CHARACTERIZING THE SURFACE DYNAMICS FOR LAND COVER MAPPING: CURRENT ACHIEVEMENTS OF THE ESA CCI LAND COVER

Céline Lamarche (1), Sophie Bontemps (1), Astrid Verhegghen(1), Julien Radoux(1), Eric Vanbogaert(1) Vasileios Kalogirou (2), Frank Martin Seifert(2), Olivier Arino (2), Pierre Defourny (1)

(1) Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium,
Email: celine.lamarche@uclouvain.be, sophie.bontemps@uclouvain.be, astrid.verhegghen@uclouvain.be, eric.vanbogaert@uclouvain.be, julien.radoux@uclouvain.be , pierre.defourny@uclouvain.be
(2) European Space Agency, Frascati, Italy, Email: vasileios.kalogirou@esa.int, olivier.arino@esa.int

ABSTRACT

Land Cover (LC) was listed as an Essential Climate Variable by the Global Climate Observing System and included the ESA Climate Change Initiative (CCI) that aims at providing global long-term satellite-based products tailored to the need of the climate modelling community. In the framework of the CCI-LC project, the LC concept was revisited in order to reconcile the LC users' divergent needs for both stable/consistent global LC products over time and more dynamic information related to the dynamic processes of the land surface. This paper aims first at describing the three global products generated in response to this need for more dynamic information, namely the condition products. These products characterize globally the green vegetation phenology, the burnt areas and snow occurrences. The main challenge beyond the production of these datasets refers to the spatio/temporal consistency between the stable and dynamic components of the LC. The second objective of this paper is therefore to address the work on-going on the characterization of this consistency.

1. INTRODUCTION

In order to define the information needs in support to climate science, the Global Climate Observing System established a list of Essential Climate Variables (ECV), selected to be critical for a full understanding of the climate system and currently ready for global implementation on a systematic basis. In response to the ECV list, ESA initiated a new program - namely the Climate Change Initiative (CCI) –to develop global monitoring datasets to contribute in a comprehensive and timely manner to the need for long-term satellite-based products in the climate domain. Among the 14 ESA CCI components respectively addressing the atmospheric, oceanic and terrestrial domains, the ESA CCI Land Cover (CCI-LC) project is dedicated to land cover (LC) characterization. LC is indeed referred to as one of the most obvious and commonly used indicators for land surface and the associated human induced or naturally occurring processes, while also playing a significant role in climate forcing. This project builds on the ESA-GlobCover projects experiences [1, 3]. It aims at revisiting all algorithms required for the generation of a global LC product from various Earth Observation (EO) instruments that matches the needs of key users' belonging to the climate modelling community.

First, a user requirements analysis was completed with this community to identify its specific needs in terms of satellite-based global LC products. This analysis highlighted a set of requirements in terms of thematic content, spatial and temporal resolution, stability and accuracy that are not met by existing global products.

A first finding of particular interest was the priority for both stable and consistent LC products over time. A second interest was also expressed for more dynamic information reflecting LC change and vegetation phenology. However, the most recent series of global land cover products are specifically pointing out this inconsistency issue as a quite difficult one [6, 2]. Yet, it must be recognized that the land cover cannot, at the same time, be defined as the physical and biological cover on the Earth's surface [12,5] and remains stable and consistent over time as expected by most users.

This conclusion calls for the development of a new land cover ontology, which explicitly addresses the issue of inconsistency between annual land cover products and/or of products sensitivity to the observation period. The proposed land cover ontology assumes that the land cover is organized along a continuum of temporal and spatial scales and that each land cover type is defined by a characteristic scale, i.e. by typical spatial extent and time period over which its physical traits are observed [13]. This twofold assumption requires introducing the time dimension in the land cover characterization, which contributes to define the land cover in a more integrative way.

