



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Statistik BFS

Projet d'innovation: STATENT-Flash

Étude exploratoire d'estimation d'une STATENT provisoire

Nous remercions l'équipe de méthodologie de l'OFS pour le soutien dans ce projet et la relecture du document.

EXPERIMENTAL STATISTICS

Neuchâtel, 2024

Éditeur: statent@bfs.admin.ch
Renseignements: statent@bfs.admin.ch, tél. +41 58 463 6 2 66
Contenu: Nicolas Vallon, Nadège Jacot, OFS
Domaine: O6 Industrie, services
Langue du texte original: Français

Concept de mise en page: Section PUB
Téléchargement: www.statistique.ch
Copyright: OFS, Neuchâtel 2024
La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée

TABLE DES MATIÈRES

1	Objet de ce document - Introduction.....	3
2	Situation initiale - Contexte.....	3
3	Objectifs	3
4	Méthodologie - Options	4
4.1	Modélisation agrégée	4
4.1.1	Variables dépendantes et explicatives	4
4.1.2	Modèles	6
4.2	Mesures de qualité	7
5	Résultats intermédiaires	8
5.1	Comparaison des modèles.....	8
5.2	Analyse des modèles retenus	9
6	Discussion et limites	9
7	Conclusion	10
8	Références.....	10
A	Annexe : Dimensions retenues.....	12
B	Annexe : Distribution des variables dépendantes	14
C	Annexe : Liste des modèles	15
D	Annexe : Résultats modélisation	17
D.A	Résultats selon le Root Mean Squared Forecast Error et l'écart total relatif	17
D.B	Résultats du modèle retenu et de référence pour les unités légales	24
D.C	Résultats du modèle retenu et de référence pour les emplois	27

1 OBJET DE CE DOCUMENT - INTRODUCTION

Ce document dresse les grandes lignes de la méthodologie utilisée pour réaliser des premières estimations dans le cadre du projet d'innovation STATENT-Flash. Les principaux éléments de la modélisation y sont décrits. Les modèles d'estimations développés ont pour but de fournir aux utilisateurs des résultats provisoires agrégés sur la structure des entreprises plus rapidement que ce qui est possible dans le cadre de la production classique de la Statistique structurelle des entreprises (STATENT). Les données sur la structure de l'économie mises à disposition par la Statistique structurelle des entreprises sont actuellement diffusées à t+20 mois (par exemple : données de décembre 2022¹ disponibles en août 2024). L'objectif est d'avoir 8 mois après la date de référence des informations sur le nombre d'emplois et d'unités légales avec une ventilation par classe de taille d'emploi, division économique de la nomenclature générale des activités économiques (NOGA) et canton.

2 SITUATION INITIALE - CONTEXTE

Le processus d'intégration standard de la STATENT mélange les données des sources suivantes :

1. Registre de l'Assurance-vieillesse et survivants (AVS)
2. Registre des entreprises et des établissements (REE)
3. Les enquêtes de mise à jour du Registre des entreprises et des établissements (ERST, Profiling et Profiling-Light)
4. Statistique de l'emploi (STATEM)
5. Relevé des structures agricoles (STRU)
6. Statistique des institutions médico-sociales (SOMED)
7. Statistique de l'aide et des soins à domicile (SPITEX)

Selon le calendrier de diffusion prévu pour la STATENT-Flash (t+8 mois), ces sources sont toutes entièrement disponibles à l'exception des données AVS. Comme ces dernières constituent la source principale de la STATENT, leurs délais de livraison, d'intégration et de traitement ne permettent pas une mise à disposition des résultats dans un laps de temps inférieur à t+20 mois. En effet, il y a plusieurs livraisons des données par les caisses AVS. Les dernières données les plus complètes pour une année de référence t sont livrées par cette source à l'OFS avec un décalage de 11 mois (soit en t+11 mois). Les 9 mois suivants sont utilisés pour le traitement des données AVS, la construction des données STATENT et sa diffusion à t+20 mois.

3 OBJECTIFS

Le but ici est la construction d'une STATENT-Flash offrant des données sur la structure de l'économie suisse à t+8 mois (par exemple : données de décembre 2023 disponible en août 2024) en conformité avec les standards internationaux en matière de « rapid estimates »². Le cycle de production de la STATENT-Flash est significativement réduit par rapport à la STATENT habituelle. Ainsi, par exemple, la diffusion de la STATENT portant sur l'année 2022 et la diffusion de la STATENT-Flash portant sur l'année 2023 seraient synchronisées.

¹ La période de référence de la STATENT est le mois de décembre.

² « Handbook on Rapid Estimates », Eurostat, 2017 :

<https://ec.europa.eu/eurostat/documents/3859598/8555708/KS-GQ-17-008-EN-N.pdf>

Suite à une analyse des besoins des utilisateurs, l'objectif est d'avoir des informations sur le nombre d'emplois et d'unités légales avec une ventilation par classe de taille d'emploi, division NOGA OFS 50 et canton. (Annexe A) Les divisions économiques NOGA OFS 50 sont un regroupement des codes NOGA selon les deux premiers chiffres du code. Pour des raisons de simplicité, le terme « division NOGA » correspond aux divisions selon la NOGA OFS 50.

4 MÉTHODOLOGIE - OPTIONS

La population de la STATENT est définie par les unités légales ayant au minimum un emploi en décembre avec plus de 2'300 CHF de revenu annuel soumis à une cotisation AVS. Cette population est inconnue, à très court terme, en raison de livraisons incomplètes des données par les caisses AVS. Deux options ont été identifiées pour construire la STATENT-Flash. La première consiste à prédire les unités légales qui seront présentes. Cela représenterait une population estimée. Dans un deuxième temps, les emplois seraient prédits pour cette population estimée. Il y a dans ce cas deux modèles : un sur la présence d'une unité légale et un sur les emplois de ces unités légales. Cette approche est dite « individuelle » car la modélisation serait au niveau des unités légales. Elle n'a pas encore été étudiée et n'est donc pas commentée ici. La deuxième option, présentée ici, teste l'option d'une modélisation directement au niveau des agrégations les plus demandées par les utilisateurs. Ainsi, une observation du modèle correspond au croisement entre la classe de taille, le canton et la division NOGA. Il y a dans ce cas également deux modèles : un pour le nombre d'emplois et un pour le nombre d'unités légales. Ces deux modèles sont indépendants et peuvent potentiellement donner des informations contradictoires, c'est-à-dire par exemple pour une observation un nombre d'emplois positif et 0 unité légale ou inversement (voir section 6).

4.1 MODÉLISATION AGRÉGÉE

Le secteur primaire (NOGA 01 à 03) et l'administration publique (NOGA 84) proviennent principalement du relevé STRU et des enquêtes Profiling/Profiling-Light respectivement. L'ensemble des données de ces sources sont disponibles au moment de l'établissement des estimations. Ces deux branches auront donc une prédiction individuelle et sont exclues de la modélisation agrégée, présentée dans ce document.

Les modèles reposent sur les dimensions suivantes (Annexe A) : 46 divisions NOGA, 4 classes de taille d'emploi et 26 cantons. Ainsi, chaque année, il y a 4'784 observations (46x4x26). La construction de cette base de données utilisée lors de l'estimation est expliquée plus en détail à la section 4.1.1.

Les résultats sont présentés dans une perspective de production. C'est-à-dire que pour chaque prédiction, uniquement les informations disponibles en accord avec une diffusion prévue en t+8 mois sont utilisées. Ainsi, pour prédire la variable dépendante (voir ci-dessous) de l'année 2019, les variables dépendantes et explicatives de l'année 2018 sont utilisées pour l'estimation du modèle. Les paramètres estimés de ce modèle sont ensuite appliqués aux variables explicatives de 2019, ce qui donne une prédiction pour la variable dépendante pour l'année 2019. Cette prédiction pour 2019 est ensuite comparée à la STATENT effective. Par analogie, les années 2020 et 2021 sont traitées de manière similaire.

4.1.1 Variables dépendantes et explicatives

Les variables dépendantes sont le nombre d'emplois et le nombre d'unités légales. Une grande partie des observations de la variable dépendante ont une valeur nulle, c'est-à-dire qu'il n'y a pas d'unité légale à l'intersection de certains canton, division NOGA et classe de taille. L'annexe B montre la distribution du nombre d'unités légales et d'emplois pour le croisement de ces dimensions. Ces distributions sont fortement inclinées vers la droite, possèdent une

plage de valeurs positives et présentent une accumulation d'observations de valeur 0. Les approches présentées dans le cadre de ce document ne tiennent compte de ces faits que partiellement et de manière limitée. Si les approches agrégées décrites ici devaient être poursuivies, la modélisation pourrait être développée de manière à ce que les modèles utilisés correspondent mieux aux distributions effectives des données.

Les variables explicatives sont développées à partir des sources de la STATENT, de l'AVS, du REE de fin d'année ou de la STATEM. Ces variables seront utilisées dans les modèles décrits à la section 4.1.2 ci-dessous. Chaque variable est basée sur des données au niveau unité légale, où le croisement division NOGA, canton et classe de taille est calculé puis les données sont agrégées pour avoir une observation par croisement. Pour chaque unité légale, les dimensions division NOGA et canton proviennent du REE au 31.12. Toutes les variables seraient disponibles dans le cas d'une publication en t+8 mois. Cinq variables explicatives sont utilisées pour la prédiction de l'emploi, à savoir:

1. Basé sur la deuxième livraison de l'AVS³ :
 - a. Total AVSQ2 : Nombre d'emplois présents dans la deuxième livraison de l'AVS pour les unités ayant au moins un emploi en décembre avec plus de 2'300 CHF de revenu annuel.
 - b. Changement emploi AVSQ2 : Changement du nombre d'emplois pour les unités légales présentes dans les deux dernières livraisons AVSQ2. Par exemple, pour 2020, la population correspond aux unités légales présentes dans les deuxièmes livraisons en 2019 et en 2020.
2. Basé sur le REE :
3. Basé sur la STATENT t-1 :
 - a. Total STATENT t-1 : Nombre d'emplois publié dans la STATENT de l'année précédente.
4. Basé sur la STATEM :
- a. STATEM : Nombre d'emplois pondéré⁴ selon la STATEM Q4 (dernier trimestre de l'année) en utilisant les poids d'extrapolation de l'enquête. Cette pondération a pour but de représenter les divisions NOGA et les grandes régions au niveau des établissements. Une unité légale est composée d'un à plusieurs établissements. L'emploi est pondéré au niveau des établissements. Puis, les emplois pondérés sont agrégés par unité légale.

Seulement trois variables sont disponibles pour estimer le nombre d'unités légales. La plupart des sources externes se focalisent sur les emplois et non sur le nombre d'unités légales. Voici les trois variables :

1. Basé sur la deuxième livraison de l'AVS :

³ Il s'agit de la dernière livraison disponible pour une publication de la STATENT-Flash à t+8 mois.

⁴ Plus d'informations sur la pondération des données STATEM :

<https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/publications.assetdetail.344658.html>

- a. Total AVSQ2 : Nombre d'unités légales présentes dans la deuxième livraison de l'AVS ayant au moins un emploi en décembre avec plus de 2'300 CHF de revenu annuel.
2. Basé sur le REE :
- a. Total REE : Nombre d'unités légales présentes dans le REE de fin d'année ayant au minimum un emploi. La source des emplois dans cette mesure est mixte entre principalement la STATENT, le Profiling, le Profiling-Light et la ERST.
3. Basé sur la STATENT t-1 :
- a. Total STATENT t-1 : Nombre d'unités légales publiées dans la STATENT de l'année précédente.

4.1.2 Modèles

Les modèles seront différents selon leur type (voir ci-dessous) et les variables incluses dans l'estimation. Au total, il y a 16 modèles pour le nombre d'unités légales et 30 modèles pour le nombre d'emplois. Comme indiqué précédemment, les distributions des variables dépendantes sont fortement asymétriques. Ces distributions impactent fortement l'estimation. Cet aspect des données est partiellement pris en compte à l'aide du modèle *TOBIT* de type 1 ci-dessous.

