtles",  "Vulpes", "Meles")] <- 0  #remove dates from bulk samples.  data$vet[like (data$sample, "bulk")] <- 0  data\_vet <- filter(data, data$vet==1)  data\_full <- data |

The impact of vetting is demonstrated in the script ‘SupplMat/scripts/3 Vetting impact.R’. Hinz[[2]](#footnote-2) has tested the sensitivity of SPDs through modelling them at different sample sizes and event intensity, and found earlier critiques that call the method’s ability to reflect past changes into question[[3]](#footnote-3) are only partly justified. Detectability of changes depends on the data density (dates per century), the strength of the signal (i.e. the severity of the change) and the certainty of identification. When measures are taken to reduce the chance of false positives (which is done in this study through rcarbon’s modelTest() function), a variety of events are detectable at realistic sample densities (see table below, after Hinz 2020: 247, table 3). SPDs with a sample density of ca. 100 dates per centuries for example, will reliably detect a decrease in human activity by 70%, detect a 50% reduction 9/10 times, and a 40% reduction around 7/10 times. These modelling results show what levels of trust and expectation are appropriate for SPDs, and can be used to ascertain whether a sample size is sufficient.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Density: | Event signal strength – activity reduced by: | | | | | | |
| Dates/year | 70% | 60% | 50% | 40% | 30% | 20% | 10% |
| 0.143 | 0.560 | 0.390 | 0.275 | 0.195 | 0.170 | 0.145 | 0.090 |
| 0.286 | 0.810 | 0.590 | 0.495 | 0.350 | 0.215 | 0.175 | 0.120 |
| 0.429 | 0.915 | 0.735 | 0.610 | 0.390 | 0.255 | 0.140 | 0.095 |
| 0.571 | 0.970 | 0.900 | 0.770 | 0.480 | 0.325 | 0.230 | 0.125 |
| 0.714 | 0.980 | 0.950 | 0.830 | 0.635 | 0.310 | 0.280 | 0.125 |
| 0.857 | 0.995 | 0.980 | 0.855 | 0.625 | 0.380 | 0.260 | 0.140 |
| 1.000 | 1.000 | 0.985 | 0.895 | 0.680 | 0.525 | 0.240 | 0.105 |

Sample density is calculated below for dates as well as site phases. We compare the full and vetted dataset across climate chronozones and countries. Density was calculated with an aoristic method, first calibrating the dates, then assigning them proportionally to each climate chronozones, for example:

* Say a date’s calibrated values span 11700 – 11500 cal BP.
* the chronozone boundary between the Younger Dryas and the Preboreal is at 11653 cal BP.

**>** 47/200 years are assigned to the Younger Dryas, i.e. 0.235

**>** 153/200 years are assigned to the Preboreal, i.e. 0.765.

A radiocarbon date that falls wholly within a chronozone counts as 1. The total aoristic number of dates per chronozones is divided by the duration of the chronozone in centuries, to arrive at the amount of dates or sites phases per century. See the script ‘2 Aoristic calculation.R’ for the full procedure. Climate events boundaries are based on Rasmussen 2014, Bond et al. 2001.

The tables below show the aoristic count and weight and are the basis for the aoristic weight plot below and in the main text. The percentage of dates and site phases lost per chronozones is visualised in red to demonstrate the non-uniform impact of vetting on the dataset.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | full | | | | vetted | | | | vetting impact | |
|  |  |  |  | dates | | phases | | dates | | phases | | dates | phases |
| name | llim | ulim | duration | count | weight | count | weight | count | weight | count | weight | %lost | %lost |
| O. Dryas | 16000 | 14650 | 1350 | 268.2 | 0.2 | 176.5 | 0.13 | 182.5 | 0.14 | 116.7 | 0.09 | 31.9% | 33.9% |
| Bølling | 14650 | 13904 | 746 | 263.6 | 0.35 | 177.5 | 0.24 | 168.4 | 0.23 | 108.0 | 0.14 | 36.1% | 39.1% |
| Allerød | 13904 | 12846 | 1058 | 544.3 | 0.51 | 427.6 | 0.4 | 314.6 | 0.3 | 237.6 | 0.22 | 42.2% | 44.4% |
| Y. Dryas | 12846 | 11653 | 1193 | 344.9 | 0.29 | 289.6 | 0.24 | 188.0 | 0.16 | 152.9 | 0.13 | 45.5% | 47.2% |
| Preboreal | 11653 | 10800 | 853 | 475.4 | 0.56 | 323.1 | 0.38 | 317.4 | 0.37 | 204.1 | 0.24 | 33.2% | 36.8% |
| Early Boreal | 10800 | 9190 | 1610 | 1449.8 | 0.9 | 891.7 | 0.55 | 1035.8 | 0.64 | 627.6 | 0.39 | 28.6% | 29.6% |
| Late Boreal | 9190 | 8090 | 1100 | 1180.9 | 1.07 | 822.9 | 0.75 | 799.2 | 0.73 | 540.1 | 0.49 | 32.3% | 34.4% |
| Atlantic | 8090 | 7500 | 590 | 657.6 | 1.11 | 503.1 | 0.85 | 395.2 | 0.67 | 284.1 | 0.48 | 39.9% | 43.5% |

