Unsupervised Domain Adaptation to Mitigate Out-of-Distribution Problem of Spatial Radiometer Images: Application to Quantitative Precipitation Estimation

Vibolroth Sambath, Natanaël Dubois-Quilici, Nicolas Viltard, Audrey Martini, Cécile Mallet

Unsupervised Domain Adaptation to mitigate out-of-distribution problem of spatial radiometer images: Application to Quantitative Precipitation Estimation

Vibolroth Sambath, Natanaël Dubois-Quilici, Nicolas Viltard, Audrey Martini, Cécile Mallet

***Abstract*—A major issue limiting the successful deployment of deep learning algorithms in geophysical applications is their inability to generalize to new contexts. Regarding the quantitative precipitation estimation (QPE) from the Global Precipitation Mission (GPM) satellite constellation, the GPM Microwave Imager (GMI) contains enough co-located brightness temperatures and rain rates data to train a deep learning inverse model. However, the difference in instrumental configurations makes it impossible to directly apply this inverse operator to another space-borne radiometric imager. A domain adaptation is thus necessary to solve the domain shift problem encountered when applying the model trained on one satellite (GPM Microwave Imager, GMI) to another satellite (F18, SSMI/S). Our goal is thus to learn a mapping from SSMI/S to GMI images. In the absence of sufficient paired images between the two satellites, we applied a Cycle consistent Generative Adversarial Network (CycleGAN), which allows for an Unsupervised Domain Adaptation approach. Evaluating the quality of adapted images is a complex problem. This paper employs two tactics: a brief evaluation of adapted radiometric images and a qualitative/quantitative evaluation of rain retrieval. Over several case studies, the domain adaptation step produces adapted SSMI/S images that retain the majority of the rain structure. The rain detection score and intensity bias are then compared using 847 overpasses. The same analysis is carried out over the French territory by comparing the results with rainfall products supplied by Météo-France. In both cases, the adapted images allow the inverse operator to provide a better score in rain detection and intensity.**

***Index Terms*—Precipitation estimation, Unsupervised domain adaptation, Passive Microwave Radiometer, Satellite.**

I. Introduction

A homogeneous and fine-scaled global precipitation map is essential to a wide range of scientific research and societal applications [1]. For instance, global precipitation estimates lead to a better understanding of the water budget on a regional and global scale, to monitoring the changes in the water cycle due to climate change, and to following floods, storms, and dry spells. However, constructing such a map is very challenging due to the intrinsically intermittent nature of precipitation and the scarcity of precipitation observations. Rain intensity varies drastically over time and space. Ground-based instruments alone can only cover small regions and are insufficient for complex terrain, over the oceans, or in developing countries. Therefore, the estimation of rainfall fields from space using remote sensing satellites is a key issue both for the global Quantitative Precipitation Estimates (QPE) and for its evolution in the context of climate change. The Global Precipitation Measurement (GPM) Mission, an international collaboration started in 2014, aims to provide global precipitation products from a heterogeneous constellation of passive microwave radiometers. The consistency of the constellation is reached through the support of the Core Observatory (GPM-Core), whose orbit was chosen so that it would cross the path of all the other satellites. The Core Observatory carries a conical-scanning passive microwave radiometer (GPM Microwave Imager, GMI) and a Dual-frequency Precipitation Radar (DPR). In contrast, the other satellites in the constellation carry mainly a passive microwave radiometer for rain retrieval [1].

Recently, a supervised deep learning model of the U-Net type for rain retrieval has achieved excellent performance in the field [2]. The images of brightness temperatures from four GMI channels are used as the sole input, and the precipitation deduced from the collocated radar observations (DPR) are used as the targets for a quantile regression. The resulting algorithm is called Deep-RAIN (DRAIN), and its complete evaluation can be found in [3]. The DRAIN model, which has learned to generalize well on GMI data, performs poorly on another radiometer due to the differences in instrument characteristics (channel frequency, angle of view, pixel resolution). Using co-located data between GPM-Core and the other satellites of GPM is, in theory, possible to build a supervised learning approach to deploy a version of DRAIN for each constellation member. However, these overpasses are not evenly distributed in latitude, with a much higher probability of occurrence near 60° North or South. An alternate solution is to set up a transfer learning step to convert brightness temperature images of any satellite of the constellation into a GMI-like image on the closest channel basis.

Other rain retrieval algorithms for the GPM constellation face the same challenge due to the differences in characteristics between the various passive microwave radiometer instruments. The most prominent among these algorithms is the

Manuscript submitted 03/08/2023

Vibolroth Sambath (corresponding author) is with the Laboratoire Atmosphères, Milieux et Observations Spatiales (LATMOS), Université de Versailles-Saint Quentin-en-Yvelines Paris-Saclay, 78280, Guyancourt, France. (email: vibolroth.sambath@latmos.ipsl.fr)

Natanaël Dubois-Quilici is with LATMOS. (natanael.dubois-quilici@latmos.ipsl.fr)

Nicolas Viltard is with LATMOS. (nicolas.viltard@latmos.ipsl.fr)

Audrey Martini is with LATMOS. (audrey.martini@latmos.ipsl.fr)

Cécile Mallet is with LATMOS. (cecile.mallet@latmos.ipsl.fr)

operational NASA passive microwave precipitation retrieval algorithm, GPROF [4], which is a Bayesian-based retrieval algorithm. In a recent effort to use neural networks for rain retrieval, a new version, GPROF-NN, has been proposed. The authors rely on simulated brightness temperatures for the other sensors in the constellation [5]. The neural networks for each instrument use the same dataset as the original version of GPROF to train. In GPROF-NN, there are two approaches to estimating rain intensity: a one-dimensional model (in spectral dimension) where a single pixel of brightness temperature is considered and a three-dimensional model (the spectral dimension plus the two spatial dimensions in the horizontal plane) where two-dimensional brightness temperatures (with multiple channels of brightness temperature images) are considered. In both cases, GPROF-NN uses ancillary data coming from global numerical models' (re-)analyses, such as total column water vapor, surface type, and surface temperature. For the one-dimensional model, a new training database of brightness temperatures is simulated from the observations by taking into account the different viewing geometries and resolutions. On the other hand, for the three-dimensional model, a convolutional neural network model is trained to simulate a GMI swath from another satellite swath. Then, this simulated image is mapped into the GMI viewing geometries. The authors noted that the simulation does not guarantee that the distribution of brightness temperatures for the database matches the observations. Therefore, they are corrected with surface types and total column water vapor. It also should be noted that both DRAIN and the 3D version of GPROF-NN are convolutional models.

Transfer learning is an emerging field in machine learning that deals with the scarcity of data. It relies on the transfer of knowledge of a model obtained from a larger dataset to a slightly different but smaller dataset. Transfer learning can take many different approaches, such as fine-tuning the trained model, combining a domain adaptation step and the predictive model within one algorithm, or adapting only the data [6]. Fine-tuning is a supervised approach where a small labeled dataset is needed in order to retrain some or all hyper-parameters of the model. The dataset size depends on the complexity of the task and the number of hyper-parameters to retrain. Models guided by the targeted task, for instance, [7], try to adapt the new data to the known domain under the constraint of the target task's error. However, while this approach sounds advantageous, it still requires a set of labeled data. Domain adaptation aims to solve the domain shift problem encountered and is extensively studied in many applications, such as computer vision, speech recognition, and medical imaging. In [8], the authors presented a review of Deep Learning domain adaptation from theoretical and practical points of view and mentioned some applications in the geoscience field. These applications concerning discriminative tasks (classification, segmentation) deal with satellite images in the visible and IR domain. [9] provides an overview of unsupervised domain adaptation on remote sensing data. However, to our knowledge, no application on microwave radiometric images, whose multi-frequency and multi-scale structures are very specific, and no evaluation of the ability of unsupervised domain adaptation to fill the gap of distribution for regression task application has been made. The challenge for domain adaptation in regression tasks is that pixel values in the adapted images need to be as precise as possible in order to obtain a good performance of the downstream regression task.

Image-to-image translation refers to a class of problems that aims to learn the correspondence between a source domain image and a target domain image using a training set of image pairs [11]. In the application considered in this paper, co-located images between two satellites are scarce and do not provide a representative dataset; paired training images are thus not available. Our goal is to learn a mapping from source to target domain without paired images. Consequently, we turned to unsupervised domain adaptation, which does not assume the availability of labeled target data for model adaptation [12].

The choice of architecture here is based on the previous feasibility study [13], where CycleGAN was used to transform between the 89 GHz horizontal and 89 GHz vertical polarisation of the GMI. CycleGAN [14] is an architecture developed for computer vision problems and trained on conventional RGB images. The previous feasibility study analyzed the applications and limitations of CycleGAN on brightness temperature images and rain retrieval in a simple case.

This article explores the domain adaptation between two different satellites and an in-depth validation of rain retrieval. The second satellite, the DMSP-F18 Defense Meteorological Satellite Program, carries a microwave imager SSMI/S that measures brightness temperatures at frequencies similar to the GMI. An in-depth analysis of a domain adaptation model between these two radiometers, which enables the brightness temperature measurements of all GPM constellation satellites to be converted into GMI-like observations, is a crucial step toward the deployment of a single precipitation restitution algorithm for all constellation satellites. The main objective of this study is to extend the use of DRAIN to all satellites in the GPM constellations. As a result, the choice of channels for DRAIN and for the domain adaptation here depends on the common or similar channels aboard all satellites in the constellations, i.e., the 19, 37, and 89 GHz. In the DRAIN paper [3], it was shown that the 36.6 and 89 GHz channels in vertical and horizontal polarizations are sufficient for rain retrieval. Therefore, in this paper, we choose to adapt these two similar channels of the SSMI/S.

The organization of this paper is as follows. Section 2 describes the characteristics, similarities, and differences between the two instruments, GMI and SSMI/S. It also includes the details of the training and validation database. Section 3 describes the unsupervised domain adaptation model, CycleGAN, and its implementation details. The results of the domain adaptation are shown in sections 4 and 5. First, section 4 offers a visual evaluation of the adapted brightness temperatures. Next, section 5 presents an in-depth evaluation of the rain retrieval performance on the adapted data. Section 6 discusses the advantages and limitations of using CycleGAN as a domain adaptation algorithm for brightness temperatures. Finally, section 7 provides the conclusion and perspective.

