Internship Report

Estimating the 3D structure of boundary-layer clouds from point measurements of solar irradiance

Intern :
Benoît Bielooseroff (ISAE SUPAERO)

Mentors :
Quentin Libois (CNRM)
Fleur Couvreux (CNRM)

Gap year internship : from 03/01/21 to 08/29/2021

Centre National de Recherches Météorologiques (CNRM)

Acknowledgments

I would like to thank my two mentors Quentin LIBOIS and Fleur COUVREUX for giving me the opportunity to work on this internship, for the time they dedicated to support me during the internship, and for sharing their knowledge. They gave me enough independence during those 6 months for me to thrive and were always available for answering my questions and following my work.

I am grateful to Didier RICARD for facilitating my arrival in the team, and always making sure the internship was going well.

I would also like to thank Najda VILLEFRANQUE for the help she provided me with the use of the High-Tune-Renderer tool.

Finally, I would like to express my appreciation to all the members of the phynh and precip teams I worked with and saw daily, and particularly Marc MANDEMENT and Marie-Adèle MAGNALDO for the help and useful tips they provided me.

Contents

Chapter 1

Introduction

The study of clouds is something very important when studying the atmospheric physics, as clouds affect earth climate, can lead to precipitation and of course modify the amount of sun radiation received at the earth's surface. Being able to measure clouds macro-physical properties would lead to a better understanding of cloud processes, and give a way to compare the already existing cloud Large Eddy Simulation with the clouds they aim to reproduce, and therefore, validate or invalidate the used models.

Macro-physical properties of cloud fields are often hard to obtain from measurements as some variables, such as liquid water path and geometrical thickness cannot be directly measured. On the contrary, solar surface irradiance can easily be measured using pyranometers. Since clouds interact with solar radiation, some of these cloud macro-physical properties may be deduced from solar surface irradiance measurements, and investigating such possibility was the aim of this internship. To test this idea, Large Eddy Simulations (LES) are used, as they fully resolve cloud processes, and can therefore be considered as realistic representations of cloud fields. From the LES simulations, the interesting 3D macro-physical variables can be computed, and the LES output are also used as the input of the High-Tune-Renderer (htrdr) 3D radiative transfer tool that computes the surface radiation accounting for cloud properties, atmospheric properties, sun position and ground characteristics [8].

First we identify the relevant 3D parameters that both represent the 3D properties of the could field and have a significant effect on the solar surface irradiance (Chapter 2). Those parameters will be selected in order to then train a neural network (Chapter 3) aimed at retrieving them from the solar surface irradiance's probability density function, obtained from the entire surface flux map coming from the htrdr simulations.

Chapter 2

Impact of cloud physical properties on the distribution of surface solar irradiance

2.1 Presentation of Htrdr

Htrdr is a 3D radiative transfer tool that computes the solar surface radiation using atmospheric properties, ground geometry and a cloud field. It can be used to render images (Figure 2.1) or flux maps (Figure 2.2). I will mainly be using this tool in its flux map rendering configuration. Htrdr is based on a Monte Carlo algorithm, consequently a sample parameter is asked, representing the number of photons that will be launched during the computation (the result of the computation will be the average result of the different launch). The higher the sample number is, the more accurate the simulation will be, but the more computational time cost there is. Therefore I had to do preliminary tests to determine a sample value that is high enough to give accurate results, but that doesn't induce a too long computational time. For all simulations in this report, the used sample will be 15000, which gives a standard deviation of 3 W/m^2 for the flux values. For the entire study, the ground will be plane, and cloud droplets have an effective radius of 10 μm.

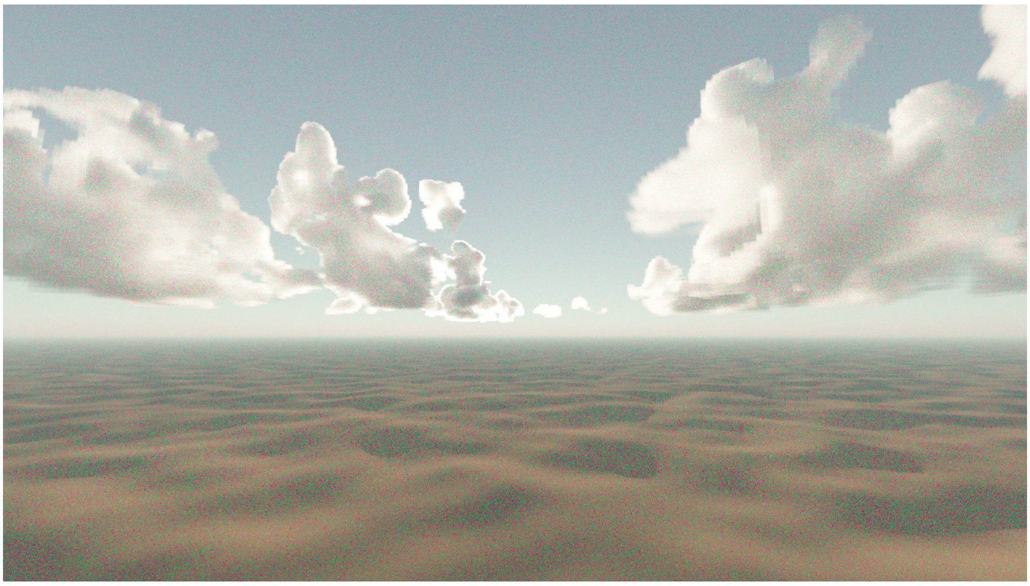
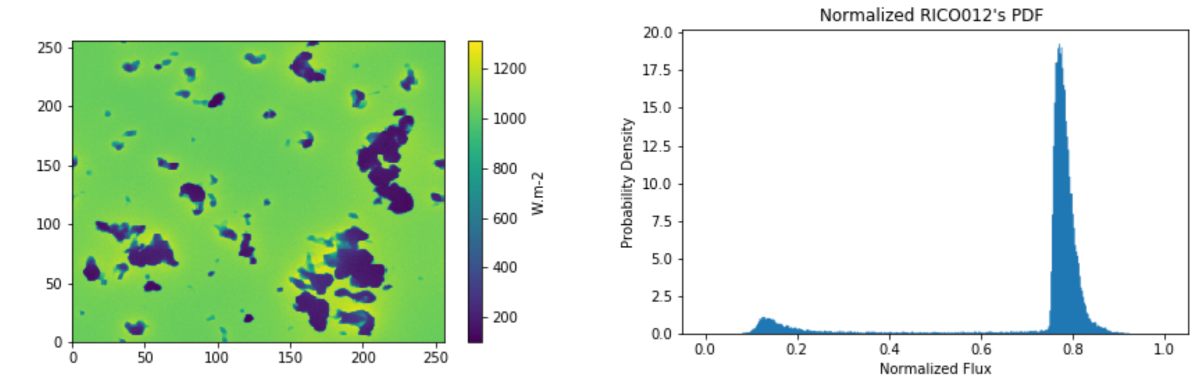


Figure 2.1: Image obtained with Htrdr

2.2 Probability Density Function (PDF) of the solar surface irradiance

Using htrdr it is possible to create a surface flux map representing the radiation from the Sun that reaches the ground (2.2.a). These flux maps can then be turned into PDF of the flux spatial distribution as shown in 2.2.b. These PDFs usually have the same shape in the presence of clouds: they have two peaks, one corresponding to pixels under clouds shadows, and one corresponding to pixels in between clouds shadows [3]. Both peaks appear to have a tail toward high values. This is one of the results of what we call the 3D effects that will be explained in a further section.

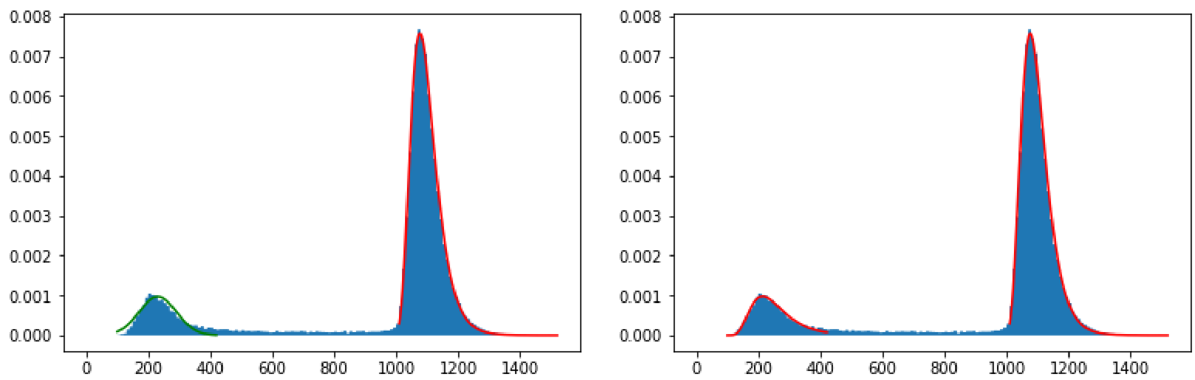


(a) Flux map under a RICO cloud field (b) Flux pdf under a RICO cloud field

Figure 2.2

As it was studied in [2], the peaks seem to be respectively following Gaussian and 3 parameters log-normal distributions, but I found that using 3 parameter log-normal distribution (Equation 2.1) for both irradiance modes better fitted the curves (2.3).

$$f(x) = \frac{1}{(x-\theta)s\sqrt{2\pi}} \exp\left(-\frac{ln^2(\frac{(x-\theta)}{m})}{2s^2}\right). \qquad (2.1)$$



(a) (b)

Figure 2.3: PDF under a RICO cloud field, with the fits of normal (green) and log-normal (red) distributions

Therefore, 6 parameters are expected to be enough to characterize the entire PDF (3 parameters for each irradiance mode, θ a location parameter, s a shape parameter and m a scale parameter).

