TEHISINTELLEKTIL PÕHINEVAD LAHENDUSED MERENDUSTEADUSTES

Prof. Urmas Raudsepp
Department of Marine System

TALLINN UNIVERSITY OF TECHNOLOGY

AI IS A DATA-DRIVEN APPROACH THAT LEARNS FROM DATA AND EXPERIENCE RATHER THAN RELYING SOLELY ON PREDEFINED RULES.

- **Theory: Rule-based simulation vs. data-driven emulation**

Many scientific and engineering problems can be expressed as:

$$\mathbf{y} = \mathbf{F}(\mathbf{x})$$

where:

x = inputs

y = outputs or observables

F = the underlying process or model that maps inputs to outputs

- **Simulation:**
  In simulation, **F** is known or defined from physical laws, equations, or rules.
  Given **x**, we compute **y**.

- **Emulation:**
  In emulation, AI learns a fast approximation of the underlying function **F** from records of data pairs **(x, y)**.

TAL TECH

MOTIVATIONS AND CONCEPTUALIZATIONS

• Data availability Historical records (x,y) may already exist, allowing us to learn the relationship directly from data.

Simulation → **exact but expensive**

Emulation → **approximate but fast**

Fast prediction Once trained, the emulator function can estimate outputs almost instantly.

- A big system simulation can produce accurate outputs, but it often obscures internal relations:

too many coupled modules, nonlinear feedbacks, high-dimensional inputs/outputs, …

- A partial emulator focuses on one target and a subset of inputs: This can remove the curtain from a certain window, reveals which variables matter, shows direction and strength of influence, exposes nonlinearities / interactions, etc

Simulation gives the whole movie;

Emulation can spotlight one scene and explain why it happens!

TAL TECH

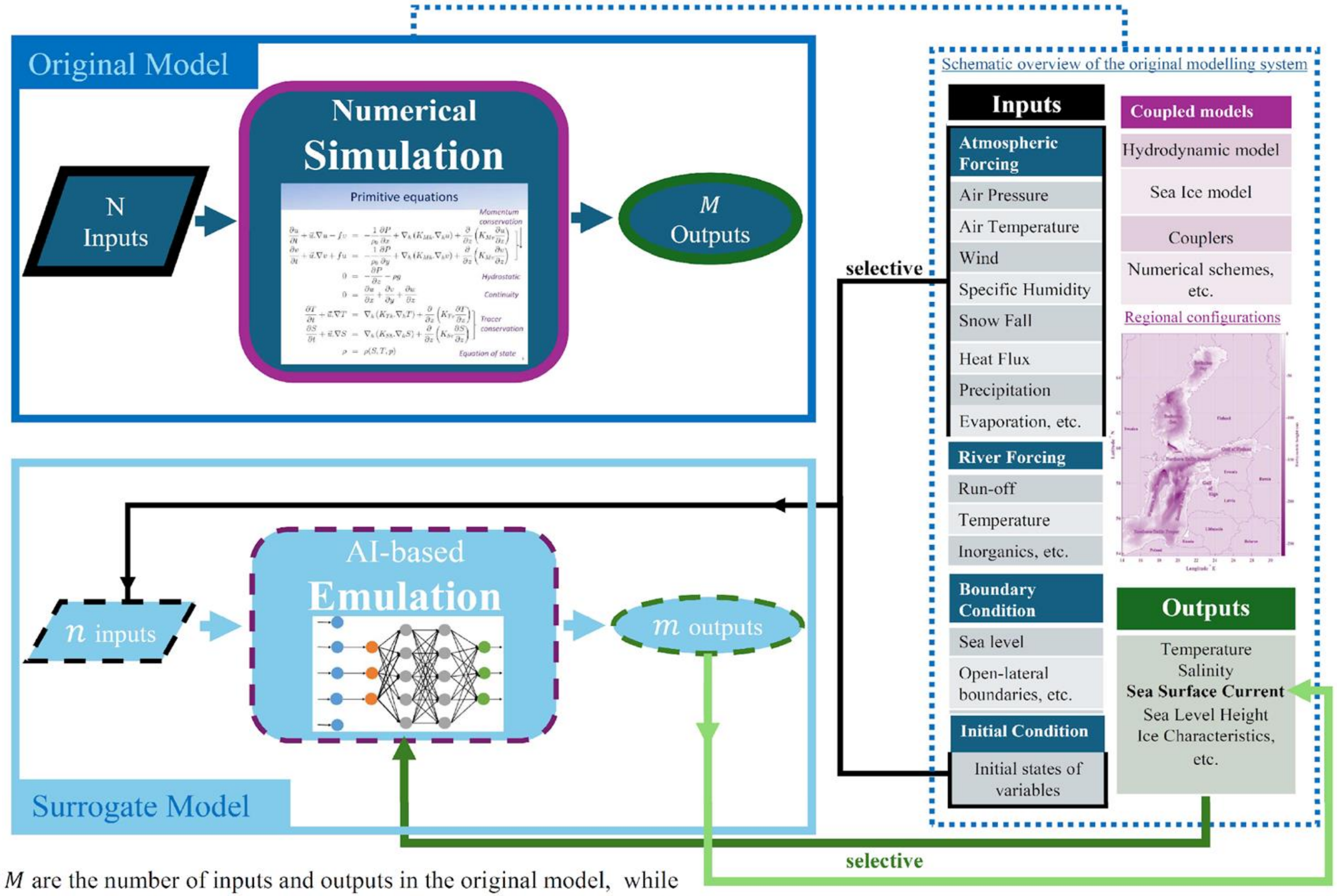
OCEAN MODELLING CASE STUDIES

Original model

- Physics-based numerical simulation
- Represents full system dynamics
- Computationally expensive for long-term runs

Surrogate model

- AI-based (data-driven) emulator
- Learns the input–output behavior of the original model from past simulations records
- Uses key input variables to estimate selected target outputs
- Reduces the need for repeated full-model runs in future

