

Le logiciel G-Series version 2.00

Les documents qui suivent servent de guides pour l'utilisation du système SAS G-Series. Ces guides donnent des explications sur les options et énoncés des procédures et sur les paramètres des macros de G-Series. Chaque guide comprend aussi des exemples que vous pouvez copier et coller dans l'éditeur de votre session SAS et exécuter tel quel.

Aperçu

G-Series est le nom sous lequel différentes procédures et macros SAS reliées aux séries chronologiques sont regroupées. Ces outils sont développés à Statistique Canada par la Division de l'ingénierie de systèmes en collaboration avec le Centre de recherche et d'analyse en séries chronologiques de la Division des méthodes d'enquêtes auprès des entreprises.

Deux procédures et une macro sont disponibles :

- **PROC BENCHMARKING** : Cette procédure sert à restaurer la cohérence entre des données de séries chronologiques d'une même variable cible mais mesurées à des fréquences différentes (ex : infra-annuelle et annuelle). Cette procédure est disponible dans les versions de production v1.04 et v2.0 de G-Series.
- **PROC TSRAKING** : Cette procédure sert à restaurer les règles d'additivité dans un ensemble de séries chronologiques mesurées à la même fréquence. Cette procédure est disponible dans les versions de production v1.04 et v2.0 de G-Séries.
- **Macro GSeriesTSBalancing** : Cette macro effectue l'équilibrage (la réconciliation) d'un système de séries chronologiques mesurées à la même fréquence étant donné un ensemble de contraintes linéaires. Cette procédure est disponible dans la version de production v2.0 de G-Séries.

Pour plus de renseignements, veuillez contacter l'équipe de soutien de G-Series à l'adresse courriel [G-Series](mailto:G-Series@statcan.gc.ca) (G-Series@statcan.gc.ca) ou consultez le site web [G-Series](#) (sur l'intranet de Statistique Canada seulement).

La procédure **BENCHMARKING**

Aperçu

Cette procédure assure la cohérence entre des séries chronologiques d'une même variable cible mesurée à différents intervalles (ex : infra-annuelle et annuelle). L'étalonnage consiste à imposer le niveau de la série d'étalons tout en minimisant, autant que possible, les révisions au mouvement observé dans la série infra-annuelle. La procédure permet également l'étalonnage non contraignant où la série d'étalons peut aussi être révisée.

La procédure peut aussi être utilisée à d'autres fins reliées à l'étalonnage, telles que la distribution temporelle (désagrégation de la série d'étalons en observations plus fréquentes), la calendarisation (cas particulier de distribution temporelle) et le raccordement (connexion de différents segments de séries chronologiques en une série unique cohérente).

Syntaxe de la procédure

PROC BENCHMARKING <option(s)>;

VAR var1 </alt1> <...varN </altN> >;

WITH var1 </alt1> <...varN </altN> >;

BY *variable(s)*;

Pour faire ceci	Utilisez cet énoncé
Procéder à un traitement séparé pour les observations de chaque groupe BY	BY
Identifier, au besoin, la ou les séries infra-annuelles à étalonner ainsi que les coefficients d'altérabilité associés	VAR
Identifier, au besoin, la ou les séries d'étalons ainsi que les coefficients d'altérabilité associés	WITH

Énoncé **PROC BENCHMARKING**

PROC BENCHMARKING <option(s)>;

Pour faire ceci	Utilisez cette option
Spécifier l'ensemble de données d'entrée qui contient les étalons	BENCHMARKS=
Spécifier une valeur pour le biais	BIAS=
Spécifier l'option pour l'estimation du biais	BIASOPTION=
Spécifier la valeur du paramètre λ (lambda)	LAMBDA=
Spécifier l'ensemble de données de sortie qui contient les étalons utilisés par la procédure	OUTBENCHMARKS=

Spécifier l'ensemble de données de sortie qui contient des données supplémentaires utiles à la production de tableaux analytiques et de graphiques	OUTGRAPHTABLE=
Spécifier l'ensemble de données de sortie qui contient la ou les séries infra-annuelles étalonnées	OUTSERIES=
Spécifier la valeur pour le paramètre ρ (rho)	RHO=
Spécifier l'ensemble de données d'entrée qui contient la ou les séries infra-annuelles à étalonner	SERIES=
Spécifier le niveau de tolérance pour le test ultime, en pourcentage ou en valeur absolue	TOLERANCEPERCENT= TOLERANCEVALUE=
Spécifie le seuil pour l'identification des valeurs négatives	TOLNEGRESULT=
Afficher des résultats intermédiaires et le temps de CPU requis dans le journal (SAS log)	VERBOSE
Afficher un message d'avertissement dans le journal (SAS log) lorsqu'une valeur négative est trouvée (voir option TOLNEGRESULT=)	WARNNEGRESULT NOWARNNEGRESULT

Options

BENCHMARKS=*ensemble-de-données-SAS*

spécifie l'ensemble de données d'entrée SAS qui contient les étalons. Cette option est obligatoire. L'ensemble de données doit contenir les variables numériques qui définissent la période de couverture de chaque étalon, soit: STARTYEAR, STARTPERIOD, ENDYEAR et ENDPERIOD. De plus, une variable numérique nommée VALUE (par défaut) contenant la valeur de l'étalon doit obligatoirement se trouver dans le fichier de données. Pour spécifier un autre nom de variable ou plus d'une série d'étalons, utilisez l'énoncé VAR et/ou WITH.

BIAS=*nombre réel*

spécifie la valeur du biais défini par l'utilisateur à utiliser pour la correction de la série infra-annuelle avant l'étalonnage. Le biais est ajouté à la série infra-annuelle avec un modèle additif (lorsque LAMBDA=0.0) et multiplié autrement (lorsque LAMBDA≠0.0). L'option BIAS= n'est pas obligatoire. La valeur par défaut est BIAS=0.0 lorsque LAMBDA=0.0 et BIAS=1.0 autrement. Il est à noter que l'option BIAS= n'est pas utilisée lorsque l'option BIASOPTION=3 (voir ci-dessous).

BIASOPTION=1, 2 ou 3

spécifie l'option pour l'estimation du biais. Cette option est obligatoire.

Si vous spécifiez la valeur suivante pour l'estimation du biais	La procédure
1	N'estimera pas le biais. Le biais utilisé pour corriger la série infra-annuelle sera la valeur spécifiée avec l'option BIAS= ou la valeur par défaut (voir l'option BIAS=).
2	Estimera le biais, imprimera le résultat dans le journal mais ne l'utilisera pas. Le biais utilisé pour corriger la série infra-

	annuelle sera la valeur spécifiée avec l'option BIAS= ou la valeur par défaut (voir l'option BIAS=).
3	Estimera le biais, imprimera le résultat dans le journal et appliquera la correction à l'aide de la valeur estimée du biais. Une valeur spécifiée à l'aide de l'option BIAS= sera imprimée dans le journal mais ne sera pas utilisée.

LAMBDA=*nombre réel (valeurs suggérées entre -3 et 3)*

spécifie la valeur du paramètre d'ajustement du modèle λ . Cette option est obligatoire. Les valeurs fréquemment utilisées sont LAMBDA=0.0 pour un modèle additif et LAMBDA=1.0 pour un modèle proportionnel. On obtient un ajustement au prorata en spécifiant LAMBDA=0.5 et RHO=0.0.

OUTBENCHMARKS=*ensemble-de-données-SAS*

spécifie l'ensemble de données de sortie SAS qui contiendra les étalons utilisés par la procédure. Si le nom de l'ensemble de données de sortie n'est pas spécifié, alors PROC BENCHMARKING le créera en utilisant les règles d'affectation des noms DATA n . L'ensemble de données contiendra les mêmes variables que l'ensemble de données BENCHMARKS=. Si des variables BY sont spécifiées, elles se retrouveront aussi dans cet ensemble de données.

OUTGRAPHTABLE=*ensemble-de-données-SAS*

spécifie l'ensemble de données de sortie SAS qui contiendra des données supplémentaires utiles à la production de tableaux analytiques et de graphiques dans le cas d'étalons sans chevauchement. Cette option n'est pas obligatoire et aucun ensemble de données de sortie ne sera créé si elle n'est pas utilisée. Le tableau suivant décrit les variables qui se trouvent dans cet ensemble de données. Si des variables BY sont spécifiées, elles apparaîtront aussi dans cet ensemble de données.

