Evidence-Driven Computational Modeling[1]

Ravi Bhavnani, Karsten Donnay and Mirko Reul

AGENT-BASED COMPUTATIONAL MODELS

Computational modeling is a powerful, versatile tool for the analysis of complex social phenomena. Historically, scholars used computational modeling to investigate abstract causal relationships in artificial settings, highlighting simple but counter-intuitive dynamics. Seminal examples include work by Thomas Schelling on the drivers of segregation (Schelling, 1971), Robert Axelrod on the evolution of cooperation (Axelrod, 1984), Joshua Epstein and Robert Axtell on artificial societies (Epstein and Axtell, 1996), and Palmer et al. on artificial stock markets (Palmer et al., 1999). These early applications influenced subsequent research, including notable studies on the formation and dissolution of nation-states after the end of the Cold War (Cederman, 1997), the dynamics of ethnic violence and genocide (Bhavnani and Backer, 2000), and more recently civil violence in Baghdad and Jerusalem (Bhavnani et al., 2014; Weidmann and Salehyan, 2013).

In contrast to *consolidative models*, which typically involve the development of 'model' systems to represent 'real-world' settings with measurable physical characteristics (for weather forecasts, see Gneiting and Raftery, 2005; Skamarock and Klemp, 2008), *exploratory* computational models stop short of formalizing the complexity of social systems (Bankes, 1993). Given the difficulty of fully observing, theorizing and validating processes in social and natural systems, our approach builds on work that is exploratory, not consolidative, in nature. One class of exploratory computational models used in the social sciences is agent-based computational modeling (ABM) (for an overview, see de Marchi and Page, 2014).

A key property of ABM is the specification of simple rules from which complex outcomes *emerge*. As such, an ABM may be specified as a non-linear function that relates combinations of inputs and parameters to outcomes. ABMs are typically composed of agents, decision-making heuristics, an interaction topology and a non-agent environment (Epstein, 1999). Agents in an ABM can represent individuals (Bhavnani et al., 2008; Epstein, 2002), groups (Bhavnani et al., 2009; Kollman et al., 1992) or institutions (Cederman, 1997), to name a few possibilities. In this regard, the approach provides a high degree of flexibility or **granularity**, given the ability to integrate phenomena specified at different scales. ABMs are process-oriented and lend themselves well to studying dynamics, in contrast to approaches that tend to be more equilibrium-centered.

In most formulations of ABM, agents are endowed with a range of characteristics and decision-making heuristics. Individual agents may learn or adapt their behavior based on their own experiences, driven by heuristics or imitation, or change may be effected for a population of agents by means of evolutionary selection (Kollman et al., 1992; Laver, 2005; Mitchell, 1996). The interaction topology specifies how agents interact with each other and their environment, the latter being composed of physical features such as geography or topography (Axtell et al., 2002; Epstein, 2002) or various states of the world (Axelrod, 1984; Nowak and May, 1992; Tullock and Campbell, 1970). These elements constitute the key components of an ABM, which is run repeatedly to identify causal mechanisms, observe relationships, patterns and emergent outcomes, and explore counterfactual scenarios.

ABMs lend themselves well to the analysis of complex social phenomenon, in particular where ostensibly simple decisions have unexpected consequences (Epstein, 1999). Yet, while agent-based models have notable strengths, they are not immune to criticism (Richiardi et al., 2006). A notable weakness of ABM is the tendency to include too many factors and interactions, given the ease with which these may be specified. As a rule of thumb, a model becomes too complicated when comprehensive exploration of the comparative statistics for each model parameter is infeasible (see Lustick et al., 2004). Under these circumstances, it is virtually impossible to determine what is driving model results. Yet another flaw is the lack of relevant theoretical and empirical anchors, which result in unrealistic or even arbitrary model specifications. These anchors are essential to address the *identification problem* – the notion that multiple plausible mechanisms may explain a given outcome (Fisher, 1966). This chapter provides practical advice for designing, implementing and using computational models that are evidence-driven and designed to address these shortcomings.

EVIDENCE-DRIVEN COMPUTATIONAL MODELS

The evidence-driven modeling (EDM) framework rests on three methodological pillars: agent-based modeling (ABM), contextualization using geographical information systems (GIS) and empirical validation. EDM harnesses the strengths of ABM in capturing both social complexity (e.g., the heterogeneity of actor beliefs, preferences, attitudes and behaviors as well as the characteristics of specific institutional settings and local environments) and causal complexity (including questions about who interacts with whom, when, where and with what effects), while simultaneously achieving a high degree of real-world correspondence and resonance. The combination places EDM squarely at the intersection of theory and empirics.

Notable examples of EDM include studies about civil violence in Jerusalem and Baghdad (Bhavnani et al., 2014; Weidmann and Salehyan, 2013), neo-patrimonial networks (Geller and Moss, 2008), social inequality in pastoralist societies (Rogers et al., 2015), the rise and fall of the Anasazi people in what is now the Southwestern United States

(Axtell et al., 2002), social capital and civic culture in Italy (Bhavnani, 2003), legislative politics (Laver et al., 2011), party competition (Laver and Sergenti, 2011) and the occurrence of burglary (Malleson and Birkin, 2012). The diversity of research demonstrates the utility of combining computational models with rich empirical data – data that are spatially and temporally disaggregated – to analyze the links between micro-level dynamics and emergent, macro-level outcomes.

Where other, more aggregate analyses yield inconclusive results, EDM enables researchers to adjudicate between alternative explanations and reveal complex, conditional relationships and inter-dependencies that would otherwise be difficult to detect (Camerer, 2003; Kim et al., 2010).[2] In particular, EDM makes it possible to specify causal mechanisms in ways that are sufficiently intricate, conditional and, thus, ultimately realistic, while maintaining the ability to go beyond purely exploratory modeling by means of rigorous empirical testing. And in contrast to experimental approaches that control for contagion and spillover, EDM explicitly incorporates these ostensible threats to validity as part of the causal chain, for example by endogenizing the effects of geographical proximity and the heterogeneity of covariates across spatial units. The more specific benefits of EDM include the following:

1 *Model Topography*: The ability of EDM to harness GIS, in conjunction with empirical data, enables realistic topographies to be substituted for the abstract grids characteristically used in ABM. As such, the landscapes used in EDM more closely represent actual physical or social inter-dependencies, capturing complex, often endogenous relationships among adjacent units, rather than controlling for these relationships statistically or by means of experimental design.
2 *Agent Granularity*: EDM can simultaneously accommodate data of different spatial and temporal resolution, whereas other methodologies are often wedded to the use of specific, fixed units. In contrast to ABM, these different units of analysis correspond to empirical observations and capture dynamics at meaningfully interlinked levels, e.g., individual decision-makers interacting with groups.
3 *Data Imputation*: EDM is typically used in data-rich contexts but also excels in data-poor contexts, where information on relevant indicators suffers from incompleteness, a lack of synchronization, mismatched units of observation, and differing levels of detail. Imputation in EDM works by seeding a model with potentially sparse empirical data, and then permitting model dynamics to evolve endogenously. The closer simulated outcomes are to empirical trends, the better the imputation. The estimation of different parameters across contexts, using the same model, is one way to increase model robustness.
4 *Identification*: As with ABM, EDM can be used to explore relationships between or adjudicate among competing micro-level explanations, relying on methods for data construction, such as participant observation, expert or field interviews. Insights from these methods help ground a model, ensuring that researchers `get the story right' and tailor the model to the specificities of a given context.
5 *Counterfactual Analysis*: Once a model is calibrated and empirically validated, counterfactuals can be devised by adjusting values of certain parameters, including those capturing micro-level dynamics and the empirical context, or by introducing new parameters. The results, produced under an assortment of `what-if' scenarios, offer an indication of what the world could look like if empirically observed trends were to change. In essence, this option enables experimentation through simulation. Short of true out-of-sample forecasts, counterfactual experiments make it possible to undertake evidence-driven forecasting.

