A Calibration to Properly Design a Model Integrating Residential Mobility and Migration in a Rural Area

Sylvie Huet, Nicolas Dumoulin, and Guillaume Deffuant

Abstract We propose a hybrid microsimulation and agent-based model of mobility integrating migration and residential mobility. We tested it on the evolution of the population of the Cantal, a French "département" with 150,000 inhabitants. We calibrated it using various data sources from 1990 to 2006, and tested its predictions on other data of the same period and on the period 2007–2012. The spatial heterogeneity of the evolution is well reproduced and the model makes surprisingly correct predictions despite numerous simplifying assumptions. From this calibration we learnt more about how to model residential mobility and migration considering an agent-based model approach.

In the literature [1], most dynamic microsimulation models of population evolution do not include spatial dynamics like the residential mobility. For example the microsimulation model of the French population called DESTINIE only includes the migration process, implemented by adding (or removing) individuals corresponding to the migratory balance. On the contrary, agent-based models and cellular automata easily integrate spatial dynamical processes. However, agent-based models which take into account all main demographic processes have not yet been developed [1] even if they have been involved in particular dynamics where they appear as more convenient, especially partnership formation and spatial mobility. More recently hybrid models combining advantages of the two modelling approaches are the subject of a growing interest [2] review the existing hybrid approaches as for example SVERIDGE [3] which integrates inter and intra migration [4] is a very interesting review on residential choice and household behaviour; it outlines several shortcomings of the current approaches:

S. Huet (✉) • N. Dumoulin • G. Deffuant

Irstea, UR LISC Laboratoire d'ingénierie des systèmes complexes, 9 Avenue Blaise Pascal, Aubière 63178, France

e-mail: sylvie.huet@irstea.fr; guillaume.deffuant@irstea.fr

- The difficulty to connect the decision to move with the residential choice per se, and more generally the lack of retroaction between the demographic and the mobility processes;
- The decision to move is undoubtedly the most neglected aspect in the residential process, most models putting much more emphasis on the location choice;
- The location is generally decided using a discrete choice model which includes housing prices and housing and neighbourhood characteristics, then it is unclear whether this is a direct or indirect utility function;
- Regarding the location, the subset of alternatives is randomly and uniformly drawn from the whole set of vacant housings, disregarding any strategic consideration in the search process of the household;
- Migration and residential mobility are considered apart from each other, as independent decisions while migration and residential mobility cannot be distinguished for people living close to the border.

Thus, to overcome these limitations we propose a model coupling agent-based to microsimulation approaches and integrating demographic, migration, and residential mobility processes. The main features of the processes are derived from the French literature on residential mobility and migration. From [5, 6], we know that in France the main reasons to move are firstly related to family events [7, 8]. Established from the analysis of various French surveys, creating or splitting up a couple are such family events explaining most of the residential mobility. The second set of reasons is professional [8]. Notice that moving decreases with age and point out that the short distance mobility is rather linked to the modification of the family structure while the long distance mobility is more often associated to professional changes. The third type of reasons concerns the change in the tenure (mainly between renters and owners) [9]. From this literature review, we retained that a decision to move is due to: (1) the formation of a new couple; (2) the split of a couple; (3) a too long commuting time (higher than a threshold called *proximity* parameter) after a change of job, a new partnership, or a move; (4) a decrease in the housing satisfaction level due to a family event: we decide to capture this change through the compatibility of the housing size with the family size. As in [10], these events are taken into account in the decision function to move that we propose, while a second function is dedicated to the location choice of the new dwelling. Most of the demographic attributes and dynamics are parameterized from data extracted from various surveys, mainly the national Census and the Labour Force Survey. But some of the model parameters, especially those in relation with the mobility, have no documented value in the literature; we thus need to calibrate them. We argue that this calibration provides some insights on the relevance of our integrated model for mobility. Moreover, our model of residential mobility aims at reproducing the evolution of the population of the Cantal, a French "département" from 1990 to 2006. This problem shows some specific difficulties: (1) the population globally decreases but locally increases and shows a strong spatial heterogeneity; (2) the size of the basic spatial element, the municipality, is less than 1000 agents for 75 % of its 260 municipalities.

The next section presents the model, shortly for most of the dynamics for which the parameter values are directly derived from data, with more details for dynamics related to residential mobility which have to be calibrated. Then, the following section describes the calibration process and the final section proposes a discussion.

1 The Model

We have adopted a hybrid micro-modelling and agent approach. The purpose of the model is to determine how different dynamics at household level determine the population evolution in a network of rural municipalities. We assume that this evolution depends, on the one hand, on commuting flows and service, and on the other hand, on the number of jobs in various activity sectors (supposed exogenously defined by scenarios). Existing literature [11–14] stresses the importance of the different types of mobility between municipalities, commuting, residential mobility (short range distance), migration (long range distance) [4] and the local employment offer generated by the presence of the local population. Moreover, obviously, it appears also essential to include the demographic evolution of the municipality considering the strands explaining the local natural balance.

