

La méthode des comparables à l'épreuve des données : validation empirique à grande échelle sur 390 996 ventes résidentielles québécoises

Sélection, nombre, ajustements, pondération et réconciliation des ventes comparables
— rétro-évaluation point-dans-le-temps sur les transactions Vrai Prix 2021-2026

Simon-Pierre Boucher — version 2026-10-10

10/10/2026

Table des matières

Résumé	1
1. Données	2
1.1 Source et nettoyage	2
1.2 Répartition	3
1.3 Caractéristiques disponibles et limites	3
2. Protocole	3
2.1 Rétro-évaluation point-dans-le-temps	4
2.2 Configurations testées	4
2.3 Mesures	4
3. Résultats	4
3.1 Combien de comparables ? À quelle distance ? De quelle ancienneté ?	5
3.2 Quels ajustements sont justifiés ?	7
3.3 Comment choisir les comparables ? (H1)	8
3.4 Pondération et réconciliation (H4)	8
3.5 Compétition de modèles	9
3.6 Où la méthode tient-elle, où défaille-t-elle ?	11
3.7 Intervalles de valeur	12
3.8 Vers un système de recommandation des comparables	13
4. Discussion	13
5. Reproductibilité	14

Résumé

Question. La méthode de comparaison des ventes, telle qu'enseignée et pratiquée en évaluation immobilière, produit-elle des estimations fiables, et lesquelles de ses règles (choix des comparables, nombre, rayon, ancienneté, ajustements, pondération, réconciliation) sont empiriquement justifiées ?

Données et protocole. 390 996 ventes de résidences unifamiliales du Québec (registre foncier jumelé au rôle d'évaluation, Vrai Prix, 2021-01-05 → 2026-09-29), après un nettoyage journalisé de 456 950 transactions brutes. Pour chacune des **189 354 ventes cibles** de 2024 (validation, 64 795) et de 2025-2026 (test hors échantillon, 124 559), le prix est masqué et la valeur est estimée avec les seules ventes conclues **avant** la date de la cible, un indice de prix ré-estimé chaque mois avec les seules ventes antérieures, et des grilles d'ajustement et modèles ré-estimés chaque trimestre. **2 256 configurations** de la méthode des comparables et treize modèles de référence sont appliqués aux mêmes cibles ; les différences sont testées par bootstrap par grappes (municipalité × trimestre).

Résultats principaux (test 2025-2026).

- 1. Les ajustements sont le facteur dominant.** Prendre la moyenne des cinq ventes les plus proches, ajustées seulement pour le temps (M2), donne une erreur absolue médiane (MdAPE) de **18,1 %** ; une grille d'ajustement hédonique (M7) la ramène à **13,0 %** (–5,1 points, IC 95 % de la différence de MdAPE [–5,24 ; –4,87]). Une grille conventionnelle stylisée (M6) fait presque aussi bien (13,4 %) : ce qui compte, c'est d'ajuster, plus que la finesse de la grille.
- 2. Choisir des comparables semblables vaut autant qu'ajuster.** Sans aucun ajustement physique, sélectionner les cinq ventes les plus **semblables** (aire, terrain, âge, lien physique, étages) plutôt que les cinq plus **proches** fait passer la MdAPE de 18,1 % à **13,0 %** (H1 confirmée : un comparable un peu plus éloigné mais semblable vaut mieux qu'un voisin dissemblable).
- 3. Le nombre optimal de comparables est bien supérieur à trois.** Avec la grille hédonique, la MdAPE décroît de 16,5 % ($k = 1$) à 13,8 % ($k = 3$), 13,0 % ($k = 5$), 12,4 % ($k = 10$), atteint son minimum à **$k = 20$** (12,1 %) et n'augmente plus jusqu'à $k = 50$ (12,1 %), quelle que soit la densité du marché ; sans ajustement physique, l'optimum est d'environ dix et l'erreur remonte ensuite (H2 confirmée seulement pour la méthode non ajustée).
- 4. La fenêtre temporelle doit être large ; le rayon importe peu.** La meilleure configuration de la surface Erreur(k , d , t) choisie en 2024 est « compromis distance / similarité, $k = 30$, ≤ 25.0 km, 24 mois, grille hédonique » (11,9 % en validation → **11,1 %** en test). Les fenêtres de 3 mois sont dominées par 12-24 mois à tous les rayons ; au meilleur k , le rayon (0,5 à 25 km) ne change presque rien à la précision, car k est un maximum et un grand rayon n'ajoute des ventes que là où les voisins manquent.
- 5. La pondération et le mode de réconciliation comptent peu** (effets marginaux moyens de 13,1 à 13,6 points selon la pondération, médiane 13,2 contre moyenne 13,6), bien moins que la sélection (12,9-14,4) et l'ajustement (11,8-15,4).
- 6. La meilleure méthode des comparables, entièrement transparente et choisie en 2024, atteint 11,1 % en 2025-2026** (M0 : compromis distance / similarité, $k = 30$, ≤ 25.0 km, 24 mois, grille hédonique, pondération égale, médiane), contre **9,0 %** pour un gradient boosting ré-entraîné chaque trimestre avec la valeur au rôle (M11b) et 10,7 % sans le rôle (M11a). L'hybride comparables + GBM (M12) atteint **8,9 %**. Le coût de la transparence est donc de 2,2 points de MdAPE face au meilleur modèle opaque (0,4 point sans le rôle), et il se réduit à 1,4 point dans le tiers des sous-marchés les plus homogènes (H5).
- 7. Le ratio au rôle est un comparable redoutable** : rôle de la cible $\times$ ratio prix/rôle médian des cinq plus proches ventes du même rôle (M13) donne 11,5 %, mieux que toute grille d'ajustement physique, mais seulement pour 90,3 % des cibles.
- 8. La fiabilité dépend d'abord de la liquidité du sous-marché** : la MdAPE de M7 passe de 10,2 % là où plus de 200 ventes ont eu lieu à moins de 2 km dans l'année à 25,3 % là où il y en a eu moins de 5 ; des règles choisies strate par strate (système de recommandation) obtiennent 11,3 % avec 93,4 % de couverture, sans faire mieux que la meilleure règle unique (11,1 %, couverture complète) : une règle « filet large » qui retient les ventes les plus semblables dans un grand rayon s'adapte d'elle-même à la densité.

Portée. Nous validons la précision prédictive de règles professionnelles *simulées* sur des caractéristiques observables (aire, terrain, âge, étages, lien physique, rôle) ; les ajustements non observables (garage, piscine, chambres, qualité, rénovations) restent dans l'erreur résiduelle, et le jugement réel d'un évaluateur n'est pas reproduit.

1. Données

1.1 Source et nettoyage

Instantané de la base de production Vrai Prix pris le 2026-10-10 : 766 596 transactions résidentielles et 3,75 M d'unités d'évaluation. Nous retenons les ventes de type « unifamilial » jumelées à une unité du rôle, et appliquons des règles de nettoyage explicites et journalisées — rien n'est éliminé silencieusement.

