

---

## Response to Hans-Werner Jacobi

---

Do the authors use the term “equivalent black carbon” as defined regarding atmospheric black carbon measurements (i.e. linked to aethalometer measurements) or do the authors imply that all absorbing impurities in the snow are represented by black carbon?

Until now, in the radiative model TARTES, all absorbing impurities are represented by black carbon. That’s why we use the term « equivalent black carbon ».

What are the simulated impurity concentrations in the snowpack? What are the calculated dry deposition fluxes? Do the concentrations and fluxes agree with what can be expected from observations? What atmospheric concentrations are required to maintain the assumed fluxes?

The dry and wet deposition fluxes are perturbed parameters of the ensemble. This is explained section 3.4 “Perturbation of impurity deposition rate”. The dry deposition flux varies from 0 to  $0.5 \text{ ng g}^{-1} \text{ s}^{-1}$  with a median value is  $0.015 \text{ ng g}^{-1} \text{ s}^{-1}$  in concentration which roughly corresponds to  $3.5 \text{ kg m}^2 \text{ s}^{-1}$ . This is the same order of magnitude than the dry deposition rates obtained for this location using ALADIN-Climat (Nabat et al., 2015) simulations.

The simulated impurity content varies a lot, for example fresh snow values varies from 0 to  $500 \text{ ng g}^{-1}$  BC equivalent with a median value of  $100 \text{ ng g}^{-1}$ . Measurement in the field performed at col de Porte on the 2014/02/11 on fresh snow shows a value of  $20 \text{ ng g}^{-1}$  BC equivalent at the surface. The evolution of impurity content in the simulated snowpack has to be investigated using more field measurements but this is not the main objective of the present study.

What is the basis for distributing the deposited impurities in the top 5 cm?

The exponential decay is used to mimic the dry deposition of impurity at the surface of the snow layer. The value of 5 cm was chosen almost arbitrarily. Changing this value (within a reasonable range) does not significantly impact the simulations.

Since the absorbing impurities can have a large impact on the albedo, can the MODIS data be used to constrain the parameters used for the deposition?

Yes definitely, the assimilation scheme can be used to provide updated value for deposition rates and that’s one of the thing we plan to do in future work.

*The authors are very grateful for this in depth review of the paper. All the reviewer comments have been taken into account in the new version of the manuscript. This is described in the detailed response below each comment. Proposed changes in the new version of the manuscript are highlighted in bold.*

---

## ANONYMOUS REVIEWER, REVIEW 1

---

### GENERAL COMMENTS

*- I think the authors have conducted a valuable and interesting analysis. To my knowledge, the assimilation of reflectance data is novel and the authors demonstrate a clear benefit for this practice. They have established a basis for future work that may have a greater impact, including the use of real-time remote sensing data from MODIS and operational implementations. I think the paper should be published pending attention to a suite of minor revisions.*

*- The manuscript lacks substantive discussion of the results, and this is the main weakness of the study in my opinion. I think the authors need to place some attention on contextualizing their results, comparing their results to prior research, and anticipating future work.*

*We completely agree on this point and for that matter, we modified the introduction section, the section 5 “assimilation of MODIS-like reflectances” and also the section 6.4 “Combining reflectance and snow depth assimilation”.*

*All modifications are indicated here in bold.*

#### Introduction

-----NEW-----

Seasonal snowpack modeling is a crucial issue for a large range of applications, including the forecast of natural hazards such as avalanches or floods, or the study of climate change (e.g. \citealp{durand1999,lehning2006,bavay2013}). The most sophisticated detailed snowpack models represent the evolution of snow microstructure and the layering of snow physical properties \citep{brun1989,brun1992,jordan1991,Bartelt2002,vionnet2012} in response to meteorological conditions. Despite constant efforts to improve these models, large uncertainties remain in the representation of the snow physics, as well as in the meteorological forcings \citep{carpenter2004,essery2013,raleigh2014}. These uncertainties are highly amplified when propagated to avalanche hazard models \citep{vernay2015}. For operational applications, the assimilation of observations can help reduce the impact of the model and forcing uncertainties in the snowpack simulations \citep[e.g.]{dechant2011}.

Satellite observations are becoming an essential component of snow modeling and forecasting systems.

\textit{In situ} measurements are the most detailed and accurate observations of the snowpack, but their spatial distribution is far too scarce to capture the high spatial variability of the seasonal snowpack properties and improve snowpack simulations through their assimilation. For this reason, the assimilation of satellite observations of snow is an active area of research.

Snow remote sensing is primarily performed in the microwave (passive and active), visible and near-infrared spectra. Since the direct assimilation of such data requires the use of radiative transfer models, a common and simple approach consists in using satellite-based snow products. In particular, the assimilation of snow cover fraction (SCF) estimates derived from optical sensors (such as MODIS) and Snow Water Equivalent (SWE) or Snow Depth (SD) estimates derived from passive microwave sensors (such as AMSR-E) has been investigated extensively \citep{sun2004, andreadis2006, clark2006, dong2007, de2012, liu2013}.

These studies have suggested that, most of the time, assimilating snow observations may be useful to improve snowpack estimation. SWE or SD assimilation generally outperforms the assimilation of SCF only, except from \citep{andreadis2006} because of large errors in the AMSR-E SWE products. The assimilation of both combined revealed larger benefit by mitigating sensors limitations. Recently, \citep{navari2015} investigated the assimilation of (synthetic) ice surface temperature while \citep{dumont2012} also experimented the assimilation of albedo retrievals, both from optical sensors. \citep{dumont2012} obtained a mass balance RMSE decrease of up to 40\% assimilating albedo data. However, satellite snow products are derived using retrieval algorithms which are not perfect and, perhaps more importantly, not physically consistent with the snowpack model used for the data assimilation. For this reason, and as advocated by \citep{durand2009} who tested the assimilation of in situ microwave radiance observations, assimilating the original satellite radiance data should be preferred when possible.

#### % Active microwave data

The potential of assimilating passive microwave radiances (in the form of brightness temperature) collected by AMSR-E satellite have been examined by \citep{dechant2011} and \citep{che2014}. Significant improvements in the SWE/SD predictions occurred but only during the accumulation period. Though the melt period, when the snowpack is wet, liquid water alters the microwave signal resulting in a lower performance. Moreover, for small-scale applications in mountainous areas, the coarse spatial resolution of these data considerably reduces their usefulness \citep{foster2005, cordisco2006, dong2007, tedesco2010}. As for active microwave measurements, several tests have been conducted to assimilate the satellite signal (e.g. \citealp{stankov2008, phan}). These tests were however limited by the accuracy of the forward electromagnetic models and by the current lack of satellite data at a daily or even weekly time frequency.

#### % VIS/NIR microwave data

Visible and near-infrared reflectances from satellite observations have never been assimilated into snowpack models despite their great sensitivity to the snowpack properties \citep{warren1982}. Even if cloud cover might limit their utility, medium and high spatial resolution data are available at daily resolution from several optical sensors (e.g. MODerate Imaging Spectrometer, Visible Infrared Imaging Radiometer Suite) and seem to be quite

suitable for complex topography \citep{sirguey2009}. In particular, the MODIS sensor, onboard TERRA and AQUA satellites, offers a daily coverage and provides reflectance measurements in seven bands distributed in the visible (at 250 to 500 m spatial resolution), near and short-wave infrared wavelengths. Surface bi-hemispherical reflectances corrected from complex topographic effects in mountainous areas can be computed \citep{sirguey2009} and have been evaluated and used in several rugged areas \citep{dumont2012,brun2015}.

## 5. Assimilation of MODIS reflectances:

-----NEW-----

Figure \ref{RMSE\_8009} shows the time evolution of the RMSE with assimilation at every observation time, at the end of the forecast step (blue solid line) and just after the filter analysis (blue dotted line). These results are compared to the RMSE without assimilation (grey lines). The RMSE of the ensemble with assimilation is always lower than the RMSE without assimilation. Averaged over the season, a reduction of 46\% was obtained for SD and 44\% for SWE, (Table \ref{RMSE\_result}: seasonal RMSE for SD: 0.07 m; SWE: 19.7 kg\,m<sup>-2</sup> compared to 0.13 m and 35.4 kg\,m<sup>-2</sup> from the ensemble without assimilation). These results indicate the usefulness of using spectral optical radiance rather than albedo data since \cite{dumont2012} obtained an improvement in SD estimate of only 14\% when assimilating albedo retrievals from MODIS sensor. It is remarkable that, despite the significant RMSE reduction in our experiment, there is most of the time no strong reduction of the RMSE from a single analysis. The reduced RMSEs with assimilation are consequently due to the successive observations throughout the season, highlighting the role of model dynamics.

[...]

The remarks stated above for the season 2010/2011 hold for the other seasons. Figure \ref{RMSE\_5seasons} reports the time evolution of the SD and SWE RMSEs for all the selected seasons, in the experiments without assimilation (red lines) and with assimilation of reflectances (blue line; the experiments shown in green and black are discussed in the next section). On average, SD and SWE RMSEs are reduced by 45\% and 48\%, respectively. This is comparable with results of \cite{che2014}, who assimilate radiances in the microwave spectrum from AMSR-E, and reduce the SD RMSE by 50\%. However, passive microwave observations are very sensitive to liquid water. Consequently, the performance of the assimilation during the melting period is reduced (\cite{che2014} reduce the SD RMSE up to 61\% from January to March, during only the dry snow period). In contrast, our results show a well-marked reduction of errors near the end of the seasons (Figure \ref{RMSE\_result}, grey lines and blue dotted lines). Our results are also consistent with those from \cite{liu2013} assimilating MODIS-derived Snow Cover Fractions (SCF), after a processing of the retrieval to improve accuracy of cloud coverage and snow mapping. Without this processing, the performance of SCF assimilation falls, with a SWE RMSE reduction near 10-20\%, similarly to \cite{andreadis2006}.

