David Causeur, Ph.D.   Curriculum Vitæ

French citizen

Director of the Department of Statistics and Computer Science, Agrocampus, Rennes, France.
Current position: Professor in Applied Statistics, Agrocampus, Rennes, France.
Head of the Master of Science in Applied Statistics for Agricultural and Agri-food applications, Agrocampus, Rennes, France.

Contact

Department of Statistics and Computer Science
Agrocampus Ouest
CS 84215 - 65 rue de Saint-Brieuc
35042 Rennes cedex

Phone: +33.223.485.884
Fax: +33.223.485.871
Email: david.causeur@agrocampus-ouest.fr
Web: http://www.agrocampus-ouest.fr/math/causeur/

Education

2005   **Research habilitation** in Mathematics, University of South Brittany (Vannes, France).
Title : Gauss-Markov type estimation in linear model with incomplete data & Spatial interaction between homologous factors
Committee
Michel Carbon, Michel Delecroix, Thierry Dhorne, Ion Grama
Reviewers: Jean-Marc Azaïs (University Paul Sabatier, Toulouse), Subir Ghosh, (University of California, Riverside) and Christian Robert (University Paris IX Dauphine)

1997   **Ph.D.** in Mathematics and Applications, University Rennes I (Rennes, France).
Title : Statistical inference in linear model with incomplete data.
Committee
Michel Bonneu, Jean-Baptiste Denis, Jean Deshayes
Supervisor: Thierry Dhorne
Reviewers: Jean-René Mathieu (University Paul Sabatier, Toulouse), Jean-Jacques Téchené (University of Pau)

1993   **Master of Science** in Pure Mathematics and Applications, University Rennes I

Positions

2007-	**Professor** in Applied Statistics at Agrocampus, Rennes, France.
2005-	**Research fellow** at IRMAR (National Center for Scientific Research), Rennes, France.
2002-2005	**Research fellow** at the Statistics and Modeling Department (Research Center for Economics and Statistics), Rennes, France.
1999-2001	**Assistant professor** in Statistics at University of South Brittany, Vannes, France.
1998-2001	**Research fellow** at the Applied Statistics Department of University of South Brittany, Vannes, France.
1998-2007	**Assistant professor** in Applied Mathematics at Agrocampus, Rennes, France.
1996-1997	**Teaching and Research Assistant** in Applied Mathematics at Agrocampus, Rennes, France.

Professional activities

Since 2012	Director of the Department of Statistics and Computer Science at Agrocampus, Rennes, France.
2010-2012	Co-head of the Undergraduate program in Biology and Mathematics, University Paris VI, Paris, France.
2011	President of the organizing committee of the 8th Rennes Statistical Meeting, Rennes, France
2008	President of the organizing committee of the 6th workshop on Statistical Methods for Post-Genomic Data, Rennes, France.
2006-2008	Elected member of the scientific council of University Rennes II.
Since 2005	Elected member of the board of directors at Agrocampus, Rennes, France.
Since 2005	Head of the Master of Science in Applied Statistics for Agricultural and Agri-food applications, Agrocampus, Rennes, France.
2006-2008	Elected member of the scientific council of Agrocampus, Rennes, France.
2005-2010	Co-organizer of the weekly statistics seminar at IRMAR (National Center for Scientific Research), Rennes, France.
2005	President of the organizing committee of the 2nd Rennes Statistical Meeting, Rennes, France

Refereeing: Annals of Applied Statistics, Biostatistics, Behavior Research Methods, BMC Bioinformatics, Communications in Statistics, Computational Statistics and data Analysis, Electronic Journal of Applied Statistics, International Journal of Statistics and Management Systems, Journal of Statistical Computation and Simulation, Journal of Statistical Planning and Inference, Metabolites, Metron, Scandinavian Journal of Statistics, Statistical Applications in Genetics and Molecular Biology, Wiley (books).

Teaching experience

Topics:
1st year undergraduate program Biology & Mathematics
    Differential equations and functional analysis.
1st year Master of Science
    Data analysis, Elementary inferential statistics, Linear modeling, Generalized linear modeling.
2nd year Master of Science
    Compartment modeling, Functional data analysis, High-dimensional data analysis, Nonlinear modeling, Spatial statistics, Statistical genetics, Statistical learning.

Research interests

Topics:
Inference for high-dimensional data
    Post-genomic data, ERP data, Large-scale simultaneous inference, Factor analysis, Stability selection.
Incomplete observations
    Estimation of variance components, Optimal sampling designs.
Spatial interaction between homologous factors
    Generalized biadditive models, Bradley-Terry model, Random matrix theory, Dimensionality.