Règle	n
R0 sans jumelage au rôle (id_provinc / code_mun)	9 255
R1 doublons (même unité, même date)	355
R2 unité du rôle non unifamiliale (chalet, plex, terrain...) ou > 1 logement	29 509
R3 coordonnées absentes ou hors Québec	0
R4 aire habitable absente ou hors [40, 600] m ²	3 797
R5 superficie de terrain absente ou hors [80, 20 000] m ²	4 278
R6 année de construction absente, < 1850 ou postérieure à la vente (vente de terrain)	2 393
R6b vente l'année même de la construction (prix du terrain ou d'une construction inachevée : médiane 346 k\$ et ratio au rôle 1,15 contre 450-480 k\$ et 1,3 aux âges 1-4)	5 427

Règle	n
R7 prix hors [50 k, 3M]	528
R8 rôle absent ou ratio prix/rôle hors [0,33 ; 3] (transactions présumées non représentatives)	8 074
R9 revente $\leq$ 180 jours avec variation > 40 % (flip)	794
R10 lien physique ou genre de construction absent au rôle	1 512
R11 municipalité absente du répertoire MAMH	32

Jeu propre : **390 996 ventes** (2021-01-05 → 2026-09-29) ; prix médian 400 000 \$, moyenne 449 464 \$.

Deux règles méritent un commentaire. R8 écarte les ratios prix/rôle extrêmes ($< 0,33$ ou > 3) : faute d'indicateur de lien de dépendance entre les parties, c'est notre proxy de « transaction non représentative » — un évaluateur les écarterait aussi. R6b écarte les ventes conclues l'année même de la construction : leur prix médian (346 108 \$) est leur ratio au rôle (1, 15) sont ceux de terrains sous construction sinacheves, non de maisons sacheves (450—480k et 1,3 aux âges 1-4 ans).

1.2 Répartition

Période	Rôle	Ventes
2021-01 → 2023-12	développement	201 642
2024	validation (choix des configurations)	64 795
2025-01 → 2026-09	test hors échantillon	124 559

Régions les plus représentées : Montérégie (78 077), Laurentides (43 254), Capitale-Nationale (37 618), Lanaudière (36 436), Estrie (28 964), Montréal (27 934).

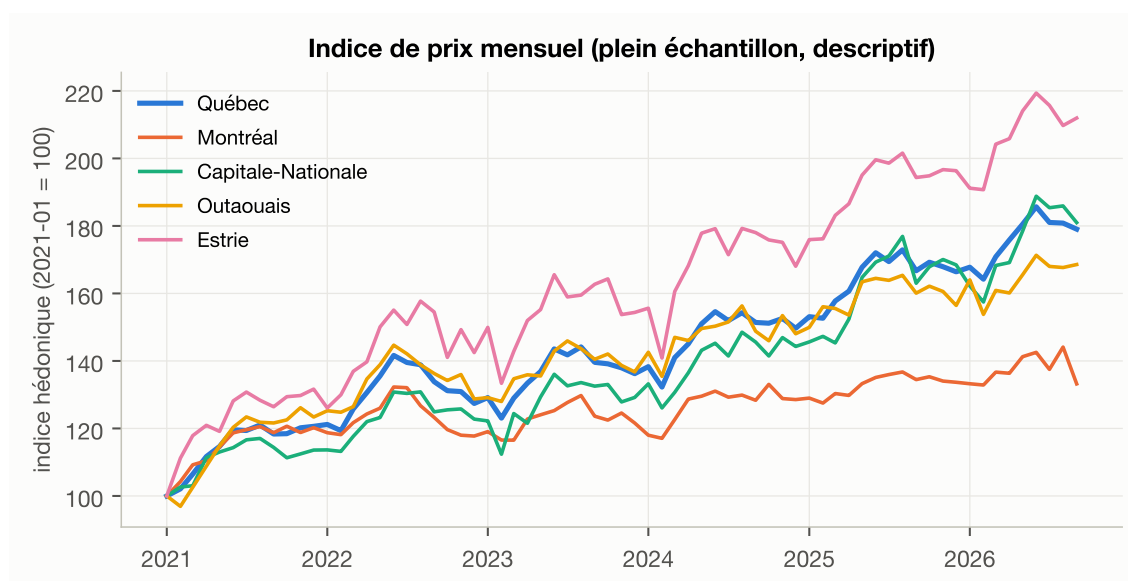


Fig. 1 : Indice de prix mensuel par région (descriptif, plein échantillon).

1.3 Caractéristiques disponibles et limites

Observées pour chaque vente : date, prix, coordonnées, municipalité, unité de voisinage, aire habitable, superficie du terrain, année de construction, nombre d'étages, lien physique (détaché, jumelé, en rangée), genre de construction, valeur au rôle en vigueur à la vente et sa part terrain. Non observées : garage, piscine, nombre de chambres et de salles de bains, qualité, état, rénovations. Les grilles d'ajustement testées ne portent donc que sur les premières ; les secondes sont une source d'erreur commune à toutes les méthodes.

2. Protocole

2.1 Rétro-évaluation point-dans-le-temps

Pour chaque cible : (1) le prix est masqué; (2) le bassin de comparables est constitué des ventes conclues strictement avant la date de la cible (la cible elle-même, si elle a déjà été vendue, est exclue : on évalue par comparables, non par vente répétée), dans un rayon de 25 km et une fenêtre de 24 mois; (3) l'indice temporel régional utilisé pour ramener chaque comparable au niveau du dernier mois connu est ré-estimé chaque mois avec les seules ventes antérieures (régression hédonique à effets fixes de municipalité et effets mensuels, par région administrative, rétrécie vers l'indice provincial quand les ventes sont rares); (4) les coefficients d'ajustement hédoniques, les grilles régionales et les modèles d'apprentissage sont ré-estimés chaque trimestre avec les seules ventes antérieures au trimestre; (5) chaque configuration produit une valeur et, le cas échéant, une dispersion des prix ajustés.

Le niveau de prix du mois de la cible n'est jamais connu : toutes les méthodes travaillent au niveau du mois précédent, ce qui reproduit la situation réelle d'un évaluateur.

2.2 Configurations testées

E1 — surface Erreur(k, d, t) : $k \in \{1, 3, 5, 7, 10, 15, 20, 30, 50\}$, $d \in \{0,5; 1; 2; 5; 10; 25 \text{ km}\}$, $t \in \{3, 6, 12, 24 \text{ mois}\}$, sélection $\in \{\text{plus proches, plus semblables, compromis}\}$, ajustement $\in \{\text{temps seul, grille hédonique}\}$, médiane. 1 296 configurations.

E2 — plan factoriel des règles aux ancrs $(k, d, t) = (3, 2 \text{ km}, 6 \text{ mois}), (5, 5 \text{ km}, 12 \text{ mois}), (10, 10 \text{ km}, 24 \text{ mois})$: sélection (4) $\times$ ajustement (8 : aucun, temps seul, grille conventionnelle, hédonique globale, hédonique régionale, régression locale, GBM, ratio au rôle) $\times$ pondération (5) $\times$ réconciliation (2). 960 configurations.

