



land surface
temperature
cci



CCI Land Surface Temperature

Algorithm Development Plan

WP2.1 – DEL-D2.4

Ref.: LST-CCI-D2.4-ADP

Date: 16-Mar-2023

Organisation: Consortium CCI LST





Signatures

	Name	Organisation	Signature
Written by	Darren Ghent	ULeic	
	Mike Perry	ULeic	
	Karen Veal	ULeic	
	Carlos Jimenez	Estellus	
Reviewed by	Darren Ghent	ULeic	
Approved by	Darren Ghent	ULeic	
Authorized by	Simon Pinnock	ESA	

Change log

Version	Date	Changes
1.0	24-Jun-2019	First version
2.0	19-Dec-2020	Second version detailing changes for Processing Cycle 2
3.0	01-Dec-2021	Third version detailing changes for Processing Cycle 3 (Phase-2 Cycle 1)

List of Changes

Version	Section	Changes
2.0	3.1.2	Details on the single-channel (SMW) algorithm
2.0	3.2.2	Description of auxiliary dataset changes, intercalibrations and time correction approach
2.0	4.1.2, 4.1.3	Summary of the MW uncertainty work
3.0	3.1.3	Details on the Optimal Estimation (OW) algorithm



3.0	3.2.3	Description of auxiliary dataset changes, orbital drift correction approach, and high resolution products
3.0	4.1.3, 4.2.3	Details on new MW algorithm approach incorporating Optimal Estimation
4.0	3.1.4	Details on the new Hybrid Algorithm
4.0	3.2.4	Updates to the cloud masking, intercalibration and orbital drift correction

Table of Content

1. EXECUTIVE SUMMARY	1
2. INTRODUCTION	5
2.1. Purpose and Scope	8
2.2. Reference Documents	8
2.3. Glossary	12
3. ALGORITHM DEVELOPMENT PLAN FOR INFRARED LST PRODUCTS	14
3.1. Current Status of Infrared LST Products	14
3.1.1. Version 2.00	14
3.1.1.1. UOL Algorithm	14
3.1.1.2. Generalised Split Window (GSW) Algorithm	15
3.1.2. Version 3.00	15
3.1.2.1. Single-channel (SMW) Algorithm	15
3.1.3. Version 4.00	16
3.1.3.1. Optimal Estimation (OE) Algorithm	16
3.1.4. Version 5.00	18
3.2. Algorithm Development Plan for Infrared Sensors	18
3.2.1. Version 2.00	18
3.2.1.1. Radiative Transfer Modelling	18
3.2.1.2. Calibration Database for Determining Retrieval Coefficients for the TIR Algorithms	19
3.2.1.3. Algorithm Stratification	20
3.2.1.4. Identification of Observations Valid for Land Surface Temperature Estimation from Thermal Infrared Sensors	20
3.2.1.5. Auxiliary Datasets for Thermal Infrared Retrievals	22
3.2.1.6. Climate Data Records	24
3.2.1.7. Uncertainty Model for Thermal Infrared Algorithms	27
3.2.2. Version 3.00	27
3.2.2.1. Identification of Observations Valid for Land Surface Temperature Estimation from Thermal Infrared Sensors	28
3.2.2.2. Auxiliary Datasets for Thermal Infrared Retrievals	28
3.2.2.3. Climate Data Records	29
3.2.2.4. Uncertainty Model for Thermal Infrared Algorithms	34
3.2.3. Version 4.00	34
3.2.3.1. Auxiliary Datasets for Thermal Infrared Retrievals	34
3.2.3.2. Climate Data Records	35
3.2.3.3. Prototype Products	36
3.2.4. Version 5.00	37
3.2.4.2. Climate Data Records	37
4. ALGORITHM DEVELOPMENT PLAN FOR MICROWAVE LST PRODUCTS	40
4.1. Current status of Microwave LST Products	40
4.1.1. Version 2.00	41
4.1.2. Version 3.00	41
4.1.3. Version 4.00	41
4.1.4. Version 5.00	42
4.2. Algorithm Development Plan for Microwave Sensors	42
4.2.1. Version 2.00	42
4.2.2. Version 3.00	43
4.2.3. Version 4.00	43
4.2.4. Version 5.00	43

List of Figures

Figure 1: Data flows for LST_cci product prototype production system for thermal infrared sensors. -----	5
Figure 2: Data flows for the SSM/I and SSMIS LST ECV prototype production system. -----	6
Figure 3: Retrieval flow chart for the OE method applied to TOA BTs and external LSE data. -----	17
Figure 4: Schematic representation of the development and implementation of the ATSR-SLSTR CDR. -	27
Figure 5: Monthly mean IASI minus Terra MODIS 11 micron (left) and 12 micron (right) BT difference (K) in the Antarctic. The red line is a straight line fit to data from January 2009 to December 2018. An exponential fit to the same data (black) and a 11 (blue) and 23 (yellow) month running means are also shown.-----	30
Figure 6: Monthly mean IASI minus Aqua MODIS 11 micron (left) and 12 micron (right) BT difference (K) in the Antarctic. The red line is a straight line fit to data from January 2009 to December 2018. An exponential fit to the same data (black) and a 11 (blue) and 23 (yellow) month running means are also shown.-----	31
Figure 7: Schematic representation of the development and implementation of the Merged IR CDR.---	33
Figure 8: AVHRR Equator Crossing Times. -----	35
Figure 9: Data flows for the SSM/I and SSMIS LST ECV prototype production system. -----	40

List of Tables

Table 1: Algorithm Developments enumerated and described. The status of each development is indicated by colour: white (not started), orange (in progress), green (implemented in next Processing Cycle), grey (cancelled – with reason given). New developments are indicated in red. -----	1
Table 2: Processing Cycle and Product Numbering nomenclature. -----	6
Table 3: Land Cover CCI biome definition.-----	22

1. Executive Summary

The European Space Agency Climate Change Initiative on Land Surface Temperature (hereafter LST_cci) aims to provide Land Surface Temperature (LST) LST Essential Climate Variable (ECV) products and validate these data to provide an accurate view of temperatures across land surfaces globally over the past 20 to 25 years.

This Algorithm Development Plan (ADP) provides the details on the expected evolutions to the current and expected Land Surface Temperature (LST) products available from processing Cycle 1 or pre-cursor studies. It includes planned developments to:

- ❖ The algorithms themselves
- ❖ The necessary auxiliary data for best implementation of the algorithms
- ❖ The calibration database for determining retrieval coefficients
- ❖ The radiative transfer models
- ❖ The cloud masking schemes

The algorithms under development in the next processing cycles will be those selected from the Round Robin intercomparison exercise [RD-30]. These were identified as the best algorithms for a future climate quality operational system. Nonetheless, it is expected that a combination of evolutions to each of the points above will improve the implementation of these algorithms. The evolutions outlined in this document will be implemented in an end-to-end system to generate the first LST_cci climate data records.

It is also important to note that this document will be regularly updated as new information comes to light from the feedback expected from the validation and intercomparison, and from the climate assessment of the LST CCI products.

Table 1: Algorithm Developments enumerated and described. The status of each development is indicated by colour: white (not started), orange (in progress), green (implemented in next Processing Cycle), grey (cancelled – with reason given). New developments are indicated in red.

Algorithm Development Number	Description	Status	Comments
LST-CCI-ADP-1	Upgrade to RTTOV 12.3 for development of all LST_cci LEO TIR ECV Products and IR Climate Data Records (CDRs)		Implemented for V2.00 Products
LST-CCI-ADP-2	An extended version of the Benchmark database using ERA5 Atmospheric profile Data, CAMEL Emissivity Data and ESA CCI Land Cover data will be used to determine retrieval coefficients for TIR algorithms		Implemented for V2.00 Products
LST-CCI-ADP-3	All GEO products will be reprocessed with a GSW calibrated with the benchmark dataset constructed for the Round Robin, in order to provide harmonized data for the Merged Product		Superseded by LST-CCI-ADP-18

Algorithm Development Number	Description	Status	Comments
LST-CCI-ADP-4	Temperature-dependent coefficients will be explored in detail for the UOL algorithm for the next processing cycle to minimise any possible impact of non-linearity		Investigation still be explored – now scheduled for Phase-2
LST-CCI-ADP-5	Development for the GSW algorithm will focus on improved derivation of the retrieval coefficients using the new calibration database, and improved input auxiliary datasets		Implemented for V2.00 Products
LST-CCI-ADP-6	All development will move to using ERA5 data, which will minimise any non-linearity between adjacent time steps since the profiles are only 1-hour apart		Implemented for V3.00 Products
LST-CCI-ADP-7	All retrieval algorithm and cloud masking developments will move to using land cover information from Land Cover CCI		Implemented in V2.00 Products
LST-CCI-ADP-8	Class 33 (bare soil) of the Land Cover CCI classification will be sub-divided into distinct sub-classes based on soil taxonomy		Implemented in V2.00 Products
LST-CCI-ADP-9	We will continue to use FCOVER output from the Copernicus Global Land Cover Services. Since there is no FCOVER output from the CCI programme using such a dataset is not inconsistent with the wider CCI programme		
LST-CCI-ADP-10	In the next Cycle the Combined ASTER and MODIS Emissivity for Land (CAMEL) database will be used to calculate LST		Implemented for V2.00 Products
LST-CCI-ADP-11	It is proposed to move towards using ESA Snow Cover CCI products for snow masking providing they become available for future Cycles. If these are not available in time for re-processing of the LST_cci products then the IMS remains the default		ESA Snow Cover CCI products they are scheduled for Phase-2
LST-CCI-ADP-12	In the processing of the ATSR-SLSTR CDR the overlap analysis for the next Cycle will include SLSTR		Implemented in V3.00 Products
LST-CCI-ADP-13	Diurnal information from the high temporal resolution geostationary satellites for each land cover class will be incorporated into the time difference assessment for the ATSR-SLSTR CDR		The time correction has focused on a wide temporal window of overlap between Terra-MODIS and ATSR/SLSTR respectively

Algorithm Development Number	Description	Status	Comments
LST-CCI-ADP-14	The ATSR-SLSTR CDR will implement the best performing cloud mask from WP2.4 developments		Delays to WP2.4 mean the probabilistic cloud mask has been implemented
LST-CCI-ADP-15	The Merged IR CDR will implement the GSW algorithm for all input sensors to ensure consistency		Superseded by LST-CCI-ADP-18
LST-CCI-ADP-16	The next production Cycle 2 will begin to implement concepts from the standardised uncertainty approach adopted for IT sensors to the microwave sensors		Implemented in V3.00 Products
LST-CCI-ADP-17	The Quality flag Processor for microwave sensors will be improved for: (a) the inundation flag, by using actual estimates of inundation; (b) the snow flag, by using the ESA CCI snow product if it becomes available; and (c) the convection flag, by revising the current cloud detection algorithm		Implemented for V2.00 Products
LST-CCI-ADP-18	All LEO and GEO products will be reprocessed with a single-channel (SMW) algorithm calibrated with the benchmark dataset constructed for the Round Robin, in order to provide harmonized data for the Merged Product		The agreed approach is to now utilise the maximum information available per instrument rather than degradation
LST-CCI-ADP-19	Apply intercalibrations to the Level-1 BTs for LEO data which is input to the ATSR-MODIS-SLSTR CDR		Implemented in V3.00 Products
LST-CCI-ADP-20	For input to the ATSR-MODIS-SLSTR CDR apply time corrections to the Level-3U LST for AATSR and SLSTR to bring to same nominal local overpass time as ATSR-2 and Terra-MODIS using Terra-MODIS as the reference sensor		Implemented in V3.00 Products
LST-CCI-ADP-21	Implement the uncertainty component due to TCWV on the application of the coefficients for all LEO and GEO infrared products		Implemented in V3.00 Products
LST-CCI-ADP-22	Exploit the capabilities of advanced algorithms, such as Optimal Estimation, to confront the challenges of higher resolution sensors and scale change		Being implemented in the prototype high resolution products



Algorithm Development Number	Description	Status	Comments
LST-CCI-ADP-23	All LEO and GEO products will be reprocessed with the best algorithm to maximise the capabilities of each instrument rather than degrade to common algorithms		Implemented in V4.00 (Phase-2)
LST-CCI-ADP-24	Implement an orbital drift correction method for NOAA-AVHRR LST products and other TIR satellites which may experience orbital drift during part of their lifetimes		Different correction methods are being evaluated and tested in the implementation
LST-CCI-ADP-25	Utilise Sentinel-2 data and the Optimal Estimation approach to downscale LST from moderate resolution instruments		Being implemented in the prototype high resolution products
LST-CCI-ADP-26	Implement the Optimal Estimation approach in developing the microwave product		Being implemented in V4.00 (Phase-2)
LST-CCI-ADP-27	Maximise the value of the Split-Window approach and the OE approach in a Hybrid solution		To be implemented in V5.00
LST-CCI-ADP-28	Implement the 1-hourly ERA5 profile data for cloud masking all LEO data		To be implemented in V5.00
LST-CCI-ADP-29	Update the intercalibration using ocean matchups to extend the dynamic range		To be implemented in V5.00
LST-CCI-ADP-30	Add the Orbital Drift Correction as an additional field to the product, keeping the original uncorrected LST as the default LST field		To be implemented in V5.00

2. Introduction

The European Space Agency Climate Change Initiative on Land Surface Temperature (hereafter LST_cci) aims to provide Land Surface Temperature (LST) LST Essential Climate Variable (ECV) products and validate these data to provide an accurate view of temperatures across land surfaces globally over the past 20 to 25 years.

This Algorithm Development Plan (ADP) provides a description of the expected improvements to be made to both the infrared (IR) and microwave (MW) algorithms over the course of each processing Cycle **after** Cycle 1. The algorithms themselves are described in detail in the Algorithm Theoretical Basis Document (ATBD) [RD-29]. This document instead details the upcoming developments to be undertaken to improve the output products so as to better address the climate requirements for LST as presented in the 2016 GCOS Implementation Plan [RD-19]. The algorithms identified in [RD-29] have been recommended as the best algorithms for a future climate quality operational system. The retrieval algorithms were selected during an open algorithm intercomparison round-robin which assessed the performance of a number of different LST retrieval algorithms for a set of specific thermal infrared and microwave satellite sensors [RD-29].

The evolutions to these form part of the end-to-end system being implemented to generate the first LST_cci climate data records. A flow chart summarising the algorithm processing is provided in Figure 1 for thermal infrared sensors and in Figure 2 for microwave sensors.