Nom de la variable	Description
ALTBENCHMARKS	Le nom de la variable utilisée pour spécifier les coefficients d'altérabilité de la série d'étalons
ALTBENCHMARKSVALUE	Les valeurs des coefficients d'altérabilité de la série d'étalons
ALTSERIES	Le nom de la variable utilisée pour spécifier les coefficients d'altérabilité de la série infra-annuelle
ALTSERIESVALUE	Les valeurs des coefficients d'altérabilité de la série infra-annuelle
AVGBENCHMARK	La valeur de l'étalon divisée par le nombre de périodes couvertes; valeur manquante s'il n'y a pas d'étalon
AVGBENCHMARKSUBANNUALRATIO	Les ratios** entre les valeurs des variables AVGBENCHMARK et AVGSUBANNUAL; valeur manquante s'il n'y a pas d'étalon
AVGSUBANNUAL	La moyenne des valeurs de la série infra-annuelle initiale par période de couverture de l'étalon; valeur manquante s'il n'y a pas d'étalon
BENCHMARKED	Les valeurs de la série infra-annuelle étalonée

BENCHMARKEDSUBANNUALRATIO	Les ratios** entre les valeurs des variables BENCHMARKED et SUBANNUAL
BIAS	La valeur du BIAIS utilisée (calculée par la procédure ou fournie par l'utilisateur)
DATE	Variable de type caractère qui combine les valeurs des variables YEAR et PERIOD
GROWTHRATEBENCHMARKED	Les taux de croissance** dans la série infra-annuelle étalonnée
GROWTHRATESUBANNUAL	Les taux de croissance** dans la série infra-annuelle initiale
LAMBDA	La valeur du paramètre λ (lambda)
M	Les identificateurs de période de couverture des étalons, $m = 1, \dots, M$; valeur manquante s'il n'y a pas d'étalon
PERIOD	Valeurs des périodes
PERIODICITY	Le nombre maximum de périodes à l'intérieur d'une année (ex : la périodicité est 4 pour une série trimestrielle)
RHO	La valeur du paramètre ρ (rho)
SUBANNUAL	Les valeurs de la série infra-annuelle initiale
SUBANNUALCORRECTED	Les valeurs de la série infra-annuelle corrigée (données rectifiées avec le biais)
T	Les identificateurs des périodes infra-annuelles, $t = 1, \dots, T$
VARBENCHMARKS	Le nom de la variable utilisée pour spécifier la série d'étalons
VARSERIES	Le nom de la variable utilisée pour spécifier la série infra-annuelle
YEAR	Valeurs des années

** Les termes « ratios » et « taux de croissance » réfèrent ici à une terminologie couramment utilisée dans le domaine des séries chronologiques et signifient en fait « différences » et « premières différences » respectivement lorsque le modèle additif est utilisé (LAMBDA=0.0).

OUTSERIES=*ensemble-de-données-SAS*

spécifie l'ensemble de données de sortie SAS qui contiendra la ou les séries infra-annuelles étalonnées. Si le nom de l'ensemble de données de sortie n'est pas spécifié, alors PROC BENCHMARKING le créera en utilisant les règles d'affectation des noms DATA n . L'ensemble de données contiendra les mêmes variables que l'ensemble de données SERIES=. Si des variables BY sont spécifiées, elles se retrouveront aussi dans cet ensemble de données.

RHO=*nombre réel entre 0 et 1 (inclusif)*

spécifie la valeur du paramètre autorégressif ρ dans l'intervalle [0,1]. Cette option est obligatoire.

SERIES=*ensemble-de-données-SAS*

spécifie l'ensemble de données d'entrée SAS qui contient la série infra-annuelle à étalonner. Cette option est obligatoire. L'ensemble de données doit contenir les variables numériques YEAR et PERIOD. De plus, une variable numérique nommée VALUE (par défaut) contenant la valeur de la série infra-annuelle à étalonner doit obligatoirement se trouver dans le fichier de données. Pour spécifier un autre nom de variable ou plus d'une série infra-annuelle, utilisez l'énoncé VAR.

TOLERANCEPERCENT= TOLP= <i>nombre réel positif (incluant 0)</i>		TOLERANCEVALUE= TOLV= <i>nombre réel positif (incluant 0)</i>
--	--	--

spécifie le niveau de tolérance, en pourcentage ou en valeur absolue, à utiliser pour le test ultime dans le cas d'étalons contraignants (coefficient d'altérabilité = 0.0 pour les étalons). Ce test compare les étalons fournis à l'entrée avec ceux calculés à partir de la série infra-annuelle étalonée. Ce nombre est optionnel.

Par défaut TOLERANCEVALUE=1E-3.

:

Exemple : Pour une tolérance de 1%, spécifiez TOLERANCEPERCENT =0.01
Pour une tolérance de 10, spécifiez TOLERANCEVALUE=10

TOLNEGRESULT=
TOLN=*nombre réel négatif (excluant 0)*

spécifie le seuil pour l'identification des valeurs négatives. Une valeur est considérée négative lorsqu'elle est inférieure à ce seuil. Ce nombre est optionnel. La valeur par défaut est -1E-3.

VERBOSE

affiche des résultats intermédiaires et le temps de CPU requis dans le journal (SAS log). Avec cette option, les matrices intermédiaires seront imprimées dans le journal ainsi que le temps de CPU utilisé pour inverser les matrices, effectuer les multiplications de matrices, etc.

WARNNEGRESULT | NOWARNNEGRESULT

spécifie si la procédure doit afficher un message d'avertissement dans le journal (SAS log) lorsqu'une valeur négative créée par la procédure est inférieure au seuil tel que spécifié par l'option TOLNEGRESULT=. L'option par défaut est WARNNEGRESULT.

Énoncé VAR

VAR var1 </alt1> <...varN </altN> >;

Arguments requis

var1...varN

spécifie la ou les variables qui contiennent les valeurs des séries infra-annuelles à étalonner. Les variables doivent être de type numérique. L'énoncé VAR est optionnel. Si cet énoncé n'est pas spécifié, la procédure cherchera la variable VALUE dans le fichier d'entrée SERIES=.

alt1...altN

spécifie au besoin la ou les variables qui contiennent les coefficients d'altérabilité applicables aux séries infra-annuelles. Les variables de coefficients d'altérabilité doivent être de type numérique. Dans le cas où une variable de coefficients d'altérabilité n'est pas spécifiée pour une série infra-annuelle donnée, la procédure utilisera la valeur par défaut d'un coefficient d'altérabilité pour une série infra-annuelle, soit 1.0.

Les coefficients d'altérabilité entrent en jeu seulement après que la série infra-annuelle initiale ait été corrigée pour le biais (le cas échéant). Par exemple, ça signifie que la spécification d'un coefficient d'altérabilité de 0.0 pour période donnée de la série infra-annuelle *ne* résultera *pas* en une valeur inchangée après étalonnage s'il y a correction pour le biais (se référer aux options BIAS= et BIASOPTION=). La spécification de coefficients d'altérabilité n'est pas permise lorsque RHO=1.

Énoncé WITH

WITH var1 </alt1> <...varN </altN> >;

Arguments requis

var1...varN

spécifie la ou les variables qui contiennent les valeurs des séries d'étalons. Les variables doivent être de type numérique. L'énoncé VAR est optionnel.. Si cet énoncé n'est pas spécifié, la procédure cherchera la variable VALUE dans le fichier d'entrée BENCHMARKS=.

alt1...altN

spécifie au besoin la ou les variables qui contiennent les coefficients d'altérabilité applicables aux étalons. Les variables de coefficients d'altérabilité doivent être de type numérique. Dans le cas où une variable de coefficients d'altérabilité n'est pas spécifiée pour une série d'étalons donnée, la procédure utilisera la valeur par défaut d'un coefficient d'altérabilité pour un étalon, soit 0.0 (étalons contraignants). La spécification de coefficients d'altérabilité n'est pas permise lorsque RHO=1.

Énoncé BY

BY variable-1 <... variable-n>;

Arguments requis

variable(s)

spécifie la ou les variables que la procédure utilise pour former les groupes BY. Les variables BY doivent être présentes dans les deux fichiers d'entrée (BENCHMARKS= et SERIES=) et elles apparaîtront dans les trois fichiers de sortie (OUTBENCHMARKS=, OUTGRAPHTABLE= et OUTSERIES=). Plus d'une

variable peut être spécifiée. Cet énoncé est optionnel. Les variables peuvent être de type caractère ou numérique.

