

Sujets de thèse proposés par le Laboratoire MASAFEQ

2026 - 2027

Sujets proposés pour la formation Statistique –Actuariat

Sujet 1

Estimation de l'économie informelle au Maroc : approche hybride combinant les modèles d'équations structurelles et l'apprentissage automatique

Sujet proposé :

Pr. Ahmed OUAAZA

Présentation :

L'économie informelle constitue un enjeu majeur pour l'économie marocaine, en raison de ses implications sur la productivité, les recettes fiscales et le financement de la protection sociale. Toutefois, son ampleur demeure difficile à mesurer, les méthodes d'estimation existantes pouvant conduire à des résultats divergents. Cette thèse vise à développer une approche hybride combinant les modèles d'équations structurelles (SEM) et les techniques d'apprentissage automatique (Machine Learning). L'objectif est d'identifier les principaux déterminants de l'informalité, d'exploiter différents indicateurs observables et d'améliorer la précision et la fiabilité des estimations en tenant compte des relations non linéaires et des effets de seuil. Les performances de l'approche proposée seront évaluées et comparées à celles des méthodes traditionnelles. Les résultats attendus contribueront à une meilleure compréhension de la dynamique de l'économie informelle au Maroc et à l'élaboration de recommandations susceptibles d'éclairer les politiques publiques de formalisation

Sujet 2

Analyse fonctionnelle et modélisation statistique des trajectoires de retrait des dépôts non échéancés : application à l'ALM bancaire

Sujet proposé par :

Pr. Ahmed OUAAZA

Présentation :

Les dépôts non échéancés constituent un enjeu majeur de la gestion actif-passif (ALM), car leur maturité contractuelle ne reflète pas nécessairement leur comportement réel. La problématique de cette thèse consiste à mieux modéliser et prévoir les trajectoires de conservation et de retrait de ces dépôts, en exploitant l'information contenue dans leur évolution temporelle et leur hétérogénéité.

L'objectif est de développer une approche statistique fondée sur l'analyse de données fonctionnelles, complétée par des méthodes d'apprentissage statistique, afin d'identifier les principaux profils comportementaux, d'étudier leurs relations avec les taux et les conditions économiques, et d'améliorer la prévision des sorties de dépôts.

La contribution attendue est à la fois **méthodologique**, à travers le développement et l'évaluation de modèles adaptés à la nature fonctionnelle, non linéaire et potentiellement hétérogène des trajectoires, et **opérationnelle**, à travers leur traduction en échéanciers comportementaux utilisables en ALM. Une attention particulière sera accordée à l'évaluation de l'incertitude des prévisions et à la fiabilité des modèles face aux changements de régime et aux situations de stress.

Sujet 3

Contribution aux méthodes d'estimation non paramétriques intégrant des algorithmes de machine learning pour les modèles d'équations structurelles

Sujet proposé par :

Pr. Ahmed OUAAZA

Présentation :

Les modèles d'équations structurelles (SEMs) constituent une famille de méthodes statistiques permettant d'analyser les relations de causalité entre plusieurs variables observées et non observées (latentes). Ils sont largement utilisés en sciences sociales, en économie, en gestion et en psychologie pour l'étude de concepts théoriques inobservables tels que la satisfaction, la confiance, l'engagement, la motivation....

L'estimation des SEMs constitue un enjeu majeur. Dans la littérature, plusieurs méthodes paramétriques, semi-paramétriques et non paramétriques ont été proposées afin d'estimer les paramètres du modèle et les liens entre les variables.

Les méthodes non paramétriques, telles que les estimateurs à noyau ou les splines, ainsi que certaines approches de machine learning, permettent de dépasser plusieurs hypothèses restrictives des méthodes paramétriques classiques, notamment la normalité des distributions, la linéarité des relations et l'indépendance des erreurs.

Parmi les objectifs des travaux qui doivent être réalisés dans le cadre de cette thèse :

- ✓ Etudier de manière approfondie les méthodes d'estimation non paramétriques appliquées aux modèles d'équations structurelles.
 - ✓ Explorer l'apport des algorithmes de machine learning dans l'amélioration des méthodes d'estimation des SEMs.
 - ✓ Proposer de nouvelles approches hybrides combinant non paramétrique et machine learning pour l'estimation des SEMs.
-

Sujet 4

Efficient and Reliable Instruction-Following Rerankers

Sujet proposé par :

Pr. Hicham JANATI

Context and research question:

Modern search engines and retrieval-augmented generation (RAG) systems rely on neural rerankers: models that take a user query and a set of candidate documents returned by a first-stage retriever, and reorder them by relevance. Increasingly, relevance is not fixed but specified by a natural-language instruction (e.g. “prioritize evidence of productivity improvements” versus “. . . productivity declines”) that changes which documents should be ranked first [1, 2, 3]. Instruction-aware rerankers based on language models now perform well on such tasks [5, 6], but they are expensive. Deploying them in practice requires compression: smaller distilled students [7, 8, 9] or low-precision quantized weights [10, 11]. Recent work shows that quantization can silently damage rankings [12], but it remains unclear whether compression specifically destroys the model’s ability to respond correctly to a change of instruction, beyond ordinary small-margin ranking fragility. The central question of this thesis is: when relevance is defined by a user instruction, how can a compact or low-precision reranker retain the correct response to changes in that instruction, with a defensible quality–cost trade-off? The project will proceed in three connected stages:

(i) Diagnosis: build a controlled, paired evaluation protocol (criterion-changing versus meaning preserving instruction rewrites) that separates errors newly caused by compression from pre-existing model weaknesses and from generic ranking instability, building on existing benchmarks [1, 3, 4].

(ii) Learning: design budget-aware supervision that selects the most informative instruction-document interventions to transfer controllability from a strong teacher to a compact student, compared against standard ranking distillation and counterfactual training [7, 9, 13, 14].

