



**MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE**

*Liberté
Égalité
Fraternité*



**Conseil général de l'Environnement
et du Développement durable**

Etude économétrique sur les facteurs de disparités géographiques de la ponctualité des trains de voyageurs en France sur la période 2012-2019

Alexandre BARBUSSE

Février 2022

Contexte et objectifs

- ❑ Les données du taux de retard des trains de voyageurs en France par activité ferroviaire de la base de l'AQST montrent des disparités géographiques sur la période 2012-2019
- ❑ Afin de savoir si cette variabilité géographique peut s'expliquer par des différences de certaines variables influençant la ponctualité : constitution de bases de données par activité ferroviaire à partir de données publiques
- ❑ Identification des variables dont l'impact sur le taux de retard est statistiquement significatif parmi l'ensemble des variables des bases de données
- ❑ Ensuite, implémentation de modèles économétriques simples pour décrire et interpréter la corrélation entre la ponctualité des trains par activité et les principales variables identifiées permettant d'expliquer pour partie sa variabilité géographique sur la période 2012-2019

Méthodologie

Données utilisées : ponctualité des trains sur la période 2012-2019

- ☐ Taux de retard annuel des trains de voyageurs en France par activité : TGV, Intercités, TER, RER et Transilien en Ile-de-France

Service	Période	Seuil de retard en minutes	Granularité géographique
TGV	2012-2019	Composite 5-10-15 min	Par liaisons ferroviaires
Intercités	2012-2019	Composite 5-10-15 min	Par liaisons ferroviaires
TER	2013-2019	5 min	Par régions administratives*
RER et Transilien d'Ile-de-France	2013-2019	Ponctualité voyageurs à 5 min (indicateur Ile-de-France Mobilités)	Par lignes
RER et Transilien d'Ile-de-France	2013-2018	Ponctualité voyageurs à 5 min (indicateur Ile-de-France Mobilités)	Par branches (voyageurs descendant à l'une des gares de la branche)

*Pour disposer d'une information plus détaillée, le découpage des régions administratives d'avant la réforme de 2016 a été utilisé. Le taux de retard des TER par ancienne région administrative pour les années 2017-2019 a été estimée sur la base des données de la période 2012-2016.

- ☐ Données de 2020 non intégrées à l'étude => année atypique du fait des conséquences de l'épidémie de Covid-19 et des périodes de confinement national

Données utilisées : variables explicatives

- ❑ Utilisation de données et documents publics pour constituer les bases de données des variables pouvant avoir un effet sur la ponctualité :
 - OpenData et sites Internet des autorités organisatrices (AOM) et des opérateurs de transport
 - Rapports d’audit et d’études technique
 - Schémas Directeurs et présentations par les AOM et les transporteurs lors des comités de ligne
 - Fiches horaires, etc...

- ❑ Détail des sources de données utilisées :

Activité	Sources des données utilisées
TGV et Intercités	Base des données de ponctualité de l'AQST
	Mémento statistique de la SNCF - 2018
	Rome2Rio
	Base des données de l'ART sur les services ferroviaires de voyageurs (2015-2018)
	https://www.horaires-de-trains.fr/ d'après l'OpenData SNCF
	Atlas SNCF Réseau 2020
	Carte SNCF - TRAFIC TOTAL 2000 SNCF / Direction du Contrôle de gestion (2000)

Données utilisées : variables explicatives

Activité	Sources des données utilisées
TER	Base ART ferroviaire (2015-2018)
	ART - Bilan du marché ferroviaire de voyageurs 2015/2016, 2017, 2018 volumes 1 et 2
	Atlas SNCF Réseau 2020
	https://www.horaires-de-trains.fr/ d'après l'OpenData SNCF
	Chambres régionales des Comptes, rapports d'observations définitives, Enquêtes Transports Express Régionaux, Exercices 2012 et suivants
	Fiches Wikipédia sur les TER par anciennes régions, et sur le parc de matériel roulant TER par anciennes régions
Transilien	OpenData SNCF
	Sites officiels des lignes RER et Transilien
	Présentations d'Ile-de-France Mobilités, de la RATP ou de la SNCF lors des Comités de ligne RER et Transilien, sites officiels des lignes
	Documentation technique de la SNCF et de la RATP
	Plans des lignes
	Fiches horaires des lignes
	OMNIL
	Fiches Wikipédia sur les lignes de Transilien et sur les différentes séries du matériel roulant composant leurs parcs
	Articles de la presse spécialisée sur les mises en service et radiations du matériel roulant
	Rapport de l'Audit stratégique de la performance du système, Panos Tzieropoulos, Yves Putallaz, EPFL, 2015
Normes de tracé horaire sur les lignes nationales pour le SA 2021, Document opérationnel, SNCF Réseau, 2019	

Données utilisées : variables explicatives

- Pour constituer la base de données de l'âge du matériel roulant des TER et RER et Transilien, des modèles de parcs ont été construits à partir des données suivantes :
 - Chambres Régionales des Comptes, rapports d'observations définitives, Enquêtes Transports Express Régionaux, Exercices 2012 et suivants
 - Fiches Wikipédia par ancienne région sur les TER et le parc de matériel roulant TER, et par ligne sur le parc de matériel roulant RER et Transilien
 - Articles de la presse spécialisée portant sur les mouvements de parc de matériel roulant (mises en service et radiations)

Données utilisées : variables explicatives

□ Pour constituer la base de données de l'âge du matériel roulant des TER et RER et Transilien, des modèles de parcs ont été construits :

- Exemple = Estimation des effectifs du matériel roulant TER Aquitaine en fin d'année

Séries	31/12/2013	31/12/2014	31/12/2015	31/12/2016	31/12/2017	31/12/2018	31/12/2019
B 81500	27	27	27	27	27	27	27
B 84500					0	3	5
X 72500	17	17	17	17	17	18	18
X 73500	18	18	20	20	20	20	20
Z 51500		10	14	22	22	22	22
Z 56300		2	9	16	24	24	24
X 2200	23	23	14	8	0		
Z 7300	22	22	15	8	0		
BB 7200	6	6	6	5	5	3	2
Z 21500	6	0					
TOTAL	119	125	122	123	115	117	118

Forme des modèles statistiques

☐ Modèles statistiques testés à l'aide du logiciel R : par variable explicative

- Régressions linéaires simples : $y_{l,t} = a_0 + b \times x_{l,t} + \varepsilon_{l,t}$ (liaison l année t)

☐ Modèles statistiques testés par activité ferroviaire

- Régressions linéaires multiples : $y_{l,t} = a_0 + \sum_{i=1}^n a_i \times x_{i,l,t} + \varepsilon_{l,t}$ (liaison l année t variable i)
- Modèles exponentiels : $y_{l,t} = A \times \prod_{i=1}^n e^{a_i \times x_{i,l,t}} + \varepsilon_{l,t}$ (liaison l année t variable i)
- Modèles log-linéaires : $y_{l,t} = a_0 + \sum_{i=1}^n a_i \times \log(x_{i,l,t}) + \varepsilon_{l,t}$ (liaison l année t variable i)
- Modèles log-log : $\log(y_{l,t}) = a_0 + \sum_{i=1}^n a_i \times \log(x_{i,l,t}) + \varepsilon_{l,t}$ (liaison l année t variable i)

Résultats

Éléments de statistique

□ Test de Student pour décrire la significativité de l'effet d'une variable explicative de la ponctualité :

▪ Hypothèse nulle $H_0 : a_i = 0$

▪ Hypothèse alternative $H_1 : a_i \neq 0$

Avec a_i le coefficient associé à la i -ème variable explicative de la ponctualité

- La p-value associée au test de Student représente la probabilité de conclure que l'effet de la variable explicative sur la ponctualité est significatif alors qu'il ne l'est pas en réalité
- Plus la p-value est faible, plus la significativité statistique de l'effet est grande

□ En pratique, on compare la p-value avec un seuil de significativité (généralement 5%)

⇒ Si la p-value est inférieure au seuil, on peut considérer que l'effet de la variable explicative sur la ponctualité est statistiquement significatif.

