


Deep-Learning-Based Approach in Imaging Radiometry by Aperture Synthesis: An Alias-Free Method

Richard Faucheron , Eric Anterrieu , Louise Yu , Ali Khazaal ,
and Nemesio J. Rodríguez-Fernández , *Member, IEEE*

Abstract—A new approach based on deep learning methods is presented for reconstructing L -band brightness temperature images from the inversion of interferometric data, namely, complex visibilities here simulated from observations of the Soil Moisture and Ocean Salinity (SMOS) interferometric radiometer. A specific deep neural network (DNN) architecture composed of a fully connected layer followed by a contracting and expansive path is proposed to learn the relationship between the simulated visibilities and the brightness temperature maps. The performances of the DNN are compared with those of algebraic inversions based on Fourier theory, which are all affected by strong aliases in the synthesized field of view (FOV), as a consequence of the spacing between the elementary antennas of SMOS, which does not satisfy the Nyquist criteria. In the alias-free FOV (AFFOV) of the algebraic reconstructions, these latter are outperformed by the DNN reconstructions: average mean absolute error (MAE) of about 0.7 K for the DNN instead of 3.7 K. Outside the AFFOV of the algebraic reconstructions, the DNN reconstructions do not show significant signs of field aliasing although the MAE increases: average MAE of about 1.5 K in the whole FOV. An analysis of the role of different neurons in the hidden layers is presented, and it is shown that some neurons are specialized in reconstructing what corresponds roughly to the AFFOV region of algebraic approaches, while other neurons are specialized in dealing with the external regions of this AFFOV.

Index Terms—Aperture synthesis, deep learning (DL), imaging radiometry, inverse problem.

I. INTRODUCTION

WITHIN the frame of imaging radiometry by aperture synthesis with arrays of elementary antennas, the problem of retrieving the brightness temperature distribution of an observed scene from the corresponding visibility samples has

Manuscript received 20 October 2023; revised 22 January 2024 and 29 February 2024; accepted 3 March 2024. Date of publication 6 March 2024; date of current version 20 March 2024. This work was supported by the Centre National d'Etudes Spatiales through a Research and Technology program. (Corresponding author: Richard Faucheron.)

Richard Faucheron, Eric Anterrieu, Ali Khazaal, and Nemesio J. Rodríguez-Fernández are with the Centre d'Etudes Spatiales de la Biosphère, Université de Toulouse, Centre National D'Etudes Spatiales, Centre National de la Recherche Scientifique, Institut de Recherche pour le Développement, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, 31401 Toulouse, France (e-mail: richard.faucheron@univ-tlse3.fr; eric.anterrieu@cnrs.fr; nemesio.rodriguez-fernandez@cnrs.fr).

Louise Yu is with the Centre National D'Etudes Spatiales, 31401 Toulouse, France (e-mail: louise.yu@cnes.fr).

Digital Object Identifier 10.1109/JSTARS.2024.3373875

been widely addressed essentially with algebraic approaches. It has been demonstrated that this inverse problem is an ill-posed one, which has to be regularized in order to provide a unique and stable solution. Moreover, depending on the geometry of the antenna array and on the spacing between the elementary antennas as well, the part of the field of view (FOV) useful for scientific purposes can be significantly reduced by the presence of aliasing artifacts.

During the last decade, deep learning (DL) has become an essential paradigm in signal and image processing [1]. However, very few examples of DL-based methods applied to aperture synthesis problems in imaging radiometry can be found in the literature [2], [3], [4]. These works are further presented in Section II. The current study aims at presenting a new imaging approach that relies on a deep neural network (DNN) to learn inverse mapping between the input visibility samples and the brightness temperature distribution of the observed scene. On the one hand, it does not share any element with the algebraic paradigm that led to the regularized methods found so far in the literature [5], as on the mathematics side, there is no explicit regularization principle [6]. On the imaging side, there is no flavor of Fourier synthesis [7]. On the other hand, this new approach proposes a distinct DNN architecture from previous articles, as it addresses the specific configuration of the Soil Moisture and Ocean Salinity (SMOS) mission [8], [9], regarding spectral content, dynamic range, and underconstrained quality of the inverse problem. The presented DNN has been trained for inverting visibility samples provided by a sparse array of 69 equally spaced elementary antennas working at low frequency such as the interferometric array of the SMOS mission.

SMOS was launched in November 2009, and for more than 14 years, it has provided accurate L -band brightness temperature maps with a spatial resolution between ~ 25 and ~ 60 km, depending on the position in the FOV. These maps have been used for retrieving surface soil moisture [10] as well as sea surface salinity [11]. Fig. 1 shows the interferometric array of the single payload of SMOS. It is a Y-shaped array populated with 69 equally spaced elementary antennas operating in the protected band 1400–1427 MHz.

The DL-based reconstruction approach presented in this article has been compared to the algebraic method commonly used for inverting the visibility samples provided by SMOS. The first results obtained over a wide variety of scenes

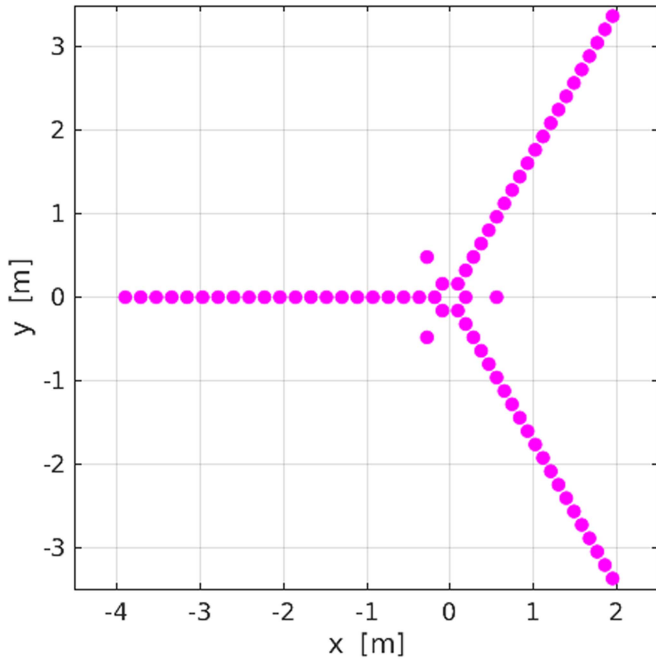


Fig. 1. Interferometric array of SMOS is populated with 69 elementary antennas spaced every $d \simeq 18.6$ cm along the three arms of a Y. According to the central frequency of observation in the protected L-band from 1400 to 1427 MHz, the shortest spacing between the antennas normalized to the central wavelength of observation d/λ is here equal to $= 0.875$.

demonstrate reconstructed brightness temperature maps free from any aliasing artifact. Moreover, the quality of the retrieved brightness temperatures shows a reduction of the mean absolute error (MAE) of 60% and over a wider useful synthesized FOV than that available with the algebraic approach. Compared to the results of the three aforementioned articles, the proposed DNN demonstrates equivalent or superior performance in terms of reconstruction error metrics.

The rest of this article is organized as follows. The relationship between the complex visibility samples and the brightness temperature distribution of the observed scene is recalled in Section II, which is also devoted to the algebraic regularized reconstruction method chosen by the European Space Agency for SMOS and to the existing DL-based reconstruction approaches. Section III describes how pairs of brightness temperature maps and complex visibility samples have been generated, as well as how these data have been split into three subsets for training, validating, and testing. Section IV is devoted to the architecture of the DNN as well as to the design of the training and of the evaluation steps. Within the frame of the SMOS mission, Section V illustrates the performances of the DL-based approach with emphasis on the absence of aliasing artifacts. The performances of the DL-based approach are compared with those of the algebraic one in Section VI. Finally, Section VII concludes this article.

II. PROBLEM FORMULATION

The measurement equation of imaging radiometry by aperture synthesis is recalled in this section together with the algebraic

image reconstruction and statistical image reconstruction approaches using the DNN.

A. Measurement Equation

The cross correlation of the signals measured by any pair on elementary antennas of an interferometric array provides a sample of the visibility function. Without polarimetric considerations, the complex visibility V_{pq} for a pair of antennas \mathcal{A}_p and \mathcal{A}_q is given by [12]

$$V_{pq} = \frac{1}{\sqrt{\Omega_p \Omega_q}} \int_{\|\xi\| \leq 1} \int \mathcal{F}_p(\xi) \mathcal{F}_q^*(\xi) \tilde{r}_{pq}(-t) (T_b(\xi) - T_{\text{rec}}) e^{-2j\pi \frac{\mathbf{b}_{pq} \cdot \xi}{\lambda}} \frac{d\xi}{\sqrt{1 - \|\xi\|^2}} \quad (1)$$

where $\mathbf{b}_{pq} \equiv \mathbf{r}_q - \mathbf{r}_p$ is the baseline vector associated with the two antennas \mathcal{A}_p and \mathcal{A}_q located in \mathbf{r}_p and in \mathbf{r}_q , respectively, the components $\xi_1 = \sin \theta \cos \phi$, $\xi_2 = \sin \theta \sin \phi$, and $\xi_3 = \cos \theta$ of the angular position variable ξ are direction cosines in the reference frame of the array, θ and ϕ are the traditional spherical coordinates (the colatitude and the azimuth, respectively), $\mathcal{F}_p(\xi)$ and $\mathcal{F}_q(\xi)$ are the normalized voltage patterns of the two antennas with equivalent solid angles Ω_p and Ω_q , $T_b(\xi)$ is the brightness temperature distribution of the scene under observation, T_{rec} is the physical temperature of the receivers (assumed to be the same for all the receivers), $\tilde{r}_{pq}(t)$ is the fringe-washing function, which accounts for spatial decorrelation effects with $t = (\mathbf{b}_{pq} \cdot \xi)/c$ the spatial delay, and $\lambda = c/f$ is the central wavelength of observation. After discretization of the double integral over an appropriate sampling grid in the direction cosine domain, the relationship between the complex visibilities V and the brightness temperature distribution T_b of the scene under observation can be written in the linear algebraic form as follows:

$$V = \mathbf{G}T \quad (2)$$

where \mathbf{G} is the linear modeling matrix of the instrument.

B. Algebraic Reconstruction

Referring back to the direct problem (2), the inverse problem aims at retrieving $T = T_b$ (the constant T_{rec} is canceled out from the visibilities with the aid of the response to a flat target [13], whatever the method used for the inversion [6]).

As the number of unknowns (the number of pixels in T) might be larger than the number of equations (the number of samples of V), this linear system might be underconstrained with an infinite number of solutions for T . This is precisely the case of SMOS. On the one hand, the cross correlation of the signals measured by the 69 elementary antennas of the Y-shaped interferometric array, as depicted in Fig. 1, provides $\frac{1}{2} \times 69 \times (69 - 1) = 2346$ complex visibility samples. Thanks to the Hermitian character of the Fourier transform of a real function, other 2346 complex visibilities can be estimated as the conjugates of those actually measured by the total number of independent real quantities remain 4692. In addition, three antennas out of 69 are also used to measure the total flux of

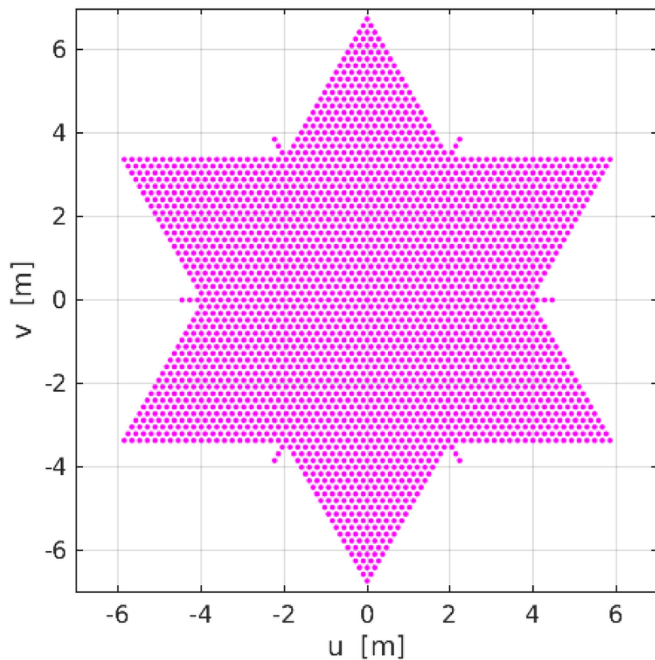


Fig. 2. Experimental frequency coverage of SMOS. Although the number of antenna pairs is equal to $69 \times (69 - 1) = 4692$, only 2790 of them are unique as a consequence of redundancies.

the image, leading to a total of 4695 measurements. Due to the redundancies of some baselines, the number of Fourier frequencies that are sampled is smaller than the number of visibilities. In the case of SMOS, as illustrated in Fig. 2, the number of Fourier frequencies in the star-shaped frequency coverage is equal to 2791, including the zero-frequency one.