Grille conventionnelle stylisée (M6) : aire habitable ajustée à la moitié du prix au m² du comparable; terrain à la moitié de la valeur foncière implicite au m² (plafond $\pm 15\%$); âge à $-0,5\%$ par année d'écart (plafond $\pm 20\%$); $\pm 5\%$ entre détaché et attaché. *Grille hédonique* (M7) : coefficients d'une régression du log du prix sur les caractéristiques, à effets fixes de municipalité (within), ré-estimée chaque trimestre. *Régression locale* (M8a) : mêmes coefficients ré-estimés sur les ventes à moins de 10 km des 24 derniers mois (≥ 40 ventes, ridge vers la grille régionale). *GBM* (M8b) : l'ajustement est la différence des prédictions d'un gradient boosting entre la cible et le comparable. *Ratio au rôle* (M13) : valeur au rôle de la cible $\times$ ratio prix/rôle du comparable, pour les comparables de la même municipalité et du même rôle triennal.

2.3 Mesures

Erreur absolue en pourcentage $APE = |\hat{P} - P| / P$; MdAPE (médiane), MAPE (moyenne), biais médian signé, parts à ± 5 , ± 10 et $\pm 20\%$, couverture (part des cibles pour lesquelles la configuration produit une valeur — une configuration trop restrictive ne «rate» pas, elle s'abstient, et cela doit se voir). Les différences entre méthodes sont des différences d'erreurs appariées sur les mêmes cibles, avec intervalles de confiance à 95 % par bootstrap par grappes municipalité $\times$ trimestre (400 répliques).

3. Résultats

3.1 Combien de comparables? À quelle distance? De quelle ancienneté?

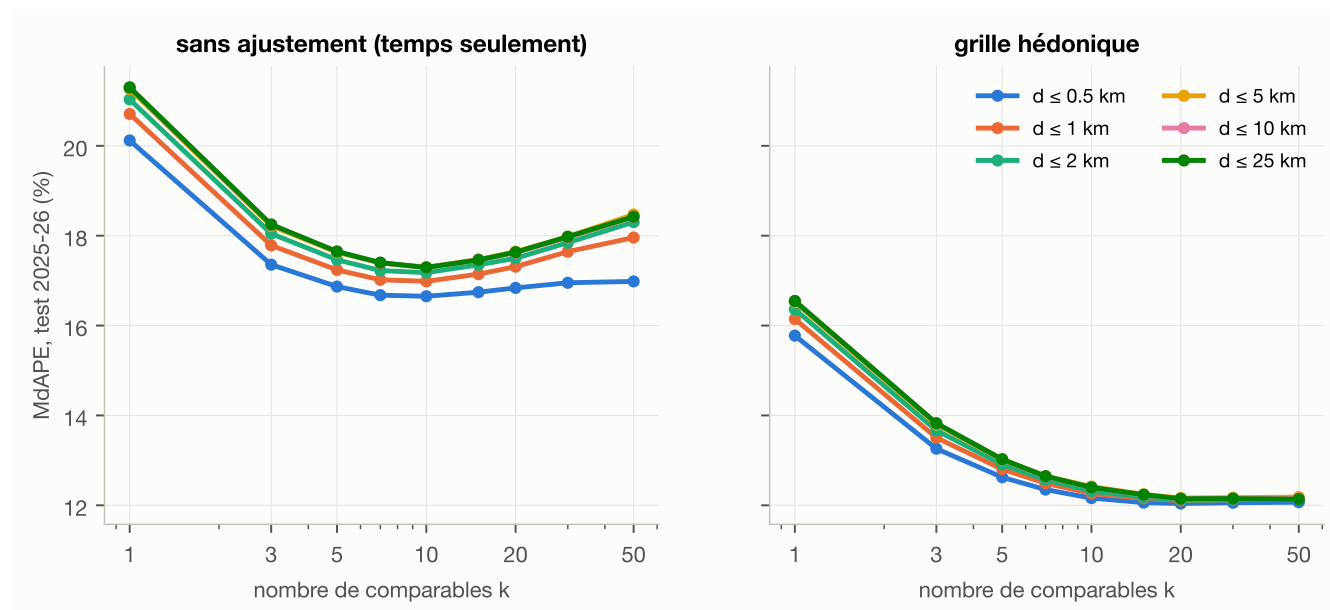


Fig. 2 : MdAPE (test 2025-2026) selon le nombre de comparables, par rayon, fenêtre de 12 mois, sélection par distance.

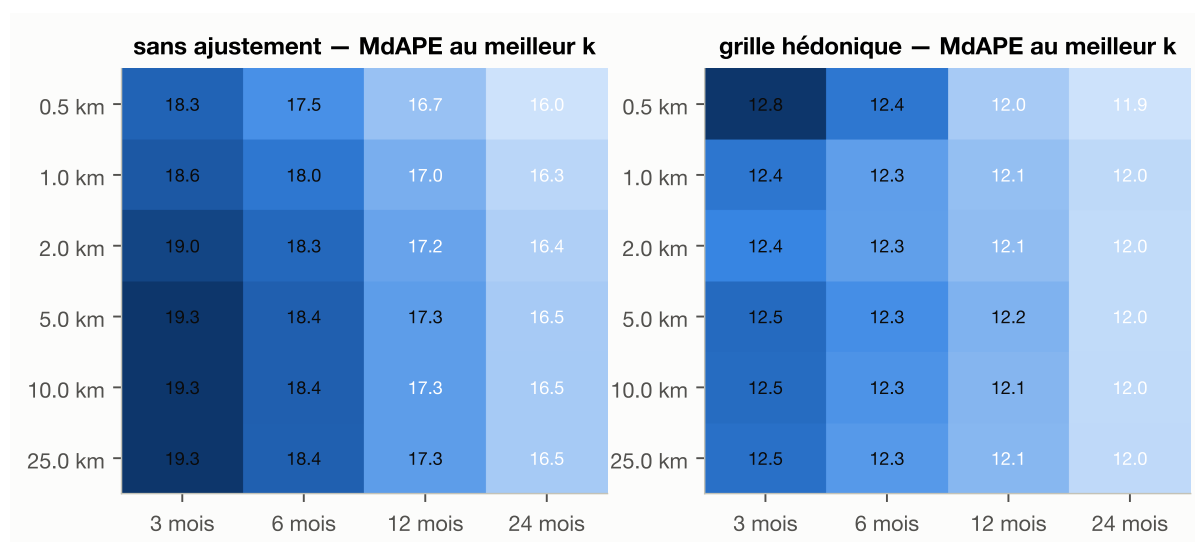


Fig. 3 : MdAPE au meilleur k pour chaque couple rayon × fenêtre.

