Statistics Iceland

Statistical Series

Hagtíðindi

Working paper

15 August 2023

Methodology of population projections based on hierarchical Bayesian models

Aðferðafræði við mannfjöldaspá Hagstofu Íslands byggð á bayesískum tölfræðilíkönum

Samantekt

Mannfjöldaspáin sem Hagstofa Íslands birti árið 2022 byggir á nýjum tölfræðilíkönum um frjósemi, dánartíðni og búferlaflutninga. Framreikningarnir eru byggðir á stöðluðum tölfræðilegum og lýðfræðilegum aðferðum. Með nýja tölfræðilíkaninu er hægt að búa til staðbundnar spár þar sem komið er í veg fyrir ofmat íbúafjölda, t.d. vegna þeirra sem flytja af landi brott án þess að tilkynna búferlaflutninga en áætlaður fjöldi þeirra íbúa sem eru með búsetu á Íslandi er um 2,5% minni en skráður fjöldi. Niðurstöður nýju mannfjöldaspárinnar sýna bæði gildi og óvissumælingar.

Framreikningsaðferðin felur ekki í sér áhrif hugsanlegra áfalla af náttúrulegum, félagslegum eða efnahagslegum orsökum og heldur ekki áhrif fjölda flóttamanna. Sem dæmi má nefna að miðgildisspáin fyrir fólksfjölgun vegna ársins 2022 er 2,02% á meðan skráð fólksfjölgun árið 2022 var 3,06%. Munurinn skýrist að öllu leyti af óvenju miklum flóttamannastraumi sem nam 0,96% aukiningu á mannfjölda. Mannfjöldaspáin verður uppfærð um leið og nýjar upplýsingar verða tiltækar. Notendur eru þess vegna hvattir til að veita Hagstofunni endurgjöf í formi eigindlegra og/eða megindlegra upplýsinga.

Summary

The population projections published by Statistics Iceland in 2022 are based on new statistical models of fertility, mortality and migration and their combined predictions. The projections are built by applying standard statistical and demographical methods. The new models and methods described in this paper allow us, in addition, to produce local projections and to incorporate the probability of overestimating the resident population. The overestimation effect is due to a lack of deregistration, e.g. the estimated resident population is about 2.5% smaller than the registered population. The results of the new population projections consist of predicted values as well as their associated uncertainty measures.

The projection method does not include any effects due to possible crises caused by natural, social or economic factors, nor the number of refugees. For instance, the predicted population growth for the year 2022 is 2.02% while the observed, register based growth during 2022 was 3.06%. The difference is entirely explained by the unusually high refugee flow which accounted for a 0.96% additional change in population. The national and local projections will be updated as soon as auxiliary, prior information becomes available and users, planning and administration factors and policy makers are invited to provide feedback of qualitative and/or quantitative type to Statistics Iceland for that purpose.

Contents

1. Introduction

In this paper we describe the new models and methods introduced by Statistics Iceland in order to generate population projections at national level as well as preliminary local (municipality level) projections, starting with the year 2022. The emphasis is on the interpretation of results, novelty of methodology, main models and open software sharing, while technical details are kept to a minimum.

Until 2022, Statistics Iceland employed a mixture of several classes of methods for producing population projections. These were based on: (i) functional models [1, 2], for predicting fertility and mortality rates (ii) Bayesian models [3] and econometric/ARDL models [4] for short term migration numbers (iii) assumptions/scenarios (as used by numerous statistical institutes) for long term migration predictions. In addition, modeling the time correlation between emigration and lagged immigration has been used in order to further improve the predictions. Only national projections were produced until now.

We have been confronted with several difficulties while using these methods. Firstly, due to the separate migration models: The short-term models performed well according to our evaluation over the last ten years but they could not be extended to the 50 years horizon required by population projections, since they included economic and social predictors which are only available for five years periods in advance. While the observed migration counts were reasonably stable through time, the short and long-term predicted migration regimes were smoothly connected in a natural way. However, when extremely high migration values occurred, as during recent years, the two regimes became disjoint. Additional, intermediate regime modelling became necessary. Secondly, incorporating prior information and uncertainty sources, e.g. related to resident versus register population, were also challenging tasks in the previous set-up. Thirdly, local projections produced according to same methods were not published until now.

The new models described in the present paper are able to encode in an efficient and systematic way the key components needed for reliable projections:

- smooth functional *relationships* between multiple predictors (individual demographical characteristics) and response (demographic events) which in turn can vary by groups/clusters of similar observations (e.g. location attributes but also cross-classifying characteristics such as citizenship, direction of international migration) and

- *correlations* across these dimensions: not only temporal but also higher dimensional ones, e.g. time-age or space-time-age correlations which may also vary by cross-classifying characteristics.

In the statistical demography literature, the following models for demographic components such as migration, mortality or fertility were recently tested [1-3, 5-8]: generalized linear models (GLMs), generalized additive models (GAMs), state space models (SSM) or even Gaussian Process (GP) models. The proposed models for these components are simpler than the ones we employ for our projections, since they include either: (i) time and age smooth dependencies, when using GP or SSM, but no other attributes, or (ii) more characteristics in addition to time and age, such as gender and region in the case of hierarchical GLMs, although in that case the rates were linear functions of these attributes/interactions and the time priors were damped linear trends (e.g. in [5]). In addition, not all studies update the exposed population at each out-of-sample time-step when calculating the predicted demographic rates based on GAMs/GLMs. Examples of studies which do this operation are based on functional models or on expert assumptions (see [1-3]).

Note that population projections have also been built without involving the cohort component method. They were based on assumptions and on Bayesian time series modeling (even model averaged), e.g. using auto-regressive and stochastic volatility models[1]. Bayesian and stochastic projections have been published by UN [17], Statistics New Zealand [5, 18] and Statistics Netherlands [19, 20]. The demographic

[1] Abel G., Bijak J., Raymer J. (2010) A comparison of official population projections with Bayesian time series forecasts for England and Wales, E.S.R.C., http://cpc2.geodata.soton.ac.uk/docs/2010_WP7_Comparison_of_Population_Projections_with_Bayesian_Time_Series_Forecasts_Abel_et_al.pdf

components involved in these stochastic population predictions are based both on assumptions and/or on forecasts of total demographic rates (as time series) while usually not accounting for the influence of extra individual or grouping characteristics.

The novelty of our contribution consists of the following:

- all demographic rates are modeled using the same class of, although not identical, up-to-date models[1] which perform well for solving the type of difficulties associated with our data, as described in the following sections

- the models are flexible: their complexity can be gradually increased and they can be systematically improved when more information and more data become available

- the exposed population is updated at each time step of the calculations and this is done according to the algorithm for stochastic cohort component method for population projections described in the following sections.

To our knowledge, there is no paper which combines hierarchical Bayesian models of demographic rates, via a stochastic version of the cohort component method into consistent population predictions.

Although this is the first time we produce population projections based on these models, and although this is only their first and simplest version, the results are already very useful and realistic. Better addapted models could be built in the future, by tuning their structure and by incorporating useful prior information from experts, local administration, policy makers.

What are population projections

Population projections describe the future population trends and associated degree of uncertainty, i.e. the trends in the number of people by age, gender, time and other demographic or spatial characteristics. The projections are based on past observed values over a reasonably long time and/or expert quantitative/qualitative information.

Why the population projections are important

There is a huge demand, as well as very good feedback in some cases[2], for estimated trends of future population from:

(i) National and local administration and policy makers
(ii) Planning of: constructions, heating, water and electricity supply systems
(iii) Planning of: health, pensions and education sectors
(iv) Researchers from academic and industry domains
(v) Individual users

Solution principle

The solution is based on statistical modeling and combining prior/expert information when possible and available.

[1] Hierarchical, generalised additive models with Gaussian Process priors components.

[2] Aknowledgement: SI is grateful to our colleagues from Veitur, HMS, Reykjavik Municipality, Byggdastofnum, and many actuary colleagues in Iceland, for useful discussions.

What is new

(i) the *type* and *interpretation* of results

- total (national) projections and local (municipality) projections refer to the resident population (as opposed to registered population). This is done by incorporating the information about de-registration delays and omissions[1]. In the 2021 Census of the population, Statistics Iceland discovered an overestimation of approximately 2.7%, albeit some was due to definitional differences[2].

- building a unique model for migration, unlike previous projections which used different models for short and long-term migration

- the predicted migration series is more conservative and smoother than results of the previous methods while the predictions for the fertility and mortality rates are almost identical

- the main projection variants are now defined by the median, upper and lower bounds of the 90% credible intervals[3] defined as follows:

a. - the probability to find the true value of the population below the median projection value (the 50-th percentile) is equal to the probability of finding the true value above the median

b. - the bounds of the 90% credible interval indicate that the probability of finding the true value of the population numbers either below or above these lower/upper limits is 5%

- the uncertainty measures have thus a more intuitive interpretation in this case, i.e. the probability that the actual value of the predicted variable is contained within the Credible Intervals' limits (denoted as the 5% Percentile and the 95% Percentile values), with probability 90%, according to and conditioning on the model

(ii) *cautionary* remarks

Population projections are not exact predictions of future values but attempts to describe the future size and structure of the population based on past values and additional expert information.