In the remainder of this chapter, a step-by-step discussion guides the reader through the use of EDM in the *Modelling Early Risk Indicators to Anticipate Malnutrition (MERIAM)* project. We provide further detail on the building blocks for EDM, as well as on the choices and practical challenges of using the approach. Our discussion is intended to serve as a point of departure for conducting research with EDM.

EVIDENCE-DRIVEN MODELING OF MALNUTRITION

The MERIAM project illustrates how the EDM approach can be applied, from initial conception to final, policy-relevant application. MERIAM is a four-year project funded by the UK government, which brings together an inter-disciplinary team of experts across four consortium partners: Action Against Hunger, the Graduate Institute of International and Development Studies, Johns Hopkins University and the University of Maryland. MERIAM's primary aim is to develop, test and scale up models to improve the prediction and monitoring of undernutrition in countries that experience frequent climate- and conflict-related shocks.[3]

In 2017, the number of undernourished people was estimated at 821 million; this is closely associated with the spread of armed conflict (FAO et al., 2018). Regions across Nigeria, South Sudan, Somalia and Yemen face severe food insecurity, related in no small measure to their exposure to conflict as well as a host of other characteristics that increase vulnerability to famine. The gravity of these situations and high interest among stakeholders serve as inspiration for devising effective means of forecasting risks to better anticipate crises and guide appropriate responses.[4]

The research team at the Graduate Institute is tasked with the development of an EDM to analyze the effect of household-level decisions on nutrition-related outcomes (e.g., acute malnutrition and resilience), accounting for variation in household characteristics; local, contextual factors; and more macro- or aggregate-level covariates. How households adapt their behavior, changing or diversifying sources of household income in response to stressors and shocks, may serve to improve or worsen a household's resilience to food insecurity over time. We develop and validate our EDM based on the case of Karamoja, Uganda, and subsequently expand our model framework to other cases in sub-Saharan Africa.

We use this project – and, in particular, the application of our approach to study malnutrition in Karamoja – as an example of applying cutting edge computational modeling techniques to a highly relevant policy issue. The 'entry-point' for the use of EDM is an abundance of theoretical knowledge on the issue, the complexity of interactions of the numerous factors that influence malnutrition outcomes at the household level and the need for a systematic, reliable and transparent forecasting technique. At present, practitioners and policymakers tasked with anticipating changes in the risk of acute malnutrition need to combine expert knowledge on malnutrition, including a deep understanding of its causes in particular cases, with statistical analysis of data from various sources on a regular basis. A prominent example in the context of sub-Saharan Africa is the Famine Early Warning Systems Network (FEWS NET), which classifies a country's risk of acute food insecurity, relying on expert discussions and analysis mainly of remote-sensing, market price and trade data. The outcome is an indicator for food insecurity on a five-point scale, ranging from 'Minimal' to 'Famine', with 'near'-term and 'medium'-term forecasting windows of up to seven months (see *IPC 2.0*[5]).

There are several potential weaknesses inherent to such an approach. First, the link between data and projected food insecurity lacks formalization (the scenario-building process is interpretive) and transparency (it is unclear how a particular prediction is made). In a related vein, expert discussions underlying the data analysis are undocumented for end-users, making comparisons of forecasts by different experts problematic should interpretations vary.[6] And finally, predictions of food insecurity by livelihood zone obscure the relative weight of various risk factors and their effects at more disaggregated spatial units.

Our EDM approach attempts to address these weaknesses. First, we harness available expertise on food security and malnutrition, relying explicitly on expert surveys, to develop a theoretically grounded computational model. Second, we use existing, household-level data and household surveys conducted in the field to empirically contextualize and validate the model for a set of subnational regions that vary in terms of their incidence and prevalence rates, livelihood zones, climate conditions and history of conflict. Third, we provide a tool that stakeholders can use to construct acute malnutrition scenarios across diverse contexts, exploring the relation between shocks and stressors, on the one hand, and more immediate and long-term outcomes on the other. In the section that follows, we provide an in-depth example of the EDM approach, beginning with why we believe this is an appropriate methodological choice. We then provide an overview of the model development process, data construction, model implementation, refinement and validation. Lastly, we discuss how the validated EDM can be used for scenario-based analyses and how its results can be presented to expert users and relevant stakeholders.

A STEP-BY-STEP GUIDE TO THE EDM APPROACH

The MERIAM project seeks to identify how, in response to conflict and climate shocks, household-level decisions affect nutrition-related outcomes – effectively unpacking the 'black box' of household behavior. At the household level, our EDM analysis is motivated by a set of fundamental questions that link household characteristics and behavior to acute malnutrition outcomes:

- Holding the context constant, why are some households affected by risk factors while others are not?
- To what extent do households within the same context react to risks in the same way?
- Is the same segment of a population recurrently affected, or is substantial flux observed?
- Do the risk factors for those affected remain the same year after year, or do they change over time?
- Does a given risk factor have the same effects across diverse contexts?

Our model uses *resilience*, the ability to cope with or adapt to various shocks and stressors, as a conceptual frame to investigate variation in nutritional outcomes in a manner that resonates with development stakeholders (e.g., Boukary et al., 2016; Food and Agriculture Organization of the United Nations, 2016; United States Agency for International Development, 2012; see Béné et al., 2015). In this particular domain, a resilient household – and in the aggregate, resilient communities, regions and countries – is better positioned to cope with the unfavorable effects of an exogenous shock, or has a greater ability to recover (say to pre-crisis intake levels of nutrition) in the aftermath of such a shock. Stakeholders may design interventions to boost endowments, moderate constraints, facilitate learning or strengthen systems for crisis management (Béné et al., 2015), all of which should contribute to greater resilience to nutritional crises.

The development of our EDM may be broken down into the six steps illustrated in Figure 4.1.

1. *Theoretical Grounding*: the EDM is grounded in existing theoretical and empirical knowledge of the subject matter.
2. *Data Construction*: relevant data to seed and validate the EDM are collected, analyzed and formatted.
3. *Model Implementation*: a preliminary version of the EDM is implemented as a computer simulation.
4. *Model Refinement and Cross-Case Validation*: the model is refined through expert interviews, fieldwork and out-of-sample testing.
5. *Counterfactual Analysis*: a valid EDM is extended by implementing counterfactual, 'what-if' experiments to explore how simulated trends are altered under different conditions.

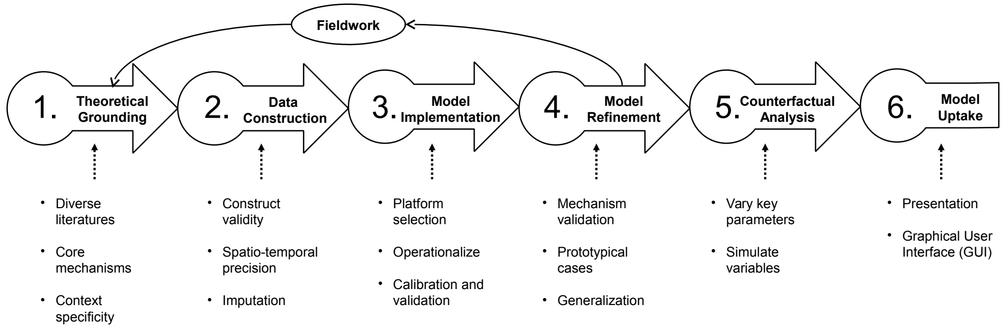

Figure 4.1 Model development process

6 *Model Uptake*: suggestions to visualize and present EDM results are put forth in an effort to make the model accessible to relevant end-users and stakeholders.

