Keywords: Demographic Application, DEVS, Microsimulation

Abstract

Microsimulation is increasingly applied in demography to project the development of populations. A stochastic model is being introduced that describes individual life courses on a continuous time base. Life courses are determined by sequences of demographic events. We show how this demographic multi-state projection model can be specified as an atomic model in DEVS. Thereby, strengths but also limitations of this approach are revealed. The limitations are addressed by exploiting a DEVS variant that supports variable structures, i.e., DYNPDEVS, and by modeling individuals as atomic models and the population as a network model. An example projection of a synthetical population based on the population of Italy shows the plausibility and the feasibility of the developed model.

1. INTRODUCTION

Currently multi-state projection models, in which people move between demographically relevant states, substitute increasingly demographic standard projection models that forecast population changes only according to people’s age and sex. One of these models is the multi-state cohort component model which produces projections of subpopulations sharing certain characteristics like year of birth, sex, race, marital status [12, 23]. However, multi-state models do commonly not regard individual, demographic behavior and intra-cohort variation of individuals [10, 26].

Microsimulation is a methodology that accounts for the latter. It complements demographic projections by population subgroups with projections of how people live their lives. The central unit of demographic microsimulation is the individual life course which is characterized by a sequence of demographic events, such as birth, marriage, childbirth, divorce, retirement and finally death [10, 32]. An individual life course is appropriately modelled by a stochastic process [2, 16]. The propensity for experiencing certain demographic events is usually age and calendar time dependent (e.g., decreasing fertility rates and increasing life expectancy in the Western World). Consequently, a demographic microsimulation is a dynamic model, with age and calendar time as essential time scales [4, 8]. These time scales can either be discrete (usually in units of years or months) or continuous [6]. [6, 21, 28] discuss in detail advantages and disadvantages of the discrete and the continuous treatment of time. A main advantage of continuous-time models is that associated simulation algorithms are found to be more efficient than algorithms to run discrete-time microsimulations [21]. Although they feature several beneficial properties compared to discrete-time models, only a limited number of continuous-time microsimulation models exist [36]. E.g., they are successfully applied in various areas, like for an analysis of the retirement income implications of US government policies [9], and projections of the Canadian population regarding ethnocultural diversity [15].

In the context of the MicMac project (www.micmac-projections.org) we have developed a software tool that comprises both a continuous-time microsimulation (Mic) and a macro projection (Mac) routine [35]. The aim of the project was to offer a bridge between aggregate projections of cohorts (Mac) and projections of the life courses of individual cohort members (Mic). [18]. The Mac part has been implemented using the cohort component model approach [26, 29, 35]. A discrete event specification has been employed to set up the Mic part.

2. DEMOGRAPHIC MICROSIMULATION MODEL

The MicMac microsimulation model is based on a continuous-time multi-state model [1, 27]. Its state space $\Psi$ is a set of states a person can occupy. As an individual’s state usually is a combined characteristic, given by the combination of several attributes, we define so called state variables. All the unique combinations of values of these state variables determine the state space. To give an example, we consider the state variables sex and marital status. Possible values of sex are ‘female’ and ‘male’. The marital status of a person can be ‘never married’, ‘married’, ‘divorced’, ‘widowed’. A state is given by a combination of values, one for each state
variable (e.g. female, married or male, widowed).

The state space is determined by the problem to be studied and contains the relevant demographic states that need to be considered. An obligatory part of the state space is death which is a risk to which each individual is always exposed. Any demographic event, like getting divorced or childbirth, changes an individual’s state. For instance, an unmarried man experiences a demographic event by marrying his common-law spouse.

Three different options exist for an individual to enter the population: (1) being member of the initial population, (2) birth, or (3) immigration (if the model includes immigration into the population). The initial population contains the distribution of population members according to the states that they occupy at simulation starting time. Individuals leave the population either by death or by emigration (if emigration is allowed).

We formulate each individual life-course as a trajectory of a semi-Markov process \( Z(t) \), \( t \geq 0 \). A counting process \( N(t) \) corresponds to \( Z(t) \), that gives the total number of state transitions an individual has experienced until time \( t \). The semi-Markov process \( Z(t) \) can alternatively be written as \( \lambda N(t) \) where \((J_n)_{n \in \mathbb{N}_0}\) is a non-homogeneous Markov Chain that maps all states that an individual occupies. In accordance, \((T_n)_{n \in \mathbb{N}_0}\) reports the sequence of the corresponding transition times along the model time scale \( t \geq 0 \). Hence, \( Z(t) \) is fully defined by the two-dimension process \((J_n, T_n)_{n \in \mathbb{N}_0}\). The waiting time that an individual remains in state \( J_n \) is then determined by \( W_n = T_{n+1} - T_n \).