Sélection par distance, grille hédonique, fenêtre 12 mois — MdAPE test (%)

rayon	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
≤ 0.5 km	15,8	13,3	12,6	12,3	12,2	12,1	12,0	12,1	12,1
≤ 1.0 km	16,1	13,5	12,8	12,5	12,3	12,2	12,1	12,1	12,1
≤ 2.0 km	16,4	13,7	12,9	12,6	12,3	12,2	12,1	12,1	12,1
≤ 5.0 km	16,5	13,8	13,0	12,7	12,4	12,2	12,2	12,2	12,2
≤ 10.0 km	16,5	13,8	13,0	12,6	12,4	12,2	12,1	12,1	12,1
≤ 25.0 km	16,5	13,8	13,0	12,6	12,4	12,2	12,1	12,1	12,1

Sélection par distance, temps seulement, fenêtre 12 mois — MdAPE test (%)

rayon	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
≤ 0.5 km	20,1	17,4	16,9	16,7	16,7	16,7	16,8	17,0	17,0
≤ 1.0 km	20,7	17,8	17,2	17,0	17,0	17,1	17,3	17,6	18,0

rayon	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
≤ 2.0 km	21,0	18,0	17,5	17,2	17,2	17,4	17,5	17,8	18,3
≤ 5.0 km	21,3	18,2	17,6	17,4	17,3	17,5	17,6	18,0	18,5
≤ 10.0 km	21,3	18,2	17,6	17,4	17,3	17,5	17,6	18,0	18,4
≤ 25.0 km	21,3	18,2	17,6	17,4	17,3	17,5	17,6	18,0	18,4

Sélection par similarité physique, grille hédonique, fenêtre 12 mois — MdAPE test (%)

rayon	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
≤ 0.5 km	14,8	12,6	12,0	11,8	11,7	11,8	11,9	12,0	12,1
≤ 1.0 km	15,2	12,7	12,0	11,8	11,6	11,6	11,6	11,7	11,9
≤ 2.0 km	15,5	12,9	12,1	11,8	11,5	11,5	11,4	11,4	11,6
≤ 5.0 km	16,1	13,3	12,5	12,1	11,7	11,6	11,6	11,6	11,6
≤ 10.0 km	16,4	13,6	12,7	12,3	12,0	11,8	11,7	11,7	11,8
≤ 25.0 km	16,6	13,7	12,8	12,4	12,1	12,0	11,9	11,9	12,0

Lecture. (i) Sans ajustement physique, l'erreur décroît avec k jusqu'à une dizaine de comparables puis remonte (les ventes supplémentaires sont de plus en plus dissemblables); (ii) avec la grille hédonique, elle décroît nettement plus, jusqu'à $k \approx 20$, et ne remonte plus; (iii) la fenêtre de 3 mois est systématiquement dominée par 12-24 mois, alors qu'au meilleur k le rayon ne change presque rien. Meilleures configurations de E1 par sélection × ajustement (choisies en 2024, évaluées en 2025-2026) :

sélection	ajustement	k	d	T	MdAPE val.	MdAPE test
geo	hedo	30	1,0	24	12,6	12,0
geo	none	10	1,0	24	17,3	16,3
mix	hedo	30	25,0	24	11,9	11,1
mix	none	10	25,0	24	12,8	12,0
sim	hedo	20	2,0	24	12,0	11,2
sim	none	20	25,0	24	12,7	11,8

k optimal selon la densité (H2)

densité	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
< 5	31,3	26,9	25,3	24,4	24,1	23,7	23,8	23,6	23,7
5-19	25,4	21,6	20,4	19,6	19,1	18,7	18,4	18,4	18,2
20-49	20,1	16,6	15,4	14,9	14,5	14,3	14,1	14,1	13,9
50-199	14,7	12,2	11,5	11,2	11,0	10,9	10,9	10,9	10,9
≥ 200	12,9	10,8	10,2	9,9	9,7	9,6	9,6	9,6	9,7

(sélection par distance, ≤ 10 km, 12 mois, grille hédonique; MdAPE test en %, par nombre de ventes disponibles à moins de 2 km au cours des 12 mois précédents)

Sans ajustement :

densité	k = 1	k = 3	k = 5	k = 7	k = 10	k = 15	k = 20	k = 30	k = 50
< 5	39,2	34,7	33,2	32,2	32,1	32,1	31,7	31,7	31,5
5-19	33,8	29,4	28,3	27,7	27,2	27,2	27,4	27,6	27,8
20-49	26,6	23,0	22,3	21,9	21,4	21,7	21,7	22,2	22,4
50-199	18,5	15,9	15,4	15,3	15,3	15,6	15,7	16,1	16,6
≥ 200	16,0	13,7	13,1	13,1	12,9	13,0	13,2	13,7	14,3

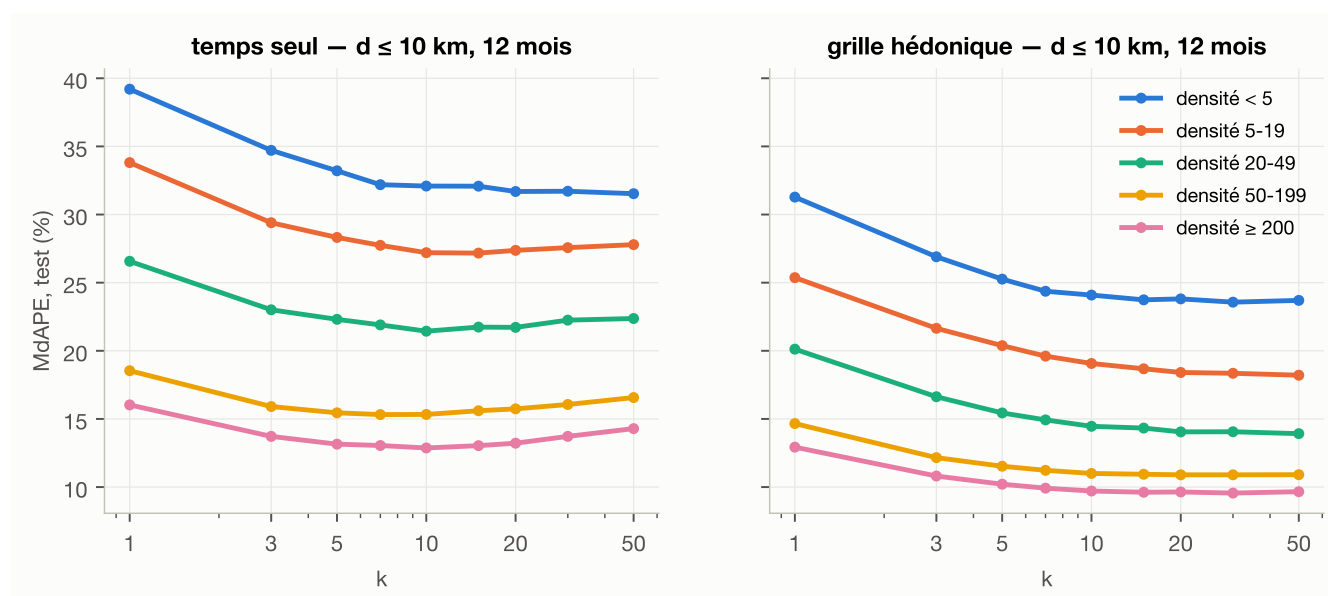
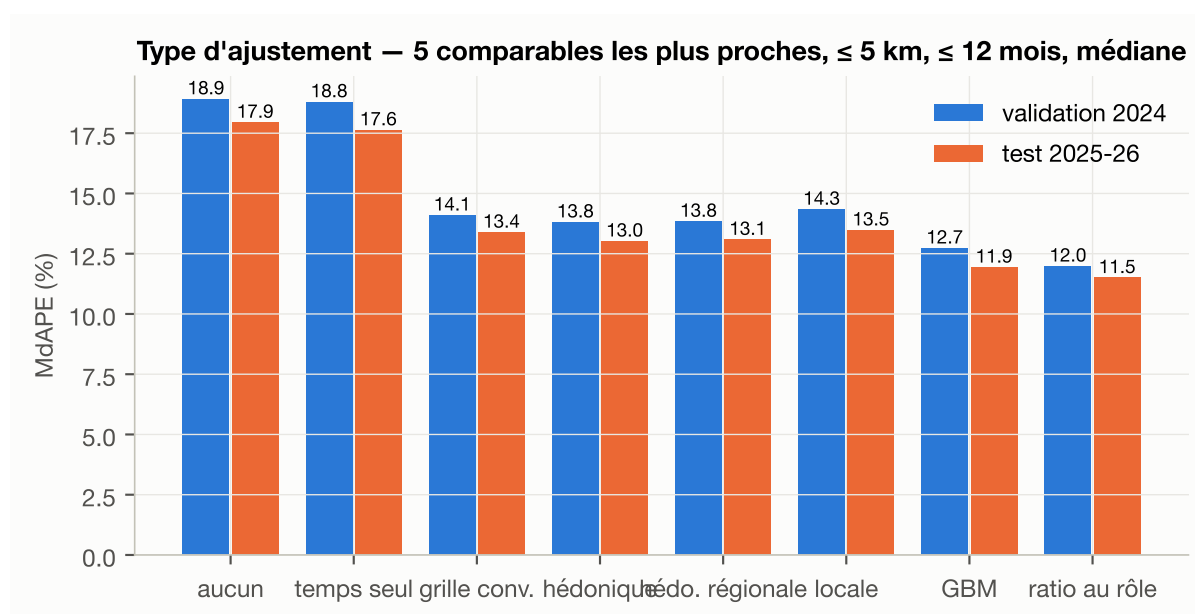