Dans le cas contraire, on ne peut pas conclure avec certitude à l'existence d'un effet (même très faible) de la variable explicative

Éléments de statistique

□ Pour caractériser la pertinence du modèle et savoir s'il décrit correctement la variabilité géographique du taux de retard des trains par service sur la période 2012-2019, on utilise :

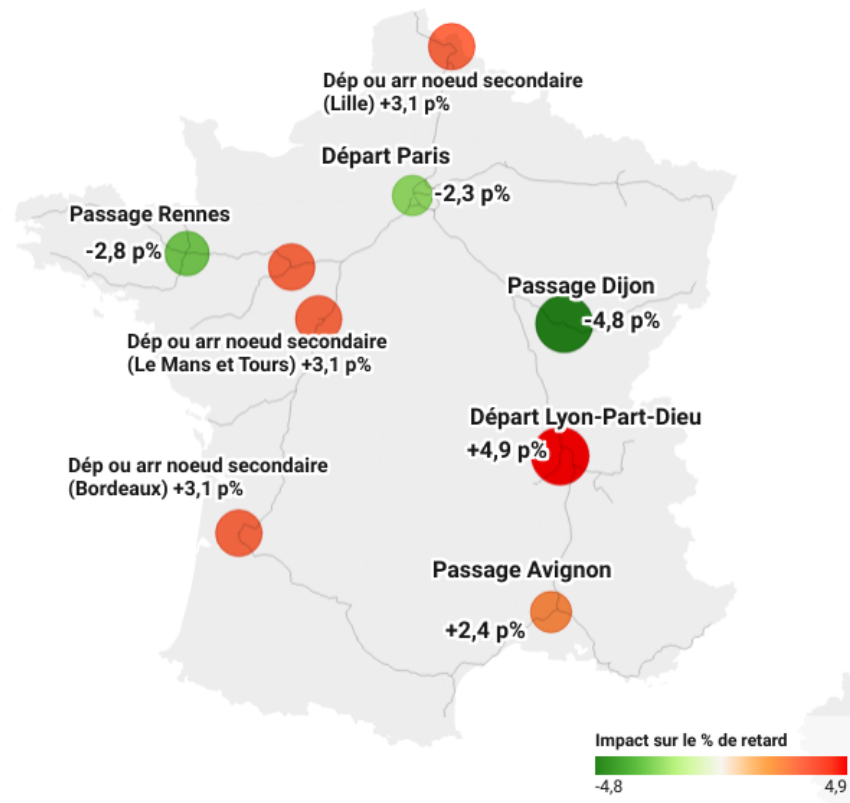
1. Le coefficient de détermination ajusté « R^2 ajusté » : quantifie la part du taux de retard expliquée par le modèle
2. Le test de Fisher qui permet de décrire la significativité globale du modèle
 - La p-value associée au test de Fisher représente la probabilité de conclure que le modèle décrit correctement la variabilité géographique du taux de retard alors que ce n'est pas le cas en réalité
 - Plus la p-value est faible, meilleur est le modèle
 - En pratique, on compare la p-value avec un seuil de significativité (ici 0,1%) :
 - ⇒ Si la p-value est inférieure au seuil, on peut considérer que le modèle retenu explique correctement les disparités géographiques observées du taux de retard sur la période 2012-2019.

Dans le cas contraire, on ne peut pas conclure avec certitude quant à la pertinence du modèle.

TGV : modèle retenu

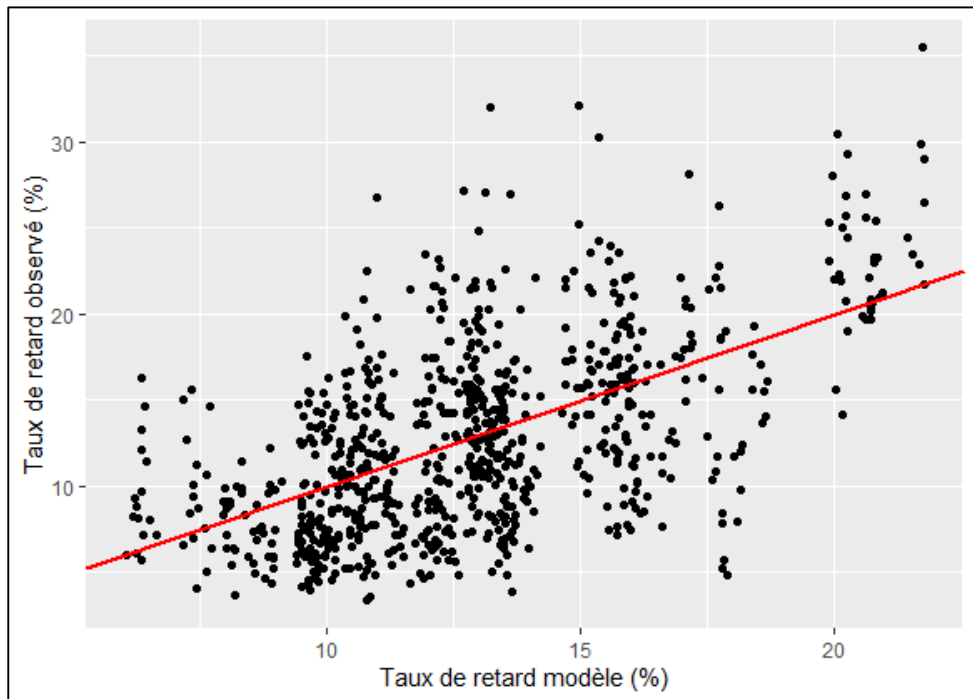
Activité ou service ferroviaire	Granularité des données utilisées	Période des données utilisées	R ² ajusté régression (p-value test de Fisher)	Variables explicatives	p-value test de Student variables explicatives	Impact sur le taux de retard
TGV	Par liaisons ferroviaires	2012-2019	0,32 ($< 0,001$)	La liaison part de Paris	$< 0,1\%$	-2,3 pts de retard
				La liaison part de Lyon-Part-Dieu	$< 0,1\%$	+4,9 pts
				La liaison traverse le nœud ferroviaire d'Avignon	$< 0,1\%$	+2,4 pts
				La liaison traverse Rennes	$< 0,1\%$	-2,8 pts
				La liaison traverse Dijon	$< 0,1\%$	-4,8 pts
				La liaison part ou arrive au niveau d'un nœud ferroviaire secondaire du réseau TGV	$< 0,1\%$	+3,1 pts
				Nombre moyen d'arrêts effectués par les TGV sur la liaison	$< 1\%$	+0,3 pts / arrêt
Trafic moyen journalier annuel (TMJA) des TGV sur la section la plus intensément utilisée de la liaison en 2019	$< 2\%$	+0,5 pts / 100 circulations TGV/jour				

Estimation économétrique de l'effet sur la ponctualité du passage des TGV par certains nœuds ferroviaires



Carte: Alexandre Barbusse/AQST • Créé avec Datawrapper

Modèle TGV : Taux de retard modèle vs taux de retard observé



Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

En moyenne, le taux de retard observé augmente lorsque le taux de retard du modèle augmente

Même si de nombreux points s'écartent fortement de la droite, le modèle passe le test de Fisher au seuil 0,1%.

Intercités : modèle retenu

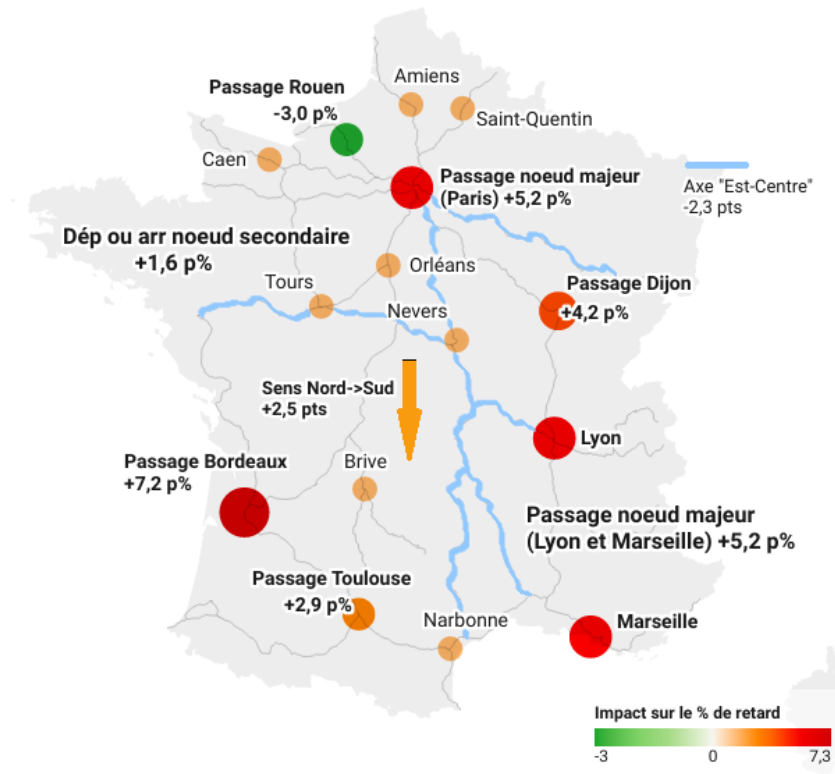
Activité ou service ferroviaire	Granularité des données utilisées	Période des données utilisées	R ² ajusté régression (p-value test de Fisher)	Variables explicatives	p-value test de Student variables explicatives	Impact sur le taux de retard
Intercités	Par liaisons ferroviaires	2012-2019	0,33 (< 0,001)	Liaison de nuit	< 0,1%	-3,9 pts de retard
				Nombre de nœuds majeurs (Paris, Lyon, Marseille) traversés sur la liaison	< 0,1%	+5,2 pts / nœud majeur traversé
				La liaison traverse le nœud ferroviaire de Bordeaux	< 0,1%	+7,3 pts
				La liaison traverse le nœud ferroviaire de Toulouse Matabiau	< 0,1%	+2,9 pts
				La liaison traverse le nœud ferroviaire de Dijon	< 2%	+4,2 pts
				La liaison traverse la gare de Rouen	< 1%	-3,0 pts
				La liaison part ou arrive au niveau d'un nœud ferroviaire secondaire du réseau classique	< 1%	+1,6 pts
				La liaison appartient à l'axe "Est – Centre" du réseau Intercités de 2016	< 1%	-2,3 pts
Le trajet sur la liaison s'effectue en sens "Nord -> Sud"	< 0,1%	+2,5 pts				

