MIIDAPS-AI: An Explainable Machine-Learning Algorithm for Infrared and Microwave Remote Sensing and Data Assimilation Preprocessing - Application to LEO and GEO Sensors

Eric S. Maddy and Sid A. Boukabara

***Abstract*—In this article, we leverage and apply state-of-the-art artificial intelligence (AI) techniques to satellite remote sensing of temperature, moisture, surface, and cloud parameters in all-weather, all-surface conditions, from both microwave and infrared sensors. The multi-instrument inversion and data assimilation preprocessing system, artificial intelligence version, or MIIDAPS-AI for short, is valid for both polar and geostationary microwave and infrared sounders and imagers as well as for pairs of combined infrared and microwave sounders. The algorithm produces vertical profiles of temperature and moisture as well as surface temperature, surface emissivity, and cloud parameters. Additional products from hyperspectral infrared sensors include selected trace gases. From microwave sensors, additional products such as rainfall rate, first year/multiyear sea ice concentration, and soil moisture can be derived from primary products. The MIIDAPS-AI algorithm is highly efficient with no noticeable decrease in accuracy compared to traditional operational sounding algorithms. The automatically generated Jacobians from this deep-learning algorithm could provide an explainability mechanism to build trustworthiness in the algorithm, and to quantify uncertainties of the algorithm's outputs. The computation gain is estimated to be two orders of magnitude, which opens the door to either 1) process massively larger amounts of satellite data, or to 2) offer improvements in timeliness and significant saving in computing power (and therefore cost) if the same amount of data is processed. Here, we present an overview of the MIIDAPS-AI implementation, discuss its applicability to various sensors and provide an initial performance assessment for a select number of sensors and geophysical parameters.**

***Index Terms*—Artificial Intelligence (AI), atmosphere, earth observing system, machine learning, neural networks, remote sensing, satellite.**

Manuscript received May 23, 2021; revised July 9, 2021; accepted August 3, 2021. Date of publication August 5, 2021; date of current version September 9, 2021. This work was supported in part by the U.S. Department of Commerce under NOAA/NESDIS Office of Projects, Planning, and Acquisition (OPPA) Technology Maturation Program (TMP) under Grant ST133017CQ0058/1332KP20FNEED0029. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce *(Corrresponding author: Eric S. Maddy.)*

Eric S. Maddy is with the Riverside Technology, Inc., College Park, MD 20770 USA (e-mail: eric.maddy@noaa.gov).

Sid A. Boukabara is with the National Oceanic and Atmospheric Administration Center for SaTellite Applications and Research (STAR), College Park, MD 20770 USA (e-mail: sid.boukabara@noaa.gov).

Digital Object Identifier 10.1109/JSTARS.2021.3104389

I. Introduction

THE multi-instrument inversion and data assimilation preprocessing system-artificial Intelligence (MIIDAPS-AI) leverages modern artificial intelligence (AI) techniques in remote sensing to efficiently emulate traditional remote sensing algorithms such as the microwave integrated retrieval system (MiRS) [1], [2] for microwave sensors or the NOAA Unique CrIS/ATMS Processing System (NUCAPS) [3], [4] for infrared sensors. This is important because data volumes are growing quickly for a number of reasons including 1) the imminent dawn of an era of small Earth observation satellites (smallsats and cubesats) with the potential to significantly increase the challenge of processing larger volumes of data, 2) an increase in the number of satellites being deployed by an increasing number of international partners, and 3) an increase in new sensors capabilities to measure the environment at higher spatial, temporal, and spectral resolutions[5], [6]. Neural networks techniques have been used for decades in remote sensing [7], [8]. Essentially, neural network algorithms, as well as new AI and/or machine learning (ML) techniques, offer a way to approximate any function be it linear or nonlinear without any presumptions on its inputs errors characteristics.

The timing of the MIIDAPS-AI development however leverages the extraordinary progress made recently in the field of AI and specifically ML. This progress is in terms of 1) the significant increase in availability of AI-specific compute nodes [i.e., Graphical Processing Units (GPUs) using Compute Unified Device Architecture (CUDA)], 2) ease of use of high-level languages to develop AI/ML algorithms (such as Keras/Tensorflow [9] and PyTorch), and 3) ability to easily develop deep-learning algorithms that can handle more sophisticated activation functions and input types. MIIDAPS-AI is therefore a new generation, enterprise remote sensing algorithm, based on modern AI and ML techniques, with capability to infer a number of geophysical products from a number of space-borne sensors. As an extension of traditional 1D variational approaches such as the MiRS (and its extension to IR satellite observations, MIIDAPS), MIIDAPS-AI can be applied to Infrared (IR) and Microwave (MW) polar and GEO sounders and imagers and is valid for any sensor with valid Community Radiative Transfer Model

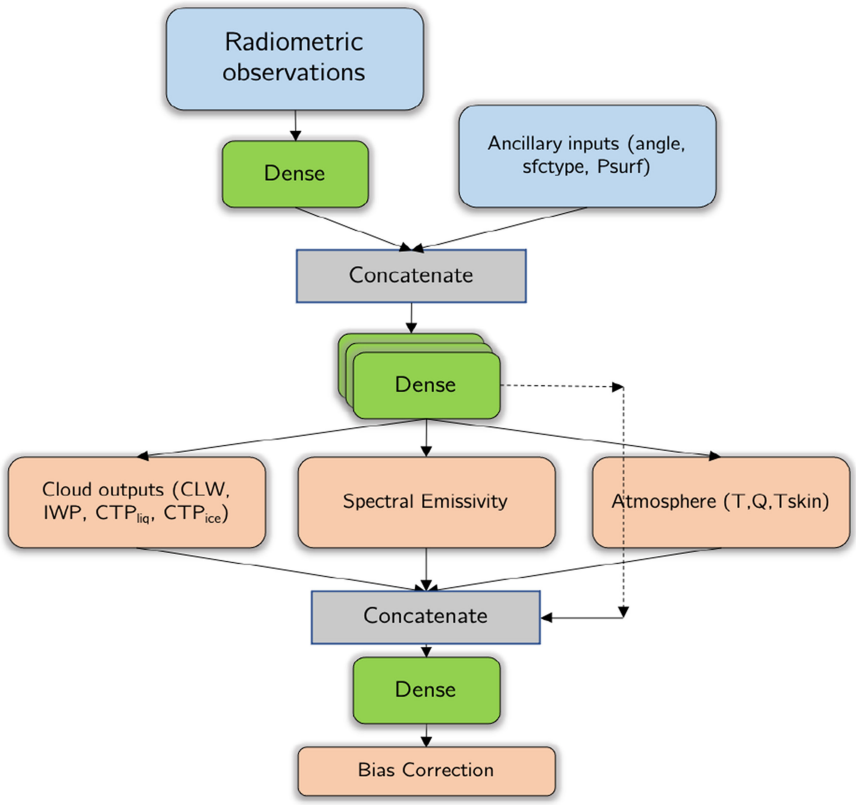


Fig. 1. Flow diagram for the MIIDAPS-AI retrieval algorithm. Algorithm inputs (blue) pass through several trainable fully connected dense layers (green) to produce outputs (orange). Untrainable layers which concatenate features at various depths in the algorithm flow (grey) are also shown. See text for a more detailed description of algorithm inputs and outputs.

(CRTM) coefficients [10]. MIIDAPS-AI can additionally serve as a preprocessor for data assimilation and data fusion [11].

The rest of this article is organized as follows. Section II describes the theoretical basis and implementation of MIIDAPS-AI. Section III, provides an overview of MIIDAPS-AI products and valid sensors. Statistical assessments of MIIDAPS-AI products versus NWP models, radiosondes, traditional remote sensing algorithms are shown in Section IV. Section V discusses the interpretability and trustworthiness of MIIDAPS-AI.

