

MÉMOIRE DE MASTER DE STATISTIQUE PUBLIQUE

Dans quelle mesure l'intégration d'informations
sous-sectorielles permet-elle d'améliorer la qualité
de la prévision de la production manufacturière ?

ÉLÈVE

Alain Quartier La Tente

MAÎTRE DE STAGE

Pierre-Damien Olive

JURY

Pierre-Damien Olive
Olivier Sautory
Lionel Truquet

Table des matières

Remerciements	ix
Introduction	1
Glossaire	3
Présentation de l'unité d'accueil	5
1 Description des données disponibles	9
1.1 Nomenclature et calendrier de diffusion	9
1.1.1 La nomenclature d'activité française	9
1.1.2 Différents calendriers et exemple de référence de la note de juin	10
1.2 Deux enquêtes de conjoncture.	11
1.2.1 ... La première réalisée par l'Insee	12
1.2.2 ... La seconde réalisée par la Banque de France	12
1.3 Indicateurs quantitatifs de production et comptes nationaux	13
1.3.1 Les indices de production industrielle	14
1.3.2 Différences entre indices de production industrielle et production indus- trielle	14
1.3.3 La production manufacturière	14
1.3.4 Volatilité des différents sous-secteurs de l'industrie manufacturière	15
2 Choix méthodologiques adoptés	17
2.1 Une sélection de modèles réalisée avec l'algorithme Gets	17
2.1.1 Modèles recherchés par l'algorithme	17
2.1.2 Tests utilisés pendant l'algorithme	18
2.1.3 Principe de l'algorithme	19
2.1.4 Détection des points atypiques	21
2.2 Trois spécifications utilisées pour la construction des étalonnages	22
2.2.1 Soldes utilisés pour les différentes spécifications	22
2.2.2 Méthode 1 : étalonnages avec les soldes bloqués et l'acquis d'IPI	23
2.2.3 Méthodes 2 et 3 : en prévoyant l'IPI sur tout le trimestre et sans l'IPI	23
2.3 Différentes façons d'agréger les données.	24
2.3.1 ... Par une approche comptable	25
2.3.2 ... Par une approche naturelle	25
2.3.3 ... Par une approche statistique	25
2.4 Outils de comparaison des modèles	26

3	Résultats	29
3.1	Résultats des différents étalonnages	29
3.1.1	Description générale des équations de prévision	29
3.1.2	Performances prédictives des étalonnages au niveau A 17	30
3.1.3	Performances prédictives des étalonnages au niveau A 64	31
3.2	Comparaison des méthodes d'agrégation des prévisions	32
3.3	Retour sur les prévisions depuis 2014	34
3.3.1	Prévisions depuis 2014	34
3.3.2	Retour sur les prévisions du deuxième trimestre	37
4	Discussion	39
4.1	Comment améliorer les prévisions?	39
4.1.1	En utilisant d'autres sources et d'autres types de variables	39
4.1.2	En prolongeant les soldes et en réalisant un étalonnage de l'IPI spécifique à chaque mois	39
4.1.3	En exploitant les informations des modèles terminaux	40
4.1.4	En prenant en compte les liens entre les secteurs par la méthode SURE	40
4.2	Limites de l'étude	41
4.2.1	Utilisation d'autres tests de spécification	41
4.2.2	Limite de la prévision des soldes	41
	Conclusion	43
A	Organigrammes de la Dese et de l'Insee	45
B	Nomenclature d'activité française détaillée et questionnaire de l'enquête de conjoncture dans l'industrie	47
C	Principes sur les enquêtes et les soldes d'opinion	51
D	Contribution à la variance	55
E	Retour sur les pondérations	57
F	Équations d'étalonnage	61

Liste des tableaux

1.1	Découpage de la nomenclature de l'activité française de l'industrie manufacturière utilisé pour les prévisions de la production manufacturière.	10
1.2	Poids en terme de production et contribution à la variance du secteur automobile (CL1) et des autres matériels de transport (CL2) dans l'industrie manufacturière et dans les matériels de transport (C4)(en %)	16
1.3	Provenances des variables utilisées dans les différents étalonnages	16
3.1	Contribution à la variance (en %) de l'évolution de la production estimée de l'IPI et des soldes d'enquête dans les étalonnages avec l'IPI connu sur tout le trimestre (méthode 2)	30
3.2	RMSE complet pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation	30
3.3	RMSE dans l'échantillon pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation	31
3.4	RMSE pseudo temps-réel pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation	32
3.5	Différents RMSE pour les méthodes 1 et 3 de construction des étalonnages au niveau A 64 ainsi que pour l'industrie manufacturière et le secteur des matériels de transport (C4) obtenus par agrégation	32
3.6	Comparaison des différentes méthodes d'agrégation au niveau A 17 pour la méthode de construction des étalonnages par les soldes bloqués (méthode 1) en terme de RMSE	33
3.7	Comparaison des différentes méthodes d'agrégation au niveau A 64 pour la méthode de construction des étalonnages par les soldes bloqués (méthode 1) en terme de RMSE	33
3.8	Taux de croissance (en %) réel, retenu dans les différentes notes de conjoncture, et prévu par agrégation des prévisions au niveau A17 de la production manufacturière du premier trimestre 2014 au deuxième trimestre 2015	34
3.9	Taux de croissance (en %) réel et prévus des comptes de production trimestriels du premier trimestre 2014 au deuxième trimestre 2015	35
3.10	Contribution à la variance de l'erreur de prévision de la production manufacturière (en %) des prévisions effectuées au niveau A 17 du premier trimestre 1991 au deuxième trimestre 2015	36
3.11	Taux de croissance (en %) réels et prévus au niveau A 17 et contribution à l'évolution (en point) de la production manufacturière de chaque secteur	38
B.1	Découpage de l'industrie manufacturière aux niveaux A 17, A 38 et A 64.	47

C.1	Opinion des industriels par dans l'industrie manufacturière	52
C.2	Évolution mensuelle de la production industrielle (en %)	52
C.3	Opinion des industriels par sous-secteurs de l'industrie manufacturière	53
E.1	Soldes et variables provenant de l'enquête de conjoncture dans l'industrie de l'Insee utilisés pour les étalonnages de la production.	61
E.2	Soldes provenant de l'enquête de conjoncture de la Banque de France utilisés pour les étalonnages de la production.	61
E.3	Étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 par la méthode des soldes bloqués (méthode 1) du 3 ^e trimestre 1990 au 4 ^e trimestre 2014	65
E.4	Résultats des étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 sans l'IPI (méthode 3) du 3 ^e trimestre 1990 au 4 ^e trimestre 2014	65
E.5	Résultats des étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 avec l'IPI connu sur tout le trimestre coïncident (méthode 2) du 3 ^e trimestre 1990 au 4 ^e trimestre 2014	66
E.6	Résultats des étalonnages de l'IPI au niveau A 17 et A 64 de la Naf rév. 2 par la méthode des soldes bloqués du 1 ^{er} trimestre 1991 au au 4 ^e trimestre 2014	66

Liste des graphiques

1	De la création de l'Insee à aujourd'hui : histoire de l'enquête de conjoncture dans l'industrie	7
1.1	Part de la production manufacturière et contribution à la variance des différents secteurs de l'industrie manufacturière au niveau A 17 (en %)	15
2.1	Estimation et test du modèle général	19
2.2	Réduction du modèle GUM et recherche des modèles terminaux	20
2.3	Détermination du modèle final	21
3.1	Taux de croissance de la production manufacturière réels et prévus par la méthode des soldes bloqués au niveau A 17	36
E.1	Liens entre les secteurs pour l'agrégation des prévisions réalisées à l'année a avec la méthode « d'approximation »	59
E.1	Prévisions en pseudo temps-réel de la production manufacturière du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 obtenu par agrégation des étalonnages au niveau A 17 par la méthode des soldes bloqués.	62
E.2	Prévisions en pseudo temps-réel de la production dans l'agro-alimentaire (C1) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	62
E.3	Prévisions en pseudo temps-réel de la production dans la cokéfaction-raffinage (C2) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	62
E.4	Prévisions en pseudo temps-réel de la production dans les biens d'équipement (C3) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	63
E.5	Prévisions en pseudo temps-réel de la production dans les matériels de transport (C4) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	63
E.6	Prévisions en pseudo temps-réel de la production dans l'automobile (CL1) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	63
E.7	Prévisions en pseudo temps-réel de la production dans les autres matériels de transport (CL2) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	64
E.8	Prévisions en pseudo temps-réel de la production dans les autres industries (C5) du 1 ^{er} trimestre 2000 au 2 ^e trimestre 2015 par la méthode des soldes bloqués.	64

Liste des encadrés

1.1	Propriétés économétriques des soldes d'opinion	11
1.2	L'indicateur de surprise	13
2.1	La correction de Bonferroni	18
2.2	Le critère d'information d'Akaike	20
2.3	Trimestrialisation des soldes	24
2.4	La validation croisée et le RMSE pseudo temps-réel	26
2.5	Le test de Diebold et Mariano	27
3.1	Décompositions de l'erreur de prévision de la production manufacturière obtenue par agrégation	34

Remerciements

Je voudrais ici prendre le temps de remercier toutes les personnes qui m'ont aidé à réaliser ce mémoire.

Avant tout, je tiens à remercier mon tuteur, Pierre-Damien. Depuis le début de cette étude, il m'a été d'une grande aide grâce à ses connaissances théoriques, pratiques et bibliographiques. Toujours disponible pour éclairer mes doutes et répondre à mes questions, il m'a aidé à acquérir de nombreuses notions scientifiques. Il a notamment contribué à rendre ce mémoire très formateur.

Je souhaiterais également remercier toutes les autres personnes de la division des enquêtes de conjoncture. Que ce soit à travers leurs sourires ou leur compagnie, ils m'ont permis de réaliser ce mémoire dans un cadre conviviale et agréable. En particulier, j'aimerais remercier Vladimir, le chef du département de la conjoncture, Catherine, la cheffe de division des enquêtes de conjoncture, et Yaëlle, la cheffe de section des enquêtes de conjoncture dans l'industrie. À travers leurs relectures, remarques et suggestions, ils m'ont été d'une grande aide et je leur en suis très reconnaissant.

Enfin, je remercie tous mes camarades qui m'ont soutenu jusqu'au rendu de ce mémoire. En particulier je remercie Romain, pour ses nombreuses relectures et remarques. Pour finir, je remercie Kim, dont les cycles d'humeur m'ont aidé à comprendre les cycles de la conjoncture, pour ses corrections et sa patience.

Introduction

« Le long terme est un mauvais guide pour les affaires courantes. À long terme, nous serons tous morts. Les économistes se fixent une tâche peu utile s'ils peuvent seulement nous dire que, lorsque l'orage sera passé, l'océan sera plat à nouveau. »
John Maynard Keynes, *A Tract on Monetary Reform*, 1923.

À travers ce message, J.M. Keynes visait aussi bien les décideurs politiques que les économistes dont les théories s'appuyaient sur les effets de long terme des politiques. Selon Keynes, le présent permet de guider le futur et d'atténuer son incertitude : les diagnostics économiques et l'orientation des politiques doivent donc s'appuyer sur l'examen des évolutions récentes et des prévisions à court terme de l'activité économique. Parmi les différents indicateurs économiques disponibles, celui qui est principalement utilisé est le produit intérieur brut (PIB) trimestriel. Il permet de mesurer la production économique réalisée à l'intérieur d'un pays et son évolution correspond à la croissance économique du pays.

En France, l'institut chargé de la production et de la publication des statistiques officielles est l'Insee. Le délai de publication du PIB par l'Insee est d'environ 45 jours après la fin du trimestre. Il est alors important de pouvoir établir un diagnostic conjoncturel de l'économie française avant la publication de ces résultats.

Dans la prévision du PIB interviennent des anticipations d'évolution de production dans plusieurs secteurs de l'économie française, dont le secteur des industries manufacturières. Les diagnostics conjoncturels effectués par l'Insee s'appuient principalement sur des données conjoncturelles calculées à un niveau fin de la nomenclature d'activité française. Toutefois, pour l'exercice de prévision de la production manufacturière, ces informations sont uniquement utilisées au niveau le plus agrégé (ensemble de l'industrie manufacturière).

Comment intégrer les informations sous-sectorielles dans la prévision de la production manufacturière ? Une approche en deux étapes peut être réalisée. Dans la première étape, les évolutions des productions à un niveau plus fin de la nomenclature sont estimées. La seconde étape consiste à agréger ces estimations afin de prévoir l'évolution de la production dans l'industrie manufacturière. L'avantage de cette décomposition est de fournir des informations quantitatives permettant une analyse plus détaillée de la situation économique du secteur manufacturier. Comment alors construire, pour chaque sous-secteur, le modèle de prévision ? Dans quelle mesure cela permettrait-il d'améliorer la qualité de la prévision de la production manufacturière ?

Pour répondre à ces questions, nous reviendrons tout d'abord sur les trois principales familles d'indicateurs disponibles pour décrire l'état de l'économie : les données issues d'enquêtes de conjoncture, les indicateurs quantitatifs de production et les comptes trimestriels de production. Ensuite, nous exposerons les différents choix méthodologiques adoptés. Ces choix portent à la fois sur la construction des modèles de prévision et sur la façon d'agréger les estimations des sous-secteurs de l'industrie manufacturière. Puisque trois méthodes de construction et

trois méthodes d'agrégation ont été retenues, nous détaillerons les différents moyens employés pour comparer les résultats. Enfin, tout cela sera mis en application dans le cadre de la Note de conjoncture, note trimestrielle publiée par l'Insee présentant la situation et les perspectives à court terme de l'économie française. Nous verrons alors comment seront intégrées les nouvelles méthodes pour établir le diagnostic conjoncturel dans l'industrie manufacturière en revenant sur les résultats relatifs au deuxième trimestre 2015.

Glossaire

Acquis (de croissance)	L'acquis de croissance d'une variable pour un trimestre T correspond au taux de croissance de la variable entre le trimestre $T - 1$ et le trimestre T que l'on obtiendrait si la variable demeurait jusqu'à la fin du trimestre T au niveau du dernier mois connu.
Acquis (de croissance) au mois 0	L'acquis de croissance au mois 0 d'une variable correspond à l'acquis de croissance de cette variable lorsque le dernier mois connu est le troisième mois du trimestre précédent.
Étalonnage	Un étalonnage est un modèle de régression linéaire. Lorsque nous parlons d'étalonnage de la production dans un secteur, nous faisons référence au modèle final utilisé pour prévoir le taux de croissance de la production dans ce secteur.
Garde fou	Le Garde fou est le rassemblement, dans un cadre comptable, de toutes les prévisions réalisées au niveau A 10 de la Naf rév. 2.
Gets	Gets est le nom de l'algorithme de sélection des variables utilisé dans ce mémoire.
IPI	L'indice de production industrielle (IPI) est un indicateur quantitatif mensuel de la production.
Naf rév. 2	La Naf rév. 2 correspond à la seconde révision de la nomenclature d'activité française. C'est la nomenclature actuellement utilisée pour décomposer l'activité française en niveaux d'agrégation, nommés « A xx » où xx représente le nombre d'agrégats.
Niveaux A 17 et A 64	Le niveau A 17 de la Naf rév. 2 permet un découpage de l'industrie manufacturière en 5 secteurs : l'agro-alimentaire (C1), la cokéfaction-raffinage (C2), les biens d'équipement (C3), les matériels de transport (C4) et les autres industries (C5). Le niveau A 64 est utilisé dans ce mémoire pour un découpage du secteur des matériels de transport (C4) en deux sous-secteurs : l'automobile (CL1) et les autres matériels de transport (CL2).
RMSE	Le RMSE – <i>root mean square error</i> – correspond à la racine carrée de l'erreur quadratique moyenne d'un modèle de prévision.
Réunion d'étalonnage	Réunion pendant laquelle les premières prévisions sont réalisées. C'est à ce moment que sont présentés les différents résultats des étalonnages.

Solde ou solde d'opinion	Un solde d'opinion est un indicateur quantitatif tiré d'informations qualitatives provenant de données d'enquêtes. Il est défini comme la différence entre la proportion de répondants ayant exprimé une opinion positive et la proportion de répondants ayant exprimé une opinion négative.
Solde bloqué	Nous parlerons de solde bloqué pour faire référence à une méthode de trimestrialisation des soldes d'enquêtes. Un solde mensuel peut être découpé en trois séries trimestrielles en fonction de la place du mois dans le trimestre. Par exemple, le solde bloqué au mois 1 correspond à la série obtenue en gardant les valeurs prises par le solde à chaque premier mois du trimestre.
Test de Diebold Mariano	Le test de Diebold Mariano est le test utilisé pour différencier les différents étalonnages en terme de qualité de prévision.
Trimestre coïncident	Le trimestre coïncident est le premier trimestre pour lequel la production n'a pas encore été publiée. Dans ce mémoire il s'agira du trimestre en cours.

Présentation de l'unité d'accueil

Ce mémoire a été réalisé dans le cadre de ma prise de poste depuis mars 2015 dans la division des enquêtes de conjoncture du département de la conjoncture, au sein de la direction des études et des synthèses économiques (Dese) à l'Institut national de la statistique et des études économiques (Insee), en tant que responsable du suivi conjoncturel dans l'industrie. À ce titre, je suis responsable de trois enquêtes de conjoncture réalisées par l'Insee : l'enquête mensuelle et l'enquête trimestrielle sur la situation et les perspectives dans l'industrie et l'enquête sur la situation de la trésorerie dans l'industrie. Un des rôles importants relatif à ce poste est notamment de participer à l'élaboration du scénario macroéconomique à court-terme et à la rédaction des notes de conjoncture publiées chaque trimestre ¹. Dans celles-ci, des prévisions d'activité dans l'industrie sont réalisées pour le trimestre en cours (ou « trimestre coïncident ») et le(s) suivant(s) (« trimestre(s) futur(s) ») ², qui serviront à construire les prévisions du taux de croissance trimestriel du PIB.

1 De l'Insee à la division des enquêtes de conjoncture

L'Insee est l'institut chargé de la production, de l'analyse et de la publication des statistiques officielles en France. Créé le 27 avril 1946, il dispose d'une indépendance vis-à-vis du gouvernement et son actuel directeur est Jean-Luc Tavernier. Son objectif principal est d'éclairer le débat économique et social et il est également chargé de veiller au respect du secret statistique. Pour y contribuer, l'Insee s'est donné six missions : collecter et produire, analyser, diffuser, coordonner, enseigner et développer la recherche et contribuer à l'édification d'un espace statistique international. Ce mémoire a été réalisé dans la direction des études et des synthèses économiques, qui est une des directions de l'Insee qui permet d'assurer ces six missions.

1.1 Direction des études et des synthèses économiques (Dese)

La Dese élabore et analyse des informations conjoncturelles, en fait périodiquement la synthèse et produit les comptes nationaux annuels et trimestriels. En particulier, elle réalise des études et prévisions économiques. Elle est composée d'une cellule administration des ressources et de trois départements : le département de la conjoncture, le département des études économiques et le département des comptes nationaux. Des organigrammes plus précis de l'Insee et de la Dese sont présents dans l'annexe [A](#).

1.2 Département de la conjoncture

Le département de la conjoncture a la responsabilité de l'ensemble des analyses conjoncturelles de l'Insee, et notamment des notes de conjoncture qui font périodiquement la synthèse sur

1. Sauf en octobre où un Point de conjoncture est publié.

2. Par trimestre en cours ou coïncident on entend le premier trimestre pour lequel la production n'a pas encore été publiée.

la situation de l'économie européenne et française et sur les perspectives à court terme. Ce département comporte une cellule et deux divisions :

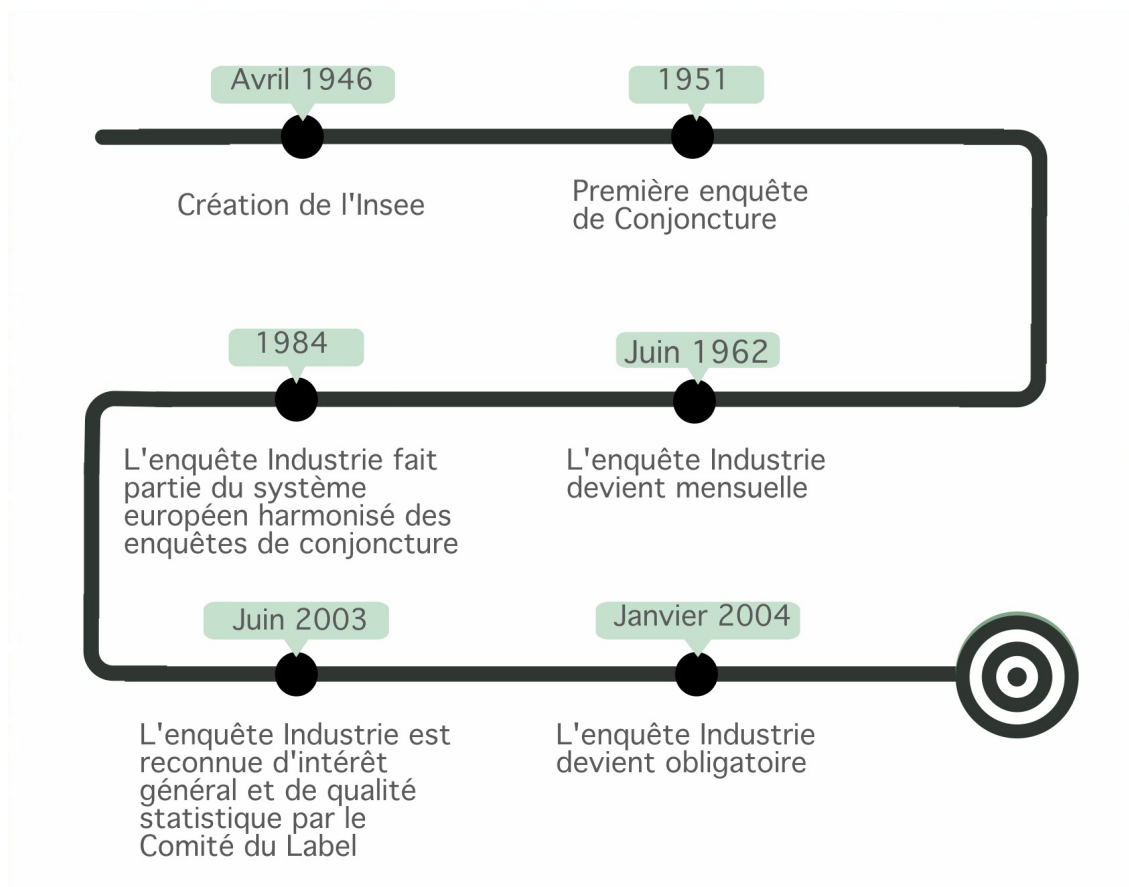
- La cellule « publications conjoncturelles » a pour mission de mettre en forme l'ensemble des publications conjoncturelles de l'Insee et d'en suivre la fabrication et la diffusion.
- La division « synthèse conjoncturelle » effectue en permanence la synthèse de l'ensemble des informations conjoncturelles relatives à l'économie française dans son environnement international. Tous les trimestres, elle élabore les notes de conjoncture présentant la situation et les perspectives à court terme de l'économie française.
- La division « enquêtes de conjoncture » gère l'ensemble des enquêtes nationales de conjoncture effectuées régulièrement auprès des entreprises (industrie, bâtiment, commerce de gros et de détail, services, investissement, trésorerie). Elle saisit et traite leurs réponses, puis présente les résultats dans les Informations Rapides, documents de deux pages résumant les résultats les plus récents des enquêtes de conjoncture. Elle participe aussi à l'élaboration des notes de conjoncture, notamment en élaborant les principales prévisions pour le scénario France de ces notes.

2 Le métier de responsable d'enquête

En tant que responsable du suivi conjoncturel dans l'industrie, j'ai à ma charge trois enquêtes, donnant toutes lieu à l'écriture d'Informations Rapides :

- L'enquête mensuelle de conjoncture dans l'industrie vise à recueillir les opinions des industriels sur l'évolution de leur activité à court terme ainsi que sur celle de leurs effectifs, de leurs carnets de commande et de leurs stocks en produits finis. Tout d'abord créée de manière trimestrielle, c'est l'une des premières enquêtes de conjoncture de l'Insee. Le graphique 1 retrace ses évolutions importantes depuis la création de l'Insee. Avec l'enquête mensuelle de conjoncture de la Banque de France, ce sont les deux enquêtes d'où proviendra la majorité des données utilisées lors de ce mémoire.
- L'enquête trimestrielle de conjoncture dans l'industrie est complémentaire de l'enquête mensuelle de conjoncture dans l'industrie. Son objectif est d'obtenir des informations supplémentaires auprès des industriels sur divers sujets : leurs capacités de production (taux d'utilisation des capacités de production, facteurs limitant la production, etc.), l'emploi (évolution des effectifs, de la durée du travail, du taux de salaire, etc.), les conditions de trésorerie, la position compétitive, les prix de vente, les délais de livraison.
- L'enquête semestrielle sur la situation de trésorerie dans l'industrie recueille les opinions des industriels sur les divers facteurs influençant leur situation de trésorerie actuelle et leurs résultats d'exploitation. Elle permet également d'étudier leurs conditions de financement, leurs anticipations à court terme concernant leur trésorerie et leurs résultats d'exploitation, ainsi que les divers usages futurs que les industriels comptent faire de leur trésorerie (hausse de l'investissement et de l'emploi, baisse des prix de vente, désendettement, etc.).

En parallèle du travail de gestion d'enquête, d'autres tâches récurrentes me sont également confiées : la rédaction d'une note mensuelle à l'intention du directeur général de l'Insee relative à la situation conjoncturelle de l'industrie ; la rédaction d'un « cahier Production » trimestriel faisant la synthèse du scénario conjoncturel retenu pour le secteur de l'industrie manufacturière ; et la réalisation d'une prévision trimestrielle de l'activité et de l'emploi dans l'industrie manufacturière amenant à la rédaction d'une partie la fiche production dans la Note de conjoncture. De manière plus précise :



GRAPHIQUE 1 – De la création de l'Insee à aujourd'hui : histoire de l'enquête de conjoncture dans l'industrie.