Fig. 4 : k optimal par strate de densité.

3.2 Quels ajustements sont justifiés ?

Fig. 5 : Type d'ajustement, cinq comparables les plus proches, ≤ 5 km, ≤ 12 mois, médiane.

Ajustement	MdAPE val.	MdAPE test	MAPE test	Biais	± 10 %	Couverture
aucun ajustement	18,9	17,9	30,4	-4,2	31,0	99,8
temps seulement	18,8	17,6	31,1	-0,8	31,7	99,8
grille conventionnelle	14,1	13,4	26,4	0,7	39,7	99,8
grille hédonique	13,8	13,0	25,1	-0,5	40,4	99,8
grille hédonique régionale	13,8	13,1	25,2	-0,5	40,4	99,8
régression locale	14,3	13,5	25,7	-0,4	39,4	99,8
ajustements GBM	12,7	11,9	24,1	-0,6	43,6	99,8
ratio au rôle	12,0	11,5	22,5	-0,1	40,6	90,3

Trois enseignements. **L'ajustement temporel seul** vaut 0,3 point; **n'importe quelle grille physique** vaut 4,2 à 5,7 points de plus; **les raffinements** (grille régionale, régression locale, ajustements non linéaires) apportent peu face à la grille hédonique globale, et la régression locale est même moins bonne (13,5 % contre 13,0 %).

Grilles locales et petits échantillons (H3)

MdAPE test (%) par type d'ajustement et densité, ancre (5, 5 km, 12 mois) :

densité	grille conventionnelle	grille hédonique	grille hédonique régionale	régression locale	ajustements GBM
< 5	27,1	25,3	25,3	25,9	24,5
5-19	21,4	20,4	20,1	20,8	19,7
20-49	16,2	15,4	15,3	15,9	14,6
50-199	11,7	11,5	11,5	11,9	10,4
≥ 200	10,4	10,2	10,3	10,7	9,2

Ancre (3, 2 km, 6 mois) :

densité	grille conventionnelle	grille hédonique	grille hédonique régionale	régression locale	ajustements GBM
< 5	30,0	28,6	28,6	29,5	28,0
5-19	23,8	22,4	21,9	22,8	21,4
20-49	17,7	16,9	16,6	17,4	15,5
50-199	12,8	12,5	12,6	12,9	11,1
≥ 200	11,1	10,8	11,0	11,4	9,8

3.3 Comment choisir les comparables ? (H1)

Ancre (5, 5 km, 12 mois), grille hédonique, médiane :

Sélection	MdAPE val.	MdAPE test	Biais	±10 %
les plus proches (distance)	13,8	13,0	-0,5	40,4
les plus semblables (physique)	13,3	12,5	-0,1	42,1
compromis distance / similarité	12,9	12,1	-0,1	43,0
les moins ajustés	13,1	12,2	-0,1	42,8

Test direct de H1 (k = 5, 12 mois, temps seulement) : les cinq voisins à moins de 1 km donnent une MdAPE de 17,2 % (couverture 96,0 %) ; les cinq ventes les plus semblables à moins de 10 km, 13,0 % (couverture 100,0 %) ; les cinq plus proches à moins de 10 km, 17,6 %. Avec la grille hédonique : plus proches 13,0 %, plus semblables 12,7 %, compromis 12,1 %.

3.4 Pondération et réconciliation (H4)

Ancre (5, 5 km, 12 mois), sélection par distance, grille hédonique :

Pondération	Réconciliation	MdAPE val.	MdAPE test	Biais	±10 %
égale	médiane	13,8	13,0	-0,5	40,4
égale	moyenne	14,6	13,7	-1,8	38,4
1/distance	médiane	13,9	13,1	-0,5	40,3
1/distance	moyenne	14,6	13,7	-1,8	38,6
récence	médiane	14,5	13,7	-0,5	38,8
récence	moyenne	14,8	13,9	-1,6	38,0
les plus semblables (physique)	médiane	13,8	12,9	-0,3	40,8
les plus semblables (physique)	moyenne	14,3	13,3	-1,5	39,7
distance × récence × similarité	médiane	14,6	13,6	-0,3	39,2
distance × récence × similarité	moyenne	14,5	13,5	-1,4	39,3

Selon la densité (MdAPE test, %) :

densité	égale / médiane	égale / moyenne	1/distance / moyenne	récence / moyenne	les plus semblables (physique) / moyenne	distance × récence × similarité / moyenne	distance × récence × similarité / médiane
< 5	25,3	25,1	25,0	25,9	25,2	25,8	28,1
5-19	20,4	20,6	20,5	21,0	20,3	20,6	21,8
20-49	15,4	16,0	16,1	16,4	16,0	16,2	16,7
50-199	11,5	12,3	12,3	12,5	11,9	12,0	12,0
≥ 200	10,2	10,9	10,9	11,0	10,5	10,6	10,5

Effets marginaux moyens (MdAPE test moyenne sur toutes les autres modalités de E2) : sélection les moins ajustés 13,1, les plus proches (distance) 14,4, compromis distance / similarité 12,9, les plus semblables (physique) 13,2 ; ajustement grille conventionnelle 13,2, ajustements GBM 12,6, grille hédonique 13,1, grille hédonique régionale 13,1, régression locale 13,3, temps seulement 14,7, aucun ajustement 15,4, ratio au rôle 11,8 ; pondération distance × récence × similarité 13,6, 1/distance 13,4, égale 13,2, récence 13,6, les plus semblables (physique) 13,1 ; réconciliation moyenne 13,6, médiane 13,2.

