# Preparing GSWP3 V1 Official forcing to be read by ORCHIDEE

Salma Tafasca, Agnès Ducharne, April 2018

## 1- Formatting the GSWP3 V1 forcing

GSWP3 V1 is the official v1 release of GSWP3 forcing [Kim et al. in prep]. In this document, we describe how we formatted the GSWP3 V1 forcing in order to be read by ORCHIDEE.

The following forcing notations will be used in this document:
**F0** : GSWP3 v0 forcing formatted by Jinfeng Chang for ISIMIP2 simulations.
**F1** : GSWP3 v1 official forcing after being formatted to be used by ORCHIDEE.
**D1**: dim2driver interpretation of **F1** forcing.

**F1** forcing has been prepared based on the official V1 release of GSWP3 at the global 0.5° resolution. It covers the period 1901-2010 and the files are stored in the shared repositories of ORCHIDEE. The script used to prepare the forcing is located in the same directory, it was adapted from Jinfeng Chang's script for preparing **F0**, based on the reflexions of Agnès Ducharne and Nicolas Vuichard.

The Rscript creates netcdf-3 forcing files, which were converted to netcdf-4 using the command *ncks -4 GSWP3_halfdeg_3h_YYYY.nc GSWP3_v1_halfdeg_3h_YYYY.nc.* Except for the first 19 years (from 1901 to 1919), which were created originally as netcdf-4 files.

### 1-1-Time axis

**F1** forcing is prepared to be read by the old driver (dim2driver), which assumes that scalar variables (Tair, Qair, Psurf and Wind) are instantaneous, while flux variables (Rainf, Snowf, SWdown, LWdown) are averaged over the previous time interval. Furthermore, dim2driver considers that the first value of the year corresponds to the time instant: 00:00 UTC+dt_forcing, regardless of the time axis. Since the original GSWP3 forcing provided by Kim (2017) starts every year at 00:00 UTC, the time axis was shifted in **F1** forcing, for all the forcing variables, to respect the assumption of dim2driver. Figure 1 shows the transformation in time axis of GSWP3 v1 to obtain F1. For every yearly file, we discard the first record which corresponds to 00:00, and we add the first record of the next year, which corresponds to 24:00. For the last year file, since we don't have the record of 24:00, the last record (21:00) is repeated twice.

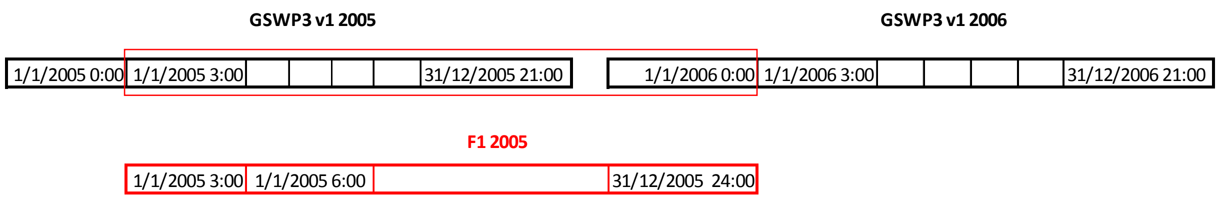



*Fig.1 Time axis transformation from F0 to F1*

In order for the F1 forcing to be read by the new driver, an attribute was added to all variables explaining that the values represent backward averaging, which means that a record at 00:00 is a mean of the time interval [-3:00, 00:00] (cf. mail of Hyungjun Kim on 17/01/2018) .The attribute is the following:

*cell_method time : mean (end)*

This attribute determines if the variable is an instantaneous value or an average over a time step. If it is an average, then the term between () determines if the specified date corresponds to the start, the center or the end of the interval over which the variable is averaged [Jan Polcher, Content of ORCHIDEE forcing files, May 2014] (http://forge.ipsl.jussieu.fr/orchidee/raw-attachment/wiki/Documentation/Forcings/Description_Forcing_Files.pdf).

Note that the old driver (dim2driver) does not consider attributes, it assumes that all scalar variables are time : instantaneous, instead of time : mean (end), which leads to a small error on the scalars. Figure 2 illustrates how dim2driver interprets F1 scalar variables (dashed blue line), while the correct interpretation is represented by the dashed red line, since the values provided by GSWP3 v1 are averaged values over the previous time step.

According to Jan Polcher, when cell method = time: mean (end), the new driver positions the scalar values at the center of the interval, leading to a more correct interpolation.