II. Theory and Implementation

Fundamentally, MIIDAPS-AI is a deep fully connected neural network that defines a nonlinear mapping between instrument measurements (IR radiances or MW brightness temperatures) and geophysical parameters such as temperature and moisture profiles, integrated cloud parameters [including cloud liquid water (CLW) and ice water path (IWP)], spectral surface emissivities as well as bias corrections between observations and scenes simulated using the CRTM. The MIIDAPS-AI network architecture (units, layers, etc.) varies based on instrument or instrument combination but the general architecture is shown in Fig. 1. For a given sensor, MIIDAPS-AI is trained to find a functional mapping, f_θ, between observations, y and desired geophysical parameters x, i.e., $x = f_\theta(y)$, where θ are parameters determined offline using millions of training samples pairs. These samples consist of colocated instrument measurements, numerical weather prediction (NWP) fields (e.g., from the National Oceanic and Atmospheric Adminstration's (NOAA's) Global Data Assimilation System (GDAS), or the European Centre for Medium-range Weather Forecasting's (ECMWF's) Reanalysis (ERA5)[12]) spatially (bilinear) and temporally interpolated to the observation locations, surface spectral emissivity, and CRTM simulations using those geophysical fields. Over land and for microwave instruments, the Tool to Estimate Land Surface Emissivities at Microwave frequencies (TELSEM) model is used to analytically calculate the emissivities for all surfaces using NWP data collocated with satellite brightness temperatures [13], while over ocean, the FAST microwave Emissivity Model (FASTEM) model [14] is used to prescribe emissivity. For infrared instruments, the University of Wisconsin Baseline Fit Emissivity Database [15] is used to prescribe emissivity over land and the surface-leaving radiance model developed in [16] is used over ocean.

Multiple years and seasons, diverse surface types, geographic coverage, and all-weather conditions are included in the training sample such that the relationships determined are representative of all cases. For instance, at least two years of data with a minimum of eight days in each year and covering each of the four seasons are first selected. Data are first selected for each day by randomly subsampling the full day of quality controlled (basic QC flags) observations by a factor five. A

second pass is then performed on each day to ensure that each 10 degree by 10 degree latitude/longitude box contains at least one observation. The dataset is then split into two parts (one-fold cross-validation) by selecting chronologically the first 80% for training *and last* 20% for testing (selection of best weights). In our initial testing of MIIDAPS-AI performances, we found networks *optimized with an 80/20 training/testing split* performed better for generalization tasks (validation on samples not seen during training/testing). Optimization of internal network parameters or weights is performed using the Python implementation of the Keras deep learning software [9] and the Adam optimizer. The weights corresponding to the epoch with the best (smallest) validation loss on the testing samples is selected out of a maximum number of epochs (usually between 200 and 1000 and depending on the instrument). In addition, an offline optimization of network hyperparameters (network architectures, layers/nodes) is performed for each instrument using a coarse grid search over network layers (linearly between 2 to 4) and nodes (logarithmically between 32 and 1024) and using a 20% (random subsample) of both the full training and testing datasets. We have found that in general similar optimized network hyperparameters are obtained for similar instruments [e.g., MW sounders such as the Advanced Microwave Sounding Unit/Microwave Humidity Sounder (AMSU/MHS) pair and the Advanced Technology Microwave Sounder (ATMS)].

As mentioned above, the algorithm produces not only geophysical parameters, but also spectral radiative bias corrections between observations and outputs. These bias corrections are included in the optimization for several reasons: 1) to account for possible differences between "true" geophysical state observed by the instrument and prescribed geophysical state from NWP models, and 2) to account for any instrument (or forward model) biases. This allows the system to determine the air-mass dependent bias on an individual case basis, as opposed to applying a wide-ranging bias correction computed offline statistically. This is similar in principle to the variational bias correction used in NWP where the bias is computed on the fly—on an individual case basis—using multiple predictors [17]. While bias corrections are estimated by the network over all surfaces, unlike other parameters, the bias correction is trained with a custom loss function, which masks losses (sets to zero) over nonocean. Therefore, the network weights determined during optimization are constrained only by biases computed for ocean surfaces. This is because we would expect models described in [14], [16] to agree better with actual sea surface emissivities and therefore expect a better agreement between CRTM simulations and real IR and MW observations.

A. MIIDAPS-AI Network Structure

Fig. 1 shows an example of the general neural network architecture and flow of MIIDAPS-AI. Inputs to the algorithm are shown in blue and include radiometric observations (IR radiances or MW brightness temperatures) as well as ancillary information such as instrument view/solar angle (or secant), forecast surface pressure and categorical variables such as surface type (one-hot encoded in ML language). Outputs from the algorithm are shown in orange and include cloud parameters, spectrally resolved emissivity, surface temperature and profiles of temperature and moisture, and a bias correction as described above. Dense/fully connected trainable layers whose weights and biases are optimized during training are shown in green and are either single layers or stacks of several hidden layers as shown in the center of the diagram before the main output layers. Those dense layers have between 64 and 1024 nodes, the actual number depending on the instrument, and utilize the Rectified Linear Unit (ReLU) activation. Concatenate layers are shown in gray and indicate a linking of all the outputs from the preceding layers in the network to the inputs to the following layer. In general, arrows in the figure indicate layers that are directly connected. Arrows from the stacks of dense layers in the middle of the architecture and bypassing the main outputs of cloud, emissivity, and atmosphere are skip connections, which help avoid the vanishing gradient problem typically seen in deep neural networks [18] and help preserve input information encoded in preceding layers to produce the bias correction outputs.

III. MIIDAPS-AI Applicability

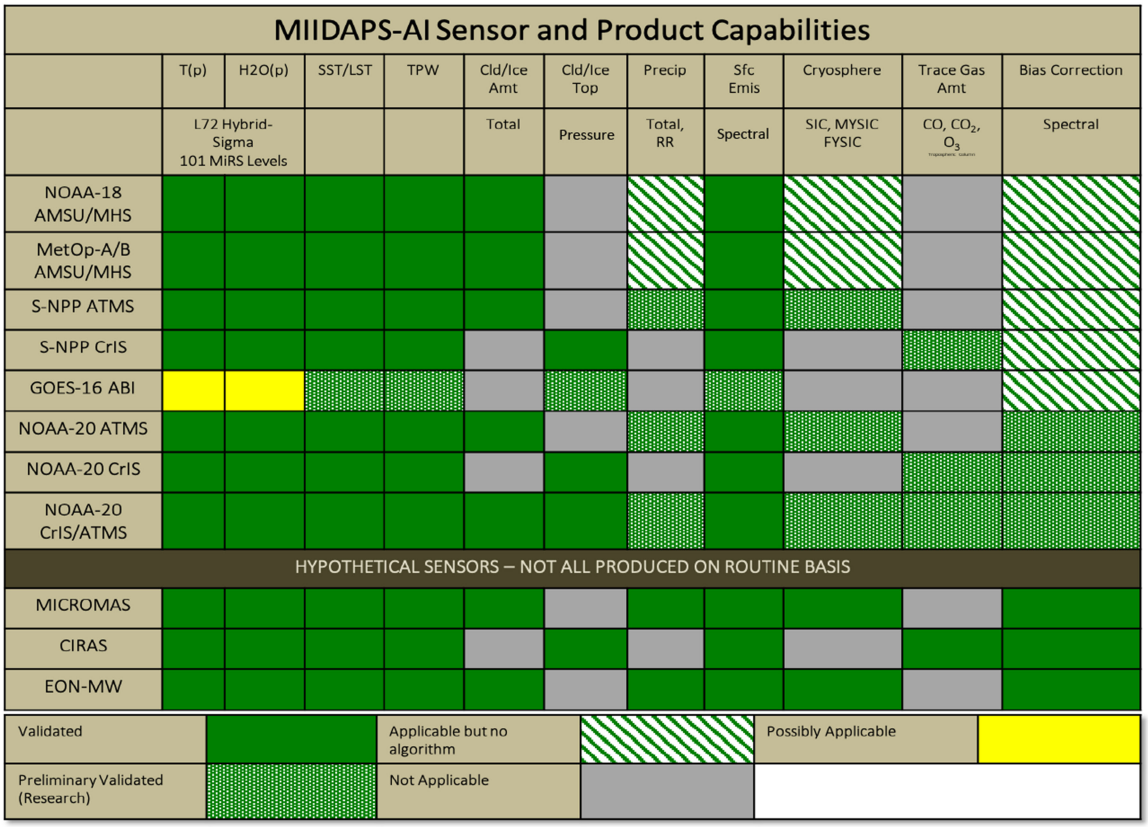
MIIDAPS-AI is an enterprise algorithm, i.e., an algorithm that extracts the maximum information content from data measured by a wide range of sensor types. MIIDAPS-AI generates several geophysical parameters from a number of sensors (or sensor pairs), as described in Table I. In principle, an instance of MIIDAPS-AI could be generated for any sensor for which the CRTM model or other forward model can be used to simulate the sensor; however, in practice, algorithms have been produced and tested on a smaller set of polar and GEO infrared and microwave imagers and sounders. In addition to real operational sensors currently flying, MIIDAPS-AI has been applied to emerging technologies and hypothetical sensors (esp. smallsats and cubesats) to assess the value of those sensors to infer information about the atmosphere and surface. This enables assessment of the performances we should expect from these proposed sensors.

