Deep Learning techniques for the generation of satellite-borne thermal Infra-Red Spectra and their applications to remote sensing retrievals.

Elisa Castelli[1,*], Enzo Papandrea[1], Alessio Di Roma[1], Ilaria Bloise[2], Mattia Varile[2], Hamid Tabani[3], Lorenzo Feruglio[2]
[1] *Istituto di Scienze dell'Atmosfera e del Clima (ISAC-CNR),Via Piero Gobetti, 101, 40129, Bologna,*
[2] *AIKO S.R.L., via dei Mille, 22, 10123, Turin, Italy,*
3 BSC, Jordi Girona 29, Barcelona, Spain

** Corresponding author. Tel: +39 051 6398040, E-mail: e.castelli@isac.cnr.it*

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Abstract: Generative Artificial Intelligence is a branch of Machine Learning algorithms. Generative Models such as Generative Adversarial Networks (GANs) and Autoencoders (AE) can be used to generate large amount of data at reduced computing costs with respect to traditional Forward Models (FM) and to test new configurations of variables.

In the frame of the ESA **DeepLIM project**, GANs and AE are used to simulate Thermal Infra-Red (TIR) spectra acquired by satellite-borne instruments. The spectra are generated starting from a dataset of spectra simulated with a Radiative Transfer Model (RTM) or measured during an aircraft campaign. The simulated dataset consists on both Top Of the Atmosphere (TOA) and Bottom Of the Atmosphere (BOA) radiances. The generated dataset is composed by the same fields. To simulate the satellite measurements, both original and newly generated radiances are convolved with spectral response functions in the 5 TIR channel of the Sentinel Land Surface Temperature Monitoring (LSTM) mission (at 8.6, 8.9, 9.2, 10.9, 12.0 µm) and converted into Brightness Temperatures (BTs). Then, the extended dataset is processed to retrieve Land Surface Temperature (LST). The retrieval is performed using both "state-of-the-art" algorithms and Neural Network (NN).
Here we report the results of this study.

Deep Learning for generation of Far-Infrared Spectra: input datasets

High spectra resolution radiances are the starting point for the development of satellite remote sensing retrieval chain.
In the frame of DeepLIM project, we aim to augment the dataset of available TIR spectra through the use of Deep Learning algorithms.
Two TIR input datasets were used as input for the generation of new signatures:

DART dataset: This dataset consist on spectra generated by the DART model (https://dart.omp.eu). Two agricultural scenario were used in this work: Maize and Wheat fields.

TASI dataset: The TASI dataset is made of aircraft measurements acquired during the SurfSense 2018 Grosseto campaign on July 2018 in support of LSTM studies.

TASI measurements were acquired at flight altitude (3 km). For this reason the dataset cannot be used as it is to infer TOA radiances. We used a RTM FM and Land Surface Temperature (LST) and Land Surface Emissivity (LSE) retrieved from TASI to simulate satellite TOA, BOA and downward irradiance at ground (Fig. 2). RTM simulations were performed for some selected scenario of the 18th July 2018 Grosseto flight (Fig.1)