- La note relative à la situation conjoncturelle de l'industrie vise à donner des éléments d'éclairage sur les soldes d'opinion publiés dans l'Information Rapide du mois concerné. En particulier, elle fait figurer les contributions des différents sous-secteurs de l'industrie à l'évolution des différents soldes, lesquelles sont complétées par quelques éléments microéconomiques qualitatifs (faisant mention des réponses des plus grandes entreprises répondantes à l'enquête ou des commentaires que certaines d'entre elles nous communiquent). Enfin, cette note s'achève par une actualisation de la prévision d'activité dans le secteur calculée sur la base des nouveaux résultats apportés par l'enquête.
- Le cahier *Conjoncture de l'industrie manufacturière* vise d'une part à faire la synthèse des nombreux indicateurs disponibles sur le secteur industrie provenant de trois sources : Insee, Banque de France et PMI Markit³. D'autre part, il intègre les informations microéconomiques concernant les grandes entreprises du secteur qui sont rassemblées par la Division indicateurs conjoncturels d'activité en charge du calcul de l'IPI (Indice de Production Industrielle). Cet indicateur permet de donner des éléments d'appréciation sur la conjoncture du secteur et d'étayer le scénario prévisionnel retenu pour la Note de conjoncture.
- Dans le cadre de la Note de conjoncture, la prévision de l'activité et de l'emploi au sein du

3. Markit est l'une des sociétés leader dans l'information financière dont les résultats des enquêtes mensuelles de conjoncture concernent plus de 30 pays et régions et sont suivis dans le monde.

secteur manufacturier consiste à estimer l'évolution de la production manufacturière (en variation trimestrielle) pour le trimestre coïncident et les deux trimestres suivants. Il en va exactement de même pour la variation trimestrielle de l'emploi manufacturier (hors intérim). L'horizon temporel de la prévision est donc court et l'exercice de prévision résulte d'une extrapolation des valeurs futures de la variable d'intérêt à partir de l'information passée (soldes d'opinion, acquis d'IPI, valeurs passées de la variable d'intérêt).

Techniquement ces prévisions sont calculées sur la base de modèles de régression linéaire (appelés *étalonnages*) et de modèles Vectoriels Auto-Régressifs (VAR).

L'ensemble de ces prévisions permet alors d'élaborer l'« histoire » conjoncturelle du secteur et celles-ci sont rassemblées au sein de cahiers appelés *Cahiers d'étalonnage*. Ces cahiers sont présentés et discutés chaque trimestre à l'occasion des *Réunions d'étalonnage*, à l'issue desquelles un premier scénario est fixé pour amorcer le Garde Fou⁴. À chaque exercice de prévision, un travail d'actualisation des modèles est effectué. Tout ce travail se concrétise par la rédaction des quelques paragraphes concernant le secteur manufacturier dans la fiche Production de la Note de Conjoncture.

C'est dans ces trois axes que s'inscrit ce mémoire : une meilleure prévision de la production manufacturière permet d'avoir des estimations plus précises de l'évolution du PIB français et une estimation par secteur permet de mieux comprendre les perspectives d'activité de l'industrie manufacturière.

4. Le Garde Fou consiste à mettre en cohérence toutes les prévisions dans un cadre comptable.

CHAPITRE 1

Description des données disponibles

L'ensemble des séries étudiées dans ce mémoire sont publiées à plusieurs niveaux d'agrégation : au niveau de l'industrie manufacturière mais également à un niveau plus fin. Ce découpage du secteur manufacturier est réalisé grâce à la nomenclature d'activité française (Naf). Pour réaliser les exercices de prévision de l'évolution des comptes trimestriels de production dans l'industrie manufacturière et ses sous-branches, deux types de variables mensuelles sont utilisés : les variables dites « *hard* » et les variables dites « *soft* ». Les premières sont des indicateurs quantitatifs de l'activité réelle. Ici, les seules variables de ce type qui sont utilisées sont les indices de production industrielle. Les secondes sont des données d'enquêtes réalisées auprès des chefs d'entreprises. Par rapport aux premières, elles ont l'avantage d'être disponibles plus rapidement, d'être peu révisées et d'apporter des indications sur le passé récent et les perspectives d'évolution à court terme dans chaque secteur. Ici, nous utilisons les données provenant de deux enquêtes : l'enquête de conjoncture dans l'Industrie publiée par l'Insee et celle publiée par la Banque de France.

1.1 Nomenclature et calendrier de diffusion

La nomenclature d'activité française (Naf) a été élaborée pour faciliter l'organisation de l'information économique et sociale. Sa finalité est donc essentiellement statistique. Dans ce mémoire nous étudierons la branche des industries manufacturières et ses différents agrégats. Les industries manufacturières sont définies par l'Insee comme « des industries de transformation des biens, c'est-à-dire principalement des industries de fabrication pour compte propre mais elles concernent aussi la réparation et l'installation d'équipements industriels ainsi que des opérations en sous-traitance pour un tiers donneur d'ordres ».

Pour les exercices de prévision des productions dans l'industrie manufacturière, différentes sources de données sont utilisées et avec différentes dates de diffusion. Ce sont les réunions d'étalonnage que nous utiliserons comme référence pour la configuration de disponibilité des données pour les prévisions de la production du trimestre coïncident. Nous utiliserons comme référence les réunions d'étalonnage de la Note de conjoncture de juin.

1.1.1 La nomenclature d'activité française

Révisée au 1er janvier 2008, la Naf rév. 2 succède à la nomenclature économique de synthèse (Nes) associée à la Naf rév. 1 et permet désormais de réaliser des comparaisons internationales. Elle est disponible en sept niveaux d'agrégation, nommés « A xx » où xx représente le nombre de postes du niveau, le regroupement le moins détaillé étant le niveau A 10.

Dans les Informations Rapides de l'enquête de conjoncture dans l'industrie, les principaux soldes de l'enquête sont publiés pour l'industrie manufacturière, mais également à un niveau plus fin (voir tableau 1.1 pour les niveaux étudiés et l'annexe B pour les autres secteurs) :

- au niveau A 17 qui permet le découpage de l'industrie manufacturière en cinq postes ;
- au niveau A 38 pour les principaux secteurs des autres industries (chimie, pharmacie, caoutchouc et plasturgie et la métallurgie) ;
- au niveau A 64 pour le secteur des matériels de transport, ce qui permet l'ajout du secteur de l'automobile et celui des autres matériels de transport.

Toutefois, les soldes d'opinion des enquêtes de conjoncture ne sont pas publiés dans le secteur de la cokéfaction-raffinage (C2) par l'Insee et la Banque de France du fait du faible nombre d'entreprises dans cette branche et pour respecter les règles du secret statistique. Ils sont néanmoins utilisés pour le calcul des soldes au niveau de l'industrie manufacturière. Dans ce secteur, nous utiliserons donc uniquement les informations provenant des enquêtes de l'Insee.

Dans ce mémoire, les étalonnages de la production seront étudiés au niveau A 17 et au niveau A 64 uniquement pour le secteur des matériels de transport (C4). Les comptes de production trimestriels ne sont diffusés que jusqu'au niveau A 17 : ces données ne sont donc pas diffusées et ne sont pas publiques pour les secteurs de l'automobile (CL1) et des autres matériels de transport (CL2).

Niveau A 17		Niveau A 64 utilisé	
Code	Libellé	Code	Libellé
C1	Industries agro-alimentaires		Non étudié
C2	Cokéfaction-raffinage		Non étudié
C3	Biens d'équipement		Non étudié
C4	Matériels de transport	CL1	Automobile
		CL2	Autres matériels de transport
C5	Autres industries		Non étudié

TABLEAU 1.1 – Découpage de la nomenclature de l'activité française de l'industrie manufacturière utilisé pour les prévisions de la production manufacturière.

1.1.2 Différents calendriers et exemple de référence de la note de juin

Les délais de publication des variables peuvent changer en fonction des sources. Ainsi, alors que les résultats des enquêtes de conjoncture de l'Insee sont disponibles autour du 22 du mois étudié, ceux des enquêtes de la Banque de France ne sont disponibles qu'environ 15 jours après la fin du mois concerné. L'indice de production industrielle (IPI) est, quant à lui, disponible environ 40 jours après la fin de celui-ci. Les premiers résultats de la production des comptes nationaux trimestriels sont disponibles environ 45 jours après la fin du trimestre étudié.

Ainsi, à une date fixée, toutes les informations ne sont pas disponibles pour le(s) même(s) mois. Dans ce mémoire, nous nous placerons dans le cas de référence de la Note de conjoncture du mois de juin. Pour cette note, les réunions d'étalonnage, où les premières prévisions de production manufacturière pour le trimestre coïncident et les trimestres futurs sont notamment décidées, ont lieu autour du 20 mai. À cette date, la production de l'industrie manufacturière est connue jusqu'au premier trimestre et l'IPI est connu jusqu'en mars. Les soldes de l'enquête Insee, non encore publiés, sont proches ou égaux à ceux qui seront publiés pour le mois de mai. Les soldes de l'enquête Banque de France ne sont connus que jusqu'en avril. Nous dirons que pour la prévision du trimestre coïncident l'IPI est connu jusqu'au mois 0¹. Nous dirons

1. Par mois 0 on entend que l'IPI n'est connu que jusqu'au dernier mois du trimestre précédent.

également que les soldes de l'enquête Banque de France sont connus jusqu'au mois 1 (premier mois du trimestre) et que les soldes de l'enquête Insee sont connus jusqu'au mois 2 (deuxième mois du trimestre) : nous parlerons d'étalonnage au mois 2.

Par ailleurs, tous les soldes d'enquêtes ne sont pas disponibles à partir de la même période :

- les productions des comptes trimestriels sont disponibles à partir du premier trimestre de l'année 1949 ;
- les indices de production industrielle et les principaux soldes de l'enquête de conjoncture dans l'industrie de l'Insee aux niveaux A 17 et A 64 sont disponibles à partir de l'année 1990 (respectivement à partir de janvier et février) ;
- les principaux soldes de l'enquête de conjoncture de la Banque de France sont disponibles à partir de janvier 1987.

Afin d'harmoniser les dates de début de toutes ces sources, nous considérerons par la suite que nos données sont disponibles à partir du 2^e trimestre de l'année 1990 pour les variables trimestrielles. Les variables trimestrielles sont les comptes de production trimestriels, à prévoir, et les soldes d'opinion lorsqu'ils sont trimestrialisés. Les variables mensuelles seront utilisées à partir d'avril 1990. De plus, afin d'avoir des données des comptes nationaux trimestriels qui seront peu modifiées, les étalonnages sont construits jusqu'au 4^e trimestre 2014.

1.2 Deux enquêtes de conjoncture...

Les indications sur le passé récent et les perspectives à court terme des enquêtes de conjoncture sont tirées d'informations de nature qualitative fournies par les entreprises interrogées. Ces informations sont ensuite transformées en variables quantitatives appelées soldes d'opinion. L'encadré 1.1 décrit les propriétés économétriques de ces soldes. Leur méthode de construction est détaillée dans l'annexe C. Dans ce mémoire deux enquêtes mensuelles de conjoncture ont été utilisées : l'enquête mensuelle de conjoncture dans l'industrie de l'Insee et l'enquête mensuelle de conjoncture de la Banque de France. Les indices PMI — *Purchasing Managers' Index* — provenant des enquêtes PMI Markit ne sont pas utilisés car il ne sont pas disponibles à un niveau plus fin que l'industrie manufacturière.



ENCADRÉ 1.1 - Propriétés économétriques des soldes d'opinion

Les soldes d'opinion ne doivent pas être interprétés comme représentatifs du taux de croissance des variables auxquelles ils se réfèrent. On ne peut les interpréter qu'en référence aux soldes antérieurs ou à leur moyenne de longue période car les soldes d'opinion sont supposés stationnaires (la loi d'un solde inconditionnellement à l'état de la conjoncture est invariante au cours du temps) ^a [1].

Par ailleurs, le mécanisme engendrant les soldes est supposé « indépendant » des agrégats macroéconomiques, notamment de la production dans les différentes branches de l'industrie manufacturière et l'indice de production industrielle (IPI). C'est-à-dire que les réponses des entreprises ne décrivent que leur situation propre et ne dépendent pas des chiffres publiés par la comptabilité nationale. Nous supposons donc que les étalonnages admettent une spécification qui rend les soldes fortement exogènes ^b.

^a. Cette stationnarité est également vérifiée par des tests de Dickey-Fuller.

^b. Pour les mêmes raisons, les étalonnages de la production sur l'IPI et les soldes admettent une spécification qui rend également l'IPI fortement exogène.

1.2.1 ... La première réalisée par l'Insee

L'enquête mensuelle de conjoncture dans l'industrie de l'Insee recouvre les secteurs de l'industrie manufacturière et celui des industries extractives (non étudiées ici). La collecte est effectuée de manière centralisée, par la voie postale ou par internet. L'échantillon est constitué d'environ 4 000 entreprises² et ne sont interrogées que celles d'au moins 20 salariés. Leurs opinions des chefs d'entreprise sont traduites sur deux séries de questions :

- Dans le premier groupe de variables, on cherche à appréhender l'évolution au cours des trois derniers mois et l'évolution prévue pour les trois prochains mois. Trois réponses qualitatives sont possibles (↘, → et ↗).
- Dans le deuxième groupe de variables on cherche à déterminer la situation par rapport au niveau jugé comme normal par les chefs d'entreprise. Trois réponses qualitatives sont possibles (supérieur à la normale, normale et inférieur à la normale).

Le questionnaire mensuel actuellement envoyé aux chefs d'entreprise pour répondre à cette enquête est présent dans l'annexe B.

Toutes les séries publiées sont corrigées des variations saisonnières, à l'exception de celles dans du secteur de la cokéfaction-raffinage (C2) du fait de leur forte volatilité.

Dans le cadre de ce mémoire, quatre soldes d'opinion différents provenant de l'enquête de conjoncture de l'Insee sont utilisés : deux correspondent à des évolutions (production passée et production prévue) et deux correspondent à la situation actuelle (niveau des carnets de commandes globaux et niveaux des stocks de produits finis). Par ailleurs, une autre variable calculée à partir des réponses individuelles des entreprises est utilisée : l'indicateur de surprise (voir encadré 1.2).

1.2.2 ... La seconde réalisée par la Banque de France

L'enquête mensuelle de conjoncture de la Banque de France couvre l'activité de plusieurs unités : l'industrie, les services, le négoce de gros et la construction (bâtiment et travaux publics). La collecte est effectuée de manière décentralisée par les unités du réseau de la Banque auprès d'un échantillon composé d'environ 8 500 entreprises ou établissements. Les informations sont recueillies au cours d'entretiens téléphoniques avec les chefs d'entreprise et permettent notamment de retracer l'activité dans l'industrie. Comme dans l'enquête de conjoncture de l'Insee, les opinions des chefs d'entreprise sont traduites sur deux séries de questions :

- Dans la première série de questions, on cherche à appréhender l'évolution au cours du mois écoulé par rapport au mois précédent et l'évolution prévue pour le prochain mois. Sept réponses qualitatives correspondant à une échelle de graduations sont possibles : forte augmentation, augmentation, stabilité, légère diminution, diminution, forte diminution. Cela permet de nuancer l'ampleur de l'évolution de la variable étudiée. Chaque réponse fait l'objet d'une pondération accordant un poids double aux réponses jugées « fortes » et un poids moitié moindre aux réponses jugées « légères » par rapport aux variations jugées « normales ».
- Dans la deuxième série de questions, c'est la situation par rapport à un niveau jugé normal que l'on cherche à appréhender. Sept réponses qualitatives sont également possibles avec le même jeu de pondérations que pour la première série de questions.

Toutes les séries publiées sont désaisonnalisées.

2. Environ 20 000 entreprises sont interrogées par l'ensemble des enquêtes de conjoncture de l'Insee.

Dans le cadre de ce mémoire, quatre soldes d'opinion différents provenant de l'enquête de la Banque de France sont utilisés : trois correspondent à des évolutions (production passée, production prévue et évolution des carnets de commandes) et un correspond à une situation par rapport au niveau jugé normal (situation actuelle de l'état des carnets de commandes).

ENCADRÉ 1.2 - L'indicateur de surprise

Pour établir des diagnostics conjoncturels, certains analystes s'appuient sur la différence entre les soldes d'opinion sur la production passée et les perspectives de production de la période précédente, ce qui peut s'interpréter comme la surprise qu'ont enregistrée les entreprises par rapport à leurs anticipations.

L'indicateur de surprise, défini dans la Note de conjoncture de mars 2015 [11], se base sur le même principe mais utilise les réponses individuelles des entreprises. Pour cela, la réponse de chaque entreprise sur ses perspectives de production au mois de septembre, par exemple, est comparée à sa réponse sur la production passée au mois de décembre. Si une entreprise a modifié sa réponse entre les deux enquêtes, c'est probablement qu'un événement imprévisible, une surprise conjoncturelle, est apparue entre-temps. On s'attend à ce que plus les entreprises sont surprises à la hausse, plus la conjoncture s'est améliorée au cours des trois derniers mois. La comparaison des réponses sur les perspectives de production au mois m et la production passée au mois $m + 3$ permet de diviser les entreprises en neuf groupes en croisant les modalités possibles (hausse, stable et baisse) des réponses à ces deux questions.

Pour calculer l'indicateur de surprise, la méthode qui a été retenue consiste à effectuer une moyenne pondérée des neuf groupes précédents avec un poids :

- nul pour le groupe Stable-Stable ;
- de 1 (resp. -1) pour le groupe Stable-Hausse^a (resp. Stable - Baisse) ;
- de 2 (resp. -2) pour le groupe Hausse - Hausse (resp. Baisse - Baisse) ;
- de 3 (resp. -3) pour le groupe Baisse - Stable (resp. Hausse - Stable) ;
- de 4 (resp. -4) pour le groupe Baisse - Hausse (resp. Hausse - Baisse).

Ce système de pondération est plus performant pour prévoir lors des étalonnages de la production manufacturière qu'un système de pondération fondé sur une analyse en composantes principales.

Actuellement l'indicateur de surprise n'est pas publié à un niveau plus fin que celui de l'industrie manufacturière. Toutefois, il est également calculé au niveau A 17, sauf pour le secteur des matériels de transport (C4), et également dans l'automobile (CL1) et les autres matériels de transport (CL2).

^a. Une entreprise est dans le groupe « Stable - Hausse » si elle a répondu au mois m que sa production serait stable et si elle a répondu au mois $m + 3$ que sa production passée était en hausse.

1.3 Indicateurs quantitatifs de production et comptes nationaux

Les indices de production industrielle sont les seuls indicateurs quantitatifs de l'activité réelle que nous utilisons. Publiés de façon mensuelle à une échéance plus longue que les enquêtes de conjoncture, ils sont notamment utilisés par les comptes nationaux pour évaluer la production dans l'industrie manufacturière et ses sous-secteurs.

1.3.1 Les indices de production industrielle

L'Insee définit les indices de production industrielle (IPI) comme des « instruments statistiques qui permettent de suivre l'évolution mensuelle de l'activité industrielle de la France ». Utilisés pour évaluer la production manufacturière dans la majorité des branches de l'économie française, ce sont des indices mensuels, corrigés des variations saisonnières et des jours ouvrables. Les IPI sont calculés à partir des enquêtes mensuelles de branches et sont les indicateurs de production en volume privilégiés pour les comptes trimestriels.

1.3.2 Différences entre indices de production industrielle et production industrielle

L'évolution de la production manufacturière peut différer de celle de l'IPI principalement pour quatre raisons :

- Les corrections des variations saisonnières et des jours ouvrables de l'IPI et des indicateurs de production des comptes de production sont effectués indépendamment et à des niveaux d'agrégation différents. Elles peuvent donc différer.
- Pour obtenir les comptes trimestriels en volume, les IPI sont étalonnés et calés sur les comptes annuels. Le mécanisme du calage peut donc introduire une dynamique qui diffère de celle de l'IPI.
- L'IPI de la branche manufacturière est obtenu par agrégation des indices à un niveau plus fin pondérés par leur poids en terme de valeur ajoutée³ dans la production manufacturière. Pour les comptes de production, les pondérations utilisées sont les poids en terme de production.
- Pour certaines sous-branches manufacturières, à un niveau de détail fin, l'indicateur utilisé n'est pas l'IPI mais les indices de chiffre d'affaires.

1.3.3 La production manufacturière

La production est définie par l'Insee comme l' « activité exercée sous le contrôle et la responsabilité d'une unité institutionnelle qui combine des ressources en main-d'œuvre, capital et biens et services pour fabriquer des biens ou fournir des services, et résultat de cette activité ». Ainsi, la production manufacturière permet de mesurer les quantités produites par les entreprises exerçant leur activité dans l'industrie manufacturière. Utilisée pour le calcul du produit intérieur brut (PIB), elle représente environ 20 % de la production totale de biens et services.

Dans les comptes nationaux, annuels et trimestriels, la production exprimée en volume est publiée au prix de l'année précédente, chaînée et avec comme année de référence l'année 2010. Le principe des indices chaînés est de conserver les évolutions des volumes au prix de l'année précédente d'une année sur l'autre et de chaîner ces évolutions à partir de l'année de référence. Cela permet de prendre en compte la déformation de la structure des prix relatifs à l'année de référence et d'utiliser les estimations dans le cadre des séries temporelles. En effet, si les niveaux des volumes aux prix de l'année précédente n'étaient pas chaînés, l'évolution entre deux années consécutives comprendrait à la fois une évolution des prix et une évolution des volumes. Toutefois, les avantages liés au chaînage ont un coût : les séries ne sont plus sommables. Une méthodologie spécifique devra donc être réalisée pour obtenir une prévision de la production manufacturière à partir des étalonnages à un niveau plus fin (voir partie 2.3).

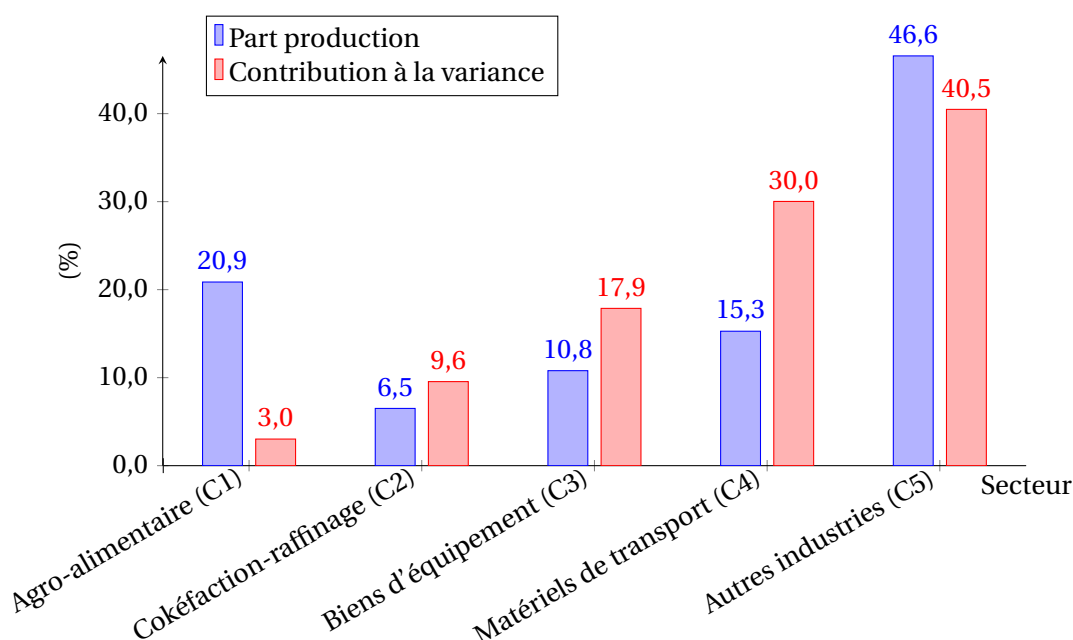
Pour le chaînage des séries trimestrielles, la méthode utilisée par l'Insee est la méthode du recouvrement annuel, où les comptes trimestriels sont évalués au prix moyens de l'année

3. La valeur ajoutée est égale à la différence entre la valeur finale de la production et la valeur des biens qui ont été consommés par le processus de production (appelée consommation intermédiaire).

précédente. Cette méthode a l'avantage de préserver la propriété d'additivité infra-annuelle : la somme des quatre comptes trimestriels d'une année est égale à son compte annuel. Cependant, une faible discontinuité est présente au premier trimestre car l'évolution en volume comprend à la fois un changement de volume et un changement d'année de référence pour les prix.

1.3.4 Volatilité des différents sous-secteurs de l'industrie manufacturière

Les niveaux A 17 et A 64 permettent un découpage de l'industrie manufacturière en fonction des activités des entreprises avec des poids, en terme de production, d'inégale importance. Un secteur dont le poids en terme de production est élevé dans l'industrie manufacturière peut donc avoir, compte tenu de faibles fluctuations, une faible contribution à la volatilité de l'évolution de la production manufacturière, et inversement. Le graphique 1.1 résume les poids en terme de production en valeur des différents secteurs par rapport à celui de l'industrie manufacturière ainsi que leurs contributions à la variance de l'évolution de la production manufacturière (voir annexe D pour la façon dont ces contributions). Bien que la branche de l'agro-alimentaire (C1) représente 21 % de la production manufacturière, elle ne contribue qu'à 3 % à la variance de cette dernière. À l'inverse, le secteur des matériels de transport (C4), qui présente des fluctuations relativement plus importantes que les autres, contribue à hauteur de 30 % à la variance de l'évolution de la production manufacturière alors qu'il ne représente que 15 % de la production du secteur manufacturier. De plus, dans les matériels de transport (C4), la volatilité est en grande partie due au secteur automobile (CL1) (voir tableau 1.2).



GRAPHIQUE 1.1 – Part de la production manufacturière et contribution à la variance des différents secteurs de l'industrie manufacturière au niveau A 17 (en %).

Sources : Insee et calculs de l'auteur.

Ainsi, si l'on veut améliorer la précision des étalonnages de la production manufacturière obtenus à partir des prévisions effectuées à un niveau plus fin, il faut améliorer les étalonnages dans les secteurs où la contribution à la variance est la plus forte : le secteur des autres industries (C5) et celui des matériels de transport (C4). En particulier, améliorer la précision des étalonnages

dans le secteur de l'agro-alimentaire aura peu d'influence sur la volatilité des prévisions de la production manufacturière.

Secteur	Poids en terme de production		Contribution à la variance	
	Dans l'ind. Manuf.	Dans le C4	Dans l'ind. Manuf.	Dans le C4
C4	15,3	100,0	31,2	100,0
CL1	7,3	48,0	26,4	84,5
CL2	7,9	52,0	4,8	15,4

TABLEAU 1.2 – Poids en terme de production et contribution à la variance du secteur automobile (CL1) et des autres matériels de transport (CL2) dans l'industrie manufacturière et dans les matériels de transport (C4) (en %).

Note de lecture : le secteur de l'automobile (CL1) contribue à hauteur de 26,4 % à la variance de l'évolution de la production manufacturière et à hauteur de 84,5 % à celle de la production des matériels de transport (C4).

Sources : Insee et calculs de l'auteur.

Conclusion

Pour décrire l'état de l'économie dans l'industrie manufacturière trois familles d'indicateurs conjoncturels sont disponibles : les soldes d'opinion issus des enquêtes de conjoncture, les indicateurs quantitatifs de production (les indices de production industrielle) et les comptes trimestriels de production. Tous ces indicateurs sont calculés et disponibles pour la branche de l'industrie manufacturière mais également à des niveaux plus fins. Ce découpage de l'industrie manufacturière est réalisé à partir de la nomenclature d'activité française. Le but de ce mémoire est d'utiliser ce découpage pour :

- détailler la prévision de la production manufacturière. Cela permet d'aider à la décision du scénario conjoncturel dans l'industrie manufacturière lors des notes de conjoncture ;
- améliorer les étalonnages actuels de la production manufacturière en utilisant des prévisions sectorielles.

