FOREMA StatisticsLab Colloquium

Understanding Change Between Observations

AQUASCOPE

Integrating high resolution monitoring and trait-based modelling to understand and predict phytoplankton dynamics

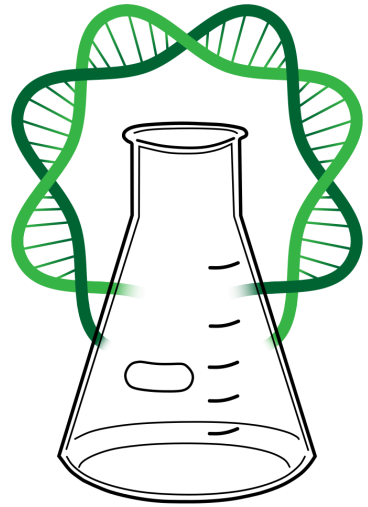
Dr. Leonardo Capitani

Department of Aquatic Ecology, Eawag

21 October 2026 · WSL Birmensdorf

Slides and links

Scan to download



open science

Take home message

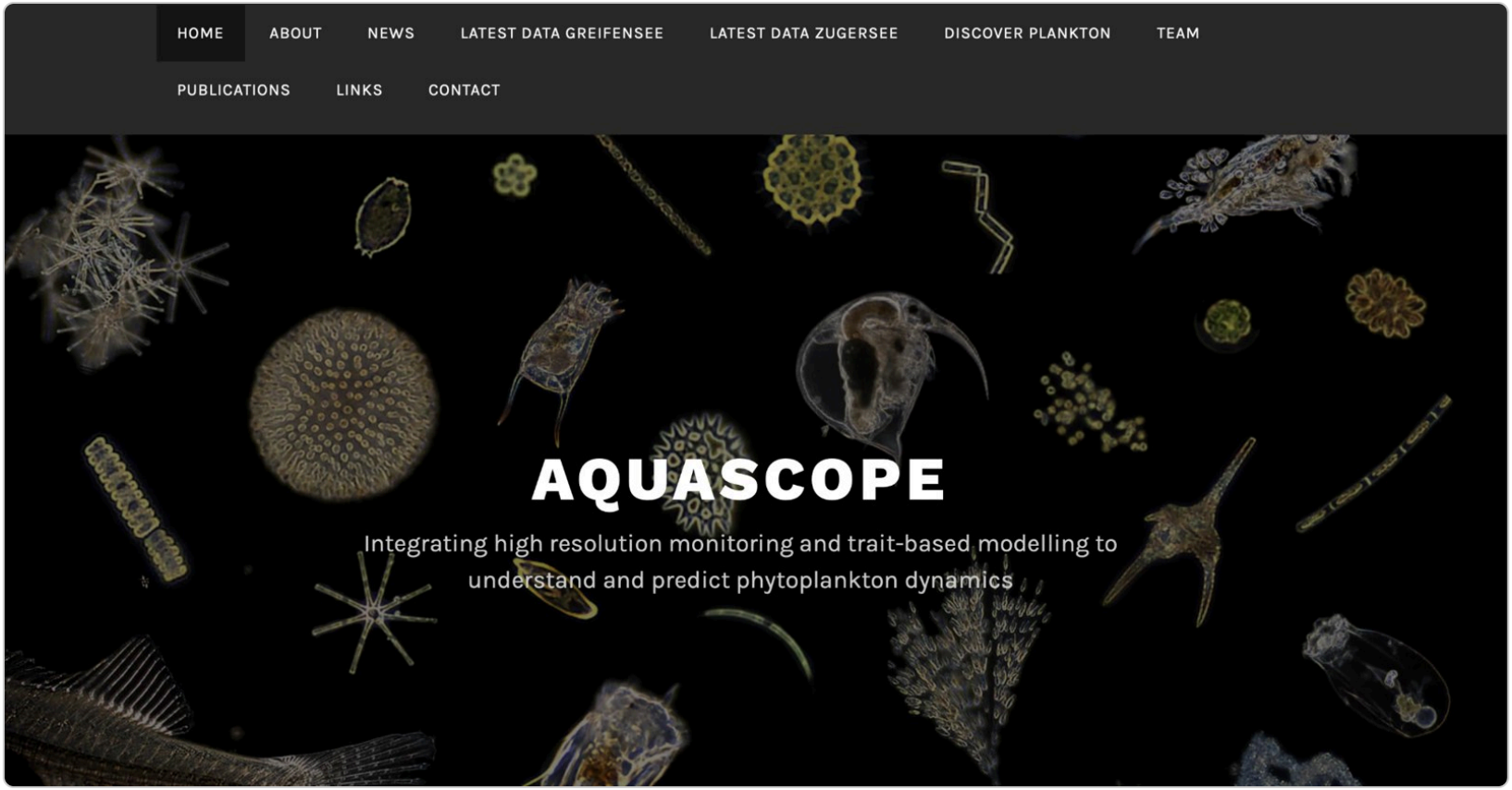


eawag
aquatic research

- PI Dr. Francesco Pomati
  (Plankton Ecology Group)
- PI Marco Baity-Jesi
  (Machine Learning & Complex Systems)

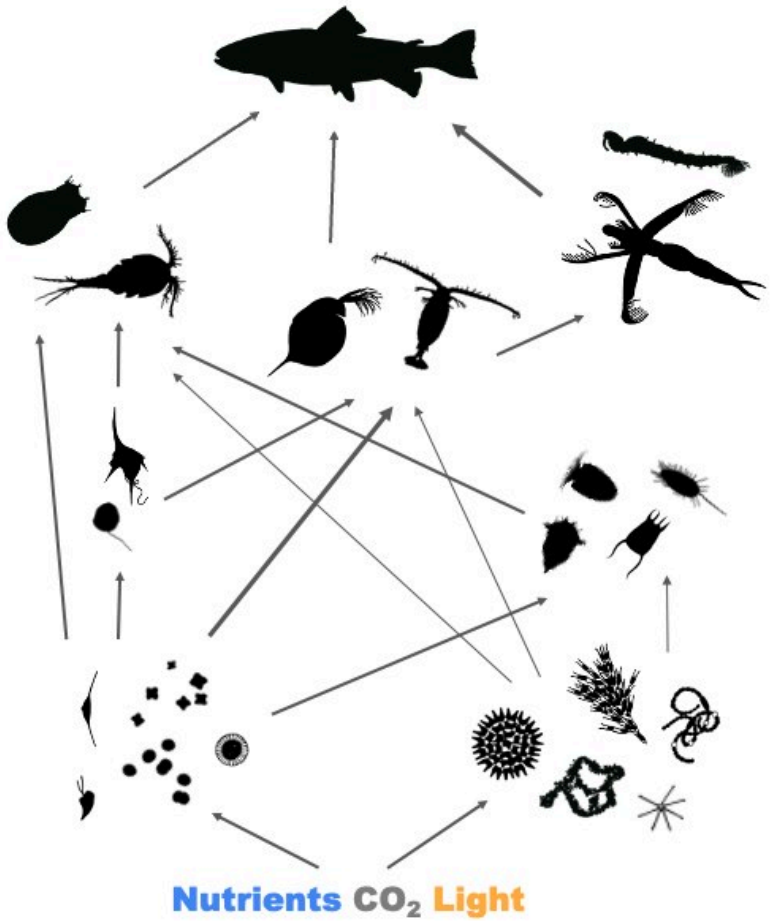
> We collaborate. Therefore we overcome research limitations.

Meet Aquascope



www.aquascope.ch

Why plankton?