3.5 Compétition de modèles

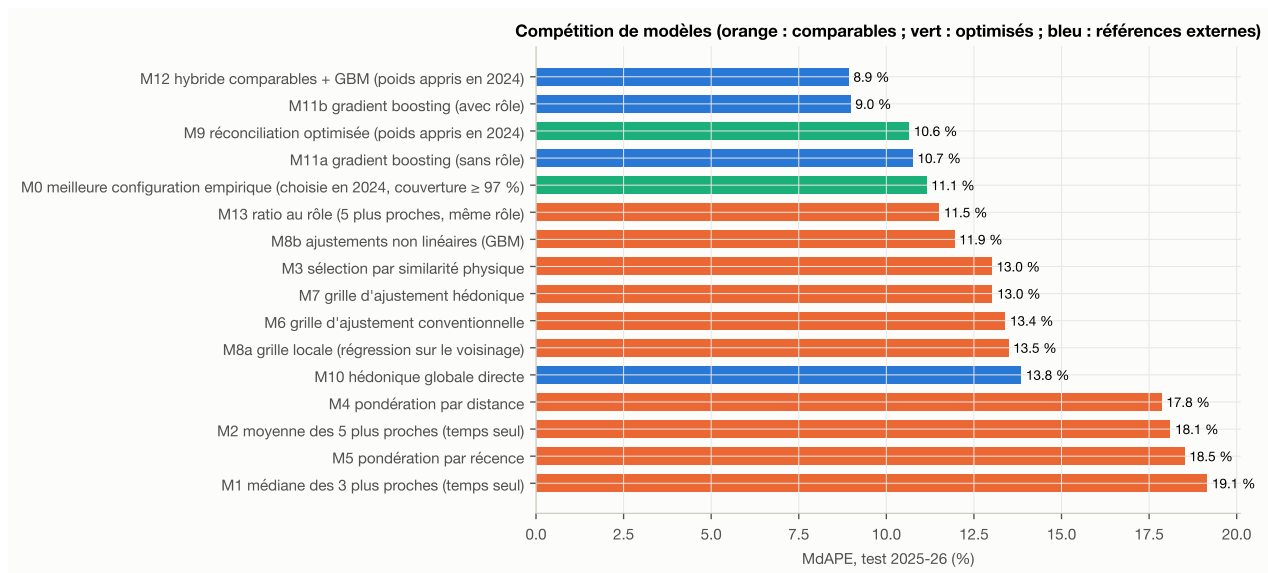


Fig. 6 : Compétition de modèles, test 2025-2026.

Modèle	MdAPE val.	MdAPE test	MAPE	Biais (médian)	±5 %	±10 %	±20 %	Couverture
M12 hybride comparables + GBM (poids appris en 2024)	9,7	8,9	19,9	-0,1	30,7	54,4	78,8	100,0
M11b gradient boosting (avec rôle)	9,8	9,0	19,9	0,1	30,5	54,1	78,7	100,0
M9 réconciliation optimisée (poids appris en 2024)	11,3	10,6	21,8	-0,8	26,2	47,6	73,0	99,8
M11a gradient boosting (sans rôle)	11,5	10,7	23,1	-0,4	26,1	47,5	72,2	100,0
M0 meilleure configuration empirique (choisie en 2024, couverture ≥ 97 %)	11,9	11,1	23,6	0,3	25,1	46,1	71,5	100,0

Modèle	MdAPE val.	MdAPE test	MAPE	Biais (médian)	±5 %	±10 %	±20 %	Couverture
M13 ratio au rôle (5 plus proches, même rôle)	12,0	11,5	22,5	-0,1	22,1	40,6	63,9	90,3
M8b ajustements non linéaires (GBM)	12,7	11,9	24,1	-0,6	23,9	43,6	68,8	99,8
M3 sélection par similarité physique	13,9	13,0	26,2	-0,2	22,1	40,8	65,3	99,8
M7 grille d'ajustement hédonique	13,8	13,0	25,1	-0,5	21,7	40,4	66,4	99,8
M6 grille d'ajustement conventionnelle	14,1	13,4	26,4	0,8	21,3	39,7	65,0	99,8
M8a grille locale (régression sur le voisinage)	14,3	13,5	25,7	-0,4	21,2	39,4	64,9	99,8
M10 hédonique globale directe	14,7	13,8	25,9	0,4	19,8	37,9	65,0	100,0
M4 pondération par distance	19,1	17,9	31,4	-0,2	16,0	30,7	54,3	99,8
M2 moyenne des 5 plus proches (temps seul)	19,3	18,1	31,8	-0,1	15,7	30,3	53,7	99,8
M5 pondération par récence	19,7	18,5	32,3	-0,2	15,4	29,8	52,8	99,8
M1 médiane des 3 plus proches (temps seul)	20,3	19,1	32,8	-0,8	15,2	28,5	49,8	96,6

M0 : compromis distance / similarité, $k = 30$, ≤ 25.0 km, 24 mois, grille hédonique, pondération égale, médiane (meilleure des 2 256 configurations en 2024 parmi celles qui couvrent ≥ 97 % des cibles ; sans cette contrainte, la meilleure est « compromis distance / similarité, $k = 10$, ≤ 10.0 km, 24 mois, ratio au rôle » à 10,7 % mais pour 90,3 % des cibles seulement). Poids appris de M9 : temps seulement 0,07, ajustements GBM 0,33, ratio au rôle 0,59. Poids de M12 : gbm_role 0,91, comps_gbm 0,03, comps_hedo 0,07.

Différences appariées (test ; bootstrap par grappes municipalité × trimestre)

Comparaison	Δ MdAPE (pts) [IC 95 %]	Δ MAPE (pts) [IC 95 %]	A plus précis (% des cibles)	n
M7 — M2	-5,08 [-5,24 ; -4,87]	-6,65 [-6,87 ; -6,43]	61,1	124 310
M7 — M6	-0,37 [-0,45 ; -0,29]	-1,32 [-1,41 ; -1,23]	52,3	124 310
M8a — M7	0,46 [0,33 ; 0,57]	0,59 [0,47 ; 0,73]	48,0	124 310
M8b — M7	-1,08 [-1,16 ; -1,00]	-1,04 [-1,11 ; -0,97]	54,3	124 310
M11b — M0	-2,17 [-2,33 ; -2,03]	-3,69 [-3,90 ; -3,52]	59,1	124 555
M11a — M0	-0,39 [-0,49 ; -0,33]	-0,50 [-0,57 ; -0,44]	51,8	124 555
M12 — M11b	-0,06 [-0,08 ; -0,03]	-0,07 [-0,09 ; -0,06]	52,6	124 559
M13 — M7	-1,40 [-1,56 ; -1,26]	-2,39 [-2,62 ; -2,17]	54,5	112 423
M3 — M2	-5,10 [-5,27 ; -4,95]	-5,58 [-5,79 ; -5,41]	59,8	124 310
M9 — M7	-2,38 [-2,50 ; -2,26]	-3,33 [-3,50 ; -3,18]	59,7	124 310
M1 — M2	1,41 [1,25 ; 1,61]	1,82 [1,68 ; 1,97]	45,8	120 372

Un Δ négatif signifie que A est plus précis que B ; l'intervalle qui exclut zéro signale une différence robuste aux dépendances spatiales et temporelles.