Table I thus lists the current sensors for which MIIDAPS-AI has been validated and also describes the geophysical capability of those sensors. See the caption in Table I for the color coding indicating, which sensor(s) and which geophysical parameters MIIDAPS-AI has been applied to, tested and/or validated for. Rows in the column correspond to polar microwave instruments (AMSU/MHS, ATMS, Micro-sized Microwave Atmospheric Satellite (MICROMAS), and Earth Observing Nanosatellite-Microwave (EON-MW)) and infrared (Cross-track Infrared Sounder (CrIS), CubeSat Infrared Atmospheric Sounder (CIRAS)) sounders as well as geostationary imagers [Advanced Baseline Imager (ABI)].

IV. Qualitative and Quantitative Assessment of MIIDAPS-AI Performance

In this section, we highlight selected examples from our comprehensive assessment of MIIDAPS-AI applied to all the sensors and associated generated products listed in Table I. Note that additional validation datasets will also be employed in future performance assessments.

TABLE I
LIST OF PARAMETERS (IN THE COLUMNS) THAT MIIDAPS-AI GENERATES WHEN APPLIED TO THE MICROWAVE AND INFRARED SENSORS FROM BOTH GEO AND LEO PLATFORMS LISTED IN THE ROWS



Note: The sensors are stratified by operational and proposed/research types (top and bottom section respectively). This table shows the applicability of MIIDAPS-AI and what it has been tested for. The color coding refers to the degree of confidence we have in the validity of the MIIDAPS-AI capability and is highlighted in the bottom of the table: green for fully validated, Dotted green for preliminarily validated. When a sensor is not capable of retrieving a parameter, the cell is colored gray. Yellow indicates that the sensor has some sensitivity to the parameter (i.e., some nominal information content exists), but no capability has yet been demonstrated. Left hashed green line indicates that the sensor is known to have the capability but MIIDAPS-AI has not been generated yet for this sensor/parameter pair. Columns in the table correspond to: atmospheric parameters - temperature, T(p) and moisture, H2O(p), profiles, sea surface/land temperature (SST/LST), TPW; cloud and precipitation parameters - cloud liquid/ice amount (CLD/ICE AMT), cloud top pressure of cloud liquid and ice (CLD/ICE TOP), precipitation (Precip) rain rate (RR); surface parameters - spectral emissivity (SFC EMIS), total sea ice concentration (SIC), multi and first year SIC (MYSIC, FYSIC); chemistry parameters - trace gas tropospheric column amounts (CO2, CO, CH4 AMT); and radiometric parameters - spectral bias correction.

A. Visual Assessment of MIIDAPS-AI Product Suite From Multiple Sensors

Fig. 2 shows a visual assessment of some geophysical products generated by MIIDAPS-AI from a select number of sensors. These products and sensors include total precipitable water (TPW) from SNPP/ATMS, IWP and CLW from NOAA-18 AMSU/MHS pair, Carbon Monoxide (CO) in the atmosphere from the CrIS sensor onboard NOAA-20 satellite, the emissivity at channel 1 (23GHz) from SNPP/ATMS, Liquid Cloud Top Pressure from GOES-16 ABI, sea ice concentration (SIC) from NOAA-20/ATMS and finally Skin temperature from the CrIS/ATMS pair onboard NOAA-20.

B. Quantitative Assessment of MIIDAPS-AI Versus NWP and Operational Algorithms

To illustrate an example of the performances of MIIDAPS-AI, quantitatively, using microwave and infrared sensors, we compare in Fig. 3, the retrieval of the TPW, and the temperature and moisture profiles, as generated by the MIIDAPS-AI algorithm, for April 7, 2019, to the ECMWF analysis for the same period. These retrievals use real data from JPSS ATMS and CrIS sounders. We also include in Fig. 3 the statistical performances of these retrievals and compare them to those obtained by operational algorithms as a point of reference. What Fig. 3 mainly shows is that the MIIDAPS-AI performances for temperature and moisture profiling, are overall comparable to those of operational algorithms, in this case NUCAPS.

In addition, MIIDAPS-AI is at least two orders of magnitude more computationally efficient. It can for instance process 720 days of ATMS microwave data as opposed to 3.6 days by MiRS or 400 days of CrIS infrared data as compared to 3.9 days by NUCAPS in one wall-clock day of computing using the exact same computing resource. From these considerations, we can say that MIIDAS-AI's main advantage is the great efficiency gain achieved with little noticeable degradation in performance. It is worth noting that the algorithm was trained using

