Introducing Machine Learning in the Ground RFI Detection System (GRDS)

R. Onrubia[(1)], R. Oliva[(1)]

(1): Zenithal Blue Technologies, Barcelona, Spain

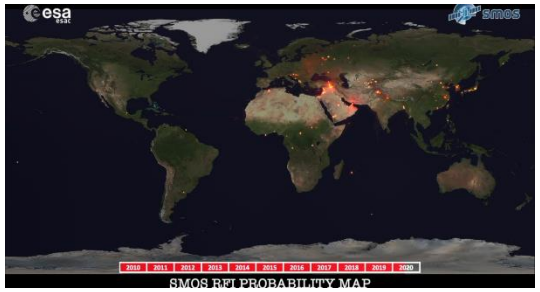
Outline

- Introduction
- GRDS System Description
- Results from validation campaign
- Adding machine learning to GRDS
- Conclusions

Introduction

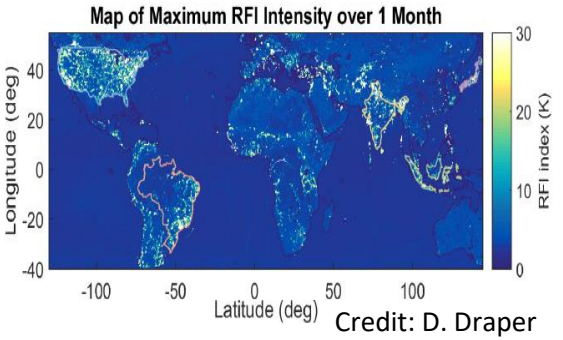
- Microwave passive remote sensing has been experiencing more and more instances of RFI
- RFI can vary widely in terms of intensity, extension, geographical regions, polarization, bandwidth, active duty cycle, etc.
- No single RFI algorithm is capable of detecting all RFI
- Best strategy is to use a combination of techniques and use the most of external information available

