UNIVERSITY OF COLOGNE

DOCTORAL THESIS

Development and Application of a Climate-driven Human Dispersal Model

---

Autor:
Christian WEGENER
aus Oberhausen, Deutschland

Betreuer:
Prof. Dr. Yaping SHAO

Inaugural-Dissertation zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln

20. August 2021

Berichterstatter: (Gutachter)	Prof. Dr. Yaping Shao Prof. Dr. Frank Schäbitz
Vorsitz der Prüfungskommission:	Prof. Dr. Michael Staubwasser
Beisitzender:	Dr. Mark Reyers
Tag der mündlichen Prüfung:	20.10.2021

UNIVERSITY OF COLOGNE

Abstract

Faculty of Mathematics and Natural Sciences

Institute of Geophysics and Meteorology

Development and Application of a Climate-driven Human Dispersal Model

by Christian WEGENER

The origin and spread of the Anatomically Modern Human (AMH) is still debated and regularly renewed as new findings challenge the current consensus. While the fossil remains suggest a gradual evolution and spread, molecular genetic studies propose an outburst of AMH that spread out and became dominant. Human population and migration models are used to bridge over ranges of time and space that lack sufficient data, utilizing diffusion-reaction models to simulate the spread and population dynamics. The aim of this study is to expand these migration models with a directed dispersal rather than the isotropic spread through diffusion. To achieve this, spatial and temporal climate simulation data is used in conjunction with archaeological data to estimate the Human Existence Potential (HEP), a measure for the survivability for a specific culture of AMH. The directed dispersal is proportional to the gradient of this potential field and the population distribution, combining different factors that can trigger migration.

The dispersal model is presented in detail as part of the Human Modeling Framework, which combines several models and applications that build upon the HEP. Idealized and realistic test scenarios are broad out to evaluate the performance of the dispersal model, which show reasonable results in regards to numeric stability and consistency. The realistic simulations include a representative Dansgaard-Oeschger event for the Aurignacian technocomplex and a simulation of monthly changing HEP for the Last Glacial Maximum. The former produces a reasonable response, while the latter illustrates the limitations of the dispersal model. As the final application, a complete simulation of AMH dispersal from the Levant to Europe from 45-25 kyr BP shows promising results regarding possible dispersal routes and the timing of first arrivals. The uncertainties in radiocarbon dating make it difficult, however, to evaluate the estimation of the timings. The presented model is a feasible foundation that can be further expanded, for example, by using more than one population and include social interactions in a simplified formulation.

Kurzzusammenfassung

Der Ursprung und die Ausbreitung des anatomisch modernen Menschen (AMH) ist Gegenstand kontinuierlicher Diskussionen, die regelmäßig aufflammen sobald neue Ergebnisse vorliegen, die den bisherigen Konsens infrage stellen. Während archäologische Funde eine schrittweise Entwicklung und Ausbreitung nahelegen, kommen genetische Studien zu dem Ergebnis, dass AMH sich schlagartig entwickelten und verbreiteten, bis sie sich schließlich als die dominierende Art durchsetzten. Populations- und Migrationsmodelle für den modernen Menschen werden benutzt, um Zeitperioden und Regionen abzuschätzen, in denen keine oder nicht ausreichende Daten vorliegen. Das Ziel dieser Studie ist es, ein Migrationsmodell zu entwickeln, dass nicht nur eine isotrope Ausbreitung durch Diffusion beinhaltet, sondern eine gerichtete Ausbreitung besitzt. Um dies zu erreichen werden paleoklimatische Simulationen mit archäologischen Datensätzen verbunden, um das menschliche Existenzpotential (HEP) zu berechnen, eine Maßeinheit für die Überlebensfähigkeit einer ausgewählten Kultur von AMH. Die gerichtete Ausbreitung ist proportional zum Gradienten dieses Potenzialfeldes und der Verteilung der Population. Diese Formulierung beinhaltet verschiedene Faktoren, welche die Migration auslösen können.

Das entwickelte Migrationsmodell ist Teil eines Frameworks zur Modellierung des modernen Menschen, welches verschiedene Methoden und Anwendungen vereint, die auf dem HEP basieren. Es wird in dieser Abhandlungen im Detail beschreiben und getestet. Dazu gehören idealisierte und realistische Szenarien, um das Verhalten des Modells zu testen und zu beurteilen. Letztere beinhalten die Simulation eines Dansgaard-Oeschger-Ereignisses für das Aurignacian, einer frühen Kultur des AMH in Europa, und eine Darstellung des letzten eiszeitlichen Maximums mit monatlichen wechselnden HEPs. Erstere liefert angemessene Ergebnisse, die den Erwartungen entsprechen, während letztere die Grenzen des Modells aufzeigen. Als letztes Experiment wird eine umfängliche Simulation der Migration des AMH von der mediterranen Ostküste nach Europa durchgeführt. Diese Simulation entspricht dem Zeitraum von 45 bis 25 kyr BP und liefert vielversprechende Ergebnisse in Hinsicht auf Ausbreitungsrouten und Zeiten der ersten Ankunft von AMH. Die Unsicherheiten in der C-14-Datierung von archäologischen Fundstücken machen es jedoch schwierig, die Genauigkeit der Modellergebnisse abzuschätzen. Das präsentierte Migrationsmodell ist eine gute Grundlage für weiterführende Forschung und kann weiter ausgebaut werden, beispielsweise durch das Betrachten von mehr als einer Gruppierung von AMH gleichzeitig und deren sozialen Interaktionen untereinander in einer vereinfachten mathematischen Formulierung.

Acknowledgements

This work is the result of four years of collaborative research and work, which would not have been possible without a multitude of people. Here, I would like to express my gratitude to all of them.

I cannot begin to express my thanks to my Doctoral Committee, which accompanied me through the whole process. This committee includes my supervisor Prof. Dr. Yaping Shao, for the meteorological core of this work, Prof. Dr. Andreas Zimmermann, for the archaeological and anthropological background and application, and Dr. Masoud Rostami, for the daily mentoring.

I would also like to extend my deepest gratitude to Dr. Sven Ulbrich and my colleague Konstantin Klein, who were heavily involved in the project. Sven helped me to settle in the workgroup in the first year. With Konstantin, we laid the foundation of our framework and reflected on the work of each other to further improve our models.

I would like to thank several members of the CRC806 for the closely knitted collaboration. Prof. Dr. Andreas Hense helped me create the model, especially in the early conceptual phase. Dr. Isabell Schmidt provided us with an extensive database of archaeological sites and gave us insight into the mechanisms of hunter-gatherer dispersal. Special thanks to Dr. Christian Willmes, who maintained a GitLab within the CRC, allowing us to easily manage different versions of our programs and make the programming phase much more bearable.

Many thanks to my colleagues and fellow Ph.D. students in the Institute for Geophysics and Meteorology, which provided helpful feedback and much needed distraction several times. I'd like to acknowledge the assistance of my colleagues in the CRC806, including fellow Ph.D. students, postdocs, and principal investigators, providing fruitful discussions on many occasions.

Special thanks to the supervisors of both Gradual Schools: Dr. Karin Boessenkool from the GSGS; Dr. Daniela Holst and Dr. Hannar Hartung from the IRTG. They provided us students with many occasions to learn new skills and improve ourselves.

Last but not least, many thanks to my family, especially to my wife Tabea, for encouraging and supporting me through the long journey of academic education.

This thesis was created as part of the CRC 806 "Our Way to Europe" funded by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG, project ID 57444011). The computation time used for the creation of the presented simulations was provided by the German Climate Computing Center (Deutsches Klimarechenzentrum DKRZ, Hamburg) as part of Project 965.

Contents

List of Figures

List of Tables

List of Abbreviations

AcHEP	**Ac**cessible **H**uman **E**xistence **P**otential
AMH	**A**natomically **M**odern **H**umans
AUC	**A**rea **U**nder a receiver operation characteristics **C**urve
AvHEP	**Av**ailable **H**uman **E**xistence **P**otential
BP	**B**efore **P**resent
BSS	**B**rier **S**kill **S**core
calBP	**Cal**ibrated time **B**efore **P**resent
CCSM4	**C**ommunity **C**limate **S**ystem **M**odell version **4**
CMIP5	**C**oupled **M**odel **I**ntercomparison **P**roject **5**
CNRM-CM5	**C**entre **N**ational de **R**echerches **M**étéorologique - **C**limatique **M**odèle **5**
CRC806	**C**ollaborative **R**esearch **C**enter 806
CRWM	**C**onstrained **R**andom **W**alk **M**odel
CTA	**C**lassification **T**ree **A**nalysis
D-O event	**D**ansgaard-**O**eschger event
EHC	**E**nvironmental **H**uman **C**atchment
EHEP	**E**nvironmental **H**uman **E**xistence **P**otential
GAM	**G**eneralized **A**dditive **M**odelling
GARP	**G**enetic **A**lgorithm for **R**ule-Set **P**rediction
GBM	**G**eneralized **B**oosting **M**ethods
GLM	**G**eneralized **L**inear **M**odelling
HDM	**H**uman **D**ispersal **M**odel
HEP	**H**uman **E**xistence **P**otential
HMF	**H**uman **M**odeling **F**ramework
ICON	**ICO**sahedral **N**onhydrostatic (model)
LASSO	**L**east **A**bsolute **S**election and **S**hrinkage **O**perator
LCP	**L**east **C**ost **P**ath
LGM	**L**ast **G**lacial **M**aximum
LOVECLIM	**LO**ch–**V**ecode-**E**cbilt-**CL**io-ag**I**sm **M**odel
MIS	**M**arine **I**sotope **S**tadium
NCAR	**N**ational **C**enter for **A**tmospheric **R**esearch
PMIP3	**P**aleoclimate **M**odeling **I**ntercomparison **P**roject 3
RF	**R**andom **F**orest
RK4	**R**unge-**K**utta **4**-step (method)
SDM	**S**pecies **D**istribution **M**odels
SVM	**S**upport **V**ector **M**achines
TVD	**T**otal-**v**ariation **D**iminishing (property)
VIF	**V**ariance **I**nflation **F**actor
WGS84	**W**orld **G**eodetic **S**ystem 19**84**

List of Symbols

Latin Symbols

A	Area of a grid cell	km
a^x, a^y	Local numerical speed	$\mathrm{km\,yr^{-1}}$
B	Sources or "Birth"-term	$\mathrm{Humans/100km^2yr}$
b_i	Predictor variable	variable
c	Constant (used for carrying capacity)	$\mathrm{Humans/100km^2}$
D	Sinks or "Death"-term	$\mathrm{Humans/100km^2yr}$
f_f	Forest fraction	$\mathrm{m^2\,m^{-2}}$
$f_i(\vec{p}$	HEP model / conversion function	1
h	Elevation	m
H^x, H^y	Advection fluxes	$\mathrm{Humans \cdot km/gridcell \cdot yr}$
$\vec{j}$	Flux of Human density	$\mathrm{Humans \cdot km/100km^2yr}$
j, k	Location indices	1
K	Diffusion parameter	$\mathrm{km^2\,yr^{-1}}$
$k_1, ..., k_4$	Runge-Kutta sub-steps	$\mathrm{Humans/yr \cdot gridcell}$
L	Borderloss	$\mathrm{Humans/100km^2yr}$
Δl	Characteristic length scale	km
N, M	Size of the used grid	1
P^x, P^y	Diffusion fluxes	$\mathrm{Humans \cdot km/cell \cdot yr}$
$\vec{p}$	Vector of second-order polynomials	1
r_B	Growth parameter	$\mathrm{yr^{-1}}$
R_{hs}	Right hand side of the numeric equation	$\mathrm{Humans/yr \cdot gridcell}$
S	Conversion parameter (Weibull distribution)	$\mathrm{Humans/100km^2}$
t	Time	yr
Δt	Numerical time step	yr
$\vec{v}$	Dispersal velocity vector	$\mathrm{km\,yr^{-1}}$
v^{crt}	Critical velocity	$\mathrm{km\,yr^{-1}}$
v^{local}	Local averaged velocity	$\mathrm{km\,yr^{-1}}$
v_{max}	Maximum velocity	$\mathrm{km\,yr^{-1}}$
V^x, V^y	Velocity in x/y direction	$\mathrm{km\,yr^{-1}}$
$\Delta x, \Delta y$	spatial spacing of the grid	km

Greek Symbols

α	Scaling parameter (Weibull distribution)	1
β	Scaling parameter (Weibull distribution)	1
β	Coefficients (HEP logistic regression)	1
γ	Velocity scaling parameter	$\mathrm{km^2 100km^2/Humans \cdot yr^2}$

δ	Difference between the sum of population change and the population dynamics	Humans/100km^{2}yr
ϵ	Advection-Diffusion Fraction	1
η	Flux-Limiter parameter	1
θ	Scaled density (Weibull distribution)	1
θ_{max}	Position of the maximum of the Weibull distribution	1
κ	Reproduction Onset	Humans/100km^2
λ	Absolute number of humans per grid cell	Humans/gridcell
$\lambda^{\pm}$	Intermediate values of λ between two grid points	Humans/gridcell
μ	Friction parameter	yr^{-1}
μ_i	Fractional vegetation types	m^2 m^{-2}
ν_i	Carrying capacity per vegetation type	Humans/100km^2
ρ	Human density	Humans/100km^2
ρ_{min}	Minimum Human density	Humans/100km^2
σ_{b_i}	Standard deviation of the predictor variable b_i	variable
σ_{ele}	Elevation parameterization	1
σ_{elestd}	Standard deviation of elevation	1
σ_{for}	Forest and Vegetation parameterization	1
σ_h	Standard deviation of elevation	m
σ_{ice}	Ice and glacial mask	1
σ_{water}	Water mask	1
$\Delta\tau$	Reaction time of Humans to changing Potential	yr
Φ_{acc}	Accessible Human Existence Potential	1
Φ, Φ_{av}	Available Human Existence Potential	Humans/100km^2
Φ_E	Environmental Human Existence Potential	1
Φ_{max}	Carrying capacity	Humans/100km^2
$\left(\vec{\nabla}\Phi\right)_{max}$	Maximum gradient of potential	Humans/100km^3
ξ	Angle between gradient direction and x direction	1

Chapter 1

Introduction

One of the crucial questions of archaeology and human biology is the origin and spread of the modern human. Different scientific fields try to answer this question and have created controversy in the last decades. Molecular genetic studies suggest a single burst of emigrating *Homo sapiens* out of Africa, which ultimately conquered the world and preserved their genes for the future (Oppenheimer, 2012; Soares et al., 2012). Fossil human remains, however, suggest a more gradual evolution over time (Benazzi et al., 2011). New evidence usually renews the discussion, e.g., newly discovered remains were dated earlier than any previous finds (Hublin et al., 2017).

Growing populations of anatomically modern humans (AMH) are partly responsible for past ecosystem changes and influenced our cultural and genetic evolution. Large pools of interacting individuals create and maintain adaptive skills and phonological variation. The archaeological record allows to estimate the human population range and size dynamics, but it does not offer quantitative population size data. Only favorable circumstances such as exceptional observation intensities and preservation conditions result in an estimate. For that, the calculated human group or settlement sizes are generalized through upscaling for larger areas. The result is used as a lower limit, based on the assumption that it is impossible to find all remains of humans at all possible sites for the given time frame (Zimmermann et al., 2009). Genetic reconstruction methods are still debated and deliver deviant temporal resolution datasets, but they allow a different approach on the topic. The number of available DNA samples from different times and locations increased in the recent past, which pushes the quality and quantity of DNA-based studies. Combinations of DNA gathered from excavations and modern populations seem to be promising to pin down the waves and timing of dispersing modern humans (López et al., 2015; Tallavaara et al., 2015).

The contradicting findings, results, and hypotheses are difficult to quantify due to the lack of sufficient findings of the hunter-gatherer populations and their dispersal for several timeframes (Richter et al., 2012). Human population and expansion models can be used to fill this gap and simulate the behavior of early humans based on initial population densities. Usually, these models are used for smaller time frames, for example, the retreat and reoccupation of

Europe during the Last Glacial Maximum, but work has been done to simulate more extensive time frames (Maier et al., 2016; Timmermann and Friedrich, 2016). Both examples use primarily climate data to either estimate the human population density or force the dispersal over the globe. The basis for most dispersal models is a reaction-diffusion equation, often referenced as the Fisher equation, Fisher-Skellam model, or Fisher-KPP (Steele, 2009).

Several other studies deal with the concept of human habitats or the mobility of hunter-gatherers in specific regions or continental scale, utilizing different approaches and datasets (Maier et al., 2016; Timmermann and Friedrich, 2016; Burke et al., 2017; Benito et al., 2017; Tallavaara et al., 2015; Martino et al., 2007; Steele et al., 1998). A more in-depth overview of the topic is given in Chapter 2.

This study aims to extend the toolset of climate-based estimations of habitability for humans and their dispersal over time, focusing on hunter-gatherer cultures. The first tool is the Human Existence Potential (HEP), an estimation of the habitability based on climate simulations, archaeological site distribution, and additional environmental data. It can be used on its own for several applications to study areas of human habitation for specific time frames, e.g., climate adaption and extrapolation to areas that are missing archaeological evidence. In addition, the HEP is used for the second tool, the Human Dispersal Model (HDM), to simulate the dispersal over time dynamically. This model extends the often used reaction-diffusion equation with a drift term forced by the HEP, directing the human dispersal to areas of high habitability.

These tools are part of the Human Modeling Framework (HMF) developed during the last few years within the workgroup of Prof. Dr. Yaping Shao, with contributions by Dr. Masoud Rostami, Konstantin Klein, and myself. This thesis will focus on the Human Dispersal Model, distinguishing my contributions from my colleagues. An overview of required elements of the framework is given with clear attributions of their primary author.

This thesis is divided into 8 Chapters. Chapter 2 will cover the relevant research and the motivation of this work. Chapter 3 presents the different parts of the HMF in detail. In Chapter 4, the dispersal model is tested with idealized situations to investigate if the numerical formulation is consistent with the analytical formulation. Chapter 5 includes case studies with time-variant HEPs that check the model's performance in semi-realistic cases. The largest experiment is carried out in Chapter 6, the dispersal of modern humans into Europe. The thesis is closed with an overarching discussion in Chapter 7 and a conclusion paired with an outlook in Chapter 8.

1.1 Collaborative Research Centre 806 - Our Way to Europe

This research is part of the Collaborative Research Centre 806 (CRC 806) - "Our Way to Europe" funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG).

The CRC 806 started in 2009, focusing on the cultural-environmental context of the spread of anatomically modern humans from their origin in Africa to one of their destinations (Central Europe). Studies include secondary expansions and retreats within the last 200,000 years due to climate, environmental, or cultural changes.

The project unites experts from several natural and sociological sciences disciplines, such as archaeologists, anthropologists, and geoscientists. In its third and last phase, the focus lies on interdisciplinary research and collaboration between the different groups rather than contributing pieces to individual works. This approach inspired the creation of the HMF, combining meteorologists' physical and numerical expertise with the knowledge of archaeologists and anthropologists. Every part of the framework was created with current knowledge of modern human behavior in mind, trying to compromise between physical formulations and conceptual models (CRC806, 2021).

Chapter 2

Literature Review

2.1 Dispersal of Modern Humans from Africa to Europe

For a long time, archaeological records and fossils together with stratigraphy and anatomy studies served as the only source of reliable information about the origin of modern humans. Radiometric dating allowed for approximations of the age of artifacts, allowing the construction of the first timelines. Since the successful sequencing of DNA, several genetic methods arose, which, combined with the former excavations, confirmed the African origin of modern humans. However, this advancement opened up several questions about the details of our ancestry (López et al., 2015): Where in Africa did the modern human evolve? When did they leave Africa and dispersed around the globe? Was this achieved by a single outburst, or were there several waves of dispersal together with backtracking populations? Which path did they take? Answers to these questions are highly debated, and new evidence, such as remains that were dated earlier than before, shift the overall favor for an answer within the scientific community (Hublin et al., 2017).

Several models for the dispersal of modern humans exist. The most present hypotheses in the discussion are the candelabra model and the out of Africa model (Templeton, 2007). The former suggests that *Homo erectus* dispersed out of Africa roughly 1.9 million years ago and spread to Asia and Europe. Each population evolved independently into modern humans. The ladder includes an additional wave of dispersal out of Africa, now with the genetically modern human, which gradually replaced other situated homini in Asia and Europe. Genetic studies pushed the "out of Africa" hypotheses to be the most accepted, but more recent studies suggest more complex models, including genetic variations and mixing with situated populations instead of complete replacements (Harding and McVean, 2004; Watkins et al., 2003; Eriksson et al., 2012).

Regarding the locality, extensive support is provided for an East African origin by modern human fossils in Ethiopia, mainly Omo I dated to 190 to 200 kya and Herto dated to 160 to 154 kya (McDougall et al., 2005; McCarthy and Lucas, 2014). A southern origin is supported by

genetic studies rather than archaeological findings (Henn et al., 2011). However, it proved difficult to separate southern and eastern populations due to common ancestry in central Africa. More clear differentiation between Northern and Southern populations can be observed (Pickrell et al., 2012; Kim et al., 2014). The possibility of a Northern origin was seen as less likely, up until more fossil finds came up (Smith et al., 2007; Hublin et al., 2017) and a genetic study traced a Y chromosome back to central and Northwestern Africa (Cruciani et al., 2011). An important note to make is that many genetic applications rely on present-day samples, which results in poor representativeness due to the mobility of modern humans. The ancestors of a now situated population are likely to originate from other localities than the present one, especially in Northern Africa (Fadhlaoui-Zid et al., 2013).

The exact time frame of the out of Africa dispersal is controversial as well. Overall, the dispersal is supposed to happen between 130 - 40 kya ago. Archaeological studies suggest an earlier dispersal, focusing on 120 - 80 kya (Rampino and Self, 1992; Liu et al., 2015; Bae et al., 2014). The calculations of genetic methods favor a later wave of expansion, roughly 60 kya to 40 kya (Soares et al., 2012; Shi et al., 2010). This discrepancy between the archaeological record and the genetic reconstructions is likely based on the use of modern DNA for most models. There are, however, genetic studies that favor the early migration (Li and Durbin, 2011).

Several models and studies investigate whether a single outburst out of Africa is likely or multiple waves of dispersal need to be assumed. Groucutt et al. (2015) gives an overview of them. For the single outburst, a bottleneck is proposed that reduced the genetic diversity of the migrating group followed by a rapid expansion in population size, allowing new groups to disperse further (Deshpande et al., 2009). Both early and late dispersal models exist for the single outburst, differing in their dispersal route as well (Klein, 1990; McDermott et al., 1993). There are, however, doubts regarding a single outburst due to conflicting archaeological and genetic results (Gutenkunst et al., 2009; Bowler et al., 1970). If more than one dispersal wave is considered, the two major time frames of dispersal can be woven together. An initial migration is proposed between 100 and 50 kya followed by a second wave till 40 kya (Mirazón Lahr and Foley, 1998). Genetic models back up the multiple wave assumption, often finding more than two possible waves (McEvoy et al., 2011; Reyes-Centeno et al., 2015). In addition to multiple migration waves, there is evidence for returning populations to Africa through high portions of non-African ancestry in modern populations. Their timing and number are not clear because both archaeological and genetic studies lack agreement.

Lastly, the exact path of dispersal is debated as well. Possible candidates for leaving Africa are the Northern Route through Egypt and Sinai and the Southern Route through Etopia, crossing the Bab el Mandeb strait to the Arabian Peninsula. The northern route features evidence through excavations (Hovers, 2009) and genetic studies (Pagani et al., 2015). There is a significant gap in time to the next dated remains (roughly 40 kya apart), creating the debate if this first expansion failed and led to extinction in this region or if the population there moved on to

other regions. The southern route is mainly proposed by mtDNA studies, allowing a rapid dispersal through Southeast Asia to Oceania (Forster, 2005). There is questionable evidence from the archaeological side, mainly due to limited evidence of maritime activity (Rose, 2010).

Several answers to the questions regarding the dispersal out of Africa of modern humans remain unclear. Some models, such as the out of Africa model and the multiple dispersal waves model, are more accepted and less debated than the region of origin, the exact timing of dispersal waves, and the travel route. Further research might tip the scale to favor one thesis over another.

2.2 Human Habitability Estimations

One obvious way to estimate if a region was habitable for modern humans is to date fossil remains and organize corresponding stone tool fragments into technocomplexes, which are cultural groups sharing techniques in creating their tools and used materials. Linking several excavation sites with corresponding data on a spatial map allows one to identify favored regions for specific times. Adding paleoclimatic data to the mix, either through simulations or reconstructions based on proxies, allows defining preferred environments of specific groups or technocomplexes. This approach depends heavily on the quality of the used artifacts and their investigation, those requiring to filter which sites to use (Schmidt and Zimmermann, 2019).

Estimates of human density based on archaeological data are rarely done. Only favorable circumstances such as exceptional observation intensities and preservation conditions result in an estimate. For that, the calculated human group or settlement sizes are generalized through upscaling for larger areas. The result is used as a lower limit, based on the assumption that it is impossible to find all remains of humans at all possible sites for the given time frame. One method for estimations is the cologne protocol, a series of geometric and mathematical methods such as the creation of Thiessen Polygons and the determination of site density (Maier and Zimmermann, 2017; Zimmermann et al., 2009).

An increasingly used set of tools for estimating the habitability of regions for modern humans are Species Distribution Models (SDMs). These correlate either presence or presence and absence data with environmental variables, called predictors, to derive an equation describing the preferences of a species regarding these variables. This equation can be further used to extrapolate distributions onto regions with low or absent data of presence/absence. Traditionally, SDMs are used for specific species of fauna or flora, but the usage for human species followed suit despite the more complex nature of *homini*. SDMs can utilize a large number of different algorithms, including mathematically based approaches (such as generalized linear or additive modeling, probability density functions) and machine-learning algorithms (e.g., support vector machines, random forest) (Guillera-Arroita et al., 2015; Tallavaara et al., 2015; Maier et al., 2016).

Mathematical algorithms often have the advantage that they allow to deduce how predictors contribute to habitability, making it possible to deduce connections between the occurrence of a species and specific environments. However, caution needs to be applied because the used set of predictors can miss important information, resulting in false conclusions. Machine Learning algorithms can find unknown connections and patterns between different datasets but lack a physical or mathematical description of this link. The following paragraphs give an overview of different studies utilizing some form of SDM.

Banks et al. (2008) use an eco-cultural niche model paired with a machine-learning Genetic Algorithm for Rule-Set Prediction (GARP) to estimate the habitability of Europe during the Last Glacial Maximum (LGM). They combine archaeological and geographic data with representative paleoclimate simulations for that time frame. With their method, they defined a northern boundary of human presence for the Solutrean and Epigravettian technocomplexes. The former population seems to be more adapted to colder and wetter environments than the Epigravettian, allowing them to exploit regions further north.

Benito et al. (2017) investigate Neanderthal distributions during the Last Interglacial optimum - corresponding to the Marine Isotope Stage (MIS) 5e - with generalized linear models fitted to several combinations of predictors. The importance of the different predictors is defined on continental and local scales by using Area under the curve scores and local regression. This method resulted in annual rainfall and winter temperatures as the most important predictors for large scales, followed by topography and summer rainfall on local scales. The Mediterranean coastline is an optimum habitat, opposing the notion of Neanderthals favoring cold conditions.

In the work of Boria et al. (2014), the possible overfitting to spatial presence data is discussed, and a reason is given to filter spatial information before using it for SDMs. Here, the Cowan's shrew tenrec (Microgale cowani) is the investigated species. The used algorithm is Maxent, which performs well compared to other techniques, but is prone to sampling biases, making it a decent candidate for the study. With spatial filtering, the overfitting was drastically reduced, improving the model's performance, especially in regions of decreased density of presence data. Kondo et al. (2018) utilize Maxent as well, but take the results of the SDM a step further. They use the inverse of the occurrence probability as a friction value and derive cumulative costs to disperse from one point to another, called the Least Cost Path (LCP). With the initial Upper Paleolithic population group as input, they derive three main migration routes from the Levant to central Europe, the Russian steppe, and the Altai region.

Burke et al. (2014, 2017) focus on the impact of climate variability on human occupation in Europe during the LGM. The first paper utilizes logistic regression between the archaeological record and the downscaled climate simulation for the Iberian Peninsula. The results confirm

that humans responded to climate variability, mainly on interannual variability of precipitation. The second paper expands the research to include Western and central Europe. It solidifies the importance of small-scale climate variability in accessing suitable habitats, linking regions of high variability with increased mobility of occupants.

Maier et al. (2016) deal with population dynamics during the last glacial maximum in Europe. Datasets are combined from climate modeling, geomorphology, and archaeology. The time frame is set to 25,000 to 20,000 calBP, which allows the unification of climatic time frames and the archaeology era of similar cultures. The used excavation sides are heterogeneous and clusters of sites alternate with empty spaces. The total number of people and their density is estimated, as well as their adaption to different environments. A relationship between the location of sites and the palaeoenvironmental conditions is drawn. Logistic regression is used as well as data of permafrost, forests and loess distributions.

A different approach is taken by Tallavaara et al. (2015), by modeling indecently of archaeological and genetic data. They hypothesize that a correlation exists between climate and hunter-gatherer population density in Europe during the LGM. Ethnographic data on terrestrially adapted mobile hunter-gatherers and their climatic space is used to construct a calibration model that predicts hunter-gatherer presence and population density. The main climatic predictors chosen for this model are: potential evapotranspiration, water balance, and mean temperature of the coldest month, which is a measure to estimate winter conditions. A large set of algorithms were used and later combined to minimize the impact of one specific algorithm's flaws on the results. This set includes:

- Generalized linear modeling (GLM)
- Generalized additive modeling (GAM)
- Support vector machines (SVM)
- Classification tree analysis (CTA)
- Random forest (RF)
- Generalized boosting methods (GBM)

The results show that the simulated range and size of populations mostly agree with the archaeological record. Population numbers were calculated for the time frame of 30 to 13 kya.

Overall, SDMs prove to be a valuable tool to estimate the habitability of humans. There is potential to extend the use of the resulting occurrence probability for additional investigations. However, SDMs have some pitfalls depending on the used data to train the model and the type of algorithm used, possibly resulting in over- or under-fitting. Spatial presence data should be filtered to avoid over-representation of specific regions.

2.3 Human Dispersal Models

Computational modeling of dispersing humans is a relatively minor topic within the already small field of numeric archaeology. Nevertheless, active development and research are going on since computers' increasing availability in the early 1970s. This section will give an overview, following the review articles of Steele (2009) and Lake (2014) and take a look into recent works of Axel Timmermann (Timmermann and Friedrich, 2016; Timmermann, 2020).

The most widespread approach for the simulation of large-scale human dispersal is a reaction-diffusion equation, which combines the spatial spread of humans with the population growth (Steele, 2009). This system was defined independently by R.A. Fisher and Kolmogorov, Petrovsky and Piskuniv, and is therefore often known as KPP or Fisher-KPP. It can be written as

$$\frac{\partial \rho}{\partial t} = f(\rho, r_B, \Phi_{max}) + K\nabla\rho \tag{2.1}$$

with the population density $\rho(x, y, t)$ at the time t and the position (x, y). The first term is the population growth function f, in this case a logistic growth function with the growth rate r_B and an upper limit Φ_{max} called carrying capacity:

$$f(\rho, r_B, \Phi_m ax) = r_B \rho \left(1 - \frac{\rho}{\Phi_{max}}\right) \tag{2.2}$$

The second term of the equation is the spatial diffusion with the diffusion constant K and the second derivative of ρ. Based on this term, the propagation speed v of the wavefront can be approximated: $v = 2\sqrt{K \cdot r_B}$.

This system was used to model the spread of hunter-gatherers in several studies, including the spread of modern humans into Europe (Mellars, 2006), the late glacial recolonization of northern Europe (Fort et al., 2004), and the spread of Clovis spear point technology in North America (Hamilton and Buchanan, 2007). In addition, Campos et al. (2006) used such a model to estimate the spread of American settlers westward from 1790 to 1910.

Some further extensions of the model can be considered and are discussed in Steele (2009). For example, the Alle Effect should be considered: At low population densities, the risk for extinction can be higher, resulting in local negative growth rates. The diffusion is symmetric in the proposed model (Eq. 2.1), but it might be preferable to make it dependent on the population density, accelerating the dispersal towards lower densities. In addition, some model parameters might depend on time and space, reflecting external factors such as orography and resource availability. Such heterogeneity has a strong influence on the model and allows for more realistic simulations.

There is the possibility of extending the model for more than one population by using a copy

of the equation for each population and coupling them with an additional term. As all populations probably compete over the same space and resources, a competition system that influences the reproduction of the other populations suggests itself, which can include coexistence.

For this thesis, the consideration of expanding this model with an advection term is of particular interest. This addition allows for anisotropic movement and the inclusion of preferential directions, which could be gathered from the gradients of resources. However, the advection was rarely considered or used in studies. One example is the work of Davison et al. (2006), which used the advection locally to accelerate the dispersal along rivers and coastlines.

While Steele (2009)'s review focuses on the mathematical approaches and the considerations behind them, Lake (2014) gives a more general overview of the history of archaeological simulations. In principle, three groups of simulations can be observed: various methods, often technically simple and derived from other sciences, dynamical systems, and agent-based modeling. One example of a dynamical system is the diffusion-reaction equation discussed above. Agent-based modeling is less deterministic and resembles systems used in artificial intelligence. These models consist of individuals acting in a given environment, either based on predetermined behavior patterns or letting them react and learn from interacting and changing the environments.

Lake (2014) divides the history of simulations into four parts: The Pioneer Phase (1960-1980s), the Hiatus (1980s), the Renaissance (1990s), and the Expansion (2000+). After an initial excitement regarding the possibilities of computer models, the limitations of the computational power halted the research for some time. The models were either too simple to be of use or missed most of their input data. With the further increasing availability and computation power in the 1990s, computation models became more interesting, but the willingness to deal with complex mathematical formulations was low. This results in the fairly small but diverse community of scientists working in this topic, including archaeologists, mathematicians and geoscientists.

Focusing on the expansion, most of the models deal with human evolution issues, and approximately one-tenth utilize a reaction-diffusion equation. These models have a trend to include environmental realism, human impact on the environment, and heterogenic populations regarding their survival strategies, hinting at some convergence between reaction-diffusion and Agent-based models.

One very recent model is the numeric human dispersal model from Timmermann and Friedrich (2016). The model is a reaction/diffusion model for the human density ρ(x,y,t). The reaction terms (growth G and mortality M) are dependent on three key parameters: net primary productivity N(x,y,t) (controls the availability of carbon-based food sources), desert fraction d(x,y,t), and temperature T(x,y,t). They determine here the mortality of Homo sapiens in the model. A quadratic mortality term is included to avoid exponential growth.

The model is forced by estimates of climate and sea-level changes over the last 125 thousand years using the LOVECLIM climate model data. It is hypothesized that the exodus of Homo sapiens out of Africa and into Eurasia occurred in several orbital-paced migration episodes between 50-120 thousand years ago. The first climate forced HDM experiment accounts for a numerical representation of gradual human adaption to environmental conditions. This scenario features the "early exit" of Homo sapiens, resulting in arrival in Europe, India, Southeast Asia, and southern China in low densities between 100-70 ka. It is consistent with earlier findings but in conflict with different interpretations of genetic and archaeological data. For instance, the early arrival in Europe around 90-80 ka challenges fossil evidence that dates this event around 45 ka. This outcome might be due to the assimilation of these groups by the Homo neanderthalensis in this region.

A second scenario was run to quantify the effect of a late exodus from Africa and the Arabic Peninsula, the "late single southern exit" model. Due to drought conditions and humans' chosen higher temperature sensitivity, the first exodus is delayed up to 84 ka and a second event around 60-47 ka. Instead of several immigration events, single vital events are dominant in this scenario. This simulation conforms with the reconstructed arrival times in the other continents but does not explain the early MIS5 presence near east and southern China. Additional sensitivity experiments with only perpetual climate conditions show substantial differences in the dispersal patterns, which arise through very different climate conditions and sea levels.

The second iteration of the model (Timmermann, 2020) consists of two populations, Homo Neanderthalensis and Homo Sapiens, by coupling two reaction-diffusion equations as discussed in Steele (2009), allowing interbreeding and competition between both populations. This model was used to simulate the interactions of both populations during the last ice age, investigating possible reasons for the extinction of the Neanderthals. The results include that the extinction is only explainable by assuming that the modern human is at least twice as efficient in utilizing resources as the Neanderthals. With faster reproduction rates and mobility, modern humans can adapt more quickly to changing environments, allowing them to reoccupy the landscape faster after cold climate periods.

This model has gained more complexity over time, as more and more environmental forcings are included in the parameters, creating a large set of equations to solve each time step. It allows for more realism in the model but increases the difficulty for interpretations regarding the impact of each calculation. Therefore, several simulations with changed parameters and assumptions are required to estimate their impact.

In summary, numerical modeling of the dispersal of humans or hunter-gatherers is a helpful tool to combine several possible drivers of dispersal. While the basic methodology is established, numerous extensions exist depending on the application, adding further complexity and realism to the models, making them less comprehensible for researchers outside of mathematical fields.

2.4 Motivation

As pointed out in the previous sections, several methods were established to estimate the habitability and dispersal of hunter-gatherers and early farmers. These methods, however, can be further extended to include additional environmental data and processes to increase the realism of these models. In addition, some of these models have become quite complex and difficult to use and interpret, particularly for archaeologists and anthropologists who do not provide an extensive mathematical background. The habitability of humans is difficult to pinpoint, as the required direct measurements (food, water, shelter) have to be derived from environmental variables such as temperature, rainfall, and orography. Furthermore, even this data is uncertain, as it is derived either by reconstructions through proxies or calculated through coarse paleoclimate simulations. The dispersal of humans is often simplified through diffusion processes, lacking discrete travel directions and social-cultural interactions.

This thesis has two precise aims: To extend and improve the established methods by using additional environmental data, including assumptions about dispersal patterns, and making these methods available in further research. For the former, the habitability for humans is estimated and shall serve as the basis for a dispersal model that includes advection processes. The ladder will be achieved by incorporating the methods into a framework, allowing a multitude of calculations to be done with the same set of required input data.

Chapter 3

Methods

This chapter explains the methods used for this thesis and the developed processes for the human existence potential and dispersal model. The first section (3.1) will give an overview of the Human modeling framework, highlighting which parts of the joined research I contributed and the dependencies of the different items. After that, the relevant modules are presented, including the Human Existence Potential (HEP, 3.2), the Accessible HEP (3.3), and the Human Dispersal Model (HDM, 3.4).

3.1 Human Modeling Framework - Overview

The Human Modeling Framework (HMF) results from an interdisciplinary approach to calculate human survivability, mobility, and dispersal, focusing on hunter-gatherer populations during the neolithic. It combines the developed methods and modules that Konstantin Klein and myself designed and programmed, supervised by Prof. Dr. Yaping Shao and Dr. Masoud Rostami. The archaeological and anthropological knowledge was provided within the CRC806, mainly by Prof. Dr. Gerd-Christian Weniger, Prof. Dr. Andreas Zimmermann, and Dr. Isabell Schmidt.

It can be divided into three different parts: The required datasets to utilize the different parts of the framework, the creation of the HEP to estimate suitable regions for human survival, and the further applications that use the HEP as their basis (Figure 3.1). The required datasets include spatial paleoclimate data for at least one point in time, created either by reconstructions based on proxy data or numeric climate simulations. If the ladder includes at least monthly seasonal data, the calculation of Bioclimatic variables is encouraged. In addition, the archaeological site record for the time and technocomplex of interest is required to create the HEP. Additional data such as topography and vegetation can be included in the accessible HEP, but the only requirement for further usage of the HEP in the applications is the use of a water mask.

The second part consists of two potentials. The Human Existence Potential combines the paleoclimate data with the archaeological record to estimate the suitability of the environment

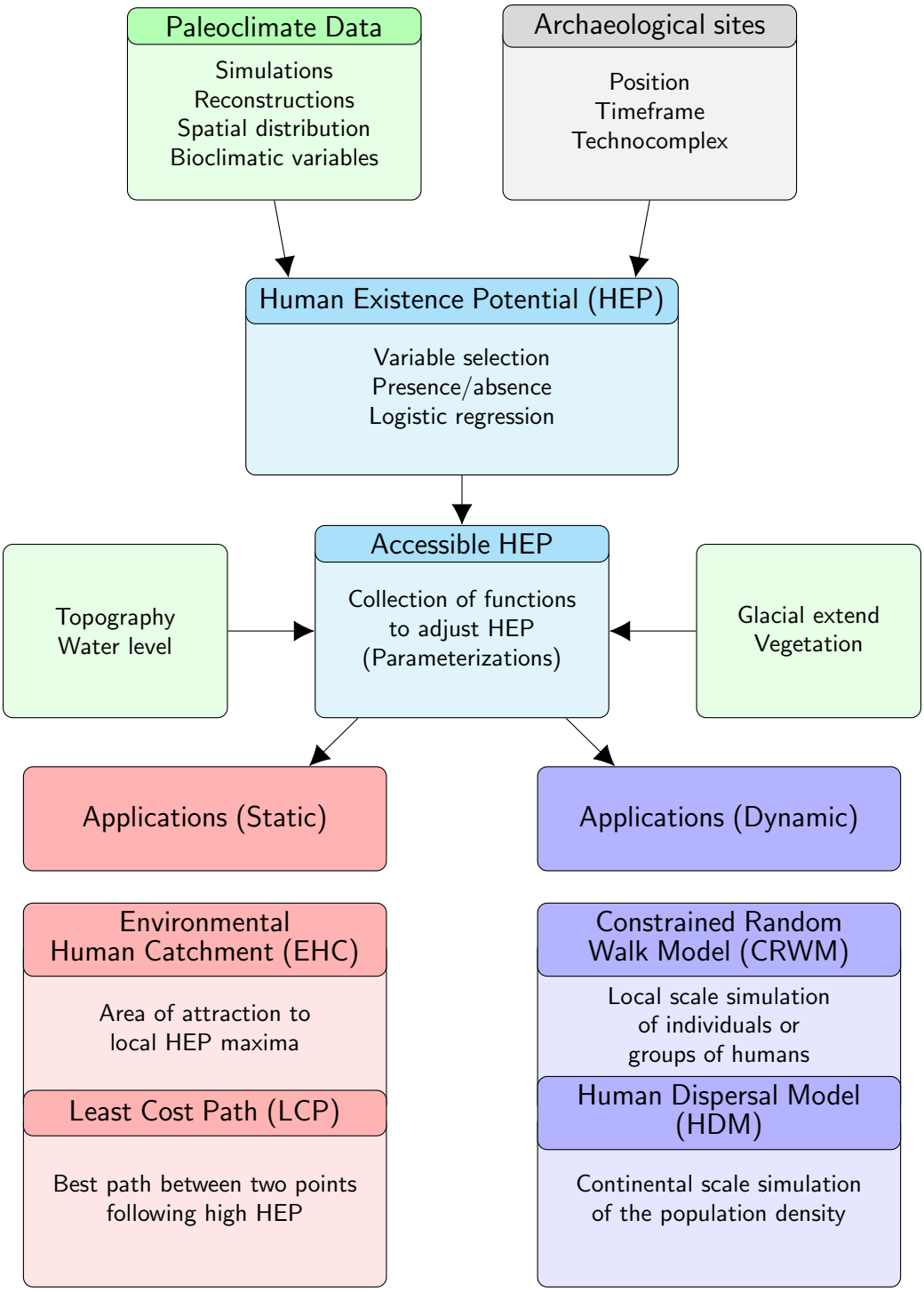

FIGURE 3.1: Overview of the Human Modeling Framework. Green backgrounds indicate required meteorological or environmental datasets. Grey indicates the archaeological data and categorization. Cyan stands for the Human Existence Potentials. The applications are separated into static (red) and dynamic (blue) methods.

for humans for a specific time step or frame inherited by the provided data. It includes selecting relevant variables, dividing the region of interest into fields of presence, absence, or "unknown", and the actual suitability calculations through logistic regression. The Accessible HEP builds on the results of the HEP and includes other factors that alter the estimate, either by masking out regions that cannot sustain humans or by altering the values with given mathematical expressions, called parameterizations. While K. Klein and I worked out both concepts, we split the actual development between us, with my colleague focusing on the HEP and myself on the Accessible HEP. Both will be laid out in detail in the following sections (3.2 and 3.3).

The further applications that the HEP enables are divided into two categories: static and dynamic. The former uses a single HEP representing a specific time step or frame. The Environmental Human Catchment (EHC) allows identifying core areas of high suitability for human survival and estimating the spatial range of influence, resembling a catchment area for hunter-gatherers. The Least Cost Path (LCP) utilizes the reverse of the potential and calculates the travel cost between two points of interest, determining optimal paths between the points. These applications are not further discussed in this thesis, but details can be found in the publication of Klein et al. (2020). The dynamic applications are mobility and dispersal models that simulate the spread and population change based on at least one HEP. It is possible to create time-dependent simulations with them by linking several HEPs together, representing the change of the environment or the adaption of humans over time. The Constrained Random Walk Model (CRWM) is developed by K. Klein and utilizes a Lagrangian view to simulate individuals' movement in small spatial scales. It is part of his doctoral thesis and is therefore not further discussed here. The Human Dispersal Model (HDM) is developed by myself and focuses on the simulation of the population density instead, using an Eulerian view and is more suitable for larger spatial scales. This model's specifications, tests, and first applications are the core of this thesis and are presented in detail in the following chapters.

3.2 Human Existence Potential

Humans are a highly adaptive species, capable of surviving in different environmental conditions as long as food and water are available. Both of these resources can be estimated through temperature and precipitation. By adapting to a specific environment, humans might not occupy all potential regions and rather stay within their adapted conditions. This culturally-induced selection of regions allows the use of spatial archaeological data in the training of the HEP.

The HEP is the result of logistic regression with second-degree polynomials, using a presence and absence record through spatial archaeological data and suitable climate predictors for the experiment (Figure 3.2). It is, therefore, a function of the predictors showing the least to most

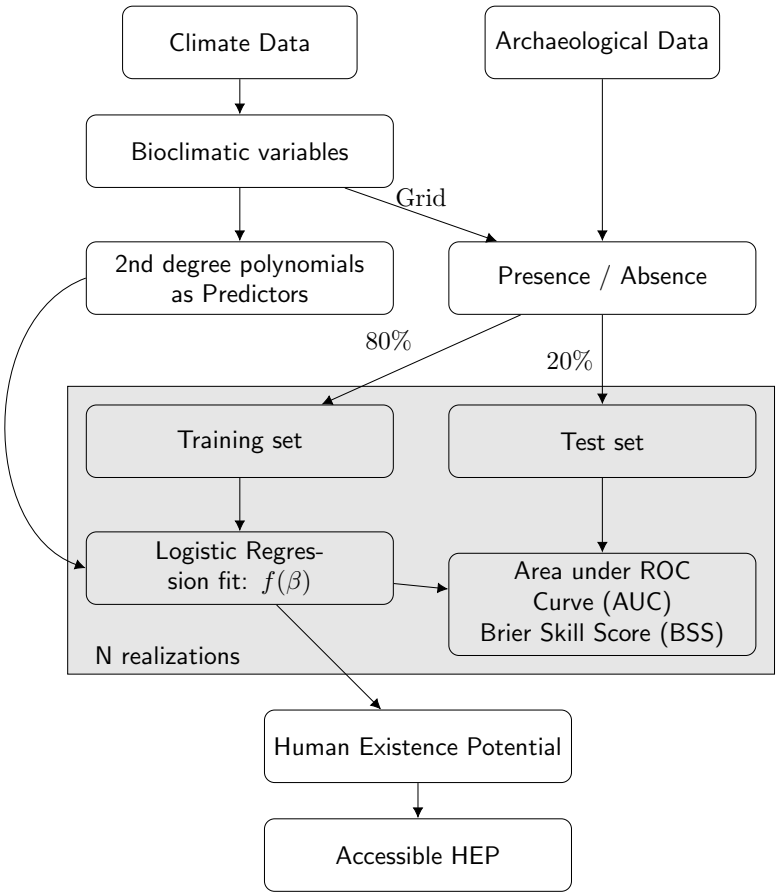

FIGURE 3.2: Overview of the required steps to calculate the HEP. Simplified version of Figure 2 in Klein et al. (2020).

favorable conditions for human survival with scores from zero to one. The predictors should include at least the average temperature and precipitation and can be extended with essential variables for the application - for example, seasonal variations of temperature and precipitation or topography through elevation, roughness, and slope. For the climate predictors, up to 19 different bioclimatic variables can be derived from a dataset of at least 30 years containing monthly mean temperature, daily maximum temperature, daily minimum temperature, and monthly precipitation. Several of these bioclimatic variables focus on the seasonality within temperature and precipitation. An overview of each variable and their calculation can be found in the Appendix 8.

The logistic regression is susceptible to collinearity between the predictors. A selection of the variables should be considered to minimalize errors and reduce the overestimation of variance by using too many variables. For this, the Variance Inflation Factor (VIF) (Dormann et al., 2013) can be used to estimate the inflation of the model. The VIF score is calculated for each variable and should be below 10 for each of them. It can be achieved by taking out the highest-ranking variable and recalculating the scores until the conditions are met. There might be some

variables of high interest or impact in the application desired to be kept in the model. These can be skipped in the selection, and the second-highest score is removed, but their VIF score should still be calculated to meet the condition.

If the archaeological record is quite small, further selection might be required to avoid the under-parameterization of the model. For this, the correlations between the remaining variables can be calculated within a distance matrix. These distances can be repeatedly clustered together by their correlation until only one group of variables is left. The result can be visualized in a dendrogram, allowing to see the distances (and correlation) between clusters and choosing variables with the highest distance (lowest correlation) accordingly. An example for this is given in Klein et al. (2020), utilizing the fastcluster package of Müllner (2013).

The now selected variables b_i are standardized to reduce numerical errors and stabilize the calculations of the regression. For this, the average ($\langle b_i \rangle$) and standard deviation σ_{b_i} for each variable is used:

$$b_i^s = \frac{b_i - \langle b_i \rangle}{\sigma_{b_i}} \tag{3.1}$$

Now, second-order polynomials of the standardized variables are calculated and then used as predictors $\vec{p}$ for the logistic regression (Equation 3.2). This includes polynomials which are pairs of two variables. For a set of three variables a, b, c, the second-order polynomials used as predictors are $\vec{p} = \langle a, b, c, ab, ac, bc, a^2, b^2, c^2 \rangle$.

In addition to the climate-derived predictors, a human presence and absence record is required for the logistic regression. We achieve this by assuming human presence within a 50 km radius around each archaeological excavation site. The radius can be changed accordingly to get a sufficiently large number of presence points, ensuring the convergence of the model. Using the grid on which the predictors are located, we can interpret each grid point within a site's range as a presence and each point outside of it as an absence. Regions that are not inhabitable, such as large water bodies or ice-covered regions, should be excluded from the absence record. The absence record is often noted as pseudo-absence, as we cannot be sure if regions without an archaeological record were indeed avoided by humans or are not sufficiently examined. The total number of presence points is noted as N_{pre}, the total number of absence points as N_{abs}.

For a later evaluation of the trained model one thousand iterations are done with randomized sets of both presence and absence, using 80% of each set for one iteration and the remaining 20 % as a validation set. In each run, the coefficients $\vec{\beta}$ and the intercept β_0 are determined by the logistic regression function of the training data:

$$y_i^{train} = \left\{ 1 + \exp\left[-(\beta_0 + \vec{\beta} \cdot \vec{p}_i^{\,train}) \right] \right\}^{-1} \tag{3.2}$$

with $y_i^{train} = 1$ for a presence grid point and $y_i^{train} = 0$ for an absence grid point, and for

$i = 1, 2, ..., (N_{pre} + N_{abs}) \cdot 0.8$. This set of equations determines the coefficients β and determines the model $f(\vec{p})$. The presence and absence points are weighted by their number during the calculations. In addition, a LASSO (Least Absolute Selection and Shrinkage Operator) regularization is included during the training processes to omit irrelevant terms of the polynomials.

The thousand iterations result in thousand models $f_j(\vec{p})$, which can be applied on the whole study area using the (normalized) climate variables, and the average of the thousand iterations give the HEP:

$$\text{HEP} = \frac{1}{1000} \sum_{j=1}^{1000} f_j(\vec{p}) = \frac{1}{1000} \sum_{j=1}^{1000} \left\{ 1 + \exp\left[-(\beta_{0,j} + \vec{\beta}_j \cdot \vec{p}) \right] \right\}^{-1} \tag{3.3}$$

Several different skill scores can be used to estimate the uncertainty of the model, utilizing the test set left out during the calculations. Two examples are the Area Under a receiver operation characteristics Curve (AUC) (Hanley and McNeil, 1982) and the Brier Skill Score (Hernández-Orallo et al., 2011). Calculating these scores for each iteration allows us to take both scores' average and standard deviation for all thousand runs. In addition, the variability of the HEP for all runs gives a measure of the robustness of the model.

3.3 Accessible Environmental Human Existence Potential

The HEP covers the main climate drivers to estimate the habitability through temperature and precipitation variables, but other important factors such as orography and resource availability may not be used as primary predictors. They still influence the HEP, however, and should be taken into account. The Accessible HEP (AcHEP) uses the HEP and alters it with several parameterizations, which change the suitability score of the HEP:

$$\Phi_{Acc} = \Phi_E \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3 \cdot ... \tag{3.4}$$

In the following subsections, each parameterization is presented and discussed. The additional data sets might feature different grids than the one used for the HEP calculations. In these cases, the additional data is interpolated on the HEP grid, utilizing linear (for floats, e.g., real numbers) or nearest neighbor algorithms (for boolean, e.g., masks of water or ice).