Les modèles ci-dessous ont principalement les deux hypothèses sous-jacentes suivantes :

1. Les paramètres ne dépendent pas des croisements. Ainsi, le lien entre une variable explicative et la variable dépendante est constant entre les différents croisements. Il peut y avoir des violations de cette hypothèse si par exemple pour la variable *Total AVSQ2* (voir section 4.1.1), les unités légales de certaines divisions économiques livrent systématiquement en retard comparativement à d'autres branches. Des coefficients par division NOGA, canton ou classe de taille peuvent potentiellement être développés.
2. Les paramètres estimés sur une année sont également valides pour l'année suivante. L'hypothèse est que le lien entre les variables explicatives et la variable dépendante est constant sur plusieurs années. Cette hypothèse serait violée par exemple pour la variable *Total AVSQ2* (voir section 4.1.1) si les livraisons des caisses AVS varient fortement d'une année à l'autre, ce qui a déjà été constaté par le passé.

Quatre types de modèles sont implémentés⁵, où j est le croisement des dimensions division NOGA, canton et classe de taille et t est l'année. Un tableau récapitulatif des modèles est disponible en annexe C :

1. Modèle *TOTAL* : Régression linéaire sur le nombre absolu d'unités légales ou d'emplois :

$$\text{STATENT}_{j,t} = \alpha_0 + \alpha_1 \text{Variables explicatives}_{j,t} + \alpha_2 \text{Variables explicatives}_{j,t-1} + \varepsilon_{j,t}$$

où $(\alpha_0, \alpha_1, \alpha_2)$ représentent les coefficients de régression correspondants.

2. Modèle *DIFF* : Régression linéaire sur la différence du nombre d'unités légales ou d'emplois entre deux années consécutives:

⁵ D'autres variantes de ces modèles ont été estimées, comme l'introduction d'effets fixes qui génère une grande quantité de prédictions négatives. Ces modèles sont provisoirement rejetés.

$$\begin{aligned} \text{STATENT}_{j,t} - \text{STATENT}_{j,t-1} \\ = \beta_0 + \beta_1(\text{Variables explicatives}_{j,t} - \text{Variables explicatives}_{j,t-1}) \\ + \beta_2(\text{Variables explicatives}_{j,t-1} - \text{Variables explicatives}_{j,t-2}) + \varepsilon_{j,t}, \end{aligned}$$

où $(\beta_0, \beta_1, \beta_2)$ représentent les coefficients de régression correspondants. Pour avoir la prédiction finale, la différence prédictée est ajoutée à la valeur de la STATENT en t-1.

3. Modèle *TOBIT* de type 1 sur le nombre absolu d'unités légales ou d'emplois :

$$\text{STATENT}_{j,t} = \begin{cases} 0 & \text{si } \text{STATENT}_{j,t}^* < 0 \\ \text{STATENT}_{j,t}^* & \text{si } 0 \leq \text{STATENT}_{j,t}^* \end{cases}$$

$$\text{STATENT}_{j,t}^* = \gamma_0 + \gamma_1 \text{Variables explicatives}_{j,t} + \gamma_2 \text{Variables explicatives}_{j,t-1} + \varepsilon_{j,t},$$

où $(\gamma_0, \gamma_1, \gamma_2)$ représentent les coefficients de régression correspondants. Ici, le résultat de la régression est borné à 0. Ce modèle *TOBIT* de type 1 est prévu à l'origine pour les variables censurées mais il peut être utilisé également dans ce cas, où la variable dépendante est nulle ou positive⁶.

4. Modèle *LOG* sur le nombre absolu d'unités légales ou d'emplois avec la fonction lien logarithmique :

$$E(Y_{j,t}|X_{j,t}, X_{j,t-1}) = e^{\delta_0 + \delta_1 \ln(X_{j,t}) + \delta_2 \ln(X_{j,t-1})}$$

$$\text{Var}(Y_{j,t}|X_{j,t}, X_{j,t-1}) = \sigma^2,$$

où Y est la STATENT, X sont les variables explicatives et $(\delta_0, \delta_1, \delta_2)$ représentent les coefficients de régression correspondants. À noter que dans le cas d'une variable explicative avec une valeur nulle, celle-ci est remplacée par 1 pour avoir un logarithme calculable. De plus, la variance est indépendante de l'espérance.

Les types de modèle 1 et 2 peuvent potentiellement produire des prédictions négatives. Dans les résultats, ces prédictions ne sont pas bornées à 0. Pour contourner ce potentiel problème, le modèle *TOBIT* de type 1 et le modèle *LOG* ont été développés.

Les modèles ci-dessus seront comparés à un modèle de référence nommée *PAST*. Ce modèle consiste à reprendre simplement les données de l'année précédente. Par exemple, la STATENT-Flash de 2023 correspondrait aux données de la STATENT de 2022. Ce modèle de référence définit les performances minimales pour les modèles développés. Ce modèle implique l'hypothèse qu'en agrégé, l'économie suisse est stable, bien que cela soit possible qu'il y ait des mouvements démographiques tels que des créations d'unités légales et d'emplois, des dissolutions d'unités légales, des pertes d'emplois, etc. Mais de manière globale, ces différents mouvements se compenseraient pour donner des totaux agrégés stables sur les années successivement.

4.2 MESURES DE QUALITÉ

Trois mesures de qualité sont calculées pour chaque modèle afin d'évaluer la qualité des prédictions, où j est le croisement des dimensions division NOGA, canton et classe de taille et t est l'année :

⁶ Jeffrey M. Wooldridge, 2001. "Econometric Analysis of Cross Section and Panel Data," MIT Press Books, The MIT Press, édition 1, volume 1, number 0262232197, December.

1. Un indicateur global: Le “Root Mean Squared Forecast Error” (RMSFE)

$$RMSFE_t = \sqrt{\sum_j \frac{1}{n} (\text{Prédiction}_{j,t} - \text{STATENT}_{j,t})^2},$$

2. L'écart total relatif de prédiction

$$\text{Écart}_t = \frac{\sum_j \text{Prédiction}_{j,t} - \sum_j \text{STATENT}_{j,t}}{\sum_j \text{STATENT}_{j,t}},$$

3. Une mesure basée sur la distribution des écarts relatifs entre la prédiction et la valeur effectivement observée. (Résultats sur demande).

$$\text{Écart}_{j,t} = \frac{\text{Prédiction}_{j,t} - \text{STATENT}_{j,t}}{\text{STATENT}_{j,t}},$$

L'indicateur global, qui est toujours positif, mesure la dispersion des écarts de prédiction (différences entre les valeurs observées et leur prédictions). Une valeur plus basse indique un meilleur modèle. L'écart total relatif de prédiction donne de l'information sur la capacité d'un modèle à estimer le total d'emplois (respectivement d'unités légales) de l'économie suisse. Cet écart peut être positif ou négatif. La dernière mesure est utile pour des analyses plus fines des modèles et afin de comprendre où est-ce que le modèle performe mieux.

5 RÉSULTATS INTERMÉDIAIRES

5.1 COMPARAISON DES MODÈLES

Les résultats des différents modèles présentés ici sont disponibles dans les tableaux de l'annexe D.A. Dans les tableaux de résultats, il y a les cinq types de modèles (*TOTAL*, *DIFF*, *TOBIT*, *LOG* et *PAST*) ainsi que les variations des variables explicatives utilisées. Une explication pour la lecture des tableaux est disponible au début de l'annexe.

Sur la prédiction du nombre d'**unités légales** (tableaux Tableau 4 à Tableau 6), le nombre de variables explicatives développées est faible et la qualité de ces variables, en tant que variables explicatives, n'est pas suffisante. Le modèle de référence (*PAST*) a un indicateur global entre 10.48 et 14.76. Du au COVID, l'année 2020 a été particulièrement stable en termes de nombre d'unités légales, ce qui explique la bonne performance en reprenant simplement l'année précédente comme prédiction. Pour les autres modèles (*TOTAL*, *DIFF*, *TOBIT* et *LOG*), l'ajout des variables explicatives améliore sensiblement la prédiction. À l'exception de 2020, le modèle *TOBIT* semble montrer plus d'écart au total que les autres modèles. Pour les meilleurs modèles (RMSFE) de chaque année, l'indicateur global est meilleur en 2019, moins bon en 2020 et similaire en 2021 que le modèle de référence. Au vu de ce qui précède, l'ajout de complexité (*TOBIT* et *LOG*) dans le type de modèle n'améliore pas grandement l'estimation. De plus, les modèles *DIFF* et *TOTAL* avec intercept ont parfois un grand nombre de prédictions négatives. Le modèle retenu est ainsi *TOTAL – NOINTREG3* pour le nombre d'unités légales.

Sur la prédiction du nombre d'**emplois** (tableaux Tableau 7 à Tableau 9), le modèle de référence (*PAST*) a un indicateur global entre 163.29 et 205.82 ainsi qu'un écart total relatif entre -2.42% et 0.95%. Les modèles *DIFF* du nombre d'emplois performent (indicateur global) moins bien que les modèles *TOTAL* en 2019 et 2020, de manière similaire en 2021. Le modèle *TOBIT* montre une très bonne prédiction en 2020 mais ce modèle semble moins stable pour l'écart total relatif. En effet, il y a une sous-estimation de 4% de l'emploi total en 2021. Les

modèles *LOG* montrent des résultats similaires (2019) ou moins bons (2020 et 2021) en termes d'écart total relatif relativement aux modèles *TOTAL*. Un critère supplémentaire est le nombre de prédictions négatives. Les modèles *TOTAL* avec intercept et *DIFF* montrent un grand nombre de prédictions négatives en 2021. Ces différents éléments favorisent ainsi le choix de sélectionner le modèle *TOTAL - NOINTREG4*.

Le meilleur modèle retenu jusqu'à présent est le modèle *TOTAL - NOINTREG3* pour le nombre d'unités légales et le modèle *TOTAL - NOINTREG4* pour le nombre d'emplois. Ces modèles montrent de la stabilité et de bons résultats sur les trois années (2019, 2020 et 2021). En ayant supprimé l'intercept, il y a également peu de prédictions négatives.

5.2 ANALYSE DES MODÈLES RETENUS

Les annexes D.B et D.C montrent pour les deux modèles retenus les coefficients des régressions et les résultats agrégés par dimension. On y voit ainsi par exemple pour chaque canton la différence entre ce qui est prédit et ce qui est publié dans la STATENT de 2019 à 2021. Il est prévu de publier trois tableaux (un par dimension) en août 2024 avec le nombre total d'emplois et d'unités légales.

Sur la prédiction du nombre d'**unités légales** (tableaux Tableau 10 à Tableau 13), les coefficients de la régression sont tous significatifs à 1%. Le coefficient lié à la variable de la STATENT t-1 est proche de 1 pour toutes les années. Les résultats agrégés par canton montrent bien que ce modèle surestime les unités légales en 2020 et les sous-estime en 2021. Ainsi le modèle ne semble pas totalement capable de prendre en compte des grands chocs sur l'économie comme le COVID-19. Les écarts sont plus répartis entre positif et négatif en 2019. L'écart relatif oscille entre -3.92% en 2021 et +3.16% en 2020. Ces deux valeurs sont observées dans le canton de Glaris. En regardant la répartition selon la division NOGA, les écarts oscillent entre -7.61% pour l'industrie pharmaceutique en 2019 et +7.41% pour le transport terrestre et transport par conduites en 2020.