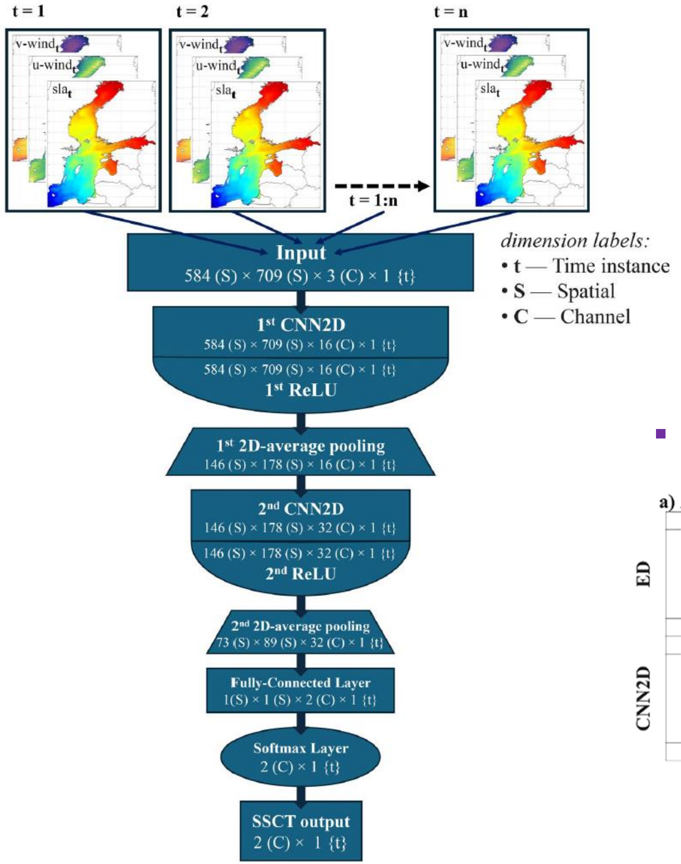


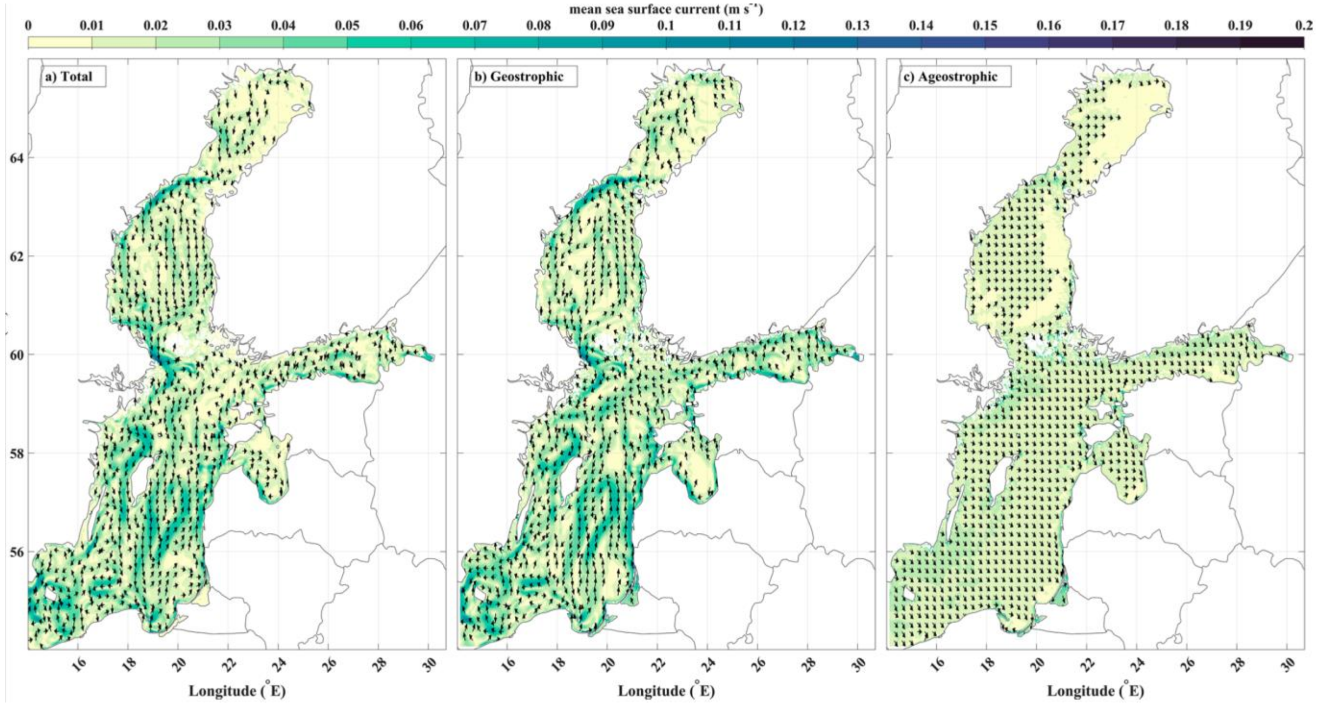
N and M are the number of inputs and outputs in the original model, while $n \ll N$ & $m \ll M$ denote the reduced number of "selected" inputs and outputs in the surrogate model.

TAL TECH

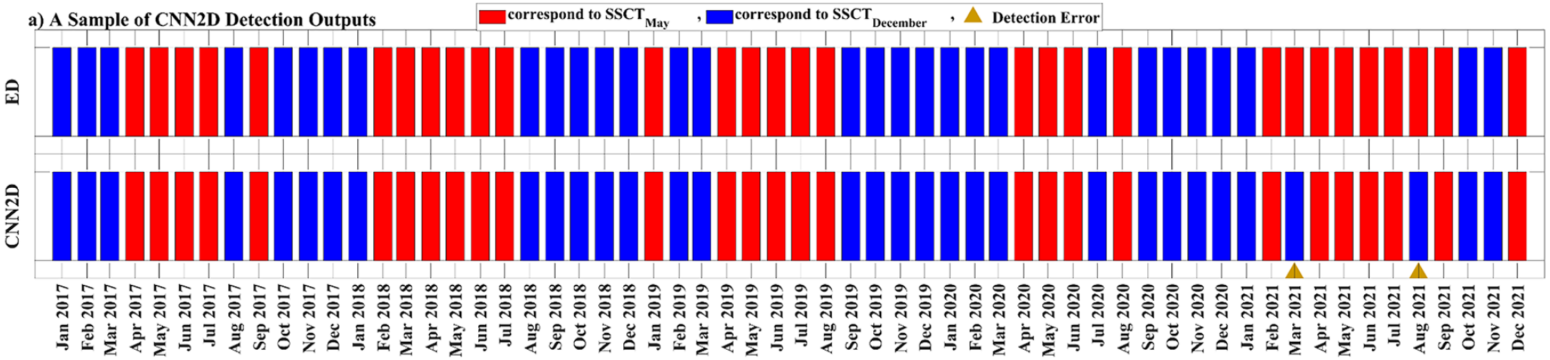
CASE STUDY 1:
PATTERN RECOGNITION OF SEA SURFACE CIRCULATION

- Uses Baltic Sea records of sea surface circulation, together with wind and sea level height
- A convolutional neural network (CNN) is trained to learn how sea surface circulation type (SSCT) respond to wind forcing and sea level variability





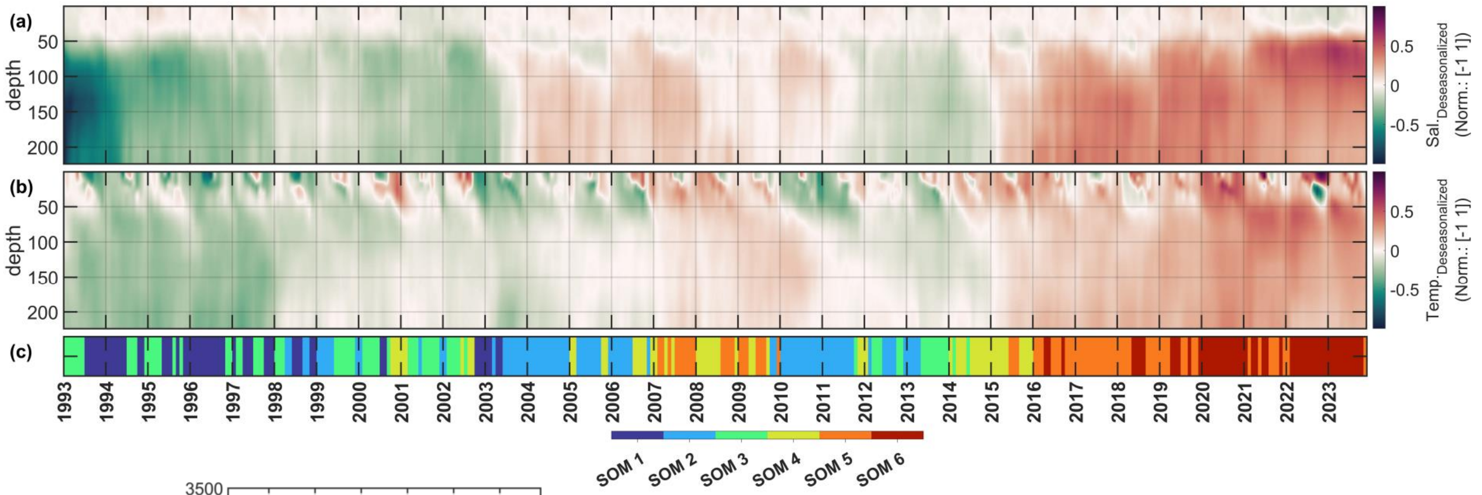
- Enables the identification of general circulation patterns using only wind and sea level data

