ecsat

Atlas Building
Harwell-Oxford
Didcot
Oxfordshire OX11 0QX
United Kingdom
T +44 (0)1235 567900
HarwellCentre@esa.int
www.esa.int

DOCUMENT

Data Engineering Working Group (DEWG)

Report on Artificial Intelligence (AI) in the Climate Change Intiative (CCI)

Prepared by **EOP-SC**
Reference **ESA-ECO-EOPS-MNGT-PL-19-0214**
Issue **1**
Revision **0**
Date of Issue **06/02/2019**
Document Type **Technical Note**

Table of contents:

ACKNOWLEDGEMENTS

The ESA Climate Office thanks the facilitators and contributors to the survey -

Manuel Arias (Sea Surface Salinity CCI, ARGANS)
Annett Bartsch (Permafrost CCI, b.geos)
Victoria Bennett (CCI Open Data Portal, STFC)
Jérôme Bruniquel (LST CCI, ACRI-ST)
Lorenzo Bruzzone (HR Land Cover CCI, University of Trento)
Michael Buchwitz (GHG CCI, University of Bremen)
Rafael J. Catany (Sea Surface Salinity CCI, ARGANS)
Emilio Chuvieco (Fire CCI, University of Alcalá)
Wouter Dorigo (Soil Moisture CCI, TU Wien)
Matthias Forkel (Soil Moisture CCI, TU Wien)
Kevin Halsall (CCI Open Data Portal, Telespazio Vega)
Hugh Kelliher (SST CCI, Space Connexions)
Phil Kershaw (CCI Open Data Portal, STFC)
Frank Paul (Glaciers CCI, University of Zurich)
Thomas Popp (Aerosol CCI, DLR)
Stengel Martin (Cloud CCI, DWD)
Thomas Lavergne (Sea Ice CCI, Met.no)
Jean-Francois Legeais (Sea Level CCI, CLS)
Paul van der Linden (Met Office, CMUG)
Alison Waterfall (CCI Open Data Portal, STFC)

ESA Climate Office participants and authors -

Pascal Lecomte
Paolo Cipollini
Paul Fisher
Ed Pechorro
Simon Pinnock
Stephen Plummer
Nele Reyniers

EXECUTIVE SUMMARY

The European Space Agency (ESA) Climate Change Initiative (CCI) *Report on AI in CCI* documents the preliminary analysis and findings of a study on AI use across the CCI community, undertaken in the form of a survey of CCI projects through the CCI Data Engineering Working Group (DEWG). Work on this study began in December 2018. The purpose of this study is to -

1. Identify *current applications* of AI in individual CCI projects
2. Identify *future applications* of AI in individual CCI projects
3. Identify *cross-ECV* scientific challenges which may benefit from use of AI
4. Identify *resource gaps* to achieving of meaningful scientific work through AI

Survey

29 CCI projects were invited to participate in the survey. 14 took the survey, 14 did not respond, and 1 declined the invitation.

		1. Current	2. Future	3. Cross-ECV	4. Res-Gaps
ECVs Theme 1	Aerosol				
	Cloud		-	-	-
	Fire				
	GHG				
	Glaciers				
	IS Antarctic	d			
	IS Greenland	x			
	Landcover	x			
	Ocean Colour	x			
	Ozone	x			
	Sea Ice	x			
	Sea Level			-	-
	SST				
	Soil Moisture				
ECVs Theme 2	Biomass	x			
	HR Land Cover				
	Lakes	x			
	LST				
	Permafrost			-	-
	Salinity				
	Sea State	x			
	Snow	x			
	Water Vapour	x			
Cross-ECV	CMUG				
	RECCAP	x			
	SLBC	x			
Knowledge Exchange	KE - ODP				
	KE - Toolbox	x			
	KE - Vis. Corner	x			

	AI used, activity investigated or use identified & anticipated.
	AI not used, not planned or non-specific general interest only.
-	Question not responded.
x	No response from CCI project.
d	Invitation declined by CCI project.