Seront alors construits des modèles de prévision des évolutions des productions des comptes nationaux trimestriels uniquement pour le trimestre coïncident, sur lequel le scénario de la Note de conjoncture s'engage le plus. Pour ce faire, puisque les différentes sources de données sont utilisées avec des calendriers de diffusion différents nous nous placerons à chaque fois dans une configuration similaire aux réunions d'étalonnages de la Note de conjoncture de juin. Pour les différents étalonnages réalisés, le tableau 1.3 rappelle pour chaque secteur les variables utilisées.

Code	Libellé	Provenance des variables utilisées
C1	Industries agro-alimentaires	Insee, indicateur de surprise, Banque de France
C2	Cokéfaction-raffinage	Insee, indicateur de surprise
C3	Biens d'équipement	Insee, indicateur de surprise, Banque de France
C4	Matériels de Transport	Insee, Banque de France
CL1	Automobile	Insee, indicateur de surprise, Banque de France
CL2	Autres matériels de transport	Insee, indicateur de surprise, Banque de France
C5	Autres industries	Insee, indicateur de surprise, Banque de France

TABLEAU 1.3 – Provenances des variables utilisées dans les différents étalonnages aux niveaux A 17 et A 64.

Avant de pouvoir étalonner les modèles finaux aux niveaux A 17 ou A 64 de la Naf rév. 2 plusieurs choix méthodologiques ont été réalisés. Tout d’abord, au vu du grand nombre de variables explicatives disponibles, un algorithme approprié de sélection de variables doit être utilisé. Pour cela, la méthode choisie, dite du *General to specific* — général au spécifique — (Gets) permet d’obtenir un modèle final licite du point de vue économétrique, avec un pouvoir prédictif comparable à celui du modèle avec toutes les variables retenues, et le plus parcimonieux possible. Ensuite, puisque les soldes d’enquêtes sont obtenus mensuellement et que la production n’est connue que trimestriellement, la méthode de trimestrialisation des soldes joue un rôle important dans les étalonnages. Trois méthodes ont été retenues et diverses statistiques ont été utilisées afin de les comparer. Enfin, le but de ce mémoire étant également d’obtenir des prévisions de la production manufacturière, plusieurs méthodes d’agrégation des prévisions obtenues à un niveau plus fin sont utilisées.

2.1 Une sélection de modèles réalisée avec l’algorithme Gets

La procédure automatique de sélection a été développée par Hoovers et Perez [12] et mise en application par Krolzig et Hendry [14]. Elle repose sur la théorie que les données proviennent d’un processus générateur de données — *data generation process* — (DGP) non observé. L’idée de l’algorithme Gets est d’obtenir une approximation valide de ce DGP : on parle de modèle « congruent ».

L’ensemble de l’étude menée dans le cadre de ce mémoire a été réalisé avec le logiciel R. Pour la sélection des modèles, nous avons utilisé le package « gets » créé par Pretis, Reade et Sucarrat. En particulier, l’algorithme de sélection de ce package, proposé par Escibano et Sucarrat [16], est une version légèrement modifiée de l’algorithme initial proposé par Krolzig et Hendry.

2.1.1 Modèles recherchés par l’algorithme

Le modèle général recherché par l’algorithme de sélection est un modèle pouvant contenir à la fois une composante autorégressive et des variables exogènes. Mathématiquement, le modèle général peut être décrit à la date t à travers les équations (2.1) et (2.2).

$$y_t = \beta_0 + \sum_{m=1}^M \beta_m y_{t-m} + \sum_{n=1}^N \zeta_n x_{n,t} + \varepsilon_t \quad (2.1)$$

$$\varepsilon_t \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma^2) \quad (2.2)$$

Dans l’équation (2.1), M est le nombre de termes autorégressifs et N le nombre de variables explicatives fortement exogènes, qui peuvent être contemporaines ou retardées.

Pour simplifier l’interprétation économiques des modèles, tous ceux retenus dans notre étude ont des erreurs non autocorrélées, homoscédastiques et suivant une loi normale.

Sous des hypothèses standard (stabilité des coefficients, résidus gaussiens indépendants et identiquement distribués) tous les paramètres des équations (2.1) et (2.2) peuvent être estimés de manière consistante par les moindres carrés ordinaires.

2.1.2 Tests utilisés pendant l'algorithme

Le point de départ de la sélection automatique est un *general unrestricted model* — modèle général sans contrainte — (GUM) : c'est le modèle où toutes les variables explicatives sont présentes. Avant de trouver le modèle final, plusieurs tests sont mobilisés :

- Les tests de spécification : ils sont utilisés pour s'assurer que les hypothèses standard de spécificité du modèle sont vérifiées. Ici, trois tests sont utilisés :
 - le test de Jarque-Bera permettant de tester la normalité des résidus ;
 - le test de Ljung-Box sur les résidus pour vérifier l'absence de corrélation ;
 - le test de Ljung-Box sur les carrés des résidus pour vérifier l'homoscédasticité (*i.e* : variance des résidus constante dans le temps) de ces derniers.
- Le test de nullité des coefficients : il s'agit du test de Student qui est employé pour savoir s'il faut ou non supprimer une variable explicative.
- Le test englobant parcimonieux (aussi appelé *backtest*) : c'est un test joint du modèle final contre le modèle GUM. Il teste si les régresseurs supprimés sont conjointement significativement nuls. C'est un test de Wald qui est utilisé.

Les tests de spécification sont des tests importants lors d'une analyse économétrique. En effet, la présence d'autocorrélation ou d'hétéroscédasticité (*i.e* : variance non constante dans le temps) peut notamment être le signe de l'oubli d'une variable explicative. De plus, ce sont ces tests qui garantissent que les estimations sont robustes¹.

Tous ces tests ont été effectués au seuil de 5 %. Pour les tests de spécification, la correction de Bonferroni est réalisée (voir encadré 2.1) : chaque test est réalisé au niveau 1,7 %. De plus, tous les tests de Ljung-Box sont effectués à l'ordre 5 [5].

ENCADRÉ 2.1 - La correction de Bonferroni

La correction de Bonferroni est une méthode pour corriger les seuils de significativité des tests lors de comparaisons multiples.

Déoulant de l'inégalité de Bonferroni, cette correction tire naturellement son nom du mathématicien du XX^e siècle Carlo Emilio Bonferroni. Celle-ci est un théorème stipulant que la probabilité d'une conjonction d'événements, A_i , ne peut dépasser la somme des probabilités des événements conjoints :

$$\mathbb{P}\left(\bigcup_i A_i\right) \leq \sum_i \mathbb{P}(A_i)$$

Si l'on suppose maintenant que l'on a k événements et la probabilité de chaque événement est majorée par γ , l'inégalité permet de déduire que la probabilité de l'union de tous ces événements est inférieure à $k\gamma$.

Ainsi, pour valider un test global au seuil α comportant k tests distincts, il suffit que chacun de ces tests soit validé au seuil $\frac{\alpha}{k}$.

1. Une estimation robuste est une estimation qui n'est pas altérée par une petite modification dans les données ou dans les paramètres du modèle choisi pour l'estimation.

2.1.3 Principe de l'algorithme

La sélection automatique des variables peut se résumer en 3 grandes étapes.

- 1ère étape : Estimation et test du modèle général (voir diagramme 2.1)
 - Si un test de spécificité n'est pas validé pour le GUM alors l'algorithme s'arrête et aucun modèle n'est sélectionné.
 - Si toutes les variables explicatives sont significatives alors le GUM est le modèle final et l'algorithme s'arrête.
 - Sinon le nombre de modèles admissibles est égal au nombre de variables non significatives dans le GUM (noté k dans les diagrammes 1 et 2). Le modèle admissible n°1 est construit en enlevant la première variable non significatives du modèle GUM. Le modèle admissible n°2 est construit en enlevant la deuxième variable non significatives du modèle GUM, etc.

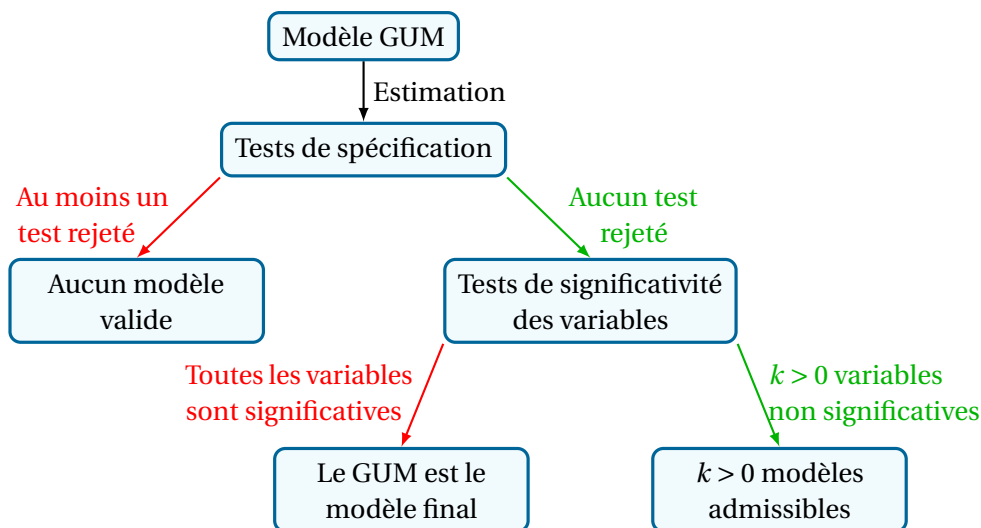


DIAGRAMME 2.1 – Estimation et test du modèle général.

- 2ème étape : Réduction du modèle GUM et recherche des modèles terminaux (voir diagramme 2.2).

Pour chaque modèle admissible les simplifications suivantes sont réalisées :

- Si la suppression de la première variable (comme définie dans l'étape 1) entraîne un problème de spécification (autocorrélation, hétéroscédasticité ou non normalité des résidus) alors le modèle étudié n'est pas retenu².
- Sinon, on enlève une par une les variables non significatives restantes dont la p-valeur est la plus grande.
- Si la suppression d'une variable explicative introduit un problème de spécification ou le rejet du *backtest* (voir partie refsubsect :testUseAlgo) alors la variable est réintroduite.
- Lorsque plus aucune variable ne peut être supprimée, alors le modèle est considéré comme terminal.

2. Notons qu'ici le *backtest* est forcément accepté. En effet, le test de Wald pour la suppression d'une variable est équivalent au test de Student de nullité d'un coefficient.

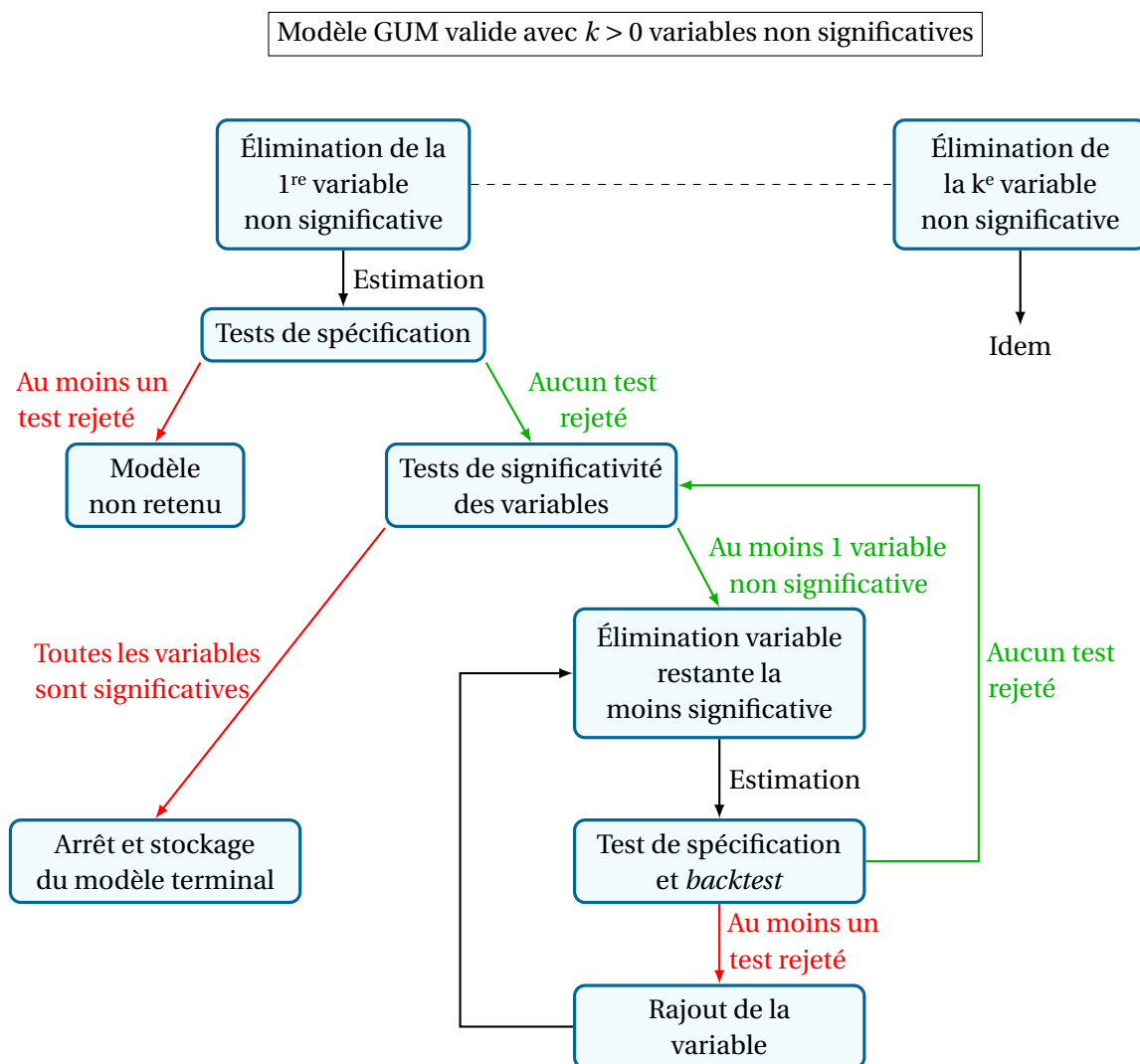


DIAGRAMME 2.2 – Réduction du modèle GUM et recherche des modèles terminaux.

— 3^{ème} étape : Détermination du modèle final (voir diagramme 2.3)

- Si aucun modèle terminal n'est trouvé alors le modèle final sera le modèle GUM.
- Si un seul modèle terminal est sélectionné alors ce sera le modèle final.
- Sinon la sélection du modèle final se fait avec un critère d'information. Ici, nous utilisons le critère d'information d'Akaike (AIC). Les modèles terminaux étant mutuellement englobants³, ils ont tous, heuristiquement, un critère d'information proche : le choix de ce dernier ne changera donc pas fondamentalement le modèle final. Toutefois, l'AIC est un critère intéressant pour la sélection de modèle lorsque l'on a un objectif de prévision (voir encadré 2.2).

3. L'approche de Gets est basée sur la théorie de l'englobement – *encompassing* en anglais –. Un modèle englobe un autre si la prévision du second modèle n'apporte aucune information conditionnellement à la prévision du premier modèle, ou à l'inverse explique les erreurs du premier modèle.

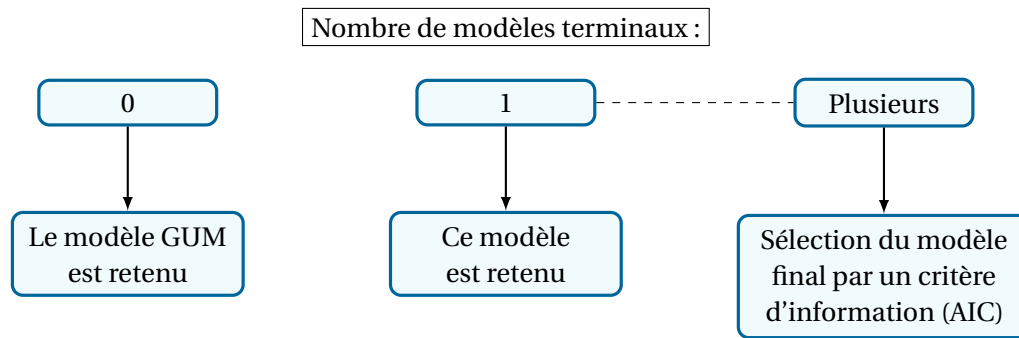


DIAGRAMME 2.3 – Détermination du modèle final.

ENCADRÉ 2.2 - Le critère d'information d'Akaike

Le critère d'information d'Akaike (AIC) est une mesure de la qualité d'un modèle statistique proposée par le statisticien japonais Hirotugu Akaike en 1973. Il est défini de la façon suivante :

$$AIC = -2\ln(\mathfrak{L}) + 2p$$

Où \mathfrak{L} est le maximum de vraisemblance du modèle et p le nombre de paramètres utilisés. Ce critère permet donc un compromis entre qualité d'ajustement et complexité du modèle (en terme de nombre de variables explicatives). De plus, asymptotiquement, minimiser l'AIC est équivalent à minimiser la statistique de validation croisée, qui est une mesure de la robustesse des prévisions (voir encadré 2.4). L'une des propriétés de l'AIC, permettant de justifier son emploi dans l'algorithme Gets, est qu'il peut être utilisé pour comparer plus de deux modèles et que les modèles comparés ne doivent pas nécessairement être emboîtés ^a[3].

^a. Deux modèles sont emboîtés s'ils contiennent les variables explicatives et que l'un a au moins un terme additionnel.

2.1.4 Détection des points atypiques

Un point atypique, ou *outlier*, est une observation qui se trouve « loin » de la tendance générale des autres observations. Leur présence peut provenir de chocs extérieurs (e.g : la crise financière de 2008) et avoir un impact important sur la production manufacturière, comme sur les variables d'enquête. Les effets des *outliers* peuvent parfois être captés par la présence de certaines variables explicatives dans les étalonnages. Ces points n'auraient alors pas d'impact important sur les estimations. Ce n'est toutefois pas toujours le cas et leur présence peut notamment être la cause des problèmes de spécification et introduire des biais dans les prévisions. Ainsi, un traitement spécifique doit être réalisé et cela se fait simplement en ajoutant dans les variables explicatives une indicatrice associée à la date de l'*outlier*.

Le nombre, la date et l'ampleur des ruptures dans un modèle étant généralement inconnus, il est nécessaire d'avoir une procédure permettant de détecter ces ruptures à tout point de l'échantillon parmi un grand nombre de candidats. La méthode que nous avons utilisée s'appelle *impulse-indicator saturation* (IIS). Le principe est simple : ajouter dans les régresseurs une indicatrice temporelle pour chaque date de la période d'estimation, effectuer l'algorithme Gets sur ces indicatrices. Toutefois, puisqu'il y a alors plus de paramètres à estimer que d'observations, l'ajout de ces indicatrices se fait en plusieurs étapes : tout d'abord, on partitionne l'ensemble des

indicatrices en plusieurs « blocs » dépendant de la période d'estimation. Ensuite, pour chacun de ces blocs l'algorithme Gets est réalisé et les indicatrices éventuellement retenues dans chaque bloc sont enregistrées⁴. Enfin, l'algorithme Gets est appliquée une nouvelle fois avec toutes les indicatrices retenues dans chaque bloc⁵.

La méthode IIS a la capacité de détecter et corriger une mauvaise spécification d'un modèle due à la non constance des coefficients dans le temps [8][10]. Elle a également la capacité de détecter des ruptures dues à la présence d'*outliers*. C'est pourquoi cette méthode est utilisée avant la procédure de sélection automatique des variables. Par ailleurs, la construction des tests d'échec prédictif et de stabilité de Chow, qui sont des tests de stabilité des coefficients, est basée sur une méthode similaire à la méthode IIS. Ainsi, l'utilisation des indicatrices permet une stabilité dans et hors de l'échantillon.

2.2 Trois spécifications utilisées pour la construction des étalonnages

Afin d'effectuer des prévisions à court terme de variables macroéconomiques, de nombreuses méthodes sont proposées dans la littérature (en grande partie résumées par Castle *et al*[6]). Néanmoins, deux approches ayant de bonnes performances prédictives sont principalement utilisées. La première consiste à effectuer des prévisions à partir de modèles à facteurs dynamiques. Dans celle-ci, la dynamique d'un nombre élevé de variables est résumée en un nombre plus réduit de facteurs communs utilisés comme variables explicatives dans le modèle. L'inconvénient de cette méthode est que les résultats sont difficiles à interpréter économiquement. La seconde approche, également connue sous le nom de « *bridge models* », consiste à construire des équations linéaires d'étalonnage à partir d'informations mensuelles. Plus populaire dans les banques centrales et les instituts nationaux de statistique du fait de la facilité des interprétations des fluctuations économiques, c'est la méthode que nous avons retenue.

2.2.1 Soldes utilisés pour les différentes spécifications

Les modèles présentés dans ce mémoire seront utilisés uniquement pour des prévisions du trimestre coïncident. Comme précisé dans la partie 1.1.2, les différentes variables explicatives ne sont pas disponibles jusqu'à la même date. En nous plaçant dans une configuration similaire aux réunions d'étalonnage de la Note de conjoncture de juin, pour les prévisions du trimestre coïncident nous supposons donc que :

- les IPI sont connus jusqu'au troisième mois du trimestre précédent ;
- les soldes de l'enquête Banque de France ne sont connus que jusqu'au premier mois du trimestre coïncident ;
- les soldes de l'enquête Insee ne sont connus que jusqu'au deuxième mois du trimestre coïncident.

Les différents étalonnages de la production doivent donc prendre en compte que, lors de l'exercice de prévision, ces variables ne sont pas connues sur l'ensemble du trimestre coïncident.

Seuls les soldes de l'enquête Insee sont utilisés dans tous les étalonnages (voir conclusion du chapitre 1). Lors de la sélection des modèles, certains soldes seront également présents de manière « retardée ». C'est-à-dire que, pour les prévisions au trimestre T , sera par exemple à la fois inclus dans la sélection de modèles le solde calculé au trimestre T et celui calculé au

4. Effectuer une estimation avec une indicatrice temporelle sur une sous-période est équivalent à effectuer une estimation en ignorant cette sous-période.

5. Si l'on souhaite utiliser l'algorithme de sélection Gets et que l'on a plus de variables explicatives que d'observations, la même méthode est à utiliser en divisant ces variables en plusieurs catégories.

trimestre $T - 1$. Nous avons choisi d'inclure des retards uniquement pour les soldes décrivant des anticipations :

- les soldes sur les niveaux des carnets de commandes ;
- les soldes sur les perspectives d'évolution de la production ou des carnets de commandes.

Lorsque que dans un modèle final apparaissent à la fois une variable et son retard, nous avons testé l'égalité en valeur absolue des estimations des paramètres associés à ces deux variables. Si l'égalité est vérifiée, ces deux variables sont remplacées par la variable en différence : nous noterons alors par exemple Δoscd_T la variable en différence à la date T du solde sur les carnets de commandes définie par :

$$\Delta \text{oscd}_T = \text{oscd}_T - \text{oscd}_{T-1}$$

De plus, dans les étalonnages au niveau A 17 et A 64, des retards du taux de croissance de la production, d'au plus deux trimestres, sont parfois également retenus comme variables explicatives pour palier des problèmes de spécification.

Pour faciliter l'analyse et l'interprétation des prévisions, pour les étalonnages dans un secteur nous utiliserons uniquement les variables de ce secteur⁶. C'est-à-dire que pour les étalonnages de la production dans les matériels de transport, par exemple, nous utiliserons uniquement les soldes agrégés au niveau des matériels de transport sans utiliser les soldes calculés dans l'automobile ou les autres matériels de transport.

2.2.2 Méthode 1 : étalonnages avec les soldes bloqués et l'acquis d'IPI

Dans les étalonnages actuellement utilisés dans le processus de prévision de la production manufacturière, les soldes sont trimestrialisés sans modifier les niveaux mensuels. Ainsi, pour chaque solde mensuelles, trois séries trimestrielles sont construites en fonction de la place du mois dans le trimestre. La première série est constituée des valeurs prises par le solde initial lors de chaque premier mois dans le trimestre : nous parlerons alors de solde bloqué au mois 1. Le même raisonnement est utilisé pour les autres mois du trimestre. Les soldes d'opinion seront bloqués au dernier mois connu lors de l'exercice de prévision : c'est le premier mois du trimestre pour ceux publiés par l'enquête Banque de France et c'est le deuxième mois du trimestre pour ceux publiés par l'Insee.

L'IPI, quant à lui, est transformé en acquis de croissance sur tout le trimestre : le taux de croissance trimestriel est calculé en prolongeant le niveau de l'IPI mensuel sur le trimestre par la dernière valeur connue. Dans notre cas, puisque le dernier point connu de l'IPI est celui du dernier mois du trimestre précédent, nous parlerons d'acquis d'IPI au mois 0. C'est-à-dire que l'acquis d'IPI au mois 0 pour le trimestre T correspond au taux de croissance de l'IPI entre le trimestre $T - 1$ et le trimestre T que l'on obtiendrait si l'IPI demeurait jusqu'à la fin du trimestre T au niveau du troisième mois du trimestre $T - 1$.

Nous appellerons par la suite cette méthode « méthode 1 ».

2.2.3 Méthodes 2 et 3 : en prévoyant l'IPI sur tout le trimestre et sans l'IPI

Deux autres méthodes pour l'étalonnage de la production ont également été retenues.

La deuxième méthode est basée sur la construction d'une prévision en deux étapes :

6. À un niveau d'agrégation donné, cela est cohérent avec le fait que l'information fournie par les entreprises d'un secteur ne reflète que leur opinion et non celle des autres secteurs.

- La première étape consiste à étalonner le taux de croissance mensuel de l'IPI sur tout le trimestre coïncident. Cet étalonnage est construit avec les mêmes soldes que ceux utilisés pour prévoir les comptes de production. Les soldes décrivant des anticipations seront également considérés avec un retard cohérent avec l'horizon temporel de la question (un mois pour les soldes de la Banque de France et trois mois pour les soldes de l'Insee). Des retards de l'IPI d'au plus trois mois sont également ajoutés pour palier des problèmes de spécification. De ces prévisions nous déduisons un taux de croissance trimestriel de l'IPI (voir encadré 2.3).
- Dans la seconde étape, un étalonnage de la production est construit. Celui-ci utilise le taux de croissance trimestriel de l'IPI, connu sur tout le trimestre coïncident, et les soldes d'opinion trimestrialisés (avec une formule homogène à celle utilisée pour transformer le taux de croissance mensuel de l'IPI en taux de croissance trimestriel, voir encadré 2.3).