For instance, the population projections **do not** include any effects due to possible crises/interventions caused by natural, social, political or economic factors. The size and effects of refugees' migration are not predicted nor included as this are impossible to predict. Since migration has a strong influence on population projections, important changes of international conditions (with respect to past observations) may have unforseen effects on the future population.

For instance, the predicted population growth for the year 2022 was 2.02% while the observed, register based growth during 2022 was 3.06%. The difference is entirely explained by the unusually high refugee flow which accounted for a 0.96% additional change in population[4].

(iii) *Up-to-date, improved methods and models*: why are they better

- the available time series data are not big enough to extract the needed patterns by unsupervised machine learning. In addition, very old data have only a small impact on future trends. Therefore, modeling and any additional knowledge need to be employed

- the projection method employed by Statistics Iceland satisfies the following conditions: it is based on statistical modeling and, due to the type of models employed, it is able to efficiently solve small

---

[1] Calian V., Zuppardo M., Hardarsson O. (2023) Random Forest algorithm to adjust for Census population over-counts, NTTS-2023, http://hagstofan.s3.amazonaws.com/media/public/2023/2bc02eaf-e8cb-417a-927a-7cfb77d99a88.pdf

[2] Hardarson, O. (2022). Census and housing census 1st January 2021. Statistical series, 14. November 2022. http://hagstofan.s3.amazonaws.com/media/public/2022/1a02f09f-c76d-40db-9bb9-2239c3c3bd27.pdf

[3] The probability value is chosen according to the purpose and utility of the study, 90% is not a unique choice.

[4] In addition to the standard register and modeling errors.

area/small population and rare events issues as well as to account for complex (auto-) correlation structures and to incorporate qualitative and quantitative prior information and/or expert assumptions.

- the models we use are more complex than GLMs, GAMs or functional models, since they include:

- smooth (non-linear and non-/ parametric) terms (just as GAMs do),

- interactions,

- hierarchical/multilevel and (temporal, spatial, grouping) correlation structures. They are mostly classified as hierarchical, general additive models (HGAMs)

- a powerful and performant model component, especially appropriate for treating the time series content of the problem, which consists of Gaussian processes (GP) (see the following section). GPs provide both a complex and consistent use of data and a tractable modeling solution [10].

When fitted in a Bayesian framework, the models may incorporate prior information in a natural way. This type of information is critical for predictions out of sample, while model fitting is mainly driven by data. The priors also become important when using short time series for model training.

Open source code and open data principle

We share the R-code developed for the population projections as open source code via the github-repository[1] titled *SIPP* (Statistics Iceland's Population Projections) which is evolving and improving constantly.

The models are easy to adapt and modify according to the availability in data input. The simplest type of data structure for births, deaths and migration counts may be easily downloaded from Statistics Iceland web-site. These data-sets can be used with the simplest versions of the models we provide.

For the production of our most recent population projections we have used a slightly more complex data structure, involving demographic events over time by age, gender, citizenship, type of migration and municipality. During the preliminary stages of model building and for exploratory and testing purposes, we have also included information concerning education levels of individuals, type of family they belong to, municipality related attributes. The results of these investigations are described in the Appendix 1 and are the basis of the most parsimonious models. For instance, these analyses showed that mortality rates do not change significantly between municipalities, fertility rates depend on citizenship and less on location while migration rates vary with both citizenship and region/municipality.

These more complex data sets are not available as open data yet, due to statistical disclosure control issues: when cross-classifying observations according to multiple criteria, the risk of disclosing sensitive and confidential information about individuals becomes very high. This limitation may be overcome in the future, by building synthetic data or creating useful perturbed versions of the true data set.

Prior information needs to be collected in order to improve the local projections and to produce high quality official statistics in the future. Statistics Iceland will identify the best method to interact with local administration, policy makers, planning authorities and other economic and social partners, for this purpose.

[1] https://github.com/violetacln/SIPP

2. Description of the method and implementation

Main concepts

There are two main parts of any population projection exercise.

(i) The first part consists of defining the future development of fertility, migration and mortality by predicting (as in model based methods) or by making assumptions (as in scenario based methods) about the numbers of births, deaths and migrants and their distributions by various attributes of interest (gender, citizenship and many others in principle).

(ii) The second part consists of combining (in a deterministic or stochastic way) all these future values, according to the well-known demographic balance equations such that the future population numbers and their distributions according to the attributes of interest can be derived.

We formulate the problem of estimating and predicting all demographic components as a unique statistical problem in terms of demographic rates (of events such as death, giving birth and migration) per exposed population and time unit, i.e. when the input is *count* data. This is the main working set-up for the purpose of population projections.

In addition, we have explored binary response data, when modelling the same type of events described in terms of *microdata*. This option was used mostly for analysis and hypothesis testing purposes, e.g. testing whether fertility depends on education of mothers or varies by municipality.

Our solution is to build (structural) generalised additive models [9] with hierarchical structures to account for any clustering effects (by location or by other characteristics such as citizenship, gender or education) and for differences between/within groups of observations. In addition, we impose the condition that the correlation structures of the models account for temporal effects [10]. This type of models (denoted in this paper HGAM-T) are also most powerful for solving small counts/small areas as well as missing data issues since they borrow strength across observations in an efficient way.

The models are estimated in Bayesian framework, with priors chosen according to data exploratory analysis and expert knowledge when available. The Gaussian process priors (over time and/or age) are particularly suitable components due to their ability to describe complex dynamics and phenomena specific to time series, including non-stationarity or periodic and trend components. They can also treat time and age dimensions on equal footing and incorporate the corresponding correlation structures accordingly.

The expected value of the response is therefore written, via a link function when needed, as a sum of *functions* over a set of predictors:

(i) time, age, gender, citizenship, municipality in the case of modelling aggregated counts of events and

(ii) education, family size, municipality size, in addition to time, age, gender, citizenship, municipality, for modelling binary response type of data.

These functions may be smooth (additive or multi-dimensional) ones, such as splines and their tensor products or unknown functions defined by a prior stochastic process (e.g. Gaussian process), with given types of correlation decay (for observations close values of time and/or age) and updated by the observed data points.

The implementation was made straightforward by using reliable R-packages such as *mgcv* [11] (which accommodates a wide selection of smoothers and Gaussian process kernels), *lme4* [12] (for frequentist fast estimates and testing multilevel models), and finally *brms* [13, 14] (Bayesian, based on a *stan* engine [15]).

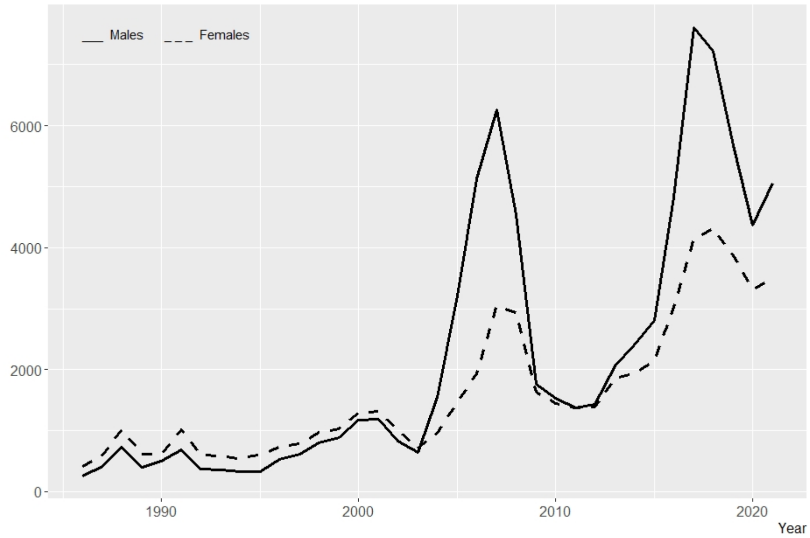
Importance of temporal effects

We emphasize and illustrate in this section the importance of accounting for temporal effects when modeling and predicting the demograhic rates, in a more intuitive way.

Several phenomena compete for influencing the number of births, deaths and migrants, such as the changes in individuals' characteristics but also in social and economic conditions. While the changes in fertility and longevity are rather smooth and take place over long periods of time, the changes in migration patterns are often very sudden and correspond, mathematically, to non-stationary type of time series, with extreme values.