1.4 GHz



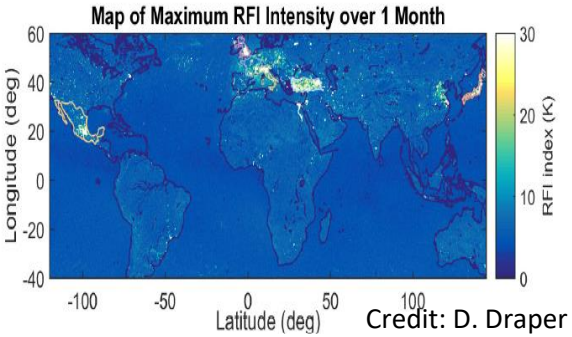
Credit: SMOS RFI team. ESAC

6.9 GHz



Credit: D. Draper

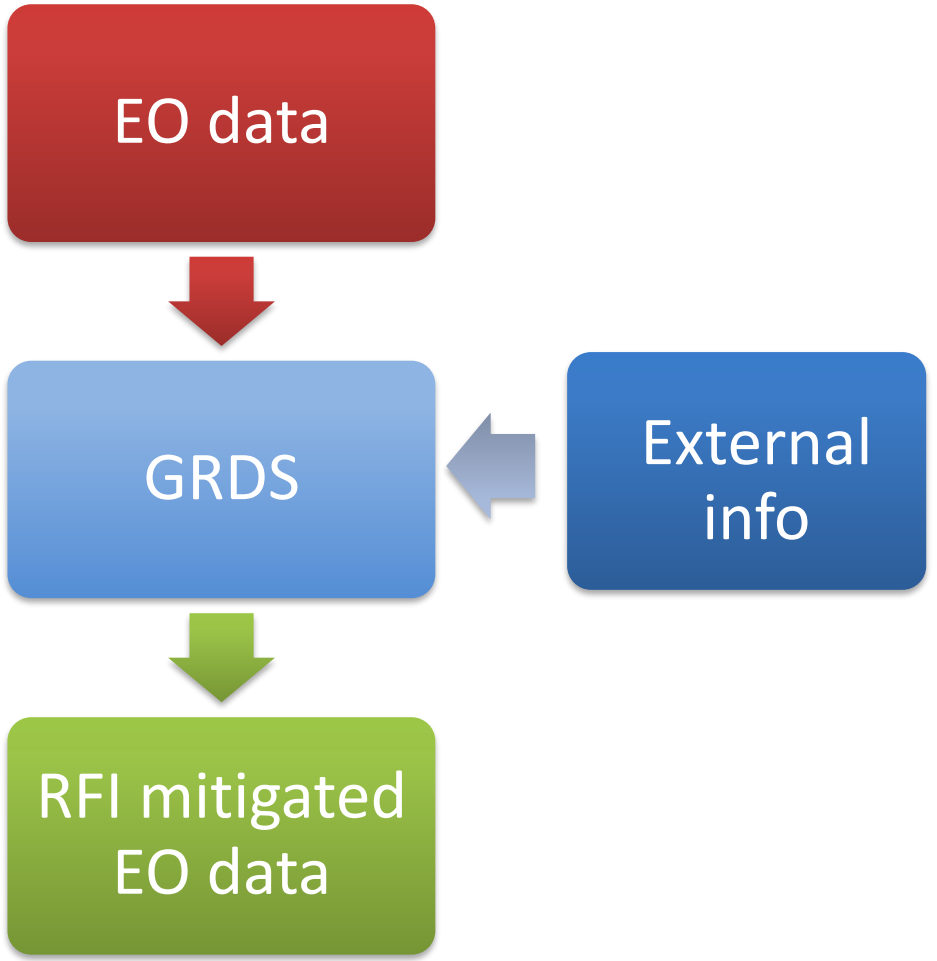
10.65 GHz



Credit: D. Draper

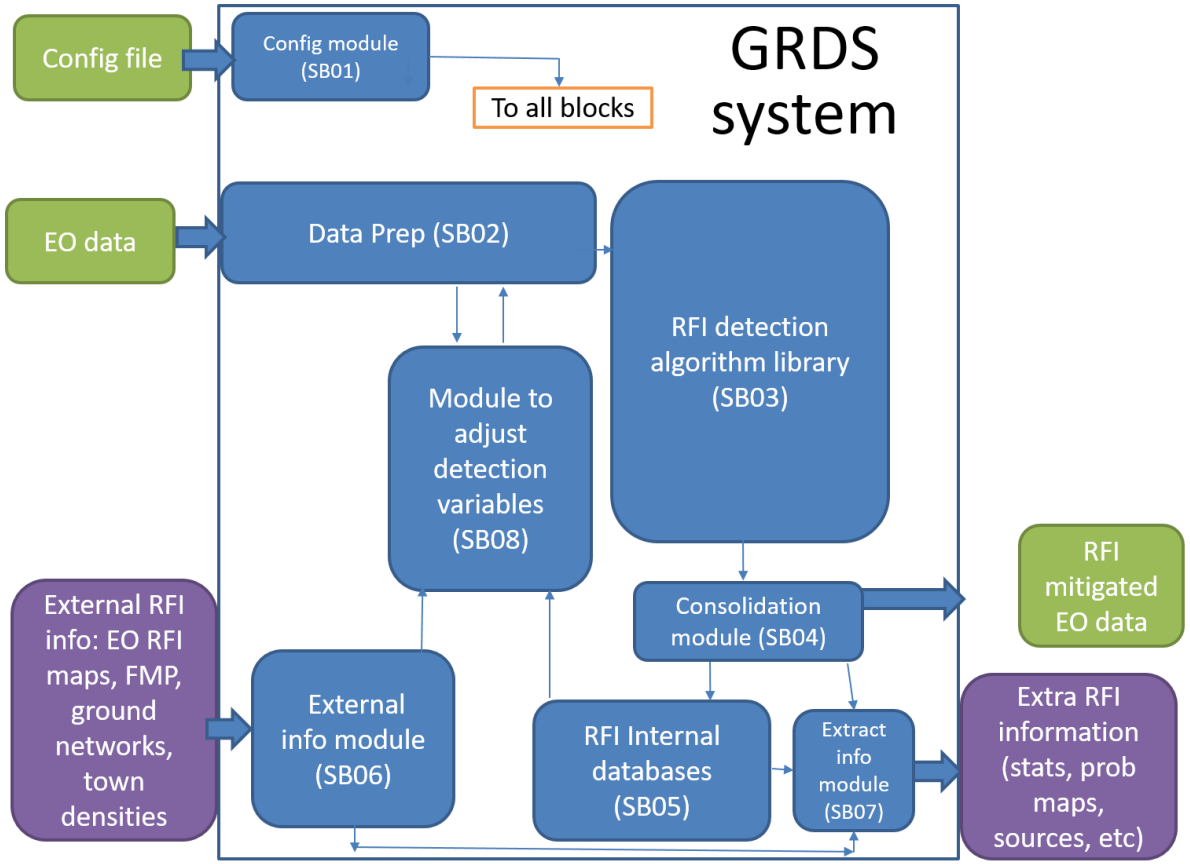
GRDS system description

- GRDS is a system developed by ZBT whose purpose is to detect low level RFI on EO data.
- The system has been developed with a flexibility in the design to be able to ingest data from any EO microwave missions (currently SMOS and AMSR2).
- It makes use of external and internal information to detect presence of RFI.
- It allows to configure the level of flagging to set the user preference between missing RFI and false alarm detection.