In each of these steps, the researcher makes consequential decisions that affect the outcome of the EDM process. Yet by design, EDM makes these choices explicit and transparent. And while computer simulations are not as well understood among social scientists and policy makers, the basic intuition underpinning EDM is relatively simple, perhaps more so than a statistical model that addresses similar questions.

Theoretical Grounding

First, we *surveyed a diverse body of research* on malnutrition, including detailed qualitative case studies (e.g. Hatløy et al., 1998; Manners, 2014; Parker et al., 2009), comprehensive 'broad-brush' approaches that integrate a wide range of mechanisms to explain malnutrition (e.g., Young and Marshak, 2017) and statistical analyses (e.g., Ajieroh, 2009; Ehrhardt et al., 2006; Fotso, 2007). Like any computational approach, the internal validity of EDM depends, in no small measure, on prior, often qualitative work that describes social processes in their requisite complexity.

Second, we reviewed this work to map the relations between leading and underlying indicators, in an effort to identify the **core mechanisms** that characterize malnutrition dynamics. The two defining categories into which these indicators fall are *shocks* and *stressors*. The first category of shocks includes the onset of a conflict, which usually has a sudden impact that is unanticipated by households. The second category of stressors accounts for the effects of longer-term or more gradual, recurring changes such as a lack of rainfall, which may vary in intensity, including its most extreme manifestation as drought.

Third, we examined *context-specific mechanisms*. In Karamoja, conflict has been endemic given the 20-year insurgency of the Lord's Resistance Army (LRA), as well as more recent pastoralist conflicts that involved cattle raiding or rustling (DCAF-ISSAT, 2017; FEWS NET, 2005). This has had many negative effects on the Karamojong, including the loss of human lives, displacement, reduction in livestock, and the progressive spread of small arms used by herders for protection, indirectly contributing to an increase in violence (DCAF-ISSAT, 2017).

Our model evolved as we relaxed simplifying assumptions and improved our understanding of malnutrition in Karamoja. Moving beyond an initial specification of households endowed with an unbounded ability to adapt behavior, we defined households as boundedly rational actors (Arthur, 1994), focusing primarily on food provision. The latest version of the model is depicted in

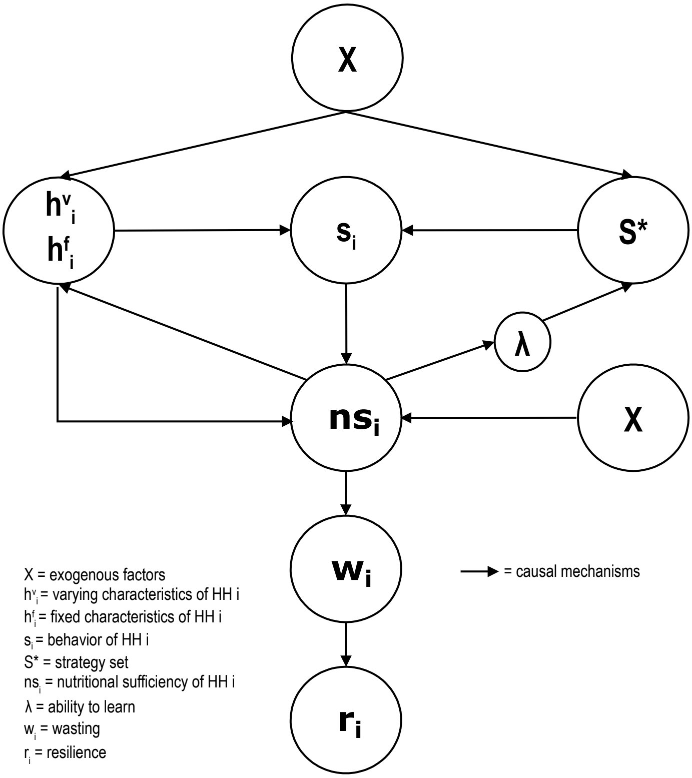

Figure 4.2 Flow diagram for MERIAM EDM

Figure 4.2. In each iteration and for every household, we calculate nutrition levels (*ns*) based on previous actions, a set of health- and food-intake factors ($h^{v,f}$), and exogenous constraints (X). Proportional to changes in ns, households adapt their behavior based on learning (λ): if *ns* is stable and sufficient to feed household members, the household continues to behave the way it did before. But if radical changes occur, households adapt their behavior either (a) randomly, (b) by copying a locally optimal strategy from their neighbors, or (c) by combining existing strategies to create a new one (Holland, 1975; Kollman et al., 1992; Krakauer and Rockmore, 2015; Mitchell, 1996; Urbanowicz and Moore, 2009).

In the current formulation, sub-optimal behavior is the rule rather than the exception, and status-quo behavior is effectively reinforced if households attribute a worsening of their situation to previous behavioral changes rather than exogenous factors (see ‘probe and adjust’ in Huttegger et al., 2014). The approach permits us to account for structural impediments to adaptation in a context like Karamoja, while still allowing for household-level change, e.g., from pastoral to agro-pastoral food production (e.g., Mercy Corps, 2016; see also Stites and Huisman, 2010). Note that the model, at this stage, may still be classified as an ABM. The next step is to construct the data necessary to enable the empirical contextualization of the model.

The process of theoretically grounded model development as is described here is prototypical, including the iterative refinement of model mechanisms and their operationalization as the modeler’s understanding develops. At this stage, choices are made based on the best available insights on the case, bearing in mind the need for further refinement following fieldwork and empirical validation in the next stage.

Data Construction

The EDM approach relies on data for empirical contextualization. The quality of the data – its accuracy, resolution and coverage – ultimately shapes our ability to seed and validate the model. A careful examination of data availability and quality is thus of paramount importance. In this section, we describe the main strengths and weaknesses of our household-level data for Karamoja, although our considerations may be generalized beyond the specifics of this case.

Our EDM is fundamentally about household characteristics and behavior, but it requires information exogenous to households as well, from district-level statistics about health facility capacities to more granular data at the grid or point level. Figure 4.3 presents an overview of the data we use to seed and validate the MERIAM EDM.

For our analysis of household characteristics and behavior in Karamoja, we utilize nutrition survey data provided by Action Against Hunger (2013). The dataset has two distinct advantages compared to other nutrition surveys. First, it contains behavioral variables at the household level. Second, the dataset is longitudinal: it consists of six survey rounds between August 2010 and May 2012, allowing us to validate and align the timescales of simulations against empirical outcomes repeatedly over this time period.

Our household-level data exhibits three principal weaknesses with respect to construct validity, spatio-temporal precision and completeness. First, some household-level variables do not measure the specific household attributes and behaviors we seek to model. For example, we use the variable 'food source' as a measure for how households obtain food. But only the '*most important'* food source was measured in the survey, which means that we cannot observe whether households use other means to obtain food.

Second, the data are imprecise. They capture longitudinal trends, rather than following the same households over time. With panel data, we could seed and validate our EDM against the decisions and characteristics that each particular household makes over time. With trend data, this information exists at an aggregate level to the extent that we know, on average, households changed on the measured variables between samples.

