

Towards a Multi-Scale Theory on Coupled Human Mobility and Environmental Change

1 Identification of the research and issues

Tens of millions of people are migrants: many fled conflicts; others sought better economic opportunities. Such population movements can be caused by—as well as affect—natural systems: droughts may lead to social conflicts; migrants may overwhelm the natural resources and in turn social fabric at destinations. The magnitudes and problems associated with migration are of grave concern, but a *satisfactory, mechanistic and predictive* theory of the interdependence between human mobility and environmental changes is still lacking. Such a theory is necessary for successful intervention and contingency plans.

Forces that drive migration are often described as “pushes” and “pulls.” There must be *pushes* at the origin, e.g., conflicts, economic hardships, natural disaster, or environmental degradation, that create a group of people who wish to migrate. There are then *pulls* of different strengths from different destinations. Several theories and models have been proposed to understand and describe these pushes and pulls (Massey et al., 1993; Greenwood, 2005; Klabunde & Willekens, 2016), but many of them are piecemeal and fragmented (Willekens et al. 2016). Gravity-like models (e.g., Cohen et al. 2008; Simini et al. 2012) are simple, parsimonious, and exhibit some exploratory power for migration patterns, but they are statistical and non-mechanistic in nature and do not explicitly incorporate environmental factors. Recent conflict research has increasingly linked these conflicts to environmental and climate changes (Van Holt et al. 2016), but most climate-conflict models (e.g., Hsiang, Burke, & Miguel 2013; Hsiang, Meng, & Cane, 2011) are non-dynamical and do not directly address the human mobility aspect, i.e., focusing on the pushes but not the pulls (but see, e.g., Hauer 2017). Beyond a recent focus on social networks, roles of social institutions such as norms, cultures, and ethnic enclaves have been either overlooked or not systematically incorporated in these models, let alone the interdependence of these social institutions with the natural systems. Without explicit connections between the *dynamics* of human and natural systems, it is difficult to understand the nature of migration-inducing tipping points, potential secondary pushes forcing new migration patterns to emerge, and how these system properties change when multiple stressors occurring simultaneously. Furthermore, findings from migration analyses are often scale-dependent (e.g., Gober-Meyers, 1978), and recent advances in multiple disciplines have not been effectively incorporated (e.g., De Domenico et al., 2013, Boccaletti et al. 2014, Kivelä et al. 2014, Yu et al. 2015, Muneeppeerakul & Anderies 2017; Meekan et al. 2017). There is a need for a richer theory that fills knowledge gaps between these useful, but fragmented, building blocks identified above and integrate them.

Thus, the overall objective of this project is to develop a modeling platform that is flexible enough to explore different approaches at multiple spatiotemporal scales as well as to strike the right balance of predictive power and facilitation of the development of an integrative theory. Such a platform must retain the useful features of existing models, while filling the gaps identified above. To guide our approach, it is instructive to consider the desired properties of such a modeling platform. With respect to these properties, our specific objectives are:

- [A] To include explicit coupling between environmental changes and human systems.
- [B] To provide the capability to simulate spatial movements of humans.
- [C] To produce models that are mechanistic and process-based, while accounting for uncertainties. This enhances the reliability and predictive power of the models, furnishes us with insights into the nature of tipping points that trigger human migration.
- [D] To develop analyses that are seamlessly linked across scales: from the local to global scales as well as for short (refugee flows) and long (permanent migration) time scales.
- [E] To provide the ability to capture decision making at different levels of aggregation in population.
- [F] To integrate representation of the roles played by social institutions.
- [G] To provide for simplicity and parameter-parsimony. This will increase model applicability and likelihood of clear analytical relationships that would serve as building blocks to develop a general theory.
- [H] To include the capacity to generate scenarios of environmental changes and assess their likelihood.

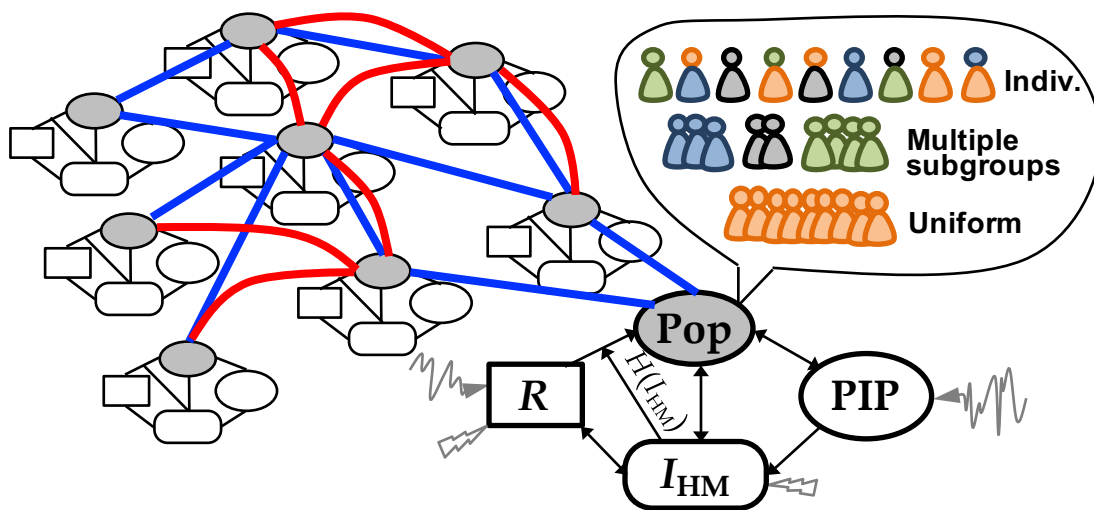


Figure 1. Schematic diagram of the proposed approach. The node-level dynamical models will be based on a widely used framework for social- ecological systems: Pop = population, R = resource, I_{HM} = human-made infrastructure, PIP = public infrastructure providers, and function $H(I_{HM})$ captures the influence of the infrastructure on the ability of population to make use of the critical resource. The critical resource and corresponding infrastructure can vary from node to node. Aggregation level of a population ranges from uniform, multiple subgroups, and individuals. The nodes are linked in a multilayer network: a pair of nodes can be linked by a set of different relationships, e.g., familial, geographical proximity, natural barriers, shared religious belief, cultural or geographical affinity, diplomatic ties (illustrated by lines of different colors). The wiggly arrows and lightning symbols represent fluctuations and shocks, respectively.

2 Proposed technical approaches

To construct models with the above properties, we draw from existing models and theories as well as inventing new components. Specifically, we will model **the migration process as movements of people within a multilayer network of populations, with each location having its own coupled dynamics between environmental and social components and each pair of sending and receiving locations connected by multiple types of linkages (Fig. 1)**. We will apply the models to **case studies of different spatial and temporal scales driven by different environmental changes**. We will consider two general classes of environmental changes, namely, sudden shocks and gradual changes, as they likely activate different sets of migration-inducing factors. In particular, we will investigate at least four case studies (Section 2.5): one with a natural disaster (a hurricane); one with degrading socio-economic conditions punctuated by a natural disaster; one with a persistent drought, and one as an example of “secondary pushes” and emerging migration patterns. Close attention will be paid how model structure changes across these scales and classes of environmental changes. (Hereinafter, we refer to the objectives identified above by square brackets.)

Sudden shocks and gradual changes in environmental forcing call for different approaches. Migration in response to sudden shocks, such as earthquakes, wildfires, and hurricanes, is akin to forced or involuntary migration caused by war or civil unrest; people do not have much time to plan or carefully consider different destination. Migration in response to gradual changes, on the other hand, may allow for more elaborate calculation of potential economic gains; for example, a period of prolonged droughts may cause populations to consider moving to other locations with more promising economic opportunities. Likewise, gradual declines in groundwater levels (e.g., Wada et al., 2010; Dalin et al., 2017) will lead to greater pumping costs and water scarcity, changing the potential for agricultural profits. As such, we will be borrowing ideas from or implementing existing migration theories differently or at different degrees for sudden shocks and gradual changes.

