MEmilio HELMHOLTZ mFUND Das Startkapital für die Mobilität 4.0

Gefördert durch: Bundesministerium für Digitales und Verkehr

aufgrund eines Beschlusses des Deutschen Bundestages

Graph Neural Network Surrogates to leverage Mechanistic Expert Knowledge towards Reliable and Immediate Pandemic Response

Agatha Schmidt[1], Henrik Zunker[1], Alexander Heinlein[2], Martin J. Kühn[1,3]

[1] Institute of Software Technology, Department of High-Performance Computing, German Aerospace Center, Cologne, Germany
[2] Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands
[3] Life and Medical Sciences Institute and Bonn Center for Mathematical Life Sciences, University of Bonn, Germany

GNN (2025)

Henrik.Zunker@DLR.de

BACKGROUND & PROBLEM

- Human ecological disruption and unsustainable consumption drive pandemic risk[1]
- Mathematical modeling can help to explore potential outcomes to guide decision makers
- Many different modeling approaches exist[2,3,4,5,6,7]
- Decision makers need reliable assessment of the situation immediately or in very short time frame
- Infectious disease spread is often heterogeneous on a spatial scale, stressing the need for correspondingly resolved models and local information

APPROACH & SOLUTION

- Use complex models as a simulator to generate high-quality and mobility-aware input data for neural network-based models
- Vary model parameters and interventions schemes to generate a large space of potential outcomes
- Precompute thousands to millions of simulator runs on a supercomputer overnight
- Train a Graph Neural Network (GNN) to a spatially resolved model that can be executed in a fraction of a second
- Use the GNN within a low-barrier webpage to allow decision makers explore potential outcomes