TABLE I
TECHNICAL INFORMATION ON THE GMI CHANNELS OF INTEREST [15]

Central Frequency	Bandwidth	Polarisation	NEDT	IFOV	Pixel	Incidence angle
36.64 GHz	1000 MHz	V	0.65 K	8.6 × 14 Km	6.0 × 13.4 Km	52.8°
36.64 GHz	1000 MHz	H	0.65 K	8.6 × 14 Km	6.0 × 13.4 Km	52.8°
89.0 GHz	6000 MHz	V	0.57 K	4.4 × 7.2 Km	3.0 × 13.4 Km	52.8°
89.0 GHz	6000 MHz	H	0.57 K	4.4 × 7.2 Km	3.0 × 13.4 Km	52.8°
Conical: 53° zenith angle; useful swath: 850 km - Scan rate: 32 scan/min = 13.4 km/scan						
Near-global coverage in 2 days; high latitudes ($> 70°$) not covered.						

TABLE II
TECHNICAL INFORMATION ON THE SSMI/S CHANNELS OF INTEREST [16]

DMSP-F18 Defense Meteorological Satellite Program - SSMI/S						
Central Frequency	Bandwidth	Polarisation	NEDT	IFOV	Pixel	Incidence angle
37.0 GHz	1580 MHz	H	0.24 K	27.5 × 44.2 Km	25.0 × 12.5 Km	53.1°
37.0 GHz	1580 MHz	V	0.24 K	27.5 × 44.2 Km	25.0 × 12.5 Km	53.1°
91.655 GHz	2829 MHz	H	0.19 K	13.1 × 14.4 Km	12.5 × 12.5 Km	53.1°
91.655 GHz	2829 MHz	V	0.19 K	13.1 × 14.4 Km	12.5 × 12.5 Km	53.1°
Conical: 53.1° zenith angle, swath 1700 km – Scan rate: 31.9 scan/min = 12.5 km/scan						
Sun-synchronous orbit, Altitude 850 km, Global coverage once/day.						

II. DATA

A. Brightness temperature images

The application of CycleGAN on brightness temperature scenes is not straightforward due to the differences in nature from the conventional images used by the computer vision community in previous applications. In passive microwave remote sensing, a radiometer records the radiant energy arising from the earth's surface and atmosphere. The signal is detected by an electro-optical scanner using a rotating mirror and a set of detectors. The power of the received signal is expressed as a temperature in Kelvin but is called "brightness temperature" (TB) to be distinguished from the physical temperature. The TB in Earth observation results from three main effects: emission and scattering from the atmosphere after reflection by the earth's surface, emission by the earth's surface, and emission and scattering from the atmosphere. The resulting TB is a frequency-dependent complex function of the vertical profiles of the atmosphere and the surface characteristics.

Around 90 GHz, TB responds mainly to scattering from atmospheric ice in the upper part of the clouds; the TB decreases somewhat proportionally in the presence of atmospheric ice over both ocean and land. Near 37 GHz, TB responds on the one hand to emission from atmospheric liquid water (rain or cloud below freezing level) and, when the ice content aloft increases, to scattering by the latter. Depending on precipitation characteristics, the TB of a precipitating cloud will appear either warmer (emission-dominated) or colder (scattering-dominated) than the environment.

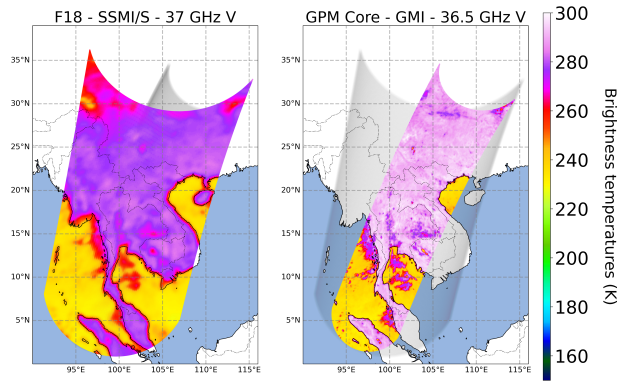
In addition, passive microwave radiometers observe the atmosphere with an Earth incidence angle near 50°; the observed atmospheric volume is thus a tilted column. The signal measured by the radiometer appears in a pixel corresponding to the beam intercepting the precipitation column. The higher up in the atmosphere for a given channel, the more horizontal shift should be expected. This angle effect, called parallax shift, is well documented in several publications [17] [18] [19] [20]. Consequently, a rainy system with a particular 3D spatial structure observed with the same sensor but from different directions may give different measured TB and a frequency-dependent localization of the rainy pixels.

B. Differences between the GMI and SSMI/S

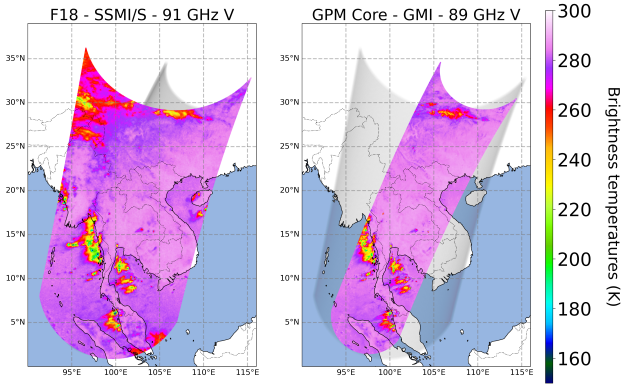
The Instantaneous Field-Of-View (IFOV) is determined by the size of the main reflector, the altitude of the satellite, and the considered channel. However, a fixed receiver integration time and sampling period are used for each radiometric channel. The resulting samples are averaged in the along-scan direction to improve the Noise Equivalent Difference Temperature (NEDT), which characterizes the radiometric sensitivity. The relative motion of the IFOV thus reduces the effective spatial resolution because of the integration time.

The GPM-Core is a non-Sun-synchronous satellite orbiting at an altitude of 407 km and at 65° inclination at the equator [21]. Its main instrument, the GMI, is a conical-scanning microwave radiometer with channels at 10.65, 18.7, 23.8, 36.6, 89.0, 166.0, 183.3+/-3 and 183+/-7 GHz. Among these channels, only the 36.5 GHz and 89 GHz in their horizontal (H) and vertical (V) polarizations are chosen for the domain adaptation. This choice is motivated, first by the performance of these four channels in the training of the deep learning rain retrieval algorithm DRAIN [3] and second, by the number of conically scanning radiometer instruments in the GPM constellation measuring at similar frequencies.

The Defense Meteorological Satellite Program-F18 (DMSP-F18), operated by the National Oceanic and Atmospheric Administration (NOAA), has a set of channel frequencies very similar to GPM-Core. It is, however, a Sun-synchronous satellite at 98.7° inclination at the Equator, carrying on-board an instrument called Special Sensor Microwave Imager/Sounder (SSMI/S) with channels at 19.35, 22.235, 37, 91.67, 150, 183+/-1, 183+/-3, 183+/-6.6 GHz plus some additional channels in the 50-60 GHz for temperature sounding (www.eoportal.org).



(a) SSMI/S - 37 V and GMI - 36.5 V



(b) SSMI/S - 91 V and GMI - 89 V

Fig. 1. Overpass example between GMI and SSMI/S that occurred on 29th May, 2017 over South-East Asia. The first row (a) shows the 37 V channel of SSMI/S and the 36.5 V of GMI. The second row (b) shows the 91 V of SSMI/S and the 89 V of GMI. The geometry at the Earth's surface is taken into account (see text for explanation).

Tables I and II give a summary of the characteristics of the channels relevant to the present study for the two instruments. It can be seen that SSMI/S and GMI are very close in terms of central frequency channels (36.6 GHz instead of 37 GHz and 89 instead of 91.6 GHz) with similar bandwidths. Though there is a slight degradation of the radiometric sensitivity of GMI compared to SSMI/S (0.6 K instead of 0.2 K), DMSP-F18 being about twice higher in altitude than GPM-Core, SSMI/S pixel sizes are thus about twice as large as GMI ones.

This difference in pixel size, combined with the highly intermittent and varying nature of the underlying precipitating processes, has important consequences for the observed TB values. The beam-filling effect is a well-known effect [22], [23] introduced when the field of view of the radiometer is not filled with a uniform rain field. Due to the complex relationships between the atmospheric parameters, the ground parameters, and the observed TB, the difference in spatial resolution necessarily introduces non-linear relationships between TB observed at different resolutions, making it difficult to estimate the relationships between the two scales.

Another side effect of this higher altitude is the much broader swath of SSMI/S when compared to GMI. This induces an adverse effect on the learning database. Because the training is performed on geometrically uncorrected images, the deformation due to Earth sphericity is not exactly the same for the two instruments. Figure 1 shows an overpass case between the two satellites over Southeast Asia while coastal convection is developing. The time difference between the two overpasses is less than 5 minutes. It illustrates the different swath sizes and slight differences in TB for the SSMI/S and GMI channels. One should note that these types of highly favorable overpasses in the Tropics are relatively rare due to the respective orbit inclination of the two satellites. The TB in the 36-37 GHz (figure 1.a) are in the 280-300 K range over land and 240 K over the ocean. The rain systems appear as warm spots at about 260-270 K along the coast over the ocean. This increase in TB is due to emissions by the liquid water. In the 89-91 GHz channels (figure 1.b), the contrast between the land and the ocean background is not as clear as for the 36-37 GHz channels. The overall TB is near 280 K, and the rain systems appear as colder spots going down to 210-220 K, which is due to ice scattering in the cloud. In figure 2, when this overpass is flattened and plotted in a regular grid instead of on a map, the deformation effect due to the instrumental differences as described above is particularly visible when looking at the land masses (yellow-green at 36-37 GHz, darker purple at 89-91 GHz).

