Algorithm Theoretical Basis Document Version 1.0

Simplified Level 2 Vegetation Processor – Distributed (SL2P-D) for Estimating Biophysical Variables using Sentinel-2 Multispectral Imager Data

Richard Fernandes and Najib Djamai

Canada Centre for Remote Sensing

# Executive Summary

There is a consensus requirement to globally monitor vegetation canopy biophysical variables at medium resolution (<1ha) and a frequency of <=10days (World Meteorological Organization , 2016) . Multispectral satellite based imagers designed to satisfy measurement requirements for this task are and will continue to be available (<https://gcos.wmo.int/en/essential-climate-variables/requirements>). The Simplified Level 2 Processor – D (SL2P-D) produces estimates of canopy biophysical variables given inputs of either a top-of-atmosphere (TOA) or top-of-canopy (TOC) bi-directional multispectral reflectance together with the illumination, view and relative azimuth angles. Separate non-linear regression models are used to estimate the expected value and the expected root mean square error of each variable. The regression estimators are optimized for multi-spectral (i.e. <10 bands with >10nm bandwidth) reflectance inputs but can be applied to arbitrary spectra as long as a radiative transfer model with sufficient accuracy to simulate such spectra is included in the processor. The parameterization, algorithm and sample results of SL2PD are presented and compared with its predecessor, SL2P (Weiss and Baret, 2016).

Contents

[Executive Summary 1](#_Toc19015270)

[Introduction 2](#_Toc19015271)

[Background 2](#_Toc19015272)

[Scope and Objectives 3](#_Toc19015273)

[Content of the Document 3](#_Toc19015274)

[Symbols and Acronyms 5](#_Toc19015275)

[Overview of SL2P-D 5](#_Toc19015347)

[Parameter File 8](#_Toc19015348)

[Inversion Algorithm Definition 13](#_Toc19015435)

[NNET Inversion Algorithm Definition 13](#_Toc19015436)

[NNETP Inversion Algorithm Definition 15](#_Toc19015437)

[Canopy and Atmosphere Parameters 15](#_Toc19015438)

[Process Flow 20](#_Toc19015439)

[Performance 25](#_Toc19015440)

[Regression Input Variables 29](#_Toc19015441)

[Regression Output Variables 29](#_Toc19015442)

[Validation Performance 32](#_Toc19015443)

[Quality Indicators 40](#_Toc19015444)

[Algorithm Implementation 41](#_Toc19015481)

[Conclusions 41](#_Toc19015482)

[References 42](#_Toc19015483)

[Appendix I SL2P-D Canopy biophysical variable parameterizations 43](#_Toc19015484)

[Appendix II – Physical variables 49](#_Toc19015485)

[Appendix III 51](#_Toc19015608)

[Figure 1. SL2P-D architecture. Soil lines indicate data flow during calibration; dashed lines indicate data flow during application. 8](file:///H:\atbdnew_v3.docx#_Toc19015609)

[Figure 2. Comparison of sampling designs for the nominal SL2P canopy architecture distribution using only 1296 samples (in contrast to the nominal 43782 samples) for demonstration purposes. 21](#_Toc19015610)

[Figure 3. Geometric Laws for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (middle) and SL2P (lower). 32](#_Toc19015611)

[Figure 4. Canopy Laws for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (middle) and SL2P (lower). 33](#_Toc19015612)

[Figure 5. Input R for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (lower) 35](file:///H:\atbdnew_v3.docx#_Toc19015613)

[Figure 6. Output variables for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (lower). 36](file:///H:\atbdnew_v3.docx#_Toc19015614)

[Figure 7. Prairie cropland (closed) fCOVER validation for SL2P-D (left) AND SL2P (right). 39](file:///H:\atbdnew_v3.docx#_Toc19015615)

[Figure 8. Prairie cropland (closed) albedo validation for SL2P-D (left) AND SL2P (right). 39](file:///H:\atbdnew_v3.docx#_Toc19015616)

[Figure 9. Prairie cropland (closed) LAI validation for SL2P-D (left) AND SL2P (right). 40](file:///H:\atbdnew_v3.docx#_Toc19015617)

[Figure 10. Prairie cropland (closed) fAPAR validation for SL2P-D (left) AND SL2P (right). 40](file:///H:\atbdnew_v3.docx#_Toc19015618)

[Figure 11. Prairie cropland (closed) LAI.Cw validation for SL2P-D (left) AND SL2P (right). 41](file:///H:\atbdnew_v3.docx#_Toc19015619)

[Figure 12. Prairie cropland (closed) LAI.Cab validation for SL2P-D (left) AND SL2P (right). 41](file:///H:\atbdnew_v3.docx#_Toc19015620)

[Figure 13. Taiga Cordillera Polar Grassland fCOVER validation for SL2P-D (left) AND SL2P (right). 42](file:///H:\atbdnew_v3.docx#_Toc19015621)

[Figure 14. Taiga Cordillera Polar Grassland albedo validation for SL2P-D (left) AND SL2P (right). 42](file:///H:\atbdnew_v3.docx#_Toc19015622)

[Figure 15. Taiga Cordillera Polar Grassland LAI validation for SL2P-D (left) AND SL2P (right). 43](file:///H:\atbdnew_v3.docx#_Toc19015623)

[Figure 16. Taiga Cordillera Polar Grassland fAPAR validation for SL2P-D (left) AND SL2P (right). 43](file:///H:\atbdnew_v3.docx#_Toc19015624)

[Figure 17. Taiga Cordillera Polar Grassland LAI.Cab validation for SL2P-D (left) AND SL2P (right). 44](file:///H:\atbdnew_v3.docx#_Toc19015625)

[Figure 18. Taiga Cordillera Polar Grassland LAI.Cw validation for SL2P-D (left) AND SL2P (right). 44](file:///H:\atbdnew_v3.docx#_Toc19015626)

# Introduction

## Background

The Simplified Level 2 Processor – D (SL2PD) produces estimates of canopy biophysical variables (outputs, Table 1) given inputs of either a top-of-atmosphere (TOA) or top-of-canopy (TOC) bi-directional reflectance spectrum together with the illumination, view and relative azimuth angles. Separate non-linear regression models are used to estimate the expected value and the expected root mean square error (RMSE) of each output. The regression estimators are optimized for multi-spectral (i.e. <10 bands with >10nm bandwidth) reflectance inputs but can be applied to arbitrary spectra as long as a radiative transfer model with sufficient accuracy to simulate such spectra is included in the processor.

Table . SL2P-D output variables.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Acronym | Units | Description | Requirement | Range |
| Fraction Canopy Cover | fCOVER | fraction | Fraction of horizontal surface area covered by vegetation | 20% | 0-1 |
| Fraction Absorbed PAR | fAPAR | fraction | Fraction of absorbed PAR by vegetation | 5% | 0-1 |
| Leaf Area Index | LAI | m2.m-2 | Half the total live foliage area per horizontal ground area | Max(1,20%) | 0-20 |
| Canopy Chlorophyll Content | CCC | ug.m-2 | Mass of chlorophyll A+B per horizontal ground area | Max(20,20%) | 0-1000 |
| Canopy Water Content | CWC | g.m-2 | Mass of H20 per horizontal ground area | 20% | 0-1000 |
| Albedo, black sky | α | fraction | Ratio of top of canopy upper hemisphere upwelling radiance to top of canopy incident direct irradiance | 10% | 0-1 |
| Directional area scattering factor, black sky | D | fraction | Scattering coefficient for foliage single scattering albedo of one, zero boundary reflectance and direct irradiance | None | 0-1 |

SL2P-D has three phases: calibration of the regression models, validation of the regression models with simulations and application of the regression models to measurements. Users are encouraged to include a prior phase corresponding to benchmarking of the simulation model and quantification of input measurement error and a posterior phase corresponding to validation of estimates using internationally accepted protocols (Widlowski et al. 2007, Fernandes et al., 2014).

SL2P-D is a modification of the SL2P processor (Weiss and Baret, 2016) and shares the same general approach for producing simulations, for using non-linear regression estimators of output expected values and expected RMSE, and for visualization of cross-validation results. SL2P-D has three major differences from SL2P: i. it allows additional radiation transfer models (RTMs), sampling schemes, and regression models ii. it allows for reuse of simulations across classes in a database, iii. it allows for a distributed regression estimates for expected values of outputs. The first two differences increase flexibility and efficiency when applying SL2P to local conditions. The third difference allows users to refine the initial regression estimate by calibrating multiple regression models stratified on one output. This stratification reduces the effective dimensionality of each of the multiple models and places implicit prior constraints on the covariance of both canopy parameters and outputs, at least in terms of bounds. The user has the flexibility of the initial stratification output. As Section xx indicates, cascaded models using either fCOVER or D can potentially reduce uncertainty of estimates of parameters such as LAI at high values of LAI.

SL2P-D is intended both as an evolution of SL2P, in that it implements the same theoretical basis of SL2P as one of its processing options, and an improvement in that it is engineered in a modular manner with respect to calibration procedures and leverages advances in machine learning and statistical tools over the decade since SL2P was written. SL2P-D may also offer actual theoretical advantages in the use of cascaded estimation. SL2P, as implemented for the Sentinel 2 Multispectral Imager (MSI), specified priori distributions for canopy parameters within the calibration database that match global priors from field experiments. As Djamai et al. (2019) show, such a strategy leads to regression estimates biased to these priors even though they do not necessarily represent either local priors or indeed the area weighted global prior. We hypothesize that the cascaded network with the D constraint calibrated using uniform priors may in many cases perform with lower bias than SL2P when applied locally.

## Scope and Objectives

The objective of this document is to provide a theoretical basis for the implementation of SL2P-D available at (<https://github.com/rfernand387/SL2PD>). To satisfy this objective the inputs to the SL2P-D algorithm are defined and the algorithm is discussed on a functional level. At each stage a theoretical justification is provided. In many instances, the justification simply follows from SL2P and is noted as such. Although, especially with respect to additional features such as sampling , new RTMs , new regression approaches and the use of D within a cascaded regression, a detailed theoretical basis is provided.

After reading this document one should be able to produce a new calibration of SL2P-D, develop regression estimators for outputs and their RMSE, and assess the performance by cross-validation. However, this document is not an engineering manual for the SL2P-D code or for modification of the code. The reader is referred to <https://github.com/rfernand387/SL2PD> for the code release and documentation. Moreover, this release of SL2P-D is not intended for high-performance application to arbitrary input images; in part because of the complexity of image formats and in part because of varying user production requirements. A version of SL2P-D has been implementaed in Google Earth Engine () for this purpose. Readers interested in modifying the code are encouraged to do so and to contact either the current authors or the SL2P authors for advice but should be aware they are entering terra incognito and walking into the realm of science (as Fred and Marie would say ‘bonne chance!’).

## Content of the Document

The document begins with a list of symbols and acronyms. The document is then separated into five sections: i. a general overview of the SL2P-D approach with a focus on new features compared to SL2P, ii. A description of parameterization of SL2P-D, iii. A description of the functional execution of SL2P-D (essentially a walk through of the MATLAB implementation), iv. A description of cross-validation and a sensitivity study comparing various options for typical canopy conditions starting with a close approximation of the original SL2P solution and ending with a recommended SL2P-D solution, v. a description of the procedure to implement the SL2P-D calibration and subsequently apply SL2P-D using Google Earth Engine.

# Symbols and Acronyms

|  |  |  |
| --- | --- | --- |
| Symbol/Acronym | Name | Units |
| AOD550 | AOD 550nm | DIM |
| b | Crown base height | m |
| B | Soil brightness | % |
| Bryo | Moss&Lichen | Nominal |
| Cab | Leaf chlorophyll content | µg.cm-2 |
| Cbd | Leaf brown pigment content | g.cm-2 |
| Ccar | Leaf carotenid content | g.cm-2 |
| CCC | Canopy chlorophyll content | µg.cm-2 |
| Cdm | Leaf dry matter content | g.cm-2 |
| CH20 | Water vapour concentration | g.cm-2 |
| CO | ozone concentration | mol |
| Crop | Cropland | Nominal |
| Cw | Leaf water content | µg.cm-2 |
| CWC | Campy water content | g.cm-2 |
| Cx | Leaf Xanthophyll content | µg.cm-2 |
| D | Canaopy directional scattering factor | DIM |
| d | Leaf diameter | cm |
| DBF | Deciduous Broadleaf Closed Forest | Nominal |
| DNF | Deciduous Needleleaf forest | Nominal |
| DOY | Day of year | Day number |
| EBF | Evergreen Broadleaf Forest | Nominal |
| ENF | Evegreen Needleleaf forest | Nominal |
| fAPAR | Fraction absorbed photonsynthetically active radiation | 0-1 |
| fCOVER | Fraction ccover | 0-1 |
| Gr | Grassland or pasture | Nominal |
| GrPolar | Polar grassland | Nominal |
| h | Overstory Height | m |
| LAD | Leaf angle distribution | probability histogram |
| LAI | Leaf area index | DIM |
| Lat | Latitude | °N |
| Long | Longitude | °E |
| MF | Mixed forest | Nominal |
| MSI | Multispectral Imager | Nominal |
| N | Leaf layers | Fractional count |
| Patm | Atmospheric pressure | KPa |
| R | Bi-directional reflectance | 0-1 |
| RTM | Radiative transfer model | Nominal |
| Sh | Shrubland | Nominal |
| ShPolar | Polar shrubland | Nominal |
| Toa | Top of atmosphere | Nominal |
| Toc | Top of canopy | Nominal |
| Veg | Vegetated | Nominal |
| w | Crown width | m |
| Wet | Herbaceous wetland | Nominal |
|  | Black sky albedo | DIM |

# Overview of SL2P-D

SL2P-D uses an input parameter file to perform three sub-processes with corresponding outputs (Figure 1). During calibration, all three processes are executed (in sequence of Simulation, Regression Calibration, and Cross-Validation). During application, only cross-validation is executed using input from existing regression predictors and a simulation database whose inputs satisfy the predictors.

Figure 1. SL2P-D architecture. Soil lines indicate data flow during calibration; dashed lines indicate data flow during application.

Regression Calibration

Simulation

Cross-Validation

Parameter File

Simulation

Database

Performance Visualization

Regression Predictors

SL2P-D is essentially a database with associated processes that operate on the data. SL2P-D is not object oriented so there are no explicit associations between processes and data. In principle one could replace elements such as the simulation database or the regression predictors using versions generated by external processes. This strategy is employed during application of SL2P-D where the input data to be processed is formatted as a database with (in most cases unless validation data exists) null output values.

Execution of SLP2-D is controlled by the parameter file. The parameter file is currently produced manually by an expert user while the other three steps are produced using published algorithms verified using international standards (in the case of RTMs) or through quality assurance from vendors (in the case of MATLAB code). Substantial detail and guidance is provided here with respect to defining a parameter file as it requires expert input. In contrast, the other components are described more in terms of how their production and specification differs from SL2P. The application for SL2P-D closely follows SL2P with the exception of additional algorithm complexity when using distributed networks. This document describes the application within the MATLAB prototyping environment.

Similar to SL2P, SL2P-D is organized around the concept of a algorithm database. This database encodes the parameter file, the simulation database(s), the regression predictors, and performance visualizations using MATLAB data structures or .png images in a unique user designated database directory. In contrast to SL2P, all simulations and regression predictors in a SL2P-D directory share the same atmosphere and soil reflectance library parameters. This allows for reuse of simulations between classes. Users who wish to apply the same land cover class to different atmosphere or soil conditions should make a new database directory.

The Simulation process produces input-output samples for calibration, cross-validaton and validation of regression estimators. Here, calibration samples are used to optimize prediction model parameters, cross-validation samples are used to determine the expected prediction model error if the model is applied to the population containing the calibration samples and validation samples are used to determine the prediction model error for a specific population. Samples are produced for a specified number of land cover classes (Classes) based on applying a RTM to a set of input Laws. The Laws for acquisition geometry and atmosphere are the same across all Classes; users should define a new parameter file and associated databases if these Laws change such as when applying SL2P-D on a regional basis or a validation database is required. The joint distribution of desired geometry and atmosphere Laws is randomly sampled and shared among all land cover classes. The Laws for canopy parameters are Class specific and based on physical quantities (Appendix I). For each Class, the joint distribution of desired canopy Laws is sampled using a user designated method. SL2P used Full-Orthogonal sampling that, while unbiased in terms of fitting regression models, required approximately a doubling of sample size for each new stratum. This is problematic when varying strata for different classes or networks in a cascade. SL2P-D provides both quasi-randomized Halton set and Sobol set sampling that offer no restrictions on sample size while also being unbiased and potentially resulting in greater precision of estimates of regression errors ( <https://www.mathworks.com/help/stats/haltonset.html> and <https://www.mathworks.com/help/stats/sobolset.html>). Other sampling schemes such as Full Orthogonal, Latin Hypercube and Monte Carlo are provided for cross-validation studies.

During the Simulation process, SL2P simulated an input-output pair for each sampled canopy Law using the ‘sail3’ RTM and a randomly sampled geometry Law. SL2P-D differs in two aspects. First, PROSAILD and FLIGHT RTMs have been added as options to supplement the ‘sail3’ RTM used in SL2P. These new models have been evaluated in the RAMI exercise and are within the ensemble of benchmark models for simulating homogenous canopies (for both) and heterogeneous (for FLIGHT) canopies over flat terrain and nadir view angle (Widlowski et al., 2007). Second, samples already simulated for other classes in the database are used to replace a desired sample for the current class if they differ by less than a user specified tolerance. To avoid duplicate matches the replacement is performed by selecting the best possible valid matches over all desired samples using the classical ‘Assignment Problem’ solution (<https://www.mathworks.com/help/matlab/ref/matchpairs.html>) . New RTM simulations are performed only on unmatched desired samples.

SL2P used the calibration database to fit single hidden layer multi-layer perceptron ‘neural network’ regressions to estimate one output (or the output error) given user specified input. SL2P-D generalizes this approach in three aspects. First, SL2P used a fixed hold out (1/3) proportion of samples from the calibration database for cross-validation of the final regression estimate. This may result in a bias due to the permanent loss of certain samples from the calibration data. Rather, SL2P-D randomizes the hold-out over multiple iterations (batches) of each regression. Second, SL2P-D allows for additional regression estimators including Gaussian Process Regression and Support Vector Regression. These are not currently employed for production algorithms but are included for scientific studies. Finally, SL2P-D allows for a cascaded regression estimator as previously described.

Currently, SL2P-D application corresponds to an operation on a database of inputs formatted as a simulation database. A sample function for parsing input Sentinel 2 MSI satellite products into such an input database format is provided. Parsing output into image products is beyond the scope of this document. Users should make use of the SL2P-D Google Earth Engine application if product generation is required.

# Parameter File

The parameter file is a Microsoft Excel .xls file (see sample at <https://github.com/rfernand387/SL2PD>) that defines algorithm calibration, including specification of the input measurements and outputs, the regression model for estimating outputs, and the nominal validation database. The parameter file has a number of worksheets that are described below. Worksheet formatting must be preserved during calibration. In most instances the MATLAB implementation will indicate an error if formatting or errors in data entry result in inconsistent parameter values or naming of resources (database or classes) that do not exists.

## Start

This sheet contains the title of the algorithm, a database, the name of the validation database and a flag (CopyFlag) indicating if simulations can be reused between land cover class (Table 2).

Table . Description of Start worksheet entries.

|  |  |
| --- | --- |
| Database directory name | Name of the directory that will contain the report |
| Comments | Comments |
| Inversion Algorithm Name | Inversion Algorithm Name |
| Name of Validation Database Directory | Name of Validation Database Directory |
| CopyFlag | CopyFlag |

By convention, uses should tile databases as:

{Sensor input}\_{geographiczone}\_{soildatabase}\_{terraincomplexity}\_{size}\_{sampling scheme}.xls

Phrases in brace brackets are selected from valid values from tables described below. The title is assigned to the database andused to determine if this database can be used to validate regression models from another database.

## Learning Data

The Learning Data worksheet contains global parameters for the algorithm as defined in Table 3.

Table 3. Learning Data worksheet descripton.

|  |  |  |
| --- | --- | --- |
| Name | Description | Range |
| Toc\_Toa | Switch between TOA and TOC reflectance | ‘Toc’ or ‘Toa’ |
| Terrain | Terrain complexity for input noise | ‘Simple’ or ‘Complex’ |
| Classification | Land cover classification for documentation | ‘CCRS’ , ‘SL2P’ |
| Nb\_Classes | Number of classes used; each requires a canopy/atmosphere description worksheet | <= #classes in Classification |
| FAPAR\_Time | Time used for fAPAR computation | Local Time hh:mm |
| RTM | Radiative transfer model used for all simulations | ‘sail3’; ‘ProsailD’; ‘Flight’ |
| Max\_Sims | Maximum # simulations per class; used to simulate geometry. | >=Nb\_Sims of all of the classes |

‘Toc\_Toa’ specifies the input reflectance measurements as either top of canopy or top of atmosphere. If top of atmosphere is specified the SMAC radiative transfer model is used with a sampled distribution of atmosphere parameters specified under the Canopy\_Atmosphere\_Class1 worksheet to convert produce a calibration database of top of atmosphere reflectance; else the output of the selected canopy RTM is used.

‘Terrain’ specifies the terrain complexity in the sense defined by the SEN2COR atmospheric correction algorithm. Complex terrain implies the contribution from surrounding land surfaces and the local slope cannot be neglected when determining TOA reflectance. Since only 1D RTMs are used to model reflectance this contribution is considered noise and specified as such in the ‘Sensor’ worksheet. Currently the results of Djamai and Fernandes (2018) are used to quantify the noise levels.

‘Classification’ specifies the land cover classification associated with each Class. This is only for documentation purposes since the parameters for each Class are specified in individual worksheets.

‘Nb\_Classes’ specifies the number of classes in this database. Each class must be specified in a separate Canopy\_atmosphere\_Class# worksheet.

‘FAPAR\_Time’ species the local time at with black sky fAPAR is computed in the simulation database.

‘RTM’ specifies the canopy radiative transfer model used for all simulations. The same model is used for all classes to allow for copying of simulations between classes. Separate databases are required if the RTM is not the same but all RTMs make use of the same input parameters.

‘Max\_Sims’ specifes the maximum allowable simulations for a class. It is used to produce a database of randomly sampled geometric acquisition variables based on the spatiotemporal extents specified in the ‘Configuration’ worksheet. The database is randomized before use.

‘Validation\_Name’ specifies the name of the database containing validation data. The database may be i. the same as the current database in which case all calibration data is used for validation ii. A different complete database or iii. A partially specified database containing only input and output values. In all cases, the database must include a definition (Def\_Base) specifying the number of classes and corresponding input and output data for each class.

‘Copy\_Flag’ is set to ‘yes’ if simulated input-output pairs for a class should be first copied from existing simulated classes in the same database. This avoids having to redo similar simulations but slightly modifies the initial sampling scheme for the class since a desired simulation law may not have an exact match in any other class. The maximum tolerated mismatch is determined by the most constraining Nb\_levels parameter specified in the Canopy\_Atmosphere\_Class# worksheet for the class being populated.

## Sensor

The ‘Sensor’ worksheet specifies the Sensor, input bands and level of additive and multiplicative noise for each band and acquisition geometry parameter under simple and complex terrain conditions. The Sensor must be selected from a pre-defined list indicated in the worksheet for which orbit and SMAC parameters are pre-tabulated in files ’Orbito\_Sensor.m’ and ‘Flitres\_Smac.mat’ respectively. The user should update these files and the ‘Bandnames’ worksheet with new sensors or if sensor specifications change; ideally using a calibration update date for the latter as a suffix for the sensor name. Table 4 indicates the input noise specification according to Djamai and Fernandes (2018).

Table 4. Input noise specification for simple and complex terrain.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Input | Simple Terrain | | Complex Terrain | |
|  | Additive Noise | Multiplicative Noise | Additive Noise | Multiplicative Noise |
| %refl or ° | % | %refl or ° | % |
| MSI B02 | 2 | 2 | 3 | 2 |
| MSI B03 | 2 | 2 | 3 | 2 |
| MSI B04 | 2 | 2 | 3 | 2 |
| MSI B05 | 2 | 2 | 3 | 2 |
| MSI B06 | 2 | 2 | 4 | 4 |
| MSI B07 | 2 | 2 | 4 | 4 |
| MSI B8a | 2 | 2 | 4 | 4 |
| MSI B11 | 2 | 2 | 4 | 4 |
| MSI B12 | 2 | 2 | 4 | 4 |
| Solar Zenith Angle | 0 | 0 | 10 | 0 |
| View Zenith Angle | 0 | 0 | 10 | 0 |
| Relative Azimuth Angle | 0 | 0 | 0 | 0 |

## Bandnames

The ‘Bandnames’ worksheet specifies the allowable band names that can be selected in the ‘Sensors’ worksheet as inputs to the calibration database. Each new sensor (or newly calibrated sensor) is specified in a new column in this worksheet. The name of the sensor at the top of the column and the band names in lower rows must match the name of a structure and associated structure members in the ‘Orbito\_Sensor.m’ and ‘Filters\_Smac.mat’ file. Following Djamai et al. (2018) the following bands used for Landsat OLI and Sentinel 2 MSI are indicated in Table 5 and Table 6.

Table 5. Landsat OLI bands used within SL2P-D.

|  |  |  |  |
| --- | --- | --- | --- |
| OLI Band | Central wavlength | Nominal Width | Spatial resolution |
| name | nm | nm | m |
| 2 | 482.5 | 66 | 30 |
| 3 | 562.5 | 76 | 30 |
| 4 | 655 | 51 | 30 |
| 5 | 865 | 41 | 30 |
| 6 | 1610 | 101 | 30 |
| 7 | 2200 | 201 | 30 |

Table 6. Sentinel 2 MSI bands used within SL2P-D.

|  |  |  |  |
| --- | --- | --- | --- |
| MSI Band | Central wavlength | Nominal Width | Spatial resolution |
| name | nm | nm | m |
| B3 | 560 | 35 | 10 |
| B4 | 665 | 30 | 10 |
| B5 | 705 | 15 | 20 |
| B6 | 740 | 15 | 20 |
| B7 | 783 | 20 | 20 |
| B8a | 865 | 20 | 20 |
| B11 | 1610 | 90 | 20 |
| B12 | 2190 | 180 | 20 |

## Configuration

The ‘Configuration’ worksheet specifies the geometric configuration, spatial and temporal extents for sampling the acquisition geometry for all simulations in the database. One of a number of possible extent specifications can be specified (Table 7) by filling required information (Table 8). The algorithm uses the first specification if multiple are entered inadvertently. Configurations corresponding to Multiple Dates, Locations and Configurations depending on Canadian ecozones as indicated in Appendix II. However, sensitivity analysis indicates that a global parameterization is sufficient.

