i-CONN Network

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 859937

Deliverable D1.2 Report

Delivery and reporting of refined minimal models to each application scenario, incorporating relevant details about the respective systems in the minimal models from D1.1

Deliverable D1.2

Delivery and reporting of refined minimal models to each application scenario, incorporating relevant details about the respective systems in the minimal models from deliverable D1.1.

Marc-Thorsten Hütt, Jacobs University, Bremen, Germany

Background

A strength of the i-CONN project is the coexistence of minimal models, detailed models and data-driven research. In this report the activities within i-CONN concerning the design, analysis and usage of minimal models are summarized.

A minimal model (also: 'toy model' or 'stylized model') is the mathematical representation of a dynamical (and often complex) system, which employs the smallest number of parameters. By capturing the stylized facts (Buchanan, 2012) of a system, it can serve as the simplest mathematical model displaying a particular form of universal behavior. Examples are phase oscillators as a model of synchronization phenomena (Arenas et al., 2008), the Bak-Tang-Wiesenfeld (BTW) model for avalanches (Bak et al., 1987) and the SER model (or 'forest fire model', as it has been called in the context of spatiotemporal pattern formation; Drossel and Schwabl (1992)) for excitable dynamics (Müller-Linow et al., 2008; Garcia et al., 2012; Fretter et al., 2017).

Refined minimal models

Minimal models of dynamical processes on graphs helped shape our understanding of different categories of relationships between structural and functional connectivities. This was described in deliverable D1.1 (Voutsa et al., 2021).

Building up on these results, in i-CONN such minimal models have been refined and adapted to specific application scenarios. Table 1 summarizes these model refinements.

In the original BTW model (Bak et al., 1987), where discrete numbers of quantities ('grains of sand') are distributed according to a formal update rule, a local conservation law is built into the dynamics. In this **BTW model with dissipation** we allow vegetation patches to block/absorb cascading sediment, thus creating a minimal model capable of describing, in a stylized manner, how vegetation affects the size distribution of sediment avalanches. These findings can be validated by detailed numerical simulations with candidate vegetation patterns serving as input to realistic, more detailed sediment transport simulation models (e.g., Mahleran, Wainwright et al. (2008b,a)). Results for this investigation have been obtained and a manuscript is currently written up for publication (envisioned for April 2023).

The SER model is a stochastic three-state cellular automaton which uses the recovery probability and a rate of spontaneous excitation as parameters. In the past it has been instrumental in understanding self-organized excitation waves in networks (Moretti and Hütt, 2020) and aspects of SC/FC correlations in excitable dynamics (Messé et al., 2015). In a series of secondments of ESRs 3 and 11 a refinement of the SER model have been studied by allowing neighboring excitations only to act after a delay time τ (**time-delayed SER model**). We computed the time-delayed mutual information within this model to compare with observations from Ioannides (2007). In the end, this model refinement, while helpful, did not provide deep mechanistic insight and hence no dedicated publication is planned.

Within i-CONN we study procurement data employing a range of modeling techniques. In particular we have developed a minimal model, called biased preferential attachment model (**BPAM**), creating test instances of procurement data, which will allow us quantitatively analyze the signatures of corruption.

Table 1: Refinements of minimal models in i-CONN. Original minimal models (first column). Refinements and modifications included for use within i-CONN (second column). Application domain of the refined model (third column). Early-stage researchers (ESRs) involved (fourth column; i-CONN project numbers are used). The fifth column indicates whether a dedicated publication will be devoted to this model refinement.