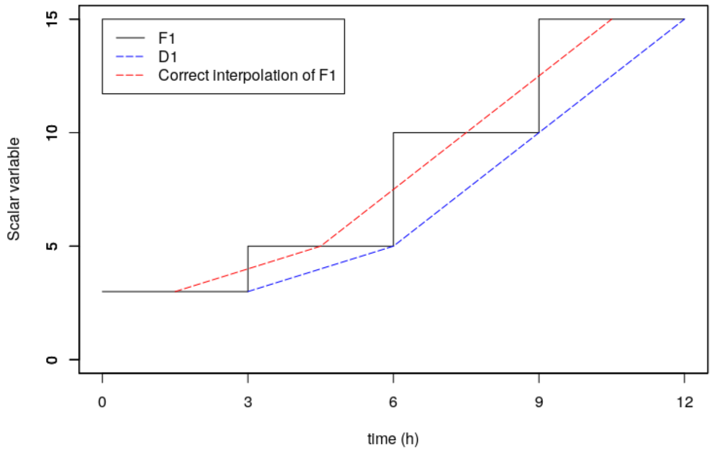


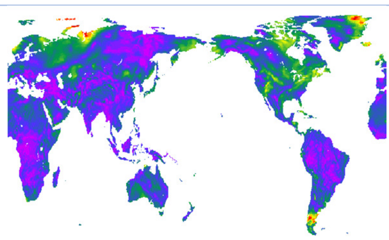
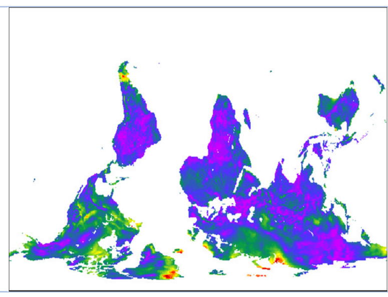

*Fig.2 dim2driver interpolation of F1 forcing (dashed blue) VS correct interpolation (dashed red)*

### 1-2-Mask

In the original GSWP3 V1 official forcing, data covers the whole globe including oceans except for the variable Wind, where data covers only land surface points. Wind data was therefore used to mask all variables. The formatted variables are kept 2D.

Table 1 summarizes the main differences between the original GSWP3 v1 forcing provided by Kim (2017), and the processed GSWP3 (referred to as F1). Note that GSWP3 v1 has leap years.

*Table 1 Differences between original GSWP3 v1 official forcing and formatted one (F1)*

| | Original GSWP3 v1 (Kim, 2017) | Processed GSWP3 (F1) |
|---|---|---|
| Organizing files | One file for every year for every variable | One file for every year containing all variables |
| Masking | Only Wind is masked | All variables are masked |
| Time axis | Year starts at January 1st 00:00 UTC and ends at December 31st 21:00 UTC | Year starts at January 1st 03:00 UTC and ends at December 31st 24:00 UTC |
| Map orientation |  |  |

## 2- Testing the F1 forcing

A test has been performed with the trunk revision of ORCHIDEE using F1 forcing, to make sure it is read correctly by the old driver of ORCHIDEE. Description of the protocol of the test and results are described in section 2.1 and 2.2, respectively.

### 2.1 Protocol of the test

#### 2-1-1 ORCHIDEE simulation

ORCHIDEE was run using F1 forcing, under the following conditions:

- Period of simulation: 5 days, from 01/01/2005 to 05/01/2005
- Ouput frequency: highest possible frequency, ie, 30 minutes
- LWdown conserved. By default, ORCHIDEE performs an interpolation of LWdown between two time instants, which does not conserve the amount of energy over a time step interval. This configuration can be changed by adding in run.def file: LWDOWN_CONS = y.

#### 2-1-2 The diurnal cycle

All forcing variables are plotted in section 2.2 on 01/01/2005 over the point 10°N and 0°E (north Togo), except for rainfall and snowfall variables which were examined at north Australia (13.5°S, 130°E) and north India (31°N, 80°E), respectively, where rainfall and snowfall values are important. For each variable, we plot GSWP3 V1 forcing curve (in purple), F1 forcing curve (in green), which both vary at a 3h time step, and D1, which is the interpretation of F1 by dim2driver at a 30min time step (in red). In order to show if dim2driver conserves the values while interpolating over a forcing time-step, an additional curve (red dashed line) is plotted, representing the average of D1 over a 3h time-step (= forcing time-step).

### 2.2 Results

#### 2-2-1- The variables Tair, Qair and Psurf

dim2driver linearly interpolates the scalar variables Tair, Qair and Psurf, which are provided at 2m height.

