ALGORITHM VALIDATION PLAN FOR SMOS LEVEL 2 SMPPD

Project code: SO-TN-CBSA-GS-0015-

Version: draft 1.g

Date: 09/11/06

	Role	***Name***	***Date and signature***
Written by :		P Waldteufel	
Approved by:			
Approved by:			

DOCUMENT STATUS SHEET

Version / Rév.	Date	Sections	Main changes and comments	Origin	Visa
005	18/04/2006	all	Partial predraft	CBSA	
005MJT	20/04/2006	all	Partial predraft	CBSA + RU	
007	31/05/2006	all	Pre-Draft	CBSA	
009	02/06/2006	all	Pre-Draft	CBSA	
010/**1a**	07/06/2006	all	009 + ARRAY comments & answers	ARRAY + CBSA	
011	20/06/2006	All	010 + progresses; introduce drafts for Table 3 & Table 4	CBSA	
012/**1b**	10/07/2006	All	011 + updating tools & tables	CBSA + ESL remarks	
014/**1b**	20/07/2006	All	012 + more account of ESL remarks	CBSA + ESL remarks	
014/1c	30/08/2006	6.3	014/1b + dealing with PreQR # 10 & 11	CBSA	
015/**1e**	05/09/2006	All	Integrate section 6 to 2 to 4 Develop phase 1. Remove annex	Following PM6	
015/**1f**	27/09/2006		Revised According to QR item 53		
015/**1f**	27/09/2006	Status sheet	Delete notes	QR RID 69 1	
015/**1f**	27/09/2006	Distribution	Corrections	QR RID 69 2	
015/**1f**	27/09/2006	6	Added for plan showing feasibility	QR RID 69 3	
015/**1f**	27/09/2006	2	All TBC and TBD removed in phase 1	QR RID 69 4	
015/**1f**	27/09/2006	1.2 & 6	Provide for open issues in phases 2 & 3	QR RID 62	
015/**1f**	27/09/2006	2.2.3.1	State L1PP needs for phase 1	QR RID 61	
015/**1f**	27/09/2006	2.2.3.1	Size of WA consistently = 123 km	QR RID 60 1	
015/**1f**	27/09/2006	2.2.3 & 4	Move Faraday & rotation biases to auxiliary data for retrieval	ESL & ESA	
015/**1f**	27/09/2006	2.2.3 & 4	Clarifications and modifications	QR RID 60 2 & 3	
015/**1f**	27/09/2006	4.3.5	Discuss validation on test sites	QR RID 59 4	
015/**1f**	27/09/2006	2.4 & 3.5	Outline success criteria	QR RID 58 3	
015/**1f**	27/09/2006	6	Plan	QR RID 58 1, 2, 4	
016/**1g**	09/11/2006	2.2.2	Update table	ESL	
016/**1g**	09/11/2006	2.2.4	Update tables and bias conditions	ESL	
016/**1g**	09/11/2006	6.1	Update plan		

DISTRIBUTION LIST

ESTEC:	
A. Hahne	
S. Delwart	
M. Zundo	
J.-C. Debryun	

ESRIN:	
N. Wright	
S. Pinori	
ESTEC:	

CNES	
F. Bermudo	
R. Moreno	
SAG	
All	
OTHER	
ESL	

REFERENCES

Reference	Documents
RD1	SO-TN- ESL-GS-0001_5.a Level 2 SM ATBD
RD2	SO-TN-ARR-L2PP-0005_TGRD_v0.3 Level SM TGRD
RD3	SMOS-FR-ACR-SA-007 Soil moisture retrieval for the SMOS mission, Final Report, 88 pp
RD4	SO-TN-CBSA-GS-0014 PRELIMINARY STATISTICS ON SMOS LAND WORKING AREAS
RD5	SO-IS-DME-L1PP-0002 SMOS L1 Product Format Specification

ACRONYMS

Acronyms	
ALC	Aggregated land cover fraction
AlgoValP	Algorithm Validation Plan
ASTD	*A priori* uncertainty (floating parameters)
ATBD	Algorithm Theretical Basis Document
BB	BreadBoard
BT	Brightness temperature
CARD A	Dielectric constant index obtained from cardioid model
DAP	Data Analysis Product
DFFG	Discrete Flexible Fine Grid
DGG	Discrete Global Grid (SMOS grid)
DQX	Theoretical retrieval incertainty
DT, DT1, DT2	Decision Tree; Decision Tree stage 1, stage 2
DTB	Radiometric uncertainty
EAF	Extended Alias Free
ECOCLIMAP	Land cover map on high resolution grid
ESL	Expert Support Laboratories
FOV	Field of View
GQX	Overall retrieval quality figure
L1 TS	Level 1c product (test scenario for acceptance review)
ISEA	Icosahedral Snyder Equal Area
L1PP	Level 1 processor prototype
LC	Land cover (high resolution)
LUT	Look Up Table
M11, M21, M31, M41, M51, M61	Short for auxiliary data
M12, M22? M32, M42, M52, M62	Short for biased auxiliary data
M SWATH	Maximum allowed swath extent
MVAL	Index for data information content
NIT	Number of iterations
RFI	Radio Frequency Interferences
SL	Short for list of values in SM loops
SM	Soil Moisture
SMPPD	Soil Moisture Processor Prototype Development
T SCENE	Scene (or antenna) temperature
TAU	Opacity Optical Thickness
TBL1 to TBL6	Designation of simulated datasets
TEC, TECu	Total Electronic Content and unit
TGRD	Table Generation Requirement Document
TH 23, TH 34	TAU thresholds driving DT2
TH MIN1, TH MIN3	MVAL thresholds driving DT2
TL1, TL2	Short for list of values in TAU loops
TX	Pol TB at antenna level
TY	Pol TB at antenna level
UDP	User Data Product
UPF	User's Parameter File (LUT)
WA	Working area enclosing SMOS pixel
WEF	Weighing Function
X SWATH	Location (abscissa) across the swath
XL	Short for list of values in X SWATH loops

TABLE OF CONTENTS - PART B

1. SCOPE AND OUTLINE

1.1 SCOPE

This document describes the Algorithm Validation Plan (AlgoValP) for SMOS Level 2 Soil Moisture Prototype Processor Development (SMPPD). It provides a description of tests to be performed in order to check the processor's performances with respect to SMOS mission's science requirements.

It is assumed that software verification operations have been completed before undertaking validation, i.e. the prototype has been verified to operate identically to the breadboards (BB). Since both are built independently, this supplies strong evidence for both being also consistent with the ATBD [RD1].

The overall validation for the SMOS L2 retrieval algorithm is a quite heavy and sophisticated process, involving many steps, extending well into the spacecraft lifetime and probably beyond.

We keep in mind that significant work in the field of algorithm validation has already been carried out in the framework of ESA studies. These studies [RD3] have been extensively used for (i) building the SMOS retrieval software, which may now be used either as a breadboard or for benchmarking various breadboard modules, (ii) designing the ATBD for the prototype. Results of [RD3] have also been used to build the preliminary error budget in the ATBD.

Therefore it seems logical to begin the validation by **consolidating the results of previous studies for the benefit of the SM prototype**.

On the other hand, the present plan addresses the steps of the prototype's validation which extends until actual SMOS data become available, which will occur during the SM L2 product verification part of the commissioning phase. It is understood that the latest steps may serve as a **rehearsal** of the SM L2 product verification operations.

1.2 OUTLINE

The core of the validation plan consists of three parts or **phases**, which differ mainly by the data used as input.

- Phase 1 (section 2) is mainly devoted at assessing the retrieval validity domain on conditions similar to those of previous sensitivity studies, i.e. use fully simulated academic data[1] in view of carrying out a systematic exploration of observing conditions for **homogeneous** scenes (i.e. a single fraction). Note that this part allows, however, testing **non uniform scenes**, i.e. variable parameters (optical thickness, soil moisture).

[1] i.e., very schematic, simplified, not necessarily bearing a relation with real life data but elaborated to assess a given point, and explore the whole range of values for surface characteristics.

This is carried out on artificial scenes. What is not done in step 1 is addressing **inhomogeneous** scenes, i.e. several fractions. The main reason is practical: it will take further effort and thus time to generate inhomogeneous scenes in the proposed scheme.

> Phase 1 is meant to be carried out in the frame of the current contract, while parts 2 & 3 (phases 2 & 3) should belong to the implementation phase.

- Phase 2 (section 3) is devoted to assessing the SMOS performance through performing the retrieval on a large sample as provided by actual, **realistic** auxiliary data. The assessment will be carried out through statistical analysis of the results.
- Phase 3 (section 4) addresses the specific **test sites** which will be selected for product verification and prepares this verification.
- In addition, section 5 addresses specific subsets of the retrieval algorithm such as updating the current maps. This section should collect pending validation issues.

Optimization of various operating parameters is foreseen in both phase 1 & 2. The assessment of retrieval efficiency is foreseen in phase 2.

Phase 2 & 3 are meant as **rehearsals** for the product verification phase. Inasmuch as they are properly designed, they should simply be repeated using actual SMOS data.

Most of phase 1 (Sections 2.1 and 2.2) corresponds loosely to the preliminary version of the draft.

