

# MODELING PROCESS CHAIN OF METEOROLOGICAL REANALYSIS PRECIPITATION DATA USING WORK CONTEXT

Aimé Richard Hajalalaina<sup>1,2</sup>, Angelo Raherinirina<sup>2,3</sup>, Adolphe Ratiarison<sup>4</sup> and  
Thérèse Libourel<sup>5</sup>

<sup>1</sup>Ecole de Management et d'Innovation Technologique  
Université de Fianarantsoa, Fianarantsoa, Madagascar

<sup>2</sup>Laboratoire d'Informatique et de Mathématiques Appliqués aux Deveppements,  
Université de Fianarantsoa, Fianarantsoa, Madagascar

<sup>3</sup>Ecole Normale Supérieure, Université de Fianarantsoa, Fianarantsoa, Madagascar

<sup>4</sup>Laboratoire de la Dynamique de l'Atmosphère, du Climat et de l'Océan, Université  
d'Antananarivo, Antananarivo, Madagascar

<sup>5</sup>LIRMM, Université de Montpellier II, Montpellier, France

## **ABSTRACT**

*In this paper, we propose a models of process chain and knowledge-based of meteorological reanalysis datasets that help scientists, working in the field of climate and in particular of the rainfall evolution, to solve uncertainty of spatial resources (data, process) to monitor the rainfall evolution. Indeed, rainfall evolution mobilizes all research, various methods of meteorological reanalysis datasets processing are proposed. Meteorological reanalysis datasets available, at present, are voluminous and heterogeneous in terms of source, spatial and temporal resolutions. The use of these meteorological reanalysis datasets may solve uncertainty of data. In addition, phenomena such as rainfall evolution require the analysis of time series of meteorological reanalysis datasets and the development of automated and reusable processing chains for monitoring rainfall evolution. We propose to formalize these processing chains from modeling an abstract and concrete models based on existing standards in terms of interoperability. These processing chains modelled will be capitalized, and diffusible in operational environments. Our modeling approach uses Work-Context concepts. These concepts need organization of human resources, data, and process in order to establish a knowledge-based connecting the two latter. This knowledge based will be used to solve uncertainty of meteorological reanalysis datasets resources for monitoring rainfall evolution.*

## **KEYWORDS**

*Modeling; Process chain; Climatology; Reanalysis; Work-context*

## **1. INTRODUCTION**

Our concerns are on the application of computer science in the field of rainfall evolution. Specialists in climatology produce hypotheses which they validate from experimental protocols. Now, climatology and computer experts work together to automatize these protocols in processing chains. This automatization poses many problems due, on the one hand, to the

volumes of meteorological reanalysis datasets coming from different sources, ECMWF [1], NCEP [2], MERRA [3] and, in the other hand, to the proliferation of more or less complex meteorological data processing methods required by environment climatology.

In this context, it is necessary to put in place systems which help to store and manage important meteorological data-streams as well as their processing, by taking into account their various origins (different formats and sources, different spatial and temporal resolutions). The restitution and the exchange of these pieces of information are a real challenge in terms of interoperability. This later has an advantage in reducing the uncertainty of meteorological reanalysis data-processing methods through the capitalization and the mutualization of experiments between climatology specialists.

A lot of research-work, which represent high accuracy of result, has been published on the study of rainfall evolution using meteorological reanalysis datasets [4][5], which has contributes to improving the knowledge on rainfall evolution. But a data and processing formalization is necessary to ensure much larger spread of the research results to the scientific community. Concerning our work in formalization of processing chain, we have published in [6] and [7].

The formalization of data and processing is realized in a Work-Context acquired by the MDWeb platform put into practice by Desconnets [8]. This platform proposes a view of human resource organization. It allows the platform administrator to reference the future users from predefined roles and rights, and to reference the ressources (data/processings) within a metadatabase.

We propose to use this of this formalization to coordinate the collection of meteorological reanalysis data and their processing, necessary for monitoring rainfall evolution. The objective is to capitalize, harmonize and spread the resources which allow a better understanding of rainfall evolution at different spatial and temporal resolutions. It is part if several achievable actions within different period of time. The first one is the formalization of meteorological reanalysis data processing to ensure their sharing, their re-use and their interoperability. The second one, which is a medium-term objective, is setting-up of a platform for the sharing a mutualization of experiments on meteorological reanalysis data processing methods on the study of rainfall evolution.