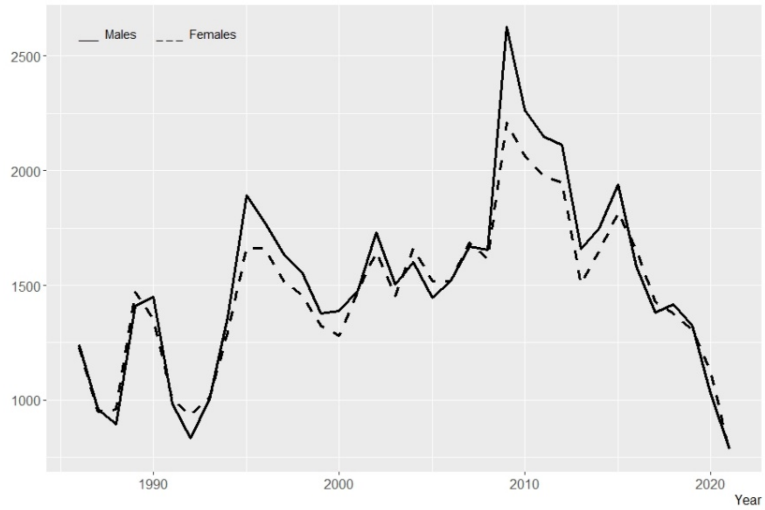
A typical example is the immigration of foreign citizens into Iceland for the past 35 years (see the figure below). We may identify not only the extreme values characterising two recent economic booms, but also the change in the relation between the migration of males and females: over 1986-2003 the women with foreign citizenship immigrating to Iceland were more numerous than the men, while over the following period, 2004-2021, this relation has been completely reversed. Only in the years following the finacial crisis of 2008 the two migration flows had close values, but only for a brief period.

Figure 1a. Immigration of foreign citizens by gender, 1986-2021



Similar patterns and few years lagged were observed for the emigration of foreign citizens, while the migration of Icelandic citizens showed yet another type of evolution, much more balanced and more stable. Only in years of economic pressure were male emigrants dominating the flow and the immigration is similar while time lagged and correlated to emigration.

Figure 1b. Emigration of Icelandic citizens by gender, 1986-2021



These types of temporal effects, non-stationarity and complex correlations over multi-dimensional predictors' surfaces (e.g. age-time or space-age-time[1]) which in addition depend on grouping variables (such as citizenship or gender and their interaction as exemplified by migration patterns) are captured by our HGAM-T models described in what follows.

Modelling the demographic rates

The models used for fitting the past data have a hierarchical structure in two respects:

(i) the observations can be grouped into multiple levels / clusters due to sharing attributes (e.g. gender, citizenship, municipality) and being observed at successive points in time.

(ii) the model itself is formulated in "layers" of complexity: the observations are distributed according to some law (e.g. Poisson[2]), which is defined by a set of parameters (e.g. rate) depending on certain variables (e.g. time, age) via known functions (e.g. linear, splines) or unknown ones (non-parametric, functional). This dependency may in turn be "stratified" by values of some grouping attributes (e.g. gender or municipality). The hyperparameters involved in defining all the functions may be distributed according to some prior distributions and the modelling process ends by finding their posterior distribution via the standard combination of the priors and the data-likelihood. If the functions are **unknown**, we may assume that they are sampled from an infinite function space and impose conditions referring to the decay of the correlations between neighbouring observations (e.g. close in time, in space, as age) with the distance between them. This kind of specification requires priors like the Gaussian *process* priors (*distributions over **functions***) employed in this study.

[1] With a special treatment of cohort effects via interaction terms as shown in the next section.

[2] This is by no means a unique choice and one may test various other distribution functions. Our choice is motivated by model fit but also by its interpretability and simplicity.

The common structure of all demographic rates' models is described by a unique expression:

$$N(x, \dots) \sim Poisson(e^{f(x, \dots)}, N_0(x, \ \dots))$$

$$f(x, \dots) = \sum f_m^{\lambda}(x, \dots) \,,$$

where $N_0(x, \dots)$ are exposed population counts, $N(x, \dots)$ are the numbers of births, migration or deaths at each combination of attributes $(x, \dots)$, $f(x, \dots)$ are the (underlying) Poisson rates, while $f_m^{\lambda}(x, \dots)$ denote one or a combination of the following:

- linear, spline or tensor product type of functions of the attribute(s) and their significant interactions

- unknown functions with Gaussian *process* priors

$$f_m^{\lambda}(x, \dots) \sim GP(\ \mu(x, \dots),\ \Gamma(\{x, \ \dots\},\ \{x', \ \dots\})\ )$$

depending on time and age ($x = (time, age)$), for each combination (λ) of categorical variables such as gender, citizenship, type of migration (i.e. immigration or emigration) and their interaction. The correlation function $\Gamma(\{x \dots\}, \{x' \dots\})$ is defined via its functional form (e.g. spherical, power exponential, Matérn) with (a minimum of) hyperparameters determined during model fitting. The mean function $\mu(x, \dots)$ (defining the trend of the process) may in turn be built by combining (spline, linear or non-linear) functions of the attributes and it plays an important role when using the model in forecasting, where the trend dominates.

A Gaussian process (GP) may be seen as a connection between machine learning and modeling based approaches to statistical inference since it is not able to make anynassumptions about the shape *of the relation between variables on the one hand but it is able to learn from the data the* strength *of the relation between observations, on the other hand.*

Gaussian processes are related to multiple other statistical methods such as reproducing kernel Hilbert spaces, (related) relevance vector machines, spatial modeling and even neural networks of specific types. In addition, most smooth terms like cubic splines or tensor product terms of generalised additive models may be formulated as particular types of maximum a posteriori estimates of Gaussian processes (see [10] monograph or useful online materials such as M. Clark's workshop[1]). An interesting connection exists between GPs and recent deep learning tools such as the use of dropout and its variants in neural networks. This connection actually provides a method for representing model uncertainty in deep learning [16].

The choice of **models** was based on model selection and automatic performance optimisation (implemented by the *brms* and *mgcv* R-packages) illustrated in Appendix 1 and on model evaluation techniques as described in section 5. They were fitted in both frequentist and Bayesian settings and contain the following terms:

- birth rates: the smooth (unknown) functions are determined by GPs depending on age, one for each value of citizenship of mothers (Icelandic and non-Icelandic), the model has age and time interaction terms (similar to including cohort effects), in addition to a citizenship grouping effect.

- mortality rates: the smooth functions are determined by GPs depending on age, one for each gender value, the model contains age and time interaction terms (similar to including cohort effects), in addition to a gender (grouping) effect.

- migration rates: smooth functions are determined by GP/tensor product depending on age and time, the model contains main effects and interaction terms depending on gender, citizenship and direction of international migration.

The **training data** used for fitting the main models employed in the production of the national population projections consists of:

- the number of births by age and citizenship of mother, over the period 1998-2021 and the number of "exposed" (with the chance of giving birth, i.e. in standard fertility age brackets) women by age and citizenship.

---

[1] https://m-clark.github.io/workshops/stars/

- the number of deaths by age and gender, over the period 1998-2021 and the number of "exposed" (with the chance of dying) population by age and gender.

- the number of migrants by age, gender, citizenship and direction of external migration over the period 1986-2021 and the number of "exposed" individuals (i.e. either the population-source of emigration or, less obvious, the population-target of the immigration process) by age, gender, citizenship.

The approach to modelling immigration is simplified as it only includes information about the "target" of the immigration process but not about its source, i.e. the population of other countries. We thus only model the "pull" factor. There are multiple studies in literature concerning attempts of including this type of effects with mixed results (see [5], [19] for a detailed discussion). Until now, the approach we adopted here has been proven most efficient. The reason why it performs rather well is that the detailed past migration patterns capture the effects of the external influences which are assumed to be similar in the future.

Microdata analysis, as illustrated[1] in Appendix 1, has proved that: (i) fertility rates depend on age, citizenship, education of mothers as well as on municipality and type of family the mothers belong to (ii) mortality rates do not vary significantly by municipality but do vary by gender and (iii) migration rates vary by age, gender, citizenship and municipality. The most significant attributes were included in the first generation of models used for population projections in 2022, which are based on aggregated counts and not on microdata.

Stochastic cohort component method

Versions of stochastic cohort component method have been previously employed although for different types of the demographic components' assumptions/models. In our case, it accompanies the models described in the previous section. Its main steps consist of:

- the rates of migration, fertility and mortality are sampled from their posterior distributions at each time step $(t+1)$ by using the corresponding models and the exposed population numbers at time step t

- these samples from the posterior distributions of the demographic components are combined, at each time step according to the standard balance equation $P_{t+1} = P_t + M_{t,t+1} - D_{t,t+1} + B_{t,t+1}$, written for all cross-classified groups according to the set of used attributes (e.g. age, gender, citizenship). Here P denotes population numbers while M-migration, D-death and B-birth numbers. The values P_{t+1} provide the exposed population numbers at time step $(t+1)$

- the process is repeated a large number of times (e.g. K=10,000) in order to generate the posterior distributions of population numbers at each value of $t > t_0$ where t_0 is the starting time of projections.

