

# Postdoc ML Ecodep

Application possible until July 1, 2022

position beginning October 1<sup>st</sup>, 2022

(12 months renewable once)

**Post-Doc position in the research project Ecology and Dependence**

**ECODEP funded by CY Initiative of Excellence**

(grant "Investissements d'Avenir" ANR-16-IDEX-0008).

Laboratory AGM-UMR 8088, CY Cergy Paris Université

**gross revenue monthly 3081€, net before income tax 2500€**

## How to apply

Please send all application package to: [doukhan@cyu.fr](mailto:doukhan@cyu.fr) and [jennifer.denis@cyu.fr](mailto:jennifer.denis@cyu.fr)

Needed documents:

- PhD or equivalent diploma
- cover letter
- detailed curriculum vitae
- copy(s) of the diploma(s)
- thesis defense report

**Profile** This post within the EcoDep project <https://doukhan.u-cergy.fr/ecodep.html> will contribute from both a theoretical and applicative point of view. This postdoc support will consider various issues in Machine Learning for application to ecology data sets and their interpretation. Beyond the standard ML tools, Oracle inequalities, variable selection, individual sequences or model based predictive methods, the project ecodep want to set a special emphasis on IML interpretable Machine Learning to consider issues of high dimensional time series, eg for abundance data sets or for various models with covariates as meteorological data. Specific skills in time series analysis are welcome. IML methods either analyze model components, model sensitivity, or surrogate models.

The post-doc position will be based in Cergy Pontoise, CY Cergy Paris Université, with collaborations in France and abroad for collaborative work within the Ecodep community. English is necessary: a candidate should thus be able to travel by invitations. Travel to Columbia (NYC), Santiago (Chile), Iena (Germany) and several other locations connected to ECODEP are envisioned (depending on the evolution of the pandemic) for research work on:

- 1- Extensions and applications of Taylor's law
- 2- Modeling of abundance
- 3- Population dynamics
- 4- Time series issues: isotonicity, causality, covariates, selection
- 5- Partly observed processes and applications

- 6- Random fields, space time models and their use
- 7- Panel data studies
- 8- Risks and data-based studies

Potential locations will depend on the skills and interests in those initial important questions which are not considered in all the Ecodep labs, but CYU will be a fixed point in this position.

We are looking for a statistician wishing to be involved in issues and population dynamics. This is why we report below several questions of importance in the project. The modelling of population dynamics is of paramount importance in many areas of application.

### **Qualification** PhD in Mathematical Statistics

**Recap of the aims of the project:** In ecology, it is a matter to understand the dynamics and life history of various species through different environments. Indeed, environmental changes can generate rapid changes in the composition of a given population, its length, its phenotypic character or also its genotype distribution. Demography is generally concerned with predicting human lifespan as well as the population structure with critical involvement in pension systems and public policy decision making. However, these dynamics raise a number of problems to which historical experience does not provide an answer.

The research proposal is to model population growth and to predict biodiversity using innovative stochastic models with a specific focus on ecological problems. The relevant aspects are related to Taylor's laws [C1998]. In addition, the difficult problems facing marine ecology, in particular those related to the evolution of the environment and its impact on marine species will be of interest. Finally, the applications will be devoted, among others, to the effects of climate change on coral reefs, ecological abundance modelling and the prediction of marine ecosystems [CFM2006], [R2019], [GH2019]. The biostatistical models are also of interest for the project [DFL2017] and they deserve additional attention.

Some of the stylized facts encountered when working with real-world datasets will be enlightened in the postdoc practice. New modelling frameworks for populations dynamics will incorporate, for instance, covariates and we will investigate their statistical properties [GX2017], [KT2019]; see e.g. observation models as in [GK2020]. These problems involve isotonic models, parsimony in the presence of non-linearity and non-stationarity [FLN2018], [D2018]. To conclude, causality [W1954] is of importance also to reduce dimensions as noticed in [Z2018]. Publications appeared in the frame of the project ECODEP, see <https://doukhan.u-cergy.fr/publications.html>.

### **Bibliography**

[CFM2006] Champagnat N, Ferrière R, Méléard S (2006). Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. Theoretical population biology.

- [C1998] Cohen JE (1998). How Many People Can the Earth Support? Bulletin of the American Academy of Arts and Sciences
- [D2018] Doukhan P (2018). Stochastic Models for Time Series. Springer, Mathematics and Applications 80.
- [DMT2020] P. Doukhan, M. H. Neumann, L. Truquet (2020). Stationarity and ergodic properties for some observation-driven models in random environments. Preprint.
- [FLN2018] Fokianos K, Leucht A, Neumann MH (2018). On Integrated  $L^1$  Convergence Rate of an Isotonic Regression Estimator for Multivariate Observations. arXiv:1710.04813v2.
- [GH2019] Goncalves B, Huillet T (2019). Scaling Features of Two Special Markov Chains Involving Total Disasters. Journal of Statistical Physics.
- [GK2020] Gorgi P, Koopman SJ (2020) Beta observation-driven models with exogenous regressors: a joint analysis of realized correlation and leverage effects. TI 2020-004/III, Discussion Paper; <https://www.tinbergen.nl>.
- [GX2017] Gouriéroux C, Zakoian JM (2017). Local explosion modelling by noncausal process. Journal of the Royal Statistical Society: Series B.
- [M2018] P Marquet et al. (2018). Modelling the current and future biodiversity distribution in the Chilean Mediterranean hotspot. Diversity and Distributions.
- [M2021] C Molnar (2021) Interpretable Machine Learning – A Brief History, State-of-the-Art and Challenges. Preprint.
- [R2019] Rebolledo R et al. (2019) Open-System Approach to Complex Biological Networks. SIAM
- [Z2018] Zhang Z (2018) Multivariate Time Series Analysis in Climate and Environmental Research, Springer.

### **Mobility in Laboratories associated to Ecodep project**

#### ***Laboratories associated to Paris Seine Initiative***

Department of Economics, University of Warwick, UK

Department of Economics and Statistics (FSSH), University of Mauritius

#### ***Laboratories in France***

ISFA, University Lyon 1, Lyon

ENSAI/CREST, Rennes

LMBA, UBS, University of South Brittany, Vannes

#### ***Laboratories abroad***

Laboratory of Populations, Rockefeller University and Columbia University, USA

Department of Statistics, Columbia University, New-York, USA

Pontificia Universidad Católica de Chile, Department of Ecology, Santiago, Chile

Pontificia Universidad Católica de Chile, Department of Statistics, Valparaiso, Chile

CIMFAV, University of Valparaiso, Valparaiso, Chile

Friedrich Schiller University Jena, Department of Mathematics, Germany