

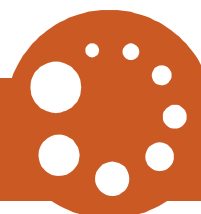


SAE-BESTA: approfondissement des estimations au niveau des cantons et analyse de la pertinence de l'extension aux branches économiques

Projet pilote dans le cadre de la stratégie d'innovation des données de l'OFS

Rapport final

EXPERIMENTAL STATISTICS



Neuchâtel, 2024

Éditeur: Office fédéral de la statistique (OFS)
Renseignements: besta@bfs.admin.ch
Rédaction: Francis Saucy, Alberto Columbano,
Daniel Kilchmann, Monika Ferster,
Anne Massiani, Djalel-Eddine Meskaldji,
Sophie Schmassmann
Domaine: 00 Bases statistiques et généralités
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

Sommaire

1	Objet de ce document	2
2	Résumé	2
3	Situation initiale	2
4	Objectifs	2
5	Processus et méthodes	3
5.1	« Business Understanding Overview »	3
5.2	« Data Understanding Overview »	3
5.3	« Data Preprocessing Overview »	4
5.4	« Modeling Overview »	5
5.5	Estimation de la variance des estimateurs SAE par le bootstrap	5
5.6	« Evaluation Overview »	5
5.6.1	Principes de la validation	5
5.6.2	Détermination des domaines validés sur la base du DIR2 et de ceux validés sur la base du CAS2	6
5.6.3	Benchmarking et facteurs de redressement	8
5.6.4	Gain de variance	8
6	Résultats régionaux	9
6.1	Grandes régions	9
6.2	Cantons et ville de ZH	10
6.2.1	Validation: critères locaux	10
6.2.2	Validation: critères globaux	15
6.2.3	Variance des estimations SAE	19
6.3	Discussion	19
7	Résultats pour les branches économiques	21
7.1	Secteurs et sections économiques	21
7.2	NOGA2	22
7.2.1	Validation : critères locaux	22
7.2.2	Synthèse des analyses de validation du modèle pour les branches économiques	25
8	Conclusions et recommandations	25
9	Annexes	26
	Annexe 1 : modèle SAE retenu (saenb)	26
	Annexe 2 : séries SAE (cantons et ville de Zurich)	27
	Annexe 3 : CV des estimations SAE cantonales, sans benchmarking	36
	Annexe 4 : validation du modèle pour les branches économiques	40
	Annexe 5 : résumé des CV des estimations directes (DIR2) pour les regroupements de divisions économiques NOGA_OFS50	41
	Annexe 6 : séries temporelles SAE pour les NOGA2	42
	Annexe 7 : résumé des résultats de la validation au niveau des NOGA2	69
	Annexe 8 : séries temporelles SAE pour les NOGA_OFS50	76
	Annexe 9 : résumé des résultats de la validation au niveau des NOGA_OFS50	92
	Annexe 10 : SAE-BESTA. Estimation de variance	96

1 Objet de ce document

Ce document a pour objet de résumer les travaux et résultats complémentaires du projet pilote « **NDS-WI-SAE-BESTA** » réalisés depuis la publication du Rapport intermédiaire de 2019¹.

2 Résumé

Sur la base des résultats préliminaires du projet initial, l'objectif était d'approfondir le développement d'un modèle SAE au niveau des cantons et d'évaluer son application au niveau des divisions économiques. Les résultats obtenus montrent qu'il est raisonnable d'envisager de publier des résultats au niveau cantonal sous forme de statistiques expérimentales sur la base de l'échantillon trimestriel de la BESTA actuelle, donc avec densification de la taille de l'échantillon demandée et financée par des cantons et des villes. Les analyses ont montré que l'extension au niveau des divisions économiques NOGA2 était globalement pertinente mais certaines divisions nécessitent encore des améliorations. Des recommandations sont proposées pour la publication de séries SAE cantonales, ainsi que pour l'amélioration des estimations directes dans le cadre de la production BESTA.

3 Situation initiale

Pour des raisons de précision, la Statistique de l'emploi (BESTA) ne propose plus de résultats cantonaux depuis la révision 2000. Les cantons ou les villes qui le souhaitent ont la possibilité de financer des compléments d'échantillon, afin d'obtenir des résultats régionaux d'une précision satisfaisante. Ces augmentations d'échantillon :

- **sont relativement chères** pour les cantons,
- **vont à l'encontre de l'objectif de réduction de la charge** auprès des entreprises mis en place à l'OFS dans le cadre du programme général des statistiques auprès des entreprises,
- **sont inévitables** : certains cantons, formant une grande région obtiennent gratuitement des résultats ou peuvent les obtenir indirectement.

Pour remédier à cette situation, la possibilité de produire des estimations du niveau absolu de l'emploi par cantons à l'aide de données existantes et des estimations sur « petits domaines » ou « small area estimation » (SAE) a été testée dans le cadre d'un projet pilote de la stratégie d'innovation des données de l'OFS. Des résultats prometteurs et encourageants ont été obtenus dans le cadre de ce projet pilote¹.

4 Objectifs

Les résultats prometteurs et encourageants obtenus dans le cadre du projet mentionné ci-dessus nous ont incité à approfondir cette approche et d'étudier la possibilité de l'étendre aux branches économiques avec les objectifs suivants :

1 Saucy, F. et al. (2019) SAE_BESTA : Evaluation du potentiel d'amélioration de la précision dans le cadre de la Statistique de l'emploi (BESTA),

- a) Approfondissement de l'approche pour les cantons
- b) Extension aux branches économiques au niveau NOGA2
- c) Recherche d'un modèle unique pour les cantons et les NOGA2
- d) Approfondissement des méthodes de validation du modèle
- e) Approfondissement des méthodes d'estimation de variance

5 Processus et méthodes

5.1 « Business Understanding Overview »

La Statistique de l'emploi (BESTA) est une statistique conjoncturelle. Elle repose sur une enquête par échantillon. Elle fournit des résultats trimestriels sur le niveau absolu de l'emploi, les places vacantes, les difficultés de recrutement de personnel selon le niveau de qualification et sur les prévisions d'évolution de l'emploi. Elle est intégrée dans le système des statistiques d'entreprises de l'OFS et applique les mêmes définitions de l'emploi que la STATENT et le Registre BUR. Les résultats sont ventilés par branches économiques (regroupements NOGA OFS50) et par grandes régions. Le plan d'échantillonnage et les méthodes d'estimation, sont conçus afin d'assurer les objectifs de précision décrits aux section 5.6 et annexe 2 du Rapport 2019¹.

5.2 « Data Understanding Overview »

Dans le contexte du précédent rapport, nous avons utilisé les données collectées sur la base des échantillons tirés dans le cadre de sondage 2015, résultant de l'adaptation de la BESTA au nouvel univers AVS/STATENT. Ce cadre de sondage a été utilisé du 2^{ème} trim. 2015 au 1^{er} trim. 2018. Dans le présent rapport, nous avons également utilisé les échantillons tirés dans le cadre de sondage 2018 pour la période du 1^{er} trim. 2018 au 4^{ème} trim. 2019, soit en tout 20 échantillons trimestriels.

La **Figure 1** présente l'organisation des travaux avec ses principaux processus, le flux des données, ainsi que les processus de validation des résultats, soit :

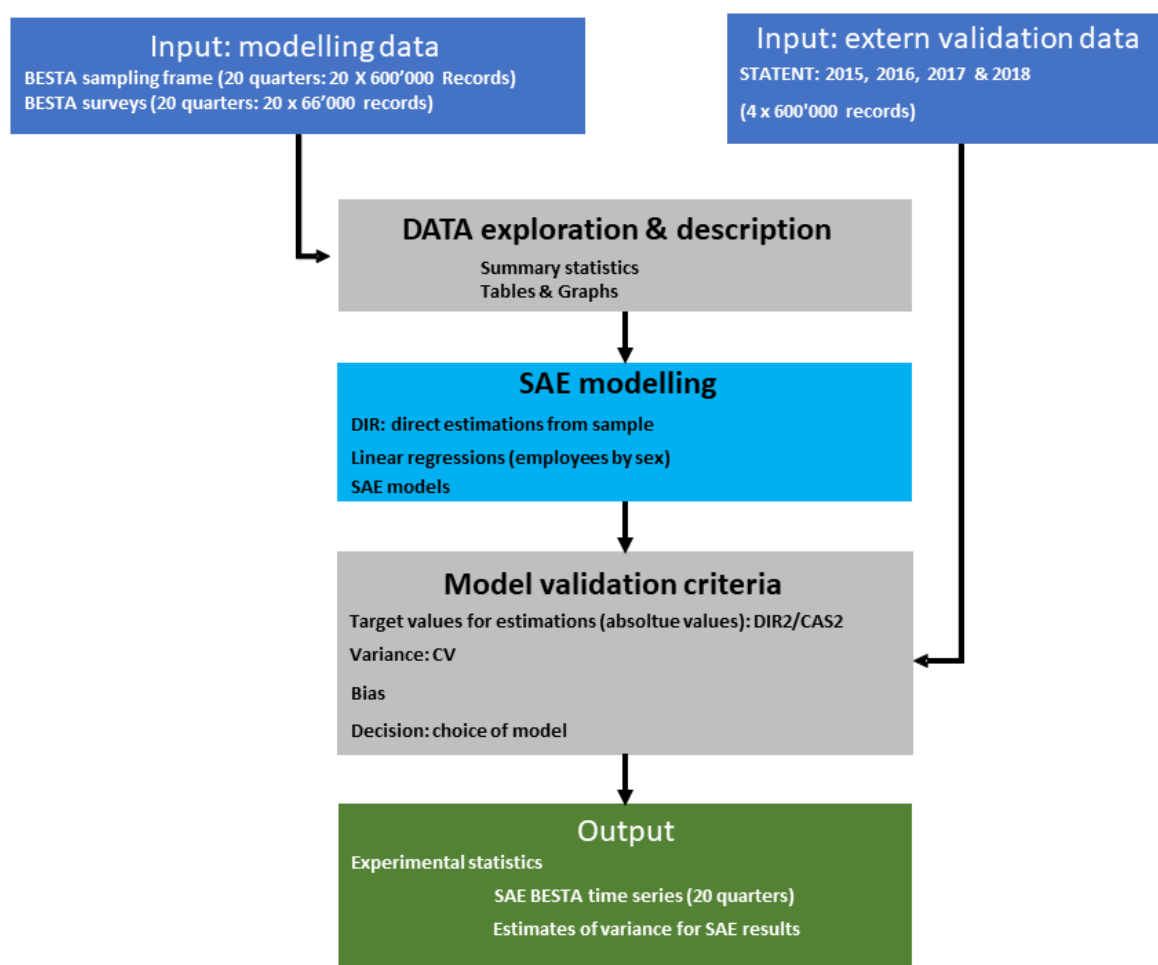
- **En input pour la modélisation**, nous disposons de 20 trimestres de données collectées dans des conditions comparables et sur deux échantillons qui ont été tirés au premier trimestre 2015 et au premier trimestre 2018 respectivement. Ces données ont été explorées et décrites pour en comprendre la structure dans le précédent rapport.

Comme dans le rapport précédent, la **modélisation** a été effectuée à l'aide de modèles de régression linéaire mixtes. Les résultats obtenus précédemment nous ont permis de raffiner l'un de ces modèles (s0) et de nous concentrer sur ce dernier (dénommé modèle saenb). Le détail du modèle est décrit à l'annexe 1. De ce modèle on tire des prédictions pour l'ensemble de la population. Cela permet d'obtenir des estimations SAE que l'on compare ensuite avec les données de validation.

- **En input pour la validation** : nous disposons
 - o des estimations directes basées sur les échantillons BESTA
 - o de la STATENT annuelle (4^{ème} trim. de l'année correspondante)
 - o du cadre de sondage trimestriel
 - o des objectifs de précision
 - o des estimations de variance des estimations directes
 - o des estimations de variance obtenues par une adaptation de la méthode du bootstrap aux estimateurs SAE utilisés. Cette partie a été développée dans le cadre de la présente phase du projet.

- Lors de la validation des résultats, les prédictions des différents modèles ont été comparées aux valeurs de référence retenues pour la validation. Les critères de validation ont été revus et approfondis.
- Enfin, en output, nous avons les résultats du modèle testé, évalué et validé (saenb) et les estimations SAE qui en découlent par canton et pour la ville de Zurich, notée « zs » ainsi qu'au niveau NOGA2. Les résultats sont synthétisés dans des tables et des graphiques. Dans ce qui suit, les estimations SAE sont obtenues en utilisant le modèle retenu, c.à.d. le modèle saenb.

Figure 1 Données et processus



5.3 « Data Preprocessing Overview »

Les données proviennent des environnements de production (BESTA et STATENT). Il s'agit de données déjà traitées, plausibilisées et validées. Elles n'ont donc pas nécessité de travaux préparatoires particuliers, si ce n'est les appariements habituels réalisés dans les processus de production. Cette préparation a été effectuée en SAS et les fichiers résultants ont été chargés dans l'environnement de travail (cloud Atlantica). Tous les calculs ont été effectués dans cet environne-

ment à l'aide du logiciel « R ». Les scripts et les procédures des calculs sont décrits à l'annexe 10 (chapitre 3).

5.4 « Modeling Overview »

La modélisation a été réalisée dans l'environnement de travail (cloud Atlantica) avec le logiciel « R ». Elle se base sur le modèle saenb pour modéliser l'emploi des entreprises de l'échantillon (environ 66'000 établissements). L'emploi du cadre (environ 600'000 établissements), ainsi que d'autres variables contenues dans le cadre, sont utilisés comme variables explicatives.

Deux familles d'estimateurs sont considérées dans ce rapport :

- Les estimateurs SAE
- Les estimateurs directs qui sont notamment utiles pour la validation (cf. section 5.6 pour plus de détails). Nous utilisons dans ce projet deux estimateurs directs :
 - o L'estimateur direct calculé comme simple somme pondérée par les poids d'extrapolation (DIR).
 - o L'estimateur par le ratio, qui multiplie l'estimateur DIR par le ratio dans le domaine entre l'emploi du cadre et l'estimation de celui-ci sur la base de l'échantillon (DIR2). En d'autres termes, l'estimateur DIR2 est calé sur l'emploi du cadre au niveau du domaine considéré, et par sexe.

Les estimations publiées par la section sont basées sur l'estimateur DIR. A noter que DIR et DIR2 sont très proches pour des domaines pour lesquels les poids d'extrapolation sont calés, par exemple les grandes régions pour l'emploi total. DIR et DIR2 peuvent en revanche différer sensiblement au niveau des cantons. L'estimateur DIR est utilisé dans la suite pour ce qui a trait aux questions de cohérence entre les estimations SAE et les résultats déjà publiés (par exemple le benchmarking), tandis que DIR2 est utilisé dans la phase de validation consistant à comparer pour des domaines suffisamment grands les estimations SAE à des estimations directes. Pour ces comparaisons, il semble en effet pertinent de se baser sur un estimateur direct calé sur le domaine considéré, plutôt que sur le DIR qui ne peut par principe pas être calé sur tous les domaines potentiellement considérés dans cette analyse, car des choix doivent être faits lors de la pondération en fonction des objectifs principaux de publication.

5.5 Estimation de la variance des estimateurs SAE par le bootstrap

L'application de la méthode SAE est généralement accompagnée d'une réduction de la variance par rapport aux estimations directes basées uniquement sur les données de l'échantillon. En revanche, avec cette méthode on s'attend à un risque de biais des estimations, qui dépend de la performance et des défauts du modèle.

L'estimation de la variance pour des plans d'échantillons complexes comme celui de la BESTA est notoirement très complexe. Nous avons retenu et développé ici une approche fondée sur un ré-échantillonnage avec remise des données, dite de « bootstrap ». Cette approche est décrite à l'annexe 10.

5.6 « Evaluation Overview »

5.6.1 Principes de la validation

Pour l'évaluation, l'approche de validation des résultats du modèle a été reconsidérée et affinée. L'objectif est de prendre prioritairement en compte des critères internes, à savoir les résultats des estimations directes lorsque leur précision (exprimée en termes de coefficient de variation) est jugée satisfaisante. Lorsque ce n'est pas le cas, la méthode utilisée dans le précédent rapport et fondée sur des critères externes a été retenue. En d'autres termes, lorsque l'estimation DIR2 n'atteint pas une précision suffisante, les estimations SAE sont comparées aux valeurs du cadre

adapté à la STATENT² (CAS, disponible uniquement pour les quatrièmes trimestres). La raison pour laquelle les estimations SAE sont comparées prioritairement aux estimations directes pour les domaines suffisamment grands est que les valeurs de référence basées sur le « cadre adapté à la STATENT », ne peuvent être considérées que comme des approximations des valeurs cibles de la BESTA, du fait des différences existant entre les deux univers. Afin de réduire ces différences, la version du cadre adapté à la STATENT utilisée dans le précédent rapport (CAS) est modifiée ici de la façon suivante : pour les unités répondantes, la valeur enregistrée dans le cadre adapté à la STATENT est remplacée par la valeur relevée lors de l'enquête. Cette adaptation du cadre adapté à la STATENT est appelée CAS2 dans le reste du document. Le CAS2 peut donc être perçu comme une énumération de la population cible, qui contient la valeur de la variable d'intérêt pour les unités répondantes, et une approximation de celle-ci pour les unités qui ne font pas partie de l'échantillon des répondants. Les valeurs de référence basées sur le CAS2 pour un domaine donné sont obtenues en additionnant les valeurs du CAS2 pour le domaine considéré. A noter que malgré le passage du CAS au CAS2, les valeurs de référence basées sur le CAS2 restent des approximations des valeurs cibles de la BESTA.

Pour les validations fondées sur les estimations directes, nous disposons de valeurs trimestrielles, alors que pour le CAS2, nous n'avons que celles du 4^{ème} trimestre par année (et pas pour 2019 pour des raisons de ressources). Les résultats sont synthétisés dans des tables et des graphiques.

De plus, afin de réaliser une évaluation la plus complète possible, deux types de critères sont examinés :

a) Critères locaux

L'impact de la méthode SAE sur les estimations est analysé en détail (valeurs extrêmes ou aberrantes, écarts importants par rapport à des valeurs de comparaison, évolutions particulières et notables telles que baisses ou augmentations spectaculaires de l'emploi pour un canton ou une NOGA2 donné(e)). Les résultats sont présentés sous forme de tables de synthèse et graphiquement dans les annexes.

b) Critères globaux

L'impact des méthodes SAE est également analysé plus globalement, par exemple en s'assurant que la somme des estimations SAE des cantons d'une même grande région est proche de l'estimation directe obtenue pour cette grande région et considérée comme fiable. Cela est lié au benchmarking décrit à la section 5.6.3.

Considérer à la fois des critères locaux et globaux est essentiel en cas d'estimation SAE. La quantification du gain global permis par l'application de méthodes SAE est en effet une information très utile, mais il est également primordial de s'assurer que ces méthodes ne fournissent pas de résultats déraisonnables pour certains petits domaines particuliers.

5.6.2 Détermination des domaines validés sur la base du DIR2 et de ceux validés sur la base du CAS2

Critères de précision retenus pour les estimations directes

Pour rappel, la précision des estimations directes déjà diffusées est jugée satisfaisante dans le cadre de ces travaux si leurs coefficients de variation (CV) respectent les critères suivants :

Emploi total : $CV < 0.5\%$

²Il s'agit d'un appariement entre le cadre de sondage et la STATENT. Pour les unités du cadre retrouvées dans la STATENT, la STATENT fournit une approximation de la valeur cible. Pour les autres, on utilise l'emploi total du cadre.

Emploi total par sections économiques (NOGA « lettres ») :	CV < 2.0%
Emploi total par grandes régions :	CV < 2.0%
Emploi total par grandes régions et par sexe :	CV < 3.0%
Emploi total par sections économiques et par sexe	CV < 3.0%

En termes de demi-longueur de l'intervalle de confiance à 95%, on obtient approximativement les critères suivants (le CV est remplacé par $1.96 \cdot CV$) :

Emploi total :	$\pm 1.0\%$
Emploi total par sections économiques :	$\pm 4.0\%$
Emploi total par grandes régions :	$\pm 4.0\%$
Emploi total par grandes régions et par sexe :	$\pm 6.0\%$
Emploi total par sections économiques et par sexe	$\pm 6.0\%$

Ecarts tolérés entre les estimations SAE et les valeurs de référence pour la validation des modèles :

Par analogie nous avons fixé des seuils du même ordre de grandeur pour l'écart entre les estimations SAE et les valeurs de référence, soit :

Emploi total par divisions économiques (NOGA2) :	$\pm 4.0\%$
Emploi total par divisions économiques (NOGA2) et par sexe	$\pm 6.0\%$
Emploi total par canton ou ville :	$\pm 4.0\%$
Emploi total par canton ou ville ventilés par sexe :	$\pm 6.0\%$

Ces seuils, s'appliquent aussi bien aux critères internes (estimations directes) qu'aux valeurs externes (STATENT, CAS et CAS2)

Détermination des domaines validés sur la base de l'estimateur direct par le ratio (DIR2) et de ceux validés sur la base du CAS2

Cette classification dépend des CV de l'estimateur par le ratio (DIR2) et des critères de précision rappelés ci-dessus. Idéalement, les critères devraient être respectés pour tous les trimestres pour que la valeur de référence pour un domaine soit donnée par l'estimateur direct par le ratio (DIR2). Pour la mise en œuvre de ces critères nous avons cependant défini des limites un peu plus souples afin de ne pas pénaliser excessivement un domaine qui serait proche de ces limites ou qui ne les excéderait que rarement.

Cette opérationnalisation a été mise en œuvre comme suit :

Pour les estimations tous sexes confondus, le modèle sera validé contre les estimations directes si :

- a) CV des estimations directes < 2% pour au moins un trimestre **and** < 3% pour tous les trimestres **(Critère and)**

Pour les estimations par sexe :

- b) CV des estimations directes < 3% pour au moins un trimestre **and** < 4.5% pour tous les trimestres **(Critère and)**

Les résultats obtenus pour les cantons sont présentés dans la table ci-après. La couleur verte signifie que les estimations SAE sont validées sur la base du DIR2 (estimateur direct par le ratio), tandis qu'elles le sont sur la base du CAS2 (approximation de la valeur cible basée notamment sur la STATENT) pour les cas apparaissant en rouge.

Critère AND	AG	AI	AR	BE	BL	BS	FR	GE	GL	GR	JU	LU	NE	NW	OW	SG	SH	SO	SZ	TG	TI	UR	VD	VS	ZG	ZH	zs
total	1	0	0	1	0	1	0	1	0	1	0	0	1	0	0	1	0	1	0	1	0	0	1	1	0	1	1
hommes	1	0	0	1	0	1	0	1	0	1	0	1	1	0	0	1	0	1	0	1	1	0	1	1	1	1	1
femmes	1	0	0	1	1	1	1	1	0	0	0	0	1	0	0	1	0	1	0	0	0	0	1	1	0	1	1

Exemple de lecture : pour le canton d'Argovie, les estimations SAE sont validées sur la base du DIR2 pour les hommes, les femmes et les deux sexes confondus. Pour le canton de Glaris, les trois estimations SAE sont validées sur la base du CAS2. Pour le canton des grisons, les estimations SAE pour les deux sexes confondus et pour les hommes sont validées sur la base du DIR2, tandis que les estimations SAE obtenues pour les femmes le sont sur la base du CAS2.

Les résultats obtenus pour les branches d'activité économique se trouvent en annexe 4, p. 40.

5.6.3 Benchmarking et facteurs de redressement

Le benchmarking consiste à ajuster la somme des estimations SAE au sein d'un agrégat de domaines à une valeur de référence considérée comme fiable et publiée pour cet agrégat. Dans notre cas, nous avons pour objectif de préserver les valeurs des estimations directes lorsque ces dernières satisfont aux objectifs de précision. Ainsi pour les cantons, les séries SAE seront ajustées aux estimations directes (DIR) publiées pour les grandes régions.

Exemple : pour les cantons GE, VD, VS, les facteurs de redressement sont calculés par rapport à l'emploi de la grande région Lémanique, dont les estimations DIR satisfont le critère de précision.

Cet ajustement permet de garantir une cohérence entre les estimations diffusées pour les différents niveaux de ventilation. L'analyse de l'ampleur des facteurs d'ajustement peut également être prise en compte pour la validation. Les facteurs d'ajustement utilisés sont conçus de façon à garantir également une cohérence entre les estimations SAE obtenues au niveau NOGA2 et les estimations DIR diffusées au niveau de la section économique.

D'un point de vue opérationnel, ces facteurs sont obtenus sur la base d'un calage :

- des sommes des prédictions (ou des valeurs observées pour les unités de l'échantillon) sur le cadre au sein des grandes régions et des sections économiques
- sur les estimations directes DIR pour ces mêmes agrégats,
- avec des poids de départ égaux à 1.

Ce calage est réalisé sur la base du package sampling de R. Les g-weights ainsi obtenus sont des facteurs d'ajustement individuels pour chaque unité (établissement) qui garantissent la cohérence souhaitée.

5.6.4 Gain de variance

Enfin, comme mentionné dans la section 5.5, la technique du bootstrap est utilisée pour estimer la variance des estimateurs SAE. Cette variance est ensuite comparée à celle des estimateurs directs afin d'évaluer le gain apporté par les méthodes SAE. Idéalement, les CV des estimations SAE cantonales devraient se rapprocher de ceux fixés dans le cadre de ces travaux pour les estimations directes au niveau des grandes régions, cf. section 5.6.2. De même, les CV des estimations SAE au niveau des Noga2 devraient se rapprocher de ceux fixés ici pour les estimations directes au niveau des sections économiques.

6 Résultats régionaux

6.1 Grandes régions

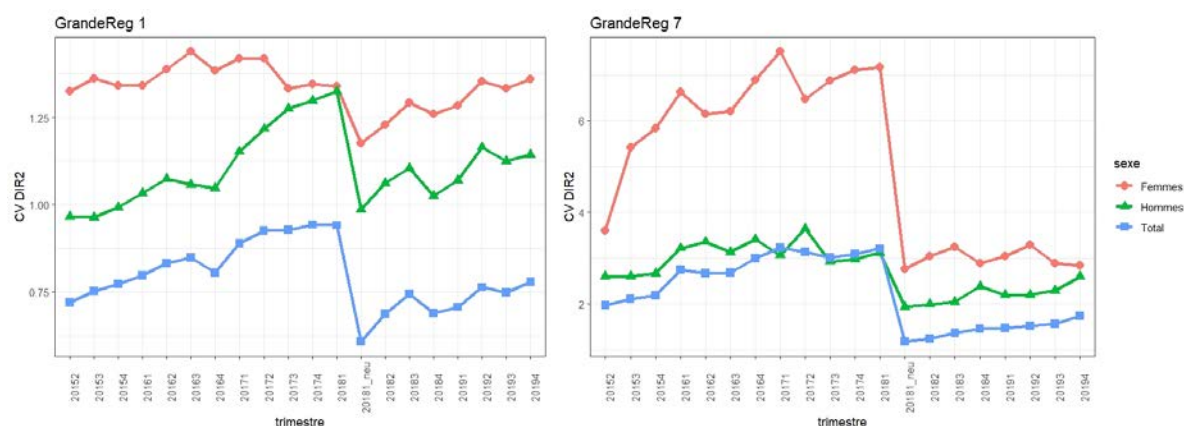
Avant de nous intéresser aux détails des résultats cantonaux, nous nous penchons sur la précision des estimations directes au niveau des grandes régions, qui sont considérées comme valeurs de référence pour le benchmarking. Dans ce cas, les objectifs de précision visés sont des CV inférieurs à 2% pour l'emploi total et de 3% pour l'emploi par sexe. La Table 1 indique que ces objectifs sont globalement respectés pour les estimations directes DIR2 (estimateur direct par le ratio). Une analyse détaillée indique que les critères sont satisfaits en moyenne pour toutes les grandes régions à l'exception de la catégorie des femmes au Tessin. Notons qu'il s'agit d'une grande région formée d'un seul canton, de plus de d'assez faible taille.

Table 1 Résumé des CV des estimations directes (DIR2) pour les sept grandes régions (période 2015T2-2019T4)

Grande région	Femmes (3%)			Hommes (3%)			Total (2%)		
	CV moy	CV min	CV max	CV moy	CV min	CV max	CV moy	CV min	CV max
1. Lemanique	1.34	1.18	1.44	1.1	0.96	1.32	0.79	0.61	0.94
2. E.Mittelland	1.2	1.07	1.33	1.27	1.05	1.41	0.7	0.55	0.8
3. NordwestCH	1.76	1.62	1.93	1.52	1.29	1.69	1.1	0.92	1.26
4. Zürich	1.15	1.05	1.26	1.06	0.93	1.2	0.71	0.58	0.87
5. OstCH	1.55	1.33	1.83	1.32	1.16	1.43	0.88	0.73	1.02
6. ZentralCH	2.41	1.86	3.84	1.71	1.55	2.04	1.34	0.99	2.38
7. Tessin	4.99	2.76	7.5	2.72	1.94	3.64	2.23	1.18	3.22

La situation du Tessin est illustrée à la Figure 2, et comparée à celle de la Région Lémanique. Ce graphe indique que la précision des estimations est fortement dépendante de l'échantillon. Elles sont globalement satisfaisantes pour le total et les hommes sur l'ensemble des 20 trimestres considérés, mais pas pour les femmes dans le cas de l'échantillon de 2015.

Figure 2 Graphique des CV (coefficients de variation) des estimations directes (DIR2) pour deux grandes régions (Lémanique et Tessin) illustrant l'instabilité des estimations de variance, surtout pour le canton du Tessin, canton de faible taille qui constitue à lui seul une grande région



6.2 Cantons et ville de ZH

6.2.1 Validation: critères locaux

L'annexe 2 contient les résultats détaillés pour toutes les séries temporelles par canton. A titre d'illustration, nous présentons et commentons les résultats obtenus pour trois cantons.

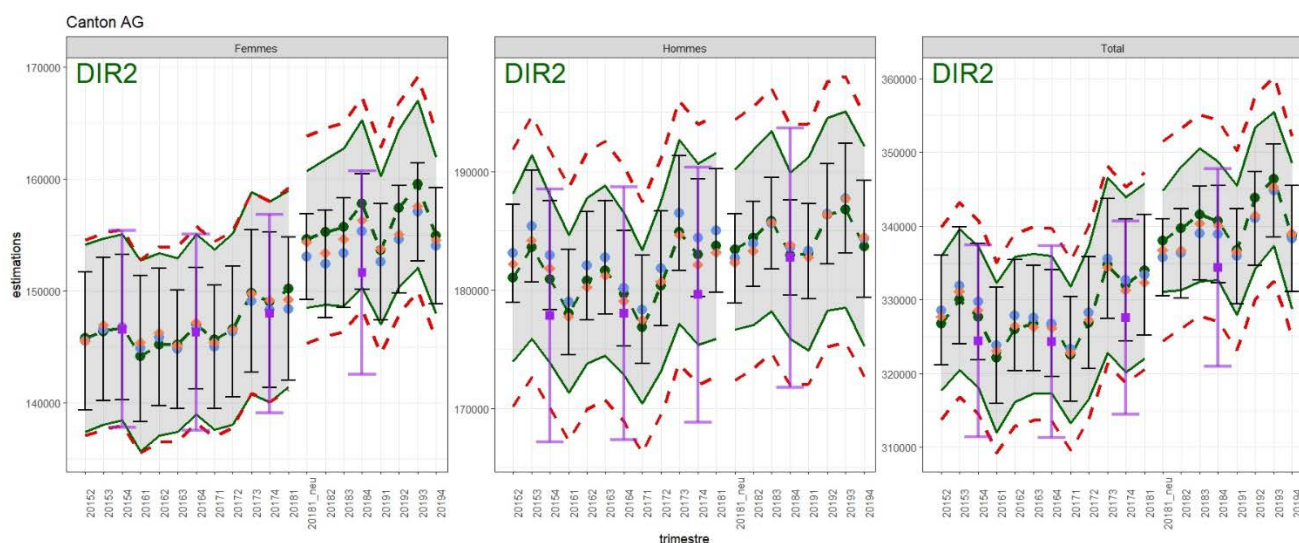
Exemple du canton AG, séries temporelles (critère « and » satisfait, cf. section 5.6.2, pour les hommes, les femmes et les deux sexes confondus → validation par le DIR2)

La Figure 3 ci-dessous représente les résultats sous forme de séries temporelles pour les estimations directes par le ratio (DIR2 en vert) et ses intervalles de confiance à 95% (bande grise), le modèle saenb (en bleu), les seuils d'écarts au DIR2 tolérés pour la validation ($\pm 6\%$ pour les hommes et les femmes, $\pm 4\%$ pour les deux sexes confondus, traitillés rouges).

On constate que les intervalles à 95% du DIR2 satisfont les objectifs de précision (la bande grise est incluse dans la zone délimitée par les traitillés rouge), justifiant une validation selon les estimations DIR2. Par ailleurs, les estimations SAE (en bleu) ne franchissent jamais les seuils d'écarts au DIR2 tolérés (traitillés rouges). Les estimations SAE sont donc considérées comme validées pour le canton d'Argovie.

En marge de la validation, on constate également que pour ce canton, l'impact du renouvellement d'échantillon en 2018 est relativement faible.

Figure 3 Estimations trimestrielles (DIR2 et SAE) pour le canton AG



Légende :

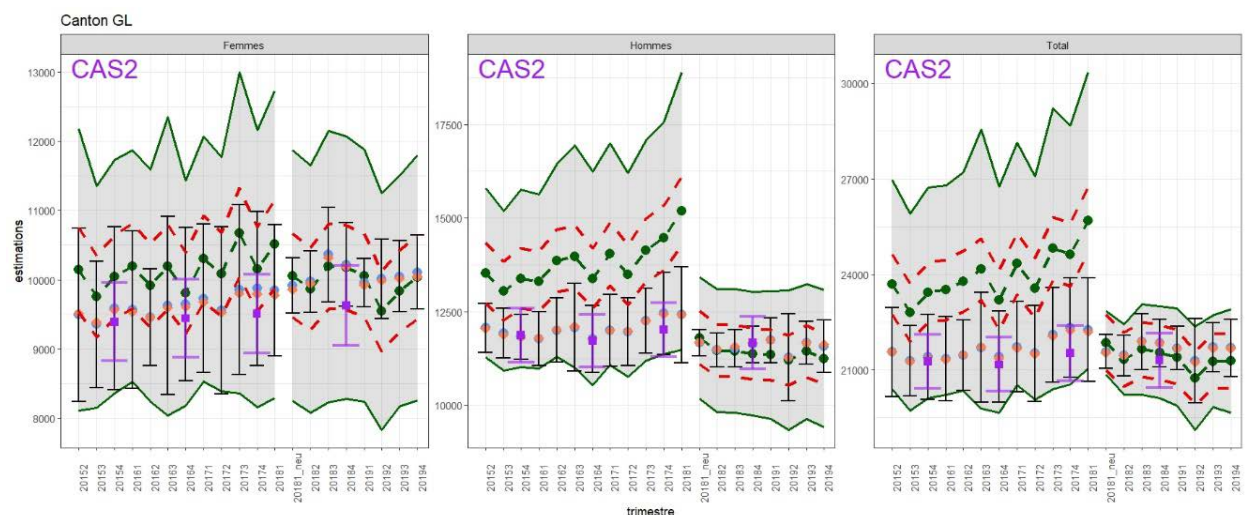
- Points verts : estimations directes DIR2 ;
- Ruban en gris : intervalle de confiance à 95% pour les estimations directes DIR2 ;
- Points (ronds) bleus : estimations SAE selon le modèle saenb ;
- Barres noires : intervalle de confiance à 95% pour les estimations SAE ;
- Lignes en rouge traitillées : écarts tolérés au DIR2, en cas de validation sur la base du DIR2 ($\pm 4\%$ sexes confondus et $\pm 6\%$ par sexe) ;
- Points (carrés) violets : totaux calculés sur la base du CAS2 ;
- Barres violettes : écarts tolérés au CAS2, en cas de validation sur la base du CAS2 ($\pm 4\%$ sexes confondus et $\pm 6\%$ par sexe) ;
- Points (losanges) oranges : estimations SAE ajustées (benchmarking) ;

La méthode de validation est indiquée sur chaque panel (DIR2 ou CAS2) ;
 La discontinuité des séries au 1^{er} trim. 2018 indique l'ampleur du renouvellement de l'échantillon avant le raccordement des séries.

Exemple du canton GL, séries temporelles (critère « and » non satisfait → validation sur la base du CAS2 pour les hommes, les femmes et les deux sexes confondus.)

La Figure 4 représente les résultats obtenus pour le canton GL. Dans ce cas, l'intervalle de confiance à 95% des estimations directes DIR2 ne satisfait pas les objectifs de précision (la bande grise dépasse largement la zone délimitée par les traitillés rouges). La validation s'effectue donc selon les valeurs du CAS2 (en violet, seulement pour les quatrièmes trimestres). On constate que l'estimateur SAE (points bleus) est toujours situé entre les seuils d'écarts tolérés par rapport au CAS2 (intervalle violet), sauf pour les femmes en 2018, l'écart entre l'estimateur SAE et le seuil restant néanmoins très léger dans ce cas. On observe par ailleurs un impact important du renouvellement de l'échantillon sur les estimations directes pour les hommes et les deux sexes confondus, alors qu'il n'a que peu d'ampleur pour les estimations SAE. Ceci illustre le gain de stabilité potentiellement permis par les SAE pour les petits domaines. Compte tenu de l'écart observé entre l'estimateur SAE et le CAS2 pour les femmes au quatrième trimestre 2018, les estimations SAE ne sont en principe pas validées pour ce trimestre. Cependant, l'estimation SAE restant dans ce cas très proche du seuil de validation, les résultats sont très encourageants. Une éventuelle reprise de ce projet pourrait notamment permettre soit d'améliorer le modèle, soit d'expliquer certains écarts observés.

Figure 4 Séries temporelles pour le canton GL

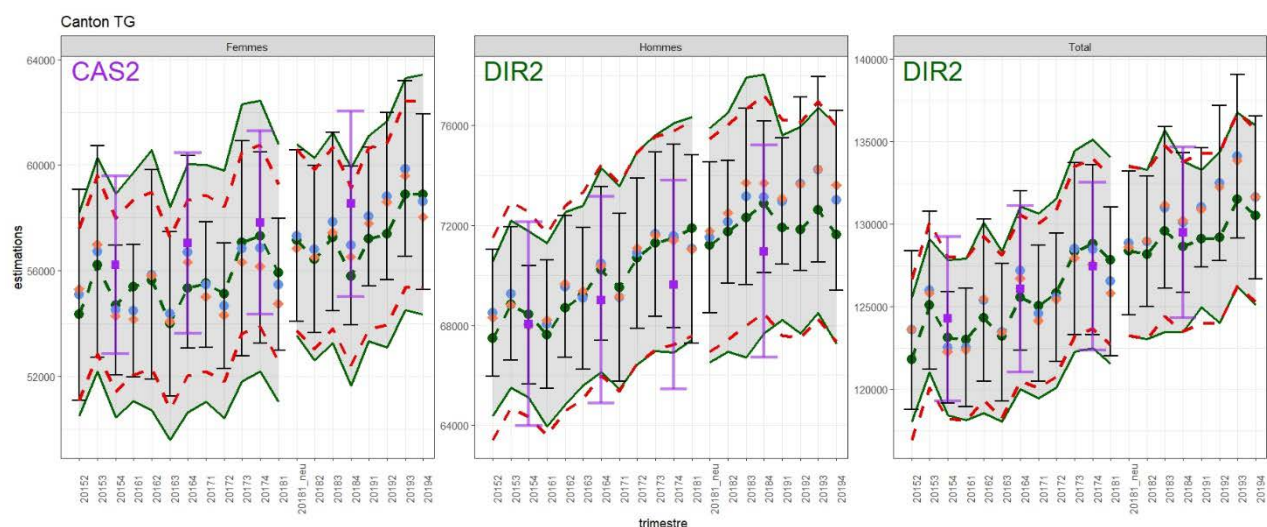


Légende : Voir figure 3.

Exemple du canton TG, séries temporelles (validation sur la base du CAS2 pour les femmes, et sur le DIR2 pour les hommes et les deux sexes confondus.)

Dans le cas du canton TG, le cas des femmes est validé par les données CAS2 alors que pour le total et pour les hommes, la validation se fait avec les estimations directes, voir figure 5. Les estimations SAE sont validées pour ce canton.

Figure 5 Séries temporelles pour le canton TG



Légende : Voir figure 3.

La table 2 résume les résultats de la validation au niveau des cantons. La table donne, pour chaque canton et pour chaque sexe ou sexes confondus, le critère utilisé pour la validation (DIR2 ou CAS2), le nombre et le pourcentage de trimestres valides, ainsi que l'écart (en valeur absolue) relatif moyen³ entre les estimations SAE (sans ou avec benchmarking) et les valeurs de référence pour la validation (DIR2 ou CAS2 selon le cas). Ces quantités sont désignées par EAM (sans benchmarking) et EAM_BM (avec benchmarking) dans la Table 2. La Table 3 donne plus de détails pour les croisements canton*sexe non validés selon ces critères. Il contient en effet les valeurs EAM et EAM_BM calculées spécifiquement sur les trimestres non-validés. A partir de ces deux tables on peut constater que les estimations SAE pour le total (les deux sexes confondus) sont validées pour tous les cantons à l'exception d'Uri, pour lequel l'estimation SAE s'écarte très légèrement du CAS2 pour l'un des trimestres. De plus, la grande majorité des croisements canton*sexe sont validés. Les croisements non validés concernent généralement l'emploi des femmes dans d'assez petits ou petits cantons des grandes régions 5 et 6 (Suisse orientale et Suisse centrale), et les écarts entre l'estimation SAE et le seuil de validation sont souvent assez faibles. Le cas le plus problématique est celui du Tessin. Les résultats sont donc globalement très encourageants, et des améliorations sont probablement encore possibles. Il a par exemple été constaté lors de ce projet que des valeurs extrêmes étaient source de problèmes, également pour les estimations directes (notamment pour le Tessin avant le renouvellement de l'échantillon en 2018). La procédure appliquée dans le cadre de ce projet n'est pas une procédure robuste, le package R utilisé ne le permettant pas pour le moment (lme4). En cas de reprise du projet, une solution serait de détecter des valeurs extrêmes par exemple à l'aide de la distance de Cook dans une régression sans effet aléatoire, puis d'utiliser le package « lme4 ».

Des investigations complémentaires pourraient par ailleurs permettre de mieux comprendre certains écarts observés entre les SAE et les seuils de validation.

³ Pour un canton donné et un type de résultat (hommes, femmes, ou les deux sexes confondus), on calcule pour chaque trimestre (ou chaque quatrième trimestre en cas de validation par le CAS2) la différence relative en valeur absolue entre l'estimation SAE et la valeur de référence pour la validation. On calcule ensuite la moyenne à travers les trimestres de ces écarts.

Table 2 Résumé des résultats de la validation au niveau des cantons. La table donne le nombre et le pourcentage de trimestres validés, ainsi que l'écart relatif (en valeur absolue) moyen entres les estimations SAE (sans ou avec benchmarking) et les valeurs de référence pour la validation (DIR2 ou CAS2 selon le cas).

Cant	Sexe	Valid	#val tri	%val trim	#val tri bm	%val tri bm	EAM (%)	EAM bm (%)
AG	Femmes	DIR2	20	100%	20	100%	0.75%	0.46%
AG	Hommes	DIR2	20	100%	20	100%	0.57%	0.33%
AG	Total	DIR2	20	100%	20	100%	0.44%	0.28%
AI	Femmes	CAS2	1	25%	1	25%	6.99%	7.35%
AI	Hommes	CAS2	4	100%	4	100%	0.40%	0.28%
AI	Total	CAS2	4	100%	4	100%	3.49%	3.59%
AR	Femmes	CAS2	3	75%	4	100%	4.63%	4.21%
AR	Hommes	CAS2	4	100%	4	100%	2.30%	2.36%
AR	Total	CAS2	4	100%	4	100%	3.43%	3.26%
BE	Femmes	DIR2	20	100%	20	100%	1.17%	1.01%
BE	Hommes	DIR2	20	100%	20	100%	0.31%	0.28%
BE	Total	DIR2	20	100%	20	100%	0.67%	0.37%
BL	Femmes	DIR2	20	100%	20	100%	2.29%	2.57%
BL	Hommes	CAS2	4	100%	4	100%	1.72%	1.50%
BL	Total	CAS2	4	100%	4	100%	2.25%	2.11%
BS	Femmes	DIR2	20	100%	20	100%	0.79%	1.02%
BS	Hommes	DIR2	20	100%	20	100%	0.94%	2.74%
BS	Total	DIR2	20	100%	20	100%	0.78%	1.46%
FR	Femmes	DIR2	20	100%	20	100%	1.22%	1.22%
FR	Hommes	CAS2	4	100%	4	100%	3.14%	3.24%
FR	Total	CAS2	4	100%	4	100%	2.83%	2.91%
GE	Femmes	DIR2	20	100%	20	100%	1.17%	0.77%
GE	Hommes	DIR2	20	100%	20	100%	1.30%	0.80%
GE	Total	DIR2	20	100%	20	100%	1.14%	0.58%
GL	Femmes	CAS2	3	75%	4	100%	3.54%	3.11%
GL	Hommes	CAS2	4	100%	4	100%	1.20%	1.05%
GL	Total	CAS2	4	100%	4	100%	2.03%	1.86%
GR	Femmes	CAS2	4	100%	4	100%	1.32%	1.20%
GR	Hommes	DIR2	18	90%	20	100%	3.36%	3.09%
GR	Total	DIR2	20	100%	20	100%	1.69%	1.45%
JU	Femmes	CAS2	4	100%	4	100%	1.47%	1.46%
JU	Hommes	CAS2	4	100%	4	100%	2.19%	2.29%
JU	Total	CAS2	4	100%	4	100%	1.86%	1.91%
LU	Femmes	CAS2	3	75%	3	75%	4.28%	4.54%
LU	Hommes	DIR2	20	100%	20	100%	0.82%	0.83%
LU	Total	CAS2	4	100%	4	100%	2.65%	2.29%
NE	Femmes	DIR2	20	100%	20	100%	1.42%	1.31%
NE	Hommes	DIR2	20	100%	20	100%	0.48%	0.81%
NE	Total	DIR2	20	100%	20	100%	0.52%	0.43%

Table 2 (suite).

Cant	Sexe	Valid	#val tri	%val_trim	#val_tri_bm	%val_tri_bm	EAM (%)	EAM_bm (%)
NW	Femmes	CAS2	3	75%	3	75%	4.40%	4.27%
NW	Hommes	CAS2	3	75%	3	75%	3.03%	3.70%
NW	Total	CAS2	4	100%	4	100%	1.19%	1.80%
OW	Femmes	CAS2	1	25%	1	25%	7.63%	7.82%
OW	Hommes	CAS2	4	100%	4	100%	1.35%	1.70%
OW	Total	CAS2	4	100%	4	100%	2.69%	2.58%
SG	Femmes	DIR2	20	100%	20	100%	0.93%	1.27%
SG	Hommes	DIR2	20	100%	20	100%	0.72%	1.00%
SG	Total	DIR2	20	100%	20	100%	0.77%	0.49%
SH	Femmes	CAS2	4	100%	4	100%	3.17%	2.67%
SH	Hommes	CAS2	4	100%	4	100%	1.51%	1.75%
SH	Total	CAS2	4	100%	4	100%	1.12%	1.15%
SO	Femmes	DIR2	20	100%	20	100%	2.83%	2.95%
SO	Hommes	DIR2	20	100%	20	100%	1.33%	1.05%
SO	Total	DIR2	20	100%	20	100%	1.32%	1.29%
SZ	Femmes	CAS2	4	100%	4	100%	3.44%	3.87%
SZ	Hommes	CAS2	4	100%	4	100%	3.04%	3.35%
SZ	Total	CAS2	4	100%	4	100%	0.79%	0.85%
TG	Femmes	CAS2	4	100%	4	100%	1.99%	2.75%
TG	Hommes	DIR2	20	100%	20	100%	0.93%	1.09%
TG	Total	DIR2	20	100%	20	100%	0.87%	0.95%
TI	Femmes	CAS2	1	25%	1	25%	7.32%	9.63%
TI	Hommes	DIR2	20	100%	20	100%	0.70%	1.29%
TI	Total	CAS2	4	100%	3	75%	3.65%	4.21%
UR	Femmes	CAS2	1	25%	0	0%	8.12%	8.69%
UR	Hommes	CAS2	4	100%	4	100%	1.90%	1.56%
UR	Total	CAS2	3	75%	3	75%	4.67%	4.68%
VD	Femmes	DIR2	20	100%	20	100%	0.37%	0.73%
VD	Hommes	DIR2	20	100%	20	100%	0.36%	0.36%
VD	Total	DIR2	20	100%	20	100%	0.26%	0.47%
VS	Femmes	DIR2	20	100%	20	100%	1.13%	1.17%
VS	Hommes	DIR2	20	100%	20	100%	1.55%	1.65%
VS	Total	DIR2	20	100%	20	100%	1.05%	1.11%
ZG	Femmes	CAS2	2	50%	2	50%	5.29%	5.89%
ZG	Hommes	DIR2	20	100%	20	100%	1.46%	1.21%
ZG	Total	CAS2	4	100%	4	100%	0.88%	0.93%
ZH	Femmes	DIR2	20	100%	20	100%	0.67%	1.34%
ZH	Hommes	DIR2	20	100%	20	100%	0.16%	1.12%
ZH	Total	DIR2	20	100%	20	100%	0.37%	0.01%
zs	Femmes	DIR2	20	100%	9	45%	1.46%	2.50%
zs	Hommes	DIR2	20	100%	20	100%	0.38%	1.25%
zs	Total	DIR2	20	100%	20	100%	0.87%	0.49%

Table 3 Résumé des résultats de la validation pour les cantons non validés. En plus des informations contenues dans la Table 2, cette table donne les écarts relatifs moyens EAM et EAM_BM également spécifiquement pour les trimestres non-validés. Ces valeurs sont notées EAM_n_val et EAM_n_val_bm. Par exemple, pour l'emploi des femmes dans le canton de Zug (validation sur la base du CAS2), 2 trimestres sur 4 sont validés. Pour les deux trimestres non validés, l'écart (en valeur absolue) relatif moyen est de 6.82%.

Cant	Sexe	Valid	#val tri	%val trim	#val tri bm	%val tri bm	EAM (%)	EAM bm (%)	EAM.n.val (%)	EAM.n.val bm (%)
AI	Femmes	CAS2	1	25%	1	25%	6.99%	7.35%	8.03%	8.39%
AR	Femmes	CAS2	3	75%	4	100%	4.63%	4.21%	6.25%	-
GL	Femmes	CAS2	3	75%	4	100%	3.54%	3.11%	6.11%	-
GR	Hommes	DIR2	18	90%	20	100%	3.36%	3.09%	6.04%	-
LU	Femmes	CAS2	3	75%	3	75%	4.28%	4.54%	7.75%	7.60%
NW	Femmes	CAS2	3	75%	3	75%	4.40%	4.27%	6.22%	6.10%
NW	Hommes	CAS2	3	75%	3	75%	3.03%	3.70%	6.65%	6.61%
OW	Femmes	CAS2	1	25%	1	25%	7.63%	7.82%	8.87%	8.75%
TI	Femmes	CAS2	1	25%	1	25%	7.32%	9.63%	9.49%	12.21%
UR	Femmes	CAS2	1	25%	0	0%	8.12%	8.69%	8.91%	8.69%
UR	Total	CAS2	3	75%	3	75%	4.67%	4.68%	6.58%	6.58%
ZG	Femmes	CAS2	2	50%	2	50%	5.29%	5.89%	6.14%	6.82%

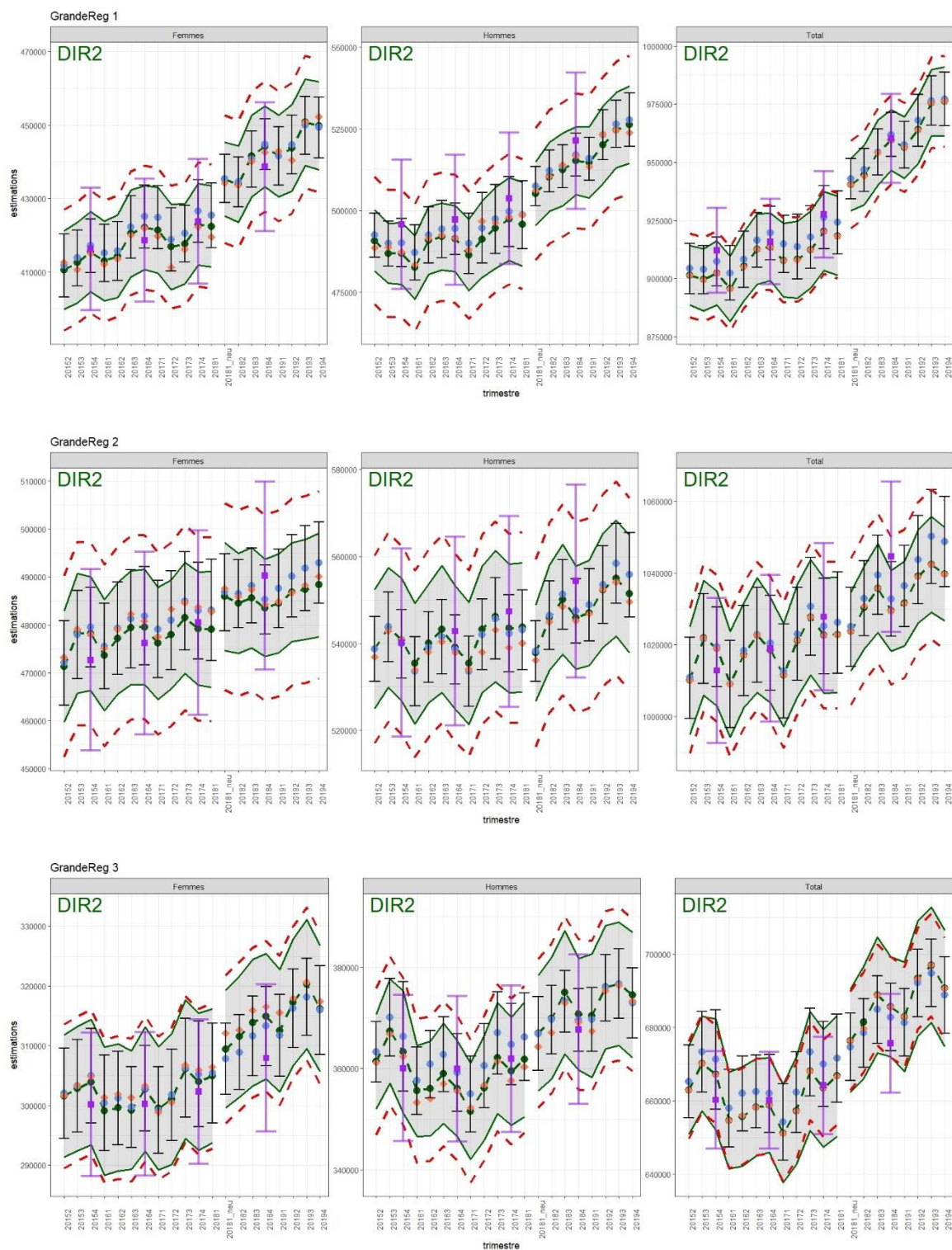
6.2.2 Validation: critères globaux

Les graphiques ci-après représentent la somme des estimations SAE au sein des différentes grandes régions (points bleus), l'estimateur direct DIR2 et son intervalle de confiance (en vert), ainsi que les valeurs du CAS2 (violet). La légende⁴ est la même que celle de la page 10. On peut faire les constatations suivantes.

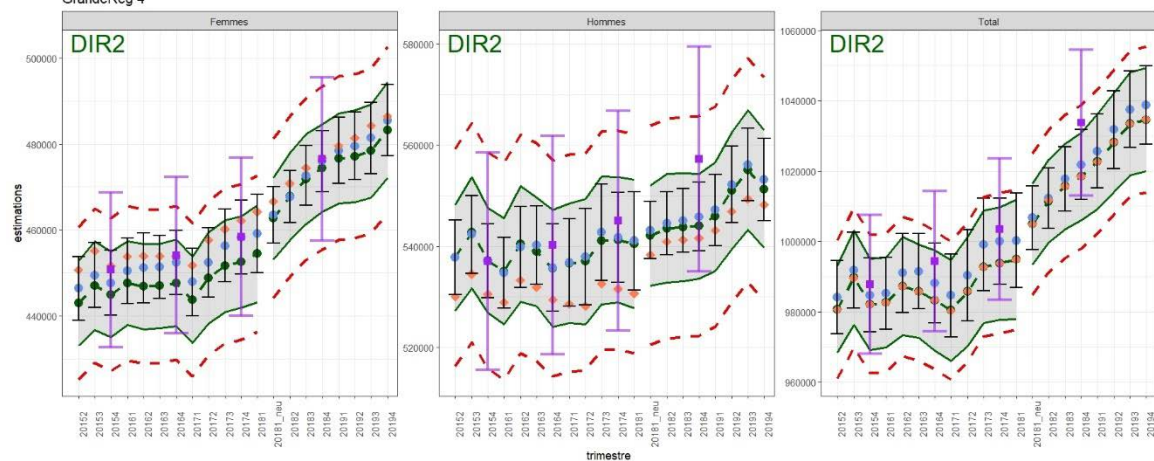
- La somme des estimations est souvent très proche de l'estimateur DIR2, et se situe toujours dans son intervalle de confiance. L'effet du renouvellement de l'échantillon en 2018 est globalement le même pour les estimations SAE et pour le DIR2, les estimations SAE réduisant toutefois cet effet pour les femmes dans la grande région 2 (espace Mittelland). A noter que l'effet stabilisateur des SAE est bien plus marqué au niveau des cantons, comme nous l'avons par exemple constaté pour le canton de Glaris (cf. section 6.2.1).
- Les valeurs du CAS2 sortent de l'intervalle de confiance du DIR2 pour les hommes et le total dans la grande région 4 (Zürich) en 2018, illustrant le fait que ces valeurs ne peuvent être utilisées que comme des approximations de la valeur cible.
- Comme déjà mentionné précédemment, le cas du Tessin est problématique. L'effet du renouvellement de l'échantillon est très fort pour les femmes et le total, et pourrait être dû à la présence de valeurs extrêmes dans l'échantillon tiré en 2015.

Néanmoins, à l'exception du Tessin, ces résultats confirment l'évaluation très encourageante du paragraphe précédent.

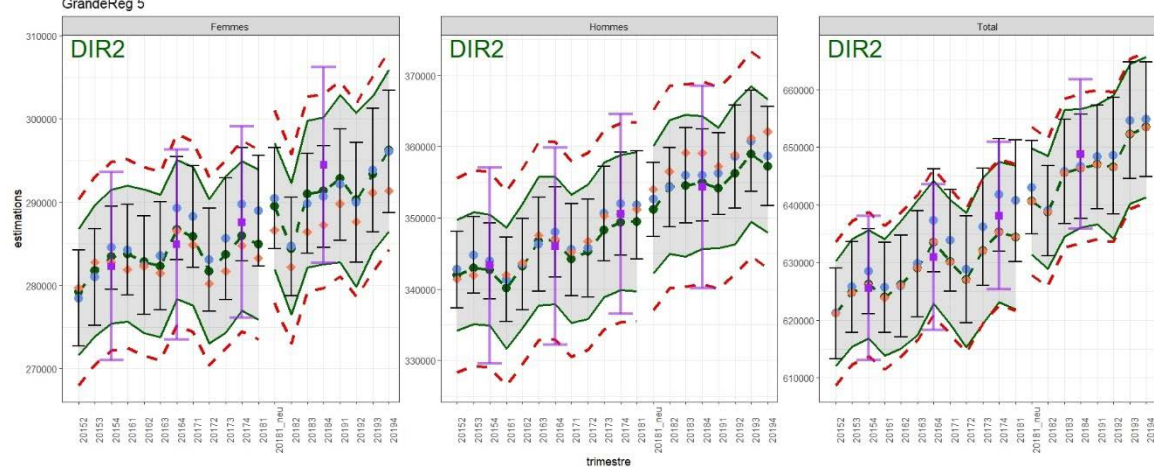
⁴ Les points orange représentent les estimations SAE avec benchmarking. Comme le benchmarking est effectué sur la base du DIR (qui est l'estimateur publié au niveau des grandes régions), la valeur des points orange coïncide par principe à celle de l'estimateur DIR. Les écarts entre les points verts (DIR2) et les points orange correspondent donc aux différences entre le DIR et le DIR2.