First, we have collected the current norms and formalization of spatial data and processing, and then proposed Work-Context models for the formalization of meteorological reanalysis data processing chains for monitoring rainfall evolution.

## **2. RESOURCE FORMALIZATION**

The formalization corresponds with the description of the resources used by respecting the norms and standards which are in force. They are in the size of metadata. Various formalizations respecting the norms in force deal with syntactic and semantic aspects for the metadata. This definition and the application of these norms help to resolve the heterogeneity of the resource used. Dublin Core in [9] BNF (2008) and ISO 19115 [10] (2003) are the two norms which are valid and most used for the formalization of georeferenced data, like meteorological reanalysis data, while ISO 19119 [11] (2005) allows us to describe their processing. The formalization of processing chains use notion of Work-Context. The Work-Context is closely linked to the field of expertise. It can then be envisaged that it is built, through consensus, by experts in the area. Every

scientist (the final user) will have at their disposal this context which they can develop or enrich in their turn. The construction of any Work-Context corresponds with the progressive on gradual organization of useful data and processing references, following three steps: human resource organization, data organization, and processing organization according to Libourel [12]. Moreover even if this approach has not been standardized yet, it uses a relatively simple graphic symbolic language, called Simple Workflow Model (SWM) suggested by Lin [13], which allows the user scientists to handle easy appropriation concepts in a simple language.

The Work-Context of meteorological reanalysis data in field of monitoring rainfall evolution is made up of three models bellow: human resource organization (Figure 1) which manages the description of the platform users as well as that of their different roles and access rights, data organization (Figure 2) which manages the description of meteorological reanalysis data category according to their sources, processing organization (Figure 3), which manages the description of meteorological reanalysis data processing categories, Work-Context knowledge-base (Figure 4) and the abstract model of processing chain of meteorological reanalysis data to monitor rainfall evolution (Figure 5). This abstract model can be implemented by designing concrete model by using R or MATLAB statistics and mapping toolbox.

From the organizational formalism of human resources using the notion of Work-Context, we identify the essential roles of users who intervene in the use of resources to study the rainfall evolution (Figure1).

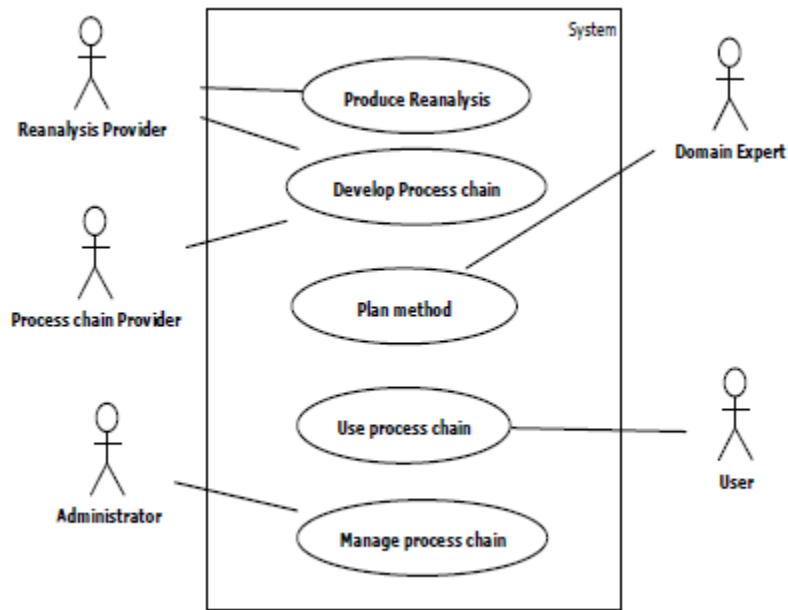


Figure1. Human resource organization model

Using the formalism of the hierarchization of data categories of Work-Context notion, we propose the following category model of meteorological reanalysis data to monitor rainfall evolution (Fig. 2).