Sensitive sentinels

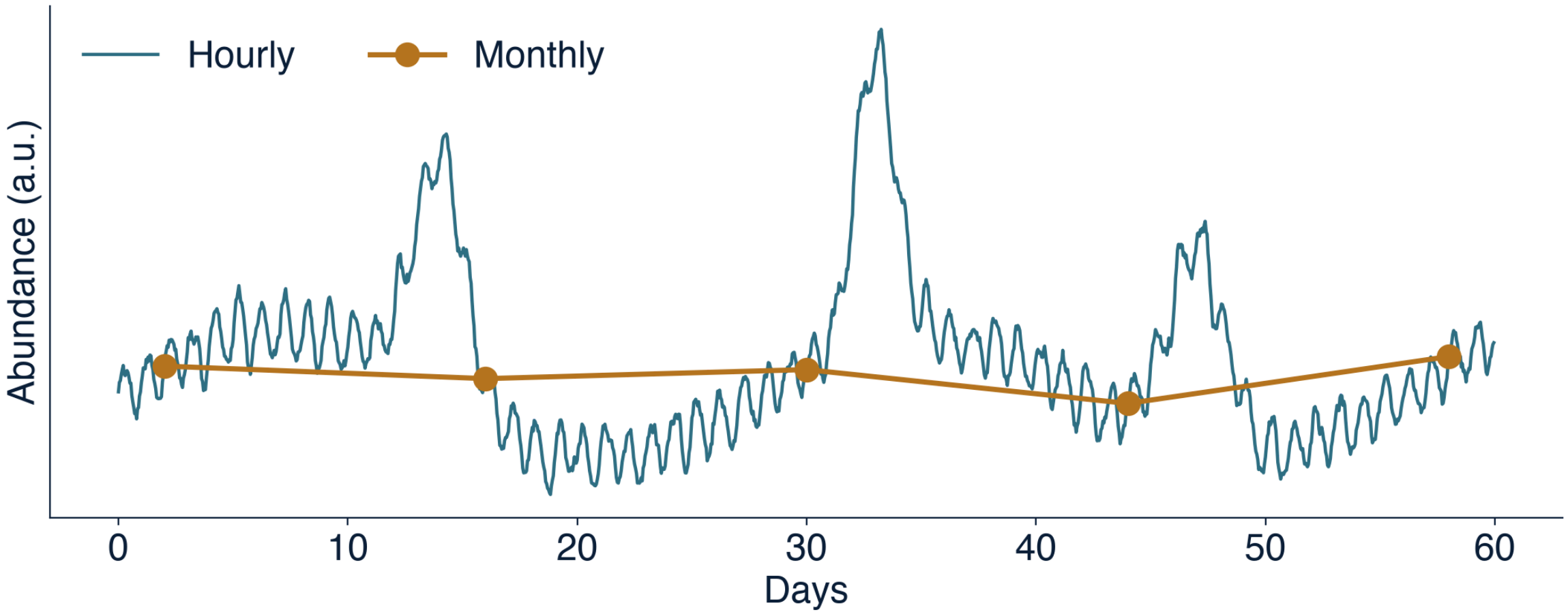
Fast turnover

Harmful blooms

Whole food web

Fast plankton turnover

Monitoring mismatch



Monthly misses blooms

Schematic, simulated data

Three-step pipeline

1

Sampling

Aquascope-Camera

→

2

Classification

Deep learning

→

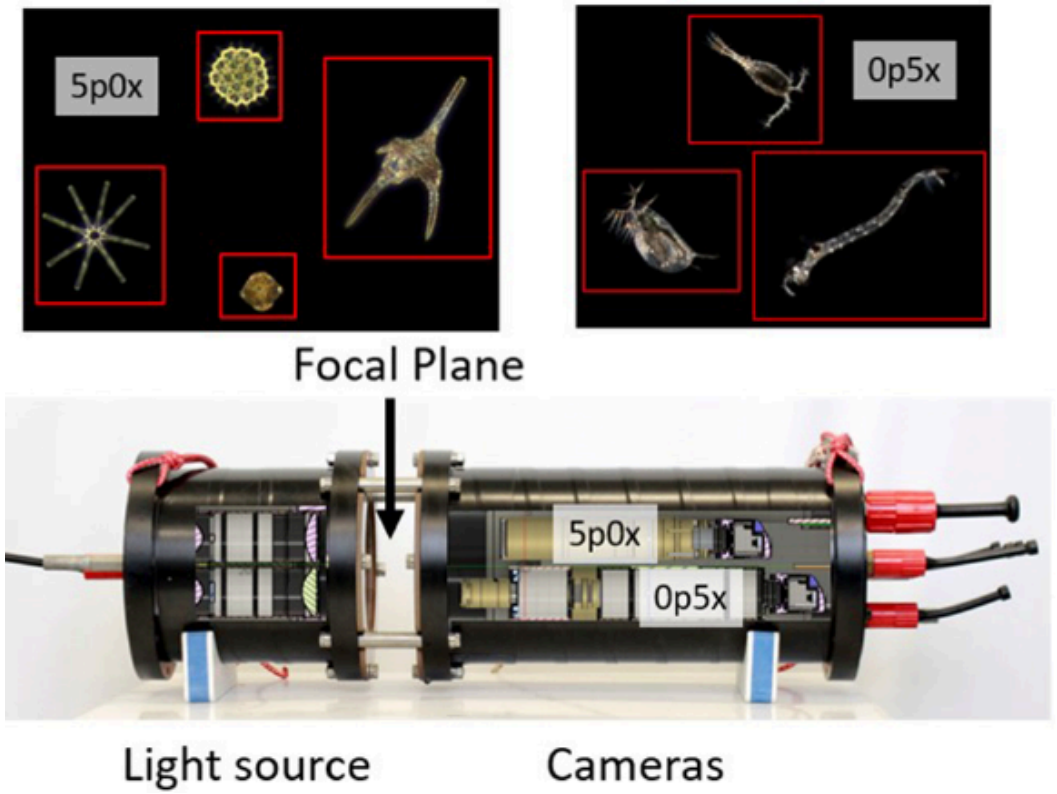
3

Modelling

Empirical Dynamic Modelling

> Every step assumes

Step 1 · Aquascope-Camera



- **Dark-field microscope**
- **Two magnifications**
- **1 μm – 1 mm**
- **Every few minutes**

ASSUMPTION

- **Imaged volume representative**
- **Undisturbed organisms**
- **Stable detection**

