Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

A META-ALGORITHM FOR VALIDATING AGENT-BASED SIMULATION MODELS TO SUPPORT DECISION MAKING

Warren Volk-Makarewicz
Catherine Cleophas

Research Area Operations Research and Management
RWTH Aachen University
Kackertstraße 7
Aachen 52072, GERMANY

ABSTRACT

By explicitly modeling the decision making of heterogeneous individuals, agent-based models can compute the resulting emergent phenomena on the micro-level. This lets planners evaluate new planning approaches for problems that depend on individual decisions. Examples include airline revenue management or traffic control. However, when decision support relies on agent-based modeling, its applicability to real-world problems depends on the model's validity. This paper introduces a novel methodological concept to decompose agent-based models for calibration and validation. This concept enables modelers to isolate agents from the evolution of the model's state variables, allowing greater choice of calibration and validation approaches. The approach first parameterizes and validates individual agents, and subsequently re-calibrates the agent-collective within the entire model.

1 INTRODUCTION

According to Macal and North (2014), agent-based modeling is a framework for simulating dynamic processes that involve autonomous agents. As agent-based models can effectively scale and incorporate detail they support the evaluation of planning solutions for complex systems. Both constructing agents and modeling their interaction are relevant for that purpose. Agents decide based on multiple criteria and different degrees of information. Their interactions combine the effect of individual decisions to compute emergent new system states. The idea that individual decisions and aggregate actions link microscopic actions to macroeconomic concepts is a frequently cited benefit of agent-based models (Macal 2016). Formulating agent-based models through separate modeling of agents and their interaction more intuitively represents dynamical behavior in the underlying system. Often, simulation is the only approach to reproduce this behavior in computer experiments, (Axelrod 1997).

However, modeling the decisions of agents and their interaction, i.e. through a utility function or an optimization program, account for only a part of the model. Another significant aspect is the data incorporated into the model. Data connects the model to the real-world system it represents. Such data is used to parameterize agents, their interactions, and environment to the system. For agents that adapt their decisions to observations, data provides the foundation for their actions. Where empirical observations are missing, expert knowledge, from managers, practitioners, and end-users has to serve as a stand-in.

Here, we focus on methodological considerations of validating and calibrating agent-based models. Following Friedman, Hastie, and Tibshirani (2009), any model that aims to provide actionable decision support needs to be validated based on empirical data of the real-world system. For agent-based models Midgley, Marks, and Kunchamwar (2007) point out that validation has to be conducted on at least two levels: At the agent-level to match model parameters with empirical data for individual agents as well as the match of model responses to the empirical data for the complete model. The authors define the first concept as *input validation*, the second as *output validation*. For both types of validation, hypothesis testing can ensure that the model is a faithful representation of the system. We denote the process of parameterizing the system to enable such validation as *calibration*.

The necessity of validating agent-based models is widely acknowledged. However, the current state of research lacks clear methodological approaches to building validated models, as existing approaches consider agents as only a collection of state variables and their calibration and validation is collectively conducted. In addition, most contributions that use agent-based models to evaluate planning approaches do not focus on their validation. This motivates us to propose a meta-algorithm to calibrate and validate agent-based models by isolating the agents from the model's state variables. Employing a meta-algorithm enables a greater range of approaches and scale to model individual agents and the environment and a greater range of methods for calibrating and validating individual modeling components. Furthermore, it lets validation assure that a valid agent-based model consists of valid representations of individual components.

Before considering related research in a brief literature review, we consider the application of airline revenue management in which we illustrate the meta-algorithm. In the next section, the proposed meta-algorithm is presented. In the final section, we highlight relevant benefits and challenges of the meta-algorithm.

1.1 Example: Airline Revenue Management

We illustrate agent-based simulation modeling and their calibration and validation through airline revenue management. In this example, an airline agent interacts with a large number of customer agents in pursuit of divergent interests. The airline optimizes inventory controls to maximize revenue given customer choice behavior.

Revenue management systems compute optimal offers for a given demand forecast, which in turn is computed from historical bookings and offers. The success of revenue management depends on dynamic markets, which are influenced by heterogeneous groups of competitors and customers. As argued in (Marks 2007), such dynamic markets are a subject well-suited for agent-based modeling.

As pointed out, for example in (Currie and Rowley 2010), the success of revenue management relies on both automated processes and human analysts. These analysts and their managers require decision support to predict the possible consequences of their interventions. This has motivated the design of several airline revenue management simulations. Frank, Friedemann, and Schröder (2008) even offer a set of guidelines for creating revenue management simulations. Predominantly, such simulations as described by (Carrier and Weatherford 2015) and (Abdelghany and Abdelghany 2008) are used to test the success of automated revenue management algorithms.

A revenue management approach as described in (Vulcano, van Ryzin, and Ratliff 2012) can incorporate airlines and customers as agents. Customers can flexibly decide whether or not to buy the tickets offered and whether and when to cancel. Customer choice can depend on a range of factors, including the offered price, competitors offers, and brand loyalty. Individual customers are generated stochastically from a preference function and a set of distributions describing demand segments. All customers share a joint behavioral model and differ in the parameters underlying this model.

Airlines, on the other hand, offer a set of supply that can be generated from empirical data, for instance on flight schedules and fleet assignments. Their decision making follows the process of a revenue management system, which includes algorithms that can be implemented in the simulation. The interaction of agents depends on another set of parameters, for instance describing the number of requests that arrive over the sales horizon per customer type.