The node-level dynamics will first be studied by low-dimensional, dynamical system models; these models clarify the nature of migration-inducing tipping points and provide guidelines for how to add more realistic complexity to the models. Relationships between locations/populations will be represented by different layers in a multilayer (multiplex) network; these relationships embody different migration theories. The structure between these relationships and migration will be inferred in a Bayesian framework, which can incorporate qualitative inputs and expert opinions and explicitly model uncertainties; in this way, our multiplex network approach both evaluates and integrates different migration theories. The models will be driven by environmental changes, including climate extremes. In each case study, the model complexity will be grown carefully—more model complexity does not always mean better results or understanding. Global sensitivity analysis will be used to analyze how the model precision in predicting migration outcomes changes with model complexity, thereby identifying the right level of model complexity for prediction purposes and for theory development. Annual workshops will be held that invite experts on migration theories, environmental modeling, and other related fields to provide feedbacks and constructive criticism on the development of models and integrative theory.

2.1 Node-level dynamics

For the coupled node-level dynamics, we will develop a model based on a widely used conceptual framework put forth by Anderies, Janssen, & Ostrom (2004) and Anderies (2015). This conceptual framework was based on many case studies of social-ecological systems (e.g., Ostrom 1990) [A, C]. Here, *infrastructure* is defined broadly to include both hard (e.g., canals, bridges) and soft (e.g., rules, norms) infrastructure—the roles of social institutions will be incorporated naturally as part of the soft infrastructure [D]. Recent work has mathematically operationalized this framework into low-dimensional dynamical system models [G] (Yu et al. 2015; Muneeppeerakul & Anderies, 2017). Such models mechanistically capture the interdependencies among the environmental and social components and yield conditions that separate sustainability and collapse – namely, tipping points or critical transitions (e.g., Scheffer et al., 2012) – in terms of mathematical expressions that link biophysical and social factors [C].

To provide a clearer picture of the mathematical operationalization of the framework, a sketch of model development is briefly discussed here. The sketch below also illustrates how concepts from different disciplines may enter the model. Consider the following generic system of equations:

$$\begin{aligned} R_{j,t+1}^r &= G(R_{j,t}^r; \mathbf{S}) - A(R_{j,t}^r, I_{j,t}^r, N_{j,t}) \\ N_{j,t+1}^n &= F(\mathbf{N}_t, \mathbf{R}_t, \mathbf{I}_t, f_{ij}(\{M_{ij,t}^l\}; \{\beta_{j,t}^l\}); \mathbf{S}) \\ I_{j,t+1}^r &= H(I_{j,t}^r; D_{j,t}, \mathbf{S}) \end{aligned} \quad (1)$$

Subscripts t and $t + 1$ are time indices; R_j^r = Resource of type r at location j ; I_j^r = infrastructure related to resource type r at location j ; N_j^n = Population or subpopulation or user group n at location j ; \mathbf{S} = a vector of shocks that impact different components of the model. On the right-hand sides of the equations include the followings: $G(R_{j,t}^r; \mathbf{S})$ captures the natural processes of the critical physical and environmental resources—as the model become more realistic and complex this function $G(R_{j,t}^r; \mathbf{S})$ will essentially be subsumed by the type of model in Section 2.3; $A(R_{j,t}^r, I_{j,t}^r, N_{j,t})$ captures the appropriation of the resources by the populations as mediated by infrastructure functionality—concepts from economics, game theory, and governance and institutional analyses can enter this function; $F(\mathbf{N}_t, \mathbf{R}_t, \mathbf{I}_t, f_{ij}(\{M_{ij,t}^l\}; \{\beta_{j,t}^l\}); \mathbf{S})$ captures the population dynamics, which includes function $f_{ij}(\{M_{ij,t}^l\}; \{\beta_{j,t}^l\})$ that translates the relationships $\{M_{ij,t}^l\}$ in the multiplex network into flows of migration (see Section 2.2.); and $H(I_{j,t}^r; D_{j,t}, \mathbf{S})$ captures the balance between the maintenance efforts and decay/depreciation of the infrastructure (be it hard or soft).

It is instructive to consider a concrete example of how this approach leads to clear mathematical expressions of tipping points in a social-ecological system. To this end, let us briefly consider recent work by Muneeppeerakul & Anderies (2017) who developed a model based on the above framework with an irrigated system in mind (R = water, N = farmers, PI = irrigation canals, and PIP = an agency tasked with maintaining the canals). G takes the simple form of $g - dR$, where g is the replenishment rate of water and d is the natural loss rate due to evapotranspiration and seepage (physical sciences). The amount of time farmers spend in working in this irrigated system (rather than working outside to earn labor wage) is modeled by replicator dynamics (evolutionary game theory). The agency imposes fee/taxation level of C on the farmers and then

invests a fraction y of their revenue in infrastructure maintenance, while having an exit option to leave this particular system for another system (economics, governance, institution). The system can be sustained in a long run if it stays within boundaries—tipping points. One such tipping point condition is

$$y > \frac{\theta}{C} \left(\frac{(1-C)\phi}{(1-C)\phi - \rho} \right),$$

where θ , ϕ , and ρ are dimensionless groups that reflect the relative decay rate of the canals relative to maintenance efforts (physical factors and governance), the potential income that farmers derive from the system relative to labor wage earned from working outside (socio-economic factors), and the natural loss rate of water relative to maximum extraction rate (environmental and social factors). The above expression is a clear example of how independence between human and natural systems may be combined to define a tipping point in a coupled natural-human system. It is one thing to stipulate that the interplay between social and environment factors is nonlinear; it is quite another to clearly specify what the nature of that nonlinearity looks like. Anticipating objections, let us be the first to say that migration process is complex and such neat expressions of its tipping points may not be possible. But that is not to say that one cannot extract insights and guidelines on the actual tipping points from these clear expressions. That is how we intend to use these low-dimensional, dynamical system models: to provide guidelines on how to meaningfully add realistic complexity to the models, while maintaining some of these nonlinearities, and theory and to develop potential metrics from a migration early warning system.

In addition to raising the prospect of more concrete identification and better understanding of the tipping points in migration processes, the type of node-level dynamical system model described above can be helpful in those tipping points that cause emergence of new migration patterns. For example, by modeling the coupled dynamics at different locations, we will be able to detect the “saturation” at the initial receiving locations, be it political (e.g., migrant quota) or environmental (e.g., lack of critical resource), which causes a “secondary push” that forces migrants to seek new destinations, hence a new migration pattern. This modeling approach also enables us to explore the dynamical behavior of the modeled system beyond the system’s historical range, including how the system responds—including migration—when subject to multiple, simultaneous environmental stressors. Such ability is instrumental to designing intervention and contingency plans for migration in the changing environment we are in.

2.2 Linkages

Each pair of sending and receiving locations are connected with different types of linkages, such as distance, natural barrier, political barrier, historical patterns, and social ties. Let M_{ij}^l denote relationship l between sending location j and receiving location i . These linkages will be combined with the node-level internal dynamics and effects of environmental forcing and—through analyses from economics, game, and network theories, including recent advances in multilayer networks (e.g., De Domenico et al. 2013; Kivelä et al. 2014; Boccaletti et al. 2014)—translated into pushes and pulls that determine population movements [B]. Besides environmental changes, migration is driven by economic, political, social, and demographic drivers (Black et al. 2011). We will determine the relative importance of these different linkages (e.g., weight distribution, nonlinear relationships and/or substitutability among them), which can

vary from node to node. This knowledge will be instrumental to our ability to anticipate how new migration patterns may emerge.

Clarification on the term ‘network.’ In the migration literature, the term ‘network’ often refers to a social network or a network of social and interpersonal ties. In this project, unless it is specified as a social network, the term ‘network’ will retain its original, more general definition, which describe a relationship (edges or links) between entities (nodes—in this case, locations).

In this project, we will use a multiplex network as an integrator and evaluator of existing theories. Each layer—or type of relationship among considered locations—captures a factor related to an existing theory or theories into the models to be developed in this project. For example, wage differences will be used to examine predictions made by the neoclassical economics of migration; trade networks will be used to examine predictions made by the world systems theory of migration; existing migrant population in receiving locations will be used to examine prediction made by the network theory of migration; and so on.

These layers of the multiplex network will be analyzed in conjunction with other layers that are known to be important/relevant to migration. These include, but are not limited to, physical proximity, immigration policies/agreements, shared religious belief, history of past and present migration, diplomatic ties, existence of humanitarian organizations to aid migrants, etc.