Accounting for the time dimension allows distinguishing between the stable and the dynamic component of land cover. The stable component refers to the set of land elements which remain stable over time and thus define the land cover independently of any sources of temporary or natural variability. Conversely, the dynamic component is directly related to this temporary or natural variability that can induce some variation in land observation over time but

Proc. 'ESA Living Planet Symposium 2013', Edinburgh, UK
9–13 September 2013 (ESA SP-722, December 2013)

without changing the land cover feature in its essence. This dynamic component, named hereafter the LC Condition, is typically driven by biogeophysical processes and encompasses different observable variables such as the green vegetation phenology, water presence, burnt scars and snow coverage.

Finally, the users' survey highlighted the superior importance of consistency between different model parameters (derived from the LC state and conditions) to the accuracy of individual parameters themselves. Discrepancies between products remain a source of inaccuracy and lead to interoperability problems in climate models.

This paper describes first the three condition products generated in the framework of the CCI-LC project to address the need for a dynamic component in the LC with regards to processes that are of interest for the Climate Modelling Community (CMC): the Normalized Difference Vegetation Index (NDVI), Burnt Areas (BA) and Snow conditions. It presents then the next steps foreseen to characterize the consistency between the stable LC and the conditions.

2. THE CONDITION PRODUCTS

On a per pixel basis, the condition products reflect, along the year, the average trajectory and the intra-annual variability of a land surface feature over the 1998-2012 period. They are built from existing long-term global datasets with high temporal frequency and moderate spatial resolution (250-1km) (Tab. 1).

They result from a compilation a 13 years of 7d instantaneous observations into 1 temporarily aggregated profile depicting, along the year, the reference behaviour for vegetation greenness, snow and BA. All conditions are expressed as 7-day time profiles of the mean and standard deviation for continuous variables (NDVI) or as temporal series of percentage of occurrence for discrete variables (snow and BA).

This condition concept is illustrated in Fig.1 through annual temporal series of the snow occurrence of an Alpine forest pixel covering the 2000-2009 period. Each coloured square from the upper figure correspond to the status flag of the pixel which can be snow (black), land (grey) or no data (white). From these observations, the ratio of the sum of snow detections to the number of valid observations is inferred at each time steps and therefore forms the snow condition product. This snow climatology represents the reference snow behaviour of this particular pixel for the 2000-2009 period.

The products are delivered in time series of 52 files (1 file per 7-day time interval), each file being made of measurements and quality flag layers. They are available in GeoTiff or NetCDF format in a Plate-Carrée projection.

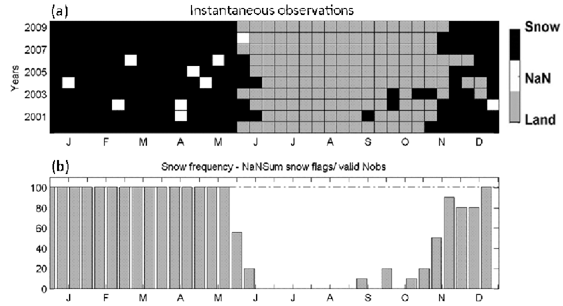

Figure 1. Illustration of the condition concept. Instantaneous observations of the snow presence of an Alpine forest pixel during 10 years from 2000 to 2009 (a) are temporarily aggregated at each 7-day time step to generate the corresponding snow condition product defined as a reference snow frequency (b).

Table 1.Overview of the condition products delivered in the CCI-LC project and their main characteristics.

Condition product	Reference period	Spatial coverage and resolution	Data source
NDVI	1999-2011	Global 1000m	SPOT-VGT S1
Burnt Areas	2000-2012	Global 500m	MCD64A1 product
Snow	2000-2012	Global 500m	MOD10A2 product

1.1. NDVI

The NDVI condition product describes globally the yearly reference dynamic of the vegetation greenness characterizing the 1999-2011 period. It is therefore a valuable reference dataset for phenology studies and phenological metrics extraction at global scale [15].

It is built from 13 years of SPOT-VGT daily top of canopy SR syntheses (S1 products) and of related quality flags.