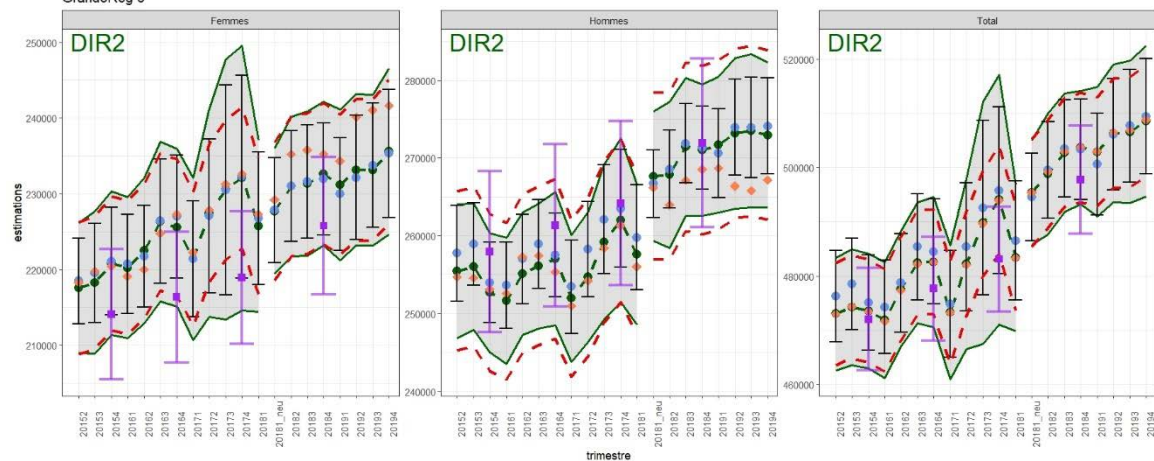
GrandeReg 4

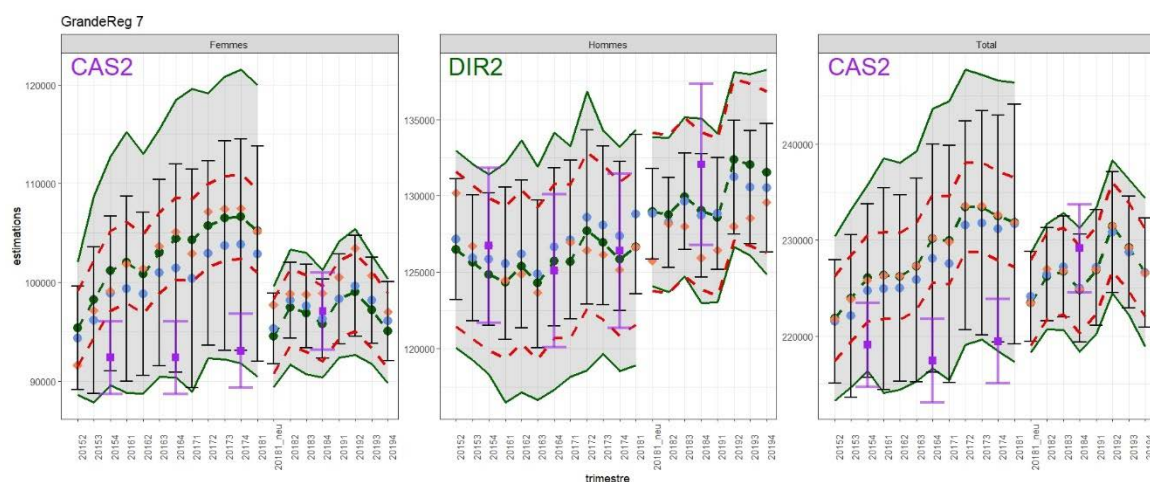


GrandeReg 5



GrandeReg 6





Benchmarking

Les estimations SAE sont ajustées aux estimations directes (DIR) des grandes régions et des branches économiques (sections) simultanément, de manière à assurer une cohérence globale (cf. section 5.6.3). La Table 4 contient quelques statistiques descriptives sur la différence entre l'estimation SAE et l'estimation DIR, au fil des différents trimestres. Plus précisément, on calcule cette différence pour chaque trimestre, puis on en prend la moyenne, le minimum et le maximum au fil des trimestres.

Table 4 Statistiques sur les valeurs absolues des différences relatives entre l'estimateur SAE (sans et avec BM) et le DIR, en pourcentage.

Canton	Femmes			Hommes			Total		
	moy	min	max	moy	min	max	moy	min	max
AG	0.40	0.00	0.87	0.52	0.02	1.23	0.27	0.11	0.43
AI	0.33	0.01	0.97	0.09	0.00	0.25	0.12	0.01	0.34
AR	0.38	0.08	0.82	0.18	0.00	0.48	0.13	0.01	0.43
BE	0.39	0.02	1.04	0.43	0.01	1.13	0.30	0.01	0.95
BL	0.29	0.02	0.50	0.39	0.00	0.96	0.20	0.08	0.37
BS	1.18	0.01	2.96	2.00	0.25	4.50	1.14	0.27	2.39
FR	0.15	0.01	0.35	0.34	0.01	1.01	0.17	0.01	0.62
GE	0.76	0.22	1.91	0.62	0.05	1.33	0.57	0.00	1.05
GL	0.38	0.02	0.75	0.17	0.01	0.42	0.13	0.00	0.42
GR	0.40	0.08	0.78	0.34	0.01	0.98	0.24	0.00	0.77
JU	0.08	0.01	0.18	0.24	0.00	0.73	0.14	0.01	0.47
LU	1.43	0.01	4.50	1.34	0.11	2.95	0.67	0.24	1.27
NE	0.15	0.00	0.33	0.40	0.00	1.49	0.25	0.01	0.96
NW	0.34	0.01	1.08	2.91	0.04	8.42	1.56	0.01	4.36
OW	1.76	0.04	8.07	0.72	0.03	2.04	0.66	0.01	2.77
SG	1.60	0.17	3.11	0.51	0.00	1.37	0.50	0.01	1.39
SH	0.74	0.07	1.63	0.27	0.02	0.67	0.25	0.00	0.86
SO	0.29	0.02	0.62	0.34	0.06	0.93	0.26	0.06	0.78
SZ	0.56	0.00	1.28	0.49	0.07	1.08	0.11	0.01	0.26
TG	0.66	0.06	1.32	0.29	0.01	0.84	0.20	0.00	0.69
TI	2.42	0.25	4.10	1.35	0.06	2.45	0.45	0.01	1.01
UR	0.63	0.05	1.54	0.60	0.13	1.43	0.09	0.01	0.22
VD	0.82	0.12	2.19	0.30	0.00	0.84	0.41	0.02	0.78
VS	0.19	0.00	0.62	0.25	0.01	0.69	0.12	0.01	0.28
ZG	0.88	0.09	2.32	1.45	0.18	3.55	0.43	0.02	1.00
ZH	0.67	0.18	1.27	1.25	0.66	1.91	0.37	0.05	0.64
zs	1.02	0.23	1.78	1.60	0.89	2.36	0.38	0.13	0.62

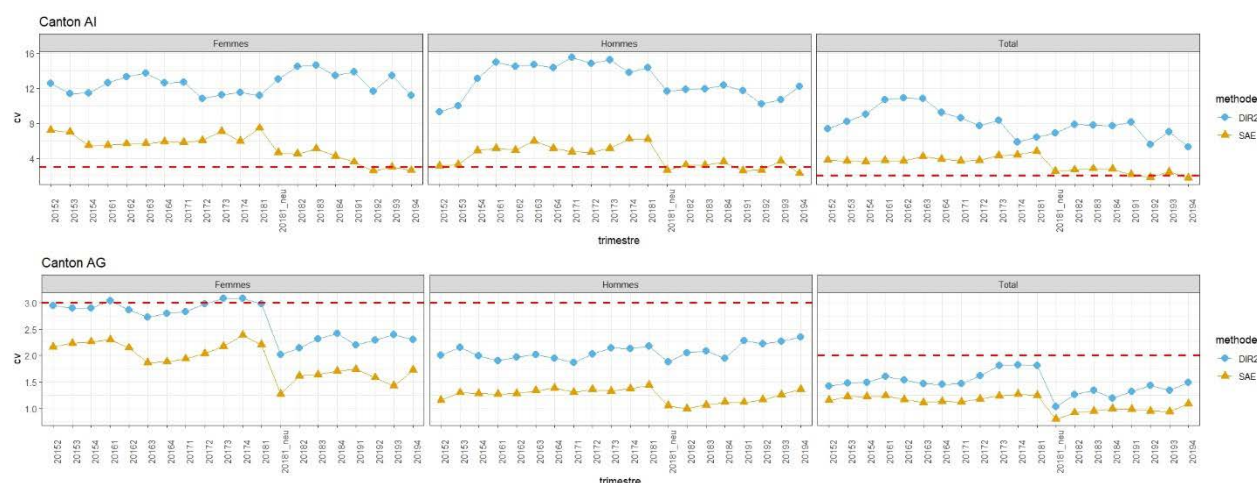
On constate que les facteurs moyens sont très faibles la plupart du temps. Pour les femmes, ils varient de 0.08% pour le Jura à 2.42% pour le Tessin. Ils se situent entre 0.09% (AI) et 2.91% (NW) pour les hommes, et entre 0.09% (UR) et 1,56% (NW) pour le total. Les maximas se situent aux alentours de 8% (par exemple 8.42% pour les hommes pour NW) et restent donc modérés. Ces résultats vont donc également dans le sens de l'évaluation très encourageante du paragraphe 6.2.1.

6.2.3 Variance des estimations SAE

Les coefficients de variation de l'estimateur SAE (selon le modèle saenb) et de l'estimateur DIR2 sont estimés par la méthode du bootstrap, puis comparés entre eux. Les résultats cantonaux figurent en annexe 3. Par souci de concision, nous nous concentrons dans ce rapport

A titre d'illustration, nous présentons ci-après les résultats obtenus pour deux cantons (AI et AG). On constate comme attendu que les CV des estimations SAE sont bien inférieurs aux CV des estimations directes (DIR2), en particulier pour le canton d'AI. Pour ce canton, les estimations SAE permettent de se rapprocher des objectifs d'un CV de 2% pour l'emploi total et de 3% pour l'emploi par sexe. De façon générale, les SAE permettent un très important gain de variance, hormis pour les très grands cantons tels que Zürich pour lequel le gain est plus faible mais l'objectif de précision est déjà atteint avec l'estimation directe. Le gain de précision est plus marqué pour les petits cantons. A noter que l'application de méthodes SAE n'a pas réglé la totalité des points surprenants que l'on peut observer pour l'estimateur direct. On constate par exemple un saut des CV du DIR2 et de l'estimation SAE pour les hommes et le total au quatrième trimestre 2015 pour le canton BL. Un phénomène similaire se produit pour le canton LU, le CV de l'estimation SAE étant néanmoins nettement inférieur à celui du DIR2 dans ce cas. D'autre part, l'instabilité des CV des estimations SAE pour le canton NW reste à clarifier. De façon résumée, l'utilisation de méthodes SAE a cependant en règle générale un effet extrêmement bénéfique sur la variance et la stabilité des estimations.

Figure 5 Comparaison entre les CV des estimations directes et ceux des estimations SAE, sans benchmarking, obtenus à l'aide de la méthode du bootstrap



6.3 Discussion

Cas du Tessin : Le cas du Tessin, canton de relativement faible taille et qui constitue à lui seul une grande région reste problématique. On remarque que la situation est également dépendante de l'échantillon, avec des CV insatisfaisants pour l'échantillon 2015 et une situation nettement meilleure avec l'échantillon de 2018. Une solution serait d'adapter le plan de sondage afin d'assurer des estimations directes satisfaisantes pour toutes les grandes régions et produire des SAE

cantonaux avec le modèle saenb.

Cas des cantons et villes avec densification d'échantillon : on constate que les cas des cantons et villes avec densification d'échantillon produisent des estimations directes dont la précision est conforme aux objectifs souhaités. Même si les résultats sont très encourageants, il n'a pas été possible dans le cadre de ce travail d'évaluer objectivement s'il sera possible à l'avenir de renoncer à ces densifications d'échantillon. En effet, il faudrait pour cela disposer de deux sets de données, l'un avec et l'autre sans densification d'échantillons. Cet essai pourrait être réalisé lors du prochain renouvellement de l'échantillon, soit avec les échantillons de 2024.

A noter que le présent travail ne concerne que l'emploi total. Les densifications d'échantillons offrent également des informations sur la répartition par taux d'occupation, les EPT, les places vacantes, ainsi que des variables qualitatives.

7 Résultats pour les branches économiques

7.1 Secteurs et sections économiques

Le 2ème volet de cette analyse a pour but d'évaluer la possibilité d'appliquer les méthodes SAE aux ventilations par branches économiques à des niveaux de détail inférieurs à ceux actuellement disponibles, qui descendent jusqu'à un niveau de regroupements de divisions économiques nommés NOGA OFS50. Ces regroupements permettent d'éviter une charge disproportionnée pour certaines divisions économiques de très faible taille et une augmentation conséquente de la taille de l'échantillon qui permettrait des estimations selon le critère de qualité ci-dessus pour toutes les divisions.

L'objectif est donc ici de déterminer si la méthode SAE permet de fournir des résultats au niveau NOGA2. En vue du benchmarking, nous examinons en premier lieu la précision des estimations directes à un niveau d'agrégation supérieur. A noter que le modèle saenb est un modèle unique qui fournit des estimations SAE en même temps pour les cantons et les divisions NOGA2.

- a) **Secteurs économiques** : au niveau des secteurs économiques, avec des CV de l'ordre de grandeur de 0.5%, proches de ceux de l'emploi total toutes divisions confondues, une précision satisfaisante des estimations est toujours assurée à ce niveau d'agrégation.
- b) **Sections économiques** : la situation est en revanche différente pour les sections économiques. Notons que certains NOGA_OFS 50 représentent directement des divisions NOGA2.

La Table 5 présente les CV moyens avec les extrema pour les estimations directes pour les sections économiques. On constate que dans 7 cas sur 19, les CV moyens sont supérieurs à l'objectif de 2% pour l'emploi total. Dans la ventilation selon le sexe, les CV excèdent nettement plus souvent le seuil de 3% chez les femmes que chez les hommes. Les CV moyens les plus élevés dépassent parfois les 10%, par exemple pour les femmes dans la section E (production et distribution d'eau ; assainissement, gestion des déchets et dépollution). Ces résultats montrent que les estimations directes actuellement diffusées ne satisfont pas totalement les objectifs de précision pour le total, et peuvent prendre des valeurs assez élevées pour les résultats par sexe. L'application d'un benchmarking au niveau des sections économiques (cf. section 5.6.3), qui est pourtant nécessaire pour garantir une cohérence entre les estimations SAE et les estimations directes déjà diffusées, est donc problématique. La poursuite de l'objectif d'obtenir des estimations SAE pour les branches économiques nécessiterait donc peut-être certaines réflexions sur une stratégie de publication alternative.

Table 5 Résumé des CV des estimations directes (DIR2) pour les sections économiques (période 2015-2019). Les cellules marquées en rouge correspondent à celles dont les valeurs excèdent la limite de 2% pour l'emploi total et de 3% par sexe.

Section	NOGA	Femmes (3%)			Hommes (3%)			Total (2%)		
		CV moy	CV min	CV max	CV moy	CV min	CV max	CV moy	CV min	CV max
B	5-9	9.77	7.84	11.81	2.54	1.59	3.48	2.56	1.62	3.54
C	10-33	1.5	1.37	1.63	0.74	0.64	0.87	0.54	0.41	0.72
D	35	3.74	2.85	4.61	1.44	0.99	1.92	1.22	1	1.68
E	36-39	11.37	9.47	17.28	2.43	1.78	3.41	2.89	1.86	5.76
F	41-43	4.23	3.75	5.08	1.32	1.08	1.59	1.21	0.9	1.54
G	45-47	2.12	1.8	2.43	2.08	1.75	2.24	1.08	0.75	1.35
H	49-53	2.58	1.99	3.05	1.62	1.34	1.98	1.49	1.15	1.8
I	55-56	2.72	2.37	3.14	3.73	2.94	5.36	2.14	1.7	2.94
J	58-63	3.22	2.89	3.58	2.17	1.52	3.39	1.88	1.41	2.88

K	64-66	1.42	1.3	1.58	1.26	1	1.55	0.84	0.63	1.11
L	68	5.51	3.96	7.03	4.84	3.92	5.63	3.98	2.6	5.25
M	69-75	1.95	1.69	2.16	1.57	1.27	1.9	1.12	0.79	1.4
N	77-82	3.72	3.23	4.09	3.13	2.43	3.66	2.19	1.45	2.68
O	84	2.5	1.32	4.55	1.23	0.92	1.87	1.29	0.58	2.87
P	85	1.89	1.71	2.3	2.23	2	2.57	1.41	1.19	1.55
Q	86-88	1.28	0.86	1.65	2.42	2.24	2.59	0.98	0.68	1.21
R	90-93	4.3	3.44	5.11	4.47	3.79	5.48	3.07	2.32	3.98
S	94-95	3.31	3.09	3.6	5.95	5.22	6.78	2.39	2.1	2.56

A titre indicatif, un résumé des CV des estimations directes (DIR2) pour les regroupements de divisions économiques NOGA OFS50 est également donné en annexe 5.

7.2 NOGA2

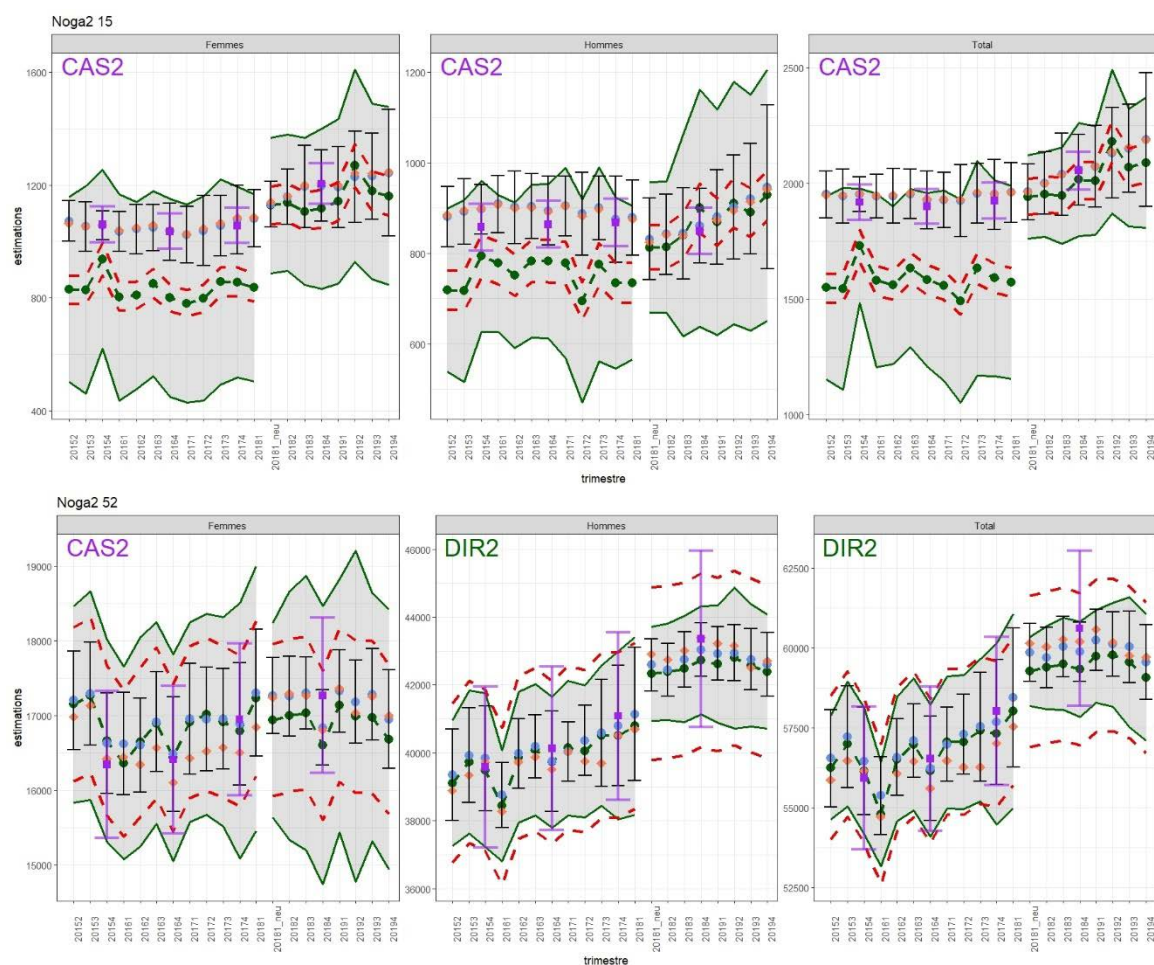
7.2.1 Validation : critères locaux

L'annexe 6 contient les résultats détaillés pour toutes les séries temporelles par NOGA2. Les tables de l'annexe 7 contiennent un résumé des résultats de la validation (équivalent pour les NOGA2 des Table 2 et 3 relatives aux cantons, par exemple pourcentages de trimestres validés). De façon résumée, les résultats sont plus contrastés que pour les cantons, mais demeurent encourageants en vue d'un approfondissement de ce projet. On observe également que lorsque la validation se base sur une comparaison avec l'estimateur direct DIR2, les estimations SAE sont toujours validées⁵. A titre d'illustration, nous présentons et commentons ci-après les résultats obtenus pour quelques NOGA2.

Exemples de cas pour lesquels les méthodes SAE fournissent de bons résultats

Nous examinons les cas des NOGA2 15 (industrie du cuir et de la chaussure) et 52 (entreposage et services auxiliaires des transports). Pour ces deux cas, les estimations SAE sont validées pour les trois types de résultats (hommes, femmes, et deux sexes confondus). On constate par ailleurs que, pour la NOGA2 15, l'estimation SAE permet un très important gain de précision et de stabilité, qui ne suffit toutefois pas à atteindre complètement les objectifs de précision visés (cf. section 5.6.4). Dans cette situation, il peut être légitime de se demander s'il est réellement pertinent de se fixer les mêmes objectifs de précision pour des NOGA2 comptant très peu d'emplois que pour les cantons. Dans le cas de la NOGA2 52, le gain de précision permet d'atteindre l'objectif pour les trois types de résultats, y compris pour les femmes, ce qui n'était pas le cas avec l'estimateur direct.

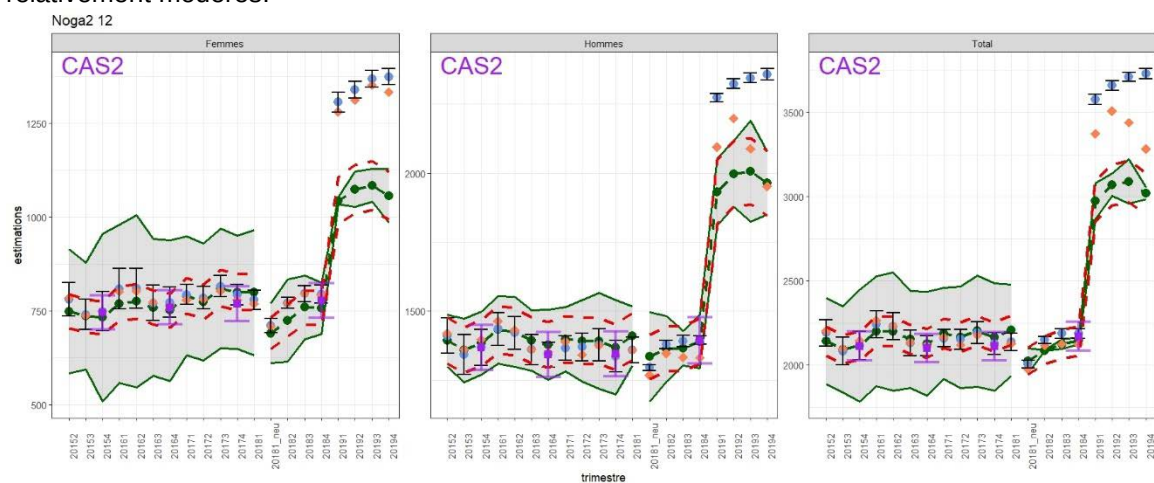
⁵ Avec une exception pour le total de la NOGA2 64 (activités des services financiers hors assurance et caisses de retraites), pour lequel l'un des 20 trimestres n'est pas validé, l'écart entre l'estimation SAE et le seuil restant dans ce cas très faible.

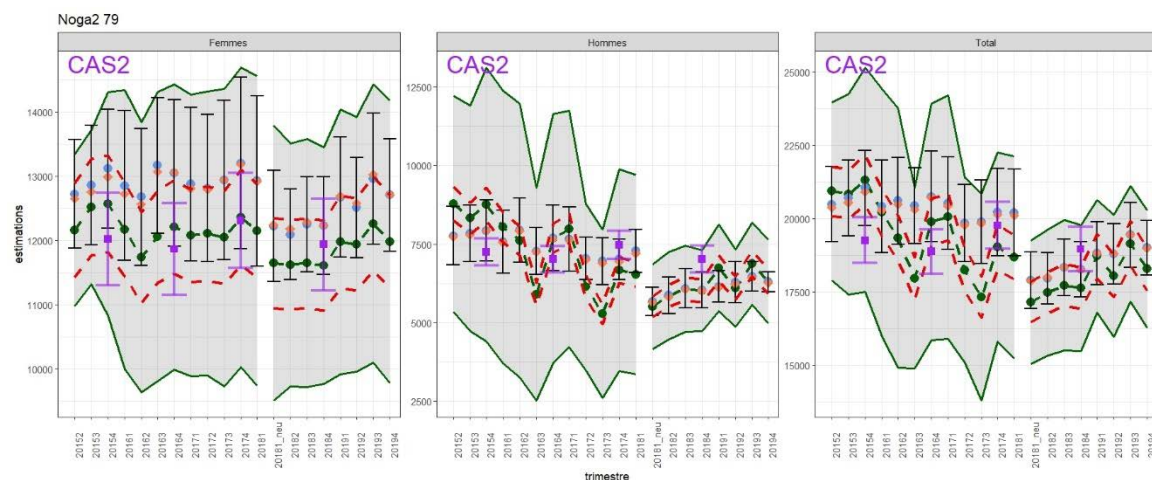


Exemples de cas problématiques

Dans la NOGA2 12 (fabrication à base de tabac), on observe à partir du premier trimestre 2019 un saut très surprenant pour l'estimateur direct, qui est amplifié par les estimations SAE. Ce cas devrait faire l'objet d'investigations approfondies.

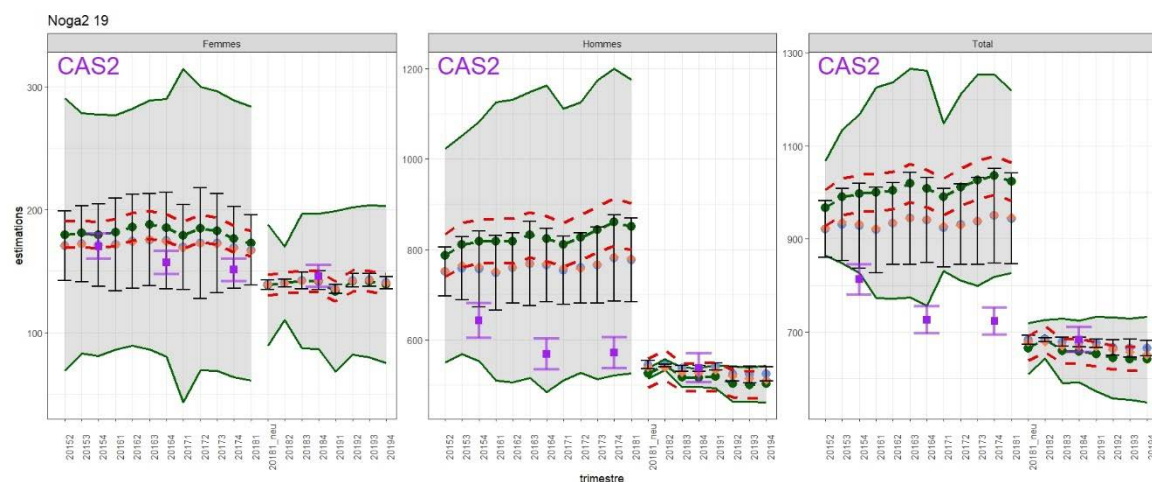
Dans le cas de la NOGA2 79 (activités des agences de voyage, voyagistes, services de réservation et activités connexes), les estimations SAE dépassent les seuils d'écart tolérés au CAS2 pour trois trimestres sur quatre pour les hommes, ainsi que pour les femmes, et pour deux trimestres pour les deux sexes confondus. On observe toutefois que les écarts aux seuils restent relativement modérés.



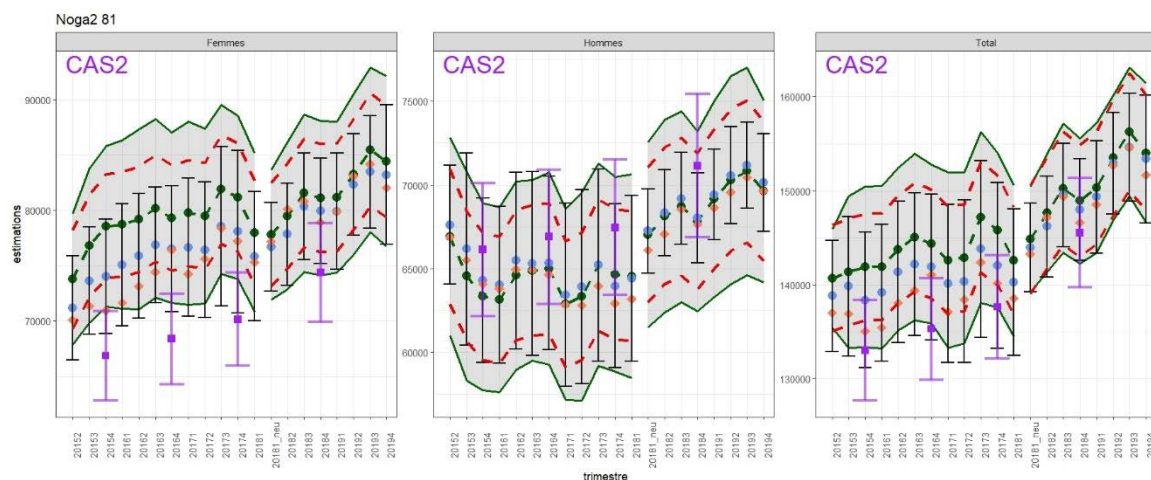


Discussion sur quelques cas pour lesquels le CAS2 et l'estimateur direct DIR2 diffèrent fortement

Dans quelques cas, le CAS2 sort de l'intervalle de confiance de l'estimateur direct DIR2, par exemple dans la NOGA2 19 (cokéfaction et raffinage) pour les deux sexes confondus. Compte tenu du saut de l'estimateur direct lors du renouvellement de l'échantillon, ce cas laisse plutôt suspecter la présence de valeurs extrêmes dans l'échantillon tiré en 2015. A noter que l'intervalle de confiance résulte d'une estimation (également sensible aux valeurs extrêmes) et est donc lui-même entaché d'une erreur statistique pouvant expliquer en partie le phénomène observé. On observe également que le manque de stabilité de l'estimateur direct est bien détecté par le critère « and » (cf. section 5.6.2), ce qui a pour conséquence que la validation se fait ici sur la base du CAS2. Par ailleurs, l'estimateur SAE, qui peut également être influencé par des valeurs extrêmes, n'est pas validé. En cas de reprise du projet, il faudrait observer si un meilleur traitement de ces valeurs permet de régler ce problème.



La situation se présente en revanche différemment dans la NOGA2 81 (services relatifs aux bâtiments et aménagement paysager). Le CAS2 sort de l'intervalle de confiance pour trois des quatre quatrièmes trimestres pour les femmes, sans que l'estimateur direct ne montre une si grande instabilité que dans le cas précédent. Ce type de cas devrait faire l'objet d'analyses approfondies afin de comprendre des différences observées. Le cas échéant, la validation pourrait se faire à un niveau plus agrégé ou sur la base d'un pooling de trimestres.



Enfin, on note que cette problématique n'était pas apparue pour les cantons. Pour tous les cantons, le CAS2 se situe en effet dans l'intervalle de confiance du DIR2 (ou, lorsque l'intervalle de confiance est très précis, le CAS2 reste dans les écarts tolérés autour du DIR2, cf. section 5.6.2).

7.2.2 Synthèse des analyses de validation du modèle pour les branches économiques

Les analyses effectuées indiquent que le modèle saenb produit des résultats encourageants, malgré des divergences importantes dans quelques cas qui devraient être approfondis. Il serait intéressant de tester si un meilleur traitement des valeurs extrêmes permet d'améliorer certains de ces cas. Le processus de validation devrait éventuellement être adapté (cf. discussion ci-dessus sur la NOGA2 81). Enfin un benchmarking au niveau des sections économiques, qui est nécessaire pour garantir une cohérence entre les estimations SAE et les estimations directes déjà diffusées, est problématique du fait du manque de stabilité des estimations directes pour certaines sections.

8 Conclusions et recommandations

Au terme de cette analyse, nous arrivons à la conclusion que le dernier modèle étudié (saenb) permet la diffusion sous la forme de statistiques expérimentales d'estimations SAE cantonales ajustées (benchmarking) sur les estimations directes au niveau des grandes régions, pour le total de l'emploi et l'emploi ventilé par sexe. Ces estimations SAE se basent sur l'échantillon trimestriel de la BESTA actuelle, donc avec l'augmentation de la taille de l'échantillon demandée et financée par des cantons et des villes. En revanche, nous ne recommandons pas, dans l'état actuel des choses, la diffusion de séries pour les branches économiques à des niveaux plus détaillés que ceux actuellement produits à l'aide des estimations directes.

Même si les résultats sont très encourageants, il n'a pas été possible dans le cadre de ce travail d'évaluer objectivement s'il sera possible à l'avenir de renoncer à ces densifications d'échantillon. En effet, il faudrait pour cela disposer de deux sets de données, l'un avec et l'autre sans densification d'échantillons. Cet essai pourrait être réalisé lors du prochain renouvellement de l'échantillon, soit avec les échantillons de 2024.

9 Annexes

Annexe 1 : modèle SAE retenu (saenb)

Sur l'exemple de l'emploi total des femmes (totf).

L'emploi total des hommes (totm) est estimé de façon analogue. L'estimation de l'emploi total pour les deux sexes confondus est obtenue en sommant les estimations de l'emploi des hommes et des femmes.

Principe

- Effet fixes : BETOT_F, secteur_gr (8 modalités : secteur secondaire + subdivision du secteur tertiaire par grande région), section_bis (sections avec regroupement des sections du secteur secondaire),
- Effet aléatoire pour les cantons et les noga2
- Effet aléatoire sur l'intercept ou sur la pente

Création de trois classes de taille

cl_taille = 0 : BETOT_F=0

cl_taille = 1 : 0<BETOT_F<25

cl_taille = 2 : BETOT_F ≥ 25

Pour les établissements avec BETOT_F = 0, qui sont assez nombreuses, nécessité d'estimer correctement leur intercept, ce qui intuitivement ne semble pas être garanti si ces entreprises sont trop mélangées avec des entreprises contenant un ou plusieurs emplois pour lesquelles l'intercept joue potentiellement un rôle plus important.

Pour les établissements avec cl_taille = 1 ou 2, effet aléatoire sur la pente.

→ un modèle spécifique pour chaque classe de taille

Mise en œuvre

cl_taille = 0

```
modlin0 <- lmer(totf ~ secteur_gr + section_alt + (1 | cant) + (1 | noga2), weight=poidsFin, data=reponds[reponds$cl_taille==0,])
```

cl_taille = 1

```
modlin1 <- lmer(totf ~ BETOT_F*secteur_gr + BETOT_F*section_alt + (BETOT_F | cant) + (BETOT_F | noga2), weight=poidsFin, data=reponds[reponds$cl_taille==1,])
```

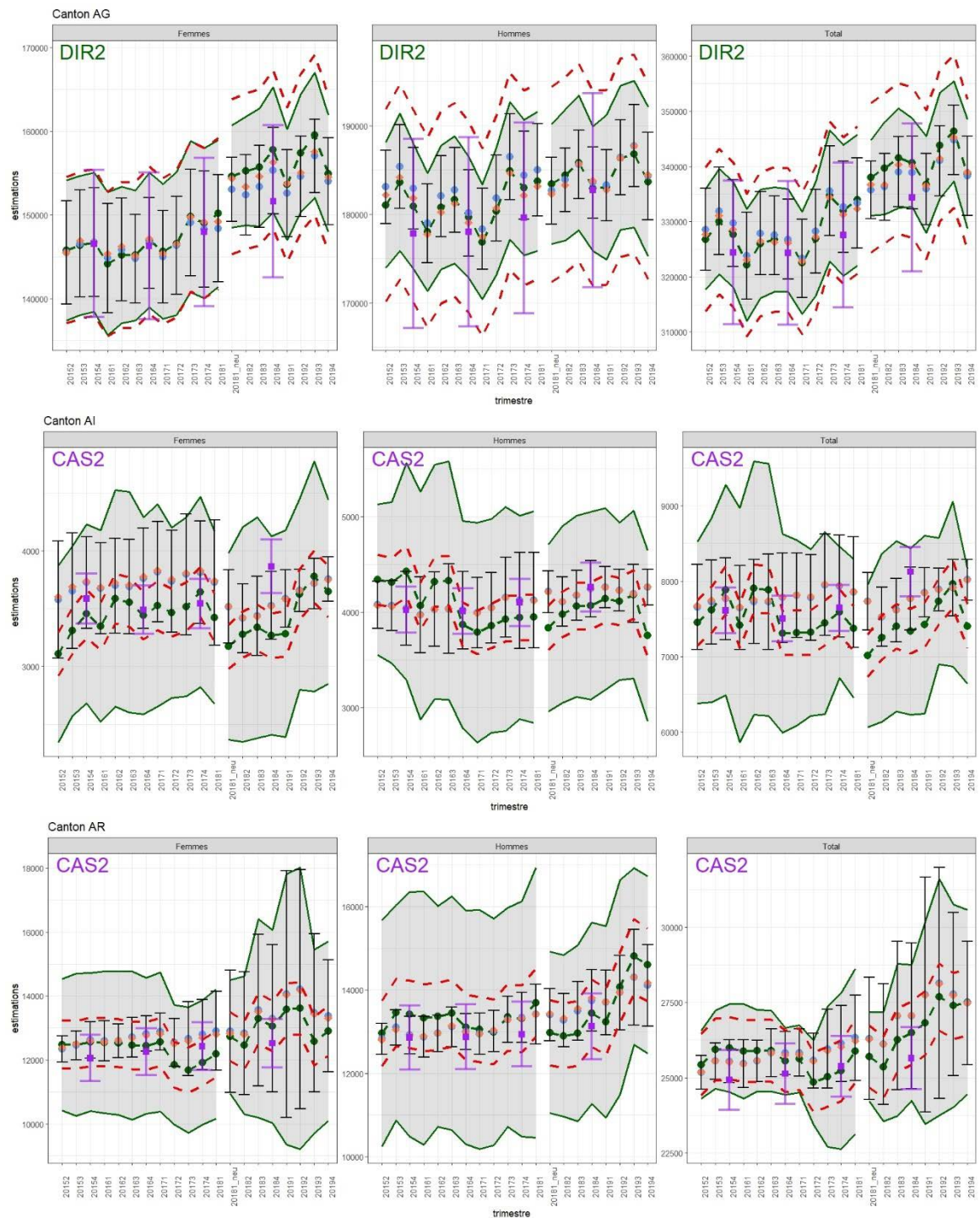
cl_taille = 2

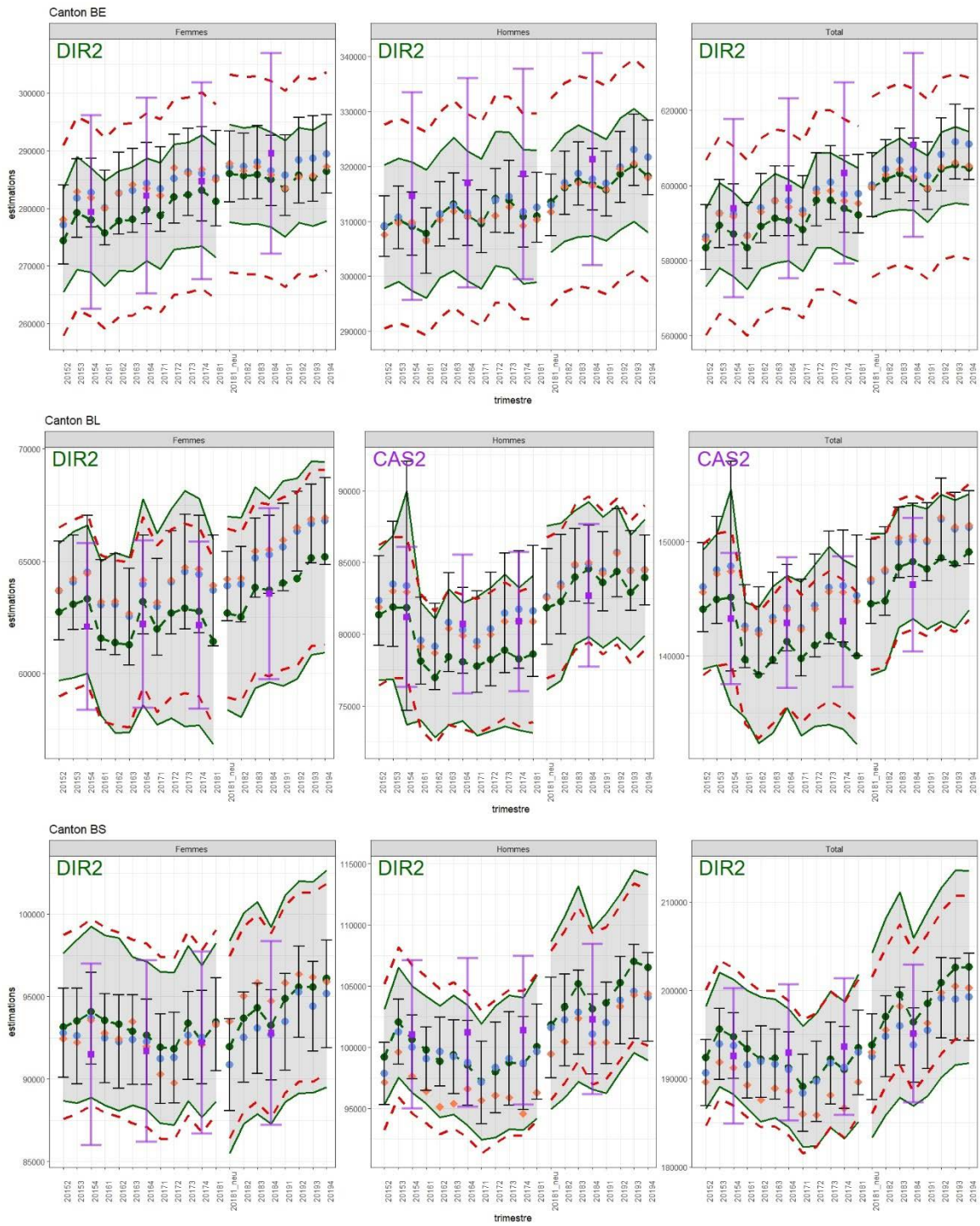
L'algorithme de la fonction lmer telle qu'elle est utilisée pour la classe de taille 1 ne converge pas. D'où l'idée de la procédure suivante.

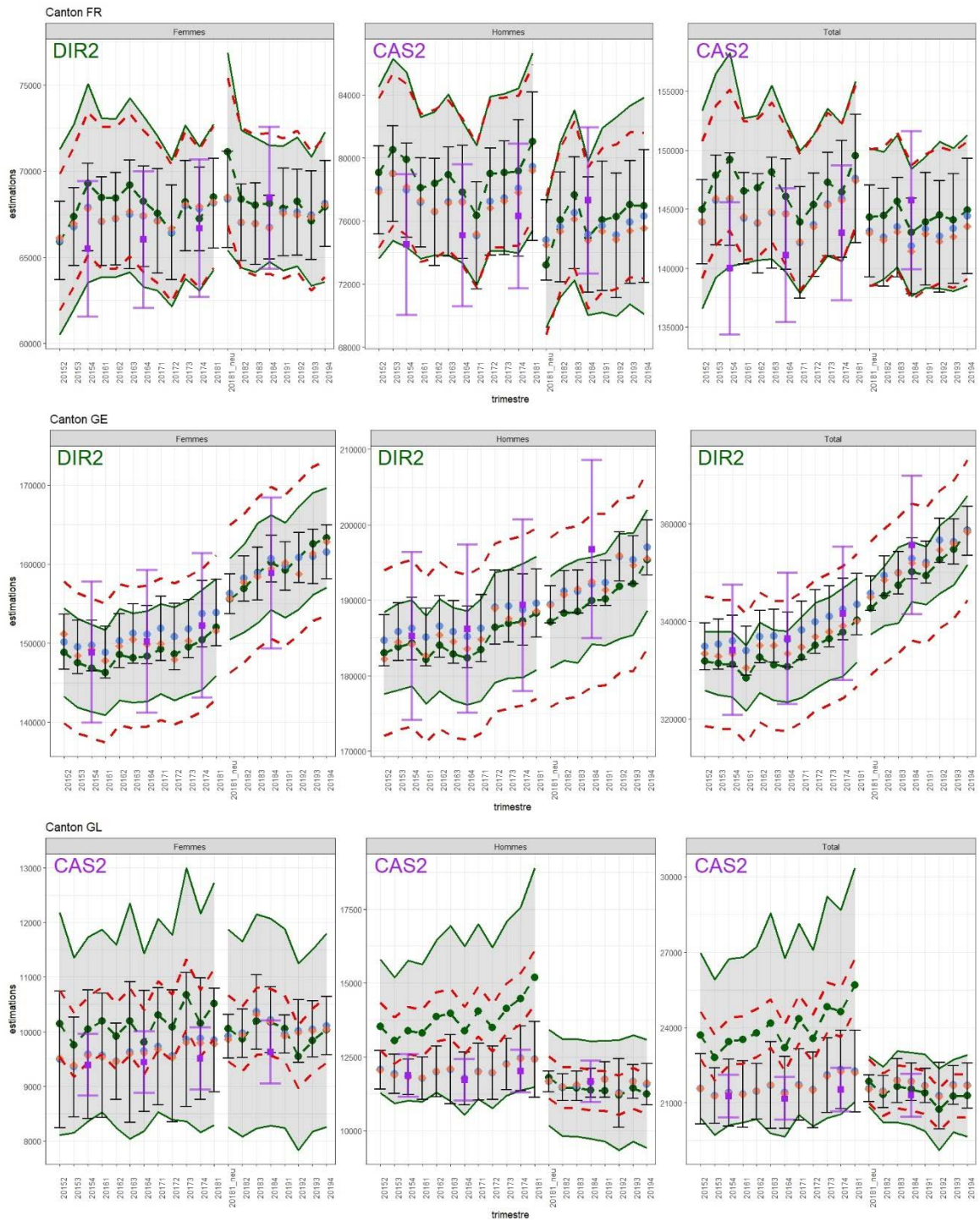
- On pose $R = \text{totf} / \text{BETOT\_F}$
- ```
modlin2 <- lmer(R ~ secteur_gr + section_alt + (1 | cant) + (1 | noga2), weight=poidsFin, data=reponds[reponds$cl_taille==2,])
```

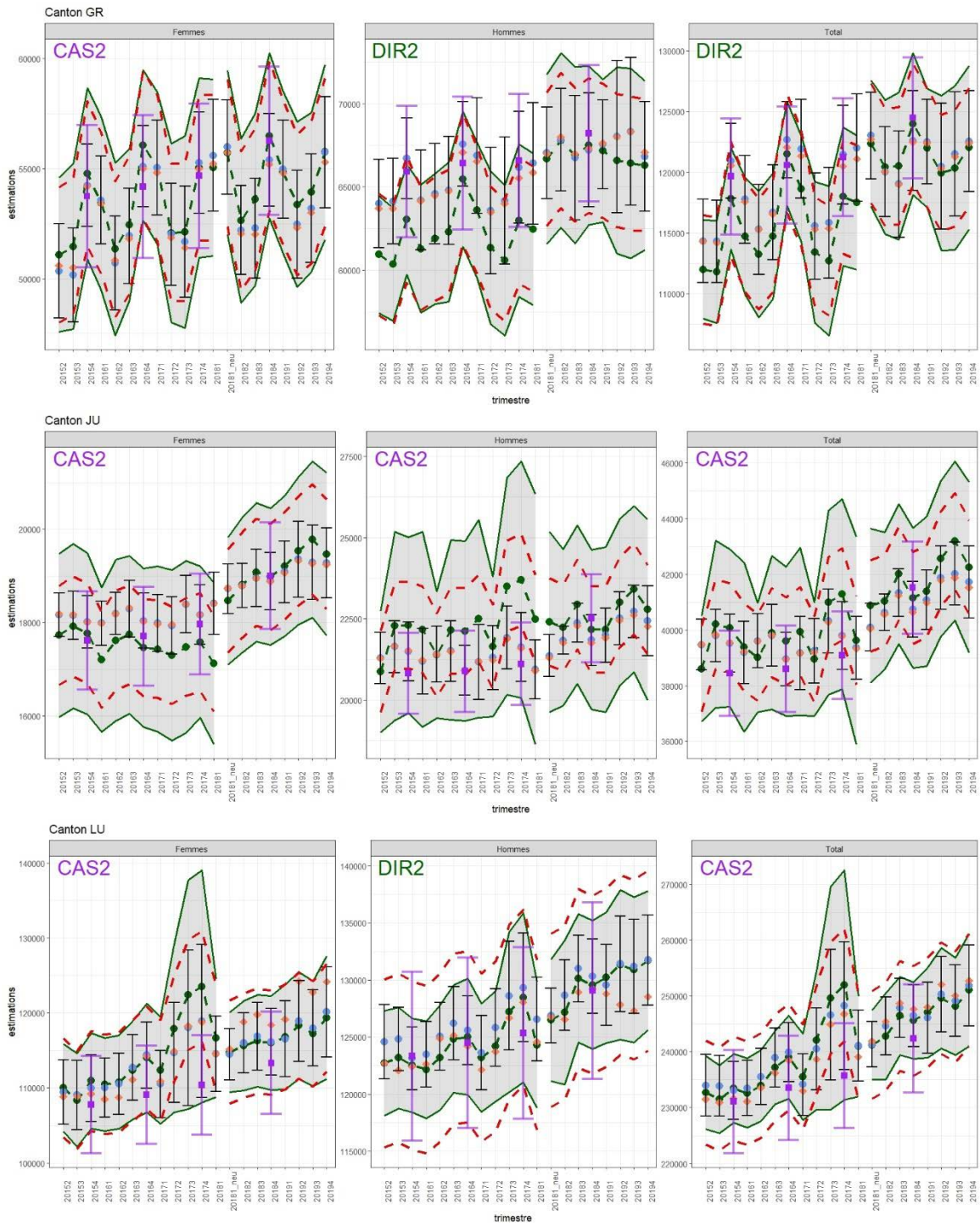
Prédictions de totf en multipliant les prédictions de R par BETOT\_F

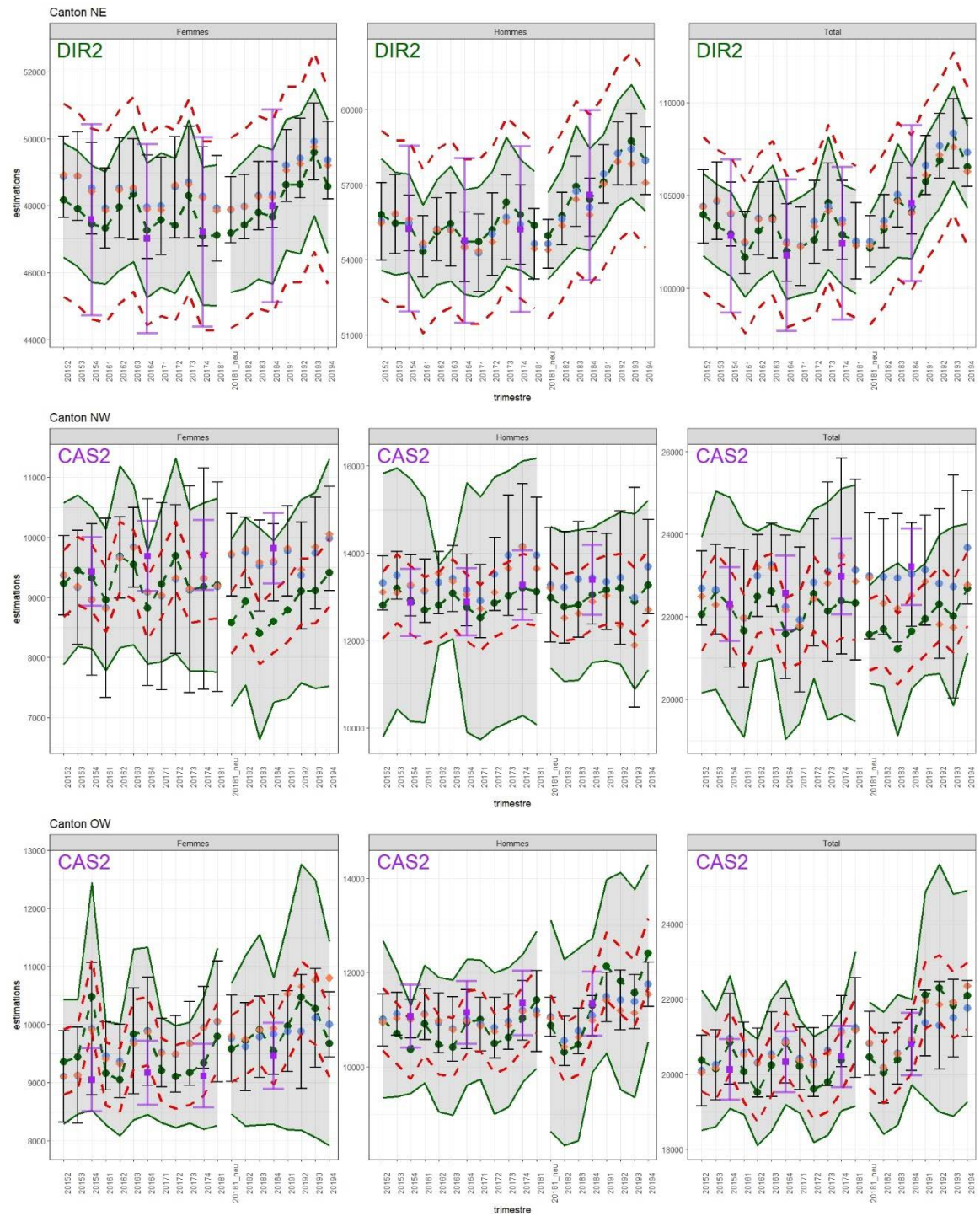
## Annexe 2 : séries SAE (cantons et ville de Zurich)