Table 7. Cases considered for sensor configuration, spatial and temporal extents.

|  |  |  |
| --- | --- | --- |
| Case | Definition | Example |
| Case Single Date, Location and Configuration | The sun position can be specified either or from time and location. | One experiment at one date, nadir viewing. |
| Case Single Date, Location and Multiple Configuration | The view directions are specified in a file. The sun position can be specified either directly or from time and location. | Acquisition at a given date and location, but over a large extent (a SPOT or TM image) or with several view directions (CHRIS/POLDER) |
| Case Multiple Dates, Locations and Single Configuration | The locations can be specified by a range of lat/lon of by a list of coordinates. | Processing of global data from directionnally normalized data (CYCLOPES) |
| Case Multiple Dates, Locations and Configurations | The locations can be specified by a range of lat/lon of by a list of coordinates. | Processing of global data from instantaneous observations (MERIS, Landsat OLI, S2MSI) |
| Case Multiple Dates, Single Location and Single Configuration | The locations can be specified by a range of lat/lon of by a list of coordinates. | Processing of images acquired with the same view angle, over the same target, at multiple dates (formosat) |

Table 8. Template datasheet for specification of a selected configuration, spatial and temporal extents. Allowable entries, identified in yellow, vary with configuration indicated in Table 2.

|  |  |  |  |
| --- | --- | --- | --- |
| Selected | **Min** | **Max** | **File** |
| **Sun zenith (°)** | 999 | 999 |  |
| **Sun Azimuth (°)** |  |  |  |
| **View zenith (°)** |  |  |  |
| **View Azimuth (°)** |  |  |  |
| **Year** |  |  |  |
| **Day of year** | 1 | 365 |  |
| **Hour (UT)** |  |  |  |
| **Minute** |  |  |  |
| **Lat (°)** | -56 | 83 |  |
| **Lon (°)** | 0 | 360 |  |

## Inversion Algorithm Definition

The inversion algorithm corresponds to a list of outputs and a specification for an estimation algorithm. Multiple inversion algorithms can be applied to a database by repeatedly executing the calibration process as long as the list of desired outputs does not change. This restriction is applied to minimize the size and data processing demands on producing a calibration database since otherwise all possible outputs would always have to be included in the calibration database. Table 9 lists the supported inversion algorithms. Each inversion algorithm requires parameters specified in sub-sections below. Currently the NNETP algorithm is used.

Table 9. Supported inversion algorithms.

|  |  |  |
| --- | --- | --- |
| Name | Description | Implementation |
| NNET | backpropagation network | MATLAB net |
| NNETP | partitioned backpropagation network | MATLAB net applied first to estimate one output variable followed by multiple MATLAB net networks each for small range of this variable |
| LUT | Look up table | Not implemented |
| GPR | Gaussian process regression | Not implemented |
| SVR | Support vector regression | Not implemented |

### NNET Inversion Algorithm Definition

Following the theoretical basis provided in SL2P, a single hidden layer multilayer perceptron neural network regression (implemented in MATLAB using the ‘train’ function) is defined to estimate a specified output given input reflectance measurements, cosine of view zenith angle, cosine of solar zenith angle and cosine of relative azimuth angle using fixed and used defined parameters provided in Table 10.

SL2P randomly allocated 2/3 of the calibration database to training and 1/3 to cross-validation. This approach will result in a network optimized to estimate outputs assuming they are randomly distributed over the prior distribution of simulation laws. This is not realistic since many canopy and soil parameters are often locally correlated. Cross validation using samples where certain parameters are correlated is likely more realistic and hence provide a less biased estimator and performance estimates. For SL2P-D, the calibration database is partitioned into clusters of approximately 100 samples using the k-means algorithm (<https://www.mathworks.com/help/stats/kmeans.html>) applied to input distributions for ALA, Cab, N, Cdm, Cw\_rel and Bs with uniform initialization. Canopy structure variables were not included since they cannot be assumed to be locally correlated. Ideally the network would be trained using a batch of different clusters held out for validation at each epoch. As this feature is not implemented in MATLAB replicates (user specified as Num\_networks) network is trained using a batch corresponding to a random sample of 1/3 of the clusters and initial weights from the previous batch (or randomly initialized for the first batch). The performance of each network is quantified using a regularized metric corresponding to the sum of the mean square prediction error over the cross-validation samples and a faction of the mean square weights.

Table 10. Regression network parameters.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Description | Nominal Value | Range |
| Number of networks | Number of replicate networks | 5 | >=1 |
| Number of hidden layers | Number of hidden layers per network | 1 | 1,2 |
| Transfer function L1 | Node transfer function for layer 1 | tansig | ‘poslin’ for positive linear; ‘tansig’ for hyperbolic tangent sigmoid |
| Number of neurones L1 | Number of neurons for layer 1 | 10 | >=1 |
| Transfer function L2 | Node transfer function for layer 2 | 0 | ‘poslin’ for positive linear; ‘tansig’ for hyperbolic tangent sigmoid |
| Number of neurones L2 | Number of neurons for layer 2 | 0 | >=1 |
| Tolerance | Limits for considering out of range retrievals still possible | 0.1 | varies with output variable |
| Time of instantaneous fAPAR (HH.MM) | Used to document time of the fAPAR used during calibration | 10:00 | Local time hh:mm |
| Performance regularization | Additional error contribution from mean square magnitudes of weights. | 0.1 | 00.1 |
| Epochs | Maximum number of sweeps through training database | 250 | 250 |
| Performance function | Objective fuction metric used | mse | ‘mse’ for mean square error |
| goal | Stopping error level. | 1e-3 | 1e-3 |
| Update | Algorithm used to update network | Levenberg-Marquardt | Levenberg-Marquardt |

### NNETP Inversion Algorithm Definition

The NNETP Inversion Algorithm uses an initial NNET inversion of a single output (P) to partition the calibration database and subsequently train separate NNET estimators for each partition. The theoretical basis for the NNETP algorithm relies on the possibility of retrieving P in a well-posed and well-conditioned manner without substantial prior information. In this case, P can be used to constrain entries in the training database. The improvement of performance over a standard NNET estimate will depend on the uncertainty of the estimate of P together with the constraint that knowledge of P has on the prior distribution of canopy parameters relevant to each output. For example, Fernandes and Dajami (2019) showed that it is possible to estimate the canopy reflectance if foliage albedo was one (D) without prior information for all but very low (<1) LAI. Further, the showed that canopy structure was strongly correlated with D; resulting in a >50% reduction in uncertainty in LAI with uniform priors and a small improvement in estimates of biochemistry (fAPAR and fCOVER were well estimated irrespective of priors for homogenous canopies).

Once P is estimated, the maximum range of P for the class is partitioned uniformly into intervals with width equal to the RMSE for predicting P based on cross-validation. For each interval, a subset database corresponding to all calibration samples matching the value of P within the interval and its neighbouring intervals (to allow for smooth interpolation across interval boundaries). A NNET estimator is then calibrated for the interval using the subset database. The cascaded network is applied using the same approach for new measurements.

### Canopy and Atmosphere Parameters

Canopy and atmosphere parameter pdf distributions required to produce the calibration database (Table 11) are specified for each class in separate worksheets named ‘Canopy\_Atmosphere\_Class#’ (e.g. Appendix III). For simplicity sheets must be follow increasing natural numbers corresponding to the class names identified in the ‘Learning’ worksheet.

Canopy and atmosphere distributions are assumed independent. There are currently 10 canopy parameters and 4 atmosphere parameters. The number of canopy parameters can be increased but this should be avoided if possible to reduce the dimensionality of the distribution of parameters to be sampled. Each parameter is specified using a truncated two parameter distribution. This differs from SL2P that performed truncation after generating distributions that can result in oversampling at the truncation bounds. Refer to <https://www.mathworks.com/help/stats/prob.normaldistribution.pdf.html> for valid distributions and definition of the two parameters (P1 and P2). The MATLAB ‘trunc’ option (<https://www.mathworks.com/help/stats/prob.normaldistribution.truncate.html>) is used with the provided lower and upper bound parameters.

Table . Canopy and atmosphere parameters for all RT models shown here for SL2P default class (the age parameter is not used for 1D RT Models).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Variable** | **Lower**  **Bound** | **Upper**  **Bound** | **P1** | **P2** | **Nb\_Class** | **Law** | **LAI\_**  **Conv** | **Var\_min (0)** | **Var\_max (0)** | **Var\_min (LAI\_Conv)** | **Var\_max (LAI\_Conv)** |
| **Canopy** | **LAI** | 0.0 | 15.0 | 2.0 | 3.0 | 6 | Normal | 1000 | 0 | 20 | 0 | 20 |
| **ALA (°)** | 30 | 80 | 60 | 30 | 3 | Normal | 10 | 30 | 80 | 55 | 65 |
| **Crown\_**  **Cover** | 1.0 | 1.0 | 1.0 | 1.0 | 1 | Uniform | 10 | 1 | 1 | 1 | 1 |
| **HsD** | 0.1 | 0.5 | 0.2 | 0.5 | 1 | Normal | 1000 | 0.1 | 0.5 | 0.1 | 0.5 |
| **Leaf** | **N** | 1.20 | 2.20 | 1.50 | 0.30 | 3 | Normal | 10 | 1.20 | 2.20 | 1.30 | 1.80 |
| **Cab (µg.m-2)** | 20 | 90 | 45 | 30 | 4 | Normal | 10 | 20 | 90 | 45 | 90 |
| **Cdm (g.m-2)** | 0.0030 | 0.0110 | 0.0050 | 0.0050 | 4 | Normal | 10 | 0.0030 | 0.0110 | 0.0050 | 0.0110 |
| **Cw\_Rel** | 0.60 | 0.85 | 0.60 | 0.85 | 4 | Uniform | 10 | 0.60 | 0.85 | 0.70 | 0.80 |
| **Cbp** | 0.00 | 0.20 | 0.00 | 0.30 | 3 | Normal | 10 | 0.00 | 0.20 | 0.00 | 0.20 |
| **Soil** | **Bs** | 0.50 | 3.50 | 1.20 | 2.00 | 4 | Normal | 10 | 0.50 | 3.50 | 0.50 | 1.20 |
|  | **Age** | 10.00 | 150.00 | 90.00 | 40.00 | 1 | Normal | 10 | 10.00 | 150.00 | 10.00 | 150.00 |
| **Atmosphere** | **P (mbar)** | 950.00 | 1080.00 | 950.00 | 1080.00 | 1 | Uniform | 1000 |  |  |  |  |
| **t550** | 0.00 | 0.80 | 0.00 | 0.80 | 1 | Uniform | 1000 |  |  |  |  |
| **H2O (cm)** | 1.00 | 6.50 | 1.00 | 6.50 | 1 | Uniform | 1000 |  |  |  |  |
| **O3 (dbs)** | 0.00 | 0.80 | 0.00 | 0.80 | 1 | Uniform | 1000 |  |  |  |  |

Three dimensional RT models require additional parameters. To reduce the dimensionality of he sampled pdf these additional parameters are determined by random sampling species dependent pdfs conditional on the canopy parameters indicated in Table 11. An additional worksheet is required for each 3D model with each row specifying pdfs for additional parameters for each possible species. For each sample, the parameter values sampled from Table 11 are used to sample from the pdf for the additional parameters for each species resulting in multiple possible 3D parameter realizations (e.g. Appendix III). Subsequently, only physically valid realizations (satisfying the species maximum within crown LAI) are retained and their probability is determined as the normalized conditional probability that the species matches the sampled Age. The pdf of species given age is then used to randomly select one species only. The parameters corresponding to the realization for this species are assigned to the already sampled parameters from Table 11. This approach ensures consistent physically valid Laws that are weighted towards likely species given age. As the likelihood of species given age are relatively flat pdfs this approach essentially prevents unrealistically old cases for certain species rather than actually biasing the selection among species that normally would be present at a given age.

One of five sampling schemes, listed in Table 12, are provided to sample the joint distribution of either canopy or atmosphere parameters. All of these schemes are unbiased so the calibrated regression will be unbiased if tested with samples from the same distribution. However, the precision of the calibrated network will vary with sample size. Table 12 indicates the theoretical precision as a function of training sample size and the effective dimensionality of the problem (which will lie somewhere between 1 and 10 for the canopy). Full orthogonal sampling guarantees a reduction rate better than Monte Carlo and potentially as good as most of the other sampling schemes. Full orthogonal has been used for single class applications by SL2P but is problematic if the sample size must vary since increases in sample size must be performed in steps of ~ and resampling, for example when reusing simulations for another class, must be done carefully to avoid large gaps between samples. Scrambled Halton sampling offers better precision than Monte Carlo sampling when the effective dimensionality of the problem space, , is less than 3 but does not always guarantee better performance than the best case Full Orthogonal. However, scrambled Sobol sampling gives better precision the Full Orthog sampling for and better precision than Monte Carlo sampling for . Scrambled Sobol sampling is used as default for calibration. However, estimation of the regression performance is performed using Monte Carlo sampling to ensure there is no correlation between residuals due to the validation sample design.

Table 12. Sampling designs for calibration and validation.

|  |  |  |  |
| --- | --- | --- | --- |
| Sampling Design | Description | Theoretical Precision | MATLAB Implementation |
| Monte Carlo | Random sampling |  | rand |
| Full Orthogonal | Random sampling within full factorial sampling |  | fullfact |
| Latin Hypercube | Random sampling within Latin hypercube |  | lhsdesign |
| Quasi-randomized Halton Set | Halton set sampling followed by reverse-radix scrambling |  | haltonset followed by scramble with ‘rr2’ option |
| Quasi-randomized Sobol Set | Sobol set sampling followed by affine scrambling | < | Sobolset followed by scramble with 'MatousekAffineOwen' |

As in SL2P, for efficiency, sampling is initially performed assuming parameters are independent. In reality, canopy parameters exhibit covariation. Following SL2P only the covariation of the truncation bounds of each parameter with respect to LAI is specified. This is performed by using linear bounds as a function of LAI. Out of bounds samples must be resampled as discarding a sample will adjust the marginal distribution of all canopy parameters. SL2P resampled by reflecting out of bounds values into valid bounds (e.g. if Cab was out of bounds by 10gcm-2 it would be adjusted to the minimum of 30gcm-2 or the upper Cab bound for the current LAI). This heuristic tends to increase sampling near boundaries and can result in a pdf that deviates from the shape of the user specified pdf for the out of bound parameter (typically resulting in heavier tails). SL2P-D resamples using a pdf specified with the new bounds. This ensures that the conditional distribution of parameters does not change for a given LAI although it does potentially increase the uncertainty of the sampling design somewhat since the final sampling design will now deviate from a hypothetically perfect design.

The combination of constraints on bounds and the sampling method results in sampled canopy parameters that exhibit trends in terms of range with respect to LAI and clumps (Figure 2). Typically, the sample size is set at a level where sensitivity of the accuracy of estimated canopy variables to the sampling design is small. However, both Sobol and Halton sampling have fewer clumps suggesting that they may provide more uniform estimation performance for predicting algorithm uncertainty over the range of canopy parameters.

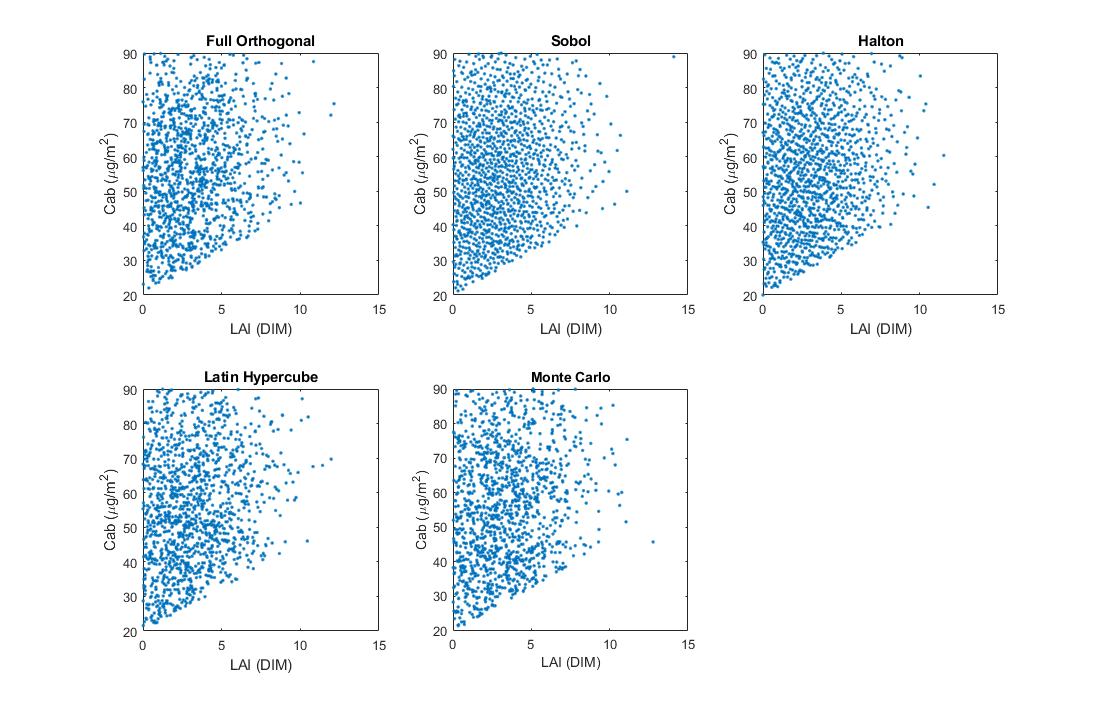


Figure . Comparison of sampling designs for the nominal SL2P canopy architecture distribution using only 1296 samples (in contrast to the nominal 43782 samples) for demonstration purposes.

Calibration database input and output values are required for each sample. While some of these values (e.g. angles) are computationally trivial others (e.g. fAPAR and albedo) require simulating spectral bi-directional reflectance which is non-trivial, especially when using 3D RT models (e.g. FLIGHT). To reduce computation, a calibration sample for a current class is first compared to the calibration sample set for other source classes that already have simulations. Each canopy parameter is normalized to range between -1 and 1 for the current class. The same scaling is then applied to the source class. LAI is stratified equally using a specified number of levels (as would also be the case for Full Orthogonal sampling). The difference in scaled canopy parameters for each pairwise combinations between current and source samples for a given LAI strata is computed. Pairs with differences that exceed the minimum sampling interval of each parameter are discarded. The remaining source samples (parameters, inputs and outputs) are uniquely assigned to current samples so as to minimize the overall distance using the MATLAB solution (<https://www.mathworks.com/help/matlab/ref/matchpairs.html>) to the ‘Assignment Problem’. This process is repeated for the unmatched current samples with the next available database having inputs and outputs. When all available database are exhausted the radiative transfer model is applied to simulate the required inputs and outputs for each of the final unmatched current samples. If current classes are processed from largest to smallest sample size, in the best case this strategy coupled will result in having to simulate only as many samples as the largest sample size. In practice the actual number of samples simulated will be somewhat larger as the truncation bounds for classes will result in areas of the canopy parameter space that do not overlap between the current class.

The same canopy and atmosphere parameters used in SL2P are used here but with different distributions based on land cover class listed in Table 13. Radiative transfer model and soil reflectance database for each land cover class.. Canopy parameters for evergreen needle leaf forest.. For each ecozone, a separate calibration database is produced for each of the 11 unique combinations of reflectance model and soil. Canopy parameters vary with land cover class based on literature survey of the bounds of each parameter (Appendix I ).

Table 13. Radiative transfer model and soil reflectance database for each land cover class.

|  |  |  |  |
| --- | --- | --- | --- |
| Land Cover Class | Model | Primary Soil | Special Soils |
| Vegetated | FLIGHT and PROSAIL | Global | Global |
| Evegreen Needleleaf forest | FLIGHT | Forest | Moss |
| Deciduous Needleleaf forest | FLIGHT | Forest | Moss |
| Evergreen Broadleaf Forest | FLIGHT | Forest | Sandy |
| Deciduous Broadleaf Forest | FLIGHT | Forest | Sandy |
| Mixed forest1 | FLIGHT | Forest | Moss |
| Shrubland | FLIGHT | Global | Sandy |
| Herbaceous wetland | PROSAIL | Global | Moss, water |
| Moss&Lichen | PROSAIL | Tundra | Moss, Lichen |
| Bare/sparse vegetation2 | PROSAIL | Global | Tundra |
| Cropland | PROSAIL | Cropland | Sandy |
| Polar shrubland | FLIGHT | Tundra | Moss, Lichen |
| Grassland or pasture | PROSAIL | Global | None |
| Polar grassland | PROSAIL | Tundra | Lichen |

1Mixed forest uses combined databases for all forests.

2Uses grassland or pasture database.

Atmosphere distribution parameters are fixed for all simulations following SL2P (Atmosphere parameters.).

Table 14. Atmosphere parameters.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Lower Bound** | **Upper Bound** | **P1** | **P2** | **Nb\_Class** | **Variable** |
| **P (mbar)** | 0.5 | 3.5 | 0.5 | 3.5 | 1 | Uniform |
| **t550** | 950 | 1080 | 950 | 1080 | 1 | Uniform |
| **H2O (cm)** | 0 | 0.8 | 0 | 0.8 | 1 | Uniform |
| **O3 (dbs)** | 1 | 6.5 | 1 | 6.5 | 1 | Uniform |

# Process Flow

This section describes the SL2P-D calibration and cross-validation process as implemented in the MATLAB prototype script using the in-line documentation. The SL2P-D application process is beyond the scope of this document.

The SL2P-D processor is executed using a Version\_Name and optional Overwrite and Debug arguments. Version\_Name is the name of a .\xls file including, if need be, a path. The processor returns status 1 if executing nominally or zero otherwise.

Overwrite must be set to a value of 1 if existing calibration data or regressions for classes are to be overwritten.

Debug must be set to a value of 1 if nominal debug output is required.

% Simplified Level 2 Processor D

% status = SL2PD(Version\_Name,Overwrite Debug) returns status of SL2PD

% when provided with a parameter file corresponding to Version\_Name.

%

% Inputs

%

% Version\_Name : Name of Processor Parameter File.

% {Overwrite} : 1 overwriting of existing class database or regressions

% {Debug} : Debug level, 1 for debug output

%

% Outputs

%

% status : 1 no exceptions, 0 otherwise

%

% Created: Fred Baret and Marie Weiss, January 2006

% Modified: Richard July 2019

The processor must be executed in the same directory as the source code and ancillary data as it currently operates using uncomplied MATLAB source.

%% Initialisation of Matlab Environment

% .\CODE holds matlab source code and other libraries and executables

% ..\Tools holds generic MALTAB source code and libraries

%

% .\DATA holds data files requires by files in .\code

% ..\Filtres\_Smac\_SL2PD.mat - spectral response functions and smac coefficients

% ..\struct\_Orbito\_Sensor.m - code that generates orbital parameters

%

A directory containing both all output data and figures is created and the regression algorithm to be applied and validation database is identified. A regression algorithm must be given since it defines the required outputs in the calibration database. Validation results are uniquely identified by both the regression algorithm and validation database name. The validation results are not produced if the regression algorithm name or validation database name is omitted.

%% Definition of Dataset Name ,Regression Algorithm, Validation Database

% Read options for current simulation

function Def\_Base = Read\_Start\_Data(Version\_Name)

%% Read current simulation options

% Richard July 2019

% Read global description of calibration database

function Def\_Base = Read\_Learning\_Data(Def\_Base)

%% Déclaration de la structure permettant de caractériser l'ensemble de la

% base d'apprentissage pour le développement de réseaux de neurones.

% Fred 12/12/2005

% Modif Fred Septembre 2007

% Modif Avril 2008

% Fred Mai 2008

% Richard July 2019

The regression algorithm parameters are read. All regression algorithms must have the same outputs.

function Def\_Base = Read\_Alg\_Archi(Def\_Base,Debug)

% Def\_Base = Read\_Alg\_Archi(Def\_Base,Debug) defines an algorithm.

%

% Inputs

% Def\_Base : parameter definitions

% Debug : Debug level, currently 0 for no debug or 1 for debug output

%

% Outputs

% Def\_Base : parameter definitions , will include a new algorithm if it can

% be created. Else, the algorithm definition is null.

% Richard July 2019

If a Sensor sampling Law does not already exist with sufficient samples the Sensor geometry is randomly sampled to produce sufficient simulations for the largest calibration database required. T

%% Create sensor sampling law for maximum number of simulations if it does not exist

% Verify the size of the created Sensor sampling Law is same as maximum

% number simulations

Sensor geometry sampling is performed by parsing the ‘Sensor’ information that specifies the Sensor as well as the geographic and temporal extents of the sampling domain. The sensor spectral response function database is read.

function Def\_Base=Read\_Observations(Def\_Base)

% Define sensor conditions including spectral response functions

% Fred 06/03/2007 (from Fred & Kathy 21/12/2005, modifié Marie, Novembre 2006)

% Fred 07/09/2009

% Richard June 2019

The orbits are randomly sub-sampled without replacement for the maximum number of simulations and the resulting sampling Law is saved in the Report Directory root.

function Law=Create\_Law\_Obs(Def\_Base)

%% Sampling of orbital geometry

% Sampling is exhaustive for point locations, time intervals or

% configurations and random for ranges of each quantity.

% Fred et Marie 28/11/2005

% Modif Fred Juillet 2007

% Richard June 2019

The current regression network and sampling law are saved.

%% Save definition and sensor sampling Law globally

If classes are not to be overwritten the simulation database for each class is read from the Report directory. A new simulation database is produced for a classes if classes are to be overwritten or the simulation database does not exist for a class.

%% Loop through all classes

%% ensure class database if present

%% load current database if it exists and overwrite not specified

%% Produce this class

The production of a new simulation database requires (re)creating the Class sub-directory and adding a definition of the class from the ‘Canopy\_Atmosphere’ worksheet to the global definition Def\_base.

% Make directories for this class

% Création des distributions de la base d'apprentissage

function Def\_Base = Read\_Canopy\_Atmos(Def\_Base,Class)

%% Definition of class parameters for simulation database creation.

% Fred 12/12/2005

% Modif Fred Septembre 2007

% Modif Avril 2008

% Fred Mai 2008

% Richard July 2019

The soil library reflectance and sensor spectral functions are plotted and saved in the Class sub-directory. SL2PD requires all classes share the same soil library and sensor specifications.

% graphiques de définition du sol et des capteurs

function Plot\_Sol\_Bandes(Def\_Base)

%% Edition des figures de spectres de sol et de sensibilité spactrale des bandes

% Fred 04/10/2005

% Modif Fred Aout 2007

% Modif Fred Avril 2008

% Richard June 2019

The updated global definition is vaved in the the root directory and in the class sub-directory to allow for later verification that the global and class specific definitions match.

% save definition for this class and globally

The sampling Law for canopy and atmosphere parameters for the class is produced and saved.

% create and save canopy and atmsopeher parameter sampling Law for this class

function [Law,Nb\_Sims]=Create\_Law\_Var(Def\_Base,Class,Soil,Law,Nb\_Sims)

% Création des distributions de variables d'entrée SAIL+SMAC pour la

% création de la base d'apprentissage

% Fred et Marie 01/08/2005

% Modif Fred Avril 2008

% Modif Marie septembre 2010: ajoût de la loi log normale, modification des

% contraintes sur les lois de co-distribution

% Richard July 2019

The simulated input-output is produced for each sample in the sampling Law using the specified RTM. If copyFlag is set to one, samples in the global database will be reused in which case the Sensor geometry Law will correspond to the copied sample so the Sensor laws are now randomly sampled with replacement rather than without replacement. If this is not desired copyFlag should be set to zero.

% simulate using Law and specified RTM

function [NewInput,NewOutput,NewLaw] = Build\_Database( Def\_Base, Law, Class, copyFlag)

%% builds database for a class by either copying or perfoming simulations

% Richard July 2019

The laws are clustered into clusters of size ~100 based on Law.ALA Law.Cab Law.N Law.Cdm Law.Cw\_Rel Law.Bs for use during batch training and cross-validation and the new global definition is saved within this class and in the root directory.

% cluster Laws for block training

The noise free simulation database is visualized.

% plot the inputs and outputs for this class

function Plot\_Matrix\_InOut(Def\_Base, Law, Input, Output, Class)

%% Edition des distributions et co-distributions des variables In et Out

% Fred Septembre 2009

% Marie, septembre 2010: matrice des résidus (bruit - valeur nominale) en

% fonction des valeurs nominales

% Richard July 2019

If a regression algorithm is specified an attempt is made to load the algorithm from the class sub-directory. If the attempt fails but the algorithm is a valid method the algorithm is calibrated using the calibration database for this class. The convex hull of the inputs and outputs used for calibration is determined and saved. Otherwise, an exception is indicated that the regression algorithm was invalid and execution returns with status=0.