The methodology consists in four steps. Firstly, a moving average window filter spatially aggregates the reflectance values in order to reduce the noise. Second, a NDVI vegetation index is computed from the Red and NIR reflectance values. Third, the NDVI profiles are smoothed with a Wittaker filter. Finally, the resulting 13-year time series are aggregated over years, 7-day period by 7-day period.

It results in 2 main time series of measurements. First, the annual behaviour of the vegetation is characterised by 52*7-day smoothed average NDVI time series and second, the inter-annual variability is defined by its

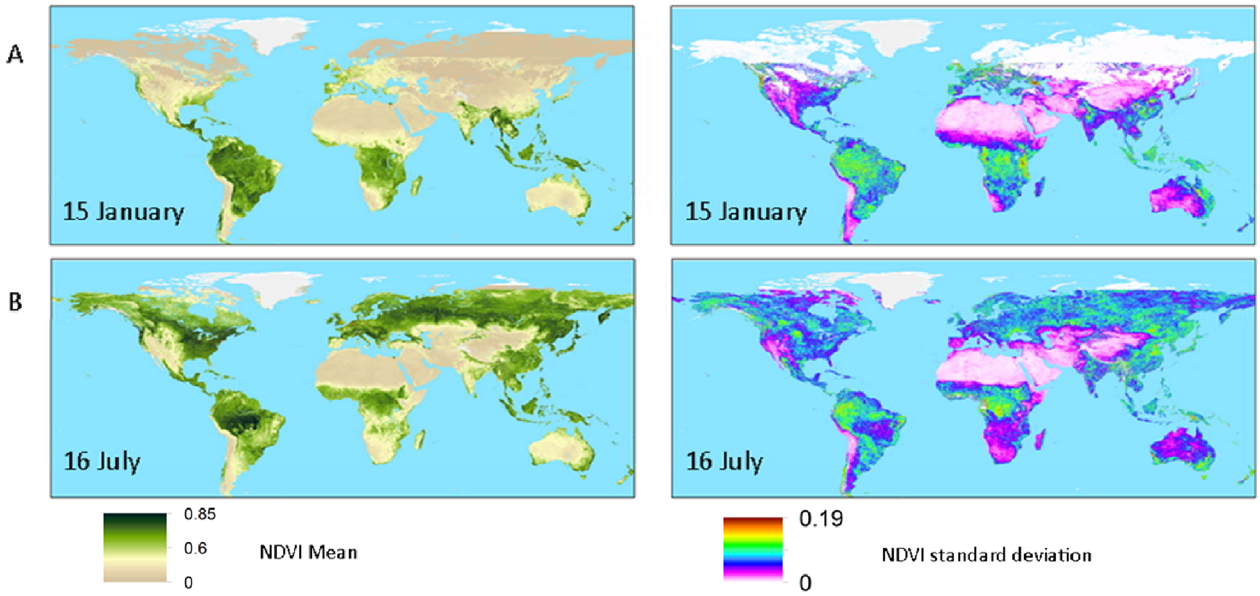

Figure 2: Mean (left) and standard deviation (right) components of the NDVI Condition Product at 4 different seasons. The dates indicated in Figure A, B, C and D correspond to starting day of the 7-day composite period. White colour situated in high latitudes corresponds to NoData values [15].

respective 52*7-day standard deviation. In addition, 2 quality flags are provided at the pixel level: the number of valid cloud-free observations used to generate the NDVI averages and the status qualifying the pixel. The mean component of the resulting NDVI condition is illustrated for the 4 seasons of the year in Fig.2.

The climatologic product clearly captures the spatial pattern of many land features, including in the cloudiest regions of the world like the equatorial areas (Fig.3). Three NDVI condition profiles, extracted on (i) a mosaic class of cropland, (ii) a mosaic class of tree and shrub cover types and (iii) a land cover class made of broadleaved deciduous trees, demonstrate the spatial consistency of the product and its capacity to depict the intra-annual variability of the vegetation greenness.