(iii) Efficient deployment: develop compression schemes that protect the components responsible for instruction control (sensitivity-based precision allocation, small control-specific residual modules), and optionally statistically calibrated policies that detect risky inputs and escalate them to a stronger model [12, 15]. The thesis combines statistical learning (experimental design, sensitivity analysis, risk calibration) with modern deep learning engineering. Experiments will use open compact rerankers (0.6B–4B parameters) on a single high-end GPU, with careful measurement of memory, latency and ranking quality. Expected outputs include a reproducible evaluation protocol, new training and compression methods for controllable ranking,

open-source code, and publications in machine learning and information retrieval venues.

Candidate profile

- Engineering degree (Diplôme d'Ingénieur) or Master 2 in applied mathematics, statistics, statistical learning / machine learning, data science or artificial intelligence.
- Solid foundations in probability, statistics, optimization and linear algebra.
- Strong programming skills in Python and hands-on experience with a deep learning framework (PyTorch preferred); familiarity with the Hugging Face ecosystem, Transformers and GPU computing is a strong asset.
- Interest in natural language processing, information retrieval or large language models; prior experience with model fine-tuning, distillation or quantization is a plus.
- Rigor in experimental methodology, autonomy, curiosity, and good scientific writing in English (French and/or Arabic is a plus).

Contact

Pr. Hicham Janati

(hjanati@insea.ac.ma), INSEA – MASAFEQ Laboratory.

Key references

- [1] O. Weller et al. FollowIR: Evaluating and teaching information retrieval models to follow instructions. *NAACL*, 2025.
- [2] W. Sun et al. MAIR: A massive benchmark for evaluating instructed retrieval. *EMNLP*, 2024.
- [3] Lu et al. MultiConIR: Towards multi-condition information retrieval. *Findings of EMNLP*, 2025.
- [4] O. Weller, D. Lawrie, B. Van Durme. NevIR: Negation in neural information retrieval. *EACL*, 2024.
- [5] Y. Zhang et al. Qwen3 Embedding: Advancing text embedding and reranking through foundation models. *arXiv:2506.05176*, 2025.
- [6] O. Weller et al. Rank1: Test-time compute for reranking in information retrieval. *arXiv:2502.18418*, 2025.
- [7] S. Hofstätter et al. Improving efficient neural ranking models with cross-architecture knowledge distillation. *arXiv:2010.02666*, 2020.
- [8] Xu et al. Distillation versus contrastive learning: How to train your rerankers. *Findings of IJCNLP-AAACL*, 2025.
- [9] Prabhakar et al. On-policy distillation meets off-policy GRPO: Training compact instruction-following rerankers. *arXiv:2609.01947*, 2026.
- [10] E. Frantar et al. GPTQ: Accurate post-training quantization for generative pre-trained transformers. *ICLR*, 2023.
- [11] J. Lin et al. AWQ: Activation-aware weight quantization for on-device LLM compression and acceleration. *MLSys*, 2024.
- [12] Zhou et al. The undetected damage of quantization on retrieval and how to fix it. *arXiv:2609.24322*, 2026.
- [13] Zheng et al. Counterfactual constraint-conditioned on-policy distillation for multi-constraint instruction following. *arXiv:2609.27421*, 2026.
- [14] S. Srinivas, F. Fleuret. Knowledge transfer with Jacobian matching. *ICML*, 2018.
- [15] A. N. Angelopoulos et al. Conformal risk control. *ICLR*, 2024.

Sujets proposés pour la formation Economie Quantitative - Finance

Sujet 1

Femme et transformation économique au Maroc, une analyse par un MEGC genré

Sujet proposé par :
Pr. Driss EFFINA.

Problématique. Quel serait l'effet d'une plus forte participation des femmes à l'activité économique sur la croissance et la transformation structurelle du Maroc ?

Objectifs

- Construire une matrice de comptabilité sociale (MCS) désagrégée par sexe.
- Simuler des politiques ciblées : hausse de l'emploi féminin, formation, accès au crédit, services de garde.
- Mesurer les effets sur le PIB, la productivité, les revenus des ménages et la structure sectorielle.

Méthodologie. Modèle d'Équilibre Général Calculable (MEGC) intégrant le travail rémunéré et non rémunéré, avec un marché du travail segmenté par sexe.

Données. Enquête nationale sur l'emploi et enquête nationale sur l'emploi du temps (HCP), comptes nationaux, tableau entrées-sorties.

Sujet 2

Déficit commercial structurel et stabilité de l'économie marocaine sur le long terme

Sujet proposé par :
Pr. Driss EFFINA.

Problématique. Le déficit commercial persistant du Maroc menace-t-il la stabilité macroéconomique à long terme, ou reflète-t-il une phase d'investissement et de montée en gamme productive ?

Objectifs

- Analyser la soutenabilité du déficit et ses liens avec la dette extérieure et les réserves de change.
- Évaluer le rôle des transferts des MRE, du tourisme et des IDE dans son financement.
- Identifier les seuils au-delà desquels le déficit pèse sur la croissance et la stabilité des prix.

Méthodologie. Modèles ARDL et VECM, tests de cointégration, analyse des chocs par fonctions de réponse impulsionnelle.

Données. Office des changes, Bank Al-Maghrib, HCP, bases du FMI et de la Banque mondiale.

Sujet 3

Fondements microéconomiques de la stabilisation macroéconomique au Maroc, un modèle DSGE néo-keynésien pour le ciblage d'inflation

Sujet proposé par :
Pr. Driss EFFINA.

Lien avec le prix Nobel 2026. Le prix Nobel d'économie 2026 a été attribué le 12 octobre 2026 à Robert E. Hall (Stanford) et Michael Woodford (Columbia) pour leurs travaux sur les fondements microéconomiques de la stabilisation macroéconomique ([Al Jazeera](#), [Britannica](#)). Leurs modèles servent aujourd'hui aux banques centrales et aux gouvernements pour atténuer les fluctuations économiques.