%% calibrate and/or validate regression algorithm if requested

% parse regression method name

% define regression method

% check if algorithm already exists for this class

% if method does not exist for this class calibrate new

% add noise to the input data

% Determine convex hull of inputs used for this network

% plot the noisy inputs amd outputs for this class and regression

% check the regression method against available methods

% method NNT selected

% Check if single or cacsading regression

% calibrate cascading regression

% save the resulting performance

% save the resulting calibrated regression

Cross validation is attempted after regression calibration is completed. If unsuccessful an exception and execution returns with status=0.

%% perform cross validation

% load in cross validation database

% add noise to the input validation data

% determine input out of range flag

% do validation and plot results

% try to load results from other validation

% validate and plot results

% calibrate and plot regressions for incertitudes

% estimate rmse for sims sharing same inputs

% calibrate regression for each incertitude and plot results

% save cross validatio

# Performance

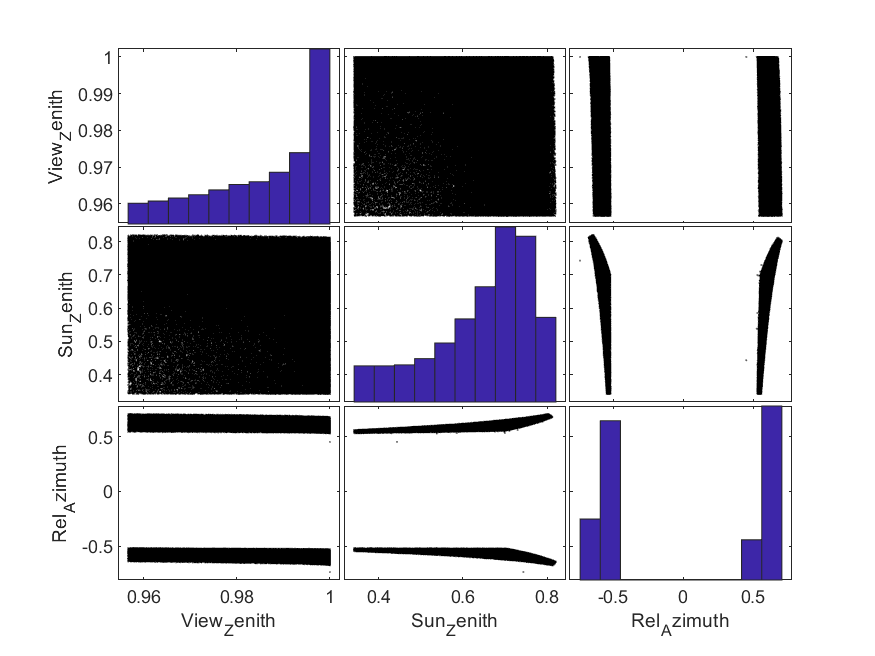
The performance of the regression models is evaluated using cross-validation with the calibration database and then validation with a test database. The test database corresponds to a Monte Carlo sampled database using the same canopy architecture distributions as the calibration database. Monte-Carlo sampling ensures that the validation is applied to clumps that appear in a different pattern than the calibration database (Figure 2). The nominal SL2P model is also implemented and applied to each test database for reference purposes. Due to the large number of comparison combinations only a single result over Cropland and over Polar Grassland classes are shown. The former corresponds to a relatively broad range of canopy architecture and acquisition conditions while the latter corresponds to a much more specific range. In addition to the SL2P-D for each class, results for the nominal SL2P are included.

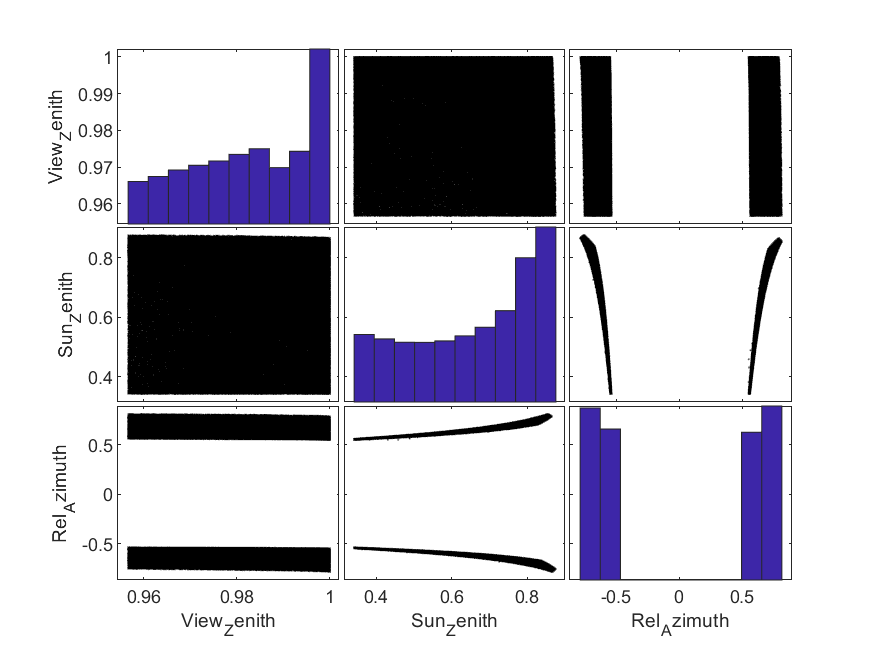
Geometric Laws

Geometric Laws (Figure 3) are controlled by the geographic extent and data range of retrievals for a database associated with an Ecozone (Table 9). Modal angles are essentially unchanged. The range of SZA is substantially greater than VZA so differences in VZA between databases are unlikely to impact performance. However, the larger variance of SZA for ecozone specific databases in comparison to SL2P suggests that they may be more difficult to calibrate but that SL2P will perform worse for samples taken from these databases.

Canopy Laws

Canopy Laws (Figure 4) are user specified as truncated marginal distributions whose bounds are constrained by LAI. The truncation scheme used by SL2P has been corrected so results from our implementation of SL2P will differ somewhat from Weiss and Baret (2016). Both the ecozone Laws indicate generally uniform distributions but over narrower ranges than SL2P. This follows the philosophy that SL2P-D should use physics and only priors regarding the range of possible variables to constraint retrievals; in contrast SL2P uses a mix of both strategies. The Polar Grassland Laws provide an extreme case of very narrow ranges for variables such as LAI, ALA and Bs. All Laws assume that leaf chlorophyll and moisture content lies in the upper half of the possible range as LAI tends to the maximum LAI for a Class. While SL2P uses orthogonal sampling and SL2P-D uses Sobol sampling the difference is not evidence in Figure 4 due to the large number of samples.





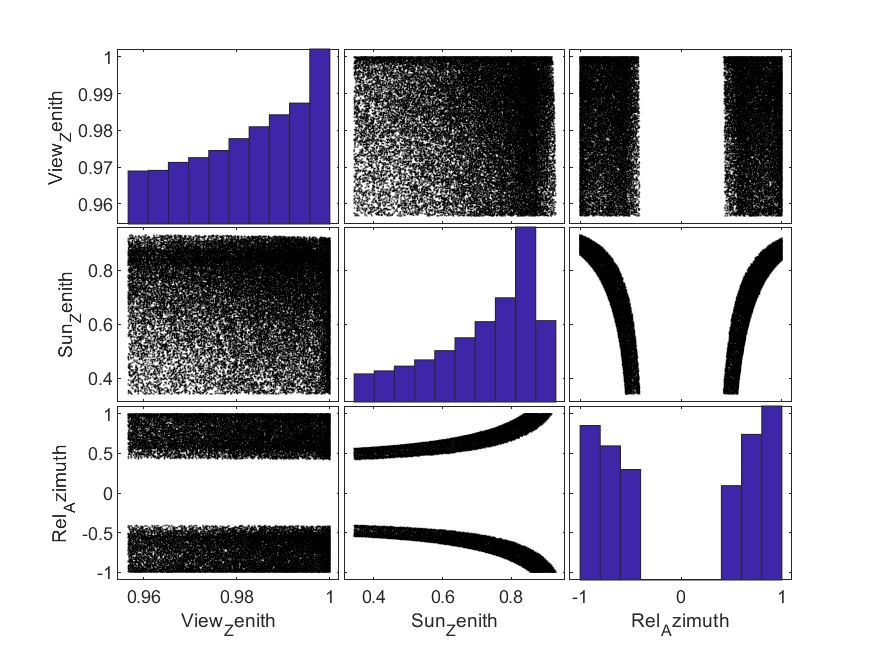
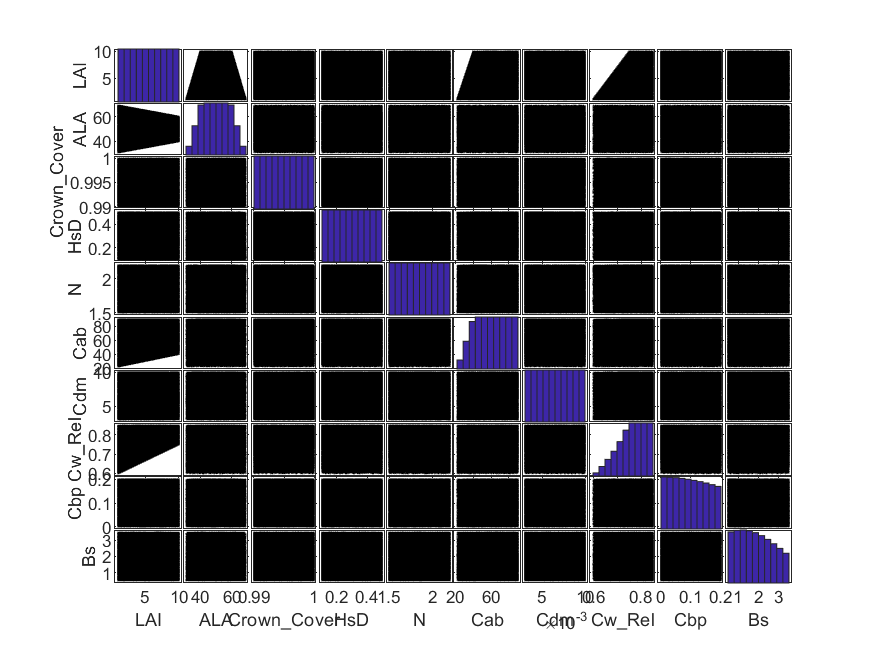
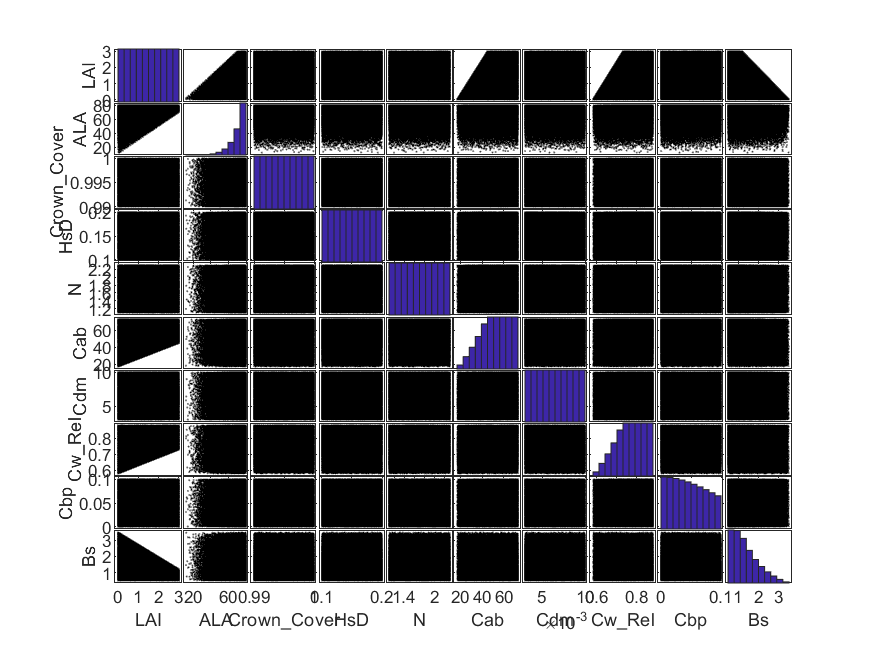


Figure . Geometric Laws for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (middle) and SL2P (lower).



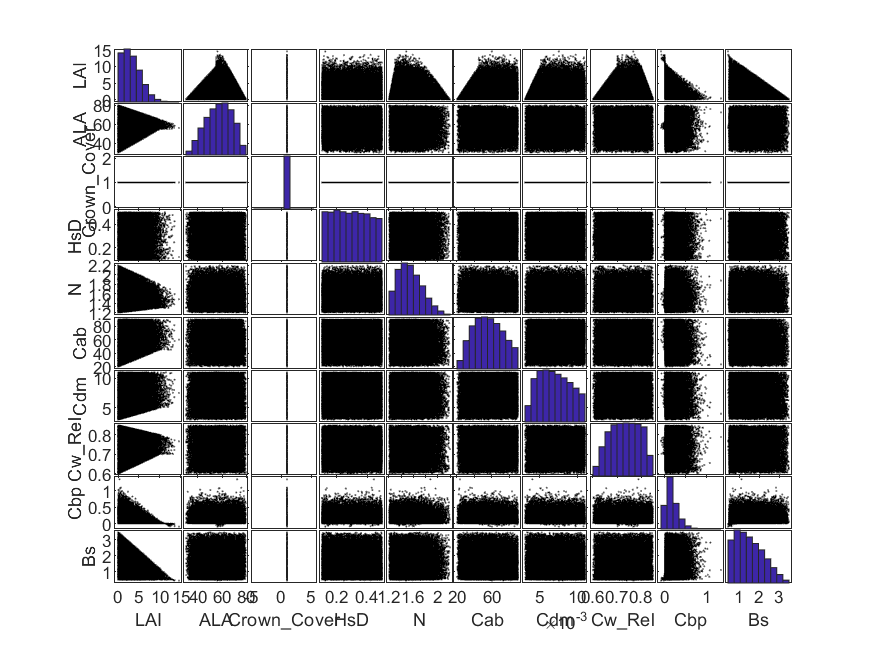


Figure . Canopy Laws for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (middle) and SL2P (lower).

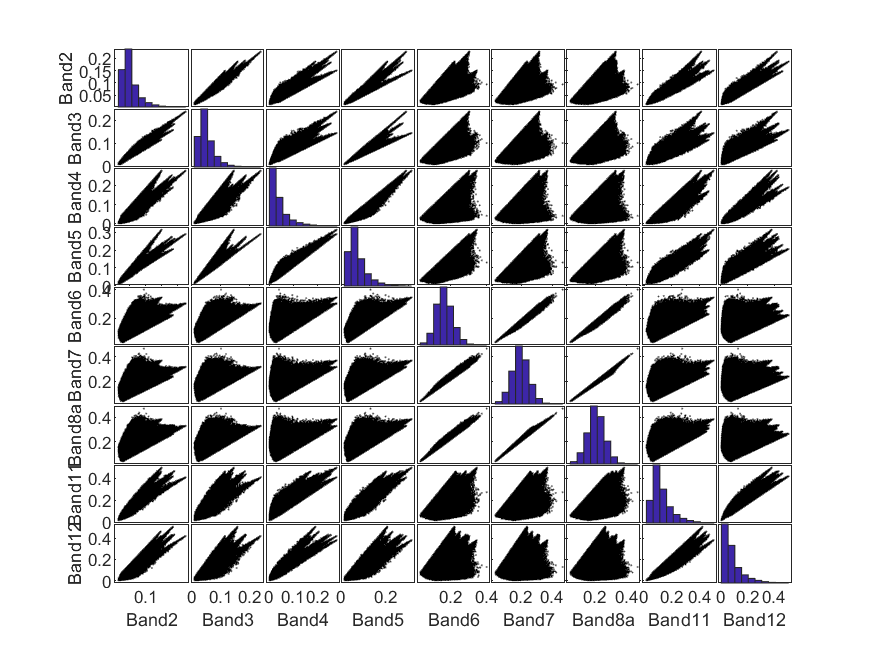
## Regression Input Variables

The input variables correspond to a combination of provided Geometric Laws and R derived from the RTM given the Laws and sensor characteristics (Figure 5). The joint distribution of multispectral reflectance as a function of these Laws has been extensively studied and documented elsewhere. It is sufficient to note that there is substantial redundancy of inputs and that the nature of the joint distribution between bands is non-trivial in many cases. Strategies based on Class and even image specific input band reduction using linear or non-linear band combinations may be effective to increase the robustness of retrievals in the presence of noise.

## Regression Output Variables

Regression output variables (Figure 6) correspond to canopy variables or derived quantities as well as an associated prediction RMSE. The prediction RMSE output variables corresponds to the validation RMSE for each variable based on a provided validation database.

The output canopy variables =correspond to a combination of provided Canopy Laws (e.g. for LAI) and quantities derived from the RTM using input Laws and sensor information. All variables other than Albedo indicate a relatively high level of pairwise correlation that is even more evident when examining density plots ( ) versus the scatter plots shown in (xx). Moreover, as the density plots and histograms in Figure 6 indicate, both the marginal and joint distribution of these variables are generally skewed. This means that that regression calibration will likely be more precise near the modal values of these distributions and will suffer at the tails – such behaviour will likely be exacerbated by SL2P which has non-uniform distributions for most Laws that drive these variables. Active learning strategies may be required if the purpose of the produced estimates is to track disturbances that may correspond to tails of these output distributions.



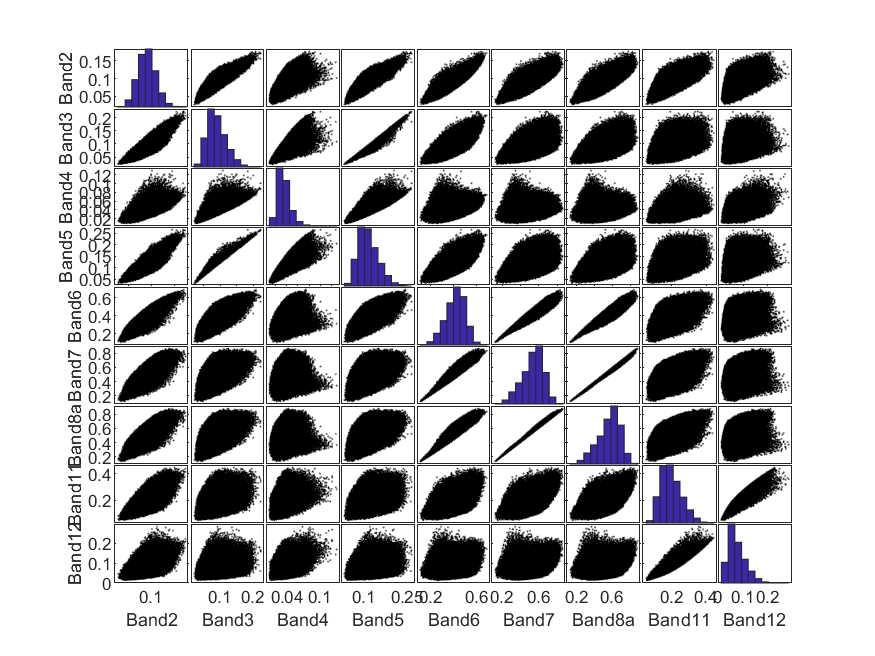


Figure 7. Input R for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (lower)

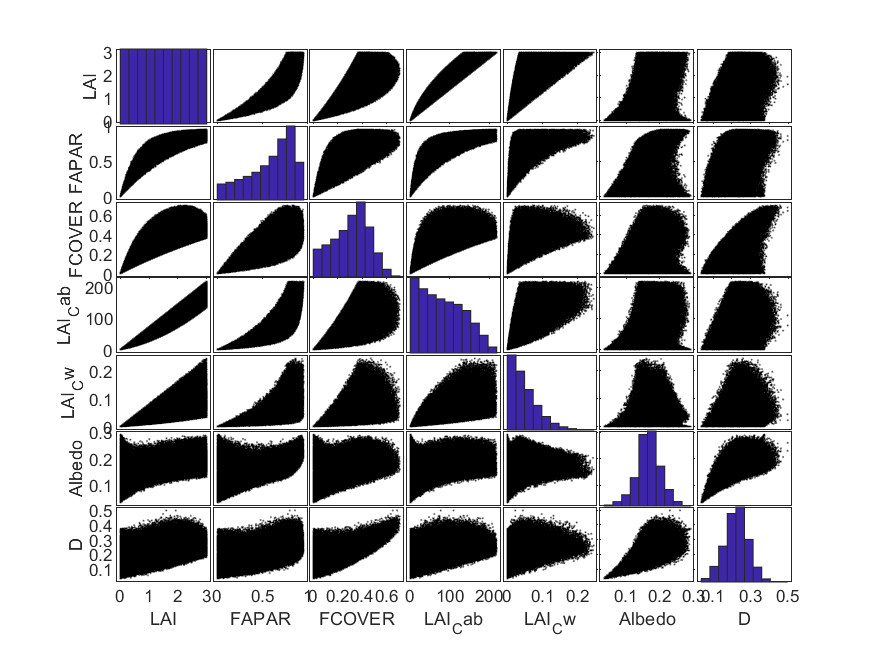
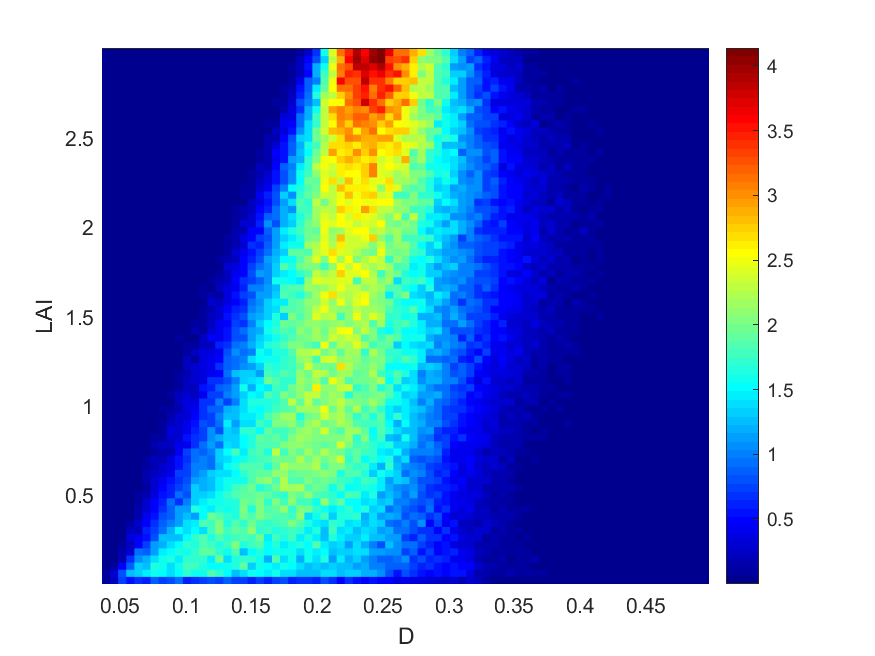
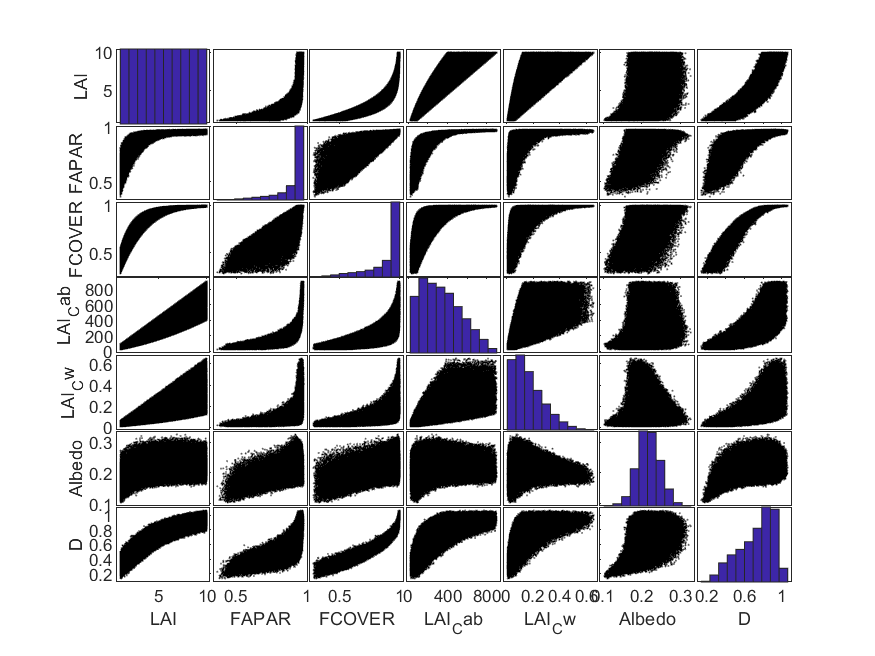
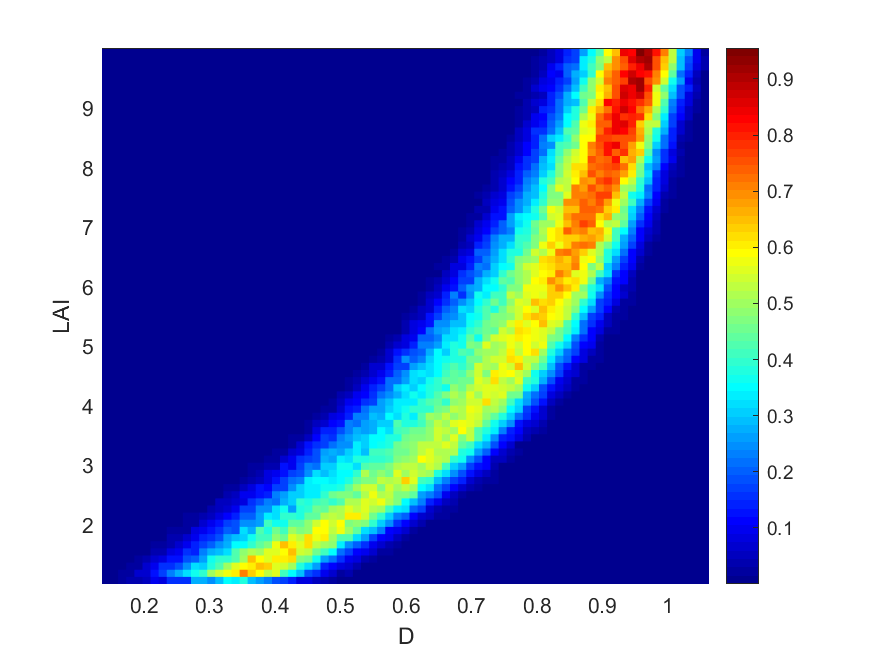


Figure 8. Output variables for Taiga Cordillera Polar Grassland (upper), Prairie Cropland Close (lower).



## Validation Performance

The overall RMSE was determined for each variable using both SL2P-D and SL2P but is of limited use considering that the actual distribution of canopy variables will differ from the validation distribution on a regional basis. For example, the upper panels of each figure indicate that the distribution of a given variable in the validation dataset is often quite clumped so that the overall RMSE may be weighted to typical conditions of the ecozone and Class rather than actual conditions during a retrieval. Rather, the uncertainty and bias for each variable are quantified using RMSE of residuals (solid lines in lower panels) and the expected bias (red line of lower panels) conditional on an estimated value. Furthermore, since some applications may wish to focus on more stringent uncertainty bounds box plots of residuals conditional on estimated values are also provided (lower panel). Finally, we identify a useful range of retrievals corresponding to the interval where the bias is less than the conditional RMSE. For brevity we discuss only the useful range of retrievals compared to the range of validation data, and the conditional RMSE, bias and 95%ile confidence intervals at the mode of the validation data for albedo and LAI or the midpoint of the retrieved range for other quantities. Furthermore we identify results that are within threshold, near threshold, or far from threshold requirements for RMSE (Table 1) or bias (threshold:bias < 33%RMSE; near threshold: bias<66%RMSE;far from threshold:bias>66%RMSE).