Détails

- Si une valeur manquante apparaît dans une des variables du fichier d'entrée BENCHMARKS= (autre que les variables BY), les observations avec valeurs manquantes sont retirées, un message d'avertissement est affiché dans le journal et la procédure poursuit son exécution.
- Si une valeur manquante apparaît dans les variable YEAR et/ou PERIOD du fichier d'entrée SERIES= et que des variables BY sont spécifiées, le groupe BY correspondant est retiré, un message d'avertissement est affiché dans le journal et la procédure traite les autres groupes BY. Si aucune variable BY n'est spécifiée, aucun traitement n'est effectué et un message d'avertissement apparaît dans le journal.
- Si une valeur manquante apparaît dans une variable désignant une série infra-annuelle dans le fichier d'entrée SERIES= (voir l'énoncé VAR) et que des variables BY sont spécifiées, le groupe BY correspondant est retiré, un message d'avertissement est affiché dans le journal et la procédure traite les autres groupes BY. Si aucune variable BY n'est spécifiée, aucun traitement n'est effectué sur la série infra-annuelle en question, un message d'avertissement apparaît dans le journal et la procédure poursuit son traitement avec la série infra-annuelle suivante (le cas échéant).
- La procédure ne permet pas l'étalonnage proportionnel ($LAMBDA \neq 0$) d'une série indicatrice qui est nulle sur toute la période de couverture d'un étalon non-nul. Une telle situation génère un message d'erreur dans le journal (SAS log).
- Avec une valeur de $RHO=1$, seuls les coefficients d'altérabilité par défaut (0 pour un étalon et 1 pour une série infra-annuelle) sont valides. La spécification de variables contenant les coefficients d'altérabilité n'est donc pas permise. Si tel est le cas, la procédure les ignore et affiche un message d'avertissement dans le journal (SAS log).
- Les coefficients d'altérabilité entrent en jeu seulement après que la série infra-annuelle initiale ait été corrigée pour le biais (le cas échéant). Par exemple, ça signifie que la spécification d'un coefficient d'altérabilité de 0 pour période donnée de la série infra-annuelle *ne* résultera *pas* en une valeur inchangée après étalonnage s'il y a correction pour le biais.
- Des noms à deux niveaux tels que *libref.ensemble-de-données-SAS* peuvent être utilisés pour les fichiers d'entrée et pour rendre permanent les fichiers de sortie.
- Si des variables BY sont spécifiées, celles-ci apparaîtront dans tous les fichiers de sortie.

Exemples

```
/* Exemple 1: Cas simple avec une seule série trimestrielle à étalonner à des valeurs annuelles */
```

```
/* Série infra-annuelle (trimestrielle) */
```

```
data mySeries;
```

```
input year  
      period  
      value;
```

```
datalines;
```

```
1998 1 1.9  
1998 2 2.4  
1998 3 3.1  
1998 4 2.2  
1999 1 2.0  
1999 2 2.6  
1999 3 3.4  
1999 4 2.4
```

```

2000 1 2.3
;

/* Étalons annuels */
data myBenchmarks;
input startyear
      startperiod
      endyear
      endperiod
      value;
datalines;
1998 1 1998 4 10.3
1999 1 1999 4 10.2
;

/* Étalonnage en utilisant...
- la valeur recommandée de RHO pour des séries trimestrielles (RHO=0.729)
- un modèle proportionnel (LAMBDA=1)
- une série infra-annuelle corrigée pour le biais estimé par la procédure
  (BIASOPTION=3)
*/
proc benchmarking
  benchmarks=myBenchmarks
  series=mySeries
  outbenchmarks=outBenchmarks
  outseries=outSeries
  outgraphtable=outGraphTable
  rho=0.729
  lambda=1
  biasoption=3;
run;

proc print data=outseries;
run;

/* Exemple 2: Deux séries trimestrielles à étalonner à des totaux annuels, avec des
   groupes BY et des coefficients d'altérabilité définis par l'utilisateur. */

/* Séries infra-annuelles (trimestrielles) où les 1er et 2e trimestres de 1999 pour le
   groupe BY « A » de la série van_sales_quarterly sont non-altérables (voir la variable
   alt_van) */
data mySeries;
length group $1;
input group /* Identificateur des groupes BY */
      year
      period
      car_sales_quarterly
      van_sales_quarterly
      alt_van @@; /* Coefficients d'altérabilité pour van_sales_quarterly */
datalines;
A 1998 1 1851 1900 1    A 1998 2 2436 2200 1
A 1998 3 3115 3000 1    A 1998 4 2205 2000 1
A 1999 1 1987 1900 0    A 1999 2 2635 2500 0
A 1999 3 3435 3800 1    A 1999 4 2361 2500 1
A 2000 1 2183 2100 1    A 2000 2 2822 3100 1
A 2000 3 3664 3650 1    A 2000 4 2550 2950 1
A 2001 1 2342 3300 1    A 2001 2 3001 4000 1
A 2001 3 3779 3290 1    A 2001 4 2538 2600 1

```

```

A 2002 1 2363 2010 1    A 2002 2 3090 3600 1
A 2002 3 3807 3500 1    A 2002 4 2631 2100 1
A 2003 1 2601 2050 1    A 2003 2 3063 3500 1
A 2003 3 3961 4290 1    A 2003 4 2774 2800 1
A 2004 1 2476 2770 1    A 2004 2 3083 3080 1
A 2004 3 3864 3100 1    A 2004 4 2773 2800 1
A 2005 1 2489 3100 1    A 2005 2 3082 2860 1
B 1998 1 1851 1900 1    B 1998 2 2436 2200 1
B 1998 3 3115 3000 1    B 1998 4 2205 2000 1
B 1999 1 1987 1900 1    B 1999 2 2635 2500 1
B 1999 3 3435 3800 1    B 1999 4 2361 2500 1
B 2000 1 2183 2100 1    B 2000 2 2822 3100 1
B 2000 3 3664 3650 1    B 2000 4 2550 2950 1
B 2001 1 2342 3300 1    B 2001 2 3001 4000 1
B 2001 3 3779 3290 1    B 2001 4 2538 2600 1
B 2002 1 2363 2010 1    B 2002 2 3090 3600 1
B 2002 3 3807 3500 1    B 2002 4 2631 2100 1
B 2003 1 2601 2050 1    B 2003 2 3063 3500 1
B 2003 3 3961 4290 1    B 2003 4 2774 2800 1
B 2004 1 2476 2770 1    B 2004 2 3083 3080 1
B 2004 3 3864 3100 1    B 2004 4 2773 2800 1
B 2005 1 2489 3100 1    B 2005 2 3082 2860 1

```

```
;
```

```
/* Étalons annuels */
```

```
data myBenchmarks;
```

```
length group $1;
```

```
input group /* Identificateur des groupes BY */
```

```
startYear
```

```
startPeriod
```

```
endYear
```

```
endPeriod
```

```
car_sales_annual
```

```
van_sales_annual @@;
```

```
datalines;
```

```

A 1998 1 1998 4 10324 12000    A 1999 1 1999 4 10200 10400
A 2000 1 2000 4 10582 11550    A 2001 1 2001 4 11097 11400
A 2002 1 2002 4 11582 14500    A 2003 1 2003 4 11092 16000
B 1998 1 1998 4 10324 12000    B 1999 1 1999 4 10200 10400
B 2000 1 2000 4 10582 11550    B 2001 1 2001 4 11097 11400
B 2002 1 2002 4 11582 14500    B 2003 1 2003 4 11092 16000

```

```
;
```

```
/* Étalonnage en utilisant...
```

```
- la valeur recommandée de RHO pour des séries trimestrielles (RHO=0.729)
```

```
- un modèle proportionnel (LAMBDA=1)
```

```
- une série infra-annuelle non corrigée pour le biais (BIASOPTION=1 et option  
BIAS= non spécifiée)
```

```
*/
```

```
proc benchmarking
```

```
benchmarks=myBenchmarks
```

```
series=mySeries
```

```
outbenchmarks=outBenchmarks
```

```
outseries=outSeries
```

```
outgraphtable=outGraphTable
```

```
rho=0.729
```

```
lambda=1
```

```
biasoption=1;
```

```
var car_sales_quarterly van_sales_quarterly/alt_van;
```

```
with car_sales_annual van_sales_annual;
```

```
by group;  
run;
```

Notes

Ce document est un guide pour l'utilisation de la procédure BENCHMARKING. PROC BENCHMARKING fait partie du logiciel G-Series de Statistique Canada, précédemment appelé Forillon.

Pour plus de renseignements, veuillez vous adresser à l'équipe de soutien de G-Series à l'aide de l'adresse courriel de [G-Series](mailto:G-Series@statcan.gc.ca) (G-Series@statcan.gc.ca) ou consultez le site web de [G-Series](#) (Intranet de Statistique Canada seulement).

Bibliographie

Bloem, A. M., R. J. Dippelsman, et N. Ø. Mæhle (2001). **Quarterly National Accounts Manual, Concepts, Data Sources and Compilation**, International Monetary Fund, Washington DC.

Dagum, E. B. et P. Cholette (2006). **Benchmarking, Temporal Distribution and Reconciliation Methods of Time Series**, Springer-Verlag, New York, Lecture Notes in Statistics, Vol. 186

Fortier, S. et B. Quenneville (2007). « Theory and Application of Benchmarking in Business Surveys », recueil de la conférence **ICES III**, Introductory Overview Lecture at the International Conference on Establishment Surveys III, juin 2007.

Latendresse, E., M. Djona et S. Fortier (2007). « Benchmarking Sub-Annual Series to Annual Totals – From Concepts to SAS® Procedure and SAS® Enterprise Guide® Custom Task », recueil de la conférence **SAS Global Forum**, avril 2007.

