A novel approach for a Digital Twin to explore future climate extremes to assess impacts

EGI2024 Conference, Sep 30 – Oct 2, 2024
Lecce, Italia

Christian Pagé, CERFACS
Use Case Lead
Anne Durif, CERFACS
Research Engineer

Funded by the European Union
The interTwin project is funded by the European Union - Grant Agreement Number 101058386

The Challenge: Climate Extremes

Generic detection algorithm:

- Intense rainfall
- Drought
- Heatwave
- Cold spell
- High wind

Different methods for one-time events (intense rainfall, high wind) vs. longer-term events (droughts,...)

Characterization (What-if Scenarios):

- Frequency of occurrence
- Spatial extent
- Intensity (if relevant)
- Duration





Pictures: The Guardian, World Meteorological Organization, WMO, Patch, Direct Energy

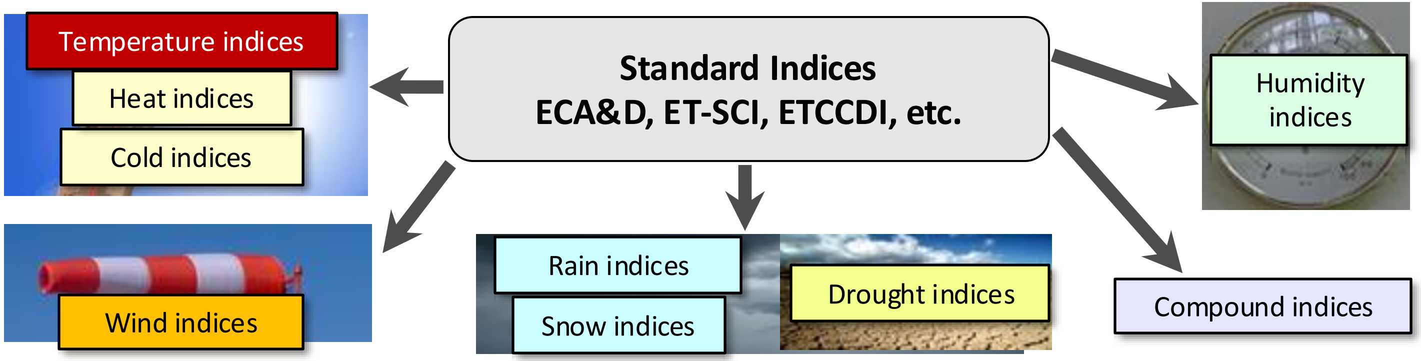
The Challenge: Climate Extremes

2021 Germany Erftstadt, southwest of Cologne

2020 Hurricane Delta causes damage to Louisiana's Gulf Coast

- Urgent needs of impact assessments
- Identify mitigation solutions
- Extreme events attribution
- Multiple domains: infrastructures, urban, agriculture, transportation, etc.
- Flexible tools needed for very diverse users