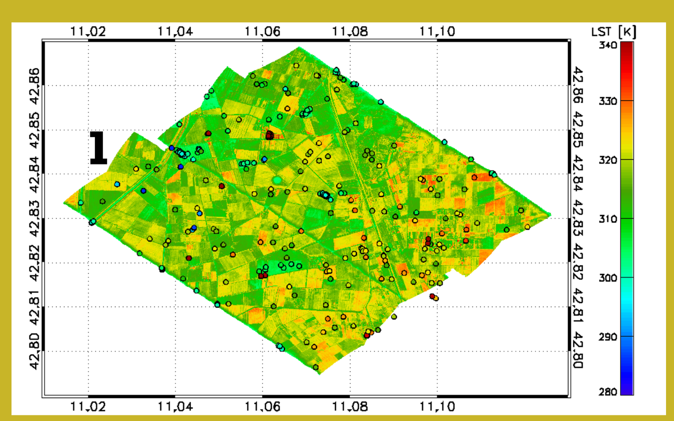


Figure 1: Selected points for the 18th July 2018

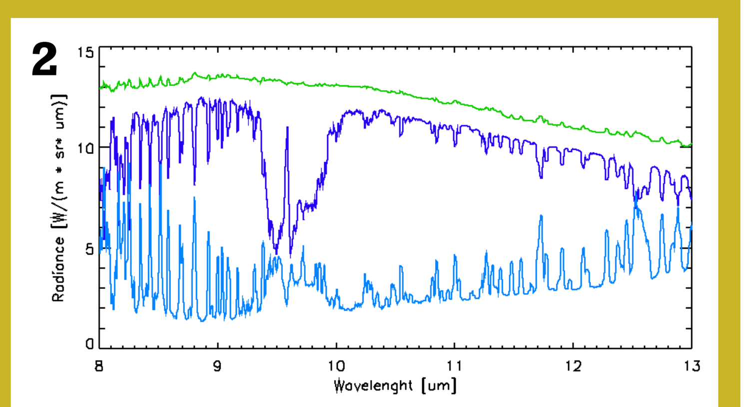


Figure 2: Simulated TOA, BOA, Irradiance

Deep Learning for generation of Far-Infrared Spectra: output datasets

The TASI dataset consist on about 1000 spectra while the DART dataset on about 50000. Due to the different size of dataset two different generation algorithms were used (autoencoders for TASI dataset and GANs for DART data).
An example of signature generated from DART TOA spectra (Fig. 3) is reported in Fig.4.

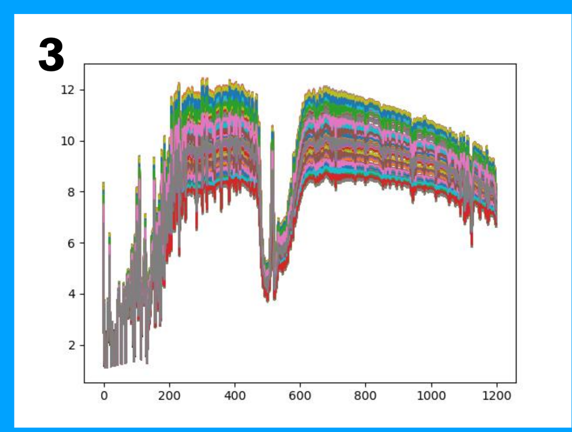


Figure 3: DART original TOA signatures *

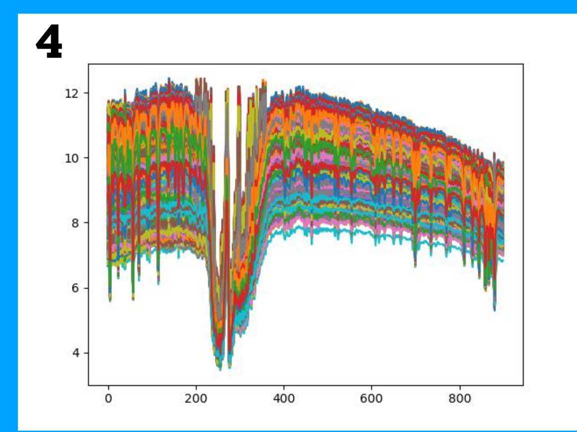
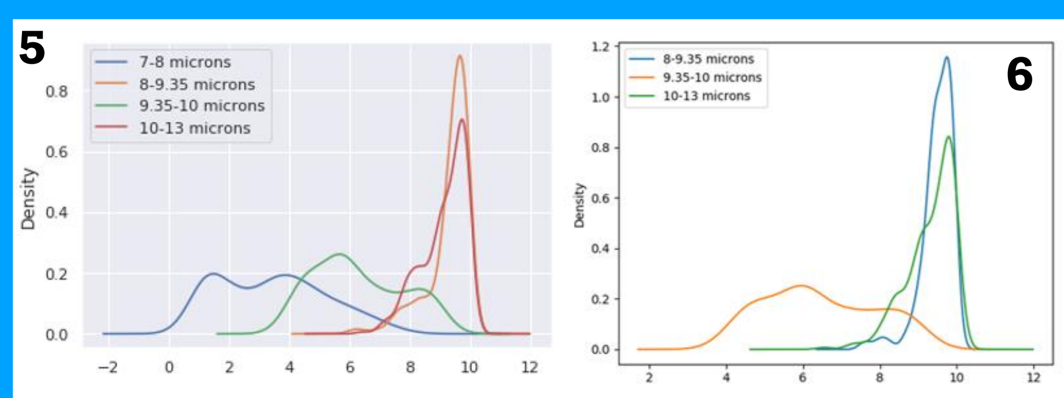


Figure 4: DART generated TOA signatures *

* as a function of point number (not wavelength as in Fig.2)



Figures 5 and 6: distributions of original (Fig.5) and generated (Fig. 6) TOA signatures.

The generated dataset was compared to the original one to evaluate if the generation procedure preserves the characteristics of original spectra.
An example of the distributions of TOA radiances for original and generated data in different spectral regions is shown in Fig.s 5 and 6.

Applications to remote sensing surface properties retrievals:

state of the art vs Deep Learning retrieval algorithms

The first step is the convolution of original and generated signatures with LSTM channels Spectral Response Functions (SRF), in green and blue in Fig.7. Then the predicted noise level is added.