Estimation économétrique de l'effet sur la ponctualité du passage des trains Intercités par certains nœuds ferroviaires*

*Remarque :

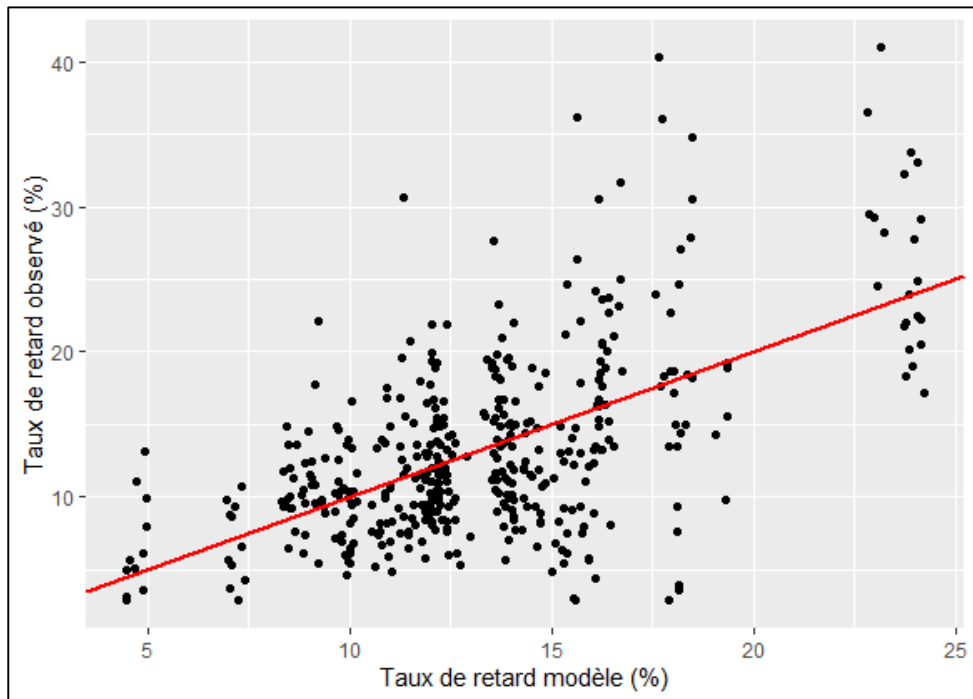
L'étude portant sur les données de la période 2012-2019, de nombreuses liaisons Intercités sont désormais dans le périmètre des services régionaux TER.

Le diagnostic de l'effet sur la ponctualité du passage par certains nœuds ferroviaires reste toutefois instructif malgré ces changements de périmètres.



Carte: Alexandre Barbusse/AQST · Créé avec Datawrapper

Modèle Intercités : Taux de retard modèle vs taux de retard observé



Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

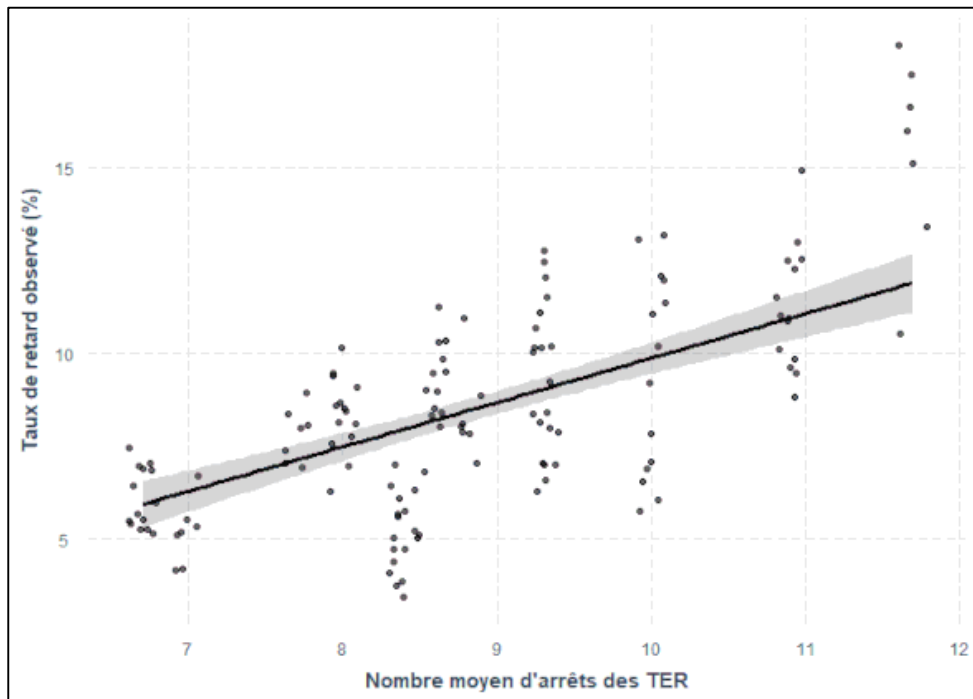
En moyenne, le taux de retard observé augmente lorsque le taux de retard du modèle augmente

Même si de nombreux points s'écartent fortement de la droite, le modèle passe le test de Fisher au seuil 0,1%.

TER : modèle retenu

Activité ou service ferroviaire	Granularité des données utilisées	Période des données utilisées	R ² ajusté régression (p-value test de Fisher)	Variables explicatives	p-value test de Student variables explicatives	Impact sur le taux de retard
TER (trains régionaux de province)	Par anciennes régions administratives	2013-2019	0,62 (< 0,001)	Nombre moyen d'arrêts par région en 2018 (nouvelles régions administratives)	< 0,1%	+1,2 pts / arrêt
				Distance moyenne parcourue par les trains TER	< 2%	+0,25 pts / 10 km
				Nombre de voyageurs-km transportés (en millions) dans l'année	< 0,1%	+1,2 pts / milliard voy-km
				Densité des « petites lignes » (catégories UIC 7 à 9 AV)	< 1%	-0,46 pt/ 1km de ligne/100 km ² de superficie

Modèle TER : Effet du nombre d'arrêts effectués par les TER (+1,2 pts de retard / arrêt)

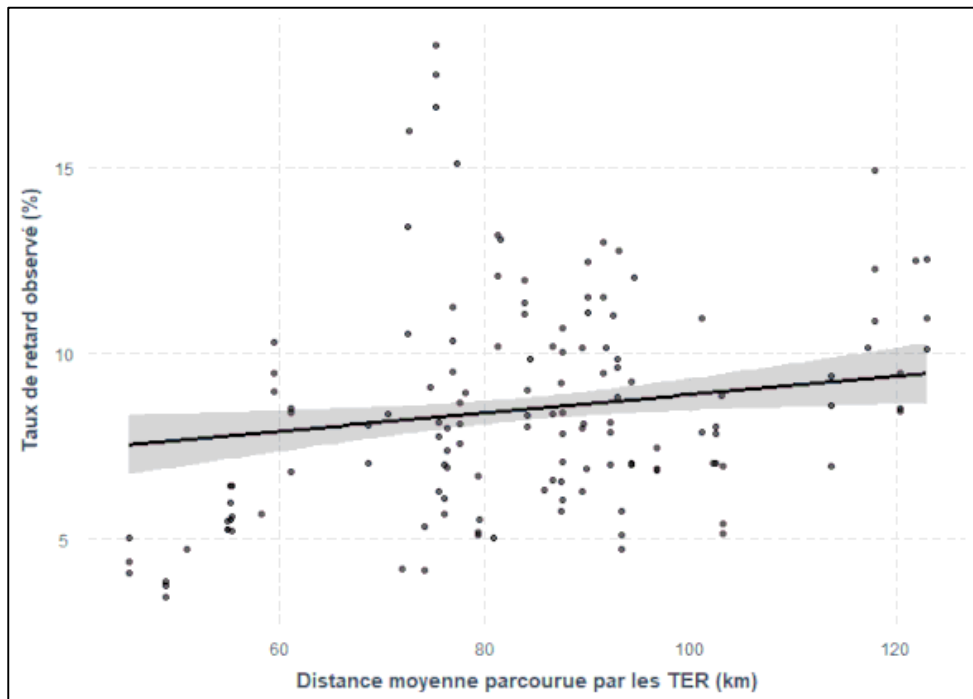


Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

On remarque que les distributions régionales (anciennes régions) observées présentent des moyennes qui s'ajustent plutôt bien avec la droite de régression linéaire, ce qui confirme la significativité élevée de l'effet d'après le test de Student