2.3 3D effects and consequences on the PDF

When solar radiation travels across the atmosphere, there are 2 main physical phenomena that can happen: absorption and scattering. Because of the atmospheric absorption, the closer you are to the ground (and the larger the distance the radiation traveled in the atmosphere), the lower the received flux, depending on the molecules the radiation will meet and on the wavelength of the radiation. Scattering depends on the wavelength too, and it's a phenomenon that alters the direction of the radiation.
Those two phenomena also happen when radiation crosses a cloud. First, the radiation can be absorbed or reflected to space, which is why the solar irradiance is generally lower in the shadow of a cloud than under a clear sky. But because of scattering, radiation crossing the clouds does not escape the cloud with the same direction it had when entering it (See figure 2.6). And clouds scatter way more than the atmosphere. Thus, for a sensor on the ground, clouds may appear brighter than the blue sky when not observing in the sun direction since cloud scattering is higher than atmospheric scattering, under an optical thickness threshold (which I wasn't able to precisely found due to computational time restriction, but which is higher than 240). The optical thickness τ is defined from the Liquid Water Path (LWP) and the effective radius (r_{eff}) (See Equation 2.2). Consequently, when there are clouds in the sky, radiation around clouds is enhanced compared to a clear sky situation [3], because the received flux involves both the direct flux coming from the sun above, and the diffused flux coming from the clouds around. This enhancement is what is called 3D effects (Figure 2.4, [6]). Those 3D effects can easily be observed, by juxtaposing the pdf from a simulation with clouds and a simulation without clouds (Figure 2.5).

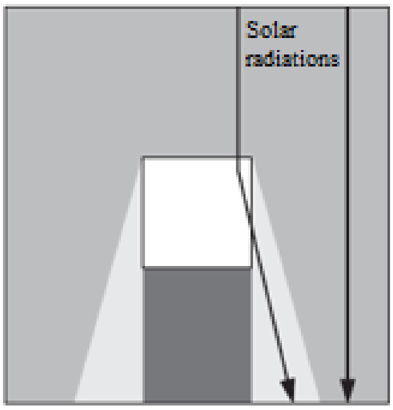


Figure 2.4: Flux enhancement near clouds with 3D radiation simulation

$$\tau = \frac{3LWP}{2\rho_{\text{water}} r_{\text{eff}}} \tag{2.2}$$

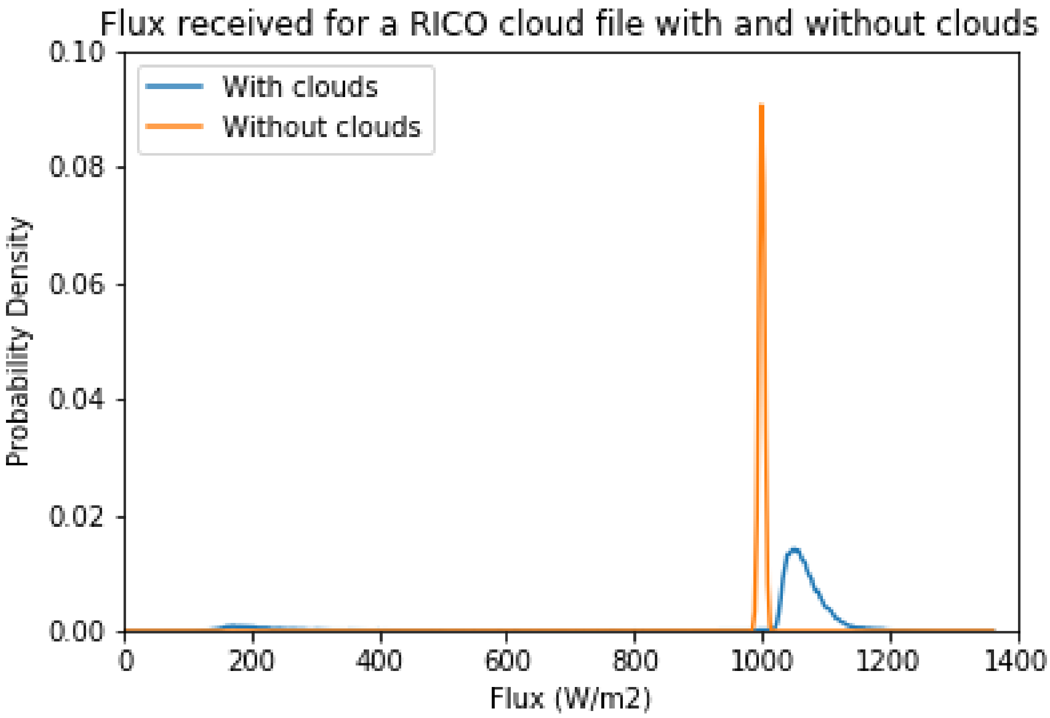


Figure 2.5: Pdf with and without clouds

Without 3D effects, we would expect both pdf to have a large irradiance mode at the same flux values. The only differences would be that for the case with clouds, there would be a first small irradiance mode, resulting from the reduced flux under the cloud, and that the large irradiance mode wouldn't have the same probability values. But in figure 2.5, we can see that the large irradiance mode is both shifted to the right and has a tail towards high flux values, which is a consequence of the 3D effects, enhancing received flux around clouds.

After noticing that the flux is increased around the clouds, the question is how is this enhancement related to the distance with the clouds. In order to study that, I created a theoretical cloud field, with only one single cloud. This cloud is a 30x30 pixels square shaped (with each pixel being themselves squares whose side size are 25 meter wide), has its base 1000 meters above the ground, and has a constant Liquid Water Path (LWP) of 0.6 kg/m^2, split into 10 layers. Then, I simulated the solar surface irradiance in one band going through the cloud in 3D and 1D and then plotted the difference next and under the clouds (Figure 2.7).
For 1D simulation I used the Independent Column Approximation (ICA), I take each column of the original cloud field and create a new, uniform cloud field with this column. After that I compute the radiation simulation with htrdr, which gives me the received surface flux for the column. In that way, each photon that is in the column at the top

of the atmosphere, will either be absorbed, be reflected to space, or reach the ground in that same column (and cannot be diffused to the neighbour columns) (Figure 2.6 from [7]). Consequently, the simulation for every column of the original cloud field is independent.

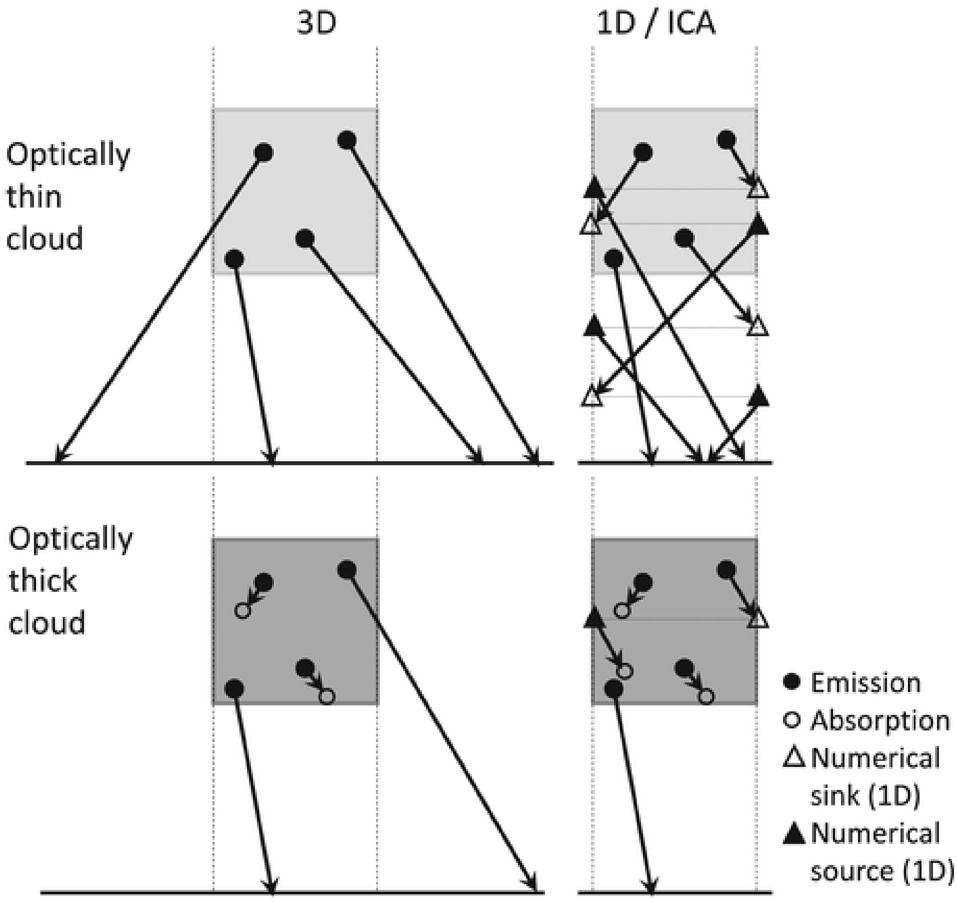
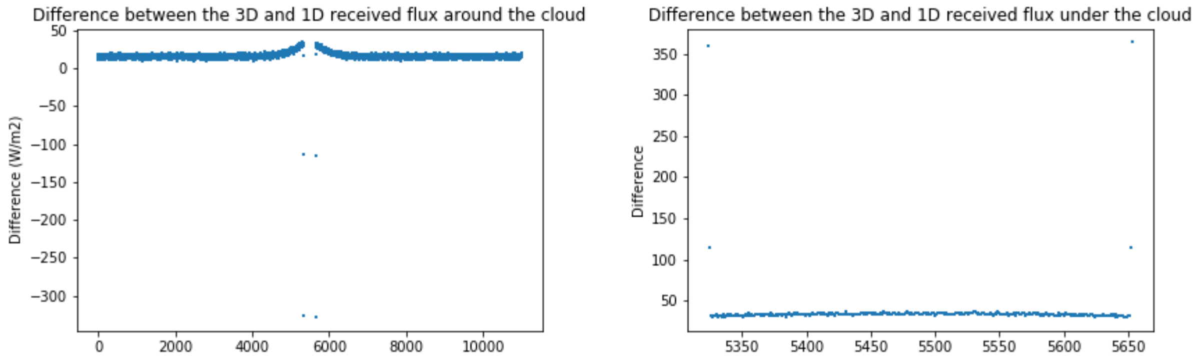


Figure 2.6: Difference between 1D and 3D radiation computation

What is remarkable, is that the flux received is enhanced, not only next to the clouds, but in the entire domain (the edge of the plot are around 25 km away from the cloud). We can also see that the 3D effects decrease really fast with the distance in the area close to the cloud. Under the cloud, the flux is enhanced too, but what is remarkable, is that the enhancement is highest in the center of the cloud (Figure 2.7). As explained in the beginning of this section, clouds appear brighter than the blue sky when not looking in the direction of the sun. And when you are at the center of the cloud, the solid angle representing the cloud is bigger than the one when you are at the edge of the cloud. This is why the received flux is more important at the center of the cloud. This means, the bigger a cloud will be, the more flux overflow there will be in the center of the cloud. Consequently, the average size of the clouds in the cloud field, might be an interesting parameter, since it is a parameter characterizing the cloud field that can be deduced from

