Distribution: Restricted

Delivery report

TerrA-P: Development and validation of a global GPP/NPP model using MERIS and Sentinel-3 data

Requirement Baseline Document

Else Swinnen, Roel Van Hoolst, Colin Prentice, Ivan Janssens, Sara Vicca, Manuela Balzarolo

Study accomplished under the authority of ESA

May 2017

DISTRIBUTION LIST

Philippe Goryl, ESA-ESRIN

Else Swinnen, VITO
Roel Van Hoolst, VITO
Bruno Smets, VITO

Iain Colin Prentice, ICL
Rebecca Thomas, ICL

Ivan Janssens, UA
Sara Vicca, UA
Manuela Balzarolo, UA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ACRONYMS

ABP	Aboveground Biomoss Production
ATBD	Algorithm Theoretical Basis Document
BP	Biomass Production
C	Carbon
CCI	Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
CGLOPS	Copernicus Global Land Service
DM	Dry Matter
DMP	Dry Matter Production
ECMWF	European Centre for Mid-range Weather Forecast
EO	Earth Observation
ESA	European Space Agency
FAPAR	Fraction of absorbed photosynthetically active radiation
FLEX	Fluorescence Explorer
GIS	Geographic Information System
GPP	Gross Primary Production
ICL	Imperial College London
LAI	Leaf Area Index
LPV	Land Products Validation
LUE	Light Use Efficiency
MARSOP	Monitoring Agricultural ResourceS OPerations contract
MODIS	Moderate Resolution Imaging Spectroradiometer
NPP	Net Primary production
PSN	Net Photosysthesis
RMSE	Root mean squared error
SPOT	Satéllite Pour l'Observations de la Terre
UA	Universiteit Antwerpen
VGT	VEGETATION instrument
VITO	Vlaamse Instelling voor Technologisch Onderzoek

CHAPTER 1 INTRODUCTION

1.1. SCOPE AND OBJECTIVES

The objective of the ESA SEOM-funded TerrA-P project is the development, prototying and validation of global land vegetation production products based on Sentinel-3 data. The targeted production products are Gross Primary Production (GPP) and Aboveground Biomass Production (ABP). The products will be developed based on the MERIS Global Vegetation Index (GVI) in order to have sufficient calibration and validation in situ data available. Later in the project, the method is applied on the Sentinel-3 fAPAR data for further validation.

The purpose of this document is twofold: (1) to describe the user requirements based on a consultation of the potential user community, and (2) to define the methods for validation and evaluation of the model and the resulting products.

1.2. CONTENT OF THE DOCUMENT

The RBD is organized in the following way:

- **Chapter 2** describes the set-up and the results of the User Requirements Survey.
- **Chapter 3** details the validation approach using in situ data and the benchmarking against similar EO-derived data sets.
- **Annex A** lists the detailed answers of the User Requirements Survey.

CHAPTER 2 USER REQUIREMENTS SURVEY

2.1. INTRODUCTION

In order to collect the user requirements for the GPP and ABP products, a survey among potential users was conducted. A list of questions was prepared and agreed upon with ESA. The survey was done using the free available surveytool SurveyMonkey (www.surveymonkey.net).
The question to complete the survey was e-mailed to the attendents of the ESA FLEX workshop held in January 2017, and a number of identified potential users. We also requested the Copernicus Global Land Service to send out the request among their users of the DMP product.
The survey was anonymous.

2.2. USER REQUIREMENTS SURVEY QUESTIONS

The survey covered the following groups of questions that are further detailed in the figures below:

- General
- Product specifications
- Timeliness, periodicity and spatial resolution
- Accuracy
- Other

TerrA-P user requirements

General

*** 1. You are from**

- ◯ Europe
- ◯ Africa
- ◯ Asia
- ◯ Oceania
- ◯ North America
- ◯ Latin America

*** 2. Background expertise**

- ◯ Climate modeller
- ◯ Earth observation scientist
- ◯ Agronomist
- ◯ Forester
- ◯ Other (please specify)

