International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Address: IIASA, Schlossplatz 1, A-2361 Laxenburg, Austria
Email: repository@iiasa.ac.at **Telephone:** +43 (0)2236 807 342

YSSP Report
Young Scientists Summer Program

Demography and heat stress: the role of population dynamics in climate risk projections

Iulia Marginean, iulia.marginean@cicero.oslo.no

Approved by

Supervisor: Roman Hoffmann, Jesus Crespo Cuaresma, Raya Muttarak, & Jing Gao
Program: Population and Just Societies (POPJUS)
September 30, 2021

This report represents the work completed by the author during the IIASA Young Scientists Summer Program (YSSP) with approval from the YSSP supervisor.

It was finished by *September 30, 2021* and has not been altered or revised since.

This research was funded by IIASA and its National Member Organizations in Africa, the Americas, Asia, and Europe.

ZVR 524808900

Table of Contents

Abstract

The impact of climate change on human health is not distributed evenly with some demographic groups e.g. older adults, people with a low socio-economic status, city inhabitants being more susceptible to increasing temperatures. High-resolution spatial projections of population vulnerability are however scarce. Using Bayesian Model Averaging (BMA), this paper downscales population structure to intra-urban level for the case study of Madrid. First, empirical data from the municipal register of inhabitants are used to map where vulnerable groups reside. Demographic heterogeneity is captured through disaggregating the input data by age groups (below and above 65 years), sex and educational attainment (as a proxy for socio-economic status). Second, BMA is employed to project shares of these groups to 2050, under selected Shared Socioeconomic Pathways.

The resulting distributions of population groups and their projected changes are relevant for local decision-making regarding adaptation strategies to decrease the health burden heat stress might impose in the future.

Acknowledgments

Participating in the YSSP was a fantastic experience, for which I am grateful to many people. First, I would like to thank my supervisors from the Population and Just Societies Programme (POPJUS) Roman Hoffmann, Jesus Crespo Cuaresma, Raya Muttarak and Jing Gao for their warm and encouraging support through the virtual meetings as well as during my stay in Vienna.

A big thank you to Tanja Huber, Aleksandra Cofala, Brian Fath and the YSSP organizing team for making both the virtual programme and our stay in Vienna a fun and memorable experience. Thank you to my fellow YSSPers, especially to the social committee for organizing the online gatherings, to the adaptation group for their commitment and engaging conversations and for all the useful comments and feedback. I would also like to thank Fabian Wagner and all the contributors to the systems analysis reading group.

A special thanks goes to Michaela Potancokova for the warm welcome when I arrived at IIASA, her efforts to organize all the hybrid meetings with the POPJUS YSSPers and for the insightful discussions and feedback on my presentations.

I am grateful to Paloma Yanez for facilitating the access to data from the statististical office of Madrid and to my PhD supervisors Anne Sophie Daloz and Kristin Aunan for their support and contributions throughout the YSSP programme and beyond.

This research is part of my PhD project, Embodying Climate Change: Transdisciplinary Research on Urban Overheating (EmCliC) – www.emclic.com. The EmCliC project is funded from the Norway and EEA grants 2014–2021 under the Basic Research Programme operated by the Polish National Science Centre in cooperation with the Research Council of Norway (grant no 2019/35/J/HS6/03992)

Part of the research was developed in the Young Scientists Summer Program at the International Institute for Applied Systems Analysis, Laxenburg (Austria) with financial support from the Research Council of Norway National Member Organization.

1. Introduction

The direct impact of heat stress associated with the chaging climate manifests through increasing disease (morbidity) and premature death (mortality) rates (Ebi et al., 2018; Gasparrini et al., 2017; Honda et al., 2014). Heat waves are considered the deadliest extreme weather events between 1991 and 2015 in Europe (EEA, 2017; Kovats et al., 2014). In particular, urban populations face amplified health risks to the elevated temperatures in the built environment. The urban heat island (UHI) effect, whereby urban areas are generally warmer that the surrounding suburban and rural areas, increases heat risk and has direct implications for human health (Heaviside et al., 2017). With urbanization being projected to increase drastically, reaching 83.7% by 2050 in Europe (European Commission), the future health risks from exposure to heat are expected to elevate in a warming climate.

Heat stress affects population sub-groups disproportionately. Epidemiological studies find that population groups with certain socio-economic and demographic characteristics such as age, sex, education, and income are associated with increased risk of mortality and morbidity during heat events (Conlon et al., 2020). Older adults (aged ≥ 65 years), children (aged < 5 years), infants (aged <1 year), women or those with low socio-economic status or pre-existing health conditions are more vulnerable to the health risks of climate change (Basu, 2009; Calleja-Agius et al., 2021). It is also expected that health vulnerability decreases with higher degrees of education (Lutz et al., 2014; Muttarak and Lutz, 2014). Research on the impacts of heat on health however consists mostly of studies carried out at the regional, national or city level. Employing these coarse resolutions may mask differences in vulnerabilities of certain population subgroups. Ascertaining how, for example certain age groups or socio-economic statuses differ in the risk of heat related mortality and morbidity is relevant for estimating future health impacts and effective planning of adaptation measures in urban areas (Arbuthnott et al., 2016). Assessing and projecting demographically differentiated vulnerability are essential to improve our understanding of current and future climate risks from heat stress (Muttarak, 2021).

To address the research gaps related to scale (Gao, 2017; Jones and O’Neill, 2016; Wear and Prestemon, 2019; Zoraghein and O'Neill, 2020) and demographically differentiated vulnerabilities (Basu, 2009; Calleja-Agius et al., 2021; Conlon et al., 2020; Jurgilevich et al., 2017; Muttarak, 2021; Rohat, 2018), this paper provides a framework for quantifying differential vulnerability based on demographic variables at very high resolution (census tract level – approximately 2400 observations) for the case study City of Madrid. Spatial and temporal projections using Bayesian model averaging (BMA) are developed in line with three scenarios of the Shared Socioeconomic Pathways (SSPs). Here, educational attainment is used as a proxy for socio-economic status as education is considered a good indicator, especially in health studies (Winkleby et al., 1992). This paper advances on research gaps highlighted in literature with 1) an application of innovative, advanced methods to 2) for the first time, include demographically differentiated vulnerability to climate risk assessments by preserving the heterogeneity of relevant population groups by age, sex and education and 3) by providing novel evidence on downscaled population projections for Madrid. This research framework is applicable to other locations and at different scales. However, the results are most useful at the highest resolution for which input data are available, so that policy making can take into consideration the specificities of different areas within the city.

Socioeconomic scenarios play a pivotal role in the assessment of heat stress impact on human health as well as in the design of adaptation policy responses (Crespo Cuaresma, 2017). The SSPs are a set of five coherent narratives, describing plausible future changes in demographics, human development, economy and lifestyle, policies and institutions, technology, and environment and natural resources (O’Neill et al., 2017). The demographic component of the SSPs provides a powerful resource for studying population dynamics according to different assumptions about future trends in fertility, mortality and migration. Age, sex and education are important sources of population heterogeneity and their changing composition is directly relevant for vulnerability reduction and adaptation to climate change (KC and Lutz, 2017). In this paper we select three of the five SSP narratives (i.e. SSP1, SSP2 and SSP4) for the vulnerability-relevant spatial population projections. The *SSP2: middle of the road* scenario combines all countries medium fertility, mortality and migration and the Global Education Trend (GET) scenario. It is considered to describe the population dynamic trends of the recent past and it assumes similar future trends. The *SSP1: sustainable development* scenario assumes investments in health and education, and generally decreasing fertility and mortality that lead to a relatively low global population. SSP1 is considered the "best case scenario" and it has been selected because, compared to other scenarios, it features low challenges for climate mitigation and adaptation. By contrast, *SSP4 describes a world of high inequality* with increasing stratification between the well-educated and internationally connected population subgroups and the poorly educated who work in labor intensive industries. Regarding adaptation, this narrative contrasts the most with SSP1, featuring high challenges for climate change adaptation due to persistent poverty and inequality trends (Hausfather, 2018).

