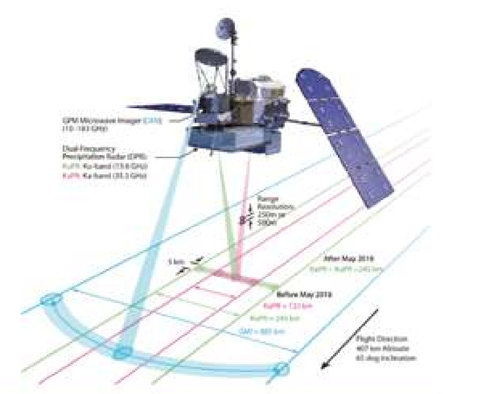
SAMBATH VIBOLROTH

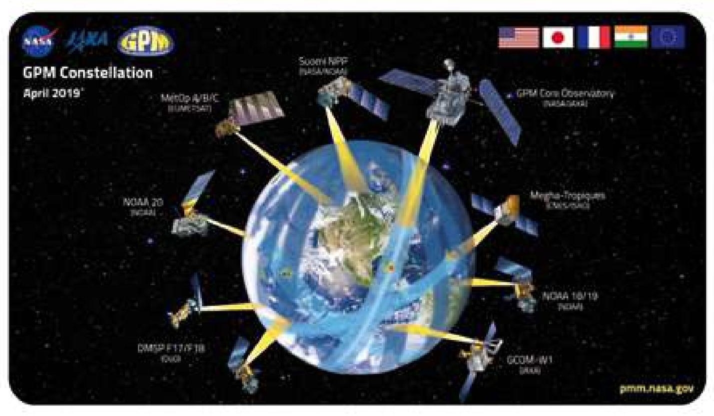
INFERRING THE RAIN RATE FROM REMOTE SENSING SPACE MEASUREMENT USING MACHINE LEARNING ALGORITHM

Encadrement : Nicolas Viltard, Cécile Mallet

Introduction

- Rain retrieval from space-borne passive microwave is a very challenging task:
  - Spatial and temporal intermittence of rain
  - Low resolution of instruments (5 – 10 Km)
  - Non-linear relationship between the brightness temperature and the surface rain



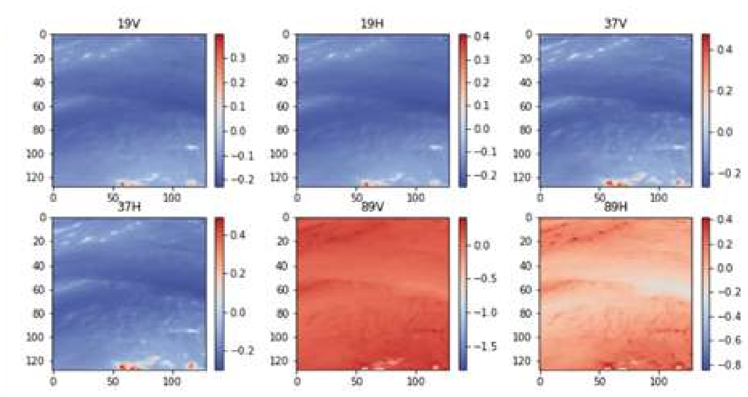


- Global Precipitation Measurement (GPM): a constellation of satellites for the purpose of providing global precipitation product. It offers a Precipitation Radar (PR and DPR) and radiometer (TMI and GMI) on the same platform.
  → Deep learning model taking as inputs the brightness temperature and producing the surface rain rate. The model trained could also provide rain rate outside the DPR band.
  → Application on the constellation of satellites in the GPM

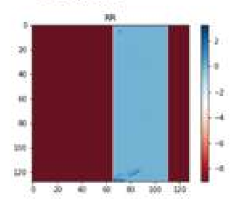
Database

- Roughly 44 000 images taken over two years consisting of:
  - Inputs: 19 GHz, 37 GHz, 89 GHz (Horizontal and vertical) brightness temperature from GMI. The brightness temperatures are normalised by their mean and standard deviation. They are divided by 3 times the standard deviation to get the majority of data between -1 and 1.
  - Outputs: surface rain product from DPR which is collocated with GMI pixel. The images chosen for the datasets have to contain either more than 100 pixels with at least 1 mm/h of rain or more than 10 pixels with at least 10 mm/h of rain.
- The images have 128x128 pixels. They are cropped and rotated randomly from the original images.