In addition, our household data is spatially imprecise insofar as it is representative at the ADM1 district. Short of obtaining representative samples at lower levels of analysis, there are still ways to mitigate spatial imprecision using *imputation*. For Karamoja, census data is only available at sparse intervals – no census was conducted between 2002 and 2014, a period that saw a 2.4% population growth in Karamoja (Uganda Bureau of Statistics, 2017). An alternative is to use remote-sensing data to estimate population numbers at the grid level.

The challenges to data quality from construct validity, (lack of) precision and completeness recur across empirical settings and are by no means specific to the EDM approach. The ability to accommodate data at varying levels of granularity, though, is a strength of the approach. As such, we need not match empirical and simulated household characteristics at the same level of granularity, given that we measure other contextual factors related to conflict, climate and market prices (see Figure 4.3). Instead of joining data at the lowest common level granularity, each model component can be simultaneously specified/matched with empirical data at the maximal level of granularity permissable.

Model Implementation

Several platforms are suitable for programming an EDM, with a trade-off between

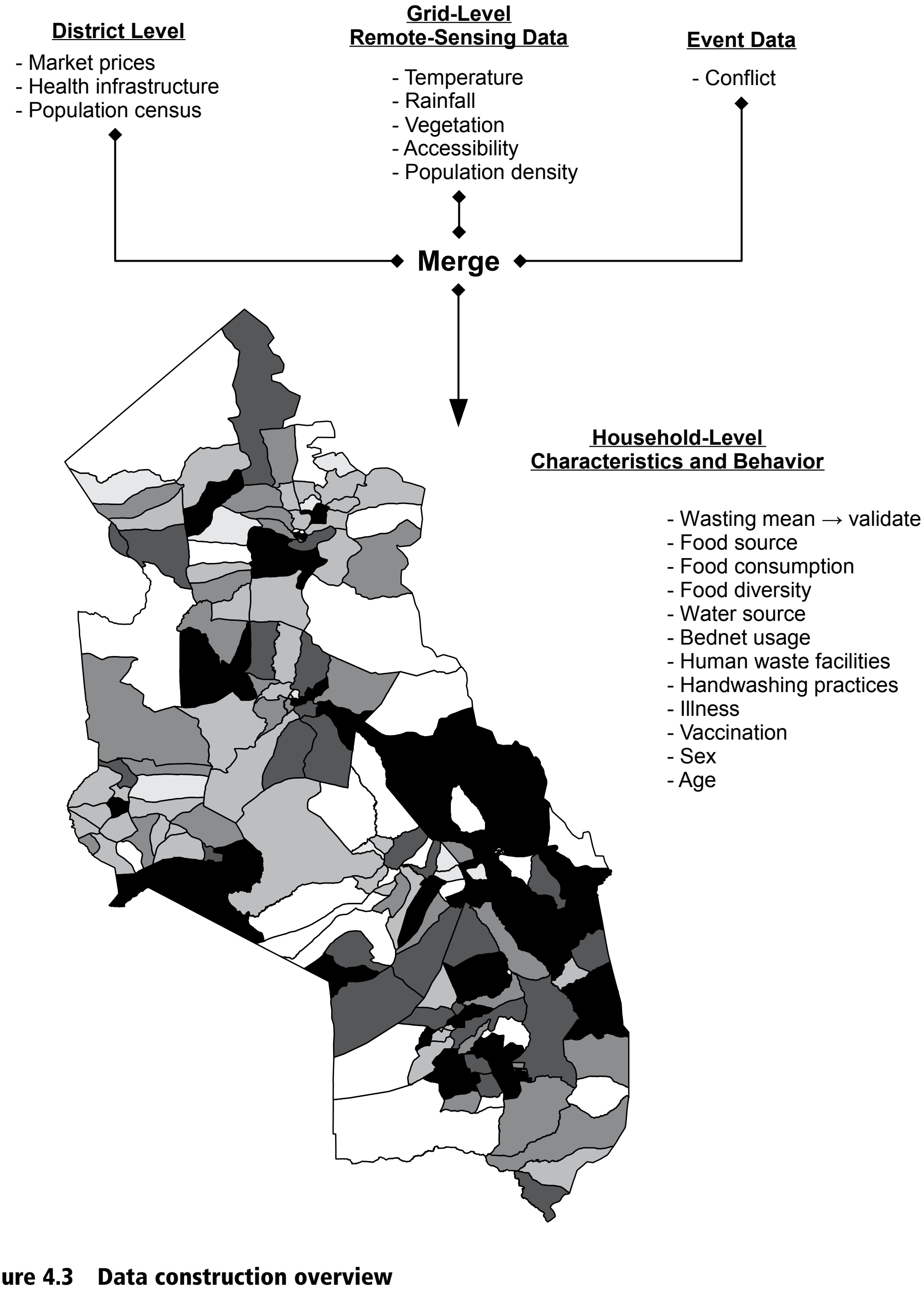

Figure 4.3 Data construction overview

simplicity of use and performance, i.e., the ability to handle complex modeling setups. A popular and widely used solution is the NETLOGO platform (Wilensky, 1999), which is easy to learn but more restrictive in its capabilities. At the other end of the spectrum are powerful libraries such as the Repast framework in Java (North et al., 2013) or MESA for Python (Masad and Kazil, 2015), both of which require greater customization.

The EDM discussed here uses a custom class-based implementation in Python. We note that the choice of programming language and framework do not affect model outcomes – only the ease of use and the ultimate runtime of the simulation.

In contrast, the *operationalization* of the model can have a profound impact. To operationalize model dynamics, it is necessary to specify the logical order of progression, sequence of actions and updates within the model – who does what, where and when and based on which information – paying close attention to attendant implications (Caron-Lormier et al., 2008). It follows that model heuristics can be universal or conditional, fixed or subject to change over time.

Consider, first, the problem of defining time progression within the model – the number of actions and updates that occur within a time step (e.g., an hour, day or year) to correctly reflect the timescales of the processes the model seeks to represent. A common solution is to make this correspondence explicit in the definition of a model time step, i.e., define time progression in terms of the fraction of possible actions or updates performed. For example, we consider a time step to have ended after updating the state of all households once. For the timescales of simulated and empirical outcomes to align, any time dependent parameters in the model that have empirical equivalents (e.g., the rate at which households adapt their strategies) have to be scaled such that their timescale aligns with that of the observed empirical process.

Second, EDM use geographical information for contextualization that ensures a high degree of correspondence between geography and the model topology. Exactly how space is operationalized within the model reflects an explicit choice to be made by the researcher. A common choice of implementation that reduces computational complexity is to discretize physical space. For example, in the model for Karamoja, household locations are defined on an underlying regular grid that is dynamically generated using actual settlement locations and their associated densities. In order to account for both low and high population densities, we use data on population densities at the grid level such that the number of households in a grid approximates the population density in the corresponding area in Karamoja.

Choices related to model operationalization are by no means simple or straightforward. To avoid influencing simulation outcomes or unwittingly introducing errors and artifacts, competing operationalizations of the same model mechanisms should be tested to ensure that a specific operationalization is not driving simulation outcomes (see also Galán et al., 2009).

Analogous to testing in- and out-of-sample predictive power for statistical models, EDM are formally validated and calibrated to maximize the correspondence between simulation results and real-world outcomes. Figure 4.4 shows the full modeling cycle, from model operationalization and contextualization to enumeration and calibration. For the Karamoja case, we identify the degree to which households are able to adapt to changing conditions that, all else equal, best explain the observed patterns of malnutrition. The closer the calibrated model approximates empirical outcomes, the greater the validity of the model predictions. Yet, quantitative agreement is not the only important measure. The parameters that best predict empirical outcomes must also reflect plausible dynamics on the ground. Should this fail, further refinement and validation of the model are necessary.