Problématique. Comment adapter ces modèles aux spécificités marocaines pour éclairer la transition de Bank Al-Maghrib vers le ciblage d'inflation et un régime de change plus flexible ?

Objectifs

- Construire un modèle DSGE néo-keynésien calibré et estimé sur données marocaines.
- Mesurer l'efficacité de la communication de la banque centrale (forward guidance), idée centrale des travaux de Woodford.
- Intégrer les rigidités du marché du travail et le secteur informel, en lien avec les travaux de Hall.
- Étudier la coordination entre politique monétaire et politique budgétaire face aux chocs.

Méthodologie. Modèle DSGE estimé par méthodes bayésiennes (Dynare), simulations de règles de politique monétaire.

Données. Bank Al-Maghrib, HCP, ministère de l'Économie et des Finances.

Sujet 3

Chocs climatiques, stress hydrique et stabilité macroéconomique au Maroc

Sujet proposé par :
Pr. Driss EFFINA.

Problématique. Comment les sécheresses récurrentes et le stress hydrique affectent-ils la croissance, l'inflation et les finances publiques du Maroc ?

Objectifs

- Quantifier l'effet des chocs pluviométriques sur la valeur ajoutée agricole, l'emploi rural et l'inflation alimentaire.
- Mesurer leur transmission au reste de l'économie et aux finances publiques.
- Évaluer les politiques d'adaptation : dessalement, irrigation, assurance agricole.

Méthodologie. Modèles VAR et économétrie de panel régional, avec un module climatique pouvant s'intégrer au modèle DSGE du thème 3.

Données. Direction de la Météorologie Nationale, ministère de l'Agriculture, HCP, Bank Al-Maghrib.

Ce thème complète le thème 3 : les sécheresses sont des chocs d'offre majeurs que la politique monétaire doit intégrer.

Sujets proposés pour la formation Mathématiques Appliquées

Sujet 1

Optimisation multi-objectif pour la planification énergétique urbaine au Maroc.

Sujet proposé par :

Pr. Mohamed OUZINEB, INSEA.

Possibilité de financement ou de soutien financier par des partenaires industriels.

Résumé du sujet :

L'objectif de cette thèse est de développer une méthodologie d'optimisation multi-objectif pour la gestion efficace de l'énergie dans des systèmes urbains marocains, en tenant compte à la fois de la fiabilité, des contraintes économiques et environnementales, et de la disponibilité des ressources énergétiques. Les algorithmes proposés viseront une adéquation optimale entre la disponibilité des sources (renouvelables et conventionnelles) et la consommation énergétique, tout en minimisant les coûts totaux de production. Ces algorithmes s'appuieront sur des prévisions de la demande énergétique et la disponibilité des différentes technologies, avec une capacité à modéliser différentes configurations du système énergétique. L'intégration d'une technique efficace d'estimation de la fiabilité globale, telle que le facteur de capacité, permettra d'identifier les configurations les plus performantes au regard des critères de coût, de fiabilité et de durabilité.

Afin de garantir la pertinence et l'applicabilité des résultats, l'étude sera illustrée par des cas concrets issus de firmes situées dans des régions urbaines marocaines, pour lesquelles des données de conception et d'exploitation sont disponibles.

Axes de recherche :

1. **Modélisation du système énergétique hybride** (solaire, éolien, cycle combiné, etc.) adapté au contexte marocain.
2. **Définition d'un indicateur global de performance énergétique** basé sur le **facteur de capacité**.
3. **Optimisation multi-objectif** (fiabilité vs coût vs impact environnemental).
4. **Comparaison de différents scénarios de développement énergétique** (100% renouvelable, mix énergétique, etc.).
5. **Cas d'étude réel** : Application à un centre urbain marocain (Casablanca, Marrakech, Rabat, etc.).

Outils et Méthodes :

- Programmation mathématique / linéaire / non linéaire
- Métaheuristiques pour l'optimisation
- Simulation de profils de charge urbaine (courbes de charge cumulées)

- Analyse de sensibilité et évaluation des incertitudes.

Mots-clés :

Mix énergétique, Algorithmes d'optimisation, Planification énergétique, Modélisation mathématique, Heuristique, Machine learning.

Sujet 2

Analyse et optimisation de la fiabilité des systèmes miniers pour la réduction du coût global de production

Sujet proposé par :

Pr. Mohamed OUZINEB, INSEA.

Possibilité de financement ou de soutien financier par des partenaires industriels.

Résumé du sujet :

L'objectif de cette thèse est de développer une méthodologie intégrée pour analyser et optimiser la fiabilité des systèmes de production dans les opérations minières, en tenant compte de la performance individuelle de chaque composante (machines, convoyeurs, équipements de traitement, etc.). L'approche proposée vise à minimiser le coût total de production en équilibrant les coûts d'investissement, d'entretien, et les pertes de production dues aux défaillances.

L'étude portera sur la modélisation du système de production en chaîne (système en série, en parallèle, ou mixte) et utilisera des outils d'analyse probabiliste et d'optimisation pour évaluer les compromis entre fiabilité et coûts.

Axes de recherche :

1. **Modélisation des systèmes de production miniers** selon leur structure (série, parallèle, hybride).
2. **Analyse de fiabilité** des composantes (MTBF, taux de défaillance, réparabilité, disponibilité).
 - **Estimation des coûts** : coût d'investissement initial, coût de maintenance et d'entretien, Coût des arrêts de production
3. **Méthode d'évaluation globale** (fonction objectif) intégrant fiabilité et coût.
4. **Optimisation globale du système** pour atteindre un compromis optimal entre fiabilité et coût total.
5. **Étude de cas** sur un site minier réel.