1.2 Related Research on Building Validated Agent-Based Models

Macal and North (2014) and (Macal 2016) provide introductory tutorials to agent-based modeling. The authors focus on the agents themselves and their properties, the relationships, and environment. The concept presented here utilizes this delineation in thought to propose an approach to decompose agent-based models for calibration and validation. Macal and North (2014) also discuss the process of constructing agent-based models, focusing on the scope of the model, the design, and data considerations.

A variety of review articles describes existing examples of agent-based models in different fields. Examples include Hommes (2006) for economics and finance, Kulasekera et al. (2011) in the area of energy microgrids, Shen et al. (2006) for intelligent manufacturing, Bazzan and Klügl (2014) for transportation, Negahban and Yilmaz (2014) for marketing, and Kiesling et al. (2012) for innovation diffusion. An early, seminal review article of the benefits of simulation, and the consideration it requires, focusing on agent-based modeling is (Axelrod 1997).

The methodological literature on calibrating agent-based models is relatively recent but growing. Several methodological reviews consider how to formulate, construct, and calibrate these models. Furthermore, modeling and estimation approaches from other fields can be applied to agent-based modeling.

Existing methodological reviews within agent-based simulation modeling focus on conceptual development, describing the process, highlight on important questions, and discuss key aspects. Two relevant examples are Gilbert (2008) and by Rand and Rust (2011). Both papers emphasize the need for appropriate modeling mechanisms and for ensuring valid inputs and of validating outputs through data. The second paper also focus on model validation. In (Midgley, Marks, and Kunchamwar 2007), the authors use a supermarket example to discuss various methodological challenges in the validation of agent-based models. Other than noting the need of validation at multiple levels, the authors recommend the design of models in view of improving transparency.

In addition, methodological approaches to agent-based modeling can be found in the domains of economics and finance. Such efforts represent agents via variables within structural equation models in which they calibrate and validate the entire model. For instance, the articles (Grazzini and Richiardi 2015) and (Grazzini, Richiardi, and Tsionas 2017) utilize a state-space representation, calibrating the same structural validation model by minimizing simulated and empirical output via statistical moments and respectively through a Bayesian approach. Lamperti (2016) utilize the Jensen-Shannon metric, a generalization of the Kullback-Leibler divergence, to minimize the distance of probability distributions between simulated and empirical data.

Looking further, metamodel approaches, for instance via polynomial regression, (Kleijnen 2008), or by kriging, (Kleijnen 2009), (Ankenman, Nelson, and Staum 2010), can calibrate agent-based models by a directed search on the graph formed by the parameters and objective function. This is achieved through successive interpolations of function valuations, where each valuation is conducted by simulation at a parameter vector. Validation of the model is then performed on the optimized choice of parameters. By isolating the agents and the environment from the evolution of the state variables, a hybrid modeling approach enables the application of existing calibration and validation methods to agent-based modeling. Such methods can stem from machine learning, (Friedman, Hastie, and Tibshirani 2009), non-linear estimation, (Motulsky and Christopoulos 2004), or from metaheuristics, (Simon 2013).

2 META-ALGORITHM

Agent-based models can be calibrated entirely by iterative adjustment of parameters from separately modeled agents and the environment where they mutually interact.

We propose a meta-algorithm via the stochastic hill-climbing algorithm, Section 4.1.1 in Russell and Norvig (2010), to calibrate agent-based simulation models with N types of heterogeneous agents. In this approach, we assume that each agent and the environment has a parameter vector that needs estimation. This approach isolates the agents and the environment from the state variables. This simplifies the validation problem, by treating it as separate sub-problems. This simple local search algorithm only accepts a proposed solution if it is strictly better than the incumbent. The transition function combined with multiple runs enable the algorithm to find improved solutions; it attempts this in searching randomly in different 'coordinates'. Our coordinates are the different types of agent and the environment.

This approach assumes that agents' behavior follows an objective function, for which an optimization algorithm calibrates parameters. Each agent's objective function utilizes a single expression, which may incorporate multiple criteria. The meta-algorithm validates in isolation each agent and, indirectly via the agent-based model, the environment via a hypothesis test or other statistical comparison. Additionally for each agent component this approach obtains a set of parameter vectors, which can be considered to validate the agent, by inverting the validation procedure. The set of valid parameter vectors can be governed by a constraint, which assists the subsequent calibration of the whole model.

The next section illustrates the proposed meta-algorithm. In this exposition, the agents' objective function and the agent-based model are calibrated by minimizing sum-of-squares error. To solve the optimization problems, we employ the NSGA-II evolutionary algorithm, (Deb et al. 2002). The extra sum-of-squares F-test, (Motulsky and Christopoulos 2004), serves as the means to validate each simulation model component and to acquire the validity constraint for each agent.

To highlight our approach, we utilize concepts from the airline revenue management application, Section 1.1. Here, the agents can represent business and holiday customers as well as the airline. Examples of agent parameters are coefficients for seat availability and demand elasticity. Example environment parameters could include the arrival rates of potential customers.

