

COMMENT JUGER DE LA QUALITÉ D'UNE DÉSAISONNALISATION ?

Stéphanie HIMPENS ()*

() Insee, Unité Méthodologie d'Entreprises*

Résumé

Depuis les années 90, un effort d'harmonisation des pratiques européennes en matière de désaisonnalisation a été mené. Le développement, la maintenance et la promotion d'un logiciel européen de désaisonnalisation, Demetra, en constituent l'un des aspects.

Afin de poursuivre cet effort, un groupe d'experts européens a été constitué : le groupe de pilotage sur les ajustements saisonniers. Ce dernier est à l'origine de la publication en 2008 du code de bonnes pratiques européennes de la désaisonnalisation (ESS guidelines for seasonal adjustment). Ce document donne un cadre afin de juger de la qualité de désaisonnalisation d'une série.

Certaines préconisations de ce guide se révélaient impossibles à mettre en œuvre avec le logiciel de désaisonnalisation existant (Demetra v2.2). Cet outil posait de plus des problèmes de maintenance et semblait difficile à faire évoluer. Le groupe de pilotage a donc décidé la mise en place d'une nouvelle solution logicielle. **Demetra+** a été développé par la banque nationale de Belgique. Ce nouveau logiciel donne accès à un ensemble de diagnostics plus large que dans le précédent logiciel.

Demetra+ est utilisé depuis janvier 2011 afin de désaisonnaliser les séries du DCST (Département des Statistiques de Court Terme) avec la méthode X-12-ARIMA. Ce département s'efforce autant que possible de mettre en œuvre les préconisations des lignes directrices européennes.

Dans ce département, plusieurs centaines de séries sont désaisonnalisés tous les mois afin de calculer l'indice de production industrielle (IPI - 91 séries au niveau NACE3 de nomenclature) et les indices de chiffres d'affaires (env. 440 séries au niveau NACE3 de nomenclature dans l'industrie et NACE4 dans les autres secteurs). Les modèles servant à la désaisonnalisation sont fixés une fois par an. Seuls les coefficients sont ensuite ré-estimés (ajustement de type concurrent partiel) tous les mois. Les points aberrants (outliers) sont également recherchés sur la fin des séries.

Lors de la recherche de modèles sur un grand nombre de séries, il est important de pouvoir comparer rapidement la qualité des différents modèles possibles afin de sélectionner le modèle le plus approprié. Les séries dont la qualité se dégrade et pour lesquelles les spécifications ne sont plus appropriées doivent également être détectées rapidement.

La nouvelle application **Demetra+** fournit de nombreux diagnostics très détaillés pour toutes les séries. Ces diagnostics sont de plus pour la plupart communs aux deux principales méthodes de désaisonnalisation (**TRAMO-SEATS** et **X-12-ARIMA**). Le logiciel permet également de changer interactivement certaines spécifications et de visualiser immédiatement le résultat.

Le DCST a décidé d'aller plus loin : un ensemble de macros SAS fonctionnant avec **Demetra +** permet de noter la qualité des séries. Outre l'implémentation de diagnostics propres au département, ce bilan qualité donne une vue globale de la qualité des séries d'un indice via l'édition de notes globales et détaillées suivant différentes catégories de diagnostics.

Les indices du DCST font l'objet d'une désaisonnalisation indirecte : les agrégats sont calculés comme des sommes pondérées des séries élémentaires les constituant qui font l'objet d'une désaisonnalisation. Il a été mis en place un système permettant d'évaluer la qualité des séries agrégées. Il est ainsi possible, en cas de détérioration de la qualité d'un agrégat de visualiser les séries qui y contribuent le plus (analyse descendante).

Des programmes de comparaison de la qualité de désaisonnalisations obtenues avec différents jeux de spécifications ont également été créés.

1. Sommaire

RÉSUMÉ	1
1. SOMMAIRE	2
2. CADRE GÉNÉRAL	34
2.1. BREF RAPPEL SUR LA SAISONNALITÉ	34
2.2. LE CADRE EUROPÉEN	45
2.3. LOGICIELS ET PROGRAMMES	45
2.4. UNE MULTITUDE DE DIAGNOSTICS MOBILISABLES	46
2.5. DE LA NÉCESSITÉ DE RÉSUMER L'INFORMATION	56
3. AMÉLIORER LA QUALITÉ DES SÉRIES D'UN INDICE	67
3.1. ÉVALUATION DE LA QUALITÉ D'UNE SÉRIE	67
3.1.1. LE NOEUD « MAIN RESULT » DE DEMETRA+	67
3.1.1.1. Limites de l'évaluation de la qualité dans Demetra+	910
3.1.2. LE RÉSUMÉ DE LA QUALITÉ DANS LE BILAN QUALITÉ	1011
3.1.2.1. Qualité des points aberrants	1112
3.1.2.2. Qualité des résidus du modèle reg-ARIMA	1314
3.1.2.3. Qualité de décomposition	15
3.1.2.4. Absence d'effets résiduels	16
3.1.2.5. Récapitulatif des diagnostics pris en compte dans le bilan qualité	1617
3.1.2.6. Révisions	1718
3.2. COMPARAISON DES DIAGNOSTICS PRÉSENTS DANS LE RÉSUMÉ DE DEMETRA+ ET LE BILAN QUALITÉ	1718
3.2.1. QUALITÉ DES PRÉ-AJUSTEMENTS	1718
3.2.2. QUALITÉ DE LA DÉCOMPOSITION	1819
3.2.3. QUALITÉ DES RÉVISIONS	1819
3.3. UTILISATIONS PRATIQUES DES RÉSUMÉS DE LA QUALITÉ DE DEMETRA+ ET DU BILAN QUALITÉ	1819
3.3.1. ANALYSE DE LA QUALITÉ DE CHAQUE SÉRIE	19
3.3.2. TRIS SUIVANT UN CRITÈRE DE QUALITÉ	1920
3.3.3. COMPARAISON DE QUALITÉ	1920
3.3.4. ÉVALUATION DE LA QUALITÉ D'UN INDICE DÉSAISONNALISÉ DE FAÇON INDIRECTE	1920
3.4. ORGANISATION DU TRAITEMENT DES CVS AU DSCT (DÉPARTEMENT DES STATISTIQUES DE COURT TERME)	2021
3.4.1. SCHÉMA D'UNE CAMPAGNE ANNUELLE DE DÉSAISONNALISATION DU POINT DE VUE DE LA QUALITÉ.	2122
4. COMPARAISON DE QUALITÉS ET DE RÉSULTATS	2324
4.1. COMPARAISON DES RÉSULTATS DE DEUX BILANS QUALITÉ	2324
4.1.1. ORGANISATION DU FICHIER EXCEL DE COMPARAISON DES DIAGNOSTICS	2324
4.2. COMPARAISONS GRAPHIQUES DE RÉSULTATS	2627
5. ÉVALUATION DE LA QUALITÉ D'UN INDICE DÉSAISONNALISÉ DE FAÇON INDIRECTE	2728
5.1. ÉVALUATION DE LA QUALITÉ DU MODÈLE REG-ARIMA	2728
5.2. ÉVALUATION DE LA QUALITÉ DE DÉCOMPOSITION	2930
5.3. ÉVALUATION DE LA QUALITÉ GÉNÉRALE DES SÉRIES	3132
5.4. ÉVALUATION DE LA QUALITÉ DES NIVEAUX D'AGRÉGATION SUPÉRIEURS	3132

2. cadre général

2.1. Bref rappel sur la saisonnalité

La plupart des séries économiques présentent des fluctuations qui se répètent d'année en année. Par exemple, de nombreuses séries du secteur de l'industrie baissent au mois d'août. Ces variations saisonnières ont des causes diverses qu'il n'est pas possible de modéliser de façon classique, avec une modélisation paramétrique faisant intervenir différents facteurs explicatifs¹. La plupart du temps, ces fluctuations masquent le message économique que l'on souhaite appréhender. Par exemple, ce que l'on souhaite savoir n'est pas que les séries de l'industrie baissent au mois d'août mais si elles baissent plus fortement que lors d'un mois d'août « moyen ».

Les méthodes de désaisonnalisation s'efforcent d'extraire la composante saisonnière des séries afin de rendre évidentes les fluctuations plus intéressantes (d'un point de vue économique).

Souvent, il est considéré que la série présente des évolutions perturbatrices aléatoires non saisonnières et parfois dues à des événements exceptionnels n'ayant rien à voir avec la dynamique propre de la série (grèves, accidents climatiques etc.). Il est parfois utile de pouvoir séparer cette composante d'évolutions de plus long terme (tendance et cycle).

Voilà pourquoi la plupart des méthodes de désaisonnalisation s'attachent à décomposer les séries en :

$$Y_t = S_t + T_t + Irr_t$$

où Y_t est la série initiale (éventuellement transformée), S_t la saisonnalité, T_t la tendance-cycle (le cycle est souvent indissociable de la tendance compte tenu de la longueur des séries étudiées) et Irr_t la composante irrégulière.

Intuitivement, on s'attendrait à ce que ces composantes vérifient certaines propriétés telles que :

- La saisonnalité s'annule (ou presque) sur une année afin que le niveau de la série désaisonnalisée ne soit pas trop différent de la série initiale
- La saisonnalité présente un profil n'évoluant pas trop d'une année sur l'autre
- L'irrégulier soit une composante aléatoire et donc ne présente pas de structure
- La tendance présente des évolutions lentes et non erratiques.
- etc.

Le problème est que les différentes composantes extraites ne peuvent être observées en pratique (modèles à composantes inobservables). Leurs définitions restent donc assez vagues ce qui complique l'établissement de diagnostics. Par exemple, certaines séries présentent des saisonnalités marquées mais semblent évoluer fortement sur la période d'analyse. Il semblerait logique que la saisonnalité extraite évolue ce qui semble contradictoire avec l'idée d'un schéma répétitif d'une année sur l'autre. La plupart des séries économiques sont affectées par des effets déterministes provoqués par la structure du calendrier (présence de jours ouvrables etc.). Ces effets perturbent le plus souvent la décomposition. Il est usuel de les corriger avant tout traitement de la saisonnalité et de les placer dans une composante séparée.

De multiples méthodes de désaisonnalisation existent. Elles peuvent être regroupées en deux grandes familles : les méthodes paramétriques et non paramétriques. Les méthodes non paramétriques présentent le désavantage de ne pas posséder de modèles explicites et en conséquence ne permettent pas de tests sur les paramètres ou la forme des composantes. Il est en revanche possible de tester l'adéquation des modèles des méthodes paramétriques aux données. Mais rien ne garantit que ces modèles sont les plus appropriés...

¹ Bien qu'il soit parfois possible de modéliser une saisonnalité très stable avec des indicatrices.

2.2. Le cadre européen

La diversité des méthodes et des façons d'appréhender la saisonnalité se révèle problématique en particulier lorsque l'on souhaite comparer et agréger les résultats comme c'est le cas entre les différents pays européens. L'harmonisation des pratiques est alors essentielle à la qualité des résultats obtenus. Depuis les années 90, des efforts ont été menés dans ce sens au sein du système statistique européen.

Le code de bonnes pratiques européennes en matière de désaisonnalisation (ESS guidelines for seasonal adjustment) publié en 2008 donne un cadre afin de juger de la qualité d'une désaisonnalisation. Il fixe un ensemble de règles, de préconisations vers lesquelles les pays doivent tendre. Ce document fixe en outre une hiérarchie des pratiques existantes en indiquant les meilleures pratiques vers lesquelles chaque pays doit tendre mais aussi les pratiques acceptables et celles devant absolument être évitées.

Le code de bonnes pratiques concerne en majorité les deux principales méthodes de désaisonnalisation principalement utilisées au sein du système statistique européen (**TRAMO-SEATS** (méthode paramétrique développée par la banque nationale d'Espagne²) et **X-12-ARIMA** (méthode non paramétrique issue du bureau du recensement américain³)). Les principaux diagnostics détaillés dans ce document sont donc ceux utilisés afin de juger de la qualité de désaisonnalisation obtenue avec ces deux méthodes.