Modèle TER : Effet de la distance moyenne parcourue (+0,25 pts / 10 km)

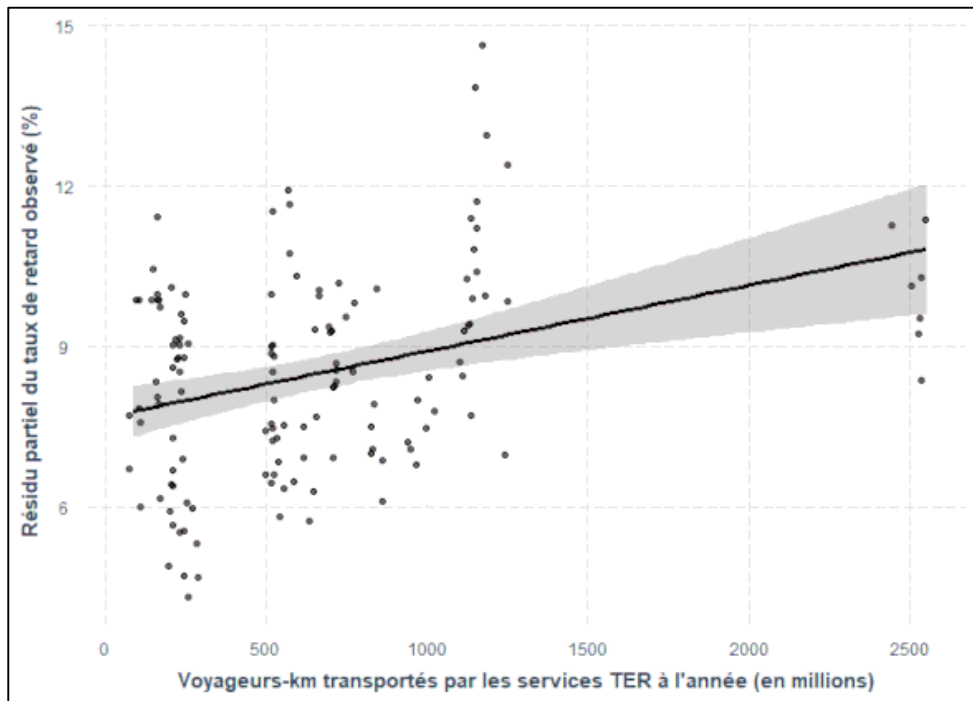


Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

Forte dispersion du taux de retard observé autour de la droite de régression à distance parcourue fixée
⇒ explique pourquoi l'effet de cette variable ne passe pas le test de Student au seuil 1% mais seulement au seuil 5%

Modèle TER : Effet du volume de voyageurs-km transportés (+1,2 pts / milliard voy-km)



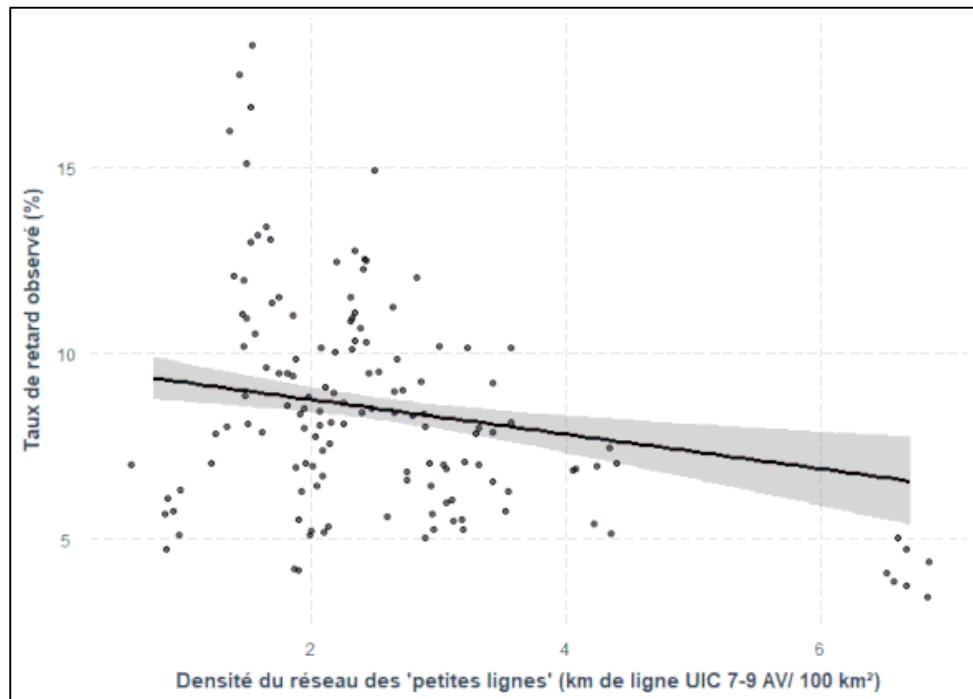
Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

⇒ Incertitude importante pour les valeurs les plus élevées du volume de voy-km transportés (ancienne région Rhône-Alpes)

Néanmoins, pour chaque « paquet » observé au sein de la distribution, la moyenne du résidu s'ajuste relativement bien au modèle linéaire

Modèle TER : effet de la densité des « petites lignes » (-0,46 pt/ 1km de ligne/100 km² de superficie)



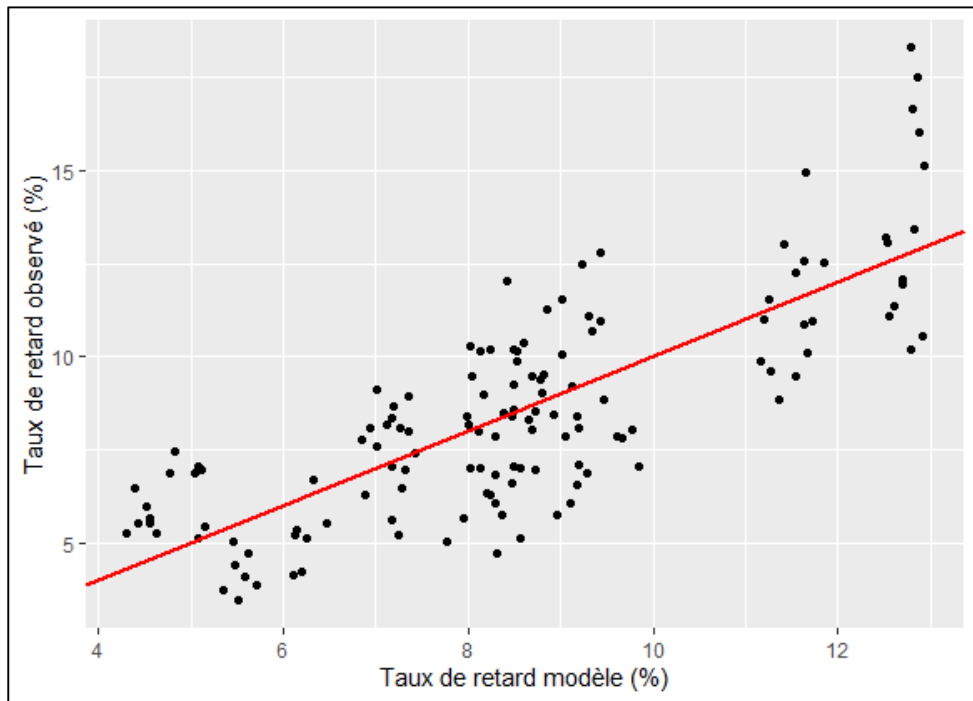
Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

Forte dispersion du taux de retard observé autour de la droite de régression à densité des « petites lignes » fixée.