The presented meta-algorithm is a general approach to calibrate and validate agent-based simulation models. The modeler is not restricted to this choice of methods and can use their own set of methods to calibrate and validate the modeling agents and environment. In the example discussed here, calibration and validation are only conducted once. However, the meta-algorithm can be applied in an iterative fashion at the beginning of each time-period within the model given additional data. In the remainder of the section, Section 2.1 describes each of the components, providing insight into the approach. We present the meta-algorithm and elaborate via the airline revenue management example in Section 2.2.

2.1 Methodological Components

In this section, we describe in turn the components that form the meta-algorithm approach to calibrate and validate the agent-based simulation models.

2.1.1 Terminology

For each of the N heterogeneous agents, let (A_i), $i \in \{1,\ldots,N\}$, be the model representing each agent. The input data that is used for each agent are $x_i = (x_{i,j})$, $j \in \{1,\ldots,n\}$. The corresponding output data is denoted by $y_i = (y_{i,j})$. The equal number of datum for all modeling components is chosen for exposition. The agent-based model evolves over a number of discrete time steps, $t = 1,\ldots,T$. The parameter vectors to be estimated are denoted by (θ_i). Hence the model output of an agent can be depicted, for $i \in \{1,\ldots,N\}$, $j \in \{1,\ldots,m\}$, by $A_i(x_{i,j},\theta_i)$ which is calibrated against empirical output.

We represent by M the function of the agent-based model. In this model, the input data for the environment is denoted by $x_{N+1} = (x_{N+1,j})$, $j \in \{1,\ldots,n\}$. The environment parameter vector is denoted by θ_{N+1}. The representation of the model output for the multi-agent model is a function of the agent models and the environment they interact in, that is, $M := M(A_1,\ldots,A_N,x_{N+1,j},\theta_{N+1})$. The corresponding empirical output data is given by $y_{N+1} = (y_{N+1,j})$. For example, this includes the bookings per time period and the price of the ticket per fare class for both types of customers.

2.1.2 Objective Function

The objective function compares model output data to empirical data including relevant estimation considerations. For an agent model component, here A_1, with parameter vector θ_1 existing in a convex domain Θ_1, the sum-of-squares objective function that is optimized is given by

$$\mathrm{Opt}_1 = \min_{\theta_1 \in \Theta_1} \sum_{j=1}^{n} (y_{1,j} - A_1(x_{1,j}, \theta_1))^2. \tag{1}$$

The meta-algorithm initially calibrates the agents in isolation. In the next stage, the meta-algorithm calibrates the agent-based model M. Here, it optimizes firstly with regard to the environment parameter θ_{N+1}, while keeping the agent parameter vectors fixed. Then, the meta-algorithm iteratively optimizes with regard to a parameter vector θ_i, $i \in \mathscr{I}_{-(N+1)} = \{1, \ldots, N\}$, selected according to a discrete uniform distribution, analogously while keeping all other parameter vectors fixed. The set $\mathscr{I}$ denotes entire index set of modeling components for which optimization has not yet terminated; the set $\mathscr{I}_{-(N+1)}$ is similarly defined, excluding the environment index.

Suppose m iterations of the hill-climbing algorithm has occurred, with $m = \sum_{i=1}^{N+1} m_i$, where the parameter vector θ_i has been re-estimated m_i times, $i \in \mathscr{I}$. For the optimization of the agent-based model with regard to (w.r.t.) the environment θ_{N+1} in convex domain Θ_{N+1}, the representation is, suppressing notation

$$\mathrm{Opt}_{N+1|N+1}(\theta_1^{(m_1)}, \ldots, \theta_N^{(m_N)}) = \min_{\theta_{N+1} \in \Theta_{N+1}} \sum_{j=1}^{n} (y_{N+1,j} - M(A_1, \ldots, A_N, x_{N+1,j}, \theta_{N+1}))^2, \tag{2}$$

Let Θ_i^v to be the convex set that results from the parameter vectors within Θ_i as considered valid via a statistical comparison. Alternatively, when optimizing w.r.t. agent parameter vector, i.e. θ_1, the related optimization is conducted in Θ_1^v, which has the expression

$$\mathrm{Opt}_{N+1|1}(\theta_2^{(m_2)}, \ldots, \theta_{N+1}^{(m_{N+1})}) = \min_{\theta_1 \in \Theta_1^v} \sum_{j=1}^{n} \left(y_{N+1,j} - M\left(A_1(\theta_1), \ldots, A_N, x_{N+1,j}, \theta_{N+1}^{(m_N)}\right)\right)^2. \tag{3}$$

Rather than a sum-of-squares objective function, the objective function can give weight to different considerations. For example, the mean absolute deviation function emphasizes less on fitting all data points equally. In the LASSO regression, (Friedman, Hastie, and Tibshirani 2009), emphasis is placed on preventing extremely sensitive estimates of parameters. Alternatively, the likelihood function can be maximized, fitting the parameters of hypothesized input distributions within the model to data.

2.1.3 Optimization Algorithm

In this contribution, an evolutionary algorithm exemplifies the variety of models that can be calibrated via our approach. Only function evaluations are needed from a model to find an optimal solution. In an evolutionary algorithm, the parameter space is searched via a population of bit-strings, encoding parameter vector evaluations. Better solutions are kept after each iteration via comparison or combination of two bit-strings, or modifying bit-strings with random number generations. A textbook in evolutionary optimization is (Simon 2013). However, if a model is differentiable, then numerical techniques such a Singular Value Decomposition can be used, see (Press et al. 2013), to minimize sum-of-squares error. For linear, normally distributed models, the Kalman filter can maximize the objective function.