GRDS system description

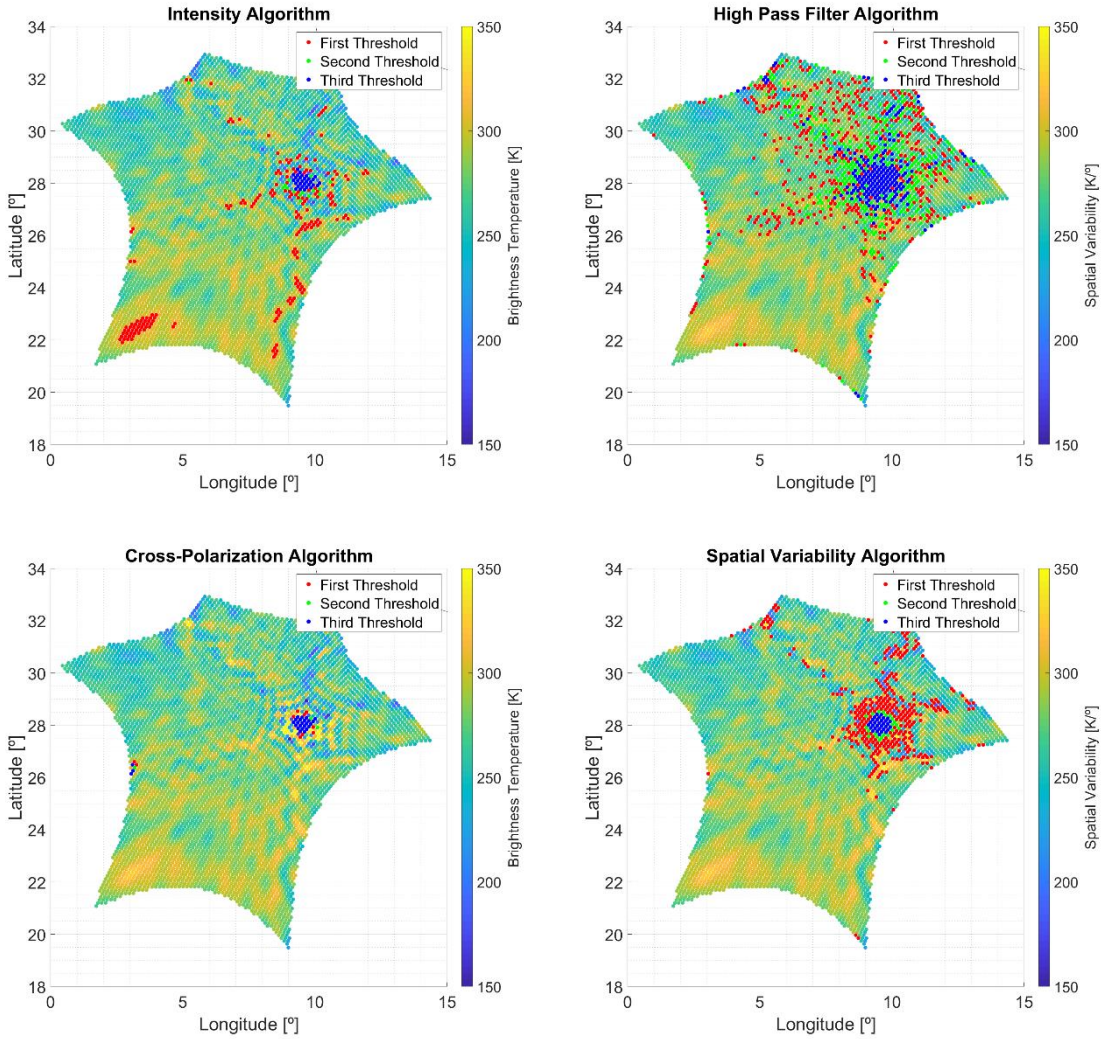
- Data ingested from multiple EO sensors and converted to a common data format