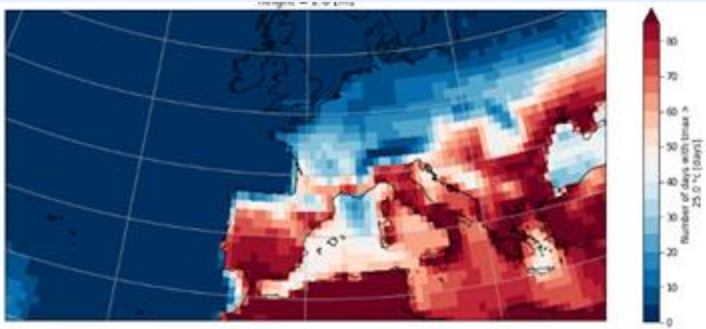
Climate Indices and Indicators



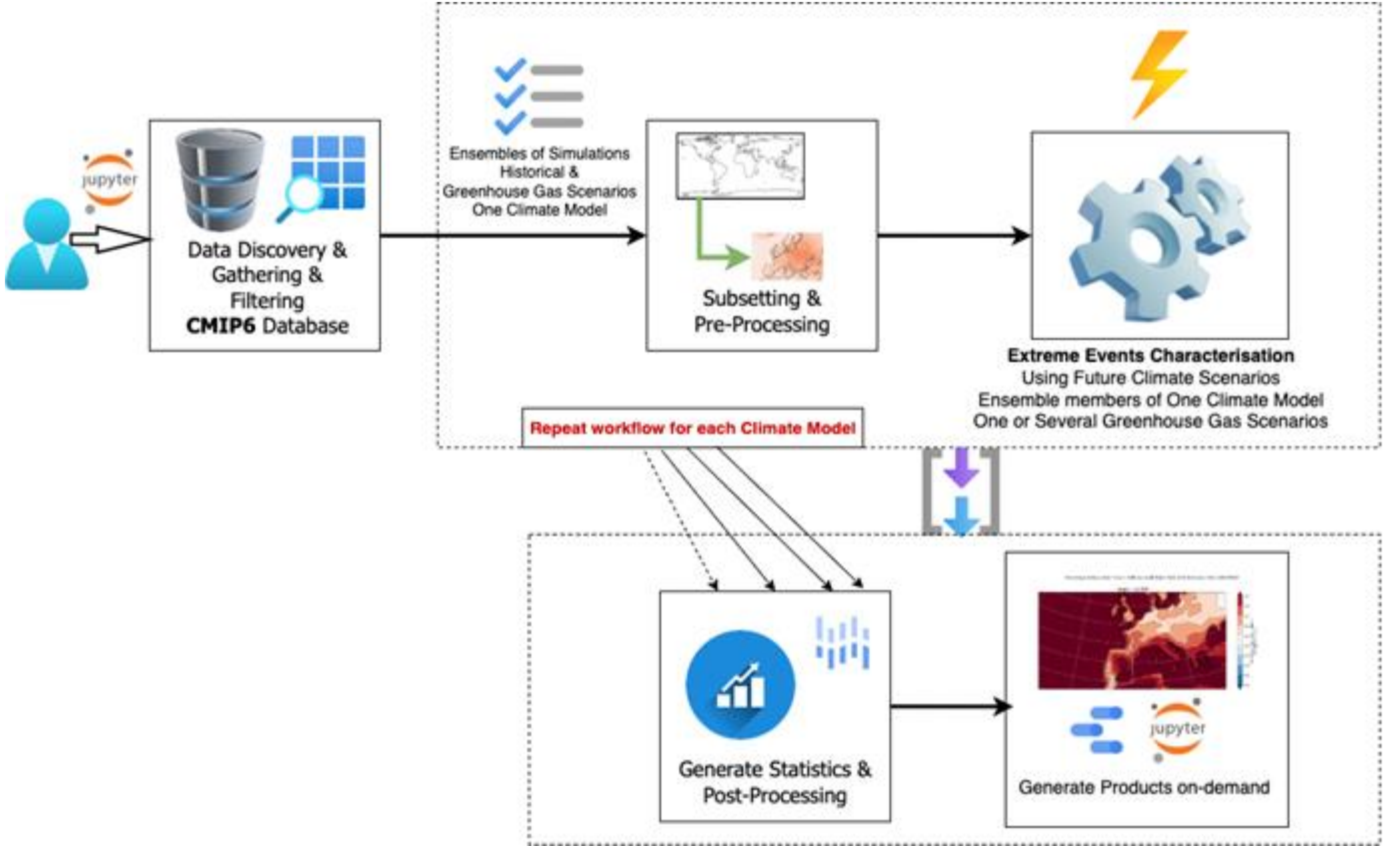
- Intra-period extreme temperature range [°C] - **ETR**
- Warm days (days with mean temperature > 90th percentile of daily mean temperature) - **TG90p**
- Summer days (days with max temperature > 25 °C) - **SU**
- **...**