C'est cette méthode en deux étapes que nous appellerons « méthode 2 ».

La troisième méthode est proche de la première mais n'utilise pas l'IPI pour les étalonnages de la production. Les soldes d'enquêtes sont trimestrialisés de la même façon que pour la deuxième méthode. C'est la méthode que nous appellerons « méthode 3 ».

ENCADRÉ 2.3 - Trimestrialisation des soldes

Pour passer d'un taux de croissance mensuel à un taux de croissance trimestriel, nous utilisons l'approximation suivante résultant d'un développement limité :

$$IPI_T^{GT} = \frac{1 \times IPI_{T-1,m2}^{GM} + 2 \times IPI_{T-1,m3}^{GM} + 3 \times IPI_{T,m1}^{GM} + 2 \times IPI_{T,m2}^{GM} + 1 \times IPI_{T,m3}^{GM}}{3}$$

Où IPI_T^{GT} est le taux de croissance trimestriel de l'IPI au trimestre T et $IPI_{T,m1}^{GM}$ le taux de croissance mensuel de l'IPI au premier mois du trimestre T .

Dans les étalonnages de la production des méthodes 2 et 3, les soldes d'enquêtes sont trimestrialisés avec la même formule :

$$solde_T = \frac{1 \times solde_{T-1,m2} + 2 \times solde_{T-1,m3} + 3 \times solde_{T,m1} + 2 \times solde_{T,m2} + 1 \times solde_{T,m3}}{3}$$

Prolongement des soldes

Pour les prévisions des comptes de production trimestriels et de l'IPI et pour la trimestrialisation des soldes, ces derniers doivent être connus sur tout le trimestre coïncident. Puisque ce n'est pas le cas dans le cadre des prévisions réalisées dans ce mémoire, les soldes sont prolongés sur le trimestre coïncident par la dernière valeur connue.

2.3 Différentes façons d'agréger les données...

Afin d'agréger les différentes prévisions réalisées au niveau A 17 ou A 64 de la Naf rév 2 pour obtenir une prévision de l'évolution de la production manufacturière, plusieurs méthodologies sont possibles. Nous avons décidé de comparer trois d'entre elles :

- La première est une approche comptable basée sur la construction de la production en volumes chaînés et permet notamment de prendre en compte le changement de prix dans le chaînage de la production aux premiers trimestres. Toutefois, cette approche manque de flexibilité pour des prévisions à un horizon plus grand que le trimestre coïncident. Cette méthode sera par la suite qualifiée de méthode « comptable ».

- La deuxième est une approximation de la première formule et consiste à pondérer simplement chaque secteur son poids en terme de production. Cette méthode sera par la suite appelée « l'approximation ».
- La troisième est une approche plus statistique proposée par Hyndman basée sur une combinaison plus cohérente de toutes les prévisions à des niveaux d'agrégation différents. Nous parlerons de la méthode « d'Hyndman ».

La première méthode d'agrégation fournissant des résultats comptablement exacts, c'est celle que nous utiliserons pour confronter nos trois méthodes de construction des étalonnages. Les deux autres façons d'agréger les prévisions seront comparées avec la méthode de construction des étalonnages que nous conseillons d'adopter. Ces trois méthodes d'agrégation sont également appliquées pour obtenir une prévision de la production dans le secteur des matériels de transport (C4) à partir des prévisions obtenues dans le secteur de l'automobile (CL1) et des autres matériels de transport (CL2) lors des étalonnages réalisées au niveau A 64.

2.3.1 ... Par une approche comptable

L'approche comptable est la méthode d'agrégation recommandée par les comptes nationaux pour l'analyse statistique et économique des évolutions en volumes. Elle utilise à la fois les comptes de production en volumes chaînés et en valeurs, leurs propriétés respectives et les relations entre les différents comptes de production pour calculer la contribution de chaque agrégat à l'évolution de la production manufacturière en volumes chaînés. La somme de ces contributions permet d'obtenir l'évolution de la production manufacturière en volumes chaînés. De plus, pour la contribution de chaque agrégat, un terme spécifique permet de corriger le changement de pondération des prix au premier trimestre. Ce terme additionnel s'annule si l'on additionne toutes les contributions.

La formule exacte utilisée est présentée dans l'annexe E. Dans cette formule apparaissent les productions en valeurs et en volumes chaînés au trimestre précédent. Si l'on souhaite réaliser des prévisions à un horizon plus élevé que le trimestre coïncident, cela suppose de réaliser des inférences sur ces comptes.

2.3.2 ... Par une approche naturelle

Une autre façon d'agréger les prévisions est de pondérer simplement les évolutions de chaque secteur par leur poids en terme de production dans l'industrie manufacturière. Ce poids est égal à la part de la production annuelle en valeur de l'année précédente du secteur étudiée dans l'industrie manufacturière⁷. Dans les faits, ce choix de pondération est une approximation de la méthode d'agrégation conseillée par les comptes trimestriels (voir annexe E).

Cette méthode d'agrégation est plus souple que la méthode comptable car elle ne fait intervenir que les comptes de production annuels. Toutefois, cette méthode ne permet pas une étude précise des contributions d'un secteur. Néanmoins, si l'objectif est d'avoir une prévision de la production manufacturière et non pas d'étudier précisément les contributions, ce qui est ici le cas, cette approximation ne génère que très peu d'erreur.

2.3.3 ... Par une approche statistique

La méthode proposée par Hyndman n'est pas une méthode d'agrégation des prévisions des sous-secteurs de l'industrie manufacturière mais plutôt une méthode de mise en cohérence de toutes les prévisions [13]. En d'autres termes, nous entendons notamment que la prévision de

7. Les comptes de production étant utilisés en valeurs, la somme des pondérations est égale à 1.

la production manufacturière n'est pas obtenue par agrégation des prévisions réalisées à un niveau plus fin mais par un étalonnage spécifique. Ensuite, toutes les prévisions sont combinées en exploitant les interactions entre tous les secteurs. Cela permet de s'assurer de la cohérence entre les prévisions obtenues pour la production manufacturière et celles obtenues pour ses différents sous-secteurs. Par exemple, la somme pondérée des prévisions dans les secteurs de l'automobile (CL1) et des autres matériels de transport (CL2) serait alors égale à la prévision dans le secteur des matériels de transport (C4). En outre, cela permet une construction facilitée des intervalles de confiance.

L'annexe E fournit plus de détails concernant cette méthode et son application dans ce mémoire.

2.4 Outils de comparaison des modèles

Dans le cadre des exercices de prévision, les modèles finaux ne doivent pas avoir de problème de spécification. Aucun des étalonnages ne présente de problème de spécification d'après l'algorithme de sélection des variables utilisé. Ces modèles doivent également fournir de bonnes prévisions « dans l'échantillon » et « hors échantillon » c'est-à-dire en dehors de la période utilisée pour l'estimation des modèles. Un indicateur de performance des prévisions dans et hors échantillon est la racine carrée de l'erreur quadratique moyenne — *root mean square error* — (RMSE).

ENCADRÉ 2.4 - La validation croisée et le RMSE pseudo temps-réel

La validation croisée est un moyen de mesurer la qualité prédictive d'un modèle statistique. Une méthode simple de validation croisée est le *leave-one-out cross-validation*. Dans le cadre des séries temporelles, en supposant que l'on possède T observation y_1, \dots, y_T , cette méthode est réalisée en 3 étapes :

1. Pour une date t , le modèle est estimé à partir de y_1, \dots, y_t et on note \hat{y}_{t+1} la prévision obtenue pour la date $t + 1$. On calcule alors l'erreur de prévision associée à cette date ($e_{t+1}^* = y_{t+1} - \hat{y}_{t+1}$).
2. La première étape est répétée pour t variant de m à $T - 1$, où m est le nombre minimum d'observations nécessaires pour évaluer le modèle.
3. On calcule l'erreur quadratique moyenne associée à e_{m+1}^*, \dots, e_T^* , qui correspond à la statistique de validation croisée :

$$\frac{1}{T - (m + 1)} \sum_{t=m+1}^T e_t^{*2}$$

La racine carrée de cette statistique est également appelée RMSE pseudo temps-réel.

Pour évaluer les qualités prédictives des modèles, plusieurs scénarios sont envisagés :

- Une prévision statique à partir du modèle estimé sur toute la période étudiée avec des prévisions réalisées en supposant que toutes les données sont disponibles. Dans le cadre de ce mémoire, ce type de prévision n'est pas représentatif de la réalité car pour la prévision du trimestre coïncident nous ne connaissons pas les valeurs des soldes (et en particulier de l'IPI) sur tout le trimestre. Ce scénario ne sera donc pas utilisé pour comparer les modèles. Le RMSE associé à ces prévisions sera appelé « RMSE complet ».
- Une prévision statique à partir du modèle estimé sur toute la période étudiée en supposant que pour chaque trimestre, les prévisions sont réalisées à partir des données connues dans la configuration similaire à la Note de conjoncture de juin (voir partie 2.2.1). C'est

dans ce cadre là que nous parlerons de prévision dans l'échantillon et donc de « RMSE dans l'échantillon ». Pour les étalonnages construits par la méthode des soldes bloqués, les prévisions associées aux RMSE complet et dans l'échantillon sont les mêmes.

- Une prévision dynamique permettant d'appréhender les comportements des étalonnages lors de plusieurs exercice de prévision depuis le premier trimestre 2000. Pour cela, la première prévision est réalisée de la façon suivante : les modèles sont estimés jusqu'au quatrième trimestre 1999 et des prévisions du trimestre coïncident (le premier trimestre 2000), sont réalisées. Cette étape est répétée jusqu'au dernier trimestre connu. On parle alors de prévision en pseudo temps-réel. En effet, les soldes utilisés pour prévoir le premier trimestre 2000, par exemple, sont des soldes d'enquêtes définitifs alors qu'en février 2000, les dernières valeurs des soldes étaient provisoires. Les soldes d'enquête étant peu révisés, les prévisions ainsi réalisées sont proches de celles qui auraient été obtenues à l'époque. Le RMSE ici calculé sera égal à la statistique de validation croisée (voir encadré 2.4) et sera appelé « RMSE pseudo temps-réel ».

Le RMSE complet, bien qu'il ne sera pas utilisé pour comparer les modèles, donnera une information sur la qualité d'ajustement des modèles sur le passé. Le RMSE dans l'échantillon permet quant à lui d'avoir un ordre de grandeur de l'intervalle de confiance des prévisions alors que le RMSE pseudo temps-réel donne un ordre de grandeur de l'écart-type de l'erreur de prévision lors des différents étalonnages associés aux notes de conjoncture.

Toutefois, ces statistiques ne suffisent pas pour décider qu'un modèle a significativement un meilleur pouvoir prédictif qu'un autre. Pour comparer les performances prédictives des différentes méthodes d'étalonnages pour les différents types de prévision, nous avons utilisé le test proposé par Diebold et Mariano (voir encadré 2.5). Il a été effectué au seuil de 5 %.

ENCADRÉ 2.5 - Le test de Diebold et Mariano

Le test de Diebold et Mariano (DM) est un test proposé en 1995 par les statisticiens américains Francis Diebold et Roberto Mariano. Il a été conçu pour comparer des prévisions et non pas pour comparer des modèles [7].

Ce test se base sur la comparaison des performances prédictives entre deux séries d'erreurs en terme d'erreur quadratique moyenne. Plus précisément, si l'on note $(e_{1t})_{t=1\dots T}$ et $(e_{2t})_{t=1\dots T}$ deux séries d'erreurs de prévision à la même horizon de prévision (égale à 1 dans notre cas) issues de deux modèles de prévision, l'hypothèse nulle d'égalité des performances prédictives s'écrit :

$$(H_0) : \mathbb{E}[d_t] = 0 \text{ avec } d_t = e_{1t}^2 - e_{2t}^2$$

Plusieurs hypothèses alternatives sont possibles : les deux modèles n'ont pas les mêmes performances prédictives ($\mathbb{E}[d_t] \neq 0$) ou un des deux modèles a de meilleures performances prédictives que l'autre ($\mathbb{E}[d_t] < 0$ ou $\mathbb{E}[d_t] > 0$).

Diebold et Mariano proposent la statistique asymptotiquement gaussienne $S = \frac{\bar{d}}{\hat{\sigma}_{\bar{d}}}$ où

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t \text{ et où } \hat{\sigma}_{\bar{d}} \text{ est une estimation consistante de l'écart-type de } \bar{d}.$$

Un des avantages du test de Diebold et Mariano est qu'il est robuste aux hypothèses sur les résidus des modèles : il reste valide lorsque les résidus sont autocorrélés, hétéroscédastiques et ne suivent pas une loi normale.

Conclusion

L'objectif de ce mémoire est double : obtenir des prévisions de la production des comptes trimestriels à un niveau plus fin que l'industrie manufacturière et regarder dans quelle mesure ces nouvelles prévisions permettraient d'améliorer la qualité de la prévision de l'évolution de la production manufacturière.

Pour prévoir l'évolution de la production manufacturière, les étalonnages des productions des sous-secteurs sont réalisées à deux niveaux de la nomenclature d'activité française : au niveau A 17 et au niveau A 64 pour le secteur des matériels de transport (C4). Le niveau A 17 permet un découpage de l'industrie en 5 secteurs (appelés C1 à C5) et le niveau A 64 pour le secteur des matériels de transport permet un découpage du secteur des matériels de transport (C4) en deux secteurs (CL1 et CL2). Les différentes prévisions sont agrégées pour obtenir une prévision de la production manufacturière et également une prévision pour le secteur des matériels de transport (C4) au niveau A 64. Pour cela, trois méthodes sont utilisées : une méthode « comptable » permettant d'avoir un agrégation comptablement exacte ; une méthode d'« approximation » qui est une approximation de la méthode « comptable » et qui consiste à pondérer les évolutions de chaque secteur par leur poids en terme de production annuelle dans l'industrie manufacturière ; et une méthode plus complexe, proposée par Hyndman, effectuant une combinaison cohérente de toutes les prévisions obtenues à des niveaux d'agrégation différents.

Pour la construction des étalonnages des productions des comptes trimestriels du trimestre coïncident, trois méthodes ont été retenues. La première méthode est la méthode des « soldes bloqués » : pour chaque trimestre, les indicateurs d'enquête sont étudiés pour un mois fixé du trimestre. Les soldes d'opinion de l'enquête Banque de France sont bloqués au mois 1 et ceux de l'enquête Insee sont bloqués au mois 2. Puisqu'il n'est pas connu sur le trimestre coïncident, l'IPI est transformé en acquis de croissance au mois 0 : un taux de croissance trimestriel est construit en le prolongeant sur le trimestre coïncident par la dernière valeur connue. La deuxième méthode est réalisée en deux étapes : l'IPI est tout d'abord étalonné mensuellement sur tout le trimestre coïncident à partir des soldes d'enquête, la production est ensuite étalonnée à partir du nouvel IPI et des soldes d'enquêtes. Ici, tous les soldes sont prolongés sur le trimestre coïncident par la dernière valeur connue. Dans la troisième méthode, les soldes sont prolongés comme dans la deuxième méthode mais l'IPI n'est plus utilisé pour les étalonnages des productions.

Lorsque qu'une méthode de construction des étalonnages a été choisie, le choix des variables explicatives est réalisée en deux étapes :

- la méthode IIS de détection des *outliers* est utilisée ;
- l'algorithme Gets est appliqué pour sélectionner les variables explicatives et les éventuelles indicatrices temporelles permettant de corriger les *outliers*.

À l'issue de ces deux étapes, les modèles d'étalonnage finaux sont obtenus et toutes les prévisions peuvent être réalisées.

Pour chaque secteur et pour chaque méthode de construction des étalonnages, l'utilisation successive de la méthode IIS et de l'algorithme Gets permet d'avoir un modèle final correctement spécifié. Ces différents modèles seront tout d'abord évalués et comparés par rapport à leurs qualités prédictives à l'aide de l'analyse des différents RMSE et du test de Diebold et Mariano. Ils seront ensuite confrontés dans un premier temps pour les étalonnages obtenus au niveau A 17 ainsi que pour la prévision de la production manufacturière obtenue par agrégation ; et dans un second temps pour les étalonnages obtenus au niveau A 64 permettant une nouvelle prévision de la production dans le secteur des matériels de transport (C4) et dans l'industrie manufacturière. Cela nous permettra de conseiller une démarche de construction des étalonnages à adopter pour les prochaines prévisions de la production manufacturière lors des réunions d'étalonnage. C'est cette méthode qui sera ensuite utilisée pour opposer nos trois moyens d'agréger les prévisions. Enfin, nous présenterons plus en détails les résultats en terme de prévision pour la méthode de construction des étalonnages et la méthode d'agrégation que nous recommandons d'utiliser. Pour cela, après un rapide aperçu des prévisions obtenus depuis le premier trimestre 2014, nous reviendrons sur la prévision retenue pour la production manufacturière du deuxième trimestre 2015 dans la Note de conjoncture de juin 2015.

3.1 Résultats des différents étalonnages

Après une description générale des équations de prévision nous comparerons ici les trois méthodes de construction des étalonnages aux niveaux A 17 et A 64. Les évolutions des productions dans le secteur manufacturier (aux niveaux A 17 et A 64) et dans celui des matériels de transport (C4) (au niveau A 64) sont calculées par la méthode d'agrégation comptable. Les différentes équations d'étalonnage obtenues avec la méthode Gets sont présentées dans l'annexe F.

3.1.1 Description générale des équations de prévision

Dans tous les modèles toutes les variables sont significatives. Il n'y a donc pas de problème de spécification dû à l'absence d'une variable explicative importante et corrigé par la présence d'une variable non significative.

La méthode IIS de détection de la présence de points atypiques est cohérente avec les grands mouvements de l'économie : la présence d'*outliers*, corrigée par l'ajout d'indicatrices temporelles, se retrouvent pendant les périodes de crise. En effet, ces points se retrouvent au début de l'année 1991 à la fin de la guerre du Golfe, où la production est sous-estimée par les étalonnages, et pendant la dernière crise financière, où la production est surestimée par les étalonnages.

De manière générale, l'IPI au mois 0 donne peu d'informations sur la production du trimestre coïncident : dans la méthode avec les soldes bloqués (méthode 1), le solde sur l'acquis d'IPI au mois 0 n'est sélectionné que dans les secteurs de l'agro-alimentaire (C1) et des matériels de

transport (C4). De plus, lorsqu'il est connu sur tout le trimestre coïncident, les soldes d'enquête influent très peu sur l'évolution de la production. En effet, pour tous les étalonnages construits de cette façon, les variables provenant d'enquêtes de conjoncture contribuent peu à la variance du taux de croissance estimé de la production (tableau 3.1) : la variabilité de la prévision est surtout expliquée par l'IPI. Cela provient de la méthode de construction de la production trimestrielle par les comptes trimestriels qui est obtenue à l'aide du seul IPI par une méthode d'étalonnage-calage.

Secteur étudié		C1	C2	C3	C4	CL1	CL2	C5
Contribution	de l'IPI	84,3	98,9	74,0	94,0		96,6	82,3
à la variance	des soldes	15,7	1,1	26,0	6,0		3,4	17,7

TABLEAU 3.1 – Contribution à la variance (en %) de l'évolution de la production estimée de l'IPI et des soldes d'enquête dans les étalonnages avec l'IPI connu sur tout le trimestre (méthode 2).

Note de lecture : aucun modèle correctement spécifié n'a pu être trouvé pour la prévision de l'IPI dans l'automobile (CL1).
Sources : Insee, Banque de France et calculs de l'auteur.

Dans le secteur de la cokéfaction-raffinage (C2), la volatilité de la série, qui se traduit par de forts à-coups, n'est pas captée par les soldes d'enquêtes et l'acquis d'IPI au mois 0. Cependant, la connaissance de l'IPI sur le trimestre coïncident permet de reproduire ces à-coups : cela s'observe par la présence de retard du taux de croissance de la production dans les étalonnages de ce secteur avec la méthode des soldes bloqués (méthode 1) et celle sans l'IPI (méthode 3). La même analyse se retrouve lors de l'analyse des étalonnages dans le secteur de l'agro-alimentaire (C1) : cette fois-ci, les à-coups sont captés par la présence de l'acquis d'IPI au mois 0.

3.1.2 Performances prédictives des étalonnages au niveau A 17

Les performances des différents étalonnages en terme de RMSE au niveau A 17, ainsi que pour la prévision de la production manufacturière obtenue par agrégation de ces résultats, sont présentées dans les tableaux 3.2 à 3.4. À titre de comparaison, avec la même configuration de disponibilité des soldes d'enquête, le meilleur modèle en terme de RMSE actuellement utilisé pour étalonner la production manufacturière a un RMSE dans l'échantillon de 0,80 et un RMSE pseudo temps-réel de 1,05.

Secteur	Méthode 1	Méthode 2	Méthode 3	Écart-type
Manuf	0,66	0,33	0,74	1,45
C1	0,77	0,46	0,73	0,86
C2	3,04	1,06	3,67	4,60
C3	1,63	1,20	1,61	2,61
C4	2,81	1,24	3,04	4,14
C5	0,54	0,27	0,56	1,35

TABLEAU 3.2 – RMSE complet pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation.
Le plus petit RMSE complet des étalonnages actuels de la production manufacturière est de 0,80.

Note de lecture : pour un secteur donné, le ou les RMSE en gras sont ceux significativement meilleurs d'après le test de Diebold Mariano au seuil de 5 %. Si aucun chiffre est en gras, c'est qu'aucun RMSE n'est significativement meilleur.
Sources : Insee, Banque de France et calculs de l'auteur.

Lorsque les données sont connues sur toute la période, la méthode utilisant l'IPI sur tout le trimestre coïncident (méthode 2) fournit des RMSE significativement meilleurs pour tous les secteurs (tableau 3.2). Ce résultat ne sera pas utilisé pour comparer les différentes méthodes de construction puisque cette situation n'est pas représentative de la réalité (voir partie 2.4).

Au vu des résultats présents dans les tableaux 3.3 et 3.4, la méthode que nous recommandons pour prévoir la production manufacturière est la méthode par les soldes bloqués (méthode 1). De plus, le test de Diebold-Mariano conclut que cette méthode a un pouvoir prédictif de la production manufacturière dans l'échantillon meilleur à celui des étalonnages actuels (tableau 3.3). En terme de pouvoir prédictif en dehors de l'échantillon, ce test conclut à une équivalence entre les deux modèles en terme de RMSE pseudo temps-réel (tableau 3.4).

Dans tous les secteurs de l'industrie manufacturière, les prévisions dans l'échantillon sont significativement meilleures ou équivalentes avec la méthode des soldes bloqués (méthode 1) que pour les deux autres méthodes (tableau 3.3). Si l'on regarde plus en détail les prévisions dans la cokéfaction-raffinage (C2), le RMSE associé à l'étalonnage en passant par l'IPI (méthode 2) est plus grand que l'écart-type de la série : prévoir une évolution de la production égale à la moyenne à chaque trimestre fournirait de meilleurs résultats.

Secteur	Méthode 1	Méthode 2	Méthode 3	Écart-type
Manuf	0,66	0,72	0,87	1,45
C1	0,77	0,77	0,74	0,86
C2	3,04	6,34	3,65	4,60
C3	1,63	1,71	1,64	2,61
C4	2,81	2,78	3,15	4,14
C5	0,54	0,67	0,56	1,35

TABLEAU 3.3 – RMSE dans l'échantillon pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation. Le plus petit RMSE dans l'échantillon des étalonnages actuels de la production manufacturière est de 0,80.

Note de lecture : pour un secteur donné, le ou les RMSE en gras sont ceux significativement meilleurs d'après le test de Diebold Mariano au seuil de 5 %. Si aucun chiffre est en gras, c'est qu'aucun RMSE n'est significativement meilleur.

Sources : Insee, Banque de France et calculs de l'auteur.

Si l'on s'intéresse désormais à la qualité des étalonnages en terme de prévision en dehors de l'échantillon (tableau 3.4), le test de Diebold-Mariano conclut que les trois méthodes fournissent des résultats équivalents dans tous les secteurs y compris dans l'industrie manufacturière dans son ensemble, à l'exception de la cokéfaction-raffinage (C2) où les prévisions par la méthode des soldes bloqués (méthode 1) sont meilleures.

Les étalonnages construits sans l'IPI sont ceux qui possèdent en général le moins bon pouvoir prédictif.

3.1.3 Performances prédictives des étalonnages au niveau A 64

Les performances des différents étalonnages en terme de RMSE au niveau A 64, ainsi que les prévisions pour le secteur des matériels de transport (C4) et l'industrie manufacturière, sont présentés dans le tableau 3.5. Pour la méthode d'étalonnages de la production après prévisions de l'IPI, ces résultats ne peuvent être utilisés car aucun modèle de prévision valide n'a pu être trouvé pour les étalonnages de l'IPI (première étape de la méthode 2) dans le secteur de l'automobile (CL1).

Secteur	Méthode 1	Méthode 2	Méthode 3	Écart-type
Manuf	1,00	0,97	1,05	1,45
C1	0,80	0,88	0,81	0,86
C2	4,32	6,51	5,17	4,60
C3	1,81	1,78	1,73	2,61
C4	3,20	3,10	3,34	4,14
C5	0,91	0,91	0,90	1,35

TABLEAU 3.4 – RMSE pseudo temps-réel pour les trois méthodes de construction des étalonnages au niveau A 17 ainsi que pour l'industrie manufacturière obtenue par agrégation.

Le plus petit RMSE pseudo temps-réel des étalonnages actuels de la production manufacturière est de 1,05.

Note de lecture : pour un secteur donné, le ou les RMSE en gras sont ceux significativement meilleurs d'après le test de Diebold Mariano au seuil de 5 %. Si aucun chiffre est en gras, c'est qu'aucun RMSE n'est significativement meilleur.

Sources : Insee, Banque de France et calculs de l'auteur.

Ici, agréger les prévisions de la production au niveau A 64 ne permet pas d'améliorer les prévisions de la production manufacturière. En effet, les RMSE dans l'échantillon et les RMSE pseudo temps réel associés à la prévision du taux de croissance de la production manufacturière sont supérieurs ou égal lorsqu'ils sont calculés par agrégation des prévisions au niveau A 64 par rapport à l'agrégation au niveau A 17. Cet écart n'est pas significatif d'après le test de Diebold Mariano. Toutefois, les RMSE pseudo temps-réel pour la prévision du secteur des matériels de transport (C4) est significativement plus grand lorsqu'elle est obtenue à partir de l'agrégation des prévisions dans l'automobile (CL1) et dans les autres matériels de transport (CL2) que lorsqu'elle est obtenue directement à partir des soldes du secteur. Pour la prévision de la production manufacturière et du secteur des matériels de transport (C4) nous recommandons donc d'utiliser les étalonnages obtenus au niveau A 17 plutôt que ceux obtenus au niveau A 64.