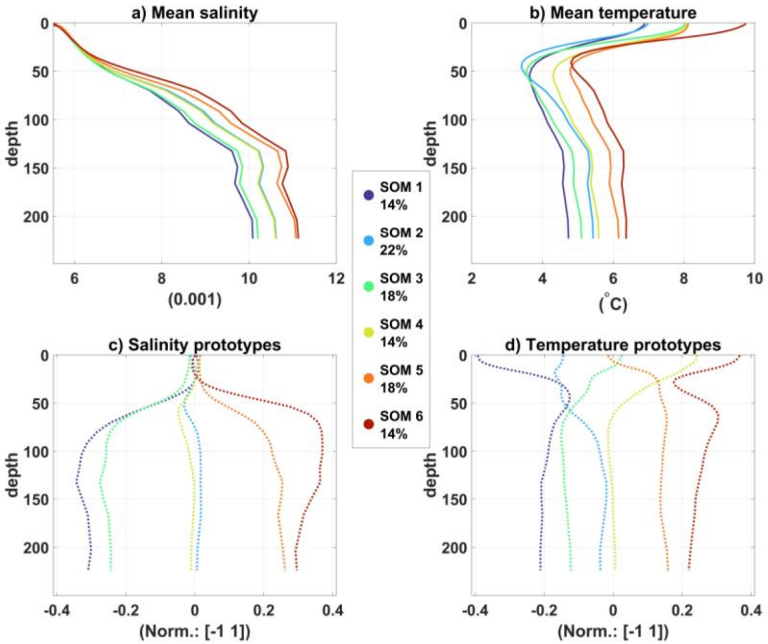


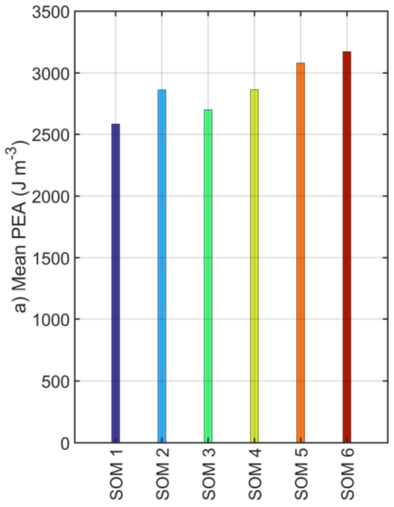
TAL TECH

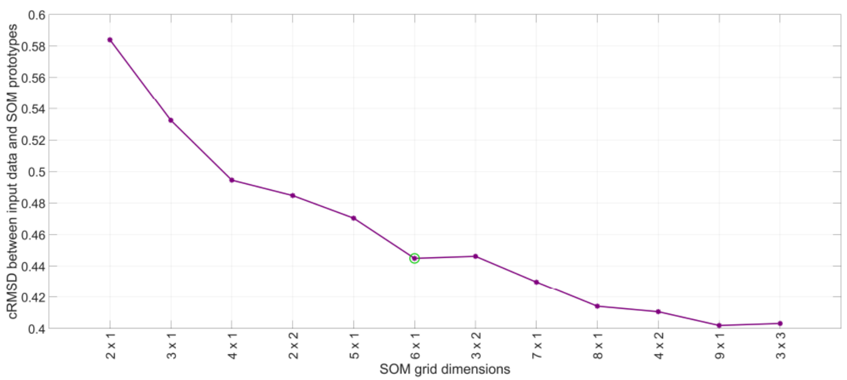
CASE STUDY 2:
BALTIC SEA HYDROGRAPHICAL REGIME IDENTIFICATION

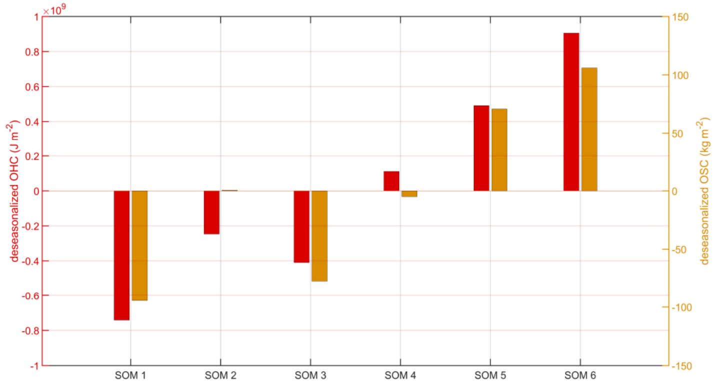
- A Self-Organizing Map (SOM) was trained on multivariable temperature and salinity profiles across the Baltic Sea
- Based on the last 30 years of data, six distinct hydrographical regimes were identified
- Each regime shows clear differences in water mass structure and hydrographic conditions