Given a population size P, the NSGA-II evolutionary algorithm can optimize a single objective function in $O(P^2)$ operations. Suppose that a parameter vector consists of k elements. From the paper by Alander (1992), the optimal choice of population size is $O(log_2(p))$ where p is the number of combinations a bit-string can represent. Given that double precision has a binary representation of 64 bits, a bit-string would, for example, be of length $64k$ encoding $p = 2^{64k}$ combinations. In terms of the number of parameters to estimate, the NSGA-II algorithm can optimize a single objective function in $O(k^2)$ operations.

2.1.4 Statistical Comparison

In our discussion, the optimization problem for each model component depicts a non-linear regression. The extra sum-of-squares F-test is a hypothesis test to compares nested non-linear models that minimize sum-of-square error. This test is equivalent to the F-test in comparing linear models. To simplify, we also label this test by the same name. In validating the modeling components we set the null hypothesis, H_0, to be a constant value. For agent one, this is estimated by the sample average $\bar{y_1}$. The alternative hypothesis, H_1, is the agent model with estimated parameter vector $\theta_1^{(1)}$. In context, a proposed agent model can be the probability of a customer purchasing a ticket based on price and seat availability. The null model could be estimated probability of purchasing a ticket from historical data.

In this test, we compare the sum-of-squared residual error, SSR, between the output data and the model output. With s^2 denoting the sample variance under the null hypothesis, $SSR(H_0) = (n-1)s^2$. For the alternative hypothesis, this value is

$$SSR\big(H_1;\theta_1^{(1)}\big) = \sum_{j=1}^{n} \big(y_{1,j} - A_1(x_{1,j},\theta_1^{(1)})\big)^2. \tag{4}$$

The test statistic F is given by

$$F = \frac{n-k_1}{k_1-1} \cdot \frac{SSR(H_0) - SSR\big(H_1,\theta_1^{(1)}\big)}{SSR\big(H_1,\theta_1^{(1)}\big)}. \tag{5}$$

Under the null hypothesis, this test statistic is F-distributed with $k_1 - 1$ degrees of freedom for the numerator and $n - k_1$ degrees of freedom for the denominator. We reject the null hypothesis for large values of F. From Seber and Wild (1989) this approach is robust to errors that are not normally distributed. From (Gallant 1975) the power of this test can be obtained from Equation (5), by computing probabilities larger than the test statistic F. Given a parameter $\lambda \geq 0$, which is a measure of model non-linearity, the statistic F under H_1 is a non-central F-distribution with parameter λ with $k_1 - 1$, $n - k_1$ degrees of freedom for the numerator and denominator. The paper (Gallant 1975) provides the definition and its estimation of the parameter λ.

2.1.5 Optimization Constraint of Agents

Other than ensuring reasonableness of the modeling components, assuming that the statistical comparison can be inverted, we can use the confidence interval or valid parameter set as a constraint in the iterative optimization of the agent-based model when the optimization is conducted w.r.t. an agent parameter vector. For example, the F-test can be inverted to obtain an upper bound. Given a probability level $\gamma \in (0,1)$, we denote for each modeling component, $i \in \{1,\ldots,N\}$ by $f_{\gamma,k_i-1,n-k_i}$ as the γ-quantile of the F-distribution with degrees of freedom $n - k_i$ and $k_i - 1$ for the numerator and denominator.

At a significance level α, denoting Type I error, from Equation (4), we reject the null hypothesis for an agent model, for instance A_1, if $F \geq f_{1-\alpha,k_1-1,n-k_1}$. Then from Equation (1), this corresponds to upper bound, v_1, for the sum of square residual error $SSR(H_1;\theta_1^{(1)})$. With s^2 denoting the sample variance of the output data y_1, our constraint is

$$SSR\big(H_1;\theta_1^{(1)}\big) \leq v_1 \;:=\; \frac{(n-1)(n-k_1)}{k_1-1} \cdot \frac{s^2(y_1)}{f_{1-\alpha,k_1-1,n-k_1}+1}. \tag{6}$$

The comparison of models via the Bayesian Information Criterion, from (Friedman, Hastie, and Tibshirani 2009), can also be utilized as a constraint since the criterion value of a new model against an incumbent is diminished with sub-optimal estimates of the parameter vector.

In our example, evolutionary algorithms do not specifically restrict the search space due to the presence of constraints. From Deb (2000), a preference relation is used to compare solutions of bit-strings via penalization of the objective function.

2.1.6 Stochastic Hill-Climbing Algorithm

The hill-climbing algorithm, (Russell and Norvig 2010), historically has been associated with the traveling salesman problem, (Steiglitz, Weiner, and Kleitman 1969) or in incorporation with genetic algorithms, (Mühlenbein 1993). In optimizing w.r.t. an agent or the environment at each iteration, we propose to utilize a discrete uniform transition function. This is due to the simplicity of this variant and because it avoids unnecessary optimization steps. This contrasts with another common approach of weighting probabilities according to the improvement in function value, requiring the computation of multiple optimizations at each iteration.