On the other hand, when discretizing the brightness temperature domain with appropriate hexagonal sampling grids satisfying the Shannon criterion, it turns out that the total number of pixels in the FOV in front of any antenna is equal to 34 087, of which only 23 042 point to the earth, while the 11 045 remaining ones point to the sky. However, because of the spacing between the elementary antennas of the interferometric array, reconstructed maps obtained with such a Fourier-based method are suffering from aliasing artifacts. The extent of the alias-free FOV (AFFOV) is shown in Fig. 3, which also shows the extended AFFOV (EAFFOV) that is free from earth aliases but not from aliasing of the sky emission. In the case of SMOS, the number of pixels in the EAFFOV is only 8981, but the inverse problem of retrieving the brightness temperature over those pixels from 4695 measurements remains ill-posed, and it has to be regularized in order to provide a unique solution. The regularization used for SMOS [7] is referring to a physical concept: the limited resolution of the instrument. This approach aims at retrieving the brightness temperature map, which has its Fourier transform confined to the experimental frequency coverage of the antenna array. Although the number of sampling points can be reduced, according to [14], it is more comfortable to keep this number as a power of 2 while expecting the Shannon sampling criterion, as illustrated in [15]. As a consequence, rephrasing the inverse problem in terms of retrieving 2791

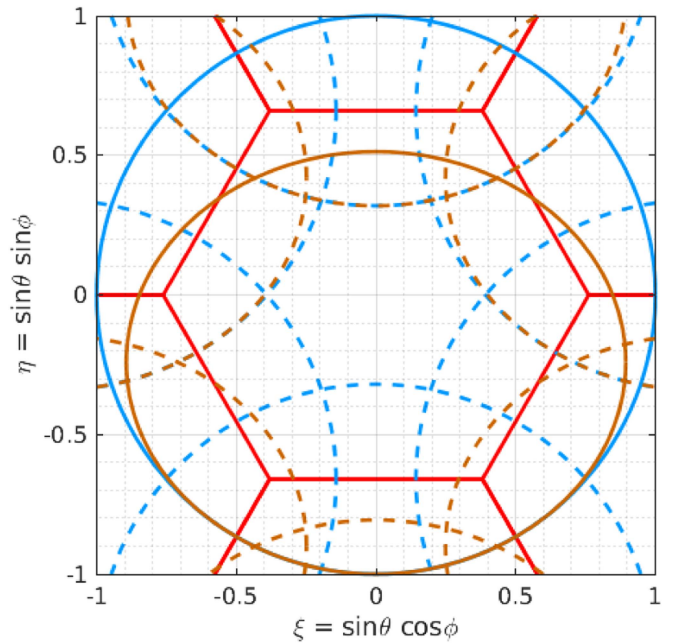


Fig. 3. FOV synthesized by SMOS is subject to aliasing because of the spacing between the elementary antennas of the Y-shaped array, as shown in Fig. 1, where $d = 0.875\lambda$. The solid ellipse (maroon) is the limit of the earthly horizon as seen from an elevation of 755 km when the platform is tilted off Nadir from 31.2° and the solid circle (blue) is the limit of the front space seen by every antenna. The hexagonal FOV is drawn with its neighbors (red) responsible for the six alias counterparts of the earthly horizon (dashed maroon) and of the sky (blue dashed). The AFFOV is that part of the FOV exempt from sky and earth aliases; the EAFFOV is free from only earth aliases.

unknown Fourier components, instead of 8981 unknown values of the brightness temperature, from 4695 measurements, leads now to an overconstrained problem.

C. DNN Reconstruction

Statistical approaches have also been discussed in the literature as an alternative for imaging by aperture synthesis.

An alternative data-driven approach has first been conducted by Zhang et al. [2] to reconstruct microwave synthetic aperture interferometric radiometer (SAIR) images from interferometric measurements with the aid of a deep convolutional neural network (CNN) framework as a decoding model. The CNN was trained with 55 000 images from ImageNet [16] labeled as “geo” and “plant.” The reconstructed images showed an improvement of the root-mean-square error (RMSE) performances from 16.1% to 8.2% with normalized measurement bias of 0.05% compared with those obtained with the conventional fast Fourier transform. In addition, visual inspections showed some noise suppression as well. The simulated instrument is a microwave SAIR equipped with a compact array of small elementary antennas working at high frequency.

In 2022, Xiao et al. [3] proposed a deep CNN that learns mapping between interferometric measurements and nonuniform synthetic aperture radiometer images. The dataset is made of 10 000 processed examples from the UC Merced dataset [17]. The UC Merced dataset is a 21-class land use image dataset, with classes such as “agricultural,” “forest,” and “tennis court.” The

visibility samples were generated from a program simulating a randomly distributed nonuniform antenna array populated with 51 elementary antennas operating at 33.5 GHz, with a maximum baseline of $50\lambda \simeq 45$ cm. The dimension of the processed T maps and the number of visibility samples V leads to a highly underconstrained problem. The authors quantitatively compared the performance of their proposed CNN with both the array factor forming method [18] and the grid method [19]. When evaluated on a testing subset, the CNN achieved an average RMSE value of 9.47 K, while the array factor forming method and the grid method recorded average RMSE values of 14.88 and 15.33 K, respectively.

More recently, Dou et al. [4] have presented a CNN that takes the samples of 1-D visibility functions to reconstruct brightness temperature profiles. Their CNN is trained, validated, and tested over 21 000 pairs of T/V , distributed in the proportions of 70%, 20%, and 10%, respectively. The brightness temperature images are processed from the UC Merced data, and the corresponding visibility function are simulated from a program reproducing the antenna array configuration of a 1-D airborne microwave interferometric radiometer operating in the C -band [4]. Their approach yields to a better image quality, lower sidelobes and smaller Gibbs oscillation than the brightness temperature profiles reconstructed by the inverse fast Fourier transform method. It also shows smaller RMSE, 0.597 K, than that of the DNN proposed by Zhang et al. [2], which is 1.611 K.

These three efforts illustrate that DNNs are a promising approach to solve inverse problems in synthetic aperture imaging radiometry with interferometric arrays and that they should deserve more attention. This study presents a DNN trained and tested over a dataset of 166 557 samples representative of interferometric L -band data, with regard to spectral content, dynamic range, and underconstrained quality of the inverse problem (4695 real-valued measurements to reconstruct 23 042 image pixels belonging to the earth, i.e., a problem underdetermined by a factor close to 5). As of the modeling instrument, we used a sparse array of equally spaced elementary antennas working at low frequency, such as the interferometric array of the SMOS mission

Referring back to the previous studies, the network described by Zhang et al. [2] has twice as many input values than output pixels. The CNN proposed by Dou et al. [4] is also trained with twice as much input values than output pixels (202 real-valued measurements to reconstruct 101 brightness temperatures). In contrast, the network described by Xiao et al. [3] is highly underconstrained since it produces output images with a number of pixels larger than the number of input interferometric data by a factor of 11.6.

III. REPRESENTATIVE DATASET

This section describes how brightness temperature distributions T and visibility samples V have been simulated, as well as how these T/V pairs have been split into three subsets for training, validating, and testing the DNN reconstruction.

A. Brightness Temperatures

Dual-polarization views of the earth at instrument level have been simulated from monthly SMOS Level 3 brightness

temperatures in H and V polarizations at ground level [20], using the highest spatial resolution available, ranging between ~ 25 and ~ 60 km depending on the position in the FOV. In order to keep the size of the dataset to a reasonable size while accounting for seasonal effects, one snapshot every 12 s for the first three days of every 12 months of year 2012 have been simulated by interpolating the multi-incidence brightness temperatures sampled over a regular geographical grid covering the whole earth. This tridimensional interpolation takes place at each sampling node of the hexagonal grid in the direction cosine domain, for which the corresponding latitude, longitude, and incidence angle are calculated with the aid of the telemetry data. T_H and T_V distributions are, thus, obtained at ground level, and a rotation angle is applied for calculating the corresponding T_X and T_Y distributions at instrument level. As mentioned in Section II, these brightness temperature maps can be expressed as vectors of 34 087 real-valued components, from which 23 042 belong to the earth and 11 045 to the sky.

B. Visibility Samples

The instrument shown in Fig. 1 has been used for simulating the visibility samples V_X and V_Y corresponding to every T_X and T_Y distributions according to (1). The modeling instrument is configured with different antenna element patterns, from the MIRAS database, used in the operational processing of SMOS. Referring back again to SMOS conditions in Section II, these visibility samples are vectors of 4695 real-valued components.

The dataset contained between 27 572 and 27 927 T/V pairs each month, totaling 332 791 pairs for the year 2012, in both X and Y polarizations: 166 557 for ascending passes and 166 234 for descending ones. Assuming that each real-valued component is coded with 8 bytes, this dataset requires about 140 GB of memory, regardless of any additional data necessary for processing it. In this work, only the 166 557 data from the ascending passes with polarizations X are used during the training, validation, and testing phases.

C. Splitting Method

The dataset described in the previous section has been split into three subsets.

- 1) 60% of the dataset has been dedicated to the *training subset* of the DNN, during which it is exercised to the relation between V and T .
- 2) 20% of the dataset has been used for the *validation subset*, during which DNN learning is monitored on T/V pairs not used for the training.
- 3) 20% have been devoted to the *testing subset*, which aims at comparing the performances of the DL-based reconstruction to those of the algebraic one, again on not previously used T/V pairs.

The main challenge in machine learning is to avoid overtraining. The model must perform well on new and never-seen inputs, not just on those on which it has been trained. This is why two different methods have been tested for splitting the dataset into these three subsets: a random split and a k -means split. Whatever the method, the proportions between the three subsets are maintained to the aforementioned values.

For the first approach, T/V pairs are randomly chosen in the dataset. An example of such a distribution is shown in Fig. 4(a). As the pairs have been simulated every 12 s, successive FOVs may overlap. As a consequence, similarity between the T/V pairs of the training and testing phases cannot be excluded. This could facilitate the prediction of brightness temperature maps from visibility samples of the testing subset and thus increasing the model performances. This is why another method of splitting has been tested in order to reduce that risk by limiting the overlapping between FOVs of T/V pairs used in the training phase and those of the testing subset.

With the k -means method [21], T/V pairs are distributed into subsets that are as independent as possible with respect to a geographical distance. To this end, clusters of snapshots have been created using the k -means algorithm, with respect to the latitude and longitude of each snapshot. The number of clusters has been fixed to 200, so that it is big enough to maintain the distribution of the T/V pairs with respect to the snapshot ocean/land composition of the k -means split method as close as possible to the distribution used for the random split method. At the same time, it is small enough to limit the number of overlapping snapshots as much as possible. An example of the resulting clusters is displayed in Fig. 4(b). Then, training, validation, and testing subsets have been randomly selected in the same proportions as before, among the clusters. An example of such a distribution can be seen in Fig. 4(c). Consequently, the only images that can overlap between the training and testing subsets are those located in the periphery of the clusters, which greatly reduces the occurrence of similar T/V pairs between the training and testing phases.

IV. METHODS: DNN

This section describes the proposed architecture, the framework used for the training and the methodology used to evaluate the DNN reconstruction performances, which includes the comparison with the algebraic approach.

A. Architecture of the DNN

The proposed DNN architecture is shown in Fig. 5. It results from several trials and is the one that achieved the best reconstruction performances in preliminary tests. It is composed of a fully connected layer followed by a contracting and expansive path with residual connections for a total of 336 million parameters.

The first layer is fully connected to the input visibility vector and creates a first estimation of the brightness temperature distribution by mapping the relationship from the input to the output space. This is the only layer that must cope with the problem's underconstraint, going from 4695 visibilities to 23 042 estimated brightness temperatures; therefore, its output is noisy. This estimation is then reshaped into a 2-D sparse regular grid so that each predicted temperature is set to its geographic place in the FOV, and the spatial information about the neighborhood of each pixel is, therefore, introduced accordingly.

The following contracting path is a succession of convolution layers, each followed by a rectified linear unit (ReLU) and a

max-pooling operation that reduces the dimensionality of the input representation by keeping only the maximum value of a sliding window of size 2×2 . The number of feature channels is doubled at each downsampling step. The expansive path consists of the repeated application of transposed convolutions with stride 2 to upsample the feature maps, followed by convolution layers and a ReLU operation. The number of channels is halved at each upsampling step. The convolutional layers from the contracting and expanding paths extract high-level features from the data and force the image to be represented sparsely in the convolutional feature space. At this stage, the layers are no longer subject to the subconstrained nature of the problem. Their task is to build a mask of 23 042 pixels that will be added to the first hidden layer output (also 23 042 pixels) for the purpose of correcting reconstruction imprecision like sharpening edges or smoothing homogeneous areas.