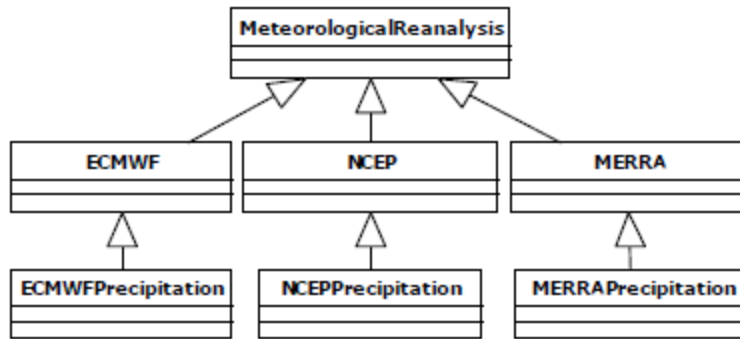


Figure 2. Meteorological reanalysis data organization model

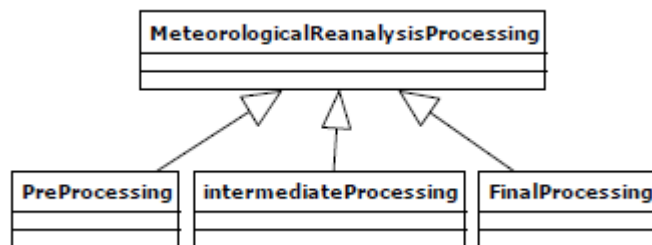


Figure 3. : Meteorological reanalysis data processing organization model

Work-Context model is built (Figure 4).

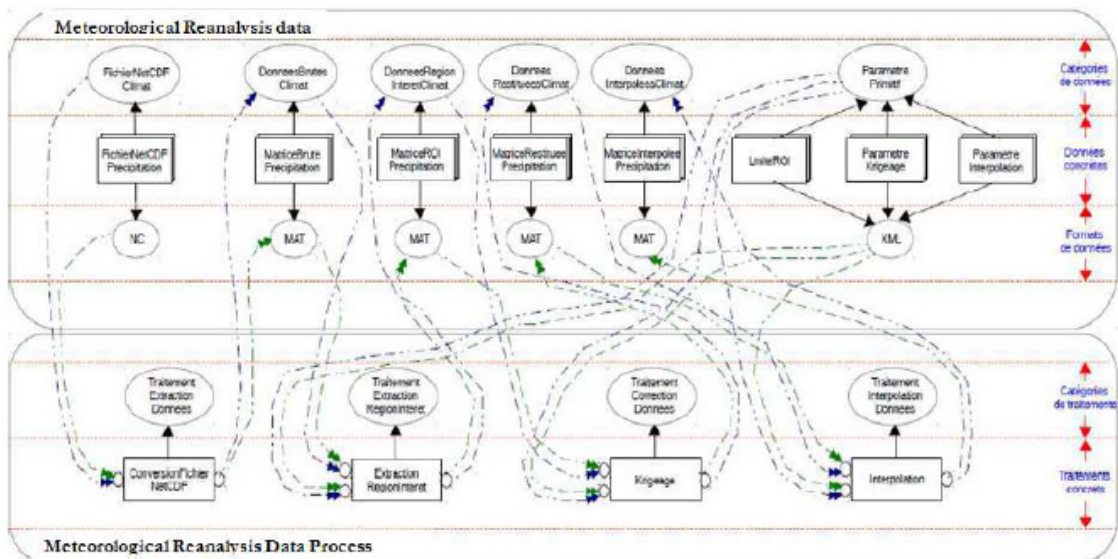


Figure 4. Extract of Work-Context knowledge-base of meteorological reanalysis data model