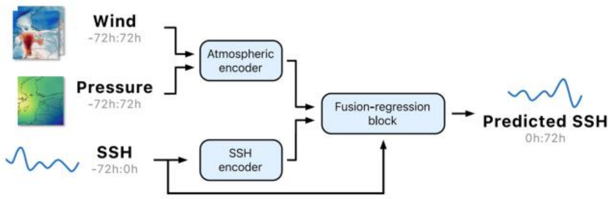


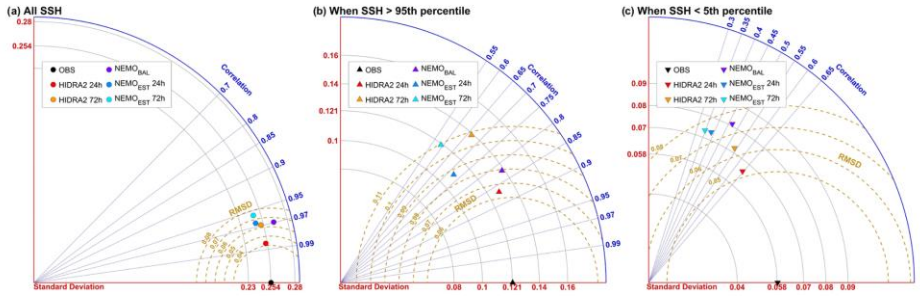


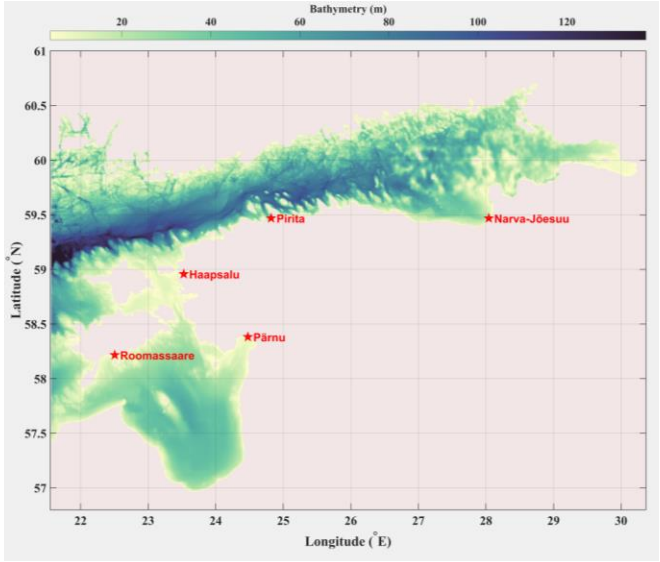
- These regimes support the interpretation of variations in heat and salt content, inflow–outflow activity, and thermodynamically driven changes in stratification

CASE STUDY 3:
SEA LEVEL FORECASTING ALONG THE ESTONIAN COAST

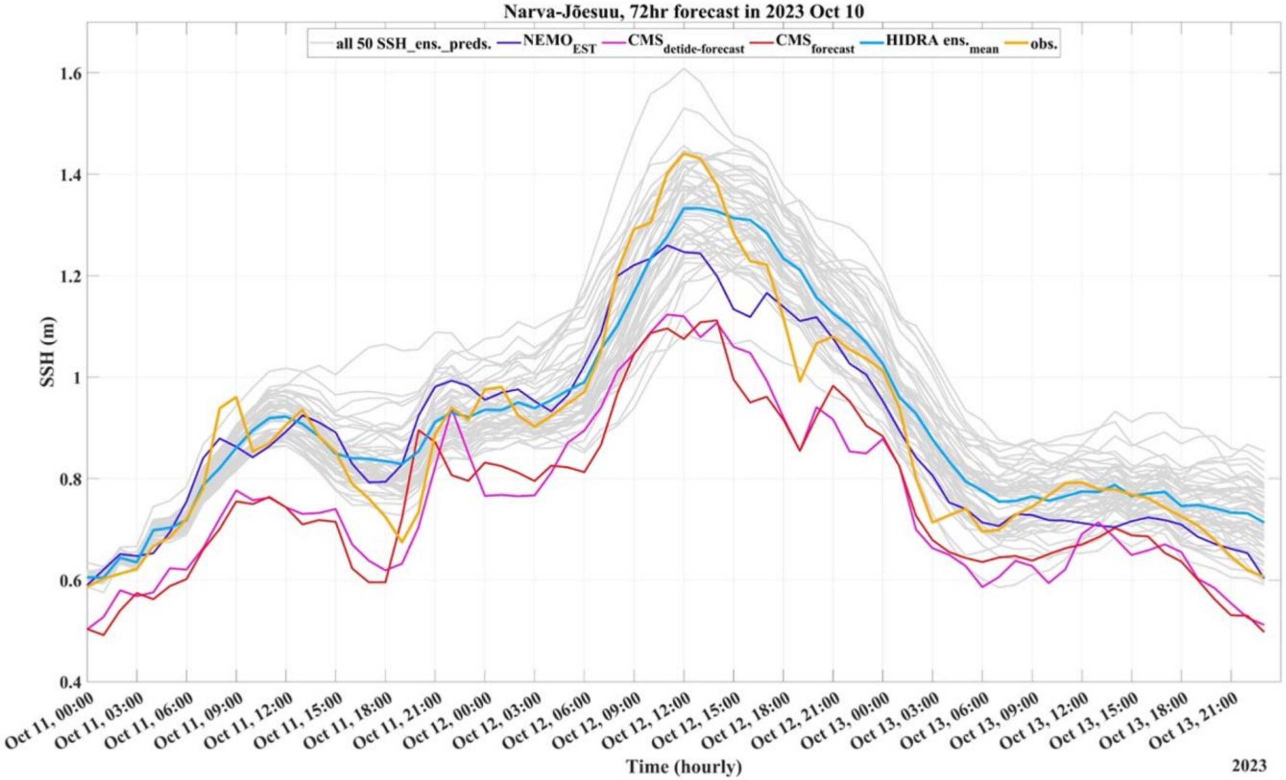
- Ensemble-based deep learning

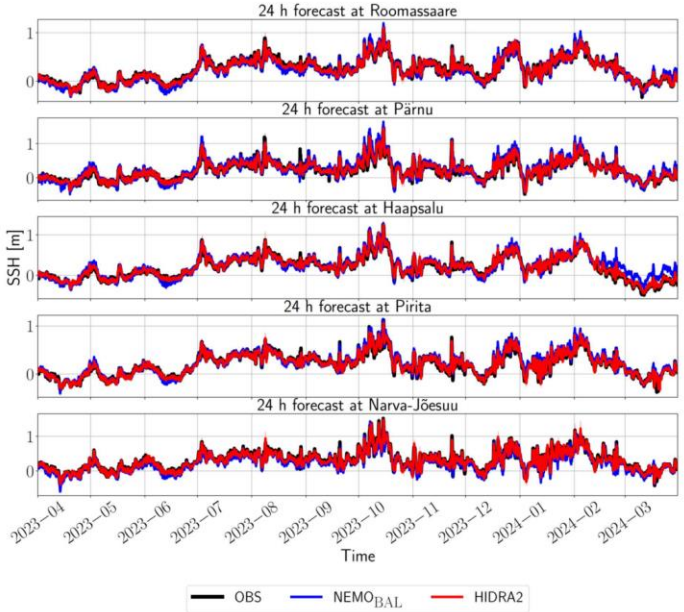






- Forecasts observed sea level from 24 hours to 72 hours ahead
- Delivers reliable storm surge forecasts

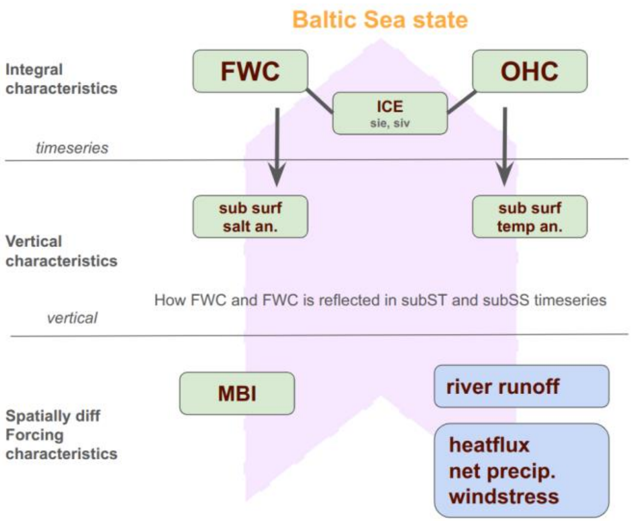




TAL TECH

- Outperforms traditional hydrodynamic models
- Supports operational forecasting, hazard warning, and coastal management