3.3.1 Ice and Glaciers

The permanent occurrence of sea ice and glaciers act as natural borders for hunters and gatherers. Without vegetation and scarce to non-existent wildlife to sustain humans, they have a

suitability score of 0 regardless of other implications by the bio-climatic variables. They might be used as bridges to access other regions but will not provide any food themselves, limiting the distance that may be crossed through them. If the used climate data can provide information about the extend of ice, it can be used in this part.

$$\sigma_{ice} = \begin{cases} 0, & \text{if the environment is dominated by ice} \\ 1, & \text{otherwise} \end{cases} \tag{3.5}$$

If a fraction of ice cover is used for the input instead of a mask, then the result in some grid points may lie between zero and one, indicating that only a fraction of the HEP may be used.

3.3.2 Water bodies

Larger bodies of water such as oceans and lakes may be favorable for hunter and gatherers to have nearby, but they act as borders for traveling humans. Therefore, they are accounted for by setting the score to 0 at their positions. This water mask is essential for the dispersal model because it determines which grid points are not included in the calculations. It prohibits the existence of any humans on grid points that were masked as water bodies, which differs from ice and glaciers that set the potential to zero (or close to it, depending on the application) but can technically be passed over. If reconstructions of the coastline exist for the time of interest, they can be used here.

$$\sigma_{water} = \begin{cases} 0, & \text{if the position is in a body of water} \\ 1, & \text{otherwise} \end{cases} \tag{3.6}$$

Other than the fractional ice cover, only a clear mask of 'water' and 'land' cells are allowed here.

3.3.3 Forest and dense Vegetation

If the present vegetation gets too dense, it becomes more difficult to pass through while gathering, hunting, or traveling. Humans can still survive in rain forests, but the number that can be sustained is reduced compared to more mobile groups on patches of grassland and more occasional forests. The impact of this factor is not as strong as the previous 'all or nothing' parameterizations. It requires information about the past vegetation, which can be reconstructed based on climate simulations or proxy data, in the unit of surface fraction (km^2/km^2). Currently, the potential is not altered for a forest fraction (f_f) below 0.3. For larger patches of

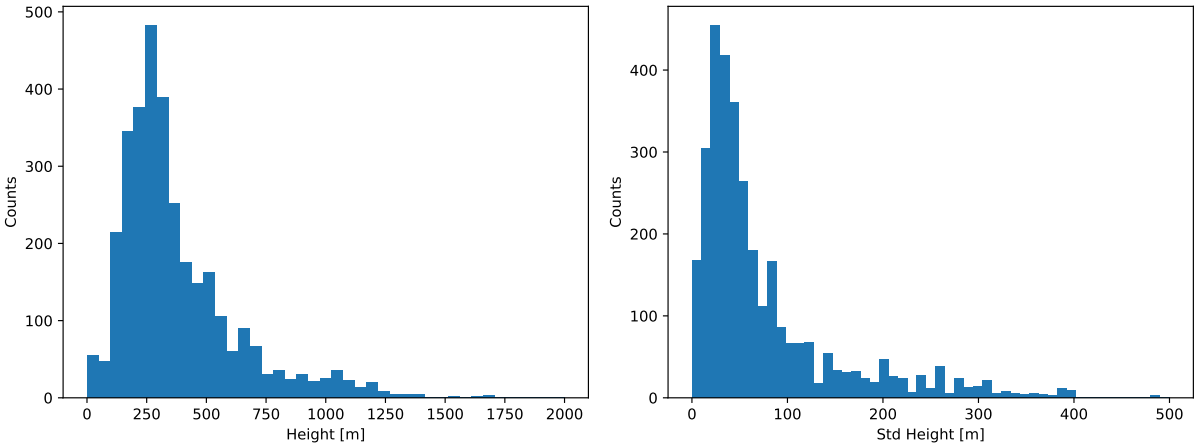

FIGURE 3.3: Histograms of sites accumulated within their 50m bin of elevation (left) and 10m bin of the standard deviation of elevation (right).

forests, the potential is lowered with a linear function starting at 0.3 with and ending at a full cover of forest:

$$\sigma_{for} = \begin{cases} 1.0, & f_f < 0.5 \\ 1.0 - (f_f - 0.5) \cdot 0.4, & 0.5 \leq f_f < 1.0 \\ 0.8, & f_f = 1.0 \end{cases} \tag{3.7}$$

3.3.4 Orography

One of the most decisive limiting factors is orography. Higher altitudes and complex terrain are usually less favorable than flat terrain, partly through more considerable variations in climatic variables and partly due to long walking distances. In a more complex terrain with mountains, valleys, and cliffs, detours must be found to reach a destination, making hunting and gathering more time-consuming than on flat land. The overall accessibility of resources is reduced in these regions. However, some complexity may be more favorable than flat terrain for hunter and gatherers to find shelter and raw materials more frequently. Histograms (Figure 3.3) of sites within Europe of several technocomplexes (or time frames) were created, based on the elevation of their location and the standard deviation of the elevation around each grid point. These were used to define the borders of preferred conditions, as well as the shape of decline. 3256 sites were used from the Last Glacial Maximum, Aurignacian, Magdalenian, Hamburgian, Creswellian, Epigravettian, and Gravettian technocomplexes. The ordinate shows the counts of sites that are within the respective bin of 50m.

The distributions show that elevations (h) of 200 m to 300m are the most common, while sites with higher altitudes decline exponentially. Combined with the assumption that coastlines and

low elevations were preferred by hunter-gatherer migrations (Kelly, 2010), we set the interval for the decline in the potential to 350 m and 2000 m and get the following linear function:

$$\sigma_{ele} = \begin{cases} 1.0, & \mathrm{h} < 350\,\mathrm{m} \\ 1.0 - (\mathrm{h} - 350) \cdot 0.1212 \cdot 10^{-3}, & 350\,\mathrm{m} \leq \mathrm{h} < 2000\,\mathrm{m} \\ 0.8, & \mathrm{h} \geq 2000\,\mathrm{m} \end{cases} \tag{3.8}$$

For the standard deviation of elevation σ_h the interval is chosen to be 50 m and 400 m, resulting in the following equation:

$$\sigma_{elestd} = \begin{cases} 1.0, & \sigma_h < 50\,\mathrm{m} \\ 1.0 - (\sigma_h - 50) \cdot 0.5714 \cdot 10^{-3}, & 50\,\mathrm{m} \leq \sigma_h < 400\,\mathrm{m} \\ 0.8, & \sigma_h \geq 400\,\mathrm{m} \end{cases} \tag{3.9}$$

3.3.5 Additional factors

Additional limiting factors can be added if required or feasible. One example is the technological advance of modern humans, which allows them to increase their suitability under certain conditions or cross water bodies to a certain distance through seafaring. These additions can either be multiplied like the ones before or change specific values (e.g., water close to coasts) and set them to a fixed value. Currently, these functions are not used in the calculations but allow the user of the model to include his additions for their application.

3.4 Human Dispersal Model

One of my collaborative research's primary goals is to create a new dispersal model that takes the Human Existence Potential as the main driver for an advection-diffusion equation. Hunter-gatherers are highly mobile as they respond to variations in the distribution of food and raw materials. Current known dispersal patterns center around the local resource availability because the social and ecological drivers for movement are still poorly understood. Therefore, groups of hunter-gatherers are forced to move when they cannot sustain themselves in a region. This point is typically reached before all available resources are utilized in a region, allowing the regeneration of the environment (Kelly, 2010).

The dispersal of hunter-gatherers is usually observed for individuals or groups or modeled as frontiers of expansion. This research aims at creating a dispersal model that focuses on human densities rather than small isolated groups of humans. This approach is closer to the view of expanding frontiers of different human species, e.g., the expansion and succession of the homo

sapiens over the globe. The main features that the HDM should include are as follows: directed dispersal into regions with higher chances of survival; indirect dispersal into neighboring regions; and reacting to the overall population numbers in a region by either moving together, splitting up, or controlling the population growth.

3.4.1 Available Human Existence Potential

The AcHEP $\Phi_{acc}(x, y)$ gives us information about regions with preferable or unsuitable conditions for the survival of hunter-gatherers. This derived environmental score does not account for the amount of acquirable food and, therefore, the density of humans that can be sustained in the area. It also ignores how much of these resources are already occupied by humans. For that, the Available Human Existence Potential $\Phi_{av}(x, y)$ (AvHEP, only Φ in the following equations) is formulated. First, vegetation data is used, if available, to estimate the carrying capacity of humans for every region. To do this, the fractional vegetation types μ_i are listed against typical carrying capacities ν_i, see Table 3.1. This conversion is adapted from Martino et al. (2007) and modified for different environment types. For every region, the cover fraction

TABLE 3.1: Vegetation type and carrying capacity conversions, adapted from Martino et al. (2007).

Landuse/vegetation type	Carrying capacity in Humans/100km^2
evergreen needleleaf forest	8
evergreen broadleaf forest	16
deciduous needleleaf forest	8
deciduous broadleaf forest	13
mixed forests	13
closed shrublands	12
open shrublands	12
woody savannas	18
savannas	0 (included in woody savannas)
grasslands	14
permanent wetlands	0 (none in this dataset)

of one vegetation type is multiplied with their corresponding carrying capacity. The sum of these results give the regional carrying capacity:

$$\Phi_{max}(x, y) = \sum_{i=1}^{N} \mu_i(x, y) \cdot \nu_i \tag{3.10}$$

This capacity is the upper limit of humans that can be sustained with renewable resources in mind. It serves as the upper limit in the dispersal model as well. Other approaches can be used if we do not have sufficient vegetation data for the time and space of interest. For example, estimating the carrying capacity as equal in all regions (Eq.3.11); changing gradually

based on the latitude (Eq. 3.12); multiplying a fixed capacity value with the HEP score to gain a surface pattern (Eq. 3.13); or regionally set by reconstructions based on proxy data (Eq. 3.14). These approaches should be integrated during the preprocessing for the human dispersal model.

$$\Phi_{max} = c \tag{3.11}$$

$$\Phi_{max} = c \cdot f(y),\ 0 \leq f(y) \leq 1 \tag{3.12}$$

$$\Phi_{max} = c \cdot \Phi_{acc} \tag{3.13}$$

$$\Phi_{max} = g(x, y) \tag{3.14}$$

Humans, which are already in the system and use up resources, diminish the number of available resources, lowering the overall potential left at their location. However, complete absence of possible mates lowers the effective potential drastically as well. To evaluate the available HEP based on the carrying capacity and the population density present, we assume a Weibull distribution shaped connection between the density and the AvHEP:

$$\Phi = S \cdot \left(\frac{\alpha}{\beta}\right)\left(\frac{\theta}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{\theta}{\beta}\right)^{\alpha}\right) \tag{3.15}$$

with the shape parameter $\alpha = 2$, the scale parameter $\beta = 0.4$ and the dimensionless $\theta(x, y)$ derived with the density ρ by using a minimum and maximum possible density:

$$\theta = \frac{\rho - \rho_{min}}{\Phi_{max} - \rho_{min}} \tag{3.16}$$

ρ_{min} defaults to zero, allowing regions without any human occupation, and Φ_{max} defaults to the maximum feasible human density, e.g. the carrying capacity derived above.

The conversion parameter $S(x, y)$ is required to convert the dimensionless Weibull distribution into a density. It can be derived by finding the position θ_{max} for which the Weibull distribution reaches its maximum:

$$\theta_{max} = \left(\frac{\alpha - 1}{\alpha} \cdot \beta^{\alpha}\right)^{1/\alpha} \tag{3.17}$$

The scaling parameter S is then obtained by dividing the carrying capacity by the maximum of the Weibull distribution, using the location θ_{max} derived above. If the carrying capacity changes in an application over time, this scaling needs to be recalculated at every time step that the distribution changes.

$$S = \Phi_{max} \div \left(\left(\frac{\alpha}{\beta}\right)\left(\frac{\theta_{max}}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{\theta_{max}}{\beta}\right)^{\alpha}\right)\right) \tag{3.18}$$

The result of Equation 3.15 is the AvHEP, which functions as the main driver of the advection in

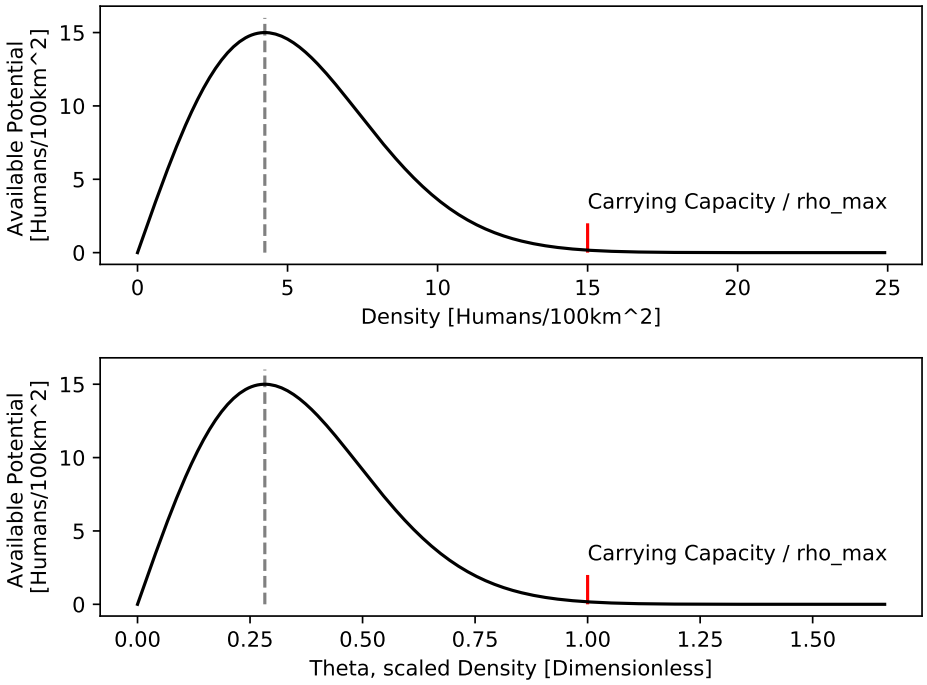

FIGURE 3.4: Weibull distribution with $\alpha = 2$ that is used for the Available HEP. The upper plot uses the density of humans, the lower plot use the density scaled by the carrying capacity.

the model. This approach assumes that hunter-gatherers prefer landscapes rich in accessible resources and are already occupied by enough humans to serve as mates. However, if the region gets crowded enough to enable competition, hunter-gatherers will control their population by migrating to emptier regions or lowering their reproduction.

The shape parameter $\alpha > 0$ allows to change the Weibull distribution between a monotonically decreasing function ($\alpha < 1$), an exponential function ($\alpha = 1$), and a distribution similar to the Rayleigh distribution ($\alpha > 1$) with diminishing skew for higher values. With $2.0 < \alpha < 2.5$ we get the desired shape: A strong rising potential with increasing population, up to an optimal size with a slighter decrease to zero for higher population numbers. The scale parameter $\beta > 0$ depicts the spread of the Weibull distribution. Lower values emphasize the turning point and increase the slope of the decay; higher values flatten the turning point and spread out the decay. With $\beta = 0.4$ the optimal potential is achieved at roughly below one third (0.28284) of the carrying capacity for $\alpha = 2$ (Figure 3.4). This selection of parameters conforms to the assumption that humans try to preserve resources by keeping the population density below the carrying capacity. The slow decay reflects that higher populations are still more desirable than lower ones. Both parameters can be adjusted if other assumptions are required. Changing α to 2.5 gives the optimal potential at roughly one-third (0.32608) of the carrying capacity, creating a distribution more similar to a normal distribution with a more substantial decay (Figure 3.5). This distribution makes medium-sized population ranges more desirable than larger ones.

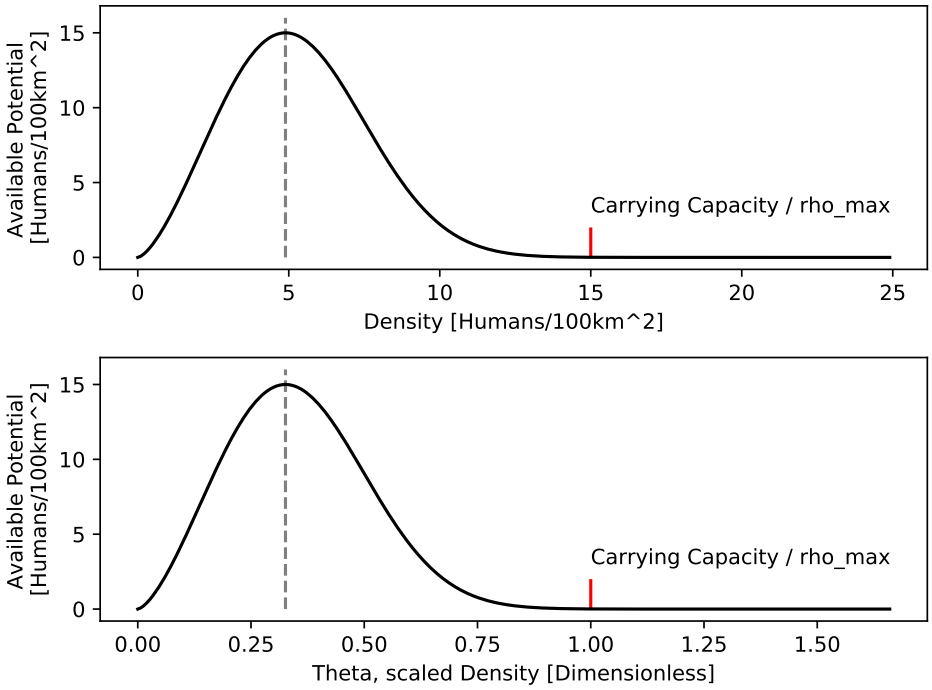

FIGURE 3.5: Weibull distribution with $\alpha = 2.5$ that is used for the Available HEP. The upper plot uses the density of humans, the lower plot use the density scaled by the carrying capacity.

Figure 3.6 shows the distribution of the Weibull function for different carrying capacities ranging from 1 to 15 Humans/100km^2, using the shape $\alpha = 2.0$ and $\beta = 0.4$ as an example. We can see that the preferred or optimal density shifts to higher values with increasing capacity and the slope $m \approx 3.54$. This shift impacts the model's behavior, as areas of lower potential are less attractive and reach their highest available potential with lower population sizes. It also induces that, given a choice, humans are drawn stronger to high HEP regions for a larger range of population sizes compared to low HEP regions, which are favorable only for a small specific range of density.

3.4.2 Dispersal Model

The actual Human Dispersal Model (Figure 3.7) is based on the AcHEP and the AvHEP derived from it. The governing equation is an advection-diffusion-reaction equation with the birth-death model as sources and sinks. In vector form it can be written as follows:

$$\frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot (\vec{v}\rho) + \vec{\nabla}(K\vec{\nabla}\rho) + B - D \tag{3.19}$$

with the human density $\rho(x, y)$, time t, movement speed and direction $\vec{v}(x, y)$, diffusion coefficient $K(x, y)$, and the sources $B(x, y)$ and sinks $D(x, y)$. This equation describes the change of human density over time by accounting for directed (advection) and undirected (diffusion)

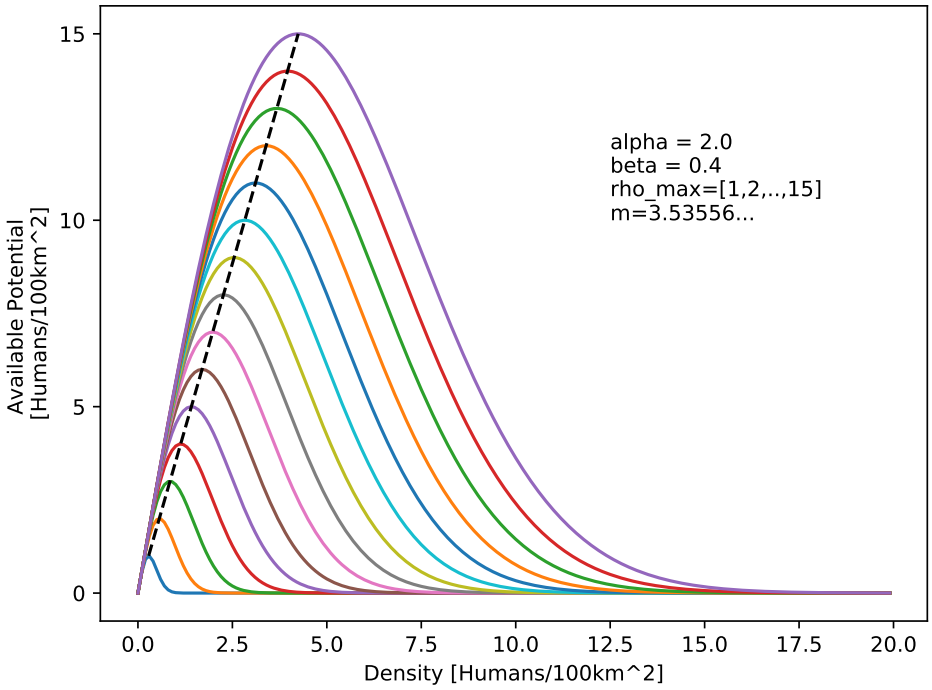

FIGURE 3.6: Weibull distributions (colored) with different capacities Φ_{max}, ranging from 1 to 15. The dashed line indicates the shift of the maximum available potential. Shape parameters and slope of the dashed line are listed within the plot.

movement and the reproduction or extinction of humans (sources/sinks). This equation is derived from the conservation equation for human density and is similar to the description of mass conservation in fluid dynamics:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = B - D \tag{3.20}$$

Therefore, this model's calculations and behavior of human density over time are similar to fluids in corresponding models. Humans live on the surface of earth, hence it is sufficient to consider Equation 3.19 in two dimensions. Due to the curvature of the earth' surface, equidistant or equal-area 2D-grid spacing is only possible in regional scales. To account for distortions in the grid in larger scales, the human density ρ, which depends on the surface area of the grid cell, is converted to an absolute number of Humans. The local area A that each grid point represents is multiplied with the density to gain the absolute number of humans per grid cell $\lambda(x, y) = \rho(x, y) \cdot A(x, y)$. This reduces the equation to:

$$\frac{\partial \lambda}{\partial t} = -\frac{\partial (v_x \lambda)}{\partial x} - \frac{\partial (v_y \lambda)}{\partial y} + \frac{\partial (K \cdot \frac{\partial \lambda}{\partial x})}{\partial x} + \frac{\partial (K \cdot \frac{\partial \lambda}{\partial y})}{\partial y} + B - D \tag{3.21}$$

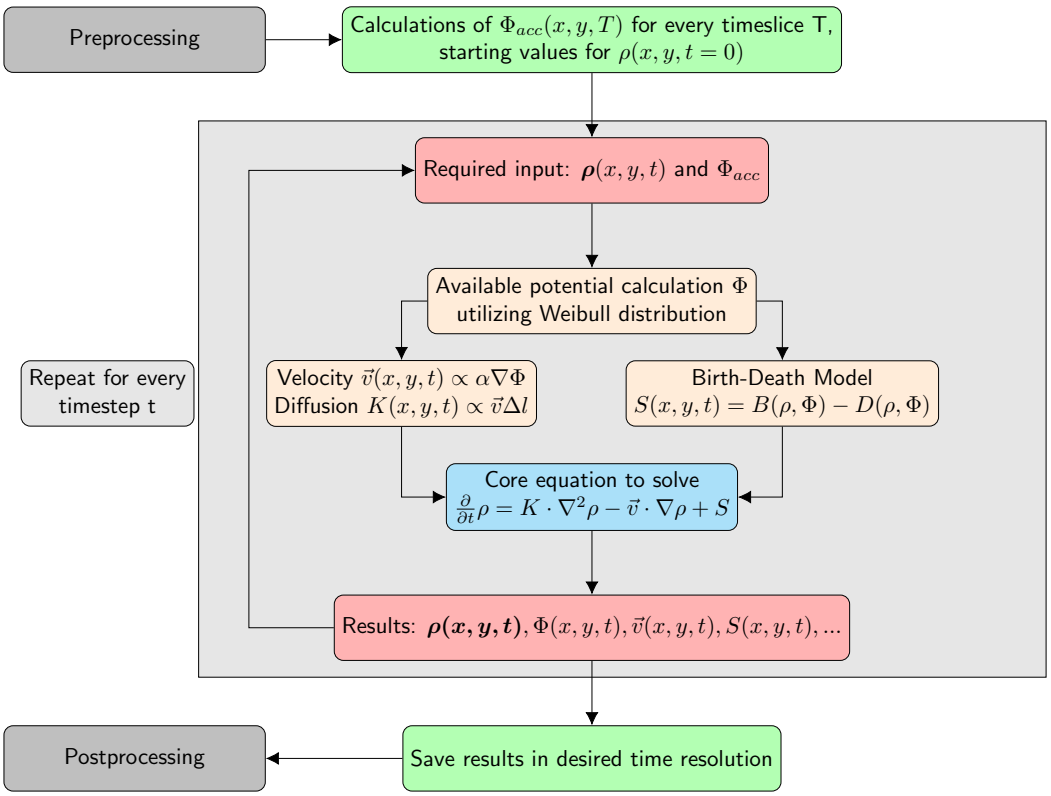

FIGURE 3.7: Overview of the Human Dispersal Model. The main required inputs, among others, are the Accessible HEP and a starting distribution of humans. The model calculates necessary parts for the core equation with the input data. The results of the core equation are used for the next timestep. Intermediate results are saved at a specified time resolution.

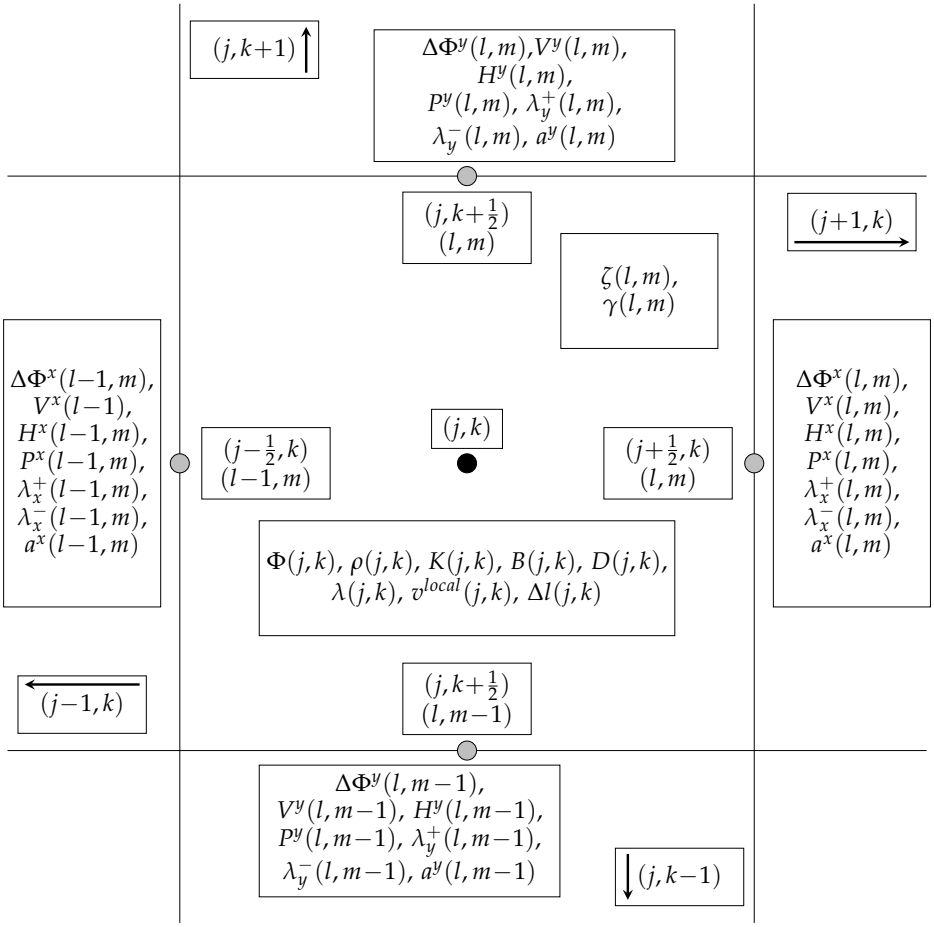

FIGURE 3.8: List of used variables in the model and their position on the grid (black dot) and sub-grid (gray dots). The central box lists the variables on the main grid, while the other lists depict the ones on the sub-grid. ζ and γ are defined on the sub-grid as well, but are combinations of x and y components, so there shift is technically $(j+1/2, k+1/2)$.

3.4.3 Numerical solver

The governing Equation 3.19 can be numerically solved on an equidistant two-dimensional grid with size $N \times M$ and the spatial spacing Δx and Δy for equidistant grids or $\Delta x(x,y)$ and $\Delta y(x,y)$ otherwise. The numerical time stepping Δt is fixed and the location indices j and k (e.g. $\Phi_{j,k}$) stand for the x and y coordinates on the grid. We define the human population $\lambda_{j,k}(t)$ and the diffusion coefficient $K_{j,k}$ on the grid, the individual time-steps as t. The velocity components $V^{x,y}$ and the gradients of the AvHEP $\Phi^{x,y}$ are defined on the intermediate steps between two grid points, the grid cell edges, which follows the finite volume approach. The resulting sub-grid has the size $(N-1) \times (M-1)$ and is shifted by $\frac{\Delta x}{2}$ or $\frac{\Delta y}{2}$, so that $V^x_{l,m} = V^x_{j+\frac{1}{2},k}$ and $V^y_{l,m} = V^y_{j,k+\frac{1}{2}}$. An overview of all used symbols and their position on the grid and sub-grid is given in Figure 3.8.

Gradient of the Available Potential

The AvHEP is the main driver for several processes in the dispersal model. It's gradient is required for the estimation of velocity and diffusion fluxes, and the direction of the gradient determines the asymmetry of the diffusion flux. The gradients are calculated on the edges of the grid cells:

$$\nabla\Phi^{x}_{l,m} = (\Phi_{j+1,k} - \Phi_{j,k})/\Delta x \tag{3.22}$$

$$\nabla\Phi^{y}_{l,m} = (\Phi_{j,k+1} - \Phi_{j,k})/\Delta y \tag{3.23}$$

The angle between the direction of the gradient and the Cartesian direction of x is required for the diffusion fluxes and is calculated for every pair of $\Phi^{y/x}$ gradients:

$$\xi_{l,m} = \arctan(-\frac{\nabla\Phi^{y}_{l,m}}{\nabla\Phi^{x}_{l,m}}) \tag{3.24}$$

Velocity

To determine $\vec{v}$, we assume that the equation of motion is proportional to the gradient of the AvHEP Φ and a scaling parameter γ:

$$\frac{d\vec{v}}{dt} = \frac{\partial\vec{v}}{\partial t} + \left(\vec{\nabla}\vec{v}\right)\vec{v} = \frac{\partial\vec{v}}{\partial t} + v_x\frac{\partial v_x}{\partial x} + v_y\frac{\partial v_y}{\partial y} = \gamma\vec{\nabla}\Phi \tag{3.25}$$

We further assume that the convection part of the total derivative is small compared to the change with time. In addition, a friction term is introduced to counteract the acceleration through the potential, leaving the following equation:

$$\frac{\partial\vec{v}}{\partial t} = \gamma\vec{\nabla}\Phi - \mu\vec{v} \tag{3.26}$$

The solution of Equation 3.26 can be written as:

$$\frac{\vec{v}(t_{i+1}) - \vec{v}(t_i)}{\Delta t} = \gamma\vec{\nabla}\Phi - \mu\vec{v}(t_i) \tag{3.27}$$

$$\vec{v}(t_{i+1}) = \Delta t\gamma\vec{\nabla}\Phi + \vec{v}(t_i) - \Delta t\mu\vec{v}(t_i) \tag{3.28}$$

with

$$\gamma = \frac{v_{max}}{\left(\vec{\nabla}\Phi\right)_{max}\cdot\Delta\tau} \tag{3.29}$$

The equation implies that the next velocity depends on the former velocity $\vec{v}(t_i)$, on the scaling parameter $\gamma(x,y)$, the friction parameter μ and on the gradient field $\vec{\nabla}\Phi$ of the available

HEP. $\Delta\tau$ is the reaction time of humans, which describes the time needed to adjust to the new conditions around them and change their travel speed and direction accordingly.

When the reaction time $\Delta\tau$ becomes very small compared to the time step Δt, e.g. the humans are able to quickly change their traveling speed and direction, we can neglect the impact of the former velocity $\vec{v}(t_i)$ and the friction term, and write the velocity as:

$$\vec{v}(t_{i+1}) = \gamma\vec{\nabla}\Phi \tag{3.30}$$

The friction term is necessary to ensure that the velocity breaks down when the gradient of the HEP closes to zero, halting the advection-induced flow of humans between the grid points if their level of potential is equal. The only exchange would be through diffusion in the absence of a gradient in the potential.

The velocity is limited by the maximal allowed speed, v_{max}, to ensure numeric stability in the model. If $|\vec{v}(t_{i+1})|$ surpasses the maximum velocity then its value is set equal to the maximal allowed speed. The direction of the velocity is conserved in this case. These calculations are done at every numerical time step.

The vector is split up into each dimension for further calculations within the numeric model:

$$V^x_{l,m}(t+1) = \Delta t \cdot \gamma_{l,m} \cdot \nabla\Phi^x_{l,m} + V^x_{l,m}(t) - \Delta t\mu V^x_{l,m}(t) \tag{3.31}$$

$$V^y_{l,m}(t+1) = \Delta t \cdot \gamma_{l,m} \cdot \nabla\Phi^y_{l,m} + V^y_{l,m}(t) - \Delta t\mu V^x_{l,m}(t) \tag{3.32}$$

The maximal gradient of potential $\left(\vec{\nabla}\Phi\right)_{max}$ is required to bring the first term of the right hand side of Equation 3.28 to the physical dimensions of a velocity (km yr^{-1} or m s^{-1}). It also scales the gradient of potential to the velocity, with the maximal gradient resulting in the maximal velocity within the parameter γ. The maximal gradient of potential was recalculated at every time step in a former version of the model to account for population evolution and to encourage dispersal within populated areas in times of slow or halted expansion. This version was used for the idealized simulations (Chapter 4). In the newest version, which is used for the application examples of the model (Chapter 5 & 6), the maximal gradient is predetermined within the configuration, as the model was prone to induce expansion in time frames of expected halt or retreat into certain regions.

Diffusion

The diffusive part of the model covers the exploration of the landscape by hunter-gatherers. In an area of constant survival conditions, humans move around to explore the region and gather resources from several places, allowing renewable sources to reface. Even in directed dispersal, some humans may stray away from the path to explore and gather. In the model

this behavior is simulated by applying a diffusion that depends on the direction and strength of the velocity $\vec{v}(x, y)$. The strongest diffusion happens perpendicular to the velocity, i.e. in a 90 degree angle to the travel direction. The diffusion parameter $K(x, y)$ (here $K_{j,k}$) is set by the following equation and recalculated every numeric time step:

$$K_{j,k} = \begin{cases} \epsilon \cdot \Delta l \cdot v_{j,k}^{local} & v_{j,k}^{local} \geq v^{crt} \\ \epsilon \cdot \Delta l \cdot v^{crt} & v_{j,k}^{local} \leq v^{crt} \end{cases} \tag{3.33}$$

ϵ is a model parameter that determines the relative strength of the diffusion term compared to the advection term, the Advection-Diffusion Fraction. It can be seen as a measure of the standard deviation of $\vec{v}$. The characteristic length $\Delta l_{j,k}$ is the mean distance between the grid point to its neighbors:

$$\Delta l_{j,k} = (\Delta x_{j,k} + \Delta x_{j-1,k} + \Delta y_{j,k} + \Delta y_{j,k-1}) \cdot \frac{1}{4} \tag{3.34}$$

$$\text{or for equidistant grids: } \Delta l = (\Delta x + \Delta y) \cdot \frac{1}{2} \tag{3.35}$$

$v_{j,k}^{local}$ is the average velocity of a grid point, calculated by averaging the fluxes on all four edges of the grid cell:

$$v_{j,k}^{local} = \frac{|V_{j,k}^{x} + V_{j-1,k}^{x}| + |V_{j,k}^{y} + V_{j,k-1}^{y}|}{4} \tag{3.36}$$

v^{crt} is the critical velocity. It ensures an active diffusion in regions of low or zero velocity. If the local velocity is lower than the predefined critical velocity, than the critical velocity is used in the calculations of the diffusion parameter. The default value of the critical velocity is set to 50% of the maximal velocity.

Time stepping

We use the Runge-Kutta 4-step method (RK4) for time and the high resolution central scheme from Kurganov and Tadmor (2000) for space to solve Equation 3.21. To apply RK4, we rewrite the equation to simplify the right hand side R_{hs}:

$$\begin{aligned} \frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) &= R_{hs}(t, \lambda(t)) \\ \lambda(t+1) &= R_{hs}(t, \lambda(t)) \end{aligned} \tag{3.37}$$

The next time step is calculated by merging the sub-steps k_1 to k_4:

$$\begin{aligned}
\lambda(t+1) &= \lambda(t) + (k_1 + 2k_2 + 2k_3 + k_4)/6 \\
k_1 &= \Delta t \cdot R_{hs}(t, \lambda(t)) \\
k_2 &= \Delta t \cdot R_{hs}\left(t + \frac{\Delta t}{2}, \lambda(t) + \frac{\Delta t}{2} \cdot k_1\right) = \Delta t \cdot \left(R_{hs}(t, \lambda(t)) + \frac{k_1}{2}\right) \\
k_3 &= \Delta t \cdot R_{hs}\left(t + \frac{\Delta t}{2}, \lambda(t) + \frac{\Delta t}{2} \cdot k_2\right) = \Delta t \cdot \left(R_{hs}(t, \lambda(t)) + \frac{k_2}{2}\right) \\
k_4 &= \Delta t \cdot R_{hs}\left(t + \Delta t, \lambda(t) + \Delta t \cdot k_3\right) = \Delta t \cdot \left(R_{hs}(t, \lambda(t)) + k_3\right)
\end{aligned} \tag{3.38}$$

The right-hand side R_{hs} is defined by the advection fluxes H, the diffusion fluxes P, and the sources and sinks $B - D$. The fluxes describe the exchange of human population through the cell edges between two grid points, the sources and sinks depict the birth and death rate of humans in a cell:

$$\begin{aligned}
R_{hs}(t, \lambda) = &- \frac{H^x_{l,m} - H^x_{l-1,m}}{\Delta x} - \frac{H^y_{l,m} - H^y_{l,m-1}}{\Delta y} \\
&+ \frac{P^x_{l,m} - P^x_{l-1,m}}{\Delta x} + \frac{P^y_{l,m} - P^y_{l,m-1}}{\Delta y} \\
&+ B - D
\end{aligned} \tag{3.39}$$

Fluxes

The advection fluxes are defined as:

$$\begin{aligned}
H^x_{l,m} &= \frac{1}{2} \cdot \left(V^x_{l,m} \cdot \left(\lambda^+_{x;\,l,m} + \lambda^-_{x;\,l,m}\right) - a^x_{l,m} \cdot \left(\lambda^+_{x;\,l,m} - \lambda^-_{x;\,l,m}\right)\right) \\
H^y_{l,m} &= \frac{1}{2} \cdot \left(V^y_{l,m} \cdot \left(\lambda^+_{y;\,l,m} + \lambda^-_{y;\,l,m}\right) - a^y_{l,m} \cdot \left(\lambda^+_{y;\,l,m} - \lambda^-_{y;\,l,m}\right)\right)
\end{aligned} \tag{3.40}$$

with the intermediate values of $\lambda^{+/-}_{x/y;\,l,m}$ using a minmod flux-limiter. $\eta \in [1, 2]$ depicts a model parameter in this equation that adjusts the limiter in its impact ($\eta \to 1$ creates the most dissipative results):

$$\begin{aligned}
\lambda^{\pm}_{x;\,l,m} &= \lambda_{j+1,k} \mp \frac{\Delta x}{2} \cdot \mathrm{minmod}_x\left(\lambda_{j+\frac{1}{2}\pm\frac{1}{2},k}\right) \\
\lambda^{\pm}_{y;\,l,m} &= \lambda_{j,k+1} \mp \frac{\Delta y}{2} \cdot \mathrm{minmod}_y\left(\lambda_{j,k+\frac{1}{2}\pm\frac{1}{2}}\right)
\end{aligned} \tag{3.41}$$

$$\begin{aligned}
\text{minmod}_x(\lambda_{j,k}) &= \begin{cases} \min(\eta\frac{\lambda_{j,k}-\lambda_{j-1,k}}{\Delta x}, \frac{\lambda_{j+1,k}-\lambda_{j-1,k}}{2\Delta x}, \eta\frac{\lambda_{j+1,k}-\lambda_{j,k}}{\Delta x}) & \text{, if all terms are positive} \\ \max(\eta\frac{\lambda_{j,k}-\lambda_{j-1,k}}{\Delta x}, \frac{\lambda_{j+1,k}-\lambda_{j-1,k}}{2\Delta x}, \eta\frac{\lambda_{j+1,k}-\lambda_{j,k}}{\Delta x}) & \text{, if all terms are negative} \\ 0 & \text{, otherwise} \end{cases} \\
\text{minmod}_y(\lambda_{j,k}) &= \begin{cases} \min(\eta\frac{\lambda_{j,k}-\lambda_{j,k-1}}{\Delta x}, \frac{\lambda_{j,k+1}-\lambda_{j,k-1}}{2\Delta x}, \eta\frac{\lambda_{j,k+1}-\lambda_{j,k}}{\Delta x}) & \text{, if all terms are positive} \\ \max(\eta\frac{\lambda_{j,k}-\lambda_{j-1,k}}{\Delta x}, \frac{\lambda_{j,k+1}-\lambda_{j,k-1}}{2\Delta x}, \eta\frac{\lambda_{j,k+1}-\lambda_{j,k}}{\Delta x}) & \text{, if all terms are negative} \\ 0 & \text{, otherwise} \end{cases}
\end{aligned} \tag{3.42}$$

The local speeds $a^{x/y}$ are less complex than in the general case presented by Kurganov and Tadmor (2000). The Jacobian of the flux form simplifies to the first derivative of the flux form due to the scalar case of λ. This results in the following equations for the local speed:

$$\begin{aligned} a^x_{l,m} &= |V^x_{l,m}| \\ a^y_{l,m} &= |V^y_{l,m}| \end{aligned} \tag{3.43}$$

The diffusion fluxes $P^{x,y}$ do not require the complex flux-limiting approach, which enables the direct calculation from the population λ. The direction of the diffusion and its strength is determined by the angle ξ (Equation 3.24). The directional diffusion parameters are obtained through the sine and cosine of $xi_{l,m}$ and the local diffusion parameter $K_{j,k}$:

$$K^x_{j,k} = |K_{j,k} \cdot \cos(\xi_{l,m})| \tag{3.44}$$

$$K^y_{j,k} = |K_{j,k} \cdot \sin(\xi_{l,m})| \tag{3.45}$$

$$\begin{aligned} P^x_{l,m} &= \frac{1}{2} \cdot \left(K^x_{j,k} \cdot \frac{\lambda_{j+1,k} - \lambda_{j,k}}{\Delta x} + K^x_{j+1,k} \cdot \frac{\lambda_{j+1,k} - \lambda_{j,k}}{\Delta x} \right) \\ P^y_{l,m} &= \frac{1}{2} \cdot \left(K^y_{j,k} \cdot \frac{\lambda_{j,k+1} - \lambda_{j,k}}{\Delta y} + K^y_{j,k+1} \cdot \frac{\lambda_{j,k+1} - \lambda_{j,k}}{\Delta y} \right) \end{aligned} \tag{3.46}$$

This numeric approach satisfies the scalar total-variation diminishing (TVD) property. It can capture sharp shocks in λ that appear at the expansion front and on discontinuities (coastline, mountains, glaciers) without creating artificial oscillations. Usually, a fine grid is required to capture these shocks, which drastically increases computation time. Using a coarse grid with discrete schemes can create false predictions or be dominated by oscillations (over- and under-estimations). The approach by Kurganov and Tadmor (2000) allows using coarser grids while still capturing these shocks, avoiding the oscillations at fixed borders.

The numeric model is stable as long as the following criteria is maintained:

$$\begin{aligned} &\max\left(\frac{\Delta t}{\Delta x}\max|V^x_{l,m}|, \frac{\Delta t}{\Delta y}\max|V^y_{l,m})| \right) \leq \frac{1}{8} \\ \Leftrightarrow \quad &\max\left(\frac{\Delta t}{\Delta x}|v_{max}|, \frac{\Delta t}{\Delta y}|v_{max}| \right) \leq \frac{1}{8} \end{aligned} \tag{3.47}$$

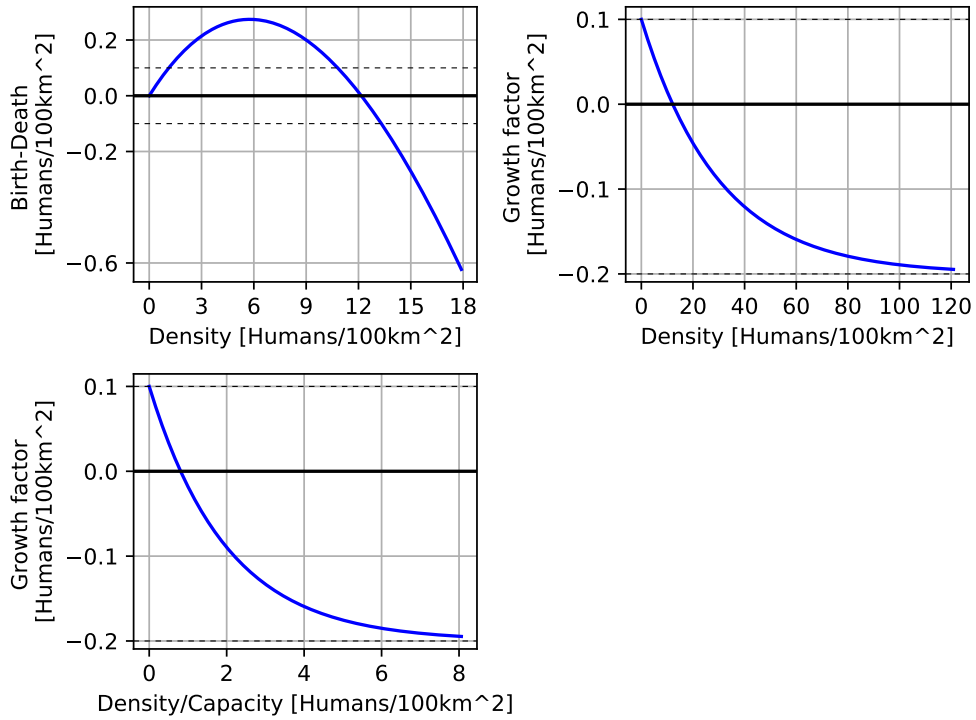

FIGURE 3.9: Different visualizations of the Birth-Death function and the related growth factor. The bottom one uses a density scaled by the carrying capacity. Further explanations are found within the text.

Birth-Death-Model

The sources and sinks in Equation 3.19 are covered in the numeric solver as well. The basis is the growth parameter r_B which depicts the yearly ratio of reproduction, including the loss by death through old age or injury. The upper limit is defined by the carrying capacity Φ_{max} (see Equation 3.10).

$$B - D = \lambda_{j,k} \cdot r_B \cdot \left[3 \cdot \exp\left(\frac{-\lambda_{j,k}}{2 \cdot \Phi_{j,k}^{max;}} \right) - 2 \right] \tag{3.48}$$

Figure 3.9 shows an example of the impact that the Birth-Death term can have. For this case, $r_B = 0.1\ \mathrm{yr}^{-1}$ and $\Phi_{max} = 15\ \mathrm{Humans}/100\mathrm{km}^2$. On the top left is the net change of population plotted based on the given density of humans. The top right shows the growth factor, which is the result of Equation 3.48 without the multiplication of the population λ. The bottom left depicts the same curve as shown in the top right, but the density was normalized with the carrying capacity.

These plots show that the population increases as long as it stays below the carrying capacity. The growth comes to a halt when roughly 80 % of the carrying capacity is reached, which mirrors the population control behavior of hunter-gatherers. This control ensures that natural and renewable sources are not depleted strong enough to create permanent depletion or damage.

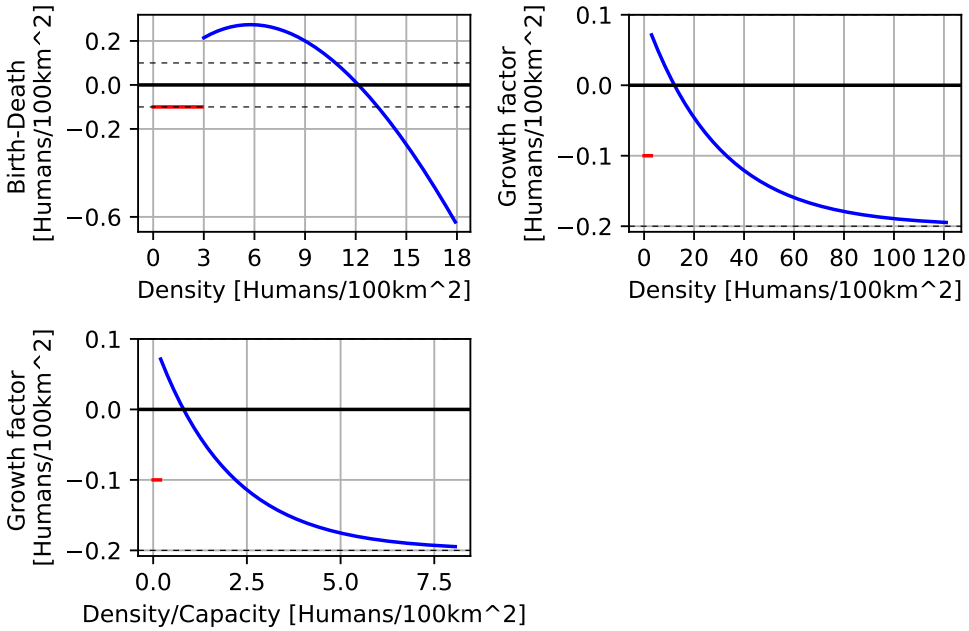

FIGURE 3.10: Different visualizations of the Birth-Death function and the related growth factor, modified with the reproduction onset for very low populations. The bottom one uses a density scaled by the carrying capacity. Further explanations are found within the text.

To defer this threshold from the actual carrying capacity, we define this value as the cultural capacity. The carrying capacity is, therefore, only reached if larger populations migrate into already crowded areas. If that happens, the birth-death rate becomes strongly negative, mirroring conflict over resources and starvation. This effect increases drastically when the density overshoots the capacity by several factors. These cases are solved by further migration to avoid prolonged decay of the total population (Kelly, 2010).

The Birth-Death term works well for most population sizes. The dynamics for low populations are, however, more challenging to represent. Smaller populations are more likely to die out over time, partly due to the lack of mates for reproduction and fewer people being able to forage food and resources. This problem is evaded by increased internal movement between groups of hunter-gatherers, either forming larger groups or exchanging members to allow for new offspring. An additional parameter is introduced into the model to account for this behavior: The reproduction onset κ. This parameter describes the lower boundary of a population to be stable enough to grow over time. If the current population is below this threshold and does not receive an inflow of other humans, then it goes extinct. In the model, it controls at which density of humans within a grid cell the switch is made between Equation 3.48 and a flat decrease of -0.1 total humans per year. Figure 3.10 shows the impact of this change in red for an example reproduction onset of $\kappa = 3\,\text{Humans}/100\text{km}^2$. This value is quite high and should be changed depending on the experiment and the predefined assumptions.

3.4.4 Requirements and Workflow of the script

This section covers the technical requirements and the workflow of the script after the accessible HEP. The source code for the Accessible HEP, the pre- and post-processing, and the HDM can be found within my GitHub repository (Wegener, 2021).

Preprocessing

Before the accessible HEP can be used as an input, some additional preparations need to be done to account for the requirements of the Fortran solver. First, the uneven grid is taken into account by calculating the distance between each grid point. Second, this information is used to calculate the corner points of each grid cell, which are used again to calculate the area of the cell. For both steps, the python library geographiclib is used, utilizing the WGS84 ellipsoid standard.

With the grid area, we can calculate the density of humans (and the carrying capacity) into an absolute number of humans specific for each cell. This approach ensures that exchange between grid points is calculated with the actual number of humans rather than their density, minimizing the error through different grid-cell sizes. The starting distribution of humans is set during the preprocessing as well.

Configuration

The first step is the configuration of the model. The data path and variable names are declared here, as well as the output names. Model-specific values, such as the time-stepping and total run-time, are set here, as well as the save interval in the output. Critical are the allowed maximum velocity, the velocity factor, and the diffusion coefficient. These have a significant impact on the model and on the numerical stability. Additionally, the growth and reproduction onset parameters are set here, and individual parts of the governing equation can be turned off.

Essential are two switches: The switch to define the incoming grid as equidistant or distorted and the one that distinguishes between a single potential slice and several in the input file. The former changes the dx/dy from assuming a single value to individual values between neighboring grid cells. The latter allows the change of the potential over time. The script closes with the name definitions of the output file.

