

Avant-propos

Aris Xanthos & Romain Loup

Université de Lausanne

{aris.xanthos,romain.loup}@unil.ch

L'édition d'un *Festschrift* est un exercice qui présente des points communs avec la pêche au chalut. On lance ses filets sans jamais savoir au juste ce que l'on y recueillera : poissons déjà repérés ou ramenés par des courants inattendus ; espèces familières ou venues de récifs lointains ; analyses approfondies, reprises de dialogues anciens ou échos de remous intellectuels que la personne honorée a contribué à faire naître. On sait en tout cas que la récolte, quelle qu'en soit la composition, sera le reflet d'un écosystème particulier – ici, celui dans lequel François Bavaud a évolué, enseigné, publié, influencé. Comme le chalut qui racle les fonds marins pour en rapporter la diversité vivante, ce volume ramène à la surface la richesse d'un environnement intellectuel, académique et institutionnel marqué par la présence de François.¹

Espace, langage, réseaux, science : voici, nous semble-t-il, une énumération concise des grandes thématiques qui structurent cette collection d'articles et reflètent des aspects centraux, récurrents et interconnectés de l'univers intellectuel de François Bavaud. Espace et langage sont naturellement des objets fondamentaux dans ses deux facultés de rattachement à l'Université de Lausanne (UNIL) – Géosciences et Lettres – et dont il a maintes fois exploré les aspects formels et mathé-

1 L'édition d'un *Festschrift* est aussi un travail d'équipe. A ce titre, les éditeurs tiennent à remercier chaleureusement Laura Delaloye Saillen pour son aide et sa relecture précieuses, Guillaume Long pour le dessin de la couverture, le Centre de linguistique et des sciences du langage, la Section des sciences du langage et de l'information et l'Institut de géographie et durabilité (tous trois à l'Université de Lausanne) pour le financement de cet ouvrage, enfin l'ensemble des auteur.e.s pour leurs contributions et leur participation à la relecture.

matiques. En témoignent notamment ses travaux autour de l'autocorrélation spatiale (cf. Bavaud, 2005, 2008b, 2013, 2024; Ceré & Bavaud, 2017; Loup & Bavaud, 2025), de l'analyse de flux (cf. Bavaud, 2008a; Bavaud et al., 2018; Guex & Bavaud, 2015; Guex et al., 2023) et de l'application aux données textuelles ou linguistiques de méthodes factorielles, markoviennes et basées sur la théorie de l'Information (cf. Bavaud, 2018; Bavaud et al., 2006; Bavaud & Xanthos, 2002, 2005; Cocco et al., 2011; Egloff & Bavaud, 2018). Considérations spatiales et linguistiques se rejoignent d'ailleurs dans certaines publications (Bavaud et al., 2012, 2015; Ceré et al., 2018). Comme on peut s'y attendre, les deux domaines sont bien représentés dans ce volume, au travers des contributions de Bivand, Jeanson, Kanevski et Rimaz & Métrailler (espace); Goldsmith & Ermolaeva, Jacquin, Margot et Meyer (langage); enfin Guex et Loup, dont les contributions présentent deux manières différentes de conjuguer spatial et linguistique.

En tant que mode de structuration de l'information, les réseaux ou graphes sont également au cœur des intérêts de recherche de François Bavaud et bon nombre de ses publications explorent leurs propriétés et applications (cf. Bavaud, 2006, 2010; Bavaud & Guex, 2012; Bavaud & Métrailler, 2023; Ceré & Bavaud, 2019, ainsi que plusieurs des publications déjà mentionnées au paragraphe précédent). Dans ce volume, les contributions de Courtain & Saerens, Pante, Rimaz & Métrailler, Rochat et Rozenblat & Rogov donnent une place fondamentale aux réseaux et à leur analyse. Il est d'ailleurs remarquable que trois d'entre elles aient aussi en commun d'être centrées sur le jeu (notamment vidéo), objet d'étude en plein essor à la faculté des Lettres de l'UNIL et mettent ainsi en évidence la pertinence des réseaux pour la modélisation et l'analyse formelles des objets (vidéo-)ludiques de multiples façons.

