MCG-Rad: Monte-Carlo Global Radiative Forcings Computation

Summary table of persons involved in the project:

Partner	Name	First name	Current position	Role & responsibilities in the project (4 lines max)	Involvement (person.month) throughout the project's total duration
LMD/IPSL	DUFRESNE	Jean-Louis	Senior researcher	Scientific coordinator of the project, Leader of tasks 0 and 2	48 PM
LAPLACE	FOURNIER	Richard	Professor	Leader of task 3 Monte-Carlo applied to engineering and atmospheric sciences	12 PM
IRIT	PAULIN	Mathias	Professor	Leader of task 1 Computer graphics and image rendering	12 PM
LMD/IPSL	CREVOISIER	Cyril	Senior researcher	Radiative transfer, spectroscopy data and remote sensing	4 PM
LMD/IPSL	HOURDIN	Frédéric	Senior researcher	Climate modelling, atmospheric transport and radiative transfer	8 PM
LMD/IPSL	ARMANTE	Raymond	Research engineer	Radiative transfer, spectroscopy data and remote sensing	10 PM
LMD/IPSL	SIMA	Adriana	Research engineer	Climate modelling and aerosol modelling	12 PM
LAPLACE	BLANCO	Stéphane	Senior Lecturer	Monte-Carlo applied to engineering and atmospheric sciences	12 PM
IRIT	MELLADO	Nicolas	Researcher	Computer graphics and multiscale analysis	12 PM
LAPLACE & Méso-Star	COUSTET	Christophe	Research Engineer	Software engineering	8 PM
LAPLACE & Méso-Star	EYMET	Vincent	Engineer	Radiative transfer	8 PM
LAPLACE	LAPEYRE-DUBOCS	Paule	PhD Student	Domain-deformation sensitivity estimates	8 PM
LAPLACE	PENAZZI	Léa	PhD Student	Implicit Monte Carlo	8 PM
LAPLACE	TREGAN	Jean-Marc	PhD Student	Nonlinear Monte Carlo and senstivity estimates	4 PM

Any changes that have been made in the full proposal compared to the pre-proposal

No significant changes compared to the pre-proposal

The reviews of the pre-proposal were globally positive and rise two questions mainly due to the limited length of the text. The first is relative to the IRIT computer graphics group which appears to be less connected to the others. The second is the lack of perspective on how we will obtain the derivatives for the sensitivity analysis. More information are now provided in this document, in particular section II-a-1 and I-b-4.

I. Proposal's context, positioning and objective(s)

a. Objectives and scientific hypotheses

I-a-1 Objectives and the research hypotheses

Radiative exchanges are the primary drivers of climate and climate changes. An accurate computation of the radiative flux is thus a necessary step for many scientific studies on climate and climate change. For instance, radiative forcing is a metric which allows to estimate and to compare the amplitude of the various natural and anthropogenic perturbations on climate (Myhre et al. 2013). Global indexes which are used for international negotiations such as the follow-up of the Paris agreement on climate change are based on radiative forcing estimates. The radiative forcing for a change of a greenhouse concentration is defined as the change of the flux at the top of atmosphere when only the concentration of this greenhouse gas is modified, i.e. before any other atmospheric variable adjusts. Computing radiative forcing at global scale requires to compute the radiative flux for a large set of atmospheric profiles and no reference model has been used to compute the radiative flux at global scale as they are by fare too expensive to run. The usual method is a multi step approach. Firstly line-by-line reference models compute the radiative exchanges for an ensemble of representative atmospheric profiles. Secondly more simplified and less expensive models (narrow band or broad band models) are elaborated and validated against these reference results. Finally these models are used to compute the radiative exchanges and the radiative forcings at the global scale by a 4 dimensional integration (1 time and 3 space dimensions).

This approach as several limitations. The adjustment and validation of these simplified models is demanding, made only from time to time, creating a delay of several years between the improvement of a spectral database and its use for radiative forcings estimate. This validation is based on a sample of a few atmospheric profiles (one to a few tens), which is sufficient for clear sky conditions but questionable for cloudy conditions. Concerning clouds, their 3D structure is highly simplified in all current estimates and the radiation is computed as a 1D phenomena, not 3D (plan parallel approximation). Finally, the calculation of radiative forcing for many gases by individually modifying the concentration of each gas requires a number of simulations which increases with the number of gas.

The primary objective of this project is to develop and implemented innovative Monte Carlo methods and computer techniques that have been developed in the engineering science community to accurately estimate the radiative forcings with a reference model that directly uses molecular-transition database and computes the radiative fluxes, as well as their sensitivities to a change in the concentration of absorption gases, at the global scale in one single step.

I-a-2 Scientific and technical barriers to be lifted

Since the origin of numerically-simulating radiative transfer, it was claimed that statistical approaches were the only practical way towards the simultaneous handling of all the optico-geometric complexity of radiation in 3D realistic systems. "Monte Carlo is the only numerical tool that passes infinite dimension". The meaning of this sentence is quite clear for algebraic *linear* systems, for instance, and radiative transfer is indeed linear-transport physics. But linear-transport means that the extinction of a beam is exponential and the nonlinearity of this exponential-extinction (Beer law) is at the origin of severe difficulties. As far as the present project is concerned, the millions of absorption lines that we need to handle define the atmospheric-opacity and appear within the exponential, whereas the multiple scattering optical paths are "outside" the exponential (they define the line of sight along which Beer extinction is applied). Spectral integration and optico-geometric integration are therefore combined *nonlinearly*, and this nonlinearity is sufficient to cancel out large parts of Monte Carlo benefits.

But M. Galtier (Galtier et al. 2016, Eymet et al. 2013) and J. Dauchet et al. (2016) have recently proposed a way to bypass this nonlinearity without any loss in accuracy, and this proposition was shown to allow extremely fast simulations of top-of atmosphere fluxes. In few words, the main idea is to make use of "null colliders" (null-absorbers, null-scatterers) that one adds to the field in order to make it look homogeneous. But when a photon encounters one of these null-colliders, it simply continues its path as if the collision had no impact. The numerical solution remains identical, but in the underlying statistical formulation the line spectrum has been shifted from inside to outside the exponential. The main consequence is that the spectral integration is no longer specific compared to other integrations. Now all the integrals are combined linearly and we can benefit of Monte Carlo easily handling complex dimensions. Research groups in the main animation-studios of the cinema industry have received the proposition of Galtier et al. and Dauchet et al. with a strong and visible attention: 3 articles in the quite prestigious SIGGRAPH and EUROGRAPHICS conferences (Novák et al. 2014, Kutz, et al. 2017, Szirmay-Kalos et al. 2017) and the launching of Méso-Star, a start-up devoted to Monte Carlo handling nonlinearities in complex systems.

Galtier et al. (2016) established the practical feasibility of sampling molecular-transitions directly within the radiative transfer Monte Carlo; but the probability set they used was not optimized. This first set was “found” essentially by trial and errors. A truly efficient practice of line-sampling Monte Carlo will therefore require that we better understand the geometrical structure of the high-dimension space that we need to sample. This will required that the first phase of the present project be organized in such a way that two communities work together on the revisiting of Galtier et al.’s proposition.

To compute the radiative flux at global scale and for climatic time scales, the optimized sampling of a location at top of the atmosphere, of a time within the considered period, or a multiple-scattering path within a cloudy atmosphere are sources of no theoretical difficulty: they belong to the standard Monte Carlo practice. But the combined samplings of frequency, absorption species and molecular transitions is truly difficult and unexplored apart from Galtier’s proposition. An intense spectral line far from the considered frequency can of course have more impact on the radiative transfer at this frequency than a closer but less intense one. But how can we organise this information considering that we deal with millions of lines and that they change with temperature, pressure and absorber amounts? How can we best pre-calculate and add to the molecular-database a limited number of parameters allowing a fast construction of well-adjusted sampling probabilities? Again, a first solution is already available and realistic simulations of top-of-atmosphere fluxes could be achieved in few minutes for a single atmospheric profile on very standard computers, but scalability is still a very serious issue: considering the accuracy required (of the order of 0.1%) we evaluate that switching to integrated radiative-forcing without specific developments, the computation time would rise up to a day. Variance reduction is therefore essential to the project. It will mean that our tools will remain truly practical for all climatologists whatever their computer equipments.