Quenneville, B., S. Fortier, Z.-G. Chen et E. Latendresse (2006). « Recent Developments in Benchmarking to Annual Totals in X-12-ARIMA and at Statistics Canada », recueil de la conférence **2006 Eurostat Conference on Seasonality, Seasonal Adjustment and Their Implications for Short-Term Analysis and Forecasting**, Luxembourg, mai 2006.

La procédure TSRAKING

Aperçu

Cette procédure restaure des contraintes d'agrégation transversales dans un système de séries chronologiques. Les contraintes d'agrégation peuvent provenir d'une table à 1 ou 2 dimensions. Optionnellement, des contraintes temporelles peuvent aussi être préservées.

Syntaxe de la procédure

PROC TSRAKING <option(s)>;

ID *variable(s)*;

Pour faire ceci	Utilisez cet énoncé
Identifier les variables du fichier d'entrée qui seront transférées dans le fichier de sortie.	ID

Énoncé PROC TSRAKING

PROC TSRAKING <option(s)>;

Pour faire ceci	Utilisez cette option
Spécifier l'ensemble de données contenant les coefficients d'altérabilité	ALTERABILITY=
Spécifier la valeur par défaut du coefficient d'altérabilité pour les totaux annuels des séries composantes (contraintes temporelles)	ALTERANNUAL=
Spécifier la valeur par défaut du coefficient d'altérabilité pour les variables SERIES (séries composantes)	ALTERSERIES=
Spécifier la valeur par défaut du coefficient d'altérabilité pour les variables TOTAL1 (totaux de contrôle de la 1 ^{ère} dimension)	ALTERTOTAL1=
Spécifier la valeur par défaut du coefficient d'altérabilité pour les variables TOTAL2 (totaux de contrôle de la 2 ^e dimension)	ALTERTOTAL2=
Spécifier l'ensemble de données d'entrée	DATA=
Spécifier l'ensemble de données contenant les métadonnées	METADATA=
Spécifier l'ensemble de données de sortie	OUT=
Spécifier le niveau de tolérance souhaité pour le test ultime, en pourcentage ou en valeur absolue	TOLERANCEPERCENT= TOLERANCEVALUE=
Spécifier le seuil pour l'identification des valeurs négatives	TOLNEGRESULT=
Afficher des résultats intermédiaires et le temps de CPU requis dans le journal (SAS log)	VERBOSE

Afficher un message d'avertissement dans le journal (SAS log) lorsqu'une valeur négative est trouvée (voir option TOLNEGRESULT=)

WARNNEGRESULT |
NOWARNNEGRESULT

Options

ALTERABILITY=*ensemble-de-données-SAS*

spécifie l'ensemble de données SAS qui contient les variables de coefficients d'altérabilité. Toute variable de coefficient d'altérabilité doit correspondre à une série composante ou un total de contrôle, c'est-à-dire qu'une variable avec le même nom doit être présente dans l'ensemble de données d'entrée DATA=. L'option ALTERABILITY= n'est pas obligatoire. Si elle est utilisée, les coefficients d'altérabilité de cet ensemble de données remplaceront ceux spécifiés avec les options ALTERxxx (voir plus bas). Lorsque l'ensemble de données DATA= contient plusieurs observations et que l'ensemble de données ALTERABILITY= n'en contient qu'une seule, les coefficients d'altérabilité qu'il contient sont utilisés (répétés) pour toutes les observations de l'ensemble de données DATA=. Alternativement, l'ensemble de données ALTERABILITY= peut ne contenir le même nombre d'observations que l'ensemble de données DATA=.

ALTERANNUAL=*nombre réel positif (incluant 0)*

spécifie la valeur par défaut du coefficient d'altérabilité pour les contraintes temporelles (totaux annuels) des séries composantes. Ce nombre est optionnel et la valeur par défaut est 0.0 (totaux contraignants). Cette valeur s'appliquera aux contraintes temporelles pour lesquelles un coefficient d'altérabilité n'a pas déjà été spécifié à l'aide de la variable ALTERANNUAL de l'ensemble de données METADATA=.

ALTERSERIES=*nombre réel positif (incluant 0)*

spécifie la valeur par défaut du coefficient d'altérabilité pour les séries composantes. Ce nombre est optionnel et la valeur par défaut est 1.0. Cette valeur s'appliquera aux séries composantes pour lesquelles un coefficient d'altérabilité n'a pas déjà été spécifié à l'aide de l'ensemble de données ALTERABILITY=.

ALTERTOTAL1=*nombre réel positif (incluant 0)*

spécifie la valeur par défaut du coefficient d'altérabilité pour les totaux de contrôle de la 1^{ère} dimension. Ce nombre est optionnel et la valeur par défaut est 0.0 (totaux contraignants). Cette valeur s'appliquera aux totaux de contrôle pour lesquels un coefficient d'altérabilité n'a pas déjà été spécifié à l'aide de l'ensemble de données ALTERABILITY=.

ALTERTOTAL2=*nombre réel positif (incluant 0)*

spécifie la valeur par défaut du coefficient d'altérabilité pour les totaux de contrôle de la 2^e dimension. Ce nombre est optionnel et la valeur par défaut est 0.0 (totaux contraignants). Cette valeur s'appliquera aux totaux de contrôle pour lesquels un coefficient d'altérabilité n'a pas déjà été spécifié à l'aide de l'ensemble de données ALTERABILITY=.

DATA=*ensemble-de-données-SAS*

spécifie l'ensemble de données SAS qui contient les variables représentant le système de séries chronologiques, c'est-à-dire les séries composantes et totaux de contrôle transversaux à réconcilier. Lorsque l'ensemble de données DATA= contient plusieurs observations, la somme des valeurs de chaque

variable est également préservée, définissant des contraintes temporelles implicites. L'option DATA= est obligatoire.

METADATA=ensemble-de-données-SAS

spécifie l'ensemble de données SAS qui contient les métadonnées décrivant les contraintes d'agrégation (règles d'additivité) utilisées par la procédure. L'option METADATA= est obligatoire. L'ensemble de données doit contenir deux variables de type caractère : SERIES et TOTAL1. Deux autres variables sont optionnelles : TOTAL2 (de type caractère) et ALTERANNUAL (de type numérique). Les valeurs de la variable SERIES représentent les noms de variable des séries composantes dans l'ensemble de données d'entrée DATA=. Pareillement, les valeurs des variables TOTAL1 et TOTAL2 représentent les noms de variable des totaux de contrôle des 1^{ère} et 2^e dimensions respectivement dans l'ensemble de données d'entrée DATA=. La variable ALTERANNUAL quant à elle, contient le coefficient d'altérabilité pour la contrainte temporelle (total annuel) associée à la série composante.

OUT=ensemble-de-données-SAS

spécifie l'ensemble de données de sortie SAS qui contient les séries composantes réconciliées, les totaux de contrôle réconciliés et toute variable spécifiée par l'énoncé ID. L'option OUT= n'est pas obligatoire. Si le nom de l'ensemble de données de sortie n'est pas spécifié, alors PROC TSRAKING le créera en utilisant les règles d'affectation des noms DATA_n.

TOLERANCEPERCENT=	TOLERANCEVALUE=
TOLP=nombre réel positif (incluant 0)	TOLV=nombre réel positif (incluant 0)

spécifie le niveau de tolérance, en pourcentage ou en valeur absolue, à utiliser pour le test ultime dans le cas où les totaux de contrôle sont contraignants (coefficient d'altérabilité = 0.0 pour les totaux de contrôle). Ce test compare les totaux de contrôle contraignants avec ceux calculés à partir des séries composantes réconciliées. Ce nombre est optionnel.

Par défaut : TOLERANCEVALUE=1E-3.

Exemple : Pour une tolérance de 1%, spécifiez TOLERANCEPERCENT =0.01
Pour une tolérance de 10, spécifiez TOLERANCEVALUE=10

TOLNEGRESULT=
TOLN=nombre réel négatif (excluant 0)

spécifie le seuil pour l'identification des valeurs négatives. Une valeur est considérée négative lorsqu'elle est inférieure à ce seuil. Ce nombre est optionnel. La valeur par défaut est -1E-3.

VERBOSE

affiche des résultats intermédiaires et le temps de CPU requis dans le journal (SAS log).

WARNNEGRESULT | NOWARNNEGRESULT

spécifie si la procédure doit afficher un message d'avertissement dans le journal (SAS log) lorsqu'une valeur négative créée par la procédure est inférieure au seuil tel que spécifié par l'option TOLNEGRESULT=. L'option par défaut est WARNNEGRESULT.

Énoncé ID

ID *variable-1 ... variable-n;*

Utilisez l'énoncé ID pour identifier des variables additionnelles à transférer du fichier d'entrée DATA= au fichier de sortie OUT=. Par défaut, le fichier de sortie ne contient que les variables identifiées dans les métadonnées (se référer à l'option METADATA=). Les variables spécifiées dans l'énoncé ID doivent être présentes dans le fichier d'entrée DATA=.