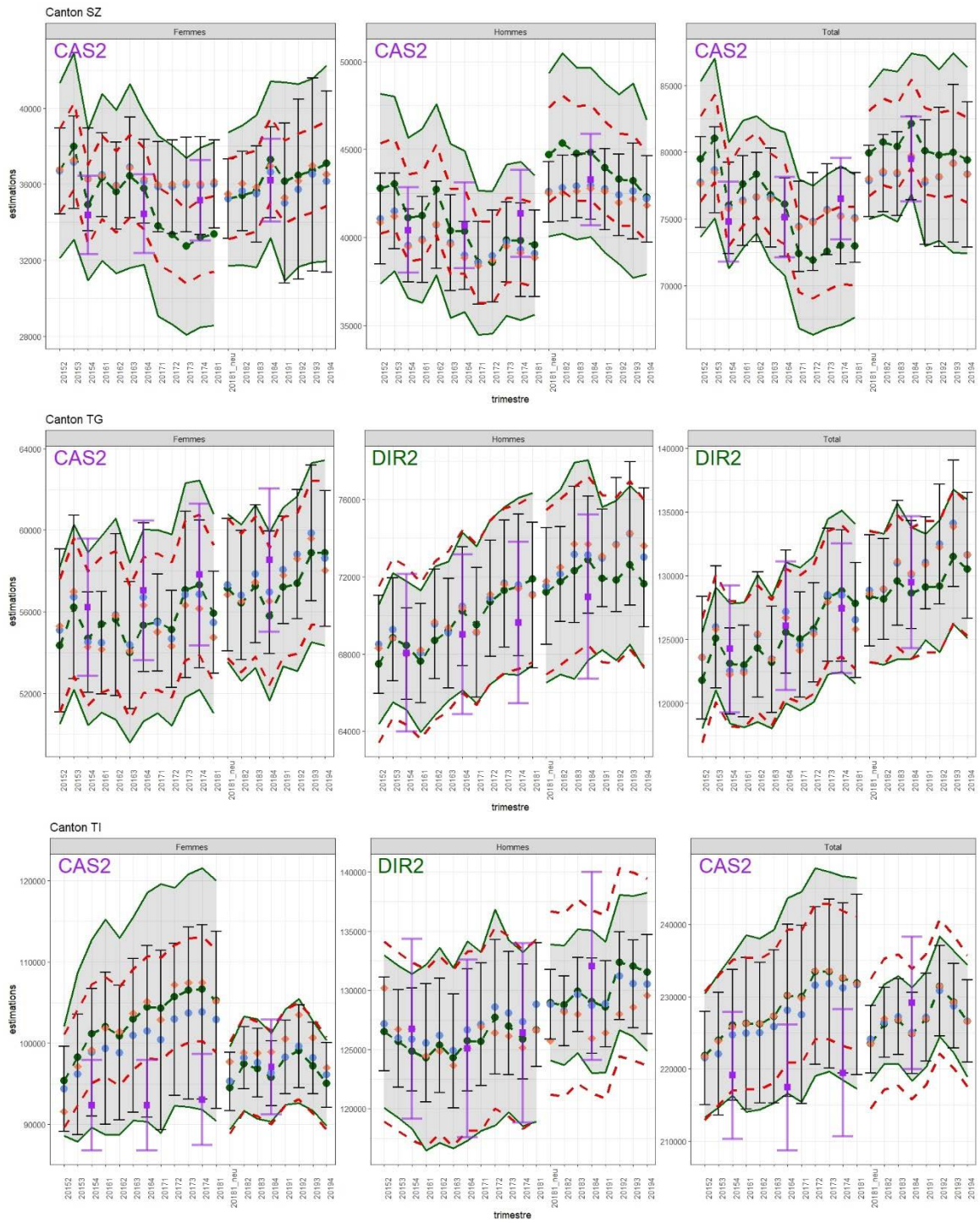


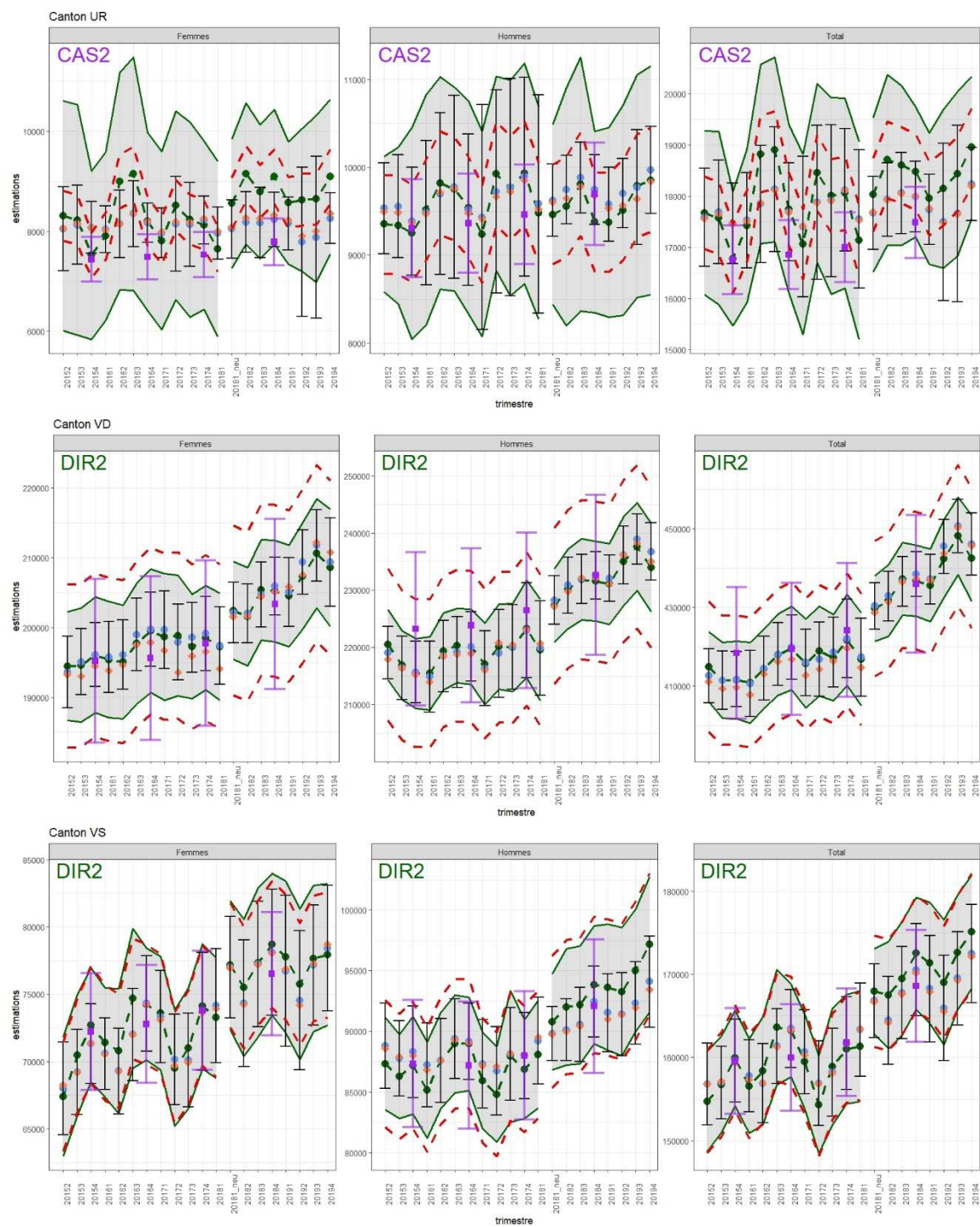


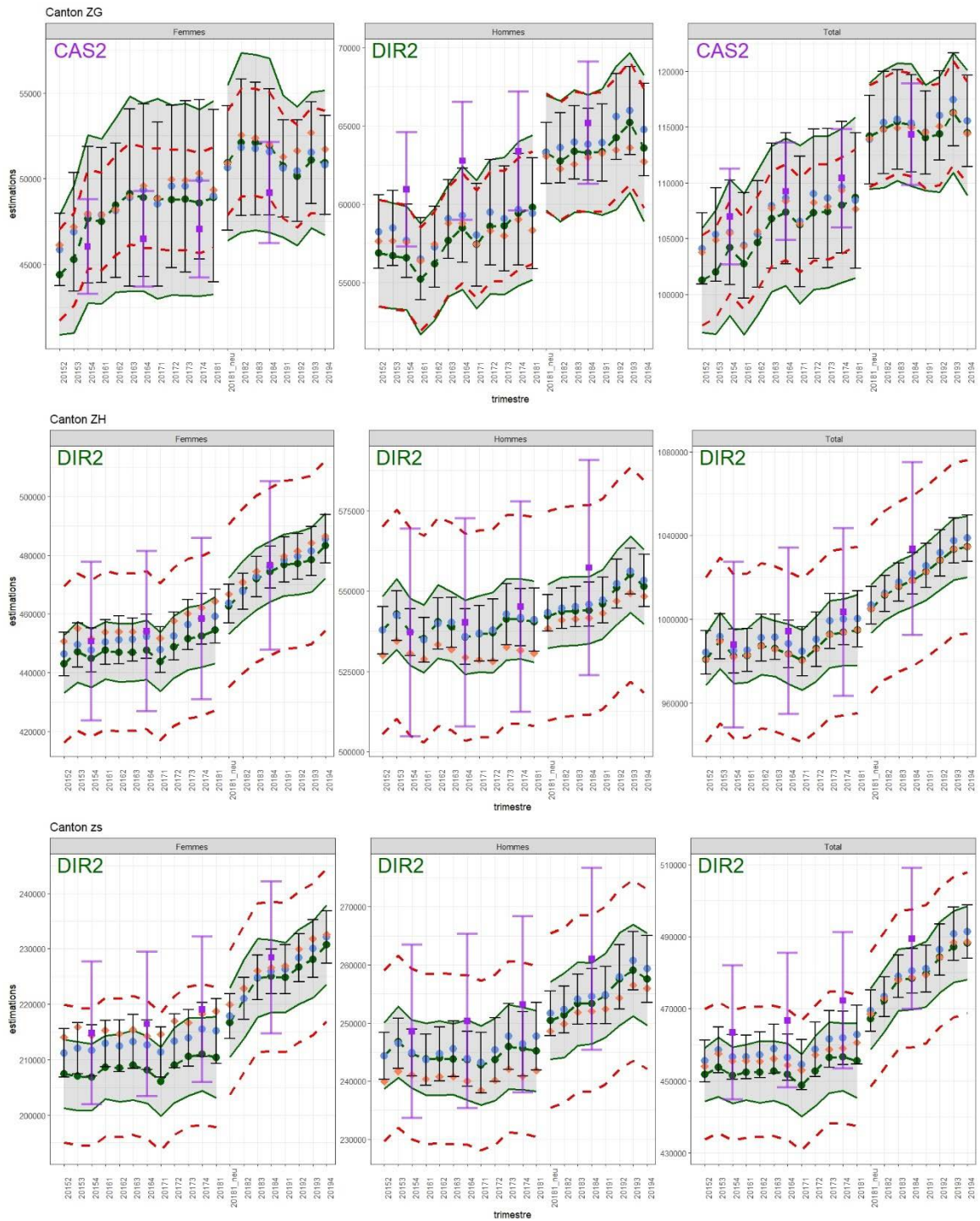




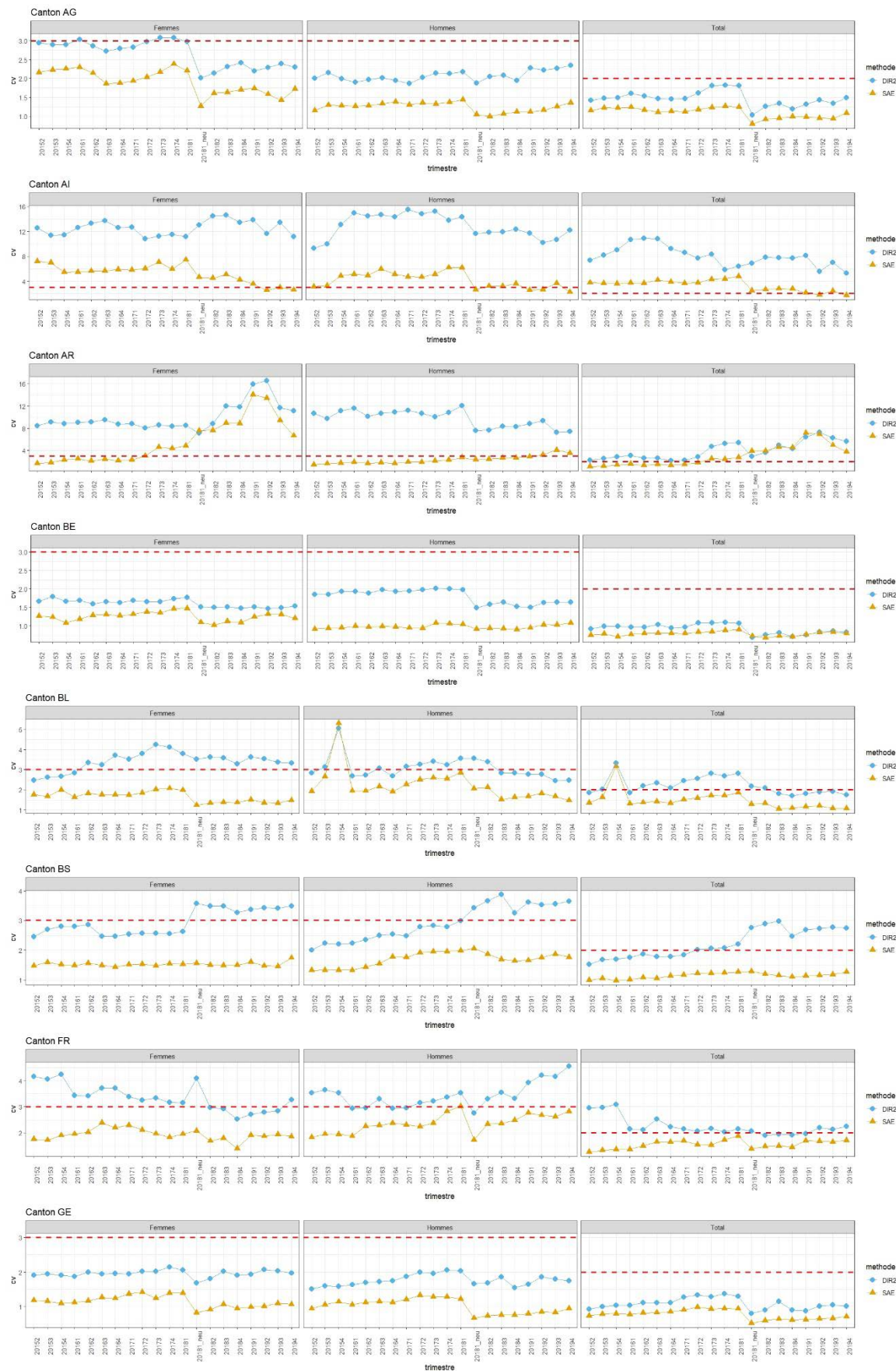


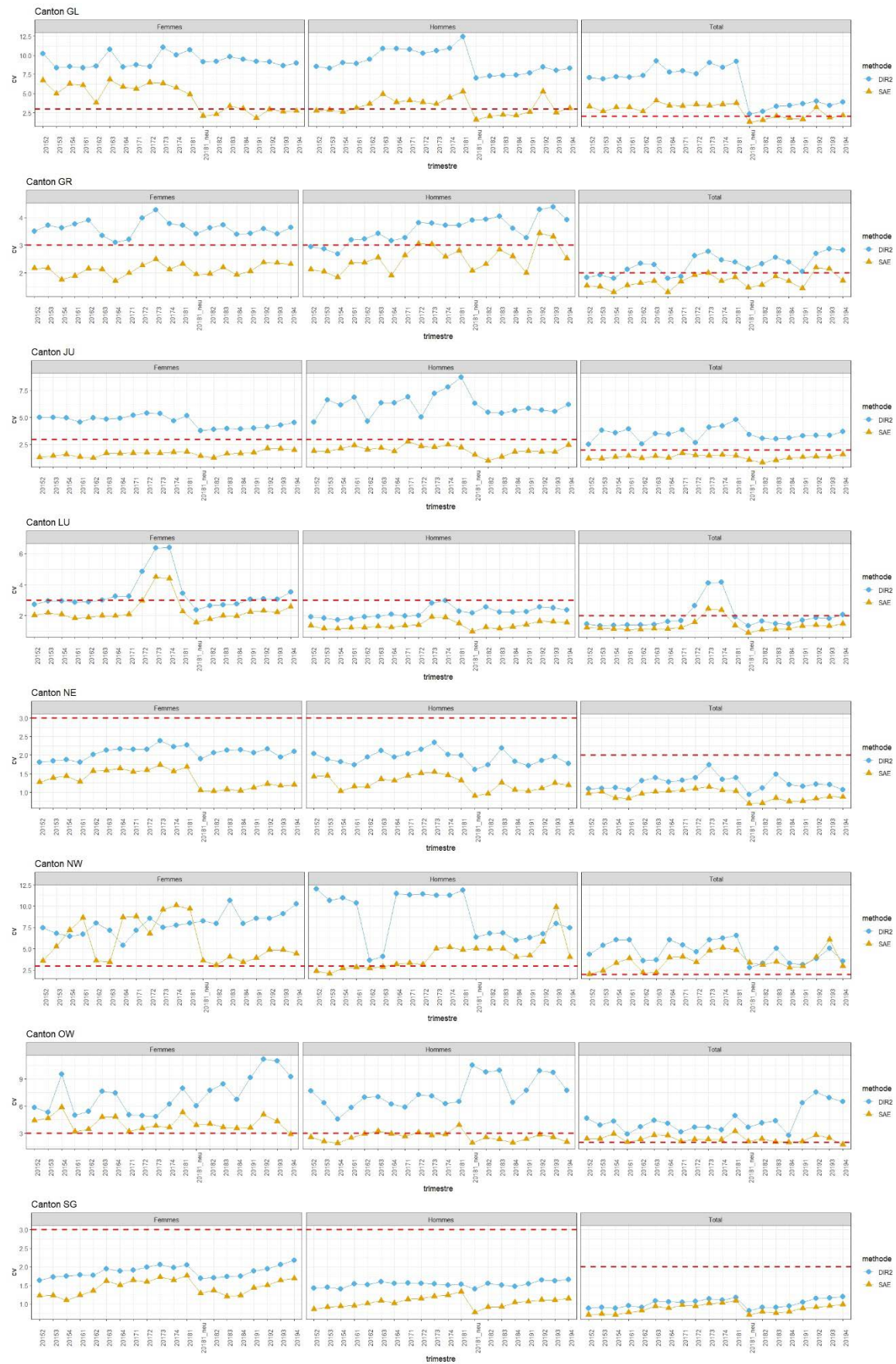


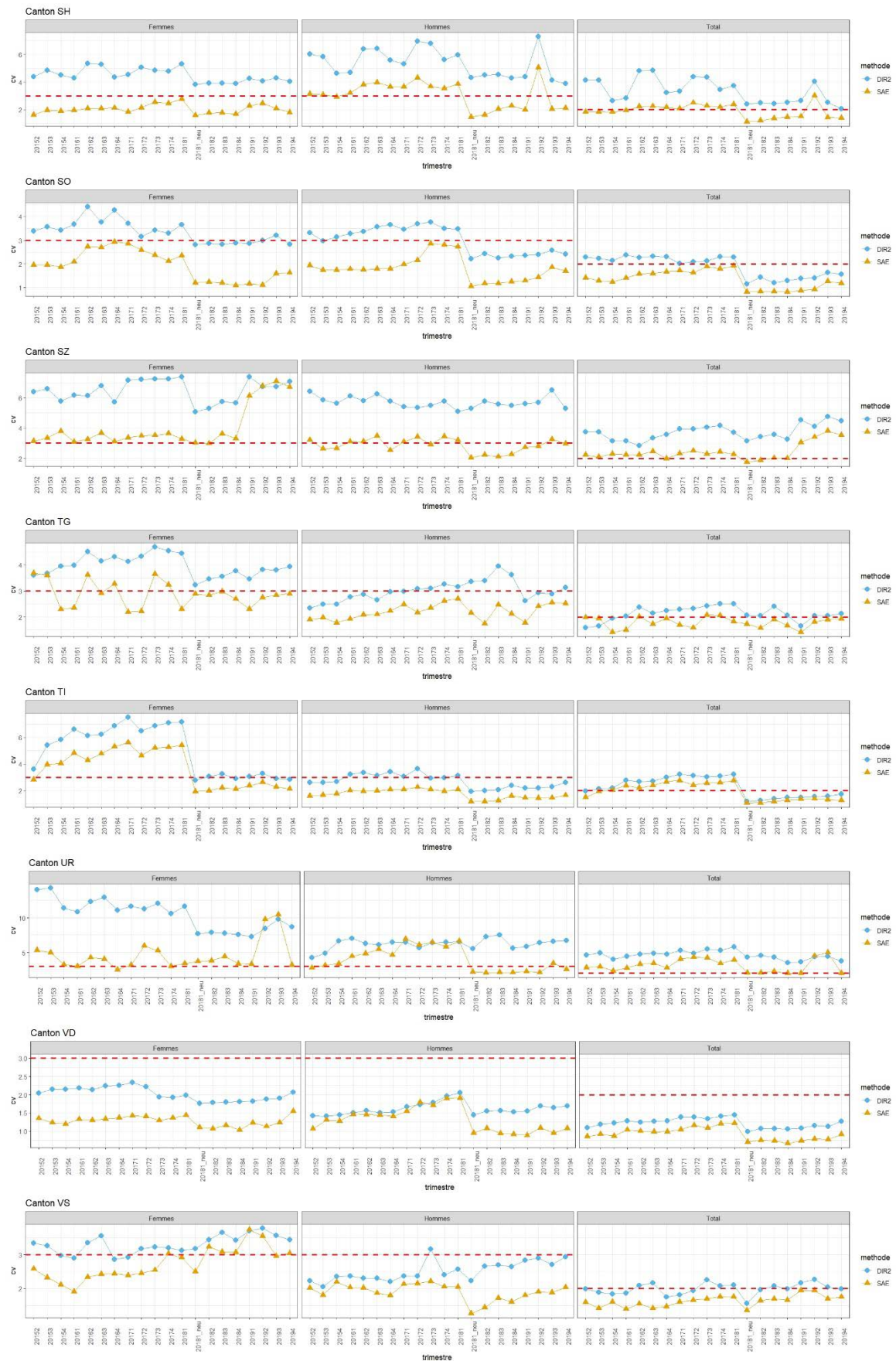




## Annexe 3 : CV des estimations SAE cantonales, sans benchmarking









## Annexe 4 : validation du modèle pour les branches économiques

Application du critère AND pour les sections économiques, les NOGA (NOGA\_OFS50)  
et les NOGA2

| Section | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|---------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total   | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Hommes  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| Femmes  | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |

| NOGA   | 5.9 | 10.2 | 13.5 | 16.8 | 19.2 | 21   | 22.3 | 24.5 | 26 | 27 | 28 | 29.3 | 31.3 | 35 | 36.9 | 41.2 | 43   | 45 | 46 | 47 | 49 | 50.1 | 52   |      |
|--------|-----|------|------|------|------|------|------|------|----|----|----|------|------|----|------|------|------|----|----|----|----|------|------|------|
| Total  | 0   | 1    | 0    | 1    | 1    | 1    | 1    | 1    | 1  | 1  | 1  | 0    | 1    | 1  | 0    | 1    | 1    | 1  | 1  | 1  | 1  | 0    | 1    |      |
| Hommes | 1   | 1    | 0    | 1    | 1    | 1    | 1    | 1    | 1  | 1  | 1  | 1    | 1    | 1  | 1    | 1    | 1    | 1  | 1  | 0  | 1  | 0    | 1    |      |
| Femmes | 0   | 0    | 0    | 0    | 0    | 1    | 0    | 0    | 0  | 0  | 0  | 0    | 0    | 0  | 0    | 0    | 0    | 0  | 1  | 1  | 0  | 0    | 1    |      |
| NOGA   | 53  | 55   | 56   | 58.6 | 61   | 62.3 | 64   | 65   | 66 | 68 | 69 | 70   | 71   | 72 | 73.5 | 78   | 79.2 | 84 | 85 | 86 | 87 | 88   | 90.3 | 94.6 |
| Total  | 0   | 1    | 0    | 0    | 0    | 0    | 1    | 1    | 0  | 0  | 0  | 0    | 1    | 0  | 0    | 0    | 1    | 1  | 1  | 1  | 1  | 1    | 0    | 0    |
| Hommes | 0   | 1    | 0    | 1    | 0    | 0    | 1    | 1    | 1  | 0  | 0  | 1    | 1    | 1  | 0    | 0    | 0    | 1  | 1  | 0  | 0  | 0    | 0    | 0    |
| Femmes | 0   | 1    | 0    | 0    | 0    | 0    | 1    | 1    | 1  | 0  | 0  | 0    | 0    | 0  | 0    | 0    | 0    | 0  | 1  | 1  | 1  | 1    | 0    | 0    |

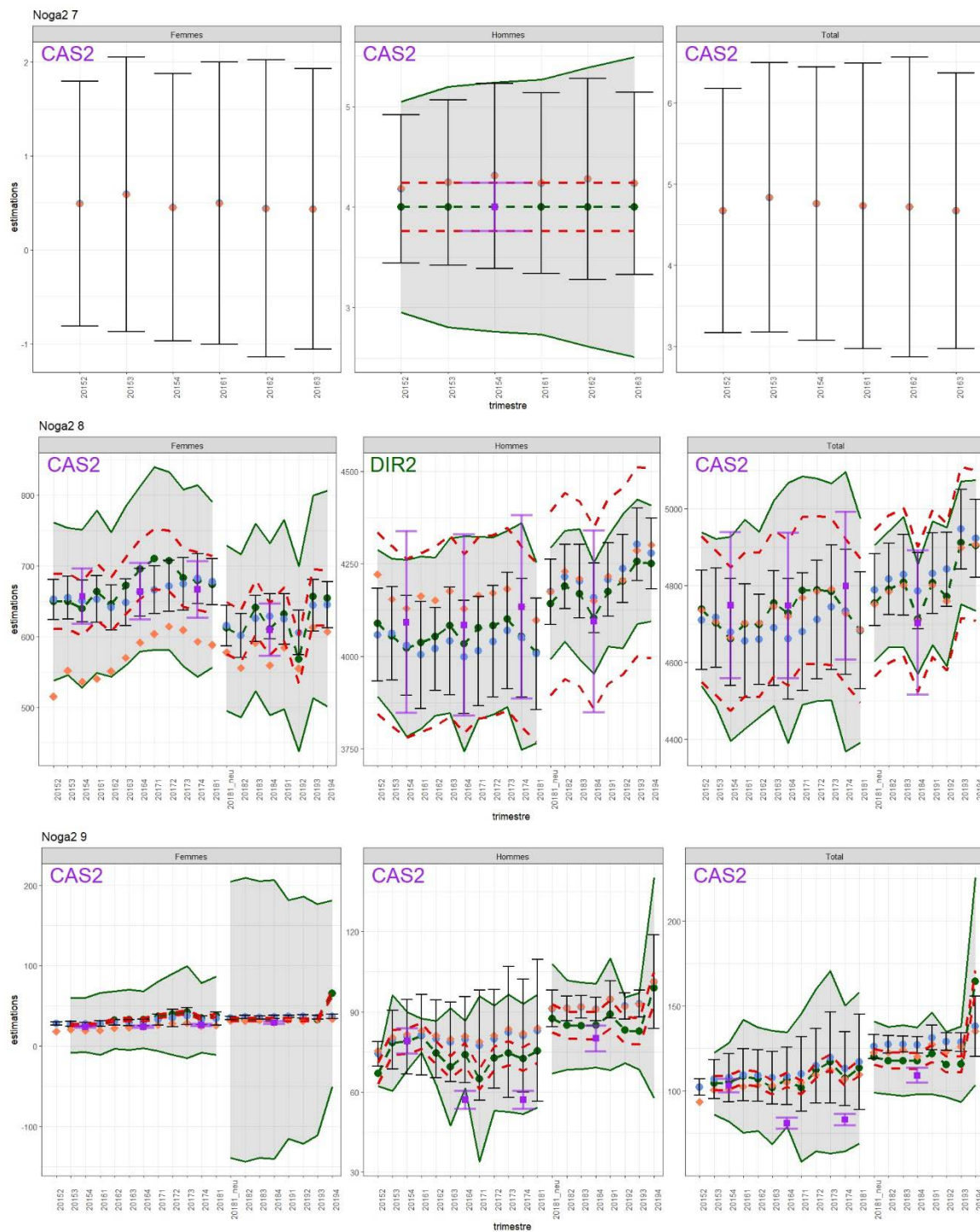
|        |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|--------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| NOGA2  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 35 | 36 | 37 | 38 | 39 | 41 | 42 | 43 | 45 | 46 | 47 | 49 | 50 |
| Total  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 0  | 1  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 0  | 1  | 1  | 1  | 1  | 1  | 0  |
| Hommes | 0  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 0  | 0  | 1  | 0  | 0  | 1  | 1  | 0  | 0  | 1  | 0  | 1  | 1  | 1  | 1  | 0  | 1  | 0  |    |
| Femmes | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  |
| NOGA2  | 51 | 52 | 53 | 55 | 56 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 77 | 78 | 79 | 80 | 81 | 82 | 84 | 85 | 86 | 87 | 88 | 90 | 91 | 92 | 93 | 94 | 95 | 96 |
| Total  | 0  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| Hommes | 0  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| Femmes | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |    |

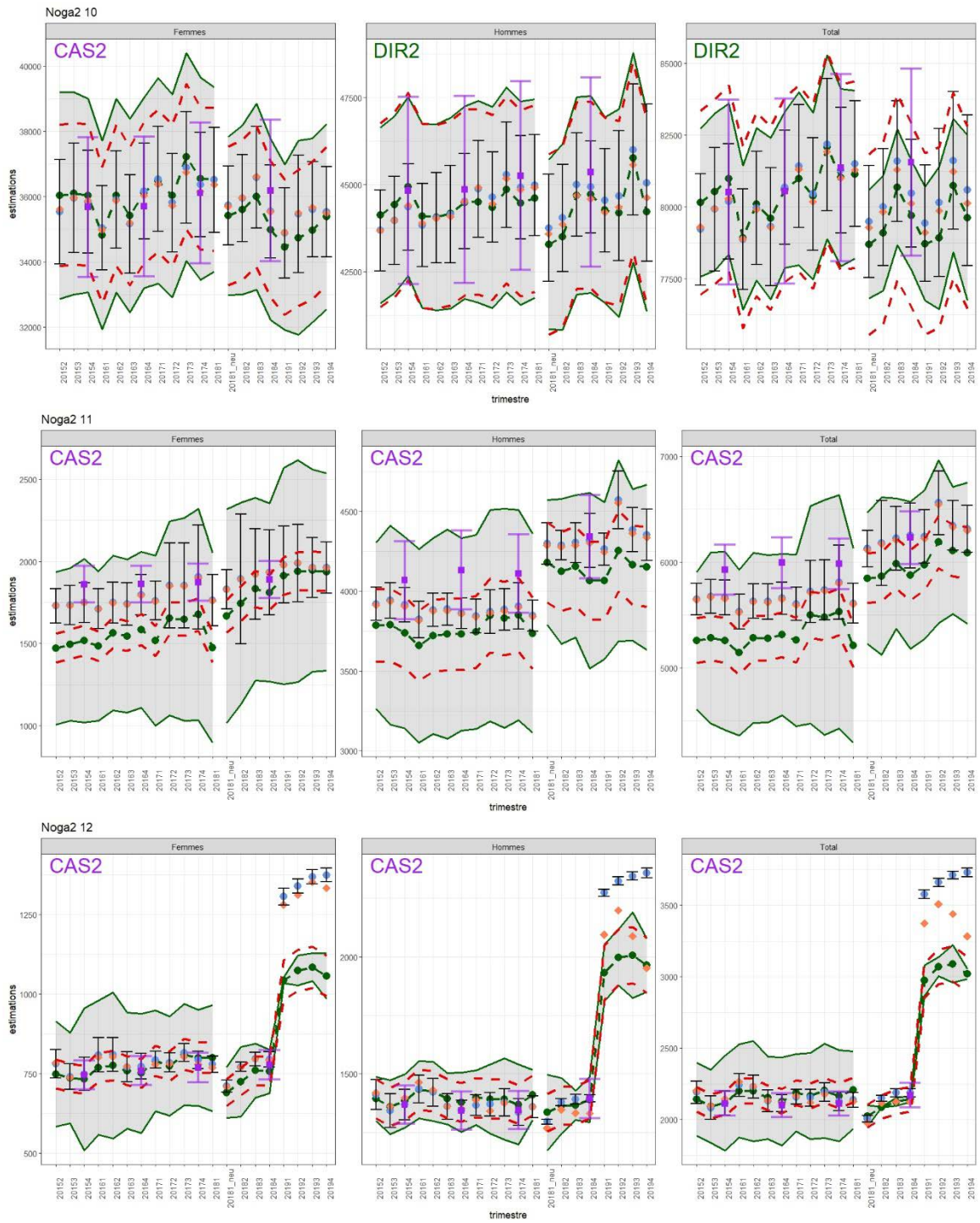
## Annexe 5 : résumé des CV des estimations directes (DIR2) pour les regroupements de divisions économiques NOGA\_OFS50

Période 2015-2019. Les cellules marquées en rouge correspondent à celles dont les valeurs excèdent la limite de 2% pour l'emploi total et de 3% par sexe.

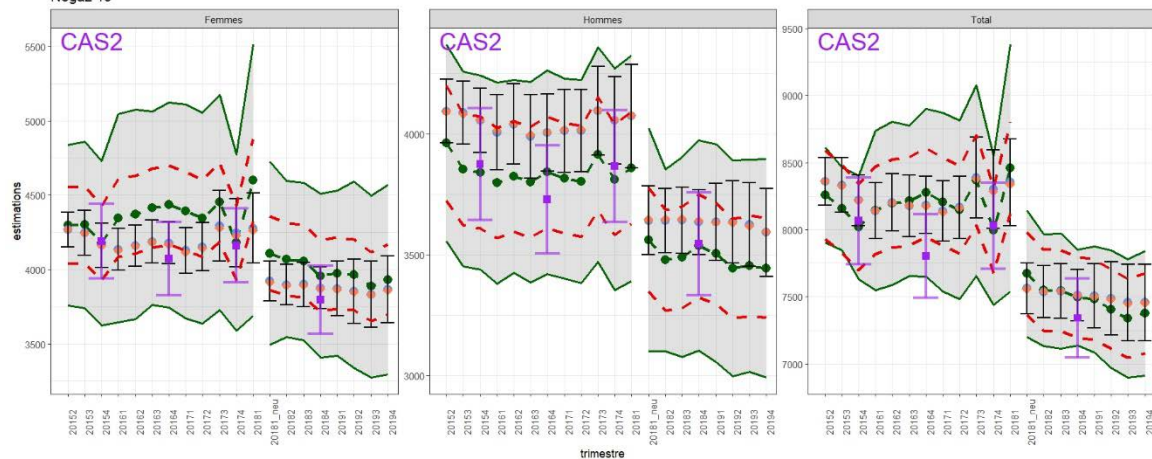
| NOGA50          | Femmes (3%) |        |        | Hommes (3%) |        |        | Total (2%) |        |        |
|-----------------|-------------|--------|--------|-------------|--------|--------|------------|--------|--------|
|                 | CV moy      | CV min | CV max | CV moy      | CV min | CV max | CV moy     | CV min | CV max |
| 5-9             | 9.77        | 7.84   | 11.81  | 2.54        | 1.59   | 3.48   | 2.56       | 1.62   | 3.54   |
| 10-12           | 3.85        | 3.4    | 4.08   | 2.74        | 2.52   | 2.94   | 1.49       | 1.1    | 1.8    |
| 13-15           | 4.79        | 3.64   | 6.06   | 5.41        | 4.67   | 6.32   | 2.57       | 1.67   | 3.91   |
| 16-18           | 5.01        | 4.42   | 5.73   | 2.35        | 2.04   | 2.69   | 1.77       | 1.39   | 2.29   |
| 19-20           | 5.48        | 4.76   | 6.34   | 2.73        | 2.1    | 3.44   | 1.9        | 1.16   | 2.52   |
| 21              | 2.51        | 2.19   | 2.92   | 2.17        | 1.35   | 2.78   | 1.37       | 1.08   | 1.64   |
| 22-23           | 5.12        | 4.49   | 5.89   | 2.14        | 1.7    | 2.54   | 1.53       | 1.08   | 1.98   |
| 24-25           | 5.06        | 4.82   | 5.33   | 1.81        | 1.61   | 2.05   | 1.44       | 1.06   | 1.72   |
| 26              | 3.91        | 3.17   | 4.85   | 2.26        | 1.85   | 3.16   | 1.78       | 1.25   | 2.89   |
| 27              | 7.61        | 6.71   | 8.83   | 2.48        | 2.1    | 2.99   | 1.89       | 1.12   | 2.8    |
| 28              | 6.37        | 5.46   | 7.76   | 2.1         | 1.66   | 2.86   | 1.62       | 1.08   | 2.38   |
| 29-30           | 6.55        | 5.25   | 9.01   | 2.28        | 1.53   | 2.7    | 2.33       | 1.38   | 3.16   |
| 31-33           | 4.98        | 4.13   | 5.55   | 2.38        | 1.99   | 2.68   | 1.68       | 1.18   | 2.01   |
| 35              | 3.74        | 2.85   | 4.61   | 1.44        | 0.99   | 1.92   | 1.22       | 1      | 1.68   |
| 36-39           | 11.37       | 9.47   | 17.28  | 2.43        | 1.78   | 3.41   | 2.89       | 1.86   | 5.76   |
| 41-42           | 7.5         | 6.05   | 10.78  | 1.67        | 1.38   | 1.97   | 1.6        | 1.23   | 2.16   |
| 43              | 5.05        | 4.55   | 5.67   | 1.79        | 1.43   | 2.18   | 1.61       | 1.19   | 2.04   |
| 45              | 4.99        | 4.67   | 5.58   | 1.94        | 1.6    | 2.2    | 1.72       | 1.34   | 2.02   |
| 46              | 3.6         | 2.95   | 4.07   | 2.26        | 1.97   | 2.6    | 1.67       | 1.14   | 2.07   |
| 47              | 2.74        | 2.31   | 3.21   | 5.34        | 4.43   | 5.78   | 1.7        | 1.06   | 2.16   |
| 49              | 5.04        | 4.15   | 6.29   | 1.88        | 1.26   | 2.23   | 1.74       | 1.2    | 2.08   |
| 50-51           | 8.01        | 5.94   | 11.03  | 5.13        | 3.65   | 6.76   | 4.49       | 3.08   | 7.07   |
| 52              | 3.17        | 2.65   | 3.61   | 2.08        | 1.33   | 3      | 1.49       | 0.8    | 2.56   |
| 53              | 5           | 2.41   | 6.79   | 6.64        | 4.64   | 10.43  | 5.51       | 3.29   | 7.23   |
| 55              | 3.05        | 2.67   | 3.39   | 2.95        | 2.51   | 3.25   | 2.22       | 1.76   | 2.51   |
| 56              | 3.67        | 3.13   | 4.35   | 5.16        | 4.02   | 7.39   | 2.91       | 2.25   | 4.04   |
| 58-60           | 4.23        | 3.71   | 4.77   | 3.16        | 2.65   | 4.03   | 2.59       | 2.01   | 3.29   |
| 61              | 3.81        | 1.51   | 8.05   | 3.29        | 1      | 6.39   | 2.69       | 0.92   | 5.2    |
| 62-63           | 5.44        | 4.88   | 6.56   | 3           | 2.06   | 4.75   | 2.76       | 2.12   | 4.24   |
| 64              | 1.73        | 1.52   | 2.01   | 1.31        | 1.08   | 1.52   | 0.83       | 0.62   | 1.09   |
| 65              | 2.58        | 1.9    | 3.56   | 2.52        | 2.11   | 3.05   | 1.22       | 0.75   | 1.89   |
| 66              | 3.62        | 2.81   | 4.33   | 3.28        | 2.56   | 4.33   | 2.42       | 1.59   | 3.53   |
| 68              | 5.51        | 3.96   | 7.03   | 4.84        | 3.92   | 5.63   | 3.98       | 2.6    | 5.25   |
| 69              | 3.67        | 3.34   | 4.21   | 3.66        | 3.35   | 4.06   | 2.47       | 2.08   | 2.92   |
| 70              | 4.05        | 3.35   | 4.55   | 3.2         | 2.22   | 4.26   | 2.54       | 1.5    | 3.46   |
| 71              | 3.96        | 3.53   | 4.78   | 2.65        | 2.28   | 3.35   | 1.87       | 1.42   | 2.51   |
| 72              | 4.54        | 3.61   | 6.91   | 3.38        | 2.85   | 3.88   | 2.66       | 1.84   | 3.49   |
| 73-75           | 4.81        | 3.83   | 5.49   | 4.27        | 3.54   | 4.97   | 2.52       | 1.74   | 3.04   |
| 78              | 7.24        | 5.47   | 8.52   | 5.36        | 3.18   | 7.1    | 4.6        | 2.64   | 5.82   |
| 77-82 (sans 78) | 4.32        | 3.52   | 4.67   | 3.65        | 3.32   | 3.96   | 2.25       | 1.64   | 2.96   |
| 84              | 2.5         | 1.32   | 4.55   | 1.23        | 0.92   | 1.87   | 1.29       | 0.58   | 2.87   |
| 85              | 1.89        | 1.71   | 2.3    | 2.23        | 2      | 2.57   | 1.41       | 1.19   | 1.55   |
| 86              | 1.98        | 1.25   | 2.69   | 3.58        | 3.12   | 3.93   | 1.53       | 1.07   | 1.94   |
| 87              | 1.51        | 1.26   | 1.8    | 3.96        | 3.47   | 4.34   | 1.31       | 0.93   | 1.64   |
| 88              | 2.65        | 2.2    | 3.06   | 4.84        | 4.4    | 5.21   | 1.84       | 1.2    | 2.33   |
| 90-93           | 4.3         | 3.44   | 5.11   | 4.47        | 3.79   | 5.48   | 3.07       | 2.32   | 3.98   |
| 94-96           | 3.31        | 3.09   | 3.6    | 5.95        | 5.22   | 6.78   | 2.39       | 2.1    | 2.56   |