Sur la prédiction du nombre d'**emplois** (tableaux Tableau 14 à Tableau 17), les coefficients de la régression sont tous significatifs à 1%, à l'exception des variables STATEM et le nombre d'emplois selon l'AVSQ2 en 2019. Pour les résultats agrégés par dimension, la surestimation ou la sous-estimation est globalement similaire à la prédiction du nombre d'unités légales. Cependant, l'amplitude de l'écart semble plus élevée que pour le nombre d'unités légales. Pour les divisions NOGA, les écarts oscillent entre -8.96% pour la recherche et développement scientifique en 2020 et +14.99% pour l'hébergement en 2020. Pour les cantons, les écarts oscillent entre -3.42% pour le Valais en 2021 et +4.79% pour Zoug en 2020.

Avec de tels écarts, les résultats agrégés sont ainsi insuffisants pour une publication standard. Excepté l'ajout d'autres variables explicatives, le modèle agrégé ne semble, à priori, pas être en mesure d'être suffisamment amélioré. Les autres sources, non utilisées ici (ERST, Profiling, Profiling-Light, STRU, SOMED et SPITEX) ne permettent pas de créer des variables explicatives pour toute l'économie suisse. En effet, ces sources donnent des informations uniquement sur des branches économiques spécifiques et sont donc moins représentatives de l'ensemble de la population STATENT.

6 DISCUSSION ET LIMITES

La modélisation agrégée retenue jusqu'à présent a plusieurs limites.

Étant donné que les modèles sur l'emploi et le nombre d'unités légales sont indépendants, il est possible qu'un modèle prédise un nombre d'emplois positif et l'autre modèle un nombre d'unités légales inférieur ou égal à 0 (et inversement). Pour l'année 2019, 1.6% des

croisements montrent cette incohérence. Cette incohérence et les prédictions négatives ne sont cependant pas visibles dans les tableaux agrégés par dimension.

Les résultats présentés ici sont le fruit d'une estimation. Ces valeurs sont incertaines et il serait avantageux de pouvoir fournir en plus du point estimé, l'incertitude de l'estimation à travers un intervalle de confiance. Ce point reste ouvert pour l'instant.

Certains modèles proposés ici reposent sur des hypothèses concernant les distributions. Ces hypothèses n'ont pas été testées et pour l'instant il n'est pas prévu de mener des travaux qui vont dans cette direction, au vu des résultats insuffisants. Une analyse des résidus a été effectuée (graphique sur demande). La distribution des résidus est fortement asymétrique indiquant probablement que les hypothèses de certains modèles ne sont pas vérifiées. Ce point peut potentiellement être amélioré. (Voir point suivant).

La gestion de la borne inférieure à zéro, la masse d'observations avec de petites valeurs et l'asymétrie de la distribution des valeurs dépendantes ne sont pas totalement pris en compte. Il existe plusieurs possibilités pour remédier à ces éléments. Une possibilité serait d'étudier des modèles linéaires généralisés avec des distributions qui correspondent mieux aux distributions des variables cibles observées ici. Par exemple, la distribution zéro-inflated Poisson pour le nombre d'entreprises ou la distribution exponentielle pour le nombre d'employés. Une deuxième possibilité serait de prendre le logarithme des variables dépendantes mais la masse de valeurs nulles devra être également modélisée. L'étude de ces possibilités n'est pour l'instant pas priorisée.

Une analyse des observations influentes a montré qu'en supprimant certaines observations, les paramètres des modèles changeaient fortement. Cela résulte probablement de l'asymétrie des variables dépendantes. Ce fait souligne la problématique du point précédent.

L'approche agrégée est, en l'état, insuffisante pour des résultats suffisamment justes et précis. Au vu des limites ci-dessus, les fortes hypothèses et les résultats présentés ici, il a été décidé de donner la priorité au développement de modèles dits individuels. Plus que la forme fonctionnelle, la pondération des observations ou autres éléments, c'est les variables explicatives qui permettent d'améliorer l'estimation. Une plus grande variété de sources de données pourra être exploitée avec l'approche individuelle et potentiellement mieux prendre en compte des chocs structurels comme le COVID-19. Les modèles développés dans le futur seront toujours comparés à ceux présentés ici, en espérant que les résultats agrégés pourront être améliorés à l'aide des autres approches.

7 CONCLUSION

Le projet d'innovation STATENT-Flash, dont le but est d'augmenter l'offre d'informations pour le public, est un défi méthodologique. La population étant inconnue au moment de l'estimation, il est difficile de produire de bonnes estimations. Un travail pour mieux comprendre et anticiper la composition de cette population permettra, à la fois d'améliorer potentiellement les résultats ici, mais également de mieux comprendre l'évolution de la Statistique structurelle des entreprises.

8 RÉFÉRENCES

« Handbook on Rapid Estimates », Eurostat, 2017 :

<https://ec.europa.eu/eurostat/documents/3859598/8555708/KS-GQ-17-008-EN-N.pdf>

Jeffrey M. Wooldridge, 2001. "Econometric Analysis of Cross Section and Panel Data," MIT Press Books, The MIT Press, edition 1, volume 1, number 0262232197, December.

SAS Institute Inc. SAS/STAT 14.3 User's Guide, The QLIM Procedure. Cary, NC, 2017. URL: https://documentation.sas.com/api/collections/pgmsascdc/9.4_3.3/docsets/etsug/content/qlim.pdf?locale=en#nameddest=etsug qlim_syntax

SAS Institute Inc. SAS/STAT 14.2 User's Guide, The REG Procedure. Cary, NC, 2017. URL: https://documentation.sas.com/api/collections/statcdc/14.2/docsets/statug/content/reg.pdf?locale=en#nameddest=statug_reg_references

SAS Institute Inc. SAS/STAT 15.1 User's Guide, The GENMOD Procedure. Cary, NC, 2018. URL:

https://documentation.sas.com/api/collections/pgmsascdc/9.4_3.4/docsets/statug/content/genmod.pdf?locale=en#nameddest=statug_genmod_syntax01

Hocking, R. R. "A Biometrics Invited Paper. The Analysis and Selection of Variables in Linear Regression." *Biometrics* 32, no. 1 (1976): 1–49. <https://doi.org/10.2307/2529336>

A ANNEXE : DIMENSIONS RETENUES

Tableau 1: Description des dimensions utilisées

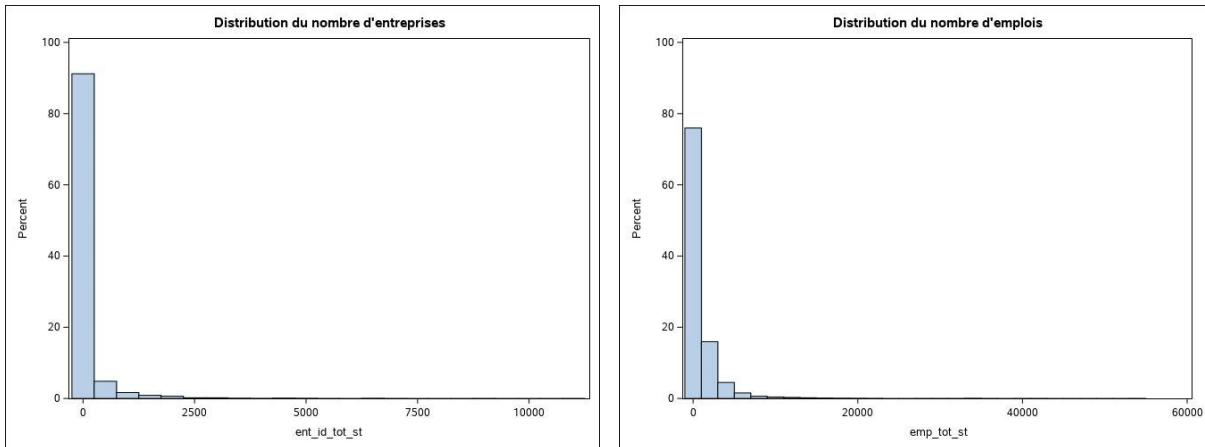
NOGA OFS 50	Description
01.3	Non modélisé ici : Agriculture, sylviculture et pêche
05.9	Industries extractives
10.2	Industries alimentaires et du tabac
13.5	Industries du textile et de l'habillement
16.8	Industries du bois et du papier ; imprimerie
19.2	Cokéfaction, raffinage et industrie chimique
21	Industrie pharmaceutique
22.3	Industries du caoutchouc et du plastique
24.5	Fabrication de produits métalliques
26	Fabrication de produits informatiques et électroniques ; horlogerie
27	Fabrication d'équipements électriques
28	Fabrication de machines et équipements n.c.a
29.3	Fabrication de matériels de transport
31.3	Autres industries manufacturières; réparation et installation
35	Production et distribution d'énergie
36.9	Production et distribution d'eau; gestion des déchets
41.2	Construction de bâtiments et génie civil
43	Travaux de construction spécialisés
45	Commerce et réparation d'automobiles et de motocycles
46	Commerce de gros
47	Commerce de détail
49	Transports terrestres et transport par conduites
50.1	Transports par eau, transports aériens
52	Entreposage et services auxiliaires des transports
53	Activités de poste et de courrier
55	Hébergement
56	Restauration
58.6	Édition, audiovisuel et diffusion
61	Télécommunications
62.3	Activités informatiques et services d'information
64	Activités des services financiers
65	Assurance
66	Activités auxiliaires de services financiers et d'assurance
68	Activités immobilières
69	Activités juridiques et comptables
70	Activités des sièges sociaux ; conseil de gestion
71	Activités d'architecture et d'ingénierie
72	Recherche-développement scientifique
73.5	Autres activités spécialisées, scientifiques et techniques
79.2	Activités de services administratifs et de soutien
78	Activités liées à l'emploi
84	Non modélisé ici : Administration publique
85	Enseignement
86	Activités pour la santé humaine
87	Hébergement médico-social et social
88	Action sociale sans hébergement
90.3	Arts, spectacles et activités récréatives
94.6	Autres activités de services
Classes de taille	Description
Microentreprises	0-9 emplois
Petites entreprises	10-49 emplois
Moyennes entreprises	50-249 emplois
Grandes entreprises	+ de 249 emplois
Cantons	Description
ZH	Zürich
BE	Bern / Berne
LU	Luzern
UR	Uri

SZ	Schwyz
OW	Obwalden
NW	Nidwalden
GL	Glarus
ZG	Zug
FR	Fribourg / Freiburg
SO	Solothurn
BS	Basel-Stadt
BL	Basel-Landschaft
SH	Schaffhausen
AR	Appenzell Ausserrhoden
AI	Appenzell Innerrhoden
SG	St. Gallen
GR	Graubünden / Grigioni
AG	Aargau
TG	Thurgau
TI	Ticino
VD	Vaud
VS	Valais / Wallis
NE	Neuchâtel
GE	Genève
JU	Jura

B ANNEXE : DISTRIBUTION DES VARIABLES DÉPENDANTES

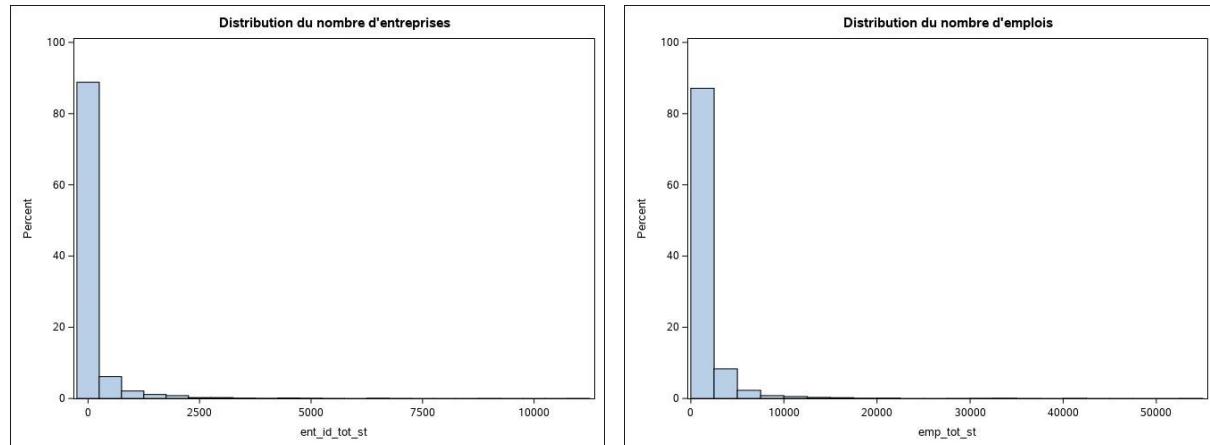
Les variables prédites (nombre d'emplois et d'unités légales) sont des variables appartenant à l'ensemble des nombres naturels. Les graphiques ci-dessous montrent leurs distributions pour l'année 2019.