Quiconque est familier·ère du travail de François Bavaud conviendra sans doute que la science et sa méthodologie – en particulier la démarche hypothético-déductive, envisagée comme ensemble de pratiques et procédures formelles visant à construire des modèles à partir de données, jauger des théories à l'aune de faits empiriques, sélectionner des hypothèses sur la base d'observations (Bavaud, 1998a) –

sous-tendent la majeure partie de sa recherche et son enseignement. Ce bagage méthodologique initialement acquis dans sa formation en physique et mathématiques à l’Université de Fribourg et à l’EPFL, François l’a transporté en sciences humaines et sociales (sciences sociales et politiques à l’UNIL et psychologie à l’Université de Genève), pour enfin le déposer en Lettres et Géosciences à l’UNIL, contribuant depuis lors à y développer une perspective computationnelle sur la construction des savoirs en SHS qui constitue aujourd’hui encore l’une des particularités remarquables de cette institution. Rien d’étonnant, donc, à ce que les réflexions sur la science et les corollaires institutionnels que sont les notions de disciplinarité et d’interdisciplinarité aient trouvé leur place dans les contributions qui composent cet ouvrage, en particulier celles de Jacquin, Jeanson, Pante, Piotrowski et Rozenblat & Rogov, qui offrent plusieurs points de vue sur la place des méthodes quantitatives en sciences humaines.

Espace, langage, réseaux et science sont donc, assurément, des catégories qui permettent de structurer une part importante des travaux de François Bavaud, au moins depuis sa stabilisation à l’UNIL en 2002. Du point de vue de leur adéquation à la description des quinze contributions qui forment le présent volume, en tout cas, il faut les envisager comme des catégories *floues*, dans la mesure où la plupart des contributions se rattachent à plus d’une de ces thématiques. Une telle configuration implique qu’un partitionnement *dur* des contributions ne pourrait être effectué qu’au prix d’une réduction considérable de l’information initiale, compression *avec perte* à laquelle nous n’avons pu nous résoudre. C’est donc dans l’ordre alphabétique du nom des auteur·e·s qu’elles sont finalement présentées dans ce volume et brièvement résumées dans la suite de cet avant-propos.

La contribution de Roger Bivand (Norwegian School of Economics), « Implementing a weighted measure of multivariate spatial autocorrelation », explore la mise en œuvre de la mesure pondérée d’autocorrélation spatiale multivariée δ développée par Bavaud (2024). Bivand met en avant l’importance des poids régionaux dans l’analyse de l’autocorrélation spatiale, présente une implémentation de δ en R. Il l’applique à des données électorales françaises et examine différentes

méthodes de pondération spatiale. Sa contribution valide ainsi l'approche originale en reproduisant ses résultats et en comparant δ à des mesures classiques comme l'indice de Moran.

Dans leur contribution intitulée « A bag-of-paths graph model with Poisson-distributed path lengths », Sylvain Courtain et Marco Saerens (Université catholique de Louvain) étendent le modèle « bag-of-paths » en imposant une distribution de Poisson sur la longueur des chemins dans un réseau. Cette approche permet de contrôler l'impact de la longueur des chemins, avec des applications en classification et détection de communautés. Courtain et Saerens présentent une nouvelle distance entre nœuds (*Poisson surprisal distance*), basée sur la probabilité de choisir un chemin entre deux points. Cette distance est évaluée en classification (semi-)supervisée et obtient de bonnes performances relativement aux distances existantes.