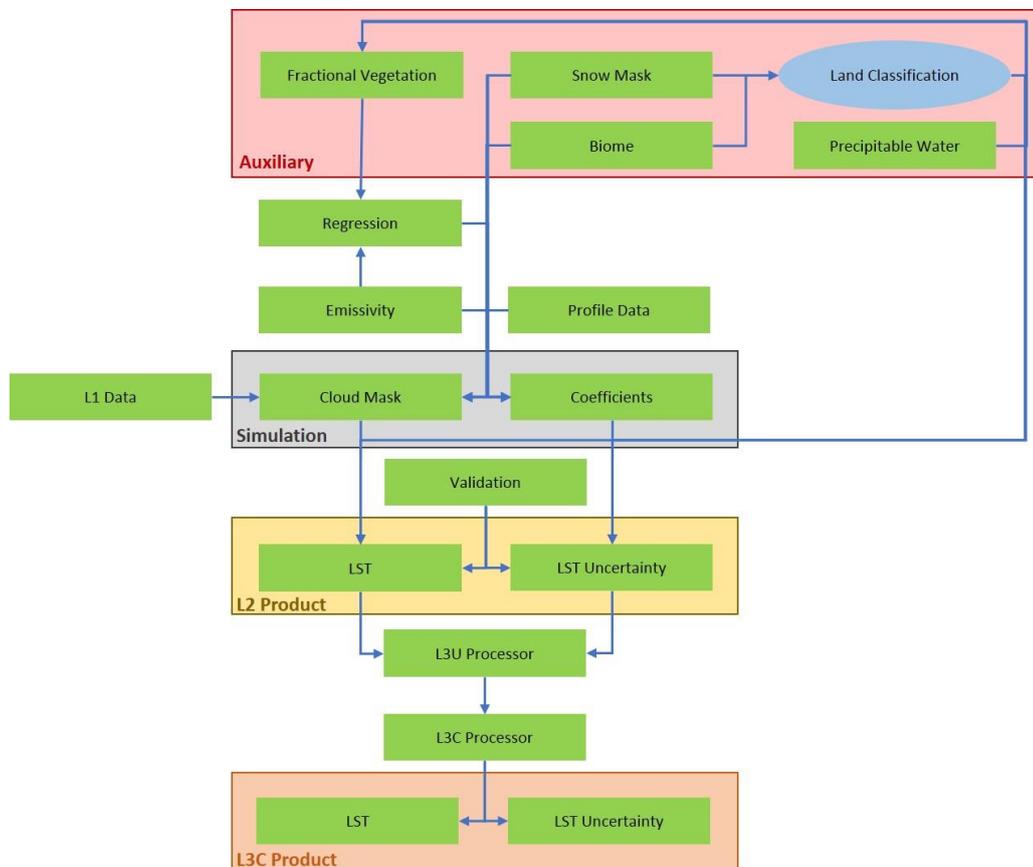


Figure 1: Data flows for LST_cci product prototype production system for thermal infrared sensors.

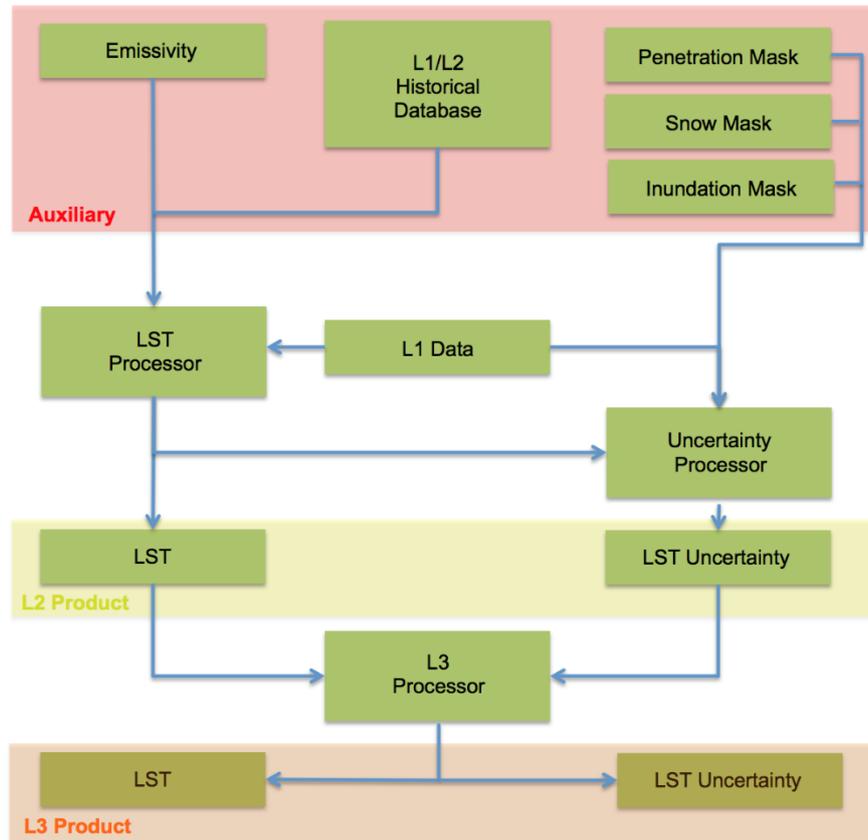


Figure 2: Data flows for the SSM/I and SSMIS LST ECV prototype production system.

At each subsequent reprocessing it is expected that ongoing algorithm assessment will be carried and evolutions made to ensure the best performing algorithm is always implemented. This will aim to produce the most accurate LST retrieval for each LST_cci product. Therefore, this document is updated as at each processing cycle with the proposed improvements for the following cycle. Table 2 indicates the Processing Cycle and Product Numbering nomenclature. **It is assumed, unless otherwise stated, that all later Product Versions contain the advancements described in this document from earlier Product Versions.**

Table 2: Processing Cycle and Product Numbering nomenclature.

Year	Processing Cycle	Product Version	Instrument	Satellite(s)
1	1	1.00	ATSR-2	ERS-2
			AATSR	Envisat
			MODIS	Terra
				Aqua
			SLSTR	Sentinel-3A
			SEVIRI	MSG-1-4
			SSM/I	DMSP F-13,17
ATSR CDR	ERS-2, Envisat			

Year	Processing Cycle	Product Version	Instrument	Satellite(s)
2	1.5	2.00	ATSR-2	ERS-2
			AATSR	Envisat
			MODIS	Terra
				Aqua
			SLSTR	Sentinel-3A
				Sentinel-3B
			SEVIRI	MSG-1-4
			Imager	GOES 12-16
JAMI	MTSAT-2			
SSM/I	DMSP F-13,17			
3	2	3.00	ATSR-2	ERS-2
			AATSR	Envisat
			AVHRR/3	NOAA-15-19
				Metop-A-C
			MODIS	Terra
				Aqua
			SLSTR	Sentinel-3A
				Sentinel-3B
			SEVIRI	MSG-1-4
			Imager	GOES 12-16
			JAMI	MTSAT-2
SSM/I	DMSP F-13,17			
ATSR-MODIS-SLSTR CDR	ERS-2, Envisat, Terra, Sentinel-3			
Merged IR CDR	LEO+GEO IR			
Ph2- Y1	3	4.00	ATSR-2	ERS-2
			AATSR	Envisat
			AVHRR/3	NOAA-15-19
				Metop-A-C
			MODIS	Terra
				Aqua
			SLSTR	Sentinel-3A
				Sentinel-3B
SEVIRI	MSG-1-4			
Imager	GOES 12-16			

Year	Processing Cycle	Product Version	Instrument	Satellite(s)
			JAMI	MTSAT-2
			SSM/I	DMSP F-13,17
			ATSR-MODIS-SLSTR CDR	ERS-2, Envisat, Terra, Sentinel-3
			Merged IR CDR	LEO+GEO IR
			AMSR-E	Aqua
			ETM+ / TIRS	Landsat-7 / -8

2.1. Purpose and Scope

This document presents the algorithm development plan for the algorithms to be used for LST data products provided by LST_cci.

2.2. Reference Documents

The following is a list of documents with a direct bearing on the content of this report. Where referenced in the text, these are identified as RD-xx, where 'xx' is the number in the table below.

Id	Reference
[RD-1]	Prata, F., Land Surface Temperature Measurement from Space: AATSR Algorithm Theoretical Basis Document. 2002.
[RD-2]	Wan, Z. and Dozier, J., (1996) A generalised spilt-window algorithm for retrieving Land-Surface Temperature from space, IEEE Trans. GeoSci. Remote Sens.
[RD-3]	Ghent D, Trigo I, Pires A, Sardou O, Bruniquel J, Gottsche F, Martin M, Prigent C, Jimenez, C, and Remedios. J., 2016. ESA DUE GlobTemperature Product User Guide V2
[RD-4]	Merchant, C., Ghent. D., Kennedy, J., Good., E., and Hoyer, J. 2016. Common approach to providing uncertainty estimates across all surfaces, H2020 EUSTACE Report
[RD-5]	Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate . Copernicus Climate Change Service Climate Data Store (CDS), date of access. https://cds.climate.copernicus.eu/cdsapp#!/home
[RD-6]	Ghent, D., Corlett, G., Gottsche, F., & Remedios, J. (2017). Global land surface temperatures from the Along-Track Scanning Radiometers. Journal of Geophysical Research: Atmospheres, 122
[RD-7]	Le Gleau, H. (2019) Algorithm Theoretical Basis Document for the Cloud Product Processors of the NWC/GEO. http://www.nwcsaf.org/Downloads/GEO/2018/Documents/Scientific_Docs/NWC-CDOP2-GEO-MFL-SCI-ATBD-Cloud_v2.1.pdf
[RD-8]	Camacho, F., Cernicharo, J., Lacaze, R., Baret, F. and Weiss, M., 2013. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over

Id	Reference
	existing products. Part 2: Validation and intercomparison with reference products. <i>Remote Sensing of Environment</i> , 137, pp.310-329.
[RD-9]	Combined ASTER and MODIS Emissivity database over Land (CAMEL) Emissivity Monthly Global 0.05Deg V001 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MEaSURES/LSTE/CAM5K30EM.001
[RD-10]	Saunders, R., RTTOV-7 - SCIENCE AND VALIDATION REPORT, 2001, EUMETSAT.
[RD-11]	Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kaliberg, P., Kobayashi, S., Uppala, S., Simmons, S. (2011). The ERA-Interim archive Version 2.0. ERA Report Series, old.ecmwf.int/publications/library/do/references/show?id=902776.
[RD-12]	Kirches, G., Brockman, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., Schlerf, M., and Santoro, M.(2016), Land Cover CCI PRODUCT USER GUIDE VERSION 2 http://data.ceda.ac.uk/neodc/esacci/land_cover/docs/ESACCI-LC-PUG-v2.5.pdf
[RD-13]	Fennig, K., Schroder, M., and Hollmann, R.: Fundamental Climate Data Record of Microwave Imager Radiances, Edition 3, 10.5676/EUM SAF CM/ FCDR MWI/V003, 2017.
[RD-14]	Prigent, C., Jimenez, C, and P. Bousquet, Satellite-derived global surface water extent and dynamics over the last 25 years, <i>J. Geophys. Res.</i> , in review.
[RD-15]	Prigent, C., I. Tegen, F. Aires, B. Marticorena, and M. Zribi (2005), Estimation of the aerodynamic roughness length in arid and semi-arid regions over the globe with the ERS scatterometer, <i>J. Geophys. Res.</i> ,110, D09205, doi:10.1029/2004JD005370.
[RD-16]	Favrichon, S., Prigent, C., Jimenez, C., and Aires, F.: Detecting cloud contamination in passive microwave satellite measurements over land, <i>Atmos. Meas. Tech.</i> , 12, 1531-1543, https://doi.org/10.5194/amt-12-1531-2019 , 2019.
[RD-17]	Sembhi, H., and Ghent, D. (2017). University of Leicester Thermal Infrared Probabilistic Cloud Detetion for Land Algorithm Theoretical Basis Document for Sentinel-3, ESA document, S3-ATBD_UOL_CLOUD_V3
[RD-18]	Seemann, S.W., Borbas, E.E., Knuteson, R.O., Stephenson, G.R., & Huang, H.L. (2008). Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements. <i>Journal of Applied Meteorology and Climatology</i> , 47, 108-123
[RD-19]	GCOS. (2016). The Global Observing System for Climate: Implementation Needs (GCOS-200)
[RD-20]	Goldberg, M., G. Ohring, J. Butler, C. Cao, R. Datla, D. Doelling, V. Gärtner, T. Hewison, B. Iacovazzi, D. Kim, T. Kurino, J. Lafeuille, P. Minnis, D. Renaut, J. Schmetz, D. Tobin, L. Wang, F. Weng, X. Wu, F. Yu, P. Zhang, and T. Zhu. (2011). The Global Space-Based Inter-Calibration System. <i>Bull. Amer. Meteor. Soc.</i> , 92, 467–475, https://doi.org/10.1175/2010BAMS2967.1

Id	Reference
[RD-21]	Göttsche, F. M., & Olesen, F. S. (2001). Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data. <i>Remote Sensing of Environment</i> , 76(3), 337-348.
[RD-22]	Hewison, T. J. (2013). "An Evaluation of the Uncertainty of the GSICS SEVIRI-IASI Intercalibration Products," in <i>IEEE Transactions on Geoscience and Remote Sensing</i> , vol. 51, no. 3, pp. 1171-1181. doi: 10.1109/TGRS.2012.2236330 URL:
[RD-23]	Chander, G., Hewison, T. J., Fox, N., Wu, X., Xiong, X., & Blackwell, W. J. (2013). Overview of intercalibration of satellite instruments. <i>IEEE Transactions on Geoscience and Remote Sensing</i> , 51(3), 1056-1080.
[RD-24]	Vinnikov, K.Y., Yu, Y., Goldberg, M.D., Tarpley, D., Romanov, P., Laszlo, I., & Chen, M. (2012). Angular anisotropy of satellite observations of land surface temperature. <i>Geophysical Research Letters</i> , 39
[RD-25]	Lagouarde, J.P., Irvine, M., (2008). Directional anisotropy in thermal infrared measurements over Toulouse city centre during the CAPITOUL measurement campaigns: First results. <i>Meteorol. Atmos. Phys.</i> 102, 173–185. doi:10.1007/s00703-008-0325-4
[RD-26]	Ghent, D., Veal, K., Trent, T., Dodd, E., Sembhi, H., and Remedios, J. (2019). A New Approach to Defining Uncertainties for MODIS Land Surface Temperature. <i>Remote Sensing</i> , 11, 1021. doi: 10.3390/rs11091021
[RD-27]	Freitas, S.C., Trigo, I.F., Bioucas-Dias, J.M., & Gottsche, F.M. (2010). Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI/Meteosat. <i>IEEE Transactions on Geoscience and Remote Sensing</i> , 48, 523-534
[RD-28]	S3MPC (2016): Sentinel-3 Mission Performance Centre. http://s3mpc-intranet.acri-cwa.fr/
[RD-29]	LST_cci (2019) Algorithm Theoretical Basis Document, Reference LST-CCI-D2.2-ATBD
[RD-30]	LST_cci (2019) Product Validation and Algorithm Selection Report, Reference LST-CCI-D2.1-PVASR
[RD-31]	LST_cci (2019) End-to-End ECV Uncertainty Budget Report, Reference LST-CCI-D2.3-E3UB
[RD-32]	LST_cci (2019) Product Specification Document, Reference LST-CCI-D1.2-PSD
[RD-33]	Duguay-Tetzlaff, A., Bento, V. A., Göttsche, F. M., Stöckli, R., Martins, J. P. A., Trigo, I., Olesen, F., Bojanowski, J. S., da Camara C., and Kunz, H. (2015) Meteosat land surface temperature climate data record: Achievable accuracy and potential uncertainties, <i>Remote Sensing</i>
[RD-34]	Freitas, S., Trigo, I., Macedo, J., Barroso, C., Silva, R., and Perdigao, R. (2013) Land surface temperature from multiple geostationary satellites, <i>International Journal of Remote Sensing</i> , 34

