How to facilitate hybrid model development by rebuilding the demographic simulation model

Jacek Zabawa[1[0000-0003-4027-8253]]

[1] Department of Operations Research and Business Intelligence, Faculty of Management, Wroclaw University of Science and Technology,
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
Jacek.Zabawa@pwr.edu.pl

Abstract. Demographic information can be used to analyze processes occurring in a wide range of human social activity: in the area of management, healthcare, social security systems. One of the leading methods of demographic modeling is continuous simulation based on the system dynamics approach. Reducing computation time for fast-running continuous model implementations enables efficient development of hybrid models. In the hybrid simulation model, the discrete-event simulation approach was used as one of the components. Technical and conceptual solutions to the observed problems were presented, experiments based on demographic data from the Wrocław region, Poland were performed and it was indicated that an effective hybrid simulation will allow to include additional cause-and-effect relationships in the models.

Keywords: Operations research, Decision support, Improving the simulation performance.

1 Introduction

Knowledge of the state of demography, i.e. at least the number of people who can be classified into cohorts (groups distinguished on the basis of the age and sex) and trends in phenomena affecting these quantities, is useful for studying issues such as: prediction of premature mortality [2], support needs in certain age groups in the coming decades [8], demand for housing [9], economic reform scenario analysis [3], consumption levels study [7], long-term care expenditure analysis [10], methods of improving neonatal care [13], long-term care capacity analysis [1], changes in the pension system [15].

Among the approaches used to model demographic dependencies, simulation in accordance with the system dynamics (SD) approach (continuous simulation) is widely used. In addition to dividing the population into cohorts, this approach distinguishes main elements as input flows: births, immigrations and ageing/growing up and output flows: deaths, ageing/growing up, emigrations. Additional elements are: birth coefficients, death coefficients, transition time between cohorts. Studies using the SD approach are usually performed in accordance with the dynamic synthesis methodology

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

presented e.g. in [13], i.e. creating a descriptive model (causal loop diagrams, CLD) and then building a computer system dynamics model (SDM), composed of flows (rates) and stocks (levels). CLD (a conceptual model) and SDM (a computational model) form the modeling and simulation cycle [4].

2 Conclusions from previous studies. Discovered issues

This paper is one of a series of works on simulation modeling applications, i.e. the study of demographic changes and the use of the discrete-event system (DES) modeling approach to modeling the demand for hospital services [11], the prospects of replacing the SD approach with discrete rate modeling (DRM) in demographic research and how cohort grouping affects these prospects [16], studying the impact of demographic scenarios on the intensity of the demand for medical services [12].
The usefulness of our models will be determined by measuring the differences between historical and simulated cohorts. In the models built so far, it was possible to prevent the problems noticed in [6], by assigning an increase in cohorts in the model to the number of cohorts equal to the maximum age of people in the model, for each gender (2*105). However, we have observed problems in the use of models built on the Extendsim platform [5], such as: e.g. assigning initial states in stock blocks (Holding Tank) and slowing down the performance of models as the number of cohorts increases. The simulation run lasted too long (about 1 hour) for our needs. We have decided to apply a DES to pass values between streams and resources. All of the above problems result in the fact that the discrete - continuous hybrid seemed difficult to implement.
Population data from 2006 were obtained from Statistics Poland [14] and concern the Wrocław region. The data used relate to the cohorts of human populations mostly 5 years wide (i.e. 17 cohorts from 0_5 to 80-84 years old) and one cohort supposedly infinite width (the model included a 20-year cohort for persons 85-104 years old).
We built separate models: SD and DES. They must therefore be run separately. DES has to wait for simulation results from SD model. It became necessary to construct an additional structure (interface between models). Purpose of SD is to save the state of the cohorts in a output file (e.g. a spreadsheet or a database file) at successive points in time that are important from the point of view of calculations.
The main task of the SD module in hybrid is to generate the status (values) of individual cohorts. The state of each cohort affects the parameters that control DES module, i.e. the flux of objects representing events, the number of which depends on the number (expressed in persons) of each cohort. An example would be the generation of objects (representing patients) with characteristics that may indicate particular types of diseases. It is assumed that these conditions may have a frequency depending on cohort.

Below is illustration of structure of hybrid simulation:

- Input: Description of demographic processes
- SD model
- Output of SD – cohort predictions
- Input: Description of the relationship between cohorts and studied phenomenon

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

- DES model
- Output of DES – objects (individuals with properties)
- Results analysis – data resource usage, finance, well-being etc.