B. Training

During the training stage, batches of complex visibilities are given as input to the network, from which it must compute the corresponding brightness temperature maps. For each batch, the error between the estimated T_e and the true T maps is measured and retropropagated through the network by gradient descent, to update the model weights.

The objective function to minimize is the mean squared error, and the optimizer is the RMSprop algorithm [22] set with the following parameters: the learning rate $\gamma = 10^{-4}$, the smoothing constant $\alpha = 0.99$, the weight decay $\lambda = 0$, and the momentum factor $\mu = 0.9$. The model has been trained over 100 epochs with mini-batch sizes set to 32, implemented in the Python programming language using the PyTorch library version 1.8 [23]. The training of the network is performed on the train subset described in Section III. The validation subset is used to monitor the training of the network on data that have not been used to update its parameters.

The training time of the proposed network on the 99 774 training T/V pairs and the 33 257 validation pairs (random splitting case) over 100 epochs is 182 h (≈ 7.6 days) on two TESLA V100 SXM2 32-GB GPUs working in parallel. During the testing phase, the network weights are not updated. In this case, the resulting reconstruction time of a single snapshot is reduced to about 114 ms on a Tesla V100 SXM2 32-GB GPU and less than 20 ms on a NVIDIA A100 SXM4 40-GB GPU. In comparison, the mean inversion time of the algebraic method is only of a few microseconds [24], and it does not require prior training. Nevertheless, the computation time would not be a bottleneck for an operational implementation since it is still much less than the integration time (1.2 s for SMOS) and the computing time allocated to this phase of the ground processing.

The work done for the present study is entirely reproducible. The python environment and the package used for splitting the data, training, and evaluating the model have been conducted under a seed value that allows us to reproduce every result. Finally, operations on GPUs have been run deterministically so that the order of execution is guaranteed and can be reproduced, at the cost of some additional computational time.

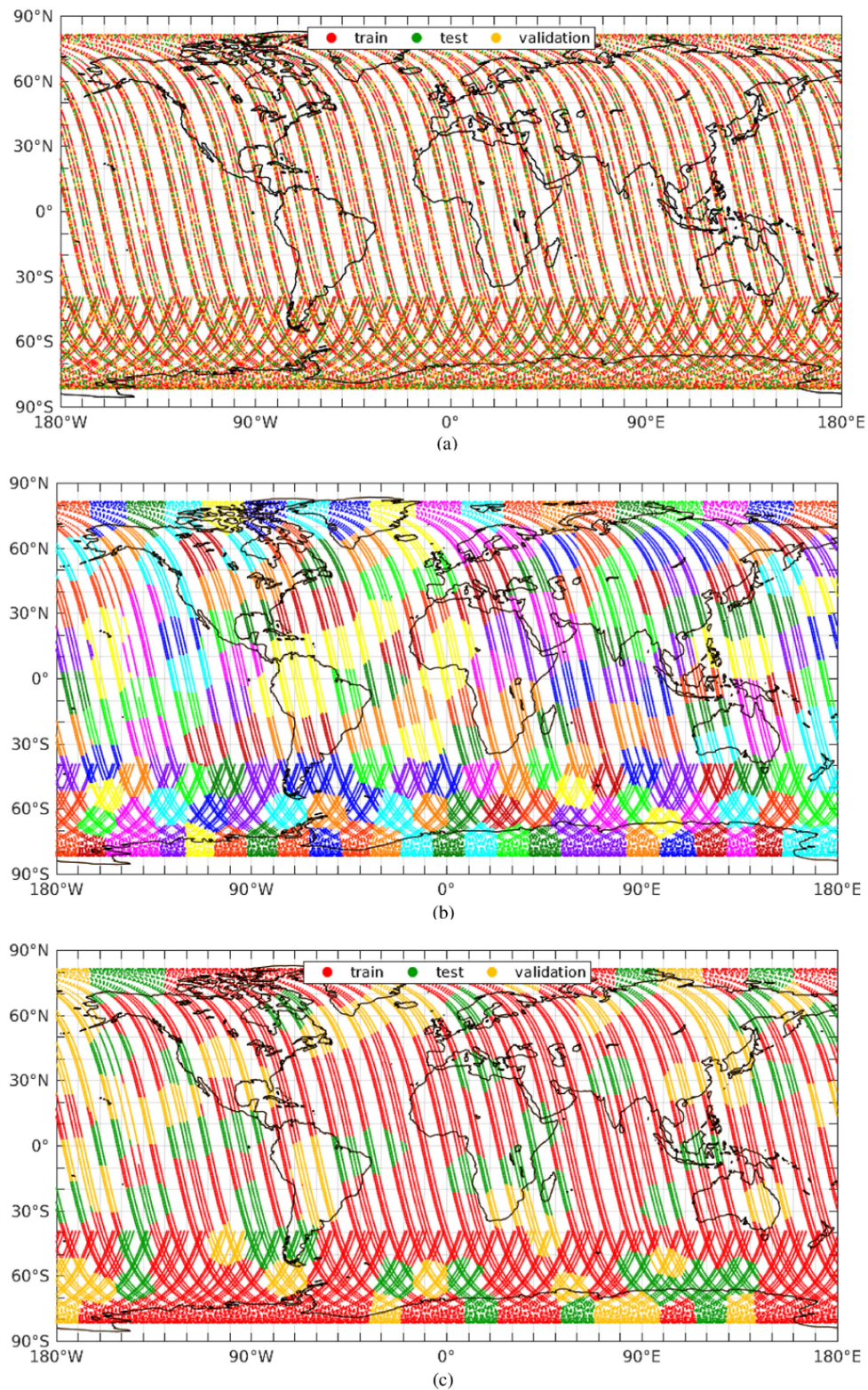


Fig. 4. Distribution of the snapshots into training, validating and testing subsets, illustrated on 25% of the total amount of data for the sake of readability. Panel (a) presents the snapshot distribution with a random sampling. Panel (b) presents the snapshot distribution into 200 clusters using the k -means algorithm. Panel (c) presents the snapshot distribution with a random selection of the k -means clusters. In both splitting methods, 60% of the data are used for training, 20% for validating, and the last 20% for testing. (a) Random splitting of the snapshots. (b) k -means clustering of the snapshots. (c) k -means splitting of the snapshots.

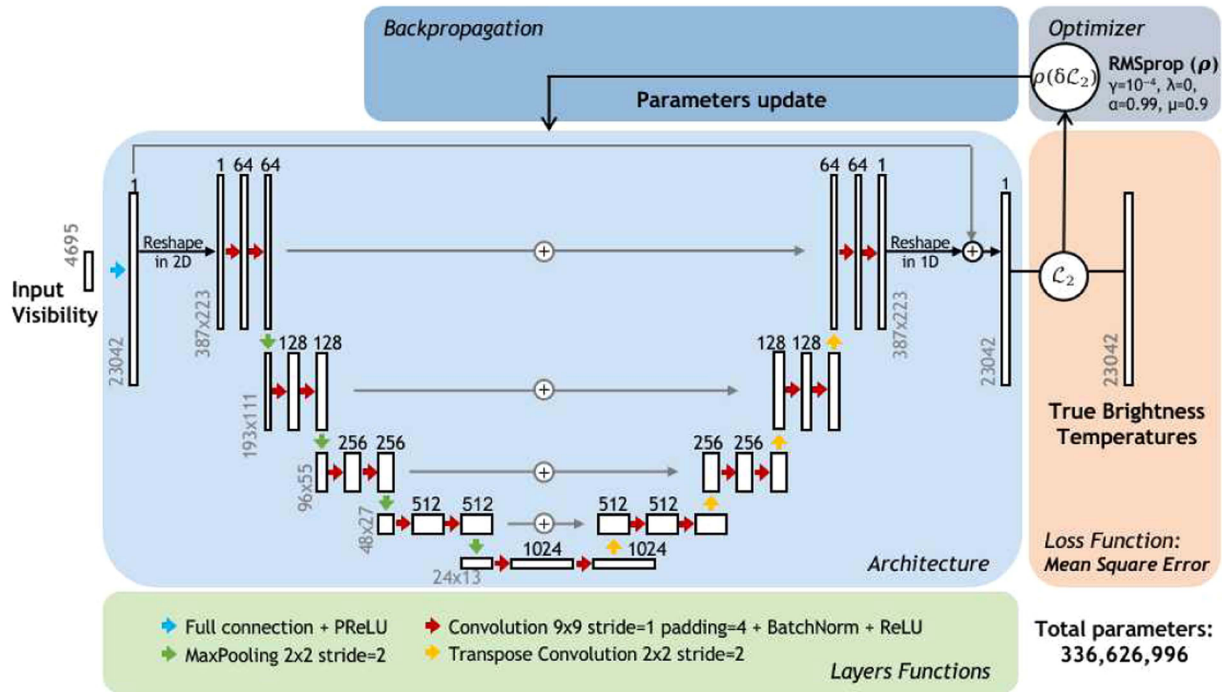


Fig. 5. Scheme of the proposed DNN architecture.

C. Evaluation

In order to evaluate the DNN performances, metrics are measured on the testing subset created from the remaining 20% of the dataset that was used neither for training nor for validation.

The complex visibilities of the testing subset were inverted with the algebraic regularized approach as well as with the proposed data-driven DNN approach. As explained in Section II, the algebraic method only reconstructs brightness temperatures in the EAFFOV whose shape is like a distorted hexagon [25]. However, it turns out that the proposed DNN is able to reconstruct brightness temperatures over the entire FOV. As a consequence, the DNN metrics are computed over the entire FOV to grasp its reconstruction performances, as well as only over the EAFFOV to allow a comparison with the algebraic method.

The error metrics used to compare the two approaches are the MAE, the standard deviation of error (STD), and the RMSE of the initial brightness temperature maps with respect to the reconstructed maps. Every values are in kelvin. In addition, the error metrics used to compare the proposed approach to previous work are the average normalized root-mean-square error (NRMSE-A) and the min-max normalized root-mean-square error (NRMSE-M). Both the metrics are expressed as percentages. The NRMSE-A is the RMSE that has been divided by the average value of every true T maps of the testing subset, while the NRMSE-M is the RMSE divided by the difference between the average maximum value of the true T maps of the testing subset and the average minimum value.

V. RESULTS

This section is devoted to the analysis of the reconstruction performances of the proposed DNN over the testing subset of

data. A qualitative description is made from two reconstructed maps. Then, a quantitative description is presented using error metrics averaged over the full testing subset.

A. Qualitative Description of Two Scenes

Figs. 6 and 7 show two representative reconstructions of brightness temperature maps. Figs. 6(a) and 7(a) display the brightness temperature of reference to retrieve. Figs. 6(b) and 7(b) show the DNN reconstruction, and Figs. 6(c) and 7(c) show the algebraic performance. Figs. 6(d) and 7(d) display the DNN reconstruction error map (that is the difference map between the DNN reconstruction and the reference map) over the full scene, Figs. 6(e) and 7(e) show the DNN error map over the EAFFOV area, and, finally, Figs. 6(f) and 7(f) display the algebraic method error map (only measured on the EAFFOV, as explained in Section II).

Fig. 6(b) displays a DNN reconstruction whose MAE is representative of the average MAE measured on the testing subset, on the entire reconstructed field, so a DNN MAE is equal to 1.5 K. Fig. 7(b) displays a DNN reconstruction whose MAE is representative of the average MAE measured on the testing subset, only over the EAFFOV reconstructed field, so a DNN EAFFOV MAE is equal to 0.7 K.

The comparison of the panels *a* and *b* of Figs. 6 and 7 shows that the DNN can reproduce the full extent of the input scene without aliasing (23 042 pixels), while the algebraic method [see Figs. 6(c) and 7(c)] reconstructs a smaller FOV with important aliasing effects (8981 pixels). The total number of reconstructed T s provided by the DNN is 2.5 times greater.