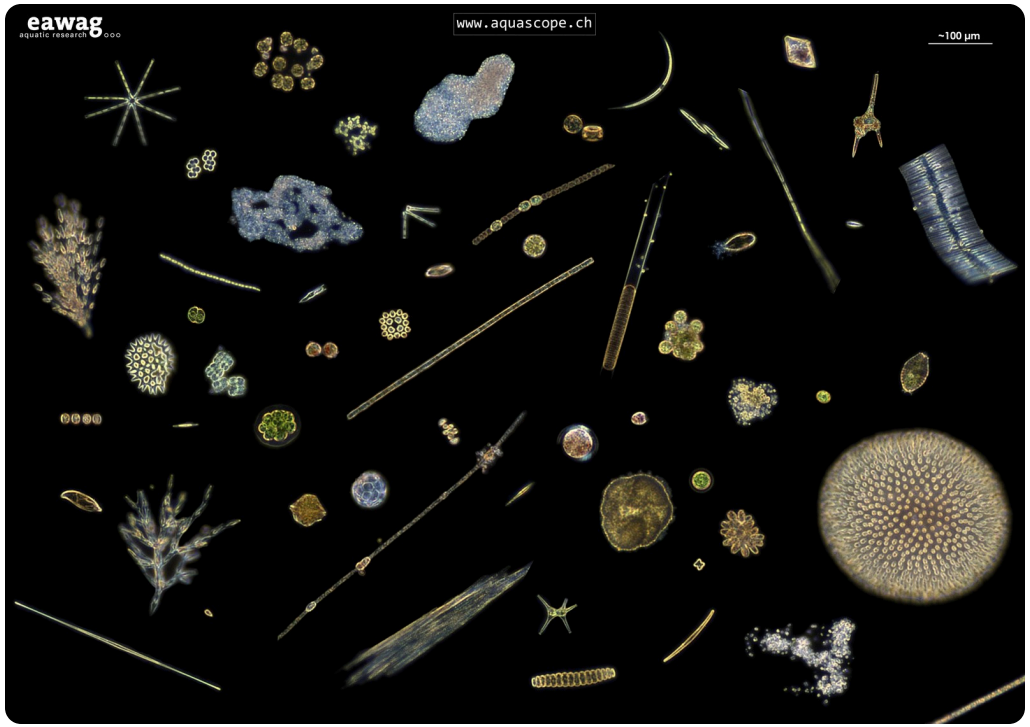
Seeing without sampling

Merz et al. 2021, Water Research 203, 117524, Fig. 1

Step 1 · Lake Greifensee station



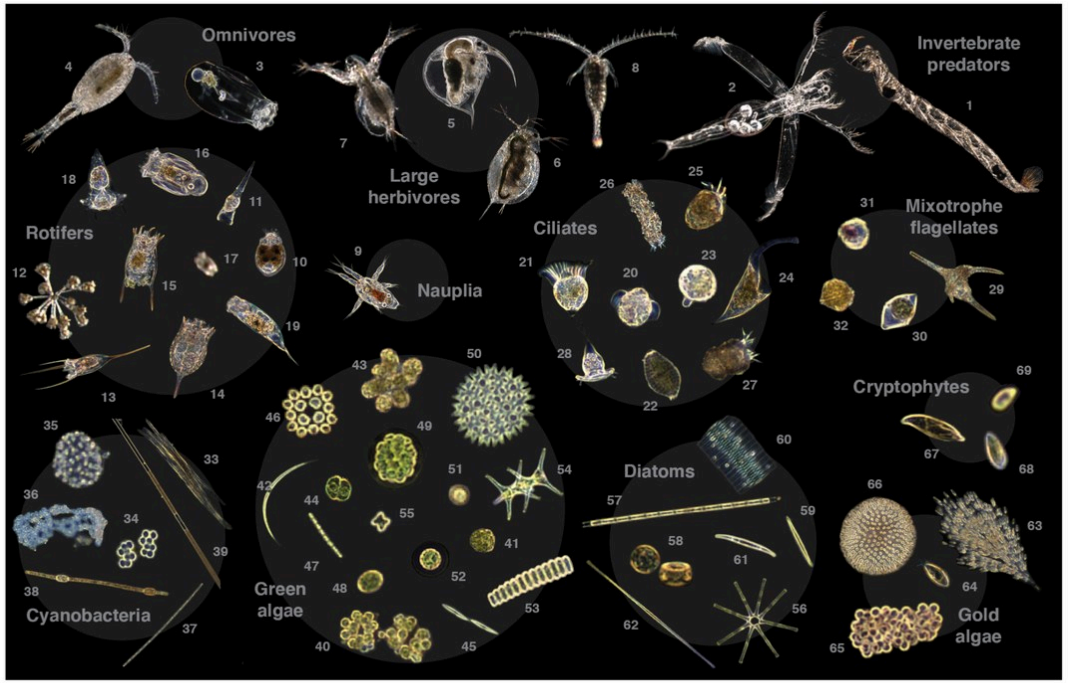
Since 2018 **3 m depth**



> Whole plankton community

Step 2 · Image classification

Image → **Deep learning** → Taxon → Abundance



ASSUMPTION

- **Deployment resembles training**
- **Expert labels correct**

> Images become counts

Step 2 · Imaging and classification performance

Camera vs microscopy

0.96
R^2 body size

0.89
R^2 phytoplankton density

0.73
R^2 zooplankton density

Classifier

35
Classes

17,900+
Annotated images

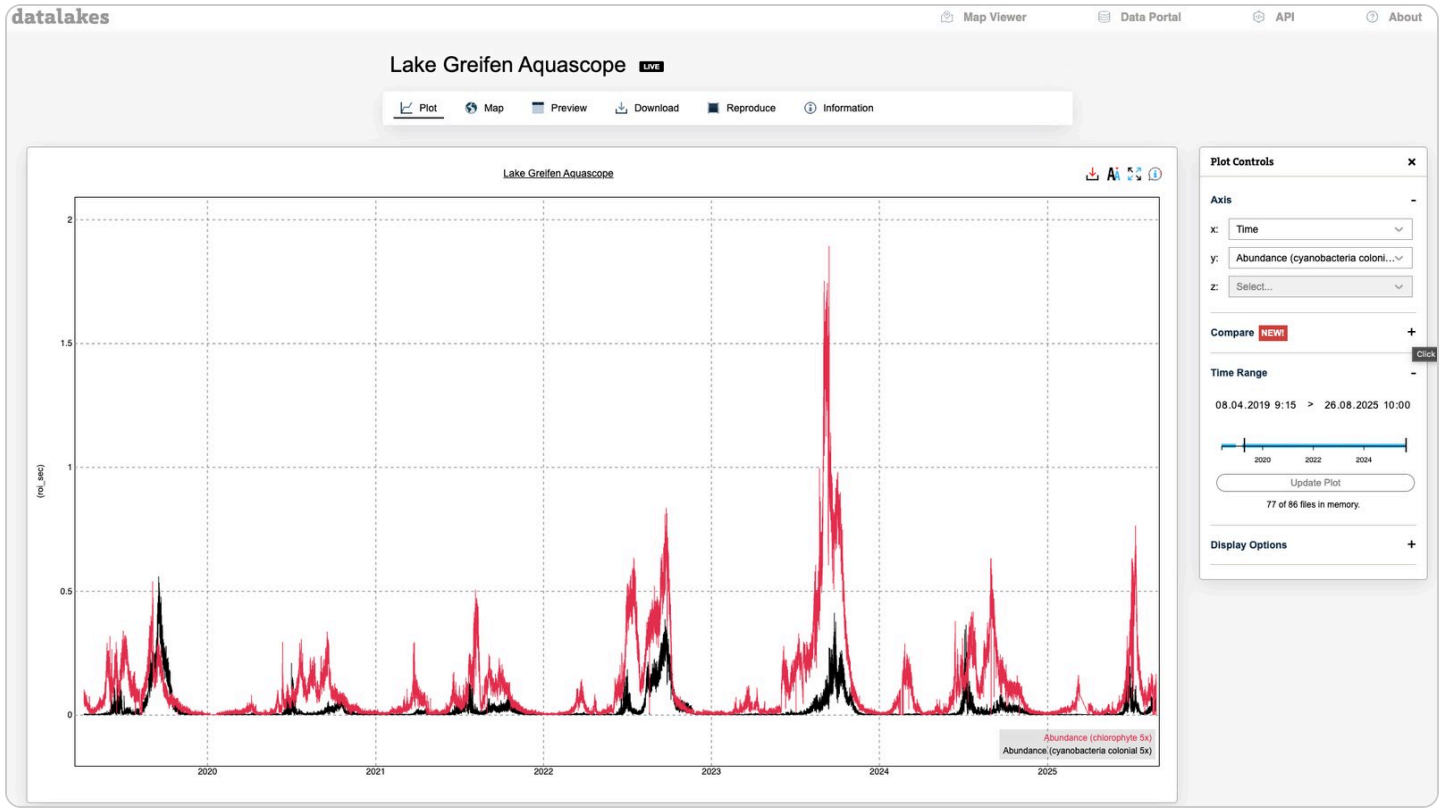
98%
Accuracy

93%
F1 score

> Validated against microscopy

Merz et al. 2021, Water Research 203, 117524 · Kyathanahally et al. 2021, Front. Microbiol. 12, 746297

