



Guillaume Marois and Michaela Potančoková

QuantMig microsimulation tool

Deliverable 8.1



QuantMig has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 870299.

History of changes

Version	Date	Changes
1.0	28 Feb 2022	First version submitted as official deliverable to the EC
1.1	17 July 2023	Dissemination to public

Suggested citation

Marois, G. and Potančoková, M. (2022) QuantMig-mic microsimulation tool. QuantMig Project Deliverable D8.1. International Institute for Applied Systems Analysis (IIASA)

Dissemination level

CO Confidential: Consortium only, including the European Commission services until release of deliverable 8.2

PU Public together with D8.2 from month 42

Acknowledgments

This document reflects the authors' view and the Research Executive Agency of the European Commission are not responsible for any use that may be made of the information it contains. We are very grateful to Patrick Sabourin for his contributions.

Cover photo: [iStockphoto.com/Guenter Guni](https://www.iStockphoto.com/GuenterGuni)

Table of Contents	
Acknowledgments.....	i
List of Tables	1
List of Acronyms.....	1
1. Introduction.....	2
2. General description of the QuantMig-mic microsimulation model.....	2
2.1 QuantMig-mic structure and modules.....	4
2.1.1 Base population.....	5
2.1.2 Modules.....	6
2.1.3 Model calibration.....	9
3. Deliverable description.....	9
3.1 Technical information for the user	9
3.2 Modgen.....	10
3.3 Model variable codes in the user interface.....	11
References.....	13

List of Tables

Table 1 Overview of the events modelled in QuantMig-mic and their determining status variables

List of Acronyms

EU LFS – European Union Labour Force Survey -

<https://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey>

Eurostat - Eurostat is the statistical office of the European Union - <https://ec.europa.eu/eurostat>

Frontex – European coast guard and border agency - <https://frontex.europa.eu/>

Research projects

CEPAM - The Centre of Expertise on Population and Migration <https://migration-demography-tools.jrc.ec.europa.eu/catalogue/dataset/pj00207>

QuantMig - Quantifying Migration Scenarios for Better Policy

1. Introduction

This is the accompanying technical cover document to the delivered code of the QuantMig-mic, a microsimulation population projection model for 31 European countries developed within the scope of the QuantMig project. This cover document is describing in the nutshell the model and provides information about its structure and its modules. It also walks the user through the delivered files with the model code and provides the users with the information about how to handle these files, how to open the executable file and view the model in the user interface. This document is not a full model documentation which will be delivered in the following deliverable 8.2 and will also include detailed description of data sources and statistical models employed to derive the model parameters. This forthcoming deliverable is cross referenced at relevant points. It also is not a detailed tutorial on how to install and use Modgen, for which we refer the users to the relevant webpages and the support of the Statistics Canada which maintains Modgen and can assist users with installation issues upon request.

The deliverable includes preliminary results of an example scenario and allows the user to view the results and model parameters in a user interface upon installation of Modgen Prerequisites. In support of the replicability of the research findings from the WP10 of the QuantMig project the model code will be open via ZENODO repository at the end of the project and will be accompanied by full detailed documentation of the model, its methodology, parameter estimates, details about the baseline population estimation and projection assumptions used in the scenarios. While we do not expect any major changes to the model code, some parameters estimates and especially the data inputs into the migration module from other work packages which are work in progress. The current version integrates preliminary migration flows data from the WP6 and will be updated in the next months. The results which are part of this deliverable should, thus, be treated as illustrative and preliminary. Therefore, the deliverable is confidential until the end of the project and will become public thereafter.

The sections below give general description of the microsimulation population projection model, the conceptualisation of the model and its components (Modules section), and information about the programming language and tools used to code the model (Modgen section) and briefly explains how the user can run the executable file and view model content and results (Technical information for the user).