Setup

After the configuration, the model is set up by defining the required variables, loading the data, and allocating space for the dynamic variables. An error routine catches the netCDF calls

if they fail to work for various reasons and give a simplified error message. The total number of time steps is calculated, and different counters are set up to catch the current time step, year, time slice, and loss of humans that move out of the system. The available potential is calculated once as the difference between the accessible potential and the starting distribution of humans. Later it will be calculated every time step depending on the current distribution of humans.

The numeric solver includes two grid points in each direction to ensure stability and the capturing of sources, sinks, and sharp gradients. Therefore, two rows and columns of grid points are considered as borders around the simulated area of interest. These points will have potential values but will not allow human density on them. Instead, humans that are directly next to these borders (3rd column or row of grid points) are taken out of the system after each time step. It is possible to simulate closed borders by setting the masks for these points accordingly in the preprocessing or use the appropriate switch in the configuration file.

If more than one slice of potential is used, a linear interpolation is calculated between each slice. It allows for a gradual change of the potential between two slices rather than sudden changes when the next potential is reached. Lastly, the functions that cover some parts of the numeric solver (Section 3.4.3) are formulated during the setup, easing the implementation of parallel processing of the script.

Run

When the setup is finished, the solver can start the calculations. The central part is a loop over all numerical time steps. It starts with setting the variables that contain the number of humans for the last and the current time step. Both are required for the RK4 scheme to function. The next step is the calculation of the advection speed, see Equation 3.28 and following. The borders are excluded, as their speed is not required during the process. The advection velocity is calculated between two neighboring grid points unless one of them is masked as invalid, resulting in a zero speed between them. If the advection velocity exceeds the maximum allowed speed, it is set to that maximum to ensure numeric stability, keeping its original direction. The diffusion parameter requires fewer steps to be calculated, see the Equation 3.33 and following. Again, the borders are excluded. The local grid and speed are calculated first for each grid point. If the local speed is higher than the critical velocity, then the ladder one is used instead in the calculations of the diffusion parameter. With the advection speed and diffusion parameter known for the time step, we can initiate the RK4 solver. Again, we exclude the border points and in advance. For every other point, we invoke the solver function. This function checks first if the given grid point is valid or not, skipping the calculations for the latter case. For a valid point the intermediate values (Equation 3.41) are estimated using the minmod function (3.42), followed by the local speeds (Equation 3.43). Now the convection fluxes on the borders between the grid point and each neighbor can be calculated with Equation 3.40. Combining these

fluxes results in the Advection term of the governing Equation (3.19). The diffusion fluxes can be calculated directly, providing when combined the diffusion term. Finally, the Birth-Death term is calculated (Equation 3.48).

The numeric solver distinguishes between equidistant and non-equidistant grids and uses different values for each grid. Which grid to use can be changed in the configuration. Additionally, all three terms can be turned off individually.

During the RK4 scheme, the solver function is called four times. Each result is temporarily stored and combined afterward, following the first line of Equation 3.38. This new distribution of humans is now further changed by removing the population on the border, simulating that they moved out of the system. The total loss is stored and saved in the output. For temporal changing potentials, the linear interpolation is calculated again when the next time slice is reached. The model assumes that the potential changes every nth year, as provided in the configuration script. Every time step, we recalculate the available potential by using the new distribution of humans. Afterward, the results are saved in specific intervals as defined in the configuration. The progress of the model is prompted and some temporary values are cleared. The temporal loop starts again at this point, solving the next time step and repeating the process until the simulation time is reached.

Chapter 4

Idealized dispersal runs

A series of experiments have been conducted to test the capabilities of the HEM, each one focusing on different parts of the model. The idealized dispersal runs were designed to target the numeric structure of the model and evaluate their behavior to ensure that the model results are consistent with the assumptions and expectations. The convergence of the used discretization is not further discussed and can be found in the work of Kurganov and Tadmor (2000). In this chapter, idealized HEPs and distributions of humans were used to simplify the calculations and interpretations of the results.

4.1 Experiment Setup

A series of nine idealized experiments have been done to test the model and its different parts. The first eight experiments are combinations of two different HEP distributions (a) & (b) and four different population initializations (c)-(f) (see Figure 4.1). The first HEP (a) is spatial equal with a fixed value of 10 Humans/100km^2, the second (b) is a linearly increasing HEP in x-direction reaching the previous value. The four starting distributions are a block with a low population (c, 2 Humans/100km^2), a block with a high population (d, 10 Humans/100km^2), a block with a high population surrounded by a low population in the rest of the domain (e), and dense population (f) surrounded by low population with an isolated island in the domain. These distributions are useful to investigate different behaviors of the model. For the spatial equal HEP simulations (a), both initializations (d) and (e) use a doubled population density (20 Humans/100km^2) to kick-start the reactions of the model, making the visualizations clearer.

For all experiments, an equidistant grid consisting of 80 points in the x-direction and 60 points in the y-direction is used, with each grid point being 50 km apart from the other. The outermost three grid points are masked to ensure an enclosed domain.

The dispersal model consists of 3 significant terms: advection, diffusion, and birth-death. Each term needs to be investigated independently. In addition, the combination of advection and

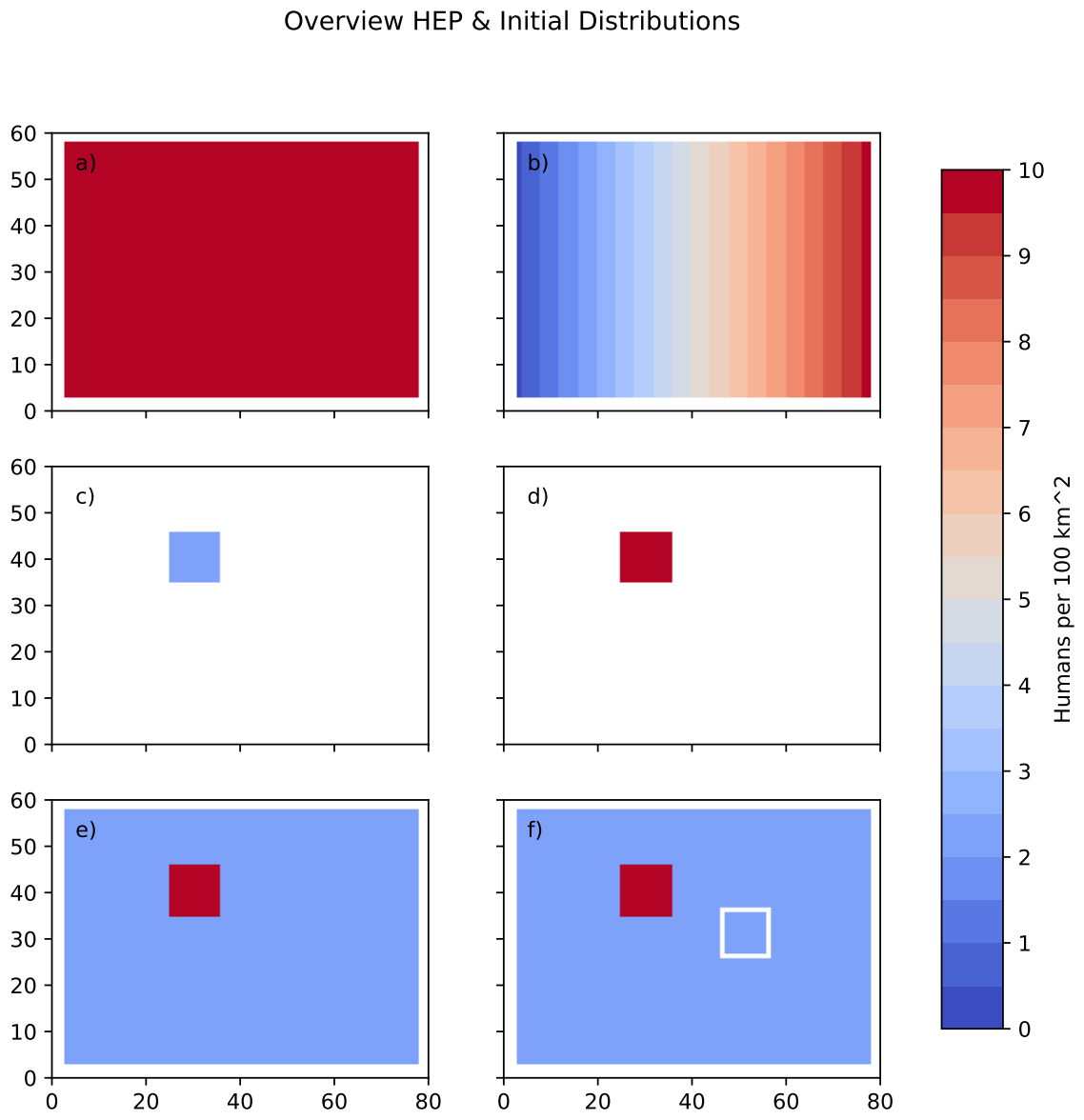

FIGURE 4.1: Overview over the different HEPs and distributions used to initialize each test. (a) Spatially equal HEP, (b) Eastward linearly increasing HEP, (c) Low population initialization, (d) High population initialization, (e) High population surrounded by low population initialization, (f) as (e) with an isolated island of low population.

diffusion introduces a different dynamic that is required to be examined. Using all three terms at once as the last combination results in five iterations for each experiment: advection, diffusion, birth-death, dispersal only, and all terms. Multiplying the eight experiments with five iterations results in 40 different runs that target different model parts within different starting conditions. Table 4.1 gives an overview and lists the abbreviation of each run.

TABLE 4.1: Overview of the realized experiments and their abbreviations. Some iterations in Experiment 9 consist of two simulations, one with a higher and one with a lower value than the reference case. (pop.: population, sur.: surrounding)

Experiments		Iterations				
HEP type	Starting Distribution	Advection	Diffusion	Birth-Death	Dispersal only	All terms
Spatial equal	Low pop.	1a	1b	1c	1d	1e
	High pop.	2a	2b	2c	2d	2e
	High pop. sur.	3a	3b	3c	3d	3e
	High pop. sur. & island	4a	4b	4c	4d	4e
Eastward linear increasing	Low pop.	5a	5b	5c	5d	5e
	High pop.	6a	6b	6c	6d	6e
	High pop. sur.	7a	7b	7c	7d	7e
	High pop. sur. & island	8a	8b	8c	8d	8e

Experiment 9 - Parameters		
Reference Case (9a)	Advection-Diffusion-Fraction (9b)	Maximum Velocity (9c)
Critical Velocity (9d)	Flux Limiter (9e)	Growth Rate (9f)
Reaction Time (9g)	Reproduction Onset (9h)	Spatial Resolution (9i)
Temporal Resolution (9j)	Open Borders (9k)	Gaussian Distribution (9l)

The first eight experiments use the parameter values listed in Table 4.2. Each iteration runs for 1000 years, with fifty time-steps each year. Some experiments were continued with an additional 1000 years to look for possible steady-state solutions. With a maximum velocity of 5 km yr^{-1} and a spatial resolution of 50 km, this fulfills the stability constraint of Equation 3.47. A slightly higher reproduction rate of 0.015 is chosen so that the population number can adjust quickly enough to be covered within the simulation time. In addition, a higher reproduction onset is used to enlarge the population sink within the linear cases, but not high enough to trigger a large-scale extinction within the "low surrounding" population cases.

TABLE 4.2: List of parameters and their values used in the first eight experiments.

Parameter	Symbol	Value	Parameter	Symbol	Value
Gridpoint Distance	$\Delta x \& \Delta y$	50 km	Time Step	Δt	0.02 yr
Simulation Time	T	1000 yr	Reaction Time	$\Delta \tau$	1yr
Maximum Velocity	v_{max}	5 km yr^{-1}	Critical Velocity	v_{crt}	2.5 km yr^{-1}
Weibull Shape	α	2.5	Weibull Scale	β	0.4
Advection-Diffusion Fraction	ϵ	0.5	Flux-Limiter	η	1.5
Growth Parameter	r_B	0.015 yr^{-1}	Reproduction Onset	κ	1.9 H/100km^2

The ninth experiment focuses on the parameters used in the model and investigates their impact. A reference run is conducted with typical values for each parameter and compared to runs with higher and lower values of each parameter. Instead of a square starting distribution of humans, an additional run was made with a Gaussian starting distribution to catch any significant differences between these conditions. In addition, runs with higher spatial and temporal resolution were created to investigate if re-scaling creates additional differences. Again, each iteration runs for 1000 years but with slightly different parameters, which are listed in Table 4.3. The Reference Case uses, as the name suggests, the Reference Values. The same is true as well for the experiments Gaußian Distribution and Open Borders. In addition, note that the time step is increased to 100 steps per year for the Experiments "Maximum Velocity" and "Resolution" to satisfy the stability constraint. The parameters add 19 additional runs, resulting in 59 idealized runs total.

TABLE 4.3: List of parameters and their values used in the ninth experiment.

Experiment and Changed Parameter (Symbol)	Low Value	Reference Value	High Value
Advection-Diffusion Fraction (ϵ)	0.1	0.5	0.9
Maximum Velocity (v_{max})	1 km yr^{-1}	5 km yr^{-1}	10 km yr^{-1}
Critical Velocity (v_{crt})	$0.25 \cdot v_{max}$	$0.5 \cdot v_{max}$	$0.75 \cdot v_{max}$
Flux Limiter (η)	1	1.5	2
Growth Parameter (r_B)	0.005	0.015	0.05
Reaction Time ($\Delta \tau$)	0.1 yr	1 yr	5 yr
Reproduction Onset (κ)	1 H/100km^2	2 H/100km^2	5 H/100km^2
Resolution ($\Delta x, \Delta y$)		50 km	25 km
Time Step (Δt)		0.02 yr	0.01 yr

It should be noted that the idealized simulations are based on an older iteration of the model, which featured a dynamic determination of the maximal gradient of potential $\left(\vec{\nabla}\Phi\right)_{max}$. The model was changed for the applications in the following chapters as briefly described in the Methods (Section 3.4.3), making this variable a model parameter as well. In addition, the friction term and its parameter μ were introduced later as well. These parameters are not further discussed or tested as part of the ninth experiment.

4.2 Results

In this section, the results of the nine experiments are shown, focusing on one experiment at a time. Visualizations of the first eight experiments are given at the end of the Chapter (Section 4.4) showing the propagation of human density, while individual time slices or additional plots of particular interest are used within each section. The ninth experiment consists of repetitions of experiment 8e with changed parameters. Therefore, individual plots are used to investigate further, as the general propagation is similar to that experiment.

4.2.1 Experiment 1

The first experiment uses the spatial equal HEP and the low population initialization (corresponding to a) and c) in Figure 4.1). It focuses on the response of the model to a straightforward case of a low human population. Figure 4.24 visualizes some time steps for each iteration.

For the pure advection iteration (1a), no actual movement occurs in the model, and the starting distribution stays the same for the whole duration. This result was expected, as the homogeneous distribution ensures zero gradients of the available potential, resulting in zero velocities and no contribution of the advection term.

The pure diffusion (1b) shows a slow dispersion into all directions, lowering the overall density but keeping the population size constant, resulting in a Gaussian-shaped distribution that keeps flattening throughout the simulation. This outcome is expected as well due to the critical velocity (Section 3.4.3) forcing the diffusion in the absence of velocity at the first few time steps. Once the distribution of humans is no longer spatial equal, a velocity is calculated even if the advection term is turned off. If this velocity is greater than the critical velocity, then it enhances the diffusion speed.

The single low population initial distribution features a dense enough population to allow the onset of reproduction. It can be seen in the iteration of pure birth-death (1c), where the population increases rapidly within the first 400 years and then stagnates at roughly 80 % of the carrying capacity. The population does not increase further than that due to the growth factor reaching zero at this threshold. If the threshold is exceeded, the factor will turn negative, forcing the population to reduce and balance out.

A combination of different processes can be observed in the dispersal only iteration (1d). The population quickly forms clusters forced by the available potential and advection to reach the optimal group size, creating gaps between the clusters with deficient populations, some of which are low enough to be considered abandoned areas. A slow diffusion smears these clusters out and creates bridges between them with a sparse population. A significant expansion, however, is not achieved due to reproduction being disabled.

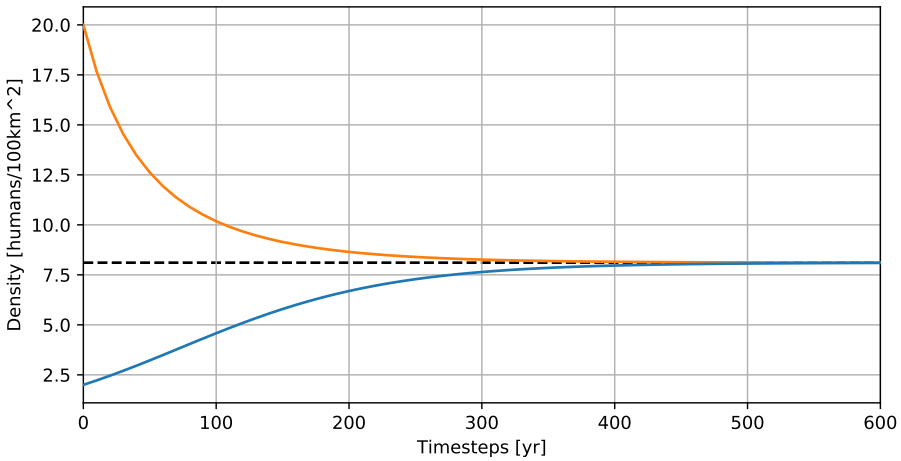

FIGURE 4.2: Population Evolution over simulation time of Experiments 1c (bottom, blue) and 2c (top, orange), with the cultural capacity as the dashed black line. For both time series, the gridcell (40,30) is used.

With all terms active (1e), the combinations of patterns described above appear: the clustering of the low-density population at the start, the growth of the population that encourages dispersal into lower populated regions, and the diffusion which makes uninhibited areas accessible. Combining these effects creates the overall population expansion, while avoided locations at the start fill up over time. With the current set of parameters, the domain is not fully covered within the simulation time of one thousand years. An extension of an additional 1000 years of simulation shows that the domain is covered with humans after 1550 years, with minor fluctuations of density in the range of 0.05 Humans/100km^2 at 2000 years of simulation, indicating a closure to a steady-state.

4.2.2 Experiment 2

The second experiment uses the spatial equal HEP and a high population initialization (corresponding to a) and d) in Figure 4.1). It mirrors the first experiment but with a much larger population at the start. Figure 4.25 shows the overall evolution of the different iterations.

Starting with the advection iteration again (2a), the same behavior as before can be seen (1a): The distribution stays the same for the whole simulation time. However, the diffusion (2b) shows a more vital expansion than 1b while maintaining the Gaussian shape. There is some potential for further expansion after the simulation time of 1000 years as the central population density is relatively high compared to the previous experiment.

With an initial population density above the carrying capacity, this iteration (2c) shows the halt of reproduction embedded in the model (see Equation 3.48, Figure 3.10). The population quickly drops exponentially within the first 100 years to half of its size while taking additional

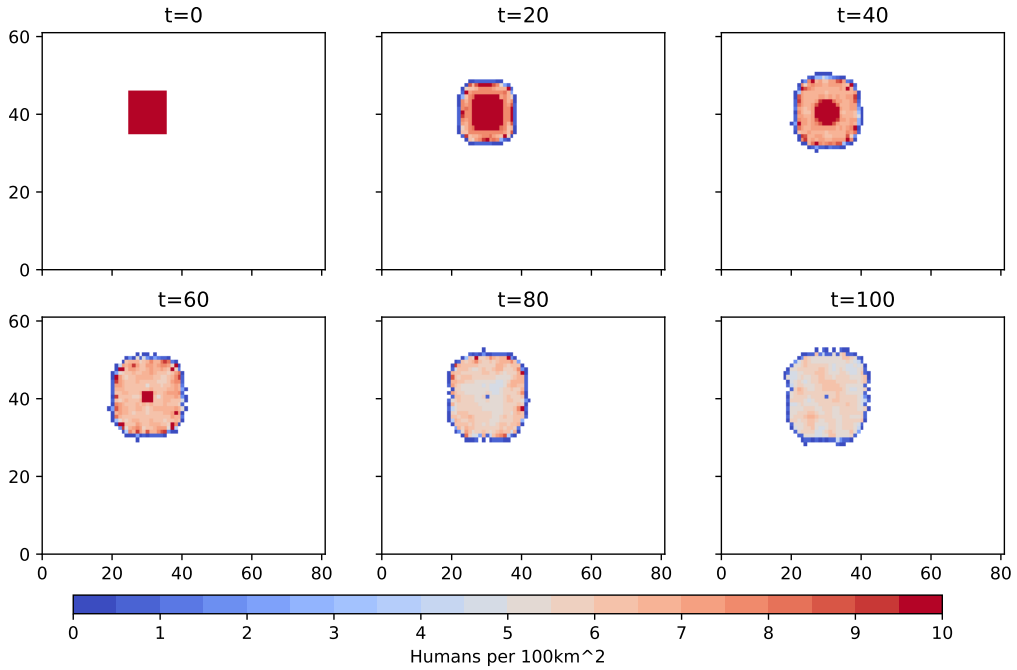

FIGURE 4.3: Snapshots of density distribution for the first 100 years of simulation, experiment 2c.

300 years to stabilize at roughly 80 % of the upper bound. Figure 4.8 compares the population evolution of experiment (1c) and (2c). It appears that the population growth in the former experiment takes more time to adjust to the cultural capacity than the declining population in the latter, despite the smaller difference in the capacity.

The increased population initialization activates the spatial spread in the pure dispersal iteration (2d). A dispersal into previously empty territory is visible instead of clustering together to stabilize the population above the Reproduction Onset. Most of the dispersal happens in the first 100 years of the simulation (see Figure 4.3), roughly quadrupling the occupation area (doubling the diameter). The initial spread results in the abandonment of the central grid point of the initial distribution, caused by the rapid migration to all four sides. After the migration burst outwards, further dispersal slows down quickly and is halted as the center regions reach the optimal group size. Some internal migration continues but slows down in the later parts of the simulation.

Combing all terms with the higher population (2e), a similar initial spread as in the previous iteration (2d) is visible, but the dispersal continues due to the now active population dynamics, allowing new offspring to populate the occupied regions. During the spread, strong dispersal within the occupied grid cells is observable, as well as an increasing population. As the population grows, the expansion accelerates as higher potential gradients at the borders induce higher velocities and stronger spread through diffusion. The domain is not fully covered within the simulation time but comes close to it. In the second run of this iteration with 2000 years of

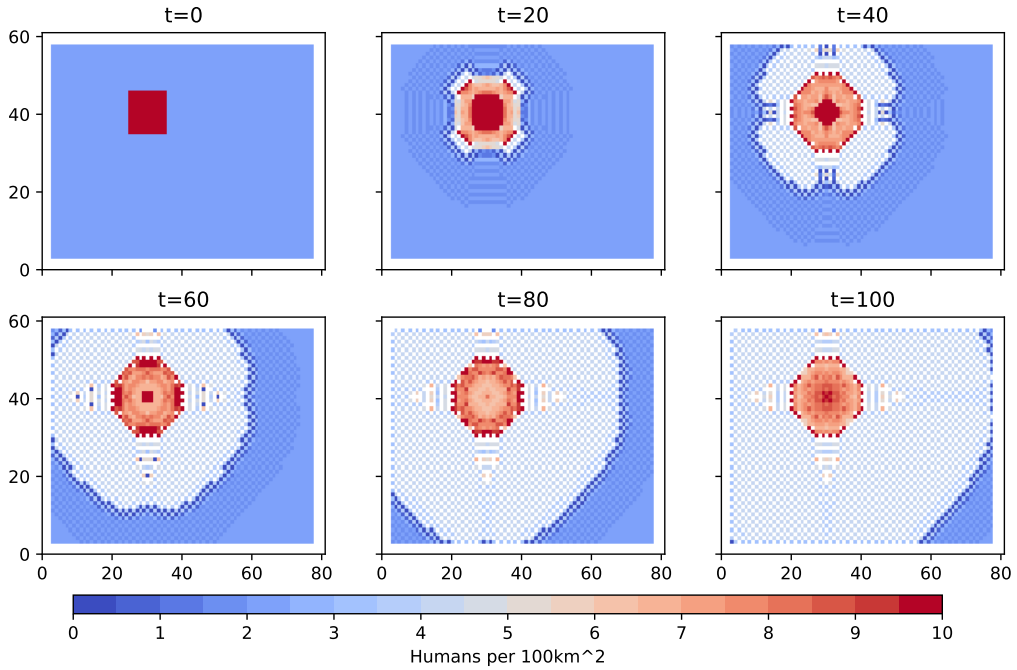

FIGURE 4.4: Snapshots of density distribution for the first 100 years of simulation, experiment 3a.

simulation time, the domain is fully conquered at 1220 years, while the population stabilizes at 1700 years with <0.02 Humans/100km^2 spatial difference, resulting in a steady state of the simulation.

4.2.3 Experiment 3

The third experiment uses the spatial equal HEP with a high and low population initialization (see a) and e) in Figure 4.1). This time, the domain is already occupied by a low population, and the reaction of the model is investigated when a large population starts in the middle of it. Again, an overview of population density for each iteration is given (Figure 4.26).

This iteration (3a) shows an unexpected model behavior, which the construction of the model can explain. In the first few time steps a dispersal of the high-density population into the lower density regions is observable, which comes to a halt after 50 years of simulation time (Figure 4.4). The reason for this is a numerical wave that propagates in front of the dispersal, triggered by the initial movement. It creates minor disturbances in the otherwise equal low population density, which in return excite further dispersal. Because of the low population, the dispersal favors the abandonment of regions and the clustering of the population, as seen in previous iterations (1d/e). The abandoned regions can not be crossed, creating a checkerboard-like pattern of population distribution. Along the cardinal directions, stripes form instead. The now trapped population in the center evens out over time.

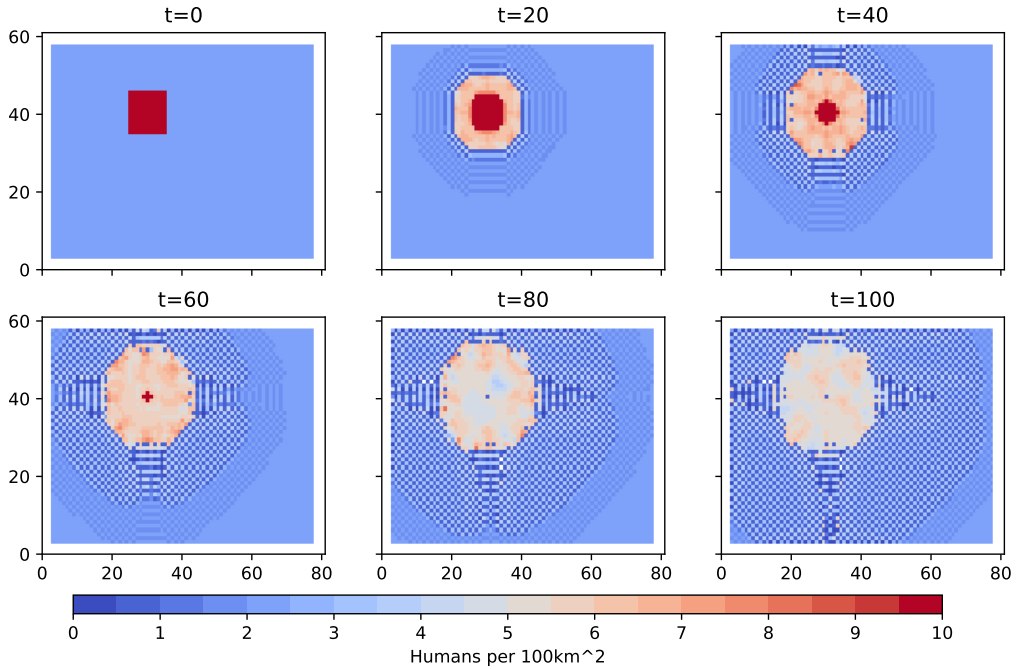

FIGURE 4.5: Snapshots of density distribution for the first 100 years of simulation, experiment 3d.

The pure diffusion (3b) is only minimally affected by the additional population. The spread is similar to 2b, albeit a bit slower and featuring a higher population in the center at the end of the simulation time of 1000 years. Focusing on the birth-death term (3c), we see the same behavior as in 2c and 1c for the two types of initial populations. Both adjust to one another over the simulation time.

The first time steps of the dispersal iteration (3d, Figure 4.5) mirrors the simulation of 3a, but the noted barrier at roughly 50 years is broken down by the diffusion term, allowing for a slightly further expansion until the optimal group size slows it down again. The checkerboard pattern is less distinctive in this iteration, as the diffusion fills the gaps until it equalizes with the advection term. Along the cardinal directions centered around the initial high distribution, the pattern is broken apart, and larger clusters emerge, although slow and sporadic at later simulation times.

The last iteration (3e, Figure 4.6) combines, once more, the previous observations: The initial dispersal that is halted by a border of low population areas, the strong internal dispersal, and for a short time, a checkerboard pattern. However, this time, the pattern is broken apart by the advection and diffusion into the lower density region, assisted by the population growth. The gaps close within the first 200 years of simulation time, creating an oscillating pattern of movement that breaks down at 500 years as the population reaches the equilibrium between birth and death.

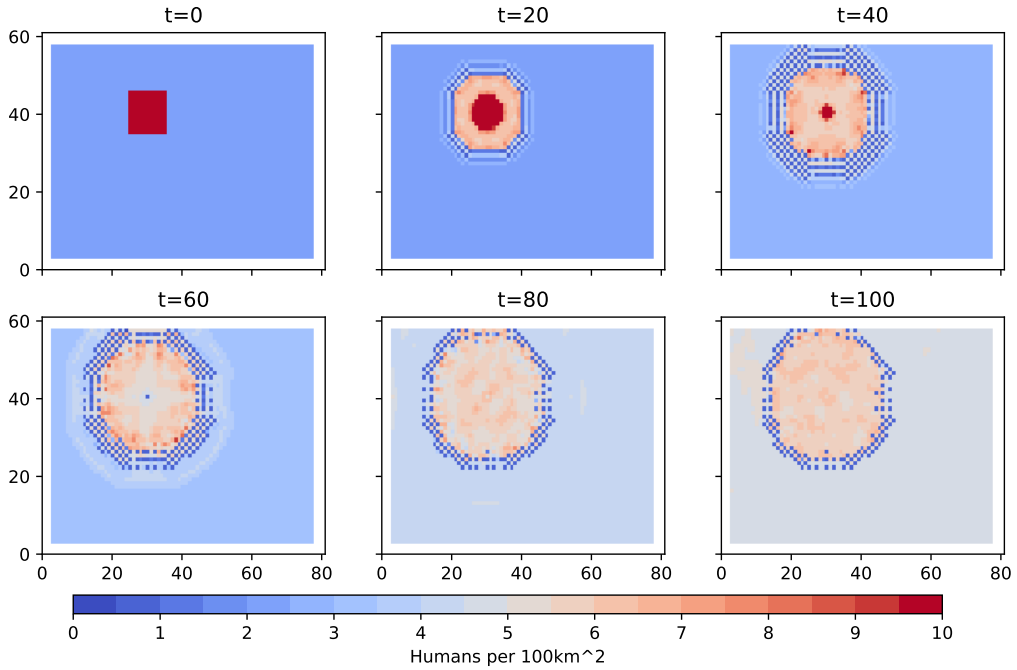

FIGURE 4.6: Snapshots of density distribution for the first 100 years of simulation, experiment 3e.

4.2.4 Experiment 4

The fourth experiment is the last one using the spatial equal HEP. It is nearly identical to the third experiment but features a lower population and a disturbance in the potential field by isolating an area from the rest by a zero potential mask, creating an island (refer to a) and f) in Figure 4.1). This island should not be disturbed by any development outside of it and serves as an indicator of the model's integrity. It is an obstacle that might affect some behavior of the dispersal around it. See Figure 4.27 for some time slices.

Overall, the addition of an isolated obstacle does not alter the results extensively. The pure advection iteration (4a) shows the same patterns as the previous experiment (3a), with a slightly disturbed numerical wave due to the obstacle island and overall lower population in the central region. This behavior can be observed in the advection-diffusion iteration (4d) as well. The pattern does not extend onto the island, showing that migration can not cross over invalid or masked grid points. The diffusion (4b) is weaker as there is less population to delude in the surroundings. The birth-death iteration (4c) is the closest to the previous iteration (3c), showing that populations well above the capacity break down to more stable numbers if they do not disperse before. In the last iteration (4e), the isolated island population makes the mobility waves within populated areas more apparent, as its growth is more uniform than the open region fluctuations.

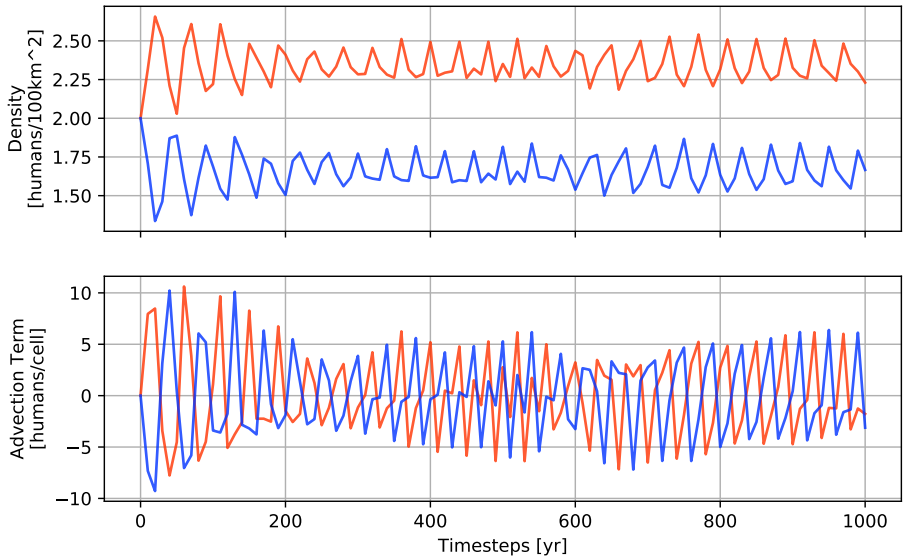

FIGURE 4.7: Time series of density and advection term for two points within the populated area of experiment 5a, one in lower potential (blue, (26,40) and one in higher potential (orange, (34,40)).

4.2.5 Experiment 5

In the fifth experiment the switch to an eastward linear increasing field is made. Again, the simulation starts with an isolated low population in this field (case b) and c) in Figure 4.1), with an overview of the density given in Figure 4.28.

The pure advection case (5a) shows a continuous fluctuation of the population over the simulation time, which does not reach a steady-state or significantly break down over time. As before, the lack of diffusion blocks any further expansion into empty regions. The now inhomogeneous HEP induces velocities to the east at the start, resulting in a slightly higher population to the east compared to the west of the starting distribution. As this wave hits the eastern border, the western parts of the block become more attractive, resulting in a backflow. This pattern repeats indefinitely, creating an oscillating dispersal (Figure 4.7).

The pure diffusion (5b) is nearly identical to the equal HEP case (1b). Only a few grid cells indicate a neglectable faster dilution of the population. The impact of the direction of velocity on the diffusion (Equation 3.24 & 3.44) seems to be insignificant for low population densities.

Looking at the response of the birth-death term (5c), the population adjusts quickly to the different HEP fields. There is not much population growth occurring as the initial region features low HEP of 3-4 humans/100km^2 which is quickly reached (see Figure 4.8).

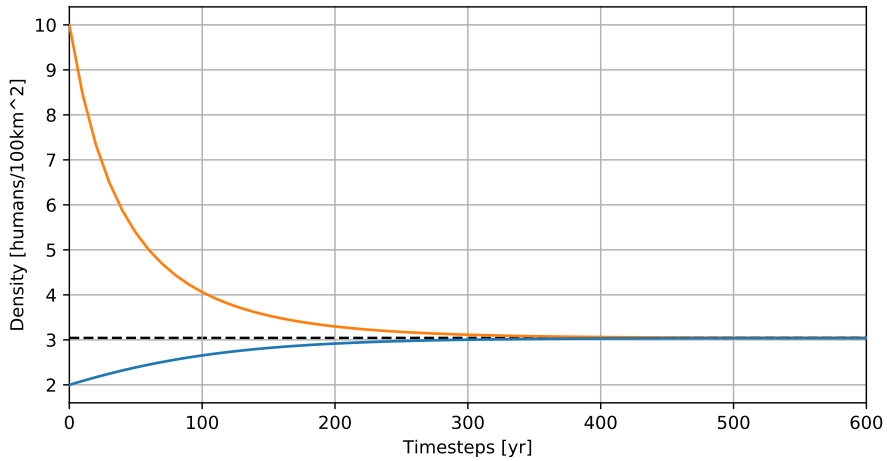

FIGURE 4.8: Population Evolution over simulation time of Experiments 5c (bottom, blue) and 6c (top, orange), with the cultural capacity as the dashed black line. For both time series, the gridcell (40,30) is used.

In the pure dispersal case (5d), a different picture to the homogeneous case (1d) is observable. The clustering of the previous experiment is not visible here, as the lower potential decreases the optimal density, which prevents population clustering. There is some intention of dispersal to be noticed by the dark blue outline of the initial distribution, but the lack of population growth hinders further expansion.

Combing all terms (5e), an expansion towards similar conditions (north and south) and to better conditions (east) occurs. The former is induced by the latter, as previously mentioned, through velocity-powered diffusion. Until the birth-death term gives additional population to accelerate the expansion, only sparse expansions into the western regions (t=400) can be observed, which are later occupied as the eastern dispersal front moves further away. This behavior stabilizes the population in the center, making the western regions attractive enough to explore. The domain is fully covered at t=1700, while a steady state is not reached at t=2000.

4.2.6 Experiment 6

The sixth experiment continues with the linear increasing HEP, utilizing now a high-density initialization (b) and d) in Figure 4.1. See Figure 4.29 for the density distributions during the simulation.

The pure advection case (6a) mirrors the behavior of experiment (5a), albeit with higher population densities. There is, again, a strong oscillation present which does not seem to dampen over time. A steady-state is not reached in the absence of diffusion in this case. The pure diffusion (6b) with an initially increased density shows differences to the homogeneous case (2a), as the dilution is accelerated to the north, south, and east, distorting the shape slightly.

Focusing on the birth-death (6c), the quick reduction of the overpopulated region becomes visible once more. Figure X shows the development for a central grid point for both the lower population (5c) and higher population (6c) cases. The low population can adjust quicker to the low cultural capacity than the high population, despite the much stronger decay forced by the birth-death function.

With advection and diffusion active (6d), a similar result as in (2d) can be observed, with a stronger expansion eastward and an extension of the eastern front to the north and south. As the overall HEP is lower in the starting region, the optimal density is reached quicker, allowing the additional expansion into more favorable regions in the east.

Activating all terms (6e) creates the same distribution as in (6d) within the first 100 time steps, followed by an enhanced dispersal to the east as the population dynamics enable replacements for migrated humans. The domain is covered after 1500 years, with minor fluctuations that mostly even out at 2000 years of simulation time. The importance of the diffusion term for empty regions becomes apparent again, as the advection cannot fill these alone. On the border of the occupied area, a robust opposing velocity from the empty area towards the populated area halts further expansion until enough humans move towards the empty area by diffusion to weaken and finally overturn the velocity to the new region.

4.2.7 Experiment 7

With the seventh experiment, more complex interactions of the different terms and parts of the model become apparent. The linear HEP (b) and a wholly covered domain (e) enable the mixing of different aspects observed in the experiments before (Figure 4.30).

For example, the pure advection (7a) in a linear HEP environment now features the initial clustering of low population in high HEP areas in the east as well as the abandonment of unfavorable regions in the west. The initial high-density region disperses to the east, quickly dissolving in the surrounding population but initiating a bow-shock within the first 300 years. This shock is responsible for the lasting fluctuations, similar to the pure advection experiments before. It also triggers the abandonment of previously occupied regions, mainly parts of the initial high-density region and areas west and northwest. With the absence of diffusion and population dynamics, the fluctuations and trenches formed stay until the end of the simulation.

The diffusion term (7b) continues as in the previous observations (6b) with a more substantial distortion eastwards than before, induced by the surrounding low population. The northern part of the expansion reaches the domain's border, forcing the diffusion to turn west or east, which distorts the shape further to the northeast.

The birth-death (7c) acts as expected and adjusts to the HEP within the first 400 years, while the higher potentials take nearly the entire 1000 years of simulation to fully adjust. In the west,

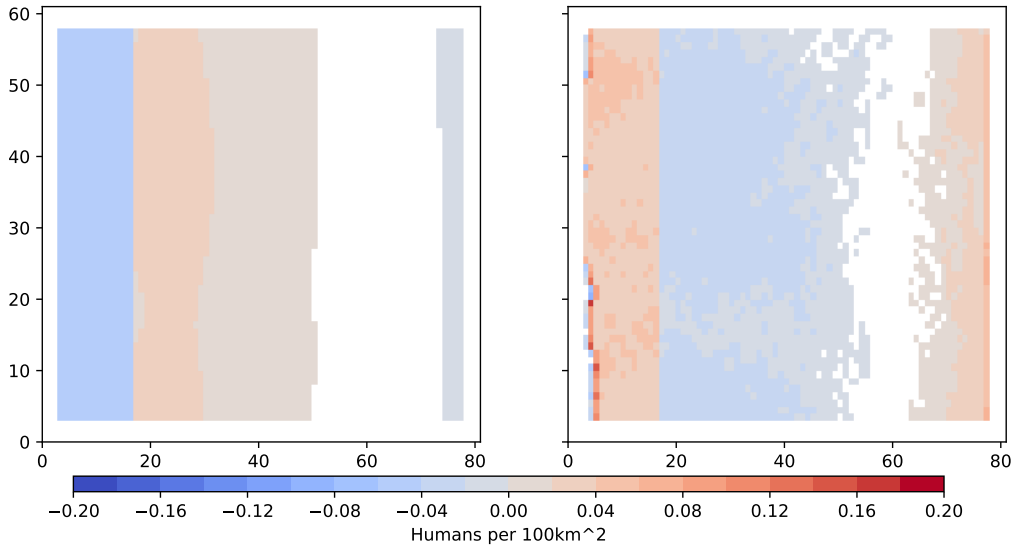

FIGURE 4.9: Accumulation of Birth-Death term (left) and Advection term (right) over ten years of simulation (time step 900 of the simulation, e.g. accumulation of the years 890-900), using density units.

the population diminishes due to the drop below the Reproduction Onset. There is a fine cut here, as no additional population can move in from the east and balance the extinction.

Combining advection and diffusion (7d), there is the same behavior at the start as in the pure advection case (7a). The trenches, however, start to fill up through the diffusion and equalize through the advection in most places. However, to achieve that, most of the western population abandons the region, creating large zones of empty places. The initial shock waves create migration routes that slowly drain the west, visible at t=400 and thin out at t=700. At the end of the simulation, most of these routes are growing short and break down, while there are plenty of high potential regions in the east with low populations left, unable to fill up.

With all terms active (7e), the different aspects overlap each other. The trenches fill up quicker, while the bow shock of the initial distribution quickly dissolves. There is a slight oscillation of the population in the second half of the simulation that dampens with further time steps, as the different terms adjust to each other to form an equilibrium. Note here that the western regions are not sharply abandoned as in the birth-death case (7c). This difference is due to the continuous dispersal into that region from the central area, which continuously replenishes the population that moves out. As the extinction takes time, the migrating humans can reach the western end of the domain before dying. In the east, the death term reduces the population while new humans migrate from the center into the favorable region (Figure 4.9).

4.2.8 Experiment 8

The eighth experiment is the last one of the term-focused ones, adding the isolated island (f) to the linear HEP (b). An overview can be found at the end of this Chapter (Figure 4.31).

As in Experiment 4, only minor changes are visible in the different iterations, mainly induced by the isolated island. The advection term (8a) is directly impacted by the obstacle, resulting in more trenches east of the island in a region that was continuously occupied in the previous experiment (7a). In addition, there are fewer abandoned regions in the northwest, as some of the energy of the initial bow shock is absorbed by the bow shocks created during the formation of the trenches. The diffusion (8b) is mostly the same but shows the same distortion close to the island as the one at the northern border of the domain, showing that the diffusing population has to move along the border. The birth-death (8c) is identical, as expected. Using both advection and diffusion (8e) repeats the same behavior seen in Experiment 7e, with the addition of the new trenches behind the island and different shapes and positions of the dispersal routes. Activating all terms (8e) shows several known behaviors in a combined version and does not differ significantly from the previous iteration (7e). The island does impact the patterns of dispersal of the surrounding domain while maintaining its behavior, ignoring any incoming shocks or induced migration of the large domain.

4.2.9 Experiment 9

9a: Reference Case

The Reference case 9a, as stated before, is quite similar to the experiment 8e, with the only difference being in the slightly higher reproduction onset of 2 Humans/100km^2. It does not have any significant impact on the simulation nor does it alter the results in a meaningful way, allowing the usage of the visualization of 8e in Figure 4.31 for the reference case as well. There are only minor differences in the density distribution, which are negligible.

To test if the discretization of Equation 3.21 works correctly, the total population change in this experiment is calculated with a sum of the birth-death terms $B - D$ over the domain and compare that with the sum of the population change $\frac{\partial \lambda}{\partial t}$ for each time step, see Equation 4.1. The population change includes all terms (advection, diffusion, birth-death), as well as the number of humans leaving the domain (marked as the loss $L(t)$) if the borders are open for outflow, making δ an indicator for the precision of the model.

$$\sum_{j,k}^{N,M} \frac{\partial \lambda}{\partial t}(t) - \left(\sum_{j,k}^{N,M} B(j,k,t) - D(j,k,t) \right) + L(t) = \delta(t) \tag{4.1}$$

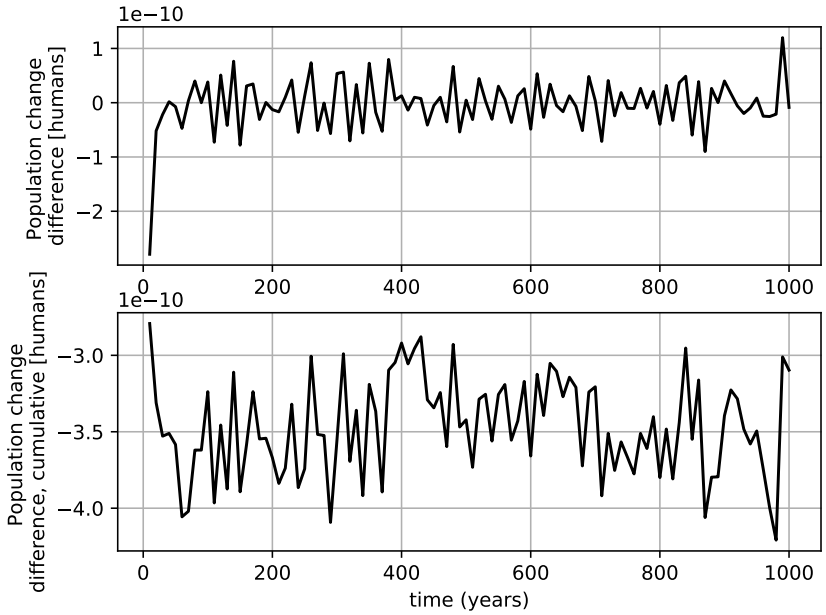

FIGURE 4.10: Difference of the sums of the population change and the population dynamics, see Equation 4.1. The top shows the difference for every time step, the bottom the cumulative difference over time.

Ideally, δ is zero. However, due to the numerical precision of the used discretization and the unavoidable floating-point errors, this number will be tiny compared to typical values of the investigated variable and close to zero. Figure 4.10 shows this difference for Experiment 9a over the simulation time with $\delta(t)$ at the top and the cumulative sum of $\delta(t)$ at the bottom. The values are in the range of 10^{-10} humans, while the cumulative difference is in the same range. If the absolute value is used of differences rather than accounting for the sign, the discrepancy lies in the range of 10^{-8} humans (not shown). To compare, the total population is in the range of 10^5 humans, with individual grid cells being in the range of 0 to 200 humans. The error is, therefore, quite small compared to the values of interest, ranging in the worst case from 10^{-15} for ten-year intervals and 10^{-13} for a simulation of 1000 years.

9b: Advection-Diffusion Fraction

The Advection-Diffusion Fraction ϵ is a parameter in Equation 3.33 that impacts the strength of the diffusion parameter. It is one of the parameters that control the strength of the diffusion term compared to the advection term and impacts the density distribution within the model.

With a high value of 0.9, which indicates that the diffusion is nearly as large as the advection, the model behaves differently (Figure 4.11). First, the trenches that form in the high HEP region are much faster filled with humans than in the reference case. The empty areas that form

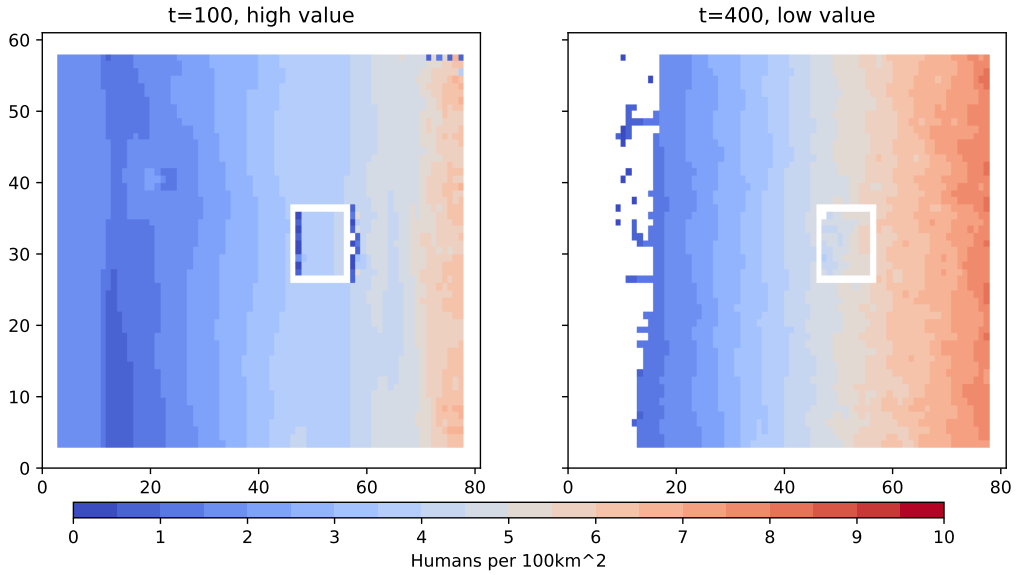

FIGURE 4.11: Snapshots of the human density at t=100 of the high value case (left) and the low value case (right) of Experiment 9b. Compare with 8e as the reference case in Figure 4.31.

in the west after the initial adjustment to the HEP are faster occupied. With a low value of 0.1, the opposite effect is achieved: the trenches are more persistent (not shown), empty areas are longer abandoned, and the overall population dispersal is longer active, leading to more extended oscillating patterns in the density field. In addition, a weak diffusion enhances the impact of the reproduction onset, allowing for larger abandoned areas in the west of the domain due to the lack of diffuse inflow from the established regions. The border between occupied and empty areas is much sharper as well.

Figure 4.12 shows the domain average of the absolute values of the advection and diffusion term, as well as the domain average of the birth-death term. The high-value case has a more robust average diffusion part at the start of the simulation that falls off quicker than the reference case, showing again that the unoccupied region fills up quicker, while the average advection term compensates with lower values overall. In contrast, the lower value case has a near-constant average diffusion that slowly decays over time, mirroring the longer time needed to fill empty regions. The average advection is lower in the first 150 years of simulation but higher after that due to the longer time needed to fill the gaps and abandon the western regions. The birth-death term is less affected by the change, only indicating that the high-value case has a slightly faster growth and the low-value case a slightly slower growth than the reference.

9c: Maximum Velocity

The Maximum Velocity v_{max} is an important parameter (Equation 3.29) that determines how fast humans can disperse from one point to the next. It is crucial for the numeric stability of

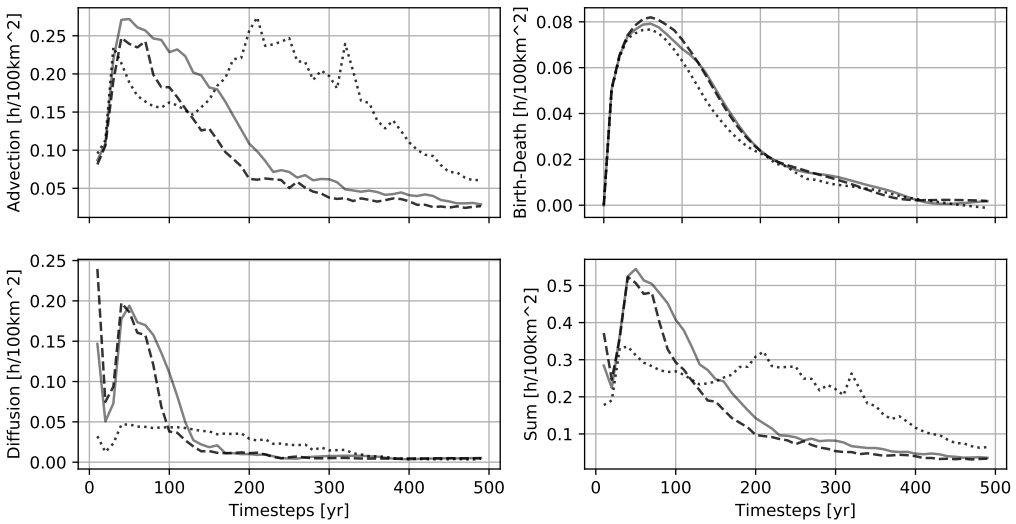

FIGURE 4.12: Domain averages over time of Experiment 9b for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

the model as well (Equation 3.47), serving as the upper border of achievable dispersal speed. It affects the strength of the advection and diffusion term.