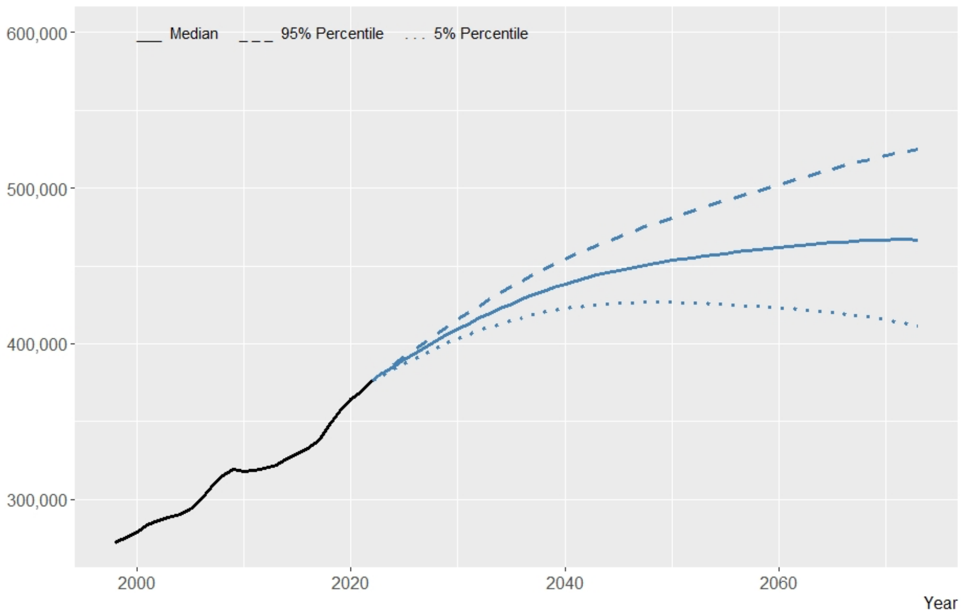
[1] Several credible intervals and conditional effects plots were generated from the microdata models tested for this purpose.

3. Illustrative results: national projections 2022-2073

In this section we illustrate the new method by showing the results obtained for all demographic components (fertility, mortality and migration) as well as their stochastic combination into the population predictions in 2022.

The *population of Iceland* is projected to grow from 376 thousand (registered population 1st of January 2022) to an estimated value of resident population between 412 and 525 thousand in the next 50 years, with 90% probability, according to the models. The median population value in 2073 is 467 thousand. The upper limit (525 thousand) indicates that there is a 5% chance that the actual number of inhabitants will be higher than this value and a 95% chance that it will be lower. The lower limit (412 thousand) indicates that there is a 5% chance that the actual population will be smaller than this value and a 95% probability that it will be higher.

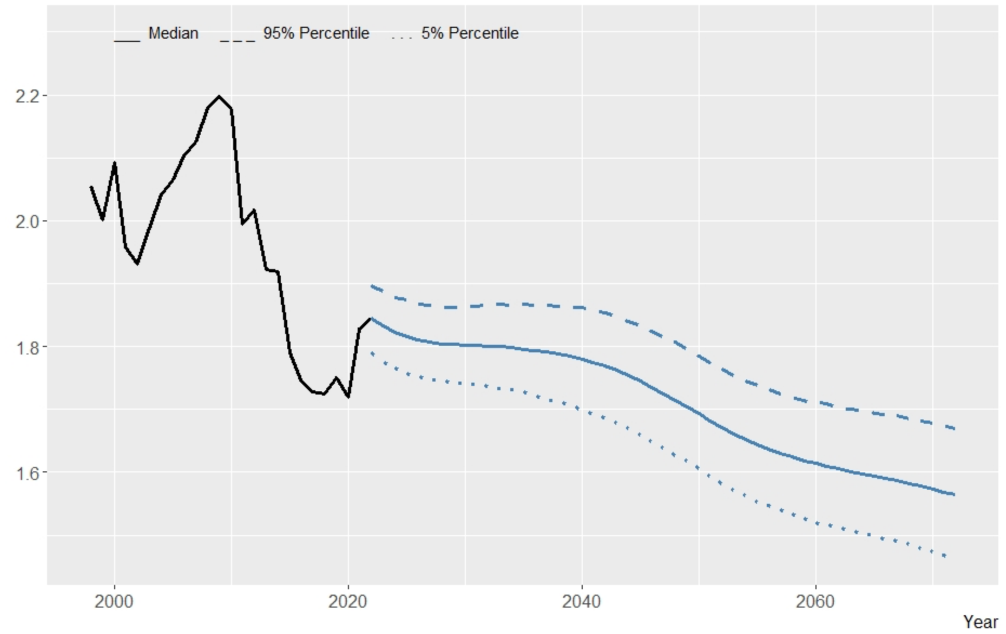
Figure 2. Population 1998-2073



The *total fertility rate* should reach 1.6 children per woman (of 13-55 years of age) in 2073, according to the median projection. Total fertility rate could have values between 1.5 and 1.7, with 90% probability, in 2073. The lowest predicted rate (1.5) for Iceland in 2073 is yet higher than the average EU total fertility rate of 1.4 which was reached in 2022.

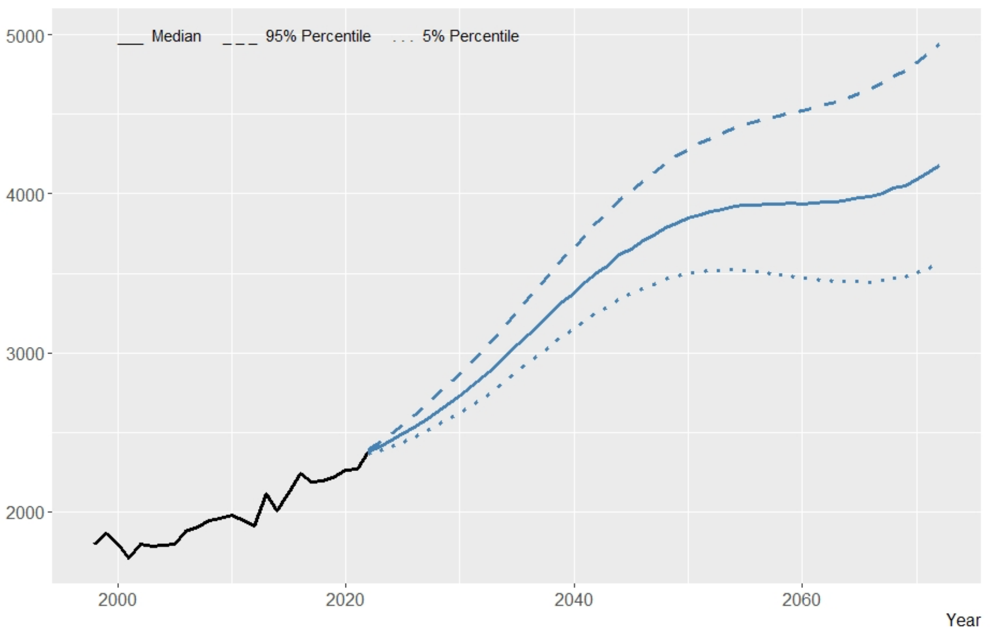
The following results (detailed in Appendix 1) are important when trying to understand why the total fertility decreases: (i) women with foreign citizenship tend to have lower fertility rates and their number is increasing with time and (ii) the strongest decreasing trend in fertility is recorded for women in younger age groups of both Icelandic and foreign citizenship.

Figure 3. Total fertility rate 1998-2073



Life expectancy at birth will grow from 84 years in 2022 to 89 years in 2073 for women and from 81 to 84 years for men, according to the median projection, i.e. an increase of about 0.1 years of life per year.

Figure 4. Number of deaths 1998-2073



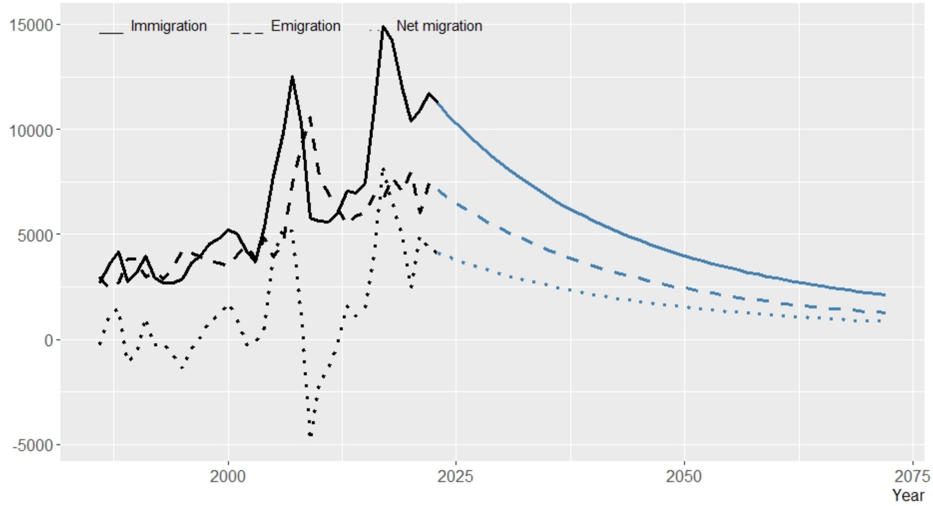
Migration 1998-2073

The number of immigrants will be higher than the number of emigrants for the whole period, due primarily to the migration of foreign citizens. The net migration of Icelandic citizens (i.e. the difference between arrivals and departures) will preserve its past character, varying around zero for the next 50 years.

