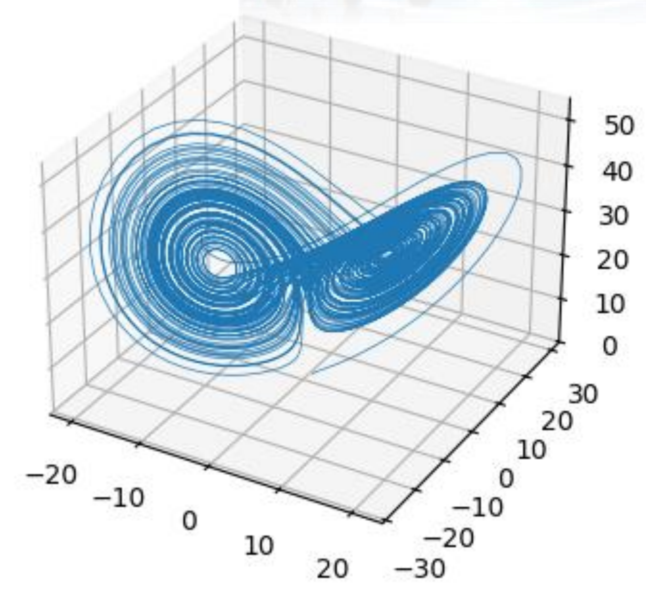
Python and Machine Learning for Weather, Climate and Environment

Roland Potthast

with contributions by Stefanie Hollborn, Jan Keller, Thomas Deppisch, Mareike Burba, Matthias Mages, Sarah Heibutzki, Marek Jacob, Florian Prill, Tobias Göcke, Nastaran Najari