Ph.D. director:

Defence	Topic
2016	Stability of model selection in high-dimensional supervised classification (E. Perthame).
2015	Assessment of statistical learning methods in genomic data analysis (R. Hornung)
2014	Gauss-Markov estimation in survey sampling with non-response (B. Gelein).
2014	Statistics for the monitoring of surveillance data in epidemiology (A. Lehebel).
2012	Modeling gene regulatory networks using transcriptomic data - Application to eQTL detection. (Y. Blum).
2009	Linear gaussian state-space model with irregular sampling - Application to Sea Surface Temperature (P. Tandeo).
2009	Impact of dependence in large-scale multiple testing (C. Friguet)

Selected Ph.D. committee memberships:

Defence	Topic
2015	Incorporating the linkage disequilibrium (LD) information in Genome-Wide association studies. (A. Dehman, PhD).
2015	Geometric approaches of genetical data analysis. (D. Laloe, Research Habilitation).
2015	Contributions to environmental, ecological, medical and archeological data exploration methodology. (L. Bellanger, Research Habilitation).
2014	Nonparametric Short Term Forecasting Electricity Consumption. (L. Hatton, PhD).
2014	Vietnamese food choice model: an approach from food products and consumers. (Nguyen Ba Thanh, University of Ho-Chi-Minh city, PhD).
2014	Optimization of association genetics and genomic selection strategies for populations of different diversity levels. (R. Rincent, PhD).
2014	Dealing with uncertainties in a complex marine ecosystem model. (L. Gasche, PhD).
2013	Accounting for the growing pigs' variability in models of nutritional needs. (B. Vautier, PhD).
2013	Statistics for genomic data analysis (Habilitation) (M.-L. Martin-Magniette, Research Habilitation).
2012	Statistical methods for robust analysis of transcriptome data by integration of biological knowledge. (M. Jeanmougin, PhD).
2012	Phenotypic prediction and variable selection in high dimensional linear and linear mixed models. (F. Rohart, PhD).
2012	The repercussions of statistical properties of interval mapping methods on eQTL detection (X. Wang, PhD).
2011	High-dimensional statistical learning of molecular genetics data (A.-L. Boulesteix, Research Habilitation).
2010	Identification of the key elements in lipid metabolism and regulation (P. Blavy, PhD)
2007	Regularized canonical analysis for highly multidimensional data (I. Gonzalez, PhD).
2007	Exploratory and analytical epidemiology of the Developmental Orthopaedic Disease in the limbs of foals (J. Lepeule, PhD)
2005	Deposition of dietary fatty acids, de novo synthesis and anatomical partitioning of fatty acids in finishing pigs. (M. Kloareg, PhD)

Selected Research grants:

2012-2014	MOLoSSE: Statistical modeling of dependence in GWAS (CNRS project).
2012-2013	Statistical Testing of Intensive, High-Dimensional Data in Behavioural Sciences (CampusFrance project with NCKU, Tainan, Taiwan).
2012-2014	Assessment of statistical learning methods in genomic data analysis (French-Bavarian project with University of Munich, Germany).
2011-2014	Identification of key regulators of lipid plasticity by combined high-throughput technologies, statistics, bioinformatics and phylogenics approaches. (ANR project with with INRA)
2000-2003	EUPigClass: Pig carcass classification in Europe (European project, 5th FP).

Articles in refereed journals:

1. Gondret, F., Vincent, A., Houée-Bigot, M., Siegel, A., Lagarrigue, S., Louveau, I. and **Causeur, D.** (2015). Communalities and particularities in the molecular responses of porcine adipose tissues at two anatomical locations to a high fat high fiber diet. *Submitted to BMC Genomics*.
2. Hornung, R., Boulesteix, A.-L. and **Causeur, D.** (2015). Combining location-and-scale batch effect adjustment with data cleaning by latent factor adjustment. *Under 2nd review in BMC Bioinformatics*.
3. Blum, Y., Houée-Bigot, M. and **Causeur, D.** (2015). Sparse factor model for co-expression networks with an application using prior biological knowledge. *Under 2nd review in Statistical Applications in Genetics and Molecular Biology*.
4. Sheu, C-.F., Perthame, E., Lee, Y-.S., **Causeur, D.** (2015). Accounting for time dependence in large-scale multiple testing of event-related potential data. *Accepted for publication in Annals of Applied Statistics*.
5. Perthame, E., Friguet, C. and **Causeur, D.** (2015) Stability of feature selection in classification issues for high-dimensional correlated data. *Statistics and computing*. Doi 10.1007/s11222-015-9569-2. pp. 1-14.
6. Gelein, B., Haziza, D. and **Causeur, D.** (2014) Preserving relationships between variables with MIVQUE based imputation for missing survey data. *Journal of Multivariate Analysis*. 131, 197–208.
7. Nyemb, K., Jardin, J., **Causeur, D.**, Guérin-Dubiard, C., Dupont, D., Rutherfurd, S.M. and Nau, F. (2014) Investigating the impact of ovalbumin aggregate morphology on in vitro ovalbumin digestion using label-free quantitative peptidomics and multivariate data analysis. *Food Research International*, 63(B), 192-202.
8. **Causeur, D.**, Chu, M.-C., Hsieh, S. and Sheu, C.-F. (2012) A factor-adjusted multiple testing procedure for ERP data analysis. *Behavior Research Methods*. **44**, 635–643.