The modeling cycle illustrated for Karamoja (Figure 4.4) serves as a template for identifying parameters that yield the closest correspondence to real-world outcomes, a process that constitutes the core of the evidence-driven approach: given a model specification that formalizes our theoretical understanding of a process and data to seed the model (possibly at varying levels of granularity), what is the model with maximal explanatory power for our outcome of

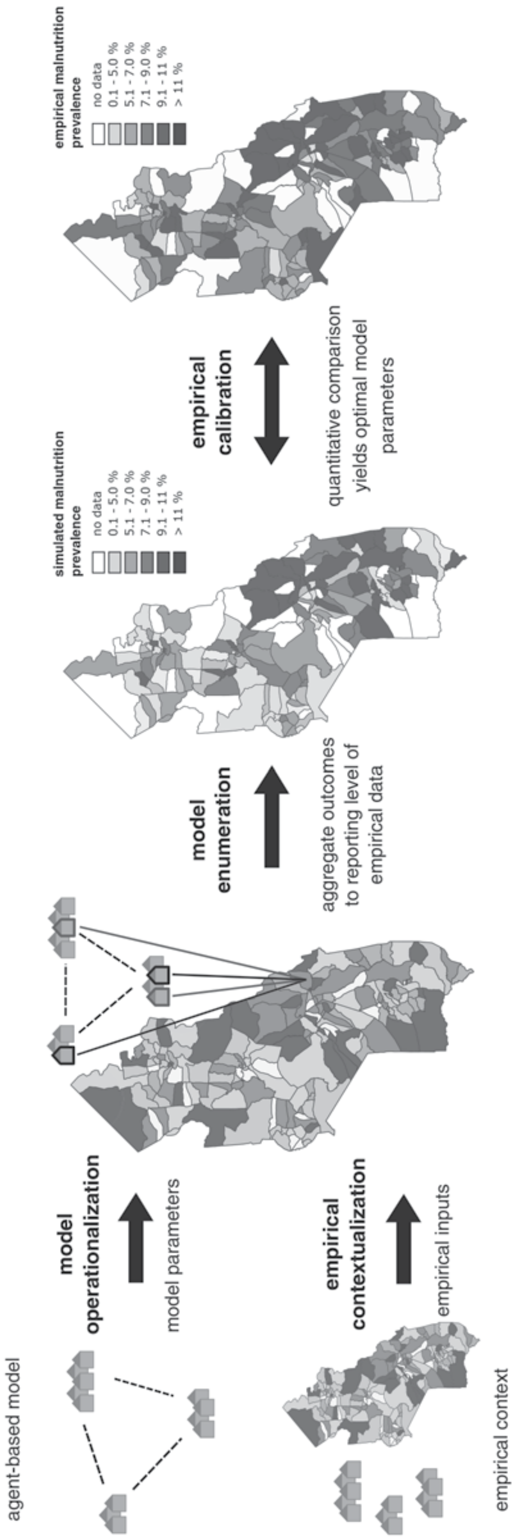

Figure 4.4 EDM modeling cycle

interest? The parameter constellation of the 'best-fit' model indicates which specific mechanisms in which constellation yields maximal agreement with empirical outcomes. Clearly, there is no guarantee that the constellation is empirically plausible, which makes the step of model refinement and validation an essential part of the process.

Model Refinement and Cross-Case Validation

While our model framework is derived from the existing literature, in addition to which we have undertaken a limited set of interviews with experts outside of the region, ground-truthing the model in Karamoja is an integral part of our EDM approach. Absent fieldwork, we run the risk of misrepresenting the dynamics at the core of our model, in particular, those that concern agent decision-making and behavior:

- Household decisions (ns, $\lambda \rightarrow S^*$): Who are household decision-makers? To what extent are 'households' independent decision-makers? Can they make decisions in the way we envision them? Which household characteristics are adaptable? If household characteristics change, what is the time scale? If households learn new strategies, what strategies are learned and how?
- Nutritional sufficiency (ns, $h^{f,v}$): How do household decision-makers evaluate the nutritional intake of their children? What general knowledge/awareness can we assume? What factors explain why some households cope/adapt successfully, while others do not?
- Household and exogenous variables ($h^{f,v}$, X): Is the relative importance of household and systemic constraints in our model appropriately balanced? What amount of agency can we attribute to households?
- Resilience (r): What types of resilience (coping vs. adaptation) can we expect in a context like Karamoja?

These and other questions pertain to core components of our model. They needed to be addressed with experts and government officials working on the ground, and more importantly with Karamojong households whose nutritional situation we seek to understand.

The relationship between model development and field research can and should be treated as an iterative process. Multiple stages of model development can be interspersed with multiple rounds of field research. As such, the EDM approach is agnostic to the timing of fieldwork. A possible starting point is the development of a theoretically grounded model framework. A first round of field research can then be used to refine the model, identify relevant causal linkages and supply additional empirical input for the model. For MERIAM, we received some of our most valuable input through modeling exercises where we probed experts and households to make their 'mental models' explicit, through surveys and focus group discussions. Depending on one's epistemological orientation and practical considerations, it is plausible to conceptualize a modeling framework inductively from preliminary field research.

Whereas 'getting the story right' in Karanoja was important, it is evident that an EDM initially developed and tested in one region need not be applicable to other contexts. Households may respond differently to the same exogenous shocks and stressors in ways that cannot be accounted for comprehensively in a single model specification. That said, re-building the entire model for each context analyzed is also unnecessary. Rather, a given model can be modified to incorporate contextual differences in a systematic manner, by identifying model *prototypes* that exhibit meaningful variation across key dimensions, and *validating and refining* the model for each prototype.

A necessary first step involves the identification of similar patterns across contexts, with respect to either (causal) drivers, mechanisms, behaviors or outcomes, as a means of building a set of computational model prototypes. For example, prototypes for malnutrition dynamics could be constructed to account for similar patterns of incidence and prevalence in wasting across contexts. The selection of prototypical cases that exhibit consequential variation, both geographically and temporally, is essential to ensure that an EDM generalizes beyond the confines of a particular case. More specifically, a different case enables us to refine the model with a view toward maximizing external validity, while a similar case allows us to verify and strengthen the internal validity of the model in another context. For the MERIAM project, analyzing sub-national cases over finer-grained temporal units was preferable to selecting country-years as units of observation.

As a second step, case-specific grounding, data construction, and fieldwork are repeated for every model prototype to which the EDM will apply, building on prior work where possible, given that much of the data on exogenous factors is constructed uniformly across contexts (e.g., remote-sensing and conflict data). For our second prototypical case, West Pokot in Kenya, we assessed historical and cultural differences relative to Karamoja, as well as changes in salient causal mechanisms. Finally, we adapted the field survey used in the Karamoja case for the particularities of this context, making only the minimal changes required.

So while the classic trade-off between external and internal validity applies in no small measure, EDM can be systematically extended to produce valid results across contexts, while retaining internal validity for specific cases.

Counterfactual Analysis

Counterfactual analysis can be divided into two types of ‘what-if’ scenarios for EDM: those that relate to model parameters – determinants of model dynamics that have no direct empirical referent and were inferred from the model – and those that relate to model inputs specified by empirical data.