The remainder of the paper is organized as follows: Section 2 describes the methodological approach based on Bayesian model averaging (BMA). Section 3 outlines the input data and empirical analysis; Section 4 presents the preliminary results and Section 5 concludes.

2. Methods

This paper provides a robust framework for modelling and projecting population dynamics with high granularity at the intra-urban level. BMA is used to assess the driving factors of differences in age structure and education for males and females and project their changes over time, following the three different SSPs: SSP1 – the sustainable development pathway, SSP2 – middle of the road and SSP4 – the inequality pathway.

The research proceeded in two steps. First, empirical data from the municipal register of inhabitants were used to analyze the correlates of the share of two age groups (25 to 64 years and 65+ years old) separately for males and females and the share of males and females with primary, secondary and tertiary educational attainment. For each dependent variable (e.g., the share of women aged 25-64 with tertiary education in 2020) a large set of explanatory variables were evaluated by BMA. Explanatory variables are represented by weighted fractions of fertility, mortality and internal migration from the baseline year 2010, education variables from the previous time step (i.e., if the dependent variable was the share of women with tertiary education in 2020, then the explanatory variables were shares of primary, secondary and tertiary education from 2010) and a series of district dummy variables (see Dummy variables subsection under Input data for BMA), as well as interaction terms among these covariates. Using BMA and data from 2010 structured by age, sex and education, we derived projections for 2020 which were in turn compared with the observed data for 2020 to assess model performance (see Baseline subsection under Analysis). Second, we developed decadal projections of differential vulnerability for Madrid, until 2050 for the three SSPs.

Bayesian Model Averaging (BMA)

The literature is ambiguous regarding the set of explanatory variables that should be included in modelling the share of different populations groups (i.e., age, sex and educational attainment groups) over time. Thus, a problem of model uncertainty arises, concerning the explanatory variables that should be included in the model and how to perform inference in the presence of uncertainty about the nature of the effect of the explanatory covariates. BMA addresses this issue by explicitly integrating this source of uncertainty in the modeling. With a linear model structure (1) where y is the dependent variable, X is the matrix of independent variables, α_Υ is a constant, β_Υ are the coefficients and ε is a normal error term with variance σ^2, BMA estimates models of all possible combinations of explanatory variables in X and constructs a weighted average over them based on their posterior model probabilities (Zeugner, 2011).

$$y = \alpha_\Upsilon + X_\Upsilon \beta_\Upsilon + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2) \tag{1}$$

This can be done with a large size of X. If X contains *k* potential variables, BMA will estimate 2^k variable combinations, thus 2^k models contained in the model space, formed by all specifications M_Υ of the type given by (1) that can be constructed. The model weights for this averaging stem from posterior model probabilities that arise from Bayes' theorem:

$$p(M_\gamma|y, X) = \frac{p(y|M_\gamma, X)p(M_\gamma)}{p(y|X)} = \frac{p(y|M_\gamma, X)p(M_\gamma)}{\sum_{s=1}^{2^k} p(y|M_s, X)p(M_s)} \tag{2}$$

Bayesian approaches differ from frequentist methods, among other features, through the existence of model priors. The model prior reflects prior knowledge or beliefs on the suitability of certain specifications and is elicited by the researcher before looking into the data (y, X). Conducting an estimate of the effect of an explanatory variable or set of explanatory variables on y in the presence of model uncertainty implies evaluating the posterior distribution given by Bayes' theorem (2), where $p(y|M_\gamma, X)$ is the posterior distribution of y conditional on specification M_Υ and $p(M_\Upsilon)$ is the prior model probability.

This paper employs Bayesian projection methods that quantitatively describe the narratives of the three SSPs (see subsection on Model Projections and SSP integration as prior knowledge) to project future trajectories of age, sex and education structure. The importance of different variables in explaining the data is represented in the posterior inclusion probabilities (PIP) which is the sum of posterior model probabilities (PMPs) for all the models wherein an explanatory variable was included. The PMP is the product of the marginal likelihood of the model (i.e. the probability of the data given model M_Υ) and the prior model probability (Crespo Cuaresma, 2021).

This Bayesian framework completes the sampling model with a prior distribution for parameters in M_Υ, namely α, β and σ (Fernández et al., 2001). The standard choice for BMA applications is to use an improper non-informative prior for the variance of the error term σ and prior over the slope coefficients in the parameter vector β_Υ given by Zellner's g-prior (Zellner, 1986). Zellner's g-prior mimics the structure of the variance-covariance matrix of the ordinary least squares, scaled by the g parameter (Crespo Cuaresma, 2021). A small g in the model setup implies little prior coefficient variance, thus the researcher is conservative in assuming that the coefficients are indeed zero. In contrast, a large g implies that the researcher is very uncertain that the coefficients are zero. Prior model probabilities can be elicited by assuming a flat or uniform prior. A flat prior implies a model probability of $p(M_\Upsilon) = 2^{-k}$ with a prior expected model size of k/2, that is very informative on model size (Zeugner, 2011). To overcome this issue, Ley and Steel (2009) propose a binomial-beta hyperprior for the inclusion of explanatory variables in a given model leads to flexible distributions for model size with uninformative priors on the number of explanatory variables (Ley and Steel, 2009).

Here, we use Bayesian model sampling and averaging (bms) algorithms to run BMA projections for population structure based on age groups, sex and educational attainment. The preliminary results are based on 10 000 iterations after 1000 burn-ins for the test runs. For the final runs, we will use 10 000 burn-ins and five million iterations. We used the benchmark prior suggested by Fernandez et al., 2001 where $g=\max(N, k^2)$, N being the number of observations and k the number of covariates. For SSP1, SSP2 and SSP4 integration we will use suitable assumptions about the overall development of demographic and socioeconomic variables which correspond to the narratives of these scenarios. Computations underlying this paper are performed in the statistical environment of R, version 3.6.1.

Model projections and SSP integration

Climate change research often employs scenarios to understand long-term consequences of near-term decisions allowing to explore different parallel futures in the context of uncertainty. Furthermore, scenarios provide a common basis for exploration of climate impacts and adaptation options, playing a pivotal role in the integration of different systems (Riahi et al., 2017). The latest generation of global development scenarios focuses on uncertainty in future societal conditions and how this can be combined with climate change projections and economic growth and population development assumptions. The so-called Shared Socioeconomic Pathways (SSPs) describe plausible alternative changes in demographic, economic, technological, social, governance and environmental aspects of societal development. They include qualitative descriptions or narratives of broad development trends as well as quantification of key variables (O'Neill et al., 2017). The three SSP narratives selected for this paper are summarized below, based on their general description in O'Neill et al. (2017) and the population development outlined in KC and Lutz (2017).