CASE STUDY 4:
AI-BASED ASSESSMENT OF THE PHYSICAL STATE OF THE BALTIC SEA



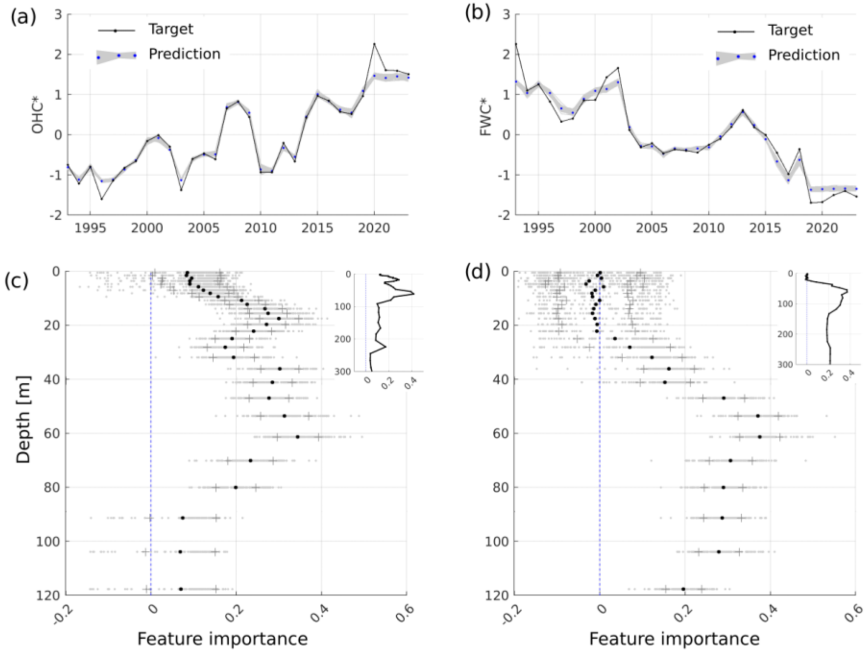
- A machine learning method based on an ensemble of decision trees
- Used to detect statistical relationships between physical state variables and forcing drivers
- Identifing the most influential drivers of Baltic Sea variability

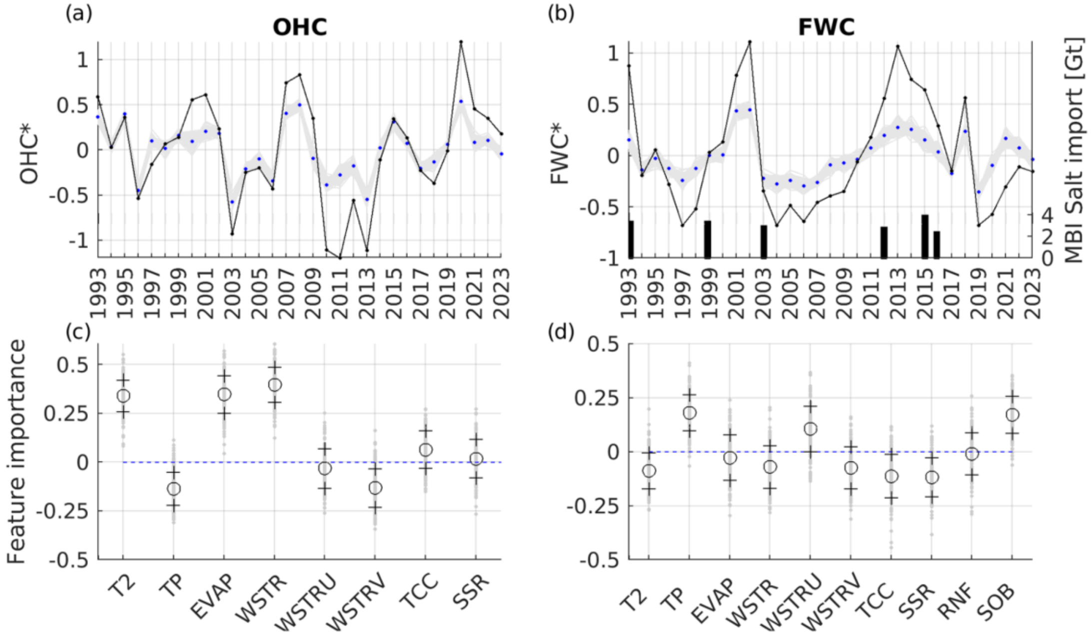
State variables:

Ocean Heat Content (OHC), Freshwater Content (FWC), subsurface temperature, subsurface salinity

Drivers:

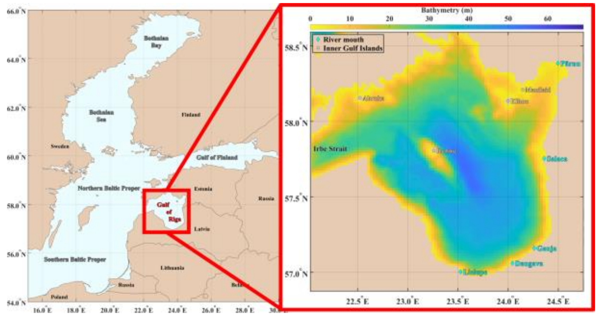
air temperature, evaporation, wind stress, precipitation, salt inflow, boundary salt transport, river runoff