9. Mach, N., Blum, Y., Bannink, A., **Causeur, D.**, Houée-Bigot, M., Lagarrigue, S. and Smits, M.-A. (2012) Pleiotropic effects of polymorphism of the gene diacylglycerol-O-transferase 1 (DGAT1) in the mammary gland tissue of dairy cows. *Journal of Dairy Science.* **95** (9), 4989–5000.
10. Blum, Y., Le Mignon, G., **Causeur, D.**, Filangi, O. Désert, C. Demeure, O. Le Roy, P. and Lagarrigue, S. (2011) Complex trait subtypes identification using transcriptome profiling reveals an interaction between two QTL affecting adiposity in chicken. *BMC Genomics.* 12:567.
11. **Causeur, D.**, Friguet, C., Houée-Bigot, M. and Kloareg, M. (2011). Factor Analysis for Multiple Testing (FAMT): An R package for Large-Scale Significance Testing Under Dependence. *Journal of Statistical Software.* **40** (14), 1–19.
12. Friguet, C. and **Causeur, D.** (2011). Estimation of the proportion of true null hypotheses in high-dimensional data under dependence. *Computational Statistics and Data Analysis*, **55** (9), 2665–2676 .
13. Caffier V., Didelot F., Pumo B., **Causeur D.** and Parisi L. (2010). Aggressiveness on a susceptible apple cultivar of eight Venturia inaequalis isolates virulent or avirulent towards the major resistance gene Vf. *Plant Pathology.* **59**(6). 1072–1080.
14. Blum, Y., Le Mignon, G., Lagarrigue, S. and **Causeur, D.** (2010). A factor model to analyse heterogeneity in gene expressions. *BMC Bioinformatics.* 11:368.
15. Friguet, C., Kloareg, M. and **Causeur, D.** (2009). A factor model approach to multiple testing under dependence. *Journal of the American Statistical Association.* **104** (488) 1406–1415.
16. **Causeur, D.**, Kloareg, M. and Friguet, C. (2009). Control of the FWER in multiple testing under dependence. *Communications in Statistics - Theory and Methods.* 38 (16 & 17) 2733–2747 .
17. Ablain, W., Hallier Soulier, S., **Causeur, D.**, Gautier, M. & Baron, F. (2009) A simple and rapid method for the disruption of Staphylococcus aureus, optimized for quantitative reverse transcriptase applications: Application for the examination of Camembert cheese. *Dairy Science and Technology.* 89 (1) p 69.
18. Cutullic, E., Delaby, L., **Causeur, D.**, Michel, G. & Disenhaus, C. (2009) Hierarchy of factors affecting behavioural signs used for oestrus detection of Holstein and Normande dairy cows in a seasonal calving system. *Animal Reproduction Science.* 113 (1-4), 22–37
19. Kloareg, M. & **Causeur, D.** (2009). Double-sampling designs to reduce the non-discovery rate. Application to microarray data. *Journal of Data Science.* 7 (2), 219–234
20. **Causeur, D.** and Husson, F. (2008). Asymptotic power of double-sampling tests for general linear hypothesis. *Statistics.* 42, 2. 115-125.
21. Kloareg, M. and **Causeur, D.** (2007). Improving Type-II error rates of multiple testing procedures by use of auxiliary variables. application to microrray data. in Recent advances in Stochastic Modelling and data Analysis, Ed. Christos Skiadas, World Scientific Publishing, Co Pte Ldt.
22. **Causeur, D.** (2006) MIVQUE and maximum likelihood estimation for multivariate linear models with incomplete observations. *Sankhyà.* **68** (3), 408-434.
23. **Causeur, D.** (2005). Optimal sampling from concomitant variables for regression problems. *Journal of Statistical planning and Inference.* **128**, 289-301.
24. **Causeur, D.**, Dhorne, T. & Antoni, A. (2005). A two-way analysis of variance model with positive definite interaction for homologous factors. *Journal of Multivariate Analysis.* **95** (2), 431-448.
25. **Causeur, D.** & Husson, F. (2005). Causeur, D. & Husson, F. (2005). A 2-dimensional extension of the Bradley-Terry model for paired comparisons. *Journal of Statistical planning and Inference* **135**, 245-259.
26. Ledauphin, S., Vigneau, E. & **Causeur, D.** (2005). Functional approach for the analysis of Time Intensity curves using B-splines. Journal of Sensory Studies. **20** (4), 285-300.
27. **Causeur, D.** & Dhorne, T. (2003). Linear Regression Models under Conditional Independence Restrictions. *Scandinavian Journal of Statistics* **30** (3), 637-650.
28. Kamp, J.F., Marin-Mamellet, C., Forzy, J.F. and **Causeur, D.** (2001). HMI Aspects of the Usability of Internet Services with an In-car Terminal on a Driving Simulator. *IATSS Research.* **25**, 2, 29-40.
29. **Causeur, D.** (1999). Exact distribution of the regression estimator in double-sampling. *Statistics* **32**, 297-315.
30. **Causeur, D.** & Dhorne, T. (1998). Finite-sample properties of a multivariate extension of double-regression. *Biometrics* **54** (4), 1591-1601.
31. **Causeur, D.** (1998). Plan d'échantillonnage en plusieurs phases pour la réduction des coûts expérimentaux en régression linéaire. *Revue de Statistique Appliquée.* **XLVI** (4), 59-73.