Another critical difference in the context of domain adaptation is the difference in orbitography since SSMI/S aboard the DMSP-F18 has global coverage, whereas the GMI aboard the GPM-Core does not overfly high latitudes. One consequence is that the areas overflown are different; even considering long observation periods, the statistical distributions of the observed meteorological situations are different. Thus, the observed TB distributions will reflect the difference in the precipitation distribution. A second consequence is that although the two satellites fly over the same area, their trajectories are different, and the same scene is not observed from the same azimuthal view angle. Thus, the three-dimensional variability of the precipitation systems will be reflected in different ways in the observed two-dimensional TB fields.

C. Training and validation datasets

Due to the relatively rare occurrence of rain in an image, the set of GMI images goes through a pre-selection step to only keep those images with non-zero rain rates from the DPR [2]. The selection criteria are either images with at least 100 pixels containing more than 10 mm/h rain or images with at least 10 pixels containing more than 100 mm/h rain rates. The pre-selection ensures that the transformations could work on rain pixels even though they are quite rare in terms of proportions. On the other hand, there is no rain pre-selection of rain images for F18 as it does not carry a precipitation radar instrument. In addition, as the F18 overfly higher latitudes than the GPM core observatory, F18 measurements from polar regions (latitude greater than 70° N or lower than 70° S) are excluded from the dataset to maintain the similarity of the types of meteorological situations observed. After the pre-selection, the training data set consists of 12,000 images for each domain, making 24,000

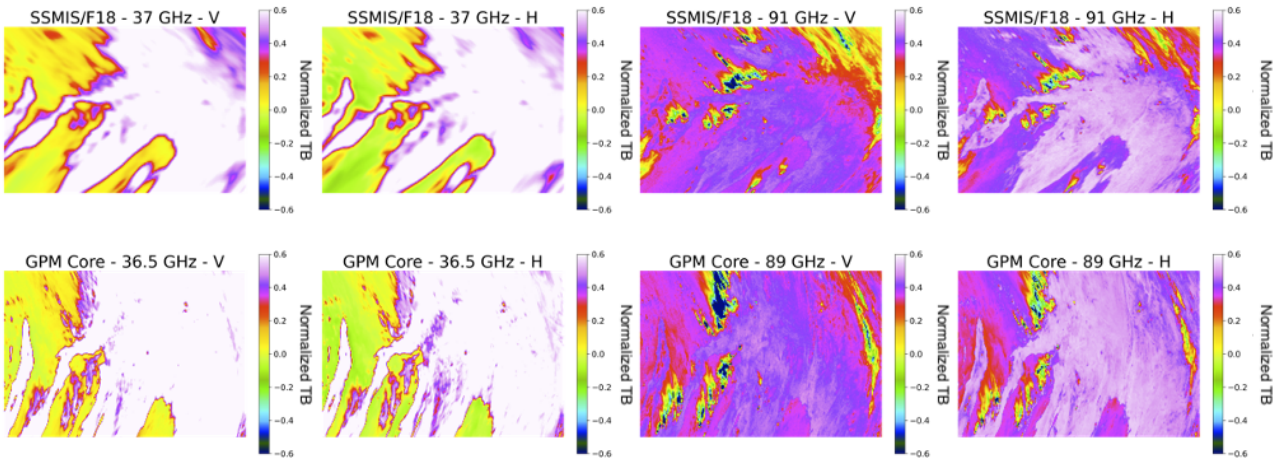


Fig. 2. The same overpass case as in figure 1, but for the four channels of interest for each sensor respectively: top row SSMI/S, bottom row GMI. The brightness temperatures are normalized, and the latitude/longitude geometry is disregarded.

TABLE III
SUMMARY OF TRAINING AND VALIDATION DATASET

	GPM-GMI		F18-SSMI/S	
	Period of observation	Number of images	Period of observation	Number of images
Training dataset	2017	12 000	2017	12 000
Validation dataset	2017	4 000	2017	4 000

images in total of different observations measured during 2017 for both satellites. The validation set comprises 4,000 images for each domain taken from the same period. Table III summarizes the training and validation datasets. None of the images are a corresponding pair. These images contain 221 x 256 pixels and 180 x 256, respectively, for the GMI and SSMI/S images. Next, the training and validation sets undergo the pre-processing step: data normalization, random crop (to 128 x 128 pixels), and random rotation. The random crop (cropping randomly within the image) and random rotation (choosing an angle at random to rotate the image) are added as a data augmentation method to increase the difficulty of the task for the CycleGAN model.

III. METHOD

The deep learning rain retrieval algorithm DRAIN is trained on the GMI TB as input and DPR rain rates as targets [3]. In the DRAIN paper, it was shown that only four channels of TB, 36.6 GHz and 89 GHz in vertical and horizontal polarization, were sufficient (figure 3a). The architecture in DRAIN is based on the U-Net architecture, which is a fully convolutional neural network first introduced for medical imagery semantic segmentation [24]. The cross-entropy loss for image segmentation is changed to quantile regression loss to estimate the rain intensity. The complete evaluation of DRAIN can be found in [3].

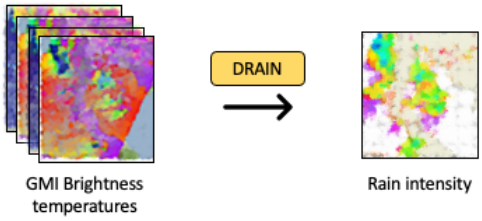
The DRAIN model was trained in a supervised approach, which is optimized on the TB from GMI. Consequently, it performs best on the unseen TB drawn from the same distribution as its training set. However, due to the differences in SSMI/S and GMI described in section II, the TB from the two satellites is similar but not enough to apply DRAIN directly. We chose an unsupervised domain adaptation approach to solve this problem for two main reasons. Firstly, the dataset of SSMI/S is unlabeled, which is to say, without the reference rain rates. As a result, neither fine-tuning nor retraining DRAIN on SSMI/S is possible. Secondly, the number of colocated overpasses between the two satellites is limited, which eliminates the supervised domain adaptation approach. We thus proposed to test a Cycle-consistent Generative Adversarial Network (CycleGAN) approach [14] for the unsupervised domain adaptation between GMI and SSMI/S (figure 3b).

Once the CycleGAN is trained, only the generator that transforms SSMI/S to GMI images is used. The adapted images are then used as inputs to DRAIN, which was previously trained on GMI data (figure 3c). The resulting precipitation estimates will then be compared to the colocated DRAIN rain intensity, the DPR rain intensity, as well as the Mosaic product from Météo-France [25].

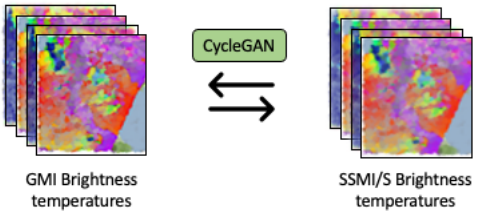
A. CycleGAN architecture

CycleGAN is an unsupervised domain adaptation method based on the Generative Adversarial Networks (GANs) [14]. It contains two GANs working together, one for each domain. Given two domains X and Y, the first generator G transforms images from its source domain X to its target domain Y. The second generator F works inversely by transforming data from its source domain Y to its target domain X. To complete the GAN structure, two discriminators, D_X and D_Y, try to classify real and generated images of their respective domains.

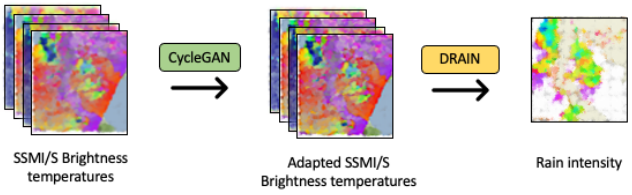
To train the two GAN generators in an unsupervised approach, the weights of the models are updated using the sum



(a) The rain retrieval algorithm DRAIN is trained with GMI brightness temperature as inputs and DPR rain rates as targets.



(b) The CycleGAN has two generators that can transform back and forth between the GMI and SSMI/S brightness temperatures.



(c) The SSMI/S brightness temperatures need to be adapted to the domain of the GMI before applying the DRAIN model for rain retrieval.

Fig. 3. The summary of the proposed approach. (a) shows the DRAIN rain retrieval on GMI. (b) shows the training of CycleGAN. (c) shows the domain adaptation and precipitation estimation on SSMI/S.

of identity loss, cycle consistency loss, and adversarial loss. The identity loss is the difference between an image and a transformation to its domain, for example, $|x - F(x)|$ for x in X. Based on the seminal paper [14], the identity loss helps to regularize the generator to be near an identity mapping; it preserves the color of the input paintings in computer vision applications. In the same way, the feasibility study [13] shows that the identity loss allows the model to conserve the correct intensity of the TB. On the other hand, cycle consistency is the difference between the original image and its complete transformation; that is, $|x - F(G(x))|$ for x in X. Considering that a complete transformation of an image should be close to the original image, the training of the generators does not require paired images. The adversarial loss is based on the performance of the discriminators and operates on the same principle as in a conventional GAN.

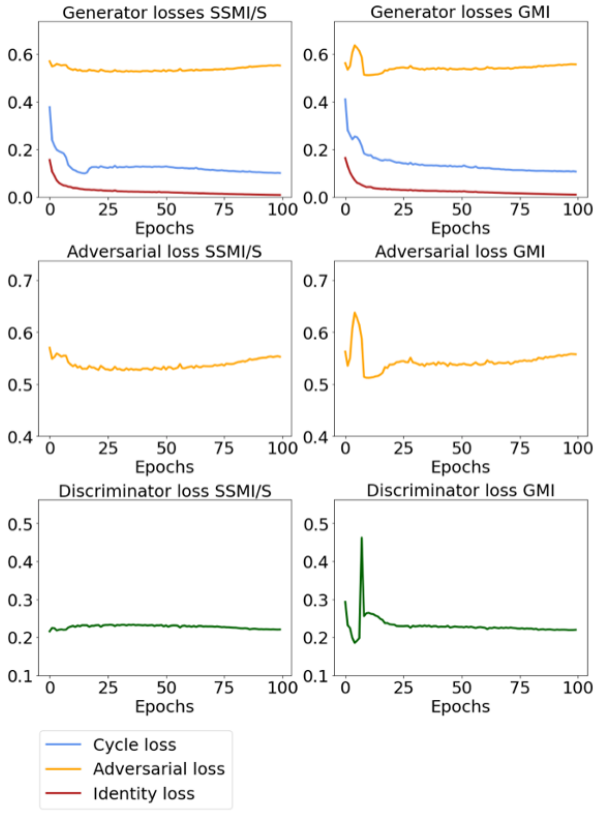


Fig. 4. Training loss for the generators and discriminators for the two domains (left column for the GMI and right column for the SSMI/S). The first row shows the three losses of the generator, while the last row shows the loss of the discriminator. The middle row shows the adversarial loss of the generators alone for better comparison to the discriminator loss.