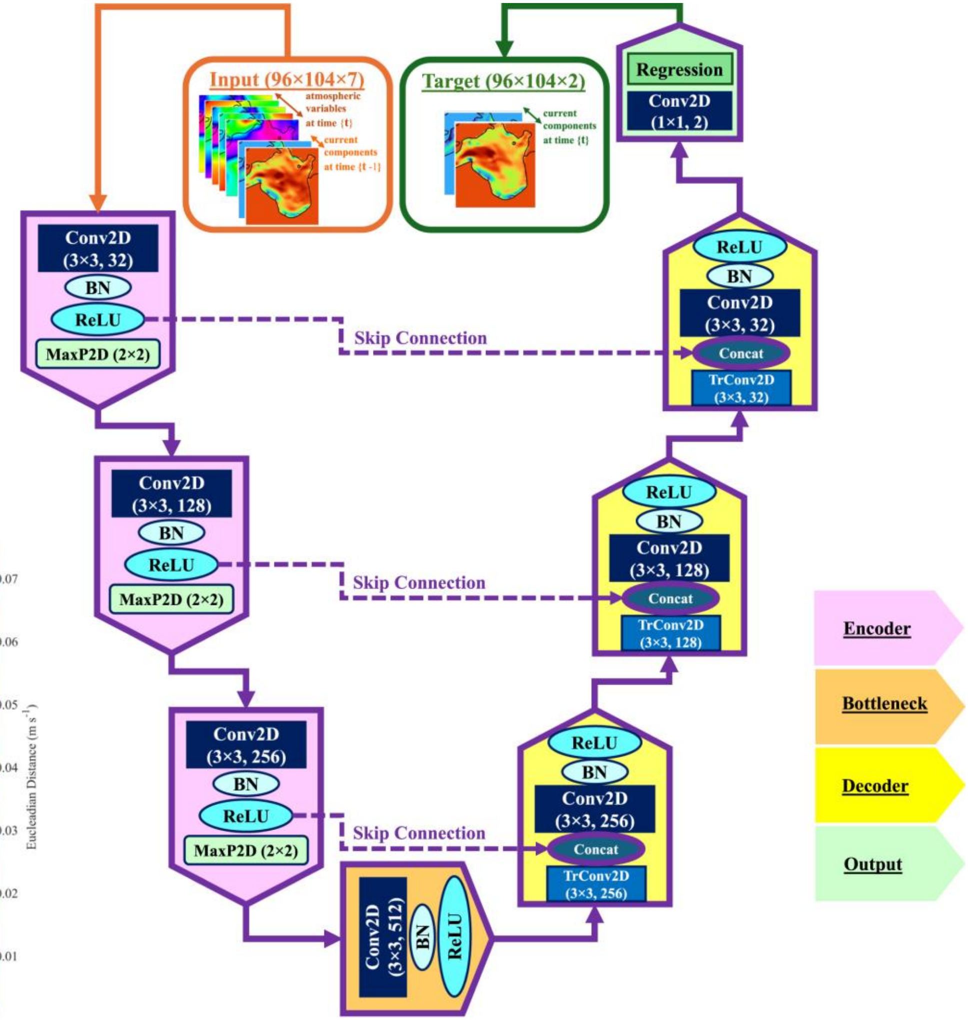


TAL TECH

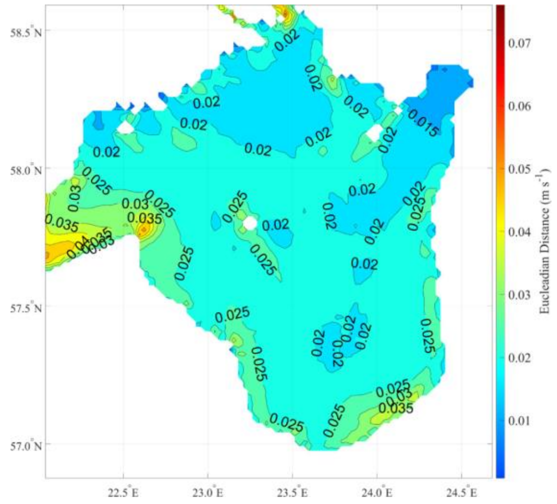
CASE STUDY 5:
AI-BASED PREDICTION OF SEA SURFACE CURRENTS IN THE GULF OF RIGA

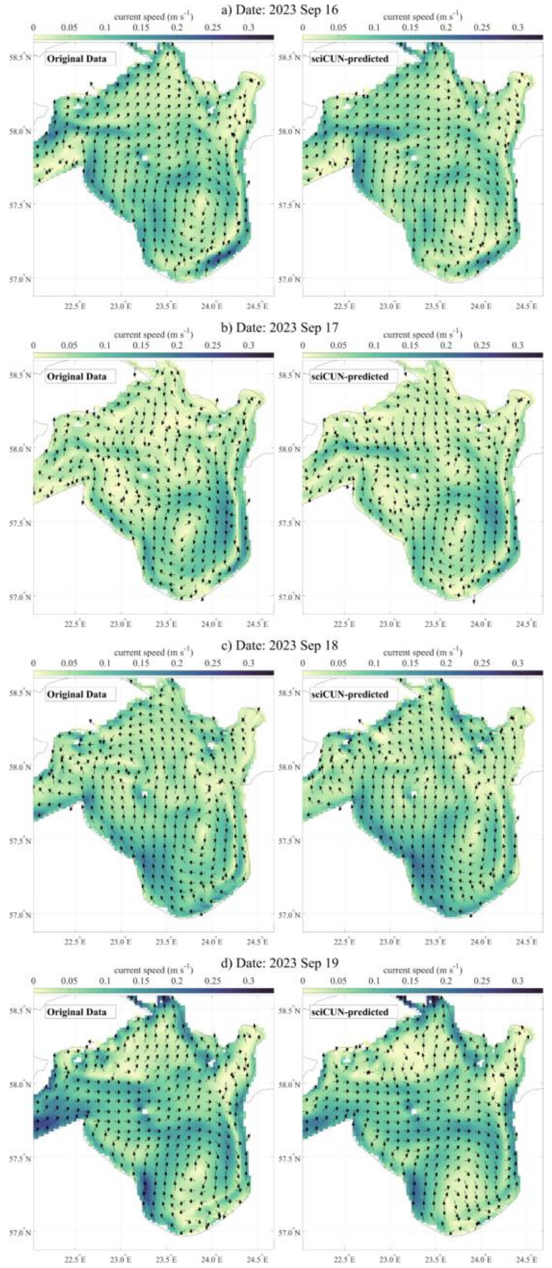


- surrogate deep learning model (Emulator) designed to learn what a computationally expensive hydrodynamic model delivers
- High-accuracy and fast prediction of sea surface current fields in the Gulf of Riga.



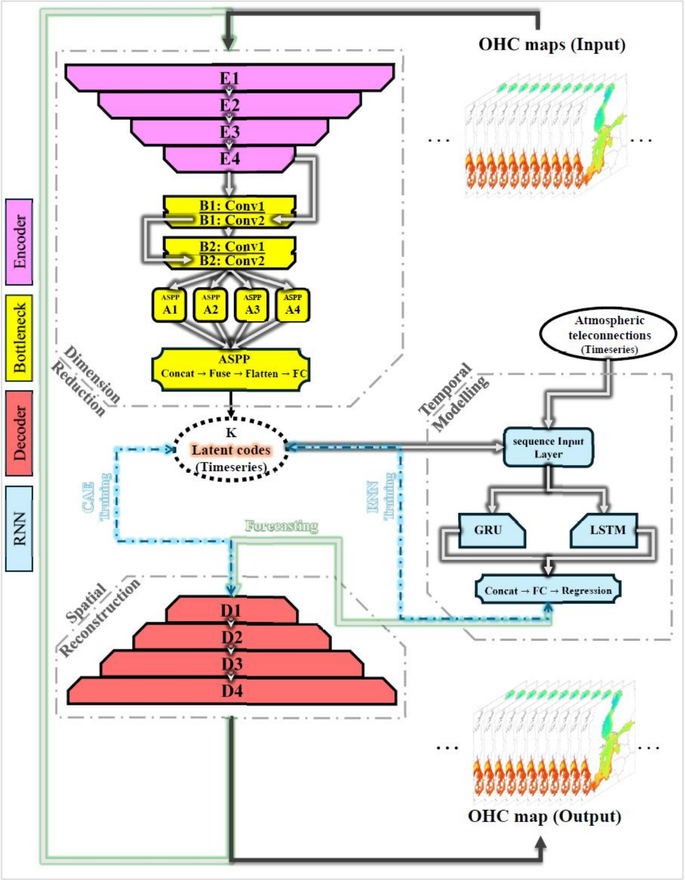
- Supports coastal and navigation-related decision-making
- Helps marine research and environmental studies
- Provides fast synthetic data for larger models and enables rapid generation of current fields for coupled modeling

