# Climate Modelling User Group [CMUG]

# Deliverable 2.3f v1

# Technical report on optimisation of the ORCHIDEE snow model parameters with the CCI SCF and SWE products

Centres providing input: IPSL (LSCE)

| Version nr. | Date | Status |
|---|---|---|
| 0.1 | April 2025 | Input from partners |
| 1.0 | April 2025 | Submitted to Met Office |
| 1.1 | June 2025 | Minor revisions |
| | | |

# Deliverable 2.3f v1

# Technical report on optimisation of the ORCHIDEE snow model parameters with the CCI SCF and SWE products

## Contents

# Technical report on optimisation of the ORCHIDEE snow model parameters with the CCI SCF and SWE products

## 1. Purpose and scope of this report

This document is the technical report on the work carried out in the WP5.6 study of the CCI-CMUG project. Our study aims to improve the representation of snow cover dynamics in temperate and boreal zones within the IPSL climate model, using the snow cover fraction and snow water equivalent products recently released by the CCI Snow (covering the last four decades), to assess the impact of snow cover dynamics and atmospheric feedback on regional to continental climate.

Snow is a critical cryosphere component of the climate system. Its high albedo gives rise to the positive snow-albedo feedback that amplifies global climate variability and is thought to be a driver of the observed Arctic amplification of the current global warming and the observed amplification of global warming at high latitudes. Widely varying treatments of the vegetation masking of snow in forested areas are suspected to be a major reason for large inter-model variations in the intensity of the snow albedo feedback (Krinner et al., 2018).

In the IPSL climate model, the land surface processes and the interactions with the atmosphere are covered by the ORCHIDEE Land Surface Model (LSM) which includes a snow model (Wang et al., 2013; Charbit et al., 2024) able to represent the main physical processes occurring in a snowpack and its evolution according to the boundary conditions at both soil and atmosphere interfaces. Snow-vegetation interactions are not explicitly described, but snow albedo and sublimation depend on the grid vegetation through the specific optical properties and surface roughness of the grid cell. SWE products such as Globsnow (Takala et al., 2011; Luojus et al., 2021) have been used in previous works to evaluate the temporal dynamics of the snowpack, mainly over Siberia (Dantec et al., 2017; Guimberteau et al., 2018). Still, the evaluation has to be extended to the global scale. Furthermore, it is well-known that land cover strongly influences some key processes that drive the evolution of the snowpack, such as lateral transport, melting, refreezing, and sublimation. Therefore, the new snow products from the CCI-Snow project, especially the snow cover extent differentiating ground and viewable fractions, are valuable for improving and calibrating snow models.

Through this project, we will thus investigate how using the CCI Snow products in combination with the CCI medium/high-resolution land cover-derived products and possibly other datasets, can improve the snow dynamics in the IPSL climate model, mainly through optimizing land cover-specific snow parameters. The methodology that we will develop here, should also apply to other land surface models.

Snow modelling in climate models presents specific issues compared to stand-level modelling relative to the spatial and temporal scales. In a climate model, the size of the grid cell may reach several hundred kilometres, which, in general, means heterogeneous landscapes, reliefs and meteorological conditions within the grid. In cold weather, this results in heterogeneous snow coverage, large uncertainties in snowfall estimation and the need to consider the impacts of

vegetation and topography in the modelling. In ORCHIDEE, these interactions are not explicitly accounted for, but they are indirectly represented through the representation of a time-varying Snow Cover Fraction (SCF) dependent on the grid cell snow amount (SWE), and the modelling of the grid cell albedo, key model parameter driving the energy balance and the resulting processes. Given the strong interdependency of SCF, SWE and Albedo, these snow products are not enough to calibrate all parameterizations. Therefore, we are developing in this project a methodology based on the synergistic use of the three kinds of observations, to improve the modelling of the snowpack dynamics at the global scale, accounting for vegetation impacts, through an improved simulation of the snow albedo and resulting energy and water budgets.

In this report, we present the work performed on the two main tasks defined for this project, which are:

- The analysis of the CCI Snow products, the comparison with modelled variables and the preliminary assessment of their potential to evaluate the ORCHIDEE snow model jointly with other datasets such as albedo and land cover products;
- The use of the CCI Snow products to improve the ORCHIDEE snow parameterizations and better simulate the snowpack dynamics and the snow-atmosphere interactions.

# 2. ORCHIDEE Land Surface Model

## *2.1 General description*

ORCHIDEE is the continental part of the Earth System Model (ESM) of the Institut Pierre-Simon Laplace (IPSL). In this study, we worked with the ORCHIDEE-V3.0 version. The model simulates the energy and water transfers in the soil-atmosphere continuum and at the surface-atmosphere interface, as well as the carbon and nitrogen cycles and their interactions (Vuichard et al., 2019). Vegetation processes are parameterized for 15 different Plant Functional Types (PFTs) presented in Table 1, related to their phenology, leaf type, physiological activity (e.g. C3/C4 photosynthetic pathways) and climate (Harper et al., 2023). Soil mineral composition is defined at the grid-cell level, based on the 12-classes USDA soil classification (Forbes et al., 1987). While energy and snow processes are computed at the grid-cell scale (one energy budget for the whole cell), water processes are resolved by accounting for the PFTs present in the grid cell, with a maximum of three soil columns separating bare soil from high (e.g., forests) and low (e.g., grasslands) vegetation. Carbon stocks and fluxes are resolved for each PFT present within the grid cell. Different vertical grids are considered depending on the biogeophysical cycles considered. Soil hydrology and thermics share the same grid consisting of 11 layers from the surface and down to 2 m with a geometrically increasing internode distance. Thermal processes are resolved deeper, up to a maximum depth of 90 m, with 7 additional layers and a number of layers of 18 in the standard version. All thermal and hydrological properties are calculated according to soil composition (mineral, carbon and water components), accounting for soil freezing and snow thermo-hydric processes.

ORCHIDEE can be run at various scales ranging from local to global. For the present study, the model requires the prescription of the atmospheric conditions (air pressure, temperature and humidity, wind speed and incoming radiation), a description of the land surface (fraction of the different PFTs, dominant soil texture) and the initialization of the energy, water, carbon and

nitrogen stocks. These conditions are derived from global datasets (i.e., atmospheric reanalysis, land cover and soil texture classifications) or from stand-level measurements.

| |
|---|
| PFT1: Bare Soil |
| PFT2: Tropical Evergreen |
| PFT3: Tropical Raingreen |
| PFT4: Temperate Needleleaf Evergreen |
| PFT5: Temperate Broadleaf Evergreen |
| PFT6: Temperate Broadleaf Summergreen |
| PFT7: Boreal Needleleaf Evergreen |
| PFT8: Boreal Broadleaf Summergreen |
| PFT9: Boreal Needleleaf Deciduous |
| PFT10: Temperate Natural Grassland (C3) |
| PFT11: Natural Grassland (C4) |
| PFT12: Crops (C3) |
| PFT13: Crops (C4) |
| PFT14: Tropical Natural Grassland (C3) |
| PFT15: Boreal Natural Grassland (C3) |

*Table 1: The 15 Plant Functional Types used in the ORCHIDEE model to describe vegetation*

## 2.2 ORCHIDEE Snow model

The work performed in WP5.6 aims to use the CCI snow products to improve the modelling of snow dynamics in ORCHIDEE. In our model, the snow processes are represented with a physically-based approach where the main processes driving the temporal evolution of the snowpack such as ageing, compaction, sublimation, melting and refreezing, are represented with a 1D-physical system neglecting the lateral transfers. In the original version of the model (Wang et al., 2013), the snowpack is vertically discretized in 3 layers for which snow temperature, density and liquid water content are prognostic variables. Snow is uniformly distributed over the grid cell regardless of vegetation distribution, and the snow cover fraction (SCF) ranges between 0 and 1 according to the snow amount. SCF is parameterized following the formulation of Niu and Yang (2007) which has been shown to better represent the seasonal variation of the relationship linking snow mass and SCF thanks to its dependence on snow density:

$$SCF = \tanh\left(\frac{Z_{snow}}{2.5 z_{0g} \times \left(\frac{\langle \rho_{snow} \rangle}{\rho_{min}}\right)^{m}}\right) \qquad (1)$$

where $\rho_{snow}$ is the snow density averaged over the total thickness of the snowpack, $\rho_{min}$ is the minimum value of snow density (set to 50 kg m$^{-3}$), that is the density of fresh snow, $z_{0g}$ is the ground roughness length (set to 0.01 m) and $m$ (set to 1.0 in ORCHIDEE) is an adjustable parameter.

The energy balance at the snow-atmosphere interface is solved at the model time step (30 mn), mostly driven by the snow albedo, which is a key parameter in the model, dependent on snow age and temperature.

Compared to the earlier version presented by Wang et al. (2013), the albedo scheme has been modified and snow albedo is now computed following the formulation of Chalita and Le Treut (1994):

$$\alpha_{snow} = A_{aged} + B_{dec} exp\left(-\frac{\tau_{snow}}{\tau_{dec}}\right) \quad (2)$$

where $A_{aged}$ represents the albedo of a snow-covered surface after snow aging (old snow) and $B_{dec}$ is defined so that the sum of $A_{aged}$ and $B_{dec}$ represents the albedo of fresh snow (i.e., maximum snow albedo). Equation (2) is used to calculate both the visible and the near-infrared (NIR) albedo, with different parameters $A_{aged}$ and $B_{dec}$. $\tau_{dec}$ is the time constant of the albedo decay and $\tau_{snow}$ is the snow age and is parameterized as follows:

$$\tau_{snow}(t+dt) = \left[\tau_{snow}(t) + \left(1 - \frac{\tau_{snow}}{\tau_{max}}\right) \times dt\right] \times exp\left(-\frac{P_{snow}}{\delta_c}\right) + f_{age} \quad (3)$$

where $\tau_{max}$ is the maximum snow age, $P_{snow}$ is the amount of snowfall during the time interval $dt$ and $\delta_c$ is the critical value of solid precipitation necessary for reducing the snow age by a factor 1/e. As the ORCHIDEE time step is fixed to 30 mn, the snow age is almost zero in a few time steps. In addition, low surface air temperatures found in arctic regions slow down the metamorphism. This effect is accounted for with the function $f_{age}$ expressed as:

$$f_{age} = \left[\frac{\left(\tau_{snow}(t) + \left(1 - \frac{\tau_{snow}}{\tau_{max}}\right) \times dt\right) \times exp\left(-\frac{P_{snow}}{\delta_c}\right) - \tau_{snow}(t)}{1 + g_{temp}(T_{surf})}\right] \quad (4)$$

$$g_{temp}(T_{surf}) = \left[\frac{max\,(T_0 - T_{surf},\ 0)}{\omega_1}\right]^{\omega_2} \quad (5)$$

Where $\omega_1$ and $\omega_2$ are tuning constants. The albedo is computed for the visible and near-infrared spectral bands. However, to compute the upward shortwave radiation, an arithmetic mean between the visible and the near-infrared albedo is considered.

Recently, Charbit et al., (2024) extended this snow scheme to ice sheets and for that purpose, modelled the snow-ice interface and used new datasets to validate the snow mass balance. In this work, they showed the improvements brought to the snow temperature prediction by replacing the 3-layer discretization with a 12-layer one, following Decharme et al., 2016. Therefore, the last and shared version of ORCHIDEE (Trunk version) now includes this updated scheme which has still not been fully validated and calibrated over continental surfaces. Furthermore, the Trunk version uses an updated scheme for the computation of the vegetation albedo which still requires calibration.