1.1 Main Entities, State Variables, and Scales

The model represents a network of municipalities and their population. The distances between municipalities are used to determine the flows of commuting individuals but also the flows of people making residential mobility as well as those looking for a partner. Each municipality comprises a list of households, each one defined as a list of individuals. The municipalities also include the offers of jobs, of residences and their spatial coordinates. Here is the exhaustive list of the main model entities with their main attributes and dynamics.

MunicipalitySet and Municipality. The set of municipalities can be of various sizes. It can represent a region of type NUTS 2 or NUTS 3,[1] or more LAU or intermediate sets of municipalities such as "communauté de communes" in France. Municipality corresponds to LAU2.[2] The municipality is the main focus of the model. It includes: A set of households living in the municipality; the set of jobs

---

[1]Eurostat defines the NUTS (Nomenclature of Territorial Units for Statistics) classification as a hierarchical system for dividing up the EU territory: NUTS 1 for the major socio-economic regions; NUTS 2 for the basic regions for the application of regional policies; NUTS 3 as small regions for specific diagnoses; LAU (Local Administrative Units 1 and 2) has been added more recently to allow local-level statistics.

[2]Consists of municipalities or equivalent units.

existing on the municipality (i.e. without those occupied by people living outside the modelling municipality set); the distribution of residences, or housings, on the municipality.

Each municipality has rings of "nearby" municipalities (practically every 3 km of Euclidian distance) with a maximum Euclidian distance of 51 km where individuals can find out jobs and partners while households can find lodgings. It is computed from the spatial coordinates of the municipalities. A threshold distance called "proximity" is the distance beyond which the municipalities are considered too far from each other for commuting between them.

There is a particular municipality, called "Outside," which includes available jobs accessible from municipalities of the set, but which are located outside this set. The job offer of Outside is infinite and the occupation is defined by the process ruling the probability of individuals to commute outside the municipality network.

The job and the residence. A job has two attributes, a profession and an activity sector in which this profession can be practiced. It is available in a municipality and can be occupied by an individual. The profession is an attribute of the individual at the same time it defines a job. In France, it takes six values. There are four activity sectors: Agriculture, Forestry and Fishing; Industry; Building; Services and Commerce. Overall, considering the six professions for four activity sectors, we obtain 24 jobs to describe the whole diversity of jobs in the Cantal "département." The residence has a type which is classically its size expressed in number of rooms. A residence is available in a municipality and can be occupied by 0, one or more households. Indeed several households can live in one residence for instance when a couple splits up and one of the partners remains in the common residence for a while. It is also the case in some European countries where it is customary for several generations to live under the same roof. The job and residence offers are initialised from Census data (see [15] for details).

Household. It corresponds to a nuclear family[3] and includes a list of individuals who have an occupation inside or outside the municipality. The first households located in each studied municipalities are built using algorithm described and studied in [16, 17]. At the initialisation time, households are associated randomly with residences. Then, before running, a first phase iterates moves in the same municipality in order to find a good fit between residences and households.

Individual. A new individual is instantiated from one of the adults of a household having the "couple" status in the birth method, or directly from the initialisation of the population, or by immigration. Her age of death, of entrance into the labour market, and of retirement are attributed when she is created. The values of these ages are drawn from distributions derived from French Census data. The age at which the agent becomes an adult and creates her own household is when she finds her first job; or she is chosen by a single adult as a partner; or she remains the only child in a household after her parents left or died while her age is higher than

---

[3]A nuclear family corresponds to the parents and the children; that is a reductive definition of the family corresponding to the most common one in Europe nowadays.

parameter firstAgeToBeAnAdult (15 years old for the French or other European National Statistical Offices). This change can lead her to move from the parental residence because of a low housing satisfaction level, but it's not always the case. The activity status of the individual defines her situation regarding employment, especially if she is looking for a job. The individual can quit a job, search for and change jobs ... Her profession is an attribute indicating at the same time her skills, level of education and the occupation she can look for.

1.2 Process Overview and Scheduling

The main loop calls processes ruling the demographic evolution, the migrations, the job changes, and their impact on some endogenously created services and/or jobs. First, the scenarios (i.e. external drivers) are applied to the municipalities. Then, demographic changes are applied to the list of households. Time is discrete with a time step of 1 year. The households are updated in a random order. The following gives more details about the dynamics.

Dynamics of offer for jobs, services and housing. In the municipality objects, changes in housing and job offers are specified in scenarios which are parameterised from data of the French censuses. Various dwelling sizes are considered in order to match the needs of households.

Dynamics of labour status and job changes. A newly born individual is initialised with a student status that she keeps until she enters the labour market with a first profession. Then, she becomes unemployed or employed with the possibility to look for a job. She may also become inactive for a while. When she gets older, she becomes a retiree. We describe rapidly these dynamics to situate them in the global picture of the model. They have been parameterised from data. The model and its parameterisation are described with more details in [10, 15, 18].

Demographic dynamics. A household can be created when an individual becomes an adult or when a new household arrives in the set of municipality (i.e. in-migration). The main reasons for household elimination are out-migration and death. Three main dynamics change the household type (single, couple, with or without children and complex[4]): makeCouple; splitCouple and givingBirth. Moreover, households change locations and that is the particular challenge of our modelling approach presented in this paper. The implementation of the "change location" related dynamics for Cantal is presented in the following (see also [10, 15]).

