

Système macro d'utilité révolutionnaire

Un modèle dynamique du potentiel de révolte, micro-fondé sur des fonctions d'utilité par classe sociale — note de cadrage scientifique

Version du 10 septembre 2026

1 Problème scientifique

On cherche à représenter, sous forme d'un **système dynamique déterministe**, la manière dont des **conditions matérielles** et des **états d'opinion**, différenciés par **classe sociale**, engendrent et font varier au cours du temps un **potentiel de révolte** agrégé.

Le modèle est **macro** — ses sorties sont des trajectoires d'indicateurs nationaux — mais **micro-fondé** : les équations agrégées se déduisent de **fonctions d'utilité par classe**, elles-mêmes estimées sur données individuelles. Il est **historique** : calé sur des séries longues du système statistique public, il sert ensuite à produire des scénarios de moyen terme.

L'objectif n'est pas la prévision d'un événement, mais l'étude de la **dynamique** — seuils, effets de retard, boucles de rétroaction — qui fait monter ou refluer la conflictualité sociale, classe par classe.

2 Cadre méthodologique

2.1 Filiation : dynamique des systèmes

La démarche reprend celle de Meadows *et al.* (*The Limits to Growth*, 1972, modèle World3) : un **grand système d'équations différentielles** de moyen terme, organisé en **variables d'état** (stocks), **flux** et **boucles de rétroaction**, intégré numériquement pour reconstituer une évolution historique puis la prolonger. On en retient la logique de couplage et de rétroaction, appliquée ici à des variables sociales et non démo-économiques.

2.2 Formalisation

Soit C classes sociales indexées par c et, pour chaque classe, un vecteur d'état $x_c(t)$ regroupant ses variables (précarité, malnutrition, colère, défiance, conscience de classe, solidarités). Chaque classe est dotée d'une fonction d'utilité

$$U_c(t) = U_c(x_c(t); \beta_c)$$

où β_c est le vecteur de **poids** propres à la classe. La dynamique s'écrit comme un système différentiel couplé de dimension $C \cdot \dim(x)$:

$$\frac{dx}{dt} = f(x(t), x(t - \tau); \theta)$$

avec retards τ (effets différés), termes de couplage inter-classes (une dégradation subie par une classe affecte les autres via l'emploi, les prix, l'imitation) et paramètres structurels θ . Le **potentiel**

de révolte agrégé est une fonctionnelle de l'état, liée aux écarts entre utilité de référence et utilité courante, pondérés par la capacité d'action collective :

$$R(t) = \sum_c \omega_c \cdot g(U_c^* - U_c(t), S_c(t))$$

où S_c agrège les solidarités — elles peuvent *amplifier* l'action collective ou *amortir* le mécontentement, le signe étant laissé à l'estimation.

2.3 Reconstruction des variables latentes

Les variables sans mesure statistique directe (colère, défiance, conscience de classe, solidarités) sont traitées comme **latentes** et reconstruites à partir d'indicateurs observés « matériels » et déclaratifs, par un **modèle de mesure** :

- analyse factorielle confirmatoire / modèles à équations structurelles (SEM), ou modèles *MIMIC* (indicateurs multiples, causes multiples) ;
- estimation d'un score latent par classe et par période, avec contrôle de l'**invariance de mesure** dans le temps (comparabilité des millésimes) ;
- fusion d'enquêtes hétérogènes par **appariement statistique** sur variables communes (PCS, âge, revenu, région).

2.4 Estimation des poids et calage

- **Poids β_c des fonctions d'utilité** : régressions sur données individuelles empilées (panel répété), par classe, de la satisfaction / des attitudes sur les composantes matérielles.
- **Paramètres structurels θ** (retards, gains de boucle, élasticités) : calage du système sur les **séries historiques** de la variable-cible et des variables d'état — moindres carrés sur trajectoires, ou *inférence indirecte* lorsque la vraisemblance est intraitable.
- **Validation** : ajustement hors échantillon (*backtesting*), analyse de sensibilité et d'identifiabilité, robustesse au choix de la nomenclature et de la variable-cible.
- **Scénarios** : variation des entrées exogènes (conjoncture, politiques sociales) et lecture des trajectoires de $R(t)$.