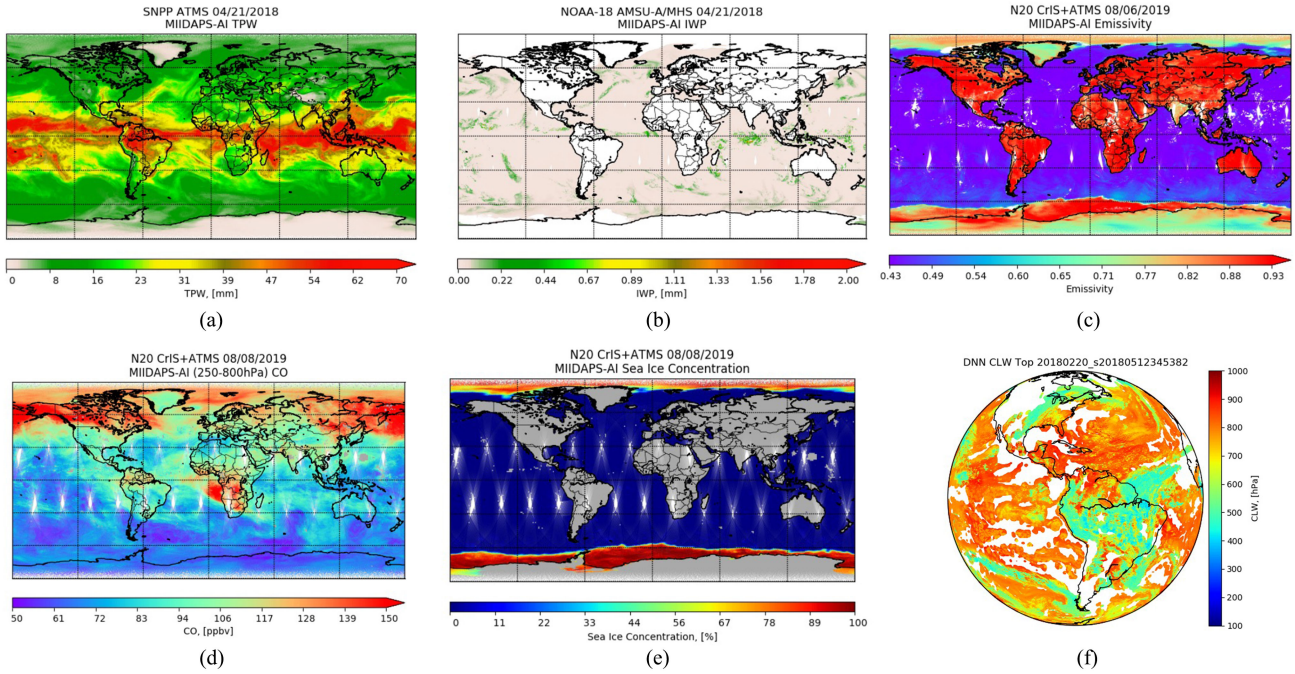


Fig. 2. Illustrations of examples of geophysical parameters generated by MIIDAPS-AI, from a select number of sensors, both microwave and Infrared, from LEO and/or GEO platforms. See text for details.

15 days' worth of ECMWF and JPSS ATMS and CrIS co-locations covering two years and multiple seasons.

C. Spatial Variability and Interparameters Correlation Assessment

Additionally, the fields of moisture and temperature obtained by MIIDAPS-AI were assessed in terms of their spatial variability and interparameter correlation structure. These assessments demonstrate that the fields of geophysical parameters determined by MIIDAPS-AI are geographically consistent in terms of their spatial variability (see Fig. 4) and that the calculated geophysical state vector is self-consistent (see Fig. 5).

This assessment of the spatial and interparameters consistencies is as important as assessing the performances in terms of accuracy and precision. To assess spatial consistency, ECMWF and MIIDAPS-AI retrievals are first colocated. The two datasets are then spatially averaged onto a 1×1 degree latitude/longitude grid and a spectral decomposition is performed. The power spectrum of the coefficients of each is plotted for several layers and for temperature and moisture. Fig. 4 shows that the variability of the MIIDAPS-AI temperature field is consistent with that of ECMWF, for the different layers assessed, which correspond to 925, 750, and 250 mb. The solid and the dashed lines are indeed almost on top of each other except at the short wavelengths. For the moisture spatial variability, Fig. 4 indicates that the variability is also mostly consistent between the ECMWF field and MIIDAPS-AI generated field of humidity. Differences at smaller spatial scales is somewhat expected given that the NWP field is an analysis that is bound to smooth the high-frequency features, while MIIDAPS-AI inversions, since they are based exclusively on real observations, capture the higher-frequency features of spatial variability that occur in reality.

Fig. 5 shows that the parameters inverted by MIIDAPS-AI, at least those assessed here, namely atmospheric temperature and moisture profiles, along with skin temperature, are correlated to each other in roughly the same manner as these parameters are found to be correlated in the NWP field of ECMWF. As would be expected from an inversion of atmospheric profiles from a MW/IR sounder with finite vertical resolution, the structure of the vertical correlation of MIIDAPS-AI temperature and moisture is similar, but shows some differences as compared to ECMWF (e.g., in the temperature/moisture correlation block in Fig. 5).

The combination of the assessments of the MIIDAPS-AI inversions, in terms of precision and accuracy (see Fig. 3), spatial variability (see Fig. 4), and interparameters correlations (see Fig. 5), demonstrates the good performance and statistical consistency of MIIDAPS-AI. In the next sections, we 1) perform an independent assessment of the MIIDAPS-AI system (by an independent team and using a different and an independent set) and 2) assess whether we can trust the outcome of MIIDAPS-AI on an individual retrieval basis.

D. Independent Assessment of MIIDAPS-AI Versus Radiosondes

In the previous section, we focused on developing and testing the MIIDAPS-AI algorithm using NWP models. In this section, we present the results obtained from an independent assessment of the performances of MIIDAPS-AI—over a different time

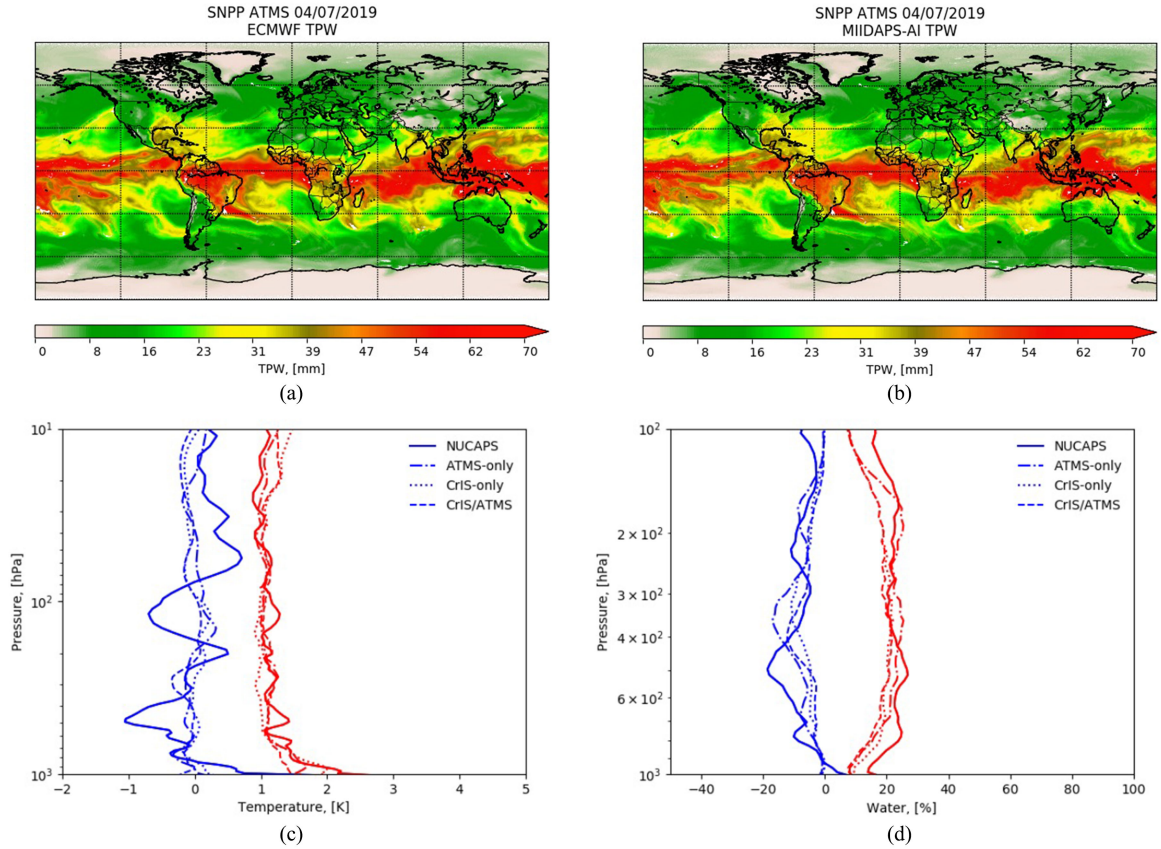


Fig. 3. (a) ECMWF TPW compared to (b) MIIDAPS-AI based TPW using SNPP/ATMS real data. (c) and (d) Statistics computed using ECMWF as a reference for MIIDAPS-AI CrIS-only, ATMS-only, and combined CrIS/ATMS as indicated by the different dashed patterns. These statistics (mean difference in blue and standard deviation of difference in red) are compared to those obtained when using the combined NUCAPS CrIS/ATMS operational algorithm shown with a solid line (NUCAPS QC applied to all retrievals). Temperature performance is highlighted in (c) and moisture performance, in percent, is presented in (d). See text for more details.