Détails

- À l'exception des variables spécifiées dans l'énoncé ID, une valeur manquante dans le fichier d'entrée DATA= entraînera l'arrêt de la procédure
 - Une valeur manquante ou négative dans le fichier d'entrée ALTERABILITY= entraînera l'arrêt de la procédure.
 - Des noms à deux niveaux tels que *libref.ensemble-de-données-SAS* peuvent être utilisés pour les fichiers d'entrée et pour rendre permanent le fichier de sortie.
-

Exemples

```
/* EXEMPLE 1 */
```

```
/* Dans cet exemple à 1 dimension, la somme des valeurs de "cars" et "vans" doit être  
égale à la valeur de "total" */
```

```
data myMetadata;  
input series $5.  
      Total1 $5.;  
datalines;  
cars total  
vans total  
;
```

```
data myData;  
input cars  
      vans  
      total;  
datalines;  
25 5 40  
;
```

```
proc tsraking  
  metadata=myMetadata  
  data=myData  
  out=outData;  
run;
```

```
/* EXEMPLE 2 */
```



```

/* Dans cet exemple à 2 dimensions, nous disposons d'estimés de ventes de "cars" et
"vans" pour les régions "alb", "sask" and "man". La somme des 3 régions pour "cars" doit
être égale à son total de contrôle (cars_alb + cars_sask + cars_man = cars_total). La
somme des 3 régions pour "vans" doit aussi être égale à son total de contrôle (vans_alb +
vans_sask + vans_man = vans_total). De plus, dans chaque region, la somme "cars" et
"vans" doit être égale au total de contrôle régional (cars_alb + vans_alb = alb_total;
cars_sask + vans_sask = sask_total and cars_man + vans_man = man_total). Finalement, les
ventes de "vans" dans la région "sask" (vans_sask) ne sont pas altérables (coefficient
d'altérabilité = 0) */

```

```

data myMetadata;
input series $9.
      Total1 $11.
      Total2 $11.;
datalines;
cars_alb cars_total alb_total
cars_sask cars_total sask_total
cars_man cars_total man_total
vans_alb vans_total alb_total
vans_sask vans_total sask_total
vans_man vans_total man_total
;

data myData;
input cars_alb cars_sask cars_man vans_alb vans_sask vans_man
      alb_total sask_total man_total cars_total vans_total;
datalines;
12 14 13 20 20 24 30 31 32 40 53
;

data myalter;
input cars_alb cars_sask cars_man vans_alb vans_sask vans_man
      alb_total sask_total man_total cars_total vans_total;
datalines;
1 1 1 1 0 1 0 0 0 0 0
;

proc tsraking
  metadata=myMetadata
  data=myData
  alterability=myAlter
  out=outData;
run;

```

```

/* EXEMPLE 3 */

```

```

/* Dans cet exemple à 2 dimensions avec contraintes temporelles, nous avons les
contraintes transversales  $A_1 + A_2 + A_3 = A$ ;  $B_1 + B_2 + B_3 = B$ ;  $A_1 + B_1 = \_1$ ;  $A_2 + B_2 = \_2$  et  $A_3 + B_3 = \_3$ . Nous avons 4 observations (trimestres) pour chaque valeur et
nous voulons préserver les totaux annuels. Pour chaque variable, la somme des 4 valeurs
trimestrielles doit demeurer la même après l'étape de réconciliation. */

```

```

data mymetadata;
input series $3.
      total1 $3.
      total2 $3.;
datalines;
A_1 _1 A

```

```

B_1 _1 B
A_2 _2 A
B_2 _2 B
A_3 _3 A
B_3 _3 B
;

data mydata;
input A_1 A_2 A_3 B_1 B_2 B_3 A B _1 _2 _3;
datalines;
12 14 13 20 20 24 40 53 30 31 32
10 9 15 21 29 20 25 80 35 35 35
12 8 17 15 20 30 40 59 23 32 44
9 9 14 17 24 23 37 71 28 35 45
;

proc tsraking
  metadata=myMetadata
  data=myData
  out=outData;
run;

```

Notes

Ce document est un guide pour l'utilisation de la procédure TSRAKING. PROC TSRAKING fait partie du logiciel G-Series de Statistique Canada, précédemment appelé Forillon.

Pour plus de renseignements, veuillez vous adresser à l'équipe de soutien de G-Series à l'aide de l'adresse courriel de [G-Series](mailto:G-Series@statcan.gc.ca) (G-Series@statcan.gc.ca) ou consultez le site web de [G-Series](#) (Intranet de Statistique Canada seulement).

Bibliographie

Bérubé, J. et S. Fortier (2009). « PROC TSRAKING: An in-house SAS procedure for balancing time series », recueil de la conférence **JSM 2009**, Business and Economic Section. Alexandria, VA: American Statistical Association.

Dagum, E. B. et P. Cholette (2006). **Benchmarking, Temporal Distribution and Reconciliation Methods of Time Series**. Springer-Verlag, New York, Lecture Notes in Statistics, Vol. 186

Fortier, S. et B. Quenneville (2009). « Reconciliation and Balancing of Accounts and Time Series », recueil de la conférence **JSM 2009**, Business and Economic Section. Alexandria, VA: American Statistical Association.

Macro GSeriesTSBalancing

Aperçu

Cette macro effectue l'équilibrage (la réconciliation) d'un système de séries chronologiques étant donné un ensemble de contraintes linéaires. La solution d'équilibrage est obtenue en résolvant un problème d'optimisation quadratique (voir l'annexe) à l'aide de la procédure OPTMODEL. Lorsqu'une solution existe au problème d'équilibrage spécifié, la macro produit un ensemble de séries chronologiques pour lesquelles les valeurs respectent les contraintes d'équilibrage pour chacune des périodes de temps. La spécification de contraintes d'égalité et d'inégalité est permise. La préservation de totaux temporels peut également être spécifiée.

Cette macro requiert la composante SAS/OR et est compatible avec les versions 9.3 et plus récentes de SAS.

Syntaxe de la macro

Voici l'ensemble complet des paramètres de la macro incluant les valeurs par défaut lorsqu'applicable:

```
%GSeriesTSBalancing
(
  /* Paramètres obligatoires */
  inTS                = ,
  inProblemSpecs      = ,
  outTS               = ,

  /* Paramètres optionnels */
  timeVarName         = ,
  periodInterval      = ,
  temporalGrpInterval = ,

  optModelSolverOptions = WITH QP,
  displayLevel         = 2,

  outOptModelSummary   = ,
  outDetailedResults   = ,
  outEvaluatedConstraints = ,
  outTemporalTotals    = ,
  outProcessingGrps    = ,

  alterPos            = 1,
  alterNeg            = 1,
  alterMix            = 1,
  alterTemporalTotal  = 0,

  lowerBound          = ,
  upperBound          = ,

  toleranceValue       = ,
  toleranceValueTemporal = ,
  tolerancePercentTemporal = ,

  TSFormat            = WIDE,
  tallTSIDVarName      = _NAME_,
  tallTSValueVarName   = _VALUE_,
  tallTSAlterVarName   = ,
```

language = EN
)

Paramètres

Le tableau qui suit fournit une brève description de tous les paramètres de la macro.

Nom du paramètre	Brève description
inTS	Obligatoire. Fichier de données d'entrée.
inProblemSpecs	Obligatoire. Fichier de spécification du problème.
outTS	Obligatoire. Fichier de données de sortie (équilibrées).
timeVarName	Variable de temps (variable numérique de date, temps ou date-temps SAS) dans le fichier de données d'entrée.
periodInterval	Intervalle de date, temps ou date-temps SAS associé aux périodes du fichier de données d'entrée.
temporalGrpInterval	Intervalle de date, temps ou date-temps SAS associé aux totaux temporels à préserver.
optModelSolverOptions	Options du solveur pour la procédure OPTMODEL.
displayLevel	Niveau d'information à afficher dans le journal (SAS <i>log</i>) et les destinations ODS actives (ex : fenêtre « Sortie »).
outOptModelSummary	Fichier de sortie du <i>résumé d'OPTMODEL</i> .
outDetailedResults	Fichier de sortie des <i>résultats d'équilibrage détaillés</i> .
outEvaluatedConstraints	Fichier de sortie des <i>contraintes d'équilibrage évaluées</i> .
outTemporalTotals	Fichier de sortie des <i>totaux temporels</i> .
outProcessingGrps	Fichier de sortie des <i>groupes de traitement</i> .
alterPos, alterNeg et alterMix	Coefficients d'altérabilité réguliers.
alterTemporalTotals	Coefficients d'altérabilité des totaux temporels.
lowerBound	Borne inférieure pour les valeurs des séries chronologiques.
upperBound	Borne supérieure pour les valeurs des séries chronologiques.
toleranceValue	Seuil de tolérance pour les contraintes d'équilibrage.
toleranceValueTemporal et tolerancePercentTemporal	Seuil de tolérance pour les contraintes temporelles implicites spécifié en valeur absolue ou en pourcentage.
TSFormat	Définition du format des fichiers de données d'entrée et de sortie.
tallTSIDVarName	Variable du fichier de données d'entrée contenant les noms des séries chronologiques (lorsque TSFormat=TALL).
tallTSValueVarName	Variable du fichier de données d'entrée contenant les valeurs des séries chronologiques (lorsque TSFormat=TALL).
tallTSAlterVarName	Variable du fichier de données d'entrée contenant les coefficients d'altérabilité des séries chronologiques (lorsque TSFormat=TALL).
language	Langue d'affichage des messages dans le journal (SAS <i>log</i>).