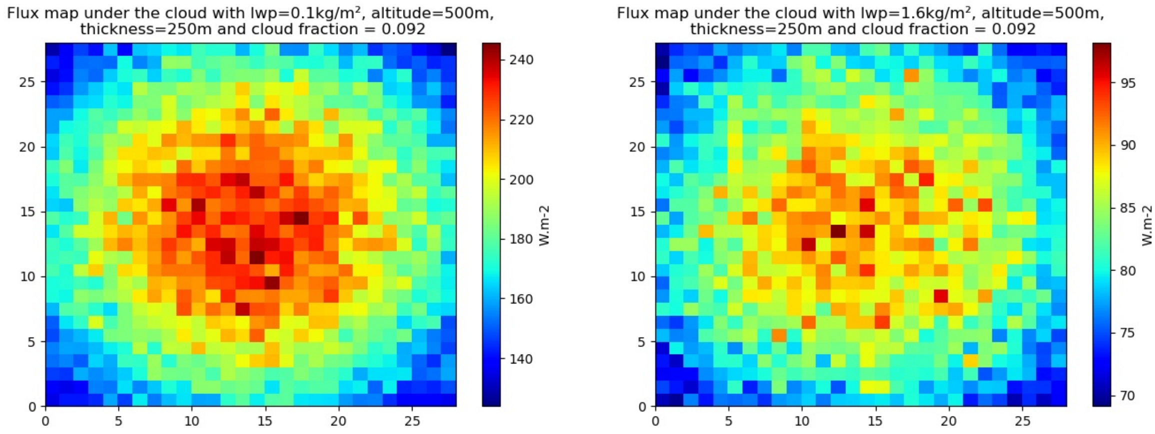


(a) (b)

Figure 2.7: Differences between 3D and 1D simulation, around (a) and under (b) the cloud

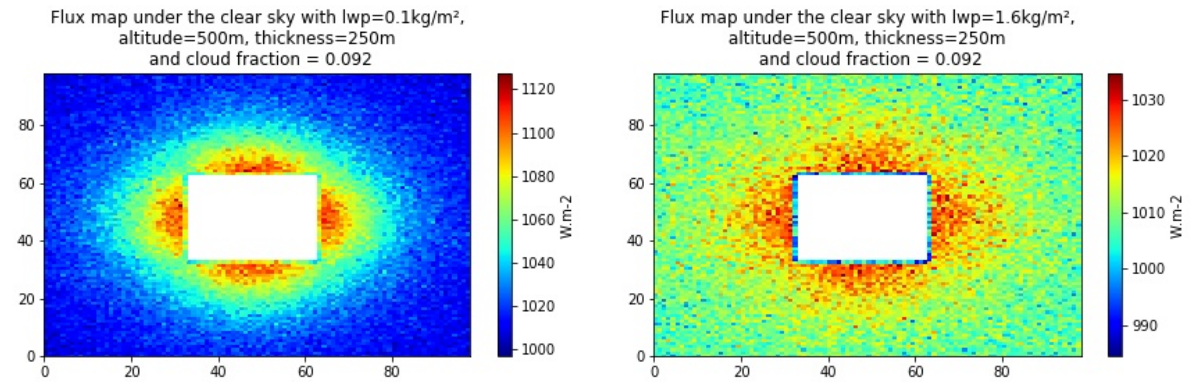
the PDF.

As said before, under a certain optical thickness value, the cloud looks brighter than the blue sky as long as you are not looking at blue sky in the sun direction. Consequently, for a cloud field with Sun at zenith, the highest flux value under the cloud will be at the center of the cloud as shown in Figure 2.8. The thicker the cloud (optically speaking), the lesser this effect will be, because their will be more radiation reflection at the top of the cloud, and therefore less scattering, meaning less photons will leave the cloud by its base or side. The reason why there is more received flux at the center of the cloud is that this area sees more cloud. But when the optical thickness of the cloud is higher, the cloud is darker, meaning that seeing more cloud will not result in as much extra radiation. And around the cloud, the highest values will be directly around the cloud, since the closer to the cloud, the bigger it appears (Figure 2.9).



(a) (b)

Figure 2.8: Flux under a cloud with sun at zenith made with a artificial cloud field, containing one square cloud

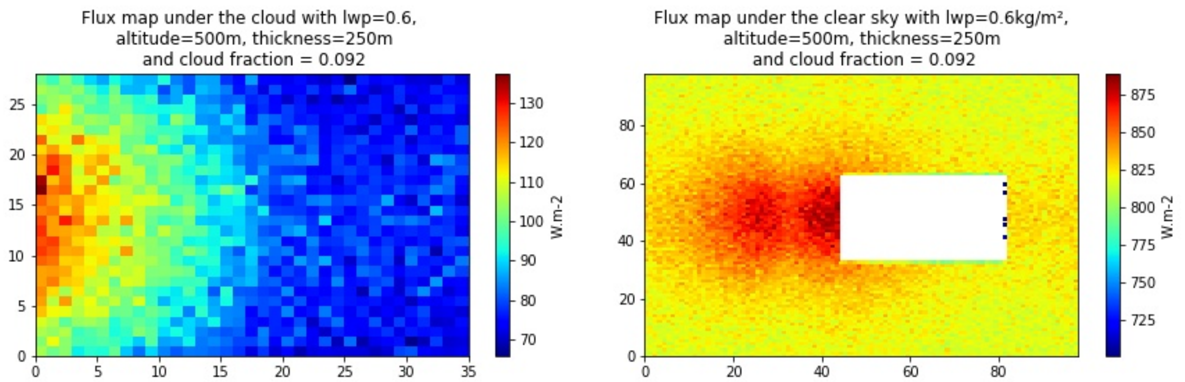


(a) (b)

Figure 2.9: Flux around the cloud with sun at zenith, made with a artificial cloud field, containing one square cloud

But when the sun is not at zenith, the behavior is quite different (Figure 2.10). Note that here the Sun is on the left. In the shadow of the cloud, the enhancement is no longer in the center of the area, but is now at the front of the shadow. This enhancement is caused by 2 phenomena. First, sun radiation is no longer coming vertically onto the cloud. Consequently, all the photons are not crossing the same amount of liquid water path. Imagine columns starting on the ground and having an angle with the ground equal to the solar zenith angle. When this angle is 90°, the column are vertical, and since the cloud used here is homogeneous, all columns have the same amount of liquid water path. But imagine this angle is now 55°. Some of the column will cross only a small part of the cloud. And those columns are the one crossing the cloud and its bottom left and top right part. Therefore, the projected optical thickness is lower at the extreme left and extreme right of the shadow of the cloud. Thus, at this point, we expect the enhancement zone to be in those areas (left and right of the shadow). But now we have to take 3D effects into account. As said before, for the used values of liquid water path, the cloud scatters more radiation than the atmosphere. And since the sun here is on the left of the cloud field, the pixels located on the left of the cloud shadow are closer to the cloud (and can even be under the cloud depending on the altitude of the cloud and the solar zenith angle) and hence, see more cloud than the pixels on the right. This is why the enhancement zone is located on the left of the cloud's shadow.

In the clear sky areas, there are now 2 enhancement zones in front of the cloud shadow. One is caused by reflection on the cloud edge and the second one is caused by scattering (3D effects again). The one caused by reflection is the first one starting from the left of the map. Since sun radiation is coming from the left, some of it hits the left side of the cloud and is therefore reflected not to space but to a ground area located on the left of the cloud shadow. The second enhancement zone is a result of 3D effect. In fact, this area is qualified as a clear sky zone, but is located under the cloud. But since the sun is low on the horizon, this area is receiving direct radiation from the sun. But this area is under the cloud, so the cloud represents a huge solid angle and therefore, there is a lot of diffused radiation in this area, resulting in the observed enhancement.
To be sure of the identification of the different areas, we see what happens to each area when LWP is changed. Because one area is caused by reflection, which will be higher if there is more LWP, and one by scattering, which will be reduced if the LWP is higher, we will be able to verify that the first area is a consequence of the reflection and that the second one is a consequence of the diffusion. This is verified in Figure 2.11



(a) In the shadow of the cloud

(b) Where the sun is seen directly

Figure 2.10: Flux when the sun has a solar zenith angle of 55 degrees

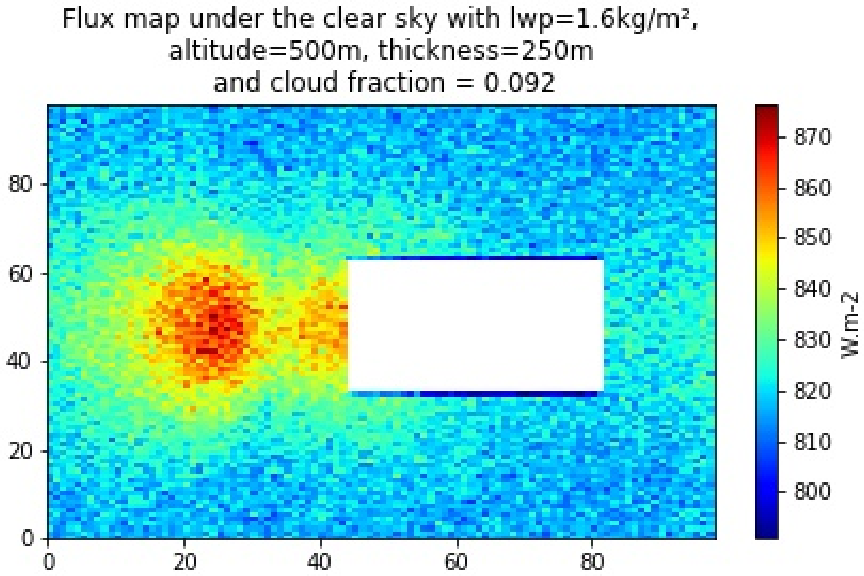
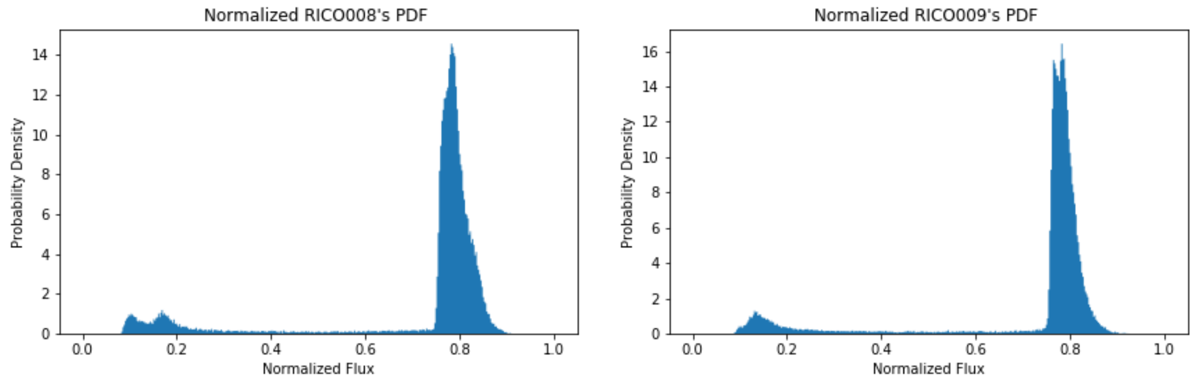


Figure 2.11: Flux map under clear sky, with optical thickness of 240

2.4 Pdf with other shapes

Until now, every cloud field that I used had a 2 peaks PDF over which I could fit 3 Parameters log-normal distribution. The idea was to then use the 6 obtained parameters and see how they evolve when different cloud field parameters are modified. But when I was doing new simulation and applying the fit on them, I noticed that sometimes, the fit did not work because the PDF had not the expected shape (Figure 2.12).