Our intent is, therefore, to benefit from the new CCI Snow products to revise the calibration of the snow model in the Trunk version (that will be used for the CMIP7 model simulations of the IPCC Assessment Report (AR7)). But, given the strong links between SCF and the albedo grid cell, snow albedo calibration requires the prior calibration of the vegetation albedo which is a priority for the ORCHIDEE team and should be performed in spring 2025. Meanwhile, we have worked on the previous ORCHIDEE-V3 version (Vuichard et al., 2019) to develop the

calibration methodology and intend to apply it to the most recent version of ORCHIDEE (Trunk) in the future.

Besides, parallel work is ongoing in our group on the representation of soil thermics in Arctic ecosystems and the impacts of soil organic carbon on thermal soil properties (Gaillard et al., 2025; Cuynet et al., submitted). In particular, Cuynet et al., (submitted) developed new parameterisations, including a better interpretation of the soil dataset Soilgrids 2.0 (Poggio et al., 2021), used in ORCHIDEE to prescribe soil properties.

We can note also that the last ORCHIDEE versions include new PFT maps derived from the ESA CCI MRLC project (PFT V3.0 product, Harper et al., 2023) which show significant improvements compared to the former ones (Lurton et al., 2021), especially in Arctic areas, where shrublands were overestimated at the expense of grasslands and bare soils (Harper et al., 2023; Ottlé et al., internal reports).

All these developments are now merged in the Trunk version and will benefit our further developments.

# 3. Observation products description

## *3.1 Presentation of the CCI Snow products*

The Snow CCI products (Solberg et al., 2021) provide several global daily time series of essential climate variables (ECV) related to snow:

- Snow Cover Fraction Viewable (SCFV)
- Snow Cover Fraction Ground (SCFG)
- Snow Water Equivalent (SWE)

The SCFV corresponds to snow on top of open areas and vegetation like forest canopies, while the SCFG is the snow on the ground for open land, corrected for masking by trees in forested areas. The snow cover for each grid cell is given as a percentage. The SWE indicates the amount of snow accumulated on land surfaces, as an equivalent height of water.

In this task, we worked with the CCI Snow V2.0 product until December 2024 and switched to the V3 versions released in December 2024 (V3.1 for SWE and V3.0 for SCF). More details on the products are provided in Table 2.

## *3.2 Presentation of the other products: albedo and PFT maps*

The MODIS daily global albedo product (doi:10.5067/MODIS/MCD43C3.061) is also used in this analysis, since SCF and albedo are interrelated. The MODIS Albedo products MCD43A3 available at a resolution of 0.05° and over the period 2000-2020 was chosen to evaluate the model predictions, analyse model errors and calibrate the model parameters.

In order to describe the vegetation, we used the new PFT maps derived from the ESA CCI MRLC project (PFT V3.0 product, Harper et al., 2023).

| | SCFG | SCFV | SWE |
|---|---|---|---|
| Product name | Snow Cover Fraction Ground | Snow Cover Fraction Viewable | Snow Water Equivalent |
| DOI | 10.5285/8847a05eeda646a29da58b42bdf2a87c | 10.5285/ebe625b6f77945a68bda0ab7c78dd76b | 10.5285/4647cc9ad3c044439d6c643208d3c494 |
| Description | Daily snow cover fraction (0-100%) **on the ground** per pixel for global land areas (excluding pixels containing more than 50% of permanent snow and ice, and pixels containing more than 30% of water) | Daily snow cover fraction (0-100%) **on top of the forest canopy** per pixel for global land areas (excluding pixels containing more than 50% of permanent snow and ice, and pixels containing more than 30% of water) | Daily snow water equivalent (in mm) per pixel, representing the amount of water stored in the snowpack<br><br>Covers the Northern Hemisphere land areas, excluding mountainous regions, glaciers, and Greenland |
| Cloud handling | Clouds are masked | | Not affected by clouds |
| Data Source | Medium-resolution optical satellite data | | Low-resolution passive microwave satellite data combined with in-situ snow depth measurements |
| Instrument | MODIS / AVHRR (separate products) | | SMMR, SSM/I and SSMIS PMR (merged product) |
| Spatial resolution | ~1 km per pixel for MODIS (coarser resolution for AVHRR) | | 0.10-degree latitude-longitude grid |
| Uncertainty | Unbiased Root Mean Square Error (RMSE) per pixel | | Statistically derived accuracy estimates on each pixel level |
| Auxiliary datasets | - **ESA CCI Land Cover 2000**: Masks water bodies and permanent snow/ice (aggregated to the pixel size of SCF product)<br>- **Forest Canopy Transmissivity Map**: Derived from ESA CCI Land Cover 2000 and Landsat tree cover density map (Hansen et al., 2013, DOI: 10.1126/science.1244693) for canopy correction | - **ESA CCI Land Cover 2000**: Masks water bodies and permanent snow/ice (aggregated to the pixel size of SCF product) | - **ESA CCI Land Cover 2000**: Masks water bodies (aggregated to the pixel size of SWE product)<br>- **ETOPO5-based Mountain Mask**: Used to exclude mountainous regions in the Northern Hemisphere (developed for the ESA GlobSnow project) |
| Other information | Forest canopy correction is applied based on the forest canopy transmissivity map | | |

*Table 2: Snow CCI products description*

# 4. Methods: Processing of the observation datasets

## *4.1 Upscaling of the CCI Snow maps*

The SCF maps are provided at 0.01° resolution, which is very refined for the purpose of our study. In order to be able to compare the SCF to the albedo product, while still keeping a fine resolution, the products are upscaled to a resolution of 0.05°. To do so, for each day, all the pixels belonging to a 0.05° pixel are averaged together to provide the SCF of the pixel. If there is a NaN value in this set of pixels, the resulting value is a NaN. It enables maintaining a good reliability on the product and does not let isolated valid pixels result in an erroneous mean value at a larger scale. These are the maps that are used for site selection, since they provide the best resolution.

Another use of these maps is to validate as well as optimize the ORCHIDEE LSM. Since the resolution of the simulations is coarser (0.5° resolution), the raw maps are also upscaled to that resolution. However, the process is slightly different because of the scarcity of valid values in a complete 0.5° pixel each day. In order to avoid this issue, a bilinear interpolation is performed from the original data at 0.01° after pre-selecting only the pixels with a valid range of values.

| Resolution | 0.01° | 0.05° | 0.1° | 0.5° | 2° |
|---|---|---|---|---|---|
| SCFV/SCFG (Snow CCI) | Original product<br><br>Daily file | Daily<br>Average if no NaN values<br>Else: NaN<br>From 0.01° map | | Daily<br>Bilinear interpolation from 0.01° map | Monthly<br>Piecewise linear time filling and average from 0.05° map |
| SWE (Snow CCI) | | | Original product<br>Daily file | Daily<br>Average if no NaN values<br>Else: NaN<br>From 0.1° map | Monthly<br>Piecewise linear time filling and average from 0.5° map |
| Albedo (MODIS) | | Original product<br>Daily file | | | Monthly<br>Piecewise linear time filling and average from 0.05° map |
| PFT (PFT V3.0 product, Harper et al., 2023) | | Aggregated PFT map from 300 m resolution map | | | |

*Table 3: Products' resolutions and upscaling methods*

Finally, since the outputs of the model simulations that were used for the comparison are at a monthly resolution, daily values cannot be compared directly. However, the observation time

series are generally incomplete due to cloud cover or polar night for some days, and the mean of the remaining values can be erroneous. To limit this effect, a gap-filling is performed on the timeseries, by performing a simple piecewise linear interpolation, and the monthly mean value is determined afterwards. The different methods used for each product are summarised in Table 3.

## *4.2 Valid data computing*

For each year of CCI Snow data (SCF products), we generate maps that contain, for each pixel, the number $n_{snow}$ of valid points during the year, since quite a lot of days can be missing, due to the night or the presence of clouds. The use of these maps is further described in section 4.4, as part of the site selection algorithm to perform ORCHIDEE site optimisations. The selection algorithm can vary depending on the needs, but the most used condition here is the following: only valid days (i.e., observations provided from the initial product), where the SCF values are strictly above 0%. That way, we can make sure that we select points where snow is present. For global optimisation, the threshold is set to a higher value (above 50% or 75%).

## *4.3 PFT maps processing*

Here, the PFT maps are needed to study the products by vegetation type. The PFT maps provide the repartition of each PFT and exist at multiple resolutions, based on the 300m CCI Land Cover Product (Harper, 2023). The aggregated product at a 0.05° resolution is used.

Since the study is mainly performed during the decade 2011-2020, we want to make sure that the sites we select did not change in terms of vegetation during this period. To do so, an Euclidian distance is calculated between the maps of each year to identify sites subject to changes. If the value of the calculated distance reaches a value above 0.1, the site is removed since it changed too much during the decade.

Moreover, since we want sites that are representative of a certain PFT, the regions where there is no dominant PFT (i.e. one PFT represents more than 50% of the pixel) are removed.

## *4.4 Site selection based on the PFT maps and the SCF product*

Since one of the goals of the study is to assess the influence of the vegetation on the snowpack temporal dynamics, it is essential to be able to select sites with homogeneous vegetation. The method developed here relies on the PFT maps which were pre-processed as presented earlier.

To select the most suitable sites for the analysis, it is necessary to find locations with the highest fraction of the considered PFT. Since the analysis focuses on the retroaction between snow and vegetation, a site of interest would be covered by snow for a long period of time, and data would be available for most days. The selection approach deals with all these aspects.

It is based on the concept of acquisition functions used in Bayesian optimization, which aims to estimate the potential gain in information by exploring a spatial domain. The idea is to assign a score to each point of this domain and then select those with the highest scores.

This score, which is further called the "desirability" of the point, must consider the dominance of the relevant PFT as well as the number of available measurements for days when the pixel is almost completely snow-covered. For each pixel $x$, the desirability function $V(x)$ is calculated as a function of the PFT fraction and the number of days where snow data is available $n_{snow}(x)$ with the SCF of the pixel over a certain threshold (0% in the study).

$$V(x) = n_{snow}(x) \times PFT\ fraction(x) \quad (7)$$

For each PFT, the $N = 100$ points that maximize the desirability function are pre-selected. Additionally, the selection must ensure that the chosen points are not all concentrated in the same region. Therefore, the final points are selected sequentially among this set of points, with the following process.

We consider the set of $N$ points $P = \{p_1, p_2, \dots, p_N\}$, that were selected in the previous phase, where each point $p_i$ is defined by its coordinates $(lat_i, lon_i)$ and its desirability function value $v_i$.

The first selected point $p^*$ is the one with the highest desirability function value such that $p^* = \arg\max_{p_i \in P} v_i$.
We initialize the set of selected points $S = \{p^*\}$.

Then, we select iteratively the most distant points. At each iteration, we select the point $p_{new}$ that is the farthest from the set of already selected points. Distance is measured using the Manhattan distance:

$$d(p_i, p_j) = |lat_j - lat_i| + |lon_j - lon_i|$$

The next point is chosen by maximizing the minimum distance to the selected set:

$$p_{new} = \arg\max_{p \in P \setminus S} \min_{q \in S} d(p, q)$$

Once selected, this point is added to the set: $S \leftarrow S \cup \{p_{new}\}$
The process continues until all points have been selected, ensuring that each newly added point is as far as possible from the previously selected ones.

For each PFT, 10 sites are selected, if the desirability function is high enough.