SSP1: Sustainability **– taking the green road**
The world shifts gradually towards a sustainable path with a focus on sustainable development within the perceived environmental boundaries. Cooperation and collaboration between local, national and international governments and organizations, the private sector and civil society leads to slow but pervasive management of the global commons. Education and health investments accelerate the demographic transition leading to a low global population due to the implied low mortality and an increase in general educational attainment, especially for women. The emphasis is shifting from economic growth to human wellbeing and quality of life. Migration levels are assuming to be medium. This scenario implies low challenges to both mitigation and adaptation
SSP2: Middle of the road
This world does not shift markedly from the historical patterns. Most economies are politically stable, while development and income growth proceed unevenly, with some countries making relatively good progress while others fall short of expectations. Environmental systems continue to experience degradation although there are some improvements which lead to a decline in the overall intensity of energy and resources. In term of demographic development, this scenario combines all countries medium fertility with medium mortality, medium migration and the Global Education Trend (GET) education scenario. Moderate challenges are assumed for both mitigation and adaptation.
SSP4: Inequality – a road divided
Increasing disparities in economic opportunity and political power accompanying with highly unequal investments in human capital lead to increasing inequalities and stratification across and within countries. The gap widens between the well-educated and internationally connected social groups that contribute to knowledge- and capital-intensive sector, and a fragmented collection of low-income and poorly educated population groups who work in low-tech industries. Current fertility patterns do now change markedly from historical patterns, hence countries with high fertility today will continue a similar trajectory whereas countries with low fertility will also maintain a comparable pattern. Also, high fertility countries are expected to see high mortality rates while others have medium mortality. Migration is assumed to be medium for all countries. This world is assumed to face low challenges for mitigation and high challenges for adaptation.

We chose to focus on these three scenarios because we want to compare different vulnerability aspects and their evolution in time. We ask what is the difference between following the current trajectory (SSP2), compared with the "best case scenario" (SSP1) and the high inequality scenario (SSP4). SSP4 was selected as the counterpart scenario over *SSP5: fossil-fuel development* because the latter includes very high radiative forcing levels, implying an unprecedented fivefold increase in coal use by the end of the century, which are deemed increasingly implausible (Hausfather and Peters, 2020).

3. Empirical Analysis

Data

The data used for this analysis are available at census tract level (approximately 2400 observations) from the municipal register in Madrid, for 2010 and 2020. The data are divided into the following subcategories: 5-year age groups (e.g. 0-4, 5-9, etc.) from 0 to 100+ for male, female and total population and 13 educational attainment categories (educational attainment categories are translated from Spanish and listed in *Table 1*). Education data are available for male, female and total population above the age of 25.

Mortality, fertility and migration data were obtained from the statistical office of Madrid. Mortality rate data for males and females are available between the years 1975 to 2019 at city level (not census tract) for 11 age categories (infant mortality – below 1 year-old, 1 to 9, 10 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, 60 to 69, 70 to 79, 80 to 89 and 90 to 99). Fertility data are classified into nine age categories for women below the age of 15, 15 to 19, 20 to 29, 30 to 34, 35 to 39, 40 to 44, 45 to 49 and above 50, from 1975 to 2019. Census tract weighted mortality and fertility from 2010 have been recalculated for the age groups of interest and included in the analysis. Internal migration data between the 21 districts of Madrid are available for 2018. Census tract weighted migration fractions have been computed separately for males and females and for the education and age groups of interest and included in the analysis.

Input data to BMA

This methodological framework can be used for any other city. However, the accuracy of such a study is highly dependent on the granularity of data available. To facilitate the applicability of this framework in other locations, the structure of the Spanish education system was coded according to the International Standard Classification of Education (ISCED).

Education (ISCED codes)

ISCED belongs to the family of economic and social classifications of the United Nations, which are applied in statistics to facilitate assembling, compilation and cross national comparability of data. ISCED has been designed to serve as framework to classify national educational attainment and resulting qualifications into internationally agreed categories. There are nine ISCED levels, from ISCED level 0 and up to ISCED level 8 (UNESCO, 2011). *Table 1* outlines the coding of the variables from the Spanish education system in which the data were originally recorded to ISCED levels.

Table 1. Correspondence between Spanish education systems and ISCED codes

<table>
<tr><th>No.</th><th>Education categories in raw data</th><th>ISCED</th></tr>
<tr><td>1.</td><td>Illiterate</td><td>ISCED 0</td></tr>
<tr><td>2.</td><td>No education</td><td>ISCED 0</td></tr>
<tr><td colspan="3">ISCED 0 or early childhood education refers to early childhood programmes that have an intentional education component within an institutional context and aim to develop socio-emotional skills for later participation in school and society.</td></tr>
<tr><td>3.</td><td>Incomplete primary education</td><td>ISCED 1 (incomplete)</td></tr>
<tr><td colspan="3">ISCED 1 or primary education is part of compulsory education and typically designed to provide students with literacy and numeracy while establishing a foundation for learning and understanding core areas of knowledge. Main criteria for ISCED 1 is systematic instruction in fundamental knowledge, skills and competences, usually between the age of 5 and 7 with instruction organized typically by one main class teacher.</td></tr>
<tr><td>4.</td><td>Elementary bachelor graduate</td><td>ISCED 2</td></tr>
<tr><td>5.</td><td>Professional formation, first degree</td><td>ISCED 2</td></tr>
<tr><td colspan="3">ISCED 2 or lower secondary education is designed to build upon learning outcomes in ISCED 1. It is characterized by a transition to more subject oriented instruction for pupils entering this category at the age 10 to 13.</td></tr>
<tr><td>6.</td><td>Professional formation, second degree</td><td>ISCED 3</td></tr>
<tr><td>7.</td><td>Superior bachelor graduate</td><td>ISCED 3</td></tr>
<tr><td>8.</td><td>Other medium degrees</td><td>ISCED 3</td></tr>
<tr><td colspan="3">ISCED 3 or upper secondary education is the final stage of general or vocational training. It is characterized by more differentiated programmes with an increased range of options and streams and is typically designed to prepare for tertiary education. Pupils enter this category usually between the age of 14 and 16.</td></tr>
<tr><td>9.</td><td>University School Diploma (e.g. BA, B.Sc.)</td><td>ISCED 6</td></tr>
<tr><td colspan="3">ISCED 6 refers to Bachelor’s or equivalent degrees and its duration typically varies between 3 to 4 years. This level of education, though often theoretically oriented is designed to provide participants with intermediate academic and/or professional knowledge, skills and competencies, leading to a first degree or equivalent qualification.</td></tr>
<tr><td>10.</td><td>Architecture and Technical Engineering</td><td>ISCED 7</td></tr>
<tr><td>11.</td><td>University degree (e.g. MA, M.Sc.)</td><td>ISCED 7</td></tr>
<tr><td>12.</td><td>Non-university higher studies</td><td>ISCED 7</td></tr>
<tr><td colspan="3">ISCED 7 is the Master’s degree or equivalent with a duration typically varying from 1 to 4 years when following ISCED level 6 and 5 to 7 years when following ISCED level 3 directly.Provides advanced academic and/or professional knowledge, skills and competences leading to a second degree or equivalent qualification.</td></tr>
<tr><td>13.</td><td>Doctoral and postgraduate studies</td><td>ISCED 8</td></tr>
<tr><td colspan="3">ISCED 8 refers to doctoral degrees, with a duration of minimum 3 years. ISCED 8 leads to an advanced research qualification with programmes devoted to advanced study and original research and are typically offered only by research-oriented tertiary educational institutions such as universities. Doctoral programmes exist in both academic and professional fields.</td></tr>
</table>

Cross-tabulations between age and education data

Education and age group data for Madrid are provided separately. However, to capture the multidimensional characteristics of the population which are relevant to vulnerability and adaptation (Lutz and Muttarak, 2017), data that are stratified by age, education and sex are required. The intersection between the two datasets was computed based on input from EUROSTAT population by educational attainment level, sex and NUTS2 regions (%) for Comunidad de Madrid (Eurostat - Data Explorer (europa.eu)). These data are available for the period between 2016 and 2020, for age group 25 to 65 years. Attaining the values for 2010 involved computing the average change per year and assume these values are constant back to 2010. The EUROSTAT data for educational attainment are clustered based on ISCED 2011 as follows: less than primary, primary and lower secondary education (ISCED 0-2); upper secondary and post-secondary non-tertiary education (ISCED 3-4) and tertiary education (ISCED 5-8). ISCED categories 4 and 5 were not represented in the observations (see Table 1 for coding framework). Therefore, ISCED fractions calculated from the raw data were clustered to match the EUROSTAT framing as follows: ISCED 0-2, ISCED 3 and ISCED 6-8.