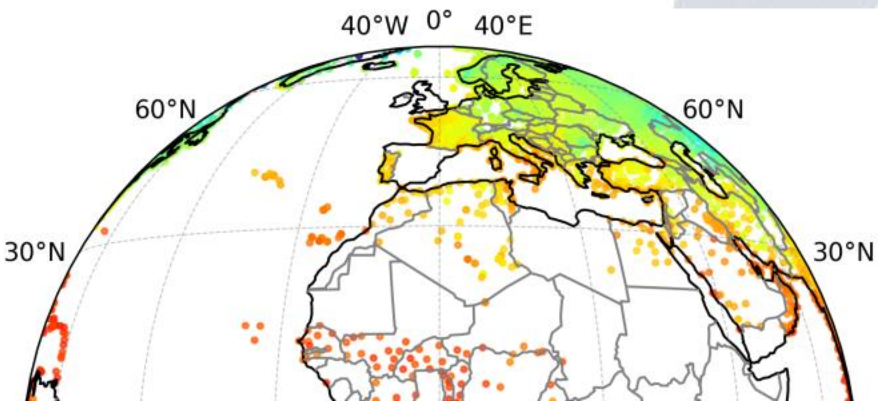
EUMETNET

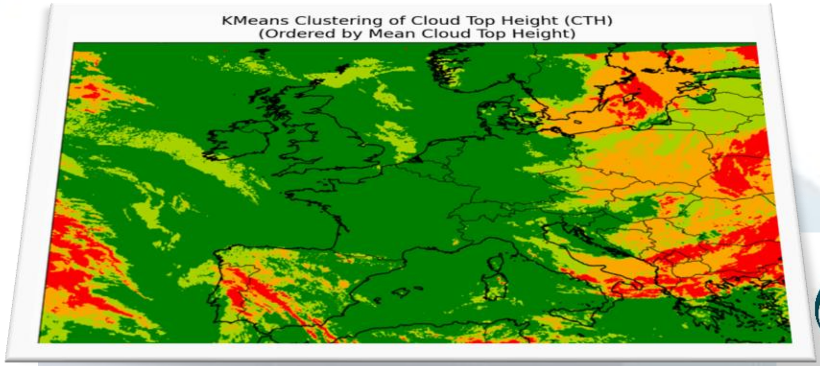
Contents Day 1



EUMETNET

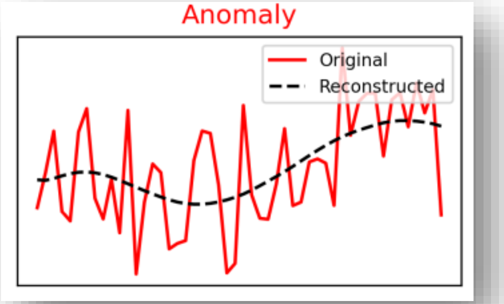
Contents Day 1

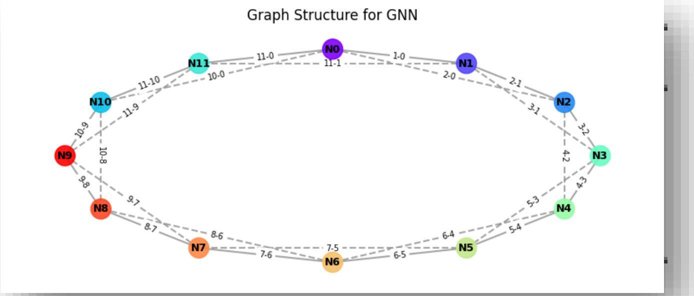




EUMETNET

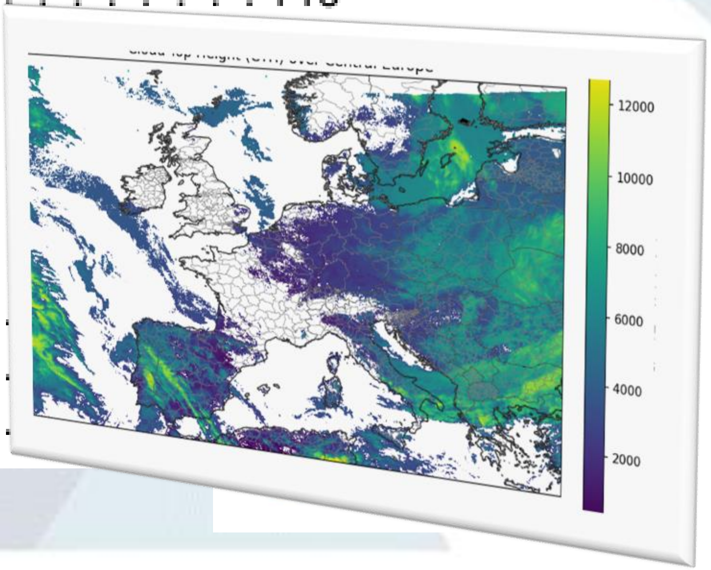
Contents Day 2





NET

Contents Day 3



Contents Day 4-6

Day 4: ML Architectures, Agents and LLM Framework DAWID

Day 5: MLFlow, MLOps, Fine Tuning LLMs

Day 6: Anemol, AICON and AI Data Assimilation

EUMETNET

Tutorial Philosophy

1) Learn to master things yourself
2) Work in small groups & help each other
3) Work from elementary understanding towards larger packages

4) Learn simple examples first

5) We need to understand the whole range of ML/AI techniques today in context

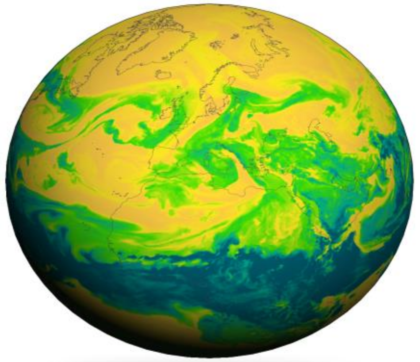
Building AI Expertise with Python and AI Workflows

To effectively integrate AI, we must ensure our teams are skilled in both AI methods and operational

- **Establishing structured learning paths for key AI techniques relevant to weather and climate modeling.**
- **Using Python libraries such as numpy, eccodes, netcdf, and xarray for handling large meteorological datasets.**
- **Training teams in machine learning frameworks such as TensorFlow, PyTorch, and Hugging Face Transformers.**
- **Setting up end-to-end AI workflows in Jupyter-based environments, covering data ingestion, training, validation, and inference.**
- **Encouraging collaboration between meteorologists, model developers, and AI experts to foster cross-disciplinary innovation workflows.**

EUMETNET

Replacing and Hybridizing Forecasting Systems with AI



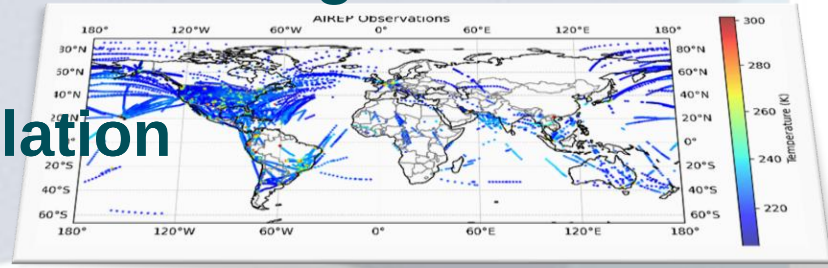
**Some forecasting components will be fully replaced by AI,
while others will integrate AI as a hybrid solution. Key shifts include:**