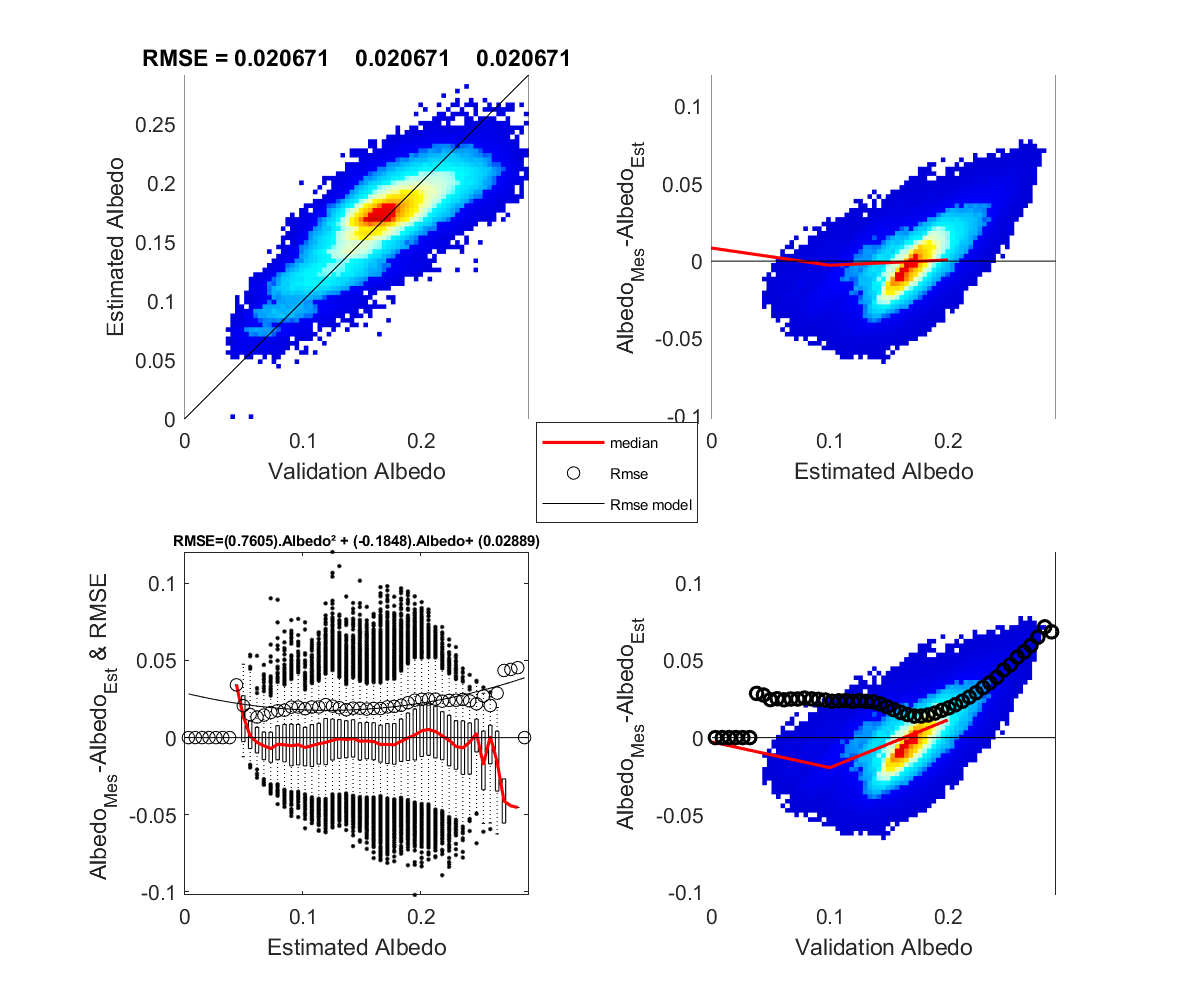
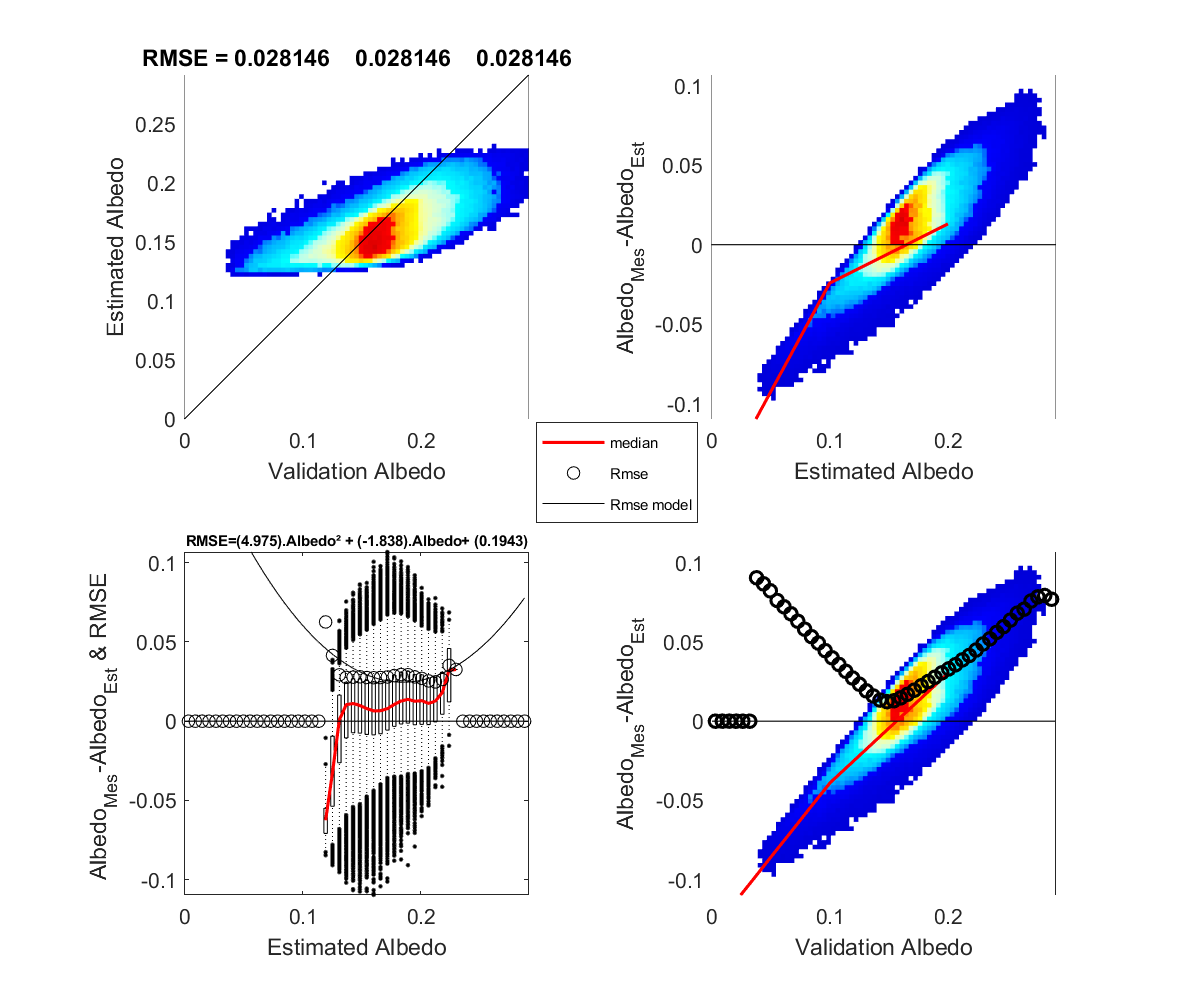
For Prairie Cropland – Closed, SL2PD met threshold requirements of residuals at 67.5%ile interval and bias or all parameters except LAI.Cw; although in most cases the residuals were still near the threshold (Table 18 ;Figure 7,Figure 8,Figure 9,Figure 10,Figure 11,Figure 12 ). In contrast, SL2P was above the threshold requirements for both RMSE and bias with bias far from threshold. Requirements for Albedo, fCOVER and fAPAR. Both algorithms were far from threshold for 95%ile residuals. The range of valid retrievals was similar between algorithms for all parameters but Albedo where SL2P had a very high lower bound of 0.14 (contract with 0.05 for SL2PD). For both algorithms the 67.5%ile residuals were relatively constant within valid ranges (indicating decreasing relative error with increasing variable magnitude) for all parameters except LAI.Cab and LAI.Cw. LAI.CWV and LAI.Cab showed increasing residuals with increasing variable magnitude indicating a constant relative error. SL2PD showed small relatively constant levels of bias within the valid range of retrievals. In contrast SL2P indicated biases increasing with magnitude for most variables and, for fAPAR and variables involving LAI, relatively large changes of sign in the bias. Similar results were observed for Taiga Cordillera Polar Grassland although SL2P improved substantially for Albedo and fCOVER residuals (Table 19; Figure 13,Figure 14,Figure 15,Figure 16). However, the lower bound of valid retrievals increased substantially for SL2P in comparison to SL2PD for fAPAR and fCOVER and both of these quantities also showed increasing bias at low magnitudes corresponding to large relative biases.

Table . Validation summary statistics for Prairie Cropland - Closed. Green:Within Threshold;Orange:Near Threshold;Red-Far from threshold. Range corresponds to reference values within average RMSE.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | SL2PD | | | |  | SL2P | | |
| Variable | Range | RMSE | bias | 95%ile | Range | RMSE | bias | 95%ile |
| Albedo | 0.05,0.25 | 10% | -2.7% | 22% | 0.14,0.23 | 13% | 11% | 38% |
| fCOVER | 0,0.65 | 8% | 1% | 17% | 0,0.60 | 13% | 8% | 32% |
| fAPAR | 0,1 | 5% | 1% | 19% | 0,.1 | 11% | 10% | 20% |
| LAI | 0,2.8 | 16% | -1% | 46% | 0.2.5 | 28% | 16% | 55% |
| LAI.Cab | 0,180 | 20% | -1% | 55% | 0,180 | 35% | 20% | 80% |
| LAI.Cwc | 0,0.17 | 31% | 10% | 85% | 0,0.17 | 30% | 10% | 80% |

Table . Validation summary statistics for Taiga Cordillera Polar Grassland. Green:Within Threshold;Orange:Near Threshold;Red-Far from threshold. Range corresponds to reference values within average RMSE.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | SL2PD | | | |  | SL2P | | |
| Variable | Range | RMSE | bias | 95%ile | Range | RMSE | bias | 95%ile |
| Albedo | 0.11,0.3 | 9% | -1% | 18% | 0.12,0.24 | 6.5% | -3% | 14% |
| fCOVER | 0.25,1 | 7% | 1% | 15% | 0.3,1 | 7% | 4% | 22% |
| fAPAR | 0.38,1 | 5% | -1% | 14% | 0.6,.95 | 7% | 1% | 16% |
| LAI | 0.8,9.2 | 20% | -1% | 60% | 0.8,9.2 | 30% | -5% | 70% |
| LAI.Cab | 0,750 | 14% | -1% | 50% | 0,750 | 33% | 13% | 80% |
| LAI.Cwc | 0,0.46 | 17% | 1% | 65% | 0,0.36 | 50% | 20% | 100% |



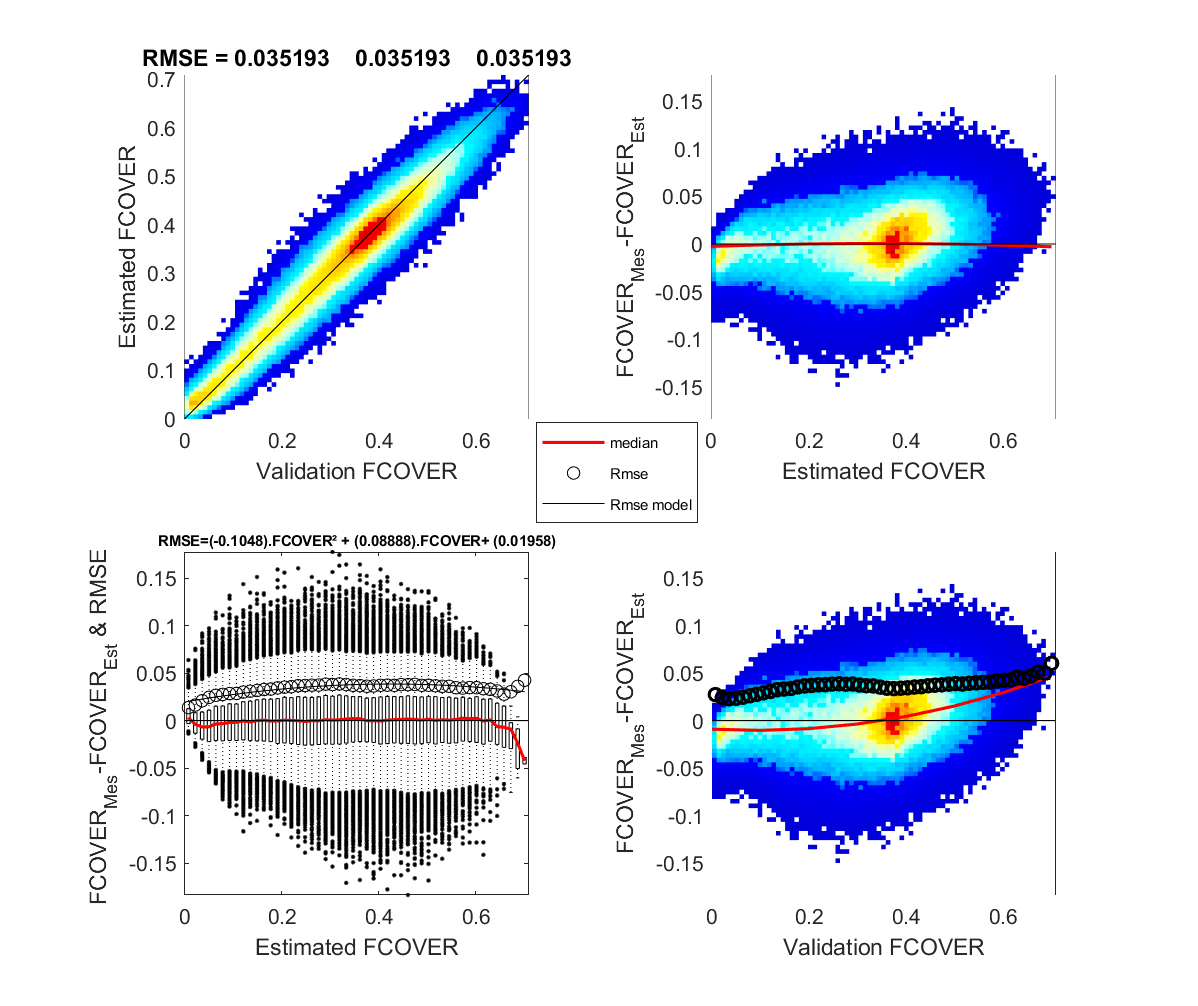
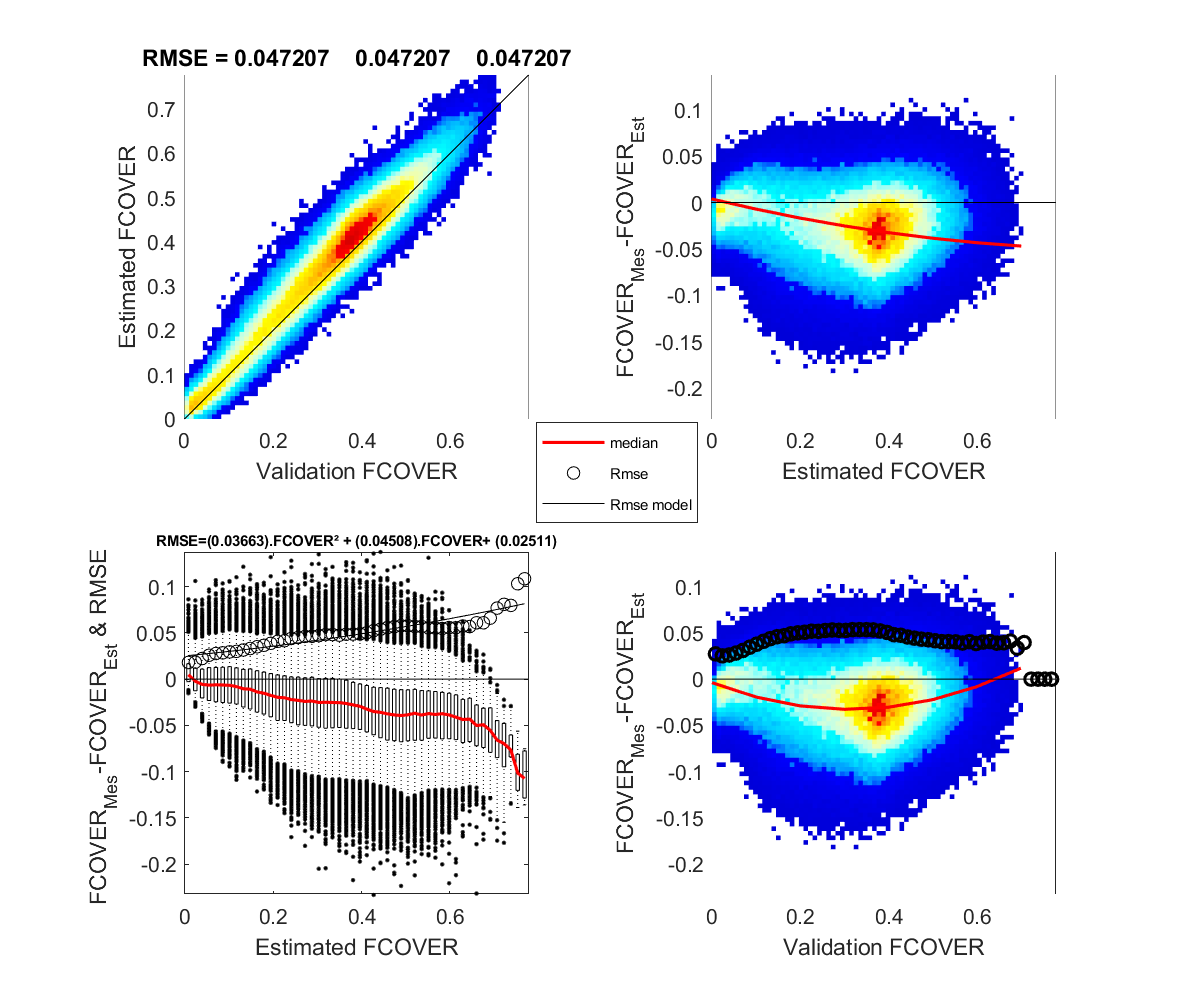


Figure 9. Prairie cropland (closed) fCOVER validation for SL2P-D (left) AND SL2P (right).

Figure 10. Prairie cropland (closed) albedo validation for SL2P-D (left) AND SL2P (right).

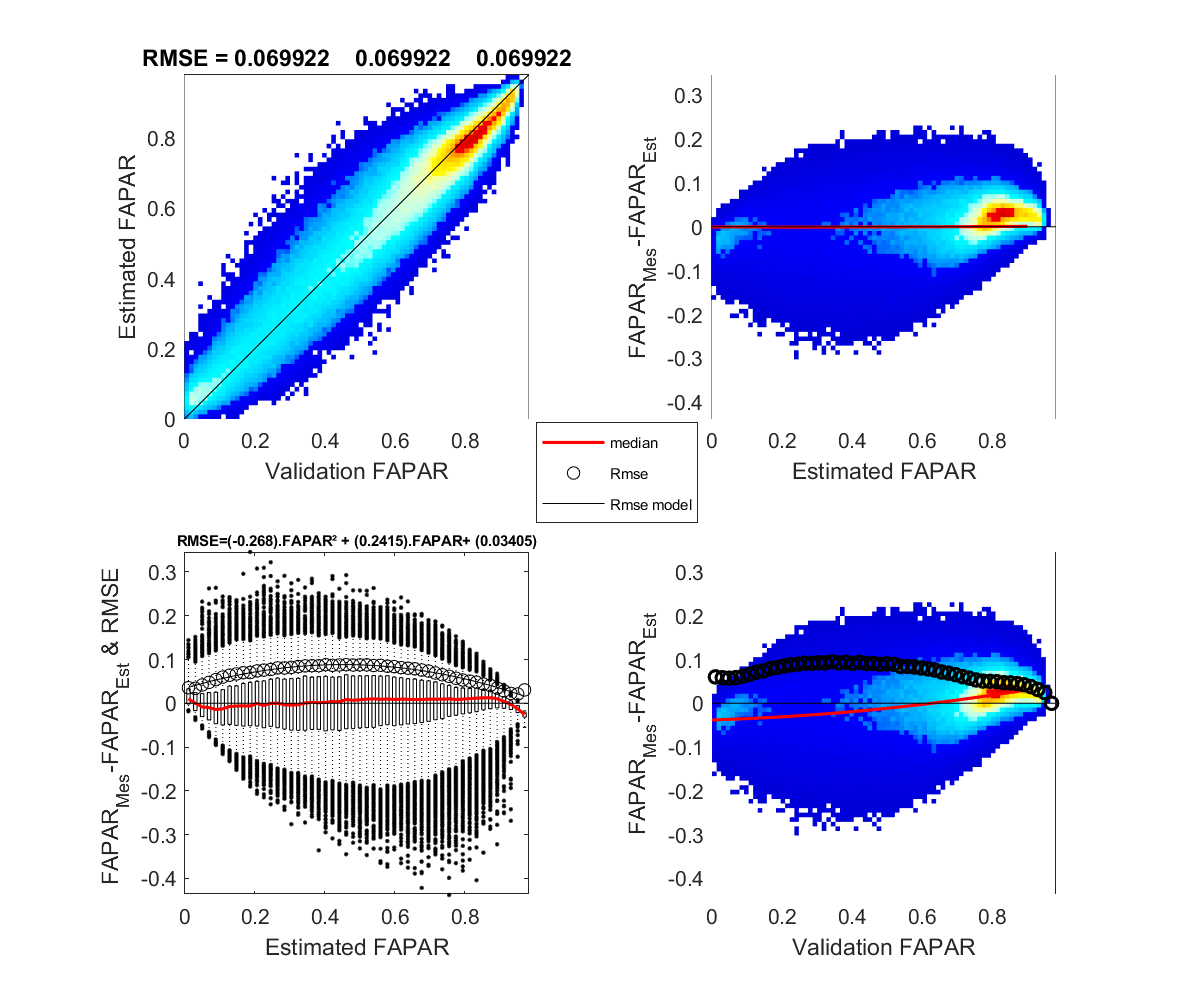
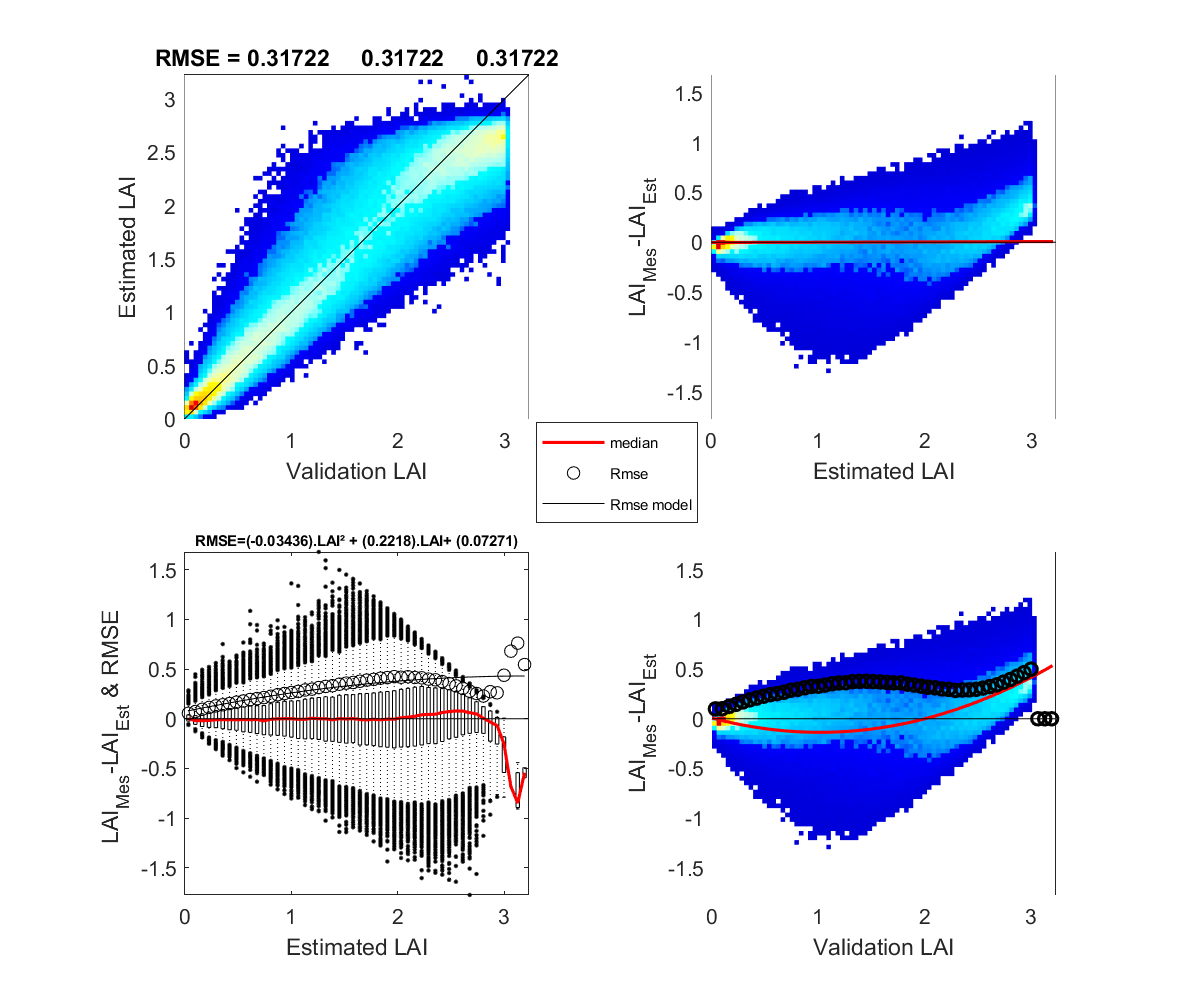
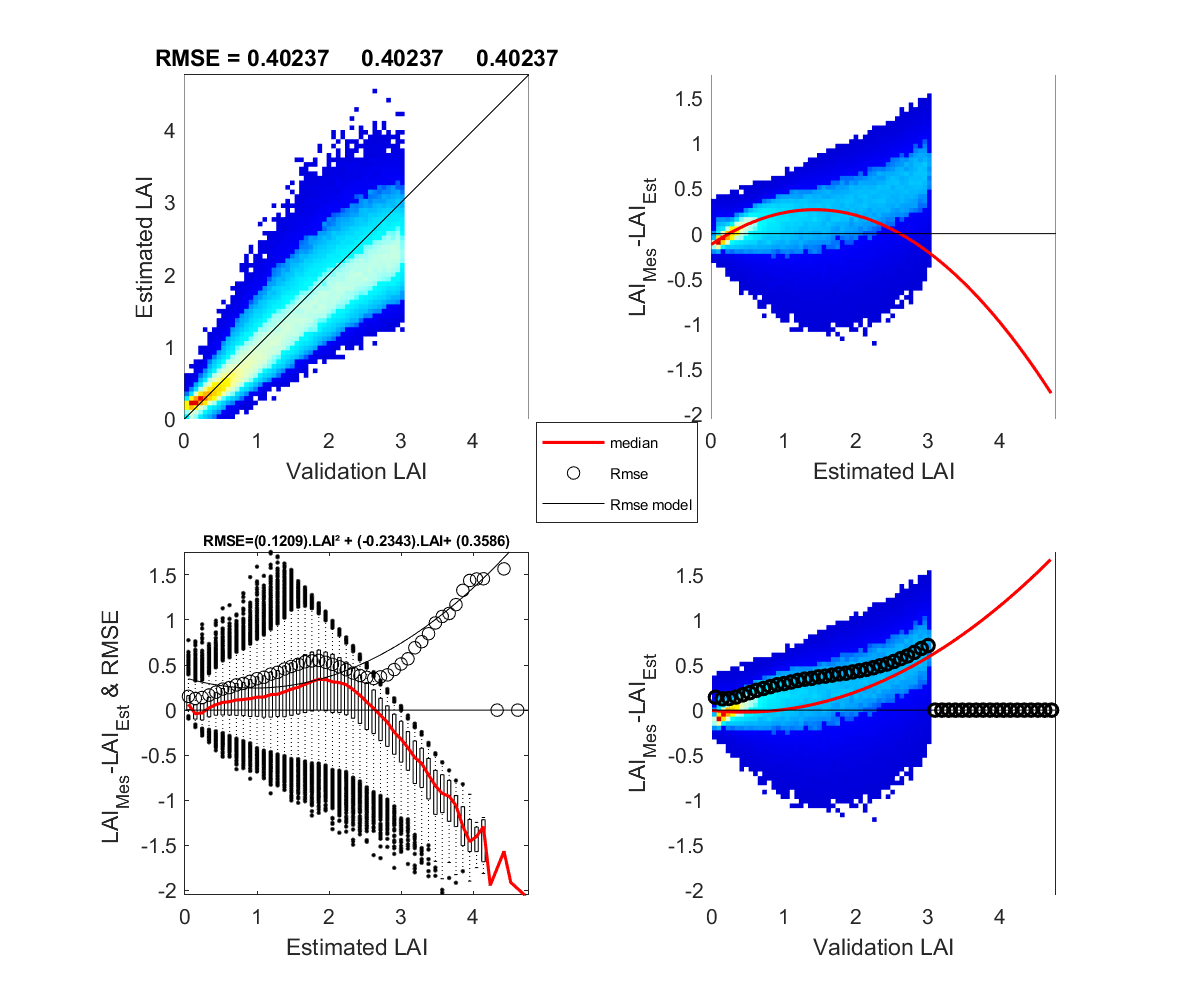
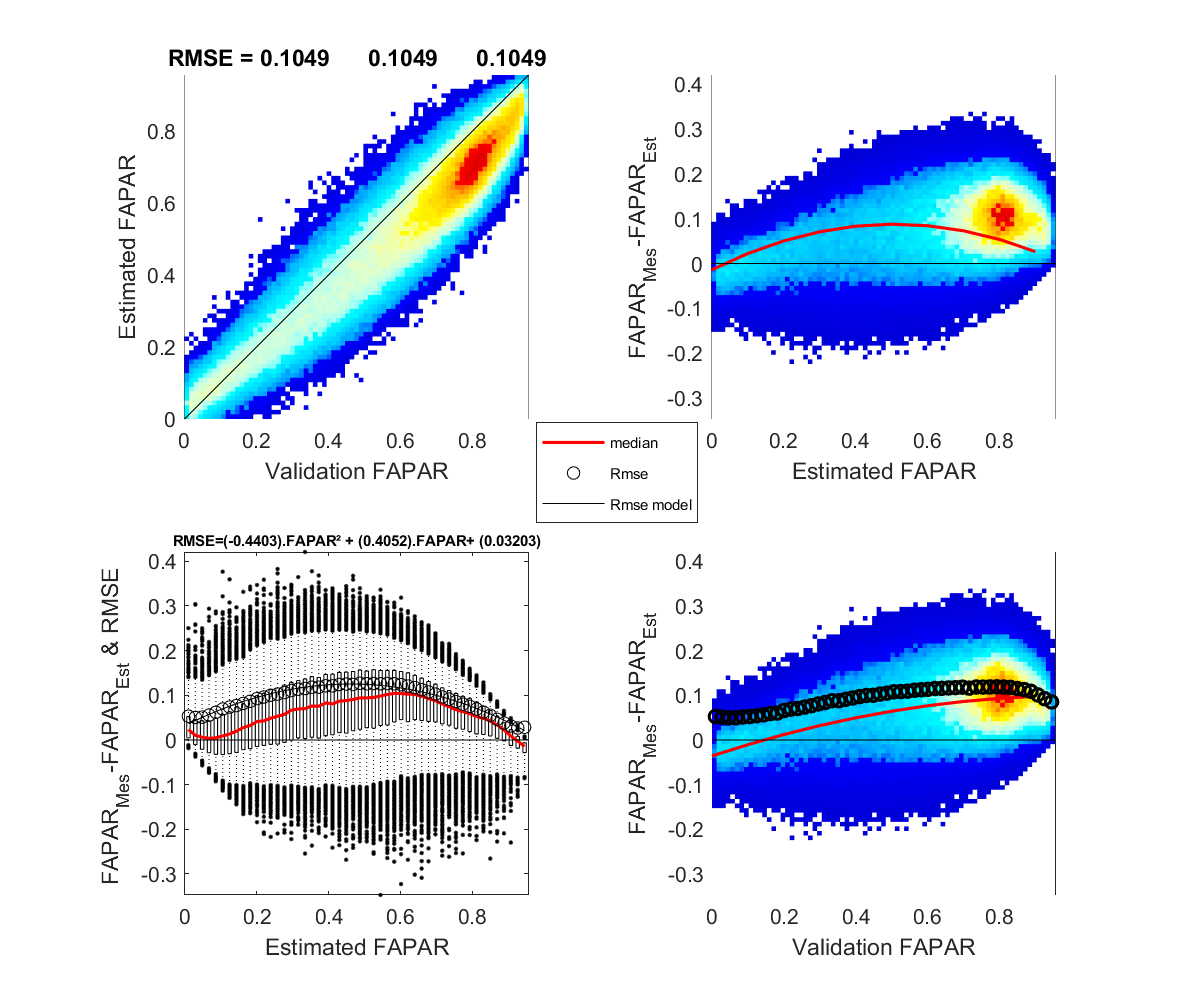