Counterfactual analysis for model parameters is equivalent to considering comparative statistics. Here, one sets all model parameters to their optimal values, except for a particular parameter whose influence we seek to assess. For example, the effect of a household’s propensity to adapt could be analyzed, all the more tellingly if the effect is non-linear, i.e., if small changes in household behavior produce significant changes in nutrition outcomes.

Counterfactuals may also be conducted for model inputs. Instead of seeding all inputs with empirical data, one can use exogenously determined values for one or more inputs, treating them as equivalent to parameters. Examples include testing the impact of climatic and economic shocks, as well as the source, timing, location, type and scope of interventions (e.g., food imports, humanitarian assistance from international sources and education designed to shape household behavior) on household behavior and malnutrition. Stakeholders can then use these insights to understand the likely effects of different interventions under a variety of conditions across different contexts. While this type of counterfactual analysis is best constrained to the period for which the model was optimized – in other words, within sample – out-of-sample counterfactual analysis, including forecasting, is feasible when one clearly specifies how parameters and empirical inputs might change in the future.

It follows that EDM are well suited to providing data-driven, scenario-based analyses, with the caveat that underlying assumptions are transparently communicated. The EDM developed as part of MERIAM forms the basis for a tool to make scenario-based forecasts of malnutrition and explore the efficacy of various interventions in response to climate- or conflict-related shocks. A concern, then, is to develop an effective means to communicate the methodology beyond a purely academic or expert audience.

Model Uptake

The EDM approach requires a combination of technical knowledge and relevant domain expertise. As such, the specification, contextualization, validation and refinement of an EDM is typically undertaken by academic researchers or trained experts. On the other hand, if done correctly, the kinds of scenario-based analysis for which EDM is suited may be of immediate interest to practitioners, subject experts or policy makers largely unfamiliar with the methodology. The question then concerns how the EDM methodology can be pitched to stakeholders beyond a mere visualization of results, to provide a 'feel' for the model, its specificity and generalizability, and its use as a tool for making evidence-driven policy decisions.

With an expert and practitioner audience in mind, we plan to complement an academic research paper with a policy brief written for a general, non-specialist audience. Rather than using technical jargon, a brief would explain the steps involved in model construction, much in the way that this chapter does, highlighting key policy-relevant insights. The brief would clearly communicate the limits and uncertainties associated with the EDM scenario-based forecasts, considering the effect of specific interventions on malnutrition outcomes discussed above.

While a brief will certainly help communicate policy-relevant insights, it falls short of providing a true 'feel' for the EDM. The only way to achieve this is by developing a tool with an interactive graphical user interface (GUI). Such a tool would preserve the full complexity of the EDM while allowing non-expert users to easily engage with the model and translate output. The tool may also require an expert to revise a particular model specification for a new case, after which the GUI will perform its intended function.

Figures 4.5 and 4.6 depict a mock-up of the Simulating Acute Malnutrition Toolkit (SAMT). At the center of the GUI is a trend-based forecast and a map of the region being analyzed. Users can switch between outcomes (e.g., resilience or wasting prevalence) and select the time point for which outcomes on the map would be displayed. The heat map shows normalized levels of the selected outcome (here, resilience) at a fine-grained level of analysis for the selected time point, with the slider below the time

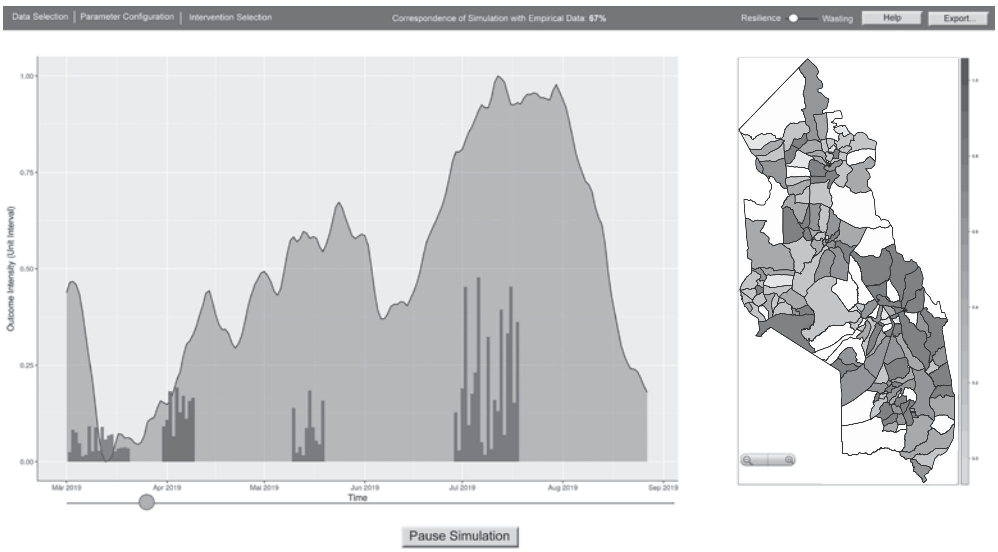

Figure 4.5 Mock-up of SAMT (main window)

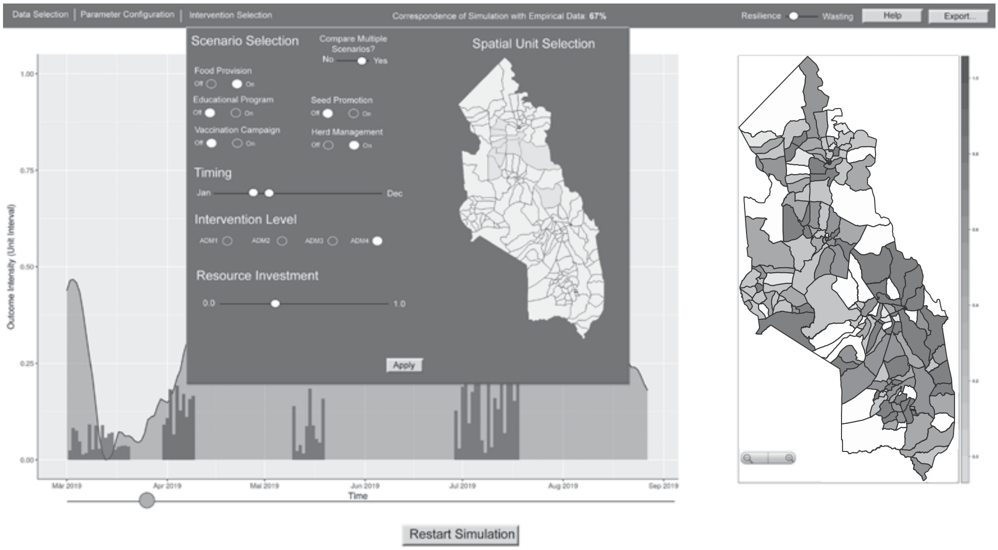

Figure 4.6 Mock-up of SAMT (intervention configuration)

series graph. Bars in the time series indicate the intensity of shocks and aid interventions, respectively. The degree to which the simulated results correspond to empirical data is specified at the top of the interface.

GUIs and software tools for decision support are generally starting to gain traction. These include a large array of domain-unspecific tools for data handling and visualization, as well as more specific policy support tools in diverse domains such as epidemiology (den Broeck et al., 2011) and public safety (Chooramum et al., 2016). Making EDM accessible to others requires this kind of explicit engagement with stakeholders, allowing them to develop a better intuition for the approach.