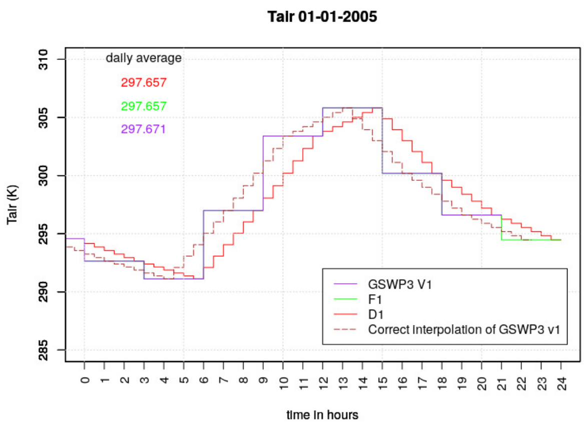



*Fig.3 Air temperature on landpoint (x=0°, y=10°N)*

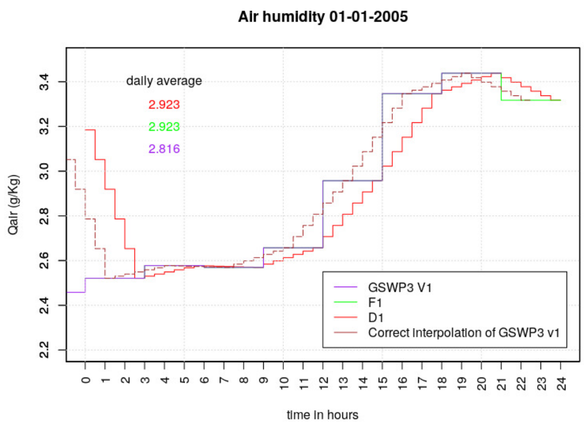



*Fig.4 Air humidity on landpoint (x=0°, y=10°N)*

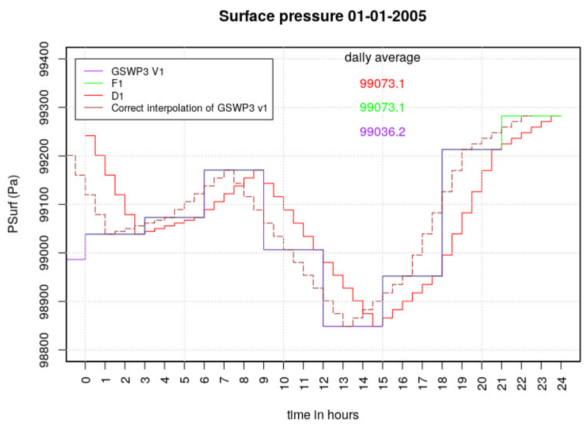



*Fig.5 Surface pressure on landpoint (x=0°, y=10°N).*

To interpolate the scalar variables over the first time step [0:00, 3:00] of the first year, and since no value is provided for the time instant 0:00, dim2driver uses the record of the last time step of the day (24:00) in order to complete the interpolation. This explains why the red curve does not follow the purple curve in the first forcing time step.

At a time instant t0, GSWP3v1 provides mean records over the time interval preceding the time instant t0. dim2driver treats the scalar records as if they were instantaneous, which leads to an error (red curve VS brown dashed curve in fig.3) (see section 1.1).

