

# S5P-NPP Cloud Processor ATBD



document number : S5P-NPPC-RAL-ATBD-0001  
authors : Richard Siddans  
CI identification : CI-7430-ATBD  
issue : 1.0.0  
date : 2016-02-12  
status : Released

## Document approval record

	<b>digital signature</b>
<b>prepared:</b>	
<b>checked:</b>	O. Hasekamp, I. Aben, H. Sihler
<b>approved PM:</b>	
<b>approved PI:</b>	

## Document change record

issue	date	item	comments
0.1	2012-10-30	All	Initial draft version
0.2	2012-12-07	Section 1.4	Possibility of using SRF introduced following comments from SRON
		Section 5	Need for ellipsoid reference geolocation noted
0.3	2013-06-04	Section 4	Added
		Section 6.4	Flow diagram added
		Various	Minor changes in response to SRR/PDR
0.5	2013-06-21	Various	Minor updates following internal review
0.6	2013-06-21	Section 1 and 5	Updates following external review
0.7	2013-11-29	Section 5	Added information on VIIRS bow-tie deletion
		Section 7	Computational resource estimates updated
0.9		Section 4	Updated instrument description
0.10	2014-04-15	Section 4	TROPOMI description now located in separate document
0.11	2014-09-30	Section 7	Recognised need for VIIRS solar zenith angle
		Section 3	Added some terms and definition
		Section 3	Added tables of input/output/internal parameters
		Section 1.4	Resolved previous uncertainties based on current knowledge
		Section 6.4	Added section describing approach to account for difference between nominal FOV and L1GPC
		Section 6.5-6	Added description on computation of SRF averages and introduced this into the step-by-step procedure
		Section 6.6	Modified step-by-step procedure to select region for coordinate transform by first finding nearest VIIRS pixel to S5P centre, rather than using range of view zenith angle.
		Section 6.6	Added approach to determine S5P/NPP time difference
		Section 6.7	Added section on treatment of missing / invalid data
		Section 8	Added reference to validation plan, with brief summary of scope
0.13	2015-09-18		Update for limited release to S5p Validation Team
1.0	2015-02-02		Update following internal review

## Contents

<b>Document approval record</b> .....	<b>2</b>
<b>Document change record</b> .....	<b>3</b>
<b>List of Tables</b> .....	<b>4</b>
<b>List of Figures</b> .....	<b>4</b>
<b>1 Introduction</b> .....	<b>6</b>
1.1 Overview .....	6
1.2 Motivation .....	6
1.3 High-level requirements .....	6
1.4 Basic definitions and assumptions .....	7
1.5 Relationship to possible activities to assess / monitor the Geometric and Radiometric performance of TROPOMI .....	8
1.6 Heritage .....	8
1.7 Document overview .....	8
<b>2 Applicable and reference documents</b> .....	<b>9</b>
2.1 Applicable documents .....	9
2.2 Reference documents .....	9
2.3 Electronic references .....	10
<b>3 Terms, definitions and abbreviated terms</b> .....	<b>11</b>
3.1 Terms and definitions .....	11
3.2 Acronyms and Abbreviations .....	13
3.3 Algorithm Input / Output .....	14
<b>4 TROPOMI instrument description</b> .....	<b>14</b>
<b>5 VIIRS</b> .....	<b>15</b>
<b>6 Algorithm description</b> .....	<b>23</b>
6.1 Overview .....	23
6.2 Coordinate systems .....	23
6.3 Transformation to normalised FOV coordinate system .....	25
6.4 Relationship between L1B ground pixel corners and the nominal FOV .....	27
6.5 Calculation of spatial response function weighted quantities .....	30
6.6 Step-by-step procedure for generating the S5P-NPP records .....	36
6.7 Treatment of missing / invalid data .....	37
<b>7 Feasibility</b> .....	<b>38</b>
7.1 Scaling parameters .....	39
7.2 Input data requirements .....	39
7.3 Input file sizes .....	40
7.3.1 S5P .....	40
7.3.2 NPP .....	40
7.4 Computation time .....	40
7.5 Computer memory requirements .....	40
7.6 Output file size .....	40
<b>8 Validation</b> .....	<b>40</b>

## List of Tables

1 Algorithm input data. ....	14
2 Algorithm output data. ....	15
3 Parameters used internally within the algorithm. ....	16
4 VIIRS spectral bands. ....	17

## List of Figures

1	Location of example VIIRS scene .....	18
2	VIIRS false colour index for example scene .....	19
3	VIIRS M9 band (1.385 micron) reflectance for example scene .....	19
4	VIIRS M15 band (10.7 micron) brightness temperature (BT / K) for example scene .....	20
5	VIIRS cloud mask for example scene .....	21
6	VIIRS thin cirrus index for example scene .....	21
7	VIIRS cloud phase index for example scene .....	22
8	Coordinate system for VIIRS pixel selection .....	24
9	Illustration of coordinate frames .....	25
10	Relationship between nominal FOV and L1GPC, as a function of distance to ground. ....	29
11	Relationship between nominal FOV and L1GPC, as a function of across-track position. ....	29
12	Off-nadir angles used to model S5P SRF .....	30
13	Illustration of VIIRS data integration over an S5P SRF. Extreme East of swath. ....	32
14	Illustration of VIIRS data integration over an S5P SRF. Centre of swath. ....	33
15	Gridded SRFs in normalised FOV coordinates for nadir view geometry .....	34
16	Gridded SRFs in normalised FOV coordinates for edge of swath view geometry .....	35
17	Flow diagram illustrating the main processing steps and I/O .....	38

# 1 Introduction

## 1.1 Overview

The Sentinel 5 Precursor (S5P) mission [AD2] acts as very important transition mission in the time frame between ENVISAT and Sentinel 4 and Sentinel 5. As full GMES mission, S5P shall fulfil demanding requirements of continuous product quality, near-real-time capability as well as dissemination and archiving of level 1 and level 2 data. The single payload instrument TROPOMI jointly developed by NSO and ESA will bring a significant improvement in precision and resolution of derived atmospheric composition products covering lower tropospheric pollutants.

S5P level 2 products, particularly those derived from the short-wave infra-red (SWIR) channel are dependent on having information on cloud at spatial resolution finer than that achievable from TROPOMI itself. It is therefore foreseen to develop a level 2 auxiliary product describing cloud in the TROPOMI field of view, derived from co-located observations of VIIRS (Visible Infra-red Imaging Radiometer Suite) on NPP (National Polar-orbiting Partnership).

This S5P-NPP Cloud product is developed by the Rutherford Appleton Laboratory (RAL) through an ESA project [AD1], in close collaboration with the teams developing the main TROPOMI level 2 products (BIRA, IUP, SRON, KNMI and DLR).

This document describes the theoretical basis of the algorithm which will be used.

## 1.2 Motivation

The methane ( $\text{CH}_4$ ) column retrieval algorithm for S5P (which makes use of the SWIR channel) is developed to work for cloud free conditions. Errors on the retrieved methane column of a few tenths of a percent can have large impact on source/sink estimations from inverse transport modelling. Therefore, an accurate cloud screening procedure is of utmost importance for methane retrieval from S5P. The baseline is to use NPP-VIIRS data for cloud screening, because its spatial resolution and spectral bands provide superior cloud screening than would be possible with S5P itself. Other level 2 (L2) products may also benefit from the cloud information provided by VIIRS, however currently the only strict requirement for this information in the L2 processor is for methane. In the development of this product, the requirements of the methane L2 processor are therefore prioritised.

## 1.3 High-level requirements

- The S5P-NPP Cloud product will contain cloud information relevant to each TROPOMI field-of-view (FOV) derived from observations made by the VIIRS instrument on NPP.
- It is expected that the product will contain the following information:
  1. A statistical summary for each S5P FOV of the NPP-VIIRS L2 Cloud Mask (VCM). As described in section 5, the VCM classifies each VIIRS pixel into one of four categories: confidently cloudy, probably cloudy, probably clear, or confidently clear. For each S5P FOV, the number of VIIRS pixels assigned to each category will be reported. From this the fraction in each class can be determined.
  2. The mean and standard deviation of the VIIRS measured level 1 (L1) reflectances in the 1.385 micron channel.

Statistics for additional conditions defined in the VCM (e.g. thin cirrus, cloud shadow, aerosol flag, cloud adjacency) could also be useful (but are not currently planned to be included in the product).

- Results are primarily required for the FOV corresponding to each reported L1 observation (i.e. measured spectrum) of the SWIR and NIR bands. It is however planned to code the processor flexibly so that results for all S5P bands which have a distinct geolocation reported at L1 could be included if required.
- For each geolocated L1 observation, in each relevant band, a nominal FOV is defined which best represents the spatial response (see below). S5P-NPP Cloud quantities will be evaluated for the nominal FOV and for additional FOVs in which both dimensions of the nominal FOV are increased by factors of (e.g., tbc) 1.1, 1.5 and 2.0.
- The product will be produced in a format complying with the agreed standards for S5P L2 data.

- The S5P-NPP Cloud processor will be a stand-alone code, separate from the L2 processing codes for the other TROPOMI L2 products. There will therefore be a stand-alone set of documents developed for this product including this Algorithm Theoretical Baseline Document (ATBD), together with a Software Development Plan (SDP), Product Assurance plan (PAP), Science Verification Plan (ScVP) and report (ScVR), Input/Output Definition Document (IODD), Software Requirements Document (SRD), Software User Manual (SUM) and Software Release Note(s) (SRN).
- Common elements (functions, interfaces etc) between the various L2 processors will of course be exploited to the fullest extent possible to ensure consistency of operation and minimise duplicated effort.

It is also identified that it may be of additional benefit to implement a dedicated retrieval of cirrus optical thickness based on the VIIRS 1.38 micron channel, for situations with thin cirrus (optical thickness < 0.20). Such a product can potentially be used to improve the baseline methane retrieval algorithm. Here, it would be important that the algorithm uses the same assumptions as the methane retrieval algorithm (e.g. same phase function, cirrus height, layer thickness). Work to assess the potential for developing this dedicated retrieval product is to be carried out. At present it is not baselined for the S5P-NPP Cloud processor and is not considered further in this ATBD.

#### 1.4 Basic definitions and assumptions

- In this document we use the term FOV to denote a quadrilateral area on the ground, considered to approximately correspond to the area within which light is collected in a single observation. This is distinguished from the more complete description provided by the spatial response function (SRF) which describes the relative response of the instrument to light originating from a given point on the ground. We also distinguish the FOV from the quadrilateral area on the ground reported in the S5P L1 geolocation records, which is obtained by interpolating between the centres of adjacent observations. This area is referred to as the L1 ground pixel corner box (L1GPC).
- The FOV and L1GPC are assumed to have straight edges in Cartesian coordinates. I.e. the effects of Earth curvature on the edges is neglected (though Earth curvature and detailed effects of the instrument spatial response are fully accounted for in determining the corner positions).
- It is noted that the departure of the SRF from an ideal box-car function is greatest in the along-track direction because of the instantaneous FOV (IFOV) and the along-track motion of the satellite during the integration time are comparable (around 7 km). This will cause the nominal FOV of TROPOMI to overlap particularly in the along-track direction.
- The nominal FOV and L1GPC differ mostly in the along-track direction, particularly towards the edge of the swath.
- Each TROPOMI spectral pixel is assumed co-located within band, so that only one set of NPP derived values will be provided for each TROPOMI across/along-track sample.
- It is expected to produce a single S5P-NPP product file for each S5P L1B file, covering the same set of observations and with a one-to-one correspondence between the S5P-NPP data and the S5P L1B geolocation records.
- It is assumed that the TROPOMI L1B file will provide as a minimum the following geolocation information:
  - The geodetic latitude and longitude of the centre of the nominal FOV.
  - The geodetic latitudes and longitudes of the four corners of the L1GPC.
  - The geodetic latitude, longitude and height of the satellite at the time of each observation.
  - The acquisition time of each scan line (since a reference date fixed for the mission).
- The L1 geo-location of TROPOMI and VIIRS will be assumed to be correct for the purposes of generating this product. (This processor will not attempt to improve upon or correct the nominal co-location of the two sensors, though it may be used to diagnose issues - see below).
- The code will be assumed to run on a system which has access to the necessary input data, namely

- S5P file(s) containing the geo-location of all scenes (in all Bands). These may be L1B files, or similarly formatted files containing the necessary sub-set of L1B data. (The code will not itself estimate S5P geolocation via an orbit / instrument model.)
- VIIRS file(s) containing VCM data and the required L1 reflectances (and possibly brightness temperatures).
- The processor will run within the S5P PDGS. PDGS will acquire the necessary NPP VIIRS input files, prior to the NPP-Cloud processor being run.
- It is assumed that the VIIRS data has characteristics similar to that of the products currently archived at the NOAA Comprehensive Large Array-data Stewardship System (CLASS) archive [ER2].

## **1.5 Relationship to possible activities to assess / monitor the Geometric and Radiometric performance of TROPOMI**

It is recognised that the S5P-NPP cloud processor performs certain tasks which could strongly support commissioning and in-flight monitoring of S5P. In particular the consistency between the geolocation and radiometric calibration of S5P and VIIRS can be verified by comparing TROPOMI reflectances spectrally averaged over VIIRS channel spectral responses, with the corresponding VIIRS reflectances averaged over the TROPOMI ground pixel (e.g. see [RD1] for analogous work related to GOME-2 and AVHRR). The correlation between the averaged spectrometer and imager data can be used to check, monitor and optimise the relative geolocation between the two sensors. Comparison of reflectances also provides information on the relative radiometric calibration of the two sensors. Use of such information to verify / monitor S5P is beyond the scope of the processor to be developed within this project. However, provided this does not add significant complication to the processor, we will seek to develop the code so that it can also generate some of the information needed. In particular, the code should be able to generate averaged radiances for channels (other than 1.38 micron) which cover regions spectrally sampled by S5P in the UV/VIS/NIR and SWIR bands. It should also be capable of accepting geolocation offsets from nominal so that reflectance averages can be calculated for ground-scenes shifted along/across-track from the nominal (L1B) geolocation. In the following this capability is included by defining FOV relative to the nominal one in terms of ranges in a normalised across / along-track coordinate system. In this way both shifts and enlargements of the nominal FOV can be simulated by the same code.

## **1.6 Heritage**

The development of this product builds on similar work carried out at RAL ([RD1]) to provide cloud information for GOME-1 using co-located measurements by the ATSR-2 and for MetOp GOME-2 using observations of MetOp AVHRR. Analogous work has also been conducted to derive cloud information for OMI from MODIS ([RD2]).

## **1.7 Document overview**

Section 5 provides a description of relevant aspects of VIIRS.

Section 6 defines the algorithm to be used to generate the S5P-NPP product.

Section 7 considers the computational feasibility of the processor.

Section 8 discusses the validation of the product.

## 2 Applicable and reference documents

### 2.1 Applicable documents

- [AD1] Sentinel-5P Level 2 Processor Development Statement of Work.  
**source:** ESA/ESTEC; **ref:** S5P-SW-ESA-GS-053; **issue:** 1.
- [AD2] GMES Sentinel-5 Precursor – S5p System Requirement Document (SRD).  
**source:** ESA/ESTEC; **ref:** S5p-RS-ESA-SY-0002; **issue:** 4.1.
- [AD3] Algorithm theoretical basis document for TROPOMI L01b data processor.  
**source:** KNMI; **ref:** S5P-KNMI-L01B-0009-SD; **issue:** 1.0.0; **date:** 2013-03-27.
- [AD4] Sentinel-5 Precursor PDGS Detailed Design Document Processing.  
**source:** DLR; **ref:** S5P-PDGS-DLR-DES-3041; **issue:** 1.0draft; **date:** 2014-04-24.
- [AD5] Sentinel-5 Precursor PDGS NPP Auxiliary Data Provider ICD.  
**source:** DLR; **ref:** S5P-PDGS-DLR-ICD-3012; **issue:** 1.0; **date:** 2014-09-12.
- [AD6] Sentinel-5P Ground Segment Master ICD.  
**source:** ESA; **ref:** S5P-PL-ESA-SY-070; **issue:** 1r3; **date:** 2014-04-23.
- [AD7] Ground Segment Requirements Document.  
**source:** ESA; **ref:** S5P-RS-ESA-GS-092; **issue:** 1.0; **date:** 2013-02-18.