B. *Implementation details*

Since the U-Nets architecture has proven suitable for the rain/TB problems [2], two such networks are used as the two generators (F and G) in the CycleGAN. The skip-connections in U-Net are very efficient in preserving the fine-scaled structure within the brightness temperature images, as previously stated in [2]. As described in [14], the discriminators are based on the PatchGAN discriminator [26], whose architecture is a series of convolution layers.

The results of our experiments show that the convergence of the CycleGAN is not trivial. The adapting domain performed with unpaired data does not rely on any task-specific, and the problem is thus under-constrained. Given the specificity of our images, unlike computer vision images, visually determining whether the result is realistic is impossible. In our case, the resulting TB fields must retain the essential properties of the precipitation characteristics that cannot be assessed by simple visual analysis. During training, the optimization of the hyperparameters relies on the learning curves to obtain the behavior shown in figure 4. An empirical approach based on the following principles is used: keeping similar loss for both domains, keeping discriminator loss slightly lower than generator loss, and stabilizing the training curve.

In order to obtain similar losses for both domains, different learning rates are necessary. The optimization of the generator

depends in parts on its ability to deceive the discriminator. Therefore, the quality of the generated images depends on the performance of the discriminator. Each time the generator deceives the discriminator, the discriminator needs to find another criterion to differentiate between real and generated images. The rivalry principle is then essential. Based on the results of our experiments, the discriminator should always be the best of the two models. Still, the generator must sometimes be able to catch up to stimulate the discriminator. In numerous runs of the model, the discriminator converges quite easily and quickly. To solve this problem, the generator's and the discriminator's learning rates are set differently due to the two-time-scale update rule (TTUR) [27]. It has been shown that GANs converge better and with fewer epochs using this method. The learning rate of the generator is twice that of the discriminator due to the complexity of its task and to increase the rivalry. In addition, the task for the generator and discriminator pair for one domain (generating a distribution corresponding to a higher spatial resolution image from one of a lower resolution image, i.e., generating GMI domain from SSMI domain) is more complicated than the opposite (from GMI to SSMI domain). For this reason, the learning rate for the SSMI to GMI pair is twice as high as for the GMI to SSMI pair. Therefore, the learning rate for the generator and discriminator pair for GMI to SSMI/S is 2e-4 and 1e-4, respectively, while the pair for SSMI to GMI is 4e-4 and 2e-2, respectively. This results in a similar loss for both domains, as shown in the training loss in figure 4.

The choice of cost functions, normalization, and weighting coefficients is fundamental to achieving a balanced reduction in the three terms of loss. Many different configurations were tested before an acceptable one was found.

To train CycleGAN, the L1 loss is used for the cycle and identity losses, while the Mean Squared Error (MSE) loss is used for the discriminator and the adversarial loss. For the discriminator, due to the value of the labels (0 for generated images and 1 for real images), these errors are, therefore, between 0 and 1, as shown in figure 4. With MSE, squaring the error will then lead to a decrease in the value used for backpropagation. To increase the impact of the adversarial loss in the generator, the MSE loss has been replaced by the L1 loss. Results from multiple training in this study using this approach show better-generated satellite images. On the other hand, replacing the MSE loss with the L1 loss for the discriminator did not result in any improvement, as the discriminator encountered more difficulty in converging. As there are three losses for training the generators, they are balanced by three coefficients. In both the feasibility study [13] and the proposed model here, the same coefficients as proposed by [14] were used (10-5-1 as the coefficients for the cycle-identity-adversarial losses, respectively).

Therefore, our full loss function for the generators is $L_{generator} = 10L_{cycle} + 5L_{identity} + L_{adversarial}$, where:

$$\begin{aligned} L_{cycle} =& E_{x\sim p_{data}(x)}[\|F(G(x)) - x\|_1] + \\ & E_{y\sim p_{data}(y)}[\|G(F(y)) - y\|_1] \end{aligned}$$

$$\begin{aligned} L_{identity} =& E_{x\sim p_{data}(x)}[\|F(x) - x\|_1] + \\ & E_{y\sim p_{data}(y)}[\|G(y) - y\|_1] \end{aligned}$$

$$\begin{aligned} L_{adversarial} =& E_{x\sim p_{data}(x)}[\|D_Y(G(x)) - 1\|_1] + \\ & E_{y\sim p_{data}(y)}[\|D_X(F(y)) - 1\|_1] \end{aligned}$$

In order to stabilize the training of CycleGAN, spectral normalization introduced by [28] was used to replace the batch normalizations. In [28], spectral normalization is only proposed for the discriminator. However, according to [29], spectral normalization can also be applied to the generator. In the present study, spectral normalization was thus used to replace the normalization methods for the two U-Nets (generators) and the discriminators. The effect of the spectral normalization results in a relatively smoother loss curve, as shown in figure 4. This normalization results in the best score in the validation by rain intensity shown in section V. Spectral normalization is a method for controlling the variations of the model function under a constant by normalizing the weights to satisfy the 1-Lipschitz constraint. This greatly reduces the large variations and thus makes the training more stable, limits the impact of local minima, and, as a result, allows the model to converge faster. It has also been shown that this method avoids the vanishing gradient problem.

After defining the implementations using the empirical approach described above that makes the training loss smoother and retains a balanced loss between the two domains, CycleGAN is still very sensitive to initialization. Several models have been trained with different initial conditions. As shown in figure 4, the discriminator and adversarial loss for the two domains exhibit extremely slow convergence. In each training, the current model was saved along with the best generator model. The best generator here is defined as having the lowest cycle and identity losses on the validation set. As the problem is under-constrained, the downstream rain retrieval task results differ for each training. Therefore, the best generator in each training (corresponding to different initial conditions) was tested on the rain retrieval task to evaluate the whole performance of the process as defined in figure 1.c. The model presented here was trained over 100 epochs. However, the best generator on the validation set appeared at around epochs 20. It is also the generator that gives the best score in the rain retrieval task. More discussions on that aspect are presented in section VI.

IV. Evaluation of adapted brightness temperature images

The main difficulty in the present application of CycleGAN on TB fields is the assessment of the quality of the adaptation since the images that are used are somewhat abstract. Unlike in the original paper [14], one can not easily recognize an object inside the image and immediately and subjectively measure how well it was transformed. Concerning the spatial properties, CycleGAN is a model that cannot make any drastic changes to the structure of the image [14]. It contains the Cycle Loss and the Identity Loss, which are pixel-to-pixel differences between the images. Cycle loss is the pixel-to-pixel differences between an image and its complete cycle adaptation; for example, an SSMI/S adapted to GMI then adapted back to SSMI/S. It quantifies the dissimilarity between an image and its complete cycle adaptation. That means that the dissimilarity between

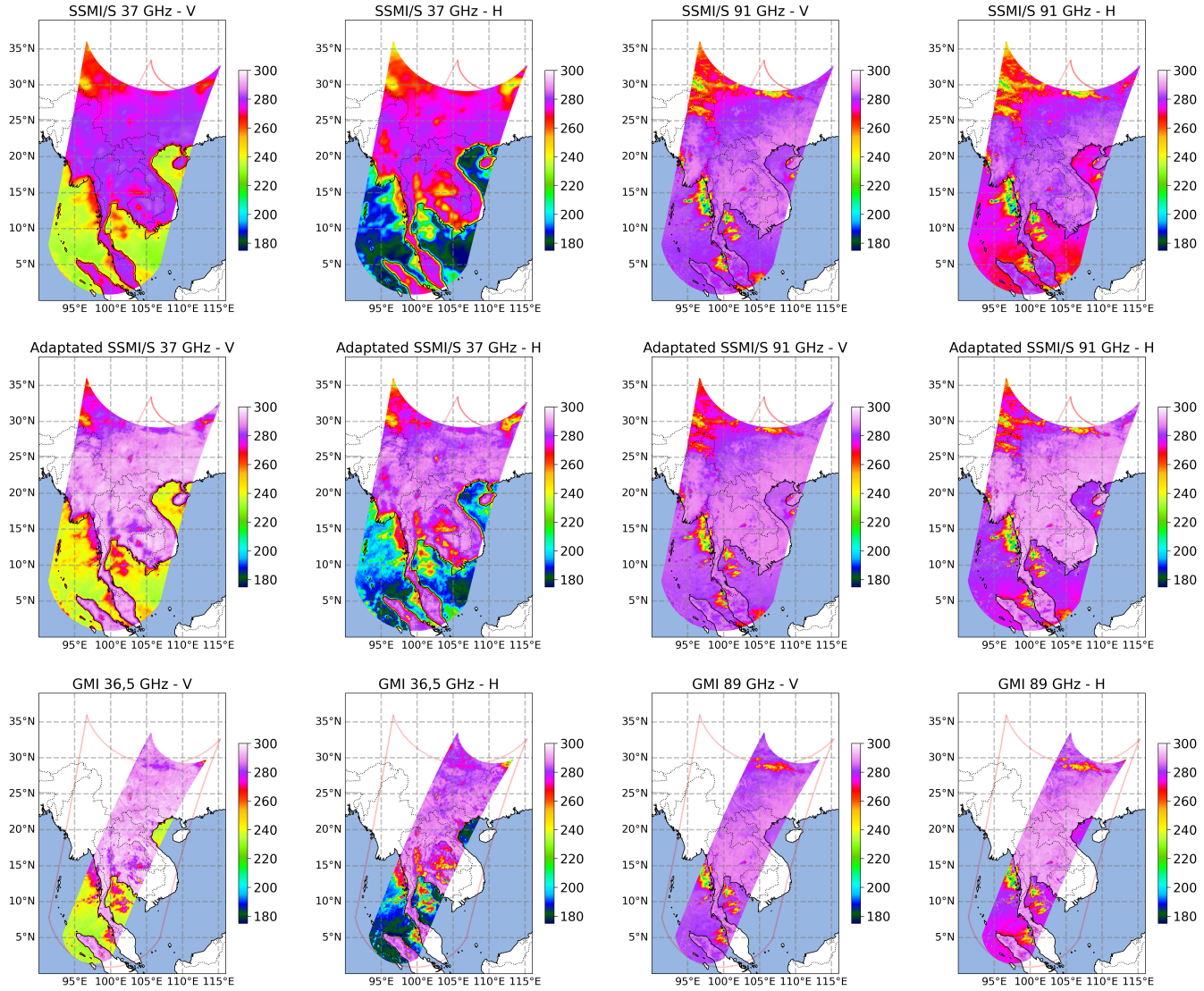


Fig. 5. Brightness temperatures in Kelvin in the same observation as figure 2 but middle row shows the adapted SSMI/S channels (SSMI/S transformed into GMI).