The higher maximum velocity requires more numerical time steps to ensure the stability of the model. The same behavior but accelerated can be seen, including the bow shocks and the forming of trenches in the east. The former travel faster through the domain and take more time to be diminished. The latter are filled up in less time (Figure 4.13). The higher speed allows the western part of the population to abandon the region to fill up the higher potential areas in the east until the population growth is strong enough to reoccupy these regions later. By reducing the maximum velocity, both advection and diffusion respond much slower to the different HEPs. Due to this, the clustering of the population does not occur so that no trenches form. The initial bow shock is weaker and disperses much quicker as well.

By increasing the maximum velocity, faster exchange between the grid points is achieved, allowing humans to move faster from one place to the next. This results in a more significant average advection term (Figure 4.14) as the dispersal is accelerated. The diffusion term, which is dependent on the direction and speed of the velocity, also has a more substantial impact. Both terms dominate over the birth-death term, diminishing its influence in the dynamics of this simulation. The peak in the sum reflects the rapid dispersal at the start of the simulation, while the reoccupation of the abandoned western part of the domain is visible as a slightly higher value than the reference between the year 150 to 300. A lower maximum velocity lets the population dynamics excel in the budget, creating the more diffusive and slower adjustment to the potential over time.

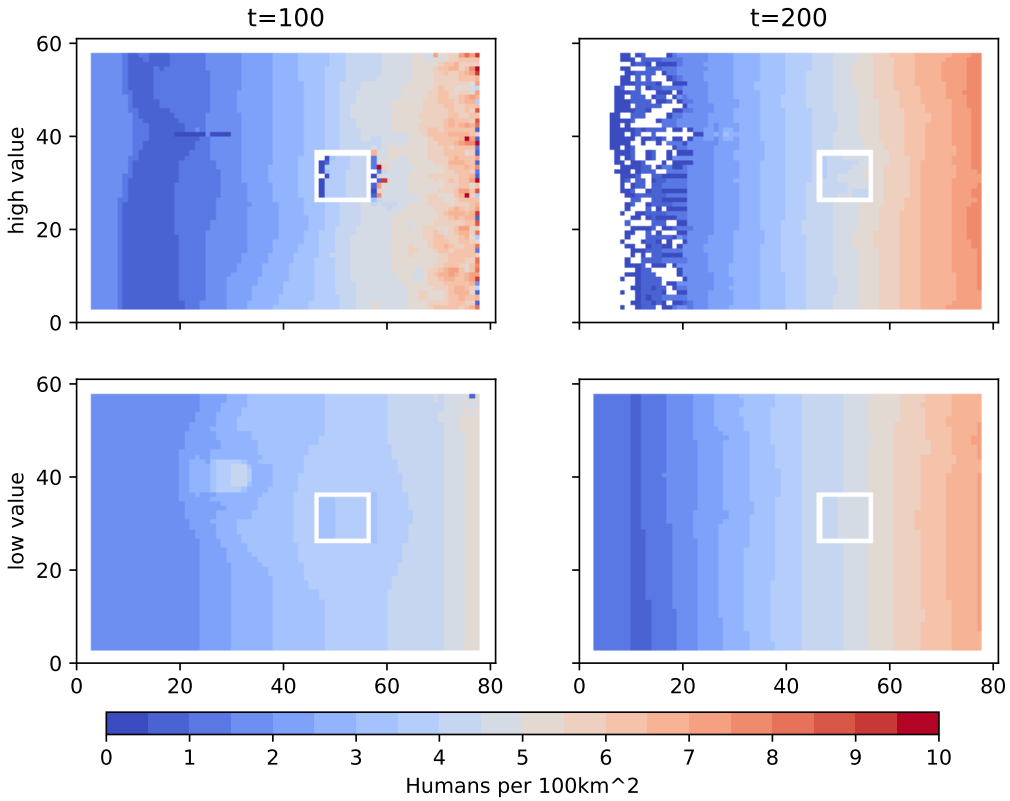

FIGURE 4.13: Snapshot of the human density at two times of the high value case (top) and the low value case (bottem) of Experiment 9c. Compare with 8e as the reference case in Figure 4.31.

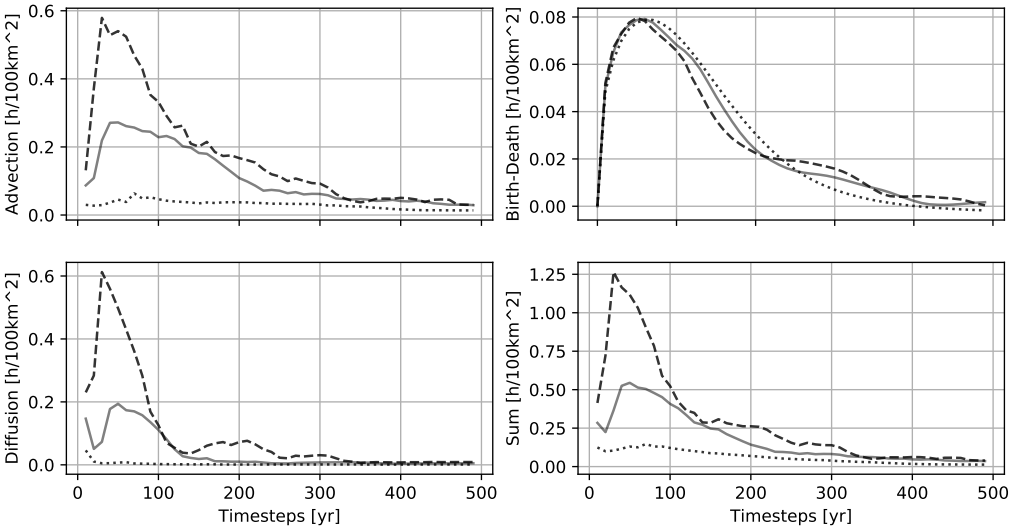

FIGURE 4.14: Domain averages over time of Experiment 9c for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

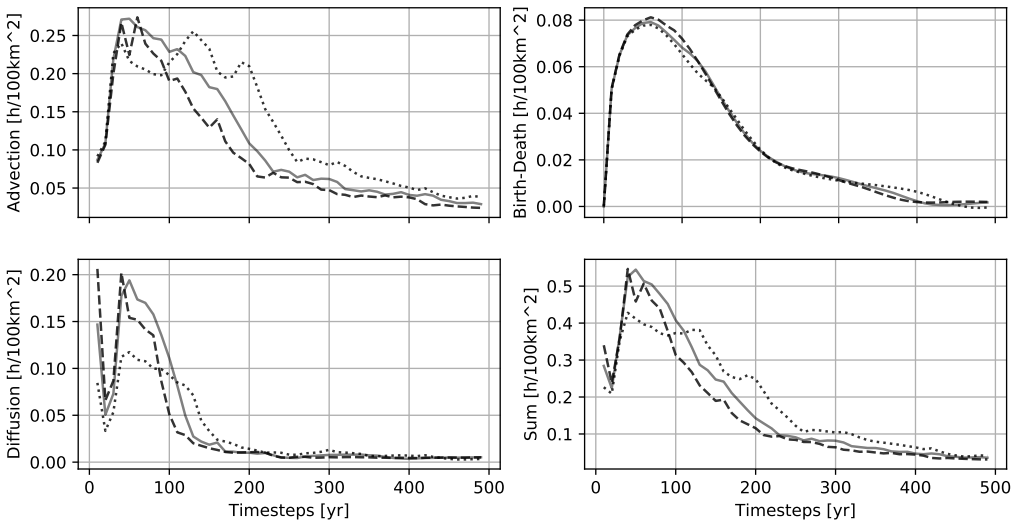

FIGURE 4.15: Domain averages over time of Experiment 9d for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

9d: Critical Velocity

The Critical Velocity v^{crt} gives the minimal velocity used to drive the diffusion term (Equation 3.33). It determines how strong the diffusion acts for cases of low velocity, ensuring that at all times, at least some exchange between the occupied regions can occur.

Figure 4.15 shows the different time series of each term. Increasing the critical velocity gives the diffusion a jump start at the beginning of the simulation, which counteracts the clustering and formation of trenches by quickly filling these regions to allow for more inflow from the adjacent areas. This effect allows the birth-death to act a bit sooner, creating a slightly higher population growth compared to the reference case. Both advection and diffusion diminish a bit quicker over the simulation time, as most adjustments by dispersal are achieved sooner. A lower critical velocity shows the reversed effect: a lower diffusion at the start, allowing more clusters that take longer to fill up. The decay of the dispersal terms takes more time as the exchange is spread out. The advection term shows higher values overall, mainly due to the large number of shocks induced by the numerous clusters that fill up. The impact of a change in this parameter compared to the maximum velocity is much lower, but it does significantly affect the shape of the dynamics.

9e: Flux-Limiter

The parameter ν is a model parameter within the minmod flux-limiter used in the discretization of the model (Equation 3.42). It regulates how sharp or dissipative shocks and high gradients

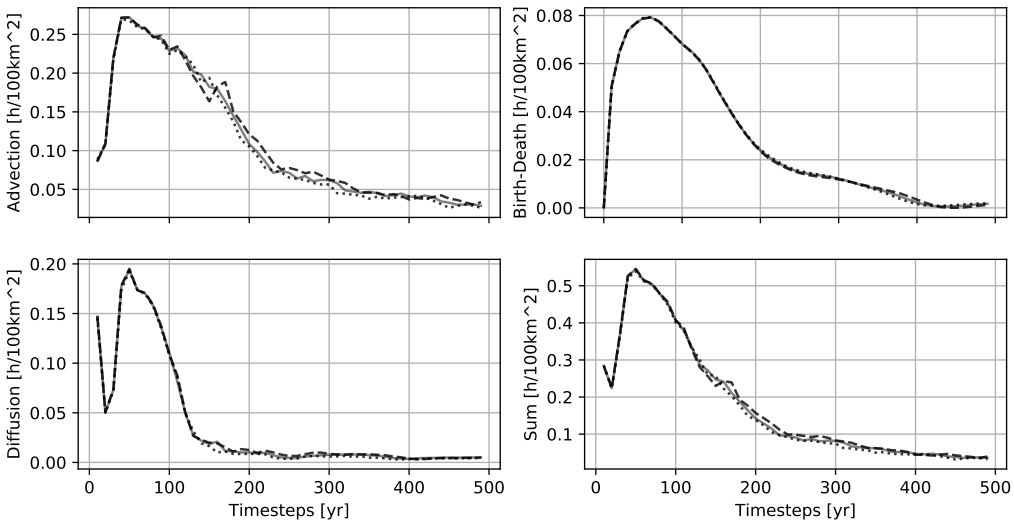

FIGURE 4.16: Domain averages over time of Experiment 9e for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

of the population density are captured.

Using a high value is less dissipative, resembling shocks more precisely and creating less smooth results. A low value has the opposite effect, allowing more dissipation and less sharp representations of discontinuities. Within this experiment, a direct impact on the domain is only occasionally observable and relatively limited. Figure 4.16 shows that the differences in the average advection term are pretty low. The high-value case tends to have a slightly stronger derivation from the reference than the low-value case, but both are quite similar. Both diffusion and birth-death are near identical with the reference case. The spatial representation of the shocks does not indicate a strong influence of this parameter on the distribution of the population density in the large domain. However, the isolated island shows different patterns, as the local shocks are more often reflected on the border than in the large domain. Despite these differences, the population stabilizes around the same time in all three cases. This parameter does not have a substantial impact on the model compared to the previous.

9f: Growth Rate

The Growth Rate r_B is, together with the carrying capacity Φ_{max}, the main parameter that determines the population dynamics of the model (Equation 3.48). It affects both the growth and decay of the population, as long as it is above the reproduction onset.

The reproduction rate is a crucial model parameter with a high impact on the results, as shown in the time series of Figure 4.17. Increasing it from 1.5 % to 5 % leads to a simulation driven

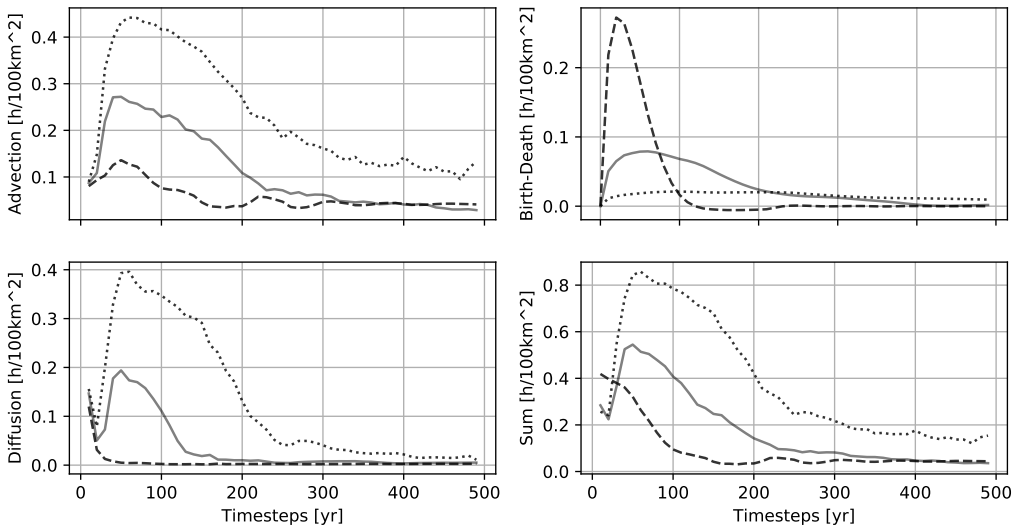

FIGURE 4.17: Domain averages over time of Experiment 9f for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

mainly by the population dynamics with much lower dispersal terms as secondary drivers than the reference case. Significantly the diffusion quickly shrinks down, while the advection, albeit weak, does contribute to the budget. The strong growth creates a slight overshoot of the population that is regulated down later, indicated by the negative values of the birth-death budget between 100-220 years of simulation. A low parameter of only 0.5 % enhances the dispersal dynamics of the model, creating more substantial advection and diffusion terms with a prolonged and gradual increase of the population.

Adjusting the parameter affects the balance of the different terms, altering the shape and patterns of dispersal visible in the population density. Figure 4.18 shows the impact of this. Increasing the growth prevents the clustering in high potential areas visible in previous experiments, in a comparable magnitude to the maximum velocity. Reducing the growth enhances the clustering effect, creating even more abandoned or empty areas that take longer to fill up or be reoccupied as the population grows much slower.

9g: Reaction Time

The Reaction Time $\Delta\tau$ describes how fast the humans can adjust their dispersal velocity and change their direction of travel, representing an inertness of the moving population (Equation 3.29). This parameter directly impacts the advection term and implicitly the strength of the diffusion term.

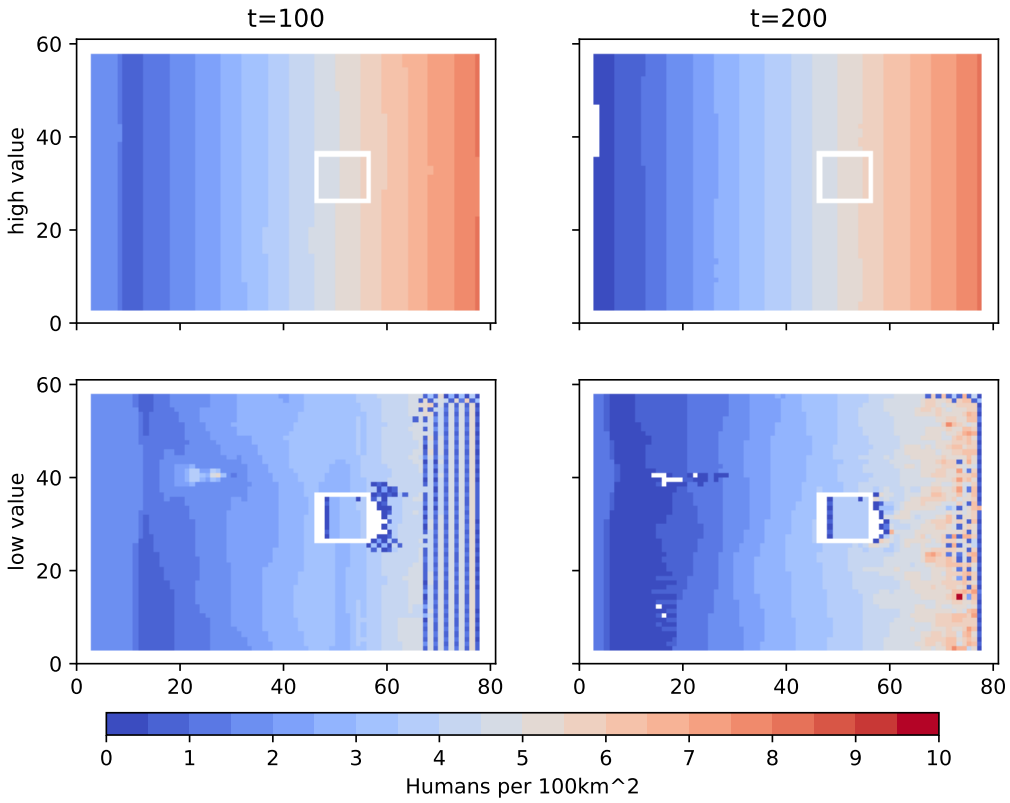

FIGURE 4.18: Snapshot of the human density at two times of the high value case (top) and the low value case (bottem) of Experiment 9f. Compare with 8e as the reference case in Figure 4.31.

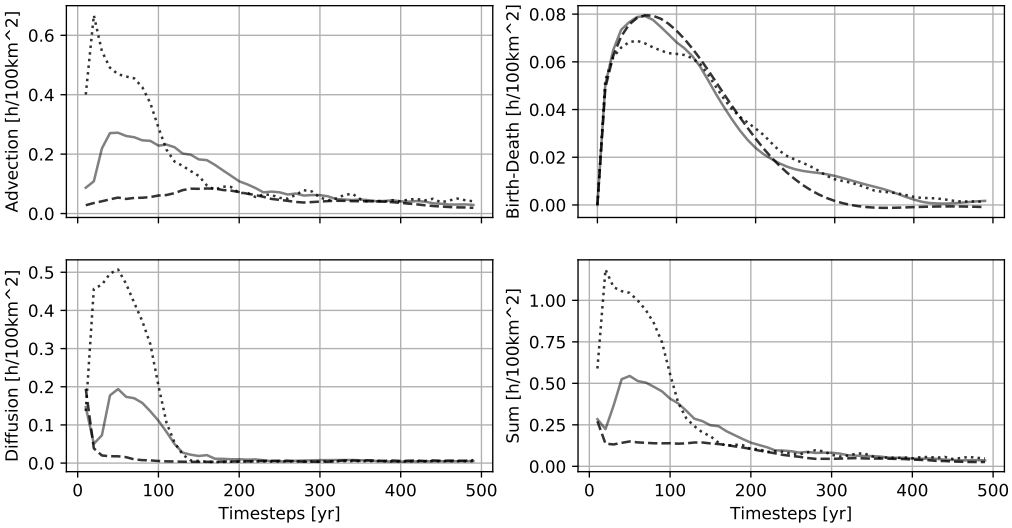

FIGURE 4.19: Domain averages over time of Experiment 9g for the absolute advection term (top left), the absolute diffusion term (bottom left), the birth-death term (top right), and the sum of all three terms (bottom right). Includes the reference case (gray, solid), the high value case (dashed) and the low value case (dotted).

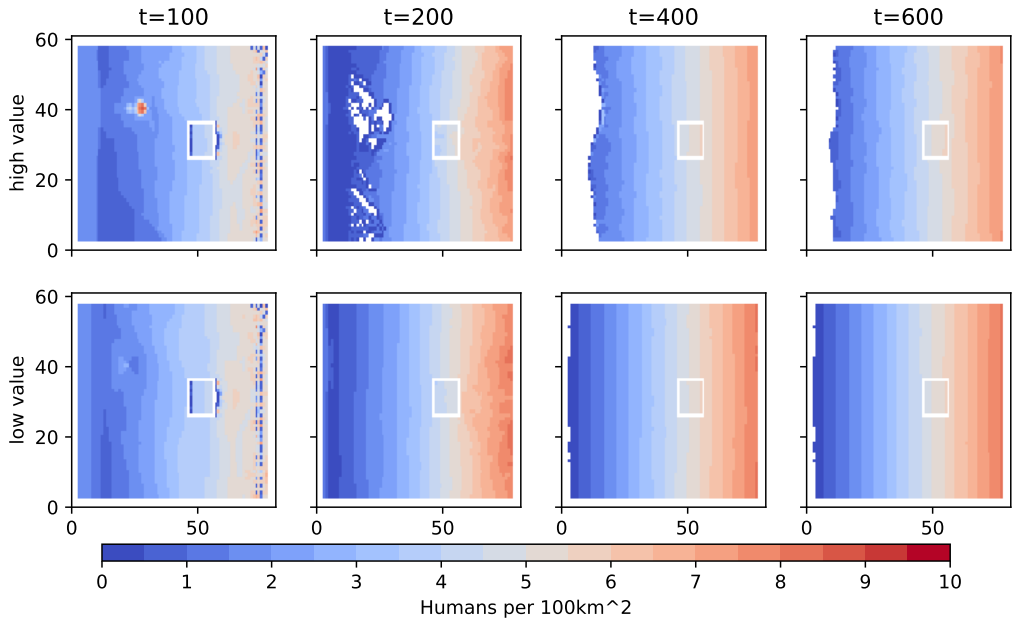

FIGURE 4.20: Snapshot of the human density at different times of the high value case (top) and the low value case (bottem) of Experiment 9h. Compare with 8e as the reference case in Figure 4.31.

This parameter has a similar impact to the model as the maximum velocity and the reproduction rate, but the effect is inverse again (Figure 4.19). A higher value represents additional time is needed to adjust the velocity, both in initial acceleration and later change of direction and magnitude. It does affect the diffusion as well, but not in the initial phase of the motion, as the high velocities trigger a strong diffusion to diminish the gradients. However, the diffusion is quickly overturned by the population dynamics, which take up the role as the main driver of the model. It reduces the amount of clustering again and gives a smoother change of density over time. Using a lower parameter, resembling a faster reaction to new circumstances, allows for a rapid increase of dispersal velocity, making both advection and diffusion stronger. The clusters are more prominent in this case again but diminish at a similar time than the reference case. The difference to the other parameters is that the magnitude of the initial dispersal changes drastically, but the breakdown of all three curves is quite similar to each other and only slightly skewed in time.

9h: Reproduction Onset

The Reproduction Onset κ regulates at which population density the birth-death term usually activates. Below the threshold, a continuous decay of the population is assigned in the model (Section 3.4.3). It represents the lower boundary of a stable population that can maintain its size.

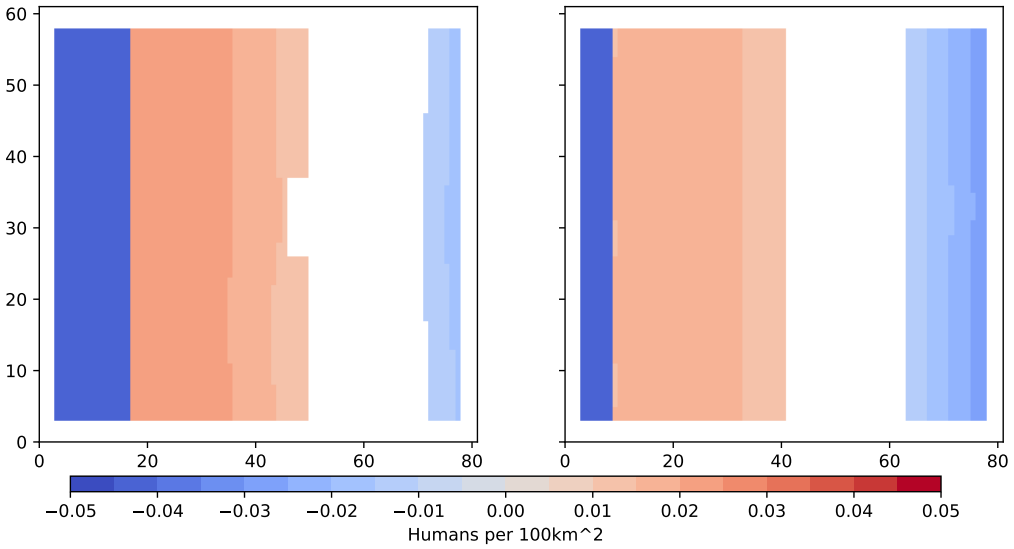

FIGURE 4.21: Accumulation of Birth-Death term for the reference case (left) and the low reproduction onset case 9h (right) over ten years of simulation (time step 900 of the simulation, e.g. accumulation of the years 890-900), using density units.

Figure 4.20 gives an overview of the density evolution for high and low values of this parameter. The impact on the dispersal patterns is quite limited, as the clustering occurs normally in both cases and diminishes simultaneously. Only the initial high-density distribution, which lies below the high parameter threshold, seems to be affected. It is, however, not the case, as the slower decay of this population can be derived from the birth-death term. The initial population is much higher than the potential, tipping the birth-death term into negative values to lower the population closer to the potential. This part of the equation is not active in the high-value case, as the high population lies in an area below the threshold. It does decay slowly over time, allowing the population to move rather than die out straight away. In addition, the western part of the domain is briefly abandoned and later slowly reoccupied, as the eastern population continuously grows and stimulates the population to expand back westwards.

The lower value does not defer significantly density-wise from the reference, but the western area has a larger area with an increasing population and the eastern area a stronger death rate. Both areas extend further west than in the reference case (Figure 4.21).

9i: Spatial Resolution

The Spatial Resolution of the model is often subject to the resolution of the HEP. Higher resolutions might affect the model's behavior, creating spatial enhanced representations of the dispersal by incorporating more details or changing the results significantly due to the new information.

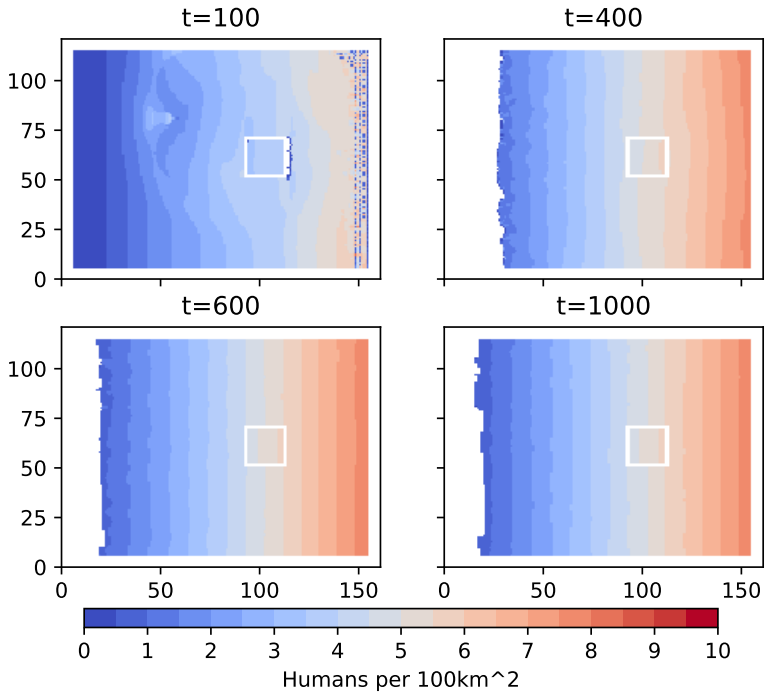

FIGURE 4.22: Snapshot of the human density at different times of Experiment 9i in a similar style to the reference case (8e, Figure 4.31).

For this experiment, the same HEP distribution and initial population are used as before, but with a resolution of 25 km, effectively quadrupling the model's size in space (now 160x120 grid cells). The numerical time step is further refined to 0.01 yr to ensure the stability of the model. Figure 4.22 gives an overview of the density distribution similar to the reference case. The overall behavior of the model shows the same effects, but the higher resolution affects their visualization. The bow shock by the dense population dispersing is sharper represented compared to the lower resolution case. More and smaller trenches are forming in the east, which is no surprise due to the clustering being directly tied to the grid cells numerics. The reproduction onset cutoff is much clearer in the west, and it takes more time for the population to push into this region. The death term is absent in the east at the end of the simulation, indicating that the expansion westwards is not finished yet. A steady-state solution is not achieved in the thousand years of the simulation time. It seems that the higher resolution is more affected by the reproduction onset and does not extend as far to the west as the lower resolution. It might be due to a slightly slower dispersal or a different adjustment to the HEP, which limits the extension westwards.

9j: Time Resolution

This experiment increases the numerical time step to 0.01 yr but keeps the exact spatial resolution to see if that affects the model as drastically as the resolution does. It does not seem to be

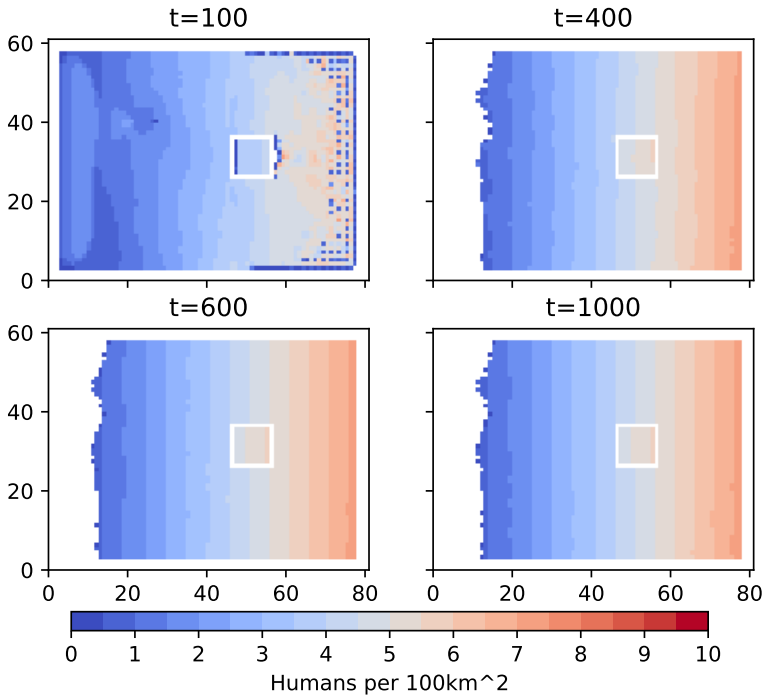

FIGURE 4.23: Snapshot of the human density at different times of Experiment 9k in a similar style to the reference case (8e, Figure 4.31).

the case, as the model shows the same behavior and takes the same time to close on a steady-state solution. The only differences that can be observed are minimal (0.1 Humans/100km^2) and are likely due to the accumulation of floating-point errors and the slightly different propagation of shock waves at higher temporal resolutions.

9k: Open borders

The model includes the possibility of outflow at the domain's border, acting as an artificial sink. Open borders are included mainly for real-world applications that focus on a particular region or continent. In this experiment, all four sides of the domain are open, effectively removing the population at every time step from the outer grid cells.

Figure 4.23 gives an overview of the density evolution over time. Noticeable are the increased number of trenches and the clustering close to the borders, which take more time to fill up. The initial population shrinks at the start of the simulation, which is not the case in the reference simulation, as many humans move out of the domain. The two outer grid cell rows and columns are primarily affected by that, mainly in the eastern half of the domain, but are reoccupied two hundred years into the simulation. The population starts to stabilize after that, but a large portion of the west is abandoned and not reoccupied until then. The population continues to move out mainly by diffusion, as the population is taken out of the system before the next time step can calculate a velocity facing into the border region. However, the advection

does fuel the layer that provides the border with the migrating population. The impact of the border reduces gradually further away, so it is recommended for a realistic application to have several rows of grid cells between the area of investigation and the border of the domain.

9l: Gaussian Start

The last experiment switches out the block of high population with a Gaussian-shaped distribution to see if such a starting condition can reduce the bow shocks of the initial movement. It is indeed the case. The initial shock is reduced in the first tens of years of the simulation. However, this does not affect the results in the long run due to stronger and more numerous shocks induced by the clustering in the east of the domain. The Gaussian distribution is only slightly faster dissolved, mainly having a local impact on the distribution.

4.3 Discussion

This chapter aimed at the evaluation of the dispersal model, focusing on its numerical consistency and the assurance that the model behaves as designed and expected. For that, a series of nine idealized experiments were conducted that aim at the different terms of Equation 3.21 and the parameters introduced during the numerical conversion (see Section 3.4.3).

The advection term is an essential addition to the established diffusion-reaction equations used for dispersal models. It should not be active if there is no gradient in the available potential between two grid points, which can be observed in Experiments 1a and 2a. If there is a gradient, the advection term is activated through the induced velocity, and the population is encouraged to balance out the difference through dispersal. As used in Experiment 3a and 4a, a significant gradient can create a bow shock that propagates through the domain, provided there is a population that can carry on the shock. If this is the case, in the pure advection the clustering of the population can be seen in lines or a checkerboard pattern, depending on the origin and distance of the wave. This behavior results from using the Weibull curve for the AvHEP, which makes it more favorable for lower population numbers to merge in one place. It requires an initial motion such as the bow shock to be triggered, as there is no dispersal in a homogeneous distribution of both potential and population density.

Without the surrounding background of low population density and only the advection term active, an oscillating pattern arises in the density, as seen in Experiment 5a and 6a. As the eastward-moving population stacks up at the highest HEP, this region becomes unfavorable through the available HEP, resulting in a backflow. The same happens at the western border once the population moved far enough back, resulting in continuous oscillation. The fact that the oscillations do not decay over time shows the absence of numerical diffusion within the discretization of the advection term, at least for equidistant grids. Obstacles can affect the patterns

of advection, as the 8th experiment shows. The clustering behind the obstacle increases as the population moves east without a replacement from the west. The HEP forces the population to stick together instead of moving far, resulting in the increased clustering.

The parameters that influence the advection are the maximum velocity, the reaction time, and the flux limiter. While the latter's impact is relatively small, the former two greatly influence the quantity of the advection term. The maximum velocity directly affects the diffusion term as well when the critical velocity is passed, while the reaction time affects the diffusion indirectly due to the changing velocity field.

The diffusion term compliments the advection term by filling previous empty grid cells with a low number of humans over time, increasing the available potential and allowing more humans to move with the advection term. On its own, the term results in a prolonged process to disperse into other regions. The different Experiments 1-8b show that the diffusion works as designed: The term is always active even in the absence of HEP gradients due to the critical velocity and acts stronger perpendicular to the velocity if the critical velocity is exceeded.

For settling in empty terrain, the diffusion is the main force to move initially, so an expansion looks entirely driven by diffusion. It is not the case in this model, as the diffusion is enhanced both by higher velocities and the advection term itself, which transports the population to the new frontline. The advection increases the diffusion part of the equation while bringing the population over the reproduction onset, activating the birth term to accelerate population growth at the front further. Combined with the AvHEP, advection and diffusion allow the different push and pull factors to enable or halt the expansion. The diffusion is affected secondarily by the advection-focused parameters, as mentioned above. However, the advection-diffusion fraction and critical velocity affect the diffusion directly, with the former having the more decisive impact on the model. The advection adjusts to different strengths of the diffusion accordingly, usually by having increased values for more extended time frames if the diffusion is weaker and prolonged over time.

The population dynamics are essential for dispersal. On their own, the dynamics regulate the population towards the carrying capacity provided by the potential (see Experiments 1-8c). Without them, the population might only expand briefly before reaching the optimal distribution and stay at their location if the HEP does not change over time. It depends on the size of the population, with larger population numbers forced by more complex HEPs taking more time to adjust than smaller populations with simpler HEPs. Experiments 1-8d targeted these cases. At stagnant conditions, the only force to trigger further expansion is a form of population pressure by overcrowding the settled areas. The main parameters for the birth-death term are the growth rate and the reproduction onset. The growth rate drastically changes the dynamics for all terms, accelerating or slowing the adjustment to the HEP primarily and influencing the speed of dispersal secondarily. A higher reproduction rate typically increases the speed at which the human population can spread, while a lower rate slows it down.

The different terms, in combination, unify the different processes, showing patterns of their behavior in unison or create entirely new patterns through their interactions. The clustering forced by the advection, the spread into abandoned areas by the diffusion, and the population adjusting to the HEP through the birth-death term are noticeable all at once. In addition, the exchange between these terms creates new patterns: The clusters forced by the advection fill up through the diffusion over time, and the reproduction onset can be overcome by the diffusion and advection, which provide the unfavorable regions with a new population from higher HEP regions.

The scale of the model has a direct impact on the results presented. Increasing the spatial resolution allows for nuanced representations of the human dispersal, including more details of the HEP field, provided that the HEP can be created with a higher resolution without interpolation. However, the results can be quite different from the lower resolution case, either due to the better-resolved processes or due to different interactions between the model parts that are affected by the resolution. The temporal resolution does not seem to affect the model in a significant way, as long as it fulfills the stability criteria. Allowing the outflow of the model has a drastic impact on the regions close to the borders, as these regions have an increased birth term to compensate for the continuous loss of population at the border. The further a grid cell is from the border of the model, the less impact can be observed. The shape of the starting distribution can have local effects during the first time steps of the model but does not influence the model results in the long run. A Gaussian shape does reduce the initial bow shock, making it preferable for practical applications compared to sharp gradients of the population. The numerical error shows a reasonable size for the idealized simulations. Even in "worst cases", in which errors stack instead of canceling each other, they do not seem to reach a size that could impact the results in a significant way. This is true for equidistant grids that are used in the idealized simulation. The application on the distorted latitude-longitude grid might result in more significant errors. This behavior will be investigated further in the following chapters.

The presented Human Dispersal Model works as intended and designed. It combines the directed dispersal to better conditions with the indirect slow spread around occupied regions and simplified population dynamics. The numerical formulation of the analytical problem is precise and stable, as long as the required conditions are met. The introduced parameters differ in their impact on the model, both in strength and distribution among the different terms.

4.4 Overview Figures

This section contains the graphs depicting the evolution of human density over time in the idealized experiments. Each figure features the five iterations of each experiment, showing the time steps for 100, 400, 700, and 1000 years after starting the simulation. An overview of the starting distributions and the underlying HEP can be found in Section 4.1, Figure 4.1.

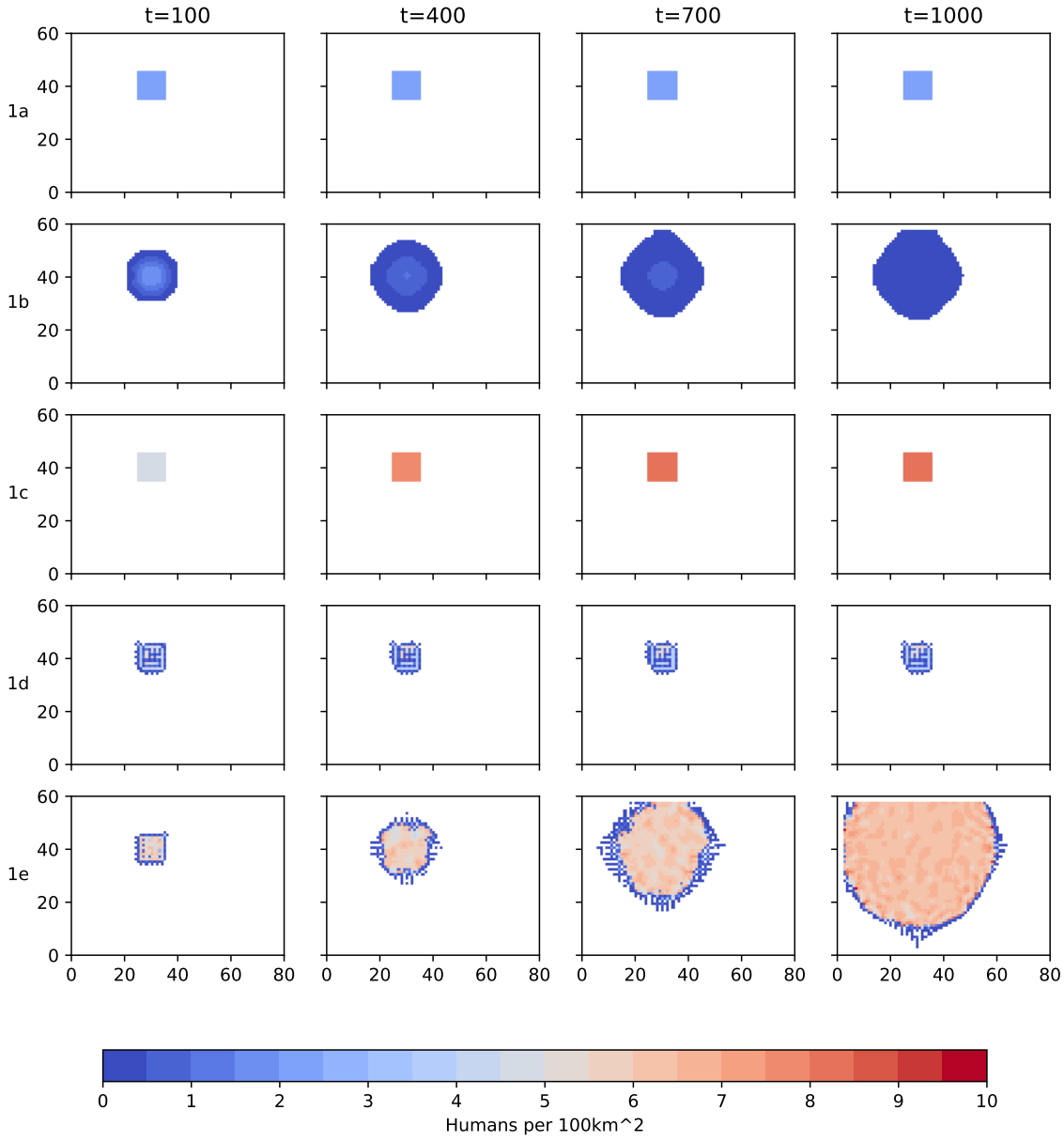

FIGURE 4.24: Human Density for each iteration (a-e) and different time steps of Experiment 1. The axes depict the grid size.

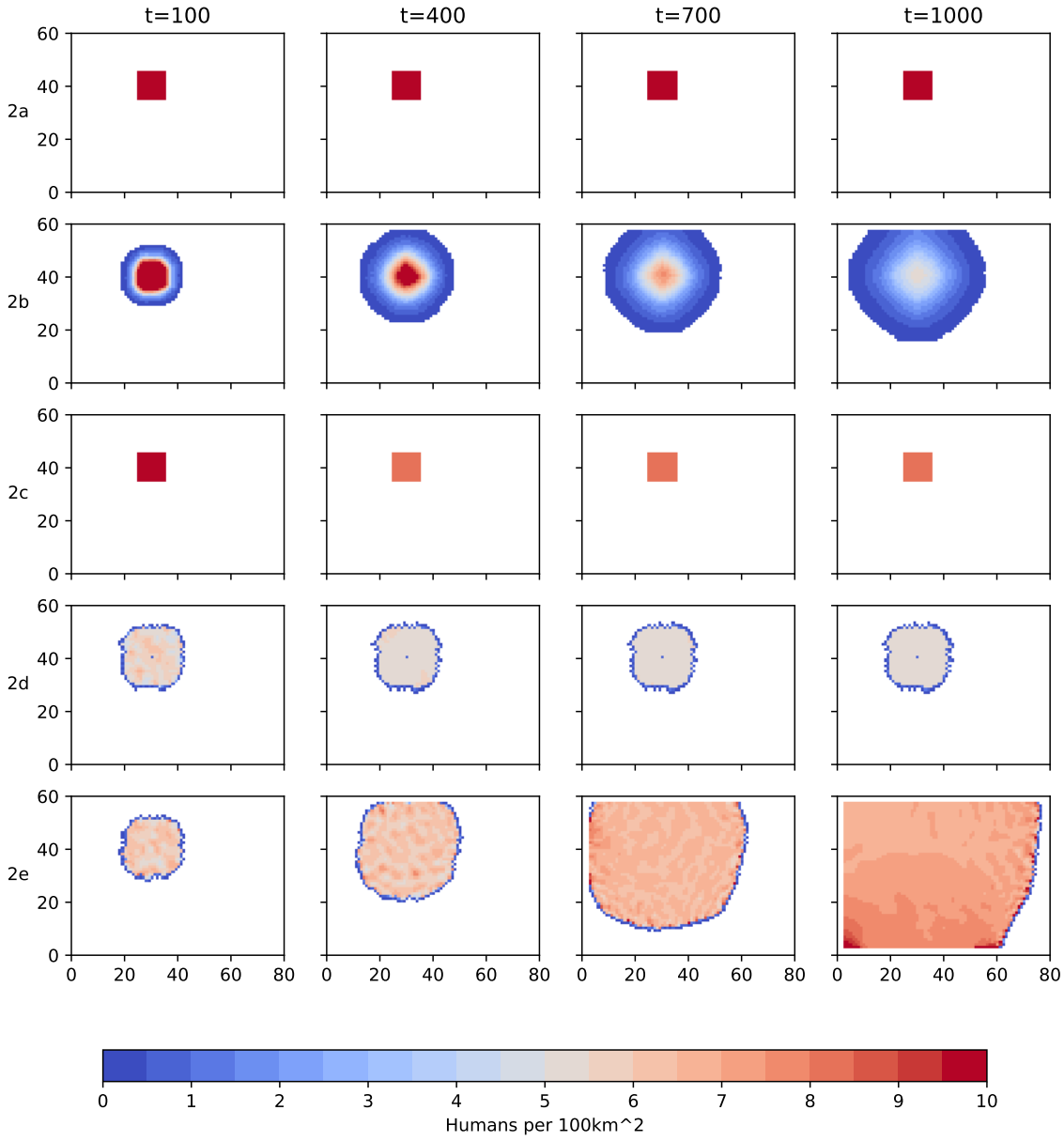

FIGURE 4.25: Human Density for each iteration (a-e) and different time steps of Experiment 2. The axes depict the grid size.

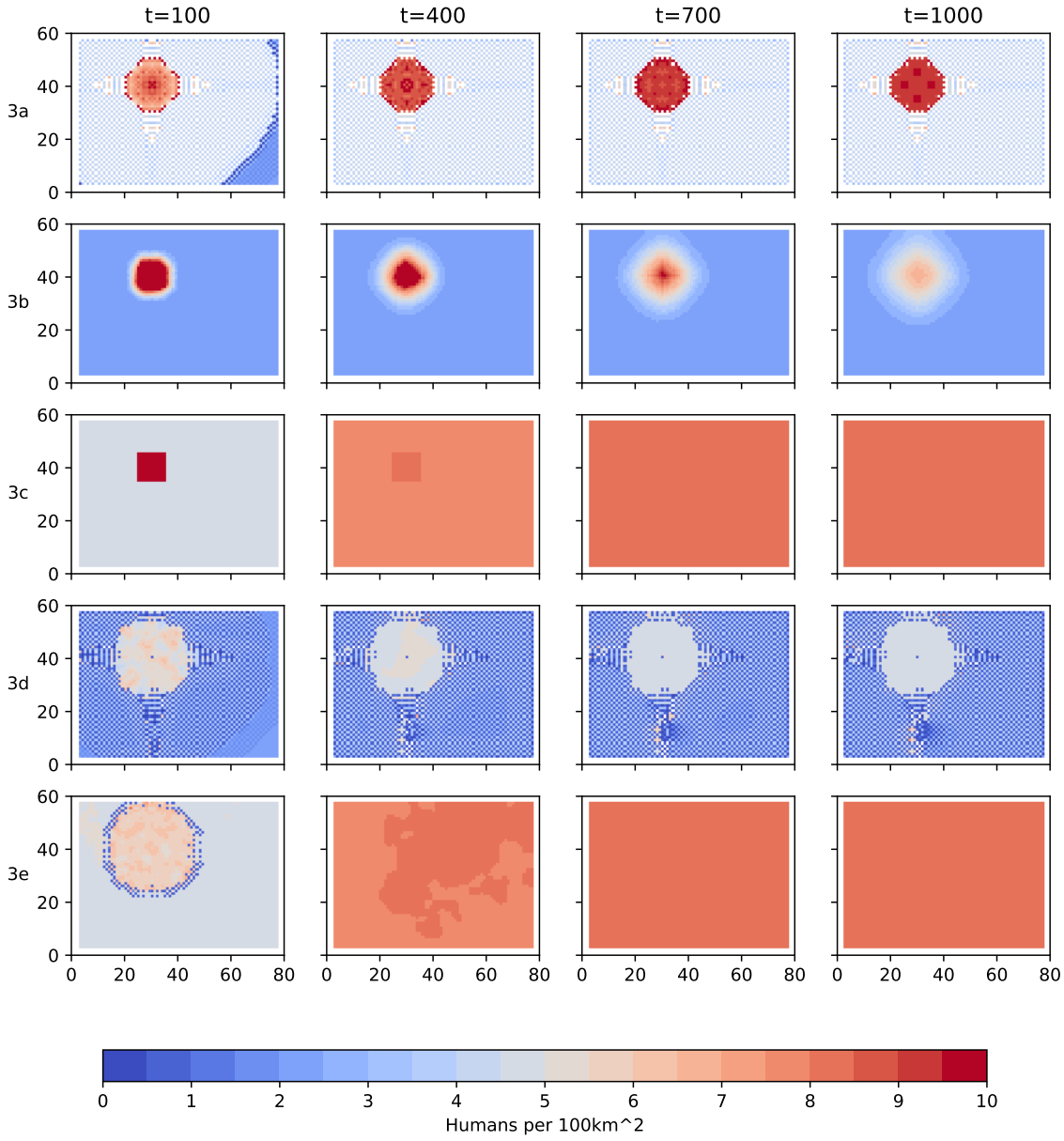

FIGURE 4.26: Human Density for each iteration (a-e) and different time steps of Experiment 3. The axes depict the grid size.

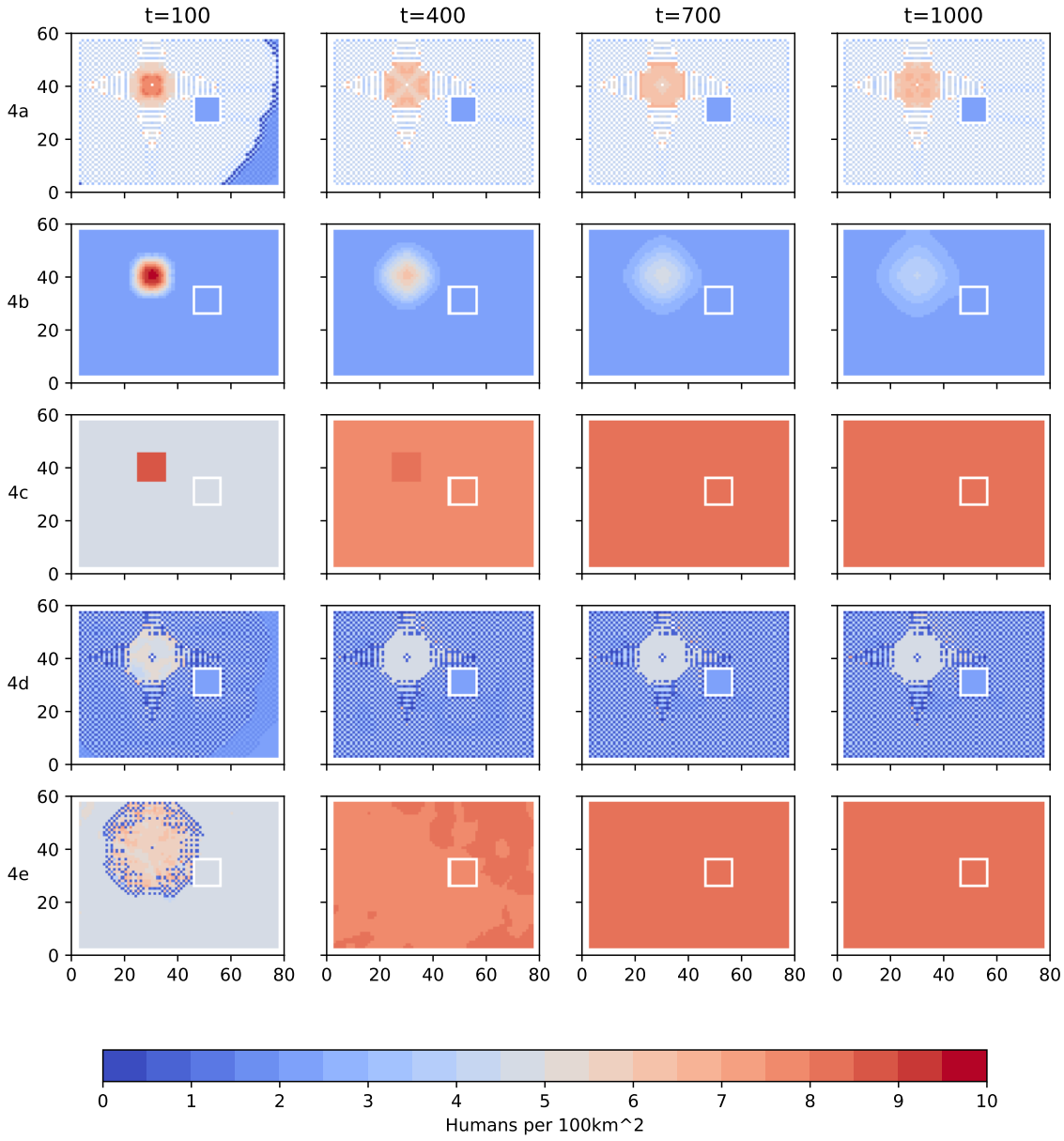

FIGURE 4.27: Human Density for each iteration (a-e) and different time steps of Experiment 4. The axes depict the grid size.