Let $\mathscr{I}_{-i}$ be the index set of remaining parameter vectors to be optimized after the current iteration, $i \in \{1,\ldots,N\}$, and let $|\mathscr{I}_{-i}|$ be the number of elements in this set. The transition function probabilities, for $l \in \mathscr{I}_{-i}$ are then $p_{i,l} = 1/|\mathscr{I}_{-i}|$. In the meta-algorithm, the initial calibration of agents and environment yields the parameter vectors $(\theta_1^{(1)},\ldots,\theta_{N+1}^{(1)})$. We propose to begin each run in optimizing the agent-based model at these estimated vectors, as it provides a good initial value. Commonly for the hill-climbing algorithm a randomly chosen initial value is used. As an option, each run can begin at a different value within the valid parameter set of the agent-based model, denoted by the Cartesian product $(\times_{i=1}^{N}\Theta_i^{\nu}) \times \Theta_{N+1}$.

We define the termination condition per modeling component to be the number of successive instances, denoted by *maxTol*, of non- or minimal improvement of the agent-based model objective function. The counter for each modeling component is denoted by s_i. When the threshold occurs for a modeling component, the component is removed from the index set $\mathscr{I}$.

2.2 Algorithm

Given the description of the different parts of the meta-algorithm in Section 2.1, the stochastic hill-climbing approach consists of the following steps:

1. Estimate agent parameter vectors in isolation, optimizing their objective function, obtaining parameter vector estimates similar to Equation (1).
2. Validate each agent model via a statistical comparison from which an optimization constraint can be obtained. If the model is valid, determine the optimization constraint associated with the statistical comparison, similar to Equation (6).
3. Optimize the objective function of the agent-based model w.r.t. the environmental parameter vector, keeping all agent-based parameter vectors fixed, similar to Equation (2).

For each $r = \{1,\ldots,nRuns\}$

4. Set $\mathscr{I} = \{1,\ldots,N+1\}$ of available parameter vectors yet to be optimized. Repeat the following four steps until the tolerance requirement is met for each agent and the environment, i.e. $\mathscr{I} = \{\emptyset\}$
   (a) Choose an index $\mathscr{I}_{-i}$ of remaining unoptimized parameter vectors other than i via a discrete uniform distribution, where i is the last modeling component that was recalibrated.
   (b) Let l be the chosen index. If l designates an agent optimize the agent-based model w.r.t the parameter vector related to the agent similar to Equation (2). If $l = N+1$, optimize the parameter vector w.r.t. the environment similar to Equation (3).
   (c) If the value of the objective function improved, record this value and the new value of θ_l.
   (d) Ascertain whether the improvement is within a certain tolerance criterion. If this is the case, $s_i = s_i + 1$. If $s_i = tolMax$ times, remove the parameter vector from $\mathscr{I}$ and the agent-based objective function is considered to be optimized w.r.t. the parameter vector. Otherwise, reset counter s_i.
5. From the r runs choose the value of the objective function with the optimized value and the associated value of the parameter vectors $(\theta_i), i = \{1,\ldots,N\}$. Test the validity of the the agent-based model from the chosen run according to the statistical comparison, similar to Equation (5).

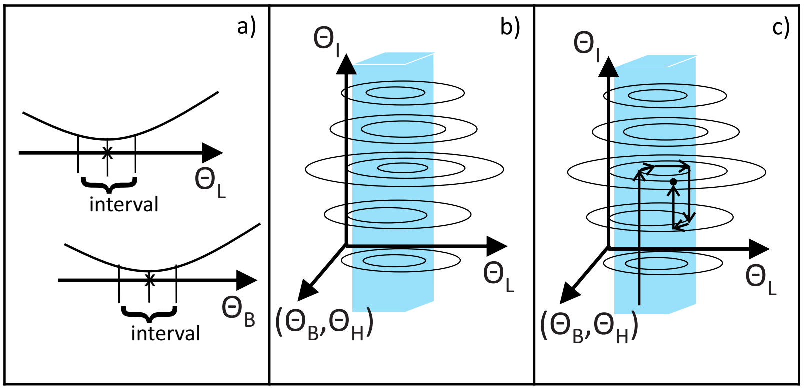

Figure 1: The stochastic hill-climbing algorithm as presented via the airline revenue management example.

2.3 Example: Airline Revenue Management

We illustrate the proposed approach on an airline revenue management example based on models described in Vulcano, van Ryzin, and Ratliff (2012) and Bartke, Kliewer, and Cleophas (2017). An airline determines revenue-maximal inventory controls for a flight consisting of $S = 100$ seats and $C = 5$ fare classes. The flight departs six months in the future. The booking horizon consists of $T = 5$ time slices of different lengths. During the booking horizon, business and holiday customers can purchase individual tickets. The agents are the two customer types, who choose from amenable fare classes, and the airline. The environment is described by the number of booked and unbooked seats on the aircraft and by the arrival rates for both types of potential customers.

At each time slice of the booking horizon, the airline determines the number of seats that are available to be sold for each fare class and its associated price. This calculation relies on historical transaction data, such as the number of bookings per time slice and fare class. The airline does not know the exact number of potential customers. If, to optimize inventory controls, the airline employs a dynamic program, exemplary estimated parameters per time slice, denoted as θ_L, would include coefficients w.r.t. the ticket price and seat availability. Estimated parameters of customers' from the business segment, θ_B, and from the holiday segment, θ_H, may describe customers' utilities per fare class or their price elasticity. The same transaction data is also used to calibrate both types of customer.