Néanmoins, l'effet est significatif au seuil 1% selon le test de Student

Modèle TER : Taux de retard modèle vs taux de retard observé



Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

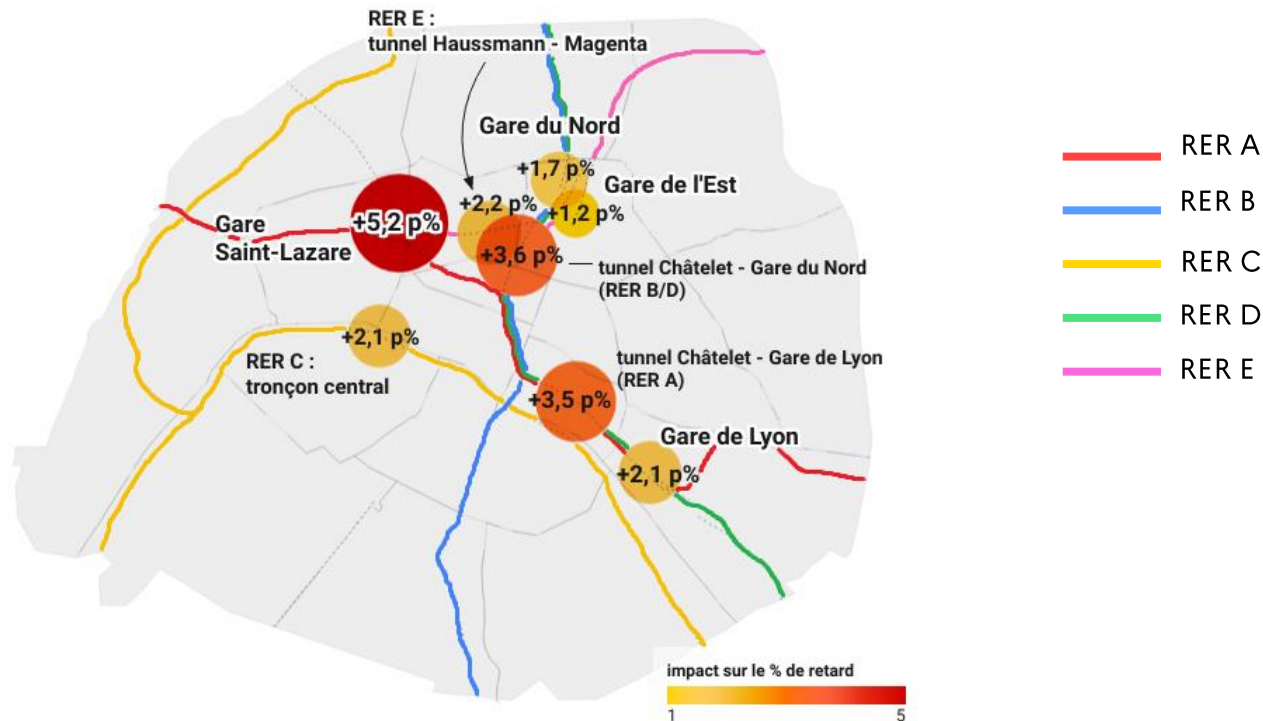
Les points s'alignent plutôt bien sur la droite en comparaison des modèles retenus pour les TGV et Intercités

⇒ illustre la meilleure capacité du modèle retenu pour les TER à expliquer la variabilité géographique et temporelle du taux de retard

RER et Transilien par ligne : modèle retenu

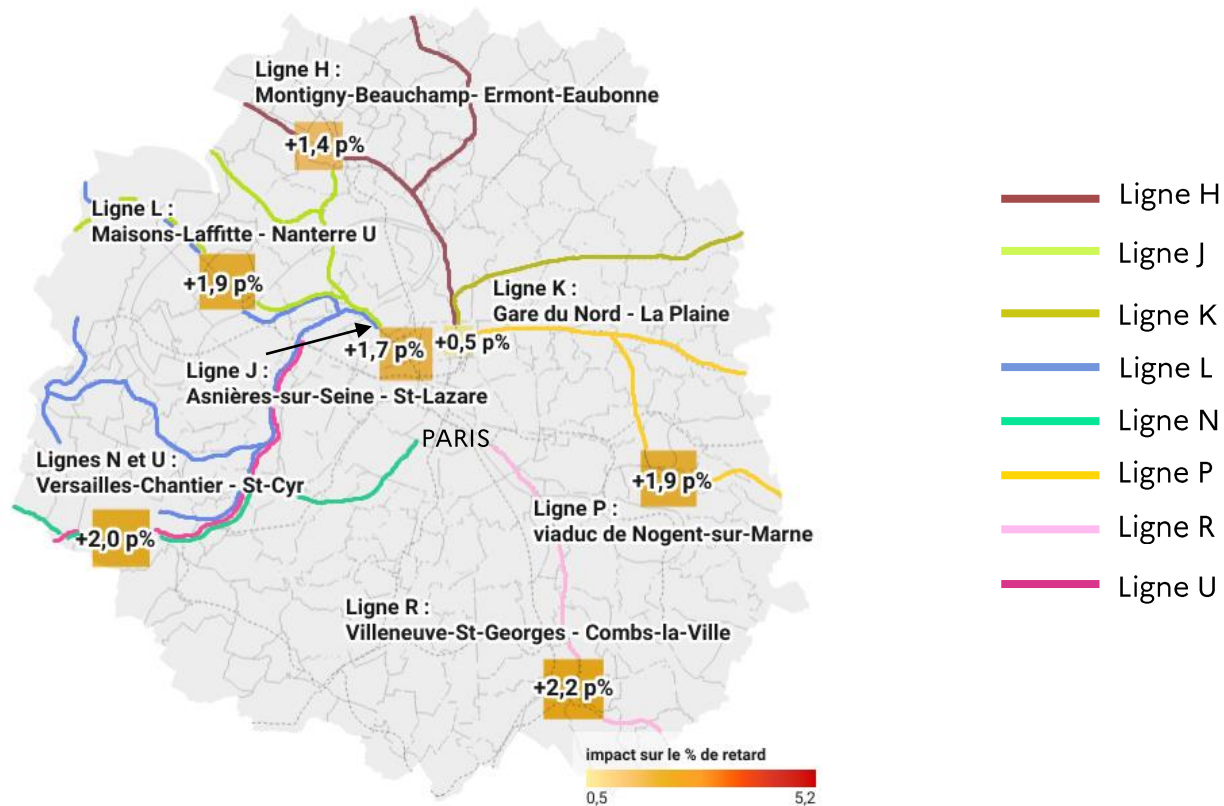
Activité ou service ferroviaire	Granularité des données utilisées	Période des données utilisées	R ² ajusté régression (p-value test de Fisher)	Variables explicatives	p-value test de Student variables explicatives	Impact sur le taux de retard
RER et Transilien (trains régionaux d'Ile-de-France)	Par lignes	2013-2019	0,63 ($< 0,001$)	Nombre de circulations quotidiennes de trains (toutes activités ferroviaires) par voie sur la section la plus chargée de la ligne	$< 1\%$	+1,1 pts / 100 circulations/voie/jour
				Distance moyenne parcourue par les trains RER ou Transilien	$< 0,1\%$	+1,1 pts / 10 km
				Estimation de l'âge moyen du matériel roulant en début d'année	$< 0,1\%$	+1,2 pts / 10 ans
				La ligne passe par la gare de Lyon	$< 5\%$	+2,1 pts
				La ligne passe par la gare Saint-Lazare	$< 0,1\%$	+5,3 pts
				La ligne passe par la gare du Nord	$< 0,1\%$	+1,7 pts
				La ligne passe par la gare de l'Est	$< 12\%$	+1,2 pts

Estimation économétrique de l'effet du passage par certaines grandes gares parisiennes et certains tronçons sur la ponctualité des RER



Carte: Alexandre Barbusse/AQST • Créé avec Datawrapper

Estimation économétrique de l'effet du passage par certains tronçons sur la ponctualité des Transilien



Carte: Alexandre Barbusse/AQST • Créé avec Datawrapper

Modèle RER et Transilien par ligne : Effet de la distance moyenne parcourue par les trains (+1,1 pts / 10 km)

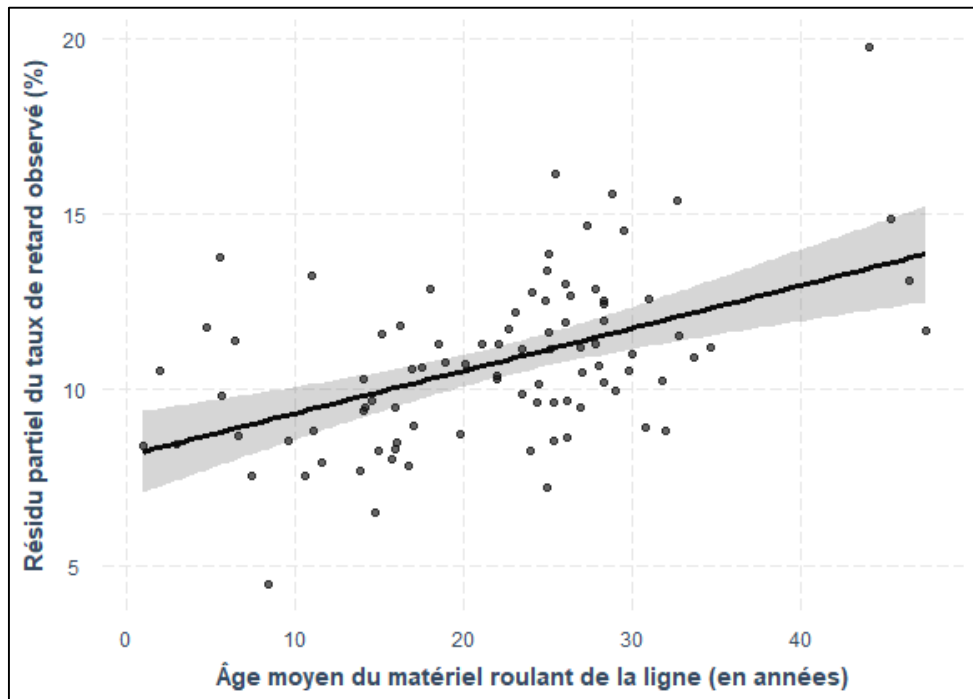


Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

La distribution du résidu partiel du taux de retard observé s'ajuste assez bien avec la droite de régression linéaire