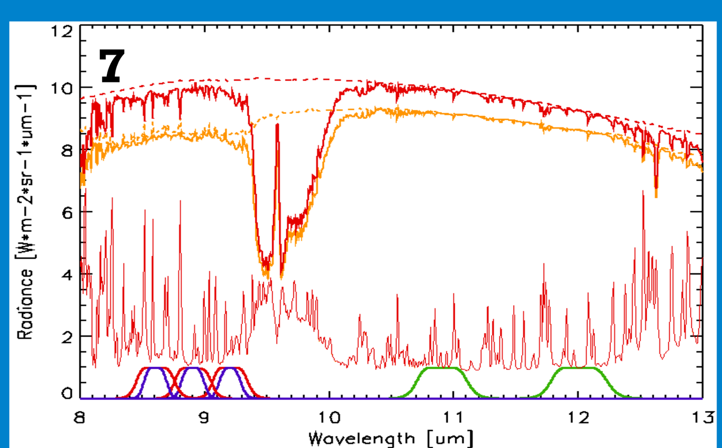


Figure 7: BOA, TOA and irradiance from DART and positions of LSTM SRFs.

Once the LSTM Brightness Temperature (BTs) original and generated datasets in the 5 channels are obtained, we can use them for the retrieval of LSE and LST.

LSE and LST retrieval was performed with "state-of-the-art" code and one code based on Deep Learning techniques.
The "sate-of-the-art" code is a TES (Temperature Emissivity Separation) algorithm (Fig. 8), while the Deep Learning based one is a Neural Network (NN) code.

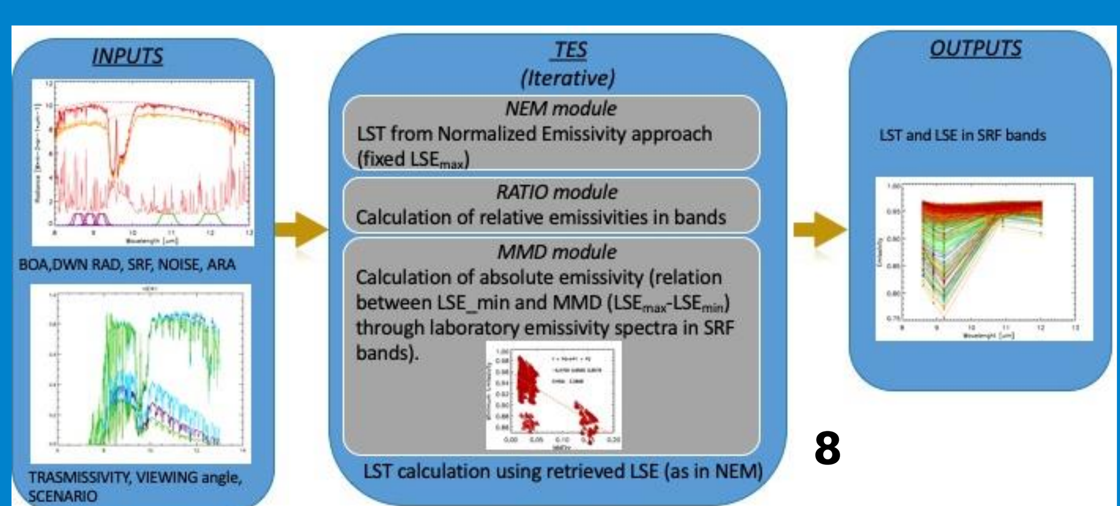


Figure 8: TES algorithm retrieval scheme

Retrieval Results

First of all the NN has been trained over a part of the training dataset (70%) and then tested on the remaining one (30%). An example of the output of this procedure is given in Fig.9. (For this process the LST values have been normalized to fit the [-1,1] scale, allowing the NN to work with similar numbers for LSE and LST.)