The propensity of an individual to change his/her current state \( s_i \) at age \( a \) and calendar time \( x \) is given by the transition rate (or hazard rate) \( \lambda_{s_i, s_j}(\tau | x, a) \), \( s_i, s_j \in \Psi \), associated with process \( Z(t) \):

\[
\lim_{h \to 0} \frac{1}{h} \mathbb{P}[J_{n+1} = s_j, W_n \in (\tau, \tau + h) | J_n = s_i, C(T_n) = x, A(T_n) = a, W_0 > \tau] \]

Here we include \( \Lambda(T_n) \) and \( C(T_n) \) to express the age- and time-dependence of the process. The function \( \Lambda(T_n) \) maps individuals’ (biological) age at \( T_n \) and the function \( C(T_n) \) the calendar time at \( T_n \).

If the transition rates of the process are known, the distribution functions of the waiting times in the distinct states of the state space can be derived. We denote by \( F(w_{s_i, s_j} | a, x) \) the distribution function of the waiting time in state \( s_i \) until moving on to state \( s_j \), conditioned on age \( a \) and calendar time \( x \). Employing the corresponding set of transition rates, \( F(w_{s_i, s_j} | a, x) \) can be expressed:

\[
F(w_{s_i, s_j} | a, x) = P(W_{s_i, s_j} \leq w_{s_i, s_j} | a, x) = 1 - \exp\{-\Lambda_{s_i, s_j}(w_{s_i, s_j} | a, x)\},
\]

where \( s_i, s_j \in \Psi, s_i \neq s_j \) and

\[
\Lambda_{s_i, s_j}(w_{s_i, s_j} | a, x) = \int_0^{w_{s_i, s_j}} \lambda_{s_i, s_j}(v | a, x) \, dv.
\]

is the so called integrated (or cumulated) hazard.

In a continuous-time multi-state model each individual life course is specified by a sequence of state transitions and the time spans between these transitions. Alternatively, the corresponding sequence of waiting times between state transitions can be used to determine an individual life course. Following this approach, in order to construct synthetic life courses, we generate sequences of waiting times.

If an individual enters at calendar time \( x \) and age \( a \) a state \( s_i \) that is not the death or the emigration state, then the individual will eventually experience a further transition in the future. The waiting time in state \( s_i \) until moving to another state \( s_j \) can be simulated by the well known inversion theorem [19]. If we denote by \( u \) a standard uniformly distributed random number, we yield a random waiting time \( w_{s_i, s_j} \) from the correct distribution by

\[
w_{s_i, s_j} = \Lambda^{-1}(w_{s_i, s_j} | a, x) \left[ -\ln(1 - u) \right]. \tag{1}
\]

According to a competing risk setting [13], we compute for all possible destination states waiting times and pick the shortest one \( w_{s_i, s_j} \) to state \( s_j \). Consequently the individual under consideration experiences his/her next state transition to \( s_j \) at age \( a + w_{s_i, s_j} \) and calendar time \( x + w_{s_i, s_j} \). This computation of the ‘shortest’ waiting times is repeated for each individual of the model population until the simulation stop time has been reached or all individuals have left the population (due to death or emigration). A more detailed description of Mic’s stochastic model is given in [36].

In equation (1) the inversion of the integrated hazard \( \Lambda(w_{s_i, s_j} | a, t) \) is required. For some basic distributions the inverse of the integrated hazard can be computed analytically, for others this can only be done numerically. The application of piecewise linear cumulative hazards is a suitable approximation to \( \Lambda^{-1}(w_{s_i, s_j} | a, x) \) is [7].

In order to forecast population numbers using the described approach, the user has to provide an initial population and transition rates for all feasible transitions that individuals may experience during their lives. Please note, that demographic forecasts are only performed along a restricted calendar time range \( T = [x_0, x_e] \), e.g., starting from now up to July 5th, 2012. In Figure 1 a part of a synthetic example life course is depicted.