2.3. Logiciels et programmes

Les deux principales méthodes de désaisonnalisation peuvent être mises en œuvre grâce à des exécutables dos écrit en FORTRAN mis à disposition sur les sites de la banque nationale d'Espagne et du bureau du recensement américain. Des interfaces ont été développées afin de rendre leur utilisation plus conviviale : **Win-X12** pour **X-12-ARIMA** et **TSW** pour **TRAMO-SEATS**.

Afin que les pays du système statistique européen puissent plus aisément implémenter les lignes directrices, il a été décidé le développement d'un nouveau logiciel **Demetra+** par la banque nationale de Belgique (en remplacement de l'ancienne solution logicielle Demetra v2.2) permettant en outre d'utiliser indifféremment l'une ou l'autre des deux méthodes précédentes. Ce logiciel fournit des diagnostics très détaillés et communs aux deux principales méthodes de désaisonnalisation. Ces diagnostics sont accessibles pour chacune des séries désaisonnalisées.

Depuis janvier 2011, Demetra+ est utilisé afin de désaisonnaliser les séries du DSCT (Département des Statistiques de Court Terme) de l'INSEE et plus particulièrement au sien de la division ICA (Indicateurs Conjoncturels d'Activité) à l'aide de la méthode X-12-ARIMA.

2.4. Une multitude de diagnostics mobilisables

Une multitude de diagnostics sont disponibles au sein des trois logiciels (**Win-X12**, **TSW** et **Demetra+**) afin d'évaluer les performances des deux principales méthodes de désaisonnalisation. Certains sont communs, d'autres sont spécifiques à l'une ou à l'autre des méthodes.

Ces diagnostics n'ont généralement pas la même importance et cette importance dépend de la méthode mise en œuvre (par exemple, l'oubli d'un point atypique lors de la phase de pré-ajustements n'aura pas la même incidence sur une désaisonnalisation effectuée avec **X11** qui dispose d'une méthode interne de détection des valeurs atypiques qu'avec **TRAMO-SEATS** dont la décomposition repose entièrement sur l'estimation du modèle ARIMA).

² voir le site <http://www.bde.es/webbde/en/secciones/servicio/software/econom.html>.

³ voir le site http://www.census.gov/srd/www/x12a/x12downv03_pc.html.

Les diagnostics peuvent être regroupés en différents ensembles (selon qu'ils concernent la phase de pré-ajustements, la décomposition de la série, la présence d'effets résiduels, les révisions etc.). Chaque groupe concerne une phase précise ou un aspect du processus de désaisonnalisation. Les informations fournies par ces différents types de diagnostics restent fortement corrélées (par exemple, une absence de modélisation d'effets cjo alors qu'ils sont requis peut avoir des conséquences sur les résidus du modèle reg-ARIMA dans la phase d'ajustements préalables mais aussi sur la série ajustée). Ils sont à considérer simultanément lorsque l'on cherche à améliorer la désaisonnalisation.

2.5. De la nécessité de résumer l'information

Il peut être difficile de se repérer dans cette multitude de diagnostics et juger de la qualité générale de désaisonnalisation d'une série.

Parfois, l'amélioration d'une statistique en dégrade une autre... et certains diagnostics restent mauvais. Quel est alors l'impact de l'échec à un ou plusieurs tests ou de certains diagnostics insatisfaisants sur la qualité globale de la désaisonnalisation ? Les résultats sont-ils tout de même utilisables ou faut-il modifier les paramètres ? Lors de l'analyse d'une multitude de séries dans un temps très court, il est difficile de procéder à une recherche de paramètres en profondeur dans tous les cas. Il peut alors être utile de s'intéresser en priorité aux traitements les plus problématiques et donc de disposer d'une évaluation de la qualité globale de désaisonnalisation des séries.

Demetra+ offre une liste de commentaires concernant la qualité pour un ensemble de diagnostics jugés les plus importants. Une synthèse générale de la qualité de chaque série est effectuée. Cette dernière est immédiatement visible pour l'ensemble des séries constitutives de l'indice. Il est alors facile de repérer les cas les plus problématiques et de concentrer les efforts sur ces derniers.

Un tel résumé de la qualité est toutefois insatisfaisant : il n'est pas possible de modifier les diagnostics pris en compte dans le résumé (il est seulement possible d'en exclure certains mais pas d'en ajouter d'autres qui sembleraient plus pertinents à l'utilisateur), le poids des différents diagnostics est difficilement modifiable. Il n'est pas possible de visualiser les commentaires concernant la qualité des différents groupes de diagnostics de façon immédiate pour toutes les séries (sous forme d'une table par exemple), etc.

Un ensemble de programmes propre au DSCT a été écrit en SAS afin d'évaluer la qualité de désaisonnalisation des différentes séries avec **Demetra+**. Ce travail reprend, en l'améliorant, une version antérieure écrite pour fonctionner avec l'ancien logiciel Demetra⁴. Dans cet ensemble de programmes SAS la qualité de différents groupes de diagnostics fait l'objet d'une notation sous forme numérique. Cet ensemble de programmes sera désigné sous le terme de « **bilan qualité** » dans la suite de ce document. Le **bilan qualité** permet de palier les manques constatés dans **Demetra+**. Il introduit également d'autres diagnostics non présents dans l'application comme des diagnostics sur la concentration des points aberrants etc.

La partie 3.1 détaille les modalités des synthèses de qualité effectuées par le logiciel **Demetra+** et les programmes SAS du bilan qualité.

Il peut parfois également être plus intéressant d'évaluer la qualité non pas en niveau mais en comparaison. La question est alors : lequel de ces deux jeux de paramètres conduit aux meilleurs résultats ou encore est-ce que le nouveau jeu de paramètres choisi à l'issue de la recherche de modèle ne conduit pas à une détérioration de la qualité ? Le bilan qualité permet donc non seulement de juger des qualités des désaisonnalisations de chaque série une à une mais également de comparer les qualités de différentes désaisonnalisations. Il est possible de comparer les qualités de désaisonnalisations obtenues avec des paramètres différents pour une même série ce qui aide à choisir le meilleur modèle. Cette fonctionnalité du bilan qualité sera abordée en partie 4 du présent document.

Dans le cas d'une désaisonnalisation indirecte, les séries peuvent présenter des poids différents dans les agrégats de niveaux supérieurs. L'amélioration de la qualité de traitement de certaines séries a

⁴ Voir travaux de Johara Khelif, Cynthia Faivre et Emmanuel Gros (INSEE)

alors plus d'impact sur certains agrégats de niveau supérieur que d'autres. Le statisticien ne disposant pas du temps suffisant pour analyser en détail toutes les séries aura alors intérêt à rechercher en particulier des paramètres satisfaisants pour ces séries. Des programmes SAS ont donc été écrits afin d'observer la qualité aux différents niveaux d'agrégation d'un indice désaisonnalisé de façon indirecte et ainsi de repérer les séries les plus contributrices à la détérioration de la qualité globale (analyse descendante de la qualité). L'évaluation de la qualité des agrégats fait l'objet de la partie 5.

3. Améliorer la qualité des séries d'un indice

3.1. Évaluation de la qualité d'une série

Pour chaque série, nous avons vu dans le chapitre introductif qu'il existait une multitude de diagnostics parfois redondants en termes d'informations. Quand les séries sont nombreuses et que le temps d'analyse est très court, il est difficile de les consulter tous et il est facile de négliger un diagnostic important.

Les diagnostics se répartissent en grands groupes ce qui permet de cibler les parties de la modélisation les plus problématiques. Un résumé de l'information de l'ensemble des diagnostics est élaboré au niveau de chaque série par **Demetra+**.

Il est alors possible de cibler les séries les plus problématiques du point de vue de la qualité au sein d'un ensemble, la qualité des diagnostics des différents groupes apparaissant de façon visuelle. Ces séries peuvent ensuite faire l'objet d'une analyse plus approfondie à l'aide du détail des diagnostics.

Les diagnostics utilisés dans la confection du résumé de **Demetra+** ne sont que peu paramétrables (il n'est pas possible par exemple de choisir un diagnostic en dehors d'une liste très restreinte ou de pondérer les diagnostics selon leur importance). Il a donc été décidé d'aller plus loin en élaborant un diagnostic synthétique propre au DSCT au sein du **bilan qualité**.

3.1.1. le nœud « Main Result » de Demetra+

La qualité est un commentaire.

Les diagnostics synthétiques de **Demetra+** reposent sur des critères de type **qualitatif** : il existe plusieurs niveaux de qualité suivant la réussite ou l'échec à certains diagnostics. Le but est de fournir une indication qualitative de l'information fournie par le test (est-ce que la valeur du test est acceptable ?). Cette information qualitative présente l'avantage d'être directement compréhensible par l'utilisateur.

Signification des différents labels de qualité

Valeur	Signification
Undefined	Échec du calcul du label de qualité : la qualité est indéfinie (non-calcul d'un test etc.)
Error	Il y a une erreur dans les résultats (certaines contraintes ne sont pas respectées (cf. test de validité de la procédure) ou présence de valeurs aberrantes extrêmes)
Severe	Le calcul est correct mais les données sont insatisfaisantes en termes de qualité.
Bad	La qualité est mauvaise (certains tests sont rejetés).
Uncertain	Certains tests de qualité ont des valeurs limites.
Good	La qualité est bonne

Ces différents « commentaires » sur la qualité sont édités pour les principaux diagnostics dans la fenêtre de résumé des diagnostics. Il est ainsi possible d'avoir un aperçu global de la qualité de désaisonnalisation d'un indice.

Détails des commentaires de qualité du nœud Main results sous Demetra+

Pre-processing (RegArima)

Estimation span: [1-1990 : 2-2011]
Series has been log-transformed
Trading days effects (6 variables)
No easter effect
5 pre-specified outliers
No outliers found

Diagnostics

summary

Good

basic checks

definition: Good (0.000)
annual totals: Good (0.007)

visual spectral analysis

spectral seas peaks: Good
spectral td peaks: Good

regarima residuals

normality: Bad (0.002)
independence: Uncertain (0.043)
spectral td peaks: Good (0.244)
spectral seas peaks: Uncertain (0.082)

residual seasonality

on sa: Good (0.734)
on sa (last 3 years): Good (0.265)
on irregular: Good (0.874)

outliers

number of outliers: Good (0.020)

m-statistics

q: Good (0.856)
q without m2: Good (0.909)

La qualité globale de la désaisonnalisation de la série est acceptable bien que certains diagnostics sur la qualité des pré-ajustements aient des valeurs limites (Uncertain) voire soient rejetés (test de normalité).

Un commentaire sur la qualité globale de la désaisonnalisation est confectionné à partir des commentaires sur les différentes catégories de statistiques. Ce commentaire indique si oui ou non la désaisonnalisation est acceptable.

Seule la qualité globale de traitement de la série fait l'objet d'un commentaire synthétique. Aucune évaluation synthétique de la qualité de certains groupes de diagnostics ou de certaines parties de la désaisonnalisation n'est pratiquée.