Figure 11. Prairie cropland (closed) LAI validation for SL2P-D (left) AND SL2P (right).

Figure 12. Prairie cropland (closed) fAPAR validation for SL2P-D (left) AND SL2P (right).

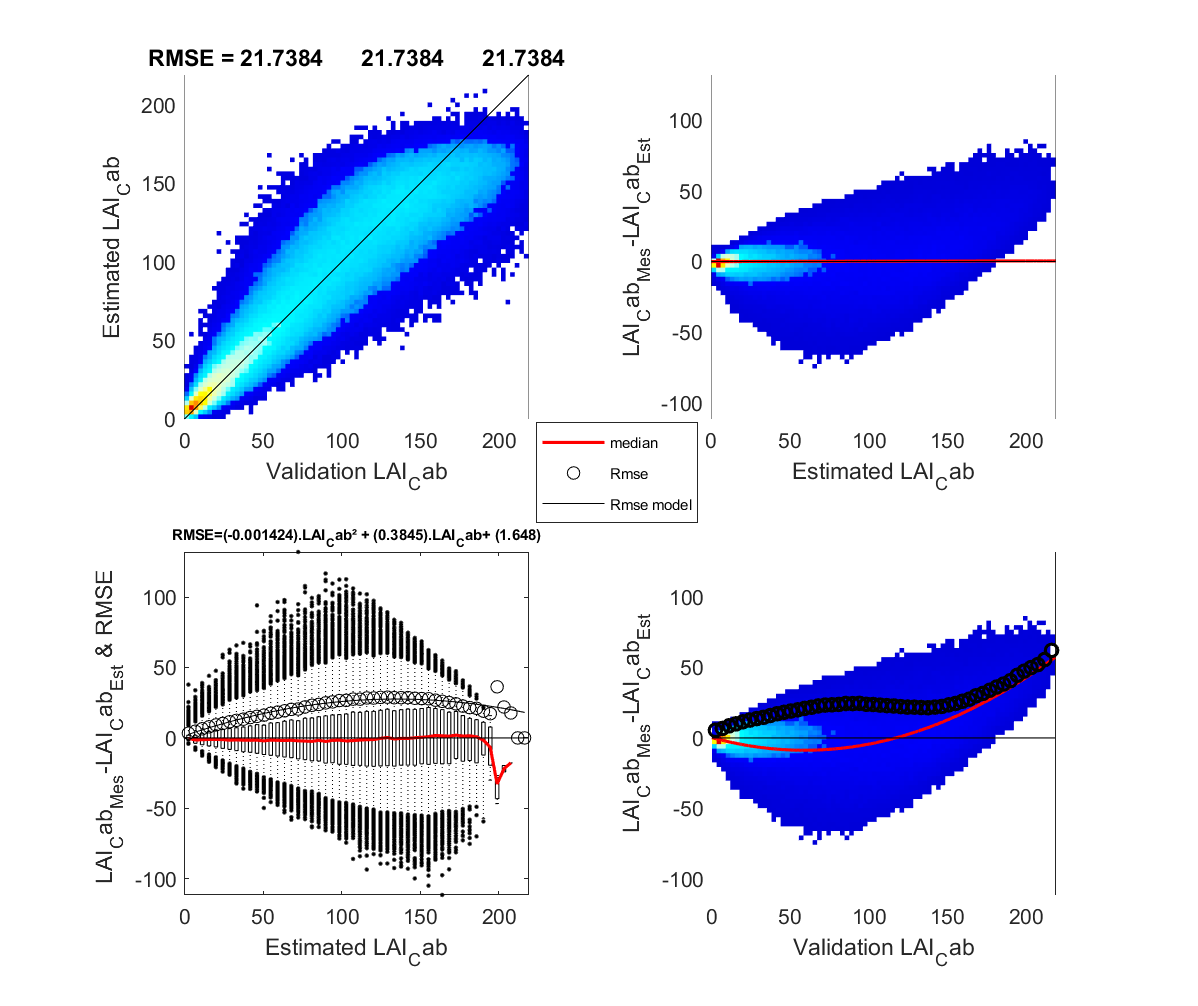
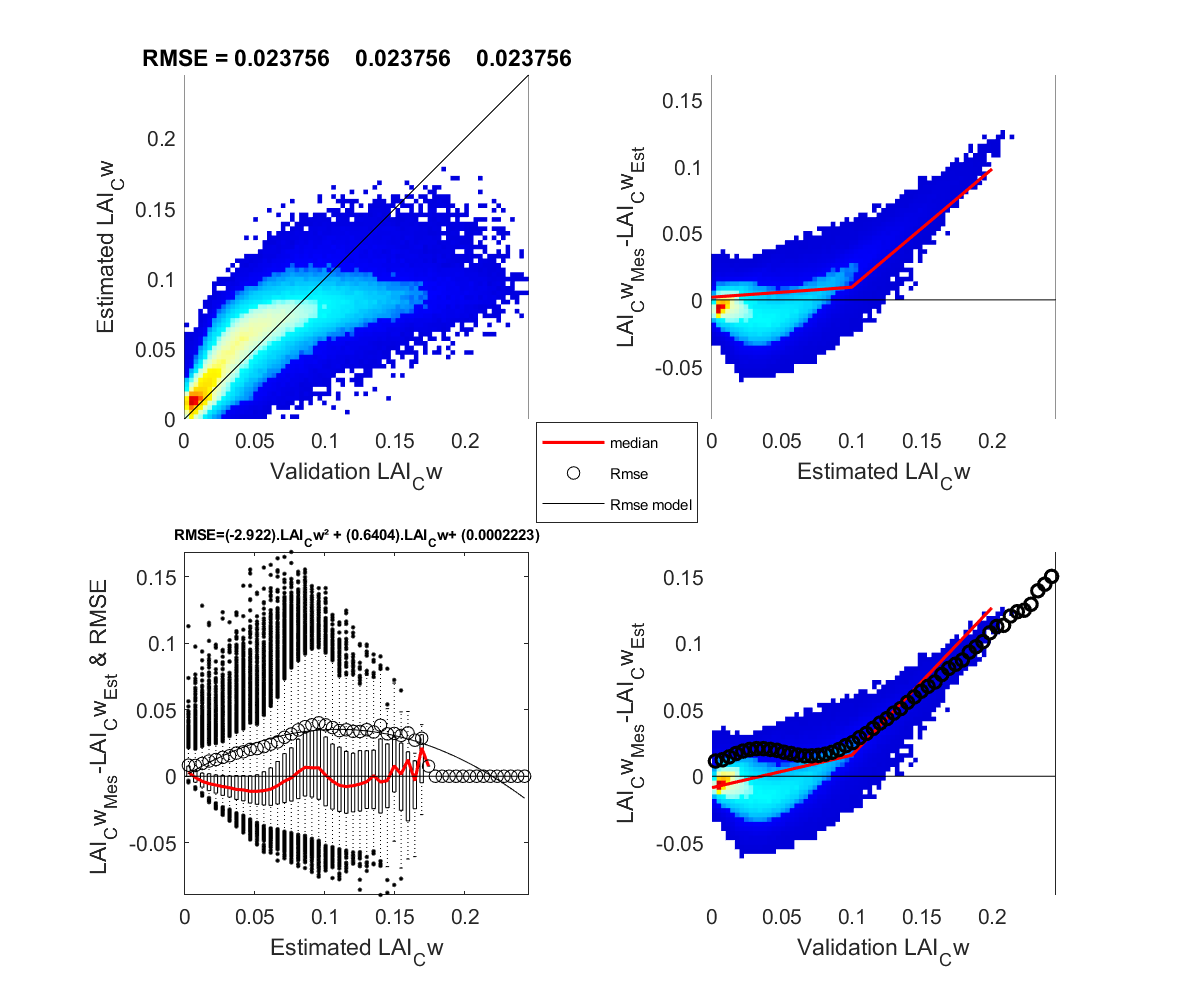
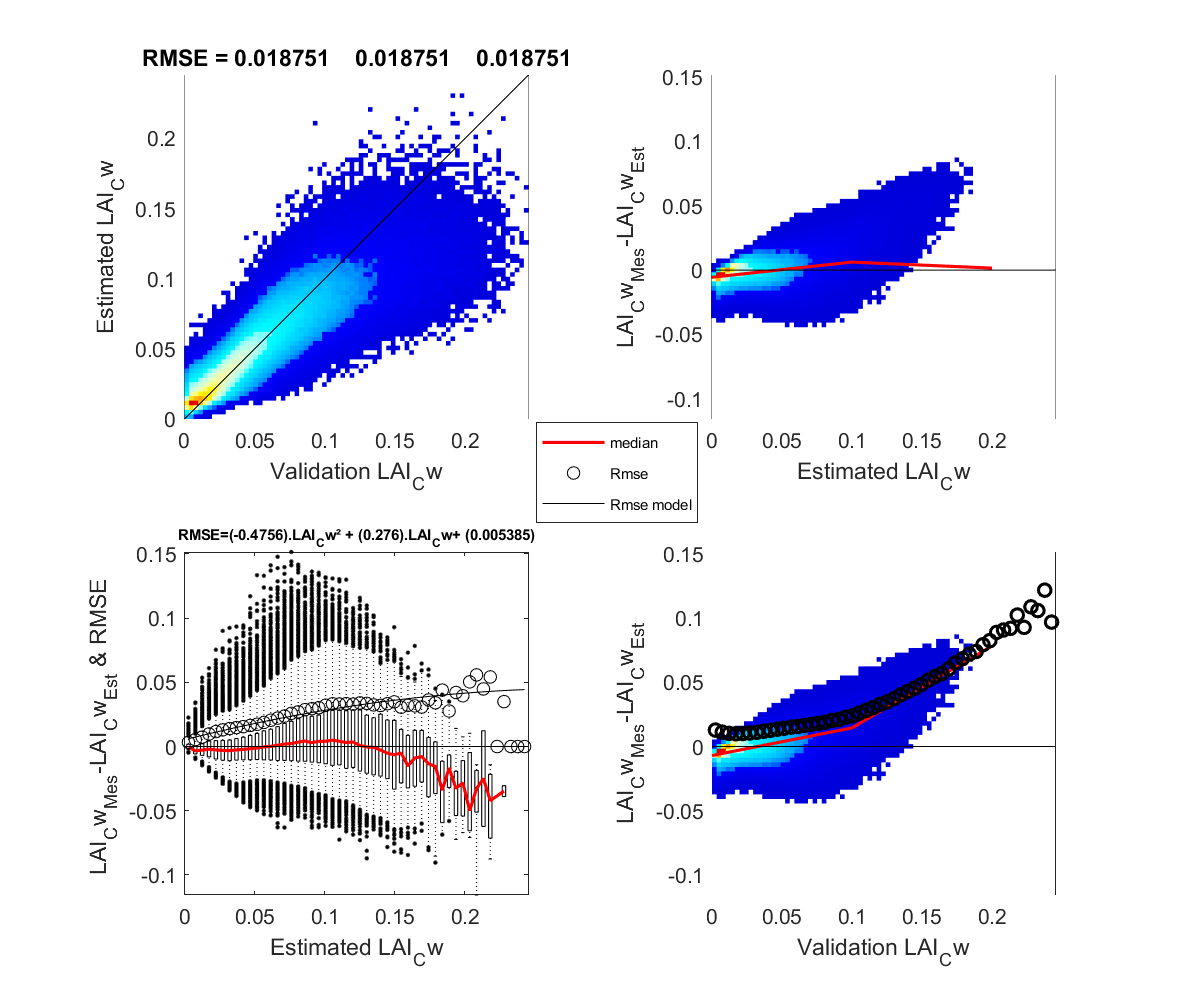
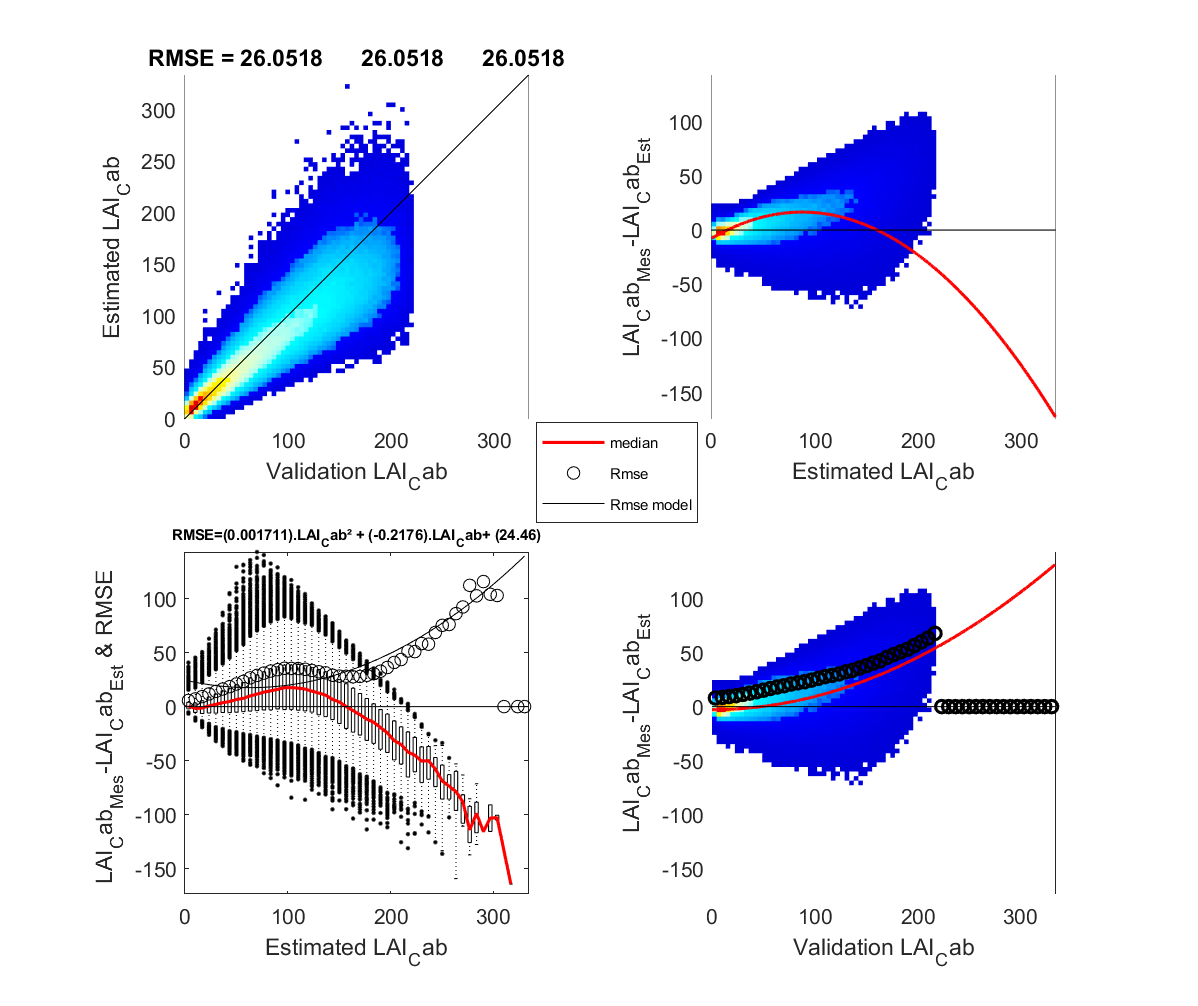


Figure 13. Prairie cropland (closed) LAI.Cw validation for SL2P-D (left) AND SL2P (right).

Figure 14. Prairie cropland (closed) LAI.Cab validation for SL2P-D (left) AND SL2P (right).