Current Applications of AI

Six CCI projects are using AI as part of their CCI work - Aerosol (Neural Network to retrieve dust Aerosol Optical Depth from thermal infrared spectra), Cloud (cloud detection algorithm using Machine Learning), Fire (Machine Learning classifier used to detect burned areas for periods without hotspots), Sea Level (proofs of concept of Machine Learning algorithms for rain cells detection in satellite altimetry measurements), Soil Moisture (Machine Learning to estimate Level-4 products, Machine learning to evaluate global models, and identification of alternative structures for global land surface models) and Permafrost (Machine Learning for the post-processing of trends from land-surface properties).

Future Applications of AI

Eight CCI projects plan to use AI as part of their future work, or have identified specific use cases of AI in their work which they anticipate implementing - Cloud (continuing Neural Network usage for dust Aerosol Optical Depth record), Fire (plan to continue the use of Random Forest for burned area detection and the creation of validation data, plus testing other classification tools based on Deep Learning), GHG (use of Machine Learning for quality flagging of the S5P scientific WFMD algorithm XCH4 product is under investigation), Sea Level (AI & Machine Learning algorithms are anticipated to provide new insights and methods to tackle complex classification problems such as outlier detection), Soil Moisture (exploration of use of Machine Learning for data merging, filtering and gap filling for soil moisture and vegetation data), High Resolution Land Cover (AI and ML will be considered for the classification module of the processing chain and possibly for change detection), Permafrost (Machine Learning for a science case study, the plan being to integrate the CCI+ permafrost extent product with existing CCI+ LST products and Hot Spot Regions of the Permafrost Change Product from ESA GlobPermafrost), and Sea Surface Salinity (AI can have many potential applications, such as the connecting of different ECVs to obtain derived variables).

Cross-ECV Scientific Challenges Which May Benefit From AI

Five CCI projects expressed ideas on broader, cross-ECV, scientific challenges potentially beneficial through AI - Aerosol (analyzing the multi-decadal records for extreme events including those where several ECVs show extreme values, and testing the use of machine learning for a variety of integrated products), Fire (interaction between ECVs through cross-classification, and to improve efficiencies in archiving and visualization techniques), Soil Moisture (cross-CCI activities such as estimation of other variables, like photosynthesis, by combining various different CCI ECV datasets within an ML algorithm), High Resolution Land Cover (AI and Machine Learning methodologies to enhance data processing capabilities, in validation procedures, in climate modeling, and in defining querying systems for improving the interaction between the users and the generated products), and Sea Surface Salinity (in terms of science, the possibility to resolve multi-parametric models in which there is limited scientific knowledge, and in terms of engineering the deployment of AI for higher performance in data handling).

Resource Gaps in Achieving Meaningful Scientific Work Through AI

Three CCI projects identified potential resource gaps towards using AI - Fire (statistics expertise to support the environment sciences expertise within the CCI community), High Resolution Land Cover (computational power required to perform AI and Machine Learning) and Sea Surface Salinity (access to expertise within the CCI community).

1 DOCUMENT SCOPE

This document reports on the preliminary analysis and findings of a study on AI use across the CCI community, undertaken in the form of a survey of CCI projects through the CCI Data Engineering Working Group (DEWG).

The document is organised as follows -

- An *Executive Summary* expressing the response to the survey, and bringing together the responses across CCI projects to each of the survey questions.
- This *Document Scope* (§1) defining the purpose of the report and describing document structure.
- A *section per CCI project* (§2 to §30) providing verbatim responses for each of the survey questions.

2 AEROSOL

2.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

One of the four IASI retrievals (by ULB - Universite Libre Brusseles) used in Aerosol_cci2 applies a neural network to retrieve dust AOD from the IASI thermal infrared spectra; this approach showed best validation results against AERONET coarse mode AOD, but provides no other physical aerosol parameters as output as the three other physical retrievals do (effective layer height, effective dust size, mineralogical composition).

2.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

Confirmed future use is continuing the ULB dust AOD record in C3S_312b_Lot2 (not co-funded in Aerosol_cci+).

I would wish to work on combining the neural network with a physical retrieval to integrate the best Dust AOD accuracy with the additional parameters – but such work is so far not funded or committed.