Sometimes, the small irradiance mode does not show one but two maxima. I tried to explain this shape, but I couldn't find a reason. There is no particularities in the LWP distribution for the cloud field that is used to plot the 2.12.a PDF. And in other cases, it is the large irradiance mode that has 2 peaks (See Figure 2.12b). In figure 2.13, we can see where the pixels belonging to each peak are located.



(a) (b)

Figure 2.12: Example of pdf with unexpected shapes

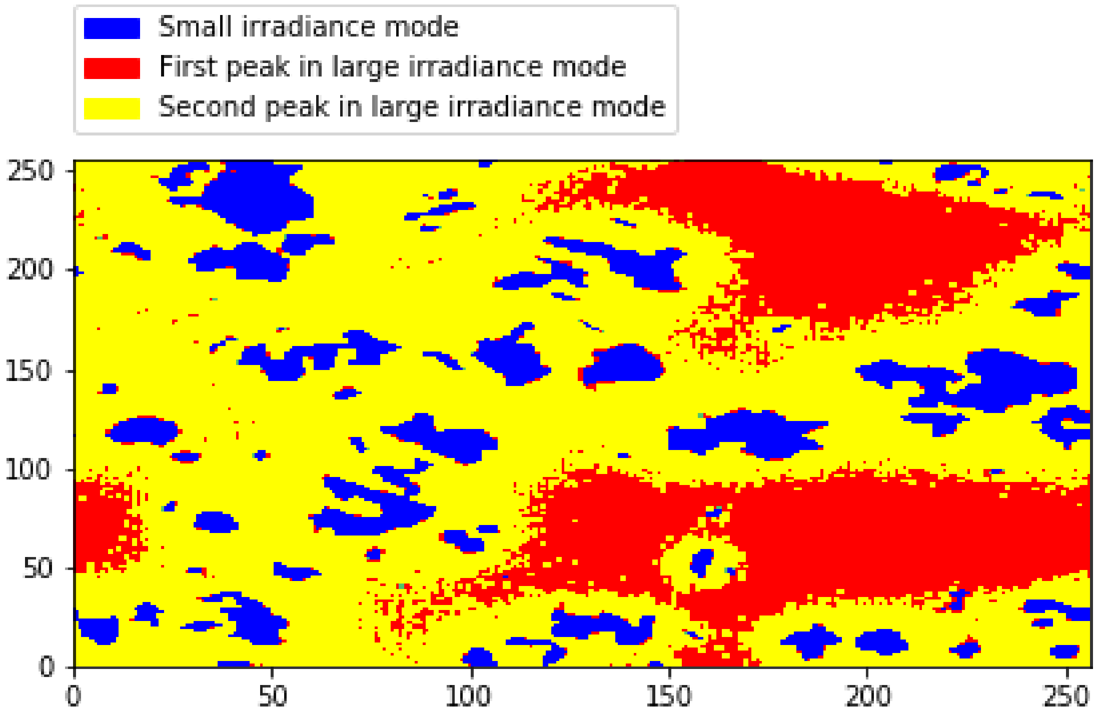


Figure 2.13: Spacing of the pixels belonging to each peak of the PDF

With Figure 2.13, it appears that the existence of 2 peaks in the large irradiance mode has to do with the 3D effects. As shown in section 2.2, the 3D effects decrease really fast with the distance to the edge of the clouds. Thus, the areas that are far from the edge of clouds will receive less radiation (the difference is approximately 20 W m^{-2}) than the areas close to the clouds. And with Figure 2.13 we can see that most of the pixels that make the first peak of the large irradiance mode are located in large areas where there are no clouds around. Therefore, we can conclude that two peaks for large irradiance mode occur when there are in the cloud fields large regions with no clouds in it (i.e. a large inter-distance between the clouds).

2.5 Impact of the cloud field parameters

We now need to find cloud field parameters that well describe the cloud field, but also have an impact on the radiation so that the pdf is different if the parameter is evolving. To do so, I will be using the square cloud presented before. The quantity of water inside the cloud, the size of the cloud, the cloud geometrical thickness and its cloud base height will be modified, to know how that affects the pdf of the surface solar irradiance. LWP dispersion and specific surface (that will be defined later) will be discussed as well.

2.5.1 Liquid Water Path (LWP)

The main driver of the received flux

The liquid water path is the amount of water integrated over one column of the cloud field. Since water absorbs and scatters solar radiation, it is intuitive that the received flux on the ground will depend on the quantity of water that radiation has to cross. In order to see how is this relation, I took a cloud field and the resulting solar surface irradiance field with the sun at zenith, and plotted for each value of the measured flux, the mean LWP (Figure 2.14).

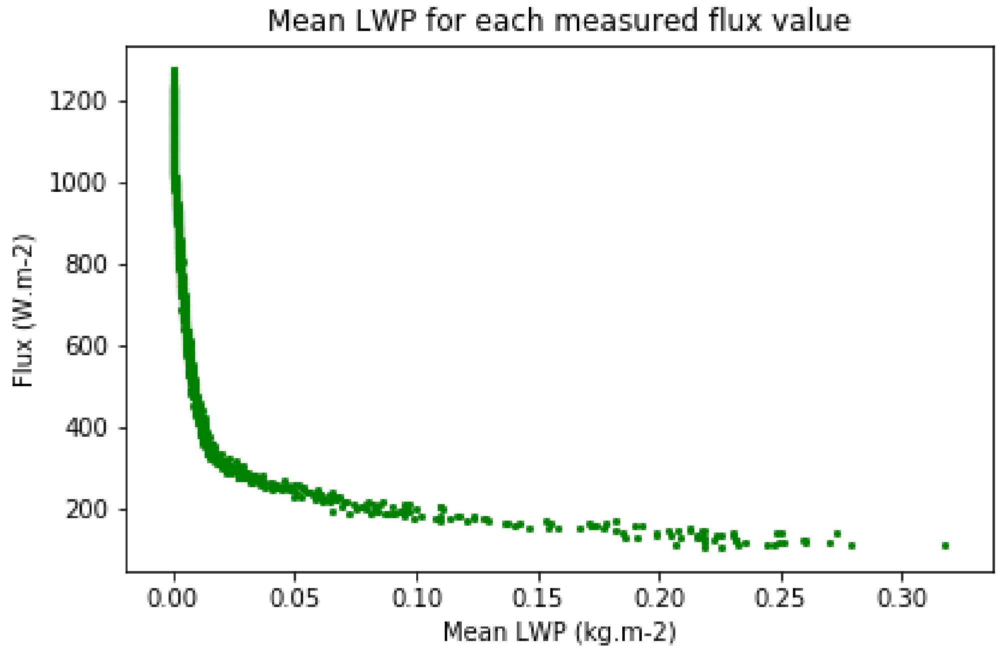


Figure 2.14: Relation between measured flux and the mean LWP above

As expected, there is definitely a correlation between the mean LWP and the flux value. This shows that the flux received under the clouds depends mainly on the quantity of water there is in the clouds. Therefore, the mean LWP of the cloud field is the first parameter that I will keep and try to predict with the neural network.

Impact of the LWP on the pdf

LWP of the clouds has an impact on the flux reaching the surface. But how can it be observed in terms of pdf? To learn more about that, I made several simulations with a theoretical single cloud in which I changed the LWP and then plotted the resulting pdf in the same figure (Figure 2.15).

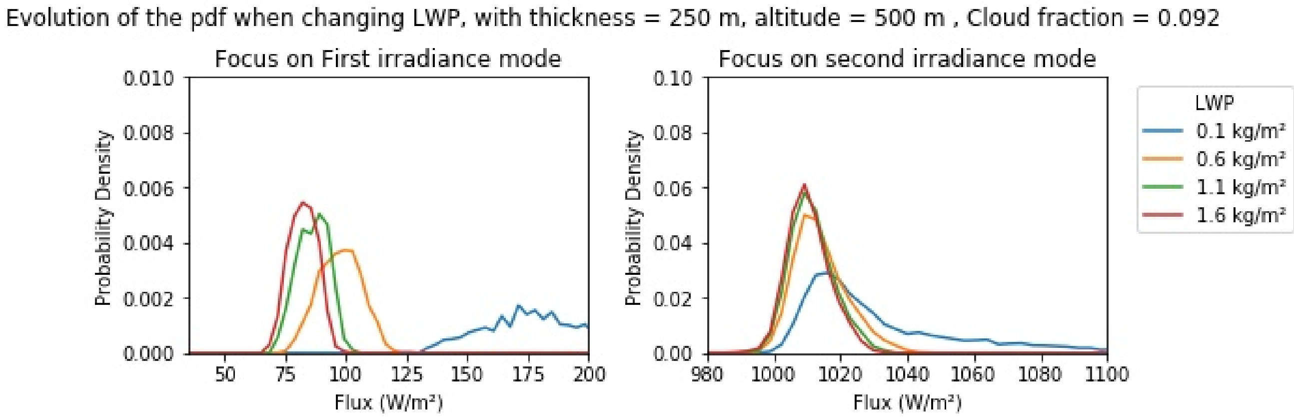


Figure 2.15: Evolution of both irradiance modes of the pdf when changing LWP, with sun at zenith

As the LWP (and therefore the optical thickness) increases, the reflection of the radiation at the top surface of the cloud (and the side surface when the sun is not at zenith) is increasing, leading to diffusion reduction, reducing the radiation crossing the cloud. Consequently, the received flux under the cloud is reduced, leading to a shift towards the left of the first irradiance mode. Moreover, since there is less diffusion, the 3D effects decline, leading to a narrowing of both irradiance modes (there is a narrowing because less 3D effects means less flux enhancement, and therefore it tends to uniform the values under and around the clouds), and to a shift on the left of the large irradiance mode too. This behavior is observed regardless of the solar zenith angle.