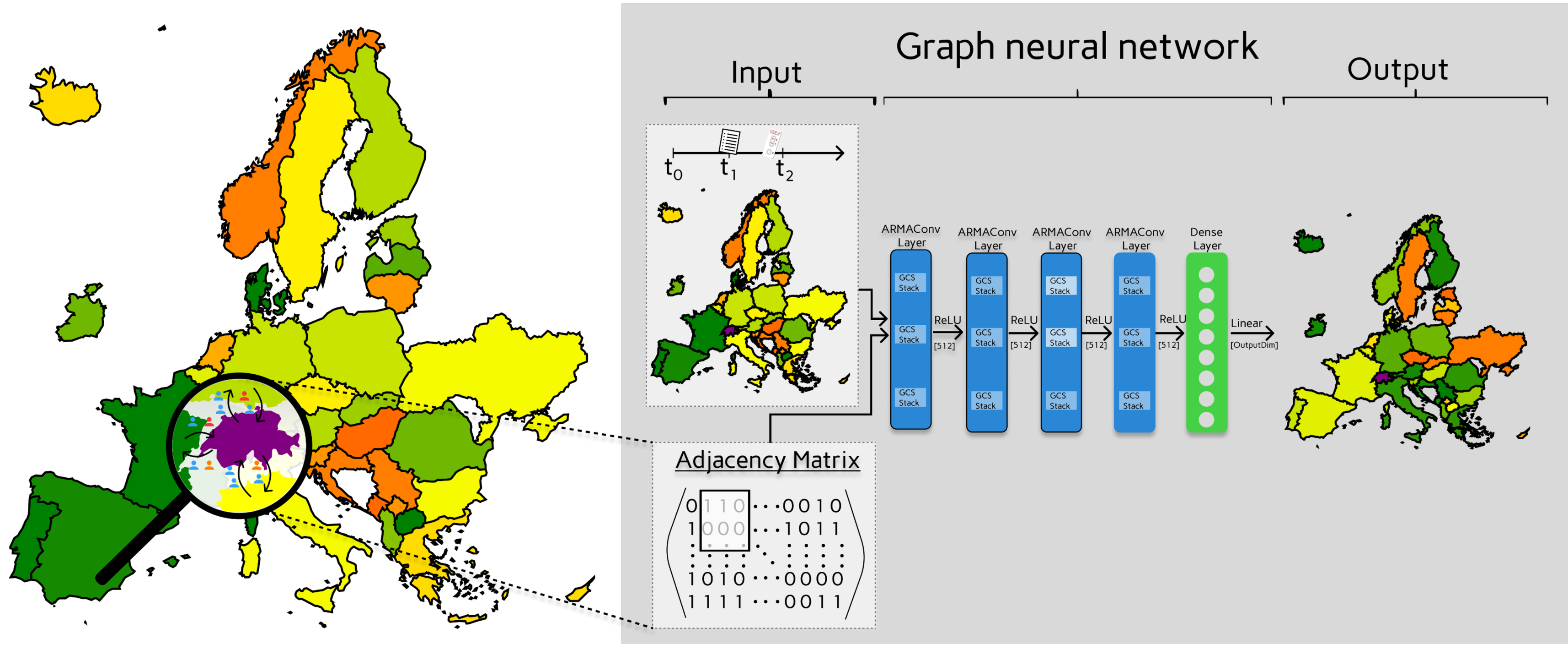

Fig 1. Schematic picture of the GNN prediction process for an exemplary resolution of European countries.

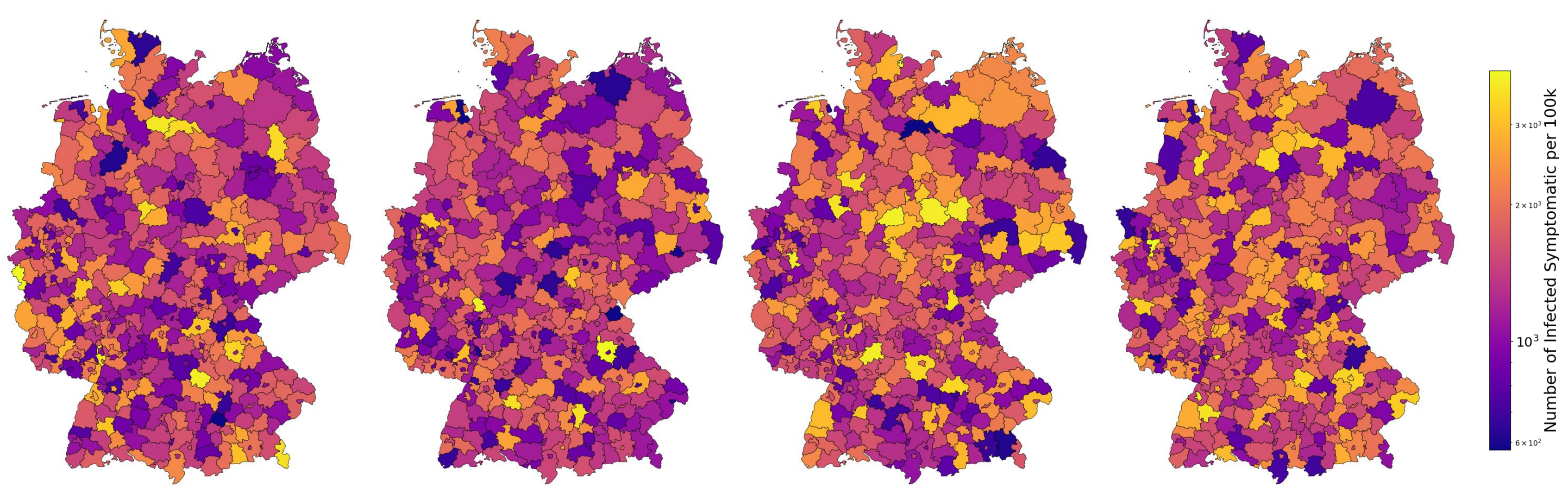

Fig 2. Randomized data input corresponding to spatially heterogeneous disease spread on the German counties.