## Annexe 6 : séries temporelles SAE pour les NOGA2

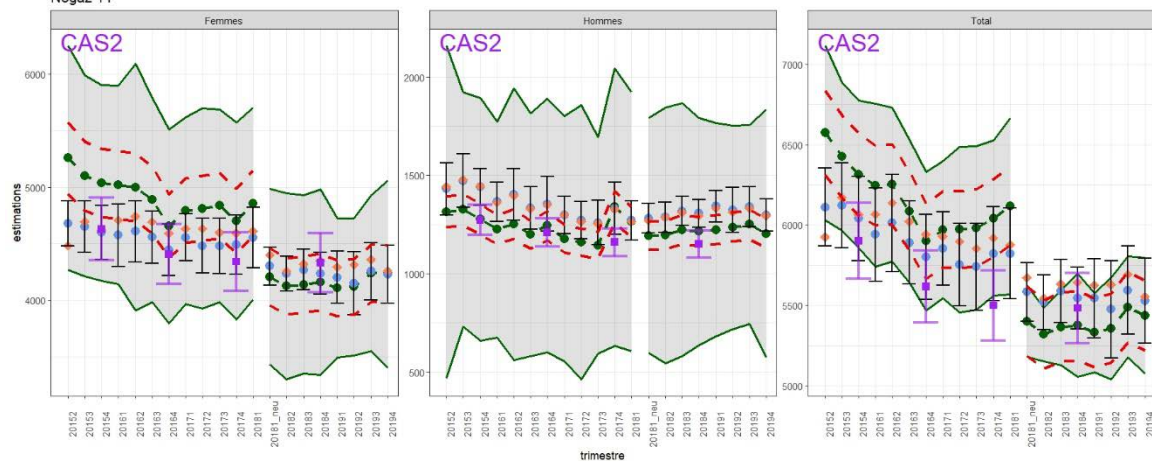




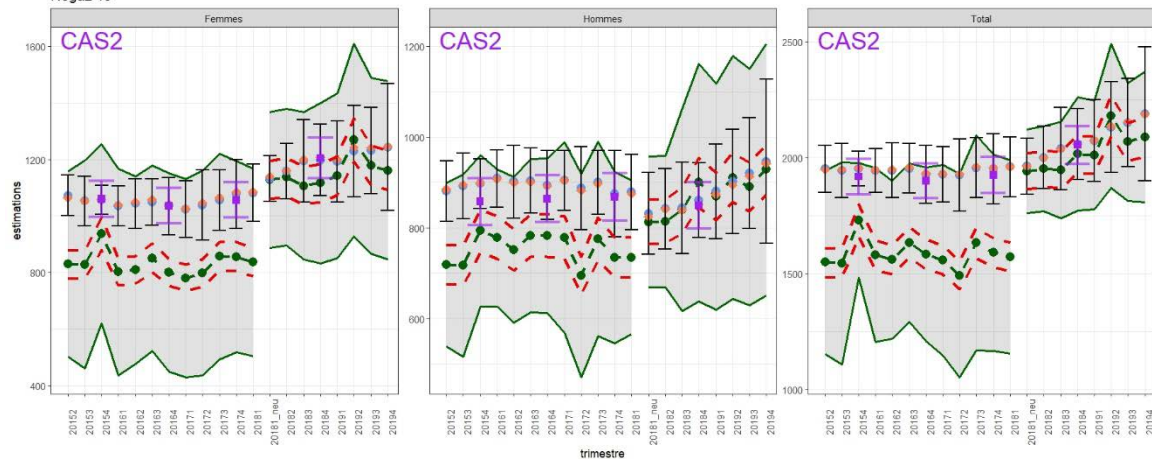
Noga2 13



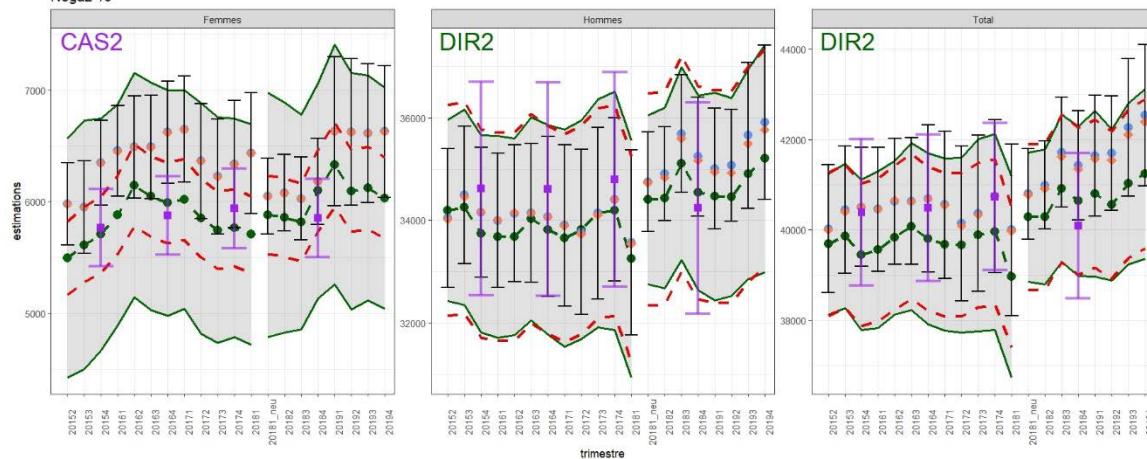
Noga2 14



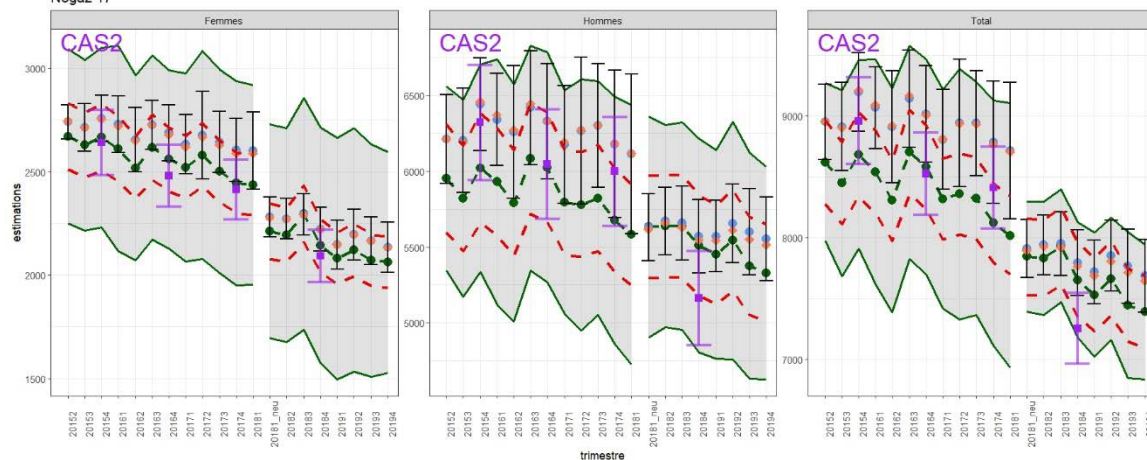
Noga2 15



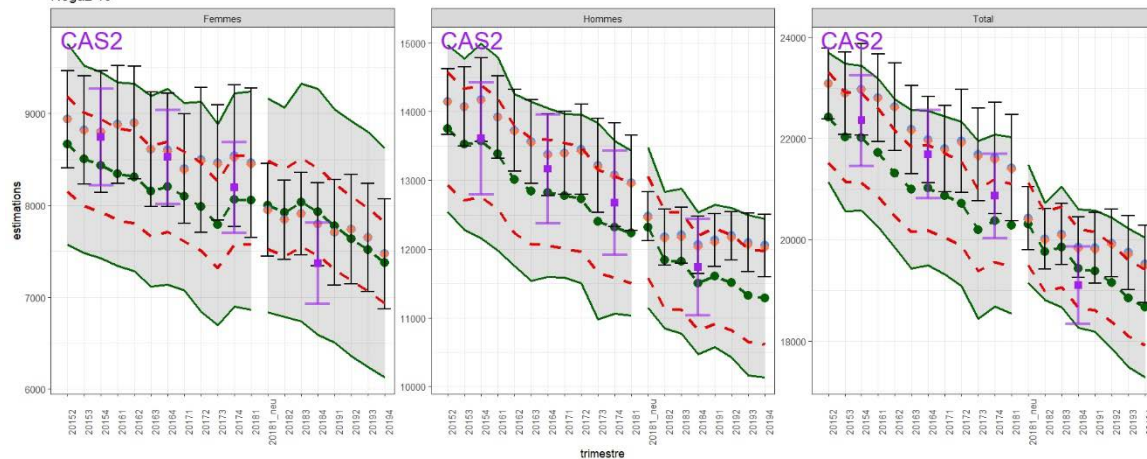
Noga2 16



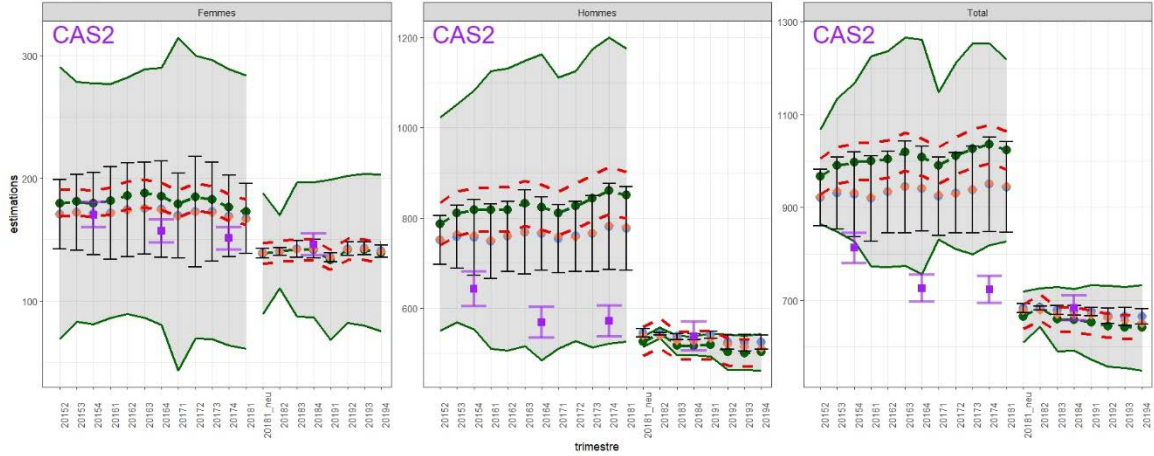
Noga2 17



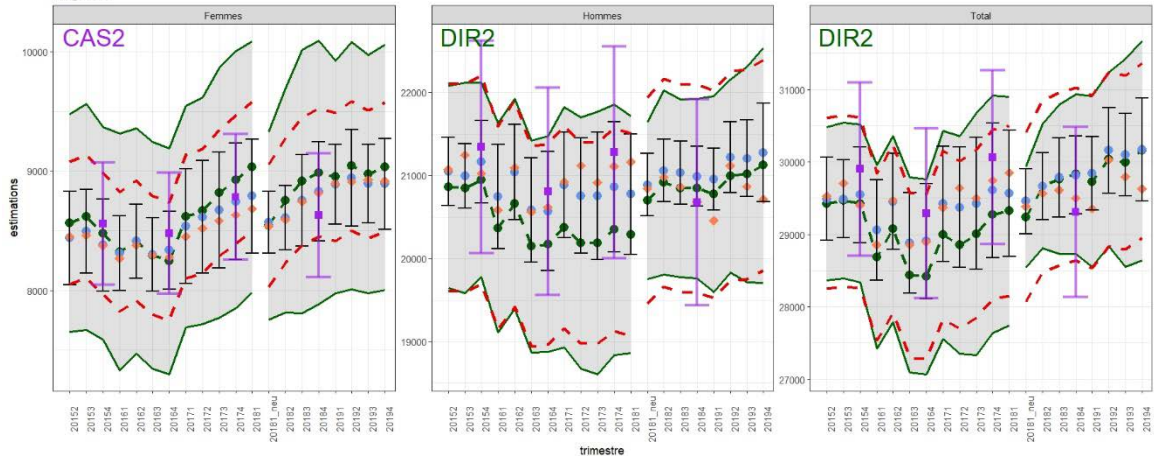
Noga2 18



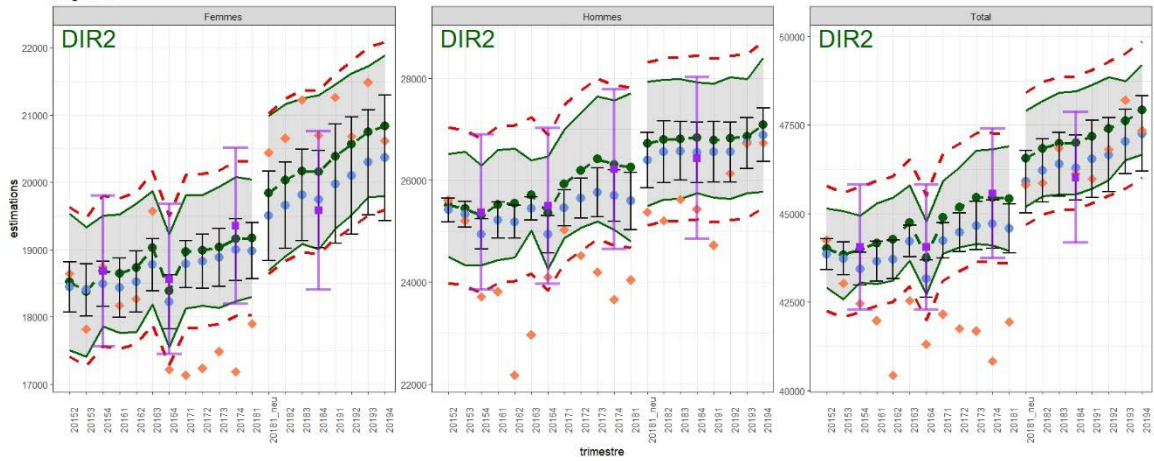
Noga2 19



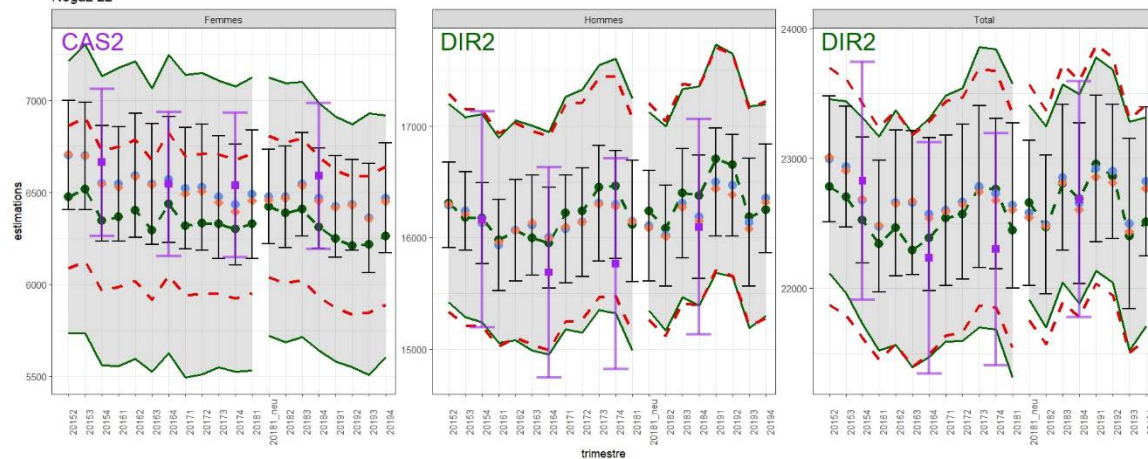
Noga2 20



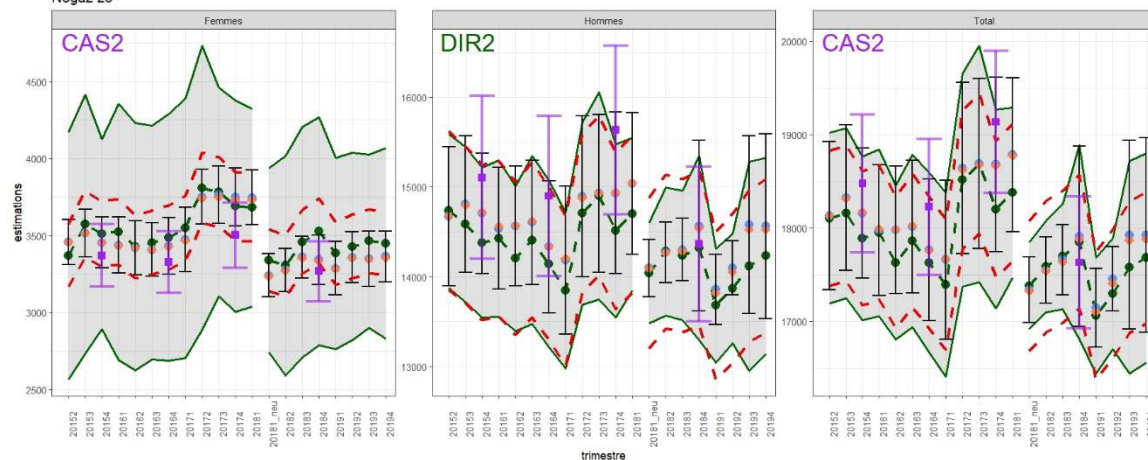
Noga2 21



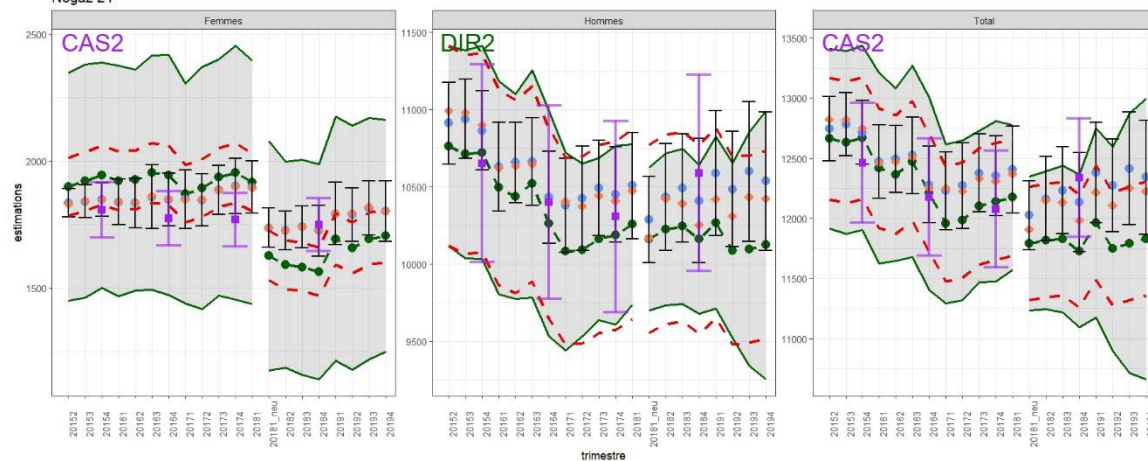
Noga2 22



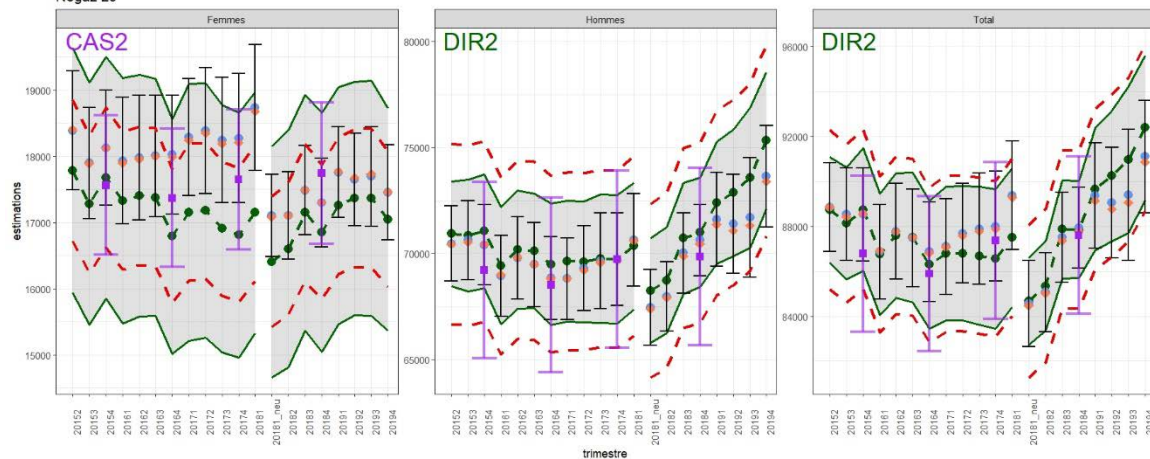
Noga2 23



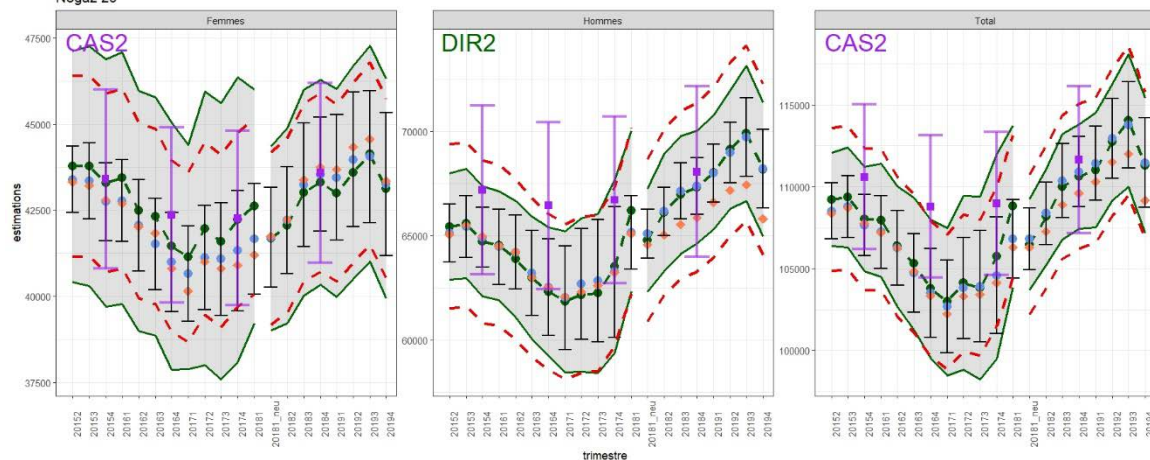
Noga2 24



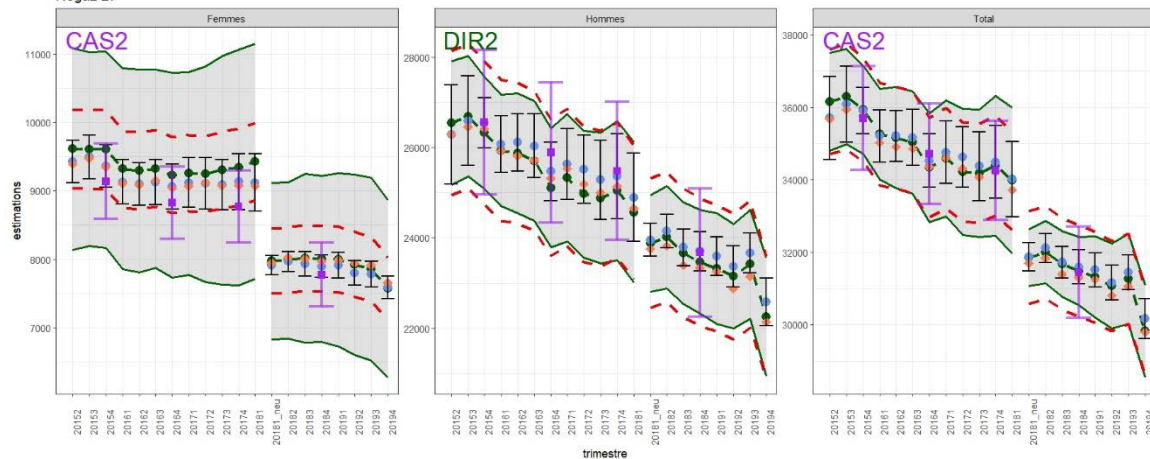
Noga2 25



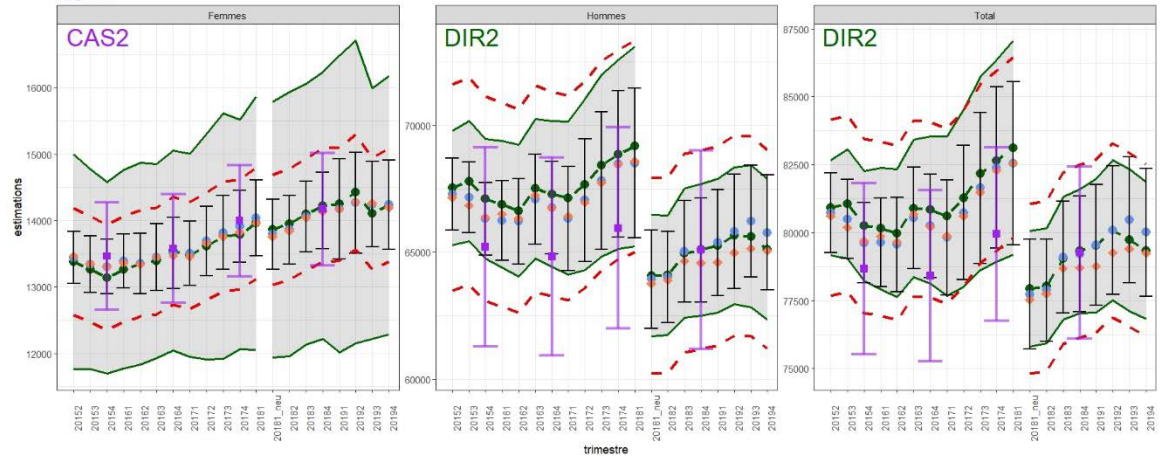
Noga2 26



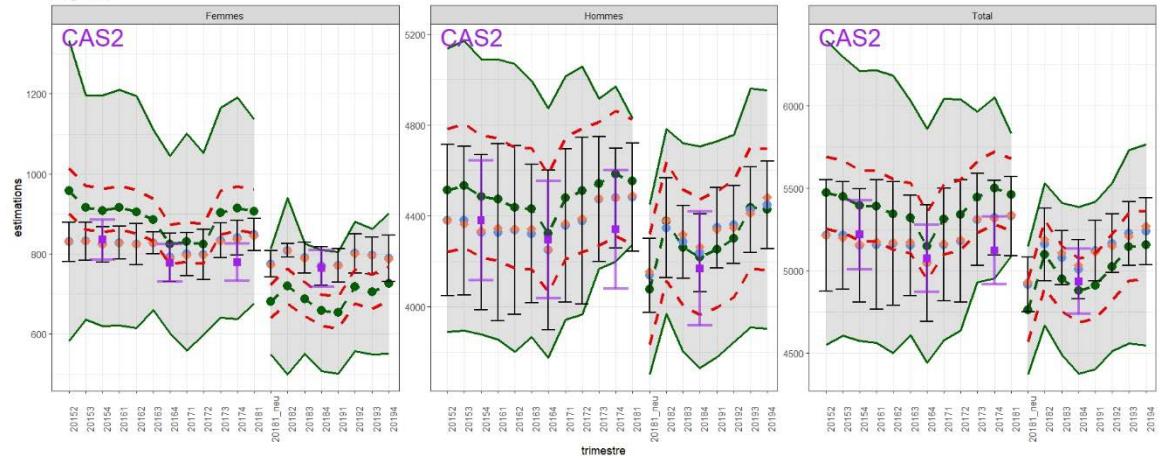
Noga2 27



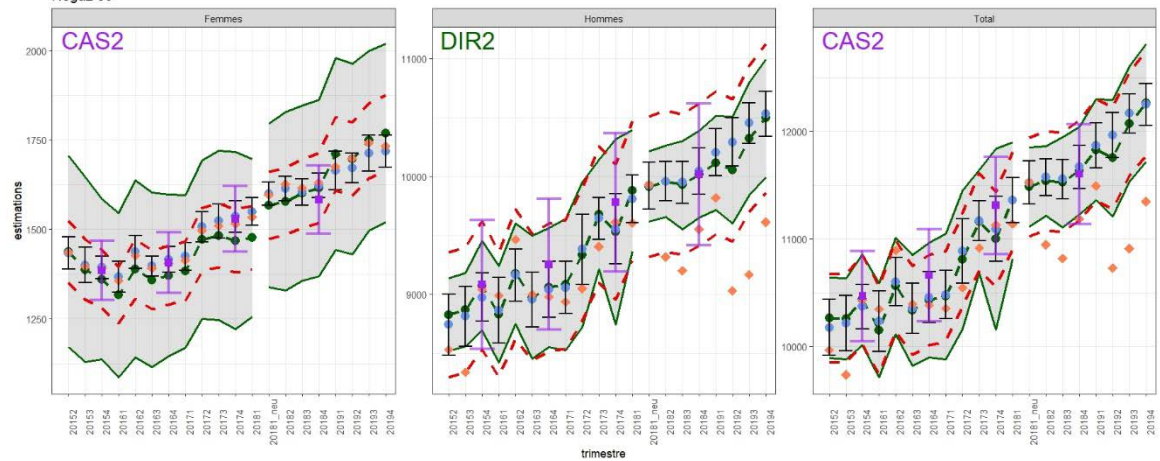
Noga2 28



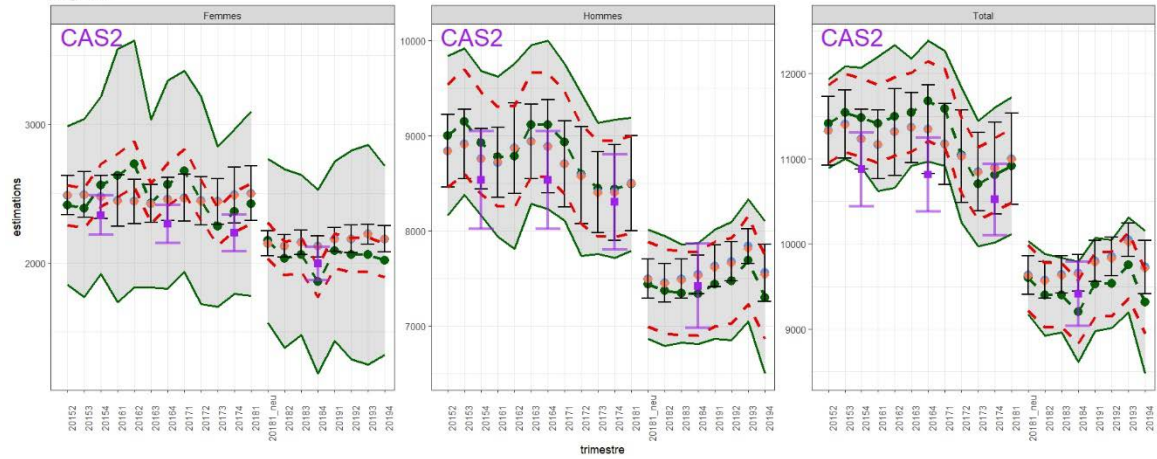
Noga2 29



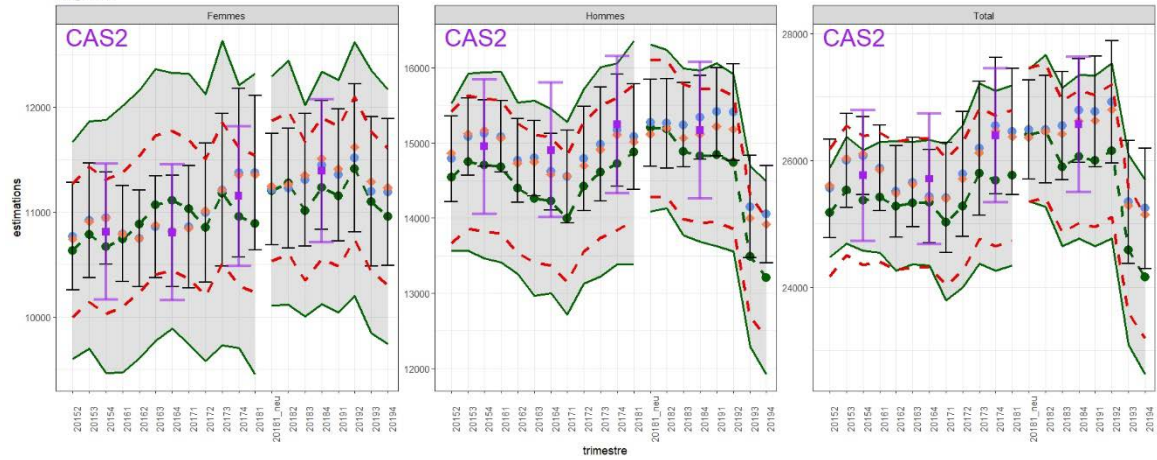
Noga2 30



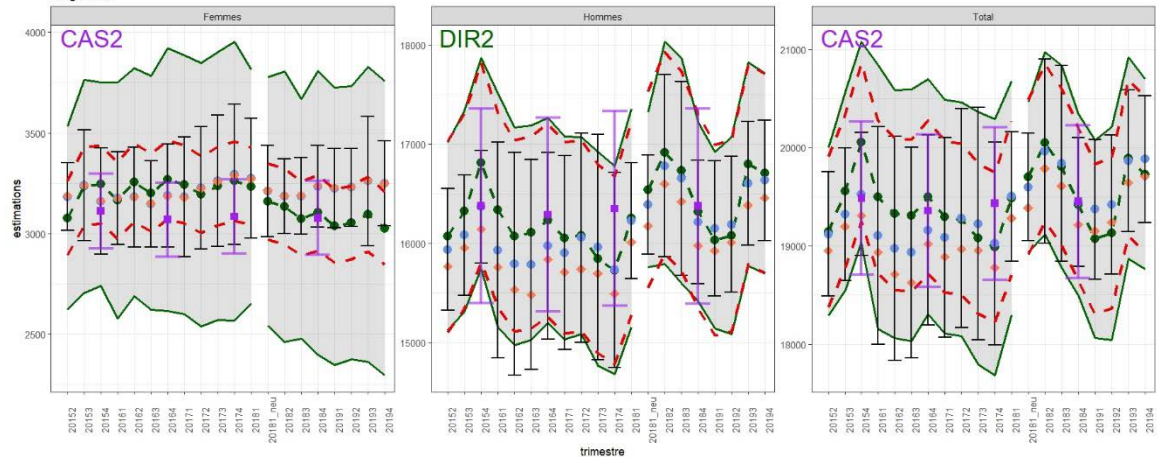
Noga2 31



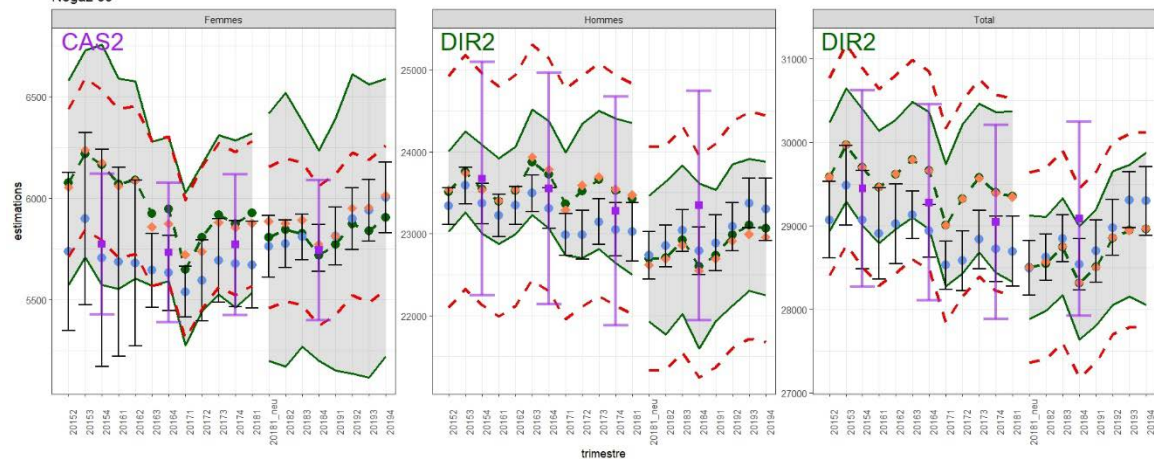
Noga2 32



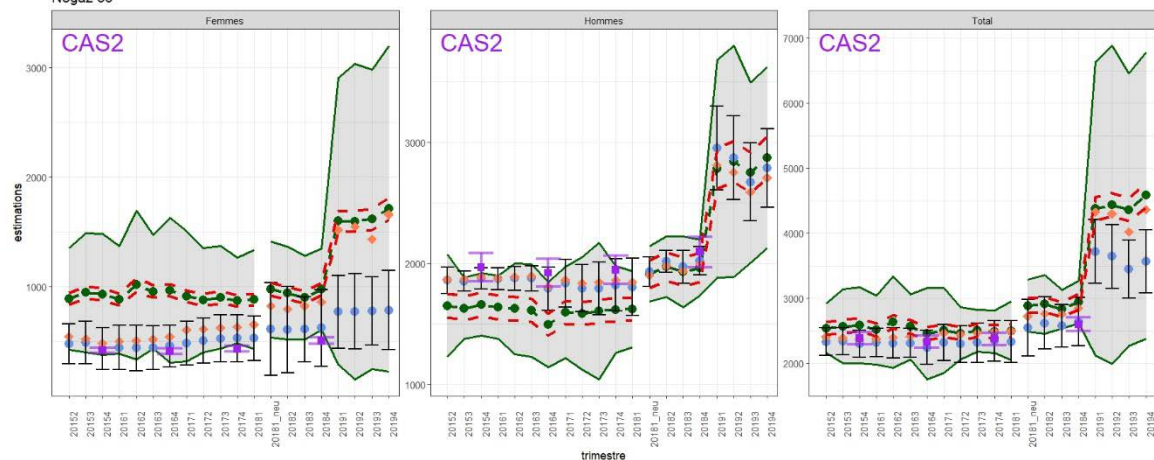
Noga2 33



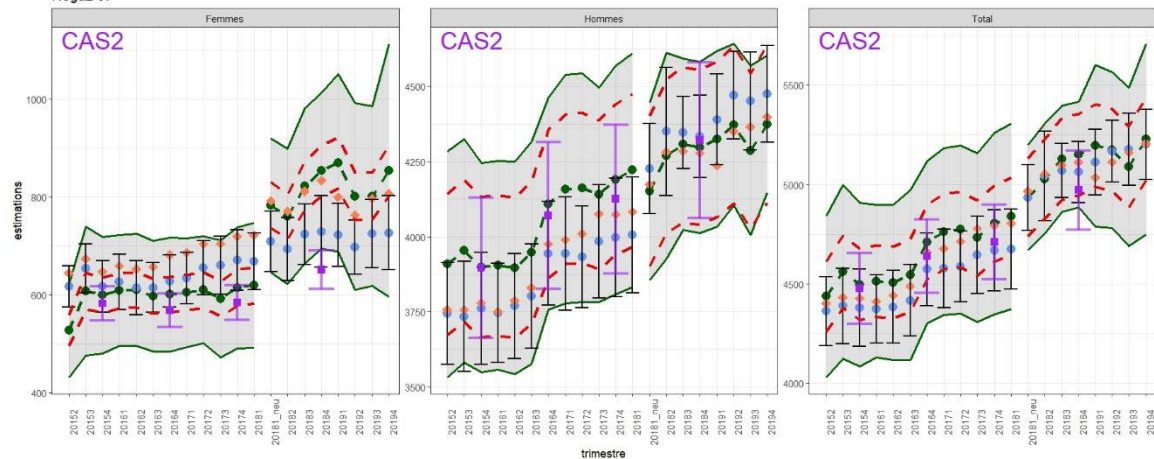
Noga2 35



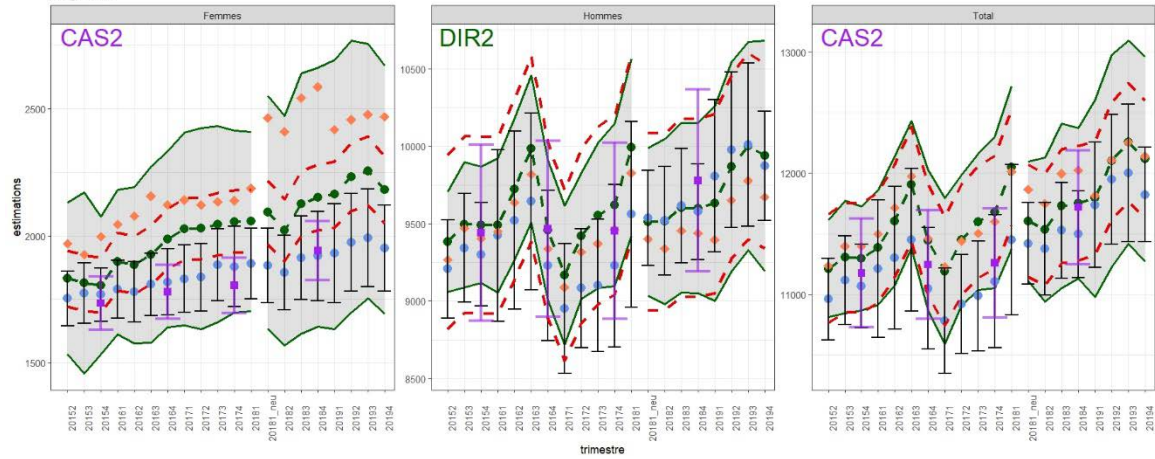
Noga2 36



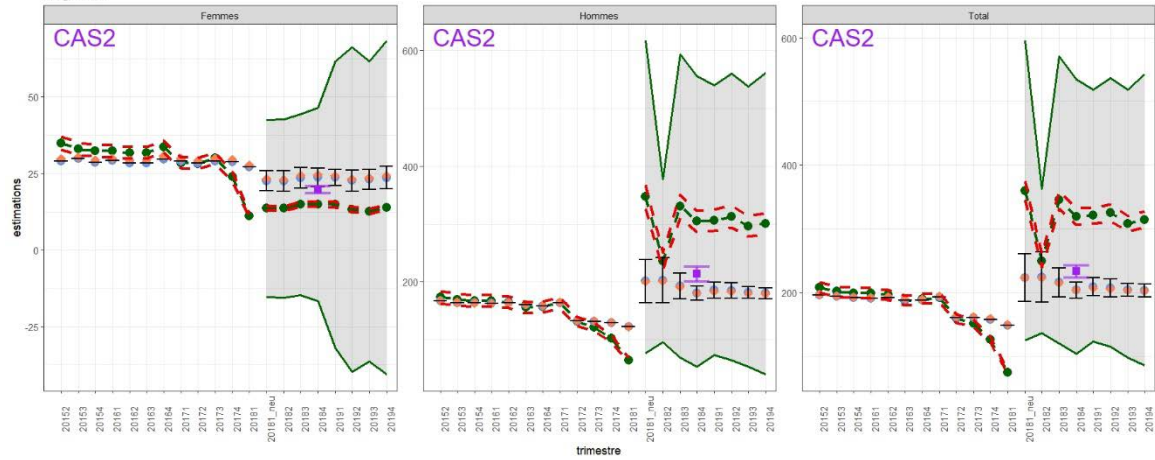
Noga2 37



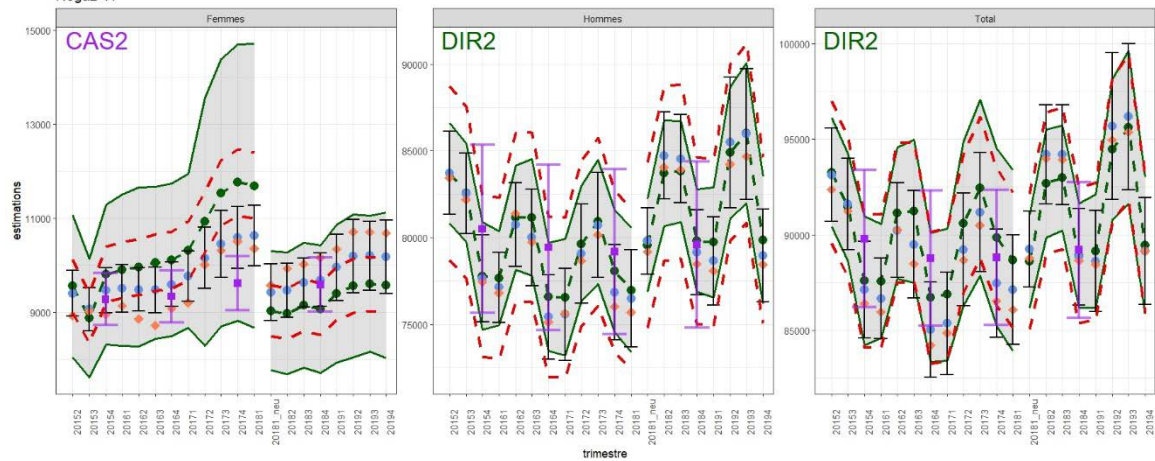
Noga2 38



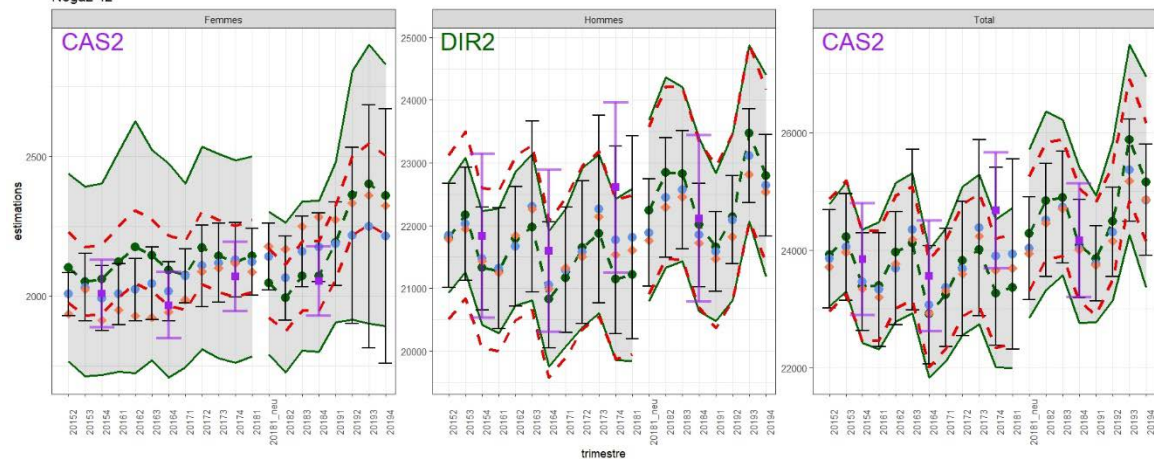
Noga2 39



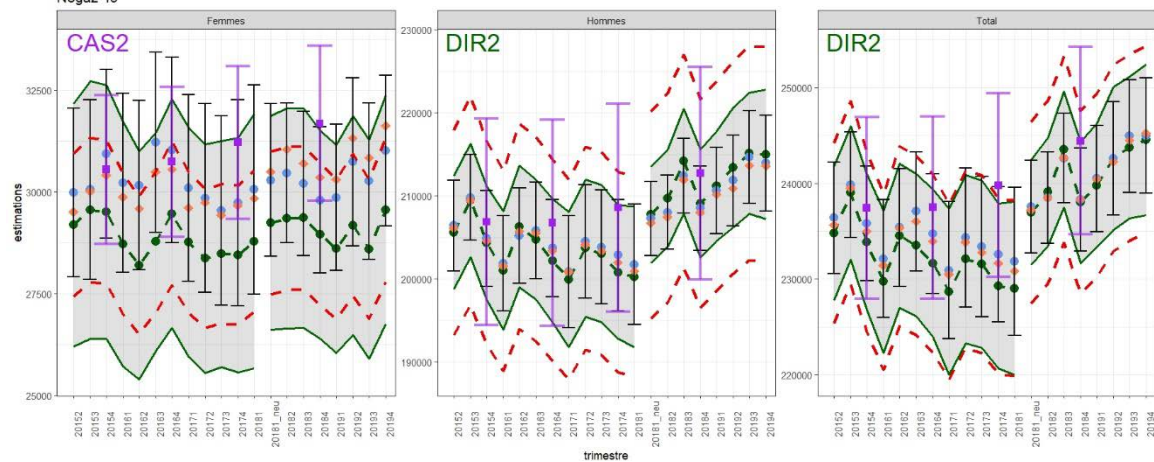
Noga2 41



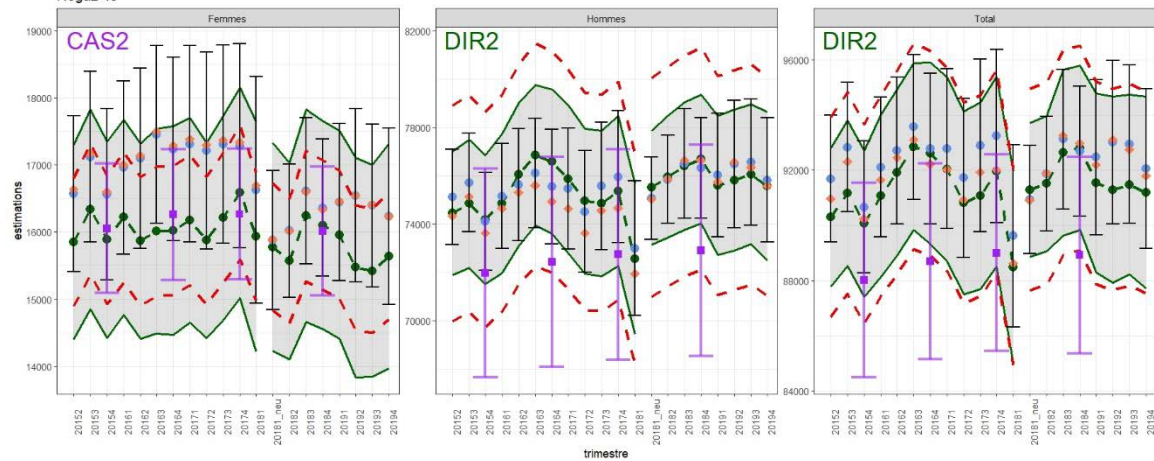
Noga2 42

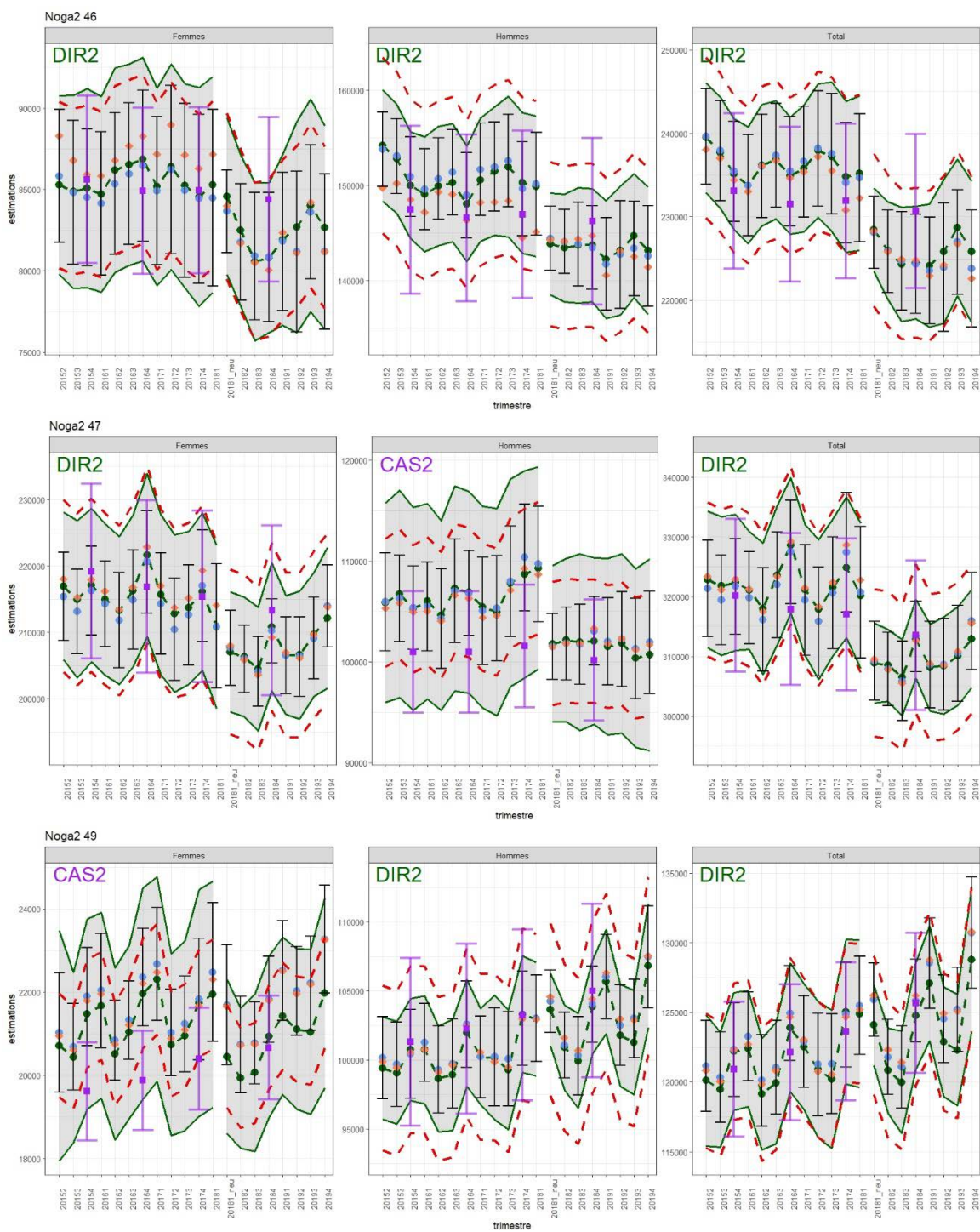


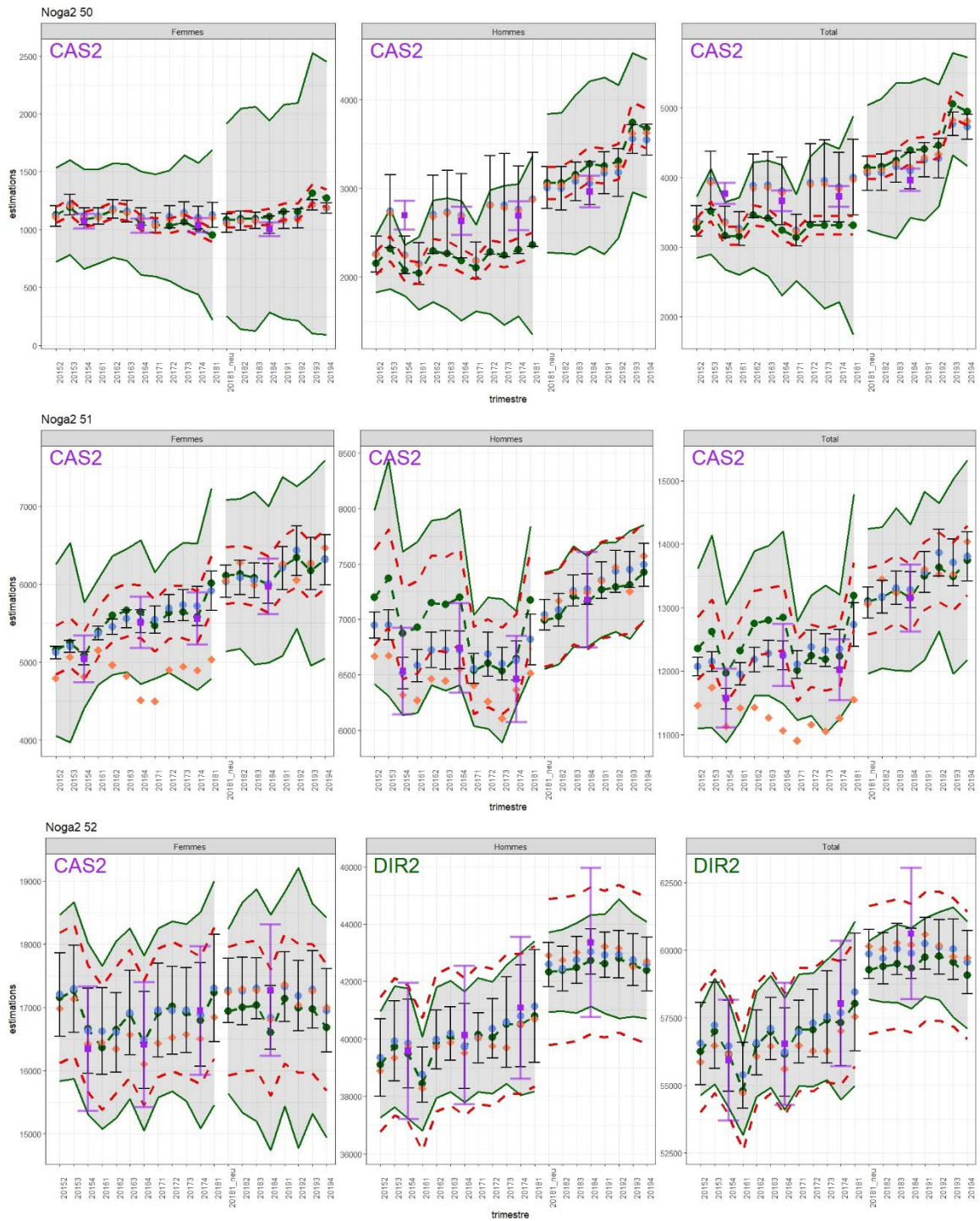
Noga2 43

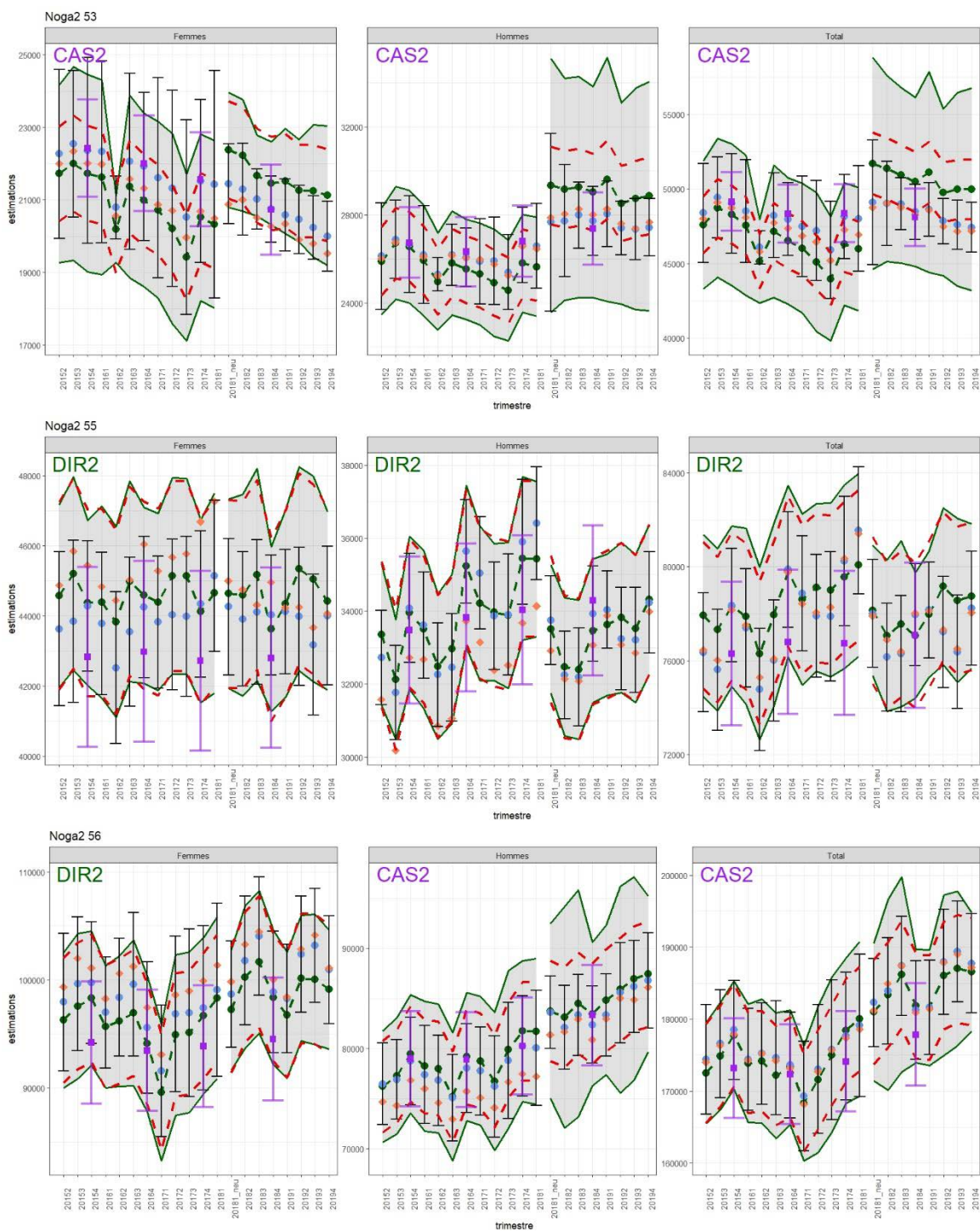


Noga2 45

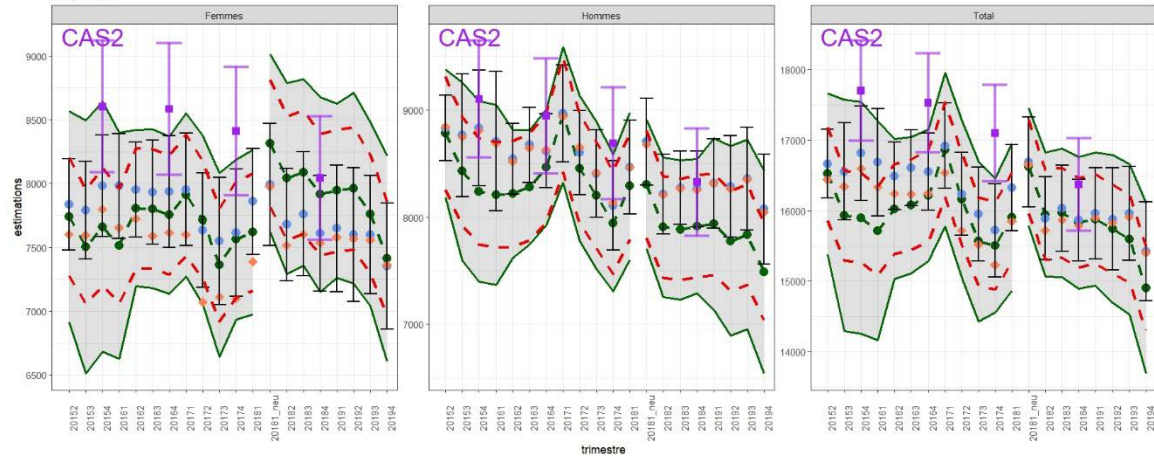




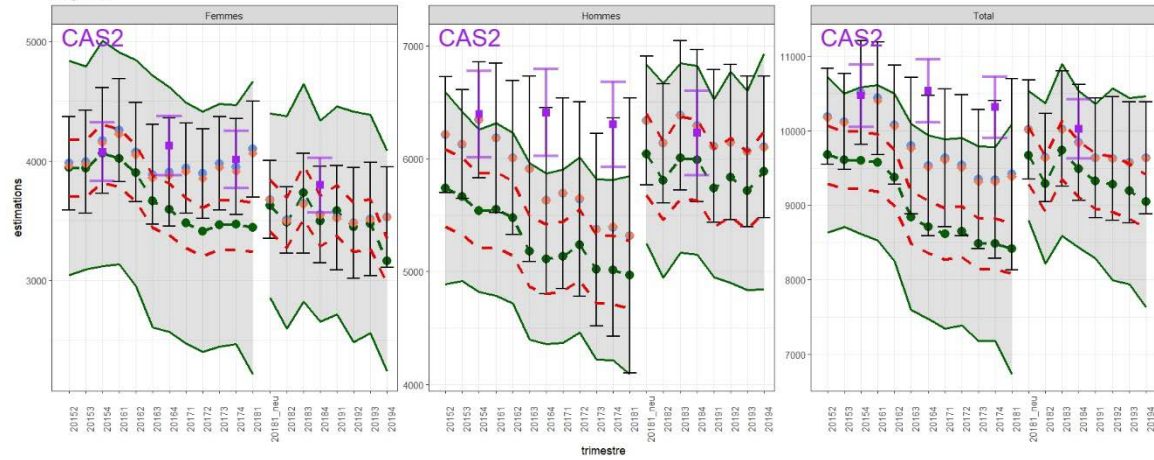




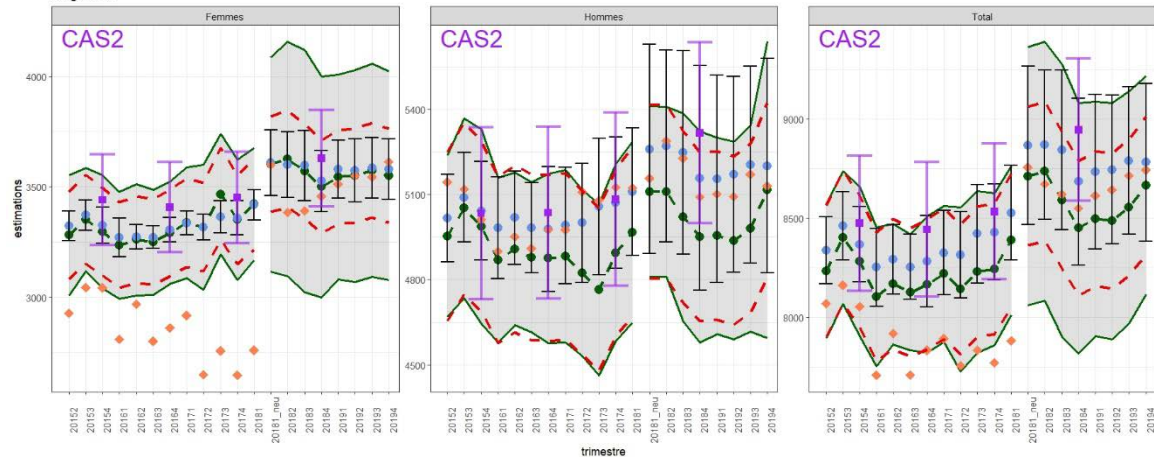
Noga2 58

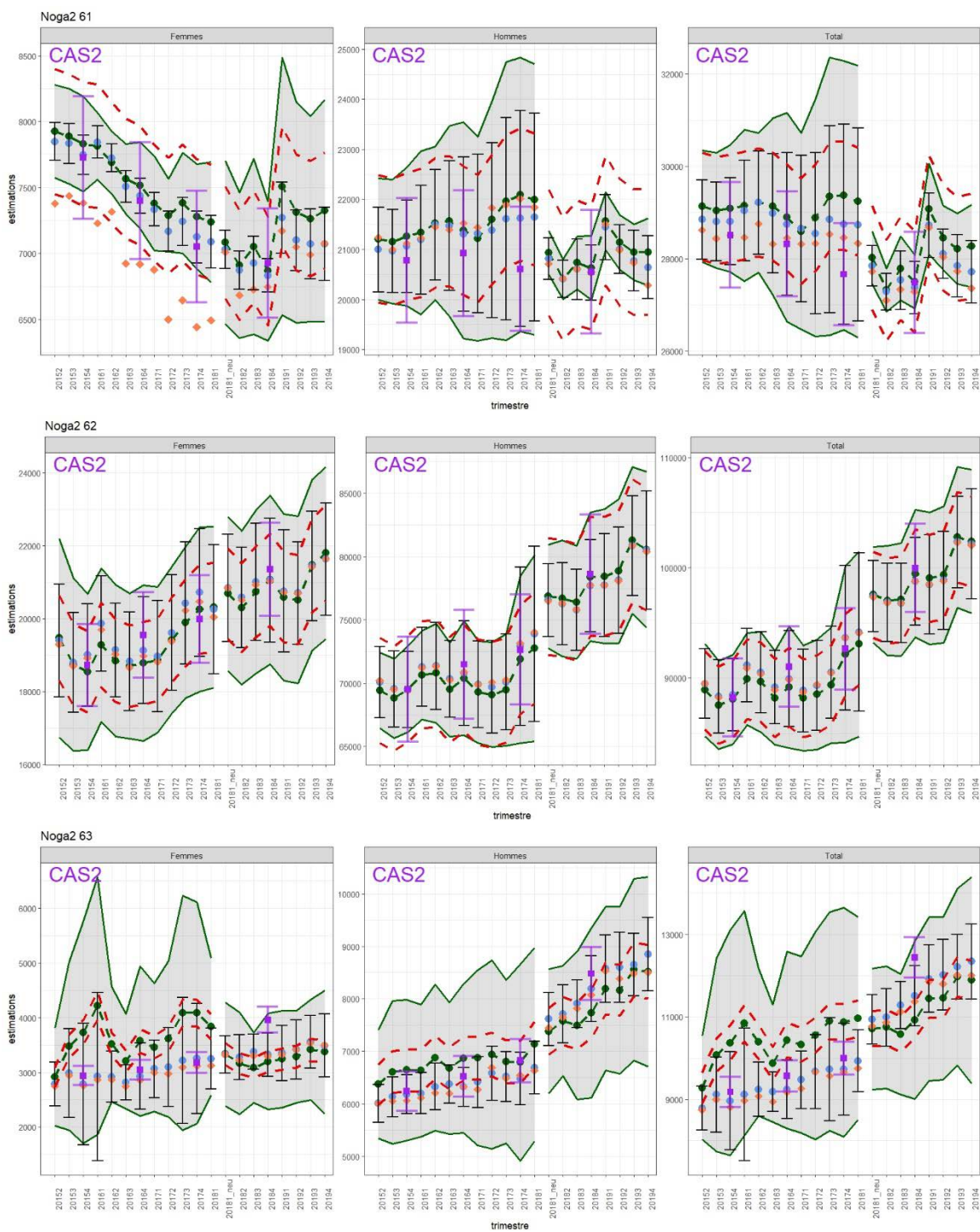


Noga2 59

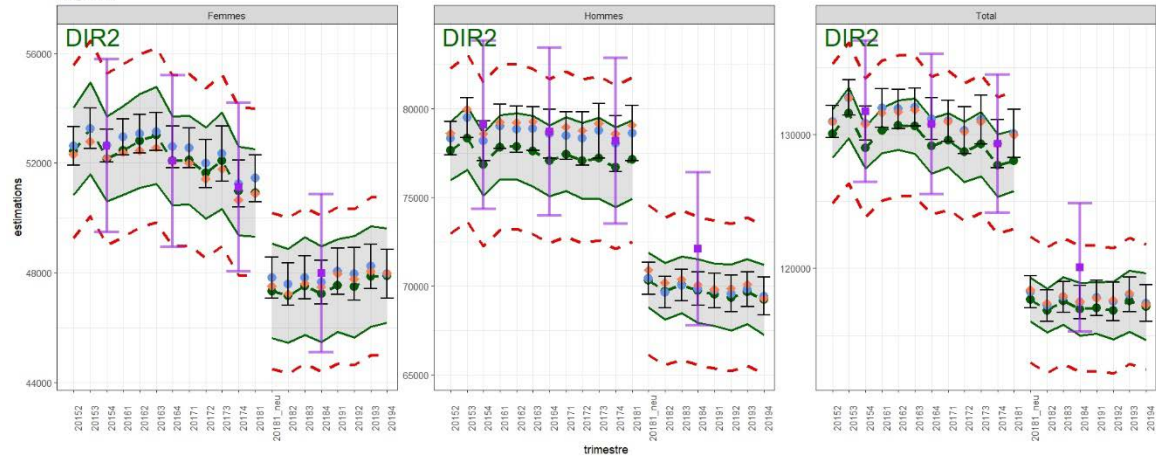


Noga2 60

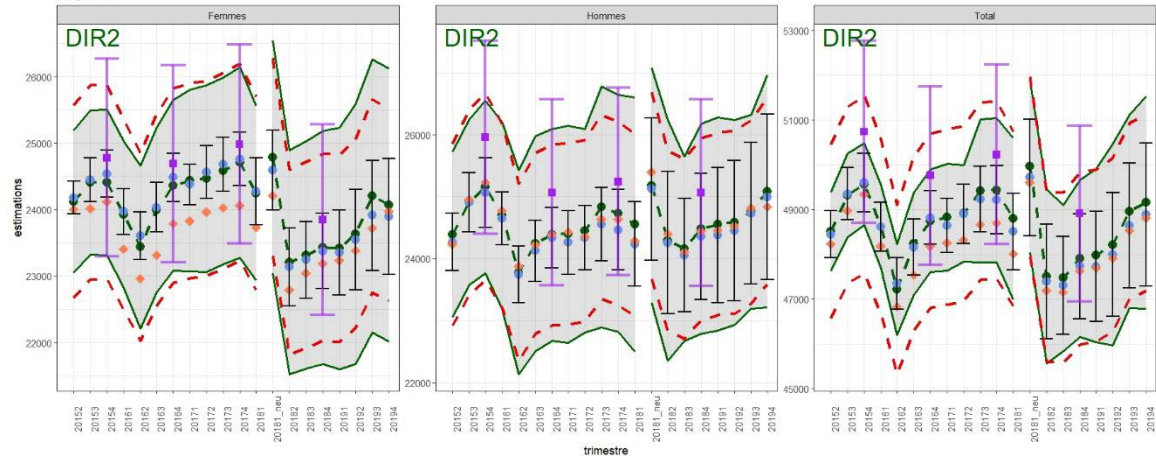




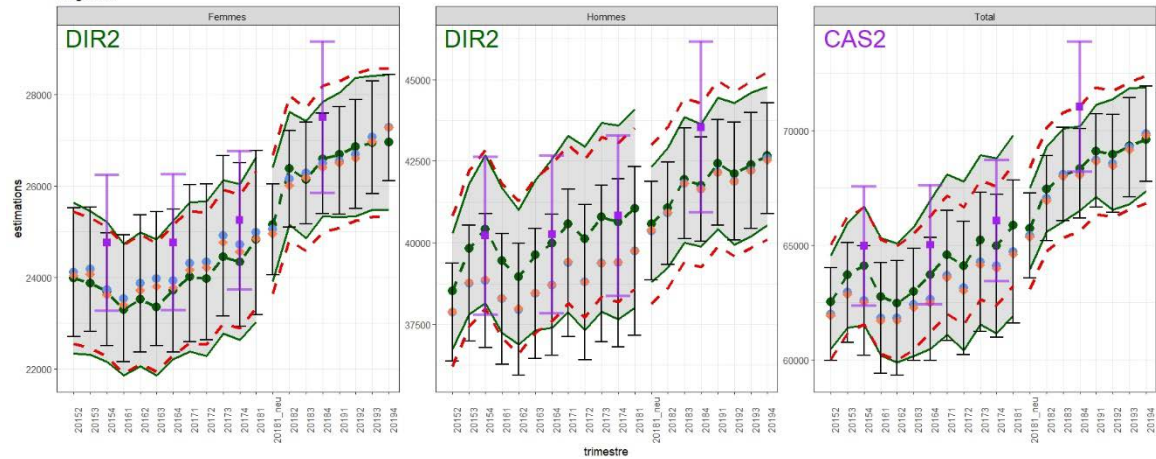
Noga2 64



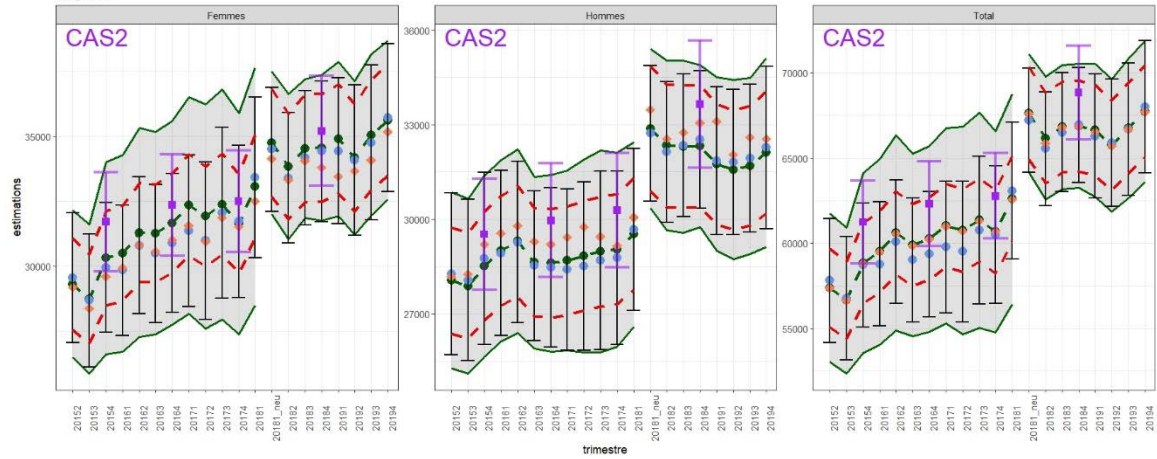
Noga2 65



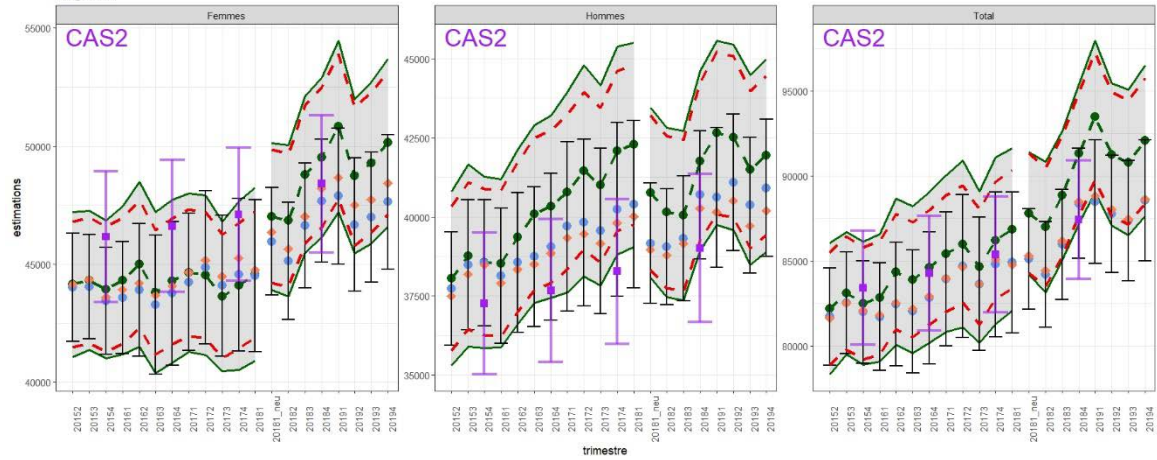
Noga2 66



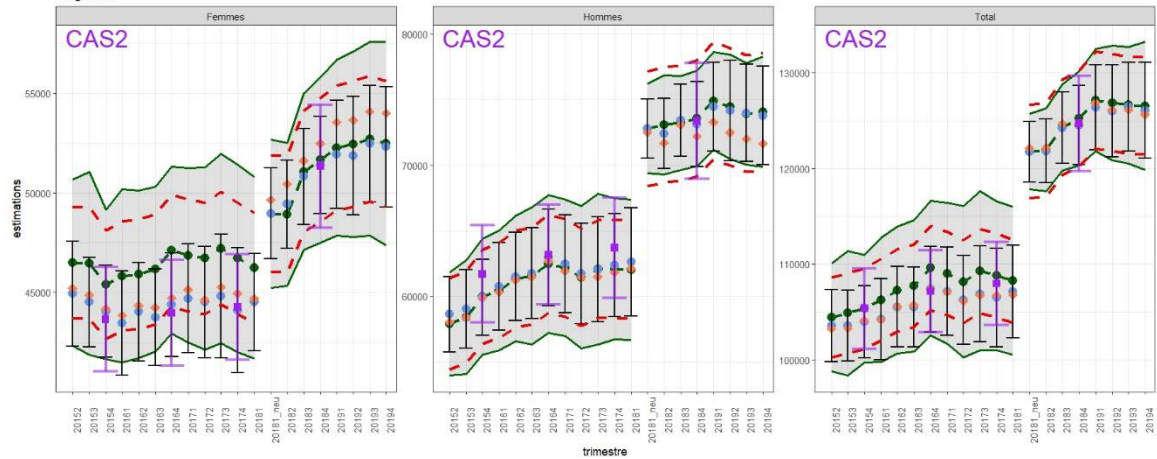
Noga2 68



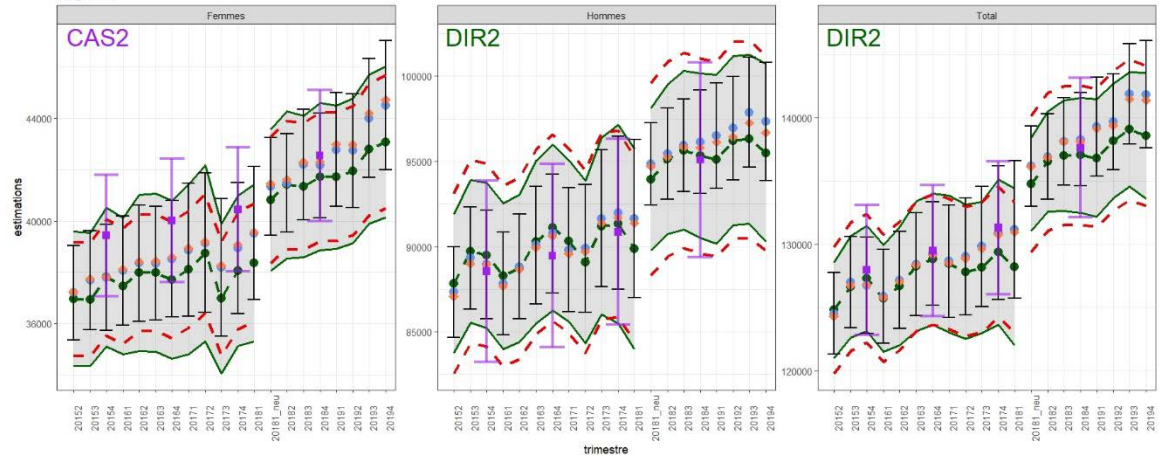
Noga2 69



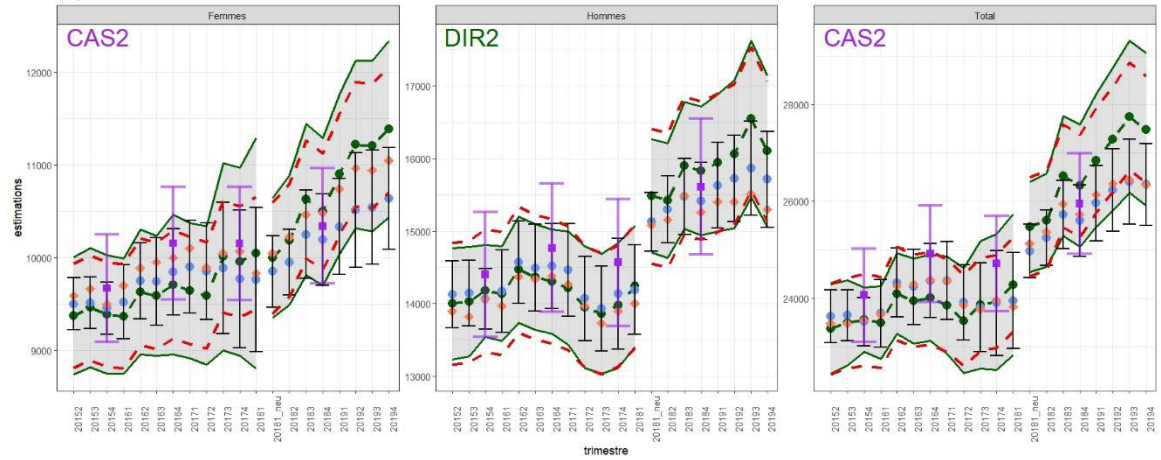
Noga2 70



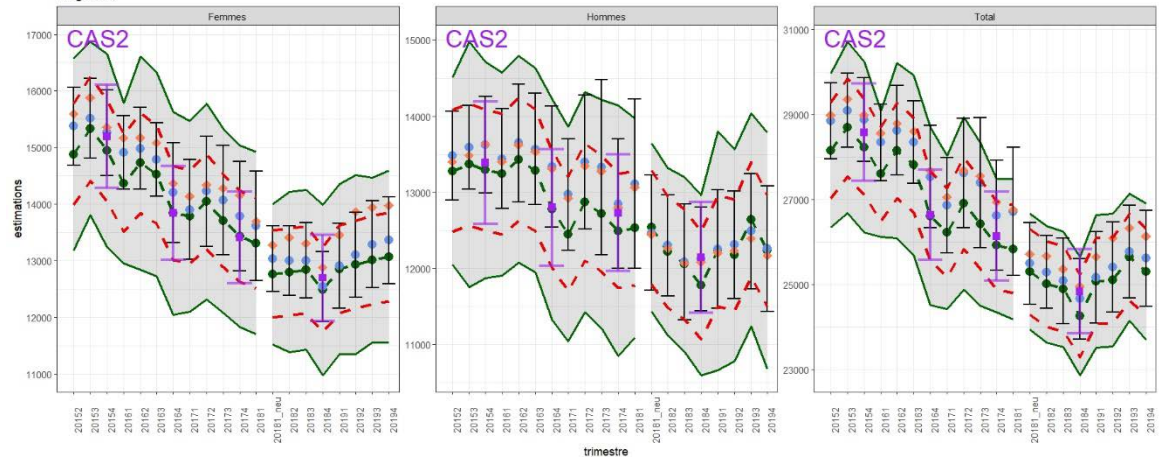
Noga2 71

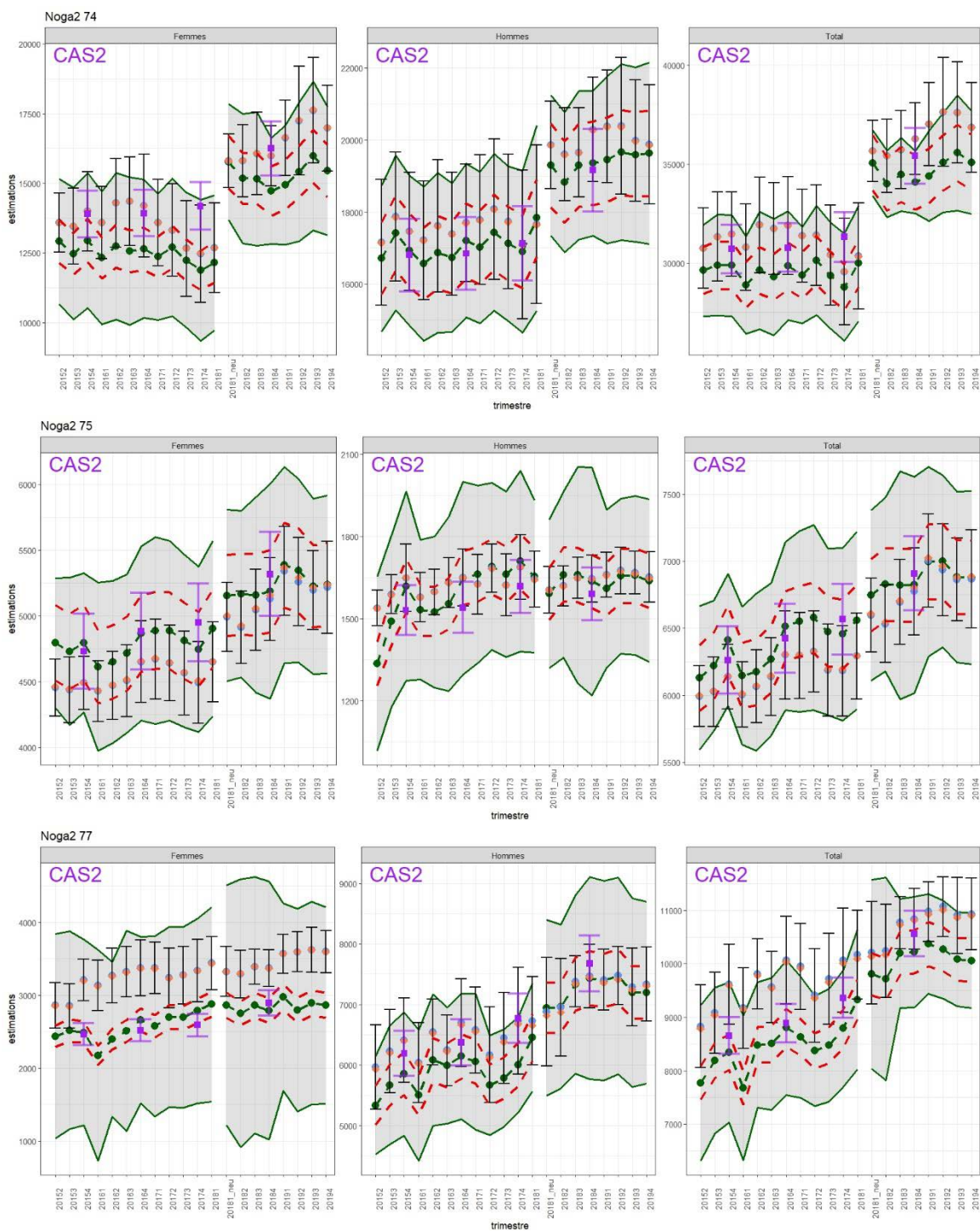


Noga2 72

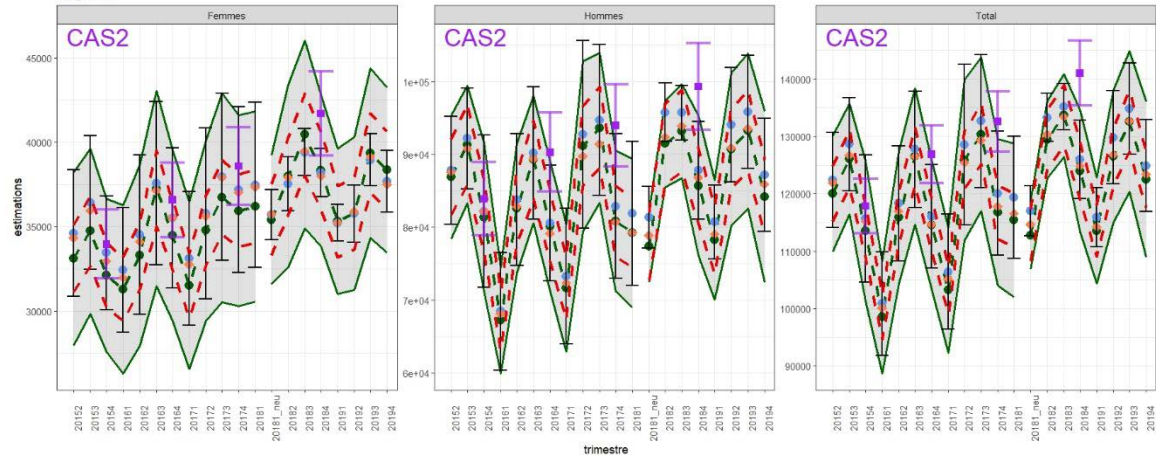


Noga2 73

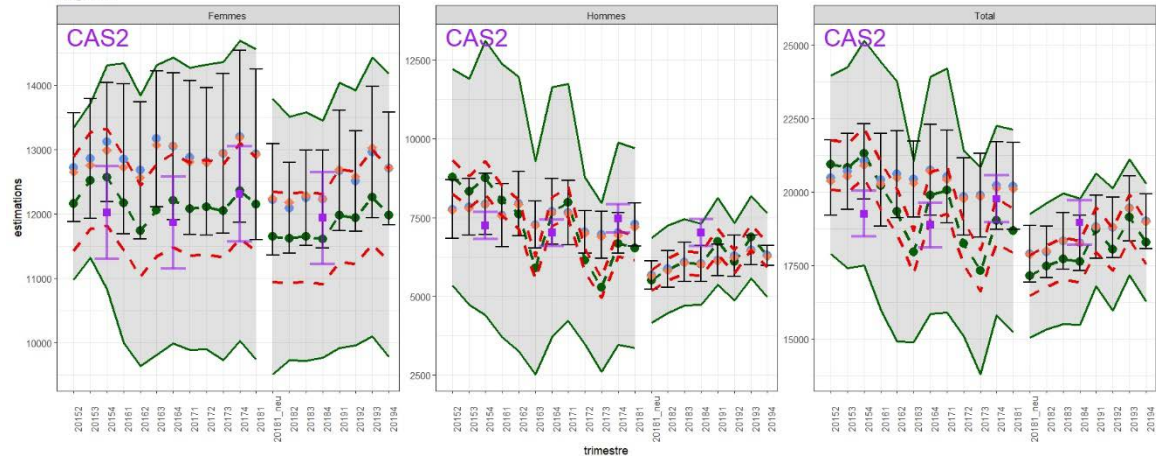




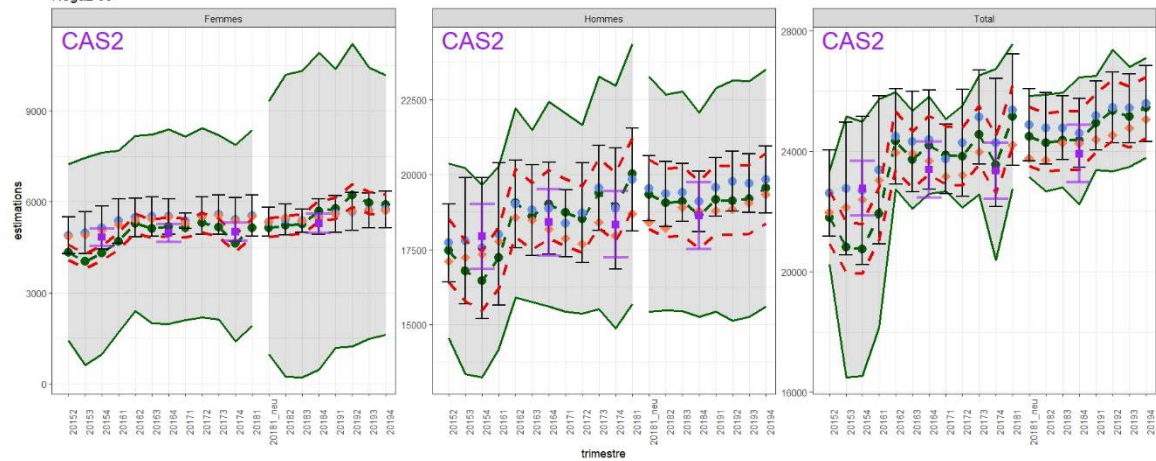
Noga2 78



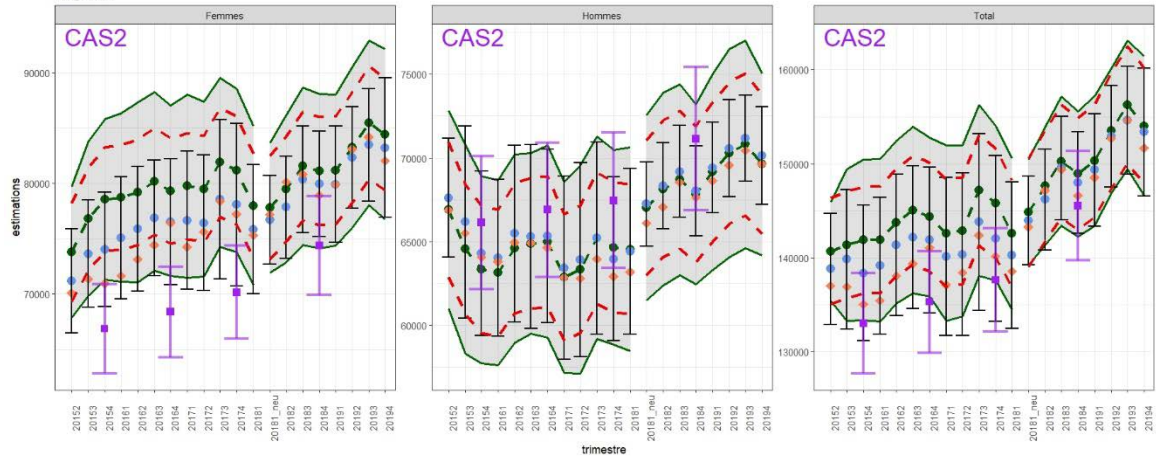
Noga2 79



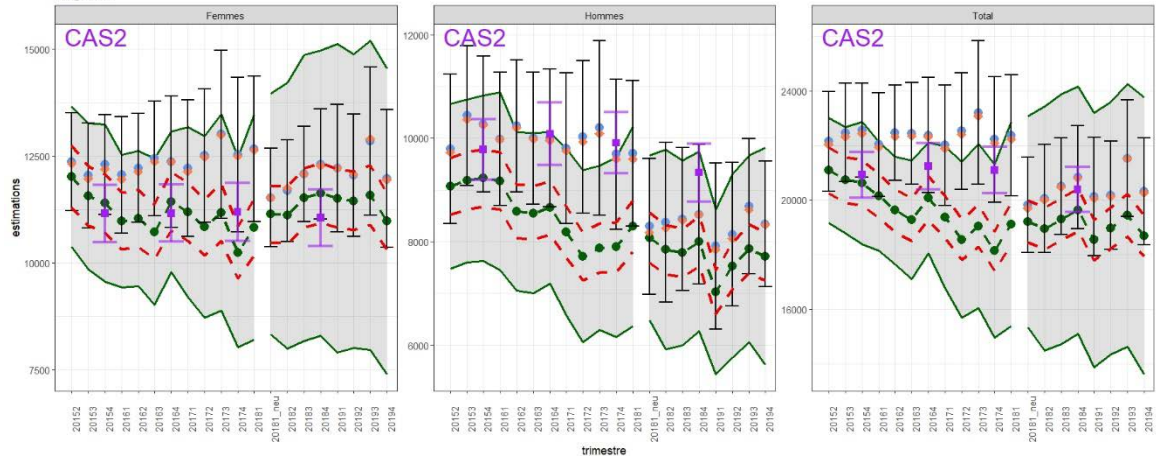
Noga2 80



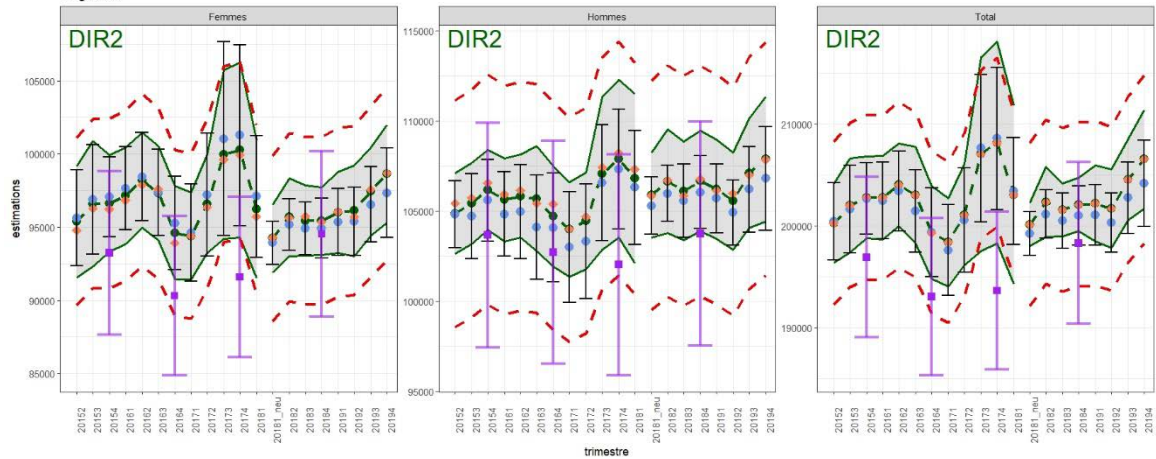
Noga2 81



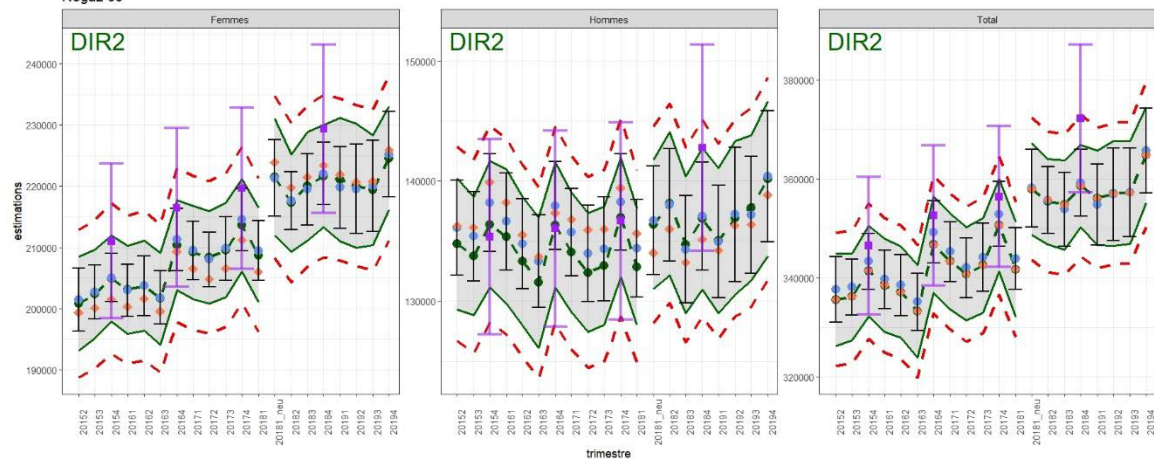
Noga2 82



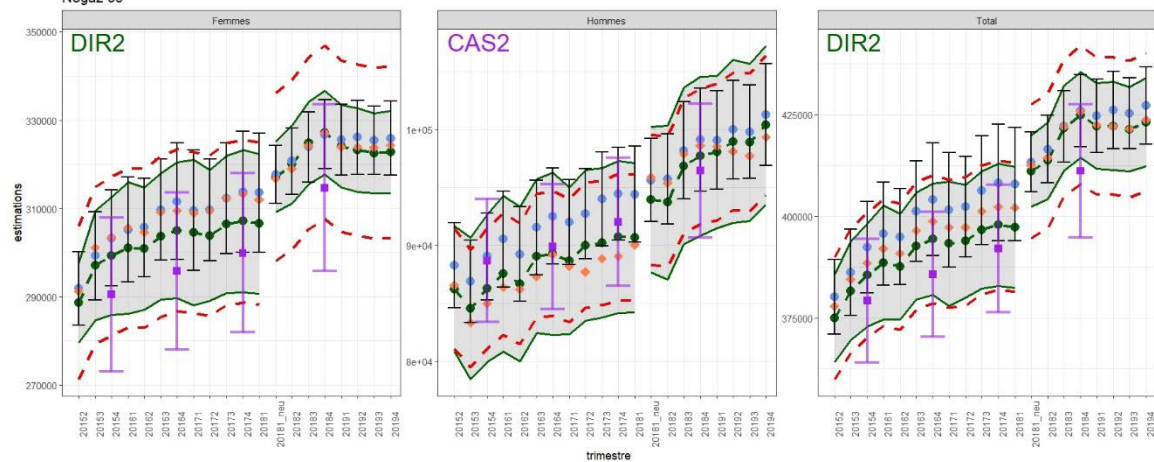
Noga2 84



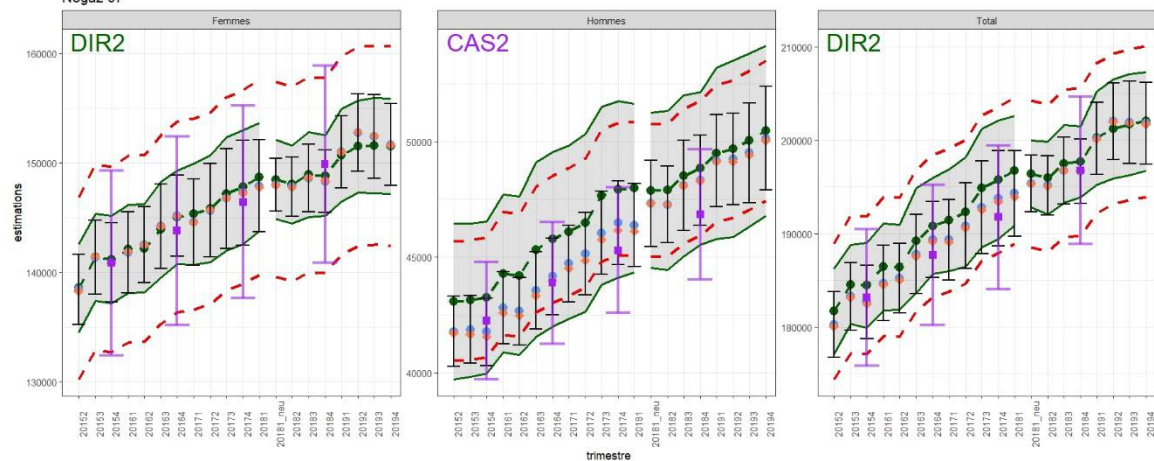
Noga2 85



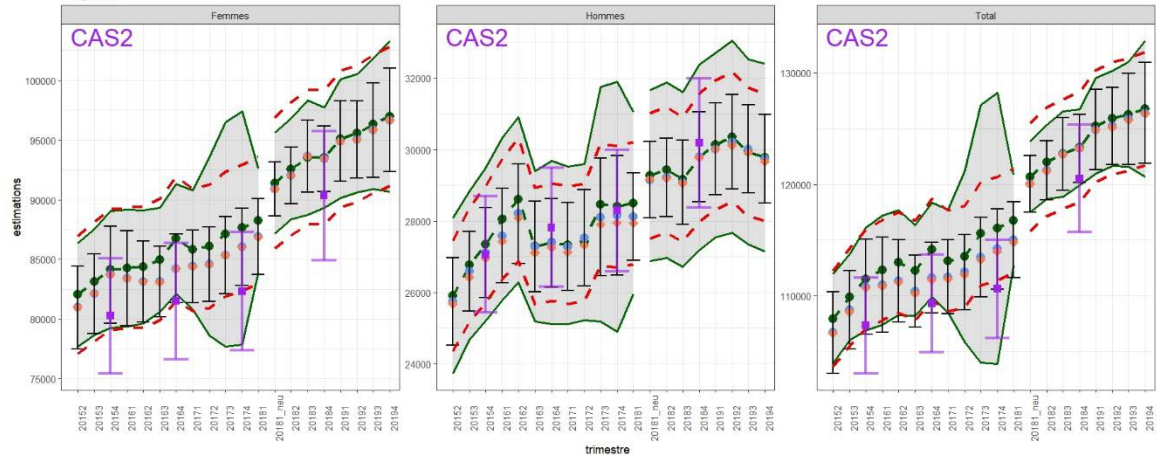
Noga2 86



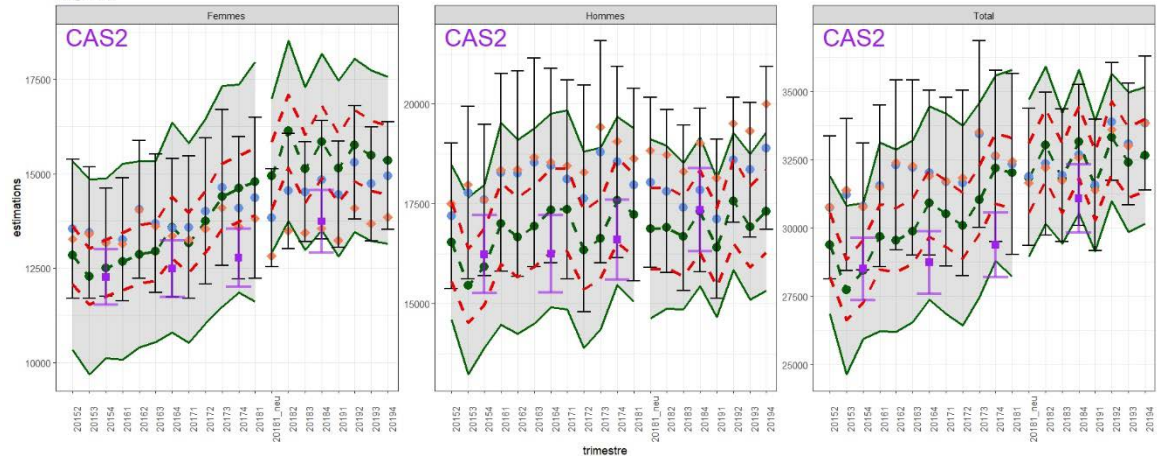
Noga2 87



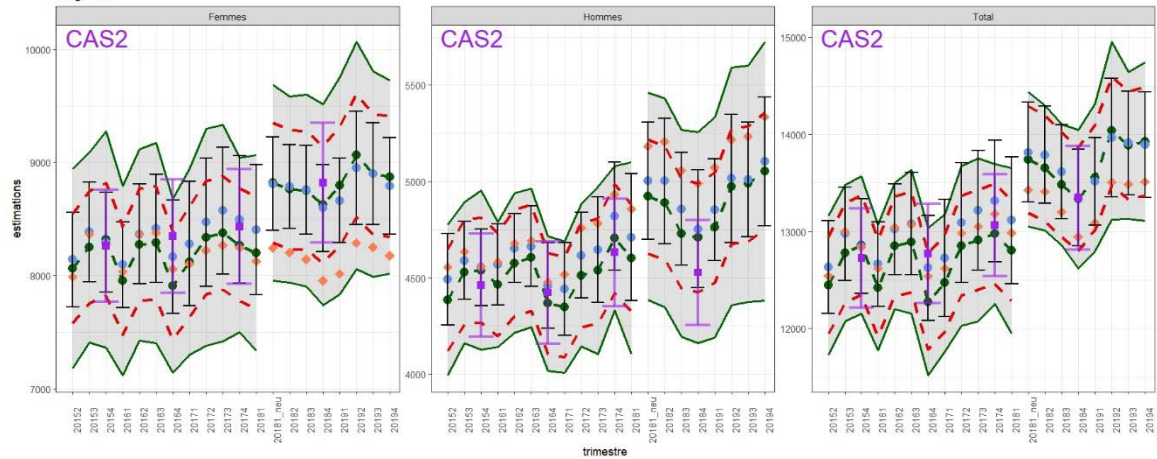
Noga2 88



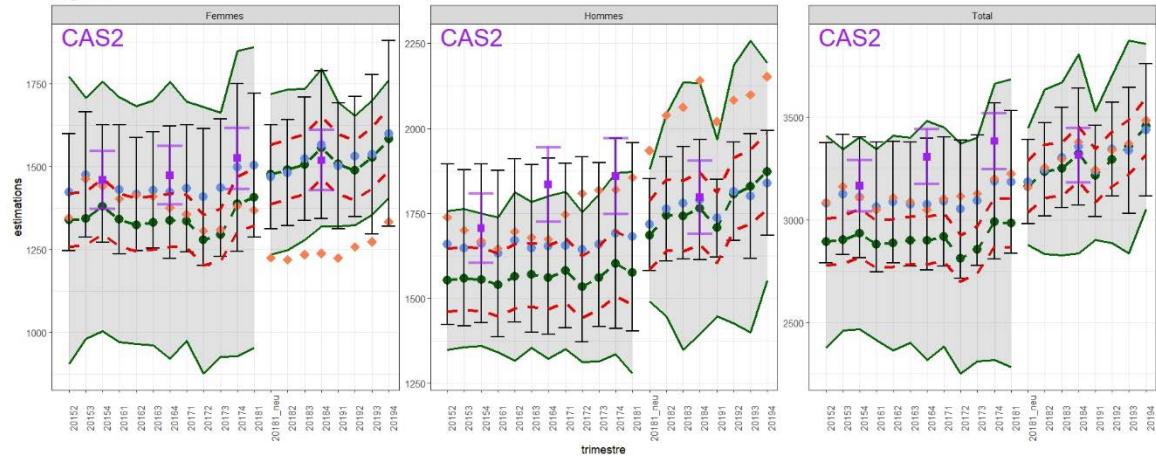
Noga2 90



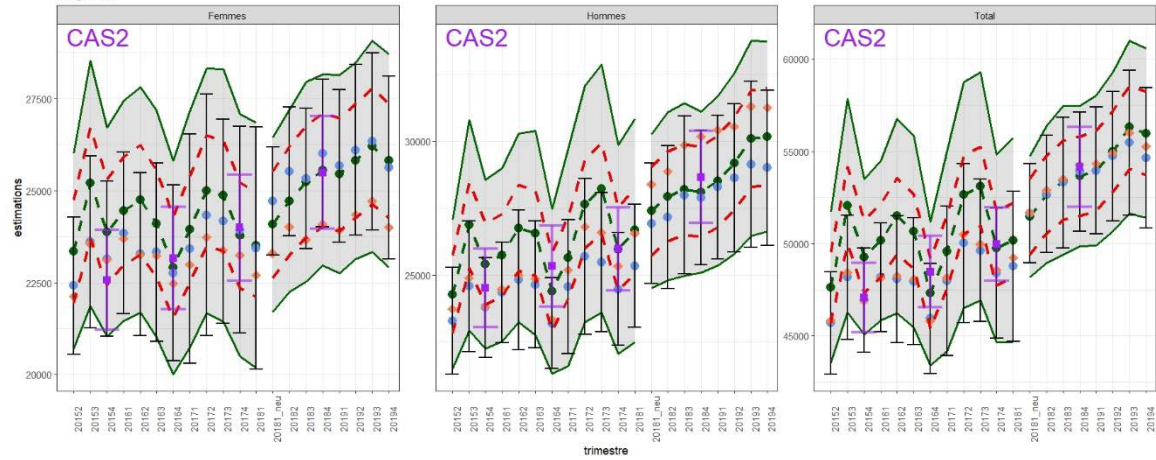
Noga2 91



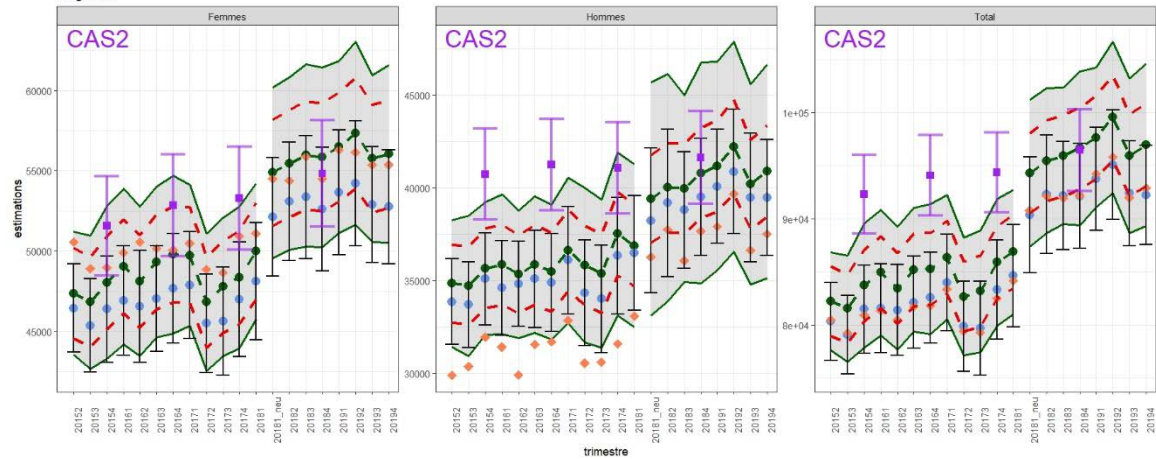
Noga2 92



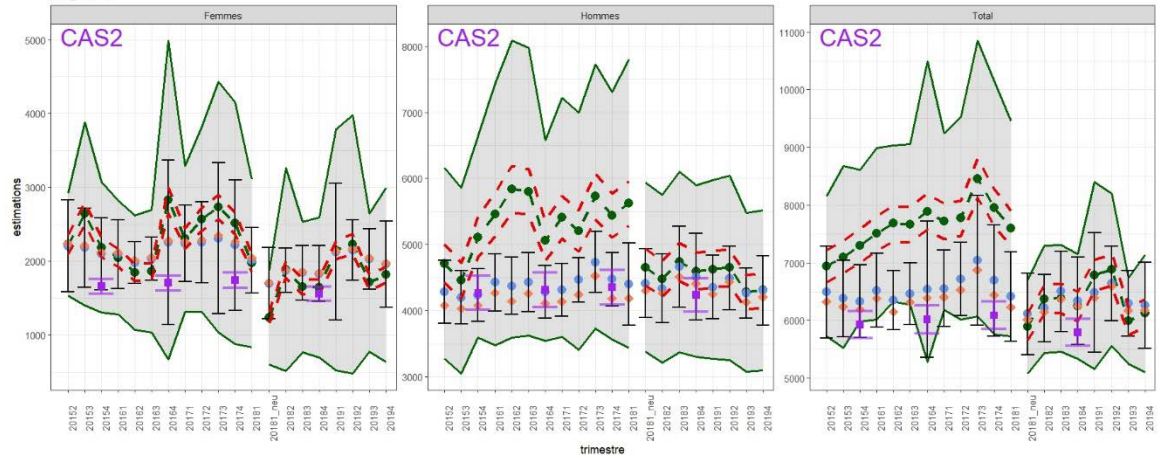
Noga2 93



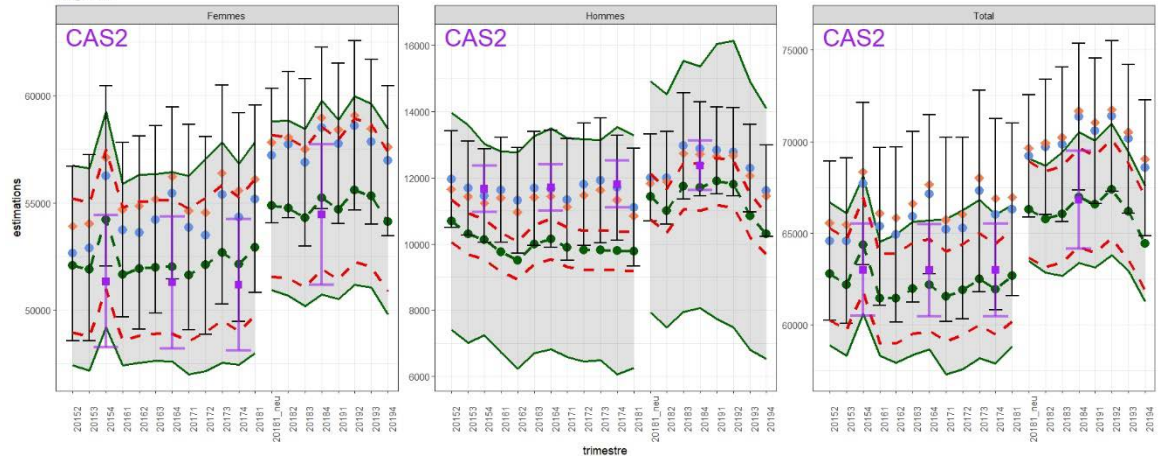
Noga2 94



Noga2 95



Noga2 96



## Annexe 7 : résumé des résultats de la validation au niveau des NOGA2

La table ci-dessous donne le nombre et le pourcentage de trimestres validés, ainsi que l'écart relatif (en valeur absolue) moyen entre les estimations SAE (sans ou avec **benchmarking**) et les valeurs de **référence pour la validation** (DIR2 ou CAS2 selon le cas).

| NOGA2 | Sexe   | Valid | # val_tri | % vali_tri | # val_tri_bm | % vali_tri_bm | EAM (%) | EAM_bm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|------------|
| 8     | Femmes | CAS2  | 4         | 100%       | 0            | 0%            | 1.66%   | 12.13%     |
| 8     | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.75%   | 1.79%      |
| 8     | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.62%   | 0.99%      |
| 9     | Femmes | CAS2  | 0         | 0%         | 3            | 75%           | 20.35%  | 6.13%      |
| 9     | Hommes | CAS2  | 1         | 25%        | 1            | 25%           | 24.41%  | 25.65%     |
| 9     | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 22.89%  | 17.69%     |
| 10    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.06%   | 0.86%      |
| 10    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.81%   | 0.56%      |
| 10    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.70%   | 0.52%      |
| 11    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 3.36%   | 3.34%      |
| 11    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 3.96%   | 4.18%      |
| 11    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.32%   | 3.37%      |
| 12    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.87%   | 1.05%      |
| 12    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 0.18%   | 1.91%      |
| 12    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.56%   | 1.18%      |
| 13    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.85%   | 1.65%      |
| 13    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 4.88%   | 4.94%      |
| 13    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.13%   | 3.05%      |
| 14    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.79%   | 2.65%      |
| 14    | Hommes | CAS2  | 0         | 0%         | 0            | 0%            | 13.40%  | 13.02%     |
| 14    | Total  | CAS2  | 4         | 100%       | 3            | 75%           | 3.14%   | 4.77%      |
| 15    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 0.78%   | 0.91%      |
| 15    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.54%   | 2.20%      |
| 15    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.24%   | 1.41%      |
| 16    | Femmes | CAS2  | 1         | 25%        | 1            | 25%           | 8.77%   | 8.75%      |
| 16    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.09%   | 0.98%      |
| 16    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 2.00%   | 1.89%      |
| 17    | Femmes | CAS2  | 1         | 25%        | 1            | 25%           | 6.74%   | 6.44%      |
| 17    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 4.38%   | 4.33%      |
| 17    | Total  | CAS2  | 3         | 75%        | 3            | 75%           | 5.07%   | 4.94%      |
| 18    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 2.89%   | 2.80%      |
| 18    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.91%   | 2.80%      |
| 18    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.90%   | 2.80%      |
| 19    | Femmes | CAS2  | 2         | 50%        | 2            | 50%           | 6.64%   | 6.46%      |
| 19    | Hommes | CAS2  | 1         | 25%        | 1            | 25%           | 22.27%  | 23.10%     |
| 19    | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 18.96%  | 19.58%     |
| 20    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.64%   | 2.11%      |
| 20    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.51%   | 1.82%      |
| 20    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.76%   | 1.08%      |

| NOGA2 | Sexe   | Valid | # val tri | % vali tri | # val tri bm | % vali tri bm | EAM (%) | EAM_bm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|------------|
| 21    | Femmes | DIR2  | 20        | 100%       | 14           | 70%           | 1.33%   | 4.29%      |
| 21    | Hommes | DIR2  | 20        | 100%       | 6            | 30%           | 1.29%   | 5.64%      |
| 21    | Total  | DIR2  | 20        | 100%       | 10           | 50%           | 1.31%   | 4.08%      |
| 22    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.38%   | 1.51%      |
| 22    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.59%   | 0.65%      |
| 22    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.56%   | 0.56%      |
| 23    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 3.75%   | 3.53%      |
| 23    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.52%   | 1.47%      |
| 23    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.05%   | 2.01%      |
| 24    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 3.83%   | 3.80%      |
| 24    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 2.50%   | 1.96%      |
| 24    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.71%   | 1.94%      |
| 25    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 3.26%   | 3.10%      |
| 25    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.95%   | 1.07%      |
| 25    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.71%   | 0.75%      |
| 26    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.77%   | 2.17%      |
| 26    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.39%   | 1.18%      |
| 26    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.05%   | 3.46%      |
| 27    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 2.70%   | 2.71%      |
| 27    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.05%   | 0.62%      |
| 27    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.57%   | 0.44%      |
| 28    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 0.60%   | 0.87%      |
| 28    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.64%   | 0.74%      |
| 28    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.49%   | 0.61%      |
| 29    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 2.93%   | 2.69%      |
| 29    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.76%   | 1.85%      |
| 29    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.80%   | 1.87%      |
| 30    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.03%   | 1.09%      |
| 30    | Hommes | DIR2  | 20        | 100%       | 13           | 65%           | 0.55%   | 3.92%      |
| 30    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.35%   | 2.05%      |
| 31    | Femmes | CAS2  | 1         | 25%        | 1            | 25%           | 7.97%   | 7.86%      |
| 31    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.39%   | 2.35%      |
| 31    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.57%   | 3.52%      |
| 32    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 0.95%   | 1.05%      |
| 32    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.18%   | 1.21%      |
| 32    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.92%   | 0.75%      |
| 33    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 4.34%   | 4.34%      |
| 33    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.03%   | 2.13%      |
| 33    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.83%   | 1.81%      |

| NOGA2 | Sexe   | Valid | # val tri | % vali tri | # val tri bm | % vali tri bm | EAM (%) | EAM_bbm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|-------------|
| 35    | Femmes | CAS2  | 4         | 100%       | 3            | 75%           | 1.19%   | 2.84%       |
| 35    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.11%   | 0.21%       |
| 35    | Total  | DIR2  | 19        | 95%        | 20           | 100%          | 1.51%   | 0.01%       |
| 36    | Femmes | CAS2  | 1         | 25%        | 0            | 0%            | 14.66%  | 40.77%      |
| 36    | Hommes | CAS2  | 2         | 50%        | 4            | 100%          | 5.39%   | 4.80%       |
| 36    | Total  | CAS2  | 4         | 100%       | 3            | 75%           | 2.49%   | 3.99%       |
| 37    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 10.83%  | 20.58%      |
| 37    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.48%   | 1.88%       |
| 37    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.59%   | 1.53%       |
| 38    | Femmes | CAS2  | 4         | 100%       | 0            | 0%            | 2.29%   | 21.52%      |
| 38    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.87%   | 1.48%       |
| 38    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.49%   | 2.39%       |
| 39    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 20.26%  | 24.56%      |
| 39    | Hommes | CAS2  | 0         | 0%         | 0            | 0%            | 15.48%  | 15.78%      |
| 39    | Total  | CAS2  | 0         | 0%         | 0            | 0%            | 12.46%  | 12.38%      |
| 41    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 3.96%   | 5.34%       |
| 41    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.77%   | 1.14%       |
| 41    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 1.11%   | 1.46%       |
| 42    | Femmes | CAS2  | 4         | 100%       | 3            | 75%           | 3.03%   | 4.86%       |
| 42    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.06%   | 1.21%       |
| 42    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.86%   | 2.40%       |
| 43    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 3.22%   | 2.54%       |
| 43    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.49%   | 0.49%       |
| 43    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.67%   | 0.50%       |
| 45    | Femmes | CAS2  | 3         | 75%        | 2            | 50%           | 4.39%   | 4.57%       |
| 45    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.64%   | 0.73%       |
| 45    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 1.01%   | 0.59%       |
| 46    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.64%   | 1.54%       |
| 46    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.42%   | 1.41%       |
| 46    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.26%   | 0.49%       |
| 47    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.44%   | 0.51%       |
| 47    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 5.42%   | 4.99%       |
| 47    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.38%   | 0.28%       |
| 49    | Femmes | CAS2  | 1         | 25%        | 1            | 25%           | 9.17%   | 8.64%       |
| 49    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.51%   | 0.54%       |
| 49    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.87%   | 0.84%       |
| 50    | Femmes | CAS2  | 2         | 50%        | 3            | 75%           | 5.53%   | 4.18%       |
| 50    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 6.17%   | 6.54%       |
| 50    | Total  | CAS2  | 3         | 75%        | 3            | 75%           | 5.42%   | 5.49%       |

| NOGA2 | Sexe   | Valid | # val tri | % vali tri | # val tri bm | % vali tri bm | EAM (%) | EAM_bm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|------------|
| 51    | Femmes | CAS2  | 4         | 100%       | 2            | 50%           | 0.98%   | 8.66%      |
| 51    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.11%   | 2.12%      |
| 51    | Total  | CAS2  | 4         | 100%       | 2            | 50%           | 0.97%   | 5.09%      |
| 52    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.26%   | 1.92%      |
| 52    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.51%   | 0.77%      |
| 52    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.57%   | 0.95%      |
| 53    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 0.23%   | 2.77%      |
| 53    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 0.78%   | 1.09%      |
| 53    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.54%   | 1.38%      |
| 55    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 1.79%   | 1.96%      |
| 55    | Hommes | DIR2  | 20        | 100%       | 19           | 95%           | 1.08%   | 3.19%      |
| 55    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 1.29%   | 1.12%      |
| 56    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 1.78%   | 3.24%      |
| 56    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 0.64%   | 3.29%      |
| 56    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.93%   | 1.71%      |
| 58    | Femmes | CAS2  | 1         | 25%        | 0            | 0%            | 7.40%   | 10.62%     |
| 58    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 3.51%   | 3.50%      |
| 58    | Total  | CAS2  | 3         | 75%        | 1            | 25%           | 5.41%   | 6.99%      |
| 59    | Femmes | CAS2  | 3         | 75%        | 2            | 50%           | 3.92%   | 4.25%      |
| 59    | Hommes | CAS2  | 2         | 50%        | 2            | 50%           | 7.11%   | 7.07%      |
| 59    | Total  | CAS2  | 2         | 50%        | 2            | 50%           | 5.29%   | 5.40%      |
| 60    | Femmes | CAS2  | 4         | 100%       | 1            | 25%           | 2.96%   | 13.91%     |
| 60    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.13%   | 1.67%      |
| 60    | Total  | CAS2  | 4         | 100%       | 2            | 50%           | 1.82%   | 6.37%      |
| 61    | Femmes | CAS2  | 4         | 100%       | 2            | 50%           | 0.80%   | 5.55%      |
| 61    | Hommes | CAS2  | 4         | 100%       | 3            | 75%           | 2.04%   | 2.86%      |
| 61    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.72%   | 1.01%      |
| 62    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 2.17%   | 1.98%      |
| 62    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 0.62%   | 0.69%      |
| 62    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.94%   | 0.92%      |
| 63    | Femmes | CAS2  | 3         | 75%        | 1            | 25%           | 6.92%   | 7.70%      |
| 63    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 3.13%   | 3.70%      |
| 63    | Total  | CAS2  | 3         | 75%        | 3            | 75%           | 3.94%   | 4.97%      |
| 64    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.70%   | 0.41%      |