icclim python package

https://github.com/cerfacs-globc/icclim

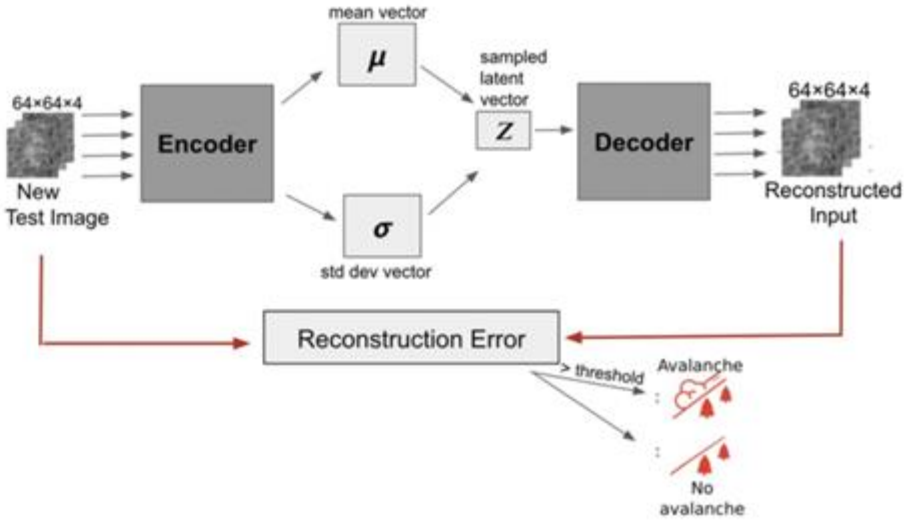


Extreme Workflow: the User Perspective



Why use a AI based method?

- To analyze a very large database of climate scenarios with a good performance
- Use efficiently new architectures (GPUs)
- Scalability in cloud-based environments
- Extreme Events spatial structures are similar to avalanches
  - Variational Autoencoder: Deep Learning Technique