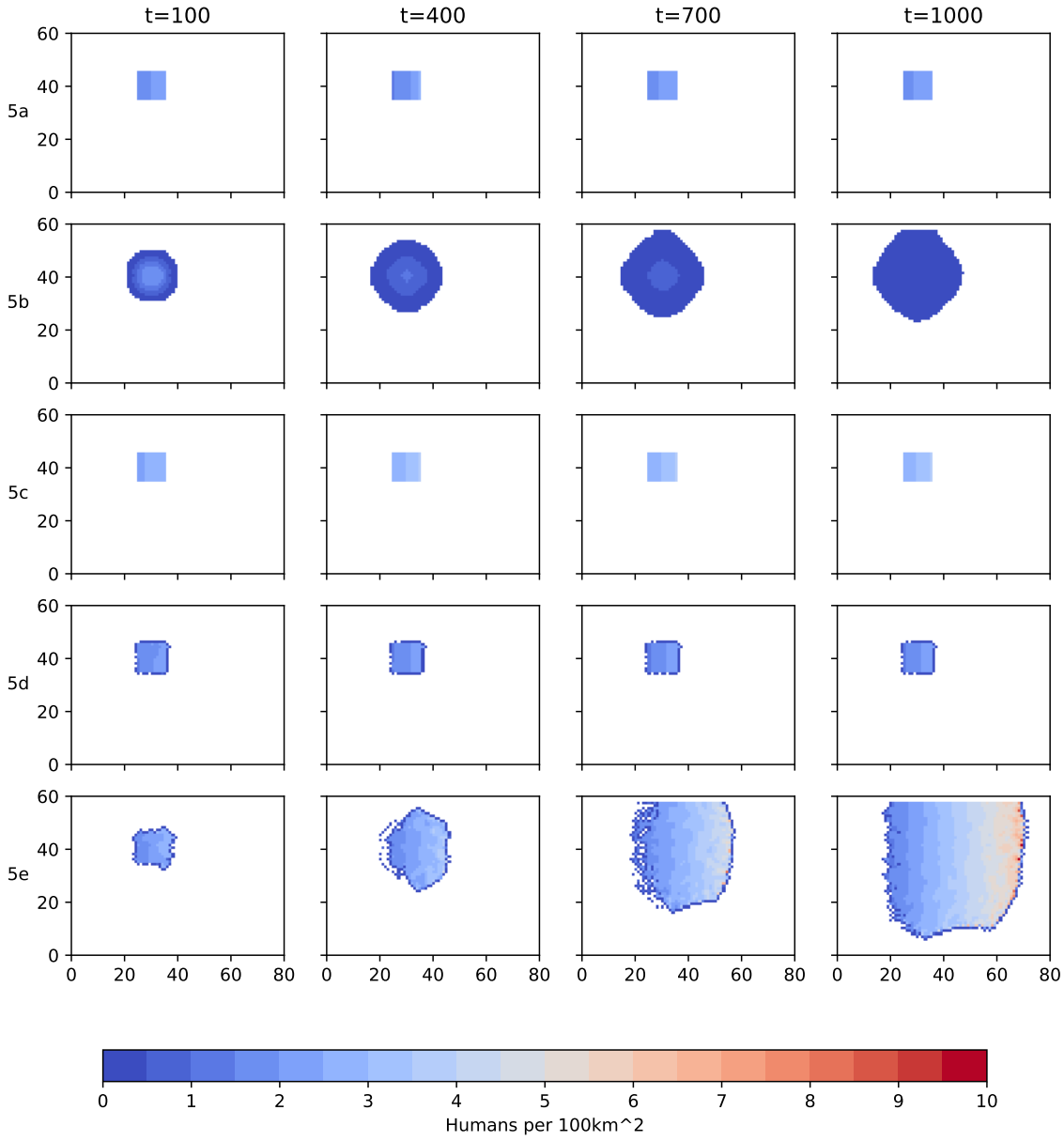

FIGURE 4.28: Human Density for each iteration (a-e) and different time steps of Experiment 5. The axes depict the grid size.

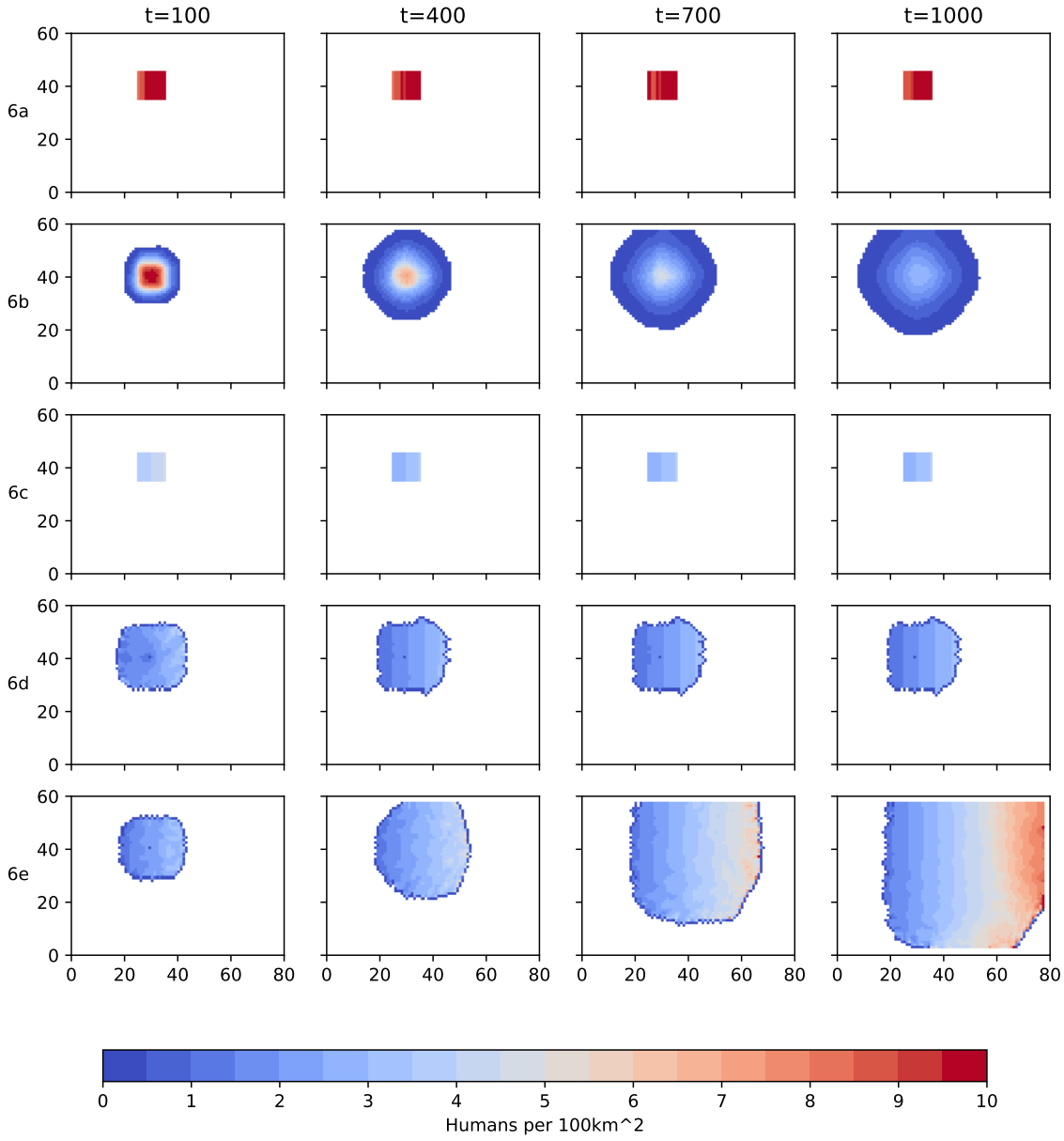

FIGURE 4.29: Human Density for each iteration (a-e) and different time steps of Experiment 6. The axes depict the grid size.

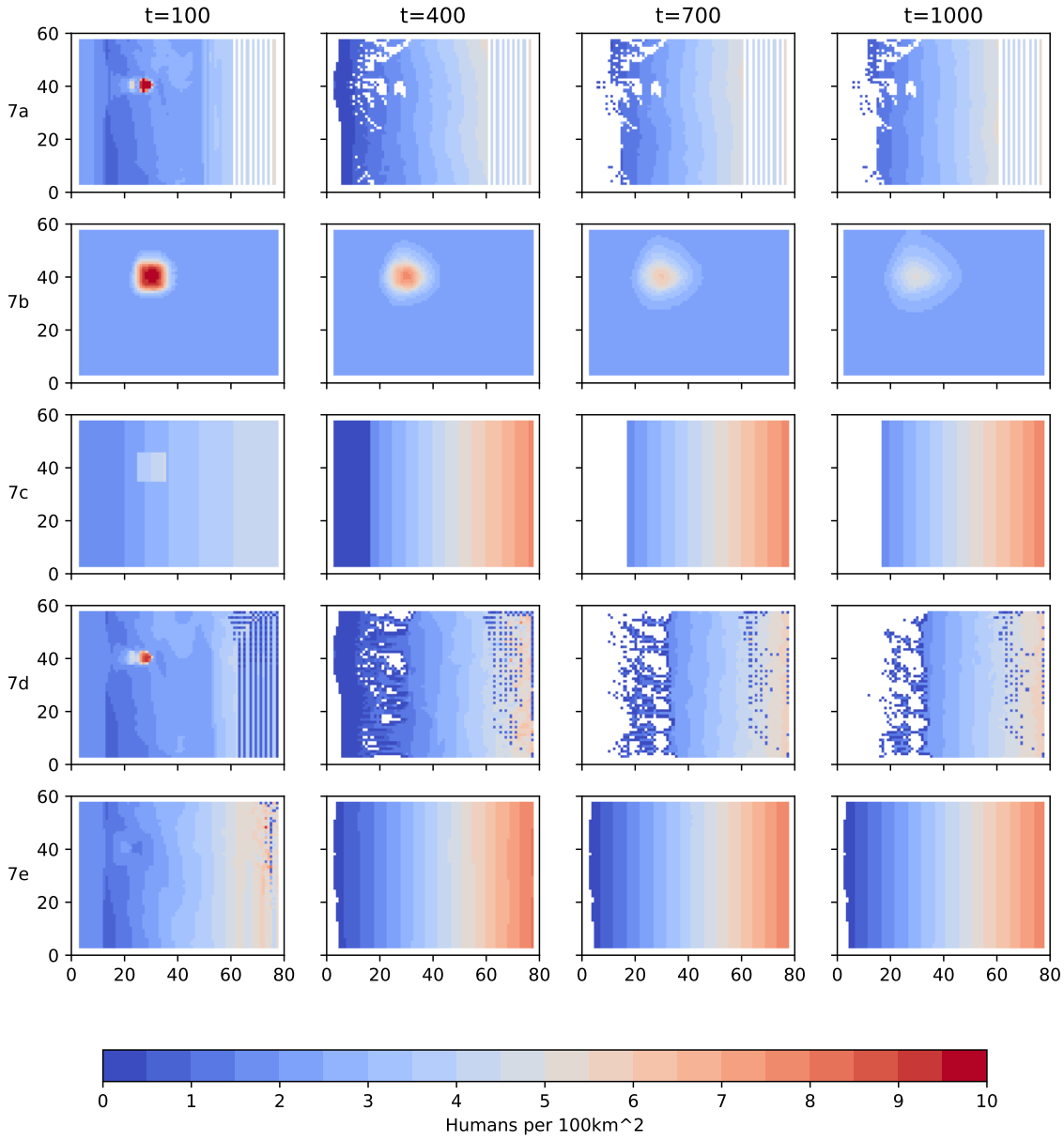

FIGURE 4.30: Human Density for each iteration (a-e) and different time steps of Experiment 7. The axes depict the grid size.

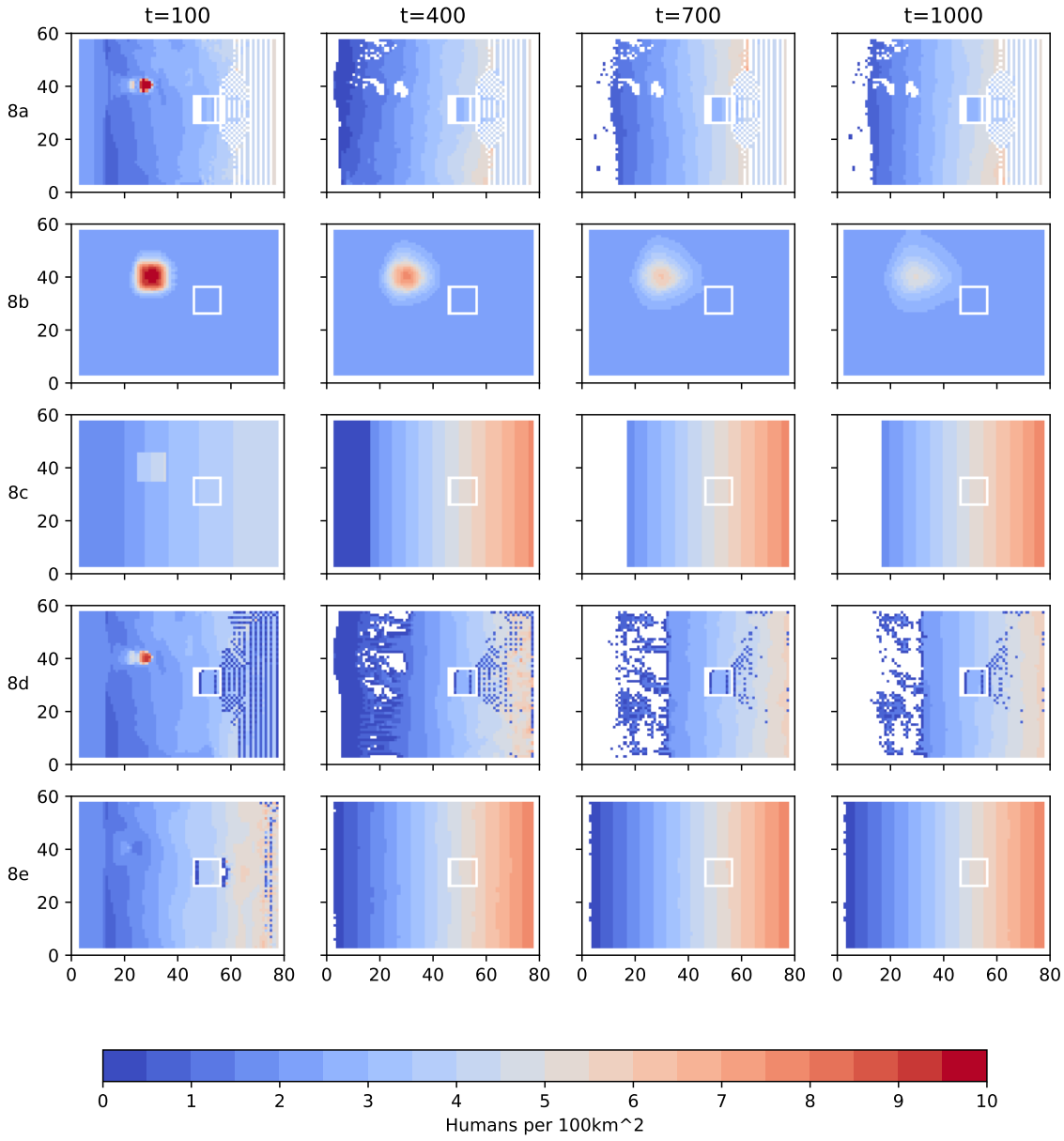

FIGURE 4.31: Human Density for each iteration (a-e) and different time steps of Experiment 8. The axes depict the grid size.

Chapter 5

Realistic Case Study

The realistic case studies are the first step towards the simulation of the dispersal to Europe. These studies aim to inspect the model further and its capabilities in more complex setups, including a non-equidistant grid and changing HEP over time. In this chapter, two types of simulations are presented: A representation of a Dansgaard–Oeschger event for the Aurignacian and Gravettian technocomplexes and a 199 year simulation of the LGM with monthly changing HEP.

5.1 Experiment Setup

In the following sections, the different setups for each experiment are presented. It includes the showcase of the used HEPs, initial distributions, and the selected parameter values. Again, the creation of the HEPs is not shown in detail, but references are given for further reading.

5.1.1 Dansgaard-Oeschger event

Dansgaard-Oeschger events (D-O events) are more or less periodic climate fluctuations that are most prominently known from the ice core records gathered in Greenland, which cover the last glacial cycle (Rasmussen et al., 2014). These fluctuations affect mainly the northern hemisphere and consist of rapid warming within decades, followed by slow and gradual cooling in the next hundreds to thousands of years. The interstadials (warm periods) lead to higher temperatures and increased precipitation, while the stadials (cold periods) create a cold and dry climate. D-O events are one of the most studied climate change events, but a complete explanation of their effects and onset has not been achieved yet (Li and Born, 2019). Their impact on the global climate through teleconnection processes is debated as well (Markle et al., 2017).

D-O events, together with occasional Heinrich events, which depict large groups of icebergs breaking off the glaciers in the northern parts of the Atlantic and floating southward, affected the climate in Europe. The changes in temperature and precipitation differ locally, e.g., the

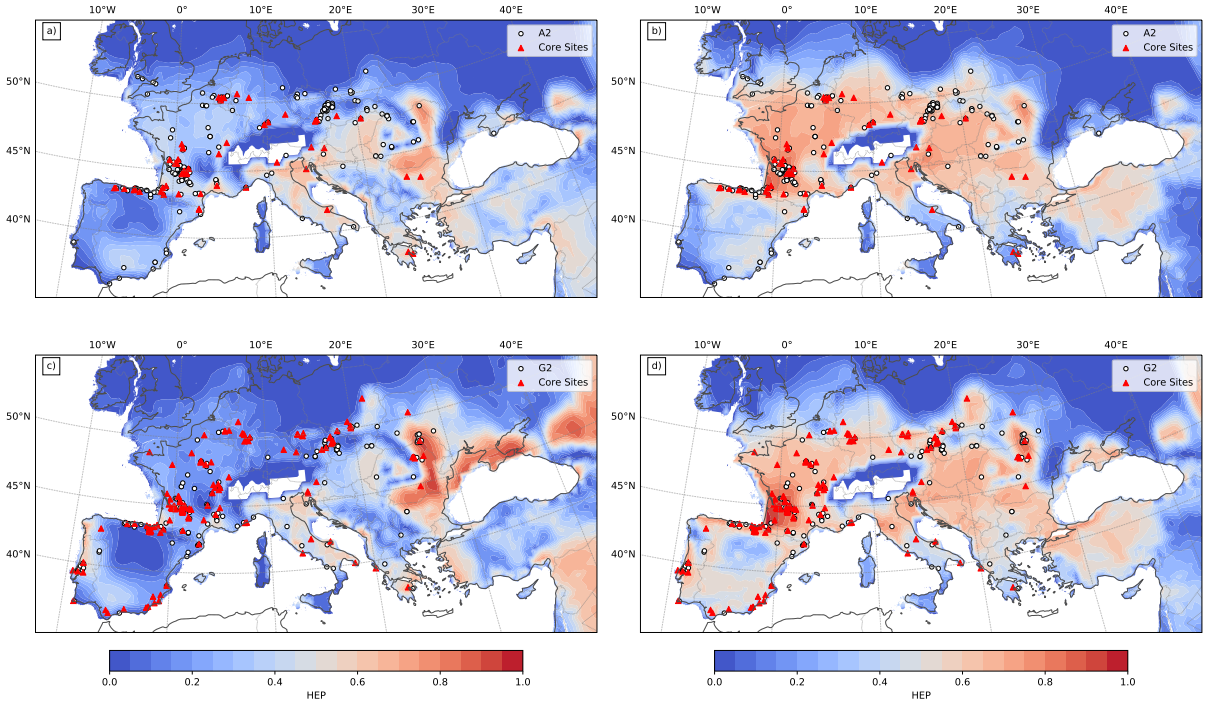

FIGURE 5.1: Overview of the HEPs for the stadial (a and c) and interstadial (b and d) conditions, with the late Aurignacien at the top (a and b) and the late Gravettian at the bottom (c and d).

western Mediterranean has increased rainfall and denser vegetation during the stadials of D-O events (Budsky et al., 2019). The rapid climate changes are likely major drivers for hunter-gatherers in general and the Aurignacian in particular, one of the technocomplexes of interest in this experiment (Banks et al., 2013).

For this case study, a combination of two separate HEPs are used to represent the extremes of a D-O event: cold, stadial, and warm, interstadial times. These HEPs were created by combining precipitation and temperature distributions with the archaeological record. The climate data consists of two simulations that mirror the extreme conditions during a prolonged stadial or interstadial cycle, which are used as prototypical representations of stadial and interstadial times in the following. These datasets were created with the Community Climate System Model version 4 (CCSM4) by the National Center for Atmospheric Research (NCAR). The glacial cover over northern Europe and the Alps are included through the AcHEP, using modified PMIP3 data. A complete overview of the details can be found in Shao et al. (2021), Section 2.2. The resulting distributions of temperature and precipitation are used, in combination with the distribution of archaeological sites of the evolved and late Aurignacian (A2) (Schmidt and Zimmermann, 2019) and the late Gravettian (G2) (Maier and Zimmermann, 2017; Schmidt and Zimmermann, 2018), to calculate the prototypical HEPs for stadial and interstadial times. The final four predictors to estimate the HEP are the Annual Mean Temperature (Bio1), the Temperature Seasonality (Bio4), the Precipitation of the wettest Month (Bio13), and the Precipitation

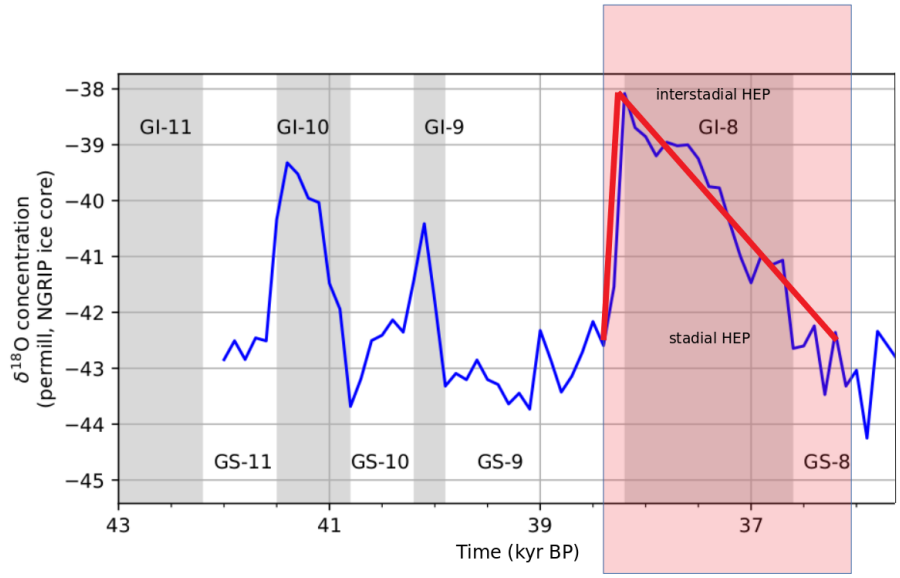

FIGURE 5.2: Greenland Ice Core δ^{18}O concentration (shortened, data from: Wolff et al. (2010)) showing glacial and interglacial time frames. Shaded in red is a typical D-O event, with a linear approximation line between the stadial and interstadial HEPs.

Seasonality (Bio15). For more details regarding the HEP predictor selection and calculation, see Shao et al. (2021), Section 2.3.

Figure 5.1 shows the resulting HEPs together with the distribution of archaeological sites. The core sites, marked as red triangles, feature both early and late Aurignacian occupation, while the white dots are sites occupied in the later phases of the Aurignacian. For the Gravettian HEPs, the same division is used. In general, the core sites can be seen as continuously occupied regions that serve as refugia in stadial times, while the remaining sites are occasionally used and are part of the expansion during interstadial times. The HEP is resolved on a 0.5° grid.

Both stadial and interstadial HEPs are chained together to simulate a D-O event. It follows an idealized curve of a long D-O event within the Greenland ice core δ^{18} Oxygen concentration curve. The curve starts with stadial conditions, followed by rapid warming to interstadial conditions within 100 years. The rest of the simulation time covers a slow return to stadial conditions by linear interpolation between the interstadial and stadial conditions. See 5.2 for a visualization of this. This process is done separately for the Aurignacian and Gravettian technocomplexes. The resulting chains of HEP consist of twenty-one time slices, spanning 2000 years whereby the HEP changes every 100 years. During the simulations, the interpolation between two HEP slides is achieved by the HDM for each numerical time step as discussed in Section 3.4.4. The parameters used for the simulations are listed in Table 5.1, the state of the simulation is saved every ten years. The carrying capacity reaches the stated value at a HEP of one, following Equation 3.13. For this experiment, the borders are open to allow a flow

out of the model. This setting mainly affects Turkey at the southeastern part of the domain, with the borders set at 36° N latitude and 38.5° E longitude. The initial population distribution mirrors the HEP at the first time step, which is the cold case but is set to be 70 % of the carrying capacity at each grid point. This setup is close to the 80 % at which the population growth switches to decay, as described in Figure 3.9. By choosing this fraction, the impact of the initial adjustments to the HEP in the first years of simulations is reduced, which further decreases artifacts that occurred in the idealized simulations (Chapter 4).

TABLE 5.1: List of parameters and their values used in the D-O experiment simulation.

Parameter	Symbol	Value	Parameter	Symbol	Value
Gridpoint Distance	$\Delta x \& \Delta y$	$0.5° \approx 50$ km	Time Step	Δt	$0.01\overline{6}$ yr
Simulation Time	T	2000 yr	Reaction Time	$\Delta\tau$	1yr
Maximum Velocity	v_{max}	5 km yr^{-1}	Critical Velocity	v_{crt}	2.5 km yr^{-1}
Weibull Shape	α	2.5	Weibull Scale	β	0.4
Advection-Diffusion Fraction	ϵ	0.4	Flux-Limiter	η	1.5
Growth Parameter	r_B	0.01 yr^{-1}	Reproduction Onset	κ	0.5 H/100km^2
Friction parameter	μ	0.1 yr^{-1}	Carrying Capacity	Φ_{max}	5 H/100km^2
Maximal potential gradient	$\left(\vec{\nabla}\Phi\right)_{max}$	0.25 H/km			

5.1.2 Last Glacial Maximum

The Last Glacial Maximum (LGM) is a period of time during the last glacial cycle that featured the global maximum of ice land volume. It is estimated to be present between 22 to 19 ka B.P. During this time, the climate in Europe was colder and drier compared to the recent climate (Bartlein et al., 2011), having a drastic impact on the human distribution of Europe (Tallavaara et al., 2015; Maier et al., 2016).

TABLE 5.2: Predefined parameters for the Gaussian distribution used for the creation of the HEP.

Variable	mean μ	standard deviation σ
Monthly Average temperature	283.15 K	6.67 K
Yearly Precipitation sum	800 mm	250 mm

For this case study, a 199 year simulation of the European climate during LGM conditions is used to create monthly HEP distributions. The data was provided by the Coupled Model Intercomparison Project 5 (CMIP5) Community Storage Server (Ekholm, 2013). This set of climate

simulations for the LGM were achieved using the CNRM-CM5 Model, following the CMIP5 framework and model setup (Voldoire et al., 2013).

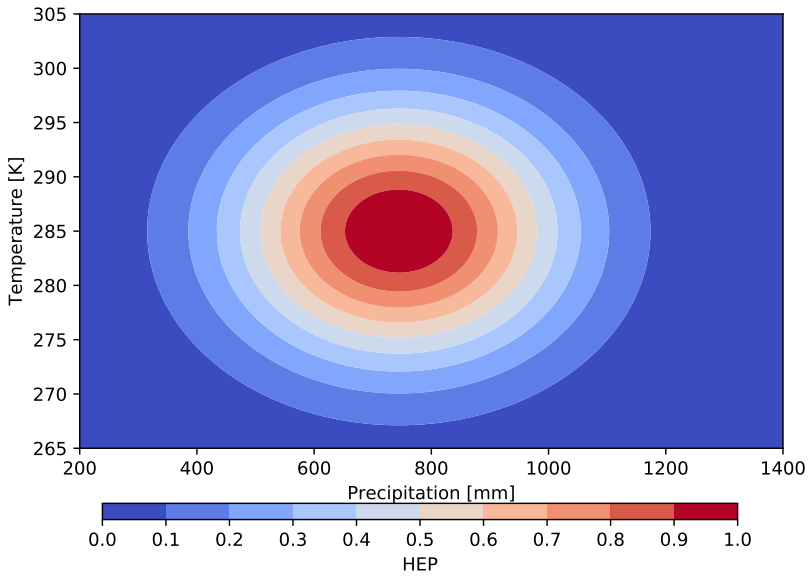

FIGURE 5.3: Conversion between Monthly Average Temperature and Yearly Precipitation Sum to HEP values, using the predefined parameters listed in Table 5.2.

In a first attempt, the average monthly distributions of precipitation and temperature were used to determine a monthly HEP. It includes a set of archaeological sites that can be assigned to the LGM time frame (Maier and Zimmermann, 2015) and that are reasonably close together to form Core Areas (Zimmermann et al., 2004). However, the approach does not provide any reasonable results here. Good regions within the monthly HEPs were limited to a narrow area and the monthly differences and fluctuations were too pronounced. Instead, a new approach was developed for this experiment, which is similar to a previous study of the CRC806: The HEP is estimated by using a predefined probability distribution for averaged temperature and precipitation (Maier et al., 2016). Instead of directly comparing the sites with the climate data, I extracted the average and the standard deviation for both variables and adjusted them to be a bit broader, ending up with the values shown in Table 5.2. Note that the temperature is included on a monthly basis, while the precipitation is a sum of the last twelve months. This approach reduced the fluctuations far enough to apply the HEP. See Figure 5.3 for a visualization of the conversion between climate variables and the HEP. The accessible HEP was also applied to the potentials using a high-resolution representation of today's topography corrected with the sea level for this time frame. The glacial cover is taken from the Paleoclimate Modeling Intercomparison Project 3 (PMIP3) model standard for simulations of the LGM (Braconnot et al., 2012).

The monthly climate values are converted into HEPs using the conversion, creating a series

TABLE 5.3: List of parameters and their values used in the LGM experiment simulation.

Parameter	Symbol	Value	Parameter	Symbol	Value
Gridpoint Distance	$\Delta x \& \Delta y$	$0.5^\circ \approx 50$ km	Time Step	Δt	$0.01\overline{6}$ yr
Simulation Time	T	2388 yr	Reaction Time	$\Delta\tau$	1yr
Maximum Velocity	v_{max}	$0.41\overline{6}$ km yr^{-1}	Critical Velocity	v_{crt}	$0.20\overline{8}$ km yr^{-1}
Weibull Shape	α	2.5	Weibull Scale	β	0.4
Advection-Diffusion Fraction	ϵ	0.4	Flux-Limiter	η	1.5
Growth Parameter	r_B	0.01 yr^{-1}	Reproduction Onset	κ	0.5 H/100km^2
Friction parameter	μ	0.1 yr^{-1}	Carrying Capacity	Φ_{max}	5 H/100km^2
Maximal potential gradient	$\left(\vec{\nabla}\Phi\right)_{max}$	0.25 H/km			

of 199 years of monthly HEPs that are used to initiate the HDM. The model, however, is not designed to use values that change monthly. The minimum time for one HEP slice is one year. This problem is countered by substituting the months in the input to be years within the HDM, requiring to change some parameters accordingly. See Table 5.3 for the corrected configuration. The model effectively computes 2388 "years", with each "year" being one month in reality. The state of the simulation is saved every actual year, resulting in 200 time steps, including the initial distribution.

Figure 5.4 shows the average and standard deviation over all monthly HEPs. The average HEP shown is used as the starting distribution for the simulation, following Equation 3.13. Comparing this with the HEP derived in the original way (Klein et al. (2020): Figure 10), several differences become visible. First, the overall HEP is lower, with the highest average values reaching 0.7 in Brittany, compared to 0.9 in central and western France in Klein et al. (2020). Second, the distribution appears similar but shows some differences that stand out. For example, there is no clear distinction through a trench in HEP between western and eastern populations, as there is no differentiation between the western Solutrean and the eastern Epigravettian technocomplex with this approach. The Iberian Peninsula and the East European areas are much less favorable at the northern shoreline. Third, the standard deviation is area-wide high with values between 0.2 and 0.3, half as large as the actual HEP in many regions. This variability shows the substantial differences in the monthly HEPs, which shift drastically to the north during summer and southward in the winter months. The HEPs used for this simulation are highly variable, and they are likely to show the limitations of the HDM model.

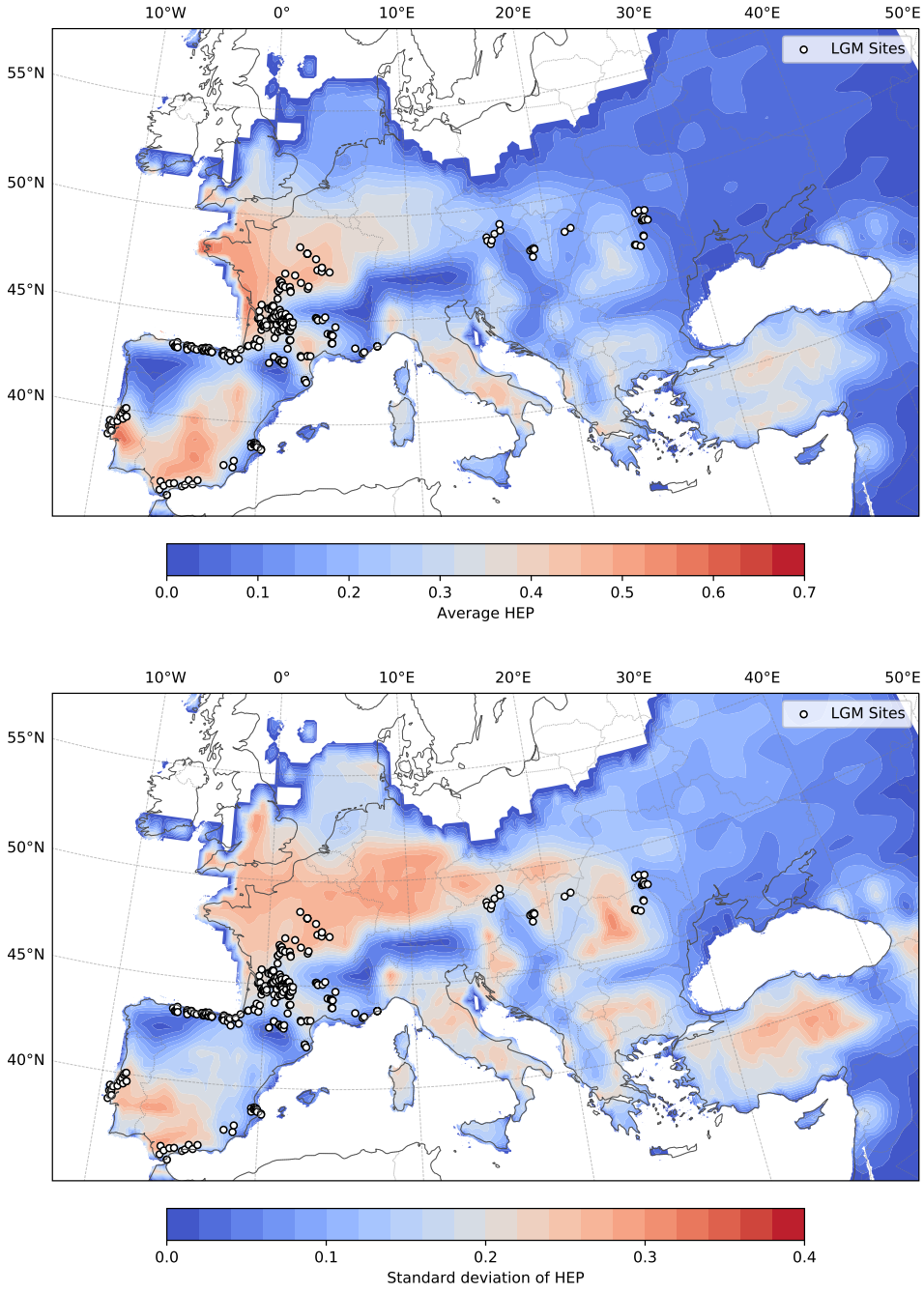

FIGURE 5.4: Average and standard deviation of the monthly HEPs for the LGM case. LGM sites are marked as white dots.

5.2 Results

In this section, the results are shown for both experiments. The focus lies on the total population evolution within the domain, the density distribution, the dispersal flux of humans, and the population dynamics. A selection of snapshots for different time slices is presented that include the main features of each simulation.

5.2.1 Dansgaard-Oeschger Event

Figure 5.5 shows the evolution of the total population within the domain for both the Aurignacian at the top and the Gravettian at the bottom as a continuous line. In addition, the total carrying capacity of the domain is shown as a dashed line. Both populations have a similar evolution: a short decrease of the population as the initial distribution adapts to the cold conditions; increasing population numbers after the sudden switch to warm conditions that feature a higher carrying capacity; the encroachment to a peak of the population roughly 500 years after the onset of the warm conditions; and gradual decay of population as the climate returns gradually back to cold conditions. Nuanced differences between both technocomplexes can be noticed: the population growth for the Gravettian lags slightly behind the Aurignacian, placing the population maximum approximately 100 years later. The slightly higher carrying capacity of the Gravettian does create a slightly higher total population. The total carrying capacity follows the step-wise switch of HEP.

For a further investigation of the model, six time slices were chosen for both experiments. Starting with 30 (a1, g1) and 80 years (a2, g2) into the simulation, the focus lies on the initial response of the population to the cold conditions. The population growth and expansion is pictured at 200 years (a3, g3), with the peak of the population size and extend at 600 years (a4, g4). The decay and a late state of the decay are visible at 1000 years (a5, g5) and 1750 years (a6, g6) of simulation time.

Starting with the Aurignacian, the density distribution over time for the different times of interest are shown in Figures 5.6 and 5.7. The arrows indicate the direction and relative strength of dispersal (Advection + Diffusion Flux) to the maximum of the simulation. In addition, the core sites of the technocomplex are shown. The population dynamics (Birth-Death) are shown in Figure 5.8 and 5.9, again with the arrows indicating the relative strength of dispersal. In the following, the focus lies on the Aurignacian experiment, as the Gravettian has a similar evolution over time with local differences forced by the different HEP. The same Figures as mentioned above for the Gravettian can be found in the Appendix (Section 8).

The initial response to the cold conditions differs for the western and eastern parts of Europe. While the population retreats southward in Western Europe, a strong northward movement in

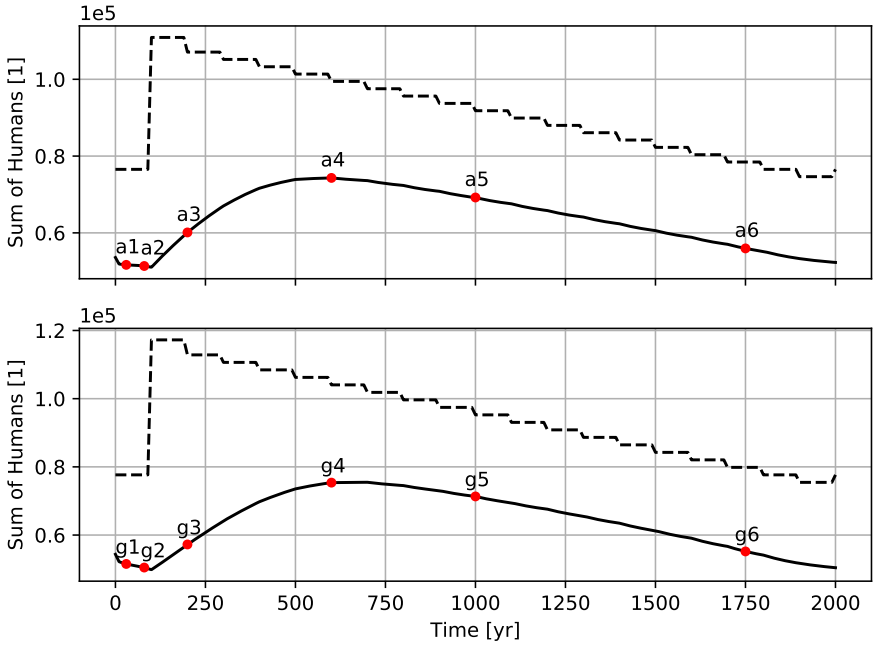

FIGURE 5.5: Total population (solid) and carrying capacity (dashed) of the domain of the Aurignacian D-O simulation (top) and the Gravettian D-O simulation (bottom). Times of interest are marked for both experiments at 30, 80, 200, 600, 1000 and 1750 years of simulation time.

the south and east becomes apparent (a1, a2). It results from a combination of different processes: the available potential is relatively low in the southern regions compared to the north, creating a solid advection term to the north. This dispersal is provided with a new population from southern Italy and Greece, as the migrating population is replaced with new offspring, indicated by the high growth rate. The same behavior can be observed in Romania as well. In the west of Europe, a population contraction is visible, which ends in a slight overpopulation in these regions that results in a decay (a2) after the initial growth in the first years of simulation (a1). The northern parts of Spain are abandoned, despite several core sites at the coast that indicate a long-term occupation of AMH. The population north of the Alps retreats and decreases, creating gaps similar to the clustering observed in the idealized simulations (Chapter 4). The northern frontline retracts as the population leaves, resulting in a high death term as the population becomes unstable, forcing the remaining population to leave as well or die out.

Once the warm conditions are in place (a3), the dispersal shifts in direction and strength drastically. In western Europe, a repopulation of the British Isles and central Germany occurs, the latter being a bit slower as the first arriving humans tend to move back as indicated by the arrows. Some gaps that formed in the first 100 years are closing again, e.g., north of the Alps, while the empty region in Spain takes more time to fill up again. The population growth is very

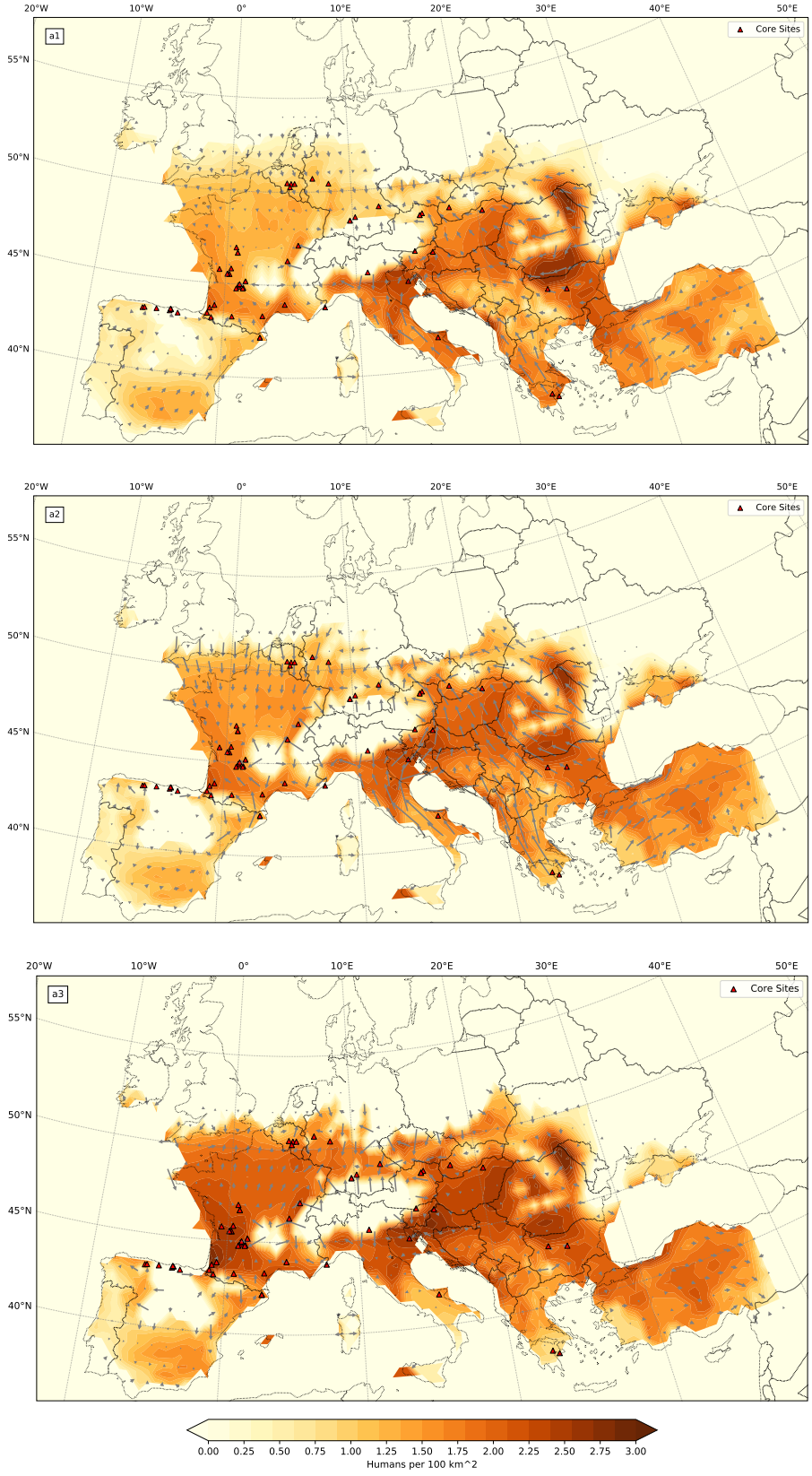

FIGURE 5.6: Density distribution (colored) and dispersal direction and strength (arrows) for the Aurignacian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

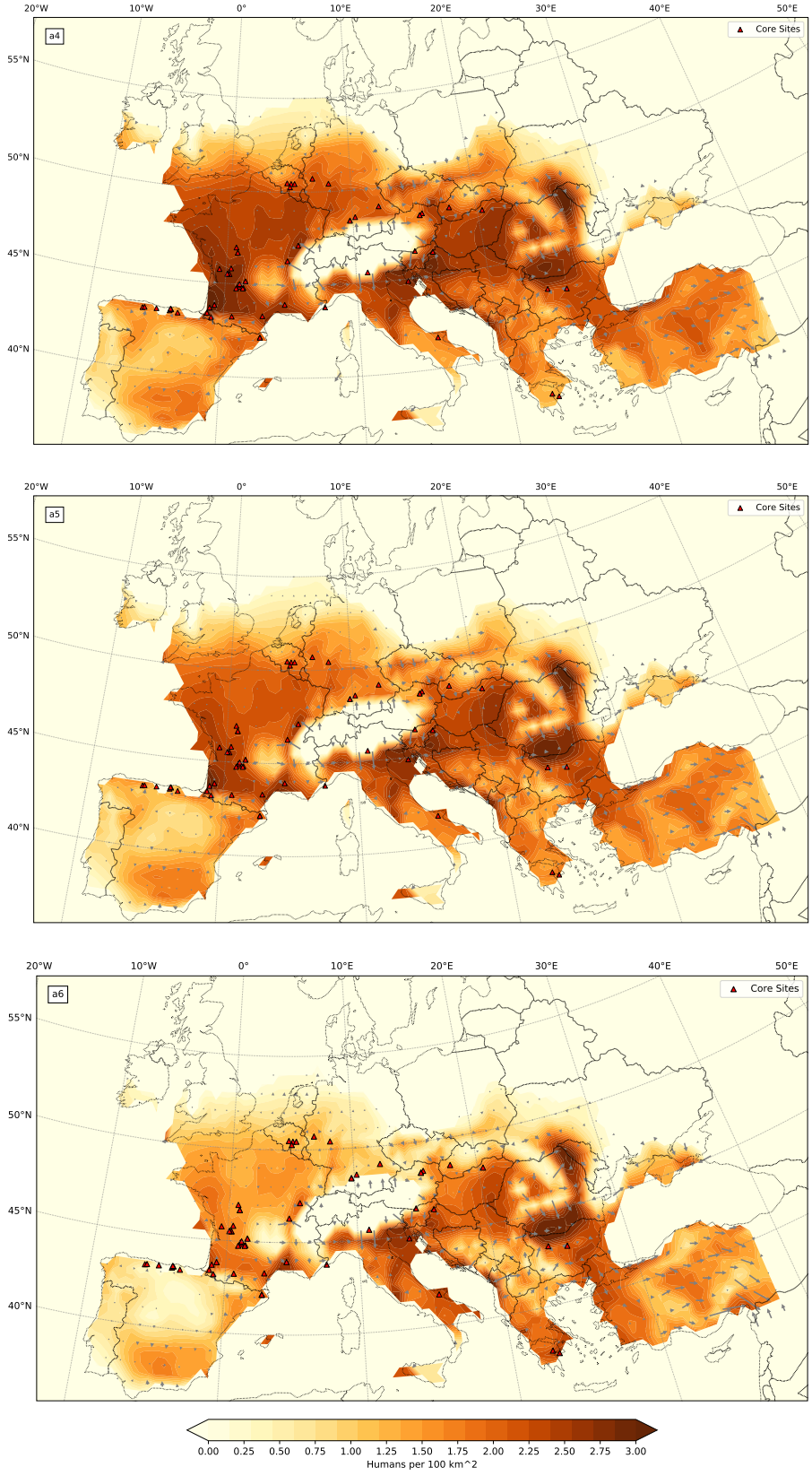

FIGURE 5.7: Density distribution (colored) and dispersal direction and strength (arrows) for the Aurignacian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

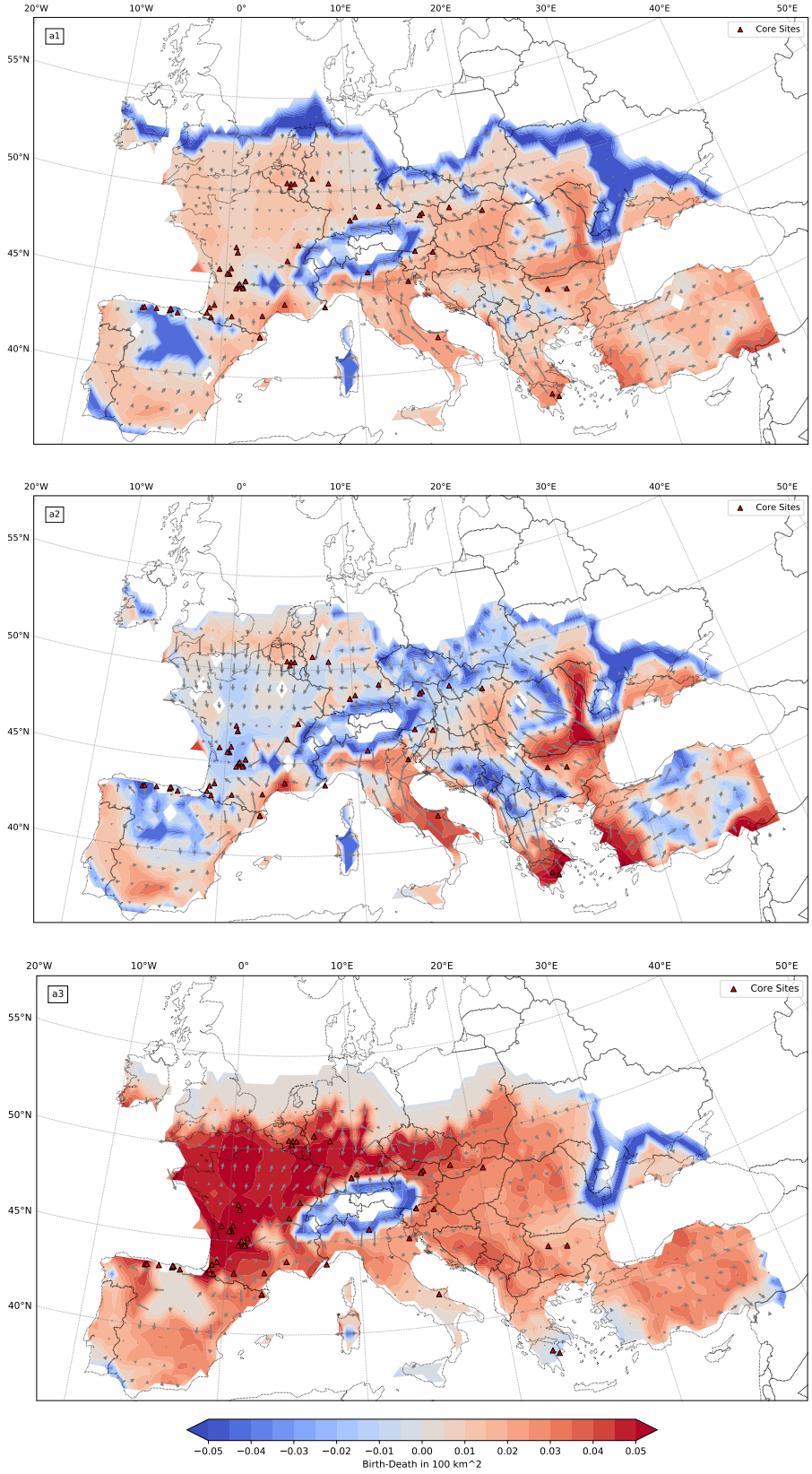

FIGURE 5.8: Birth-Death Term result (colored) and dispersal direction and strength (arrows) for the Aurignacian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

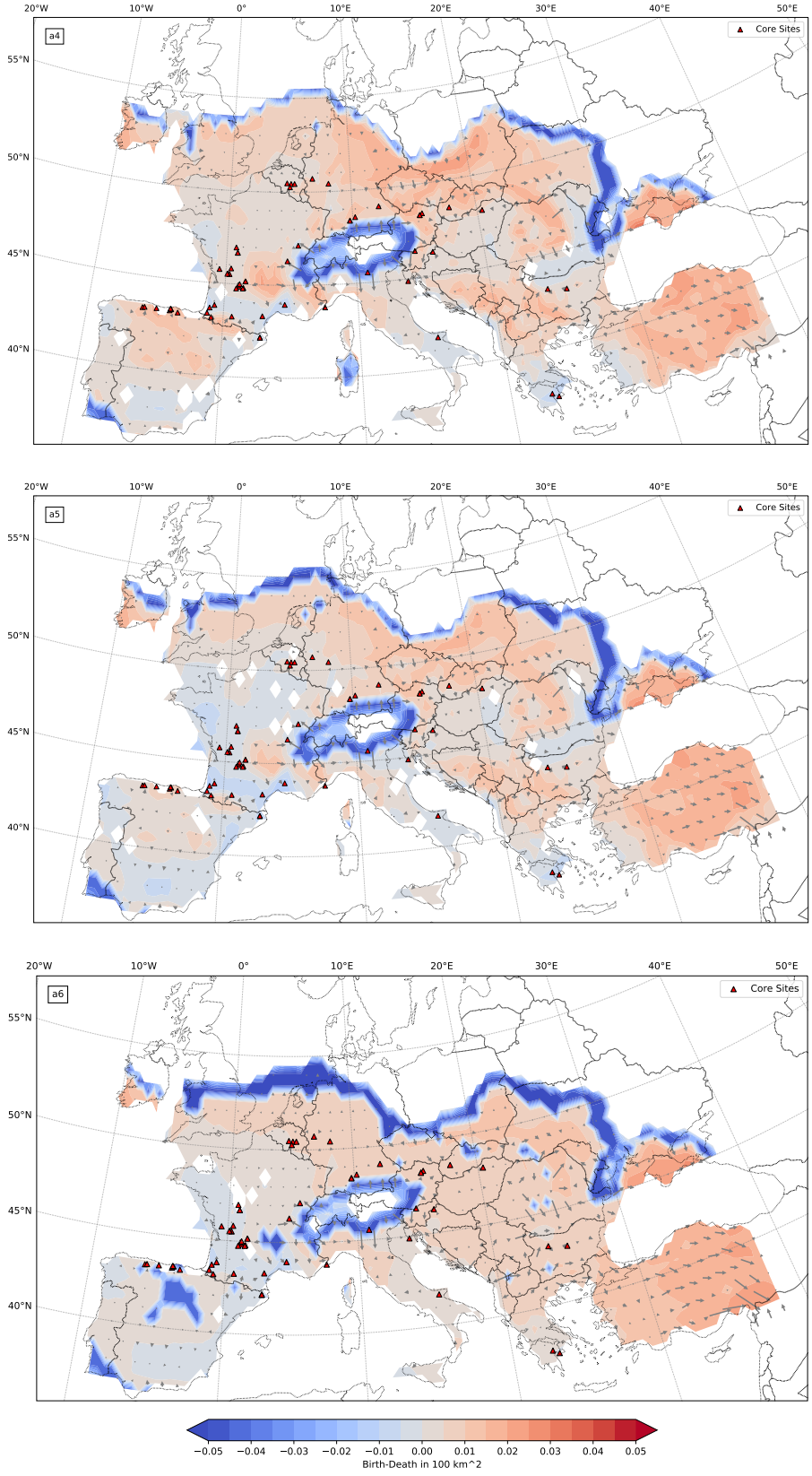

FIGURE 5.9: Birth-Death Term result (colored) and dispersal direction and strength (arrows) for the Aurignacian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

high in western and central Europe, enhancing northward expansion in the region. An expansion in the east of Europe does not take place during this time frame. Instead, slow growth of population in all regions is visible. The lack of dispersal shows that the available HEP is mostly uniform, which slowed down the dispersal that prevailed in the first 100 years of the simulation. The population decay at the front in the north is replaced by a more prominent front of meager population growth, showing that the first individuals moved into these regions. In Turkey, once the population adjusted to the potential, the influence of the open border becomes obvious. A continuous slow drift of the population close to the border equalizes the number of humans that leave the domain. About ten humans per year are lost this way.