3. DEVS MODEL FOR MICROSIMULATION

The MicMac microsimulation model specifies individual life courses as sequences of discrete events that occur along a continuous time line. For implementation purposes, we deemed it therefore best to formulate the microsimulation as a DEVS model [33]. The DEVS formalism offers all the functionalities that are required to formulate population dynamics in the described way (cp. section 2.). It is proven, tested, and widely established. A remarkable benefit of DEVS is
Figure 1: By being part of the initial population, at January 1, 2000, a female enters the virtual population. At this time she is $a_0 = 20.9$ years old, never married and childless. After a waiting time of 3.6 years at the age of 24.5 she experiences marriage event. When she is 30.5 years old, at July 6, 2009, a birth event is simulated and she becomes mother.

the existence of reliable verification and validation methods [14, 24, 34]. As several DEVS based tools exist (e.g., DEVSJava, JDEVS, CD++, DEVS variants in JAMES II), implementing a DEVS model does not pose serious problems.

To expand the classes of problems that can be represented, various DEVS variants have been developed [25, 33]. For a demographic microsimulation we employ two different DEVS model variants: first, we use a classical atomic DEVS model to formulate population dynamics over time. As we face here some difficulties capturing time-varying population sizes, we have additionally developed a DYNDEVS network that allows a variable number of individual model components.

3.1. Atomic Population Model

We formulate the MicMac microsimulation model as an atomic DEVS model $\text{Pop}^+$. Its state is defined by the attributes of individual population members. Population dynamics are caused by state transitions and the arrival of immigrants. The model time $t$ to which we refer in the subsequent, maps the calendar time horizon $T$ of our population projection. $\text{Pop}^+$ is given by the structure:

$$\text{Pop}^+ = \{ X_{IM}, Y_{EM}, \Phi, \phi_0, \delta_{int}, \delta_{ext}, \lambda, ta \}$$

where

$$X_{IM} = \begin{cases} \{ \Psi_0, \ldots, \Psi_{im} \} & \text{if im immigrants enter} \\ \emptyset & \text{otherwise} \end{cases}$$

$$\Psi_i (i = 1, \ldots, im) \text{ comprises the state } s_0 \text{ and the age of an immigrant when he/she enters the population, } (s_0 \text{ is the initial state of that process } Z(t) \text{ that refers to an immigrant’s life course})$$

$$Y_{EM} \text{ is a output port for emigrants to leave the population}$$

$$\Phi = \times_{k=1}^n \Psi_k \times \mathbb{R}_0^+ \text{ is the state space of } \text{Pop}^+, \text{ where }$$

$n$ is the number of all individuals that ever belong to $\text{Pop}^+$ along model time $t$

$I = \{ I_1, \ldots, I_k \}$ is the set of all individuals that $\text{Pop}^+$ comprises along model time $t$

$$\Psi_k = \begin{cases} \Psi \times \Psi \times \mathbb{R}_0^+ & \text{if } I_k \text{ is member of population at } t \\ 0 & \text{otherwise} \end{cases}$$

$\Psi$ is the state space of semi-Markov process $Z(t)$ that maps the individual life course (cp. Section 2.)

$$\psi_k = [s_j, s_c, a]$$

$\Psi_k \in \Psi_k$ describes the last state $s_j$, the current state $s_c$ ($s_j, s_c \in \Psi$) and the age $a$ of individual $I_k$ (current state of that process $Z(t)$ that refers to $I_k$’s life course) or indicates its non-existence: $\psi_k = 0 \text{ (} k = 1, \ldots, n \text{)}$

The model time $t$ is the state of $\text{Pop}^+$, where

$$\phi = ([\psi_1, \ldots, \psi_n], \tau) \text{ is the state of } \text{Pop}^+, \text{ where }$$

$$\phi = ([\psi_1, \ldots, \psi_n], 0 \text{ (if } n_i \text{ belongs to population)}} \text{ and } n_i \text{ is the number of individuals that belong or have ever belonged to population until model time } t$$

$\tau$ indicates the time to the next event

$\phi_0$ initial state of $\text{Pop}^+$ (is determined by initial population of microsimulation)

$I_{k*}$ is the individual with the shortest waiting time$^1$:

$$I_{k*} = \arg \min_{l \in I | \psi_l \neq 0} \{ \min_{j \in \Psi} \{ (w_{k, s_j}, s_c) \} \}$$

$s_{j*}$ is the next state of $I_{k*}$

$\tau = \min (w_{k, s_j}) \text{, is the related waiting time of } I_{k*}$

$w_{k, s_j} \text{ (random) waiting time between the states } s_c \text{ and } s_j \text{, defined by equation (1)}$

$$\psi_{k*} = (s_c, s_j, a + \tau) \text{ gives } I_{k*}’s \text{ next state and the corresponding age at transition}$$