Détails des seuils des différentes statistiques (valeurs par défaut)

		Undefined	Error	Severe	Bad	Uncertain	Good	
	Définition du modèle (basic checks)							
	Q		> 0.000001				<= 0.000001	
	Totaux annuels (basic checks)							
	Q		> 0.5	[0.1, 0.5]	[0.05, 0.1]	[0.01, 0.05]	<=0.01	
Pré-ajustements	Qualité des résidus	Test de normalité (regarima residuals)						
		p-value de la statistique de Doornik-Hansen				<0.01	[0.01, 0.1[>=0.1
		Test d'indépendance des résidus (regarima residuals)						
		p-value de la statistique de Ljung-Box(24)				<0.01	[0.01, 0.1[>=0.1
		Test de présence de fréquences saisonnières dans le spectre des résidus (regarima residuals)						
		p-value de la statistique de test			<0.001	[0.001, 0.01[[0.01, 0.1[>=0.1
	Test de présence de fréquences liées à des effets de calendriers dans le spectre des résidus (regarima residuals)							
	Outliers	Valeurs extrêmes (outliers)						
% d'outliers				>0.1	[0.05, 0.1[[0.03, 0.05[<=0.03	
Décomposition	Décomposition	Tests de spécifications (adéquation du modèle aux résultats obtenus) (Seats)						
		p-value du test sur la variance de la composante saisonnière				<0.01	[0.01, 0.05[>=0.05
		p-value du test sur la variance de l'irrégulier				<0.01	[0.01, 0.05[>=0.05
	p-value du test sur la corrélation entre la composante saisonnière et l'irrégulier				<0.01	[0.01, 0.05[>=0.05	
	Statistiques M sur la qualité de décomposition (m-statistics)							
	q			>2	[1, 2]			<=1
q sans m2			>2	[1, 2]			<=1	
Effets résiduels	Tests de saisonnalité stable résiduelle (Residual seasonality)							
	dans la série CVSCJO			<0.01	[0.01, 0.05[[0.05, 0.1[>=0.1	
	dans les trois dernières années de la série CVSCJO			<0.01	[0.01, 0.05[[0.05, 0.1[>=0.1	
	dans l'irrégulier			<0.01	[0.01, 0.05[[0.05, 0.1[>=0.1	
Test visuel sur le spectre des composantes (Visual spectral analysis)*								
présence de pics saisonniers				Irreg+sa	Irreg ou sa		aucun	
présence de pics liés à des effets cjo				Irreg+sa	Irreg ou sa		aucun	

*Il est possible d'imposer à ce test d'être **Severe** dès qu'un effet résiduel apparaît dans l'une des deux composantes.

Le seuil de détection des pics est également paramétrable ainsi que le nombre d'années utilisées afin de tracer le spectre des séries (par défaut, seules les 8 dernières années sont utilisées comme sous X-12-ARIMA).

Tous les seuils du tableau précédent peuvent être modifiés dans **Demetra+** (menu **options>Diagnostics**).

Il est également possible de retirer certains groupes de diagnostics de l'analyse (regroupements figurant de façon visuelle dans les sorties de **Demetra+**).

L'évaluation de la qualité est effectuée sous forme de labels. Une synthèse de la qualité à un niveau plus agrégé (qualité globale de désaisonnalisation de la série) nécessite l'établissement de règles d'agrégation de ces différents commentaires sur la qualité.

Critères d'agrégation de la qualité sous Demetra+

Diagnostic au niveau agrégé	État des diagnostics détaillés
Undefined	Tous les diagnostics sont Undefined
Error	Au moins un des diagnostics est en erreur (Error).
Severe	Au moins un des diagnostics est à Severe mais aucun n'est en erreur (Error).
Bad	Aucun diagnostic à Error ou Severe. La moyenne des diagnostics (hors Undefined) est <1.5 en utilisant la pondération : Bad=1, Uncertain=2, Good=3.
Uncertain	Aucun diagnostic à Error ou Severe. La moyenne des diagnostics (hors Undefined) est >=1.5 et <2.5 en utilisant la pondération : Bad=1, Uncertain=2, Good=3.
Good	Aucun diagnostic à error ou Severe. La moyenne des diagnostics (hors Undefined) est >=2.5 en utilisant la pondération : Bad=1, Uncertain=2, Good=3.

Remarques :

- Les diagnostics Severe ou Error sont absorbants : la simple présence d'un diagnostic présentant cette mention au niveau détaillé provoque la mise en Severe ou Error du diagnostic au niveau agrégé.. Le rejet de la série (qualité inférieure à Severe) est donc immédiat dès que certains diagnostics présentent une qualité insuffisante quels que soient les résultats obtenus par ailleurs.

D'après le tableau sur les niveaux de qualité possibles pour chacun des diagnostics, ces derniers n'ont pas tous le même niveau de gravité :

- Seuls les tests de respect des contraintes (basic checks) peuvent mettre en erreur la série. Dans ce cas, une des grandes équations de la désaisonnalisation liant les différentes composantes entre elles n'est pas vérifiée (par exemple, la série cvs-cjo est égale la composante saisonnière plus la composante irrégulière plus les effets des autres régresseurs sur la série brute (=régresseurs placés à part). Le lecteur intéressé par le détail des contraintes testées par Demetra+ pourra se référer au document de Necmettin Alpay KOÇAK cité en bibliographie). Cela se produit en particulier lorsque de grandes valeurs aberrantes sont présentes ou en cas d'échec de la procédure.
- Certains tests ne peuvent produire un diagnostic sévère (= Severe donc de diagnostic absorbant qui influe directement sur la statistique générale). Ces tests ne peuvent à eux seuls provoquer le rejet de la série (diagnostic <Severe).
- Certains diagnostics ne possèdent pas de valeurs limites. Dès qu'ils sont rejetés, la qualité est considérée comme mauvaise (= bad tests visuels sur les spectres, M-statistics).

Ces différences dans l'impact des statistiques reviennent à pondérer leur importance ou au moins à contrôler leurs impacts. La façon dont les différents diagnostics influent sur la qualité globale de la série n'est pas paramétrable dans le logiciel (= il n'est pas possible de définir quels sont les niveaux de qualité de chaque statistique (par exemple, un test sur les résidus du modèle ARIMA (indépendance ou normalité) ne sera jamais commenté Severe)).

3.1.1.1. Limites de l'évaluation de la qualité dans Demetra+

Avantages :

- L'édition de la qualité sous forme de label directement compréhensible par l'utilisateur permet de repérer très vite les séries à problèmes
- Il est possible d'accéder très rapidement à une liste très importante de diagnostics (notamment en cas de mauvais diagnostics dans le nœud « Main Result »)
- Les paramètres de désaisonnalisation d'une série peuvent être modifiés interactivement. La qualité résultant des modifications des spécifications est visible immédiatement dans le logiciel.

Inconvénients :

- La comparaison des diagnostics (ainsi que des composantes issues de la désaisonnalisation) est difficile au sein de l'application.
- Il est difficile de voir la qualité des différents groupes de diagnostics en un seul coup d'œil pour toutes les séries (pas de table de diagnostics détaillés pour l'ensemble des séries).
- Pas de tri possible des séries suivant des critères extérieurs (poids des séries dans l'indice par exemple) ou suivant un critère de qualité particulier.
- La pondération et les statistiques prises en compte dans l'évaluation de la statistique de qualité globale sont peu modifiables.

Un ensemble de macros SAS (**bilan qualité**) est donc utilisé afin de :

- pouvoir visualiser immédiatement la qualité des séries au niveau des différents groupes de diagnostics
- calculer des diagnostics de qualité propres au département
- prendre en compte des informations extérieures comme la pondération des séries dans l'indice
- comparer les diagnostics et pouvoir mettre côte à côte les spécifications qui ont conduit à ces diagnostics
- pondérer les diagnostics plus finement dans l'évaluation de la qualité globale.

3.1.2. Le résumé de la qualité dans le bilan qualité

La qualité est une note.

Les appréciations concernant la qualité dans le **bilan qualité** reposent sur des **notes (=valeurs numériques)**.

La qualité est évaluée au niveau de grands ensembles de diagnostics.

Comme dans **Demetra+**, les diagnostics sont regroupés par grands groupes.

Mais, dans le **bilan qualité**, à la différence de **Demetra+**, le regroupement n'est plus seulement visuel mais des notes sont calculées au niveau de chaque grand groupe de diagnostics.

Il est possible de connaître pour chaque série, la note obtenue selon les différents axes de qualité :

- Qualité des points aberrants
- Qualité des résidus du modèle ARIMA
- Qualité de décomposition
- Absence d'effets résiduels
- Qualité des révisions

Conséquences : il est possible de voir quel(s) groupe(s) de diagnostics pose(nt) problème pour une série. L'utilisation de notes (sommables) et de moyennes rend cohérente la qualité aux différents niveaux d'agrégation. Il est possible de pondérer l'impact des statistiques (une bonne statistique n'aura pas forcément toujours le même impact sur la qualité générale selon le diagnostic auquel elle renvoie).

Exemple : la normalité des résidus du modèle reg-ARIMA peut sembler moins importante que l'absence de corrélations au sein de ces derniers.

Le détail des différentes notes est consultable dans deux tables SAS :

- ➔ 1 table résumant la qualité de la désaisonnalisation sur la totalité de la série

Exemple de la table SAS ipi_qualite (série de l'IPi finissant en janvier 2011 désaisonnalisée avec une modélisation entièrement automatique)

Diagnostiques sur les points aberrants

Diagnostiques sur les résidus du modèle ARIMA

	name	note_aber_totale	note_aber_mois	note_aber_6mois	note_aber_an	note_tot_aber	note_totale_arima	box_pierce	ljung_box	normalite_asymetrie	ljung_box_var	box_pierce_var	normalite
1	RF061	98	100	100	86	100	78	100	50	100	100	100	100
2	RF062	80	79	100	75	50	89	100	100	100	100	100	50
3	RF081	100	100	100	100	100	67	100	100	0	100	100	50
4	RF089	82	50	100	93	100	100	100	100	100	100	100	100
5	RF101	100	100	100	100	100	78	100	100	100	50	50	100
6	RF103	72	69	100	43	50	44	100	100	0	0	0	0
7	RF104	67	100	100	0	100	100	0
8	RF105	100	100	100	100	100	56	50	50	100	100	100	100
9	RF106	44	100	100	0	0	0	50

Note : Les séries n'ayant pas de points aberrants affichent des valeurs manquantes pour les diagnostics les concernant.

→ 1 table sur la qualité des révisions.

C'est cet ensemble de deux tables qui sert à l'amélioration des qualités des modèles lors des campagnes annuelles. Le détail du calcul des notes est donné dans les points suivants.

3.1.2.1. Qualité des points aberrants

L'évaluation de la qualité des points aberrants repose dans le bilan qualité sur un ensemble de 4 notes :

- **note_aber_mois** : note évaluant la concentration des points aberrants sur les différents mois
- **note_aber_an** : note évaluant la concentration des points aberrants sur les différentes années
- **note_aber_6mois** : note évaluant la concentration des points aberrants sur la fin de la série
- **note_tot_aber** : note évaluant le nombre total de points aberrants

Le tableau ci-dessous donne le détail du calcul de ces différentes notes :

Statistiques	Conditions	Notes
Calcul des notes élémentaires au niveau de chaque mois (de janvier à Décembre)		
% de mois M ayant un point aberrant (années complètes = de 12 mois)	si > 20%	0
	si dans]10%, 20%]	50
	si <=10% ou 1 seul point aberrant	100
note_aber_mois : La note globale de concentration des points aberrants sur chaque mois est calculée comme la moyenne des notes de chaque mois possédant des points aberrants.		
Calcul des notes élémentaires au niveau de chaque année complète de la période de modélisation		
% de mois de l'année A ayant un point aberrant (années complètes = de 12 mois)	si > 25%	0
	si dans]10%, 25%]	50
	si <=10%	100
note_aber_an : La note globale de concentration des points aberrants sur chaque année est calculée comme la moyenne des notes de chaque année possédant des points aberrants.		

Statistiques	Conditions	Notes
Calcul de la note de concentration des points aberrants sur la fin de la série (note_aber_6mois)		
Nombre de points aberrants sur les 6 derniers mois	si ≥ 3	0
	si = 2	50
	si < 2	100
Calcul de la note sur le total de points aberrants (note_tot_aber)		
% de points aberrants (Nombre de points aberrants / nombre de points de la série)	si $> 10\%$	0
	si inclus dans]5%, 10%]	50
	si $\leq 5\%$	100

Guide de lecture : La note de la qualité des points aberrants survenant un mois de janvier est égale à :

- 0 si plus de 20% des mois de janvier des années complètes de la période de modélisation ont un point aberrant
- 50 si entre 10% exclus et 20 % inclus des mois de janvier des années complètes de la période de modélisation ont un point aberrant
- 100 si 10% ou moins des mois de janvier des années complètes de la période de modélisation sont modélisés par un point aberrant.

Les notes sur les concentrations des points aberrants sur les différents mois/années sont calculées comme des moyennes des notes obtenues sur chaque mois/année possédant des points aberrants.

