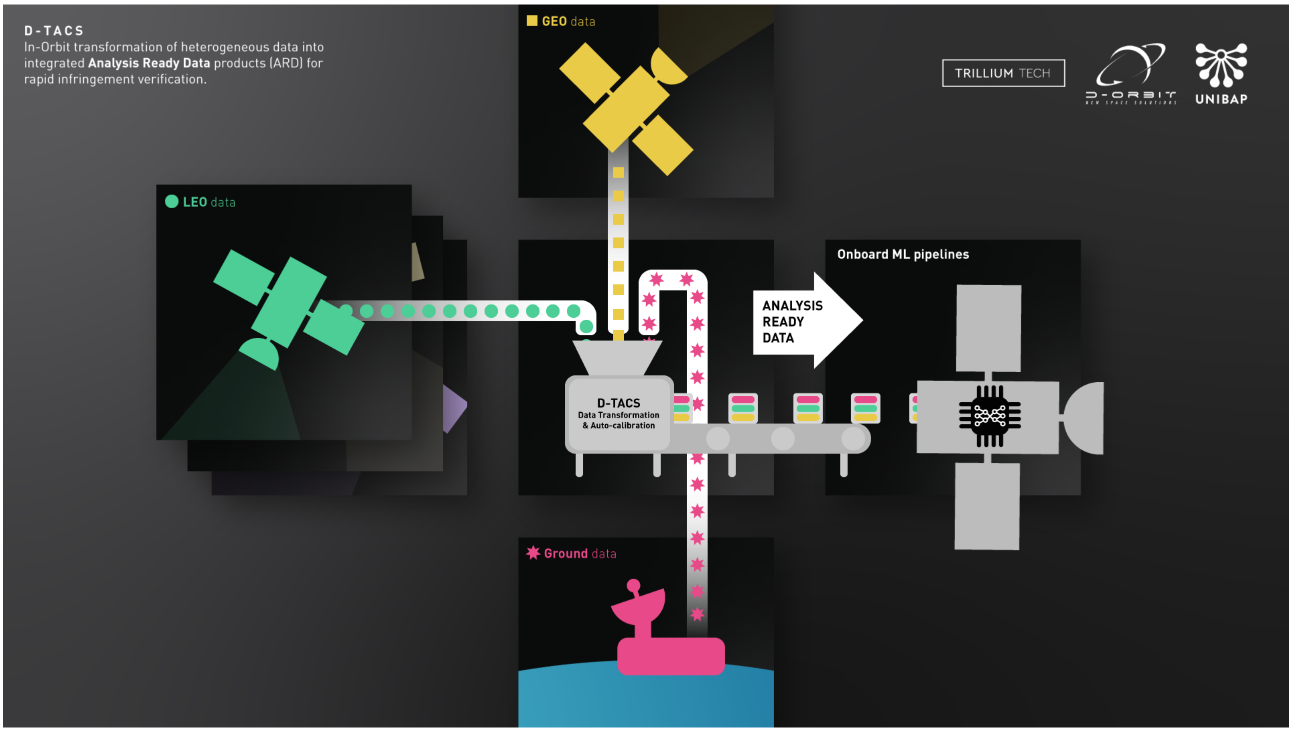
TRILLIUM TECH

Networked Intelligence in Orbit



D-TACS: on demand data transformations and auto-calibration in-orbit

Executive summary
Activity type: Early Technology Development

[ESA Cognitive Cloud Computing in Space Campaign](https://www.esa.int)
Project lead: Gonzalo Mateo-García[1 2]
Researchers: Giacomo Acciarini[1 3], Cesar Aybar[1 2]
Advisors: James Parr[1], Vit Ruzicka[1 5], Luis Gómez Chova[2], Simon Reid[4]
Affiliations: [1] Trillium Technologies, [2] University of Valencia, [3] University of Surrey, [4] D-Orbit, [5] Oxford University

Activity summary:
Many use-cases for insight generated in orbit require the integration and calibration of raw data from multiple instruments and vantage-points –a future vision sometimes referred to as ‘hybrid

Publishing Date: 23/11/2022
Project: D-TACS I-2022-00380
Agreement/PO number: 4000138109
Implemented as ESA Initial Support for Innovation

observation'. In this project, we developed DTACSNet, an on-board scene classification and atmospheric correction processor for calibration of Sentinel-2 images based on lightweight Neural Networks emulators. DTACSNet provides results 7x faster than the operational Sentinel-2 processor under identical CPU settings and 30x faster when a GPU accelerator is provided. Our results show that DTACSNet efficiently replicates Sentinel-2 surface reflectance values with an absolute error below 2%. In addition, the scene classification module obtains the highest F2-score (0.92) compared to Fmask (0.7) and Sen2Cor (0.65) for cloud and cloud shadow masking.