## 6.4 Combining reflectances and snow depth assimilation

-----NEW-----

These results indicate the usefulness of combining these two datasets in operational applications. \cite{liu2013} reached a similar conclusion by combining the assimilation of SCF and SD (with an SWE RMSE reduction up to 72\%; up to 74\% in our study). However, given the strong spatial variability of the snow cover, the spatial representativity of punctual SD measurements may make their assimilation questionable. This issue should be addressed with experiments over two-dimensional, realistic domains.

*- I think it would be beneficial for the authors to discuss what is necessary to include a radiative transfer model like TARTES in a snow model (instead of an albedo parameterization). This is important because it seems that the only way to assimilate remotely sensed reflectance into an existing model is to ensure that it has the capability of outputting reflectance data at different wavelengths. Many, if not most, existing snow models do not have this capacity. So it would be useful to have some discussion about the changes required in the model structure, runtime, and operation, and what level of complexity is needed in order to achieve the methodologies demonstrated with Crocus.*

*The section 2.3 has been modified accordingly.*

Given that satellite observations indirectly relate to the quantities of interest, an observation operator is required to link the satellite observation and the model state variables \citep{reichle2008}. This operator transforms the model variables into diagnostic variables to allow a direct comparison with satellite observations, preserving the physical consistency of the satellite signal with the snow model.

To this end, a new radiative transfer model was recently implemented in Crocus to calculate spectral reflectances that can be used for the comparison and the assimilation of satellite data such as MODIS. This model, named TARTES (Two-streAm Radiative TransfEr in Snow, Libois et al., 2013, 2014), simulates the absorption of solar radiation within the stratified snowpack using the  $\delta$ -Eddington approximation, with a spectral resolution of 20nm. This contrasts with the original version of Crocus, where albedo was computed for three large spectral bands only and from the properties of the first two layers (Brun et al., 1992; Vionnet et al., 2012).

TARTES is implemented as an optional module to be called instead of the original Crocus albedo scheme. This implementation has no significant impact on the model structure but increases the computation time of roughly a factor 10 depending on the number of snow layers and the snow depth.

TARTES makes use of four Crocus prognostic variables (specific surface area -- SSA, density, snow layer thickness and impurity content) and the angular and spectral characteristics of the incident radiance (e.g. the solar zenith angle and the presence of cloud cover). The computation of SSA has recently been implemented by \citet{carmagnola2014}.

The use of a full radiative transfer model embedded within the snowpack model enables the assimilation of the satellite reflectance data, therefore avoiding the introduction of uncertainties from an external retrieval algorithm. And beyond its use for the assimilation of

reflectances, TARTES also provides a more accurate calculation of light absorption parameters, leading to better simulations of the snowpack.

*- In several places in the manuscript, the authors use the word "envelops" as a noun, but it is a verb. I think they mean "envelopes" instead in some (but not all) of these cases. Please rectify this word usage.*

*This has been modified in the new version of the manuscript. Thank you.*

## SPECIFIC COMMENTS

*1. The second paragraph of the abstract (page 6830, lines 11-22) does not make it consistently clear that MODIS data are not actually used in the study. The first sentence suggests MODIS reflectance data are used, but the subsequent sentences refer to MODIS-like data. The authors need to include additional clarification here.*

*Both reviewers indicate the need to clarify that data are synthetic and/or this study is based on twin experiments. Some modifications were consequently done in the abstract and in the introduction (see below).*

-----NEW-----: abstract

**This paper examines the ability of optical reflectance data assimilation to improve snow depth and snow water equivalent simulations from a chain of models with the SAFRAN meteorological model driving the detailed multilayer snowpack model Crocus now including a two-stream radiative transfer model for snow, TARTES.**

The direct use of reflectance data, allowed by TARTES, instead of higher level snow products, **mitigates** uncertainties due to commonly used retrieval algorithms.

Data assimilation is performed with an ensemble-based method, the Sequential Importance Resampling Particle filter, to represent simulation uncertainties. **In snowpack modeling, uncertainties of simulations are primarily assigned to meteorological forcings. Here, a method of stochastic perturbation based on an autoregressive model is implemented to explicitly simulate the consequences of these uncertainties on the snowpack estimates.**

Through twin experiments, the assimilation of synthetic spectral reflectances matching the MODerate resolution Imaging Spectroradiometer (MODIS) spectral bands is examined over five seasons at the Col du Lautaret, located in the French Alps. Overall, the assimilation of MODIS-like data reduces by 45\% the root mean square errors (RMSE) on snow depth and snow water equivalent. At this study site, the lack of MODIS data on cloudy days does not affect the assimilation performance significantly. The combined assimilation of MODIS-like reflectances and a few snow depth measurements throughout the 2010/2011 season further reduces RMSEs by roughly 70%. This work suggests that the assimilation of optical reflectances has the potential to become an essential component of spatialized snowpack simulation and forecast systems. The assimilation of real MODIS data will be investigated in future works.

-----NEW-----: Intro

**The work presented in this article examines the possibility, the relevance and the limitations of assimilating visible and near-infrared satellite reflectances into a multilayer snowpack model.**

A convenient approach, known as twin experiment, uses synthetic data in the same spectral bands than the real data, to examine the content of information of the observations, and the impacts we can expect from their assimilation. In twin experiments, the model used to create the synthetic data is the same as the model used for the assimilation. The synthetic observations are extracted from a member of the ensemble considered as the “true” state. Twin experiments are preferred in this first study in order to focus on the information content of the observations and to avoid the problem of observational biases.

Data assimilation is performed with a particle filter and a Sequential Importance Resampling (SIR) algorithm \citep{gordon1993,vanleeuwen2009,vanleeuwen2014}. The particle filter is easy to implement, free of hypotheses about the nature of the model and the observations, and provides uncertainties in the estimation of the snowpack state.

*2. The biases in shortwave and longwave radiation are of opposite sign (Table 1, left column). Does this reflect some specific shortcoming of SAFRAN, such as problems with a low bias in cloud conditions or a high bias in atmospheric transmissivity? In other words, are the radiation errors linked in some physical way, or is it just by chance that the biases are positive for shortwave and negative for longwave? Does vegetation/topographic shading at the CdP site factor into the positive SAFRAN shortwave bias, and does this influence the longwave estimation in any capacity? I recommend addressing these questions in the paragraph that discusses the discrepancies between SAFRAN and observations (page 6838, Lines 9-17). This paragraph currently focuses on temperature, precipitation, and wind speed but could be improved with more attention to the radiation components.*

*Some additional checks have revealed a higher bias (but of opposite sign) for longwave and shortwave radiation in case of clear sky, that may reflect a shortcoming in the atmospheric transmissivity calculation.*

*In addition, measured longwave fluxes are affected by vegetation and slopes while that is not considered in the SAFRAN longwave estimates. The altitude of clouds might also be too simplistically simulated in SAFRAN.*

*Concerning the shortwave radiation, a shading mask is applied to SAFRAN estimates to account for topographic and vegetation shading effects but the seasonal evolution of nearby vegetation is not taken into account in this correction.*

*Modifications were done in section 3.2 concerning this point (in bold below).*

*According also to review 2, specific comments 13 to 17, section 3.2 was rephrased as follows. (Answers to review 2, specific comments 13 to 17 are located directly bellow each comment.)*

**-----NEW----- Section 3.2: Quantification of meteorological forcing uncertainties**

To quantify and calibrate the meteorological forcing uncertainties, we compare 18 years of surface meteorology from SAFRAN reanalysis with \textit{in-situ} observations at the CdP. A long time-series from 1993 to present \citep{morin2012} being available at this site, uncertainties in the SAFRAN meteorological reanalysis can be estimated.

Table \ref{perturbed\_forcing\_stat} (left column) reports the bias and the standard deviation of the difference between SAFRAN and the observations carried out at the CdP site, for each meteorological variable (the right column reports other data discussed later). The table reflects differences between SAFRAN and \textit{in-situ} observations, resulting from the different spatial representativities of both sources, the intrinsic errors of the analysis system and measurement errors.

As highlighted by \cite{quintana2008} who conducted an extended evaluation of SAFRAN reanalysis but over a shorter period (one year), the large discrepancies between the model and the observations can be explained by local effects due to orography and vegetation and, for the precipitation and wind speed, by the time interpolation necessary to obtain hourly forcing fields from the daily analysis. For example, the precipitation analysis is performed on a daily basis in order to include in the analysis the numerous rain gauges observations.

**Radiation fluxes uncertainty might be attributed to biases in cloud coverage and altitude estimates, effects of vegetation and surrounding slopes that are not taken into account for longwave estimates. Finally, the shading mask for shortwave radiation does not account for vegetation evolution that can also lead to shortwave flux discrepancies.** \cite{durand2009} carried out, only on a limited set of variables, a more systematic evaluation of SAFRAN for the 1958-2002 period using 43 sites in the French Alps. Averaged over all locations, the RMSE on air temperature are similar to the one computed in our study. However, their results also highlight the spatial variability of SAFRAN performance (site RMSE ranges from -0.8 to +1.5 \textdegree C). Nevertheless, this will not have a strong impact in this study since it is based on twin experiments.

