Development of a NOWcasting MachineLearning module for X-band and C-band weather radars (DeNOW) @GRID'5000

Company: SkyEcho v.o.f

Rotterdam, The Netherlands

contact: Igor Stepanov

igor.stepanov@sky-echo.eu

X-band weather-radar hardware, deployed at 153m tall building in Rotterdam, Netherlands

GRID’5000

Hardware used in the DeNOW project

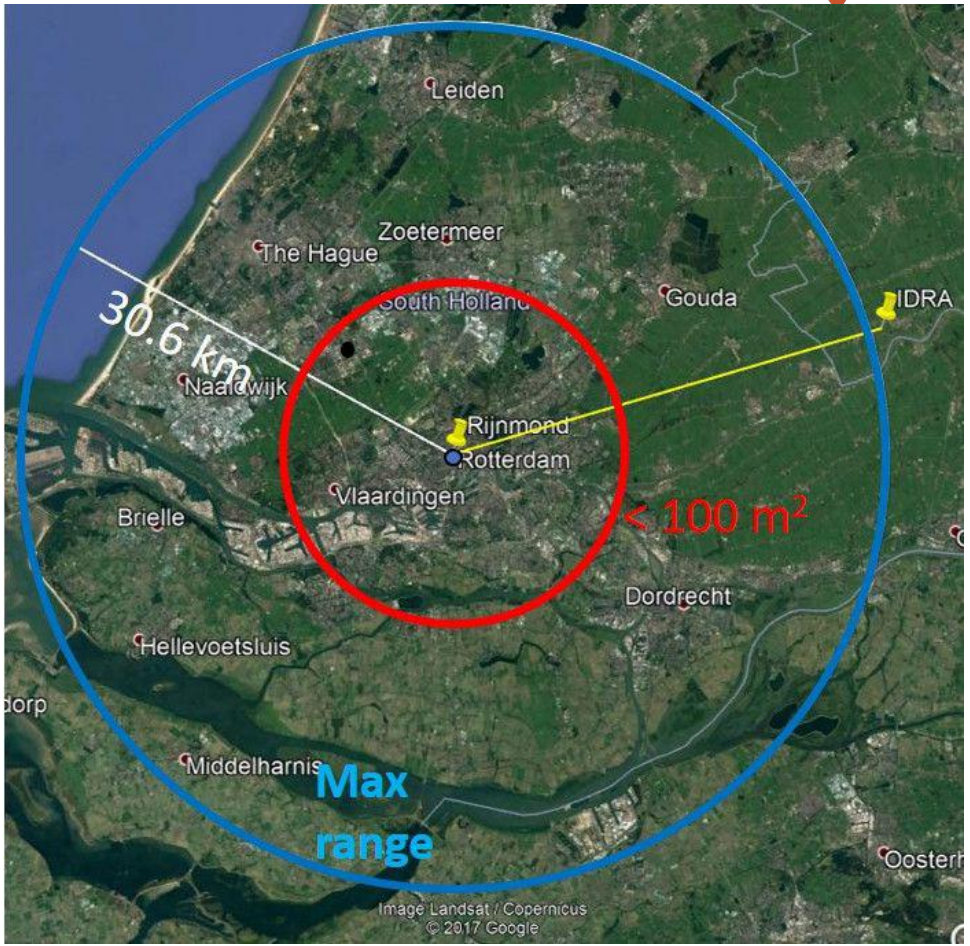
- **Experiment description (max. 4 slides)**
  - Concept and objectives
  - Background and motivation
  - Experiment set-up
- **Project results (max. 3 slides)**
  - Measurements
  - Lessons learned
- **Business impact (min. 4 slides)**
  - Impact on your business, .. how did Fed4FIRE helped you ?
  - Value perceived, .. why did you come to Fed4FIRE ?
- **Feedback (min. 4 slides)**
  - Used resources and tools
  - Added value of Fed4FIRE