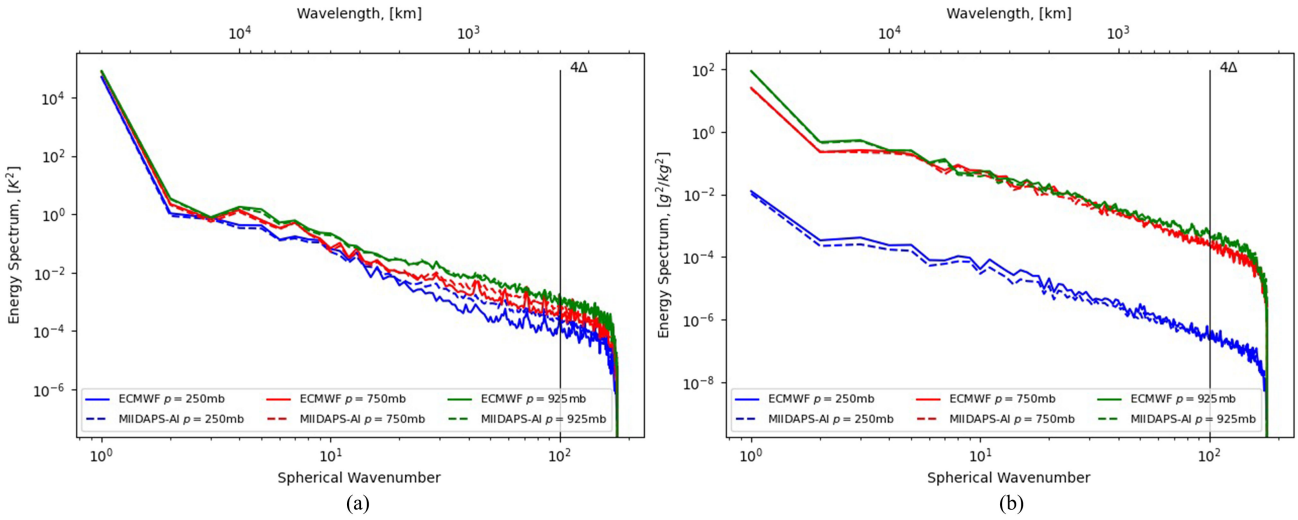


Fig. 4. Spatial variability assessment of the temperature (left) and the moisture (right) for a single day: September 7, 2020. The power spectrum, a measure of how frequent a particular variability occurs in the geophysical field, is shown for the ECMWF field at different levels of the atmosphere in blue, red, green, (solid lines), as well as for the MIIDAPS-AI field (dashed lines). The vertical line at 4Δ indicates the spherical wavenumber and corresponding grid spacing of the highest frequency computed for MIIDAPS-AI and ECMWF which in this case is 100 km.

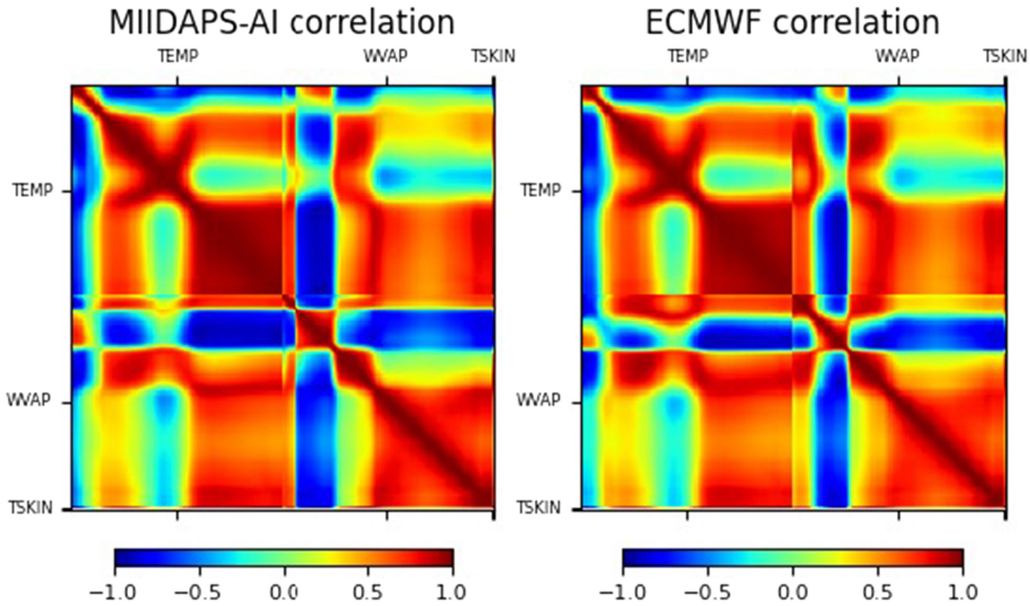


Fig. 5. Assessment of NOAA-20 Combined CrIS/ATMS MIIDAPS-AI Inter-parameters intercorrelation for a single day September 7, 2020. The left panel shows the correlation matrix of the MIIDAPS-AI retrievals, representing the correlation between skin temperature, atmospheric moisture and temperature profiles. Red represents positive correlation and dark blue represents anti-correlation. Temperature and moisture profile layers are ordered such that pressures higher in the atmosphere are toward the left of the Figure. For comparison, the right panel represents the same matrix but computed based on ECMWF NWP analysis.

period, using a different reference data type (radiosondes), and performed independently by a different team.

Fig. 6 shows statistical assessments of 12 days of colocated MIIDAPS-AI and NUCAPS using operationally launched radiosondes as a reference. This independent assessment was performed by the NOAA Products Validation System (NPROVS) [19] using their software for retrieval/radiosonde colocation, statistical computation, and QC. We note that due to the relative difference in vertical resolution of radiosonde and satellite measurements/retrievals, a rigorous assessment should account for the averaging kernels of both NUCAPS and MIIDAPS-AI (e.g., [20], [21]) however, that comparison will be left for a future publication.

To provide insight into the complexity of scenes assessed, the comparison statistics are computed using two QC schemes - panels (a) and (b) include cases, where NUCAPS combined IR+MW retrievals converged, while for panels (c) and (d), for cases where the IR+MW failed and if the MW-only retrieval was successful, those MW-only retrievals were included. NUCAPS utilizes a technique known as cloud-clearing to enable sounding in partially cloudy scenes and the main reason the IR+MW algorithm fails to converge is a failure of the cloud-clearing algorithm (i.e., cloudier or scenes which violate cloud-clearing assumptions).

Generally speaking, the results are consistent with global comparison to ECMWF even though radiosondes observations are less spatially uniform and are mostly land cases. Compared to NUCAPS, MIIDAPS-AI retrievals retrieval performance is similar in the top and bottom panels and suggest MIIDAPS-AI performance is not a strong function of cloudiness, while NUCAPS IR+MW QC removes more difficult cases. For both QC schemes and for temperature and moisture, MIIDAPS-AI provides similar sounding performance to NUCAPS above 700 hPa and somewhat improved skill near the surface.

V. Model Interpretation and Visualization for Understanding and Transparency

Ideally, we would like to ensure that the ML-based algorithms are performing well, statistically and on an individual retrieval basis. We also would like to ensure that the results are generating the right answer for the right reasons. In other words, we should be able to *explain* the results or at least have access to a metric, generated along with the individual retrieval itself, that supports our trust in the results. In this study, we explore using the ML Jacobians, generated automatically while developing the MIIDAPS-AI retrieval algorithm itself, to explain the results.