*** 3. Which GPP/NPP products do you use now?**

- ☐ MODIS NPP: MOD17A3
- ☐ MODIS GPP: MOD17A2/MYD17A
- ☐ Copernicus Global Land Service - SPOT-VGT/Proba-V based Dry Matter Productivity (DMP)
- ☐ Modelled data

Other (please specify)

*** 4. Domain / application**

- ☐ Climate modelling
- ☐ Monitoring productivity of CROPS
- ☐ Monitoring productivity of PASTURE
- ☐ Monitoring productivity of FORESTS
- ☐ Early warning for food security

Other (please specify)

5. What is your region of interest?

- ☐ Global
- ☐ Continental
- ☐ Regional
- ☐ National
- ☐ Subnational

Please specify your region of interest.

TerrA-P user requirements

Product specification

Please read the text below.

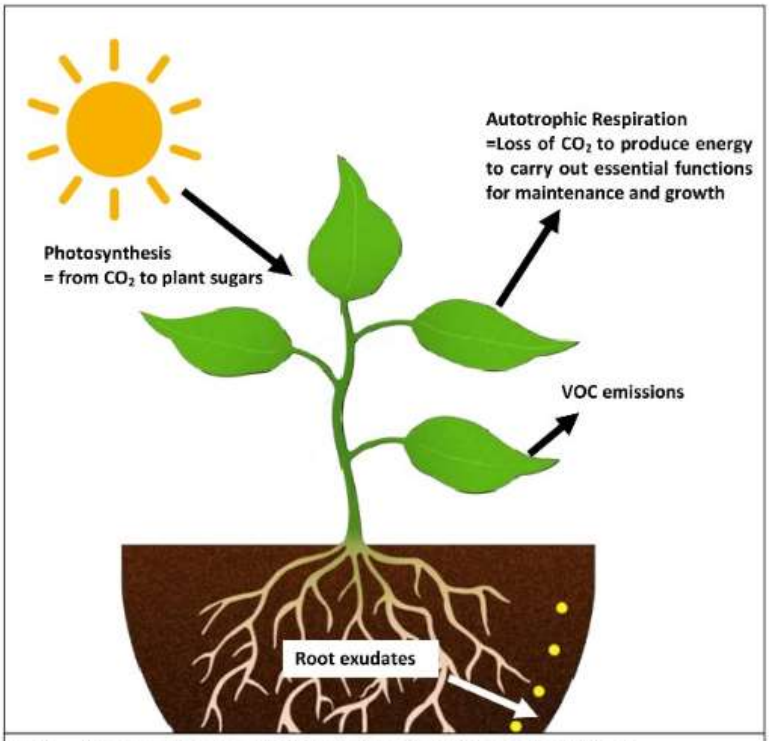
The objective of this section of the survey is to gather requirements on the product specification. An important question is what the content of the product should be. To express ecosystem productivity, different terms and variables can be used. These are **often confused**. Below we describe and comment four commonly used terms.

GPP Gross primary production – the rate of total carbon fixation (photosynthesis) by the ecosystem. This is the most fundamental measure of primary production, as all ecosystem functions depend on it. Also, thanks to flux measurements, GPP data are available – at time scales from half-hourly up to multi-annual – for some hundreds of locations worldwide (albeit with a bias towards temperate regions) and for crops as well as natural and managed ecosystems.

ABP Above-ground primary production – the rate of production of plant matter, excluding roots. This is a practically important measure, because – for example – this is the production rate of forage for grazing animals; it is closely related to the production rate of timber for harvest; and it can be converted (through data on the harvest index of different crops) to estimates of crop yield. There are data on ABP, occasionally at a monthly time scale but more commonly at the annual time scale, for many ecosystems, especially crops and managed forests but also for natural ecosystems.

BP Biomass production – the rate of production of plant matter, including roots. For most crops the root production is of less interest than the above-ground production. (Even for root crops, there are data on the harvest index, i.e. the ratio of yield to ABP.) There are data on BP but in most cases the root production has not been measured directly, but rather inferred from above-ground measurements using standard conversions, thereby increasing the associated uncertainty.