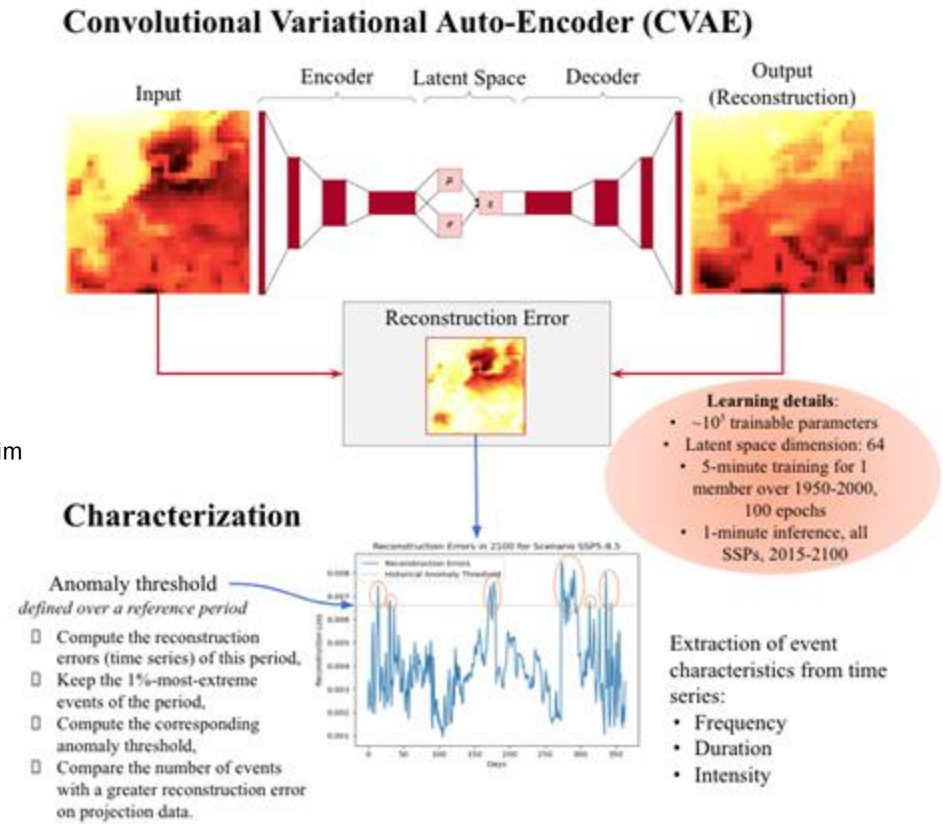


Sinha, Saumya & Giffard-Roisin, Sophie & Karbou, Fatima & Deschatres, Michael & Karas, Anna & Eckert, Nicolas & Coléou, Cécile & Monteleoni, Claire. (2020). Variational Autoencoder Anomaly-Detection of Avalanche Deposits in Satellite SAR Imagery. 113-119. 10.1145/3429309.3429326.

Deep Learning Method

xtclim python package

https://github.com/cerfacs-globc/xtclim



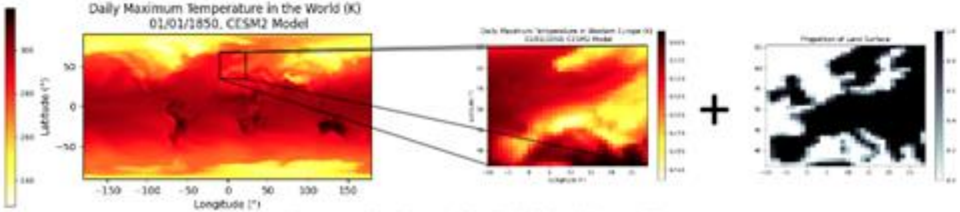
Data…

Raw Data

- Coupled Model Intercomparison Project, phase 6 (CMIP6)
- General Circulation Models (e.g. CMCC-ESM2)
- 1°×1° resolution (~125km spatial grid)
- Daily data from 1850 to 2100
- Climate variables: temperature, precipitations, wind…
- Various carbon emission scenarios (IPCC):
    - SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5

Preprocessing

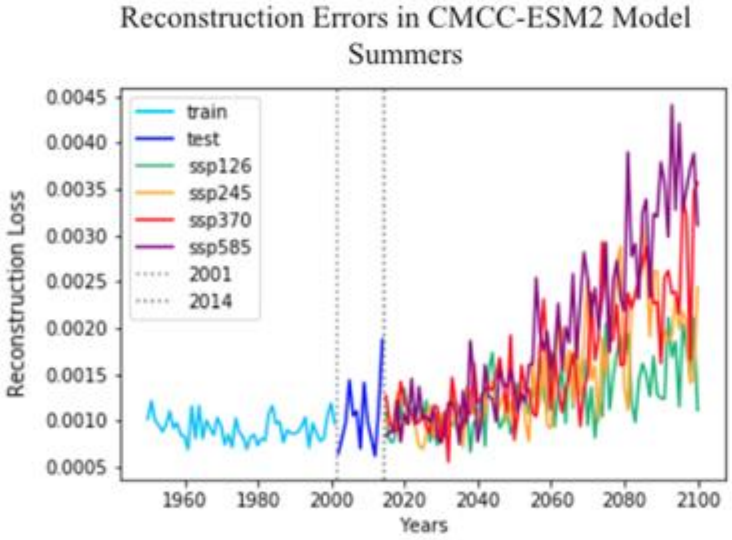
- From NetCDF files to numpy tables
- 32×32 square over Western Europe
- Season split
- Min-max normalization