| 64    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.03%   | 1.47%      |
| 64    | Total  | DIR2  | 19        | 95%        | 19           | 95%           | 0.89%   | 0.85%      |
| 65    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.40%   | 1.74%      |
| 65    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.48%   | 0.41%      |
| 65    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.32%   | 0.93%      |

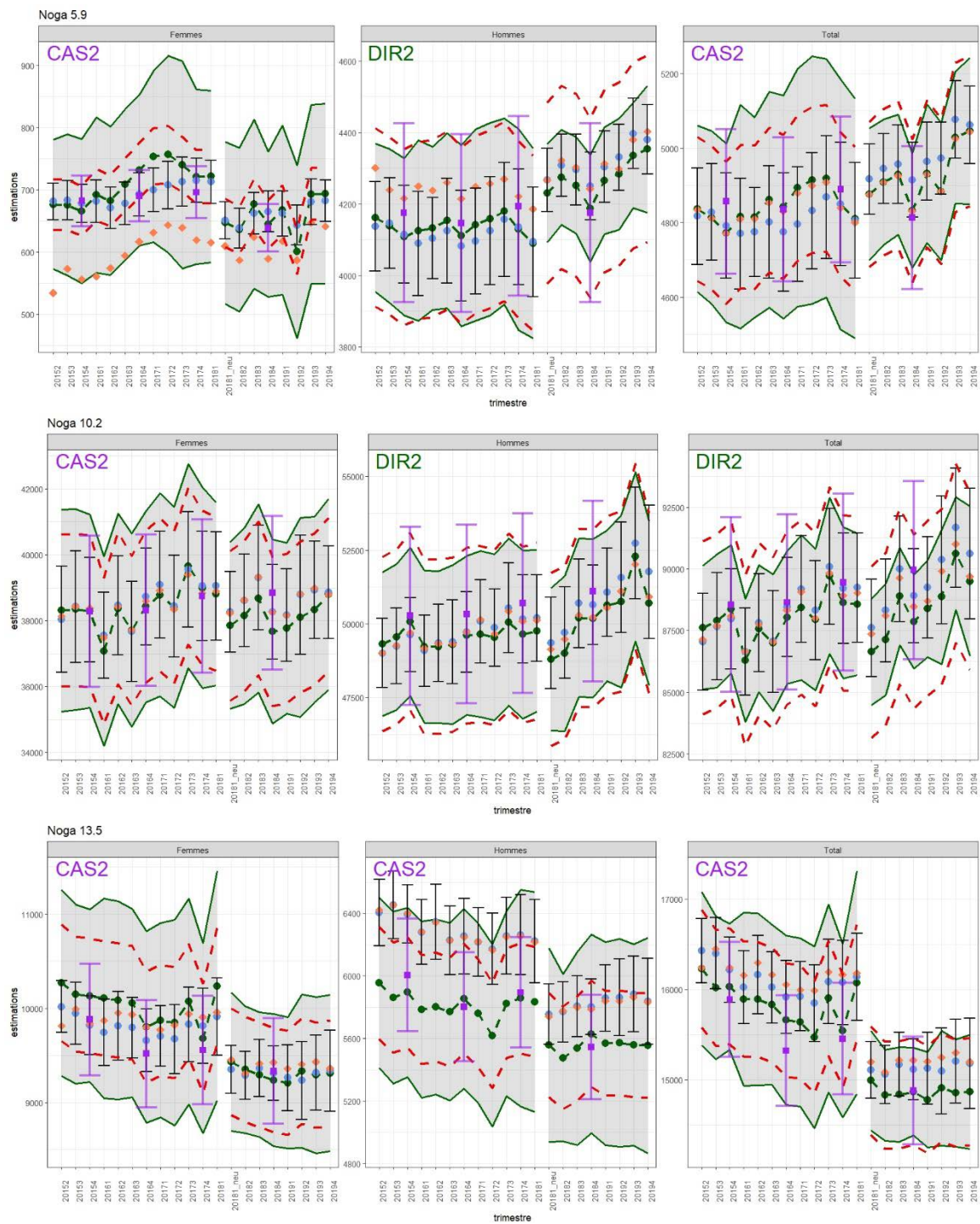
| NOGA2 | Sexe   | Valid | # val tri | % vali tri | # val tri bm | % vali tri bm | EAM (%) | EAM_bm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|------------|
| 66    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.99%   | 0.73%      |
| 66    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.96%   | 1.93%      |
| 66    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.60%   | 3.76%      |
| 68    | Femmes | CAS2  | 4         | 100%       | 3            | 75%           | 3.66%   | 4.42%      |
| 68    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 3.96%   | 2.26%      |
| 68    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.81%   | 3.38%      |
| 69    | Femmes | CAS2  | 3         | 75%        | 4            | 100%          | 4.73%   | 3.83%      |
| 69    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 4.14%   | 3.41%      |
| 69    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.31%   | 1.20%      |
| 70    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 0.56%   | 1.66%      |
| 70    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.40%   | 2.05%      |
| 70    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 0.76%   | 0.64%      |
| 71    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 3.14%   | 2.92%      |
| 71    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.78%   | 0.64%      |
| 71    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.91%   | 0.80%      |
| 72    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 2.61%   | 1.37%      |
| 72    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 1.47%   | 1.93%      |
| 72    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.29%   | 2.04%      |
| 73    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.77%   | 2.96%      |
| 73    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.77%   | 1.70%      |
| 73    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.76%   | 2.22%      |
| 74    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 4.08%   | 4.05%      |
| 74    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 3.77%   | 3.76%      |
| 74    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.53%   | 3.53%      |
| 75    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 5.67%   | 5.41%      |
| 75    | Hommes | CAS2  | 2         | 50%        | 2            | 50%           | 5.70%   | 5.66%      |
| 75    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.92%   | 2.73%      |
| 77    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 27.39%  | 27.22%     |
| 77    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.97%   | 3.06%      |
| 77    | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 8.63%   | 8.28%      |
| 78    | Femmes | CAS2  | 3         | 75%        | 3            | 75%           | 4.12%   | 4.69%      |
| 78    | Hommes | CAS2  | 1         | 25%        | 1            | 25%           | 9.05%   | 10.25%     |
| 78    | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 7.62%   | 8.64%      |
| 79    | Femmes | CAS2  | 1         | 25%        | 1            | 25%           | 7.21%   | 6.92%      |
| 79    | Hommes | CAS2  | 1         | 25%        | 0            | 0%            | 9.79%   | 9.91%      |
| 79    | Total  | CAS2  | 2         | 50%        | 2            | 50%           | 6.26%   | 5.95%      |
| 80    | Femmes | CAS2  | 1         | 25%        | 2            | 50%           | 7.36%   | 6.51%      |
| 80    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.52%   | 1.86%      |
| 80    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.86%   | 1.07%      |

| NOGA2 | Sexe   | Valid | # val tri | % vali tri | # val tri bm | % vali tri bm | EAM (%) | EAM_bm (%) |
|-------|--------|-------|-----------|------------|--------------|---------------|---------|------------|
| 81    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 10.36%  | 8.52%      |
| 81    | Hommes | CAS2  | 4         | 100%       | 3            | 75%           | 3.66%   | 4.54%      |
| 81    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 3.45%   | 2.09%      |
| 82    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 11.12%  | 10.68%     |
| 82    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 4.17%   | 4.49%      |
| 82    | Total  | CAS2  | 3         | 75%        | 3            | 75%           | 5.17%   | 4.74%      |
| 84    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.60%   | 0.28%      |
| 84    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.69%   | 0.25%      |
| 84    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.42%   | 0.00%      |
| 85    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.22%   | 0.99%      |
| 85    | Hommes | DIR2  | 20        | 100%       | 20           | 100%          | 0.76%   | 1.55%      |
| 85    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.39%   | 0.02%      |
| 86    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 1.21%   | 1.03%      |
| 86    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 2.18%   | 2.61%      |
| 86    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 1.48%   | 0.63%      |
| 87    | Femmes | DIR2  | 20        | 100%       | 20           | 100%          | 0.28%   | 0.29%      |
| 87    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.89%   | 1.71%      |
| 87    | Total  | DIR2  | 20        | 100%       | 20           | 100%          | 0.66%   | 0.74%      |
| 88    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 3.90%   | 3.88%      |
| 88    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 0.88%   | 1.26%      |
| 88    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.71%   | 2.58%      |
| 90    | Femmes | CAS2  | 0         | 0%         | 1            | 25%           | 8.62%   | 5.52%      |
| 90    | Hommes | CAS2  | 1         | 25%        | 0            | 0%            | 9.14%   | 11.78%     |
| 90    | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 8.93%   | 8.73%      |
| 91    | Femmes | CAS2  | 4         | 100%       | 3            | 75%           | 1.46%   | 3.93%      |
| 91    | Hommes | CAS2  | 4         | 100%       | 2            | 50%           | 2.98%   | 5.02%      |
| 91    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 1.02%   | 1.66%      |
| 92    | Femmes | CAS2  | 4         | 100%       | 1            | 25%           | 2.26%   | 8.95%      |
| 92    | Hommes | CAS2  | 2         | 50%        | 2            | 50%           | 5.48%   | 8.09%      |
| 92    | Total  | CAS2  | 3         | 75%        | 3            | 75%           | 3.94%   | 4.25%      |
| 93    | Femmes | CAS2  | 4         | 100%       | 4            | 100%          | 1.65%   | 3.50%      |
| 93    | Hommes | CAS2  | 3         | 75%        | 3            | 75%           | 4.97%   | 4.60%      |
| 93    | Total  | CAS2  | 4         | 100%       | 4            | 100%          | 2.28%   | 2.21%      |
| 94    | Femmes | CAS2  | 1         | 25%        | 4            | 100%          | 8.92%   | 3.83%      |
| 94    | Hommes | CAS2  | 1         | 25%        | 0            | 0%            | 11.46%  | 19.31%     |
| 94    | Total  | CAS2  | 1         | 25%        | 1            | 25%           | 10.03%  | 10.60%     |
| 95    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 25.93%  | 27.44%     |
| 95    | Hommes | CAS2  | 3         | 75%        | 4            | 100%          | 2.69%   | 4.28%      |
| 95    | Total  | CAS2  | 0         | 0%         | 2            | 50%           | 8.71%   | 6.13%      |
| 96    | Femmes | CAS2  | 0         | 0%         | 0            | 0%            | 7.84%   | 9.45%      |
| 96    | Hommes | CAS2  | 4         | 100%       | 4            | 100%          | 1.75%   | 3.18%      |
| 96    | Total  | CAS2  | 1         | 25%        | 0            | 0%            | 6.43%   | 7.36%      |

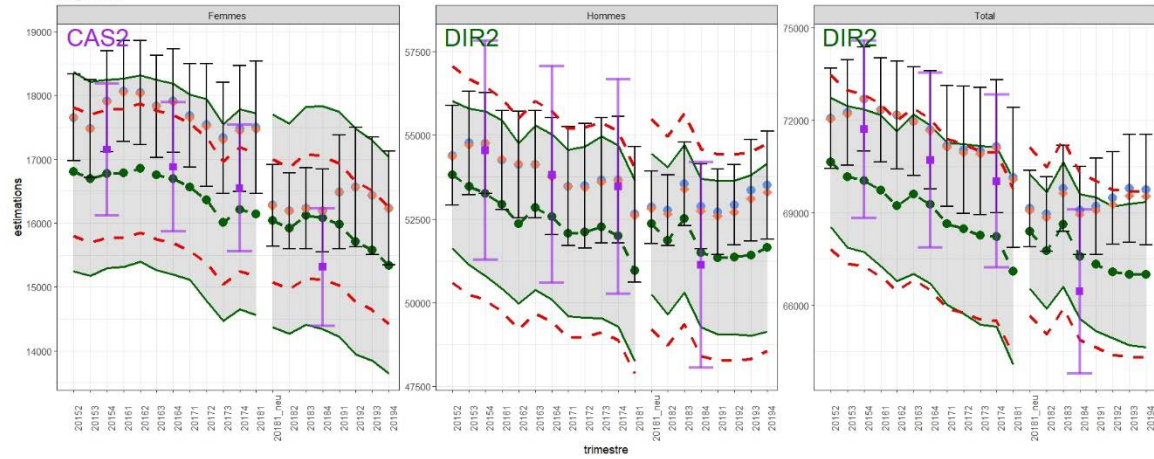
La table suivante donne les écarts relatifs moyens spécifiquement pour les trimestres non-vali-dées. Elle est l'équivalent pour les NOGA2 de la Table 3 de la section 6.2.1

| NOGA2 | Sexe   | Valid | #val tri | %val trim | #val tri bm | %val tri bm | EAM (%) | EAM bm (%) | EAM.n.val (%) | EAM.n.val bm (%) |
|-------|--------|-------|----------|-----------|-------------|-------------|---------|------------|---------------|------------------|
| 9     | Femmes | CAS2  | 0        | 0%        | 3           | 75%         | 20.35%  | 6.13%      | 20.35%        | 19.48%           |
| 9     | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 24.41%  | 25.65%     | 31.85%        | 33.09%           |
| 9     | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 22.89%  | 17.69%     | 29.00%        | 22.92%           |
| 11    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 3.96%   | 4.18%      | 6.54%         | 6.65%            |
| 13    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 4.88%   | 4.94%      | 7.40%         | 7.44%            |
| 14    | Hommes | CAS2  | 0        | 0%        | 0           | 0%          | 13.40%  | 13.02%     | 13.40%        | 13.02%           |
| 16    | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 8.77%   | 8.75%      | 9.83%         | 9.79%            |
| 17    | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 6.74%   | 6.44%      | 7.52%         | 7.14%            |
| 17    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 4.38%   | 4.33%      | 7.92%         | 7.40%            |
| 17    | Total  | CAS2  | 3        | 75%       | 3           | 75%         | 5.07%   | 4.94%      | 7.42%         | 7.03%            |
| 19    | Femmes | CAS2  | 2        | 50%       | 2           | 50%         | 6.64%   | 6.46%      | 11.67%        | 11.25%           |
| 19    | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 22.27%  | 23.10%     | 29.57%        | 30.14%           |
| 19    | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 18.96%  | 19.58%     | 25.01%        | 25.40%           |
| 23    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 3.75%   | 3.53%      | 7.10%         | 6.57%            |
| 24    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 3.83%   | 3.80%      | 7.52%         | 7.41%            |
| 29    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 2.93%   | 2.69%      | 7.77%         | 7.32%            |
| 31    | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 7.97%   | 7.86%      | 8.76%         | 8.62%            |
| 33    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 4.34%   | 4.34%      | 6.79%         | 6.65%            |
| 35    | Total  | DIR2  | 19       | 95%       | 20          | 100%        | 1.51%   | 0.01%      | 2.44%         | -                |
| 36    | Femmes | CAS2  | 1        | 25%       | 0           | 0%          | 14.66%  | 40.77%     | 18.62%        | 40.77%           |
| 36    | Hommes | CAS2  | 2        | 50%       | 4           | 100%        | 5.39%   | 4.80%      | 6.69%         | -                |
| 37    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 10.83%  | 20.58%     | 10.83%        | 20.58%           |
| 39    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 20.26%  | 24.56%     | 20.26%        | 24.56%           |
| 39    | Hommes | CAS2  | 0        | 0%        | 0           | 0%          | 15.48%  | 15.78%     | 15.48%        | 15.78%           |
| 39    | Total  | CAS2  | 0        | 0%        | 0           | 0%          | 12.46%  | 12.38%     | 12.46%        | 12.38%           |
| 41    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 3.96%   | 5.34%      | 10.07%        | 9.25%            |
| 45    | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 4.39%   | 4.57%      | 6.24%         | 6.45%            |
| 47    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 5.42%   | 4.99%      | 8.66%         | 7.58%            |
| 49    | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 9.17%   | 8.64%      | 10.36%        | 9.68%            |
| 50    | Femmes | CAS2  | 2        | 50%       | 3           | 75%         | 5.53%   | 4.18%      | 7.20%         | 6.07%            |
| 50    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 6.17%   | 6.54%      | 16.63%        | 16.73%           |
| 50    | Total  | CAS2  | 3        | 75%       | 3           | 75%         | 5.42%   | 5.49%      | 10.79%        | 11.19%           |
| 58    | Femmes | CAS2  | 1        | 25%       | 0           | 0%          | 7.40%   | 10.62%     | 8.06%         | 10.62%           |
| 58    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 3.51%   | 3.50%      | 6.68%         | 6.38%            |
| 58    | Total  | CAS2  | 3        | 75%       | 1           | 25%         | 5.41%   | 6.99%      | 8.04%         | 8.14%            |
| 59    | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 3.92%   | 4.25%      | 6.43%         | 6.32%            |
| 59    | Hommes | CAS2  | 2        | 50%       | 2           | 50%         | 7.11%   | 7.07%      | 13.31%        | 13.24%           |
| 59    | Total  | CAS2  | 2        | 50%       | 2           | 50%         | 5.29%   | 5.40%      | 9.46%         | 9.75%            |
| 63    | Femmes | CAS2  | 3        | 75%       | 1           | 25%         | 6.92%   | 7.70%      | 16.14%        | 9.70%            |
| 63    | Total  | CAS2  | 3        | 75%       | 3           | 75%         | 3.94%   | 4.97%      | 7.48%         | 8.57%            |
| 64    | Total  | DIR2  | 19       | 95%       | 19          | 95%         | 0.89%   | 0.85%      | 1.40%         | 1.36%            |
| 69    | Femmes | CAS2  | 3        | 75%       | 4           | 100%        | 4.73%   | 3.83%      | 6.11%         | -                |
| 74    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 4.08%   | 4.05%      | 12.04%        | 11.89%           |
| 75    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 5.67%   | 5.41%      | 9.22%         | 8.90%            |
| 75    | Hommes | CAS2  | 2        | 50%       | 2           | 50%         | 5.70%   | 5.66%      | 7.41%         | 7.44%            |
| 77    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 27.39%  | 27.22%     | 27.39%        | 27.22%           |
| 77    | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 8.63%   | 8.28%      | 10.64%        | 10.23%           |
| 78    | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 4.12%   | 4.69%      | 8.42%         | 8.86%            |
| 78    | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 9.05%   | 10.25%     | 11.40%        | 13.00%           |
| 78    | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 7.62%   | 8.64%      | 9.53%         | 10.76%           |
| 79    | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 7.21%   | 6.92%      | 8.80%         | 8.41%            |
| 79    | Hommes | CAS2  | 1        | 25%       | 0           | 0%          | 9.79%   | 9.91%      | 11.06%        | 9.91%            |
| 79    | Total  | CAS2  | 2        | 50%       | 2           | 50%         | 6.26%   | 5.95%      | 9.57%         | 9.12%            |
| 80    | Femmes | CAS2  | 1        | 25%       | 2           | 50%         | 7.36%   | 6.51%      | 8.45%         | 9.03%            |
| 81    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 10.36%  | 8.52%      | 10.36%        | 8.52%            |
| 82    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 11.12%  | 10.68%     | 11.12%        | 10.68%           |
| 82    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 4.17%   | 4.49%      | 8.66%         | 8.70%            |
| 82    | Total  | CAS2  | 3        | 75%       | 3           | 75%         | 5.17%   | 4.74%      | 7.84%         | 7.23%            |
| 90    | Femmes | CAS2  | 0        | 0%        | 1           | 25%         | 8.62%   | 5.52%      | 8.62%         | 6.89%            |
| 90    | Hommes | CAS2  | 1        | 25%       | 0           | 0%          | 9.14%   | 11.78%     | 11.23%        | 11.78%           |
| 90    | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 8.93%   | 8.73%      | 10.19%        | 10.05%           |
| 92    | Hommes | CAS2  | 2        | 50%       | 2           | 50%         | 5.48%   | 8.09%      | 9.47%         | 13.99%           |
| 92    | Total  | CAS2  | 3        | 75%       | 3           | 75%         | 3.94%   | 4.25%      | 7.01%         | 7.83%            |
| 93    | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 4.97%   | 4.60%      | 8.43%         | 7.80%            |
| 94    | Femmes | CAS2  | 1        | 25%       | 4           | 100%        | 8.92%   | 3.83%      | 10.54%        | -                |
| 94    | Hommes | CAS2  | 1        | 25%       | 0           | 0%          | 11.46%  | 19.31%     | 13.58%        | 19.31%           |
| 94    | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 10.03%  | 10.60%     | 11.88%        | 12.64%           |
| 95    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 25.93%  | 27.44%     | 25.93%        | 27.44%           |
| 95    | Hommes | CAS2  | 3        | 75%       | 4           | 100%        | 2.69%   | 4.28%      | 6.44%         | -                |
| 95    | Total  | CAS2  | 0        | 0%        | 2           | 50%         | 8.71%   | 6.13%      | 8.71%         | 7.08%            |
| 96    | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 7.84%   | 9.45%      | 7.84%         | 9.45%            |
| 96    | Total  | CAS2  | 1        | 25%       | 0           | 0%          | 6.43%   | 7.36%      | 6.96%         | 7.36%            |