# 5. Analysis of the CCI Snow products

## *5.1 Comparison of the MODIS SCF and the AVHRR SCF (CCI Snow V2)*

Snow CCI provides SCF products from different sensors, including MODIS and AVHRR (among others), which were chosen because of their global resolution and the long time series they provide. The SCF values have been compared in a small region and are presented in Figures 1 (SCFV) and 2 (SCFG) with their uncertainties. The patterns are similar, but differences between the two products can be observed. Moreover, the uncertainties provided with AVHRR are superior to those provided with MODIS data. The differences between the MODIS and AVHRR SCF are calculated and shown in Figure 3. The SCFG differences are small, because on the selected date, the ground is entirely covered by snow (SCFG close to 100%), while bigger differences can be noticed between the SCFVs. When considering Figure 4 which shows the most common PFT per pixel in the same area, it is interesting to note that the patterns in the SCFV differences seem to be related both to latitude and to the dominant PFT map. Indeed, in the presence of grasslands or PFT9, MODIS SCFV is below that of AVHRR, while when PFT7 is mostly represented, the MODIS SCFV is bigger. Also, larger differences are observed in the northern part of the domain with a clear threshold around 55°N. These features seem to be related to the different characteristics of MODIS and AVHRR sensors and orbital parameters as well as methods used to correct bi-directional reflectances.

From this preliminary analysis, the choice was made to use the SCF CCI products derived from MODIS sensors, for data analysis and model optimization, since the resolution is finer and the uncertainties are smaller.

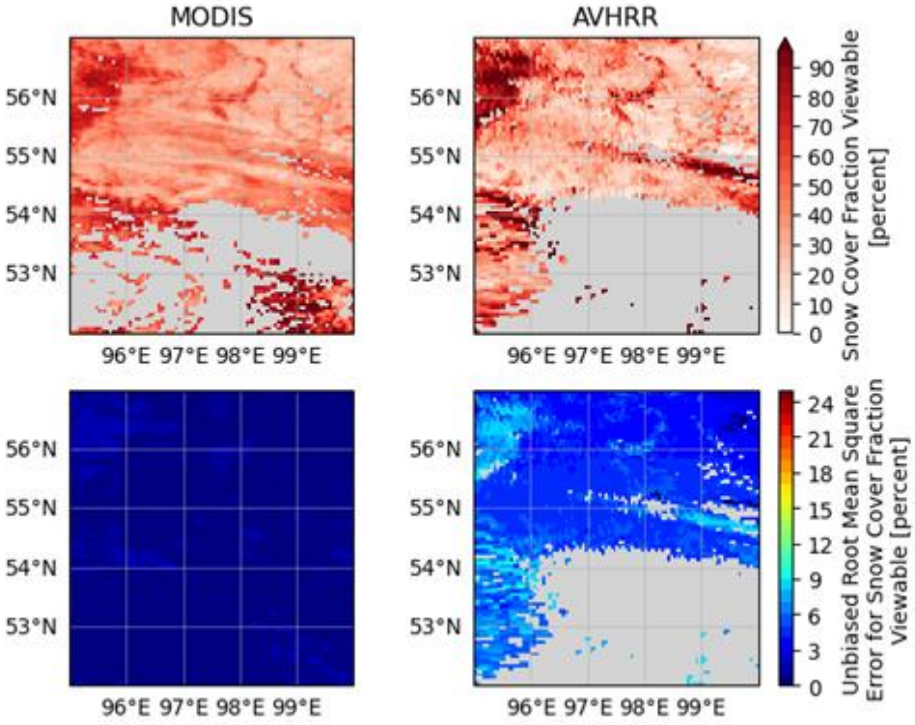


*Figure 1: SCFV from MODIS (upper left) and AVHRR (upper right), and the corresponding unbiased RMSE (lower panels) – day: 02/02/2010*

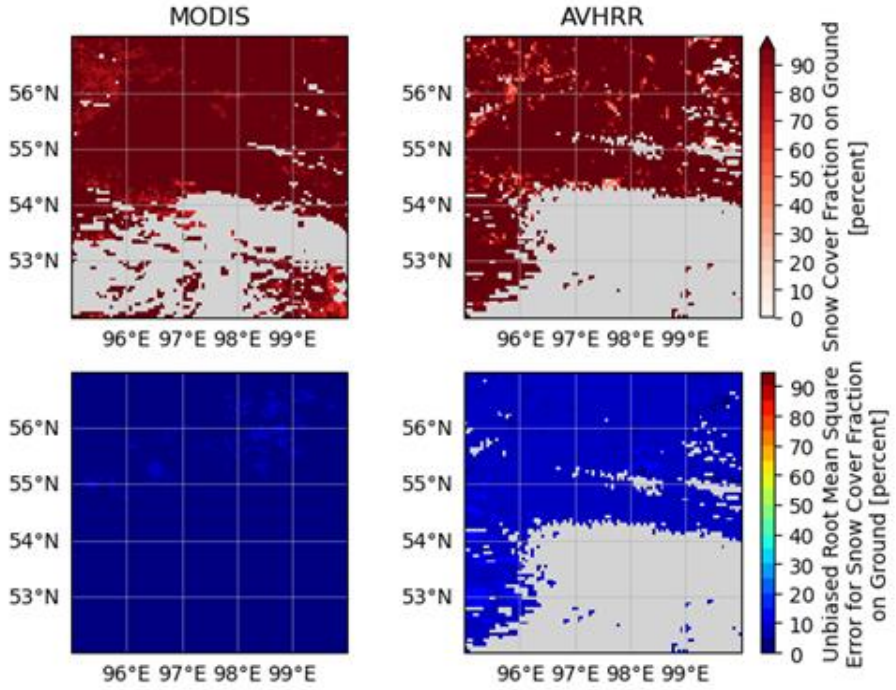


*Figure 2: SCFG from MODIS (upper left) and AVHRR (upper right), and the corresponding unbiased RMSE (lower panels) – day: 02/02/2010*

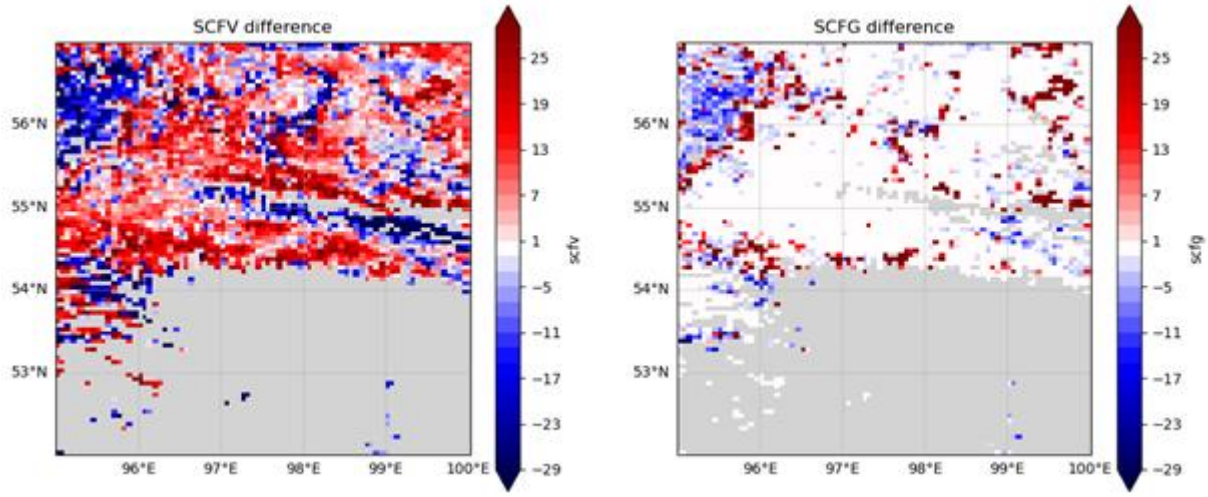


*Figure 3: Differences between MODIS and AVHRR SCFV (left) and SCFG (right) – date: 02/02/2010*

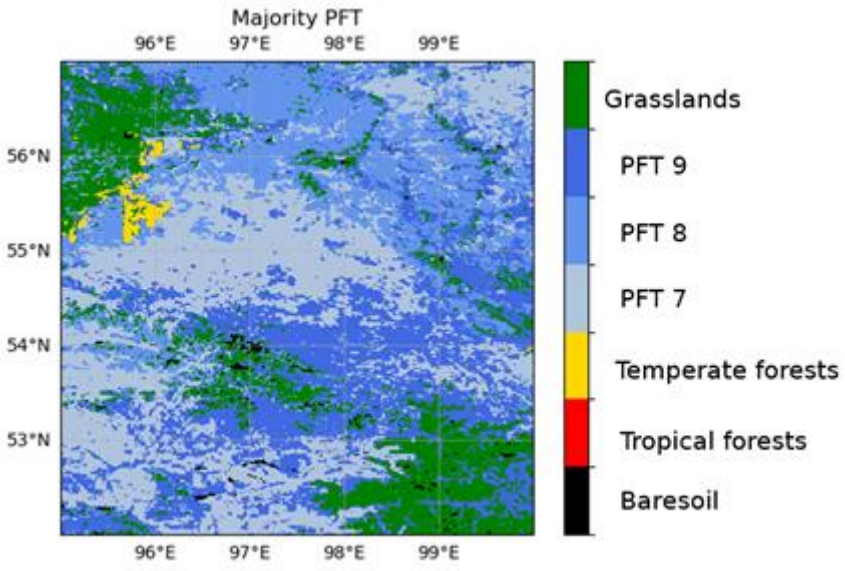


*Figure 4: Most common PFT per pixel in the studied region extracted from the ORCHIDEE PFT map derived from the CCI – Medium Resolution Land Cover dataset*

## 5.2 Site evolution of SCF and SWE

Time series of SCF and SWE on different sites have been studied. An example is provided in Figure 5, with the evolution of the SCFG, SCFV and SWE, with their uncertainties. The evolution of the different time series is consistent, and uncertainties related to the SCF are not too wide, while the SWE presents larger uncertainties. However, improvements were noted in v3.1 of the SWE product in terms of uncertainties compared to the earlier versions of the product. The time series presented in Figure 5 are derived from v3.1, and show lower uncertainties than in the previous version.

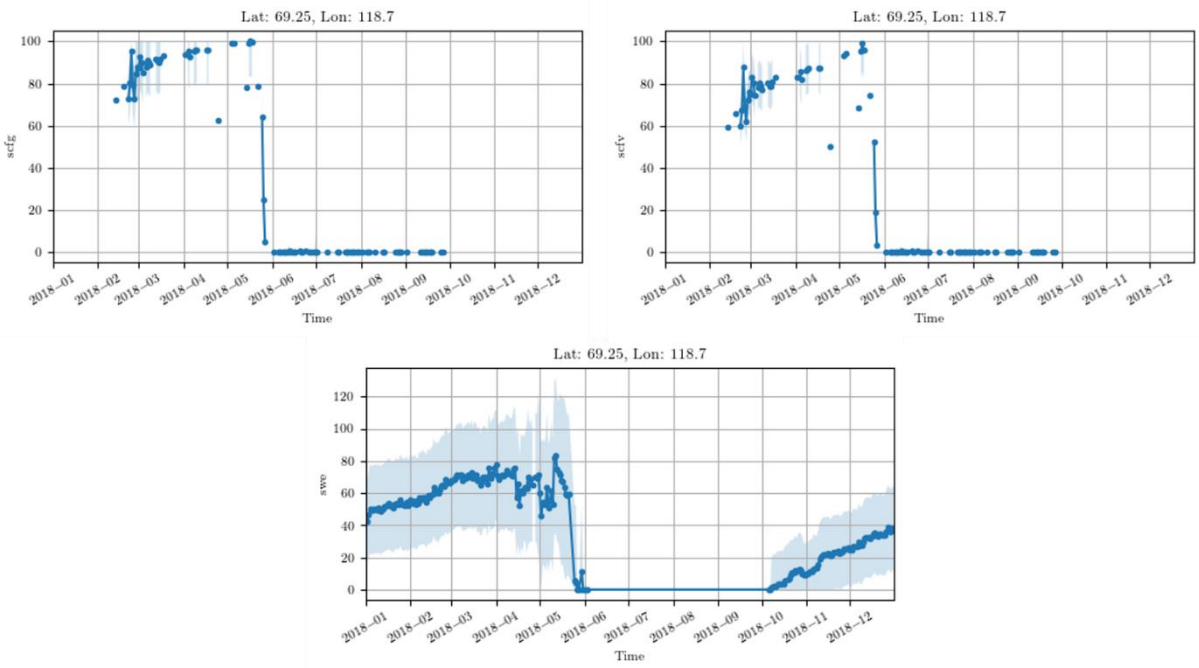


*Figure 5: From left to right, and top to bottom: time series of MODIS SCFG [%], MODIS SCFV [%] and SWE [mm] with the provided uncertainties at a selected point in Siberia (69.25°N, 118.70°E)*

## 5.3 Influence of vegetation type on snow dynamics

It is complex to attribute the evolution of the Snow Cover Fraction to the vegetation, because of the huge spatial variability of the meteorological conditions at the scale of the remote sensing products (i.e., a few kilometres at best). It is therefore not possible to compare the results obtained at sites that are too distant from each other since the meteorological conditions may vary widely. A first attempt at identifying the role of the vegetation on the snow behaviour was thus made by selecting a small part of a region that was considered to be fairly homogeneous in terms of snowfall and topography.