2. General description of the QuantMig-mic microsimulation model

Microsimulation is an alternative approach to the deterministic macro-level population projection models that use aggregate level data to project future population size and composition using cohort-component method. Although multidimensional population projections, which are an extension of the classic cohort-component model, can project populations by larger number of characteristics than age and sex (Rogers 1986), microsimulation is a powerful tool that can replace traditional multistate projections when the number of dimensions becomes large (Van Imhoff and Post 1998). Microsimulation allows for including a larger set of dimensions (characteristics) than the standard multistate population projection models because the handling multiple variables and states is more flexible. Microsimulation also allows for the use of complex statistical models to project life-course transitions and events. Microsimulation the modelling is based on individual level data

and the transitions between the states are determined stochastically.

QuantMig-mic is a continuous time dynamic microsimulation projection model developed for projecting population of the 31 countries including the current European Union (EU) 27 Member States, the United Kingdom, Iceland, Norway and Switzerland. We further refer to this country grouping as the EU+. The model simultaneously projects demographic (age, sex, country of residence, region of birth, immigrant status), socioeconomic (educational attainment, labour force participation) and socio-cultural (religion, language spoken at home) characteristics of the population of all 31 countries. Individuals from the base population are simulated one by one and their characteristics are modified through scheduled events whose timing is determined by Monte-Carlo process, meaning that a random number is generated and compared to the probability of experiencing given transition. If this number is larger than the given probability, then transition happens and the person's state changes.

The model is built following the framework proposed by Belanger et al. (2019) to study population changes in a context of relatively high immigration and low fertility. Simultaneous simulation of individual life courses allows the model to dynamically update the risks of various events based on an individual's state. QuantMig-mic allows for changes in individual characteristics over the life course as well as for intergenerational transfers of some characteristics from mother to child. This interaction is used to model educational attainment of individuals following evidence on intergenerational transmission of education which differs across social groups (Marois et al. 2019a). Educational attainment, in turn, is instrumental in modelling of demographic events and labour market participation of population groups (Marois et al. 2019b).

QuantMig-mic is a microsimulation model programmed in Modgen. It is case-based, meaning that each individual is simulated separately from other individuals and that no interactions between individuals are allowed (with the exception of interactions between mother and children). The model is also dynamic and in continuous time, meaning that characteristics of individuals are modified continuously in "real time", in contrast to discrete-time models where characteristics are changed within predefined time units (typically one year such as in population projections of the EUROSTAT or five years such as in population projection of UN).

QuantMig-mic microsimulation tool is developed for the analysis of the future structure and composition of European populations according to different migration scenarios with respect to volume, composition and socio-economic integration of foreign-born populations. The modelling starts with a baseline population which represent population of all 31 countries at the time of 2011 censuses. The horizon of the simulation is set to 2060 but the tool allows the users to set later projection horizons (given that projection assumptions for later years are available and provided into the model).

QuantMig-mic microsimulation tool allows the study of alternative scenarios and their consequences for future population trends in the European Union (at aggregate level as well as at the level of individual Member States). These scenarios are not part of this report, an illustrative scenario is provided as part of the deliverable. This scenario only includes preliminary migration data which will be updated with deliverable 6.4 at a later point. QuantMig-mic model allows analysing the implications of various immigration flows – in terms of their volume and composition by region of origin of immigrants – and their implications on population level indicators as well as on EU's and individual countries' labour force.

The base population corresponding to 2011 census is used in pre-simulation to 1st January 2020. Presimulation is calibrated to demographic events (births and deaths) as reported by Eurostat for 2011-2019, labour force participation from the LFS and migration will be calibrated to D6.4.

Simulation of projection scenarios starts from 2020 and is set to 2060.

The model was built at IIASA and is an extension of the CEPAM model which was developed at IIASA and the Joint Research Centre within the scope of the Centre of Expertise on Population and Migration (CEPAM) project during 2017-2019. QuantMig-mic extends the previous model with respect to the number of countries that are being simulated and includes now EFTA countries (with population of at least 200000) which have sizeable foreign-born populations and which are part of the European Economic Area. We also extended previous work by modifications to the base population, and the coding and estimation of model parameters in fertility, labour force and migration modules (see the Modules section). These modifications were made in order to analyse the demographic impacts and labour force integration outcomes of immigrants under alternative migration scenarios and for subgroups of migrants by their region of birth. The migration module has been developed to account for modelling of migration by individual country of birth for those born in 31 European countries and region of birth for those born outside Europe¹ (subject to the availability of migration flows by country of origin, country of destination and country of birth). All model parameters were updated using the most up to date data from the EU-LFS and ESS. Detailed description of the statistical models used to estimate the parameters, the input data and documentation of all modifications will be part of D8.2 deliverable.