1.2. Burnt Areas

The BA condition product presents the frequency at which burnt areas have been detected along the year on given pixel, based on observations over the 2000-2012 period. It provides the percentage of BA occurrence with a 7-day temporal resolution and depicts the seasonal dynamic of the fire impact on the surface.

The BA input data are currently derived from the MODIS Direct Broadcast Monthly Burned Area Product (MCD64A1) [10] being part of the Global Fire Emissions Database version 3 (GFED.v3) products [11]. It is available from 2000 to present, at a 500m spatial resolution.

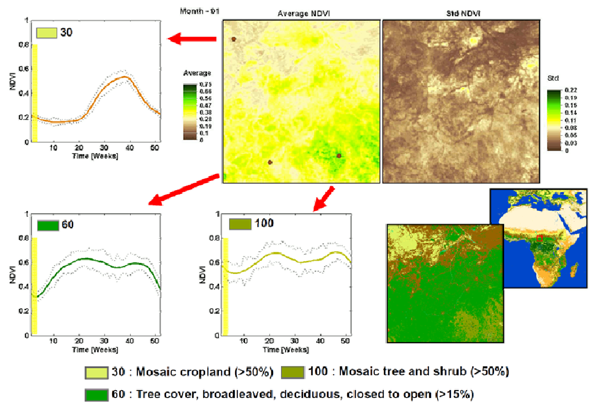

Figure 3. Detailed spatial example of NDVI condition profiles - mean (plain line) and standard deviation (dotted line) - extracted in region of Central Africa. The profiles are extracted from 3 pixels belonging to 3 land cover classes of the 2010 land cover state map product. The variety of the dynamic of vegetation is clearly captured.

As soon as available, the Level 3 CCI-Fire disturbance product generated by the ESA CCI Fire project [2] will be used as input. It is important to mention that such a condition is built on products time series but does not substitute these time series as the condition only aims to provide an average behaviour and not the fire activity from one year to another.

The BA condition product is composed of two series of 52 layers. The first one represents the percentage of BA occurrence, on a 7-day basis, calculated as the sum of

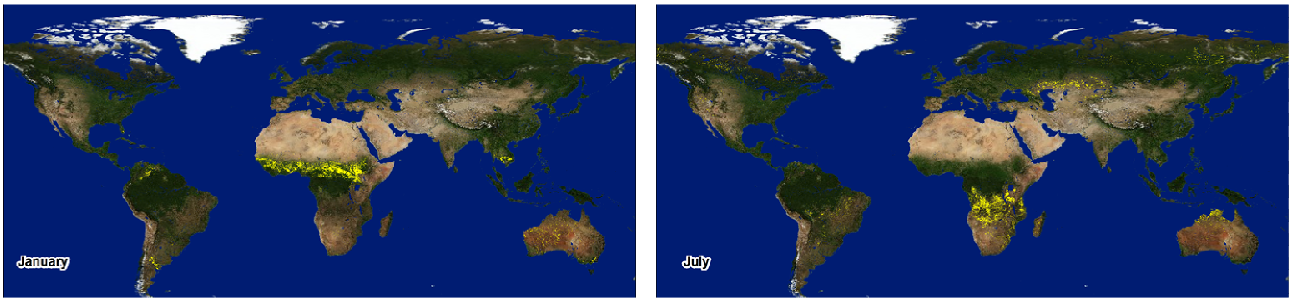

Figure 4. Illustration of the global reference BA occurrence for January (left) and July (right). Each yellow dot represents a percentage of Burnt Areas occurrences different from 0.

BA detections over the number of years in the aggregation period (currently 13 years – 2000 to 2012). It is expressed between 0 and 100. The second layer gives, on a 7-day basis, the number of valid years contributing to each 7-day period. It stands for a quality indicator of the occurrence values and is expressed between 0 and 13. Fig.4 globally shows the percentage of BA occurrence characterizing the reference BA occurrence of January (left) and February (right).