*3. I am unsure why the shortwave perturbations are additive while the longwave perturbations are multiplicative (page 6839, lines 8-10). Please clarify the logic behind this decision.*

*The reviewer is actually spotting a typo here: the perturbations was multiplicative for SW and additive for LW. The logic behind this is clarified in the new version of the manuscript, section 3.3, also reported after the answer to comment 4 below. Note that Section 3.3 has been entirely modified after comment 3, 4, and comments from Reviewer 2.*

*4. The authors rightfully attempt to maintain physical consistency amongst the meteorological variables (page 6839, Lines 16-19). Are any efforts made to examine the physical consistency between shortwave and longwave radiation? For example, a high shortwave value and a high longwave value might not be physically realistic because the high shortwave implies no cloud cover whereas the high longwave value can imply cloud presence.*

*No, there is actually nothing done in the ensemble to preserve the physical consistency between the longwave and shortwave radiation. For real data assimilation, the physical consistency between the meteorological variables will be crucial: we plan to use forcing ensembles from meteorological ensemble simulations.*

*The limitations of our meteorological ensemble are further enlightened and discussed in the new section 3.3. "Building the ensemble of meteorological forcings".*



### New section 3.3 : “Building the ensemble of meteorological forcings”

The sample of meteorological forcings is formed by perturbing the original SAFRAN reanalysis with a random noise commensurate with the actual uncertainty. We thus build an ensemble of meteorological forcings with a negligible bias with respect to the SAFRAN reanalysis and a standard deviation close to the one computed from CdP statistics (Table \ref{perturbed\_forcing\_stat}, left column).

To keep the procedure simple and preserve physically consistent time variations of the forcings, the random perturbations are computed using a first-order autoregressive (AR(1)) model \citep{deodatis1988} for each variable:

$$X_t = \varphi X_{t-1} + \epsilon_t,$$

with  $X$  being the perturbation value at time  $t$  and  $t-1$ .  $\varphi$  is the AR(1) model parameter and can be written  $\varphi = e^{-\frac{\Delta t}{\tau}}$ ,  $\Delta t$  being the time step and  $\tau$  the decorrelation time.

Parameter  $\tau$  is adjusted for each variable, so that the perturbed variable exhibits the same frequency of temporal variations than the original variable (Fig. \ref{Supp-Snowfall} bottom, in blue).

The amplitudes of the meteorological uncertainties are introduced with  $\epsilon_t$ , a white noise process with zero mean and constant variance  $\sigma^2$ . Variance  $\sigma^2$  is computed from each standard deviation of the residuals between the reanalysis and observations at CdP ( $\sigma_{CdP}$ ) : Table \ref{perturbed\_forcing\_stat}, left column) following this equation:

$$\begin{equation} \sigma^2 = \sigma_{CdP}^2 \times (1 - \varphi^2) \end{equation}$$

Finally, for each meteorological variable, the selection of an additive or multiplicative perturbation method is driven by (i) the nature of the variable (ii) the dependency of the model-measurement difference to the measured values as detailed below.

For precipitation rates, shortwave radiations and wind speed, the choice of a multiplicative method is motivated by the following reasons:

- SAFRAN reanalysis captures well the occurrence of precipitation (since it assimilates surface observation network) but are more subject to errors in the amount of precipitation;
- Regarding shortwave flux and wind speed, the model biases exhibit a linear dependency to the value of the variable (not shown). Consequently, a multiplicative method was selected.

For longwave radiation and air temperature, given that there is no dependency between the model biases and the field values, an addition method is chosen.

At every time step the perturbation  $X_t$  is applied as follows:

For the additive method,  $variable_t = variable_t + X_t$ ; For the multiplicative method, the perturbation is centered on 1 ( $Y_t$ ) before multiplying the variable.

$$Y_t = X_t + 1$$

$$variable_t = variable_t \times Y_t$$

For the multiplicative method, the perturbations are bounded by 0.5 and 1.5 to avoid extreme values. The result from this perturbation method is illustrated by Fig. \ref{Supp-Snowfall} which shows the snowfall rates over a one week period, as described by SAFRAN reanalysis, a realization of the perturbed analysis, and the full ensemble of perturbed analysis.

To maintain further physical consistency between the meteorological variables, snowfall is changed to rainfall if air temperature is higher than 274.5 K and the shortwave radiation is bounded to  $200 \text{ W m}^{-2}$  in case of rain or snow fall due to the inherent cloud cover. This behavior is consistent with the CdP statistics where over 18 years, during a precipitation period, the measured in situ shortwave radiation rarely exceeds  $W \text{ m}^{-2}$ .

Ensembles are generated with model errors coming from the statistics of the CdP site but as explained previously, the assimilation framework is based on the CdL area. Some adjustments in the building of ensembles are also required to take into account differences between these two areas.

In particular, the forest at CdP affects the local wind field and the radiative fluxes \citep{morin2012}, which explains a large part of the variability of SAFRAN errors at CdP. At CdL, an open meadow area, such variability is unlikely. To limit the overspreading of the forcing ensemble, the standard deviation used in the equation (3) for wind speed, short and longwave radiation are reduced to  $0.6 \text{ m s}^{-2}$ , 70 and  $7 \text{ W m}^{-2}$ , respectively, against  $1.12 \text{ m s}^{-2}$ , 79 and  $24.5 \text{ W m}^{-2}$  (Table \ref{perturbed\_forcing\_stat}, left column, values in brackets).

As shown in Table 1, the standard deviations computed from the generated ensemble (right column) are close to the ones prescribed to generate it (left column).

In the end, this stochastic method of perturbations makes possible the construction of an ensemble of perturbed forcings which are required when using ensemble methods.

The calibration of the perturbations are based on the CdP statistics while their temporal correlation is ensured by the AR(1) model.

The perturbation method exhibits some obvious limitations. Inter-variable correlations are indeed not taken into account in the ensemble except from the precipitation phase and the maximum value of short-wave radiation in case of precipitation. Adjustments to CdL are somewhat subjective, but this is not crucial in our twin experiment context since the considered truth will be simulated running Crocus with one forcing member drawn from this generated ensemble. A more physically consistent ensemble will be required when real data assimilation is investigated.

## TECHNICAL CORRECTIONS

-1. Page 6830, Line 5: *Uncertainties can never be ruled out in any type of dataset; they can only be identified and reduced through improved datasets. Please rephrase.*

*Correction done in Specific Comment 1*

-2. Page 6830, Line 11 and Page 6832, Line 11: *The full name of MODIS is “MODerate resolution imaging spectroradiometer”. Please correct these lines.*

*Thanks (both reviewers) to have seen this mistake.*

*Correction done in Specific Comment 1*

- 3. Page 6831, Line 9: *Replace “reducing” with “reduce”.*

*Done*

-4. Page 6833, Line 6: *Should read “statistics”.*

*Done*

-5. Page 6833, Line 25: *Replace “including” with “includes”.*

*Done*

- 6. Page 6835, Line 21: *Replace “first” with “top” to be more specific to the location of these layers.*

*Done*

-7. Page 6837, Line 15: *The phrasing would sound better as “other physical parameterizations”.*

*Done*

- 8. Page 6839, Line 21: *Should read “captures”.*

*Done*

- 9. Page 6840, Line 4: *Based on Table 1, I think this should read “58.3” instead of “70”.*

*No, 58.3 corresponds to the standard deviation obtained from the difference between the SAFRAN reanalysis with the generated ensemble for the shortwave flux. 70 is the value used in the normal distribution to generate it.*

*We added in brackets the adjusted standard deviation values (used to generate perturbations) in the table 1. The caption was modified as well. Details of these adjustments are given in the new section 3.3*

Variables	CdP: Reanalysis - Observations		CdL: Reanalysis - Ensemble	
	Bias	std: $\sigma_{CdP}$	Bias	std: $\sigma_{CdL}$
Air temperature (C)	0.28	1.08	5.0 e-03	1.07
Wind speed (m s <sup>-1</sup> )	0.2	1.12 (0.6)	4.0 e-04	0.4
Shortwave radiation (W m <sup>-2</sup> )	22.4	79 (70)	-3.1 e-03	58.3
Longwave radiation (W m <sup>-2</sup> )	-14.0	24.5 (7)	2.0 e-02	7.0
Snowfall rate (kg m <sup>-2</sup> h <sup>-1</sup> )	-2.0 e-02	0.4	5.0 e-03	0.1
Rainfall rate (kg m <sup>-2</sup> h <sup>-1</sup> )	7.2 e-03	0.5	-5.0 e-03	0.1

**Table 1.** Bias and standard deviations (std) of the differences between SAFRAN reanalysis and *in-situ* observations (left) and the differences between SAFRAN reanalysis and the ensemble built up in the present study (right), for the perturbed meteorological forcings. The first set of statistics is derived from 18 years (1993-2011) of observations and reanalysis at the CdP and the second set is derived from our 300-members ensemble over the 2010/2011 hydrological season. The values in brackets correspond to the adjusted standard deviation used to generate the ensemble at CdL site.

-10. Page 6841, Line 17: What do you specifically mean here by the “spread” in the melt-out date? Is this the range (max-min) or the variance or some other statistic? A definition of uncertainty in melt-out date appears later (page 6844, Lines 13-14), so it might be helpful to bring this definition earlier in the text.

We meant the range. Since this was unclear, the following modification were done in the text:

-----NEW-----

“The spread of the SD and SWE ensembles (Fig. 1b-c) is the largest at the end of the season, leading to a range of 24 days from the first to the last member to fully melt.”

Done

- 11. Page 6843, Line 16: Add “to” before “presenting”.

Done

- 12. Page 6846, Line 1: Should read “simpler” instead of “simplest”.

Done

-13. Page 6847, Line 8: Add “to” after “according”.

Done

- 14. Page 6849, Line 6: Should read “cloud coverage” instead of “clouds coverage”.

Done

- 15. Page 6850, Line 13: Rephrase to say “The ensemble spread retrieves almost the same value as . . .”

Following suggestions of both Reviewers (see review 2, Specific Comment 68), we modified the sentence this way:

-----NEW-----

The ensemble spread at the end of the season returns to almost the same value than the experiment without assimilation.

Done

- 16. Page 6854, Line 21: Add “as” before “MODIS”.

Done

- 17. Page 6854, Lines 22-23: Rephrase to say “Combining reflectance assimilation with SD assimilation at 4 dates during the snow season leads to. . .”

-----NEW-----

Combining reflectance assimilation with SD assimilation at 4 dates during the snow season leads to a decrease of SD and SWE RMSE by a factor close to 3.