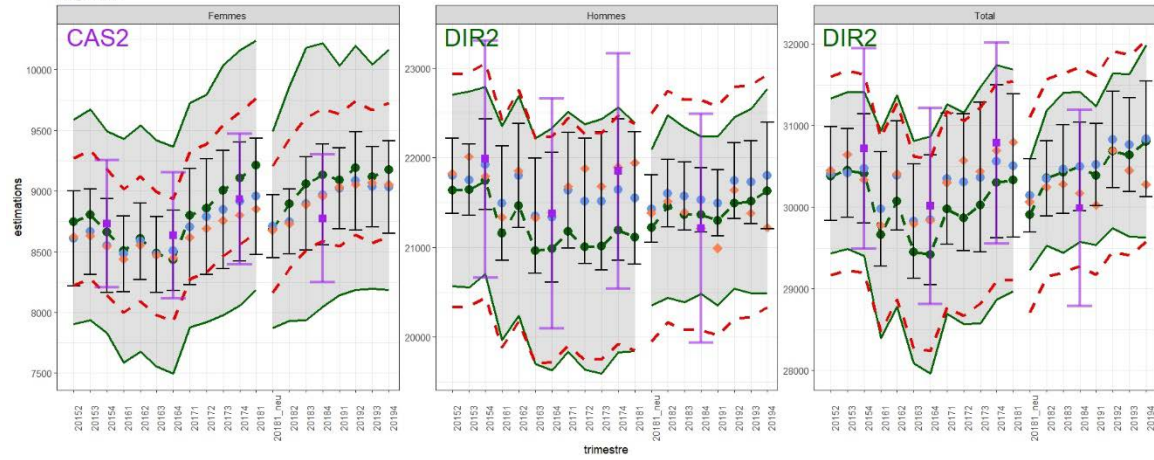
## Annexe 8 : séries temporelles SAE pour les NOGA\_OFS50