Household migration and mobility. In the process of changing residence, we include both residential migration and mobility without differentiating them on the distance of move, as it is often the case [4] in the literature. The submodel we

---

[4]A complex household is a household which is not a single, a couple with or without children.

propose directly manages both types of moving. However, it turned out easier for us to distinguish two categories of migration: the migration from outside to the set; the migration from inside the set.

The immigration into the set is an external forcing. Each year, a number of potential immigrants from outside the set are added to the municipalities of the set. These potential immigrants can really become inhabitants of the set if they find a residence by themselves or by being chosen as a partner by someone already living in the set in case they are single. That is the only actions they can do. Until becoming real immigrants, they are temporarily located into a municipality with a probability depending on its population size and its distance to the frontier of the set. A particular attraction of young agents for larger municipalities is also taken into account.

The mobility of agents already living inside the set of municipalities is mainly endogenous. Such a mobility can lead the household simply to change residence, municipality or to quit the set of studied municipalities. Overall, a household decides to look for a new residence when:

- a couple splits: one of the partners has to find out another residence even if she remains for a while in the same residence (creating her own household);
- one of its adults finds a job away from the current place of residence (beyond the proximity parameter which has to be calibrated) and she is the one deciding for the household (the leader is chosen at random every time a decision should be taken);
- it is not anymore satisfied by the residence. This satisfaction is assessed through a function possibly considering two dimensions identified in the literature (see our introduction): the difference of size between the occupied size and an ideal size for this household; the average age of the household. The calibration procedure has to determine if both dimensions should be taken into account to reproduce available data about residential mobility and migration.
- a student or a retiree decides to move outside the set (parameterised by data);

These are the events leading to a decision for changing residence. This is important to notice that other events possibly imply these events and lead to a decision for a residential mobility. This is the case of most of the family events with have an impact on the distance to commuting, the size of the household...

When a household has decided to move, the principle for the search of a new residence is the same as the one for searching a place to work. The leader of the household (chosen at random among the adults each time a decision has to be taken for the household) looks from the closest to the furthest job offer, considering successive rings of range of distance[5] of size 3 km for France (i.e. the same basic distance used to search for a job). She starts from her place of work (or residence if she does not work), meaning at a distance at most 3 because the first range is 0–3 km (i.e. first ring for search). If she can't find a satisfying place to live in, she

---

[5]The distance definition depends on the parameterization of the model.

continues looking from 3 to 6 km. She iterates the procedure, from 6 to 9 and so on, until finding a not empty list of possible residence offers or reaching a ring which is too far and the research stops. Before accepting to consider an offer, she checks the residence offer is not too far from the place of work of her partner (if she has one). She can also move outside. The decision moving outside depends on the parameterisation. The searching procedure finished, if the agent had not found out a residence in her current municipality and if she has found possibilities elsewhere, she decides to move outside using the probability to move outside knowing her municipality of living. Finally, if she does not move outside the set of municipalities and has found a residence, she chooses at random a municipality of residence in the list of collected residence offers. The probability to move out of the set of municipalities varies with the age of the individual. What is an acceptable residence offer to collect during the search procedure depends on the size of the housing. The level of acceptation of a possible size is 0 if the size does not respect the fact that the household wants to increase or decrease the current size of its house. If the offered size respects this tendency its probability to be collected for the list of choices decreases with the difference of size between the offered and ideal size.

The way a household decides if a residence offer is satisfying has to be parameterised too. The level of satisfaction of a possible place of residence depends on two dimensions: the municipality where the proposal is; the size of the proposed housing. We have noticed from the literature [9] that most young couples want to become owners and quit renting. In rural areas, they often prefer a house instead an apartment. Such housing is much more rare and expensive in town than in the countryside. Thus, a municipality is examined as a place to reside if it satisfies the need for a house (and not for an apartment) for the household of size higher than one. Practically, a municipality has a ratio of house offer over its residence offer and it is considered in the research procedure with a probability equal to this ratio for every household larger than one member. Only the three main towns of the Cantal have a ratio possibly lower than 1. This ratio should be parameterised by the calibration process. The result of the calibration indicates how useful is considering a more detailed model of the offer. The second dimension influencing the place of residence is the size of the residence offer: we simply state it has at least to respect the need for increasing size or decreasing size of the household.

Overall, the parameters for immigration are: yearly migration rate; number of out of the set migrants in year $t^0 - 1$; probabilities for characteristics of the immigrants (size of the households, age of individuals . . .); distance to the frontier of the region of each municipality. Regarding the endogenous residential mobility, we have to parameterise or calibrate: a function defining the satisfaction of the size of housing and the satisfaction of job and workplace; migration rates of different agent categories extracted from data (as for example in France, for students and retirees).

Most of the events related to couples (formation, splitting, birth) can impact on residential mobility. Moreover they are poorly known from a modelling point of view and should be calibrated considering possibly different dimensions. They are presented in the following.

