Reports on projects with SMEs.

This appendix includes short progress reports on progress with SMEs, for those projects having started more than 6 months ago.

- Project 1: Climate projection post-processing with TCDF
- Project 2: Analogue seasonal forecast with ARIA-Technology
- Project 3: Transfer of LMD-Z with ARA-Technology

Generation of bias corrected CMIP5 and CORDEX essential climate variables data set

Preformed by The Climate Data Factory

1. Goals

The goal of this 18 month project is to develop a post-processing chain software to handle statistical post-processing operations on climate model data sets and generate a first data set of bias corrected climate model simulations.

The post-processing chain is based on the SYNDA software that is an advanced ESGF downloade manager developed by the IPSL PRODIGUER team. The SYNDA SDT (Synda Data Transfer) component handles the synchronisation and download of large data files between the *Earth System Grid Federation* climate model data repository on local resources. The Synda SDP (Synda Data Processing) component handles a workflow engine - it orchestrates complex distributed interdependent tasks triggered upon download completion. Principal data sets on ESGF are the CMIP5 and CORDEX projects that are a set of global (CMIP) and regional (CORDEX) reference climate simulations under several atmospheric CO2 future conditions (The so called *Representative Concentration Pathways).* During the project several computational modules are developed to perform the calculations as well as apply quality control procedures both on technical specifications (ESGF file standards) as well as a data check (outliers detection).

The statistical post-processing that is developed is "bias correction". Indeed, because climate models are not perfect, simulated climatologies differ from observed ones making them unfit for direct use in Climate Change impact studies. A significant part of those differences are due to systematic errors that can be corrected by statistical techniques called bias correction. Such techniques include regression, analogues or quantile-quantile matching. The method implemented during this project is a quantile-quantile transformation developed at LSCE (Vrac et al. 2016).

These developments will allow to generate a bias corrected data set of climate simulations for a number o essential climate variables based on the CMIP5 an Euro-Cordex climate model simulations.

2. Expected deliverables

The deliverables are a set of software modules and a data set. Software module will be orchestrated by the SDT component.

There are three modules: a statistical module (to apply the bias correction method), a technical quality control module (for checking ESGF standards of data sets) and physical quality control module (to check for outliers in the data).

The bias corrected data set is based on 6 daily variables from 2 experiments. The variables are surface mean temperature, surface minimum temperature, Surface maximum temperature, precipitations, surface wind and surface radiation. The experiments are the CMIP5 and Euro CORDEX experiments, including all available models and RCP's published on ESGF.

3. Progresses so far

The project started in may 2016 with the development of the software modules and their inclusion in the SYNDA software. SDP module has been extended to fulfil TCDF requirements, especially bias correction related requirements. As of December 2016, development and technical testing of the chain is almost over and the development phase will be finalized by January 2017. What remains to be done is the so-called "production" phase that is the generation of the data set.

4. Envisaged activities, potential markets for the company

The climate data factory (TCDF) is a climate service provider of post-processed climate change model data to the climate change impact and adaptation communities of both scientist and practitioners. Indeed climate change impact studies require data that are becoming difficult to access and need correction from systematic errors in order to be used in impact models. The data management segment of the climate services market was evaluated at 850M euro in 2015.

Un Générateur de Temps pour l'Innovation (GTI)

Projet Valorisation LABEX IPSL

Report of GTI project

Pascal Yiou (LSCE) and Armand Albergel (ARIA Technologies)

Goals

The statistical method of atmospheric circulation analogues was devised in the 1960s as a means for weather prediction. Its performances as such were soon superseded by numerical weather prediction. It has mostly been used as a *statistical downscaling* tool (e.g. Vautard and Yiou, 2009). Recently, this method found innovative applications: the *reconstruction* of atmospheric fields that are compatible with surface atmospheric structures (Yiou et al., 2013, 2014), and the *detection* of extreme events (Yiou and Cattiaux, 2012, 2013, 2014).