GRDS system description

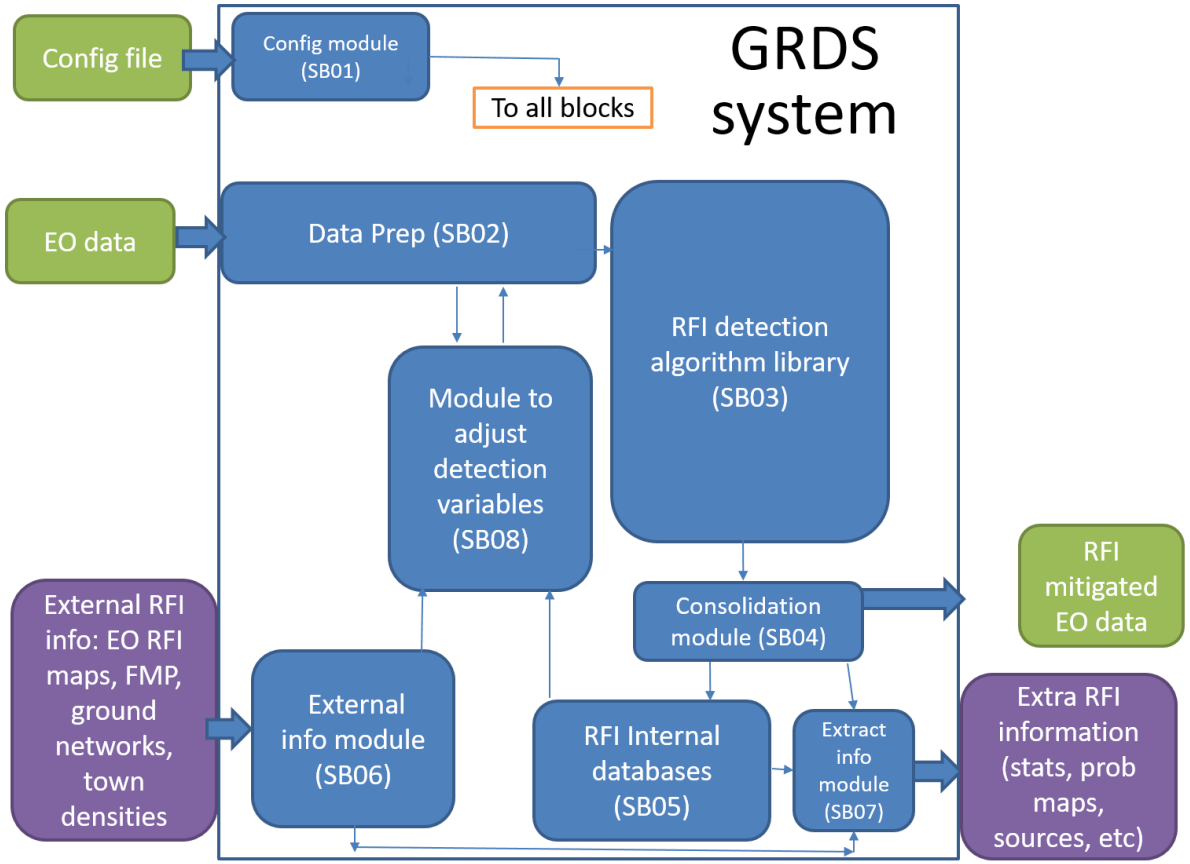
- GRDS scans all observations using a library of RFI detection algorithms, including:
  - Intensity (HH, VV, ST3, ST4)
  - Outlier
  - Cross-polarization
  - Cross-frequency channels (RFI Index)
  - Spatial variability
  - Image Enhancement using High Pass filter
  - Kurtosis
  - Skewness

- Up to three different threshold levels
  - Statistically determined using products over mostly clean regions and removing SMOS RFI flagged data



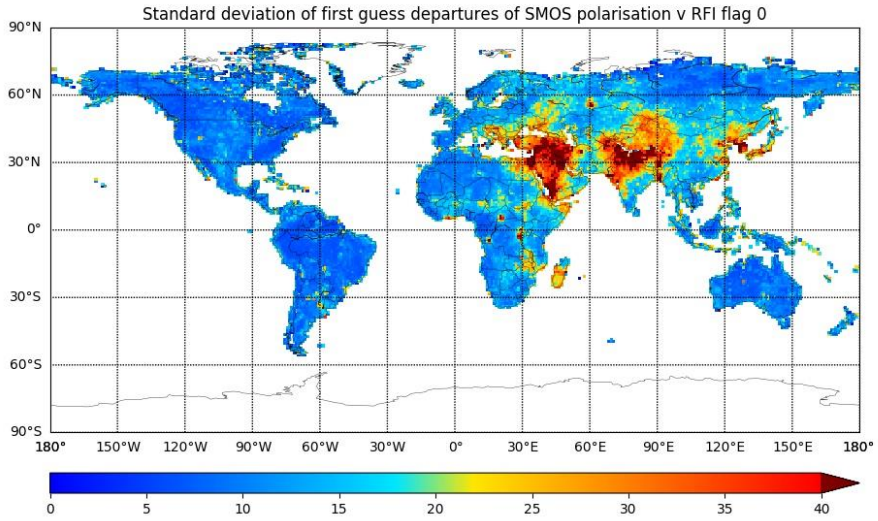
GRDS system description

- Flags are combined and added to the original EO file
- RFI instances are added to an internal RFI databases
- RFI Detection thresholds are adjusted based on previous detections and external information, such as:
  - IEEE GRSS RFI Database
  - Population density
  - Airports and Air navigation aids
- GRDS on purpose does not use NWP models in order to be NWP-independent.