Done

- 18. Page 6855, Line 11: Replace “what” with “which”.

Done

- 19. Page 6855, Line 18: Should read “through” instead of “though”.

Done

#### TABLE AND FIGURE COMMENTS

- Figure 1c: The upper limit of the vertical axis cuts off the SWE ensemble. Please extend so the entire ensemble can be seen.

Done

- Table 2: The understandability of this table would be improved if the column headings not only included the Figure reference, but also a brief description of what is represented in each experiment. For example, the “Fig. S2” column should also have a heading that says something like “Reflectance – all days” while the “Fig. S7” column should have a heading saying “SD – clear sky days”, etc. This will help the reader by not requiring them to keep searching for what is tested in each scenario.

Thank you for this welcome suggestion.

So we did some modifications, adding new lines (and a new column according review 2, specific comment 73) as displayed below:

Results reported in	Fig. 1	Fig. 3	Fig. S7	Fig. S2	Fig. S3	Fig. S4	Fig. S5	Fig. S6	Fig. S8	Fig. 5
Variable assimilated	No AD	Baseline	Clear sky days	All days	Accu.	Melt	Before Snowf	Refl.	Refl. + SD	Refl.
Assimilation timing	No AD	Baseline	Clear sky days	All days	Accu.	Melt	Before Snowf	After Snowf	Refl. + SD	All seasons
SD (m)	0.13	0.07	0.03	0.05	0.05	0.12	0.07	0.12	0.04	0.07
SWE (kg m <sup>-2</sup> )	35.4	19.7	7.4	14.4	12.9	35.5	21.8	37.2	9.6	20.2

**Table 2.** SD and SWE seasonal averaged RMSE computed with respect to the synthetic truth for all experiments.

---

## ANONYMOUS, REVIEW 2

---

### General Comments:

The study is important and relevant in that it has showed the potential usefulness of assimilating reflectance data into a snow model. The design of experiments appears to be well thought out, and the results of the sensitivity experiments are particularly interesting. The presentation of procedures carried out is sometimes unclear. The study is scientifically sound, and most of the comments relate to making the presentation more clear. I feel the paper should be published after relatively minor changes discussed below:

1. The abstract be specific as to the procedures that were carried out. It should make clear from the outset that the authors are examining the usefulness of assimilating reflectance data, but are not using real reflectance data during assimilation. The Crocus model should be mentioned as the model used to calculate snow depth, and the source of meteorological inputs (SAFRAN), as well as the method used to generate an ensemble of input forcing should be mentioned.

*The abstract has been modified accordingly.*

*See details provided in reviewer 1, Specific Comment 1*

2. It should be clear throughout the paper wherever “observations” are synthetic observations derived from model results. In these instances “truth” should be changed to “synthetic truth” and “observations” to “synthetic observations”, etc.

*The 2 reviewers attest the need to clarify that this study is based on twin experiments with synthetic data. To this end, we also added a paragraph in the introduction section to present what the twin experiment method is. Also we changed “truth” to “synthetic truth” and “observations” to “synthetic observations”.*

*See details provided in reviewer 1, Specific Comment 1*

3. A paper describing the potential for assimilating MODIS data in a distributed way over the Greenland Ice Sheet has been recently published (Navari et al., 2015), but discusses assimilation of ice surface temperature data derived from MODIS (i.e. in a different part of the electromagnetic spectrum). Since that study also uses MODIS data (albeit in a different way with far infrared measurements from MODIS) it could be referred to in the introduction. Navari, M. Margulis, S. A., Bateni, S. M., Tedesco, M., Alexander, P., and Fettweis, X.: Feasibility of improving a priori regional climate model estimates of Greenland ice sheet surface mass loss through assimilation of measured ice surface temperatures, *The Cryosphere*, 10, 103-120, doi: 10.5194/tc-10-103-2016, 2016.

*Thank you for the reference.*

*The reference has been added in the introduction:*

*See review1, General Comment 1.*

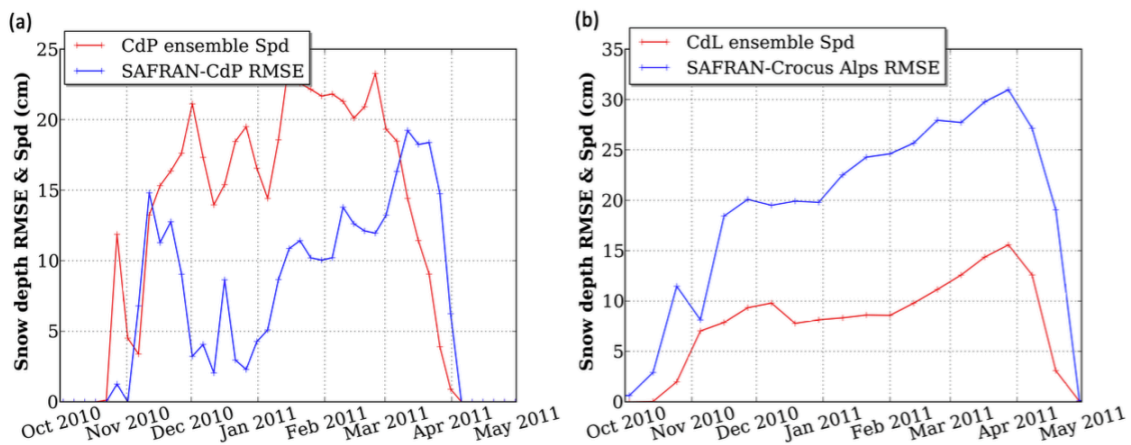
4. The application of errors from Col de Porte at Col de Lautaret needs more discussion. The methods used to adjust errors from one site to the other site should be discussed in more detail, perhaps in the supplement.

A more detailed discussion on this adjustment has been added in the new section 3.3. (see modifications in response of Review 1, Specific Comments 3 & 4.

5. The comparison between the ensemble of simulations at CdL and RMSE at multiple locations seems unfair in that the spatial distribution of snow depth errors is not necessarily the same as the distribution of errors associated with errors in input forcing. As the authors mentioned, errors at CdP and CdL are different partly because of differences between the sites (forested vs. open). It seems likely that errors in SND at CdL will be smaller than those at other locations because of the lack of forest cover. Therefore I don't see the purpose of the comparison described in Section 3.5, except perhaps to illustrate that errors may be larger when considering multiple locations, and therefore, perturbations may need to be adjusted spatially in future use of "real" data assimilation. It may make more sense to create an ensemble at CdP, and to see whether the ensemble method captures the RSME of SND at that particular location, given the known uncertainties in forcing data at CdP. This would somewhat validate the method being used to generate errors in SND and other variables. Following your suggestion, we created an ensemble at CdP and reported the results on new Figure 2.

The section 3.6 "Dispersion of the ensemble of Crocus simulations" has become as follows.

-----NEW-----



**Figure 2.** Time evolution over the 2010/2011 season of (in red) the SD ensemble Spd with respect to the ensemble mean and (in blue) the SD RMSE between SAFRAN-Crocus estimates and *in-situ* observations. (a) for the CdP site and (b) for the CdL ensemble compared to the multiple Alps stations at the same elevation than CdL.

Here we assess whether our ensemble represents a realistic spread of SD over time with respect to previous evaluations of the model through a spread-skill plot.

Given that no SD measurements were systematically carried out at the CdL site, we were not able to evaluate our ensemble spread from SAFRAN-Crocus simulations with a time series of *in-situ* measurements at this site. But, as demonstrated by [fortin2014](#), the ability



of the ensemble spread to depict the simulation error can be evaluated by the comparison of the root-mean-squared error (RMSE) and the ensemble spread (Spd) with respect to the ensemble mean.

Firstly, using the method previously described, an ensemble of Crocus simulations was carried out at the CdP site with no adjustment on CdP statistics to evaluate the relevance of our perturbation method by comparing the RMSE between SAFRAN-Crocus simulation and in situ measurements with the Spd of the CdP ensemble. Then, we compare the Spd of our ensemble simulation at the CdL site with a SAFRAN-Crocus RMSE computed from the difference between SD Crocus estimates with in situ SD measurements across multiple stations (at the same elevation than CdL). We used roughly 60 daily snow depth measurements stations from the M\et\eo-France observation stations network (only stations within the same altitude range as the CdL site (1800 - 2200m a.s.l.)).

The multiple station RMSE and Spd terms are defined as follows, for a variable  $X$ ,

$$\text{RMSE}(X) = \left( \frac{1}{M} \sum_{t=1}^M \frac{1}{N_k} \sum_{k=1}^{N_k} (X_{t,k}^{\text{model}} - X_{t,k}^{\text{in situ}})^2 \right)^{1/2},$$

$$\text{Spd}(X) = \left( \frac{1}{M} \sum_{t=1}^M \frac{1}{N_e} \sum_{n=1}^{N_e} (X_{t,n} - \bar{X}_t)^2 \right)^{1/2},$$

where  $M$  represents the number of time steps,  $N_e$  the size of the ensemble and  $N_k$  the number of in situ measurements. The SD value of the ensemble member  $n$  at the date  $k$  is  $X_{t,n}$  and  $\bar{X}_t$  is the mean of the ensemble at the date  $t$ .

The value from SAFRAN-Crocus simulation at the site measurement  $k$  and at the date  $t$  is given by  $X_{t,k}^{\text{model}}$ , and  $X_{t,k}^{\text{in situ}}$  is the value from the in situ SD measurement. RMSE and Spd are computed at observation times.