Executive Summary

Processing data onboard Earth observation satellites will enable more reactive, agile and autonomous systems, e.g. by discarding on-real-time cloud contaminated imagery [1] or reactively selecting relevant targets (algae blooms, oil spills, methane leaks, crop damage, fires, etc.). It can also help us to optimize data to be sent back to Earth by reducing the size of the products to be downlinked [2] or even prioritize the tiles to download [3].

Nevertheless, onboard processing requires the integration and calibration of raw data before it is used by most applications. While this is normal practice on the ground, preparation of data to create Analysis Ready Data (ARD) products is not straightforward since the processes running on the ground are less constrained than those running onboard. For instance, onboard hardware has lower processing capability and memory access and the access to ancillary data of onboard processes is very restrictive (if any).

One ubiquitous process to create Analysis Ready Data (ARD) products for optical sensors is atmospheric correction. Atmospheric correction is a sophisticated procedure involving two core steps: a) identifying contaminated pixels where the the signal from the surface cannot be recovered (thick clouds), and b) removing the perturbations introduced by the atmosphere in surface reflectances (conversion from Top-of-Atmosphere (TOA) to surface reflectances (SR)). These perturbations are caused by absorption and scattering of atmospheric constituents (thin clouds, aerosols, water vapor, ozone...) and by occlusions (cloud and terrain shadows). In this project we have developed an atmospheric corrector processor that can be run onboard with tight requirements of memory and processing capabilities. For developing this processor, that we call DTACSNet, we have taken the Sentinel-2 mission as a reference taking advantage of its publicly available TOA and SR data (level 1C and level 2A products respectively).

We compare DTACSNet with the operational Sentinel-2 atmospheric correction processor: the Sen2Cor [4] software which is also available at the ESA web page. Sen2Cor produces accurate surface reflectances when compared with ground truth data according to the recent Atmospheric Correction Intercomparison Exercise (ACIX) [6]. However its cloud detection is significantly worse than other approaches [7,8]. When we look at critical variables for onboard processing (memory consumption and processing time), we have found that the current implementation of Sen2Cor (v2.10) is too demanding to be run onboard.

Publishing Date: 23/11/2022
Project: D-TACS I-2022-00380
Agreement/PO number: 4000138109
Implemented as ESA Initial Support for Innovation

DTACSNet is a Convolutional Neural Network (CNN) model for semantically segmenting Sentinel-2 Level-1C images into cloud semantic classes and to atmospherically correct TOA pixel values. DTACSNet is trained in the recently published CloudSEN12 dataset [8]. DTACSNet is intended to work on remote embedded systems with limited hardware resources. We validated DTACSNet in a large dataset of independent locations of CloudSEN12 (not used for training the models) and over a year-long time series over the same locations used in the ACIX [6] work. We show that our scene classification models are significantly more accurate than Sen2Cor and that the atmospheric correction models have errors in the same order of magnitude as those of Sen2Cor in the ACIX exercise (1.8%). Additionally, we show that the atmospheric correction models work also with less multispectral bands mimicking other existing and prospective satellite configurations. Finally we have compared DTACSNet and Sen2Cor on different hardware configurations and found improvements between x4-11 in running time.

Project achievements

1. We developed **a scene classification model** with manually labeled ground truth data that **is more accurate than Sen2Cor** (0.92 F2 score vs 0.65).
2. We developed an **atmospheric correction model that emulates Sen2Cor atmospheric correction.** The model **error with respect to Sen2Cor is below 2%**. **Inference speed is 4 to 11 times faster.** Memory consumption can be tuned which allows it to run in constrained devices.
3. We found that if the **emulator** is **trained with less bands mimicking other existing and prospective sensors the atmospheric correction error barely changes**. This is a very promising result that shows that we could do **accurate atmospheric correction on sensors with few bands without requiring ancillary data**.
4. We implemented the models in operational onboard hardware provided by Unibap. We benchmark our models in terms of inference time and memory consumption against Sen2Cor.

Results

In this section we show the results supporting these claims.

Scene classification results

Figure 1 shows the results of the scene classification network. The scene classification network classifies each pixel of the image in 4 classes: thick cloud, thin cloud, cloud shadow and clear. The figure compares the F2 score measured in the CloudSEN12 dataset (y-axis) against the approximate inference speed in a full Sentinel-2 scene (10,980x10,980 pixels). We see that DTACSNet is more accurate than Sen2Cor. These results agree with [7,8] and it is mainly because the DTACSNet network is trained over a very large and representative dataset (CloudSEN12). Figure 2 shows some representative outputs from chips of the CloudSEN12 test dataset. We see

Publishing Date: 23/11/2022
Project: D-TACS I-2022-00380
Agreement/PO number: 4000138109
Implemented as ESA Initial Support for Innovation

that DTACSNet (third row) captures much better thin clouds and cloud shadows and it is less confused by the presence of snow (see e.g. third column).