I-a-3 Expected results

Achievement: *a community-distributed reference radiative tool for climate and climate change analysis*

The primarily goal of this project is **to design and realize a radiative transfer simulation tool that connects directly the output of a climate simulation to a molecular-transition database** such as the widely used GEISA (developed at LMD) or HITRAN. This tool will allow to compute the radiative balance and the radiative forcing of the greenhouse gases at the global scale over one or a few years of atmospheric profiles produced at high frequency (typically 1 to 3 hours) by meteorological centres or climate modelling groups. This computation at global scale with a reference model has never been achieved, is currently considered as infeasible within the existing resources and techniques, and will be a focal point in atmospheric radiative computation. The model will be made freely available with an open-source license. We will produce and publish an ensemble of **new estimates of the radiative forcing of forty greenhouse gases at global scale with a model of unmatched accuracy.**

Beyond the estimate of radiative forcing, the developed model will allow fast computing of radiative transfer for a large variety of atmospheric profiles and over very large spectral domains with **an accurate representation of gas absorption and three-dimensional (3D) propagation of radiation** in cloudy atmospheres. This tool may be of great utility for many atmospheric studies and can change current paradigms in atmospheric radiation transfer studies. This project benefit from the Monte Carlo codes developed in the frame of the ANR High-Tune and PNTS-2016-05 projects, two projects primarily focussed on radiation scattering in very high resolution atmospheric models for the former, in satellite observation for the later. In turn, our project will allow these two communities to benefit from the developments made to accurately represent the absorption and emission by gases.

b. Originality and relevance in relation to the state of the art

MCG-Rad addresses a computation that, until recently, was presumed infeasible. The reason why this has become meaningful is a set of theoretical advances (most of them published in Galtier et al. 2016) associated with feasibility tests clearly indicating that earth-atmosphere is within the range of practical applicability. In Galtier et al, first-level numerical choices have been made in order to prove this feasibility: these choices were clearly presented as "non yet optimized" and an essential part of the present project is precisely this optimization (see Task 1). But here, in the present section, we start by describing the published approach as is. The existence of an already-practicable solution is indeed at the heart of our software-development choices (see Task 3).

I-b-1: Monte Carlo and spectroscopic databases

The main strength of Monte Carlo is its ability to deal with **high-dimension spaces**. Here, we typically use Monte Carlo because we have to compute quantities such as

$$E_{[t_1,t_2]} = \int_{t_1}^{t_2} dt \int_{-\infty}^{+\infty} d\nu \int_{Globe} d\vec{x} \int_{\Gamma} d\gamma \varphi(t,\nu,\vec{x},\gamma) \qquad (1)$$

φ is the monochromatic radiative power transported at time t, at frequency ν, along a multiple-scattering path γ starting at location $\vec{x}$.The addressed climatic quantity is the corresponding energy $E_{[t1,t2]}$ when we integrate this power over all times within the climatic period, all frequencies throughout both infrared and shortwave, all locations on the globe, and all radiative paths.

The radiative-transfer community has **a long practice of handling such combined integrals** and of anticipating which of them will be the source of convergence difficulties (i.e. statistical variance). Here for instance, without any doubt the main source of variance will be the frequency-domain. The spectral lines of molecular gases as so sharp (so localised) that a naive-sampling of frequency must be avoided. The second source will be multiple-scattering. Although quite impressive at first glance, the integration over the whole planet and over all century-long periods will be easy to handle (statistically speaking, i.e. leaving aside the practical question of addressing the data). Overall we know how to deal with these "pure radiative transfer" spaces.

The reason why global radiative forcings could not yet be computed using 3D fields, multiple-scattering and high-resolution spectra is elsewhere: in the handling of spectroscopic data. This reason is quite simple. Let us consider one single radiative path γ at a given set of time, frequency and location. Along this path, we need to deal with the exponential of Beer-extinction law. For clarity, let us reduce φ to this simple extinction:

$$\varphi = \exp\left(-\int_0^{l_\gamma} k_a(s)ds\right) \qquad (2)$$

The integral over the path-length within the exponential is the absorption optical thickness. It accounts for the non-homogeneous field of the absorption coefficient k_a and this is where spectroscopic-databases are used: the absorption-coefficient is modelled as additive contributions of **a very large number N_t of molecular-state transitions**, $k_a = \sum_{j=1}^{N_t} h_{a,j}$. Databases such as GEISA or HITRAN include all the parameters required to compute $h_{a,j}$ as function of the local composition and thermodynamic state of the atmosphere. Altogether,

$$\varphi = \exp\left(-\int_0^{l_\gamma} ds \sum_{j=1}^{N_t} h_{a,j}(s)\right) \qquad (3)$$

and compared to Eq. 1 we now have to face two more integrals/sums, one over the length of the path and one over the molecular-state transitions. This would induce **no major difficulty if these two new integrals were linearly combined to the four preceding ones** (time, frequency, location, path). **But they are combined non-linearly via the exponential**.

I-b-2: Null-Collisions and nonlinear Monte Carlo

"*The extension of Monte Carlo methods to nonlinear processes may be impossible.*" (Cutiss, 1953) "*Monte Carlo methods are not generally effective for nonlinear problems mainly because expectations are linear in character.*" (Kalos ea. 2008) "*A nonlinear problem must usually be linearized in order to use Monte Carlo technique.*" (Chatterjee ea. 2014).

Such statements are typical of how the extension to nonlinearly-combined spaces has been perceived by the Monte Carlo community since the origin of the method. However, several successful attempts have been reported, the most successful of which is undoubtly the use of Null-Collision algorithms. A review of these techniques is available in (Galtier ea. 2013) and (Dauchet ea. 2016).

The starting point of Galtier et al. (2016) is precisely the use of a Null-Collision algorithm in order to suppress the nonlinearity of the exponential in Eq. 3 :

1- This exponential-nonlinearity is Taylor-expanded (Galtier ea. 2013), at the price of introducing a new infinite-dimension integration-space.

2- The power of each monome is deployed as a product of integrals (that can be sampled independently) (Dauchet ea. 2016)

Practically, the negative sign in the exponential is first suppressed by introducing N_t free parameters with the $\hat{h}_{a,j}$ constraint $\hat{h}_{a,j} > h_{a,j}$ (they will have an impact on the variance of the estimator and on the computational costs, but they introduce no bias). Then the exponential is Taylor-expanded:

$$E_{[t_1,t_2]} = \int_{t_1}^{t_2} dt \int_{-\infty}^{+\infty} d\nu \int_{Globe} d\vec{x} \int_{\Gamma} d\gamma \sum_{n=0}^{+\infty} \int_0^{l_\gamma} ds_1 \int_0^{l_\gamma} ds_2 ... \int_0^{l_\gamma} ds_n \sum_{j_1=1}^{N_t} \sum_{j_2=1}^{N_t} ... \sum_{j_n=1}^{N_t} \left[\frac{\exp(-\hat{\tau})}{n!} \prod_{i=1}^{n} \tilde{h}_{a,j_i}(s_i) \right] \quad (6)$$

where $\tilde{h}_{a,j} = \hat{h}_{a,j} - h_{a,j}$ and $\hat{\tau} = \int_0^{l_\gamma} ds \sum_{j=1}^{N_t} \tilde{h}_{a,j}(s)$. The deployment leads to

$$\varphi = \sum_{n=0}^{+\infty} \int_0^{l_\gamma} ds_1 \int_0^{l_\gamma} ds_2 \int_0^{l_\gamma} ds_n \sum_{j_1=1}^{N_t} \sum_{j_2=1}^{N_t} ... \sum_{j_n=1}^{N_t} \left[\frac{\exp(-\hat{\tau})}{n!} \prod_{i=1}^{n} \tilde{h}_{a,j_i}(s_i) \right] \quad (5)$$

and this is reported into Eq. 1 to give an exact formulation of the global radiative-forcing in which all integrals are now combined in a strict-linear manner:

$$E_{[t_1,t_2]} = \int_{t_1}^{t_2} dt \int_{-\infty}^{+\infty} d\nu \int_{Globe} d\vec{x} \int_{\Gamma} d\gamma \sum_{n=0}^{+\infty} \int_0^{l_\gamma} ds_1 \int_0^{l_\gamma} ds_2 ... \int_0^{l_\gamma} ds_n \sum_{j_1=1}^{N_t} \sum_{j_2=1}^{N_t} ... \sum_{j_n=1}^{N_t} \left[\frac{\exp(-\hat{\tau})}{n!} \prod_{i=1}^{n} \tilde{h}_{a,j_i}(s_i) \right] \quad (6)$$

Because of its linearity, Monte Carlo can address this integral formulation by sampling the corresponding infinite-dimension space and this is how Galtier et al. managed to evaluate radiative transfer quantities without precomputing the fields of absorption-coefficient at all locations and all frequencies. The spectroscopic data are handled by the Monte Carlo itself: they are sampled together with time, frequency, location and path.

By comparison with standard Monte Carlo approaches to radiative transfer, the main originality is therefore the repeated sampling of transition-index $j_1, j_2 \ldots j_n$. But which sampling probabilities should be used in order to minimize the variance ? In M. Galtier's PhD, this question was not addressed in depth (as will be done in Task 1), but still sufficiently to

- identify that the variance is highly sensitive to this choice of transition-index probabilities: importance sampling is essential;
- identify that this importance sampling is highly dependent on frequency: it is therefore difficult to rely on precomputed importance-estimates as the amount of corresponding data would be impracticable;
- find a first, non-optimized, but already practical solution: typical computation times for top-of-atmosphere fluxes are 0.1s to 12s on a standard laptop (Intel Core i7 2.8 GHz).

I-b-4: Evaluating sensitivities

One of the strengths of Monte Carlo methods is the ability to evaluate sensitivities simultaneously with the main quantity. In the present project, we will address sensitivities to a fractional change in the concentration of absorption gases. This means that when computing the global radiative-forcing for given atmospheric profiles, we will simultaneously produce the vector of its derivatives with respect to the mean concentration of each gas. The principle of Monte Carlo evaluating parametric sensitivities (or Jacobians) in radiative-transfer is quite simple (Lataillade ea. 2002, Roger ea. 2005). With Monte Carlo, the addressed quantity is viewed as the expectation of a random variable. Formally speaking, this expectation is an integral and this integral can be derived as function of any parameter to give another integral that can again be evaluated using Monte Carlo. The sensitivity evaluation is only **simultaneous** if the very same random sampling algorithm can be used for both the addressed quantity and its sensitivity. In Roger et al. (2005) this was shown to be possible whatever the parameter-type, and even when the parameter affects the integration domain.