$\text{bde : } \Psi \times \Psi \rightarrow \{ \emptyset, n_i \text{ newborns}, death, emigration} \}$ identifies birth, death, and emigration events (birth events imply $n_i \in \mathbb{N}$ newborns)

$$\delta_{int}(\phi) = \begin{cases} ([\psi_1, \ldots, \psi_{k*}, \ldots, \psi_n, 0, (n-n_i)], \tau), & \text{if } \text{ bde}(s_c, s_{j*}) = \emptyset \\ ([\psi_1, \ldots, \psi_{k*}, 0, \psi_{k*+1}, \ldots, \psi_n, 0, (n-n_i)], \tau), & \text{if } \text{ bde}(s_c, s_{j*}) \in \{ \text{death, emigration}}} \\ ([\psi_1, \ldots, \psi_{k*}, \psi_n, 0, \psi_{n_i}, 0, (n-n_i-n_i)], \tau), & \text{if } \text{ bde}(s_c, s_{j*}) = \{n_i \text{ newborns}\} \\ \text{not defined, otherwise.} \end{cases}$$

$^1 \arg \min_{a \in A \phi} f(a)$ gives the argument (element of $A$) that minimizes the value of function $f(\cdot)$
Immigrants will be generated by DEVS generator model \( \text{Gen} \), to which the population model is coupled. The generation of these events relies on empirical data about immigration dates and number of immigrants. The model’s outputs serve as an input to our population model. Please note that several immigrants may enter the population simultaneously, e.g., family members or couples.

Attributes and behavior of individuals are captured by \( \text{Pop}^+ \)’s state \( \phi \). Each of its components refers to an individual who was, already is, or will at some time be part of the model population. A zero value is assigned to individuals that are not yet born or immigrated. The same applies to individuals who have already left the population (due to death or emigration) at time \( t \).

Changes in individual attributes (provoked by state transitions of the corresponding semi-Markov processes) result in an internal event of \( \text{Pop}^+ \). Here the rule applies that the next event is triggered by the earliest state change of an individual. Relying on the algorithm given in Section 2, waiting times are computed for all individuals that are part of the model population. Subsequently, the individual with the shortest waiting time is picked and the associated transition is performed. Let \( I_{k^*} \) denote the individual with the shortest waiting time. If \( I_{k^*} \)’s state change does not imply that he/ she will leave the population or that (new) individuals will enter it (because \( I_{k^*} \) experienced a birth event), just the affected component of \( \phi \) will be replaced by \( I_{k^*} \)’s new state and the respective age at transition.

However, if \( I_{k^*} \) dies or emigrates, he/ she will not any longer be considered in the model state. The related component of \( \phi \) is set to zero. If a (female) individual experiences a birth event, we add as many children to the population as the woman gives birth to. The model state is accordingly updated by assigning initial attributes to the components that refer to the newborns. The arrival of immigrants causes an external event which leads to a similar state update.

As individual state changes occur along a continuous time line, emigration events occur one at a time. In addition these emigration events refer to solely one individual in the current model. However, it is far from being realistic, e.g., that a seven-year-old boy emigrates without his mother. To let families emigrate linkages between individuals have to be introduced.

In Figure 2 the basic setup of the \( \text{Pop}^+ \) model and the related experimental frame are illustrated.

- Atomic Model: \( \{X_{\text{IM}}, Y_{\text{EM}}, \Phi, \cdots\} \)
- Experimental frame: Gen \( \xrightarrow{X_{\text{IM}}} \) Multi-State Model \( \xrightarrow{Y_{\text{EM}}} \) Emigrants

Figure 2: The experimental frame holds the conditions under which the model population is observed, i.e., the parameterization of the model. Moreover, it covers the objectives that motivate its modelling and simulation; in our case, the simulation of life-courses of a representative sample of an entire population over a restricted time range. The atomic model \( \text{Pop}^+ \) comprises the multi-state model described in section 2. A generator model \( \text{Gen} \) creates the attributes of immigrants that enter the population. Via the output port \( Y_{\text{EM}} \) the attributes of emigrants are forwarded.

Whereas the DEVS atomic model allows a one to one translation of the stochastic model of Section 2 and provides additional structure by distinguishing nicely between immigration and the internal dynamics of our population, also certain problems remain. As individuals might enter and leave the population (due to migration, birth or death), the number of population members changes over model time. However, the classical DEVS approach requires a state space that is time invariant. Accordingly, at the time of model initialization the number of all individuals who might join the population up to time \( t_e \) needs to be assessed. As childbirth, emigration, or death rely on random processes, this is rather difficult. To make it worse, immigrants might enter the population at any time during simulation. The problem can be addressed by setting \( n \) (the amount of potential population members) to a large number that is unlikely to be exceeded during the simulation up to time \( t_e \). However, this is neither a general nor a (space-) efficient solution.