For the 65+ age group, the fraction for the three education categories were calculated based on data from the Wittgenstein Center Human Capital Data Explorer (Wittgenstein Centre Human Capital Data Explorer) for population size by education (000's) indicator by sex (male/female), in 2010 for Spain. The education categories from this dataset were coded to match the EUROSTAT categories as follows: no education, incomplete primary, primary and lower secondary education were coded into the ISCED 0-2 category; upper secondary education was coded into ISCED 3-4 category and post secondary education was coded into ISCED 5-8. Age and ISCED categories fractions from the total population were multiplied with EUROSTAT and Wittgenstein data for 2010 to obtain weighted population structure by the three education categories, two age groups (25 to 64 and 65+) for males and females.

Dummy variables

Finally, to capture the differences between the 21 districts in Madrid, a set of dummy variables was created. For each district in turn, a value of 1 was assigned to the district of interest with all other districts taking a value of 0. Then, the district dummy was multiplied with the independent education variables when future education was the y variable and with the independent variables for age groups to project future age structure.

Baseline

Future projections for age, sex and education structure were developed with input from raw data on a decadal basis. Basic population change variables (i.e. mortality, fertility and internal migration) along with education and a district dummy from 2010 were used as input to project the education and age structure in 2020. To assess model performance, the projections derived for 2020 were compared with observations from the same year. *Figure 1* shows and example of model performance, i.e. the comparison between the projection of ISCED 0-2 education category (red line) and the raw data for the same education category (blue line) in 2020.

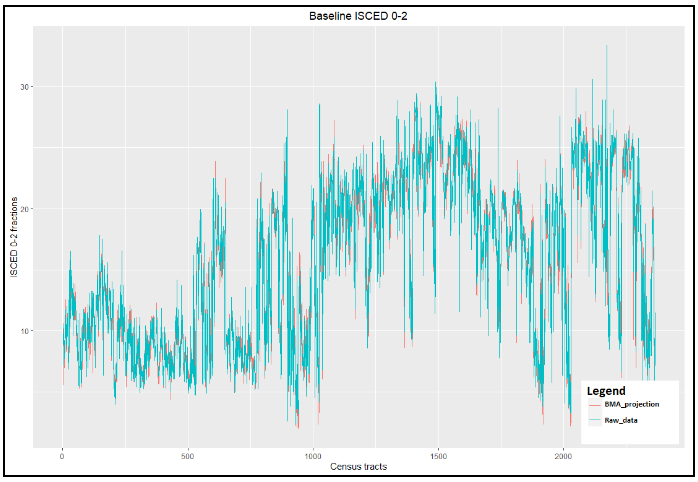

Figure 1. Baseline ISCED 0-2 2020

4. Preliminary results

In this section we present some of the preliminary results obtained through model testing. So far, we have performed Bayesian model sampling and averaging in R, for three education categories (i.e. ISCED 0-2, ISCED 3 and ISCED 6-8) for females, aged 25 to 64. The preliminary results are based on 10 000 iterations after 1000 burn-ins with a benchmark prior where g=max(N, k^2) as suggested by Fernandez et al., 2001. We conducted Bayesian model sampling and averaging over approximately 2400 observations (*N*) and 86 (*k*) covariates. The final runs will be conducted with 10 000 burn-ins and five million iterations and with integration of the narratives from the three selected SSPs (i.e., *SSP1: sustainable development*, *SSP2: middle of the road* and *SSP4: inequality*). We will use suitable assumptions about the overall development of demographic and socioeconomic variables which correspond to the narratives of these scenarios.

BMA runs yield posterior model size distributions and coefficient results. The prior and posterior model size distributions for the three education categories (i.e. ISCED 0-2, ISCED 3 and ISCED 6-8) for females between the age of 25 and 64 years in 2020 are outlined in ANNEX 1. The preliminary coefficient results are presented in *Figure 2* and *Figure 3* provides a visual representation of the coefficients for the best 100 models.

The final results will consist of downscaled population projections at decadal intervals from 2010 to 2050 and disaggregated in subgroups by age, sex and education. These projections will show the population dynamics and allow for comparison between the three trajectories of the SSP narratives. The corresponding datasets of the distribution of population groups and their projected changes will be openly available.

Coefficient results

Figure 2 represents the coefficient results matrix for females, aged 25 to 64 with education level ISCED 0-2. The posterior inclusion probabilities (PIP) column shows the importance of the variables in representing the data. With migration, ISCED 0-2 fractions from 2012 and a set of dummy variables for the same education category at 100%, the posterior model mass relies on models that include these variables (see Bayesian Model Averaging (BMA) subsection under Methods for a more detailed interpretation of *Figure 2*).

Post Mean shows the coefficients (weighted) averages over all the models. Among the variables with PIP close to 1.000 - i.e. variables with the highest importance in explaining the data, isced02_2012 covariate has a comparatively large coefficient and thus seems to be the most important. The Post SD column displays the coefficients posterior standard deviations, Cond.Pos.Sign is the posterior probability of a positive coefficient expected value conditional on inclusion and the last column idx represents the index of the variables appearance in the original data.

Figure 3 brings a comprehensive overview of the best 100 models and their variables, included based on PIP. Blue represents a positive coefficient, red corresponds to a negative coefficient and white refers to non-inclusion. The x-axis shows the PMP – i.e. posterior model probabilities, for the best 100 models.