Therefore, a region of 0.5°×0.5° has been selected in Siberia (latitudes between 55.5 and 56°N, and longitudes between 96.5 and 97°E), which is mainly covered by two PFTs: boreal needleleaf evergreen vegetation (PFT7) and boreal broadleaf summergreen vegetation (PFT8), as seen on the right panel in Figure 6, that indicates the most common PFT on 0.05° large pixels when it represents more than 50% of the area. In this particular region, there are 30 pixels that mostly contain PFT7, and 25 pixels containing PFT8. To produce the time series appearing in Figure 6, that show the evolution of SCFV and SCFG for both PFTs, a spatial averaging is performed for each day and each PFT. The number of points where data are available vary from one day to another because of the possible presence of clouds, that sometimes prevents the measurement of the SCFs.

Figure 6 shows that the SCFV for PFT7 is lower than for PFT8, while the differences between the SCFG are less important. Based on this first result, the impact of the vegetation on the SCF can already be emphasized. It means that it is not possible to treat all PFTs the same way when it comes to snow cover fraction. The parameters to determine the SCF in ORCHIDEE need to be adapted to each PFT.

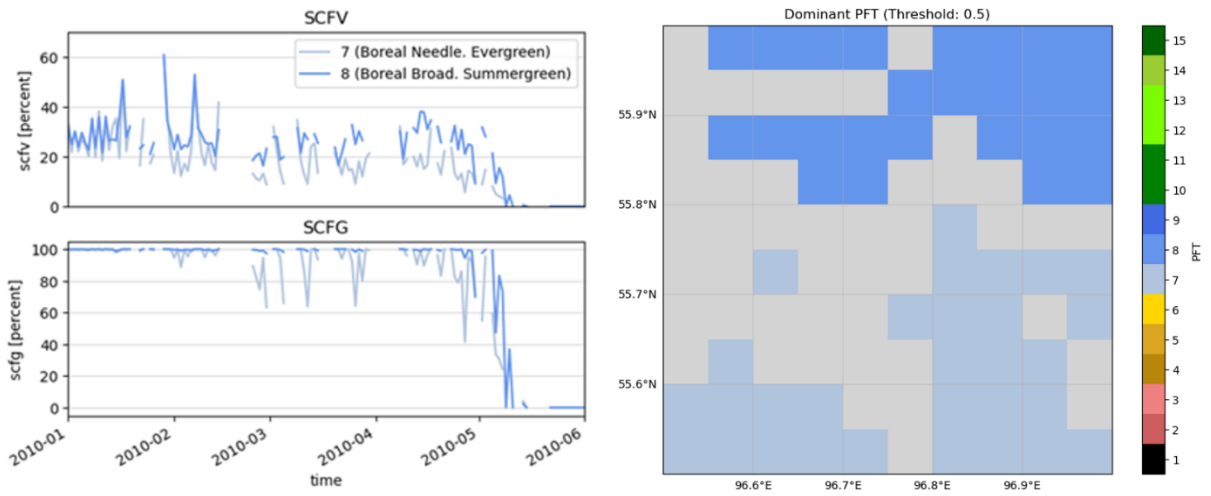


*Figure 6:Comparison of the SCFV and SCFG extracted over two types of vegetation (boreal needleleaf evergreen and boreal broadleaf summergreen) in the same region (Siberia) (left), and repartition of the most common PFT in the region derived from the CCI – Medium Resolution Land Cover dataset (right)*

# 6. Comparison of the SCF, SWE and albedo products to ORCHIDEE simulations

In this section, the outputs of two global reference simulations where the snow parameters were not optimized, performed respectively with ORCHIDEE v3 and the Trunk version were compared to the Snow CCI products and the MODIS albedo product. Only the regions above 30°N are shown here.

## *6.1 Comparison with ORCHIDEE v3*

Figure 7 shows the mean albedo Root Mean Square Error (RMSE) over the period 2011-2019 between the monthly averaged albedo from ORCHIDEE V3 and the MODIS albedo product. The RMSE values are generally around 0.05 but can reach values as high as 0.15 in high latitudes, especially in Eastern Siberia and the North of Canada, regions where snow generally tends to have a large impact on the albedo. Figure 8 provides a representation of the monthly averaged differences of albedo between ORCHIDEE v3 and MODIS over the period 2011-2019. Overall, the albedo tends to be slightly underestimated in winter, except in the mountainous areas and in particular the Himalayan range, where the observations might not be completely reliable due to the relief of that region and therefore limiting our ability to draw conclusions. From June to September, biases remain very low, with the exception of regions that are very far north. The global RMSE is therefore largely impacted by the winter month differences. It means that the errors linked to the albedo calculations in this version of the model are mostly linked to an erroneous estimation of the snow albedo, therefore showing the importance of better parameterizing the snow processes.

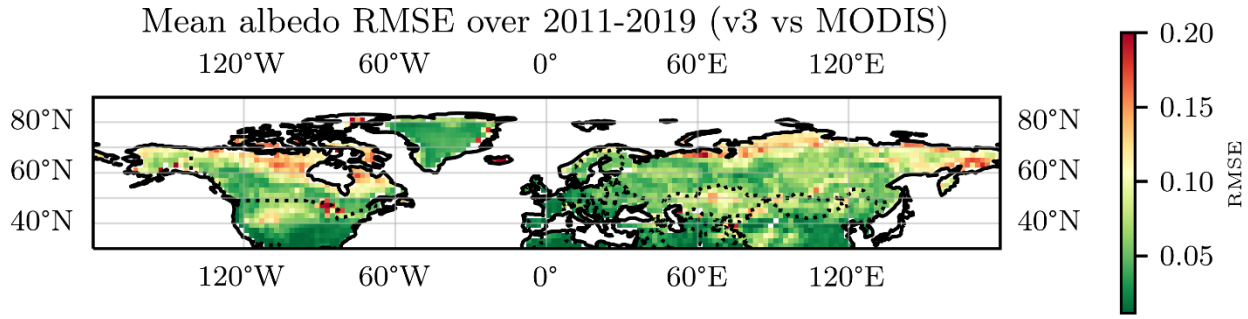


*Figure 7: Mean albedo RMSE over 2011-2019 between ORCHIDEE v3 and MODIS albedo.*

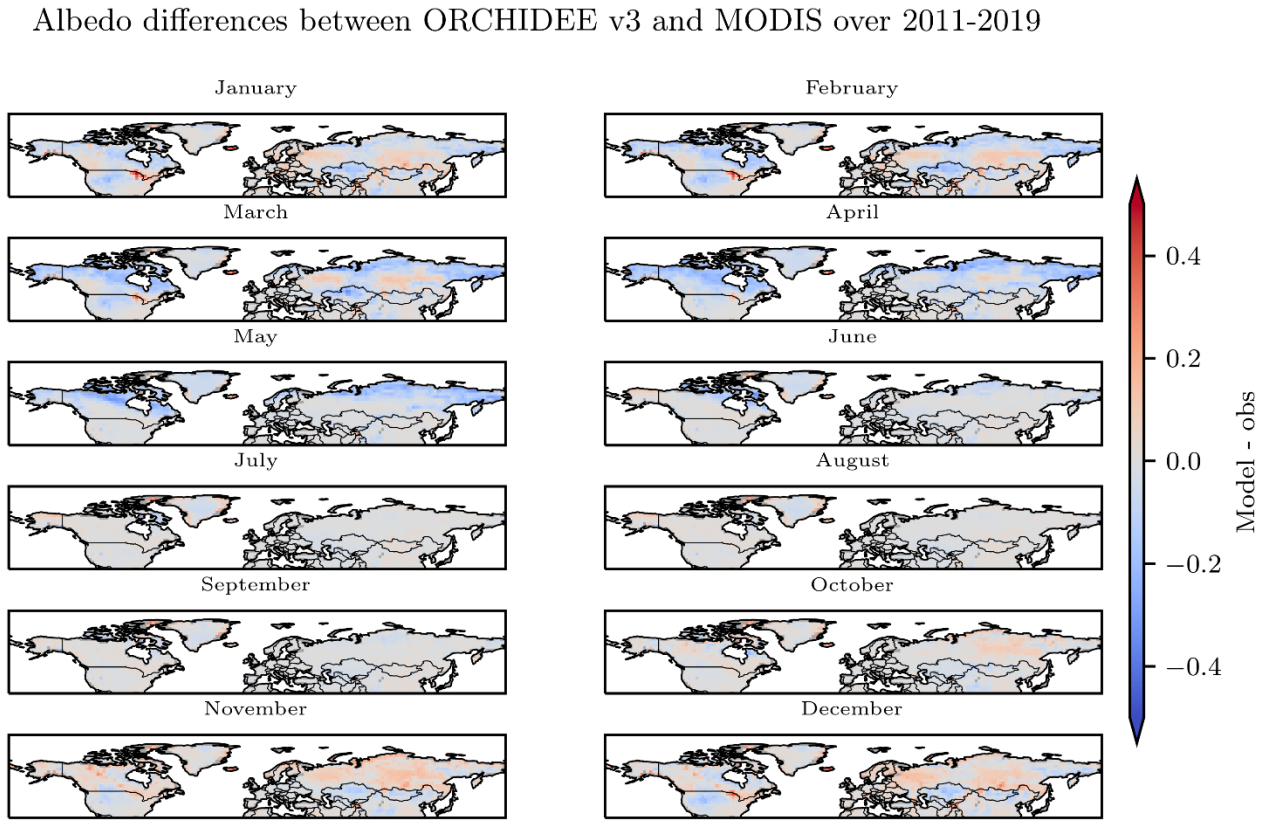


*Figure 8: Differences between the albedo computed in ORCHIDEE v3 and the albedo measured by MODIS (monthly averages)*