However, despite of the complete generality of this proposition, depending on the contexts, the practice can be quite difficult, in particular as far as numerical convergence is concerned: nothing insures indeed that the relative variance of the sensitivity-estimate is comparable to that of the main quantity. Here, we will benefit of two ongoing works centred on evaluating sensitivities with Monte Carlo in 3D multiple scattering fields (both are described in El Hafi ea. 2018):
Preliminary work 1: In the frame of the High-Tune project, LAPLACE and CNRM produced updated versions of the star-engine library dealing with both shortwave and infrared radiation in cloudy atmospheres. In her ongoing PhD, N. Villefranque already added and tested sensitivity-estimates (using a simplified model for gaseous absorption, RRTM).
Preliminary work 2: In his ongoing PhD on nonlinear Monte Carlo, for heat-transfer in engineering applications, J.M. Tregan recently explored convergence issues associated to the use of null-collision algorithms. He observed that the variance of the sensitivity-estimate increases when reducing the amount of null-colliders (accelerating the sampling procedure) and found a way to bypass this difficulty by working on the integral-formulation.

b. Methodology and risk management

I-c-1 : Methodology

MCG-Rad deals with
- accurate spectral integration of radiative-transfer using GEISA or HITRAN spectral data bases;
- combination of such reference radiative-transfer integrations with space and time integration (whole atmosphere, years of GCM climatic simulations);
- evaluation of sensitivities to the amounts of absorbing gases.

This is possible thanks to the Monte Carlo algorithm of Galtier et al that bypasses the step of evaluating the gaseous absorption coefficient at each point in the 4D GCM-ouputs field: radiative-forcing can be computed by sampling jointly the spectroscopic databases and the GCM-outputs. **The whole project is therefore structured around Galtier's proposition: one objective will be its implementation** (development, testing and distribution of a code computing radiative-forcing for any GCM-output), **the other will be to explore means of accelerating it.**

When choosing the project organisation, the main three points were the following:
- Galtier et al's transition-sampling algorithm is available and was fully tested for application to clear-sky earth atmosphere (one time, one location on the globe, no-scattering);
- the star-engine library (developed by LAPLACE and Meso-Star) includes efficient path-sampling algorithms that have already been implemented for simulation of 3D multiple-scattering atmospheres at both shortwave and infrared frequencies, including first sensitivity-estimates;
- when combining these two sampling strategies it is expected that the strongest source of variance remains the sampling of transitions: the time required to compute globally and temporally integrate radiative forcings should should remain practical (which is not necessarily true for sensitivity estimates).

This means that implementation of Galtier et al's algorithm within the star-engine library can be started at the launching of the project and rapidly used for first climatic explorations. In the meanwhile the gathered team of radiative-transfer physicists and computer-graphics specialists will revisit extensively this algorithm, for which there is still a lack of theoretical perspective, in order to evaluate all acceleration potentials and explore the convergence of sensitivity estimates.

Maintaining a close relationship with neighbour communities: The initial proposition of Galtier et al. was not restricted to atmospheric sciences and most of the feasibility tests were oriented toward engineering questions, mainly heat transfer in combustion devices. We will therefore develop our radiative-transfer tools within the EDStar (http://edstar.lmd.jussieu.fr/), an easily usable open source environment that is commonly developed by LAPLACE, RAPSODEE and Meso-Star (https://www.meso-star.com). These two last partners are not officially part of the present project, but they are directly interested: RAPSODEE is a research laboratory involved in combustion heat-transfer and concentrated-solar engineering, in which gaseous radiation is essential; Meso-Star is start-up applying computer graphics technologies to the engineering of complex systems. A contract between LAPLACE and Méso-Star engages Méso-Star to provide an engineering support to LAPLACE for the development of the EDStar platform via the star-engine library (https://gitlab.com/meso-star/star-engine). Therefore, the Meso-Star engineers will be effectively involved in the present project (16 person.month), which is a guaranty that all developments will be capitalized and designed in coherence with needs expressed outside the climatology community.

Organizing our second circle of scientific expertise: In addition to RAPSODEE and Méso-Star, several partners have officially expressed their interest in the present project (see the letters of intention at http://www.lmd.jussieu.fr/~jldufres/MCG-Rad/Letters/ login:MCG-Rad passwd:ANR2018). In the climate community these are CNRM, LOA and LaMP/OPGC as typical representatives of DEPHY and partners of the projects High-Tune and PNTS « Nouvelles approches en transfert radiatif ». In the engineering community these are CETHIL, PROMES and Institut Pascal that are involved in both nonlinear Monte Carlo theory and in the development of the EDSTAR platform. We will also invite the Planéto team at LMD for application of line-sampling Monte Carlo algorithms to planetary atmospheres. In addition to the participation of these 10 to 15 researchers to the project meetings, we will devote two sessions of the annual Roffiac Seminar (http://edstar.lmd.jussieu.fr/roffiac) to a close scientific interaction with them.

I-c-2 : Risk assessement

We do not anticipate a fundamental difficulty when adding scattering and spatio-temporal integration to the algorithm proposed by Galtier et al. We are therefore confident that we will be able to develop and provide a code that allows to compute integrated radiative fluxes and radiative forcings at the top of atmosphere for any GCM-output with a “reasonable” computer time. In Galtier et al. (2016) these times were of the order of 10s on a laptop for a single clear-sky atmospheric profile. Without any change in the algorithm, we expect an increase up to 100000s when adding scattering and time/space integration, including the computer cost of accessing large GCM outputs. However, we cannot be that confident concerning the sensitivity-estimates (see preliminary-task 2 in section I-b-4). The major difficulty of this project and the main risk is therefore to make too little progress in the sampling strategies and in reducing the computer costs. We would then still observe that accurate radiative-forcing computations are accessible with first-level computation resources, but evaluating gaseous-aboundance sensitivities with a similar accuracy would be computationally very demanding. This would reduce the size of the community interested in our work and would not create a paradigm shift, but that would not prevent the computations at the center of this project (Task 2). To limit this risk we chose to address the improvement of the sampling strategy by two approaches, one more “physically based” (benefiting of three ongoing PhDs dealing with the convergence of sensitivity-estimates) and one more “computer graphics based” (Task 1). The work of the new PhD student, recruited as part of this project, will be an important piece to successfully conduct this interdisciplinary research. The beginning of the thesis is planed in September 2019 and we believe that we have enough time and we are enough involved in four “écoles doctorales” to find a student that will be interested and will be able to conduct this challenging task.

II. Project organisation and means implemented

a. Scientific coordinator and its consortium / its team

II-a-1 : Complementarity of the consortium

The project includes both consolidation of novel existing methods, development of new methods, software implementation of new techniques and use of these developments to advance climate science questions. It is a multidisciplinary project that gather experts in Monte-Carlo methods, radiative transfer, computer graphics, geometric modelling, spectroscopy data, climate modelling and climate change. This allows to build an end-to-end project, from fundamental developments to software tools for high performance simulations applied to climate challenges. The consortium is composed of 3 internationally recognized laboratory and gathers scientists which expertises cover the whole chain of competency required to implement this ambitious project. These partners have been working together for more than a decade now. This fruitful and ongoing collaboration already resulted in co-supervised thesis, cosigned papers and the development of radiative transfer codes for the atmosphere of the Earth and other planets (Mars, Venus): LMD/LAPLACE: 2 PhD, 10 publications; LAPLACE/IRIT: 2PhD, 4 publications, Méso-Star created by doctors from both laboratories.

<u>Partner 1: LMD/IPSL, Laboratoire de Météorologie Dynamique</u> (UMR8539, CNRS/ Sorbonne Université / Ecole Polytechnique / Ecole Normal Supérieur). LMD is a leading laboratory in the theory, observation and numerical modelling of atmospheric dynamics and physics. It has a strong expertise in atmospheric radiation, in particular for remote sensing and numerical modelling of other planetary atmospheres. Two teams are involved in the project. The first, EMC3, hosted by “Sorbonne Université”, develops the general circulation model LMDZ, the atmospheric component of the IPSL Earth-System-Model, and studies the processes that govern climate variability and change. The second team, ABCt, is hosted by Ecole Polytechnique and specializes in the study of interaction between atmosphere and radiation in order to establish long time series of essential climate variables observed by remote sensing. It has developed several radiative transfer codes and manages the GEISA spectroscopic databases.

<u>Partner 2: LAPLACE, Laboratoire PLAsma and Conversion d'Energie</u> (UMR5213, CNRS, Université Paul Sabatier /Institut National Polytechnique de Toulouse) is the first French concentration of research in the field of Electrical Engineering and Plasma nationally with 160 fulltime researchers and a similar number of PhD students and postdocs. Within LAPLACE, the research activities of GREPHE (Groupe de Recherche Energétique, Plasma et Hors-Equilibre) lie at the interface between engineering, physics and fundamental technological issues are also addressed. Most of the physical systems that are studied are non-equilibrium or far from equilibrium. Modelling and simulation are an essential part of the group’s activities and the group has developed an expertise in the modelling of non equilibrium plasmas (for a wide variety of applications) and radiation transport. LAPLACE has a collaboration contract with the Meso-Star start-up to provide an engineering support for the development of the EDStar platform.