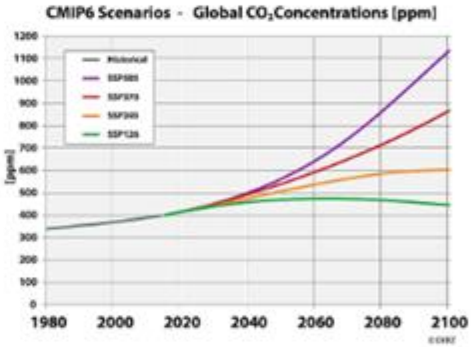


Preprocessing Steps: from World to Western Europe
Normalized and Enhanced with Land-Sea Proportion

Fast comparison between GES

Comparison between IPCC Scenarios





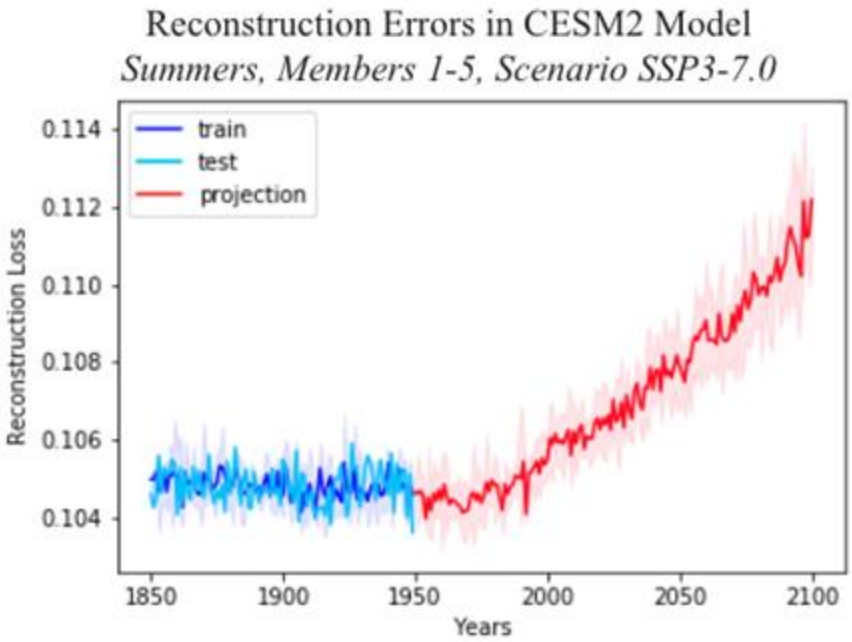
Anomaly Analysis in CMCC-ESM2 Model (Summers)

Detection when the reconstruction error exceeds a threshold

	2001-2014	2015-2100			
Scenario	Test data	SSP1-2.6	SSP2-4.5	SSP3-7.0	SSP5-8.5
Proportion of unusual days	1,00%	3,78%	8,09%	11,07%	17,83%
Maximum spike	0,0051	0,0053	0,0089	0,0087	0,0095
Maximum duration (days)	5	22	27	53	55
Average duration (days)	2	3,44	5,66	6,17	7,51

Large Ensemble Data Analysis

Comparison between ensemble members



Anomaly Analysis in CESM2 Model

Member Comparison for Scenario SSP3-7.0

	1850-1950	1950-2100				
Member n°	Test Data - 5	SSP3-7.0 - 1	SSP3-7.0 - 2	SSP3-7.0 - 3	SSP3-7.0 - 4	SSP3-7.0 - 5
Unusual days	92	3530	3614	3931	3455	3798
Proportion of unusual days	1,00%	25,41%	26,02%	28,30%	24,87%	27,34%
Maximum spike	0,112	0,123	0,119	0,120	0,119	0,121
Average maximum	0,1094	0,1108	0,1108	0,1110	0,1109	0,1110
Maximum duration (days)	7	83	77	81	81	78
Average duration (days)	2,4	9,7	9,9	10,4	9,7	10,5
Proportion of spikes	39	91	91	94,25	88,75	90,25