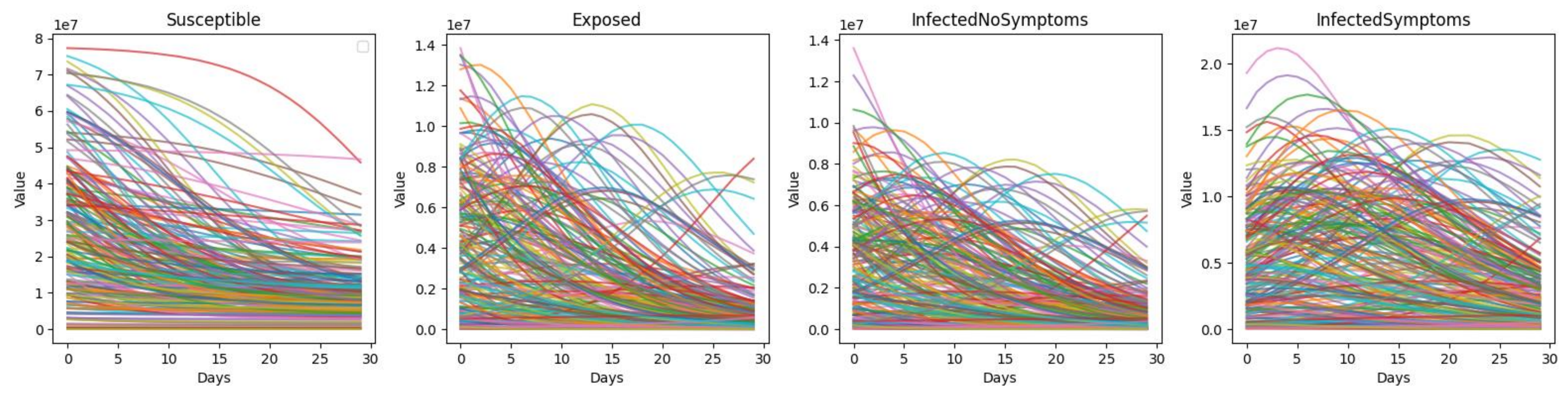

Fig 3. Randomized trajectories of infectious disease spread in one spatial entity.