2.5.2 Cloud Fraction

Impact on the irradiance mode integrals

The cloud fraction represents the proportion of the sky that is covered by clouds. This is a parameter that very well describes the global cloud field. And this parameter can be easily obtained from studying the PDF of the solar surface irradiance. Indeed, all we need is to be able to count the number of pixel of the solar surface irradiance field that are covered with clouds. And this is really easy to do when the sun is high in the sky. If the sun is at zenith, we simply have to count the proportion of pixels that are receiving a flux inferior to the clear sky value. And in the PDF, this information is located in the integral of the first irradiance mode, as shown in figure 2.16.

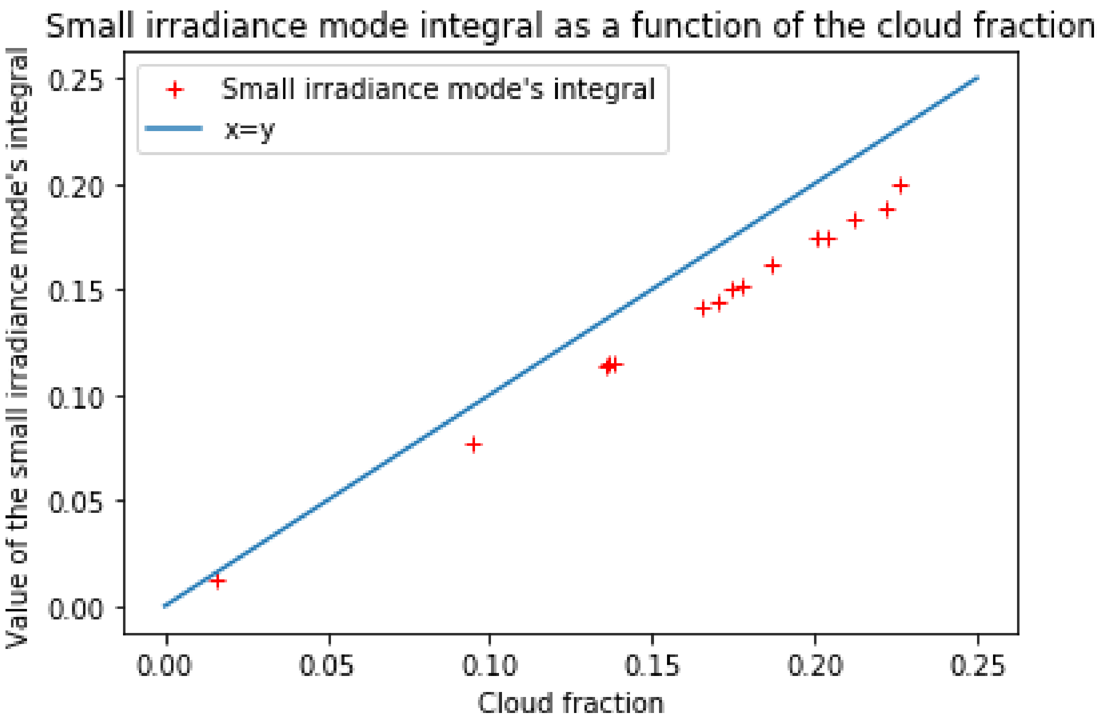


Figure 2.16: Integral of the first irradiance mode, as a function of the cloud fraction

Consequently, the cloud fraction will be a parameter that I will try to predict too, using the neural network.

Impact on the position of the pdf

The cloud fraction changed the proportion of the two irradiance modes but we can wonder if it has other consequences on the pdf. To answer this I made a few simulations with a theoretical single cloud changing only the horizontal extent (and by doing so, changing the cloud fraction).

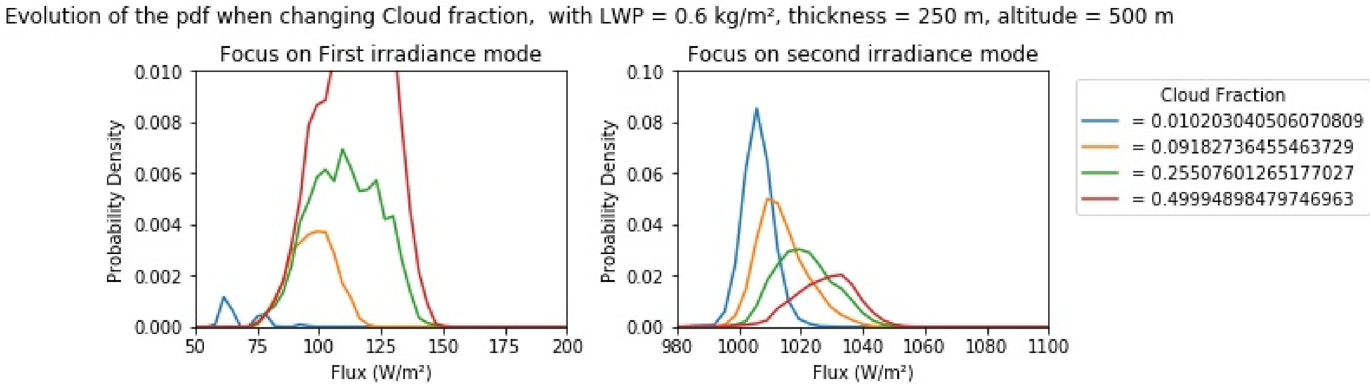


Figure 2.17: Evolution of both irradiance modes of the pdf when changing CF, with sun at zenith

As we can see in Figure 2.17, both irradiance modes are shifted to the right when the cloud fraction is increased. This is explained with the rise of the 3D effects. Indeed, when the cloud fraction is larger, the solid angle corresponding to the cloud is increasing for every measurement point at the surface, under the cloud and under the clear sky, because since the cloud is bigger, it appears bigger. And as said in a previous section, the cloud is scattering more radiation than the blue sky when looking in any direction other than the sun, making the clouds look brighter than blue sky. Therefore, the more cloud you see the more diffuse radiation you will receive. Consequently, the higher the cloud fraction, the more 3D effects, leading to the observed rightwards shifts.

2.5.3 Cloud base height

The cloud base height is important to describe a cloud field. But has it an effect on the solar surface irradiance? To answer this I used the same method as previously, making several simulations and only changing the cloud base height.

With Sun at zenith

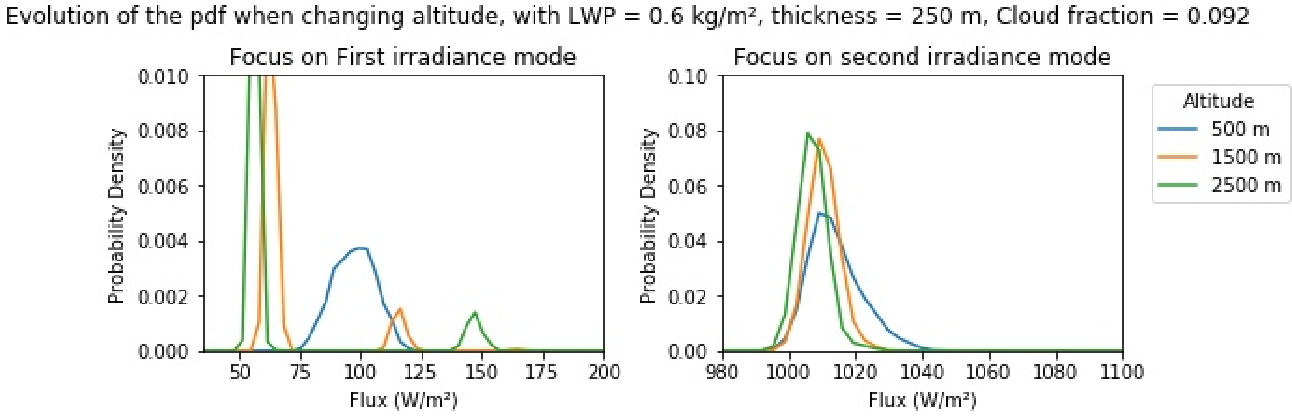


Figure 2.18: Evolution of both irradiance modes of the pdf when changing cloud base height, with sun at zenith

Figure 2.18 shows that both irradiance modes are shifted to the left and narrowed when the cloud base height is increased. This behavior is expected, since the higher the clouds, the smaller the cloud looks for any pixel on the ground, and therefore, the less important the 3D effects. Nevertheless, when verifying that 3D effects where indeed at the origin of this behavior, I realized that it wasn't the only phenomenon at stake for the first irradiance mode. Indeed, ground albedo and water vapor absorption are both playing a role too. For ground albedo, this is explained by the multiple reflections between the ground and on the cloud base, that will be reduced if the altitude of the cloud base raises. With Figure 2.19, which is a 1D simulation with no water vapor and a ground albedo of 0, we see that there is still a left shift, meaning that an other phenomenon is interfering, but I wasn't able to identify it.