Outils et Méthodes :

- Programmation mathématique / linéaire / non linéaire
- Calcul de la fiabilité des systèmes de production
- Outils logiciels : MATLAB, Python, R, ou logiciels spécialisés
- Analyse de sensibilité et évaluation des incertitudes

Mots-clés : Fiabilité, Algorithmes d'optimisation, Modélisation mathématique, Heuristique, Machine learning.

Sujet 3

Data-Driven Logistics for Sustainable Development: Environmental Challenges and Econometric Approaches

Logistique pilotée par les données pour le développement durable : enjeux environnementaux et approches économétriques

Sujet proposé par :

Pr. Mohamed OUZINEB, INSEA.

Pr. Ahmed El Ghini, UM5.

1. Contexte

La logistique et le transport sont au cœur de l'activité économique : ils relient les sites de production, les ports et les marchés. Leur croissance rapide s'accompagne cependant d'une forte consommation d'énergie, d'émissions de gaz à effet de serre, de congestion et d'autres pressions environnementales. Pour des pays en pleine expansion logistique comme le Maroc (ports, corridors ferroviaires et routiers, chaînes d'exportation industrielles), concilier performance économique et durabilité devient un enjeu stratégique. Les décideurs ont besoin de preuves chiffrées et d'outils d'aide à la décision fiables, plutôt que d'intuitions.

2. Problématique

Comment mesurer rigoureusement l'impact environnemental des activités logistiques et de transport, et comment exploiter ces mesures pour concevoir des systèmes logistiques à la fois économiquement efficaces et moins émetteurs ?

3. Objectifs

- **Mesurer** : quantifier, par des modèles économétriques, la relation entre activité logistique (volumes transportés, modes de transport, infrastructures) et émissions ou consommation d'énergie.
- **Évaluer** : estimer l'effet réel de politiques publiques ou d'investissements (report modal vers le rail, normes d'émissions, tarification carbone) sur la performance environnementale du transport.
- **Optimiser** : intégrer ces résultats dans des modèles d'optimisation multi-objectifs qui arbitrent entre coût, niveau de service et émissions.
- **Recommander** : produire des recommandations concrètes et transférables pour les décideurs publics et les industriels.

4. Positionnement scientifique

La littérature traite souvent séparément l'analyse économétrique des émissions du transport et l'optimisation verte des réseaux logistiques. L'originalité visée est de relier les deux : utiliser des estimations empiriques robustes comme paramètres d'entrée de modèles d'optimisation, afin que les décisions proposées reposent sur des relations

mesurées et non sur des hypothèses théoriques. L'apport de l'intelligence artificielle sera exploré pour la prévision des flux et des émissions.

5. Méthodologie envisagée

- **Économétrie** : modèles en données de panel (données suivies sur plusieurs pays ou régions et plusieurs années), méthodes d'évaluation d'impact de politiques (différence de différences, contrôle synthétique).
- **Optimisation** : programmation linéaire en nombres entiers et optimisation multi-objectifs (coût et CO₂) appliquées à un réseau de transport ou une chaîne logistique réelle.
- **Machine learning** : prévision de la demande de transport et des émissions associées, en complément des modèles économétriques.
- **Données** : bases internationales (Banque mondiale, Agence internationale de l'énergie, OCDE), statistiques nationales du transport, et si possible données opérationnelles d'un partenaire industriel ou logistique.

6. Résultats attendus

- Des estimations fiables des facteurs qui déterminent les émissions logistiques dans le contexte étudié.
 - Un cadre méthodologique combinant économétrie et optimisation, réutilisable pour d'autres pays ou secteurs.
 - Des scénarios chiffrés et des recommandations pour une logistique plus verte, efficace et résiliente.
 - Des publications dans des revues scientifiques, dont une contribution possible au numéro spécial associé à ce thème.
-

Sujet 4

Apport de l'optimisation multiobjectif aux modèles d'apprentissage automatique; application aux machines à vecteurs de support (SVM).

Sujet proposé par :

Pr. Mohamed OUZINEB, INSEA.

Pr. Fatima zahra RAHOU, INSEA.

Résumé du sujet :

Cette thèse propose d'étudier l'intégration des méthodes d'optimisation multiobjectif (MOO) dans l'entraînement et la configuration des modèles d'apprentissage automatique, en se concentrant sur les machines à vecteurs de support (SVM) pour des problèmes de classification.

Axes de recherche / Plan de travail :

- État de l'art
 - Optimisation multiobjectif : théorie, méthodes classiques.

- Apprentissage automatique et SVM : principe, formulation mathématique, hyperparamètres critiques.
- Travaux existants sur la combinaison MOO + SVM, lacunes et opportunités.
- Formulation du problème multiobjectif
 - Identification des critères à optimiser simultanément (précision, complexité, robustesse, etc.).
 - Modélisation mathématique et conditions d'existence d'optima.
- Développement des méthodes
 - Mise en place d'approches MOO pour SVM (algorithmes évolutionnaires, métaheuristiques hybrides).
- Applications et expérimentations
- Perspectives
 - Extension à d'autres modèles de ML (réseaux de neurones, arbres de décision).
 - Applications à des problèmes réels de grande dimension et temps réel.

Méthodes et outils :

- Méthodes : Optimisation multiobjectif, SVM supervisé, sélection de variables.
- Outils informatiques : Python (scikit-learn, DEAP, pymoo), MATLAB, R (analyses statistiques et visualisation Pareto).
- Données : jeux de données publics (UCI, Kaggle), éventuellement données locales pour applications spécifiques.

Mots-clés :

- Optimisation multiobjectif (MOO)
 - Machine Learning / Apprentissage automatique
 - Support Vector Machines (SVM)
 - Classification supervisée
 - Sélection de variables / hyperparamètres
 - Algorithmes évolutionnaires
 - Frontière de Pareto
 - Applications réelles (santé, finance, énergie)
-

Sujet 5

Digital Twin intelligent pour l'Industrie 5.0 basé sur la vision par ordinateur, l'IoT et l'IA générative

Sujet proposé :

Pr. Mohamed OUZINEB, INSEA.