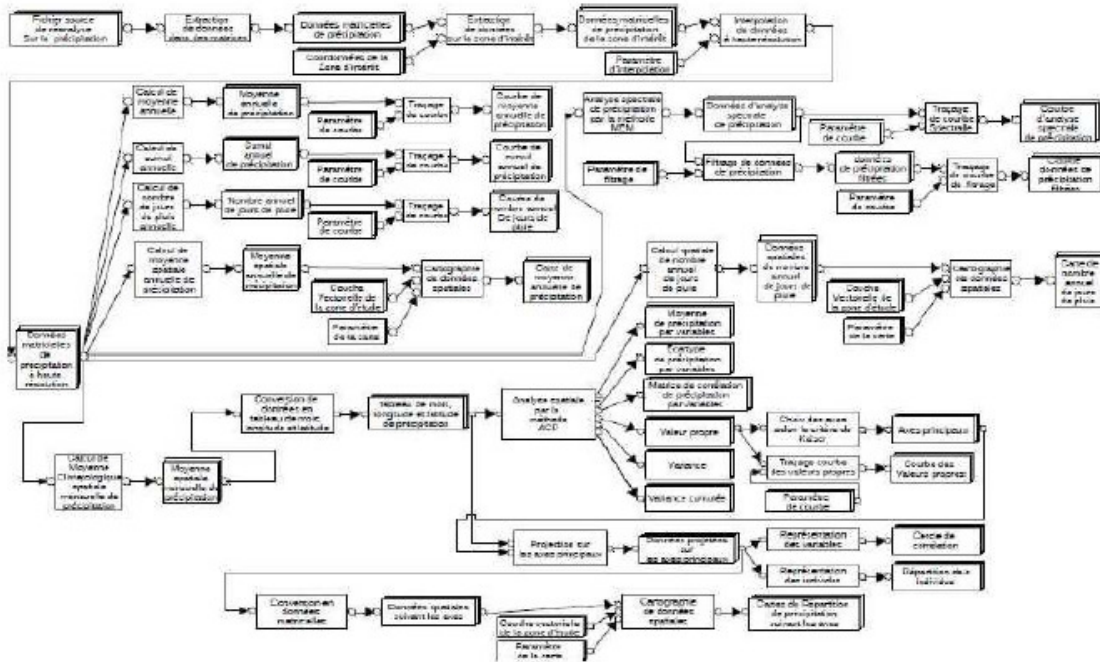


Figure 5. Abstract model of multi-resolution and multi-temporal ECMWF meteorological reanalysis data processing chain to monitoring rainfall evolution

### 3. ASSESSMENT

The manual processing chains of meteorological reanalysis data will always remain a tough intellectual process. As a matter of fact, this type of chain requires the help of a climatology expert all along the realization process. In addition, new meteorological reanalysis data entering the chain require the re-start of the process from the beginning till the end. This makes their reuse difficult for users who are not specialists in the field and it does not make it possible to capitalize the experts' knowledge to assist the occasional users in the valorization of new reanalysis data. We can say that the manual processing chains are not adapted to the meteorological reanalysis data actualizing critical phenomena like the monitoring of rainfall evolution.

Faced with this advantages of the manual processing chains, it interest of our proposals on formalizing processing chains lies in the automatization of different processing on meteorological reanalysis datasets.

Processing chain re-use: Once processing chain is formalized, its execution (implementation) is carried out automatically by taking into account new reanalysis data. This is what makes the chain re-usable and also makes several repetitions of the processing possible without climatology expert assistance for using the new meteorological reanalysis data to carry out their analyses corresponding to their needs. Those properties are actually adapted to the monitoring of the rainfall evolution, in view of changing and evolving character of meteorological reanalysis datasets. It presents a major interest for capitalizing the experiments on meteorological reanalysis data processing methods in order to overcome the problems of uncertainty in the processing of rainfall evolution.

**Interoperability of processing chains:** The formalized chain can therefore, a priori, be implemented in a distributed environment of Grid or Cloud type. Thus, it would then now be interesting to display these chains in a distributed manner in order to profit from this possible interoperability. As a matter of fact, in the contexts of rainfall evolution monitoring, the distributed dimension will bring about an important advantage to this approach for it allow to divide up the experiments and spread on the whole set of the system to put in place, the contribution of separate elements to the knowledge-bases of Work-Context. In addition, normalized exchange protocols allow distributed requests which can be interesting for simultaneous research of meteorological reanalysis data processing methods used in monitoring rainfall evolution, which makes it possible to fills up the gaps on the uncertainty of meteorological reanalysis data processing methods.

**Cataloging processing chains:** Making catalogs is the first step that will lead to an exchange or a spread of resources. Often the catalog is the origin of the initiatives for the rapprochement between the actors who wish to share their meteorological reanalysis data processing chains for the monitoring of rainfall evolution. We would like to stress that it does not answer a move to share the processing chains but the one to share the metadata. We take profit from this possibility that the processing chains modeled and formalized in this way can be included into a catalog.