La contribution de John Goldsmith (University of Chicago) et Marina Ermolaeva (Moscow State University), « Exploring oppositions in morphology », se penche sur la question – relativement peu étudiée hors de la tradition structuraliste de la première moitié du XX^e siècle – des méthodes et procédures formelles permettant d'identifier les unités élémentaires qui composent les productions langagières et les systèmes linguistiques sous-jacents. En particulier, Goldsmith et Ermolaeva proposent une réflexion sur les manières de formaliser la notion de différence entre des séquences de symboles et démontrent leur utilité pour la découverte des unités morphologiques et des processus de variation allomorphique.

Dans sa contribution intitulée « A framework for spatial clustering of textual objects: applications in topic clustering and text segmentation », Guillaume Guex (UNIL) présente un formalisme statistique dont dérivent deux méthodes pour le clustering spatial d'objets textuels. Ces méthodes, aux performances compétitives relativement à l'état de l'art, équilibrent similarité sémantique et proximité des objets dans un texte pour effectuer simultanément la segmentation textuelle et l'extraction de topics. Elles constituent ainsi des outils de lecture distante pour extraire de façon non supervisée des informations thématiques à partir d'un document unique.

La contribution de Jérôme Jacquin (UNIL), « La pragmatique de corpus comme lieu d’expérimentation des méthodes mixtes : pour une interdisciplinarité focalisée », explore les développements récents de la pragmatique de corpus avec un accent particulier sur les méthodes combinant perspectives quantitative et qualitative. À travers l’exemple de l’analyse de correspondances multiples et des tests d’accords inter-annotateurs, Jacquin montre comment la statistique et les méthodes computationnelles peuvent enrichir la recherche actuelle en pragmatique et plaide en faveur d’une collaboration plus étroite entre ces disciplines.

Loïc Jeanson (UNIL) retrace, dans sa contribution intitulée « Analyse spatiale, émergence en géographie, en archéologie et en histoire », l’évolution de l’analyse spatiale et son adoption graduelle dans plusieurs disciplines. Il explore la transformation de la géographie dès les années 1950 avec l’introduction des méthodes quantitatives, marquant un tournant entre géographie descriptive et analytique. Il montre ensuite comment l’archéologie puis l’histoire se sont progressivement emparées des outils d’analyse spatiale dans le cadre du *spatial turn*, pour conclure sur l’importance croissante de ces méthodes dans les sciences humaines et sociales et le rôle central du numérique dans ces évolutions.

La contribution de Mikhail Kanevski (UNIL), « On machine learning from environmental data », propose une revue exhaustive de l’application des algorithmes de machine learning à l’analyse, la modélisation et la visualisation des données environnementales. Il présente une méthodologie complète, allant de la collecte et l’exploration des données à la sélection de variables et l’entraînement des modèles. En conclusion, il insiste sur la nécessité d’une expertise combinée en science des données et en géosciences pour exploiter efficacement le machine learning dans l’étude des phénomènes environnementaux et la prise de décision.

Dans son étude intitulée « Entre toponymes et situation géographique : cartographie et autocorrélation spatiale des suffixes communaux suisses », Romain Loup (UNIL) combine analyse spatiale et analyse linguistique pour explorer la distribution géographique des suf-

fixes dans les noms de communes suisses. Les résultats mettent en évidence une forte concentration de certaines terminaisons toponymiques dans des zones géographiques spécifiques. Cette étude propose ainsi un nouvel éclairage sur la formation des identités territoriales en Suisse et contribue à une meilleure compréhension de la diversité linguistique et culturelle nationale à travers le prisme des noms de lieux.

La contribution de Cédric Margot (UNIL), « La vilaine TINA et le capitalisme autoritaire », revisite l'argument principal de sa thèse sur le présupposé de mesure dans le discours de l'information. Margot explique que ce présupposé invalide de nombreuses affirmations reposant sur des indicateurs statistiques, en particulier celles qui, en sciences économiques, prétendent qu'il n'y a pas d'alternative (*There is no alternative – TINA*). Il critique l'utilisation des statistiques comme outils de légitimation des politiques néolibérales, propose une réflexion épistémologique sur la dépolitisation de la quantification et appelle à une prise de conscience des fondements idéologiques qui sous-tendent les propos chiffrés dans les discours médiatiques et politiques.