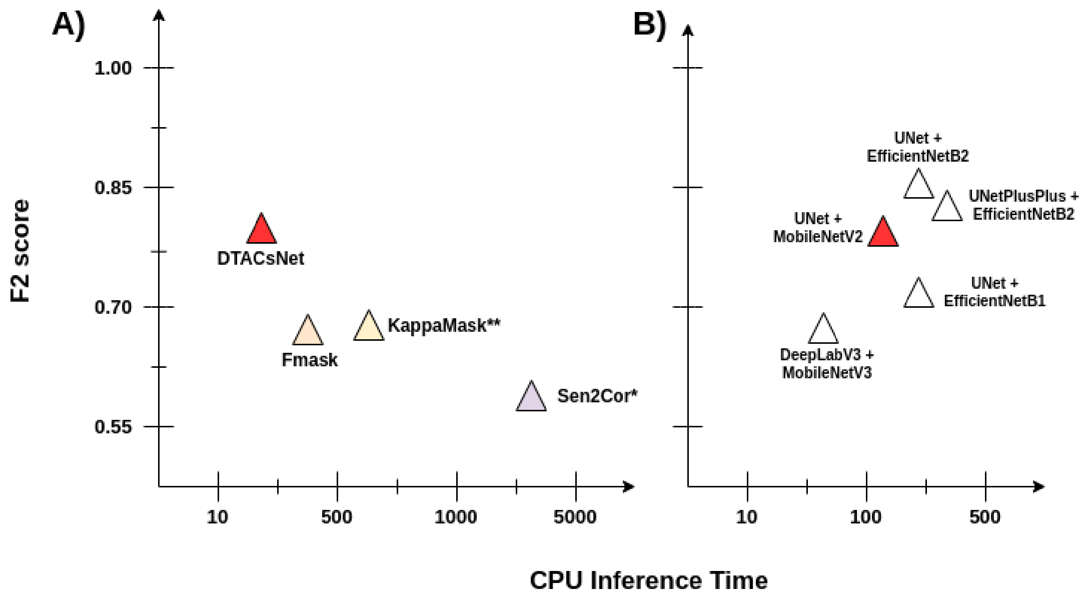


Figure 1. Scatterplot of F2 score vs CPU inference time of cloud detection models. A) Compared DTACSNet with other state-of-the-art models. B) Compared DTACSNet with other encoder/decoder architectures. The F2 score value is computed as the median across all test IPs. Figure 1A, the DTACSNet is better performing (higher F2 score) and faster (lower inference time) than all competing methods. Figure 1B UNet with MobileNetV2 encoder was chosen for final implementation as a good compromise between speed and performance.