Secteur étudié	RMSE Complet		RMSE dans l'échantillon		RMSE pseudo temps-réel		Écart-type
	Méthode 1	Méthode 3	Méthode 1	Méthode 3	Méthode 1	Méthode 3	
Manuf	0,71	0,81	0,66	0,78	1,06	1,03	1,45
C4	2,79	3,00	2,79	3,39	3,47	3,24	4,14
CL1	3,45	3,90	3,45	3,93	5,06	4,56	5,88
CL2	4,24	4,07	4,24	4,11	4,08	4,01	4,50

TABLEAU 3.5 – Différents RMSE pour les méthodes 1 et 3 de construction des étalonnages au niveau A 64 ainsi que pour l'industrie manufacturière et le secteur des matériels de transport (C4) obtenus par agrégation.

Les étalonnages actuels de la production manufacturière donnent un RMSE complet et dans l'échantillon de 0,80 et un RMSE pseudo temps-réel de 1,05.

Note de lecture : pour un secteur et un type de RMSE donné, le RMSE en gras est celui significativement meilleur d'après le test de Diebold Mariano au seuil de 5 %. Si aucun chiffre n'est en gras, c'est qu'aucun RMSE n'est significativement meilleur.

Sources : Insee, Banque de France et calculs de l'auteur.

3.2 Comparaison des méthodes d'agrégation des prévisions

Les tableaux 3.6 et 3.7 présentent les différents RMSE associés aux 3 méthodes d'agrégation pour les étalonnages aux niveaux A 17 et A 64 réalisés à partir des soldes bloqués (méthode 1). Dans le

tableau 3.6, pour la méthode « comptable » et l'« approximation », les étalonnages utilisés pour calculés les différents RMSE dans les secteurs autres que celui de l'industrie manufacturière sont ceux obtenus dans la partie 3.1 avec la méthode des soldes bloqués. De même dans le tableau 3.7 pour les secteurs autres que l'industrie manufacturière et celui des matériels de transport (C4). Seule la méthode d'Hyndman utilisant les combinaisons de toutes les prévisions peut modifier les prévisions réalisées dans ces secteurs.

Secteur étudié	RMSE dans l'échantillon			RMSE pseudo temps-réel			Écart-type
	Comptable	Approx.	Hyndman	Comptable	Approx.	Hyndman	
Manuf	0,66	0,66	0,66	1,00	0,99	0,99	1,45
C1	0,77	0,77	0,78	0,80	0,80	0,80	0,86
C2	3,04	3,04	3,04	4,32	4,32	4,32	4,60
C3	1,63	1,63	1,63	1,81	1,81	1,80	2,61
C4	2,81	2,81	2,81	3,20	3,20	3,20	4,14
C5	0,54	0,54	0,58	0,91	0,91	0,92	1,35

TABLEAU 3.6 – Comparaison des différentes méthodes d'agrégation au niveau A 17 pour la méthode de construction des étalonnages par les soldes bloqués (méthode 1) en terme de RMSE.

Sources : Insee, Banque de France et calculs de l'auteur.

Secteur étudié	RMSE dans l'échantillon			RMSE pseudo temps-réel			Écart-type
	Comptable	Approx.	Hyndman	Comptable	Approx.	Hyndman	
Manuf	0,66	0,76	0,65	1,06	1,02	1,02	1,45
C1	0,77	0,77	0,78	0,80	0,80	0,80	0,86
C2	3,04	3,04	3,04	4,32	4,32	4,32	4,60
C3	1,63	1,63	1,63	1,81	1,81	1,80	2,61
C4	2,79	2,83	2,60	3,47	3,55	3,26	4,14
CL1	3,45	3,45	3,28	5,06	5,06	4,02	5,88
CL2	4,24	4,24	4,20	4,08	4,08	4,02	4,50
C5	0,54	0,54	0,57	0,91	0,91	0,91	1,35

TABLEAU 3.7 – Comparaison des différentes méthodes d'agrégation au niveau A 64 pour la méthode de construction des étalonnages par les soldes bloqués (méthode 1) en terme de RMSE.

Sources : Insee, Banque de France et calculs de l'auteur.

Au niveau A 17, toutes les méthodes sont équivalentes et les RMSE de tous les secteurs sont très proches. Au niveau A 64, seule la méthode proposée par Hyndman permet d'améliorer certains étalonnages en terme de RMSE. En effet, avec cette méthode, les RMSE dans l'échantillon et pseudo temps-réel sont réduits de manière significative dans le secteur des matériels de transport (C4) et dans l'automobile (CL1) par rapport aux autres méthodes et par rapport aux étalonnages obtenus au niveau A 17.

Cependant, la méthode que nous conseillons d'adopter dans le cadre des réunions d'étalonnages de la Note de conjoncture est la méthode utilisant la part annuelle dans l'industrie manufacturière de chaque secteur en production en valeur. Le niveau d'agrégation utilisé pour la prévision de la production manufacturière est le niveau A 17, les étalonnages au niveau A 64 seraient uniquement utilisés comme aide à l'établissement du diagnostic conjoncturel.

La méthode conseillée est préférée à la méthode exacte comptablement car elle est plus souple, plus simple et donne des résultats équivalents. Par ailleurs, nous ne recommandons pas d'adopter de l'approche d'Hyndman principalement pour des raisons économiques. En effet, même si cette approche permet d'améliorer significativement les résultats dans le secteur des matériels de transport (C4) et dans l'automobile (CL1), les gains restent faibles. L'approche d'Hyndman utilise des cales automatiques aux étalonnages non maîtrisées par l'économiste : des interprétations économiques des différents étalonnages sont alors difficilement réalisables. De plus, dans tous les secteurs, les erreurs de prévision des différents étalonnages restent élevées : les améliorations à réaliser doivent être faites plutôt à travers les étalonnages que dans la méthode d'agrégation.

3.3 Retour sur les prévisions depuis 2014

Dans cette partie nous nous intéressons aux prévisions des comptes de production réalisées par la méthode des soldes bloqués (« méthode 2 ») et aux évolutions réelles observées depuis le premier trimestre 2014.

3.3.1 Prévisions depuis 2014

Les tableaux 3.8 et 3.9 effectuent un retour sur les prévisions qui auraient été réalisées avec la méthode des soldes bloqués présentée dans ce mémoire depuis le premier trimestre 2014. Le tableau 3.8 présente également les étalonnages obtenus pour la production manufacturière par l'étalonnage utilisé à titre de comparaison dans la partie 3.1. Est également présent dans ce tableau les prévisions du trimestre coïncident retenus durant les différentes notes de conjoncture.

Taux de croissance	2014 T1	2014 T2	2014 T3	2014 T4	2015 T1	2015 T2
Réel	0,1	-0,6	0,6	0,2	1,3	-0,7
Retenu	0,9	0,4	0,2	0,2	0,4	0,3
Prévu par agrégation	0,1	-0,2	-0,2	0,2	0,1	0,9

TABLEAU 3.8 – Taux de croissance (en %) réel, retenu dans les différentes notes de conjoncture, et prévu par agrégation des prévisions au niveau A17 de la production manufacturière du premier trimestre 2014 au deuxième trimestre 2015.

Sources : Insee, Banque de France et calculs de l'auteur.

Les prévisions retenues pour l'industrie manufacturière ne se basent pas uniquement sur l'étude d'un étalonnage mais sur :

- l'étude de plusieurs étalonnages faisant intervenir divers soldes d'opinion ;
- l'étude des différents indicateurs conjoncturels produits par l'Insee, la Banque de France et PMI Markit ;
- l'étude des réponses des plus grosses entreprises aux enquêtes de conjonctures ainsi que de leurs éventuels commentaires (pouvant par exemple indiquer une grève ou un incident ayant un impact sur leur production).

Après la prise en compte de tous ces éléments, un équilibrage comptable ressources-emplois est également réalisé : c'est le Garde Fou. Les résultats des nouveaux étalonnages seront donc uniquement utilisés pour l'aide à la décision et ne prévaudront pas de l'évolution de la production manufacturière retenue dans le cadre des réunions d'étalonnages. Tout cela explique en partie les écarts entre les prévisions obtenues par les étalonnages et la valeur finale retenue pour le taux de croissance de la production manufacturière dans les notes de conjoncture.

Secteur	Prévision	2014T1	2014T2	2014T3	2014T4	2015T1	2015T2
C1	Réel	-0,4	-0,2	0,3	-0,6	0,2	0,6
	Prévu	0,3	0,2	0,1	0,2	0,4	0,4
C2	Réel	5,3	-3,3	5,2	4,0	6,5	-9,6
	Prévu	1,6	-1,3	-1,2	1,2	-2,7	2,3
C3	Réel	-1,5	-2,1	0,0	2,0	-2,1	2,1
	Prévu	-0,1	0,3	-0,7	0,6	-0,3	1,8
C4	Réel	-1,9	0,2	1,2	-2,3	5,5	-1,3
	Prévu	-1,4	-0,2	1,3	0,4	1,5	2,2
CL1	Réel	-3,7	-1,2	4,7	-5,4	7,1	0,7
	Prévu	0,7	3,0	1,6	1,8	2,7	2,5
CL2	Réel	-0,3	1,5	-1,9	0,7	4,0	-3,2
	Prévu	1,0	1,0	2,0	1,9	2,0	0,9
C5	Réel	0,5	-0,3	0,0	0,4	0,5	-0,4
	Prévu	0,3	-0,4	-0,5	-0,1	-0,1	0,2

TABEAU 3.9 – Taux de croissance (en %) réel et prévus des comptes de production trimestriels du premier trimestre 2014 au deuxième trimestre 2015.

Sources : Insee, Banque de France et calculs de l'auteur.

ENCADRÉ 3.1 - Décompositions de l'erreur de prévision de la production manufacturière obtenue par agrégation

Pour le trimestre T notons respectivement Y_T et $Y_{i,T}$ les taux de croissance trimestriels de la production dans le secteur manufacturier et du secteur i . Notons également $\omega_{i,T}$ la part de la production annuelle en valeur, dans l'industrie manufacturière, du secteur i à l'année précédant le trimestre T . Le taux de croissance \hat{Y}_T est obtenu par agrégation des résultats des étalonnages au niveau A 17 avec la méthode « d'approximation » :

$$\hat{Y}_T = \sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} \hat{Y}_{i,T}$$

L'erreur de prévision de la production manufacturière au trimestre T , $Y_T - \hat{Y}_T$, peut donc s'écrire :

$$\begin{aligned} Y_T - \hat{Y}_T &= \left(Y_T - \sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} Y_{i,T} \right) + \left(\sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} Y_{i,T} - \hat{Y}_T \right) \\ &= \left(Y_T - \sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} Y_{i,T} \right) + \left(\sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} Y_{i,T} - \sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} \hat{Y}_{i,T} \right) \end{aligned}$$

D'où la décomposition suivante :

$$Y_T - \hat{Y}_T = \underbrace{\left(Y_T - \sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} Y_{i,T} \right)}_{\text{erreur d'agrégation}} + \left(\sum_{i \in \{C1, \dots, C5\}} \omega_{i,T} \underbrace{(Y_{i,T} - \hat{Y}_{i,T})}_{\text{erreur de prévision du secteur } i} \right)$$

En pratique, l'erreur d'agrégation est négligeable.

L'erreur de prévision de la production manufacturière obtenue par agrégation des étalonnages au niveau A 17 par la méthode « d'approximation » peut se décomposer en la somme de six termes (voir encadré 3.1) :

- un terme dû à l'erreur réalisée par la méthode d'agrégation employée, qui n'est pas comptablement exacte. Cette erreur est en pratique négligeable ;
- cinq termes proportionnels aux erreurs réalisées dans les sous-secteurs de l'industrie manufacturière.

Le graphique 3.1 résume, depuis le premier trimestre 2015, l'ensemble de ces cinq derniers termes ainsi que le taux de croissance réel et prévu de la production manufacturière.

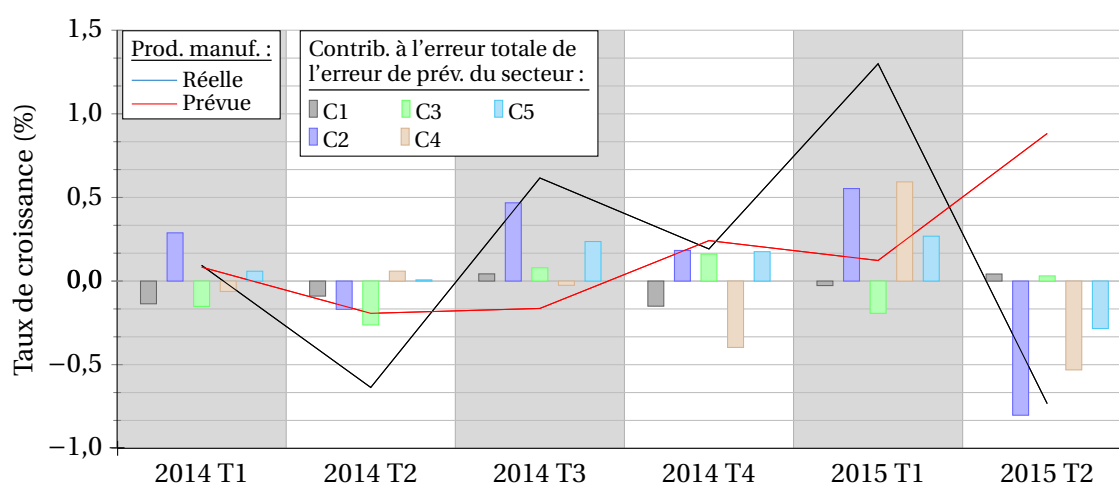
La variance de l'erreur de prévision, donnant un ordre de grandeur de l'intervalle de confiance, peut également être décomposée en plusieurs termes spécifiques à chaque sous-secteur de l'industrie manufacturière. Les contributions de chaque secteur à la variance totale sont présentées dans le tableau 3.10. La contribution de l'erreur d'agrégation à la variance totale de prévision est de 0,1 % : elle est donc bien négligeable.

Secteur	C1	C2	C3	C4	C5
Contribution	3,7	12,7	18,1	42,2	23,0

TABLEAU 3.10 – Contribution à la variance de l'erreur de prévision de la production manufacturière (en %) des prévisions effectuées au niveau A 17 du premier trimestre 1991 au deuxième trimestre 2015.

Sources : Insee, Banque de France et calculs de l'auteur.

Dans la partie 1.3.4, nous avons vu que le secteur contribuant le plus à la variance de la production manufacturière était le secteur des autres industries (C5). Ce n'est plus le cas pour la variance de l'erreur de prévision : le secteur ayant la plus grande contribution est celui des matériels de transport (C4) à hauteur de 42 %. La contribution du secteur des autres industries (C5) est proche de celle du secteur des biens d'équipement (C4) : elle est de 23 % pour le premier et de 18 % pour le second. Ainsi, si l'on souhaite améliorer la précision des prévisions de la production manufacturière, il est préférable de chercher à améliorer les prévisions dans le secteur des matériels de transport (C4) plutôt que dans les autres secteurs.



GRAPHIQUE 3.1 – Taux de croissance de la production manufacturière (en %) réels et prévus par la méthode des soldes bloqués au niveau A 17.

Sources : Insee, Banque de France et calculs de l'auteur.

3.3.2 Retour sur les prévisions du deuxième trimestre

Nous revenons ici sur la prévision de la production manufacturière retenue pour le 2ème trimestre 2015 dans la Note de conjoncture de juin 2015 (+0,3 %).

Pourquoi +0,3 % de croissance de la production manufacturière au deuxième trimestre 2015 ?

Pour comprendre ce choix, revenons tout d'abord sur la situation au premier trimestre 2015. D'après les comptes nationaux, la production manufacturière a accéléré au premier trimestre 2015 (+1,3 % après +0,1 %). C'est un peu plus que l'évolution trimestrielle de l'IPI sur le même champ (+0,8 %) et nettement plus que ce qui avait été retenu dans la Note de conjoncture de mars (+0,4 %). Le dynamisme provient principalement des matériels de transport (C4) (+5,5 % après -2,3 %) et de la branche cokéfaction-raffinage (C2) (+6,5 % après +4,0 %). Pour le deuxième trimestre 2015, il faut donc s'attendre à un à-coup négatif sur la production manufacturière puisqu'aucun des étalonnages n'avait prévu une si forte accélération pour le premier trimestre.

Les différents étalonnages actuellement utilisés, construits à partir des soldes d'enquêtes (Insee, Banque de France) et de l'IPI disponible jusqu'en mars 2015, indiquent une prévision variant de +0,2 % à +0,9 %. La valeur retenue pour la croissance de la production manufacturière au deuxième trimestre 2015 est de +0,3 %. Cela est cohérent d'un côté avec un acquis d'IPI élevé en mars (+0,4 %) et une augmentation entre mars et mai du solde d'opinion sur la production passée (+25 points)¹ ; mais de l'autre côté avec un léger recul du solde d'opinion sur les perspectives de production depuis février.

Ce qu'il s'est passé au deuxième trimestre 2015

Dans les résultats détaillés des comptes trimestriels, l'activité manufacturière est en nette baisse au deuxième trimestre (-0,7 %). Dans la branche cokéfaction-raffinage (C2), cette détérioration de l'activité est particulièrement importante (-9,6 % après +6,5 %), du fait notamment d'arrêts de plusieurs raffineries pour maintenance. L'activité a également baissé dans les autres industries (C5) (-0,4 %) et dans les matériels de transport (C4) (-1,3 %), après un premier trimestre en hausse (respectivement +0,5 % et +5,5 %). Dans ce dernier, la baisse est principalement due à une baisse de l'activité dans les autres matériels de transport (CL2). À l'inverse la production a accéléré dans l'agro-alimentaire (C1) (+0,6 % après +0,2 %) et dans les biens d'équipement (C3) (+2,1 % après -2,1 %).

Retour sur les prévisions sectorielles

La surévaluation du taux de croissance dans l'industrie manufacturière au deuxième trimestre par les nouveaux étalonnages présentés dans ce mémoire sont dus à une surévaluation dans trois secteurs (voir graphique 3.1 et tableau 3.11) : la cokéfaction-raffinage (C2), les matériels de transport (C4) et les autres industries (C5).

Dans la cokéfaction-raffinage (C2) la baisse de la production est principalement due à des événements non conjoncturels (maintenance de raffineries), non pris en compte par les soldes d'opinion. Toutefois, ces événements sont connus lors de l'exercice de prévision : un à-coup négatif dans ce secteur était attendu.

Dans le secteur des matériels de transport (C4), la baisse de l'activité est due à la baisse de la production dans le secteur des autres matériels de transport (CL2). Les variations conjoncturelles de ce sous-secteur du des matériels de transport sont difficilement prises en compte par les

1. La valeur des différents soldes d'opinion sur le premier semestre 2015 sont résumés dans l'annexe C.

	Secteur étudié					
	Manuf	C1	C2	C3	C4	C5
Évolution réelle	-0,7	0,6	-9,6	2,1	-1,3	-0,4
Contribution à l'évolution		0,1	-0,6	0,2	-0,2	-0,2
Évolution prévue	0,9	0,4	2,3	1,8	1,8	0,2
Contribution à l'évolution		0,1	0,2	0,2	0,3	0,1
Contribution à l'erreur de prévision		0,0	-0,8	0,0	-0,5	-0,3

TABLEAU 3.11 – Taux de croissance (en %) réels et prévus au niveau A 17 et contribution à l'évolution (en point) de la production manufacturière de chaque secteur.

Sources : Insee, Banque de France et calculs de l'auteur.

soldes d'opinion utilisés ici. En effet, dans les étalonnages de ce secteur, seule une variable est retenue pour expliquer l'évolution de la production (le solde sur l'évolution des carnets de commandes de la Banque de France). Par ailleurs, la surévaluation de l'évolution de la production dans le secteur des matériels de transport (C4) est également due à la surestimation du taux de croissance trimestriel de l'IPI (-1,9 %) par son acquis calculé avec la valeur de mars (+1,1 %).

Enfin, le secteur des « autres industries » (C5) est le secteur ayant le plus grand poids en terme de production dans l'industrie manufacturière, avec un grand nombre de sous-secteurs (sept au niveau A 38). Les indicateurs conjoncturels utilisés pour l'étalonnage de ce secteur ne reflètent donc qu'un mouvement d'ensemble et prennent peu en compte les disparités d'évolutions entre les différents sous-secteurs. Au deuxième trimestre 2015, la baisse de l'activité des « autres industries » ne résulte pas d'une baisse générale de l'activité dans ce secteur. Elle est principalement due à la baisse de l'activité dans le secteur de la chimie, en partie compensée par une hausse de la production dans la plasturgie. Des prévisions réalisées à un niveau plus fin de la nomenclature pour le secteur des « autres industries » auraient peut-être pu prendre en compte ces disparités et améliorer la prévision de ce secteur.

Conclusion

En conclusion, pour les prochaines réunions d'étalonnage, nous recommandons d'effectuer, en plus des étalonnages actuellement utilisés, une prévision de la production manufacturière réalisée à partir des étalonnages construits à un niveau plus fin. Pour chaque sous-secteur de l'industrie manufacturière, les étalonnages seraient construits avec la méthode des soldes bloqués (méthode 1) et la prévision de l'évolution de la production manufacturière serait obtenue par agrégation des résultats obtenus au niveau A 17. Les étalonnages au niveau A 64 seraient utilisés uniquement pour l'aide au diagnostic conjoncturel du secteur manufacturier et du secteur des matériels de transport (C4) : ils ne seraient pas employés pour la construction des prévisions de la production manufacturière. La méthode d'agrégation que nous conseillons d'adopter est « l'approximation » qui consiste à pondérer les prévisions de chaque secteur par leur part, dans l'industrie manufacturière, de la production annuelle en valeur de l'année précédente. C'est la méthode la plus souple et elle fournit des résultats équivalents aux autres façons d'agréger. Les performances prédictives des étalonnages ainsi construits sont meilleurs que les étalonnages de la production actuellement utilisés, avec une précision des estimations significativement améliorée.

Dans ce chapitre, nous revenons sur les démarches que nous avons effectuées et sur les points faibles de nos méthodes afin de proposer de possibles améliorations de cette étude.

4.1 Comment améliorer les prévisions ?

4.1.1 En utilisant d'autres sources et d'autres types de variables

Pour effectuer les étalonnages, nous nous sommes essentiellement restreints à des variables provenant d'enquêtes. Toutefois, d'autres sources de données auraient pu être utilisées. En effet, des variables pluviométriques, météorologiques ou encore sur la consommation électrique pourraient, par exemple, avoir un pouvoir explicatif sur la production dans le secteur de l'agro-alimentaire. Nous aurions par exemple pu insérer dans nos variables explicatives une variable sur l'écart à la température moyenne saisonnière.

Pour les soldes d'enquête de la Banque de France et Insee, seuls les principaux soldes mensuels ont été employés. D'autres soldes (mensuels ou trimestriels) tels que le taux d'utilisation des capacités de production auraient également pu être utilisés. En particulier, l'étalonnage retenu pour la prévision de la production dans le secteur des autres matériels de transport (CL2) ne contient qu'une seule variable explicative. L'usage d'autres variables pourrait permettre de mieux capter les fluctuations conjoncturelles et de réduire les erreurs de prévision.

Dans les étalonnages actuels de la production manufacturière, une des variables utilisée est l'indicateur synthétique du climat des affaires. Celui-ci résume l'opinion des chefs d'entreprise sur la conjoncture dans l'industrie : plus sa valeur est élevée, plus les industriels considèrent que la conjoncture est favorable. Il est calculé par analyse factorielle à partir des cinq soldes d'opinion de l'enquête de conjoncture Insee suivants : production passée, perspectives personnelles de production, perspectives générales de production, stocks, carnets de commandes globaux et carnets de commandes globaux. Cet indicateur synthétique pourrait être calculé de manière sectorielle et introduit dans les différents étalonnages.

4.1.2 En prolongeant les soldes et en réalisant un étalonnage de l'IPI spécifique à chaque mois

Pour les méthodes 2 et 3 de construction des étalonnages après prévision de l'IPI et sans l'IPI, les soldes d'enquêtes ont été prolongés de façon simple sur le trimestre coïncident : par la dernière valeur connue. Une autre façon, plus complexe, qui aurait pu être utilisée est de prolonger les soldes à l'aide d'un processus auto-régressif (AR).

Pour la prévision de l'IPI sur le trimestre coïncident, la même équation d'étalonnage a été utilisée pour tous les mois du trimestre. Toutefois, rien ne garantit qu'utiliser cette même équation pour tous les mois du trimestre exploite au mieux l'information disponible. Par exemple, au premier

mois de chaque trimestre, le solde Insee sur la tendance prévue de la production représente la meilleure prévision que puissent faire les industriels et peut donc être un indicateur important pour prévoir le niveau de l'IPI de la fin du trimestre. À l'inverse, à la fin du trimestre, le solde sur la tendance passée de la production sera un meilleur indicateur pour prévoir l'IPI au début du trimestre que le solde sur la tendance prévue de la production du trimestre précédent.

Ainsi, tenter de prévoir, en fin de trimestre, l'IPI à l'aide du solde sur la production passé revient à ne pas utiliser correctement l'information disponible. D'autant plus que ce solde serait prolongé puisque, lors de l'exercice de la prévision, il n'est pas connu sur tout le trimestre coïncident. Il est donc plus pertinent de construire un étalonnage de l'IPI spécifique à la prévision de chaque mois du trimestre.

4.1.3 En exploitant les informations des modèles terminaux

Lors de la dernière étape de l'algorithme de sélection des variables Gets, lorsque plusieurs modèles terminaux sont retenus, la sélection du modèle final a été réalisée à l'aide d'un critère d'information. Les autres modèles terminaux n'ont donc pas été étudiés. Ils pourraient être utilisés en prenant comme prévision une pondération des prévisions des différents modèles. Cela permettrait de réduire l'écart-type de l'erreur de prévision. Une moyenne arithmétique simple pourrait être appliquée. Cette pondération simple possède des performances comparables à des approches plus complexes. Par exemple, les trois approches suivantes sont souvent employées :

- Pondérer par la probabilité que le modèle soit le « vrai modèle » (on parle également de la probabilité *a posteriori* du modèle). Si l'on a m modèles notés $(M_i)_{i \in [1, m]}$, la probabilité que le modèle i soit le « vrai modèle » est égale à [15] :

$$\frac{\exp\left(-\frac{1}{2}BIC_i\right)}{\sum_{j=1}^m \exp\left(-\frac{1}{2}BIC_j\right)}$$

Où BIC_i est le critère d'information bayésien du modèle i . Cette approche est également appelée approche BMA (*Bayesian Model Averaging*).

- Pondérer par l'inverse de l'écart-type de l'erreur de chaque modèle. Cette méthode n'est pas issue de fondements théoriques mais donne empiriquement de bons résultats.
- Pondérer par une moyenne arithmétique simple.

Pour les modèles terminaux, les critères d'information BIC et les écart-types des erreurs sont heuristiquement très proches. Les pondérations de ces deux premières approches sont donc voisines de la pondération utilisée pour la moyenne arithmétique simple.