Pr. Faycal MIMOUNI, INSEA

Contexte et problématique

L'Industrie 5.0 place l'humain, la durabilité et la résilience au cœur des systèmes de production [1]. Le jumeau numérique, représentation virtuelle actualisée d'un équipement ou d'un procédé, constitue un support pour observer, simuler et anticiper son fonctionnement [2]. Son exploitation reste toutefois confrontée à l'hétérogénéité des données, aux défauts de synchronisation et à la difficulté de produire des recommandations compréhensibles pour les opérateurs.

Cette thèse propose de concevoir et d'évaluer un Digital Twin intelligent associant vision par ordinateur, Internet des objets (IoT) et IA générative. La question centrale est la suivante : comment fusionner les observations visuelles et les mesures des capteurs pour maintenir une représentation fiable du système industriel, puis fournir une assistance contextualisée et vérifiable à la décision ? Les verrous scientifiques concernent l'alignement temporel des modalités, la robustesse aux données manquantes et aux changements de conditions, ainsi que la maîtrise des réponses non fondées de l'IA générative.

La démarche portera sur une cellule de production pilote, instrumentée par des caméras et des capteurs de vibration, de température ou de consommation électrique. Des modèles de détection d'anomalies et de fusion multimodale alimenteront le jumeau numérique. Un assistant génératif, enrichi par recherche documentaire (RAG) [3], exploitera l'état du jumeau et les documents techniques pour expliquer les alertes et proposer des interventions soumises à validation humaine. Des scénarios seront simulés avant toute action sur l'équipement. Les données synthétiques pourront être étudiées pour les défauts rares, avec une validation distincte sur des observations réelles.

L'évaluation comparera les approches visuelles seules, IoT seules et fusionnées, ainsi que l'assistance avec et sans RAG. Elle mesurera la précision des alertes, les faux positifs, la latence de mise à jour, la fidélité des réponses aux sources et le temps de diagnostic. Des séparations par période et par équipement limiteront les fuites de données. Les essais analyseront aussi les pertes de capteurs, la dérive des données, l'utilité perçue par les opérateurs et les effets sur les rebuts et l'énergie par pièce. Les contributions attendues sont une méthode de fusion robuste, une assistance traçable et un prototype reproductible ; les gains restent à démontrer expérimentalement.

Mots-clés

Jumeau numérique intelligent ; Industrie 5.0 ; Vision par ordinateur ; IoT industriel ; Fusion multimodale ; IA générative ; Aide à la décision ; Résilience.

Profil recherché

Master à orientation recherche ou diplôme d'ingénieur en informatique, intelligence artificielle, automatique ou génie industriel. Bonnes bases en apprentissage profond, traitement d'images, séries temporelles et programmation Python. Un intérêt pour les systèmes embarqués et l'interaction humain-machine est souhaité.

Références indicatives

1. Commission européenne (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. DOI : 10.2777/308407.

2. Tao, F., Zhang, H., Liu, A. et Nee, A. Y. C. (2019). Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics, 15(4), 2405-2415. DOI : 10.1109/TII.2018.2873186.
 3. Lewis, P. et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS, 33, 9459-9474. arXiv:2005.11401.
-

Sujet 6

Application de l'Industrie 4.0 pour l'amélioration de la qualité et de la sécurité dans l'industrie aéronautique

Sujet proposé :

Pr. Mohamed OUZINEB, INSEA.

Pr. Faycal MIMOUNI, INSEA

Contexte et problématique

L'industrie aéronautique est caractérisée par des exigences très strictes en matière de qualité, de fiabilité et de sécurité, notamment dans les activités de maintenance, réparation et révision (MRO). Les processus sont complexes et fortement réglementés, et les organisations doivent répondre à des contraintes simultanées de performance, de conformité et de sécurité.

L'Industrie 4.0, définie par l'intégration de technologies numériques avancées telles que l'Internet des Objets (IoT), l'intelligence artificielle (IA), la robotique collaborative, la réalité augmentée et la gestion de données massives, offre un potentiel pour transformer la manière dont la qualité et la sécurité sont gérées.

Cependant, l'adoption de ces technologies pose plusieurs questions :

1. Quels outils numériques peuvent réellement renforcer la qualité et la sécurité ?
2. Comment concilier exigences réglementaires et innovations technologiques ?
3. Quelles sont les limites liées à l'organisation, à la formation du personnel ou à la cybersécurité ?
4. Comment l'intégration des technologies de l'Industrie 4.0 peut-elle contribuer à l'amélioration de la qualité et de la sécurité dans l'industrie aéronautique, et quelles conditions facilitent ou limitent cette contribution ?

Objectifs de la recherche

Objectif principal :

Analyser et évaluer l'impact de l'Industrie 4.0 sur la qualité et la sécurité dans l'industrie aéronautique.

Objectifs spécifiques :

1. Identifier les technologies 4.0 les plus pertinentes pour la gestion de la qualité et de la sécurité dans un contexte aéronautique.
2. Étudier les processus existants et leurs points critiques, afin de déterminer où l'innovation numérique peut apporter des améliorations.
3. Proposer un cadre méthodologique pour l'intégration des technologies numériques dans un environnement réglementé.
4. Évaluer les effets de l'implémentation de l'Industrie 4.0 sur des indicateurs de qualité et de sécurité.

Revue de littérature :

1. Concepts et principes de l'Industrie 4.0.
2. Cadres de gestion de la qualité et de la sécurité dans l'industrie aéronautique.
3. Études sur l'impact de la digitalisation sur les performances organisationnelles et la prévention des risques.

Bibliographie préliminaire

Kagermann, H., Wahlster, W., & Helbig, J. (2013). Recommendations for implementing the strategic initiative Industrie 4.0.

Lee, J., Bagheri, B., & Kao, H.A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3.