LSCE developed a *stochastic weather generator* (AnaWEGE) that simulates time series of climate variables (temperature and precipitation) that are related to the atmospheric circulation, with a daily time step (Yiou, 2014). The ingredients of this weather generator are random selections of analogues of the atmospheric circulation over the North Atlantic region. This weather generator was tested for European temperatures, with a couple of test cases (summer 2003, winter 2009/2010).

In its present version (free software, with CECILL license, protected by a depository at the French APP), AnaWEGE simulates large ensembles of sequences (for instance, seasons) that are initialised by observed conditions between 1st January 1948 and 31st December 2012. It allows to determine the probability distribution of daily climate variables (temperature, precipitation) after a known initial state. It appears that this tool can be modified to produce simulations in “prediction” mode with an appropriate choice of circulation analogues. Hence, we plan to test the performance of this tool to simulate the spread of climate trajectories at a time scale between days and seasons.

This tool is not intended to replace predictions from meteorological centers, but to generate large ensembles (at least several hundreds of members, on a desktop computer in a few minutes) of climate variables, and compute their probability distributions. The scientific added value is a focus on small regions, which is not reachable by numerical weather prediction products. Its ultimate use by ARIA Technologies is to increase the client awareness on uncertainty of weather prediction, and hence constitute a useful pedagogical tool. ARIA Technologies also showed an interest of the AnaWEGE tool (as is) for its own needs.

The scientific tasks of the project will be to adapt the AnaWEGE programme to prediction mode (based on regularly updated reanalyses). We will test the stochastic prediction mode against observations (and available operational analyses) in order to:

- compare the skill with “persistence”, “climatology” and “observations”,
- determine the upper bound of a prediction window (between 2 days and 4 weeks),
- evaluate the synoptic conditions that lead to poor/acceptable prediction.

The developments and tests should provide the material for a scientific paper. ARIA Technologies made several market studies on the impact of climate change. It was found that most potential clients are interested in subseasonal climate prediction and uncertainty. If the tests convey useful results for ARIA, a formal agreement of further software commercial use will be built between CEA and ARIA at the end of the project. In the present project, ARIA will get acquainted with the modified code of AnaWEGE and will conduct the tests, with the assistance of P. Yiou.

Expected deliverables

D1: Automated procedure of ensemble stochastic prediction of temperature, using NCEP reanalysis updates. The computing system has been installed on the ARIA computing cluster and works in a semi-operational way.

D2: Analysis of skill of ensemble temperature predictions for selected stations. We selected Paris and Toulouse. Tests are made for the 2000-2015 period.

The first two deliverables were achieved within the training (stage M2) of Mariette Lamige (May-September 2015).

D3: Extension of analysis to temperature in all stations in Europe (from ECA&D): To be done

D4: Analysis of skill for precipitation: to be done

Progress

A Masters student (Mariette Lamige; Université de Lyon 1) was hired for 5 months for an internship at ARIA and LSCE (May 2015 to September 2015).

We first started to modify the code of the stochastic weather generator AnaWEGE code into a predictive mode. This is done simply by selecting random analogues outside of the period to be simulated. The code was adapted to make ensembles of predictions from updated NCEP reanalysis SLP data. The analogues are taken from the continuously updated computations at:
https://a2c2.ipsl.fr/

The code is written in R and linux shell scripts. It can run on processors with several cores to increase the speed of the calculations.

The second goal was to test the stochastic predictions against observations. We selected temperature datasets (mean daily temperature) in Paris and Toulouse from ECA&D. The predictions are done in hindcast for winters between 2000 and 2010 (for which ECA&D was available at the time). For each day, we also used the North Atlantic weather regime decomposition, also available at: https://a2c2.ipsl.fr. Weather regimes are used to determine the weather types associated with high/low predictability.

We tested how the stochastic prediction behaves with respect to persistence (no variation after the initial condition) or climatology (no variation around the climatological mean of temperature around the initial calendar day). An example of prediction in December 2009 is shown in Figure 1. This example is chosen because it is the beginning of the major cold winter that struck Europe in 2009/2010.