Inputs

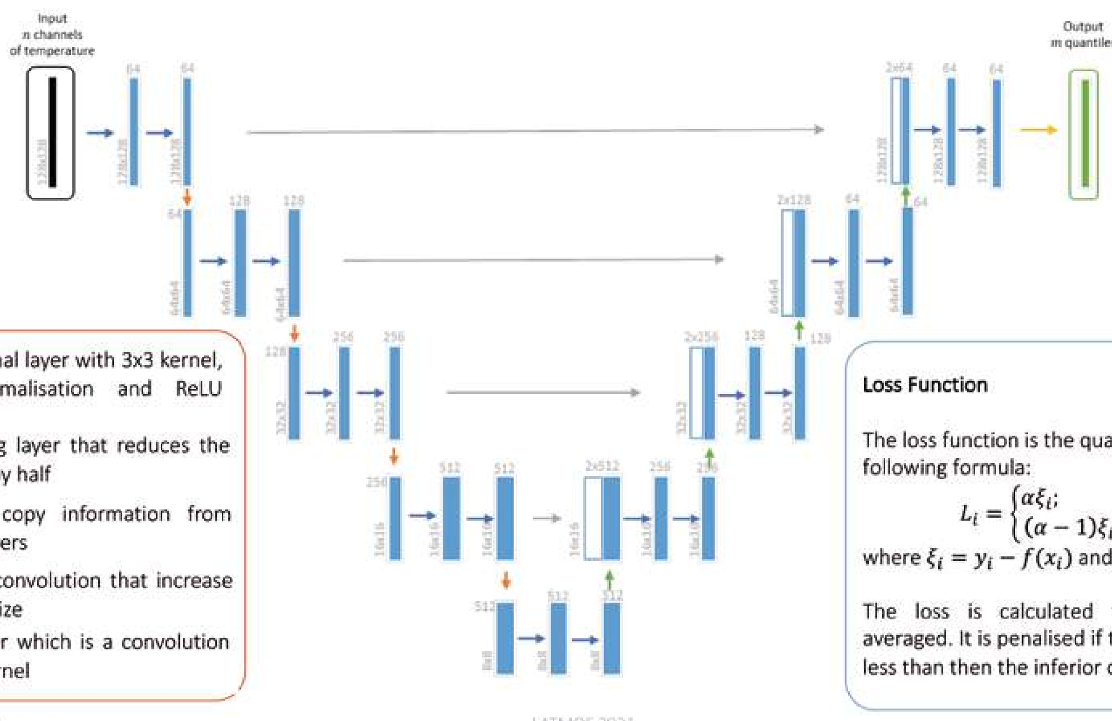


Outputs



The rain band of the DPR is smaller than the brightness temperature images. The pixel where there is no data is replaced by a negative value (-9)

Model Architecture: UNet



- Convolutional layer with 3x3 kernel, batch normalisation and ReLU activation
- Max pooling layer that reduces the image size by half
- Crop and copy information from previous layers
- Transpose convolution that increase the image size
- Output layer which is a convolution with 1x1 kernel

Loss Function

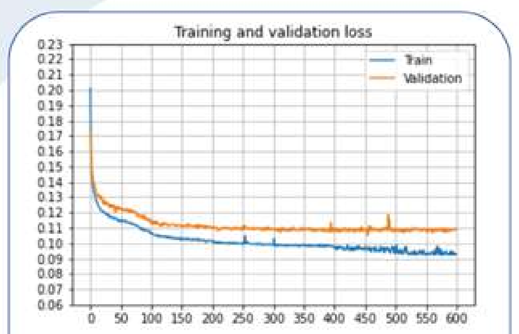
The loss function is the quantile loss with the following formula:

$$L_i = \begin{cases} \alpha\xi_i; & if\ \xi_i \geq 0 \\ (\alpha - 1)\xi_i; & if\ \xi_i < 0 \end{cases}$$

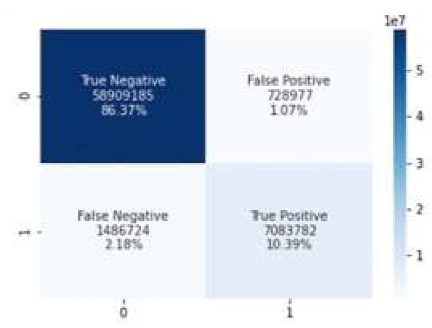
where $\xi_i = y_i - f(x_i)$ and α is the quantile.

The loss is calculated for each pixel then averaged. It is penalised if the superior quantile is less than then the inferior quantile.

DRAIN model preliminary results

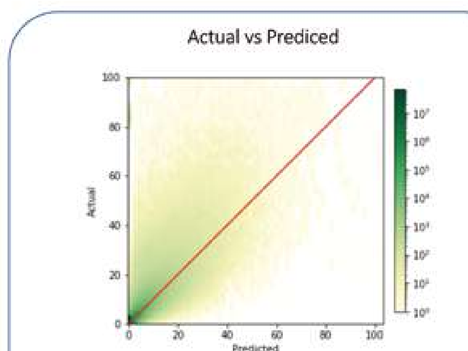


- The model initialised weights are drawn from a uniform distribution. The optimization used is from the Pytorch library (Adam).
- Retro-propagation: since the DPR rain rates have smaller dimension than the GMI images, the loss is not calculated on the unavailable grid point (-9).
- The validation loss does not decrease after 200 epochs.