Real-time dashboard



aquascope.ch

datalakes-eawag.ch/datadetail/1387

Live plankton composition

Step 3 · Workshop on Empirical Dynamic Modelling

Advanced time series analysis using EDM

25–27 August 2026 · Eawag

Guests: Steve Munch, Tanya Rogers







eawag aquatic research

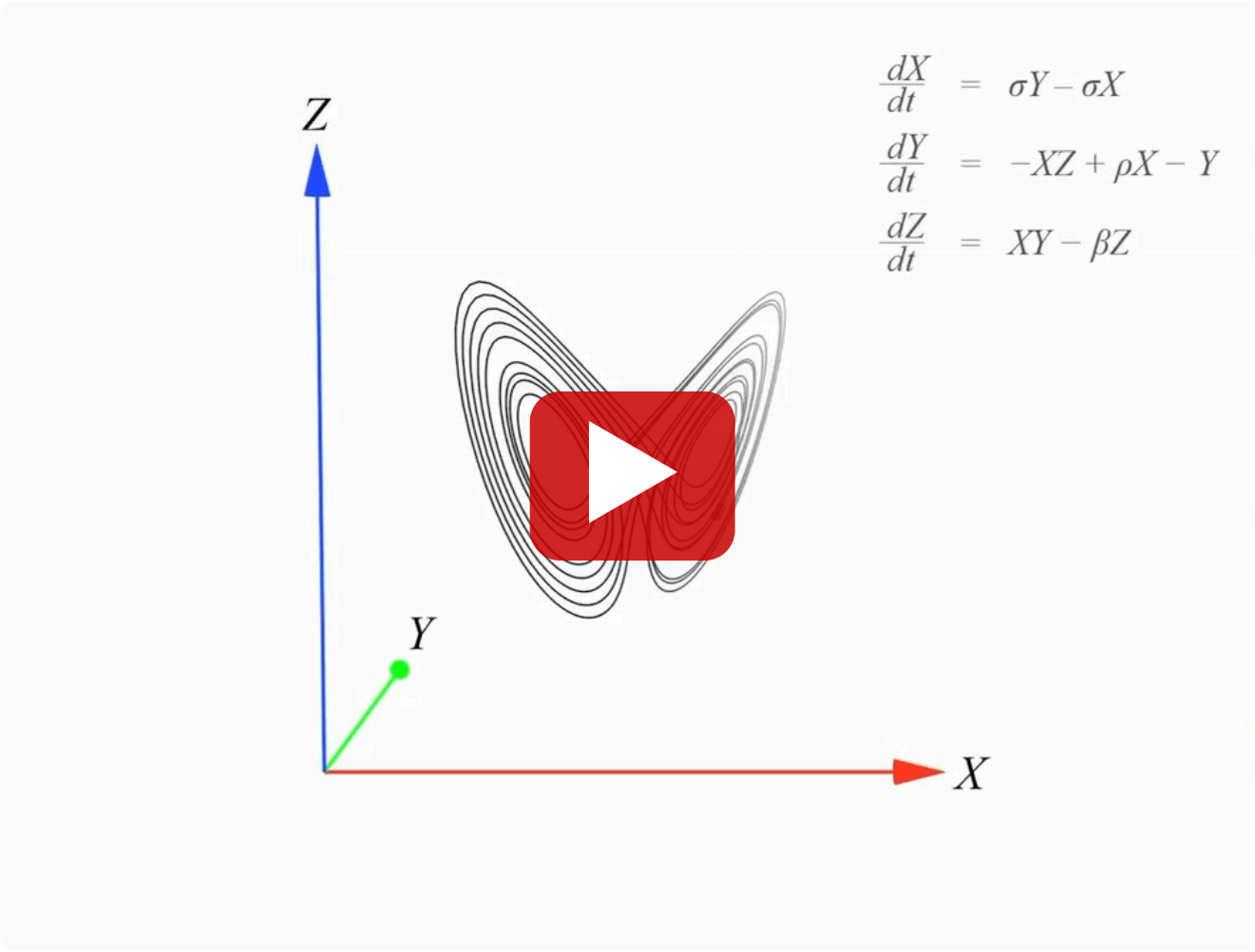
WSL

[Supported by: Biodiversity Center WSL–Eawag]

[Workshop material: leomarameo7.github.io]

The importance of knowledge exchange in situ

Step 3 · Empirical Dynamic Modelling



Video: Introduction to Empirical Dynamic Modeling, Sugihara Lab (YouTube) · Takens 1981; Sugihara & May 1990

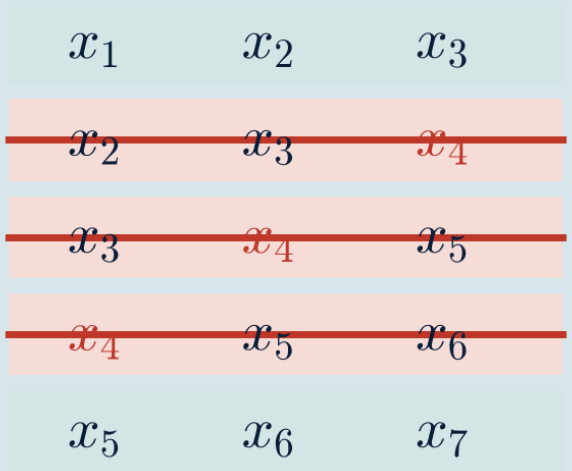
Step 3 · EDM assumptions

- Equal sampling intervals
- Stationary time series
- Enough data points (≥ 35–40)
- Moderate measurement and process error
- Low-dimensional, deterministic dynamics
- No smoothing; seasonality can mask

Chang, Ushio & Hsieh 2017, Ecological Research 32, 785–796 (Data issues; errors; time series length)