Cautionary remark:

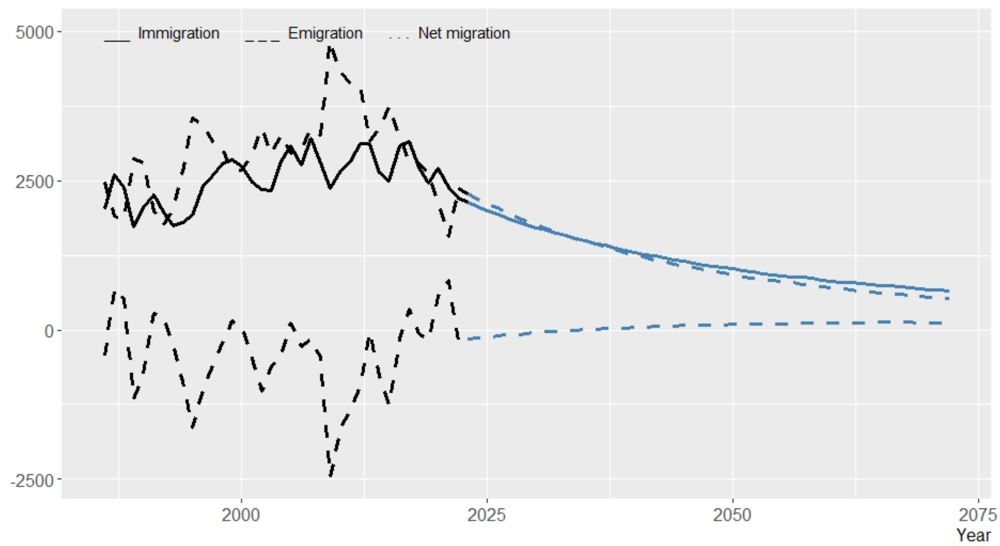
The predictions of the migration patterns are very conservative in the current release, even considering de-registration corrections and the fact that only main trends are predicted by the models. They do not include the unusual high number of refugees which could not be predicted. Only non-/weakly- informative priors were used in 2022 and we emphasize again that the priors' role is much more important when forecasting than when fitting a model. This is why, although the long-term values of predicted migration numbers coincide with expert assumptions (as they were described in [4] and yearly publications of Statistics Iceland), we formulate here the conjecture that these forecast values are too low. This is primarily because external conditions have changed dramatically during the very recent past with respect to the rather stable conditions of the previous few decades. More prior information needs to be incorporated into the models for a more realistic outcome.

Figure 5. Total immigration, emigration, net migration (the difference between immigration and emigration) 1998-2073, median values



The main difference between the predictions of the current and previous models is that the Icelandic net migration should reach positive, although very low, values over long term.

Figure 6. Migration of Icelandic citizens 1986-2073, median values



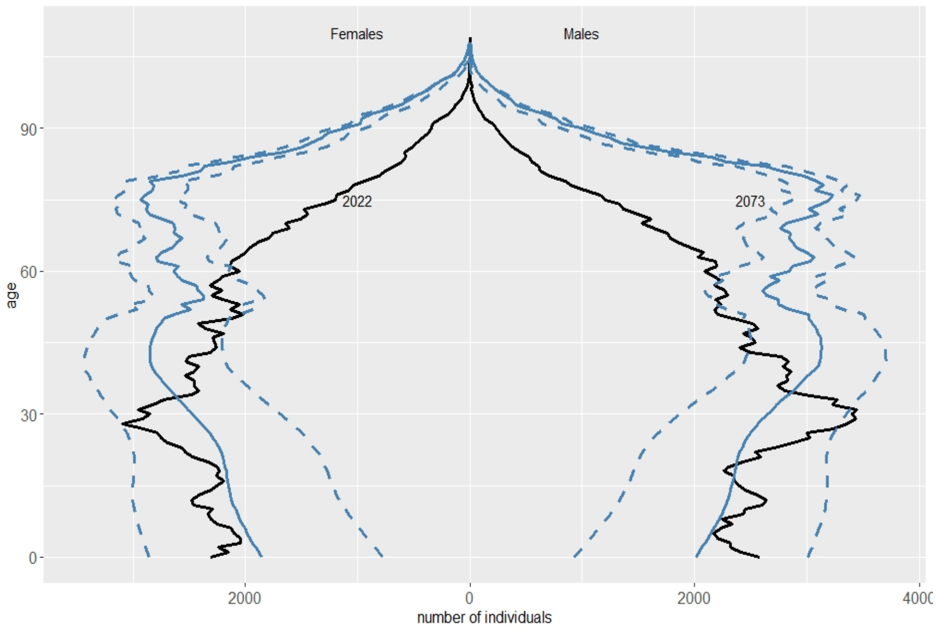
As a consequence of the evolution of birth, death and migration numbers, the total *population structure* is predicted to change as follows:

(i) The proportion of the working age population (defined as 16 to 74 years old) will decrease from 74% in 2022 to 68% in 2073, according to the median projections

(ii) After 2055, the older age population (over 65 years) will become more numerous than the youngest (less than 20 years old), according to the median projections

The change in population age-distribution is illustrated usually by comparing the population pyramids of the starting and end years of the projection horizon.

Figure 7. Population age distributions 2022 and 2073 (median, 5% and 95% percentiles)



Population ageing is caused by decreasing fertility and increased life expectancy. However, this development is still slower in Iceland than in the EU countries, due to high migration and relatively high fertility compared with the EU average. The median age of the Icelandic population was 36 years in 2021 and is predicted to be 46 years in 2073. In comparison, the median age in the EU countries was already 44 years in 2021, a value which will not be reached in Iceland until 2058, according to the median projection.

4. Building the local projections 2022-2073

In order to produce the local population projections, one may start by modelling the demographic rates as described in the sections above while including the dependence on municipality into the model as well. This would provide local projections and therefore national projections as (stochastic) combinations of all local ones. It would also be computationally intensive and resource demanding, with increased uncertainty of estimates due to high number of imputed values and stochastic addition of multiple time series.

Therefore, we use a different, more robust and simpler approach. The local population "rates", defined by the local counts and the "exposed" total population of Iceland, are modelled as smooth functions of several attributes and/or time. The complexity of this type of model can be increased, while the priors can be made more informative, as additional information or new analysis become available. The current simple model of local population rates includes multiple time dependent GPs, one for each municipality as well as municipality-level smooth intercepts.

Remarks on strengths and *limitations* of the current, beta version of the local projections (see Appendix 2):

- the local rates are modelled in this first version only by using the national median, *resident* population projection, therefore the widths of the credible intervals are underestimated. If we sampled from the posterior distribution of total population values for each time point of the forecasting horizon, the generated credible intervals of the local projections would be wider.

- additional information should be collected from users, experts, policy makers. It could be incorporated as *prior* information in order to improve the model predictions in the future.

- these is only the first, simplest and agnostic, in terms of priors, version of the local projections. The projections will be published as **experimental** statistics[1] as soon as the additional information is collected. This will allow us to build informative priors and update both the total population and the predictions of population numbers by municipality.

- some local predictions have large uncertainty (e.g. many of the municipalities with small populations, especially if the counts have been oscillating in the recent past) while others look rather confident (e.g. municipalities with larger population, but also some of the much smaller ones if showing very clear past development trend). Note also that the reported uncertainty reflects at this point only the predictive abilities of the *models*.

- a preliminary evaluation of results by using subjective opinions of various experts showed that some of the main trends predicted by us are too optimistic while others are too conservative (usually smaller ones). These cases are also characterised by wide uncertainty limits which show that much higher and much lower values than the median predictions are also rather likely to occur.

[1] „Experimental statistics is intended to support official statistics and improve their quality by facilitating innovation in the production of official statistics and give users faster access to interesting statistical information." – see https://www.statice.is/publications/experimental-statistics/

5. Evaluation of the new method

In order to evaluate the performance of the population projection methods, the models of all demographic components as well as their stochastic combination need to be carefully scrutinised.

The evaluation of these models was performed in two steps:

(i) the model *fitting and validation*, as exemplified by the posterior checks in Appendix 1 (while optimization of model components is already included in the automatic fitting and model selection procedures) and

(ii) the "in-sample forecast" method for evaluating the model *forecasting* performance

The observations used for fitting these models are time-correlated. Therefore, the typical methods for measuring accuracy, such as cross-validation statistics, should be calculated in a special way[1] since leaving out observations does not remove their influences on other data points. The method applied for such cases is known as the "in-sample" forecast and consists of generating independent (time step=1, …) forecasts based on the model which needs to be evaluated, the forecasts being thus performed for shorter and shorter time series of input data. The errors measured by comparing the predicted and observed values are then combined into standard measures such as RMSE, MSE, MAE[2].