2.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

- Analyzing the multi-decadal records for extreme events, in particular those where several ECVs show extreme values.
- Testing the use of machine learning for integrated products of (1) visible radiometers (e.g. AATSR/SLSTR which can provide good AOD and fine mode AOD), UV spectrometers (e.g.

OMI, which has sensitivity to absorbing aerosols), and thermal infrared spectrometers (e.g. IASI with sensitivity to mineral dust / coarse AOD).

2.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

No dedicated funding for testing such new approaches has been foreseen so far (and if so it would compete with continuing other development lines).

3 CLOUD

Response from Technical Officer -

DWD cloud detection algorithm which uses Machine Learning.

ESA Cloud_cci. Algorithm Theoretical Baseline Document v5.1. (Applicable to Cloud_cci version 2.0 products). Issue 5 Revision 1. 22 February 2018. Deliverable D-2.1.

ESA Cloud cci. Algorithm Theoretical Baseline Document v5.1. Community Cloud retrieval for Climate (CC4CL). (Applicable to Cloud cci version 2.0 products). Issue 5 Revision 1. 28/8/2017. Deliverable D-2.1.

4 FIRE

4.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

Response from Technical Officer -

In Fire CCI AI was used for SAR processing. The Sentinel-1 data was processed using open-source libraries available in the Orfeo ToolBox (OTB), an image processing software developed by the National Centre for Space Studies (CNES), France (Inglada and Christophe 2009). The OTB-based processing chain uses Ground Range Detected (GRD) Sentinel-1 images with the SAR data being tiled to 100 km. The chain is highly scalable and autonomous, once few parameters are set, and includes data download from Sentinel-1 repositories. The SAR data processing may be grouped in several steps including, pre-processing, geocoding and temporal filtering.

The SAR BA algorithm used as input temporal series of C-band dual polarized (VV and VH) data. The algorithm was locally adaptive, autonomous, and was designed to cope with variable data availability. The algorithm applied the Reed-Xiaoli detector (RXD) to distinguish anomalous changes of the backscatter coefficient. Such changes were related to fire events using information on thermal anomalies (hotspots) acquired during the detection period by different sensors including MODIS and VIIRS. Land cover maps were used to account for changing backscatter behaviour as the RXD is class dependent. A machine learning classifier (i.e., Random forests) is used to detect burned areas for periods without hotspots (Belenguer-Plommer et al. 2018).

Belenguer-Plommer, M.A., Tanase, M.A., Fernández-Carrillo, A., & Chuvieco, E. (2018). Burned area detection and mapping using Sentinel-1 backscatter and thermal hotspots: a global approach. Remote Sensing of Environment, in review.

Miguel A. Belenguer-Plomer, Emilio Chuvieco, Mihai A. Tanase. (2017). Sentinel-1 based algorithm to detect burned areas. 11th EARSeL Forest Fires SIG Workshop.

Inglada, J., & Christophe, E. (2009). The Orfeo Toolbox remote sensing image processing software. In, Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009 (pp. IV-733-IV-736): IEEE.

Response from team; response which follow in §4.2, §4.3, and §4.4 are also from team.

Within Fire_cci, we have used the Random Forest (RF) classifier, which is generally considered as a machine learning tool, for different applications:

1. Generation of Reference BA perimeters used for validation purposes within Fire-CCI Phase 2. The objective was to provide BA perimeters at 30 m spatial resolution in order to validate different global BA products, as well as the SAR BA product generated within Option 3 over the Amazon basin. RF classification was complemented with visual inspection to detect and correct errors. Different spectral bands were used to train and classify the Landsat and Sentinel-2 images.
2. Classification of the BA from Sentinel-1 radar images as part of the classification algorithm (which included other steps). Part of Fire_cci Option 3.