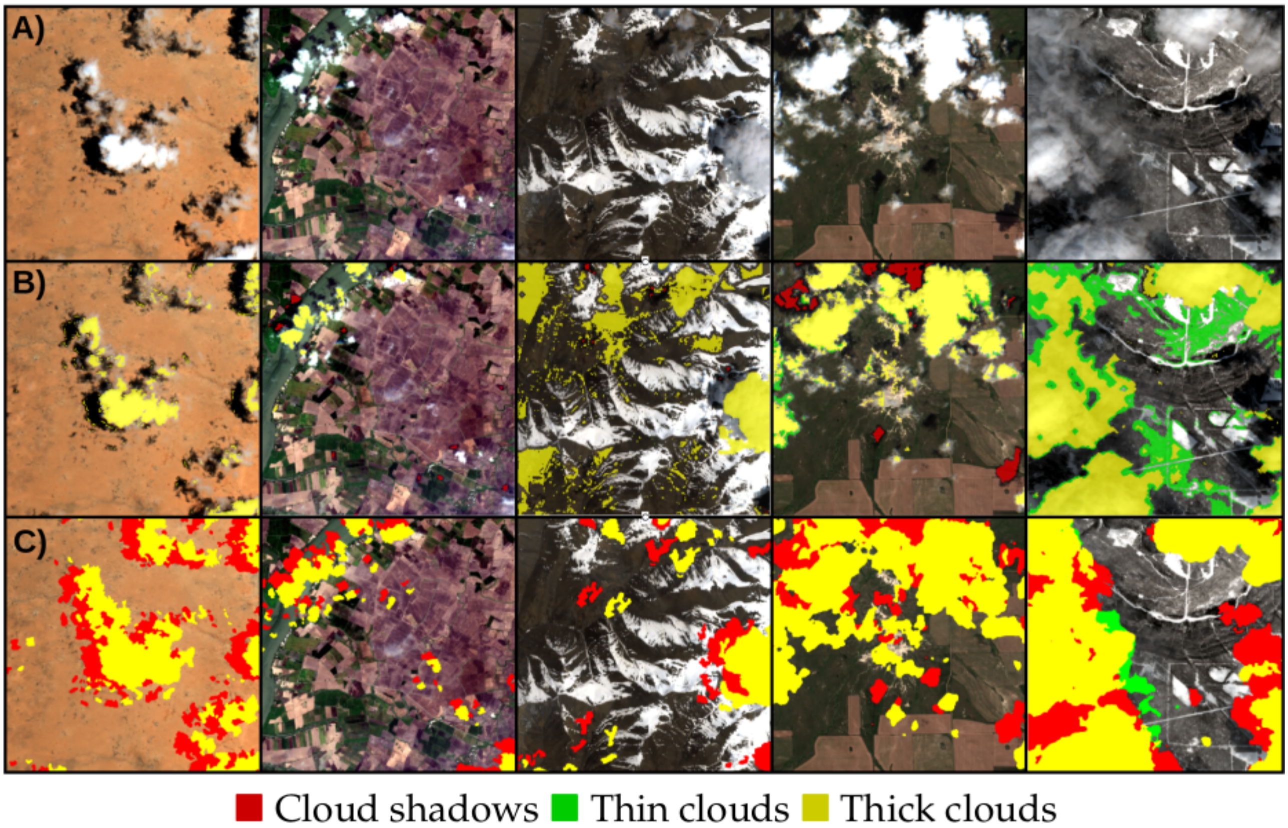


Figure 2. Illustration of the cloud detection methods. (B) Sen2Cor v2.8 and (C) DTACSNet. Clear pixels are displayed in a transparent state, while pixels representing thick clouds are yellow, thin clouds are green, and cloud shadows are red. Compared to Sen2Cor (B) DTACSNet (C) classifies better the thin clouds and cloud shadows as shown in columns 1, 2 and 4. Also it is not confused with snow as shown in columns 3 and 5.

Atmospheric correction results

In figure 3 we show the root mean square error (RMSE) of Sen2Cor surface reflectances compared with DTACSNet and two baselines (no atmospheric correction and a linear model). The error is computed for each of the different bands of Sentinel-2 (x-axis) in two different test datasets (CloudSEN12 and AERONET locations of ACIX [6]). We show big gains of the model against the no atmospheric correction baseline. Notice that not doing atmospheric correction is the default practice in all missions with onboard processing capabilities (to our knowledge). We show that the DTACSNet average is far below the 2% line for all bands except B9. In the ACIX paper [6] Sen2Cor is compared against ground truth reflectances obtained from the AERONET towers. In that work the average RMSE of Sen2Cor is 1.8%. That means that our emulator is almost within noise of the capacity to emulate Sen2Cor surface reflectances. Figure 4 shows several examples from the CloudSEN12 dataset of DTACSNet compared with Sen2Cor. Three first columns show the TOA L1C product, the atmospheric correction of DTACSNet and Sen2Cor (L2A) respectively. Last two columns show the average difference across all bands of DTACSNet against the TOA Sentinel-2 image (L1C) and against the Sen2Cor image (L2A). We see a high agreement between DTACSNet and Sen2Cor with biggest differences occurring in cloud shadows and terrain occluded areas. Figure 5 shows the RMSE of DTACSNet networks trained using only subsets of bands of Sentinel-2 mimicking the band configuration of three *small satellite* missions. As mentioned, the RMSE of the model using all bands is almost the same as the RMSE of the model using only the subset. This is a very surprising result since Sen2Cor atmospheric correction relies on bands B1, B9 and B10 to estimate aerosols and water vapour that are then used to perform the atmospheric correction. On the other hand, DTACSNet here is only using the information of the bands to correct. This result is very encouraging since for these satellites the atmospheric correction relies on gathering ancillary data (to obtain estimations of aerosols, water vapour and ozone). Here we show that it is not needed at least to obtain surface reflectances with an average error below 2% in reflectance units.