Modèle RER et Transilien par ligne : Effet de l'âge moyen du matériel roulant (+1,2 pts / 10 ans)

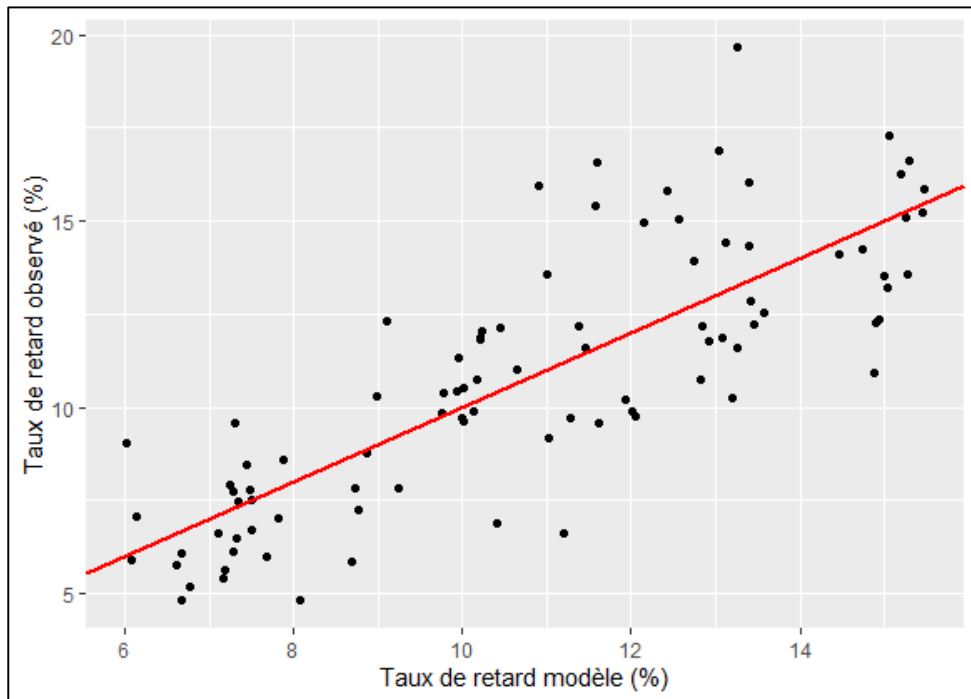


Droite de régression linéaire modélisant l'effet de la variable

Zone ombrée = plage d'incertitude statistique sur les paramètres de la droite

La dispersion du résidu partiel autour de la droite de régression est assez limitée par rapport aux variables précédentes, ce qui se traduit par une significativité statistique élevée de l'effet de l'âge moyen du matériel roulant

Modèle RER et Transilien par ligne : Taux de retard modèle vs taux de retard observé



Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

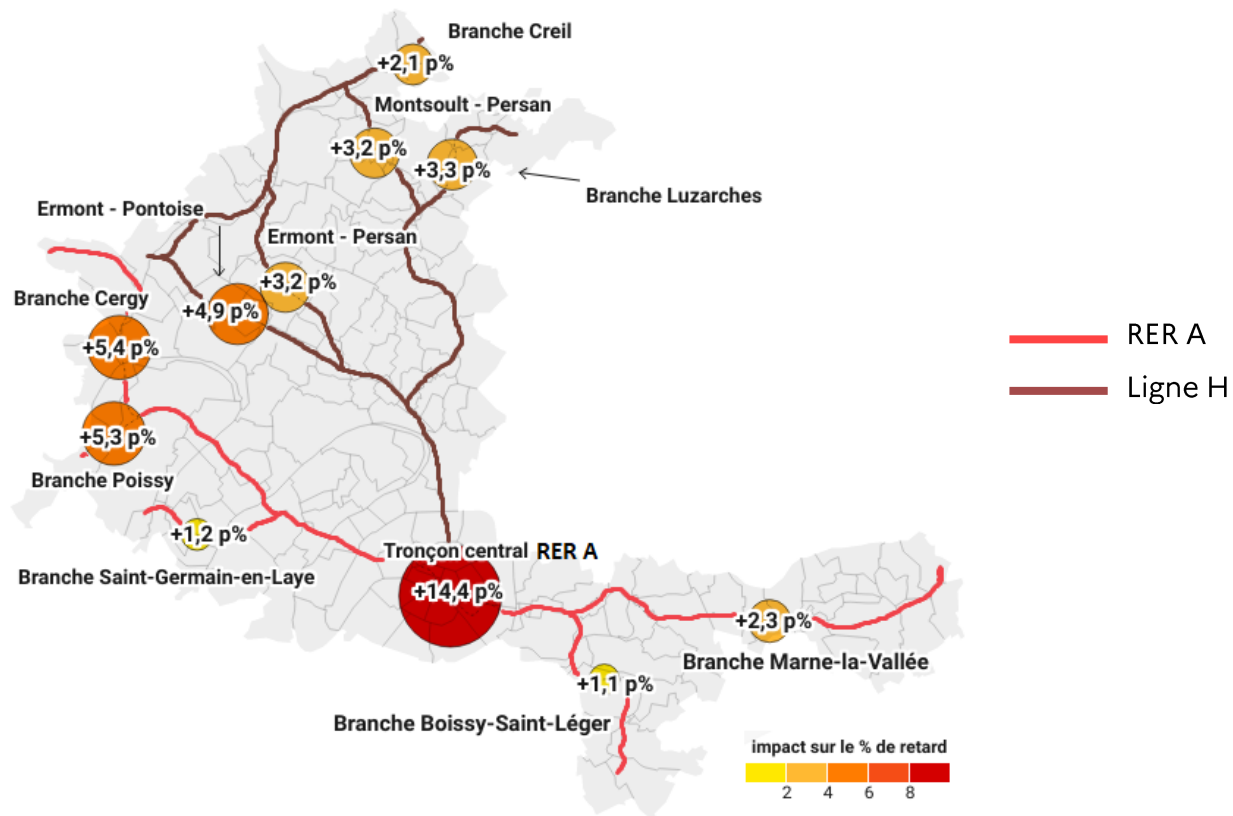
Les points s'alignent plutôt bien sur la droite – notamment pour les taux de retard prédits les plus faibles et les plus élevés - même si certains points s'écartent fortement de la droite

⇒ illustre l'assez bonne capacité du modèle par ligne retenu à expliquer la variabilité géographique et temporelle du taux de retard des RER et Transilien