La note globale de qualité des points aberrants est obtenue en effectuant une moyenne pondérée des notes du tableau précédent.

La pondération de la note synthétique (variable : note_aber_totale) est la suivante :

Variables	libellés	pondérations
note_aber_6mois	Note de qualité des points aberrants sur la fin de la série	2
note_aber_an	Note de qualité de la concentration des points aberrants sur les différentes années	1
note_aber_mois	Note de qualité de la concentration des points aberrants sur les différents mois (années complètes de la période)	2
note_tot_aber	Note sur le pourcentage de points aberrants	1
note_aber_totale	Note synthétique points aber.	Ens. = 6

Les pondérations les plus fortes concernent la qualité des points aberrants sur la fin de la série et la mesure de la concentration des points aberrants sur un type de mois.

Une trop forte concentration des points aberrants sur la fin de la série peut masquer une évolution importante (prise en compte d'une évolution de la tendance dans l'irrégulier par exemple ou dans le coefficient saisonnier (sous TRAMO-SEATS en particulier)) et être source de révisions importantes dans le futur. Par exemple, une évolution du coefficient saisonnier du mois de janvier peut d'abord être prise en compte comme une suite de valeurs extrêmes (donc modélisée par des indicatrices « Additive Outliers ») et pas comme une évolution de la saisonnalité. D'une façon générale, les outliers ou valeurs extrêmes apparaissant en fin de série doivent faire l'objet d'une analyse particulière. La pertinence et le type des variables détectées (en particulier sur le dernier point) sont à considérer attentivement (Par exemple, le changement de niveau détecté sur le dernier point est-il judicieux compte tenu du fait que l'on ne dispose que de peu d'information sur le futur de la série et risque-t-il de se maintenir lors de l'ajout de nouveaux points ?).

La répétition de points aberrants sur un même type de mois plusieurs années de suite peut faire craindre une prise en compte de la saisonnalité par des points aberrants.

3.1.2.2. Qualité des résidus du modèle reg-ARIMA

La notation repose sur 3 ensembles de notes :

- 2 notes sur l'absence d'autocorrélation des résidus : l'une s'appuyant sur la statistique de Ljung-Box(24) (**ljung_box**) et l'autre sur la statistique de Box-Pierce(24) (**Box-Pierce**)
- 2 notes sur l'absence d'autocorrélation du carré des résidus (test d'homoscédasticité) : l'une s'appuyant sur la statistique de Ljung-Box(24) (**ljung_box_var**) et l'autre sur la statistique de Box-Pierce(24) (**box_pierce_var**)
- 3 notes sur la normalité des résidus : statistiques du Kurtosis (aplatissement = **normalite_applatissement**), du Skewness (asymétrie = **normalite_asymetrie**) et statistique de Doornick-Hansen (**normalite**) résumant les deux précédentes.

En **bleu** figurent les noms des variables des tables SAS.

Détails du calcul des notes sur les résidus du modèle reg-ARIMA (bilan qualité)

Statistiques	Conditions	Notes
Tests du bruit blanc (absence de lien linéaire dans les résidus du modèle reg-ARIMA)		
Calcul de la note sur la statistique de Ljung-Box sur les résidus (ljung_box)		
p-value de la statistique du Ljung-Box(24) (1)	si <=1% si inclus dans]1%, 5%] si >5%	0 50 (compté comme 0 dans la note récapitulative)* 100
Calcul de la note sur la statistique de Box-Pierce(24) sur les résidus (box_pierce)		
p-value de la statistique de Box-Pierce(24)	si <=1% si inclus dans]1%, 5%] si >5%	0 50 (compté comme 0 dans la note récapitulative)* 100
Tests d'homoscédasticité des résidus (absence de lien dans le carré des résidus)		
Calcul de la note sur la statistique du Ljung-Box(24) sur le carré des résidus (ljung_box_var)		
p-value de la statistique du Ljung-Box(24) sur le carré des résidus	si <=1% si inclus dans]1%, 5%] si >5%	0 50 (compté comme 0 dans la note récapitulative) 100
Calcul de la note sur la statistique de Box-Pierce sur le carré des résidus (box_pierce_var)		
p-value de la statistique de Box-Pierce(24) sur le carré des résidus	si <=1% si inclus dans]1%, 5%] si >5%	0 50 (compté comme 0 dans la note récapitulative)* 100

Statistiques	Conditions	Notes
Tests de normalité des résidus		
Calcul de la note sur l'aplatissement de la distribution (normalite_aplatissement)		
p-value de la statistique du Kurtosis (aplatissement)	si <=5%	0
	si >5%	100
Calcul de la note sur l'asymétrie (normalite_asymetrie)		
p-value de la statistique du Skewness (asymétrie)	si <=5%	0
	si >5%	100
Calcul de la note de normalité globale (normalite)		
p-value de la statistique de normalité globale (Doornick-Hansen)	si <=1%	0
	si inclus dans]1%, 5%]	50 (compté comme 0 dans la note récapitulative)
	si >5%	100

Guide lecture (1) :

La note sur la statistique de Ljung-Box(24) sur les résidus est égale à :

- 0 si la p-value est <=1%
- 50 si la p-value est compris entre 1% exclu et 5% inclus
- 100 si la p-value est >5%

* Les notes égales à 50 (qualité limite) apparaissent dans la table récapitulative mais sont comptabilisées à 0 dans la note globale. Il a en effet été choisi de conserver les notes calculées dans l'ancienne version du bilan qualité et ainsi d'attirer l'attention sur les séries dès qu'un test était rejeté à 5% mais de faire apparaître tout de même les seuils pris en compte dans l'évaluation de la qualité sous Demetra v2.2 (mise en bleu ou en rouge des statistiques selon leur rejet à 5% ou à 1% et influence différente sur le critère de qualité global).

Le détail des pondérations est le suivant :

Variabes	libellés	pondérations
box_pierce	Note liée au test de Box-Pierce sur les résidus	2
box_pierce_var	Note liée au test de Box-Pierce sur le carré des résidus	1
ljung_box	Note liée au test de Ljung-Box sur les résidus	2
ljung_box_var	Note liée au test de Ljung-Box sur le carré des résidus	1
normalite	Note sur le test de normalité (Doornick-Hansen)	1
normalite_aplatissement	Note sur le test du Kurtosis	0
normalite_asymetrie	Note sur le test du Skewness	2
note_total_arima	Note globale sur la qualité des résidus du modèle ARIMA	ens. = 9

L'absence de liens linéaires entre les résidus est considérée comme plus importante que leur éventuelle hétéroscédasticité.

Pour la normalité, l'asymétrie est mise en avant. De forts résidus peuvent déstabiliser la modélisation et être liés à un oubli de modélisation de valeurs atypiques. Ils peuvent également la conséquence d'une mauvaise spécification du modèle ARIMA (présence de résidus présentant une structure, non homogènes sur toute la période, etc.). Les séries présentant une asymétrie doivent donc être repérées.

3.1.2.3. Qualité de décomposition

Aucune statistique d'adéquation des modèles des composantes utilisés par SEATS aux résultats obtenus n'est reprise dans le bilan qualité (tests de comparaison des variances théoriques des différents filtres avec les variances empiriques des composantes). Ceci est lié au fait que le DCST utilise la méthode X-12-ARIMA qui ne repose pas sur une modélisation a-priori des composantes (méthode non paramétrique).

Les statistiques M1 à M11 calculés dans Demetra+ sont récupérées afin de servir au calcul des notes dans le bilan qualité. La qualité est considérée comme bonne si la statistique récapitulative Q est inférieure à 1.

Demetra+ ne donne malheureusement pas accès à ces statistiques pour des séries désaisonnalisées avec TRAMO-SEATS. Une analyse empirique de la variance est donc pratiquée dans ce cas. La qualité est considérée comme bonne si le poids de l'irrégulier dans la variance totale n'est pas trop important et le poids de la saisonnalité évolutive est strictement inférieur à la moitié du poids de la saisonnalité stable (cf. tableau ci-dessous pour plus de précisions).

Détails du calcul des notes pour la qualité de décomposition

Statistiques	Conditions	Notes
Qualité de la décomposition (X11)		
Calcul de la note de la qualité de décomposition (à partir des statistiques M1 à M11) : (qualite_decomposition)		
Statistique Q ⁵	si ≥ 1 si < 1	0 100
Qualité de la décomposition (TRAMO-SEATS)		
Calcul de la note de la qualité de décomposition (à partir d'une analyse empirique de la variance) : (qualite_decomposition)		
Le calcul de la note repose sur l'analyse des covariances avec la somme de l'irrégulier + la composante saisonnière (Irr+S).		
Calcul du critère sur le poids de l'irrégulier		
Le pourcentage de la covariance de l'irrégulier avec la somme des composantes irrégulière et saisonnière n'excède pas 80% de la variance de la somme des composantes saisonnière et irrégulière.	si $cov(Irr, Irr+S) \geq 80\% V(S+Irr)$ si $cov(Irr, Irr+S) < 80\% V(S+Irr)$	Qualité mauvaise Qualité ok
Calcul du critère sur le poids de la saisonnalité évolutive par rapport à la saisonnalité stable		
Le pourcentage de la covariance de la saisonnalité évolutive avec la somme des composantes irrégulière et saisonnière est inférieur de moitié à celle de la saisonnalité stable avec la somme des composantes irrégulière et saisonnière.	si $cov(eseas, Irr+S) \geq 1/2 cov(mseas, Irr+S)$ si $cov(eseas, Irr+S) < 1/2 cov(mseas, Irr+S)$	Qualité mauvaise Qualité ok
La note de qualité de composition basée sur l'analyse de la variance est égale à 100 si les deux critères sur le poids de l'irrégulier et le poids de la saisonnalité évolutive sont considérés comme bons.		

Guide de lecture : la qualité de décomposition d'une série désaisonnalisée avec X-12-ARIMA sera considérée comme bonne si sa statistique Q est inférieure strictement à 1.

⁵ Une attention particulière pourra également être portée à la statistique M7 (présente dans les tables SAS du bilan qualité) qui repose sur le test de saisonnalité identifiable de X-12. En cas de doute, la consultation des tests de saisonnalité de la série brute dans Demetra+ pourra se révéler fructueuse (nœud Diagnostics>Seasonality tests de l'arbre de diagnostics détaillés).

3.1.2.4. Absence d'effets résiduels

Un ensemble de 4 notes indiquant la présence ou l'absence d'effets résiduels (de calendriers et saisonniers) sur les séries de l'irrégulier et CVS-CJO figurent dans le [bilan qualité](#).

Ces notes ne font l'objet d'aucune synthèse dans la table SAS résumant la qualité de toutes les séries.