**Processing chain exchange:** We can take profit from the processing chain exchange by taking into account the different human resources which are proposed for monitoring change in rainfall evolution. As a matter of fact, in contract to the spread which can be made through a downloading website, the exchange of meteorological reanalysis data processing chains generally requires an interaction between the actors involved. This encounter is often initiated for legal reasons in order to draw up an exchange convention that respects the objectives of each party, or each of the parties. The notion of exchange is sometimes mixed with that of partnership, much broader. A partnership is a process of strategy cooperation between at least to actors and whose objective is to achieve a goal through the common use of material, intellectual, human, financial means. It often has a political dimension. The exchange of processing chains is interesting in setting up a process concerning the physical realization of one of the aspects of partnerships between the different actors. We stress that the notion of exchange implies a bilateral relation (the producer versus the receiver) which focuses on a flow while the notion of sharing also allows considering all the related aspects, in particular the appropriation and sharing of the knowledge which can result the exchange.

**Spread of processing chains:** More modest on the level of interactions, the objectives on the spread are not a challenge for the capacities of appropriation of processing chains by the actors with who the producer will have little or even no interaction. It can be pointed out here that there is a marked opposition with the co-production objective which, instead of relying on work groups, prefers to spread the produced processing chains without any discussion. It's the logic of the action of the researchers who wish to make their processing chains accessible to a great number of people.

**Mutualization of processing chains:** The interest in the formalization of chain-processing is to encourage the mutualization of experience experiments between the researchers and is justified by various reasons. The reason is first of all to avoid useless duplication of efforts in the constitution of chain processing, avoiding at the same time redundancy of the result fragmentation which comes out. Considering the compartmentalization related to the

multiplication of autonomous and non-coordinated chain-processing, the formalization of these chains represents an adequate solution to these defects.

Finally, our proposals on the formalization of processing chains constitute necessary conditions for the mutualization of experiments for monitoring rainfall evolution. In fact, the mutualization of data implies formalizing the production and spread supervision modalities, according to an interoperable mode of a piece of information ready for use and whose quality is checked, for reason of common usage. By reinforcing the cooperation between producer and users of processing chains, the mutualization also favours the exchange of experience and good practice through networking.

Processing-chain sharing: At last, it can be said that our proposals on the processing-chain formalization facilitate the sharing of experiments to monitoring change in rainfall evolution. As a matter of fact, the data sharing aims at providing the community of experiment-users with coherent formats and structures, which will enable them to fulfill their missions better and produce their experimentations according to common and pre-defined requirements. Conversely, this interoperability and the improvement of processing-chain availability can contribute to reinforcing the relationships between the producer and user organizations and therefore can help to fill up some gaps on the use of meteorological reanalysis datasets for the monitoring of rainfall evolution.

Experiment-sharing platform: Starting from the advantages offered by over proposals on the modeling and formalization of meteorological reanalysis data processing chains for the monitoring of rainfall evolution we are going to present other advantages which come out. We can state that the formalization of processing chains can be valued by the setting-up of a platform for sharing, mutualization re-use and the spread of experiments on the valuation of meteorological reanalysis datasets for the monitoring of rainfall evolution.

The knowledge on the meteorological reanalysis data processing methods for the monitoring of rainfall evolution are modeled then formalized in the form of processing-chains. The knowledge acquired on meteorological reanalysis data processing has to be made available to the actors. This capitalization is conceived so that everyone's experiment does not remain continued to individual level, but serves the collective in a move for knowledge sharing allowing to reduce the uncertainty in concerned field. The preservation and transmission of experiments and knowledge through formalized processing-chains facilitates the implementation of new experiment protocols using meteorological reanalysis data. More other, the capitalization and valuation of experiments using processing-chains is part of knowledge management. It means that the starting point is strong hypothesis that all experience or knowledge can be organized, referenced, enriched on supports which are adapted and exchanged as knowledge that other people can make their own. This puts the stress on the re-use, interoperability and sharing, the re-use, the spread of processing-chains to value the new meteorological reanalysis datasets available as well as the new innovating methods for the monitoring of rainfall evolution.