2.1 QuantMig-mic structure and modules

This section describes the conceptualisation of the model events modules of the QuantMig-mic. The model can dynamically project the population of EU+ countries with the following 13 demographic, socio-economic and socio-cultural variables:

- age²
- sex [male, female]
- country of residence [31 European countries]
- country or region of birth [31 European countries and 8 broad world regions³]
- labour force participation [active, inactive]
- employment [employed, unemployed]
- age at immigration⁴
- duration of residence [0-4, ..., 20+ years]
- educational attainment (3 categories based on ISCED 2011 classification)⁵

¹ CEPAM model allowed to compare only between EU28-born and born outside EU28.

² The model runs in continuous time; 5 year age groups are used in the output [0-4, ..., 95+].

³ The world regions are determined by the classification used the EU-LFS: Europe Outside EU-28 and EFTA (includes Turkey, Russian federation and other Eastern European countries); North Africa; Other Africa; Near Middle East; East Asia; South South-East Asia; North America & Australia Oceania; Latin America. For the country grouping see LFS documentation:

<https://ec.europa.eu/eurostat/documents/1978984/6037342/Country-codification-from-2012-onwards.pdf>

⁴ The variable generation status is derived from the age at immigration. The categories of this variable are the real age. Then we create 3-categories more general categories of the variable for the generation status: G1 – immigrants who arrived as adults, G1.5 – immigrants who arrived as children (by age 15), and G2 – descendants of immigrant parents.

⁵ Low: none or ISCED1-2, i.e. lower secondary or below; Medium: ISCED 3, i.e. completed upper secondary;

-
- education of the mother (3 categories based on ISCED 2011 classification)
 - student status [student, not student]
 - religion [Christian, Unaffiliated, Muslim, Other]
 - language spoken at home [Official Language in Country of Residence, Other EU official language, Other languages]

2.1.1 Base population

Any microsimulation model requires high-quality, reliable, and accurate individual-level data for the base population in all its simulated characteristics. Such data usually come from a survey or census microdata and the simulation starts from these baseline population stocks. For QuantMig-mic we build a synthetic baseline that represents populations in 31 simulated countries by 12 characteristics. Since there is no single data source that would contain all these variables for all countries, we combine census 2011, EU-Labour Force Survey (LFS) and European Social Survey (ESS) data to derive the baseline population. We start with census 2011 population stocks which to date provide the most reliable information on population counts by age, sex, country of birth and educational attainment. The EU-LFS was used to derive the labour force characteristics, duration of stay and the age of arrival to the country of residence. Socio-cultural variables (religion, language spoken at home) are not included in the EU-LFS and were imputed using ESS.

The base population of the QuantMig-mic model is different from the one in the previous CEPAM which used EU-LFS microdata. This change was made due to deficiencies in the EU-Labour Force Survey data especially with respect to foreign-born populations, and reduced information in the EU-LFS for the EFTA countries which were added and for which LFS files do not provide person records younger than adult population. The synthetic base population was prepared with sampling rules that assure oversampling of actors from little represented population groups which is particularly important since the focus of the Quantmig project is on migration and the aim is to allow for analyses for specific subgroups of foreign born populations, for example by combination of demographic characteristics and region of birth.

The base population takes a form of a database in which each line contains a record of an individual (simulated actor) and each column gives a value of a variable in a given record. Each actor (a line in the database) represents a certain population with a given set of characteristics. Generating a fully synthetic baseline population requires a) counts of the real population by a set of characteristics of interest and b) case generation rules which set how big subpopulation an actor represents.