Timeline

Description automatically generated

The vetting procedure also impacts different regions within the dataset in a non-uniform manner.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | dates | | | phases | | | sites | | |
| country | full | vetted | % lost | full | vetted | % lost | full | vetted | % lost |
| Britain | 1350 | 893 | 33.85% | 974 | 631 | 35.22% | 478 | 339 | 29.08% |
| Belgium | 794 | 541 | 31.86% | 596 | 376 | 36.91% | 291 | 169 | 41.92% |
| Netherlands | 1915 | 1110 | 42.04% | 1334 | 695 | 47.90% | 764 | 370 | 51.57% |
| Germany | 1037 | 785 | 24.30% | 691 | 526 | 23.88% | 349 | 263 | 24.64% |
| Denmark | 319 | 217 | 31.97% | 207 | 159 | 23.19% | 130 | 109 | 16.15% |
| **Total** | **5415** | **3546** | **34.52%** | **3802** | **2387** | **37.22%** | **2012** | **1250** | **37.87%** |

Difference between the full and vetted SPDs are visualised below with the mean (0.19) and 1sd range (±0.076). It shows again the vetting impact is not uniform across the temporal frame. The periods before ca. 14850 calBP and after ca. 9000 calBP are respectively least and most affected by vetting.

Chart, line chart

Description automatically generated

The effect of bandwidth size on the overall pattern of the KDE was explored as well. This has a the smoothing effect on century scale irregularities in the temporal pattern, but retains the overall pattern of increased summed probability during interstadial regimes and decreased summed probability during stadial regimes.

A picture containing chart

Description automatically generated

# Custom climate modeltest

See ‘SupplMat/scripts/Modeltest climate.R’

The Summed Probability Distribution (SPD) for Northwest Europe was tested using rcarbon’s modeltest(). Besides testing the data against a standard model of exponential growth, a custom climate based modeltest was used. The model is based on d18O data from NGRIP, a reliable proxy for northern hemisphere temperature changes (North Greenland Ice Core Project members 2004; data obtained from Bazin et al. 2013). To test how closely the SPD as a proxy for human activity corresponds to temperature, we standardized and interpolated the NGRIP d18O data and used this as a model in modeltest(). Here graphs and statistics are provided to supporting the model choice and standardisation process.

Chart

Description automatically generated

The overall relative patterns of NGRIP (red, above) and SPD (black, below) appear to show synchronous periods of high and low temperature and human activity respectively. To demonstrate the significance of the correlation between the two proxies we provide the scatterplot below with Spearman’s rho (r­­s =0.902, p<0.000) using cor.test().

Chart, scatter chart

Description automatically generated

To use the climate dataset as a model, it first needed to be interpolated using approx(), so that there are values for every year, and subsequently the data was standardized so that both datasets are at the same scale with mean = 0. This was done using:

standardized <- df %>% mutate\_at(c('variablename'), ~(scale(.) %>% as.vector))

Following this, the data was transformed again to fit the original distribution of the SPD. The result of standardisation is demonstrated in the below frequency distributions of relative values across both datasets, with higher values on the x axes representing higher temperature and higher probability density respectively. It is clear that both distributions are bimodal, caused by difference in temperature and human activity during stadials (GS2, GS1) and interstadial (GI-1, Early Holocene).