Step 3 · Handling missing data

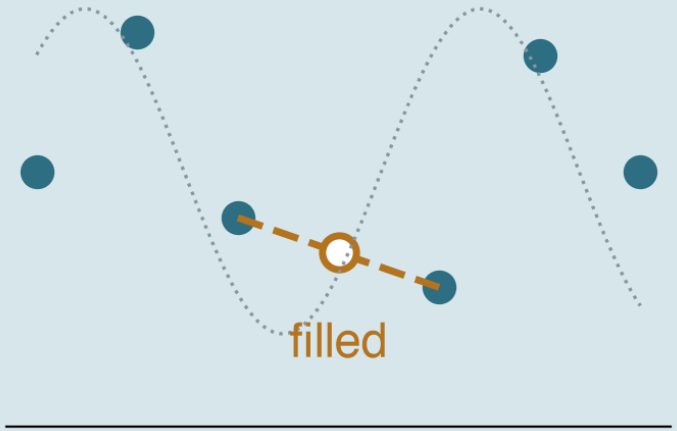
1
Exclusion



Drop incomplete vectors

Wastes data

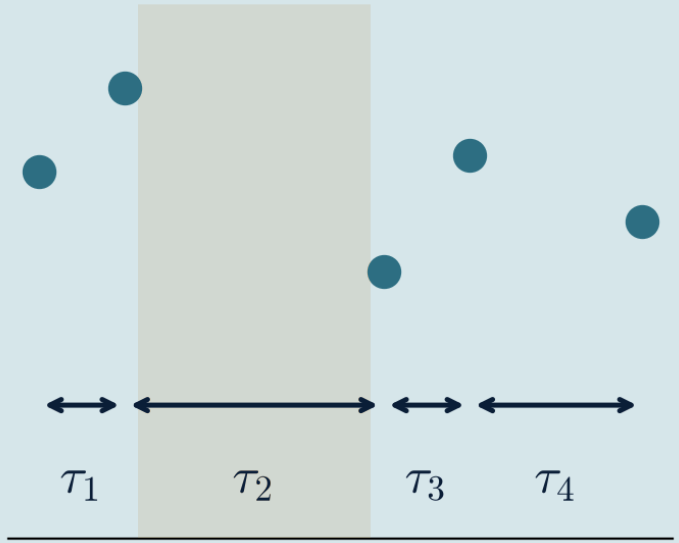
2
Interpolation



Fill, then standard EDM

Adds bias

3
Variable step-size



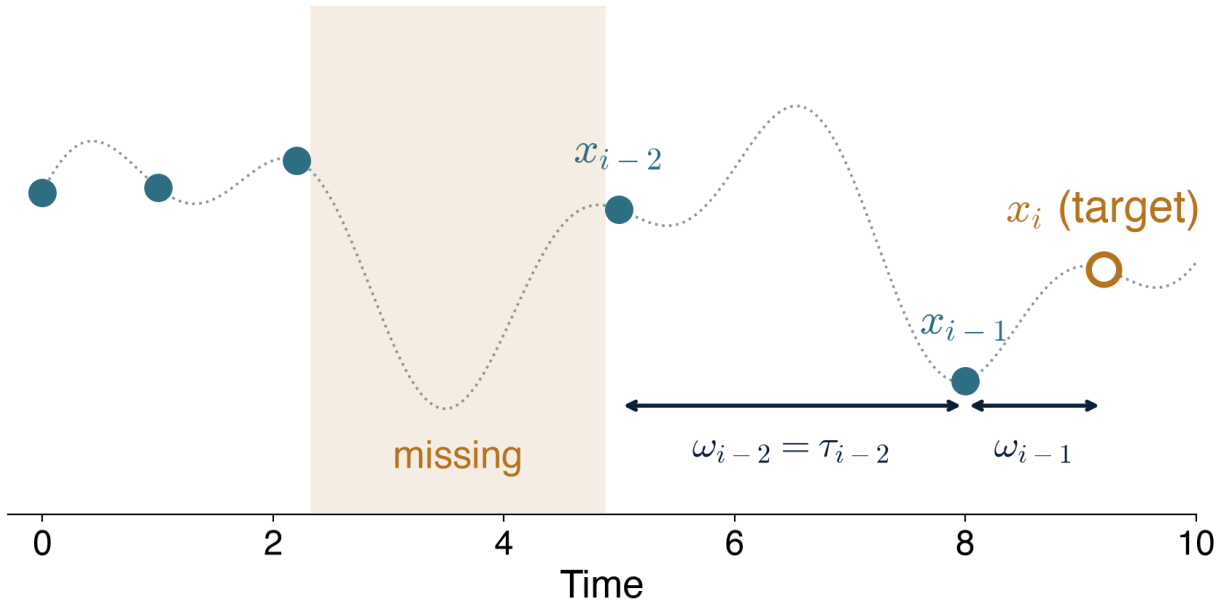
Add gaps as inputs

Doubles inputs

Three ways, three trade-offs

Johnson & Munch 2022 (complete-case: Little 1992; linear interpolation: Ye et al. 2015; bundle embedding: Stark et al. 1997; Stark 1999)

Step 3 · Variable step-size EDM



Add time gaps → **Dimension E → 2E**

Bundle-embedding theorem

Extends Takens to forced systems: add forcing lags ω

$$\{x_{i-1}, \ldots, x_{i-E},\ \omega_{i-1}, \ldots, \omega_{i-E}\}$$

Environmental drivers **Harvesting**

Stark: ω = τ (time between samples)

$$\omega_i = \tau_i \ \Rightarrow\ \{x_{i-1}, \tau_{i-1}, \ldots, x_{i-E}, \tau_{i-E}\}$$

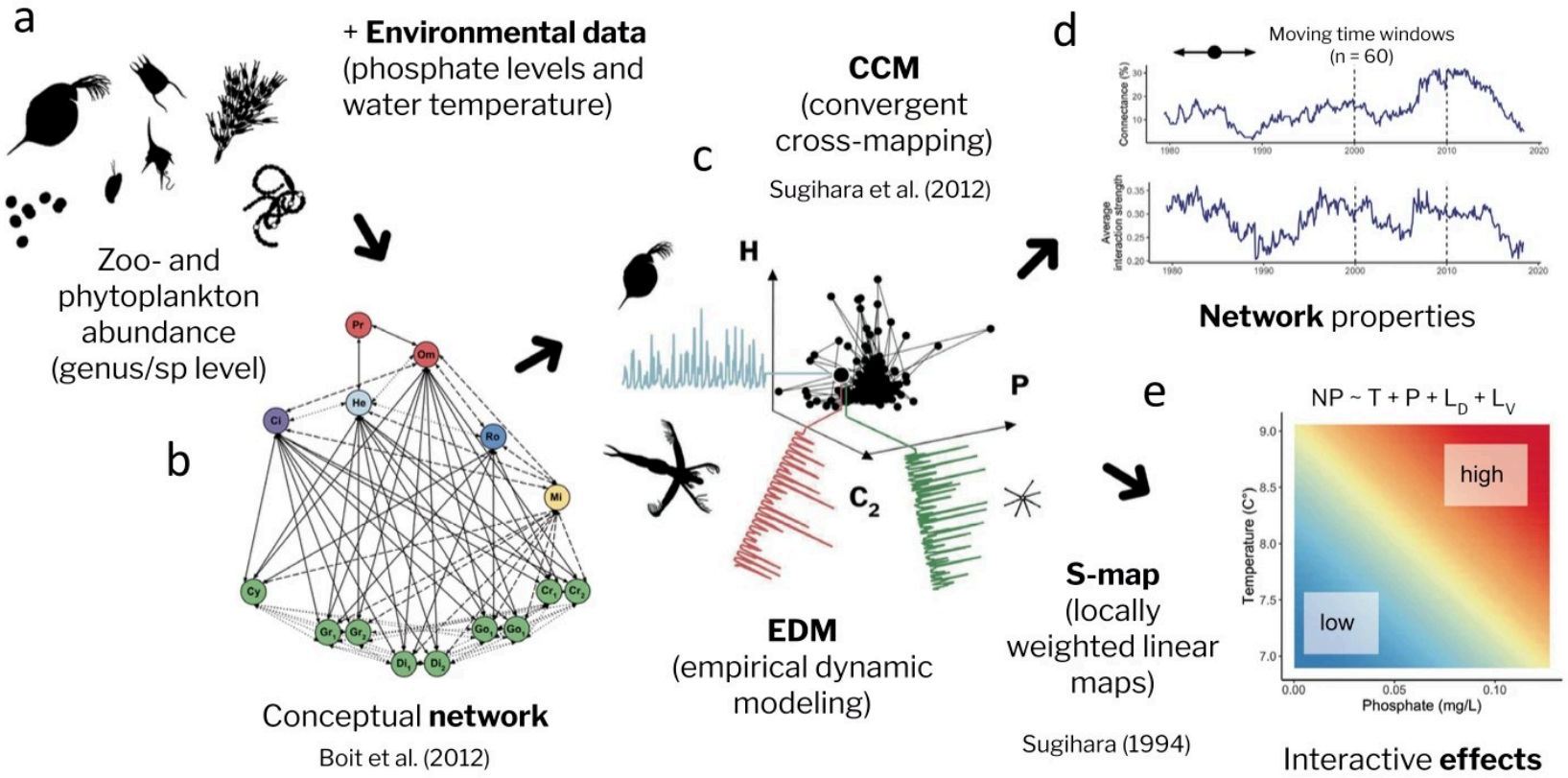
ω = τ: the next state depends on the current state and on how long we waited.