### 2.2 Reference documents

- [RD1] B.L. Latter; Cross comparison of GOME-2, AVHRR and AASTR Reflectance. *GSICS Quarterly Newsletter*; **5** (2011) (3). URL [http://www.star.nesdis.noaa.gov/smcd/GCC/documents/newsletter/GSICS\\_Quarterly\\_Vol5No3\\_2011.pdf](http://www.star.nesdis.noaa.gov/smcd/GCC/documents/newsletter/GSICS_Quarterly_Vol5No3_2011.pdf).
- [RD2] Stammes, P., M. Sneep *et al.*; Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation. *Journal of Geophysical Research*; **113** (2008) (D16S38).
- [RD3] Terms, definitions and abbreviations for TROPOMI L01b data processor.  
**source:** KNMI; **ref:** S5P-KNMI-L01B-0004-LI; **issue:** 1.0.0; **date:** 2011-05-18.
- [RD4] Terms and symbols in the TROPOMI Algorithm Team.  
**source:** KNMI; **ref:** SN-TROPOMI-KNMI-049; **date:** 2011-09-28.
- [RD5] TROPOMI Instrument and Performance Overview.  
**source:** ; **ref:** S5P-KNMI-L2-0010-RP; **issue:** 0.10.0; **date:** 2014-03-15.
- [RD6] B Gao and Y.J. Kaufman; Selection of the 1.375 micron MODIS channel for remote sensing of cirrus clouds and stratospheric aerosols from space. *J. Atm. Sciences*; **52** (1995) (23), 4231. URL [http://www.star.nesdis.noaa.gov/smcd/GCC/documents/newsletter/GSICS\\_Quarterly\\_Vol5No3\\_2011.pdf](http://www.star.nesdis.noaa.gov/smcd/GCC/documents/newsletter/GSICS_Quarterly_Vol5No3_2011.pdf).
- [RD7] Joint Polar Satellite System (JPSS) VIIRS Cloud Mask (VCM) Algorithm Theoretical Basis Document JPSS VIIRS CM ATBD.  
**source:** NASA; **ref:** 474-00033; **date:** 2011-07-31.
- [RD8] Keith D. Hutchison, B.D. Lisagerb and B. Hauss; The use of global synthetic data for pre-launch tuning of the VIIRS cloud mask algorithm. *International Journal of Remote Sensing*; **33** (2012) (5).
- [RD9] Keith D. Hutchison, R.L. Mahoney, E.F. Vermote *et al.*; A geometry-based approach for identifying cloud shadows in the VCM algorithm for NPOESS. *J. Atmospheric and Oceanic Technology*; **26** (2009), 1388.
- [RD10] World geodetic system 1984.  
**source:** The Defence Mapping Agency; **ref:** A-TR-8350.2 second edition; **date:** 1991.
- [RD11] Algorithm theoretical basis document for TROPOMI L01b data processor.  
**source:** KNMI; **ref:** S5P-KNMI-L01B-0009-SD; **issue:** 1.0.0; **date:** 2013-03-27.

- [RD12] Joint Polar Satellite System (JPSS) Common Data Format Control Book – External (CDFCB-X) Volume V – Metadata Formats.  
**source:** NASA; **ref:** 74-00001-05-B0124; **date:** 2013-12-4.
- [RD13] On the Conversion from Radiances to Equivalent Brightness Temperatures.  
**source:** Eumetsat; **ref:** ; **date:** 2005-02-25.
- [RD14] Joint Polar Satellite System (JPSS) Common Data Format Control Book - External (CDFCB-X) Volume III - SDR/TDR Formats.  
**source:** NASA; **ref:** 474-00001-03-B0123; **date:** 2012-08-23.
- [RD15] Algorithm theoretical basis document for the TROPOMI L01b data processor.  
**source:** KNMI; **ref:** S5P-KNMI-L01B-0009-SD; **date:** 2012-12-03.
- [RD16] Joint Polar Satellite System (JPSS) VIIRS Geolocation Algorithm Theoretical Basis Document (ATBD).  
**source:** NASA; **ref:** 474-00053; **date:** 2011-07-31.
- [RD17] S5P-NPP Cloud processor Software Validation Plan.  
**source:** RAL; **ref:** S5P-NPPC-RAL-SVP-0001; **issue:** 0.5.2; **date:** 2014-08-29.

### 2.3 Electronic references

- [ER1] URL <http://npp.gsfc.nasa.gov/science/documents.html>.
- [ER2] URL <http://www.class.noaa.gov>.
- [ER3] URL <ftp://ftp-npp.class.ngdc.noaa.gov>.
- [ER4] URL <http://www.spec.org/cpu2006/results/res2007q3/cpu2006-20070806-01685.html>.

### 3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L0-1b data processor are described in [RD3]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L2 data processors are described in [RD4]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

#### 3.1 Terms and definitions

S5P-NPP processor	The processor which implements the algorithm defined in the ATBD to generate the S5P-NPP cloud product.
Defined constant	Parameter expected to remain constant and whose value may be explicitly set in the S5P-NPP code
Job order	File used by PDGS to control execution of the S5P-NPP processor
SRF file	Auxiliary file used by the S5P-NPP processor, which contains information related to the S5P spatial response function (SRF) which is calculated off-line.
L1GPC	Quadrilateral defined area associated with each S5P observation, defined by the ground pixel corners stored in the S5P L1B geolocation record
Nominal FOV	Quadrilateral area associated with each S5P observation, representative of the area on the ground from which light is collected.
Normalised FOV coordinates	Coordinate system on the ground with origin at the centre of the S5P pixel and scaled such that the corners of the nominal FOV have values of $y = \pm 1$ and $z = \pm 1$



## 3.2 Acronyms and Abbreviations

ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
BIRA-IASB	Belgian Institute for Space Aeronomy
CLASS	NOAA's Comprehensive Large Array-data Stewardship System
DLR	Deutsches Zentrum fur Luft-und Raumfahrt
EO	Earth Observation
ESA	European Space Agency
FOV	Field-of-view. In this document this refers to a quadrilateral area on the ground considered representative of spatial sensitivity of an individual S5P observation
GB	Gigabyte
GMES	Global Monitoring for Environment and Security
GOME	Global Ozone Monitoring Experiment (on ERS-2)
GOME-2	Second Global Ozone Monitoring Experiment (on MetOp)
IDL	Interactive Data Language
IODD	Input/Output Data Definition document
IUP	Institute of Environmental Physics, University of Bremen
JPSS	Joint Polar Satellite System
L0	Level 0 Product
L1	Level 1 Product
L1GPC	Level 1 ground-pixel corner box: The quadrilateral area on the ground corresponding to a single S5P observation, as recorded in the S5P L1B files.
L1B	Calibrated, geolocated level 1 Product
L2	Level 2 Product
LUT	Look-up-table
KNMI	Koninklijk Nederlands Meteorologisch Instituut
MB	Megabyte
NIR	Near infra-red
NOAA	National Oceanic and Atmospheric Administration
NPP	National Polar-orbiting Partnership
NRT	Near Real Time
PAP	Product Assurance plan
PDGS	Payload Data Ground Segment
S5P	Sentinel-5 Precursor
ScVP	Scientific Verification Plan
ScVR	Science Verification Report
SDP	Software Development Plan
SRD	Software Requirements Document
SRF	Spatial Response Function
SRN	Software Release Note
SUM	Software User Manual
SVP	Software Verification Plan
SoW	Statement Of Work
SW	Software
SWIR	Short wave infra-red
TIR	Thermal infra-red
TBC	To Be Confirmed
UV	Ultraviolet
VCM	VIIRS Cloud Mask
VIIRS	Visible Infrared Imaging Radiometer Suite

### 3.3 Algorithm Input / Output

Input parameters used by the algorithm are shown in table 1. Refer to section 3.1 for the definition of some of the terms used (particularly in the source column).

Output data is summarised in table 2.

Other parameters used internally by the algorithm are listed in table 3.

Description	Symbol	Unit	Source
Semi-major axis of the WGS'84 ellipsoid (=6378.1370 km)	$a$	km	Defined constant
Semi-minor axis of the WGS'84 ellipsoid (=6356.7523 km)	$b$	km	Defined constant
Eccentricity of the WGS'84 ellipsoid (=0.081819191)	$e$	-	Defined constant
Window of tolerance in time used to identify potentially relevant NPP files for a given S5P scan line (=100 s).	$\Delta t_{lw}$	s	Defined constant
Number of VIIRS cloud classifications (=4)	$N_C$	-	Defined constant
Number of defined S5P FOVs ( $\approx 4$ )	$N_F$	-	Job order
Number of VIIRS moderate resolution channels used in S5P-NPP product ( $\approx 3$ )	$N_M$	-	Job order
Maximum and minimum values of transformed coordinates $y, z$ , which define a specific S5P FOV.	$y_k^{min}, y_k^{max}, z_k^{min}, z_k^{max}$	-	Job order
Instantaneous spatial response function	$S_I$		On-ground calibration
Latitude of S5P pixel corner ( $l = 0, 1, 2$ or $3$ ), centre ( $l = C$ ) or sensor ( $l = S$ )	$\phi_{sl}$	$^{\circ}N$	S5P L1B
Longitude of S5P pixel corner ( $l = 0, 1, 2$ or $3$ ), centre ( $l = C$ ) or sensor ( $l = S$ )	$\phi_{sl}$	$^{\circ}N$	S5P L1B
Altitude of S5P sensor	$h_{SS}$	km	S5P L1B
View zenith angle of S5P pixel.	$\theta_s$	$^{\circ}$	S5P L1B
Sensing time of S5P scan line.	$t_{sl}$	s	S5P L1B
Sensing start time of S5P L1 file.	$t_{s0}$	s	S5P L1B
Sensing stop time of S5P L1 file.	$t_{s1}$	s	S5P L1B
Latitude of VIIRS pixel	$\phi_v$	$^{\circ}N$	VIIRS Geolocation
Longitude of VIIRS pixel	$\lambda_v$	$^{\circ}N$	VIIRS Geolocation
Line-of-sight zenith angle of a VIIRS pixel.	$\theta_v$	$^{\circ}$	VIIRS Geolocation
Sensing time of a VIIRS scan line.	$t_v$	s	VIIRS Geolocation
Sensing start time of NPP L1 file.	$t_{v0}$	s	VIIRS Geolocation
Sensing stop time of NPP L1 file.	$t_{v1}$	s	VIIRS Geolocation
Sensing start time of NPP L1 granule.	$t_{vg0}$	s	VIIRS Geolocation
Sensing stop time of NPP L1 granule.	$t_{vg1}$	s	VIIRS Geolocation
Sensing start centre-scan latitude of NPP L1 granule.	$\phi_{vg0}$	$^{\circ}$	VIIRS Geolocation
Sensing stop centre-scan latitude of NPP L1 granule.	$\phi_{vg1}$	$^{\circ}$	VIIRS Geolocation
Sensing start centre-scan longitude of NPP L1 granule.	$\lambda_{vg0}$	$^{\circ}$	VIIRS Geolocation
Sensing stop centre-scan longitude of NPP L1 granule.	$\lambda_{vg1}$	$^{\circ}$	VIIRS Geolocation
VIIRS pixel level geolocation data quality flag	$Q_G$	-	VIIRS Geolocation
VIIRS band averaged spectral radiance	$I_m$	$sr^{-1}$	VIIRS Radiance
VIIRS pixel level radiance data quality flag	$Q_m$	-	VIIRS Radiance
VIIRS cloud mask	$C_v$	-	VCM
VIIRS pixel level cloud mask quality flag	$Q_C$	-	VCM
Ratio of nominal FOV extent to that of L1GPC in across- and along-track directions	$f_y, f_z$	-	SRF file
Spatial response function (SRF), including effects of satellite motion	$S$	-	SRF file

**Table 1:** Algorithm input data.

## 4 TROPOMI instrument description

A description of the TROPOMI instrument and performance, referred to from all ATBDs, can be found in [RD5].

Description	Symbol	Unit	Data type	Number of values per pixel
Latitude of S5P pixel corner ( $l = 0, 1, 2$ or $3$ ), centre ( $l = C$ ) or sensor ( $l = S$ )	$\phi_{sl}$	°N	float	6
Longitude of S5P pixel corner ( $l = 0, 1, 2$ or $3$ ), centre ( $l = C$ ) or sensor ( $l = S$ )	$\lambda_{sl}$	°N	float	6
Altitude of S5P sensor	$h_{sS}$	km	float	1
Sensing time of S5P scan line.	$t_{si}$	s	float	1
View zenith angle of S5P pixel.	$\theta_s$	°	float	1
Number of VIIRS pixels with a valid radiance in a given moderate resolution bands, within a S5P FOV.	$M_m$	-	int	$N_M \times N_F$
Number of VIIRS pixels with a given cloud classification, within a S5P FOV.	$M_l$	-	int	$N_C \times N_F$
Mean radiance in a given VIIRS moderate resolution bands, considering all pixels within each S5P FOV.	$\bar{I}_m^{FOV}$	-	float	$N_M \times N_F$
Standard deviation of the radiances in a given VIIRS moderate resolution bands, considering all pixels within each S5P FOV.	$\Delta I_m^{FOV}$	-	float	$N_M \times N_F$
Mean radiance in a given VIIRS moderate resolution bands, weighted by the S5P SRF.	$\bar{I}_m^{SRF}$	-	float	$N_M$
Line-of-sight zenith angle of nearest VIIRS pixel to S5P pixel centre.	$\theta_v$	°	float	1
Time difference between S5P and VIIRS observation	$\Delta t_v$	s	float	1

**Table 2:** Algorithm output data.

## 5 VIIRS

VIIRS is an imaging instrument integral to the Joint Polar Satellite System (JPSS). One VIIRS instrument is currently flying on the Suomi-NPP platform which is in sun-synchronous orbit with 13:30 local time of ascending node crossing. It is planned to operate S5P such that the orbit tracks of the two platforms will be closely aligned, ensuring co-located observations of VIIRS and TROPOMI with a time difference of less than 5 minutes.

Detailed information on VIIRS and the products derived from its observations can be found at [ER1].

VIIRS observes radiation in bands in the UV, visible, NIR, SWIR and TIR spectral ranges, see table 4. It has five “imagery” (I) bands with a spatial resolution of  $\approx 375\text{m}$  at nadir and sixteen “moderate” resolution (M) bands with  $\approx 750\text{m}$  resolution. The instrument observes with contiguous spatial coverage over a 3000km wide swath.