<u>Partner 3: IRIT-STORM, Institut de Recherche en Informatique de Toulouse</u> (UMR5505 CNRS, UT3/UT1/UT2/INPT). IRIT is the lead center of research in Computer Science in Midi-Pyrenees and one of the bigger computer science research center in France. With 270 researchers and research professors, on a global workforce of 700 people, research activities cover the whole range of computer science, from theoretical to finalized research. Within IRIT, the STORM research group, develops its activities in the field of Computer Graphics. From geometric modeling to lighting simulation for realistic rendering, the STORM people aims at developing computationally efficient models and tools for digital content creation and edition. By developing robust mathematical models, data structure and low complexity algorithms, the STORM research group address theoretical and practical bottlenecks on efficient sampling and reconstruction strategies for Monte Carlo simulation algorithms as well as multi-scale, feature guided, geometric reconstruction from acquired data.

II-a-2 : The coordinator : involvement and experience

The project will be coordinated by Jean-Louis Dufresne, 56-year old, senior scientist (DR1) at CNRS, LMD/IPSL. He is expert in climate modelling, climate change studies and radiative transfer computation. He has been involved in more than 10 national and European projects, has been task leader in several of them, and has coordinated the large Numerical Model ANR project “Convergence en Science du Climat à l'ère du Big Data et des challenges de l'Exascale” (2013-2018). He was head of the Climate Modelling and Climate Change team at LMD during 8 years (2004-2011), head of the IPSL Climate Modelling Centre and deputy director of IPSL during 9 years (2008-2016). IPSL is a federation of 9 laboratories in the Paris area whose research topics concern the global environment. He acts as lead author of the chapter on “Long-term Climate Change” in the last IPCC-AR5 report (2013). He contributed to about 90 publications in peer-reviewed journals, h-index 36, named Highly Cited Researcher (Clarivate Analytics & Web of Science) 2015, 2017.

After coordinating the development of the IPSL climate model for 9 years, he now focus his research on the role of the radiative exchanges in the climate system. He will act as lead author of the chapter on «The Earth's energy budget, climate feedbacks, and climate sensitivity » of the next IPCC assessment report (IPCC-AR6), he coordinates the IPSL contribution to the radiative forcing inter-comparison project (RFMIP). He is familiar with multidisciplinary approaches. He was pioneer in coupling a climate model with a carbon model and analysing the climate-carbon feedbacks, an approach that is now at the heart of the so-called “Earth-System Models (ESM)”. In the ANR “Convergence” project he coordinated, experts of climate sciences, high performance computing, data analysis and management join their efforts to develop a platform capable of running large ensembles of simulations with a suite of models, handling the complex and voluminous datasets generated and facilitating the evaluation and validation of the models. This platform is now operational on two national computer centres (TGCC and IDRIS). His long-standing collaboration with the Laplace laboratory has already made possible successful collaborations between engineering and atmospheric sciences research community.

II-a-3 : Curriculum of most involved persons

Richard Fournier, 52-year old, professor at University Paul Sabatier and researcher at LAPLACE-GREPHE (Research Group in Energetics, Plasma and Non-Equilibrium Phenomena). He focuses on the statistical approaches to complex-systems engineering, for energetic, atmospheric and biological applications. Together with Mouna El Hafi (Mines Albi), and in collaboration with Meso-Star (http://www.meso-star.com), he animates the EDStar Plateform that makes available to engineers a set of open-source libraries inheriting of recent statistical-physics and computer-graphics advances (http://edstar.lmd.jussieu.fr/). He also animates the axe "Approche statistique du rayonnement" of the GDR ACCORT, federating Thermal Radiation research at the national scale (http://www.gdr-accort.cnrs.fr). 55 publications in peer-reviewed international journals.

Stéphane Blanco, 48-year old, senior Lecturer at University Paul Sabatier and researcher at LAPLACE-GREPHE. He has 25 years research in non-equilibrium thermodynamics, radiative transfer computation, and statistical physics for energetics and biological applications. He was head of the Master of teaching physics at the university of Toulouse until 2014. He has contributed to 45 publications in peer-reviewed rank A journals.

Mathias Paulin, 50-years old, Professor at Université Paul Sabatier - Toulouse 3 and researcher at IRIT-STORM (Structural models and tools for computer graphics). Mathias Paulin is the leader of the IRIT-STORM research group. His research focus on physically based rendering for digital image synthesis with two main approaches : Monte Carlo methods and path-space light transport simulation on the one hand, realtime rendering for interactive applications on the other hand. He his deputy director of the GDR IG-RV from CNRS-INS2I (http://http://icube-web.unistra.fr/gdr-igrv) and director of the computer science master program at the university of Toulouse. He was a former member of the French National research Council (section 7) from 2012 to 2016. He his the leader of the ANR project CaLiTrOp (Comprehensive Analysis of Light Transport Operators for image synthesis - ANR 16-CE33-0026-01). 60 publications in peer-revieweed international journals.

Nicolas Mellado, 30-years old, CNRS Researcher at IRIT-STORM , University of Toulouse. His research interests include point cloud processing, multiscale analysis and 3D registration. Nicolas Mellado received a PhD degree in Computer Science from the University of Bordeaux, France, December 2012. He worked as a post-doctoral researcher at University College London with Niloy J. Mitra on 3d point-cloud registration and reconstruction, and at University of Toulouse on developing color processing techniques using point-based geometric approaches. 13 publications in peer-revieweed international journals.

Cyril Crevoisier, 40 years-old, CNRS researcher at LMD. His reasearch deals with the analysis of measurements of atmospheric radiation made from space by various instruments, especially in the thermal and shortwave infrared, in order to derive information on the evolution and variability of climate variables. He is involved in the design of new space missions, such as the new generation of hyperspectral sounders (IASI-NG of CNES), the active mission Merlin of CNES/DLR, the passive SWIR mission Microcarb of CNES. Since 2007, he has been teaching atmospheric radiative transfer at Sorbonne University. He leads the ABCt team of LMD and chairs the Atmospheric Science Advisory Group of CNES.

b. Means of achieving the objectives

II-b-1 Project structure and description:

The project is organized in 4 major tasks. Task 0 ensures the coordination of the project, while the three other tasks are devoted to the scientific questions and will deliver the main results:
Task 1 (“*Transition-sampling and computer-graphics*”) deals with all the theoretical aspects of Monte Carlo integration, from the detailed geometrical understanding of the integration domain, to the design of efficient sampling strategies.
Task 2 (“*Radiative forcing from 4D fields*”) gathers the climatic dimension of the project, from deciding and evaluating the successive test-cases, to the detailed picturing of how these new simulation potentials may be used by the climate-change community.
Task 3 (“*Software development*”) deals with the coding and testing of a succession of Monte Carlo solvers with gradually enhanced calculation potentials and interfaces.

Task 0 : Management and outreach.

Task leader: J.-L. Dufresne

Because the tasks are interconnected and because of the interdisciplinary nature of this project (climatology, radiative transfer numerics, computer graphics and software engineering), the scientific animation is crucial. In particular, the three scientific tasks extend in parallel along most of the project and the closer we manage to relate them, the better in terms of both technical achievement and conclusion strength. The project coordinator, partner PIs and task leaders will form the project steering committee. This committee will be convened to assess the advancements of the project at least three time a year (by teleconference twice a year and concomitantly with the project meeting) and a status report, mentioning the main results, the difficulties encountered in the work progress will be established.

Formal project meetings will be organized on a yearly basis. They will be organized in Paris or in Toulouse. The 1st formal meeting will take place in Toulouse. All participants will be required as possible to physically participate to build a tight group since the beginning of the project despite the varied expertise. Colleagues with whom we already work in other projects on radiative transfer computation (see section I-c-1) will be invited. The time frame for each formal meeting is one full day plus half a day for the steering committee. Two workshop will be organised in the frame of the Roffiac Seminar. Intermediate less formal meetings will take place in-between annual formal meeting concerning each scientific individual task.

This task is also in charge of the dissemination of the MCG-Rad results. Results of the project will be regularly presented to the radiative transfer community during the course of the project, to the DEPHY community that developed the physical parameterization of the French atmospheric climate and weather forecast models and to the radiative forcing international community in the frame of the RFMIP meeting.

Milestones and deliverables: (M0 = 1/11/2018)

Number	Description	Date
	TASK 0: Management and outreach	
D0.1	Kick-off meeting, project mailing list and web site	M1
M0.1	Steering committee meetings	3 time a year
D0.2	Project meetings	M12, M24, M36
D0.3	Workshops	M10, M34
D0.4	Project reports	M18, M30, M48
D0.5	End of project meeting	M47

Success criteria: The good achievement of the project

Risks and envisaged solutions: No risk is foreseen as most of the partners have a long experience of cooperation. During these common past works, various practical difficulties have appeared and have always been solved.

Task 1: Transition-sampling and computer-graphics

Task leader: Mathias Paulin

Summary : *Data analysis and modeling for revisiting of the radiative-forcing formulation. Sampling strategies and variance reduction techniques. Spectral database modeling, processing and storage for efficient line-sampling strategies.*

This task deals with the optimization of Galtier et al's Monte Carlo algorithm, including sensitivity estimates. In this context, when we refer to a **modeling** effort, this never means that we change the physical model itself. We only model the multi-dimension integral of Eq. 6 in order to reduce the variance of the corresponding statistical estimate. **The choice that will be finaly made of one model or another will not change the estimated quantity**: it will introduce no bias. Only the convergence speed (the number of required samples) and the computation requirement (the computation time associated to each sample) will be impacted.