For comparisons based on only one point, the RMSE equation for a variable  $X$  becomes:

$$\text{RMSE}(X) = \left( \frac{1}{M} \sum_{t=1}^M (X_t^{\text{model}} - X_t^{\text{in situ}})^2 \right)^{1/2},$$

Figure \ref{dispersion} (a) shows that at the CdP site the SD dispersion (Spd) of the ensemble is consistent with the RMSE between SAFRAN-Crocus simulation with respect to in situ measurements at this site. This suggests that our perturbation method is able to represent the forcing uncertainties on snowpack simulations. Nevertheless, concerning the CdL area over the 2010/2011 season, the SAFRAN-Crocus RMSE is roughly two times higher than the SD

dispersion (Spd) of our ensemble (**Fig. \ref{dispersion} (b)**). This means that our ensemble is under-dispersive in terms of SD. This may be partly explained by the calibration of perturbations, based on statistics at a location (CdP) which is not highly affected by wind erosion/accumulation in contrast to many other measurement sites. In addition, only meteorological errors are considered in our ensemble whereas the other model errors also contribute to the simulation error.

Nonetheless, given that experiments in the present work are twin and that the observations are selected within the ensemble (synthetic observations), the impact of this under dispersion is not crucial, but will be considered when using real data.

## Specific Comments:

1. P. 6830, Lines 7-8: This sentence is unclear. I'm not sure what is meant by "essentially ascribed". Inclusion of details discussed in the general comments section may allow this sentence to be modified or replaced.

*Some modifications were carried out in the abstract, see review 1: Specific Comment 1.*

*Also, more details are given at the beginning of the section 3.1:*

*"As shown in Raleigh et al. (2015), the meteorological forcings are the major source of uncertainty in snowpack simulations when a meteorological model is used to drive the snow model."*

2. P. 6830, Line 11: Perhaps this sentence can be modified to make it more clear that the reflectances are not real: "The assimilation of synthetic spectral reflectances, designed to match the spectral resolution of the MODerate resolution Imaging Spectroradiometer..." Also note the correct name for MODIS.

*Done*

*See review 1: Specific Comment 1 & Technical Correction 2*

3. P. 6830, Lines 19-21: Since real data have not been assimilated yet, perhaps this statement is slightly too strong. "Should become" could be changed to "has the potential to become" or something similar.

*Done*

4. P. 6833, Lines 8-10: Perhaps here, the forcing data used at Col de Lautaret can be introduced, and the method of generating an ensemble of forcing can be briefly noted as well.

*Done*

-----NEW-----

For this reason, assimilation experiments are carried out at the Col du Lautaret (CdL) located (2058 m a.s.l.) in the Ecrins area, France, which exhibits a large flat open area, above treeline, more suitable for remote sensing. **Consequently, an ensemble of perturbed forcing was generated in order to represent the range of possible weather conditions at the CdL area. To this end, we developed a stochastic method using a first-order autoregressive model based on the estimated meteorological uncertainty.**

5. P. 6833, Line 18: Change "reflectance observations" to "synthetic reflectance observations". Perhaps also change "one point" to "CdL".

*Done*

6. P. 6833, Line 19: Suggest changing "this previous experiment" to "the reference experiment"

*Done*

7. P. 6834, Lines 26-27: Over what period is this threshold applied, a single model time step? *The procedure of layering is activated at the beginning of each time step (15 minutes in our case)*

-----NEW-----

For a snowfall on an existing snowpack, fresh snow is incorporated into the top layer if (i) snow microstructure characteristics are similar, (ii) the top layer is thinner than 1cm and (iii) the snowfall intensity is inferior to 0.03 kg.m<sup>-2</sup>.h<sup>-1</sup>. If one of these criteria is not met or change during the time step, a new top layer is created.

*Done*

Does the threshold change if the time step also changes?

*Actually it does not, but we do not usually change the time step either.*

8. P. 6835, Line 1: Change “identical layers” to “a set of identical layers” for clarity.

*Done*

9. P. 6835, Lines 5-7: Change “layer that is too small relatively” to “layers that are too small relative”, and change “is aggregated with an adjacent one” to “are aggregated with adjacent ones”. *Done*

Is the “optimal vertical profile” an optimal profile of layer thicknesses?

How is this optimal profile determined?

*We directly included the answer into the paragraph.*

-----NEW-----

In absence of snowfall, the model seeks first to merge two thin and adjacent layers with similar microstructure characteristics, or inversely, split the thick ones.

When the number of layers has reached a maximum of 50, layers that are too small relative to a prescribed optimal vertical profile are aggregated with adjacent ones. **This idealized thickness profile depends on the current snow depth and on the user-defined maximal number of layers. To reach the optimal vertical profile, the model first seeks to thin the top layers, most subject to the exchange of energy, and then to keep an appropriate thickness ratio between adjacent snow layers to prevent numerical instabilities in the resolution of the heat diffusion equation through the snowpack**

10. P. 6836, Line 24: Please spell out the acronym SAFRAN.

*Système d’Analyse Fournissant des Renseignements Atmosphériques à la Neige*

*This has been added in the new version of the manuscript.*

Also a few more details about SAFRAN would be appreciated, for instance, what kinds of observations go into the product, and what is its spatial resolution?

*Done, see modification below*

-----NEW-----

The snowpack evolution strongly depends on near-surface meteorological forcings. These forcings are provided by the meteorological downscaling and analysis tool SAFRAN (**Système d’Analyse Fournissant des Renseignements Atmosphériques à la Neige** \citep{durand1993}). SAFRAN is used to drive snowpack simulations in the French mountains because it is designed to operate at the geographical scale of meteorologically homogeneous mountain ranges, varying from 400 to 2000km<sup>2</sup>. The model combines vertical profile estimates from the ERA-40 re-analysis with observed weather data from the automatic surface observations network at different elevations, the French Snow/Weather network, rain radars, and rain gauges. As

outputs, SAFRAN provides meteorological data to the snowpack model with an hourly time step for all slopes and aspects, and a 300 m-elevation step.

11. P. 6837, Line 9: The phrase “simulate the errors” is a bit confusing... perhaps you mean that you need to first simulate the impact of errors on the simulation of the snowpack?

*Yes that is what it means. This has been clarified in the new version of the manuscript as reported below.*

-----NEW-----

In view of assimilating observations to reduce snowpack simulation uncertainties, we first need to **represent them**.

12. P. 6837, Lines 14-16: Perhaps provide further explanation as to why these errors are not considered. I would imagine these errors are difficult to evaluate as they may vary by location and may be difficult to separate from other sources of error. Can the authors briefly comment on how their inclusion might affect the results presented?

*We follow Raleigh et al. (2015) by considering the forcings as the only source of uncertainty. Obviously, this is an approximation. We did not consider model errors because they are indeed more difficult to characterize and represent, notwithstanding their spatial variations that would be ignored here. The manpower needed to implement their effects has been spared for this first study focused on the relevance of reflectance assimilation. But of course, this question will have to be addressed with real data. This is (more briefly) stated in the new version of the manuscript, section 3.1.*

-----NEW-----

Snowpack model errors introduced by metamorphism and other parameterizations of physical laws are not taken into account here. The characterization and representation of these errors, notably in the perspective of real data assimilation, will be addressed in a future and dedicated work. An identified option is to use multi-physics ensemble simulations.

13. P. 6838, Line 3: I think it would be better to refer to RMSE rather than the “standard deviation of the difference”, for consistency with other portions of the paper.

*The standard deviation of the difference is computed by subtracting the mean, contrary to the RMSE. Since these are different quantities, we suggest to keep the text as is.*

14. P. 6838, Line 5: Does “significant” refer to statistical significance? Please clarify.

*No, actually we did not make statistical tests. Significant has been consequently removed from the sentence.*

“The table **reflects differences** between..”

*Done, See Review 1, Specific Comment 2.*

15. P. 6838, Lines 6-8: Differences can also occur because of measurement errors at CdP.

*True, this has been added in the new manuscript.*

*Done, See Review 1, Specific Comment 2.*

16. P. 6838, Line 13: Hourly interpolation of the daily analysis wasn't discussed earlier. Please elaborate.

*Done, See Review 1, Specific Comment 2.*

17. P. 6838, Lines 15-17: This sentence is unclear. What is the average RMSE or range of RMSE values at the stations?

Which study highlights the spatial variability of SAFRAN? I presume it is the Durand et al. (2009) study, but this is not clear from the sentence. How much do the RMSEs change across stations?

*All these information are provided in Durand et al., 2009. Averaged on 43 sites, the values of RMS temperature observation error range from 1° to 1.5°C over the 1959 – 2001 period.*

*But at the same time, over the whole period, the statistics between SAFRAN and in situ air temperature can be quite different from a site to another. For example, at the Lus La Croix Haute site the bias and rms are respectively, +0.2 and 1.2°C while at the barcelonnette site we note +1.1 and 5.8°C.*

What are the implications of this spatial variability for this study; i.e. can the uncertainty estimates at CdP really be used as indicators of the uncertainty at CdL?

*Regarding the above arguments, it should have been more convenient to quantify uncertainty directly at the CdL area but this site does not provide a time-series of meteorological forcing. Consequently, since the CdP site is the most instrumented area overs 20 years in the French Alps, the use of the CdP statistics was the best option to get an estimate of the meteorological uncertainty. Given that CdP and CdL are not located in the same mountain range (respectively, Chartreuse mountain range and Grandes Rousses mountain range) and at the same altitude, some adjustments were carried out on the perturbations, as explained in section 3.3.*

*The possible error introduced from this variation will not have a strong impact in this study since it is based on twin experiments. A more physically consistent ensemble should be built for future work with real data.*

*This has been added in the new version of the manuscript.*

*The Section 3.2 was rephrased and presented in Review 1, Specific Comment 2. The modifications concerning this comment are reported below:*

-----NEW-----

**\cite{durand2009} carried out, only on a limited set of variables, a more systematic evaluation of SAFRAN for the 1958-2002 period using 43 sites in the French Alps. Averaged over all locations, the RMSE on air temperature are similar to the one computed in our study. However, their results also highlight the spatial variability of SAFRAN performance (site RMSE ranges from -0.8 to +1.5 \textdegree C). Nevertheless, this will not have a strong impact in this study since it is based on twin experiments.**

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

18. P. 6839, Line 2: Mention how tau is chosen here rather than later on.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

19. P. 6839, Line 6: Some formulas should be included describing how a given variable at a given time step is perturbed (through either multiplication or addition).