Id	Reference
[RD-35]	Perry, M., Remedios, J., Ghent, D., Veal, K., & Gottsche, F. An inverse method for the retrieval of LST and LSE from ASTER, in preparation
[RD-36]	Rodgers, C. D. (2000). <i>Inverse methods for atmospheric sounding - Theory and Practise</i> . (Vol. 2). World Scientific Publishing Co. Pte. Ltd.
[RD-37]	Perry, M. J. S. (2017). High Spatial Resolution Retrieval of LST and LSE for the Urban Environment (Doctoral dissertation, Department of Physics and Astronomy).
[RD-38]	Dybkjær, G., R. Tonboe, and J.L. Høyer, Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data. <i>Ocean Science</i> , 2012. 8(6): p. 959.
[RD-39]	Key, J. and M. Haefliger, Arctic ice surface temperature retrieval from AVHRR thermal channels. <i>Journal of Geophysical Research: Atmospheres</i> , 1992. 97(D5): p. 5885-5893.
[RD-40]	Good, S.; Fiedler, E.; Mao, C.; Martin, M.J.; Maycock, A.; Reid, R.; Roberts-Jones, J.; Searle, T.; Waters, J.; While, J.; Worsfold, M. The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses. <i>Remote Sens.</i> 2020, 12, 720.
[RD-41]	Julien, Y. & Sobrino, J. A. (2012). Correcting Long Term Data Record V3 estimated LST from orbital drift effects, <i>Remote Sensing of Environment</i> , 123 (2012) 207-219.
[RD-42]	Landsat 8 OLI and TIRS Calibration Notices https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8-oli-and-tirs-calibration-notice
[RD-43]	Sobrino, J. A., Jimenez-Munoz, J. C., Soria, G., Romaguera, M., Guanter, L., Moreno, J., Martinez, P. (2008). Land surface emissivity retrieval from different VNIR and TIR sensors. <i>IEEE Transactions on Geoscience and Remote Sensing</i> , 46(2), 316-327. doi:10.1109/TGRS.2007.904834
[RD-44]	Hersbach, H, Bell, B, Berrisford, P, et al. (2020), The ERA5 global reanalysis. <i>Q J R Meteorol Soc.</i> , 146: 1999– 2049, https://doi.org/10.1002/qj.3803 .
[RD-45]	Aires, F., C. Prigent, F. Bernardo, C. Jiménez, R. Saunders, and P. Brunel (2011), A Tool to Estimate Land-Surface Emissivities at Microwave frequencies (TELSEM) for use in numerical weather prediction, <i>Q. J. R. Meteorol. Soc.</i> , 137, 690–699.
[RD-46]	Prigent, C., & Rossow, W. R. (1999). Retrieval of surface and atmospheric parameters over land from SSM/I: Potential and limitations. <i>Quarterly Journal of the Royal Meteorological Society</i> , 125(559), 2379-2400.
[RD-47]	Jiménez, C., et al., Ocean and Sea Ice retrievals from a simulation of the 1.4 to 36.5 GHz measurements of the Copernicus Imaging Microwave Radiometer (CIMR), in review, <i>J. Geophys. Res.</i>
[RD-48]	(A)ATSR 4th Reprocessing Plan, 2020, v1.2, AATSRESL.PLN.003

2.3. Glossary

The following terms have been used in this report with the meanings shown.

Term	Definition
ATSR	Along-Track Scanning Radiometer
ATSR-2	Along-Track Scanning Radiometer-2
AATSR	Advanced Along-Track Scanning Radiometer
ALB2	ATSR Land Biome Classification
ATBD	Algorithm Theoretical Basis Document
BOA	Bottom of Atmosphere
BT	Brightness Temperature
C3S	Copernicus Climate Change Service
CAMEL	Combined ASTER and MODIS Emissivity for Land
CCI	Climate Change Initiative
CDR	Climate Data Record
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
Envisat	Environmental Satellite
ERA5	ECMWF Re-analysis 5
ERS	European Remote-Sensing Satellite
ESA	European Space Agency
GEO	Geostationary Orbit
GSW	Generalised Split Window
IGBP	International Geosphere–Biosphere
ISRF	Instrument Spectral Response Function
LEO	Low Earth Orbit
LSE	Land Surface Emissivity
LST	Land Surface Temperature
LST_cci	ESA CCI on LST
MODIS	Moderate Resolution Imaging Spectroradiometer
MW	Microwave
NDVI	Normalised Difference Vegetation Index
NN	Neural-Network
NWC SAF	Satellite Application Facility on Support to Nowcasting & Very Short Range Forecasting
RTM	Radiative Transfer Model



Term	Definition
RTTOV	Radiative Transfer for TOVS
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SLSTR	Sea and Land Surface Temperature Radiometer
SSM/I	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager Sounder
SST	Sea Surface Temperature
SW	Split Window
TCWV	Total Column Water Vapour
TIR	Thermal Infrared
TOA	Top of Atmosphere
UOL	University of Leicester
VCM	Vegetation Cover Method

3. Algorithm Development Plan for Infrared LST Products

3.1. Current Status of Infrared LST Products

In the LST_cci open algorithm intercomparison round-robin, the performance of different LST retrieval algorithms for a set of specific thermal infrared and microwave satellite sensors was assessed to identify the best algorithms for a future climate quality operational system. The algorithms chosen for the thermal infrared (IR) were:

- ❖ the University of Leicester (UOL) Algorithm
 - for the Advanced Along-Track Scanning Radiometer (AATSR) LST ECV dataset
 - for the AATSR / Sea and Land Surface Temperature Radiometer (SLSTR) / Moderate Resolution Imaging Spectroradiometer (MODIS) CDR
- ❖ the Generalised Split Window (GSW) Algorithm
 - for the MODIS and Spinning Enhanced Visible and InfraRed Imager (SEVIRI) LST ECV datasets
 - for the Merged Dataset (AATSR / MODIS / SEVIRI)

These algorithms are the ones which will continue to be improved over the course of the processing Cycles. The full descriptions are provided in [RD-29], but we present a brief summary here.

3.1.1. Version 2.00

3.1.1.1. UOL Algorithm

The standard algorithm ([RD-1], for (A)ATSR and SLSTR) uses a nadir-only (SW) algorithm with classes of coefficients for each combination of biome-diurnal (day/night) condition. Non-linearity is parametrised across the swath. The full form of the algorithm is presented as follows:

$$\begin{aligned}
 LST = & d(\sec(\theta) - 1)pw + (fa_{v,i} + (1 - f)a_{s,i}) + (fb_{v,i} \\
 & + (1 - f)b_{s,i})(T_{11} - T_{12})^{1 / (\cos(\theta / m))} \\
 & + ((fb_{v,i} + (1 - f)b_{s,i}) + (fc_{v,i} + (1 - f)c_{s,i}))T_{12}
 \end{aligned}$$

where the six retrieval coefficients $a_{s,i}$, $a_{v,i}$, $b_{s,i}$, $b_{v,i}$, $c_{s,i}$ and $c_{v,i}$ are dependent on the biome (i), fractional vegetation cover (f) - the retrieval coefficients $a_{s,i}$, $b_{s,i}$ and $c_{s,i}$ relate to bare soil (f = 0) conditions, and $a_{v,i}$, $b_{v,i}$ and $c_{v,i}$ relate to fully vegetated (f = 1) conditions. The fractional vegetation cover (f) and precipitable water (pw) are seasonally dependent whereas the biome (i) is invariant [RD-6].

The retrieval parameters d and m are empirically determined from validation and control the behaviour of the algorithm for each zenith viewing angle (θ) across the nadir swath. The parameter d resolves increases in atmospheric attenuation as the water vapour increases, an effect that is accentuated with increasing zenith viewing angle. The parameter m is supported by previous studies [RD-6], which suggest a non-linear dependence term on the BT difference T11 - T12 would elicit improvement in the accuracy of the LST retrievals. The rationale here is that the BT difference increases with increasing atmospheric water vapour, since attenuation due to water vapour is greater at 12 μ m than at 11 μ m.

The nature of the algorithm means that land surface emissivity is implicitly dealt with through the regression of retrieval coefficients to biome and bare soil / fully vegetated states. In other words, while

LSE is not an estimated output the algorithm still uses LSE knowledge, any uncertainty of which is propagated in the LST derivation. This knowledge is passed to the algorithm through the biome and fractional vegetation states, which themselves are regressed to emissivity states in the coefficient generation. Dynamic Fractional Vegetation Cover (FVC) ancillary data will be retrieved from auxiliary data.

For the generation of the retrieval coefficients for each biome–diurnal (day/night) combination vertical atmospheric profiles of temperature, ozone, and water vapour, surface and near-surface conditions and the surface emissivities are required. These are input, in addition to specifying the spectral response functions of the instrument, into a radiative transfer model in order to simulate TOA BTs. Retrieval coefficients are determined by minimizing the l2-norm of the model fitting error (Δ LST).

3.1.1.2. Generalised Split Window (GSW) Algorithm

The generalised split window algorithm is a view-angle dependent split-window algorithm proposed for LST retrieval by [RD-2]. It is based around channels in the 11 and 12 μ m regions.

The success of the generalized split-window LST algorithm depends on knowledge of the band emissivities for real land surfaces. In the LST_cci GSW method, emissivity information will be used explicitly rather than incorporating this information implicitly through biome coefficients.

Having determined the emissivity of the pixel coefficients these can be applied to derive an LST estimate similar to that given below:

$$T_s = C + \left(A_1 + A_2 \frac{1 - \varepsilon_{mean}}{\varepsilon_{mean}} + A_3 \frac{\Delta\varepsilon}{\varepsilon_{mean}^2} \right) \frac{T_1 + T_2}{2} + \left(B_1 + B_2 \frac{1 - \varepsilon_{mean}}{\varepsilon_{mean}} + B_3 \frac{\Delta\varepsilon}{\varepsilon_{mean}^2} \right) \frac{T_1 - T_2}{2}$$

Where C, A and B are coefficients derived from linear regression using simulated data as done for the UOL algorithm but adapted for the GSW. T1 and T2 are the 11 and 12 μ m brightness temperatures. ε_{mean} is the mean emissivity of the two thermal channels used in the GSW algorithm:

$$\varepsilon_{mean} = 0.5 (\varepsilon_{11} + \varepsilon_{12})$$

$\Delta\varepsilon$ is the difference between the two thermal channels, calculated as:

$$\Delta\varepsilon = \varepsilon_{11} - \varepsilon_{12}$$

The coefficients for GSW are dependent on satellite viewing angle and water vapour. The bands for water vapour are of width 15 $\text{kg}\cdot\text{m}^{-2}$ so that the first water vapour band is from [0,15) $\text{kg}\cdot\text{m}^{-2}$. The bands for satellite zenith angle are of width 5°. The retrieval coefficients are linearly interpolated between viewing angle and water vapour bands to minimise step changes.

3.1.2. Version 3.00

3.1.2.1. Single-channel (SMW) Algorithm

While the Round Robin intercomparison correctly identified the UOL and GSW algorithms as the best algorithms for the respective products, this did not take account of the lowest common denominator in the multi-sensor CDRs, which is the number of channels. For some single-sensor products (early GOES and

(MTSAT) only one thermal infrared channel can be used. This impacts the Merged Product, and so for these products a single-channel algorithm will be applied.

The specific algorithm chosen is that used by the CM SAF for MVIRI retrievals and has been successfully applied to SEVIRI with good results [RD-33] and is close to the formulation for GOES [RD-34]:

$$LST = A + \frac{T_x}{\varepsilon_x} + B \frac{1}{\varepsilon_x} + C$$

Where A, B and C are coefficients derived from linear regression using simulated data as done for the split window algorithms but adapted for the SMW. T_x is the single channel brightness temperature and ε_x is the spectral emissivity for this channel.

The coefficients for GSW are dependent on satellite viewing angle and water vapour. The bands for water vapour are of width $15 \text{ kg}\cdot\text{m}^{-2}$ so that the first water vapour band is from $[0,15) \text{ kg}\cdot\text{m}^{-2}$. The bands for satellite zenith angle are of width 5° . The retrieval coefficients are linearly interpolated between viewing angle and water vapour bands to minimise step changes.

3.1.3. Version 4.00

3.1.3.1. Optimal Estimation (OE) Algorithm

For higher resolution instruments we need to exploit the capabilities of advanced algorithms. We will start with the Optimal Estimation (OE) approach which has successfully been applied to high resolution LST data [RD-35]. This approach is sometimes referred to as Inverse Theory, and is a technique developed to help solve problems which are either *over-* or *under- constrained* and there is some degree of uncertainty in the measurements or formulation. It has proved to be of great use in Earth Observation, specifically in resolving multispectral measurements to give multilevel atmospheric profiles.

As an example consider an EO satellite which observes TOA BTs at i different wavelengths. The measurements are collated into the single observation vector (\mathbf{y}). The objective is to retrieve a profile (\mathbf{x}), containing n different atmospheric and/or surface parameters. If the actual state of the atmosphere and surface was known at the time of observation then:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \boldsymbol{\epsilon},$$

where \mathbf{F} represents the *Forward Model* which computes the radiative transfer process based on the input profile (\mathbf{x}), and $\boldsymbol{\epsilon}$ is the combined forward model and measurement error. The maximum probability solution can be found by minimising the cost function, C , as demonstrated by [RD-36], that is:

$$C(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) + (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}))$$

where \mathbf{x}_a is an *a priori* estimate of the true state with a covariance described by the matrix \mathbf{S}_a and the observation vector has a covariance matrix described by the matrix \mathbf{S}_y . [RD-36] essentially describes how the maximum probability solution is calculated based on a compromise between the confidence associated with the predicted state and the confidence associated with the observation. The retrieval process is illustrated in Figure 3.

The observations given to the retrieval can apply the instrument specific SRF for the desired number and configuration of channels. The main iterative loop provides a configurable platform to simulate the TOA

brightness temperature and retrieve the surface parameters. The flexibility in this module extends to multiple algorithms, related to the degree on non-linearity in the system and the speed of the convergence.

In comparison to SW techniques, OE methods ensure greater consideration of the true physics. Generally, this ensures that the retrieved state is more likely to reproduce the observation vector when used as input in a forward model. In contrast, for deployment in an ECV production chain sensitivity to atmospheric profiles of temperature and water vapour will be assessed through the round robin exercise. A further advantage of the OE methodology is in the retrieved uncertainties and radiance residuals inherent to the retrieval process. The retrieval automatically produces an uncertainty which is back traceable, allowing the determination of the contributions of the underlying variables [RD-37].

LST-CCI-ADP-10: Exploit the capabilities of advanced algorithms, such as Optimal Estimation, to confront the challenges of higher resolution sensors and scale change.

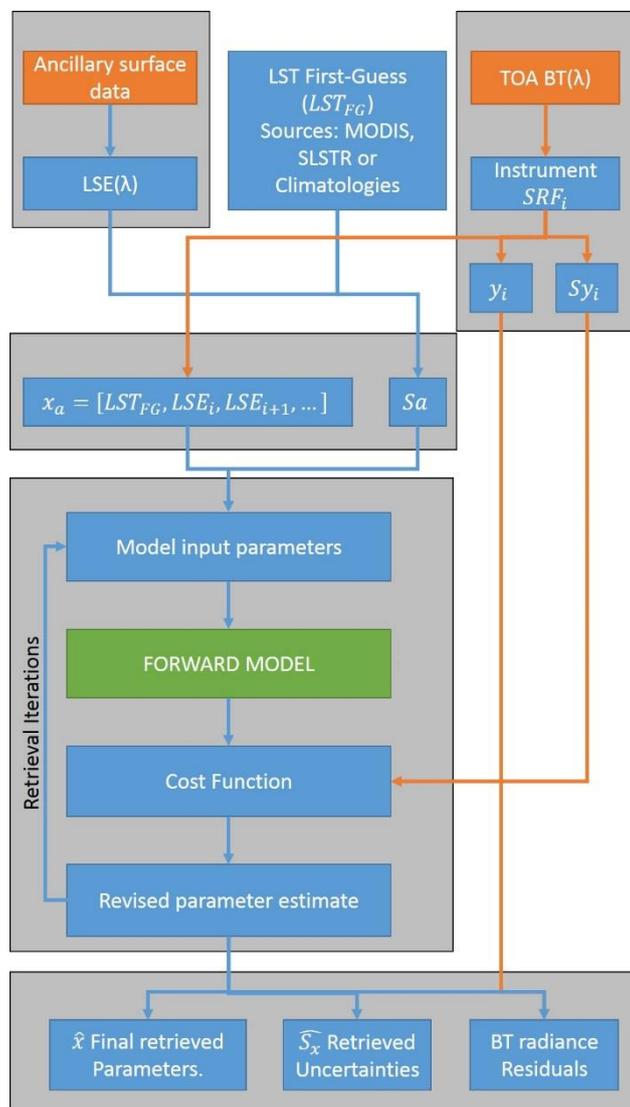


Figure 3: Retrieval flow chart for the OE method applied to TOA BTs and external LSE data.