1.3. Snow

The Snow condition product presents the proportion of snow occurrence detected, along the year based on observations over the 2000-2012 period. It provides the seasonal dynamic behaviour of snow coverage with a 7-day temporal resolution.

The snow input data originates from the 8-day maximum snow extent product (MOD10A2) [14]. Data are respectively available freely from 2000 and 2002 up to now at global scale. All products come in tiles of 1200*1200 km and share the same technical properties: a sinusoidal projection and HDF format.

The generation of the snow condition relied on 3 main steps. A spatial filter was first applied at each 8-day period to reduce the occurrence of clouds. It works as a conservative erosion filter as only small clusters of “cloud” pixels received the valid label of their surrounded neighbours, provided they all belonged to an identical and valid class. The second step calculates, for each period, the percentage of snow occurrence as the sum of snow detection over the number of valid observations. Finally, the 8-day time series is converted into 7-day time series by linear interpolation at the mid-period location. The day of year is computed for the middle of each 7 day period and the corresponding previous and next eight-day periods are selected. The new observed value is then linearly interpolated between these two dates.

The snow condition product is composed of two series of layers. The first one represents the percentage of snow occurrence, on a 7-day basis, calculated as the sum of detections over the number of valid observations in the aggregation period. It is expressed between 0 and 100. The second layer gives, on a 7-day basis, the number of valid years contributing to each 7-day period. It stands for a quality indicator of the occurrence values and is expressed between 0 and 13. Fig.5 shows the snow condition product for the first week of January.

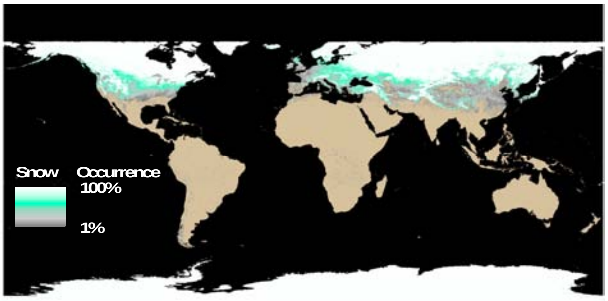

Figure 5. Global percentage of snow occurrence characterizing the 1st week of January. “No snow coverage” is presented in beige and “NoData” in black. Optical data are no longer available at latitudes above 68°N due to the absence of solar illumination.

These three condition products are complementary to the three CCI-LC global maps products characterizing the stable component of the land cover for the same period (Fig. 6) (see Bontemps et al. – also submitted to the same conference).

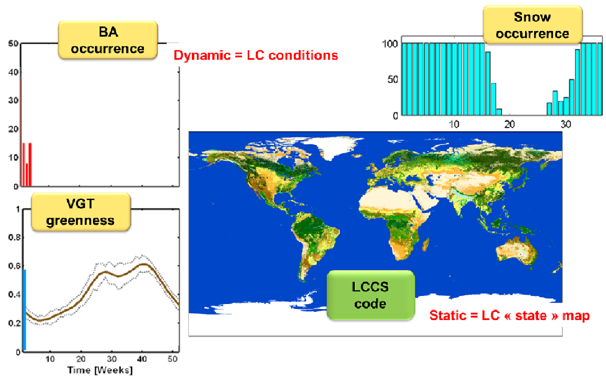

Figure 6. Overview of the condition products generated in the CCI-LC project and their complementarity with the LC map.