In addition, in the current model, no relationships between individuals are considered, implicitly assuming that the individual life courses can be calculated independently from each other. Whereas this is in line with current approaches in demography and the stochastic model we presented in Section 2, it is increasingly seen as the reason why simulation results fail in explaining societal patterns that are found in reality [17, 20]. Thus, a model should provide suitable means to take interrelations between individual life-courses into ac-
models. This is not the case, if a population is described as one monolithic atomic DEVS model. Using a DEVS atomic model for each individual would enable us to describe the relation between individuals by explicit couplings between atomic models. However, neither disappearing and appearing atomic models nor changing interactions between models are supported in classical DEVS.

Both observations ask for a different approach.

### 3.2. Population Network

DYNDEVS has been developed to support variable structure models, i.e., models that change their own composition, interaction, and behavior pattern [22]. In DYNDEVS a model’s description entails the possibility to change its own state and behavior pattern. Models are interpreted as a set of models successively generating themselves by model transitions which map the current state of a model into a set of models the model belongs to. Thereby, a sequence of model incarnations is produced. In following we employ a revised and parallel version of DYNDEVS, DYNPDEVS, [25] to describe the model population. The network or coupled model represents the population. Associated network components represent a population’s individual members. In doing so, we account for both the discrete event conception of the microsimulation and for the temporal variation in the corresponding model population. Individual behavior and therefore population dynamics emerge due to state transitions of individual models.

The dynamic DYNPDEVS network Pop describes the model population of the MicMac microsimulation. Its atomic components are ‘individual models’ denoted by I. At model time t, t ≥ 0, the population consists of nI, nI ∈ N, individuals. The set I comprises all individual models of Pop. We formulate the Pop network as structure

\[
\text{Pop} = \langle X_{IM}, Y_{EM}, \text{pop}_{\text{init}}, \Gamma(\text{pop}_{\text{init}}) \rangle
\]

where

\[
X_{IM} = \begin{cases} IM = [\psi^0_1, \cdots, \psi^0_{\text{im}}] & \text{if im immigrants enter} \\ \emptyset & \text{otherwise} \end{cases}
\]

\[
\psi^0_i (i = 1, \cdots, \text{im}) \text{ comprises the state } s_0 \text{ and the age of an immigrant when he/she enters the population, } (s_0 \text{ is the initial state of that process } Z(t) \text{ that refers to an immigrant’s life course)}
\]

Y_{EM} is a output port for emigrants to leave the population

\[
\text{pop}_{\text{init}} \in \Gamma(\text{pop}_{\text{init}}) \text{ describes the initial population model}
\]

\[
\Gamma(\text{pop}_{\text{init}}) \text{ is the set of all possible incarnations of the population model } \langle CP, I, \rho_{\text{pop}} \rangle
\]

\[
CP = \{ (i, Y_{em}, Y_{EM}) | i \in I \} \text{ couples the output ports } Y_{em} \text{ of all individuals } i \in I \text{ to the output port } Y_{EM} \text{ of the network}
\]

I is the set of all individual DYNPDEVS models \langle I_0, I, Y_{sc}, Y_{em} \rangle

I_0 is the initial individual model

Y_{em} is the output port of an individual which is connected to the emigrant output port of the network model and via which the internal state of the emigrant is sent

Y_{sc} is the structural output port via which the structural changes like death, emigration, or birth are signalized to the network

I has the structure \langle \Psi_I, \Psi_0, \delta_{\text{int}}, \rho_{\lambda}, \lambda, \tau \rangle where

\[
\Psi_I = \Psi \times \Psi \times \mathbb{R}_0^+ \text{ set of possible states that } I \text{ can occupy}
\]

\[
\Psi \text{ is the state space of a semi-Markov process that maps the individual life course}
\]

\[
\Psi = [s_1, s_2, a], \quad \Psi \in \Psi_I, \text{ is a tuple that comprises } I \text{’s last state } s_1 \in \Psi, I \text{’s current state } s_2 \in \Psi, \text{ and } I \text{’s current age } a
\]

\[
\Psi_0 = [s_0, s_0, a_0], \text{ comprises } I \text{’s initial state } s_0 \text{ and age } a_0
\]

\[
\lambda(s_1, s_2, a) = \begin{cases} \Psi & \text{if emigrate}(\text{argmin } s_j \in \Psi \{s_c\}(w_{s_c, s_j})) \\ \emptyset & \text{otherwise} \end{cases}
\]