### 2-2-2- The variables SWdown and LWdown

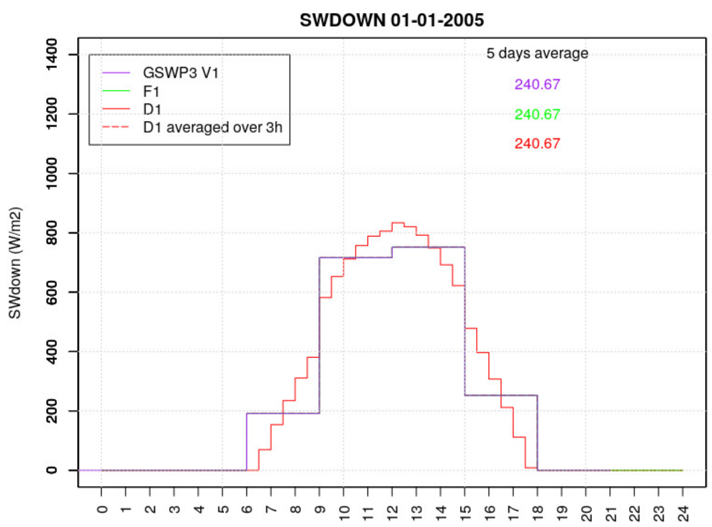



*Fig.6 Downward short wave radiation. (x=0°, y=10°N).*

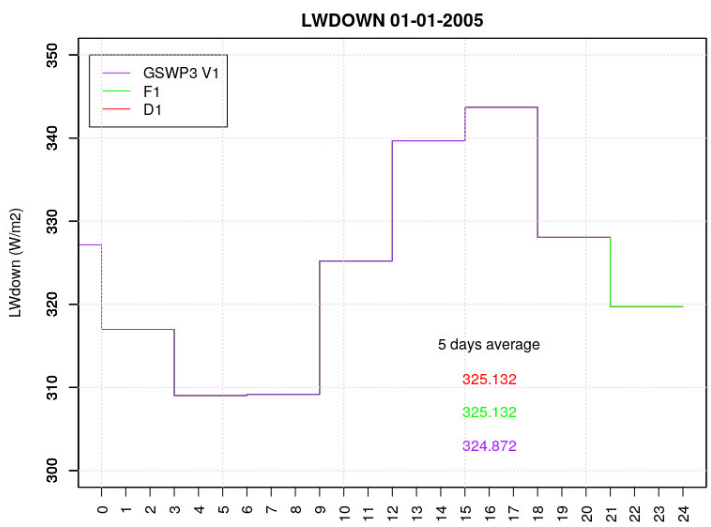



*Fig.7 Downward long wave radiation. (x=0°, y=10°N)*

dim2driver respects two conditions while disaggregating SWdown variables:

1- The provided average flux should be preserved during a time step interval.

2- SWdown should be distributed proportional to the zenith angle.

In the absence of noise induced by clouds, the diurnal cycle of SWdown should be relatively smooth and well centered at noon. Figure 6 shows that the downward solar radiation for D1 is consistent with the zenith angle (presence of the mid-day peak) and that SWdown is conserved, over the land point 10°N, 0°E. However, a global analysis points out a small energy conservation violation over some particular regions of the globe (Figure 8).

Figure 8a shows that the 5-day average value of SWdown in F1 differs from the one in D1, in some regions of the globe. The shape of these zones reminds the shape of a SWdown plot at the sunset zones (Fig 8b and c).

For a better understanding of what is happening in these zones, we selectedthe landpoints in which SWdown is not conserved (41°W, 14°S) for which we analysed SWdown curves (Fig. 9). F1 and D1 forcings seem to be overlapped, but a visualization of the difference (Fig. 9b) reveals a tiny difference of around $1W/m^2$ occurring at the last forcing time step of the day. This difference is barely visible so we did not work on a correction in dim2driver to fix it.