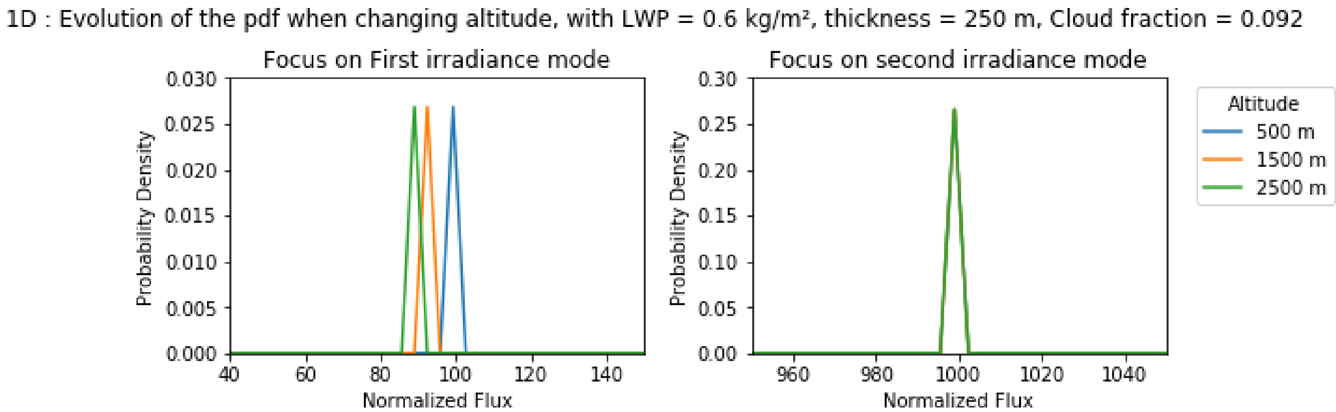


Figure 2.19: 1D Simulation : Evolution of both irradiance modes of the pdf when changing cloud base height, with sun at zenith, ground albedo at 0 and no water vapor

With a solar zenith angle of 55°

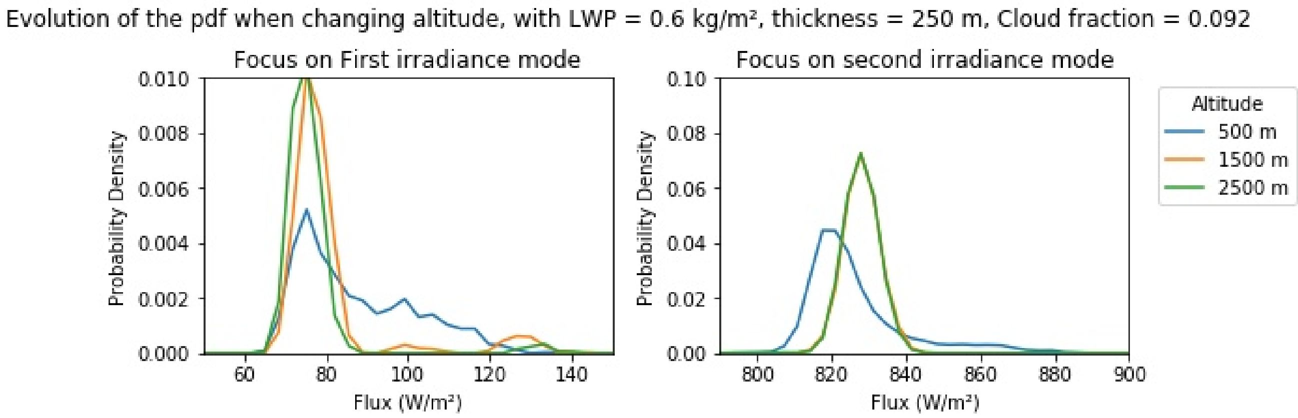


Figure 2.20: Evolution of both irradiance mode of the pdf when changing cloud base height, with sza = 55°

For the first irradiance mode, there is still the left shift for the same reasons as before (Figure 2.20). But for the large irradiance mode, we can see that the dynamic is different, with the pdf for an altitude of 500 m that is way wider and takes both smaller and larger values. For the larger values, it is expected, because again as the cloud is closer to the ground it appears bigger, resulting in both more received diffusion and reflection for the concerned pixels. But for the smaller values, this is probably a consequence of the small size of the cloud field. Since the cloud field is only 99x99 pixels wide, and that this cloud field is periodically and infinitely repeated for the htrdr simulation, we might observe higher flux values for high cloud case because of the influence of the neighbouring clouds. When using a larger cloud field, this behavior is highly reduced.

2.5.4 Geometrical thickness

The same study can be done with the thickness of the clouds. Since thickness is directly linked to the size of the sides surfaces of the cloud, and therefore might have an effect of the strength of the 3D effects, thickness is highly likely to have an impact on the pdf. Once again I used the theoretical cloud to see the geometrical thickness' effects on the PDF of the solar surface irradiance.

With Sun at zenith

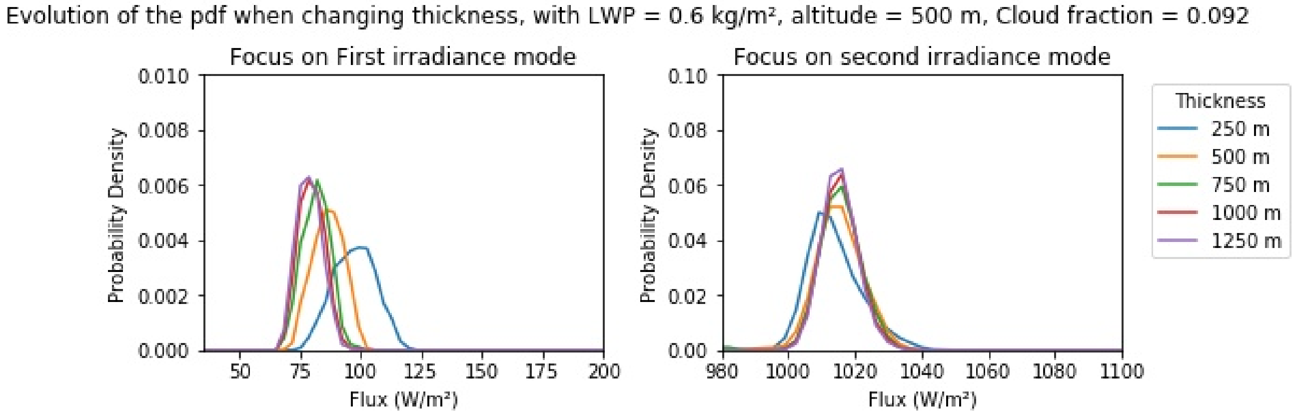


Figure 2.21: Evolution of both irradiance modes of the pdf when changing thickness, with Sun at zenith

For the small irradiance mode, a left shift is observed when for the large irradiance mode, there is a right shift. The explanation is simple, the thicker the cloud is, the more likely the radiation are to leave the cloud by the side and not by the base of the cloud. For that reason, the thicker the cloud is, the darker the cloud base appears and the brighter the cloud sides are, meaning less 3D effects for pixels under the clouds and more 3D effects around the clouds.

With a solar zenith angle of 55°

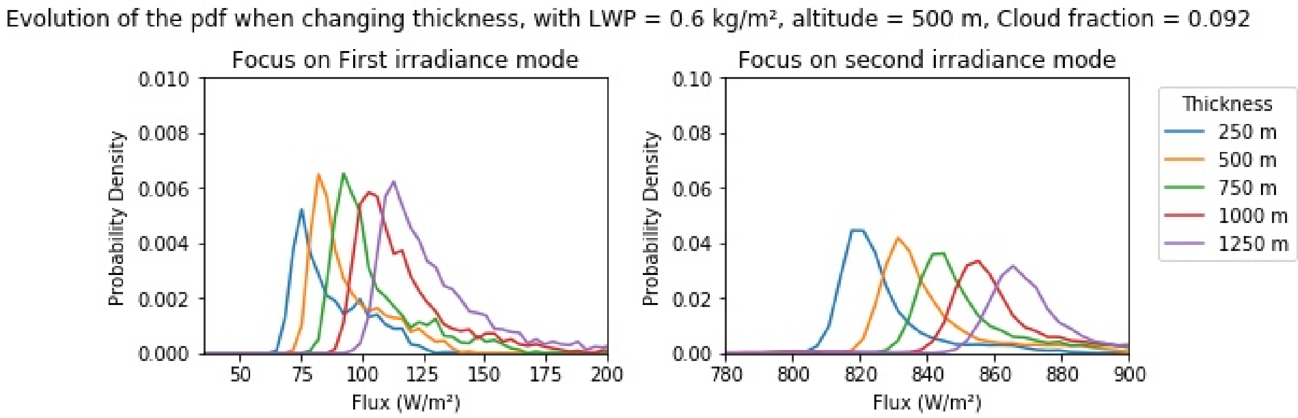


Figure 2.22: Evolution of both irradiance mode of the pdf when changing thickness, with sun at zenith

When the Sun is not at zenith, the results change for the small irradiance mode. There is now a right shift. There are 2 reasons for this. First, the apparent surface of the cloud receiving solar radiation is now bigger if the cloud is thicker, since the cloud side is receiving radiation too. Therefore, the cloud will scatter more radiation, leading to that flux enhancement. Second, as we explained it when talking about the flux when Sun is not at zenith (section 2.2), some columns starting from the ground and reaching the Sun would not cross the entire cloud. And if the cloud is thicker, with the same amount of liquid water path, that means that for the same distance crossed in the cloud you will be crossing less water. That leads to an enhancement too. On the other side, the thicker the cloud is, the bigger the cloud shadow will be, leading to the growth of the first mode integral.

For the large irradiance mode, the shift is the same but is much stronger. As we said before, the right shift observed is caused by 3D effects. Now that the Sun is not at zenith, the surface of cloud receiving radiation is larger, since the side of the cloud is concerned too, and it gets even larger when the thickness of the cloud is increasing. Therefore, the 3D effects are more important, this is why we observe a stronger shift.

2.5.5 Specific Surface and LWP relative dispersion

The specific surface is defined as the total surface of an object divided by its volume. To calculate it, we can use the auto-correlation of the cloud field . Indeed, the specific surface is equal to the reverse of the opposite of the derivative of the auto-correlation function in 0. Therefore, I calculated the 2D auto-correlation of the LWP map, obtained from the vertically projected cloud field, and deduced the specific surface, by fitting an exponential function [1] on the obtained auto-correlation function. This specific surface contains information on both the size of the cloud and the space between them which has an impact on the measured pdf as discussed earlier, and will consequently be one of the parameters that we will try to retrieve.

The LWP relative dispersion contains information about the clouds heterogeneity, which is an important parameter to characterize the cloud field and therefore will be a parameter of interest for the rest of the study.

Chapter 3

Neural Network

Now that we have selected 6 parameters (Cloud fraction, mean liquid water path, cloud base height, liquid water path dispersion, cloud geometrical thickness and specific surface) that both give information on the cloud field and have an impact on the pdf of the received surface flux, the objective is to predict those parameters describing the cloud geometry from the corresponding pdf of surface fluxes, using a neural network.