The case generation rules are implemented in the model as a parameter. This parameter allows to flexibly set how many actors will be generated in the synthetic baseline population per each subpopulation based on its size in the real population. The case generation rules are designed to overrepresent small population groups because the QuantMig model is intended to project migrant populations. This is because for some combinations of place of residence – place of origin – age – sex – educational attainment only a very small stock of individuals exists in the respective populations represented by the census 2011 data. These case generation rules are specified as follows:

High: ISCED 4 or higher, i.e. post-secondary non-university or tertiary (university) education.

2 actors for a subgroup of up to 30 individuals

10 actors for a subgroup counting between 30 and 100 individuals

30 actors for a subgroup counting between 100 and 1000 individuals

40 actors for a subgroup counting between 1000 and 10 000 individuals

0.05% of the subgroup size for subgroups counting more than 10 000 individuals

The composition of the base population by all 13 characteristics listed in the previous section can be viewed in the Modgen user interface. Detailed information about the exact data sources and methods used to derive the base population will be included in the upcoming D8.2 deliverable.

2.1.2 Modules

Table 1 provides an overview of the simulated events and their determining status variables. The model includes nine modules corresponding to the events that are being modelled. Availability of comparable data for all 31 countries sets limits to the interlinkages that can be captured and incorporated in the model. This is particularly true for the EU-LFS where different countries do not provide some variables or provide them in deferring category break downs, in which case the least granular coding determines the use. The work on adding place of birth as determining status for emigration rates is ongoing. It can be added in more or less detailed granularity of place of birth variable depending on the results of the migration modelling in WP6.

Table 1: Overview of the events modelled in QuantMig-mic and their determining status variables

<i>Events</i>	<i>Age</i>	<i>Sex</i>	<i>Place of residence</i>	<i>Place of birth</i>	<i>Age at immigration</i>	<i>Duration of stay (immigrants)</i>	<i>Language</i>	<i>Religion</i>	<i>Education</i>	<i>Mother' s education</i>
<i>Fertility</i>	X		X	X	X	X			X	
<i>Mortality</i>	X	X	X						X	
<i>Intra-European migration</i>	X	X	X	(x)						
<i>Emigration</i>	X	X	X	(x)						
<i>Educational attainment</i>	X	X	X	X			X	X		X
<i>Labour force participation</i>	X	X	X	X	X	X			X	
<i>Employment</i>	X	X	X	X	X	X			X	
<i>Language used at home</i>					X	X	X			
<i>Religion</i>	X	X	X					X		

Fertility module

Fertility event initiates a new actor inside the model. Fertility rates are derived from the SSP2

scenario of the Wittgenstein Human Capital projection⁶ which vary according to age, country of residence, level of education and projection year. These rates are further adjusted to incorporate differentials for native-born and foreign-born women. The age-, country-, and education specific rates are adjusted by contrasted parameters which take into account the place of birth and duration of stay and generation status of women. The version implemented in the QuantMig-mic is an update of the method explained in Potančoková and Marois (2020). In the updated version, the education variable used is the “highest level of education that will be reached” rather than the “current level of education and student status”. In other words, we introduce fertility differentials between native-born and different groups of immigrant women depending on their migrant characteristics on top of the previously mentioned dimensions. The rationale and the importance of capturing these differentials is discussed in Potančoková and Marois (2020) and the statistical modelling and actual parameters will be available in D8.2 deliverable. No fertility differentials based on other cultural characteristics are included due to lack of data.

Mortality module

Mortality rates are derived from the Wittgenstein Human Capital projection parameters⁷. They vary according to age, sex, country of residence, level of education and projection year. No mortality differentials based on ethno-cultural characteristics are included. The final version of the model will take into account covid-19 mortality impacts through modelling of reduction of life expectancy gains which is still work in progress in collaboration with IIASA colleagues.

Education module

The education module of QuantMig follows the same methodology as the one of CEPAM-Mic (Marois, Sabourin, and Bélanger 2019a). The highest level of education that an individual will reach over the life course is set probabilistically at birth (or at arrival for immigrants who arrived during childhood) using parameters estimated from an ordered logit regression. The model explicitly considers the influence of personal characteristics and the highest educational attainment of the mother. Sex- and country-specific cohort parameters are also included and extrapolated to establish assumptions for future cohorts.