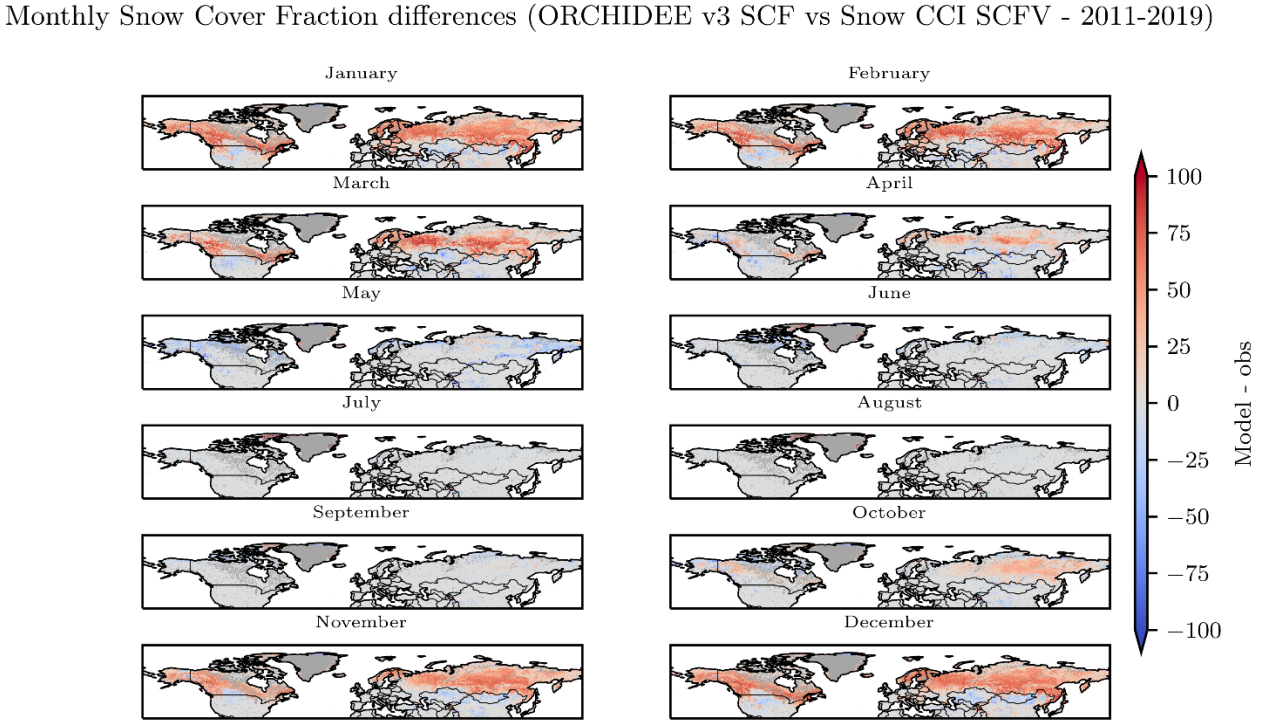


Figure 9 provides the monthly snow cover fraction differences between ORCHIDEE v3 and the

SCFV based on the MODIS instruments for each month averaged over 2011-2019. Most of the year, a large overestimation of the SCF is visible compared to the SCFV product, that does not reproduce the patterns seen with the albedo. Indeed, the albedo of snow being larger than that of the vegetation, a large SCF should imply larger albedo values. It means that the albedo parameters tend to be very underestimated by the model and should be reoptimized when the SCF is important, while the SCF should also be optimized separately.

Finally, Figure 10 shows the Snow Water Equivalent monthly-averaged differences between ORCHIDEE v3 and the CCI Snow SWE product over 2011-2019. June, July, August and September are not shown since SWE data is not provided during these months. Differences between the model and observations rarely exceed 40 mm. The analysis of this Figure corroborates the conclusions regarding the SCF patterns, with a larger SWE computed by ORCHIDEE most of the time. However, during the melting period, the SWE tends to be underestimated in some regions suggesting that ORCHIDEE models the snow melting too early in the season, indicating a potential issue in the melting dynamics.

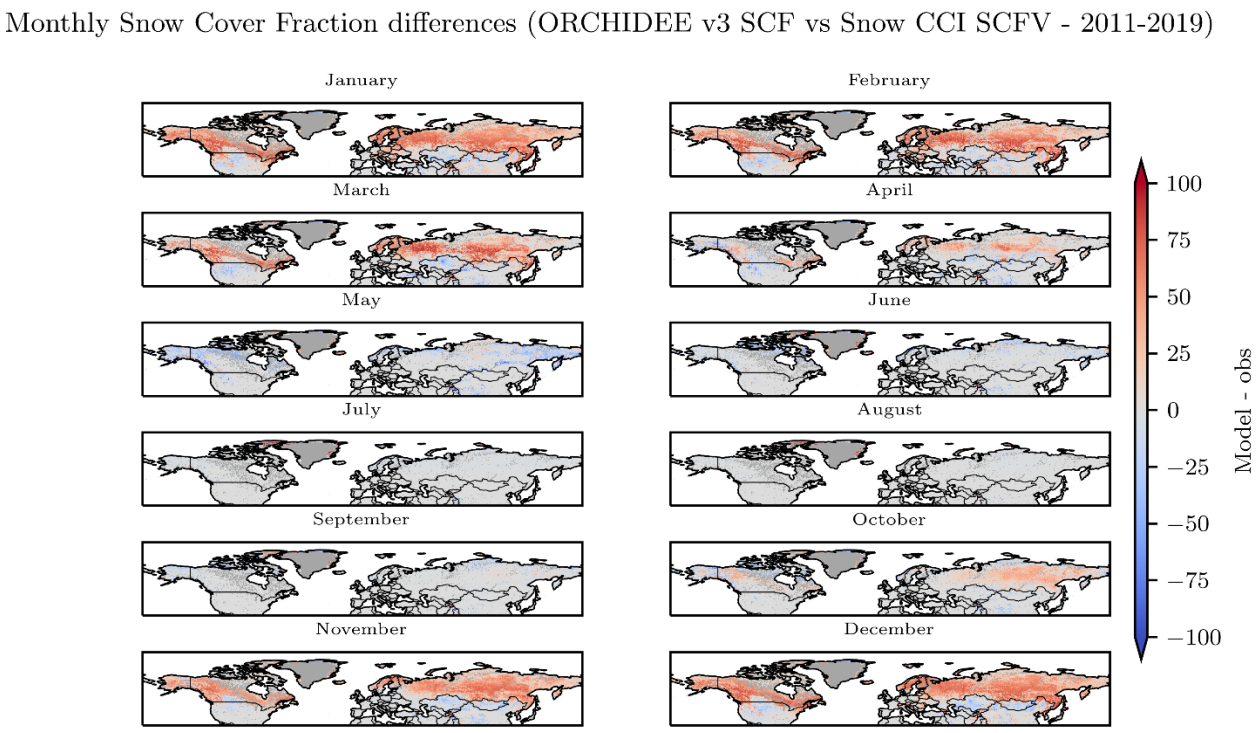


*Figure 9: Difference between the SCF calculated in ORCHIDEE v3 and the observed SCFV (monthly averages)*

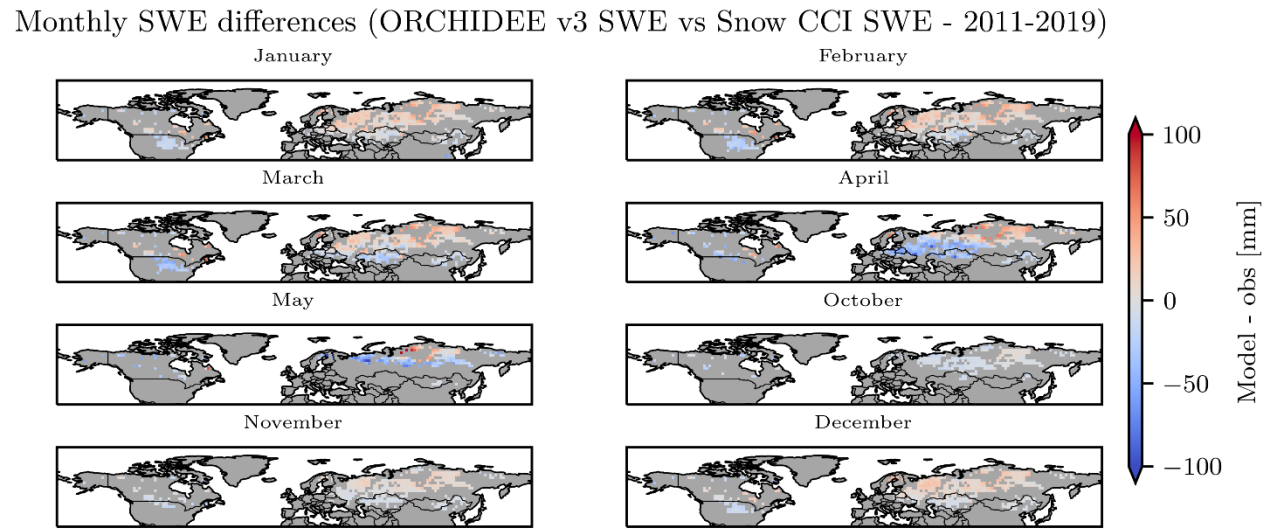


*Figure 10: SWE monthly differences [mm] between ORCHIDEE v3 and Snow CCI - 2011-2019. June, July, August and September are not shown since the SWE product generally does not provide data during this period. The regions in grey correspond to pixels where data were unavailable.*

## 6.2 Comparison with ORCHIDEE V4 (Trunk)

A global simulation was run with ORCHIDEE v4 at a 2° resolution. The albedo parameters were set to their standard values in that version. The outputs of the model were compared to the mean MODIS albedo, as well as to the MODIS-derived SCFV and SWE provided by Snow CCI (V3).

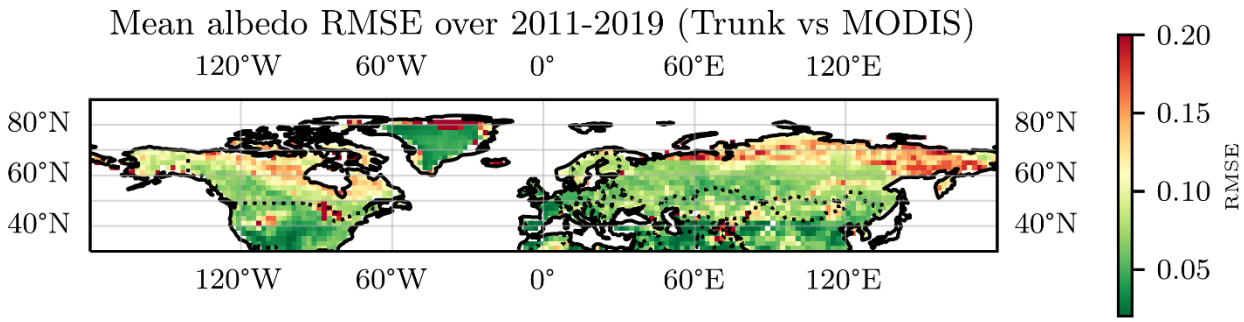


*Figure 11: Mean albedo RMSE over 2011-2019 between ORCHIDEE Trunk and MODIS albedo.*

Figure 11 shows the mean albedo RMSE between the Trunk and the MODIS product over 2011-2019. The RMSE values are higher in the Trunk than they were in v3, sometimes attaining 0.20 in particular in high-latitude regions and eastern Siberia, as well as in the North of Greenland. Additionally, RMSE values are elevated in a region of North America that might correspond to the Great Lakes region (modelled as bare soils in this version of ORCHIDEE).

Figure 12 provides more detailed information regarding the seasonal behaviour of the model, by showing the monthly-averaged differences between the Trunk and the MODIS albedo. The model exhibits a strong underestimation of albedo in winter and at high latitudes. Conversely, an overestimation of albedo occurs in summer and in areas where snow has melted, highlighting the need for vegetation albedo calibration.