Original model	**Refinement**	**Application**	**ESRs**	**Publ.**
BTW model (Bak et al., 1987)	BTW with dissipation	geomorphology	1, 3	yes
SER model (Müller-Linow et al., 2008)	time-delayed SER model	neuroscience	3, 11	no
preferential attachment (Barabasi and Albert, 1999)	BPAM	procurement contract dynamics	3, 4	yes
ant trails (Camazine et al., 2001)	settlement-driven trail dynamics	settlement and route systems	2, 13	no
ERGM (Robins et al., 2007; Cranmer et al., 2020)	business kinship political influence network	procurement contract dynamics	4	yes
leaky integrate and fire neuron (Perkel et al., 1967)	model version with stochastic threshold	neuroscience	5	Lima et al. (2021)

The minimal model serving as the starting point for BPAM is the preferential attachment model of network growth from Barabasi and Albert (1999). Here we adapted this concept to the application domain of procurement by assuming that a company's probability of obtaining a contract from a public institution is proportional to the relative number of contracts this company has already received from this institution in the last T years (where T, the 'memory' of the system, is a parameter of the model). We can now ask two types of questions in this model of contract network dynamics: (1) Assuming a public institution i deviates from this scheme (e.g., by using a shorter memory $T_i < T$), how do network properties deviate from the standard case of a uniform T. This line of investigation can help develop predictors of corruption. (2) One can analyze contract networks compiled from data and their dynamics and extract the effective memory parameters, i.e., the time windows for each public institution for which a preferential attachment model best describes the data.

Results for these two lines of investigations have been obtained and a manuscript is currently written up for publication (envisioned for June 2023).

Regarding the settlement and trail data from the Bronze age investigated in the research project of ESR 13, the question has arisen how well a simple pheromone-based ant trail model using the known settlements as generalized 'food sources' (i.e., a model of **settlement-driven trail dynamics**) can serve as a data generator and how realistic the resulting trail patterns can be. This generative model can serve two purposes: (1) Data with varying levels of completeness can be created, in order to calibrate detection algorithms of settlements and completion algorithms of trails. (2) This initial minimal model can pave the way towards more detailed, parameter-rich agent-based models, in order to test, which features of the data are indicative of certain behavioral patterns and underlying rules.

First results for this application, obtained during a secondment of ESR 13, looked promising, but more urgent algorithmic challenges have been given priority in this part of the i-CONN project.

Exponential-family random graph models (ERGMs) (Robins et al., 2007) explicitly model network interdependencies and predict the probability of tie formation within an observed network based on an attribute, dyadic, and structural model terms. ERGM uses a Markov Chain Monte Carlo based algorithm to estimate parameter values. In this segment of i-CONN we use an extension for bipartite networks described in Cranmer et al. (2020). In the subproject of i-CONN devoted to political influence networks, the ERGM framework is augmented by a bipartite homophily term, leading to a **business kinship political influence network**. Results for this investigation are expected to lead to a manuscript in the first half of 2023.

In one segment of i-CONN devoted to computational neuroscience, we include an extension of the standard leaky integrate and fire model (Perkel et al., 1967) of a neuron. Usually, the model has a fixed threshold to define an action potential or 'spike'. In a work prior to his i-CONN involvement (Lima et al., 2021) ESR 5 added a stochastic threshold that follows a sigmoidal probability with parameters fitted from electrophysiological recordings of neurons *in vitro*.

Currently, many of these investigations are still ongoing, even though the refined models are in place (in fulfillment of the requirements of deliverable D1.2).

This set of investigations, where minimal models are challenged by the requirements of the data and the domain knowledge in the various applications has led to an important insight: Often the relationships between network architecture and dynamics need to be considered as a multiscale phenomenon. This insight and its further implications are briefly discussed at the end of this report.

Use of existing models for data interpretation

In addition to the dedicated projects to modify and adapt minimal models described above, several mathematical or computational models are used in ongoing i-CONN investigations to interpret empirical data. These efforts are summarized in Table 2.

Table 2: Use of existing models for data interpretation within i-CONN. The model (first column) and the references summarizing it (second column), the application domain (third column) and the ESR employing the model for their data interpretation (fourth column) are summarized here.