- Finally, section 6 describes the implementation plan for each phase in terms of actions, schedule, allocation of tasks.
- Some choices in phase 1 need tuning. This task is to be scheduled in the plan.
- Many aspects of phases 2 and 3 are not fully defined. The open issues have to be addressed while carrying out phase 1. This task has to be scheduled in the plan.

2. PHASE 1: ASSESSMENT OF VALIDITY DOMAIN

2.1 RESTRICTIONS

- While the L2 algorithm is designed to deal with inhomogeneous scenes, we consider in this section homogeneous scenes, in the sense of aggregated (ALC) fractions. Hence no issue linked to the weighting function WEF will be addressed at this stage.

  This is a strong limitation, because in the real world a vast majority of SMOS scenes will consist in mixed scenes (most often including minor added contributions). Mixed pixel cases are considered in later validation steps (next section) prior to the launch.

  However, possible errors due to **non uniformities** (i.e. spread of the vegetation optical thickness TAU, which is the main parameter inducing non linear effects) should be investigated for homogeneous scenes.

  Homogeneous scenes for which SM is not retrieved are considered at the end of the present section.
- This stage of the validation will be restricted to **dual polarization** operation. This is meant to save time and is justified by the fact that from previous studies no major difference concerning the validity domain was detected among polarization modes. However, there is obviously no difficulty to considering both polarization modes, at least once they are available.

2.2 NOMINAL SCENES

2.2.1 OBJECTIVES

2.2.1.1 ASSESSING SCIENCE PERFORMANCES

The science performances should be assessed with respect to the mission requirements, i.e. the capability to achieve a 3-day global coverage with an uncertainty over retrieved soil moisture (SM) better than 4%, over homogeneous flat scenes, for the whole range of physically realistic soil moisture values, and over a range of the vegetation optical thickness TAU extending up to a limit which has been estimated close to 0.6.

This task will require building **synthetic performance indexes** taking into consideration both the (SM, TAU) domain and the observing conditions, which mainly vary according to the abscissa X_SWATH across the instrument field of view (FOV), inasmuch as this latter factor impacts on the actual revisit period through its influence on the actual swath over which the retrievals are compliant.

2.2.1.2 ASSESSING TECHNICAL PERFORMANCE

It is necessary to assess the ability of the iterative retrieval to provide converging solutions and to provide them at the expense of an acceptable number of iteration steps. However, this will be done in a further step using more realistic simulated scenes with all kind of noise added. In the present section, the only option left is to add noise to initial values for free or almost free parameters

2.2.1.3 OPTIMIZATION

Since mixed scenes are not considered, the present validation step practically shortcuts the **first** stage of the algorithm decision tree (DT).

On the other hand, it should validate the **second DT stage** (DT2) through ensuring consistency between various processing options described by this stage. This may imply optimization with respect to the pre-selected thresholds which drive the retrieval options.

2.2.2 METHODOLOGY

2.2.2.1 RETRIEVAL SCHEMES

The expected science performances have been estimated assuming a **dual step retrieval** (i.e. using, as a priori value and uncertainty for TAU, the results retrieved from previous SMOS visits). Although it is by no means certain that this procedure will be applied in the initial stages of the mission, the science validation will thus have to consider the dual step scheme, **along** with the single step one[2].

2.2.2.2 RETRIEVAL UNCERTAINTY VERSUS RETRIEVAL BIAS

Same as the ESA soil moisture retrieval study [RD3], the principle is a **sensitivity analysis**, the raw output of which is the result of the retrieval for SM, characterized either by the theoretical output uncertainty (DQX) propagated from input uncertainties, or by an output bias propagated from an input bias.

The DQX method is to be preferred because it offers several advantages.

If one retrieves soil moisture using a bias method i.e. forcing a wrong value on e.g. the proportion of sand, one will get indeed a bias which expresses the sensitivity. But if one retrieves soil moisture assuming for the sand proportion the right value with an input error equal to the difference (wrong value-correct value) in the above trial, one will get a contribution to the SM DQX which is equal to the bias found formerly.

The DQX takes full advantage of the prototype capabilities. It provides directly what is needed (i..e. retrieval uncertainty). It can deal with several parameters at the same time. It might finally account for possible correlations due to the functional model (inasmuch as it is conceivable that the variances do not add exactly).

[2] Opinions may differ there and we expect to get some feed back from ESA on this point.

However, one must take into consideration the flexibility in the prototype in which, while some parameters may be left free for retrieval, some others are assumed fixed for every retrieval mode.

Therefore there will be 2 ways to carry out this validation step.

- For parameters which the prototype allows to leave free, use will be made of the prototype and of the output DQX (**DQX runs**).
- For other cases, there are again two possibilities. Either one can modify the BB in order to allow floating and obtain DQX figures, or one can classically carry out a retrieval where the input value of the concerned parameter is modified, in which case the useful result is a retrieval bias.

  The second possibility (**bias runs**) is preferred, because it is simpler and avoids a deviation between customized BB versions and prototype. In this case, however, there are shortcomings because adding the (squared) obtained retrieval bias will not account for correlations between variables and thus induce a pessimistic estimate. This effect is minimized by carrying out the retrieval with respect to other parameters as stipulated by the decision tree stage 2.

  This second case will concern, in this validation step, the TEC, atmospheric attenuation, soil structure parameters (sand and clay fractions), sky contribution, possibly the rotation angle from surface to antenna level.

  Finally, the impact of strong lack of uniformity of TAU or SM within the scene can only be tested through the bias approach.

2.2.3 BUILDING SIMULATED DATA

2.2.3.1 GENERATING SIMULATED DATASETS

For phase 1, in order to obtain the test data as quickly as possible, the following scheme is suggested:

- The simulated SMOS data are supplied to the L2 prototype in the L1c format using the L1PP test scenario for a half orbit (L1_TS).
- The actual antenna TB are provided by ESLs using an ad hoc simulator which uses the breadboards for computing upwelling brightness temperatures, adds the rotations to obtain antenna level TB, and computes radiometric uncertainties DTB.
- To this end, it is necessary that ESL have a list of the DGG node present in the L1_TS, (grid_point_ID) together with, for each node, the list of incidence, rotation and Faraday angles.
- For computing DTB, applying the radiometric equation would require knowing the element directional antenna gain, which is not part of the L1c record but can be modeled simply. It is proposed to re-compute the DTB in the simulated runs; to this end, the **direction cosines** (or at least the angle with respect to bore-sight) must be supplied for each DGG node.
- Additionally, it is necessary that for each DGG the **value of the X_SWATH** is supplied also to ESL. No high accuracy is required on this figure (a few km), and therefore there should be a simple way to estimate it from the distance from the DGG to the track assumed to be a straight line between extreme satellite locations.
- Finally, a convenient **format** should be given to ESL in order to provide the simulated antenna frame TB and radiometric accuracies.

In summary, generating the necessary datasets for phase 1 essentially requires that a L1 test scenario L1_TS be available with reliable geometry data (angles, locations); this is considered to be the case. The L1_TS geometry data must be completed by information concerning X_SWATH and director cosines for the DGG nodes. Then, simulated values for TB and DTB in the antenna frame of reference will be computed and patched in the L1c test product. The resulting modified L1_TS will be used as inputs for phase 1.

It is proposed to include experiments in order to assess the effect of biases on parameters which are supplied by the L1c product, that is rotation and Faraday angles. To that purpose, it is necessary in addition that the actual values for these angles be extracted, biased, and patched again in the L1c product.

2.2.3.2 DESCRIPTION OF NECESSARY DATASETS

From the list of DGGs, the ESL will anyway edit a **restricted list of DGG nodes** which are spaced over the Earth in such a way that homogeneous working areas can be built independently over Earth surface in the auxiliary data for the retrieval. This is mandatory in order **to build a consistent auxiliary LAI map for the L2 retrieval**.

Assuming the X_SWATHs are provided together with the DGG coordinates, the ESL **will optimize the organization of input parameters over the ISEA grid for the simulation**. It is stressed that for assessing the validity domain, what matters is the ability to carry out a representative average along the swath, rather than using exactly defined and regularly spaced values of sampled X_SWATH.

According to the Product Format L1 document ([RD5], § 5.5.1), there are **78650** DGGs in a half-orbit L1c product. Since the count given in the L1 product includes a conservative 800 km width across the alias-free FOV, this figure is probably somewhat on the low side. If the spacing is extended from the 15 km ISEA spacing to 123 km for independent working areas, the remaining number is about **1100**.

Table 1 lists the cases for which input data sets are required for both DQX and bias runs: the total number is estimated to be close to 1600.

Loops (pointed by a short notation in the "loop" column) are carried out on SM, TAU and X_SWATH; in order to limit the number of cases (although the resulting economy is modest), the step of the loops is chosen looser for bias runs focused on other parameters. The "**x/+**" column indicates whether the number of cases in the N column should be summed or multiplied in order to obtain the total figures. The last r.h.s. column includes short notations for the datasets.