A. Sensitivity of MIIDAPS-AI to Radiometric Observations

In Fig. 7 above, daily ML based MIIDAPS-AI Jacobians (of temperature with respect to the channel brightness temperatures), based on MIIDAPS-AI applied to ATMS, were averaged over a single day. This result shows that the ML Jacobians are overall consistent with the expectation that the information that led to the temperature retrieval, at a certain layer, originates from channels that peak at those layers. For example, channels labeled 0, 1, 2, 3, 4, and 5 corresponding to (23, 31.4, 50.3, 51.76, 52.8GHz frequencies) are sensitive to the surface mainly and the lower troposphere. Channels labeled 6 through 13 are known to be sensitive to higher tropospheric and stratospheric layers and results of this MIIDAPS-AI ML Jacobian is consistent with this. Channels 15 through 21 are either surface sensitive channels or atmospheric moisture sounding channels and therefore have their sensitivity peak at the lower layers, consistent with the

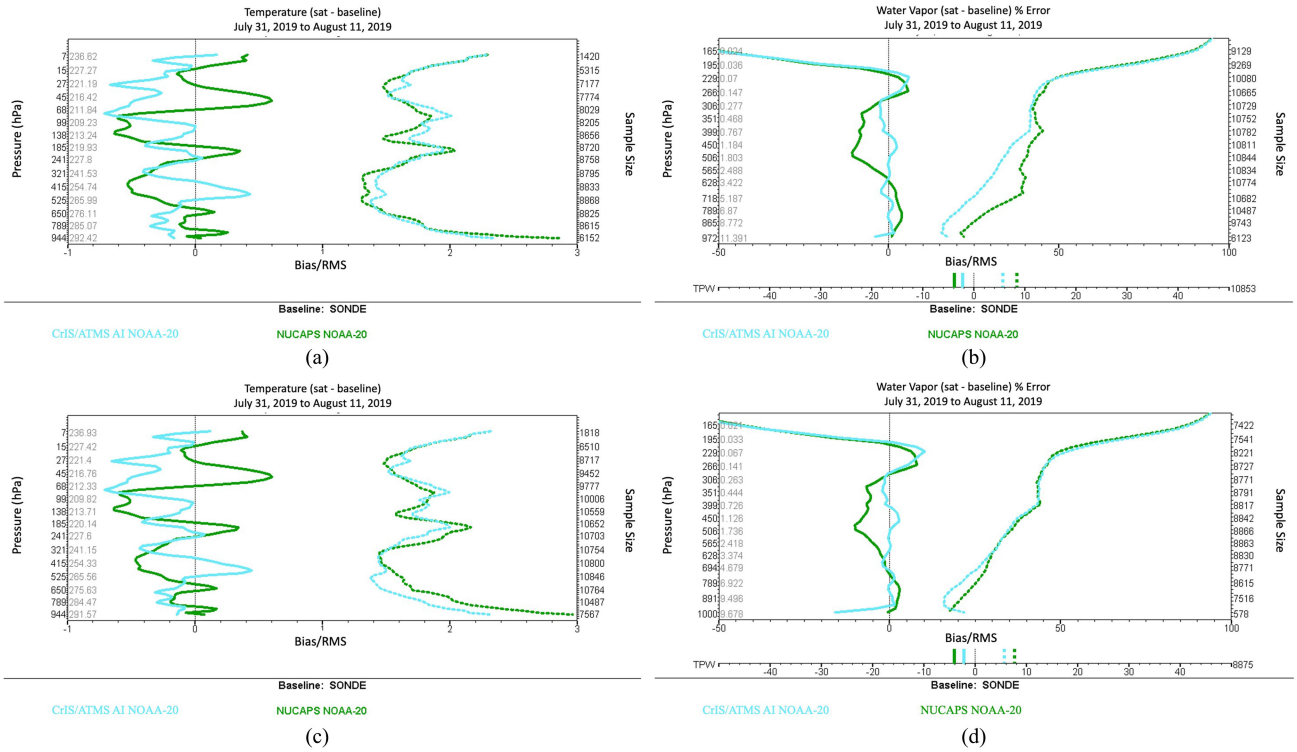


Fig. 6. Bias (solid) and RMSE (dashed) statistics computed using all operational radiosondes (July 31,2019-August 11,2019) as a reference for NUCAPS and MIIDAPS-AI for NOAA-20. (a) Temperature and (b) moisture statistics of MIIDAPS-AI combined CrIS/ATMS and NUCAPS (IR+MW QC = 0) for operational radiosondes. (c) and (d) These same statistics computed using NUCAPS QC (IR+MW QC = 0 or MW-only QC = 0). Bottom vertical bars on the water vapor panels show the TPW vapor bias and RMSE. Sample size is shown on the right of each panel. See text for more details.

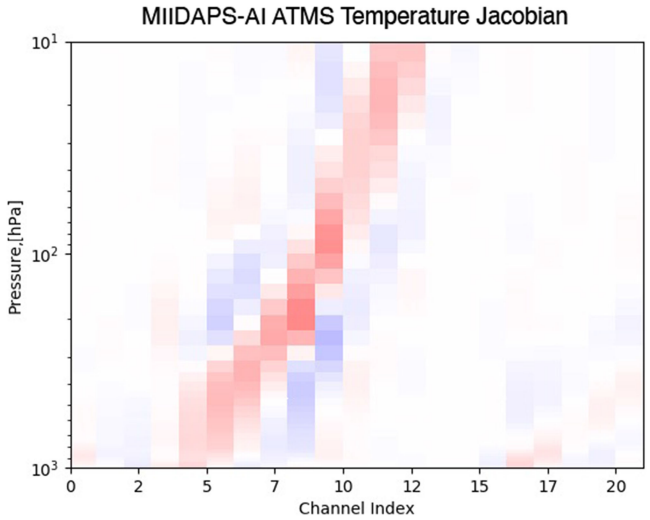


Fig. 7. Daily Average MIIDAPS ML-Jacobians normalized between -1 (blue) and 1 (red). See text for details.

findings in Fig. 7. This builds confidence that the performances obtained statistically shown in the previous section, are anchored in a physically based sensitivity of the temperature to the appropriate sensor's channels. In other words, that the algorithm is performing well for the appropriate reasons: appropriate sensitivities of observations -brightness temperatures- to the appropriate layers in the atmosphere.

B. Error Quantification

An important feature of the Bayesian physical retrieval methodology is the ability to estimate the a posteriori uncertainty of the resultant retrieval. This can be very useful for users when using the retrievals for their applications. Can MIIDAPS-AI do the same?

We propose two ways of predicting the errors of MIIDAPS-AI individual retrievals due to random noise. In the next Figure, we assess these two approaches and compare them to the true errors, obtained by simply comparing the retrievals to the true state of the environment. This is done using simulated data to know exactly the truth and therefore the true errors. These two approaches are described as follows.

1) Nonlinear Monte Carlo Error Estimation Approach: In this approach, we take advantage of the fact that MIIDAPS-AI is very efficient and, therefore, we can run it multiple times, each time randomly perturbing the inputs (sensors measurements of MW brightness temperatures). For each realization, $\boldsymbol{y^i}$, where the index $\boldsymbol{i}$ denotes the index, we perturb the sensor MW brightness temperatures, $\boldsymbol{y}$ by $\boldsymbol{\epsilon_r^i} \sim \boldsymbol{NE\Delta T} \cdot \boldsymbol{N}(0,1)$, where $\boldsymbol{NE\Delta T}$is the instrument noise equivalent delta temperature,