Observations in each moderate resolution band are acquired using sixteen detectors, organised to sample along-track. Across-track sampling is obtained by scanning. Thus, in a single across-track scan 16 along-track samples are acquired and the scan rate is such that consecutive scans give effectively contiguous coverage at nadir. Towards the edges of the swath, the observing geometry is such that consecutive scans (consisting of 16 “rows” of data) overlap along-track (e.g. detector 16 of the first scan samples further along-track than detector 1 of the second scan). This is referred to as the “bow-tie” effect (see [RD16]). L1 data are provided in 2-d arrays, with rows corresponding to individual detectors. Geolocation information is provided for all these rows, however radiance values in the over-lapping rows are not reported, but filled with a “missing value” indicator. This results in all valid data in the 2-d arrays being ordered progressively along-track.

The inherent across-track sampling and resolution of the moderate resolution bands is relatively fine, such that L1 data is provided with variable co-adding of across track scenes to maintain a much more consistent spatial resolution across track. At the centre of the swath reported radiances correspond to an approximately square 750 x 750m scene, while at the edge of the swath, the across track resolution is degraded to around 1.6km. At the sampling at which L1B is reported, there are 3200 across-track moderate resolution pixels.

The 1.385 micron channel (channel M9) is particularly sensitive to cirrus cloud as water vapour strongly absorbs in this spectral range, such that under most atmospheric conditions there is negligible transmission of light to the surface [RD6]. Hence the signal is dominated by scattering from cloud or aerosol layers which lie above the surface<sup>1</sup>. The presence of thin cirrus in scenes otherwise classified as clear may well be identifiable by inspecting the signal in this channel. Because of the importance of detecting thin cirrus in order to accurately retrieve CH<sub>4</sub> from the S5P SWIR band, the VIIRS M9 channel radiances are to be summarised for the S5P FOV in the S5P-NPP-Cloud product.

<sup>1</sup> The VIIRS M9 band has a width of only 15 nm, a factor two finer than that of the corresponding MODIS band, making the VIIRS sensitivity to cirrus more distinct than is the case for MODIS.

Description	Symbol	Unit	Reference
East-West radius of curvature of the Earth.	$R_{EW}$	km	Equation 1
Position vector in Earth Centre frame of S5P pixel corner ( $l = 0, 1, 2$ or $3$ ), centre ( $l = C$ ) or sensor ( $l = S$ )	$\vec{p}_{s_l}^{EC}$	km	Equation 1
Earth centred cartesian coordinates of S5P pixel corners.	$y_{s_l}^{EC}, z_{s_l}^{EC}$	km	Equation 1
Position vectors in frame with origin at S5P pixel centre, oriented with respect to North	$\vec{p}_{s_l}^{L10}$	km	Equation 7
S5P cartesian coordinates in frame with origin at S5P pixel centre, oriented with respect to North	$y_{s_l}^{L10}, z_{s_l}^{L10}$	km	Equation 7
Cartesian coordinates in frame with origin at S5P pixel centre, rotated so $z$ is in along-track direction	$y_{s_l}^{L1z}, z_{s_l}^{L1z}$	km	Equation 8
Mid point of S5P L1GPC in along- and across-track directions	$\vec{y}_m^{L1z}, \vec{z}_m^{L1z}$	km	Equation 11,14
Width of S5P L1GPC in along- and across-track directions	$\vec{y}_w^{L1z}, \vec{z}_w^{L1z}$	km	Equation 10,13
Width of S5P nominal FOV (and SRF) in along- and across-track directions	$\Delta y_{FOV}, \Delta z_{FOV}$	km	Equation 22,22
Position vector in Earth Centre frame of VIIRS pixels which may fall within a given S5P FOV.	$\vec{p}_v^{EC}$	km	Equation 1
VIIRS cartesian coordinates in frame with origin at S5P pixel centre	$y_v^{L10}, z_v^{L10}$	km	Equation 7
VIIRS cartesian coordinates in frame with origin at S5P pixel centre, rotated so $z$ is in along-track direction	$y_v^{L1z}, z_v^{L1z}$	km	Equation 8
VIIRS normalised L1GPC coordinates	$y_v^{L1n}, z_v^{L1n}$	-	Equation 12, 15
VIIRS Normalised FOV coordinates	$y_v, z_v$	km	Equation 16, 16
Angle to rotate S5P pixel centre frame (L10) such that new $z$ coordinate aligned with along-track direction (L1z)	$\theta_x$	°	Equation 8
Off-bore-sight angles in across and along-track directions	$\theta_y, \theta_z$	°	Figure 12
Along-track distance travelled during integration of S5P observation	$\Delta z_{L1}$	km	Equation 19
L1GPC across-track extent in distance on the ground	$\Delta y_{L1}$	km	Equation 22
L1GPC across-track extent in off-nadir angle from the sensor	$\Delta \theta_y$	°	Equation 20
Distance from sensor to L1GPC centre on the ground	$d_C$	km	Equation 18
SRF weighted average of $I(y, z)$ for specific scene, $i$ .	$I_i$	$sr^{-1}$	Equation 23
Mean VIIRS radiance for S5P scene $i$ in SRF grid cell $j, k$ .	$I_{i,jk}$	-	Equation 26
Coverage of SRF by VIIRS observations for S5P scene $i$ .	$C_i$	-	Equation 27
Gridded, normalised SRF value, for S5P scene $i$ and SRF grid cell $j, k$ .	$s_{i,jk}$	-	Equation 25
Time difference between node crossing of S5P and NPP sub-satellite point	$\Delta t_{sat}$	s	Section 6.6
Window of tolerance in time used to identify potentially relevant NPP files for a given S5P L1 file.	$\Delta t_{fw}$	s	Section 6.6
Index to scan line within S5P L1 file	$i$	-	Section 6.6
Index to particular definition of S5P field of view.	$k$	-	Section 6.6
Total number of S5P observations in an orbit	$N_s$	-	Section 6.6
Index to across-track pixel within S5P scan line	$j$	-	Section 6.6
Index to a VIIRS moderate resolution channel	$m$	-	Section 6.6
Index to a VIIRS cloud classification	$l$	-	Section 6.6

**Table 3:** Parameters used internally within the algorithm.

The VIIRS Cloud Mask (VCM) is described in [RD7]. It determines whether each moderate resolution scene is obstructed by a cloud and assigns each a classification of confidently cloudy, probably cloudy, probably clear, or confidently clear. These assignments are based on applying a series of tests, to observations in bands M1,4,5,7,9,10,11,12,13,14,15 and 16, as well as I1,2,4 and 5, which each generate a probability of whether a scene is cloudy or not [RD8]. The individual test results are combined to generate the cloud-confidence classification. This includes a threshold test on the M9 band radiance. A separate (day-time) thin cirrus test is performed with a lower threshold for the M9 test, which is set to be more conservative (i.e. provide a margin of safety for applications which are particularly sensitive to thin cirrus). This test does not affect the assigned cloud confidence (but is reported elsewhere in the mask product).

The mask also contains a flag indicating the presence of cloud shadows which may well also be important for S5P retrievals. The potential to include statistics on cloud shadows based on this flag will be considered in future. The algorithm described below does not exclude the possibility of adding this information to the product if required. It is noted that the shadow flag has some limitations (particularly at high latitude) e.g. see [RD9].

The precise form of the data as it will be made available for this processor is assumed to be that currently

Spectral Range	Band name	Centre wavelength microns	FOV along x across-track / km		Driving Application
			@ nadir	@ swath edge	
Visible	M1	0.412	0.742 x 0.259	1.60 x 1.58	Ocean Color / Cloud detection over desert
	M2	0.445	0.742 x 0.259	1.60 x 1.58	Ocean Color
	M3	0.488	0.742 x 0.259	1.60 x 1.58	Ocean Color
	M4	0.555	0.742 x 0.259	1.60 x 1.58	Ocean Color / Snow/ice detection
	I1	0.640	0.371 x 0.387	0.80 x 0.789	Imagery / Cloud edge detection
NIR	M5	0.672	0.742 x 0.259	1.60 x 1.58	Ocean Color
	M6	0.746	0.742 x 0.776	1.60 x 1.58	Atmospheric Corr.
	I2	0.865	0.371 x 0.387	0.80 x 0.789	NDVI
	M7	0.865	0.742 x 0.259	1.60 x 1.58	Ocean Color
	DNB	0.7	0.742 x 0.742	0.742 x 0.742	Imagery
SWIR	M8	1.24	0.742 x 0.776	1.60 x 1.58	Cloud Particle Size
	M9	1.378	0.742 x 0.776	1.60 x 1.58	Cirrus/Cloud Cover
	I3	1.61	0.371 x 0.387	0.80 x 0.789	Binary Snow Map
	M10	1.61	0.742 x 0.776	1.60 x 1.58	Snow/cloud identification/ cloud phase
	M11	2.25	0.742 x 0.776	1.60 x 1.58	Clouds vs aerosol
	I4	3.74	0.371 x 0.387	0.80 x 0.789	Imagery Clouds
	M12	3.70	0.742 x 0.776	1.60 x 1.58	Snow/ice/cloud detection
	M13	4.05	0.742 x 0.259	1.60 x 1.58	SST Fires
TIR	M14	8.55	0.742 x 0.776	1.60 x 1.58	Cloud Top Properties
	M15	10.763	0.742 x 0.776	1.60 x 1.58	Thin cirrus/cloud detection (night)
	I5	11.450	0.371 x 0.387	0.80 x 0.789	Cloud Imagery
	M16	12.013	0.742 x 0.776	1.60 x 1.58	Thin cirrus/cloud detection (night)

**Table 4:** VIIRS spectral bands.

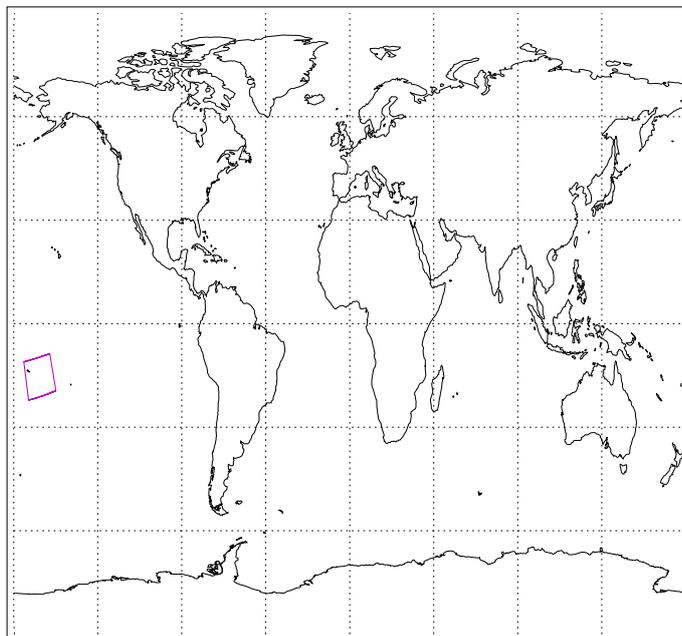
available at CLASS. VIIRS data on CLASS is available in sections of 3072 along-track scan lines by 3200 across. This corresponds to a orbit section of approximately 340s or 2400km in length. Each file typically contains data from four VIIRS *granules*, each containing 85 s of data. The CLASS VIIRS moderate resolution L1 files and VCM files cover identical scenes, with a one-to-one correspondence of pixel data within, so at this stage it is does not appear necessary to separately identify VCM and L1 files or separately geolocate the relevant values within them.

This implies separate geolocation information should not be needed for both VCM and L1 radiances (a significant reduction in the volume of data required).

It also implies a certain convenience in handling both files together in the S5P-NPP processor, generating at the same time the S5P FOV summarised quantities for both channel radiance(s) and cloud mask, rather than running two processes to generate information from L1 and VCM separately.

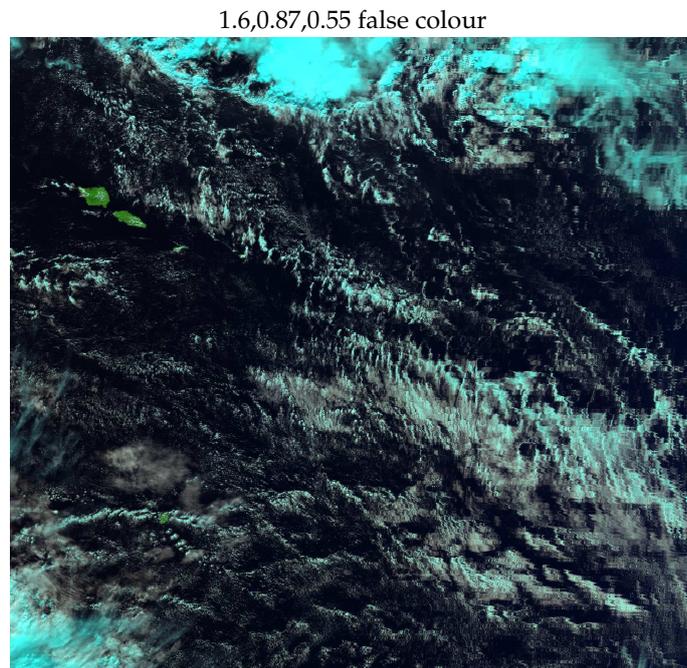
It is also noted that on CLASS, two sets of geolocation information is available, one based on the intersection of the line-of-sight with the assumed Earth ellipsoid and a second set which is based on the intersection with an assumed terrain height model. The ellipsoid intersection is needed for this project as this is expected to correspond precisely to the geolocation provided in the S5P L1B files.

VIIRS data for an example scene (over Samoa on 27 July 2012) are shown in figures 1 to 7. Figure 3 (note logarithmic scale) clearly shows the capability of the M9 channel to observe thin ice cloud which is very difficult to detect by other means. Structures which are presumably due to thin ice cloud are observed at reflectance levels between approximately 0.001 and 0.02. These structure are observed in pixels flagged as confidently clear by the VCM (5) and even where thin cirrus is not identified by the dedicated flag (figure 6). Note also that the islands and low altitude clouds are not visible in the M9 image (due to the absorption by water vapour). It is also noted that noise and striping is evident in the M9 channel around the 0.001 reflectance level.