with emigrate : \Psi \rightarrow \text{ boolean}

\[
\delta_{\text{int}}(s_1, s_c, a) = (s_c, \text{argmin } s_j \in \Psi \{s_c\}(w_{s_c, s_j}), a + \tau) \text{ is the internal state transition function}
\]

\[
\rho_{\lambda}(s_1, s_c, a) = bde(s_1, s_c), \text{ with } bde : \Psi \times \Psi \rightarrow \{ \emptyset, n_c \text{ newborns, death, emigration} \}
\]

\[
\tau a(s_1, s_c, a) = \tau \text{ where } \tau = \min \{w_{s_c, s_j}\} \text{ is the corresponding waiting time}
\]

\[
w_{s_c, s_j} \text{ is the (random) waiting time between the states } s_c \text{ and } s_j \text{ defined in equation (1)}
\]

\[
\rho_{\text{pop}} : \times_{i \in I} \times_{j \in Y_{sc}} Y_{IM} \rightarrow N(\text{pop}_{\text{init}}) \text{ is the network transition function}
\]

\[
\rho_{\text{pop}} \text{ takes into account}
\]

(i) the arrival of immigrants: \( X_{IM} \neq \emptyset \) (causes the creation of new individual models)

(ii) the announcement of childbirth(s) (and therefore the creation of \( n_c \in N \) new individual model(s))

(iii) the announcement of the extinction of a model
The Pop network exhibits one input port $X_{IM}$ for receiving information about the arrival of immigrants. It further possesses an output port $Y_{EM}$ to let emigrants leave the population.

The Pop network comprises a set $\Gamma$ of all possible model incarnations, including the initial populations $\text{pop}_{\text{init}}$ of the model.

Each individual model is equipped with an internal transition function $\delta_{\text{int}}$ that handles all of its state transitions. It incorporates furthermore the time advance function $\tau$ that determines the corresponding series of waiting times in states. If and when events (state transitions) occur during an individual life course, and therefore during the life time of an individual model, is specified by the stochastic model that is described in Section 2. Thereby, the time of next event and the event to occur are specified.

The simulation algorithm of DYNPDEVS ensures that among all possible next state transitions always the one is realized for which the shortest waiting time has been determined.

If an upcoming transition of an individual model is related to a birth event that causes $n_c$ newborns, function $\rho_c$ signifies to output port $Y_{sc}$ the implied structural network changes. Likewise, $\rho_c$ forwards to $Y_{sc}$ an imminent death or emigration event.

In the later case the activation of the emigration event is associated with the activation of an individual’s output port and the concurrent transfer of his/her data via $Y_{EM}$ to the outside world.

Subsequent changes in the structure of Pop and, accordingly, the creation of new model components and the deletion of existing ones is executed by the model transition function $\rho_{\text{pop}}$. The network transition function $\rho_{\text{pop}}$ is activated for eliminating or adding model components.

Whereas the elimination of components is only caused by information received via individuals’ output ports (death and emigration events), the adding of components is triggered by both information received via individuals’ output ports (birth events) and external events (arrival of immigrants). The network transition function possesses accordingly two arguments: (1) information received via individuals’ structural output ports and (2) the immigrants that enter the population. If $\rho_{\text{pop}}$ is called due to a structural change in an individual model, as second argument, an empty set is handed over to $\rho_{\text{pop}}$. Figure 3 illustrates the role of $\rho_{\text{pop}}$. An example situation is depicted in Figure 4.

The above model only exploits part of the functionality provided by the DYNPDEVS formalism.

Referring to the typical network model in DYNPDEVS, the population model does not have any structural input and output ports. Consequently, the definition of the population does neither need the $\rho_h$ function whose sole role it is to charge the structural output port. Both can be used in case hierarchical populations comprising other populations shall be modeled. The coupling structure of the model is rather simplistic as well, currently the individuals are only coupled via their output ports to the network output port.

Looking at the individuals, we see further simplifications of the typical atomic DYNPDEVS model. As the dynamics of each individual is currently solely driven by its internal dynamics, neither input ports nor an external transition function are needed. Of course those can and will be added, as soon as our evaluation of the information referring to linkages has been completed. In addition it is not foreseen that the individuals themselves change their own behavior. Thus, the set of incarnations (that typically represents an atomic DYNPDEVS model) boils down to one individual. Consequently, the model transition function $\rho_{\alpha}$ of the original DYNPDEVS becomes obsolete, as its only role is to produce new model incarnations.