The interests in formalizing processing-chains can be valued by the proposal on setting up a platform while allows the sharing, the mutualization the spread, the naming of meteorological reanalysis data processing-chains through catalogs for the monitoring of rainfall evolution. In addition, the existing platforms at the world level propose functionality for knowledge

management following this view; we have proposed platforms for sharing meteorological reanalysis datasets to deal with the uncertainty due to the lack of climatological data of this type.

Knowledge management consists in grouping, sharing, and updating knowledge. It requires not only setting up mechanisms and procedures for uniting, organizing, presenting and spreading meteorological reanalysis data processing-chains which are modeled and partners, but also proceeding to the assessments of these operation-results.

We have demonstrated how to produce meteorological reanalysis data processing-chains which are reusable, sharable, spreadable, interoperable, and mutualizable. The step which can follow is implementation of a platform for sharing experiments. In view the interest the valorization of our proposals, we can draw up a deduction that the solutions we found represent a help in the formalization of meteorological reanalysis data processing-chains to capitalize the knowledge of the monitoring of the evolution of rainfall evolution.

#### **4. CONCLUSION AND PERSPECTIVES**

This article shows the possibility for sharing and mutualizing the experiments or existing meteorological reanalysis data processing to deal with the uncertainty of the monitoring of rainfall evolution using the technique of climatology. The proposal on resource modeling and formalization for their sharing is one of the efficient techniques for resolve the problems of heterogeneity of meteorological reanalysis data processing. In fact the experiments or the processing methods of these reanalysis data used for the monitoring of climate change will be reusable for non-climatology-expert-users. This decreases the uncertainty of the use of meteorological reanalysis data processing in the rainfall evolution area.

Our work is limited to modeling, formalization and design of a resource-knowledge base related to the Work-Context. This latter has its advantage because it can be enriched by new recently-produced resources. It helps integrate new meteorological reanalysis datasets and recent processing methods for analysing rainfall evolution.

Next, the perspective of this work is the integration of resources formalized in that way. This formalization makes the resources interoperable, which makes it possible to integrate the resources easily into the MDWeb tools for the sharing and mutualization and WPS (Web Processing Services), for the execution. For the execution of processing-chains via the web, we propose the web WPS Service server OGC standard [14] since 2005. The specification of the WPS is in the form of a generic interface allowing describe and carry out or execute spatial data processing chains of meteorological satellite data, according to [15] and [16], which we produced in this work. The WPS is based on the http protocol and XML language. The processing-chains presented in this work will be converted into XML files, and will be integrated into WPS for execution. It is also possible to implement the processing chain formalize using different kind of processing server according to [17] which propose a consolidated model of procedures for workflow management.