3.6 Où la méthode tient-elle, où défaille-t-elle ?

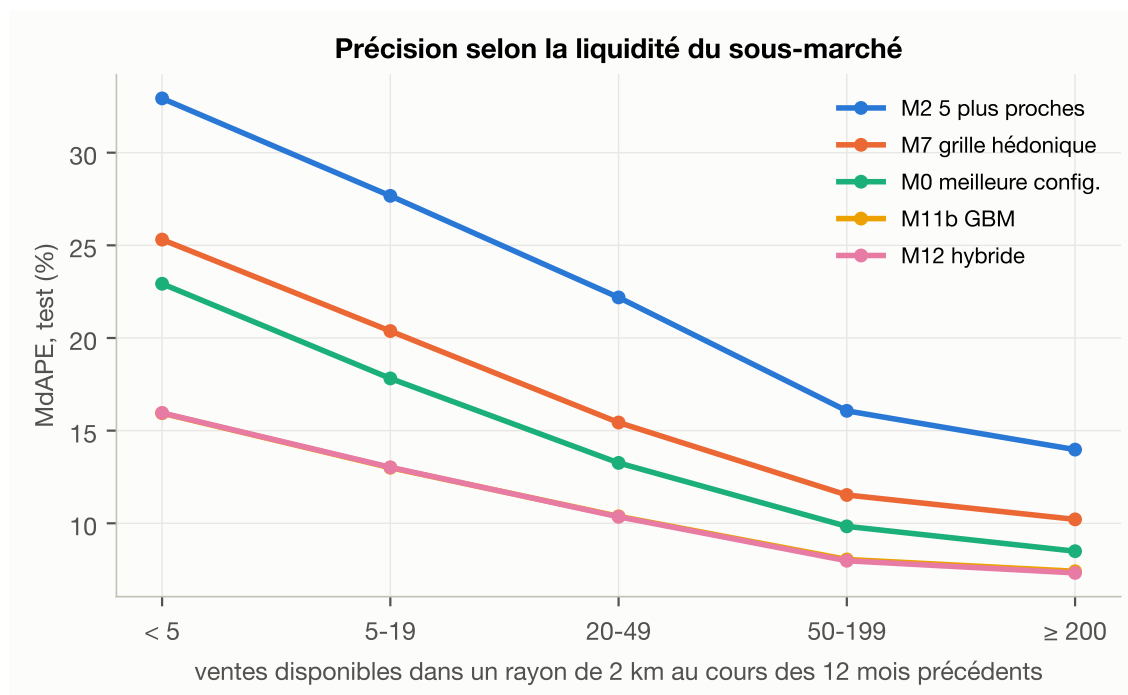


Fig. 7 : Précision selon la liquidité du sous-marché.

Par densité (ventes à moins de 2 km dans les 12 mois précédents) — MdAPE test (%) :

densité	n	M2	M7	M8a	M13	M0	M10	M11b	M12
< 5	9 177	32,9	25,3	25,9	19,6	22,9	23,6	15,9	16,0
5-19	15 149	27,7	20,4	20,8	16,5	17,8	18,9	13,0	13,0
20-49	15 553	22,2	15,4	15,9	13,4	13,3	15,1	10,4	10,4
50-199	56 442	16,1	11,5	11,9	10,4	9,8	12,9	8,0	8,0
≥ 200	27 989	14,0	10,2	10,7	9,3	8,5	11,3	7,4	7,3

Par région administrative :

région	n	M2	M7	M8a	M13	M0	M10	M11b	M12
Montréal	25 133	16,1	11,3	11,9	10,5	9,5	11,7	8,0	7,9
Laurentides	13 629	18,0	12,8	12,8	11,4	10,6	12,9	8,5	8,5
Capitale-Nationale	11 711	18,7	13,4	14,2	11,4	11,6	13,8	9,1	9,1
Lanaudière	11 377	17,2	12,9	12,6	11,6	10,6	13,0	8,6	8,6
Montréal	9 496	16,0	12,0	13,5	11,2	10,7	21,4	9,2	9,1
Estrie	9 437	21,0	14,6	15,2	12,7	12,5	14,7	9,7	9,6
Chaudière-Appalaches	8 135	22,0	15,5	16,6	13,4	13,6	14,7	10,4	10,4
Outaouais	7 426	14,3	10,6	9,8	9,4	9,1	13,0	7,6	7,5
Laval	5 243	14,5	9,6	10,3	8,1	8,1	10,6	6,9	7,0
Saguenay-Lac-Saint-Jean	5 030	20,7	16,0	16,0	12,8	14,2	15,6	10,2	10,2
Centre-du-Québec	4 681	20,5	13,7	14,0	12,4	12,1	13,1	9,6	9,6
Mauricie	4 670	21,6	16,3	16,7	14,3	14,2	16,4	11,4	11,3
Bas-Saint-Laurent	3 260	23,3	18,2	18,7	15,4	15,9	16,5	12,3	12,2
Abitibi-Témiscamingue	2 412	22,2	17,0	16,6	12,4	15,3	17,1	9,8	10,0
Côte-Nord	1 227	23,9	19,3	19,5	16,9	17,7	19,1	13,7	13,6
Gaspésie-Îles-de-la-Madeleine	1 157	29,0	24,0	23,5	22,6	21,5	22,9	17,1	17,4
Nord-du-Québec	286	27,1	21,8	22,4	15,7	20,7	22,0	13,9	13,6

Par quintile de prix :

quintile de prix	n	M2	M7	M8a	M13	M0	M10	M11b	M12
Q1 (bas)	21 659	50,4	36,7	37,0	28,7	35,8	36,7	26,4	26,2
Q2	24 680	17,5	13,7	14,2	12,7	11,7	13,5	9,4	9,4
Q3	23 834	13,6	10,4	10,6	9,7	9,0	11,5	7,1	7,1
Q4	26 783	13,3	10,0	10,1	8,9	8,2	11,3	6,7	6,7
Q5 (haut)	27 354	19,3	11,6	12,4	9,7	10,1	13,2	8,2	8,2

Par âge du bâtiment :

âge	n	M2	M7	M8a	M13	M0	M10	M11b	M12
16-40	39 475	16,6	11,1	11,3	9,8	9,2	11,9	7,5	7,4
4-15	17 344	17,0	9,9	10,1	8,8	8,4	12,1	6,7	6,6
41-70	48 592	17,7	14,4	14,9	13,0	12,5	14,5	10,2	10,2
> 70	16 526	25,1	19,7	21,6	16,9	18,0	20,9	14,2	14,2
≤ 3 ans	2 373	20,1	11,9	12,1	9,8	9,8	12,9	8,0	7,8

Par lien physique (0 détaché, 1 jumelé, 2 en rangée, 3 intégré) :

lien physique	n	M2	M7	M8a	M13	M0	M10	M11b	M12
0,0	100 999	19,6	14,3	14,7	12,4	12,2	14,2	9,8	9,7
1,0	15 936	12,9	9,1	9,4	8,6	7,7	11,8	6,6	6,4
2,0	7 297	11,9	8,2	8,6	8,0	7,8	14,5	6,2	6,1
3,0	78	34,3	22,3	18,0	10,3	15,3	31,0	10,9	9,9