Model	References	**Application domain**	**ESRs**
Network Sediment Transporter	Czuba (2018); Pfeiffer et al. (2020)	geomorphology	6
Mahleran	Wainwright et al. (2008b,a)	geomorphology	1, 10
Limburg Soil Erosion Model (LISEM)	(De Roo et al., 1994, 1996)	geomorphology	10
CCHE-2D	Jia and Wang (1999); Thakur et al. (2018)	freshwater ecology	8
SAOM	Snijders et al. (2010); Steglich et al. (2010)	social-ecological systems	4, 14
ERGM	Robins et al. (2007); Cranmer et al. (2020)	social-ecological systems	4, 13, 14

In the geomorphology segment of i-CONN several existing models have been employed to interpret project data: The **Network Sediment Transporter** (Czuba, 2018; Pfeiffer et al., 2020) is a tool for simulating how sediment moves through a river network on the level of 'parcels' of sediment components with the same geometrical and physical properties. It is part of the Landlab software framework.

Mahleran (Wainwright et al., 2008b,a) (see also above) is a connectivity-based model. In i-CONN it is used to simulate water and sediment fluxes over $10 \times 30\text{m}^2$ plots located in the southwest USA under specific rain events at a spatial resolution of 0.5 m^2 and temporal resolution of 1 sec.

The **Limburg Soil Erosion Model** (LISEM) (De Roo et al., 1994, 1996) incorporates multiple modules with its own set of differential equations representing various processes to dynamically simulate soil erosion. The model calculates for both flow and sediment transfer iteratively for each unit (i.e., a pixel) on a given space (i.e., hillslope or a watershed).

In another segment of i-CONN, the simulation model CCHE-2D (Jia and Wang, 1999; Thakur et al., 2018) is considered, a two dimensional finite element model (2-D hydrodynamic surface water model.

Furthermore, two larger model families are employed to interpret i-CONN data, stochastic actor-oriented models (**SAOMs**) (Snijders et al., 2010; Steglich et al., 2010) and exponential-family random graph models (**ERGMs**) (Robins et al., 2007) (see also above). They are both in the focus of the i-CONN components studying social-ecological systems. In the case of SAOM, the model for two-mode networks (Koskinen and Edling, 2012; Snijders et al., 2013) and one-sided initiative ((Ripley et al., 2022) is used to analyze procurement networks. The ERGM framework is employed in a range of i-CONN projects to study interdependences of modalities within the various network data.

Novel model analysis techniques

Beyond the application of minimal models to specific domains, this phase of the i-CONN project has also enabled a transdisciplinary debate between modelers and domain experts within i-CONN, which has led to the development of more sophisticated strategies of analyzing minimal models:

- *Analysis of coupled map lattices via symbolic encoding*

  In Voutsa et al. (2021) we observed that for coupled logistic maps, a famous model system of chaotic dynamics, SC/FC relationships vary strongly with coupling strength. Using symbolic encoding, the mapping of the dynamics onto a cellular automaton and the subsequent analysis of the resulting attractors, we show that this behavior can be mechanistically understood.

  On a methodological level, we here introduce cellular automata as a data analysis tool, rather than a simulation model of dynamics on graphs.

  This work is near submission to a journal (planned for February 2023).

- *Prediction of dynamics from rule numbers of cellular automata*

  As a formal test scenario of predictive powers of machine learning devices and with the goal to better understand the organization of the (often very large) rule space of cellular automata (CA) of dynamics (i.e., Wolfram classes, Wolfram (1984)) from the rule number of cellular automata. We envision that this investigation allows us to better understand mechanisms of feature extraction in machine learning and, furthermore, help us contribute to the fundamental question of CA classification (Vispoel et al., 2021).

  This investigation has just started.

- *Estimating missing edges from observed collective patterns*

  Many real-life networks are incomplete. Dynamical observations can allow estimating missing edges. Such procedures, often summarized under the term 'network inference', typically evaluate the statistical correlations among pairs of nodes to determine connectivity. Within this i-CONN investigation, we have been able to offer an alternative approach: completing an incomplete network by observing its collective behavior. We illustrated this approach for the case of patterns emerging in reaction-diffusion systems on graphs, where collective behaviors can be associated with eigenvectors of the network's Laplacian matrix.