4.1.4 En prenant en compte les liens entre les secteurs par la méthode SURE

Dans ce mémoire, tous les étalonnages ont été réalisés de façon indépendante. En d'autres termes, nous avons considéré que la production d'un secteur n'était expliquée que par les informations conjoncturelles de ce secteur. En particulier, nous avons considéré que la production d'un secteur n'était pas influencée par la production des autres secteurs. Ce n'est toutefois en pratique pas le cas. Statistiquement, cette dépendance entre les étalonnages se vérifie avec la corrélation entre les erreurs de prévision dans les différents secteurs. Dans notre étude cette corrélation est significativement non nulle. Toutefois, estimer les équations de prévision de manière indépendantes ne biaisent pas les résultats : les estimations sont toujours non biaisées, consistantes¹ mais ne sont plus efficaces².

1. Un estimateur consistant est un estimateur qui converge asymptotiquement (en probabilité) vers la vraie valeur.

2. Un estimateur efficace est un estimateur qui est de variance minimale.

Une façon de prendre en compte cette dépendance entre les différents étalonnages est la méthode SURE – *Seemingly Unrelated Regression Equations* (systèmes d'équations apparemment non reliées) –. Elle pourrait être appliquée sur les modèles finaux obtenus par l'algorithme de sélection des variables.

4.2 Limites de l'étude

4.2.1 Utilisation d'autres tests de spécification

Dans l'algorithme de sélection des variables utilisés dans ce mémoire, la bonne spécification des modèles est vérifiée par trois test :

- le test de Jarque-Bera pour la normalité des résidus ;
- le test de Ljung-Box sur les résidus pour l'autocorrélation ;
- le test de Ljung-Box sur les carrés des résidus pour l'homoscédasticité.

D'autres tests plus robustes, non implémentés dans l'algorithme utilisé, auraient toutefois pu être employés. En effet, le test d'homoscédasticité de Ljung-Box est également appelé test d'absence d'hétéroscédasticité conditionnelle ou test ARCH (*AutoRegressive Conditional Heteroscedasticity*) : il vérifie l'absence d'hétéroscédasticité des résidus sur leur passé. Cependant, la présence d'hétéroscédasticité peut également être due à une dépendance entre la variance des résidus et les variables explicatives, ce qui n'est pas pris en compte par ce test. Ainsi, en parallèle du test ARCH, le test de White (étudiant la dépendance le carré des résidus, les variables explicatives et le carré de ces variables) pourrait donc également être utilisé.

Outre la normalité, l'autocorrélation et l'hétéroscédasticité, d'autres tests de spécifications auraient pu être mobilisés, comme par exemple :

- Le test de linéarité de la forme fonctionnelle, appelé test RESET (*Regression Equation Specification Error Test*). Celui ci permet d'étudier si la régression linéaire est bien correcte.
- Le test CUSUM (*cumulative sum*) permettant d'étudier la stabilité du modèle au cours du temps.

4.2.2 Limite de la prévision des soldes

Dans les méthodes 2 et 3 de construction des étalonnages des comptes de production après une prévision de l'IPI et sans l'IPI, la sélection de variables a été réalisée en supposant que les soldes étaient connus sur tout le trimestre coïncident. Toutefois, ces derniers étaient, lors de la prévisions, prolongés par la dernière valeur connue. Or, comme précisé avec les étalonnages de l'IPI dans la partie 4.1.2, cette façon de procéder n'utilise pas au mieux l'information disponible [9].

Conclusion

Afin de décrire l'état de l'économie dans l'industrie manufacturière et ses différents sous-secteurs, trois familles d'indicateurs conjoncturels sont utilisés :

- Les indicateurs mensuels issus des enquêtes de conjoncture : ils ont l'avantage d'être disponibles rapidement et fournissent des indications sur le passé récent et les perspectives à court terme tirées des réponses des entreprises. Ces indicateurs sont également appelés soldes d'opinion.
- Les indices de production industrielles (IPI) : ce sont des indicateurs quantitatifs mensuels de l'activité réelle, mais ne sont disponibles que tardivement.
- Les comptes trimestriels de production sont utilisés pour établir le diagnostic conjoncturel des différents secteurs. Leurs évolutions sont estimées à partir des deux premières familles d'indicateurs.

Disponibles à un niveau fin, un diagnostic conjoncturel de chaque sous-secteur de l'industrie manufacturière peut donc être réalisé. Celui-ci est établi à l'aide de modèles de prévision des productions, appelés étalonnages. Ces derniers sont construits dans les sous-secteurs de l'industrie manufacturière à différents niveaux de la nomenclature : le niveau A 17 (permettant un découpage en cinq sous secteurs) et le niveau A 64 pour le sous-secteur des matériels de transport.

Pour construire les différents étalonnages, trois méthodologies sont envisagées prenant en compte que, lors de la prévision, les variables ne sont pas connues sur tout le trimestre à prévoir. La première méthode consiste à « bloquer » les soldes d'opinion au dernier mois connu pour les trimestrialiser. Si le dernier mois connu est le premier mois du trimestre, le solde « bloqué » associé correspond à la série trimestrielle obtenue ne gardant que la valeur du solde au premier mois de chaque trimestre. L'IPI est également transformé en acquis : son taux de croissance trimestriel est calculé en prolongeant son niveau mensuel par la dernière valeur connue. La deuxième méthode est réalisée en deux étapes : l'IPI est tout d'abord prévu mensuellement sur tout le trimestre à partir des soldes d'opinion et cette prévision est ensuite utilisée dans l'étalonnage des comptes de production. Tous les soldes sont prolongés sur le trimestre coïncident par la dernière valeur connue. Dans la troisième méthode, les soldes sont également prolongés sur le trimestre par la dernière valeur connue mais l'IPI n'est plus utilisé pour prévoir la production.

Puisque de nombreuses séries sont disponibles, un algorithme de sélection, Gets, est utilisé. Celui-ci permet d'avoir un modèle licite économétriquement, le plus parcimonieux possible et avoir un pouvoir prédictif comparable au modèle où toutes les variables seraient retenues. Lorsque les prévisions aux niveaux A 17 et A 64 sont réalisées, elles sont agrégées pour obtenir une prévision de la production manufacturière. Trois méthodes d'agrégation sont également envisagées : une méthode comptablement juste mais ne permettant pas d'effectuer des prévisions à une horizon plus grande que le trimestre en cours ; une méthode plus souple que la précédente qui consiste à pondérer simplement les prévisions de chaque secteur par son poids en terme

de production dans l'industrie manufacturière ; et une méthode plus complexe utilisant les pondérations précédentes et en réalisant une mise en cohérence de toutes les prévisions.

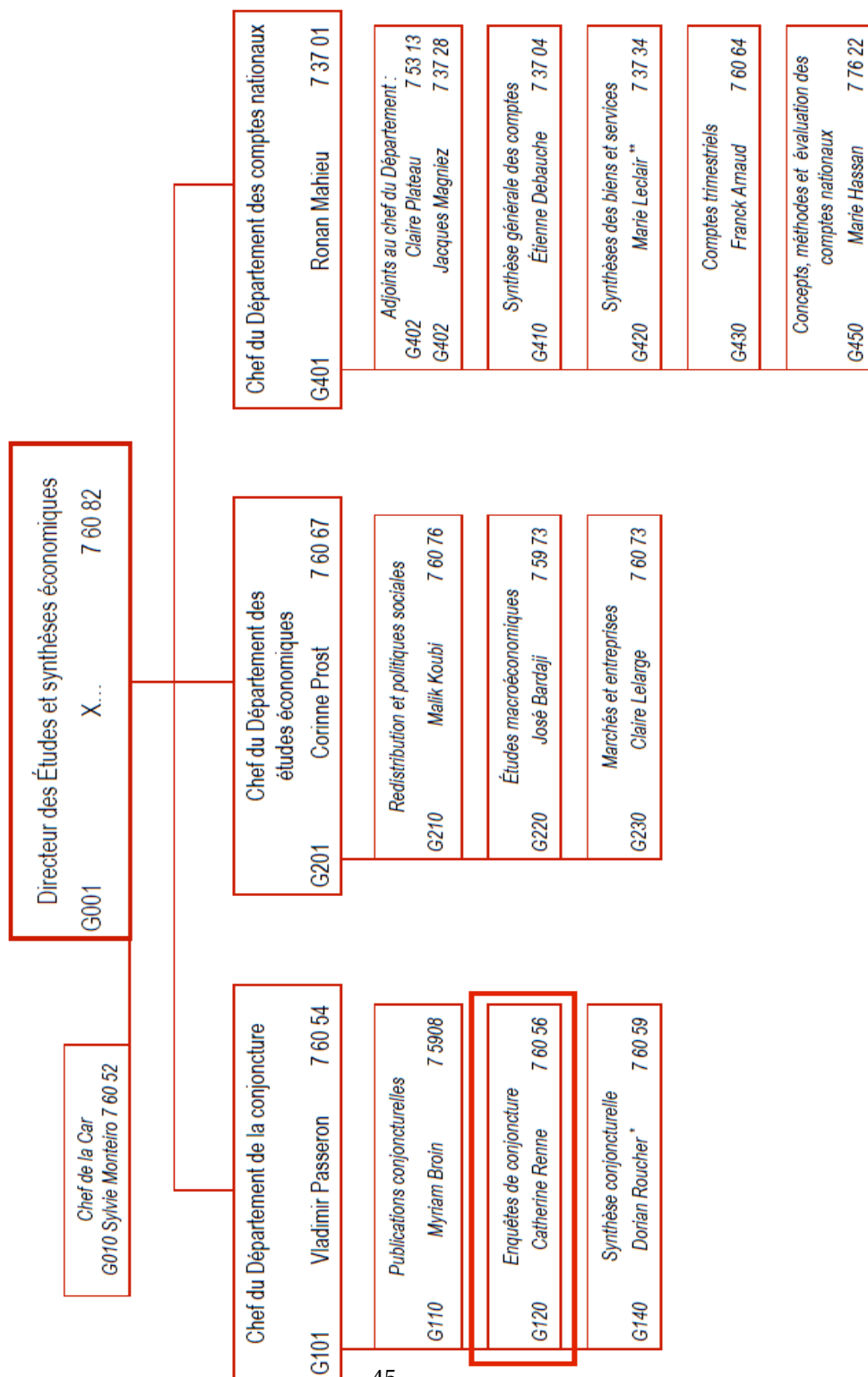
En terme de qualités prédictives, c'est la méthode de construction des étalonnages utilisant les soldes « bloqués » qui fournit les meilleurs résultats. De plus, les prévisions de l'évolution de la production manufacturière obtenues par agrégation des prévisions des sous-secteurs sont plus précises que celles obtenues par les étalonnages actuellement utilisés. Par ailleurs, les trois méthodes d'agrégation des résultats donnent des résultats proches. Cependant, utiliser les étalonnages au niveau A 64 ne permet pas d'améliorer les prévisions de la production manufacturière et produit des moins bonnes prévisions pour la production des matériels de transport que lorsqu'elle est étalonnée directement au niveau A 17.

En somme, pour les prochains exercices de prévision, nous recommandons donc d'effectuer des prévisions sous-sectorielles à partir de soldes « bloqués ». Une prévision de la production manufacturière serait obtenue par agrégation des prévisions au niveau A 17 par la méthode la plus souple, qui est proche de la méthode comptablement exacte. Les résultats des différents étalonnages effectués aux niveaux A 17 et A 64 seraient utilisés pour aider à la construction du scénario conjoncturel de l'industrie manufacturière. En particulier, ces nouveaux modèles de prévision pourront aider à l'élaboration du Point de conjoncture d'octobre concernant la conjoncture du secteur manufacturier du troisième trimestre 2015.

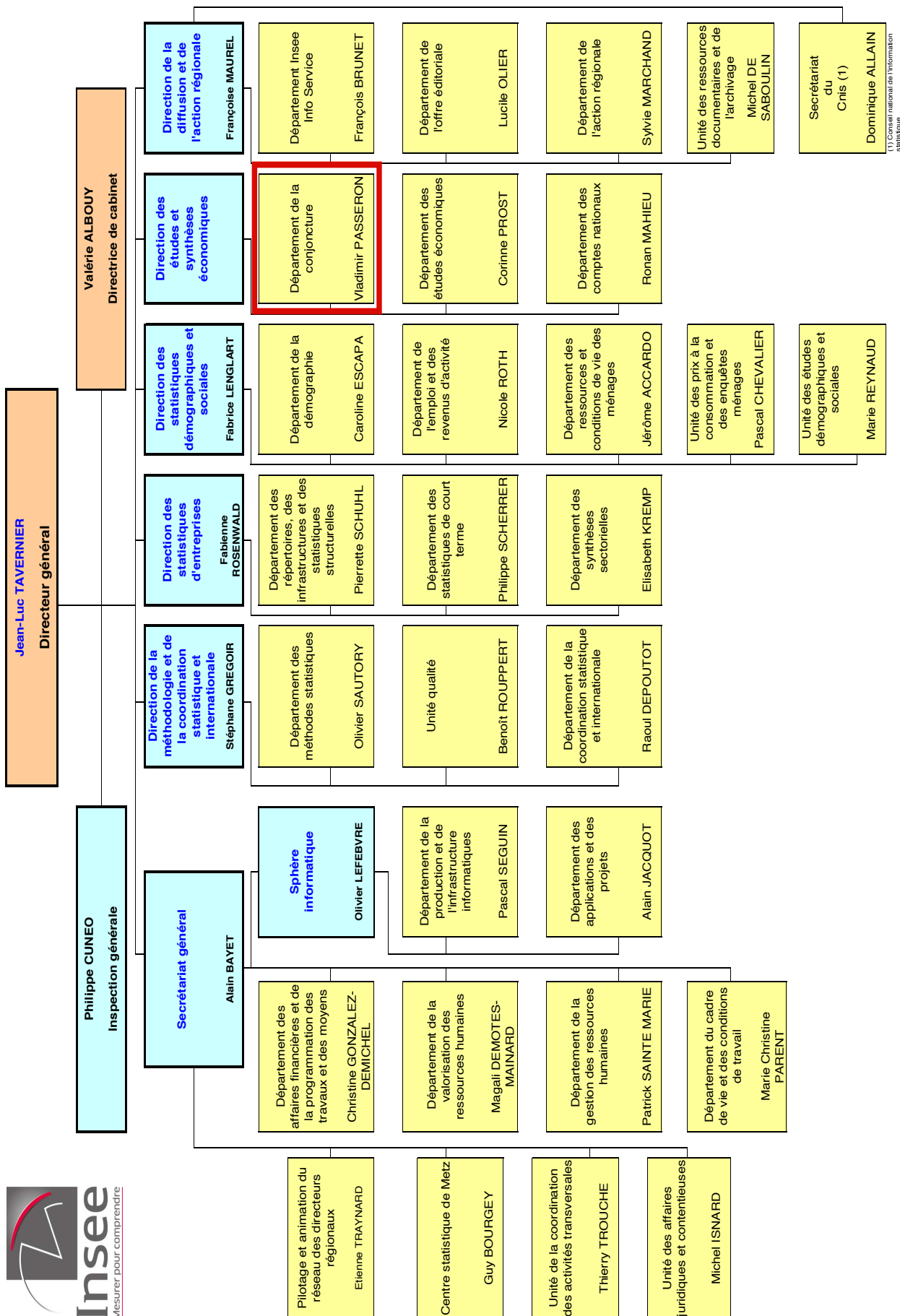
Dans ce mémoire, nous avons vu que réaliser une approche sectorielle permet d'améliorer la prévision de l'évolution production manufacturière du trimestre coïncident. Qu'en est-il pour les prévisions des trimestres suivants ? Et pour les autres grands secteurs de l'économie française tels que celui des services marchands ?

Organigrammes de la Dese et de l'Insee

LA DIRECTION DES ÉTUDES ET SYNTHÈSES ÉCONOMIQUES



Institut national de la statistique et des études économiques (Insee) - direction générale



Annexe B

Nomenclature d'activité française détaillée et questionnaire de l'enquête de conjoncture dans l'industrie

A 17	Intitulé A 17	A 38	A 64	Intitulé A38 ou A64
C1	Industries agro-alimentaires			
C2	Cokéfaction et raffinage			
C3	Biens d'équipement	CI		Fabrication de produits informatiques, électroniques et optiques
		CJ		Fabrication d'équipements électriques
		CK		Fabrication de machines et équipements n.c.a.
C4	Matériels de transport	CL		Fabrication de matériels de transport
			CL1	Industrie automobile
			CL2	Fabrication d'autres matériels de transport
C5	Autres industries	CB		Fabrication de textiles, industries de l'habillement, industrie du cuir et de la chaussure
		CC		Travail du bois, industries du papier et imprimerie
			CC1	Travail du bois et fabrication d'articles en bois et en liège, à l'exception des meubles ; fabrication d'articles en vannerie et sparterie
			CC2	Industrie du papier et du carton
			CC3	Imprimerie et reproduction d'enregistrements
		CE		Industrie chimique
		CF		Industrie pharmaceutique
		CG		Fabrication de produits en caoutchouc et en plastique ainsi que d'autres produits minéraux non métalliques
			CG1	Fabrication de produits en caoutchouc et en plastique
			CG2	Fabrication d'autres produits minéraux non métalliques
		CH		Métallurgie et fabrication de produits métalliques à l'exception des machines et des équipements
			CH1	Métallurgie
			CH2	Fabrication de produits métalliques, à l'exception des machines et des équipements
		CM		Autres industries manufacturières ; réparation et installation de machines et d'équipements
			CM1	Autres industries manufacturières n.c.a.
			CM2	Réparation et installation de machines et d'équipements

TABLEAU B.1 – Découpage de l'industrie manufacturière aux niveaux A 17, A 38 et A 64.



Direction des Etudes et Synthèses Economiques



000008 / 000015

Cadre réservé à l'INSEE

QN01H/V 0008
SIRET

Dossier suivi par :

Tél :
Mél :

Paris, le

ENQUETE MENSUELLE DE CONJONCTURE DANS L'INDUSTRIE

Septembre 2015

Questionnaire à retourner avant le

Statistiques obligatoires loi du 7 juin 1951

Nous vous invitons à répondre à cette enquête sur le site www.conjoncture.entreprises.insee.fr.
Ce questionnaire est accessible depuis le

Madame, Monsieur,

Je vous remercie de bien vouloir remplir ce questionnaire. Votre réponse contribuera à l'évaluation de la conjoncture dans votre secteur d'activité. La synthèse des résultats de l'enquête, « Informations Rapides », sera mise en ligne sur le site internet de l'Insee le jour de la publication de l'enquête (www.insee.fr, cliquer en haut, cadre rouge sur « Thèmes » puis à gauche dans la liste des thèmes sur « Conjoncture » puis sur « Indicateurs de Conjoncture »).

Le Chef du Département de la Conjoncture
V. PASSERON

Vu l'avis favorable du Conseil National de l'Information Statistique, cette enquête, reconnue **d'intérêt général et de qualité statistique, est obligatoire**. Visa n° 2015M021EC du Ministre des finances et des comptes publics et du Ministre de l'économie, de l'industrie et du numérique, valable pour l'année 2015.

Aux termes de l'article 6 de la loi n° 51-711 du 7 juin 1951 modifiée sur l'obligation, la coordination et le secret en matière de statistique, les renseignements transmis en réponse au présent questionnaire ne sauraient en aucun cas être utilisés à des fins de contrôle fiscal ou de répression économique. L'article 7 de la loi précitée stipule d'autre part que tout défaut de réponse ou une réponse sciemment inexacte peut entraîner l'application d'une amende administrative.

Questionnaire confidentiel destiné à la direction générale de l'Insee.

La loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés, garantit aux personnes physiques concernées un droit d'accès et de rectification pour les données les concernant. Ce droit peut être exercé auprès de la direction générale de l'Insee.

COMMENTAIRES

Nom et fonction de la personne qui a répondu : -

Téléphone :

Mél. :

Cette adresse électronique permettra à l'Insee de vous signaler la disponibilité des enquêtes de conjoncture sur le site de collecte par internet : www.conjoncture.entreprises.insee.fr. Elle ne sera utilisée qu'à cette seule fin. Merci de corriger l'adresse de l'entreprise si elle est erronée.

Ce questionnaire, une fois rempli, doit être replié de façon à ce que l'adresse de l'Insee (en bas à droite) apparaisse dans la fenêtre de l'enveloppe-retour ci-jointe.



INSEE DIRECTION GENERALE
Département de la Conjoncture
AUTORISATION 31714
92249 MALAKOFF CEDEX

REPUBLIQUE FRANÇAISE
INSEE - Département de la Conjoncture
15 Bd Gabriel Péri - 92245 MALAKOFF CEDEX

Certaines questions peuvent être sans objet au niveau de votre entreprise. C'est le cas, par exemple, si celle-ci est donneur d'ordre, sous-traitante ou intégrée dans un groupe. Nous vous prions néanmoins de bien vouloir renseigner les questions pertinentes pour votre activité.

QUESTIONS RELATIVES AUX PRODUITS DE VOTRE ENTREPRISE		(le cas échéant, mettre à jour la liste de produits pré-imprimés, SVP)			
DESIGNATION DES PRODUITS
L'ensemble des questions posées ci-dessous concernent vos unités de production localisées en France : Veuillez cocher d'une croix la case qui convient ou entourer la flèche correspondant à votre réponse.					
1. VOTRE PRODUCTION					
a. Evolution au cours des 3 derniers mois	↗	↑	↘	↗	↗
b. Evolution probable au cours des 3 prochains mois	↗	↑	↘	↗	↗
2. LES COMMANDES (OU LA DEMANDE) GLOBALE(S) (toutes provenances)					
Considérez-vous que, compte tenu de la saison, votre carnet de commandes (ou votre demande) est actuellement	☐	☐	☐	☐	☐
	supérieur(e) à la normale	normal(e)	supérieur(e) à la normale	inférieur(e) à la normale	normal(e)
3. LES COMMANDES (OU LA DEMANDE) ETRANGERE(S)					
Considérez-vous que, compte tenu de la saison, votre carnet de commandes étrangères (ou votre demande étrangère) est actuellement	☐	☐	☐	☐	☐
	supérieur(e) à la normale	normal(e)	supérieur(e) à la normale	inférieur(e) à la normale	normal(e)
4. VOS STOCKS DE PRODUITS FINIS (produits prêts pour la vente)					
a. Si la nature de votre production fait que vous travaillez toujours sans stocks de produits fabriqués, cochez la case ci-contre	☐ jamais de stocks				
b. Considérez-vous que, compte tenu de la saison, vos stocks actuels de produits fabriqués sont	☐	☐	☐	☐	☐
	supérieurs à la normale	normaux	supérieurs à la normale	inférieurs à la normale	normaux
5. VOS PRIX DE VENTE					
a. Evolution probable de vos prix de vente (hors taxes) au cours des 3 prochains mois	↗	↑	↘	↗	↗
QUESTIONS RELATIVES A LA MAIN D'ŒUVRE Veuillez entourer la flèche correspondant à votre réponse. 7. LES EFFECTIFS TOTAUX DE VOTRE ENTREPRISE : a. Evolution au cours des 3 derniers mois b. Evolution probable au cours des 3 prochains mois					
VOS PREVISIONS SUR L'INDUSTRIE FRANCAISE DANS SON ENSEMBLE Veuillez entourer la flèche correspondant à votre réponse. (Il s'agit ici de votre opinion sur l'ensemble de l'industrie française) 6. EVOLUTION PROBABLE AU COURS DES 3 PROCHAINS MOIS : a. Du volume de la production industrielle b. Du niveau général des prix des produits industriels					

TEMPS DE RÉPONSE

Combien de temps avez-vous mis pour répondre à cette enquête ?
 (recherche des données + remplissage du questionnaire)

--- heure(s) --- minute(s)

Principes sur les enquêtes et les soldes d'opinion

Un solde d'opinion est une valeur synthétique des réponses des entreprises à une question donnée. Il exprime la différence de proportion entre les entreprises ayant une opinion positive de la situation et celles ayant une opinion négative. Ils sont tout d'abord calculés à un niveau très fin appelé strate élémentaire (qui correspond au croisement de la taille de l'entreprise par l'activité ou à l'activité selon la question étudiée) puis obtenus à un niveau plus agrégé à l'aide d'un système de pondérations [1]. Le calcul se fait donc en deux étapes : l'agrégation primaire et l'agrégation secondaire.

C.1 Agrégation primaire des résultats

En notant i la strate élémentaire et α_k le coefficient de pondération relatif à l'entreprise (ou à un produit en fonction de la question posée) k , le solde d'opinion de la strate élémentaire est calculé de la façon suivante :

$$solde_i = 100 \times (\%Hausse_i - \%Baisse_i)$$

Avec :

$$\%Hausse_i = \frac{\sum_{k \in i} \alpha_k \mathbb{1}_{\{k \text{ répondant « en hausse »\}}}}{\sum_{k \in i} \alpha_k} \quad \%Baisse_i = \frac{\sum_{k \in i} \alpha_k \mathbb{1}_{\{k \text{ répondant « en baisse »\}}}}{\sum_{k \in i} \alpha_k}$$

$$\text{et } \mathbb{1}_{\{k \text{ répondant « en hausse »\}} = \begin{cases} 1 & \text{si l'entreprise } k \text{ a répondu « en hausse »} \\ 0 & \text{sinon} \end{cases}$$

Pour l'enquête de la Banque de France le coefficient de pondération est fonction des effectifs moyens et de l'importance relative de chaque entreprise dans sa branche. Pour l'enquête de l'Insee il dépend de la question posée : il s'agit de l'effectif de l'entreprise pour les questions relatives à la main d'œuvre ; du chiffre d'affaires annuel du produit fabriqué pour les questions d'activité relatives aux produits des entreprises ; et du chiffre d'affaires annuel total pour les questions relatives aux capacités de production et aux facteurs limitant la production.

C.2 Agrégation secondaire des résultats

L'agrégation secondaire correspond à une moyenne pondérée des soldes d'opinion élémentaires permettant de tenir compte l'importance relative de la strate élémentaire dans le niveau d'agrégation. En notant $solde_i$ le solde issu de l'agrégation primaire de la strate élémentaire i et J le niveau d'agrégation, le solde d'opinion agrégé s'écrit :

$$solde_J = \frac{\sum_{i \in J} \beta_i \times solde_i}{\sum_{i \in J} \beta_i}$$

Où β_i correspond au niveau de redressement. Pour l'enquête de la Banque de France il s'agit du poids de la strate dans le niveau d'agrégation en termes de valeur ajoutée. Pour l'enquête de l'Insee, il varie là encore selon les questions : il dépend de l'effectif de la strate ventilé par tranche de taille pour les questions relatives à la main d'œuvre ; de la production par branche, ventilé part tranche de taille, pour les questions d'activité relatives aux produits des entreprises ; et de la production par branche pour les questions relatives aux capacités de production et aux facteurs limitant la production.