- **AI-Based Nowcasting: AI-driven short-term weather predictions using real-time observational data (e.g., satellite, radar, sensors), enhancing or replacing conventional nowcasting techniques.**
- **Neural Weather Models: Deep learning models trained on historical data can generate competitive forecasts at lower computational cost.**
- **Hybrid AI-NWP Models: AI enhances physics-based forecasting through bias correction, uncertainty quantification, and ensemble optimization.**
- **Machine Learning for Subgrid Processes: AI could improve or replace empirical parameterizations in turbulence, cloud physics, and convection models.**
- **Automated Impact Forecasting: AI-driven models provide direct risk assessments for extreme weather events, minimizing reliance on manual interpretation.**

EUMETNET

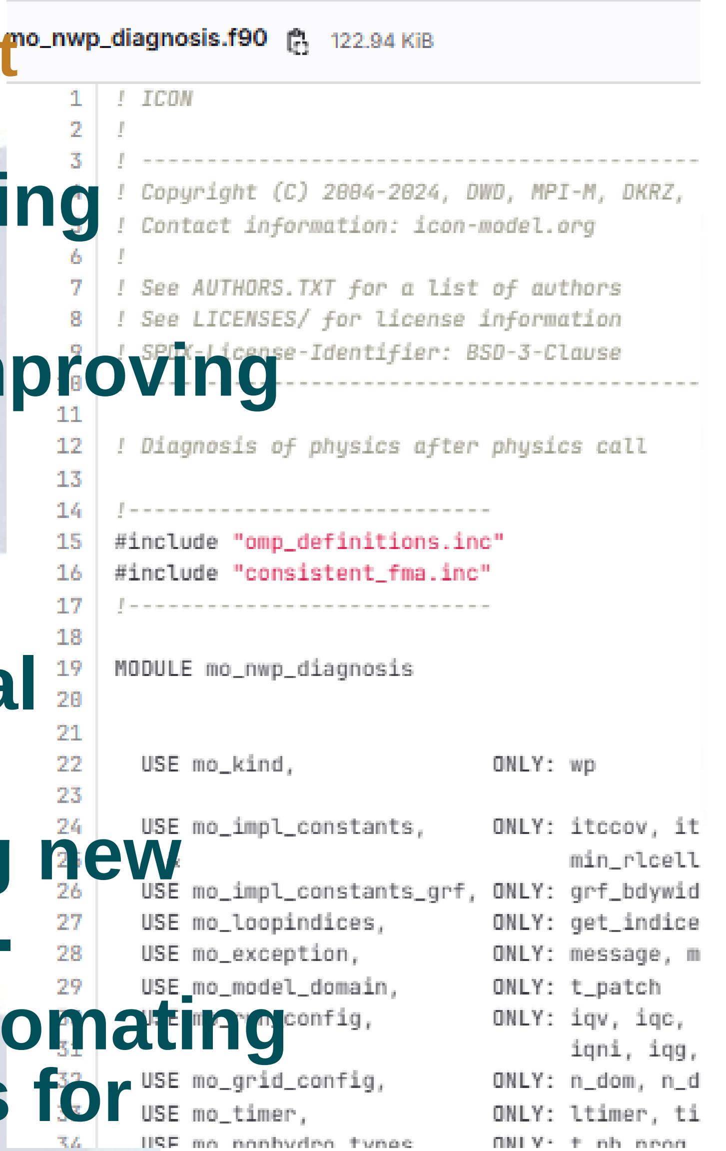
AI in Data Assimilation and Learning Directly from Observations

- **Machine Learning for Observation Processing: AI-driven quality control of observational data, filling data gaps and detecting sensor anomalies.**
- **AI-Based Data Assimilation: AI improving assimilation processes by optimizing observation ingestion.**



- **Deep Learning for Data Assimilation: AI learning complex relationships between observations and model states, accelerating assimilation workflows.**
- **End-to-End AI Data Ingestion: Future AI models trained directly on observational datasets, potentially reducing reliance on classical assimilation techniques.**
- **Self-Learning Systems: AI dynamically adjusting to new data, improving continuously without manual recalibration.**