3.1 Architecture

A Neural network is a computing system inspired from biological neurons. It is composed out of neurons that can be linked. Each neurons when receiving a signal is processing it and then transmitting the resulting signal to the neurons he is linked with. The signal is a real number, that is obtained from non linear function applied to the sums of the inputs a neurons has. Neural network are used in a lot of disciplines because of their ability to reproduce nonlinear processes. It can be used for classification, data processing, and regression analysis, which is why it seems to be a good choice for retrieving cloud field parameters from the solar surface irradiance's pdf. First I need to define the neural network I have been using. The opposite process (predicting the pdf from the cloud field parameters) had already successfully been tested [2], so I decided to use the same architecture : one input layer, several hidden layers, and then an output layer, with between each layers connections between each neurons. To make the neural network, I used keras [4], which has a hyper parameter optimizer included. Hyper parameters are the number of layers, the number of neurons in each layer and the learning rate. The final architecture of the neural network can be seen in Figure 3.1. This neural network takes the pdf of solar radiation (split in 400 equal bins) and the corresponding solar zenith angle as inputs, and returns the 6 cloud parameters that I described before as outputs.

```
Model: "model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 400)]             0
_________________________________________________________________
layer1 (Dense)               (None, 570)               228570
_________________________________________________________________
layer2 (Dense)               (None, 230)               131330
_________________________________________________________________
layer3 (Dense)               (None, 700)               161700
_________________________________________________________________
layer4 (Dense)               (None, 330)               231330
_________________________________________________________________
layer5 (Dense)               (None, 200)               66200
_________________________________________________________________
dense (Dense)                (None, 6)                 1206
=================================================================
Total params: 820,336
Trainable params: 820,336
Non-trainable params: 0
```

Figure 3.1: Neural Network architecture

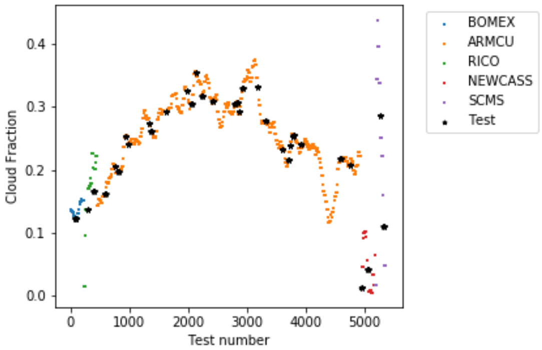
I then needed to choose the right loss function. The loss function is what the neural network will try to minimize during the training. It is supposed to show how far the prediction made is far fro the expected value. Here I use the relative mean error (for every parameter, the relative error is computed, and then the mean is made which gives the loss to minimize) which has the advantage to give the same weight to every parameters.

3.2 Data allocation

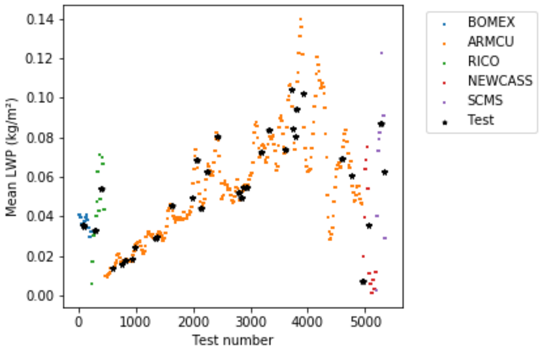
I will now describe the data I have been using for the learning and the test of the neural network. I used 5 different types of cloud fields. For each cloud field file, 15 different simulations were run with Htrdr, for various solar zenith angles, from 20° to 90° with a step of 5°. For the cloud fields, I have:

- 12 RICO files [5]
- 16 NEWCASS files
- 300 ARMCU files
- 15 BOMEX files
- 11 SCMS files

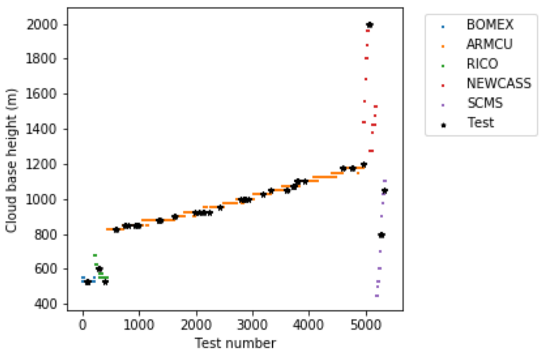
So I have a total of 354 cloud fields which makes 5310 simulations. I then have to divide my entire sample into two parts: one group for training and one for testing. To do so, I randomly take 10% of the simulations and put them in the test group. The rest goes for the training. In Figure 3.2 you can see data allocation with the black dots localizing the test sample :



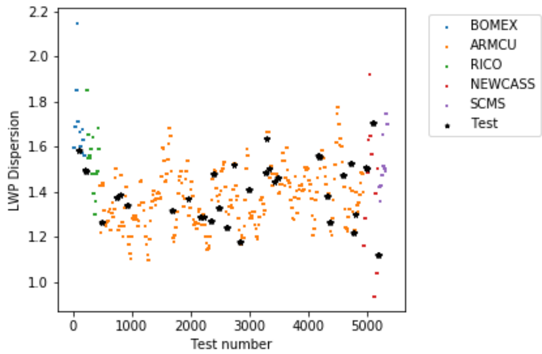
(a) Cloud Fraction distribution



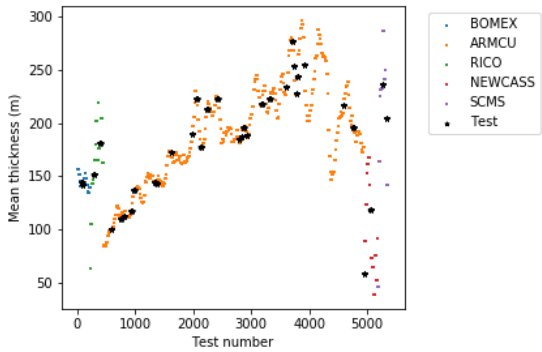
(b) Mean LWP distribution



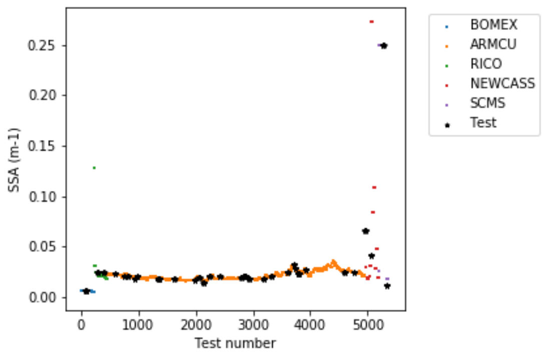
(c) Cloud base height distribution



(d) LWP Relative Dispersion distribution



(e) Mean thickness distribution



(f) SSA distribution

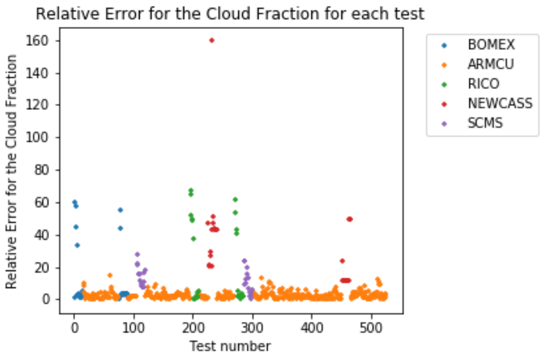
Figure 3.2: Parameters distribution, with in black symbols the one kept for the test sample

3.3 Results

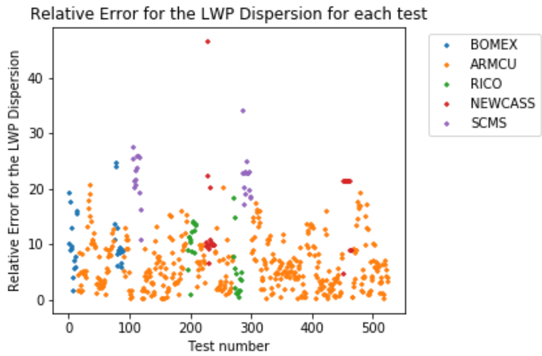
After separating the sample in two, we need to train the neural network and then test it. In Figure 3.3, you can see the errors obtained for the test. It is important to know that for every cloud file there are 15 simulations, which means for every selected file for the test, there are 15 tests (one for each solar zenith angle). On the following figures, the 15 tests of a file are represented in a row (for example, in Figure 3.4, the 15 first points correspond to one BOMEX file)

Figure 3.3: Mean relative error during the Test

The overall mean relative error for the test is 9.8% but we can see that there are a lot of variation, depending on the type of clouds. It appears that prediction for ARMCU cloud files are way better than the other, which makes sense since there are a lot more ARMCU files in the training sample than the other files. The other striking thing is that the quality of the prediction seems to depend on the solar zenith angle, which is unexpected since the solar zenith angle is given as an entry parameter, but yet the neural network seems to have difficulties to adapt to every solar zenith angles. In order to understand why some errors may occur, we can look at the relative error for each parameters (Figure 3.4) and the difference between expectation and prediction for poorly predicted files (meaning every file where the mean relative error is above 30%) (Figure 3.5).