The sampling loops in Table 1 have been tailored down in order to obtain a situation where **a single L1c product** is needed for the DQX simulation. This situation is indeed highly convenient for the retrieval runs. This estimation assumes that most filtered DGG nodes are relevant from the point of view of the X_SWATH sampling. While a loop XL is indicated in Table 1, it will certainly have to be tuned further. This task has to be scheduled in the plan.

Table 1: simulations for validity domain exploration, using nominal model

Direct simulation		Lo-op	Selection	x / +	N	data set
Data options for DQX runs					**792**	TBL1
Surface scene						
	Scene		homogeneous nominal ; suggested code 166			
	WEF		none necessary			
Input parameters for simulated data						
	SM	SL	0.0 to 0.5 step 0.1	x	6	
	TAU	TL1	0, 0.01, 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8	x	12	
	Sand, clay		0.4, 0.3			
	Tsurf		288			
	Tsky		3.7			
	T0 atm, P0, WVC		288, 1013, 3			
	ro_s, alpha, SAL, eps_pa		From UPF LUT			
	HR, NR_H, etc.		from LANDCOVER_CLASS LUT for selected code			
Input observation conditions						
	Radiometric uncertainty		simulated based on L1c test scenario			
	X_SWATH (km)	XL	0, 100, 200, 300, 350, 400, 425, 450, 475, 500, 525	x	11	
	Incidence & rotation angles		from L1c test scenario			
	Faraday angles		from L1c test scenario			
	L1c flags		from L1c test scenario			
Additional data options for bias runs					**330**	
	Radiometric uncertainty		Zero			
	non uniform TAU		Mixing 2 direct models with TAU=0 and 0.6	+	1	TBL3
	non uniform SM	TL2	Mixing 2 direct models with 2 SM TBD around loop value, for TAU=0 to 0.6 step 0.2	+	4	TBL4
	SM	SL	0.0 to 0.5 step 0.1	x	6	
	X_SWATH	XL	0, 100, 200, 300, 350, 400, 425, 450, 475, 500, 525	x	11	
Total					**1122**	

The last column of the table identifies 3 independent datasets (TBL) . In addition, some bias runs will require a subset TBL2 of the TBL1 dataset (aimed at DQX retrieval runs). Therefore the simulated data should be, for convenience, organized into **4 distinct L1c products**.

Only the restricted list of DGG nodes (in the sense of the 123 km spacing) will be useful, further restricted to each relevant dataset for TBL2 to TBL6. The best would be to keep only this data in each L1c product. In case this is complicated, at least the other DGG should be excluded from the retrieval. A simple way to achieve this is to assign, to those spurious DGG, TB values outside the accepted range stipulated by the UPF file.

2.2.3.3 AUXILIARY DATA AND PARTICULAR CASES

Concerning the auxiliary data for the TB simulator, the fastest way is to **assign them directly** rather than using auxiliary files. Care will have, however, to be exercised in order to achieve consistency with auxiliary data provided to the L2 prototype for the retrieval, which will have by definition to be supplied as LUT and auxiliary files.

- Input parameters for surface (excluding SM and TAU) are selected with representative values. For the parameters normally available from the land cover class LUT, values corresponding to a specific code will be chosen. A plausible choice is ECOCLIMAP code 166 ("temperate crops").

- Input parameters for external factors (atmospheric quantities) are selected using representative values according to climatology.
- It is proposed to assume the sky contribution is **constant** in the direct simulation.
- The non uniformity tests for TAU and SM are not fully defined and this requires finalizing with RU. This task has to be scheduled in the plan. Note the numbers of cases are comparatively small wrt the overall number of simulations. For these tests, the simplest is to compute the upwelling TB by averaging the results of two direct models.

Other specific issues are discussed below in the description of **retrieval options**.

2.2.4 RETRIEVAL TESTS

2.2.4.1 VERIFICATIONS OF THRESHOLDS

Prior to the runs, a couple of basic verifications must be made. They begin through comparing the MVAL index (see ATBD) to the TH_MMIN* thresholds[3] and inspecting its variation with the X_SWATH

The threshold TH_MIN1, which helps to eliminate the DGG nodes for which the available number of views is too small, must be checked. The requirements for this threshold are that it should filter out data for which the number of views is smaller than 8 (this is the maximum number of floating parameters here) and that it should then retain about the maximum value M_SWATH allowed for the swath extent (somewhere around 500 km).

The quantity which expresses the weighted number of available views mainly depends on the location across the swath. Other factors are artifacts (L1c flags, RFI, outliers), but here we ignore them as a 1rst step.

Thus, if we specify TH_MIN1 in absence of artifacts, we get a M_SWATH value. If M_SWATH was smaller than about 470 km, we would have a problem, but this is not likely to happen. If M_SWATH is larger, a choice must be made: either keep a large M_SWATH, or increase TH_MIN in order to decrease M_SWATH. Probably the optimal choice lies in between, so as to make allowance for RFI and the like.

The impact of threshold TH_MIN3, beyond which the most demanding retrieval option is chosen (R4, see below), must be fully assessed by tabulating the concerned swath extent. The value should be tuned in such a way that this extent is a narrow zone near the satellite track. The exact value will be chosen after validation. Further requirements for TH_MIN3 are discussed below.

The impact of the thresholds depends on the set of radiometric uncertainties, which in turn depend on (i) the views for which this uncertainty is augmented on the basis of L1 flags, and (ii) the scene temperature T_SCENE[4]. Some allowance will be made for (i) (see retrieval options); concerning (ii), a T_SCENE value on the high side, e.g. 220 K, will be selected for this validation step in the direct simulation (to be discussed again for next section).

2.2.4.2 STAGE 2 RETRIEVAL OPTIONS

These options are given in the user's parameter file UPF, tables #40 (thresholds) and #41 (ASTD), reproduced below for the nominal scene (Table 2), as well as in the LAND_COVER_CLASSES LUT.

- For nominal scenes, the DT stage 2 lists 3 retrieval strategies Ret_2, Ret_3, Ret_4, depending on MVAL as compared to 3 thresholds TH_MMIN1, 2, 3. The MVAL parameter describes very roughly the information content of BT data; when L1c data are not eliminated due to artifacts, it is closely related to the location in the swath as given by the X_SWATH abscissa.
- For each of these strategies, the DT stage 2 again distinguishes 3 cases (TAU_small, TAU_medium, TAU_high) depending on the range of the initial/expected optical thickness TAU as compared to 2 thresholds TH_23 and TH_34.

[3] There is a difficulty here, since the **MVAL index is not reported** in the DAP. It must be computed in the simulation.

[4] The formula for radiometric uncertainty includes a factor (Tnoise+Tscene). Tnoise is estimated to 215 K. Tscene (also called Tantenna) depends on the scene! It cannot be higher that around 220 K (physical blackbody for high optical thickness, minus ¼ because ¼ of the field of view sees the cold sky). Over land, it can be lower for high SM and low TAU, down to a minimum of about 150 K.

Then, in each of the 9 cases, the nature of floating parameters and their a priori uncertainties ASTD are defined.

Of course, the retrieval options are selected in such a way that the retrieval makes the best of the available information, with simplifying assumptions made accordingly. However, it must be borne in mind that even though in a given retrieval some parameters are assumed perfectly known, this is not the case in the real world; uncertainties on frozen parameters are still there!

Therefore one is led to the conclusion that, while every retrieval option according to DT2 must be simulated, the relevant benchmark is a **retrieval scheme Ret_5 where every uncertainty** is kept[5] (see r.h.s. column). This will be implemented easily through modifying the ASTD in the UPF.

Table 2: ASTD for nominal retrieval (modified TGRD table #41)

Data richness c_V:	MVAL		>V2	2	>V3	3	>V4	4	>V5	5
Vj foreseen values			3?		30		45		80	
Thickness c_T:	TAU			1	>T23	2	>T34	3		
Tij foreseen values					0.05		0.5			

Parameter	Unit												
VT=10*c_V + c_T		**21**	**22**	**23**	**31**	**32**	**33**	**41**	**42**	**43**	**51**	**52**	**53**
σ0_A_CARD	-	80	80	80	80	80	80	80	80	80	80	80	80
σ0_SM	%	80	80	80	80	80	80	80	80	80	80	80	80
σ0_TAU	-	0.2	0.5	0.5	0.2	0.5	0.5	0.2	0.5	0.5	0.2	0.5	0.5
σ0_TPHYS	K										2.5	2.5	2.5
σ0_TTH	-										1	1	1
σ0_DIFF_O	-										0.1	0.1	0.1
σ0_OMH	-								0.1	0.1	0.1	0.1	0.1
σ0_HR	-				0.5			0.5		0.1	0.1	0.1	0.1
σ0_RTT	-									2	2	2	2

In this table the TH_MMIN are labeled Vj (genering an index c_V) and the TAU thresholds are labeled Tij (genering an index c_T). The combined index is labeled VT.

In order to check the consistency of DT2 retrieval options, the retrieval runs to be listed below should include each of them for relevant ranges of the optical thickness.

2.2.4.3 RETRIEVAL DATA SETS

The retrieval tests will be carried out over the ranges of

- SM realistic values (from 0 to 0.5),
- TAU realistic values for the SMOS requirement (0 to 0.6) extended by some margin in order to reach the limit of the validity domain.