Let $f_{ij}(\{M_{ij}^l\}; \{\beta_j^l\})$ be the function that maps the set of relationships $\{M_{ij}^l\}$ to a *migration predictor*, where β_j^l is a “coefficient” that population j “attaches” to relationship of type l when making a decision to migrate. This is where we will investigate how these coefficients β_j^l will be attached to each type of relationship (i.e., each layer of the multiplex network), ranging from standard techniques and novel ones.

We will start with such a standard technique as f being a linear combination, that is,

$$f_{ij}(\{M_{ij}^l\}; \{\beta_j^l\}) = \sum_{l=1}^L \beta_j^l M_{ij}^l.$$

This could be entered into a multinomial logistic regression, which translates these migration predictors into the probabilities that a person in location j will migrate to location i . Despite its being a standard technique, we can learn initial insights from applying it to the case studies (Section 2.5). The index j in β_j^l accommodates the (likely) possibility that the coefficients vary from place to place. The coefficients may also vary among different types of environmental forcing (e.g., sudden shocks *vs.* gradual changes). Even within the same episode of migration, these coefficients may vary through time, i.e., $\beta_{j,t}^l$, which can be used to infer the interplay between different migration theories (Section 2.6).

After this initial investigation, we will explore other nonlinear combinations of these relationships. For example, some population may see one type of relationship seen as, to some extent, substitutable for another. For expositional clarity, let us consider two types of relationships ($l = 1, 2$). In such a case, we may have

$$f_{ij}(\{M_{ij}^l\}; \{\beta_j^l\}, \{\gamma_j\}) = [\beta_j^1 (M_{ij}^1)^{\gamma_j} + (1 - \beta_j^1) (M_{ij}^2)^{\gamma_j}]^{1/\gamma_j},$$

where γ_j reflects the degree of substitutability between M_{ij}^1 and M_{ij}^2 . Other nonlinear combinations, including general basis functions such as splines, will be explored based on

literature review and expert opinions that will be gathered from experts on our team and others during our annual workshops (Section 2.6).

More complicated mapping between $\{M_{ij}^l\}$ and migration outcomes may be obtained by Bayesian inference techniques. This approach offers needed flexibility because while some structural relationships are directly inferred from available data, others must be obtained through expert elicitation. Recognizing the inherent uncertainties in both approaches in determining the relationships, it is useful to think of the causal model as a dynamic Bayesian network, where the state of the population system evolves across space and time given an underlying model. For state variables that are categorical or ordinal, their dependence on other variables can be prescribed through conditional probability tables or multinomial regression; for continuous variables, Gaussian process models (Neal, 1997; Gelman et al, 2014, Datta et al, 2016) that provide a flexible and parsimonious way to represent linear and nonlinear functional dependence will be used.

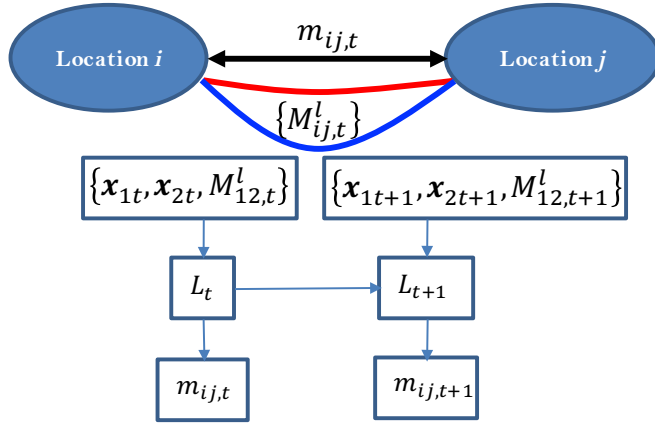


Figure 2. Illustration of an underlying model structure for a migration pair: $m_{ij,t}$ is the migration rate between locations i and j in year t ; \mathbf{x}_{jt} a suite of demographic, environmental, economic, conflict and related variables for location j , which may have structural relations among themselves (e.g., economic outcomes depending on environmental variables); and L_t a latent variable, which can be used in uncover structural dependence between migration and other variables, including the tipping points that trigger migration

Fig. 2 illustrates potential use of latent variables in our analysis. In this example, we consider a Hidden Markov Model for the \mathbf{x}_{jt} , such that the joint distribution of the \mathbf{x} variables for both locations informs a multi-state latent variable L_t which in turn is assumed to have Markovian dynamics. The migration $m_{ij,t}$ is then modeled as a function of the latent state assigned to the current condition.

Consider a model with 2 latent states, i.e., $L_t = 0$ or 1. We can think of the first state as one in which there is a very low migration rate with modest variance, while the second state corresponds to conditions in which a high migration rate may occur from location 1 to location 2, as determined by the \mathbf{x}_1 and the \mathbf{x}_2 values at that time. Inference on the entire dynamics is feasible given data (complete or incomplete) in a Bayesian framework by recognizing that the joint probability distribution of observations may be written as:

$$f(m_{ij,t+1}|L_{t+1})f(L_{t+1}|L_t, \mathbf{x}_{1,t+1}, \mathbf{x}_{2,t+1}, \{M_{12,t+1}^l\})f(\mathbf{x}_{1,t+1}, \mathbf{x}_{2,t+1}, \{M_{12,t+1}^l\})$$

In this example, it is useful to think of the Latent variable states as underlying regimes for the migration rates and their covariates and the Markovian process for the latent variables as informing us about conditions under which a high probability of persistence in the same state exists, or a state transition is likely to happen. A high probability of state transition could then

indicate a tipping point for the existing state, and the associated α and migration rates then inform us about the conditions that we may hypothesize from the low-dimensional models.

The example model described here can be extended to consider more complex dependence structures across the variables and inference on model complexity can be done in a structured manner. Model parameter estimation can be done in a Hierarchical Bayesian framework across the potential network of locations that are valid migration pairs, thus reducing the uncertainty in estimation and providing for a structured exploration of model structure

Expert elicitation of the conditional probability tables and/or of potential relationships between causal variables and migration can be done in a Bayesian framework using the Delphi process following Bijak and Wiśniowski (2010) and Abel et al. (2013). The Delphi process will be conducted during the annual workshops (Section 2.6). Consistent with the formulation of the low-dimensional models, we will generalize the Bayesian networks described by these conditional probabilities to a dynamic network so that temporal evolution of the population dynamics over a topological network can also be modeled.

For the estimation of a causal model regarding a decision to migrate, recent economics literature (Alem et al, 2016) provides examples of the use of instrumental variables with probit regression. These methods can also be put in a Bayesian inference framework as shown by Lancaster (2004), Kleibergen and Zivot, E. (2003), and Lopes and Polson (2014). An accessible example of fitting a Hierarchical Bayesian model for population movements inside Japan as determined by a set of variables including shocks, population demographics and economic variables is provided by Oblander (2017). A more detailed Bayesian model for probabilistic population projections considering international migration that incorporates a constraint that the net global migration is zero is provided by Azose and Raftery (2015). The key point of the Hierarchical Bayesian models as applied to multiple locations for structural estimation of the coefficients of a postulated relationship is that if a multi-level model is properly defined, a “partial pooling” of information across all locations is possible such that one can infer the posterior distribution of the model parameters at each location, while constraining them to an average and a variance of the parameters estimated across locations. We have considerable experience developing such models in the climate context (Wang et al., 2017; Lima and Lall, 2009, 2010; Yuan et al 2016; Sun et al, 2015 and Steinschneider and Lall, 2015), and will explore how to develop such models for the case studies where adequate data to parameterize the relationships is indeed available.

2.3 Environmental forcing

We will develop a general spatio-temporal model for environmental changes that are likely triggers of a migration episode. These environmental changes will impact both the nodes and linkages. It is important to note that the response of a population to gradual changes (e.g., shifting rainfall patterns induced by climate change) and sudden natural disasters (e.g., hurricanes and earthquakes) will likely differ. The former can be seen as fluctuating disturbances with an overall trend, whereas the latter a big shock to the coupled system. Therefore, the models and analyses to capture the population response in these two cases will be different.