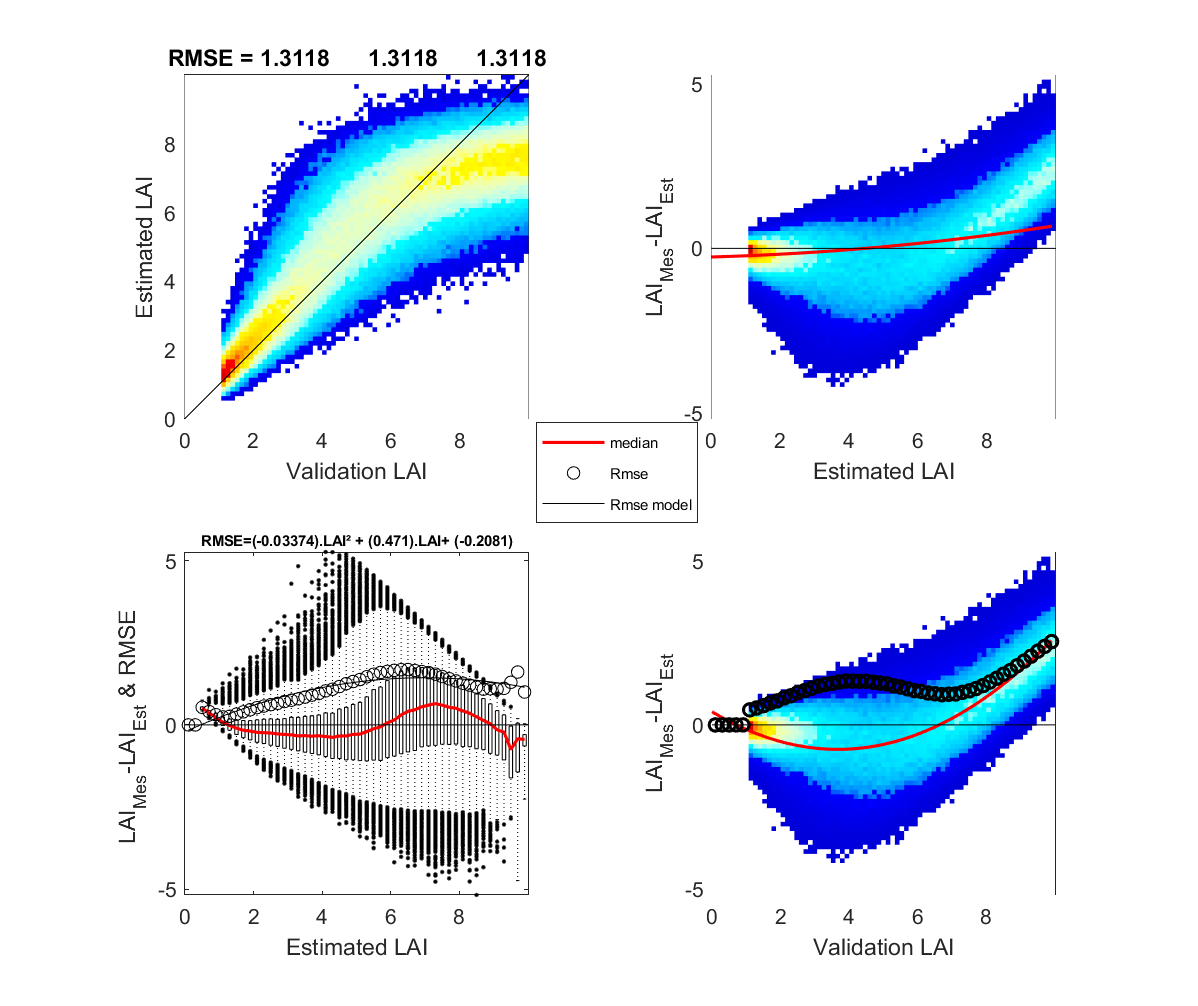
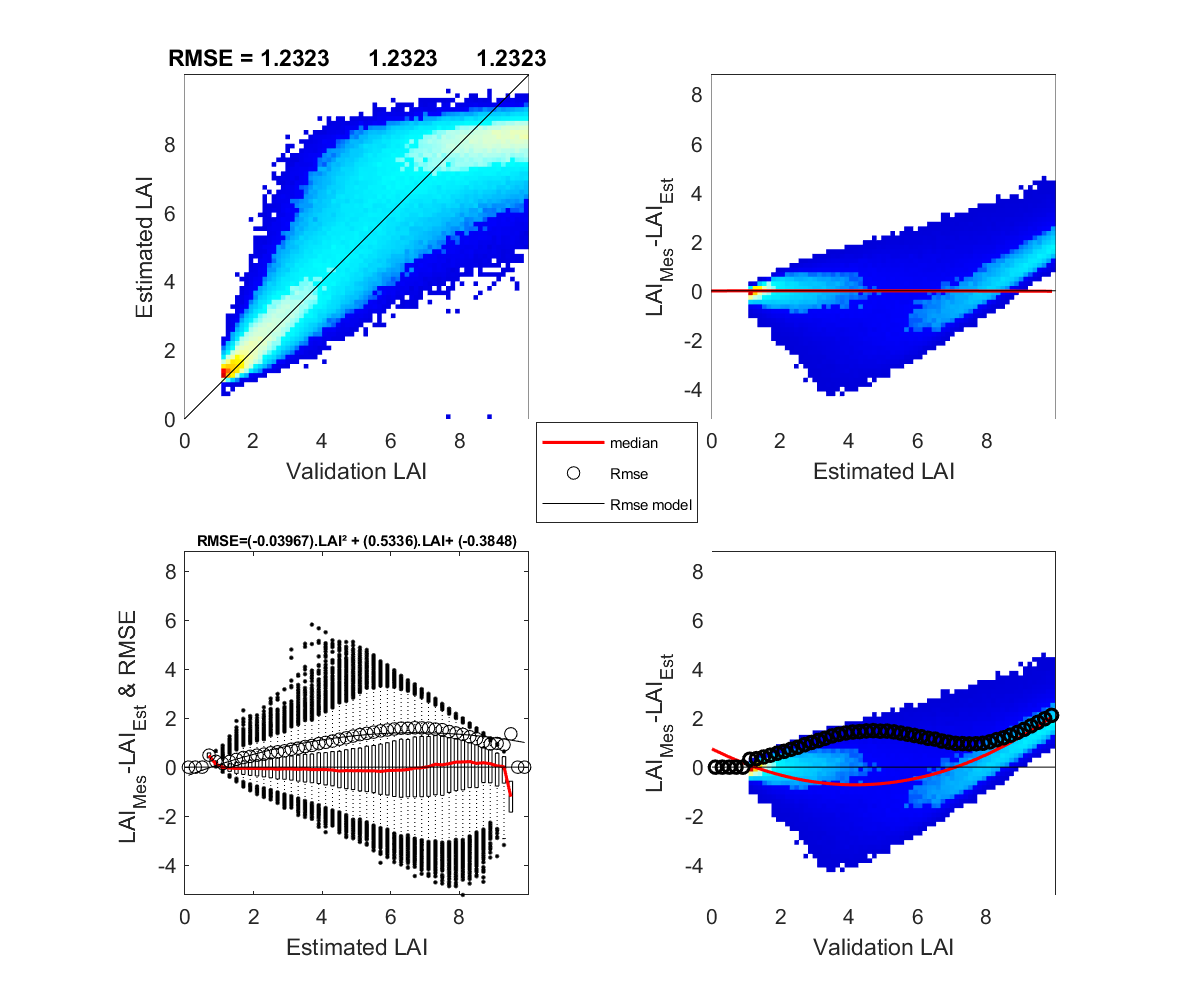
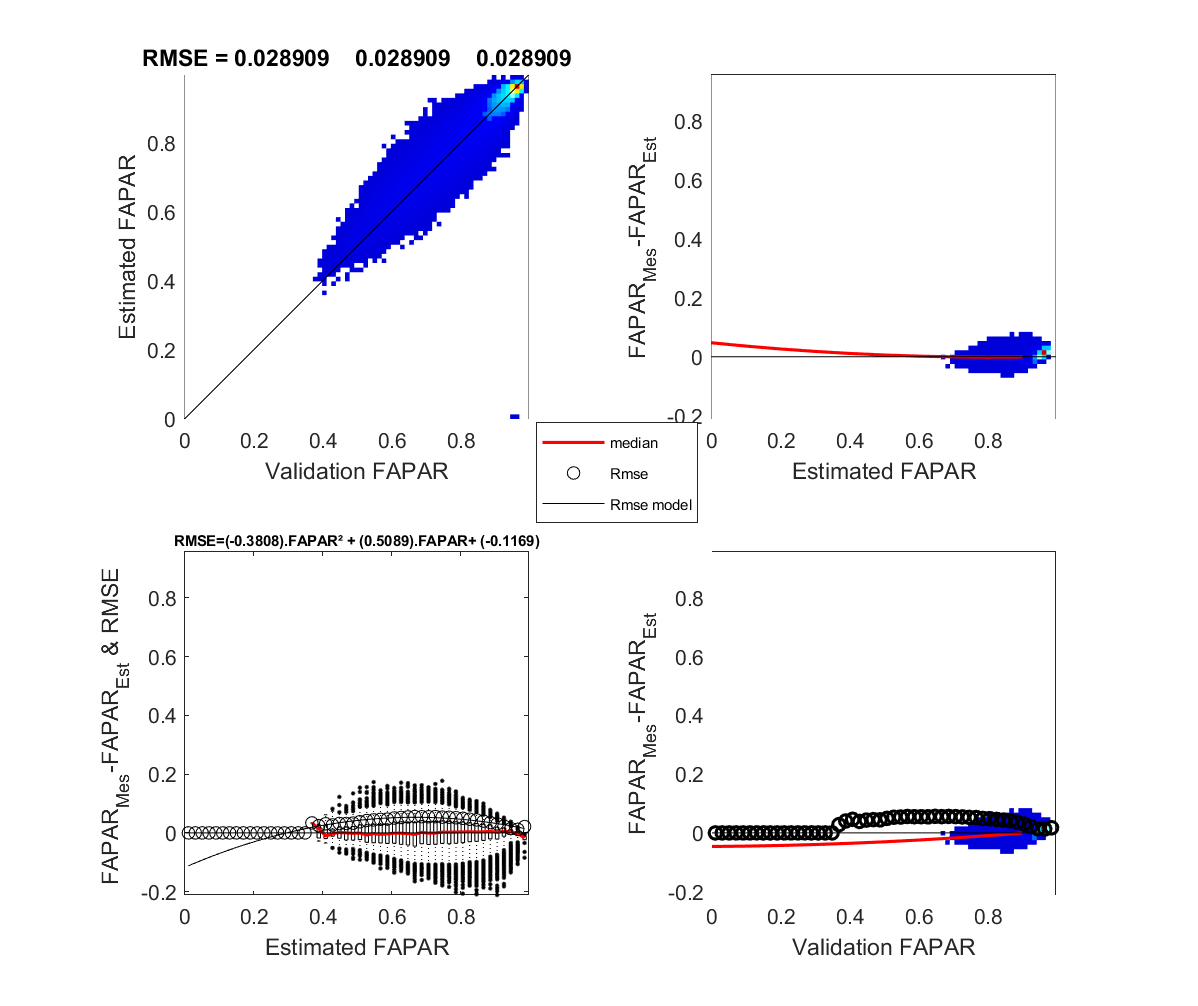
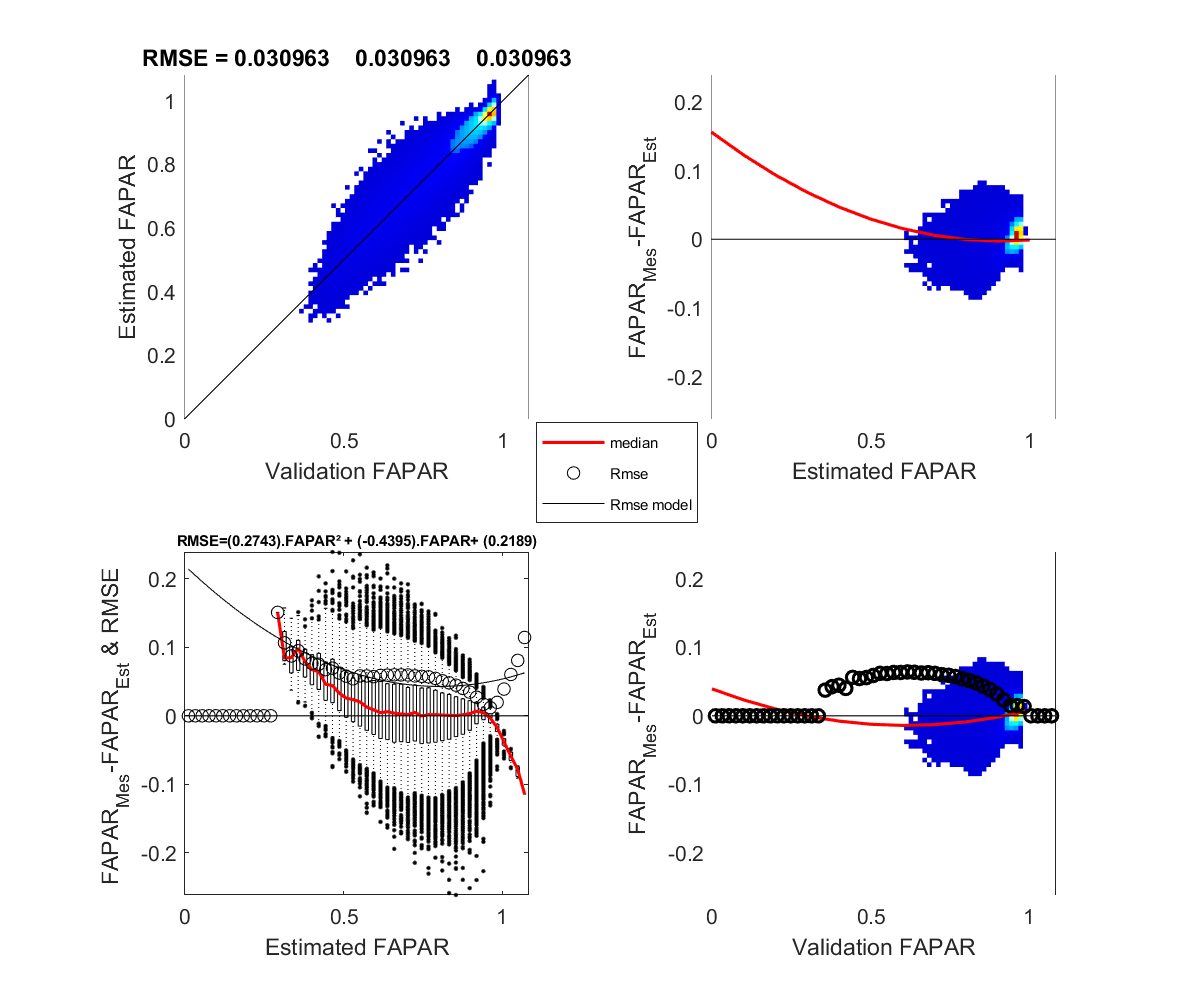
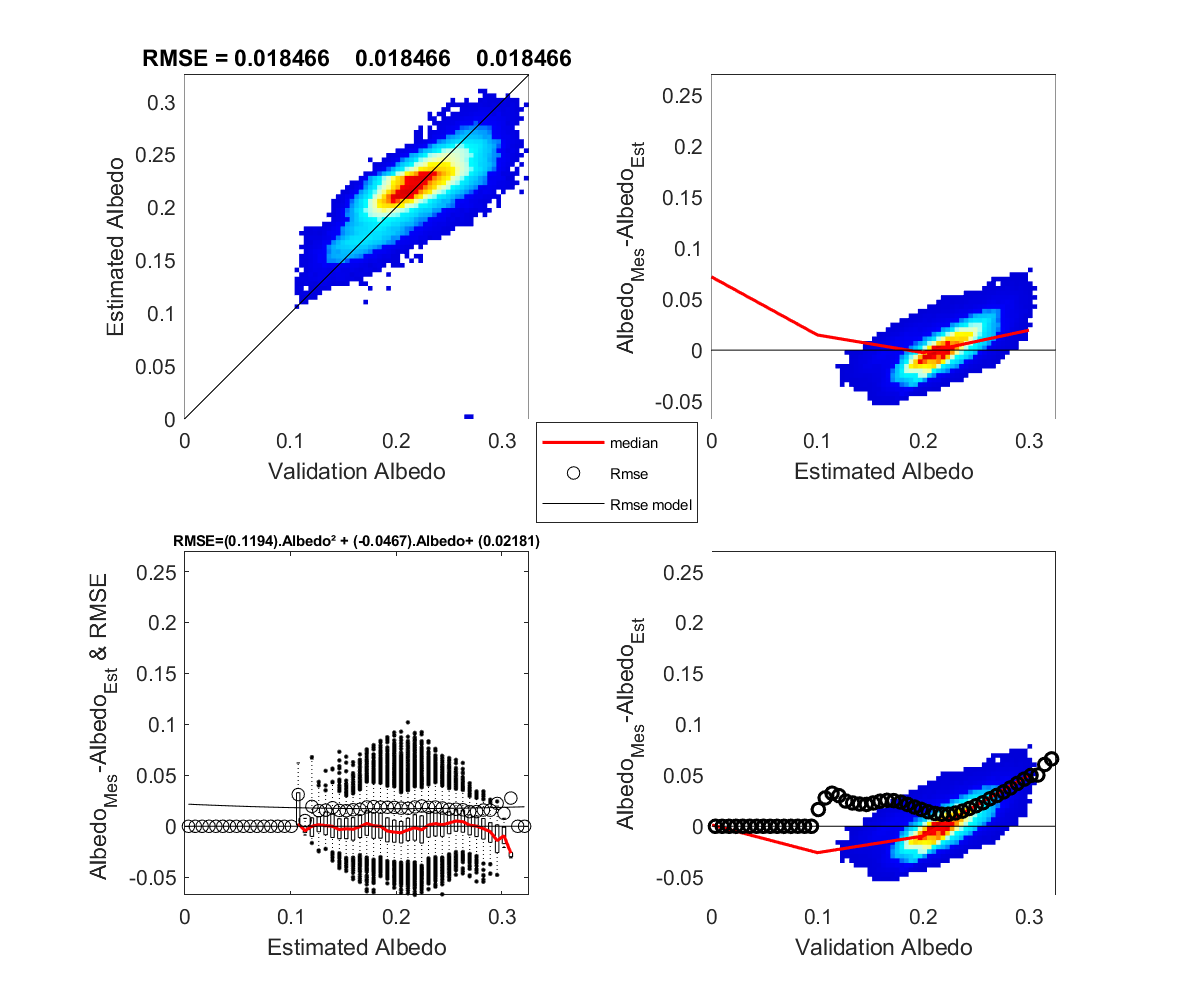
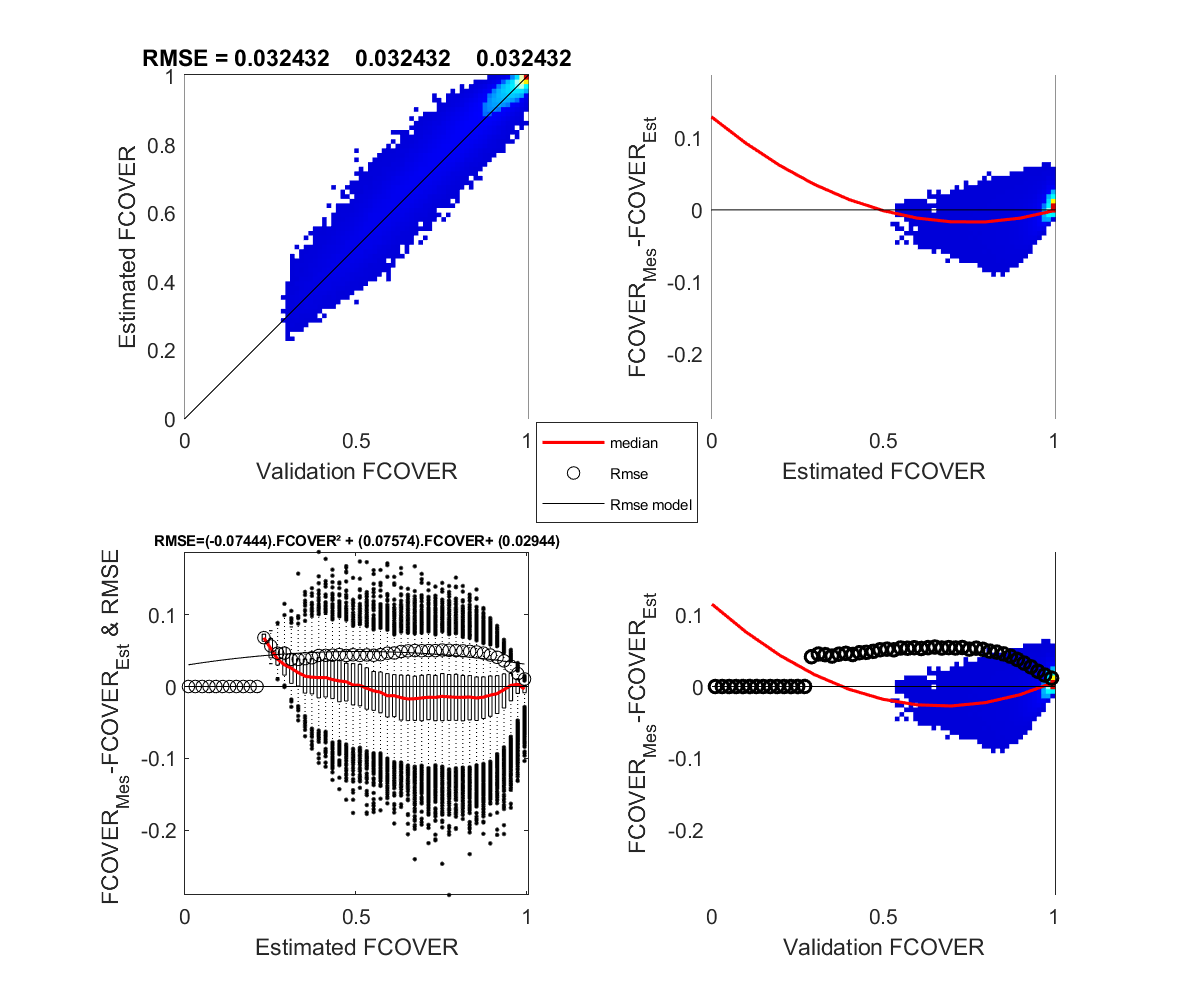
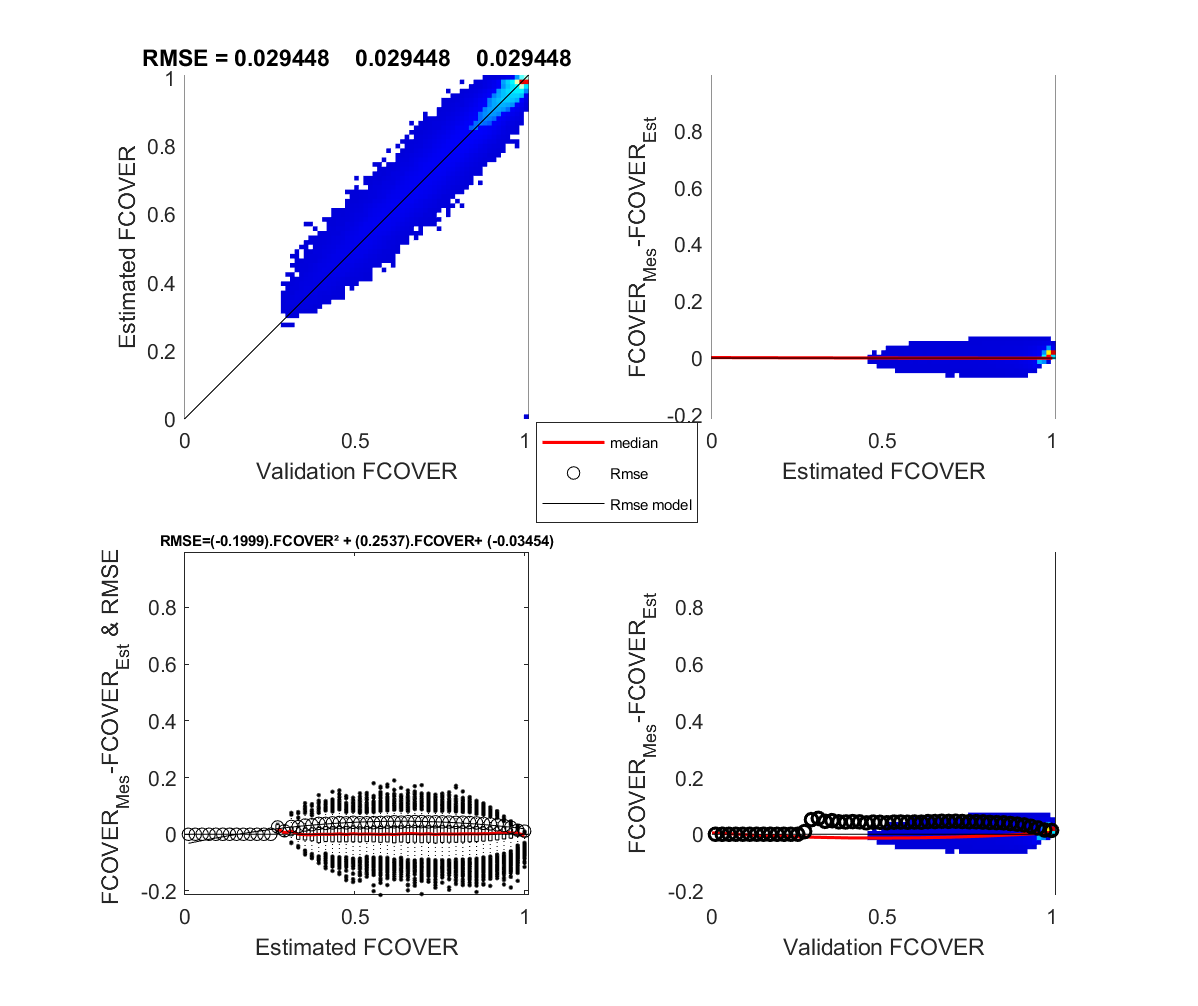
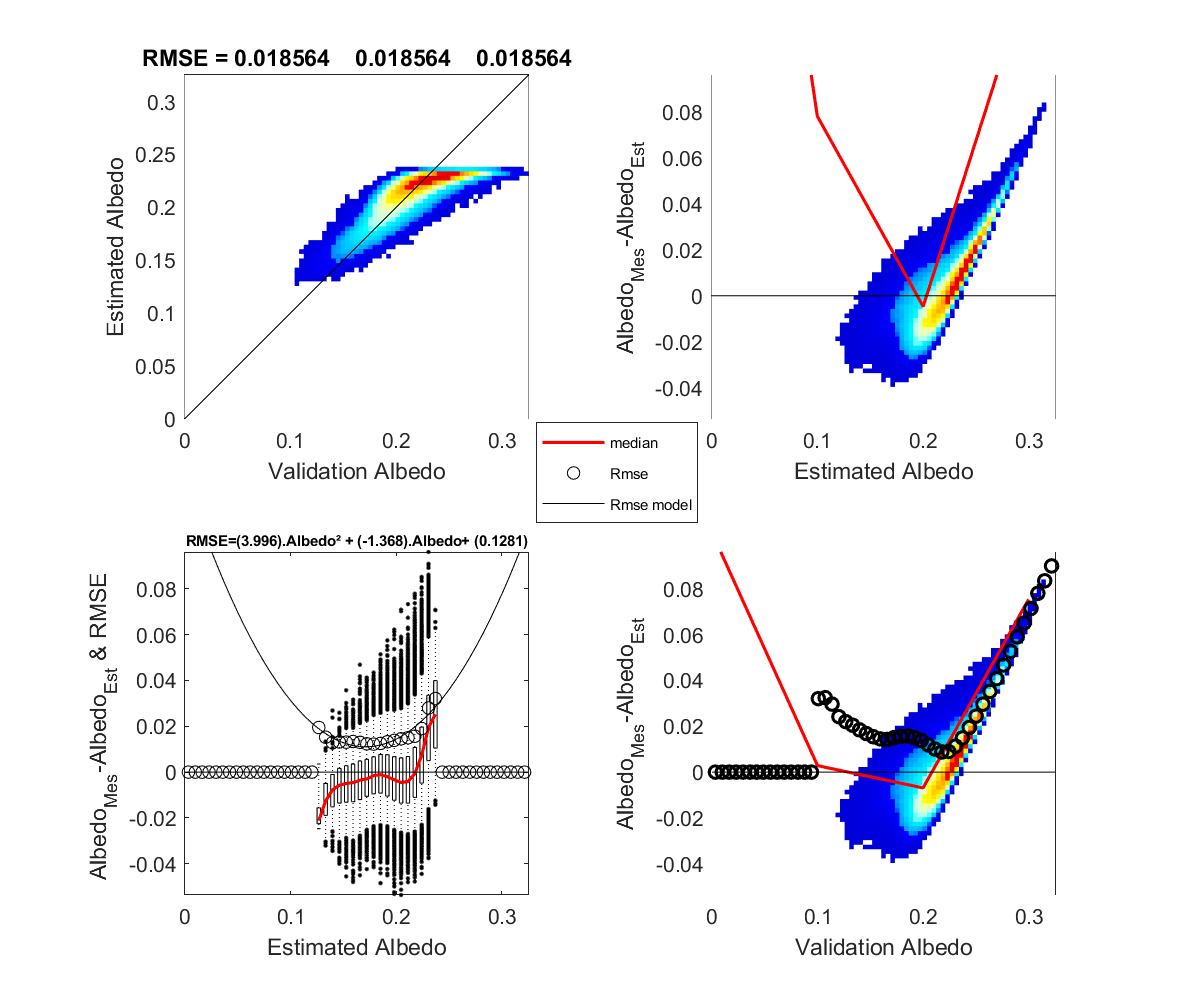


Figure 15. Taiga Cordillera Polar Grassland fCOVER validation for SL2P-D (left) AND SL2P (right).

Figure 16. Taiga Cordillera Polar Grassland albedo validation for SL2P-D (left) AND SL2P (right).

Figure 17. Taiga Cordillera Polar Grassland LAI validation for SL2P-D (left) AND SL2P (right).

Figure 18. Taiga Cordillera Polar Grassland fAPAR validation for SL2P-D (left) AND SL2P (right).