Integration with interTwin DTE

- Integration with itwinai and DTE Core Components (unlocks additional functionalities)
  - Extensive Logging
  - Hyper-Parameters' Optimization
- Containerization

itwinai / use-cases / xtclim / pipeline.yaml

r-sarma Bug fixes and addition of CERFACS use-case (#151)

Code Blame 31 lines (29 loc) · 922 Bytes

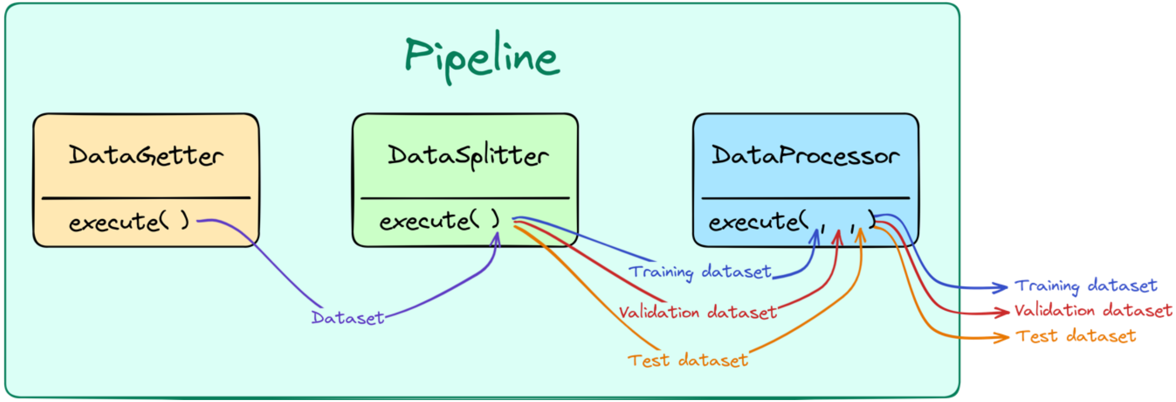
```yaml
# General configuration
dataset_root: '/p/scratch/intertwin/datasets/cerfacs/'
epochs: 3
batch_size: 10
lr: 0.001
scenario: '245'
strategy: 'ddp'

# Workflows
pipeline:
   class_path: itwinai.pipeline.Pipeline
   init_args:
     steps:
         preprocessing-step:
           class_path: preprocessing.preprocess_functions_2d_ssp.PreprocessData
           init_args:
               dataset_root: ${dataset_root}
               scenario: ${scenario}
         preprocessing-split-step:
           class_path: preprocessing.preprocess_2d_seasons.SplitPreprocessedData
           init_args:
               scenario: ${scenario}
         training-step:
             #class_path: src.trainer_dist.XTClimTrainer
           class_path: src.trainer.TorchTrainer
           init_args:
               epochs: ${epochs}
               batch_size: ${batch_size}
               lr: ${lr}
               #strategy: ${strategy}

```

Integration with interTwin DTE

- Advanced workflow composition (CWL)
- Integration with the interTwin Data Lake
  - Rucio
  - STAC Catalog
  - ENES Data Lake



Source: itwinai Documentation

https://itwinai.readthedocs.io/latest/how-it-works/workflows/explain_workflows.html

Integration with interTwin DTE

Perspectives

- Exploitation of **geospatial information**
- Implementation of a **severity index**
- Exploitation of the **latent space** of the Neural Network
- **Validation** against analytical method for robustness assessment (icclim)
- **Extensions** to other climate variables

Next Steps

Take-Home Messages

- interTwin DTE enables fast Digital Twin Application Development and Exploitation
- The Convolutional **Variational** Auto-Encoder (CVAE) achieves **Unsupervised Anomaly Detection**
- Events can be **characterized** with various indicators
- Results are **consistent**
- Performance of this approach unlocks the ability to better **quantify climate change impact uncertainties**

Thank you! Questions?

Christian Pagé
christian.page@cerfacs.fr

LinkedIn LI @pagechristian