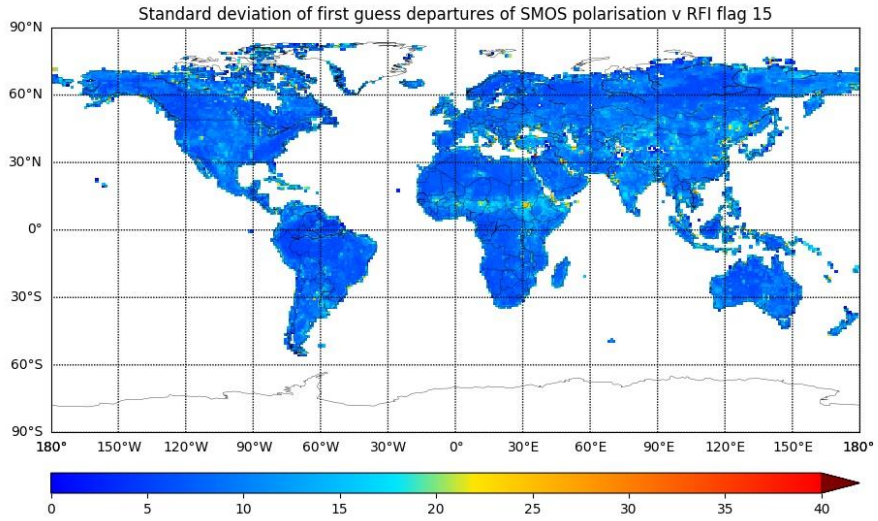


Results

- Validation conducted independently by ECWMF using std(First Guess Departures)
  - Brightness Temperature collocated to 40x40 km grid
  - Standard deviation of error w.r.t. expected values from numerical models
    - Use of physical temperature, atmospheric pressure, precipitation, etc.