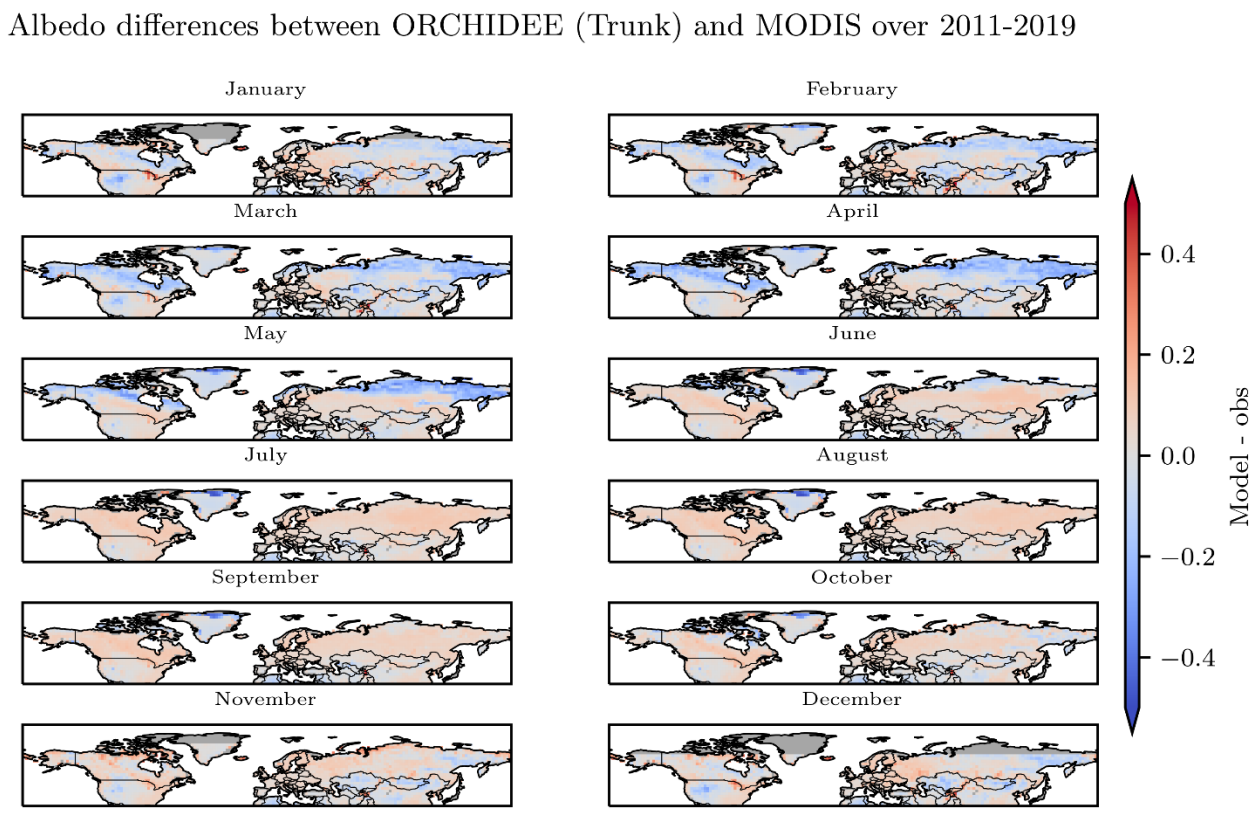


*Figure 12: Mean albedo monthly differences over 2011-2019 between ORCHIDEE Trunk and MODIS albedo. The regions in grey correspond to pixels where data were unavailable.*

Figure 13 provides the SCFV monthly differences between the Trunk and the Snow CCI SCFV. Similar to v3, SCF remains overestimated in regions where snow is present throughout the year. However, there is a more widespread underestimation of SCF in May due to early melting.

Figure 14 shows the SWE differences between the Trunk and the Snow CCI product. The pattern observed in this figure is very similar to that of v3. Early melting at the end of the snow season is apparent, even though the SCF differences are less significant in April and May.

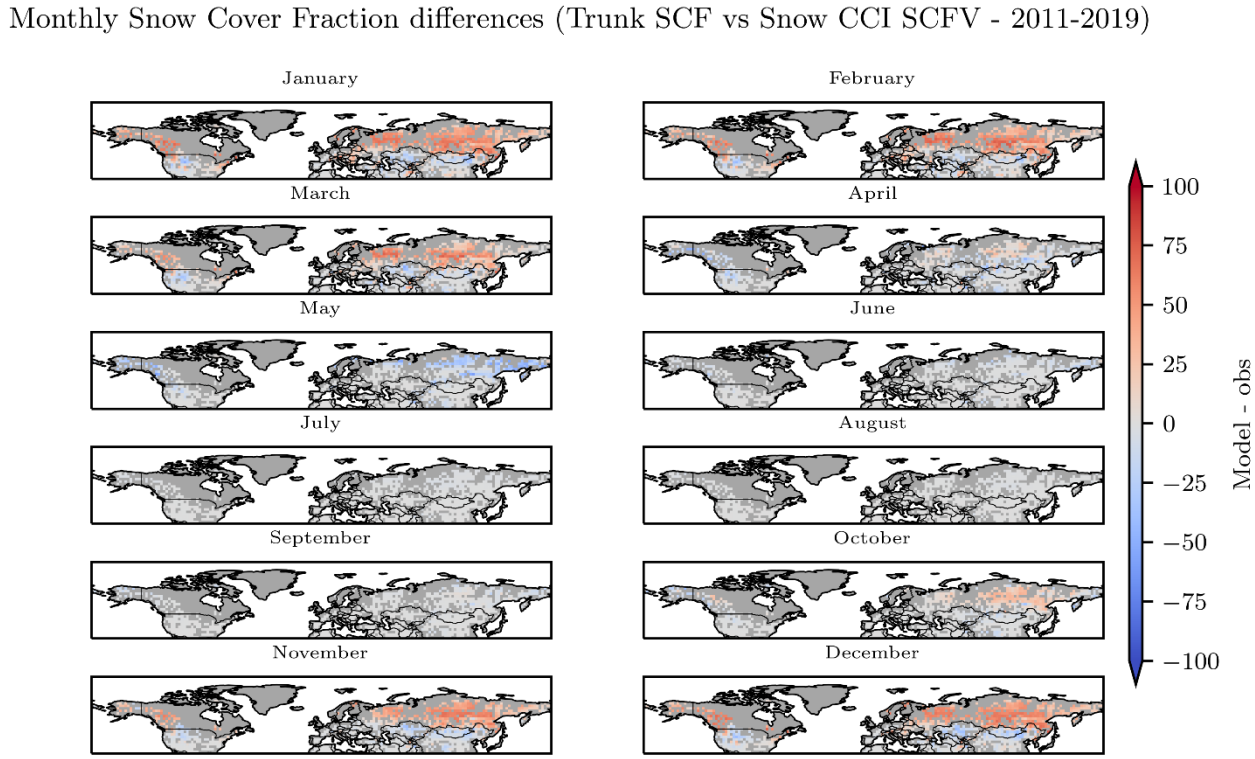


*Figure 13: SCFV monthly differences [%] between ORCHIDEE v4 and Snow CCI - 2011-2019. The regions in grey correspond to pixels where data were unavailable.*

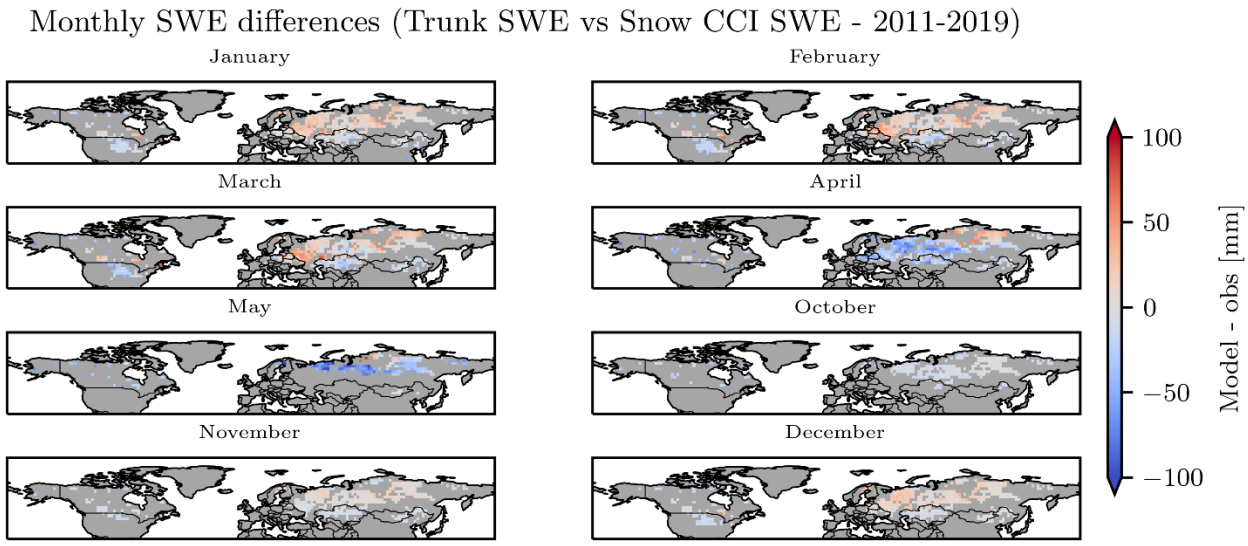


*Figure 14: SWE monthly differences [mm] between ORCHIDEE v4 and Snow CCI - 2011-2019. June, July, August and September are not shown since the SWE product generally does not provide data during this period. The regions in grey correspond to pixels where data were unavailable.*

## *6.3 Synthesis of the models and observations comparisons*

The ORCHIDEE v3 and Trunk versions exhibit distinct behaviours in albedo modelling. In v3, background and vegetation albedo appear to be well-calibrated overall, with albedo errors concentrated during winter. However, the SCF shows weak correlation with albedo, while the SWE representation remains relatively consistent with the SCF. In contrast, Trunk follows a dipole pattern: it strongly underestimates albedo in snowy conditions, particularly in high-latitude regions, while overestimating it in non-snow-covered areas and in southern regions. This suggests a strong constraint linked to vegetation albedo.

SCF is slightly less overestimated in Trunk compared to v3, yet the SWE and SCF patterns remain quite similar between the two versions. The differences between v3 and Trunk indicate that the new albedo scheme present in the Trunk and not in v3 impacts the overall albedo computation at all seasons, and that both the vegetation and snow albedo dynamics should be analysed and optimised.

Optimising albedo parameters in both versions is necessary to address these disparities, ensuring improved consistency in albedo, SCF, and SWE modelling across different regions and seasons.

# 7. ORCHIDEE model developments

In this study, the focus was on the use of the Snow CCI and the MODIS albedo products in order to improve the albedo simulation in ORCHIDEE. A stepwise approach is used to optimise the albedo parameterisation in ORCHIDEE v3. The albedo is therefore optimised via three steps:

1) Tuning of the global vegetation albedo parameters
2) Tuning the snow albedo parameters from equations (1) and (2) ($A_{aged}$, $B_{dec}$ and $\tau_{dec}$)
3) Tuning the snow cover fraction parameters as it plays a huge role in the albedo calculation

## *7.1 Global vegetation albedo optimisation*

Step 1 was performed on ORCHIDEE v2 (which has the same albedo scheme as v3), and the MODIS albedo was used for optimisation and validation. The optimisation was performed on the period 2011-2020 at a monthly frequency. The optimisation was made on data where the observed snow fraction and the calculated snow fraction were equal to 0%. This led to significant global RMSE reductions when comparing the model outputs to the MODIS albedo product (22.3% for the visible albedo, and 8.3% for the near-infrared albedo).