2.5 Construction des classes sociales

Nomenclature à **quatre axes croisés**, puis résumé statistique :

Axe	Contenu et variables sources
A — Position socioprofessionnelle	PCS INSEE 2020 (6 à 42 postes) ; inactifs ventilés par ancienne PCS ; table de passage PCS 2003 ↔ 2020 pour la profondeur historique
B — Rapport au capital	statut (employeur / indépendant / salarié), encadrement, part des revenus du patrimoine, détention d'entreprise — sépare bourgeoisie économique et salariat à revenu égal
C — Niveau de vie	déciles / vingtiles de revenu disponible par unité de consommation ; seuils D1, médiane, D9, top 1 %
D — Capital scolaire	diplôme le plus élevé ; diplôme des parents si disponible (reproduction, déclassement)

Deux objets complémentaires : une **grille a priori** de 7 à 9 classes nommées (bourgeoisie économique, cadres supérieurs, petite bourgeoisie indépendante, professions intermédiaires, salariat stable d'exécution, précaires, hors-emploi durable, retraités), pour l'interprétation ; et une **typologie statistique** obtenue par classification ascendante hiérarchique pondérée ou *analyse en classes latentes* sur $A \times B \times C \times D$ ($k \approx 8-12$, choisi par critères d'information), pour le calage des utilités.

3 Outils

Fonction	Outil retenu
Langage principal	Python 3 (écosystème scientifique)
Manipulation de données	<code>pandas</code> , <code>polars</code> , <code>pyarrow</code> (Parquet), DuckDB pour les fichiers volumineux (recensement : > 17 M lignes)
Pondérations d'enquête, appariement	<code>samplics</code> ; appariement statistique (<i>hot-deck</i> , plus proche voisin)
Classification / typologie	<code>scikit-learn</code> (CAH, k-means), <code>scipy.cluster</code> , classes latentes (<code>stepmix</code>)
Variables latentes / SEM	<code>semopy</code> (Python) ou <code>lavaan</code> (R) ; modèles MIMIC ; analyse factorielle
Économétrie	<code>statsmodels</code> , <code>linearmodels</code> (panel), <code>pyfixest</code>
Intégration du système différentiel	<code>scipy.integrate</code> (<code>solve_ivp</code> , schémas raides) ; équations à retard via <code>ddeint</code> ou formulation augmentée
Dynamique des systèmes (maquettage)	PySD, BPTK-Py pour l'articulation stocks / flux
Calage et sensibilité	<code>lmfit</code> / <code>scipy.optimize</code> ; SALib (indices de Sobol) ; inférence indirecte, ABC (<code>pyABC</code>) si besoin
Reproductibilité	environnement conda/venv figé, pipeline scripté (<code>Snakemake</code> ou <code>make</code>), suivi Git, données brutes isolées et versionnées à part
Restitution	<code>matplotlib</code> / <code>plotly</code> ; trajectoires historiques passé + projeté sur un même graphe

4 Données

Règle : sources **exclusivement publiques**, admises si (1) accessibles en open source et en version complète, (2) documentées, (3) produites par une institution statistique fiable. Premier lot constitué le 10 septembre 2026 ($\approx 1,4$ Go, formats Parquet et CSV, dictionnaires et questionnaires joints).

4.1 Bases acquises (téléchargées, licence ouverte)

Source	Structure	Contenu / variables du modèle
INSEE — Enquête Emploi 2023 (fichier détail)	348 624 individus × 83 variables	précarité (CDD, intérim, temps partiel subi, sous-emploi, halo), PCS, statut, encadrement, diplôme, secteur
INSEE — Filosofi (communal 2021 ; carroyé 200 m 2019 et 2021)	833 k lignes (SDMX) ; 2,3 M carreaux	niveau de vie, déciles, taux et intensité de pauvreté, composition du revenu disponible — axe C
INSEE — Recensement 2021 (fichiers détail)	individus : 17,4 M × 92 ; logements : 26,0 M × 69	socle des nomenclatures — PCS, diplôme, statut d'activité, occupation du logement, composition familiale, migrations
DREES — Baromètre d'opinion (panel 2000–2025)	94 338 individus × 359 variables	sentiment de justice / d'injustice sociale, défiance, redistribution, perception des inégalités — avec PCS, région, statut : estimation par classe directe
SSMSI — sécurité intérieure	séries mensuelles depuis 2016 (88 k points) ; bases communale (5,2 M) / dép. / rég.	proxy de la colère et de la conflictualité non gréviste ; sentiment d'insécurité
DARES — grèves / JINT	séries annuelles 1975–2004 et depuis 2005	conflictualité du travail (candidate variable-cible), syndicalisation, négociation