Dans sa contribution intitulée « πάντα ρε : changements sémantiques dans la terminologie mathématique », Robin Meyer (UNIL) retrace l'origine et le développement des termes « mathématique » et « statistique », en particulier la façon dont leurs significations ont changé au fil du temps. Il montre que le premier, qui recouvrail autrefois des disciplines comme la musique et l'astronomie, a subi une restriction sémantique, tandis que le second a vu son sens s'étendre de la gestion de l'État à l'analyse de données. La contribution met ainsi en lumière l'importance de connaître l'histoire de la terminologie pour comprendre le développement de la pensée dans des disciplines scientifiques.

La contribution d'Isaac Pante (UNIL), « Le doigt et la lune », défend la légitimité de l'informatique en Faculté des lettres, retraçant son histoire institutionnelle et soulignant son potentiel comme science humaine à part entière plutôt que simple auxiliaire méthodologique. Pante appelle la discipline à se doter d'un corpus propre et désigne le jeu (vidéo), objet culturel mêlant narration, esthétique et mécanique computationnelle, comme candidat privilégié. Il illustre cette perspec-

tive par l'étude des « livres dont vous êtes le héros », dont l'analyse via la théorie des graphes permet de croiser regard littéraire et analyse computationnelle. L'informatique y apparaît comme un espace d'indiscipline féconde, entre littérature, jeu et création.

L'essai de Michael Piotrowski (UNIL), intitulé « Les humanités numériques – vers une “mathématique originale” ? », offre une réflexion épistémologique sur le rôle des mathématiques dans les sciences humaines, arguant qu'il ne se limite pas à la quantification des phénomènes mais permet l'invention de nouvelles structures. Il retrace l'évolution de la conception des mathématiques, de l'art de la mesure à la science des modèles, et plaide en faveur d'une meilleure intégration des méthodes informatiques avec leurs fondements logiques et mathématiques. Il insiste enfin sur l'importance de définir rigoureusement les humanités numériques pour tirer pleinement parti des mathématiques comme fondement méthodologique de la construction de modèles en sciences humaines.

La contribution de Loris Rimaz et Coline Métrailler (UNIL), « “May you find peace on your journey” : une approche comparative des mondes de FromSoftware », propose une analyse formelle des mondes des jeux vidéo de FromSoftware à travers leur représentation en graphes orientés. Cette approche permet de comparer leur structure spatiale, de mesurer leur linéarité ou interconnexion et d'évaluer les expériences spatiales des joueurs. Les résultats montrent une tendance à la linéarisation progressive des mondes, tout en mettant en lumière la distinction entre espace perçu, conçu et vécu. S'appuyant sur des mesures de graphes et des visualisations, l'étude ouvre des pistes pour des analyses plus poussées avec des données dynamiques ou télémétriques.

Dans sa contribution intitulée « Des mathématiques et des jeux. Le caractère ludique des graphes et des réseaux », Yannick Rochat (UNIL) explore les liens entre jeux et mathématiques. Il présente un éventail de jeux reposant explicitement ou implicitement sur des graphes – du Dobble au Monopoly, en passant par Catan, Pandemic ou Scotland Yard – en distinguant trois approches : les jeux issus de théorèmes mathématiques, ceux intégrant des structures relationnelles et ceux qui, sans que les mathématiques soient au centre de leurs mécaniques, se

prêtent à une modélisation mathématique. Le texte met en évidence la valeur pédagogique et créative des graphes, jusqu'à leur usage dans la littérature ludique, notamment chez Perec et l'Oulipo.