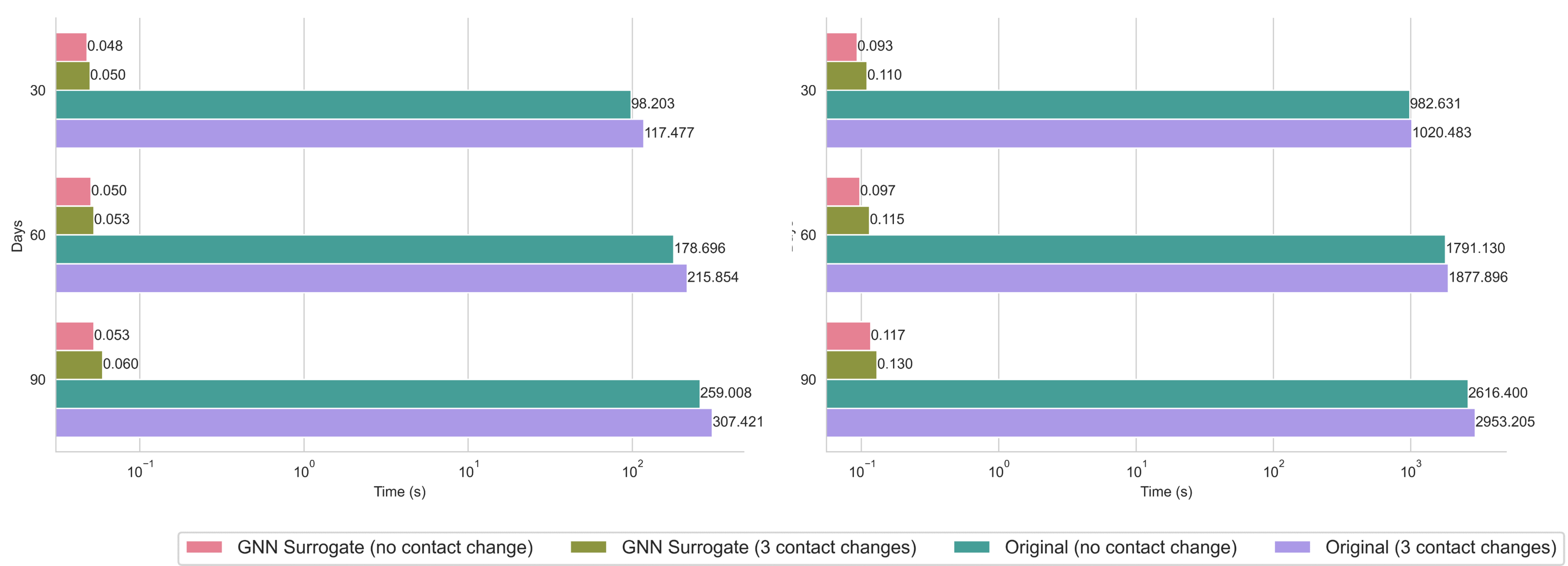

Fig 5. Execution time performance for GNN vs. complex Graph-ODE[2] model.
Comparison for 10 (left) and 100 (right) simultaneous model executions.

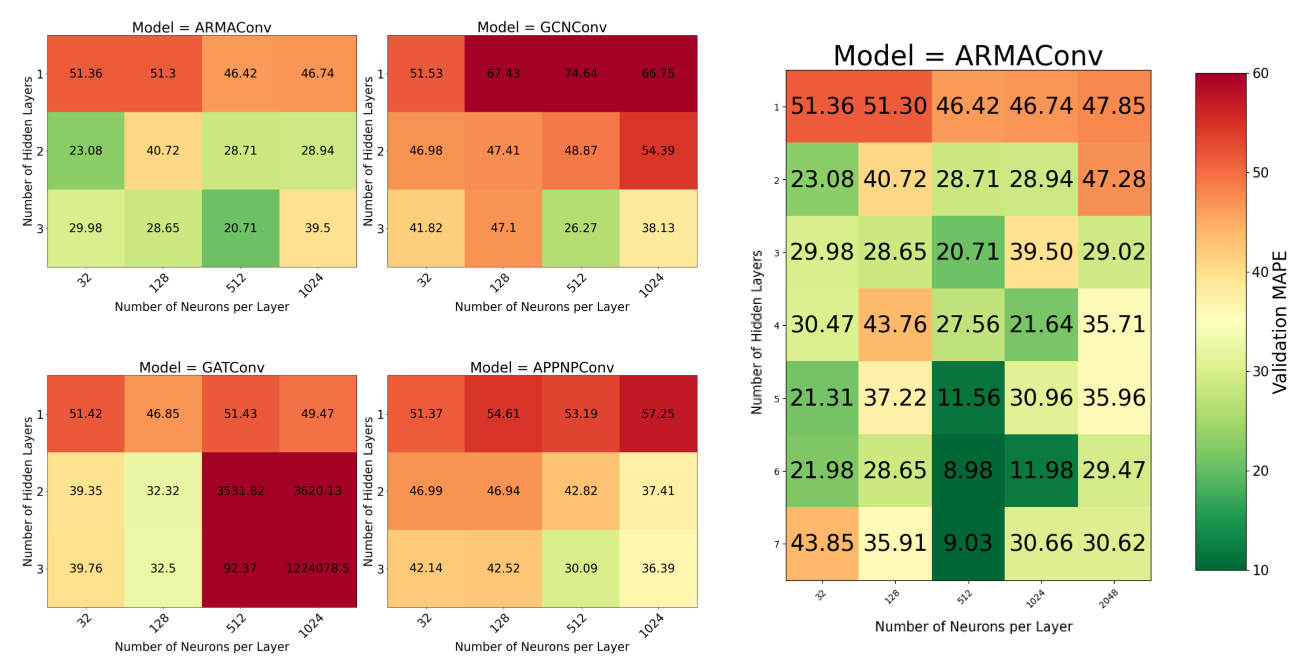

Fig 4. Grid search for GNN layers.
Different GNN layers (left) and extended grid search for best performing ARMAConv layer (right). Visualization of mean absolute percentage error (MAPE) on transformed logarithmic scale for cross-validation on input data.