Concepts and objectives

SkyEcho v.o.f operates unique high-resolution X-band weather-radar

Location: Rotterdam, NL

Purpose: **100x more accurate rainfall observations**

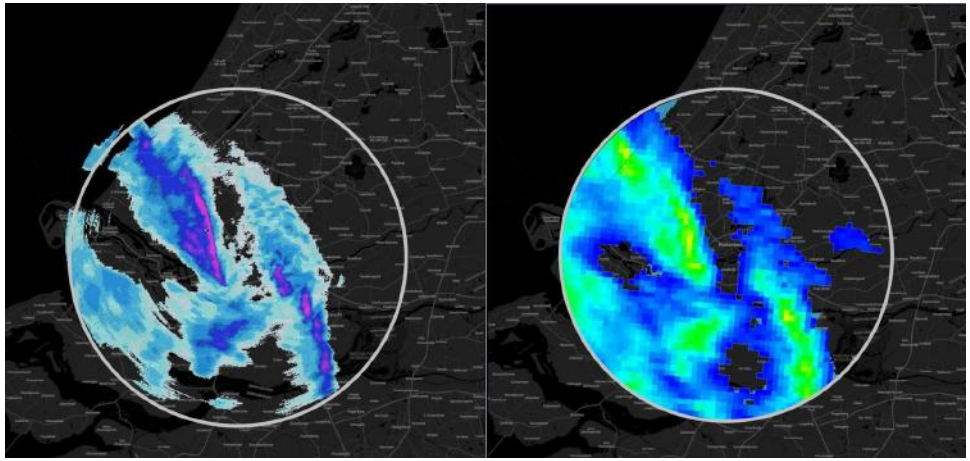


Concepts and objectives

Standard resolutions:
Spatially: 1km
Temporally: 5-minute

X-band radar resolutions:
Spatially: 0.1km
Temporally: 1-minute

100x more accurate observations! But what about the ***forecast accuracy***?



X-band HD-radar

C-band National radar

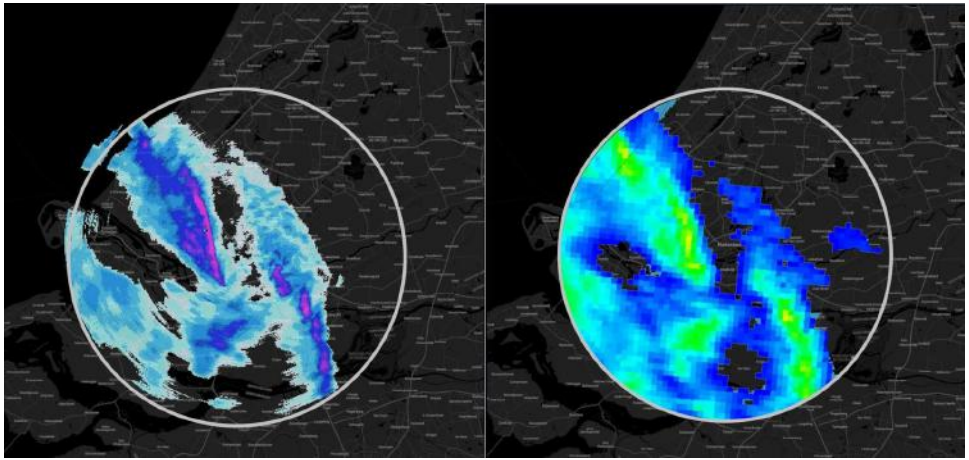
Concepts and objectives - Data processing

- 10x10 matrix, representing 1km2 of rainfall rate data (around target point)
- Removal of all non-rainy days (to avoid biased model that would prefer to predict no rain

→ **Final 3 model layers**: input layer size 100 (10x10 matrix)
: LSTM layer
: output layer size 1 (target rainfall for target point)

HPC testbed used: GRID'5000 in Nancy

- **Total model training time: 7 days** / 2-GPUs / 30 days of rain data
- Job submitted as interactive
- Model training initially carried out on 1-day & 3-days (confirming LSTM model can be used on GRID'5000 HPC)