MakeCouple. The couples are made of individuals both existing in the population (either already living in dwelling of the municipality or a potential migrant). Therefore, two "compatible" individuals have to be identified for creating a couple. The usually procedure presented in the literature [1, 19–29] appears very heavy to compute for our case, especially because we have to take some spatial constraints into account while they don't. Moreover, it generally forbids that the individual lives other events during the year she tries forming a couple, such as moving for example. Thus we opt for a close procedure, significantly lighter computationally. We decompose the couple creation dynamics in two subdynamics: the decision searching for a partner; the search for the partner. The implementation for Cantal is:

- at each time step, each single individual searches for a partner following a probability; we do not consider a pool of single who are looking for another one as it is classically done in the literature (simply, a searching single can meet another single who is not necessarily searching a partner);
- when she searches, she tries a given number of times in her place of residence before trying in every municipality close to her own and her place of work to find someone who is also single and whose age is not too different (given from the average difference of ages in couples and its standard deviation from INSEE[6]). She begins searching very close and goes to search further until reaching a distance equal to the threshold parameter "proximity" (in the same way she looks for a residence or a job). She can search among the inhabitants or the potential immigrants.

Overall the probability to search and the maximum number of trials have to be calibrated since they cannot be derived from available data. The calibration should indicate if these two parameters are really useful (see discussion). Notice that when a couple is formed, the new household chooses the larger residence (the immigrating households always go into residences of their new partners) and this move can lead one member to commute very far (at a distance higher than the MunicipalitySet parameter proximity). This situation can change only when she is becoming the leader triggered by the job search method and implying that the household will aim to move closer to her job location, or if she changes of job.

SplitCouple. Even if the union duration is often used to model the union dissolution [1, 19, 20, 23–25], it appears that observed splitting probability is more strongly correlated with the date of union creation than with the duration of the union [30]. However, we do not have in our data when the couples existing in 1990 were formed. Thus, to limit the number of parameters and the complexity of the model, we use a constant probability to split that we will calibrate. When the split takes place, the partner who works further from the residence leaves the household

---

[6]French Institute of Statistics and Economical Studies

and creates a new household, which implies that she searches for a new residence. When there are children, they are dispatched among the two new households at random.

Giving birth. Following the reviews of the literature [1, 19, 20], the most common variables used for the fertility are the age, the marital status, and the number of children already born. However, in practice the total number of variables used to define the probability of birth is often very high. The same type of approach is developed in [23]. We have not enough data related to the Cantal to envisage such a data-based approach. In our model, only "couple" households can have children, and one of the adults should be in age to procreate. A couple has an annual probability to have a child computed from an "average number of children by individual." This "average number of children" has to be calibrated because it cannot be directly derived from existing data such as the fertility rate.[7] The calibration should indicate us if this "average number of children" is a constant or increase over time. The other parameters for "giving birth" are the age bounds to give birth: they are 18 and 45 for Cantal.

2 The Calibration of the Residential Mobility Dynamics

2.1 *The Method: The Particle Swarm Optimization (PSO)*

The particle swarm optimisation (PSO) [31] is a heuristic inspired from a social metaphor of individuals using their knowledge as well as the information given by their neighbours in order to find out a good solution. PSO considers a swarm of particles moving in the space of the parameter values. Each particle knows its position and the associated value of the function to minimise, the best position found by one of its neighbours, and its best position since the beginning. At each time step, every particle moves in the space with a velocity which is a compromise between its previous velocity, the knowledge about its best position and the best position in her neighbourhood. It has been shown to successfully optimise a wide range of continuous functions [32]. Some theoretical results about its convergence and good practices for choosing its parameters are available [33]. We parameterised as follows the algorithm:

---

[7]The total fertility rate (TFR), sometimes also called the fertility rate of a population is the average number of children that would be born to a woman over her lifetime if: (1) she were to experience the exact current age-specific fertility rates through her lifetime, and; (2) she were to survive from birth through the end of her reproductive life. It is obtained by summing the single-year age-specific rates at a given time. This computation is equivalent to give an identical weight to each age range, whatever their real weight in the population. It suppresses the structural effect linked to the distribution by age ranges of women in age to procreate.

- 15 particles: it has been shown few particles are sufficient—the choice is empirical depending on the perceived difficulty of the problem (values between 10 and 60 are used in practice). Their initial positions are randomly chosen in the space of research where particles are confined during the iterations.
- A "ring" neighbourhood: each particle has itself and the two particles located before and after it on the ring for neighbours.
- 10 replicas for 700 iterations have been executed to define our best result.

2.2   The Criteria for Calibration

The function to minimise is the sum of the differences between indicators computed from simulation results and the data. The 518 indicators we used are described in the following as well as the results of the calibration process (8 at the regional level; 54 at the "canton" level; 456 at the municipality level).