Bonus: extrapolate back to continuous-time dynamics

Gaps become inputs

Johnson & Munch 2022, Ecological Modelling 468, 109948 · Bundle-embedding theorem: Stark et al. 1997; Stark 1999 · τ as forcing: Stark 1999

Step 3 · EDM: inputs and outputs examples



- **Temperature**
- **Nutrients**
- **Density dependence**

Time-varying interactions

Pomati group schematic · Boit et al. 2012; Sugihara 1994; Sugihara et al. 2012

Take home message



> We collaborate. Therefore we overcome research limitations.

Thank you

Dr. Leonardo Capitani

Department of Aquatic Ecology, Eawag

Pomati group & colleagues: Francesco Pomati · Ewa Merz · Marta Reyes · Pinelopi Ntetsika · Steffi Eyring · Stuart Dennis · Christian Ebi · Marco Baity-Jesi · Luis Gilarranz · Rudolf Rohr · Agostino Merico · Serguei Saavedra · George Sugihara · Steve Munch · Jules Jaffe · Paul Roberts · Jean-Olivier Irisson · Raphael Bossart · Gabi Koerich · Elisabeth Janssen · David Johnson · Jon Norberg

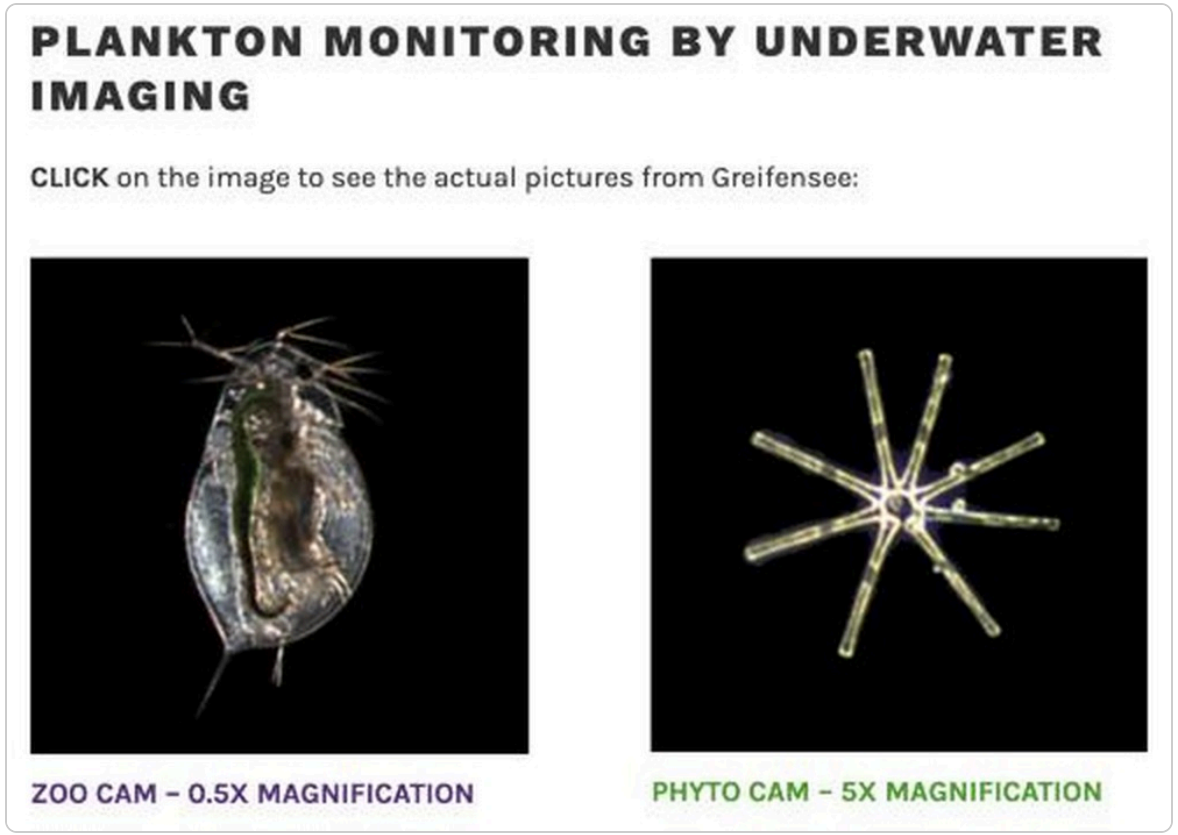
FNSNF
Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

Bundesamt für Umwelt BAFU
Office fédéral de l'environnement OFEV

leomarameo7.github.io

Supplementary materials

Who lives in Greifensee?



aquascope.ch/latest-data

Camera specifications

	5.0x	0.5x
Size range	10–150 μm	100 μm – 1 cm
Volume per frame	0.2–10 μL	4–200 μL
Target	Phytoplankton	Zooplankton
Max frame rate	10 per second	
Antifouling	UV lamp, static	

Merz et al. 2021

Camera vs microscopy

0.96
R^2 body size

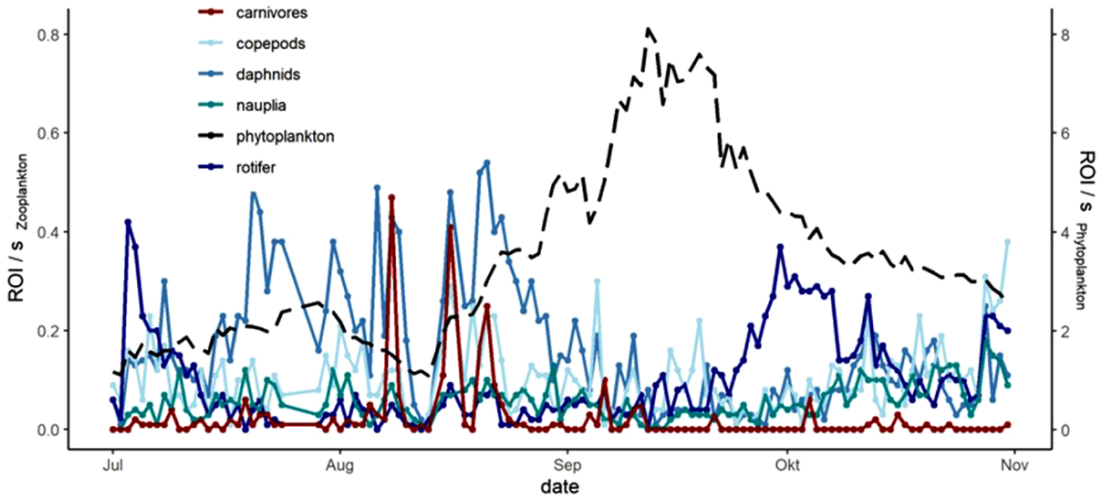
0.89
R^2 phytoplankton density

0.73
R^2 zooplankton density

Richness slightly underestimated

Merz et al. 2021 · laboratory calibration and field comparison

Why imaging, not nets?