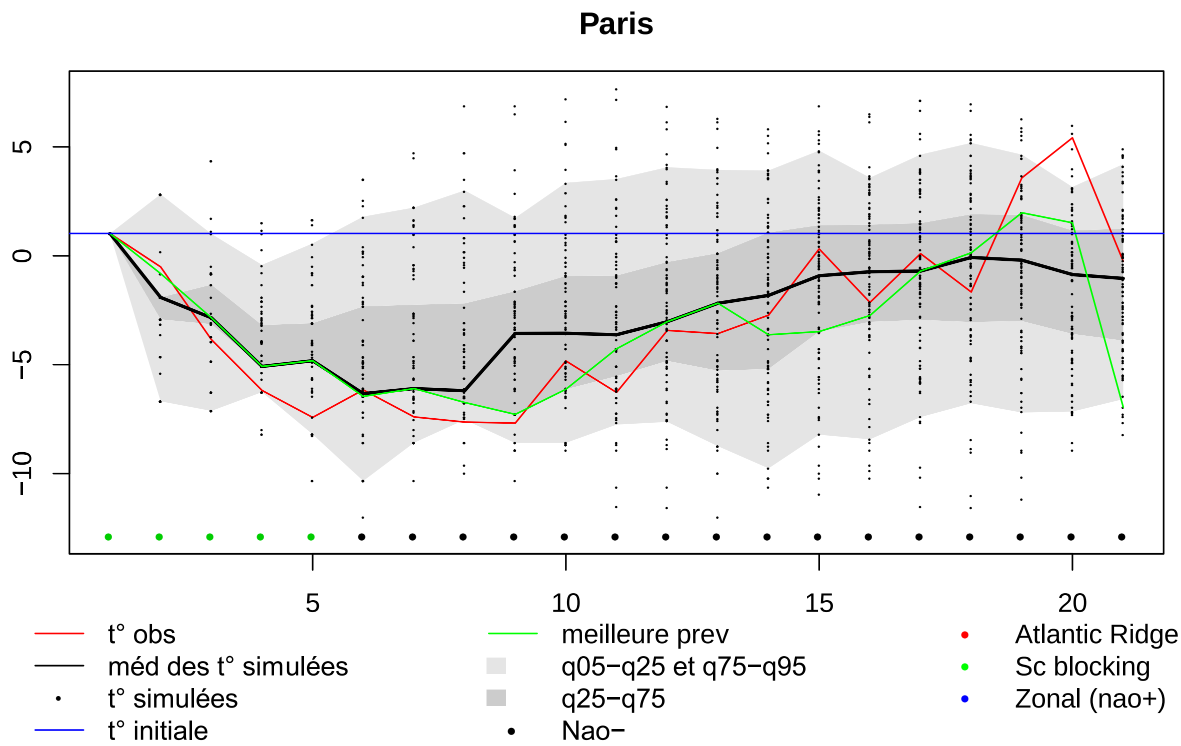

Figure 1: Comparison of stochastic predictions and observations for Paris temperature, starting on 11 December 2009. The red line is observations (from ECA&D). The dots are 100 realizations of the stochastic prediction and the black line is their median. The green line represent the best of the predictions, according to an RMSE score. The colored dots at the

bottom of the figure represent the weather regime. The horizontal blue line is the initial condition, reflecting persistence.

In this example, the prediction follows quite closely the observation for more than two weeks. After that, it converges to a climatological value. This is an extreme case where the prediction works very well for a relatively long time.

We propose a test of trend of prediction by evaluating the number of days for which difference between the prediction and the initial condition has the same sign as the difference between the observation and the initial condition. This test is evaluated for every day.

Overall, we find that the predictions are quite satisfactory for Paris temperature, with an average of 8 days in the winter (Figure 2). The skill appears to be lower in Toulouse.