RER et Transilien par branche : modèle retenu

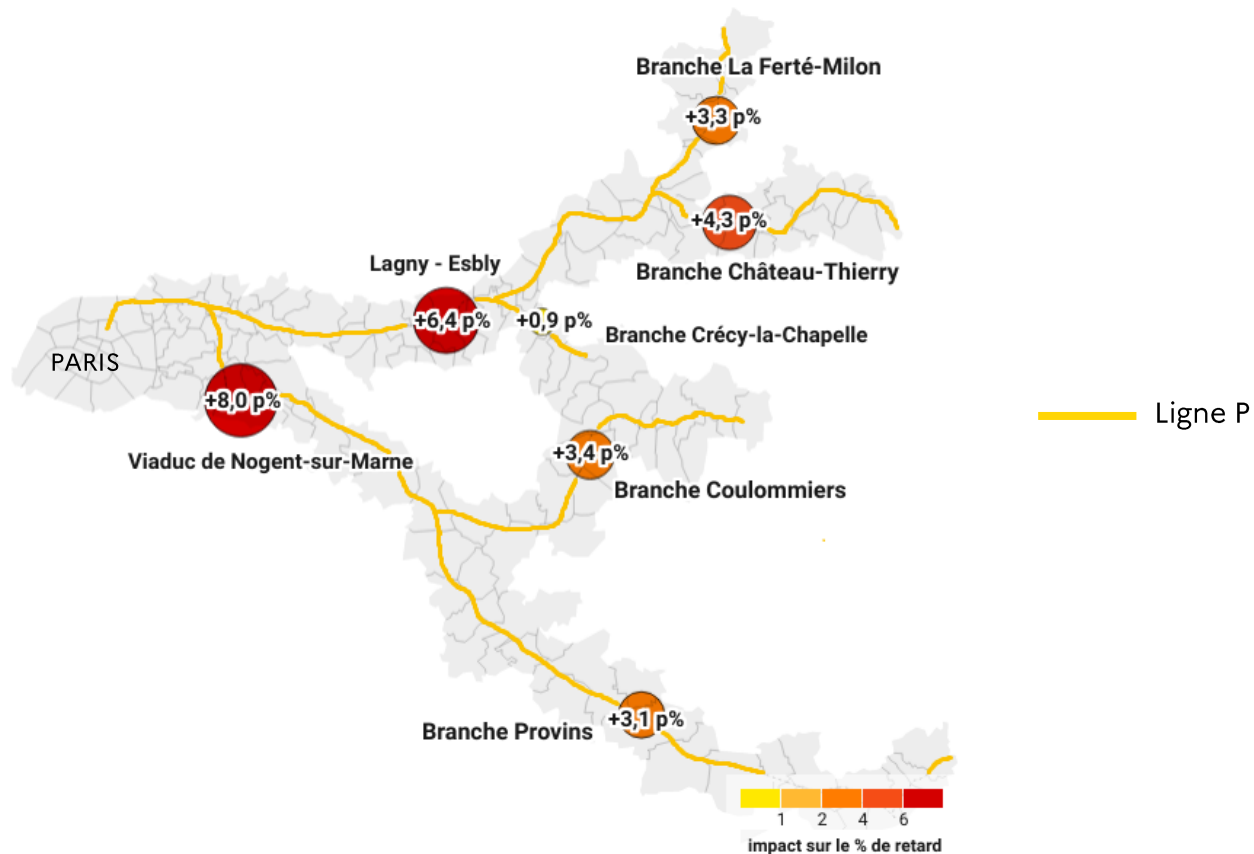
Activité ou service ferroviaire	Granularité des données utilisées	Période des données utilisées	R ² ajusté régression (p-value test de Fisher)	Variables explicatives	p-value test de Student variables explicatives	Impact sur le taux de retard
RER et Transilien Lignes A, E, N et R	Par branches	2013-2018	0,72 ($< 0,001$)	Nombre de circulations quotidiennes de trains (toutes activités ferroviaires) par voie sur la section la plus chargée de la branche	$< 2\%$	+2,0 pts / 100 circulations/voie/jour
				Nombre moyen d'arrêts effectués par les RER ou Transilien desservant la branche	$< 0,1\%$	+1,2 pts / arrêt
				Estimation de l'âge moyen du matériel roulant en début d'année	$< 0,1\%$	+3,9 pts / 10 ans
RER et Transilien Lignes A, H et P	Par branches	2013-2018	0,67 ($< 0,001$)	Nombre moyen de circulations quotidiennes de trains (toutes activités ferroviaires) par voie moyen sur la branche	$< 1\%$	+4,5 pts / 100 circulations/voie/jour
				Nombre de circulations quotidiennes de trains (toutes activités ferroviaires) par voie sur la section la plus chargée de la ligne	$< 0,1\%$	+5,4 pts / 100 circulations/voie/jour
				Estimation de l'âge moyen du matériel roulant en début d'année	$< 0,1\%$	+3,0 pts / 10 ans
				Nombre de voies de service en terminus de branche	$< 1\%$	-0,2 pt/ voie de service

Estimation économétrique de l'effet du trafic journalier moyen par branche sur la ponctualité des lignes A et H



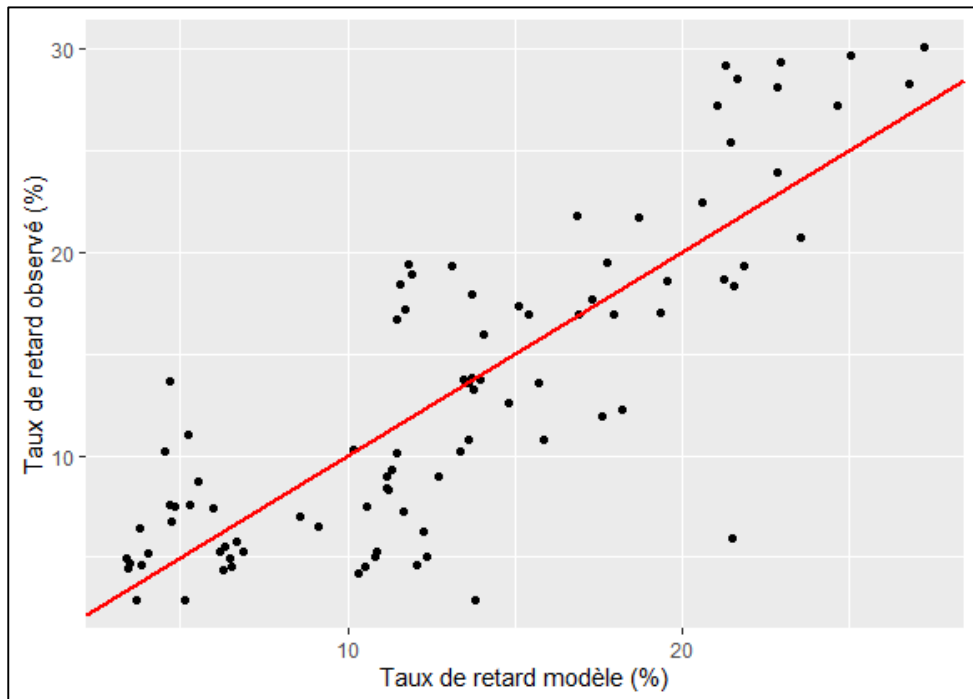
Carte: Alexandre Barbusse/AQST • Créé avec Datawrapper

Estimation économétrique de l'effet du trafic journalier moyen par branche sur la ponctualité de la ligne P



Carte: Alexandre Barbusse/AQST - Créé avec Datawrapper

Modèle RER et Transilien par branche (ligne A, H, P) : Taux de retard modèle vs taux de retard observé

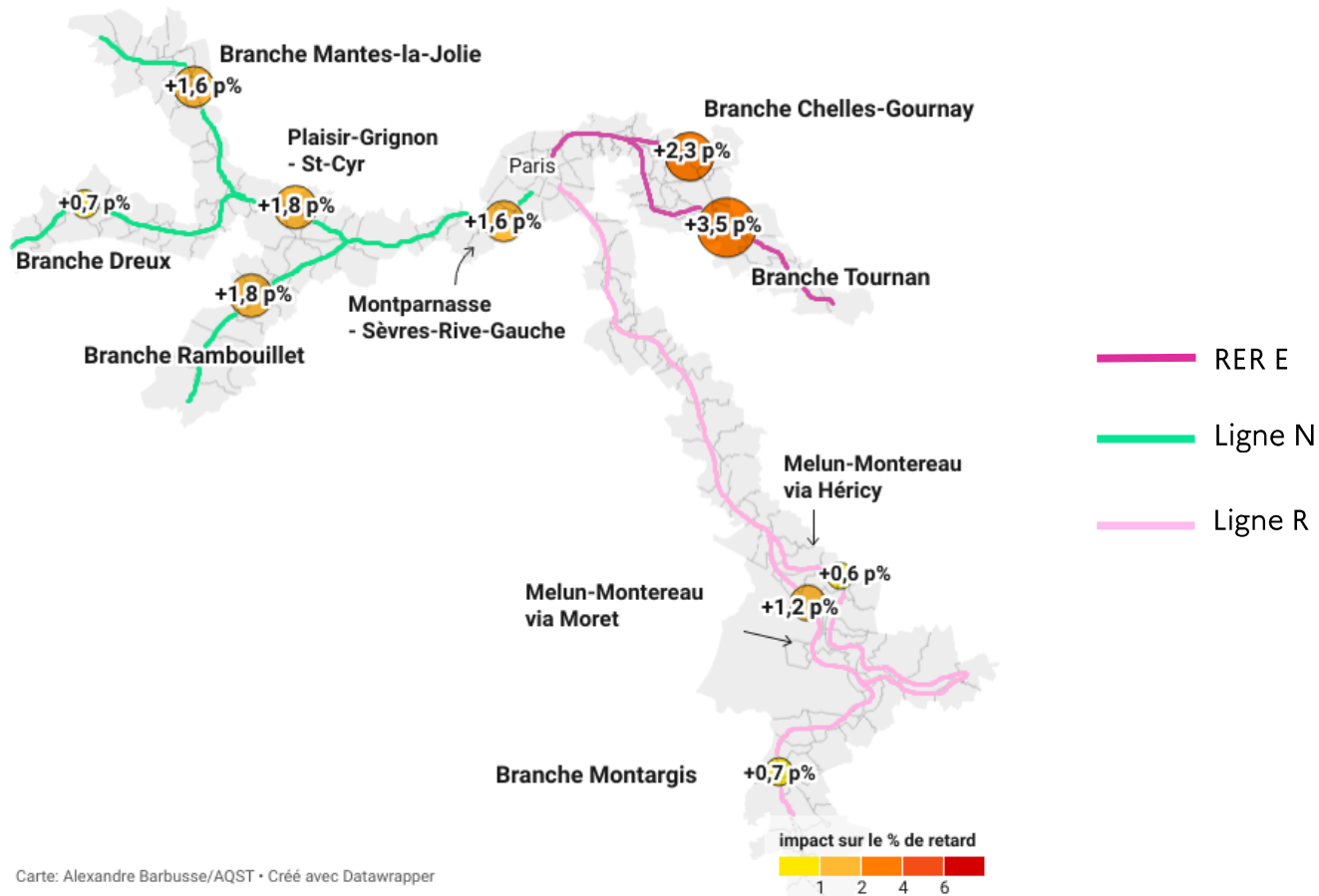


Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

Les points s'alignent plutôt bien sur la droite et leur dispersion autour de la droite est assez limitée