Figure 1: Distribution du nombre d'unités légales et d'emplois



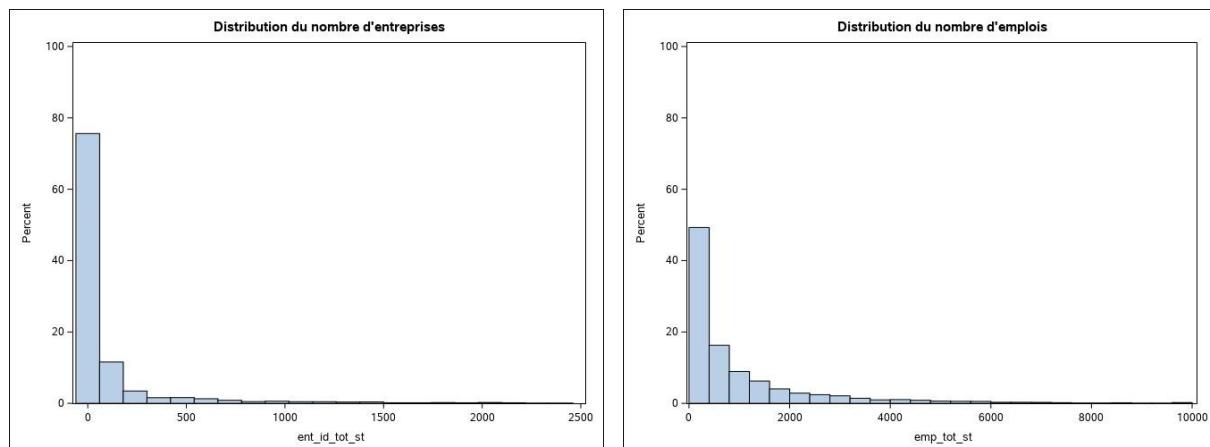
Il y a une grande masse de points à zéro (1'020 observations en 2019). En retirant ces observations nulles, l'histogramme ne changent pas d'aspect. À noter que les bornes des classes sont automatiquement modifiées.

Figure 2: Distribution du nombre d'unités légales et d'emplois (valeurs de zéro exclues)



En supprimant les valeurs extrêmes, l'asymétrie de la distribution se réduit partiellement.

Figure 3: Distribution du nombre d'unités légales]0 ;2'500[et d'emplois]0 ;10'000[



C ANNEXE : LISTE DES MODÈLES

Tableau 2: Liste des modèles sur les unités légales

Type de modèle	Nom	Intercept	Variables explicatives	Fonction Lien	Distribution	Méthode d'estimation	Procédure SAS
PAST	PAST19		TOTAL Statent t-1 avec paramètre égal à 1	IDENTITY	-		
TOTAL	REG1	Oui	TOTAL AVSQ2	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG1	Non	TOTAL AVSQ2	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG2	Non	TOTAL AVSQ2 TOTAL REE	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG3	Non	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	-	Ordinary least squares	REG
DIFF	REG1	Oui	TOTAL AVSQ2	IDENTITY	-	Ordinary least squares	REG
DIFF	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	-	Ordinary least squares	REG
DIFF	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	-	Ordinary least squares	REG
TOBIT	REG1	Oui	TOTAL AVSQ2	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	Normal	Maximum likelihood	QLIM
LOG	REG1	Oui	TOTAL AVSQ2	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG2	Oui	TOTAL AVSQ2 TOTAL REE	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	LOG	Normal	Maximum likelihood	GENMOD

Tableau 3: Liste des modèles sur les emplois

Type de modèle	Nom	Intercept	Variables explicatives	Fonction Lien	Distribution	Méthode d'estimation	Procédure SAS
PAST	PAST19		TOTAL Statent t-1 avec paramètre égal à 1	IDENTITY	-		
TOTAL	REG1	Oui	TOTAL AVSQ2	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG4	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 STATEM	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG5	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du	IDENTITY	-	Ordinary least squares	REG

			nombre d'emplois AVSQ2				
TOTAL	REG6	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	IDENTITY	-	Ordinary least squares	REG
TOTAL	REG7	Oui	TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG4	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 STATEM	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG5	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG6	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	IDENTITY	-	Ordinary least squares	REG
TOTAL	NOINTREG7	Oui	TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	IDENTITY	-	Ordinary least squares	REG
DIFF	REG1	Oui	TOTAL AVSQ2	IDENTITY	-	Ordinary least squares	REG
DIFF	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	-	Ordinary least squares	REG
DIFF	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	-	Ordinary least squares	REG
DIFF	REG4	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 STATEM	IDENTITY	-	Ordinary least squares	REG
TOBIT	REG1	Oui	TOTAL AVSQ2	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG2	Oui	TOTAL AVSQ2 TOTAL REE	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG4	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 STATEM	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG5	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2	IDENTITY	Normal	Maximum likelihood	QLIM
TOBIT	REG6	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois	IDENTITY	Normal	Maximum likelihood	QLIM

			AVSQ2 STATEM				
TOBIT	REG7	Oui	TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	IDENTITY	Normal	Maximum likelihood	QLIM
LOG	REG1	Oui	TOTAL AVSQ2	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG2	Oui	TOTAL AVSQ2 TOTAL REE	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG3	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG4	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 STATEM	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG5	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG6	Oui	TOTAL AVSQ2 TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	LOG	Normal	Maximum likelihood	GENMOD
LOG	REG7	Oui	TOTAL REE TOTAL Statent t-1 Changement du nombre d'emplois AVSQ2 STATEM	LOG	Normal	Maximum likelihood	GENMOD

D ANNEXE : RÉSULTATS MODÉLISATION

D.A RÉSULTATS SELON LE ROOT MEAN SQUARED FORECAST ERROR ET L'ÉCART TOTAL RELATIF

Les tableaux de cette section sont composés ainsi :

- *Type de modèle* indique le type de modèle utilisé. (Voir section 4.1.2)
- *Nom* : Nom utilisé pour désigner un modèle. Le nom avec une étoile correspond au modèle sélectionné par une méthode de sélection STEPWISE (FORWARD pour le *TOBIT*).
- *RMSFE niveau canton-NOGA-classe de taille* : Indicateur global de qualité calculé au niveau canton-NOGA OFS 50-classe de taille.
- *RMSFE niveau canton-NOGA*: Indicateur global de qualité calculé au niveau canton-NOGA OFS 50.
- *Écart total relatif* indique l'écart au niveau le plus agrégé entre la prédiction et la STATENT publiée.
- *Nombre de prédictions négatives* indique si le modèle prédit des nombres négatifs au niveau canton-NOGA-classe de taille.
- Les autres colonnes (*Intercept, Total AVSQ2...*) indiquent la présence ou non des variables explicatives et la significativité des coefficients associés.

Le modèle en bleu correspond au modèle de référence. En vert, les indicateurs de qualité avec une meilleure performance que le modèle de référence. En jaune, le modèle retenu.

Tableau 4: Résultats au niveau global de la modélisation du nombre d'unités légales en 2019

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercept	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Nombre de prédictions négatives
PAST	PAST19	14.76	30.46	-1.63%					0
TOTAL	REG1	40.56	88.42	3.42%	XXX	XXX			0
TOTAL	REG2	19.93	42.63	-0.90%	XX	XXX	XXX		1027
TOTAL	REG3*	10.28	21.11	-0.74%	XXX	XXX	XXX	XXX	1017
TOTAL	NOINTREG1	40.74	87.61	0.68%		XXX			0
TOTAL	NOINTREG2	19.94	42.55	-0.38%		XXX	XXX		0
TOTAL	NOINTREG3*	10.27	20.93	-0.37%		XXX	XXX	XXX	39
DIFF	REG1	14.63	29.73	-0.96%	XXX	-			0
DIFF	REG2	16.37	32.68	-1.58%	XXX	XXX	XXX		3
DIFF	REG3*	13.75	27.43	-1.53%	-	XXX	XXX	XXX	39
TOBIT	REG1	41.67	89.13	-2.54%	XXX	XXX			0
TOBIT	REG2	20.04	44.13	-2.71%	XXX	XXX	XXX		0
TOBIT	REG3*	10.44	22.55	-1.81%	XXX	XXX	XXX	XXX	0
LOG	REG1	38.92	86.11	3.98%	XXX	XXX			0
LOG	REG2	19.68	43.05	-2.09%	XXX	XXX	XXX		0
LOG	REG3*	10.38	21.49	-0.98%	XXX	XXX	XXX	XXX	0

Note : Le modèle PAST19 correspond à publier les chiffres 2018 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

Tableau 5: Résultats au niveau global de la modélisation du nombre d'unités légales en 2020

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercept	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Nombre de prédictions négatives
PAST	PAST20	10.48	20.15	0.02%					0
TOTAL	REG1	54.46	117.65	-1.92%	XXX	XXX			0
TOTAL	REG2	24.63	49.44	2.84%	XXX	XXX	XXX		1137
TOTAL	REG3*	14.66	28.68	0.96%	-	XXX	XXX	XXX	978
TOTAL	NOINTREG1	54.55	119.10	-4.79%		XXX			0
TOTAL	NOINTREG2	24.64	50.16	3.44%		XXX	XXX		0
TOTAL	NOINTREG3*	14.66	28.71	1.03%		XXX	XXX	XXX	39
DIFF	REG1	12.29	24.71	0.37%	XXX	XXX			0
DIFF	REG2	12.10	24.90	1.33%	XXX	XXX	XXX		0
DIFF	REG3	15.11	30.24	1.57%	XXX	XXX	-	XXX	27
TOBIT	REG1	54.88	121.73	-6.92%	XXX	XXX			0
TOBIT	REG2	24.80	47.62	1.15%	XXX	XXX	XXX		0
TOBIT	REG3*	14.77	28.48	-0.15%	XXX	XXX	XXX	XXX	0
LOG	REG1	53.10	114.11	-1.30%	XXX	XXX			0
LOG	REG2	24.70	47.85	2.32%	XXX	XXX	XXX		0
LOG	REG3*	14.79	28.92	1.14%	XXX	XXX	XXX	XXX	0

Note : Le modèle PAST20 correspond à publier les chiffres 2019 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

Tableau 6: Résultats au niveau global de la modélisation du nombre d'unités légales en 2021

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercept	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Nombre de prédictions négatives
PAST	PAST21	13.46	28.43	-1.72%					0
TOTAL	REG1	43.43	92.98	3.91%	XXX	XXX			0
TOTAL	REG2	19.32	39.57	-0.15%	XXX	XXX	XXX		1493
TOTAL	REG3*	13.57	28.83	-1.71%	-	X	XXX	XXX	1017
TOTAL	NOINTREG1	43.49	91.80	1.38%		XXX			0
TOTAL	NOINTREG2	19.34	39.89	0.97%		XXX	XXX		0
TOTAL	NOINTREG3*	13.58	28.79	-1.66%		X	XXX	XXX	23
DIFF	REG1	13.03	27.36	-1.51%	-	XXX			0
DIFF	REG2	12.51	26.52	-1.77%	XX	XXX	XXX		1037
DIFF	REG3*	12.24	26.16	-1.78%	XXX	XXX	XXX	XXX	1034
TOBIT	REG1	44.32	93.08	-2.27%	XXX	XXX			0
TOBIT	REG2	19.55	40.40	-1.95%	XXX	XXX	XXX		0
TOBIT	REG3*	13.68	30.97	-2.97%	XXX	X	XXX	XXX	0
LOG	REG1	43.10	93.38	5.37%	XXX	XXX			0
LOG	REG2	18.52	37.83	-1.48%	XXX	XXX	XXX		0
LOG	REG3	13.63	28.72	-1.39%	XX	X	XXX	XXX	0