The values of the RMSE on the test data, when training of the models was done for the period 1998-2012 and the (in-sample) testing over the period 2013-2022, are:

1.7 (median) for the number of deaths (conditioned on age-group and gender), with 1st and 3rd quantiles equal to 0.8 and 3.9 respectively; 6.8 (median) for the number of births (conditioned on age-group and citizenship of mothers), with 1st and 3rd quantiles equal to 2.4 and 14.8 respectively; 4.3 (median) for the number of migrants per age-group, citizenship and movement direction of migrants, with 1st and 3rd quantiles equal to 2.4 and 12.1 respectively.

The RMSE values on the training data, i.e. based on residuals/model fitting are much smaller, ranging from 1.01 for the elementary (conditioned on the other attributes, as above) number of deaths, to 1.3 for migrants and 1.6 for births, for the whole period 1998-2022.

A detailed comparison between mortality forecasts based on several methods (employing GP similar to ours and functional models) is described in [7]. It shows narrower confidence intervals of the functional models (underestimating uncertainty) and small differences between estimates of both models, without a systematic trend of these discrepancies. It confirms the fact that GPs (over time and age) are trustable and reliable forecasting tools.

The evaluation of the total population predictions based on multiple in-sample forecast of demographic components combined via the *stochastic* cohort component method is work in progress. This exercise is computationally intensive and time consuming. But since the demographic components perform very well, we work on the assumption that the stochastic combination of the predicted rates will also do that. We did however compare the results of all methods used to produce population projections during the last 10 years, although different since continuously updated, as shown in the following section.

---

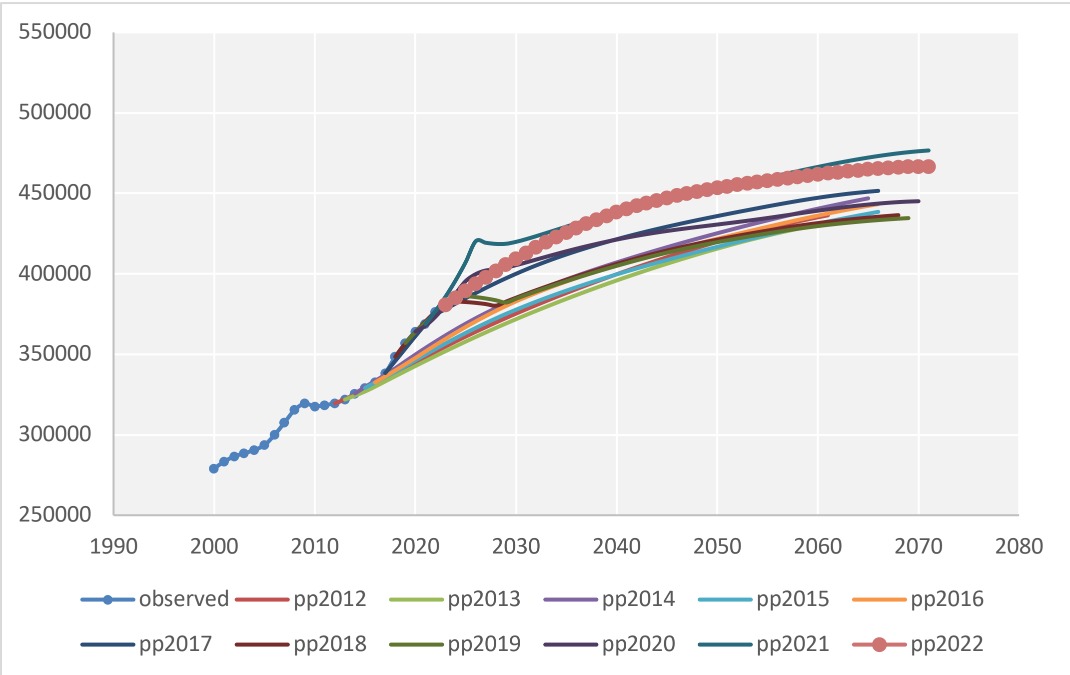
[1] https://robjhyndman.com/hyndsight/crossvalidation/

[2] Notations stand for: root mean square error, mean square error, mean absolute error. respectively.

6. Previous population projections published by Statistics Iceland

The population projections published during the past ten years have been based on multiple methods and models which were continuously improved and updated. In the following figure all predictions of 2012-2022 are plotted in addition to the observed population data. One may easily notice that all projections before 2017, when the most recent economic boom has started, have underestimated the population of the years that followed, especially short-term. The newer projections have incorporated this information and predicted higher short-term population levels, accordingly. The long-term predictions of all past projections (main variants) are not very different from each other. They span an interval of about 30 thousand inhabitants in the year 2060 for instance, i.e. less than 7% of the average predicted population total for that year. The credible interval of the 2022 predictions (not shown in this figure but in the first figure of this paper) covers this interval.

Figure 8. Population projections (pp2012, ...pp2022) published between 2012 and 2022, compared with the observed data



This comparison and the data behind it are available on the dedicated github repository[1] where any future updates will be uploaded as soon as they become available.

[1] https://github.com/violetacln/SIPP

7. Conclusions

New population projection methods

In this paper we describe the new methods for producing population projections at Statistics Iceland, including preliminary work regarding projections by municipality. They are based on hierarchical, Bayesian models of demographic rates and on a stochastic cohort component method of combining the projected rates in order to provide estimates of future population numbers and associated uncertainty measures. This is, to our knowledge a first exercise in using this particular type of models for all components and for producing corresponding stochastic projections.

New type of data, software and results

The main new features of the projections are as the following. The results are defined in terms of median values and lower/upper bounds of credible intervals (e.g. 90% probability level) instead of low/high variants based on scenarios. The predictions refer to *resident* (as opposed to register) future population. The source code is shared openly and can be easily tested on the open data on Statistics Iceland's web-site, in a straightforwardly simplified version. We estimate that if efficient solutions for confidentiality issues are found, the most detailed data-sets could become open as well.

Work in progress

We are working towards improving and developing the methods along the following lines:

- including prior information, based on feedback from experts, users, policy makers, local administration, in order to produce **informed updates** of the total population projections and publish experimental then official, local projections. This would improve the quality and reduce the uncertainty of predictions

- performing sensitivity analyses for testing the effect of changes in prior distributions

- building a unique model for local and total population projections

- incorporating new significant attributes in the production version, such as education

- testing the use of Markov Random Fields for modelling local effects more accurately

- adding more municipality-level attributes into the model

- testing rigorously the performance of a model averaging approach (still showing mixed results at this point)

In addition, technical and theoretical refinements are needed for the forecast evaluation step, as well as increased quality of implementation and use of computational resources.

References

1. Hyndman R J, Booth H (2008) Stochastic population forecasts using functional data models for mortality, fertility and migration, International Journal of Forecasting 24(3), 323-342.

2. Hyndman R J, Booth H, Yasmeen F (2013) Coherent Mortality Forecasting: The Product-Ratio Method With Functional Time Series Models, Demography 50(1), 261-283.

3. Bijak J & Bryant J (2016) Bayesian demography 250 years after Bayes, Population Studies, 70:1, 1-19, DOI: 10.1080/00324728.2015.1122826

4. Calian V, Hardarson O (2015) Methodology of population projections, Statistical Series of Statistics Iceland, https://hagstofan.s3.amazonaws.com/media/public/6210b1e3-cd70-4a7b-8bab-dc957243dc4c/pub_doc_CmyUZfa.pdf

5. Bryant, J., & Zhang, J. L. (2016). Bayesian forecasting of demographic rates for small areas: emigration rates by age, sex and region in New Zealand, 2014-2038. *Statistica Sinica*, *26*(4), 1337–1363. http://www.jstor.org/stable/44114337

6. Lam, K.K.; Wang, B. (2021). Robust Non-Parametric Mortality and Fertility Modelling and Forecasting: Gaussian Process Regression Approaches. *Forecasting* 2021, *3*, 207-227. https://doi.org/10.3390/forecast3010013

7. Ludkovski, M., Risk, J., Zail, H. Gaussian process models for mortality rates and improvement factors. ASTIN Bulletin: The Journal of the IAA , Volume 48 , Issue 3 , September 2018 , pp. 1307 – 1347, https://doi.org/10.1017/asb.2018.24

8. Zhang J. L., Bryant- J., Nissen K. (2019). Bayesian small area demography. Survey Methodology, 2019 (special issue) 13, Vol. 45, No. 1, pp. 13-29 Statistics Canada, Catalogue No. 12-001-X.