2.3   The Results of the Calibration Process

The exact values of the nine calibrated parameters are given and discussed in the following concluding section. The present subsection shows how 50 replicates of these 9 calibrated parameters fit the indicators composing the calibration objective function. Figure 1 show the fit level of the first set of 8 indicators. The value is given 1 year less for the simulated results compared to the target value of the INSEE in order to make them comparable because the simulated results are measured at the
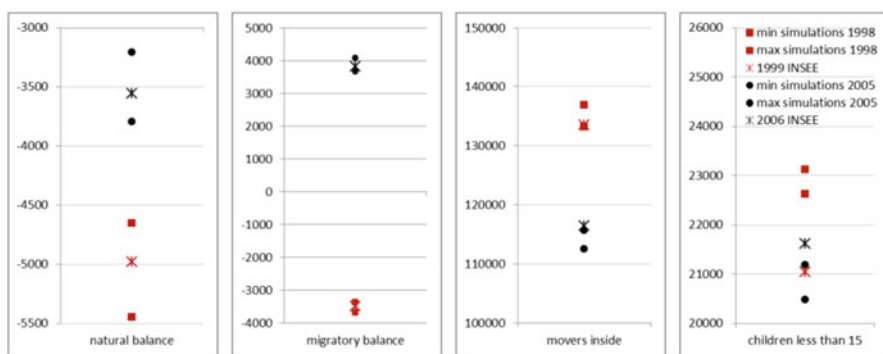

Fig. 1 Comparison between the simulated values and the target value. Target values from INSEE are represented by a cross while minimum and maximum values obtained over 50 replicas of the best parameter set are represented by *red squares* in 1998 and by *black circles* in 2005. From *left* to *right*: natural balance; migratory balance; number of movers (i.e. people changing residence) inside the Cantal; number of children less than 15

end of the year while the INSEE gives its value on the 1ˢᵗ January of a year. We can see on the figure that: the natural balances from 1990 to 1999 and from 2000 to 2006; the migratory balances from 1990 to 1999 and from 2000 to 2006; the number of individuals changing residences inside the Cantal from 1990 to 1999 and from 2000 to 2006; the number of individuals less than 15 years old in the Cantal in 1999 and 2006, are well fitted by the model since the targeted values are embedded in the simulation result ranges over 50 replicas. Except the number of individuals being less than 15 years old in the Cantal in 1999 and 2006 the values from data are within the interval of variations of the 50 replicas. It might not be relevant to try to obtain a better fit because the data giving the target value are not so stable or reliable:

- the number of children results partly from the immigrating children who are potential immigrants really come in the set of municipalities; these potential immigrants are built regarding their age on reference data measured only on 25 % of the population. Then it is possible the difference between the target and the model comes from this approximation;
- the targeted number of movers inside does not correspond, as measured in the model, to the number of agents moving from one residence to another. Indeed, in the census, this number corresponds to the number of inhabitants having a different address since the last census, but we don't know if they moved several times.
- the initial population in the simulation model has been built directly from the detailed data of the 1990 census. However, the next target in 1999 and 2006 has been built on the following census which has changed the counting method. Then, instead the 159,245 counted by the 1990 census and implemented in the simulation model, the next census which are corrected "a posteriori" assesses the population in 1990 at 158,723 (522 less).

The other 4 basic indicators are averaged errors over spatial objects. They relate to: (1) Population of each of the 27 "Cantons" of the Cantal (error based on the average of the 27 differences between the simulated value and the targeted one) in 1999 and 2006 (forming overall 54 indicators since two dates are considered); (2) Population of each of the 228 municipalities of at most 1000 inhabitants in 1990 (error based on the mean of 228 differences between the simulated value and the targeted one over the target value) in 1999 and 2006 (forming overall 456 indicators since two dates are considered).

The two first dedicated to the 27 "Cantons" of the Cantal in 1999 and 2006 are shown in Fig. 2. Only 48 % in 1998 and 44 % in 2006 of the "cantons" are given by the simulations without error (target comprised in the interval of replicas' results). For the others "cantons," the errors given by the average value over the replicas and expressed relatively to the target values are below 0.2.

Despite these errors the global evolutions of the "cantons" in terms of decrease or increase are well respected (see Fig. 3). Indeed, only one "canton" over 27 increases instead of decreasing but the error is not so large since in fact the population of the canton 1523 [Vic-sur-Cère] (on the right side of Aurillac) increases for an average
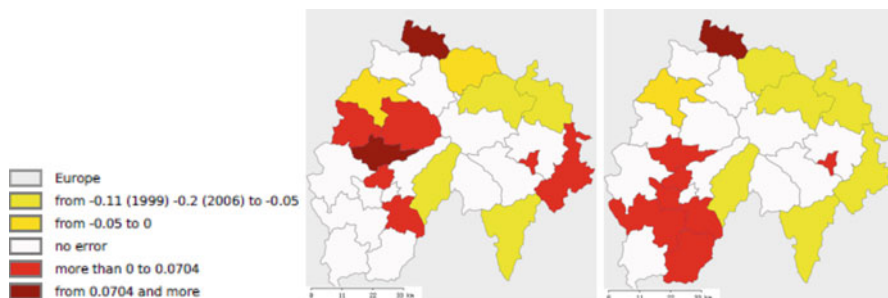
Fig. 2 "Canton" maps. Average relative errors over the targeted population of each "canton" for 50 replicas (considering the bounds defined by the minimum and maximum results over 50 replicas, the error is computed as the (larger distance between one bounds and the target)/target) in 1999 (on the *left*) and 2006 (on the *right*). Then, the value no error means that the target value is comprised in the interval defined by the minimum and the maximum results over 50 replicas. A negative value means model gives a larger population than the target while a positive value means the contrary
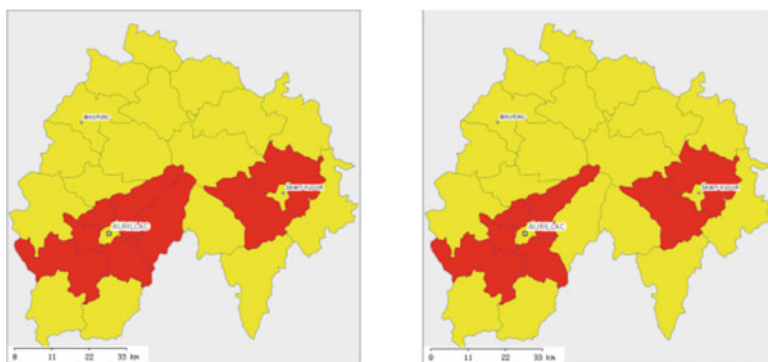