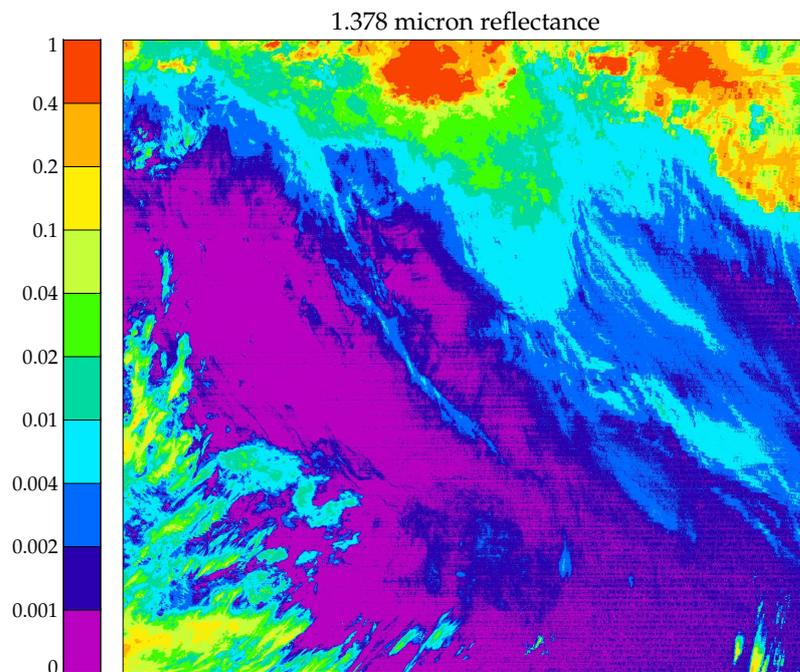


16:19:09/11/12 GMODO01020304050607080910111213141516\_npp\_d20120723\_t0057091\_e0102495\_b03811\_c20120724090240309197\_noaa\_ops.h5

**Figure 1:** Location of example VIIRS scene

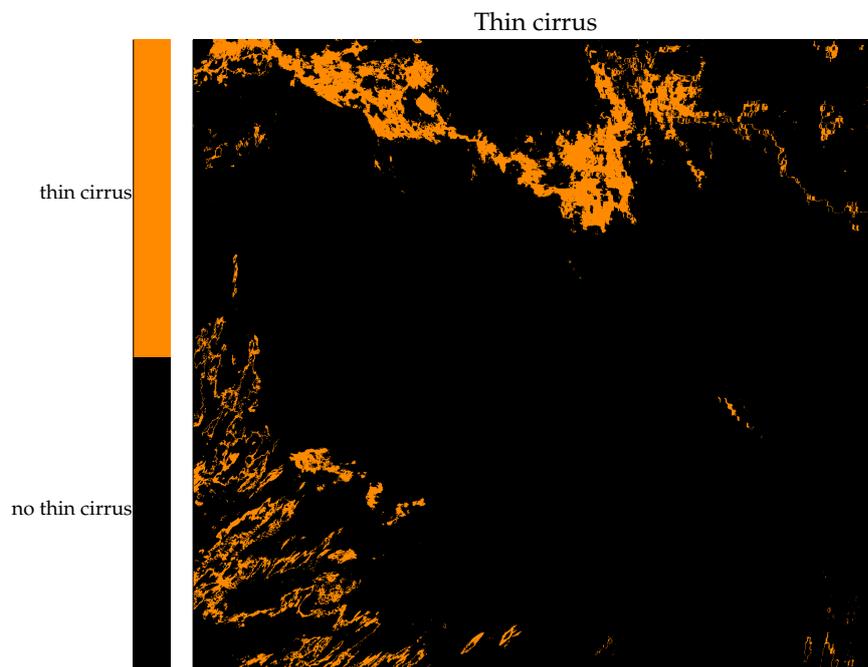


**Figure 2:** VIIRS false colour index for example scene

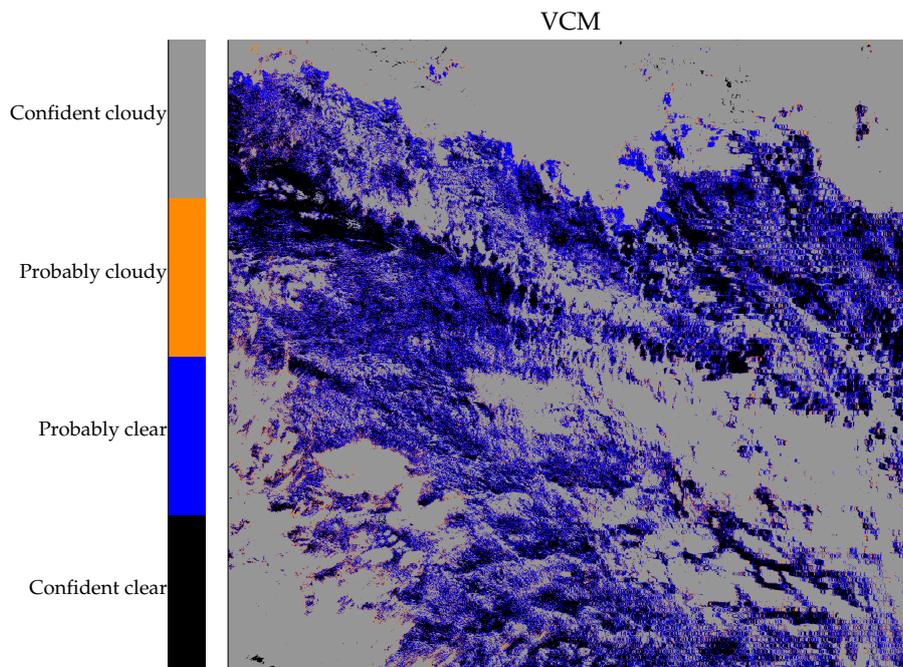


**Figure 3:** VIIRS M9 band (1.385 micron) reflectance for example scene

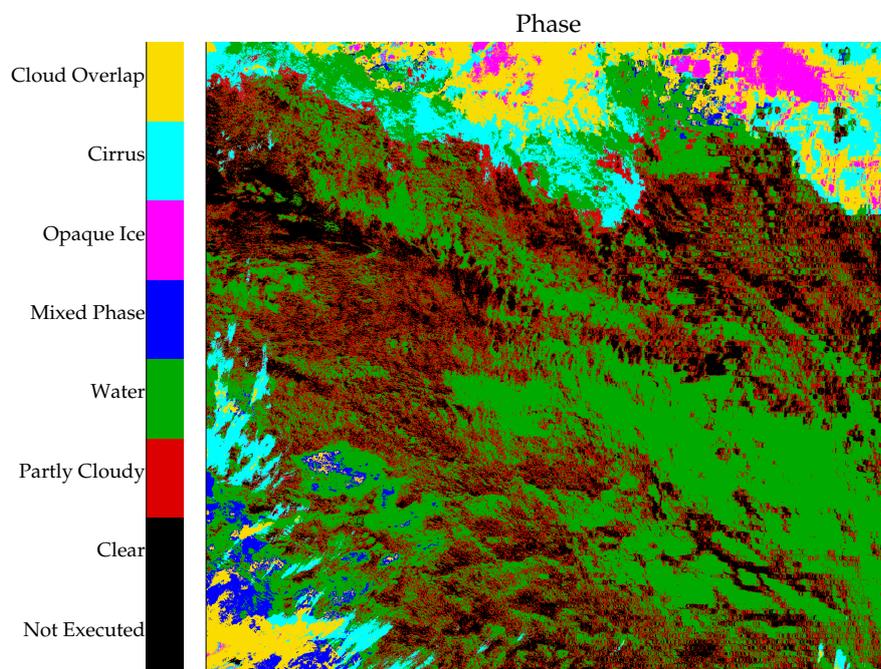
**Figure 4:** VIIRS M15 band (10.7 micron) brightness temperature (BT / K) for example scene



**Figure 5:** VIIRS cloud mask for example scene



**Figure 6:** VIIRS thin cirrus index for example scene



**Figure 7:** VIIRS cloud phase index for example scene

## 6 Algorithm description

### 6.1 Overview

S5P-NPP files will be generated for each S5P L1 file. The files will contain records summarising required parameters from the VIIRS L1 and VCM files, for each required S5P band (for which separate geolocation information is provided at L1), and for several definitions (scaled versions) of the S5P FOV.

The basis for identifying which VIIRS pixels fall within a S5P FOV will be to transform the latitude and longitude of the VIIRS pixels into the cartesian coordinate system  $y$  (across-track),  $z$  (along-track), with origin at the centre of the S5P pixel and scaled such that the corners of the nominal FOV have values of  $y = \pm 1$  and  $z = \pm 1$ , as illustrated in figure 8. FOV which are scaled (or shifted) with respect to the nominal FOV can be defined simply in terms of different ranges of  $y$  and  $z$ . The required ranges are provided as input parameters to the processor.

As noted in section 1.4, two distinct quadrilateral areas related to each scene are of relevance:

- The corners of the ground pixels provided in the S5P L1B files which are based on interpolation of the centre coordinates of the individual ground scenes, accounting for any across-track binning of detector pixels and along-track motion during integration. This region is referred to here as the *L1 ground pixel corner box (L1GPC)*. It defines areas which tessellate neatly on the ground, suitable for plotting quantities on maps. However in the S5P SRF typically extends over a significantly larger area, particularly in the along-track direction, towards the edge of the swath (see below).
- The *nominal FOV* which is defined to better represent the area on the ground from which light is collected, and is more appropriate for determination of VIIRS statistics. There are a number of possible ways to define “best” in this context. The definition could be based on consideration of the integrated energy (IE) of the spatial response function (SRF). Here we chose to define this as the (approximately) rectangular area whose boundaries in the along and across-track directions correspond to the half-maximum points of the SRF integrated in across and along-track directions (respectively)<sup>2</sup>.

It is required that the S5P-NPP product report VIIRS cloud and radiance information corresponding to the nominal FOV as defined above (as well as scaled FOV defined relative to the nominal FOV). The transformation required to map VIIRS observations into the normalised FOV coordinate system is described in detail below.

### 6.2 Coordinate systems

In the S5P L1B files, geolocation is defined using the WGS’84 model [RD10]. This defines the earth as an ellipsoid of revolution (about the North-South axis of rotation), with semi-major axis  $a = 6378.1370$  km and semi-minor axis  $b = 6356.7523$  km). With reference to this ellipsoid the following terms are defined:

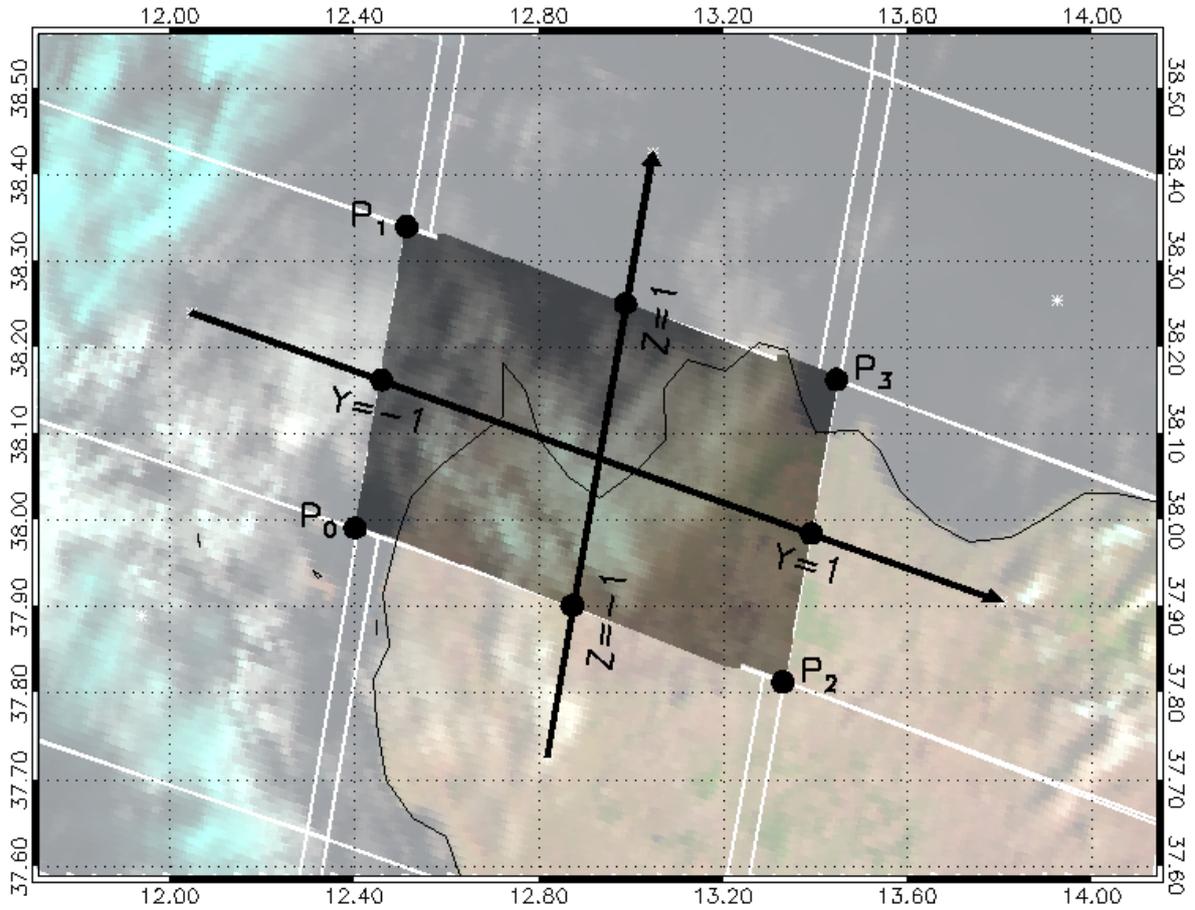
- Nadir-line: Normal-line through the ellipsoid surface to a point above it.
- Geodetic altitude,  $h$ : The distance between a point and the surface along the nadir.
- Geodetic latitude,  $\phi$ : Normal angle formed by the nadir-line and the equatorial plane.
- Geocentric longitude,  $\lambda$ : Angle on the equatorial plane between 0 longitude and the normal projection of the nadir line.

Points are geolocated at L1 in terms of their latitude, longitude and altitude, as defined above. However it is often more convenient to perform mathematical operations in terms of Cartesian coordinates as these avoid issues associated with singularities (poles) and discontinuities (longitude across the dateline).

It is conventional to adopt the following, Earth-centred frame (also illustrated in figure 9):

- $+X^{EC}$ : In equatorial plane pointing along 0 longitude.

<sup>2</sup> This choice is made because regions of fixed integrated energy are less straightforward to define over the range of SRFs which can be expected. If no across-track binning is applied and / or for nominal integration times the S5P SRF can be a relatively smooth function. In this case it would be natural to define the nominal FOV as the area enclosing a fixed integrated energy in the range of 50-70%. However, for scenes with relatively large across-track binning and along-track integration, the SRF tends to a box-car and such values of IE will lead to an area being defined which is smaller than the L1GPC (the L1GPC contains a much higher fraction of the IE). The adopted definition based on half-maxima points gives a more intuitively ‘sensible’ FOV in both cases.



**Figure 8:** Illustration of the coordinate system used to select VIIRS pixels within an S5P FOV. The example shows GOME-2 FOV and AVHRR imager data selected using the existing IDL code at RAL. White lines shows the borders of some GOME-2 FOV, which overlap in the across-track direction. The darkened image shows pixels selected for the central FOV by the algorithm. The transformed coordinate system used to select the pixels is indicated in black.

- $+Y^{EC}$ : In horizontal plane pointing along  $90^\circ$  longitude.
- $+Z^{EC}$ : Along axis of rotation pointing through north pole.

(Here the super-script  $^{EC}$  is used for consistency with notation used in section 6.3.) To convert from latitude, longitude and altitude to Earth-centred XYZ coordinates:

$$\begin{aligned}
 x^{EC} &= (R_{EW} + h) \cos(\phi) \cos(\lambda) \\
 y^{EC} &= (R_{EW} + h) \cos(\phi) \sin(\lambda) \\
 z^{EC} &= ((1 - e^2)R_{EW} + h) \sin(\phi)
 \end{aligned} \tag{1}$$

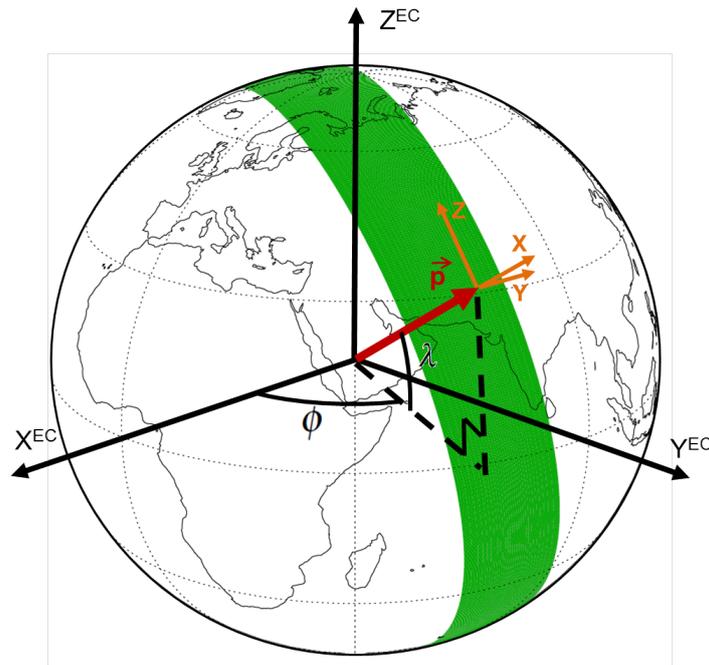
The following quantities are used below.