3. CHARACTERIZING CONSISTENCY

Analysing the consistency between the condition products and the LC maps, as well as among conditions represents an enormous challenge in a context of multiple products, built independently from each other. Input data originate from various sensors (SPOT-VGT, MODIS for the conditions and MERIS FR for the LC map), spatial resolutions (1km, 500m and 300m respectively) and algorithms. In addition, 3 LC maps were released at 300m spatial resolution and for 3 epochs centred on the year 2010 (2008-2012), 2005 (2003-2007) and 2000 (1998-2002).
Various possible approaches could be considered to create a formalized typology of observable discrepancies. This typology is expected to lay the foundations to build a pixel-based discrepancy indicator. Some discrepancies are presented here below as a brief outline of the discrepancy characterization. The list should be completed in further analyses.

A first type of discrepancies relates to the fundamental and unambiguous aspects of conformity between LC map and conditions.
A first aspect relates to the spatial coverage foreseen as global and identical for all LC products. This is currently not the case as the spatial coverage of the snow condition is affected by the lack of solar illumination at extreme latitudes. In the Northern Hemisphere, the upper average latitude including valid observations decreases from 80° N (September and October) to 60°N from November to January.
A second aspect concerns the definition of the land mask that spatially delineates the area concerned by the LC. Discrepancies were observed between the LC map and condition products. In the BA condition product, no status layer exists and coastlines are, for some weeks, undefined over large regions. There is therefore a need for a common reference land/water mask between the CCI-LC maps and conditions product. A potential candidate could the CCI-LC WB product (see Santoro et al. – also submitted to the same conference).

A second type of discrepancies could be defined for pixels that are described simultaneously by a condition and a LC map class such as permanent snow and ice (class 220). The permanency inferred by the definition of the class label should be confirmed by the condition product and vice versa. In this context, discrepancies could be tested as follows:

- the stable character of each pixel from LC class water bodies and permanent snow and ice should be confirmed, on a 7-day basis, by a high percentage of water and snow occurrence, respectively;
- accordingly, high and frequent percentage of water and snow occurrence in the condition products should be classified as LC class water bodies and permanent snow and ice.

To take into account the natural inter-annual variability of these land surface features as well as the uncertainties in the observations, discrepancies of permanency should be tested for various thresholds of percentage and duration in the condition occurrence. The key issue will be to determine the threshold at which a disagreement is no longer tolerable and therefore considered as a true discrepancy. The same methodology could be applied on water bodies (class 210) in the case a water bodies product was generated.

A third type of discrepancies could be defined for pixels that are not described simultaneously by a condition and a LC map class but whose class label infers a particular and expected behaviour with regards to the conditions. For continuous conditions such as the NDVI, the discrepancy analysis could, for example, rely on phenological metrics. For discrete conditions, categories of probabilities (null, low, medium, high) of the event occurrence over a particular LC map class could be tested.
This discrepancy characterization is less straightforward and would necessitate the implementation of a-priori test cases between the LC class and the conditions. Distinct rules could be defined according to the continuous or discrete character of the condition.

Fig.7 spatially illustrates this third type of discrepancies. It shows hot spots of discrepancies between the BA condition and some LC map classes that have little chance of burning:

- 160: Tree cover, flooded, fresh or brakish water
- 170: Tree cover, flooded, saline water
- 180: Shrub or herbaceous cover, flooded, fresh/saline/brakish water
- 190: Urban areas
- 200: Bare areas

- 210: Water bodies
- 220: Permanent snow and ice

In this case, discrepancies were flagged if the percentage of BA occurrence > 0% at least once over the reference yearly condition profile. The same test should be set with less strict thresholds in order to take into account the possible inaccuracies existing in the BA condition product.

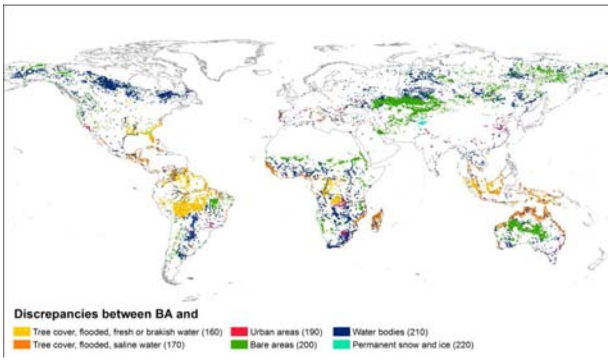

Figure 7: Illustration of discrepancies detection when BA is detected at least once on a LC class with little chance burning.