3.1.4. Version 5.00

3.1.4.1. Hybrid Algorithm

The Hybrid Algorithm follows a similar logic to the downscaling approach in Section 3.2.3.3.2. The rationale is that we utilise the power of the Split-Window technique (such as detailed in Section 3.1.1.2) to resolve the atmospheric correction and then apply a Bottom-Of-Atmosphere (BOA) OE scheme to optimise the LST and LSE outputs. So for instance we take the MODIS LST product at 1 km resolution, if this is product we are producing, and use NDVI derived *a priori* LSE data to iteratively update the LST in an OE scheme. In this methodology the OE scheme will assimilate the medium resolution MODIS LST and the Copernicus-derived LSE and attempt to minimise the difference between the MODIS calculated BOA brightness temperatures (BTs) and simulated BOA BTs generated from the MODIS LST and the Copernicus-derived LSE.

The processing first obtains the normalised difference vegetation index (NDVI) from the Copernicus Global Land Service (CGLS) and maps this onto the MODIS LST swath. The NDVI data is used with the NDVI threshold method of [RD-43] to estimate the LSE. These *a priori* LST and LSE are used with Planck's Law to estimate the BOA BT values. In order to transfer the traceable uncertainties from the Split-Window algorithm, the random component of the total uncertainty from the MODIS LST algorithm is applied as a noise on these BOA BT values using a Gaussian random distribution. The OE scheme iteratively updates the estimates for LST and LSE used in the retrieval until an empirically determined threshold is reached from where there is no longer any significant improvement to be made by further iteration.

This methodology produces a full uncertainty breakdown including both the input total uncertainty of the MODIS LST retrieval as well as the uncertainties due to the BOA OE iterations.

LST-CCI-ADP-27: Maximise the value of the Split-Window approach and the OE approach in a Hybrid solution.

3.2. Algorithm Development Plan for Infrared Sensors

3.2.1. Version 2.00

After Cycle 1, all single-sensor GEO products will be reprocessed with a GSW calibrated with the benchmark dataset constructed for the Round Robin. A natural consequence of this is that all single-sensor GEO ECV Products will also be consistent with LEO ECV Products with respect to coefficient generation.

3.2.1.1. Radiative Transfer Modelling

Radiative Transfer for TOVS (RTTOV) is a fast Radiative Transfer Model (RTM) from the NWP-SAF [RD-10]. It is an efficient radiative transfer forward model for the visible, infra-red and microwave wavelengths. In contrast to models using a line-by-line methodology, RTTOV conceptualizes the simulation in terms of channel radiances.

For Cycle 1, all the Low Earth Orbit (LEO) ECV Products have been developed using RTTOV Version 11.3 to determine the retrieval coefficients. While this is consistent with previous studies, such as in the ESA DUE GlobTemperature [RD-3] and Sentinel-3 Mission Performance Centre (S3MPC) [RD-28] it does not utilise

the most recent knowledge of global surface emissivity as provided by the Combined ASTER and MODIS Emissivity database over Land (CAMEL) [RD-9] dataset. The latest version of RTTOV, 12.3, integrates this recent dataset in its radiative transfer. Thus, all LST_cci LEO TIR ECV Products and IR Climate Data Records (CDRs) will benefit from this upgrade starting from Cycle 1.5. The expected impact here is for improved retrieval coefficients.

LST-CCI-ADP-1: Upgrade to RTTOV 12.3 for development of all LST_cci LEO TIR ECV Products and IR Climate Data Records (CDRs).

In addition to the retrieval algorithms, RTTOV is also used in the UOL_3 and Bayesian cloud masking algorithms to calculate the probability of cloud cover in the observations given the background state. Retrieval coefficients are derived using forward modelling and building regressions between the skin temperature and the TOA radiances to create a Calibration Database for determining retrieval coefficients. RTTOV is also used in the threshold tests employed in the NWC-SAF cloud masking algorithm.

3.2.1.2. Calibration Database for Determining Retrieval Coefficients for the TIR Algorithms

Globally robust, traceable retrieval coefficients for both the GSW and UOL approaches are generated using RTTOV, which allows fast processing of sufficient numbers of profiles to adequately characterise a wide range of potential atmospheric states representative of each biome (UOL algorithm) or each water vapour – viewing angle combination (GSW algorithm).

In Cycle 1 the profile data has been sourced from ERA-Interim, which provides a large number of input profiles which encompass the full range of atmospheres and surfaces observed by TIR instruments. A uniform random sampling strategy is used to select a number of clear sky profiles for each biome class. A large sample of locations are selected randomly across land and ice surface types over all latitude and longitude bands to represent the full range of surface types across land areas [RD-6]. A temporal sampling strategy ensures intra- and inter-annual coverage; profiles are selected from the 15th day of each month between 2002 and 2011 with identified profiles closest to the day and night overpass times of the satellite of interest. Representative emissivity information is extracted for the locations of the profile data from auxiliary datasets. These selected profiles are then inputs to the RTTOV forward model along with the various sensor spectral response functions. RTTOV then yields the brightness temperatures and LSTs for the given sensor and location and these are used to generate retrieval coefficients for all cases of biome type, fractional vegetation and water vapour using linear regression. The Calibration Database comprises a global set of independent profiles and emissivity values covering all land cover types and distributed across all latitude and longitudes and capturing the seasonality of the land surface, as well as the coefficients generated from these profiles.

While the same general approach to coefficient generation is to be utilised for all subsequent processing, the source of the input profiles and surface states will be upgraded from those used in Cycle 1. Specifically, an extended version of the Benchmark database constructed for the Round Robin [RD-30] will be used to determine retrieval coefficients for TIR algorithms. This extended benchmark dataset will use ERA5 Atmospheric profile Data, CAMEL Emissivity Data and ESA CCI Land Cover data. ERA5 profiles and CAMEL emissivity data are used as an input to RTTOV. Representative profiles distributed across the globe are extracted, including simulated brightness temperatures, LST, elevation and other atmospheric information.

LST-CCI-ADP-2: An extended version of the Benchmark database using ERA5 Atmospheric profile Data, CAMEL Emissivity Data and ESA CCI Land Cover data will be used to determine retrieval coefficients for TIR algorithms.

The use of ERA5 brings consistency with other ECVs across CCI. Moreover, the temporal and spatial resolution is also improved upon compared with ERA-Interim so the expectation would be for more representative atmospheric and surface states in the Calibration Database to better capture heterogeneity and thus retrieval coefficients that generate physically more robust LST estimates.

The CAMEL dataset of surface emissivity represents the most up-to-date knowledge of global surface emissivity on a sufficiently high temporal resolution – in this case monthly. This dataset is embedded in the RTTOV 12.3 functionality and at a spatial resolution more compatible with the ERA5 atmospheric and near-surface data.

3.2.1.3. Algorithm Stratification

3.2.1.3.1 UOL Algorithm

Validation and simulations from [RD-6] suggests improved accuracy and precision for the UOL algorithm could be achieved by stratifying the retrieval coefficients by temperature. The rationale here is that the onboard calibration for the ATSRs and SLSTR is driven by the need for highly accurate sea surface temperature (SST). Calibration is thus performed to cover the dynamic range of SST, and measurements outside this do not have the same traceable source of calibration while also being more non-linear. Temperature-dependent coefficients could address this and will be explored in detail for the next processing cycle. This will be performed globally for all years and sensors against the multi-sensor matchup database (MMDB). This is possible since the MMDB is indexed in such a way that the subset of data applicable for the MMDB can easily be identified and reprocessed quickly. The likely impact would be improved LST estimates with lower uncertainties.

LST-CCI-ADP-4: Temperature-dependent coefficients will be explored in detail for the UOL algorithm for the next processing cycle to minimise any possible impact of non-linearity.

3.2.1.3.2 GSW Algorithm

The parameters in the GSW algorithm for SEVIRI are estimated for different classes of total column water vapour in bands of 7.5 kgm^{-2} and for classes of viewing zenith angle of 5° , ensuring that all ranges of atmospheric attenuation within the thermal infrared are covered. Literature [RD-26; RD-27] indicates this level of stratification is suitable for global retrievals of LST using the GSW approach. Instead, algorithm development will focus on improved derivation of the retrieval coefficients using the new calibration database, and improved input auxiliary datasets. These ensure the retrieval is robust to different input conditions from the Level-1 brightness temperatures, auxiliary data, and instrument viewing geometry.

LST-CCI-ADP-5: Development for the GSW algorithm will focus on improved derivation of the retrieval coefficients using the new calibration database, and improved input auxiliary datasets.

3.2.1.4. Identification of Observations Valid for Land Surface Temperature Estimation from Thermal Infrared Sensors

Cloud screening is a fundamental step for Thermal Infrared (TIR) LST retrieval. For LST_cci products the cloud mask is given, or applied to, Level 2 and Level 3 LST products.

Traditionally, threshold based techniques have been used to detect cloud but these often fail under difficult circumstances – for example, in the detection of thin cirrus or low-level fog. Two cloud detection algorithms are considered here:

- ❖ The UOL_3 algorithm
- ❖ Bayesian algorithm
- ❖ the NWCSAF Cloud Mask Algorithm

The full details of the approaches are given in the ATBD [RD-29], but we provide a very brief summary here for the UOL_3 and NWCSAF algorithms, and in Section 3.2.2.1.2 for the Bayesian algorithm. In Cycle 1 only the UOL_3 (LEO products) algorithm and NWCSAF (GEO products) algorithm were applied. For some products, such as MODIS and SLSTR, the respective operational cloud masks were applied.

For subsequent processing cycles the best cloud masking per single-sensor product will be used. For the GEO single-sensor products this will remain the NWCSAF approach since it has been developed over many years for the operational SEVIRI products and utilises information on the diurnal cycle for the cloud detection.

For the LEO single-sensor products the UOL_3 algorithm is expected to be applied. This will be an improved implementation from that applied in Cycle 1 for the ATSRs, which will be customised for additional sensors (MODIS, SLSTR and AVHRR). The expected impact here is for a reduction in some of the known issues with the operational cloud masks which do not take account of the coincident conditions.

3.2.1.4.1 UOL_3 Algorithm

The UOL_3 algorithm is a semi-Bayesian cloud masking approach using the probability of clear-sky conditions which has been developed at University of Leicester [RD-6]. A pixel-level cloud mask is derived using a combination of simulated brightness temperatures and observational climatology. The approach is equally valid for both day and night-time retrievals as this method is independent of visible wavelength information. It has been implemented in the ESA DUE GlobTemperature project previously for ATSR data records [RD-6] and is being implemented operationally for SLSTR [RD-17].

This cloud masking algorithm uses atmospheric profile data to predict clear-sky conditions for the coincident space and time of a given satellite sensor observation. Coincident clear-sky brightness temperatures are derived by bilinear interpolation between surrounding ECMWF profile locations and a temporal interpolation between the 6-hourly analysis fields. ERA-Interim data has been used for profiles in Cycle 1 and Cycle 1.5.

3.2.1.4.2 NWCSAF Cloud Mask Algorithm

This cloud mask has been designed to be applicable to imagers on board meteorological geostationary satellites [RD-7]. It is based on a series of satellite dependent threshold tests [RD-7]. Steps in the process include: i) identifying pixels using a series of multispectral threshold tests based on factors such as viewing geometry, surface temperature and atmospheric water content, elevation, and climatological data; ii) a smaller series of multispectral tests; iii) analysis of the temporal variation in a short time period for detection of rapidly moving clouds; iv) a specific treatment targeting low clouds in twilight conditions; v) an analysis of solar channels at high spatial resolution to detect sub-pixel clouds; and vi) a spatial filtering for cold / snow areas, cloud edges, and isolated cloud pixels.

3.2.1.5. Auxiliary Datasets for Thermal Infrared Retrievals

The following section gives a description of the planned auxiliary datasets to be used for the thermal infrared retrieval algorithms.

3.2.1.5.1 Biome

In Cycle 1 biome information for the LST_cci products has been provided by the ATSR Land Biome Classification (ALB2) [RD-6]. For the next and all subsequent cycles the biome system used will be that provided by the ESA CCI Land Cover maps developed by the Land Cover CCI [RD-12] and hybridised for LST_cci (Table 3). This hybrid land cover scheme (LCCS) will be used in the UOL retrieval algorithm, in combination with other variables, to determine the most appropriate coefficients to apply. It will also be used in the UOL_3 and Bayesian cloud masking algorithms.

LST-CCI-ADP-7: All retrieval algorithm and cloud masking developments will move to using land cover information from Land Cover CCI.

While an aim here is consistency across the CCI projects, it is also critical to use the most appropriate system for thermal infrared data. In that respect we propose to modify the system to sub-divide class 33 (bare soil) into distinct sub-classes based on soil taxonomy. The rationale here is that emissivity variability is highest for different types of bare soil and for biome-based algorithms distinguishing between different bare soil types is crucial for improving robustness of the retrieval algorithms. It is consistent with the CCI programme in general while allowing for an extra dimension to support the specific needs of LST_cci. Since the existing ALB2 is a static mask there would be expected to be an impact on the LST estimates through use of more dynamic information.

LST-CCI-ADP-8: Class 33 (bare soil) of the Land Cover CCI classification will be sub-divided into distinct sub-classes based on soil taxonomy.

Table 3: Land Cover CCI biome definition.

Biome number	Definition
0	No data
10	cropland rainfed
11	cropland rainfed herbaceous cover
12	cropland rainfed tree or shrub cover
20	cropland irrigated
30	mosaic cropland
40	mosaic natural vegetation
50	tree broadleaved evergreen closed to open
60	tree broadleaved deciduous closed to open
61	tree broadleaved deciduous closed
62	tree broadleaved deciduous open
70	tree needleleaved evergreen closed to open
71	tree needleleaved evergreen closed
72	tree needleleaved evergreen open
80	tree needleleaved deciduous closed to open
81	tree needleleaved deciduous closed

82	tree needleleaved deciduous open
90	tree mixed
100	mosaic tree and shrub
110	mosaic herbaceous
120	shrubland
121	shrubland evergreen
122	shrubland deciduous
130	grassland
140	lichens and mosses
150	sparse vegetation
151	sparse tree
152	sparse shrub
153	sparse herbaceous
160	tree cover flooded fresh or brakish water
170	tree cover flooded saline water
180	shrub or herbaceous cover flooded
190	Urban
200	bare areas of soil types not contained in biomes 21 to 25
201	unconsolidated bare areas of soil types not contained in biomes 21 to 25
202	consolidated bare areas of soil types not contained in biomes 21 to 25
203	bare areas of soil type Entisols Orthents
204	bare areas of soil type Shifting sand
205	bare areas of soil type Aridisols Calcids
206	bare areas of soil type Aridisols Cambids
207	bare areas of soil type Gelisols Orthels
210	Water
220	Snow and ice

3.2.1.5.2 Fractional Vegetation

Fractional vegetation cover information for LST_cci is provided by the Copernicus Global Land Cover Services FCOVER dataset V2.0 (<https://land.copernicus.eu/global/products/fcover>). This global dataset is available at 1/112° resolution every 10 days from 1999 onwards. It is acquired using a moving temporal window of around 30 days [RD-21, RD-6]. Fractional vegetation is used in the UOL retrieval algorithm, in combination with other variables, to weight the appropriate retrieval coefficients applied in the algorithm. Since there is no FCOVER output from the CCI programme using such a dataset, which has been shown to be well validated [RD-8], is not inconsistent with the wider CCI programme.