To classify burned pixels based on AVHRR-LTDR data at 0.05 degree resolution. We computed monthly RF models, using as inputs a biophysical LTDR index for several years when both LTDR and MCD64A1 BA product were available. MCD64A1 was used as reference data. A binary classification was performed (burned/unburned), as it provided better results than using proportion intervals. The criteria to select the final probability threshold for RF classification was based on the Dice Coefficient (DC), previously used for BA assessment within the Fire_cci project. The analysis was done with the whole common time series between the MCD64A1 and LTDR (2001 to 2015). The median monthly value of the RF probability providing the best DC output for the full time series was used as a threshold for the final RF classification.

4.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

Within Fire_cci+ we plan to continue the use of Random Forest for burned area detection and the creation of validation data. Complementary, we will test other classification tools based on ML concepts, such as deep-learning neural nets. We plan also to use machine learning techniques for merging of different BA products at BA level and at reflectance level.

4.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

1. Interaction between ECV, cross-classification, temporal series.
2. It will be useful for archiving and visualization techniques, to improve efficiency and speed.

4.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

Most of our researchers come from environmental science fields, they are not statisticians neither mathematicians, and therefore we are not fully exploring the potentials of new ML tools.

5 GHG

5.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

Currently not used.

5.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

Under investigation: Machine Learning (random forest) for quality flagging of the S5P scientific WFMD algorithm XCH4 product, which is under development. Too early to give details.

5.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

I cannot give good recommendations at this stage.

5.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

General: See 3 [§5.3]. GHG-CCI+: no gaps have been identified.

6 GLACIERS

6.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

We are currently not using AI or ML in Glaciers_cci.

6.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

We are currently not planning to use AI or ML in Glaciers_cci.

6.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

There are certainly many, but our wish list has currently other priorities (e.g. accurately orthorectified Sentinel-2 images with a DEM that is freely available :)

6.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

None as we have other priorities.

7 ICE SHEET ANTARCTIC

Survey invitation declined.

8 ICE SHEET GREENLAND

No response to survey invitation.

9 LAND COVER

No response to survey invitation.

10 OCEAN COLOUR

No response to survey invitation.

11 OZONE

No response to survey invitation.

12 SEA ICE

No response to survey invitation.

13 SEA LEVEL

13.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

Regarding satellite altimetry, there is still little use of AI & ML in data processing or analysis, with the exception of radiometer data processing which has been based on neural networks for years. Studies involving AI & ML are rather proofs of concept: for example, we tested ML algorithms for rain cells detection in satellite altimetry measurements. The goal is to be able to remove measurements affected by rains from further analysis, this is a well formulated classification problem and results are promising but still need to be confirmed by in depth validation.

13.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

AI & ML algorithms are anticipated to provide new insights and methods to tackle complex classification problems that arise in remote sensing data processing such as outlier detection, attribution. With the ongoing move towards higher resolution products, methods used on low resolution data are inefficient and ML algorithms should be tested, as they may be performing well.

14 SEA SURFACE TEMPERATURE

14.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

SST CCI does not currently use AI.

14.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

We have no plans to use it in the future.

14.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

We do not think that AI is applicable to CCI in general but other projects may identify use cases that we are not aware of.

14.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

Since we see no applications for AI and machine learning in CCI, we do not believe CCI resources should be spent in these areas.

15 SOIL MOISTURE

15.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

Response from team; response which follow in §4.2, §4.3, and §4.4 are also from team.

We currently use machine learning for three purposes:

1. *Machine learning to estimate level-4 products*. We use microwave vegetation data (VOD, vegetation optical depth) to estimate photosynthesis. Thereby we use a machine learning algorithm (GAM, generalized additive models) that we train against in situ data with predictor data from the satellite observations.
2. *Machine learning to evaluate global models*. We aim to understand the reasons for good or bad performances of global climate/carbon cycle models or dynamic global vegetation models. The basic idea is that we fit machine learning algorithms to satellite and climate data, and to model outputs. Then we adapted an approach to derive response functions from the machine learning algorithm to see what kind of relationships the algorithm leaned from the data or models. In a first application we evaluated models from the Fire Model Intercomparison Project by using random forest and different CCI datasets: https://www.biogeosciences-discuss.net/bg-2018-427/
3. *Identification of alternative structures for global land surface models*. We also aim to identify better model structures (for vegetation and fire models) by investigating what relationships we can learn from satellite observations. A related paper is here: https://www.geosci-model-dev.net/10/4443/2017/

Response from ESA Climate Office member -

Papagiannopoulou et al. applied machine learning to different variables/ECV's in order to investigate non-linear relationships between them. They used many different data sources, including the CCI passive and combined soil moisture datasets as well as GlobSnow.