Using AI for Code Refactoring and Model Development

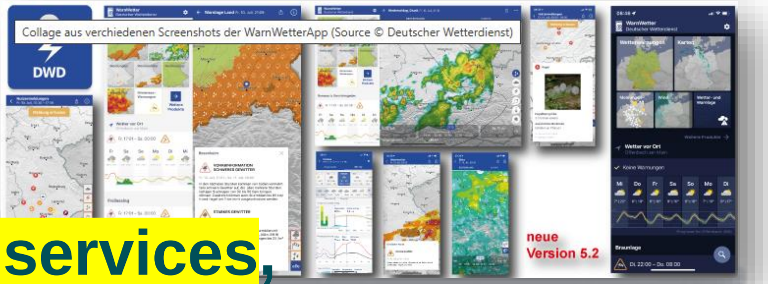


AI also modernizes modeling workflows, improving efficiency in research and development:

- **Refactoring Legacy Code: AI-assisted tools improving Fortran, C++, and Python models for better maintainability and performance.**
- **Automated Model Optimization: AI tuning hyperparameters and optimizing computational performance.**
- **AI-Assisted Scientific Discovery: AI identifying new climate and weather patterns in large datasets.**
- **AI-Generated Documentation and Testing: Automating documentation and generating validation tests for numerical models.**

EUMETNET

Transforming Services and User Interaction with AI



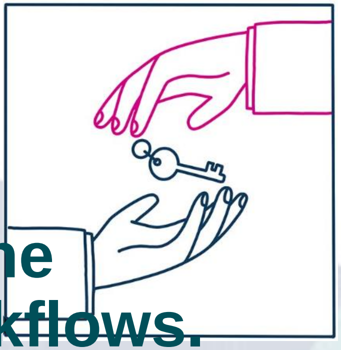
Collage aus verchiedenen Screenshots der WarnWetterApp (Source © Deutscher Wetterdienst)

AI enables new ways to deliver weather and climate services, improving automation, personalization, and accessibility:

- **AI-Generated Weather Reports: Natural language generation models translating raw data into meaningful insights for different user groups.**
- **Conversational Forecasting Assistants: AI chatbots and voice assistants allowing users to interactively query weather and climate predictions.**
- **Real-Time Impact Forecasting: AI models directly linking weather forecasts to risks in agriculture, energy, transportation, and disaster management.**
- **AI-Powered Data Visualization: Interactive AI tools allowing users to explore, manipulate, and interpret complex weather datasets.**

EUMETNET

A Clear Migration Strategy for AI Transformation



1. AI Readiness Assessment: Identify areas where AI provides the highest impact while ensuring compatibility with existing workflows.
2. Pilot AI Replacements: Test AI-based forecasting models in parallel with traditional methods before full adoption.
3. Hybrid Deployment Strategy: Introduce AI-driven improvements in stages, ensuring fallback options are in place.
4. AI Validation and Trust Building: Develop transparent evaluation metrics for AI models to ensure trust and reliability.
5. Workforce Training and Knowledge Transfer: Enable teams to transition smoothly from traditional methods to AI-driven solutions.
6. Continuous AI Governance: Establish guidelines for AI model retraining, performance monitoring, and ethical considerations.

EUMETNET

Limitations and Responsible Use of AI

While AI offers transformative opportunities in forecasting, modeling, and service delivery, it is crucial to acknowledge its current limitations and apply it with scientific caution:

- **Data Requirements**: Most AI models rely on large, high-quality datasets and perform poorly in data-sparse or non-stationary environments.
- **Lack of Physical Consistency**: AI predictions may violate conservation laws or produce unrealistic results in rapidly evolving scenarios.
- **Limited Interpretability**: Unlike traditional models, many AI systems operate as black boxes, making it difficult to understand or trace their internal reasoning.
- **Bias and Overfitting**: Biased or unbalanced training data can lead to flawed predictions, while overfitting to historical data may reduce adaptability to changing climate conditions.
- **Need for Rigorous Validation**: AI models must be continuously monitored, validated, and benchmarked to ensure stability, fairness, and scientific reliability. Validation needs metrics and scores beyond traditional forecasting metrics.
- **Complementary Role**: AI should be seen as an enhancement to—not a replacement for—physics-based modeling, supporting a hybrid approach for trustworthy innovation.

EUMETNET


Be careful, look closely!
Lets get into it now …

CONTACT DETAILS

Prof. Dr. Roland Potthast

E-AI Programme Manager

Dr. Marek Jacob

E-AI Programme Coordination

EIG EUMETNET
European Meteorological Services' Network
www.eumetnet.eu

EIG EUMETNET Secretariat
c/o L'Institut Royal Météorologique de Belgique
Avenue Circulaire 3
B-1180 Brussels, Belgium

Registered Number 0818.801.249 - RPM Bruxelles

Tel: 0049 69 8062 2713

Email: Roland.Potthast@dwd.de