Some environment parameters, such as aircraft capacities, can be considered as fixed. Environment parameters that have to estimated, denoted by θ_P, may include the arrival rates for both types of customers for the five periods. The overall model incorporates the arrival rates together with the airline optimization model and the customers' method of determining which ticket to purchase. The objective function of the agent-based model is to match the bookings obtained per time slice of the booking horizon and per fare class to the empirical data.

Following the steps in the meta-algorithm, Step 1 begins to calibrate the model by calibrating the three types of agents, business customers, holiday customers, and the airline. It obtains parameter estimates $\theta_B^{(1)}, \theta_H^{(1)}, \theta_L^{(1)}$ from the airline transaction data using the sum-of-squares objective function as described in Equation (1). The NSGA-II evolutionary algorithm applies to all agents. The evolutionary algorithm's tolerance level is that the fittest solution between consecutive iterations must be no more than 10^{-6}. When price and seat availability coefficients are only used in addition to the number of time slices per fare class, the number of parameters to estimate is $k_B = k_H = 35$ per customer type. When only using price and seat availability data for the dynamic program, the number of parameters to estimate for the airline model is $k_L = 50$. As in Deb (2000), we set the population size for each modeling to be $10k$.

Let the validation rely on the F-test in Step 2. This step sets the significance level of this test to $\alpha = 0.10$. It computes sum-of-square residuals and the F statistic according to Equations (4) and (5).

Frame a) in Figure 1 illustrates the validation consideration for the business customer and airline. Then from Equation (6) the validation constraints, $v_i, i \in \{H, B, L\}$, are obtained for all agents and ensuring the agents remain valid during the calibration of the agent-based model. The associated new validity domains for the parameter vectors are Θ_i^v.

At this stage, both customer types and the airline are calibrated and validated with estimates $\theta_i^{(1)}, i \in \{H, B, L\}$. The next task is to estimate arrival rates as environmental parameters per time slice and customer type, $k_P = 10$, to obtain an initial operating simulation model. Frame b) is a visualization of the level-sets of the optimization function for the agent-based model with both customers along one axis. The prism denotes the valid convex set in which agents remain valid. In Step 3, Equation (2), obtains the environment parameter vector estimate $\theta_P^{(1)}$. The population size for the genetic algorithm is again $10k_P$.

Step 4 recalibrates the revenue management simulation model by iteratively improving parameter vector estimates. The stochastic hill-climbing algorithm will only accept improvements if the value of the sum-of-squares objective function, Equations (2) and (3), is reduced. Frame c) in Figure 1 depicts the algorithm for determining a minima, denoted by the bullet. For example, we set $nRuns = 100$ runs of the model and set $tolMax = 6$ in which the tolerance level is within 10^{-6} for which modeling components are considered optimized. In Step 5, given a significance level $\alpha = 0.10$ we observe a valid agent-based model with our minimal function value.

If the entire revenue management simulation representing the agent-based model was optimized as a whole via the NSGA-II evolutionary algorithm, this approach would take $O((k_B + k_H + k_L + k_P)^2) = O(14400)$ operations to obtain parameter estimates. By taking advantage of the structure of the agent-based model, this method would take $O(k_B^2 + k_H^2 + k_L^2 + k_P^2) = O(5350)$ operations for calibration and validation via the meta-algorithm.

In the computation of the power we assume $n = 800$ data points to estimate the parameter vectors and agents and estimates of $\lambda = 20$ for the non-linearity of all modeling components. In addition the value of test statistic is $F = f_{0.90,k-1,n-k}$, the 0.90-quantile for each modeling component. The probabilities for the power are 0.7495 for types of customers, 0.6623 for the airline, and 0.4207 for the agent-based model. For the agent-based model, the computation includes all parameters from all modeling components. In this example, an increase in the power can occur by increasing the number of data points in obtaining estimates, by an increase in the non-centrality parameter, or a decrease in the number of parameters. The last two considerations relate to model design and development, which is not easily controllable.

3 CONCLUSION: BENEFITS AND CHALLENGES

This contribution presented a new approach to calibrate and validate agent-based models via decomposition. Valid models are a prerequisite to using agent-based models for business decision support, as illustrated through airline revenue management. Before concluding this paper with an outlook to future research, this section critically considers the benefits and challenges of the proposed concept.

3.1 Benefits

First, we recapture and highlight several benefits of the proposed approach.

Free choice of benchmarking models: The concept lets modelers select a comparative model to separately ascertain whether the effect of calibrating the agents or their interactions cause an improvement in validity. Separately calibrating the different types of agents and the environment, the mechanism governing their interaction, permits creating new agents and environments for comparison. This also motivates to the use of design patterns for constructing different components of the agent-based model, as recommended by North and Macal (2014).

Reduced dimensionality of optimization model: Via the decomposition of agents and their interactions, the concept lets modelers face multiple smaller optimization problems multiple times instead of asking them to solve a single, possibly very large problem. Following the stochastic hill-climbing algorithm, the resulting sub-optimization problems for the agent or environment modeling components include fewer parameters to be estimated. This benefit seems particularly promising for reducing the computational cost of calibrating large agent-based models, which may contain a multitude of parameters.

Enable model averaging: Model averaging, as described in Chapter 8 of Friedman, Hastie, and Tibshirani (2009), lets multiple models jointly represent a system. A weighted average merges the models' outputs to improve the fit of the results to empirical observations. The proposed meta-algorithm enables researchers to apply model averaging to single agent types or to the mechanism governing the environment.