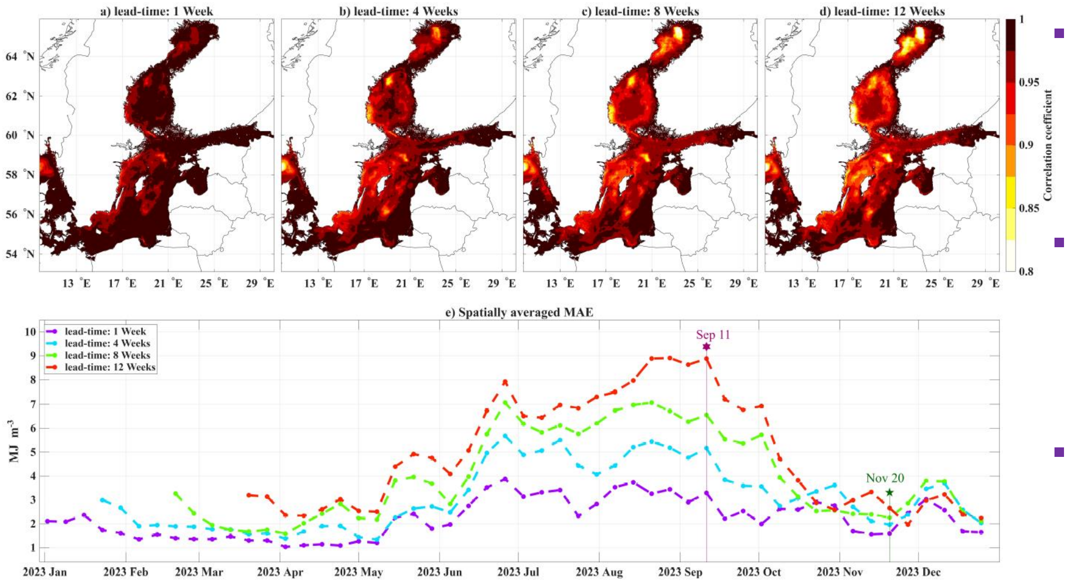


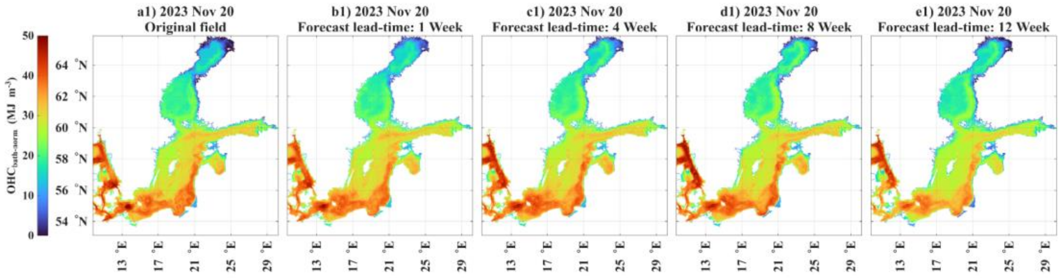


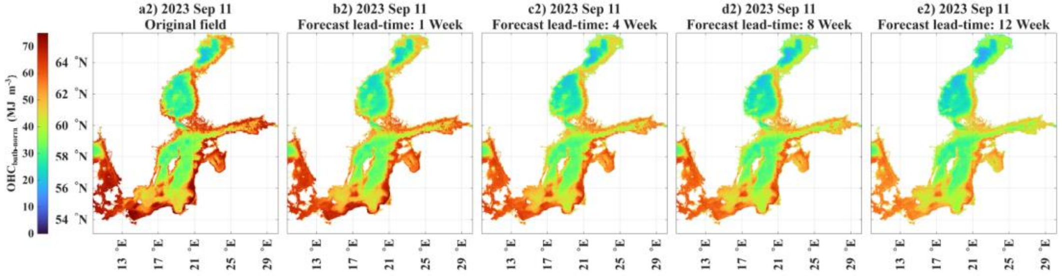
TAL TECH

CASE STUDY 6:
SUBSEASONAL FORECAST OF BALTIC SEA HEAT CONTENT

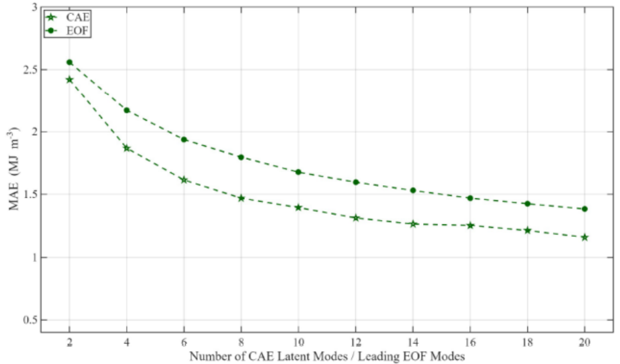






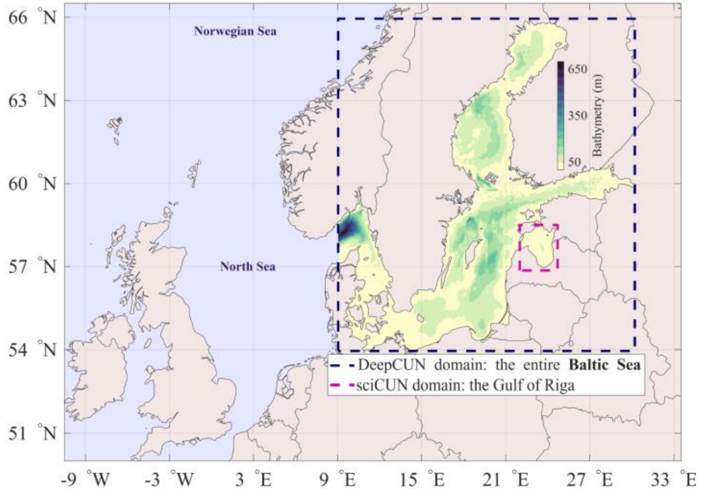


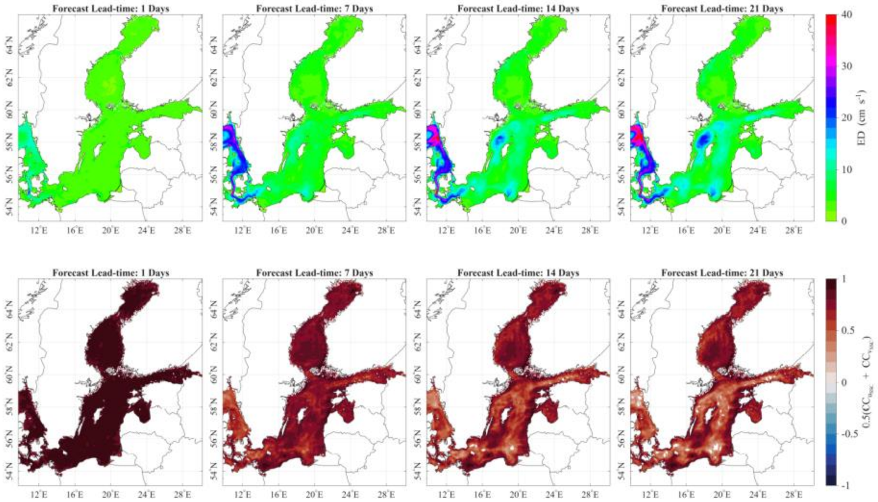
- A novel two-stage Autoencoder and Recurrent Neural Network was designed for high-dimensional spatial data
- Climate signals were integrated as conditions for forecasting mapped ocean heat content distribution
- It supports efficient forecasting of large-scale ocean heat variability and our AI-model showed acceptable accuracy up to 12 weeks ahead
- It outperforms traditional methods and is useful for climate-informed marine analysis and monitoring