SSMI/S adapted to GMI, then adapted back to SSMI/S, and the dissimilarity between GMI adapted to SSMI/S, then adapted back to GMI, are minimized during training. Identity loss, on the other hand, is the difference between an image and its identity transformation, i.e., SSMI/S adapted to SSMI/S. The first step in evaluating the generated images is to confirm that the structures are preserved throughout the cycle of transformations. To monitor that there is no transfiguration of the structure, we monitored the loss and displayed the pixel difference between the images and their transformations to ensure that there is no structural high error. The quantification of these losses during the validation process further confirms that they are also minimized when considering unseen images.

Secondly, the visual evaluation is done in complement to the monitoring of the training and validation loss. A few overpasses of the two radiometers GMI and SSMI/S over the same area are used to qualitatively analyze the TB fields. As mentioned earlier, and as shown in Figure 1, the two radiometers do not have the same swath width or the same angle of view. It is, therefore, not expected to obtain precisely the same TB field. The model is supposed to correct for systematic differences such as resolution and frequency but not for differences due to different viewing angles of the same system since this is not an image translation problem. During its flight, each satellite encounters different precipitation systems, which it sees from different angles depending on the relative positions of the latters and the satellites.

Figure 5 shows the adaptation results of SSMI/S to GMI on the overpass case presented before. The first row is the original images of the four SSMI/S channels. The second row shows the adapted images, and the third row shows the GMI channels. The most prominent feature is the small increase in the perceived resolution of the adapted images, particularly the 37-GHz channel, as some details become visible in the

TABLE IV
SUMMARY OF TEST DATASET

	Number of pixels
Overpass between GMI and SSMI/S (section V-A)	$\sim 88 \times 10^5$
Overpass between DPR and SSMI/S (section V-B)	$\sim 36 \times 10^5$
Colocation between Météo-France and SSMI/S (section V-C)	$\sim 25 \times 10^5$

adapted images. This validates the SSMI/S-to-GMI generator's capacity to increase the lower resolution of the SSMI/S to match the higher resolution of the GMI. The adapted images also conserve all precipitation structures, even if they appear somewhat deformed because of the difference in view angle.

V. Evaluation on quantitative precipitation estimation task in the target domain

The previous section shows an evaluation of the domain adaptation between F18 and GPM observations. Since the goal of the adaptation is ultimately rain retrieval, an in-depth assessment of the latter is required to evaluate the transfer learning method's potential. The rain retrieval step takes the adapted SSMI/S data from the CycleGAN generator directly as inputs to the DRAIN model. The complete details of the training and performance of DRAIN on GPM can be found in [3].

Quantitative comparisons are performed on three different data sets (Table 4). The 847 passes of the two satellites between latitudes 70° N and 70° S during the whole year 2019 are considered. A first test set is constituted from the intersection of the swaths of the two radiometers. A subset containing only the observation area covered by the narrower swath of the DPR radar constitutes the second test set. The third test is a co-location in time and space over the whole of 2019 with the Météo-France five-minute rain product at a resolution of 1 km, which is a good reference for mid-latitude QPEs.

For these three test sets, a pixel-by-pixel comparison is performed. The ability to detect rain, which is to distinguish rain and no-rain pixels, is analyzed through the contingency table. As the rainy/non-rainy classes are not balanced, we have a majority of non-rain pixels, which completely distorts the computation of accuracy. For example, due to the rare cases of rain pixels, if a model predicts non-rain for all pixels, it still has a relatively high accuracy score. Therefore, the F1-score that appears more appropriate than the latter is calculated to facilitate comparison. F1-score resolves the unbalanced data by taking into consideration both recall and precision scores. The recall score is the percentage of the True Positives out of all the positive samples in the dataset. In contrast, the precision score is the percentage of True Positives out of all the predicted positives. Combining these two metrics can deal with the rarity of the positive data points. The F1-score formula is given by,

$$F1 = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The bias and standard deviation calculated on all pixels are dominated by zero values, which correspond to the majority of true non-rain pixels. To emphasize the improvement in rain intensities, the true positives (pixels considered as rainy simultaneously by DRAIN and by the used rain reference) are then retained. The bias and the standard deviation of the rainfall rate differences are finally calculated on these pixels.

A. Comparison between overpasses

As described in section II, the resolution and swath of the two satellites are different. In order to compare the DRAIN estimation on GMI, non-adapted, and adapted SSMI/S, it is essential to colocate the pixels of GMI and SSMI/S. A KD-Tree algorithm [30] (implemented with SciPy library in Python) is used to calculate the nearest SSMI/S pixels of a given GMI pixel. The rain rate computed for that nearest pixel from the SSMI/S TB (adapted or not) is then affected to the GMI pixel in question to allow a direct comparison.

Figure 6 shows the result for the Southeast Asian case presented in Figure 1. Pixels with no rain (RR=0) are not plotted. The left-hand side image is the reference DRAIN applied directly to GMI data. The center image is the fully adapted, and the right-hand-side image is the non-adapted result. In this example, dominated by tropical convection, the non-adapted result does not perform too poorly in terms of the general structure of the rain regions. Locally, however, some rain pixels are retrieved as non-rainy, and the rain intensities are substantially off, with rain rates no greater than 20 mm/hr. On the fully adapted image, both the structure and intensities appear much closer to the reference, showing that the adaptation indeed succeeded in transforming the SSMI/S TB image into a GMI TB image. Two more overpasses are shown in the Annex.

Table V presents the contingency and base statistics of the adapted and original SSMI/S TB against DRAIN-GMI used as a reference for the 847 overpasses mentioned above. One must note that most of these overpasses occur at high latitudes where the GPM-core reaches its maximum latitude (65°). As expected in both cases, nearly 95 % of the pixels are non-rainy. However, adaptation clearly improves the fraction of true positives, which goes from 0.66 % before adaptation to 2.61 % after. This improvement comes mainly from the bad detections, which go from 3.12 % before adaptation down to 1.17 % after. Table V also shows that the adaptation clearly improves the score on both rain and non-rain predictions. The F1-score increases from 0.20 for the non-adapted images to 0.66 for the adapted data. Both the bias and standard deviation of errors are significantly reduced from 1.64 to 0.75 mm/hr and 3.05 to 1.74 mm/hr, respectively, while the number of rain pixels has increased considerably for the adapted data. The reduction in overall error may be largely due to the number of low rain-intensity pixels that the adapted images are able to capture. The improvement in intensity is comparable between

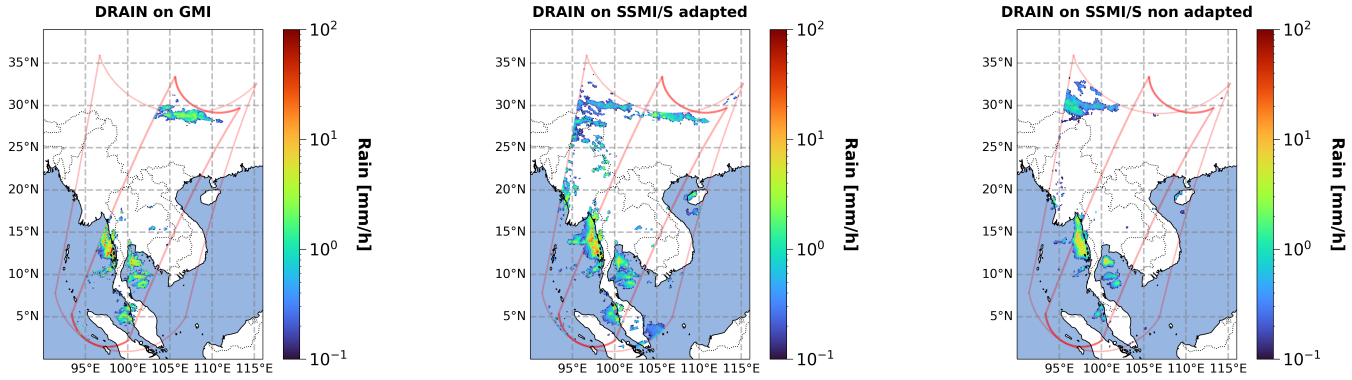


Fig. 6. Same as figure 1 but for the rain intensities retrieved by DRAIN. The left-hand side is DRAIN on GMI data. The Center is DRAIN on adapted SSMI/S data. The right-hand side is DRAIN on unadapted data. All rain rates in mm/hr. Both SSMI/S-inferred rain rates are co-located with the GMI pixels for comparisons.