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

20. P. 6839, Lines 7-8: What is meant by “the nature of the variable”? It becomes a bit clearer later on, e.g. precipitation should not be additively perturbed to avoid creating precipitation where there is none. Can the authors be more specific? The second criteria is also unclear and does not seem to be mentioned later. Please provide a more detailed description of how a method is chosen for a given variable.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

21. P. 6839, Lines 14-16: I suggest discussing how tau is adjusted when tau is introduced in the previous paragraph. More details should be provided as to how tau is chosen, i.e. how is it determined that the “temporal variation” of perturbed variables is similar to that of the original variables?

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

22. P. 6839, Lines 18-19: Suggest saying that the maximum value of shortwave radiation is set to  $200 \text{ W m}^{-2}$  for clarity. Is this Done because of the presence of clouds? Please make this clear.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

23. P. 6839, Lines 20-23: This is out of place here, and should be mentioned earlier in the previous paragraph.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

24. P. 6839, Lines 27-29: Suggest modifying this sentence for clarity: “In particular a forested area masks a portion of the shortwave radiation at the CdP site, and modifies the local wind field. The model does not account for this forested area, resulting in the large discrepancies between model and observations.”

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

25. P. 6840, Lines 1-4: The procedure for these adjustments should be provided, along with the results of the sensitivity tests if possible, in the supplementary material.

Also, the standard deviation for wind speed appears to be different. Was this also adjusted?

*Thank you to have seen this mistake.*

*This was added in the new section 3.3.*

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

26. P. 6840, Lines 5-6: Actually the standard deviations used to generate the ensemble are not provided in the left column, as they have been adjusted for the new location. I suggest adding another column showing the prescribed standard deviations for CdL. Also perhaps this is best mentioned after mentioning the use of an ensemble in the next paragraph.

*Instead adding a new column, we just add the adjusted standard deviation used to generate the ensemble in brackets for the variables for each the standard deviation was adjusted. This is now detailed in the new version of Table 1 legend.*

Variables	CdP: Reanalysis - Observations		CdL: Reanalysis - Ensemble	
	Bias	std: $\sigma_{CdP}$	Bias	std: $\sigma_{CdL}$
Air temperature (C)	0.28	1.08	5.0 e-03	1.07
Wind speed (m s <sup>-1</sup> )	0.2	1.12 (0.6)	4.0 e-04	0.4
Shortwave radiation (W m <sup>-2</sup> )	22.4	79 (70)	-3.1 e-03	58.3
Longwave radiation (W m <sup>-2</sup> )	-14.0	24.5 (7)	2.0 e-02	7.0
Snowfall rate (kg m <sup>-2</sup> h <sup>-1</sup> )	-2.0 e-02	0.4	5.0 e-03	0.1
Rainfall rate (kg m <sup>-2</sup> h <sup>-1</sup> )	7.2 e-03	0.5	-5.0 e-03	0.1

**Table 1.** Bias and standard deviations (std) of the differences between SAFRAN reanalysis and *in-situ* observations (left) and the differences between SAFRAN reanalysis and the ensemble built up in the present study (right), for the perturbed meteorological forcings. The first set of statistics is derived from 18 years (1993-2011) of observations and reanalysis at the CdP and the second set is derived from our 300-members ensemble over the 2010/2011 hydrological season. The values in brackets correspond to the adjusted standard deviation used to generate the ensemble at CdL site.

27. P. 6840, Lines 12-13: It would be helpful to remind the reader here how these are taken into account in this case.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

28. P. 6840, Lines 13-14: Explain why it is not crucial to account for intervariable correlations for the purposes of this experiment.

*Comments 18 to 28: See Review 1, Specific Comments 3 & 4.*

*The whole section 3.3 was modified taking into account these comments.*

29. P. 6841, Lines 1-3: Are the impurity concentrations consistent with any previous measurements of snow impurity content?

*The dry deposition flux varies from 0 to 0.5 ng g<sup>-1</sup> s<sup>-1</sup> with a median value is 0.015 ng g<sup>-1</sup> s<sup>-1</sup> in concentration which roughly corresponds to 3.5 kg m<sup>2</sup> s<sup>-1</sup>. This is the same order of magnitude than the dry deposition rates obtained for this location using ALADIN-Climat (Nabat et al., 2015) simulations. We however prefer not to include this discussion in the paper since the comparison required more detailed analysis and field measurements and since it not the main objective of our study.*



30. P. 6841, Lines 8-13: The details of the figure could be shortened somewhat since they are already mentioned in the figure caption.

*Done*

-----NEW-----

Figure \ref{run\_init} presents **the result of the ensemble simulation with 300 members (represented by the black lines).**

The simulation forced by the unperturbed reanalysis (**red line**) is included within the envelope of the ensemble. The spread of the ensemble reflects the consequences of possible overestimations and underestimations of meteorological data by the reanalysis.

31. P. 6841, Lines 17-18: Change “dispersion range ( $\Delta$  SWE  $\approx$  300 kg m<sup>-2</sup>)” to “dispersion range of SWE ( $\Delta$  SWE  $\approx$  300 kg m<sup>-2</sup>)”

*Done*

32. P. 6841, Lines 18-19: It would be better to refer to the snowpack here: “The snowpack in some ensemble members has just started to melt, while in other cases, the snowpack has already disappeared.”

*Done*

33. P. 6842, Line 11: It seems that the RMSE is calculated as a RMSE for modeled vs. in situ snow depths across multiple stations. Please clarify what the RMSE refers to here.

*See Review 2, General Comment 5.*

34. P. 6842, Line 14: The letter n was used in the equation for Spd as an indicator of the ensemble number. Here, I believe it refers to neither time, nor ensemble member, but to a station identifier, and M is the total number of stations, at a given time. Please clarify and change the notation to avoid confusion.

*See Review 2, General Comment 5.*

35. P. 6842, Line 18: It is not clear what the “reference” is at this point. Perhaps replace with “a synthetic true reference simulation”

*See Review 2, General Comment 5.*

36. P. 6842, Lines 19-21: Since observations are not mentioned yet, perhaps this should be moved to the next paragraph.

*See Review 2, General Comment 5.*

37. P. 6843, Line 11: The term “twin” has not been defined yet. It would be helpful to define it here rather than later on for those readers not familiar with the terminology.

*A description of twin experiment was added in the introduction section.*

*See Review 1, Specific Comment 1*

38. P. 6843, Line 19: Perhaps change “more physics details” to “further details”.

*Done*

39. P. 6843, Line 22: Again I think it would be better to mention this in section 3.6.

*The aim of section 3.6 is only to evaluate the dispersion of the ensemble from the stochastic perturbation method. We prefer keeping this sentence with its idea at the beginning of the section 4.1 General settings and diagnostics.*

40. P. 6843, Line 23: But the synthetic observations come from the same model into which data are being assimilated. The model is not independent. Please revise, and clarify that the synthetic observations come from a single Crocus ensemble member.

*Done, see modification below*

-----**NEW**-----

The assimilation experiments are twin, meaning that the observations are synthetic and come **from a single model simulation**. They are performed over five winter seasons at the CdL area.

41. P. 6844, Line 1: I suggest changing “control simulation” to “synthetic truth simulation” for clarity, here and in other places where it is mentioned.

*Done*

42. P. 6844, Line 2: Perhaps change “virtual observations” to “synthetic observations” for clarity.

*Done*

43. P. 6844, Lines 2-4: It may be helpful to mention the virtual or synthetic observations extracted from the simulations. In particular it would be helpful to mention (for understanding the following section) that synthetic reflectance observations are obtained from Crocus, through the use of the TARTES model.

-----**NEW**-----

A synthetic truth simulation is first obtained running Crocus, **through the use of the radiative transfer model TARTES**, forced by one perturbed meteorological forcing, as detailed in Section \ref{sec\_pert}.

*Done*

44. P. 6844, Line 7: Change “truth” to “synthetic truth” for clarity.

*Done*

45. P. 6844, Lines 6-8: Suggested change to the sentence: “Data assimilation performances are evaluated by comparing RMSE for ensembles with and without assimilation, and by comparing the synthetic true simulation to the 33<sup>rd</sup>, 50<sup>th</sup>, and 67<sup>th</sup> quantiles from the ensembles with assimilation.”

*Done*

46. P. 6844, Lines 16-18: Mention that these are derived using TARTES in Crocus.

-----**NEW**-----

The first set of **synthetic** observations is composed of surface reflectances of the first seven bands of MODIS (central wavelengths: 460, 560, 640, 860, 1240, 1640, 2120 nm;

\citealp{hall2007}). These synthetic observations are extracted from the synthetic truth simulations, using the radiative transfer model TARTES in Crocus.

47. P. 6845, Lines 5-6: I suggest moving this sentence to the end of the paragraph.

-----NEW-----

The setup designed in our study (one point, twin experiments) allows relevant comparisons of the benefits of assimilating separately or jointly the two above mentioned types of observations. **For future works assimilating real data, the difference in the geometrical configuration between the simulated TARTES reflectances and satellite observations will be addressed.**

48. P. 6846, Lines 24-25: The sentence is unclear. Suggested change: "Inversely, a new perturbed forcing is assigned to a duplicated particle for propagation to the next analysis."

*Done*

49. P. 6847, Line 7: How are these "clear sky days" chosen? Are these based on the real measurements at CdL, or are they from real MODIS data? Please clarify.

*Actually we use real MODIS data to do it. As other possible variables (albedo, NDSI, reflectance..) a cloud mask can be estimated from MODIS radiance data using the algorithm named MODimLAB (Sirguey et al., 2009).*

-----NEW-----

To mimic real cloud conditions, reflectances are assimilated at 34 clear sky days of the season. **We define a clear sky date according to the real cloud mask from MODIS data computed with the method of Sirguey et al. (2009).**

50. P. 6847, Lines 8-9: Are real MODIS data used in this case? If so, please provide details about MODIS data in an earlier section. If not, I think this sentence can be removed, because this study does not make use of real MODIS data.