Task 1.1 Modeling and iterating over a high-dimensional integration domain.

One of the most important question that arises when trying to integrate a function over a very high dimensional space concerns the relationship that can be highlighted between the different dimensions and how to reduce the apparent dimension to its intrinsic kernel. In our case, within the null-collisions and non linear Monte Carlo framework described in section I-d-2, this space is infinite-dimensional and we aim at defining a way of sampling and estimating all the integrals of this formulation that maximizes the overall computational efficiency while lowering the estimator variance.

Compared to standard radiative transfer, the null collision formulation of Galtier et al. introduces a new infinite dimensional space in the integration domain due to the linearization of the exponential function of Beer extinction. This space is the n-ary Cartesian power of the transition space for each degree n of the linearization expression. Exploring such an infinite dimensional space might be done naïvely, by sampling or exploring the primal space (i.e. the space of all degree of the linearization and all the transition parameters for each degree), or by exploring a subspace topologically equivalent but of lower dimension. The dimension of this subspace is called the intrinsic dimension of the primal space.

We will of course start from the current implementation that is already practical. But it strongly relies on heuristics defined thanks to the Radiation Physics expertise of the authors. What if the topological study leads us to explore structural changes such that these chosen heuristics are no more efficient? Can we afford to fully reformulate such a complex integration problem in view of designing faster and still robust exploration strategies? These are the main questions that we will study, in a first phase, using simultaneously the approaches of the radiative-transfer and computer-graphics communities:

1. Characterizing the integration space with a zero-variance approach

When radiative-transfer physicists address such questions, the starting point is the literature of all available approximate models. The integration spaces are considered one by one, successively, and at each step the physical quantity corresponding to the remaining integrals is used to defined an ideal sampling-probability: the one that would lead to a zero-variance (Assaraf and Caffarel 1999, Dauchet et al. 2013). Then, this ideal sampling-probability is approached using the best approximate model available. For illustration, let us reduce the radiative-forcing problem to only infrared frequencies and consider the integral over the globe: when picking a location on the globe, the remaining integrals define the local value of the cooling-to-space. The ideal-sampling of global location should therefore follow rigorously the dependence with location of cooling-to-space. The question then becomes: can we find a simple radiative transfer model evaluating (even roughly) this dependence with location of the cooling-to-space and can it be turned into a computationally practical location-sampler (i.e. fast enough and using a limited amount of precomputed data).

With such a simplified example, we see nothing more than the building of an importance-sampling approach using available models. But with the same example, if go back to an integration over all frequencies, including shortwave, then the quantity to model becomes the top-of-atmosphere net-flux and it takes both positive and negative values as function of location. So, the sampling probability cannot be simply designed to match this net-flux. Such difficulties can only be handled by changing the integral-formulation (changing variables or inverting integration orders) so that the optimization-objective finally matches a better defined modeling-question.

We will explore all such reformulations, at slow speed, throughout this task. One of them with the highest potential consists in inverting the frequency integral and the sum over transitions. This is possible because the space of multiple-scattering paths is independent of frequency: all the multiple-scattering paths exist whatever the frequency and only their weights depend on frequency. We can therefore chose to define the frequency-integral as the last integral, and then a quite efficient model will simply be a convolution-product of Lorentz line-profiles. This looks promising but the questions that will arise concerning the other integrals, as a consequence of this inversion, will of course be widely open.

2. Characterizing the integration space by data analysis and modeling.

Facing the very same question, the computer-graphics approach starts with geometrical considerations. What is difficult in finding a lower dimension subspace of the n-ary Cartesian power of the transition space is to apprehend how this space is filled according to the various parameters of the gases and transitions and according to the degree of freedom introduced by the null-collision formulation. As the atmospheric conditions one might encounter in the computation of Radiative Forcing could be of high variability, we might not model once for all this subspace but will need to construct a model from a given set of gases and transition parameters. Thus, we will concentrate our work on building an intuitive model of the structure of this space and define how it could be explored and sampled.

By using multi-resolution analysis of the primal transition space, in which each transition is associated to a point and each atmospheric condition is a set of points, and by carefully taking into account the needs of a line-sampling Monte Carlo estimator, we want to find a way to transform this primal space into a subspaces (e.g. manifolds) that will allow efficient representation of transitions but also of sets of transitions representing gases. This sub-manifold being then locally Euclidian, exploring the neighborhood at a point will be simplified and might serve as an efficient line-sampling strategy. To construct this multi-resolution model of the space of transitions, while allowing computationally efficient access to individual transitions at the finer scale we will study how to generalize some computer graphics related results on data analysis and modeling, such as (Mellado ea. 2014, 2016), to higher dimensional spaces. A similar method has already been developed in Computer Graphics for spaces of lower dimensions (Jakob ea. 2012), demonstrating the feasability of the approach on a simple case.

Output of this task, both from Radiation Physics modeling and from Data Analysis and Modeling will serve to define a concrete sampling algorithm in Task 1.2. Each proposal will be evaluated with respect to the resulting estimator variance, computational efficiency (both in time and memory) and scalability.

Task 1.2 Multiple Importance Sampling.

Efficiency, convergence rate and variance are central concerns when building a scalable Monte Carlo estimator such as the one needed for Radiative Forcing. In our case, it is clearly identified that the most difficult problem is due to the linearization of the exponential of Beer-extinction law. Building a low variance estimates of this integral relies on the sampling strategy used for choosing the transition to evaluate. Defining this strategy is well described in Galtier et al. (2016) but, as seen previously, lacks robustness with respect to the transition parameters.

Building on the study above (Task 1.1), we want to make the choice of the transition less critical by introducing a multiple importance sampling scheme taking advantages of the multi-resolution representation of the transition space. Importance sampling used in the Beer-extinction line-sampling method aims at lowering the variance while improving the efficiency of the estimator. The estimates of eq. 5 (equation on φ page 5) could then be written as $\hat{\varphi} = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}$ with *p(x)* the carefully constructed distribution associated to the line sampling strategy for a given set of transition parameters. *p(x)* is quite sensitive to the transition parameters and, if not adequate, could lead to high variance of the estimates.

Starting from this observation, we want to make the estimates less dependent on the definition of *p(x)* over the whole transition space. Instead of using a unique distribution, and thanks to the Multiple importance sampling -MIS- (Veach ea. 1995) method and to the results of task 1.1, we will construct several sampling strategies to make the estimator robust with respect to the transition data. The central point of MIS is to combine several sampling strategy, defined by different distributions, in a way that the estimator remains unbiased while its variance is lowered. In our case, given *n* distributions describing different properties of the transition space, we could build several estimators $\hat{F}_e$, each with an associated sampling distribution $p_e(x)$ used to generate n_e samples. Then by combining each estimator with, e.g, the balance heuristic (Veach ea. 1995), we can build a multi-sample estimator $\hat{F} = \frac{1}{N}\sum_{e=1}^{n}\sum_{j=1}^{N}\frac{f(X_{e,j})}{\sum_k c_k p_k(X_{e,j})}$ where $X_{e,j}$ is the j^{th} sample obtained by the strategy *e* with distribution $p_e(x)$, $N = \sum_e n_e$ is the total number of samples and $c_k = \frac{n_k}{N}$ is the fraction of samples from p_k.

MIS is proven to reduce the variance of an estimator while being more robust to some imprecisions on the distributions p_e. The central difficulty of MIS computation is to clearly identify all the sampling strategies and their distributions, and how to build an efficient heuristic (the balance heuristic is not always the best one) with respect to the given strategies. To build such a set of sampling strategies, several works aim at using a synthetic, precomputed or iteratively refined, views of the integrand properties (Pajot ea. 2011) and optimize the heuristic with respect to the expected variance of the estimator. Our work in this task will follow the same approach.

The results of both research directions of task 1.1 will serve as basis for constructing an efficient MIS scheme.
1. Radiative Physics modeling will define the choice of dimension-dependent distributions and their physical meaning will guide us toward meaningful MIS schemes. Unfortunately, such distributions could be difficult, or computationally expensive, to sample. Our Data Analysis and Modeling procedure might then help in developing computationally efficient sampling algorithms for these distributions.
2. In another research direction, the generalization of Globally Adaptive Control Variate (Pajot ea. 2014) to integration over high dimensional domain will also benefit from our Data Analysis and Modeling procedure. We expect that our multi-resolution approach will define the Control Variate, that can be integrated exactly while the remaining part of the integrand will be estimated using a MIS-Based Monte Carlo estimator.

Output of this task will be a set of importance sampling strategies, eventually with the associated procedure to preprocess spectral data or to generate the samples, and their associated Multiple Importance Sampling combination schemes. Each proposal will be evaluated with respect to the resulting estimator variance, computational efficiency (both in time and memory) and scalability.

Task 1.3 Storage and processing of spectral database for efficient Radiative forcing.