Note : Le modèle PAST21 correspond à publier les chiffres 2020 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

Tableau 7: Résultats au niveau global de la modélisation du nombre d'emplois en 2019

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercept	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Changement du nombre d'emplois AVSQ2	STATEM	Nombre de prédictions négatives
PAST	PAST19	163.29	312.95	-1.25%							0
TOTAL	REG1	837.90	1759.46	5.74%	XXX	XXX					0
TOTAL	REG2	174.62	360.39	0.71%	XXX	-	XXX				0
TOTAL	REG3	152.23	283.23	0.17%	XX	-	XXX	XXX			0
TOTAL	REG4	152.13	282.97	0.17%	XX	-	XXX	XXX		-	0
TOTAL	REG5	154.74	288.28	0.18%	XXX	-	XXX	XXX	XXX		0
TOTAL	REG6	154.64	287.96	0.18%	XXX	-	XXX	XXX	XXX	-	2
TOTAL	REG7	154.89	288.54	0.19%	XXX		XXX	XXX	XXX	-	2
TOTAL	NOINTREG4	152.34	283.04	-0.30%		-	XXX	XXX		-	32
TOTAL	NOINTREG5	154.94	288.30	-0.31%		-	XXX	XXX	XXX		26
TOTAL	NOINTREG6	154.83	287.98	-0.31%		-	XXX	XXX	XXX	-	26
TOTAL	NOINTREG7	155.04	288.48	-0.30%			XXX	XXX	XXX	-	11
DIFF	REG1	162.56	307.29	0.18%	XXX	XXX					0
DIFF	REG2	161.92	306.04	0.20%	XXX	XXX	-				0
DIFF	REG3	160.90	296.13	0.21%	XXX	XXX	-	XXX			14
DIFF	REG4*	160.88	296.12	0.21%	XXX	XXX	-	XXX		-	14
TOBIT	REG1	850.24	1797.99	-9.38%	XXX	XXX					0
TOBIT	REG2	182.67	393.98	-4.44%	XXX	-	XXX				0
TOBIT	REG3	155.36	296.71	-2.06%	XXX	-	XXX	XXX			0
TOBIT	REG4	155.27	296.45	-2.06%	XXX	-	XXX	XXX		-	0
TOBIT	REG5	157.66	300.99	-2.04%	XXX	-	XXX	XXX	XXX		0
TOBIT	REG6	157.56	300.67	-2.04%	XXX	-	XXX	XXX	XXX	-	0
TOBIT	REG7	157.63	300.83	-2.04%	XXX		XXX	XXX	XXX	-	0
LOG	REG1	839.72	1772.02	6.92%	XXX	XXX					0
LOG	REG2	171.91	347.03	0.45%	XXX	-	XXX				0
LOG	REG3	156.85	290.91	0.28%	XXX	XXX	XXX	XXX			0
LOG	REG4	157.13	291.96	0.34%	XXX	XXX	XXX	XXX		XXX	0
LOG	REG5	166.51	312.55	0.27%	XXX	XXX	XXX	XXX	XXX		0
LOG	REG6	166.69	313.36	0.32%	XXX	XXX	XXX	XXX	XXX	XXX	0

LOG	REG7*	166.40	312.74	0.30%	XXX		XXX	XXX	XXX	XXX	XXX	0
-----	-------	--------	--------	-------	-----	--	-----	-----	-----	-----	-----	---

Note : Le modèle PAST19 correspond à publier les chiffres 2018 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

Tableau 8: Résultats au niveau global de la modélisation du nombre d'emplois en 2020

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercept	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Changement nombre AVSQ2	du d'emplois	STATE M	Nombre de prédictions négatives
PAST	PAST20	205.82	405.08	0.95%								0
TOTAL	REG1	841.16	1756.38	-0.19%	XXX	XXX						0
TOTAL	REG2	175.66	339.58	2.38%	XXX	XXX	XXX					0
TOTAL	REG3	168.12	341.73	2.33%	XXX	XXX	XXX	XXX				0
TOTAL	REG4	167.51	340.98	2.31%	XXX	XXX	XXX	XXX			XXX	0
TOTAL	REG5	168.28	342.63	2.34%	XXX	XXX	XXX	XXX	X			0
TOTAL	REG6*	167.68	341.89	2.32%	XXX	XXX	XXX	XXX	X		XXX	0
TOTAL	REG7	168.46	343.67	2.34%	XXX		XXX	XXX	-		XXX	0
TOTAL	NOINTREG4	166.81	334.08	1.38%		XXX	XXX	XXX			XXX	0
TOTAL	NOINTREG5	167.60	335.83	1.41%		XXX	XXX	XXX	X			0
TOTAL	NOINTREG6*	167.00	335.16	1.40%		XXX	XXX	XXX	X		XXX	0
TOTAL	NOINTREG7	167.79	336.71	1.38%			XXX	XXX	-		XXX	24
DIFF	REG1	204.90	408.09	2.12%	XXX	XXX						0
DIFF	REG2	187.24	367.99	2.24%	XXX	XXX	XXX					17
DIFF	REG3	187.36	368.11	2.23%	XXX	XXX	XXX	-				18
DIFF	REG4	186.15	369.55	2.20%	XXX	XXX	XXX	-			XXX	22
DIFF	REG5	186.90	367.70	2.20%	XXX	XX	XXX	-	X			18
DIFF	REG6	185.81	369.20	2.18%	XXX	XXX	XXX	-	-		XXX	22
DIFF	REG7	184.70	367.31	2.10%	XXX		XXX	-	XXX		XXX	22
TOBIT	REG1	858.22	1833.42	-13.39%	XXX	XXX						0
TOBIT	REG2	175.54	327.98	-0.61%	XXX	XXX	XXX					0
TOBIT	REG3	167.41	330.96	0.01%	XXX	XXX	XXX	XXX				0
TOBIT	REG4	166.81	330.32	0.00%	XXX	XXX	XXX	XXX			XXX	0
TOBIT	REG5	167.62	332.14	0.02%	XXX	XXX	XXX	XXX	XX			0

TOBIT	REG6*	167.02	331.50	0.01%	XXX	XXX	XXX	XXX	XX	XXX	0
TOBIT	REG7	167.84	333.23	0.02%	XXX		XXX	XXX	-	XXX	0
LOG	REG1	831.33	1725.78	-1.43%	XXX	XXX					0
LOG	REG2	181.41	360.56	2.44%	XXX	XXX	XXX				0
LOG	REG3	167.11	347.21	2.27%	XXX	-	XXX	XXX			0
LOG	REG4	166.82	346.87	2.23%	XXX	-	XXX	XXX		X	0
LOG	REG5	167.15	347.33	2.27%	XXX	-	XXX	XXX	-		0
LOG	REG6	166.86	346.99	2.22%	XXX	-	XXX	XXX	-	X	0
LOG	REG7	167.06	347.27	2.23%	XXX		XXX	XXX	-	X	0

Note : Le modèle PAST20 correspond à publier les chiffres 2019 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

Tableau 9: Résultats au niveau global de la modélisation du nombre d'emplois en 2021

Type de modèle	Nom	RMSFE niveau canton-NOGA-classe de taille	RMSFE niveau canton-NOGA	Écart total relatif	Intercep t	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	Changement nombre AVSQ2 du d'emplois	STATE M	Nombre de prédictions négatives
PAST	PAST21	191.70	404.02	-2.42%							0
TOTAL	REG1	736.22	1584.44	2.99%	XXX	XXX					0
TOTAL	REG2	173.64	355.18	-1.20%	-	XXX	XXX				989
TOTAL	REG3	151.71	320.16	-1.91%	XXX	XXX	XXX	XXX			1038
TOTAL	REG4	151.60	320.50	-1.90%	XX	XXX	XXX	XXX		XXX	1027
TOTAL	REG5	152.92	320.77	-1.82%	XX	XXX	XXX	XXX	XXX		1022
TOTAL	REG6*	152.68	320.73	-1.80%	XX	XXX	XXX	XXX	XXX	XXX	1014
TOTAL	REG7	155.45	326.62	-1.90%	XX		XXX	XXX	XXX	XXX	1023
TOTAL	NOINTREG4	151.42	318.74	-1.36%		XXX	XXX	XXX		XXX	0
TOTAL	NOINTREG5	152.88	319.35	-1.31%		XXX	XXX	XXX	XXX		0
TOTAL	NOINTREG6*	152.64	319.44	-1.33%		XXX	XXX	XXX	XXX	XXX	0
TOTAL	NOINTREG7	155.38	325.28	-1.46%			XXX	XXX	XXX	XXX	7
DIFF	REG1	192.18	401.61	-2.55%	XXX	XXX					1068
DIFF	REG2	159.41	316.37	-3.07%	XXX	XXX	XXX				1173
DIFF	REG3	156.97	311.46	-2.92%	XXX	XXX	XXX	XXX			1157
DIFF	REG4*	153.75	303.43	-2.87%	XXX	XXX	XXX	XXX		XXX	1152

DIFF	REG5	156.21	310.03	-2.93%	XXX	XXX	XXX	XXX	-		1157
DIFF	REG6	152.85	301.74	-2.88%	XXX	XXX	XXX	XXX	-	XXX	1153
DIFF	REG7	152.99	305.35	-3.29%	XXX		XXX	XXX	XXX	XXX	1182
TOBIT	REG1	742.38	1624.92	-10.62%	XXX	XXX					0
TOBIT	REG2	177.32	376.21	-3.78%	XXX	XXX	XXX				0
TOBIT	REG3	156.18	348.86	-4.24%	XXX	XXX	XXX	XXX			0
TOBIT	REG4	156.02	348.52	-4.22%	XXX	XXX	XXX	XXX		XXX	0
TOBIT	REG5	156.52	347.81	-4.18%	XXX	XXX	XXX	XXX	XXX		0
TOBIT	REG6*	156.28	347.29	-4.16%	XXX	XXX	XXX	XXX	XXX	XXX	0
TOBIT	REG7	159.16	353.90	-4.27%	XXX		XXX	XXX	XXX	XXX	0
LOG	REG1	742.77	1594.27	3.90%	XXX	XXX					0
LOG	REG2	178.61	366.71	-1.62%	XXX	XXX	XXX				0
LOG	REG3	156.18	336.73	-2.50%	XXX	XXX	XXX	XXX			0
LOG	REG4	156.55	337.59	-2.61%	XXX	XXX	XXX	XXX		XXX	0
LOG	REG5	157.87	338.46	-2.44%	XXX	XXX	XXX	XXX	X		0
LOG	REG6*	158.94	340.49	-2.54%	XXX	XXX	XXX	XXX	XX	XXX	0
LOG	REG7	160.00	343.56	-2.58%	XXX		XXX	XXX	XX	XXX	0

Note : Le modèle PAST21 correspond à publier les chiffres 2020 comme STATENT-flash. En bleu, le modèle de référence. En vert, les modèles performant mieux que le modèle de référence. En jaune, le modèle retenu.