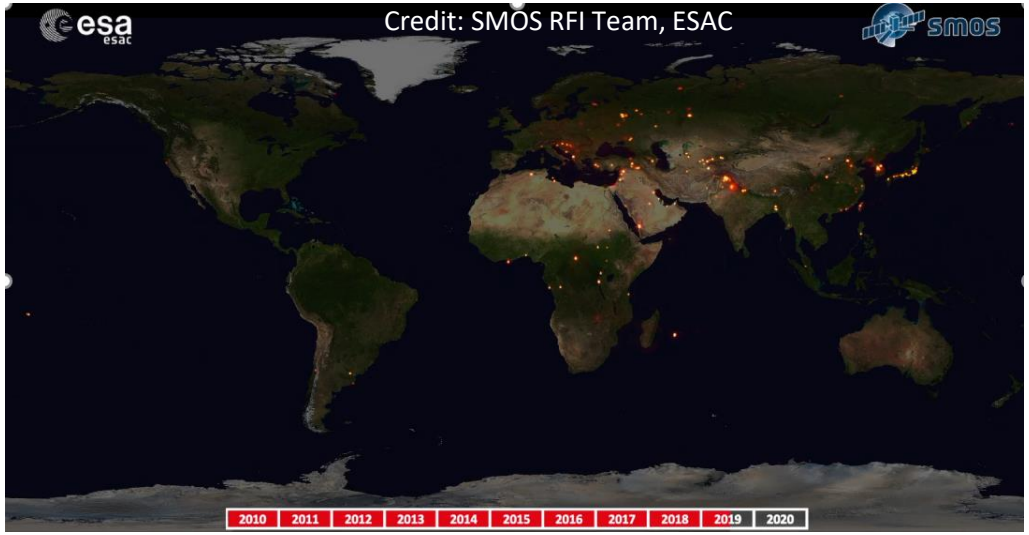
Original SMOS data

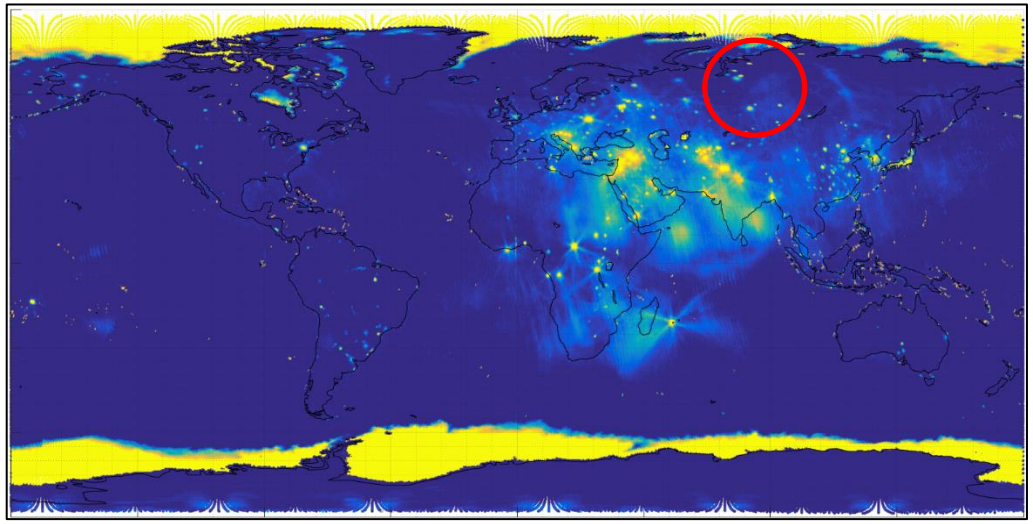


After GRDS

Results

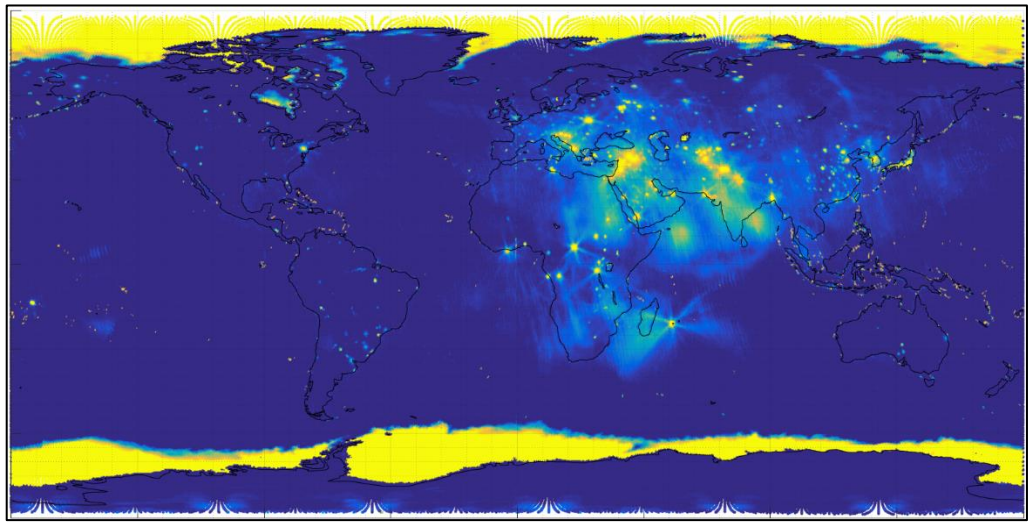
- SMOS RFI sources map vs. GRDS RFI Probability map
  - SMOS RFI sources match the areas with highest RFI probability
  - New sources found by GRDS
  - False positives near the poles caused by terrain misclassification on sea ice





Introducing ANN in GRDS

- Objective: improving the GRDS performance
  - Correcting terrain misclassification on sea ice
    - No sea ice coverage information is included in SMOS Level 1 products
    - Wrong threshold levels used for RFI detection techniques
  - RFI detection
    - No reliable ground truth