CONCLUSION

Evidence-driven computational modeling effectively harnesses the strengths of ABM, while achieving a high degree of real-world correspondence and resonance. As our discussion of the MERIAM project demonstrates, the EDM approach incorporates contextual knowledge and theoretical insight, captures complex spatio-temporal inter-dependencies, explicitly accounts for endogenous relationships, uses realistic topographies and harnesses data at varying levels of measurement. The combination places EDM at the intersection of theory and empirical work. For the MERIAM project, we harness the power of EDM to make scenario-based forecasts and undertake counterfactual analyses, developing a tool for policy makers tasked with addressing the high-stakes problem of malnutrition. The development of the MERIAM EDM has been elaborate, costly and time consuming, given that many of the standard elements of research design in political science – theory building, case selection, data collection and fieldwork – comprise the approach. We believe the contribution to evidence-driven decision-making is well worth the effort, and trust that the procedures and best practices outlined in this chapter will result in the development and use of EDM across diverse domains.

Notes

1 The authors thank Alessandra Romani from the Graduate Institute, Geneva for helpful comments on draft versions of this chapter. This document is an output from a project funded by UK Aid from the UK Department for International Development (DFID). The views expressed do not necessarily reflect the UK government's official policies.

2 Take, for example, contact theory. Researchers have found empirical evidence to support the notion that increased inter-group contact leads both to higher and lower levels of violence. EDM have been used to explore the conditions that give rise to these divergent outcomes. For a detailed discussion, see Bhavnani et al. (2014).

3 Visit https://www.actionagainsthunger.org/meriam for more information on the project.

4 This description, drawn from the MERIAM project proposal, serves as the overarching motivation for the larger project as well as our specific contribution to the same.

5 http://fews.net/IPC.

6 See Samimi et al. (2012) for an earlier critique, and http://fews.net/our-work/our-work/scenario-development for details on the FEWS NET scenario-building process we seek to complement with our approach.

REFERENCES

Action Against Hunger. 2013. 'Nutrition Surveillance Data Analysis: Karamoja, Uganda, December 2009 – May 2012.' Technical Report, https://www.actionagainsthunger.org/sites/default/files/publications/Nutrtion_Surveillance_Data_Analysis_Uganda_08.2013.pdf (Retrieved July 15, 2019).

Ajieroh, Victor. 2009. 'A Quantitative Analysis of Determinants of Child and Maternal Malnutrition in Nigeria.' *Nigeria Strategy Support Program* (NSSP10): 1–69.

Arthur, W. Brian. 1994. *Increasing Returns and Path Dependence in the Economy*. Ann Arbor, MI: University of Michigan Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axtell, R. L., J. M. Epstein, J. S. Dean, G. J. Gumerman, A. C. Swedlund, J. Harburger, S. Chakravarty, R. Hammond, J. Parker and M. Parker. 2002. 'Population Growth and Collapse in a Multiagent Model of the Kayenta Anasazi in Long House Valley.' *Proceedings of the National Academy of Sciences* 99(Supplement 3): 7275–7279.

Bankes, Steve. 1993. 'Exploratory Modeling for Policy Analysis.' *Operations Research* 41(3): 435–449.

Béné, Christophe, Derek Headey, Lawrence Haddad and Klaus von Grebmer. 2015. 'Is Resilience a Useful Concept in the Context of Food Security and Nutrition Programmes? Some Conceptual and Practical Considerations.' *Food Security* 8(1): 123–138.

Bhavnani, Ravi. 2003. 'Adaptive Agents, Political Institutions and Civic Traditions in Modern Italy.' *Journal of Artificial Societies and Social Simulation* 6(4).

Bhavnani, Ravi, and David Backer. 2000. 'Localized Ethnic Conflict and Genocide: Accounting for Differences in Rwanda and Burundi.' *Journal of Conflict Resolution* 44(3): 283–306.

Bhavnani, Ravi, David Backer and Rick R. Riolo. 2008. 'Simulating Closed Regimes with Agent Based Models.' *Complexity* 14(1): 36–44.

Bhavnani, Ravi, Karsten Donnay, Dan Miodownik, Maayan Mor and Dirk Helbing. 2014. 'Group Segregation and Urban Violence.' *American Journal of Political Science* 58(1): 226–245.

Bhavnani, Ravi, Michael G. Findley and James H. Kuklinski. 2009. 'Rumor Dynamics in Ethnic Violence.' *The Journal of Politics* 71(3): 876–892.

Boukary, Aboubakr Gambo, Adama Diaw and Tobias Wünscher. 2016. 'Factors Affecting Rural Households' Resilience to Food Insecurity in Niger.' *Sustainability* 8(3): 181.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Caron-Lormier, Geoffrey, Roger W. Humphry, David A. Bohan, Cathy Hawes and Pernille Thorbek. 2008. 'Asynchronous and Synchronous Updating in Individual-Based Models.' *Ecological Modelling* 212(3–4): 522–527.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton, NJ: Princeton University Press.

Chooramum, Nitish, Peter J. Lawrence and E. R. Galea. 2016. 'Urban Scale Evacuation Simulation Using buildingEXODUS.' *Interflam* 2: 1645–1656.

DCAF-ISSAT. 2017. 'Cattle Rustling and Insecurity in Africa: A Comparative Perspective.' Technical Report, https://issat.dcaf.ch/Learn/SSR-in-Practice2/Countries-Regions/Madagascar/Reinforcing-African-Union-SSR-Unit-support-to-national-SSR-processes-Madagascar (Retrieved July 15, 2019).

de Marchi, Scott, and Scott E. Page. 2014. 'Agent-Based Models.' *Annual Review of Political Science* 17(1): 1–20.

den Broeck, Wouter Van, Corrado Gioannini, Bruno Gonçalves, Marco Quaggiotto, Vittoria Colizza and Alessandro Vespignani. 2011. 'Implementation Results and Discussion Conclusions Availability and Requirements Declarations References Software Open Access Open Peer Review the GLEaMviz Computational Tool, a Publicly Available Software to Explore Realistic Epidemic Spreading Scenarios at the Global Scale.' *BMC Infectious Diseases* 11: 37.

Ehrhardt, Stephan, Gerd D. Burchard, Carsten Mantel, Jakob P. Cramer, Sarah Kaiser, Martina Kub, Rowland N. Otchwemah, Ulrich Bienzle and Frank P. Mockenhaupt. 2006. 'Malaria, Anemia, and Malnutrition in African Children – Defining Intervention Priorities.' *The Journal of Infectious Diseases* 194(1): 108–114.

Epstein, Joshua M. 1999. 'Agent-Based Computational Models and Generative Social Science.' *Complexity* 4(5): 41–60.

Epstein, J. M. 2002. 'Modeling Civil Violence: An Agent-Based Computational Approach.' *Proceedings of the National Academy of Sciences* 99(Supplement 3): 7243–7250.

Epstein, Joshua M., and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington D.C.: Brookings Institution Press.

FAO, IFAD, UNICEF, WFP and WHO. 2018. 'The State of Food Security and Nutrition in the World 2018: Building Climate Resilience for Food Security and Nutrition.' Technical Report, http://www.fao.org/3/I9553EN/i9553en.pdf (Retrieved July 15, 2019).

FEWS NET. 2005. 'Conflict Early Warning and Mitigation of Resource Based Conflicts in the Greater Horn of Africa: Conflict Baseline Study Report Conducted in the Karamajong Cluster of Kenya and Uganda.' Technical Report, Famine Early Warning Systems Network (FEWS NET)/GHA.

Fisher, Franklin M. 1966. *The Identification Problem in Econometrics*. New York: McGraw-Hill.