A last type of discrepancies would be considered between conditions. Some pairs of conditions are mutually exclusive as the event they represent cannot occur simultaneously on a given pixel (red arrows in Fig.8). This is the case for BA versus Water and BA versus Snow. As the conditions are Level 4 products aggregated over ~13 years, it is rather difficult to compare them precisely. A way to highlight this type of discrepancy could be resumed by testing that the sum of percentages of occurrence of mutually exclusive events does exceed 100%.

The case of the continuous conditions, such as the NDVI, is less clear. The NDVI is affected by the other conditions, but to a less precise extent (blue arrows, Fig.8).

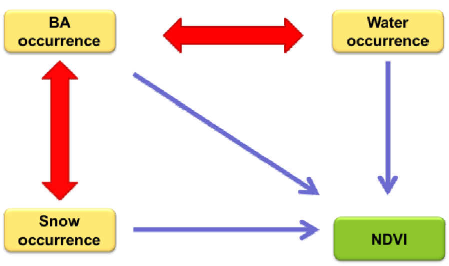

Figure 8: Illustration of the relationships between conditions in terms of discrepancies. Red arrows show mutually exclusive conditions. This is the case for the pairs: BA-Snow and BA-WB. The NDVI is affected by the other conditions to a less defined extent.

The discrepancies related to the NDVI could be characterized through tests of the temporality of events. This is illustrated in Fig.9 where the percentage of burnt areas occurrence (top) is compared to the NDVI condition (bottom) for a zone of Central Africa. The temporal evolution of both variables shows they are consistent with each other because the BA season occurs before and after the period when the vegetation is at its maximum greenness.

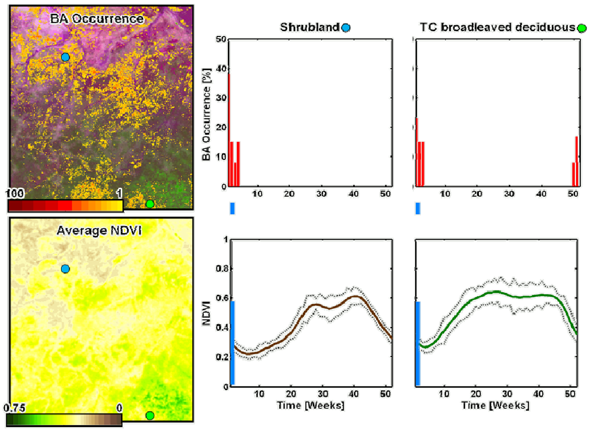

Figure 9: Spatial example of burnt areas occurrence (Top-left) and NDVI condition (Bottom-left) products characterizing the reference behaviour of the 1st week of January. Yearly discrete occurrence profiles illustrate the BA behaviour for two pixels originating from the shrubland and tree cover broadleaved deciduous LC classes extracted in region of Central Africa. The Burnt Area season and the period of maximum greenness are not concomitant, indicating that both conditions are consistent.

After the definition of a typology, the next steps will be to build a discrepancy indicator that integrates the various types of observed discrepancies. It is foreseen to concentrate our efforts on thematic discrepancies existing with the LC maps, in priority to the discrepancies between conditions. Indeed, from a climate point of view, the LC map is the integrative variable that is assimilated in the climate models and from which land surface parameters are derived. The discrepancy between conditions could possibly serve as a sort of preliminary step to discard the pixels with discrepancy.

It can be expected that each pixel undergoes a sequence of hierarchical tests that investigate the various discrepancies characterized in the typology. The discrepancy indicator could be built, by pixel and in the form of categories of discrepancies' severity, according to the type of the discrepancy detected. The proper format of the discrepancy indicator will be built in partnership with the climate modellers.