Christina Papagiannopoulou, Diego G. Miralles, Stijn Decubber, Matthias Demuzere, Niko E. C. Verhoest, Wouter A. Dorigo, and Willem Waegeman. Geosci. Model Dev., 10, 1945–1960, 2017 www.geosci-model-dev.net/10/1945/2017/, doi:10.5194/gmd-10-1945-2017.

15.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

We currently explore if we can use ML for data merging, filtering and gap filling for soil moisture and vegetation data. This is very early work, so we don't have more details.

15.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

Cross-CCI activities benefit from AI&ML. For example, we could estimate other variables (like photosynthesis) be combining various CCI datasets (soil moisture, land cover ...) within a ML algorithm.

15.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

I cannot provide you answers to this question.

16 BIOMASS

No response to survey invitation.

17 HIGH RESOLUTION LAND COVER

17.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

HRLC ECV started its activities on September 25th, 2018. Thus, currently we do not have any use of AI or machine learning because we are still in the phase of design of the processing chain.

17.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

HLRC aims at extracting land cover and land-cover change maps from both optical and SAR images. The processing chain includes several steps (pre-processing, geo-location, classification, data fusion, multitemporal time series analysis, etc.). We are currently evaluating the possibility to use AI and ML for achieving required performance with high accuracy. AI and ML will be especially considered for the classification module and possibly for the change detection module of the processing chain.

17.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

Many components could benefit from AI and machine learning, In general these methodologies can enhance the data processing capabilities and the accuracy for the extraction of various essential climate variables. Moreover, AI and machine learning can be integrated in the validation procedures, in the climate modeling, and also in defining querying systems for improving the interaction between the users and the generated products (on this we have strong know how in the team even if this is not a goal of the current project).

17.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

Recent machine learning (including deep learning) and AI are data hungry tools in terms of training samples and are characterized by a heavy computational load (especially in the training phase). In many cases GPU-based processing architectures are more suitable than standard CPU-based architectures for deep learning algorithms. Since CCI is tackling historical remote sensing time series with global coverage, both aforementioned elements are potentially highly critical and should be analyzed carefully.

18 LAKES

No response to survey invitation.

19 LAND SURFACE TEMPERATURE

Response -

We are not using AI for LST CCI and we have no plan for using AI and certainly not in the short term. We are of course interested to have a better understanding about the possible use of AI and machine learning from the presentations which will be given during the next co-location meeting.

20 PERMAFROST

20.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what was the science objective, and how did you address that objective. The more details the better!"

ML is of relevance to one of our application cases. Our products will be compared to trends from landsurface properties derived from Landsat. ML is used in postprocessing of these trends.

ML methodologies include machine-learning based classification of robust trends of multi-spectral indices of Landsat data (TM, ETM+, OLI) and object-based lake detection, to analyze and compare the individual, local and regional lake dynamics of four different study sites (Alaska North Slope, Western Alaska, Central Yakutia, Kolyma Lowland) in the northern permafrost zone from 1999 to 2014 (Nitze et al 2017, Remote Sensing).

Nitze I, Grosse G, Jones BM, Arp CD, Ulrich M, Fedorov A, Veremeeva A (2017): Landsat-based trend analysis of lake dynamics across northern permafrost regions. Remote Sensing, 9(7): 640. doi: 10.3390/rs9070640.

ML methodologies include temporal trend-analysis with machine-learning to map the spatial distribution of Permafrost Region Disturbances (lake formation, lake expansion, lake drainage, thaw slumps, and fire scars) and their relation to permafrost properties, ecological zones and climate, allowing comprehensive and unique insights into PRD distribution, abundance, and dynamics as well as permafrost vulnerability to thaw for large permafrost regions in Siberia, Alaska, and Canada (Nitze et al., 2018, Nature Communications).