	PIP	Post Mean	Post SD	Cond.Pos.Sign	Idx
migration_fem_25.64_tot	1.0000	-3.103700e-02	5.436731e-03	0.00000000	3
isced02_2012	1.0000	9.426038e-01	1.519516e-02	1.00000000	4
X01isced02	1.0000	-1.660599e-01	1.259715e-02	0.00000000	8
X02Xisced02	1.0000	-1.695163e-01	3.445145e-02	0.00000000	12
X03Xisced02	1.0000	-2.037717e-01	4.557576e-02	0.00000000	16
X04Xisced02	1.0000	-1.867405e-01	1.790783e-02	0.00000000	20
X05Xisced02	1.0000	-3.193976e-01	4.722462e-02	0.00000000	24
X07Xisced02	0.9839	-1.852583e-01	2.831876e-02	0.00000000	32
Dummy_Tetuan06	0.9743	-1.168287e+00	2.443914e-01	0.00000000	27
X05Xisced3	0.9716	3.947852e-01	9.074285e-02	1.00000000	25
X08Xisced02	0.9103	-2.356233e-02	1.035689e-02	0.00000000	36
Dummy_Moncloa.Aravaca09	0.8681	-8.790914e-01	3.747437e-01	0.00000000	39
X12Xisced02	0.8489	1.985436e-02	1.069302e-02	1.00000000	52
mortality_2010	0.7213	2.070056e+07	2.918647e+07	1.00000000	1
Dummy_Hortaleza16	0.6969	-3.995371e-01	3.024119e-01	0.00401779	67
Dummy_Ciudad.Lineal15	0.6900	-3.565912e-01	3.108951e-01	0.00000000	63
fertiliy_2010	0.6496	-4.223366e+05	5.955197e+05	0.42903325	2
X03Xisced3	0.6390	1.556595e-01	1.297109e-01	1.00000000	17
Dummy_Puente.de.Vallecas13	0.6022	2.883678e-01	2.877936e-01	0.97492527	55
X02Xisced3	0.5819	1.337011e-01	1.260361e-01	1.00000000	13
X17Xisced3	0.5565	8.188357e-02	7.911027e-02	1.00000000	73
X13Xisced02	0.3837	7.956998e-03	1.106799e-02	1.00000000	56
Dummy_Villaverde17	0.3102	2.223901e-01	3.865848e-01	1.00000000	71
X15Xisced68	0.3079	-1.104038e-02	1.758510e-02	0.00227347	66
X19Xisced02	0.2621	5.595501e-03	1.023405e-02	1.00000000	80
Dummy_Vicalvaro19	0.2054	9.321010e-02	2.001137e-01	1.00000000	79
Dummy_Retiro03	0.1961	3.351666e-01	8.078907e-01	0.98164202	15
X11Xisced02	0.1679	1.433850e-03	4.607538e-03	0.87611674	48
X16Xisced3	0.1676	-1.814339e-02	4.179037e-02	0.00000000	69
Dummy_Arganzuela02	0.1639	1.895800e-01	4.817369e-01	0.97071385	11
X09Xisced02	0.1409	-7.807785e-03	2.026484e-02	0.00000000	40
X17Xisced02	0.1347	1.643253e-03	1.078547e-02	0.75501114	72
X10Xicsed68	0.1310	-5.755153e-03	1.932914e-02	0.00000000	46
X16Xisced68	0.1300	-4.295155e-03	1.217392e-02	0.00000000	70
Dummy_Usera12	0.1262	5.377464e-02	2.085728e-01	0.92789223	51
isced3_2012	0.1091	1.063860e-02	3.441078e-02	1.00000000	5
Dummy_Carabanchel11	0.0978	2.254283e-02	7.470486e-02	1.00000000	47
X13Xisced68	0.0912	-2.090448e-03	1.263812e-02	0.01096491	58
X16Xisced02	0.0869	-5.789622e-04	7.295502e-03	0.52243959	68
Dummy_Fuencarral.El.Pardo08	0.0863	-5.419628e-02	2.096497e-01	0.04287370	35
X18Xisced3	0.0824	-7.787784e-03	2.908925e-02	0.00000000	77
X03Xisced68	0.0798	3.257215e-03	1.786172e-02	0.95614035	18
X08Xisced68	0.0758	2.821313e-03	1.163470e-02	1.00000000	38
X19Xisced3	0.0729	4.650082e-03	2.170422e-02	1.00000000	81
isced68_2012	0.0661	-2.894029e-03	1.204854e-02	0.00000000	6
X08Xisced3	0.0628	3.170931e-03	1.860329e-02	0.88694268	37
X10Xicsed3	0.0582	1.017385e-03	1.968374e-02	0.30927835	45
X10Xicsed02	0.0536	-3.430711e-04	5.586681e-03	0.06902985	44
X12Xisced3	0.0509	-1.278725e-03	1.862892e-02	0.10019646	53
X19Xisced68	0.0500	6.623268e-04	7.033407e-03	0.82800000	82
X05Xisced68	0.0471	-3.150224e-03	3.874128e-02	0.29299363	26
X17Xisced68	0.0446	4.162218e-03	2.025039e-02	1.00000000	74
X20Xisced02	0.0441	-5.889498e-04	3.186975e-03	0.00000000	84
Dummy_Centro01	0.0436	1.252266e-02	1.286241e-01	1.00000000	7
Dummy_Moratalaz14	0.0393	-2.092175e-02	1.413486e-01	0.01781170	59
X14Xisced3	0.0367	1.188605e-03	2.182667e-02	0.36512262	61
Dummy_Chamberi07	0.0350	-9.530074e-02	7.581824e-01	0.46857143	31
X11Xisced68	0.0336	3.811399e-04	4.950401e-03	0.92261905	50
X06Xisced68	0.0330	-1.548432e-03	9.863259e-03	0.00000000	30
X15Xisced3	0.0316	2.069239e-03	3.160658e-02	0.56012658	65
Dummy_Latina10	0.0314	-9.438248e-04	1.648337e-01	0.10509554	43
X07Xisced68	0.0299	3.668255e-03	2.770163e-02	1.00000000	34
X02Xisced68	0.0261	1.066091e-03	7.573547e-03	0.91570881	14
Dummy_San.Blas.Canillejas20	0.0256	-9.273614e-04	2.064712e-02	0.19140625	83
X06Xisced02	0.0248	-9.967075e-04	7.658811e-03	0.14919355	28
X15Xisced02	0.0246	-1.255133e-04	3.314534e-03	0.20731707	64
Dummy_Chamartin05	0.0196	1.416179e-01	1.342183e+00	0.74489796	23
X04Xisced68	0.0183	2.611900e-04	3.272168e-03	1.00000000	22
Dummy_Salamanca04	0.0176	1.382246e-02	1.289461e-01	1.00000000	19
X01Xisced68	0.0169	1.592802e-04	2.759686e-03	1.00000000	10
X04Xisced3	0.0157	1.567786e-03	1.527883e-02	1.00000000	21
X09Xisced68	0.0156	6.219683e-05	3.523488e-03	0.87820513	42
X14Xisced02	0.0154	-1.793437e-04	2.243545e-03	0.00000000	60
X20Xisced68	0.0149	3.896151e-05	1.435274e-03	0.87919463	86
X12Xisced68	0.0124	-6.683928e-04	7.279839e-03	0.00000000	54
X18Xisced02	0.0113	-8.196977e-05	1.162502e-03	0.00000000	76
X07Xisced3	0.0111	2.690299e-04	7.946411e-03	1.00000000	33
X14Xisced68	0.0100	-9.032898e-05	1.572657e-03	0.00000000	62
X11Xisced3	0.0072	5.139100e-04	7.854508e-03	1.00000000	49
X09Xisced3	0.0058	-7.096124e-04	1.018732e-02	0.00000000	41
X06Xisced3	0.0052	1.049955e-04	1.025863e-02	0.90384615	29
Dummy_Villa.de.Vallecas18	0.0034	-9.733036e-04	2.104769e-02	0.00000000	75
X20Xisced3	0.0022	6.203346e-05	1.737075e-03	1.00000000	85
X18Xisced68	0.0021	-5.738906e-05	1.616837e-03	0.00000000	78
X01Xisced3	0.0013	1.131923e-04	4.195778e-03	1.00000000	9
X13Xisced3	0.0009	1.535757e-05	9.536978e-04	1.00000000	57