Highlight structural model invalidity: If the agent-based model is structurally invalid, the mechanism governing the agents' interactions in the environment is inconsistent with the agent models. The meta-algorithm first tests the validity of the included agents based on empirical data. If the agents are judged valid, but subsequently the agent-based model is judged invalid, then the environment representing agent interactions, i.e. the choice of arrival process within the airline revenue management example, can be considered to be invalid. In consequence, further structural improvements are required to improve validity.

3.2 Challenges

While we clearly consider the concept proposed here as promising for calibrating and validating agent-based models, it leaves some challenges to be considered. In particular, we highlight the question of how to separate agents and their environment and the task of optimizing the (sub-)models to fit empirical data.

How to decompose agents and the environment: This is a system-dependent concern. There may be a natural choice of agent representation or multiple representations to attribute features to agent modeling components. In our example, a simpler characterization, removing the need to model agents, would be to re-interpret the arrival rate at each time-step as the rate, respectively, of business and holiday customers willing to purchase a plane ticket. These arrival parameters can be calibrated from the price and seat availability for each fare class made by the airline.

A possible general-purpose way to handle interactions amongst agents is the contextual bandit, for example (Bastani and Bayati 2015), which learns online by combining model improvement of regression estimates and the optimization of decisions.

Controlling Type I and Type II error: Following (Lehmann and Romano 2006), type I and II errors should be balanced to ensure high probabilities of correct identification under both hypotheses. This is complicated for agent-based model validation, which likely includes a large number of parameters, combining agent and environment modeling components. Increasing the number of data samples may reduce both sources of error. Additionally, improving parameter identification will also reduce both sources of error by reducing the number of parameters to be estimated. Canova and Sala (2009) provide approaches to this end.

3.3 Future Research

Having laid down the basics of the proposed concept in this paper and illustrated them with relevant examples, future research must provide working implementations. To that end, we aim to create a framework for calibration and validation that follows the described meta-algorithm. By classifying simulation models and their data prerequisites and implementing exemplary studies, the aim is to develop guidelines for implementing the concept in a variety of contexts.

REFERENCES

Abdelghany, A., and K. Abdelghany. 2008. "A Micro-Simulation Approach for Airline Competition Analysis and Demand Modelling". *International Journal of Revenue Management* 2 (3): 287–306.

Alander, J. T. 1992. "On Optimal Population Size of Genetic Algorithms". In *CompEuro 1992 Proceedings: Computer Systems and Software Engineering*, 65–70. IEEE.

Ankenman, B., B. L. Nelson, and J. Staum. 2010. "Stochastic Kriging for Simulation Metamodeling". *Operations research* 58 (2): 371–382.

Axelrod, R. 1997. "Advancing the Art of Simulation in the Social Sciences". In *Simulating Social Phenomena*, edited by R. Conte, R. Hegselmann, and P. Terna, 21–40. Berlin, Germany: Springer.

Bartke, P., N. Kliewer, and C. Cleophas. 2017. "Benchmarking Filter-Based Demand Estimates for Airline Revenue Management". *EURO Journal on Transportation and Logistics*. Available at doi:10.1007/s13676-017-0109-4.

Bastani, H., and M. Bayati. 2015. "Online Decision-Making with High-Dimensional Covariates". LEM Working Paper Series 2016/18. Available via http://ssrn.com/abstract=2807557.

Bazzan, A. L. C., and F. Klügl. 2014. "A Review on Agent-Based Technology for Traffic and Transportation". *The Knowledge Engineering Review* 29 (3): 375–403.

Canova, F., and L. Sala. 2009. "Back to Square One: Identification Issues in DSGE Models". *Journal of Monetary Economics* 56 (4): 431–449.

Carrier, E., and L. Weatherford. 2015. "Implementation of MNL Choice Models in the Passenger Origin-Destination Simulator (PODS)". *Journal of Revenue and Pricing Management* 14 (6): 400–407.

Currie, C. S. M., and I. T. Rowley. 2010. "Consumer Behaviour and Sales Forecast Accuracy: What's going on and how should Revenue Managers Respond?". *Journal of Revenue and Pricing Management* 9 (4): 374–376.

Deb, K. 2000. "An Efficient Constraint Handling Method for Genetic Algorithms". *Computer Methods in Applied Mechanics and Engineering* 186 (2): 311–338.

Deb, K., S. Agrawal, A. Pratap, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II". *IEEE Transactions on Evolutionary Computation* 6 (2): 182–197.

Frank, M., M. Friedemann, and A. Schröder. 2008. "Principles for Simulations in Revenue Management". *Journal of Revenue and Pricing Management* 7 (1): 7–16.

Friedman, J., T. Hastie, and R. Tibshirani. 2009. *The Elements of Statistical Learning*, Volume 2. New York, USA: Springer.

Gallant, A. R. 1975. "The Power of the Likelihood Ratio Test of Location in Nonlinear Regression Models". *Journal of the American Statistical Association* 70 (349): 198–203.

Gilbert, N. 2008. *Agent-Based Models*. Thousand Oaks, CA, USA: Sage Publications.

Grazzini, J., and M. Richiardi. 2015. "Estimation of Ergodic Agent-Based Models by Simulated Minimum Distance". *Journal of Economic Dynamics and Control* 51:148–165.