Individuals are set as student starting from age 5 until the age of graduation from the highest completed level (but at most to age 29). The age at graduation is determined for all levels using Eurostat distributions by ISCED levels for the latest graduated cohorts. The modelling of education thus allows distinguishing, for each individual at each projection step, a) the highest level of educational attainment that will be reached during the life course, b) the current level of education, and c) whether or not the individual is still in schooling.

Parameters used in QuantMig-mic differs from CEPAM model following changes in the ordered logit regression model and will be detailed in the deliverable D8.2.

Labour force participation and employment modules

The labor force participation module follows a similar methodology as what developed for CEPAM-Mic (Marois, Sabourin, and Bélanger 2019b). In short, at the end of each year, the module determines

⁶ Available at: <http://dataexplorer.wittgensteincentre.org/wcde-v2/> For the description of the SSP2 scenario see Lutz et al.(2018).

⁷ Available at: <http://dataexplorer.wittgensteincentre.org/wcde-v2/> For the description of the SSP2 scenario see Lutz et al.(2018).

probabilistically whether or not the actor participates in the labour force, i.e. the actor is in state active or in state inactive. Labour force participation is determined purely by individual characteristics of the actor (age, sex, place of birth, etc., see Table 1). Parameters are estimated from sex- logit regressions on a binomial variable representing participation in the labour force. The logit regressions model differs from CEPAM-Mic in many ways and will be described in detail in the D8.2 deliverable.

Actors aged less than 15 or 75 or more are considered to be inactive.

The employment module is an extension of the labour force participation and will be described and documented in detail in the deliverable D8.2.

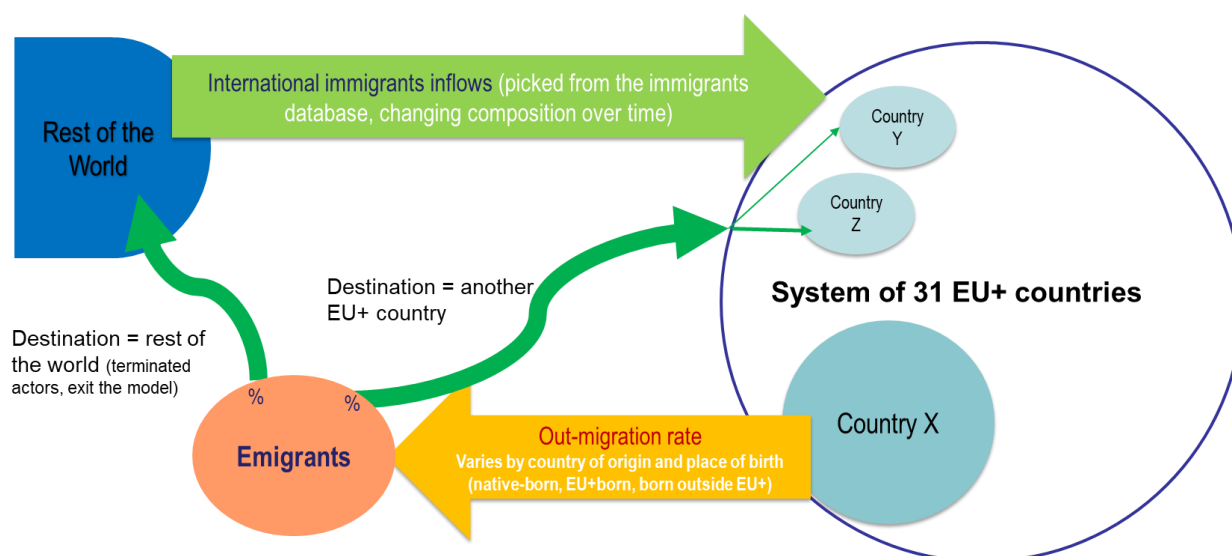
Migration module

First, a country-specific emigration rate is modelled, and a ratio is applied to this rate to determine what share of emigrants will move to another European country (thus stay in the model) and what share will move to another world region. Actors who migrate to another world region leave the model (are terminated). OD matrix determines a new destination country for the actors who move to another European country. Migration between the 31 European countries is modelled inside the model.