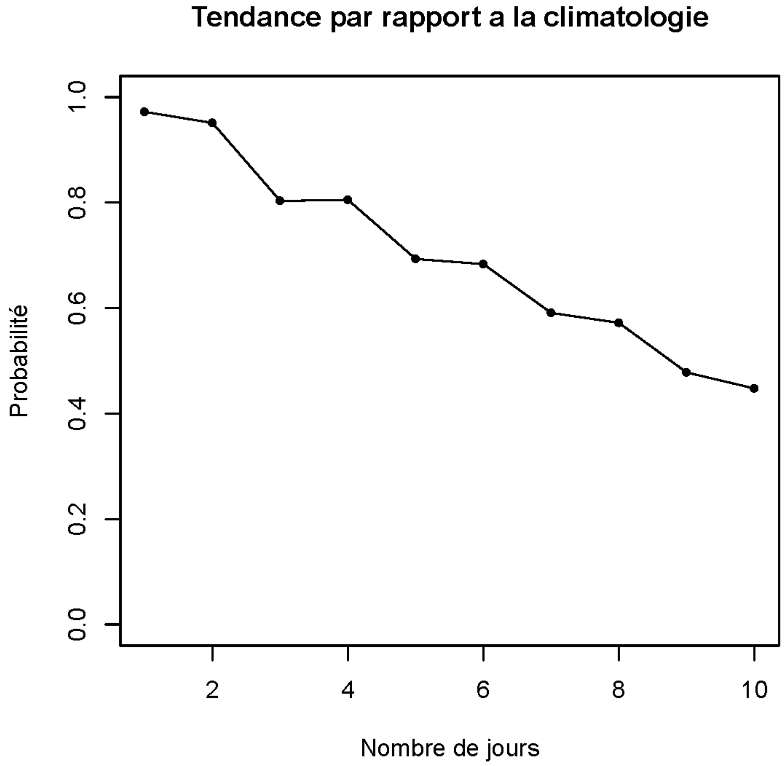

Figure 2: Probability that the prediction of winter temperature in Paris has the same trend as observation with respect to climatology. The probability exceeds 0.6 for ~6 days in advance.

The full report of Mariette Lamige (2015) is available at:
https://files.lsce.ipsl.fr/public.php?service=files&t=2a179d2d2acabba3183102605109d5c1

Those preliminary results on French temperatures are promising for the design of a low cost prediction tool. This tool now runs automatically on the computer cluster at ARIA Technologies. More tests (other temperature stations, other skill scores) need to be performed to ensure the relevance of the tool. We have advertised internship opportunities in universities.

Envisaged activities, potential markets for the companies

The analogue methods as used in AnaWEGE are promising in term of direct application in weather forecasts and especially for extreme events. ARIA Technologies and LSCE already have collaborated together in two promising Climate-KIC projects:

- E3P: Extreme Events for Energy Providers. LSCE did a downscaling of IPCC global model outputs and ARIA implemented a dedicated Website to disseminate indices (http://web.aria.fr/creator/E3P/)
- OASIS: a platform for insurance and reinsurance. This activity requires big sets of realistic data on extreme events which are obviously very few.

ARIA Technologies is presently working on an extension of OASIS called OASIS+ and it holds the lead of Wat-Ener-Cast. (WEC) where several products dedicated to both water and energy management are developed. Many forecast horizons are investigated with final customers from short range (nowcasting) to long range (climate change). This project opens new potential application of AnaWEGE as it is and suggests new perspectives for analogue techniques.

Two summarize the potential applications segments of the current works between ARIA Technologies and LSCE:

(1) Solution to better take into account the-meteorological parameters for relevant industries
   a. Direct application of Climate-KIC projects as E3P, OASIS or WEC).
   b. Perspectives : We are seriously want to investigate how AnaWEGE methodology shall improve the nowcasting with new rain monitoring methods using mobile phone signals exchanged between microwave stations (Climate KIC Pathfinder T-RAIN).

(2) Climate Change adaptation
   A key question is how to properly and efficiently downscale IPCC data set (as CMIP5) to provide most realistic figures of the future local climate. ARIA is answering to several calls and Tenders. As soon as AnaWEGE methodology is validated, it can be used is such works. An example can be the current project funded by UNDP to study rain and water stress in Middle East (Israel, Jordan and Palestine).