Grazzini, J., M. G. Richiardi, and M. Tsionas. 2017. "Bayesian Estimation of Agent-Based Models". *Journal of Economic Dynamics and Control* 77:26–47.

Hommes, C. H. 2006. "Heterogeneous Agent Models in Economics and Finance". In *Handbook of Computational Economics, Volume 2.*, edited by L. Testfastion and K. L. Judd, Chapter 23, 1109–1186. Amsterdam, The Netherlands: Elsevier.

Kiesling, E., M. Günther, C. Stummer, and L. M. Wakolbinger. 2012. "Agent-Based Simulation of Innovation Diffusion: A Review". *Central European Journal of Operations Research* 20 (2): 183–230.

Kleijnen, J. P. 2009. "Kriging Metamodeling in Simulation: A Review". *European Journal of Operational Research* 192 (3): 707–716.

Kleijnen, J. P. C. 2008. *Design and Analysis of Simulation Experiments*. 1st ed. New York, NY, USA: Springer.

Kulasekera, A., R. Gopura, K. Hemapala, and N. Perera. 2011. "A Review on Multi-Agent Systems in Microgrid Applications". In *Innovative Smart Grid Technologies-India (ISGT India)*. Institute of Electrical and Electronics Engineers, Inc.

Lamperti, F. 2016. "Empirical Validation of Simulated Models through the GSL-div: An Illustrative Application". Available via http://ssrn.com/abstract=2661896.

Lehmann, E. L., and J. P. Romano. 2006. *Testing Statistical Hypotheses*. 3rd ed. New York, NY, USA: Springer.

Macal, C. M. 2016. "Everything you need to know about Agent-Based Modelling and Simulation". *Journal of Simulation* 10 (2): 144–156.

Macal, C. M., and M. North. 2014. "Introductory Tutorial: Agent-Based Modeling and Simulation". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and A. Miller, 6–20. Institute of Electrical and Electronics Engineers, Inc.

Marks, R. E. 2007. "Validating Simulation Models: A General Framework and Four Applied Examples". *Computational Economics* 30 (3): 265–290.

Midgley, D., R. Marks, and D. Kunchamwar. 2007. "Building and Assurance of Agent-Based Models: An Example and Challenge to the Field". *Journal of Business Research* 60 (8): 884–893.

Motulsky, H., and A. Christopoulos. 2004. *Fitting Models to Biological Data using Linear and Nonlinear Regression: a Practical Guide to Curve Fitting*. Oxford, UK: Oxford University Press.

Mühlenbein, H. 1993. "Evolution in Time and Space-The Parallel Genetic Algorithm". In *Foundations of Genetic Algorithms*, edited by L. D. Whiteley, Volume 1st, 316–337. San Mateo, CA, USA: Morgan Kaufman.

Negahban, A., and L. Yilmaz. 2014. "Agent-Based Simulation Applications in Marketing Research: An Integrated Review". *Journal of Simulation* 8 (2): 129–142.

North, M. J., and C. M. Macal. 2014. "Product and Process Patterns for Agent-Based Modelling and Simulation". *Journal of Simulation* 8 (1): 25–36.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 2013. *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. Cambridge, UK: Cambridge University Press.

Rand, W., and R. T. Rust. 2011. "Agent-Based Modeling in Marketing: Guidelines for Rigor". *International Journal of Research in Marketing* 28 (3): 181–193.

Russell, S., and P. Norvig. 2010. *Artifical Intelligence: A Modern Approach*. 3rd ed. New York City, NY, USA: Pearson.

Seber, G. A. F., and C. J. Wild. 1989. *Nonlinear Regression*. Chichester, NY, USA: John Wiley & Sons.

Shen, W., Q. Hao, H. J. Yoon, and D. H. Norrie. 2006. "Applications of Agent-Based Systems in Intelligent Manufacturing: An Updated Review". *Advanced Engineering Informatics* 20 (4): 415–431.

Simon, D. 2013. *Evolutionary Optimization Algorithms*. Hoboken, USA: John Wiley & Sons.

Steiglitz, K., P. Weiner, and D. J. Kleitman. 1969. "The Design of Minimum-Cost Survivable Networks". *IEEE Transactions on Circuit Theory* 16 (4): 455–460.

Vulcano, G., G. van Ryzin, and R. Ratliff. 2012. "Estimating Primary Demand for Substitutable Products from Sales Transaction Data". *Operations Research* 60 (2): 313–334.

ACKNOWLEDGMENT

The research presented in this contribution was funded by the Deutsche Forschungsgemeinschaft, Project ID CL605/4-1.

AUTHOR BIOGRAPHIES

WARREN VOLK-MAKAREWICZ is a Postdoctoral Researcher within the Research Area Operations Research and Management at RWTH University Aachen. He received degrees in mathematics and statistics and actuarial studies from the University of Melbourne and received his PhD degree in operations research at VU University Amsterdam. His research interests lie in agent-based simulation and simulation based optimization, with applications including revenue management and financial engineering. His email address is wm.volk.makarewicz@ada.rwth-aachen.de.

CATHERINE CLEOPHAS is Associate Professor for Advanced Analytics in the Research Area Operations Research and Management at RWTH Aachen University. Her research fields include revenue management, modeling and forecasting the interaction of suppliers and demand, and agent-based, stochastic simulation systems. Her email address is catherine.cleophas@ada.rwth-aachen.de.