The current version of the model uses model migration schedules (Rogers and Castro 1981) and these will be replaced in the later version upon availability of age-specific migration flow estimates from the WP6. Sex-specific OD matrices can be added if available, the current version of the model uses a single OD matrix for both sexes.

Immigrants from other world regions are supplied to the model from an immigration database and enter the model as new actors. Immigrants in the model are given the same characteristics as recent immigrants in the base population (arrived in the country of residence five years ago or less, as measured in the LFS). The scheme below visualises the components and mechanics of the migration module.

Figure 1: Migration module visualisation



In the Modgen user interface the different components of the migration module are labeled as follows:

Mobility stores parameters for the emigration rates (“Exit rates”), the proportion for leaving

model (“Proportion leaving EU”) and the OD matrix (“OD matrix for intraEU mobility”). The immigrant flows are loaded from a separate file into “Immigration Pop” and “Immigration” stores the sampling rules for case generation.

Language and religion

These modules are identical to the ones in CEPAM model. At birth, religious affiliation and language spoken at home are taken directly from the mother, and are subsequently allowed to change during the life course. Transition rates for religious denomination were taken directly from the PEW projections on religion (Pew Research Centre 2015). Life course transition rates for language spoken at home were based on model schedules (Sabourin and Bélanger 2015) calibrated using data from the ESS. The socio-cultural variables are used to model educational transmission and to capture differentials in educational attainment on top of place of birth. We do not foresee additional work on these modules due to lack of data.

2.1.3 Model calibration

The current version of the model is not calibrated to Eurostat population stocks of 2020 and these stocks. Births are calibrated to total births reported by Eurostat 2011-2019 and mock migration flows based on preliminary version of D6.4 The calibration year and calibration parameters can be flexibly changed in the “Person core” in the user interface.

3. Deliverable description

3.1 Technical information for the user

The zipped file of the deliverable 8.1 includes folders and files with the version of the model as of 24th February 2022. For practical reasons and to allow the user to view the model in the Modgen user interface we provide not only the model code, as specified by the deliverable 8.1, but model including the input data (base population and immigration database), current version of model parameters and an illustrative “Default” scenario with preliminary results stored in the output database of the model. The code of the model can be found in “code” subfolder and can be viewed in Visual Studio also by opening “modgen-model.vcx. proj” file.

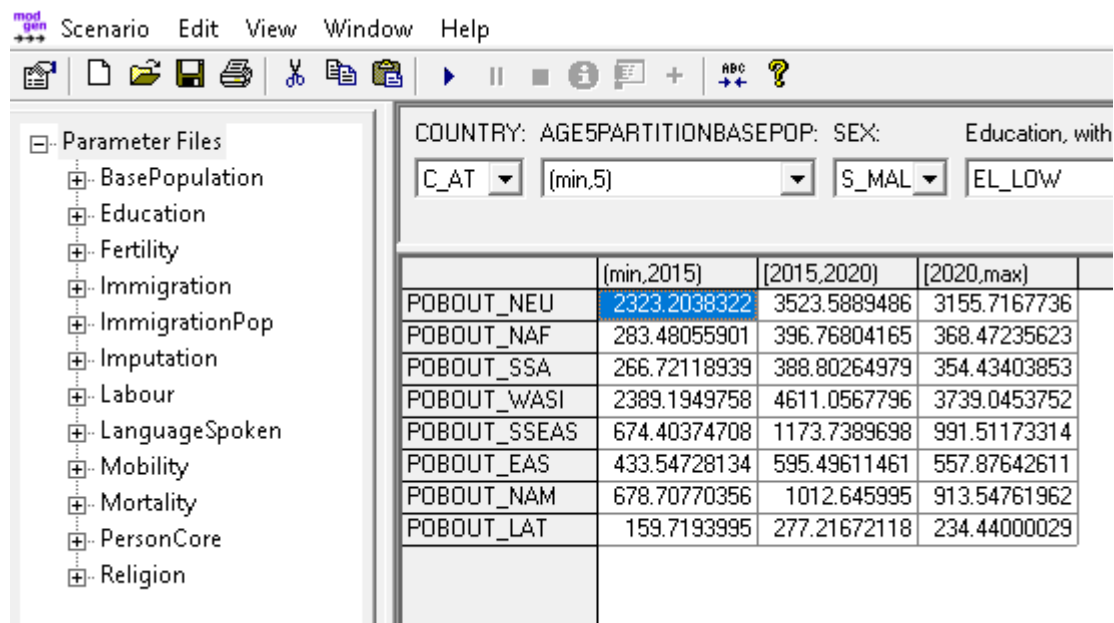
To use the QuantMig-mic, the user needs to first install Modgen Prerequisites 12. Installers for Modgen are not part of this deliverable and we refer the user to the Statistics Canada webpage: <https://www.statcan.gc.ca/en/microsimulation/modgen/download> Providing step by step tutorial on how to install Modgen and information on operating system requirements is beyond the scope of this cover document. It can be found at webpage: <https://www.statcan.gc.ca/en/microsimulation/modgen/new/introduction/introduction> In case of problems with the software installation the users are advised to contact Statistics Canada through a dedicated form at the Modgen webpage.