  This work has been published (Haj Ali and Hütt, 2022).

Implications

The study of minimal models within i-CONN has lead to a range of methodological advances on the level of *data generators* (e.g., vegetation patterns as input to different sediment transport simulation models or pheromone-based route system creation), *model analysis methods* (e.g., symbolic encoding of coupled chaotic maps) and *data completion algorithms* (e.g., the pattern-based estimation of missing edges when the underlying network is not completely known).

One important piece of evidence for this is that often the adjacency matrix itself is not a detailed enough representation of structural connectivity.

Definitions of structural connectivity, which go beyond the local scale, are path-based connectivities, as used in geomorphology (see, e.g., Wainwright et al., 2011) and the hub set orientation prevalence used to understand activation patterns in excitable dynamics on graphs ().

At the core of processes of self-organization is the bridging of scales. Local interactions give rise to large-scale patterns and collective behaviors. Historically these phenomena have been studied in (homogeneous) space or in systems with an all to all coupling.

Conceptually, in self-organized dynamics *on graphs* – which one theoretical core of i-CONN research activities – the main qualitative difference the network view contributes is that self-organized behavior arises via an interplay of microscale, mesoscale and macroscale features. And often causation seems not to follow the standard hierarchy of scales.

This insight, a range of examples illustrating it and a discussion of its implications is currently prepared as a review article (envisioned for May 2023).

References

Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., and Zhou, C. (2008). Synchronization in complex networks. *Physics Reports*, 469(3):93–153.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the $1/f$ noise. *Physical Review Letters*, 59(4):381.

Barabasi, A. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509.

Buchanan, M. (2012). It's a (stylized) fact! *Nature Physics*, 8(1):3–3.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraula, G., and Bonabeau, E. (2001). Self-organization in biological systems. In *Self-Organization in Biological Systems*. Princeton University Press.

Cranmer, S. J., Desmarais, B. A., and Morgan, J. W. (2020). *Inferential network analysis*. Cambridge University Press.

Czuba, J. A. (2018). A Lagrangian framework for exploring complexities of mixed-size sediment transport in gravel-bedded river networks. *Geomorphology*, 321:146–152.

De Roo, A., Wesseling, C., Cremers, N., Offermans, R., Ritsema, C., and Van Oostindie, K. (1994). LISEM: a new physically-based hydrological and soil erosion model in a GIS-environment, theory and implementation. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, 224:439–448.

De Roo, A., Wesseling, C., and Ritsema, C. (1996). LISEM: a single-event physically based hydrological and soil erosion model for drainage basins. I: theory, input and output. *Hydrological processes*, 10(8):1107–1117.

Drossel, B. and Schwabl, F. (1992). Self-organized critical forest-fire model. *Physical Review Letters*, 69(11):1629–1632.

Fretter, C., Lesne, A., Hilgetag, C. C., and Hütt, M.-T. (2017). Topological determinants of self-sustained activity in a simple model of excitable dynamics on graphs. *Scientific Reports*, 7:42340.

Garcia, G. C., Lesne, A., Hütt, M.-T., and Hilgetag, C. C. (2012). Building blocks of self-sustained activity in a simple deterministic model of excitable neural networks. *Frontiers in Computational Neuroscience*, 6:50.

Haj Ali, S. and Hütt, M.-T. (2022). Inferring missing edges in a graph from observed collective patterns. *Physical Review E*, 105(6):064610.

Ioannides, A. A. (2007). Dynamic functional connectivity. *Current opinion in neurobiology*, 17(2):161–170.

Jia, Y. and Wang, S. S. (1999). Numerical model for channel flow and morphological change studies. *Journal of Hydraulic Engineering*, 125(9):924–933.