As the population reaches its peak (a4), the expansion of population is completed as indicated by the lack of significant dispersal. There is some movement left in the central and eastern parts of Europe, as regions of low HEP with a small population are often less favorable than regions with high HEP and dense population. This occurrence is often the case in the range of very mountainous regions, e.g., the Alps, the Carpathian Mountains, and the Central Massif in France. The population growth declines and starts to equalize with the decay at the fronts of the distribution. Northern Spain is populated again as well.

After the peak, the gradual decay of the population sets in (a5, a6). A general, albeit slow, dispersal to the south can be observed, mainly in the eastern parts of Europe. This dispersal connects with the continuous movement out of the domain, slightly increasing the loss at the border over time. In western Europe, the decline of the population is mainly forced by population dynamics, as most of the area features a negative balance of the Birth-Death term. In the later stages of the decay, most population loss happens at the front of the distribution, fueled by the continuous diffusion term and the slightly positive population growth in established regions.

Closing this experiment, Figure 5.10 shows the numerical consistency in the same manner as in the idealized experiment (Equation 4.1, Figure 4.10). However, this time, a large discrepancy between the estimated population change and the actual population change is visible. The initial peak showing a surplus of humans is presumably the result of the initial population that is taken out of the model at the border in the first time step. After the initial time step, the difference ranges between -0.4 and 0.1 humans per 10 years, having only a minor impact on the cumulative difference over the complete simulation. The error lies in the range of 0.08 % of the total population after 2000 years of simulation. Comparing this with the idealized population difference, the impact of the distorted grid becomes apparent. Using a latitude-longitude grid that does not account for the distortion of the globe drastically reduces the precision of the calculations. The produced errors are within an acceptable range, but there is room for improvement, either by altering the calculations to account for the distortion or by avoiding non-equidistant grids at all.

The Gravettian Experiment (Appendix 8) has generally related results: Increased movement in

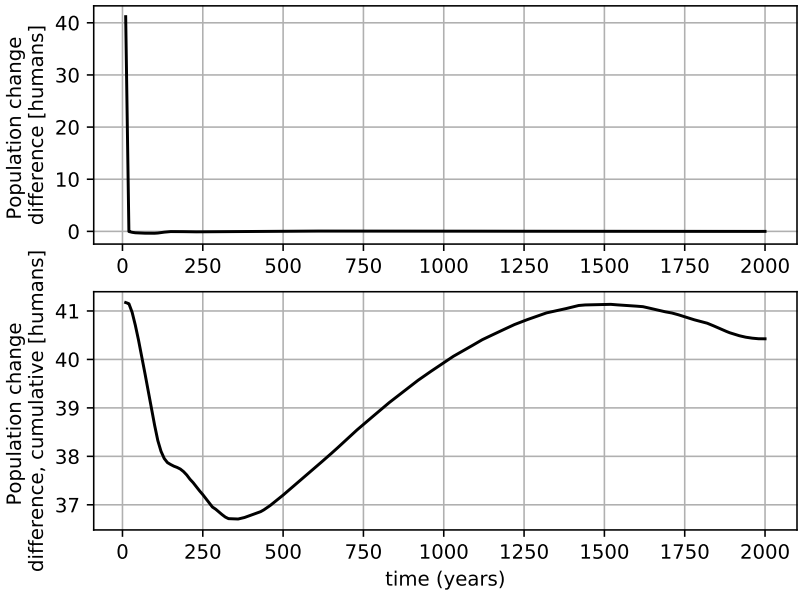

FIGURE 5.10: Difference of the sums of the population change and the population dynamics, see Equation 4.1. The top shows the difference for every time step, the bottom the cumulative difference over time.

Eastern Europe during the initial cold phase and southward retreat in the west (g1); reduction of the population and clustering close to the frontier (g2); expansion into formerly abandoned areas and high birthrate (g3); stabilization and onset of decay of the population (g4, g5, g6). The significant differences are the strong clustering in Western Europe resulting from a drastically decreased population and the cut-off of the main routes of east-west dispersal around the Alps. These connections are reestablished in the interstadial phase but are less occupied the closer the model gets to the stadial phase again. The population in Eastern Europe is more stable and does not appear to be as susceptible to the stadial conditions as the western population. The numerical difference has a higher peak during the onset of the simulation than the Aurignacian experiment, but the overall precision lies in the same range as the Aurignacian experiment, maximizing at 0.12 % of the total population within 2000 years of simulation.

5.2.2 Last Glacial Maximum

Following the previous experiment, the evolution of the total population within the domain is plotted in Figure 5.11. This time, the individual carrying capacities for each month are in the form of scattered points. The population drastically decreases within the first 200 years, effectively halving in this short period. The decrease continues until 125 years of the simulation, at which the population stabilizes at around 7500 humans. The total population lies drastically below the majority of monthly carrying capacities. Only a few months that mark the lower

end of the capacity spectrum come close to the results. The average carrying capacity for the domain lies at 21500 thousand humans, which is nearly three times the stable population size that forms at the end of the simulation.

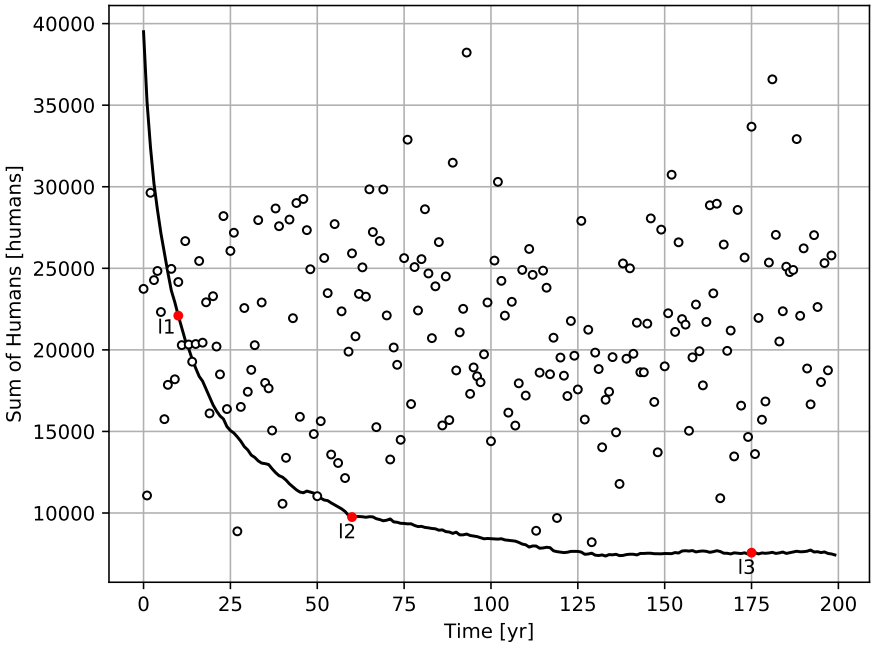

FIGURE 5.11: Total population (solid) and carrying capacity (dots) of the domain of the LGM simulation. Times of interest are marked 10, 60 and 175 years of simulation time.

Three time slices are further investigated: After 10, 60, and 175 years of simulation (marked as l1, l2, and l3). The density distribution is given in Figure 5.13 and the Birth-Death ratio is visualized in Figure 5.14. Both feature the direction and relative strength of the dispersal flux as arrows. Note here that the density range has changed to the previous plots, as the overall population is lower in this experiment.

After ten years of simulation (l1), it can already be observed that the population retreats and diminishes. The dispersal is generally directed towards higher population densities in western France, southern Iberia, and central Italy. However, the actual flux rate is shallow and only one-tenth as potent as in the D-O simulations. The presented arrows had to been magnified by one order to be readable. Most of the dynamics within this simulation is achieved through the Birth-Death term. Nearly the entirety of the domain has an unfavorable ratio of the sources and sinks term, except parts of western France at the cluster of LGM sites, forming a core area, and parts of the Iberian Peninsula. This decrease leads to drastically low populations in central and eastern Europe.

Fifty years later (l2), most of the spatial distribution of humans is gone. Eastern Europe is completely abandoned. Only a tiny population is left in southwestern Germany. Northern Iberia is

mostly evacuated. The populations in Greek and Turkey are getting close to fall below the Reproduction Onset parameter, which would trigger their extinction in the following years. The loss of population numbers does slow down significantly, as the now populated regions have a positive Birth-Death ratio at their cores. The general dispersal is still concentrated towards these center points, showing that the retreat is not yet finished.

Near the end of the simulation (l3), the population has stabilized and retreated into refugia in western France, southern Iberia, and central Italy. Central and Eastern Europe have been abandoned entirely due to the highly fluctuating potential there. The populations in Greece and Turkey are nearly extinct. Only a few humans remain, which can survive until the end of the simulation. The dispersal is still mainly directed towards the population centers, which leads to population clustering and slows down the expansion. The Birth-Death ratio equalizes in most regions, with a general surplus of population in the center and a general decay at the borders.

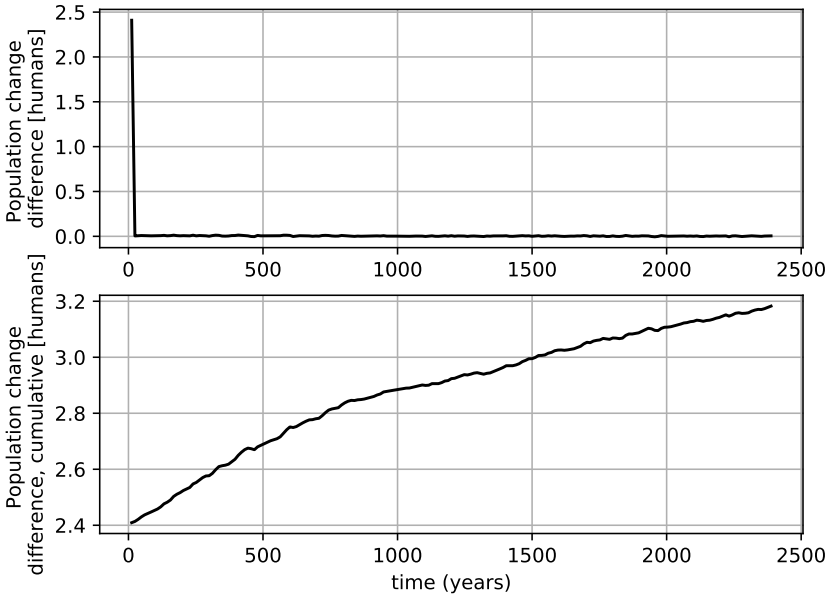

FIGURE 5.12: Difference of the sums of the population change and the population dynamics, see Equation 4.1. The top shows the difference for every time step, the bottom the cumulative difference over time.

The reduced population size and spatial spread can be noticed in the numeric precision as well (Figure 5.12). The initial discrepancy is drastically reduced, as only a tiny portion of the population is located at the border during the initialization of the simulation. The overall reduced dispersal lowers the scale of individual deviations, resulting in only half the error (0.04 %) compared to the Aurignacian D-O event. This result is remarkable, as the LGM simulation does have a significant amount of time steps to calculate.

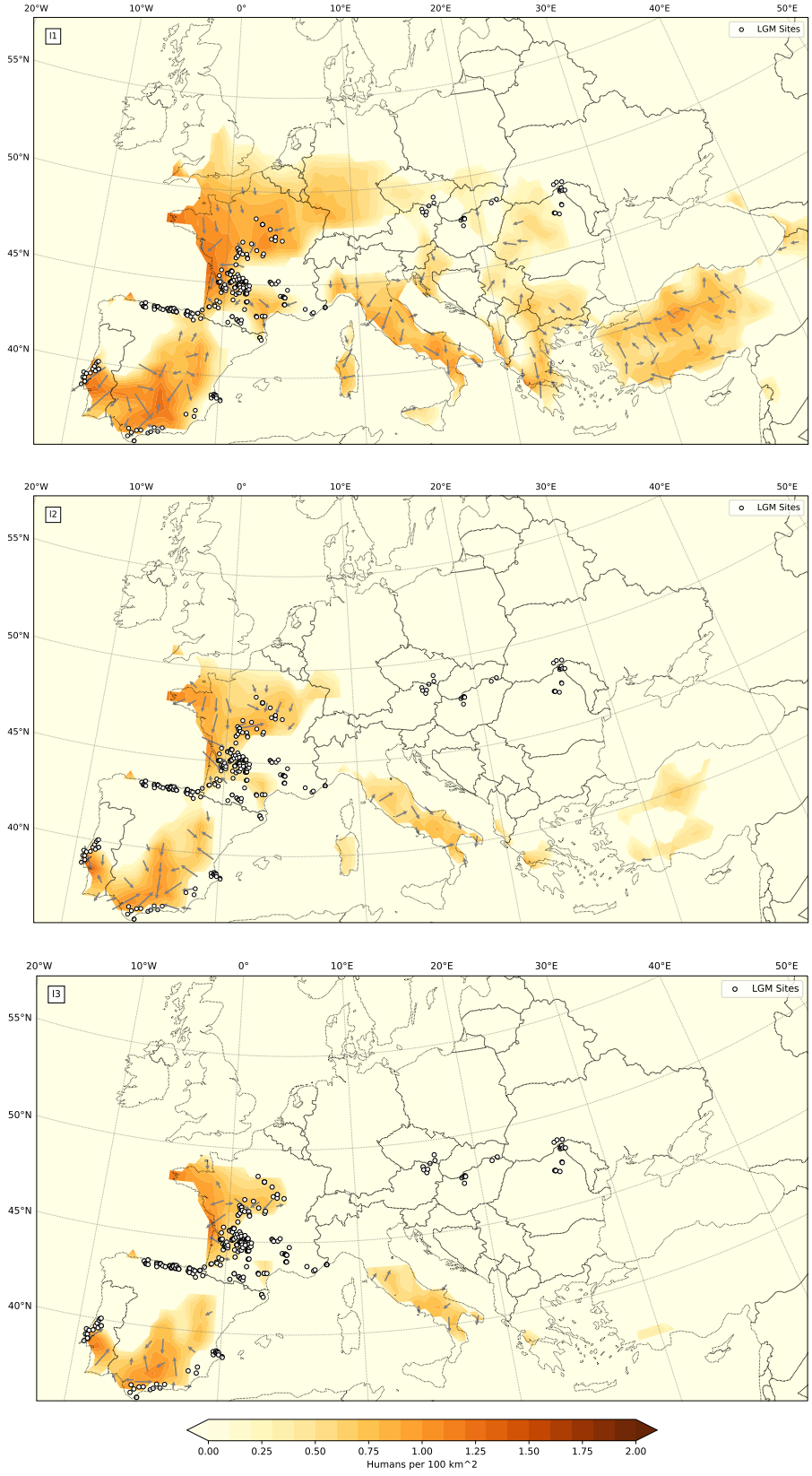

FIGURE 5.13: Density distribution (colored) and dispersal direction and strength (arrows) for the LGM experiment. The time slices follow the pattern as described in the text and Figure 5.11. Note the different color distribution to the previous plots of this kind.

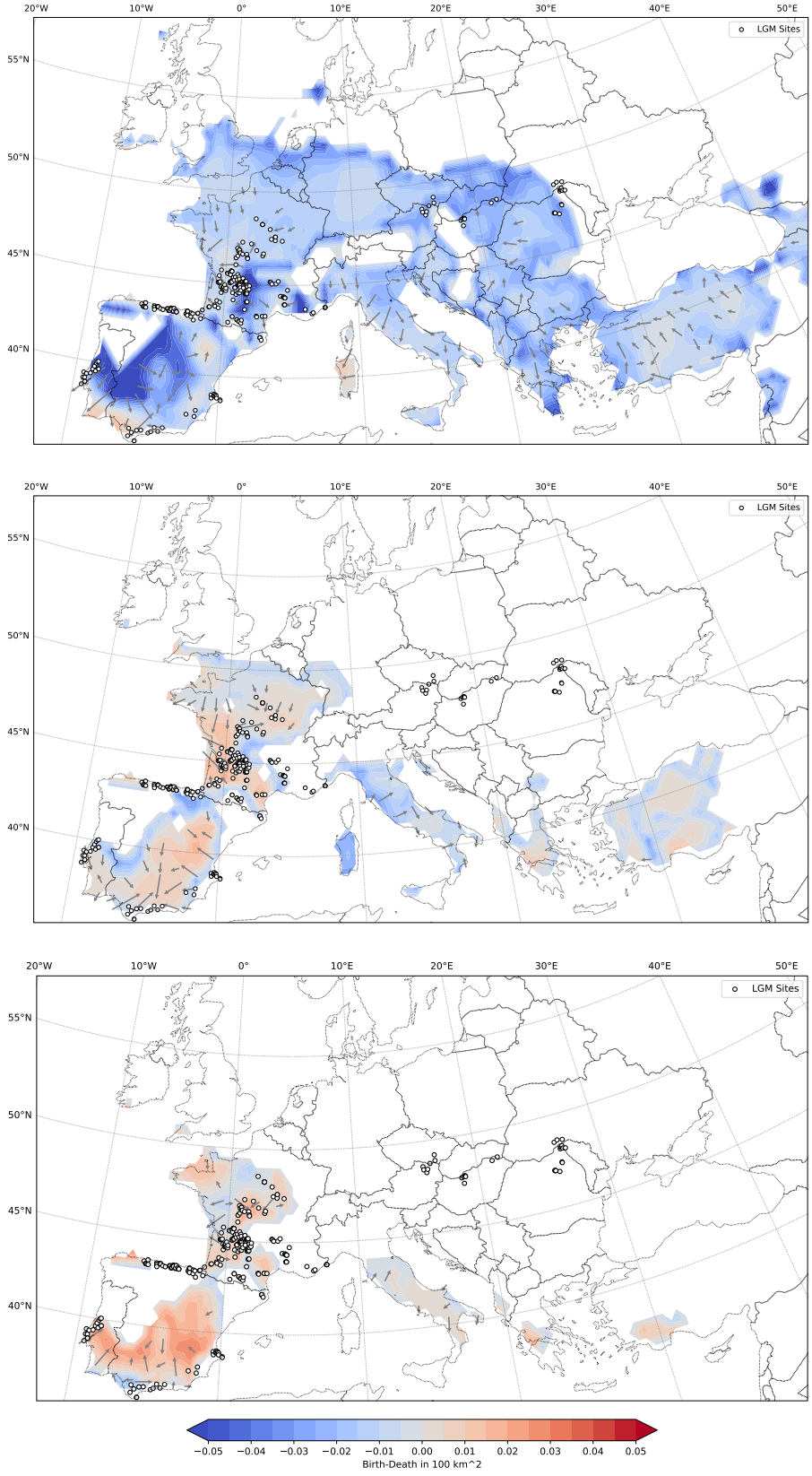

FIGURE 5.14: Birth-Death Term result (colored) and dispersal direction and strength (arrows) for the LGM experiment. The time slices follow the pattern as described in the text and Figure 5.11.

5.3 Discussion

In this chapter, more realistic simulations are shown and the response of the model to temporal changing potentials is investigated. The experiments featured an idealized representation of a Dansgaard-Oeschger Event for the Aurignacian and Gravettian technocomplexes, as well as a seasonal change of HEPs during LGM conditions.

The D-O events are reasonably represented in the HDM. After the initial adjustment to the cold conditions, dispersal and population dynamics change rapidly. The initial contraction in western Europe is reversed into an expansion, driven by a robust positive Birth-Death ratio. The expansion during stadial conditions in eastern Europe is halted in interstadial conditions, reducing the dispersal dynamic and increasing the Birth-Death ratio instead. This discrepancy can be explained by the different HEP distributions for both cases (Figure 5.1): during cold conditions, the southern parts of Italy and Greece are more favorable, while in warm conditions, the continental areas to the north are more favorable. The initial spread to the north is achieved as most of the potential in the south is occupied, triggering an advection driven dispersal towards northern areas with acceptable conditions and a less dense populations. A similar trend can be observed for the Gravettian as well. The difference in HEP is much more distinct in western Europe, which creates a stronger reaction to changing conditions in both Dispersal and population dynamics.

The LGM simulation uses monthly changing potentials, which is extremely fast compared to the gradual evolution of the previous experiment. This setup results in very different results: The population retreats to regions with sufficient HEP for all seasons and is drastically reduced in size, stabilizing at less than one-fourth of the initial population. This experiment shows the model's limitations: It takes time for the advection term to adjust the dispersal direction and velocity to the new potential distribution, but the quickly changing conditions are too fast to create practical, strong dispersal terms. The overall strength of migration is only one-tenth when compared to the D-O experiments. In addition, the quickly changing conditions considerably affect the Birth-Death ratio, resulting in concise time frames of growth in good conditions and decay in unfavorable conditions, depending on the location and season. The overall population is reduced drastically as a result, as for most regions, only a few months each year feature a high potential. These results emphasize that the HDM is designed to simulate long-term, large-scale dispersal and population evolution in the sense of continental scale and over periods of hundreds to thousands of years.

The core sites serve as a guideline for the continuous occupation of AMH. There are, however, some disagreements between them and the HDM results. This discrepancy is mainly the case for the northern Iberian Peninsula at the coast of the Atlantic. Reasons for this can be found in both the climate data used for the HEP and the general structure of the model. The Iberian

Peninsula is prone to increased precipitation during the LGM at the coast, which is often overestimated in paleoclimate models, with an underestimation of precipitation in the Mediterranean basin (Ludwig et al., 2016). This includes the used climate simulation data based on the CNRM-CM5 Model and the use of the CMIP5 setup. The CCSM4 model used for the estimations of stadial and interstadial climate extremes shares these problems, as it is based on the same setup albeit using modifications to better fit the conditions in times before the LGM (Shao et al., 2021). The model itself has a resolution of 0.5° on the latitude-longitude grid, creating distances in the domain between 50 to 40 km. This resolution might not be enough to capture the complex terrain and climate details at the Iberian coast, creating a lower HEP value in conjunction with the difficulties in the climate simulation. In addition, the population in this region might have adapted to the local "unfavorable" climate, which might not be captured in the logistic regression as most sites are located in a different climate. This could be circumvented by separation of the HEP calculations, as it is shown in Klein et al. (2020). The region might be too small in this case, however, to generate enough input data to train the logistic regression.

The numerical precision in the presented simulations does suffer under the distorted latitude-longitude grid, as the numerics are not specially designed to deal with it. From the previous discrepancy that is several magnitudes below the range of values, up to 0.12 % error margin for the Gravettian simulation is achieved. The initial distribution does play a significant role, as most of the difference occurs at the first output of the model. It is likely to be an effect of the outflow of the population at the borders of the domain, which instantaneously reduces the total population during the first numerical time step. The outflow is realized by a sum over all borders, which might not account for small populations due to the inevitable cut-off when small floating-point numbers are added to a larger one. The cut-off is not represented in the sum, which leads to an underestimation of the population that leaves the domain, creating the sizeable positive overshoot in the comparison of the total change with the individual terms. The more minor differences after the initial phase result from the distorted grid, as the numeric construction assumes equidistant distances between grid cells. This effect has been diminished by using the number of humans instead of the density, as the latter is not conserved if the grid cells have different sizes. There is, however, space for improvement, for example, by constructing an actual equidistant grid as a basis for the desired calculations or transforming the issue into another coordinate space, altering the numeric equations.

The realistic experiments show that the HDM can be reasonably used for real-world applications and highlighted the limitations of the large-scale approach of the model. As long as the temporal evolution of the HEP is not too fast or varies significantly, the different mechanisms of the model can adjust in time with a new dynamic.

Chapter 6

Dispersal to Europe

This chapter contains the most ambitious simulation done with the HDM: A full dispersal simulation for the AMH for the dispersal from Africa to Europe, spanning twenty thousand years from 45k BP to 25k BP. This experiment includes large parts of the temporal allocation of Aurignacian and Gravettian technocomplexes. The goal is to extract possible timings and dispersal routes within the initial spread. The second part of the simulation might give clues on dispersal networks within the population.

6.1 Experiment Setup

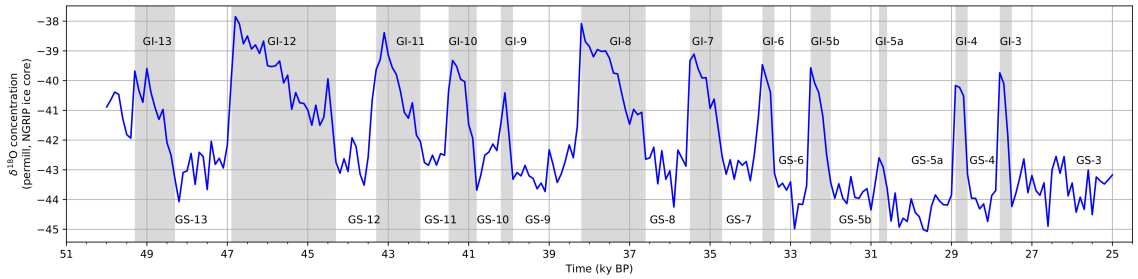

FIGURE 6.1: Greenland Ice Core δ^{18}O concentration (shortened, data from: Wolff et al. (2010)) showing glacial and interglacial time frames. Interstadials are shaded in gray, while stadials are left white. Both are numbered pairwise.

The Aurignacian and Gravettian technocomplexes are split into early and late phases: the early Aurignacian from 45k BP to 38k BP (Aur-P1); the late Aurignacian from 38k BP to 33k BP (Aur-P2); the early Gravettian from 33k BP to 29k BP (Gra-P1); and the late Gravettian from 29k BP to 25k BP (Gra-P2). For each phase, two HEPs are calculated: one for glacial and one for interglacial conditions. For the late phases, see Figure 5.1 in the previous chapter. For the early phases, see Figure 6.2. The climate data used for the creation was designed to capture the extreme climate conditions for alternating stadial and interstadial times that are present in the chosen time frame (Shao et al., 2021). If the onsets of the glacial and interglacial periods are known, the corresponding HEPs can be combined to illustrate the human adaptions to

the different climate phases. This combination allows the HDM to run, simulating a human dispersal driven by climate change. However, using fixed points to switch between cold and warm phases is a rough approximation.

Ice core samples are used to increase the resolution and capture the distinctions between glacial, interglacial, and intermediate phases. For that, the GICC05 model for the NGRIP Ice cores is used (Wolff et al., 2010). It consists of δ^{18}O values in 20-year time steps, which are a ratio of stable oxygen-18 and oxygen-16 isotopes. These can be used as a proxy of the current temperature. For the simulation, the 100 year average of the δ^{18}O values is taken. Then the minimal and maximal values of δ^{18}O are calculated for the desired time frame from 45 to 25 kyr BP.). These values are used to normalize the δ^{18}O curve between zero and one, with zero being the coldest glacial phase and one being the warmest glacial phase in the twenty thousand years of interest. With this normalization, we can now interpolate between the glacial and interglacial HEPs using the following formula for every 100 year time step t:

$$HEP_{new} = (1 - \delta^{18}O_{norm}[t]) \cdot HEP_{cold} + \delta^{18}O_{norm}[t] \cdot HEP_{warm} \tag{6.1}$$

When the time t moves forward, we will have to eventually switch the current HEP (prev) to one of the following phases or technocomplexes (next). An additional interpolation for six time steps (600 years) is carried out for cold and warm HEPs during the switch to make the transition smoother. The resulting intermediate cold and warm HEPs are then used for Equation 6.1 to calculate the values between two phases.

$$HEP_{cold/warm} = \frac{n}{6} \cdot HEP_{nextcold/warm} + \frac{1-n}{6} \cdot HEP_{prevcold/warm} \tag{6.2}$$

The δ^{18}O curve could be used as a proxy for temperature itself and to adjust the temperature that is used for the HEP calculations directly. However, this would require recalculating the HEP at every 100 year time step for the investigated time frame. The HEP does rely on predictors of annual and seasonal temperature and precipitation values, with the annual temperature as one of the fundamental predictors. Changing the temperature according to the δ^{18}Ocurve and recalculating the HEP is likely to produce similar results to the direct interpolation of HEP based on δ^{18}O. Thus, the ladder is chosen to simplify the calculations.

In this experiment, the immigration of AMH to the European continent is simulated. It is very likely that this originated in the Levant, which is why the simulation starts at that point. The distribution is Gaussian shaped and includes a starting population of approximately 5000 thousand humans, with a density of 20 Humans/100km^2 at the center. The archaeological findings from the Skhul and Qafzeh caves suggest that early AMHs already appeared in this region around 100 kyr BP (Cortés-Sánchez et al., 2019)). However, the settlement Europe by AMH is estimated much later and linked to the dispersal from this region around 45 kyr BP (Bosch

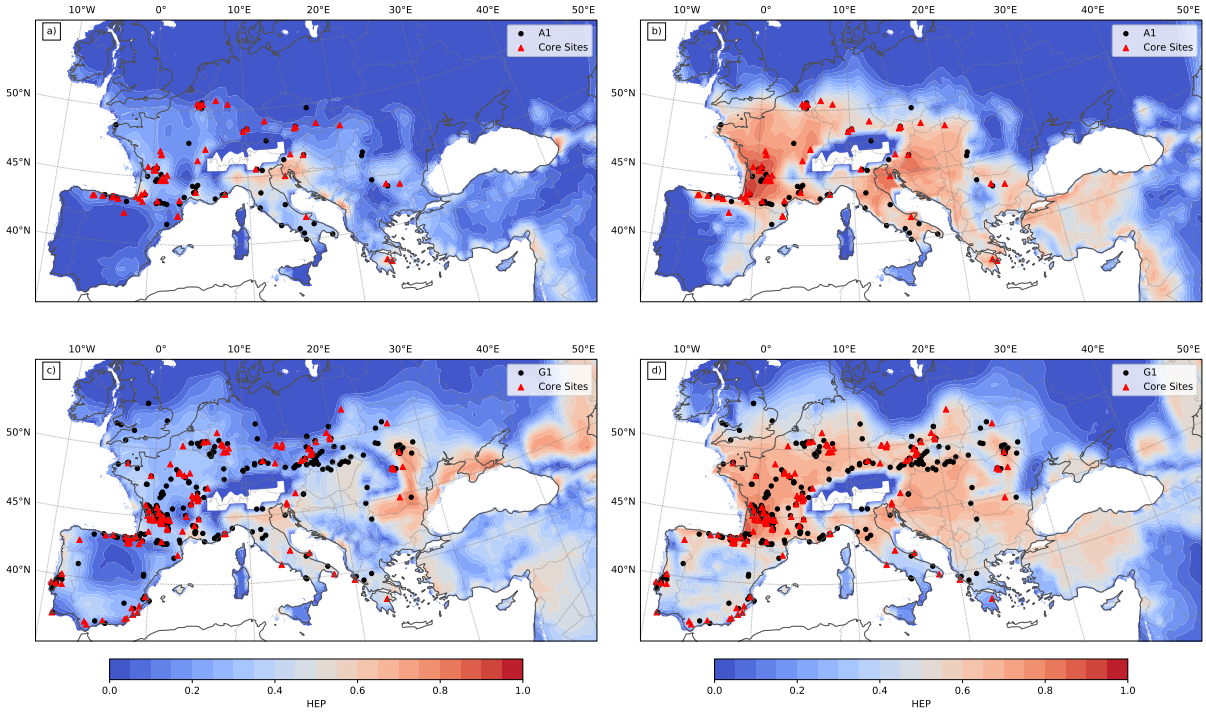

FIGURE 6.2: Overview of the HEPs for the stadial (a and c) and interstadial (b and d) conditions, with the early Aurignacien at the top (a and b) and the early Gravettian at the bottom (c and d).

et al., 2015). Therefore, the simulation excludes the first part of the way between Africa and western Eurasia and all possible human African-Eurasian interactions beforehand. The inclusion of north and central Africa in the model would require an additional dedicated study, which would include the accumulation or creation of climate simulation data and archaeological sites for the time before 45 kyr BP. In addition, there is no clear indication for the precise location of the AMH origin in Africa (Bergström et al., 2021), which includes findings of early AMH ranging from 300 kyr BP to 160 kyr BP scattered over the entirety of the African continent. For this experiment, it is sufficient to focus on the proposed time frame.

The study domain covers the area of (15° W - 49° E, 20° - 60° N) with a resolution of 0.5 degrees. The simulation starts with a small Gaussian population distributed in the area between the Levant and the Dead Sea. The used model parameters are presented in Table 6.1. Every ten years of simulation, the current state of the model is saved.

6.2 Results

Figure 6.3 shows the total Aurignacian population for the time period of dispersal from the Levant to Europe at the top and the sum of dispersal flux at the bottom. Note here that Africa is excluded in the sum, despite having the northern parts within the simulation domain. This

TABLE 6.1: List of parameters and their values used in the Dispersal to Europe experiment.

Parameter	Symbol	Value	Parameter	Symbol	Value
Gridpoint Distance	$\Delta x \& \Delta y$	$0.5^\circ \approx 50$ km	Time Step	Δt	$0.01\bar{6}$ yr
Simulation Time	T	20000 yr	Reaction Time	$\Delta\tau$	1yr
Maximum Velocity	v_{max}	5 km yr^{-1}	Critical Velocity	v_{crt}	2.5 km yr^{-1}
Weibull Shape	α	2.5	Weibull Scale	β	0.4
Advection-Diffusion Fraction	ϵ	0.4	Flux-Limiter	η	1.5
Growth Parameter	r_B	0.01 yr^{-1}	Reproduction Onset	κ	0.5 H/100km^2
Friction parameter	μ	0.1 yr^{-1}	Carrying Capacity	Φ_{max}	5 H/100km^2

exclusion is done to keep the focus on the area of interest. Time slices that are further investigated are marked as red dots.

The population adjusts to the HEP during the first 500 years and spreads over the Levant. Further dispersal is significantly slowed down due to the onset of the glacial stadial 12 and the low HEP in the region. The rough terrain in these parts enhances this as well. With the switch to the glacial interstadial 11, a short burst of dispersal expands the front to the Balkan and Rhodope mountains. The population grows afterward as the interstadial conditions settle in. Shortly before 43 kyr BP, the dispersal into the rest of Europe continues, marked by the sudden increase in population and dispersal flux. Beyond this point, the population curve correlates well with the Greenland Ice core δ^{18}O curve as expected by the design of this experiment: interstadial times feature higher populations and stadial times diminished ones. The dispersal flux has ties to the ice core data as well, but significant movements might occur in both glacial states, mainly expansion during interstadials and contraction during stadials. A full overview of the population for the complete simulation is given in the Appendix 8, Figure 8.6.

Figures 6.4, 6.5 and 6.6 show nine time slices (44, 43, 42.5, 42.0, 41.5, 41.0, 40.0, 39 and 37 kyr BP) of density distribution as well as the relative strength and direction of the dispersal flux for the early Aurignacian dispersal into Europe and the later spread over the continent. The movement can be divided into three stages: the initial slow spread up to the Balkans in the first two thousand years (45-43 kyr BP); the rapid dispersal and expansion over Europe in the following two thousand years (43-41 kyr BP); and the adaption to the cold and warm cycles in the following time of the early Aurignacian (41-38 kyr BP). The slow down in the first stage was briefly discussed before (Figure 6.4, top). The second stage features a swift expansion of the population over Europe, beginning from the Balkans westward to Italy and south into Greece. The wave splits up at the eastern Alps, with the gross of the wave taking the southern route to France and a smaller portion moving north around the Alps. Both waves meet again in central France less than a thousand years later. By the end of this stage, the Iberian Peninsula

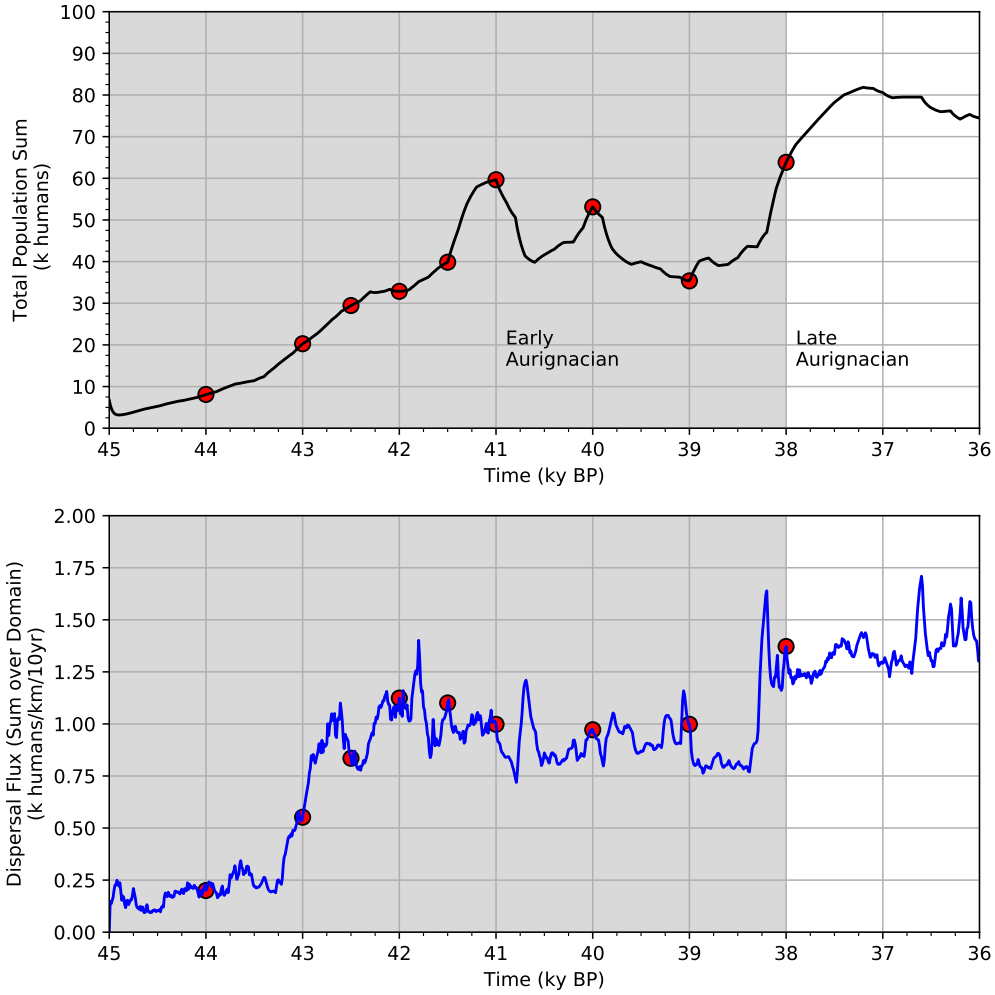

FIGURE 6.3: Total population sum and dispersal flux sum over the domain for the initial dispersal to Europe. The times which are further investigated as time slices are marked with full dots.

has the first inhabitants of AMH.

In the last stage, the population expands northwards during interstadials and contracts back again during stadials, increasing and reducing the population density accordingly. The highest population and spread of the early Aurignacian over Europe is reached at 41 kyr BP at the onset of the stage. In the last phase, the brief interstadials and strong stadial glacial cycles, especially the prolonged glacial stadial 9, reduce the population and halt further expansions until the next interstadial begins at 38 kyr BP. This switch marks the transition from early to late Aurignacian, accompanied by a large increase in population and expansion into former empty areas of Europe. These are mainly the rest of the Iberian Peninsula, the northern parts of central Europe, including the southern parts of the British Isles, and the northern coastal area of the Black Sea. The maximum extent is reached at 35 kyr BP (Figure 6.7), with the highest population densities in southwestern France and the basin below the Carpathians. The density distribution is in good agreement with the archaeological site distribution once Europe is populated.

Focusing on the expansion front, Figure 6.8 shows the timing of the first arrival of AMH in Europe. The first arrival time is defined by the population density reaching the Reproduction Onset (0.5 H/100km^2) for the first time. After an initial northward expansion, the dispersal is slowed down by the Taurus mountain chain and the Anatolian plateau, accompanied by unfavorable stadial conditions. At around 43 kyr BP, the dispersal front reached the Balkans. Due to the favorable interstadial conditions, humans quickly spread across eastern Europe, arriving in central and western Europe around 42 kyr BP. The migration continues southward into the Iberia and northward to the British Island, reaching the maximum extent of the early Aurignacian occupation at 41 kyr BP. Further expansion is achieved, as mentioned before, in the late Aurignacian.

Some sites with their radiocarbon dating are highlighted. While the first AMH have been in the Levant at least 100 kyr BP, the actual dispersal into Europe likely occurred much later between 50-40 kyr BP (Bosch et al., 2015). This is in agreement with the age estimate at the Bacho Kiro site, ranging from 45.8 to 43.6 kyr cal BP (Fewlass et al., 2020). The first occupation within this model falls on a late arrival for this site, being roughly 200 years behind the earliest estimate. The timing is more in line with the Cantabrian site at the northern coastline of the Iberian Peninsula, dated to 43.3 to 40.5 kyr BP (Marín-Arroyo et al., 2018), with the model's estimation between 41.5 and 41.0 kyr BP. While some humans reach the southern parts of the British Isles already at 41 kyr BP, most of the extension into the region happens close to 37 kyr BP, agreeing with the estimations by Dinnis (2012). Overall, the model suggests a slower expansion and a later arrival than estimated by the radiocarbon data.

After the initial occupation of large parts of Europe, the first arrival is a less useful measure for human dispersal. Instead, the dispersal flux, the multiplication of the velocity with the

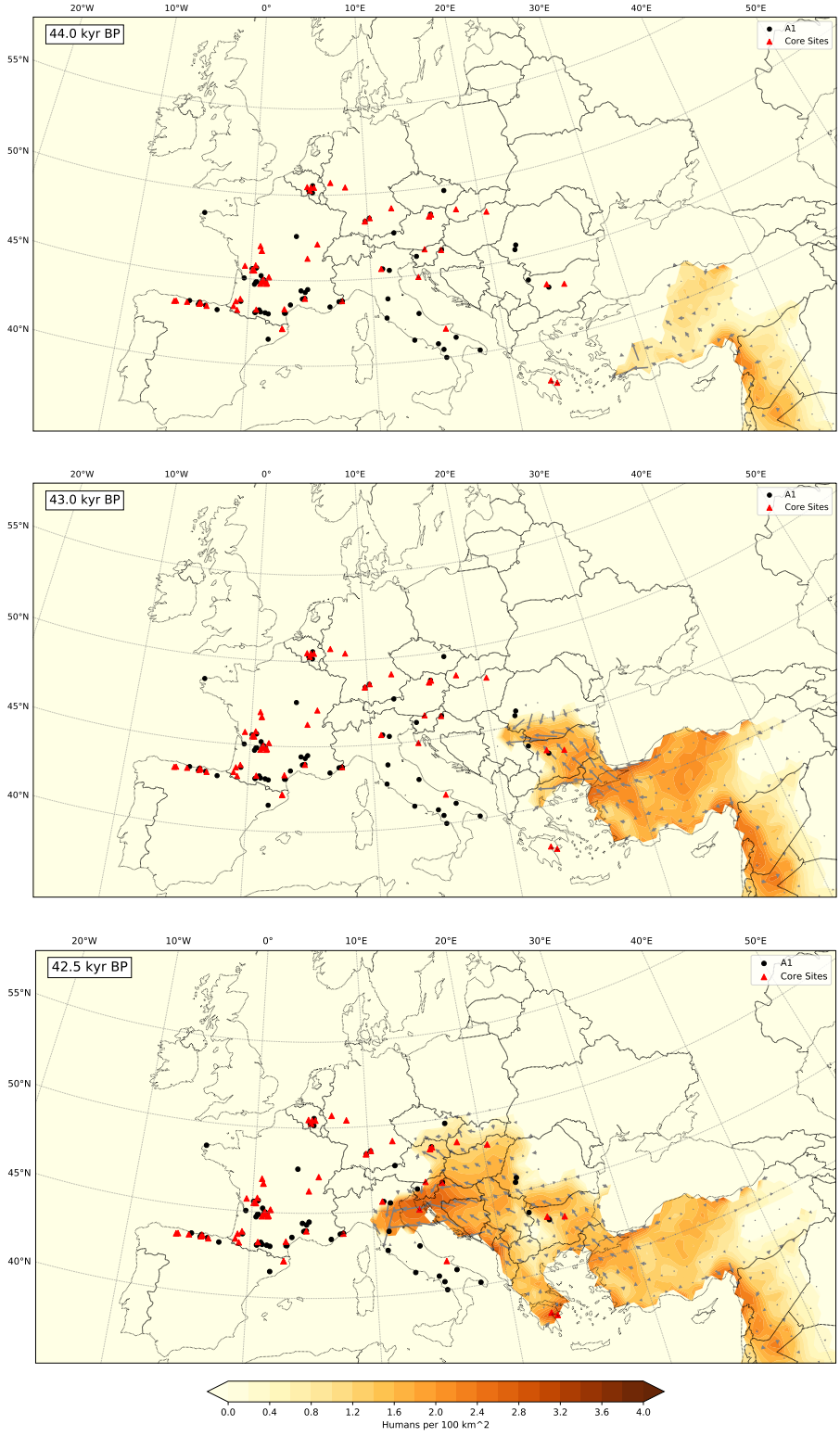

FIGURE 6.4: Simulated population density in Europe at 44.0, 43.0 and 42.5 kyr BP. The symbols represent the Aurignacian archaeological sites, and the arrows the population fluxes (relative to the maximum for the entire period).

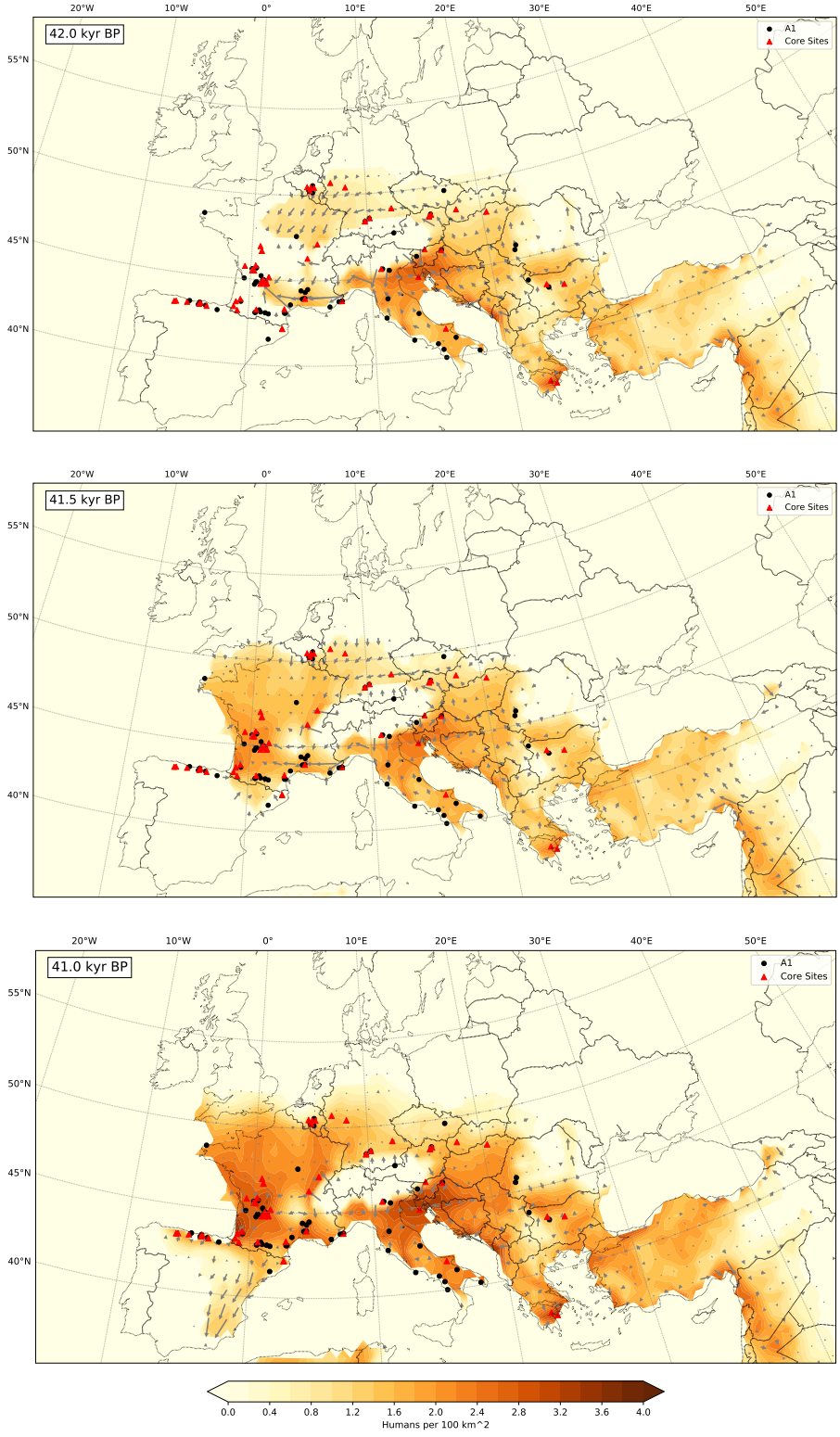

FIGURE 6.5: Simulated population density in Europe at 42.0, 41.5 and 41.0 kyr BP. The symbols represent the Aurignacian archaeological sites, and the arrows the population fluxes (relative to the maximum for the entire period).