Software development:

ERP	R package for Event-Related Potentials data analysis.
FADA	R package for supervised classification of high-dimensional correlated data.
FAMT	R package for large scale significance analysis with highly dependent data.

Selected talks since 2010:

1. **Causeur, D.** (2013) Large Scale Mutiple Comparisons in Ranking Studies. Invited Seminar, November 2013. NCKU, Tainan, Taiwan.
2. **Causeur, D.** (2013) Sparse factor modeling of high-throughput data. Invited Seminar, March 2013, University of Montpellier.
3. **Causeur, D.** (2013) Sparse factor models for high-dimensional interaction networks. Invited talk, May 2013, 7th International Workshop on Simulation, Rimini, Italy.
4. **Causeur, D.** and Sheu, C.-F. (2012) Large scale significance testing under dependence. Invited Seminar, November 2012. NCCU, Chia-Yi, Taiwan.
5. **Causeur, D.** (2012) Sparse factor model for high-dimensional data. Invited Seminar, November 2012. Academia Sinica, Taipei, Taiwan.
6. **Causeur, D.**, Blum, Y., Houée-Bigot, M. and Cadoret, M. (2012) Inference on sparse interaction networks using factor models. Statistical meeting of the French Statistical Society. Brussels, Belgium.
7. **Causeur, D.** (2012) Sparse factor models for gene interaction networks. Invited Seminar, March 2012. University of Limerick, Ireland.
8. **Causeur, D.** (2011) Sparse factor models for inference on gene regulatory networks. Invited Seminar, December 2011. IBE, University of Munich, Germany.
9. **Causeur, D.** (2011) High-throughput data analysis in systems biology. 2nd international workshop in Biostatistics, Invited talk, Bio-SI 2011. Rennes, France.
10. **Causeur, D.** & Lê, S. (2011) Integrating biological knowledge in gene expression data analysis. Invited session, ISI 2011 (International Statistical Association), Dublin, Ireland.
11. **Causeur, D.** (2011) Feature selection stability in high-dimensional heterogeneous data. Invited talk, Cladag 2011. Pavia, Italy.
12. **Causeur, D.** (2011) Stability of model selection in supervised classification. Asmda 2011, Rome, Italy.
13. Blum, Y., Lagarrigue, S. and **Causeur, D.** (2010) A factor model to analyze heterogeneity in gene expression in a context of QTL characterization. International Society for Animal Genetics, Edinburgh, 26-30 July 2010.
14. Blum, Y., Friguet, C., Lagarrigue, S. and **Causeur, D.** (2010) Inference of gene networks using factor models. Statistical meeting of the French Statistical Society, Marseille, 25-28 May 2010.