While tasks 1.1 and 1.2 are research oriented tasks, the objectives of Task 1.3 are much more practical. Integrating operational constraints on using Monte Carlo estimators for Global Radiative Forcing will lead to making compromises between time, space, accuracy and scalability of the proposed algorithm. The objectives of this task is to experiment on the proposal from previous tasks using real world data. These experiments will serve to specify and prototype an operational workflow for Global Radiative Forcing that clearly identifies the pre-computation and storage needs, the runtime constraints and the validation procedure for the obtained results.

<u>**Deliverables :**</u> (M0 = 1/11/2018)
- Report on different expressions of the Monte Carlo estimator (D.1.1 - Task 1.1.a): M18
- Report on multi-resolution analysis and modeling of spectral data (D.1.2 - Task 1.1.b): M18
- Report on Multiple Importance Sampling schemes (D.1.3 - Task 1.2): M28
- Software specification and prototype for spectral data analysis, modeling and sampling (D.1.4 - Task 1.3): M34

<u>**Success criteria:**</u>
- The data analysis and modeling procedure allows to build an efficient control variate for importance sampling.
- Cooperation between radiation-physics and data modeling results in the definition of a robust MIS scheme with bounds on maximal variance.

Risk assessment:
- Huge computational resources (both time and space) needed to achieve the data analysis and modeling work. As this will be a preprocessing step for the operational system, it will not have an impact on task 2 nor 3. The data modelling realized in this task will then get available for the users communities.
- Unable to build and efficient control variate over the (concentration pressure, temperature)-parameterized transition space. In this case, we will partly go back to Galtier's intial proposition an think of importance sampling on individual gases.

Task 2: Radiative forcing from 4D atmospheric fields

Task leader: Jean-Louis Dufresne

Summary: *Validation of the radiative forcing computed by the codes developed in task 3. Building a statistical 3D geometry of clouds from GCMs outputs.*

First, the code developed in Task 3.1 will be validated and assess for a few numbers (hundred) of atmospheric profiles with clear sky conditions (no clouds). This will be done in the frame of the RFMIP project which has a dedicated research themes on testing radiative codes (https://rfmip.leeds.ac.uk/rfmip-irf/). The atmospheric profiles are produced by general circulation models (GCMs) with a typical horizontal resolution of 50 to 200 km and a vertical resolution of a few tens to a few hundreds meters. The goal of Task 2.2 is to provide the necessary information to a statistical description of the 3D geometry of clouds for each atmospheric profile from the information provided by the GCMs. At the end of the project, the last task (T-2.3) is to perform, analyse and promote the computation of the radiative forcing over the whole globe, over climatic durations, considering the 3D radiative effects.

Task 2.1 : Validation and analysis of the radiative code in clear sky conditions or with the plan parallel assumption

The first versions of the code will allow us to compute radiative flux for any atmospheric profile with no clouds or with clouds with the plan parallel assumptions. This assumption is currently used by all GCMs and all studies that estimate the radiative forcings. This is the case for the RFMIP project in which we will participate. This will allow us to validate and promote our radiative code.

For the MCG-Rad project, the most critical part of the validation will be for clear sky conditions as it is the most innovative part. We will do a first validation by comparing the developed model with the already existing reference line-by-line model 4A developed at LMD for a few number of atmospheric profiles. Then we will participate in the IRF-GHG research themes of the RFMIP project by calculating the radiative flux and radiative forcings for the required 100 cloud free atmospheric profiles and making the results available for the climate community. This will allow us to make a precise validation with the other reference models at the international level. This will also be an opportunity to verify that the average of the 100 independent Monte-Carlo computations of these profiles is equal to the single Monte-Carlo computation that directly include these 100 profiles, although the later is almost 100 less CPU time consuming that the former.

The method we use to simulate multiple-scattering has already been fully validated for one single frequency or spectrally integrated using a simplified representation of gaseous absorption (RRTM) within ongoing projects (ANR High-Tune and PNTS, El Hafi et al. 2018 and publications in preparation). For simulations over a very wide spectral domain, we will do the same type of computation as above but with profiles that contains aerosols and will compare our results with the other reference codes within the IRF-AER research themes of the RFMIP project. In the margins of this project we will also propose to the other participants to perform some simulations with simple idealized clouds and the plan parallel approximation. We expect that our code will provide more accurate results compared to the other codes as the Monte-Carlo Methods are recognized to better handle scattering phenomena.

Task 2.2 : Building a 3D structure of clouds using GCMs outputs

In each grid cell of an atmospheric GCMs, and therefore in the input files we will use, the clouds are described by their fraction, their amount of water and some other physical (water phase, drop size, etc.) and optical (single scattering albedo, phase function, etc.) properties. But there is almost no information on the horizontal distribution of clouds, their subgrid heterogeneity and how they overlap on the vertical.

All the radiative codes used in current GCMs and global climate studies make the plan parallel assumption. However, as we wanted to take into account the 3D radiative effects, we need to build a statistical description of the 3D geometry of clouds from the information provided by the GCMs. Realistic 3D clouds fields can be generated by high resolution fluid mechanics models, by stochastic models, or by a mix of the two approaches (Szczap et al., 2014). However, to our knowledge there is currently no easy way to impose that these synthetic clouds also fulfil all the cloud properties simulated by the GCM. Monte Carlo Independent Column Approximation (McICA, Pincus et al., 2003) methods still use the plan parallel approximation but simulate a vertical overlap of clouds fully consistent with cloud properties provided by the GCMs. We will complete this approach by adding some 3D geometric aspects, even imperfect. We will use observations provided by lidar and radar instruments aboard satellites which give a sectional view of clouds in the atmosphere supplemented by some results of synthetic clouds. We will validate our results with satellite observations of the angular distribution that have been established for a large variety of cloud types.

Task 2.3 : Performing and promoting the 4D computation of the radiative forcing at global scale

This task aims at testing, assessing and promoting the final version of the code. This will be the first attempt to compute the radiative forcing of multiple greenhouse gases at global scale directly from a reference radiative model. We will consider all the forty greenhouse gases for which we have concentration data for the recent period and for the future (Meinshausen et al, 2017) and we will compute the forcing relative to the preindustrial period (year 1850).

a- <u>Current radiative forcings:</u> In a first step we will estimate the radiative forcing of greenhouse gases for current climate. For doing so, we will use the ECMWF ERA5 atmospheric reanalysis. Meteorological reanalysis are estimates of the 4D (3D geometric space + 1D time) state of the atmosphere obtained by a meteorological model that assimilate the largest possible amount of observations. Although they are model results, ERA5 reanalysis are considered as the current best estimate of the meteorological state. They will be available each hour on a horizontal resolution of 30 km on 137 levels, from the surface up to around 80 km, from 1950 to date. The amount of data is huge and the essential meteorological fields will be transferred at the IPSL data centre. We will compute the radiative forcings for the whole available period for each individual greenhouse gas. We will also estimate the radiative forcing of gas together and estimate the non-linearity due to the overlap between absorption lines.

b- <u>Recent and future radiative forcings:</u> In a second step we will estimate the time evolution of the radiative forcing of the same forty gases since 1850 to date, and in the next 100 years for the 8 scenarios of future emissions proposed by the ScenarioMIP project. This work will be done in the frame of the sixth phase of the CMIP project that will provide major inputs to the next IPCC assessment report. A large subset of the CMIP6 data will be transferred to the IPSL data centre and we will use them to compute the radiative forcings for the whole “historical” and “future” period (1850-2100) for each individual greenhouse gas. These new estimates will be published and compared to previous estimates based on less accurate approaches. As we use the atmospheric profiles of an ensemble of climate models, we will also investigate how the radiative forcing estimates depend on these data, a study which has not been done yet.

<u>**Deliverables:**</u> (M0 = 1/11/2018)
• Contributing to the RFMIP project with the first code version and report on its validation and publication exploiting these results (D2.1 - Task 2.1): M18
• Algorithm that provides a statistical description of the 3D geometry of clouds from GCM outputs and corresponding publication (D2.2 - Task 2.2): M30
• Computing the radiative forcing at global scale with the final (3A) version of the code and dissemination of the results (D2.3 - Task 2.3): M45

<u>**Success criteria:**</u>
• accuracy of the computation compared to other reference model for clear sky atmosphere
• possibility to use the developed code as a common tool to estimate the radiative forcing at global scale
• paradigm shift opened by the possibility of estimating radiative fluxes and radiative forcings with a fast Monte Carlo model that will be a reference for both the spectral aspects of gas absorption and 3D aspect of scattering.

<u>**Risk assessment:**</u> The developed code remains computationally expensive. This will limit the interest of the code for a large community but will still allow us to compute the radiative forcing with a code that is a

reference for both gas absorption and 3D radiative effects as we can have access to large computational equipments if required.

Task 3: Software development

Task leader: Richard Fournier

Summary: *Development and assessment of a radiative tool for the climate community. Recoding of Galtier's codes using the star-engine library and adding sensitivity evaluation. Interfacing with GEISA/HITRAN. Interfacing with GCM outputs. Implementation of the sampling strategy of Task 1.3.*

In a first phase (Tasks 3.1 and 3.2) the algorithm of Galtier et al will be implemented as is, that is to say independently of the better-understanding and technical solutions resulting of Task 1. The only algorithmic change will be to add the evaluation of sensitivities (algorithm already available, see preliminary work in paragraph I-b-4). In the second phase (Task 3.3), the objective will be precisely to upgrade the software according to Task 1-conclusions. This division guaranties that in any case, although not optimum, a new software will be distributed to the community and that this software will meet the principal objective (evaluating radiative forcing at the global scale, integrated over climatic durations).