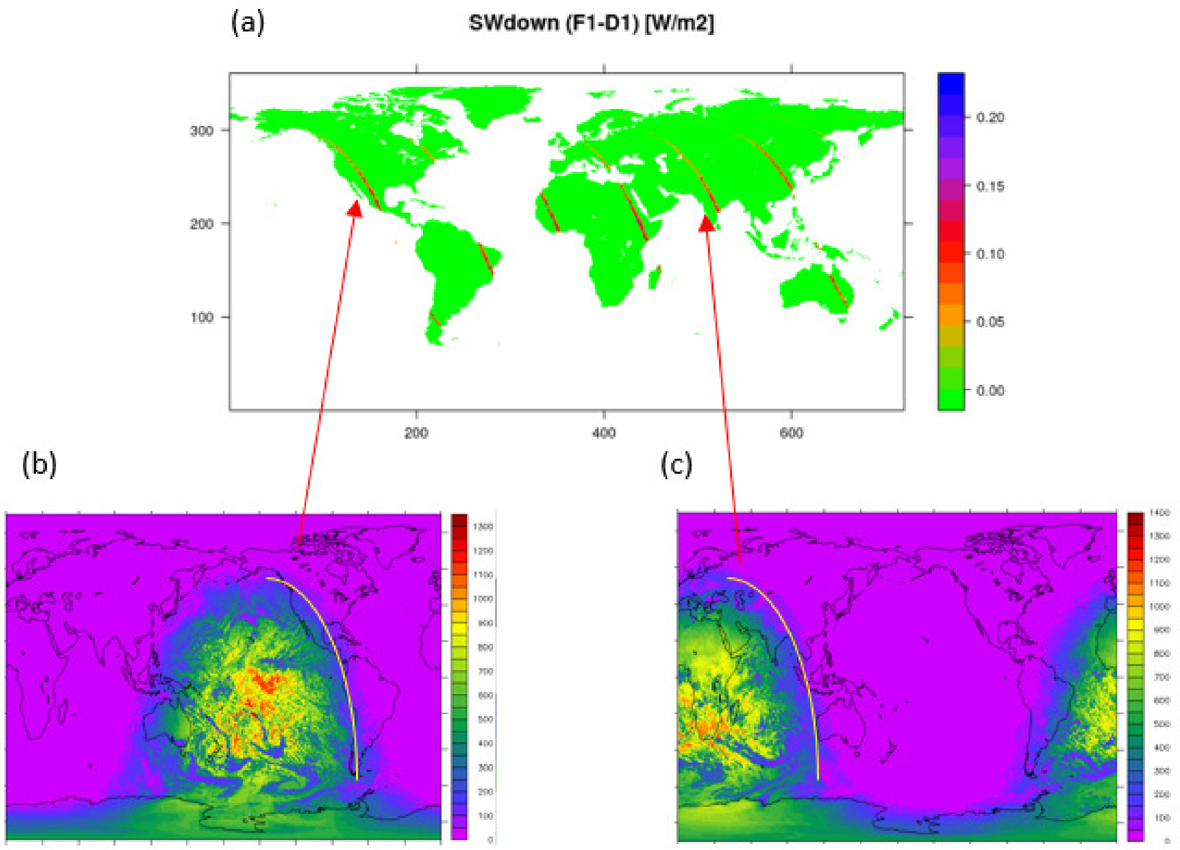



*Figure 8 (a) Difference in averaged SWdown over 5 days, between the forcing F1 and D1 (1st to 5th January 2005). (b) Global SWDOWN plot on 01-01-2005 at 00:00 and at (b) 12:00 UTC, from forcing F1.*

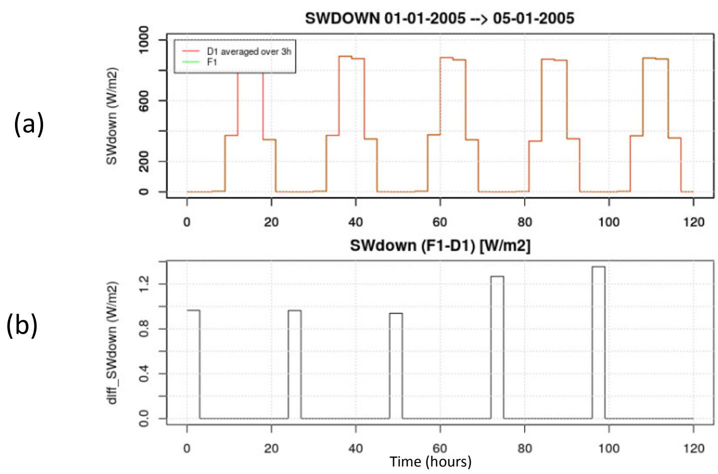



*Fig.9 Downward short wave radiation. Green and red curves are not completely overlapped, the difference between the two curves is plotted in fig (b). (x=41°W, y=14°S).*

By default, ORCHIDEE does a linear interpolation of LWdown between t and t+dt, in this case, the mean value of LWdown over the interval [t, t+dt] is (LW(t)+LW(t+dt))/2, which is different from the real mean value, which is LW(t+dt), so the amount of energy is not conserved. (for more details, refer to the discussion on the differences between the two drivers that can be found on: [https://forge.ipsl.jussieu.fr/orchidee/wiki/Branches/Driver_Improvements](https://forge.ipsl.jussieu.fr/orchidee/wiki/Branches/Driver_Improvements)).

In the present simulation, we added LW**down**_cons = y in the run.def file, in order to preserve the average flux provided by the forcing during the temporal disaggregation. It was checked that LW**down** is perfectly conserved over the whole globe.