Statistiques	Conditions	Notes
Effets cjo résiduels		
Test de nullité simultanée des coefficients des régresseurs cjo (7 régresseurs)		
Calcul de la note sur les effets cjo résiduels sur la composante irrégulière (cjo_res_dos_irreg)		
p-value de la statistique de Fisher	si <= 5%	0
	si > 5%	100
Calcul de la note sur les effets cjo résiduels sur la série CVS-CJO (cjo_res_dos_cvscjo)		
p-value de la statistique de Fisher	si <= 5%	0
	si > 5%	100
Effets CVS résiduels		
Test de saisonnalité identifiable		
Calcul de la note sur les effets CVS résiduels sur la composante irrégulière (note_cvsvsur_irreg)		
Diagnostics du test de présence de saisonnalité identifiable de X12.	si non NOT PRESENT (présence d'effets saisonniers)	0
	si NOT PRESENT	100
Calcul de la note sur les effets CVS résiduels sur la série CVS-CJO (note_cvsvsur_cvscjo)		
Diagnostics du test de présence de saisonnalité identifiable de X12.	si non NOT PRESENT (présence d'effets saisonniers)	0
	si NOT PRESENT	100

3.1.2.5. Récapitulatif des diagnostics pris en compte dans le [bilan qualité](#)

En résumé, les différents critères de qualité (hors critères concernant les révisions) résumant la qualité de l'ajustement saisonnier d'une série présents dans le [bilan qualité](#) sont les suivants :

	Variables de la table SAS résumant les qualités des séries (table suffixée par _qualite)	Critères de qualité
Qualité des pré-ajustements	Box_pierce	Note d'indépendance des résidus basée sur la statistique de Box-Pierce(24)
	Box_pierce_var	Note d'indépendance du carré des résidus (Box-pierce). Test d'homoscédasticité.
	Ljung_box	Note d'indépendance des résidus basée sur la statistique de Ljung-Box(24)
	Ljung_box_var	Note d'indépendance du carré des résidus (Ljung-Box). Test d'homoscédasticité.
	Normalite	Note de normalité globale (test de Doornick-Hansen)
	Normalite_aplatissement	Note d'aplatissement de la distribution des résidus (Kurtosis)
	Normalite_asymetrie	Note d'asymétrie de la distribution des résidus (Skewness)
	Note_total_arima	Note évaluant la qualité globale du modèle ARIMA

	Qualité des points aberrants	Note_aber_6mois	Note sur la concentration des points aberrants sur les 6 derniers mois de la période de modélisation
		Note_aber_an	Note sur la concentration des points aberrants sur les différentes années
		Note_aber_mois	Note sur la concentration des points aberrants sur les différents mois
		Note_tot_aber	Note sur le pourcentage global de points aberrants
		Note_aber_totale	Note synthétique sur la qualité de la modélisation des valeurs atypiques
Qualité de la décomposition	Décomposition	qualite_ajustement	Note sur la qualité de décomposition (note reposant sur la statistique Q avec X-12 et sur une décomposition de la variance avec la méthode TRAMO-SEATS)
	Effets résiduels	note_cvs_sur_cvscjo	Note sur la présence d'effets CVS résiduels sur la série CVSCJO
		note_cvs_sur_irreg	Note sur la présence d'effets CVS résiduels sur la composante irrégulière
		cjo_res_dos_cvscjo	Note sur la présence d'effets CJO résiduels sur la série CVSCJO.
		cjo_res_dos_irreg	Note sur la présence d'effets CJO résiduels sur la composante irrégulière.

3.1.2.6. Révisions

Il n'existe pas de critère bien établi permettant de déterminer ce qu'est une série fortement révisée. Les révisions sont fortement liées aux caractéristiques de la série initiale :

- Certaines séries instables par essence seront plus fortement révisées que d'autres. Ainsi les amplitudes moyennes de révisions varient selon les secteurs d'activité.
- La série désaisonnalisée met parfois du temps à prendre en compte les évolutions de la courbe initiale. Les révisions seront plus fortes en cas d'évolutions brutales de la série initiale.
- Les points aberrants détectés peuvent affecter les révisions (évaluation des coefficients saisonniers différente).

Il n'existe donc pas de critère au sein du bilan qualité indiquant si les révisions sont trop ou pas assez importantes en niveau. Les différentes statistiques sur les révisions sont plutôt utilisées en comparaison, en mettant côte-à-côte la stabilité des résultats obtenus avec deux jeux de spécifications différents.

3.2. Comparaison des diagnostics présents dans le résumé de Demetra+ et le bilan qualité

3.2.1. Qualité des pré-ajustements

Demetra+ et le **bilan qualité** présentent dans leurs bilans des diagnostics sur la qualité des résidus du modèle reg-ARIMA (absence de structure linéaire, normalité) et sur le nombre de valeurs modélisées par des outliers (valeurs extrêmes).

Diagnostics présents dans [Demetra+](#) et pas dans le [bilan qualité](#)

- [Demetra+](#) effectue des tests d'effets résiduels (saisonnalité, effets CJO) sur le spectre des résidus du modèle reg-ARIMA. Dans le [bilan qualité](#), les tests d'effets résiduels interviennent seulement sur les composantes à l'issue de la désaisonnalisation. De façon générale, aucun test n'est pratiqué sur des spectres dans le [bilan qualité](#).

Diagnostics présents dans le [bilan qualité](#) et pas dans [Demetra+](#)

- Pas de tests sur l'hétéroscédasticité des résidus dans le résumé de Demetra+ (indépendance du carré des résidus)
- Pas de tests sur la concentration des points aberrants dans Demetra+
- Seule la statistique résumant le diagnostic sur la normalité des résidus figure dans l'onglet **Main Results** de [Demetra+](#).

3.2.2. Qualité de la décomposition

[Demetra+](#) et le [bilan qualité](#) accordent de l'importance aux diagnostics jugeant de la présence d'effets résiduels (saisonniers et liés à des effets de calendriers) sur l'irrégulier et la série CVS-CJO.

Diagnostics d'effets résiduels présents dans [Demetra+](#) et pas dans le [bilan qualité](#)

- [Demetra+](#) offre des tests visuels d'effets résiduels sur les spectres de l'irrégulier et de la série CVS-CJO.
- Un test de saisonnalité stable est pratiqué sur les trois dernières années de la série CVS-CJO.

Diagnostics d'effets résiduels présents dans le [Bilan Qualité](#) et pas dans [Demetra+](#)

- Un test de Fisher de nullité simultanée des coefficients du jeu de 7 régresseurs est pratiqué sur l'irrégulier et la série CVSCJO. Sous [Demetra+](#), aucun véritable test d'effets CJO ne figure mais seulement des diagnostics spectraux.

Lors de l'utilisation de la méthode [X-12-ARIMA](#), la qualité de décomposition est mesurée par les statistiques M dans le [bilan qualité](#) et sous [Demetra+](#). En revanche, lors de l'utilisation de [TRAMO-SEATS](#), l'analyse du [bilan qualité](#) repose sur une décomposition empirique de la variance. [Demetra+](#) ne teste pas vraiment la validité de la décomposition opérée par [SEATS](#) mais plutôt l'adéquation des estimateurs aux résultats empiriques.

3.2.3. Qualité des révisions

Aucun diagnostic sur la qualité des révisions ne figure dans le résumé de la qualité de [Demetra+](#). Dans le [bilan qualité](#), une indication sur l'amplitude des révisions (de type History) figure. Il est également possible de consulter les points aberrants instables ou la qualité de l'ajustement du modèle ARIMA sur les séries tronquées successives.

3.3. Utilisations pratiques des résumés de la qualité de Demetra+ et du bilan qualité

Lors des campagnes annuelles, les résumés de la qualité des traitements et les diagnostics fournis par [Demetra+](#) et le [bilan qualité](#) sont utilisés conjointement. Ce paragraphe résume dans quelles circonstances. Les différents points évoqués seront détaillés dans les chapitres suivants.

3.3.1. Analyse de la qualité de chaque série

Les résumés de la qualité des différentes séries sous **Demetra+** et le **bilan qualité** reposent sur des diagnostics différents. Leurs conclusions peuvent donc diverger.

Aucun résumé global de la qualité des séries n'est effectué dans le **bilan qualité**. C'est parce qu'en réalité, le diagnostic de **Demetra+** suffit à repérer les séries les plus problématiques. Les séries présentant un diagnostic synthétique de qualité insatisfaisant sont ensuite étudiées plus en détail dans ce même logiciel.

Ce sont plutôt les notes obtenues pour les différents groupes de diagnostics qui sont étudiées dans le **bilan qualité**. L'analyse porte en outre plus sur un niveau de qualité comparativement à un autre que sur un niveau absolu (cf. chapitre 4 sur les comparaisons de qualité).

3.3.2. Tris suivant un critère de qualité

L'intérêt du **bilan qualité** réside dans la possibilité qu'il offre de **trier les séries** suivant un axe de qualité précis (par exemple, les séries possédant des effets cjo résiduels sur l'une des composantes testées). Il est alors possible de modifier automatiquement les paramètres de désaisonnalisation des séries présentant des manques évidents (par exemple pas de régresseurs cjo et des effets résiduels) dans **Demetra+**. Sous **Demetra+**, un tri des traitements suivant un critère de qualité précis ne peut être réalisé au sein de l'application et nécessite un export des diagnostics ce qui revient à créer un bilan qualité sommaire. De plus, les modifications éventuelles de spécifications doivent être entrées à la main dans le logiciel et ne peuvent faire de façon simple l'objet d'une automatisation.

3.3.3. Comparaison de qualité

Un autre intérêt est la possibilité **de comparer des niveaux de qualité**. Les comparaisons sont malaisées sous **Demetra+** (pas de possibilité de juxtaposer tous les diagnostics et de les faire défiler simplement pour toutes les séries).

Or, il est parfois plus intéressant de savoir si la qualité se dégrade ou s'améliore que de déterminer un niveau absolu de qualité. Certaines séries sont si chaotiques qu'il est difficile voire impossible de trouver des paramètres convenables dans les délais impartis. Le but sera alors de s'assurer que la qualité de la désaisonnalisation ne se dégrade pas en modifiant le modèle.

Il est également intéressant de pouvoir comparer plusieurs variantes de spécifications suivant les différents groupes de diagnostics afin de faciliter le choix de modèle.

Des fichiers Excel peuvent être produits automatiquement afin de comparer les qualités engendrées par différentes versions de spécifications. Ces fichiers mettent en parallèle les notes obtenues suivant les différents groupes de diagnostics dans le **bilan qualité** et les différents paramètres ayant permis d'obtenir ces résultats.

3.3.4. Évaluation de la qualité d'un indice désaisonnalisé de façon indirecte

L'évaluation de la qualité dans le **bilan qualité** est effectuée sous forme de notes c'est-à-dire de chiffres. L'intérêt d'une évaluation de la qualité sous forme numérique réside dans la possibilité de les agréger non seulement pour donner une évaluation de la qualité globale d'une série (résumé du résultat obtenu pour les différents grands groupes de diagnostics) mais aussi afin **d'évaluer la qualité des agrégats dans le cas d'une désaisonnalisation indirecte** (résumé de résultats obtenus par les séries constitutives de l'agrégat).

En effet, souvent, les principaux agrégats de l'indice font également l'objet de publications. Il est donc important de contrôler leur qualité. Il ne suffit donc pas d'améliorer la qualité des séries suivant leur

pondération dans l'indice. Il faut également étudier la façon dont elles se combinent afin de former les principaux agrégats.

En cas de résultat insatisfaisant au niveau d'un agrégat, il est possible de connaître les séries les plus contributrices à la détérioration de la qualité (analyse descendante de la qualité). Dans le cas d'une analyse des modèles dans un temps très court (campagne annuelle), cette méthode permet de repérer les séries les plus problématiques, celle dont l'amélioration aura le plus d'impact sur la qualité globale de l'indice et de ses principaux agrégats.