Romero, D., & Stahre, J. (2021). Towards the resilient operator 5.0: The role of digital technologies in resilient organizations. *Computers & Industrial Engineering*.

International Civil Aviation Organization (ICAO) (2020). *Safety Management Manual* (Doc 9859).

European Union Aviation Safety Agency (EASA) (2023). *Guidance on Part-145 and Safety Management Systems*.

Sujet 7

Statistical Analysis and Optimization of High-Risk Road Sections in Morocco

Sujet proposé :

Pr. Mohamed OUZINEB, INSEA.

Pr. Abdelbari Redouane, ENSMR.

PROBLEM STATEMENT

Improving road safety is a major global concern since road crashes cause human and economic costs. The WHO (n.d.) estimated that each year, approximately 1.16 million people die from road traffic crashes. Moreover, road traffic injuries are the leading cause of death among children and young adults aged 5-29 years. In addition, the WHO estimates that road crashes cost most countries 3% of their GDP (n.d.). Another study found that the mean share of GDP of the cost of road crashes differs between High-Income Countries (2.9%) and Low- and Middle-Income Countries (2.2%), with this share reaching 6% for the United States and 4.6% for Belgium (Wijnen and Stipdonk 2016).

Morocco is no exception in terms of the impact of road crashes at the human and economic levels. Indeed, the National Agency of Road Safety (NARSA n.d.-a), which is the administration in charge of road safety in Morocco, estimated that the number of deaths caused by road accidents was 3,641 in 2024, with 9,334 people severely injured and a total of 142,247 casualty accidents¹. Regarding the economic cost of road crashes in Morocco, the ITF (2024) estimated the cost of road crashes at 1.6 billion euros in 2022, representing 1.2% of the country's GDP.

This has prompted policymakers and road safety officials to take measures to improve road safety and reduce the human and economic costs of road crashes. For this purpose, the Moroccan government has implemented the first national road safety strategy (2004–2013), followed by the second national road safety strategy (2017–2026) (NARSA 2022), and is currently developing the third national strategy for road safety (2026–2030), with more than 8 billion dirhams allocated to implementing this strategy, equivalent to approximately US\$2.2 billion (MAP 2026). This substantial amount of public resources dedicated to addressing the issue of road crashes indicates a strong political commitment and provides operational actors with the necessary resources to improve road safety.

However, one of the main issues facing operational actors is optimizing resources and prioritizing actions that are likely to have the biggest positive impact on road safety, given that the allocated resources, though considerable given the limited public resources, remain insufficient. This thesis aims to contribute to prioritizing public actions by providing scientific profiling of the road safety system in Morocco. For this purpose, an analysis of road safety management and a statistical approach are combined to shape the profile of the Moroccan road safety system and to identify candidate roads to be targeted by public actions, in addition to partly identifying factors associated with the probability that a person involved in a traffic crash dies.

The analysis of road safety management would focus on five axes, namely, the legal and regulatory framework of road safety in Morocco, the public and private stakeholders involved in road safety and the synergies between them, the sector strategy and its implementation, an inventory of road safety infrastructure and resources, and an analysis of road safety processes, measures, and indicators adopted in Morocco.

The statistical analysis would focus on two main aspects: an analysis of the main road safety indicators on the Moroccan motorway network and an analysis of the factors associated with the probability that a person involved in a traffic crash dies across the entire Moroccan road network (including the motorway network).

The primary road safety variables examined in the statistical analysis are the numbers of fatalities, crashes, severe injuries, and minor injuries. These indicators are analysed using a range of count data models, including the standard Poisson, Negative Binomial, and Conway–Maxwell–Poisson (COM-Poisson) models, as well as zero-inflated models, hurdle models, and mixed-effects models.

The selection of these models is motivated by the need to account for several statistical characteristics of count data. In particular, count data often exhibit overdispersion and an excess of zeros (Wan et al. 2010). Accordingly, the standard Poisson model, which assumes equidispersion, is compared with the Negative Binomial model, which can

accommodate overdispersion, and the Conway–Maxwell–Poisson model, which can accommodate underdispersion, equidispersion, and overdispersion.

Furthermore, zero-inflated and hurdle models are examined to address the issue of excess zeros and to investigate their potential contribution to overdispersion, since an excess of zeros can be a source of overdispersion (Dean and Lundy 2016). The comparison of the Zero-Inflated Poisson (ZIP) and Zero-Inflated Negative Binomial (ZINB) models allows for an assessment of whether accounting for overdispersion, in addition to zero inflation, improves the model's ability to fit the observed data.

Dean and Lundy (2016) suggested that there are two types of models that are often used to address excess zeros, namely, the Zero-Inflated models and the Hurdle models. Both models assume different data-generating processes for count data. The Zero-inflated models distinguish between structural zeros, which arise from a subgroup not at risk for an event, and sampling zeros that occur randomly within the other subgroup, which is at risk, with sampling zeros generated from the same discrete distribution that generates positive counts. On the other hand, the Hurdle models separately model the zero counts with a point mass at zero and utilize a truncated discrete distribution on the positive counts. In the present thesis, a comparative study is also conducted between the Zero-Inflated Poisson, the Zero-inflated Negative Binomial, the Hurdle Poisson, and the Hurdle Negative Binomial models to address the issue of excess zeros, while still identifying the best-suited model to do so.

In this thesis, the issue of unobserved heterogeneity is also examined through mixed-effects models. Unobserved heterogeneity refers to variations in the effects of explanatory variables across observations that are unobserved to the analyst (Mannering, Shankar, and Bhat 2016). Mixed-effects models are a departure from the assumption that the effects of explanatory variables are the same across all observations. They introduce random-effect components to allow for such unobserved variations.