### 2-2-3 The variable Wind:

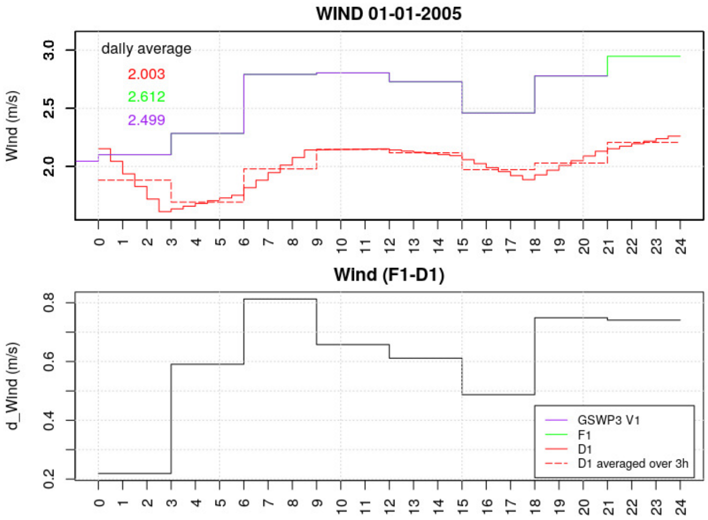



*Fig.10 Wind curve on land point (x=0°, y=10°N).*

According to Hyungjun Kim, GSWP3 relies on wind speed measured at heights above the surface of between 2 and 20 m. Measurement height varies both within and between countries, and in many cases, the heights were not specified, consequently, no corrections were made to wind data. The large majority of known heights were 10 m, and the interpolated wind field should be assumed to represent the wind speed at this height (New et al., 2002). This means we could assume GSWP3 winds as values at 10 m.

The wind is treated like a scalar, and both dim2driver and the new driver assume by default that it is measured at a height of 10m. Consequently, the drivers apply an interpolation to get wind values at a height of 2m needed for ORCHIDEE land surface variables. Figure 10 shows the difference between wind from forcing F1 (measured at 10 m) and the 2-m interpolated values (forcing D1).

### 2-2-4 The variables Rainf and Snowf:

**dt_force** : time-step of the forcing
**dt** : time-step of sechiba
**split** = int(dt_force/dt)

When dt_force is less than 3 hours, the precipitation is uniformly distributed over dt_force, otherwise, by default, the old and new drivers non-uniformly distribute precipitation at ORCHIDEE's time resolution, using the parameter nb_spread, which is the number of the first time steps of ORCHIDEE over which the precipitation is distributed, within one forcing time step. When nb_spread = split, the precipitation is uniformly distributed over the forcing time step. By default nb_spread=split/2. Ideally, to be closer to the real world, the value of *nb_spread* should be climate dependent and it is therefore not realistic to use one single value for the whole globe, as it is done in ORCHIDEE, due to lack of sufficient information (30-min precipitation data over the whole globe). Figures 11 and 12 show the effect of nb_spread on the high frequency distribution of Rainf and Snowf respectively, with good conservation of the input values (the small differences between the simulations and GSWP3 come from the initial time step).

To assess the effects of the *nb_spread* value over the water balance variables and 1-st layer soil temperature, we ran twice the same 5-day simulation with two different configurations. The first one uses the default value of *nb_spread* while the second one uses a uniform distribution of precipitation, ie, *nb_spread = split*. These simulations were run regionally over the domain between 120 and 150°E and 0 and 15°S (north Australia, Indonesia, New Guinea, Fig. 13).

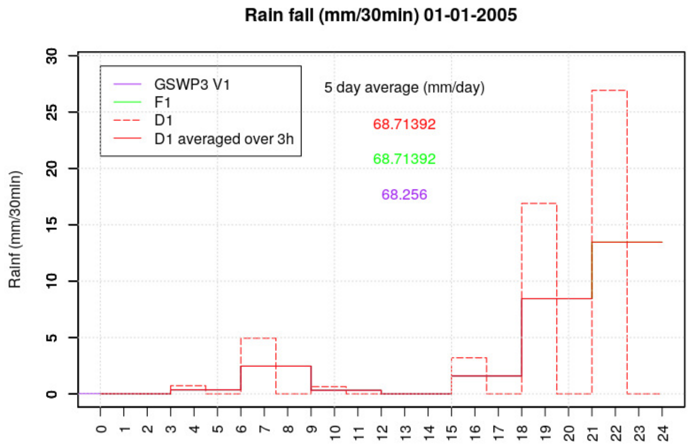



*Fig.11 Rain fall curve on land point (x=130° east , y=13.5°S) (north Australia)*

