

Fiche de poste

Métier ou emploi type* :

* REME, REFERENS, BIBLIOPHILE

Post-doc on global scale species and gene clusters distribution modelling from Tara Oceans data using machine learning *Blue Cloud project*

Catégorie :

Corps :

BAP (si ITRF) :

Les activités qui composent la fiche de poste sont appelées à évoluer en fonction des connaissances du métier et des nécessités de service.

Présentation de Sorbonne Université

Pour transmettre les connaissances, comprendre le monde et relever les défis du 21^e siècle, une nouvelle université est née le 1^{er} janvier 2018, issue de la fusion entre les universités Paris-Sorbonne et Pierre et Marie Curie. Sorbonne Université est une université pluridisciplinaire, de recherche intensive et de rang mondial. Ancrée au cœur de Paris, présente en région, elle est engagée pour la réussite de ses étudiants et s'attache à répondre aux enjeux scientifiques du 21^e siècle. www.sorbonne-universite.fr

Presentation of the laboratory and team

Description (missions, team,...) : The Laboratory of Oceanography of Villefranche-sur-mer (LOV; <http://lov.obs-vlfr.fr/>) belongs to one of the three marine stations of Sorbonne Université; it is located close to Nice, on the French Riviera. With about 40 permanent staff and 90 total staff, the LOV is a leading research laboratory in oceanography, tackling questions ranging from marine ecology to biogeochemistry, producing ~100 scientific papers per year, and receiving prestigious grants from ANR, the European Commission and various international funding agencies.

The COMPLEX (COMPUtational PLankton Ecology) team gathers ~30 members studying marine plankton. Plankton encompasses all organisms drifting with marine currents. Those organisms are responsible for producing half the oxygen we breathe, storing the carbon we emit, feeding the fish we eat; plankton is therefore a major building block of Earth's ecosystem. In COMPLEX, we use numerical methods (modelling, statistics, machine learning), quantitative imaging instruments and, increasingly, meta-omics data to further our understanding of plankton. We strongly interact with the Quantitative Imaging Platform of Villefranche (PIQV; <https://sites.google.com/view/piqv>), which oversees the daily operation of tools that the team develops. Those tools include imaging sensors, such as the Underwater Vision Profiler or the ZooScan, and software packages, such as ZooProcess or the EcoTaxa web application (<https://ecotaxa.obs-vlfr.fr/>) that uses machine learning to assist taxonomists into sorting plankton images. COMPLEX has a large network of collaborators, in oceanography, imaging, genomics, and computer sciences, in particular as part of the Tara Oceans consortium.

Localisation : Villefranche-sur-Mer, France

Missions et main activities

Mission : The scientific goal of this post-doctoral position is to predict the potential distribution of planktonic entities at the scale of the world's ocean based on a limited number of points where we know their concentrations. Knowing that plankton is strongly constrained by the conditions of the water mass it lives in, we will quantify the relationships between plankton and its environment at the points where we have data and then project those relationships at global scale (a.k.a. species distribution modelling).

This approach is classic but the main originality of this project is that the entities to predict will come from a deep analysis of the meta-omics data of Tara Oceans by a current post-doc. They may therefore be "species" but also clusters of genes associated with a function or coherent entities whose identity is unknown (i.e. genomic "dark matter"). The modelling of their distribution will therefore help in the interpretation of their nature and their role in the ecosystems.

While we will start with established methods for a few entities, to provide a baseline, further possible developments include: (i) predicting multiple entities simultaneously to exploit the information of their relative concentrations, (ii) using deep learning methods for species distribution modelling to capture environmental context, (iii) using ~3000 points from a worldwide, quantitative, imaging dataset and fine resolution satellite data to compare with the Tara Oceans one. These developments, and others, will be discussed and decided with the selected candidate.

The goal of the Blue Cloud project that funds this position is to provide a scientific infrastructure for others to use. Therefore, the work will be embedded into this EU-scale infrastructure and should be made accessible through well documented code or a web-based application.

Finally, it is expected from the post-doc to disseminate his/her work within the Blue Cloud community (especially during dedicated workshops) and beyond (international conference, scientific publication, code and product sharing, etc.).

Supervision of other staff : NO

Knowledge and skills*

Qualifications and experience :

- Hold or be close to completion of a Ph.D. in numerical ecology or applied mathematics.
- Experience with machine learning methods, ideally with deep learning (in a context of regression, not classification)
- Basic knowledge of ecology and ideally experience with species distribution models (a.k.a. niche models)
- Excellent writing and communication skills to write papers and present findings, but also write basic code documentation.

Skills :

- Good programming skills and ability to handle large datasets, with R and/or Python
- Experience with querying netCDF files would be a plus
- Scientific and technical English (B2 level for written and oral). French would be a plus, but it is not mandatory.

Human skills :

- Ability to work as part of a team, within an international project

Particular conditions :

For more information or to submit your application, please email Lionel GUIDI (guidi@obs-vlfr.fr), Sakina-Dorothée AYATA (sakina@obs-vlfr.fr), and Jean-Olivier IRISSON (irisson@obs-vlfr.fr) with "Blue Cloud Post-doc" in the title.

To apply, submit a single PDF file with:

- Curriculum vitae ; in addition to the standard fields, the CV should list publications and give contact information for at least two persons of reference
- Letter of intent explaining your interest in the position and your view on the work to be done

If some papers are small enough to be attached too, it would be appreciated.

* Conformément à l'annexe de l'arrêté du 18 mars 2013 (NOR : MENH1305559A)