Nitze I, Grosse G, Jones BM, Romanovsky VE, Boike J (2018): Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic. Nature Communications, 9: 5423. doi: 10.1038/s41467-018-07663-3.

20.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

Continuation of ML use for a science case study in CCI+ Permafrost, where we plan to integrate the CCI+ permafrost extent product with existing CCI+ LST product and Hot Spot Regions of Permafrost Change Product from ESA GlobPermafrost across 4 major continental-scale transects to identify correlations and drivers of permafrost change.

21 SALINITY

21.1 Current Use of AI

"What are your current uses of AI & Machine Learning in CCI. In each case, what what was the science objective, and how did you address that objective. The more details the better!"

In CCI+ Salinity we do not use and it is not foreseen the use of AI / Machine learning. However, at SMOS L1, there are initiatives towards potential implementation of AI algorithms for RFI detection. It is in embryonic state and not directly specific for CCI, but something to consider as example of application.

21.2 Future Use of AI

"What are your planned or anticipated uses of AI & Machine Learning in CCI. In each case, what is the science objective, and how do you anticipate addressing that objective. Again, details are good."

AI can have many potential applications from a CCI's perspective:

- Quality control of input/output datasets: Detection of data anomalies, recognition of artificial patterns into the data (e.g. to reveal not well known or unknown dependencies of auxiliary data, as wind speed)
- Information discrimination processes: Of interest for those CCIs handling large volume of data, Machine Learning can help to asses with better accuracy the weight of specific ancillary data in the quality of the final products (i.e. identifying variables of the models that could be redundant/unnecessary), or just remove data of no value for the production (e.g. cloud and shadows detection in ocean color input data sets) before they enter into the data processors.
- Specific retrieval algorithms: There are already algorithms based in NNs for some variables (e.g. soil moisture). AI could be brought to L2 to L4 algorithms to optimize computing times and improve retrieval quality in some cases.

- Retrieval of derived variables or sub-ECVs: There is an entire universe of possibilities here, including the potential to connect different ECVs to obtain derived variables.

21.3 Broader AI Challenges

"More widely, across the whole CCI programme, what other CCI science and CCI engineering challenges do you think could benefit from AI and Machine Learning."

In terms of science, as said above, the possibility to resolve multi-parametric models in which we do not have a full scientific knowledge (e.g. AI to support El Nino/La Nina forecasting based in CCI data, or extreme events early detection). In terms of engineering, AI algorithms usually deploy a much higher performance in data handling and retrieval of information. When possible, their implementation should reduce computing times and IT resources necessary for the production of the ECV fields.

21.4 Resource Gaps For AI

"What gaps do you see as existing between (i) the CCI challenges expressed in 1, 2 and 3 above, and (ii) the CCI resources for AI & Machine Learning to meet those challenges."

In my view, the major challenge is set on the access to the right expertise. Data Scientists are usually very well paid in other sectors. Academic community has expressed their concern quite often about how good students are hired by big companies (Apple, Google, Uber, banks…) before they finish their projects, and the issue is that salaries in these other sectors are considerably higher. Which means that any scientist properly trained in AI/Machine Learning can be difficult to keep within the EO community. Apart from that, there is a lack of specific calls from ESA to address questions as the ones mentioned above, so CCI –specific initiatives are more than welcome, which is of course the full point of this initiative.

22 SEA STATE

No response to survey invitation.

23 SNOW

No response to survey invitation.

24 WATER VAPOUR

No response to survey invitation.

25 CMUG

Response -

We don't have plans to use AI in the current programme of work.

26 RECCAP

No response to survey invitation.

27 SEA LEVEL BUDGET CLOSURE

No response to survey invitation.

28 KNOWLEDGE EXCHANGE - OPEN DATA PORTAL

Response out of scope.

29 KNOWLEDGE EXCHANGE - TOOLBOX

No response to survey invitation.

30 KNOWLEDGE EXCHANGE - VISUALISATION TOOL

No response to survey invitation.