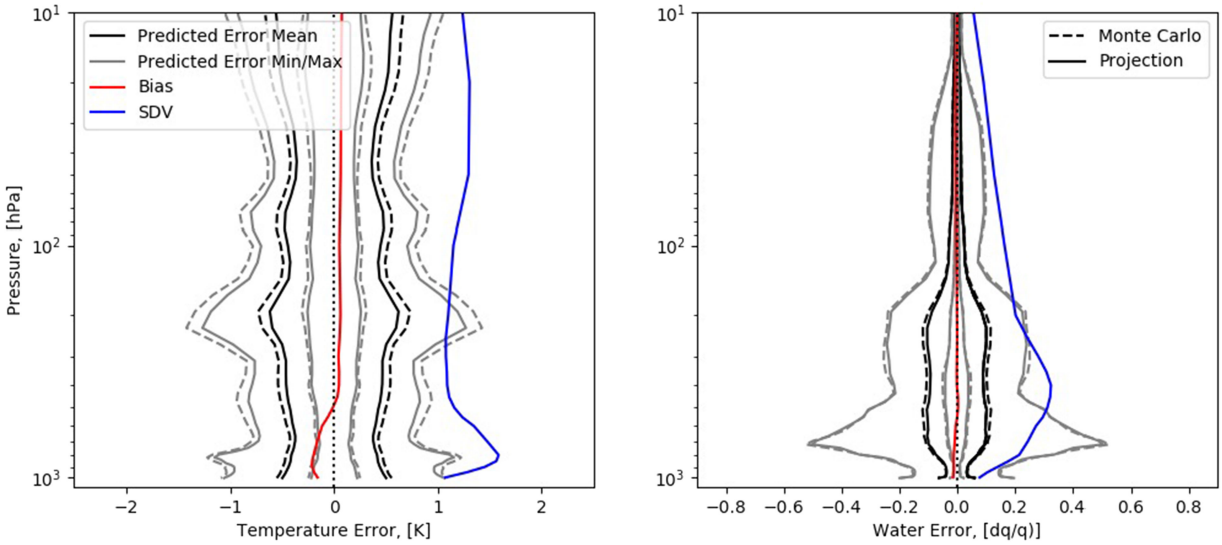


Fig. 8. Performance of temperature (left) and moisture (right) retrievals based MIIDAPS-AI, along with the errors quantification, assessed and predicted. Both biases (red) and standard deviations (blue) are shown and correspond to the true errors (retrievals minus truth). The mean predicted error is shown in black while the min-max envelope is shown in grey. The two approaches to predict the errors (Monte-Carlo and Projection) are shown in dash and solid lines respectively. See text for more details.

and $\boldsymbol{N}(0,1)$ is a normal random number with zero mean and unit standard deviation, i.e., $\boldsymbol{y^i} = \boldsymbol{y} + \boldsymbol{\epsilon_r^i}$.

We then run MIIDAPS-AI to produce a retrieval corresponding to the perturbed input, that is $\boldsymbol{x^i} = \boldsymbol{f_\theta}(\boldsymbol{y^i})$. The Monte-Carlo estimate of retrieval error can then be computed as the standard deviation of all MIIDAPS-AI realizations.

2) Linear Monte Carlo/Analytic Projection Error Estimation Approach: Linear Monte Carlo/Analytic projection error estimation approach: In this approach, we use the automatic differentiation available from the Keras/Tensorflow backprop algorithm to compute the saliency or gradient of MIIDAPS-AI outputs with respect to the inputs, $\boldsymbol{J_{f_\theta(y)}}$, and then project the noise realizations (as defined for the Monte-Carlo approach) from radiometric space to geophysical space, i.e., we multiply the matrix $\boldsymbol{J_{f_\theta(y)}}$ by the vector of perturbations $\boldsymbol{\epsilon_r^i}$. The analytical estimate of the retrieval error is then the standard deviation of those projections.

Fig. 8 shows that the errors computed with the NN-Jacobian *Projection method* and those with *Nonlinear Monte-Carlo* are very similar. However, both estimated (or predicted) errors underestimate the actual errors (assessed by comparing to truth). This might be anticipated since the errors estimated with above two approaches represent only a portion of the actual error in MIIDAPS-AI products. Other potential error components are the limited information content of the instrument measurements (i.e., finite instrument vertical resolution) and irreducible errors due to the one-to-many/ill-posedness of inverse remote sensing problems. In other words, the *projection* and *Monte-Carlo* methods only assess the random components of MIIDAPS-AI error due to instrument noise. Actual MIIDAPS-AI retrieval errors that depend on optimality of the network architecture, weights (training) as well as instrument information content, are not assessed by these methods. Nevertheless, the max/min predicted retrieval error indicates that the information content of the retrievals provided by retrieval Jacobians exhibit case dependence.

VI. Conclusion

MIIDAPS-AI is a very efficient retrieval algorithm for microwave and infrared sensors on both LEO and GEO satellites and produces simultaneously (in a single retrieval) a suite of geophysical products. This simultaneous approach is shown to force the inverted space vector to be physically consistent. This algorithm leverages modern AI (and ML) techniques and was developed and optimized using Keras/TensorFlow. Both the science and efficiency performances of the algorithms were briefly assessed, using both simulated and real data from joint polar sounding system (JPSS) and GOES satellites. More advanced evaluation, digging deeper into the performance of all the geophysical parameters (e.g., surface parameters such as sea ice and soil moisture; trace gases such as CO and ozone; and, cloud/precipitation parameters) inverted by MIIDAPS-AI was not shown but will be the subject of future reports. We have shown however that the accuracy and precision of the atmospheric profiles of temperature and moisture are overall similar to those of operational algorithms, but with a computational efficiency gain of 100. This opens the door to a multitude of possibilities in the future, chief among them, the ability to process orders of magnitude more observations, from satellites including smallsats and cubesats.

To build further confidence in the retrievals, the spatial variability (of geophysical analyses fields retrieved by MIIDAPS-AI from microwave sensor ATMS) and the inter-parameter correlation of the MIIDAPS-AI state vector were also assessed and found to be of acceptable quality. To avoid using the AI-based optimization as a *black box*, we showed how the ML retrieval process operates by computing the temperature and moisture

sensitivities to the satellite observations (applied to microwave ATMS), and found that the ML-based Jacobians are very similar to those sensitivities expected from a 1DVar approach (e.g., contribution functions), therefore, giving us trust in the results and providing evidence that the MIIDAPS-AI algorithm is extracting the right information content from the appropriate channels measurements present in the inputs. We also proposed two ways to quantify the random errors of the retrievals due to channel noise. Random errors predicted using those methods underestimated the real total uncertainty because those methods do not account for other noise sources. The assessments of the interpretability of MIIDAPS-AI presented in this article were done more extensively than what is succinctly presented in this study to keep the size of this high-level overview paper manageable. Other more dedicated publications (for more sensors, and more parameters) will describe more extensive validation work.

The MIIDAPS-AI, because it generates cloud, ice, and precipitation information, could be used as part of a preprocessing step for data assimilation to QC satellite data. Additionally, geophysical parameters currently used as fixed boundary conditions in NWP systems, such as emissivity and trace gases, could be obtained from MIIDAPS-AI, thereby providing dynamically varying *boundary conditions based on real observational data*, which would be an improvement over the fixed atlases that are sometimes used. Note that depending on the application or the NWP system, the emissivity could be used as a standalone variable to define the surface boundary or could be used as an atmosphere-cleared signal to derive geophysical surface variables such as sea-ice, soil moisture, and vegetation. MIIDAPS-AI is easily expandable to new and hypothetical sensors, as long as CRTM is also extended for these sensors. An example of this expansion is described in [22] where a configuration of many 12U smallsats of microwave sensors called EON-MW was evaluated for their capabilities and value. Significant efficiency coupled with good quality products, or at least similar quality to that obtained with operational systems, and ways to scientifically explain the results and quantify the errors, makes MIIDAPS-AI an attractive candidate algorithm for operational (i.e., in real-time) processing of microwave and infrared sensors data, from both polar, and geostationary satellites. This will be especially true in the upcoming era of swarms of smallsats and cubesats, expected to be deployed in the future, and MIIDAPS-AI provides part of the answer to address the big data challenge.

ACKNOWLEDGMENT

The authors would like to thank the STAR NPROVS team, in particular Tony Reale and Bomin Sun, for providing the independent radiosonde comparisons, Ross Hoffman for providing insightful comments on the manuscript prior to publication, and Narges Shahroudi and Adam Neiss for their work in extending MIIDAPS-AI to hypothetical sensors.