Rain vs No Rain identification

- False Positive: rain incorrectly predicted when there is no rain is at most 1 mm/h in most of the cases (99.7%)
- False Negative: undetected rain corresponds to DPR rain lower than 1 mm/h in most of the cases (98.93%)



Error on true positive cases

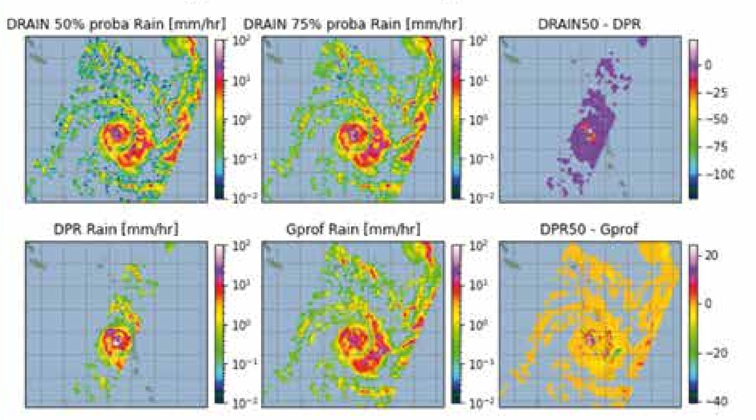
	RMSE	MAE
True Positive	1.66	0.47

Score on rain/no rain classification

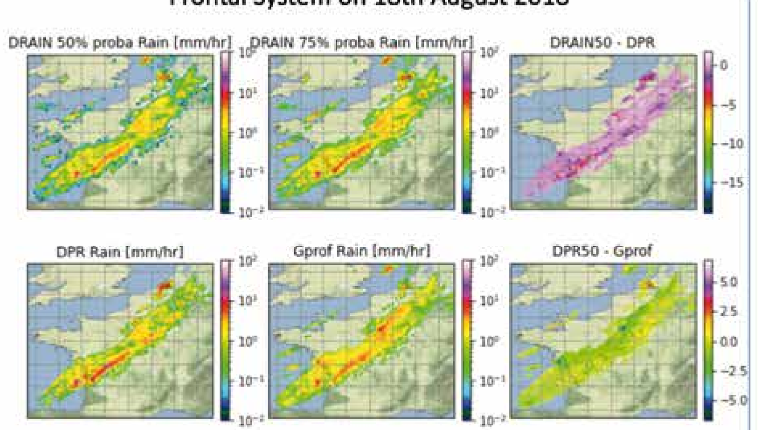
Accuary	F1 Score
0.96	0.86

DRAIN model preliminary results

Typhoon Harold on 6th April 2020



Frontal System on 18th August 2018



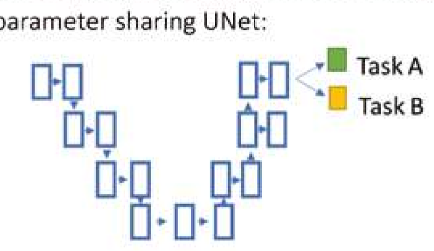
- DRAIN 50% → Median; DRAIN 75% → 3rd Quartile
- DPR → radar rain rate; Gprof → NASA Goddard reference algorithm in the community in terms of rain retrieval using auxiliary data to constrain the solution (temperature, surface type, humidity, cloud cover, etc)

→ In both cases, we observed good rain structure and intensities compared to radar rain rate (DPR) and reference algorithm (Gprof).

Perspectives

Multi-task Learning

- In Multi-task learning, the model has to predict more than one task. Learning two tasks at the same time allows the model to generalise better.
- In Multi-task learning, there is two methods: hard parameter sharing and soft parameter sharing.
- Hard parameter sharing UNet:



- Task A → Rain Rate (Quantile regression);
- Task B → Convective/stratiform flag (Classification).

Transfer Learning

- By using Transfer Learning, we seek to transfer the knowledge from the model trained using the DPR surface rain.
- In GPM, there is a constellation of satellites with slightly different configuration of radiometers. The knowledge transferred could allow us to retrieve rain rate from those satellites.
- We could use other channels of brightness temperatures to help the models be familiar with different configurations of radiometers. It is then necessary to compare the performance of the models with different inputs.
- The Multi-task learning could also help create a model that generalises better.