- **Hourly resolution**
- **Undisturbed aggregates**
- **Diel patterns visible**

Merz et al. 2021, Fig. 7

Classifier details

35
Classes

17,900+
Annotated images

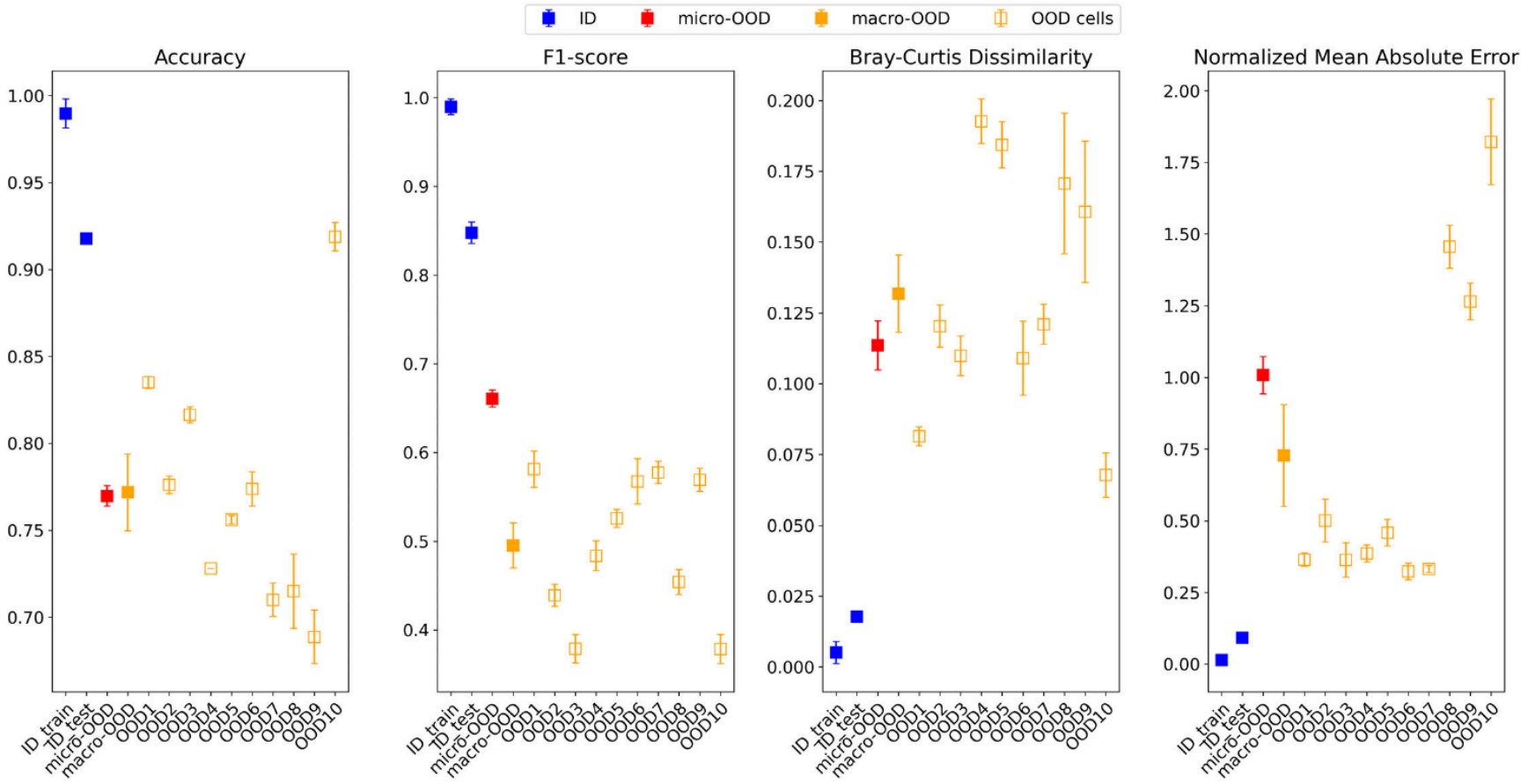
98%
Accuracy

93%
F1 score

Transfer learning + ensembling

Kyathanahally et al. 2021

Field performance drop



92%
Test accuracy

77%
New days

83%
Robust ensemble

Test scores overstate field performance; errors propagate into abundances

Chen et al. 2025, L&O Methods 23, 39–66, Fig. 5 (MobileNet; ID = test set, OOD = new deployment days). Ensemble: BEsT model

Dataset shift

Shift	Meaning
Prior probability	Taxon abundances change
Concept	Appearance changes
Causes	Turbidity, fouling, traits
Fixes	Test cells, ensembles

Chen et al. 2025

Correcting counts



> Quantification, not classification

Chen et al. 2025 (adjusted count quantification)

Gap handling options

Gap	Option
Short	Interpolate
Medium	Skip affected vectors
Long	Split series
Any	Report sensitivity

VS-EDM performance

Rotifers (R^2)	Mean
Linear interpolation	0.59
Exclusion	0.57
Piecewise constant	0.60
VS-EDM	0.68

Best: many gaps

Worse: short, smooth

Aphids: 50% missing

Gaps change method rankings

Johnson & Munch 2022, Table 2 and Fig. 3 (one-step-ahead forecasts)

EDM toolbox

Method	Use
Simplex projection	Forecasting, embedding choice
S-map	Local linear interactions
Regularised S-map	Noisy, many variables
Convergent cross mapping	Causal signal

Sugihara et al. 2012; Cenci et al. 2019; Chang et al. 2021

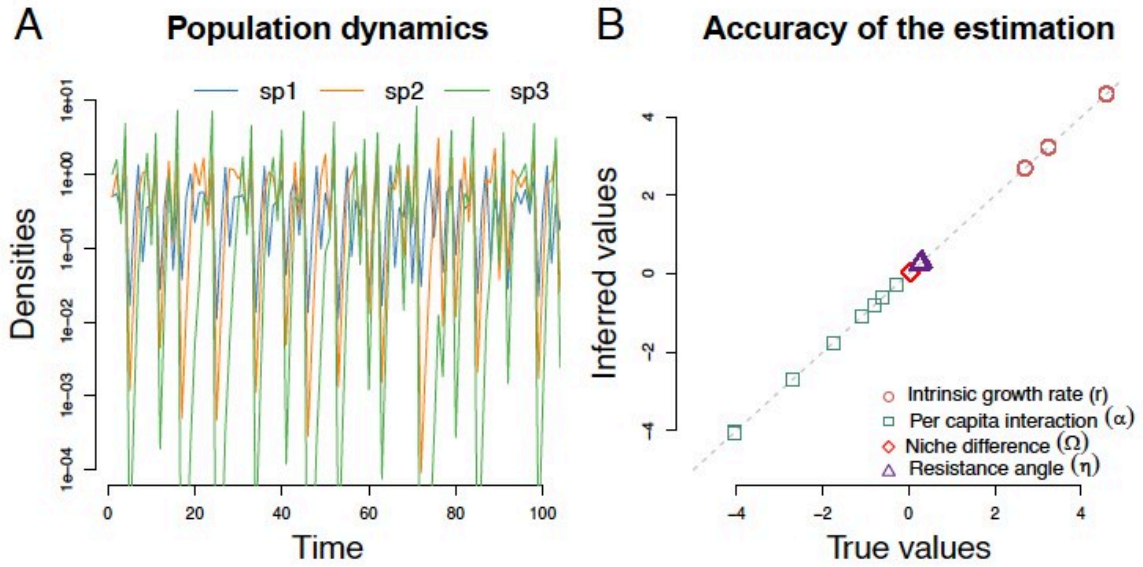
Choosing parameters

Embedding dimension E

Locality θ

Regularisation λ

Cross-validated skill



Out-of-sample decides

Simulation test: Nguyen, Pomati & Rohr 2025

Measuring forecast skill

Metric	Reads as
Correlation ρ	Pattern match
RMSE	Error size
Persistence model	Baseline to beat
Horizon	Days ahead