REFERENCES

[1] S.-A. Boukabara *et al.*, "MiRS: An all-weather 1DVAR satellite data assimilation & retrieval system," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3249–3272, Sep. 2011.

[2] S.-A. Boukabara *et al.*, "A physical approach for a simultaneous retrieval of sounding, surface, hydrometeor, and cryospheric parameters from SNPP/ATMS," *J. Geophys. Res. Atmos.*, vol. 118, pp. 12600–612619, 2013, doi: 10.1002/2013JD020448.

[3] C. D. Barnet *et al.*, "The NOAA unique combined atmospheric processing system (NUCAPS) algorithm theoretical basis document. NOAA 2020," Accessed: May 4, 2021. [Online]. Available: https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_NUCAPS_v3.0.pdf

[4] B. Sun, A. Reale, F. H. Tilley, M. E. Pettey, N. R. Nalli, and C. D. Barnet, "Assessment of NUCAPS S-NPP CrIS/ATMS sounding products using reference and conventional radiosonde observations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 2499–2509, Jun. 2017.

[5] S.-A. Boukabara *et al.*, "Leveraging modern artificial intelligence for remote sensing and NWP: Benefits and challenges." *Bull. Am. Meteorol. Soc.*, vol. 100, no. 12, pp. ES473–ES491, Dec. 2019.

[6] S.-A. Boukabara, V. Krasnopolsky, J. Q. Stewart, S. G. Penny, R. N. Hoffman, and E. Maddy, "Artificial intelligence may be key to better weather forecasts," *Eos Trans. Amer. Geophys. Union*, vol. 100, 2019. [Online]. Avilable: https://doi.org/10.1029/2019EO129967

[7] F. Aires, W. B. Rossow, N. A. Scott, and A. Che´din, "Remote sensing from the infrared atmospheric sounding interferometer instrument 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles." *J. Geophys. Res.*, vol. 107, no. D22, pp. 1–12, 2000, doi: 10.1029/2001jd001591.

[8] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community." *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 042609, 54, doi:10.1117/1.jrs.11.042609.

[9] Y. Han *et al.*, "JCSDA community radiative transfer model (CRTM)—version 1," *NOAA Tech. Rep. NESDIS*, vol. 122, 2006, Art. no. 40.

[10] E. Jones, E. Maddy, K. Garrett, and S. - A. Boukabara, "The MIIDAPS algorithm for retrieval and quality control for microwave and infrared observations," in *Proc. 23rd Conf. Integr. Observing Assimilation Syst. Atmos. Oceans Land Surf. Appl. Data Assimilation*, Phoenix, AZ, 2019. [Online]. Available: https://ams.confex.com /ams/2019Annual/webprogram/Paper352855.html

[11] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Quart. J. Roy. Meteorol. Soc.*, vol. 146, pp. 1999–2049, 2020. [Online]. Available: https://doi.org/10.1002/qj.3803

[12] D. Wang *et al.*, "Surface emissivity at microwaves to millimeter waves over polar regions: Parameterization and evaluation with aircraft experiments," *J. Atmos. Ocean. Technol.*, vol. 34, no. 5, pp. 1039–1059, 2017, Accessed: May 3, 2021. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/34/5/jtech-d-16-0188.1.xml

[13] M. Kazumori and S. J. English, "Use of the ocean surface wind direction signal in microwave radiance assimilation," *Quart. J. Roy. Meteorol. Soc.*, vol. 141, no. 689, pp. 1354–1375, 2015.

[14] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: https://www.tensorflow.org/ Software Available: tensorflow.org

[15] S. W. Seemann, E. E. Borbas, R. O. Knuteson, G. R. Stephenson, and H. - L. Huang, "Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multi-spectral satellite radiance measurements," *J. Appl. Meteor. Climatol.*, vol. 47, pp. 108–123, 2008. [Online]. Available: http://journals.ametsoc.org/doi/abs/10.1175/2007JAMC1590.1

[16] N. R. Nalli, P. J. Minnett, E. Maddy, W. W McMillan, and M. D. Goldberg, "Emissivity and reflection model for calculating unpolarized isotropic water surface-leaving radiance in the infrared. 2: Validation using Fourier transform spectrometers," *Appl. Opt.*, vol. 47, pp. 4649–4671, 2008.

[17] D. P. Dee and S. Uppala, "Variational bias correction of satellite radiance data in the ERA-Interim reanalysis," *Quart. J. Roy. Meteorol. Soc.*, vol. 135, pp. 1830–1841, 2009.

[18] K. He; X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Conf. Comput. Vis. Pattern Recognition (CVPR)*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[19] T. Reale, B. Sun, F. H. Tilley, and M. Pettey, "The NOAA products validation system (NPROVS)," *J. Atmos. Ocean. Technol.*, vol. 29, no. 5, pp. 629–645, 2012, Accessed: May 4, 2021. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/29/5/jtech-d-11-00072_1.xml

[20] C. D. Rodgers and B. J. Connor, "Intercomparison of remote sounding instruments," *J. Geophys. Res.*, vol. 108, no. D3, 2003, Art. no. 4116, doi: 10.1029/2002JD002299.

[21] E. S. Maddy and C. D. Barnet, "Vertical resolution estimates in version 5 of AIRS operational retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2375–2384, Aug. 2008.

[22] Y. Zhou *et al.*, "A preliminary assessment of the value and impact of multiple configurations of constellations of EON-MW, a proposed 12U microwave sounder CubeSat for global NWP," *Tellus A, Dynamic Meteorol. Oceanogr.*, vol. 73, no. 1, pp. 1–26, 2021, doi: 10.1080/16000870.2020.1857143

Eric S. Maddy received the B.S. degree in physics, and B.A. degree in mathematics from Frostburg State University, Frostburg, MD, USA, in 2001, and the M.S. and Ph.D. degrees in atmospheric physics from the University of Maryland Baltimore County, Baltimore, MD, USA, in 2003 and 2007, respectively.

He is currently an expert on traditional and exploratory (artificial intelligence) polar and geostationary IR/MW remote sounding and data fusion. He has more than 17 years of experience working with National Oceanic and Atmospheric Administration (NOAA) center for SaTellite Applications and Research (STAR) leading projects on the development of carbon trace gases retrievals from operational hyperspectral sounders, fusing algorithms and products from multiple instrument platforms (CrIS/ATMS/VIIRS, IASI/AMSU-A MHS/AVHRR), and leading the development of artificial intelligence versions of those algorithms.

Sid A. Boukabara received the Ingénieur degree in electronics and signal processing from the National School of Civil Aviation (ENAC), Toulouse, France, in 1994, the M.S. degree in signal processing from the National Polytechnic Institute of Toulouse, Toulouse, France, in 1994, and the Ph.D. degree in remote sensing from the Denis Diderot University, Paris, France in 1997.

Until 2005, he was a Scientist with Atmospheric Environmental Research (AER) Inc., working on various projects including MonoRTM, NPOESS CrIS, CMIS, and QuikSCAT. He then joined the National Oceanic and Atmospheric Administration (NOAA) Center for SaTellite Applications and Research (STAR), where he has led over the years several efforts including the development of variational algorithms and data assimilation approaches to all-weather conditions and over all surface backgrounds for a variety of new sensors (including JPSS and GOES-R); observing systems impact assessments of existing and future satellites using OSE and OSSE tools including for small satellites; and environment data fusion tools to combine traditional data (space-based and conventional ground-based) with emerging new sources of environmental observations. After an assignment to NOAA headquarters, where he worked as a Senior Policy Advisor on policy-making and strategic planning related to satellite systems, he came back home to NESDIS/STAR to engage with strategic new initiatives and pilot projects to explore new technologies and new numerical approaches such as cognitive leaning, to handling big data, remote sensing algorithms, data assimilation, and data fusion.