LST-CCI-ADP-9: *We will continue to use FCOVER output from the Copernicus Global Land Cover Services. Since there is no FCOVER output from the CCI programme using such a dataset is not inconsistent with the wider CCI programme.*

3.2.1.5.3 Emissivity

For initial LST_cci products in Cycle 1, the CIMSS Baseline Fit Emissivity Database [RD-18] has been used. CIMSS is a monthly dataset 0.05° with emissivities available at ten wavelengths between 3.6µm and 14.3µm. It has been derived using the MODIS operational land surface emissivity product and by applying a baseline fit method to fill in the spectral gaps between the six infrared emissivity wavelengths provided.

In the next and Cycle the Combined ASTER and MODIS Emissivity for Land (CAMEL) database will be used to calculate LST. Emissivity from CAMEL is also employed in deriving retrieval coefficients for the LST retrieval algorithm by way of the Calibration Database. The is a global monthly mean emissivity dataset spanning the years 2000 – 2016. A climatology of CAMEL data will be used after 2016 if regular updates of this dataset are not available. It assimilates both ASTER Global Emissivity Database retrieved values and University of Wisconsin-Madison MODIS Infra-red Emissivity dataset values. The CAMEL dataset contains 12 emissivity values at different wavelengths from 3.6 to 14.3 μm at a resolution of 0.5° [RD-9]. Due to the dataset originating from satellite observations, it is highly relevant to realistic materials observed from space and should remove materials in spectral libraries, which are too fine a scale to be useful. The expected impact of this change in emissivity database to be used would be improved LST estimates for the GSW algorithm.

LST-CCI-ADP-10: *In the next and Cycle the Combined ASTER and MODIS Emissivity for Land (CAMEL) database will be used to calculate LST.*

3.2.1.5.4 Atmospheric Variables

Precipitable water is used in the UOL retrieval algorithm, along with coefficients selected using biome and fractional vegetation information, to derive LST. Water vapour information and atmospheric temperature are inputs required to determine retrieval coefficients for both the algorithm selected from the Round Ronin (UOL and GSW algorithms) [RD-30]. Atmospheric profile data from reanalysis is also used in the UOL_3 cloud masking algorithm to derive clear sky probability information. Furthermore, atmospheric profiles are employed in deriving retrieval coefficients for the retrieval algorithm to create a Calibration Database for determining retrieval coefficients.

In Cycle 1 ERA-Interim has been used instead. In all sequent Cycles this will be superseded by ERA5 [RD-5]. ERA5 is a re-analysis dataset which provides hourly estimates of a significant number of land and atmospheric variables over the full globe (See **LST-CCI-ADP-11**). It is the successor to ERA-Interim. ERA5 currently has a temporal coverage similar to other reanalyses (from 1979 to present), but more years are due to be added to extend this dataset back to 1950. The improved temporal and spatial resolution of ERA5 are expected to result in improvements to the retrieved LST.

3.2.1.6. Climate Data Records

3.2.1.6.1 ATSR-SLSTR CDR

The finest temporal resolution of the final product will be daily (day / night). The diurnal dimension determined from the solar zenith angle. The spatial resolution is to be 0.05°. This spatial-temporal combination is intended to meet the corresponding GCOS requirements [RD-19].

To fill the gap between the end of the Envisat mission and the end of the Phase E1 commissioning of the Sentinel-3A mission requires an instrument of equivalent spatial resolution with a LECT close in time to AATSR and SLSTR and of sufficient high quality. The choice made here is Terra-MODIS. Not only does it meet these minimum requirements, but it also spans all three of ATSR-2, AATSR and SLSTR, and moreover has a common LECT (10:30 and 22:30) with ATSR-2 so knowledge of the temporal correction between ATSR-2 and AATSR is also applicable for Terra-MODIS and AATSR. Note, AATSR and SLSTR have the same LECTs (10:00 and 22:00).

There are several steps to generating a CDR from ATSR-2 through to SLSTR:

- ❖ Stage 1: gap bridging

- A first step for bridging the gap between the end of Envisat and the start of routine data availability from Sentinel-3 is to intercalibrate the radiances between AATSR and SLSTR. This necessarily also includes intercalibration of the other sensors in the CDR: ATSR-2 and Terra-MODIS.
- Two approaches are to be used for intercalibrating the radiances between AATSR, Terra-MODIS and SLSTR:
 - ◆ Utilising matched observations from in situ Fiducial Reference Measurements (FRMs):
 - ◆ Generating simultaneous nadir overpass (SNO) matchups between each of AATSR, Terra-MODIS and SLSTR, and a reference sensor:
- A motivation for utilising both approaches is firstly that it permits more than one independent assessment of the radiances, and secondly that a wider scope of the dynamic range of the instruments can be assessed. Due to the strict 5-minute temporal threshold for SNO matchups as recommended by GSICS [RD-20], matchups between IASI and AATSR / Terra-MODIS / SLSTR will be principally in the mid-to-high latitudes and therefore confront the intercalibrations at the low to mid temperatures of the dynamic range. The use of FRMs from Gobabeb permit intercalibrations at the mid to high temperatures.
- ❖ Stage 2: time difference correction
 - A key consideration in the derivation of a continuous CDR from ATSR-2 through to SLSTR is to harmonise the temporal differences between the instruments. This difference is more complex than for SST with the impact on surface temperature due to the temporal offset in observation times an order of magnitude greater than radiometric calibration and spectral filter response variations. The 30-minute difference in LECT between Envisat and Sentinel-3, and ERS-2 and Terra has significant implications for stability of the long-term LST CDR at the levels required by GCOS.
 - To correct ATSR-2 and Terra-MODIS to the common reference LECT of AATSR and SLSTR it is crucial to characterise the diurnal information for the different land covers encountered for a global mission. In extreme cases, such as in arid environments, a 30-minute difference between observations during local mid-morning can result in several kelvin in surface temperature [RD-21].
 - In Cycle 1 we have been deriving a dataset of corrections to apply to one or more of the input datasets to alleviate the problem of combining time series that have a 30 minute difference in LECT:
 - ◆ This uses an overlap analysis between ATSR-2, AATSR and Terra-MODIS, and has been applied for a ATSR-2 to AATSR CDR.
 - ◆ SLSTR will be added to this overlap analysis in the next Cycle.
 - ◆ Diurnal information from the high temporal resolution geostationary satellites for each land cover class will also be incorporated.

LST-CCI-ADP-12: In the processing of the ATSR-SLSTR CDR the overlap analysis for the next Cycle will include SLSTR.

LST-CCI-ADP-13: Diurnal information from the high temporal resolution geostationary satellites for each land cover class will be incorporated into the time difference assessment for the ATSR-SLSTR CDR.

- ❖ Stage 3: apply consistent Level-2 algorithms
 - The UOL algorithm was shown to be the best performing algorithm for the ATSR and MODIS combination from the Round Robin exercise as documented in [RD-30]. In the next Cycle and all subsequent cycles this algorithm will be applied to the Level-1 data to ensure consistency from an algorithm perspective across the ATSR-SLSTR CDR.
 - The best performing cloud mask from WP2.4 developments and assessment will be applied to the Level-1 data for the individual sensors to ensure consistency.

LST-CCI-ADP-14: *The ATSR-SLSTR CDR will implement the best performing cloud mask from WP2.4 developments.*

- ❖ Stage 4: restrict satellite viewing angle
 - To maintain consistency of the CDR from ATSR-2 through to SLSTR in terms of data coverage and to minimise anisotropic differences in the data record as we switch from one instrument to another the satellite zenith angle (SZA) will be restricted to the lowest common denominator. In the case, both ATSR-2 and AATSR have a maximum SZA of approximately 22°. Thus, only MODIS and SLSTR pixels with SZAs less or equal to this threshold will be processed through the CDR chain. As part of this step an assessment of the impact of restricting / not restricting the angles will be carried out.
- ❖ Stage 5: gap filling
 - The gap filling process consists of the production of the CDR using the appropriate data stream for a given day. The modularisation of the LEO processing chains, which has been mostly implemented in Cycle 1, allows for the ingestion of multiple satellite Level-1 data through module plug-ins.
 - Since a key temporal interval is expected to be monthly data, the switch between sensors will occur at the monthly interchange. So the following sensors are to be the input over the full CDR time window:
 - ◆ ATSR-2: 08/1995 to 07/2002
 - ◆ AATSR: 08/2003 to 03/2012
 - ◆ Terra-MODIS: 04/2002 to 07/2016
 - ◆ SLSTR: 08/2016 to 12/2020
 - Note, we intend to implement the full period from ATSR-2 even though this means incompletely observed months. This includes a missing period of 6 months from January 1996 to June 1996 inclusive due to a scan mirror failure, and a further period from January 2001 to June 2001 inclusive due to a gyro failure. In the L1b data (Main Product Header) an indicator is set as to the application of a yaw-correction. Where this is not set data is segregated and this flagging will be taken into account in the processing. Further periods lasting up to a few days occur throughout the data record due to instrument decontaminations or anomalies. Months where these occur are excluded from the data record.

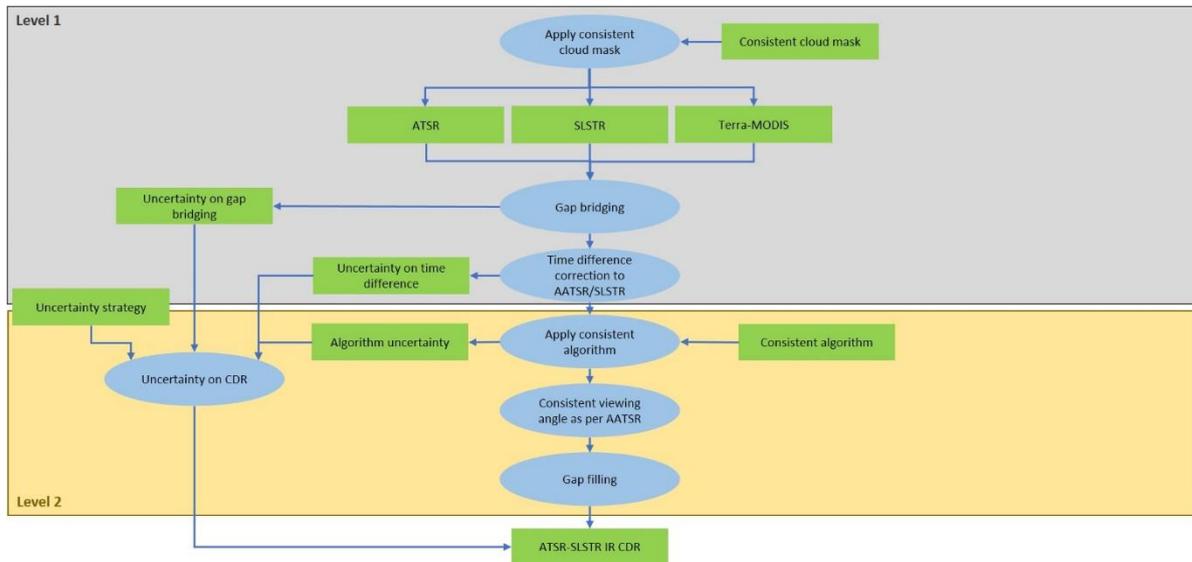


Figure 4: Schematic representation of the development and implementation of the ATSR-SLSTR CDR.

3.2.1.7. Uncertainty Model for Thermal Infrared Algorithms

Following the agreed approach being undertaken in other projects such as ESA DUE GlobTemperature [RD-3] and H2020 EUSTACE [RD-4], whereby SST, LST and IST all conform to a standardised uncertainty model. For LST this has started to be implemented for AATSR, MODIS and SEVIRI data, but further improvements are to be implemented in line with the approach being documented in the End-to-End ECV Uncertainty Budget Report (E3UB) [RD-31].

Generally, for each pixel, three components of uncertainty are provided, representing the uncertainty from effects whose errors have distinct correlation properties:

- ❖ random (no correlation of error component between cells);
- ❖ locally systematic (correlation of error component between “nearby” pixels);
- ❖ [large-scale] systematic (correlation of error component between “distant” pixels).

Locally correlated errors are modelled via spatio-temporal correlation length scales that determine how an observation influences the analysis in the vicinity of its time-space location. Systematic errors will be accounted for by allowing a bias to be determined within the analysis procedure between different sources of data, whose magnitude is conditioned by the uncertainty attributed to systematic effects.

3.2.2. Version 3.00

For Cycle 2, all input LEO and GEO data will be reprocessed with a SMW algorithm calibrated with the benchmark dataset constructed for the Round Robin, in order to provide harmonized data for the Merged Product. This is important since the impact of not bringing consistency would be potential discontinuities between the different data streams.

LST-CCI-ADP-18: All LEO and GEO products will be reprocessed with a single-channel (SMW) algorithm calibrated with the benchmark dataset constructed for the Round Robin, in order to provide harmonized data for the Merged Product.

For the Merged IR Product consistency across each component of the data development and production is critical. In this respect our objective is to use the same calibration dataset to derive retrieval coefficients for each input sensor to the output product. Thus, the extended version of the Benchmark database will be used to determine the retrieval coefficients for the SMW algorithm for both Low Earth Orbit (LEO) and Geostationary Earth Orbit (GEO) IR sensors.

3.2.2.1. Identification of Observations Valid for Land Surface Temperature Estimation from Thermal Infrared Sensors

For V3.00 products additional improvements will be made. These include enhancements to the UOL_3 algorithm, and the use of the best algorithm for the CDRs. For these two CDR developments (ATSR-MODIS-SLSTR CDR and the Merged IR CDR) either the Bayesian or the UOL_3 algorithm will be applied depending on which is most appropriate. This activity is being carried out in WP2.4.

3.2.2.1.1 UOL_3 Algorithm

We will move to ERA5 data, which will minimise any non-linearity between adjacent time steps since the profiles are only 1-hour apart. The higher spatial resolution of ERA5 is also expected to improve the outputs of the bilinear interpolation onto each sensors tie-point grid.

LST-CCI-ADP-6: All development will move to using ERA5 data, which will minimise any non-linearity between adjacent time steps since the profiles are only 1-hour apart.

3.2.2.1.2 Bayesian Algorithm

The Bayesian cloud mask, which was developed at the University of Reading, calculates the probability of clear-sky $P(c|\mathbf{y}^o, \mathbf{x}_b)$ given the observation vector (\mathbf{y}^o) and prior knowledge of the background state (\mathbf{x}_b):

$$P(c|\mathbf{y}^o, \mathbf{x}_b) = \left[1 + \frac{P(\bar{c})P(\mathbf{y}^o|\mathbf{x}_b, \bar{c})}{P(c)P(\mathbf{y}^o|\mathbf{x}_b, c)} \right]^{-1}$$

Where \bar{c} and c denote cloud and clear conditions respectively. The prior probabilities of clear and cloudy conditions ($P(c)$ and $P(\bar{c})$) are currently defined using ECMWF ERA-Interim total cloud cover [RD-11]. For subsequent cycles ERA5 will be used as per all cloud masking developments.

3.2.2.2. Auxiliary Datasets for Thermal Infrared Retrievals

The following section gives a description of the planned auxiliary datasets to be used for the thermal infrared retrieval algorithms.

3.2.2.2.1 Emissivity

It is expected by the final processing Cycle that a new and improved emissivity dataset for MODIS shall be available through the LST_cci Work Package 2.9: Temperature and Emissivity Separation from MODIS multispectral TIR data (CCN to Baseline Project). Assuming a high quality, as per the validation effort in the CCN, this would replace CAMEL as an input for other sensors using equivalent TIR channels.