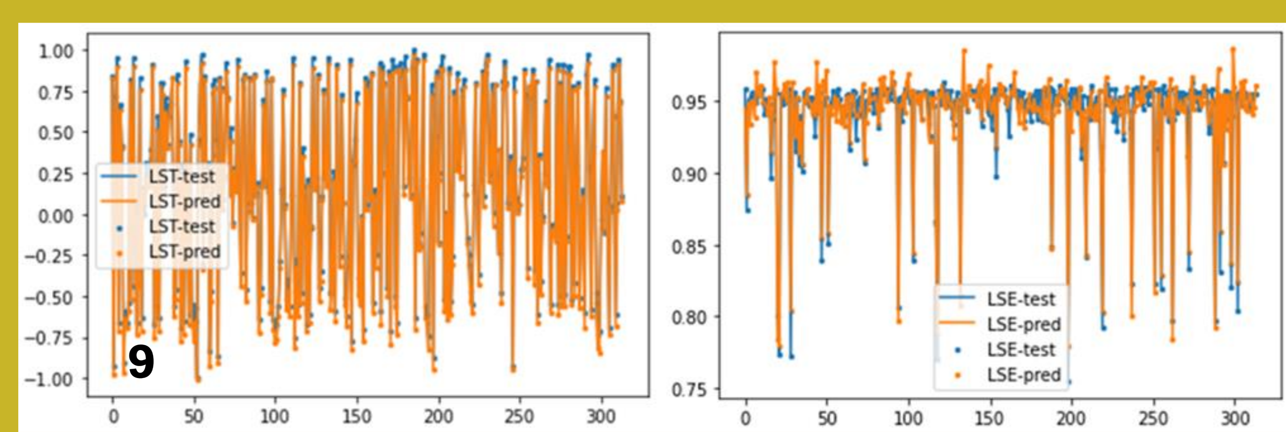


Figure 9: LST and LSE NN retrievals vs true values for the test dataset

Then the NN has been applied to a different part of the dataset not used in the previous exercise. The results have been compared to the TES one (Fig. 10 and Fig. 11). A general good agreement is obtained, the generated dataset produce realistic LST and LSE retrievals.

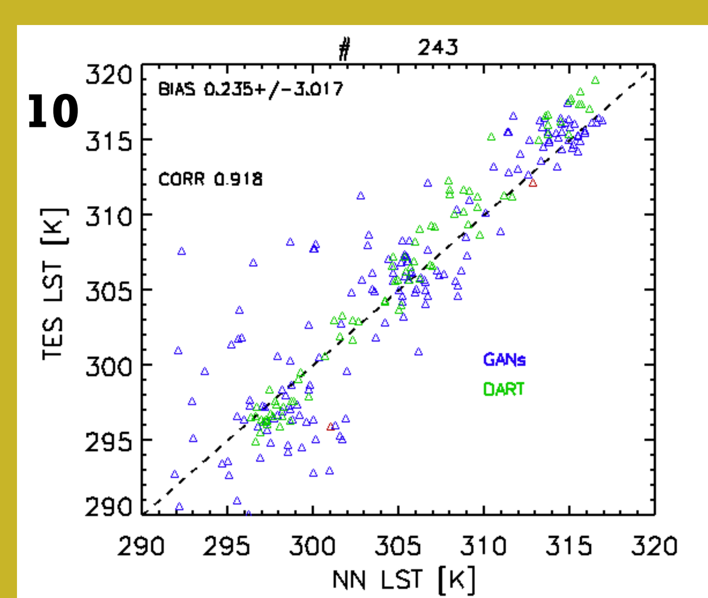


Figure 10: TES vs NN results for LST retrieval

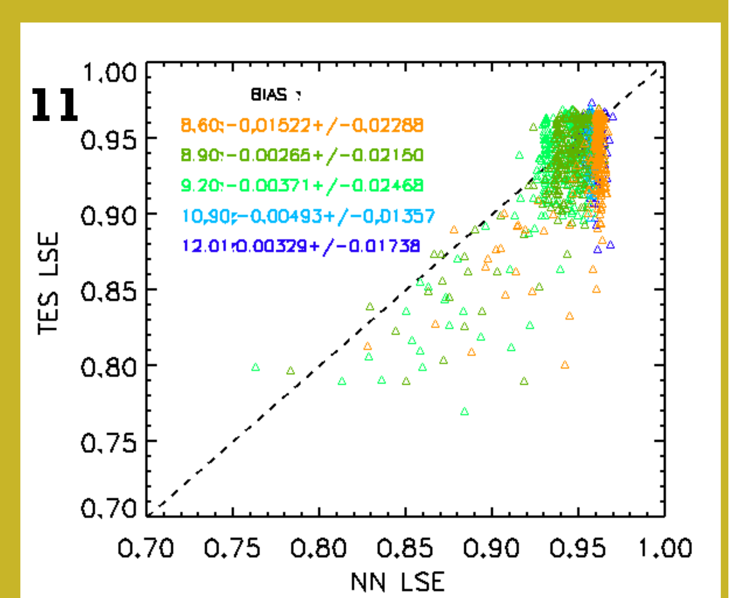


Figure 11: TES vs NN results for LSE retrieval

Conclusions

The ESA **DeepLIM project** main results are:
1) Demonstrate the feasibility of using Deep Learning techniques to emulate FMs output at low costs
2) Demonstrate the feasibility of using Deep Learning techniques to support the development of satellite retrieval chains

Further details will be found in: Castelli et al., 2021: "Deep Learning application to surface properties retrieval using TIR measurements: a fast forward/reverse scheme to deal with big data analysis from new satellite generations", submitted to Remote Sensing