The use of mixed-effects models is expected to be highly beneficial in two respects: first, it would provide a means of examining the presence of unobserved heterogeneity, as the available dataset only contains two explanatory variables, namely, the Average Annual Daily Traffic and Road Section Length, second, it would allow for an examination of another potential source of overdispersion as unobserved heterogeneity is also a common source of it, alongside excess zeros (Altinisik 2023). It is worth noting that four mixed-effects models are examined in this thesis namely, the random-intercept Poisson model, the random-intercept Negative Binomial model, the random-slope Poisson model, and the random-slope Negative Binomial model.

Overall, the comparison of all these models would provide insight into the statistical characteristics of road safety variables collected on the Moroccan motorway network in terms of dispersion, excess zeros and unobserved heterogeneity. Moreover, it would allow for the identification of the model that best fits the data, the estimated parameters of which would be adopted to capture the relationship between the response safety variables and the explanatory variables.

Ultimately, the identification of the model that provides the best fit to the data would serve two of the main purposes of the present thesis. First, the high-risk road sections

would be identified based on the predicted mean from this model, thereby serving one of the main purposes of this thesis, which is the prioritization of road sections that should be acted upon first. Second, the identification of the effect of explanatory variables on the response variables, which is captured by the estimated parameters of the best model, would provide a means to assess the effect of changes in the explanatory variables on the expected response variable.

The second aspect of the statistical analysis concerns the factors associated with the probability that a person involved in a traffic crash dies across the Moroccan road system. The analysis will be performed using the probit and logit models, which are used to model the probability of binary response variables.

Several types of explanatory variables will be examined via the logit and probit models. These variables include those capturing characteristics of the driver, such as age, sex, and profession, among others. Other variables are included in the models to capture the effect of weather conditions and road conditions, among others, on the response variable.

Eventually, this thesis would provide a set of recommendations to improve road safety based on the results obtained through the road safety management analysis and the statistical analysis. The recommendations would provide decision-makers with some practical means to improve road safety in Morocco.

REFERENCES

Altinisik, Y. 2023. "Addressing Overdispersion and Zero-Inflation for Clustered Count Data via New Multilevel Heterogenous Hurdle Models." *Journal of Applied Statistics* 50 (2): 408–433. <https://doi.org/10.1080/02664763.2022.2096875>.

Dean, C. B., and E. R. Lundy. 2016. "Overdispersion." In *Wiley StatsRef: Statistics Reference Online*, edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri, and J. L. Teugels. Wiley. <https://doi.org/10.1002/9781118445112.stat06788.pub2>.

International Transport Forum. 2024. "Morocco: Road Safety Country Profile 2023." Paris: OECD Publishing. <https://www.itf-oecd.org/sites/default/files/morocco-road-safety.pdf>. Maghreb Arab Press Agency. 2026. "Plus de 8 MMDH mobilisés pour la mise en œuvre de la stratégie nationale de sécurité routière 2026–2030". Accessed September 12, 2026. <https://maroc.ma/fr/actualites/plus-de-8-mmdh-mobilises-pour-la-mise-en-oeuvre-de-lastrategie-nationale-de-securite-routiere-2026>.

Mannering, Fred L., Venky Shankar, and Chandra R. Bhat. 2016. "Unobserved Heterogeneity and the Statistical Analysis of Highway Accident Data." *Analytic Methods in Accident Research* 11: 1–16. <https://doi.org/10.1016/j.amar.2016.04.001>.

National Agency of Road Safety (NARSA). 2022. "La sécurité routière au Maroc : Les progrès et les défis." <https://unece.org/sites/default/files/2022-09/ECE-TRANS-WP1-Presentation2022-14e.pdf>.

National Agency of Road Safety (NARSA). n.d.-a. "Fiche technique sur les indicateurs provisoires de la sécurité routière au titre de l'année 2024".

Wan Yaacob, Wan Fairos, Mohamad Alias Lazim, and Yap Bee Wah. 2010. "A Practical Approach in Modelling Count Data." In Proceedings of the Regional Conference on Statistical Sciences 2010.

https://www.researchgate.net/publication/265865874_A_Practical_Approach_in_Modelling_Count_Data.

Wijnen, Wim, and Henk Stipdonk. 2016. "Social Costs of Road Crashes: An International Analysis." *Accident Analysis & Prevention* 94: 97–106.

<https://doi.org/10.1016/j.aap.2016.05.005>.

World Health Organization. 2026. "Road Traffic Injuries." Accessed September 12, 2026. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.

1 Casualty accidents denotes accidents that caused personal injury (death or injury), as opposed to accidents that only caused property damage.

Sujet 8

Modeling Interconnections and Anticipating Systemic Risk in Financial Ecosystems

Sujet proposé :

Pr. Badreddine BENYACOUB, INSEA

Pr. Abdellah MANADIR, INSEA

Contexte et problématique

Financial ecosystems are becoming increasingly complex due to the growing interconnections between financial institutions, markets and other economic actors. These interconnected relationships create multiple channels through which financial shocks can propagate and potentially threaten the stability of the overall financial system.

Traditional systemic risk approaches often struggle to simultaneously capture the complexity and dynamic nature of financial interconnections and contagion mechanisms. This research explores how advanced statistical and artificial intelligence methods can improve the modeling, measurement and anticipation of systemic risk, with a particular focus on dynamic financial networks and Graph Neural Networks (GNNs). Using financial data and network-based approaches, the study investigates how interconnected structures evolve and how systemic risk can be identified and anticipated. The research contributes to a better understanding of financial contagion and to the development of more effective early-warning approaches for complex financial ecosystems.

Keywords: Systemic Risk – Financial Ecosystems – Artificial Intelligence – Financial Interconnections – Dynamic Networks – Risk Anticipation.