Merz et al. 2025

Caveats

Limit	Consequence
Short series	Wide uncertainty
Observational data	Not causal proof
Realised interactions	Context dependent
Image-based taxa	Genus-level resolution

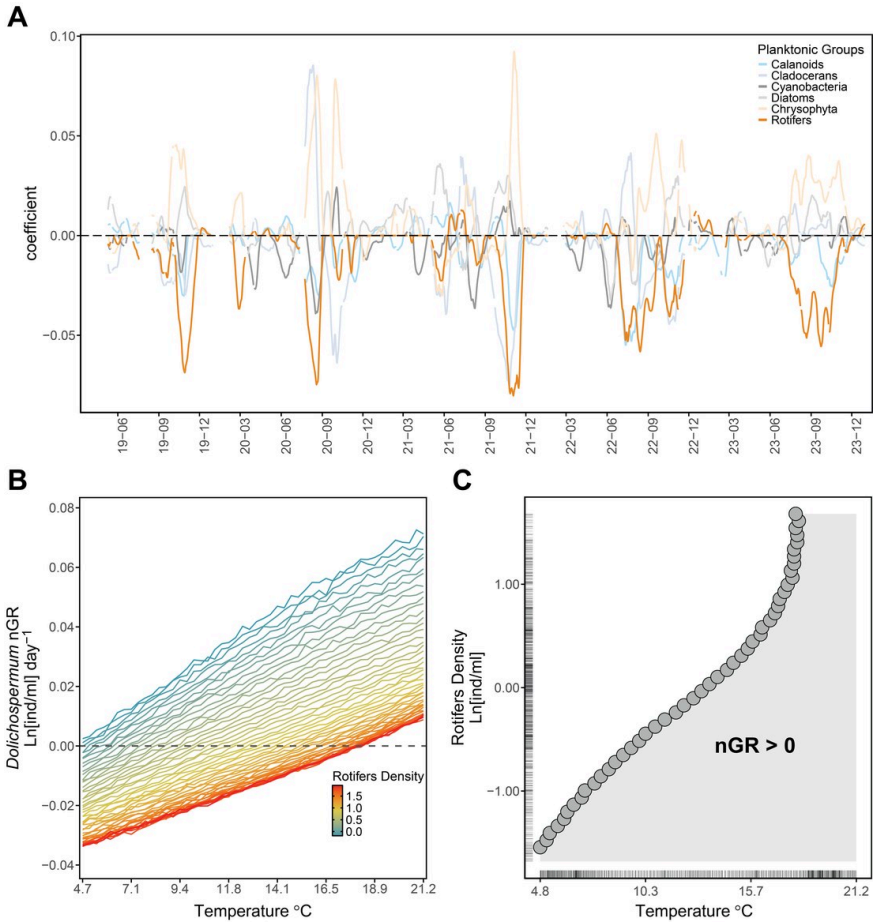
Trait-based view



- **Six size classes**
- **Temperature key**
- **Grazing: small cells**
- **Nutrients: extremes**

Eyring et al. 2025, ISME Commun. 5, ycae148 (Random Forest, six size classes)

Interactions shift niches



Ntetsika et al. 2025, bioRxiv 10.1101/2025.04.16.649070

Open data and code

Resource	Where
Time series data	opendata.eawag.ch
Live dashboards	datalakes-eawag.ch
Classifier code	github.com/mbaityje/plankifier
Annotation tool	taxonify.org
Greifensee dataset	doi 10.1038/s41597-025-04988-9

References (1/2)

- Cenci, S., Sugihara, G. & Saavedra, S. (2019). Regularized S-map for inference and forecasting with noisy ecological time series. *Methods in Ecology and Evolution*, 10, 650–660.
- Chang, C.-W., Ushio, M. & Hsieh, C. (2017). Empirical dynamic modeling for beginners. *Ecological Research*, 32, 785–796.
- Chang, C.-W. *et al.* (2021). Reconstructing large interaction networks from empirical time series data. *Ecology Letters*, 24, 2763–2774.
- Chen, C., Kyathanahally, S.P., Reyes, M., Merkli, S., Merz, E., Francazi, E., Hoege, M., Pomati, F. & Baity-Jesi, M. (2025). Producing plankton classifiers that are robust to dataset shift. *Limnology & Oceanography: Methods*, 23, 39–66.
- Eyring, S. *et al.* (2025). Five years of high-frequency data of phytoplankton, zooplankton and limnology from a temperate eutrophic lake. *Scientific Data*, 12, 653.
- Eyring, S., Merz, E., Reyes, M., Ntetsika, P., Dennis, S.R., Isles, P.D.F., *et al.* (2025). Distinct phytoplankton size classes respond differently to biotic and abiotic factors. *ISME Communications*, 5, ycae148.
- Johnson, B. & Munch, S.B. (2022). An empirical dynamic modeling framework for missing or irregular samples. *Ecological Modelling*, 468, 109948.
- Kyathanahally, S.P., Hardeman, T., Merz, E., Bulas, T., Reyes, M., Isles, P., Pomati, F. & Baity-Jesi, M. (2021). Deep learning classification of lake zooplankton. *Frontiers in Microbiology*, 12, 746297.
- Merz, E., Kozakiewicz, T., Reyes, M., Ebi, C., Isles, P., Baity-Jesi, M., *et al.* (2021). Underwater dual-magnification imaging for automated lake plankton monitoring. *Water Research*, 203, 117524.

References (2/2)

- Merz, E., Pomati, F., Saavedra, S. & Gilarranz, L.J. (2025). Biodiversity forecasting in natural plankton communities reveals temperature and biotic interactions as key predictors. *Proceedings of the Royal Society B*, 292.
- Nguyen, P.L., Pomati, F. & Rohr, R.P. (2025). Inferring coexistence likelihood in changing environments from ecological time series. *PNAS*, 122, e2417905122.
- Ntetsika, P. *et al.* (2025). Biotic interactions shape the realised niche of toxic cyanobacteria. *bioRxiv*, 2025.04.16.649070.
- Orenstein, E.C. *et al.* (2020). The Scripps Plankton Camera system: a framework and platform for in situ microscopy. *Limnology & Oceanography: Methods*, 18, 681–695.
- Stark, J. (1999). Delay embeddings for forced systems. I. Deterministic forcing. *Journal of Nonlinear Science*, 9, 255–332.
- Stark, J., Broomhead, D.S., Davies, M.E. & Huke, J. (1997). Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis: Theory, Methods & Applications*, 30, 5303–5314.
- Deyle, E.R., Maher, M.C., Hernandez, R.D., Basu, S. & Sugihara, G. (2016). Global environmental drivers of influenza. *PNAS*, 113, 13081–13086.
- Rogers, T.L. & Munch, S.B. (2020). Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. *PNAS*, 117, 479–485.
- Brias, A. & Munch, S.B. (2021). Ecosystem based multi-species management using empirical dynamic programming. *Ecological Modelling*, 441, 109423.
- Sugihara, G. (1994). Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions of the Royal Society A*, 348, 477–495.
- Sugihara, G. & May, R.M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344, 734–741.
- Sugihara, G., May, R., Ye, H., Hsieh, C.-H., Deyle, E., Fogarty, M. & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338, 496–500.