Figure 2. Coefficient results ISCED 0-2

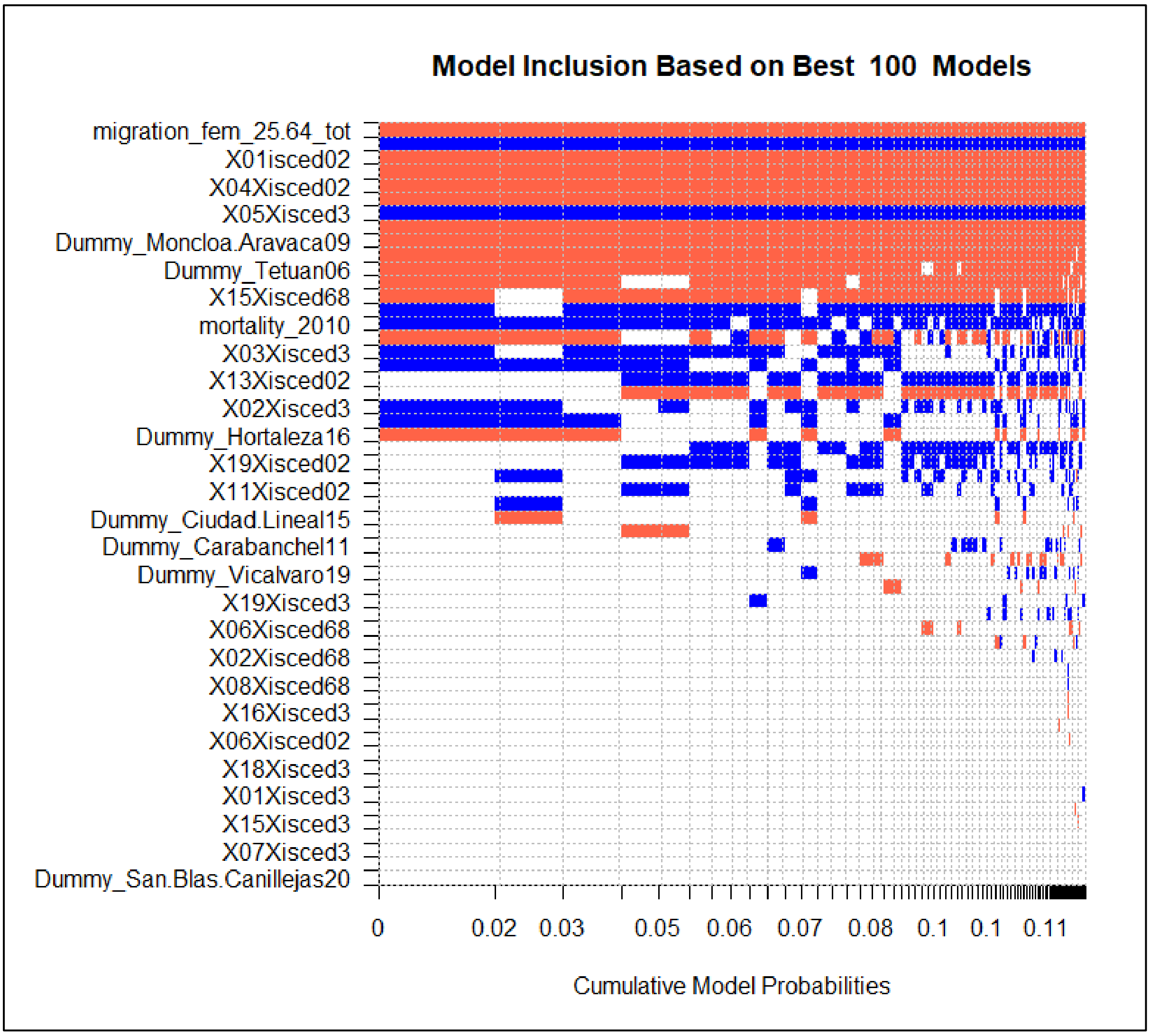

Figure 3. Model inclusion and coefficient sign for ISCED 0-2, females aged 25-64

Projected change in female education

Projections for each education category and age group have been conducted so far for SSP2. To assess population dynamics over time, we look at projected changes for each education category and age group. *Figure 5* portrays the change in a share of highly educated females (ISCED 6-8), aged 25 to 64 between 2010 and 2030 at the census tract level. Clusters of decreasing education in adult women can be observed in the far south-east area of the city. These neighborhoods – Villa de Vallecas and Vicalvaro – are extensions of the overly populated Puente de Vallecas district*. Figure 5* also depicts a belt-shaped pattern, on the south side of the city center where only a slight increase in female education is visible. These neighborhoods are characterized by high-density working-class population.

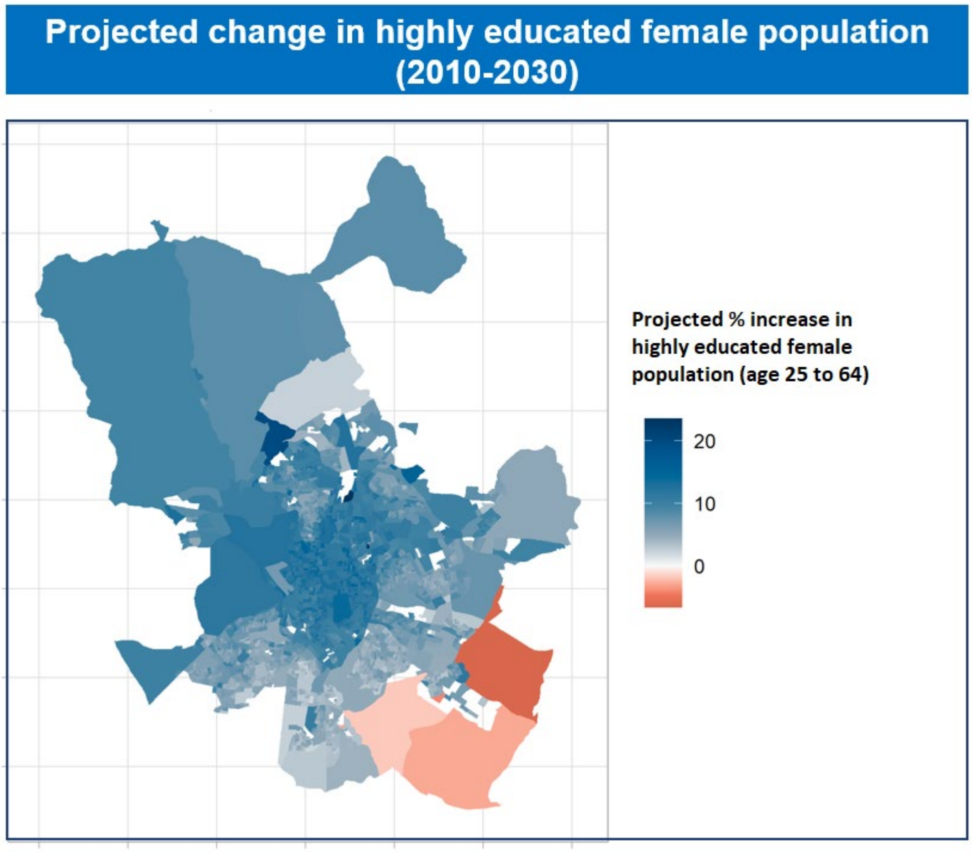

Figure 2. Change between 2010 and 2030 in highly educated females, ages 25 to 64

5. Conclusions

The negative effects of climate change are already experienced, especially by city dwellers. Rapid urbanization and the UHI effect significantly increase the exposure to heat stress in urban environments. While drastic emission reductions are urgently needed to meet the ambitions of the Paris Agreement and limit the global warming below 2º, enhancing adaptation capacity is paramount to minimize the climate impacts on the health of urban populations. Decreasing the burden that climate change is imposing on public health, adaptation mechanisms and strategies need to target and prioritize vulnerable groups.

Population dynamics play a central role in climate risk assessments and quantification of vulnerability to heat stress. High-resolution population estimates and projections at the intra-urban scale are useful in identifying clusters of vulnerable groups within the city. Using the BMA for downscaled population projections can contribute to significant advancement in assessing the risk of heat stress for human health in cities. We find that BMA performs well in integrating uncertainty within the model and the projected data falls in line with observations for all three education categories.