Sudden shocks

Natural disasters such as earthquakes and hurricanes are simply difficult to predict. For example, the National Hurricane Center clearly states that prediction of a hurricane path is not reliable

beyond 72 hours. Furthermore, the extent of the damage, especially the number of people who would be displaced by such natural disasters, very much depends on many local features, e.g., local topography and capacity of different types of local infrastructures (think levees, power grids, and hurricane shelters). As such, for this type of environmental forcing, it is more effective to develop a capacity to produce a library of migration scenarios. That is, the analysis should address the question of the type: “If N people in need of migration is created at location X , what is the resulting migration patterns and how do those patterns change over time?” The patterns corresponding to N of, say, 1,000, 10,000, or 100,000 would differ. People will exhaust the most convenient options first and then move on to the next options. As N grows, the migrants may push the limit of resources needed to support them at an initial receiving location, forcing the subsequent migrants to seek new destinations, thereby giving rise to new migration patterns that are much different from the patterns at low N . To this end, coupling people and natural resource dynamics explicitly—as discussed in Section 2.1—will offer better mechanistic understanding that underlies such emerging migration patterns and improve our predictive capability.

Gradual changes

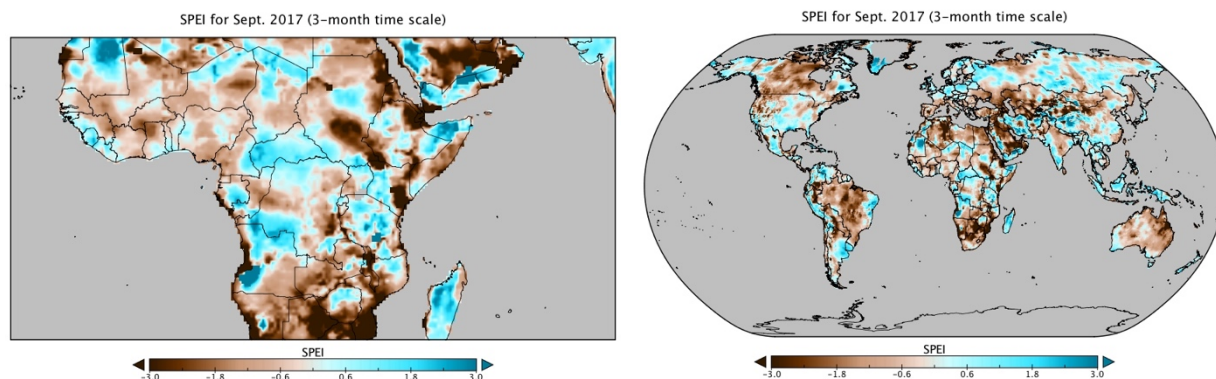


Figure 3. Maps of the Standardised Precipitation-Evapotranspiration Index (SPEI), a drought index. The data were obtained from <http://spei.csic.es/map/maps.html#months=1#month=8#year=2017>.

Gradual changes that affect migration include climatic attributes that influence regional hydrologic attributes, as well as nutrition and infrastructure loss, including the persistence, intensity and spatial extent of the episodes of these environmental changes. We have a reasonable capability to predict these trends and how they may lead to masses of migrants (e.g., Werrell & Femia, 2017). These gradual changes include shifts in extremes like drought, which are expected in the coming decades (Cook et al., 2014); see **Fig. 3** for an example of potential climate drivers of migration (i.e., a drought index) that we will use in the models. We will use climate model projections as a tool to understand these gradual changes, recognizing that they continue to undergo continuous improvement and verification (Ring et al. 2017). Retrospective and prospective global simulations from these models (e.g., the Coupled Model Intercomparison Project (CMIP) 5 and the upcoming CMIP6) will be considered as scenarios for spatio-temporal climate extremes that may be pertinent to migration dynamics. This will allow us to integrate information that suggests shifting risk associated with climate extremes, as was recently done for drought risk in the US Southwest (Cook et al., 2015). We will also examine the co-occurrence or repeated occurrence of climate and other trigger events (e.g., earthquakes) in a specified time

window. The ensemble of these events would be used to define the temporal dynamics of the potential for migration and the subsequent extent (magnitude and duration) of migration, as well as the potential for return migration [H].

The environmental forcings themselves will have impacts on population at a given node. In particular, the nutritional and health state of the population will be directly affected, which, in turn, influence the forces driving migration. In particular, we will include nutrition analyses using an agent-based, supply-demand model with food storage dynamics and trade policies (Puma et al., 2015; Marchand et al., 2016; Schewe et al., 2017; Otto et al. In prep.) to understand food energy and macronutrient availability under out-of-equilibrium conditions.

We will combine the capability to predict these gradual changes in environmental and related forcings with the dynamical systems models developed at the node level to develop quantitative metrics for *migration potential* at each node. The outcome will be a map of these migration potential metrics that can change through time, providing a basis for developing an early warning system for migration and contingency plans to address/accommodate it (Lopez-Lucia, 2015).

Climate extremes and multiple stressors

Climate change and its manifestation as a higher frequency, intensity and duration of extreme weather and its impacts (e.g., floods, droughts, heat waves, hurricanes, fires) is one of the environmental drivers of concern. While future projections of climate change impacts are quite uncertain, they have been well studied over the last two years, and uncertainty distributions accounting for GHG emissions and for climate models have been developed by the IPCC. The associated simulations are publicly available for the key variables of interest, through <http://www.ipcc-data.org/>. Extensive analyses of the historical climate simulations of each model have also been done, and significant biases in both temperature and precipitation have been noted. Many methods of correcting these biases so these simulations can be used with hydrologic, agricultural, economic and other models for future projections. A vast literature on bias correction has resulted (Iizumi et al, 2017). The typical strategy is to “correct” the biases in the basic statistics of each variable at each location, over the historical period and to then apply the same correction to the future projection. This is an intellectually unsatisfying solution to the problem of model error. Even after these bias corrections, the persistent expression and spatial correlation structure of the occurrence of climate extremes is usually not reproduced very well relative to the historical data. Consequently, the direct application of these climate projections for the future has to be approached with caution.

In the context of advancing a theory of migration induced by environmental factors, we can still use such projections or we can rely on the extensive climate re-analyses data sets that provide global coverage for the 20th century. The ECMWF and NOAA re-analyses extend back to 1900 and 1870 respectively and may provide a rich data set to explore the co-occurrence of drought or flood or other climate induced disasters—i.e., multiple stressors—at multiple locations over a time window of interest. For migration dynamics, consider that there is a potential migration network from location A to a set of other locations. Given that migration is a dynamic process, if climate induces out-migration from A, then one needs to know which of the potential in-migration locations is also likely to experience climate induced disasters over some time window. There is evidence (Bonnafous et al, 2017a,b) that climatic extremes exhibit significant multi-year clustering at a given geography and are significantly spatially correlated as well. There are well established inter-annual (e.g., ENSO), decadal (NAO, AO) and multi-decadal

(PDO, AMO) climate models with very specific spatial expression. In fact, ENSO has widespread global drought and flood impacts that translate into famine and disease over much of the world. Since we have over a century of relatively good data from re-analysis on these extremes, it is possible to do some inference on the nature of these teleconnections and simulate simultaneous time series of extreme occurrence and intensity at all locations of interest. Examples of such simulations are provided in Kwon et al (2007) and Erkihyun et al (2016, 2017). These methods can be combined with the copula based methods in Lall et al (2016) to provide appropriately correlated spatial scenarios on multiple climate variables. We propose to develop such an approach using re-analysis data and then perturb the frequency and intensity of the events based on the direction of the IPCC scenarios at the locations of interest. This would provide an approach for analyzing migration potential and scale given both the cache of historical space-time climate structure and potential changes in the future. Since we will simulate climatic time series for each variable (e.g., precipitation, temperature, drought index, extreme rainfall index, heat wave index) of interest gradual changes as well as shocks will be embedded in the simulation and used in the migration models.