4. CONCLUSION

In the framework of the CCI-LC project, three condition products were generated at global scale, to better meet the need for more dynamic information expressed by the LC data and climate user communities. These products characterize, on a 7-day frequency, the average reference behaviour for the green vegetation phenology (through the NDVI), the burnt areas and snow occurrences under the form of climatology datasets. These products were derived from global long-term existing EO datasets with moderate spatial resolution and high temporal frequency.

The major challenge but also the main added-value beyond the compilation of these already existing data sets is the consistency through space/time, with the LC and between conditions. It is indeed intended to describe the whole dynamic of the terrestrial surface in a meaningful way over time. As the respective production of the different data sets was fully independent and sometimes sensor-dependent, discrepancies and incompatibility were observed and highlighted. The methodology to characterize the consistency between LC maps and conditions as well as among conditions is currently on-going.

These condition products should be considered as a first version of an integrative description of the terrestrial surface. The sensor-independent processing algorithms of these products and a better consistency between sensors, such as the one expected among Sentinel instruments, allow expecting major improvements in consistency, spatial and temporal resolution. Of course, other conditions could be also considered such as the Leaf Area Index (LAI) or Land Surface Temperature (LST) and help enrich the land surface characterization and modelling.

5. REFERENCES

1. Arino O., Bicheron P., Achard F., Latham J., Witt R. and Weber J. L., 2008, Globcover: the most detailed portrait of Earth, ESA Bulletin 136, 24–31

2. Bontemps S., Herold M., Kooistra L., van Groenestijn A., Hartley A., Arino O., Moreau I. and Defourny P., 2012, Revisiting land cover observation to address the needs of the climate modeling community, Biogeosciences, 9, 2145–2157

2. Chuvieco, E, Calado, T, Oliva, P (2011), ESA CCI ECV Fire Disturbance (fire_cci) Product Specification Document (PSD)

3. Defourny P.et al., 2009, Accuracy assessment of a 300-m global land cover map: the GlobCover experience, in: Proceedings of the 33rd International Symposium on Remote Sensing of Environment, Stresa, Italy, 4–8 May 2009, PS-B1-3 (Ref 863)

4. Defourny P. and Bontemps S., 2012. Revisiting Land-Cover Mapping Concepts. In: Remote Sensing of Land Use and Land Cover : Principles and Applications, 2012, p. 49-63. 978-1-4200-7074-3. http://hdl.handle.net/2078/114554

5. Di Gregorio, A. (Ed.). (2005). Land Cover Classification System: Classification Concepts and User Manual: LCCS (No. 8). FAO.

6. Friedl, M. A., Sulla-menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). . Remote Sensing of Environment MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1), 168–182. doi:10.1016/j.rse.2009.08.016

10. Giglio, L., Loboda, T., Roy, D.P., Quayle, B. and Justice, C.O., 2009. An active-fire based burned area mapping algorithm for the MODIS sensor. Remote Sensing of Environment, 113(2): 408-420.

11. Giglio, L., Randerson, J.T., van der Werf, G.R., Kasibhatla, P.S., Collatz, G.J., Morton, D.C. and DeFries, R.S., 2010. Assessing variability and long-term trends in burned area by merging multiple satellite fire products. Biogeosciences, 7(3): 1171-1186.

12. Herold, M. (2009). Land Cover. Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables (p. 36). Rome: Global Terrestrial Observing System.

13. Miller, R. I. (ed.). 1994. Mapping the Diversity of Nature Chapman and Hall, London.

14. Riggs, G.A., Hall, D.K. and Salomonson, V.V., 2006. MODIS Snow Products, User Guide to Collection 5.

15. Verhegghen, A., Bontemps, S. and Defourny, P., In revision for publication. Global land-surface phenology reference dataset from 13 years NDVI and EVI SPOT VEGETATION time series analysis. International Journal of Remote Sensing.