Estimations of population heterogeneity allow for identification of vulnerable subgroups and assessment of adaptive capacity. Preservation of population heterogeneity in the projections is directly relevant for anticipating socioeconomic changes and preparing targeted adaptation plans. This paper addresses the need for more detailed assessment of vulnerability to heat stress, providing a framework for such analysis that is applicable in any other locations. However, the resolution of input data is a determinant for the detail that can be obtained in the effort to capture this heterogeneity and the use of data at highest resolution available is recommended.

References

Arbuthnott, K., Hajat, S., Heaviside, C. and Vardoulakis, S., 2016. Changes in population susceptibility to heat and cold over time: assessing adaptation to climate change. Environ Health, 15 Suppl 1: 33.

Basu, R., 2009. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. Environ Health, 8: 40.

Calleja-Agius, J., England, K. and Calleja, N., 2021. The effect of global warming on mortality. Early Hum Dev, 155: 105222.

Conlon, K.C., Mallen, E., Gronlund, C.J., Berrocal, V.J., Larsen, L. and O'Neill, M.S., 2020. Mapping Human Vulnerability to Extreme Heat: A Critical Assessment of Heat Vulnerability Indices Created Using Principal Components Analysis. Environ Health Perspect, 128(9): 97001.

Crespo Cuaresma, J., 2017. Income projections for climate change research: A framework based on human capital dynamics. Global Environmental Change, 42: 226-236.

Crespo Cuaresma, J., 2021. Uncertainty and business cycle synchronization in Europe. Applied Economics Letters: 1-7.

Ebi, K.L., Hasegawa, T., Hayes, K., Monaghan, A., Paz, S. and Berry, P., 2018. Health risks of warming of 1.5 °C, 2 °C, and higher, above pre-industrial temperatures. Environmental Research Letters, 13(6).

EEA, 2017. Climate change, impacts and vulnerability in Europe 2016. An indicator-based report., European Environment Agency, Copenhagen.

Fernández, C., Ley, E. and Steel, M.F.J., 2001. Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics, 16(5): 563-576.

Gao, J., 2017. Downscaling global spatial population projections from 1/8-degree to 1-km grid cells.

Gasparrini, A., Guo, Y., Sera, F., Vicedo-Cabrera, A.M., Huber, V., Tong, S., de Sousa Zanotti Stagliorio Coelho, M., Nascimento Saldiva, P.H., Lavigne, E., Matus Correa, P., Valdes Ortega, N., Kan, H., Osorio, S., Kyselý, J., Urban, A., Jaakkola, J.J.K., Ryti, N.R.I., Pascal, M., Goodman, P.G., Zeka, A., Michelozzi, P., Scortichini, M., Hashizume, M., Honda, Y., Hurtado-Diaz, M., Cesar Cruz, J., Seposo, X., Kim, H., Tobias, A., Iñiguez, C., Forsberg, B., Åström, D.O., Ragettli, M.S., Guo, Y.L., Wu, C.-f., Zanobetti, A., Schwartz, J., Bell, M.L., Dang, T.N., Van, D.D., Heaviside, C., Vardoulakis, S., Hajat, S., Haines, A. and Armstrong, B., 2017. Projections of temperature-related excess mortality under climate change scenarios. The Lancet Planetary Health.

Hausfather, Z., 2018. Explainer: How 'Shared Socioeconomic Pathways' explore future climate change. Carbon Brief.

Hausfather, Z. and Peters, G., 2020. Emissions – the 'business as usual' story is misleading. Nature.

Heaviside, C., Macintyre, H. and Vardoulakis, S., 2017. The Urban Heat Island: Implications for Health in a Changing Environment. Curr Environ Health Rep, 4(3): 296-305.

Honda, Y., Kondo, M., McGregor, G., Kim, H., Guo, Y.L., Hijioka, Y., Yoshikawa, M., Oka, K., Takano, S., Hales, S. and Kovats, R.S., 2014. Heat-related mortality risk model for climate change impact projection. Environ Health Prev Med, 19(1): 56-63.

Jones, B. and O'Neill, B.C., 2016. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. Environmental Research Letters, 11(8): 084003.

Jurgilevich, A., Räsänen, A., Groundstroem, F. and Juhola, S., 2017. A systematic review of dynamics in climate risk and vulnerability assessments. Environmental Research Letters, 12(1).

KC, S. and Lutz, W., 2017. The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. Glob Environ Change, 42: 181-192.

Kovats, R.S., Valentini, R., Bouwer, L.M., Georgopoulou, E., Jacob, D., Martin, E., Rounsevell, M. and Soussana, J.-F., 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. In: V.R. Barros, C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, L.L. White (Editor). Cambridge University Press, United Kingdom and New York, NY, USA, pp. 1267-1326.

Ley, E. and Steel, M.F.J., 2009. On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. Journal of Applied Econometrics, 24(4): 651-674.

Lutz, W. and Muttarak, R., 2017. Forecasting societies' adaptive capacities through a demographic metabolism model. Nature Climate Change, 7(3): 177-184.
Lutz, W., Muttarak, R. and Stiessnig, E., 2014. Universal education is key to enhanced climate adaptation. Fund more educators rather than just engineers. Science, 346(6213): 1061-1062.
Muttarak, R., 2021. Demographic perspectives in research on global environmental change, International Institute for Applied Systems Analysis, Laxenburg, Austria.
Muttarak, R. and Lutz, W., 2014. Is Education a Key to Reducing Vulnerability to Natural Disasters and hence Unavoidable Climate Change? Ecology and Society, 19(1).
O'Neill, B.C., Kriegler, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., van Ruijven, B.J., van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M. and Solecki, W., 2017. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Global Environmental Change, 42: 169-180.
Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J.C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A. and Tavoni, M., 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. Global Environmental Change, 42: 153-168.
Rohat, G., 2018. Projecting Drivers of Human Vulnerability under the Shared Socioeconomic Pathways. Int J Environ Res Public Health, 15(3).
UNESCO, I.f.S., 2011. International Standard Clasification of Education ISCED 2011. UNESCO Institute for Statistics, Montreal, Quebec H3C 3J7 Canada.
Wear, D.N. and Prestemon, J.P., 2019. Spatiotemporal downscaling of global population and income scenarios for the United States. PLoS One, 14(7): e0219242.
Winkleby, M.A., Jatulis, D.E., Frank, E. and Fortmann, S.P., 1992. Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease. American Journal of Public Health, 82(No. 6).
Zellner, A., 1986. On Assessing Prior Distributions and Bayesian Regression Analysis with G–prior Distributions, Bayesian Inference and Decision Techniques: Essays in Honor of Bruno De Finetti Elsevier Science Ltd.
Zeugner, S., 2011. Bayesian Model Averaging with BMS.
Zoraghein, H. and O'Neill, B., 2020. A spatial population downscaling model for integrated human-environment analysis in the United States. Demographic Research, 43: 1483-1526.

ANNEX 1

Running BMA for three education categories for females between the age of 25 and 64 years in 2020, yields a posterior model size distribution and coefficient results. An uninformative model prior implies a symmetric prior distribution around k/2 = 43. The posterior distribution – i.e. after the update from the data, adds more weight on parsimonious models. *Figure 1A*, 2A and 3A show the comparison between the prior and posterior model distributions for ISCED 0-2, ISCED 3 and ISCED 6-8 education categories. The posterior model size mean values, namely 23.7477 for ISCED 0-2, 26.6324 for ISCED 3 and 21.6885 for ISCED 6-8 represent the number of variables that shall be included in the model with the highest probability density. Thus, the mean is the posterior expected model size – i.e. the average number of included covariates (see Bayesian Model Averaging (BMA) subsection under Methods for a more detailed interpretation of these figures).