9. Wood S (2017). *Generalized Additive Models: An Introduction with R*, 2nd edition. Chapman and Hall/CRC.

10. Rasmussen C. E. and Christopher K. I. Williams (2006) Gaussian Processes for Machine Learning, MIT Press, ISBN-10 0-262-18253-X, ISBN-13 978-0-262-18253-9, http://gaussianprocess.org/gpml/chapters/

11. https://cran.r-project.org/web/packages/mgcv/index.html

12. Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using *lme4*. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

13. Bürkner P (2017). "brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software*, **80**(1), 1–28. doi: 10.18637/jss.v080.i01

14. Bürkner P (2018). "Advanced Bayesian Multilevel Modeling with the R Package brms." *The R Journal*, **10**(1), 395–411. doi: 10.32614/RJ-2018-017

15. Stan Development Team (2022). "RStan: the R interface to Stan." R package version 2.21.7, https://mc-stan.org/

16. Gal Y., Ghahtamani Z. (.2016) „Dropout as a Bayesian approximation: representing model uncertainty in deep learning", Proceedings of the 33rd International Conference on Machine Learning, PMLR 48:1050-1059. http://proceedings.mlr.press/v48/gal16.pdf

17. Raftery E. A., Li N., Ševčíková H., Heilig G.K. (2012) „Bayesian probabilistic population projections for all countries", PNAS 109(35)13915-13921, https://doi.org/10.1073/pnas.1211452109

18. National population projections: 2022(base)-2073, Statistics New Zealand, https://www.stats.govt.nz/information-releases/national-population-projections-2022base2073/

19. Coen van Duin (2017) „Quasi stochastic population forecasts", https://www.cbs.nl/en-gb/background/2017/04/quasi-stochastic-population-forecasts

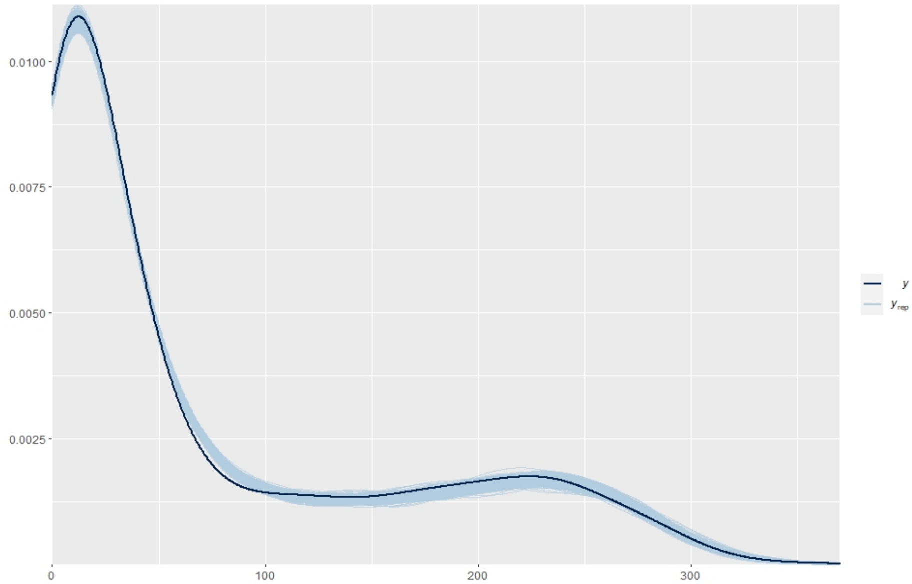
20. Statistics Netherlands, „Population in the future", https://www.cbs.nl/en-gb/visualisations/dashboard-population/population-dynamics/population-in-the-future

Appendix 1: Model fitting and model validation for demographic rates

Posterior checks for model validation

We exemplify this type of validation of the model with the results of posterior checks for the Bayesian fertility model *(brms* fitted). This means that one compares the observed data to simulated samples (or some summary statistics) generated from the model **posterior** predictive distribution.

Figure 9. Comparing the observed and predicted values (posterior checks)



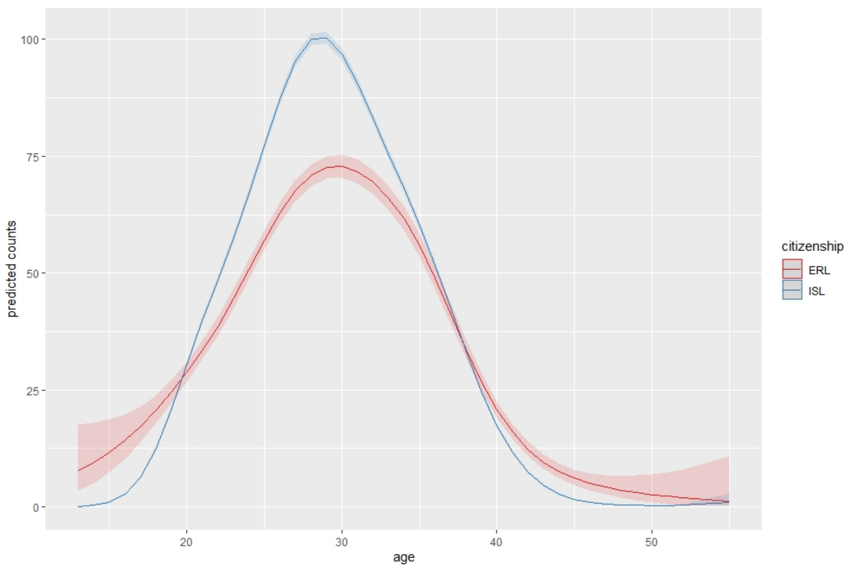
Fertility

The effects of age, time and citizenship according to the models without spatial dependence (i.e. based on total population counts of the number of births by time, age and citizenship of mothers) are illustrated in the figure below and display the main trends observed for significant age groups of mothers. The number (*mgcv* fitted, for speed) model of births shows the strongest decrease with time for the most typical young age-group (22) while it is slightly increasing with time for the higher age-groups (31 and 40). Women of all age groups and citizenships had more children around the years 2009-2010.

Figure 10. Dependence of the number of births on time by citizenship of mothers and significant age-groups of mothers

All variants of the tested fertility models show a well-known pattern of the last almost 20 years, with the most likely age of mothers close to 30. Most municipality effects (few exceptions only) are statistically non-significant. The tests also showed that women of Icelandic citizenship have higher fertility rates than the ones with foreign citizenship (see both figures in this section). The uncertainty of model predictions is highest for data corresponding to small number of events (births) such as the ones recorded for extreme ages (youngest and oldest) of mothers with foreign citizenship (see the figure below).

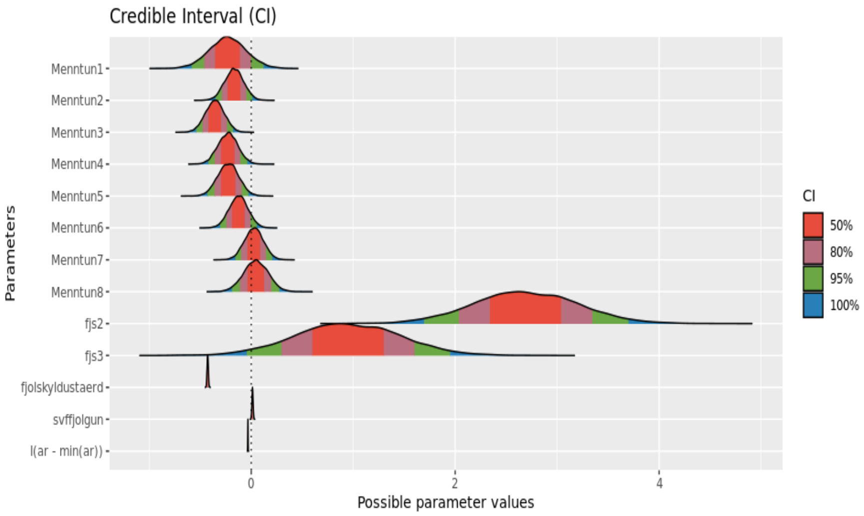
Figure 11. Dependence of the number of births on ages of mothers, by citizenship (ISL- Icelandic and ERL- foreign)



Same results for (conditional) effects were obtained with/without including spatial attributes (*municipality*) in a Bayesian hierarchical model *brms* fitted.

The preliminary study based on fertility microdata confirmed (see figure below[1]) that education levels have an effect (although rather small) on fertility, the variation of municipality population between years (seen as a characteristic of the municipality) has a non-significant impact but the main factors influencing birth events are the size of family and the type of partnership between parents.