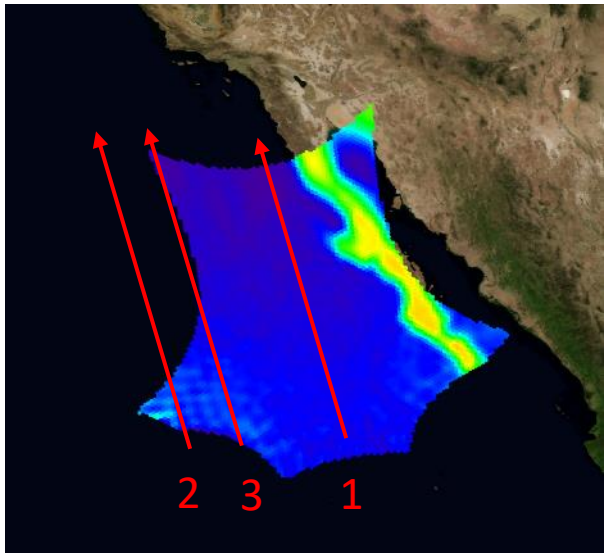


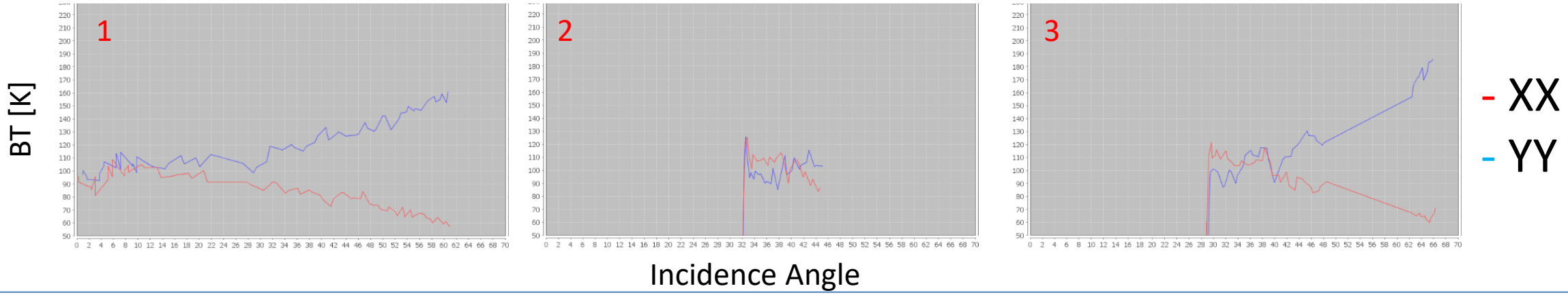
Sea Ice detection with ANN

- Objective: Prediction of the sea ice coverage
- Input data:
  - Brightness temperature
  - Incidence angle
- Ground truth:
  - Sea ice coverage daily products from AMSR2 (89 GHz)
- PCA applied for feature reduction to speed up the ANN training

SMOS data issues with ANN

- Irregular data
  - Different number of data points per ground pixel and polarization.
  - Different incidence angles per observation.
  - Solutions:
    - Collocate data in “incidence angle bins” and average it.
    - Fill the gaps with the average brightness temperature value.