Concepts and objectives

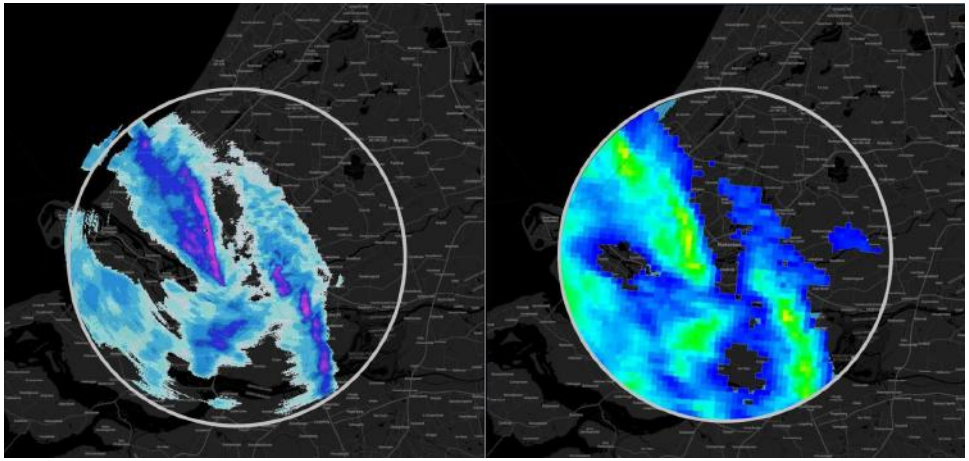
Long short-term Memory - natural fit to the problem

Input rainfall data: 30km radius, with 1-minute interval. Data cleaned to use only first 10km (high resolution) and rainfall <5mm/hr removed.

Step 1: Train the model with 1-month of archive data (cleaned)

Step 2: Evaluate accuracy, using data previously unknown to the model

Enviroment: python libraries, CUDA architecture (GPU)



Project results

Desired milestones:

- Machine Learning model for rainfall rate nowcast, with high accuracy
- Possibility to assimilate results into SkyEcho data cloud (database, API)
- Possibility to further train the model with more data from remaining year seasons
- Possibility to use the same model for radars using different frequencies (C- and S-band of national weather services)

Project results

Data output:

Format: NetCDF

Reason: Assimilation in SkyEcho database, to make DeNOW available with API-endpoints

Trained model for operational purposes:

- currently developed model on GRID’5000 can be deployed in operational cloud of SkyEcho easily

Project results - model evaluation

1. Model trained and evaluated based on same data

   RMSE = 0.0104
   Model error ~ radar accurac error

2. Model evaluated on consecu day (from same month)
   RMSE = 0.3646

3. Model evaluated with random rainfall data (same month)
   RMSE = 0.85

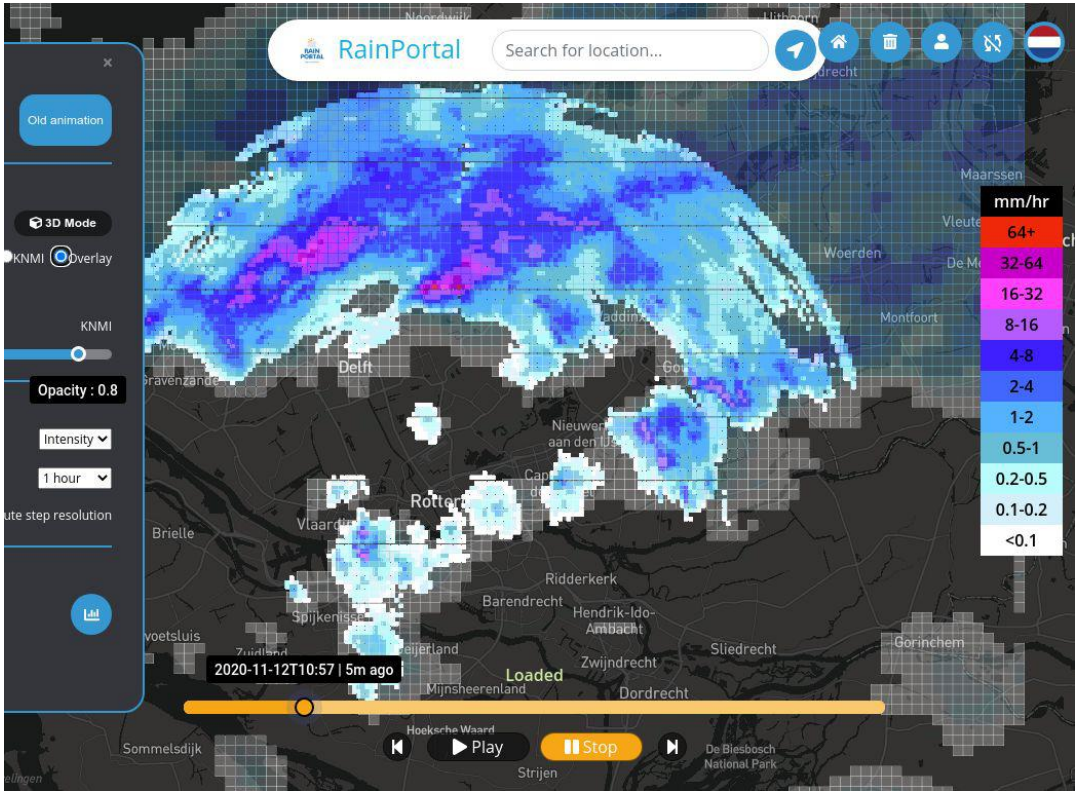
Table evaluation:

Model	RMSE on same days as training	RMSE on a consecutive day	RMSE on a random day
48h data	0.0104	0.3646	0.85

Business impact

Value of Improved forecast

1. Implementation on RainPortal frontend
2. Applications for B2B and B2G
3. Smart City applications
4. *“Pescatarian”* case study 🐟



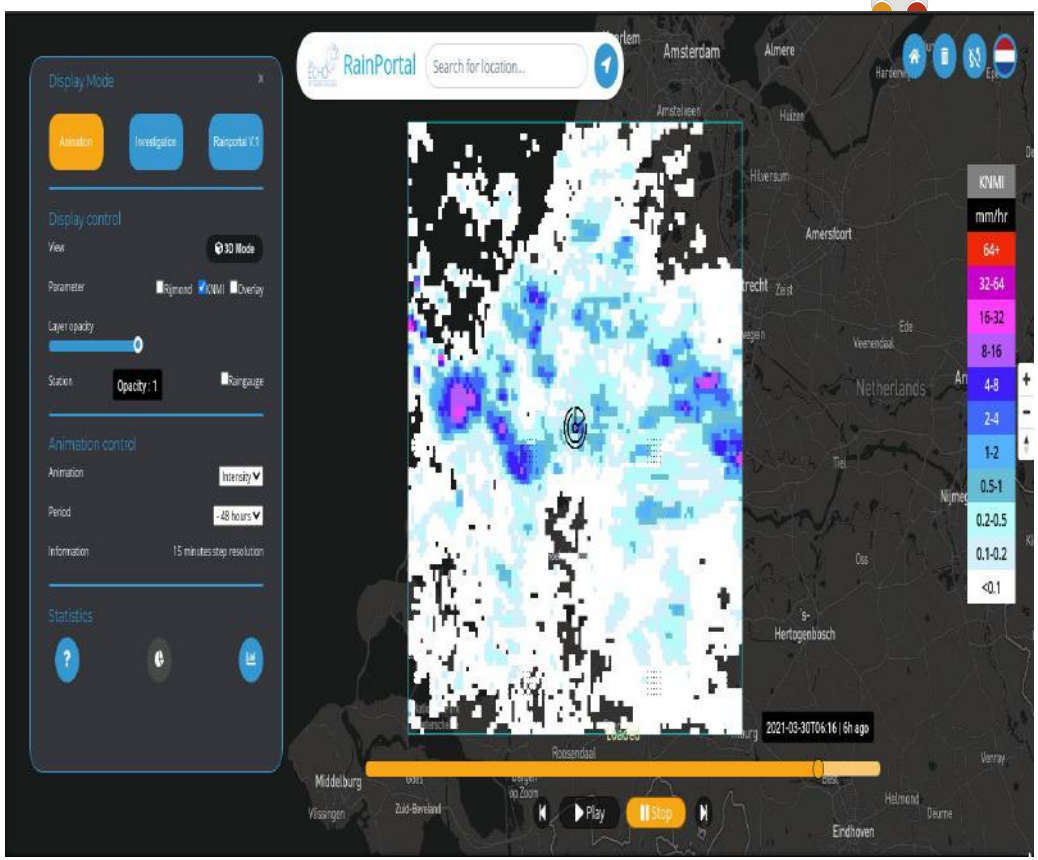
Business impact

1. RainPortal implementation

As data-providers, visibility of data is crucial

Visually showing clients accurate foracast is a **perfect DEMO for investors**

Scalable solution



Business impact

Waterboards of Netherlands

Data usage order for automatic pumps:

Weather models

Standard radar

X-band nowcast for last minute response

Photo source: https://en.wikipedia.org/wiki/Canals_of_Amsterdam

Business impact

Pescatarian case study

Clients already uses SkyEcho current C-band nowcast to visualize in fishing app

But the success depends on quality of forecast!

Photo source: pxhere.com

Business impact

Fed4FIRE+ contribution

Crucial role in providing PLATFORM for connecting with a range of testbeds

Experience in proving support from GRID'5000 HPC

Added value from F4F+

- connecting to testbed Patrons
- funding the project hours
- enabling new product that SkyEcho needed to make anyway

Feedback to Fed4FIRE+

Used resources:

Testbed: GRID'5000

Interactive nodes with dedicated GPU

Location: Nancy, France

Feedback to Fed4FIRE+

Used resources:

Storage increase over 200Gb on home directory

Python environment for model training

- Connection:

via SSH over CLI

Account on HPC was generated by GRID'5000 admin

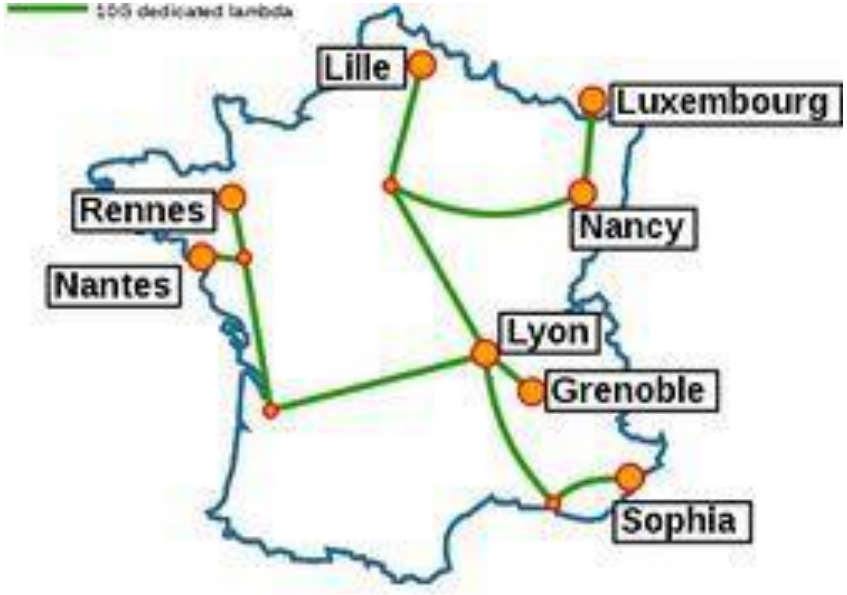
SkyEcho was included in email group of GRID'5000

Feedback to Fed4FIRE+

Added value testbed:

- **Flexible** and decentralized network of nodes
- **Interactive** and batch nodes available
- **Responsive** support

-> issue with python environment in batch jobs resolved with testbed
-> Deploy interactive graphical way to upload data for future



Feedback to Fed4FIRE+

Added value Fed$FIRE+:

Very quick evaluation of the submitted call

Risk for company is reduced by carrying out experiments with Fed4FIRE+

- **Networking:** connecting SME's and academia
- **Possibilities** for follow-up (Stage-2 experiment)
- **Visibility** of the SME via Fed4FIRE+ website / social media

This project has received funding from the European Union's Horizon 2020 research and innovation programme, which is co-funded by the European Commission and the Swiss State Secretariat for Education, Research and Innovation, under grant agreement No 732638.

Thanks for attention! Questions?

WWW.FED4FIRE.EU