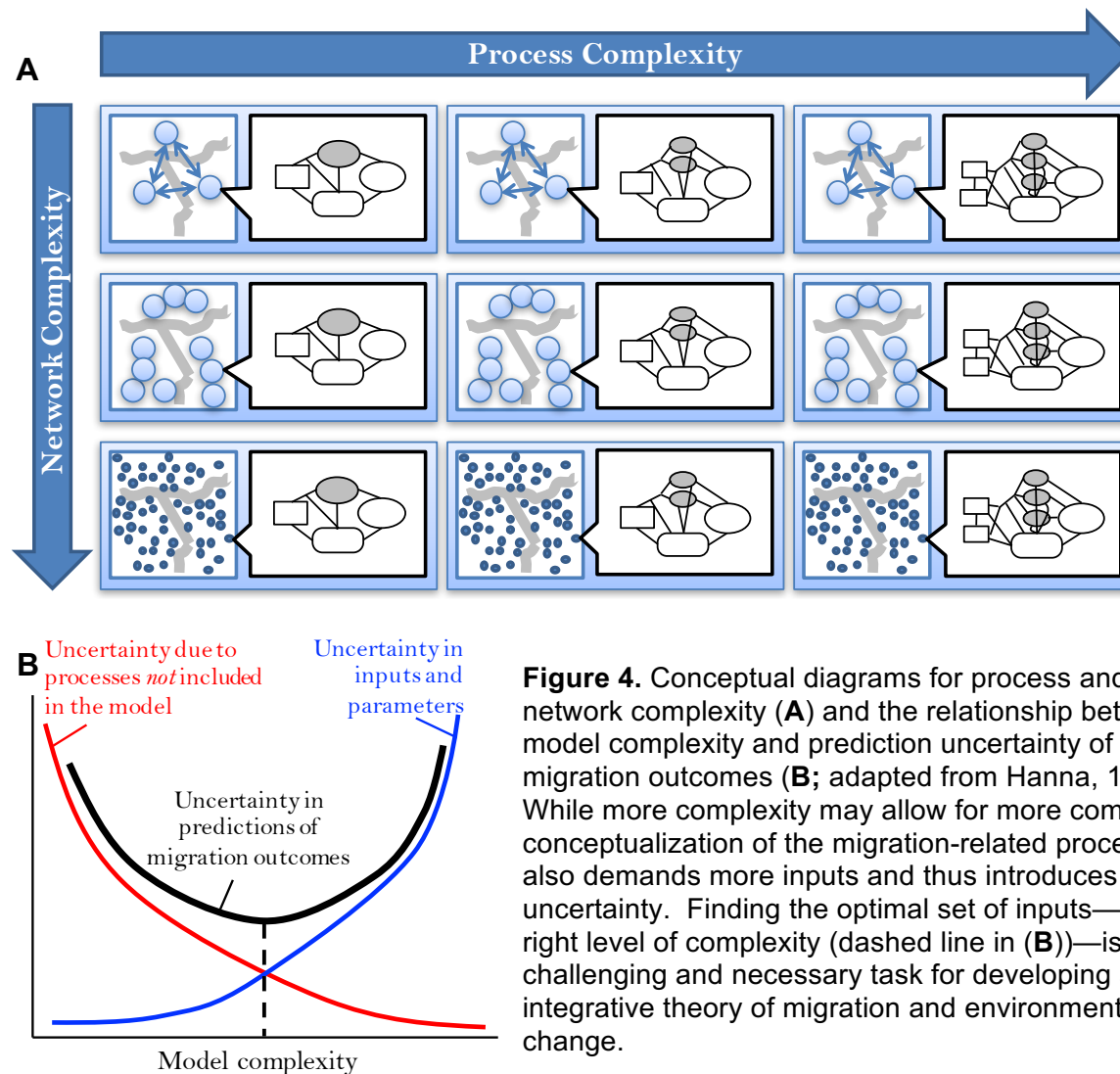
2.4 Growing complexity

Building on predictive capability and insights from the models with simple settings, we will pursue a systematic strategy for model elaboration to add to model complexity in order to evaluate the implications for uncertainty in migration outcomes. Our elaboration entails two types of complexity: “process complexity” via incorporation of more model processes (parameters) and feedbacks (e.g., multiple critical resources and interactions of subpopulations within a population [E]); and “network complexity” via spatial resolution or number of sending and receiving locations and thus the number of linkages between them (**Fig. 4**).

Finding the right level of complexity. More complexity does not always lead to better results or greater predictive capability (Nihoul, 1994; Fisher et al., 2002). While more complexity may allow for more complete conceptualization of the processes of interest, it also demands more parameters and thus introduces greater uncertainty (Hanna, 1993; Snowling and Kramer, 2001). Throughout this elaboration process, global sensitivity analysis (Saltelli et al. 2008) and Bayesian network analysis (Murphy and Russell, 2002; Pearl, 2014) will be applied to detect significance and benefits of these added details with the aim of keeping only the smallest possible number of factors in the theory. In complex models such as those integrating environmental and social factors to predict migration, there is greater potential for many sources of uncertainty with non-linear effects and interactions (Leamer 1990). In such cases, “global” sensitivity and uncertainty analysis (GSUA) is required, for only it can evaluate non-linearities and interactions among multiple sources of uncertainty (Saltelli, et al. 2000, 2004, 2008). An important benefit of GSUA is that it provides statistics of importance of each source of uncertainty on their own (direct effects) or through interactions (higher order effects), so it informs not only the overall predictive uncertainty with increasing model complexity, but how the model components control the model predictions. This in turn supports the identification the need for refined data monitoring plans around the important processes controlling the system, gaps or inconsistencies in the conceptual model, etc. Bayesian networks (Murphy and Russell, 2002; Pearl, 2014) provide a framework to investigate how different groups may react to spatio-temporal stresses and opportunities and how the spatio-temporal factors driving choices may co-evolve (Uusitalo, 2007, Madadgar and Moradkhani, 2014). They permit the integration of expert knowledge, conceptually based causal

structures as well as empirical inferences from structured data analysis (Gelman et. al., 2014). This knowledge will then guide us to meaningfully add complexity to our models.

A useful framework for identifying the optimal model complexity (Hanna, 1993; Snowling and Kramer, 2001) considers the systematic evaluation of models of increasing complexity for a given problem (**Fig. 4B**). Muller et al. (2011) operationalized and tested the framework by quantifying uncertainty based on the 95% confidence interval of the model predictions (Y-axis in **Fig. 4B**) of models of increasing complexity (represented by increasing number of input factors, X-axis in **Fig. 4B**), as obtained from multivariate global sensitivity and uncertainty analysis. Lagerwall et al. (2016) further tested the framework to identify the optimal complexity of an agent based model predicting invasive species migration at a moderate level of model complexity.



Predictive power and theoretical development. Deeper understanding of the migration theory and greater predictive power of the models that we will develop will go hand in hand—up to a point. With the advent of more computational power, big data, and machine learning techniques, it is possible to develop a model that exhibits great predictive power, but with explanatory

variables that are difficult to couch under a coherent theory. In addition, these models with great predictive power may not lend themselves well for shedding light on the nature of the tipping points that trigger migration episodes and may be of little use when underlying processes and factors change beyond their empirical/historical ranges. This is particularly true for migration induced by environmental changes under ongoing climate change. Dynamical system approaches enable us to have a better grasp on how different variables and parameters interplay and constitute such tipping points and explore theoretical ranges beyond those observed in the available dataset. It is therefore essential to combine different approaches in developing the models and the theory, while being cognizant of the strengths and weaknesses of each approach. And this is what we will do in the project.

2.5 Case studies

We will apply the modeling framework to different case studies below, which are of different spatiotemporal scales and exposed to different environmental changes. Based on the lessons learned and the models and theory developed from these case studies, we will select additional cases studies and apply the models and theory to them. Our hypothesis is that the model structure would change across different temporal and spatial scales in these cases, and this knowledge is crucial to the theory development.

Case 1. Hurricane Mitch and the development of Honduran-U.S. Migration patterns

Hurricane Mitch, the second deadliest hurricane on record, devastated Honduras and much of Central America in 1998, causing historic levels of flooding and the displacement of nearly three million people in Honduras, Guatemala, and Nicaragua (Ensor 2009). Prior to Mitch, Honduras did not have a long tradition of migration to the United States, with migrant pathways confined to shipping links between the ports of La Ceiba and New Orleans and a few regions with ties to southern California. Following Hurricane Mitch, however, Honduras-U.S. migration accelerated due to the devastation and to the U.S. government offering Temporary Protective Status (TPS) to Hondurans displaced by the hurricane. This led to new migrant pathways from Honduras, with many migrants focusing on the U.S. southeast for its abundant employment opportunities in food processing, construction, and services. States like Mississippi, Alabama, and North Carolina became new migrant destinations, where Hondurans not only found work but also established small businesses like convenience stores, hair salons, and auto repair shops. These developments created a context for continued Honduran-U.S. migration even after TPS was no longer offered to new Honduran migrants, facilitating undocumented migration through existing social networks as well as a “migration industry” of migrant smugglers, false documents and people who specialized in Individual Tax Identification Numbers, transportation, and other services. Thus, this case could help us build our model by providing details such as the role of the state in stimulating migration, the role of employers in influencing migrant destinations, and the importance of settled populations of other Hondurans, including the presence of Honduran-run businesses.