3 Modeling approach

The paper presents a description of the key places in the model built in the Extendsim simulation environment [5], where model is grafically created by placing blocks from various libraries into the working area, linking them using connectors, adjusting parameters and, to a lesser extent, creating programs in the ModL (similar to C) language. It seems that the proposed solutions to the outlined problems will also be useful when using tools with similar architecture and methods of operation (using the push-pull approach). In the author's view, Extendsim is distinguished by the intuitive structure of the models, however, obtaining results requires additional efforts of the modeller.
The following suggestions for model implementation were made:

- Supplement the simulation run with an additional year (step) preceding the initial year to assign baselines to all cohorts based on the values stored in the spreadsheet.
- Use 'Go' connectors (triggering calculations on demand) and careful select and arrange signals triggering calculations (inputs and outputs) - in Equation blocks.
- Use Equation blocks rather than elementary Math blocks (such as Add, Subtract, Multiply, Divide) when it does not reduce the readability of the models.
- Use of internal databases (increase the calculation speed) for intermediate results.
- Store initial data and state variables in the course of the simulation (dynamic birth rates) and results in spreadsheets (preferably a single .XLSX file), in separate sheets but with permanently assigned sheet columns for each cohort.
- During the test experiments, it seems useful to store the initial states of individual cohorts and the simulations results concerning the state of cohorts in one worksheet, in distant groups of rows (first the initial states of the cohorts, then the space for the simulation results).

4 Technical solutions

For the sake of simplicity, we consider a simulation model for female cohorts. The male part of the model differs only in that male cohorts do not affect the birth stream and have their coefficients (except duration of residence in the cohort).
First, the initial values are loaded from the spreadsheet cells (see Fig.1) selected by additional control module (Fig. 2).
The situation will be similar for saving the results, but the space will be shifted by several dozen (e.g. 20) rows down (the offset will depend on the size of the space reserved for input data when testing the model). The task of the lower part of the submodel is to determine which simulation step number we are dealing with; The block on the left generates auxiliary objects at fixed intervals (1 year). The first auxiliary object

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

will be generated as soon as the simulation is run. The variable signal2 changes its value from 0 to 1 only when the second helper object arrives – this fixes the row representing the initial data "set back by 1" and correctly sets the initial data in the Holding Tank (HT) blocks (representing the term "level" used in the continuous approach). The value of the signal variable is equal to 1, of course, after the arrival of the second object in the control stream, but it is calculated and used as a double "message" running calculations in "cohort" blocks for the first year actually processed (the point is that the streams incoming and outgoing from the blocks implementing the "level" concept, i.e. the HT, are updated after assigning initial values to the HT blocks (via the Init connector) and update (resulted in) also saving the new value of the HT. Without the above described action, the result of the calculation would cause the state of the HT block to change only after the arrival of a new control object, despite the fact that the streams associated with the HT are already updated (by the first message).
Secondly, the number of live births in a given step of the simulation should be calculated (see Fig. 3). The birth rate stream is directly affected by the increase in the number of people (in this case, women) aged 0 (belonging to the F0_4 cohort).

	A	C	D	E	F	G	H
2		F_0_4	F_5_9	F_10_14	F_15_19	F_20_24	F_25_29
3	Year	1	2	3	4	5	6
5	2006	24028	25019	29648	37431	51220	53727