La section qui suit donne plus de détails sur chacun des paramètres de la macro, en ordre alphabétique.

`alterMix` = *nombre réel positif (incluant 0)*

Coefficient d'altérabilité associé aux valeurs des séries chronologiques ayant un mélange de coefficients positifs et négatifs dans les contraintes d'équilibrage dans lesquelles elles sont impliquées. Les coefficients d'altérabilité fournis dans le fichier de spécification du problème ont priorité sur cette valeur.

La valeur par défaut de ce paramètre optionnel est 1.0.

`alterNeg` = *nombre réel positif (incluant 0)*

Coefficient d'altérabilité associé aux valeurs des séries chronologiques ayant un coefficient négatif dans toutes les contraintes d'équilibrage dans lesquelles elles sont impliquées. Les coefficients d'altérabilité fournis dans le fichier de spécification du problème ont priorité sur cette valeur.

La valeur par défaut de ce paramètre optionnel est 1.0.

`alterPos` = *nombre réel positif (incluant 0)*

Coefficient d'altérabilité associé aux valeurs des séries chronologiques ayant un coefficient positif dans toutes les contraintes d'équilibrage dans lesquelles elles sont impliquées. Les coefficients d'altérabilité fournis dans le fichier de spécification du problème ont priorité sur cette valeur.

La valeur par défaut de ce paramètre optionnel est 1.0.

`alterTemporalTotals` = *nombre réel positif (incluant 0)*

Coefficient d'altérabilité associé aux totaux temporels des séries chronologiques. Les coefficients d'altérabilité fournis dans le fichier de spécification du problème ont priorité sur cette valeur.

La valeur par défaut de ce paramètre optionnel est 0.0.

`displayLevel` = 0, 1, 2, 3, 4 ou 5

Niveau d'information à afficher dans le journal (SAS log) et les destinations ODS actives (ex : fenêtre « Sortie »).

Information affichée	0	1	2	3	4	5
Entête et message de fin d'exécution de la macro dans le journal	√	√	√	√	√	√
Messages d'erreur et d'avertissement dans le journal	√	√	√	√	√	√
Contenu de la variable macro <code>_OROPTMODEL_</code> dans le journal		√				√
Notes de procédure OPTMODEL dans le Journal			√	√	√	√
Les tableaux ODS suivants de la procédure OPTMODEL dans les destinations ODS actives : <ul style="list-style-type: none">« Résumé du problème »« Résumé de la solution »			√	√	√	√

Information affichée	0	1	2	3	4	5
Les tableaux ODS suivants de la procédure OPTMODEL dans les destinations ODS actives : <ul style="list-style-type: none"> • « Méthodes de calcul des dérivées » (solveur NLP) • « Options du solveur » • « Statistiques d'optimisation » • autres tables ODS spécifiques au solveur (lorsqu'applicable) 				√	√	√
Valeurs équilibrées des séries chronologiques (résultats de l'équilibrage) dans les destinations ODS actives					√	√
Problème d'optimisation détaillé (énoncé EXPAND de la procédure OPTMODEL) dans les destinations ODS actives						√

La valeur par défaut de ce paramètre optionnel est 2.

`inProblemSpecs` = *ensemble-de-données-SAS*

Nom du fichier de spécification du problème. Ce fichier sert à spécifier les contraintes linéaires (règles d'équilibrage) qui définissent les relations qui doivent être restaurées. Il permet également, si nécessaire, la spécification de coefficients d'altérabilité et de bornes inférieures et supérieures spécifiques à certaines séries qui auraient priorité sur les valeurs génériques définies par les paramètres `alterPos`, `alterNeg`, `alterMix`, `alterTemporalTotals`, `lowerBound` and `upperBound`.

Ces informations sont fournies au moyen de 4 variables obligatoires (`_type_`, `_col_`, `_row_` et `_coef_`) et d'une variable optionnelle (`_timeVal_`). Le fichier contient 2 types d'enregistrements :

- Enregistrements *définissant une étiquette* (quand `_type_` n'est pas vide)
- Enregistrements *spécifiant une information* (quand `_type_` est vide)

Variable	Type	Détails – Définition d'une étiquette	Détails – Spécification d'une information
<code>_type_</code>	Car.	Mot-clé réservé indiquant à la macro comment interpréter les <i>enregistrements spécifiant une information</i> : EQ: contrainte d'égalité (=) LE: contrainte d'inégalité ≤ GE: contrainte d'inégalité ≥ lowerBd: borne inférieure upperBd: borne supérieure alter: coefficient d'altérabilité régulier alterTmp: coef. d'altér. de total temp.	Ne s'applique pas (valeur <i>vide</i>)
<code>_col_</code>	Car.	Ne s'applique pas (valeur <i>vide</i>)	Nom d'une série chronologique ou mot-clé <code>_rhs_</code> afin de spécifier une borne de contrainte différente de 0.
<code>_row_</code>	Car.	Étiquette à associer au « mot-clé type »	Étiquette associée au « mot-clé type » identifiant le type d'information spécifiée.
<code>_coef_</code>	Num.	Ne s'applique pas (valeur <i>manquante</i>)	Borne de contrainte (<code>_rhs_</code>) ou information de série chro., ex : <ul style="list-style-type: none"> • coefficient de contrainte (0 par défaut) • borne inf. (défaut défini par le para. <code>lowerBound</code>) • borne sup. (défaut défini par le para. <code>upperBound</code>) • coefficient d'altérabilité régulier (défaut défini par les paramètres <code>alterPos</code>, <code>alterNeg</code> et <code>alterMix</code>) • coefficient d'altérabilité de total temporel (défaut défini par le paramètre <code>alterTemporalTotals</code>)

Variable	Type	Détails – Définition d'une étiquette	Détails – Spécification d'une information
<code>_timeVal_</code>	Num.	Ne s'applique pas (valeur <i>manquante</i>)	Valeur de temps optionnelle visant à restreindre l'application des coef. d'altérabilité ou des bornes à une période de temps précise. Ces coefficients et bornes s'appliquent à toutes les périodes de temps par défaut (quand <code>_timeVal_</code> n'est pas incluse dans le fichier ou quand sa valeur est <i>manquante</i>). Pour les problèmes sans variable de temps, les valeurs de <code>timeVal</code> correspondent aux numéros d'observation (à l'intérieur de la série chro.) dans le fichier de données d'entrée.

La macro n'est pas sensible à la casse en ce qui concerne l'information fournie dans ce fichier. Ce paramètre est obligatoire.

`inTS` = *ensemble-de-données-SAS*

Nom du fichier de données d'entrée contenant les valeurs des séries chronologiques impliquées dans le problème d'équilibrage. Il peut également contenir une variable de date, temps ou date-temps pour la préservation des totaux temporels (voir le paramètre `timeVarName`). Les données peuvent être fournies en format *wide* (par défaut) ou *tall* (voir le paramètre `TSFormat` pour plus de détails). En mode *tall*, les paramètres `tallTSIDVarName`, `tallTSValueVarName` et `tallTSAAlterVarName` permettent de préciser le rôle des variables dans le fichier de données d'entrée. Ce paramètre est obligatoire.

`language` = EN *ou* FR

Langue d'affichage des messages dans le journal (SAS *log*). La langue anglaise (EN) est utilisée par défaut.

`lowerBound` = *nombre réel*

Borne inférieure pour les valeurs des séries chronologiques. Si non spécifiée, la borne inférieure est $-\infty$. Les bornes inférieures spécifiées dans le fichier de spécification du problème ont priorité sur cette valeur.

Ce paramètre optionnel n'a pas de valeur par défaut.

`optModelSolverOptions` = *chaîne de caractère*

Options du solveur pour la procédure OPTMODEL. Deux solveurs sont appropriés pour les problèmes d'équilibrage de séries chronologiques : QP (« quadratic programming ») et NLP (« nonlinear programming »), chacun disposant de ses propres options (voir la documentation SAS). QP est très efficace et est le solveur recommandé (et celui utilisé par défaut par la macro). Le mot-clé NONE peut être spécifié afin de laisser le choix du solveur à la procédure OPTMODEL, qui choisira le solveur NLP (avec options par défaut) pour ce problème d'optimisation lorsqu'exécuté avec la version 9.3 de SAS.