Fig. 3 "Canton" maps. *Left*: average absolute model evolution 2006-1990 over 50 replicas, in terms of decrease in *yellow* (−1043.5 to 0), and increase in *red* (0 and more)—*Right*: INSEE evolution 2006-1990, *yellow* –1296 to 0; *red* 0 and more

of 133.26 agents while the INSEE predicts a decrease of 295 individuals. Then the error is about 430 agents while the whole population of the Cantal is about 150,000.

Concerning the two last indicators dedicated to the population of each of the 228 municipalities of at most 1000 inhabitants in 1990 in 1999 and 2006: 68 % of these municipalities in 1998 (154 municipalities over 228) and 69 % in 2005 (157 municipalities over 228) have populations for INSEE (i.e. the reference) comprised in the minimum–maximum range of the model's results given by the 50 replicas of the best set parameter values.

3 Conclusion: Discussion

We propose a model of mobility integrating commuting, family events, satisfaction of housing, job availability, and migration. We assume that the decision to move depends on: some family events as union dissolution and couple formation; the distance of commuting; the adequacy of the size of the housing expressed in number of rooms to the size of the household. Nine parameters of these dynamics cannot be derived directly from available data and we chose to calibrate them. The calibration provides values for which the results from simulations are close to the reference data. The following presents these parameters and their calibrations values for our implementation of the whole population of the Cantal region (starting with one agent "individual" for each "real" individual of the region, with virtual "individuals" composing "households" located in municipalities formed of the same number of households as "real" municipalities). They are discussed in order to define what we have learnt regarding how to model mobility.

Two parameters relate to the formation of couples, the calibration yields a probability to search of 0.22 and a number of search trials of 5. This means only one over 4 or 5 singles decide 1 year to look for a partner (not everyone), knowing that she is susceptible to meet any single agent since she has a congruent age to her and using the location rule described in this paper. This result indicates us the utility of the two parameters: indeed, considering everyone search for a partner and just limiting the number of trials is not a sufficient model to fit data of reference. A two-step process is required based on a decision phase and a search phase.

One parameter drives the decision to split for a couple for which we made the simplest hypothesis considering the average duration of a couple is constant. Our calibration yields a splitting probability of 0.107. This means on average couples have duration of about 9–10 years. Even if we know from literature a cohort approach is more relevant and then that we overestimate the dissolutions of older people, the choice for a constant probability to split is sufficient for calibrating the model properly. This is very useful since in our case, as almost always, we have not enough data to implement a cohort approach.

Regarding the birth rate, the question was to know if it is necessary to consider an increasing probability to give birth for each couple to obtain the right number of births (remaining almost constant over the time while the fertility increases during the considered periods in Cantal). Our calibration confirms that the model has to consider an increasing individual probability to give birth. Indeed, the slope of our linear function giving the average number of children of a woman a given year (and then the searched probability) is 0.063 and the intercept is 2.38.

Except for splitting of a couple or the migration of students and old people, the decision to move is associated to the satisfaction level about the current dwelling. This satisfaction depends on the size of the dwelling compared to the size of the household, and possibly from the average age of the household. How high the satisfaction increases with age whatever the dwelling size is, controlled by the

parameter *pres*. It has been parameterised to 0.045. This value confirms us older people does not want to move even if their dwelling is too large: if it were not the case, it would be 0.

We also assume a searching process for a housing, for a job or for a partner (as finding a partner implies changing residence for at least one of the partner) starting from the best location and enlarging the search progressively if the object of search (house or partner) is not found. For an occupied active individual, the best location of housing is the municipality of work. For the other individuals, the best location is the current municipality of residence. The issue of the calibration process is to define if the individual searches until she finds or if there is a maximum distance where she stops searching. The calibration suggests that the search should be bounded by 18 km (Euclidian distance). The quality of this result can be evaluated through the capacity of the model to give values describing the commuting distance and the residential mobility distance distributions close to the references. Figure 4 shows the two commuting and residential mobility distributions given by the model and the INSEE. They are close enough even if the model tends to underestimate the probability to move in the same municipality (3 km or less). The result shows a reasonable fit given the hidden complexity of the dynamics behind such a distribution.