- Eccentricity of the ellipsoid:

$$e = \sqrt{\frac{a^2 - b^2}{a^2}} \tag{2}$$

- Local East-West radius of curvature at latitude  $\phi$ ,

$$R_{EW} = \frac{a}{\sqrt{1 - e^2 \sin^2(\phi)}} \tag{3}$$



**Figure 9:** Illustration of coordinate frames used here. An S5P swath is indicated in green, and the position vector of the centre of an individual S5P FOV in red. The final transformed coordinate frame is indicated in orange, which is obtained by the procedure described in section 6.3.

The following matrix operations can be used to rotate the coordinate system in which vector  $\vec{p}^{EC}$  is defined about the  $X^{EC}$ ,  $Y^{EC}$  and  $Z^{EC}$  axes by angles  $\theta_x$ ,  $\theta_y$ , and  $\theta_z$ , respectively.

$$\vec{p}'^{EC} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & \sin(\theta_x) \\ 0 & -\sin(\theta_x) & \cos(\theta_x) \end{pmatrix} \vec{p}^{EC} \quad (4)$$

$$\vec{p}'^{EC} = \begin{pmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{pmatrix} \vec{p}^{EC} \quad (5)$$

$$\vec{p}'^{EC} = \begin{pmatrix} \cos(\theta_z) & \sin(\theta_z) & 0 \\ -\sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{p}^{EC} \quad (6)$$

### 6.3 Transformation to normalised FOV coordinate system

The transformation to the normalised FOV coordinate system is performed by first using the 4 corner coordinates provided by the L1B geolocation, for each scene independently. The FOV may be any quadrilateral shape. In practice it should be close to rectangular and very close to trapezoidal with parallel sides in the along-track direction, but it is not necessary to make assumptions about the regularity of the FOV here.

This coordinate system is then scaled by factors  $f_y$  and  $f_z$  which are the ratios of the nominal FOV extent to that of the L1GPC in the across and along-track directions, respectively. These factors are modelled as function of across-track position, along-track motion during integration and spatial binning factor as described in section 6.4.

Once in this coordinate frame, it is trivial to identify which VIIRS points fall within the S5P nominal FOV. It is also trivial to identify points within scaled (or shifted) versions of the nominal FOV defined in terms of the normalised coordinates.

Different coordinate frames are denoted by superscripts as follows:

- *EC*: Earth centred cartesian coordinates.
- *L10*: Coordinates with origin at the L1GPC centre, with *Z* axis oriented towards local North.
- *L1z*: Coordinates with origin at the L1GPC centre, with *Z* axis oriented along-track.
- *L1n*: Normalised L1GPC Coordinates: As *L1z*, with axes scaled such that  $y^{L1n}$  and  $z^{L1n}$  are +/-1 at the boundaries of the L1GPC.

No superscript on  $y, z$  implies normalised FOV coordinates, such that  $y^{L1n}$  and  $z^{L1n}$  are +/-1 at the half maxima points of SRF.

The transformation is accomplished as follows:

1. The L1GPC latitude and longitudes are converted to Cartesian coordinates, applying equation 1 to the values provided in the L1B geolocation records. These 4 positions are stored in vectors  $\vec{p}_{s0}^{EC}, \vec{p}_{s1}^{EC}, \vec{p}_{s2}^{EC}, \vec{p}_{s3}^{EC}$ , where the index 0-3 refers to a specific pixel corner as defined in figure 8. The position vectors of the S5P pixel centre,  $\vec{p}_{sC}^{EC}$ , and the sensor itself,  $\vec{p}_s^{EC}$ , are also computed.
2. The frame of the corner coordinates is rotated to give vectors  $\vec{p}_{s_l}^{L10}$  (where index  $l = 0, 1, 2, 3$  as above), with origin at the S5P pixel centre by (a) rotating about *Z* by the S5P centre longitude and (b) about *Y* by the S5P centre latitude. This is achieved combining equations 6 and 5:

$$\vec{p}_{s_l}^{L10} = \begin{pmatrix} \cos(\lambda) & 0 & \sin(\lambda) \\ 0 & 1 & 0 \\ -\sin(\lambda) & 0 & \cos(\lambda) \end{pmatrix} \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{p}_{s_l}^{EC} \quad (7)$$

3. Coordinates are rotated about the new *X* axis (through the centre of the FOV) to give  $\vec{p}_s^{L1z}$  in which side P0-1 in figure 8 is aligned with the *Z* (along-track) axis. The appropriate angle is given by

$$\theta_x = \tan^{-1}((y_{s1}^{L10} - y_{s0}^{L10}) / (z_{s1}^{L10} - z_{s0}^{L10})) \quad (8)$$

Again, indices of the points relate to the pixel corners as shown in figure 8.

4. The latitudes and longitudes of VIIRS pixels are likewise rotated about *Z, Y* and *X* to give  $\vec{p}_v^{L1z}$ .
5. The centre,  $\vec{y}_m^{L1z}$  and width,  $\vec{y}_w^{L1z}$ , of the FOV in the *Y* (across-track) track dimension are determined at the *Z* (along-track) position of each VIIRS pixel:

$$\vec{y}_w^{L1z} = y_{s2}^{L1z} - y_{s0}^{L1z} + \frac{y_{s3}^{L1z} - y_{s2}^{L1z}}{z_{s3}^{L1z} - z_{s2}^{L1z}} (\vec{z}_v^{L1z} - z_{s2}^{L1z}) \quad (9)$$

$$\vec{y}_m^{L1z} = \frac{1}{2} \left( y_{s2}^{L1z} + y_{s0}^{L1z} + \frac{y_{s3}^{L1z} - y_{s2}^{L1z}}{z_{s3}^{L1z} - z_{s2}^{L1z}} (\vec{z}_v^{L1z} - z_{s2}^{L1z}) \right) \quad (10)$$

$$= y_{s0}^{L1z} + \frac{\vec{y}_w^{L1z}}{2} \quad (11)$$

(It should be the case, for a realistic FOV, that the along-track sides are parallel to each other in this frame. If this can be assumed then the equations above simplify.)

6. Final *Y* coordinates, are now obtained for all the relevant VIIRS pixels by

$$\vec{y}_v^{L1n} = 2 \frac{\vec{y}_v^{L1z} - \vec{y}_m^{L1z}}{\vec{y}_w^{L1z}} \quad (12)$$

$y_v^{L1n}$  will have values between -1 and +1 within the L1GPC.

7. The centre,  $\vec{z}_m^{L1z}$  and width,  $\vec{z}_w^{L1z}$ , of the FOV in the *Z* (along-track) dimension are now determined at the *Y* (across-track) position of each VIIRS pixel:

$$\vec{z}_w^{L1z} = z_{s1}^{L1z} - z_{s0}^{L1z} + \frac{z_{s3}^{L1z} - z_{s2}^{L1z} - z_{s1}^{L1z} + z_{s0}^{L1z}}{2} (\vec{y}_v + 1) \quad (13)$$

$$\vec{z}_m^{L1z} = \frac{1}{2} \left( z_{s1}^{L1z} + z_{s0}^{L1z} + \frac{z_{s3}^{L1z} + z_{s2}^{L1z} - z_{s1}^{L1z} - z_{s0}^{L1z}}{2} (\vec{y}_v + 1) \right) \quad (14)$$

8. The  $z$  coordinates normalised with respect to the L1GPC are obtained by

$$\vec{z}_v^{L1n} = 2 \frac{\vec{z}_v^{L1z} - \vec{z}_m^{L1z}}{\vec{z}_w^{L1z}} \quad (15)$$

9. Finally, the normalised L1GPC coordinates are converted to normalised FOV coordinates using the appropriate scale factors in the along and across-track directions:

$$\vec{y}_v = \frac{1}{f_y} \vec{y}_v^{L1n} \quad (16)$$

$$\vec{z}_v = \frac{1}{f_z} \vec{z}_v^{L1n} \quad (17)$$

$f_y$  and  $f_z$  are determined from LUTs as described in section 6.4 below. In practice,  $f_y$  can be assumed equal to 1 (see below).

To minimise computational expense, it is important to identify (in a fast, approximate manner) a conservative subset of VIIRS pixels to be considered for the coordinate transformation. In the current context observation time and view (or satellite) zenith angle are suitable parameters to use to select potentially relevant VIIRS pixels, as described in the following section.

## 6.4 Relationship between L1B ground pixel corners and the nominal FOV

As indicated above, the relationship between the extent of nominal FOV and that of the L1GPC is described by factors  $f_y$  and  $f_z$  (the ratios of the nominal FOV extent to that of the L1GPC measured in the across and along-track directions, respectively). These factors are obtained by modelling (outside the S5P-NPP processor) the SRF as a function of across-track position, along-track motion during integration and across-track spatial binning. Note that by far the largest discrepancy between L1GPC and nominal FOV occurs in the along-track direction as the L1GPC extent in this direction is defined mainly by the satellite motion during integration (which leads to a nearly constant along-track L1GPC extent over the swath) while the projection onto the ground of the instantaneous SRF increases greatly towards the edge of the swath. Since the width of the instantaneous SRF projected onto the ground at nadir is comparable to that of along-track motion ( $\approx 7$  km), the former dominates at the edge of the swath, such that the nominal FOV is approximately two times larger than the L1GPC in the along-track dimension at edge of swath. In the across-track direction, both the L1GPC and nominal FOV extents are directly related to viewing angle and so  $f_y$  is close to 1 and varies relatively little across the swath.

In order to estimate  $f_y$  and  $f_z$  for a particular scene it is necessary to know (i) the instantaneous SRF ( $S_I(\theta_y, \theta_z)$ ), as a function of viewing angle (relative to bore-sight) in the along and across-track directions,  $\theta_y$ ,  $\theta_z$  (see figure 12); (ii) the satellite (along-track) motion during integration,  $\Delta z_{L1}$ , which is assumed to be given by the L1GPC extent in the along-track direction; (iii) the spatial binning factor applied<sup>3</sup> or (equivalently) the across-track angular extent of the L1GPC, as seen from the satellite,  $\Delta\theta_y$ . (iv) the satellite position and viewing geometry of the bore-sight.

Here, a single instantaneous SRF is assumed per S5P band (spatial and spectral variation within the band are neglected). Given this instantaneous SRF, the full SRF projection onto the ground for a range of  $\Delta z_{L1}$ ,  $\Delta\theta_y$  and bore-sight viewing geometries can be computed. It is noted that the view geometry related effects on  $f_z$  are almost entirely determined by the distance from the satellite to the centre of the FOV/L1GPC on the ground,  $d_C$ . Variations in sensor altitude, viewing geometry, Earth radius / sphericity cause negligible variations in  $f_z$  (or  $f_y$ ) for a given value of  $d_C$ . Scene specific values of  $d_C$ ,  $\Delta z_{L1}$  and  $\Delta\theta_y$  can be computed from the L1B geolocation information as follows:

$$d_C = |\vec{p}_{sC}^{EC} - \vec{p}_{sS}^{EC}| \quad (18)$$

$$\Delta z_{L1} = (z_{s3}^{L1z} - z_{s2}^{L1z} + z_{s1}^{L1z} - z_{s0}^{L1z})/2 \quad (19)$$

<sup>3</sup> The number of across-track detector pixels averaged as described in [RD11]. Measurements in the SWIR band are not expected to be binned, however other channels will normally be binned by varying (integer) factors across-track so as to reduce variations in the across-track pixel size over the swath.

$$\Delta\theta_y = \cos^{-1} \frac{(\vec{p}_{s0}^{EC} - \vec{p}_{sS}^{EC}) \cdot (\vec{p}_{s2}^{EC} - \vec{p}_{sS}^{EC})}{|\vec{p}_{s0}^{EC} - \vec{p}_{sS}^{EC}| |\vec{p}_{s2}^{EC} - \vec{p}_{sS}^{EC}|} \quad (20)$$

Figure 10 illustrates the expected variation of  $f_z$  (top-left hand panel) and  $f_y$  (top-right hand panel) as a function of the distance to the L1 ground-pixel centre, together with the across-track pixel size on the ground (bottom left). Figure 11 shows the same information as a function of across-track distance from the nadir point.  $f_z$  is also shown (as different coloured lines) for varying values of the  $\Delta z_{L1}$ .  $f_y$  and the across-track pixel size are also shown as function of across-track bin factor, where a bin-factor of 1 here corresponds to  $\Delta\theta_y = 0.2375$  degrees ( 3.46 km). In these plots, bin-factor 2 corresponds to the nominal S5P across-track pixel size at nadir (for all bands<sup>4</sup>).

These curves are computed (by code which does not form part of the S5P-NPP processor) as follows:

- The instantaneous SRF,  $S_I(\theta_y, \theta_z)$ , is defined (*in lieu* of suitable ground-based data) as the convolution of a 2-D box-car function (corresponding to the projection of the across-track binned detector pixel extent onto the ground) and a 2-D Gaussian function (corresponding to the expected point-spread function of the telescope). These widths of these function are defined in off-nadir angle such that the dimensions projected onto the ground at nadir are 3.5 (across-track) x 7 km along-track for the box-car and 1.7 km (circular) for the Gaussian. This function is modelled on a regularly spaced grid of along and across-track off-bore-sight angles,  $\theta_y, \theta_z$ , as illustrated in figure 12.
- For a given off-nadir angle of the bore-sight, the instantaneous SRF is projected onto the ground, assuming a spherical Earth with radius of 6367.3818 km and a satellite altitude of 834.78068 km. It is also assumed that the bore-sight azimuth angle with respect to flight direction is 90 degrees.
- The pixel across-track extent which would be reported for the L1GPC,  $\Delta y_{L1}$ , is calculated by projecting onto the ground the view angles corresponding the appropriate detector pixel boundaries.
- The projected  $S_I(\theta_y, \theta_z)$  is interpolated onto a fixed grid in along and across-track distance relative to the pixel centre (the bore-sight projected onto the ground), to give  $S_I(y, z)$ . The values of the interpolated SRF are divided by the solid angle extent (at the sensor) associated with each grid point (in  $y, z$ ), such that the correct angular sensitivity of the sensor is maintained.
- The  $S_I(y, z)$  projected onto the ground is convolved in the along-track direction with a box-car function representing the assumed along-track motion,  $\Delta z_{L1}$ , to give  $S(y, z)$ .
- The full-width-half maxima  $\Delta z_{FOV}$  and  $\Delta y_{FOV}$  of  $S(y, z)$  are determined from integrals of  $S(y, z)$  in the across and along-track directions, respectively.
- The pixel extent ratios are then determined as follows:

$$f_y = \frac{\Delta y_{FOV}}{\Delta y_{L1}} \quad (21)$$

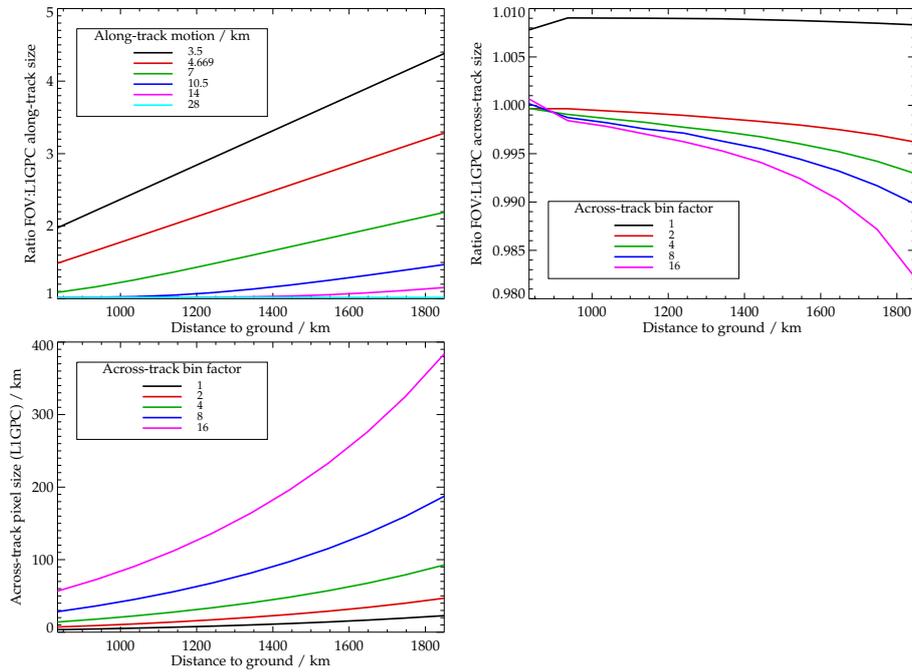
$$f_z = \frac{\Delta z_{FOV}}{\Delta z_{L1}} \quad (22)$$

From these calculations we conclude that  $f_y$  is sufficiently close to 1 to be assumed equal to 1 in the S5P-NPP processor. Variations in  $f_z$  need to be accounted for and depend almost exclusively on  $d_C$  and the along-track motion during integration. The S5P-NPP algorithm therefore makes use of look-up-tables (LUTs) containing values of  $f_z$  tabulated as a function of  $d_C$  and  $\Delta z_{L1}$ .

