

**Séries statistiques des remboursements des prestations prises en charge par le régime agricole au titre de la branche maladie**  
**Méthodologie statistique de correction des variations saisonnières (CVS), des jours ouvrables (CJO) et de complétude**

### Introduction

La CCMSA produit chaque mois des séries statistiques relatives aux remboursements des prestations prises en charge par le régime agricole au titre de la branche maladie, et qu'elle publie sur son site : <http://statistiques.msa.fr/>

Elles portent sur :

- les remboursements mensuels des soins de ville et des cliniques privées, tous risques confondus (maladie, maternité, accidents du travail et maladies professionnelles), les versements des pensions d'invalidité et des indemnités relatives au congé paternité, en date de remboursement, pour la France métropolitaine. Les données sont présentées en valeur brute.
- les remboursements mensuels des soins de ville, tous risques confondus (maladie, maternité, accidents du travail et maladies professionnelles), en date de soins, pour la France métropolitaine. Les données font d'une part, l'objet d'un traitement de complétude et d'autre part, sont corrigées des jours ouvrables et des variations saisonnières.

Les séries statistiques relatives à un mois de remboursement M sont publiées au début du mois M+2.

Les séries statistiques relatives à un mois de soins M sont publiées au début du mois M+4.

Le document définit dans une première partie la méthodologie utilisée pour la correction des jours ouvrables et des variations saisonnières et dans une seconde partie la méthode de complétude des données en date de soins.

## 1- Correction des variations saisonnière et des jours ouvrables

### 1-a Principe

Les séries statistiques relatives aux remboursements des prestations prises en charge par le régime agricole au titre de la branche maladie sont produites à partir du système d'information statistique de la CCMSA (RAAMSES) qui collecte des données brutes issues des chaînes de liquidation des prestations des organismes de MSA.

Une analyse temporelle réalisée à partir de données brutes ne rendrait pas compte des phénomènes statistiques et conjoncturels intervenant au cours de la période étudiée comme :

- les phénomènes saisonniers de consommation de soins liés aux épisodes épidémiques, à des habitudes sociologiques, etc. ;
- l'effet lié au nombre de jours ouvrables d'un mois donné, et aussi à la composante journalière du mois ;
- les mesures de revalorisations tarifaires ou des changements réglementaires ayant un impact financier sur les remboursements de soins ;
- la situation conjoncturelle : l'évolution démographique du régime agricole, l'emploi, la croissance économique, etc.

Une correction des variations saisonnières et des jours ouvrables est donc appliquée aux différentes séries afin de permettre la comparaison des différents mois.

Les séries sont constituées par régime (salariés agricoles et non-salariés agricoles) et tous risques confondus (maladie, maternité, accident du travail et maladie professionnelle).

### 1-b Méthodologie

La correction des séries des variations saisonnières et des jours ouvrables est effectuée selon la méthode X13-ARIMA-SEATS proposée par le logiciel Jdemetra+, qui combine la méthode X12 ARIMA<sup>1</sup> et la méthode SEATS<sup>2</sup>.

La méthode X12-ARIMA est constituée de deux phases :

- Une phase de pré-ajustement (REG-ARIMA) : correction préalable par régression linéaire des points aberrants, ruptures de tendance et effets de calendrier et prolongement des séries en faisant des prévisions par modélisation ARIMA.
- Une phase de décomposition de la série en tendance-cycle, saisonnalité et irrégulier à l'aide de moyennes mobiles (méthode X11).

La méthode SEATS part d'une série initiale modélisée par un modèle ARIMA, estime ce modèle, puis partitionne de façon additive le spectre de la série en spectres associés aux différentes composantes (tendance, saisonnalité et irrégulier), chaque composante étant alors modélisée par un ARIMA.

---

<sup>1</sup> ARIMA : Auto-regressive Integrated Moving Average = processus auto régressif à moyennes mobiles intégrées

<sup>2</sup> SEATS : Signal Extraction in Arima Time Series

La détermination des coefficients de correction est faite à partir d'un historique de données allant de janvier 2012 jusqu'au dernier mois dont les données sont disponibles. Le choix de cet historique est révisé chaque année à la suite de tests retenant différents scénarios de durée.