Enfin, la contribution de Céline Rozenblat et Mikhail Rogov (UNIL), « Les réseaux des travaux de François Bavaud dans tous leurs étaux ou comment aborder François Bavaud dans le style de “Cantatrix Sopranica” », propose une exploration de la bibliographie de François Bavaud bien plus détaillée que la synthèse rapide et partielle faite dans cet avant-propos. Les auteur·e·s appliquent des méthodes d'analyse textuelle et d'analyse de réseaux au corpus de ses publications, ainsi qu'à ceux des travaux qu'il cite ou qui le citent. Ils mettent en évidence le rôle central de trois articles (Bavaud, 1991, 1998b, 2011) qui illustrent ses contributions sur la mécanique statistique, les matrices spatiales pondérées et les transformations de Schoenberg. Ils soulignent en particulier la créativité mathématique de François, la rigueur de sa pensée, sa capacité à construire des ponts entre des domaines variés et, par là même, à avoir un impact sur la recherche dans ces domaines.

Au terme de ce bref aperçu des belles prises ramenées dans nos fils, il nous reste à espérer qu'elles sauront susciter l'intérêt des lecteur·rice·s, et tout particulièrement celui de François Bavaud. Ce périple intellectuel à travers la modélisation de l'espace et du langage, l'analyse des réseaux et la réflexion sur la science lui est présenté comme un témoignage de notre reconnaissance pour sa recherche et son enseignement, qui constituent pour nous une source intarissable d'inspiration. Puisse cet hommage refléter l'empreinte indélébile qu'il laisse sur nos trajectoires académiques et personnelles.

Références

- Bavaud, F. (1991). Equilibrium properties of the Vlasov functional: the generalized Poisson-Boltzmann-Emden equation. *Reviews of Modern Physics*, 63(1):129–149.
- Bavaud, F. (1998a). *Modèles et données : une introduction à la Statistique uni-, bi- et trivariée*. L'Harmattan, Paris.
- Bavaud, F. (1998b). Models for spatial weights: a systematic look. *Geographical Analysis*, 30(2):153–171.

- Bavaud, F. (2005). Using local formalism in quantitative geography: a straightforward method for taking spatial auto-correlation into account. In *14th European Colloquium on Theoretical and Quantitative Geography*.
- Bavaud, F. (2006). Spectral clustering and multidimensional scaling: a unified view. In Batagelj, V., Bock, H.-H., Ferligoj, A., & Ziberna, A. (éd.), *Data science and classification*, pages 131–139. Springer, Heidelberg.
- Bavaud, F. (2008a). The endogenous analysis of flows, with applications to migrations, social mobility and opinion shifts. *Journal of Mathematical Sociology*, 32:239–266.
- Bavaud, F. (2008b). Local concentrations. *Papers in Regional Science*, 87(3):357–370.
- Bavaud, F. (2010). Euclidean distances, soft and spectral clustering on weighted graphs. In Balcázar, J. L., Bonchi, F., Gionis, A., & Sebag, M. (éd.), *Machine Learning and Knowledge Discovery in Databases*, pages 103–118, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Bavaud, F. (2011). On the Schoenberg transformations in data analysis: Theory and illustrations. *Journal of Classification*, 28:297–314.
- Bavaud, F. (2013). Testing spatial autocorrelation in weighted networks: the modes permutation test. *Journal of Geographical Systems*, 15(3):233–247.
- Bavaud, F. (2018). Combien y a-t-il de variétés distinctes d’anglais? *Cahiers du Centre de Linguistique et des Sciences du Langage*, 56:11–30.
- Bavaud, F. (2024). Measuring and testing multivariate spatial autocorrelation in a weighted setting: A kernel approach. *Geographical Analysis*, 56(3):573–599.
- Bavaud, F., Cocco, C., & Xanthos, A. (2012). Textual autocorrelation: formalism and illustrations. In *Proceedings of JADT, 11èmes Journées internationales d’analyse statistique des données textuelles*, page 109–120. Université de Liège.
- Bavaud, F., Cocco, C., & Xanthos, A. (2015). Textual navigation and autocorrelation. In Mikros, G. K. & Macutek, J. (éd.), *Sequences in Language and Text*, pages 35–56. De Gruyter Mouton, Berlin, München, Boston.
- Bavaud, F. & Guex, G. (2012). Interpolating between random walks and shortest paths: A path functional approach. In Aberer, K., Flache, A., Jager, W., Liu, L., Tang, J., & Guéret, C. (éd.), *Proceedings of the 4th International Conference on Social Informatics (SocInfo ’12)*, volume 7710 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.
- Bavaud, F., Kordi, M., & Kaiser, C. (2018). Flow autocorrelation: a dyadic approach. *The Annals of Regional Science*, 61(1):95–111.
- Bavaud, F. & Métrailler, C. (2023). A (dis)similarity index for comparing two character networks based on the same story. In Rochat, Y., Métrailler,