## *7.2 Snow albedo parameters optimisation*

With step 1 completed, the optimised parameters for albedo vegetation were used in step 2. For step 2, the ORCHIDEE, MODIS albedo and Snow CCI SCFV datasets were used, with data from the year 2015, which was rich in snow, at a daily frequency. Two optimisations Optim.1 and Optim.2 were performed, the only difference being the threshold used for the value of the snow cover fraction to be attained. Data for optimisation were selected from day-pixel datasets where all the following conditions were met:

- MODIS is valid (MODIS extracted with Albedo_Quality $\leq 5$)
- CCI SCFV > 50% for Optim.1 and CCI SCFV > 75% for Optim.2
- Continental fraction > 0.99
- Inland water fraction < 0.1

The optimisations were run with an analytical gradient calculation (Bastrikov et al., 2018), and independently for the NIR ($A_{aged\,NIR}$, $B_{dec\,NIR}$) and visible albedo parameters ($A_{aged\,VIS}$, $B_{dec\,VIS}$).
The parameters $A_{aged}$ and $B_{dec}$ were dependent on the PFT, but were regrouped into PFT types (one value for bare soil – PFT1, six for the temperate/boreal tree PFTs – 4 to 9, and one common parameter for the grasslands and the crops – 10 to 15). $\tau_{dec}$ was also optimised on the visible and NIR runs. This parameter is global and does not depend on the PFT.

The optimisation of the snow parameters led to a 20.5% RMSE reduction for the NIR albedo and a 23.6% RMSE reduction for the visible albedo for Optim.1, while for Optim.2, the RMSE reductions were respectively of 10.6% and 29.4% reduction.

Table 4 provides the maximum albedo values (corresponding to $A_{aged} + B_{dec}$ for NIR and for VIS), as well as the mean maximum albedo values (mean between the maximum values for NIR and VIS) that can be reached per PFT. Increases are written in red, and decreases are written in blue. The optimisations tend to increase the albedo values for most PFTs, which is consistent with the conclusions drawn when comparing the ORCHIDEE default outputs to the MODIS and CCI Snow observations. All the grassland PFTs have a maximum albedo value that is increased, while the results are more nuanced with the forest, and some decreases can be noted, especially for PFT5, 6 and 9. Also, Optim.2, which is performed with a threshold of 75% of SCFV gives higher albedo values, which is consistent since snow albedo is greater than vegetation albedo.

| | $A_{aged} + B_{dec}(NIR)$ | | | $A_{aged} + B_{dec}(VIS)$ | | | Max albedo (mean between VIS and NIR) | | |
|---|---|---|---|---|---|---|---|---|---|
| **PFT** | **Default** | **Opt. 1** | **Opt. 2** | **Default** | **Opt. 1** | **Opt. 2** | **Default** | **Opt. 1** | **Opt. 2** |
| **1** | 0,63 | 0,631 | 0,634 | 0,95 | 0,98 | 0,98 | 0,79 | 0,806 | 0,807 |
| **4** | 0,2 | 0,2208 | 0,2831 | 0,22 | 0,14 | 0,25 | 0,21 | 0,180 | 0,267 |
| **5** | 0,47 | 0,2802 | 0,1541 | 0,32 | 0,209 | 0,338 | 0,395 | 0,245 | 0,246 |
| **6** | 0,24 | 0,233 | 0,2597 | 0,24 | 0,167 | 0,213 | 0,24 | 0,2 | 0,236 |
| **7** | 0,2 | 0,231 | 0,281 | 0,23 | 0,22 | 0,323 | 0,215 | 0,226 | 0,302 |
| **8** | 0,24 | 0,3959 | 0,464 | 0,24 | 0,501 | 0,616 | 0,24 | 0,448 | 0,54 |
| **9** | 0,35 | 0,18 | 0,245 | 0,42 | 0,196 | 0,325 | 0,385 | 0,188 | 0,285 |
| **10 - 15** | 0,56 | 0,577 | 0,59 | 0,72 | 0,847 | 0,872 | 0,64 | 0,712 | 0,731 |

*Table 4: Maximum albedo values for NIR, VIS and global albedo before optimisation (default), and after with a threshold of SCFV above 50% (Optim.1) and above 75% (Optim.2)*

## 7.3 Snow cover fraction optimisation

In order to have a correct estimation of the albedo in ORCHIDEE, it is also necessary to correctly assess the snow cover fraction. Indeed, the global albedo strongly depends on the SCF, as shown in equations (6) and (7), where $\alpha$ is the mean albedo, $\alpha_{snow}$ is the snow albedo, $\alpha_{snow-free}$ is the albedo of the snow-free part of the pixel (mostly vegetation and bare soil), and each pixel has a given fraction of each PFT $i$ $f_{PFT,i}$ with a PFT dependent snow albedo $\alpha_{snow}^{PFT,i}$.

$$\alpha = SCF \times \alpha_{snow} + (1 - SCF) \times \alpha_{snow-free} \quad (6)$$

$$\alpha_{snow} = \sum_{PFT} f_{PFT,i} \times \alpha_{snow}^{PFT,i} \quad (7)$$

The parameters $\alpha_{snow-free}$ and $\alpha_{snow}^{PFT,i}$ have been optimised in the two previous steps. The only variable that needs to be refined is now the SCF. The calculation of the SCF provided in equation (1) may not necessarily be applicable to all the types of vegetation and should be

adapted accordingly. The term "$2.5 \times z_{0_{ground}}$" at the denominator might not be well adapted in the presence of trees, since its value can quickly reach 100% when the snow depth increases. Indeed, $z_{0_{ground}}$ ranges from 0.001 to 0.01 m in the model, meaning that the snow cover fraction converges to 100% even when the snow depth is relatively low. To solve this issue, a parameter "div_fsnow" is created to replace the 2.5 term, and some tests are done to optimise this parameter in the model, while using the optimised vegetation and snow albedo parameters.

However, this part is not as straightforward as it seems. The SCF used in ORCHIDEE defined here should relate to the SCFV, i.e. the Snow Cover Fraction viewed from the sky. However, since the snow albedo is already optimised per type of vegetation, and the SCF used during the optimisation of these parameters was not optimised, the interpretation of this variable should be more nuanced. As it was previously said, during the optimisation, the SCF was calculated with equation (1) and tended to provide a SCF close to 100% even with low snow depths. It means that the SCF calculated by ORCHIDEE might be overestimated and closer to the SCFG values.

Preliminary site optimisation tests have been performed for some PFTs. The methodology to select the sites is depicted in section 4.4. The optimisation was performed with the albedo parameters being initialised to the values obtained with Optim.1 from the snow albedo parameters optimisation. The results from this optimisation were chosen instead of Optim.2 because the number of used data points was larger, and the averaged RMSE reductions between NIR and VIS for the two optimisations were similar. Here, the optimisation tests were performed independently for each PFT, only modifying the parameter value of the dominant PFT for the considered sites. However, for further developments, we should consider optimising first the grassland and bare soil PFTs (1 and 10-15) since these PFTs can be found alone in their pixel (representing more than 90% of the pixel), while forest PFTs are always associated with grassland PFTs in their pixel, meaning that the other PFT can affect the results.

Some optimisation results for PFT8, PFT9 and PFT15 are provided in Table 5. Only the parameter "div_fsnow" is optimised, and only for the considered PFT, although some sites can contain a mix of different PFTs, especially when forest PFTs are studied, since they are always associated with some grassland PFTs.

For grassland and bare soil PFTs, the SCF generally reaches quite easily 100% as soon as there is snowfall. The parameter "div_fsnow" should therefore be quite low, and the initial parameter used for the different sites that were tested generally tends towards the lowest possible value defined (0.5 in the presented test). The optimised albedo parameters provide very good estimations on the studied sites. Also, all the sites that were analysed for PFT15 tend to quickly reach a SCF of 100%. The selection algorithm should be refined to select mostly sites with intermediate SCF values to be able to correctly parameterise the value of "div_fsnow", to be more precise in its estimation.

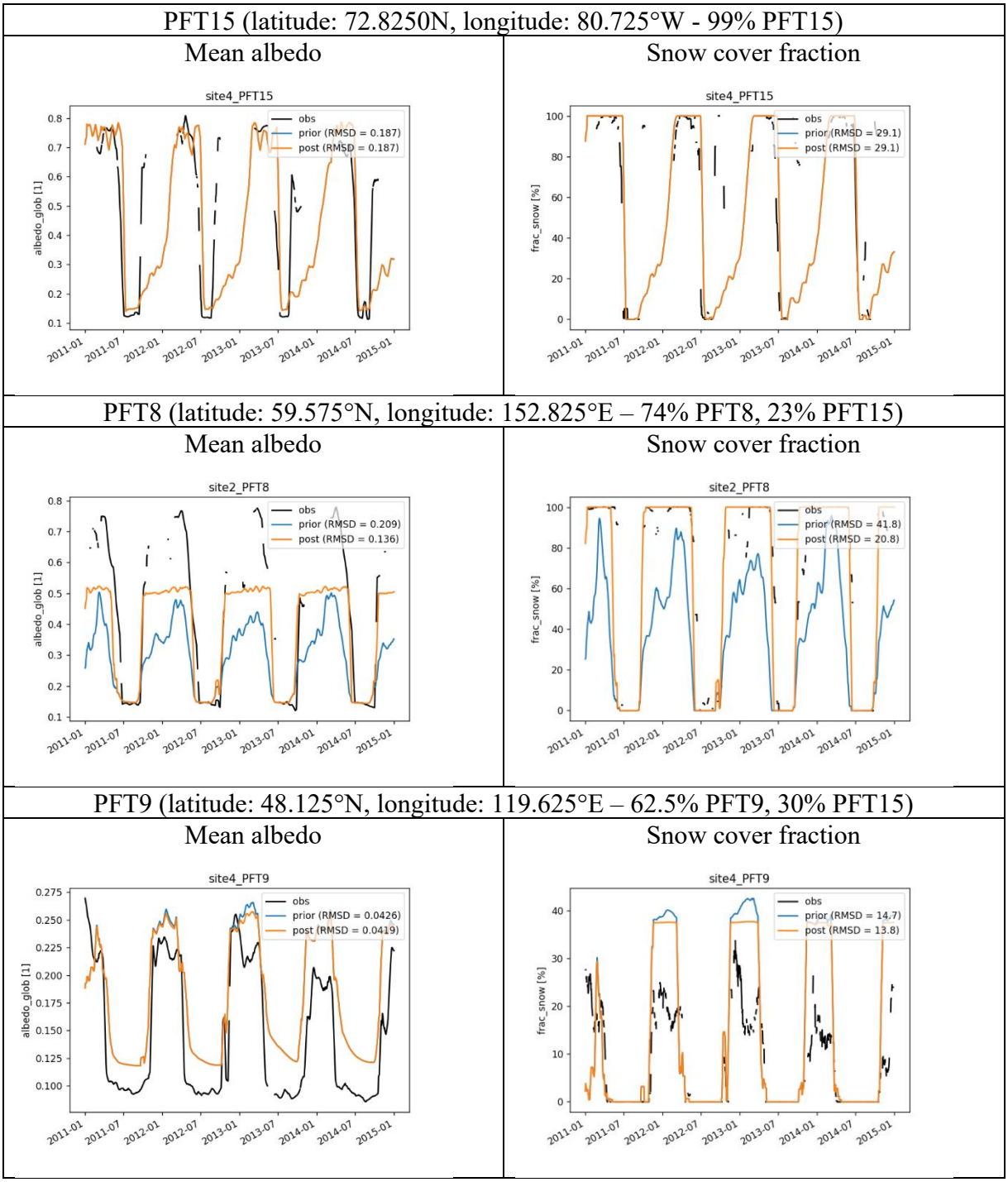


*Table 5: Optimisation of albedo and SCF for PFT15, PFT8 and PFT9 (mean albedo on the left column and snow cover fraction on the right column). The observations appear in black, the simulation results without the SCF parameter optimisation are in blue (prior) and the optimised simulation results are in orange (post)*

For PFT8, we can notice that the albedo tends to be underestimated by the model in the prior (blue), while the observed SCF tends to reach 100%. Here, the SCF of the model needs to be

refined to enable the model to show a 100% SCF during winter, which is not the case in the prior, with an initial value of "div_fsnow" = 25. The optimisation enables the adjustment of the SCF and better fits the observations by diminishing the value of "div_fsnow". However, the albedo computed by the model is unable to reach the observed albedo values, meaning that the snow albedo parameters that were optimised in the previous step are too low. Indeed, they were determined with a threshold of SCFV at 50%, and maybe a higher threshold should be used. For the next optimisations, we suggest using the results from Optim.2 for albedo parameters instead.