Les valeurs admissibles pour ce paramètre dans le contexte de la macro sont :

Solveur QP :	<code>optModelSolverOptions = WITH QP </ options></code>
Solveur NLP :	<code>optModelSolverOptions = WITH NLP </ options></code>
Choix laissé à OPTMODEL :	<code>optModelSolverOptions = NONE</code>

Cette macro requiert la composante SAS/OR et est compatible avec les versions 9.3 et plus récentes de SAS. La valeur par défaut de ce paramètre optionnel est `WITH QP`.

outDetailedResults = *ensemble-de-données-SAS*

Nom du fichier de sortie des *résultats d'équilibrage détaillés*. Ce fichier contient de l'information détaillée sur les séries chronologiques impliquées dans le problème d'équilibrage. Il sera créé seulement si ce paramètre est spécifié.

Variable	Description
<code>_NAME_</code> ¹	Nom de la série chronologique.
<code>_T_</code>	Compteur de période utilisé à l'interne par la macro.
Valeur du paramètre <code>timeVarName</code>	Variable de temps (lorsque spécifiée).
<code>_VALUE_IN_</code>	Valeur avant équilibrage.
<code>_ALTER_</code>	Coefficient d'altérabilité.
<code>_VALUE_OUT_</code>	Valeur après équilibrage.
<code>_DIF_</code>	Différence (<code>_VALUE_OUT_</code> - <code>_VALUE_IN_</code>)
<code>_RDIF_</code>	Différence relative (<code>_DIF_</code> / <code>_VALUE_IN_</code>); manquante si <code>_VALUE_IN_</code> = 0.

Ce paramètre optionnel n'a pas de valeur par défaut.

outEvaluatedConstraints = *ensemble-de-données-SAS*

Nom du fichier de sortie des *contraintes d'équilibrage évaluées*. Ce fichier contient de l'information détaillée sur les contraintes d'équilibrage spécifiées. Par exemple, il montre l'écart de la solution optimale par rapport aux bornes des contraintes d'inégalité. Il permet également d'identifier les contraintes qui ne sont pas respectées dans des cas de solutions non optimales (ex : problèmes d'équilibrage complexes infaisables). Le fichier sera créé seulement si ce paramètre est spécifié.

Variable	Description
<code>_CON_</code>	Étiquette de la contrainte.
<code>_T_</code>	Compteur de période utilisé à l'interne par la macro.
Valeur du paramètre <code>timeVarName</code>	Variable de temps (lorsque spécifiée).
<code>_TYPE_</code>	Type de contrainte (EQ, LE ou GE).
<code>_RHS_</code>	Constante « côté droit » de la contrainte (borne de la contrainte).
<code>_LHS_IN_</code>	Côté gauche de la contrainte évalué avant équilibrage.
<code>_LHS_OUT_</code>	Côté gauche de la contrainte évalué après équilibrage.
<code>_DIF_</code>	Différence (<code>_LHS_OUT_</code> - <code>_LHS_IN_</code>)
<code>_RDIF_</code>	Différence relative (<code>_DIF_</code> / <code>_LHS_IN_</code>); manquante si <code>_LHS_IN_</code> = 0.

Ce paramètre optionnel n'a pas de valeur par défaut.

¹ En mode *tall* (paramètre `TSFormat=TALL`), le nom de cette variable provient du paramètre `tallTSIDVarName`.

outOptModelSummary = *ensemble-de-données-SAS*

Nom du fichier de sortie du *résumé d'OPTMODEL*. Ce fichier contient l'information incluse dans la variable macro `_OROPTMODEL_` pour chaque groupe de traitement (état de la solution, valeur de la fonction objective, nombre d'itérations, temps d'exécution, etc.), résumant ainsi l'exécution de la procédure OPTMODEL par groupe de traitement. Le fichier sera créé seulement si ce paramètre est spécifié.

Variable	Description
<code>_PROC_GRP_</code>	Compteur de groupe de traitement utilisé à l'interne par la macro. ²
<code>_PROC_GRP_LABEL_</code>	Étiquette du groupe de traitement (description plus significative). ²
<code>terme₁</code>	1 ^{er} terme de la variable macro <code>_OROPTMODEL_</code> .
<code>terme₂</code>	2 ^e terme de la variable macro <code>_OROPTMODEL_</code> .
...	...
<code>terme_k</code>	<i>k</i> ^e terme de la variable macro <code>_OROPTMODEL_</code> .

Les détails concernant les termes inclus dans la variable `_OROPTMODEL_` pour chaque solveur se trouvent dans la documentation SAS de la procédure OPTMODEL.

Ce paramètre optionnel n'a pas de valeur par défaut.

outProcessingGrps = *ensemble-de-données-SAS*

Nom du fichier de sortie des *groupes de traitement*. Ce fichier contient de l'information détaillée sur les périodes et les groupes de traitement du problème d'équilibrage. Il sera créé seulement si ce paramètre est spécifié.

Variable	Description
<code>_T_</code>	Compteur de période utilisé à l'interne par la macro.
Valeur du paramètre <code>timeVarName</code>	Variable de temps (lorsque spécifiée).
<code>_PROC_GRP_</code>	Compteur de groupe de traitement utilisé à l'interne par la macro. ²
<code>_PROC_GRP_LABEL_</code>	Étiquette du groupe de traitement (description plus significative). ²

Ce paramètre optionnel n'a pas de valeur par défaut.

outTemporalTotals = *ensemble-de-données-SAS*

Nom du fichier de sortie des *totaux temporels*. Ce fichier contient de l'information détaillée sur les totaux temporels du problème d'équilibrage, soit les valeurs des séries chronologiques cumulées pour chacun des groupes temporels complets (ex : les années qui contiennent les 12 mois). Il sera vide si les totaux

² Lors de la préservation de totaux temporels, les périodes d'un groupe temporel complet (ex : années contenant 12 mois) partagent la même valeur de groupe de traitement avec une étiquette correspondant à la valeur formatée du groupe temporel (ex : 2011) alors que les périodes d'un groupe temporel incomplet (ex : années contenant moins de 12 mois) ont des valeurs distinctes avec une étiquette correspondant à la valeur formatée de la période (ex : 2011-01, 2011-01, etc.). Autrement, lorsque les totaux temporels ne sont pas préservés, toutes les périodes ont une valeur de groupe de traitement distincte avec une étiquette correspondant à la valeur formatée de la période. L'étiquette `_PROC_GRP_LABEL_` n'apparaît pas dans le fichier lorsque le paramètre `temporalGrpInterval=_ALL_` ou lorsqu'une variable de temps (paramètre `timeVarName`) n'est pas spécifiée.

temporels ne sont pas préservés. Pour plus de détails sur la préservation des totaux temporels, voir le paramètre `temporalGrpInterval`. Le fichier sera créé seulement si le paramètre est spécifié.

Variable	Description
<code>_NAME_</code> ¹	Nom de la série chronologique.
<code>_PROC_GRP_</code>	Compteur de groupe de traitement utilisé à l'interne par la macro. ²
<code>_PROC_GRP_LABEL_</code>	Étiquette du groupe de traitement (description plus significative). ²
<code>_TOTAL_IN_</code>	Total temporel avant équilibrage.
<code>_ALTERNP_</code>	Coefficient d'altérabilité du total temporel.
<code>_TOTAL_OUT_</code>	Total temporel après équilibrage.
<code>_DIF_</code>	Différence (<code>_TOTAL_OUT_</code> - <code>_TOTAL_IN_</code>)
<code>_RDIF_</code>	Différence relative (<code>_DIF_</code> / <code>_TOTAL_IN_</code>); manquante si <code>_TOTAL_IN_</code> = 0.

Ce paramètre optionnel n'a pas de valeur par défaut.

`outTS` = *ensemble-de-données-SAS*

Nom du fichier de données de sortie (équilibrées). Ce fichier est identique au fichier de données d'entrée à l'exception des valeurs initiales (non équilibrées) des séries chronologiques qui ont été remplacées par les valeurs équilibrées. C'est le fichier de sortie principal. Il est toujours créé par la macro.

Ce paramètre est obligatoire.

`periodInterval` = *intervalle de date, temps ou date-temps SAS*

Intervalle de date, temps ou date-temps SAS associé aux périodes du fichier de données d'entrée. Le paramètre `timeVarName` doit être spécifié. Si ce paramètre n'est pas spécifié mais qu'une variable de temps (paramètre `timeVarName`) l'est, la macro détermine cet intervalle à l'aide de la fonction `INTGET` de SAS en se basant sur les trois premières périodes (non manquantes) du fichier de données d'entrée.

Voir la section « Détails » pour de plus amples renseignements sur le concept de temps dans la macro.

Ce paramètre optionnel n'a pas de valeur par défaut.