*See response to the previous comment.*

51. P. 6847, Line 13: Clarify that the envelopes are the envelopes from SD and SWE ensembles for the baseline experiment (e.g. "SD and SWE ensembles for the baseline experiment...")

*Done*

52. P. 6847, Lines 17-18: Suggest changing sentence to: "This is observed in Fig. 3, where the baseline experiment envelopes (blue shading) are narrower than those of the ensemble without assimilation (grey shading)."

*Done*

53. P. 6847, Lines 21-23: Mention the terms "forecast" and "analysis" to be consistent with Figure 4.

*Fig 4, RMSE curves: "(Blue solid line: forecast; blue dotted line: analysis)".*

-----NEW-----

Figure \ref{RMSE\_8009} shows the time evolution of the RMSE with assimilation at every observation time, **at the end of the forecast step (blue solid line)** and just after the filter analysis (**blue dotted line**). These results are compared to the RMSE without assimilation (grey lines).

54. P. 6848, Line 1: What is meant by “the continuous flow of observations”? I think the authors mean to indicate that the observations are well distributed in time. Please clarify.

*Yes, see modification below*

-----NEW-----

The reduced RMSEs with assimilation are consequently due **to the successive observations throughout the season**, highlighting the role of model dynamics.

55. P. 6848, Lines 4-5: Change “extended and unobserved periods without precipitation” to “extended periods without precipitation and without available observations”

*Done*

56. P. 6848, Line 11: Which figure should we refer to for the example on 28 January?

*Done*

-----NEW-----

(for example, on 28/01/2011 on **Fig. 3**)

57. P. 6848, Line 19: Change “the whole ensemble to a unique set” to “all ensemble members to the same set”

*Done*

58. P. 6848, Line 20: Change “discrimination between members impossible only from the reflectances” to “discrimination between members using reflectances alone impossible”

*Done*

59. P. 6848, Line 21: Change “the analysis” to “the subsequent analysis”

*Done*

60. P. 6848, Line 22: Again, which figure should be examined here?

*Fig 3, Done*

61. P. 6848, Line 28: Change “at the end” to “towards the end”.

*Done*

62. P. 6848, Lines 2-4: This statement is not really supported by the previous analysis, although it is shown that this is important later. Either note that this will be shown later, or remove this sentence.

*Done*

*Ok, this has been reformulated as follows:*

-----NEW-----

The strongest RMSE reductions occur right after extended periods without precipitation and without available observations, when the reflectance ensemble spread is particularly

**pronounced (e.g. Fig 3a).** During these periods (e.g. from 07/12/2010 to 14/12/2010, or from 11/01/2011 to 21/01/2011), the ensemble uncertainties on reflectances, SD and SWE grow under the influence of the perturbed forcings including the perturbed impurity deposition rate.

63. P. 6849, Line 8: Suggest changing “impact of the limited number of available data” to “impact of limiting the number of available observations” for clarity.

*Done*

64. P. 6849, Lines 9-10: Change “first one” to “baseline experiment” for clarity. *Done*  
Change “...carried out but assimilating an observation” to “...carried out, but assimilation is performed...”.

*Done*

65. P. 6849, Line 15: As the envelope for the experiment is wider than the reference experiment, I think it is an overstatement to say that the fit is “perfect”.

*Done*

-----NEW-----

Obviously, in this second experiment, **concerning the 640-nm reflectance variable, the spread of the ensemble is greatly reduced, well fitting the observations (red dots) and its envelope does not show any extended periods with a large range of reflectance values anymore (Fig. \ref{Supp-FP\_8017} a).**

66. P. 6849, Line 23: Change “that the limited number” to “that assimilating a limited number” for clarity.

*Done*

67. P. 6850, Line 7-10: Is it possible to give each of these simulations a meaningful name? This would be very helpful when references are made to each simulation in the paper. Otherwise the reader forgets the details of the simulation. If the authors think it would be more confusing to name each simulation, the details should be briefly mentioned when describing the results, e.g. “In case (i), where assimilation occurs only at the beginning of the season, results show...”

*Done, see modification below*

*(Table 2, was modified accordingly. See Review1 Table and Figure Comments)*

-----NEW-----

The baseline experiment suggests that the timing of observations may largely determine the quality of the assimilation process. To explore the role of the timing, four additional assimilation tests are designed for which MODIS-like reflectances are assimilated (i) only at the beginning of the season (before 31/12/2010, Fig. \ref{Supp-FP\_8018}: **Accu**), (ii) only in the second part of the snow season (after 31/12/2010, Fig. \ref{Supp-FP\_8019}: **Melt**), (iii) only after several day-long periods without precipitations (Fig. \ref{Supp-FP\_8020}: **Before Snowf**) and (iv) only right after snowfalls (Fig. \ref{Supp-FP\_8026}: **After Snowf**).

In case (i: **Accu**), results show that even if the SD and SWE spreads are reduced during the assimilation period, the assimilation has almost no effect on the snow estimates during the snow melt. The ensemble spread retrieves to almost the same value than the experiment without assimilation. The uncertainty of the snow melt-out date is reduced to 22 days only, in comparison with 24 days without assimilation. As for case (ii: **Melt**), the spread reduction becomes quite discernible roughly 2 months after the first assimilation date and never reaches the value of the baseline experiment. The uncertainty of the snow melt-out date is however reduced to 11 days. This demonstrates that it is essential to assimilate reflectances over the entire season to compensate the fast growth of the snowpack ensemble in response to the uncertainties in the meteorological forcing.

In both cases (iii: **Before Snowf**) and (iv: **After Snowf**), reflectances are assimilated at only 7 dates of the season. Case (iii: **Before Snowf**) exhibits a more pronounced SD and SWE spreads reduction compared to case (iv: **After Snowf**). The uncertainty on the snow melt-out date drops to 9 days in case (iii: **Before Snowf**) while it stays at 23 days in case (iv: **After Snowf**). In absence of precipitation, the snow surface is aging, leading to a decrease of reflectance values and a spread of the reflectance ensemble (Fig. \ref{Supp-FP\_8020} a).

68. P. 6850, Line 13: Change “ensemble spread retrieves” to “ensemble spread at the end of the season returns”

*Done*

-----NEW-----

The ensemble spread at the end of the season returns almost the same value as the experiment without assimilation.

69. P. 6851, Lines 20-21: This sentence seems a bit out of place. Perhaps it can be moved to the end of the paragraph and expanded on a bit.

We followed your suggestion and added a new paragraph at the end of the section.

-----NEW-----

Figure \ref{RMSE\_5seasons} also shows that, all these findings obtained for the 2010/2011 season are also verified for the five studied seasons. All assimilation experiments of synthetic SD observations reduce the RMSE with respect to both the model run without assimilation (red lines) and the experiments assimilation synthetic reflectances data (blue lines). However, in case of shallow snowpack, the better performance is obtained using reflectance data.

70. P. 6852, Line 5: The snowpack is probably also more sensitive to absorbed solar radiation. We are unsure what the reviewer is meaning here. In case of thin snowpack more solar radiation may be absorbed by the soil. No change

71. P. 6852, Lines 13-14: Although this seems likely, I’m not sure what evidence from the experiments that were Done supports this statement.

*This statement is supported by the fact that every SD observations assimilation reduces the envelops independently of the precipitation events and which is not the case for reflectance (Fig S7). This is now detailed in the modified version below.*

-----NEW-----

Excepted for thin snow cover, the assimilation of SD observations outperforms reflectance assimilation in terms of SWE and SD estimates and seems to be less affected by the time distribution of the observations. **When assimilating reflectance data, the ensemble needs to sufficiently spread (from an extended period without precipitation) to observe an impact of the assimilation (Fig 3a).** Inversely, and even if that may be very small, every SD observations assimilation reduces the SD ensemble independently of the precipitation events (Fig S7b, excepted for thin snow cover).

72. P. 6852, Line 25: Suggest changing “punctual usage in time” to “low temporal frequency”  
*Done*

73. P. 6852, Line 3: It would be nice to also have an additional supplemental figure showing the impact of including both snow depth and reflectance for the 2010/11 season.

*This figure was added in the supplement part (Fig S8).*

*We modified accordingly the section 6.4 “combining reflectance and snow depth assimilation” (See review 1, General Comment 1) as well as the Table 2 (See Review1, table and figure comments)*

74. P. 6853, Line 15: Could the use of a high spatial resolution make assimilation more useful?  
*Yes indeed.*

-----NEW-----

However, given the strong spatial variability of the snow cover, the spatial representativity of **punctual SD** measurements may make their assimilation questionable. This issue should be addressed with experiments over two-dimensional, realistic domains.

75. P. 6854, Line 7: Change “improves” to “reduces”

*Done*

76. P. 6854, Line 15: Change “provides results almost as good” to “reduces RMSE almost as much as”

*Done*

77. Table 1: Suggest changing “Standard Deviations” to “RMSE” for consistency with other parts of the paper.

*See review 2, specific comments 13*

78. Table 1, caption: Mention the range of years for the 18 years of observations from CdP.

-----NEW-----

Means and standard deviations (std) of the differences between SAFRAN reanalysis and `\textit{in-situ}` observations (left) and the differences between SAFRAN reanalysis and the ensemble built up in the present study (right), for the perturbed meteorological forcings. The first set of statistics is derived from 18 years (**1993-2011**) of observations and reanalysis at the CdP and the second set is derived from our 300-members ensemble over the 2010/2011 hydrological season.

*Done*

79. Table 2: I think having names for simulations would be more useful than including figure numbers here, or both names and figure numbers could be included. Simulations could be given meaningful names, or referred to as “case (i)”, etc. as discussed in the text.