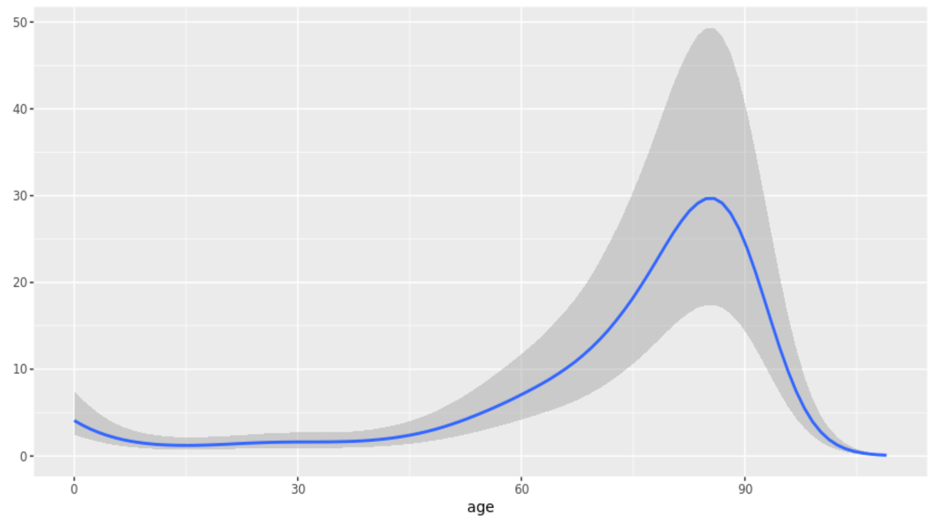
Figure 12. Effects of education, type of family and characteristics of municipality on fertility (model parameters) and their uncertainty measures (credible intervals)



Mortality

Exploratory analysis was conducted for models of mortality as well, based on microdata and on aggregated counts before validation of the final model. Bayesian models, with and without spatial dependence, brms-fitted, confirmed both the most likely ages of death and the growing estimation uncertainty with age. The tested models also proved that location (municipality) does not have a significant effect on mortality rates while the gender does.

Figure 13. Dependence of the number of deaths on age (conditional effect)

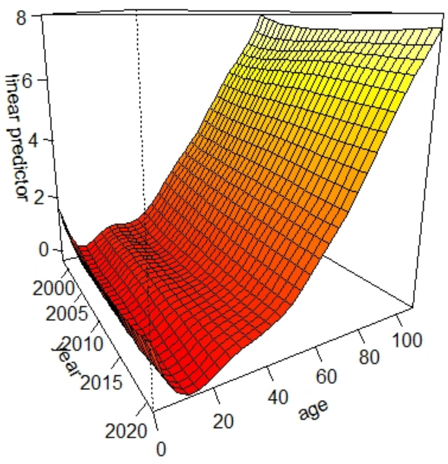


---

[1] This model was fitted using the *lme4* R-package and the credible intervals were generated by simulating with the *arm* and *bayestestR* R-packages.

The predicted surface of log-mortality rates as a function of age and time may be generated faster by using *mgcv*-package to fit the model, as shown in the following figure:

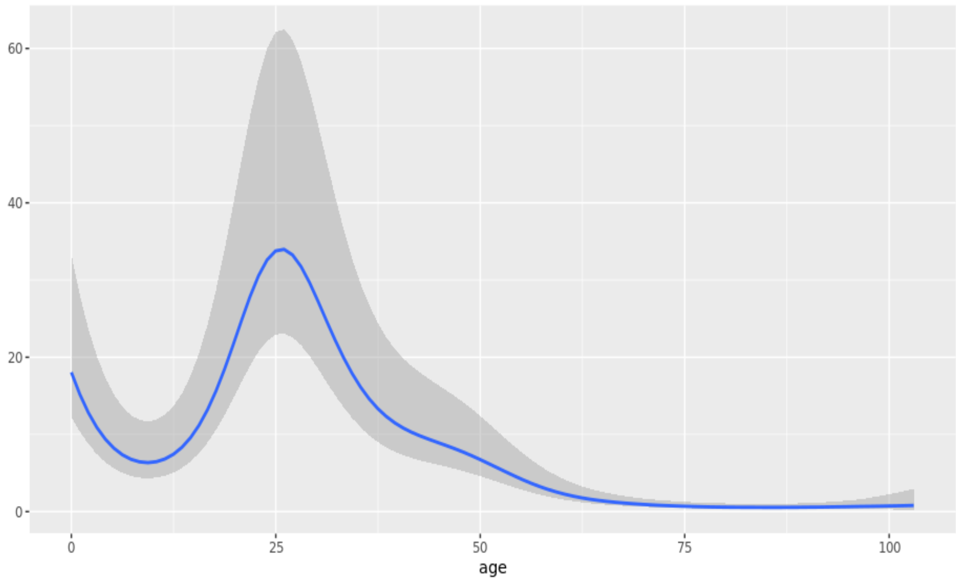
Figure 14. Mortality rates (log-scale) by age and time



Migration

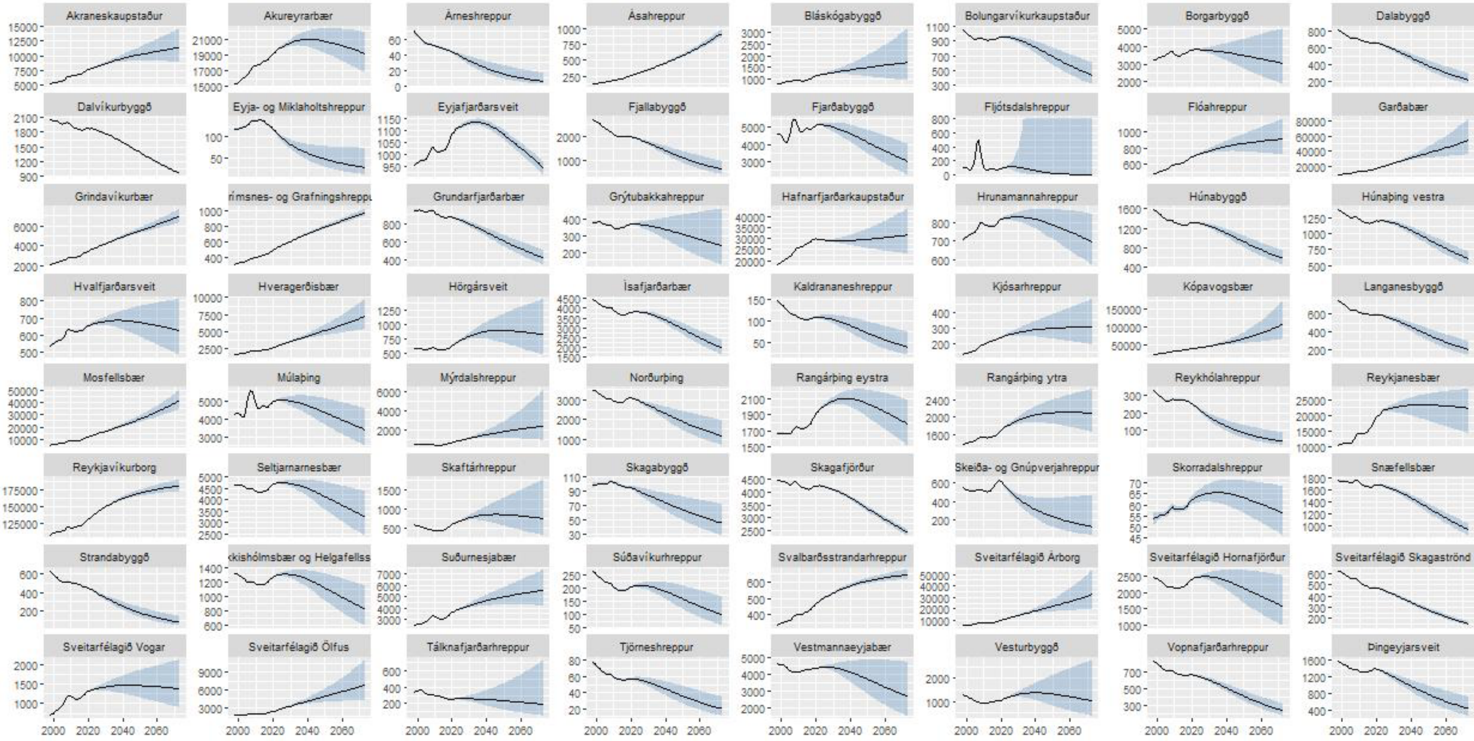
Employing a similar methodology to the one used for fertility exploratory and validation stages, Bayesian models, with and without spatial dependence, were brms-fitted for migration (by type). They show that the most likely age of migrants is around 27, followed closely by the very young ages of their children. The rates depend on municipality, in addition to age, citizenship and gender.

Figure 15. Dependence of the number of migrants on age (conditional effect)



Appendix 2: Local population projections, experimental results based on non-informative priors

Figure 16. Experimental, local population projections 2022-2073



Experimental statistics: Population by municipality 1998-2073

Hagtíðindi Mannfjöldi

Statistical Series *Population*

Vol. 108 ◆ Issue 4 15 August 2023

ISSN 1670-4770 Umsjón *Supervision*

Violeta Calian ◆ violeta.calian@hagstofa.is