According to this development strategy, five versions of the radiative-transfer code will be delivered, with graduated application and distribution potentials. The first version impacts the project strongly as both early Task 1 and early Task 2 depend on its achievement. But this first version is essentially the recoding of available softwares in a well known environment and the corresponding development-time is accurately predictable.

Task 3.1 : Rewriting M. Galtier's PhD code

a- Identically, with sensitivities: Here, our starting points are (i) the algorithm of Galtier et al. (ii) a sensitivity evaluation algorithm designed and fully tested by J.M. Tregan (see preliminary work)

The objective is to recode these algorithms under the EDSTAR environment, using the star-engine library (doing so, later extension to multiple-scattering and the interfacing with LES-outputs will be straightforward). This task will start immediately at the launching of the project. The inputs formats will be simplified as much as possible. We will fix a version of one of the spectroscopic databases and attach it to the code. The atmospheric fields will be restricted to clear-sky single columns. This rigorously corresponds to the test-cases made by M. Galtier during his PhD and Post-Docs (only the infrared results are presented in Galtier et al, but tests have been made at both infrared and shortwave frequencies). This version of the code will not be distributed [**version 1A**]. Code validation, including sensitivity evaluation, is reported to Task 2.1 (RFMIP inter-comparison).

b- Extension to 4D fields including scattering (still with the same line-sampling strategy).

This software extension should be straightforward : the star-engine library includes already the functions required to input 4D LES-fileds and sample both time and multiple-scattering optical paths. However, code validation will be time demanding. No reference is indeed available for such 3D/4D multiple-scattering simulations using high-resolution spectra. The only possible cross-validations will be

- against approximate simulations using k-distributions (simulations already available, for instance as part of the High-Tune Project),
- against exact simulations on limited parts of the spectrum (simulations already available at LOA as part of the PNTS project).

This version of the code [**version 1B**] will be distributed to the cloud community without waiting for the optimized-sampling of Task 3.3. As far as GCM applications are concerned, version 1B will also include the possibility of sampling a single-column among those of RFMIP inter-comparison according to any climatically meaningful distribution. This will simulate the sampling of longitude an latitude without requiring to address the memory constraints associated to the reading of true GCM-outputs. We will therefore start evaluating the performance of global radiative forcing computation before Task 3.2 is completed.

Task 3.2 : Accessing spectroscopic and climatic data

a- Interfacing with GEISA and HITRAN: The GEISA and HITRAN spectroscopic databases will be interfaced to our radiative-transfer code so that the user is free to choose any version of these databases. Some of the line-sampling parameters of Galtier et al are precomputed from the database: we will therefore include tools that allow to either recompute these parameters using a new database-version, or check that a previous version of the parameters is still compatible. Theoretically speaking, water-vapor continuum should also be adjusted to the database version. This question will not be addressed directly. On a standard basis, we will keep the continuum model of Clough et al (1989) without adjusting it to the retained database, but we will allow the use of external functions for computation of continuum in case new continuum versions become available. For this version of the code [**version 2A**] detailed documentation will be provided concerning all the practical steps (required pre-processing) and theoretical considerations (water vapor continuum) associated to the switching from one database version to another. The same document will also clarify the possibility of using HIGHTEMP for planetary and combustion applications (mainly in terms of memory requirements).

b- Accessing GCM-outputs: We need to be able to access the 4D fields data produced by long duration climate-simulations. The data-format question was already addressed by the international community and the tools that we will produce will be easily adaptable to most GCMs (except for the inclusion of the 3D-cloud respresentation of Task 2.2). The only true question in this task will be the radiative-path and time sampling strategy: do we sample only one radiative-path per time or do we conclude that the computational constraint associated to the reading of one atmospheric field at one date justifies that we sample several paths per date. This question is not at all trivial as it depends on how much the variance is reduced by sampling more paths versus how much time is spent accessing the fields. We will address this question on the same theoretical basis as we recently did in S. Weitz et al for temporal-integration of concentrated-solar radiation. Once interfaced with GCM-outputs, the code will be distributed to the climate community [**version 2B**]. At this stage, the original line-sampling algorithms of Galtier et al will be fully adapted to global radiative-forcing computation, but without any inclusion of the outputs of Task 1. The report evaluating the performance of global radiative forcing computation will be updated correspondingly.

Task 3.3 : Including the validated outputs of Task 1.3

a- A new optimized line-sampling strategy: Here we will implement the algorithms designed in Task 1. The interface with spectroscopic and climatic data will be unchanged but the precomputed line-sampling parameters will differ. The associated tools will therefore be recoded and the documentation about the switching from one database-version to another will be updated. The corresponding version of the code [**version 3A**] will be distributed to the climate community as the final code-version of the MCG-Rad project.

b- Distribution of the line-sampling parameters: The line-sampling parameters associated to a given database are independent of GCM-outputs and they are precomputed once for each database-version. But this pre-processing is likely to be computationally very demanding. At least, this is the case for Galtier et al's algorithm. This opens the question of associating these precomputed data directly to the GEISA and HITRAN projects: radiative-transfer users could recover the line-sampling parameters together with any GEISA/HITRAN update, without the need of any preprocessing. The practical meaning of such a data-distribution protocol will be evaluated within the present project as far as GEISA is concerned. This will also be discussed externally with the HITRAN management team.

Deliverables (M0 = 1/11/2018)

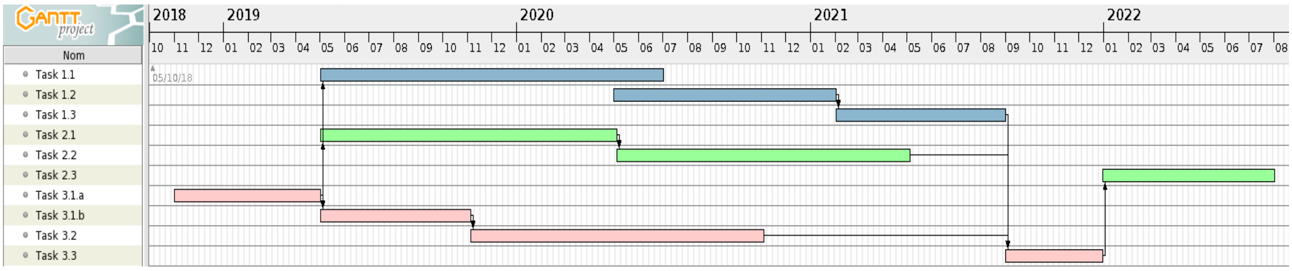
• Code and documentation, versions 1A (D3.1, M6), 1B (D3.2, M22), 2A (D3.3, M24), 2B (D3.4, 4) and 3A (D3.5, M34).
• Report "Performance evaluation : global radiative-forcing computation without accessing GCM outputs" (D3.6, M12)
• Documentation : "Switching from one database-version to another" (D3.7, M24)
• Documentation update : "Switching from one database-version to another" (D3.8, M24)
• Report: "Performance evaluation : radiative-forcing computation including access to GCM outputs" (D3.9, M24)

Success criteria:

• respecting the deadlines for the first-phase development
• adequation of the proposed user-interface with the climate-research practice
• adequation of the computer-ressource requirements with todays climate-center equipments (parallel computing, memory access, out-of-core)

Risk assessment: There is a risk that the first-phase software-design of Task 3.1 and Task 3.2 becomes incompatible with the need to include the modified algorithms of Task 1.3 : Task 3.3 would then require heavy recoding. An essential objective of the project-management team will therefore be to anticipate Task 3.3 : making sure that the software development team is fully aware of each of the partial conclusions of Task1 an that the API's are maintained open enough (even at the price of temporary efficiency-losses) to accept any of the finally retained algorithmic upgrades.

Gantt chart (M0 = 01/11/2018)



II-b-2 Funding requested:

Partner 1: LMD/IPSL

item of expenditure	description	cost
Staff	Postdoc with 2-7 year experience; Task 2.1 & 2.2 ; M6 to M30 (24 month)	132 768 €
	Gratification for 2 x Master 2 internship (2x4 month)	4 880 €
Instruments and material costs	1 workstation for CDD scientist + 1 laptop for permanent staff + small equipment	3 000 €
Travel and other costs	Project meetings and workshop (RFMIP)	7 000 €
	Attendance to international conferences	2 000 €

Partner 2: LAPLACE

item of expenditure	description	cost
Staff	PhD Student; Task 1.1 & 1.2 ; M10 to M46 (36 month)	108 816 €
Instruments and material costs	2 small workstations	1 600 €
Outsourcing / subcontracting	External service for the realisation of the software developments of Task3, using the open-source star-engine library under the EDSTAR platform	80 000 €
Travel and other costs	Project meetings	5 000 €
	Attendance to international conferences	4 800 €
	workshop organization	7 700 €

Partner 3: IRIT

item of expenditure	description	cost
Staff	Postdoc with 2-7 year experience; Task 1.1 & 2.2 ; M12 to M30 (18 month)	91 134 €
Instruments and material costs	1 powerful workstations	4 000 €
Travel and other costs	Project meetings	5 000 €
	Attendance to international conferences and workshops (SIGGRAPH, Eurographics, EGSR)	10 000 €