Population | Ageing | BirthsCoef | Migrations | DeathsCoef | Births

Fig. 1. Actual data on the initial population – excerpt from worksheet

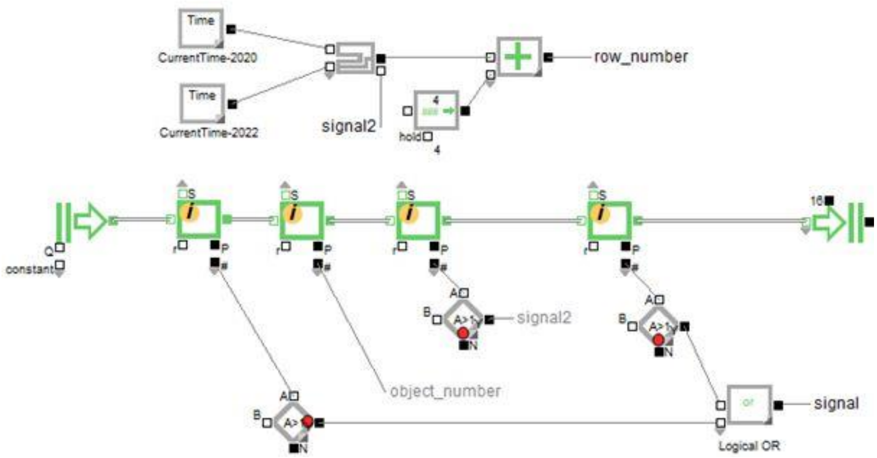

Fig. 2. Continuous model control submodule. It ensures that the initial values of the stock blocks (Holding Tank, outside the figure) and the inflow and outflow streams are correctly assigned

The non-standard settings of the Equation block are marked with red lines (see Fig. 3) on the input and output connectors. The markings denote that the calculations are

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

performed only when a signal is sent to the block about a new, upcoming object in the object stream (a submodule built according to the DES approach, but this does not apply to the first object generated at the start of the simulation). The model contains a lot of other submodules, e.g. for the total inflow (birth + growing up + immigration).

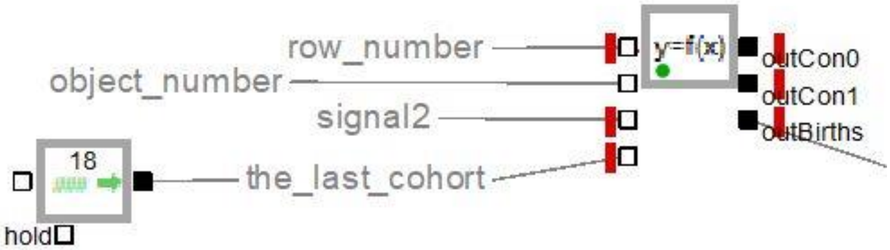

Fig. 3. Sub-module that calculates the birth rate in the first cohort (it uses data on the size of female cohorts and fertility rates (newborns, e.g. females))

5 Experiments

We want to test whether the model results are close to the real data. Therefore, deviations (differences) between the above values in individual cohorts will also be observed. From several tests, the long simulation time - 10 years was chosen to present (see Fig. 4). The actual data is the same as in the papers [11, 12, 16, 17]. The data refers to the female population of the Wroclaw region (Poland), begins at 2006, was divided into cohorts 5 years wide (from cohort 0_4 to 80_84). The model is fed by size of individual cohorts, birth rates (for individual female cohorts), death rates, migration rates.

The dependence of the average percentage deviations on the number of years of simulation was determined (Fig. 4). A monotonic increase in deviations is visible. The simulation run (15 years) takes approx. 12 seconds (Intel Core i7-10700). Shapes of graphs of real cohort size distribution and model results (simulation) are similar (Fig. 5).

Years of simulation	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Average of % |Error|	3,3	3,8	5,2	7,1	7,9	8,0	8,2	8,9	10,1	10,8	11,2	12,8	15,1	18,5

Fig. 4. The dependence of the average percentage deviation on the number of simulation length

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

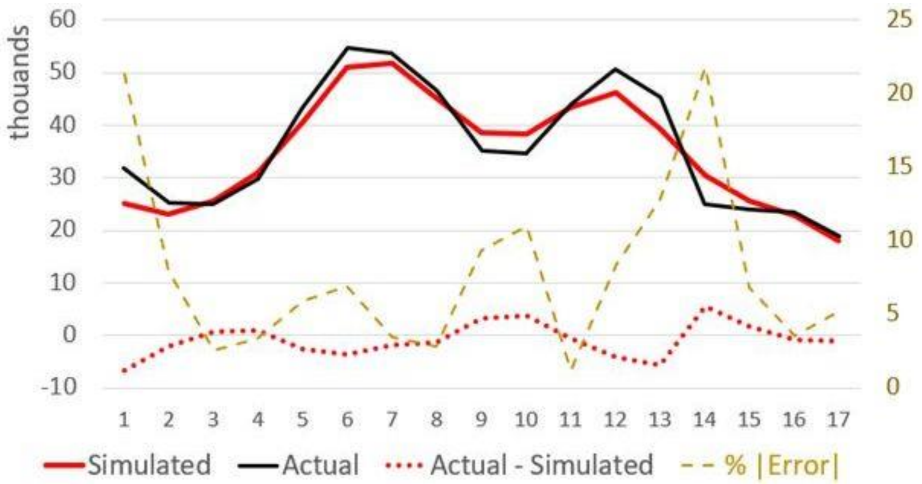

Fig. 5. Year 2016, the 10th year of simulation. Actual data, simulation results and deviations between them. The highest percentage deviations are in cohort no. 15 (F70_74), the deviations are irregular - an additional measure should be proposed