3.2.2.2.2 Snow masking

Snow masking is part of the biome information used in the UOL retrieval algorithm to determine the most appropriate coefficients to apply. It is also utilised in the UOL_3 cloud masking algorithm. In Cycle 1 snow

masking information has been provided by the Interactive Multi-sensor Snow and Ice Mapping System (IMS) Daily Northern Hemisphere snow and ice analysis.

The IMS snow maps are daily maps of Northern Hemisphere land, sea, snow and ice on an equal area polar stereographic grid at 1 km, 4 km and 24 km resolution, depending on time period. It is available from 1997 to present with higher resolution maps available for shorter time periods. For inclusion in LST_cci algorithms and products, Daily IMS maps of snow and ice presence in the northern hemisphere at a resolution of 0.01° are produced by nearest neighbour interpolation [RD-6].

The use of this data will continue in Cycle 1.5, but for the final Cycle 2 it is proposed to move towards using ESA Snow Cover CCI products for snow masking providing they become available. If these are not available in time for re-processing of the LST_cci products then the IMS remains the default. Ingestion of snow Cover CCI would better satisfy the consistency requirements of CCI, but an assessment of the applicability in comparison to IMS will be a prerequisite.

LST-CCI-ADP-11: It is proposed to move towards using ESA Snow Cover CCI products for snow masking providing they become available for future Cycles. If these are not available in time for re-processing of the LST_cci products then the IMS remains the default.

3.2.2.3. Climate Data Records

3.2.2.3.1 Intercalibration

All instruments contributing to the ATSR-MODIS-SLSTR CDR are to be cross-calibrated with IASI data. Brightness temperatures (BTs) will be produced from IASI spectra for each test instrument using the test instrument spectral response function (SRF). The IASI swath BTs will be matched with the test instrument CCI L3U gridded BTs. Analysis of the matched BTs will enable calibration coefficients to be calculated. Thus all instruments with missions overlapping IASI will be aligned with IASI. Inter-calibration of ATSR-2 BTs will rely on the inter-calibration of ATSR-2 and AATSR which is part of the 4th Reprocessing of ATSR data, since there is no overlap with IASI:

- ❖ The IASI spectra are convolved with the SRF from the test instrument and BTs calculated for every IASI pixel.
- ❖ The IASI quality indicators are used to filter the data – data are only used if all quality flags are good.
- ❖ Each IASI observation target consists of a group of four pixels. IASI pixels have an elliptical FOV, which is circular at nadir but changes shape towards the swath edges:
 - The pixel ellipse major axis is assumed parallel to the across track direction.
 - The pre-processor estimates the angle between the semi-major axis (the across track direction) and the line of latitude through the centre of the pixel using the coordinates of each of the four pixels making up target observation.
 - The pre-processor calculates the length of the semi-major and semi-minor axes.
 - The subset of L3U grid cells falling within a circle centred on the pixel centre and with radius equal to the ellipse semi-major axis is then tested to see if the grid cell centre falls within the ellipse.
 - A rectangular latitude-longitude box which includes all the grid cells falling within the ellipse is constructed.
 - The indices on the L3U grid of the minimum and maximum, longitude and latitude of this box are output along with the IASI pixel centre latitude and longitude.

- The matchup processor takes the IASI BTs on the IASI swath and matches to grid cells on the L3U grid using the pre-processed output as auxiliary data.
- The grid cells are then tested for valid data and matching in time.

The analysis of matchups is performed for each year-month. Matchups are first filtered using the following thresholds:

- ❖ fraction of cloud in L3U average ≤ 0
- ❖ satellite zenith angle (either sensor) ≤ 10 degrees
- ❖ matchup time difference: ≤ 5 minutes

As an example, for IASI v. Terra MODIS nearly all of the matchups occur at latitudes polewards of 60° . The statistics were found to be very noisy in the Northern Hemisphere. This is likely due to differences between the land-sea masks of the MODIS data and that used in the L3U binning. These differences are exacerbated by the very large number of small islands in the Arctic. Therefore, only the Antarctic matchups are used in the inter-calibration. Day and night matchups are used to ensure data are available for all seasons. BT differences due to solar heating over the matchup time difference are assumed to be negligible over ice in this region.

The statistics are used to estimate a time dependent bias by fitting a straight line function to the monthly bin means for the latitude bin 90° S to 60° S. The line is fit using least absolute deviation. In the case of Terra MODIS there are very low numbers of matchups from the start of the IASI mission to late 2008. This is due to poor quality Terra MODIS data which will be rectified in later re-processing. Only data from 01 January 2008 were used in the line fitting for Terra MODIS (Figure 5 and Figure 6).

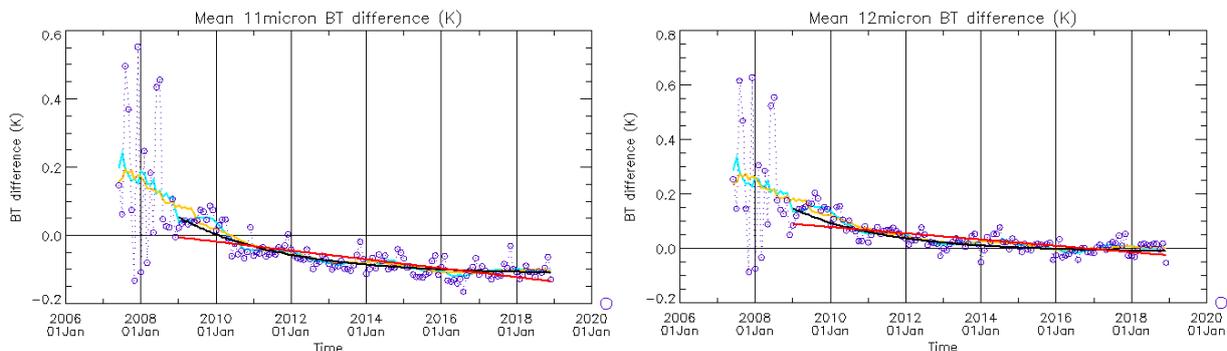


Figure 5: Monthly mean IASI minus Terra MODIS 11 micron (left) and 12 micron (right) BT difference (K) in the Antarctic. The red line is a straight line fit to data from January 2009 to December 2018. An exponential fit to the same data (black) and a 11 (blue) and 23 (yellow) month running means are also shown.

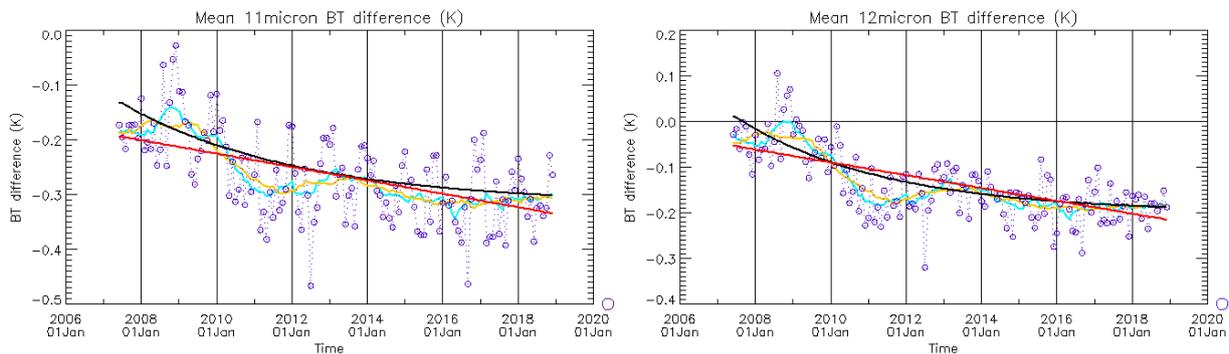


Figure 6: Monthly mean IASI minus Aqua MODIS 11 micron (left) and 12 micron (right) BT difference (K) in the Antarctic. The red line is a straight line fit to data from January 2009 to December 2018. An exponential fit to the same data (black) and a 11 (blue) and 23 (yellow) month running means are also shown.

An output look-up table (LUT) will be applied to the Level-1 BTs prior to the Level-2 retrieval algorithm processing.

LST-CCI-ADP-19: Apply intercalibrations to the Level-1 BTs for LEO data which is input to the ATSR-MODIS-SLSTR CDR.

3.2.2.3.2 Time correction

In addition to intercalibration, instruments contributing to the ATSR-MODIS-SLSTR CDR need to be harmonised to a common local overpass time. Differences in overpass time between sensors of a series result in step changes in LST. This occurs for example between ATSR-2, overpass time 10:30 and 22:30, and AATSR with overpass time of 10:00 and 22:00. Since there is a data gap between the end of AATSR mission and start of SLSTR mission, LST from Terra MODIS is being used to fill the gap but again there are differences in overpass time:

- ❖ We estimate a correction to the LST due to the time difference using Terra MODIS as the reference.
- ❖ Matchups are produced from gridded data (L3U) using a time window of 40 minutes.
- ❖ Matchups were restricted to data where the difference in satellite zenith angle was less than 10 degrees.
- ❖ The matchup LST differences are binned to multi-dimensional histograms: day or night (on reference solar angle), observation time difference (1 minute bins), latitude (10 degree bins), and landcover class.
- ❖ Multi-year histograms for each calendar month are then calculated.

The LST correction is read from the LUT (given the time of day, location, landcover class, and observation time difference between the acquisition time and the nominal time for the CDR for the given pixel) then applied at the L3U level on the 0.05° grid to the sensor to be corrected.

LST-CCI-ADP-20: For input to the ATSR-MODIS-SLSTR CDR apply time corrections to the Level-3U LST for AATSR and SLSTR to bring to same nominal local overpass time as ATSR-2 and Terra-MODIS using Terra-MODIS as the reference sensor.

3.2.2.3.3 Merged IR CDR

The Merged IR CDR will be produced on a global scale eight times daily (00h, 03h, 06h, 09h, 12h, 15h, 18h, 21h, UTC times). The nominal spatial resolution will be a 0.05° equal angle latitude-longitude grid. The individual stages are described in detail below:

- ❖ Stage 1: apply consistent Level-2 algorithms
 - The single channel (SMW) algorithm used for both earlier GOES and MTSAT is the lowest common denominator in ensuring consistency across all instruments of the Merged Product. To avoid both step-changes between satellites of the same series and discontinuities across the globe from multiple sensors for the same time window, it is decided to use a common algorithm (even if that increases uncertainty for some sensors contributing to the product).
 - The UOL_3 cloud mask will be applied to the Level-1 data for the individual sensors to ensure consistency.

LST-CCI-ADP-18: All LEO and GEO products will be reprocessed with a single-channel (SMW) algorithm calibrated with the benchmark dataset constructed for the Round Robin, in order to provide harmonized data for the Merged Product.

- ❖ Stage 2: bias correction
 - Following the approach developed in GlobTemperature, bias corrections to a single reference sensor will be determined. This is, in essence, the removal of systematic differences between instruments:
 - The bias correction is carried out by computing linear regressions between LST from a reference sensor and those of the sensor to be bias-corrected.
 - The computation is done for satellite viewing angles within a limit range ($\pm 5^\circ$ for zenith and $\pm 5^\circ$ for azimuth angle) and, in a first approach, for night-time only.
 - The matching between LEO and GEO data will be performed for a select number of days for each month in a given year.
 - The reference sensor chosen is MSG-SEVIRI, which is the only sensor in the ensemble which has an operational maturity level in terms of Global Space-based Inter-Calibration System (GSICS) calibration to the standard reference source [RD-22; RD-23]:
 - The LEO sensors will thus be corrected to SEVIRI.
 - For the remaining GEOs we will use MODIS as the transfer sensor.
 - The associated uncertainty from these bias correction adjustments will be determined in accordance with the strategy developed in the E3UB.
- ❖ Stage 3: angular correction
 - Satellite LST retrievals correspond to an integrated radiometric surface temperature within the sensor footprint and, therefore, remotely sensed LST is in general highly anisotropic. As such, a scene viewed by the same sensor from different viewing angles would lead to different LST retrievals.
 - Two methodologies are to be considered:
 - An approach based on the parametric model [RD-24].
 - A modified version of the model developed by [RD-25].
 - In both methodologies, the parametric models are calibrated using collocated LST products in space time, namely LST estimates from geostationary platforms (e.g., MSG, GOES) collocated with LST values retrieved from polar-orbiters (e.g. MODIS, SLSTR).
 - The angular correction will be provided as an additional variable rather than amending the output LST.

❖ Stage 4: GEO + LEO Merging

- The basis of the Merged Product will be the GEOs with the LEOs infilling the gaps at higher latitudes and over Central Asia.
- For the GEOs the principle area of overlap is over South America. In this case all data will be taken from only one instrument to minimise possible discontinuities (in this case GOES).
- The viewing geometry of the GEOs will be restricted to between 50 and 60 degrees (an assessment will be carried out to understand the impact on data coverage).
- For LEO vs. GEO superposition:
 - LEOs will not be used to infill inside the coverage of the GEO disks
 - LEOs can however be used as an additional filter on the cloud masking of the GEOs inside the coverage of the disks

❖ Stage 5: time difference characterisation

- Each GEO has a different scanning time of its full disk, and each has a different delay relative to the nominal time given in the respective file:
- We select SEVIRI/MSG to be merged with GOES and MTSAT corresponding to the same nominal time-slot (i.e., 03:00).
- In this manner, we aim to minimize the time differences between the GEO's actual observation times.
- The same assessment of the time differences will be made for LEOs. On a global scale, each UTC time step is only sparsely filled with LEO data. Nevertheless, the LEO data in many cases fills in areas of the globe unobserved by the GEO satellites:
- LEO data will be added to the 3-hourly UTC product where that data falls within a certain threshold of the UTC time. For the GEOs this is governed by their scans (-28 minutes for MTSAT, through to +12 minutes for MSG). To simplify we will use a ± 30 minutes inclusion threshold for the LEO data.
- To investigate correction to a common UTC we will apply the LUT from Section 3.2.2.3.2.

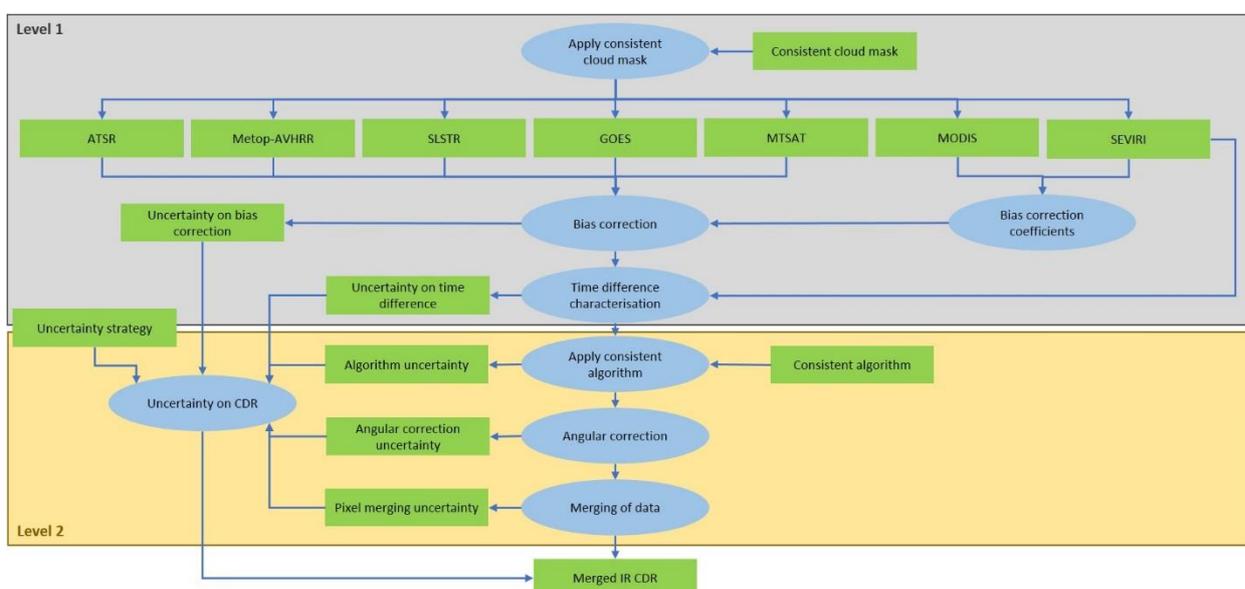


Figure 7: Schematic representation of the development and implementation of the Merged IR CDR.

 land surface temperature cci	ALGORITHM DEVELOPMENT PLAN <i>WP2 – DEL-2.4</i>	Ref.: LST-CCI-D2.4-ADP Version: 4.0 Date: 16-Mar-2023 Page: 34
--	---	---

3.2.2.4. Uncertainty Model for Thermal Infrared Algorithms

A key gap in the first version of the End-to-End ECV Uncertainty Budget Report (E3UB) [RD-31] and thus the first products was an uncertainty due to total column water vapour (TCWV) on the “application of the coefficients”. A corresponding uncertainty for the TCWV on the “derivation of the coefficients” was already implemented.