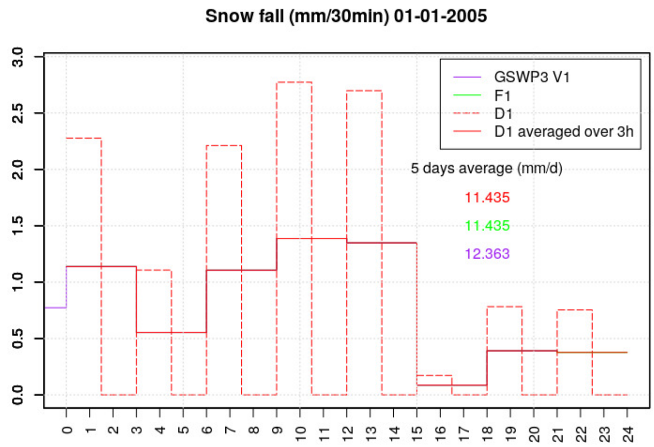



*Fig.12 Snow fall curve on land point (x=80° E , y=31°N). (north India)*

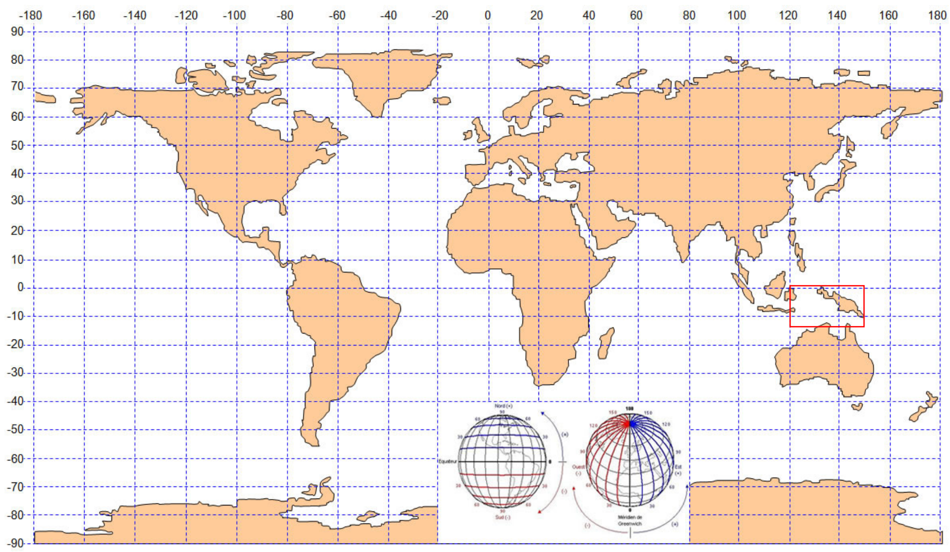


*Fig.13 Simulated region* $X \in (120°, 150°), Y \in [-15°, 0°]$

Figure 14 shows the time series of ORCHIDEE outputs at a half-hourly time step, represented at point (x=130°E, y=13.5°S → north of Australia) for the last three days of the simulations. The first two days were discarded to avoid misinterpreting effects caused by initialization from restart. For each variable, a 3-day regional average is calculated.

The high frequency variation of surface soil moisture (Fig. 14c) comes primarily from the rain(Fig. 14a), while the drainage variations do not seem to be affected by the high-frequency precipitation variation(Fig. 14e). Mean drainage, however, is reduced in case of non-uniform precipitation which leads to increased surface runoff (Fig. 14g) thus reduced surface soil moisture. In fact, soil permeability is limiting the infiltration, thus intense precipitation leads to more infiltration-excess runoff. As a result, the dependency of surface runoff on rainfall rate is highly non-linear, and a uniform distribution of precipitation reduces runoff considerably over the simulated region (from 1.845 mm/d to 1.165 mm/d, with a decrease of the ratio runoff/drainage from 44 to 28%). Note that this may not be true in regions with flat areas, where we can have reinfiltration during the second half of dt_force, during which no rainfall is simulated, and therefore, increase drainage.