6 Hybrid simulation model. Discussion of the results

After providing fast calculations for continuous simulation, we will try to attach a DES module. Two cohorts were selected for the study: F_30_34 and F_35_36. The actual cohorts size and number of corresponding objects should be similar. However, the DES module is driven by the values of the current continuous simulation results, not by the actual size of the cohorts. It is noteworthy that the results of the simulation were more and more distant from the actual values (see Fig. 6).

The 15-year simulation lasted about 1 minute and 10 seconds when the DES module was tested for the two cohorts mentioned above. Taking into account the previously given information about the length of the simulation time of the continuous module, we can conclude that obtaining results from the discrete part of the model requires much more computational time. We know that the time needed for calculations in a continuous module does not depend on the size of the studied population, but only on the complexity of the model, i.e. the number of cohorts. In DES models, we will want to observe every object (person), even if the attributes assigned to it have stochastic values. We can assume that the larger the nominal numbers of people belonging to each cohort, the longer the time will be required to perform the calculations. As we can see in the example under consideration, even with a population of 40,000 people, the time of calculations in a continuous model is negligible in relation to the time needed for calculations in DES module, and this is the situation we wanted to achieve. In the models we used previously (210 cohorts), the calculation time was approx. 40 minutes, which made it particularly difficult to perform experiments for different variants – scenarios concerning, in particular, the values of birth and death rates and migration. We estimate that it will now be possible to include modules designed to model the above indicators so that

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

they are a function of the current state of the cohorts or a function of their history. It should be noted that graphs (see Fig. 6) roughly reflect the dynamics of the actual states of the cohorts. However, as the simulation time horizon increases, the waveforms of the simulation results move away from the trajectory of the actual values and these are negative deviations (i.e., the simulation generates too few objects). These conclusions coincide with the observations on the results of the continuous model. We presume that they are due to the long, 5 years time span represented by the elementary cohort. Therefore, heterogeneities in the age distribution of individuals in a given elementary cohort are not taken into account.

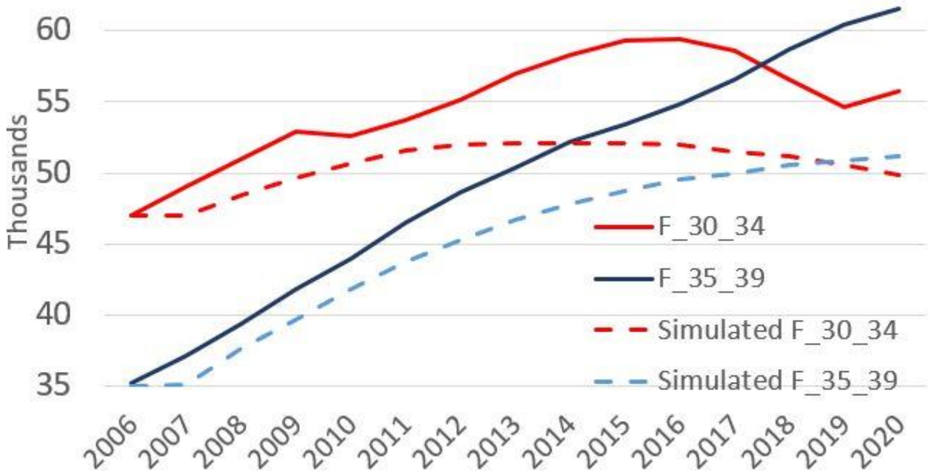

Fig. 6. A comparative chart of the actual size of cohorts and objects generated in subsequent years in DES module. Two female cohorts in the 4th decade of life. Negative deviations of the simulation results from the actual data are visible

7 Conclusions

Long calculation times and difficulties with the initial values assigning were observed in the past. These implementation shortcomings made it difficult to create effective hybrids of a continuous-discrete-event system. The architecture of the continuous model has been redesigned. Construction solutions were presented to overcome the current shortcomings of the implementation. A very high acceleration of calculations and a reduction in the simulation time were achieved. This enabled the implementation of an efficient and direct continuous-discrete connection (hybrid). The consistency of the state trajectories of the selected cohorts was also tested in terms of actual data and simulation results. The observed shortcomings prompt a modification of the model consisting in increasing the number of aggregated (composite) cohorts and shortening the time interval per elementary cohort. It corresponds to one of the recommendations [7].