Case 2. Hurricane María and Puerto Rican migration to Florida.

Following the devastation of Hurricane María, a category 4 hurricane that destroyed Puerto Rico’s communication systems and other infrastructure, most Puerto Rican families and communities had to depend on U.S. government and military aid for food, water, security, and other basic goods and services. During the initial period of recovery, as the distribution of these goods and services was slow to reach across the island, many Puerto Ricans fell back on a long-

time response to local suffering: migration to the U.S. mainland. As soon as the airports opened, thousands of Puerto Ricans fled to the traditional U.S. mainland destinations of Florida, New York, and Illinois. During the first two weeks of October 2017 alone, over 27,000 Puerto Ricans migrated to Miami, Orlando, Tampa, and other Florida cities, swelling an already large Puerto Rican population in Florida (<http://www.npr.org/2017/10/13/557108484/-get-us-out-of-here-amid-broken-infrastructure-puerto-ricans-flee-to-florida>). Over the two years prior to the hurricane, largely due to the economic crisis on the island, over 80,000 Puerto Ricans per year moved to Florida, up from around 50,000 per year before the crisis.

Puerto Rican-Florida migration thus represents a case of gradual crisis over two years followed by a single, catastrophic event that triggered a far more massive, immediate migration flow. As such, Puerto Rico-Florida represents a case of both sudden and gradual environmental events—the one natural and the other social—stimulating different levels of migration. Puerto Rico-Florida migration is also more recent than migration to other parts of the U.S. mainland, such as New York and Chicago. The shallower history of Puerto Rican migration and settlement in Florida will make it somewhat easier to locate and develop data sources to aid in modeling the relationships among gradual and immediate pressures forcing emigration from Puerto Rico and the development of social networks, economic opportunities, and other relevant phenomena in Florida.

Case 3. Famine-induced migration in Africa

Currently, a humanitarian crisis is emerging particularly in Somalia, Nigeria, South Sudan, and Yemen, where over 30 million people need food assistance and more than 10 million of them are on the brink of famine (Cadre Harmonisé, 2017; FSNAU, 2017; IPC-South Sudan, 2017; IPC-Yemen, 2017). Conflict between armed groups in a setting with low social and environmental resilience is the main common driver. In particular, Somalia continues to suffer from the effects of a protracted civil war and has endured recent extreme drought, which has already led to the displacement of 680,000 people (OCHA-Somalia, 2017). In northeastern Nigeria, the regions has been devastated by the cumulative impacts of the eight-year conflict between the Nigerian Government and militant Boko Haram, which has led to mass displacement, human rights violations, and now severe food insecurity (UNHCR, 2017). More than 5.5 million people are severely food insecure in South Sudan due to fighting that has forced more than 3 million people from their homes since the conflict erupted in December 2013 (UNHCR South Sudan, 2017). Finally, Yemen is a country on the verge of collapse. Civil war has completely undermined Yemen's food supply, with a 17 million food-insecure people (IPC-Yemen, 2017).

There is an increasing body of empirical evidence that links the largest outward flow of refugees with countries experiencing armed conflict and high food insecurity (WFP, 2017). These four countries already have substantial numbers of internally displaced persons (IDPs) in 2016: 2.0 million in Nigeria, 1.1 million in Somalia, 1.9 million in South Sudan, and 2.0 million in Yemen (IDMC, 2017).

Fig. 5 shows a baseline scenario for potential, global refugee movement. We identify likely refugee destinations using historical asylum seeker data for the years 2011-2015 (UNHCR-PSD, 2017). The assumption is that refugee flows tend to follow previous migration and refugee networks (Neumayer, 2005). In this basic example, we assume that 10% of the severely food-insecure population in each of the four countries leave as refugees. Also, we take refugees exclusively from the IDP population, which is clearly an end-member scenario. In this simple

example, refugee flows are as follows: 1.6 million from Yemen, 0.9 million from Nigeria, 0.6 million from South Sudan, and 0.3 million from Somalia. The destination for the most number of refugees is Jordan. Main destinations for refugees in Africa include Uganda, South Africa, Egypt, and Kenya. Longer distance destinations include Italy, Sweden, Germany, Malaysia, the US, and the UK. To put this scenario in perspective, the United Nations High Commissioner for Refugees has estimated that there are presently ~194,000 Nigerian refugees from Nigeria in Cameroon, Chad, and Niger (UNHCR, 2017); this is ~20% of our total Nigerian refugee outflow.

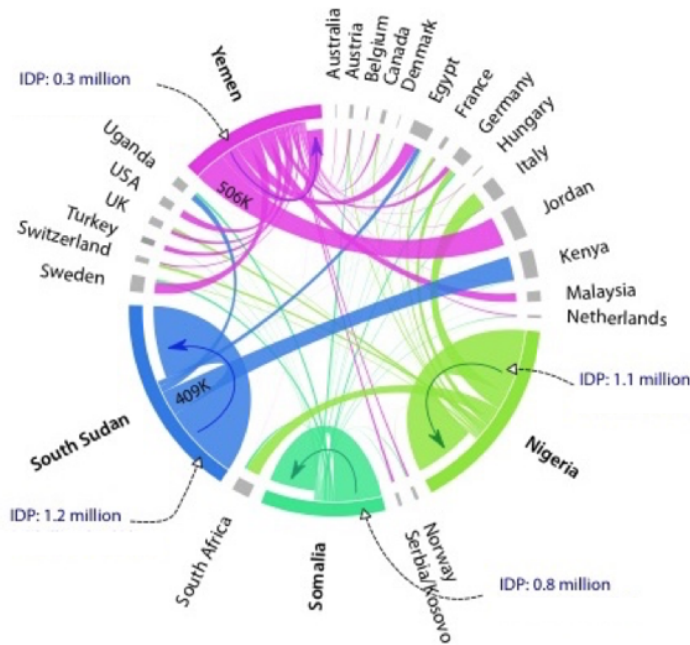


Figure 5. Potential refugee flows for a baseline scenario in which the refugee outflow rate is 10% of the severely food-insecure population in each of the four countries. We take the refugee population from the existing *internally displaced person* population (an end-member scenario) and use the average pathways of asylum seekers for the years 2011-2015. Note: Bilateral links are included only if the historical refugees flows were on average greater than 10,000 people.

This case study is an example where a region experiences both environmental stresses and conflict, which manifests itself by undermining the food supply. The consequences of this food insecurity extend beyond the regions, leading to global scale impacts. The proposed modeling approach would allow us to examine scenarios for refugee movement and its implications both regionally and globally. In fact, besides food insecurity, Nigeria and Yemen are dealing with major outbreaks of cholera. Our modeling structure is designed to be flexible enough to explore these impacts of migration and refugee movement.

Case 4. Syrian refugee crisis as a test case

The Syrian refugee crisis is an interesting case of extensive international migration that has been well documented. While there are numerous factors at play, the 2007–2010 drought is thought to have contributed to the outbreak of conflict (Kelley et al., 2015). Widespread failure of crops and mass migration from rural farming areas to urban centers occurred due to the severe drought (Kelley et al., 2015).

There is considerable data compiled from a number of sources, most notably the United Nations as noted below. One notable data set contains dyadic ties between origin countries (sources) and recipient countries of asylum or residence (sinks) with total refugee numbers from 1979-2013. Such data would facilitate the ability to model the changes in flows of refugees from sources to sinks over a considerable period of time. In addition, it would allow for the building of multi-

layer networks involving both sources, in this case Syria, and sink countries where multiplex networks can be developed involving trade relations, economic relations, treaty relations, military relations, among others. This would provide an extensive test case for benchmarking model predictions over time, including the saturation effect and secondary pushes discussed in Section 2.1.