4.2 Bases en accès chercheur (ADISP / CASD, gratuit sur inscription recherche)

Source	Contenu / variables du modèle
INSEE — Enquête Budget de famille	part du budget alimentaire, reste à vivre, ~ 900 postes de consommation, transferts monétaires entre ménages (solidarité familiale matérielle)
INSEE — SRCV / EU-SILC France	privation matérielle et sociale (13 items), restrictions alimentaires, arriérés, soutien social , isolement ; volet longitudinal
CEVIPOF — Baromètre de la confiance politique (2009→, 17 vagues)	défiance envers personnel politique / institutions / médias, confiance interpersonnelle, item colère , auto-positionnement social , rapport à la démocratie

Source	Contenu / variables du modèle
INJEP / DJEPVA — Baromètre sur la jeunesse (2016→)	défiance, sentiment d'utilité et d'isolement, engagement (partis, syndicats, collectifs), aides reçues / données, renoncements

4.3 Correspondance variables ↔ sources

Variable du modèle	Sources principales
Précarité	Enquête Emploi, Filosofi, Recensement, Baromètre DREES
Malnutrition / insécurité alimentaire	Budget de famille, SRCV
Colère	CEVIPOF, Baromètre DREES, SSMSI (proxy)
Défiance (politiques / institutions / médias)	CEVIPOF, Baromètre DREES, INJEP
Conscience de classe	CEVIPOF (auto-positionnement), Enquête Emploi + Recensement (position objective), DREES
Solidarité professionnelle	DARES (syndicalisation, grèves), Enquête Emploi
Solidarité familiale	Budget de famille (transferts), SRCV (soutien), Recensement (cohabitation)
Solidarité partisane	CEVIPOF, INJEP
Solidarité sectorielle	DARES, Enquête Emploi (secteur × statut)
Variable-cible (conflictualité)	DARES JINT, SSMSI

5 Limites méthodologiques identifiées

- **Comparabilité temporelle** : ruptures de nomenclature (PCS), d'enquête et de questionnaire ; à traiter par tables de passage et tests d'invariance de mesure.
- **Enquêtes d'opinion** hors service statistique public strict (CEVIPOF, INJEP) : retenues sous les trois critères, mais effets de maison et de mode de collecte à contrôler.
- **Identifiabilité** du système différentiel : nombre de paramètres élevé face à des séries courtes pour certaines variables — d'où priorité aux séries longues (grèves depuis 1975, Baromètre DREES depuis 2000) et régularisation.
- **Causalité** : le modèle formalise des mécanismes plausibles calés sur corrélations historiques ; les scénarios sont conditionnels, non prédictifs.
- **Endogénéité des solidarités** : signe et sens de causalité à établir empiriquement.

6 Questions ouvertes et étape suivante

À trancher : la variable-cible de calage (grèves seules, SSMSI, ou indice composite pondéré) ; l'unité d'analyse (individu ou ménage) ; le seuil de la « bourgeoisie économique » ; la forme fonctionnelle des utilités.

Prochaine étape : construire la **table individus harmonisée** (Recensement 2021 + Enquête Emploi 2023) — variables des quatre axes, poids de sondage, PCS harmonisée, décile approché via Filosofi —, socle sur lequel se branchent les fonctions d'utilité puis le système d'équations.

*Documents de référence du projet : **consignes.md** (cahier des charges), **nomenclatures.md** (les quatre axes en détail), **base de données/MANIFEST.md** (inventaire détaillé, licences, correspondance variables ↔ sources, démarches d'accès chercheur).*