All the selected infrared retrieval algorithms in Section 3.1 are dependent on the specification of the total column water vapour. The TCWV estimate at a given instant and location is obtained from Numerical Weather Prediction (NWP) models. A commonly used measure of uncertainty in NWP models is the ensemble spread that is generally available together with the variable best estimate. The ensemble systems consist in a set of model runs with perturbed initial conditions; some systems also include perturbations to the model physics, more than one model within the ensemble or different physical parametrization schemes. Processing of ensemble data can be quite demanding and, therefore, we will approximate the instantaneous TCWV spread to a climatology of this spread that depends on actual TCWV, latitude and month.

LST-CCI-ADP-21: *Implement the uncertainty component due to TCWV on the application of the coefficients for all LEO and GEO infrared products.*

3.2.3. Version 4.00

For Cycle 3 and all subsequent cycles LST single instrument ECV Products should not be constrained by the lowest common denominator when a more optimal algorithm could be applied should the instrument have a different channel configuration. In each case a consistent algorithm will be applied across all sensors where no degradation in performance is encountered:

- ❖ for single channel sensors then a consistent algorithm will be applied
- ❖ for split-window sensors then a consistent algorithm will be applied that maximises the information in the two channels rather than loss of information through use of a single channel algorithm.

LST-CCI-ADP-23: *All LEO and GEO products will be reprocessed with the best algorithm to maximise the capabilities of each instrument rather than degrade to common algorithms.*

3.2.3.1. Auxiliary Datasets for Thermal Infrared Retrievals

The following section gives a description of the planned updates to auxiliary datasets to be used for the thermal infrared retrieval algorithms.

3.2.3.1.1 Emissivity

It is anticipated that a new and improved emissivity dataset for MODIS shall be available through the LST_cci Work Package 2.9: Temperature and Emissivity Separation from MODIS multispectral TIR data (CCN to Baseline Project). For all other datasets CAMEL V2 will be retained as an input. For sea-ice, CAMEL is not available, therefore we will use ice emissivity information from the ECOSTRESS Spectral Library, as an estimate of sea ice emissivity, which are in agreement with values quoted in the literature [RD-38, RD-39].

3.2.3.1.2 Sea-ice masking

One of the biggest challenges in producing a CDR including sea-ice is identifying sea-ice observations not contaminated by cloud. We start here with classification of the surface itself using knowledge of the global sea-ice cover. Considering the relatively high spatial resolution of the IST data the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) daily sea-ice analysis [RD-40] will be used to identify snow / ice pixels over the sea. The OSTIA system is a daily global gap-free dataset which includes sea-ice area fraction produced on a 0.05° equal angle grid. The sea-ice concentration, defined as the local area fraction of a given grid point that is covered by ice, is derived from passive microwave satellite measurements. The required sea-ice concentration at 1/120° is produced by interpolating the input OSTIA data onto the LCCS biome grid. Pixels with sea-ice concentrations greater than 50% are designated as sea-ice, using a new biome class of “230” in the LCCS classification system.

3.2.3.2. Climate Data Records

3.2.3.2.1 Orbital drift correction

For the development of the AVHRR/3 products from the NOAA satellites it is necessary to address the issue of orbital drift (Figure 8). This results in an artificial gradient over time in retrieved LST. To maximise knowledge within LST_cci the method of [RD-41] will be implemented since this offers a suitable approach for bulk processing across the AVHRR archive.

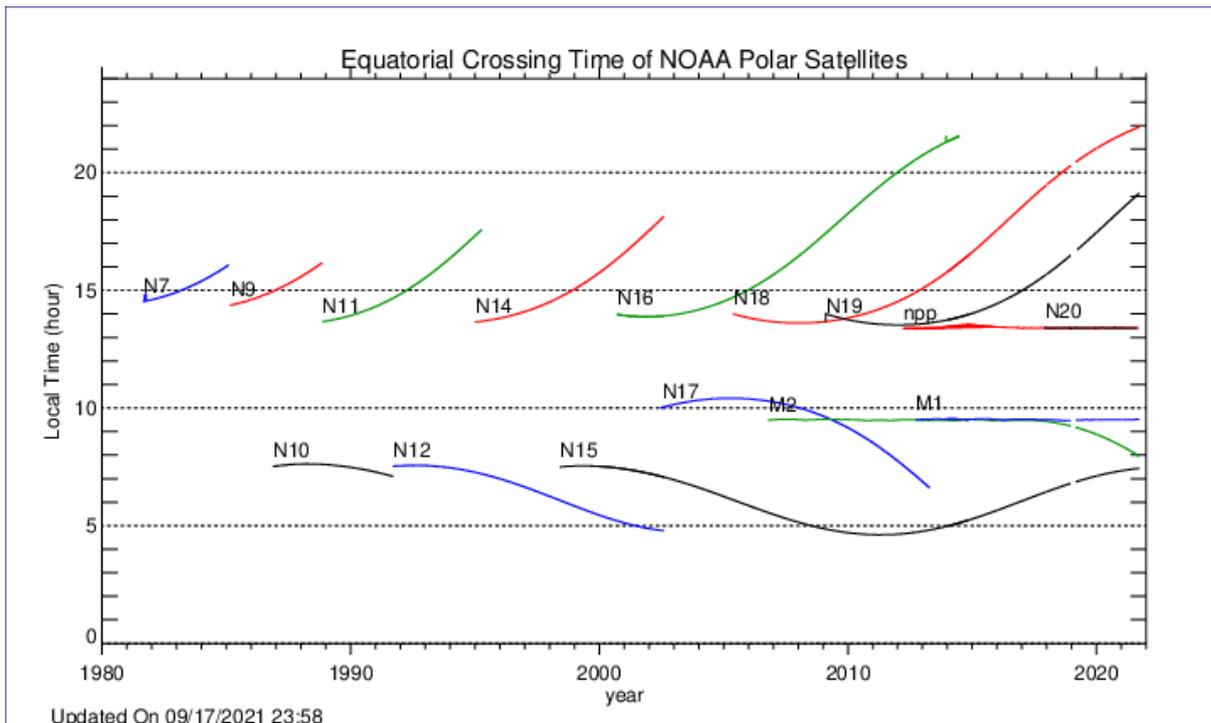


Figure 8: AVHRR Equator Crossing Times.

This method of [RD-41] involves adjusting retrieved LST time series on the basis of statistical information extracted from the time series themselves. It is simple and straightforward, in order to be implemented easily for large datasets. The correction is applied on a per-pixel basis, and relies on a second order polynomial fit of solar zenith angle (SZA) anomalies against time per satellite. If the pixel time series is identified as contaminated by the orbital drift for any of the different satellite active periods, LST

anomalies are fitted linearly against both time and the second order polynomial fit of SZA anomalies. This double fit allows for the removal of orbital drift influence without removing climate trends in the data. When applied to LST time series, the approach normalizes the distribution of LST values at the beginning and end of each satellite activity period, and visual inspection of the time series does not show any residual orbital drift in the corrected LST time series.

LST-CCI-ADP-24: *Implement an orbital drift correction method for NOAA-AVHRR LST products and other TIR satellites which may experience orbital drift during part of their lifetimes.*

3.2.3.3. Prototype Products

3.2.3.3.1 High resolution retrievals

Processing Cycle 3 will include the first retrievals made for Landsat data. This will include Landsat-7 and -8, although neither of these Landsat satellites whilst possessing thermal instrumentation have the capabilities for split-window resulting in a high reliance on ancillary data for atmospheric and surface knowledge. Landsat-8 does have two thermal bands but has suffered from significant stray light problems in the second thermal channel [RD-42], and therefore this channel will not be used here. A single channel Landsat OE retrieval will be ill-posed due to the single trusted observation TIR channel, therefore retrieval of both LST and LSE is not possible. To provide high resolution and accurate surface estimates for the emissivity in this retrieval the Visible-Near IR (VNIR) channels of Landsat are exploited.

The OE methodology allows the retrieval to use both the a priori information (from surface estimates) and the observation without being overly tied to one in particular. Radiometric issues in Landsat can be adapted to, and errors in surface state knowledge can be minimised in the iterative process. The methodology also allows uncertainty estimates in the final estimate from mixed sources without requiring detailed knowledge on the individual components a priori. The retrieval will simply ingest the best prior uncertainty knowledge we have and propagate it according to which parameters impact on the OE retrieval.

3.2.3.3.2 Downscaling

Processing Cycle 3 will exploit the LST_cci SLSTR LST Product in a new way. A downscaled LST product will be derived by taking the SLSTR LST product at 1 km resolution and using Sentinel-2 derived LSE data to iteratively update the LST in an OE scheme. In this methodology the OE scheme will assimilate the medium resolution SLSTR LST and the Sentinel-2 LSE and attempt to minimise the difference between the SLSTR calculated Bottom-Of-Atmosphere (BOA) brightness temperatures (BTs) and simulated BOA BTs generated from the SLSTR LST and the Sentinel-2 LSE. This would be an operation performed at the resolution of the Sentinel-2 data.

The processing first obtains the normalised difference vegetation index (NDVI) from the Sentinel-2 data and maps the SLSTR LST data onto the Sentinel-2 NDVI pixel grid. The NDVI data is used with the NDVI threshold method of [RD-43] to estimate the LSE.

The re-gridded SLSTR LST combined with the original coarse resolution LSE associated with the pixel in the Split-Window processing (whether this be explicit or through biome estimation) is used with Planck's Law to estimate the BOA BT values. In order to transfer the traceable uncertainties from the Split-Window algorithm, the random component of the total uncertainty from the SLSTR LST algorithm is applied as a noise on these BOA BT values using a Gaussian random distribution.

The SLSTR LST is then used with the Planck function again, but this time with the LSE derived from the Sentinel-2 NDVI threshold method. Two sets of BOA BTs are evaluated and the differences used to iteratively update the estimate for LSE used in the retrieval until an empirically determined threshold is reached from where there is no longer any significant improvement to be made by further iteration.

This methodology produces a full uncertainty breakdown including both the input total uncertainty of the SLSTR LST retrieval as well as the uncertainties due to the downscaling. The use of the SLSTR LST means that any additional uncertainties are a result of the retrieval process and the surface parameters, as the atmospheric uncertainty should have been fully captured in the SLSTR LST retrieval.

LST-CCI-ADP-25: *Utilise Sentinel-2 data and the Optimal Estimation approach to downscale LST from moderate resolution instruments.*

3.2.4. Version 5.00

3.2.4.1. UOL_3 Algorithm

For all LEO single-sensor products and LEO inputs to the Merged IR CDR we will use the best profile data for simulating the cloud state of each pixel. Thus we will use ERA5 data, which will minimise any non-linearity between adjacent time steps since the profiles are only 1-hour apart. The higher spatial resolution of ERA5 is also expected to improve the outputs of the bilinear interpolation onto each sensors tie-point grid. ERA5 is chosen rather than ERA5-Land since we need to maintain consistency across all land surfaces and sea-ice.

LST-CCI-ADP-28: *Implement the 1-hourly ERA5 profile data for cloud masking all LEO data.*

3.2.4.2. Climate Data Records

3.2.4.2.1 Intercalibration

Intercalibration differences between sensors mean that different brightness temperatures (BTs) would be derived even in the hypothetical situation where they were observing the same field of view, within a similar spectral band, at the same time and from the same view angle.

Harmonisation of Level-1 data across all sensors for the LST_cci CDRs entails adjustment of the BTs to a reference sensor. The Global Space-based Inter-Calibration System (GSICS) have used the Infrared Atmospheric Sounding Interferometer (IASI) as the reference sensor in a calibration of SEVIRI radiances. IASI is a Fourier transform spectrometer and provides infrared spectra with high resolution (0.5 cm^{-1} after apodisation, L1C spectra) between 645 cm^{-1} and 2760 cm^{-1} ($3.6\text{ }\mu\text{m}$ to $15.5\text{ }\mu\text{m}$).

Previously, we followed the approach of GSICS to determine intercalibration coefficients for MODIS for example. In the matchup process for MODIS nearly all of the matchups occur at latitudes polewards of 60° . In the next version we will develop and refine the approach across the full CDRs. Taking the ATSR-S3 CDR as an example, all contributing instruments are to be cross-calibrated with IASI data. BTs will be produced from IASI spectra for each test instrument using the test instrument SRF. The IASI swath BTs will be matched with the test instrument LST_cci Level-3 Uncollated (L3U) gridded BTs. Analysis of the matched BTs will enable calibration coefficients to be calculated. Thus all instruments with missions overlapping IASI will be aligned with IASI. Inter-calibration of ATSR-2 BTs will rely on the inter-calibration of ATSR-2 and AATSR [RD-48] which is being implemented as part of the 4th Reprocessing of ATSR data, since there is no overlap with IASI. To extend the dynamic range ocean pixels will also be included in the matchups. The following steps are taken:

- ❖ The IASI spectra are convolved with the SRF from the test instrument and BTs calculated for every IASI pixel.
- ❖ The IASI quality indicators are used to filter the data – data are only used if all quality flags are good.
- ❖ Each IASI observation target consists of a group of four pixels. IASI pixels have an elliptical FOV, which is circular at nadir but changes shape towards the swath edges:
 - The pixel ellipse major axis is assumed parallel to the across track direction.
 - The pre-processor estimates the angle between the semi-major axis (the across track direction) and the line of latitude through the centre of the pixel using the coordinates of each of the four pixels making up target observation.
 - The pre-processor calculates the length of the semi-major and semi-minor axes.
 - The subset of L3U grid cells falling within a circle centred on the pixel centre and with radius equal to the ellipse semi-major axis is then tested to see if the grid cell centre falls within the ellipse.
 - A rectangular latitude-longitude box which includes all the grid cells falling within the ellipse is constructed.
 - The indices on the L3U grid of the minimum and maximum, longitude and latitude of this box are output along with the IASI pixel centre latitude and longitude.
 - The matchup processor takes the IASI BTs on the IASI swath and matches to grid cells on the L3U grid using the pre-processed output as auxiliary data.
 - The grid cells are then tested for valid data and matching in time.
- ❖ The analysis of matchups is performed for each year-month. Matchups are first filtered using the following thresholds:
 - fraction of cloud in L3U average ≤ 0
 - satellite zenith angle (either sensor) ≤ 10 degrees
 - matchup time difference: ≤ 5 minutes
- ❖ An output look-up table (LUT) is applied to the Level-1 BTs prior to the Level-2 retrieval algorithm processing.