3.4. Organisation du traitement des CVS au DSCT (Département des statistiques de Court Terme)

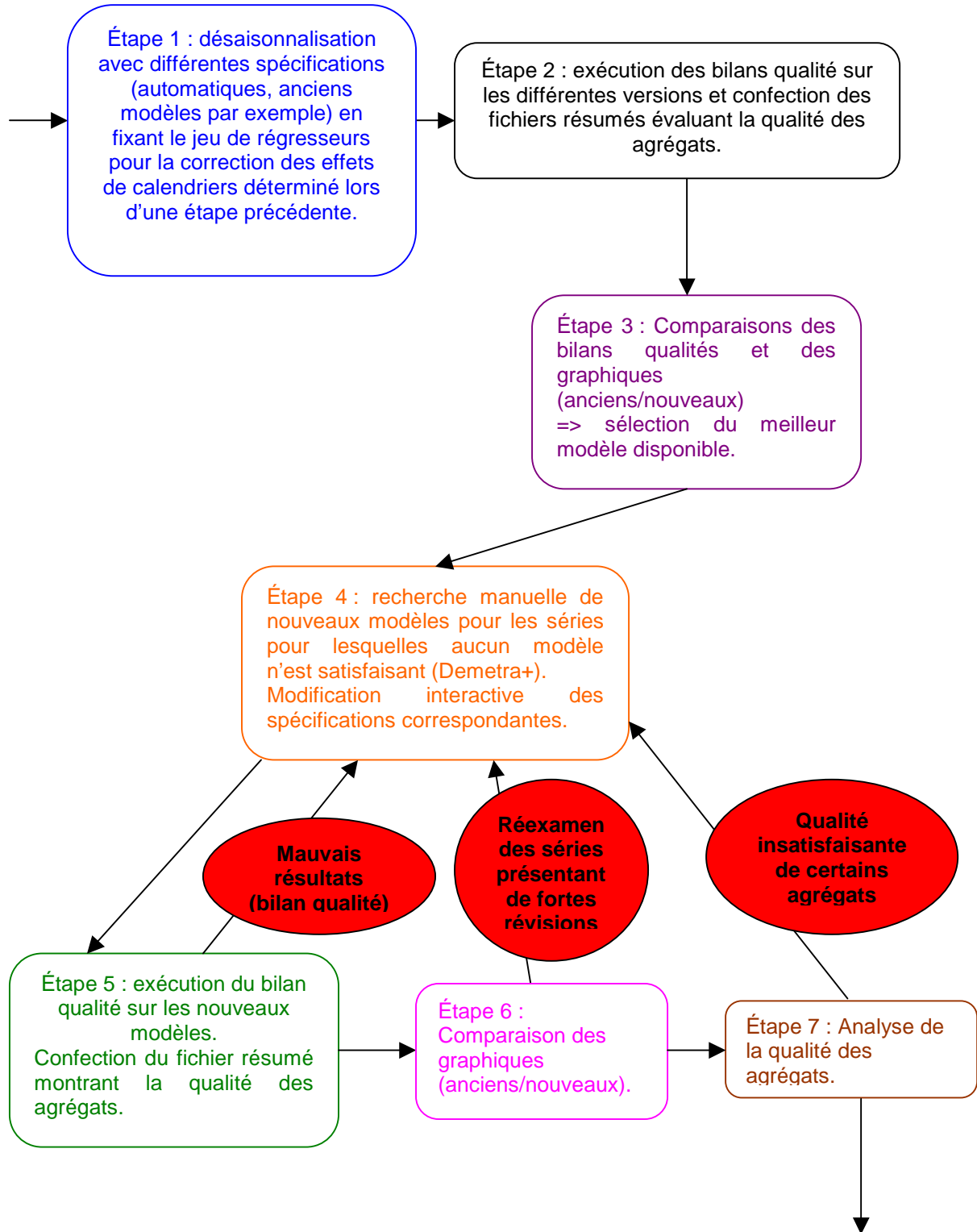
Au sein du département des Statistiques de Court Terme, ce sont 531 séries environ qui sont désaisonnalisées tous les mois.

L'ajustement est de type **concurrent partiel** c'est-à-dire que tous les modèles sont fixés une fois par an (lors de campagnes annuelles de désaisonnalisation), puis seuls les paramètres des modèles sont ré-estimés chaque mois. Les points atypiques (outliers) trouvés lors de la campagne annuelle de désaisonnalisation sont fixés s'ils sont antérieurs à une date choisie par le responsable d'indice (en général, la date en-deça de laquelle les données ne peuvent plus être révisées). Les outliers sont identifiés automatiquement par la méthode X-12-ARIMA lors de chaque campagne mensuelle pour la période au-delà de cette date. Il est tout de même possible de fixer certains outliers économiquement significatifs.

Les indices sont désaisonnalisés de façon indirecte c'est-à-dire que les séries désaisonnalisés des agrégats sont calculées comme des sommes de séries désaisonnalisés de niveau inférieur. Le niveau d'agrégation où se pratique la désaisonnalisation est le niveau NACE3 pour l'IPI (Indice de Production Industrielle), NACE3 pour les séries de Chiffres d'Affaires de l'Industrie et NACE5 pour les autres secteurs.

3.4.1. Schéma d'une campagne annuelle de désaisonnalisation du point de vue de la qualité.

Lors d'une campagne annuelle de désaisonnalisation, toutes les séries font l'objet d'une nouvelle recherche de modèle. Le **bilan qualité** est dans ce cas utilisé afin de faciliter la recherche de modèle dans un temps pouvant être très court.



Lors de l'étape 1, différentes possibilités en termes de spécifications sont testées. En particulier, une désaisonnalisation automatique de toutes les séries avec une nouvelle recherche des points aberrants sur toute la période est pratiquée.

Une désaisonnalisation avec TRAMO-SEATS peut également être envisagée afin de bénéficier de sa procédure de sélection de modèle (les modèles seront tout de même utilisés au final avec X-12-ARIMA).

Remarque : les régresseurs CJO utilisés avec chaque série sont fixés lors d'une étape préalable non détaillée dans le schéma précédent.

Étape 2 : Un bilan de la qualité est édité pour chacune des versions confectionnées lors de l'étape 1.

Pour chaque jeu de spécifications, un fichier résumant la qualité pour les différents niveaux d'agrégation est calculé.

Étape 3 : Les qualités des désaisonnalisations effectuées avec les différentes spécifications suivant les différentes catégories de diagnostics sont comparées afin de déterminer le meilleur modèle pour chaque série (macros du bilan qualité).

Comparaisons des bilans : ce sont les notes au niveau des groupes de diagnostics qui sont comparées.

- Des modifications automatiques de modèles peuvent être effectuées (mise en place de 7 régresseurs si la série présente des effets résiduels, retour à l'ancien modèle ARIMA si ce dernier est strictement meilleur etc.).
- Les éléments pouvant être mobilisés sont : la comparaison des qualités (en particulier par rapport à l'ancien), la révision par rapport à la série publiée, les commentaires de qualité sous Demetra+.

Étape 4 : les séries problématiques font l'objet d'une recherche manuelle de modèle :

- séries dont le diagnostic général est Severe sous Demetra+
- séries dont la qualité se dégrade
- séries présentant une pondération importante dans l'indice
- séries présentant des révisions importantes par rapport à la série publiée

Cette recherche repose sur les diagnostics de Demetra+. Une étude en détail de tous les diagnostics disponibles est réalisée pour ces séries problématiques. Ce logiciel offre la possibilité de modifier les spécifications de la série et de visualiser immédiatement le résultat ce qui facilite la recherche de modèle.

Étape 5 : Un résumé final de la qualité est élaboré. La qualité des agrégats est évaluée. Il est possible de pointer les séries contribuant le plus à la dégradation de la qualité (analyse descendante).

Les séries pour lesquelles la qualité se dégrade ou présentant des notes de bilan insuffisantes peuvent faire l'objet d'une nouvelle analyse dans Demetra+. Dans ce cas, un retour vers l'étape 4 est effectué. Plusieurs va-et-vient entre les étapes 4 et 5 peuvent se produire.

Étape 6 : Les graphiques superposant les séries publiées et les nouvelles séries sont réalisés sous SAS ce qui permet d'appréhender les révisions produites par les nouveaux modèles (et éventuellement les nouvelles données).

En cas de fortes révisions, un réexamen de la série est possible (et donc un retour à l'étape 4).

Étape 7 : Une comparaison du fichier résumant la qualité des séries publiées (modèles avant la campagne annuelle avec des données non révisées) avec celui résumant les notes de qualité obtenues avec les nouveaux paramètres permet d'appréhender l'évolution de la qualité introduite par la campagne annuelle de désaisonnalisation. Une analyse descendante permet d'identifier les séries contribuant le plus à la dégradation de qualité des agrégats et donc de procéder à une nouvelle analyse de ces séries sous Demetra+ (donc un retour à l'étape 4).

4. Comparaison de qualités et de résultats

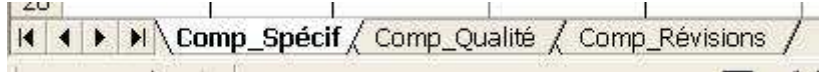
4.1. Comparaison des résultats de deux bilans qualité

Des programmes ont été écrits afin de **comparer les résultats en terme de qualité** obtenus grâce à deux jeux de spécifications.

Ces programmes créent à partir de l'exécution de deux **bilans qualité** sur deux ensembles de traitements Demetra+ (multi-processing) contenant des séries identiques désaisonnalisés avec des jeux de spécifications différents un fichier Excel comparant les notes de qualité obtenues et les diagnostics synthétiques de qualité des séries de Demetra+.

4.1.1. Organisation du fichier Excel de comparaison des diagnostics

Ce fichier contient 3 onglets :



→ **Comp_Spécif** : contient les comparaisons des spécifications de chacune des séries pour les deux jeux de spécifications.

Dans cet onglet figure les paramètres utilisés afin de désaisonnaliser la série au final c'est-à-dire à l'issue d'une éventuelle phase d'identification). Si pour une série une procédure d'identification automatique du modèle reg-ARIMA est requise et que le logiciel sélectionne un modèle Airline, le modèle Airline figurera dans le fichier Excel et pas le paramétrage « identification automatique des paramètres ».

Les colonnes à **en-têtes verts** sont des indicatrices indiquant le caractère identique ou non des paramètres. Il est ainsi possible de visualiser rapidement les éléments différents des deux jeux de spécifications à l'origine des différences de qualité observés dans les autres onglets.

Les colonnes suivantes (avec des **en-têtes bleus**) permettent de visualiser les paramètres des deux jeux de spécifications pour chacune des séries.

Les paramètres comparés et figurant dans l'onglet sont :

- le modèle ARIMA
- la transformation
- le nombre de régresseurs
- le nombre d'outliers
- la présence de moyenne dans le modèle ARIMA

Ces paramètres sont ceux de la modélisation reg-ARIMA. Ils sont communs aux deux méthodes de désaisonnalisations utilisables dans **Demetra+**.

Résumé des différences dans les spécifications.

Colonnes de visualisation des différents paramètres

	A	B	C	Résumé				Modèle			Transformation		
	Série	Pond	Tendance	Modèles ARIMA égaux	Type transformation égaux (Add/Mult)	Régress-seurs égaux	Nombre outliers égaux	Nouveau	Ancien	Nouveau	Ancien		
3	RF061	247	0	0	1	0	1	(1 1 0)(0 1 1)	(0 1 1)(0 1 1)	Logarithm	Logarithm	L	
4	RF062	144	0	0	1	0	0	(1 1 1)(0 1 1)	(0 1 1)(0 1 1)	None	None	N	
5	RF081	1286	0	0	1	1	0	(0 1 1)(0 1 1)	(2 1 1)(0 1 1)	None	None	7	
6	RF089	95	0	0	1	0	0	(1 0 1)(0 1 1)	(0 1 1)(0 1 1)	Logarithm	Logarithm	7	
7	RF101	4418	0	1	1	1	1	(0 1 1)(0 1 1)	(0 1 1)(0 1 1)	None	None	7	

Pondérations des séries

Cette colonne est une indicatrice indiquant si la série est suivie en tendance ou non (cas de séries très volatiles)

Exemple : comparaison de la qualité obtenue à l'aide des anciens modèles et des modèles automatiques pour les séries de l'IPI lors de la campagne 2011.

→ **Comp_qualité** : Les évolutions de qualité (différences de notes) sont évaluées pour tous les groupes de diagnostics du **bilan qualité** (voir évaluation de la qualité pour une série).

Les premières colonnes de cet onglet (avec des **en-têtes verts**) montrent les différences de notes : La comparaison concerne :

	A	B	C	Différences Qualité Nouveau/Ancien						Notes Q		
	Série	Pond	Tendance	Diff Note ARIMA	Diff Note Outliers	Diff Note CJO res	Diff Note CVS res	Diff Note Qualité Décomp	Diff Qualité Dem+	Note ARIMA	Note Outliers	Note C res
3	RF061	247	0	-22.2	0	0	0	0	0	78	98	
4	RF062	144	0	11.1	-1.2	0	0	100	0	89	80	
5	RF081	1286	0	-33.3	0	0	0	0	0	67	100	
6	RF089	95	0	22.2	-7.4	100	0	0	0	100	82	
7	RF101	4418	0	0	0	0	0	-100	0	78	100	

- les notes ayant trait aux qualités des modèles : **note_total_arima** (qualité des résidus du modèle ARIMA) et **note_aber_totale** (qualité des points aberrants).
- les notes sur les effets résiduels détectés dans la **série CVSCJO** : effets CJO résiduels (**cjo_res_dos_cvscjo**) et CVS résiduels (**note_cvs_sur_cvscjo**).
- les notes sur la qualité de décomposition : **qualite_decomposition**

La dernière colonne compare les diagnostics résumant la qualité sous **Demetra+**. Cette colonne est égale à **1** lorsque la qualité s'améliore (par exemple, le diagnostic de qualité globale passe de **Severe** à **Bad**) et à **-1** lorsque la qualité se dégrade (par exemple, passe de **Good** à **Uncertain**).