Figs. 6(d) and 7(d) show that the higher errors are concentrated on coastlines, and more particularly on coastlines that lie on the periphery of the images. Continental surface and ocean areas

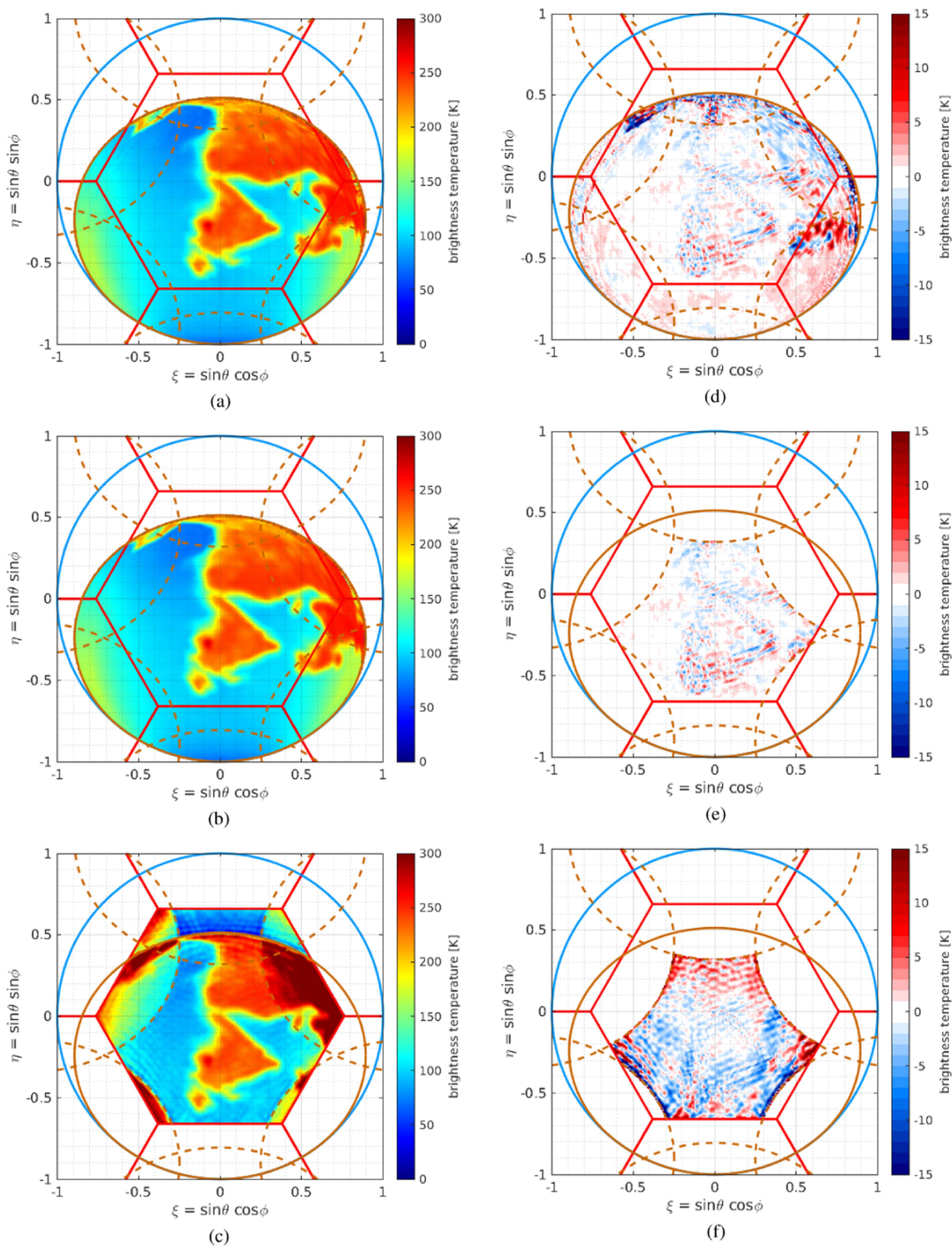


Fig. 6. Reconstruction example of our proposed DNN compared with the algebraic regularized approach. (a) Ground truth brightness temperature (BT) map calculated by the SMOS operational ground segment. (b) Inversion of the visibilities by the DNN. (c) Inversion of the visibilities by the algebraic method. (d) and (e) Difference map between the DNN reconstruction and the reference BT map, over the full reconstructed field and over the EAFFOV area, respectively. (f) Reconstructions error maps of the algebraic method on the EAFFOV. DNN MAE: 1.5 K; DNN (EAFFOV) MAE: 1.0 K; and algebraic MAE: 3.6 K. (a) Reference BT map. (b) DNN reconstruction. (c) Algebraic reconstruction. (d) DNN - Reference. (e) DNN - Reference (EAFFOV). (f) Algebraic - Reference (EAFFOV).

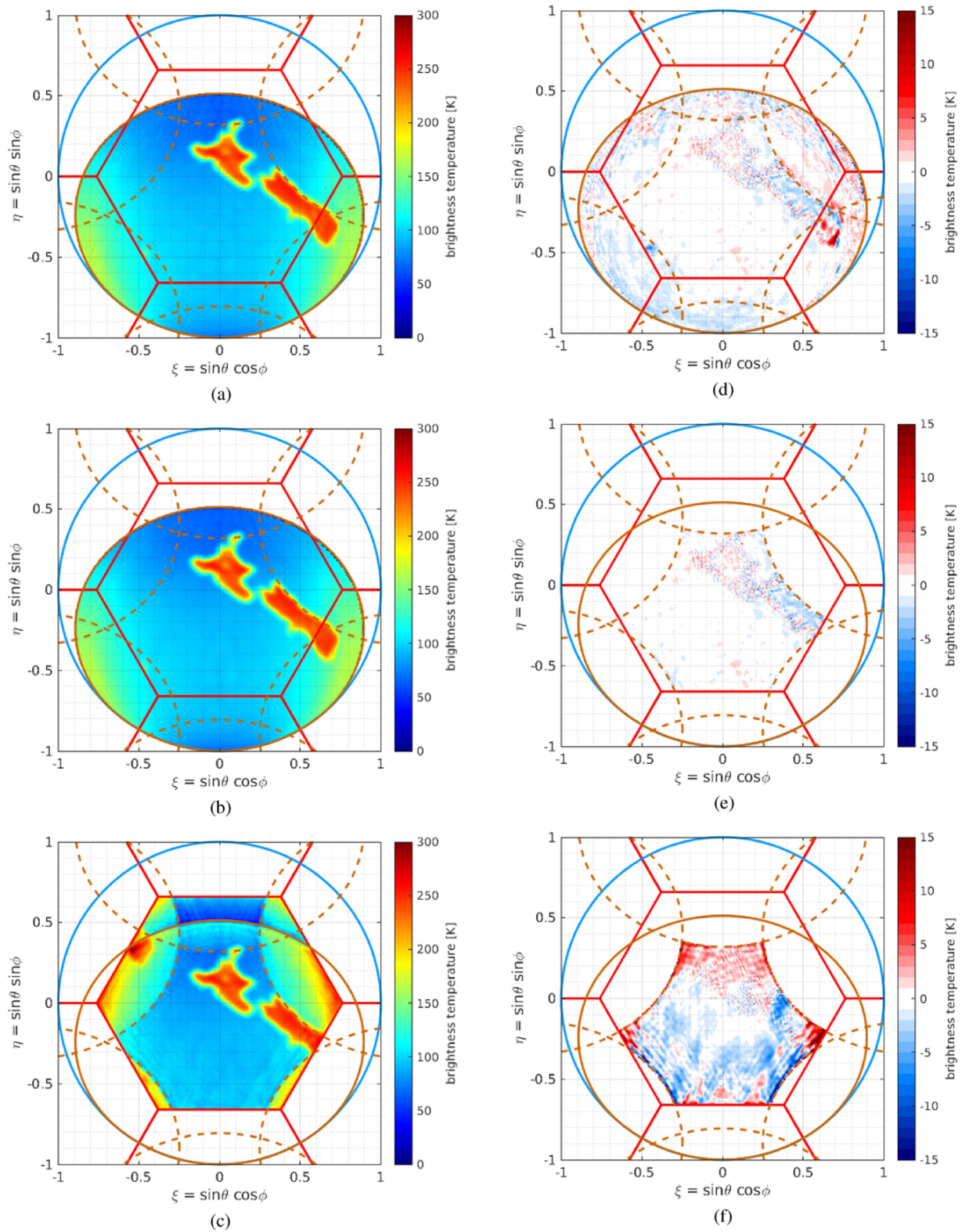


Fig. 7. Reconstruction example of our proposed DNN compared with the algebraic regularized approach. (a) Ground truth brightness temperature (BT) map calculated by the SMOS operational ground segment. (b) Inversion of the visibilities by the DNN. (c) Inversion of the visibilities by the algebraic method. (d) and (e) Difference map between the DNN reconstruction and the reference BT map, over the full reconstructed field and over the EAFFOV area, respectively. (f) Reconstructions error maps of the algebraic method on the EAFFOV. DNN MAE: 0.9 K; DNN (EAFFOV) MAE: 0.7 K; and algebraic MAE: 4.3 K. (a) Reference BT map. (b) DNN reconstruction. (c) Algebraic reconstruction. (d) DNN - Reference. (e) DNN - Reference (EAFFOV). (f) Algebraic - Reference (EAFFOV).

TABLE I
AVERAGE RECONSTRUCTION ERROR METRICS FOR THE ALGEBRAIC APPROACH (ONLY IN THE EAFFOV) AND FOR THE DNN ONE (IN THE EAFFOV AND IN THE FULL FOV)

	MAE	STD	RMSE	NRMSE-A	NRMSE-M
DNN (FOV)	1.53	2.88	2.90	1.80%	1.35%
DNN (EAFFOV)	0.70	0.88	0.98	0.61%	0.46%
Algebraic (EAFFOV)	3.75	7.75	7.78	4.84%	3.63%

Metrics are averaged over the entire testing subset. Quantities are given in Kelvin.

are otherwise well reconstructed anywhere else. The algebraic method does not seem to be influenced by the coastal lines, since no error pattern can be seen on the ocean/land transitions.

In addition, the comparison of the three error maps of Figs. 6 and 7 shows that the DNN reconstruction error within the full reconstructed field as well as in the EAFFOV is lower and more homogeneous than that of the algebraic method [see Figs. 6(f) and 7(f)].

B. Global Quantitative Results

The global performances of the DNN approach and of the algebraic method are given in Table I, where the previous errors metrics are reported for the testing subset in the EAFFOV area as well as in the entire earth FOV.

Table I shows that the reconstructions of the DNN in the EAFFOV area present an MAE of 0.7 K, which is twice as low as the MAE measured over the entire reconstructed field. Likewise, both the STD and the RMSE are about three times smaller in the EAFFOV than in the earth FOV. As a consequence, the quality of the DNN reconstructions is obviously better in the EAFFOV than in any other area of the reconstructed FOV. This confirms the qualitative observation made in the previous subsection that the DNN reconstructions present higher errors outside the EAFFOV.

When comparing the error metrics in the EAFFOV of the brightness temperature maps reconstructed by the DNN approach with those synthesized by the algebraic method in the same area, it turns out that all those of the DNN are always smaller. Considering the DNN reconstruction in the earth FOV, the algebraic approach still presents higher errors overall.

In order to take into account the actual dispersion of the quality metrics for all the studied scenes of the testing subset, histograms were computed. Fig. 8 displays the DNN's reconstruction MAE distribution on the full scene [see Fig. 8(a)] and on the EAF-FOV [see Fig. 8(b)], while the lower panel shows the algebraic method reconstruction MAE distribution on the EAFFOV [see Fig. 8(c)]. In addition, the metrics distribution was studied for different ocean/land ratios of the snapshots. The ocean/land ratio is divided into four categories: "Full Ocean" are snapshots comprising more than 90% of ocean pixels, "Full Land" are snapshots with more than 90% of land, "Mostly Ocean" are composed of 50–90% of ocean, and "Mostly Land" are composed of 50–90% of land. Under each histogram, four boxplots indicate the MAE percentiles for each ocean/land composition and are colored accordingly: dark blue for the category "Full Ocean,"

light blue for "Mostly Ocean," yellow for "Mostly Land," and green for "Full Land." The fifth boxplot, colored in red, indicates the average percentiles over every snapshots. For each boxplot, the upmost left and the upmost right ticks display the 10th and 90th percentiles of the MAE distribution, respectively. The left- and right-hand sides of the boxes display the 25th and 75th percentiles, respectively, and the bold vertical tick in the boxes displays the 50th percentile (median).

Looking at the global boxplot, colored in red in Fig. 8, it appears that the median on the full reconstructed field is below the average MAE of 1.5 K, so most of the testing T/V pairs are reconstructed with an averaged MAE below 1.5 K on the full scene. Also, the 90th percentile displayed by the furthest right vertical tick of the global boxplot [see Fig. 8(b)] shows that 90% of the snapshots are reconstructed with an MAE less than 1.1 K on the EAFFOV. This value is to be compared to the mean MAE of 1.5 K of the full reconstruction [see Fig. 8(a)]. The EAFFOV area is, thus, better reconstructed by the DNN than the rest of the scene. Finally, the error distribution of the EAFFOV reconstruction is less spread out than on the full FOV, where 50% of the reconstructions error are ranging between 0.4 and 0.8 K for the EAFFOV, against a range between 0.8 and 1.9 K for the full area. Looking now at the four boxplots relative to the ocean/land composition of the scenes, it appears that every category yields smaller reconstruction error and smaller error range in the EAFFOV than in the full area. Fig. 8 shows that the DNN reconstructions are better on pure ocean scenes than on mixed ocean/land and pure land snapshots, on both the EAFFOV area and the full FOV. This is expected as those are rather homogeneous scenes, which are easier to predict.