Additional case studies. Based on the lessons learned and the models and theory developed from the above case studies, we will select additional cases studies and apply the models and theory to them. We will explore the growing number of expanding databases on migration. Examples of these databases are listed below:

- The United Nations Global Migration Databases (UNGMD) (<https://esa.un.org/unmigration>)
- The Organisation for Economic Cooperation and Development (OECD) international migration database (www.oecd.org/els/mig/keystat.htm)
- The Migration Polity Institute (MPI) database (<http://migrationinformation.org/datahub>)
- The International Labour Organization (ILO) migration database (<http://www.ilo.org/global/statistics-and-databases>)
- The United Nations High Commissioner for Refugees (UNHCR) database (www.unhcr.org/statistics)
- World Bank's Global Bilateral Migration Database (<http://data.worldbank.org/data-catalog/global-bilateral-migration-database>)
- Internal Displacement Monitoring Center (IDMC) (<http://www.internal-displacement.org>)

2.6 Theory development

Many migration theories exist, but in a somewhat fragmented fashion: they tend to focus on certain aspects of the migration processes, e.g., the push factors that create a mass of migrants or how they are assimilated at destinations. In many of these theories, the foci have been on the social and political forces, with interdependence with the natural systems being in the background and ignored; possible causal linkages were inferred through statistical regressions, rather than dynamical couplings, making identification and/or quantification of tipping points difficult or impossible. What is needed, we argue, is a modeling platform that can integrate the existing theories—valid and applicable at different time and spatial scales under different contexts—with mechanistic and dynamic representation of environmental changes that impact populations. Only then would we be in a better position to understand and quantify tipping points that would trigger a migration episode and thus be more prepared in either alleviating the root causes of migration or allocating resources for effective humanitarian efforts.

The existing theories—fragmented as they are—provide a good starting point. One must, and we will, be strategic in how to integrate each of them in the sought-after integrative theory. One key criterion in our initial triage is the types of environmental changes under consideration. In response to sudden shocks (e.g., hurricanes, floods, earthquakes, wildfires), migration is often forced and involuntary; as such, we will be borrowing theories from the forced migration literature, with keeping in mind the differences between the environmental sudden shocks and the usual root causes in that literature (e.g., war and civil unrest). In response to more gradual environmental changes (e.g., droughts, famines, sea level rise), people have more time to plan their migration, and existing theories that are based on a variety of economic principles may be more appropriate. Then there is the intermediate cases between these two types of

environmental changes: magnitude and frequency of natural phenomena like hurricanes and wildfires will change with climate change—how should theories for forced and planned migrations be brought together in the integrative theory? For example, differential economic forces and political settings across spatial units of analysis may be seen as priming the system. Once environmental changes—be they sudden or gradual—“push the button” to unleash a mass of people in need to move, they will then propagate through the paths of *their* least resistance in the multilayer networks primed with those economic and political differences.

One of the strengths of our approach is the combination of dynamical system modeling at a node level and the multiplex network that links the nodes, which we expect will produce results exemplified by **Fig. 6**. Recent research (e.g., Qubbaj et al. 2015) has shown that when node-level dynamics is sufficiently nonlinear, many standard network metrics loses their meanings and usefulness as potential predictors. Given that the migration-related dynamics will likely be nonlinear, whatever multiplex network metrics we come up with will be carefully tested, using global sensitivity analysis and Bayesian inference techniques discussed earlier, for their contributions to predicting migration outcomes and developing the integrative theory.

Annual workshops to gather expertise. We will invite experts from relevant disciplines to attend annual workshops from years 1 through 5. Workshop participants will be based on particular themes, e.g., expertise on a given case study. Opinions of these experts will provide a “reality check” to the soundness of our model and theory development. For each workshop, we will ensure a balance of the participants; that is, collectively, they will have expertise in migration theories and modeling of the natural systems and environmental changes under consideration and/or knowledge of/familiarity with the case studies. In inviting the potential workshop participants, we will be cognizant of the nuances in expertise in migration research. As discussed in Section 2.3 above, migration in response to sudden shocks in environmental conditions share some common features with forced migration caused by war or civil unrest; accordingly, we will invite experts on forced migration and/or refugee studies to our workshops. Migration in response to gradual changes in environmental conditions may be akin to migration induced by economic forces; accordingly, we will invite experts on migration theories that are based on economic principles to our workshops. This mindful selection of participants will ensure that the workshops will provide critical and useful information of the development of our models and integrative theory that work across scales and contexts.

To provide a more concrete idea of the workshops, here are some potential invitees of the workshops—which are subject to change as the model and theory development proceeds and the needs for inputs and feedbacks change:

Douglas Massey (migration theory; Princeton University)
Michael J. Greenwood (migration modeling; University of Colorado, Boulder)
Susanne Schmeidl (refugee study; University of New South Wales)
Matthew E. Hauer (sea level rise-induced migration; University of Georgia)
Richard Black (migration theory; University of London)
Solomon Hsiang (climate-conflict modeling; University of California, Berkeley)
J. Marty Anderies (dynamical systems/institutional analysis; Arizona; State University)
Ning Lin (hurricane modeling; Princeton University)
Sally Thompson (wildfire modeling; University of California, Berkeley)
Matti Kummu (water and food security modeling; Aalto University, Finland)

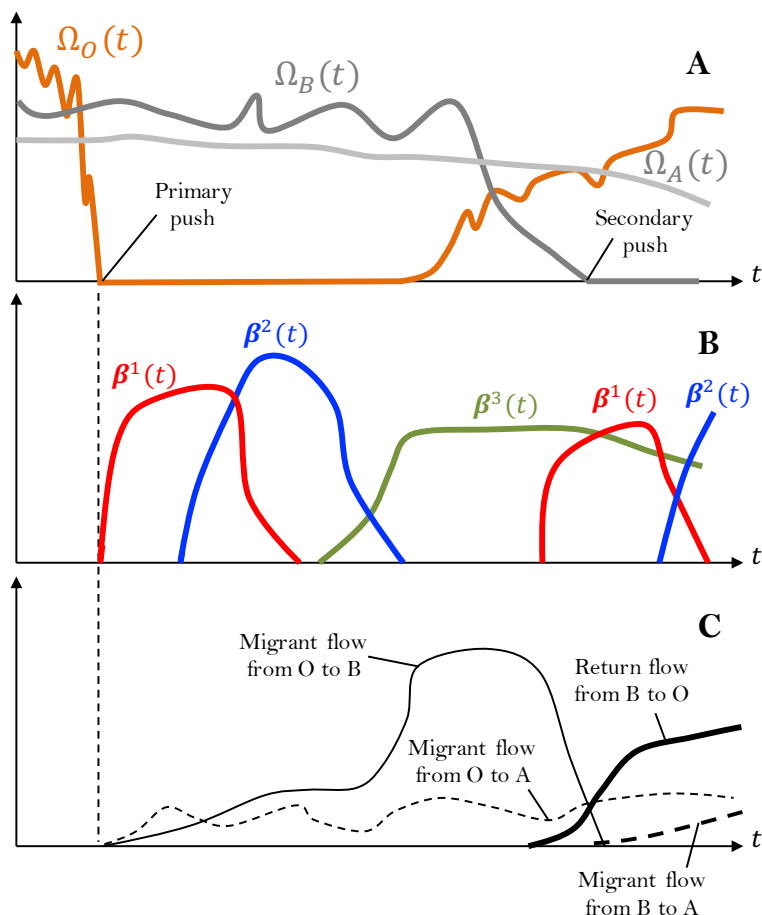


Figure 6. Schematic diagram of how different components will be analyzed together during the development of the integrative theory. $\Omega_j(t)$ represents the migration potential indicator or resilience metric of population or location j ; other variables have been defined in the preceding sections. Changes in $\beta^l(t)$ (Panel B) illustrates the possibility that validity of different theories changes over time during a prolonged episode of migration. The developed integrative theory will recognize and accommodate such temporal features.

As discussed in Section 2.2, we will use different layers of the multiplex network to embody the hypotheses that the different theories and perspectives yield. At these workshops, we will elicit hypotheses from these experts relevant to the case studies and explore how these hypotheses may interact with one another. In the context of the multiplex network, such interactions will be translated to, say, weights assigned to different layers (for linear combinations of the linkages) or, more interestingly and more likely, nonlinear combinations of these different layers.

3. Potential impact on DoD capabilities

This project focuses on fundamental theoretical issues concerning human adaptive responses to environmental change and addresses modeling issues in the integration of human and natural systems. The results of this effort will provide interested DoD entities an enhanced ability to anticipate and predict various types of population movements resulting from both extended and punctuated environmental changes. The ability to forecast such migration events will help foster better logistical responses on the part of DoD to any given event or sets of events.