Il est ensuite possible de consulter le détail des notes obtenues avec les deux jeux de spécifications (**bilan qualité**) et des diagnostics de qualité de **Demetra+** dans les colonnes suivantes (avec des **entêtes bleues**).

	A	J	K	L	M	N	O	P	Q	R	S	T	U
1		Notes Qualités Nouveau						Notes Qualités Ancien					
2	Série	Note ARIMA	Note Outliers	Note CJO res	Note CVS res	Note Qualité Décomp	Note Qualité Dem+	Note ARIMA	Note Outliers	Note CJO res	Note CVS res	Note Qualité Décomp	Note Qualité Dem+
3	RF061	78	98	100	100	100	Good	100	98	100	100	100	Good
4	RF062	89	80	100	100	100	Good	78	82	100	100	0	Good
5	RF081	67	100	100	100	100	Good	100	100	100	100	100	Good
6	RF089	100	82	100	100	100	Severe	78	90	0	100	100	Severe
7	RF101	78	100	100	100	0	Severe	78	100	100	100	100	Severe

Ci-dessus : la différence des notes de qualité des modèles ARIMA de -22.2 constaté dans les colonnes précédentes pour la série RF061 s'explique par une dégradation de la note de 100 à 78. Dans l'onglet *comp_spécif*, on s'aperçoit que le modèle $(1\ 1\ 0)(0\ 1\ 1)$ des anciens spécifications a fait place à un modèle Airline. Les régresseurs ont également été modifiés (passage de 2 régresseurs à un effet année bissextile seul dans le modèle identifié automatiquement). Aucun effet CJO résiduel n'est cependant décelable dans la série CVS-CJO. Au final, le modèle automatique a été choisi (peu de modification des résultats finaux et suppression d'une variable non significative dans le jeu de régresseurs CJO utilisé).

- ➔ **Comp_revisions**: Les statistiques concernant les révisions à prévoir avec les deux modèles figurent dans cet onglet. Aucune note ne figure dans le **bilan qualité**. Dans la feuille Excel ne figure donc que la juxtaposition des colonnes avec les différentes mesures des révisions.

Les grandeurs figurant dans le fichier sont :

- ➔ MEANSQUAREDREV : moyenne des carrés des révisions
- ➔ MEANREV : révision moyenne
- ➔ MEANABSREV : moyenne des révisions absolues
- ➔ MINREV : minimum des révisions
- ➔ MAXREV : maximum des révisions

Les révisions sont calculées 1 mois, 2 mois, 3 mois et 1 année après la première publication selon la méthode de calcul de l'OCDE..

Caractéristiques des révisions estimées 1 mois après la première publication

	A	B	C	modèles nouveaux - 1 mois				modèles anciens - 1 mois				N		
	Série	Pond	Tendance	MEAN-SQUARED-REV	MEANREV	MEANABS-REV	MINREV	MAXREV	MEAN-SQUARED-REV	MEANREV	MEANABS-REV	MINREV	MAXREV	MEAN-SQUARED-REV
3	RF061	247	0	0.2	0	0.2	-1.4	1.2	0	0	0	-0.4	0.1	0
4	RF062	144	0	7.5	0.6	1	-1.6	12.4	0	0	0.2	-0.7	0.3	7.5
5	RF081	1286	0	0.2	-0.1	0.2	-1.7	0.5	0	0	0.1	-0.3	0.8	0
6	RF089	95	0	0.5	0	0.5	-1.1	2.4	1.3	-0.2	0.3	-5.5	0.3	0
7	RF101	4418	0	0	0	0	-0.4	0.3	0	0	0	0	0	0

Les valeurs des moyennes des carrés des révisions (MEANSQUAREDREV) supérieures à 5 figurent en rouge.

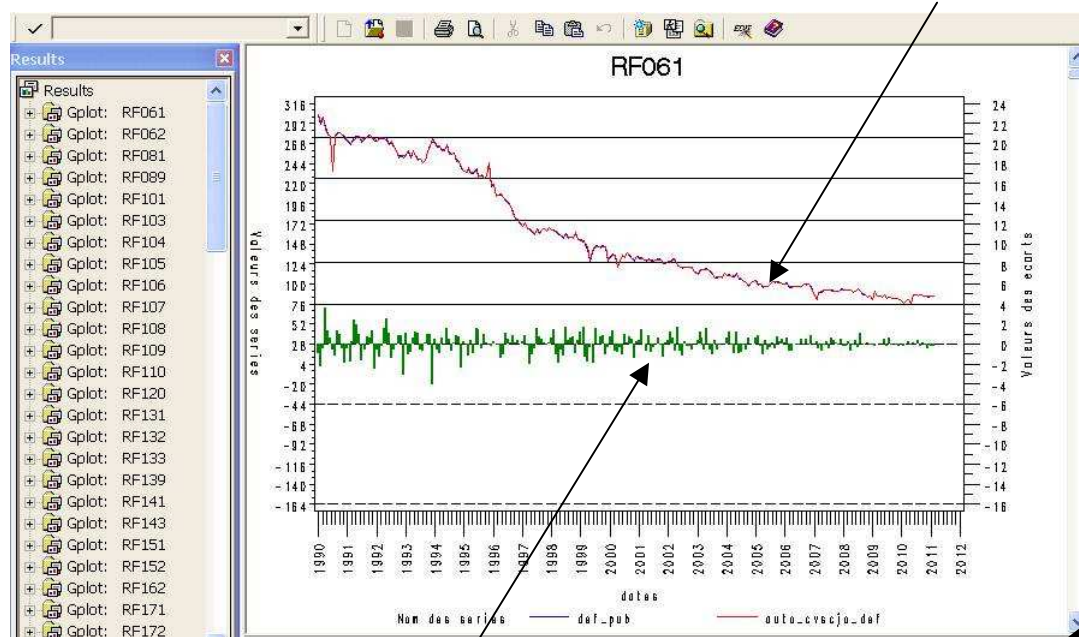
4.2. Comparaisons graphiques de résultats

Afin de choisir le meilleur modèle pour une série, il peut être utile de regarder les révisions par rapport aux séries publiées introduites par chacune des spécifications testées. Il ne s'agit pas à proprement parler d'un critère de qualité puisqu'un changement de modèle peut être justifié, s'il est meilleur même s'il est source de révisions importantes (encore faudra-t-il s'assurer que le nouveau modèle est stable dans le temps ...). Mais parmi deux jeux de spécifications présentant des qualités comparables, mieux vaut sélectionner, en terme d'affichage, celui qui sera source de révisions les plus faibles.

La comparaison de l'ensemble des composantes (séries CVS-CJO par exemple) d'un indice est malaisée sous Demetra+ (impossibilité de faire défiler simplement les courbes superposées obtenues avec deux jeux de spécifications pour un grand nombre de séries). Il a donc été décidé de développer un outil SAS traçant toutes les séries, avec un diagramme bâton attirant l'attention sur les différences entre les deux courbes et permettant de les faire défiler simplement à l'aide des flèches directionnelles.

Visualisation des graphiques des séries de l'IPJ (séries publiées / séries désaisonnalisées automatiquement) sous SAS :

Partie haute :
graphique superposant
les séries issues des
deux traitements.



Partie Basse :
Diagramme bâton
représentant les
différences entre les
deux courbes.

Ascenseur
permettant de faire
défiler les séries.

5. Évaluation de la qualité d'un indice désaisonnalisé de façon indirecte

Un fichier Excel résume la qualité de la désaisonnalisation à chaque niveau d'agrégation de l'indice.

Allure générale du fichier Excel de résumé de la qualité sur les différents niveaux d'agrégation

		NOTE GLOB	Qualité du Modèle Reg-ARIMA						Qualité de la décomposition					
			note sur 20 points	Tests du bruit blanc		Tests d'hétéroscédasticité		Examen des points aberrants	Note Reg-Arima totale de la série	Recherche d'effets résiduels sur la série cvs-cjo		Recherche d'effets résiduels sur l'irrégulier		Note globale sur la qualité de la décomposition (note sur 10)
				Ljung-Box	Box-Pierce	Ljung-Box ²	Box-Pierce ²			% de points aberrants dans la série < 5%	Présence d'un effet jours ouvrables résiduel	Présence d'un effet saisonnier résiduel	Présence d'un effet jours ouvrables résiduel	
BE - Industrie manufacturière, indus	187 283.0	18.5	3.2	3.4	3.5	3.5	3.9	17.6	3.9	3.6	4.0	4.0	4.0	19.5
C INDUSTRIE MANUFACTURIÈRE	163 304.0	18.5	3.2	3.5	3.5	3.5	3.9	17.6	3.9	3.6	4.0	4.0	4.0	19.4
<i>CZ - dont : industrie manufacturière</i>	<i>163 304.0</i>	<i>18.5</i>	<i>3.2</i>	<i>3.5</i>	<i>3.5</i>	<i>3.5</i>	<i>3.9</i>	<i>17.6</i>	<i>3.9</i>	<i>3.6</i>	<i>4.0</i>	<i>4.0</i>	<i>4.0</i>	<i>19.4</i>
C1 - Fabrication de denrées alimentaires, de boissons et de produits à base de viande	21 822.0	16.6	2.6	2.8	2.5	2.5	3.6	14.0	4.0	3.2	4.0	4.0	4.0	19.2
10.1 Transformation et conservation de la viande et préparation de produits à base de viande	4 418.0	18.0	4.0	4.0	2.0	2.0	4.0	16.0	4.0	4.0	4.0	4.0	4.0	20.0
10.3 Transformation et conservation de fruits et légumes	4 247.0	3.0	0.0	0.0	0.0	0.0	2.0	2.0	4.0	0.0	4.0	4.0	4.0	16.0
10.4 Fabrication d'huiles et graisses végétales et animales	534.0	16.0	0.0	0.0	4.0	4.0	4.0	12.0	4.0	4.0	4.0	4.0	4.0	20.0
10.5 Fabrication de produits laitiers	2 901.0	13.0	2.0	4.0	4.0	4.0	4.0	18.0	4.0	4.0	4.0	4.0	4.0	20.0

Des notes synthétiques sont d'abord évaluées au niveau de chaque série. La qualité du modèle reg-ARIMA d'une part et la qualité de décomposition font l'objet du calcul d'une note synthétique. Étant donné la difficulté à constituer des notes pertinentes sur l'amplitude des révisions, la qualité des révisions a été exclue du calcul de la note de qualité globale. **La note de qualité générale de la série** est évaluée comme une moyenne des notes de qualité du modèle reg-ARIMA et de qualité de la décomposition.

Les notes de qualité des séries désaisonnalisées peuvent ne pas apparaître dans le fichier Excel exporté par les programmes SAS. Le principe de calcul reste le même et les notes sont toujours des moyennes pondérées de notes obtenues par les traitements des séries constitutives de l'agrégat.

5.1. Évaluation de la qualité du modèle reg-ARIMA

La qualité du modèle reg-ARIMA est évaluée par une note sur 20 à partir des diagnostics sur les résidus du modèle :

- Tests de blancheur des résidus (absence de corrélations linéaires) basés sur les statistiques de Ljung-Box(24) et de Box-Pierce(24). Ces notes valent 4 en cas d'acceptation de l'hypothèse d'absence de corrélations entre les résidus.
- Tests d'homoscédasticité des résidus basés sur les statistiques de Ljung-Box(24) et de Box-Pierce(24) calculées sur le carré des résidus. Ces notes valent 4 en cas d'acceptation de l'hypothèse d'absence d'autocorrélation du carré des résidus.

et du diagnostic sur le pourcentage des points aberrants

- La note vaut 4 si le pourcentage est < 5, 2 si le pourcentage est >=5 et <10 et 0 sinon.