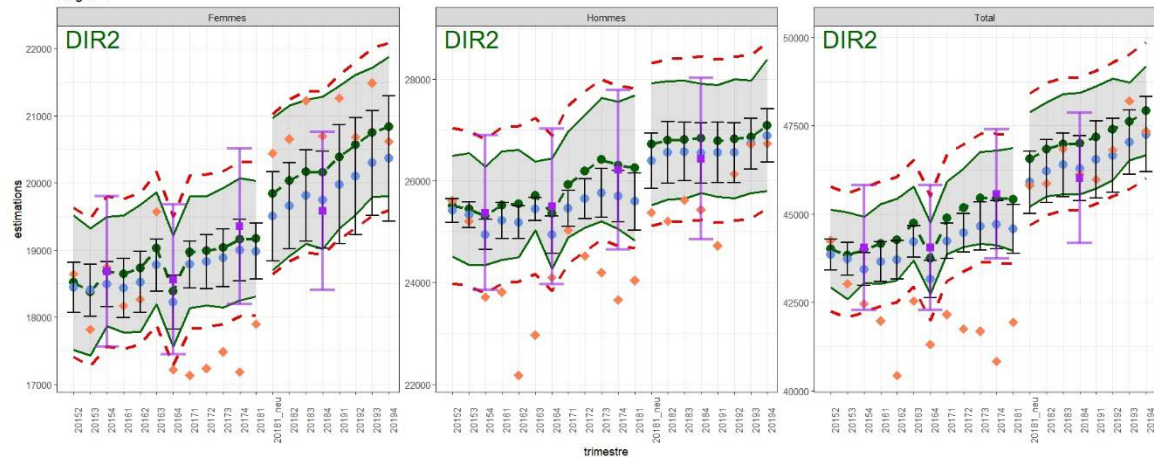
Noga 16.8



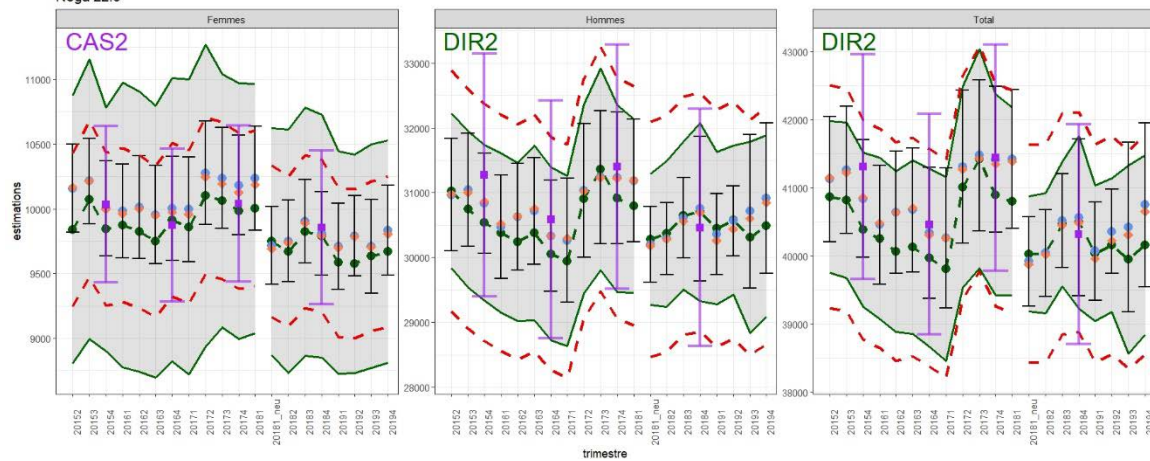
Noga 19.2



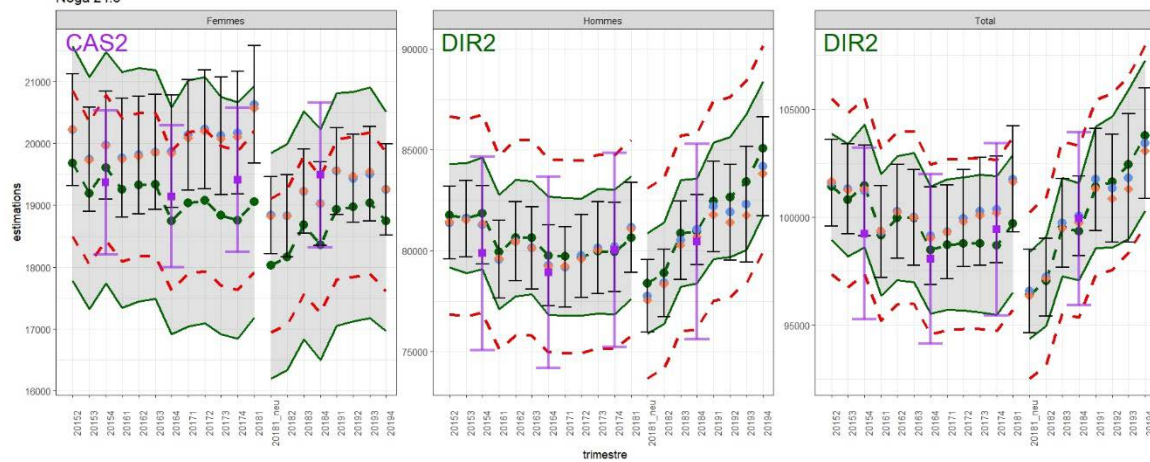
Noga 21



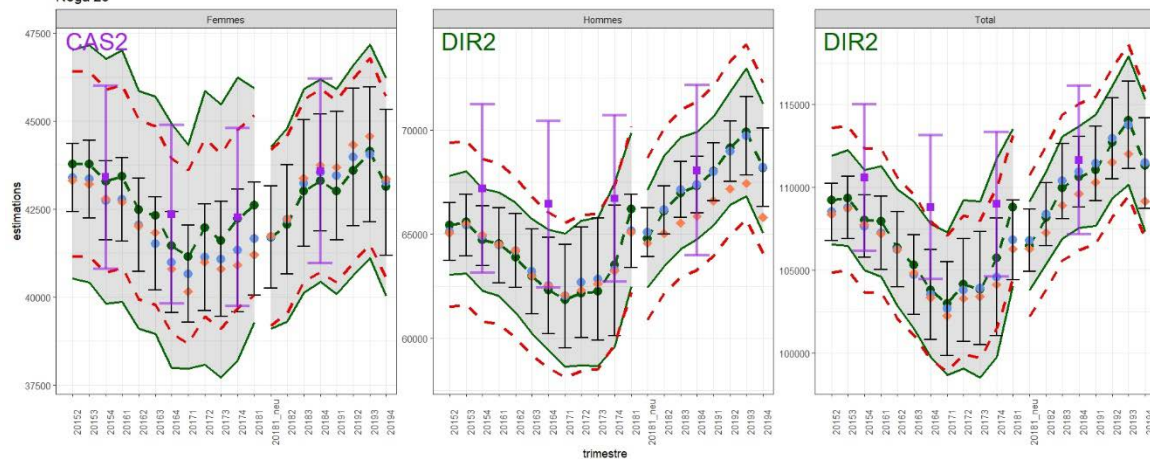
Noga 22.3



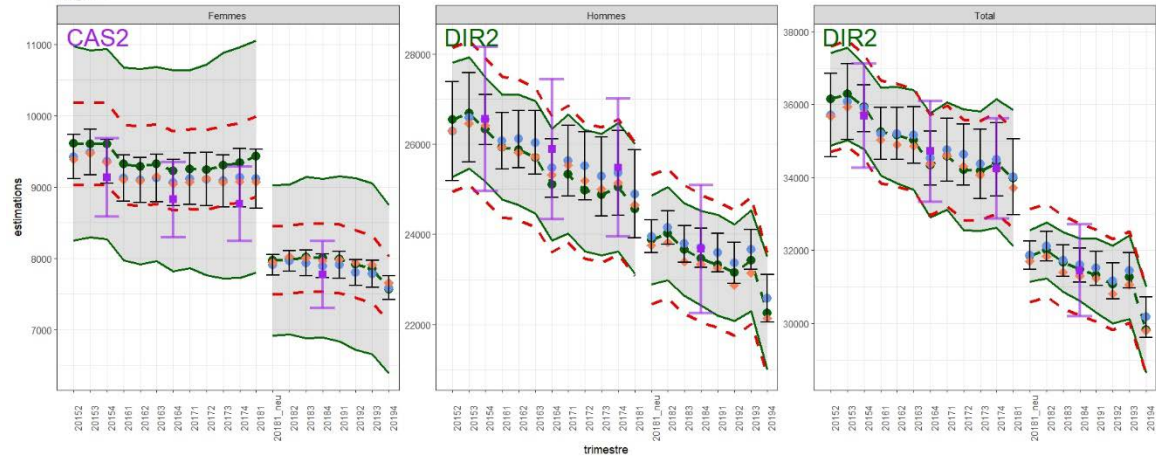
Noga 24.5



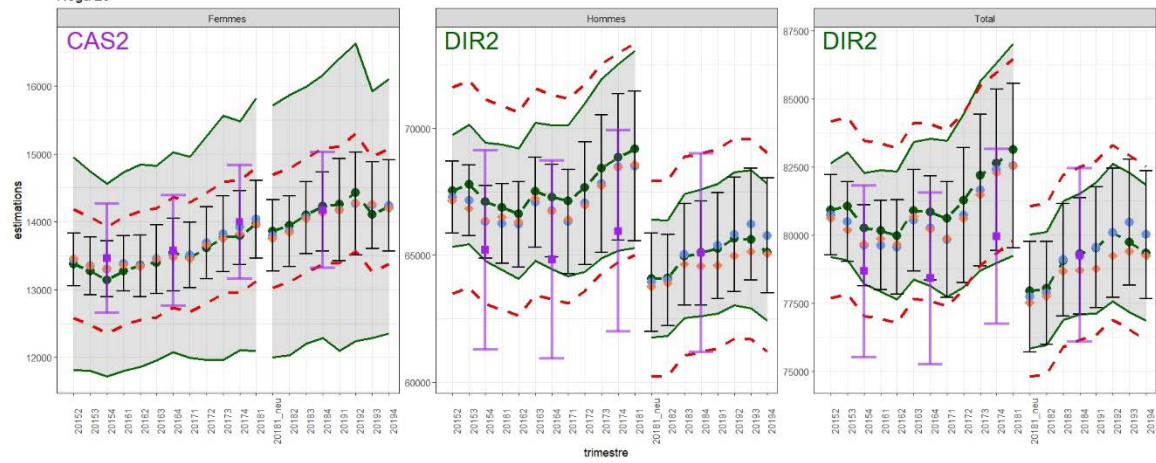
Noga 26



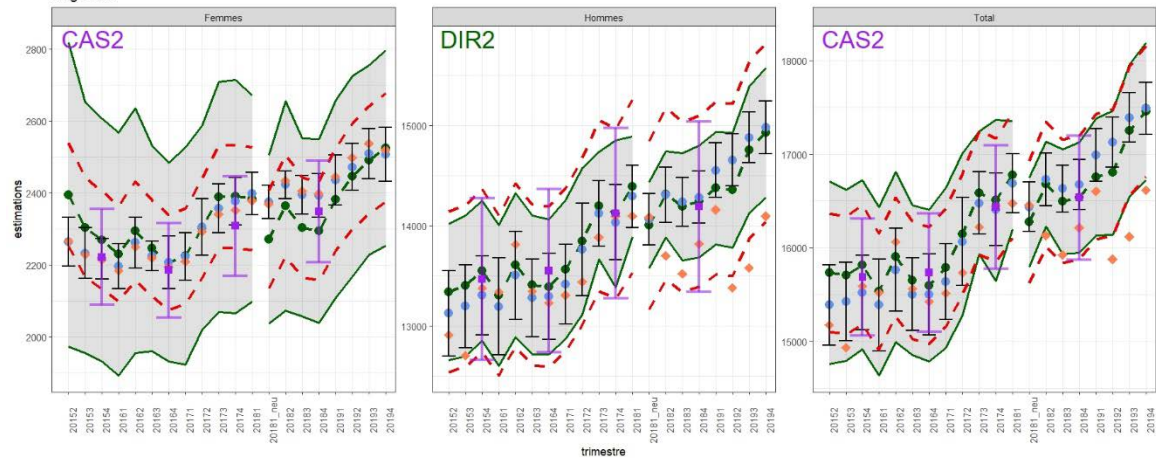
Noga 27



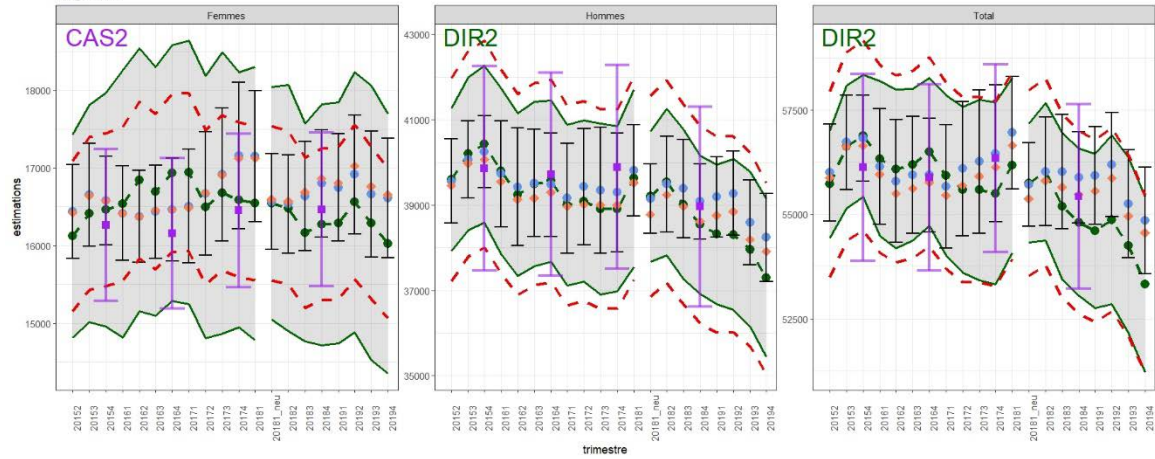
Noga 28



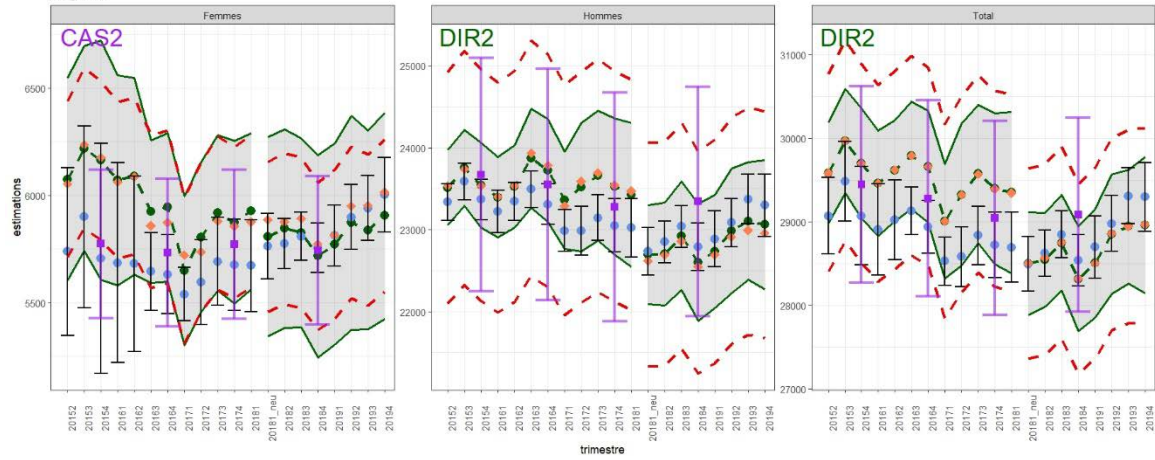
Noga 29.3



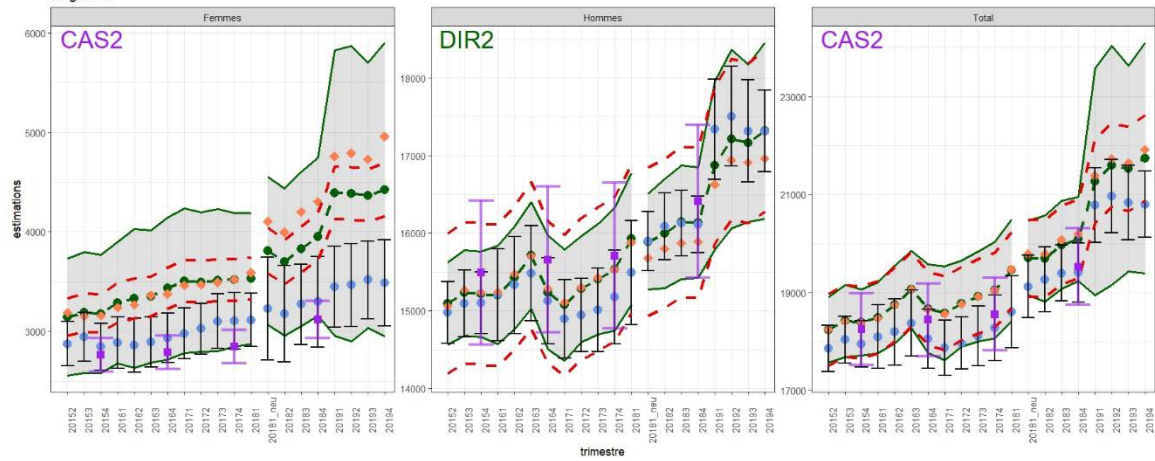
Noga 31.3



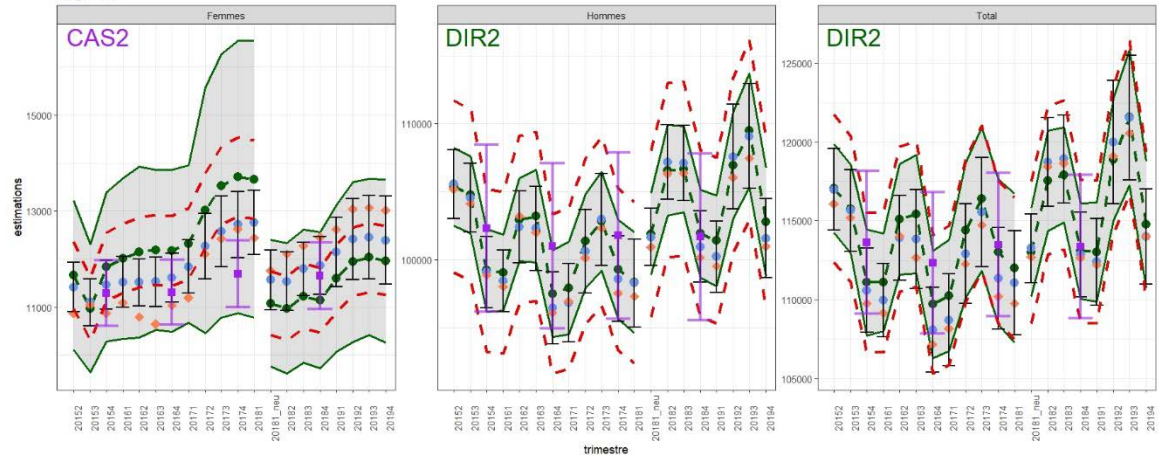
Noga 35



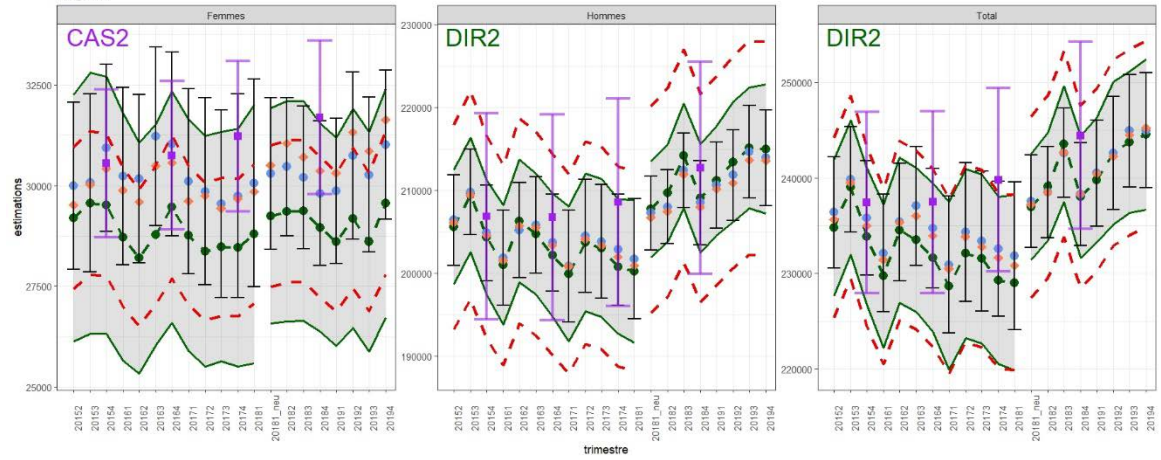
Noga 36.9



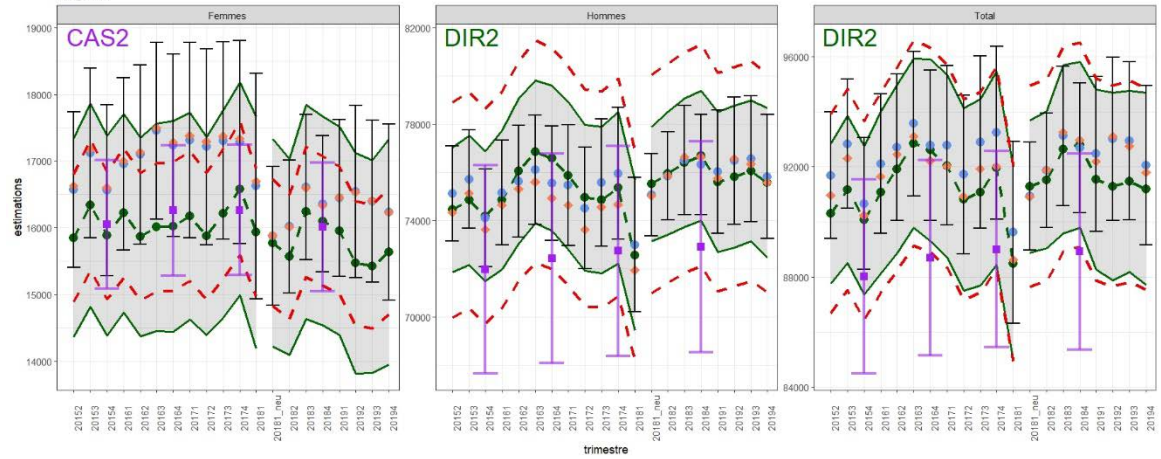
Noga 41.2



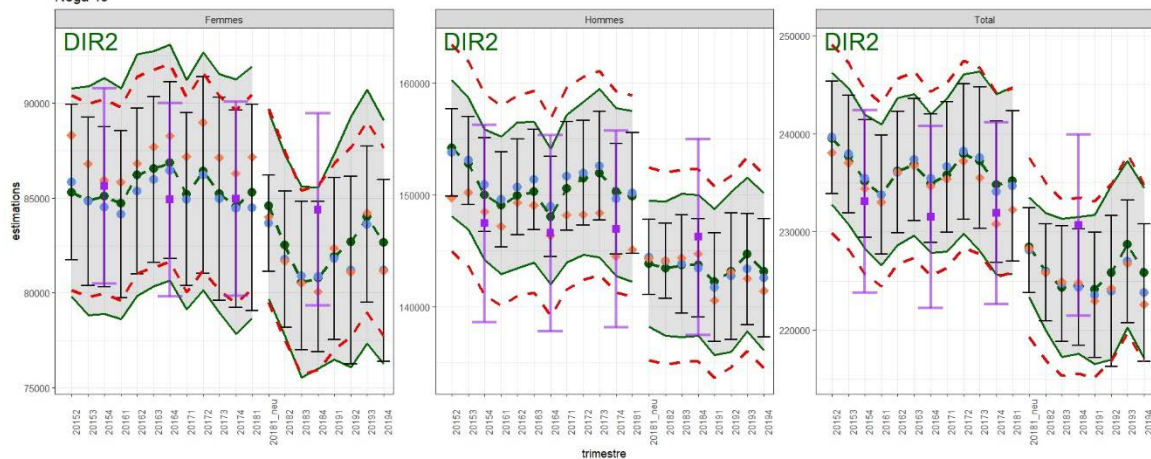
Noga 43



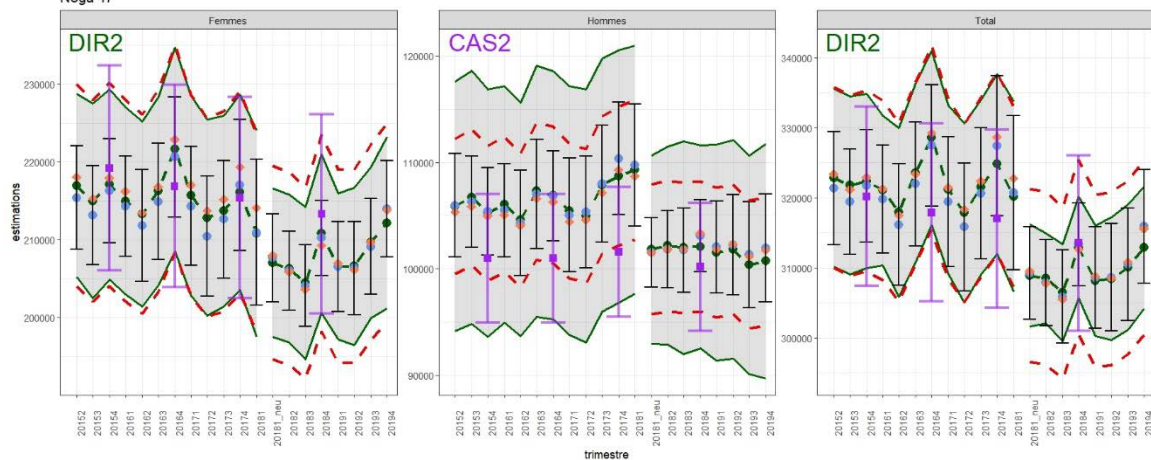
Noga 45



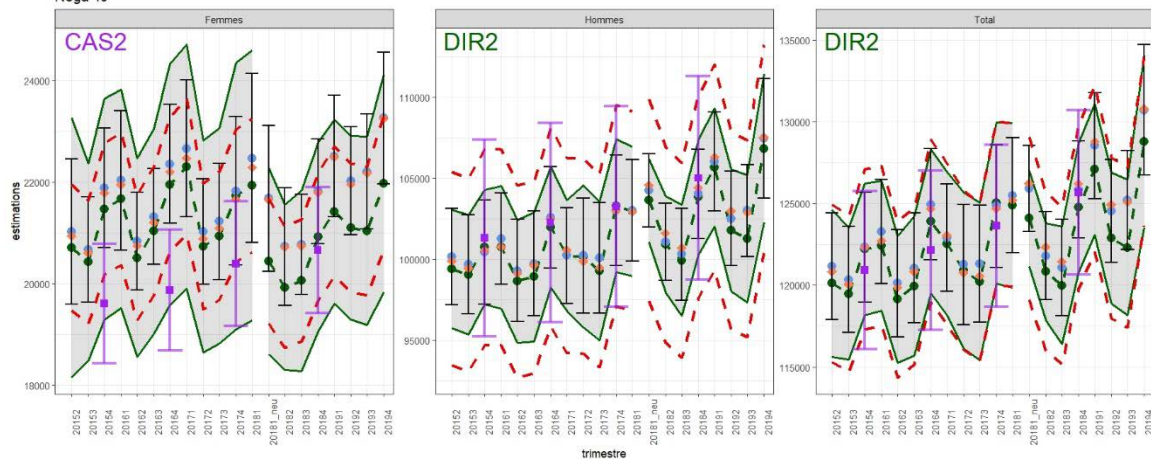
Noga 46



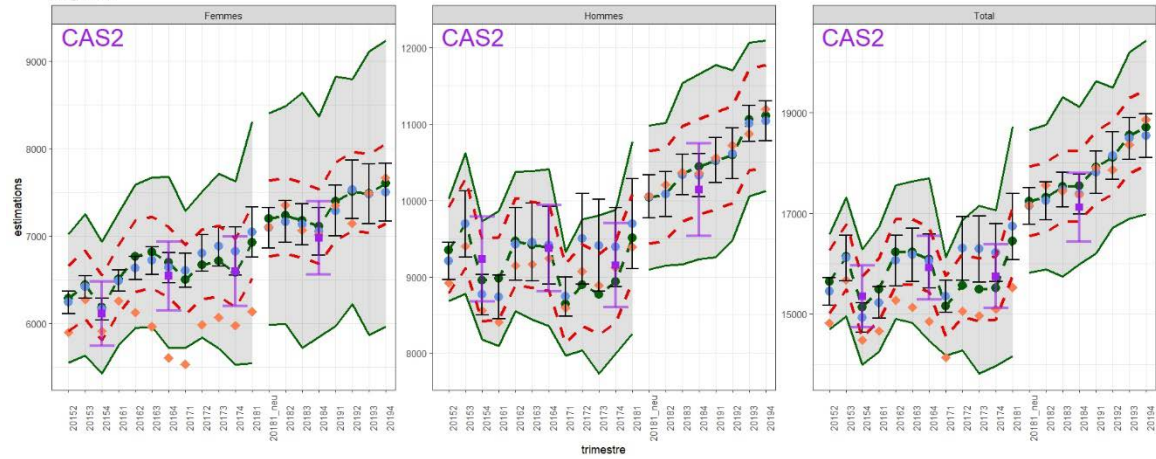
Noga 47



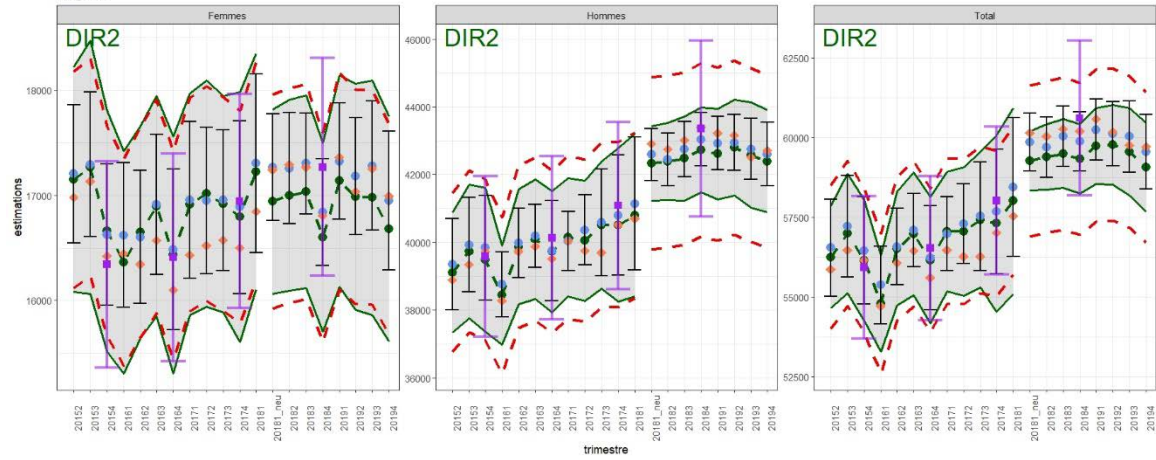
Noga 49



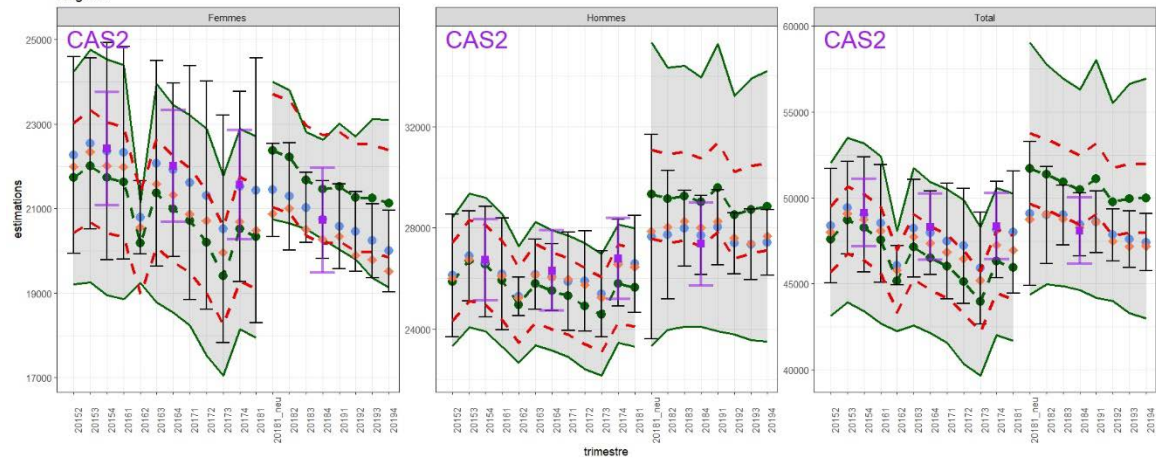
Noga 50.1

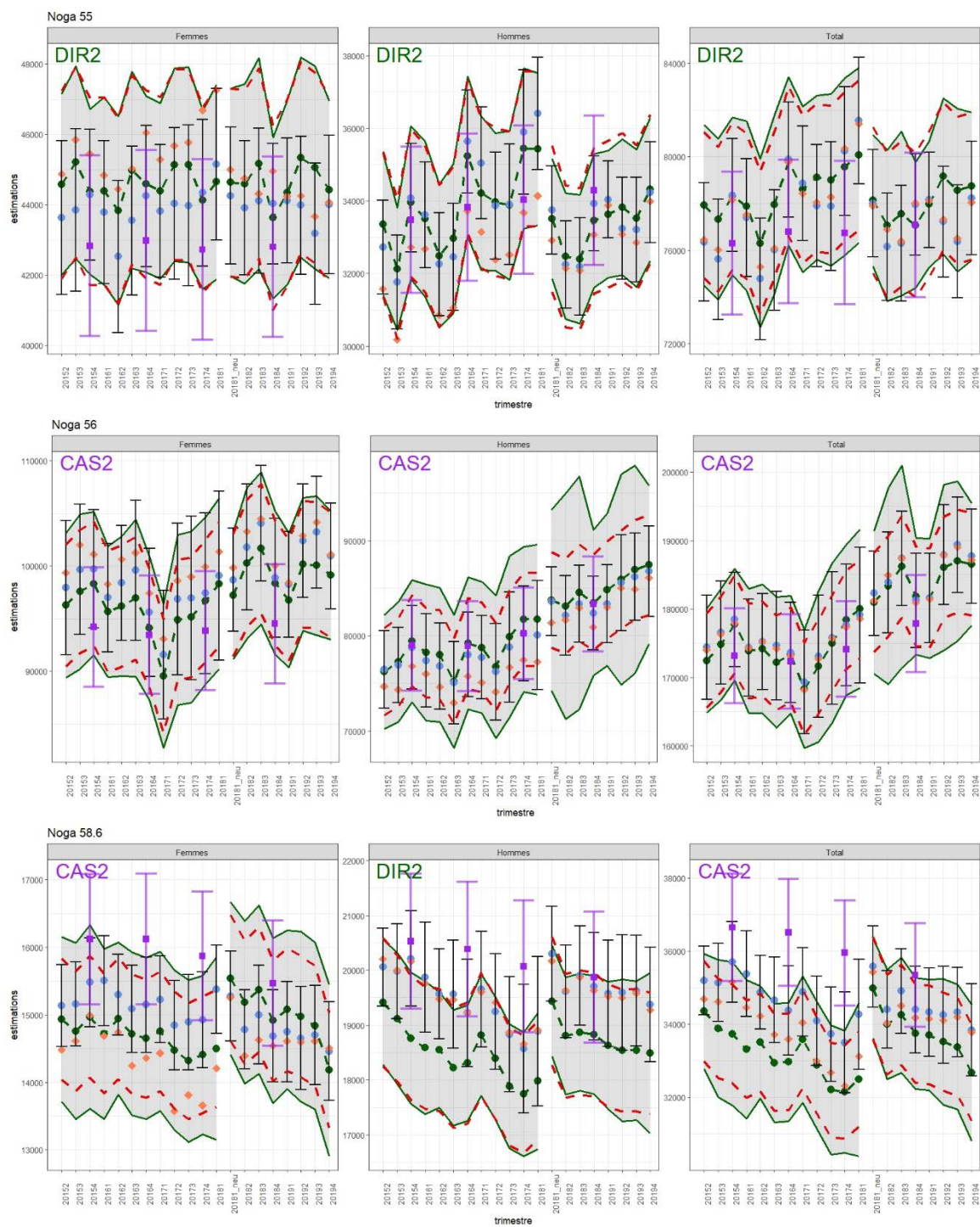


Noga 52

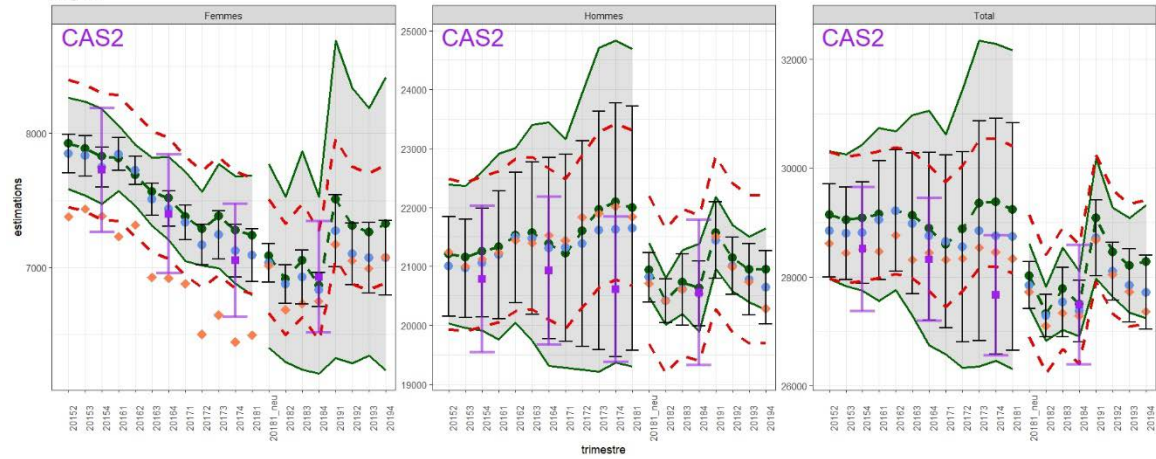


Noga 53

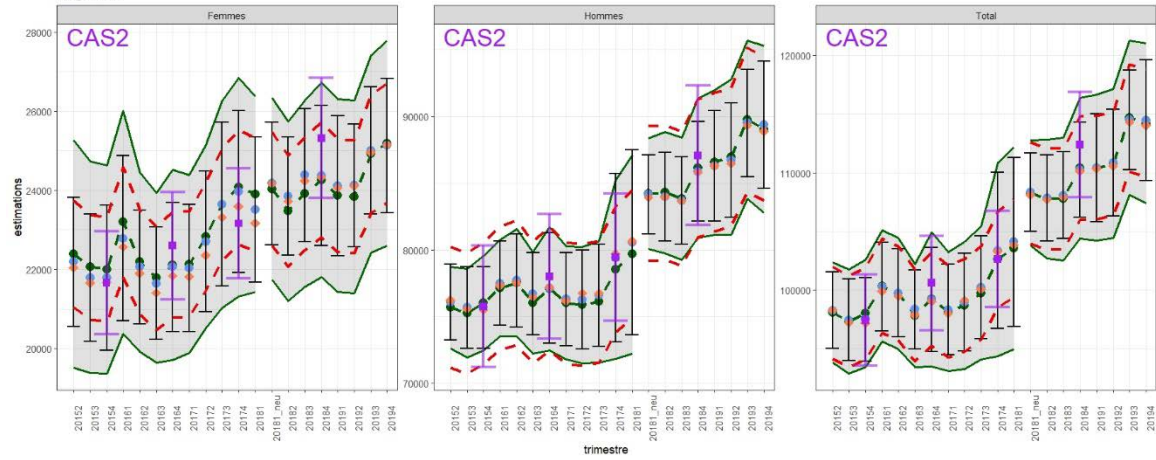




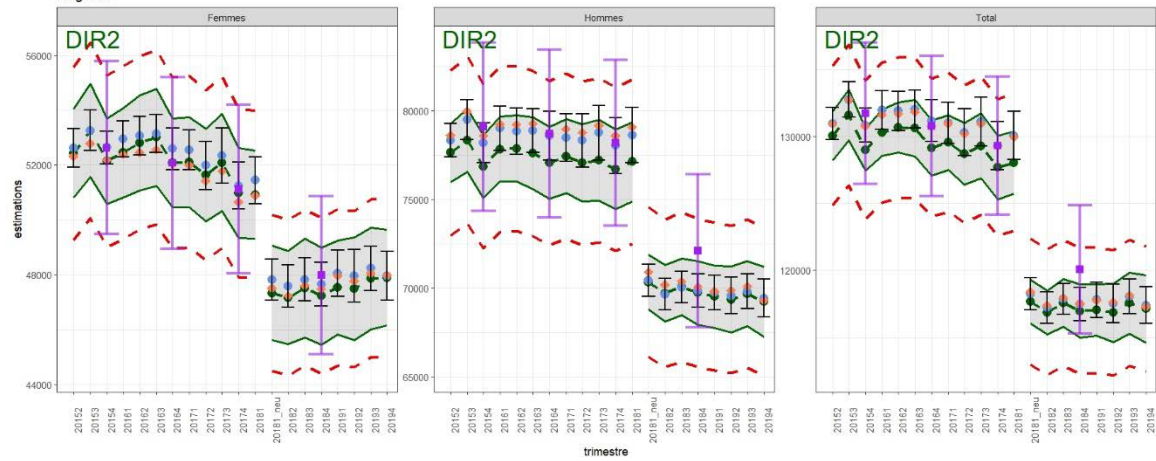
Noga 61



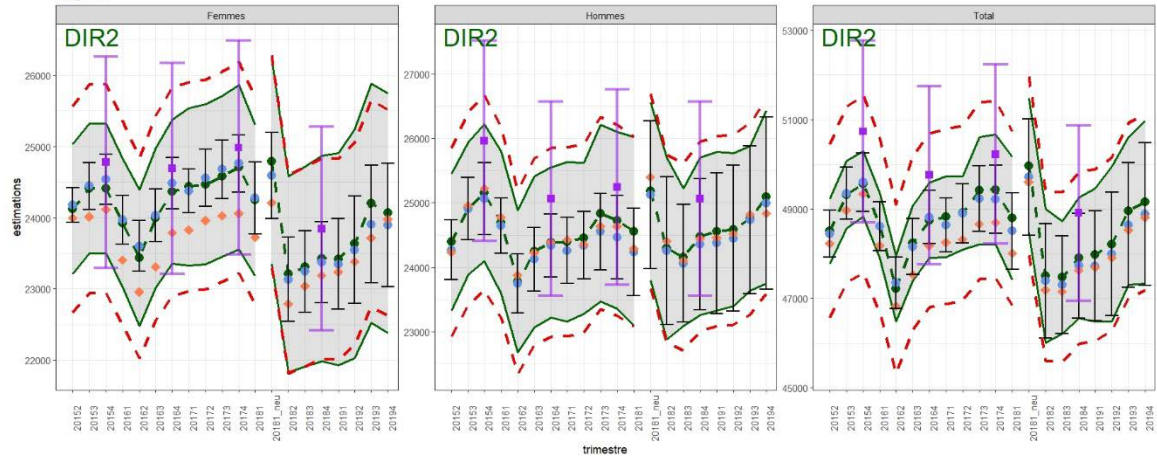
Noga 62.3



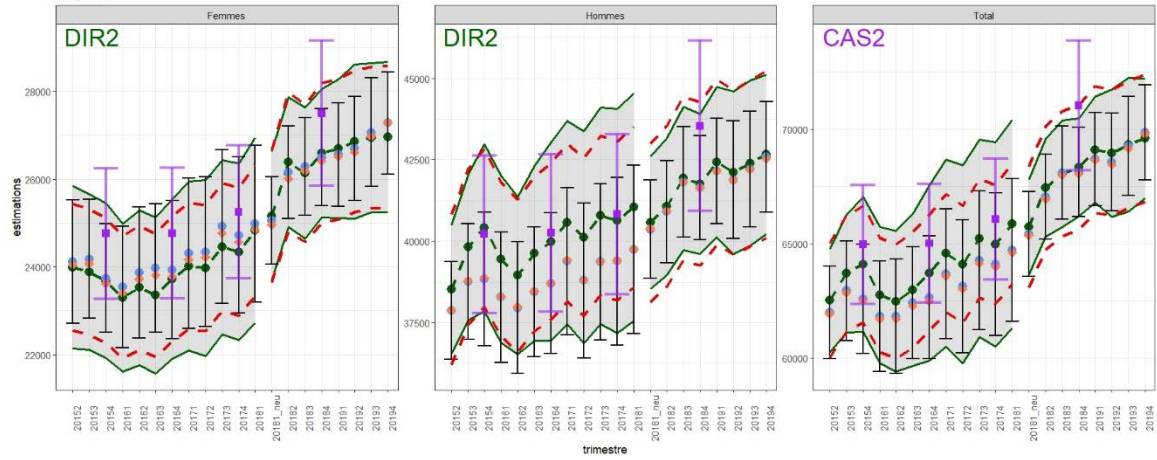
Noga 64



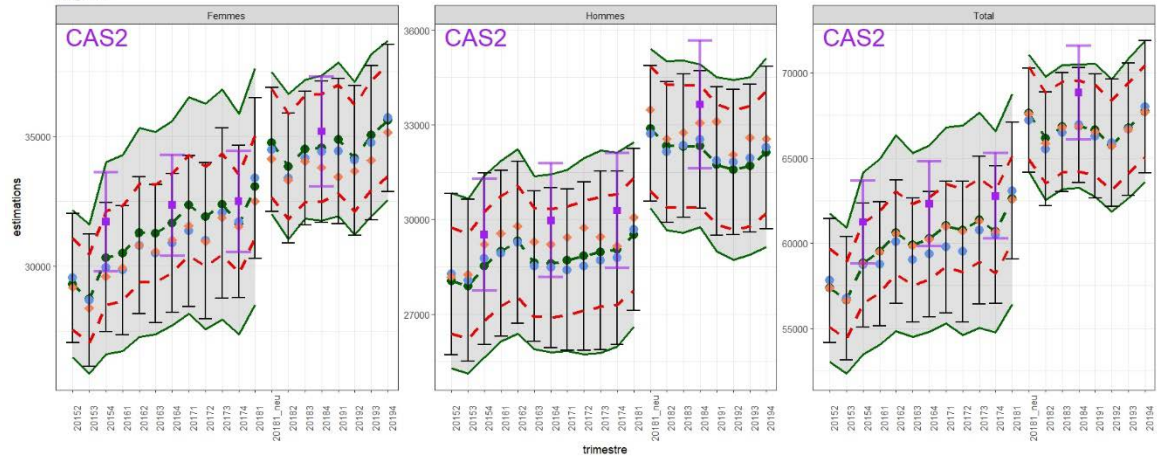
Noga 65



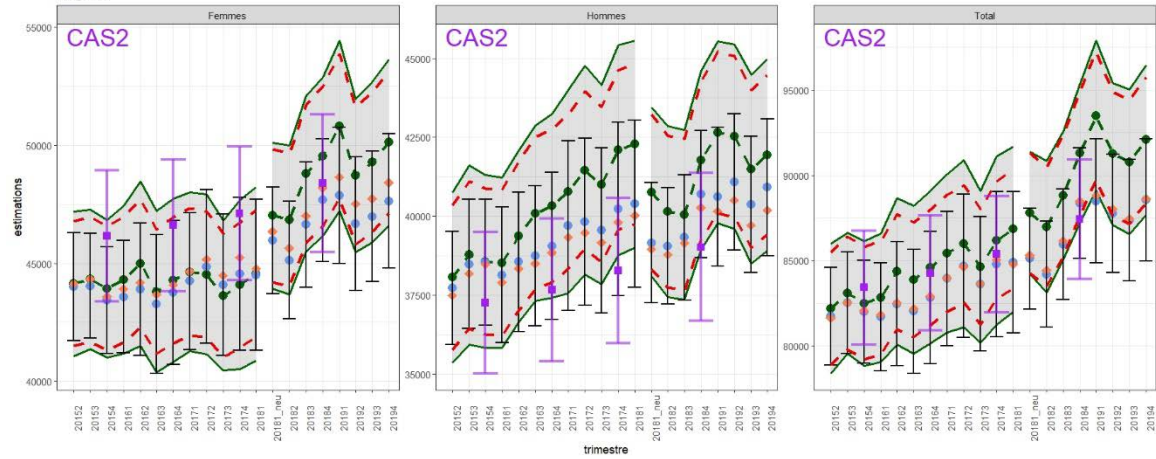
Noga 66



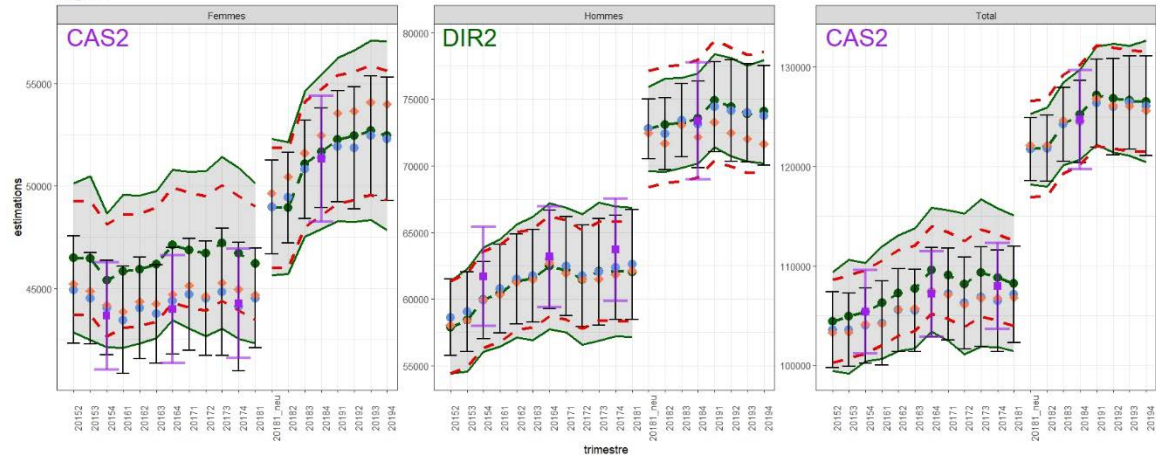
Noga 68



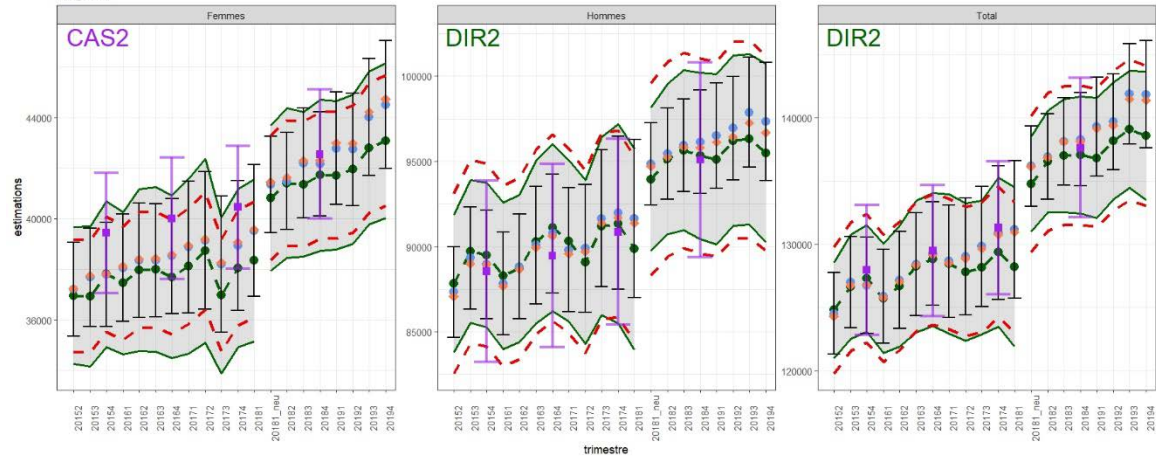
Noga 69



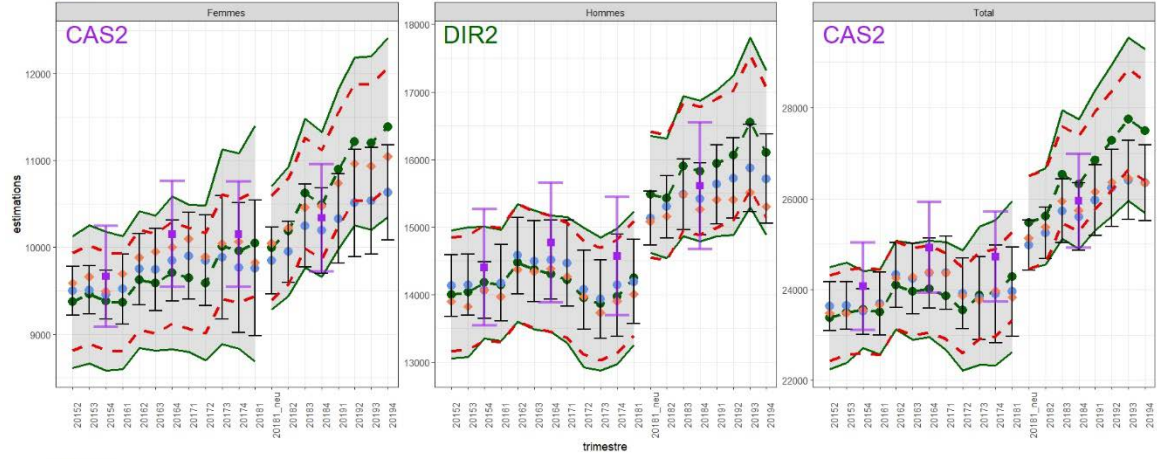
Noga 70



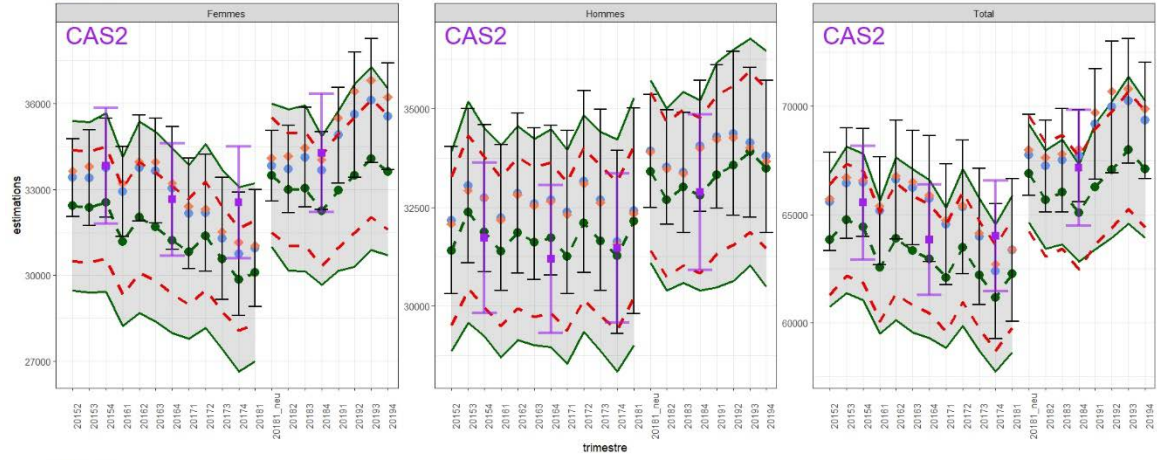
Noga 71



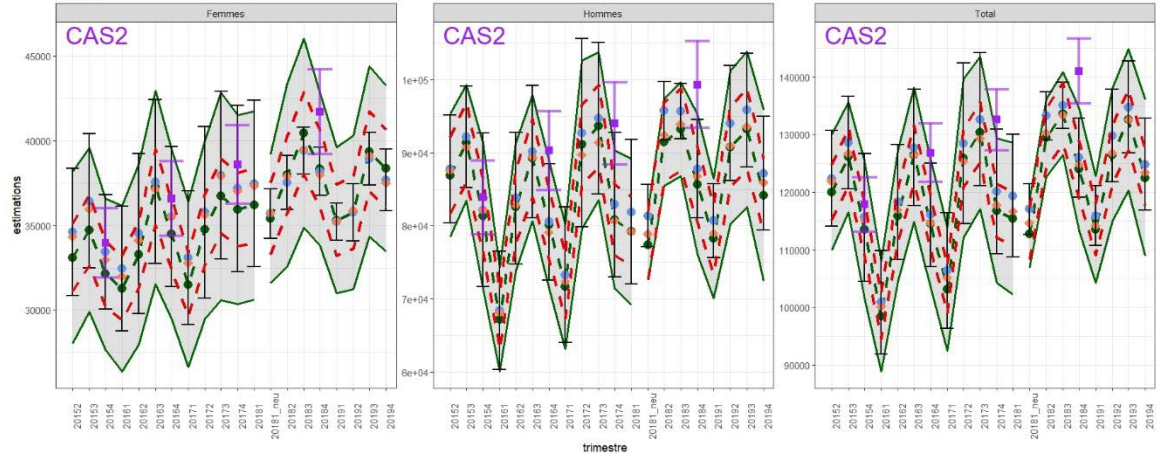
Noga 72

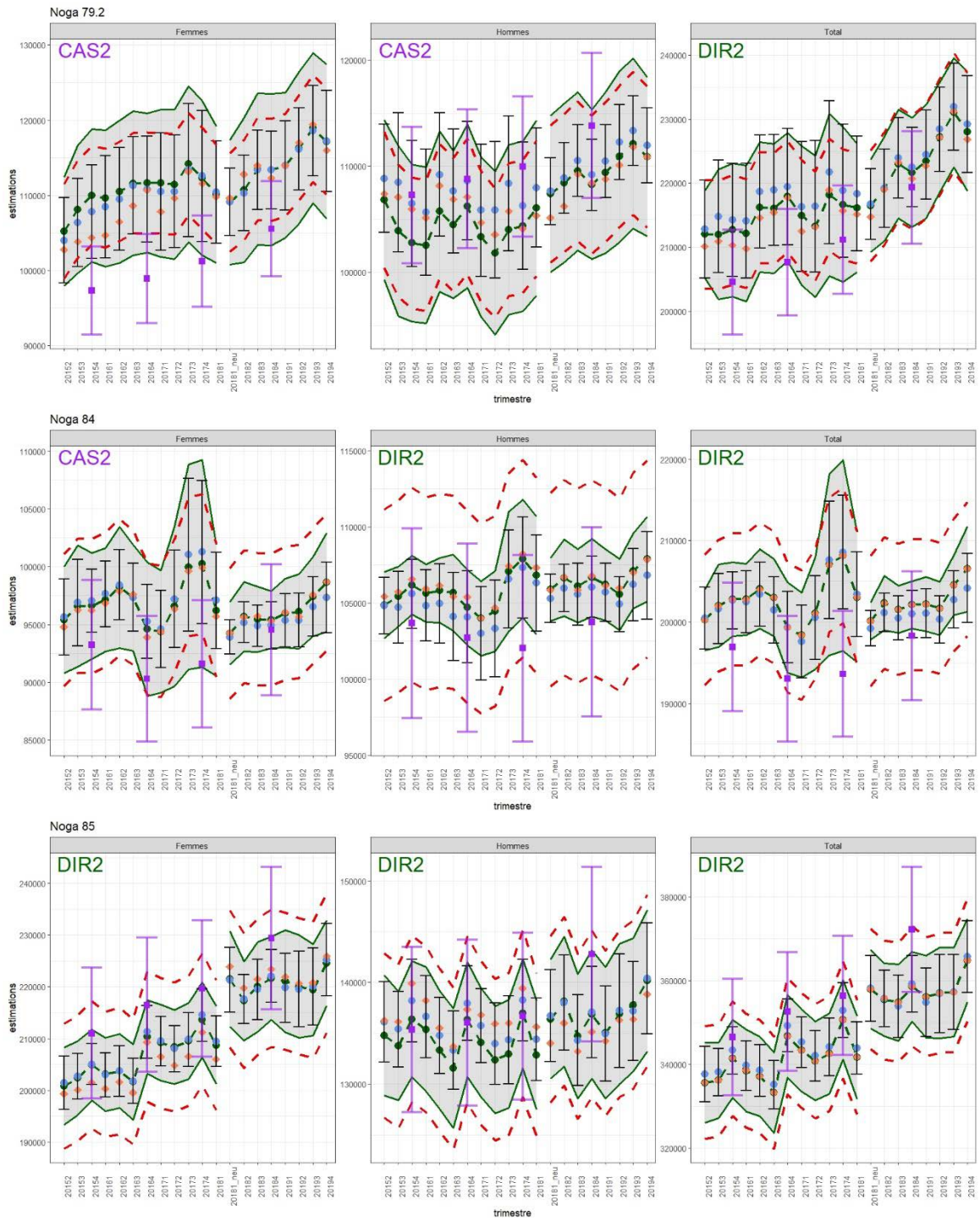


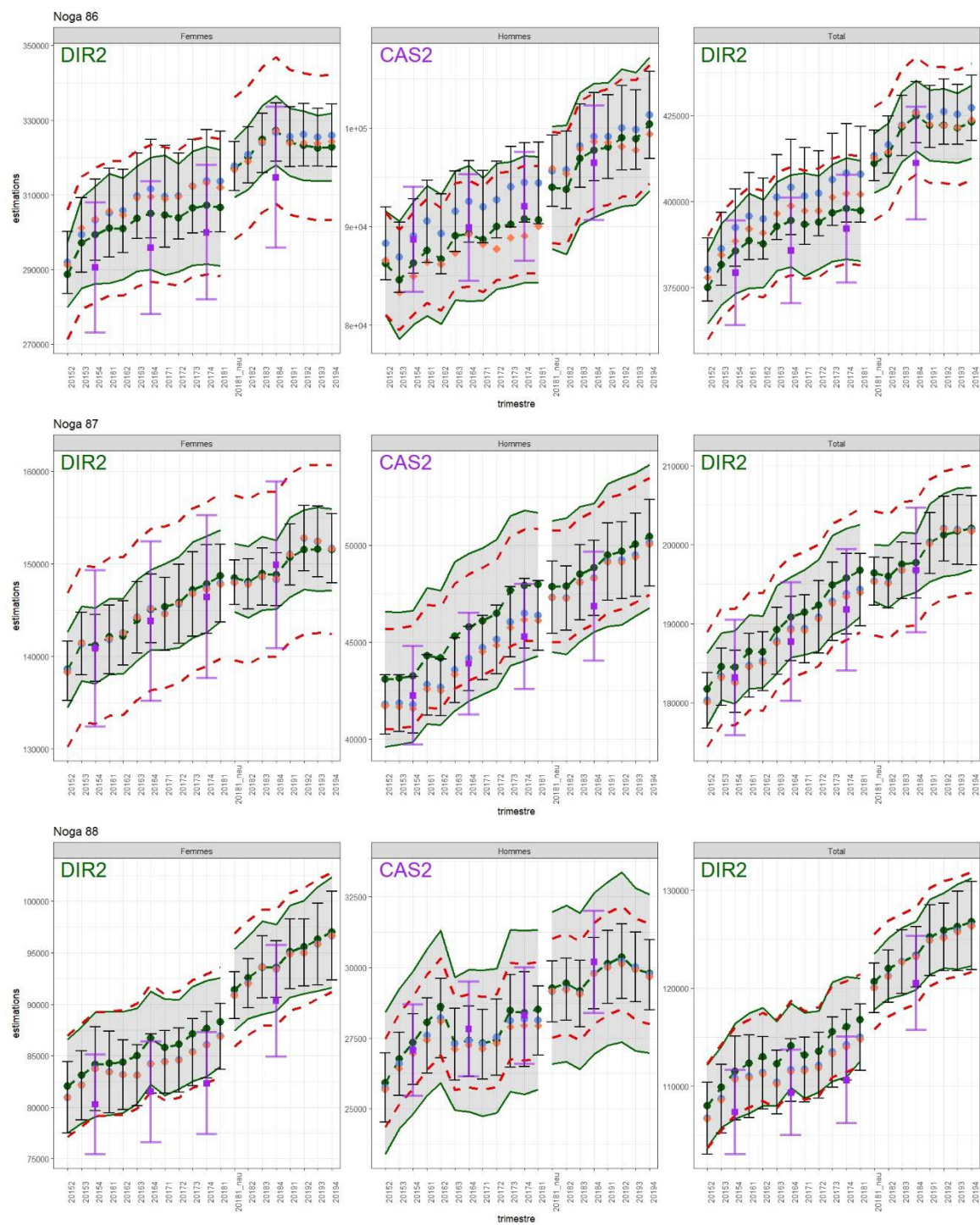
Noga 73.5



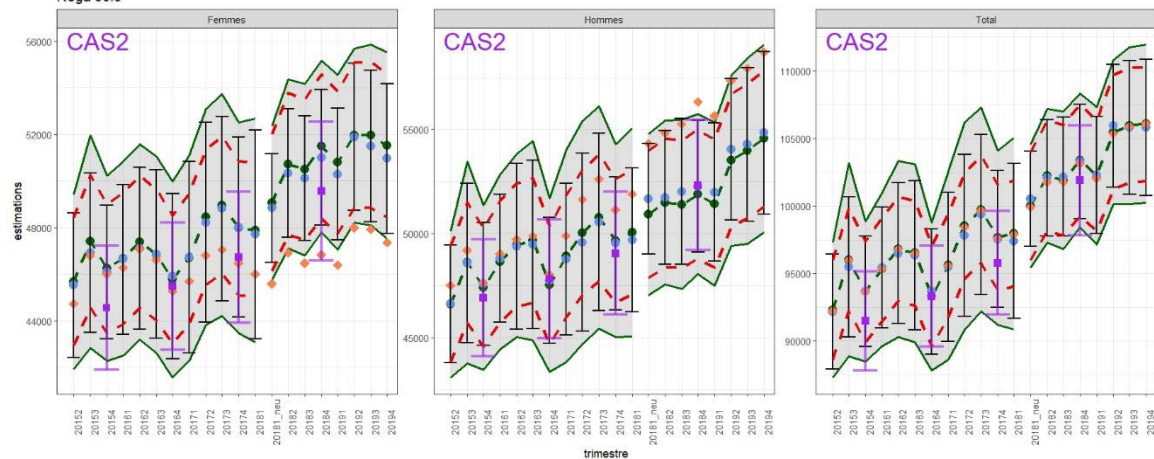
Noga 78



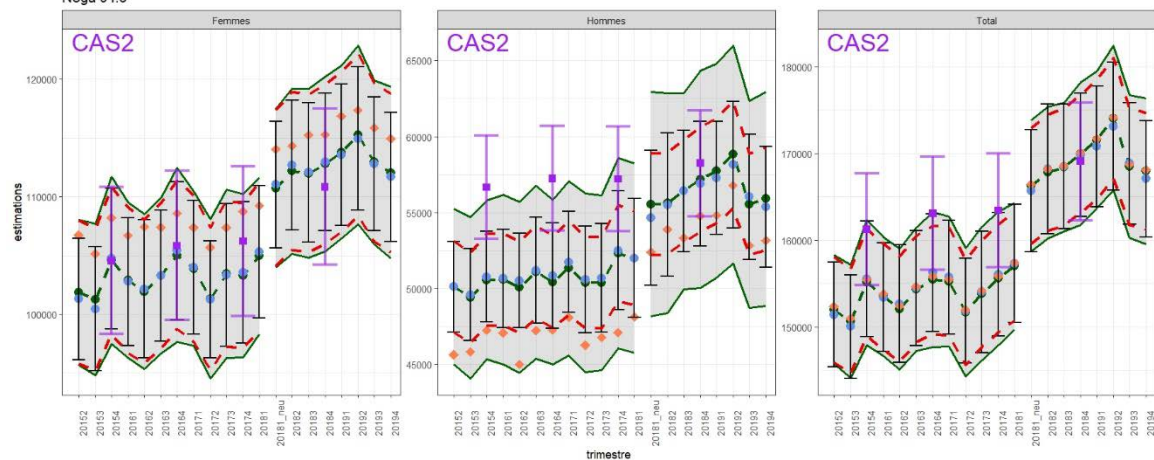




Noga 90.3



Noga 94.6



## Annexe 9 : résumé des résultats de la validation au niveau des NOGA\_OFS50

La table ci-dessous donne pour les NOGA\_OFS50 le nombre et le pourcentage de trimestres validés, ainsi que l'écart relatif (en valeur absolue) moyen entre les estimations SAE (sans ou avec **benchmarking**) et les valeurs de **référence pour la validation** (DIR2 ou CAS2 selon le cas).

| NOGA_OFS50 | Sexe   | Valid | #val tri | %val trim | #val tri bm | %val tri bm | EAM (%) | EAM bm (%) |
|------------|--------|-------|----------|-----------|-------------|-------------|---------|------------|
| 5.9        | Femmes | CAS2  | 4        | 100%      | 0           | 0%          | 1.97%   | 11.98%     |
| 5.9        | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.74%   | 1.94%      |
| 5.9        | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.37%   | 0.81%      |
| 10.2       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 0.88%   | 0.68%      |
| 10.2       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.87%   | 0.49%      |
| 10.2       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.78%   | 0.47%      |
| 13.5       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 1.26%   | 2.02%      |
| 13.5       | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 6.28%   | 6.20%      |
| 13.5       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 2.89%   | 3.46%      |
| 16.8       | Femmes | CAS2  | 3        | 75%       | 4           | 100%        | 5.49%   | 5.39%      |
| 16.8       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 2.59%   | 2.46%      |
| 16.8       | Total  | DIR2  | 12       | 60%       | 14          | 70%         | 3.26%   | 3.15%      |
| 19.2       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 1.49%   | 1.96%      |
| 19.2       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.36%   | 1.63%      |
| 19.2       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.64%   | 0.94%      |
| 21         | Femmes | DIR2  | 20       | 100%      | 14          | 70%         | 1.33%   | 4.29%      |
| 21         | Hommes | DIR2  | 20       | 100%      | 6           | 30%         | 1.29%   | 5.64%      |
| 21         | Total  | DIR2  | 20       | 100%      | 10          | 50%         | 1.31%   | 4.08%      |
| 22.3       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 0.87%   | 0.72%      |
| 22.3       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.69%   | 0.72%      |
| 22.3       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.78%   | 0.72%      |
| 24.5       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 3.31%   | 3.16%      |
| 24.5       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.52%   | 0.65%      |
| 24.5       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.60%   | 0.57%      |
| 26         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 1.77%   | 2.17%      |
| 26         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.39%   | 1.18%      |
| 26         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.44%   | 0.87%      |
| 27         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.70%   | 2.71%      |
| 27         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.05%   | 0.62%      |
| 27         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.47%   | 0.54%      |
| 28         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 0.60%   | 0.87%      |
| 28         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.64%   | 0.74%      |
| 28         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.49%   | 0.61%      |
| 29.3       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 1.51%   | 1.21%      |
| 29.3       | Hommes | DIR2  | 20       | 100%      | 15          | 75%         | 0.87%   | 2.84%      |
| 29.3       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 0.88%   | 1.15%      |
| 31.3       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.54%   | 2.59%      |
| 31.3       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.86%   | 0.61%      |
| 31.3       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 1.09%   | 0.93%      |
| 35         | Femmes | CAS2  | 4        | 100%      | 3           | 75%         | 1.19%   | 2.84%      |
| 35         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.11%   | 0.21%      |
| 35         | Total  | DIR2  | 18       | 90%       | 20          | 100%        | 1.51%   | 0.01%      |
| 36.9       | Femmes | CAS2  | 3        | 75%       | 0           | 0%          | 5.71%   | 24.23%     |
| 36.9       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.13%   | 0.73%      |
| 36.9       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.47%   | 2.00%      |

| NOGA OFS50 | Sexe   | Valid | #val tri | %val trim | #val tri bm | %val tri bm | EAM (%) | EAM bm (%) |
|------------|--------|-------|----------|-----------|-------------|-------------|---------|------------|
| 41.2       | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 3.72%   | 5.25%      |
| 41.2       | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.59%   | 1.00%      |
| 41.2       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.78%   | 1.20%      |
| 43         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 3.22%   | 2.54%      |
| 43         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.49%   | 0.49%      |
| 43         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.67%   | 0.50%      |
| 45         | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 4.39%   | 4.57%      |
| 45         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.64%   | 0.73%      |
| 45         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 1.01%   | 0.59%      |
| 46         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.64%   | 1.54%      |
| 46         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.42%   | 1.41%      |
| 46         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.26%   | 0.49%      |
| 47         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.44%   | 0.51%      |
| 47         | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 5.42%   | 4.99%      |
| 47         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.38%   | 0.28%      |
| 49         | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 9.17%   | 8.64%      |
| 49         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.51%   | 0.54%      |
| 49         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.87%   | 0.84%      |
| 50.1       | Femmes | CAS2  | 4        | 100%      | 2           | 50%         | 1.70%   | 6.86%      |
| 50.1       | Hommes | CAS2  | 4        | 100%      | 3           | 75%         | 2.48%   | 2.78%      |
| 50.1       | Total  | CAS2  | 4        | 100%      | 3           | 75%         | 2.01%   | 4.49%      |
| 52         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.84%   | 1.62%      |
| 52         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.51%   | 0.77%      |
| 52         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.57%   | 0.95%      |
| 53         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 0.23%   | 2.77%      |
| 53         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 0.78%   | 1.09%      |
| 53         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 0.54%   | 1.38%      |
| 55         | Femmes | DIR2  | 20       | 100%      | 19          | 95%         | 1.79%   | 1.96%      |
| 55         | Hommes | DIR2  | 20       | 100%      | 19          | 95%         | 1.08%   | 3.19%      |
| 55         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 1.29%   | 1.12%      |
| 56         | Femmes | CAS2  | 4        | 100%      | 2           | 50%         | 4.14%   | 5.98%      |
| 56         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 0.64%   | 3.29%      |
| 56         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.93%   | 1.71%      |
| 58.6       | Femmes | CAS2  | 3        | 75%       | 1           | 25%         | 5.25%   | 9.47%      |
| 58.6       | Hommes | DIR2  | 17       | 85%       | 18          | 90%         | 5.23%   | 5.12%      |
| 58.6       | Total  | CAS2  | 3        | 75%       | 2           | 50%         | 4.48%   | 6.36%      |
| 61         | Femmes | CAS2  | 4        | 100%      | 2           | 50%         | 0.80%   | 5.55%      |
| 61         | Hommes | CAS2  | 4        | 100%      | 3           | 75%         | 2.04%   | 2.86%      |
| 61         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.72%   | 1.01%      |
| 62.3       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.55%   | 2.30%      |
| 62.3       | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 0.68%   | 0.78%      |
| 62.3       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.02%   | 1.11%      |
| 64         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.70%   | 0.41%      |
| 64         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.03%   | 1.47%      |
| 64         | Total  | DIR2  | 19       | 95%       | 19          | 95%         | 0.89%   | 0.85%      |

| NOGA OFS50 | Sexe   | Valid | #val tri | %val trim | #val tri bm | %val tri bm | EAM (%) | EAM bm (%) |
|------------|--------|-------|----------|-----------|-------------|-------------|---------|------------|
| 65         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.40%   | 1.74%      |
| 65         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.48%   | 0.41%      |
| 65         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.32%   | 0.93%      |
| 66         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.99%   | 0.73%      |
| 66         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.96%   | 1.93%      |
| 66         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 3.60%   | 3.76%      |
| 68         | Femmes | CAS2  | 4        | 100%      | 3           | 75%         | 3.66%   | 4.42%      |
| 68         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 3.96%   | 2.26%      |
| 68         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 3.81%   | 3.38%      |
| 69         | Femmes | CAS2  | 3        | 75%       | 4           | 100%        | 4.73%   | 3.83%      |
| 69         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 4.14%   | 3.41%      |
| 69         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.31%   | 1.20%      |
| 70         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 0.56%   | 1.66%      |
| 70         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.49%   | 0.91%      |
| 70         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 0.76%   | 0.64%      |
| 71         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 3.14%   | 2.92%      |
| 71         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.78%   | 0.64%      |
| 71         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.91%   | 0.80%      |
| 72         | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.61%   | 1.37%      |
| 72         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 1.47%   | 1.93%      |
| 72         | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 2.29%   | 2.04%      |
| 73.5       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.18%   | 1.70%      |
| 73.5       | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 3.01%   | 2.88%      |
| 73.5       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.94%   | 2.03%      |
| 78         | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 4.12%   | 4.69%      |
| 78         | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 9.05%   | 10.25%     |
| 78         | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 7.62%   | 8.64%      |
| 79.2       | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 10.36%  | 8.95%      |
| 79.2       | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 2.08%   | 3.20%      |
| 79.2       | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.75%   | 0.47%      |
| 84         | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 5.14%   | 4.29%      |
| 84         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.69%   | 0.25%      |
| 84         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.42%   | 0.00%      |
| 85         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.22%   | 0.99%      |
| 85         | Hommes | DIR2  | 20       | 100%      | 20          | 100%        | 0.76%   | 1.55%      |
| 85         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.39%   | 0.02%      |
| 86         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 1.21%   | 1.03%      |
| 86         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 2.18%   | 2.61%      |
| 86         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 1.48%   | 0.63%      |
| 87         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 0.28%   | 0.29%      |
| 87         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 1.89%   | 1.71%      |
| 87         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.66%   | 0.74%      |
| 88         | Femmes | DIR2  | 20       | 100%      | 20          | 100%        | 1.12%   | 1.13%      |
| 88         | Hommes | CAS2  | 4        | 100%      | 4           | 100%        | 0.88%   | 1.26%      |
| 88         | Total  | DIR2  | 20       | 100%      | 20          | 100%        | 0.97%   | 1.08%      |
| 90.3       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 2.52%   | 2.45%      |
| 90.3       | Hommes | CAS2  | 4        | 100%      | 3           | 75%         | 0.66%   | 3.48%      |
| 90.3       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 1.50%   | 1.39%      |
| 94.6       | Femmes | CAS2  | 4        | 100%      | 4           | 100%        | 1.25%   | 3.13%      |
| 94.6       | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 8.03%   | 14.41%     |
| 94.6       | Total  | CAS2  | 4        | 100%      | 4           | 100%        | 3.17%   | 3.31%      |

La table suivante donne les écarts relatifs moyens spécifiquement pour les trimestres non-validés. Elle est l'équivalent pour les NOGA\_OFS50 de la Table 3 de la section 6.2.1.

| Noga_OFS50 | Sexe   | Valid | #val tri | %val trim | #val tri bm | %val tri bm | EAM (%) | EAM bm (%) | EAM.n.val (%) | EAM.n.val bm (%) |
|------------|--------|-------|----------|-----------|-------------|-------------|---------|------------|---------------|------------------|
| 13.5       | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 6.28%   | 6.20%      | 6.83%         | 6.80%            |
| 16.8       | Femmes | CAS2  | 3        | 75%       | 4           | 100%        | 5.49%   | 5.39%      | 6.12%         | -                |
| 16.8       | Total  | DIR2  | 12       | 60%       | 14          | 70%         | 3.26%   | 3.15%      | 4.06%         | 3.79%            |
| 35         | Total  | DIR2  | 18       | 90%       | 20          | 100%        | 1.51%   | 0.01%      | 2.33%         | -                |
| 36.9       | Femmes | CAS2  | 3        | 75%       | 0           | 0%          | 5.71%   | 24.23%     | 9.05%         | 24.23%           |
| 41.2       | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 3.72%   | 5.25%      | 8.80%         | 7.45%            |
| 45         | Femmes | CAS2  | 3        | 75%       | 2           | 50%         | 4.39%   | 4.57%      | 6.24%         | 6.45%            |
| 47         | Hommes | CAS2  | 3        | 75%       | 3           | 75%         | 5.42%   | 4.99%      | 8.66%         | 7.58%            |
| 49         | Femmes | CAS2  | 1        | 25%       | 1           | 25%         | 9.17%   | 8.64%      | 10.36%        | 9.68%            |
| 58.6       | Femmes | CAS2  | 3        | 75%       | 1           | 25%         | 5.25%   | 9.47%      | 6.03%         | 10.62%           |
| 58.6       | Hommes | DIR2  | 17       | 85%       | 18          | 90%         | 5.23%   | 5.12%      | 7.33%         | 7.12%            |
| 58.6       | Total  | CAS2  | 3        | 75%       | 2           | 50%         | 4.48%   | 6.36%      | 6.81%         | 9.04%            |
| 64         | Total  | DIR2  | 19       | 95%       | 19          | 95%         | 0.89%   | 0.85%      | 1.40%         | 1.36%            |
| 69         | Femmes | CAS2  | 3        | 75%       | 4           | 100%        | 4.73%   | 3.83%      | 6.11%         | -                |
| 78         | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 4.12%   | 4.69%      | 8.42%         | 8.86%            |
| 78         | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 9.05%   | 10.25%     | 11.40%        | 13.00%           |
| 78         | Total  | CAS2  | 1        | 25%       | 1           | 25%         | 7.62%   | 8.64%      | 9.53%         | 10.76%           |
| 79.2       | Femmes | CAS2  | 0        | 0%        | 0           | 0%          | 10.36%  | 8.95%      | 10.36%        | 8.95%            |
| 84         | Femmes | CAS2  | 3        | 75%       | 3           | 75%         | 5.14%   | 4.29%      | 10.57%        | 9.12%            |
| 94.6       | Hommes | CAS2  | 1        | 25%       | 1           | 25%         | 8.03%   | 14.41%     | 9.93%         | 17.24%           |

## Annexe 10 : SAE-BESTA. Estimation de variance

Description de l'estimation de la variance des estimateurs SAE par la méthode du bootstrap. Cette annexe contient aussi les scripts « R » et l'automatisation des processus d'analyse des données.

### Résumé

Dans ce travail, on s'intéresse à l'estimation de la variance des estimateurs petits domaines (SAE) et on propose d'utiliser la technique du "bootstrap" pour arriver à cette fin. Tous les deux ans, environ 18'000 entreprises sont tirées selon un plan de Poisson dans un cadre de sondage d'environ 600'000 entreprises, en respectant un système de coordination permettant de répartir la charge d'enquête. Une fois qu'une entreprise est tirée, tous les établissements appartenant à cette entreprise sont sélectionnés (tirage par grappes), soit environ 66'000 établissements. Les principales variables d'intérêt sont : le nombre d'employés (EMP) et les équivalents plein temps (EPT), pour les femmes, les hommes et le total. Les résultats sont publiés trimestriellement au niveau des établissements pour différentes ventilations. Les méthodes SAE sont appliquées pour l'estimation du nombre d'employés, principalement pour les cantons et les domaines d'activité NOGA2. Concrètement, des modèles linéaires à effets mixtes (LMM) ont été utilisés pour prédire le nombre d'employés pour la partie non échantillonnée de la population. Les mêmes modèles permettent à la fois des prédictions des ventilations géographiques et économiques. Afin de tenir compte de la complexité du plan de sondage et des estimateurs, la technique du bootstrap a été employée pour estimer la variance (sous le plan) des estimations SAE. En pratique, nous avons adapté l'algorithme du bootstrap pour les plans à probabilités inégales proposé par [Chauvet \(2007\)](#), qui consiste à générer une pseudo-population à partir de l'échantillon de départ, dans laquelle des échantillons bootstrap sont tirés selon le même plan de sondage que celui utilisé pour effectuer le tirage de départ. La méthode du bootstrap nous a permis de calculer les coefficients de variations (CV) des estimations SAE et de les comparer avec les CV des estimations directes, calculés par des formules analytiques, quand la taille de l'échantillon est suffisamment grande. Ceci a permis de mesurer le gain potentiel apporté par la méthode SAE. En termes de réduction de variance, les résultats sont très encourageants, sauf pour quelques NOGA2 qui nécessitent une analyse plus approfondie et spécifique. En effet, les CV des estimations SAE sont en moyenne 4 fois plus petits que les CV des estimations directes dans le cas des cantons. Cette moyenne est de 11 fois dans le cas des NOGA2.

# 1 Introduction

Ce travail s'inscrit dans le cadre du projet pilote d'innovation "SAE-BESTA : Estimations cantonales par la méthode des petits domaines appliquée à la Statistique de l'emploi (BESTA)" (OFS, 2019a,b).

Tous les deux ans, environ 18'000 entreprises sont tirées dans un cadre d'environ 600'000 entreprises), en respectant un système de coordination permettant de répartir la charge d'enquête. Lorsqu'une entreprise est tirée, tous les établissements de cette entreprise sont sélectionnés (tirage par grappes), soit environ 66'000 établissements. Le tirage au niveau des entreprises se fait sous un plan de Poisson, et les résultats sont publiés au niveau des établissements. Certaines grandes entreprises sont sélectionnées avec une probabilité égale à 1.

Le but du projet SAE-BESTA est d'estimer le total des emplois (EMP, représenté dans le code par les variables : `tot_m` pour les hommes, `tot_f` pour les femmes et `tot_t` pour le total) selon différentes ventilations, qu'il n'est pas possible d'atteindre avec suffisamment de précision avec les méthodes traditionnelles. La fréquence de publication des résultats de cette enquête est trimestrielle et les variables d'intérêt sont actuellement exploitées pour deux types de ventilations, géographique et domaine d'activité économique. Pour le premier type, les résultats sont publiés au niveau des 7 grandes régions (GR) qui sont Région lémanique, Espace Mittelland, Nordostschweiz, Zürich, Ostschweiz, Zentralschweiz et Tessin. Les estimations n'étant pas disponibles au niveau des cantons et des grandes villes, 4 cantons (GE, NE, VD et SG) et deux villes (Zürich et Winterthur) financent des compléments d'échantillons pour leurs propres besoins (pour la ville de Winterthur, ceci n'est valable qu'à partir du premier trimestre 2018). Pour le deuxième type, c.à.d le domaine d'activité, les résultats sont publiés au niveau de 18 sections, et 48 NOGA OFS50.

Nous disposons des données trimestrielles (2015/2 à 2019/4) de l'enquête des emplois STATEM, des registres (utilisées comme cadres d'extrapolation), et des données de la statistique d'entreprises STATENT (au 4<sup>ème</sup> trimestre de chaque année). Les poids de sondage ont été corrigés pour la non-réponse (variable `poidsCorrNR`) et un calage a été effectué avec une robustification pour produire des poids après calage  $w_i$  (variable `poidsFin`).

Les estimations diffusées des totaux des variables d'intérêt sont effectuées en utilisant un estimateur direct (estimateur par expansion). Pour un domaine  $d$  donné, cet estimateur s'écrit comme :

$$\hat{Y}_d^{\text{DIR}} = \sum_{k \in s_d} y_{ik} \times w_i, \quad (1)$$

où  $y_{ik}$  est le nombre d'emplois de l'établissement  $k$  de l'entreprise  $i$ , et  $s_d$  est la partie de l'échantillon correspondant aux établissements appartenant au domaine  $d$ .

Un autre estimateur direct a été considéré dans le projet SAE-BESTA. C'est l'estimateur par le ratio définit par :

$$\hat{Y}_d^{\text{DIR2}} = \frac{X_d}{\hat{X}_d} \times \hat{Y}_d^{\text{DIR}}, \quad (2)$$

où  $\hat{X}_d$  est l'estimateur par expansion du total des employés du cadre d'extrapolation dans le domaine  $d$  :

$$\hat{X}_d = \sum_{k \in s_d} x_{ik} \times w_i, \quad (3)$$

et  $X_d$  correspond au total sur l'entier du cadre au sein du domaine  $d$ . La variable auxiliaire  $x$  utilisée dans cet estimateur est le nombre d'employés dans le cadre, c.à.d. la variable BETOT. L'estimation de la variance pour des domaines non-calés revient à considérer l'estimateur DIR2 qui est ainsi utilisé dans les comparaisons dans ce rapport.

Comme le plan de sondage n'était pas prévu pour des domaines aussi petits, les estimations directes ne satisfont pas toujours les critères de précision pour la publication. Afin d'obtenir une meilleure précision, la méthode "estimation sur des petits domaines" (Small Area Estimation (SAE), en anglais) a été employée. Cette méthode exploite les données disponibles sur tous les domaines en même temps. Concrètement, des modèles linéaires mixtes (LMM) ont été utilisés pour prédire le nombre d'employés dans la partie non échantillonnée de la population. Les mêmes modèles permettent à la fois des prédictions des ventilations géographiques et économiques. La principale variable utilisée dans la modélisation est le total des employés dans le cadre (variables : `BETOT_M` pour les hommes, `BETOT_F` pour les femmes et `BETOT_T` pour le total). Un modèle spécifique a été ajusté en fonction du sexe et du nombre d'employés dans les établissements (`BETOT_M` ou

BETOT\_F : 0, 1-25 et >25), ce qui donne six modèles au total, le total des employés étant la somme des employés hommes et femmes. Les grandes entreprises sélectionnées avec une probabilité égale à 1 sont exclues de la modélisation car elles constituent une sous-population spécifique entièrement recensée. Les principales variables auxiliaires utilisées dans le plan de sondage (taille et région) sont aussi prises en compte dans la modélisation. Chaque modèle contient des effets fixes pour le croisement entre la région et le secteur économique, ainsi qu'une pente aléatoire pour le canton et la NOGA2. Cela signifie que le même modèle a été utilisé pour les différents niveaux de ventilations. De plus, l'utilisation des poids lors de l'ajustement des modèles s'est avérée absolument nécessaire à cause de leurs hétérogénéité. Cette modélisation a été donc utilisée pour prédire la partie non échantillonnée  $\bar{s}_d$  de la population du domaine  $d$ . Les valeurs prédites  $\hat{y}_k$ ,  $k \in \bar{s}_d$  ou observées  $y_k$ ,  $k \in s_d$  sont ensuite agrégées par domaine pour obtenir l'estimateur petits domaines :

$$\hat{Y}_d^{\text{SAE}} = \sum_{k \in s_d} y_k + \sum_{k \in \bar{s}_d} \hat{y}_k. \quad (4)$$

L'évaluation de la précision des estimations obtenues par la méthode SAE reste un grand défi. Celle-ci est encore plus difficile lorsque des modèles complexes sont utilisés. Dans ce rapport, nous proposons d'utiliser la technique du bootstrap pour évaluer la précision de l'estimation. Nous allons dans un premier temps présenter le principe général de cette technique et de son utilisation dans la théorie de l'échantillonnage. Ensuite, nous allons présenter une adaptation de cette technique aux estimateurs SAE utilisés dans ce projet.

## 2 Estimation de la variance par la méthode du Bootstrap

Souvent, lorsque cela est possible, une linéarisation est utilisée pour estimer la variance des estimateurs complexes. Cependant, il n'y a pas d'outil à disposition qui traite le cas de figure des enquêtes par échantillonnage correctement et qui permet en même temps d'intégrer des effets aléatoires. De plus, la linéarisation ne permet pas, du moins de manière simple, d'obtenir de bonnes approximations de la variance pour des estimateurs issus de modèles linéaires robustes à cause de la complexité de la dérivation des fonctions d'influence. Un inconvénient est que la linéarisation doit être développée pour chaque nouveau modèle contrairement à la méthode du bootstrap, pour laquelle seulement une adaptation des codes est nécessaire. En revanche, l'estimation par le Bootstrap peut parfois être très gourmande en termes de temps de calcul selon la complexité des estimateurs utilisés.

### 2.1 La méthode du Bootstrap

La technique du bootstrap a été introduite par [Efron \(1979\)](#) dans le cadre d'une population infinie. L'idée de base de cette technique consiste à reproduire le mécanisme d'échantillonnage d'origine qui a généré l'échantillon  $S$ . Le bootstrap permet d'obtenir une approximation de la distribution d'un estimateur. Il permet donc d'obtenir une estimation consistante de la variance et des intervalles de confiance, y compris pour des estimateurs non linéaires tels que les quantiles, et de prendre en compte les corrélations entre les différentes estimations. Il permet aussi d'étudier certains aspects du biais d'un estimateur.

L'adaptation de la technique du bootstrap à la théorie du sondage et au cas d'une population finie est assez problématique, et fait l'objet d'une littérature abondante, voir par exemple [Davison & Hinkley \(1997\)](#), [Davison & Sardy \(2007\)](#), [Shao & Tu \(2012\)](#) et [Mashreghi et al. \(2016\)](#).

Pour un plan de sondage à probabilités égales, [Gross \(1980\)](#) suggère d'utiliser l'échantillon  $S$  afin de recréer une pseudo-population  $U^*$ , dans laquelle on ré-échantillonne selon un tirage aléatoire simple et sans remise. La méthode proposée par Gross est intuitivement proche du principe du bootstrap introduit par [Efron \(1979\)](#). Cependant, toutes les méthodes du bootstrap proposées par la suite présentent certaines difficultés pratiques. De plus, la généralisation à un plan de sondage complexe est délicate, en dehors de quelques plans de sondage particuliers.

[Chauvet \(2007\)](#) propose une généralisation du principe de la méthode de Gross pour un tirage à probabilités inégales, en s'appuyant sur le principe d'estimation de Horvitz-Thompson : une unité  $k$  de l'échantillon représente  $1/\pi_k$  unités de la population, où  $\pi_k$  est la probabilité de sondage de l'unité  $k$ . Cette généralisation peut donc s'appliquer par exemple à un plan de Poisson. Elle peut aussi être adaptée au cas d'un tirage à plusieurs niveaux, où les unités primaires vont être utilisées pour construire la pseudo-population de ces dernières unités primaires. Les échantillons du bootstrap des unités primaires/secondaires sont sélectionnés ensuite parmi les unités primaires/secondaires de la pseudo-population, respectivement, selon le même plan de sondage qui a généré l'échantillon de départ.

### 2.2 Estimation de la variance des estimateurs SAE

Afin de tenir compte de la complexité du plan de sondage et des estimateurs, la technique du bootstrap a été employée pour développer des procédures d'estimation de la variance (sous le plan) des estimations SAE. La méthode du bootstrap a été appliquée principalement parce que la linéarisation semblait ne pas être bien adaptée aux estimateurs SAE. Dans ce travail, nous avons adapté l'algorithme proposé dans ([Chauvet, 2007](#), p. 174) à la méthode SAE.

#### Algorithme : Estimation de variance des estimateurs SAE pour un plan de sondage à deux niveaux par la technique du bootstrap

Soit  $\pi_i$  la probabilité de sondage de l'unité primaire  $i$ , et soit  $S_i$  l'échantillon d'unités secondaires dans l'unité primaire  $i$ .

Etape 1 : Construction de la partie fixe  $u_f^*$  de la pseudo-population d'unités primaires (entreprises) et la partie fixe  $U_f^*$  de la pseudo-population d'unités secondaires (établissements).

- (a) Posons  $r_i = 1/\pi_i$  et soit  $\lfloor r_i \rfloor$  la partie entière de  $r_i$ .
- (b)  $U_f^*$  s'obtient en dupliquant chaque échantillon  $S_i$  correspondant à l'unité primaire  $i$ ,  $\lfloor r_i \rfloor$  fois. Les probabilités d'inclusions correspondantes sont aussi dupliquées  $\lfloor r_i \rfloor$  fois. Ainsi, une pseudo-population  $u_f^*$  d'unités primaires est aussi créée avec des probabilités de tirages correspondantes.

**Etape 2 :** Construction de la partie aléatoire  $u_a^*$  de la pseudo-population d'unités primaires et de la partie aléatoire  $U_a^*$  de la pseudo-population d'unités secondaires.

- On tire parmi l'ensemble d'unités primaires de l'échantillon de départ  $s$  une pseudo-population  $u_a^*$  d'unités primaires selon un plan de Poisson, avec probabilités d'inclusion :  $r_i - \lfloor r_i \rfloor$ .
- La pseudo-population  $U_a^*$  s'obtient en combinant les échantillons  $S_i$  des unités primaires se trouvant dans  $u_a^*$ .
- La pseudo-population d'unités secondaires  $U^*$  s'obtient en combinant la partie aléatoire  $U_a^*$  et la partie fixe  $U_f^*$ . De même, la pseudo-population  $u^*$  d'unités primaires s'obtient en combinant la partie aléatoire  $u_a^*$  et la partie fixe  $u_f^*$ .
- La sélection de la partie aléatoire est répétée  $C$  fois pour obtenir  $C$  pseudo-populations différentes  $U_c^*, c = 1, \dots, C$ .

**Etape 3 :** Sélection des sous-échantillons et estimation des paramètres.

Pour chacune des  $C$  pseudo-populations  $U_c^*, c = 1, \dots, C$ , on répète  $B$  fois les étapes (a) et (b) suivantes :

- On sélectionne un échantillon  $s_b^*$  de la pseudo-population d'unités primaires  $u^*$  en utilisant le même procédé qui a été utilisé dans le plan de sondage de base. L'échantillon du bootstrap d'unités secondaires  $S_b^*$  est obtenu en combinant les échantillons  $S_i$  correspondants aux unités primaires dans  $s_b^*$ . Donc, chaque échantillon  $S_i$  est répliqué autant de fois que l'unité principale  $i$  se trouve dans  $s_b^*$ .
- On calcule les estimations SAE,  $\hat{Y}_{d,c,b}^*, b = 1, \dots, B$ , pour chaque domaine  $d$ , en se basant sur l'échantillon  $S_b^*$  et la population originale  $U$ , selon l'équation :

$$\hat{Y}_{d,c,b}^* = \sum_{k \in s_d} y_k + \sum_{k \in \bar{s}_d^*} \hat{y}_k. \quad (5)$$

Ceci implique que la modélisation (ajustement des modèles) se fait pour chaque pseudo-échantillon sélectionné. A la fin des  $B$  répétitions, on calcule la variance des estimations SAE obtenues en se basant sur tous les échantillons tirés de la pseudo population  $U_c^*$  :

$$\hat{V}_{d,c}^* = \frac{1}{B-1} \sum_{b=1}^B \left( \hat{Y}_{d,c,b}^* - \hat{Y}_{d,c,\cdot}^* \right)^2, \quad (6)$$

$$\text{où } \hat{Y}_{d,c,\cdot}^* = B^{-1} \sum_{b=1}^B \hat{Y}_{d,c,b}^*.$$

Finalement, la variance de chaque estimation (basée sur toutes les pseudo-populations) est obtenue en prenant la moyenne des estimations de variance correspondants à chacune des  $C$  pseudo-populations :

$$\hat{V}_d^* = \frac{1}{C} \sum_{c=1}^C (\hat{V}_{d,c}^*). \quad (7)$$

La moyenne et la distribution des estimations du bootstrap peuvent aussi être calculées pour permettre d'évaluer le biais conditionnel dû au modèle et la moyenne de la somme des carrés des erreurs (MSE), ainsi qu'un intervalle de confiance (basé sur le bootstrap) pour chaque estimation. Dans ce rapport, nous nous intéressons à l'estimation de la variance.

Cet algorithme a été implémenté en langage "R" et permet de calculer les estimations de variance et les moyennes des estimations bootstrap pour toutes les ventilations en même temps : grandes régions, cantons et grandes villes (Zurich et Winterthur), sections, NOGA et NOGA2. Concernant les probabilités  $\pi_i$ , nous avons utilisé l'inverse des poids de sondage corrigés pour la non-réponse, c.à.d. la variable `poidsCorrNR`. L'utilisation des codes se fait à partir d'un code principal qui est présenté dans l'annexe 6.4.

**Remarque .** Les unités primaires de poids  $\pi_i = 1$  sont sélectionnées d'office dans chaque échantillon bootstrap mais elles ne sont pas utilisées dans la modélisation. Celles-ci ne contribuent donc pas à la variance des estimations.

Les estimations SAE du canton de Zurich ont été calculées à l'intérieur des boucles du bootstrap en sommant les estimations SAE de la ville de Zurich, la ville de Winterthur, et du reste du canton. Ceci permet de prendre en considération les corrélations entre ces deux villes et le reste du canton dans le calcul de la variance des estimations du canton. De même, pour prendre en compte les corrélations entre les estimations des hommes

et des femmes dans l'estimation de la variance du total des deux sexes, les estimations des deux sexes sont additionnées pour chaque échantillon bootstrap. La variance du total est calculée ensuite selon les équations (6) et (7) en utilisant ces sommes des estimations des deux sexes.

**Remarque .** La taille des pseudo-populations construites selon l'algorithme décrit ci-dessus sera, en moyenne, égale à la taille de la population de départ. De plus, les unités primaires sélectionnées dans les étapes 1 et 2 de l'algorithme sont toutes considérées comme des unités différentes dans la pseudo-population. C'est à dire que nous allons attribuer des identifiants différents à toutes les unités primaires/secondaires de la pseudo-population. Ainsi, puisque le tirage des échantillons bootstrap s'effectue selon le même plan de sondage de base, les échantillons bootstrap auront en moyenne, la même taille que l'échantillon de départ. En pratique, l'attribution des nouveaux identifiants a été optimisée en évitant les boucles 'for' pour que l'attribution des nouveaux identifiants aux unités secondaires se fasse dans un temps raisonnable.

Afin d'optimiser le temps de calcul, une modification du modèle *modlin1*, correspondant au modèle pour les établissements avec un nombre d'employés BETOT (hommes ou femmes) entre 1 et 25, a été adoptée. Cette modification consiste à supprimer les termes d'ordonnées à l'origine des effets aléatoires (en ajoutant des "-1") pour les cantons et les NOGA2 dans ce modèle comme suit.

```
modlin1 = lmer(y ~ BETOT*secteur_gr + BETOT*section_alt
 + (BETOT-1|cant) + (BETOT-1|noga2),
 weight=poidsFin, data=data1)
```

Ceci a permis de réduire le temps de calcul de 3.5 à 1.7 heures pour un trimestre donnée et un sexe donné, lorsque  $C = 10$  pseudo-populations sont créées et  $B = 100$  échantillons bootstrap sont tirés de chacune de ces pseudo-populations. Ce qui revient à dire que l'estimation de la variance se base sur  $B \times C = 1000$  échantillons bootstrap. Cette modification n'a été utilisée que dans le calcul de la variance. De plus, celle-ci n'affecte que de peu les estimations SAE comme on le montre dans la section 6.2 de l'annexe.

Après des premiers tests, nous avons remarqué qu'un établissement dans le canton d'Obwald qui, une fois sélectionné dans un échantillon bootstrap, induit un effet aléatoire très large pour ce canton dans le modèle 0 ajusté. Ceci a conduit à une variance trop grande pour ce canton dû à son poids de sondage élevé par rapport à son nombre d'employés élevé également.

Même si la sélection de cet établissement n'arrive que rarement, nous avons décidé d'éliminer cet établissement et toutes les entreprises avec un coefficient de robustification  $< 0.5$  de la modélisation. Celles-ci sont, bien évidemment, toujours considérées dans le calcul des variables d'intérêt. Cette élimination a permis de résoudre le problème en question.

En plus du calcul des estimations de précision des estimateurs SAE, nous calculons aussi des estimateurs SAE redressés (benchmarking) pour que les totaux correspondent aux estimations directes des domaines plus grands comme les sections et les grandes régions. Afin de garantir la cohérence entre les résultats des estimations SAE pour différents niveaux de ventilation, qui ne sont pas nécessairement imbriqués (géographique et activité économique), le benchmarking est réalisé en utilisant un calage (type raking ratio implémenté dans la fonction "calib" du logiciel "R"). La variance des estimateurs redressés est aussi estimée par la technique du bootstrap. Cependant, cette variance est conditionnelle (étant donné les estimations directes sur lesquelles on cale), et ne peut pas être utilisée telle quelle comme base de comparaison entre les estimations directes et les estimations SAE sans considérer les variances des estimations directes sur lesquelles on cale. Enfin, les codes implémentés en langage "R" permettent d'une manière automatisée le calcul des estimations directes, des estimations SAE, et des variances des estimations directes et SAE, ainsi que la comparaison des CV des estimations SAE à ceux des estimations directes, sous forme de tables et de graphiques.

### 3 Resultats

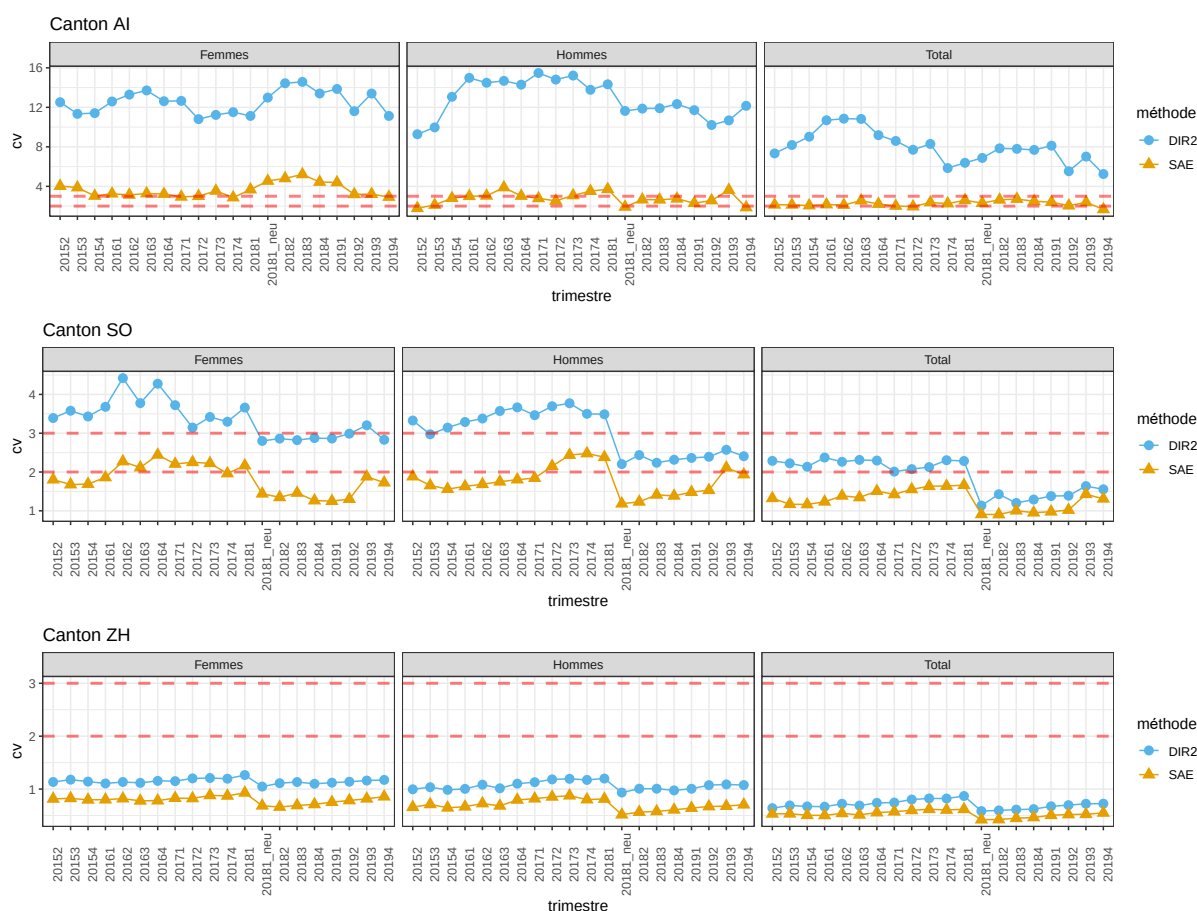


**Fig. 1:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes (DIR2) par sexe pour certains cantons (AI, SO et ZH).

La méthode du bootstrap nous a permis de calculer les coefficients de variations (CV) des estimations SAE et de les comparer avec les CV des estimations directes (DIR2), calculés par des formules analytiques, quand la taille de l'échantillon le permet. Ceci a permis de mesurer le gain potentiel apporté par la méthode SAE, et ceci pour toutes les ventilations : grandes régions, cantons, sections, NOGA et NOGA2.

Nous présentons dans l'annexe les résultats au niveau des cantons (figures 3 à 7) et des NOGA2 (figures 9 à 21), pour les hommes, les femmes, ainsi que pour le total. Les résultats pour trois des 26 cantons suisses sont présentés dans la figure 1 : un petit (AI), un moyen (SO) et un grand canton (ZH), pour les hommes, les femmes et le total. Pour tous ces résultats, le nombre de pseudo-populations créées est  $C = 10$  et le nombre d'échantillons du bootstrap sélectionnés de chaque pseudo-population est  $B = 100$ . Ce qui revient à dire que les variances sont calculées sur la base de 1000 échantillons bootstrap. Dans ces illustrations, nous comparons les coefficients de variation (CV) des estimations SAE avec ceux des estimations directes. En général (pour les 26 cantons), les CV des estimations SAE sont plus petits que les CV des estimations directes. Le gain en termes de CV est plus important pour les petits domaines (comme le canton AI, Figure 1, panel 1) par rapport à un grand domaine (comme ZH, Figure 1, panel 3). Les CV des estimations SAE sont en moyenne 4 fois plus petits que les CV des estimations directes dans le cas des cantons. Cette moyenne est de 11 fois pour le cas des NOGA2. Pour certains cantons, comme SO (Figure 1, panel 2), le CV de l'estimation SAE est inférieur à 3% pour les deux sexes et inférieur à 2% pour le total, et ceci étant valable pour tous les trimestres. Cet exemple montre que si les estimateurs SAE sont utilisés à la place des estimateurs directs, le nombre de cantons pour lesquels les critères de précision requis pour les grandes régions seraient respectés (c'est-à-dire 2% pour l'emploi total et 3% par sexe, indiqués par les lignes en tirets dans la figure 1), augmenterait considérablement. Cette observation est également valable dans le cas des NOGA2.

Concernant le problème d'instabilité de la variance de l'estimateur SAE du canton d'Obwald, nous présentons

dans la Figure 8 les CV de ce canton avant d'exclure les établissements avec un coefficient de robustification inférieur à 0.5 de la modélisation. En comparant les CV de ce canton avec ceux du même canton après avoir exclu les établissements avec un coefficient de robustification inférieur à 0.5 (voir Figure 5), nous pouvons affirmer que le problème d'instabilité de la variance de ce canton a été traité. Enfin, le canton AR présente aussi une certaine forme d'instabilité de la variance. Mais cette fois, l'estimateur direct pose aussi un problème.

## 4 Conclusions

Dans ce travail, nous avons proposé une procédure pour estimer la variance (sous le plan) des estimateurs SAE, basée sur la technique du bootstrap. Dans cette procédure, il était nécessaire de prendre en compte la complexité du plan de sondage et des estimateurs SAE utilisés. Néanmoins, la technique du bootstrap s'est révélée efficace et "facile" à utiliser. Les résultats sont très encourageants sauf pour certaines NOGA2 qui nécessitent une analyse plus approfondie et spécifique. Nous avons aussi proposé un benchmarking pour garantir la cohérence entre les résultats des estimations SAE avec les estimations directes calculées pour des domaines plus grands, et ceci pour différents niveaux de ventilation, qui ne sont pas nécessairement imbriqués (ventilation géographique versus ventilation par activité économique). Enfin, il serait potentiellement profitable d'étudier d'autres algorithmes du bootstrap que ceux proposés par [Chauvet \(2007\)](#) et leurs potentielles adaptations à l'estimation de variance des estimateurs SAE. Tous ces ingrédients réunis font de ce projet un produit "idéal" pour la statistique expérimentale.

## 5 Références

### Références

- Chauvet, G. (2007). *Méthodes de bootstrap en population finie*. Ph.D. thesis, ENSAI. (Cité pages 96, 99 et 104.)
- Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap methods and their application*. No. 1. Cambridge university press. (Cité page 99.)
- Davison, A. C. & Sardy, S. (2007). Resampling variance estimation in surveys with missing data. *Journal of Official Statistics* **23**, 371–386. (Cité page 99.)
- Efron, B. (1979). Bootstrap Methods : Another Look at the Jackknife. *The Annals of Statistics* **7**, 1 – 26. (Cité page 99.)
- Gross, S. (1980). Median estimation in sample surveys. In *Proceedings of the Section on Survey Research Methods*, vol. 1814184. American Statistical Association Alexandria, VA. (Cité page 99.)
- Mashreghi, Z., Haziza, D. & Léger, C. (2016). A survey of bootstrap methods in finite population sampling. *Statistics Surveys* **10**, 1 – 52. (Cité page 99.)
- OFS (2019a). *Rapport final Projet pilote BESTA NDS SAE BESTA WI*. (Cité page 97.)
- OFS (2019b). *SAE-BESTA : Evaluation du potentiel d'amélioration de la précision dans le cadre de la Statistique de l'emploi (BESTA)*. [www.experimental.bfs.admin.ch](http://www.experimental.bfs.admin.ch). (Cité page 97.)
- Shao, J. & Tu, D. (2012). *The jackknife and bootstrap*. Springer Science & Business Media. (Cité page 99.)

## 6 Annexes

### 6.1 Validation de la technique du bootstrap

En premier lieu, nous avons implémenté un bootstrap pour un plan de sondage de Poisson à un niveau (au niveau des établissements) et on a comparé les variances estimées par le bootstrap avec les variances estimées par des formules analytiques pour un plan de Poisson. Ceci nous a permis de valider l'approche adoptée dans ce travail. L'algorithme utilisé est le suivant :

#### Algorithme : Bootstrap pour un plan de sondage à un niveau.

Soit  $S$  l'échantillon de départ, et soit  $\pi_i$  la probabilité de sondage de l'unité  $i$ .

Etape 1 : Construction de la partie fixe  $U_f^*$  de la pseudo-population d'établissements.

- (a) Posons  $r_i = 1/\pi_i$  et soit  $\lfloor r_i \rfloor$  la partie entière de  $r_i$ .
- (b)  $U_f^*$  s'obtient en dupliquant chaque unité  $i$ ,  $\lfloor r_i \rfloor$  fois. Les probabilités d'inclusions correspondantes sont aussi dupliquées  $\lfloor r_i \rfloor$  fois.

Etape 2 : Construction de la partie aléatoire  $U_a^*$  de la pseudo-population.

- (a) On tire parmi l'ensemble d'unités de l'échantillon de départ  $S$ , une pseudo-population  $U_a^*$  d'unités (avec leur probabilité de sondage  $\pi_i$ ) selon un plan de Poisson, avec probabilités d'inclusion :  $r_i - \lfloor r_i \rfloor$ .
- (b) La pseudo-population  $U^*$  s'obtient en combinant la partie aléatoire  $U_a^*$  et la partie fixe  $U_f^*$ .
- (c) La sélection de la partie aléatoire est répétée  $C$  fois pour obtenir  $C$  pseudo-populations différentes  $U_c^*, c = 1, \dots, C$ .

Etape 3 : Sélection des sous-échantillons et estimation des paramètres. Pour chacune des  $C$  pseudo-populations  $U^*$ , on répète les étapes suivantes  $B$  fois :

- (a) On sélectionne un échantillon  $S_b^*$  de la pseudo-population  $U^*$  d'unités en utilisant un plan de Poisson (à un niveau).
- (b) On calcul les estimations directes,  $\hat{Y}_{d,b}^*$ , pour chaque domaine  $d$ , en se basant sur l'échantillon  $S_b^*$ , selon l'équation (1). A la fin des  $B$  répétitions, on calcule la variance de chaque estimation obtenue de la façon suivante :

$$\hat{V}_{d,c}^* = \frac{1}{B-1} \sum_{b=1}^B \left( \hat{Y}_{d,c,b}^* - \hat{Y}_{d,c,\cdot}^* \right)^2, \quad (8)$$

$$\text{où } \hat{Y}_{d,c,\cdot}^* = B^{-1} \sum_{b=1}^B \hat{Y}_{d,c,b}^*.$$

Finalement, la variance de chaque estimation (basée sur toutes les pseudo-populations) est obtenue en prenant la moyenne des estimations de variance correspondant à chacune des  $C$  pseudo-populations :

$$\hat{V}_d^* = \frac{1}{C} \sum_{c=1}^C (V_{d,c}^*), \quad (9)$$

La Table 1 présente les différentes estimations de variances, par des formules analytiques et par le bootstrap, des estimations directes au niveau des cantons pour le quatrième trimestre de l'année 2016 pour les femmes. On constate tout d'abord que la moyenne des estimations basées sur le bootstrap sont très proches des estimations directes, ce qui est normal puisque l'estimateur par expansion est sans biais. On remarque aussi que le rapport entre les variances estimées par des formules analytiques et les variances estimées par le bootstrap est proche de 1. De plus, les CV des estimations calculés par le bootstrap sont très proches de ceux estimés par les formules analytiques avec un léger biais positif. Ceci peut être considéré comme une validation de la technique du bootstrap pour le plan de sondage utilisé (Poisson). Il est à noter que les variances obtenues pour le plan de sondage à deux niveaux, présentées dans la section des résultats et estimées par l'algorithme de la section 2.2, ne sont pas très différentes des estimations de variance obtenues ici. Ceci s'explique par le fait que la majorité des entreprises (clusters) ne contiennent qu'un seul établissement, ce qui limite l'effet cluster considérablement.

**Tab. 1:** DIR : estimation par expansion,  $var(DIR)$  : estimation par formule analytique de la variance du DIR,  $var_b(DIR)$  : moyenne Bootstrap du DIR,  $var_b(DIR)$  : variance Bootstrap du DIR,  $var(DIR)/var_b(DIR)$  : rapport des deux variances à titre de comparaison. B = 100, C = 10, année = 2016, trimestre 4, sexe = F.

| cant | DIR       | $var(DIR)$   | $mean_b(DIR)$ | $var_b(DIR)$ | $cv_b(DIR)$ | $var/var_b$ |
|------|-----------|--------------|---------------|--------------|-------------|-------------|
| AG   | 152600.44 | 57403162.99  | 152542.90     | 53483985.53  | 4.79        | 1.07        |
| AI   | 4562.41   | 818514.59    | 4546.42       | 839684.94    | 20.08       | 0.97        |
| AR   | 12294.24  | 9756049.56   | 12313.94      | 10671827.26  | 26.57       | 0.91        |
| BE   | 290226.07 | 139389176.44 | 290321.13     | 119545116.39 | 3.77        | 1.17        |
| BL   | 59082.31  | 16634501.52  | 58829.21      | 17367414.06  | 7.05        | 0.96        |
| BS   | 91585.94  | 20578889.71  | 92044.40      | 20819936.50  | 4.98        | 0.99        |
| FR   | 62882.18  | 27462568.78  | 62766.87      | 26339788.07  | 8.16        | 1.04        |
| GE   | 150234.41 | 22242202.11  | 150273.12     | 19791057.01  | 2.96        | 1.12        |
| GL   | 12191.76  | 3652258.89   | 12322.95      | 3646238.61   | 15.66       | 1.00        |
| GR   | 51213.14  | 16297541.82  | 51336.71      | 15691354.26  | 7.73        | 1.04        |
| JU   | 15435.62  | 2318304.54   | 15405.83      | 2066417.57   | 9.31        | 1.12        |
| LU   | 113460.14 | 47226180.47  | 113286.43     | 43365304.67  | 5.80        | 1.09        |
| NE   | 46339.13  | 2673033.78   | 46313.92      | 2820239.84   | 3.62        | 0.95        |
| NW   | 9988.77   | 4972728.26   | 10001.90      | 4753135.91   | 21.83       | 1.05        |
| OW   | 10323.71  | 4367365.13   | 10155.36      | 4048043.91   | 19.49       | 1.08        |
| SG   | 129106.84 | 15897839.16  | 129186.58     | 16321308.70  | 3.13        | 0.97        |
| SH   | 19520.77  | 6755588.06   | 19480.86      | 6968782.00   | 13.52       | 0.97        |
| SO   | 65906.89  | 58511072.09  | 65764.18      | 65170785.74  | 12.25       | 0.90        |
| SZ   | 32309.72  | 14990450.95  | 32372.60      | 14531025.51  | 11.80       | 1.03        |
| TG   | 57682.55  | 16123680.46  | 57681.30      | 16652822.39  | 7.07        | 0.97        |
| TI   | 105112.49 | 262058169.42 | 105911.08     | 271875609.82 | 15.69       | 0.96        |
| UR   | 6073.53   | 1653669.91   | 6131.97       | 1621185.11   | 20.96       | 1.02        |
| VD   | 195621.88 | 46674212.61  | 195417.09     | 42408033.26  | 3.33        | 1.10        |
| VS   | 76120.61  | 30811154.26  | 75884.82      | 30741117.31  | 7.28        | 1.00        |
| ZG   | 55098.64  | 34716825.50  | 55063.44      | 35709420.93  | 10.85       | 0.97        |
| ZH   | 453926.70 | 123336712.69 | 454219.39     | 114485786.13 | 2.36        | 1.08        |
| zs   | 207686.55 | 28738089.43  | 207541.64     | 29212591.81  | 2.60        | 0.98        |
| ZHdj | 246240.15 | 94598623.27  | 246677.75     | 89588552.63  | 3.84        | 1.06        |

## 6.2 Comparaison des deux modèles : lent et rapide

Comme mentionné dans la section 2, nous avons modifié le modèle *modlin1* correspondant au nombre d'employés BETOT (hommes ou femmes) entre 1 et 25, en supprimant les ordonnées à l'origine des effets aléatoires correspondants aux cantons et aux NOGA2. Ce qui a réduit considérablement le temps de calcul. Ici, nous allons comparer les estimations SAE des deux variantes. Dans la figure 6.2 (panels de gauche), nous présentons la distribution (en forme de "Box-plot") des différences des estimations entre les deux variantes par ventilation, pour les hommes et pour les femmes, à travers tous les 20 trimestres étudiés (ce qui totalise  $2 \times 20 = 40$  différences par ventilation). Nous présentons aussi (dans les panels de droite de la même figure) les p-valeurs des 40 tests statistiques associés (t-tests de Student appariés). L'ensemble de ces comparaisons indique que les deux modèles (sans et avec modification) ne diffèrent pas de manière systématique et significative. Ceci suggère que cette modification peut être utilisée, non seulement pour l'estimation de la variance, mais aussi pour calculer les estimateurs SAE eux-mêmes, sans introduire un biais supplémentaire à l'estimation.