In addition, the X_SWATH range will be explored up to the value defined from the TH_MIN1 verification (see above).

These requirements are accounted for when building the simulated datasets. Further loops are introduced in order to tune the DT stage 2 and assess other retrieval options.

[5] In the retrievals Ret_2 to Ret_4 we assume that some parameters are frozen. This helps to carry out the retrieval. Some uncertainty exists on the frozen parameters all the same.

In the present step of the validation exercise we should not have convergence problems since there is no noise. Hence we can keep all possible parameters floating so as to come closer to realistic output uncertainties.

Table 3 (a) shows the number and characteristics of retrieval runs; they are larger than the number of simulated datasets by a factor of about 9. This difference is explained mainly by the dual-step option and by the need to test several options of the DT2.

Each of **the 30 lines** of the table calls for a specific **configuration** of the L2 prototype, i.e. a specific combination of L1c inputs, UPF and other auxiliary data. The last 2 lines aim at testing the retrieval after the DT2 thresholds have been tuned.

Table 3 a: simulated retrievals for validity domain exploration, using nominal model

	Retrieval runs	Nb	TB set	c_V (=RET_n)	c_T (=T ij)	ASTD but TAU	STEP MODE	TAU ASTD	TAU	SM	X_SWATH	RO°	FA°	Tsurf	T0, P0, WVC	SAND CLAY	SKY
	Origin		L1 _TS	UPF	UPF	UPF	CURt & UPF	CURt or UPF	LAI map	SM _ini	L2	L1 _TS	L1 _TS	EC map	EC maps	S+C map	Sky map
1	DQX	792	TBL1	2	1 & 3	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
2	DQX	792	TBL1	3	1 & 3	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
3	DQX	792	TBL1	4	1 & 3	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
4	DQX	792	TBL1	5	1 & 3	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
5	DQX	792	TBL1	2	2	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
6	DQX	792	TBL1	3	2	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
7	DQX	792	TBL1	4	2	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
8	DQX	792	TBL1	5	2	UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
9	non uniform TAU	66	TBL3	3	2	UPF	S	UPF	0.3	SL	XL	M11	M21	M31	M41	M51	M61
10	non uniform SM	264	TBL4	3	2	UPF	S	UPF	TL2	SL	XL	M11	M21	M31	M41	M51	M61
11	Rotation bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M12	M21	M31	M41	M51	M61
12	Faraday bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M11	M22	M31	M41	M51	M61
13	S+C bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M11	M21	M32	M41	M51	M61
14	Sky bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M11	M21	M31	M42	M51	M61
15	Atmo bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M11	M21	M31	M41	M52	M61
16	Tsurf bias	264	TBL2	3	2	UPF	S	UPF	TL2	SL	XL	M11	M21	M31	M41	M51	M62
17	DQX	792	TBL1	2	2	UPF	D	CUR	TL1	SL	XL	M11	M21	M31	M41	M51	M61
18	DQX	792	TBL1	3	2	UPF	D	CUR	TL1	SL	XL	M11	M21	M31	M41	M51	M61
19	DQX	792	TBL1	4	2	UPF	D	CUR	TL1	SL	XL	M11	M21	M31	M41	M51	M61
20	DQX	792	TBL1	5	2	UPF	D	CUR	TL1	SL	XL	M11	M21	M31	M41	M51	M61
21	non uniform TAU	66	TBL3	3	2	UPF	D	CUR	0.3	SL	XL	M11	M21	M31	M41	M51	M61
22	non uniform SM	264	TBL4	3	2	UPF	D	CUR	TL2	SL	XL	M11	M21	M31	M41	M51	M61
23	Rotation bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M12	M21	M31	M41	M51	M61
24	Faraday bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M11	M22	M31	M41	M51	M61
25	S+C bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M11	M21	M32	M41	M51	M61
26	Sky bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M11	M21	M31	M42	M51	M61
27	Atmo bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M11	M21	M31	M41	M52	M61
28	Tsurf bias	264	TBL2	3	2	UPF	D	CUR	TL2	SL	XL	M11	M21	M31	M41	M51	M62
29	DQX	792	TBL1	Tuned		UPF	S	UPF	TL1	SL	XL	M11	M21	M31	M41	M51	M61
30	DQX	792	TBL1	Tuned		UPF	D	CUR	TL1	SL	XL	M11	M21	M31	M41	M51	M61
		14 916															

Table 3 b defines the notations which refer to simulated Tb datasets and necessary auxiliary data.

Once the restricted list of DGG nodes has been defined, a table results which associates a specific node to every item in the (SM, TAU, X_SWATH) triple loop. The information necessary for building the LAI map is available from the TAU values in this table.

Table 3 b: simulated retrievals for validity domain exploration, using nominal model

TB sets		TBL1	uniform no bias			TBL2	subset of TBL1		TBL3, 4	autonomous
UPF for Ret_2						TH_MMIN1=3			TH_MMIN2=100	TH_MMIN3=100
UPF for Ret_3						TH_MMIN1=3			TH_MMIN2=3	TH_MMIN3=100
UPF for Ret_4						TH_MMIN1=3			TH_MMIN2=3	TH_MMIN3=3
UPF for Ret_5		shift Ret_5 column in Ret_2				TH_MMIN1=3			TH_MMIN2=100	TH_MMIN3=100
UPF for TAU 1 & 3		(short + high) TAU				TH_23=0.3			TH_34=0.3	
UPF for TAU 2		(medium) TAU				TH_23=0			TH_34=0.8	
UPF tuned						TH_MMIN1=3?			TH_MMIN2=30 ?	TH_MMIN3=50 ?
UPF tuned						TH_23=0.05 ?			TH_34=0.5 ?	
SM map		uniform SM=0.15			TBC					
LAI map		custom built to accommodate X_SWATH, TAU & SM loops on restricted DGG list								
USE_DEFLT_TAU_NAD_LV		single step:		1		dual step:		0		
CURRENT TAU		single step:		void		dual step:		fill from single step results		
		Unbiased				**Biased**				
Rotation angle		M11	from L1_TS			M12	from L1_TS plus 2° uniform			(patched in L1_TS)
Faraday angle		M21	from L1_TS			M22	from L1_TS multiplied by 2			(patched in L1_TS)
Tsurf		M31	uniform 288			M32	uniform 290			
T0, P0, WVC		M41	uniform 288, 1013, 3			M42	uniform 290, 1030, 6			
Sand + Clay		M51	uniform 0.4, 0.3			M52	uniform 0.5, 0.2			
Sky map		M61	uniform 3.7			M62	uniform 13.7			

Notations M11 to M61 and M12 to M62 stand respectively for unbiased and biased auxiliary data files.

2.2.4.4 DQX RUNS

DQX runs are numbered 9-16 and 21-28 in Table 3a.

- Retrievals defined as R2, R3, R4 are systematically carried out together with the **R5 scheme** defined above, for the TAU_medium case, according to UPF table #41. This will require each time to control and adjust the thresholds in UPF and the values in the file which correspond to table 41.
- This sequence is repeated for the TAU_small and TAU_high options, over ranges which cover the lower and higher parts of the TAU loop.
- The LAI map, used to assign in the L2 the initial TAU values, has to be built around the restricted list of DGG nodes over the relevant working areas.
- The soil moisture map is chosen as a single, uniform value. This is simpler and allows some iterations and thus a minimum testing of the way the retrieval is operating (except when the SM actually used in the simulation happens to be close to the selected initial guess). Since the soil moisture is a quasi free parameter, the retrieval result is not expected to depend on the initial value. The SM ASTD in table 2 has been increased to this end.

  On the other hand, reference values for the constrained parameters should be the exact ones, in order to obtain the theoretical retrieval standard deviation where it is wished to obtain them.

- The retrievals are be carried out for the single step and **repeated** for the dual step retrieval scenario, only for the TAU_medium option. In the latter case, it is proposed that the input ASTD for TAU at a given X_SWATH will be chosen from the single step retrieval results for abscissa (M_SWATH – X_SWATH). The M_SWATH value will be tuned depending on the TH_MMIN1 verification result. This algorithm will have to be refined (this task must be scheduled in the plan), accounting for the detailed structure of the X_SWATH sampling list.

  Finally, a specific routine is needed to organize the corresponding ASTD values in the CURRENT_TAU file in order to feed them to the L2 prototype.

2.2.4.5 BIAS RUNS

Bias runs are numbered 1-8 and 17-20 in Table 3a.

According to the previous discussion, while no noise is applied to simulated TB data, the retrieval is carried out using the same ASTD as in the DQX runs.

Then

- Either the input BTs are distorted (case of TAU or SM non uniformity, rotation angle, Faraday angle),
- Or (frozen) input parameter values for the retrieval are biased (case of surface temperature, SAND/CLAY, atmospheric values, sky) through introducing a bias in the relevant uniform auxiliary data maps.

The biases are retrieved for input errors adjusted to the estimated typical uncertainties, on the large side.

Concerning auxiliary files for bias runs, M11 (M12) and M21 (M22) are particular cases because they are provided trough the L1c product.