Koskinen, J. and Edling, C. (2012). Modelling the evolution of a bipartite network—peer referral in interlocking directorates. *Social Networks*, 34(3):309–322.

Lima, V., Pena, R. F., Shimoura, R. O., Kamiji, N. L., Ceballos, C. C., Borges, F. S., Higa, G. S., De Pasquale, R., and Roque, A. C. (2021). Modeling and characterizing stochastic neurons based on in vitro voltage-dependent spike probability functions. *The European Physical Journal Special Topics*, 230(14):2963–2972.

Messé, A., Hütt, M.-T., König, P., and Hilgetag, C. C. (2015). A closer look at the apparent correlation of structural and functional connectivity in excitable neural networks. *Scientific Reports*, 5:7870.

Moretti, P. and Hütt, M.-T. (2020). Link-usage asymmetry and collective patterns emerging from rich-club organization of complex networks. *PNAS*, 117(31):18332–18340.

Müller-Linow, M., Hilgetag, C. C., and Hütt, M.-T. (2008). Organization of excitable dynamics in hierarchical biological networks. *PLoS Computational Biology*, 4(9):e1000190.

Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967). Neuronal spike trains and stochastic point processes: I. The single spike train. *Biophysical Journal*, 7(4):391–418.

Pfeiffer, A., Barnhart, K., Czuba, J., and Hutton, E. (2020). NetworkSedimentTransporter: A Landlab component for bed material transport through river networks. *Journal of Open Source Software*, 5(53).

Ripley, R. M., Snijders, T. A., Boda, Z., Vörös, A., and Preciado, P. (2022). Manual for SIENA version 4.0 (version August 11, 2022). *Oxford University Press.*

Robins, G., Pattison, P., Kalish, Y., and Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29(2):173–191.

Snijders, T. A., Lomi, A., and Torló, V. J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, 35(2):265–276.

Snijders, T. A., Van de Bunt, G. G., and Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social networks*, 32(1):44–60.

Steglich, C., Snijders, T. A., and Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological methodology*, 40(1):329–393.

Thakur, B., Parajuli, R., Kalra, A., and Ahmad, S. (2018). Exploring CCHE2D and its sediment modelling capabilities. In *World Environmental and Water Resources Congress*, pages 65–74.

Vispoel, M., Daly, A. J., and Baetens, J. M. (2021). Progress, gaps and obstacles in the classification of cellular automata. *Physica D: Nonlinear Phenomena*, page 133074.

Voutsa, V., Battaglia, D., Bracken, L. J., Brovelli, A., Costescu, J., Diaz Munoz, M., Fath, B. D., Funk, A., Guirro, M., Hein, T., Kerschner, C., Kimmich, C., Lima, V., Messé, A., Parsons, A. J., Perez, J., Pöppl, R., Prell, C., Recinos, S., Shi, Y., Tiwari, S., Turnbull, L., Wainwright, J., Waxenecker, H., and Hütt, M.-T. (2021). Two classes of functional connectivity in dynamical processes in networks. *Journal of The Royal Society Interface*, 18(183):20210486.

Wainwright, J., Parsons, A. J., Müller, E. N., Brazier, R. E., Powell, D. M., and Fenti, B. (2008a). A transport-distance approach to scaling erosion rates: 2. Sensitivity and evaluation of MAHLERAN. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, 33(6):962–984.

Wainwright, J., Parsons, A. J., Müller, E. N., Brazier, R. E., Powell, D. M., and Fenti, B. (2008b). A transport-distance approach to scaling erosion rates: 3. Evaluating scaling characteristics of Mahleran. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, 33(7):1113–1128.

Wainwright, J., Turnbull, L., Ibrahim, T. G., Lexartza-Artza, I., Thornton, S. F., and Brazier, R. E. (2011). Linking environmental regimes, space and time: Interpretations of structural and functional connectivity. *Geomorphology*, 126(3-4):387–404.

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature*, 311(5985):419–424.