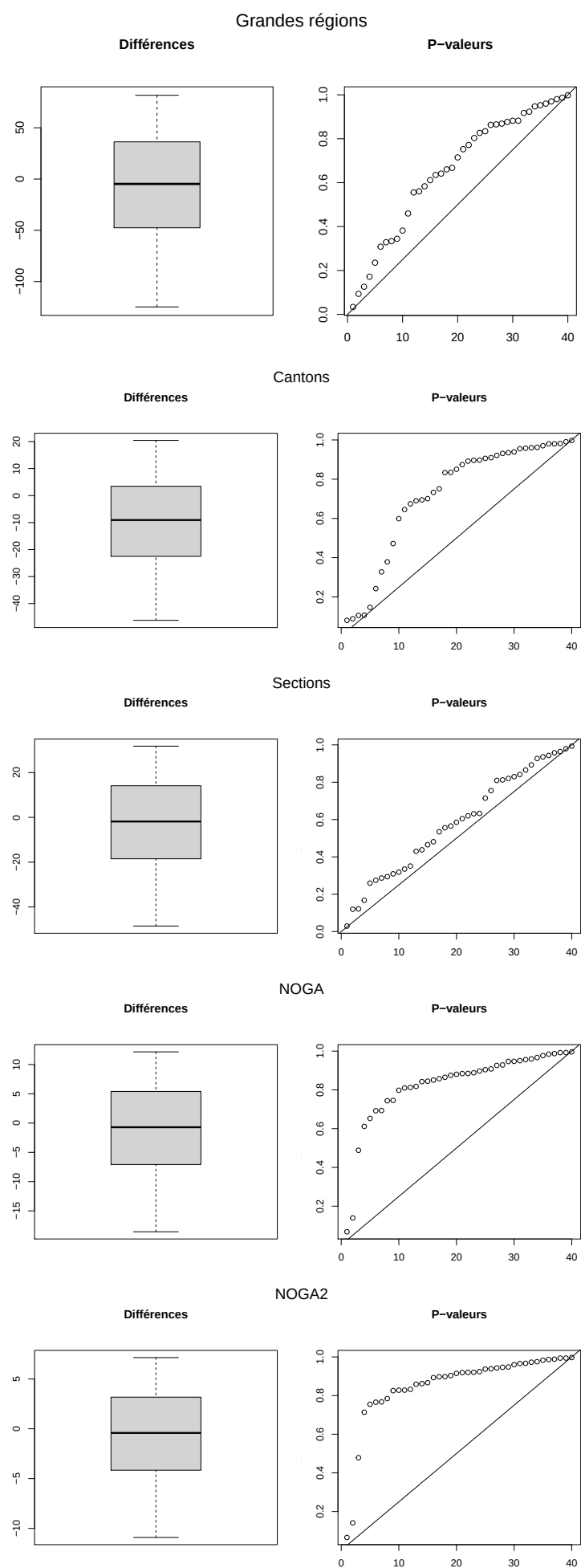


**Fig. 2:** Comparaison des estimateurs SAE sans et avec modification apportée au modèle 1.

## 6.3 Graphiques

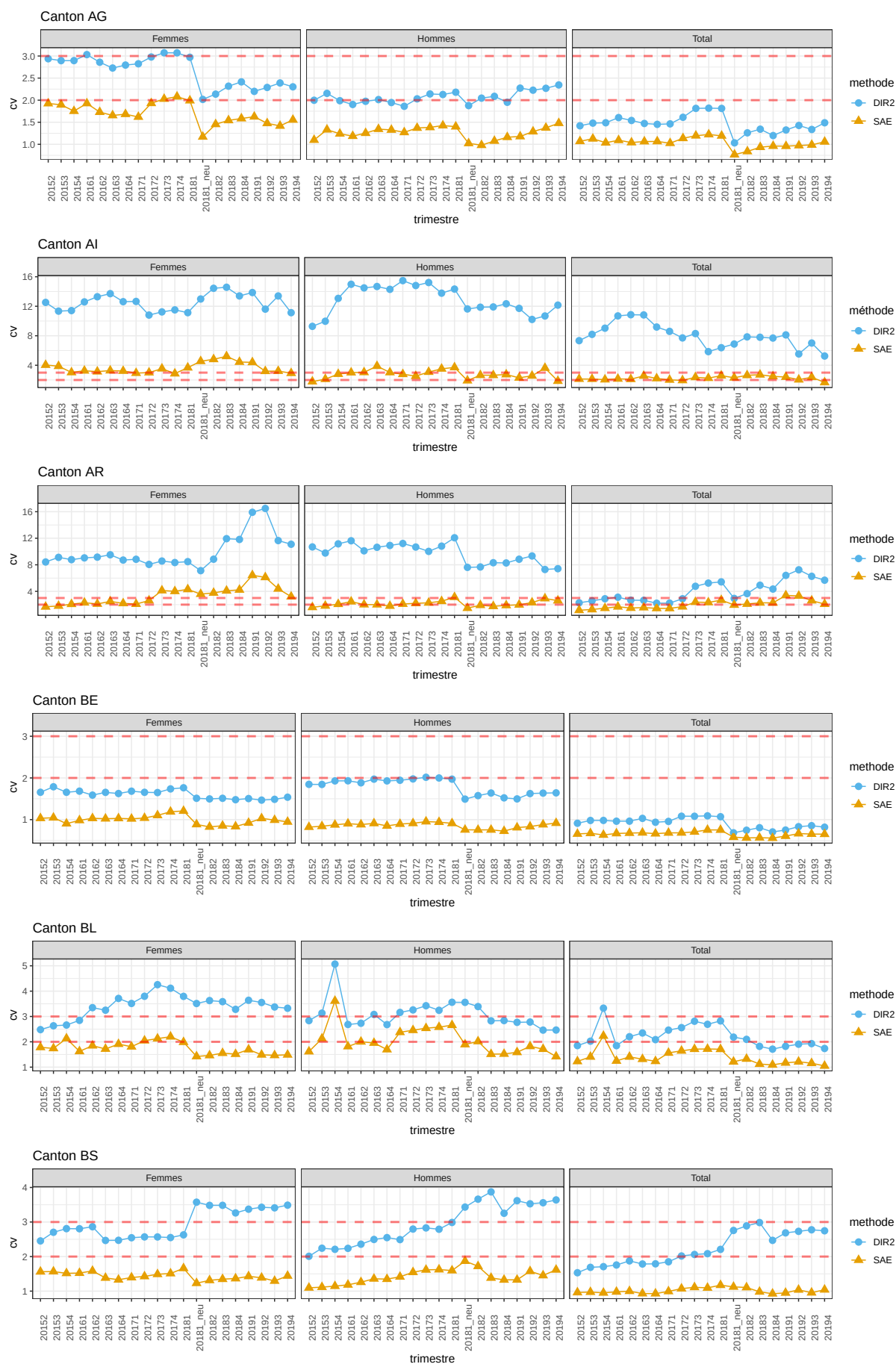


**Fig. 3:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par canton et par sexe.

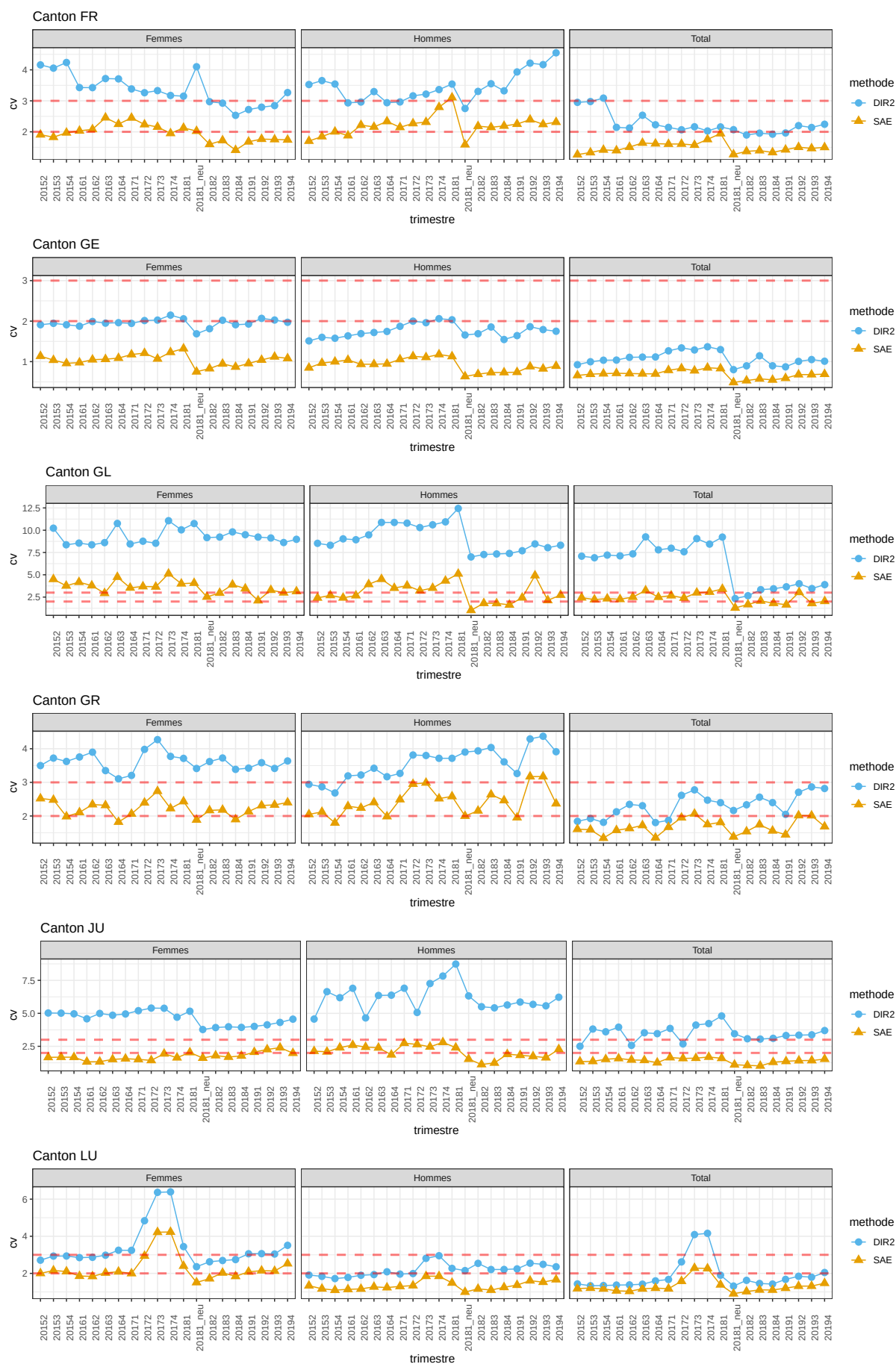


**Fig. 4:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par canton et par sexe (suite).

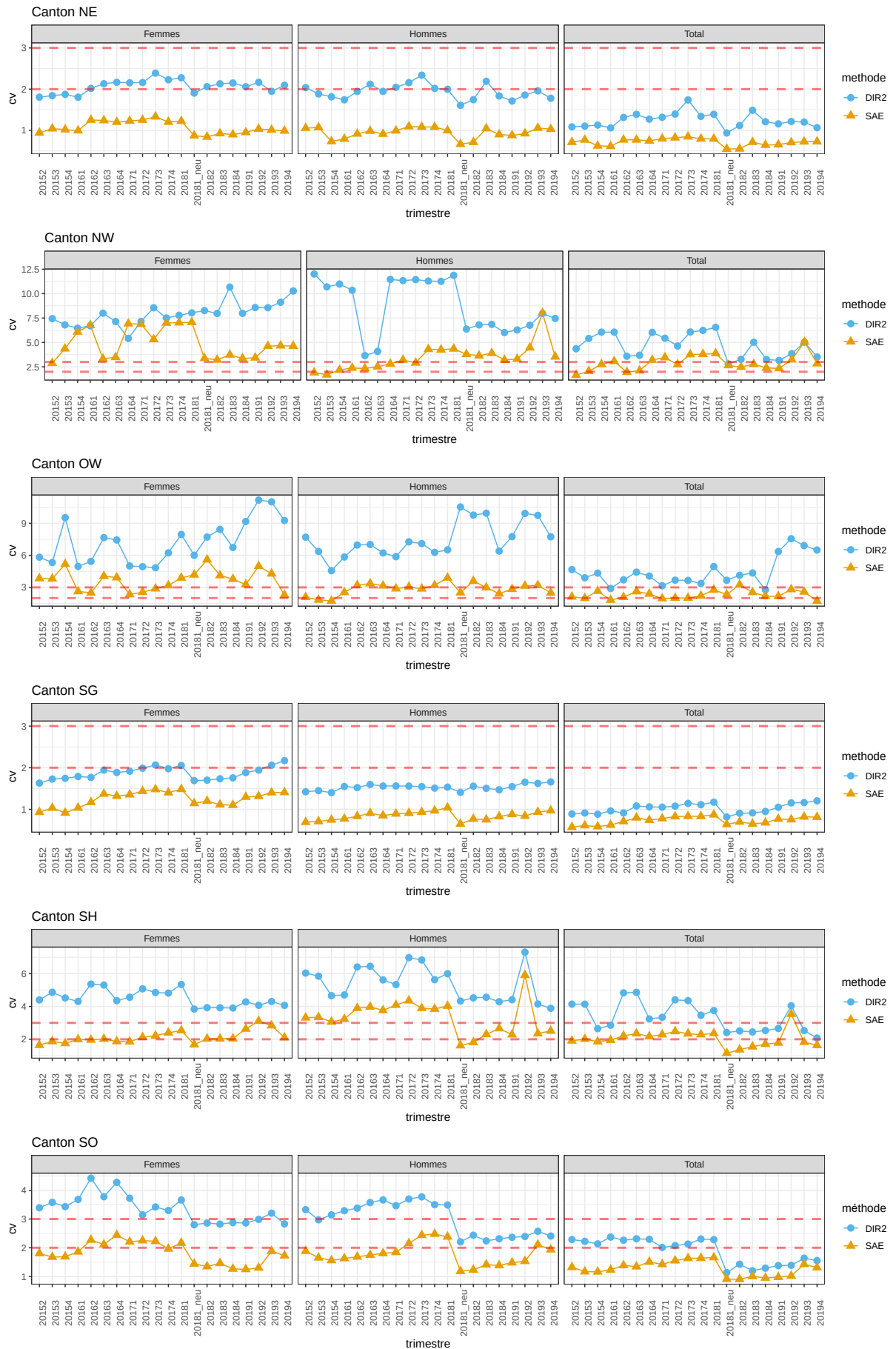


**Fig. 5:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par canton et par sexe (suite).

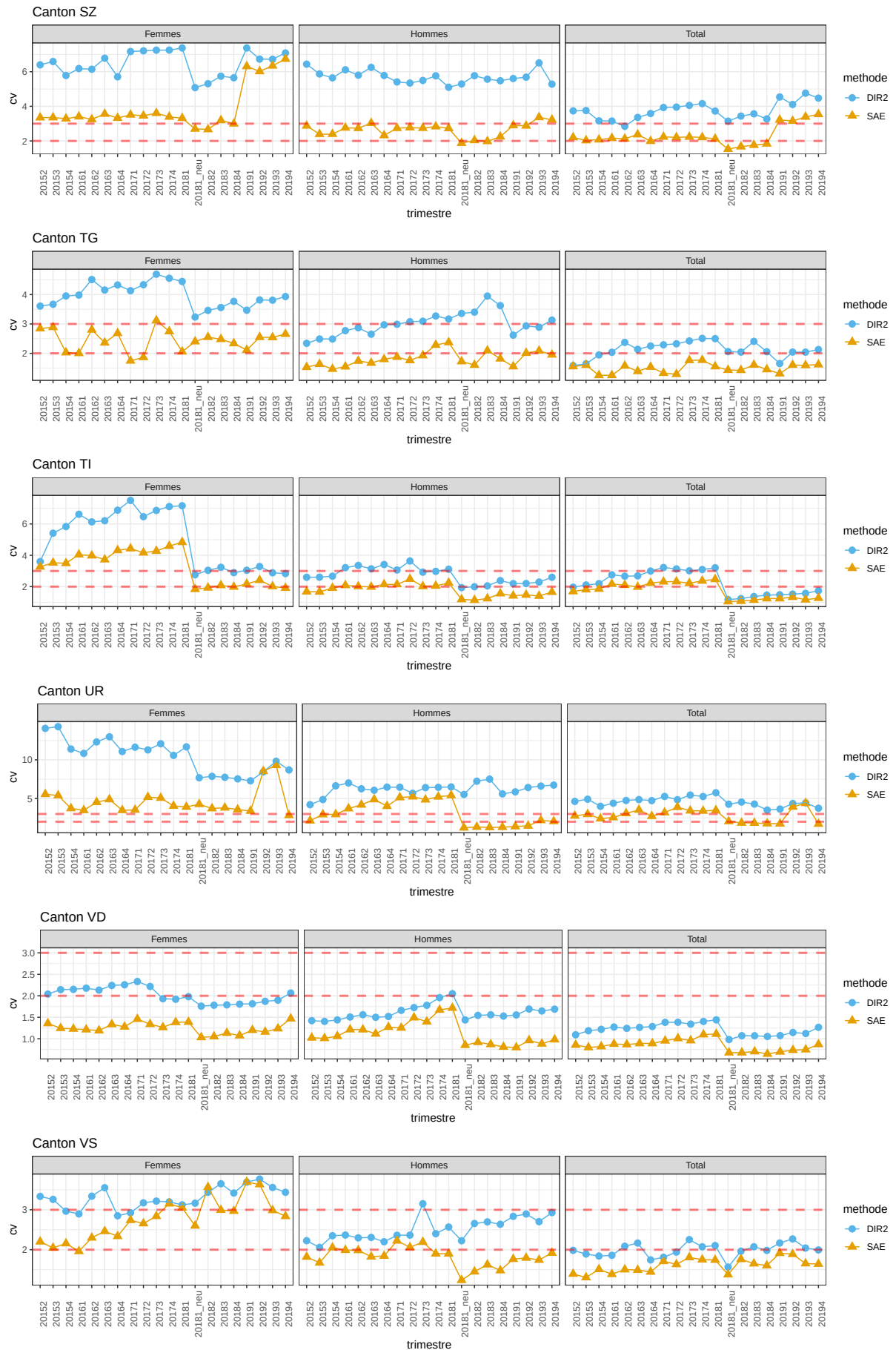


**Fig. 6:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par canton et par sexe (suite).



**Fig. 7:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par canton et par sexe (suite).

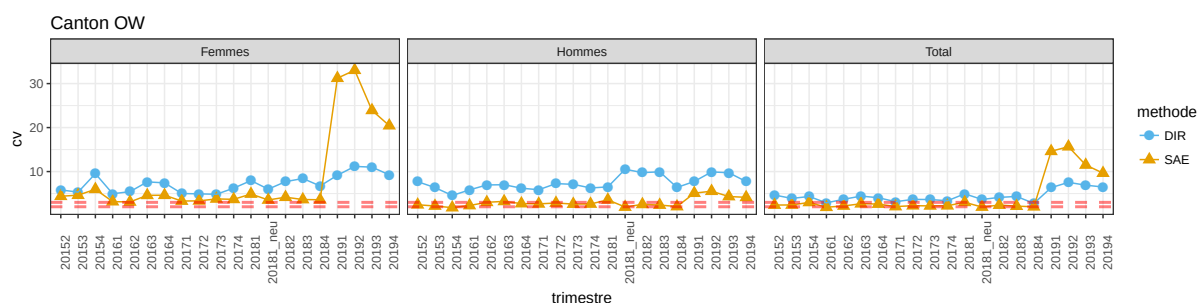


**Fig. 8:** Comparaison des CV des estimations SAE aux CV des estimations directes pour le canton de OW avant l'exclusion des établissements avec un coefficient de robustification inférieur à 0.5.



**Fig. 9:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe.

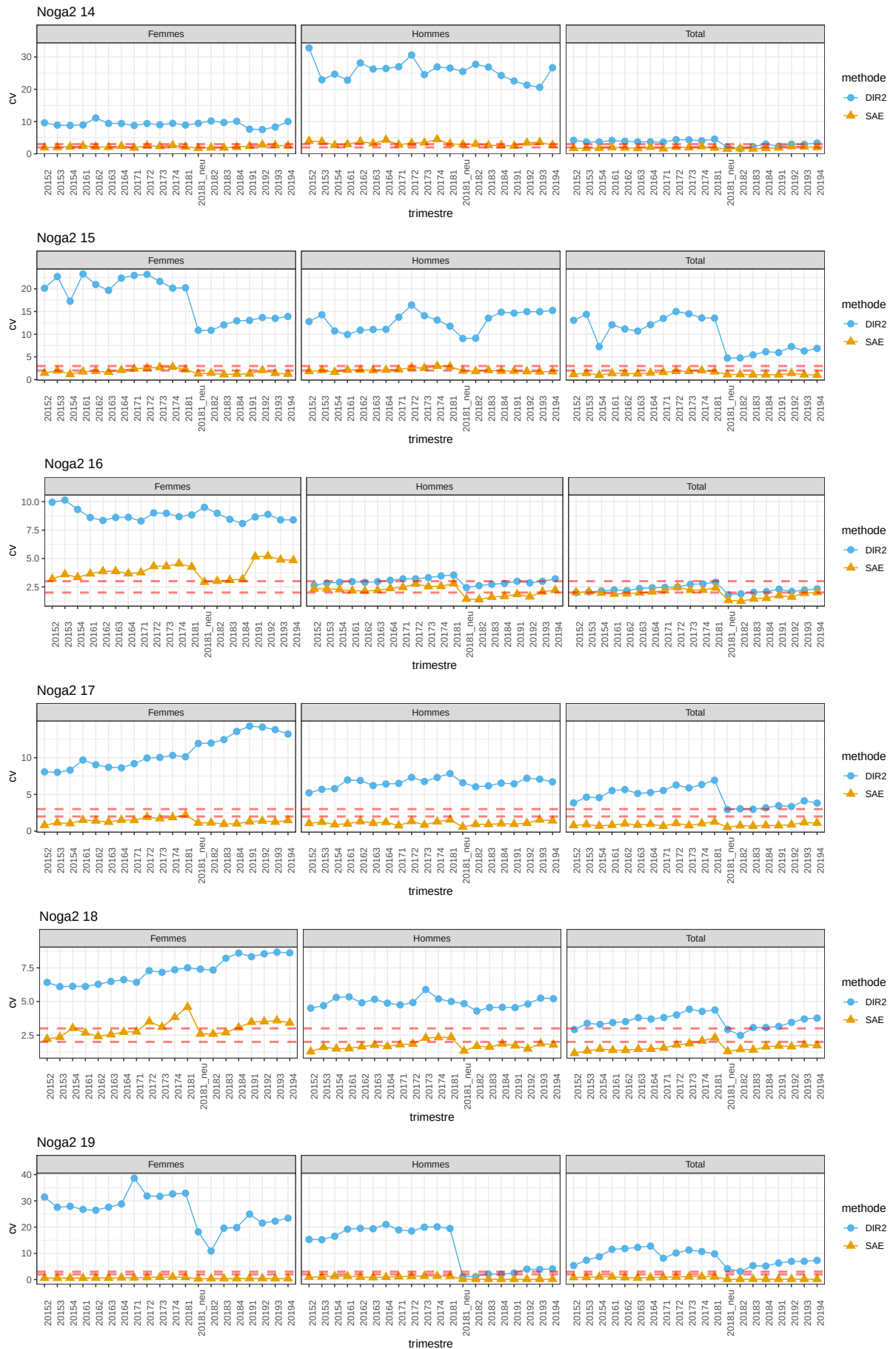


**Fig. 10:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).



**Fig. 11:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

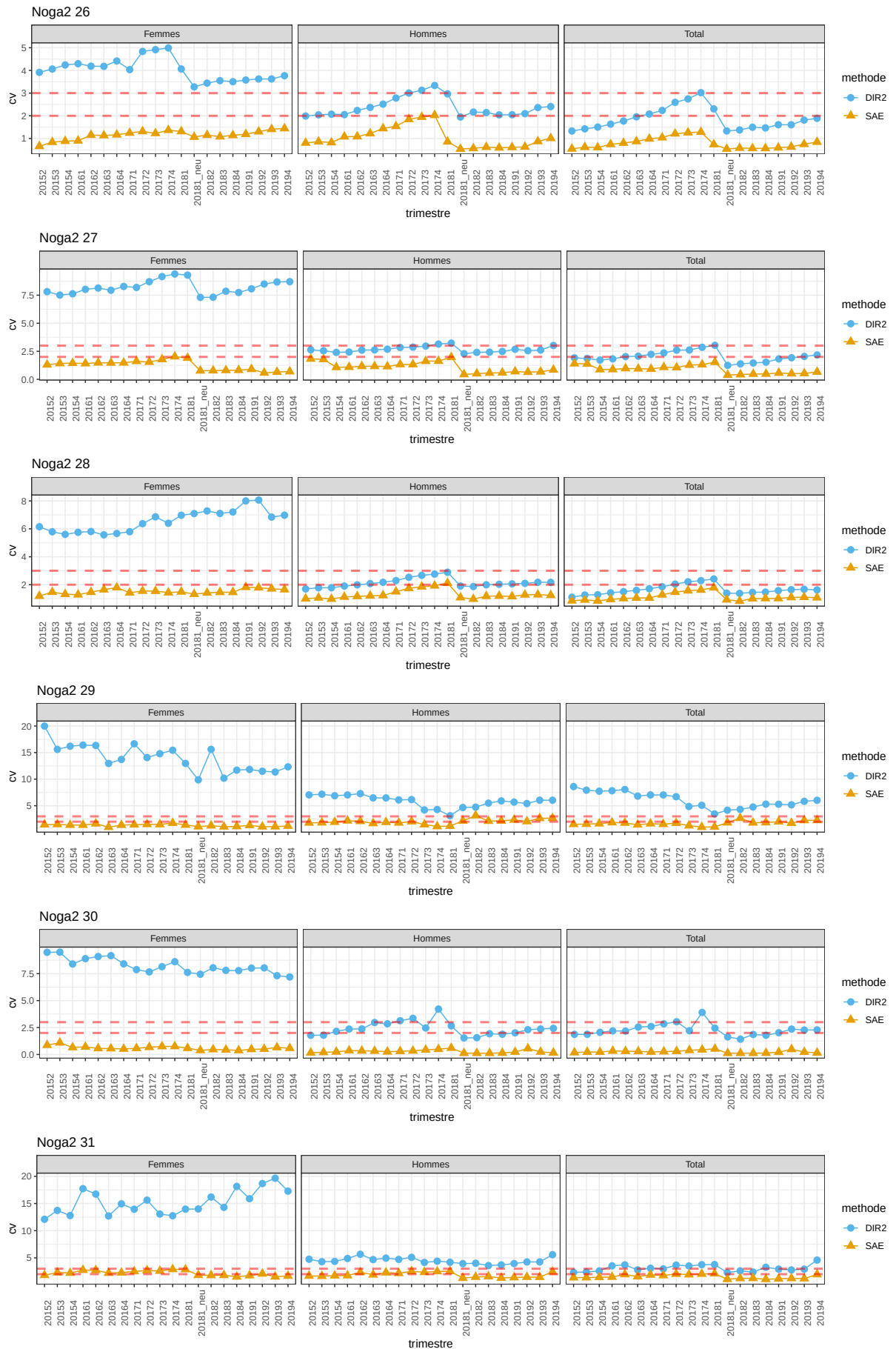


**Fig. 12:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).



**Fig. 13:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).



**Fig. 14:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

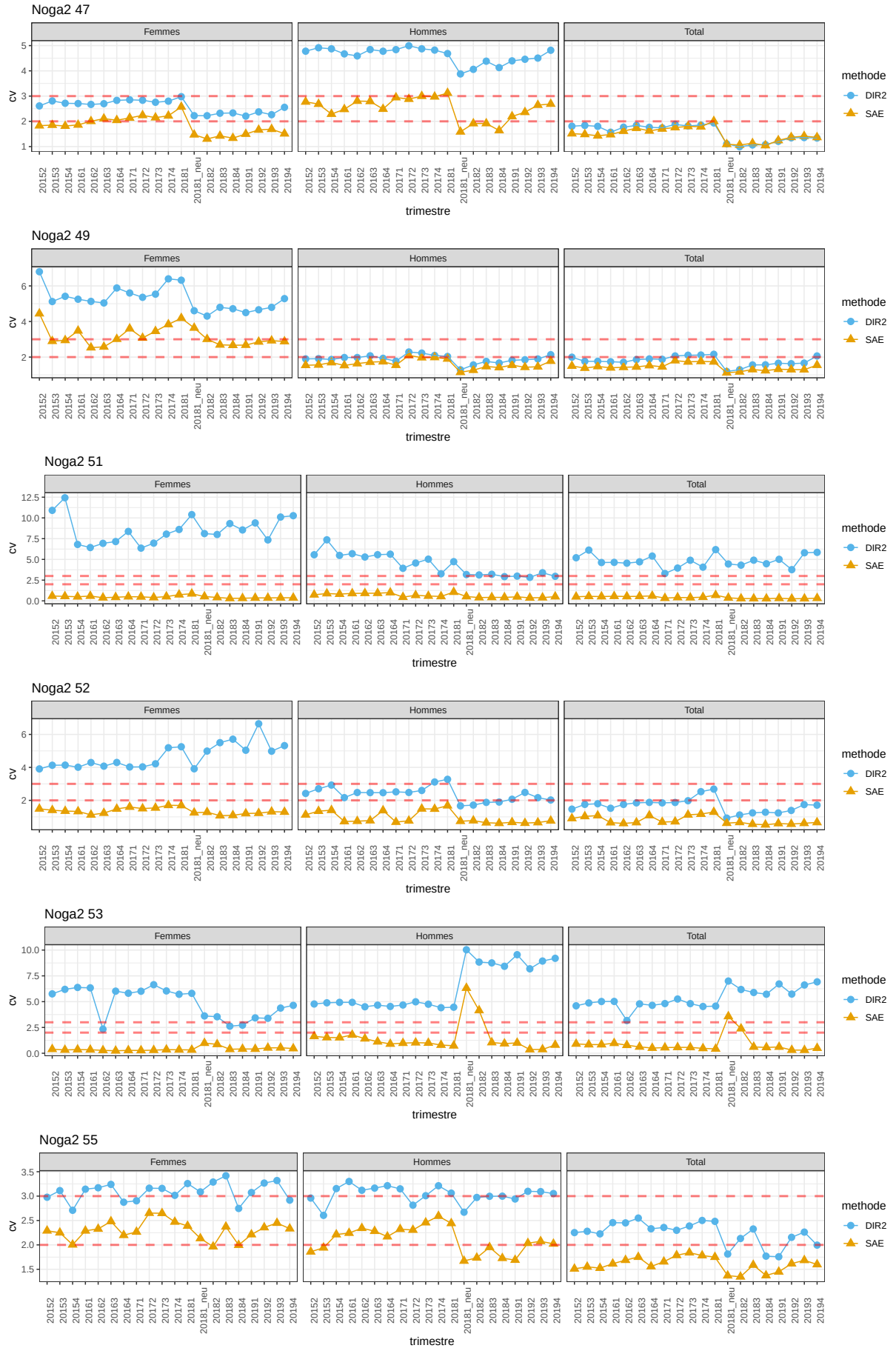


**Fig. 15:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

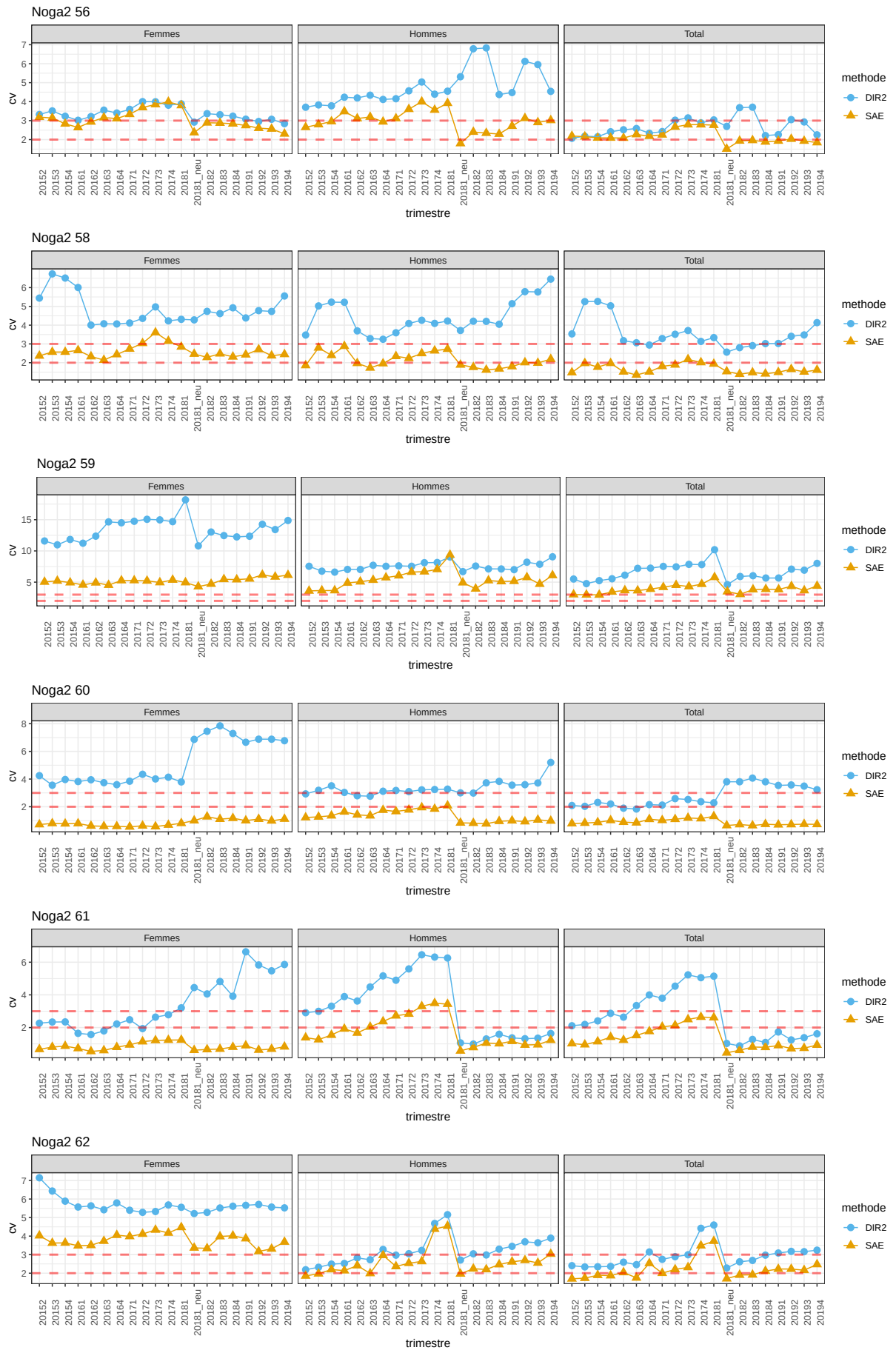


**Fig. 16:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

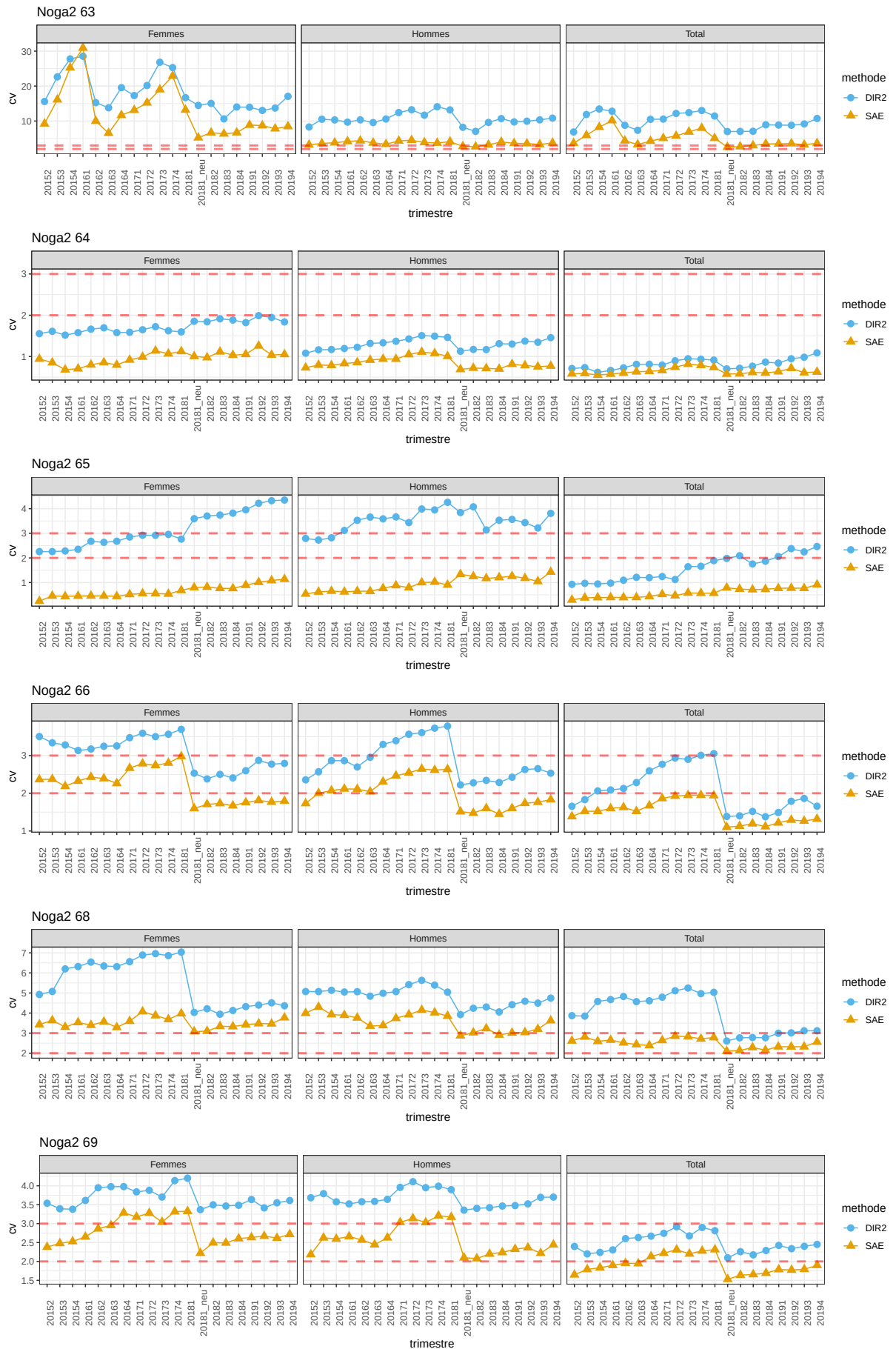


**Fig. 17:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

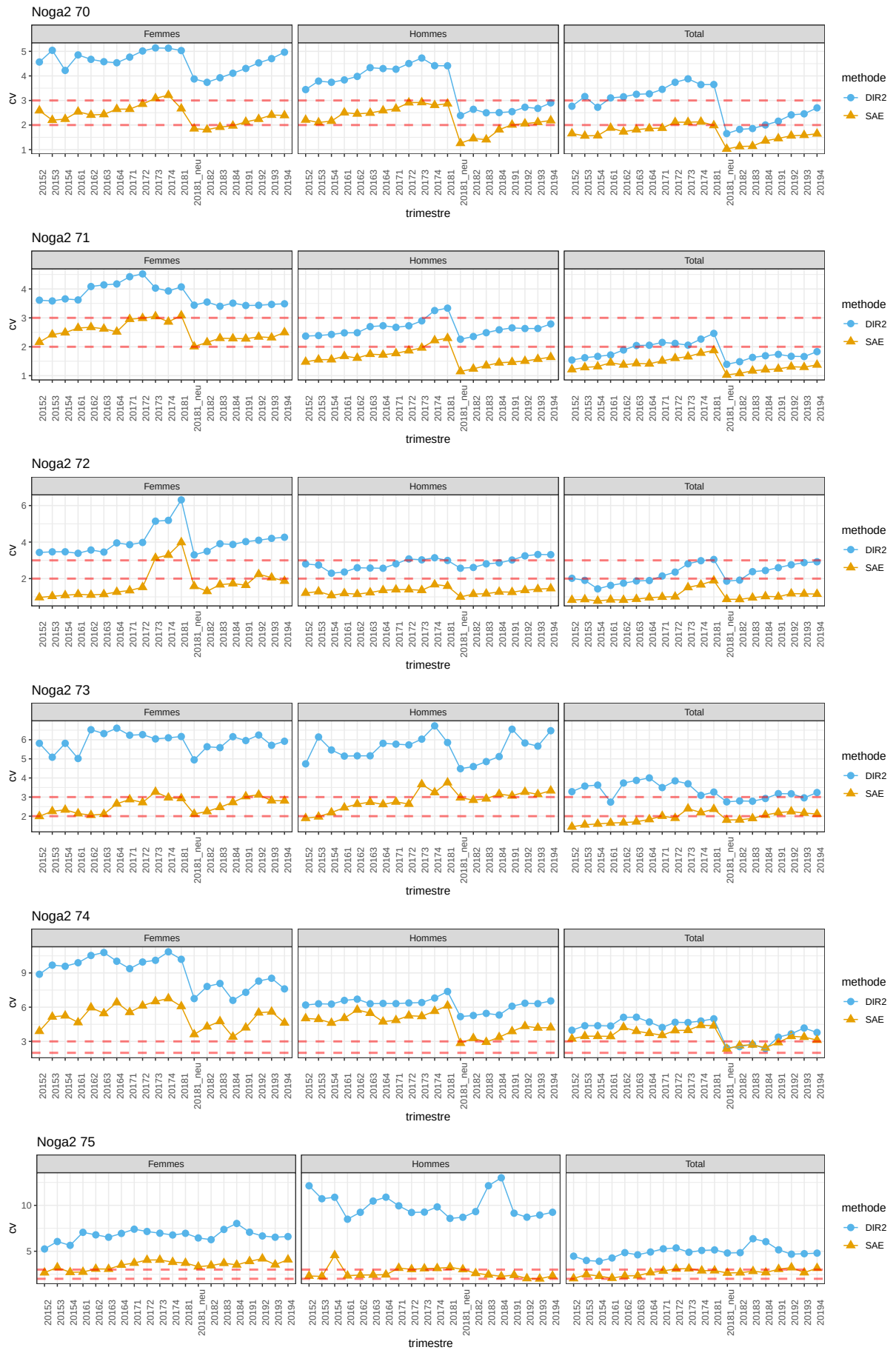


**Fig. 18: Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).**

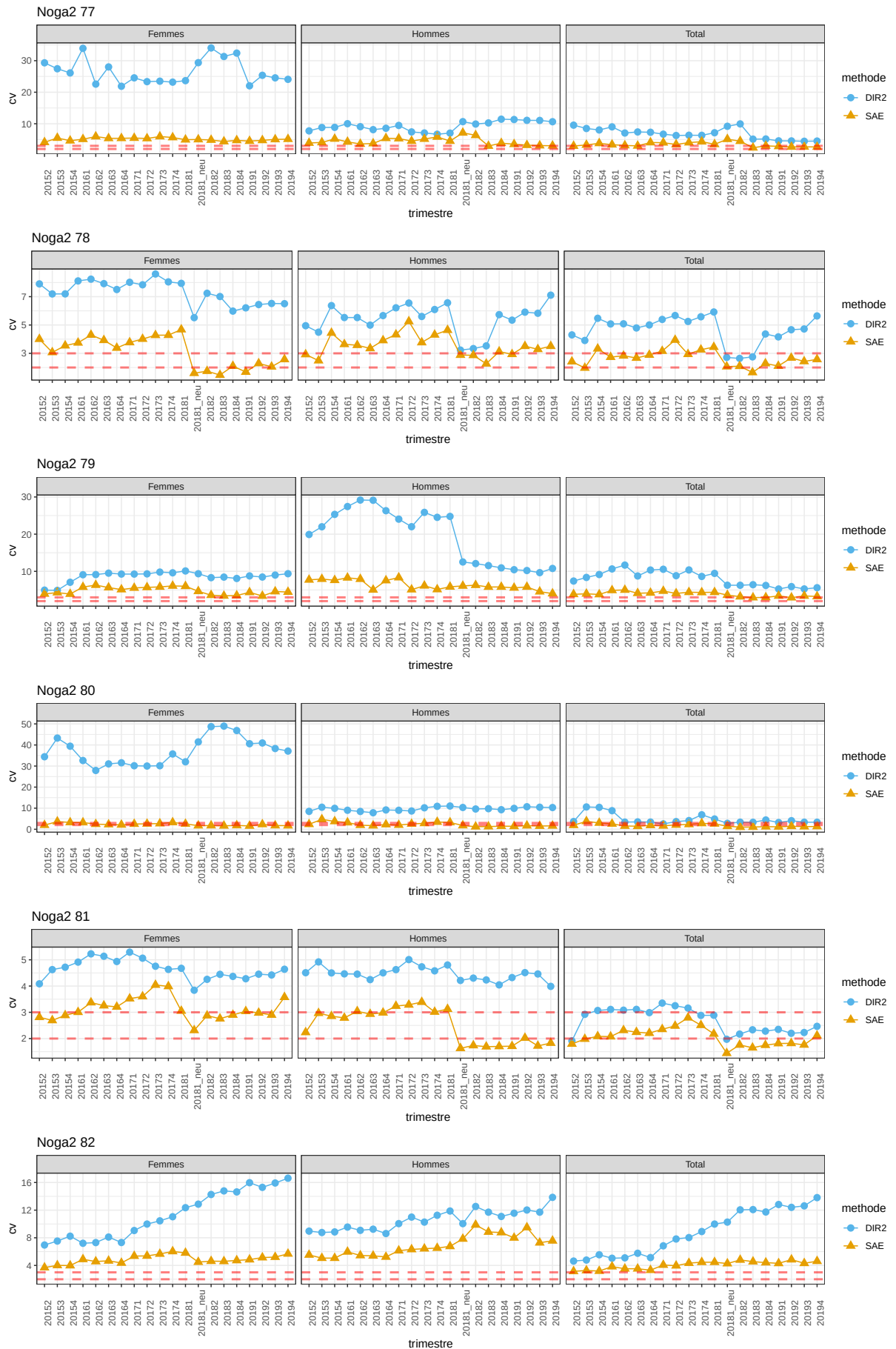


**Fig. 19:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

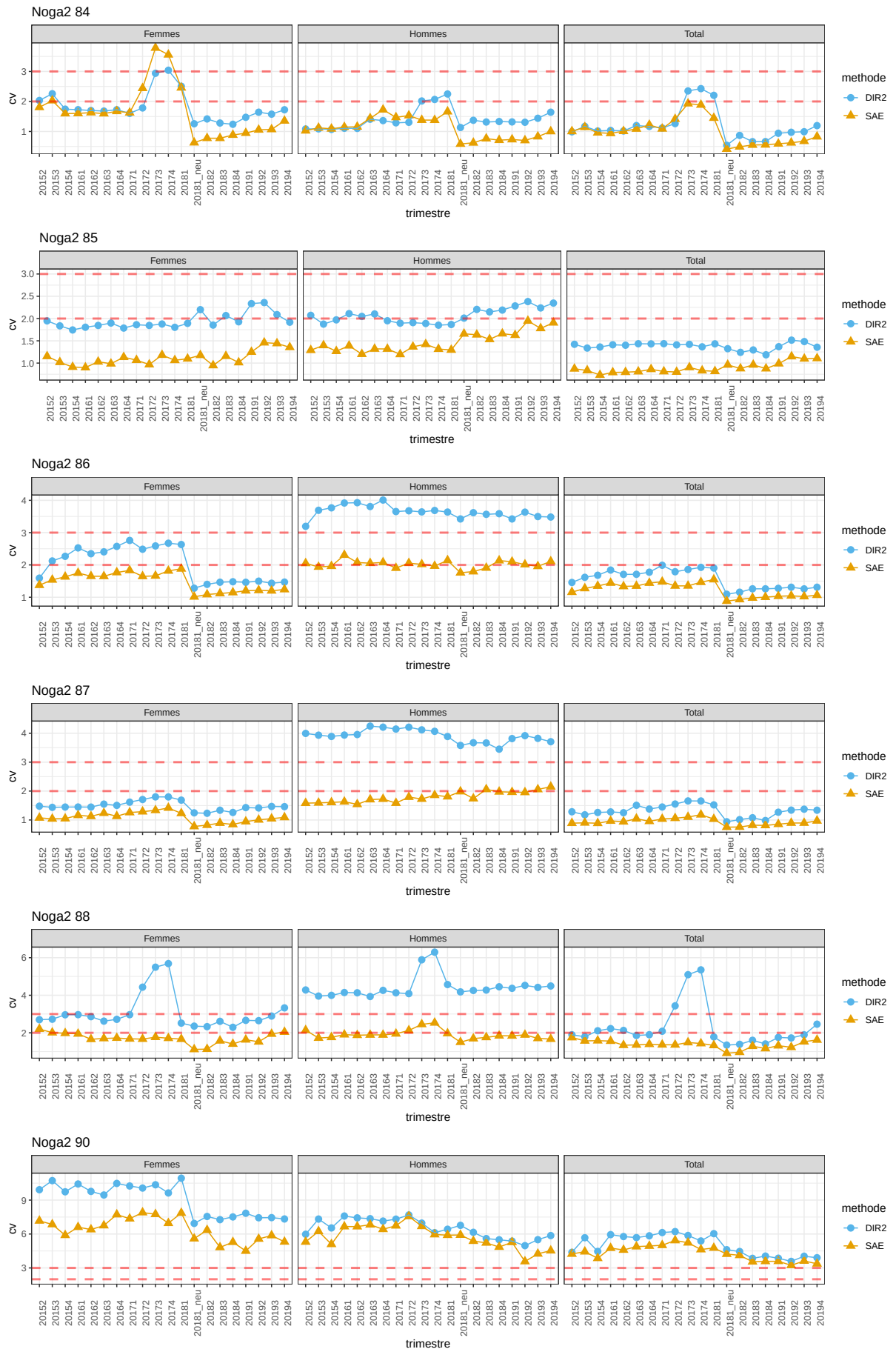


**Fig. 20:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

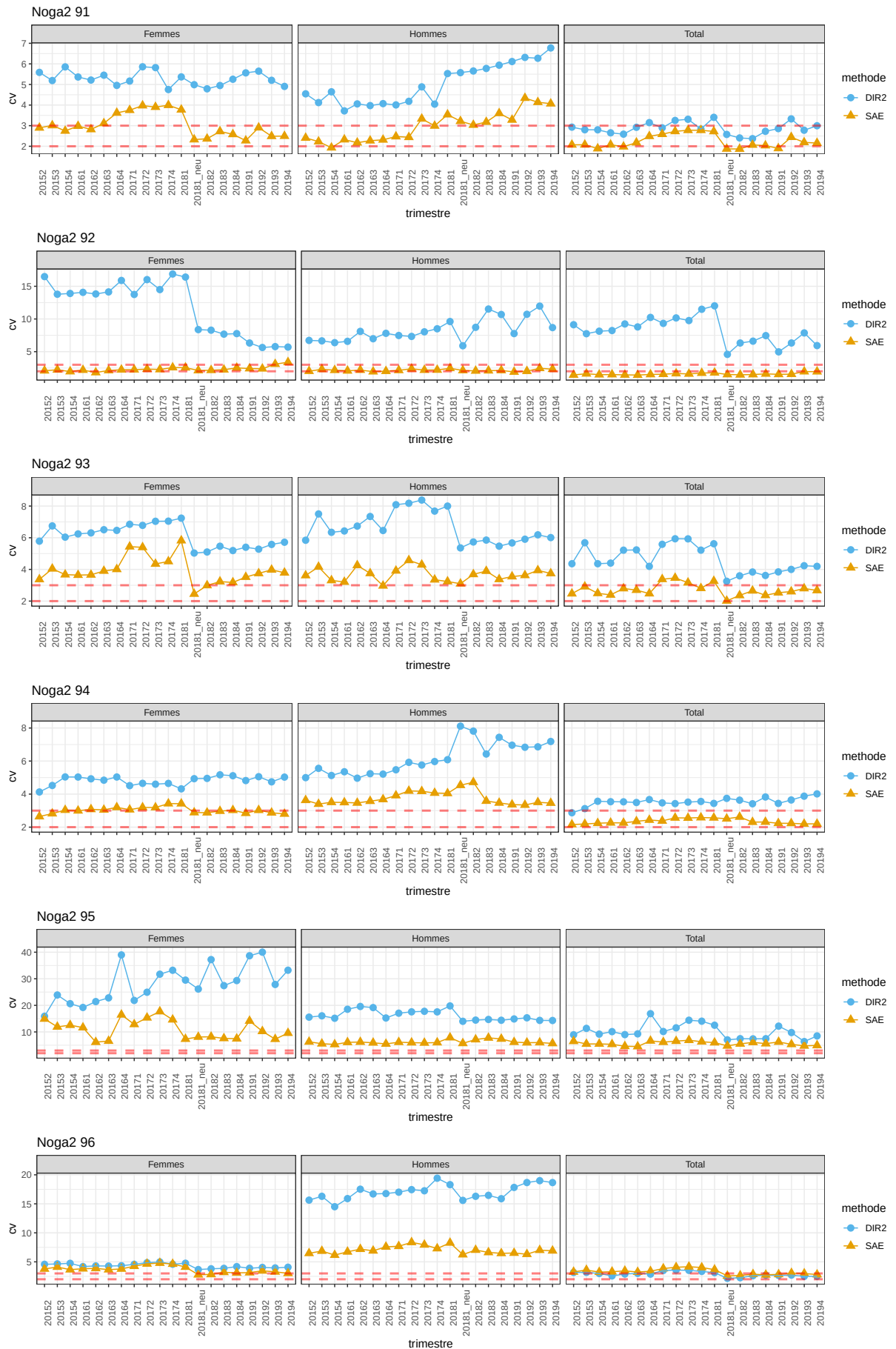


**Fig. 21:** Comparaison des CV des estimations SAE (sans benchmarking) aux CV des estimations directes par NOGA2 et par sexe (suite).

## 6.4 Codes et organisation des dossiers

L'utilisation des codes se fait à partir d'un code principal qui était le suivant (les chemins se rapportent à un environnement d'une instance virtuelle sécurisé de l'OFIT) :

```
#####
Chemin vers le dossier qui contient les codes
source_path="C:/BFS_NDS_BESTA/Code/_users/DJME/Codes_propres/"
#=====
Chargement des libraires
#=====
Ce code charge toutes les libraires nécessaires pour exécuter tous les codes

source(paste0(source_path,"01_Librairies.R"))

#=====
Chargement de fonctions utiles
#=====
Ce code charge toutes les fonctions nécessaires pour exécuter le reste des codes
source(paste0(source_path,"02_Fonctions_utiles.R"))

#=====
Choix des paramètres
#=====
source des données
pathData="C:/BFS_NDS_BESTA/Data/_raw/besta_cadrectrechrep/besta_cadrectrechrep_"

source des CV des estimateurs directs (calculés par la section KE)
pathCV_KE="C:/BFS_NDS_BESTA/Data/_raw/variances/"

dossier des résultats
pathResults="C:/BFS_NDS_BESTA/Data/Processed/DJME/variance_bootstrap/
clusterPoisson_original_population/"

Choix des trimestres

yeartrims=spl('20152,20153,20154,20161,20162,20163,20164,20171,20172,20173,20174,
20181,20181_neu,20182,20183,20184,20191,20192,20193,20194')

une autre façon pour générer les trimestres
sort(c(paste0(rep(c(2015:2019),4),c(1:4)), "20181_neu"))[-1]

Choix des paramètres du bootstrap
D=10 # nombre de pseudo-populations créées
B=100 # nombre d'échantillons bootstrap par pseudo-population

nombre total d'échantillons bootstrap = D x B

D*B

sexes=spl('F,M') # pas besoin de mettre le total ici

for(yeartrim in yeartrims){
print(yeartrim)

for(sexe in sexes){
print(sexe)
#=====

Lecture et préparation des données
```

```

#=====
Ce code charge et prépare les données nécessaires pour calculer les différents
estimateurs pour le semestre et le sexe choisis
source(paste0(source_path,"03_Lecture_et_preparation_des_donnees.R"))

#=====
Calcul des estimateurs directs
#=====
Ce code calcule les estimateurs directs pour le semestre et le sexe choisis sur
toutes les ventilations
source(paste0(source_path,"10_calcul_des_estimateurs_DIR.R"))

#=====
Calcul des estimateurs SAE
#=====
Ce code calcule les estimateurs SAE pour le semestre et le sexe choisis sur
toutes les ventilations
source(paste0(source_path,"11_1_calcul_des_estimateurs_SAE.R"))

Ce code recalcule les estimateurs SAE pour que les totaux correspondent aux estimateurs
directs des grandes régions et des sections (benchmarking)
le code 11_calcul_des_estimateurs_SAE.R doit être lancé avant celui-ci
source(paste0(source_path,"11_2_benchmarking_des_estimateurs_SAEN.R"))

#=====
Calcul de la variance des estimateurs SAE
#=====
Ce code calcule les variances des estimateurs SAE (avec et sans benchmarking) pour le
semestre et le sexe choisis sur toutes les ventilations
le code 11_calcul_des_estimateurs_SAE.R doit être lancé avant celui-ci
source(paste0(source_path,"12_calcul_de la_variance_estimateurs_SAE.R"))
source(paste0(source_path, "12_calcul_de la_variance_estimateurs_SAE_avec_BM
_original_population.R"))
}

#=====
Calcul de la variance des estimateurs SAE pour le total
#=====
Ce code calcule les variances des estimateurs SAE (avec et sans benchmarking) du total
des deux sexes pour le semestre choisi sur toutes les ventilations
la variance des deux sexes doit être calculée avant celui-ci
!!! la variance des estimateurs avec benchmarking est conditionnelle et ne peut pas être
utilisée telle quelle pour construire des intervalles de confiance. !!!
source(paste0(source_path,"13_2_calcul_de la_variance_estimateurs_SAE_pour_le_total_BM.R"))

#=====
Calcul des intervalles de confiance pour les estimateurs SAE en utilisant
les échantillon du bootstrap
#=====
source(paste0(source_path,"14_calcul_des intervalles_de_confiance.R"))
#=====
Combinaison des résultats SAE et les résultats DIR
#=====
Ce code combine les CV SAE avec les CV directs calculés par la section KE.
source(paste0(source_path,"21_Combiner_les_resultats.R"))

#=====
Graphiques de comparaison des CV SAE et des CV DIR
#=====
Ce code produit les graphiques comparant les CV SAE avec les CV directs calculés

```

# par la section KE.

```
source(paste0(source_path,"31_Graphiques_comparaison_de_CV.R"))
```

Les résultats sont sauvegardés dans des sous-dossier différents à l'intérieur du dossier « résultats » spécifié par l'utilisateur. Tous les sous-dossiers sont créés automatiquement.

Le sous-dossier suivant contient les estimateurs directs.

\DIR

Le sous-dossier suivant contient les estimateurs SAE (sans et avec benchmarking).

\SAEN

Pour  $B = 100$  et  $D = 10$  ( $B \times D = 1000$ ), le sous-dossier suivant contient les résultats détaillés de l'estimation de la variance des estimateurs SAE (sans et avec benchmarking), ainsi que les intervalles de confiance basés sur les échantillons du bootstrap.

\SAEN\1000

Le suivant compare les CV des estimateurs directs aux CV des estimateurs SAE (sans benchmarking).

\Summary

Celui-là contient les graphiques. Les graphiques sont organisés dans des sous-dossiers correspondants aux ventilations.

\Plots

Et celui-là, les tables de données utiles pour générer les graphiques.

\Donnees\_pour\_les\_graphiques.