Only for the search process for dwelling, the calibration has to indicate if a specific mechanism should be added for avoiding sometimes the three largest municipalities of Cantal as residence locations: we should know if bigger municipalities are less preferred by families for their residence. The parameter value for "dispo threshold" means they are maintained in the searching zone only for 32 % of the searches susceptible to contain them. Overall, we conclude that both mechanisms, one pushing agents to live outside the main municipalities, the other bounding the area where people search for new housing or partner, are necessary to reproduce the spatial evolution of the population. It is probable, even if the parameters change depending on case study these dynamics should always be both present to model the spatial evolution.
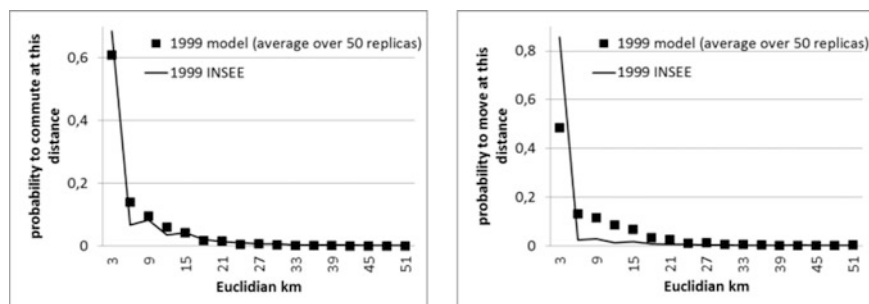

Fig. 4 Commuting (on *left*) and residential mobility (on *right*) distance distributions for INSEE (*line*) and the model based on results from calibration (*squares*) averaged over 50 replicas
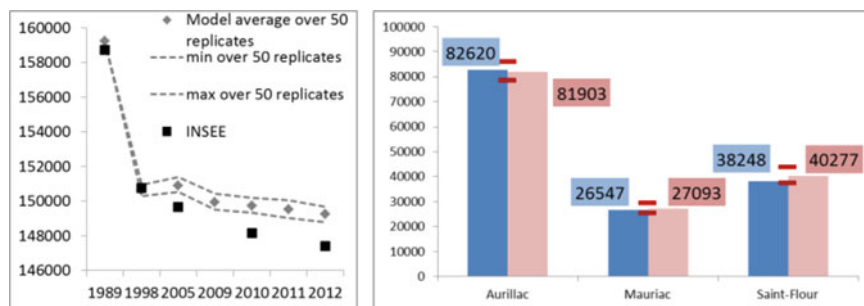
Fig. 5 Population in 2012: (1) on the *left* at the Cantal level; (2) for the three various "arrondisse-ments" composing the Cantal "département" given by INSEE (in *blue*) and the model (on average over 50 replicates in rose); The red lines correspond to the minimum and the maximum results over the 50 replicates

The last challenge we had to deal with was about the small size of our municipalities (on average 440 inhabitants without the two largest). Despite this difficulty impacting the statistical robustness, the results at a global level seem reasonably well reproduced. At a lower level, the canton is a more reliable level to look at the results since it is composed on average of 4500 inhabitants (without the two largest). Overall the model correctly reproduces the spatial heterogeneity of the evolution sign of the municipalities. Moreover, the model seems to keep some relevance to predict the "future." Indeed, while it has not been calibrated from 2006 to 2012, the total 2012 population predicted by the model exhibits 1 % error. On Fig. 5 (1), the reference of INSEE is just below the curves given by the model and we have to remind the change made by INSEE in the counting method making afterwards the starting population 522 less than the one we built from the 1990 Census. Figure 5 (2) shows the population of the three subspaces corresponding to "arrondissement" level. The error for the "arrondissements" are 1, 2 and 5 % on average but the target defined by the INSEE reference is between the minimum and the maximum results over 50 replicates.

Finally our model has coped with many difficulties related to the classical independent views of the mobility. However, its many integrated processes makes necessary a sufficient quality of related process involving a decision to move and a location choice. Reaching this quality is a challenge for each of them. Several processes, for instance the search for a job, remain to validate and detailed data are not always available, especially in rural areas.