XXX : Significatif à 1%, XX : Significatif à 5%, X : Significatif à 10%, - : Non significatif, Vide : Absent de la régression

D.B RÉSULTATS DU MODÈLE RETENU ET DE RÉFÉRENCE POUR LES UNITÉS LÉGALES

Tableau 10: Coefficients régression modèle TOTAL - NOINTREG3 sur le nombre d'unités légales

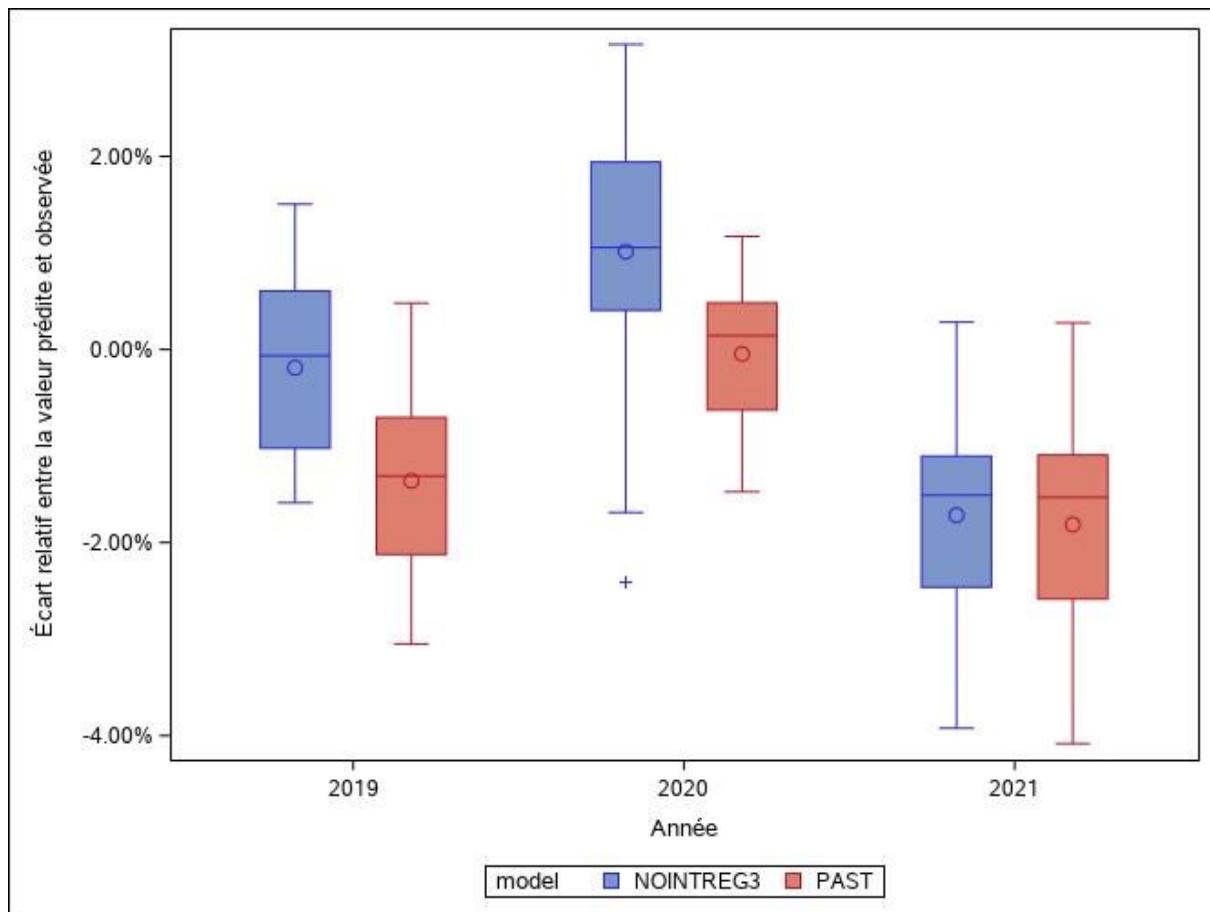
Type de modèle	Nom	Année	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1
TOTAL	NOINTREG3	2019	0.036***	-0.066***	1.045***
TOTAL	NOINTREG3	2020	0.101***	-0.115***	1.039***
TOTAL	NOINTREG3	2021	0.006***	0.053***	0.941***

*** : Significatif à 1%, ** : Significatif à 5%, * : Significatif à 10%

Tableau 11:Prédiction du nombre d'unités légales selon le canton

Canton	TOTAL – NOINTREG3						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif
AG	-0.22	0.00%	709.74	1.94%	-409.11	-1.11%	-483	-1.32%	125	0.34%	-460	-1.24%
AI	18.84	1.51%	19.57	1.57%	-19.37	-1.52%	6	0.48%	5	0.40%	-24	-1.89%
AR	-19.49	-0.48%	16.19	0.40%	11.32	0.28%	-41	-1.02%	27	0.67%	11	0.28%
BE	404.16	0.67%	542.81	0.91%	-67.42	-0.11%	-520	-0.87%	289	0.48%	-132	-0.22%
BL	69.66	0.42%	380.21	2.31%	-25.90	-0.16%	-141	-0.85%	140	0.85%	-35	-0.21%
BS	23.90	0.16%	209.06	1.40%	-349.38	-2.28%	-180	-1.21%	-37	-0.25%	-361	-2.35%
FR	-259.31	-1.51%	-112.95	-0.65%	-559.85	-3.12%	-432	-2.52%	-222	-1.28%	-580	-3.23%
GE	-371.41	-1.00%	502.15	1.35%	-871.47	-2.29%	-789	-2.12%	59	0.16%	-906	-2.38%
GL	15.67	0.60%	80.93	3.16%	-104.76	-3.92%	-23	-0.89%	30	1.17%	-109	-4.08%
GR	33.08	0.21%	169.62	1.10%	-185.91	-1.19%	-109	-0.70%	76	0.49%	-186	-1.19%
JU	-12.41	-0.27%	119.44	2.60%	-116.30	-2.46%	-74	-1.59%	45	0.98%	-122	-2.58%
LU	-202.76	-0.82%	320.93	1.29%	-321.72	-1.28%	-463	-1.87%	-84	-0.34%	-364	-1.44%
NE	-14.36	-0.13%	224.62	2.03%	-398.61	-3.47%	-165	-1.49%	14	0.13%	-415	-3.61%
NW	-35.78	-1.07%	-10.80	-0.32%	-22.03	-0.65%	-45	-1.35%	-21	-0.63%	-26	-0.77%
OW	22.74	0.82%	23.55	0.85%	-6.13	-0.22%	-16	-0.58%	-28	-1.01%	-9	-0.32%
SG	269.11	0.90%	532.45	1.78%	-359.24	-1.18%	-147	-0.49%	46	0.15%	-410	-1.35%
SH	12.80	0.25%	46.23	0.90%	-66.23	-1.27%	-32	-0.62%	-1	-0.02%	-72	-1.38%
SO	7.92	0.05%	415.40	2.87%	-268.19	-1.82%	-191	-1.31%	161	1.11%	-290	-1.97%
SZ	-138.90	-1.05%	-14.76	-0.11%	-203.56	-1.50%	-320	-2.43%	-182	-1.36%	-220	-1.62%
TG	-185.02	-1.12%	103.13	0.62%	-340.93	-2.02%	-415	-2.51%	22	0.13%	-371	-2.20%
TI	220.30	0.65%	-580.68	-1.69%	-529.48	-1.52%	-140	-0.41%	-494	-1.44%	-471	-1.35%
UR	17.23	0.93%	2.60	0.14%	-14.26	-0.76%	-3	-0.16%	-15	-0.80%	-16	-0.85%
VD	-796.96	-1.59%	1286.67	2.57%	-1665.98	-3.22%	-1531	-3.05%	205	0.41%	-1739	-3.36%
VS	-328.42	-1.40%	168.32	0.71%	-805.50	-3.30%	-503	-2.14%	-97	-0.41%	-803	-3.29%
ZG	-175.85	-1.02%	-420.46	-2.41%	-617.99	-3.42%	-423	-2.46%	-257	-1.47%	-620	-3.43%
ZH	-666.36	-0.64%	1054.46	1.01%	-1194.81	-1.13%	-2003	-1.92%	315	0.30%	-1152	-1.09%

Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Figure 4: Distribution de l'écart relatif du nombre d'unités légales selon le canton**Tableau 12: Prédiction du nombre d'unités légales selon la classe de taille**

	TOTAL – NOINTREG3						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
Classe de taille	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif
Micro-entreprises	-2210.47	-0.44%	4413.46	0.88%	-8029.46	-1.58%	-8582	-1.72%	-629	-0.13%	-8448	-1.66%
Petites entreprises	152.50	0.29%	926.21	1.78%	-1138.19	-2.14%	-427	-0.81%	421	0.81%	-1092	-2.05%
Moyennes entreprises	-21.39	-0.21%	371.46	3.79%	-294.62	-2.92%	-142	-1.41%	270	2.76%	-291	-2.89%
Grandes entreprises	-12.46	-0.66%	77.27	4.24%	-50.55	-2.70%	-32	-1.70%	59	3.24%	-51	-2.72%

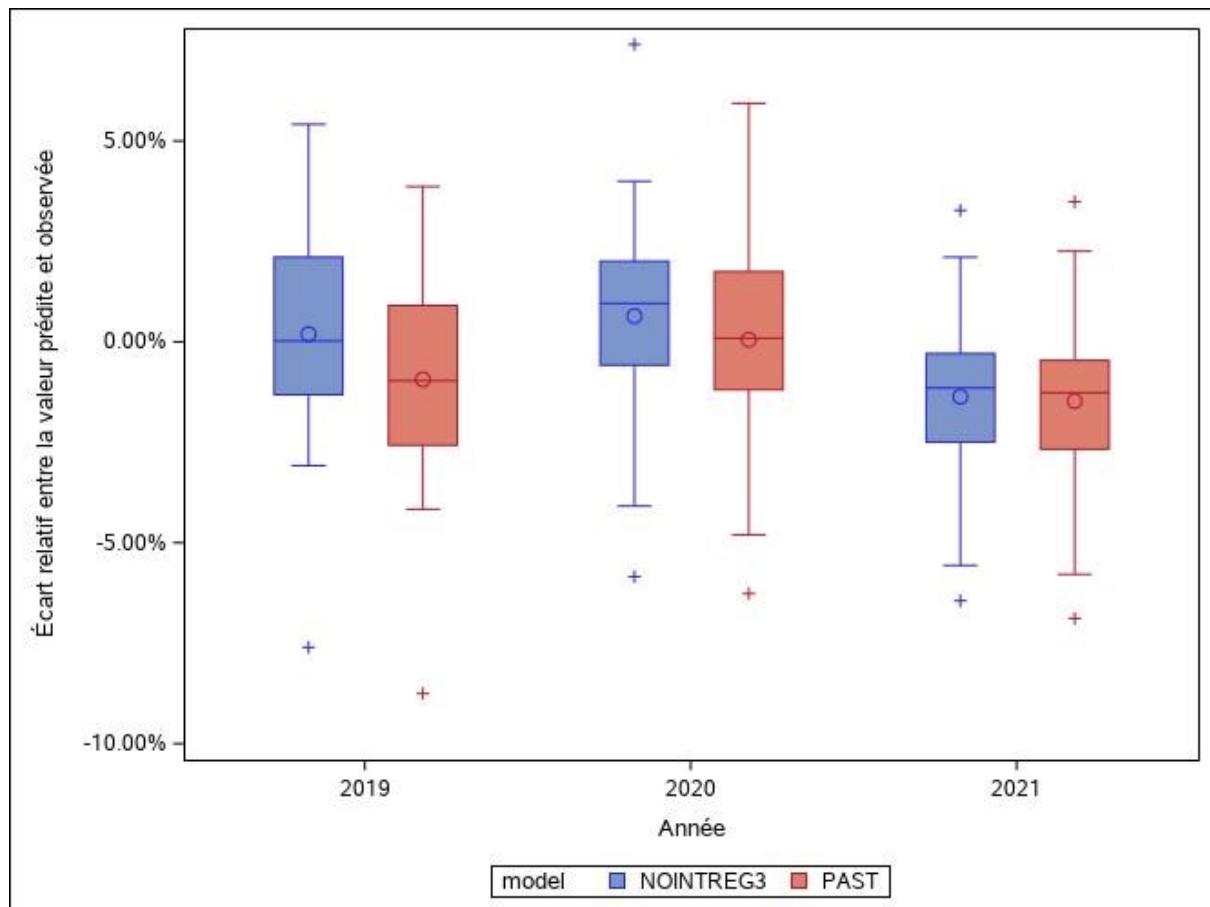
Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Tableau 13: Prédiction du nombre d'unités légales selon le code NOGA