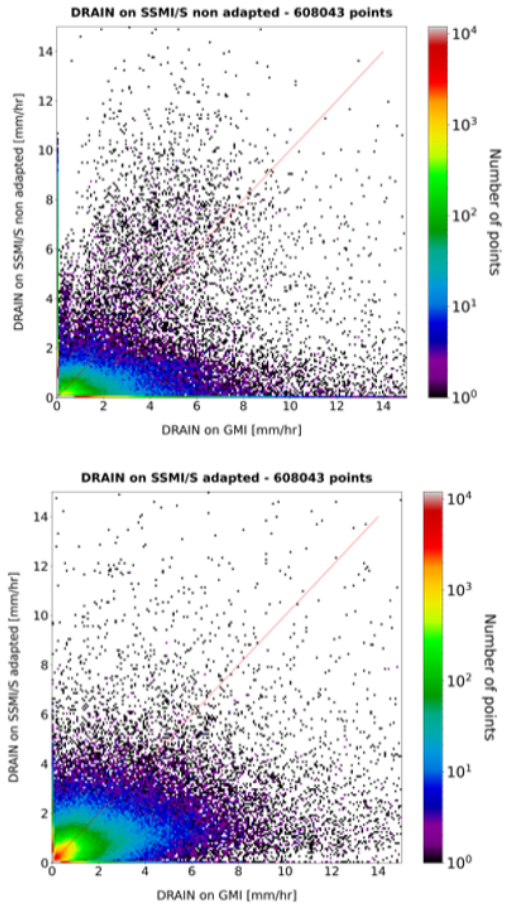


Fig. 7. Scatter plot for the 847 overpasses between GMI and SSMI/S data, after retrieval with DRAIN and co-location to the GMI pixels. The left-hand side plot is for non-adapted SSMI/S data, and the right-hand side is for adapted ones. The number of points above each plot indicates the points where at least one of the three datasets (DRAIN on GMI, on original SSMI/S, and on adapted SSMI/S) has non-zero rain rates.

adapted and non-adapted for the rain intensity between 1-3 mm/hr. The bias reduced from 1.24 mm/hr (mean target rain rates at 1.75 mm/hr) to 0.90 mm/hr (mean target rain rates at 1.66 mm/hr).

Figure 7 shows the scatter plots corresponding to the statistics of Table V. The improvement brought by the adaptation is clear, although, above 3 to 4 mm/hr, an underestimation of the intensities can still be seen. However, these overpasses comparisons must be taken with some degree of caution because first, the co-location of the SSMI/S pixels with the GMI pixels might affect the statistics, and second, most of these overpasses happen at high latitudes where rain intensity is predominantly light, and DRAIN might not be giving its best performances (see [3]). In addition, the difference between the three cases is the ability to capture low rain rates. DRAIN and the adapted case are better at estimating low rain, which leads to lower average rain rates for the colocated pixels.

B. Comparison with DPR

In this section, DRAIN estimation on colocated data between the GMI and the SSMI/S is compared to the DPR data. As in the previous section, data of the 847 overpasses between F18 and GPM-Core are considered. The SSMI/S data are used as an input in DRAIN and co-located afterward with the GMI pixels. The co-location of DPR pixels was carried out in a previous study ([2]) as follows: depending on the location of the pixels in the swath and in an orbit, 1 to 5 DPR pixels are averaged for each GMI pixel within the DPR swath. Due to the smaller swath of DPR, there are less colocated data than in Table V. As expected, the performance in Table VI is slightly worse than in Table V. The precipitation rates estimated by DRAIN in the adapted target domain (adapted SSMI/S) are closer to those estimated in the source domain (GMI) than to the labels (DPR) learned by DRAIN.

However, performance improvements are brought by the adaptation and are clearly shown in Table VI. The True Positives go from 0.67 % for the non-adapted data to 2.66 % for the adapted ones. This result is to be compared with the 4.25 % of the GMI data. True Negatives are similar for all three configurations, and, once again, it's the False Alarm rate that changes the most between non-adapted at 5.27 % and adapted at 3.28 %. These values remain higher than the optimal 1.95 % for GMI data. A substantial improvement is also seen in the F1 score, the bias, and the standard deviation

TABLE V
RAIN RETRIEVAL RESULTS WITH DRAIN ON GMI AS THE REFERENCE.

	non-adapted SSMI/S		adapted SSMI/S	
DRAIN on GMI	Rain	Non-Rain	Rain	Non-Rain
Rain	0.66%	3.12%	2.61%	1.17%
Non-Rain	1.85%	94.35%	1.45%	94.75%
F1 score	0.20		0.66	
Bias on rain (mm/hr) (TP)	1.64		0.75	
Standard deviation of error (TP)	3.05		1.74	
Number of rain pixels (TP)	58799		232164	
Mean of rain rates in mm/hr (TP)	2.19		1.21	

TABLE VI
RAIN RETRIEVAL RESULTS WITH DPR AS THE REFERENCE.

	non-adapted SSMI/S		adapted SSMI/S		GMI	
DPR	Rain	Non-Rain	Rain	Non-Rain	Rain	Non-Rain
Rain	0.67%	5.27%	2.66%	3.28%	4.25%	1.95%
Non-Rain	1.81%	92.23%	1.16%	92.88%	0.76%	93.02%
F1 score	0.16		0.55		0.75	
Bias on rain (mm/hr) (TP)	2.47		1.13		0.46	
Standard deviation of error (TP)	6.36		3.65		1.76	
Number of rain pixels	24584		96740		286884	
Mean of rain rates in mm/hr (TP)	2.98		1.61		1.17	

of the error, but the performances of DRAIN on the adapted data cannot quite match the performances on the original GMI ones. Similar to the previous evaluation, the lower mean target rain rates in Table VI are due to the increase in the capability of capturing low rain rates. However, for rain rates between 1-3 mm/hr, the non-adapted error is 1.33 mm/hr (for mean target rain rates of 1.79 mm/hr), while the adapted model error is 1.00 mm/hr (for mean target rain rates of 1.69 mm/hr). For this interval of rain rates, the error for DRAIN is 0.55 mm/hr with the mean target rain of 1.63 mm/hr.

The same caution should be used when looking at these results as for the previous ones, but they give an idea of the potential of the domain adaptation and the margin of progression left.

C. Comparison with Météo-France Mosaic product

This section presents an evaluation of the rain retrieval results of the adapted images over an independent set of validation data. As in section V-A and V-B, the results will be evaluated over the whole year 2019, which was not part of the training of either DRAIN or CycleGAN. The Météo-France mosaic [25], hereafter mosaic, provides a precipitation product at 5-minute intervals over a 1536x1536 pixel grid with a resolution of 1 km. Due to the difference in data structure between the mosaic and the satellite observation, the comparison requires a co-localization process again. All satellite and mosaic data are averaged in a 0.2 ° x 0.2 ° grid. Each satellite overpasses France at a different time and on different days over the course of the year 2019. In particular, DMSP-F18 is a geosynchronous platform, while GPM-Core is not. We can, therefore, expect their rain rate distributions to be significantly different. However, the errors between the mosaic and the three satellite-based estimates, SSMI/S non-adapted, SSMI/S adapted, and GMI, should give us robust estimates of the adaptation performances.

Table VII shows the performances of the three satellite estimators against the mosaic. As in Table V, the performances are better for the adapted data when compared to the non-adapted and are very similar to the GMI. Interestingly, the True Positive for SSMI/S adapted is slightly better than that of GMI (4.94 % vs 4.85 % resp.). It is also noticeable that the imbalance between False Alarm and Bad Detection is also similar for the three estimators. The Bad Detection starts at 13.34 % for the non-adapted and goes down to 10.80 % for the adapted and 9.81 % for the GMI. This shows that the mosaic detects rain more often than the satellite estimators about 10% of the time, and it comes as no surprise since one would expect ground-based radars to be more efficient at detecting very light rains near the surface.

Figure 8.a and .b show the scatter plots of non-adapted and adapted SSMI/S, respectively, vs the mosaic rain rates. The shades of color show the density of points from dark blue, low density, to purple, high density. For the non-adapted case, the densest area of points is clearly off the $x = y$ line with an underestimation of about 0.3 mm/hr. For the adapted case, the points are definitely closer to the $x = y$ line and are less scattered than for the non-adapted case. After the adaptation, there are also fewer outliers; even if above 1.5 mm/hr, there is still a clear signature of underestimation.

To further analyze the scatter plot, the Orthogonal Distance Regression (ODR), as proposed by **??**, is used to represent the points. This approach is chosen over the simple linear regression due to the fact that both the Mosaic rain rate and the retrieved rain rates are affected by uncertainties. The regression line for the non-adapted may appear to be closer to the identity line ($y = x$) in figure 8. However, this is misleading as the explained variance by each of these linear regressions is 0.14 before adaptation and 0.32 after. Even though the regression slope on the non-adapted case (0.81) seems closer to the x=y line, the barycentre of the histogram

TABLE VII
RAIN RETRIEVAL RESULTS WITH DRAIN ON THE MOSAIC AS THE REFERENCE.

	non-adapted SSMI/S		adapted SSMI/S		GMI	
Mosaic	Rain	Non-Rain	Rain	Non-Rain	Rain	Non-Rain
Rain	2.40%	13.34%	4.94%	10.80%	4.85 %	9.81 %
Non-Rain	2.14%	82.13%	1.49%	82.78%	0.36% %	84.98 %
F1 score on rain	0.24		0.45		0.49	
Bias on rain (mm/hr) (TP)	0.21		0.07		0.02	
Standard deviation on rain (TP)	0.33		0.21		0.33	
Number of rain pixels	2556036		2556036		6645997	

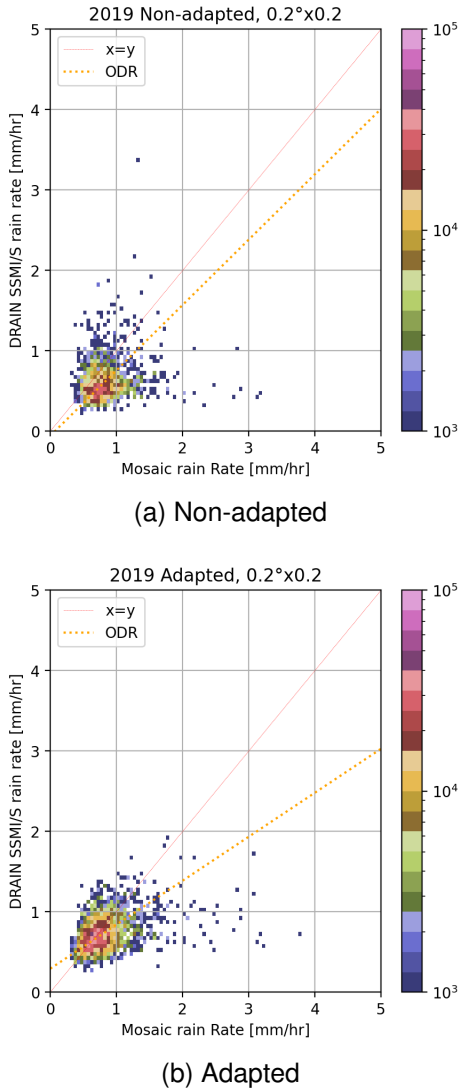


Fig. 8. (a) Scatter plot of retrieved rain from non-adapted SSMI/S vs. Météo-France Mosaic. (b) same as (a) but for adapted SSMI/S.

is off-center, and the regression is poor. The position of the regression line being closer to the identity line may be due to the wrong estimates that are coincidently almost symmetric for both the low and high rain intensity. After adaptation, the regression slope is further (0.55) from the x=y line, but the barycentre of the scattered points is spot on the x=y line, and the explained variance was multiplied by a factor of two. In addition, lower rain rates are better estimated in the adapted case.