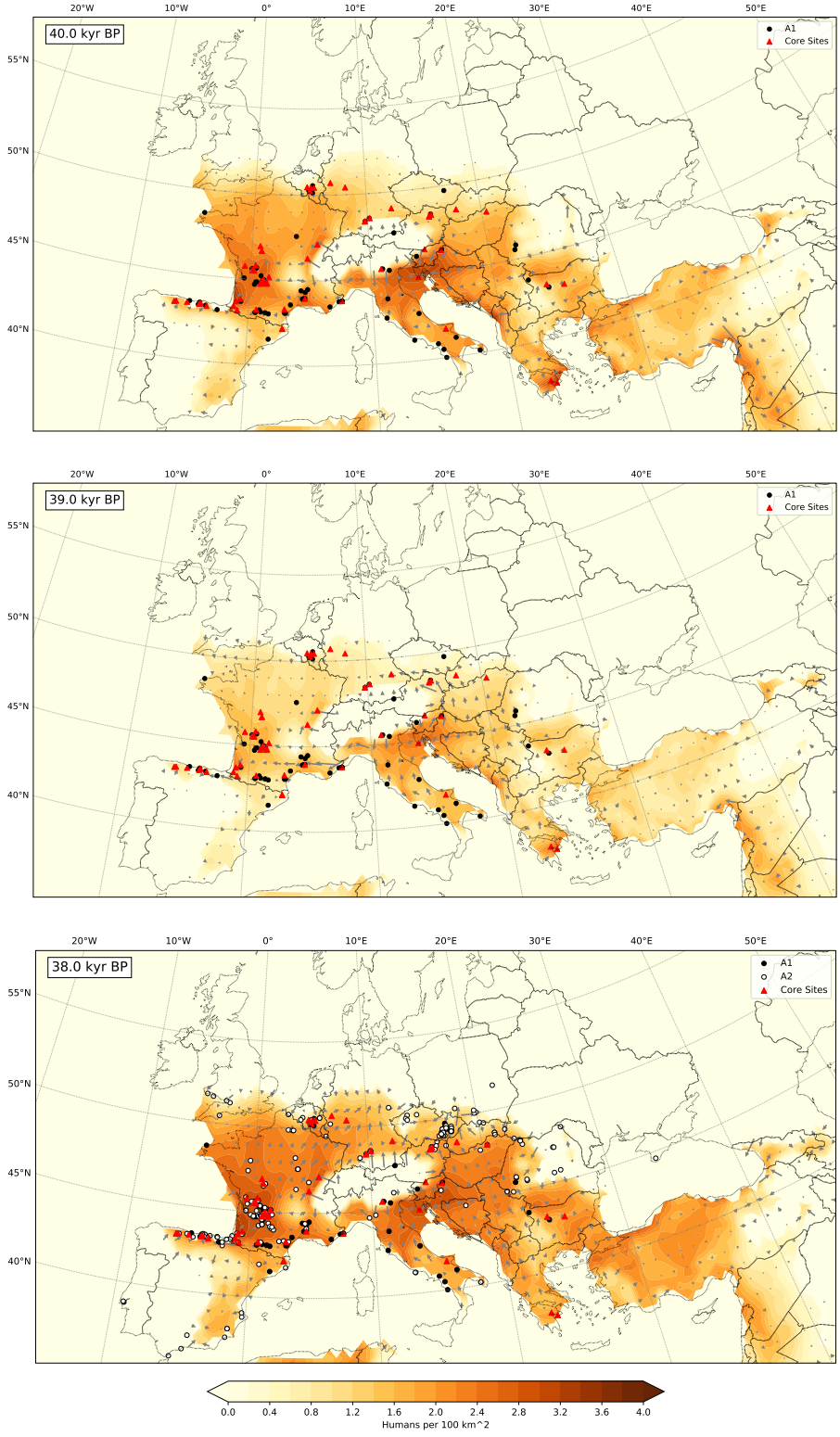

FIGURE 6.6: Simulated population density in Europe at 40.0, 39.0 and 38.0 kyr BP. The symbols represent the Aurignacian archaeological sites, and the arrows the population fluxes (relative to the maximum for the entire period).

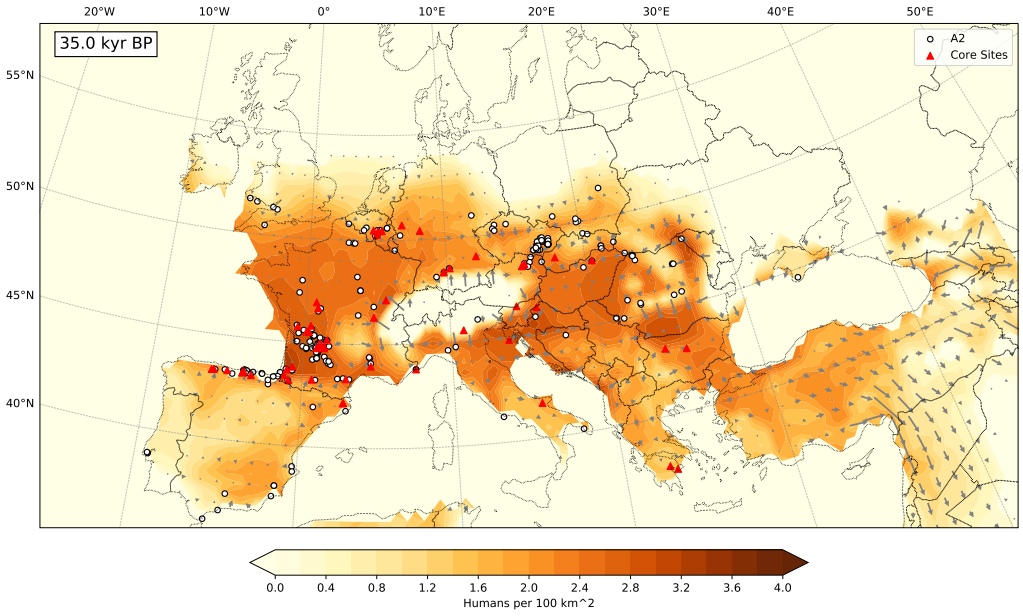

FIGURE 6.7: Simulated population density in Europe at 35 kyr BP. The symbols represent the Aurignacian archaeological sites, and the arrows the population fluxes (relative to the maximum for the entire period).

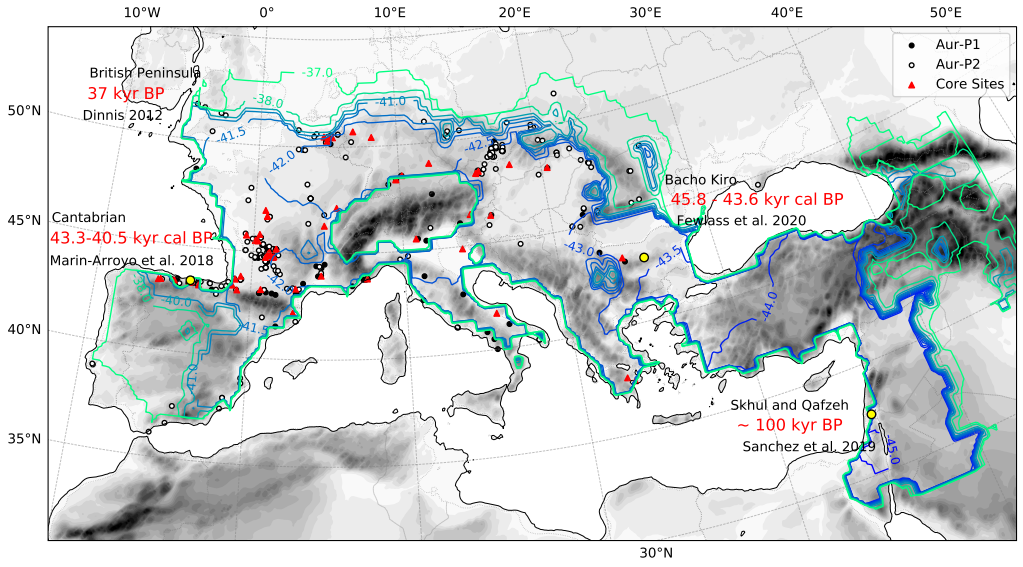

FIGURE 6.8: Front of human dispersal defined by population density reaching 0.4 P 100 km-2 for the first time. The lines indicate the year of the first arrival in kyr BP. Includes the early and late Aurignacian sites (black/white dots) and sites of interest further specified by the text annotation (yellow dots).

migrating human numbers, can be used to identify networks and corridors of human dispersal. The magnitude of this human flux (Humans/km10yr) describes the number of humans moving across a section of distance over time. Figure 6.9 shows the flux averaged over the early Aurignacian phase (44-38 kyr BP) and the late Aurignacian Phase (38-32 kyr BP). For the first phase, the model predicts a dispersal route that follows the general expansion described previously.

The network can be divided into four parts: The Mediterranean coastline (I), the Balkan (II), the eastern (III), and the western Alpine corridors (IV). The main route is the Mediterranean coastline, connecting the Levant in the east with the Iberian Peninsula in the west along the Mediterranean and the Black Sea. This route appears to be the primary path for dispersal into Europe, connecting the different parts of the network and the entirety of southern Europe. The Balkan corridor branches off early in Turkey and proceeds around the Carpathians. Note the additional connection between both routes (I and II) following the Danube and Sava rivers that appears to be the gateway used in the initial spread of AMH in the model (see Figure 6.4, 43-42.5 kyr BP). The Alpine-east connection resembles the path of the northern population of the initial dispersal wave that splitted at the Alps. It shows the route of the population dispersing to the west and connects the southern route with central Europe. The Alpine-west connection is the western counterpart to that and is responsible for the merge of the initial wave and the spread over France and into Belgium at a later time. Very noticeably is the causality between the flux and the topography in several regions as the network bends around mountain ranges and follows basins and riverbeds.

In the later stages of the Aurignacian, the extension of the network into new territory is observable, as well as new connections that arise between the different parts. The Mediterranean corridor expands further into the Iberian Peninsula, reaching the southern and western parts of it. The path from the Levant to Europe through Turkey becomes less clear, as the coastal route shows less flux than before. An inland route seems to be preferred instead of the coastal route. The Balkan corridor extends to the west and east, connecting with the eastern Alpine corridor and including the northern coast of the Black Sea into the network. This connection creates an extensive mesh that connects most parts of Eastern Europe and other parts of the continent. The Danube connection between the Mediterranean and the Balkan routes seems less critical and is only barely distinguishable from the background flux in this stage. The western Alpine corridor extends to the British Isles and is responsible for the spread over central Europe. This corridor, however, is much less distinct regarding the flux-based approach in the northern parts of it. A gradual dispersal over the area appears to be more likely, as the population is distributed mostly uniformly and the HEP lacks significant obstacles here, for example, large mountain ranges.

In both stages, the archaeological site distribution is closely connected with the network. Most sites are either within the network or are closely associated with at least one corridor. The number of sites further from the network increases towards the late Aurignacian, as the spread

appears to be less focused on the routes. The Gravettian (Figure 8.7 in the Appendix) features overall the same flux distribution. The network is therefore not added to these figures. There is, however, an increased flux in the eastern parts of Europe around the Carpathian Mountains and northeast of the Alps. The site distribution expands into regions of low flux and away from the routes of increased dispersal.

Figure 6.10 shows, as in the previous chapters, the numerical consistency of the simulation. Note here that the overall domain is larger than the one used in previous experiments, and the simulation time is ten times as long. The main difference to the D-O event graphs of this type is the lack of a significant initial error, as the population gradually extends from one region that is far enough from the border of the domain. However, the range of the total error per time slice is near twice as large (-0.2 to 0.6 humans) and highly fluctuating. After 20000 years of simulation, the total error accumulates to approximately 240 humans, which is around 0.2 % of the total population. It appears that the error is primarily negative during the initial spread and turns over to be primarily positive once Europe is occupied. The significant negative spikes at 9500, 12500, and 16500 years of simulation time (32,5 and 28,5 kyr BP) correspond to major switches of the HEP, either by the change to a different technoculture or the rapid climate change at the onset of interstadials. These events create large differences in the available potential, triggering the advection term in return and forcing large numbers of humans to move at once to adjust accordingly, which explains the sharp spikes.

A budget analysis is presented in the following to investigate the interactions of the different terms of the model. The focus is on two different sites of interest: Isturitz in Aquitaine, France (43.35° latitude, -1.21° longitude), and Goyet in Wallonia, Belgium (50.44° latitude, 5.01° longitude). The presented results and values correspond to the grid point of the simulation that includes these locations. These points are chosen due to their differences in the accessible HEP, with Isturitz being highly favorable for large parts of the simulated time frame and Goyet being less favorable over time as stadial conditions become more prevalent. Figure 6.11 shows the population evolution of both sites in numbers of humans. Goyet is slightly earlier occupied by the slightly faster but less dense wave north of the Alps, while Isturitz is first populated by the larger population wave. The population in Isturitz is overall twice as large as in Goyet and three times as much during the early Aurignacian phase until the onset of the interstadial GI-8. Both curves correlate with each other with a factor of 0.738, following Pearson's correlation coefficient, which is not surprising as both are strongly dependent on the δ^{18}O ice core curve. The Goyet population declines during the occupation of the Gravettian technocomplex due to the mostly stadial conditions during this time, as mentioned before.

Starting with Isturitz, Figure 6.12 shows the value of each term over time in humans/km^2, with advection, diffusion (a), birth, death, and the sum of all terms (b). A considerably large initial diffusion achieves the initial occupation into the region, followed closely by a huge wave of dispersing humans in the form of the advection term. The diffusion is counteracted by the advection at first but this quickly shifts around and leads to the inflow of humans. After this

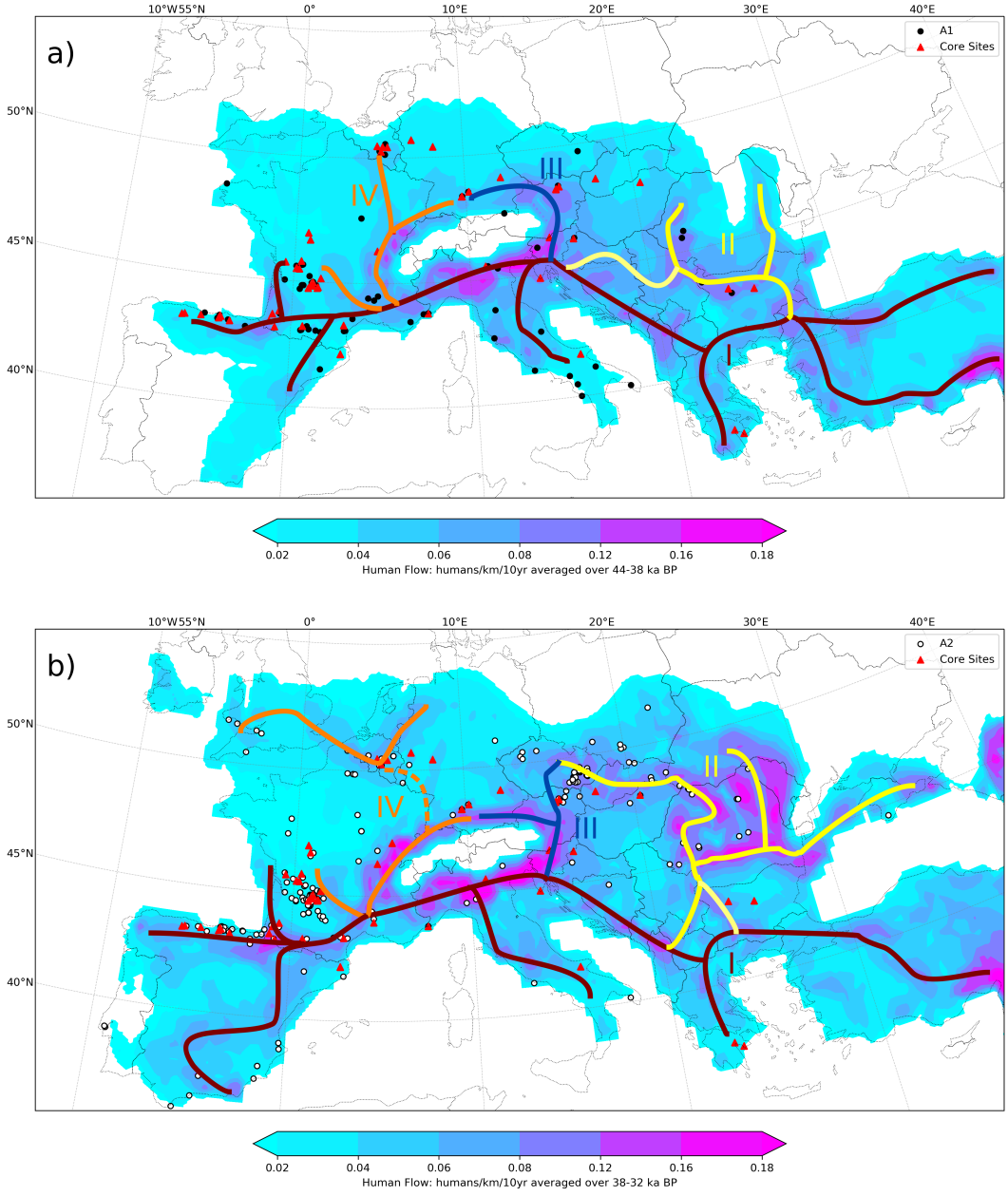

FIGURE 6.9: Population flux in humans/km/10yr averaged over the 44-38 kyr BP period (approximately Aur-P1) and the approximate human dispersal corridors, including the Mediterranean-Coast (I), Balkan (II), Alpine-East (III) and Alpine-West (IV) corridors. (b) as (a), but averaged over the 38-32 kyr BP period (approximately Aur-P2).

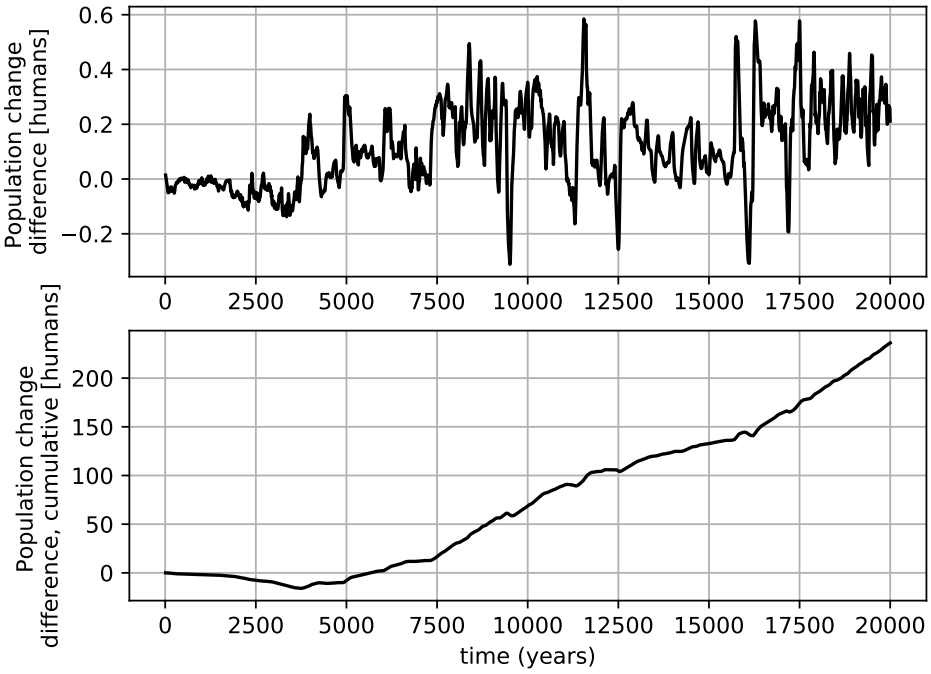

FIGURE 6.10: Difference of the sums of the population change and the population dynamics, see Equation 4.1. The top shows the difference for every time step, the bottom the cumulative difference over time.

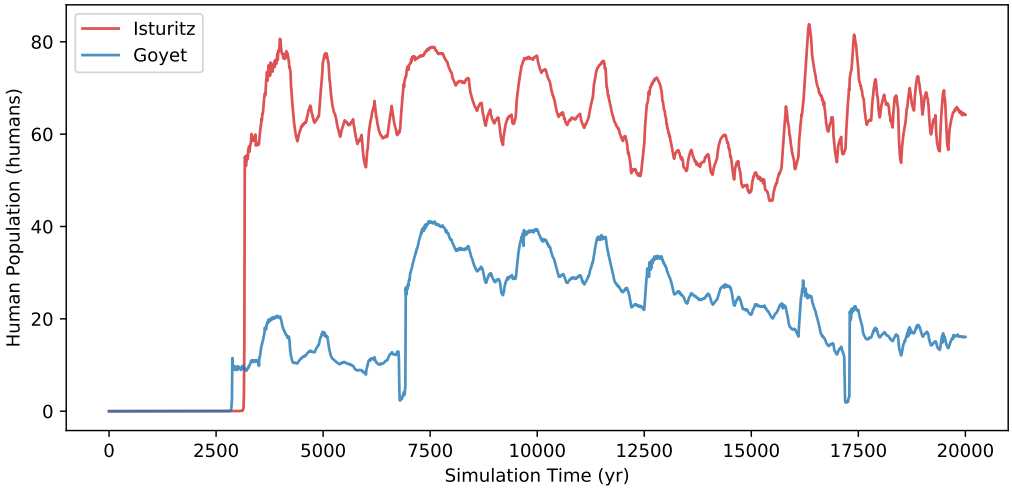

FIGURE 6.11: Population (in humans) over time for the grid cell containing the sites of interest: Isturitz (red, top) and Goyet (blue, bottom).

initial wave, both terms mostly counteract each other: new occupants move in through advection, while others leave the area through diffusion. The high potential that persists through the simulation is very high, which results in this continuous migration into the cell. There is strong mixing of the population with the neighboring grid cells: Assuming a constant overturn of 0.002 humans/km^2 every ten years (estimated by the advection and diffusion term), 4-5 humans are exchanged in the grid cell in this time. In 100 years, this accumulates to 40-50 humans, which is between 50 to 90%of the total population in the cell, depending on the time frame. The birth and death terms have less impact on the population, peaking at 0.0005 humans/km^2, which converts to one human. The actual change of the population size within the grid cell, depicted by the sum of all terms, is in a similar range, excluding the large initial peak of around 55 humans moving into the region.

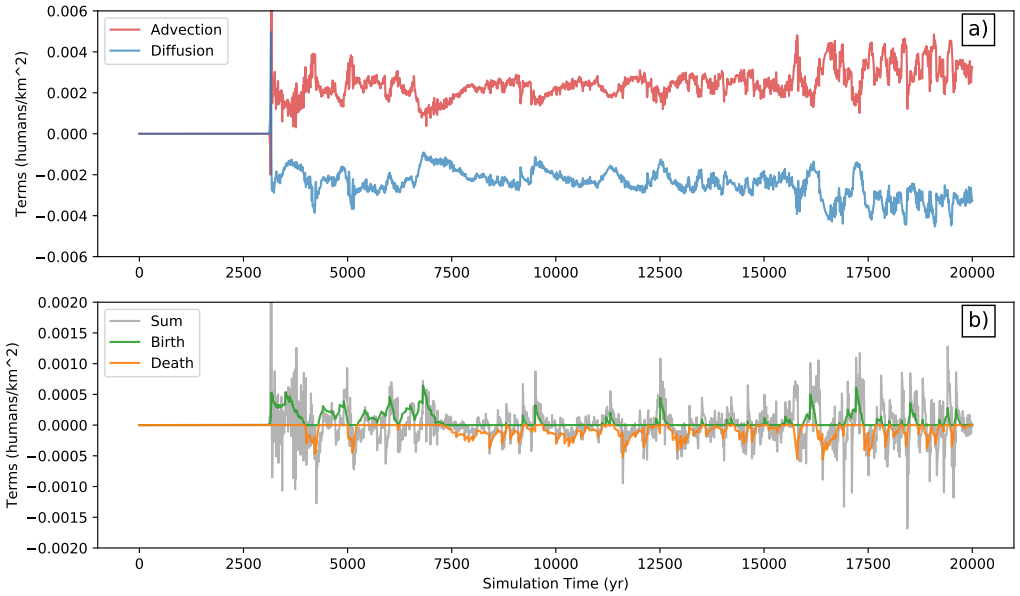

FIGURE 6.12: Impact of the different terms over time for Isturitz. (a) Advection and Diffusion, (b) Birth and Death, with the sum of all terms shaded in gray.

Goyet shows more dynamics in the different terms than Isturitz (Figure 6.13). The initial occupation is mostly diffusion-driven, and the advection is working against the dispersal into this region. It continues for most of the time frame that features the Aurignacian technocomplex, only interrupted by bursts of migration. These are tied to rapid changes in the climate, especially at the onset of interstadials. They show a dispersal out of the region, followed by more extensive migration back into it, for example, at the onsets of GI-8 and GI-3. This behavior appears to be a clustering effect, as explained before, with one of the neighboring grid cells, which is quickly overcome as additional humans move in to fill the gaps. At the onset of the time frame of Gravettian occupation, the terms shift around, having the advection pull humans in and the diffusion driving them out. This is the shift of the front line of human occupation, which lies further north during this time. Before that, the proximity to the front line made the region less favorable. The average strength of both advection and diffusion is lower compared

to Isturitz, moving around one-fourth of the strength. It results from the overall lower population size and the lower attraction to the region, reducing the overturning to less than one human, especially at later simulation times. The birth and death terms are only half as strong during the early Aurignacian phase and become a more significant part of the equation later on. They peak on the onset of solid dispersal activity but do not reach a comparable size to advection and diffusion during these times.

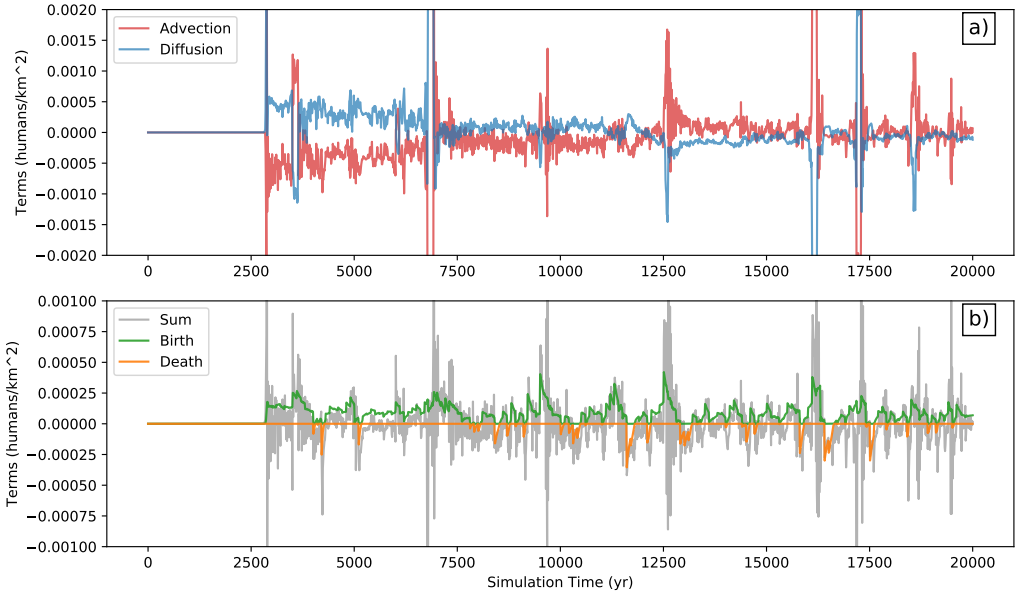

FIGURE 6.13: Impact of the different terms over time for Goyet. (a) Advection and Diffusion, (b) Birth and Death, with the sum of all terms shaded in gray.

Both sites react to the changing HEP with varying intensity and frequency, with the Goyet grid cell reacting both stronger relative to the population size and more often than Isturitz. Goyet features outbursts rather than the gradual exchange of its population with the surrounding area, with a more active birth-death dynamic. Both sites, however, are more driven by dispersal than by the population dynamics, with usual ratios of 2 to 5 times more migration than population control. The ratio of advection and diffusion is close to one for Isturitz at most times after the initial dispersal, as both terms regularly cancel each other out. At Goyet, the ratio usually tilts in the direction of one of the two processes, with a factor of 10 (or 0.1 respectively) being quite common.

6.3 Discussion

This chapter showed the results of a dispersal simulation of the AMH's migration from Africa to Europe, ranging from 45k to 25k BP. The climate change during this migration is approximated by using the Greenland δ^{18}O ice core curve to switch between stadial and interstadial representations of HEP for two technocomplexes, the Aurignacian and the Gravettian.

The total population within the simulation domain excludes the African continent westwards from the Levant but includes a relatively large area from 15° W to 49° E and 20° to 60° N. It is, therefore, no surprise that the result of the model is way above previous estimates of the population size. For the early Aurignacian, the population ranges from 30 to 60 thousand humans. From the later Aurignacian phase onward, estimates range between 70 and 90 thousand humans. This stands in significant contrast to other estimates of population, e.g., 4400-5900 humans from the Aurignacian to the LGM (95% confidence interval: 1700-37700) by Bocquet-Appel et al. (2005). A more recent study by Schmidt and Zimmermann (2019) results in even lower numbers (1500, range of 800-3300). Both studies, however, investigated a much smaller area, focusing mainly on the presence of sites. Schmidt et al. excluded several sites as well, focusing on well-studied regions. For the Gravettian, the population ranges from 1700 to 3700 for the early phase and 700 to 1550 for the late phase (Maier and Zimmermann, 2017). All these studies focus on the site distribution and raw material distribution, excluding areas that could be suitable from the climate point of view. They should considered as the lower limit of population. The HEP is able to extrapolate into these regions but does not account for the availability of raw material. In addition, the range of density used for the HEP often corresponds to the density present at the "core areas" rather than farther out, resulting in larger values. Considering that this model estimates the population for areas that lack archaeological site data but have climate conditions that can suit the survival of AMH, the proposed numbers should be seen as an upper boundary for the population.

The dispersal routes of the initial wave of occupation follow the routes described in Shao et al. (2021), with a major route along the Mediterranean coast from east to west and a split-off east of the alps that reaches central Europe, following the Danube river (see Figure 8.8 in the Appendix). In this simulation, the initial wave takes these two routes simultaneously and merges in France before further expansion to the north and south occurs. This dispersal is generally in agreement with Mellars (2011), but the mechanism is unique in this case: The dispersal is accelerated during interstadial times and slowed or even halted during stadial times. The impact of the climate and large-scale atmospheric systems on the dispersal path and strength is not yet fully understood but is investigated (Obreht et al., 2017).

The exact timing of the first arrival of AMH is challenging to pinpoint precisely for specific points, both in space and time. The comparison between the estimated times of arrival by the model and radiocarbon dates shows that the model suggests a slower expansion and later arrival than depicted by the radiocarbon dates. However, these dates are prone to significant uncertainties (Guilderson, 2005), and a complete list of dates for all used sites in the calculations has not been compiled, making it difficult to state whether the model's estimated arrival times are valid or not. Several simulations with different starting times and dispersal velocities could be run to counteract this, creating a spectrum of possible first arrival dates for specific points of interest. However, such a time-consuming task could not be done in time and is subject to future research.

The initial dispersal wave can be further investigated with the human flux. The flux creates a pattern that can be interpreted as a network of dispersal routes. This network has been created for the early and late Aurignacian (Figure 6.9), featuring the main dispersal routes discussed before and adding more details to secondary routes. The network consists of four parts, which intertwine and connect. Notable features of the presented network include that it often bends around obstacles such as extended mountain ranges and mostly follows coastlines and riverbeds, which are known to accelerate the dispersal of humans (Davison et al., 2006; Campos et al., 2006).

It should be noted that the model runs on a pretty low resolution of 0.5°, resulting in a significant loss of details that can omit obstacles that are crucial for human dispersal. The presented routes result from subjective placing that follows areas of high human flux, sometimes neglecting the fine details of the underlying orography. They should be interpreted as suggestions of dispersal routes and have room for further improvement. An example of this is the connection between the Mediterranean and the Balkan route, following the Danube and Sava rivers. The human flux barely indicates this connection at the first phase and disconnects it in the second phase, despite it being one of the likely routes as suggested by Mellars (2011).

The initial discrepancy in the numerical precision present in the previous experiments is avoided. This change is achieved by the small starting distribution, far from the borders of the model. However, the overall error grows significantly over time, showing that the open borders combined with the distorted grid create an overestimation of the population, albeit being relatively small with 0.2% of the total population after twenty thousand years of simulation. Further improvements could be achieved by altering the border condition to a radiation boundary or absorbing boundary, diminishing the impact of incoming waves of dispersal. A more straightforward approach would be to generate an actual equidistant grid for continental approaches. For global applications, a switch to a different numeric formulation might be the best action, e.g., based on an unstructured triangular grid derived from depicting the globe as a gradually refined icosahedron, as it is used for the ICON model (Zängl et al., 2015). However, this switch would require a complete rewrite of the current numerics used for the HDM to the different grid type of the ICON approach.

Analyzing the different terms for two sites of interest, the interplay of these is revealed. Advection and diffusion are often intertwined, usually balancing each other out. However, several times one of them prevails as part of larger dispersal waves during the switch from cold to warm climate conditions. The population dynamics are often reduced, albeit having their peaks during times of strong dispersal. They reach the size of the dispersal terms close to the front line of occupation and in regions of low populations. The model reacts differently to changes within the HEP, utilizing advection and diffusion firstly and rapidly to even out the available potential, followed by birth and death to adjust the population size to a balanced level over more extended time frames.

The HDM gives a good result for the dispersal of AMH to Europe, with room for improvements. The population size needs adjustment, which can be achieved by using a lower carrying capacity. A population size distribution can be extracted to estimate a range of possible values by using a variety of carrying capacities. The timing of the first arrival is generally a bit too late, which could be adjusted by either increasing the dispersal velocity or starting the simulations at earlier times. The flux network is helpful to identify pathways of continuous migration but depends on the spatial scale of the model and might miss key features of the landscape that are below the resolution. The numerical performance is good enough for most applications, but significant distortions of the grid should be avoided. All model terms are relevant for the dynamics, with advection and diffusion being the most prevalent.

Chapter 7

Discussion

The Human Dispersal Model (HDM) for Anatomically Modern Humans (AMH) has been tested using different approaches to evaluate each part of the model. Starting with nine idealized experiments (Chapter 4) and a total of 59 runs with different setups and settings, each term and parameter and its influence were examined. The results indicate that the model works as expected: the advection term directs the population to areas with higher potential, which can result in a clustered distribution until diffusion and birth terms fill the gaps; the diffusion term drives the population to uncharted terrain and increases in strength at high velocities, allowing further expansion; the birth and death term adjust the population size to the underlying Human Existence Potential (HEP) over time, leading to the extinction in unfavorable regions and increases reproduction rates under favorable conditions. All terms combined show all of the patterns mentioned above and create additional effects through their interactions, e.g., filling up the trenches between clustered populations and overcoming the reproduction onset to expand slightly further into unfavorable regions. The parameters vary in their impact on the model and have to be chosen carefully, as they determine the strength of each model dynamic. The numerical consistency is very high for equidistant grids and stable during the simulation, as long as the criteria, including velocity, spatial and temporal resolution, is met. Overall, the tests reveal a consistent and robust model to simulate the expansion of HDM.

Following the idealized simulations, semi-realistic cases were established to investigate the impact of temporal changing HEP on the model (Chapter 5). These include a distortion of the used grid, as the HEP is defined on latitude and longitude coordinates instead of an equidistant formulation. A Dansgaard-Oeschger (D-O) event is replicated, featuring rapid warming followed by slow and continuous cooling. In addition, monthly values of HEP were used for climate conditions that represent the Last Glacial Maximum (LGM), resulting in a simulation that reveals the model's limits. The former is well represented by the HDM, as the population quickly adjusts through dispersal to the rapid warming and grows over time, reaching a peak 500 years after the onset of climate change. The gradual cooling does not trigger a strong dispersal. The population adjusts mainly through the birth and death term. Some of the interactions mentioned in the idealized simulation are visible as well, mainly the clustering through the advection at high regions of high potential (Figure 8.1, b). Western Europe reacts

more decisive through dispersal to the warming climate than eastern Europe, which featured more movement during the initial cold phase and a halt in the warm phase. Reasons for that lie in the differences between the HEP for stadial and interstadial conditions. Some differences between the archaeological record and the density distribution occur, which are likely the result of discrepancies in the climate simulation used for the HEP. The low spatial resolution of the HDM might be an additional cause, as niches below the resolution in size are either not captured or not represented well enough.

The LGM simulation, however, shows the limits of the model. The temporal shift of HEP is too fast for the model to react, and the substantial differences in the spatial distribution make it difficult for the population to extend beyond their initial distribution. In fact, they decrease in both population numbers and spatial spread, mainly through the population dynamics. Moreover, the dispersal terms are only one-tenth as large as during the D-O simulations. The HDM is suitable for long-term simulations on large spatial scales and cannot capture monthly changes of the HEP. It is better to include seasonality as a variable that serves as a predictor for the generation of HEP.

Chapter 6 applies the HDM to simulate the dispersal of AMH from Africa to Europe, spanning twenty thousand years of continuous simulation and covering the time frames of Aurignacian and Gravettian technocomplexes in Europe. Typical climate conditions for the stadial and interstadial glacial times were used to create HEPs for cold and warm cases. These HEPs distinguish between the early and late phases of the two technocomplexes. This approach results in eight HEP distributions that are further refined by interpolating between them, following the Greenland Ice Core δ^{18}O curve, which can be used as a proxy for the temperature and state of the climate in the Northern Hemisphere. The results depict a feasible dispersal from the Levant into Europe, following pathways as depicted in Mellars (2011), albeit having some deviations. The total population is much higher than presented in the literature, but the simulated population's spatial extend is also larger. Therefore, the result of the HDM should be considered as an upper maximum of the population that could be reached if all of the suitable regions are occupied by AMH. Additional factors can diminish the space of suitable living conditions, e.g., the distance to raw material catchments. The timing of the first arrival appears to be rather late compared to radiocarbon dates but is within reason. It is, however, difficult to decide if the model is at fault here or if the ample time ranges of the dating even allow a distinct assessment. The budget analysis confirms the previous observations. Both advection and diffusion quickly adapt to changes in HEP, creating spikes in the budget. They often contradict each other, diminishing the actual change of the population. The population dynamics have less impact but work over longer time frames rather than spiking sharply during the switch of the HEP.

The experiments show that the numerical precision is reduced by using a distorted grid. An enormous discrepancy can be assigned to the border condition, which removes large parts of the population at the outer grid point layer in the first time step. The sum used in the program seems to create a cut-off, creating an overestimation of the population in the output, despite

being removed correctly in the calculations. The differences afterward result from the grid distortion, creating minor errors as populations move between grid cells of different sizes. These can be diminished by designing an accurate equidistant grid for continental applications. For global simulations, the discrepancies have to be accepted, or the numerics have to be changed significantly to consider the distortion in the calculations.

The HDM is a helpful tool to simulate AMH dispersal and can be theoretically used for other types of humans or animals. For that, the HEP or any other type of suitability estimate needs to be created to meet the needs of the species of interest. Adjustments to the different parameters are required in addition, reflecting the different behavior and limits. The HDM is heavily dependent on the suitability estimates and ultimately mirrors the distribution over time. In the case of a static potential, this connection can create a density distribution similar to the HEP. However, not all potential areas are populated due to the cut-off condition by the reproduction onset and the birth-death term, which leads to a balance of the population growth and decay at around 80% of the carrying capacity. Gaps with low potential can be overcome with time, but perfectly isolated areas of potential can not be settled in. Once several potentials can be linked over time, creating a string of changing conditions, the HDM becomes the only approach to estimate how the population reacts to the different conditions.

Chapter 8

Conclusion and Outlook

This study developed and tested a human population and migration model for the anatomically modern human. This human dispersal model is part of a human modeling framework, an interdisciplinary approach to estimate human survivability, mobility, and dispersal. The focus lies on hunter-gatherer populations during the Paleolithic, but the framework can be applied with suitable adjustments to other species and early farming culture. The main improvement of the human dispersal model compared to models of similar type is the advection term, which resembles a forced movement in a specific direction. The human existence potential determines this force, a measurement for the suitability of a region for the species of interest, derived through logistic regression by combining climate variables and the archaeological site record. This addition to the classic diffusion-reaction equation introduces the aspect of directed dispersal into nearby regions with higher survivability. The potential is further refined in an additional step, allowing the inclusion of the topography and masking out unfeasible regions for occupation, such as water bodies and glaciers. The analytical formulation is translated into a set of numeric equations, which serve as the basis for the calculations.

Several experiments are brought out to test the human dispersal model. The idealized simulations show that the model works as intended and features processes that mirror human behavior, e.g., clustering into stable population sizes rather than spreading out evenly and reacting to changes in both the environment and the distribution of nearby populations. The model is robust and numerically stable as long as the criterium regarding the dispersal velocity in dependency of the temporal and spatial resolution of the model is met. Special care should be taken in choosing the parameters, as they have a significant impact on the dynamics of the model.

Additional experiments were done to test the model's response to temporal changing potentials and the distortion of the grid when using latitude-longitude-based potential data. These include a representation of a Dansgaard-Oeschger event, which describes rapid warming of the climate followed by gradual cooling. Monthly changing conditions were used based on a dataset representing the climate state of the last glacial maximum, showing the limits of the human dispersal model. While the former is represented well by the model for two different

technocomplexes of humans, the early Aurignacian and early Gravettian, the ladder is not. The rapidly changing conditions are too fast for the model to adjust to, resulting in a considerable reduction of the population, which retreats to confined regions of sufficient conditions during all seasons close to the coasts of the Mediterranean Sea and the Atlantic. The grid's distortion and the initial population loss at the border limit the numeric precision but remain small enough so that the model results are tenable regardless.

The final experiment simulates the dispersal of the modern human from the Levant into Europe, spanning twenty thousand years and the Aurignacian and Gravettian technocomplexes. To approximate the climate change in this period, δ^{18}O values from the North Greenland ice core project are taken to interpolate typical HEP distributions for glacial stadial and interstadial times, resulting in approximated conditions every one hundred years. The results show a feasible dispersal along the Mediterranean coastline and north of the Alps, albeit a bit slow compared to the radiocarbon dates of selected sites along the way. A slightly earlier start of the model or an increased dispersal speed could close the gap. A more extensive database of radiocarbon dates was not available, making it difficult to estimate if the chosen sites are representative of the timing or not. In addition, the dates themselves have significant uncertainties tied to them. The estimated population size is quite significant in comparison and should be interpreted as an upper boundary, as empty areas in the archaeological record are occupied in this simulation. In conclusion, the model is suitable for the dispersal simulation but highly depends on the availability of climate and archaeological data to estimate the region's suitability for human survival. It should be used as an additional tool to quantify the dispersal in conjunction with other approaches on that matter.

Additional work can be done, both in the design of the model and the presented application. First, social structures and interactions are only indirectly considered in the formulation of the HEP by defining the preferences of a population. These could be included by differentiating between different subgroups of humans, each with an individual set of equations representing dispersal with additional terms to regulate their interactions. They share the available resources, which in return enables additional conflicts that need to be resolved between the groups. Possible resolutions include:

- sharing the space and resources but being indifferent to each other,
- mixing the populations with the possibility of creating a new subgroup,
- moving away from each other to create a gap between them,
- or driving the other population back by force.

Second, the dispersal simulation into Europe could be repeated with different starting times and dispersal velocities, allowing to visualize a distribution of first arrival times in places of interest. The collection of additional radiocarbon data is of high priority to better estimate the feasibility of the simulation. Third, several additional improvements to the model could

be made, especially at the parameterization of processes that are below the resolution of this model. For that, agent-based models could be used to determine patterns of interactions between the simulated individuals, allowing the creation of formulations for larger scales.

In conclusion, the HDM is a valuable tool for a dynamic simulation of human dispersal that combines archaeological knowledge with paleoclimate simulation data.

Appendix

Bioclim Variables

TABLE 8.1: Overview of the 19 Bioclimatic variables which serve as candidates for predictors to calculate the HEP. Adapted from Klein et al. (2020), Table 2.

Variable	Descriptive Name
Bio 1	Annual Mean Temperature
Bio 2	Mean Diurnal Range
Bio 3	Isothermality
Bio 4	Temperature Seasonality
Bio 5	Max Temperature of Warmest Month
Bio 6	Min Temperature of Coldest Month
Bio 7	Temperature Annual Range
Bio 8	Mean Temperature of Wettest Quarter
Bio 9	Mean Temperature of Driest Quarter
Bio 10	Mean Temperature of Warmest Quarter
Bio 11	Mean Temperature of Coldest Quarter
Bio 12	Annual Precipitation
Bio 13	Precipitation of Wettest Month
Bio 14	Precipitation of Driest Month
Bio 15	Precipitation Seasonality
Bio 16	Precipitation of Wettest Quarter
Bio 17	Precipitation of Driest Quarter
Bio 18	Precipitation of Warmest Quarter
Bio 19	Precipitation of Coldest Quarter

For the bioclimatic variables, the monthly temperature (T) and precipitation (P), as well as the daily maximum (Tx) and minimum temperature (Tn) are required. To estimate the climate conditions, a dataset of at least 30 years is required. The following calculations have to be done for each grid point individually (Klein et al., 2020).

Annual Mean Temperature

$$\text{Bio } 1 = \frac{1}{30} \sum_{i=1}^{30} \left(\frac{1}{12} \sum_{j=1}^{12} T((i-1) \cdot 12 + j) \right) \tag{8.1}$$

Mean Diurnal Range

$$\text{Bio } 2 = \frac{1}{10958} \sum_{i=1}^{10958} (Tx(i) - Tn(i)) \tag{8.2}$$

Isothermality

$$\text{Bio } 3 = \text{BIO } 2/\text{BIO } 7 \cdot 100 \tag{8.3}$$

Temperature Seasonality

$$\text{Bio } 4 = \frac{1}{30} \sum_{i=1}^{30} \sqrt{\frac{1}{12} \sum_{j=1}^{12} (T((i-1) \cdot 12 + j) - \overline{T_i})^2}, \tag{8.4}$$

while $\overline{T_i}$ is the mean Temperature of the year i.

Max Temperature of Warmest Month

Txm is the monthly mean maximum temperature. For $j \in [1, 12]$:

$$\text{Bio } 5 = \max_j \left(\frac{1}{30} \sum_{i=1}^{30} (\text{Txm}((i-1) \cdot 12 + j) \right) \tag{8.5}$$

Min Temperature of Coldest Month

Tnm is the monthly mean minimum temperature. For $j \in [1, 12]$:

$$\text{Bio } 6 = \min_j \left(\frac{1}{30} \sum_{i=1}^{30} (\text{Tnm}((i-1) \cdot 12 + j)) \right) \tag{8.6}$$

Temperature Annual Range

$$\text{Bio } 7 = \text{BIO } 5 - \text{BIO } 6 \tag{8.7}$$

Mean Temperature of Wettest Quarter

For each month the precipitation of the two following months are added to get 12 different quarterly precipitations (PQ). For the quarterly precipitation of November and December the precipitation of the consecutive months from the following year have been added. In the last

year the precipitation of January and February of the same year is taken instead. Then, the quarterly precipitation is averaged over the 30 years.

$$PQ_k = \frac{1}{30} \sum_{i=1}^{30} \left(\sum_{j=k}^{k+2} P\left((i-1) \cdot 12 + j\right) \right), \text{ with } k \in [1, 12], \tag{8.8}$$

for $i = 30$ and $k = 11$ and 12 as described in the text.
The quarterly temperature (TQ) is calculated accordingly, however taking the average over 3 month instead of accumulating.

$$TQ_k = \frac{1}{30} \sum_{i=1}^{30} \left(\frac{1}{3} \sum_{j=k}^{k+2} T\left((i-1) \cdot 12 + j\right) \right), \text{ with } k \in [1, 12], \tag{8.9}$$

for $i = 30$ and $k = 11$ and 12 as described in the text.

For $k \in [1, 12]$, where: $PQ_m = \max_k PQ_k$

$$\text{Bio } 8 = TQ_m \tag{8.10}$$

Mean Temperature of Driest Quarter

For $k \in [1, 12]$, where: $PQ_m = \min_k PQ_k$

$$\text{Bio } 9 = TQ_m \tag{8.11}$$

Mean Temperature of Warmest Quarter

For $k \in [1, 12]$, where: $TQ_m = \max_k TQ_k$

$$\text{Bio } 10 = TQ_m \tag{8.12}$$

Mean Temperature of Coldest Quarter

For $k \in [1, 12]$, where: $TQ_m = \min_k TQ_k$

$$\text{Bio } 11 = TQ_m \tag{8.13}$$

Annual Precipitation

$$\text{Bio } 12 = \frac{1}{30}\sum_{i=1}^{30}\left(\sum_{j=1}^{12} P((i-1)\cdot 12 + j)\right) \tag{8.14}$$

Precipitation of Wettest Month

$$\text{Bio } 13 = \frac{1}{30}\sum_{i=1}^{30}\max_{j}(P((i-1)\cdot 12 + j)), \quad \text{with } j \in [1,12], \tag{8.15}$$

Precipitation of Driest Month

$$\text{Bio } 14 = \frac{1}{30}\sum_{i=1}^{30}\min_{j}(P((i-1)\cdot 12 + j)), \text{ with } j \in [1,12], \tag{8.16}$$

Precipitation Seasonality

$$\text{Bio } 15 = \frac{1}{30}\sum_{i=1}^{30}\left(\frac{\sqrt{\frac{1}{12}\sum_{j=1}^{12}(P((i-1)\cdot 12 + j) - \overline{P_i})^2}}{\overline{P_i}}\right), \tag{8.17}$$

while $\overline{P_i}$ is the mean precipitation of the year i.

Precipitation of Wettest Quarter

For $k \in [1,12]$, where: $PQ_m = \max_k PQ_k$

$$\text{Bio } 16 = PQ_m \tag{8.18}$$

Precipitation of Driest Quarter

For $k \in [1,12]$, where: $PQ_m = \min_k PQ_k$

$$\text{Bio } 17 = PQ_m \tag{8.19}$$

Precipitation of Warmest Quarter

For $k \in [1,12]$, where: $TQ_m = \max_k TQ_k$

$$\text{Bio } 18 = PQ_m \tag{8.20}$$

Precipitation of Coldest Quarter

For $k \in [1, 12]$, where: $TQ_m = \min_k TQ_k$

$$\text{Bio } 19 = PQ_m \tag{8.21}$$

Additional Figures

The following Figures are mentioned within this work and are included here for the sake of completeness. This includes: the results of the Gravettian D-O experiment which shows mostly the same behavior as the Aurignacian but with a different HEP distribution; the total population sum over the complete simulation of the Out of Europe experiment; dispersal fluxes for the Gravettian in the later stages of the Out of Europe experiment, lacking the network analysis.