NPP Net primary production – equal to GPP minus respiration. Formerly, NPP was assumed to equal to BP, and most data that claim to be NPP are in fact BP. But it is now understood that a fraction of NPP – under some circumstances this can be as much as 20% of NPP – is "lost" from the plant in the form of volatile organic compounds (such as isoprene) and/or root exudates. Published data sets of "NPP" are of poor quality.



GPP = Amount of carbon fixated produced during photosynthesis
NPP = GPP - autotrophic respiration
BP = NPP – volatile organic compounds emissions – root exudates
ABP = the above-ground part of BP

* **6. TerrA-P will produce GPP and ABP. The description and motivation is given above. Do you agree with this approach?**

- ◯ Yes
- ◯ No, why not?

* **7. In which units should the output product be expressed?**

- ◯ Dry Matter (kg DM/ha)
- ◯ Carbon (g C/m2)

*** 8. Should we distinguish between C3/C4 plants?**

- ◯ No
- ◯ Yes

If yes, how to distinguish them spatially?

*** 9. Which additional information will you use?**

- ☐ Quality layer (info on no-data pixels due to clouds, missing input data, etc.)
- ☐ Information on uncertainty
- ☐ Other (please specify)

*** 10. Which uncertainty information should be available?**

*** 11. How will the uncertainty be used in your application?**

TerrA-P user requirements

Timeliness, periodicity and spatial resolution

*** 12. Periodicity: What is the preferred revisit for your application?**

	Daily	5-daily	7-daily	10-daily	Monthly	Annually	No-revisit (only archive)
GPP	◯	◯	◯	◯	◯	◯	◯
ABP	◯	◯	◯	◯	◯	◯	◯

*** 13. Timeliness: when would you need the products?**

	After 1 year	After 1 month	Within 5 days after acquisition	Within 3 days after acquisition
GPP	◯	◯	◯	◯
ABP	◯	◯	◯	◯

*** 14. Resolution: What should be the horizontal spatial resolution?**

- ☐ 300 m
- ☐ 1 km
- ☐ 0.25°
- ☐ 0.5°
- ☐ 1°
- ☐ Other (please specify)

TerrA-P user requirements

Accuracy

* **15. What should be the relative accuracy of the product?**

10 (%, high accuracy) Maximum % deviation 50 (%, low accuracy)

* **16. What should be the absolute accuracy of the product?**

0 (g C/m2/year, high accuracy) Maximum deviation in g C/m2/year 600 (g C/m2/year, low accuracy)

TerrA-P user requirements

Other

17. Do you have other requirements?

18. Do you have other questions?

* **19. Would you like to receive data from the project for evaluation?**

- ◯ Yes
- ◯ No

20. Would you like to receive the results of this survey?

- ◯ No
- ◯ Yes, please leave your email adress here.

2.3. RESULTS AND DISCUSSION OF THE USER REQUIREMENTS SURVEY

In this section, we discuss the answers to the survey. The full detailed answers are provided in Annex A.

2.3.1. GENERAL QUESTIONS

The survey was sent to approximately 400 persons, and 42 have responded, which is just over 10%. The majority of the respondents are from Europe (40.5%) and Africa (35.7%) and none from Oceania.
Half of the respondents are EO scientists, followed by climate modeller (14.3%). Only few respondents anwered forester (2.4%) or agronomist (4.8%). Almost a third of the respondents (28.6% or 12 persons) specified their background expertise themselves. These responses can be divided into the following categories: ecology/biology (4), GIS applications (3), agrometeorologist (1), humanitarian aid (1) and carbon cycle modeller (1).

Of the currently used data sets (multiple answers possible) the MODIS NPP (MOD17A3) data (54.8%) and the CGLOPS DMP (57.1%) are the most popular, but the MODIS GPP (MOD17A2/MYD17A2) (38.1%) and modelled data (38.1%) are also often used. Most respondents use multiple data sets. Also mentioned were: in situ fluxnet data that are extrapolated globally and other data generated by the respondent himself.

The domain in which the data sets are used are (multiple answers possible) predominantly the monitoring of production of crops (45.2%), pastures (45.2%) and forests (31%). Over 25% of the respondents used the data for climate modelling (28.6%) and early warning for food security (26.2%). Other application domains mentioned are: modelling effects of fire on vegetation, biomass burning fuel consumption, examining human impact, greenness in urban areas and various land use applications.