Figure 9 show respectively the difference Mosaic-DRAIN averaged in 0.2x0.2 boxes for all the DMSP overpasses of 2019. In order to enhance the differences, only True Positive has been kept here. In the non-adapted case (figure 9.a), almost no rain is retrieved over the ocean, a lot of rain is missed along the Loire Valley, and south of the latter, a clear underestimation of the rain intensities can be seen. Over the Alps and the Pyrennees mountains, the Météo-France mosaic might be affected by larger errors due to imperfectly filtered ground clutter despite the use of the quality flag provided in the Météo-France product: only rain estimates with a confidence greater or equal 80% were kept here. The adapted case (figure 9.b), on the other hand, shows a much more balanced distribution of the differences, even if oceanic regions seem to be slightly underestimated. The northern part of France (where rain is usually dominated by frontal systems) seems to be overestimated, while the southern part of the country appears to be more underestimated, maybe due to the difference in the nature of the rain regimes as mentioned above.

VI. DISCUSSIONS

The validation process is one of the main challenges in training CycleGAN since it is meant to be used when no paired data are available. In order to get the best performance, it is important to monitor the training loss and save the best generator. Without any paired data for comparison, cycle and identity loss on a validation set determine the best epoch to select the optimal generators. In numerous trainings, the generators that ultimately give the best F1-score on the rain retrieval step are the ones in the earlier epochs (around epoch 14) rather than the later epochs. This may be due to the different properties of the two datasets, in particular, one containing images that were selected on rain-based criteria (GMI), while in contrast, the other dataset contains images that were only filtered for latitudes (SSMI/S). It is very likely that the discriminator working with the first dataset forces its generator to generate rain systematically. Data selection for a more balanced dataset requires an algorithm, either rule-based or with a Machine Learning approach, to identify rain pixels from TB alone in the SSMI/S dataset. An in-depth study of such an algorithm is important for the future application of CycleGAN or other unsupervised domain adaptation methods.

Another challenge in optimizing a CycleGAN model is that it is very sensitive to initialization. Indeed, a common weakness of generative models is that performance can vary greatly from one learning experiment to another for a given set of hyperparameters. Although choosing the best model for

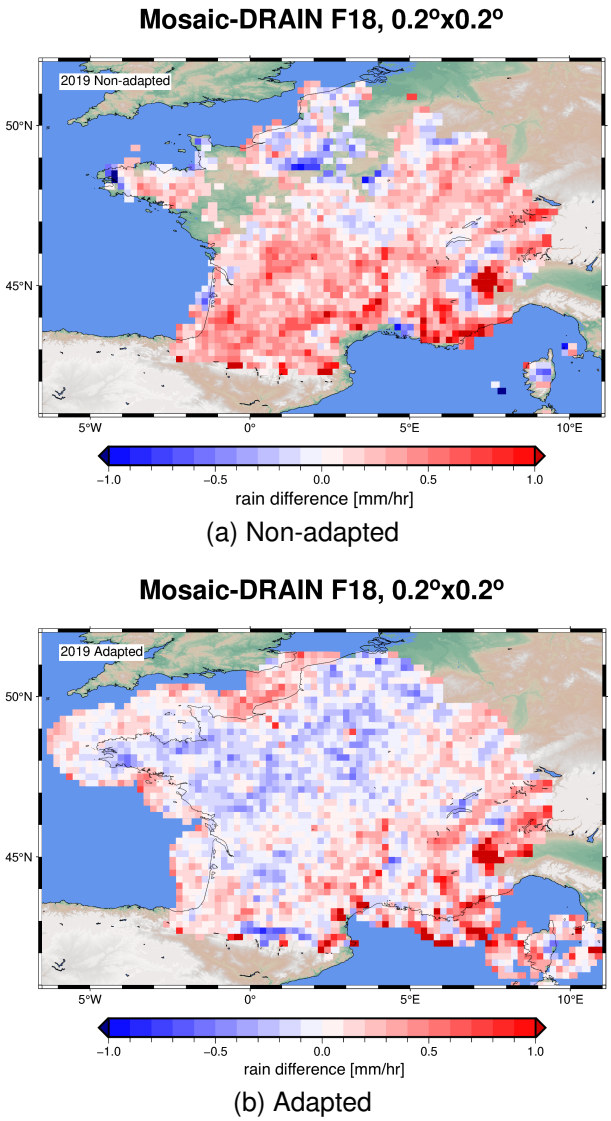


Fig. 9. Maps of the difference Mosaic-DRAIN for the True Positives. (a) for the non-adapted et (b) for the adapted. Data of all the overpasses are averaged in 0.2x0.2 boxes over the whole year 2019.

the domain adaptation between GMI and SSMI/S involved the model with the lowest validation score, as described above, the performances in terms of low rain retrieval errors also had to be taken into account. Moreover, each modification of CycleGAN's hyper-parameters leads to different performances, which means that CycleGAN has to be trained several times to select the best model.

CycleGAN is a domain-to-domain adaptation. In its application on rain retrieval for the GPM constellation, the domain of each satellite microwave imager needs to be adapted to that of the GMI as it is the only satellite that allows for the supervised training of the DRAIN algorithm. As a result, a different CycleGAN is required for each satellite in the constellation. This can lead to different rain retrieval errors for each satellite, as CycleGAN works better when the two domains are close. Nevertheless, the results presented validate the machine learning approach to domain adaptation for rain retrieval applications. In future applications, a multi-source domain adaptation approach where one model can be used for all satellites in the constellation might be a more sustainable method. For instance, an approach of domain generalization proposed by [31] using an auxiliary vision bridge domain with mappings from multiple domains. Combined with self-supervision learning techniques, it can train an effective and efficient model for generalization across multi-domain tasks. Instead of cross-domain adaptation, StarGAN is a GAN-based model capable of learning multiple domains with a single generator [32]. StarGAN comprises a discriminator and a generator. Its discriminator is trained to distinguish between real and fake images as well as predict the label of identified real images. The training of the generators contains more steps. The generator first tries to transform an input image to a target domain. Next, it transforms the fake image and its concatenated original label back to the original domain. The generator tries to fool the discriminator with fake images to train the model with adversarial loss. After optimization, the generator will have learned mapping from multiple domains.

VII. Conclusion

In this paper, CycleGAN was used to adapt the passive microwave brightness temperatures from the SSMI/S domain to the GMI domain. Ultimately, the goal is to use the adapted data as input in a deep-learning rain retrieval algorithm, DRAIN, trained on GMI data. The adaptation process is an unsupervised approach that does not rely on any paired images. One of the unique characteristics of the present study is the use of a technique originally developed for computer vision and its application to abstract pseudo-images made of geophysical measurements. Unlike RGB images, our task deals with four channels of brightness temperatures describing the condition of the atmosphere over a given scene. Multiple tests regarding hyper-parameters have thus carefully been examined to find the best configuration. This process has shown that using CycleGAN with the intention of a downstream task for the adapted images is very challenging in terms of optimization of the hyper-parameters.

However, the adapted SSMI/S data significantly increase DRAIN performances over the unadapted data. The retrieved rain field is improved both in terms of structure and intensity on the test case of coastal convection in Southeast Asia. An in-depth analysis using the whole year 2019 over three different successive references, DRAIN on GMI, DPR, and Meteo-France mosaic, also shows an improvement in rain detection and rain intensity estimation. The most noticeable improvement is reached in the comparison with the Météo-France mosaic. Although spatially limited to mainland France, the total number of overpasses is much higher than between the two satellites, providing a more robust estimate for the considered region.