Détail des notes des séries du secteur CI :

Note sur le pourcentage de points de la série modélisés par un outlier.

Notes sur la qualité des résidus du modèle (égales à 4 en cas d'absence de corrélations).

		Qualité du Modèle Reg-ARIMA					Note Reg-Arima totale de la série
		Tests du bruit blanc		Tests d'hétéroscédasticité		Examen des points aberrants	
		Ljung-box	Box-Pierce	Ljung-Box ²	Box-Pierce ²	% de points aberrants dans la série < 5%	
CI - Fabrication de produits informatiques, électroniques et optiques		2.9	3.1	3.6	3.6	4.0	17.2
	26.1 Fabrication de composants et cartes électroniques	4.0	4.0	4.0	4.0	4.0	20.0
	26.2 Fabrication d'ordinateurs et d'équipements périphériques	2.0	4.0	4.0	4.0	4.0	18.0
	26.3 Fabrication d'équipements de communication	0.0	0.0	4.0	4.0	4.0	12.0
	26.4 Fabrication de produits électroniques grand public	0.0	0.0	0.0	0.0	4.0	4.0
	26.5 Fabrication d'instruments et d'appareils de mesure, d'essai et de navigation ; horlogerie	4.0	4.0	4.0	4.0	4.0	20.0
	26.6 Fabrication d'équipements d'irradiation médicale, d'équipements électromédicaux et électrothérapeutiques	4.0	4.0	0.0	0.0	4.0	12.0
	26.7 Fabrication de matériels optique et photographique	0.0	2.0	0.0	0.0	4.0	6.0

La note sur la qualité du modèle ARIMA est obtenue par sommation des notes sur la qualité des résidus et du % de points aberrants.

La pondération est équivalente pour tous les diagnostics sur les résidus. Aucun diagnostic contrôlant la normalité des résidus n'est pris en compte dans les calculs. Aucune note sur la concentration des points aberrants ne figure dans le fichier Excel.

Comparaison des pondérations utilisées dans les fichiers SAS et dans le fichier de calcul des notes des niveaux agrégés

Variables	libellés	poids (macros SAS)	poids (fichier résumé Excel, autres onglets de résultats)	poids (macros SAS)	poids (fichier résumé Excel, autres onglets de résultats)
box_pierce	Note liée à la statistique de Box-Pierce(24) sur les résidus	2	4	22,2	20
box_pierce_var	Note liée à la statistique de Box-Pierce(24) sur le carré des résidus	1	4	11,1	20
ljung_box	Note liée à la statistique de Ljung-Box sur les résidus	2	4	22,2	20
ljung_box_var	Note liée à la statistique de Ljung-Box sur le carré des résidus	1	4	11,1	20
normalite	Note sur le test de normalité (Doornick-Hansen)	1	0	11,1	0
normalite_aplatissement	Note sur le test du Kurtosis	0	0	0,0	0
normalite_asymetrie	Note sur le test du Skewness	2	0	22,2	0
p_aberrants	note sur le nombre de points aberrants	0	4	0	20
note_total_arima	Note globale sur la qualité du modèle ARIMA	ens. = 9	ens. = 20	ens.=100	ens. = 100

5.2. Évaluation de la qualité de décomposition

L'évaluation de la qualité de décomposition repose sur :

- Les notes considérant l'absence d'effets résiduels (CJO ou CVS) sur les séries CVS-CJO et les composantes irrégulières. Dans le fichier Excel, les notes sur les effets résiduels valent 4 en cas d'absence d'effets et 0 sinon.
- La note sur la qualité d'ajustement valant 4 si la décomposition est acceptable, 0 sinon. Cette note repose sur la statistique Q résumant les statistiques M1 à M11 ([X-12-ARIMA](#)) ou sur la décomposition empirique de la variance décrite dans les chapitres précédents ([TRAMO-SEATS](#)).

Détail des notes sur la qualité de décomposition pour les séries du secteur CI (fichier Excel)

Effets résiduels sur la série CVS-CJO (notes égales à 4 en cas d'absence d'effets). Effets résiduels sur la composante irrégulière (notes égales à 4 en cas d'absence d'effets).

		Qualité de la décomposition					
		Recherche d'effets résiduels sur la série cvs-cjo		Recherche d'effets résiduels sur l'irrégulier		Qualité de la décomposition acceptable (2 = OUI, 0 = NON)	Note globale sur la qualité de la décomposition (note sur 10)
		Présence d'un effet jours ouvrables résiduel	Présence d'un effet saisonnier résiduel	Présence d'un effet jours ouvrables résiduel	Présence d'un effet saisonnier résiduel		
CI - Fabrication de produits informatiques, électroniques et optiques		4.0	3.2	4.0	4.0	3.9	19.2
	26.1 Fabrication de composants et cartes électroniques	4.0	4.0	4.0	4.0	4.0	20.0
	26.2 Fabrication d'ordinateurs et d'équipements périphériques	4.0	4.0	4.0	4.0	4.0	20.0
	26.3 Fabrication d'équipements de communication	4.0	0.0	4.0	4.0	4.0	16.0
	26.4 Fabrication de produits électroniques grand public	4.0	4.0	4.0	4.0	0.0	16.0
	26.5 Fabrication d'instruments et d'appareils de mesure, d'essai et de navigation ; horlogerie	4.0	4.0	4.0	4.0	4.0	20
	26.6 Fabrication d'équipements d'irradiation médicale, d'équipements électromédicaux et électrothérapeutiques	4.0	4.0	4.0	4.0	4.0	20
	26.7 Fabrication de matériels optique et photographique	4.0	4.0	4.0	4.0	4.0	20

Note sur la qualité de décomposition (statistiques M1 à M11 ou décomposition empirique de la variance)

La note sur la qualité de décomposition est obtenue en sommant les notes sur la présence d'effets résiduels et la note sur la qualité de décomposition.

5.3. Évaluation de la qualité générale des séries

La note de qualité globale est obtenue en effectuant la moyenne des notes sur la qualité du modèle reg-ARIMA et sur la qualité de décomposition.

5.4. Évaluation de la qualité des niveaux d'agrégation supérieurs

Les notes de qualité des niveaux agrégés sont estimées à partir de moyennes pondérées des notes de chaque série désaisonnalisée les constituant. La pondération utilisée est celle des séries dans l'indice. Ainsi, l'impact des indicateurs de qualité des séries sur la qualité des agrégats est proportionnel à leur poids dans l'indice.

Exemple : le secteur CI est composé de 7 séries : (à remplacer)

	pondérations	NOTE GLOBALE	Qualité du Modèle Reg-ARIMA					Note Reg-Arima totale de la série
		note sur 20 points	Tests du bruit blanc		Tests d'hétéroscédasticité		Examen des points aberrants	
			Ljung-Box	Box-Pierce	Ljung-Box ²	Box-Pierce ²	% de points aberrants dans la série < 5%	
BE - Industrie manufacturière, industries extractives et autres	187 283.0	18.5	3.2	3.4	3.5	3.5	3.9	17.6
C INDUSTRIE MANUFACTURIÈRE	163 304.0	18.5	3.2	3.5	3.5	3.5	3.9	17.6
<i>CZ - dont : industrie manufacturière</i>	<i>163 304.0</i>	<i>18.5</i>	<i>3.2</i>	<i>3.5</i>	<i>3.5</i>	<i>3.5</i>	<i>3.9</i>	<i>17.6</i>
[...]								
C3 - Fabrication d'équipements électriques, électroniques, informatiques ; fabrication de machines	27 581.0	18.3	3.7	3.7	3.2	3.2	3.8	17.6
CI - Fabrication de produits informatiques, électroniques et optiques	8 764.0	18.2	2.9	3.1	3.6	3.6	4.0	17.2=(
26.1 Fabrication de composants et cartes électroniques	3 395.0	20.0	4.0	4.0	4.0	4.0	4.0	20.0 x3 395.0
26.2 Fabrication d'ordinateurs et d'équipements périphériques	500.0	19.0	2.0	4.0	4.0	4.0	4.0	18.0 + 18.0
26.3 Fabrication d'équipements de communication	1 731.0	14.0	0.0	0.0	4.0	4.0	4.0	12.0 x1 731.0
26.4 Fabrication de produits électroniques grand public	122.0	10.0	0.0	0.0	0.0	0.0	4.0	4.0 + 4.0
26.5 Fabrication d'instruments et d'appareils de mesure, d'essai et de navigation ; horlogerie	2 272.0	20.0	4.0	4.0	4.0	4.0	4.0	20.0 x2 272
26.6 Fabrication d'équipements d'irradiation médicale, d'équipements électromédicaux et électrothérapeutiques	484.0	16.0	4.0	4.0	0.0	0.0	4.0	12.0 +12
26.7 Fabrication de matériels optique et photographique	260.0	13.0	0.0	2.0	0.0	0.0	4.0	6.0 +6
								x260
)8 764.0

La bonne note générale du secteur CI est surtout due aux séries 26.1 Fabrication de composants et cartes électroniques (note de 20 et pondération de 3 395.0) et 26.5 Fabrication d'instruments et d'appareils de mesure, d'essai et de navigation ; horlogerie (note de 20 et pondération de 2 272) ayant les plus grosses pondérations.

Les fichiers résumant la qualité aux différents niveaux d'agrégation peuvent être comparés entre eux afin par exemple de mesurer des évolutions de qualité avant et après une campagne de désaisonnalisation ou de comparer les qualités obtenues à l'aide de deux jeux de spécifications.

Évolution de la qualité des séries de l'IPI au cours de la campagne 2011

	pondérations	Nouvelles spécifications	Anciennes spécifications	Variation de qualité
		note sur 20 points	note sur 20 points	note sur 20 points
BE - Industrie manufacturière, industries extractives et autres	187 283.0	18.6	18.5	0.1
C INDUSTRIE MANUFACTURIÈRE	163 304.0	18.9	18.5	0.4
<i>CZ - dont : industrie manufacturière</i>	<i>163 304.0</i>	<i>18.9</i>	<i>18.5</i>	<i>0.4</i>
[...]				0
C3 - Fabrication d'équipements électriques, électroniques, informatiques ; fabrication de machines	27 581.0	17.5	18.3	-0.8
CI - Fabrication de produits informatiques, électroniques et optiques	8 764.0	18.3	18.2	0.1
26.1 Fabrication de composants et cartes électroniques	3 395.0	19.0	20.0	-1.0
26.2 Fabrication d'ordinateurs et d'équipements périphériques	500.0	18.0	19.0	-1.0
26.3 Fabrication d'équipements de communication	1 731.0	16.0	14.0	2.0
26.4 Fabrication de produits électroniques grand public	122.0	16.0	10.0	6.0
26.5 Fabrication d'instruments et d'appareils de mesure, d'essai et de navigation ; horlogerie	2 272.0	20.0	20.0	0.0
26.6 Fabrication d'équipements d'irradiation médicale, d'équipements électromédicaux et électrothérapeutiques	484.0	16.0	16.0	0.0
26.7 Fabrication de matériels optique et photographique	260.0	16.0	13.0	3.0

Pour le secteur CI, la faible amélioration de qualité constatée (+0.1 points) provient de l'évolution de la qualité de la série 26.3 Fabrication d'équipements de communication dont la note passe de 14 à 16. Dans une moindre mesure, les séries 26.4 Fabrication de produits électroniques grand public (note passant de 10 à 16) et 26.7 Fabrication de matériels optique et photographique (note évoluant de 13 à 16) participent également à cette évolution favorable de la qualité du secteur.

Bibliographie

[1] Necmettin Alpay KOÇAK, « On analysis of German Industrial Production with Demetra+ », *Turkish Statistical Institute*.

[2] Sylwia Grudkowska, «Demetra+ User Manual», National Bank of Poland