Marchés homogènes et hétérogènes (H5)

Terciles de la dispersion (écart-type du log des prix ajustés) des dix ventes les plus proches à ≤ 10 km / 24 mois — MdAPE test (%) :

sous-marché	σ médian (log)	M2	M7	M8a	M13	M0	M10	M11b	M12
homogène	0,121	13,9	10,2	10,5	9,2	9,0	12,1	7,5	7,4
moyen	0,252	19,0	13,9	14,3	12,1	11,9	14,3	9,5	9,5
hétérogène	0,400	22,8	16,4	16,9	14,3	13,6	15,7	10,4	10,4

Stabilité dans le temps

trimestre	n	M2	M7	M0	M11b	M12	biais M7	biais M11b
2025Q1	11 337	19,5	14,2	12,5	10,3	10,1	-2,4	3,0
2025Q2	24 823	17,3	12,4	10,5	8,5	8,4	-1,1	-0,3
2025Q3	20 140	18,2	13,0	11,1	8,9	8,8	1,5	-0,5
2025Q4	16 692	19,2	13,9	12,0	9,5	9,6	-0,7	-0,1
2026Q1	10 619	19,6	14,2	12,7	10,0	10,0	-2,4	-0,1
2026Q2	22 825	16,8	12,1	10,1	8,3	8,2	-1,3	0,4
2026Q3	18 123	18,3	12,9	11,1	8,8	8,8	1,7	-0,3

Le biais trimestriel commun à toutes les méthodes reflète le fait que le niveau de prix du mois courant est inconnu (toutes travaillent au niveau du mois précédent) et la saisonnalité des prix.

3.7 Intervalles de valeur

La dispersion des prix ajustés des comparables retenus fournit un intervalle de valeur — mais un intervalle large : calibré pour couvrir 80 % des prix, il s'étend sur environ ± 30 % de la valeur (largeur médiane ci-dessous). Avec $z = 1,28$ (couverture nominale de 80 % si les erreurs étaient normales), puis avec un z calibré sur la validation 2024 :

configuration	z calibré (val.)	couverture à z = 1,28	couverture calibrée	largeur médiane (%)	n
temps seulement, k = 5, ≤ 5 km, 12 mois	1,89	67,2	80,3	90,5	123 913
temps seulement, k = 10, ≤ 10 km, 24 mois	1,41	77,3	80,4	85,2	124 516
grille hédonique, k = 5, ≤ 5 km, 12 mois	1,88	67,9	79,9	65,0	123 913
grille hédonique, k = 10, ≤ 10 km, 24 mois	1,35	78,5	80,0	66,5	124 516
ajustements GBM, k = 5, ≤ 5 km, 12 mois	2,03	66,5	80,2	62,9	123 913
ajustements GBM, k = 10, ≤ 10 km, 24 mois	1,44	77,3	80,5	66,7	124 516

3.8 Vers un système de recommandation des comparables

Règle choisie en 2024 pour chaque strate de densité (parmi les 2 256 configurations couvrant ≥ 95 % des cibles de la strate), évaluée en 2025-2026 :

densité	règle	MdAPE val.	MdAPE test	couverture val.	n test
< 5	compromis distance / similarité, k = 10, ≤ 10.0 km, 24 mois, grille hédonique, pond. égale, moyenne	24,7	23,7	99,8	9 426
5-19	compromis distance / similarité, k = 30, ≤ 10.0 km, 24 mois, grille hédonique, pond. égale, médiane	18,3	17,8	100,0	15 149
20-49	compromis distance / similarité, k = 30, ≤ 10.0 km, 24 mois, grille hédonique, pond. égale, médiane	13,9	13,3	100,0	15 553
50-199	compromis distance / similarité, k = 10, ≤ 10.0 km, 24 mois, ratio au rôle, pond. égale, médiane	9,8	9,7	96,5	56 442
≥ 200	les moins ajustés, k = 10, ≤ 10.0 km, 24 mois, ratio au rôle, pond. égale, médiane	8,9	8,8	95,8	27 989

Globalement, ces règles conditionnelles atteignent une MdAPE de **11,3 %** (42,5 % des estimations à ± 10 %, couverture 93,4 %), contre 11,1 % à couverture complète pour la meilleure règle unique : elles ne font pas mieux. La meilleure règle unique (les 30 ventes les plus semblables à ≤ 25.0 km et 24 mois) s'adapte d'elle-même : elle utilise les ventes voisines là où elles existent et va plus loin seulement là où elles manquent. Ce que la stratification apporte, c'est le choix de l'ajustement (ratio au rôle dans les marchés denses où les comparables du même rôle abondent), pas des nombres ou des rayons propres à chaque strate.

4. Discussion

Ce que la méthode des comparables fait bien. Correctement paramétrée — comparables semblables, nombre d'une dizaine, rayon et fenêtre larges, ajustements hédoniques ou même conventionnels — elle atteint une précision (11,1 %) que ne dépasse le meilleur modèle d'apprentissage ré-entraîné chaque trimestre que de 2,2 points, tout en restant intégralement explicable : chaque valeur est la médiane de quelques ventes identifiées, ajustées par des coefficients lisibles.

Ce qui contredit la pratique. Trois comparables, un rayon de quelques centaines de mètres et une fenêtre de trois mois — les conventions les plus répandues — sont dominés sur toute la surface Erreur(k, d, t) : 13,6 % de MdAPE (k = 3, ≤ 1 km, 3 mois, grille hédonique, couverture 89,2 %) contre 11,1 % pour la meilleure configuration. La rareté des comparables « parfaits » est un argument pour élargir la recherche et ajuster, non pour se restreindre.

Le rôle d'évaluation comme information. Le ratio au rôle (M13) et le GBM avec rôle (M11b) dominent leurs équivalents sans rôle : la valeur au rôle encapsule des caractéristiques non observées ici (qualité, dépendances, état) et une appréciation locale que les modèles globaux ne captent pas. Les modèles globaux estimés par moindres carrés présentaient d'ailleurs un biais médian de -5 % tant qu'ils visaient la moyenne du log (distribution asymétrique) ; ils ont été ré-estimés en régression médiane (L1) ou recentrés, et ne sont plus biaisés.

Limites. (1) Règles professionnelles simulées, non dossiers d'évaluateurs. (2) Caractéristiques limitées à celles du rôle et du registre. (3) Délai de publication des ventes ignoré. (4) Un seul type de bien (unifamilial); condos et plex restent à traiter. (5) Les configurations ont été choisies sur 2024 puis évaluées sur 2025-2026 : les écarts entre validation et test (tableaux ci-dessus) mesurent la stabilité des règles, et ils sont faibles.

5. Reproductibilité

Code : `pipeline/01_dataset.py` → `06_report.py` (Python 3.14, pandas, numpy, scipy, lightgbm, matplotlib); sorties versionnées `out/2026-10-10/` (configurations `configs_E1.csv`, `configs_E2.csv`; estimations `est_E1.npy`, `est_E2.npy`; cibles `targets.parquet`; tables `tables/*.csv`; `results.json`). Exécution : `./run_all.sh` sur le nœud de calcul (M2U64), `./make_pdf.sh` sur le poste de travail.