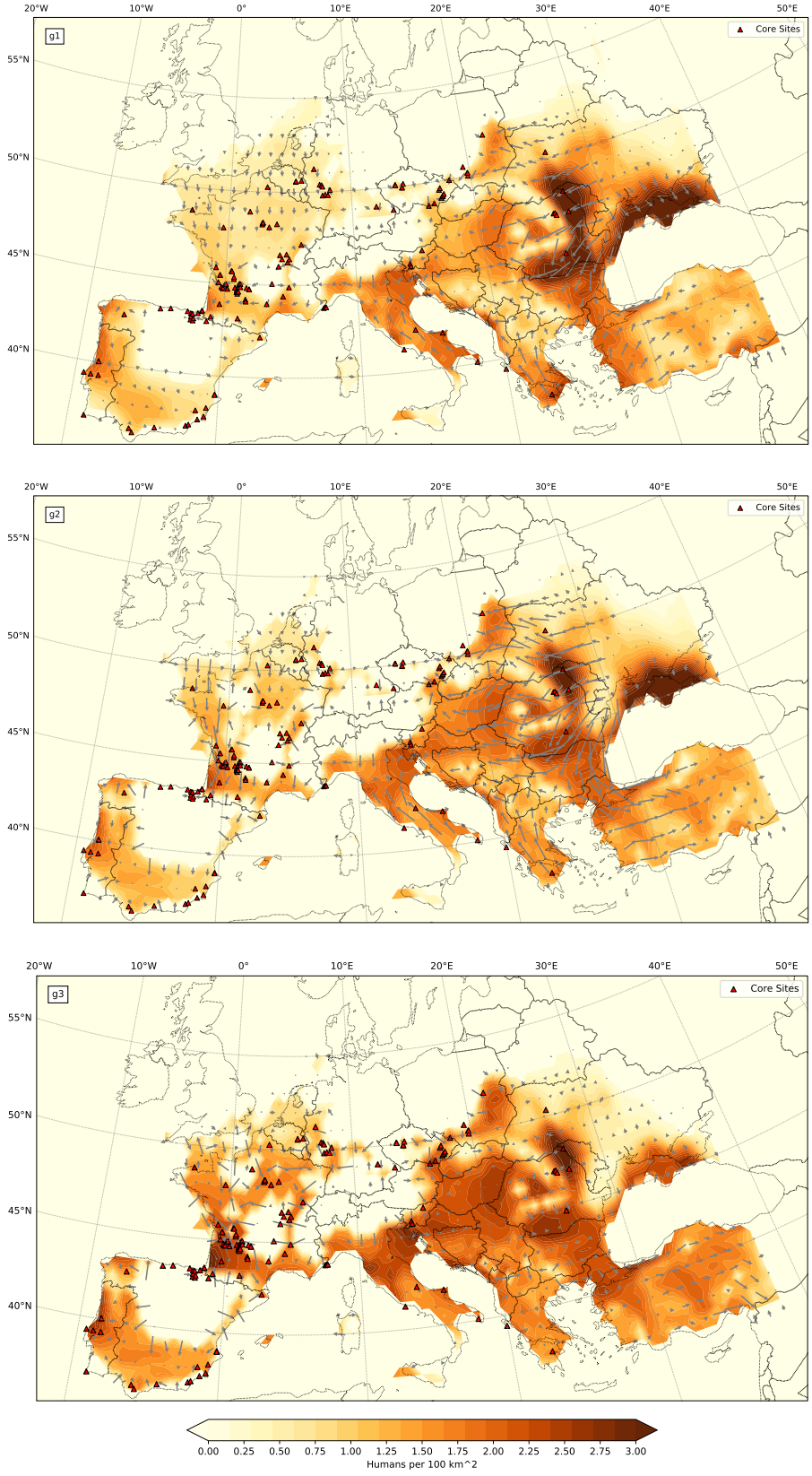

FIGURE 8.1: Density distribution (colored) and dispersal direction and strength (arrows) for the Gravettian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

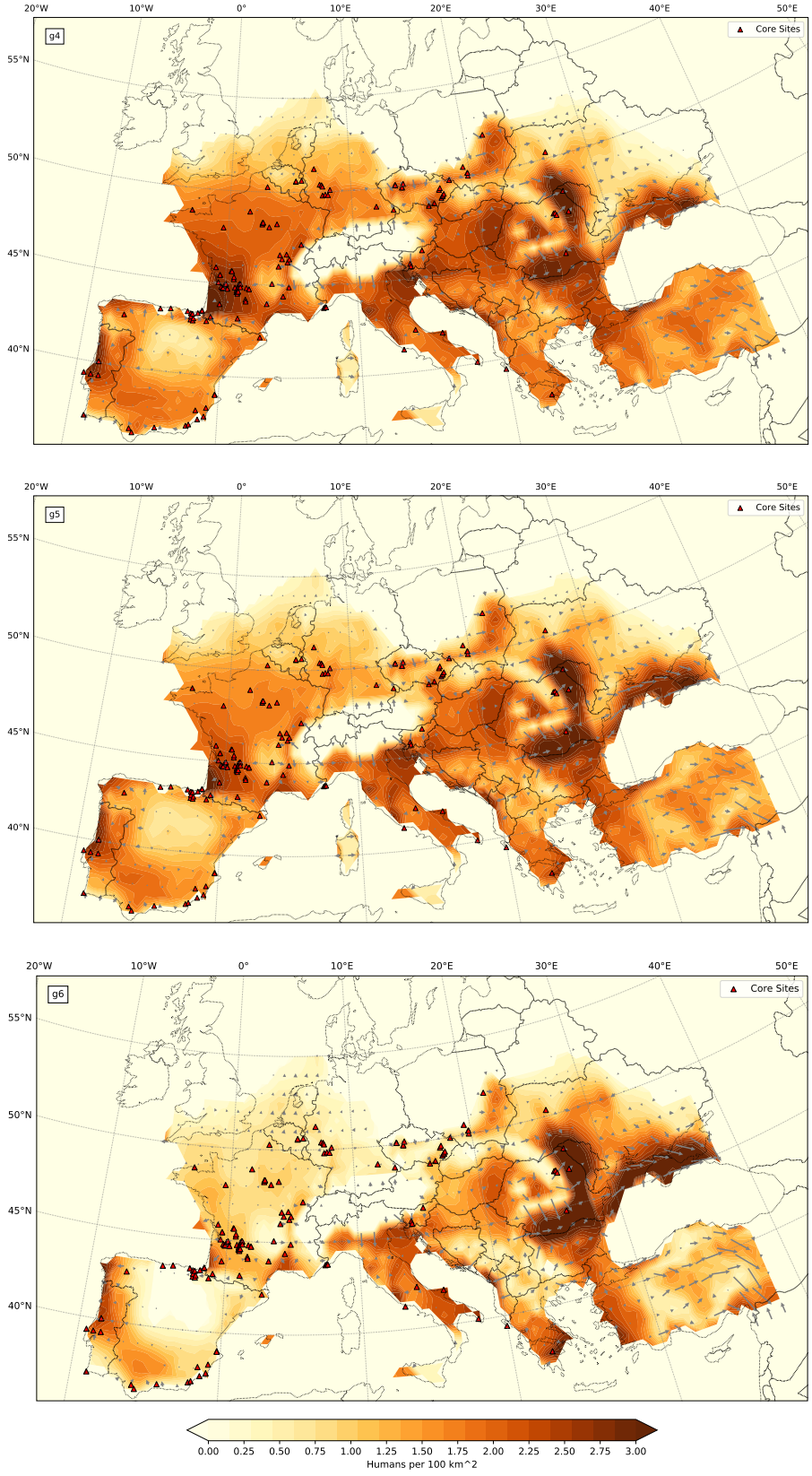

FIGURE 8.2: Density distribution (colored) and dispersal direction and strength (arrows) for the Gravettian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

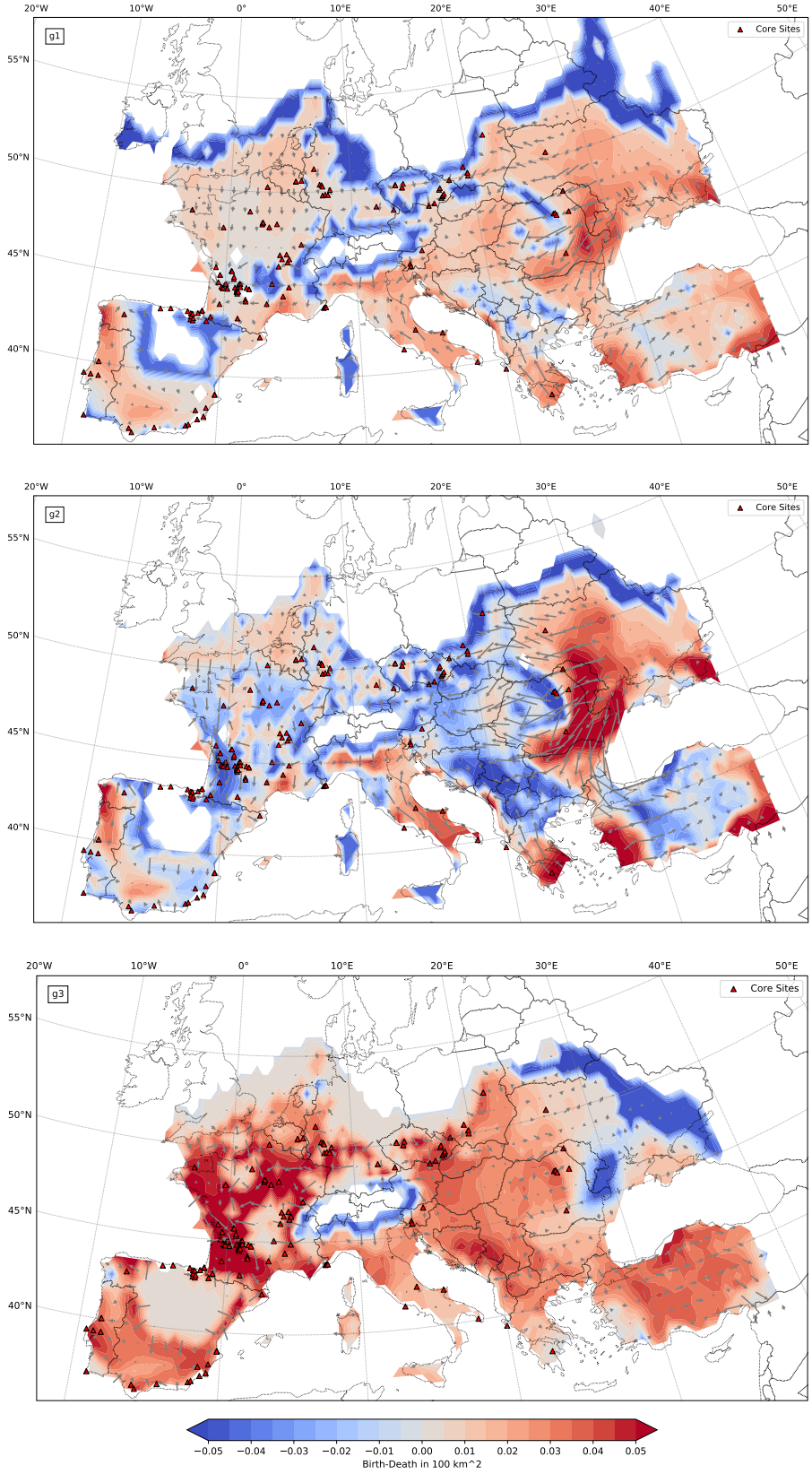

FIGURE 8.3: Birth-Death Term result (colored) and dispersal direction and strength (arrows) for the Gravettian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

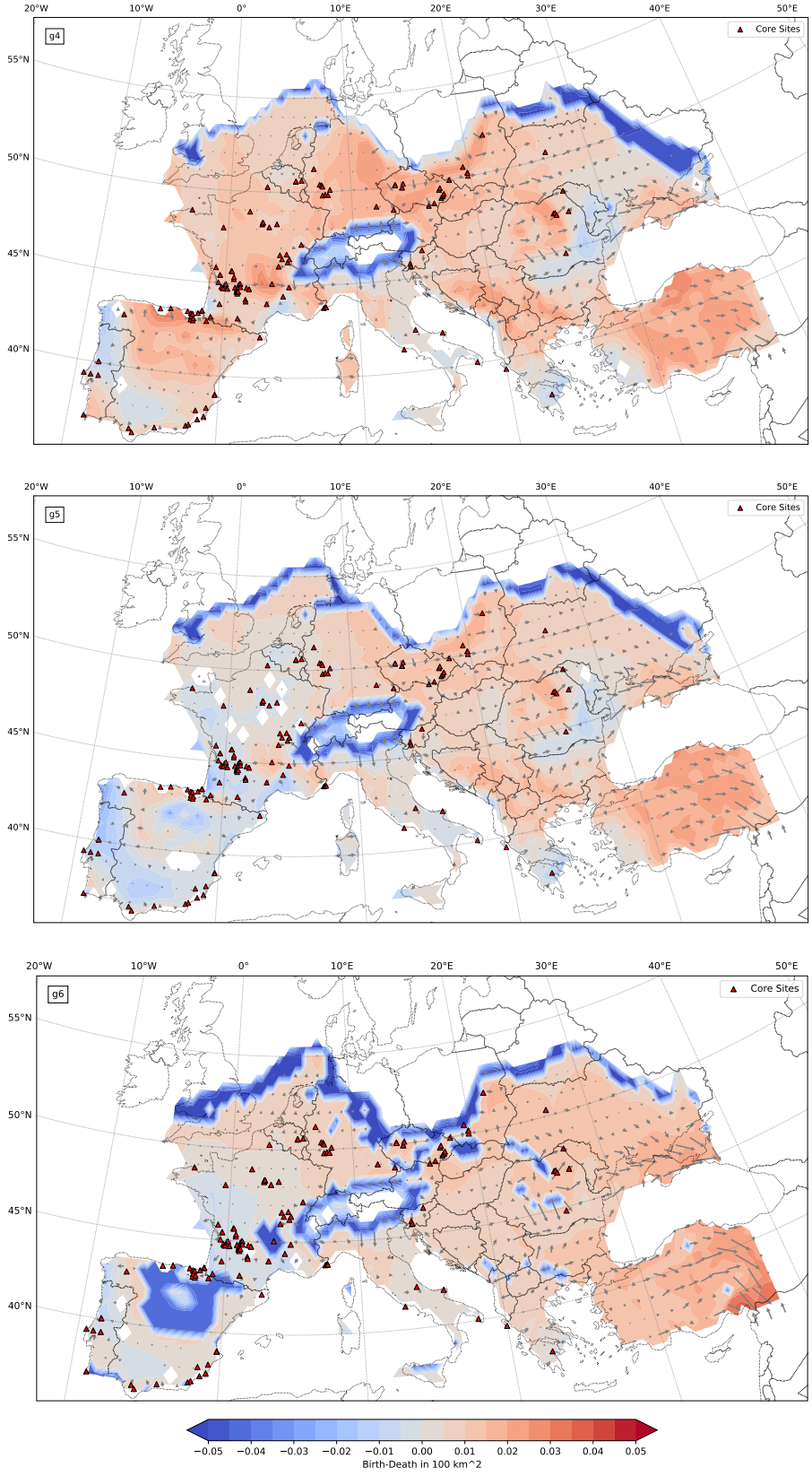

FIGURE 8.4: Birth-Death Term result (colored) and dispersal direction and strength (arrows) for the Gravettian D-O experiment. The time slices follow the pattern as described in the text and Figure 5.5.

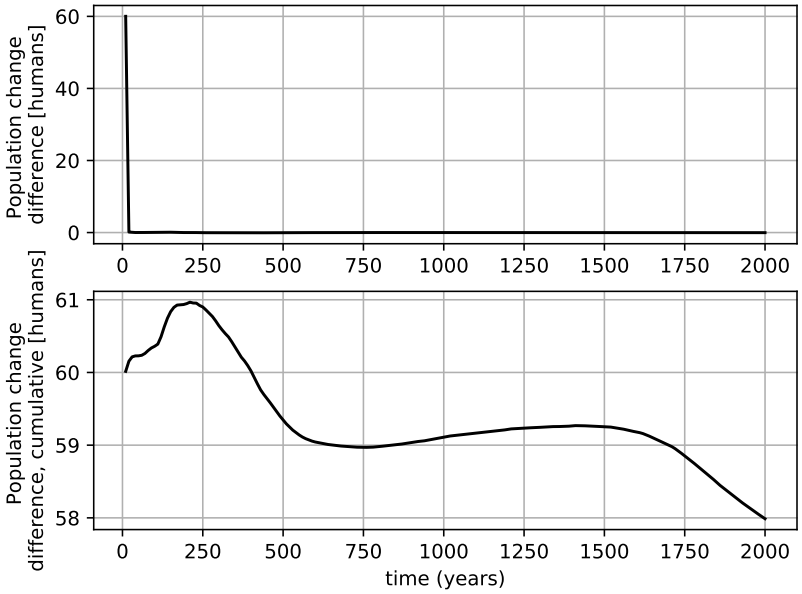

FIGURE 8.5: Difference of the sums of the population change and the population dynamics for the Gravettian D-O experiment, see Equation 4.1. The top shows the difference for every time step, the bottom the cumulative difference over time.

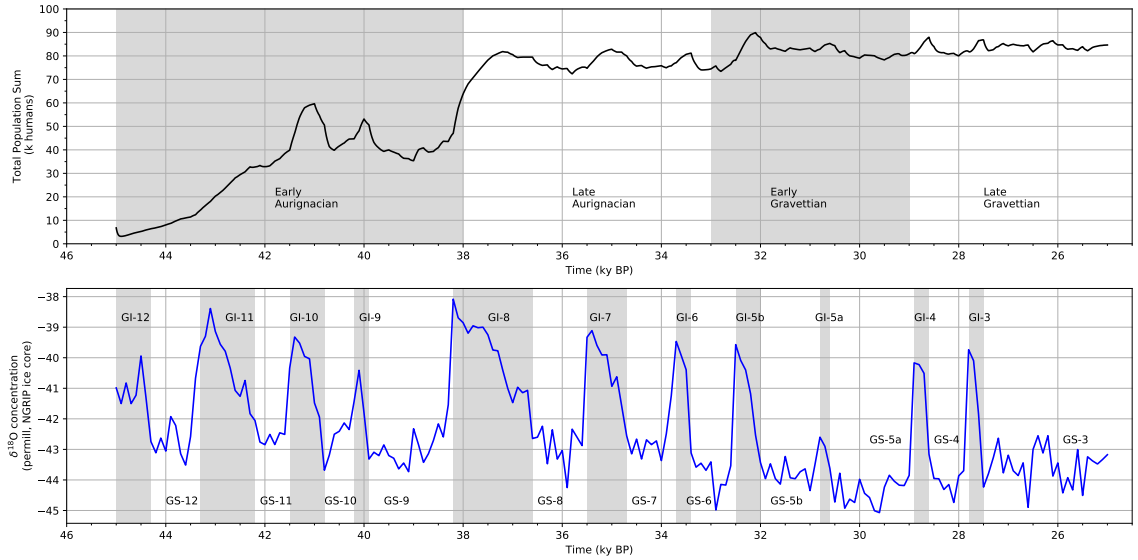

FIGURE 8.6: Total population sum of the domain of the Way to Europe experiment (excluding northern Africa) with the temporal frames of the used technomplexes (top) and, for comparison, the NGRIP δ^{18}O concentration with numbered glacial stadial and interstadial time frames.

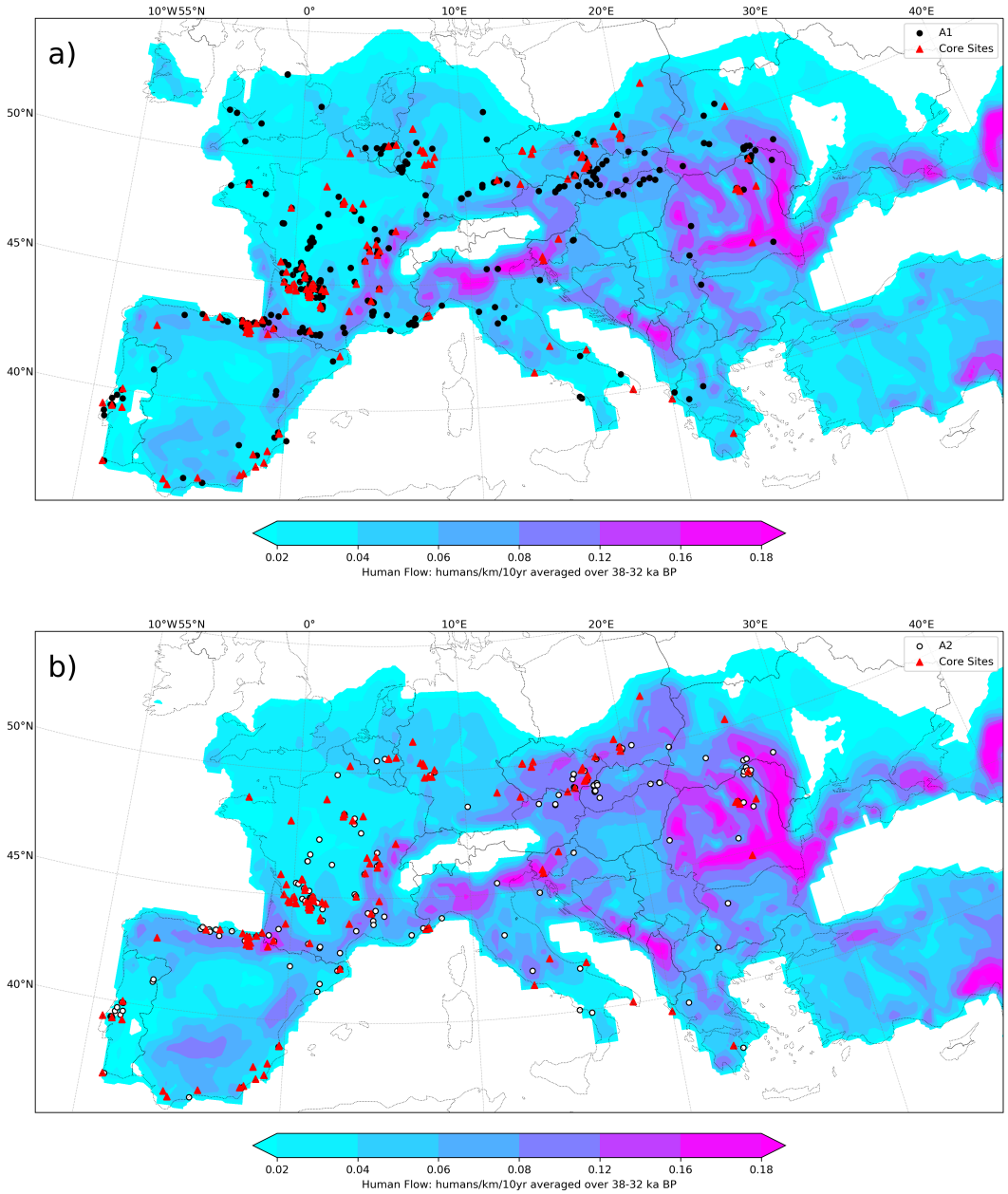

FIGURE 8.7: Population flux in [humans 100 km-1 10 yr-1] averaged over the 33-29 kyr BP period (approximately Gravettian-P1). (b) as (a), but averaged over the 29-25 kyr BP period (approximately Gravettian-P2).

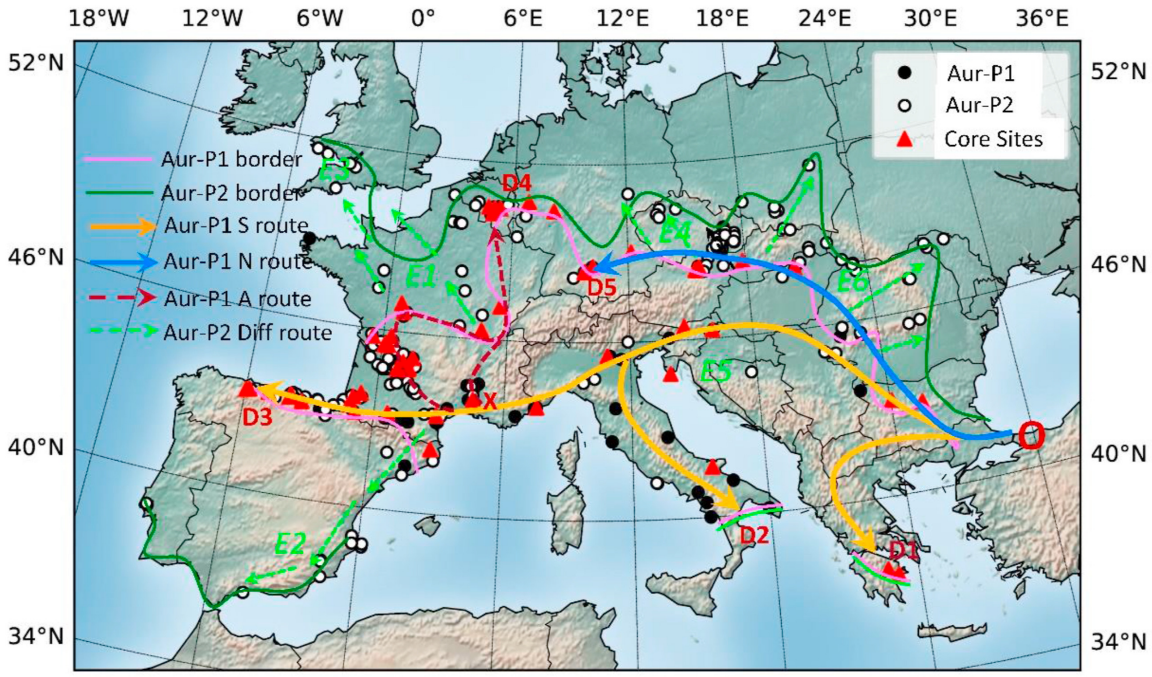

FIGURE 8.8: Dispersal routes derived from the presented HEP distributions. Includes the estimated borders of occupation for both Aurignacian Phases, the likely migration routes corresponding to the first phase, and the routes of expansion during the second phase. See Shao et al. (2021) for a full discussion and explanation.

Bibliography

C. J. Bae, W. Wang, J. Zhao, S. Huang, F. Tian, and G. Shen. Modern human teeth from Late Pleistocene Luna cave (Guangxi, China). *Quaternary International*, 354:169–183, dec 2014. ISSN 10406182. doi:10.1016/j.quaint.2014.06.051.

W. E. Banks, F. D'Errico, A. T. Peterson, M. Vanhaeren, M. Kageyama, P. Sepulchre, G. Ramstein, A. Jost, and D. Lunt. Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. *Journal of Archaeological Science*, 35(2): 481–491, 2008. ISSN 10959238. doi:10.1016/j.jas.2007.05.011.

W. E. Banks, F. D'Errico, and J. Zilhão. Human-climate interaction during the Early Upper Paleolithic: Testing the hypothesis of an adaptive shift between the Proto-Aurignacian and the Early Aurignacian. *Journal of Human Evolution*, 64(1):39–55, jan 2013. ISSN 00472484. doi:10.1016/j.jhevol.2012.10.001.

P. J. Bartlein, S. P. Harrison, S. Brewer, S. Connor, B. A. Davis, K. Gajewski, J. Guiot, T. I. Harrison-Prentice, A. Henderson, O. Peyron, I. C. Prentice, M. Scholze, H. Seppä, B. Shuman, S. Sugita, R. S. Thompson, A. E. Viau, J. Williams, and H. Wu. Pollen-based continental climate reconstructions at 6 and 21 ka: A global synthesis. *Climate Dynamics*, 37(3):775–802, aug 2011. ISSN 09307575. doi:10.1007/s00382-010-0904-1.

S. Benazzi, K. Douka, C. Fornai, C. Bauer, O. Kullmer, J. Svoboda, I. Pap, F. Mallegni, P. Bayle, M. Coquerelle, S. Condemi, A. Ronchitelli, K. Harvati, and G. W Weber. Early Dispersal of Modern Humans in Europe and Implications for Neanderthal Behaviour. *Nature*, 479: 525–528, 2011. doi:10.1038/nature10617.

B. M. Benito, J. C. Svenning, T. Kellberg-Nielsen, F. Riede, G. Gil-Romera, T. Mailund, P. C. Kjaergaard, and B. S. Sandel. The ecological niche and distribution of Neanderthals during the Last Interglacial. *Journal of Biogeography*, 44(1):51–61, 2017. ISSN 13652699. doi:10.1111/jbi.12845.

A. Bergström, C. Stringer, M. Hajdinjak, E. M. Scerri, and P. Skoglund. Origins of modern human ancestry. *Nature*, 590(7845):229–237, feb 2021. ISSN 14764687. doi:10.1038/s41586-021-03244-5.

J. P. Bocquet-Appel, P. Y. Demars, L. Noiret, and D. Dobrowsky. Estimates of Upper Palaeolithic meta-population size in Europe from archaeological data. *Journal of Archaeological Science*, 32 (11):1656–1668, nov 2005. ISSN 10959238. doi:10.1016/j.jas.2005.05.006.

R. Boria, L. Olson, S. Goodman, and R. Anderson. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275:73–77, 2014. doi:10.1016/j.ecolmodel.2013.12.012.

M. D. Bosch, M. A. Mannino, A. L. Prendergast, T. C. O'Connell, B. Demarchi, S. M. Taylor, L. Niven, J. Van Der Plicht, and J. J. Hublin. New chronology for Ksâr 'Akil (Lebanon) supports Levantine route of modern human dispersal into Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 112(25):7683–7688, jun 2015. ISSN 10916490. doi:10.1073/pnas.1501529112.

J. M. Bowler, R. Jones, H. Allen, and A. G. Thorne. Pleistocene human remains from Australia: A living site and human cremation from lake mungo, western new south wales. *World Archaeology*, 2(1):39–60, 1970. ISSN 14701375. doi:10.1080/00438243.1970.9979463.

P. Braconnot, S. P. Harrison, M. Kageyama, P. J. Bartlein, V. Masson-Delmotte, A. Abe-Ouchi, B. Otto-Bliesner, and Y. Zhao. Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, 2(6):417–424, jun 2012. doi:10.1038/nclimate1456.

A. Budsky, J. A. Wassenburg, R. Mertz-Kraus, C. Spötl, K. P. Jochum, L. Gibert, and D. Scholz. Western Mediterranean Climate Response to Dansgaard/Oeschger Events: New Insights From Speleothem Records. *Geophysical Research Letters*, 46(15):9042–9053, aug 2019. ISSN 19448007. doi:10.1029/2019GL084009.

A. Burke, G. Levavasseur, P. M. A. James, D. Guiducci, M. A. Izquierdo, L. Bourgeon, M. Kageyama, G. Ramstein, and M. Vrac. Exploring the impact of climate variability during the Last Glacial Maximum on the pattern of human occupation of Iberia. *Journal of Human Evolution*, 73:35–46, 2014. ISSN 0047-2484. doi:https://doi.org/10.1016/j.jhevol.2014.06.003.

A. Burke, M. Kageyama, G. Latombe, M. Fasel, M. Vrac, G. Ramstein, and P. M. James. Risky business: The impact of climate and climate variability on human population dynamics in Western Europe during the Last Glacial Maximum. *Quaternary Science Reviews*, 164:217–229, may 2017. ISSN 02773791. doi:10.1016/j.quascirev.2017.04.001.

D. Campos, J. Fort, and V. Méndez. Transport on fractal river networks: Application to migration fronts. *Theoretical Population Biology*, 69(1):88–93, feb 2006. ISSN 00405809. doi:10.1016/j.tpb.2005.09.001.

M. Cortés-Sánchez, F. J. Jiménez-Espejo, M. D. Simón-Vallejo, C. Stringer, M. C. Lozano Francisco, A. García-Alix, J. L. Vera Peláez, C. P. Odriozola, J. A. Riquelme-Cantal, R. Parrilla Giráldez, A. Maestro González, N. Ohkouchi, and A. Morales-Muñiz. An early Aurignacian

arrival in southwestern Europe. *Nature Ecology and Evolution*, 3(2):207–212, feb 2019. ISSN 2397334X. doi:10.1038/s41559-018-0753-6.

CRC806. Research profile of the {Collaborative Research Center 806 (Phase 3)}, 2021. URL https://www.sfb806.uni-koeln.de/index.php/about?showall=1.

F. Cruciani, B. Trombetta, A. Massaia, G. Destro-Bisol, D. Sellitto, and R. Scozzari. A revised root for the human y chromosomal phylogenetic tree: The origin of patrilineal diversity in Africa. *American Journal of Human Genetics*, 88(6):814–818, jun 2011. ISSN 00029297. doi:10.1016/j.ajhg.2011.05.002.

K. Davison, P. Dolukhanov, G. Sarson, and A. Shukurov. The role of waterways in the spread of the Neolithic. *jarcs*, 33:641–, 2006. doi:10.1016/j.quascirev.2017.04.001.

O. Deshpande, S. Batzoglou, M. W. Feldman, and L. L. Cavalli-Sforza. A serial founder effect model for human settlement out of Africa. *Proceedings of the Royal Society B: Biological Sciences*, 276(1655):291–300, jan 2009. ISSN 14712970. doi:10.1098/rspb.2008.0750.

R. Dinnis. The timing of Aurignacian occupation of the British Peninsula. *Quartar*, 59:67–83, 2012. ISSN 03757471. doi:10.7485/QU59-3.

C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. Mcclean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1): 27–46, jan 2013. ISSN 16000587. doi:10.1111/j.1600-0587.2012.07348.x.

A. K. Ekholm. CMIP5 Community Storage Server, 2013. URL http://cmip5.whoi.edu:8080/data/CMIP5/output1/CNRM-CERFACS/CNRM-CM5/lgm/mon/atmos/Amon/r1i1p1/.

A. Eriksson, L. Betti, A. D. Friend, S. J. Lycett, J. S. Singarayer, N. von Cramon-Taubadel, P. J. Valdes, F. Balloux, and A. Manica. Late Pleistocene climate change and the global expansion of anatomically modern humans. *Proceedings of the National Academy of Sciences*, 109(40): 16089–16094, 2012. ISSN 0027-8424. doi:10.1073/pnas.1209494109.

K. Fadhlaoui-Zid, M. Haber, B. Martínez-Cruz, P. Zalloua, A. B. Elgaaied, and D. Comas. Genome-wide and paternal diversity reveal a recent origin of human populations in north africa. *PLoS ONE*, 8(11):e80293, nov 2013. ISSN 19326203. doi:10.1371/journal.pone.0080293.

H. Fewlass, S. Talamo, L. Wacker, B. Kromer, T. Tuna, Y. Fagault, E. Bard, S. P. McPherron, V. Aldeias, R. Maria, N. L. Martisius, L. Paskulin, Z. Rezek, V. Sinet-Mathiot, S. Sirakova, G. M. Smith, R. Spasov, F. Welker, N. Sirakov, T. Tsanova, and J. J. Hublin. A 14C chronology for the Middle to Upper Palaeolithic transition at Bacho Kiro Cave, Bulgaria. *Nature Ecology and Evolution*, 4(6):794–801, jun 2020. ISSN 2397334X. doi:10.1038/s41559-020-1136-3.

P. Forster. EVOLUTION: Enhanced: Did Early Humans Go North or South? *Science*, 308(5724): 965–966, may 2005. doi:10.1126/science.1113261.

J. Fort, T. Pujol, and L. L. Cavalli-Sforza. Palaeolithic Populations and Waves of Advance. *Cambridge Archaeological Journal*, 14(1):53–61, 2004. doi:10.1017/S0959774304000046.

H. S. Groucutt, M. D. Petraglia, G. Bailey, E. M. Scerri, A. Parton, L. Clark-Balzan, R. P. Jennings, L. Lewis, J. Blinkhorn, N. A. Drake, P. S. Breeze, R. H. Inglis, M. H. Devès, M. Meredith-Williams, N. Boivin, M. G. Thomas, and A. Scally. Rethinking the dispersal of Homo sapiens out of Africa. *Evolutionary Anthropology*, 24(4):149–164, jul 2015. ISSN 15206505. doi:10.1002/evan.21455.

T. P. Guilderson. GEOSCIENCE: The Boon and Bane of Radiocarbon Dating. *Science*, 307(5708): 362–364, jan 2005. doi:10.1126/science.1104164.

G. Guillera-Arroita, J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentini, M. A. Mccarthy, R. Tingley, and B. A. Wintle. Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3):276–292, mar 2015. ISSN 14668238. doi:10.1111/geb.12268.

R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, 5(10):1000695, oct 2009. ISSN 15537390. doi:10.1371/journal.pgen.1000695.

M. J. Hamilton and B. Buchanan. Spatial gradients in Clovis-age radiocarbon dates across North America suggest rapid colonization from the north. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40):15625–15630, oct 2007. ISSN 00278424. doi:10.1073/pnas.0704215104.

J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982. ISSN 00338419. doi:10.1148/radiology.143.1.7063747.

R. M. Harding and G. McVean. A structured ancestral population for the evolution of modern humans. *Current Opinion in Genetics and Development*, 14(6):667–674, dec 2004. ISSN 0959437X. doi:10.1016/j.gde.2004.08.010.

B. M. Henn, C. R. Gignoux, M. Jobin, J. M. Granka, J. M. MacPherson, J. M. Kidd, L. Rodríguez-Botigué, S. Ramachandran, L. Hon, A. Brisbini, A. A. Lin, P. A. Underhill, D. Comas, K. K. Kidd, P. J. Norman, P. Parham, C. D. Bustamante, J. L. Mountain, and M. W. Feldman. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(13):5154–5162, mar 2011. ISSN 10916490. doi:10.1073/pnas.1017511108.

J. Hernández-Orallo, P. Flach, and C. Ferri. Brier curves: A new cost-based visualisation of classifier performance. In *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, pages 585–592, 2011. ISBN 9781450306195.

E. Hovers. *The lithic assemblages of Qafzeh Cave*. 2009. ISBN 9780195322774. URL https://www.academia.edu/25029303/The{_}Lithic{_}Assemblages{_}of{_}Qafzeh{_}Cave{_}text{_}and{_}graphs.

J.-J. Hublin, A. Ben-Ncer, S. E. Bailey, S. E. Freidline, S. Neubauer, M. M. Skinner, I. Bergmann, A. Le Cabec, S. Benazzi, K. Harvati, and P. Gunz. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature*, 546:289, 2017. doi:10.1038/nature22336.

R. L. Kelly. *The lifeways of hunter-gatherers: The foraging spectrum*. 2010. ISBN 9781139176132. doi:10.1017/CBO9781139176132.

H. L. Kim, A. Ratan, G. H. Perry, A. Montenegro, W. Miller, and S. C. Schuster. Khoisan hunter-gatherers have been the largest population throughout most of modern-human demographic history. *Nature Communications*, 5, 2014. ISSN 20411723. doi:10.1038/ncomms6692.

K. Klein, C. Wegener, I. Schmidt, M. Rostami, P. Ludwig, S. Ulbrich, J. Richter, G.-C. Weniger, and Y. Shao. Human existence potential in Europe during the Last Glacial Maximum. *Quaternary International*, aug 2020. ISSN 10406182. doi:10.1016/j.quaint.2020.07.046.

R. G. Klein. The human career: human biological and cultural origins. *Choice Reviews Online*, 27(09):27–5183–27–5183, 1990. ISSN 0009-4978. doi:10.5860/choice.27-5183.

Y. Kondo, K. Sano, T. Omori, A. Abe-Ouchi, W.-L. Chan, S. Kadowaki, M. Naganuma, R. O'ishi, T. Oguchi, Y. Nishiaki, and M. Yoneda. Ecological Niche and Least-Cost Path Analyses to Estimate Optimal Migration Routes of Initial Upper Palaeolithic Populations to Eurasia. *The Middle and Upper Paleolithic Archeology of the Levant and Beyond*, pages 199–212, 2018. doi:10.1007/978-981-10-6826-3_13.

A. Kurganov and E. Tadmor. New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection–Diffusion Equations. *Journal of Computational Physics*, 160:241–282, 2000. doi:10.1006/jcph.2000.6459.

M. W. Lake. Trends in Archaeological Simulation. *Journal of Archaeological Method and Theory*, 21(2):258–287, 2014. ISSN 15737764. doi:10.1007/s10816-013-9188-1.

C. Li and A. Born. Coupled atmosphere-ice-ocean dynamics in Dansgaard-Oeschger events. *Quaternary Science Reviews*, 203:1–20, jan 2019. ISSN 02773791. doi:10.1016/j.quascirev.2018.10.031.

H. Li and R. Durbin. Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357):493–496, jul 2011. ISSN 00280836. doi:10.1038/nature10231.

W. Liu, M. Martinón-Torres, Y. J. Cai, S. Xing, H. W. Tong, S. W. Pei, M. J. Sier, X. H. Wu, R. L. Edwards, H. Cheng, Y. Y. Li, X. X. Yang, J. M. B. De Castro, and X. J. Wu. The earliest unequivocally modern humans in southern China. *Nature*, 526(7575):696–699, oct 2015. ISSN 14764687. doi:10.1038/nature15696.

S. López, L. van Dorp, and G. Hellenthal. Human dispersal out of Africa: A lasting debate. *Evolutionary Bioinformatics*, 11:57–68, 2015. ISSN 11769343. doi:10.4137/EBo.s33489.

P. Ludwig, E. J. Schaffernicht, Y. Shao, and J. G. Pinto. Journal of Geophysical Research : Atmospheres. pages 1–16, 2016. doi:10.1002/2015JD024444.Received.

A. Maier and A. Zimmermann. Population Dynamics: Demographic Changes of Hunter-Gatherer Populations during the Upper Pleistocene and Early Holocene in Europe (LGM), 2015. URL `https://crc806db.uni-koeln.de/dataset/show/crc806e1lgmsitesdatabase20150313142839605 9/`.

A. Maier and A. Zimmermann. Populations headed south? The Gravettian from a palaeodemographic point of view. *Antiquity*, 91(357):573–588, 2017. ISSN 0003-598X. doi:10.15184/aqy.2017.37.

A. Maier, F. Lehmkuhl, P. Ludwig, M. Melles, I. Schmidt, Y. Shao, C. Zeeden, and A. Zimmermann. Demographic estimates of hunter–gatherers during the Last Glacial Maximum in Europe against the background of palaeoenvironmental data. *Quaternary International*, 425: 49–61, 2016. ISSN 10406182. doi:10.1016/j.quaint.2016.04.009.

A. B. Marín-Arroyo, J. Rios-Garaizar, L. G. Straus, J. R. Jones, M. de la Rasilla, M. R. González Morales, M. Richards, J. Altuna, K. Mariezkurrena, and D. Ocio. Chronological reassessment of the Middle to Upper Paleolithic transition and Early Upper Paleolithic cultures in Cantabrian Spain. *PLoS ONE*, 13(4):e0194708, apr 2018. ISSN 19326203. doi:10.1371/journal.pone.0194708.

B. R. Markle, E. J. Steig, C. Buizert, S. W. Schoenemann, C. M. Bitz, T. J. Fudge, J. B. Pedro, Q. Ding, T. R. Jones, J. W. White, and T. Sowers. Global atmospheric teleconnections during Dansgaard-Oeschger events. *Nature Geoscience*, 10(1):36–40, jan 2017. ISSN 17520908. doi:10.1038/ngeo2848.

L. A. Martino, A. Osella, C. Dorso, and J. L. Lanata. Fisher equation for anisotropic diffusion: Simulating South American human dispersals. *Phys. Rev. E*, 76(3):31923, 2007. doi:10.1103/PhysRevE.76.031923.

R. C. McCarthy and L. Lucas. A morphometric re-assessment of BOU-VP-16/1 from Herto, Ethiopia. *Journal of Human Evolution*, 74:114–117, 2014. ISSN 00472484. doi:10.1016/j.jhevol.2014.05.011.

F. McDermott, R. Grün, C. B. Stringer, and C. J. Hawkesworth. Mass-spectrometric U-series dates for Israeli Neanderthal/early modern hominid sites. *Nature*, 363(6426):252–255, 1993. ISSN 00280836. doi:10.1038/363252a0.

I. McDougall, F. H. Brown, and J. G. Fleagle. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature*, 433(7027):733–736, feb 2005. ISSN 00280836. doi:10.1038/nature03258.

B. P. McEvoy, J. E. Powell, M. E. Goddard, and P. M. Visscher. Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. *Genome Research*, 21(6):821–829, jun 2011. ISSN 10889051. doi:10.1101/gr.119636.110.

P. Mellars. Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (25):9381–9386, jun 2006. ISSN 00278424. doi:10.1073/pnas.0510792103.

P. Mellars. The earliest modern humans in Europe, nov 2011. ISSN 00280836.

M. Mirazón Lahr and R. A. Foley. Towards a theory of modern human origins: Geography, demography, and diversity in recent human evolution. *American Journal of Physical Anthropology*, 107(S27):137–176, jan 1998. ISSN 00029483. doi:10.1002/(sici)1096-8644(1998)107:27+<137::aid-ajpa6>3.3.co;2-h.

D. Müllner. Fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. *Journal of Statistical Software*, 53(9):1–18, may 2013. ISSN 15487660. doi:10.18637/jss.v053.i09.

I. Obreht, U. Hambach, D. Veres, C. Zeeden, J. Bösken, T. Stevens, S. B. Marković, N. Klasen, D. Brill, C. Burow, and F. Lehmkuhl. Shift of large-scale atmospheric systems over Europe during late MIS 3 and implications for Modern Human dispersal. *Scientific Reports*, 7(1):1–10, dec 2017. ISSN 20452322. doi:10.1038/s41598-017-06285-x.

S. Oppenheimer. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1590):770–784, 2012. ISSN 0962-8436. doi:10.1098/rstb.2011.0306.

L. Pagani, S. Schiffels, D. Gurdasani, P. Danecek, A. Scally, Y. Chen, Y. Xue, M. Haber, R. Ekong, T. Oljira, E. Mekonnen, D. Luiselli, N. Bradman, E. Bekele, P. Zalloua, R. Durbin, T. Kivisild, and C. Tyler-Smith. Tracing the Route of Modern Humans out of Africa by Using 225 Human Genome Sequences from Ethiopians and Egyptians. *American Journal of Human Genetics*, 96 (6):986–991, jun 2015. ISSN 15376605. doi:10.1016/j.ajhg.2015.04.019.

J. K. Pickrell, N. Patterson, C. Barbieri, F. Berthold, L. Gerlach, T. Güldemann, B. Kure, S. W. Mpoloka, H. Nakagawa, C. Naumann, M. Lipson, P. R. Loh, J. Lachance, J. Mountain, C. D. Bustamante, B. Berger, S. A. Tishkoff, B. M. Henn, M. Stoneking, D. Reich, and B. Pakendorf. The genetic prehistory of southern Africa. *Nature Communications*, 3(1):1–6, oct 2012. ISSN 20411723. doi:10.1038/ncomms2140.

M. R. Rampino and S. Self. Volcanic winter and accelerated glaciation following the Toba super-eruption. *Nature*, 359(6390):50–52, 1992. ISSN 00280836. doi:10.1038/359050a0.

S. O. Rasmussen, M. Bigler, S. P. Blockley, T. Blunier, S. L. Buchardt, H. B. Clausen, I. Cvijanovic, D. Dahl-Jensen, S. J. Johnsen, H. Fischer, V. Gkinis, M. Guillevic, W. Z. Hoek, J. J. Lowe, J. B. Pedro, T. Popp, I. K. Seierstad, J. P. Steffensen, A. M. Svensson, P. Vallelonga, B. M. Vinther, M. J. Walker, J. J. Wheatley, and M. Winstrup. A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy. *Quaternary Science Reviews*, 106:14–28, 2014. ISSN 02773791. doi:10.1016/j.quascirev.2014.09.007.

H. Reyes-Centeno, M. Hubbe, T. Hanihara, C. Stringer, and K. Harvati. Testing modern human out-of-Africa dispersal models and implications for modern human origins. *Journal of Human Evolution*, 87:95–106, oct 2015. ISSN 00472484. doi:10.1016/j.jhevol.2015.06.008.

J. Richter, T. Hauck, R. Vogelsang, T. Widlok, J. M. Le Tensorer, and P. Schmid. "Contextual areas" of early Homo sapiens and their significance for human dispersal from Africa into Eurasia between 200 ka and 70 ka. *Quaternary International*, 274:5–24, 2012. ISSN 10406182. doi:10.1016/j.quaint.2012.04.017.

J. I. Rose. New light on human prehistory in the Arabo-Persian Gulf Oasis. *Current Anthropology*, 51(6):849–883, dec 2010. ISSN 00113204. doi:10.1086/657397.

I. Schmidt and A. Zimmermann. Population Dynamics: Demographic Changes of Hunter-Gatherer Populations during the Upper Pleistocene and Early Holocene in Europe (Aurignacien, Gravettian), 2018. URL https://crc806db.uni-koeln.de/dataset/show/crc806e1eupsitesdatabase201503161428395931.

I. Schmidt and A. Zimmermann. Population dynamics and socio-spatial organization of the Aurignacian: Scalable quantitative demographic data for western and central Europe. *PLoS ONE*, 14(2):1–20, 2019. ISSN 19326203. doi:10.1371/journal.pone.0211562.

Y. Shao, H. Limberg, K. Klein, C. Wegener, I. Schmidt, G. C. Weniger, A. Hense, and M. Rostami. Human-existence probability of the aurignacian techno-complex under extreme climate conditions. *Quaternary Science Reviews*, 263:106995, jul 2021. ISSN 02773791. doi:10.1016/j.quascirev.2021.106995.

W. Shi, Q. Ayub, M. Vermeulen, R. G. Shao, S. Zuniga, K. Van Der Gaag, P. De Knijff, M. Kayser, Y. Xue, and C. Tyler-Smith. A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations. *Molecular Biology and Evolution*, 27(2):385–393, feb 2010. ISSN 07374038. doi:10.1093/molbev/msp243.

T. M. Smith, P. Tafforeau, D. J. Reid, R. Grün, S. Eggins, M. Boutakiout, and J. J. Hublin. Earliest evidence of modern human life history in North African early Homo sapiens. *Proceedings*

of the National Academy of Sciences of the United States of America, 104(15):6128–6133, apr 2007. ISSN 00278424. doi:10.1073/pnas.0700747104.

P. Soares, F. Alshamali, J. B. Pereira, V. Fernandes, N. M. Silva, C. Afonso, M. D. Costa, E. Musilová, V. MacAulay, M. B. Richards, V. Černý, and L. Pereira. The expansion of mtDNA haplogroup L3 within and out of Africa. *Molecular Biology and Evolution*, 29(3):915–927, 2012. ISSN 07374038. doi:10.1093/molbev/msr245.

J. Steele. Human dispersals: mathematical models and the archaeological record. *Human biology; an international record of research*, 81(2-3):121–140, 2009. ISSN 0018-7143. doi:10.3378/027.081.0302.

J. Steele, J. Adams, and T. Sluckin. Modelling Paleoindian dispersals. *World Archaeology*, 30(2): 286–305, 1998. doi:10.1080/00438243.1998.9980411.

M. Tallavaara, M. Luoto, N. Korhonen, H. Järvinen, and H. Seppä. Human population dynamics in Europe over the Last Glacial Maximum. *Proceedings of the National Academy of Sciences*, 112(27):8232–8237, 2015. ISSN 0027-8424. doi:10.1073/pnas.1503784112.

A. R. Templeton. Genetics and recent human evolution. *Evolution*, 61(7):1507–1519, 2007. ISSN 00143820. doi:10.1111/j.1558-5646.2007.00164.x.

A. Timmermann. Quantifying the potential causes of Neanderthal extinction: Abrupt climate change versus competition and interbreeding. *Quaternary Science Reviews*, 238, 2020. ISSN 02773791. doi:10.1016/j.quascirev.2020.106331.

A. Timmermann and T. Friedrich. Late Pleistocene climate drivers of early human migration. *Nature*, 538(7623):92–95, oct 2016. ISSN 0028-0836. doi:10.1038/nature19365.

A. Voldoire, E. Sanchez-Gomez, D. Salas y Mélia, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, M. Chevallier, M. Déqué, J. Deshayes, H. Douville, E. Fernandez, G. Madec, E. Maisonnave, M. P. Moine, S. Planton, D. Saint-Martin, S. Szopa, S. Tyteca, R. Alkama, S. Belamari, A. Braun, L. Coquart, and F. Chauvin. The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dynamics*, 40(9-10):2091–2121, 2013. ISSN 14320894. doi:10.1007/s00382-011-1259-y.

W. S. Watkins, A. R. Rogers, C. T. Ostler, S. Wooding, M. J. Bamshad, A. M. E. Brassington, M. L. Carroll, S. V. Nguyen, J. A. Walker, B. V. Prasad, P. G. Reddy, P. K. Das, M. A. Batzer, and L. B. Jorde. Genetic variation among world populations: Inferences from 100 Alu insertion polymorphisms. *Genome Research*, 13(7):1607–1618, jul 2003. ISSN 10889051. doi:10.1101/gr.894603.

C. Wegener. Human Dispersal Model: Sourcecode, 2021. URL `https://github.com/ChrisWege/HumanDispersalModel`.

E. W. Wolff, J. Chappellaz, T. Blunier, S. O. Rasmussen, and A. Svensson. Millennial-scale variability during the last glacial: The ice core record. *Quaternary Science Reviews*, 29(21-22): 2828–2838, oct 2010. ISSN 02773791. doi:10.1016/j.quascirev.2009.10.013.

G. Zängl, D. Reinert, P. Rípodas, and M. Baldauf. The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141(687):563–579, jan 2015. ISSN 1477870X. doi:10.1002/qj.2378.

A. Zimmermann, J. Richter, T. Frank, and K. P. Wendt. *Landschaftsarchäologie II - Überlegungen zu prinzipien einer landschaftsarchäologie*, volume 85. 2004. ISBN 3805335172.

A. Zimmermann, J. Hilpert, and K. P. Wendt. Estimations of population density for selected periods between the Neolithic and AD 1800. *Human biology; an international record of research*, 81(2-3):357–380, 2009. ISSN 0018-7143. doi:10.3378/027.081.0313.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

K. Klein, C. Wegener, I. Schmidt, M. Rostami, P. Ludwig, S. Ulbrich, J. Richter, G.-C. Weniger, and Y. Shao. Human existence potential in Europe during the Last Glacial Maximum. Quaternary International, aug 2020. ISSN 10406182. doi:10.1016/j.quaint.2020.07.046.

Y. Shao, H. Limberg, K. Klein, C. Wegener, I. Schmidt, G. C. Weniger, A. Hense, and M. Rostami. Human-existence probability of the aurignacian techno-complex under extreme climate conditions. Quaternary Science Reviews, 263:106995, jul 2021. ISSN 02773791. doi:10.1016/j.quascirev.2021.106995.

Christian Wegener, 20.08.2021