4. EXAMPLE

We illustrate the functionality and the implementation of the presented population DEVS model by forecasting population numbers of a synthetic population based on the population of Italy. The forecast is executed over 17 years, starting from January 1, 2004 up to December 31, 2020. We consider individuals aged between 0 and 99. The data we employ for this purpose have been accomplished within the MicMac project. We take data to build up an initial population from [30] and all required transition rates from [11]. The initial population consists of 283,754 males and 300,839 females (which corresponds to 1% of the real Italian population at January 1, 2004). The demographic attributes that we include are:

![Figure 3: The set $\Gamma$ contains all possible model incarnations of the population network Pop. Network transitions are handled by the function $\rho_{\text{pop}}$. Beginning with the initial population model $\text{pop}_{\text{init}}$, every time a network transition is caused, a new network incarnation is created.](image-url)
Figure 4: Function $\rho_{pop}$ allows for modeling a time-varying number of individuals included in $Pop$. Of the initial population $pop_{init}$ individual $I_2$ dies at time $t_1$. The individual model informs via its output port $Y_{sc}$ the population $Pop$. Subsequently, function $\rho_{pop}$ performs the deletion of individual $I_2$. Given that an immigrant arrives at time $t_2$, $\rho_{pop}$ triggers the creation of model $I_4$ for specifying the newly arrived individual.

State Variable Possible Values
---
gender female, male
marital status never married, married, divorced, widowed
fertility status no child, first child, ..., fourth child
mortality alive, dead
migration immigrated, emigrated

Figure 5 shows the initial population cross-classified according to age, sex, and marital status (upper graph), as well as according to age, sex, and fertility status (lower graph). The displayed age-gender pyramids are bell-shaped. At January 1, 2004, most people are aged from 30 to 50, and of these the majority are in their thirties. The number of children and adolescents aged from 0 to 20 is relatively homogeneously distributed over the corresponding age classes. At older ages, the proportion of women is higher. For Italy typical is the comparatively high proportion of very old people (over 90 years).

As nowadays the majority of Italians identify themselves as Roman Catholic, divorce is still rare. For this reason and due to a higher mortality of men, at older ages many women are widowed. At January 1, 2004, many of the women aged 20 to 40 are married, which is contrasted by a lower number of married men of the same age. This situation is mirrored by the upper population pyramid. The lower pyramid shows that at January 1, 2004, at the ages from 20 to 30, women are either childless or mother of only one child. Mothers of two or more children are older than 30.

In our example individuals can marry, they can get divorced or become widowed, and they can marry again (corresponding rates are shown in Figure 6). Females experience fertility events which implies transitions to higher order parities, e.g., from one child to a second one (cp. Figure 7). For females the probability to marry is highest around the age of 30, for males around the age of 32. At the ages from 25 to 35 married females face the highest risk to give birth to a first or a second child. The propensity of an unmarried mother to give a further birth is relatively high compared to the equivalent propensity of an unmarried childless female. Empirical findings show that the propensity to change the marital status depends on the presence of children. For instance, the probability of a divorce is lower for couples with children than for childless couples. Likewise, fertility rates depend on the marital status. The probability to become mother is generally higher for a married female than for an unmarried one. To
simplify matters in this example we assume that the propensity to change the marital status and the propensity to give birth are age-specific, but were held constant over calendar time.

We use mortality rates that are age- and sex-specific and vary over calendar time (cp. Figure 8). The risk to die gradually decreases over time. It drops stronger for men than for women. In- and out-migration to and from the model population is assumed to take place in our example. We make use of an immigration pattern and emigration rates that are typical for contemporary Western European countries. The migration data relies on the EuroPop2004 projections (baseline scenario) provided by EuroStat [5]. Figure 9 displays the number of immigrants that enter the model population during simulation (upper graph) and the propensity to emigrate (lower graph). The latter is assumed to be age- and sex-specific, but to be constant over calendar time. Emigration is more prevalent among young adults, in particular men. Mostly, young men aged 20 to 30 enter the Italian population. This is contrasted by a lower number of women at the same age. Italy’s immigration pattern is mainly shaped by the many Eastern European young men who enter the country to find profitable jobs, and who are often accompanied by their wives. Due to a lack of corresponding data in this example we set the rather unrealistic scenario that natives and immigrants as soon as they have entered the model population behave in the same way. On request the described data that has been used to run this example are available from the first author.