Food and Agriculture Organization of the United Nations. 2016. 'RIMA-II: Resilience Index Measurement and Analysis – II: Analysing Resilience for Better Targeting and Action.' Technical Report, http://www.fao.org/3/a-i5665e.pdf (Retrieved July 15, 2019).

Fotso, Jean-Christophe. 2007. 'Urban-Rural Differentials in Child Malnutrition: Trends and Socioeconomic Correlates in Sub-Saharan Africa.' *Health & Place* 13(1): 205–223.

Galán, José Manuel, Luis R. Izquierdo, Segismundo S. Izquierdo, José Ignacio Santos, Ricardo del Olmo, Adolfo López-Paredes and Bruce Edmonds. 2009. 'Errors and Artefacts in Agent-Based Modelling.' *Journal of Artificial Societies and Social Simulation* 12(1): 1.

Geller, Armando, and Scott Moss. 2008. 'Growing QAWM: An Evidence-Driven Declarative Model of Afghan Power Structures.' *Advances in Complex Systems* 11(2): 321–335.

Gneiting, Tilmann, and Adrian E. Raftery. 2005. 'Weather Forecasting with Ensemble Methods.' *Science* 310(5746): 248–249.

Hatløy, A., L. E. Torheim and A. Oshaug. 1998. 'Food Variety – a Good Indicator of Nutritional Adequacy of the Diet? A Case Study from an Urban Area in Mali, West Africa.' *European Journal of Clinical Nutrition* 52(12): 891–898.

Holland, John. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Huttegger, Simon M., Brian Skyrms and Kevin J. S. Zollman. 2014. 'Probe and Adjust in Information Transfer Games.' *Erkenntnis* 79(S4): 835–853.

Kim, Sung-Youn, Charles S. Taber and Milton Lodge. 2010. 'A Computational Model of the Citizen as Motivated Reasoner: Modeling the Dynamics of the 2000 Presidential Election.' *Political Behavior* 32(1): 1–28.

Kollman, Ken, John H. Miller and Scott E. Page. 1992. 'Adaptive Parties in Spatial Elections.' *American Political Science Review* 86(4): 929–937.

Krakauer, David, and Daniel Rockmore. 2015. 'The Mathematics of Adaptation (Or the Ten Avatars of Vishnu).' In *The Princeton Companion to Applied Mathematics*, ed. Nicholas J.

Higham. Princeton, NJ: Princeton University Press, 591–597.

Laver, Michael. 2005. 'Policy and the Dynamics of Political Competition.' *American Political Science Review* 99(2): 263–281.

Laver, Michael, and Ernest Sergenti. 2011. *Party Competition: An Agent-Based Model*. Princeton, NJ: Princeton University Press.

Laver, Michael, Scott de Marchi and Hande Mutlu. 2011. 'Negotiation in Legislatures Over Government Formation.' *Public Choice* 147(3–4): 285–304.

Lustick, Ian S., Dan Miodownik and Roy J. Eidelson. 2004. 'Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?' *American Political Science Review* 98(2): 209–229.

Malleson, Nick, and Mark Birkin. 2012. 'Analysis of Crime Patterns Through the Integration of an Agent-Based Model and a Population Microsimulation.' *Computers, Environment and Urban Systems* 36(6): 551–561.

Manners, Kristy. 2014. 'Nutrition Causal Analysis: Isiolo County, Kenya.' Technical Report, Action Against Hunger, https://www.actionagainsthunger.org/sites/default/files/publications/ACF_NCA_Kenya_Report_Feb2014.pdf (Retrieved July 15, 2019).

Masad, David, and Jacqueline Kazil. 2015. 'Mesa: An Agent-Based Modeling Framework.' https://github.com/projectmesa/mesa

Mercy Corps. 2016. 'Karamoja Strategic Resilence Assessment.' Technical Report. https://www.mercycorps.org/sites/default/files/UgandaSTRESSKaramojaFinalRep.pdf (Retrieved July 15, 2019).

Mitchell, Melanie. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

North, Michael J., Nicholson T. Collier, Jonathan Ozik, Eric R. Tatara, Charles M. Macal, Mark Bragen and Pam Sydelko. 2013. 'Complex Adaptive Systems Modeling with Repast Simphony.' doi:10.1186/2194-3206-1-3.

Nowak, Martin A., and Robert M May. 1992. 'Evolutionary Games and Spatial Chaos.' *Nature* 359(6398): 826–829.

Palmer, Richard G., William Brian Arthur, John H. Holland and Blake LeBaron. 1999. 'An Artificial Stock Market.' *Artificial Life and Robotics* 3(1): 27–31.

Parker, Dawn C., Kathryn H. Jacobsen and Maction K. Komwa. 2009. 'A Qualitative Study of the Impact of HIV/AIDS on Agricultural Households in Southeastern Uganda.' *International Journal of Environmental Research and Public Health* 6(8): 2113–2138.

Richiardi, Matteo, Roberto Leombruni, Nicole J. Saam and Michele Sonnessa. 2006. 'A Common Protocol for Agent-Based Social Simulation.' *Journal of Artificial Societies and Social Simulation* 9(1): 15.

Rogers, J. Daniel, Claudio Cioffi-Revilla and Samantha Jo Linford. 2015. 'The Sustainability of Wealth Among Nomads: An Agent-Based Approach.' In *Mathematics and Archaeology*, eds Juan A. Barcelo and Igor Bogdanovic. Boca Raton, FL: CRC Press, 431–438.

Samimi, C., A. H. Fink and H. Paeth. 2012. 'The 2007 Flood in the Sahel: Causes, Characteristics and Its Presentation in the Media and FEWS NET.' *Natural Hazards and Earth System Sciences* 12(2): 313–325.

Schelling, Thomas C. 1971. 'Dynamic Models of Segregation.' *Journal of Mathematical Sociology* 1(2): 143–186.

Skamarock, William C., and Joseph B. Klemp. 2008. 'A Time-Split Nonhydrostatic Atmospheric Model for Weather Research and Forecasting Applications.' *Journal of Computational Physics* 227(7): 3465–3485.

Stites, Elizabeth, and Carrie S. Huisman. 2010. 'Adaptation and Resilience: Responses to Changing Dynamics in Northern Karamoja, Uganda.' Briefing Paper, Feinstein International Center, Tufts University, Boston..

Tullock, Gordon, and Colin D. Campbell. 1970. 'Computer Simulation of a Small Voting System.' *The Economic Journal* 80(317): 97–104.

Uganda Bureau of Statistics. 2017. 'National Population and Housing Census 2014 – Analytical Report.' https://www.ubos.org/onlinefiles/uploads/ubos/2014CensusProfiles/National_Analytical_Report_nphc\%202014.pdf (Retrieved July 15, 2019).

United States Agency for International Development, ed. 2012. *Building Resilience to Recurrent Crisis: USAID Policy and Program Guidance*.

Urbanowicz, Ryan J., and Jason H. Moore. 2009. 'Learning Classifier Systems: A Complete Introduction, Review, and Roadmap.'

Journal of Artificial Evolution and Applications, 1–25. doi.10.1155/2009/736398.

Weidmann, Nils B., and Idean Salehyan. 2013. 'Violence and Ethnic Segregation: A Computational Model Applied to Baghdad.' *International Studies Quarterly* 57(1): 52–64.

Wilensky, Uri. 1999. 'NetLogo.' http://ccl.northwestern.edu/netlogo/.

Young, Helen, and Anastasia Marshak. 2017. 'Persistent Global Acute Malnutrition.' Technical Report, Feinstein International Center, Tufts University, Boston.