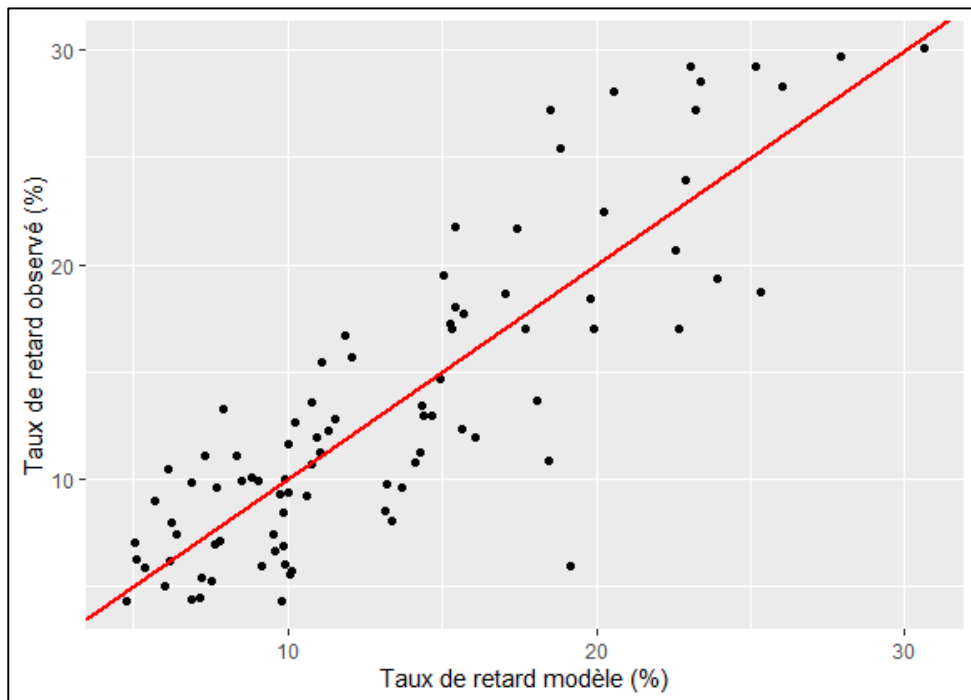
⇒ illustre l'assez bon pouvoir descriptif du modèle retenu pour expliquer la variabilité géographique et temporelle du taux de retard des branches des lignes RER A et Transilien H et P

Estimation économétrique de l'effet du trafic journalier maximal par branche sur la ponctualité des lignes E, N et R



Carte: Alexandre Barbusse/AQST • Créé avec Datawrapper

Modèle RER et Transilien par branche (ligne A, E, N, R) : Taux de retard modèle vs taux de retard observé



Ligne rouge = droite $y=x$ matérialisant le « modèle parfait »

Les points s'alignent plutôt bien sur la droite et leur dispersion autour de la droite est assez limitée

⇒ illustre une certaine capacité du modèle retenu à expliquer la variabilité géographique et temporelle du taux de retard des branches des lignes RER A et E, et Transilien N et R

Principaux résultats

- ❑ Pour les modèles retenus, la part de la variance du taux de retard sur la période 2012-2019 expliquée par le modèle est de l'ordre de 1/3 pour les services longue distance, et de l'ordre de 60% pour les trains régionaux. Les modèles retenus sont tous significatifs au seuil 0,1%.

- ❑ Dégradation de la ponctualité lors du passage par certains nœuds ferroviaires :
 - Majeurs : Paris, Lyon-Part-Dieu, Marseille Saint-Charles particulièrement congestionnés (CGEDD, BACOT, 2015)
 - ⇒ effet lors de la traversée visible pour les TGV (Lyon-Part-Dieu +4,9 pts) et Intercités (+5,2 pts/nœud) + Paris-gare de Lyon (+2,1 pts) pour les Transilien
 - Secondaires : effet au départ et au terminus des circulations pour les TGV (+3,1 pts) et trains Intercités (+1,6 pts) mais aussi lors de la traversée de certains nœuds = Bordeaux (+7,3 pts), Toulouse (+2,9 pts) et Dijon (+4,2 pts) pour les Intercités et Avignon (+2,4 pts) en TGV

- ❑ Les TGV passant par Rennes (-2,8 pts) et Dijon (-4,8 pts) sont significativement plus ponctuels
 - ⇒ Le nœud de Rennes a fait l'objet d'importants investissements de capacité lors de la construction de la LGV Bretagne-Pays de la Loire mise en service en 2017
 - ⇒ Les TGV passant par Dijon circulent en partie sur la LGV Rhin-Rhône qui est assez récente (2011) et connaît un trafic relativement faible

Principaux résultats

- ❑ En moyenne, le taux de retard des trains régionaux croît lorsque la distance parcourue augmente : plus un trajet dure longtemps, plus il a de chance « d’attraper un retard »

- ❑ En moyenne, pour l’ensemble des activités, une hausse du niveau de trafic au niveau de la section la plus chargée engendre une dégradation de la ponctualité
 - ⇒ Peut-être davantage de risques de subir un retard se propageant d’un train au suivant
 - + effet possible de la détérioration des infrastructures ferroviaires (voies, signalisation, etc.) à mesure qu’elles sont intensément utilisées

- ❑ Pour les TGV, les TER et les RER ou Transilien, chaque arrêt supplémentaire augmente le risque de retard (effet plus prononcé pour les trains régionaux)
 - ⇒ événements retardant le redémarrage du train lors d’un arrêt : malaises voyageurs, blocages des portes, dysfonctionnements du matériel roulant lors du redémarrage ou de la signalisation en sortie de gare, ...

Principaux résultats

- ❑ Pour les TER, plus le nombre de voyageurs transportés est grand, plus le risque d'occurrence d'un retard est important
 - ⇒ Peut-être lié aux retards causés par les voyageurs (malaises voyageurs, etc)

- ❑ Modèle TER : plus le réseau des « petites lignes » est dense, meilleure est la ponctualité des TER
 - ⇒ Hypothèse envisageable : une répartition plus homogène du trafic permettrait peut-être de réduire pour partie la congestion et éventuellement de limiter la propagation d'un gros retard

- ❑ Modèles RER et Transilien d'Ile-de-France : un vieillissement du parc de matériel roulant s'accompagne d'une baisse de la ponctualité (+1,2 pts de retard/ 10 ans)
 - ⇒ Un matériel roulant âgé présente un risque de panne ou d'incident matériel plus important

Principaux résultats

- ❑ Modèles RER et Transilien d'Ile-de-France : fort impact de la saturation des grandes gares parisiennes – en particulier la gare du Nord (+1,7 pts), la gare Saint-Lazare (+5,3 pts) et la gare de Lyon (+2,1 pts)
 - ➔ En accord avec la littérature technique (audit du réseau Transilien par l'EPFL, TZIEROPOULOS et PUTALLAZ, 2015 / étude de l'Institut Paris Région sur les grandes gares ferroviaires, 2013)

- ❑ Modélisation par branche pour les RER et Transilien : pour certaines lignes de RER (A et E) et la majorité des lignes Transilien : les branches les plus chargées sont moins ponctuelles, en accord avec l'intuition

- ❑ Pour certaines lignes de RER et Transilien : un nombre important de voies de service en terminus de branche contribue à l'amélioration de la ponctualité de la branche, potentiellement parce que cela offre davantage de souplesse à l'exploitation et à l'adaptation du plan de circulation, notamment en situation dégradée

Conclusion

- ❑ Les données du taux de retard des trains de voyageurs en France par activité ferroviaire sur la période 2012-2019 de la base de l'AQST montrent d'importantes disparités géographiques
- ❑ Des modèles économétriques simples pour identifier les principales variables permettant d'expliquer ces disparités géographiques du taux de retard sur la période 2012-2019
- ❑ Les modèles retenus mettent en évidence un effet significatif sur la ponctualité des trains en France du passage par certains nœuds ferroviaires (nœuds majeurs de Paris, Lyon-Part-Dieu et Marseille-Saint-Charles notamment) et par les grandes gares parisiennes proche des limites de leurs capacité pour les trains régionaux d'Ile-de-France (RER et Transilien). Les modèles montrent également que chaque arrêt supplémentaire accroît le risque de retard, en particulier pour les TER. Ils indiquent également un impact sur la ponctualité de l'intensité de trafic (toutes activités ferroviaires confondues) au niveau de la section la plus chargée de la ligne ou de la branche (RER et Transilien) empruntée par le train, et du volume de voyageurs transportés à l'échelle régionale (anciennes régions antérieures à la réforme de 2016) pour les TER. Enfin, les modèles dédiés aux trains régionaux d'Ile-de-France suggèrent que le vieillissement du parc de matériel roulant contribue à la dégradation de la ponctualité des trains.



**MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE**

*Liberté
Égalité
Fraternité*



**Conseil général de l'Environnement
et du Développement durable**

Merci de votre attention