	TOTAL – NOINTREG3						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
NOGA	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif
05.9	6.81	2.58%	-2.25	-0.84%	8.44	3.27%	4	1.52%	-3	-1.12%	9	3.49%
10.2	-27.62	-0.59%	-16.21	-0.34%	-168.66	-3.45%	-85	-1.83%	-65	-1.38%	-176	-3.60%
13.5	-1.53	-0.05%	28.19	0.99%	16.05	0.56%	-35	-1.22%	0	0.00%	6	0.21%
16.8	180.22	1.92%	287.55	3.12%	91.64	1.01%	64	0.68%	178	1.93%	88	0.97%

19.2	-15.30	-2.04%	-46.63	-5.84%	-55.23	-6.44%	-22	-2.94%	-50	-6.27%	-59	-6.88%
21	-22.59	-7.61%	-12.75	-4.09%	-15.55	-4.74%	-26	-8.75%	-15	-4.81%	-16	-4.88%
22.3	42.06	2.13%	56.16	2.90%	-6.48	-0.33%	22	1.11%	42	2.17%	-9	-0.46%
24.5	156.91	2.11%	144.20	1.96%	103.42	1.42%	67	0.90%	70	0.95%	103	1.42%
26	45.59	2.36%	17.75	0.93%	23.76	1.25%	28	1.45%	17	0.89%	22	1.16%
27	0.49	0.06%	11.92	1.49%	-7.64	-0.94%	-8	-0.99%	8	1.00%	-9	-1.11%
28	42.82	2.18%	60.69	3.15%	17.22	0.90%	19	0.97%	41	2.13%	16	0.84%
29.3	-13.32	-2.86%	-2.73	-0.58%	-5.20	-1.09%	-19	-4.09%	-7	-1.48%	-6	-1.26%
31.3	62.70	0.87%	188.08	2.65%	-73.75	-1.03%	-20	-0.28%	124	1.75%	-91	-1.27%
35	10.73	1.38%	-18.81	-2.35%	-13.95	-1.71%	3	0.39%	-22	-2.75%	-14	-1.72%
36.9	43.73	2.92%	-15.50	-1.02%	-42.56	-2.72%	30	2.00%	-20	-1.32%	-44	-2.82%
41.2	-1.50	-0.02%	-282.83	-2.88%	-133.86	-1.34%	-115	-1.21%	-333	-3.39%	-129	-1.30%
43	347.33	0.89%	238.34	0.61%	-307.93	-0.78%	-142	-0.36%	-164	-0.42%	-278	-0.70%
45	50.09	0.31%	196.63	1.21%	-198.35	-1.21%	-122	-0.75%	30	0.19%	-207	-1.26%
46	523.34	2.35%	192.49	0.88%	325.21	1.51%	375	1.69%	330	1.51%	288	1.33%
47	570.35	1.67%	233.83	0.69%	-375.24	-1.09%	217	0.64%	41	0.12%	-466	-1.35%
49	-31.07	-0.27%	805.41	7.41%	11.44	0.11%	-199	-1.73%	646	5.94%	-1	-0.01%
50.1	12.77	3.70%	4.82	1.43%	-15.61	-4.40%	10	2.90%	7	2.07%	-17	-4.79%
52	-21.81	-1.31%	0.44	0.03%	-25.93	-1.54%	-35	-2.11%	0	0.00%	-28	-1.66%
53	12.96	2.76%	10.34	2.25%	-5.76	-1.24%	8	1.71%	9	1.96%	-5	-1.08%
55	60.17	1.14%	171.27	3.31%	-12.91	-0.25%	3	0.06%	124	2.40%	-9	-0.17%
56	-88.03	-0.38%	391.09	1.70%	-136.55	-0.59%	-343	-1.47%	202	0.88%	-110	-0.47%
58.6	-67.39	-1.33%	17.89	0.35%	-299.01	-5.56%	-127	-2.51%	-7	-0.14%	-311	-5.79%
61	14.60	3.47%	12.92	3.18%	-2.19	-0.54%	11	2.61%	15	3.69%	-3	-0.73%
62.3	-329.22	-1.77%	-299.13	-1.59%	-498.07	-2.56%	-513	-2.76%	-303	-1.61%	-558	-2.87%
64	78.83	1.36%	-131.25	-2.24%	-132.62	-2.20%	13	0.22%	-70	-1.19%	-151	-2.51%
65	32.22	5.42%	5.06	0.86%	-5.57	-0.93%	23	3.87%	3	0.51%	-6	-1.00%
66	-6.43	-0.07%	-27.37	-0.28%	101.05	1.04%	-93	-0.95%	-19	-0.19%	91	0.93%
68	-388.01	-2.14%	-334.74	-1.81%	-351.52	-1.87%	-567	-3.12%	-292	-1.58%	-390	-2.07%
69	182.86	0.81%	295.52	1.30%	-160.68	-0.70%	-100	-0.44%	27	0.12%	-181	-0.79%
70	-686.80	-2.86%	-596.03	-2.41%	-763.99	-2.99%	-959	-3.99%	-668	-2.70%	-828	-3.24%
71	308.74	1.25%	190.74	0.78%	-71.95	-0.29%	51	0.21%	133	0.54%	-117	-0.48%
72	-56.12	-3.08%	-69.39	-3.67%	-106.20	-5.29%	-71	-3.90%	-72	-3.80%	-114	-5.68%
73.5	-402.85	-1.49%	338.85	1.25%	-694.49	-2.50%	-782	-2.89%	-16	-0.06%	-743	-2.67%
77.2	-246.72	-1.08%	75.65	0.33%	-514.45	-2.18%	-555	-2.42%	-187	-0.81%	-518	-2.19%
78	-6.97	-0.32%	42.24	2.01%	-16.99	-0.80%	-28	-1.30%	46	2.19%	-20	-0.94%
85	-614.39	-2.72%	396.61	1.76%	-549.05	-2.38%	-939	-4.16%	15	0.07%	-553	-2.40%
86	-712.03	-1.15%	1360.30	2.19%	-1901.46	-2.98%	-1831	-2.97%	-310	-0.50%	-1823	-2.86%
87	55.47	2.56%	44.13	2.07%	43.90	2.11%	34	1.57%	40	1.88%	47	2.26%
88	-80.68	-0.95%	111.72	1.33%	-151.84	-1.77%	-199	-2.35%	26	0.31%	-144	-1.68%
90.3	-543.12	-2.17%	968.50	4.00%	-541.14	-2.18%	-850	-3.40%	751	3.10%	-596	-2.40%
94.6	-570.11	-1.06%	744.73	1.38%	-1892.57	-3.40%	-1380	-2.57%	-181	-0.34%	-1822	-3.28%

Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Figure 5: Distribution de l'écart relatif du nombre d'unités légales selon le code NOGA

D.C RÉSULTATS DU MODÈLE RETENU ET DE RÉFÉRENCE POUR LES EMPLOIS

Tableau 14: Coefficients régression modèle TOTAL - NOINTREG4 sur le nombre d'emplois

Type de modèle	Nom	Année	TOTAL AVSQ2	TOTAL REE	TOTAL Statent t-1	STATEM
TOTAL	NOINTREG4	2019	-0.002	0.096***	0.914***	0.002
TOTAL	NOINTREG4	2020	0.011***	0.443***	0.546***	0.008***
TOTAL	NOINTREG4	2021	0.024***	0.628***	0.328***	0.023***

*** : Significatif à 1%, ** : Significatif à 5%, * : Significatif à 10%

Tableau 15:Prédiction du nombre d'emplois selon le canton

Canton	TOTAL – NOINTREG4						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	
AG	737.89	0.26%	5439.65	1.91%	-1503.71	-0.52%	-1970	-0.68%	3342	1.17%	-5996	-2.06%
AI	62.02	0.83%	122.74	1.65%	-45.91	-0.60%	-5	-0.07%	18	0.24%	-222	-2.89%
AR	234.16	0.98%	407.13	1.72%	417.59	1.78%	-22	-0.09%	260	1.10%	183	0.78%
BE	1703.37	0.28%	11987.88	2.02%	3221.66	0.54%	-4587	-0.76%	7077	1.19%	-7412	-1.23%
BL	128.44	0.10%	1733.29	1.41%	-1208.55	-0.98%	-1044	-0.84%	1338	1.09%	-939	-0.76%
BS	102.03	0.04%	3163.95	1.32%	-4563.04	-1.87%	-2280	-0.94%	2920	1.22%	-4920	-2.02%
FR	-394.61	-0.30%	1619.52	1.25%	-2157.65	-1.60%	-1607	-1.23%	991	0.76%	-4774	-3.55%
GE	-4869.07	-1.59%	4129.99	1.36%	-7683.69	-2.46%	-7588	-2.48%	3416	1.13%	-9870	-3.16%
GL	-186.15	-0.98%	570.91	3.09%	-390.86	-2.04%	-358	-1.88%	602	3.26%	-746	-3.89%
GR	22.38	0.02%	1578.58	1.54%	-2171.73	-2.06%	-817	-0.79%	1251	1.22%	-3219	-3.05%

JU	-335.71	-1.01%	586.59	1.82%	-876.50	-2.63%	-529	-1.60%	924	2.87%	-1103	-3.31%
LU	66.38	0.03%	2267.41	1.00%	-2427.15	-1.05%	-2198	-0.96%	671	0.29%	-4460	-1.92%
NE	-730.34	-0.82%	2100.89	2.42%	-2167.35	-2.42%	-1583	-1.78%	2007	2.31%	-2676	-2.99%
NW	172.15	0.78%	988.23	4.59%	198.21	0.92%	-44	-0.20%	616	2.86%	-96	-0.44%
OW	278.69	1.50%	387.58	2.12%	-44.74	-0.24%	119	0.64%	292	1.60%	-234	-1.27%
SG	265.66	0.10%	5192.65	1.91%	553.97	0.20%	-2487	-0.91%	1788	0.66%	-4693	-1.70%
SH	441.26	1.22%	495.04	1.37%	-218.42	-0.59%	76	0.21%	125	0.35%	-1008	-2.71%
SO	-182.60	-0.15%	2852.10	2.44%	-580.95	-0.49%	-1143	-0.96%	2387	2.04%	-1230	-1.04%
SZ	-582.61	-0.73%	1089.74	1.39%	-820.38	-1.01%	-1386	-1.74%	1048	1.33%	-2450	-3.02%
TG	-1203.34	-0.99%	-395.29	-0.32%	-626.39	-0.50%	-2430	-2.01%	-1734	-1.41%	-2451	-1.96%
TI	482.45	0.26%	203.41	0.11%	-4760.75	-2.47%	-1126	-0.60%	364	0.19%	-3853	-2.00%
UR	131.16	0.94%	6.55	0.05%	-260.86	-1.79%	-28	-0.20%	-159	-1.13%	-422	-2.90%
VD	-2180.76	-0.58%	574.21	0.15%	-10787.83	-2.72%	-5767	-1.53%	-1774	-0.47%	-17136	-4.32%
VS	-910.11	-0.66%	975.90	0.71%	-4934.98	-3.42%	-2241	-1.62%	507	0.37%	-5948	-4.13%
ZG	-4340.63	-3.29%	5951.83	4.79%	-3868.11	-2.94%	-6079	-4.60%	7724	6.21%	-7291	-5.54%
ZH	-3171.34	-0.31%	10735.89	1.07%	-17732.39	-1.73%	-12192	-1.21%	8351	0.83%	-23348	-2.28%

Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Figure 6: Distribution de l'écart relatif du nombre d'emplois selon le canton

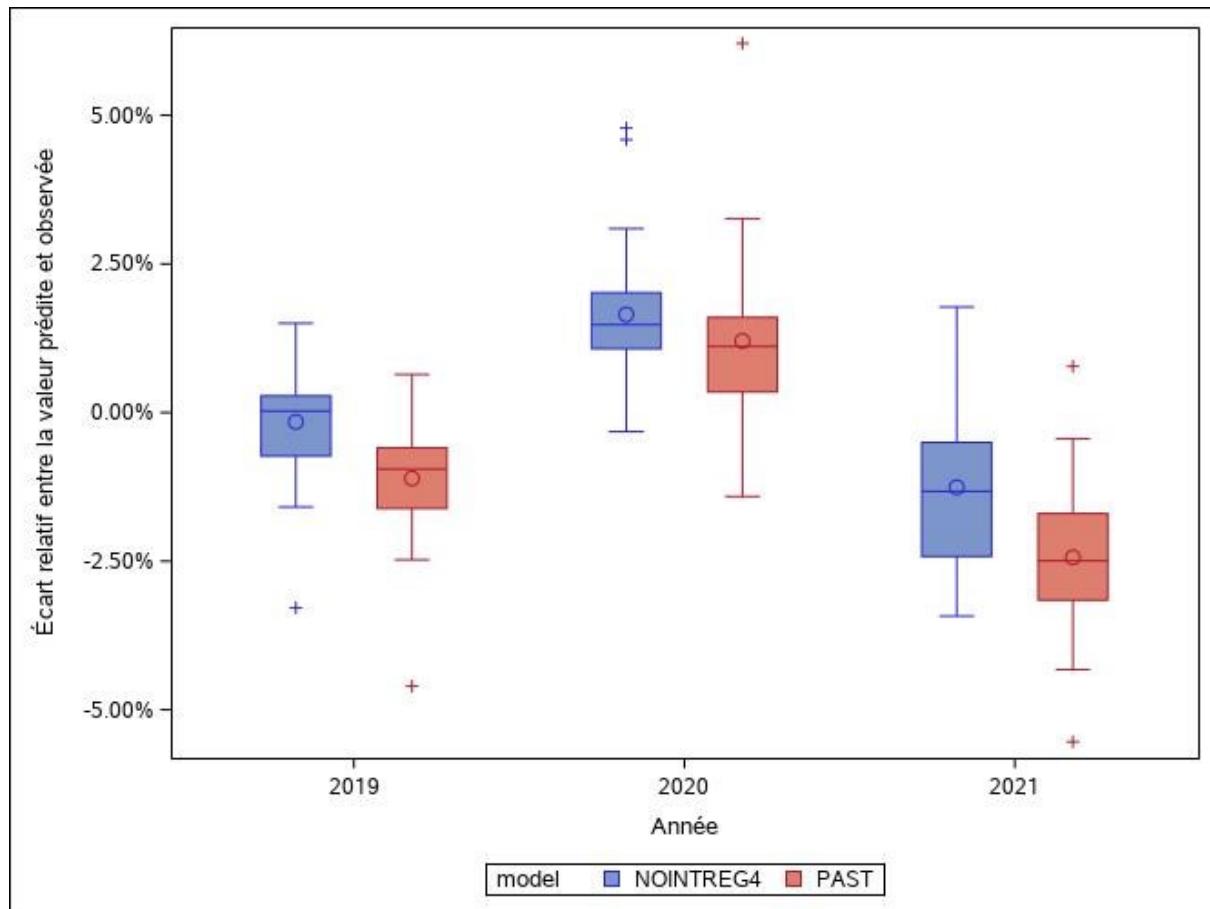


Tableau 16:Prédiction du nombre d'emplois selon la classe de taille

	TOTAL – NOINTREG4						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
Classe de taille	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif
Micro-entreprises	837.82	0.08%	13683.15	1.28%	1716.20	0.16%	-9566	-0.90%	-3187	-0.30%	-12262	-1.14%
Petites entreprises	-2412.49	-0.24%	9706.51	0.96%	-22705.52	-2.21%	-10347	-1.01%	14268	1.42%	-23269	-2.26%

Moyennes entreprises	-6612.27	-0.66%	16156.18	1.65%	-24720.15	-2.45%	-14488	-1.45%	21354	2.18%	-30808	-3.06%
Grandes entreprises	-6072.27	-0.37%	25220.54	1.54%	-19731.03	-1.17%	-24915	-1.51%	11917	0.73%	-49975	-2.97%

Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Tableau 17: Prédiction du nombre d'emplois selon le code NOGA

NOGA	TOTAL – NOINTREG4						Modèle de référence PAST					
	2019		2020		2021		2019		2020		2021	
	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif	Écart	Écart relatif
05.9	-31.93	-0.69%	18.27	0.40%	-66.42	-1.40%	-70	-1.51%	37	0.80%	-139	-2.93%
10.2	-289.05	-0.28%	2285.83	2.28%	-406.77	-0.40%	-1307	-1.28%	2002	1.99%	-1511	-1.48%
13.5	241.85	1.66%	749.98	5.33%	528.33	3.78%	79	0.54%	527	3.74%	114	0.82%
16.8	2443.38	3.82%	2902.71	4.72%	1569.88	2.56%	1791	2.80%	2427	3.94%	305	0.50%
19.2	270.94	0.89%	908.62	3.06%	-2475.14	-7.38%	-120	-0.39%	807	2.72%	-3830	-11.43%
21	-369.02	-0.77%	-1103.63	-2.23%	-837.24	-1.69%	-869	-1.81%	-1396	-2.82%	48	0.10%
22.3	857.14	2.06%	906.23	2.21%	148.96	0.36%	483	1.16%	658	1.61%	-49	-0.12%
24.5	1713.85	1.75%	3527.38	3.69%	609.95	0.63%	959	0.98%	2379	2.49%	-1569	-1.61%
26	268.67	0.24%	2195.59	1.98%	-330.71	-0.29%	-853	-0.76%	1428	1.29%	-1663	-1.48%
27	1338.61	4.46%	1528.84	5.38%	8.87	0.03%	993	3.31%	1626	5.72%	-192	-0.67%
28	593.68	0.75%	2798.45	3.63%	1407.95	1.83%	-139	-0.18%	2257	2.93%	157	0.20%
29.3	-239.10	-1.33%	1584.96	10.44%	50.66	0.33%	-388	-2.16%	2774	18.27%	-65	-0.43%
31.3	-581.61	-1.05%	-278.06	-0.49%	-402.47	-0.69%	-763	-1.37%	-1330	-2.34%	-1331	-2.29%
35	161.97	0.52%	-281.86	-0.89%	-394.12	-1.22%	-230	-0.74%	-761	-2.40%	-603	-1.86%
36.9	-247.04	-1.33%	-163.58	-0.88%	-432.81	-2.27%	-444	-2.39%	-24	-0.13%	-473	-2.48%
41.2	1916.23	1.74%	1693.34	1.54%	822.21	0.75%	772	0.70%	550	0.50%	132	0.12%
43	1446.43	0.59%	329.20	0.13%	-1291.52	-0.52%	-860	-0.35%	-1672	-0.68%	-2514	-1.01%
45	587.62	0.67%	1855.94	2.12%	255.75	0.29%	-474	-0.54%	857	0.98%	-595	-0.68%
46	1770.96	0.78%	4657.30	2.08%	-123.01	-0.05%	-591	-0.26%	2625	1.17%	-1724	-0.76%
47	3954.02	1.16%	2960.71	0.87%	-3079.81	-0.90%	455	0.13%	-210	-0.06%	-3696	-1.07%
49	-1178.72	-0.87%	1629.27	1.21%	93.26	0.07%	-2674	-1.97%	390	0.29%	-2501	-1.82%
50.1	-923.46	-4.79%	805.72	4.50%	171.21	1.00%	-1098	-5.69%	1397	7.81%	732	4.27%
52	-428.23	-1.00%	1766.55	4.29%	146.93	0.35%	-852	-1.98%	1801	4.38%	-845	-2.01%
53	1388.26	3.02%	1600.31	3.48%	1128.67	2.44%	657	1.43%	-25	-0.05%	-327	-0.71%
55	335.06	0.45%	9730.21	14.99%	852.20	1.24%	-487	-0.65%	9922	15.28%	-3959	-5.75%
56	-1520.84	-0.87%	14490.93	9.15%	517.49	0.31%	-3236	-1.86%	15641	9.88%	-6600	-4.00%
58.6	35.14	0.09%	663.00	1.78%	-1343.99	-3.54%	-115	-0.30%	1225	3.29%	-720	-1.90%
61	529.42	1.95%	720.24	2.70%	722.40	2.77%	218	0.80%	407	1.52%	608	2.33%
62.3	-5241.29	-4.39%	-3841.28	-3.11%	-7784.55	-5.95%	-6134	-5.14%	-4012	-3.25%	-7605	-5.81%
64	961.67	0.79%	1490.17	1.24%	-1261.25	-1.04%	12	0.01%	1639	1.37%	-1281	-1.06%
65	-243.21	-0.42%	-1034.02	-1.75%	-1636.80	-2.72%	-657	-1.13%	-1024	-1.73%	-1015	-1.69%
66	-1903.42	-2.41%	-2620.87	-3.20%	-2831.75	-3.40%	-2570	-3.25%	-2671	-3.27%	-1577	-1.89%
68	-3331.46	-4.33%	1492.09	2.01%	-910.32	-1.21%	-3584	-4.66%	2769	3.73%	-819	-1.09%
69	109.14	0.13%	959.44	1.11%	-840.74	-0.96%	-738	-0.86%	-83	-0.10%	-1439	-1.64%
70	-363.66	-0.38%	-243.41	-0.25%	-2934.99	-2.91%	-842	-0.88%	-927	-0.96%	-4380	-4.35%
71	-1264.85	-0.95%	-288.88	-0.21%	-1657.16	-1.20%	-2513	-1.88%	-1031	-0.77%	-3075	-2.23%
72	-1140.93	-4.58%	-2417.23	-8.96%	-608.43	-2.22%	-1337	-5.37%	-2067	-7.66%	-371	-1.36%
73.5	-80.33	-0.12%	1903.52	2.85%	-170.01	-0.25%	-808	-1.19%	893	1.34%	-1978	-2.87%
77.2	-1869.43	-0.94%	5985.85	3.08%	-1914.11	-0.95%	-4248	-2.14%	4112	2.12%	-6338	-3.16%
78	962.62	0.75%	8890.10	7.52%	-7579.96	-5.42%	-1139	-0.89%	9497	8.04%	-21731	-15.53%
85	-1923.10	-0.90%	457.01	0.21%	-5292.05	-2.38%	-3989	-1.86%	-1204	-0.56%	-6563	-2.95%

86	-3856.42	-0.91%	-9133.23	-2.09%	-14508.04	-3.24%	-7762	-1.83%	-12581	-2.88%	-11776	-2.63%
87	-870.71	-0.43%	-4286.84	-2.09%	-1994.46	-0.98%	-2747	-1.37%	-4371	-2.13%	645	0.32%
88	-2771.68	-2.51%	-621.94	-0.56%	-4671.02	-4.07%	-3792	-3.44%	-161	-0.15%	-4191	-3.65%
90.3	-2866.53	-3.06%	6362.88	7.23%	323.01	0.35%	-3799	-4.06%	5674	6.45%	-4188	-4.54%
94.6	-2609.88	-1.43%	1230.59	0.69%	-7018.61	-3.81%	-3506	-1.92%	3581	2.01%	-5822	-3.16%

Note : Les sous-estimations sont en orange, les surestimations en vert. Les bornes pour le code couleurs sont : -5%, -2%, -1%, 0%, 1%, 2% et 5%.

Figure 7: Distribution de l'écart relatif du nombre d'emplois selon le code NOGA