Acknowledgments. ExtendSim blocks copyright © 2024 ANDRITZ Inc.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

8 References

1. Ansah, J.P. et al.: Implications of long-term care capacity response policies for an aging population: A simulation analysis. Health Policy (New. York). 116, 1, 105–113 (2014). https://doi.org///dx.doi.org/10.1016/j.healthpol.2014.01.006.
2. Best A., et al.: Premature mortality projections in the USA through 2030: a modelling study. Lancet Publ. Heal. 2018, 3(8): E374–384. https://doi.org/10.1016/S2468-2667(18)30114-2
3. Colacelli, M., Corugedo, E.F.: Macroeconomic Effects of Japan's Demographics: Can Structural Reforms Reverse Them? (2018). IMF Working Paper 18(248):1
4. Crielaard, L., et al.: Refining the Causal Loop Diagram: A Tutorial for Maximizing the Contribution of Domain Expertise in Computational System. Dynamics Modeling. Psychological Methods, 2024, Vol. 29, No. 1, 169–201, https://doi.org/10.1037/met0000484
5. Diamond B., Krahl D., Nastasi A., Tag P.: ExtendSim advanced techology: Integrated simulation database, Proceedings of the 2010 Winter Simulation Conference, Baltimore, MD, USA, 2010, pp. 100-100. 32-39, doi: 10.1109/WSC.2010.5679178
6. Eberlein, RL, Thompson, JP.: Precise modeling of aging populations. Syst Dynam Rev 2013; 29(2):87–101. https://doi.org/10.1002/sdr.1497
7. Giesecke, J., Meagher, G.: Population ageing and structural adjustment. Aust J Lab Econ 2008; 11(3):227–47. DOI: https://doi.org/10.26686/lew.v0i0.1663
8. Kingston, A. et al.: Forecasting the care needs of the older population in England over the next 20 years: estimates from the Population Ageing and Care Simulation (PACSim) modelling study. Lancet Public Heal. 2018, 3(9): E447–E455. https://doi.org/10.1016/S2468-2667(18)30118-X
9. Lauf, S., et al.: Simulating demography and housing demand in an urban region under scenarios of growth and shrinkage. Environment and Planning B: Planning and Design 2012; 39(2): 229–246. https://doi.org/10.1068/b36046t
10. McGrattan, E. et al.: On Financing Retirement, Health, and Long-term Care in Japan. IMF Work. Reuters. 18, 249, 1–43 (2018). https://doi.org/10.5089/9781484384718.001
11. Mielczarek B., Zabawa J.: Modelling demographic changes using simulation: Supportive analyses for socioeconomic studies, Socio-Economic Planning Sciences, Volume 74, 2021, 100938, https://doi.org/10.1016/j.seps.2020.100938
12. Mielczarek, B. et al.: The impact of demographic trends on future hospital demand based on a hybrid simulation model. In: Proceedings of the 2018 Winter Simulation Conference, pp. 1476-1487. https://doi.org/10.1109/WSC.2018.8632317
13. Semwanga, A.R., Nakubulwa, S., Adam, T.: Applying a system dynamics modelling approach to explore policy options for improving neonatal health in Uganda. Health Res Policy Sys 14, 35 (2016). https://doi.org/10.1186/s12961-016-0101-8
14. Statistics Poland Homepage. Główny Urząd Statystyczny. https://stat.gov.pl/en/
15. van Sonsbeek, J.: Micro simulations on the effects of ageing-related policy measures. Econ Modell 2010; 27:968–979 https://doi.org/10.1016/j.econmod.2010.05.004
16. Zabawa, J., Mielczarek, B.: An Attempt to Replace System Dynamics with Discrete Rate Modeling in Demographic Simulations. In: Pashinsky, M., et al. (eds) Computational Science – ICCS 2021. Lecture Notes in Computer Science, vol. 12745. Springer, Cham. https://doi.org/10.1007/978-3-030-77970-2_21

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_21