(3) Perspectives in meteorological situation classification
   For many applications (Environmental Impact Report, Safety analysis Report…) it is not possible to cover many years of meteorological data (past and/or future). A solution is to select “typical” concerned situations. We think that this segment of work can also open new perspectives to the analogue methods.

MESOSCALE & COMPREHENSIVE DOWNSCALING with LMDz Model (MiCaDo)

Projet Valorisation LABEX IPSL

Report of MiCaDo project

Frédéric Hourdin (LMD) , Bruno Guillaume, Céline Déandreis, Armand Albergel (ARIA Technologies)

Goals

More and more applications connected with weather or Climate issues require high resolution 3D gridded data. Dynamical solutions take benefit from large scale data and a better description of the regional domain (topography, land-use,…) .The most common solution is the use of the US-WRF model nested in global output as GFS or ECMWF analyses or forecast. The use of LMDz software developed by IPSL/LMD could provide a new and innovative solution including several competitive differences.

The "z" of "LMDz" refers to this zoom capability. LMDz follows a different approach comparing to WRF, the meteorological variables outside the zoomed region are guided (called nudging in the literature) by applying a relaxation term to the variables of the larger domain. The LMDZ model can therefore be used as a limited area model. This model has advantages compared to the limited-area models because it allows not to completely cut back the feedbacks with the exterior of the area of interest (high resolution grid), which is represented in the global model (coarse grid).

The use of this model also allows to propose a 100% French solution and to directly benefit from the modelling expertise of the team that develops LMDz.

In the context of the present work, we plan to carry out and evaluate several downscaling runs and to assess their possible interest in the fields of air quality and water management. The main test here is in replaying with LMDz some of the SECIF (IPSL ANR project) cases and in comparing the results obtained with those obtained previously with the WRF model.

Expected deliverables

D1: Implementation of the LMDz model on ARIA’s cluster and preliminary tests on an interest area taken from a SECIF case

- *In progress*

D2: Replay the case with a dual global and zoomed model in present and future climate.

- *To be done*

D3: Analysis of LMDz software use as a downscaling tool (SWOT: Strengths, Weaknesses, Opportunities, Threats)

- *To be done*

Progress

LMDz model code was transferred and implemented on ARIA computers. A first training session was carried out.

The selected SECIF case is a domain around Roman –sur –Isere (Alps Mountains) and tests on heavy rains have been done. Numerical problem on our cluster are fixed but some modelling issues remain to be solved with IPSL team assistance.

The project has been delayed for two main reasons:

(1) Due to administrative issues (see Francoise Beaud 's email sept. 2016), the project was set to stand-by position these last months. A new agreement is in discussion and will propose a new planning.

(2) Celine Deandreis who is in charge of the modelling work in ARIA is currently on maternity leave (her return is expected in March 2017)

Envisaged activities, potential markets for the companies

ARIA Technologies is dedicated to the atmospheric environment with 25 years of experience in developing numerical models for forecasting the dispersion of atmospheric effluents. ARIA wants to build a bridge between research and application, initially by applying micro-scale codes (domains of 20 km, meshes of the order of 20 m) for Environmental Impact Report and Safety analysis Report. For the past ten years, ARIA has also developed air quality and mesoscale prediction systems for several cities and regions around the world, most of which are based on the CHIMERE community model developed at IPSL: Israel, Rio, Qatar, Kaliningrad, India with zoom on Delhi, Beijing, Tunisia with zoom on Tunis and other cities, Romania with zoom on Bucharest, Indonesia at 15km resolution, and recently Concepción in Chile). MM5 or WRF are generally used and it could be an interest to propose LMDz in the future releases.

In the past five years, ARIA has begun to accumulate dynamic modeling experience in several climate studies: KIC-Climat E3P, OASIS, CARBOCOUNT City, the ANR SECIF project, co-financing of Benjamin Quesada's thesis (Directed by Robert Vautard) with IPSL (study of cold waves), the co-financing of the thesis of Luiz Molina (directed by Philippe Ciais) on the estimation of the CO_2 emissions of the Amazon with reverse CHIMERE.

The Climate Change opens also new doors for LMDz as a downscaling tool using IPCC output.