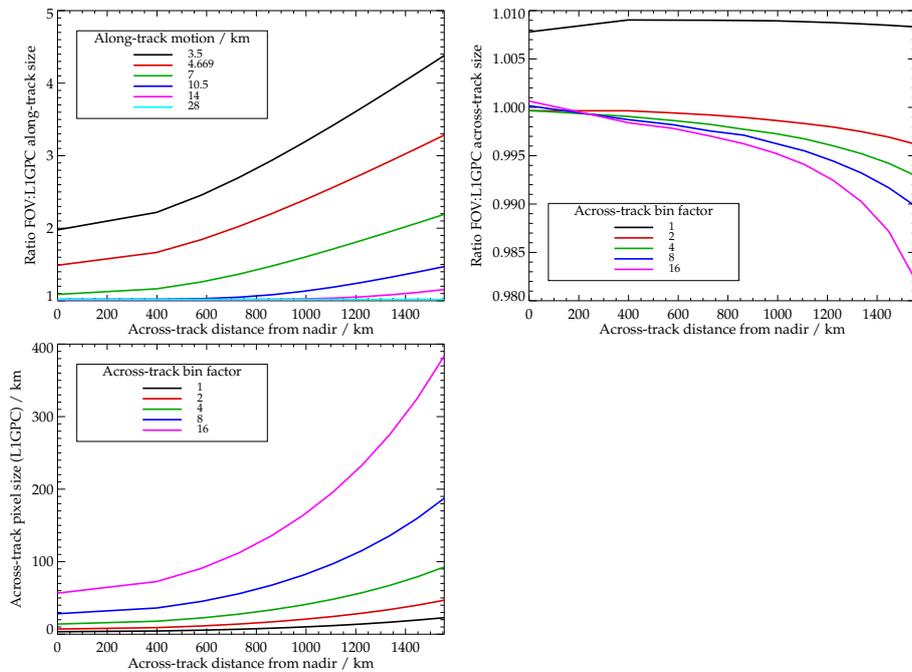
These LUTs are computed off-line (based on best available information on the S5P SRF for each band) and stored in an auxilliary file provided to the S5P processor. This file also contains a more detailed representation of the SRF as discussed further in section 6.5, below.

The S5P-NPP processor is required to read this auxilliary SRF file, and for each scene determine the ground pixel specific values of  $f_z$  by bi-linearly interpolating the LUTs to the appropriate values of  $d_C$  and  $\Delta z_{L1}$  which are determined using equations 18 and 19.

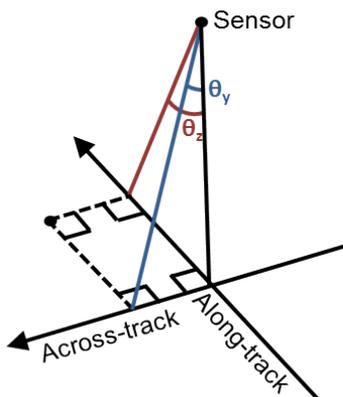
<sup>4</sup> This includes SWIR bands as the SWIR detector pixels are approximately two times larger than those of the other bands, such that although the SWIR bands are not expected to be binned, their nominal extent on the ground corresponds approximately to that of the other bands with bin-factor = 2.



**Figure 10:** Relationship between the along and across-track extent of the nominal S5P FOV and that of the L1 ground-pixel corner box (L1GPC), as a function of distance of the ground-pixel centre from the sensor.



**Figure 11:** Relationship between the along and across-track extent of the nominal S5P FOV and that of the L1 ground-pixel corner box (L1GPC), as a function of across-track position.



**Figure 12:** Off-nadir angles used to model S5P SRF

## 6.5 Calculation of spatial response function weighted quantities

The coordinate transformation described above enables VIIRS pixels within a defined quadrilateral FOV to be simply identified. It is then trivial to count or determine the simple mean and standard deviation of values associated with each of these pixels. For channel radiances (and possibly other quantities) it is desirable to compute the integral of the VIIRS observations weighted by the S5P spatial response function (SRF):

$$I_i = \int_y \int_z I(y, z) S_i(y, z) dz dy \quad (23)$$

Where

- $I_i$  is the SRF weighted average for S5P observation  $i$  of the continuously varying quantity  $I(y, z)$  (e.g. spectral radiance in a particular VIIRS band). Here index  $i$  is assumed to refer to a specific along/across-track ground scene from a specific band. The SRF is assumed not to vary spectrally within a band.
- $y, z$ , are the normalised along and across-track coordinates as defined above.
- $S$  is the SRF normalised such that:

$$\int_y \int_z S_i(y, z) dz dy = 1 \quad (24)$$

In practise the SRF will vary with band and detector pixel (spectral and across-track). It is not expected to vary significantly with time (at least not within an orbit).

Note the normalised coordinates follow the size of the quadrilateral FOV. Variations of the SRF defined in these normalised coordinates are therefore relatively small compared to the variation of the SRF in simple along/across-track distance.

Evaluating equation 23 using the discrete VIIRS samples is complicated by fact that VIIRS data does not have a perfectly regular along-track organisation, due to the “bow-tie” effect (see section 5). Sampling across-track is also somewhat complicated by the different co-adding factors used to maintain approximately constant spatial resolution (which means the spacing of samples does not vary smoothly). The problem is addressed as follows:

1. The S5P SRF (for a given scene  $i$ ) is represented on a regularly spaced grid in the normalised, transformed  $(y, z)$  space. This grid is defined by a range and sampling interval in each dimension. The values of the SRF in grid cell  $j$  (across-track index),  $k$  (along-track index),  $S_{i,jk}$ , is equal to the integral of the continuous SRF over the grid cell. The normalisation of the SRF leads to

$$\sum_j \sum_k S_{i,jk} = 1 \quad (25)$$

The grid sampling interval is chosen to correspond to around 1 km, to slightly oversample the VIIRS moderate imaging resolution, such that at nadir there will usually be one VIIRS observation within each cell.

2. Valid VIIRS data within each grid cell are identified (testing the  $\vec{y}_v, \vec{z}_v$  against the range of the cell).
3. The simple means of the valid VIIRS data within each SRF cell are determined, yielding an array of values  $I_{i,jk}$ .
4. Due to the chosen grid sampling, most cells will be filled under most circumstances, however some will not be filled even under standard observing conditions, particularly at the edge of the swath. These empty cells are filled by taking the average of the mean values in neighbouring cells which have been filled. Only the 4 directly neighbouring values along and across-track are considered (for speed), i.e. those in cells  $(j-1, k), (j+1, k), (j, k-1), (j, k+1)$  (not the diagonally adjacent cells).
5. The SRF weighted quantity is then evaluated from the filled, gridded mean values:

$$I_i = \sum_j \sum_k I_{i,jk} s_{i,jk} \delta_{jk} \quad (26)$$

Where  $\delta_{jk}$  is

- $\delta_{jk} = 1$  if there is a defined value of  $I_{i,jk}$  (including after filling based on the direct neighbours) or
- $\delta_{jk} = 0$  otherwise, i.e.  $I_{i,jk}$  is undefined because the regions of missing VIIRS data are too extensive to be filled just using neighbouring cells.

6. The coverage of the SRF by valid data (including the simple filling operation) is measured by the quantity

$$C_i = \sum_j \sum_k s_{i,jk} \delta_{jk} \quad (27)$$

By the definition of the SRF grid sampling,  $C_i = 1$  in most circumstances, unless there are specific reasons for the VIIRS data being invalid over an extensive area (e.g. at very low solar zenith angle).

The scene to scene variation in the gridded SRF (denoted by index  $i$ ) is characterised off-line and stored in the auxilliary SRF file as LUTs as a function of  $d_C, \Delta z_{L1}$  and  $\Delta \theta_y$ . I.e. For each value in the 2-D gridded SRF, a 3-D LUT is provided as a function of  $d_C, \Delta z_{L1}$  and  $\Delta \theta_y$ , which needs to be tri-linearly interpolate to the conditions appropriate to the given scene.

This process is illustrated in figures 13 and 14, for the extreme edge of the S5P swath and nadir observing conditions, respectively. The figures show the following:

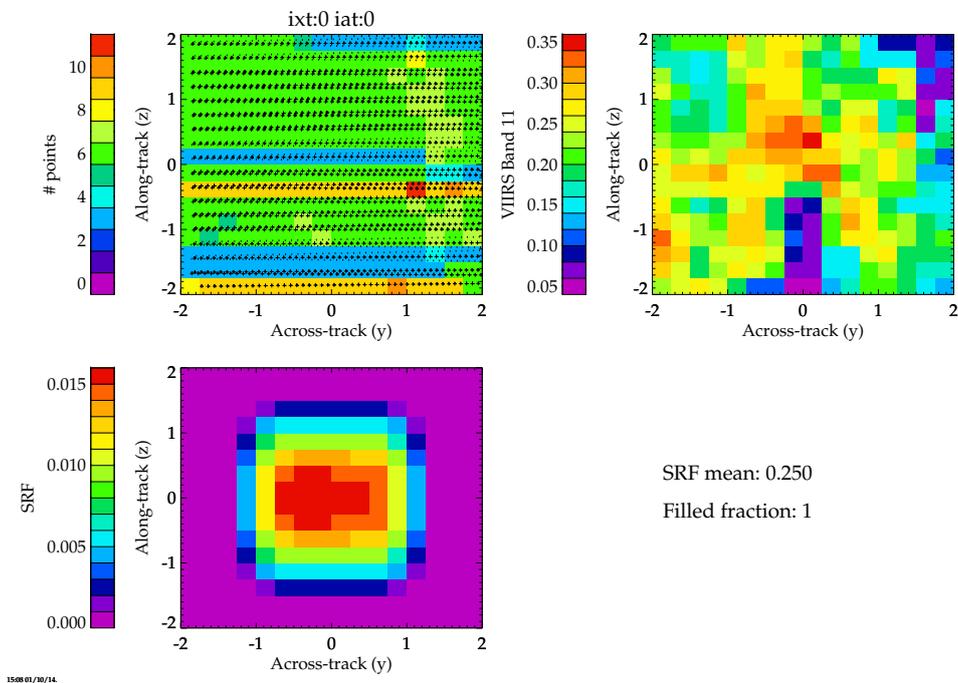
- Bottom left: An example gridded, normalised SRF. This particular example is based on the following:
  - The instantaneous SRF is assumed to be a box-car function of width 7 km on the ground (at nadir) in both directions.
  - This is convolved in both directions with a 1.7 km wide Gaussian function representing the telescope point spread function.
  - This is convolved in the along-track direction with a 7 km box-car representing the along-track motion during integration.

Results shown in these figures are obtained using simulated S5P L1B geolocation records, together with the consistently defined SRFs. The SRF grid spacing as assumed to be 0.25 in normalised FOV coordinates, which approximately corresponds to 1 km on the ground at nadir.

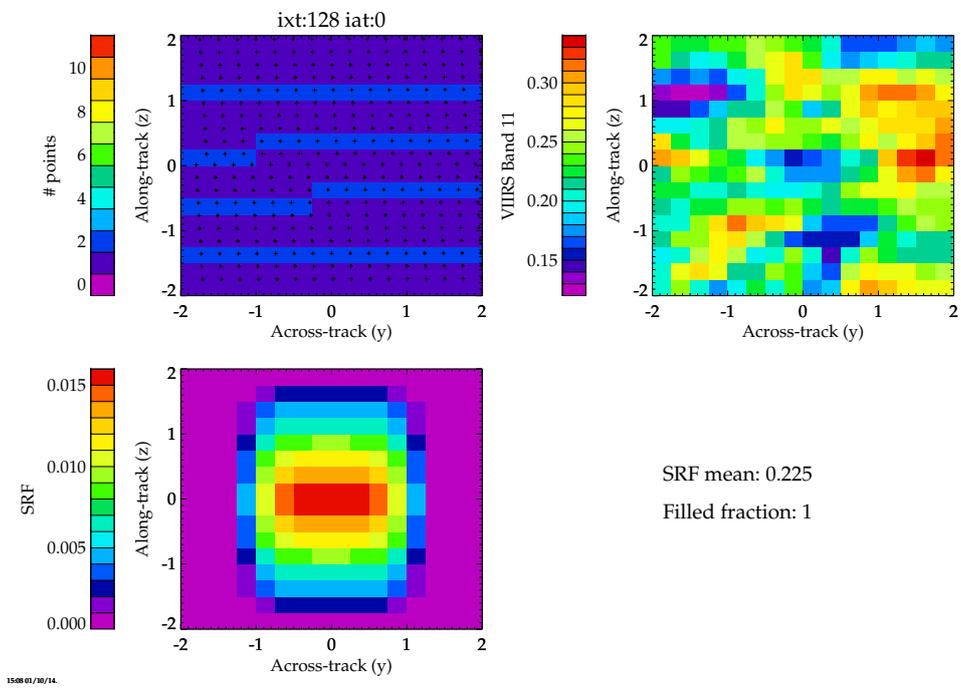
- Top left: The number of VIIRS pixels inside each of the SRF grid cells. Small crosses show locations of individual valid VIIRS pixels. Data is taken from real VIIRS measurements in band M11 at around 12:17 UT on 10 May 2013. The simulated S5P L1B file is generated to have timing and geolocation records which match the selected real VIIRS observations. At the edge of the swath (figure 13) there are several VIIRS observations across-track per SRF grid cell as the VIIRS across-track sampling remains only 1.6 km while the S5P across-track FOV on the ground is much larger than at nadir. There are also nearly duplicated scan lines of measurements arising from the bow-tie effect (successive read-outs of the along-track sampling detectors overlap).
- Top right: The mean of the valid observations (sun-normalised radiance in VIIRS band M11) in each SRF grid cell. Cells coloured in white have no valid VIIRS observations.