*See review2 specific comment 67*

80. Table 2, caption: What does “Seasonal” refer to? The 2010/2011 season, or the period when snow cover exists for all seasons?

*Modified as follows:*

**SD and SWE seasonal averaged RMSE computed with respect to the synthetic truth for all experiments.**

81. Figure 1, caption: Change “band 1 of MODIS” to “center of band 1 of MODIS” for clarity. Define SD and SWE.

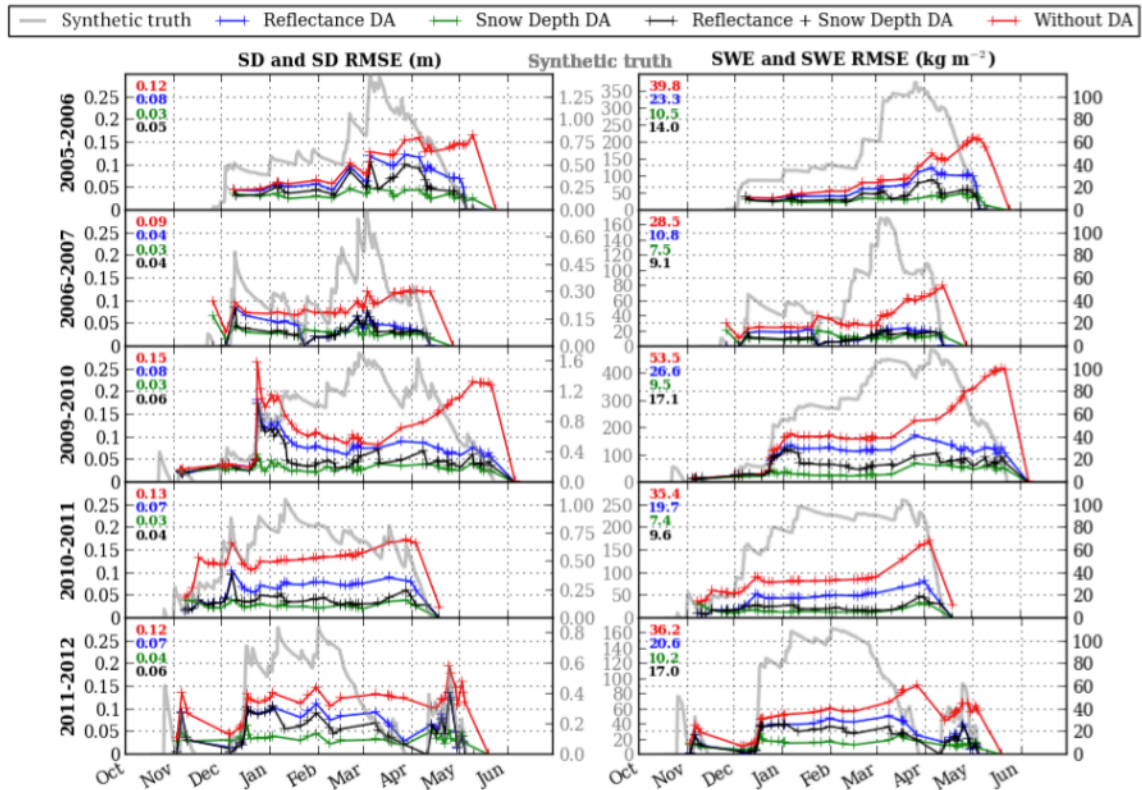
*Done*

82. Figure 3, caption: Change “patterns” to “shading”, when describing the envelope colors. Clarify whether the quantiles are for the baseline experiment or ensemble without assimilation.

*Done*

83. Figure 5: The figure is initially difficult to understand. “Model control” should be replaced by “Synthetic true snow depth”. Change the left and right titles to “SD and SD RMSE (m)” and “SWE and SWE RMSE (kg m<sup>-2</sup>)” for clarity.





**Figure 5.** Time evolution of ensemble RMSEs on SD (left) and SWE (right) for the five seasons under study, for the run without assimilation (red lines), the baseline experiment (assimilating reflectances, blue lines), the experiment assimilating SD data (green lines) and the experiment assimilating combined reflectances and SD data (black lines). Crosses indicate analysis steps. Seasonal averages are displayed in the upper left corner of each graph. The model control simulation is represented by the grey lines, scaled by the "Synthetic truth" y-axes.

84. Figure 5, caption: Define "seasonal means".

Done

Seasonal averages

### Technical Corrections:

*All these technical corrections have been accounted for and included in the new version of the manuscript. Thank you for the very detailed proofreading.*

1. Change "envelop" and "envelops" to "envelope" and "envelopes" throughout.
2. P. 6832, Line 3: Change "active microwave" to "active microwave measurements"
3. P. 6832, Line 13: Change "on board TERRA..." to "onboard the TERRA..."

4. P. 6832, Line 28: Change “snow simulations” to “snow simulation”
5. P. 6833, Line 2: Change “Moreover, 2...” to “Two...”
6. P. 6833, Line 3: Change “serie” to “series”
7. P. 6833, Line 4: Change “Indeed, the Col de Porte...” to “The Col de Porte...”
8. P. 6833, Line 6: Change “statistic” to “statistics”
9. P. 6834, Line 8: Change “of the avalanche” to “of avalanche”
10. P. 6835, Line 5: Change “its maximum” to “a maximum”
11. P. 6837, Line 15: Change “other physical laws parametrization” to “other parameterizations of physical laws”
12. P. 6837, Line 20: I believe “present section” should be changed to “following section”.
13. P. 6839, Line 2: The exponent in the expression for  $\phi$  should be raised.
14. P. 6839, Line 10: Change “longwave radiations” to “longwave radiation”.
15. P. 6839, Line 13: Change “a week period” to “a one week period”.
16. P. 6840, Lines 2 and 3: Change “longwave radiations” to “longwave radiation”.
17. P. 6840, Line 11: Change “inter-variables” to “inter-variable”
18. P. 6840, Line 14: Change “Real data assimilation...” to “A real data assimilation...”
19. P. 6840, Line 19: Change “is not” to “are not”
20. P. 6840, Line 20: Change “their dry” to “the dry”
21. P. 6841, Line 17: Change “24 days spread” to “a 24 day spread”
22. P. 6842, Line 4: Change “ensemble of simulation” to “ensemble”
23. P. 6842, Line 21: The word “consequently” can be removed.
24. P. 6842, Line 22: Remove “In that purpose,”
25. P. 6843, Line 2: Change “than the CdL” to “as the CdL”
26. P. 6843, Line 7: Change “perturbations calibration” to “calibration of perturbations”
27. P. 6843, Line 9: Change “measurements sites” to “measurement sites”
28. P. 6843, Line 10: Change “model error” to “model errors”
29. P. 6843, Line 13: Add comma after “not crucial for our study”
30. P. 6843, Line 16: Change “prior” to “prior to”
31. P. 6843, Line 18: Change “observations datasets” to “observational datasets”
32. P. 6844, Line 1: Change “obtained” to “obtained by”
33. P. 6844, Line 12: Change “RMSE” to “RMSEs”
34. P. 6844, Line 20: Change “they are mainly varying” to “they mainly vary”
35. P. 6845, Line 4: Change “measurements provides” to “measurements provide”
36. P. 6845, Line 6: Change “later” to “latter”
37. P. 6845, Line 13: Change “to 0.003 m” to “to be 0.003 m”
38. P. 6846, Line 1: Change “simplest” to “simpler”.
39. P. 6846, Line 12: Suggest changing “particular flavor of the particle filter” to “particular type of particle filter”
40. P. 6846, Line 19: Change “distances to” to “distances from”
41. P. 6846, Line 27: Change “or” to “and”.
42. P. 6847, Line 13: Change “All along the season” to “Throughout the season”
43. P. 6847, Line 25: Change “RSME SD” to “RMSE for SD”

44. P. 6848, Line 8: Change “poorly” to “not very”
45. P. 6848, Line 17: Change “On top of this” to “Moreover”
46. P. 6848, Line 22: Remove “the” before dates.
47. P. 6848, Line 27: Change “coarsely” to “roughly”
48. P. 6848, Line 2: Change “meets limitations” to “exhibits limitations”
49. P. 6849, Line 6: Change “clouds coverage” to “cloud coverage”
50. P. 6849, Lines 12-13: Change “patterns” to “shading”
51. P. 6849, Line 18: Change “uncertainties” to “uncertainty”
52. P. 6849, Line 24: Change “are not” to “is not”
53. P. 6850, Line 9: Change “precipitations” to “precipitation”
54. P. 6850, Line 10: Change “snowfalls” to “snowfall events”
55. P. 6850, Line 13: Change “the snow melt” to “the snow melt period”
56. P. 6850, Line 24: Change “stays to 23 days” to “stays at 23 days”
57. P. 6851, Line 20: Change “all along” to “throughout”
58. P. 6852, Line 12: Change “Excepted” to “Except”
59. P. 6852, Line 16: Change “a ‘surface’ information only” to “ ‘surface’ information only”
60. P. 6852, Line 20: Change “spatialized” to “spatially distributed”
61. P. 6853, Line 13: Change “encourage to combine” to “indicate the usefulness of combining”
62. P. 6853, Line 22: Change “estimate uncertainties” to “uncertainty estimates”
63. P. 6854, Line 26: Change “kind of data assimilation” to “kinds of data assimilation”
64. P. 6855, Line 12: I believe all equations in this section should be given a number.
65. P. 6855, Line 13: I believe these equations come primarily from Gordon et al. (1993) and this reference should be referred to here.
66. P. 6856, Line 4: I believe  $X_{k-1}$  just to the right of the integration sign should be  $X_k$
67. P. 6856, Line 5: Change “Bayes’rule” to “Bayes’ rule”
68. P. 6857, Line 8: Please provide a reference for “Kitagawa”.