Comparing the DNN algorithm with the algebraic method, it appears that both the approaches have shown a larger error reconstructing mixed ocean/land and pure land snapshots, due to the higher complexity of those scenes compared to pure ocean snapshots. However, the DNN exhibits less dispersion in the error distribution than the algebraic method for the ocean/land ratio distribution. Focusing on the EAFFOV, 50% of the DNN MAE ranges from about 0.4 to 0.8 K, while the algebraic MAE ranges from 2.5 K to over 4.7 K. For the DNN, 90% of the estimated T_e are reconstructed with an MAE less than 2.7 K on the full reconstructed field, and with an MAE less than 1.1 K on the EAFFOV. Those values are compared to the 25th percentile of the algebraic method, displayed by the left-hand side of the fifth boxplot of Fig. 8(c), showing that 75% of the T maps are reconstructed with an MAE superior to 2.5 K. Looking now

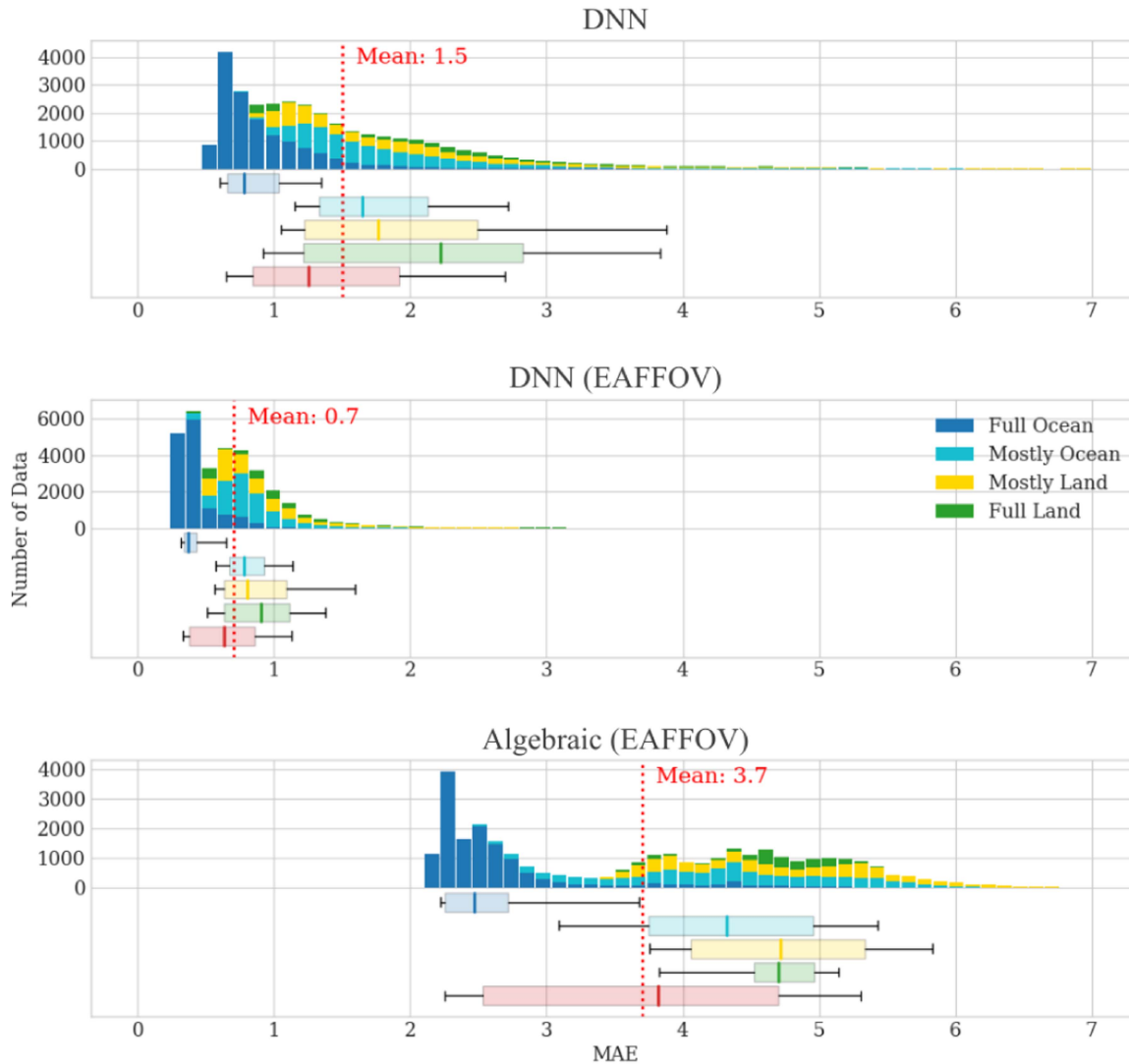


Fig. 8. Comparison of the reconstructions MAE distribution of the proposed DNN and the algebraic method, over the testing subset, with a *random split* of data, colored by ocean/land snapshot composition. Measured on the EAFFOV area for the algebraic approach and on the EAFFOV and complete areas for the DNN. “Full Ocean” are snapshots comprising more than 90% of ocean pixels, “Full Land” are snapshots with more than 90% of land, “Mostly Ocean” are composed between 50% and 90% of ocean, and “Mostly Land” are composed between 50% and 90% of land. The boxplots indicate the percentiles for each ocean/land composition and are colored accordingly. The last boxplot, colored in red, indicates the average percentiles over every snapshots. For each boxplot, the upmost left and the upmost right ticks display the 10th and 90th percentiles, respectively. The left- and right-hand sides of the boxes display the 25th and 75th percentiles, respectively. The bold vertical tick in the boxes displays the 50th percentile. The mean value of the MAE is displayed for each distribution by a vertical red dashed line.

at the four boxplots relative to the ocean/land composition of the scenes, it appears that only the MAE distribution of the full land reconstructions is about the same range between the DNN reconstructions over the EAFFOV and the algebraic reconstruction over the EAFFOV. Otherwise, every other category of ocean/land composition presents higher error range and higher error value when reconstructed by the algebraic method.

C. Results as a Function of Data Splitting

Fig. 9 displays the distribution of the MAE over the testing subsets for the DNN trained with a *k*-means split of the data (see Section IV). The color, legend, and disposition of the panels are the same as the ones detailed for Fig. 8. Using the *k*-means split, the proportion of Full Ocean scene in the testing subset is slightly lower than with the random split.

Fig. 9(a) shows that the MAE on the full reconstructed field is equal to 1.7 K when the model is trained on the *k*-means split subset, which is higher by 0.2 K than when trained on the random split subset [see Fig. 8(a)]. Also, 50% of the MAE distribution on the *k*-means split subset ranges from 0.9 to 1.7 K, with a 90th percentile at 3.6 K. This is slightly narrower than 50% of the MAE distribution on the random split subset, which ranges from 0.8 to 1.9 K, with a 90th percentile at 2.7 K. Looking now at the four boxplots relative to the ocean/land composition of the scenes, it appears that the distribution of the MAE over the full ocean scenes is similar. However, mostly ocean and mostly land categories present overall smaller errors when the model is trained on the *k*-means split subset, but with more extreme high error reconstructions. On the other hand, the full land scenes show higher error and higher error range when reconstructed by the model trained on

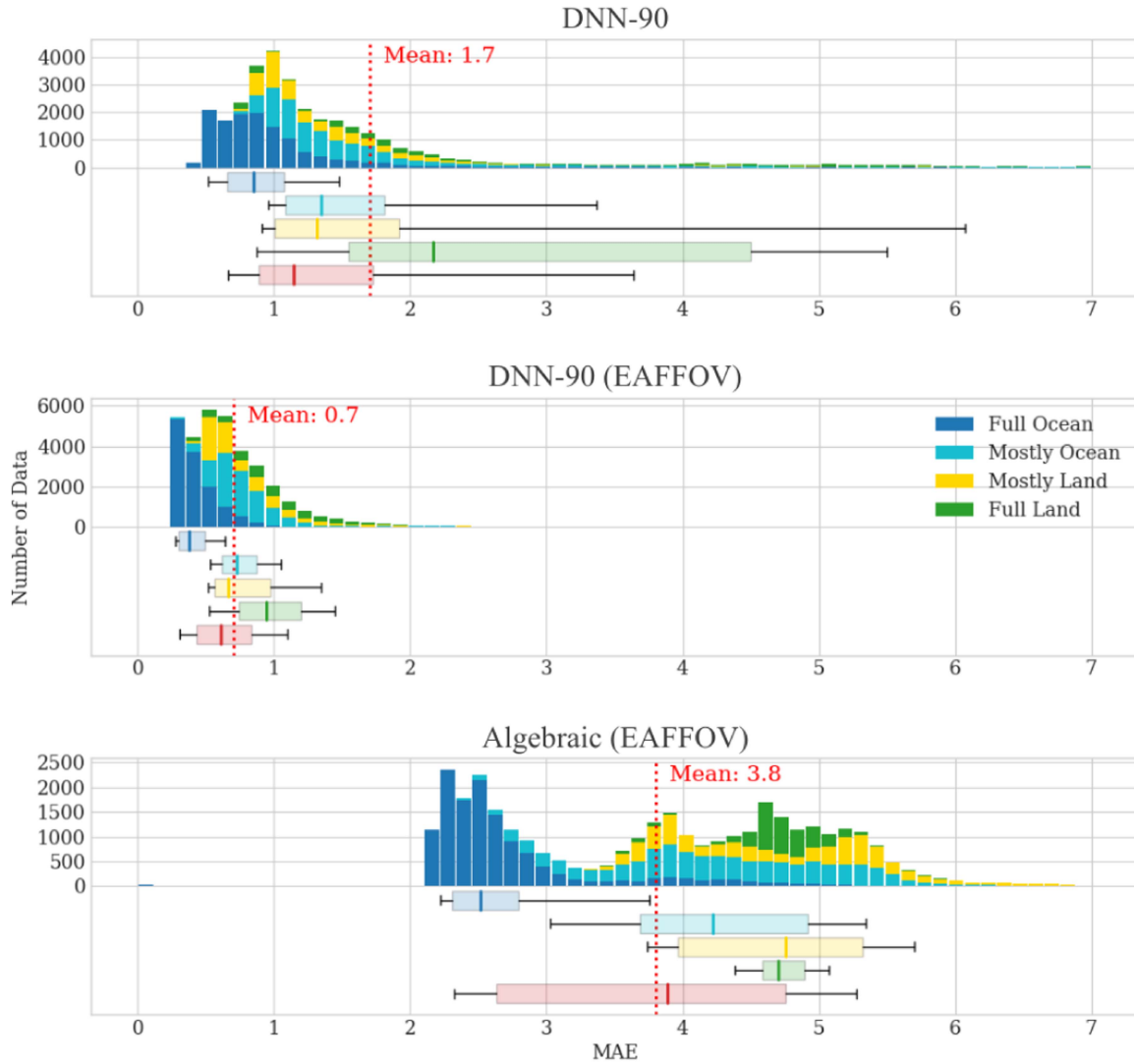


Fig. 9. Comparison of the reconstructions MAE distribution of the proposed DNN and the algebraic method, over the testing subset, with a k -means split of the data, colored by ocean/land snapshot composition. Measured on the EAFFOV area for the algebraic approach and on the EAFFOV and complete areas for the DNN. “Full Ocean” are snapshots comprising more than 90% of ocean pixels, “Full Land” are snapshots with more than 90% of land, “Mostly Ocean” are composed between 50% and 90% of ocean, and “Mostly Land” are composed between 50% and 90% of land. The boxplots indicate the percentiles for each ocean/land composition and are colored accordingly. The last boxplot, colored in red, indicates the average percentiles over every snapshots. For each boxplot, the upmost left and the upmost right ticks display the 10th and 90th percentiles, respectively. The left- and right-hand sides of the boxes display the 25th and 75th percentiles, respectively. The bold vertical tick in the boxes displays the 50th percentile. The mean value of the MAE is displayed for each distribution by a vertical red dashed line.

the k -means split subset and also present more extreme high error.

The disparity in performance becomes less significant when assessing the reconstruction inside the EAFFOV. In this context, both data splitting methods yield similar results, with an average MAE of 0.7 K and a common 90th percentile value of 1.1 K. Moreover, when inspecting the boxplots relative to the ocean/land compositions of the scenes, minimal differences are observable, both in terms of MAE value and MAE range.