Modgen Prerequisites installs all of the software required to run a model developed using Modgen. It includes a .chm (compiled help) file corresponding to the Modgen Guide to the Visual Interface. This means that the user can view and modify parameters components and run a new simulation. The coding of states, events and timing and outcome determining function cannot be change.

The executable file of the model is located in Model/modgen/bin/ subfolder of the deliverable. Executable open the user interface in Modgen. To view the model the user needs to open the scenario using a menu or an icon in the top panel:



The Default scenario is stored in the separate subfolder Model/scenarios of the deliverable. Once the scenario is loaded the model components can be viewed:



To view the model results the user can code specific table outputs which are stored in the output database of the model. To facilitate viewing the results we have prepared predefined tables in the “Tables” in the user interface labeled “Demography” and “LFP” for labour force. The first table allows the user to view modelled population by country of birth (POB_x), age, sex, county of residence and educational attainment. The user can select population subgroups from the drop down menu and can also reshape the table to view different cross tables by changing table Properties.

Modgen version for model development is not necessary to use the model and view its components in the user interface but it is necessary to make any changes to the model code. The parameters can also be altered through Visual Studio user interface, however, this is more challenging as these have to be copied into the model code in specified structure, which may not be straightforward for an unacquainted user. Therefore, we recommend that the users who wish to test or change model parameters and other components install full Modgen and make changes through the user interface.

3.2 Modgen

QuantMig-mic is programmed in Modgen which is a generic microsimulation programming language specifically designed to create and maintain dynamic microsimulation models. Modgen was developed by the Statistics Canada (STATCAN) and the required software is freely available on the Statistics Canada website⁸. QuantMig-mic is programmed in Modgen 12.

⁸ More about MODGEN: <https://www.statcan.gc.ca/en/microsimulation/modgen/modgen>

Modgen is a meta-language of C++ and microsimulation models are coded and implemented in the Microsoft Visual Studio software suite. Visual Studio 2015 or 2017 is required and Visual Studio 2017 was used to code QuantMig-mic model (see Modgen System Configuration) ⁹. Once compiled, a Modgen model takes the form of a stand-alone executable file (.exe) that allows manipulation of the model through a graphical user interface. From this interface the user is able to modify simulation parameters and create customized scenarios.

For more details on Modgen, please visit Statistics Canada's website. For a full didactical presentation of Modgen, you may consult the book "Microsimulation and population dynamics: an introduction to Modgen 12" (Bélanger and Sabourin 2017).