References

1. Morand, E., Toulemon, L., Pennec, S., Baggio, R., Billari, F.: Demographic modelling: the state of the art. INED (2010)
2. Birkin, M., Wu, B.: A review of microsimulation and hybrid agent-based approaches. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (eds.) Agent-Based Models of Geographical Systems, pp. 51–68. Springer, New York (2012)
3. Holme, E., Holme, K., Makila, L., Mattson-Kauppi, M., Mortvik, G.: The SVERIGE Spatial Microsimulation Model: Content, Validation and Example Applications. Spatial Modelling Centre, Sweden (2004)
4. Coulombel, N.: Residential choice and household behavior: state of the art. Ecole Normale Supérieure de Cachan (2010)
5. Minodier, C.: Changer de logement dans le même environnement. Données sociales. La Société Française 515–523 (2006)
6. Gobillon, L.: Emploi, logement et mobilité résidentielle. Economie et Statistique 349–350, 77–98 (2001)
7. Debrand, T., Taffin, C.: Les facteurs structurels et conjoncturels de la mobilité résidentielle depuis 20 ans. Economie et Statistique 381–382, 125–146 (2005)
8. Debrand, T., Taffin, C.: Les changements de résidence: entre contraintes familiales et professionnelles. Données sociales. La Société Française 505–513 (2006)
9. Djefal, S.S.E.: Etre propriétaire de sa maison, un rêve largement partagé, quelques risques ressentis. CREDOC **177**, 4 (2004)
10. Huet, S.: Individual based models of social systems: data driven hybrid micro-models of rural development and collective dynamics of filtering or rejecting messages. Computer Science, vol. Doctorate, p. 252. Université Blaise Pascal—Clermont II, Ecole Doctorale Sciences pour l'Ingénieur de Clermont-Ferrand, Irstea (2013)
11. Davezies, L.: L'économie locale "résidentielle". Géographie, économie, société **11**, 47–53 (2009)
12. Blanc, M., Schmitt, B.: Orientation économique et croissance locale de l'emploi dans les bassins de vie des bourgs et petites villes. Economie et Statistique **402**, 57–74 (2007)
13. Perrier-Cornet, P.: La dynamique des espaces ruraux dans la société française: un cadre d'analyse. Territoires **2020**(3), 61–74 (2001)
14. Dubuc, S.: Dynamisme rural: l'effet des petites villes. L'Espace Géographique **1**, 69–85 (2004)
15. Huet, S., Lenormand, M., Deffuant, G., Gargiulo, F.: Parameterisation of individual working dynamics. In: Smajgl, A., Barreteau, O. (eds.) Empirical Agent-Based Modeling: Parameterization Techniques in Social Simulations, vol. 1, The Characterisation and Parameterisation of Empirical Agent-Based Models, p. 22, 133–169. Springer, Heidelberg (2013)
16. Gargiulo, F., Ternès, S., Huet, S., Deffuant, G.: An Iterative Approach for Generating Statistically Realistic Populations of Households. PLoS One **5**, 9 (2010)
17. Lenormand, M., Deffuant, G.: Generating a synthetic population of individuals in households: Sample-free versus sample-based methods. Journal of Artificial Societies and Social Simulation **16**(4), 15 (2012)
18. Huet, S., Lenormand, M., Deffuant, G., Gargiulo, F.: Parameterisation of individual working dynamics. In: Smajgl, A., Barreteau, O. (eds.) Empirical Agent-Based Modeling: Parameterization Techniques in Social Simulations, p. 22. Springer, New York (2012)
19. Bacon, B., Pennec, S.: APPSIM—Modelling Family Formation and Dissolution. University of Canberra (2007)
20. O'Donoghue, C.: Dynamic microsimulation: a methodological survey. Braz. Electron. J. Econ. **4**, 77 (2001)
21. Li, J., O'Donoghue, C.: A Methodological Survey of Dynamic Microsimulation Models. Maastricht University, Maastricht Graduate School of Governance, The Netherlands (2012)
22. Abelson, R.P., Miller, J.C.: Negative persuasion via personal insult. J. Exp. Soc. Psychol. **3**, 321–333 (1967)

23. Duée, M.: La modélisation des comportements démographiques dans le modèle de microsimulation Destinie. INSEE (2005)
24. INSEE, D.R.E.P.S.: Le modèle de simulation dynamique DESTINIE. Série des documents de travail de la Direction des Etudes et Synthèses Economiques **124** (1999)
25. Robert-Bobée, I.: Modelling demographic behaviours in the French microsimulation model Destinie: an analysis of future change in completed fertility. INSEE (2001)
26. INSEE: Enquête sur l'histoire familiale. Data distributed by the Centre Maurice Halbwachs, 48 bd Jourdan, Paris (1999)
27. Blanchet, D., Buffeteau, S., Crenner, E., Le Minez, S.: Le modèle de microsimulation Destinie 2: principales caractéristiques et premeirs résultats. Economie et Statistique 441–442, 101–122 (2011)
28. Blanchet, D., Chanut, J.M.: Projeter les retraites à long terme. Résultats d'un modèle de microsimulation. Economie et Statistique, **315** (1998)
29. Poubelle, V.: Prisme, le modèle de la Cnav. Retraite et Société **48**, 202–215 (2006)
30. Vanderschelden, M.: Les ruptures d'unions: plus fréquentes mais pas plus précoces. Insee Première novembre, **4** (2006)
31. Kennedy, J., Eberhart, R.: Particle swarm optimization. IEEE 0-7803-2768-3/95, 1942–1948 (1995)
32. Poli, R., Kennedy, J., Blackwell, T.: Particle swarm optimization. An overview. Swarm Intell. 33–57 (2007)
33. Clerc, M., Kennedy, J.: The particle swarm—explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. **6**, 59–73 (2002)