# Inferring mass fluxes in diverse microbial communities

**Large-scale mechanistic microbial ecosystem models are automatically calibrated to high-resolution time-series observations using numerical optimization to infer mass fluxes**

***F. Hellweger, M. Mayerhofer, J. Hoffmann, F. Eigemann, C. Lackner****, Water Quality Engineering, TU Berlin*

## In Short

- Mass fluxes between microbes, e.g. from phytoplankton A to bacteria B, can be inferred from time series observations
- Our FluxNet method uses numerical optimization to calibrate a mechanistic microbial ecosystem model to time series data
- The method is applied to a number of different locations to learn about the ecology in these systems

Microbes are members and affect the functioning of many ecosystems, from the human gut to the global ocean, with important implications for health and climate. Components of these complex, diverse and dynamic systems, e.g. microbes and substrates, can be observed at high resolution using modern technologies. However, a critical step towards a quantitative understanding is to also characterize interactions, i.e. how mass moves through these ecological networks. For mass fluxes, observations are still limited to few samples, and bulk ecological compartments or select types. Consequently, interactions have to be inferred from observations of components, like time series data.

Past approaches to infer interactions from microbial time series data have been mostly empirical, including principal component analysis (PCA), non-metric multidimensional scaling (NMDS), empirical dynamic modeling (EDM) and various regression and correlation analyses. These empirical methods can point to possible interactions, but results can be difficult to interpret mechanistically (e.g. virus-virus interaction) and are not quantitative (e.g. do not provide carbon flux between species).

Mechanistic models describe the time evolution of components using differential mass balance equations that include specific interaction terms, like exudation of dissolved organic matter (DOM) by phytoplankton and assimilation by heterotrophic bacteria (hereafter bacteria). Parameters, like half-saturation constants, can be calibrated to observations using numerical optimization routines, but past applications have been limited to few components.

We previously developed the FluxNet method, which is based on a mechanistic model that is upscaled to hundreds of state variables and thousands of parameters. Parameters are optimized/calibrated to minimize the discrepancy between the model and observations. The optimization is challenging because of the many components, nonlinear interactions, and resulting local optima in the objective function.

A novel feature of our method is that it mimics natural speciation, where a coarse-grained model is gradually de-lumped to a finer resolution. This is illustrated in Fig. 1, which shows how the model starts with just one component in each ecological compartment (Fig. 1E) This model is optimized until a threshold is reached, and then all species are de-lumped/split into two, followed by another round of optimization and so on. During the course of the optimization, with time or model runs, the number of components and parameters increase, and the total error generally decreases, although there can be a transient increase when new species are introduced (Fig. 1A&B). This way the optimization routine works with a smaller model on average and computational effort can be directed to a smaller set of parameters corresponding to newly introduced species, and the performance increases (Fig. 1C).

In this project we apply the FluxNet method to additional locations, including English Channel (EC), Bermuda Atlantic Time Series (BATS), Müggelsee (MUG) and others. Each application constitutes an extension from the Helgoland application in different dimensions (e.g. vertical resolution) and different science objectives, as outlined below.

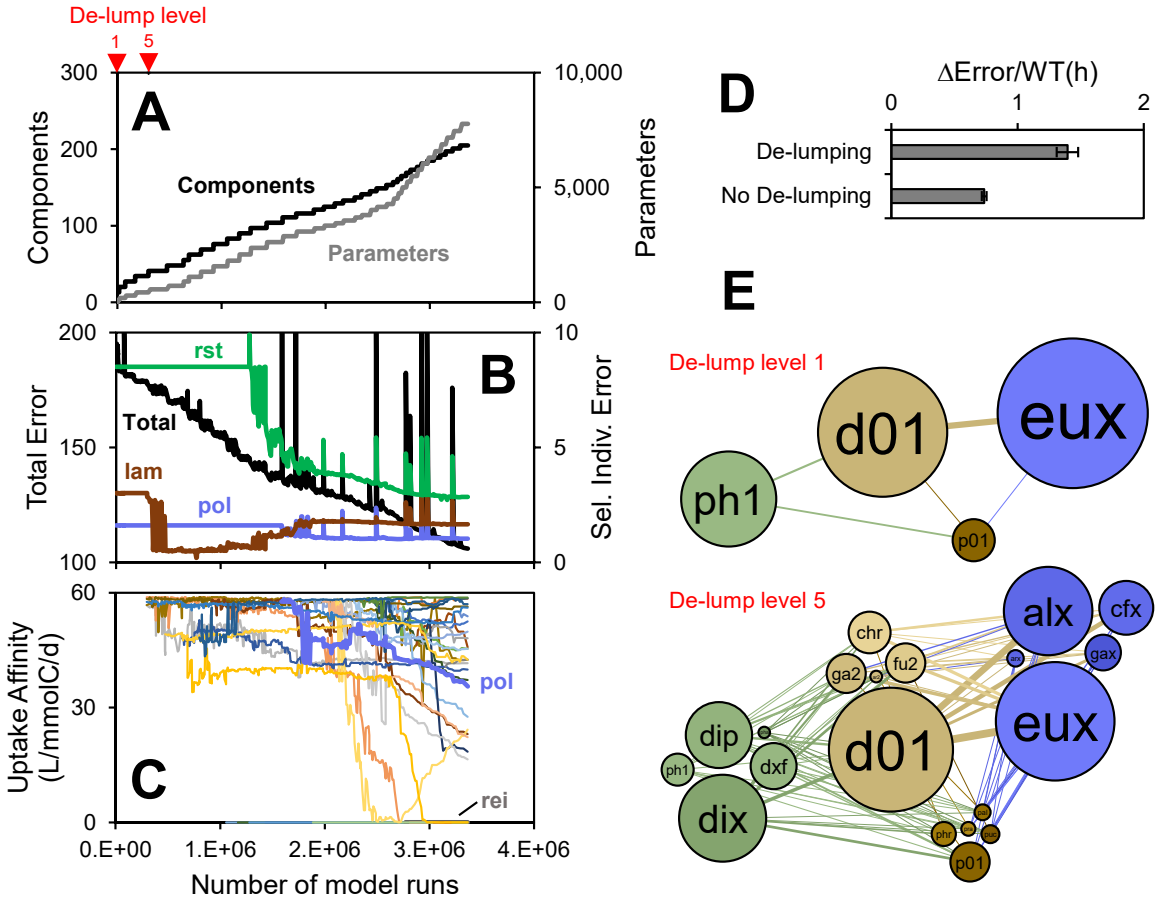


***Fig. 1. FluxNet inference method illustration.*** *(A) Number components and optimized parameters. (B) Error for entire model (Total) and selected individual observations (rst = R. styliformis, pol = Polaribacter, lam = particulate chrysolaminarin). Best of 128 replicate runs. (C) Diversification of chrysolaminarin uptake affinity (max. heterotrophy rate / half-saturation constant). (D) Method performance with and without de-lumping. (E) Network corresponding to different de-lump levels.*

## WWW

https://www.wrh.tu-berlin.de/menue/wasserreinhaltung/

## More Information

[1] M. M. Mayerhofer, F. Eigemann, C. Lackner, J. Hoffmann, F. L. Hellweger, Dynamic carbon flux network of a diverse marine microbial community, *in review*.

## Funding

DFG, Simons Foundation, NOAA and TU Berlin.