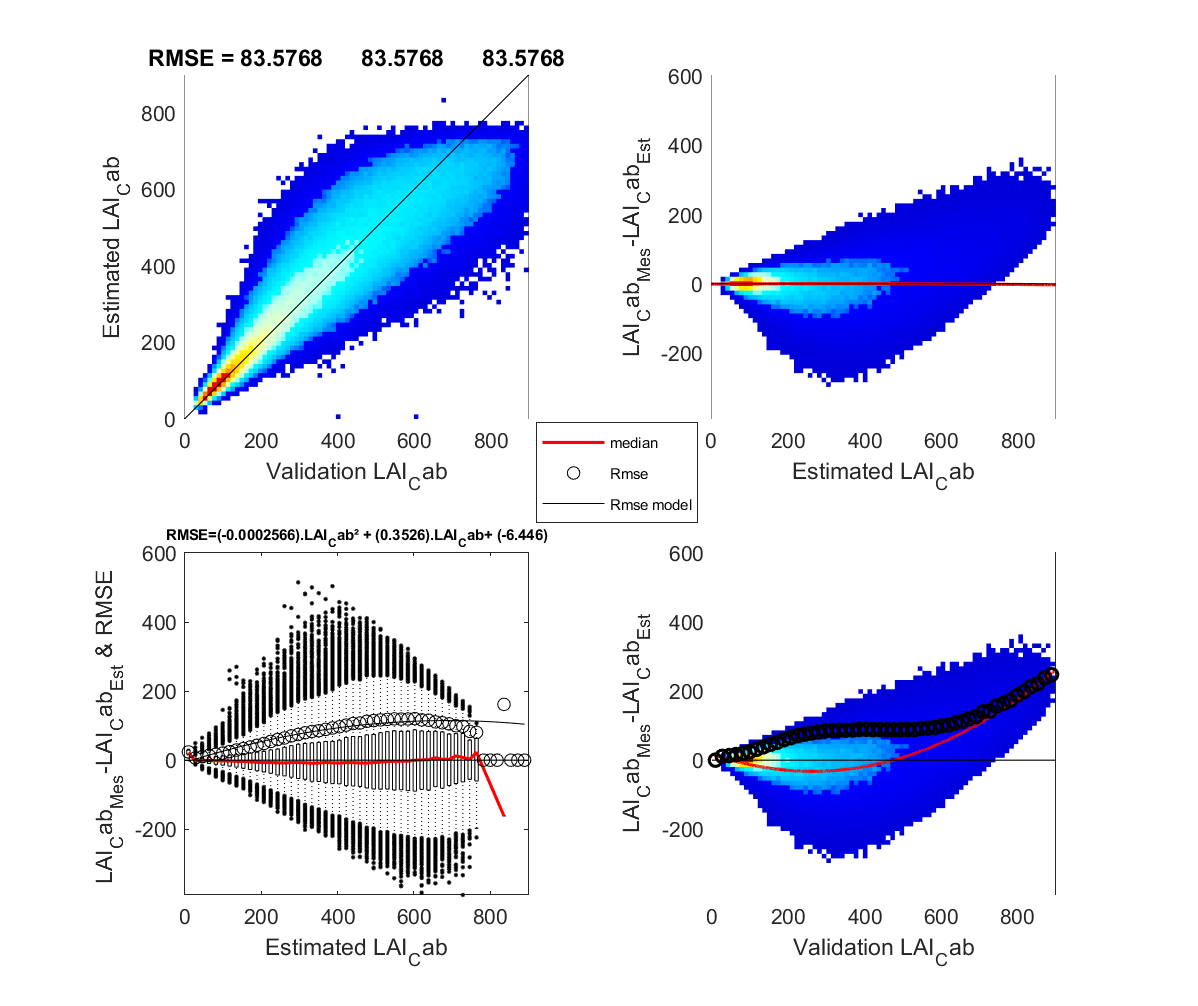
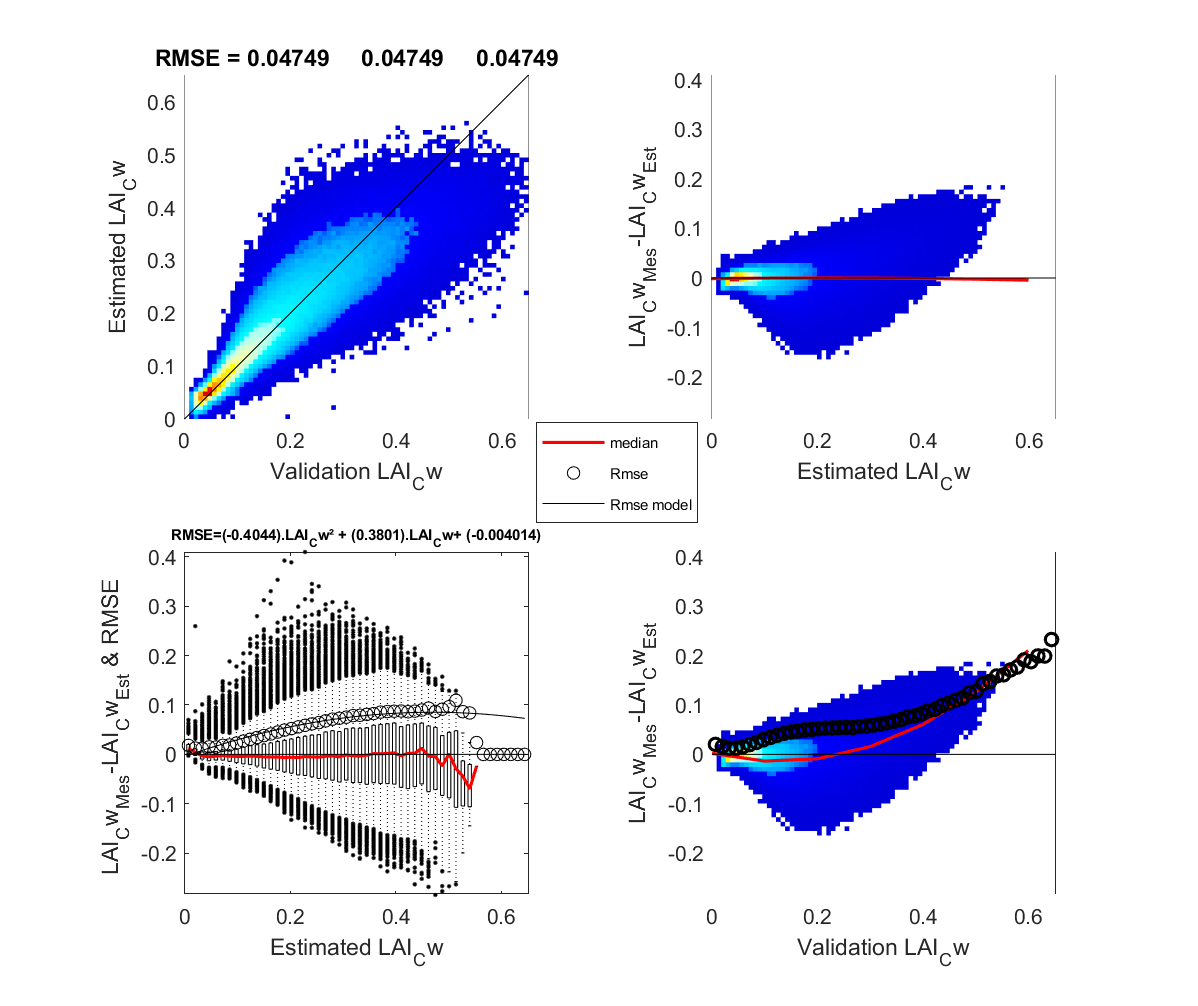
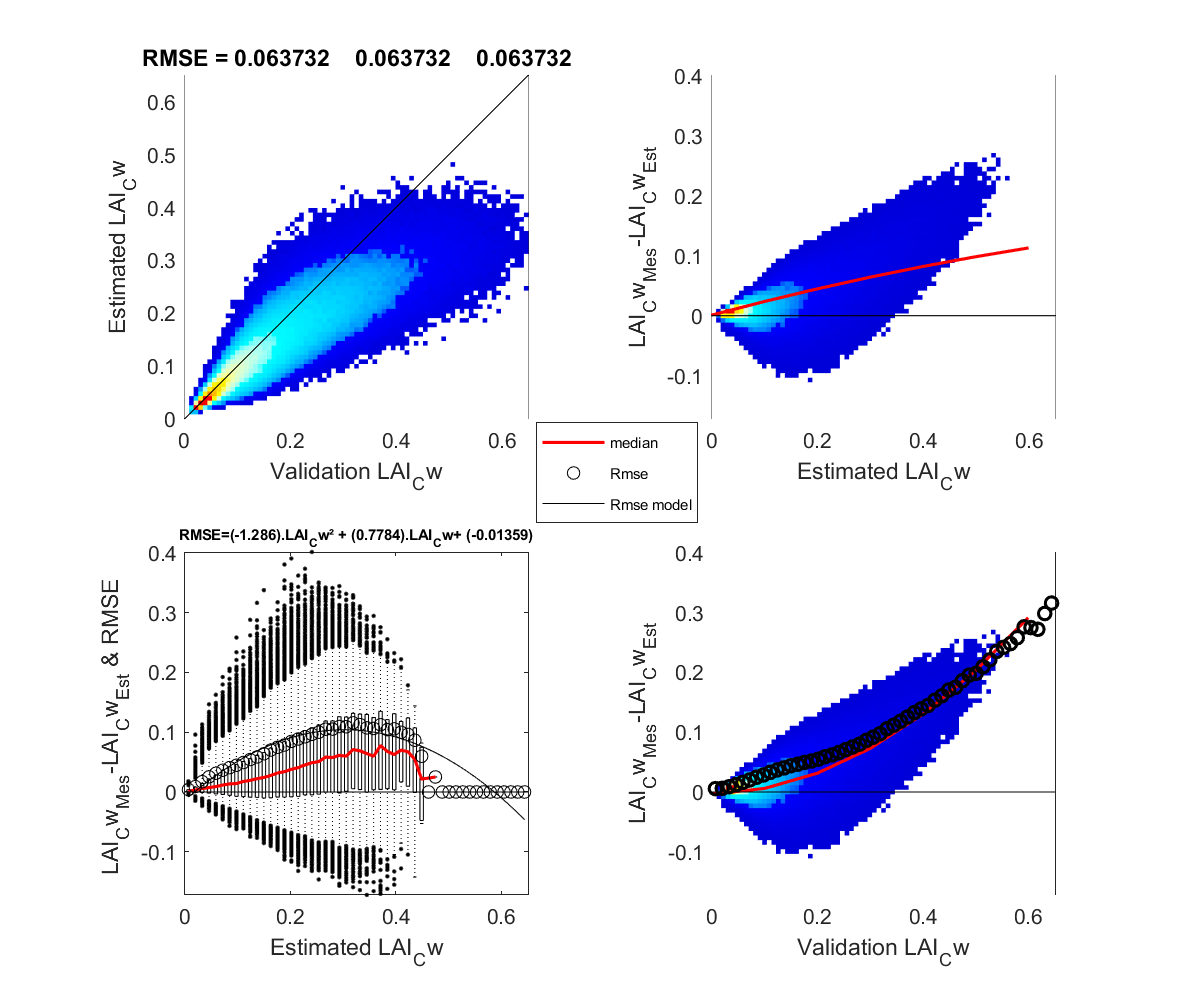
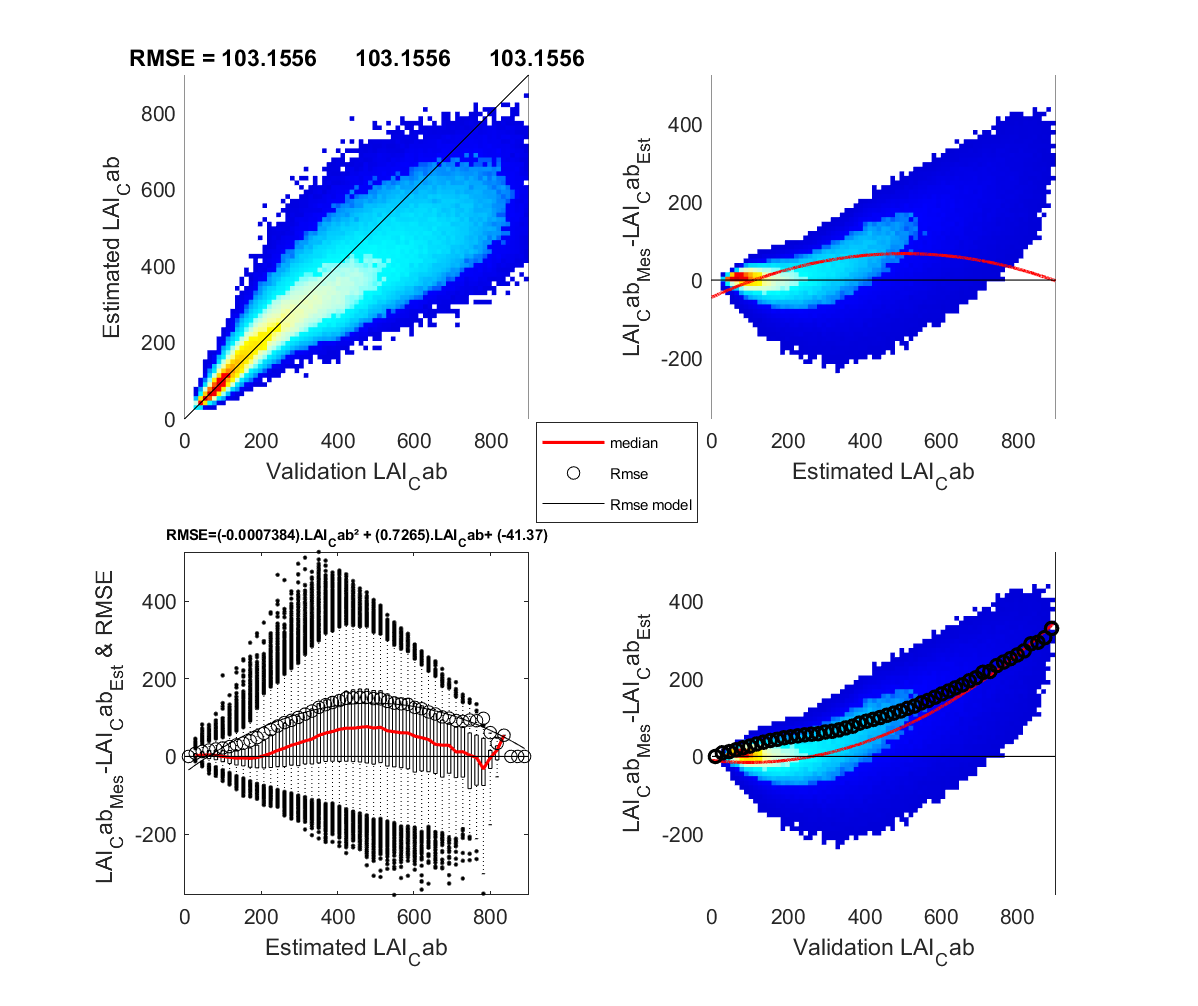


Figure 19. Taiga Cordillera Polar Grassland LAI.Cab validation for SL2P-D (left) AND SL2P (right).

Figure 20. Taiga Cordillera Polar Grassland LAI.Cw validation for SL2P-D (left) AND SL2P (right).

## Quality Indicators

Quality indicators are provided for each retrieval using a single byte to code 8 logical indictors (Table 20).

The terrain indicator identifies the designated terrain complexity. Complex terrain corresponds to approximately twice the expected noise of input measurements than simple terrain resulting in fewer out of range inputs but a larger retrieval uncertainty.

The LandCover indicator is set to Specific if a non-generic land cover is applied. By convention Class 1 is the generic vegetated land cover. Generic land cover will result in fewer out of range inputs and outputs but with a larger retrieval uncertainty if the correct land cover is known and specified.

The Input range indicator is set to Out of Range if the input measurements fall outside the convex hull of the noisy input training dataset.

The Output range indicator is set to Out of Range if the output estimates fall outside the convex hull of the output training dataset and the user specified range of LAI.

Table . Quality indictor coding.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Bit | Value=0 | Value=1 |
| Terrain | 1 | Simple Terrain | Complex Terrain |
| LandCover | 2 | Vegetated | Specific |
| Input range | 3 | In range | Out of range |
| Output range | 4 | In range | Out of range |
| Unused | 5 | Unused | Unused |
| Unused | 6 | Unused | Unused |
| Unused | 7 | Unused | Unused |
| Unused | 8 | Unused | Unused |

# Algorithm Implementation

To implement SL2P-D one requires MATLABR2019 or later with the Statistics and Machine Learning package and the Parallel Processing toolbox. One could disable parallel processing by searching and replacing all ‘parfor’ commands with ‘for’ commands.

SL2P-D executes for the first time for a provided parameter file by invoking at the Matlab command line:

status = SL2PD(‘parameter\_file\_name,1,1)

Successful execution will return a status of ‘1’ and produce a report directory, containing the calibration databases, as named in the parameter file. Validation is not essential.

SL2P-D can be executed multiple times to either calibrate additional regression algorithms or to apply a selected algorithm to a validation database. Both actions will result in additions to the Results.mat file for each class – either in terms of new networks or in terms of new validation results. If a new network is calibrated a new .xls worksheet will be added to the parameter file for the network details. If new validation is performed, will result in new visualizations for each class. One can make use of the .xls worksheet for a network to apply the regression algorithm to other datasets.

To apply SL2P-D in Google Earth Engine, the matlab function SL2P2GEE must be executed with an input database class directory specified. This function will produce one (using no partitioning) or more (using many partitions) CSV files corresponding to each network calibrated for the class. The user can then upload these CSV files as GEE assets. The LEAF Toolbox within GEE can then be configured to apply these networks based on a user specified index layer.

# Conclusions

The Simplified Level 2 Processor – Distributed for deriving vegetation biophysical variables from multispectral reflectance data is described. SL2P-D is generalization of SL2P to Allow for distributed regression algorithms as a function of a selected initial algorithm output (e.g. the ‘D’ variable). SL2P-D parameterizations for land cover found in Canada are provided. Validation of SL2P-D and SL2P using the same simulated databases indicates SL2P-D generally meets threshold requirements for bias and uncertainty while SL2P may not meet these requirements for certain variable and land cover conditions. SL2P-D and sample parameters files and output are available at <https://github.com/rfernand387/SL2PD/upload/master/Reports>. This resource can be used with Matlab Version 2019 and the Statistics and Machine Learning toolbox to apply SL2P to user databases formatted to meet the current simulation databases. Applications are under development to use SL2P-D regression algorithms within Python and Java (on Google Earth Engine). Further testing with the FLIGHT RTM (for forests) and using validation with in-situ measurements is underway.

# References

[Canada's National Forest Inventory](https://nfi.nfis.org/en/standardreports), 2013, Table 5.1 Area (1000 ha) of forest land by forest type, age class in Canada, 2006 revised baseline, Version 3, December 2013 (accessed April 7, 2017).

[Djamai and Fernandes, 2018](https://www.sciencedirect.com/science/article/pii/S0034425719301117#bbb0180), **Comparison of SNAP-derived sentinel-2A L2A product to ESA product over Europe,** Remote Sens., 10 (2018), p. 926.

Djamai, N.; Fernandes, R.; Weiss, M.; McNairn, H.; Goïta, K. Validation of the Sentinel Simplified Level 2 Product Prototype Processor (SL2P) for mapping cropland biophysical variables using Sentinel-2/MSI and Landsat-8/OLI data. Remote Sens. Environ. 2019, 225, 416–430.

[Fernandes et al., 2014](https://www.sciencedirect.com/science/article/pii/S0034425719301117#bbb0225), **Global Leaf Area Index Product Validation Good Practices,** Committee of Earth Observing Systems Working Group on Calibration and Validation (2014), p. 75

Fernandes, R.A. and Djamai. N., 2019. Well posed retrieval of canopy biophysical variables from multi-spectral imagery using the canopy directional scattering factor. In preparation.

Weiss and Baret, 2016, **S2ToolBox level 2 products. Version 1.1** , Available online [step.esa.int/docs/extra/ATBD\_S2ToolBox\_L2B\_V1.1.pdf](http://step.esa.int/docs/extra/ATBD_S2ToolBox_L2B_V1.1.pdf)

Widlowski, J-L., M. Taberner, B. Pinty, V. Bruniquel-Pinel, M. Disney, R. Fernandes, J.-P. Gastellu-Etchegorry, N. Gobron, A. Kuusk, T. Lavergne, S. Leblanc, P. Lewis, E. Martin, M. Mottus, P. J. R. North, W. Qin, M.Robustelli, N. Rochdi, R.Ruiloba, C.Soler, R.Thompson, W. Verhoef, M. M.Verstraete, and D. Xie (2007), 'The third RAdiation transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance modelling', Journal of Geophysical Research-D.

World Meteorological Organization, 2016, The Global Observing System for Climate: Implementation Needs, WMO Pub No. GCOS – 200.

# Appendix I – Physical variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fraction Canopy Cover | FCOVER | % | Fraction of horizontal surface area covered by vegetation | 0-1 |
| Leaf Area Index | LAI | DIM | Half the total live foliage area per horizontal ground area | 0-15 |
| Canopy Chlorophyll Content | CCC | g.m-2 | Mass of chlorophyll a+b per unit ground area | 0-100 |
| Canopy Water Content | CWC | g.m-2 | Mass of water per unit ground area | 0-100 |
| Albedo, black sky | A | 0-1 | Ratio of top of canopy upper hemispherical upwelling radiance to top of canopy incident direct irradiance | 0-1 |
| D | D | 0-1 | canopy scattering coefficient for a foliage single scattering albedo of one under direct irradiance | 0-1 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heterogenity | None | Nominal | Canopy heterogeneity | HomogenousHet: heterogenous |
| Senes. Frac | - | 0-1 | Fraction of matter that is senescent | 0-1 |
| Woody Frac | - | 0-1 | Fraction of matter that is woody | 0-1 |
| Age | - | years | Age since first growth | 0-100 |
| Leaf angle distribution | LAD | Discrete probability histogram | Proportion of leaves for angles: [xx),[xx),[xx),[xx),[xx),[xx] | 0-1; sum 1 |
| Leaf diameter | d | cm | Diameter of leaf | 0-100 |
| Leaf layers | N | Fractional count | Number of effective scattering layers in a leaf | 0-4 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chlorophyll a+b content | Cab | µg.cm-2 | Foliage chlorophyll a+b content | 0-100 |
| Water content | Cw | g.cm-2 | Foliage water content | 0-100 |
| Dry matter content | Cdm | g.cm-2 | Foliage dry matter content | 0-100 |
| Brown pigment content | Cbd | g.cm-2 | Foliage brown pigment content | 0-100 |
| Carotenid content | Ccar | µg.cm-2 | Foliage carotenid content | 0-100 |
| Xanthophyll content | Cx | µg.cm-2 | Foliagexanthophyll content | 0-100 |

# Appendix II SL2P-D Configurations for Canadian Ecozones

Table 18. Spatial partitions for calibration datasets. Regions correspond to Ecozones of Canada. Snow End (Onset) dates taken from earliest (latest) dates observed over CCRS, NOAA and MODIS snow cover poducts betweenb 2006 and 2010..

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Region | Area | LatMin | LatMax | Long Min | Long Max | Snow End | Snow Onset |
| 1000km2 | °N | °N | °E | °E | DOY | DOY |
| Global | 100000 | 41 | 83 | -141 | -47 | 1 | 365 |
| Canada | 10000 | 41 | 83 | -141 | -47 | 1 | 365 |
| Northern Arctic | 2886 | 58 | 83 | -128 | -62 | 122 | 294 |
| Southern Arctic | 1673 | 53 | 72 | -141 | -68 | 104 | 308 |
| Arctic Cordillera | 338 | 57 | 83 | -95 | -61 | 140 | 308 |
| Taiga Cordillera | 265 | 60 | 70 | -141 | -125 | 104 | 308 |
| Boreal Cordillera | 468 | 56 | 65 | -141 | -122 | 86 | 322 |
| Taiga Shield | 816 | 51 | 60 | -80 | -55 | 86 | 322 |
| Boreal Shield | 2072 | 44 | 60 | -112 | -47 | 86 | 350 |
| Taiga Plain | 654 | 57 | 69 | -137 | -113 | 86 | 322 |
| Hudson Plains | 447 | 50 | 59 | -96 | -75 | 86 | 336 |
| Mixedwood Plains | 169 | 41 | 48 | -85 | -69 | 1 | 365 |
| Boreal Plains | 738 | 49 | 62 | -123 | -96 | 32 | 336 |
| Atlantic Maritime | 281 | 43 | 50 | -73 | -59 | 1 | 365 |
| Prairies | 466 | 49 | 55 | -114 | -95 | 1 | 365 |
| Montane Cordillera | 448 | 49 | 58 | -131 | -112 | 1 | 365 |
| Pacific Maritime | 245 | 48 | 60 | -138 | -121 | 1 | 365 |

Day of Year, min and max correspond to Snow End and Snow Onset dates respectively

# Appendix III SL2P-D Canopy biophysical variable parameterizations

Parameterizations are provided for canopy biophysical variables representative of certain land cover classes found in Canada (Table 36). Age is parameterized for all classes but actually only used for treed classes when using 3D RT models based on distributions ( Table 34) fitted to data from the Candian Forest Service (Figure 3).

Table . Land cover legend conversions.

|  |  |  |  |
| --- | --- | --- | --- |
| Land Cover Class | Symbol | NRCan Definition | EU Land Service Definition |
| Vegetated | Veg | Vegetated area | Vegetated area |
| Evegreen Needleleaf forest | ENF | Temperate of subpolar needleaf forest | Evegreen Needleleaf forest |
| Deciduous Needleleaf forest | DNF | Temperate of subpolar needleaf forest | Deciduous Needleleaf forest |
| Evergreen Broadleaf Forest | EBF | Evergreen Broadleaf Forest | Evergreen Broadleaf Forest |
| Deciduous Broadleaf Closed Forest | DBF | Temperate of subpolar broadleaf deciduous forest | Deciduous Broadleaf Forest |
| Mixed forest | MF | Mixed Forest | Mixed forest  Unknown Forest |
| Shrubland | Sh | Temperate or sub-polar shrubland | Shrubland |
| Herbaceous wetland | Wet | Wetland | Herbaceous wetland |
| Moss&Lichen | Bryo | Subpolar or polar barren-licehn-moss | Moss&Lichen |
| Bare/sparse vegetation | LowVeg |  | Bare/sparse vegetation |
| Cropland | Crop | Cropland | Cropland |
| Polar shrubland | ShPolar | Polar or sub-polar shrubland-lichen-moss | Shrubland |
| Grassland or pasture | Gr | Temperate or subpolar grassland; | Herbaceous vegetation |
| Polar grassland | GrPolar | Polar or sub-polar grassland-lichen-moss | Bare/sparse vegetation |

Table . Age distributions for Canadian treed land cover classes based on data presented in Figure xx after Candian Forest Service, 2016.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Land Cover Class | Model | LB | Ub | P1 | P2 |
|  |  | y | y | y | y |
| Forest | Uniform | 10 | 150 | 90 | 40 |
| Evegreen Needleleaf forest | Truncated Gaussian | 10 | 150 | 80 | 30 |
| Deciduous Needleleaf forest | Truncated Gaussian | 10 | 150 | 62 | 30 |
| Evergreen Broadleaf Forest | Truncated Gaussian | 10 | 150 | 80 | 30 |
| Deciduous Broadleaf Forest | Truncated Gaussian | 10 | 150 | 62 | 30 |
| Mixed forest1 | Uniform | 10 | 110 | 10 | 110 |
| Shrubland | Uniform | 5 | 50 | 5 | 50 |

Figure . Stand age distributions across Canada and SL2P-D distribution for ENF (upper), DBF (middle) and mixed (lower) classes (from Canadian Forest Service, 2016)

Table 19. Canopy parameters for evengreen needleaf forest.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0.25 | 12 |  |  | 0.25 | 16 | 0.25 | 16 | Fern 2003 |
| ALA | ° | 20 | 70 | 60 | 20 | 20 | 70 | 20 | 70 | Zhu 2019 |
| LAD | Nom | S or U | S or U | S or U | S or U | S or U | S or U | S or U | S or U | Zhu 2019 |
| Crown  Cover | 0-1 | .2 | 1 |  |  | .2 | 1 | .2 | 1 | Latifovic2014 |
| Leaf Size | cm | 1 | 10 |  |  | 1 | 10 | 1 | 10 | Guess |
| HsD | DIM | 0.05 | 1 |  |  | 0.05 | 1 | 0.05 | 1 | Guess |
| N | DIM | 1.5 | 3.5 |  |  | 1.5 | 3.5 | 1.5 | 3.5 | ZT2004 |
| Cab | g.cm-2 | 10 | 70 |  |  | 10 | 70 | 10 | 70 | Croft2014  ZT2004 |
| Cdm | g.cm-2 | 0.005 | 0.01 |  |  | 0.005 | 0.01 | 0.005 | 0.01 | Croft2014  ZT2004 |
| Cw\_Rel | 0-1 | 0.5 | 0.65 |  |  | 0.5 | 0.65 | 0.5 | 0.65 | ZT2004 |
| Cbp | DIM | 0 | 0.6 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | Guess |
| Senes. Frac | 0-1 | 0 | 0.25 |  |  | 0 | 0.25 | 0 | 0.25 | Pine |
| Woody Frac | 0-1 | 0 | 0.5 | 0.1 | 0.2 | 0 | 0.5 | 0 | 0.5 | Guess |
| Age | cm | 5 | 30 | 10 | 10 | 5 | 5 | 5 | 30 |  |