The random split method shows lower mean reconstruction errors overall and much more less extreme high errors. However, it is nuanced by its higher MAE for mostly ocean and mostly land snapshots. Concerning the EAFFOV area, both splitting

methods present similar performances. The random split of the data allows the DNN to encounter more training T/V pairs close to the testing data, better preparing it to reconstruct those testing data. In this way, the DNN is more capable of extrapolating the T located outside the EAFFOV, reducing the occurrence of extreme errors in this area.

VI. DISCUSSION

In this section, the results of the current study are discussed in comparison to previous data-driven works and to the algebraic method. Then, some insights on the understandability of the proposed DNN and its ability to prevent aliasing artifacts are provided.

A. Comparison to Previous Works

The trained CNN of Xiao et al. (IASR-CNN), in [3], performed an average RMSE of 9.47 K measured over a testing subset of 1000 samples. Depending on the level of the Gaussian noise added to these ideal visibilities used for training as well as testing, the averaged RMSE was reported to increase up to 14 K. Referring back to Table I, as a comparison, the present DNN performs an averaged RMSE of 2.90 K on the full reconstructed FOV with realistic noiseless SMOS visibilities, as reported in Section III.

The CNN proposed by Zhang et al. [2] (SAIR-CNN) has been evaluated with different undersampling ratios of the complex visibility samples versus the image pixels to reconstruct. An RMSE of 3.6% is obtained with a 40% undersampling ratio, whereas with a 10% undersampling ratio, the RMSE is reported to be as large as 24.6%. As a comparison, the NRMSE-A of the DNN presented in the current study, that is the RMSE calculated as a percentage of the true T maps average value, turns out to be about 1.8% over the testing subset with an undersampling ratio of about 20% typical of the SMOS satellite (4695/23042).

The 1-D CNN developed by Dou et al. [4] (AS-CNN) has been tested over their 2100 testing data and achieved an average RMSE of 0.597 K, which is lower than the RMSE obtained with their implementation of the SAIR-CNN: 1.611 K on the same testing dataset. Dou et al. [4] specify in their article that the brightness temperature of the original scenes of the testing dataset is comprised between an average minimum of 59.44 K and a maximum of 179.61 K. The corresponding NRMSE-M of the AS-CNN is thus equal to $0.597/(179.61 - 59.44) \simeq 0.005 = 0.50\%$. Referring back to Table I, as a comparison, our proposed DNN performs an NRMSE-M of 0.46% on the EAFFOV and an NRMSE-M of 1.35 K on the full reconstructed FOV, over the brightness temperatures of our testing subset comprised between an average minimum of 64.08 K and a maximum of 278.64 K. The AS-CNN being trained in over-constrained conditions, its achieved NRMSE-M is most likely to be compared to the results obtained by our proposed approach within the EAFFOV.

B. Comparison of the DNN to the Algebraic Method

When comparing the reconstructions of the proposed DNN method and the algebraic method only in the EAFFOV, the DNN obviously has a better accuracy, with an average MAE of 0.70 K over the testing subset, against 3.75 K for the algebraic method over the same dataset. From the resolution point of view, when looking over the multitude of snapshots we have processed for this article, as soon as a small island appears in the FOV, and it is often the case, we have observed in the corresponding reconstructions the same features as those noticed with the algebraic inversions. This suggests that the synthetic impulse response is the same, if it is not better. Moreover, upon examining the panels a and b of Figs. 6 and 7, it appears that the spatial resolution of the DNN reconstructions closely corresponds to that of the nominal SMOS. This is totally expected since the DNN training has been conducted with realistic L -band brightness temperature maps computed from the highest spatial resolution global

L -band observations currently available, those provided by SMOS.

When considering the full reconstructed field of the DNN method (AFFOV plus EAFFOV), the measured error metrics are still higher for the algebraic method than for the DNN. Indeed, the average RMSE of the DNN measured on the entire scenes of the testing subset is about 2.90 K, whereas it is about 7.78 K for the algebraic reconstructions measured in the EAFFOV only. Moreover, the algebraic approach cannot reconstruct T outside the EAFFOV; thus, the DNN presents the advantage of reconstructing a much wider FOV, growing from 8981 pixels to 23 042 pixels, representing 2.5 times more retrieved temperatures.

Considering the reconstruction of the DNN method over the full reconstructed field, we see that a significant number of scenes composed only of ocean are reconstructed within the same error range than reconstructions over mixed scene and pure land, between an MAE of 1 and 1.5 K. On the other hand, the reconstructions of the algebraic method stand in two distinct ranges of error. The pure ocean scenes are reconstructed within an MAE range of 2–3 K, while the other types of scene present an MAE between 3.5 and 6 K. This comparison suggests that the proposed DNN is more robust to scene compositions of sea fraction, since the reconstruction error is less correlated to the ocean/land ratio of the scenes.

C. Understandability of the DNN

It has been shown in Section V that there are little signs of aliasing in the images reconstructed with the DNN method. However, in principle, an interferometric array whose antennas are spaced by more than half the wavelength of the observation should show aliasing in the reconstructed images as the Shannon criterion is not fulfilled. Therefore, it is important to understand the processes involved in the proposed DNN, which does not produce images with aliases folding.

To get further insight into this, we first visually analyzed DNN reconstruction error maps that present the higher errors. Fig. 10 presents such a reconstruction. Fig. 10(a) displays the brightness temperature map of reference, Fig. 10(b) displays the corresponding DNN reconstruction, and Fig. 10(c) shows the corresponding error map. In the latter, we can see two distinct areas defined from two different error scales. The first area presents high errors all over the periphery of the image, while lower errors draw an area in the center of the image with a hexagonal like shape. We notice that this hexagonal shape is remarkably similar to the EAFFOV of the algebraic approach. This behavior is in fact globally attested by the MAE metric visible in Table I, where the DNN MAE over the EAFFOV is of 0.70 K, against a higher MAE of 1.53 K over the full reconstructed field.

An investigation of the hidden layers among the deepest in the network shows feature maps with neurons that activates particularly in or out the same hexagonal-like shape. Two examples of those feature maps are shown in Fig. 11.

This behavior could be the evidence that the DNN is able to learn the relationship between T and V to such an extent

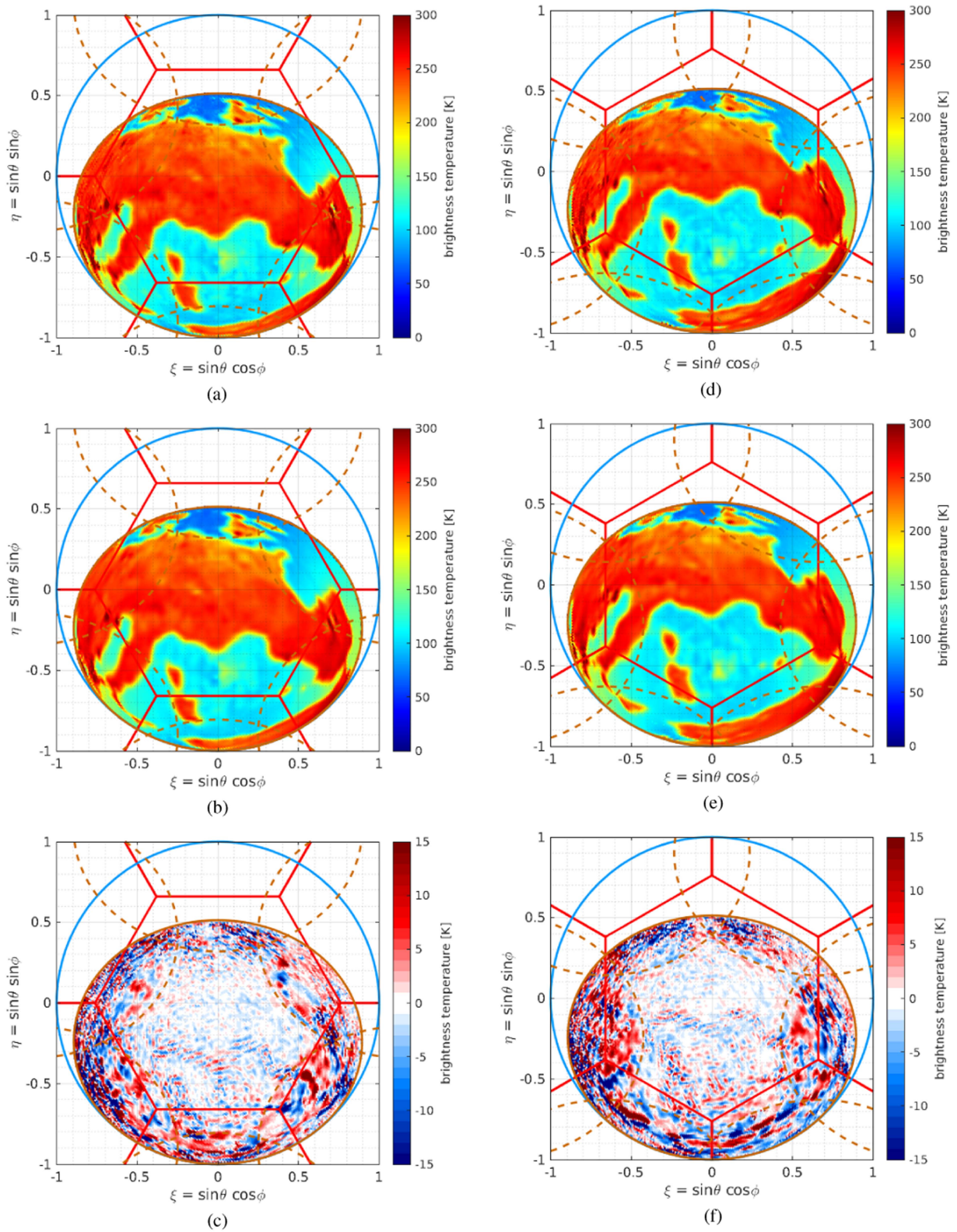


Fig. 10. Example of the DNN reconstruction on visibilities taken from the modeled instrument with the characteristics of the SMOS satellite [(a)–(c)] and on visibilities taken from the same instrument rotated by 90° [(d)–(f)]. (a) and (d) Ground truth brightness temperature map. (b) DNN reconstruction. (e) DNN-90 reconstruction. (c) DNN - Reference. (f) DNN-90 reconstruction error map. (a) Reference BT map from a SMOS-like instrument. (b) DNN reconstruction. (c) DNN - Reference. (d) Reference BT map from a SMOS-like instrument rotated by 90° . (e) DNN-90 reconstruction. (f) DNN-90 - Reference.

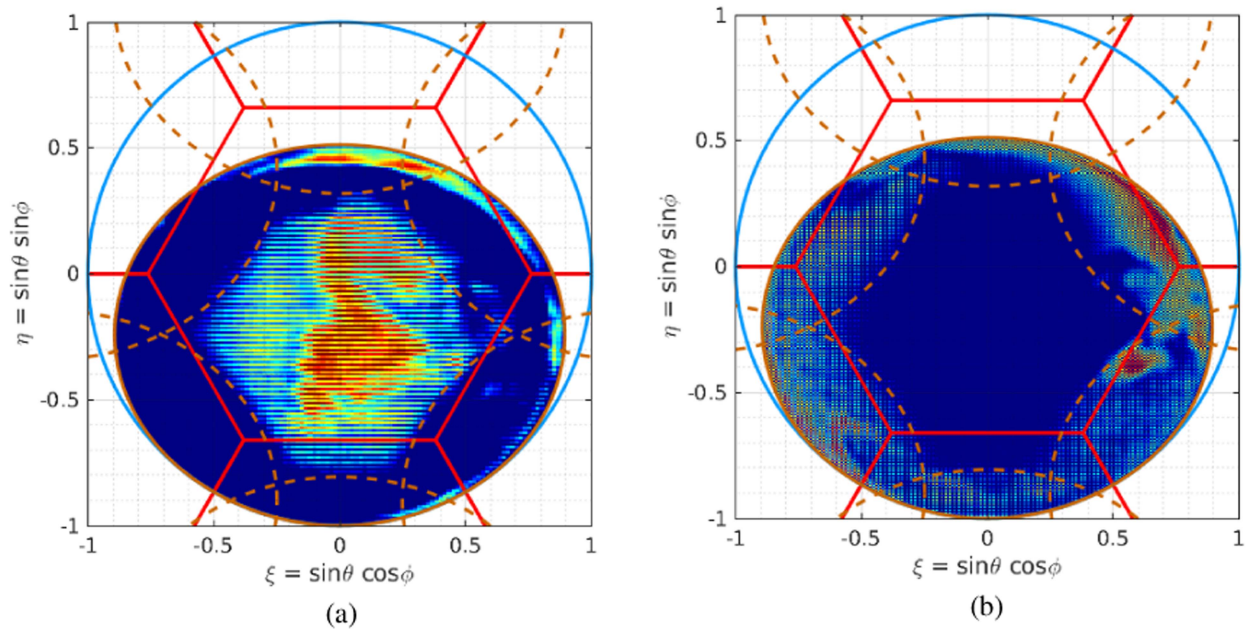


Fig. 11. Example of DNN feature maps. (a) Feature map where the neurons activate in the EAFFOV from the 24th hidden layer. (b) Feature map where the neurons activate outside the EAFFOV from the 26th hidden layer.

that it makes the distinction between areas where the brightness temperature can be accurately retrieved from the accessible visibilities (the AFFOV) and areas where the instrumental sampling would produce folding that the DNN has to learn to extrapolate (or actually *interpolate* in the spatial frequencies space). Therefore, these feature maps suggest that the two areas are processed by the DNN separately.