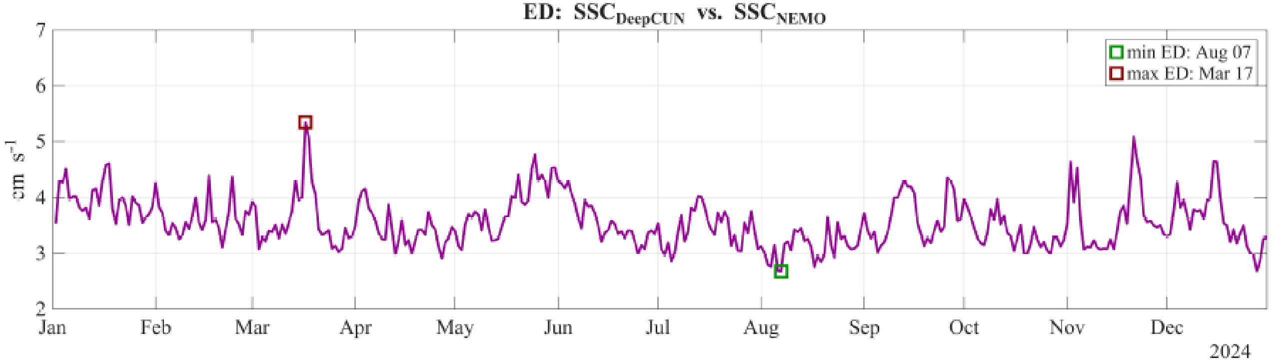
TAL TECH

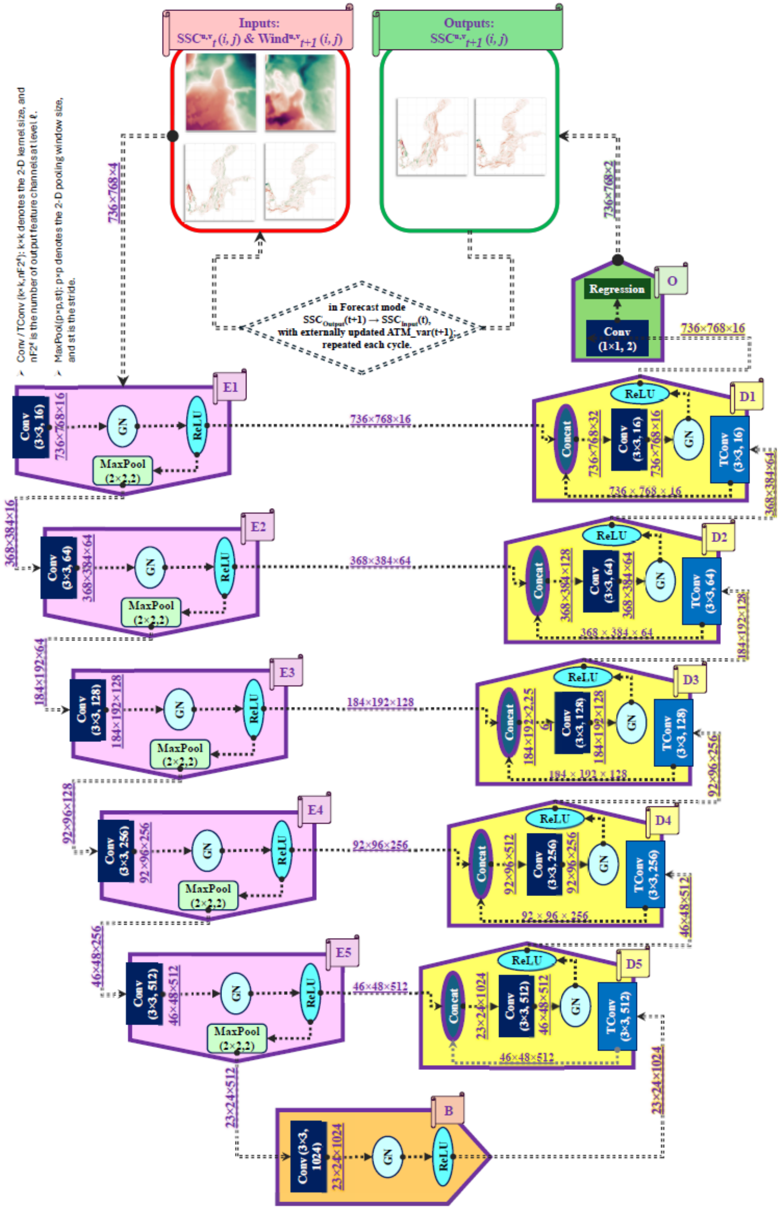
CASE STUDY 7:
EXPLAINABLE AI (XAI) FOR SEA SURFACE CURRENT FORECAST





- High-accuracy and fast prediction of daily sea surface current fields across the Baltic Sea. The forecasting framework showed satisfactory prediction skill up to 21 days
- Two innovative XAI frameworks were applied to interpret how the model transforms input data into output predictions.
- Useful for forecasting, model understanding, and marine decision support





TALTECH

REFERENCES

- Barzandeh, A., Maljutenko, I., Rikka, S., Lagemaa, P., Männik, A., Uiboupin, R., & **Raudsepp**, **U**. (2024). Sea surface circulation in the Baltic Sea: decomposed components and pattern recognition. *Scientific Reports (Nature)*, *14*(1), 18649.
- Gudimetla, V. V., Hayatijozani, A., Barzandeh, A., Maljutenko, I., Uiboupin, R., Männik, A., & **Raudsepp**, **U**. (2025). Self-Organizing Map–Derived Regimes as a Basis for Basin-Integrated Diagnostics of Baltic Sea Hydrographic Change (1993–2023). *Under Review by Scientific Reports (Nature)*
- Barzandeh, A., Ličer, M., Rus, M., Kristan, M., Maljutenko, I., Elken, J., ... & Uiboupin, R. (2025). Application of the HIDRA2 deep-learning model for sea level forecasting along the Estonian coast of the Baltic Sea. *Ocean Science*, 21(4), 1315-1327.
- **Raudsepp**, **U**., Maljutenko, I., Lagemaa, P., & von Schuckmann, K. (2025). A new conceptual framework for assessing the physical state of the Baltic Sea. *State of the Planet*, 6, 1-17.
- Barzandeh, A., Maljutenko, I., Rikka, S., & **Raudsepp**, **U**. (2026). sciCUN: A deep learning model for daily sea surface current fields inference–A case study of the Gulf of Riga. Ocean Modelling, 102693.
- Hayatijozani, A., Barzandeh, A., Maljutenko I., & **Raudsepp**, **U**. (2026). Convolutional Autoencoder Latent-Space Modeling with Climate-Index–Conditioned Recurrent Networks for Assessing Subseasonal Potential Forecast Skill of the Baltic Sea Heat Content. *Accepted by Engineering Applications of Artificial Intelligence*
- A., Barzandeh, A., Maljutenko, I., Rikka, S., & **Raudsepp**, **U**. (2026). Data-Driven Emulation of Numerically Simulated Baltic Sea Surface Currents with a Deep Convolutional U-Net: Explainability and Potential Forecast Skill. *Under Review by Journal of Geophysical Research: Machine Learning and Computation*

THANK YOU

FOR YOUR

ATTENTION

TALLINN UNIVERSITY OF TECHNOLOGY

taltech.ee/en