Finally, for PFT9, the albedo values tend to be quite consistent. However, the SCF calculated for this PFT is generally too high, meaning that the parameter "div_fsnow" needs to be increased. Here, it enables a decrease of the SCF value, and to improve the albedo as well. However, it is difficult to further diminish the SCF close to the observations, probably because of the presence of PFT15 which represents around 30% of the pixel.

However, although the results can still be refined, it seems that the use of the new snow and vegetation albedo parameters enables the model to better simulate the global albedo, and the optimisation of the parameter "div_fsnow" used to calculate the SCF can provide even better results. The evolution of the simulated albedo and SCF tend to be in phase with the observations, meaning that the previous optimisations have also enabled to limit the effects observed during the preliminary analysis, where snow tended to melt too rapidly during Spring. Based on those results, the methodology for optimisation could be updated to take these effects into account, as well as the snow/vegetation interactions that play a large role in these processes. Moreover, this analysis has shown the need to use optimised snow albedo parameters from optimisations with a high SCF threshold, in order to have sufficiently high albedo values computed by the model in the presence of snow.

# 8. Summary

The work performed during this first year allowed us to progress on the main tasks of the project which are:
- the analysis of the CCI-Snow products (SCFV, SCFG and SWE) and their preliminary comparison with the ORCHIDEE V3.0 and Trunk versions,
- the assessment of their potential to improve the modelling of the snow cover fraction and the calibration of the albedo related parameters,
- the optimization of the snow albedo parameters (in ORCHIDEE v3).

A protocol has been set up to select the best areas to perform site optimization, with an algorithm created to select the most homogeneous pixels in terms of vegetation for SCFV, SCFG, SWE and albedo observation data. The analysis of the vegetation evolution was done with the new PFT maps developed by Harper et al., 2023. In ORCHIDEE v3, the vegetation albedo parameters were recalibrated using these maps and the albedo products, and the snow albedo parameters, PFT-dependent, were also recalibrated thanks to the combination of the CCI Snow products, the MODIS albedo product and the PFT maps. This methodology has to be applied now to the current trunk version. All this framework is now set up to use the CCI Snow products and assess their contribution for improving snowpack dynamics modelling in ORCHIDEE and study atmospheric interactions in a second step.

*Special thanks*

Two internships have been dedicated to the tasks of WP5.6. Some of the works presented here are direct results of these internships (Guillermo Cossio, summer 2023, and Benoît Lecomte, fall 2023, internal reports). Moreover, this project also benefited from the work of our LSCE colleagues (Vladislav Bastrikov and Luis Olivera) on the albedo optimisation of ORCHIDEE.

# 9. References

Bastrikov, V., Macbean, N., Bacour, C., Santaren, D., Kuppel, S. and Peylin, P.: Land surface model parameter optimisation using in situ flux data: Comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1.9.5.2), *Geosci. Model Dev., 11(12),* 4739–4754, doi:10.5194/gmd-11-4739-2018, 2018.

Boone, A., and Etchevers, P.: An intercomparison of three snow schemes of varying complexity coupled to the same land surface model: Local-scale evaluation at an Alpine site. *Journal of Hydrometeorology, 2(4),* 374-394, 2001.

Chalita, S. and Le Treut, H.: The albedo of temperate and boreal forest and the Northern Hemisphere climate: a sensitivity experiment using the LMD GCM, *Climate Dynamics, 10*, 231-240, doi: 10.1007/BF00208990, 1994.

Charbit, S., Dumas, C., Maignan, F., Ottlé, C., Raoult, N., Fettweis, X., & Conesa, P.: Modelling snowpack on ice surfaces with the ORCHIDEE land surface model: application to the Greenland ice sheet. *The Cryosphere*, *18*(11), 5067-5099, 2024.

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., et al.: Improved near-surface continental climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics, *Journal of Advances in Modeling Earth Systems, 12*, e2019MS002005., doi: 10.1029/2019MS002005, 2020.

Cuynet, A., Salmon, E., Lopéz-Blanco, E., Goeckede, M., Ikawa, H., Kobayashi, H., ... & Ottlé, C.: Enhanced prescription of soil organic and mineral content in the ORCHIDEE LSM to better simulate soil temperatures: application at nine high-latitude GEM and FLUXNET sites. *In review for Journal of Geophysical Research: Biogeosciences,* 2025.Dantec-Nédélec, S., Ottlé, C., Wang, T., Guglielmo, F., Maignan, F., Delbart, N., ... & Jouzel, J.: Testing the capability of ORCHIDEE land surface model to simulate Arctic ecosystems: Sensitivity analysis and site-level model calibration. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1212-1230, 2017.

Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P. and Morin., S.: Impacts of snow and organic soils parameterization on northern Eurasian soil temperature profiles simulated by the ISBA land surface model, *The Cryosphere, 10*, 853-877, doi: 10.5194/tc-10-853-2016, 2016.

Forbes, T., Rossiter, D., & Van Wambeke, A.: *Guidelines for evaluating the adequacy of soil resource inventories* (Vol. 4). Cornell University, Department of Agronomy, 1987.

Gaillard, R., Peylin, P., Cadule, P., Bastrikov, V., Cheruy, F., Cuynet, A., Ghattas, J., Zhu, D., Guenet, B.: Arctic soil carbon insulation averts large spring cooling from surface–atmosphere feedbacks. *Proceedings of the National Academy of Sciences*, *122*(3), e2410226122, 2025.

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., ... & Ciais, P.: ORCHIDEE-MICT (v8. 4.1), a land surface model for the high latitudes: model description and validation. *Geoscientific Model Development*, *11*(1), 121-163, 2018.

Harper, K. L., Lamarche, C., Hartley, A., Peylin, P., Ottlé, C., Bastrikov, V., ... & Defourny, P.: A 29-year time series of annual 300 m resolution plant-functional-type maps for climate models. *Earth System Science Data*, *15*(3), 1465-1499, 2023.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J., Fairhead, L., Filiberti, M., Friedlingstein, P., Grandpeix, J., Krinner, G., Le Van, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrised physics with emphasis on tropical convection, *Climate Dynamics, 27*, 787–813, doi:10.1007/s00382-006-0158-0, 2006.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochemical Cycles, 19*, GB1015, doi:10.1029/2003GB002199, 2005.

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., ... & Venäläinen, P.: GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset. *Scientific Data*, *8*(1), 163, 2021.

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., ... & Boucher, O.: Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2019MS001940, 2020.

Niu, G.-Y., and Yang, Z.-L.: An observation-based formulation of snow cover fraction and its evaluation over large North American river basins, *Journal of Geophysical Research, 112*, D21101, doi:10.1029/2007JD008674, 2007.

Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B., Kempen, B., Ribeiro, E., & Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil, 7(1),* 217-240, https://doi.org/10.5194/soil-7-217-2021, 2021.

Raoult, N., Charbit, S., Dumas, C., Maignan, F., Ottlé, C., and Bastrikov, V.: Improving modelled albedo over the Greenland ice sheet through parameter optimisation and MODIS snow albedo retrievals, *The Cryosphere, 17*, 2705–2724, https://doi.org/10.5194/tc-17-2705-2023, 2023.

Solberg, R., G. Schwaizer, T. Nagler, S. Wunderle, K. Naegeli, K. Luojus, M. Takala, J. Pulliainen, J. Lemmetyinen, and M. Moisander: ESA CCI+ Snow ECV: Product User Guide, version 3.1, December 2021, 2021.

Takala,M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J.-P., Koskinen, J. and Bojkov, B.: Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, *Remote Sensing of Environment, Volume 115, Issue 12*, 15 December 2011, Pages 3517-3529. doi: 10.1016/j.rse.2011.08.014, 2011.

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V. and Peylin, P.: Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production. *Geoscientific Model Development, 12(11)*, 4751-4779, doi:10.5194/gmd-12-4751-2019, 2019.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S. and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, *Journal of Geophysical Research Atmosphere, 118*, 6064–6079, doi:10.1002/jgrd.50395, 2013.

Wang, T., Peng, S., Krinner, G., Ryder, J., Li, Y., Dantec-Nédélec, S. and Ottlé, C.: Impacts of Satellite-Based Snow Albedo Assimilation on Offline and Coupled Land Surface Model Simulations. *PLoS ONE 10(9):* e0137275, doi:10.1371/journal.pone.0137275, 2015.

# 10. Glossary

| **Terms** | |
|---|---|
| Data assimilation | Observations directly influence the model initial state considering their error characteristics during every cycle of a model. This is used for reanalysis, NWP, which includes seasonal and decadal forecasting. |
| Model validation | Observations are compared with equivalent model fields to assess the accuracy of the model. This can be on short time scales for process studies or long-time scales for climate trends. |
| Climate monitoring | This describes the use of a satellite only dataset to monitor a particular atmospheric or surface variable over a period > 15yrs to investigate whether there is a trend due to climate change. |
| Initialisation | To initialise prognostic quantities of the model with reasonable values at the beginning of the simulation but do not continuously update. |
| Prescribe boundary conditions | Prescribe boundary conditions for a model run for variables that are not prognostic (e.g. land cover, ice caps etc). |
| Accuracy | Accuracy is the measure of the non-random, systematic error, or bias, that defines the offset between the measured value and the true value that constitutes the SI absolute standard. |
| Stability | Stability is a term often invoked with respect to long-term records when no absolute standard is available to quantitatively establish the systematic error – the bias defining the time-dependent (or instrument-dependent) difference between the observed quantity and the true value. |
| Precision | Precision is the measure of reproducibility or repeatability of the measurement without reference to an international standard so that precision is a measure of the random and not the systematic error. Suitable averaging of the random error can improve the precision of the measurement but does not establish the systematic error of the observation. |
| **Acronyms** | |
| AVHRR | Advanced Very High-Resolution Radiometer |
| CCI | Climate Change Initiative |
| CMC | Climate Modelling Community |
| CMIP6/7 | Climate Model Intercomparison Project-6/7 |
| CMUG | Climate Modelling Users Group |
| ECV | Essential Climate Variable |
| EGU | European Geophysical Union |
| ENSO | El Nino- Southern Oscillation |
| ERA | ECMWF Reanalysis |
| IPCC | International Panel for Climate Change |
| LAI | Leaf Area Index |
| SCF | Snow Cover Fraction |
| SCFG | Snow Cover Fraction Ground |
| SCFV | Snow Cover Fraction Viewable |
| SWE | Snow Water Equivalent |