- C., & Piotrowski, M. (éd.), *Proceedings of the Workshop on Computational Methods in the Humanities 2022 (COMHUM 2022)*, pages 33–42. CEUR Workshop Proceedings.
- Bavaud, F., Picca, D., & Curdy, B. (2006). Non-linear correspondence analysis in text retrieval: a kernel view. In *Proceedings of JADT, 8èmes Journées internationales d'Analyse statistique des Données Textuelles*, pages 741–747.
- Bavaud, F. & Xanthos, A. (2002). Thermodynamique et statistique textuelle : concepts et illustrations. In *Proceedings of JADT, 6èmes Journées internationales d'Analyse statistique des Données Textuelles*, pages 101–111.
- Bavaud, F. & Xanthos, A. (2005). Markov associativities. *Journal of Quantitative Linguistics*, 12(2-3):123–137.
- Ceré, R. & Bavaud, F. (2017). Multi-labelled image segmentation in irregular, weighted networks: A spatial autocorrelation approach. In *Proceedings of the 3rd International Conference on Geographical Information Systems Theory, Applications and Management*. SCITEPRESS - Science and Technology Publications.
- Ceré, R. & Bavaud, F. (2019). Soft image segmentation: On the clustering of irregular, weighted, multivariate marked networks. In Ragia, L., Laurini, R., & Rocha, J. G. (éd.), *Geographical Information Systems Theory, Applications and Management*, pages 85–109, Cham. Springer International Publishing.
- Ceré, R., Egloff, M., & Bavaud, F. (2018). Geographical exploration and analysis extended to textual content. In Winter, S., Griffin, A., & Sester, M. (éd.), *10th International Conference on Geographic Information Science (GIScience 2018)*, volume 114 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:7, Dagstuhl, Germany. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- Cocco, C., Pittier, R., Bavaud, F., & Xanthos, A. (2011). Segmentation and clustering of textual sequences : a typological approach. In Angelova, G., Bontcheva, K., Mitkov, R., & Nicolov, N. (éd.), *Recent Advances in Natural Language Processing, RANLP 2011, 12-14 September, 2011, Hissar, Bulgaria*, pages 427–433. RANLP 2011 Organising Committee.
- Egloff, M. & Bavaud, F. (2018). Taking into account semantic similarities in correspondence analysis. In *Proceedings of the Workshop on Computational Methods in the Humanities 2018 (COMHUM 2018)*, volume 2314 of *CEUR Workshop Proceedings*, pages 45–51.
- Guex, G. & Bavaud, F. (2015). Flow-based dissimilarities: Shortest path, commute time, max-flow and free energy. In Lausen, B., Krolak-Schwerdt, S., & Böhmer, M. (éd.), *Data Science, Learning by Latent Structures, and Knowledge Discovery*, pages 101–111, Berlin, Heidelberg. Springer Berlin Heidelberg.

Guex, G., Loup, R., & Bavaud, F. (2023). Estimation of flow trajectories in a multi-lines transportation network. *Applied Network Science*, 8(44).

Loup, R. & Bavaud, F. (2025). Spatial autocorrelation of political opinions: A kernel approach. *Journal of Geographical Systems*.