Table 20. Canopy parameters for deciduous broadleaf forest.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0.25 | 15 |  |  | 0.25 | 16 | 0.25 | 16 | Fern2003 |
| ALA | ° | 20 | 60 | 30 | 20 | 20 | 60 | 20 | 60 | Zhu 2019  Falster 2003 |
| LAD | Nom | P or S or U | P or S or U | S or U | S or U | S or U | S or U | S or U | S or U | Zhu 2019 |
| Crown  Cover | 0-1 | .2 | 1 |  |  | .2 | 1 | .2 | 1 | Latifovic2014 |
| Leaf Size | cm | 1 | 10 |  |  | 1 | 10 | 1 | 10 | Falster 2003 |
| HsD | DIM | 0.1 | 1 |  |  | 0.1 | 1 | 0.1 | 1 |  |
| N | DIM | 1.1 | 2.3 |  |  | 1.1 | 2.3 | 1.1 | 2.3 | Lemaire2008 |
| Cab | g.cm-2 | 10 | 70 |  |  | 10 | 70 | 10 | 70 | Croft2014  ZT2004  LEmaire2008 |
| Cdm | g.cm-2 | 0.005 | 0.01 |  |  | 0.005 | 0.01 | 0.005 | 0.01 | Croft12014  ZT2004 |
| Cw\_Rel | 0-1 | 0.5 | 0.65 |  |  | 0.5 | 0.65 | 0.5 | 0.65 | ZT2004 |
| Cbp | DIM | 0 | 0.6 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | Guess |
| Senes. Frac | 0-1 | 0 | 0.25 |  |  | 0 | 0.25 | 0 | 0.25 | Pine |
| Woody Frac | 0-1 | 0 | 0.3 | 0.1 | 0.2 | 0 | 0.3 | 0 | 0.3 | ICP2016 |
| Age | cm | 5 | 30 | 10 | 10 | 5 | 5 | 5 | 30 |  |

Table 21. Canopy parameter ranges for shrubland

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 4 |  |  | 0 | 4 | 0 | 4 | Sonnentag  Maloley |
| ALA | ° | 30 | 80 | 60 | 4.1 | 30 | 80 | 30 | 80 | SL2P |
| LAD | Nom | U | U | U | U | U | U | U | U | SL2P |
| Crown  Cover | 0-1 | 0.1 | 0.8 | 0.45 | 0.058 | 0.1 | 0.8 | 0.1 | 0.8 | Tundra8 |
| Leaf Size | cm | 1 | 6 |  |  | 1 | 6 | 1 | 6 |  |
| HsD | DIM | 0.1 | 0.2 | 0.15 | 0.0083 | 0.1 | 0.2 | 0.1 | 0.1 |  |
| N | DIM | 1.1 | 2.3 |  |  | 1.1 | 2.3 | 1.1 | 2.3 | Lemaire2008 |
| Cab | Ug.cm-2 | 10 | 100 | 55 | 7.5 | 10 | 40 | 40 | 100 | Tundra7 |
| Cdm | g.cm-2 | 0.003 | 0.01 | 0.065 | 0.0058 | 0.003 | 0.01 | 0.003 | 0.01 | Tundra10  Tundra12 |
| Cw\_Rel | 0-1 | 0.58 | 0.88 | 0.73 | 0.025 | 0.58 | 0.88 | 0.58 | 0.88 | Tundra12 |
| Cbp | DIM | 0 | 0.1 | 0 | 0.05 | 0 | 0 | 0.1 | 0.1 | Guess |
| Senes. Frac | 0-1 | 0 | 0.1 |  |  | 0 | 0.1 | 0 | 0.1 | Guess |
| Woody Frac | 0-1 | 0 | 0.5 | 0.1 | 0.2 | 0 | 0.25 | 0 | 0.25 | ICP2016 |
| Age | cm | 5 | 30 | 10 | 10 | 5 | 5 | 5 | 30 |  |

Table 22. Canopy parameters for herbaceous wetland.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0.25 | 7.5 |  |  | 0.25 | 7.5 | 0.25 | 7.5 | Grass1 |
| ALA | ° | 50 | 80 |  |  | 50 | 80 | 50 | 80 | Grass2 |
| LAD | Nom | E | E | E | E | E | E | E | E | Grass2 |
| Crown Cover | 0-1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | Trundra1 |
| Leaf Size | cm | 1 | 1 |  |  | 1 | 1 | 1 | 1 | Grass5 |
| HsD | DIM | 0.05 | 0.1 |  |  | 0.1 | 0.2 | 0.1 | 0.1 | Grass2 |
| N | DIM | 1.5 | 1.9 |  |  | 1.5 | 1.9 | 1.5 | 1.9 | Grass2 |
| Cab | g.cm-2 | 15 | 55 |  |  | 15 | 55 | 15 | 55 | Grass1  Grass2 |
| Cdm | g.cm-2 | 0.005 | 0.01 |  |  | 0.005 | 0.01 | 0.005 | 0.01 | Grass2Grass3 |
| Cw\_Rel | 0-1 | 0.58 | 0.88 |  |  | 0.58 | 0.88 | 0.58 | 0.88 | Grass2  Grass3 |
| Cbp | DIM | 0 | 0.6 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | Grass4 |
| Senes. Frac | 0-1 | 0 | 0.65 | 0.1 | 0.1 | 0 | 0.25 | 0 | 0.65 | Grass4 |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Guess |
| Age | cm | 5 | 30 | 10 | 10 | 5 | 5 | 5 | 30 |  |

Table 23. Canopy parameters for sphagnum and feathermoss.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 3.5 |  |  | 0 | 3.5 | 0 | 3.5 | Moss1  Moss4 |
| ALA | ° | 57 | 57 |  |  | 57 | 57 | 57 | 57 | Moss1 |
| LAD | Nom | U | U | U | U | U | U | U | U | Moss1 |
| Crown Cover | 0-1 | 0.5 | 1 | 1 | 0 | 0.5 | 1 | 1 | 1 | Moss1 |
| Leaf Size | cm | 0.5 | 0.5 |  |  | 0.5 | 0.5 | 0.5 | 0.5 | Moss1 |
| HsD | DIM | 0.01 | 0.01 |  |  | 0.01 | 0.01 | 0.01 | 0.01 | Guess |
| N | DIM | 1 | 1 |  |  | 1 | 1 | 1 | 1 | Guess |
| Cab | g.cm-2 | 10 | 30 |  |  | 10 | 30 | 10 | 30 | Moss4 |
| Cdm | g.cm-2 | 0.005 | 0.02 |  |  | 0.005 | 0.02 | 0.005 | 0.02 | Moss1  Moss2  Moss4 |
| Cw\_Rel | 0-1 | 0.75 | 1 |  |  | 0.75 | 1 | 0.75 | 1 | Moss2 |
| Cbp | DIM | 0 | 0.1 |  |  | 0 | 0.1 | 0 | 0.1 | Guess |
| Senes. Frac | 0-1 | 0 | 0.5 | 0.1 | 0.1 | 0 | 0.5 | 0 | 0.5 | Moss1 |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Moss1 |
| Age | years | 100 | 100 |  |  | 5 | 5 | 20 | 100 |  |

Table 24. Canopy parameters for lichen and feathermoss.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 2 |  |  | 0 | 2 | 0 | 2 | Moss1  Moss4 |
| ALA | ° | 57 | 57 |  |  | 57 | 57 | 57 | 57 | Moss1 |
| LAD | Nom | U | U | U | U | U | U | U | U | Moss1 |
| Crown Cover | 0-1 | 0.5 | 1 | 1 | 0 | 0.5 | 1 | 1 | 1 | Moss1 |
| Leaf Size | cm | 0.5 | 0.5 |  |  | 0.5 | 0.5 | 0.5 | 0.5 | Moss1 |
| HsD | DIM | 0.01 | 0.01 |  |  | 0.01 | 0.01 | 0.01 | 0.01 | Guess |
| N | DIM | 1 | 1 |  |  | 1 | 1 | 1 | 1 | Guess |
| Cab | g.cm-2 | 2 | 20 |  |  | 2 | 20 | 2 | 20 | Moss6 |
| Cdm | g.cm-2 | 0.005 | 0.02 |  |  | 0.005 | 0.02 | 0.005 | 0.02 | Moss1  Moss2  Moss4 |
| Cw\_Rel | 0-1 | 0.2 | 0.8 |  |  | 0.2 | 0.8 | 0.2 | 0.8 | Moss2 |
| Cbp | DIM | 0 | 0.1 |  |  | 0 | 0.1 | 0 | 0.1 | Guess |
| Senes. Frac | 0-1 | 0 | 0.5 | 0.1 | 0.1 | 0 | 0.5 | 0 | 0.5 | Moss1 |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Moss1 |
| Age | years | 100 | 100 |  |  | 5 | 5 | 20 | 100 |  |

Table 25. Canopy parameters for cropland – closed, no senescent material.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 1 | 8 |  |  | 1 | 8 | 1 | 8 | Crop1  Fern2003 |
| ALA | ° | 30 | 70 | 30 | 20 | 30 | 70 | 30 | 70 | Crop1 |
| LAD | Nom | U,P,S | U,P,S | U,P,S | U,P,S | U,P,S | U,P,S | U,P,S | U,P,S | SL2P |
| Crown Cover | 0-1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | Crop1 |
| Leaf Size | cm | 1 | 10 |  |  | 1 | 10 | 1 | 10 | Crop1 |
| HsD | DIM | 0.1 | 0.5 |  |  | 0.1 | 0.5 | 0.1 | 0.5 | Crop1 |
| N | DIM | 1.5 | 2.2 |  |  | .5 | 2.2 | .5 | 2.2 | SL2P |
| Cab | g.cm-2 | 20 | 90 |  |  | 20 | 90 | 20 | 90 | SL2P |
| Cdm | g.cm-2 | 0.003 | 0.011 |  |  | 0.003 | 0.011 | 0.003 | 0.011 | SL2P |
| Cw\_Rel | 0-1 | 0.6 | 0.85 |  |  | 0.6 | 0.85 | 0.6 | 0.85 | SL2P |
| Cbp | DIM | 0 | 2 | 0 | 0.3 | 0 | 2 | 0 | 2 | SL2P |
| Senes. Frac | 0-1 | 0 | 0.65 | 0.1 | 0.1 | 0 | 0.25 | 0 | 0.65 | - |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - |
| Age | years | 100 | 100 |  |  | 5 | 5 | 20 | 100 |  |

Table 26. Canopy parameters for cropland – sparse, no senescent material. Same as for closed cropland but reduced max for most parameters and decreased min Crown Cover.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 2 |  |  | 0.25 | 7.5 | 0.25 | 7.5 | Crop1 |
| ALA | ° | 20 | 60 |  |  | 20 | 60 | 20 | 60 | Crop1 |
| LAD | Nom | U, P, or S | U, P, or S |  |  | U, P, or S | U, P, or S | U, P, or S | U, P, or S | SL2P |
| Crown Cover | 0-1 | 0.25 | 0.75 | 1 | 0 | 0.25 | 0.75 | 0.25 | 0.75 | Crop1 |
| Leaf Size | cm | 0.5 | 5 |  |  | 0.5 | 5 | 0.5 | 5 | Crop1 |
| HsD | DIM | 0.1 | 0.5 |  |  | 0.1 | 0.5 | 0.1 | 0.5 | Crop1 |
| N | DIM | 1.2 | 2.2 |  |  | 1.2 | 2.2 | 1.2 | 2.2 | SL2P |
| Cab | g.cm-2 | 20 | 90 |  |  | 20 | 90 | 20 | 90 | SL2P |
| Cdm | g.cm-2 | 0.003 | 0.011 |  |  | 0.003 | 0.011 | 0.003 | 0.011 | SL2P |
| Cw\_Rel | 0-1 | 0.58 | 0.88 |  |  | 0.58 | 0.88 | 0.58 | 0.88 | SL2P |
| Cbp | DIM | 0 | 1 | 0 | 0.3 | 0 | 21 | 0 | 1 | SL2P |
| Senes. Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | - |
| Age | years | 100 | 100 |  |  | 5 | 5 | 20 | 100 |  |

Table 27. Canopy parameter ranges for Polar shrubland

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 3.5 |  |  | 0 | 3.5 | 0 | 3.5 | Tundra9  Chen2009 |
| ALA | ° | 30 | 80 | 60 | 4.1 | 30 | 80 | 30 | 80 | SL2P |
| LAD | Nom | U | U | U | U | U | U | U | U | SL2P |
| Crown Cover | 0-1 | 0.1 | 0.8 | 0.45 | 0.058 | 0.1 | 0.8 | 0.1 | 0.8 | Tundra8 |
| Leaf Size | cm | 1 | 3 | 4.5 | 0.25 | 1 | 3 | 1 | 3 | Wiki; Tundra10 |
| HsD | DIM | 0.1 | 0.2 | 0.15 | 0.0083 | 0.1 | 0.2 | 0.1 | 0.1 |  |
| N | DIM | 1.1 | 2.3 |  |  | 1.1 | 2.3 | 1.1 | 2.3 | Lemaire2008 |
| Cab | g.cm-2 | 10 | 100 | 55 | 7.5 | 10 | 40 | 40 | 100 | Tundra7 |
| Cdm | g.cm-2 | 0.003 | 0.01 | 0.065 | 0.0058 | 0.003 | 0.01 | 0.003 | 0.01 | Tundra10  Tundra12 |
| Cw\_Rel | 0-1 | 0.58 | 0.88 | 0.73 | 0.025 | 0.58 | 0.88 | 0.58 | 0.88 | Tundra12 |
| Cbp | DIM | 0 | 0.1 | 0 | 0.05 | 0 | 0 | 0.1 | 0.1 | Guess |
| Senes. Frac | 0-1 | 0 | 0.1 |  |  | 0 | 0.1 | 0 | 0.1 | Guess |
| Woody Frac | 0-1 | 0 | 0.1 |  |  | 0 | 0.1 | 0 | 0.1 | Guess |
| Age | years | 5 | 100 |  |  | 5 | 5 | 20 | 100 |  |

Table 28. Canopy parameters for grassland or pasture.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 7.5 |  |  | 0 | 7.5 | 0 | 7.5 | Grass1 |
| ALA | ° | 50 | 80 |  |  | 50 | 80 | 50 | 80 | Grass2 |
| LAD | Nom | E | E | E | E | E | E | E | E | Grass2 |
| Crown Cover | 0-1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | Trundra1 |
| Leaf Size | cm | 0.4 | 1.2 |  |  | 0.5 | 1 | 0.5 | 1 | Grass5 |
| HsD | DIM | 0.05 | 0.1 |  |  | 0.1 | 0.2 | 0.1 | 0.1 | Grass2 |
| N | DIM | 1.5 | 1.9 |  |  | 1.5 | 1.9 | 1.5 | 1.9 | Grass2 |
| Cab | g.cm-2 | 15 | 55 |  |  | 15 | 55 | 15 | 55 | Grass1  Grass2 |
| Cdm | g.cm-2 | 0.005 | 0.01 |  |  | 0.005 | 0.01 | 0.005 | 0.01 | Grass2  Grass3 |
| Cw\_Rel | 0-1 | 0.58 | 0.88 |  |  | 0.58 | 0.88 | 0.58 | 0.88 | Grass2  Grass3 |
| Cbp | DIM | 0 | 0.6 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | Grass4 |
| Senes. Frac | 0-1 | 0 | 0.65 | 0.1 | 0.1 | 0 | 0.25 | 0 | 0.65 | Grass4 |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Guess |
| Age | years | 100 | 100 |  |  | 100 | 100 | 100 | 100 |  |

Table 29. Canopy parameter ranges for Polar grassland.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Units | Min | Max | Mode | Std | Var  Min | Var  Max | Var  Min | Var  Max | Reference |
| LAI | DIM | 0 | 3 | 1.5 | 0.25 | 0 | 3 | 0 | 3 | Trundra1  Tundra2 |
| LAI | ° | 50 | 80 | 65 | 2.5 | 50 | 80 | 50 | 80 | Tundra2 |
| ALA | Nom | E | E | E | E | E | E | E | E | Tundra2 |
| LAD | 0-1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | Trundra1 |
| Crown Cover | cm | 0.5 | 1 | 0.75 | 0.041 | 0.5 | 1 | 0.5 | 1 | Tundra2 |
| Leaf Size | DIM | 0.1 | 0.2 | 0.15 | 0.0083 | 0.1 | 0.2 | 0.1 | 0.1 | Wiki |
| HsD | DIM | 1.4 | 2.4 | 1.9 | 0.0833 | 1.4 | 2.4 | 1.4 | 2.4 | Tundra3  Thickness1 |
| N | g.cm-2 | 17 | 73 | 45 | 4.67 | 17 | 73 | 17 | 73 | Tundra3 |
| Cab | g.cm-2 | 0.003 | 0.01 | 0.065 | 0.0058 | 0.003 | 0.1 | 0.003 | 0.1 | Tundra3 |
| Cdm | 0-1 | 0.58 | 0.88 | 0.73 | 0.025 | 0.58 | 0.88 | 0.58 | 0.88 | Tundra3 |
| Cw\_Rel | DIM | 0 | 0.6 | 0 | 0.3 | 0 | 0 | 0 | 0.6 | Tundra2 |
| Cbp | 0-1 | 0 | 0.5 | 0.1 | 0.1 | 0 | 0 | 0 | 0.5 | Tundra2 |
| Senes. Frac | 0-1 | 0 | 0.65 | 0.1 | 0.1 | 0 | 0.25 | 0 | 0.65 | Guess |
| Woody Frac | 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |
| Age | years | 100 | 100 |  |  | 100 | 100 | 100 | 100 |  |

# Appendix IV – FLIGHT Variables suitable for Canada

The FLIGHT radiative transfer model requires additional canopy (Table 30). These quantities are tied to the current canopy parameters to minimize the dimensionality of the canopy parameter distribution using equations and pdf’s whose parameters are specified in a class specific worksheet for FLIGHT simulations (e.g. Table 31 shown in transposed form for formatting). A realization of canopy paramaters for a single species is produced as follows based on a given sample of parameters listed in Table xx.

The probability of the given age is determined using a truncated pdf. This is later used to select the species ultimately used for this sample. The site index, leaf size, fraction of senescent, fraction of bark and shoot clumping are determined using a truncated pdf specified using databases developed for allometric models or scientific studies (Table 32) . Height is determined as a function of age and site index based on the specified Age H fn and dbh is determined as a function of height based on the specified H DBH fn. Crown radius and crown semi height are determined as a function of the height. The mean height to crown base is determined as H – 2\*crown semi height. The minimum and maximum height to crown base are determined by adding a random perturbation to the mean height to the crown base. The random perturbation corresponds to a normally distributed sample with zero mean and standard deviation corresponding to the Euclidean sum of the root mean square error of the allometric model for H and the functional model for crown semi height; restricted to a minimum height of 0.1m. The validation of each realization is tested by verifying that the within crown LAI density (LAI/crown cover) falls within the specified range. Invalid species are discarded for this sample.

The condition probability of each species given age is determined by dividing the conditional probability of age given species by the sum across all species. One species is selected from the valid realizations based on the conditional probability of seeing that species for the given age.

Table . FLIGHT Canopy inputs

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | FLIGHT Symbol | Units | Description |
| Leaf Area Index | **TOTAL\_LAI** | m2 /m2 | Half all sided foliage area per unit horizontal ground area |
| Fraction green | **FRAC\_GREEN** | 0-1 | Fraction green leaves in foliage by area |
| Fraction senescent | **FRAC\_SEN** | 0-1 | Fraction of senescent/shoot material in foliage |
| Fraction bark | **FRAC\_BARK** | 0-1 | Fraction of bark in foliage |
| Leaf angle distribution | **LAD** | 0-1 | normal to leaves and vertical as fraction lying within 10 degree bins |
| Soil roughness | **SOILROUGH** | 0-1 | Lambertian soil 0, rough (mean slope 60deg) given by 1 |
| Leaf size | **LF\_SIZE** | m | radius |
| Fraction crown cover | **FRAC\_COV** | 0-1 | vertical projection of opaque crowns |
| Crown shape | **CROWN\_SHAPE** | ‘e’ or ‘c’ | 'e' for ellipsoid, 'c' for cones |
| Crown radius | **Exy** | m | Crown radius |
| Crown semi-height | **Ez** | m | Crown centre to top distance |
| Minimum crown base height | **MIN\_HT** | m | min height to crown base |
| Maxium crown base height | **MAX\_HT** | m | max height to crown base |
| Diameter at breast heigh | **dbh** | m | Diameter at breast height |
| Shoot clumping | **Gamma** | dim | Shoot clumping index |

Table . FLIGHT sampling parameters for a typical forest class (evergeen needleaf forests in Canada).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Parameter | Species | | | | | | |
| Species | Black spruce | Jack pine | Red Pine | White pine | White spruce | Larch | Balsam fir |
| Ecosystem | Ontario | Ontario | Ontario | Ontario | Western Canada  Ontario | Western Canada | Western Canada |
| Age Lb | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Age Ub | 100 | 100 | 100 | 200 | 130 | 100 | 160 |
| Age P1 | 10 | 10 | 10 | 5 | 10 | 5 | 10 |
| Age P2 | 100 | 100 | 100 | 200 | 130 | 100 | 160 |
| Age Distn | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform |
| S Lb | 3 | 3 | 9 | 15 | 3 | 10 | 11 |
| S UB | 24 | 24 | 27 | 30 | 24 | 30 | 34 |
| SI P1 | 3 | 3 | 3 | 15 | 3 | 10 | 11 |
| S P2 | 24 | 24 | 24 | 30 | 24 | 30 | 34 |
| S Distn | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform |
| rl LB | 0.03 | 0.02 | 0.0508 | 0.05 | 0.021213 | 0.02 | 0.00995 |
| rl UB | 0.08 | 0.04 | 0.0762 | 0.1 | 0.029665 | 0.05 | 0.020516 |
| rl P1 | 0.03 | 0.02 | 0.0508 | 0.05 | 0.021213 | 0.02 | 0.00995 |
| rl P2 | 0.08 | 0.04 | 0.0762 | 0.1 | 0.029665 | 0.05 | 0.020516 |
| rl Distn | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform |
| g LB | 1.4 | 1.4 | 1.4 | 1.4 | 1.3 | 1.3 | 1.4 |
| g UB | 2 | 2 | 2 | 2 | 1.5 | 1.5 | 2 |
| g P1 | 1.61 | 1.61 | 1.61 | 1.61 | 1.3 | 1.3 | 1.61 |
| g P2 | 0.19 | 0.19 | 0.19 | 0.19 | 1.5 | 1.5 | 0.19 |
| g Distn | Normal | Normal | Normal | Normal | Normal | Normal | Normal |
| fbark LB | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 |
| fbark UB | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| fbark P1 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 |
| fbark P2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| fbark Distn | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform |
| fsen LB | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| fsen UB | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| fsen P1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| fsen P2 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| fsen Distn | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform | Uniform |
| H lb | 1.42 | 2.06 | 1.97 | 1.49 | 2.5 | 1.52 | 1.52 |
| H ub | 26.8 | 28.02 | 40.68 | 38.87 | 30.78 | 26.88 | 26.88 |
| a0 | 5.9962 | 1.8458 | 1.5785 | 5.851 | 1.6086 | 1.2906 | 1.207 |
| a1 | 0.5104 | 0.8797 | 0.9729 | 0.6273 | 0.9851 | 1.0096 | 1.0383 |
| a2 | -0.0223 | -0.0381 | -0.0282 | -0.0245 | -0.0369 | -0.0401 | -0.055 |
| a3 | 50 | 50 | 50 | 50 | 50 | 2.0034 | 50 |
| a4 | 0 | 0 | 0 | 0 | 0 | 0.0182 | 0 |
| sigma H | 0.42 | 0.29 | 0.85 | 1.23 | 2 | 0.49 | 0.5 |
| Age H fn | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| DBH lb | 0.7 | 1.4 | 2.5 | 2.5 | 2.5 | 2.4 | 2.4 |
| DBH UB | 36.5 | 44.8 | 61.2 | 90.2 | 51.9 | 42.7 | 42.7 |
| b0 | 22.2124 | 22.943 | 31.5167 | 30.9183 | 31.6215 | 27.0418 | 20.1136 |
| b1 | 0.0729 | 0.0967 | 0.0487 | 0.0214 | 0.0367 | 0.067 | 0.0791 |
| b2 | 1.4633 | 1.9923 | 1.5772 | 1.1307 | 1.1899 | 1.305 | 1.5449 |
| H DBH fn | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| rxy lb | 0.4 | 1 | 1 | 0.3 | 0.3 | 0.15 | 0.15 |
| rxy ub | 2.6 | 10 | 10 | 7.4 | 4.65 | 6.25 | 6.25 |
| rxy c0 | 0.40125 | 1.1817 | 1.1817 | 1.1817 | 1.125 | 1.0275 | 1.1817 |
| rxy c1 | 0.742 | 0.0265 | 0.0265 | 0.0265 | 0.496 | 0.572 | 0.0265 |
| rxy model | 2 | 1 | 1 | 1 | 2 | 2 | 1 |
| rz d0 | 0.63 | 0.4 | 0.4 | 0.54 | 0.67 | 0.57 | 0.57 |
| rz d1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| LAI Min | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| LAI Max | 10 | 10 | 15 | 10 | 10 | 10 | 10 |

Table . Functional relationships used to determine FLIGHT canopy parameters.

|  |  |
| --- | --- |
| H DBH fn | Reference |
| 1 | http://flash.lakeheadu.ca/~chpeng/OFRI155.pdf |
| 2 | <https://ir.library.oregonstate.edu/downloads/3n2043903> |
| Age H fn | Reference |
| 1 | https://academic.oup.com/forestscience/article/64/1/33/4804510 |
| 2 | https://www.for.gov.bc.ca/hfd/pubs/Docs/Fgi/fgi06/part1.pdf |
| [3](https://www.researchgate.net/publication/222332466_Modified_site_index_equations_for_major_Canadian_timber_species) | <https://www.researchgate.net/publication/222332466_Modified_site_index_equations_for_major_Canadian_timber_species> |
| [4](https://www.sciencedirect.com/science/article/pii/S037811270900317X) | <https://www.sciencedirect.com/science/article/pii/S037811270900317X> |
| 5 | https://novascotia.ca/natr/library/forestry/reports/Report92.pdf |
| 6 | https://academic.oup.com/forestscience/article/63/3/283/4583992 |
| Crown radius | Reference |
| 1 | <https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1049&context=nrm_fac> |
| 2 | https://nefismembers.org/wp-content/uploads/2015/06/Russell-and-Weiskittel-2011-maximum-and-largest-crown-widths.pdf |
| Gamma | |  | | --- | | <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/97JD01107> | | <http://faculty.geog.utoronto.ca/Chen/Chen's%20homepage/PDFfiles/unp115_Jing6_AFM.pdf> | |
| Leaf size | <https://tidcf.nrcan.gc.ca/en/trees/factsheet/>  https://plants.usda.gov/core |
| Frac\_bark | |  | | --- | | <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/97JD01107> | | <http://faculty.geog.utoronto.ca/Chen/Chen's%20homepage/PDFfiles/unp115_Jing6_AFM.pdf> | | http://www.nrcresearchpress.com/doi/pdf/10.1139/x95-125 | |
| Crown semi height | |  | | --- | | <https://pubs.cif-ifc.org/doi/pdf/10.5558/tfc70174-2> | | https://ecampusontario.pressbooks.pub/forestmeasurements/chapter/5-4-live-crown-ratio/ | | https://pubs.cif-ifc.org/doi/pdfplus/10.5558/tfc62451-5 | | http://www.nrcresearchpress.com/doi/pdf/10.1139/x95-125 | | <https://pubs.cif-ifc.org/doi/pdf/10.5558/tfc70174-2> | | <https://www.srs.fs.usda.gov/pubs/misc/ag_654/volume_1/pinus/banksiana.htm> | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Soil type | Soil | Nominal | Saturated soil spectrum | 1-8 |
| Soil brightness | B | % | Scalar for saturated soil spectrum | 1-3 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| AOD 550nm | AOD550 | DIM | Aerosol optical depth at 550nm | 0-10 |
| Water vapour concentration | CH20 | g.cm-2 | Tropospheric water vapour concentration | 0-100 |
| ozone concentration | CO | mol | Total column ozone concentration | 0-100 |
| Atmospheric pressure | Patm | KPa | Atmospheric pressure | 90-130 |