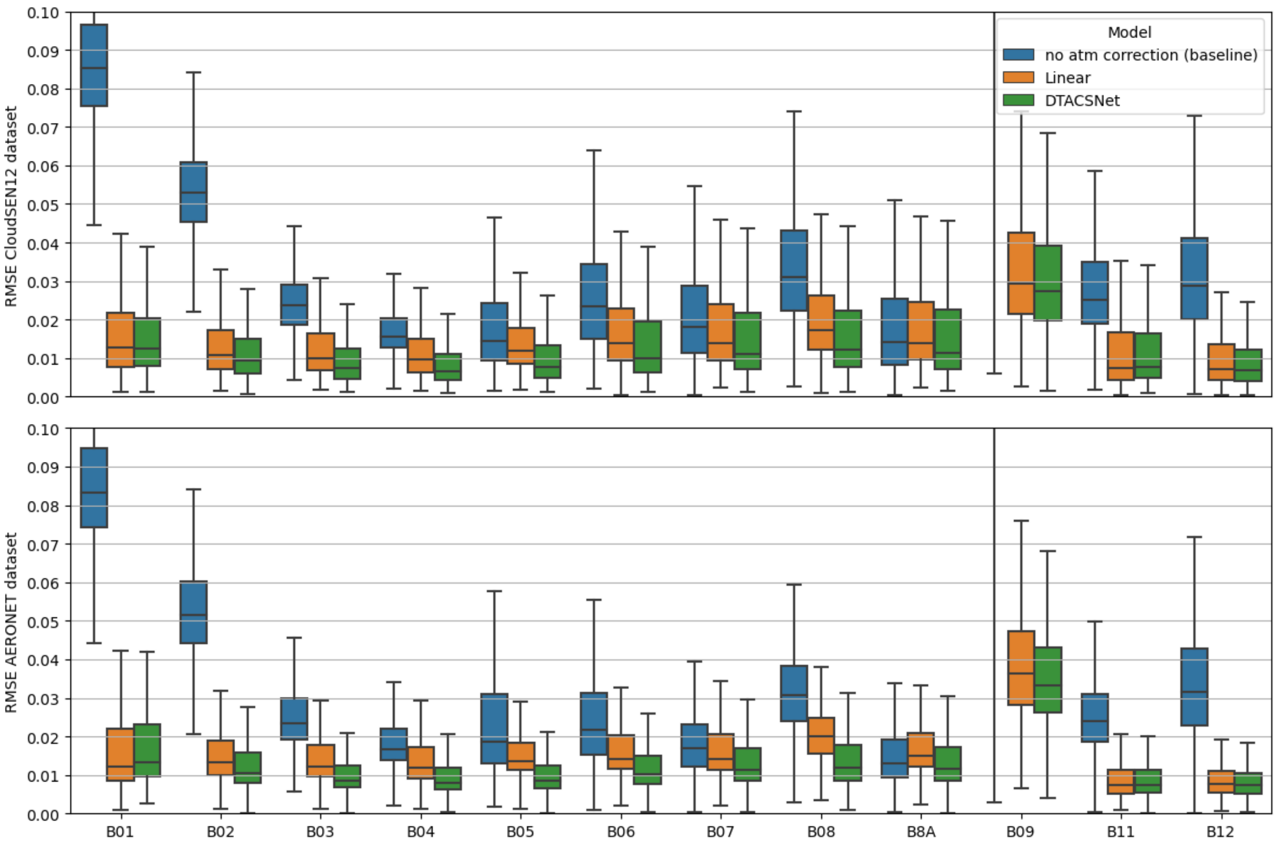