In order to ensure that the EAFFOV hexagon is not an artifact coming from the fact that the brightness temperatures used as input are coming from algebraic reconstructions of SMOS measurements, an additional experiment was performed. A new set of visibilities has been computed from the same brightness temperature maps and with the same instrument, except that the modeled instrument was rotated by 90° . Therefore, the EAFFOV of the algebraic reconstruction should have another shape in the center area of the image, as displayed by the dashed maroon line in Fig. 10(f). This rotated EAFFOV will be referred to as the EAFFOV-90. The proposed DNN has been trained over this new set of visibilities taken from a 90° -rotated instrument and the same brightness temperature maps, under the same conditions of training. The proposed DNN trained in the context of this experiment is called DNN-90.

An example of reconstruction from DNN-90 is compared to a reconstruction from the original DNN in Fig. 10, where Fig. 10(d) shows the reference T , Fig. 10(e) shows the DNN-90 reconstruction, and Fig. 10(f) shows the corresponding error map, with the new rotated EAFFOV-90 superposed in dashed maroon line. The two reconstructed scenes are geographically close, so their complexity is similar. The figure shows that the area highlighted by the lower errors in the DNN-90 reconstruction error map is very similar to the EAFFOV-90 of the algebraic approach. This behavior is globally noticeable in Fig. 12, which displays the MAE map of the DNN averaged over the testing

subset, for both training runs. The EAFFOV [see Fig. 12(a)] and EAFFOV-90 [see Fig. 12(b)] are represented by the dashed maroon lines. Deep hidden layers of DNN-90 also present some feature maps with neurons activated mostly in or out the new EAFFOV-90 shape.

In conclusion, the original EAFFOV shape is not an artifact due to the origin of the input dataset. The adaptation of the proposed DNN to the new shape of the EAFFOV attests that the DNN succeeds in finding the proper link between V and T , and that the aliasing artifacts that would be present taking into account the instrument characteristics are suppressed.

D. Influence of the Antenna Element Patterns

It is well known that disparities in the antenna patterns can have important effects on aperture synthesis microwave radiometer reconstructions such as a systematic error, referred to as spatial ripple or floor error. Using algebraic reconstruction methods, these effects exist even in the ideal case where the actual antenna pattern of every elementary antenna element is known [26], [27].

Therefore, it is interesting to assess the sensibility of image reconstruction to disparities in the antenna patterns when using statistical approaches and DNNs. In the simulations discussed in the previous sections, visibilities were computed using slightly different patterns for the different elementary antennas. The algebraic reconstruction did use this information on the individual antenna patterns. In contrast, the antenna patterns are not used at all for the DNN reconstructions. Despite this, the DNN reconstructions consistently exhibit smaller overall errors compared with algebraic reconstructions. At least to a certain extent, this can be due to a lower impact of antenna pattern differences in the DNN approach.

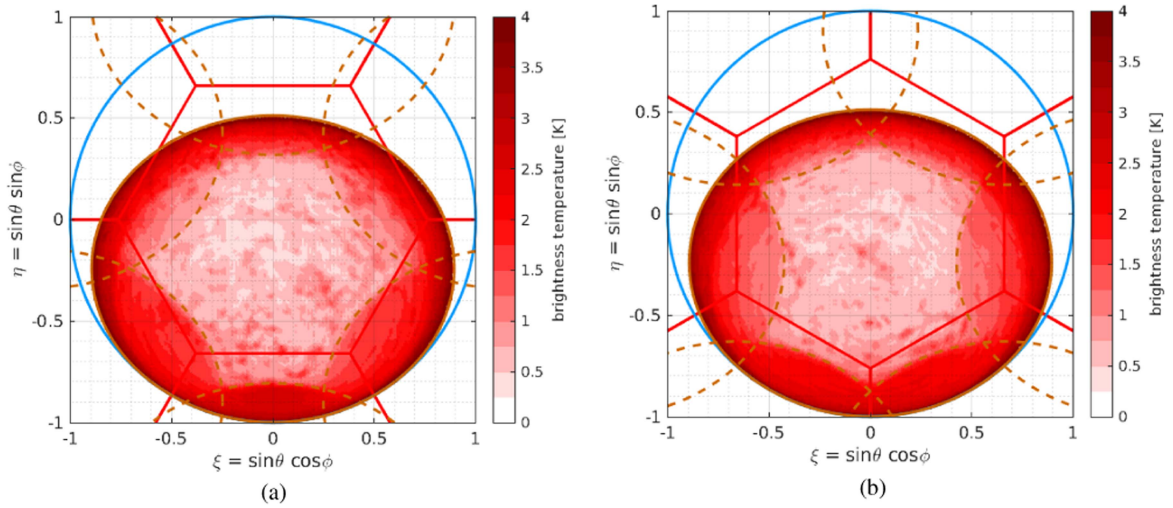


Fig. 12. MAE map of the DNN reconstructions over every snapshots of the testing subset. (a) MAE map over the testing subset of the original instrument. (b) MAE map over the testing subset of the 90° rotated instrument. In both the cases, the EAFFOV is the region of the FOV with the lowest level of the MAE. (a) MAE map on the original instrument. (b) MAE map on the instrument rotated by 90° .

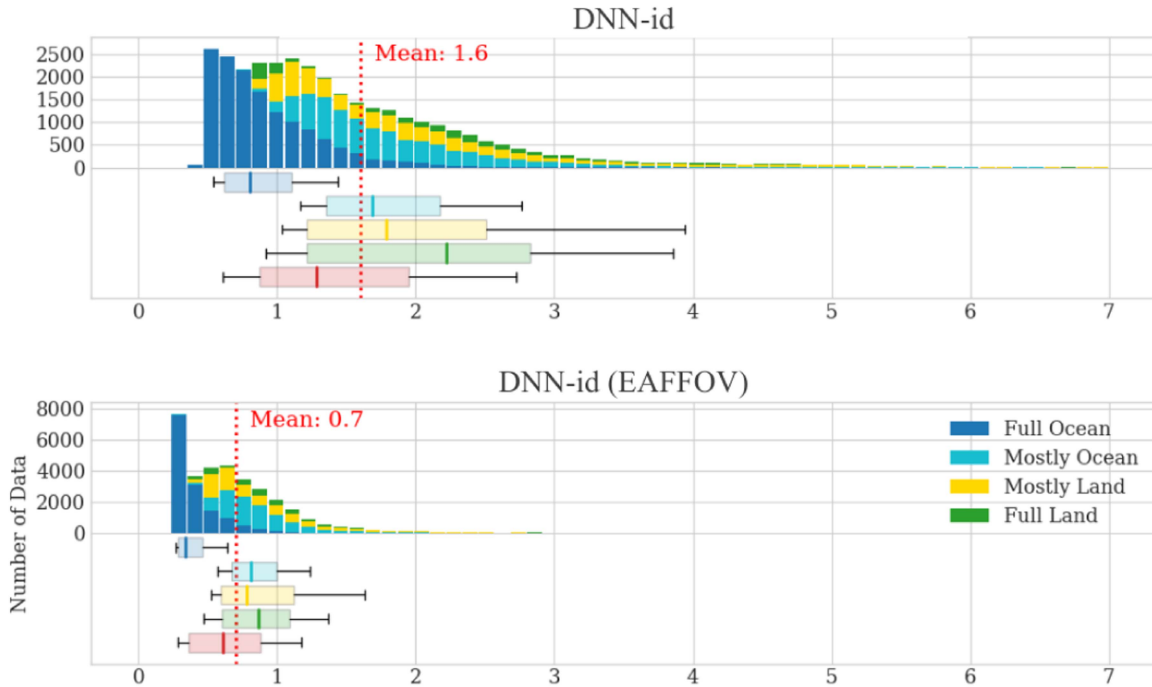


Fig. 13. Reconstructions MAE distribution of the proposed DNN, over the testing subset, with a random split of the data, colored by ocean/land snapshots composition. The DNN has been trained over visibility samples simulated with a modeling instrument with identical antenna patterns. Measured on the EAFFOV and complete areas for the DNN. “Full Ocean” are snapshots comprising more than 90% of ocean pixels, “Full Land” are snapshots with more than 90% of land, “Mostly Ocean” are composed between 50% and 90% of ocean, and “Mostly Land” are composed between 50% and 90% of land. The boxplots indicate the percentiles for each ocean/land composition and are colored accordingly. The last boxplot, colored in red, indicates the average percentiles over every snapshots. For each boxplot, the upmost left and the upmost right ticks display the 10th and 90th percentiles, respectively. The left- and right-hand sides of the boxes display the 25th and 75th percentiles, respectively. The bold vertical tick in the boxes displays the 50th percentile. The mean value of the MAE is displayed for each distribution by a vertical red dashed line.

To get further insight into the impact of antenna patterns dissimilarities, we conducted an experiment to validate that the proposed DNN can effectively learn mapping between visibility samples and brightness temperatures, irrespective of the antenna patterns used in simulating the visibility samples. We generated a new set of visibility samples, denoted as V_{id} , following the same methodology outlined in this article. However, this time,

we employed a modeling instrument with *identical* antenna patterns. Then, we trained the proposed architecture on these new T/V_{id} pairs under the same conditions of training than the proposed DNN. This newly trained model is referred to as DNN-id. The distribution of reconstruction MAE of DNN-id is displayed in Fig. 13 and should be compared to the distribution of reconstruction MAE of the proposed DNN displayed in

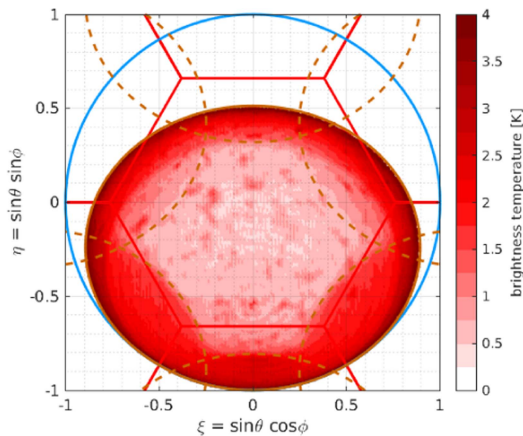


Fig. 14. MAE map of the DNN-id reconstructions over every snapshot of the testing subset with visibility samples simulated with a modeling instrument with identical antenna patterns.

Fig. 8. The reconstruction performances of DNN-id and DNN on their respective testing subsets are very similar, with an average MAE of 1.6 and 1.5 K, respectively, across the entire reconstructed field. The reconstruction performance concerning the ocean/land snapshot composition is nearly identical, except for full ocean scenes, which display slightly higher error rates when reconstructed by DNN-id. The MAE map of DNN-id reconstructions averaged over every snapshot of its testing subset is presented in Fig. 14. It reveals that errors manifest as small patches with errors inferior to 1.5 K in the EAFFOV, and the error distribution is more homogeneous outside, consistent with the pattern observed in the average MAE map of the DNN model displayed in Fig. 12(a).

In summary, the DNN reconstructions exhibit similar errors, whether visibilities are computed with an imperfect instrument with slightly different antenna patterns or if they are computed with an ideal instrument with identical antenna patterns. When using algebraic reconstructions, spatial ripples arise from the combination of antenna pattern dissimilarities and aliasing, even when the different antenna patterns are used in the inversion. Therefore, the result obtained here of a lower sensitivity of the DNN method to antenna pattern dissimilarities is promising.