- The caption in the bottom right shows the SRF weighted mean radiance ( $I_i$ ) and the coverage of the SRF with valid values ( $C_i$ ).

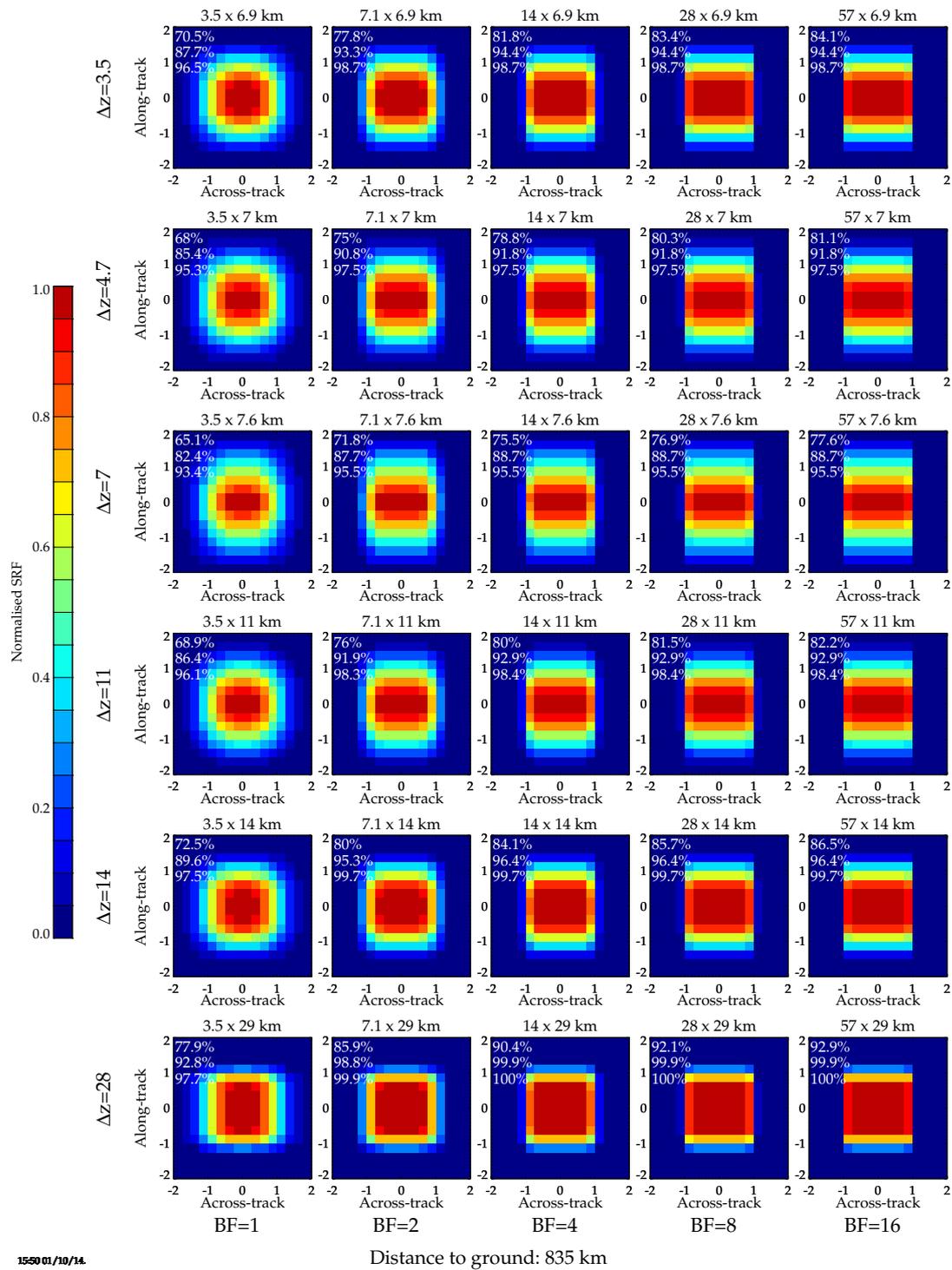
Figures 15 and 16 show modelled SRF functions (using the approach outline in section 6.4 above) for nadir and edge of swath scenes, respectively. Each figure shows the SRF (in normalised FOV coordinates) for a range of along-track motion (from top to bottom) and spatial bin factors (from left to right). Nominal conditions for a 7 x 7 km nadir pixel are shown in the 2nd column from the left of the 3rd row down in each of these plots. Note that at edge of swath (particularly for high bin factors) the SRF is weighted towards the nadir side. This is a consequence of the slant view coupled to the Earth curvature within the scene, which means that a given along-track distance on the ground corresponds to a larger solid angle subtended at the satellite on the nadir side of the scene compared to the outer edge. As stated, these SRF functions are shown in normalised FOV coordinates, so the area enclosed by the nominal FOV is that within +/-1 in both dimensions. The caption above each panel quotes the along/across-track extent of the nominal FOV in each case, together with the fraction of the integrated energy of the SRF which is enclosed by it.



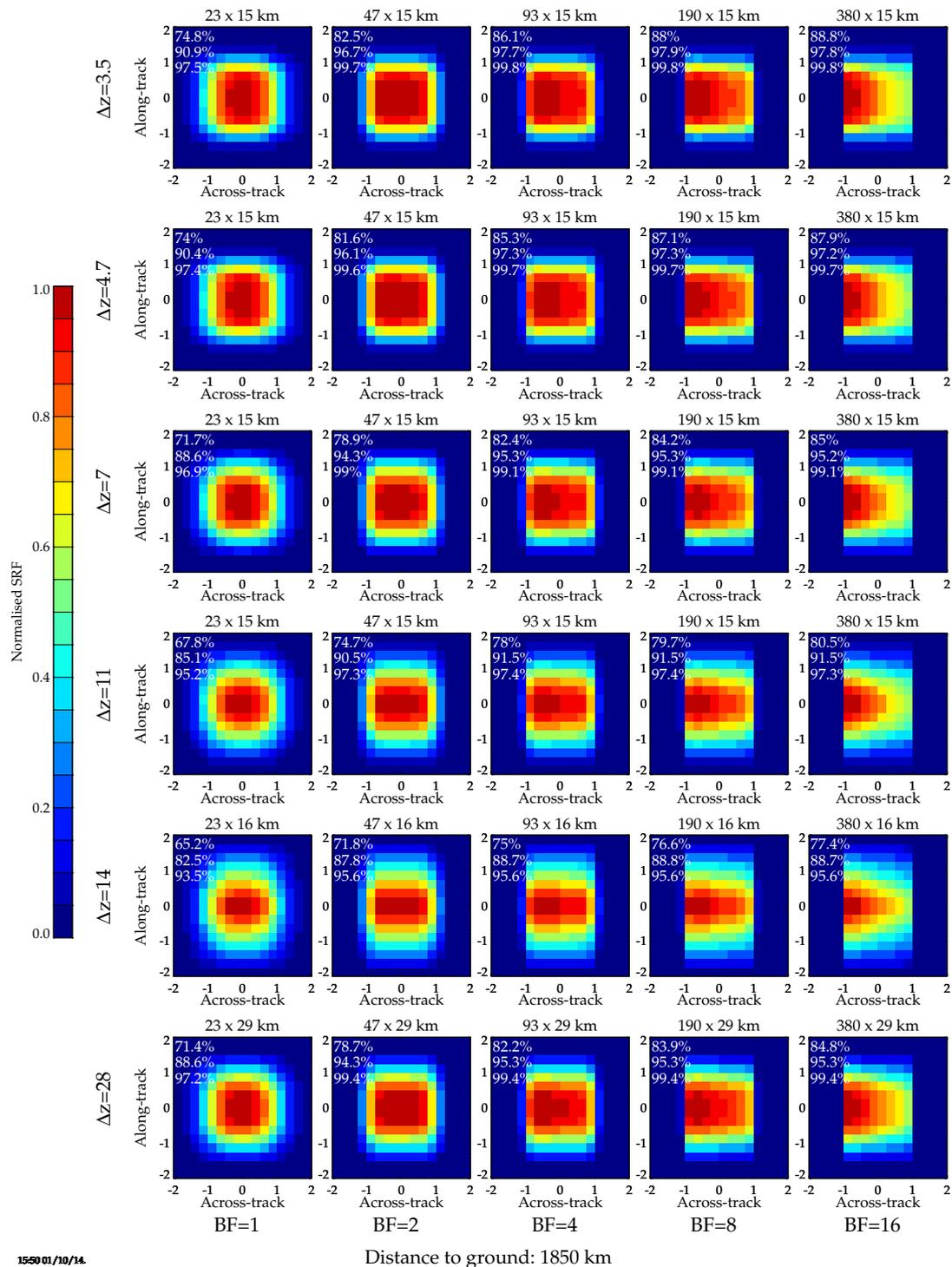
**Figure 13:** Illustration of VIIRS data integration over an S5P SRF. Extreme East of swath.



**Figure 14:** Illustration of VIIRS data integration over an S5P SRF. Centre of swath.



**Figure 15:** Illustration of gridded SRFs in normalised FOV coordinates for nadir view geometry, as function of pixel bin factor (left to right) and along-track motion (top to bottom). Caption above each panel shows the nominal FOV extent in km, and numbers in white indicate the fraction of the integrated energy contained within 1, 1.2 and 1.5 times the nominal FOV.



1550 01/10/14.

**Figure 16:** Illustration of gridded SRFs in normalised FOV coordinates for edge of swath view geometry, as function of pixel bin factor (left to right) and along-track motion (top to bottom). Caption above each panel shows the nominal FOV extent in km, and numbers in white indicate the fraction of the integrated energy contained within 1, 1.2 and 1.5 times the nominal FOV.

## 6.6 Step-by-step procedure for generating the S5P-NPP records

It is assumed that a list of all potentially relevant VIIRS files for a given S5P L1B file is given as input to the processor. Identification of relevant files will be based on time, assuming that the two orbits have similar ascending node crossing times within a tolerance of  $\pm\Delta t_{fw}$ . To cover all expected conditions we assume  $\Delta t_{fw}$  can be less than 20 minutes. Sensing start and stop times for S5P ( $t_{s0}, t_{s1}$ ) and VIIRS ( $t_{v0}, t_{v1}$ ) can be readily determined from the VIIRS file names and / or information in the file meta-data. Here we assume that all NPP files are identified for which  $t_{v0} > t_{s0} - \Delta t_{fw}$  and  $t_{v1} < t_{s1} + \Delta t_{fw}$ . All times are defined in seconds since a reference date and time which is fixed for the mission to simplify handling time ranges which span midnight.

The procedure to generate the S5P-NPP Cloud product is illustrated in the flow diagram shown in figure 17, and described in more detail below.

1. Ingest S5P L1 file containing the geolocation of each scene for which the S5P-NPP-Cloud information is required.
2. The time difference,  $\Delta t_{sat}$ , between S5P centre-swath observations and those of spatially co-located VIIRS measurements is needed to be known to an accuracy of  $\approx 1$  minute (see below).  $\Delta t_{sat}$  can be estimated to a sufficient level of accuracy (without reading all VIIRS geolocation information at this stage) using metadata in the VIIRS geolocation file (see [RD12]). Specifically, the records G\_RING\_LATITUDE, G\_RING\_LONGITUDE, N\_BEGINNING\_TIME\_IET and N\_ENDING\_TIME\_IET can be used to determine:
  - $\phi_{vg0}, \lambda_{vg0}, t_{vg0}$ : Latitude, longitude and time of the central across-track VIIRS pixel at the start of a given VIIRS granule.
  - $\phi_{vg1}, \lambda_{vg1}, t_{vg1}$ : Latitude, longitude and time of the central across-track VIIRS pixel at the end of a given VIIRS granule.

These values from all identified VIIRS files can be read and the closest S5P pixel in space to any of these locations identified<sup>5</sup>. The time difference between the closest matching pair of S5P/VIIRS observations is then taken to define  $\Delta t_{sat}$ . This approach should be robust to gaps in the coverage of S5P or VIIRS data, provided at least one of the identified VIIRS files has overlapping coverage with the S5P data. If this is not the case then the processor can produce no valid output in any case.

3. Loop over S5P scan lines. A particular scan line is referred to be index  $i$ .
  - (a) Identify VIIRS files which containing information relevant to these scan lines. These will be identified on the basis of  $t_{v0} > t_{si} + \Delta t_{sat} - \Delta t_{Iw}$  and  $t_{v1} < t_{si} + \Delta t_{sat} + \Delta t_{Iw}$ , where  $t_{si}$  is the time associated with the centre across-track pixel the given S5P scan line. Time buffer  $\Delta t_{Iw}$  will be set to  $\approx 100$ s) to ensure that at most two consecutive VIIRS files (each 5 minutes long) will be identified. Under these circumstances  $\Delta t_{sat}$  need only be known to a similar (but smaller) level of accuracy to ensure the correct files are indentified for a particular scan line (the time buffer will be large enough to accomodate such errors).
  - (b) Read and store in memory the relevant information from these VIIRS files<sup>6</sup>.
  - (c) Loop over S5P across-track pixels within the current scan line. Across track pixels are referred to with index  $j$ .
    - i. Calculate the Cartesian coordinates of the S5P FOV centre (using the centre latitude, longitude in the L1B geolocation record).
    - ii. Identify (in Cartesian coordinates), the nearest VIIRS pixel to the S5P centre. An iterative approach is adopted for this purpose to efficiently search the stored VIIRS data locations at progressively finer sampling, focusing on progressively smaller search regions as the search sample interval decreases.
    - iii. Within a defined (along, across-track ) region of VIIRS pixels about the identified pixel, apply the local coordinate transform which maps the latitude and longitude of VIIRS pixels so that the current nominal S5P pixel corners are at  $y = \pm 1$  and  $z = \pm 1$ . These region must be chosen to be large enough to cope with the somewhat irregular along-track sampling of VIIRS, caused by the bow-tie effect, and ensure that all valid VIIRS pixels within the S5P FOV are retained in this region. A region of at least 65 by 65 VIIRS pixels is found to be necessary in practice.

<sup>5</sup> If necessary sub-sampling of the S5P pixels could be performed to speed up the search, providing this does not degrade the accuracy in the estimated  $\Delta t_{sat}$  below the required level. <sup>6</sup> This should identify at most two files. For code efficiency, the necessary information from both files should be stored in memory. As the loop repeats, this step should be executed only when different files are identified at step 3(a).