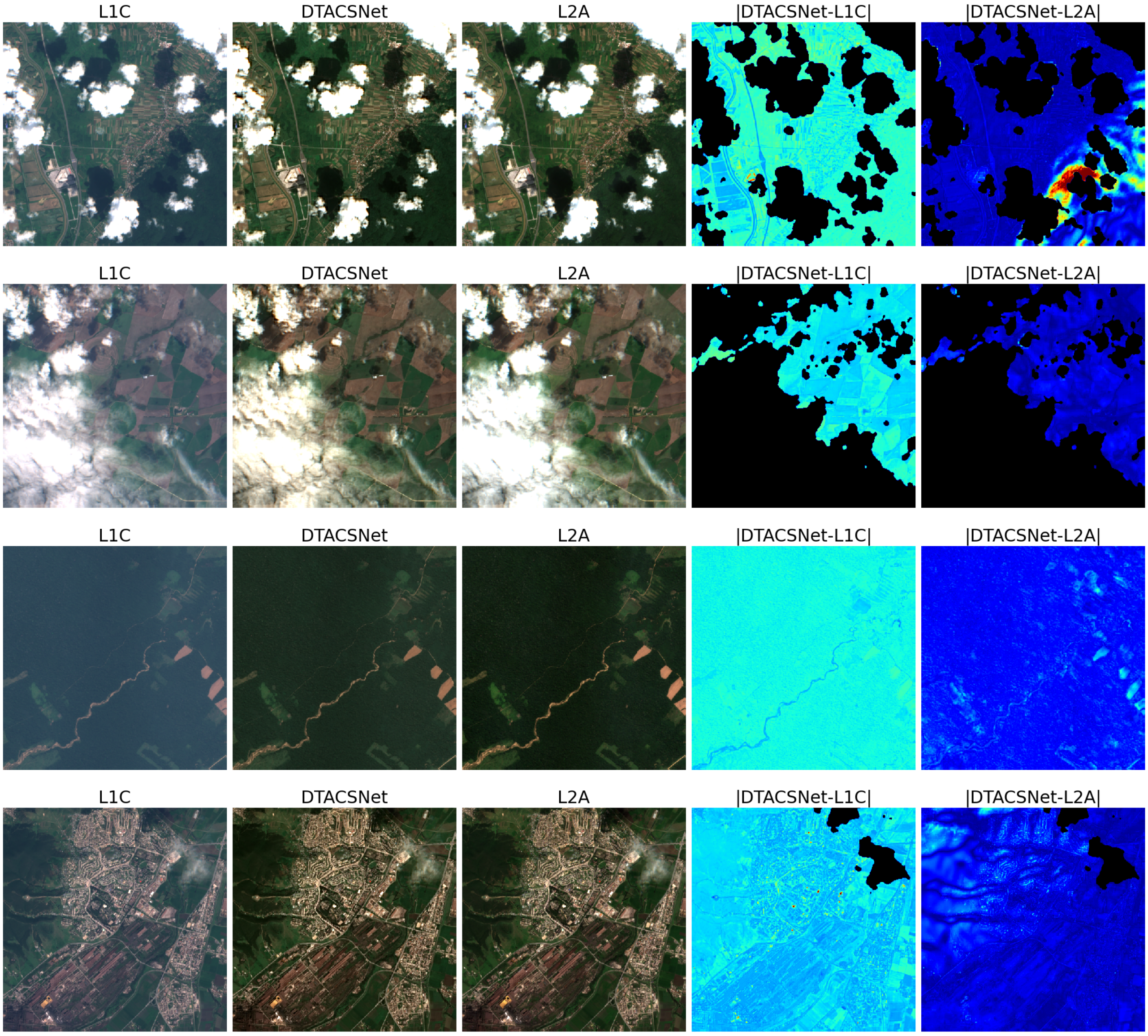