- The rotation angle is provided to L2 by the L1_TS for the actual geometry. A way to introduce the rotation angle bias consists in extracting the values from the L1_TS, modifying them and patching back the modified values in the relevant L1c simulated product.
- The Faraday angle is provided to L2 by the L1_TS for the actual TEC map. The proposed biasing method is similar.

2.2.5 USING THE RETRIEVAL RUNS

2.2.5.1 OUTPUTS

The analysis should be based exclusively on information stored in the UDP and DAP products, plus information about the test conditions themselves, concerning basically the data stipulated in Table 1 and Table 3.

2.2.5.2 TUNING THE TH_MMIN PARAMETERS

This can be done without accounting for error sources which are tested as bias generators, on the basis of retrieval uncertainties DQX only, because these biases (expected to be small anyway) should not depend on the exact retrieval scheme. When DQX is mentioned below with no specification, it refers to soil moisture.

- As X_SWATH increases away from track, for a given retrieval scheme (assuming the track abscissa is processed using R4), the DQX will increase. Crossing in turn thresholds TH_MMIN3 and TH_MMIN2 will switch to poorer retrieval schemes. Then the consistency constraint which will help to tune TH_MMIN3 and TH_MMIN2 is that the DQX should not decrease *when crossing thresholds*, as this would be unrealistic. It should only increase more slowly.

  Full information necessary for testing this is available since the retrievals R2, R3 and R4 are carried out for every case, both in dual and single step modes.
- Concerning TH_MMIN3 and the R4 scheme close to the track, a further physical constraint is that, for every retrieved parameter, the ex post DQX should be smaller than the input value. In other words, the retrieval should supply at least some information about the concerned parameters. If this is not the case for a given parameter, letting it float is useless, and the selection of free parameters for R4 should be modified.

This analysis does not apply to R5 results, which supply the benchmark in terms of fully realistic retrieval errors.

2.2.5.3 TUNING THE TAU THRESHOLDS

- The **TAU_small** threshold is anyway small, since in the corresponding class the surface is assumed bare. The value should be tuned close to the minimum of the retrieval uncertainty (DQX for TAU) for small X_SWATH values, for bare soil assumed in the direct simulation. In this way, the possible error on SM resulting from a non zero TAU value will stay within the limit of the retrieval performance.

  If the value to be selected varies somewhat with SM, an intermediate value (SM = 0.2 possibly) will be selected for tuning.

- The **TAU_high** threshold corresponds to situations close to the limit of the validity domain. Hence the question arises of a trade-off where the emphasis would rather lie with getting best estimates of TAU, at the expense of narrowing the ASTD for SM and therefore undergoing some risk of bias on the retrieved SM value.

  Setting the optimal value for TAU_high would require an assessment of the actual confidence which can be set in a priori (i.e. forecast) TAU estimates for large TAU values. Another complication is due to the fact that the limit of the validity domain actually depends on SM.

  As a first step, the TAU_high threshold will stay around 0.5, which corresponds to the limit of the validity domain for dual step processing and SM around 30%

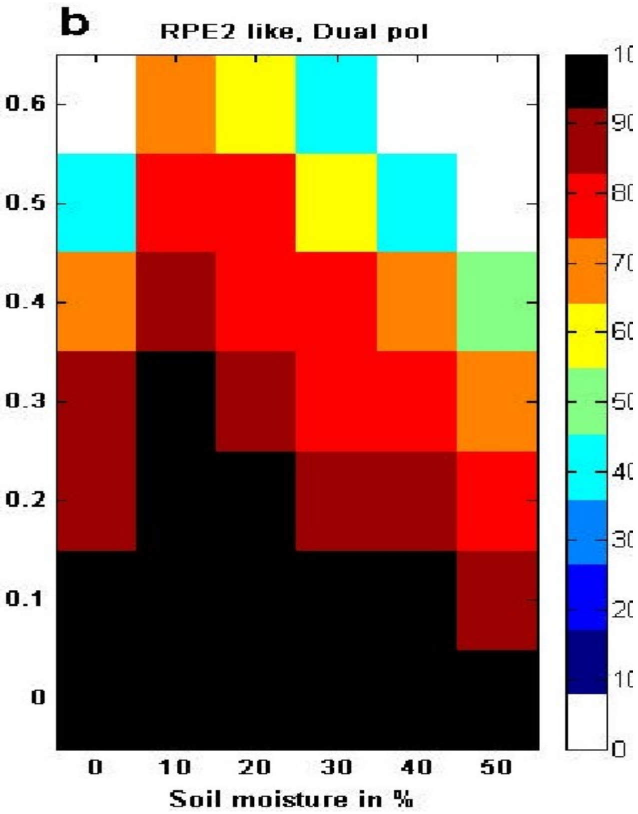
2.2.5.4 PRINCIPLE OF SCIENCE PERFORMANCE FIGURES

As explained above, the main requirement for SMOS is expressed by a validity domain for a 4% uncertainty in the retrieved SM, over the SWATH required to ensure a 3 days repeat time.

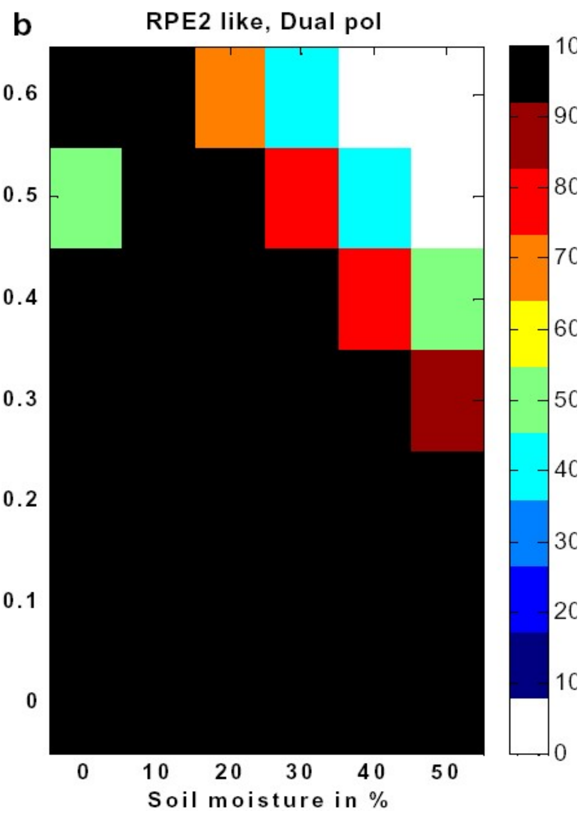
From a large number of preliminary studies it is expected that, provided that input errors are realistic, this main requirement, defined as an average over the SMOS swath, will be met up to a certain value of optical thickness, which itself depends on soil moisture. The way to check the requirement will consist of **adding quadratically all biases and retrieval uncertainties**.

The figures below illustrate these points. They represent in the (SM, TAU) plane the **percentage** of the swath (color scale) over which the **swath averaged** retrieval uncertainty is smaller than a given threshold (here 1%). On the l.h.s. TAU is left fully free; on the r.h.s. the prior TAU uncertainty is taken from the previous SMOS visit. These diagrams illustrate the results of prior ESA studies [RD3]; they are built for homogeneous scenes and account only for the radiometric uncertainty, hence they are on the optimistic side.

The area enclosed by a threshold (normally 100%) will provide a possible overall performance figure.



SM retrieval performance with free TAU



SM retrieval performance with constrained TAU

2.2.5.5 IMPLEMENTATION OF SCIENCE PERFORMANCE FIGURES

- For each cell on the (SM, TAU) plane: the **elementary** proposed figure is the proportion of the SWATH for which the SM requirement is met.

  The swath limit to be used will be taken either as the value resulting from the TH_MMIN1 threshold, or as 475 km which ensures global covering in 3 days.

  This is to be computed for dual and single step modes, for

  - the R5 retrieval
  - the planned combination of R2 to R4 options.

  The contribution of biases should be added quadratically to the SM DQX.

  The figure is to be computed either considering separately every result across the swath, or the average.

- Over the (SM, TAU) plane, **integrated performance figures** will finally be computed through **averaging**.

  This average should ideally be weighted depending on the actual climatological occurrence of SM and TAU values. In absence of accurate knowledge, no (or very raw) weighting will be applied: the SM range covers every realistic value, and the TAU range is limited to various values up to 0.6. Other partitions of the plane might be considered; listing them is a task to be scheduled in the plan.

2.3 HOMOGENEOUS NON NOMINAL SCENES

Similar approaches as described in 2.2 for the nominal scene FNO may be applied to building performance figures.

- FFO (forest): The FNO results can directly be applied, changing however the range for TAU. In this case, a map should also be drawn for TAU retrieval performances.
- FWL (wetlands), the validity domain must be redefined in (TAU, Tsur) coordinates.

- FWO (open water): since there is no TAU; the integrated science figure is not relevant.
- Other ALC classes (FEB, FEI, FEU), the validity domain must be redefined in (TAU, Card_A) coordinates, where Card_A is the dielectric constant index obtained from the cardioid model.

The DT2 tuning operations are anyway only carried out for FNO.

These tests are not considered mandatory for the phase 1 of validation.

On the other hand, at that stage we have not described tests allowing validating the "**secondary" cardioid retrievals**. A task for adding this must be scheduled in the plan. The formalism for nominal retrieval can be adapted easily.