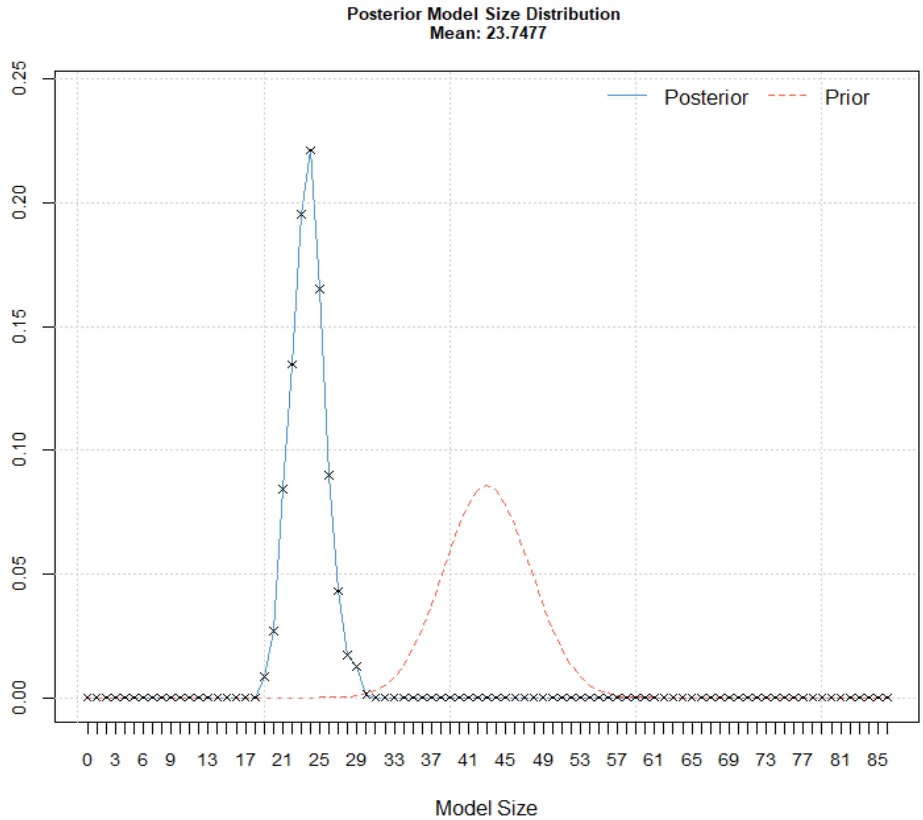

Figure 3A. Prior and posterior model distribution for females aged 25-64, education level ISCED 0-2

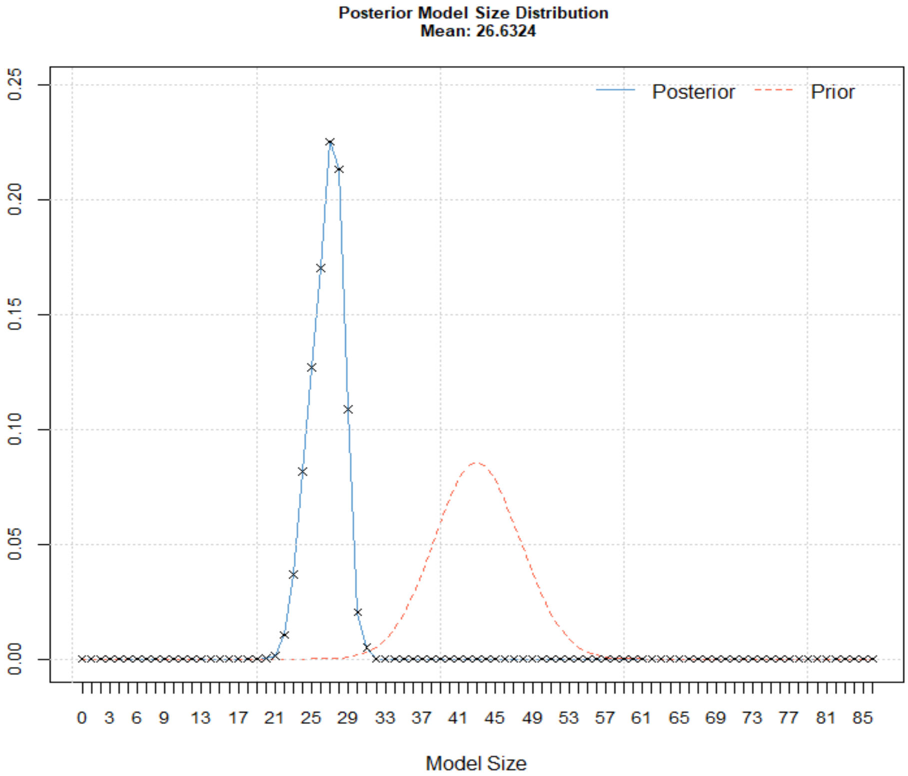

Figure 4A. Prior and posterior model distribution for females aged 25-64, education level ISCED 3

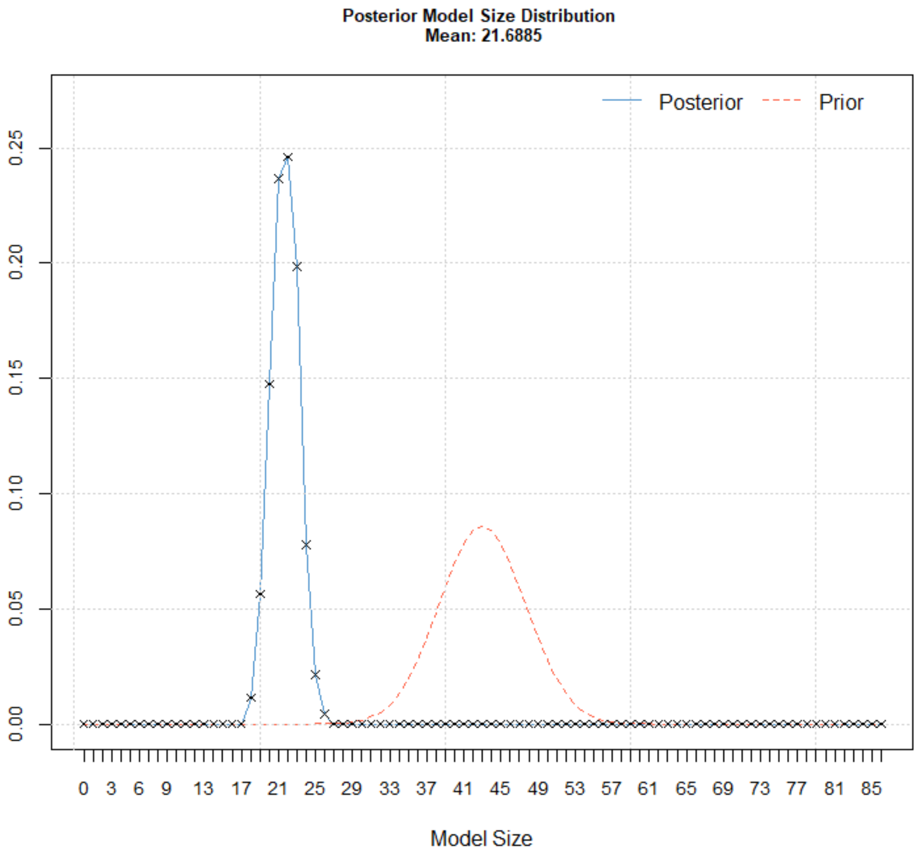

Figure 5A. Prior and posterior model distribution for females aged 25-64, education level ISCED 6-8