Figure 3 Atmospheric correction RMSE across bands for different models. The boxplot in the top shows the results in the CloudSEN12 dataset whereas the boxplot in the bottom displays the results over the AERONET data. RMSE is calculated by comparing the output of the models with Sen2Cor outputs (i.e. using Sen2Cor as *ground truth*). DTACSNet has errors within 2% for most of the bands (except band 9 which is not very relevant for surface applications). Compared to not doing atmospheric correction (blue bar) it is a significantly better choice.



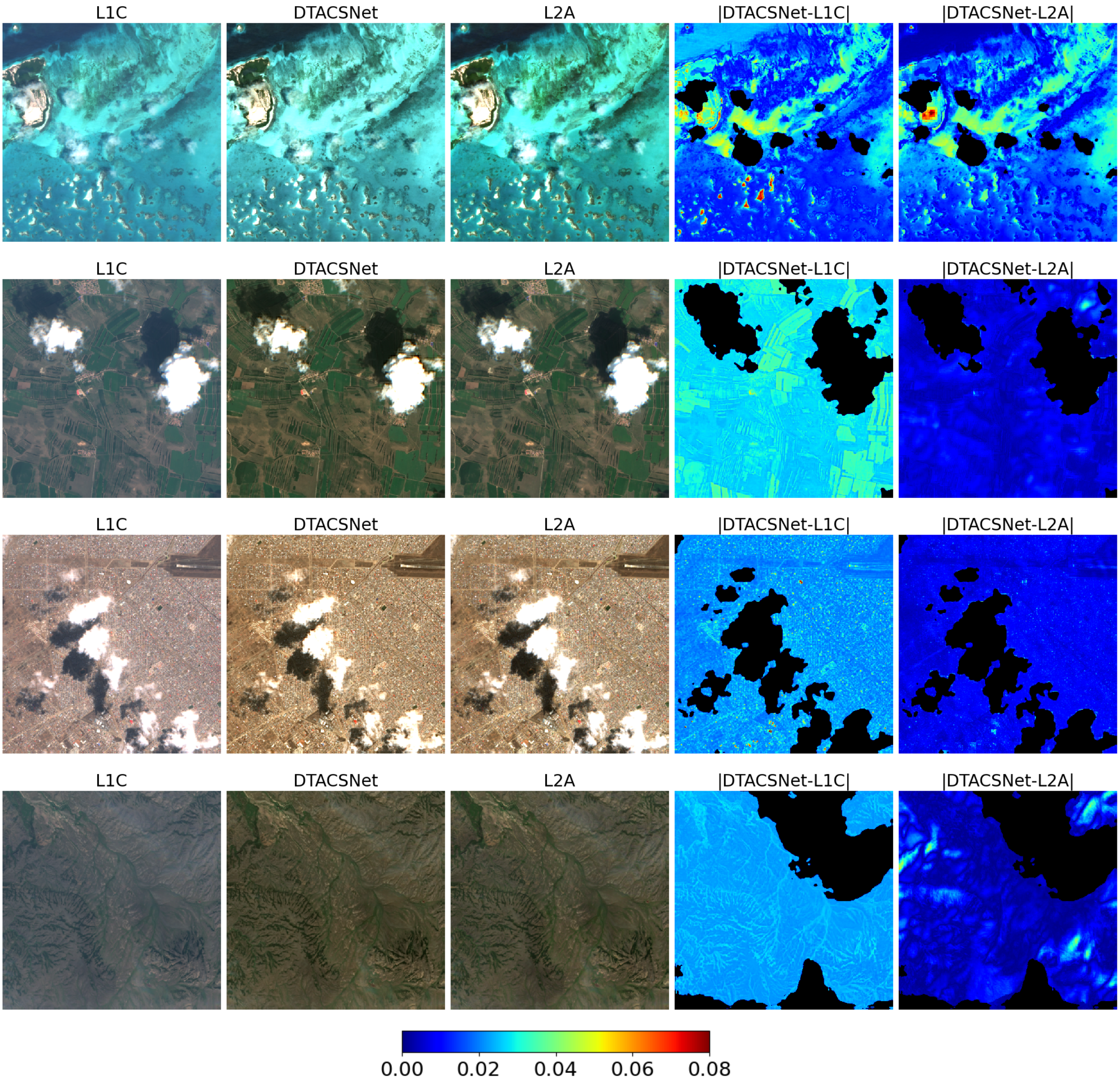


Figure 4. RGB Samples from the test set of CloudSEN12. First column SEN2 TOA image, second column image corrected with DTACSNet AC network, third column SEN2 SR image corrected with Sen2Cor, fourth column absolute differences between TOA and DTACSNet output across all bands, fifth column absolute differences between Sen2Cor and DTACSNet outputs across all bands. In the last two columns clouds and shadows are masked with the DTACSNet SC network.

DTACSNet outputs are very similar to Sen2Cor as we see in columns 2 and 3 in the RGB bands. Across all the bands errors are within 2% as we see in the last columns.

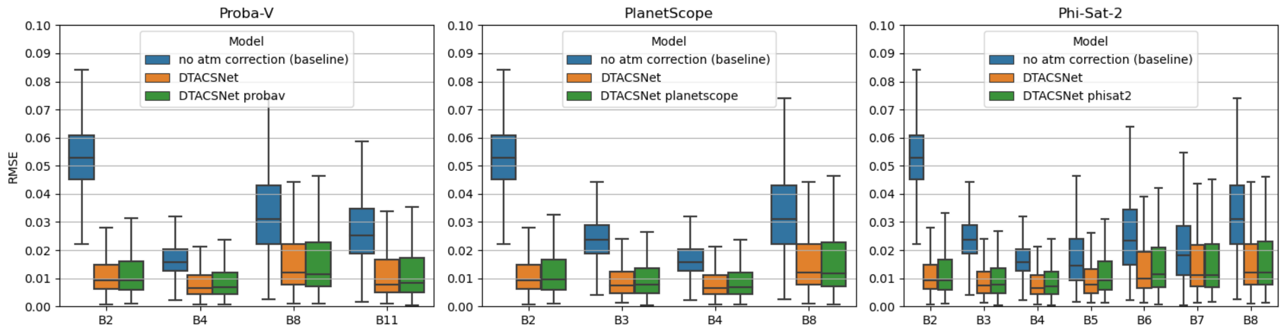


Figure 5 RMSE distribution in the CloudSEN12 dataset of models trained in band combinations of other sensors. We tested the band combination of three small satellite missions: Proba-V, PlanetScope and Φ-Sat-2. When we train DTACSNet with fewer bands, the model output is almost identical to the output using all bands.

Finally, table 1 shows the inference time of DTACSNet on a wall-to-wall comparison with Sen2Cor (including the time to read the image in memory and write back the output to disk). The results are compared in a standard work station with a NVIDIA T4 GPU and on a *SpaceCloud* flatsat provided by UNIBAP with a similar configuration of the mission **[23]**. The main takeaways are that Sen2Cor v.2.10. cannot be run at full resolution on the flatsat due to memory limitations (2GB RAM), that DTACSNet can run it at 10m but still it takes a significant amount of time (26 min) and that running the model at 20m produces speedups of x11 compared to Sen2Cor.