C.3 Quelques valeurs des soldes d'opinion et des IPI

Dans cette partie sont présentées les valeurs des différentes publiées et utilisés dans les étalonnages aux niveau A 17 et A 64. Pour rappel, les soldes d'opinion ne sont pas publiés dans le secteur de la cokéfaction-raffinage (C2) et l'indicateur de surprise ne l'est que dans le secteur manufacturier. Ces soldes sont résumés, pour le premier semestre 2015, dans le tableau C.1 pour le secteur de l'industrie manufacturière et dans le tableau C.3 pour les sous-secteurs étudiés. Dans le tableau C.2 sont présents les évolutions mensuelles des IPI dans les différents secteurs étudiés au premier semestre 2015.

	Moy*	janv-15	févr-15	mars-15	avr-15	mai-15	juin-15
Production passée (Insee ; tppa)	3	-2	0	-10	5	16	5
Production prévue (Insee ; tppre)	5	13	11	8	6	6	7
Production passée (BdF ; prodpas)	5	7	1	4	2	6	2
Production prévue (BdF ; prodpre)	5	4	4	4	4	5	3
Carnets de commandes (Insee ; oscd)	-19	-21	-25	-20	-18	-19	-22
Sit. des carnets de commandes (BdF ; sitcar)	1	0	0	2	4	4	5
Évol. des carnets de commandes (BdF ; evocar)	6	8	1	5	6	8	3
Stocks de produits finis (Insee ; ossk)	12	13	4	5	8	9	7
Indicateur de surprise (Insee ; surprise)	0	0	0	0	0	0	0

* Moyenne de longue période.

TABLEAU C.1 – Opinion des industriels dans l'industrie manufacturière.

Note de lecture : entre parenthèses figurent les sources des enquêtes dont proviennent les soldes ainsi que le code des variables utilisé dans les étalonnages.

Sources : Insee – enquête mensuelle de conjoncture dans l'industrie – et Banque de France (BdF) – enquête mensuelle de conjoncture.

Secteur	janv-15	févr-15	mars-15	avr-15	mai-15	juin-15
Industrie manufacturière	0,4	0,2	0,3	-1,0	0,7	-0,7
Agro-alimentaire (C1)	0,3	-0,4	0,8	-0,5	-0,3	1,5
Cokéfaction-raffinage (C2)	0,9	4,0	7,7	-11,1	-3,0	-13,8
Biens d'équipement (C3)	0,2	-1,0	-0,3	0,4	0,4	0,6
Matériels de transport (C4)	2,0	1,9	0,6	-3,9	2,1	-1,2
Automobile (CL1)	2,0	1,8	2,0	-1,6	-0,6	-0,2
Autres matériels de transport (CL2)	2,0	2,0	-0,5	-5,8	4,6	-2,1
Autres industries (C5)	0,1	0,3	0,0	-0,6	0,8	-1,0

TABLEAU C.2 – Évolution mensuelle de la production industrielle (en %).

Source : Insee.

		Moy*	janv-15	févr-15	mars-15	avr-15	mai-15	juin-15
Agro-alimentaire (C1)	Production passée (Insee ; tppa)	6	23	14	-4	3	14	13
	Production prévue (Insee ; tppre)	7	7	-3	10	9	11	8
	Production passée (BdF ; prodpas)	10	7	0	7	7	18	3
	Production prévue (BdF ; prodpre)	10	8	8	7	7	9	7
	Carnets de commandes (Insee ; oscd)	-18	-9	-17	-15	-11	-16	-13
	Situation des carnets de commandes (BdF ; sitcar)	1	-1	-1	4	8	4	7
	Évolution des carnets de commandes (BdF ; evocar)	11	14	-3	15	12	14	11
	Stocks de produits finis (Insee ; ossk)	12	9	8	12	13	7	9
Biens d'équipement (C3)	Production passée (Insee ; tppa)	4	-5	-2	-17	-1	-2	-7
	Production prévue (Insee ; tppre)	2	-1	-8	-9	-1	4	0
	Production passée (BdF ; prodpas)	5	3	3	2	0	-1	-2
	Production prévue (BdF ; prodpre)	5	2	2	0	3	4	5
	Carnets de commandes (Insee ; oscd)	-24	-39	-35	-33	-32	-27	-30
	Situation des carnets de commandes (BdF ; sitcar)	-4	-12	-14	-12	-9	-9	-12
	Évolution des carnets de commandes (BdF ; evocar)	5	2	6	3	3	7	-2
	Stocks de produits finis (Insee ; ossk)	17	11	14	19	22	24	17
Matériels de transport (C4)	Production passée (Insee ; tppa)	8	-4	-5	15	39	31	16
	Production prévue (Insee ; tppre)	8	45	48	38	35	34	18
	Production passée (BdF ; prodpas)	3	10	11	3	8	8	10
	Production prévue (BdF ; prodpre)	4	8	5	7	7	8	7
	Carnets de commandes (Insee ; oscd)	-14	-12	-16	5	10	1	5
	Situation des carnets de commandes (BdF ; sitcar)	4	25	28	27	32	35	37
	Évolution des carnets de commandes (BdF ; evocar)	2	22	7	4	7	11	8
	Stocks de produits finis (Insee ; ossk)	14	8	-12	-4	12	21	26
Automobile (CL1)	Production passée (Insee ; tppa)	3	-12	-24	19	56	61	22
	Production prévue (Insee ; tppre)	3	49	61	2	42	38	21
	Production passée (BdF ; prodpas)	3	26	22	-3	7	12	15
	Production prévue (BdF ; prodpre)	4	12	0	5	8	9	5
	Carnets de commandes (Insee ; oscd)	-20	-22	-23	5	10	-16	-11
	Situation des carnets de commandes (BdF ; sitcar)	-14	1	2	4	8	13	17
	Évolution des carnets de commandes (BdF ; evocar)	2	31	9	6	8	18	6
	Stocks de produits finis (Insee ; ossk)	10	3	-24	-14	9	-2	4
Autres matériels de transport (CL2)	Production passée (Insee ; tppa)	15	23	18	14	11	18	11
	Production prévue (Insee ; tppre)	15	33	42	44	44	42	7
	Production passée (BdF ; prodpas)	5	6	4	11	8	-8	10
	Production prévue (BdF ; prodpre)	5	3	8	8	5	10	3
	Carnets de commandes (Insee ; oscd)	-5	1	-7	4	11	19	22
	Situation des carnets de commandes (BdF ; sitcar)	32	51	55	54	57	58	58
	Évolution des carnets de commandes (BdF ; evocar)	6	17	9	3	7	8	4
	Stocks de produits finis (Insee ; ossk)	20	14	5	9	15	44	50
Autres industries (C5)	Production passée (Insee ; tppa)	1	1	6	-8	3	4	3
	Production prévue (Insee ; tppre)	4	9	2	7	-4	-3	4
	Production passée (BdF ; prodpas)	5	7	-1	3	-1	5	0
	Production prévue (BdF ; prodpre)	6	4	5	4	4	7	2
	Carnets de commandes (Insee ; oscd)	-20	-23	-24	-24	-25	-20	-24
	Situation des carnets de commandes (BdF ; sitcar)	0	-2	-2	-1	1	1	1
	Évolution des carnets de commandes (BdF ; evocar)	6	5	-1	3	5	7	0
	Stocks de produits finis (Insee ; ossk)	10	10	6	9	8	4	-1

* Moyenne de longue période.

TABLEAU C.3 – Opinion des industriels par sous-secteurs de l'industrie manufacturière.

Note de lecture : entre parenthèses figurent les sources des enquêtes dont proviennent les soldes ainsi que le code des variables utilisé dans les étalonnages.

Sources : Insee – enquête mensuelle de conjoncture dans l'industrie – et Banque de France (BdF) – enquête mensuelle de conjoncture.

La notion de contribution à la variation d'une a été mobilisée plusieurs fois durant ce mémoire. Dans cette annexe, nous décrivons plus en détail la façon dont ces contributions ont été calculés.

D.1 Quelques rappels statistiques

Dans cette partie nous effectuons quelques rappels statistiques sur les notions de variance et de covariance qui seront utilisés pour le calcul de la contribution à la variation.

Notons Y une variable aléatoire stationnaire admettant des moments d'ordre 2 et observée en T périodes distinctes (par exemple Y est le taux de croissance de la production manufacturière et T est le nombre de trimestre observé). La variance de Y , notée $\mathbb{V}[Y]$ est une mesure de la dispersion de cette variable : c'est-à-dire qu'elle permet d'indiquer de quelle manière Y est dispersée autour de sa moyenne. Une petite variance signifie que les valeurs de Y sont proches les unes des autres et une grande variance signifie que ces valeurs sont fortement écartées. Elle peut être estimée de la façon suivante :

$$\mathbb{V}[Y] = \frac{1}{T} \sum_{t=1}^T (Y_t - m_Y)^2 \quad \text{où } m_Y \text{ est la moyenne empirique de } Y$$

Notons maintenant Y une variable aléatoire stationnaire admettant également des moments d'ordre 2 et observée en T périodes distinctes. La covariance entre X et Y , notée $\text{Cov}(X, Y)$, est une mesure de la variation simultanée des deux séries. Plus elle sera proche de 0, plus les deux séries auront des variations indépendante. Elle peut être estimée de la façon suivante :

$$\text{Cov}(X, Y) = \frac{1}{T} \sum_{t=1}^T (X_t - m_X)(Y_t - m_Y)$$

Où m_X et m_Y sont respectivement les moyennes empiriques de X et Y .

Ainsi, la covariance de Y avec elle-même est égale à sa variance.

La propriété de la covariance qui sera utilisée et la propriété de bilinéarité. Mathématiquement, si l'on a n séries $(X_i)_{i \in \llbracket 1, n \rrbracket}$ et p séries $(Y_j)_{j \in \llbracket 1, p \rrbracket}$ de variables aléatoires admettant toutes des moments d'ordre 2, cette propriété s'écrit :

$$\text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^p Y_j\right) = \sum_{i=1}^n \sum_{j=1}^p \text{Cov}(X_i, Y_j)$$

D.2 Contribution à la variance : cas général

Soit Y une variable aléatoire pouvant se décomposer en la somme de n variables aléatoires notées X_1 à X_N . Supposons que toutes ces variables sont stationnaires et admettent des moments

d'ordre 2. On a la relation suivante :

$$Y = \sum_{i=1}^n X_i$$

La contribution de X_i à la variance de Y correspond à la part de la variance de Y expliquée par X_i . Elle est égale au rapport entre la covariance de X_i et Y sur la variance de Y :

$$\frac{\text{Cov}(X_i, Y)}{\mathbb{V}[Y]}$$

En effet, des propriétés présentés dans la partie [D](#) on obtient :

$$\begin{aligned}\mathbb{V}[Y] &= \text{Cov}(Y, Y) \\ &= \text{Cov}\left(\sum_{i=1}^n X_i, Y\right) \\ &= \sum_{i=1}^n \text{Cov}(X_i, Y)\end{aligned}$$

C'est-à-dire :

$$\sum_{i=1}^n \frac{\text{Cov}(X_i, Y)}{\mathbb{V}[Y]} = 1$$

Retour sur les pondérations

Dans cette annexe nous revenons sur les différentes méthodes d'agrégation utilisées.

E.1 La méthode comptable

Pour agréger les différentes prévisions au niveau A 17 et A 64 pour obtenir une prévision de la production manufacturière ou du secteur des matériels de transport (C4), la méthode comptable se base sur le calcul des contributions. La contribution d'un secteur i à l'évolution de la production manufacturière correspond à la part de l'évolution de la production manufacturière due au secteur i . La somme des contributions des sous-secteurs de l'industrie manufacturière est donc égale au taux de croissance de la production manufacturière.

Les comptes de production trimestriels que nous utilisons sont en volumes chaînés : ils ne sont donc pas sommables. Par exemple, la somme des productions dans les secteurs de l'automobile (CL1) et des autres matériels de transport (CL2) n'est pas égale à la production dans le secteur des matériels de transport (C4).

Nous admettons [2][4] que pour un secteur i fixé, la contribution du secteur i à l'évolution de la production manufacturière au trimestre T de l'année a s'écrit :

$$ev(Y^{i,ch}) \frac{Y_{a,T-1}^{i,ch}}{Y_{a,T-1}^{sect,ch}} \times \frac{Y_{a-1}^{i,val} / Y_{a-1}^{i,ch}}{Y_{a-1}^{manuf,val} / Y_{a-1}^{manuf,ch}} \quad (E.1)$$

$$+ \underbrace{\left(\frac{Y_{a-1,4}^{i,ch}}{Y_{a-1,4}^{manuf,ch}} - \frac{Y_{a-1}^{i,ch}}{Y_{a-1}^{manuf,ch}} \right) \left(\frac{Y_{a-1}^{i,val} / Y_{a-1}^{i,ch}}{Y_{a-1}^{manuf,val} / Y_{a-1}^{manuf,ch}} - \frac{Y_{a-2}^{i,val} / Y_{a-2}^{i,ch}}{Y_{a-2}^{manuf,val} / Y_{a-2}^{manuf,ch}} \right)}_{\text{Terme de correction du changement de pondérations des prix}} \mathbb{I}_{T=1} \quad (E.2)$$

Où $Y_a^{manuf,val}$ est la production annuelle en valeur à l'année a du secteur manufacturier, $Y_{a,T}^{manuf,ch}$ est la production du trimestre T de l'année a en volumes chaînés du secteur manufacturier. Le terme $ev(Y^{i,ch})$ correspond au taux de croissance trimestriel de la production en volumes chaînés du secteur i au trimestre T .

La première de l'équation E.2 peut également se réécrire pour faire apparaître la correction de la contribution due à l'absence d'additivité des séries en volumes chaînés :

$$ev(Y^{i,ch}) \frac{Y_{a,T-1}^{i,ch}}{Y_{a,T-1}^{sect,ch}} = ev(Y^{i,ch}) \frac{Y_{a,T-1}^{i,ch}}{Y_{a,T-1}^{sect,ch}} + \underbrace{ev(Y^{i,ch}) \frac{Y_{a,T-1}^{i,ch}}{Y_{a,T-1}^{sect,ch}} \left(\frac{Y_{a-1}^{i,val} / Y_{a-1}^{i,ch}}{Y_{a-1}^{manuf,val} / Y_{a-1}^{manuf,ch}} - 1 \right)}_{\substack{\text{Corrige l'absence d'additivité des séries en volumes chaînés} \\ \text{En tant que produit de taux, il est de second ordre par rapport au premier}}}$$

La même formule peut se transposer à la contribution du secteur l'automobile (CL1) ou des autres matériels de transport (CL2) à l'évolution de la production des matériels de transport (C4).

E.2 L'approximation

Pondérer les évolutions des productions des agrégats par leur poids annuel en terme de production dans l'industrie manufacturière est une approximation de l'approche comptable d'agrégation des prévisions présentée dans la partie E.

En effet, dans la formule des contributions d'un secteur à l'évolution de la production manufacturière, on peut tout d'abord omettre le terme spécifique à la correction du changement de pondérations des prix. Cela n'aura pas d'impact sur le calcul de l'évolution de la production manufacturière puisque ce terme s'annule. Nous ne pouvons toutefois plus analyser de manière correcte la contribution de chaque secteur. Ensuite, les deux premiers termes de l'équation E.2 peuvent se réécrire de la façon suivante :

$$ev(Y^{i,ch}) \frac{Y_{a,T-1}^{i,ch} / Y_{a,T-1}^{sect,ch}}{Y_{a-1}^{i,ch} / Y_{a-1}^{manuf,ch}} \frac{Y_{a-1}^{i,val}}{Y_{a-1}^{manuf,val}}$$

Avec i un secteur de l'industrie manufacturière. Par conséquent, l'approximation est d'autant meilleure que l'écart $\frac{Y_{a,T-1}^{i,ch}}{Y_{a,T-1}^{sect,ch}} - \frac{Y_{a-1}^{i,ch}}{Y_{a-1}^{manuf,ch}}$ est faible. En pratique, pour les secteurs étudiés et la période étudiée, cet écart est très faible et au maximum de l'ordre du centième.

En prenant les notations précédentes, la pondération associée au secteur de l'agro-alimentaire (C1) dans l'agrégation permettant d'obtenir la prévision de la production manufacturière au trimestre T de l'année a est donc :

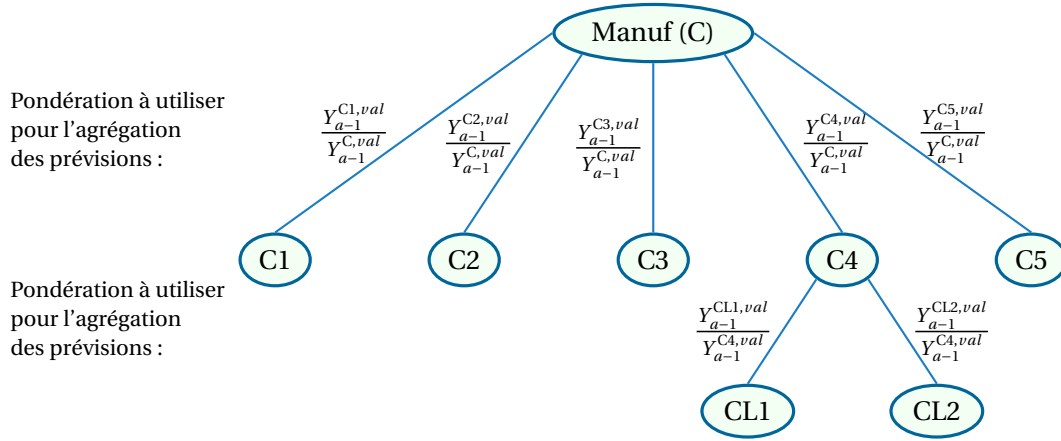
$$\frac{Y_{a-1}^{C1,val}}{Y_{a-1}^{manuf,val}}$$

À une année fixée, ce poids ne dépend pas du trimestre.

E.3 La méthode d'Hyndman

La méthode d'agrégation proposée par Hyndman [13] est différente des deux autres méthodes d'agrégation dans le sens où toutes les prévisions ne sont pas agrégées mais mises en cohérence. Ainsi, un étalonnage spécifique de la production manufacturière est utilisé et les prévisions obtenues sont mises en cohérence avec les prévisions réalisées aux niveaux A 17 et/ou A 38. C'est-à-dire que l'on s'assure, par exemple, que la somme pondérée des prévisions au niveau A 17 soit égale à la prévision de la production manufacturière. Cette sommabilité est assurée par des cales automatiques. Les pondérations utilisées dans cette méthode sont celles utilisées dans « l'approximation » (voir partie E) : pour un secteur donné il s'agit de la part dans l'industrie manufacturière annuelle de production en valeur de l'année précédente. Elle est notée ω_i^a pour le secteur i à tous les trimestres de l'année a . Notons également $Y_{a,T}^i$, le taux de croissance de la production en volumes chaînés du secteur i au trimestre T de l'année a .

Les liens entre les différents secteurs et les pondérations à utiliser pour l'agrégation sont résumées dans le graphique E.1. La méthode d'Hyndman se base sur la prise en compte simultanée de toutes ces relations.



GRAPHIQUE E.1 – Liens entre les secteurs pour l'agrégation des prévisions réalisées à l'année a avec la méthode « d'approximation ».

Au niveau A 17, le point de départ de cette méthode consiste à écrire la relation suivante :

$$\underbrace{\begin{pmatrix} Y_{a,T}^{\text{manuf}} \\ Y_{a,T}^{C1} \\ Y_{a,T}^{C2} \\ Y_{a,T}^{C3} \\ Y_{a,T}^{C4} \\ Y_{a,T}^{C5} \end{pmatrix}}_{Y_{a,T}} = \underbrace{\begin{pmatrix} \omega_{C1}^{a-1} & \omega_{C2}^{a-1} & \omega_{C3}^{a-1} & \omega_{C4}^{a-1} & \omega_{C5}^{a-1} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{S_{a,T}} \begin{pmatrix} Y_{a,T}^{C1} \\ Y_{a,T}^{C2} \\ Y_{a,T}^{C3} \\ Y_{a,T}^{C4} \\ Y_{a,T}^{C5} \end{pmatrix}$$

Notons de plus $\omega_{CL1/C4}^a$ et $\omega_{CL2/C4}^a$ les parts annuelles de production, dans le secteur des matériels de transport (C4), de l'automobile (CL1) et des autres matériels de transport (CL2). Au niveau A 64, la relation précédente s'écrit :

$$\underbrace{\begin{pmatrix} Y_{a,T}^{\text{manuf}} \\ Y_{a,T}^{C1} \\ Y_{a,T}^{C2} \\ Y_{a,T}^{C3} \\ Y_{a,T}^{C4} \\ Y_{a,T}^{CL1} \\ Y_{a,T}^{CL2} \\ Y_{a,T}^{C5} \end{pmatrix}}_{Y_{a,T}} = \underbrace{\begin{pmatrix} \omega_{C1}^{a-1} & \omega_{C2}^{a-1} & \omega_{C3}^{a-1} & \omega_{CL1}^{a-1} & \omega_{CL2}^{a-1} & \omega_{C5}^{a-1} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_{CL1/C4}^{a-1} & \omega_{CL2/C4}^{a-1} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{S_{a,T}} \begin{pmatrix} Y_{a,T}^{C1} \\ Y_{a,T}^{C2} \\ Y_{a,T}^{C3} \\ Y_{a,T}^{CL1} \\ Y_{a,T}^{CL2} \\ Y_{a,T}^{C5} \end{pmatrix}$$

Notons maintenant \hat{Y}_T le vecteur contenant toutes les prévisions des taux de croissance des productions au trimestre T . La méthode d'Hyndman consiste à corriger ces prévisions de la façon suivante :

$$\tilde{Y}_{a,T} = \underbrace{S_{a,T} (S'_{a,T} S_{a,T})^{-1} S'_{a,T}}_{W_{a,T}} \hat{Y}_{a,T}$$

La matrice $W_{a,T}$ est la matrice de correction des prévisions. Cette matrice est constante sur tous les trimestres d'une année. Au niveau A 17, pour tous les trimestres de l'année 2014 la matrice de correction des prévisions, $W_{2014,T}$, suivante est utilisée :

$$\begin{array}{c}
 \begin{array}{c} C \\ C1 \\ C2 \\ C3 \\ C4 \\ C5 \end{array}
 \begin{pmatrix}
 C & C1 & C2 & C3 & C4 & C5 \\
 0,23 & 0,16 & 0,06 & 0,08 & 0,11 & 0,36 \\
 0,16 & 0,97 & -0,01 & -0,02 & -0,02 & -0,07 \\
 0,06 & -0,01 & 1,00 & -0,01 & -0,01 & -0,03 \\
 0,08 & -0,02 & -0,01 & 0,99 & -0,01 & -0,04 \\
 0,11 & -0,02 & -0,01 & -0,01 & 0,98 & -0,05 \\
 0,36 & -0,07 & -0,03 & -0,04 & -0,05 & 0,80
 \end{pmatrix}
 \end{array}$$

Au niveau A 64, pour tous les trimestres de l'année 2014 c'est la matrice de correction des prévisions, $W_{2014,T}$, suivante qui est utilisée :

$$\begin{array}{c}
 \begin{array}{c} C \\ C1 \\ C2 \\ C3 \\ C4 \\ CL1 \\ CL2 \\ C5 \end{array}
 \begin{pmatrix}
 C & C1 & C2 & C3 & C4 & CL1 & CL2 & C5 \\
 0,22 & 0,16 & 0,06 & 0,09 & 0,04 & 0,04 & 0,04 & 0,36 \\
 0,16 & 0,97 & -0,01 & -0,02 & -0,01 & -0,01 & -0,01 & -0,08 \\
 0,06 & -0,01 & 1,00 & -0,01 & 0,00 & 0,00 & 0,00 & -0,03 \\
 0,09 & -0,02 & -0,01 & 0,99 & 0,00 & 0,00 & 0,00 & -0,04 \\
 0,04 & -0,01 & 0,00 & 0,00 & 0,33 & 0,31 & 0,35 & -0,02 \\
 0,04 & -0,01 & 0,00 & 0,00 & 0,31 & 0,85 & -0,17 & -0,02 \\
 0,04 & -0,01 & 0,00 & 0,00 & 0,35 & -0,17 & 0,81 & -0,02 \\
 0,36 & -0,08 & -0,03 & -0,04 & -0,02 & -0,02 & -0,02 & 0,83
 \end{pmatrix}
 \end{array}$$

Équations d'étalonnage

Dans ce chapitre sont présentées les différentes équations de prévision pour les trois méthodes de construction des étalonnages. Celles-ci ont été obtenues aux niveaux A 17 et A 64 de la Naf rév. 2 grâce à l'algorithme de sélection des variables Gets.

F.1 Codes des variables

Les tableaux F.1 et F.2 résument l'ensemble des soldes et des variables provenant respectivement de l'enquête de conjoncture de l'Insee et de l'enquête de conjoncture de la Banque de France, ainsi que les codes utilisés dans les différentes équations d'étalonnage.

Type de question	Code	Libellé
évolution	tppa	Tendance passée de la production sur les 3 derniers mois
évolution	tppre	Tendance prévue de la production sur les 3 prochains mois
situation	oscd	Niveau des carnets de commandes globaux
situation	ossk	Niveau des stocks de produits finis
autre	surprise	Indicateur de surprise

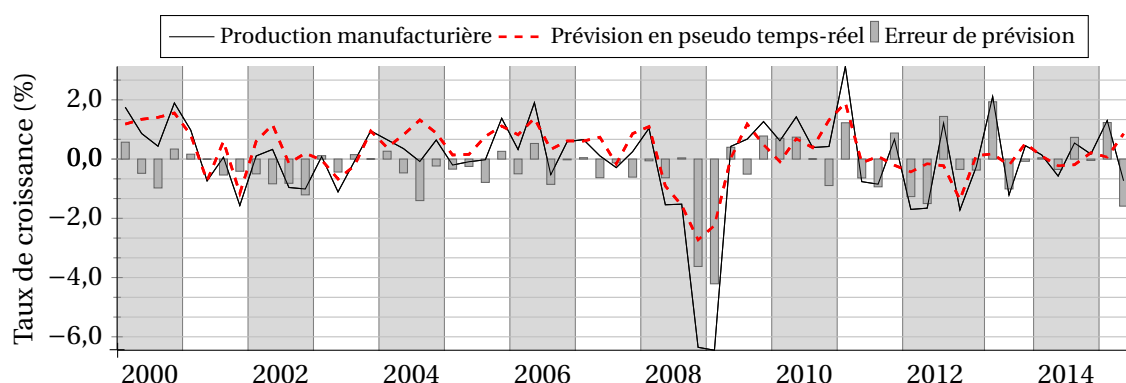
TABLEAU F.1 – Soldes et variables provenant de l'enquête de conjoncture dans l'industrie de l'Insee utilisés pour les étalonnages de la production.