The regions of interest reported (multiple answers possible) are regional (65.9%), national (51.2%), global and continental (both 31.7%) and subnational (24.4%).

The above results indicate that there is an interest for the proposed products from a wide user community, serving many applications and spatial scales. It also indicates that most people have already access to similar data sets to serve their applications.

2.3.2. PRODUCT SPECIFICATION

We proposed the respondents that the TerrA-P project focusses on the estimation of GPP and ABP and asked for their feedback on this approach. 75% of the respondents agreed with this approach. The 25% who did not agree gave a number of different reasons. It was stated that Biomass Production (BP) is either lacking or should be preferred over ABP, because it provides a more comprehensive picture especially for crops especially root and tuber crops and is important from ecosystem/climate point of view. On the other hand, the yield of these crops is certainly related to the ABP, and probably by a reasonably invariant factor.
Other remarks were that ABP is difficult to estimate for complex communities like natural grasslands and that a simple parametrization to obtain ABP will likely result in poor quality. The estimation of VOC and root exudate is particularly considered as a difficulty.

One respondent questioned how a comparison between the new products and the already existing long-term historic EO-derived NPP products could be possible if NPP is not produced.
One respondent mentioned that he/she has no interest in GPP. Furthermore, advice was given on the GPP from flux tower measurements, which are actually modelled data.

There is no clear preference for the units in which the GPP and ABP products will be expressed: the results are almost fifty-fifty between Dry Matter (kg DM/ha) (53.6%) and Carbon (g C/m²) (46.4%). When looking at the difference between the climate modellers and the the other domains, then the former requests all g C/m², and for the other domains both units are requested.

The majority of the respondents (60.7%) would like to have a distinction between C3 and C4 plants, but most of them do not know how to make the distinction at the global scale between these two types. Suggestions are: by their level of leaf greenness, annual behaviour or different spectral signatures, based on land cover products or crop classification or by solar-induced chlorophyll fluorescence.
We propose to provide the production for both C3 and C4 plants per pixel, because a given field can be under, say, maize one year and soybean the next. Also, the question of how to distinguish C3 and C4 vegetation from space is unresolved. Methods have been proposed but none that could work globally. Moreover, some natural ecosystems (e.g. prairies in the USA, and savannas all over the tropics) are mixtures. Still et al. (2003, *Global Biogeochemical Cycles*) put together a data set, from multiple lines of evidence, but only with 1˚ resolution – this is useful from a scientific point of view but not an operational one.

We also questioned which additional information layers are requested by the users. The quality layer that labels pixels based on its quality (e.g. clouds, missing, bad quality) and information on uncertainty of the products is requested by almost all respondents (resp. 89.3% and 78.6%).
Vegetation biome distribution, Light Use Efficiency layer and information on the model equations are also requested.
No respondent is interested in having the Type B uncertainty. However, in the detailed answers to the question which uncertainty information should be available, a number of respondents (8) expressed their need for an uncertainty measure that could be considered as Type B uncertainty. These answers are:

- An estimate of the observed uncertainty
- Uncertainty (via sensitivity tests) due to assumed parameters or algorithms
- Model uncertainty
- Statistical
- What the uncertainty is caused by and/or a % or grading of uncertainty
- Uncertainties in estimating cropland GPP
- an estimate of the error you make for each estimation, which includes (propagates) errors from the radiometry and the assumptions you make as you go along in your algorithm. It should be measured as a standard deviation.
- uncertainty on model estimations

Other answers are captured in the quality flag and results of the validation of the products.

On the question how the ucertainty information will be used, the following groups of answers were collected:

- To mask out estimates with high uncertainties, or the weight the derived results according to the uncertainty.
- To compare to modelled data sets
- To understand better the data sets, to have confidence in the results

Only one respondent wants to propagate the uncertainy into derived products.
Several respondents will not use uncertainty information, or are not sure for what it can be used (21.4%).