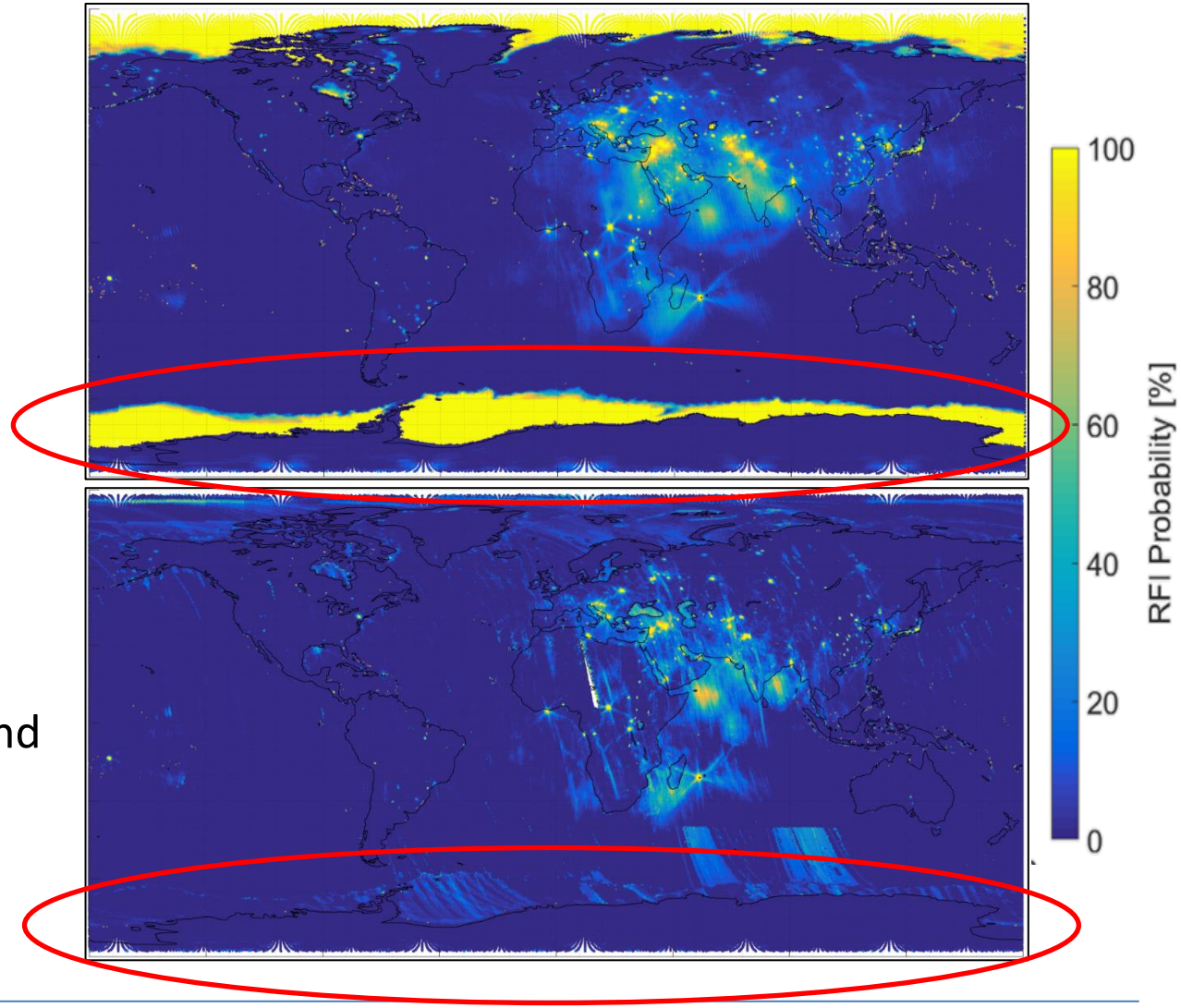




Sea Ice detection preliminary results

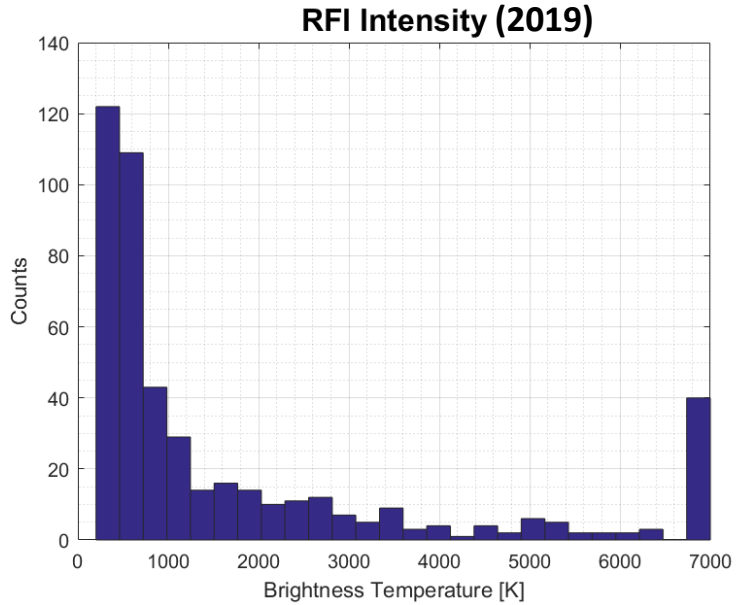
- Training
  - 3 days of data
  - Only austral data used for training
  - January (austral sea ice melting peak)
  - Results (optimum threshold)
    - Precision : 92.58%, Recall : 97.50 %
    - F1 Score : 94.98%

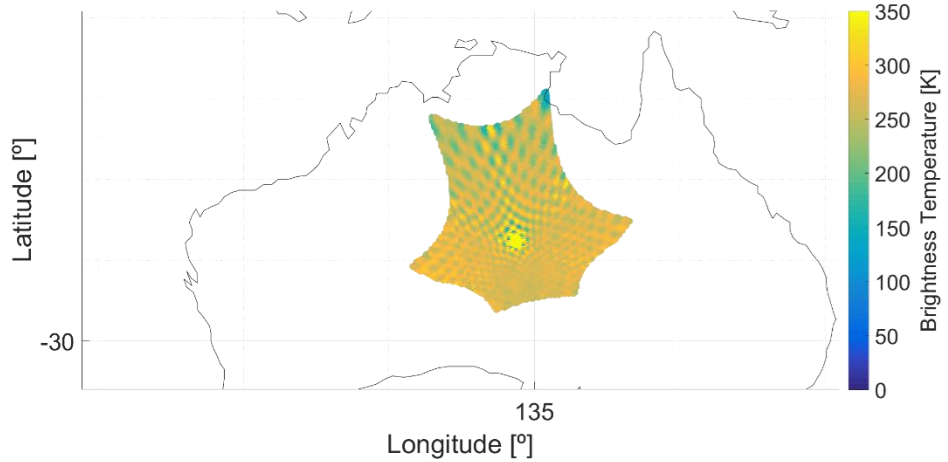
- Sea ice
  - Only 3 days of SMOS data analyzed
  - Large sea ice areas detected properly
  - Still some false negatives over sea ice and some false positives over sea
    - Some RFI related



RFI simulation for ANN training

- ANN training
  - No reliable RFI ground truth to train the ANN
    - RFI sources have been located
      - "Tails" are not tagged
      - Only strong RFI sources
    - No "clean" areas
  - RFI need to be simulated
    - Threshold determination for optimum RFI tagging under study based on radiometric resolution





Conclusions

- RFI are very diverse in nature and no single RFI detection technique is capable of properly detecting them all
- GRDS tool is a software that screens Earth Observation data for presence of RFI contamination.
- GRDS has been very successful in achieving its objectives :
  - Reducing RFI contamination to very low levels, with first guess departure variations compatible with regions with no RFI
  - At a low false alarm rate (<4% with tightest thresholds).

- ANN currently being introduced for
  - Terrain misclassification on sea ice
    - Already promising results
  - RFI detection

THANKS FOR YOUR ATTENTION

Questions?

We very much welcome feedback from the audience:

onrubia@zenithalblue.com