LST-CCI-ADP-29: Update the intercalibration using ocean matchups to extend the dynamic range.

3.2.4.2.2 Orbital drift correction

Different methods have been, and are being, evaluated in LST_cci for implementing the Orbital Drift Correction (ODC) for the NOAA satellite series. The first method of [RD-41] involves adjusting retrieved LST time series on the basis of statistical information extracted from the time series themselves. A correction is applied on a per-pixel basis, and relies on a second order polynomial fit of solar zenith angle (SZA) anomalies against time per satellite. **Unfortunately, the fit removes both the orbital drift influence and any climate trends in the data and so is not deemed suitable for use in LST_cci.**

Two feasible methods are therefore being evaluated for selection of the most appropriate approach. The selected approach will then be implemented.

Adapting the LST_cci orbital drift correction from Microwave

For the SSM/Is the orbit oscillates around 6am/6pm local time, with a maximum drift of approximately 50 mins, but the correction for change in time works across the swath so could be up to 3 hours at the poles.

The approach taken in LST_cci for the MW is to use another space craft which observes on the same day close to the reference time. We estimate the $dLST/dt$ as $LST1-LST/(t1-t2)$ then estimate delta LST as $\Delta LST = dLST/dt \times \Delta t$ where delta is the difference between corrected and observed.

This relies on LST change being linear during the period covering the LSTs from the two spacecraft and the LST at the reference time. This is unlikely to be true for the whole time series as the 6am chosen reference time for the NOAA products is close to the dawn inflexion point and the 6pm reference time is close to the time when the linear cooling during the afternoon begins to slow.

For the MW this was addressed by binning the slopes into three bins (before 5am, 5-7am and after 7am) according to the observation time of instrument to be corrected. However, the timing of the change in slope of the LST change will be seasonally varying and also likely to vary with cloud cover; these factors are not accounted for. Also, it is necessary to ensure both spacecraft observations are on the same side of the inflexion point which may mean using more than one spacecraft, if available, to estimate the slopes and complicates the processing as it is necessary to know when the inflexion point occurs for a particular location and time of year.

Utilising coincident profile data to model time corrections

One drawback with adapting the MW approach to the IR is that the NOAA drift in time is considerably greater than MW DMP drift. This matters since the temperature change is non-linear. This is particularly pertinent at the chosen reference local time for the NOAA data products (6am, 6pm) as the gradients around these points are highly variable.

An alternative approach we are experimenting with uses a radiative transfer model and simulates hourly BTs for 24 hours around the actual observation time at each pixel. The inputs to radiative transfer model include ERA-5 profiles and skin temperature, and CAMEL emissivities.

The concept is that we determine the BT_{sim} difference between pixel observation time and reference time and use as the correction. If either the observation is cloudy or it is cloudy at the reference time then no BT is calculated and the pixel "lst_correction" is set as a fill value. We can use in situ data to validate the correction and to provide a measure of the uncertainty.

Once the ODC has been determined per pixel it will be stored in a separate "lst_correction" field in the output product. This ensures the users have access to the uncorrected LST (ie the actual observation) and maintains consistency with the approach taken for the Microwave CDR.

LST-CCI-ADP-30: Add the Orbital Drift Correction as an additional field to the product, keeping the original uncorrected LST as the default LST field.

4. Algorithm Development Plan for Microwave LST Products

4.1. Current status of Microwave LST Products

In the LST_cci open algorithm intercomparison round-robin, the performance of different LST retrieval algorithms to produce the Special Sensor Microwave/Imager (SSM/I) and Special Sensor Microwave/Imager and Sounder (SSMIS) LST ECV was assessed to identify the best algorithm for a future climate quality operational system. The algorithm chosen for LST_cci Cycle 1 MW product was the Neural-Network-Emissivity-All-channels (NNEA) algorithm. The algorithm is based on a non-linear regression implemented by a neural network having as inputs the brightness temperatures from all the sensor channels and a monthly-mean climatology of the emissivities at the corresponding observing frequencies. Full details about the algorithm and the calibration database can be found in the ATBD [RD-29].

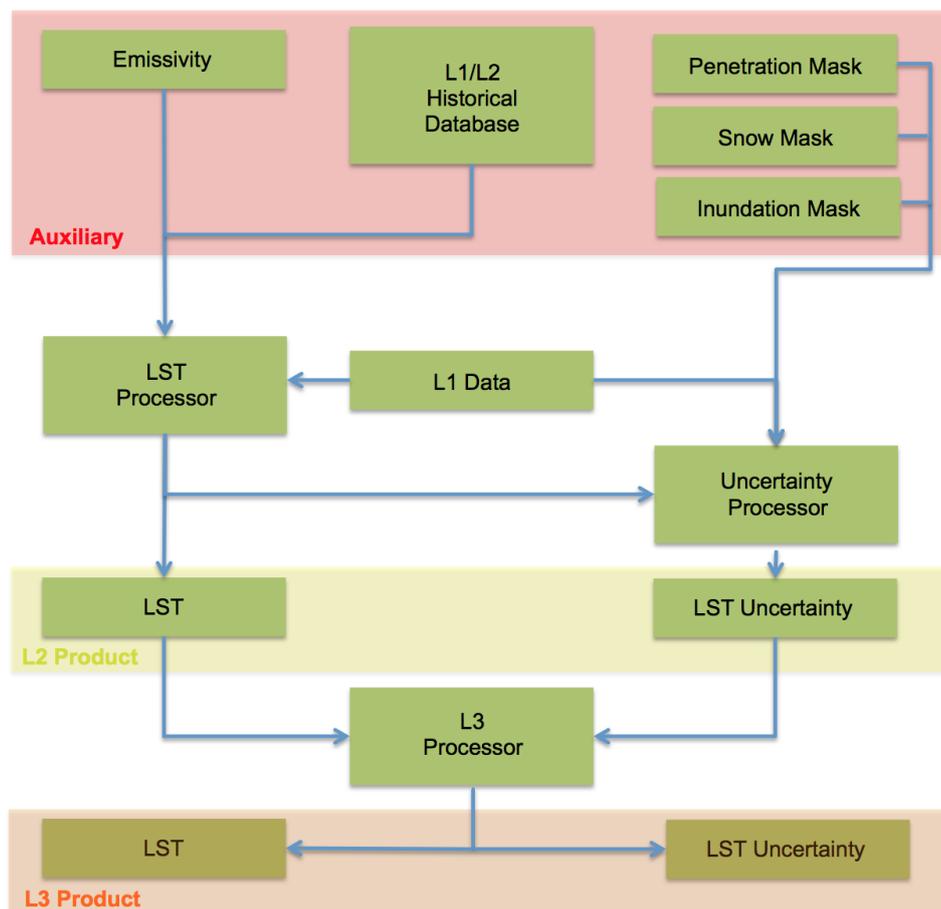


Figure 9: Data flows for the SSM/I and SSMIS LST ECV prototype production system.

Following the algorithm selection, the processing chain illustrated in Figure 9 was prototyped with a LST processor built around the NNEA algorithm. The L1/L2 Database used to calibrate the algorithm was extended to cover 4 years of real SSM/I and SSMIS observations to increase the sampling of different atmospheric and surface conditions. As source of L1 Data, the MW Imager FCDR [RD-13] was selected and the brightness temperatures of the SSM/I F13 and SSMIS F17 were downloaded from the EUMETSAT Climate SAF archive.

 land surface temperature cci	ALGORITHM DEVELOPMENT PLAN <i>WP2 – DEL-2.4</i>	Ref.: LST-CCI-D2.4-ADP Version: 4.0 Date: 16-Mar-2023 Page: 41
--	---	---

4.1.1. Version 2.00

Given the particularities of the MW LST retrievals, a quality flag processor different from the IR counterpart was implemented targeting a different set of flags. The processor produces flags to signal: (1) snow-covered surfaces, by using a snow flag currently derived from the snow density and snow depth from the ERA-Interim reanalysis; (2) inundated surfaces, by using estimates from the Global Inundation Extension from Multi-Satellites (GIEMS-2) [RD-14]; (3) coastal areas, using a database of distance to the coastline; (4) surfaces with large microwave penetration depth where emission can be emanating from subsurface layers, by using a monthly climatology of radar backscattering from [RD-15], and; (6) surfaces where there can be convection activity at the overlying atmosphere, by using the cloud flag described in [RD-16]. A final flag informing about the confidence in the retrieved LST value has also been added by looking at the convergence of multiple retrievals of the same scene by the NNEA algorithm under slightly different inversion initial conditions, as detailed in [RD-29].

Using the described processing prototype, MW LST from SSM/I and SSMIS was produced, covering the period 1998-2008 and 2009-2015, respectively, and uploaded to the esacci_lst workspace in the UK JASMIN environment in the netcdf format specified at the Product Specification Document (PSD) [AD4]. The expected period for the MW LST production for Cycle 1 was 1998-2016, so the Cycle 1 production was missing 2016 due to the MW Imager FCDR still not providing brightness temperatures for that year.

For Cycle 1.5, the expected period to be covered was 1998-2018. Following the proposal to include the full period from ATSR-2 starting in 1995, we also proposed to process the SSM/I LST starting from 1995, and to add a further year from 2018 to complete a 1995-2019 data record. Concerning the forward extension of the MW LST data record, the years 2016 and 2017 were already available as a beta version and were kindly provided to Estellus for a first assessment of data quality. A problem with one of the channels in the F17 SSMIS sensor was confirmed, so a switch to the F18 satellite for the 2016-2019 years was required. These years were also available as beta versions, and together with the remaining years used to complete the 1995-2019 production for Cycle 1.5.

4.1.2. Version 3.00

For Cycle 2 no further years are expected to be processed, so a 1995-2019 data record will be produced. There may be a possibility that the year 2020 is available at the EUMETSAT Climate SAF archive before the project ends. In that case, a voluntary extension of the data record to cover the period 1995-2020 could be considered.

Work on the uncertainty processor has started, and it is fully described in [RD-29]. It provides an estimation of theoretical uncertainty based on an analysis of the NNEA algorithm errors when inverting the brightness temperatures of the L1/L2 database. The standardised uncertainty model adopted for the retrievals of some of the IR sensors has not been fully adopted yet, and work in this direction is already planned for the next production cycle (Cycle 2).

LST-CCI-ADP-16: *The next production Cycle 2 will begin to implement concepts from the standardised uncertainty approach adopted for IT sensors to the microwave sensors.*

4.1.3. Version 4.00

For Cycle 3 AMSR-E and AMSR2 will be processed ready to be added to the merged microwave product. AMSR-E has channels at frequencies close to the SSM/I and SSMIS sensors, and observe at close angles. However, for a simultaneous inversion of SSM/-SSMIS-AMSR-E-AMSR2 observations, a consistent

 land surface temperature cci	ALGORITHM DEVELOPMENT PLAN <i>WP2 – DEL-2.4</i>	Ref.: LST-CCI-D2.4-ADP Version: 4.0 Date: 16-Mar-2023 Page: 42
--	---	---

database able to provide the required brightness temperatures and corresponding retrieved LST is not available, and it needs to first be created. To achieve this consistency two elements are critical:

- ❖ a dataset of inter-calibrated radiances
- ❖ a common setup to invert the radiances from all the instruments using a unique radiative transfer model and similar a priori information to initiate the retrieval.

It is expected that such database will greatly mitigate possible biases in the microwave LST of the merged SSM/I-SSMIS-AMSR-E-AMSR2 product.

Given these requirements two inversion algorithms are needed:

- ❖ an algorithm based on the optimal estimation (OE) methodology to produce the calibration database [RD-36]
- ❖ a neural network algorithm to be calibrated with a dataset based on the previous optimal estimation retrievals.

The OE algorithm is described in Section 3.1.3.1, and needs to be developed for the microwave application, while the neural network algorithm will be identical to the inversion algorithm of previous versions, so no further developments are required apart from a recalibration with the new dataset.

4.1.4. Version 5.00

For Cycle 4 the Optimal Estimation and neural network algorithms are in place, together with the new calibration database, and no new developments are needed in term of algorithms. The LST_cci MW data records will be extended till 2023 for SSMIS and AMSR2.

4.2. Algorithm Development Plan for Microwave Sensors

4.2.1. Version 2.00

For Production Cycle 1.5 the following activities are planned:

1. Revision of LST Processor if any findings of the validation activities show shortcomings that can be reasonably tackled by modifying the current algorithm. Given the reasonable performance of the NNEA algorithm presented in [RD-30] compared with the other algorithms, we do not foresee a change of algorithm for the next production cycles, but only minor improvements to the current NNEA algorithm if robust improvements in retrieval performance can be demonstrated.
2. Work on the Uncertainty processor to implement the algorithms specified at the End to End Uncertainty ECV Budget Model.
3. Revision of the Quality Flag Processor to improve the current flags. We foresee work carried out to improve: (a) the inundation flag, by using actual estimates of inundation from [RD-14] instead of the current estimates based on a monthly climatology derived from the same inundation product; (b) the snow flag, by using the ESA CCI snow product if it becomes available, and; (c) the convection flag, by revising the current cloud detection algorithm presented in [RD-16] for the surfaces where the detection is presently more challenging.

LST-CCI-ADP-17: The Quality flag Processor for microwave sensors will be improved for: (a) the inundation flag, by using actual estimates of inundation; (b) the snow flag, by using the ESA CCI snow product if it becomes available; and (c) the convection flag, by revising the current cloud detection algorithm.

4.2.2. Version 3.00

No major changes are foreseen in the algorithm chosen for the LST_cci Cycle 2 MW product. The NNEA algorithm works as expected and no further developments will be carried out. The revision of the quality flag processor was already implemented for Cycle 1.5, and the same processor will be used in the Cycle 2 production. Only the Uncertainty Processor will be revised again to see if the MW LST uncertainty can be further characterized for the final Cycle 2 product, as discussed above.

4.2.3. Version 4.00

The following steps will be carried out to develop the microwave product:

- ❖ Compilation of datasets required for the inversions.
 - The CM-SAF SSM/I and SSMIS L1 data used in previous versions will be replaced by the new version V4, with expected improvements regarding the intercalibration of both instruments. CM-SAF is not providing the AMSR-E and AMSR2 L1 data records, but they will be providing intercalibration coefficients to remove biases between the SSM/I-SSMIS and the AMSR-AMSR2 family of instruments.
 - The OE inversions require ancillary data to initiate the retrievals. The a priori LST and atmospheric conditions will be sourced from the ERA5 reanalysis to have a consistent description of the surface and atmosphere available for both clear and cloudy conditions [RD-44]. An a priori surface emissivity value at the different frequencies and observation angles is also needed as a starting point of the retrieval, and it will be inferred from the frequency and angular parameterizations available at TELSEM [RD-45].
- ❖ Implementation of the OE algorithm
 - The OE algorithm will be built based on previous developments [RD-46, RD-47], and will be coded to handle the different frequencies and observation angles of the four instruments.
- ❖ Production of an initial inversion dataset
 - The OE algorithm will be initially run for 4 years, e.g., 2007 for SSM/I and AMSR-E, and 2015 for SSMIS and AMSR2. The LST retrievals will be assessed in terms of errors, difference between instruments, and processing aspects.
- ❖ Evaluation of inversion strategy
 - Based on the previous analyses, a decision will be taken regarding the full-time production, i.e., whether to expand the database a few more years and implement a fast neural network processor as in Phase-1, or to carry out the full data record production with the OE algorithm.

LST-CCI-ADP-26: Implement the Optimal Estimation approach in developing the microwave product.

4.2.4. Version 5.00

No major changes are foreseen in the algorithm chosen for the Cycle 4 LST_cci MW product.



End of document