FOREMA StatisticsLab Colloquium

Understanding Change Between Observations

How do we reconstruct processes when observations are incomplete, irregular or separated across time or space? This colloquium explores methods for estimating what happens between measurements and understanding how different approaches to “filling the gaps” shape the conclusions we ultimately draw.

October 21st 2026, 14.00 – 16.00 with following Apéro, in the Englersaal, at WSL Birmensdorf.

Program

- Introduction: Understanding change between NFI measurements, **Gina Marano** (WSL)
- Data fusion for space-time data: chances and pitfalls, **Reinhard Furrer** (UZH)
- A modular, multi-tier deep learning pipeline for gap-filling and virtual reconstruction of coupled forest micro-environmental time series, **Mirko Lukovic** (WSL)
- AQUASCOPE: Integrating high resolution monitoring and trait-based modelling to understand and predict phytoplankton dynamics, **Leonardo Capitani** (Eawag)
- Apéro

Online participation via Microsoft Teams
Meeting ID: 325 545 047 970 391; Passcode: an73bW76

Register

Scan the QR code or use the link to the Registration form

Contact: Roman Flury & Jonas Stillhard
(roman.flury@wsl.ch, jonas.stillhard@wsl.ch)

WSL

Abstracts

Introduction: Understanding change between NFI measurements

Gina Marano, Forest Resources and Management, Swiss Federal Institute for Forest, Snow and Landscape Research, WSL, Switzerland.

The Swiss National Forest Inventory (NFI) measures forests at specific points in time and space. Since we do not directly observe changes between these measurements, we must use methods to estimate how and why the forests may have changed during this period. Illustrative examples why estimating these changes is important and what we can learn from them.

Data fusion for space-time data: chances and pitfalls

Reinhard Furrer, Department of Mathematical Modeling and Machine Learning, University of Zurich, Switzerland.

Data describing processes that vary over space and time increasingly come from multiple sources, often with different resolutions, coverage, and quality. Combining these sources—broadly referred to as data fusion—offers the opportunity to extract information that would not be available from any single source alone. This makes data fusion an attractive solution, and perhaps explains the
ubiquitous "Have you tried data fusion?" suggestion whenever a method does not achieve the desired quality. Yet combining data sources does not automatically lead to better results and may introduce its own difficulties.

In this talk, I will use several examples from space-time modeling to illustrate different strategies for making use of multiple data sources. Rather than focusing on mathematical or statistical details, I will discuss when such approaches can be useful, when they may not be, and some of the pitfalls that arise in practice. The talk will not provide ultimate answers, but will hopefully stimulate discussion about when—and how—different sources of information should be combined.

A modular, multi-tier deep learning pipeline for gap-filling and virtual reconstruction of coupled forest micro-environmental time series

Mirko Lukovic, Forest and Soil Ecology, Swiss Federal Institute for Forest, Snow and Landscape Research, WSL, Switzerland.

TreeNet is a growing network of forest monitoring sites designed primarily to measure tree stem radius changes using high-resolution point dendrometers. These measurements, taken every 10 minutes with micrometer precision, are complemented by other critical environmental data streams, such as temperature and relative humidity. All data are transmitted via LoRaWAN without local logging, making them vulnerable to gaps caused by transmission disturbances. As the network expands, ensuring data integrity through automated gap-filling and cleaning becomes essential—not only for dendrometer signals but for all channels. We present an approach based on machine learning to address this challenge by leveraging cross-correlation among locally available data streams and meteorological products.
We present a modular pipeline for gap-filling the data, organised around two ideas. The variables are processed in a fixed physiological order, and each stage's gap-filled product becomes an input feature for all downstream stages, so that inter-variable coupling in the soil-plant-atmosphere continuum is exploited explicitly rather than ignored. Five signal reconstruction tiers that match model complexity to gap severity, from linear interpolation of sub-hourly gaps, through a bidirectional temporal convolutional network with self-attention for gaps of up to 35 days, to the generation of complete synthetic series at sites where an instrument was never installed.

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

Leonardo Capitani, Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Switzerland.

Plankton communities turn over on scales of hours to days yet conventional lake monitoring samples once or twice a month — a mismatch that makes many ecological dynamics statistically invisible. In the Pomati group at Eawag we address this with the Aquascope-Camera a dual-magnification dark-field underwater microscope that images organisms from roughly 1 µm to 1 mm every few minutes continuously and in situ without disturbing the community it observes a purpose-trained deep learning model then classifies the images into taxonomic groups and abundances. This design solves one problem and creates others: classification uncertainty propagating into abundance estimates irregular gaps from maintenance biofouling and instrument downtime and strong non-stationarity across seasons. To analyse these series we use Empirical Dynamic Modelling a nonparametric state-space framework that assumes no fixed functional form and suits short noisy nonlinear ecological data allowing us to ask how plankton interactions change with temperature nutrients and density dependence and how far ahead community dynamics can realistically be forecast. I will describe the data collection pipeline the temporal-gap and uncertainty problems it generates and the reasoning behind our statistical choices with both technical detail and intuition for non-specialists.