In particular, we will develop a **general modeling framework with theoretical underpinnings along with guidelines and/or procedures to add complexity to the model to enhance predictive capabilities for a specific case** (which depends on temporal and spatial scales, types

of environmental changes, and other contextual variables). Along with those products, we may be able to also develop an **early warning system**, especially for the migration potential related to gradual changes in the environment for which predictive power of models is greater than for sudden shocks. There are a number of thresholds or tipping points in migration process. Not only does consideration of thresholds is critical to understanding the migration process, but it also opens up the connection between migration research to theoretical advances in early warning systems and resilience literature. These two concepts are closely tied to thresholds or tipping points of a system under focus. Applying them to migration, with an eye for an early warning system, would significantly enrich the insights and value of the proposed research project.

4. Project schedule

Research activity	Yr 1			Yr 2			Yr 3			Yr 4			Yr 5		
	Su	F	Sp	Su	F	Sp	Su	F	Sp	Su	F	Sp	Su	F	Sp
Analyze data on case studies															
Develop prototypes of node-level dynamical models															
Develop a multiplex network modeling platform															
Identify potential workshop participants															
Workshops															
Implement the models to selected case studies															
Growing complexity of the models and theory															
Link the models to an early warning system															
Synthesize findings into an integrative theory															

The tentative schedule of research activities is shown above. Brief description of the schedule of other research activities are as follows. We anticipate to publish project findings in peer-reviewed journals on a regular basis. We anticipate that our faculty members, postdoctoral researchers, and graduate students will present project findings at scientific conferences every year. Since this is a very interdisciplinary research project, the potential conferences are diverse; as such, the conferences have not been specified, and thus their schedules are not known at this point. The nature of the project findings will determine the suitable conferences. The annual in-person project meetings will be scheduled to coincide with the annual workshops; a potential format is having the project team members stay a day or two longer than external experts. Not only is this format cost-effective, but a face-to-face meeting after receiving fresh perspectives, critical feedbacks, and constructive criticism from the external experts will also ensure the relevance and novelty of the research approaches and findings. Project meetings, through such platforms as Skype or Zoom, will be held at least once a month.

5. Management approach

Overall approach to the management of this effort, facilities and subawardees

The project team members (see Project Team section below) will conduct the research and analyses according to their expertise and integrate their results and approaches with the rest of the team members through frequent planning and integration activities (see Coordination and

Interactions sections below). They will also be in charge of dissemination of results through manuscripts and talks at national meetings. The PI and key investigators will monitor implementation within their respective area and discuss any methodological changes necessary to achieve the project's objectives. PI Munepeerakul will work closely with the project team to set the overall scientific direction, coordinate project activities, and ensure integration and dissemination of results to **MURI Research Topic Chiefs and other agency personnel as required**. The PI will coordinate the development of **a draft annual report prior to the project annual meeting**. Feedback during the annual meeting will then be used to finalize the report and to set objectives for the upcoming year.

This project focuses on fundamental theoretical issues concerning human adaptive responses to environmental change and addresses modeling issues in the integration of human and natural systems. It makes use of existing available data and does not involve experiments or fieldwork. Therefore, the requirement for facilities and equipment is minimal. As shown in the supporting documents, the three participating universities, namely University of Florida, Columbia University, and East Carolina University have the resources to complete the proposed research tasks.

Subawards from the lead University (University of Florida) are proposed for Columbia University (led by co-PI Lal), and East Carolina University (led by co-PI Griffith). As described in the sections below (key investigators and Coordination), both subawardees provide complementary expertise and facilities necessary to achieve the proposed project goals.

Project team: Principal Investigator (PI) and key investigators

Rachata Munepeerakul (Principle Investigator) is a complex systems modeler. He has studied systems ranging from freshwater fish communities, urban economies, and coupled natural-human systems. His expertise on network approaches, modeling coupled natural-human systems and modeling dispersal and evolutionary process in explicitly spatial settings has much to offer the development of theory for human migration. *Time committed to this project: 2 summer months, years 1 through 5.*

Michael J. Puma's research is focused on global food security, especially understanding how susceptible the global network of food trade is to natural (e.g., megadroughts, volcanic eruptions) and man-made (e.g., wars, trade restrictions) disturbances using non-equilibrium, network based economic models. *Time committed to this project: 3 calendar months, years 1 through 5.*

Upmanu Lall's research links climate extremes, water, food and energy in a systems modeling context. He brings expertise in Bayesian methods, systems modeling, machine learning and spatio-temporal modeling of extremes to the project. *Time committed to this project: 1 summer month, years 1 through 5.*

David N. Griffith has been studying migrant populations since 1981, including work on guest workers, undocumented economic migrants, and refugees fleeing civil war, natural disasters, and collapsing states and economies. His specific area of expertise related to this project is his work on the relationships among migration, environmental degradation, and economic development. His recent work has traced relationships between labor scarcity and deforestation in Honduras following migrants fleeing the devastation from Hurricane Mitch. *Time committed to this project: 1 summer month, years 1 through 5.*

Jeffrey Johnson's work most related to this project focuses on network models of complex human and biological systems, and their integration, employing various applications of continuous time Markov chain and exponential random graph models to the study of trophic dynamics in food webs, particularly as it relates to the interplay between food web dynamics and human behavioral networks. He has also worked on understanding the drivers of conflict, both within and between human groups. *Time committed to this project:* 1 summer month, years 1 through 5.

Rafael Muñoz-Carpena is an expert in uncertainty and global sensitivity analysis of complex models, especially complex hydrological and ecological models. His expertise in global sensitivity analysis will help determine the right level of complexity of the models. *Time committed to this project:* 1 summer month, years 1 through 5.

In addition to the PI and Key Personnel, this project also supports 6 PhD students and 3 postdoctoral researchers as part of the proposed research team.

Detailed on-going activities and dedication to these are provided in the Current and Pending support documents for each co-PI submitted with this proposal. The research team members have no existing conflicts of interest with this proposal.

Division and coordination of research activities

The research activities have been built around the complementary set of expertise of the project team members. In addition to coordinating different proposed research activities, PI Muneeppeerakul will take the lead on developing and analyzing the low-dimensional, dynamical models for node-level dynamics, focusing on understanding the nature of migration-inducing tipping points and using the resulting insights to meaningfully add realistic complexity to the models. He will also take the lead in integrating findings from different research tasks into a coherent theory. Co-PIs Lall and Puma will be responsible for generating scenarios of environmental forcing that will drive the migration models. Co-PI Lall will also be responsible conducting Bayesian inference analyses for uncovering the structural relationships among variables and migration response and for model parameterization. Co-PIs Johnson and Griffith have worked together for many years on social science topics including migration and conflicts. Co-PI Johnson will take the lead in applying network approaches, which he has used to investigate other human and biological systems, to model the dynamics of human migration. We anticipate there to be great synergy between his and Lall's approaches, which may lead to novel methods for modeling and analyzing migration dynamics. Co-PI Griffith has intensive experience in research on the relationships among migration, environmental degradation, and economic development. He will provide guidance on compiling, processing, and analyzing human movement data to address various issues in different migration theories. Co-PI Muñoz-Carpena will be responsible for identification of the optimal model complexity through global sensitivity analysis of the models, as described in Section 2.4.

Plans to manage interactions among research team members

PI Muneeppeerakul will organize regular meetings, at least monthly, via platforms such as Skype or Zoom. Co-PI Puma will organize data storage on Google Drive and use Google Earth Engine/Cloud Platform to both coordinate our modeling efforts and enhance our capacity to disseminate results to partners. We will hold project meetings annually, which will coincide with the expert workshops (e.g., the project team members may stay longer than the workshop

participants). A meeting in person after receiving fresh perspectives, critical feedbacks, and constructive criticisms from the workshop participants will be a great way to revitalize and re-examine the directions and approaches of the project. This will ensure that the project will make meaningful contributions to the field. The 6 PhD students and 3 postdoctoral researchers will support the specific activities of their lead PI and participate as integral members of the team in all meetings in activities. Co-advising by co-PI and institutions will be promoted to better integrate the activities of the team and advisees, from identification of the research activities, to development and sharing of data and resources among team members. A regular item in all meetings and workshops will be presentation and critical review of work by the project mentees to ensure the fulfillment of the project goals.