Model	Computer	Device	Res	Scene Classification	Atmospheric Correction	Both
DTACSNet	Google VM	CPU	10 m	2m 26s	1m 49s	**4m 15s**
		GPU	10 m	40s	1m 37s	**2m 17s**
	SpaceCloud	CPU	10 m	21m 25s	9m 5s	30m 30s
		MYRIAD	10 m	16m 56s	15m 47s	32m 43s
		MYRIAD/CPU	10 m	16m 56s	9m 5s	**26m 1s**
		CPU	20 m	5m 21s	2m 23s	7m 25s
		MYRIAD	20 m	4m 23s	4m 1s	8m 24s
		MYRIAD/CPU	20 m	4m 23s	2m 23s	**6m 46s**
Sen2Cor v2.10	Google VM	CPU	10 m	-	-	19m 21s
	SpaceCloud	CPU	10 m	-	-	Memory crash
		CPU	20 m	-	-	1h 1m 27s

Table 1. Table with running time results over a full Sentinel-2 tile taking into account input and output compared against the Sen2Cor process. DTACSNet is faster both in a standard workstation (Google VM with 32 GB RAM, 8 CPUs and a NVIDIA T4 card) and in the SpaceCloud flatsat (1.7GB RAM, 4 CPUs, Myriad-X VPU). In the flatsat Sen2Cor cannot be run at the standard resolution (10m) due to memory constraints.

References

1. Giuffrida, G.; Fanucci, L.; Meoni, G.; Bati˘c, M.; Buckley, L.; Dunne, A.; Van Dijk, C.; Esposito, M.; Hefele, J.; Vercruyssen, N.; et al. The Φ-Sat-1 mission: the first on-board deep neural network demonstrator for satellite earth observation. IEEE Transactions on Geoscience and Remote Sensing 2021, pp. 1–1. https://doi.org/10.1109/TGRS.2021.3125567.

2. Mateo-Garcia, G.; Veitch-Michaelis, J.; Smith, L.; Oprea, S.V.; Schumann, G.; Gal, Y.; Baydin, A.G.; Backes, D. Towards global flood mapping onboard low cost satellites with machine learning. Scientific Reports 2021, 11, 7249. Number: 1 Publisher: Nature Publishing Group, https://doi.org/10.1038/s41598-021-86650-z.

3. Růžička, V.; Vaughan, A.; De Martini, D.; Fulton, J.; Salvatelli, V.; Bridges, C.; Mateo-Garcia, G.; Zantedeschi, V. RaVÆn: unsupervised change detection of extreme events using ML on-board satellites. Scientific Reports 2022, 12, 16939. Number: 1 Publisher: Nature Publishing Group, https://doi.org/10.1038/s41598-022-19437-5.

Publishing Date: 23/11/2022
Project: D-TACS I-2022-00380
Agreement/PO number: 4000138109
Implemented as ESA Initial Support for Innovation

4. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. Sentinel-2 Sen2Cor: L2A processor for users. In Proceedings of the Proceedings living planet symposium 2016. Spacebooks Online, 2016, pp. 1–8.

6. Doxani, G.; Vermote, E.; Roger, J.C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; et al. Atmospheric Correction Inter-Comparison Exercise. Remote Sensing 2018, 10, 352. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, https://doi.org/10.3390/rs10020352.

7. Skakun, S.; Wevers, J.; Brockmann, C.; Doxani, G.; Aleksandrov, M.; Batiˇc, M.; Frantz, D.; Gascon, F.; Gómez-Chova, L.; Hagolle, O.; et al. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. Remote Sensing of Environment 2022, 274, 112990. https://doi.org/10.1016/j.rse.2022.112990.

8. Aybar, C.; Ysuhuaylas, L.; Loja, J.; Gonzales, K.; Herrera, F.; Bautista, L.; Yali, R.; Flores, A.; Diaz, L.; Cuenca, N.; et al. CloudSEN12 464 - a global dataset for semantic understanding of cloud and cloud shadow in Sentinel-2 2022. https://doi.org/10.31223/X5S35G.

23. Mateo-García, G.; Veitch-Michaelis, J.; Purcell, C.; Longepe, N.; Reid, S.; Anlind, A.; Bruhn, F.; Parr, J.; Mathieu, P.P. In-orbit demonstration of a re-trainable Machine Learning Payload for processing optical imagery. Submitted to Scientific Reports 2022. https://doi.org/10.21203/rs.3.rs-1941984/v1.