2.3.3. TIMELINESS, PERIODICITY AND SPATIAL RESOLUTION

The preferred revisit time for the products are provided in .

Table 1 Requested revisit time.

	Daily	5-daily	7-daily	10-daily	Monthly	Annually	No-revisit (only archive)	Total
GPP	28.57% 8	3.57% 1	3.57% 1	35.71% 10	21.43% 6	3.57% 1	3.57% 1	28
ABP	25.00% 7	3.57% 1	7.14% 2	32.14% 9	28.57% 8	3.57% 1	0.00% 0	28

For both GPP and ABP, the 10-daily time step is the most wanted, followed by daily for GPP and monthly for ABP.

The requested timeliness or update frequency of the products is presented in Table 2. There is a higher preference for an update frequency after 5 or 3 days, but the update frequency of after 1 year or 1 month is also still high.

Table 2 Requested update frequency.

	After 1 year	After 1 month	Within 5 days after acquisition	Within 3 days after acquisition	Total
GPP	25.93% 7	18.52% 5	29.63% 8	25.93% 7	27
ABP	21.43% 6	21.43% 6	28.57% 8	28.57% 8	28

In this result, there is a clear distinction between the application domains. Respondents that are active in production monitoring need a higher update frequency than those working with climate models.

The same difference between communities is also clear for the requested spatial resolution. The climate community is satisfied with a spatial resolution ranging from 1 km to 0.5°, whereas the 300 m (and 1 km to a lesser extent) resolution is very popular for the production monitoring community. Some people even requested higher spatial resolution.

2.3.4. ACCURACY

We asked for the preferred relative and absolute accuracy of the GPP and ABP products. The answers for the relative accuracy (expressed in %) ranged between 10% - 35% (see Table 3). The minimum value that could be entered was 10%. For the absolute accuracy (expressed in g C/m²/year), the anwers ranged between 0 and 377 g C/m²/year.

Table 3 Requested accuracy calculated from all respondents

Measure	Relative accuracy [%]	Absolute accuracy [g C/m²/year]
Average	17.14	125.46
Median	20	118
Standard deviation	6.6	94.63
Minimum	10	0
Maximum	35	377

When removing the answers where 0 g C/m²/year absolute accuracy was requested (not realistic), the figures are slightly higher (Table 4).

Table 4 Requested accuracy calucalted from respondents with realistic answers.

Measure	Relative accuracy [%]	Absolute accuracy [g C/m²/year]
Average	19.09	159.05
Median	20	150
Standard deviation	6.29	80.14
Minimum	10	50
Maximum	35	377

2.3.5. OTHER REQUIREMENTS

Other requirements from the respondents are:

- Data products should be available in easy accessible formats such as netCDF (CF compliant).
- The impact of water on ABP should be described.

2.4. SUMMARY OF USER REQUIREMENTS AND PRODUCT SPECIFICATIONS

Based on the analysis of the survey responses, the requirements are summarized in Table 5.

Table 5 summary of requested and proposed product specifications and Target Requirements

Characteristic	Target Requirement/ Product Specification	
	Proposed	Requested
Definition		
Products	GPP ABP, because there is too little in situ data to accurately produce BP or NPP	GPP ABP / NPP / BP
Units	GPP: g C/m²/day ABP: kg DM/ha/day	g C/m²/day and kg DM/ha/day
Further specification	GPP and ABP will be produced for C3 and C4 plants, without distinction of where they occur.	Distinction between C3 and C4 plants
Quality & uncertainty information	Quality flags Per-pixel uncertainties Accuracy assessment of products.	Quality flags Per-pixel uncertainties Information on accuracy of product

Timeliness	Within 3-5 days after acquisition	Within 3-5 days after acquisition
Periodicity	10-daily	10-daily (daily)
Spatial resolution	300 m Provide downsampled products	300 m (or lower)
Accuracy		
Relative accuracy	20%	20%
Absolute accuracy	150 g C/m²/year	120-150 g C/m²/year

REFERENCES

Still et al. 2003, Global distribution of C3 and C4 vegetation: Carbon cycle implications. Global Biogeochemical Cycles, 17 (6) 1-14