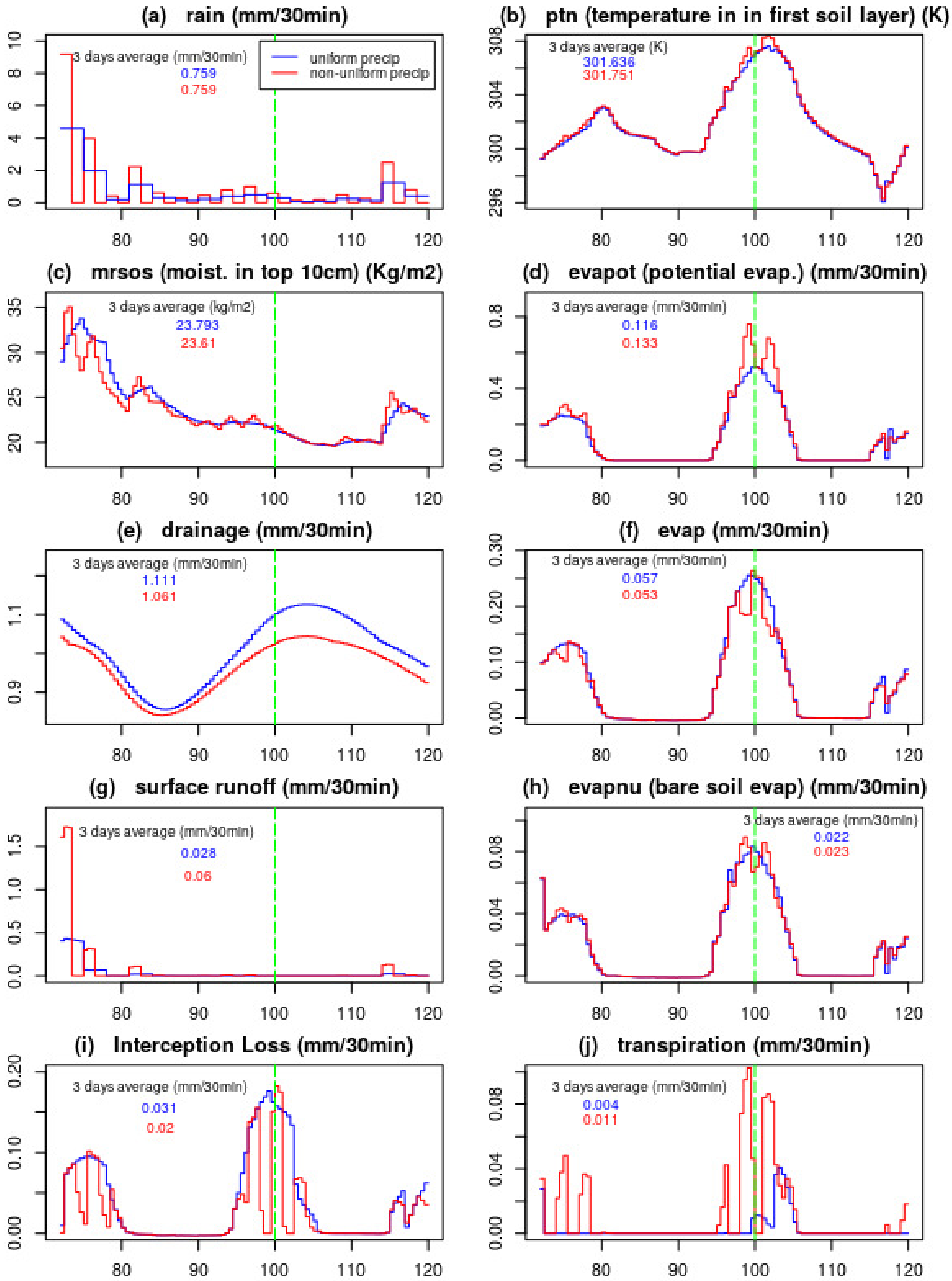



*Fig.14 time series of (a) rain, (b) soil moisture in top 10 cm of soil, (c) drainage, (d) surface runoff, (e) temperature in soil top layer, (f) potential evaporation, (g) evaporation, (h) bare soil evaporation, (i) interception loss, (j) transpiration, on landpoint X=130°E, Y=13°30'S (north Australia)*

A non-uniform distribution of precipitation also leads to peaks of potential evaporation over time steps without rainfall (Fig. 14d), thus with a higher surface temperature (Fig. 14 b). In addition, it produces discontinuities in transpiration and interception loss (Fig. 14 i and j). Evaporation (Fig. 14f) and potential evaporation are phase shifted, because the main component of evaporation (interception loss) is itself in opposite phase with potential evaporation. This is due to the fact that canopy water storage is the limiting factor of the interception loss: the water storage on the canopy is always small and can quickly become zero when precipitation ceases, thus shutting down interception loss. In contrast, the other components of evaporation (transpiration and bare soil evaporation (Fig. 14j and h) are in phase with potential evaporation because it is the main driving factor. Besides, there is no transpiration in ORCHIDEE until interception loss ceases, which explains the transpiration being phase shifted with interception loss.

In the studied case, in addition to reducing peaks of all evaporation variables, a uniform distribution of precipitation slightly increases total evaporation (from 3.242 to 3.317 mm/d), even though the potential evaporation decreases. This comes mostly from the response of interception loss, which is the main component of evaporation, and is much high under uniform precipitation loading, to the detriment of transpiration. Note, however, that these quantitative results might not be true in other places/situations.

Based on these analyzes, the simulations for the SP-MIP experiments will be run using a uniform distribution of precipitation, to prevent from spurious hourly variability of evaporation.

References:

New Mark, Lister David, Hulme Mike and Makin Ian, 2002. A high-resolution data set of surface climate over global land areas. Clim. Res. 21, 1–25.

Kim H et al. (2017) A century-long global surface meteorology for offline terrestrial simulations, in preparation.