Requested means by item of expenditure and by partner*

		Partner *LMD*	Partner *LAPLACE*	Partner *IRIT*
Staff expenses		137 648.00 €	108 816.20 €	91 134.00 €
Instruments and material costs (including the scientific consumables)		3 000.00 €	1 600,00 €	4 000.00 €
Building and ground costs				
Outsourcing / subcontracting			80 000,00 €	
General and administrative costs & other operating expenses	Travel costs	9 000.00 €	17 500.00 €	15 000.00 €
	Administrative management & structure costs**	11 971.84 €	16 633.30 €	8 810.72 €
Sub-total		161 619.84 €	224 549.50 €	118 944.72 €
Requested		**505 114.06 €**		

II. Impact and benefits of the project

II.a Benefit for the scientific community:

Beyond this project, the partners have active collaborations with many national and international groups, ensuring a good connection between this project and current research activities and several of them have already expressed their interest in the present project (see the letters of intention at http://www.lmd.jussieu.fr/~jldufres/MCG-Rad/Letters/ login:MCG-Rad passwd:ANR2018). This project may have impact in the following research topics:

Climate change: Current estimates of radiative forcing mainly rely on studies that were made 10 to 20 years ago, they have not benefited from the progress made since. For instance in the last IPCC report (e.g. Myhre et al. 2013), the well-mixed greenhouse gas radiative forcing are taken from Myrhre et al., [1998]. Many progresses have been made since, both for the spectral data and for the characteristics of the atmosphere and our model will include them. The dissemination of our results within this community will be facilitate by our involvement in the RFMIP project and by our contribution to the next IPCC assessment report.

Earth atmosphere radiation: We are strongly involved in the Dephy national project which gather the French developers of atmospheric climate and weather forecast models at global and regional scale, and which have a dedicated task on "interactions between clouds and radiation". We also have strong connection with the remote sensing community.

Planets et exoplanets: Radiation is the very starting point of any atmosphere study of planets and exoplanets. We have long-standing collaborations with research groups working on these topics at LMD and in other laboratories.

Combustions: The proposition of Galtier et al. was initially motivated by heat transfer in combustion devices. LAPLACE is deeply involved in such ongoing researches and will promote the MCG-Rad's softwares in a broad engineering context.

II.b Socio-economic benefits:

The Méso-Star company will be economically interested as the library it develops, in collaboration with LAPLACE, will be at the heart of the present project: Méso-Star will have the opportunity to communicate on the new simulation potential of the star-engine library in terms of line-sampling algorithms, way outside the climate-change community. This ensures in particular that the present project benefits to all the industrial partners of Méso-Star, mainly in the field of Energy Engineering, where the accurate handling of gaseous thermal-radiation is a major issue.

II.c Education and training:

In direct relation with the EDStar plateform, the Occitanie Region already finances a teaching program (CLE-2017 project) addressed to researchers and engineers wishing to learn today practice of Monte Carlo applied to the modeling of complex systems. The present project will directly contribute to this teaching program.

III. References related to the project

Assaraf, R, and Caffarel, M. Zero-Variance Principle for Monte Carlo Algorithms, Phys Rev Lett 83(23):4682–5 (1999)

Chatterjee, K., Roadcap, J. R. and Singh, S. A new Green's function Monte Carlo algorithm for the solution of the two-dimensional nonlinear Poisson-Boltzmann equation. J. Comput. Phys. 276, 479-485 (2014)

Clough, S.A.et al. (1989) Line shape and the water vapor continuum. Atmospheric Research 23, 229-241

Curtiss, J. H. Monte Carlo Methods for the Iteration of Linear Operators. J. Math. Phys. 32, 209-232 (1953).

Dauchet, J. (2012), Analyse radiative des photobioréacteurs, PhD thesis, Univ. Clermont Ferrand

Dauchet, J., Blanco, S., Cornet, J.-F., Hafi, M. E., Eymet, V., and Fournier, R. (2013) The practice of recent radiative transfer Monte Carlo advances and its contribution to the field of microorganisms cultivation in photobioreactors. J. Quant. Spectrosc. Radiat. Transfer 128, 52-59

Dauchet, J., Bezian, J.-J., Blanco, S., Caliot, C., Charon, J., Coustet, C., El-Hafi, M., Eymet, V., Farges, O., Forest, V., Fournier, R., Galtier, M., Gautrais, J., Khuong, A., Pelissier, L., Piaud, B., Roger, M., Terrée, G., and Weitz, S (2016).

Dufresne J.-L., R. Fournier, C. Hourdin, and F. Hourdin (2005) Net exchange reformulation of radiative transfer in the CO2 15-um band on Mars. J. Atmos. Sci., 62 :3303-3319 Monte carlo and nonlinearities. https://arxiv.org/abs/1610.02684

El Hafi, M., Blanco, S., Dauchet, J., Galtier, M., Fournier, R., Tregan, J.M. and Villefranque, N (2018) Three viewpoints on null collision Monte Carlo algorithms, In Eurotherm Seminar 110 - VI April 11-13, Cascais, Portugal

Eymet, V., Fournier, R., Dufresne, J.-L. et al. (2009) Net-Exchange parameterization of infrared radiative transfers in Venus' atmosphere, JGR Planets, 114, E11008, doi:10.1029/2008JE003276

Eymet, V. et al. 2013: Null-collision meshless Monte-Carlo. Application to the validation of fast radiative transfer solvers embedded in combustion simulators. JQSRT 129, 145-157, https://doi.org/10.1016/j.jqsrt.2013.06.004

Galtier, M., Blanco, S., Caliot, C., Coustet, C., Dauchet, J., El Hafi, M., Eymet, V., Fournier, et al.(2013) Integral formulation of null-collision Monte Carlo algorithms. J. Quant. Spectrosc. Radiat. Transfer, 125, 57–68

Galtier, M. et al. 2016: Radiative transfer and spectroscopic databases: a line-sampling Monte-Carlo approach. J. Quant. Spectrosc. Radiat. Transfer, 83-97, https://doi.org/10.1016/j.jqsrt.2015.10.016

Jakob W. and S. Marschner (2012) Manifold exploration: A markov chain monte carlo technique for rendering scenes with difficult specular transport. ACM Trans. Graph., 31(4):58:1–58:13.

Kalos, M.H. & Whitlock, P.A. Monte Carlo Methods, second ed. (Wiley-VCH, Weinheim 2008).

Kutz, P. et al. 2017: Spectral and decomposition tracking for rendering heterogeneous volumes. ACM Trans. Graph. 36, 111:1-111:16, http://doi.acm.org/10.1145/3072959.3073665

Lataillade, A., S. Blanco, Y. Clergent, J.L. Dufresne, M. El Hafi and R. Fournier (2002) Monte Carlo methods and sensitivity estimations. J. Quant. Spectrosc. Radiat. Transfer, Vol. 75, N.5, pp.529-538

Meinshausen M. et al. (2017) Historical greenhouse gas concentrations for climate modelling (CMIP6), Geosci. Model Dev., 10, 2057–2116, doi:10.5194/gmd-10-2057-2017

Mellado N., D. Aiger, and N. J. Mitra (2014). Super 4pcs fast global pointcloud registration via smart indexing. Comput. Graph. Forum, 33(5):205–215.

Mellado N., M. Dellepiane, and R. Scopigno (2016) Relative scale estimation and 3d registration of multi-modal geometry using growing least squares. IEEE Transactions on Visualization and Computer Graphics, 22(9):2160–2173

Myhre G et al. (1998) New estimates of radiative forcing due to well mixed greenhouse gases. Geophys. Res. Lett., 25(14), p. 2715-2718

Myhre G. et al. (2013) Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Fifth Assessment Report of the Intergovernmental Panel on Climate Change

Novák et al. 2014: Residual ratio tracking for estimating attenuation in participating media. ACM Trans. Graph. 33, 179:1-179:11, https://doi.org/10.1145/2661229.2661292

Pajot A., L. Barthe, M. Paulin, and P. Poulin (2011). Representativity for robust and adaptive multiple importance sampling. IEEE Transactions on Visualization and Computer Graphics, 17(8):1108–1121.

Pajot A., L. Barthe, and M. Paulin (2014) Globally adaptive control variate for robust numerical integration. SIAM Journal on Scientific Computing, 36(4):A1708–A1730

Pincus, R., H. W. Barker, and J.-J. Morcrette (2003) A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields, J. Geophys. Res., 108(D13), 4376, doi:10.1029/2002JD003322

Roger M, S Blanco, M El Hafi, R Fournier (2005) Monte Carlo estimates of domain-deformation sensitivities. Phys. Rev. Lett. 95(18), pp. 180601, https://doi.org/10.1103/PhysRevLett.95.180601

Szczap F. et al. (2014) A flexible three-dimensional stratocumulus, cumulus and cirrus cloud generator (3DCLOUD), Geosci. Model Dev., 7, 1779-1801 https://doi.org/10.5194/gmd-7-1779-2014

Szirmay-Kalos L. et al. 2017: Unbiased Estimators to Render Procedurally Generated Inhomogeneous Participating Media. Computer Graphics Forum 36 (2) https://doi.org/10.1111/cgf.13102

Veach E. and L. J. Guibas (1995). Optimally combining sampling techniques for Monte Carlo rendering. In Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95, pages 419–428