2.4 SUCCESS CRITERIA

Criteria for the success of phase 1 should be appreciated with respect to the objectives of this phase. They are qualitative and must be applied to the 2 parts of the whole exercise (see respectively § 2.2.5.2 & 3 and § 2.2.5.4 & 5). Phase 1 should be considered as successful inasmuch as:

- The tuning of DT stage 2 parameters can be achieved consistently;
- The validity domain on the (SM, TAU) space is found consistent with previous studies.

3. PHASE 2: GLOBAL STATISTICAL VALIDATION

3.1 METHOLOGY AND OBJECTIVES

While exploring the validity domain in fully controlled conditions in step 1 above is useful to ensure continuity with previous performance simulated assessment, the next step of the plan considers **realistic** simulated data.

Indeed, when considering the investigation of performances for mixed pixels, it would make no sense to explore every combination of fractions whereas in the real world the sampling is much more reduced and some limited combinations are immensely more meaningful than others.

A very large data base is simulated. This is necessary in order to cover fully the experimental cases. A major resulting advantage is that one need not rely upon the theoretical (DQX) retrieval standard deviations: adding noise to the radiometric data and adding errors to the reference values is relevant because, since the population is large, statistics over actual retrieval errors will allow an estimate of the overall retrieval uncertainty for each class (to be defined) of SMOS scenes.

Moreover, this realistic simulation of input errors will provide a realistic estimation of the technical performances of the algorithm.

This is a quite heavy task because it comes close to what is termed in the SOW "end to end" test, including scene in-homogeneities. The difficulty will be understood by sketching the simulation of input data.

3.2 BUILDING SIMULATED DATA

3.2.1 SIMULATOR

The option proposed for this phase is to follow systematically, for all steps of the validation, the "end to end" scheme. Although it seems ambitious, it may also be the most efficient.

In this scheme:

- The L2 prototype is always used (rather than L2 breadboards), as well as fully fledged auxiliary data files.
- The L2 prototype is fed by outputs from the L1 prototype formatted as L1c products.

- The L1 prototype is fed as L0 packets by a modified SEPS. Major necessary modifications to the current SEPS nominal version include:
  - Implementing direct radiative models in line with the models selected for retrieval in the L2 prototype;
  - Implementing an improved spatial resolution

3.2.2 ISSUES FOR SIMULATION

- We should have direct models accounting for both topography and mixed snow effects; the list of ALCs to be considered for simulation is FM0 rather than FM (see ATBD table #14). The model for topography exists but is resources greedy and questionable in some cases as models are not validated for high incidence angles. For melting snow, the retrieval will use the dielectric constant model. However, it is necessary to select an approach for the direct simulation.
- Simulation of Faraday or rotation angle effects in the ground to antenna level transfer is not mandatory, as it will have been tested in the previous step. Furthermore, specific L1 corrections are not considered.
- How are auxiliary data to be fed to the SEPS? The simple direct assignment proposed for phase 1 is not practical. Using data file formats **as close as possible** to those used for feeding the L2 prototype would certainly improve the robustness and practicality of the system. This calls for more analysis TBD.
- Since mixed scenes are considered, the MEAN_WEF and WEF must be used for computing the fractional content within each DFFG grid.

  It is likely, however, that, for computing power reasons, the resolution for DFFG will be chosen coarser than the nominal (4 km) one; possibly, a resolution close to the ECMWF step should be retained. This has the shortcoming of introducing **large quantification effects** on the WEF. No qualitative large distortion is however expected.

3.2.3 NOISE, FRACTION ERRORS, STEP MODE

- **Noise** must be added before retrieval to antenna level **BT** according to the computed radiometric uncertainties. Ideally, one would rather compute noise-free simulated TB (for archiving) and add random noise just before the retrieval; however, it is probably impractical TBD to patch the noise generation in between L1c outputs and L2. Therefore noise will probably have to be computed by the SEPS and incorporated to simulated TB
- Similarly, noise should be added to **every reference value**; the same question arises TBD, and probably the same answer will result.

  The magnitude of the errors for land surface parameters should be available from the LAND_COVER_CLASSES LUT.
- Finally, phase 2 includes testing a new source of error, i.e. errors in **fractions.** The same question arises again TBD.
- For dual-step validation purposes, it becomes necessary that the CURRENT_TAU map is updated together with running the L2 prototype.

Summarizing, it appears that the step mode remains the only option for retrieval, while all errors and uncertainties are added at the simulation stage. This requires further discussion keeping in mind the size of the data base.

3.3 SIZE OF THE SIMULATED DATA BASE

3.3.1 TIME DURATION

- The simulated data should first cover a time period large enough in order to both allow a variety of observing conditions for the DGG and allow the delay necessary for testing the dual step scenario in close to realistic conditions. To this end, the adequate overall time window should cover a **minimum** of 2 global coverages = 6 days.

The best time of year is near the solstice: in this way, there is hope of both local winter and summer features, with the presence of non permanent soil covers and highly variable vegetation covers on the nominal fractions.

Practically, the SMOS orbit propagation ought then to be created around June 21 or December 21st 2007. The necessary ECMWF data fields would be obtained from past analyses for the same period in 2006 or 2005. The optimal solution would probably be to **select 3 periods of 6 days** at the two solstices and one equinox.

- It is suggested to build a simulated data base covering the full extent of L1c products throughout this period. Indeed, since the tests and tuning operations will be of a statistical nature, a large data set is necessary. Moreover, obtaining data on a global scale removes any concern about rain episodes, since at global scale rainy episodes are bound to happen somewhere.

3.3.2 SORTING CONSIDERATIONS

- Actually, the total size should also be such that an **adequate** sample size is obtained for all the subsets for which statistics are needed.
- Table 4 is a draft for describing how it is wished to organize the retrieval runs into "homogeneous" subsets. Selecting options described by the above table should be made easy by the user interface discussed below.

Table 4: simulation subsets

Sorting	Select	Categories
Time windows		
	sample	Days ND1 to ND2
Space windows		
	sample or loops	Square in LAT LON
	sample or loops	Zonal bands and continent masks
	sample	NPE zone
Orbit		
	sample or loop	Morning, evening, or both
Polarization mode		
	sample or loop	Dual, full or both
SMOS land cover		
	loops	For FM0 ALC categories
	option	Account for proposed DT1 thresholds
	option	Consider topography
	option	Consider LC homogeneity
Reference parameters		
	loop	For MN: SM & TAU bins
	loop	For MD and MW: TAU bins
X_SWATH		
	loop	Min / step / max values
	option	One or both sided
Retrieval N steps		
	sample or loop	Dual, single, both
Retrieval model		
	select	MN, MW, MD, all
Time series		TBD

What is an adequate size for a subset? In order to obtain a reliable empirical estimate of the standard deviation of the retrieval errors, it is suggested that the population should include 1000 samples.

Some feeling about what is allowed by the data base is provided by Table 5, which was built using the data base described in (RD4]. This table is built over the global land cover map (930354 DGG nodes); the numbers in the N column stand for the total number of cases where the relevant fraction lies between the TH (threshold fraction) values on each side.

Table 5: population of some scene categories

	TH	N	TH	N	TH	N	TH	N	TH	N	TH	N	TH	N	TH
TH_WO	1	95 739	0.99	38 380	0.90										
TH_WO	1				0.90	38 735	0.60								
TH_FO	1	34 129	0.97	38 518	0.90	35 935	0.80	29 314	0.70	26 515	0.60				
TH_NO	1	163 439	0.99	46 057	0.97	26 907	0.95	43 188	0.90	57 162	0.80	77 933	0.60	67 770	0.40
TH_WL	1	298	0.90												
TH_EB	1	1 479	0.80												
TH_TI	1	71 289	0.99	2 046	0.90										
TH_UH	1	0	0.80												
TH_**		35 521	0.00												

Note that no account is taken of topography, snow and frost.

Furthermore, homogeneous classes should be restricted to limited ranges for SM, TAU and X_SWATH. The total number of sorting categories for these 3 parameters cannot be much smaller than 5 x 6 x 11 = 330.

It is readily seen that even including both morning and evening orbits, a 1000 population for each individual class will not be reached often.

Therefore a case is made here for **extending the simulation period to 5 full global coverages**, i.e. 15 days.

3.3.3 USER INTERFACE

This is simply an interface to select the sample. Although there is no inherent complexity, it must be thought out with care since many filtering choices should be possible, without making the interface too intricate. Elements for section include non exhaustively (see Table 4):

- A time bracket (this being many in view of the commissioning operations)
- A space bracket (regions of the planet, including for example regions where no non permanent cover is likely to occur or the reverse)
- Orbits: morning or evening or both
- Nature of land cover, to be sorted into a manageable number of categories
- Nature of the retrieval scenario
- Selection of specific errors introduced at the retrieval stage; an example concerns errors in fractions.

3.3.4 HOUSEKEEPING FILE

In a validation exercise, simulated data are first computed using "exact" input information (i.e. reference parameters); next, they are eventually modified though adding simulated errors.