Chaque mois, avec l'apport d'un mois de données supplémentaire, l'ensemble des coefficients sont ré-estimer par rafraîchissement (méthode "partial concurrent adjustment / last outliers).

Les modèles pour lesquels Jdemetra+ indique une qualité "severe" ou "uncertain" sont systématiquement revus et un meilleur modèle est recherché : changement de jeux de régresseurs (2 variables CJO au lieu de 7...) ou changement de modèles parmi ceux proposés par Jdemetra+ (méthode X13 ou tramo-seats).

Par ailleurs, une attention particulière est portée à certaines séries, dont la qualité est qualifiée de "bonne" mais qui présentent des profils très erratiques, même après désaisonnalisation (en particulier, les indemnités journalières des non-salariés agricoles, les remboursements des soins dentaires et les dispositifs médicaux).

Les différents modèles proposés par Jdemetra+ sont comparés, et est sélectionné celui dont l'amplitude d'oscillation est la plus faible.

Les traitements de corrections sont appliqués sur chaque série de données élémentaires. Les séries correspondant à des agrégats (par exemple le total des remboursements des soins effectués par les omnipraticiens libéraux) sont constituées par sommation de séries élémentaires.

L'ensemble des séries sont constituées dès l'origine "tous risques confondus" (maladie, maternité, accidents du travail et maladies professionnelles).

Afin que les changements de tarification des prestations ne viennent pas perturber les calculs de saisonnalité, et lorsque cela est possible (par exemple, pour les consultations, les visites, les actes d'infirmiers et de masseurs-kinésithérapeutes, etc.), les traitements de correction sont effectués sur des séries de dénombrement de prestation. Ensuite, pour chacune des séries, est appliqué un montant moyen mensuel qui permet ainsi de disposer de séries en montant remboursé.

A défaut (par exemple, les médicaments – il faudrait autrement disposer de séries de données pour chaque médicament afin d'éliminer l'incidence des variations de prix), les traitements de correction sont directement effectués sur des séries constituées en montant remboursé.

## 2- Méthode de complétude des données en date de soins

### 2-a Principe

Tout assuré disposant d'un délai de deux ans et trois mois pour transmettre à sa caisse d'assurance maladie ses feuilles de soins et obtenir un remboursement, il faut donc attendre ce délai pour être certain que l'intégralité des soins effectués un mois donné a fait l'objet d'un remboursement.

A défaut d'attendre 27 mois, pour analyser la conjoncture en date de soins, il convient d'estimer le montant des remboursements correspondant aux soins effectués les mois antérieurs alors qu'ils n'ont pas encore tous été mis au remboursement. La méthodologie de cette prévision est appelée « complétude ».

La date de soins correspond à la date de début effective d'exécution des soins pour les actes médicaux ou paramédicaux, à la date de délivrance pour les médicaments et les dispositifs médicaux, à la date de début d'indemnisation pour les indemnités journalières.

### 2-b Méthodologie

La complétude est réalisée en reportant sur les montants remboursés pour des soins effectués au mois  $m$ , le taux de complétude observé sur les montants remboursés du même mois  $m$  de l'année précédente.

$$\text{Mois de soins } m_{\text{complété à fin } r} = \text{Mois de soins } m_{\text{remb à fin } r} * \text{Taux de complétude}_{m-12}$$

$$\text{où } \left\{ \begin{array}{l} \text{Taux de complétude}_{m-12} = \frac{(\text{Mois de soins } m-12)_{\text{remb à fin } r}}{(\text{Mois de soins } m-12)_{\text{remb à fin } r-12}} \\ r \text{ étant le dernier mois de remboursement disponible} \\ m \text{ étant le mois de soins ; } m = \{r - 2, r - 3, r - 4, \dots, r - 25\} \end{array} \right.$$

Ce calcul ne peut se faire pour les deux derniers mois de soins dont on dispose, la part des soins de ces mois non encore remboursés étant trop importante pour être estimée.

Les taux de complétude des 25 derniers mois de soins sont revus chaque mois à la suite de la prise en compte d'un nouveau mois de données en remboursement.

Chaque mois, la DSEF produit un tableau indiquant la révision du total des soins de ville des mois précédents, induite par l'intégration d'un mois de remboursement supplémentaire.