RESULTS

- Final model speeds up predictions by a factor of up to 28,000
- Evaluation in 0.1 sec
- MAPE between 10.33 and 27.11 %

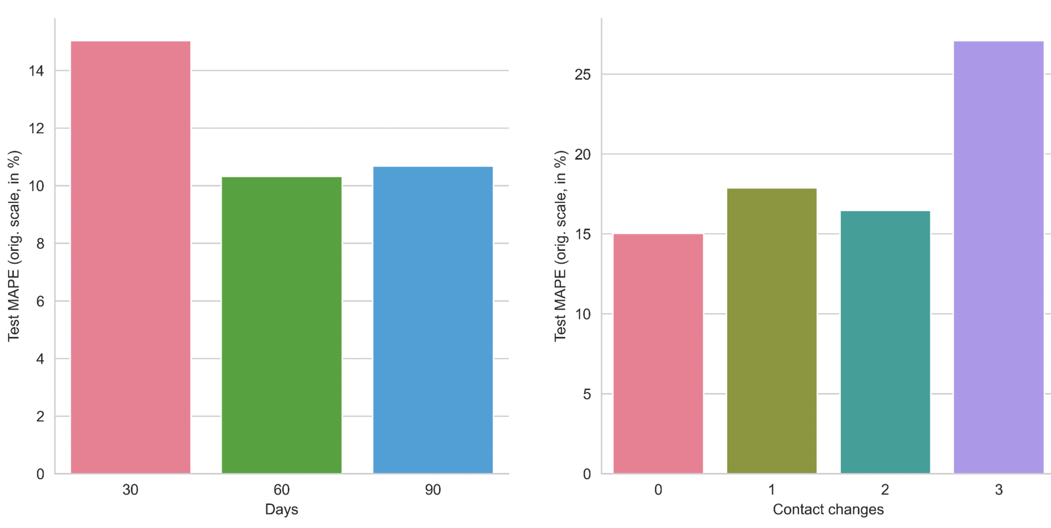

Fig 6. Test MAPE (original scale) for different prediction horizons and numbers of contact changes.
MAPE for 30, 60, and 90 days predictions (left) and 0, 1, 2, and 3 contact changes over 30 days prediction horizon (right).

REFERENCES

1. Daszak et al. Workshop *Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), 2020.*
2. Kühn et al. *Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution. Mathematical Biosciences 339 (2021): 108648.*
3. Zunker et al. *Novel travel time aware metapopulation models and multi-layer waning immunity to assess late-phase epidemic and endemic scenarios. PLOS Computational Biology 20(12): e1012630 (2024)*
4. Bicker et al. *Hybrid metapopulation agent-based epidemiological models for efficient insight on the individual scale: a contribution to green computing. Infectious Disease Modelling 10(2), pp. 571-590, (2025).*
5. Wendler et al. A nonstandard numerical scheme for a novel SECIR integro-differential equation-based model allowing nonexponentially distributed stay times. Applied Mathematics and Computation 509: 129636 (2026)
6. Kerkmann et al. Agent-based modeling for realistic reproduction of human mobility and contact behavior to evaluate test and isolation strategies in epidemic infectious disease spread. Computers in Biology and Medicine 193: 110269 (2025)
7. Plötzke et al. Revisiting the Linear Chain Trick in epidemiological models: Implications of underlying assumptions for numerical solutions. Mathematics and Computers in Simulation 239: 823–844 (2026)