Type de question	Code	Libellé
évolution	prodpas	Évolution de l'activité par rapport au mois précédent
évolution	prodpre	Prévisions de l'activité à 1 mois
évolution	evocar	Évolution des commandes reçues par rapport au mois précédent
situation	sitcar	Situation actuelle des carnets de commandes

TABLEAU F.2 – Soldes provenant de l'enquête de conjoncture de la Banque de France utilisés pour les étalonnages de la production.

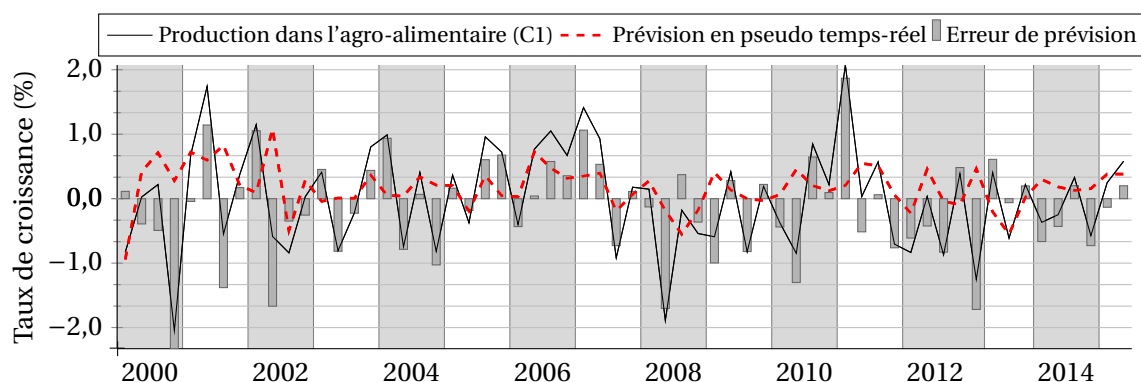
F.2 Étalonnages retenus

Les graphiques F.1 à F.8 montrent, pour la méthode des soldes bloqués, les prévisions en pseudo temps-réel. Les tableaux F.3 à F.6 présentent les différentes équations de prévision obtenues par l'algorithme Gets pour chaque méthode de construction d'étalonnage.



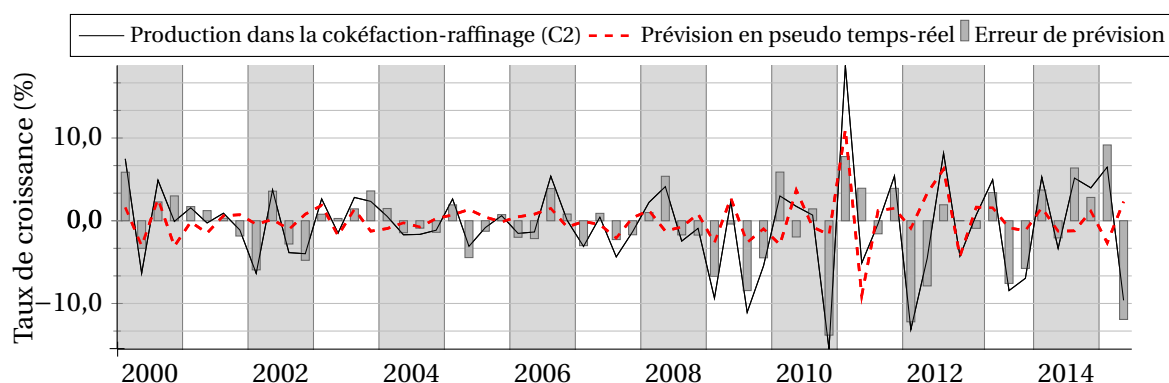
GRAPHIQUE F.1 – Prévisions en pseudo temps-réel de la production manufacturière du 1^{er} trimestre 2000 au 2^e trimestre 2015 obtenu par agrégation des étalonnages au niveau A 17 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



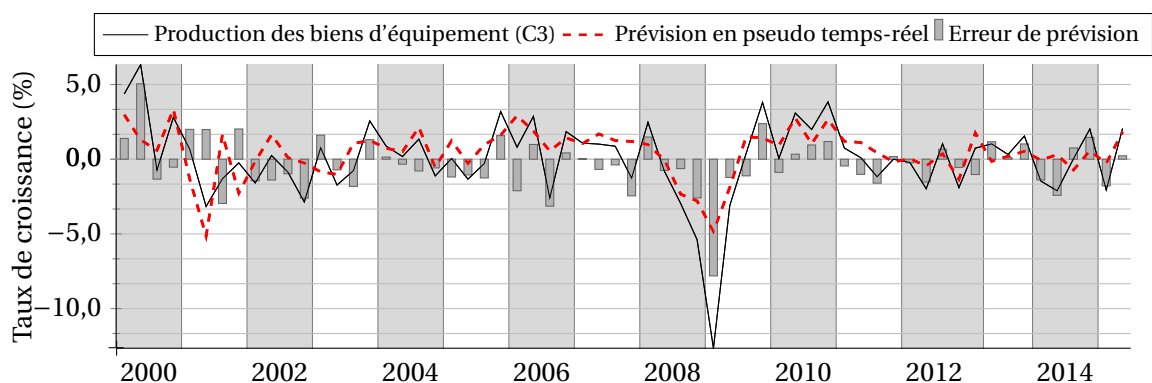
GRAPHIQUE F.2 – Prévisions en pseudo temps-réel de la production dans l'agro-alimentaire (C1) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



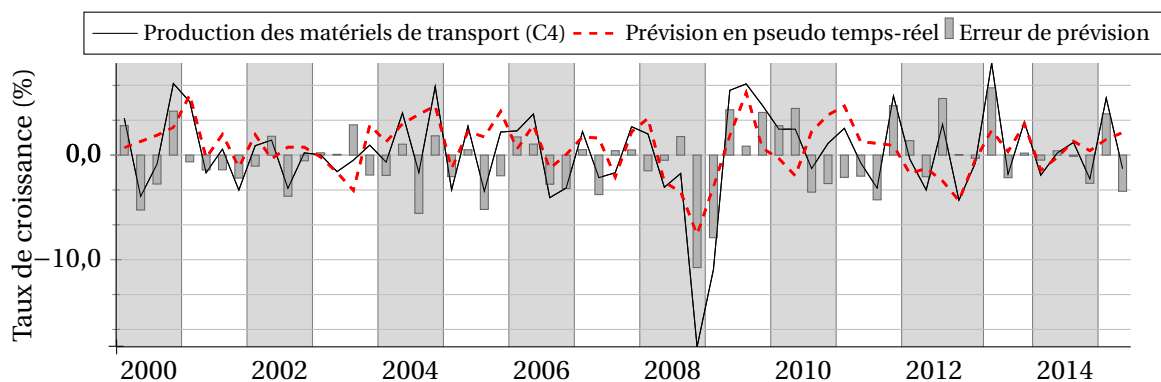
GRAPHIQUE F.3 – Prévisions en pseudo temps-réel de la production dans la cokéfaction-raffinage (C2) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



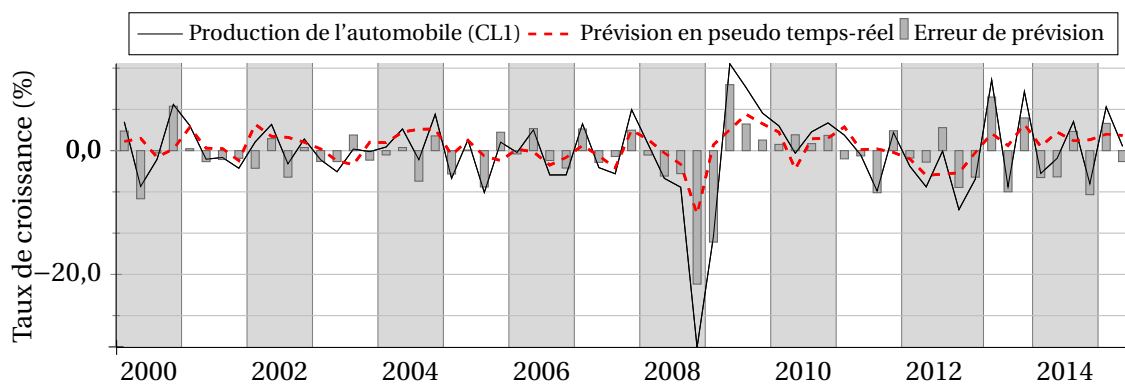
GRAPHIQUE F.4 – Prévisions en pseudo temps-réel de la production dans les biens d'équipement (C3) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



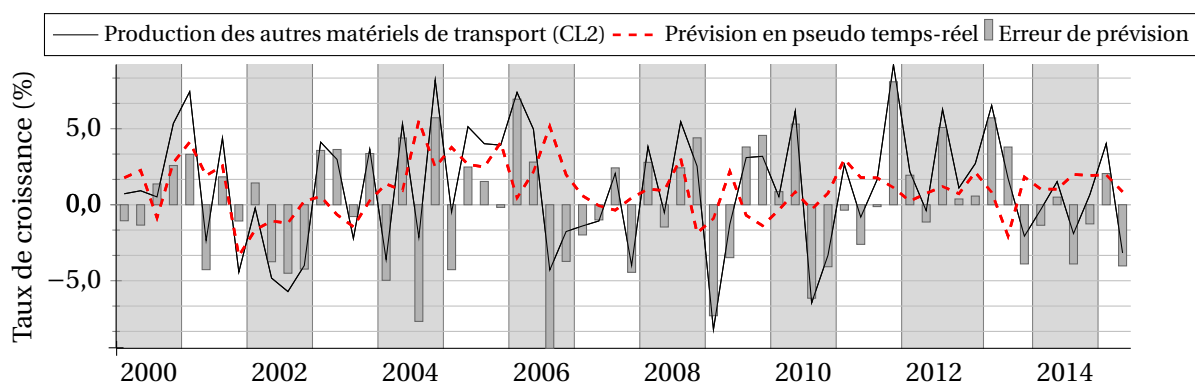
GRAPHIQUE F.5 – Prévisions en pseudo temps-réel de la production dans les matériels de transport (C4) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



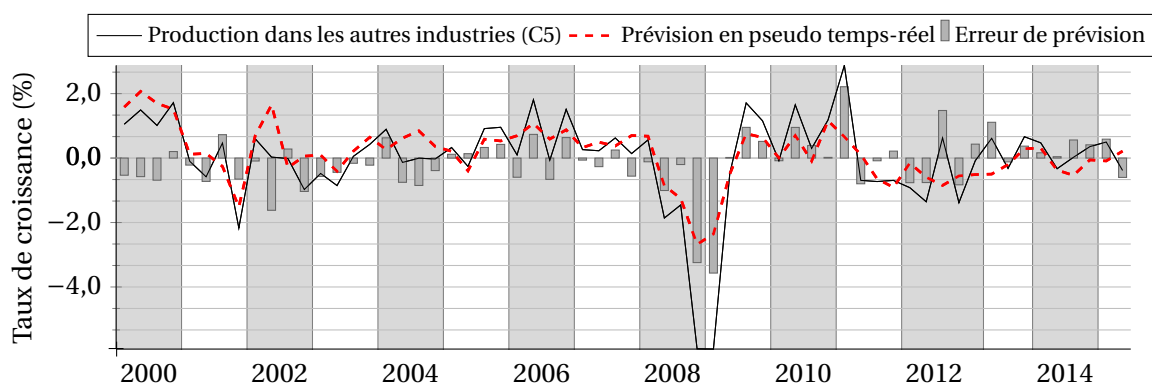
GRAPHIQUE F.6 – Prévisions en pseudo temps-réel de la production dans l'automobile (CL1) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



GRAPHIQUE F.7 – Prévisions en pseudo temps-réel de la production dans les autres matériels de transport (CL2) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.



GRAPHIQUE F.8 – Prévisions en pseudo temps-réel de la production dans les autres industries (C5) du 1^{er} trimestre 2000 au 2^e trimestre 2015 par la méthode des soldes bloqués.

Sources : Insee, Banque de France et calculs de l'auteur.

Secteur étudié	Estimations des étalonnages
C1	$\text{prod}_T = 0,21\text{acquis}_0\text{IPI}_T + 1,15\text{surprise}_{m2,T} + 0,01\text{evocar}_{m1,T}$ (0,09) (0,45) (0,01)
C2	$\text{prod}_T = -0,24\text{prod}_{T-1} + 0,45\text{acquis}_0\text{IPI}_T + 0,02\text{oscd}_{m2,T} + 0,03\text{tpa}_{m2,T} - 14,19\mathbb{I}_{2010T4}$ (0,07) (0,07) (0,01) (0,01) (3,27) $- 12,26\mathbb{I}_{2012T1}$ (3,18)
C3	$\text{prod}_T = 2,77\text{surprise}_{m2,T} + 0,07\text{prodpre}_{m1,T} + 0,11\Delta\text{tppre}_{T,m2} + 6,58\mathbb{I}_{1991T1} - 7,67\mathbb{I}_{2009T1}$ (1,01) (0,03) (0,02) (1,72) (1,75)
C4	$\text{prod}_T = 0,45\text{acquis}_0\text{IPI}_T - 0,06\text{evocar}_{m1,T} + 0,10\text{prodpas}_{m1,T} + 0,10\text{prodpre}_{m1,T}$ (0,14) (0,02) (0,02) (0,04) $+ 0,07\Delta\text{sitcar}_{T,m1}$ (0,02)
C4-CL1	$\text{prod}_T = 0,04\text{tpa}_{m2,T} - 2,34\text{surprise}_{m2,T} + 0,08\text{prodpas}_{m1,T} + 0,08\text{prodpre}_{m1,T}$ (0,01) (1,01) (0,02) (0,03) $- 0,04\text{sitcar}_{m1,T-1} - 19,85\mathbb{I}_{2008T4} - 16,68\mathbb{I}_{2009T1}$ (0,01) (3,94) (3,83)
C4-CL2	$\text{prod}_T = 0,12\text{evocar}_{m1,T}$ (0,03)
C5	$\text{prod}_T = 0,45 + 0,03\text{oscd}_{m2,T} + 0,04\text{evocar}_{m1,T} - 0,03\text{prodpas}_{m1,T} + 0,05\text{prodpre}_{m1,T}$ (0,22) (0,01) (0,02) (0,02) (0,02) $- 0,03\text{sitcar}_{m1,T-1} + 0,03\Delta\text{tppre}_{T,m2} - 3,16\mathbb{I}_{2008T4} - 3,70\mathbb{I}_{2009T1} + 2,23\mathbb{I}_{2011T1}$ (0,01) (0,01) (0,63) (0,62) (0,59)

TABLEAU F.3 – Étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 par la méthode des soldes bloqués (méthode 1) du 3^e trimestre 1990 au 4^e trimestre 2014.

Note de lecture : les chiffres entre parenthèses sont les écarts-types estimés des paramètres.

prod_T correspond au taux de croissance trimestriel de la production, en pourcentage, du secteur étudié au trimestre T .

Sources : Insee, Banque de France et calculs de l'auteur.

Secteur étudié	Estimations des étalonnages
C1	$\text{prod}_T = -0,24\text{prod}_{T-1} - 0,01\text{tpa}_T - 0,01\text{osk}_T + 0,01\text{evocar}_T + 0,48\text{surprise}_T + 0,01\Delta\text{oscd}_T$ (0,10) (0,00) (0,00) (0,00) (0,21) (0,00)
C2	$\text{prod}_T = -0,40\text{prod}_{T-1} - 0,20\text{prod}_{T-2} + 0,01\text{oscd}_T + 0,01\text{tpa}_T + 0,01\text{osk}_T - 14,34\mathbb{I}_{2010T4}$ (0,09) (0,10) (0,01) (0,01) (0,01) (3,86) $- 11,55\mathbb{I}_{2012T1}$ (3,92)
C3	$\text{prod}_T = -0,02\text{tpa}_T + 0,04\text{tppre}_T + 0,03\text{evocar}_T - 0,01\text{sitcar}_{T-1} + 6,90\mathbb{I}_{1991T1} - 7,36\mathbb{I}_{2009T1}$ (0,01) (0,01) (0,01) (0,00) (1,69) (1,74)
C4	$\text{prod}_T = 0,08\text{prodpre}_T + 0,03\text{evocar}_{T-1} - 0,02\text{sitcar}_{T-1}$ (0,01) (0,01) (0,00)
C4-CL1	$\text{prod}_T = -1,70 - 0,02\text{oscd}_T - 0,02\text{osk}_T + 0,06\text{prodpas}_T + 0,06\text{prodpre}_T$ (0,57) (0,01) (0,01) (0,01) (0,02)
C4-CL2	$\text{prod}_T = 0,05\text{evocar}_T$ (0,01)
C5	$\text{prod}_T = 0,01\text{evocar}_T + 0,02\text{prodpre}_T + 0,01\text{oscd}_{T-1} - 0,01\text{sitcar}_{T-1} + 0,02\Delta\text{tppre}_T$ (0,01) (0,01) (0,00) (0,00) (0,00) $- 3,34\mathbb{I}_{2008T4} - 3,15\mathbb{I}_{2009T1} + 2,35\mathbb{I}_{2011T1}$ (0,62) (0,63) (0,60)

TABLEAU F.4 – Résultats des étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 sans l'IPI (méthode 3) du 3^e trimestre 1990 au 4^e trimestre 2014.

Note de lecture : les chiffres entre parenthèses sont les écarts-types estimés des paramètres.

prod_T correspond au taux de croissance trimestriel de la production, en pourcentage, du secteur étudié au trimestre T .

Sources : Insee, Banque de France et calculs de l'auteur.

Secteur étudié	Estimations des étalonnages
C1	$\text{prod}_T = 0,51\text{IPI}_T - 0,01\text{tpa}_T - 0,01\text{osk}_T + 0,01\text{evocar}_T + 0,41\text{surprise}_T$ (0,04) (0,00) (0,00) (0,00) (0,13)
C2	$\text{prod}_T = 0,76\text{IPI}_T + 0,01\text{tpa}_T$ (0,02) (0,00)
C3	$\text{prod}_T = 0,70\text{IPI}_T + 0,03\text{tppre}_T + 0,02\text{evocar}_T - 0,04\text{sitcar}_T - 0,02\text{tppre}_{T-1} + 0,03\text{sitcar}_{T-1}$ (0,08) (0,01) (0,01) (0,01) (0,01) (0,01)
	$+ 4,07\mathbb{I}_{1991T1} + 4,60\mathbb{I}_{1991T2}$ (1,31) (1,26)
C4	$\text{prod}_T = 0,98\text{IPI}_T + 0,02\text{prodpre}_T - 3,57\mathbb{I}_{2009T3} - 3,31\mathbb{I}_{2010T4}$ (0,04) (0,01) (1,31) (1,27)
C4-CL1	Pas de modèle valide pour la prévisions de l'IPI (problème de normalité des résidus)
C4-CL2	$\text{prod}_T = 0,12\text{prod}_{T-1} + 1,02\text{IPI}_T + 0,00\text{osk}_T + 0,00\text{oscd}_{T-1} + 0,01\text{tppre}_{T-1}$ (0,05) (0,06) (0,00) (0,00) (0,00)
C5	$\text{prod}_T = 0,62\text{IPI}_T + 0,01\text{tpa}_T + 0,01\text{osk}_T + 0,01\text{prodpre}_T + 0,00\text{oscd}_{T-1} - 0,01\text{tppre}_{T-1}$ (0,02) (0,00) (0,00) (0,00) (0,00) (0,00)
	$+ 0,00\text{sitcar}_{T-1} + 0,01\text{prodpre}_{T-1}$ (0,00) (0,00)

TABLEAU F.5 – Résultats des étalonnages de la production au niveau A 17 et A 64 de la Naf rév. 2 avec l'IPI connu sur tout le trimestre coïncident (méthode 2) du 3^e trimestre 1990 au 4^e trimestre 2014.

Note de lecture : les chiffres entre parenthèses sont les écarts-types estimés des paramètres.

prod_T correspond au taux de croissance trimestriel de la production, en pourcentage, du secteur étudié au trimestre T .

Sources : Insee, Banque de France et calculs de l'auteur.

Nom du modèle	Estimations du modèle
C1	$\text{IPI}_t = -0,99 - 0,63\text{IPI}_{t-1} - 0,41\text{IPI}_{t-2} - 0,21\text{IPI}_{t-3} + 0,04\text{evocar}_t + 0,03\text{tppre}_t$ (0,20) (0,05) (0,06) (0,05) (0,01) (0,01)
	$- 0,02\text{oscd}_{t-3} + 4,92\mathbb{I}_{\text{mai}1992} - 6,54\mathbb{I}_{\text{aoû.1992}} + 4,18\mathbb{I}_{\text{mai}1996}$ (0,01) (1,25) (1,25) (1,25)
C2	$\text{IPI}_t = 0,30\text{IPI}_{t-1} - 0,42\text{IPI}_{t-3} + 0,02\text{tpa}_{t+1} - 59,27\mathbb{I}_{\text{oct.2010}} + 105,85\mathbb{I}_{\text{janv.2011}}$ (0,05) (0,05) (0,01) (6,49) (6,91)
	$- 35,51\mathbb{I}_{\text{fév.2011}} + 40,08\mathbb{I}_{\text{avr.2011}}$ (8,65) (8,86)
C3	$\text{IPI}_t = -0,96 - 0,54\text{IPI}_{t-1} - 0,22\text{IPI}_{t-2} - 0,03\text{oscd}_t + 0,04\text{tppre}_t + 0,04\text{prodpas}_t$ (0,19) (0,05) (0,05) (0,01) (0,01) (0,01)
	$+ 0,03\text{evocar}_t + 5,60\mathbb{I}_{\text{aoû.1997}} - 15,42\mathbb{I}_{\text{janv.2009}} - 6,05\mathbb{I}_{\text{fév.2009}}$ (0,01) (1,43) (1,44) (1,66)
C4	$\text{IPI}_t = -0,51\text{IPI}_{t-1} - 0,21\text{IPI}_{t-2} - 0,11\text{IPI}_{t-3} - 0,06\text{prodpre}_t + 0,06\text{sitcar}_t + 0,07\text{prodpas}_t$ (0,05) (0,05) (0,05) (0,03) (0,03) (0,02)
	$+ 0,13\text{prodpre}_{t-1} - 0,06\text{sitcar}_{t-1} - 13,32\mathbb{I}_{\text{aoû.1999}} - 13,56\mathbb{I}_{\text{aoû.2004}}$ (0,03) (0,03) (3,11) (3,11)
C4-CL1	Pas de modèle valide pour la prévisions de l'IPI (problème de normalité des résidus)
C4-CL2	$\text{IPI}_t = -0,71\text{IPI}_{t-1} - 0,47\text{IPI}_{t-2} - 0,21\text{IPI}_{t-3} + 0,03\text{tpa}_t + 14,81\mathbb{I}_{\text{avr.2010}}$ (0,06) (0,06) (0,06) (0,01) (4,86)
C5	$\text{IPI}_t = -0,98 - 0,56\text{IPI}_{t-1} - 0,28\text{IPI}_{t-2} + 0,03\text{tppre}_t + 0,06\text{evocar}_t - 0,02\text{oscd}_{t-3}$ (0,13) (0,05) (0,05) (0,01) (0,01) (0,00)
	$+ 3,86\mathbb{I}_{\text{aoû.1997}} - 4,35\mathbb{I}_{\text{nov.2008}} - 5,66\mathbb{I}_{\text{déc.2008}} - 4,18\mathbb{I}_{\text{janv.2009}}$ (1,04) (1,07) (1,10) (1,11)

TABLEAU F.6 – Résultats des étalonnages de l'IPI au niveau A 17 et A 64 de la Naf rév. 2 du 1^{er} trimestre 1991 au au 4^e trimestre 2014.

Note de lecture : les chiffres entre parenthèses figurent les écarts-types.

IPI_t correspond au taux de croissance mensuel de l'IPI, en pourcentage, du secteur étudié au mois t .

Sources : Insee, Banque de France et calculs de l'auteur.

Bibliographie

- [1] « L'enquête de conjoncture sur la situation et les perspectives dans l'industrie : méthodologie » – *Insee Méthodes* (2007), no. 117.
- [2] « Méthodologie des comptes trimestriels en base 2005 » – *Insee Méthodes* (2012), no. 126.
- [3] D. ANDERSON et K. BURNHAM – « AIC myths and misunderstandings », (2006).
- [4] F. ARNAUD – « Calcul des contributions en volumes chaînés », (2014).
- [5] P. BUMS – « Robustness of the ljung-box test and its rank equivalent », *Correlations and Volatilities of Asynchronous Data, The Journal of Derivatives*, 2002, p. 7–18.
- [6] J. CASTLE, D. HENDRY et O. KITOV – « Forecasting and Nowcasting Macroeconomic Variables : A Methodological Overview », Economics Series Working Papers 674, University of Oxford, Department of Economics, 2013.
- [7] F. X. DIEBOLD – « Comparing predictive accuracy, twenty years later : a personal perspective on the use and abuse of Diebold-Mariano tests », Working Paper 18391, National Bureau of Economic Research, September 2012.
- [8] J. A. DOORNIK, D. F. HENDRY et F. PRETIS – « Step indicator saturation », Economics Series Working Papers 658, University of Oxford, Department of Economics, June 2013.
- [9] É. DUBOIS et E. MICHAUX – « Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats », *Économie & prévision* **172** (2006), no. 1, p. 11–28.
- [10] N. ERICSSON et E. L. REISMAN – « Evaluating a global vector autoregression for forecasting », International Finance Discussion Papers 1056, Board of Governors of the Federal Reserve System (U.S.), 2012.
- [11] Y. HAUSEUX, P.-D. OLIVE, C. RENNE et C. BORTOLI – « De nouvelles avancées dans l'utilisation des enquêtes de conjoncture de l'Insee pour le diagnostic conjoncturel », *Note de conjoncture* (2015), p. 21–41.
- [12] K. HOOVER et S. PEREZ – « Data mining reconsidered : encompassing and the general-to-specific approach to specification search », *Econometrics Journal* **2** (1999), no. 2, p. 167–191.
- [13] R. J. HYNDMAN, R. A. AHMED, G. ATHANASOPOULOS et H. L. SHANG – « Optimal combination forecasts for hierarchical time series », *Computational Statistics & Data Analysis* **55** (2011), no. 9, p. 2579 – 2589.
- [14] H.-M. KROLZIG et D. HENDRY – « Computer automation of general-to-specific model selection procedures », *Journal of Economic Dynamics and Control* **25** (2001), no. 6-7, p. 831–866.
- [15] E. LEBARBIER et T. MARY-HUARD – « Le critère BIC : fondements théoriques et interprétation », Research Report RR-5315, 2004.
- [16] G. SUCARRAT et A. ESCRIBANO – « Automated model selection in finance : general-to-specific modelling of the mean and volatility specifications », *Oxford Bulletin of Economics and Statistics* **74** (2012), no. 5, p. 716–735.