A module is needed in order to keep track of both initial ("exact") and modified parameters. Care must be exercised as such modifications may have to be introduced at several locations in the processing.

The same module should keep track of retrieval options as stipulated by selected values in the user parameter file or in the LAND_COVER_CLASSES LUT. Some values may require revisiting (see comments by INRA in Annex);

Options selected in the user's interface (global validation) are to be copied in the housekeeping file.

3.4 USING RETRIEVAL RESULTS

3.4.1 OUTLINE

Table 6 is a draft for describing further sorting and suggested output to be used in validation and tuning operations.

Table 6: output statistics

Sorting on retrieval outcome		
		attempted R4, successful
		attempted R4, R3 successful
		attempted R4, R2 successful
		attempted R3, successful
		attempted R3, R2 successful
		attempted R2, successful
		Failure
Statistics on results		
	technical	statistics on iteration number
	fit quality	statistics on normalized CHI2
	"	statistics on normalized RMSE
	performance	stat error on retrieved param
	"	stat DQX
	"	stat GQX
	failures	stat ref values for free parameters
Comparisons		
	Orbits	morning evening
	pol mode	dual full
	retrieval scenario	single dual step

Note than for the retrieval there is no need to introduce a "reference" R5 retrieval or several R* cases: the L2 prototype is operating according to the DT stage 2, hopefully tuned with the help of the results of the previous validation phase.

3.4.2 TEST AND TUNE OPERATING PARAMETERS

See Table 6 above:

- Test number of iteration NIT,
- Test percentage of successful retrievals,
- Check success of various (R2, R3, R4) retrieval options.

Tune Levenberg Marquardt initial parameters and multiplying factors if necessary.

A preliminary run on a **limited subset** of the database should be used for this step (this is the reason why this part of validation is given first) in order to run the retrieval on the full sample only after tuning.

3.4.3 BUILD STATISTICS ON RETRIEVAL RESULTS

See Table 6 above.

This should be done on roughly sorted categories inside each branch of the decision tree; the classifications mentioned in Table 4 to Table 6 should be used again to sort the results.

It must be possible to select convenient time windows (anticipating the time windows available during commissioning), various space windows (regions on the planet), types of mixed pixels, and so on.

It is expected that **every ESL will contribute to**

- define aspects of the input data,

- as well as specific subpopulations, possibly using other criteria,
- and to define statistics to be built in order to understand and assess the retrieval performances.

Common sense choices are illustrated on Table 6:

- Test retrieved parameters against (noise free) auxiliary data, i.e. a priori guesses in order to obtain retrieval error populations and statistics. To this end, information concerning these a priori guesses, along with retrieval choices (the UPF) should be available together with retrieval results.
- Test empirical retrieval uncertainties thus obtained against theoretical estimates (DQX) and global GQX quality figures.
- Compare performances for morning against evening orbit, dual pol against full pol, dual step against single step scenario retrieval.

Overall, these statistics should provide a full assessment of the SMOS expected performance over mixed scenes.

3.4.4 TEST SOUNDNESS OF DT 1 AND TUNE THRESHOLDS

For each tree branch and subcategory, a few samples are drawn at random in each class in order to check that the DT is operating as foreseen TBD. This may not be necessary, as actually it is more of a verification issue.

However, what remains necessary is to define a way to investigate the impact of thresholds on retrieval performances in order to tune the thresholds. This can be achieved through careful sorting of the retrieval results.

3.4.5 TUNE OUTPUT QUALITY INFORMATION

- The overall quality figure GQX and its coefficients should be tuned so as to be in agreement with theoretical retrieval uncertainties.
- The fraction output flags should be tuned so as to be raised for values which are not arbitrary but such that the impact of the flagged error source becomes significant.

  It is expected that the large number of retrievals will make it possible to accomplish this tuning from statistical processing. The methodology is TBD; brute force means varying the threshold levels and deciding based upon the average retrieval bias. This tuning operation is very similar to the one concerning the DT1 threshold and thought must be given to the way of combining them.

3.5 SUCCESS CRITERIA

Criteria for the success of phase 2 should be appreciated with respect to the objectives of this phase. They are semi-quantitative and must be applied subsets of the retrievals in terms of the nature of the SMOS pixels.

The main reasons why the criteria are semi quantitative rather than quantitative are as follows:

- The mission requirements state that a 4% uncertainty on soil moisture should be reached over homogeneous low vegetation scenes. This will obviously be tested. However, it would not make sense to stipulate that the validation fails if the requirement is obtained say 99.5% of the time. Therefore a (quite high) threshold will have to be defined TBD.
- On the other hand, even if the requirement is met 100% of the time, this is not fully satisfactory because fully homogeneous nominal scenes are a very small percentage of the land surface [RD4]. Therefore it is mandatory to assess the performance over mixed pixels, probably with a few categories depending the fraction polluted by heterogeneous scenes. Again, thresholds need to be defined TBD.
- Finally, although neither estimations of the optical thickness nor retrieval of information over non nominal covers are driven by requirements, they concur significantly to the expect outcome of the mission and should be assessed. It is not clear whether success criteria need to be defined a priori in this case TBD.

4. PHASE 3: DEFINE AND RUN TEST CASES

4.1 SELECT TEST CASES

Test cases are defined here as identified DGG nodes where ground truth measurements will be available during the product verification phase, owing to the presence of operational validation sites.

Likely validation sites must be identified and so to speak "**labeled**". This is a common issue between product verification planning and AlgoVal.

On the AlgoValP side, it seems that the responsibility lies in estimating whether the validation sites encompass an adequate variety of physical situations, i.e. to decide which kind of site might be missing on the basis of answers to AO. At the same time, some redundancies will possibly be pointed out.

A minimum list of adequate sites (in terms of their cover characteristics) should be established as the result of combining statistical descriptions of the SMOS scenes and the DT stage 1. The initial step of statistical descriptions is provided by [RD4].

A **specific imaging tool** (see section 4.4 below) to assess site spatial coverage (as a representative network is needed to provide ground truth for SMOS measurements), consistency between LC and ALC classes and land cover reported from visual inspection or direct knowledge is needed. Using this tool, site candidates will be compared to the list established above.

Possibly (likely), labeling validation sites will require some extension of the sites and necessary negotiations to achieve such upgrading. The responsibility for this should lie on the product verification side[6].

4.2 DATA SIMULATION FOR TEST CASES

It may be assumed that the number of test sites will be circa 12 to 20, typically 15. For each test site, several test cases will be defined.

The basis for obtaining data might be seen as a very small subset of the data base build for the previous AlgoValP step. However, seasonal variations should be accounted for. Hence, whereas a single solstice 15-day period seems adequate for the work described in section 3, for test site the time windows may be longer (3 weeks ?) and should cover several such periods, if possible 3 (2 solstice and one equinox period).

Building the simulation dataset follows the path described above for global validation (§ 3.2 & 3.2.3). Of course it does not seem sensible to simulate the whole globe while only about 15 sites are of interest. Filtering appropriate orbit segments must be allowed by the user's interface.

Adjusting the periods in such a way that some rainy episodes are present needs to be discussed.

Test cases should include dual versus full polarization (option TBC), both orbits, single versus dual step in case of nominal retrieval. Over the periods selected for phase 2, an adequate variability of observing conditions will be obtained.

While for the global validation, a low resolution DFFG should probably be used for computing power reasons, the nominal resolution (about 4 km) ought to be used for the test sites. This requirement can certainly and should be added to the requirements below for the user interface in section 3.3.3.

4.3 TEST CASE FORMS

While it is not clear how to fill a test case template for AlgoValP steps 1 & 2, it seems well suited to test cases.

4.3.1 TEST NAME

In every case, the test is a retrieval simulation. Then the test might be designated by a series of numbers. For example: TC(ns, nd, no, np, nm),

Where ns refers to the site, nd to the day, no to orbit, np to polarization, nm to retrieval mode.

[6] This discussion will require validation by ESA

4.3.2 TEST PROCEDURE

Realistic simulated data are built for the specific site and simulated retrievals are carried out.

4.3.3 TEST CHARACTERISTICS AND INPUTS

The specifics of inputs and tests are listed in a next table to be built.

Simulated data are built for the specific site, considering the cover properties, various observation conditions (equivalent to various M_SWATH), several cases of seasons (for vegetation cover and possibly non permanent covers) and simulated soil moisture.

Realistic noise and uncertainties are added to simulated SMOS and auxiliary data

Referring to 4.3.1, the total number of cases when ns, nd, no, np, nm vary might be of the order of 15 x 30 x 2 x 2 x 2 = 3600.

4.3.4 TEST OBJECTIVES

The results are used to assess the actual retrieval error.

4.3.5 PERFORMANCE/EVALUATION CRITERIA

The SMOS requirements should be met provided the site belongs to the claimed validity domain, considering results of section 2 of AlgoValP.

- The retrieval errors should be consistent with theoretical estimated retrieval uncertainties.
- The overall quality figure should be consistent with both previous results.

These criteria are put forwards for the case where only representative input auxiliary data are available, and it is wished to validate the retrieval algorithm over test sites in a restricted sense. However, it might be desirable to investigate the adequacy of ground truth measurements for representing the surface conditions. Then, auxiliary data might be built using actual measurements.