The Mersenne Twister random number generator is used for the generation of the random waiting times in the demographic states. We perform 50 simulation runs on an Intel(R) Core(TM) 2 Duo CPU with a 1.60GHz equipped with 2GB memory. The runs last between 75.61 and 78.01 seconds. During simulation all demographic events (births and deaths, state transitions of individuals, immigration and emigration events) are tracked. The average number of simulated events is 962,353. In order to evaluate and illustrate the output of a simulation run, we apply the statistical software R (cran.r-project.org). Several interesting results that the simulation output reveals are mentioned and illustrated in the subsequent. Figures 10 and 11 show comparisons between the population structure that was given in the initial population and the simulated population on 2020-12-31. Population aging is evident. Moreover, the gain in males’
life expectancy becomes apparent. Figure 10 depicts in particular the number of women in each age class and their respective numbers of children. It is obvious that during the simulation period the number of newborns decreases. A comparison of the number of women aged from 20 to 30 having one child at 2004-01-01 with those who have one child at 2020-12-31, shows a postponement of the age of first birth. This phenomenon might be boosted by the fact that more and more women stay unmarried or marry late (cp. Figure 11). Furthermore, compared to 2004-01-01 apparently fewer women have three or more children in 2020.

Figure 12 illustrates the relative frequency distribution of females experiencing a transition into and out of the ‘married’ status. The distribution has been computed referring to the corresponding state of origin and state of destination. We find that 72 percent of the women who marry, marry for the first time (transitions from ‘NM’ to ‘M’), consequently, the remaining 28 percent remarry (transitions from ‘D’ or ‘W’ to ‘M’). Apparently, the proportion of first marriages is high. Only few women remarry after getting divorced or being widowed. Among remarriages the proportion of divorced women is higher than that one of widows. During simulation, of married women 48 percent experience a divorce (transitions from ‘M’ to ‘D’), 34 percent a transition to widowhood (transitions from ‘M’ to ‘W’), 10 percent die (transitions from ‘M’ to ’dead’), and six percent emigrate (transitions from ‘M’ to ‘rest’). An interesting outcome is that the fewer children a wife has, the more likely she is to experience a divorce event. Over calendar time the age at death distribution shows for males a slight increase in the median age at death, for females no significant effect is observable (cp. Figure 13). As for both sexes decreasing mortality rates have been assumed, this outcome is a little bit unexpected.

5. CONCLUSION

The implementation of the MicMac population in DEVS and afterward DYNPDEVS revealed limitations of the original approach but also opened up new avenues for extending the original model for a more realistic capturing of demographic dynamics and thus, for better predictions. Our next steps will be including linkage between individuals and thereby, relating life courses of different individuals. Including linkages in continuous-time microsimulation models is currently hampered by a lack of suitable modeling ap-

Figure 8: Mortality rates of females and males (rates are given on the log-scale)

Figure 9: Number of immigrants that enter the model population during simulation (upper graph) and the age- and sex-specific propensity to emigrate (lower graph)
Population Pyramids According To Children Ever Born 2004−01−01 and 2020−01−01

Figure 10: Population pyramids displaying the distribution of individuals according to children ever born at start and end date of simulation.

Population Pyramids According To Marital Status 2004−01−01 and 2020−01−01

Figure 11: Population pyramids displaying the distribution of individuals according to marital status at start and end date of simulation.

With exploiting a DEVS-based approach we deliberatively aim at an approach that lies in between. With its emphasis on states and state transitions and its modular structure, it is on the one side sufficiently close to the typical individual-based models, on the other side it supports additional structures that facilitates a successive refinement.

REFERENCES


Figure 12: Relative frequency distributions of origin and destination states to/from being married; coding: never married (NM), married (M), divorced (D), widowed (W), emigrated (rest), open age category (restCategory)


Figure 13: Boxplots of age at death distribution of males and females


AUTHOR BIOGRAPHIES

SABINE ZINN is a Ph.D candidate at the MPIDR and the University of Rostock. She received a diploma in business mathematics from the University in Jena. She is interested in microsimulation and statistical programming. Her email address is <zinn@demogr.mpg.de>.

JUTTA GAMPE is the Head of the Laboratory of Statistical
Demography of the MPIDR. She received her Ph.D in Statistics from the TU Berlin. Her research interest is in statistical demography and related topics. Her email address is <gampe@demogr.mpg.de>.

JAN HIMMELSPACH is a post doc in the Computer Science Department at the University of Rostock. He received his Ph.D. in Computer Science from the University of Rostock. His research interest is on software engineering for modeling and simulation. His email address is <jh194@informatik.uni-rostock.de>.

ADELINDE M. UHRMACHER is a Professor at the Department of Computer Science at the University of Rostock and head of the Modeling and Simulation Group. Her research interests are in modeling and simulation methodologies and their applications. Her email address is <lin@informatik.uni-rostock.de>.