Sujet 9

Modélisation des événements extrêmes dans les réseaux de transport et les chaînes logistiques

Sujet proposé :

Pr. Badreddine BENYACOB, INSEA

Pr. Fayçal MIMOUNI, INSEA

Contexte et problématique

Les réseaux de transport et les chaînes logistiques sont conçus pour fonctionner en régime normal. Leurs coûts les plus lourds proviennent pourtant d'épisodes rares : pluies intenses et inondations qui coupent des axes, accidents ou fermetures qui paralysent un corridor, congestion exceptionnelle, ou ruptures d'approvisionnement. Ces épisodes engendrent des retards, des pertes d'exploitation et des surcoûts que les indicateurs moyens ne permettent pas de décrire.

Cette thèse propose de développer et d'évaluer des méthodes statistiques pour mesurer, modéliser et anticiper les extrêmes dans les systèmes de transport et de logistique. Les travaux pourront notamment porter sur :

- La dépendance extrême entre segments d'un réseau ou entre sites logistiques ;
- La simulation de scénarios de stress plausibles ;
- L'estimation des queues des distributions des temps de parcours ou des retards et le calcul de mesures de risque ;
- La mesure de la résilience structurelle des réseaux : identification des facteurs déclenchant un effondrement du réseau, détection des nœuds et des tronçons critiques lors d'événements extrêmes, et caractérisation des phénomènes de percolation conduisant à la fragmentation du réseau.

Mots-clés : *Théorie des valeurs extrêmes ; dépendance extrême ; réseaux de transport ; logistique ; résilience ; percolation ; mesures de risque ; simulation.*

Références indicatives

1. Engelke, S. et Hitz, A. S. (2020). Graphical models for extremes. *Journal of the Royal Statistical Society, Series B*, 82(4), 871–932.
2. Fałdziński, M., Osińska, M. et Zalewski, W. (2021). Extreme value theory in application to delivery delays. *Entropy*, 23(7), 788.
3. Bhuyan, P., Jana, K. et McCoy, E. J. (2023). Causal analysis at extreme quantiles with application to London traffic flow data. *Journal of the Royal Statistical Society, Series C: Applied Statistics*, 72(5), 1452–1474.

4. Wang, H. W., Peng, Z. R. et al. (2020). Evaluation and prediction of transportation resilience under extreme weather events: A diffusion graph convolutional approach. *Transportation Research Part C: Emerging Technologies*, 115, 102619.
-

Sujet 10

Politique budgétaire, financement de l'innovation technologique et industrialisation : une analyse à l'aide d'un modèle dynamique d'équilibre général.

Sujet proposé :

Pr. Badreddine BENYACOU, INSEA

Pr. Abdellah MANADIR, INSEA

Contexte et problématique

Dans un contexte international caractérisé par l'intensification de la concurrence, la recomposition des chaînes de valeur mondiales et la transition vers des modèles de production plus durables, la question du rôle de l'État dans l'accompagnement de cette transformation devient centrale. A cet effet, la politique budgétaire apparaît comme un levier stratégique majeur de transformation productive. Au-delà de sa fonction traditionnelle de stabilisation conjoncturelle, elle influence la structure sectorielle de l'économie à travers l'investissement public, les dépenses en commande publique d'infrastructure, les incitations fiscales et les mécanismes de soutien aux secteurs productifs. La question centrale qui se pose est donc la suivante : dans quelle mesure la politique budgétaire marocaine contribue-t-elle à stimuler l'industrialisation et à renforcer en conséquence la capacité productive nationale ?

Dans le même temps, la contrainte budgétaire de l'État marocain s'est resserrée, tandis que le financement de l'innovation technologique — dépense intérieure de R&D, soutien à la R&D privée, capital-risque, incitations fiscales — demeure faible au regard des seuils requis pour soutenir une montée en gamme technologique.

Keywords:

Politique budgétaire, Financement de l'innovation technologique, Industrialisation, modèle dynamique d'équilibre général, Modèle Stock-Flow Consistent (SFC).

Références indicatives

Agénor, P.-R. (2018). *Public Capital, Growth and Welfare: Analytical Foundations for Public Policy*. Princeton University Press.

Aghion, P., Cai, J., Dewatripont, M., Du, L., Harrison, A. et Legros, P. (2015). « Industrial Policy and Competition ». *American Economic Journal: Macroeconomics*, 7(4), 1-32.

Caverzasi, E. et Godin, A. (2015). « Post-Keynesian Stock-Flow-Consistent Modelling: A Survey ». *Cambridge Journal of Economics*, 39(1), 157-187.

Nikiforos, M. et Zezza, G. (2017). « Stock-Flow Consistent Macroeconomic Models: A Survey ». *Journal of Economic Surveys*, 31(5), 1204-1239.

Rodrik, D. (2016). « Premature Deindustrialization ». *Journal of Economic Growth*, 21(1), 1-33.

Sujet 11

Conception et analyse de métaheuristiques pour la résolution de problèmes d'optimisation complexes

Pr. Badreddine BENYACOU, INSEA

Pr. Mohamed OUZINEB, INSEA

Contexte et problématique

Face à la complexité croissante des problèmes décisionnels réels mêlant variables discrètes et continues (non-linéaires, multimodalité, explosion combinatoire), les méthodes d'optimisation exactes se heurtent à d'importantes limites de passage à l'échelle. Les métaheuristiques constituent une alternative incontournable pour explorer efficacement ces espaces de recherche hétérogènes.

Ce projet doctoral vise à concevoir, modéliser et analyser de nouvelles approches métaheuristiques génériques capables de traiter conjointement les domaines combinatoires et continus. Les travaux porteront sur l'élaboration de représentations algorithmiques adaptées, l'exploration de passerelles ou transformations d'espaces (du continu vers le discret), et le développement de mécanismes adaptatifs garantissant un équilibre rigoureux entre diversification globale et intensification locale. Les contributions théoriques seront adossées à une analyse approfondie des dynamiques de convergence et validées sur des bancs d'essai académiques ainsi que sur des cas d'usage réels à grande échelle.

Keywords:

Intelligence Artificielle, Métaheuristiques, Optimisation Combinatoire & Continue, recherche opérationnelle, Algorithmique Avancée.