## REFERENCES

- [1] Uppala S., Kållberg P., Simmons A., Andrae U., Bechtold V., Fiorino M., et al, (2005) “The ERA-40 re-analysis” *Q J R Meteorol Soc*, 131, pp. 2961–3012
- [2] Saha S., Moorthi S., Pan H., Wu X., Wang J., Nadiga S., et al, (2010) “The NCEP climate forecast system reanalysis” *Bull Am Meteorol Soc*, 91, pp. 1015–1057
- [3] Rienecker M., Suarez M., Gelaro. R., Todling R., Bacmeister J., Liu E. , et al, (2011) “MERRA: NASA's modern-era retrospective analysis for research and applications” *J Clim*, 24, pp. 3624–3648
- [4] Hajalalaina A. (2015) “Aide à la formalisation des chaînes de raitement de données spatiales Capitalisation des connaissances sur les climat à Madagascar“. Phd thesis, Computer Science, University of Fianarantsoa, Madagascar. p.313, 27, november 2015,
- [5] Raheri,irina A., Hajalalaina A., (2016) “Markovian model for rainfall data. A case study on the monthly rainfall in Madagascar from 2013 to 2014” SADA'2016, 28 Nov-3 Dec 2016, Cotonou, Bénin
- [6] Hajalalaina A., Hervé D., Razafimandimby J., Delaître E., Desconnets J., Libourel T., (2015) “Formalisation des chaînes de traitements de données spatiales satellitaires sur la forêt à Madagascar“. In Hervé D., Razanaka S., Rakotondraompiana S., Rafamantanantsoa F., Carrière S. (eds.):*Transitions agraires au sud de Madagascar. Résilience et viabilité, deux facettes de la conservation. Actes du séminaire de synthèse du projet FPPSM, 10-11/06/2013, Antananarivo, IRD-SCAC/PARRUR, Ed. MYE, Article 15, pp. 237-250.*
- [7] Hajalalaina A., Hervé D., Delaitre E., Libourel T., (2016) Modeling process chain of SPOT images for resources uncertainty to monitor change in forest cover. In *Proceedings of Spatial Accuracy 2016, Spatial uncertainty in knowledge-based systems, ISBN: 978-2-9105-4510-5, Montpellier, France, pp. 38-45.*
- [8] Desconnets J., Libourel T., Clerc S., Granouillac B., (2007) “Cataloguing for distribution of environmental resources”. 10th AGILE, International Conference on Geographic Information Science, Aalborg University, Danmark.
- [9] BnF, (2008). Guide d'utilisation du Dublin Core (DC) à la BnF : Dublin Core simple et Dublin Core qualifié, avec indications pour utiliser le profil d'application de TEL, Version 2.0. Bibliothèque nationale de France /Direction des Services et des Réseaux / Département de l'Information bibliographique et numérique, France.
- [10] ISO19115, (2003) “Geographic Information Metadata, ISO 19115”. International Organization for Standardization (ISO), Genève, Suisse.
- [11] ISO19119, (2005) “Geographic Information Service, ISO 19119”. International Organization for Standardization (ISO), Genève, Suisse.
- [12] Libourel T., Lin Y., Mougnot I., Pierkot C, (2010) “A platform dedicated to share and mutualize environmental applications” In J. Filipe, J. Cordeiro J. (eds.), *Proceedings of the 12th international conference on enterprise systems. ICEIS, International Conference on Enterprise Systems, 12., Madere Funchal, 8-12 juin 2010. Setubal : SciTePress, p. 50-57.*
- [13] Lin Y., Mougnot I., Libourel T., (2008) “Un nouveau langage de workflow pour les sciences expérimentales“. *INFORSID'08 : Atelier ERTSI Evolution, Réutilisation et Traçabilité des Systèmes d'Information, Fontainebleau, France.*

- [14] OGC, (2016) “Web Processing Service 2.0 Standard Working Group”.  
<http://www.opengeospatial.org/projects/groups/wps2.0swg>. Accessed: 2016-04-29
- [15] Eberle J., Strobl C., (2012) “WEB-Based Geoprocessing and Workflow Creation for Generating and Providing Remote Sensing Products”. Geomatica, Vol.66(1), pp.13-26. Canadian Institute of Geomatics.
- [16] Machet E., Kamhi M., Jacquin M., Le Page M., Dejoux J., Dedieu G., (2008) “Web Processing Service pour le traitement des images satellites“. CESBIO, CNES, Toulouse, France.
- [17] Hadzhikoleva S., Hadzhikolev E., (2014), “consolidated model of procedures for workflow management”, International Journal of Information Technology, Modeling and Computing (IJITMC) Vol.2, No.1.

## AUTHORS

### **HAJALALAINA Aimé Richard**

Phd at Ecole de Management et d'Innovation Technologique, University of Fianarantsoa, Madagascar.

Research Area: Computer science, Computer Application, Software Engineering, business intelligence, Geomatic and Applications,



### **RAHERINIRINA Angelo Fulgence**

Phd at Ecole Normale Supérieure, University of Fianarantsoa,, Madagascar.

Research Area: Modeling and analysis of stochastic systems, Markov models Monte Carlo Method by Marko Chain (MCMC), Bayesian Statistics, Didactics of mathematics.



### **RATIARISON Adolphe**

Professor at Dept. Physic and application, laboratory of Dynamics of Atmosphere, Climate and Oceans, University of Antananarivo.

Research Area: Physic and Applications, climatology, mathematics and applications, computer science and applications

### **LIBOUREL Therese**

Professor at Laboratory of Informatic, Robotic, Mathematics, university of Montpellier, France

Research Area: Databases, Information Systems (Business Informatics), Software Engineering, Geomatic