A further possibility, not clearly foreseen, is the case where airborne TB data become available. Then, a procedure should be conceived in which these data are used to constrain the input dataset.

4.4 INTERACTIVE ANALYSIS TOOL (IAT)

4.4.1 PURPOSE

When considering the nodes of the SMOS grid (DGG) and their relationship to the relevant surrounding area for radiometric contributions, it is necessary to obtain a realistic assessment of the soundness of:

- land cover classification,
- aggregating options,
- main decision tree options,
- Representativity of test sites.

These are therefore oriented toward both algorithm validation and preparation of commissioning (inasmuch as the ultimate part of AlgoValP is understood as a rehearsal for L2 product verification, using same test cases which correspond to test sites).

4.4.2 SUMMARISED DESCRIPTION

4.4.2.1 INPUTS

1. The DGG grid (isea9).
2. The ECOCLIMAP codes on the discrete fine grid (DFG). The initial ECOCLIMAP will be enriched by incorporating a detailed map for open fresh water and topography indices.
3. The soil parameter map.
4. Addition to the ECOCLIMAP: non permanent cover features (snow and frost): TBD.
5. Table coding ECOCLIMAP land cover (LC) to aggregated land cover (ALC).

6. Land cover: a set of representative seasonal LAI and LAIMAX.
7. A "scenery navigator" (example is Google Earth).

4.4.2.2 REQUIRED FONCTIONALITIES

1. Possibility to compare directly those maps for visual inspection.
2. Navigate with LAT LON grid, zoom, and the like. Customize color code.
3. Possibility in some cases to view them simultaneously, in some cases to superimpose them.
4. Superimpose the DGG with associated contours (circle for MEAN_WEF, customized ellipses for WEF, customized contours for test site areas and features).

4.4.2.3 DEADLINES

Emergencies are:

1. Assessing soundness of decision tree in ATBD.
2. Assessing representativity of test sites considered for SM product verification.

The largest constraint comes from the second issue, inasmuch as it might appear necessary to enhance some measuring networks.

A preliminary version (some inputs e.g. 3 & 4 missing, some features missing) would be necessary for **end of September 2006.** Full version for **February 2007**.

4.4.3 DISCUSSING THE IMPLEMENTATION

Even though some of the files to be manipulated are not small (e.g. ECOCLIMAP is 0.9 Gbytes), the main difficulty probably lies with the "scenery navigator". However, images where expert eyes are able to determine unambiguously whether forest is forest (and which kind) are necessary, and this implies high resolution.

It may be that the only possibility is a mosaic where the "scenery navigator" is juxtaposed to other images. Then is it necessary that the LAT LON coordinates can be matched to the LAT LON system of the scenery navigator.

5. SPECIFIC ITEMS

5.1 TEST CURRENT MAPS

Check the way RFI map is filled.

Check the way TAU and HR maps are filled. In the particular case of the TAU map, see discussion about dual step scenario.

Check that DAP includes information about current maps being filled.

5.2 LOOSE ENDS AND MISSING ITEMS

5.2.1 ARTEFACTS

To be elaborated.

There seems to be no difficulty introducing biases either in upwelling BT over a large range of incidence angles (**simulating RFI**) or in antenna level BT (simulating outliers).

Enhancing the radiometric sensitivity for non EAF or border views is foreseen. However, there is a high uncertainty on the enhancing coefficients to adopt, as long as relevant information from L1 is not available.

5.2.2 WEF

To be elaborated

5.2.3 DT ITEMS

Check that the procedure allows testing the "winter forest" DT1 branch (see ATBD).

6. IMPLEMENTATION PLAN

6.1 PLAN FOR PHASE 1

What follows is a free indication of elements for the plan.

Here are some tasks for the plan (**copied from earlier sections**). It was not possible to remove the TBC other than stating that solving these issues is actually part of Phase 1 !

1. Some choices in phase 1 need tuning. **This task is to be scheduled in the plan**. In particular, every numerical value for relevant parameters in the UPF should be revisited.
2. The sampling loops in Table 1 have been tailored down in order to obtain a situation where a single L1c product is needed for the DQX simulation. This situation is indeed highly convenient for the retrieval runs. This estimation is still TBC because it assumes that most/every filtered DGG nodes are relevant from the point of view of the X_SWATH sampling. While a loop XL is indicated in Table 1, it will certainly have to be tuned further. **This task has to be scheduled in the plan**.
3. The non uniformity tests for TAU and SM are not fully defined, and this requires finalizing. **This task has to be scheduled in the plan**. However, note the numbers of cases are comparatively small with regard to the overall number of simulations. For these tests, the simplest is to compute the up-welling TB by averaging the results of two direct models.
4. The retrievals are be carried out for the single step and repeated for the dual step retrieval scenario, only for the TAU_medium option. In the latter case, it is proposed that the input ASTD for TAU at a given X_SWATH will be chosen from the single step retrieval results for abscissa (M_SWATH – X_SWATH). The M_SWATH value will be tuned depending on the TH_MMIN1 verification result. This algorithm will have to be refined (**this task must be scheduled in the plan**), accounting for the detailed structure of the X_SWATH sampling list.
5. Over the (SM, TAU) plane, integrated performance figures will finally be computed through averaging. This average should ideally be weighted depending on the actual climatological occurrence of SM and TAU values. In absence of accurate knowledge, no (or very raw) weighting will be applied: the SM range covers every realistic value, and the TAU range is limited to various values up to 0.6. Other partitions of the plane (TBD) might be considered; **listing them is a task to be scheduled in the plan**.
6. A **task for analysing** the "secondary" cardioid retrievals **must be scheduled in the plan.**

Actually, several of these tasks can be concatenated together in a single action, which consists of **pre-simulating the phase 1 validation**. This will be carried out by ESL at least for both DQX and bias runs and will allow addressing issues 1,3, 4, 5, 6 as well as defining accurately the procedure to use the retrievals.

Furthermore, all these tasks correspond to preparatory steps. Formally, they will be carried out between November 15 and the beginning of simulations (possibly around January 15). Actually, some of these tasks are obviously already underway.

Table 3c summarizes the tasks to be carried out. The simulated L1c products are referred to by the TBL number; the retrievals are referred to by the number in the 1rst column of Table 3a.

Table 3 c: summary of data building for Phase 1

prior to simulation			When (monthes from T0=onset)
		Simulate Phase 1on breadboards, tune parameters, prepare processing	T0-2 to T
		Obtain DGG list and X_SWATH	T0-1 to T-0.5
	TBL1	define locations of nodes for TBL1 according to X_SWATH availability	T0-0.5 to T0
	TBL2 (=4)	define restricted location of nodes	T0-0.5 to T0
	TBL3	define restricted location of nodes	T0-0.5 to T0
direct simulation			
	TBL1	assign input parameters	T0 to T0 + 0.5
		compute DTB	T0 to T0 + 0.5
		compute MVAL	T0 to T0 + 0.5
	TBL2	repeat for restricted list of nodes	T0 to T0 + 0.5
	TBL3	mix TAU values and repeat	T0 to T0 + 0.5
	TBL4	mix SM values and repeat	T0 to T0 + 0.5
	TBL1 to 6	fill spurious nodes with out of bound values; feed L1c data (TB & DTB)	T0 to T0 + 0.5
	TBL5 & 6	modify Rotation and Faraday (M12 & M22); feed L1c data	T0-0.5 to T0
		Built input	T0 to T0 + 0.5
prior to retrieval			
	TBL1	build LAI map	T0-0.5 to T0
		build uniform map for SM	T0 to T0 + 0.5
		build uniform maps M31, M41, M51, M61	T0 to T0 + 0.5
retrieval			
	1 to 8	use TBL1, adjust UPF	T0+0.5 to T0+1
		tune thresholds (repeat if necessary)	T0+0.5 to T0+1
	1 to 8	use TBL1, adjust UPF	T0+0.5 to T0+1
	9	use TBL3, adjust UPF	T0+0.5 to T0+1
	10, 11, 12	use TBL4, next TBL5, next TBL6	T0+0.5 to T0+1
	13 to 16	use TBL2; replace M32, M42, M52, M62	T0+0.5 to T0+1
		use retrievals from 6 to build CURRENT_TAU_LV map	T0+1 to T0+1.5
	17 to 28	repeat 5 to 16 for dual step	T0+1 to T0+1.5
Process results			T0+1 to T0+2
		write report	T0+2 to T0 + 2.5

• Many aspects of phases 2 and 3 are not fully defined. The open issues have to be addressed while carrying out phase 1. This task has to be scheduled in the plan.

6.2 PLAN FOR PHASE 2

- Tasks and open issues,
- Completeness and reliability of L1 algorithm,
- Discuss where noise is to be added to BT, reference values and fractions,
- Generate ECMWF maps,
- Improve cover map (topography, non permanent features),
- Devise direct simulation for melting snow, and

- Define classes for 3..4 and beyond.

6.3 PLAN FOR PHASE 3

Tasks and open issues:

Use real data from campaigns, Cal Val activities, real SMOS data.

Requires L1 to be fully validated

Validated approaches on not-so-well-known targets

Address exotic targets.