- iv. Loop over S5P FOV, identified by index  $k$  ( $= 1 \dots N_F$ ). Each FOV is defined by normalised coordinates  $y_k^{min}, y_k^{max}, z_k^{min}, z_k^{max}$ .
  - A. Identify which VIIRS pixels fall within the given FOV, i.e. those for which  $y_k^{min} \leq z_v \leq y_k^{max}$  and  $z_k^{min} \leq z_v \leq z_k^{max}$ .
  - B. Generate the S5P-NPP cloud record for this S5P scan line, across-track pixel and FOV definition:
    - $\bar{I}_m^{FOV}$ : Mean radiance in each considered VIIRS moderate resolution band identified by index  $m$ . For channels which VIIRS reports as reflectance, this will be converted to sun-normalised radiance by multiplying by the cosine of the VIIRS solar zenith angle. For thermal bands (if required, TBD), VIIRS L1 reported brightness temperatures will be converted to  $mW/cm^2/micron/sr$  using standard coefficients to perform the conversion (as in e.g. [RD13]). The result will be converted back to brightness temperatures.
    - $\Delta I_m^{FOV}$ : Corresponding standard deviation(s) of the radiance in each relevant VIIRS moderate resolution band.
    - $M_m$ : Number of valid VIIRS pixels within the FOV for each considered VIIRS band.
    - $M_l$ : Number of VIIRS pixels within each designated cloud category, identified by index  $l$  ( $= 1 \dots N_C$ ). As a baseline there will be  $N_C = 4$ , for each of the cloud confidence levels. This could however be extended, e.g. to include the thin cirrus mask, cloud shadow mask or the 8 cloud phase categories (if there is a requirement for this).

In addition scene identification (time and geolocation) information copied from the L1 file will be added to enable the record to be easily associated with other S5P products and displayed / geophysically interpreted without reference to other files. As a minimum this information would include the latitude, longitude and sensing time of the centre of the nominal FOV.
- v. Determine averages of the selected VIIRS band radiances,  $\bar{I}_m^{SRF}$ , weighted by the defined S5P SRF, following approach described in section 6.5.
- vi. Also store in the output record the view zenith angle of the identified nearest VIIRS pixel to the S5P FOV centre,  $\theta_v$ , and the difference in time between this observation and that of S5P,  $\Delta t_v$ .

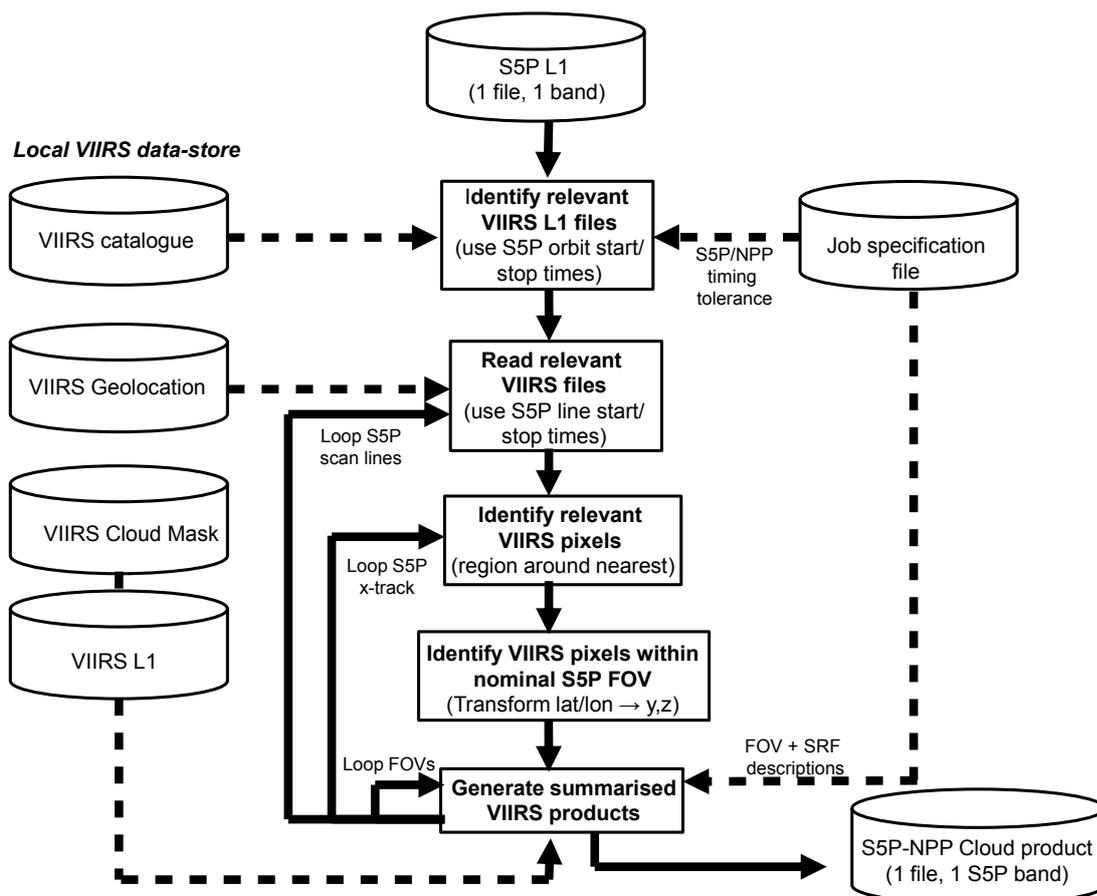
The above description does not explicitly indicate a loop over S5P bands which have distinct geolocation. Since results for different S5P bands are stored in different L1B files, the above process can simply be repeated with different input L1B files.

## 6.7 Treatment of missing / invalid data

The S5P-NPP output data records will be organised to have 1:1 correspondence with the S5P L1B geolocation records. If any of the required data from the S5P L1B geolocation records (see table 1) is missing or invalid then the corresponding S5P-NPP output records will be filled with a suitable “missing data” value.

Records will likewise be filled with “missing data” values if there is missing or invalid VIIRS data, as follows:

- The S5P-NPP processor requires VIIRS L1B geolocation files to exist with 1:1 corresponding VCM and radiance data files. The list of files provided to the processor will be checked for such 1:1 correspondence (using time stamps in the file names) and any anomalies (e.g. a VCM file is named for which there is no corresponding geolocation file) will be flagged in the processor output log. The processor will still function (provided at least one geolocation file is named), but VCM and radiance data files which do not have a corresponding geolocation file will be subsequently ignored by the processor.
- If no geolocation files are named, the processor will terminate with an error.
- Gaps in the VIIRS coverage (e.g. due to missing geolocation files) will not be directly reported as an error, rather such gaps will (naturally) result in no VIIRS pixels being identified for a particular S5P scene. In this case the values of  $M_m$  and  $M_l$  (all values for the given scene) will be set to 0 and values of other output parameters will be set to “missing data”.
- Within geolocation files, the validity of the VIIRS latitudes and longitudes are tested using the provided pixel-level geolocation data quality flag,  $Q_G$ . Values flagged as bad are not used and this will reduce the number of valid VIIRS pixels found within a scene. If this is reduced to zero, then values for the specific S5P scene will be filled as above.



**Figure 17:** Flow diagram illustrating the main processing steps and inputs/outputs.

- The VCM and radiance files also each have corresponding pixel-level quality flags,  $Q_C$  and  $Q_m$ . Again these are tested and if data is flagged bad, the result will not count towards the output statistics. The same is naturally true if entire VCM and/or radiance files are missing. If there is no valid VCM data for a given S5P scene (whether due to missing file or missing data within a file), then all  $M_l$  will be set to zero. Likewise, if there is no radiance data for a particular band then the corresponding values of  $M_m$  will be set to zero, and the corresponding values of  $\bar{I}_m^{FOV}$ ,  $\Delta I_m^{FOV}$  and  $\bar{I}_m^{SRF}$  will be set to “missing value”. In the case of radiance values, a pixel-specific “fill value” is set in the VIIRS file and this is used together with the quality flag (if either indicate no data, then the pixel is ignored).
- As implied above, the code will function if there are missing VCM and / or radiance files. This will simply result in zero counts and missing values as indicated above. Values of  $\theta_v$  and  $\Delta t_v$  are derived only from the geolocation records, so these will be filled irrespective of missing VCM / radiance band missing files or values.
- If errors occur while reading a given VIIRS input file, the processor will proceed as if that file was missing.

The “missing data” value used in the S5P output is to be stored in the output file (as a netCDF attribute to the corresponding record).

## 7 Feasibility

The algorithm described here is very simple and not challenging to code. The main challenges for the processor relate to the very large volumes of data to be accessed, processed and stored. Rough estimates of the data volumes involved are provided below. More detailed information is provided in the IODD.

## 7.1 Scaling parameters

The following parameters are relevant to estimating the computational demands of the NPP-S5P processor:

- $N_F$ : The number of defined S5P FOVs (nominal FOV + scaled versions for each observation). Assume for now  $N_F = 4$ .
- $N_M$ : The number of VIIRS moderate resolution channels used in S5P-NPP product. Assume for now  $N_M = 3$ , to include channels M9, M11 (for comparison to S5P SWIR) and M7<sup>7</sup> (for comparison to the near-ir).
- $N_C$  Number of VIIRS cloud classifications. If only the basic mask is reported then  $N_C = 4$ .
- $N_s$  Total number of S5P observations in an orbit (i.e. spectra assigned distinct location at L1B). The number of S5P scenes (per band) is expected to be approximately 850 000. Assuming data is provided for 3 different bands (SWIR, NIR and VIS) then  $N_s = 2500000$  (2.5 million).

## 7.2 Input data requirements

The following input data is strictly required to produce the S5P-NPP product:

- From S5P L1B, for each band and spectrum:
  - Ground pixel centre coordinates (latitude, longitude).
  - Ground pixel corner coordinates (latitude, longitude).
  - Observation time.
  - Satellite (or view) zenith angle at the ground.
  - Any indices used as standard to identify the scene / band.
- From VIIRS L1B, for each moderate resolution pixel:
  - Latitude and longitude.
  - Observation time.
  - Satellite zenith angle at the ground.
  - Solar zenith angle at the ground.
  - Calibrated reflectances (or brightness temperatures for any selected TIR channels) in the selected bands. As a minimum this should include channel M9. Need for other channels TBC. Channels M7 and M11 are considered the next highest priority at this stage.
  - The pixel level quality flag indicating validity of each measured value.

These records are to be found in the VIIRS Moderate Resolution Band Sensor Data records (SDRs) and associated geolocation data [RD14].

- From VIIRS VCM product:
  - Byte 1 of the 6 byte mask (containing the cloud-confidence mask, and indicator of the mask quality).

Note:

1. In all cases, longitude and latitude are expected to be given for the intersection of the LOS with the WGS'84 ellipsoid, in geodetic coordinates. View zenith angle is also expected at the ellipsoid intersection. Both relevant S5P [RD15] and NPP ATBDs [RD16] use this reference. However it is noted that the VIIRS ATBD specifies that (after computing the ellipsoid intersection) the intersection with terrain height is computed and reported. Both ellipsoid and terrain intersection data are available at CLASS. The former is required for the purpose of producing the S5P-NPP product.
2. Observation time is expected to define both the time of day (to fraction of a second precision) and date in a standard manner (e.g. Julian Date). A difference in time convention between S5P and VIIRS could be accommodated by the S5P-NPP processor.

<sup>7</sup> M6 would be ideal for comparison to the S5P near-ir band, but this VIIRS channels saturates in cloudy conditions.

## 7.3 Input file sizes

The following provides a rough initial estimate of the potential file sizes.

### 7.3.1 S5P

S5P L1B data will be produced for each of the 8 bands (UV, UVIS, NIR and SWIR bands are divided into 2 sub-bands each). The size of a complete set of L1B data is estimated to be around 32 GB per orbit [RD15]. For this purpose only a limited amount of geolocation data is required. The essential data would be the 12 values for each of the  $N_s$  observations (corner, centre locations, observation time, view zenith angle). Assuming each stored as 4 byte words this would give a minimal S5P input file of around 230 MB (uncompressed) per orbit.

### 7.3.2 NPP

VIIRS L1 data is available from [ER3] in tar files containing a single parameter for a complete day. Files for a typical reflectance channel are around 2.5 GB per day (182 MB per orbit). Geolocation parameters occupy  $\approx 2$  GB per day (150 MB per orbit). Assuming the only geolocation parameters required are latitude, longitude, time, satellite zenith angle, and solar zenith angle, and  $N_M = 3$ , this implies a minimal volume of L1 data of around 1 GB per orbit.

VIIRS VCM files from [ER3] have a size of approximately 1 GB per day (70 MB per orbit).

## 7.4 Computation time

Processing time requirements are estimated based on prototype code which generates a representative product ( $N_C = 4$ ,  $N_M = 3$ ,  $N_F = 4$ ) from a 100s section of simulated S5P L1B data and matching real VIIRS products. The prototype is mainly coded in IDL but core routines to do the computationally intensive operations of step 3(d) above are coded in C. Computational requirements are therefore expected to be realistic, though probably represent an overestimate by a factor of 1.5-3.

Processing time, in addition to that needed to read the input files, is currently 50s for the 100s section of orbit, based on an AMD Opteron(tm) Processor 2216 HE, 2400 MHz CPU (SPECint2006=12.3, SPECint\_base2006=11.1, see [ER4]).

Time to read files (dominated by VIIRS) is of order 15s (though network limited).

Processing time for an entire orbit is then estimated to be around 25 minutes (assuming 50 minutes of sun-lit data in the orbit).

Processing time will not scale directly with  $N_C$  or  $N_M$  since the time to compute statistics once the pixels are identified is relatively small. Since the coordinate transformation is only performed once for all FOV, processing time is also *not* scaled by  $N_F$ .

It noted that the L1 data will be provided in “granules” covering sub-sections of a complete orbit (in 5 minute sections), so processing of a single orbit could be carried out in parallel by distributing the individual L1 granules across multiple processors.

## 7.5 Computer memory requirements

The current IDL code requires around 800 MB of memory for the 100s section. This is dominated by storage of VIIRS data and should not be significantly larger for a full orbit or granule (since only two VIIRS images should be needed in memory at any given time in the processing).

## 7.6 Output file size

The prototype processor used to estimate CPU requirements generates a 6.5 MB output file from 100s of S5P data, i.e. 200 MB per orbit (assuming 50 minutes of sun-lit data).

## 8 Validation

There will not be a parallel development of an independent scientific algorithm to generate similar quantities to the S5P-NPP product, which could form the basis of scientific verification along the lines planned for the main

S5P L2 products. The algorithm to be implemented in the S5P-NPP product is relatively simple compared to the other L2 products, so this level of independent scientific verification is not considered necessary.

Simple scientific verification is planned within the L2 development contract to confirm the code performs according to basic expectations. This includes tests to confirm the consistency of the S5P-NPP product with the input VIIRS data. This would include generation of large-scale (e.g. 5 degree) latitude/longitude gridded averaged quantities from the (i) selected sample VIIRS datasets and (ii) the derived S5P-NPP-Cloud product. It should be expected that e.g. the fraction of scenes flagged with a given cloud confidence level within a given latitude/longitude box derived from the VIIRS data directly and the S5P-NPP Cloud product should be very similar. The same should apply to the mean radiances in the 1.385 micron channel. Tests on gridded VIIRS L1B information which can be directly compared to S5P L1B quantities would also be included e.g. observations times should be consistent.

Tests to confirm the code functions as expected prior to code delivery / acceptance are described in the Software Validation Plan [RD17]. These tests include processing synthetic orbits of S5P data, which are generated to co-locate with complete orbits of real VIIRS data. A range of scenarios are considered with a view to ensuring conditions likely to be encountered in flight are covered and the code performs satisfactorily in each case.

During operations, comparisons between the S5P-NPP product and S5P L2 cloud products will no doubt be informative. The cloud parameters will not be strictly comparable so this does not constitute direct validation, but the level of consistency between the two sources of cloud information should be investigated.

If relevant VIIRS channels are processed, then S5P-NPP averaged radiances should be compared to matching S5P radiances (which have been convolved with the VIIRS channel spectral response function). This could be done on a routine basis during operations and would enable the relative calibration and geolocation of the two sensors to be monitored as discussed in section 1.5. This would also help facilitate rapid identification of problems with the S5P-NPP processor.