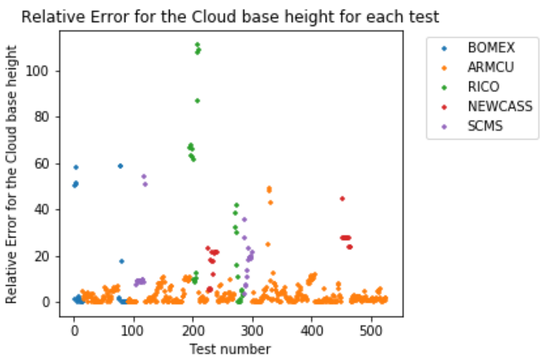
(a)



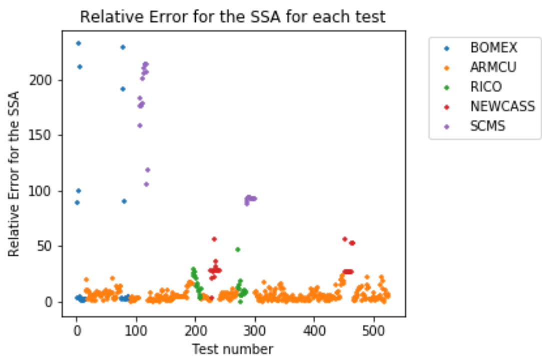
(d)

(b)

(e)



(c)



(f)

Figure 3.4: Relative Error for each parameter

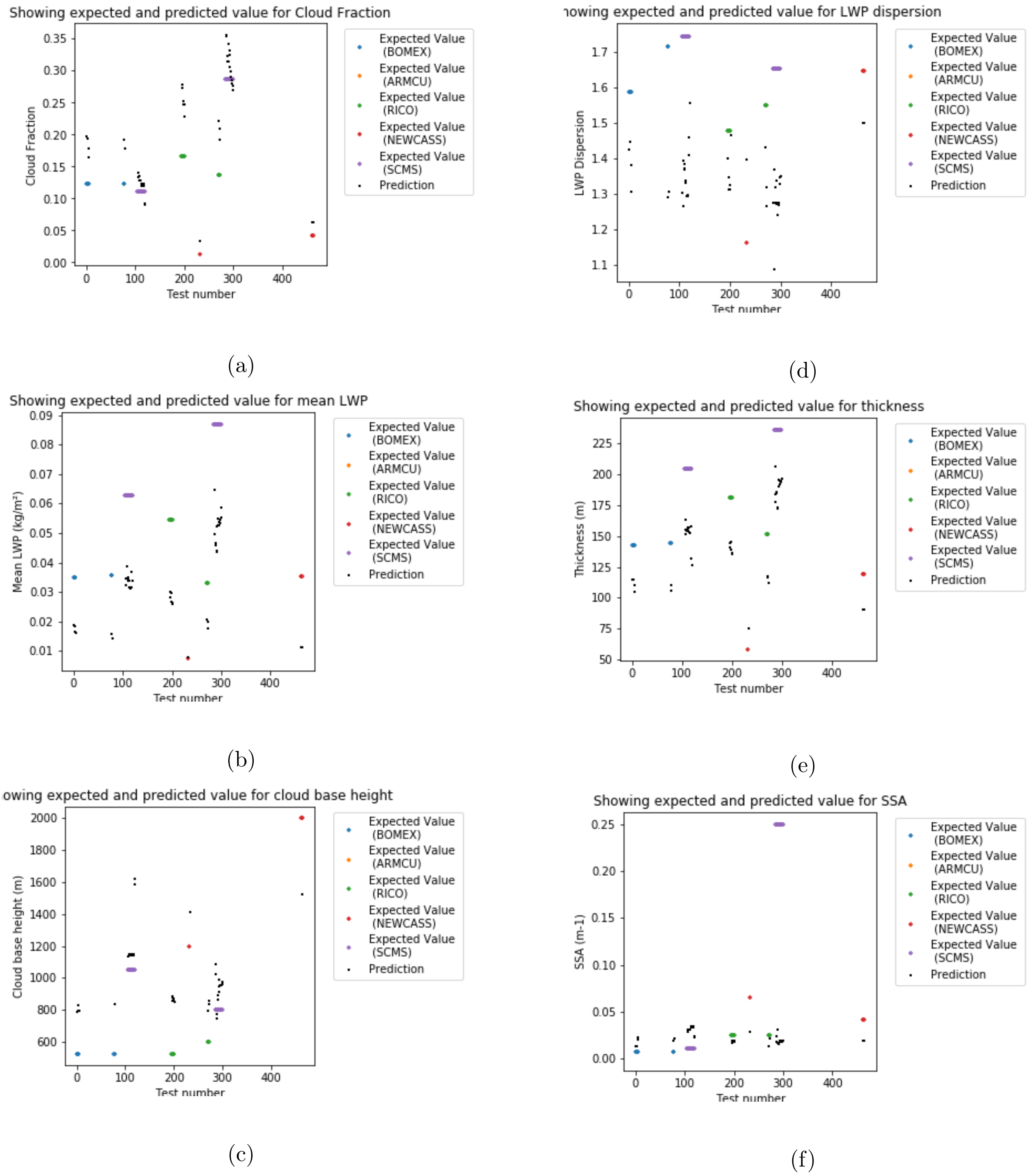


Figure 3.5: Expectation and prediction for each parameters for badly predicted files (Error >30%)

For the badly predicted BOMEX and SCMS files, with Figure 3.4 we can see that a huge part of the mean error comes from the SSA relative error. This error is huge because the absolute value is very low (Figure 3.5.f) , so a small error leads to a big relative error.

For the error on the other parameters, when we look at all variables in Figure 3.5, we can see that Cloud fraction is overestimated and mean liquid water path is underestimated. As we saw on section 2.4 a lower liquid water path and a bigger cloud fraction both lead to a right shift of the small irradiance mode. On the other hand, there is a thickness underestimation and a cloud base height overestimation. As we have seen in section 2.4, with the Sun not at zenith, a thinner cloud and a higher altitude lead to a left shift of the small irradiance mode. So maybe these errors compensate each other. Those errors might also come from the training sample. Indeed, when looking at figure 3.2 we might understand some of these errors. For example, in the badly predicted files, the cloud fraction is almost always overestimated. And when looking at Figure 3.2.a, we can see that most of the ARMCU files of the training have a bigger cloud fraction than the ones from other cloud types. The same happens with the cloud base height.

An other strange fact here is that sometimes, like for BOMEX or some RICO files only the first simulations, the one with the lowest sza are badly predicted, when the other seem to be really well predicted (less than 5% error) but this behavior is observed during training too (Figure 3.6).

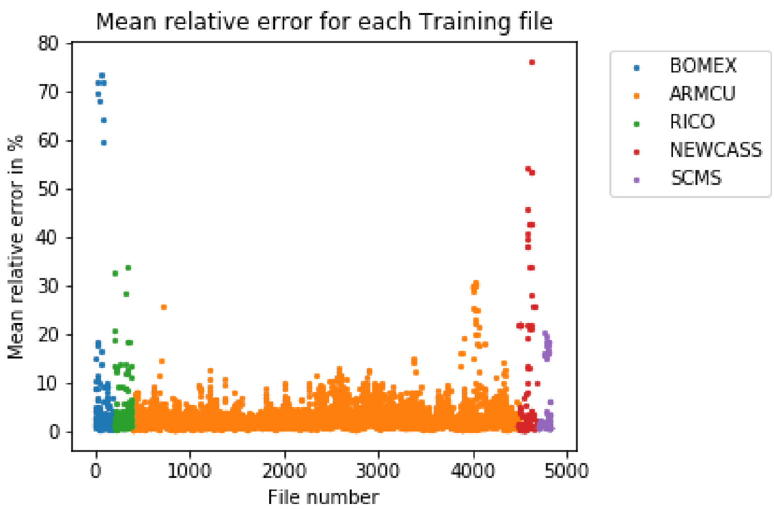


Figure 3.6: Mean relative error during the training

3.4 Further thoughts

Since the mean relative error on all test is less than 10% we could say that the neural network is good at predicting the cloud field parameters from the pdf. The issue is, that the prediction are mainly good for predicting ARMCU, but the results worsen when it comes to other cloud fields. To change that I already tried to reduced the ARMCU file number in the training sample, by picking 6 files of each cloud field type for the training.

But the results got worse, probably due to the reduced number of training data when using that method. But to improve the neural network two ways could be explored : the first one would be to modify the loss function, so that instead of being the average on the all the test, it would be averaged by cloud field type. That way, ARMCU files would have the same weight as the other files in the training phase. The other way would be to add more files of the other cloud field types in a way that ARMCU files do not represent a so big proportion of the data.

Chapter 4

Conclusion

Clouds have a strong impact on the surface radiation, not only via their optical thickness, but also by their size, their altitude or geometrical thickness as shown in section 2. Using a simplified cloud field containing only one cloud, I was able to identify how the cloud fraction, the mean liquid water path, the geometrical thickness and the cloud base altitude impact on the solar surface irradiance's pdf (Section 2.5). But the exact prediction of a real cloud field from a solar surface irradiance measure is more complex because one cloud field does not contain only one cloud. As a result of the presence of multiple clouds in a cloud field, and of the heterogeneity between the clouds and inside the clouds, we observe that sometimes the obtained pdf has not the expected shape as seen in section 2.4. Different shapes can be caused by a large distance between clouds, but also by some cloud distribution.

Moreover, the effects of clouds over the solar surface irradiance involve a lot of phenomenon, a lot of them resulting in what we called the 3D effects, which can be really counter intuitive, like the fact that under a certain LWP threshold, when the sun is at zenith, the maximum value of the received flux under the cloud is a its center.

For all this reasons, a neural network is really well suited for such a cloud parameter's prediction task. Even if the one I developed here reached decent but not perfect results (Mean relative error of 10 %), with further work on the neural network architecture and on the data available, it might come to really good prediction.

Since the phenomenon at stake for the interaction between clouds and radiation depend on the wavelength the pdf of the solar surface irradiance are very different when computing the simulation for one wavelength, using such a neural network at specific wavelengths could give extra information or be more efficient.

If such a process was to be used to determine the characteristics of actual cloud field, we would have to build a pdf from only point measurement, since it is not possible to have an entire flux map measurement for a multiple kilometers wide area. This is easy doable, by placing multiple pyranometers on the field, and doing time measures. The more pyranometers there is, and the higher the frequency of the measures are, the better, but when these aren't enough, the final pdf can be approximated by using interpolation.

Bibliography

[1] P. Debye. “Scattering by an Inhomogeneous Solid”. In: (1949).

[2] Jake J. GRISTEY. “On the relationship between shallow cumulus cloud field properties and surface solar irradiance”. In: (2020).

[3] Jake J. GRISTEY. “Surface Solar Irradiance in Continental Shallow Cumulus Fields : Observations and Large-Eddy Simulation”. In: (2020).

[4] *Keras API.* URL: `https://www.tensorflow.org/api_docs/python/tf/keras`.

[5] Robert M Rauber. “Rain in Shallow Cumulus Over the Ocean: The RICO Campaign”. In: (2007).

[6] *Robin Hogan’s presentations.* URL: `http://www.met.rdg.ac.uk/~swrhgnrj/presentations/`.

[7] Sophia A. K. Schäfer. “Representing 3-D cloud radiation effects in two-stream schemes: 1. Longwave considerations and effective cloud edge length”. In: (2016).

[8] Najda VILLEFRANQUE. “A path-tracing monte carlo library for 3D radiative transfer in highly resolved cloudy atmospheres”. In: (2019).