E. Perspectives for Applying DNN Methods for the Image Reconstruction of Real Interferometric Data

The proposed DNN has been trained on simulated interferometric data similar to that of the SMOS instrument, in terms of spectral content, dynamic range, and underconstrained quality of the inverse problem. It has demonstrated its ability to learn the mapping between visibility samples and brightness temperature, showing equivalent or superior performances in terms of reconstruction error metrics when compared to the algebraic method. It has been shown that the proposed DNN architecture manages to deal with the strongly ill-posed problem, that the alias can be almost suppressed, and that the sensitivity of disparities of the antenna patterns can be lower. Therefore, these new results are promising. And subsequent works can deal on applying the new approaches presented here to real data acquired by a microwave interferometer. However, many issues must be

addressed before applying the proposed method to the inversion of actual interferometric data such as SMOS acquisitions. Examples are contributions from radio frequency interference (RFI), Faraday rotation angle, radiometric noise, residuals of corrected sources (direct and reflected sun, sky, and moon), polarizations, and atmospheric attenuation. In contrast to the algebraic method, DNNs are nonlinear, statistical, data-driven approaches. The training subset must encompass every contribution to the actual measurements. Thus, the network processes different error sources together, making it challenging to assess the impact of a given error source on the final reconstructions and to perform potential corrections. Dedicated works should look into detail into these aspects to address if hybrid approaches should be put in place with physical modeling complementing the statistical approaches before the actual reconstruction (preprocessing) or within the DNN training phase. The results presented in this work are just the first step to put in place such a methodology.

VII. CONCLUSION

A novel DNN that learns reconstruction mapping between complex visibilities and brightness temperature maps trained on simulated interferometric data similar to that of the SMOS instrument was presented. The new method competes with the current algebraic method on every type of scene, from pure ocean to complex mixed ocean/land scenes, with an average MAE of 0.7 K over the EAFFOV, compared to 3.7 K for the algebraic method. As a consequence of alias-free capabilities, it also retrieves an exploitable FOV 2.5 times wider than the one computed by the algebraic approach, with an MAE of 1.5 K over the full scene, by learning how to extrapolate missing information from the input data as well as from the similar scenes observed during the learning phase and, thus, remove any alias contamination. This interesting property can be helpful for the design of future imaging radiometers by aperture synthesis such as SMOS-HR [28], [29] since the spacing between the antennas would no longer be a parameter of the necessary trade-offs between scientific requirements and technical constraints. Furthermore, investigations into optimizing the reconstruction performances can be explored with the aid of external information in addition to the input data. For example, water fraction masks of each scene could be implemented into the DNN to provide additional geographic information about the position of coastlines and water points. Finally, this work might only be considered as a proof of concept of a new method that cannot yet be used for processing operational data from an interferometer such as SMOS. However, it must be seen as the first step toward a practical DL-based method to retrieve brightness temperature maps from complex visibilities. Of course, additional refinements are required to take into account all the contributions from foreign sources (sky, sun, RFI, etc.) to interferometric measurements and also to enhance the robustness of the method against various errors and noise. This is precisely the road map toward the next steps that will be detailed in a forthcoming paper.

ACKNOWLEDGMENT

The authors would like to acknowledge interesting discussions with Francois Cabot. This work made use of data from the Centre Aval de Traitement des Données SMOS operated for the Centre National D'Etudes Spatiales by the Institut Français de Recherche pour l'Exploitation de la Mer.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, 2015, Art. no. 436.
- [2] Y. Zhang, Y. Ren, W. Miao, Z. Lin, H. Gao, and S. Shi, "Microwave SAIR imaging approach based on deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10376–10389, Dec. 2019.
- [3] C. Xiao et al., "Non-uniform synthetic aperture radiometer image reconstruction based on deep convolutional neural network," *Remote Sens.*, vol. 14, 2022, Art. no. 2359.
- [4] H. Dou et al., "Deep learning imaging for 1-D aperture synthesis radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5300716.
- [5] M. A. Goodberlet, "Improved image reconstruction techniques for synthetic aperture radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1362–1366, May 2000.
- [6] B. Picard and É. Anterrieu, "Comparison of regularized inversion methods in synthetic aperture imaging radiometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 218–224, Feb. 2005.
- [7] É. Anterrieu, "A resolving matrix approach for synthetic aperture imaging radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1649–1656, Aug. 2004.
- [8] H. M. J. Barré, B. Duesmann, and Y. H. Kerr, "SMOS: The mission and the system," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 3, pp. 587–593, Mar. 2008.
- [9] Y. Kerr, P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger, "Soil moisture retrieval from space: The soil moisture and ocean salinity (SMOS) mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1729–1735, Aug. 2001.
- [10] Y. H. Kerr et al., "The SMOS soil moisture retrieval algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1384–1403, May 2012.
- [11] S. Zine et al., "Overview of the SMOS sea surface salinity prototype processor," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 3, pp. 621–645, Mar. 2008.
- [12] I. Corbella, N. Duffo, M. Vall-Ilossera, A. Camps, and F. Torres, "The visibility function in interferometric aperture synthesis radiometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1677–1682, Aug. 2004.
- [13] M. Martín-Neira, M. Suess, J. Kainulainen, and F. Martín-Porqueras, "The flat target transformation," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 3, pp. 613–620, Mar. 2008.
- [14] A. Camps, J. Bara, I. Sanahuja, and F. Torres, "The processing of hexagonally sampled signals with standard rectangular techniques: Application to 2-D large aperture synthesis interferometric radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 183–190, Jan. 1997.
- [15] E. Anterrieu, P. Waldteufel, and A. Lannes, "Apodization functions for 2-D hexagonally sampled synthetic aperture imaging radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 12, pp. 2531–2542, Dec. 2002.
- [16] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, pp. 211–252, 2015.
- [17] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, 2010, pp. 270–279.
- [18] L. Feng et al., "Array factor forming for image reconstruction of one-dimensional nonuniform aperture synthesis radiometers," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 237–241, Feb. 2016.
- [19] L. Feng et al., "The gridding method for image reconstruction of nonuniform aperture synthesis radiometers," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 274–278, Feb. 2015.
- [20] A. A. Bitar et al., "The global SMOS level 3 daily soil moisture and brightness temperature maps," *Earth Syst. Sci. Data*, vol. 9, pp. 293–315, 2017.
- [21] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. Berkeley Symp. Math. Statist. Probab.*, 1967, pp. 281–297.
- [22] T. Tieleman and G. Hinton, "Lecture 6.5-RMSPROP: Divide the gradient by a running average of its recent magnitude," *COURSERA: Neural Netw. Mach. Learn.*, vol. 4, pp. 26–30, 2012.

- [23] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Int. Conf. Int. Neural Inf. Process. Syst.*, 2019, pp. 8024–8035.
- [24] E. Anterrieu, P. Lafuma, and N. Jeannin, "An algebraic comparison of synthetic aperture interferometry and digital beam forming in imaging radiometry," *Remote Sens.*, vol. 14, no. 9, 2022, Art. no. 2285.
- [25] P. Waldteufel, É. Anterrieu, J.-M. Goutoule, and Y. H. Kerr, "Field of view characteristics of a microwave 2D interferometric antenna as illustrated by the MIRAS concept," in *Proc. 6th Spec. Meeting Microw. Radiometry Remote Sens. Appl.*, 1999, pp. 477–483.
- [26] A. Camps, I. Corbella, F. Torres, N. Duffo, M. Vall-Ilossera, and M. Martín-Neira, "The impact of antenna pattern frequency dependence in aperture synthesis microwave radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2218–2224, Oct. 2005.
- [27] M. Martín-Neira et al., "SMOS instrument performance and calibration after six years in orbit," *Remote Sens. Environ.*, vol. 180, pp. 19–39, 2016.
- [28] N. J. Rodríguez-Fernández et al., "A new L-band passive radiometer for earth observation: SMOS-high resolution (SMOS-HR)," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 5978–5981.
- [29] É. Anterrieu et al., "Preliminary system studies on a high-resolution SMOS follow-on: SMOS-HR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 8451–8454.



Richard Faucheron received the M.Sc. degree in industrial and applied mathematics from both Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées, Grenoble, France, and Université Grenoble Alpes, Grenoble, in 2020.

From 2018 to 2019, he was with the Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale, Institut de Recherche pour le Développement, Nouméa, Nouvelle-Calédonie, as an Intern Scientist for 12 months, where he worked on the development of deep learning algorithms for bio-acoustic processing. Since 2020, he has been a Centre National de la Recherche Scientifique Scientist and Research Engineer with the Centre d'Etudes Spatiales de la Biosphère, Toulouse, France, within the Soil Moisture and Ocean Salinity team. His current research interests include image reconstruction, microwave interferometric radiometry for earth observation, and machine learning.



Eric Anterrieu was born in Brive, France, in 1965. He received the Engineer and M.S. degrees in solid-state physics from the Institut National des Sciences Appliquées, Toulouse, France, in 1988, and the M.S. and Ph.D. degrees in image reconstruction in astronomy from University Paul Sabatier, Toulouse, France, in 1989 and 1992, respectively.

The subject of his thesis was the image reconstruction algorithms for multiple aperture interferometry. Since 1993, he has been an Engineer of Research in Computer Science with the Centre National de la Recherche Scientifique, Paris, France. He was with the Radio and Optical Aperture Synthesis group, Laboratoire d'Astrophysique de Toulouse-Tarbes (UMR 5572), Toulouse, from 1993 to 2000; with the Signal and Image Processing Team, Centre Européen de Recherche et Formation Avancée en Calcul Scientifique (URA 1875), Toulouse, France, from 2000 to 2004; and with the Signal, Image and Instrumentation group, Institut de Recherche en Astrophysique et Planétologie (UMR 5277) from 2005 to 2016, Toulouse. Since 2017, he has been with the Observing Systems Team, Centre d'Etudes Spatiales de la Biosphère (UMR 5126), Toulouse. His current research interests include numerical analysis and image and signal processing with particular emphasis on the Soil Moisture and Ocean Salinity mission and on the future microwave missions using interferometry, aperture synthesis, and beamforming.



Louise Yu received the Ph.D. degree in astrophysics from Paul Sabatier University, Toulouse, France, in 2019.

The subject of her thesis was the detection and characterization of hot Jupiters around forming stars. Since 2020, she has been an Engineer in Data Processing for Microwave Space Missions with the Centre National D'Etudes Spatiales, Toulouse, and has participated to the Soil Moisture and Ocean Salinity High-Resolution project. Her current research interests include swath altimetry missions.



Ali Khazaal was born in Lebanon in 1981. He received the Engineer degree in telecommunication and computer science from the Faculty of Engineering, Lebanese University, Beirut, Lebanon, in 2003, the M.S. degree in signal, image, acoustic, and optimization from Institut National Polytechnique, Toulouse, France, in 2006, and the Ph.D. degree in signal/image processing and optimization from Paul Sabatier University, Toulouse, France, in 2009.

He was working with the Signal, Image and Instrumentation Group, Laboratoire d'Astrophysique de Toulouse-Tarbes, Observatoire Midi-Pyrénées, Paul Sabatier University. The subject of his thesis was the image reconstruction algorithms for the Soil Moisture and Ocean Salinity (SMOS) mission. From 2009 to 2018, he was a Research Scientist with the Centre d'Études Spatiales de la Biosphère, Toulouse. In 2019, he started his own company "RDIS Conseils," Toulouse, specialized in the research and development in spatial imaging. His current research interests include numerical analysis and signal and image processing with direct application on European Space Agency (ESA)'s SMOS mission and is a Member of ESA's SMOS Payload Calibration Committee. He is also working on the future microwave missions using aperture synthesis and in particular SMOS High-Resolution project of the Centre National D'Études Spatiales.



Nemesio J. Rodríguez-Fernández (Member, IEEE) received the Licentiate degree in fundamental physics and the Ph.D. degree in astrophysics from the Universidad Complutense de Madrid, Madrid, Spain, in 1996 and 2002, respectively.

From 2002 to 2004, he was a Marie Curie Fellow with the Observatoire de Paris, Paris, France. From 2004 to 2006, he was a Teaching Assistant with Paris Diderot University, Paris, France, and the University of Bordeaux, Bordeaux, France. From 2006 to 2011, he was an Astronomer with the Institut de Radio Astronomie Milimétrique, Grenoble, France. Since 2012, he has been a Centre National de la Recherche Scientifique Scientist and a Research Engineer with the Centre d'Études Spatiales de la Biosphère, Toulouse, France, with the exception of 11 months from 2015 to 2016 when he joined the European Center for Medium-Range Weather Forecast (ECMWF), Reading, U.K., as a Research Scientist. He designed the neural network algorithm of the European Space Agency Soil Moisture and Ocean Salinity (SMOS) soil moisture near real-time product and the one that is used for operational data assimilation of SMOS soil moisture at the ECMWF. He is also working on aperture synthesis algorithms and new satellite mission design (SMOS High Resolution and the Fine Resolution Explorer for Salinity Carbon and Hydrology). His current research interests include microwave remote sensing for earth observation, in particular machine learning approaches for the estimation of geophysical parameters, such as soil moisture and biomass and their assimilation into numerical weather prediction and carbon cycle models.