The results in this paper validate the Deep Learning approach to rain retrieval from passive microwave brightness temperatures from the GPM constellation. The success of Deep Learning Unsupervised Domain Adaptation to mitigate the discrepancy in the distribution of radiometric images observed by the different satellites proves that DRAIN can be applied to the

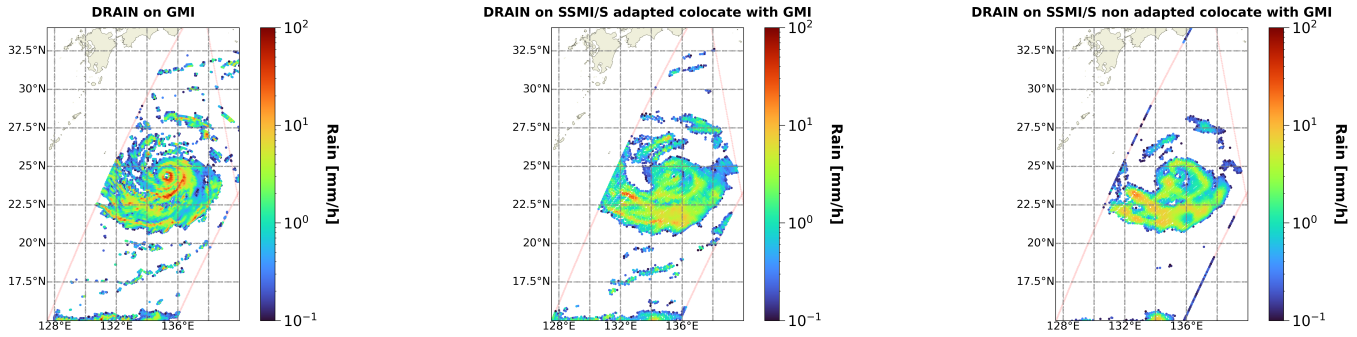


Fig. 10. Rain retrieval of an overpass on 16/09/2022. From left to right: DRAIN on GMI data, DRAIN on adapted SSMI/S, DRAIN on original SSMI/S. The rain rates are in mm/hr

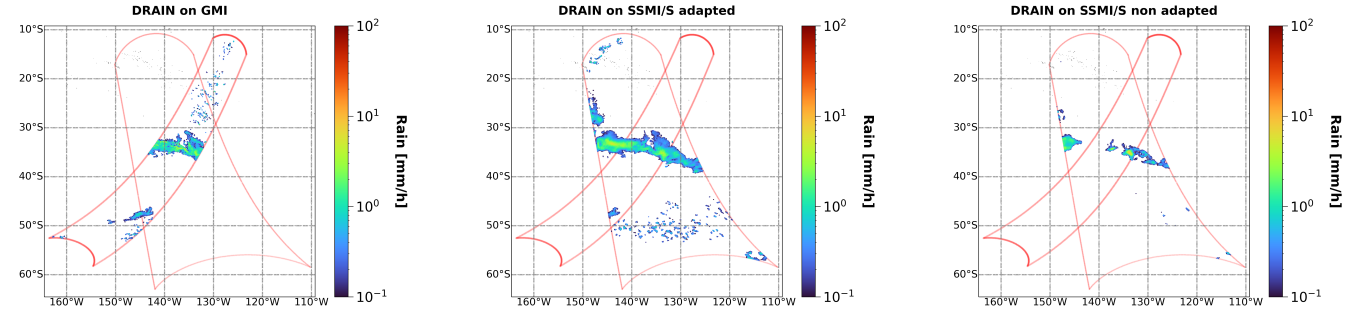


Fig. 11. Same as 01/02/2019. From left to right: DRAIN on GMI data, DRAIN on adapted SSMI/S, DRAIN on original SSMI/S. The rain rates are in mm/hr

whole constellation. However, with the difficulties in training CycleGAN described in this paper, as well as its limitation within the one-to-one domain, we believe that future methods for unsupervised domain adaptation of brightness temperatures should consider a domain generalization approach or a multi-source domain adaptation approach.

ACKNOWLEDGEMENT

This work was performed with funding from both Programme National de Télédétection Spatiale and CNES-TOSCA. The HAL GPU cluster from IPSL-ESPRI was used for the training of the CycleGAN. The GMI, DPR, and SSMI/S datasets were downloaded from NASA's Precipitation Processing System (PPS). The Météo-France datasets were downloaded from SEDOO-AERIS.

APPENDIX

As in section V-A, figure 10 and 11 show the rain retrieval by DRAIN on GMI image, adapted SSMI/S image, and original SSMI/S image on two different overpass cases. The orbit of the SSMI/S can be recognized by the larger swath. The first case, figure 10, is a cyclone between 0N-40N and 120E to 150E on the 16th of September 2022. It should be noted that there is approximately a one-hour delay between the passage of GMI and SSMI/S. The second case, figure 11, shows another rain structure in the Pacific Ocean on the 1st of February 2019. In both cases, the estimated rain structure on the adapted images shares more similarity to that of the GMI images. The difference in the first case could be due to the delay between the two passages.

REFERENCES

[1] E. A. Smith, G. Asrar, Y. Furuhama, A. Ginati, A. Mugnai, K. Nakamura, R. F. Adler, M.-D. Chou, M. Desbois, J. F. Durning, J. K. Entin, F. Einaudi, R. R. Ferraro, R. Guzzi, P. R. Houser, P. H. Hwang, T. Iguchi, P. Joe, R. Kakar, J. A. Kaye, M. Kojima, C. Kummerow, K.-S. Kuo, D. P. Lettenmaier, V. Levizzani, N. Lu, A. V. Mehta, C. Morales, P. Morel, T. Nakazawa, S. P. Neeck, K. Okamoto, R. Oki, G. Raju, J. M. Shepherd, J. Simpson, B.-J. Sohn, E. F. Stocker, W.-K. Tao, J. Testud, G. J. Tripoli, E. F. Wood, S. Yang, and W. Zhang, *International Global Precipitation Measurement (GPM) Program and Mission: An Overview*. Dordrecht: Springer Netherlands, 2007, pp. 611–653. [Online]. Available: https://doi.org/10.1007/978-1-4020-5835-6_48

[2] N. Viltard, P. Lepetit, C. Mallet, L. Barthès, and A. Martini, "Retrieving rain rates from space borne microwave sensors using u-nets," in *Climate Informatics 2020. 10th International Conference*, Sep. 2020. [Online]. Available: https://hal-insu.archives-ouvertes.fr/insu-02894942

[3] N. Viltard, V. Sambath, P. Lepetit, A. Martini, L. Barthès, and C. Mallet, "Evaluation of drain, a deep-learning approach to rain retrieval from gpm passive microwave radiometer." *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[4] C. D. Kummerow, D. L. Randel, M. Kulie, N.-Y. Wang, R. Ferraro, S. J. Munchak, and V. Petkovic, "The evolution of the goddard profiling algorithm to a fully parametric scheme," *Journal of Atmospheric and Oceanic Technology*, vol. 32, no. 12, pp. 2265 – 2280, 2015. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/32/12/jtech-d-15-0039_1.xml

[5] S. Pfreundschuh, P. J. Brown, C. D. Kummerow, P. Eriksson, and T. Norrestad, "Gprof-nn: a neural-network-based implementation of the goddard profiling algorithm," *Atmospheric Measurement Techniques*, vol. 15, no. 17, pp. 5033–5060, 2022. [Online]. Available: https://amt.copernicus.org/articles/15/5033/2022/

[6] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[7] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ser. ICML'15. JMLR.org, 2015, p. 1180–1189.

[8] X. Liu, C. Yoo, F. Xing, H. Oh, G. El Fakhri, J.-W. Kang, J. Woo *et al.*, "Deep unsupervised domain adaptation: A review of recent advances and perspectives," *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, 2022.

[9] M. Xu, M. Wu, K. Chen, C. Zhang, and J. Guo, "The eyes of the gods: A survey of unsupervised domain adaptation methods based on remote sensing data," *Remote Sensing*, vol. 14, no. 17, 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/17/4380

[10] Y. Wang and X. X. Zhu, "The sarptical dataset for joint analysis of sar and optical image in dense urban area," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 6840–6843.

[11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[12] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2066–2073.

[13] V. Sambath, N. Viltard, L. Barthès, A. Martini, and C. Mallet, "Unsupervised domain adaptation for global precipitation measurement satellite constellation using cycle generative adversarial nets," *Environmental Data Science*, vol. 1, p. e24, 2022.

[14] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[15] OSCAR. (2022) Satellite: Gpm core observatory. World Meteorological Organization. [Online]. Available: https://space.oscar.wmo.int/satellites/view/gpm_core_observatory

[16] ——. (2022) Instrument: Ssmis. World Meteorological Organization. [Online]. Available: https://space.oscar.wmo.int/instruments/view/ssmis

[17] P. Bauer, L. Schanz, and L. Roberti, "Correction of three-dimensional effects for passive microwave remote sensing of convective clouds," *Journal of Applied Meteorology and Climatology*, vol. 37, no. 12, pp. 1619–1632, 1998.

[18] Y. Hong, J. L. Haferman, W. S. Olson, and C. D. Kummerow, "Microwave brightness temperatures from tilted convective systems," *Journal of Applied Meteorology*, vol. 39, no. 7, pp. 983–998, 2000.

[19] S.-W. Kim, D.-B. Shin, and Y. Choi, "Effects of the three-dimensional hydrometeor distributions of precipitating clouds on passive microwave rainfall estimations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 1957–1966, 2016.

[20] C. Guilloteau and E. Foufoula-Georgiou, "Beyond the pixel: Using patterns and multiscale spatial information to improve the retrieval of precipitation from spaceborne passive microwave imagers," *Journal of Atmospheric and Oceanic Technology*, vol. 37, no. 9, pp. 1571 – 1591, 2020. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/37/9/jtechD190067.xml

[21] G. Skofronick-Jackson, W. Berg, C. Kidd, D. B. Kirschbaum, W. A. Petersen, G. J. Huffman, and Y. N. Takayabu, "Global precipitation measurement (gpm): Unified precipitation estimation from space," in *Remote Sensing of Clouds and Precipitation*. Springer, 2018, pp. 175–193.

[22] C. Mallet, C. Klapisz, and N. Viltard, "Effects of heterogeneous precipitating atmospheres on simulated brightness temperatures," in *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*. CRC Press, 1999, pp. 291–298.

[23] D. Harris and E. Foufoula-Georgiou, "Subgrid variability and stochastic downscaling of modeled clouds: Effects on radiative transfer computations for rainfall retrieval," *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D10, pp. 10 349–10 362, 2001.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[25] J. Figueras i Ventura and P. Tabary, "The new french operational polarimetric radar rainfall rate product," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 8, pp. 1817–1835, 2013.

[26] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[27] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.

[28] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," *arXiv preprint arXiv:1802.05957*, 2018.

[29] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *International conference on machine learning*. PMLR, 2019, pp. 7354–7363.

[30] J. L. Bentley, "Multidimensional binary search trees used for associative searching," vol. 18, no. 9, 1975. [Online]. Available: https://doi.org/10.1145/361002.361007

[31] S. Harary, E. Schwartz, A. Arbelle, P. Staar, S. Abu-Hussein, E. Amrani, R. Herzig, A. Alfassy, R. Giryes, H. Kuehne *et al.*, "Unsupervised domain generalization by learning a bridge across domains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5280–5290.

[32] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "Stargan: Unified generative adversarial networks for multi-domain image-to-image translation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8789–8797.