Chart, histogram

Description automatically generated

The model fit is shown below with the modes now corresponding to the stadial and interstadial values of both summed probability density and the climate model.

Chart, histogram

Description automatically generated

See section 5.1, figure 4.

# Calculations

In the assessment of biases the size of several subsets were calculated, which can be followed in the script ‘SupplMat/scripts/6 Calculations in text.R’.

*Landscape zones*

The permutation test dividing the dataset into landscape zones shows a significant overrepresentation of the middle zone (between 10 and 50 meters above ordinance level) during GS-1 and the Preboreal. This is possibly because the large number of dates from the Netherlands skew these results.

Of the 533 dates in this zone and period, 234 (43.9 %) are from the Netherlands.

Similarly, the low landscape zone (<=10 meters above OL) is overrepresented by the large number of Dutch dates from the 8.2ka event onwards. This is due to a combination of factors: e.g. accessibility of coastal contexts to researchers, past migration into this regions and preservation bias.

Of the 614 dates in this zone and period, 419 (68.2%) are from the Netherlands.

*Sampling procedures*

The Early and Late Boreal each show a prominent peak in the frequency distribution. The permutation tests for countries show that the Early Boreal peak is absent in the Dutch dataset while the Late Boreal peak is absent in the Belgium dataset. Additionally, the permutation test for sample material shows that the Early Boreal peak is less pronounced the charcoal subset and the Late Boreal peak is absent from the plant remains subset. While these differences could in part relate to real differences in past activity, we think differences in sampling practices between Belgium and the Netherlands should be considered.

We note the preference for dating hazelnuts on Mesolithic sites in Belgium:

Of the 338 Early Boreal dates done on hazelnuts, 34 (10.1%) are from the Netherlands, while 119 (35.2%) are from Belgium.

Hazelnut accounts for 6% of the total dates from the Netherlands in this period (569)

Hazelnut accounts for 37.3% of the total dates from Belgium in this period (319)

We note the preference for dating charcoal from pit hearths on Mesolithic sites in the Netherlands, which is not common practice in Belgium:

Of the 383 Late Boreal dates from the Netherlands, 223 (58.2%) are from pit hearths.

Of the 72 Late Boreal dates from Belgium, 7 (9.2%) are from pit hearths.

Generally we can see dates on hazelnut and dates on pit hearths contribute considerably to the prominence of the two Boreal peaks.

Chart, histogram

Description automatically generated

# Source bias SPDs

The temporal focus of different source datasets is visualised here with SPDs filtered to the spatiotemporal framework. Unfortunately the datasets in their original form cannot be supplied at this time. Some datasets were combined:

data\_Lit: combines Miscellaneous literature and Grimm & Weber 2008

data\_mesoDE: combines Gehlen EM2, Gehlen et al. 2020 CCA, Gehlen et al 2020 Mesolithic pits and Street et al. 2019.

data\_8200: combines Griffiths & Robinson 2018 and Waddington and Wicks 2017

data\_cem is Maier 2015 (Central European Magdalenian)

data\_vermeersch is the Palaeolithic DB (v26)

SPDs are given below with blue sections indicating GS-1, 11,4 event, 9.3 event and 8.2 event:

Chart, histogram

Description automatically generatedChart, timeline

Description automatically generated

Chart

Description automatically generatedChart

Description automatically generated

Chart

Description automatically generatedChart

Description automatically generated

Chart, timeline

Description automatically generatedChart

Description automatically generated

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

Chart

Description automatically generatedChart, histogram

Description automatically generated

Chart, histogram

Description automatically generatedChart

Description automatically generated

Chart, timeline

Description automatically generatedChart, histogram

Description automatically generated

1. European Environment Agency, ‘European Digital Elevation Model (EU-DEM), version 1.1’, *Copernicus Land Monitoring Service*, (2018). [↑](#footnote-ref-1)
2. M. Hinz, ‘Sensitivity of radiocarbon sum calibration’, *Journal of Computer Applications in Archaeology*, 3/1 (2020), 238–52. [↑](#footnote-ref-2)
3. D. A. Contreras and J. Meadows, ‘Summed radiocarbon calibrations as a population proxy : a critical evaluation using a realistic simulation approach’, *Journal of Archaeological Science*, (2014), 1–18. [↑](#footnote-ref-3)