3.3 Model variable codes in the user interface

Age

AGE5PARTITION (min, 5), (5,10), ..., (100+)

Sex

S_MAL = male

S_FEM = female

Country of residence (C_2ALPHAcode)

C_AT Austria

C_BE Belgium

C_BG Bulgaria

C_CH Switzerland

C_CY Cyprus

C_CZ Czech republic

C_DE Germany

C_DK Denmark

C_EE Estonia

C_ES Spain

C_FI Finland

C_FR France

C_GR Greece

C_HR Croatia

⁹ Read here about Modgen System Configuration of Visual Studio:

<https://www.statcan.gc.ca/en/microsimulation/modgen/system-configuration>

C_HU Hungary
C_IE Ireland
C_IS Iceland
C_IT Italy
C_LT Lithuania
C_LU Luxembourg
C_LV Latvia
C_MT Malta
C_NL Netherlands
C_NO Norway
C_PL Poland
C_PT Portugal
C_RO Romania
C_SE Sweden
C_SI Slovenia
C_SK Slovakia
C_UK United Kingdom

Region of birth

POBS_OWN == born in the country of residence
POBS_EU15 == born in EU15 country
POBS_NM13 == born in new members state (accession 2004-2013)
POBS_OEU == born in Other European country
POBS_NAF == born in North Africa
POBS_OAF == born in Other Africa
POBS_NME = born in Near Middle East
POBS_EAS == born in East Asia
POBS_SSEAS == born in South South-East Asia
POBS_NAM == born in North America, Austria and Oceania
POBS_LAT == born in Latin America

Labour force participation

[active, inactive]

Employment

[employed, unemployed]

Age at immigration

G_1 = generation 1, arrived at age 15+ (adult)

G_15 = generation 1.5, arrived at age 0-14 (child)

G_2P = generation 2 (descendants)

Duration of residence [0-4, ..., 20+ years]

D_NAT = born in the country of residence

D_2 == less than 5 years

D_7 == 5-9 years

D_12 == 10-14 years

D_17 == 15-19 years

D_22 == 20-24 years

D_25P == 25 or more years

Educational attainment

EL_LOW = low

EL_MED = medium

EL_HIGH = high

Student status

IS_STUD == is student

Religion

R_CHR == Christian

R_NON == Unaffiliated

R_MUS == Muslim

R_OTH == Other

References

Bélanger, A., and Sabourin, P., 2017. 'Microsimulation and Population Dynamics. An Introduction to Modgen 12.' Vol. 43 2017. The Springer Series on Demographic Methods and Population Analysis. Springer International Publishing.

- Belanger, A., Sabourin, P., Marois, G., Van Hook, J., & Vézina, S. (2019). A framework for the prospective analysis of ethno-cultural super-diversity. *Demographic Research* 41 293-330. 10.4054/DemRes.2019.41.11.
- Lutz, W., Goujon, A., Stonawski, M., Stilianakis, N. (eds.) (2018). Demographic and Human Capital Scenarios for the 21st Century: 2018 assessment for 201 countries. European Union. Doi: 10.2760/835878
- Marois, G., Sabourin, P., & Belanger, A. (2019a). Forecasting human capital of EU member countries accounting for sociocultural determinants. *Journal of Demographic Economics* 85 (3) 231-269. 10.1017/dem.2019.4.
- Marois, G., Sabourin, P., & Bélanger, A. (2019b). How reducing differentials in education and labor force participation could lessen workforce decline in the EU-28. *Demographic Research* 41 125-160. 10.4054/DemRes.2019.41.6.
- Pew Research Center (2015). "The Future of World Religions: Population Growth Projections, 2010-2050" <http://www.pewforum.org/2015/04/02/religious-projections-2010-2050/>
- Potančoková, M., Marois, G. (2020). "Projecting future births with fertility differentials reflecting women's educational and migrant characteristics." *Vienna Yearbook of Population Research* 2020. Vol. 18. Austrian Academy of Sciences, Vienna, Austria. DOI: 10.1553/populationyearbook2020.res02
- Rogers, A. (1986). Parameterized Multistate Population Dynamics and Projections. *Journal of American Statistical Association* 81(393): 48-61, DOI: 10.1080/01621459.1986.10478237
- Rogers, A., Castro, L.J. (1981). Model Migration Schedules. IIASA Research Report. IIASA, Laxenburg, Austria: RR-81-030
- Sabourin, P. and Bélanger, A. (2015). The dynamics of language shift in Canada. *Population* 70(4): 727-757. doi:10.3917/popu.1504.0771
- van Imhoff, E., Post, W. 1998. Microsimulation Methods for Population Projections. *Population: An English Selection* 10: 97-138.