`tallTSAlterVarName` = *nom-de-variable-SAS*

Nom de la variable du fichier de données d'entrée contenant les coefficients d'altérabilité des séries chronologiques en mode *tall* (paramètre `TSFormat=TALL`). Ce paramètre est ignoré lorsque les données sont en mode *wide* (paramètre `TSFormat=WIDE`). Les coefficients d'altérabilité non manquants spécifiés le fichier de données d'entrée ont priorité sur ceux spécifiés avec les paramètres `alterPos`, `alterNeg` et `alterMix` ainsi que ceux fournis dans le fichier de spécification du problème.

Ce paramètre optionnel n'a pas de valeur par défaut (pas de variable de coef. d'altérabilité par défaut).

`tallTSIDVarName` = *nom-de-variable-SAS*

Nom de la variable du fichier de données d'entrée contenant les noms des séries chronologiques en mode *tall* (paramètre `TSFormat=TALL`). Ce paramètre est ignoré lorsque les données sont en mode *wide* (paramètre `TSFormat=WIDE`).

La valeur par défaut de ce paramètre optionnel est `_NAME_`.

`tallTSValueVarName` = *nom-de-variable-SAS*

Nom de la variable du fichier de données d'entrée contenant les valeurs des séries chronologiques en mode *tall* (paramètre `TSFormat=TALL`). Ce paramètre est ignoré lorsque les données sont en mode *wide* (paramètre `TSFormat=WIDE`).

La valeur par défaut de ce paramètre optionnel est `_VALUE_`.

`temporalGrpInterval` = *intervalle de date, temps ou date-temps SAS ou le mot réservé _ALL_*

Intervalle de date, temps ou date-temps SAS associé aux totaux temporels à préserver (groupe de périodes à traiter simultanément). À noter :

- Si ce paramètre n'est pas spécifié, les totaux temporels ne sont pas préservés et chaque période est équilibrée de façon indépendante (traitement période par période).
- Le mot réservé `_ALL_` peut être utilisé pour préserver les totaux temporels correspondant à la somme des valeurs pour toutes les périodes du fichier de données (traiter toutes les périodes simultanément). La valeur `_ALL_` est applicable également aux problèmes sans variable de temps.
- Les totaux temporels peuvent être préservés seulement si le fichier de données d'entrée contient plus d'une période et si la fréquence de l'intervalle associé aux groupes temporels est inférieure à celle des périodes (ex : `periodInterval=MONTH` et `temporalGrpInterval=YEAR`). Si le fichier de données d'entrée ne contient qu'une seule période ou si la fréquence de l'intervalle associé aux groupes temporels est supérieure ou égale à celle des périodes (ex : `periodInterval=QTR` et `temporalGrpInterval=MONTH` ou `=QTR`), le concept de total temporel ne s'applique pas et le traitement sera automatiquement effectué période par période.

Voir la section « Détails » pour de plus amples renseignements sur le concept de temps dans la macro.

Ce paramètre optionnel n'a pas de valeur par défaut (pas de préservation des totaux temporels par défaut).

`timeVarName` = *nom-de-variable-SAS*

Nom de la variable de temps (variable numérique de date, temps ou date-temps SAS) dans le fichier de données d'entrée. Si une variable de temps n'est pas spécifiée, chaque période du fichier de données d'entrée sera équilibrée de façon indépendante, sauf si `temporalGrpInterval=_ALL_` (voir la description du paramètre `temporalGrpInterval`).

Voir la section « Détails » pour de plus amples renseignements sur le concept de temps dans la macro.

Ce paramètre optionnel n'a pas de valeur par défaut (pas de variable de temps par défaut).

toleranceValue = *nombre réel positif*

Seuil de tolérance pour les contraintes d'équilibrage. Quand une tolérance est spécifiée,

Les contraintes EQ	$G\theta = b$	deviennent	$b - \text{toleranceValue} \leq G\theta \leq b + \text{toleranceValue}$
Les contraintes GE	$G\theta \geq b$	deviennent	$G\theta \geq b - \text{toleranceValue}$
Les contraintes LE	$G\theta \leq b$	deviennent	$G\theta \leq b + \text{toleranceValue}$

Ce paramètre optionnel n'a pas de valeur par défaut (pas de tolérance par défaut).

toleranceValueTemporal = *nombre réel positif*

tolerancePercentTemporal = *nombre réel positif*

Seuil de tolérance pour les contraintes temporelles implicites correspondant aux totaux temporels contraignants, spécifié soit en valeur absolue ou en pourcentage. Les paramètres toleranceValueTemporal et tolerancePercentTemporal ne peuvent être spécifiés en même temps. Ils ne s'appliquent pas aux (sont ignorés pour les) totaux temporels non contraignants.

Ces paramètres modifient les contraintes temporelles implicites $\sum_{t \in T_j} \theta_{kt}^{(x)} = a_{kj}$ comme suit, où le total temporel initial a_{kj} est contraignant ($|c_{kj}^{(a)} a_{kj}| = 0$) et j désigne un groupe temporel complet:

$$a_{kj} - \text{toleranceValueTemporal} \leq \sum_{t \in T_j} \theta_{kt}^{(x)} \leq a_{kj} + \text{toleranceValueTemporal}$$

ou

$$a_{kj} (1 - \text{tolerancePercentTemporal}) \leq \sum_{t \in T_j} \theta_{kt}^{(x)} \leq a_{kj} (1 + \text{tolerancePercentTemporal})$$

Ces paramètres optionnels n'ont pas de valeur par défaut (pas de tolérance par défaut).

TSFormat = WIDE ou TALL

Définit le format des fichiers de données d'entrée et de sortie (paramètres inTS et outTS). Le format WIDE signifie que les données des séries chronologiques apparaissent dans des colonnes (variables) séparées dans le fichier de données (une colonne par série chronologique). Avec le format TALL, les données des séries chronologiques sont empilées dans une seule colonne (variable). Les paramètres tallTSIDVarName, tallTSValueVarName et tallTSAlterVarName permettent de préciser le rôle des variables dans le fichier de données d'entrée lorsque le format TALL est spécifié.

Ce paramètre est optionnel et sa valeur par défaut est WIDE.

upperBound = *nombre réel*

Borne supérieure pour les valeurs des séries chronologiques. Si non spécifiée, la borne supérieure est ∞ . Les bornes supérieures spécifiées dans le fichier de spécification du problème ont priorité sur cette valeur.

Ce paramètre optionnel n'a pas de valeur par défaut.

Détails

- La macro n'est pas sensible à la casse en ce qui concerne l'information fournie en entrée. Ceci s'applique autant à la valeur des paramètres de la macro qu'à l'information fournie dans le fichier de spécification du problème et dans le fichier de données d'entrée (ex : nom des séries chronologiques en mode *tall*).
- Le tableau suivant présente l'ordre de priorité en ce qui concerne les différentes façons de spécifier des coefficients d'altérabilité ainsi que des bornes inférieures et supérieures :

1	Coefficients d'altérabilité non manquants fournis dans le fichier de données d'entrée en mode <i>tall</i> (voir le paramètre <code>tallTSAAlterVarName</code>).
2	Coefficients d'altérabilité et bornes inférieures/supérieures datés (valeur non manquante pour <code>_timeVal_</code>) spécifiés dans le fichier de spécification du problème.
3	Coefficients d'altérabilité et bornes inférieures/supérieures non datés (valeur manquante pour <code>_timeVal_</code>) spécifiés dans le fichier de spécification du problème.
4	Coefficients d'altérabilité et bornes inférieures/supérieures définis par les paramètres <code>alterPos</code> , <code>alterNeg</code> , <code>alterMix</code> , <code>alterTemporalTotals</code> , <code>lowerBound</code> et <code>upperBound</code> .

- La notion de temps est un concept clé dans la macro **GSeriesTSBalancing**, tout particulièrement pour la préservation des totaux temporels. Heureusement ce concept est solidement implanté en SAS et la macro tire avantage des puissants et versatiles intervalles et fonctions de date et de temps qui sont disponibles en SAS. Trois types d'éléments reliés à la notion de temps sont utilisés dans la macro :
 - Périodes** : valeurs de temps de haute fréquence qui correspondent à la périodicité des séries chronologiques à équilibrer (ex : mois, trimestres). Elles correspondent aux enregistrements du fichier de données d'entrée en mode *wide*.
 - Groupes temporels** : valeurs de temps de basse fréquence associées aux totaux temporels à préserver (ex : années). Un groupe temporel donné est dit *complet* lorsque les données d'entrée contiennent toutes les périodes appartenant au groupe temporel (ex : années pour lesquelles les 12 mois sont tous présents dans les données) ou *incomplet* dans le cas contraire (ex : année en cours pour laquelle des données sont disponibles jusqu'en octobre seulement). La *cardinalité* d'un groupe temporel complet correspond au nombre de périodes qu'il contient (ex : 12 et 4 lorsqu'il s'agit d'années avec des données mensuelles et trimestrielles respectivement). Des **totaux temporels** sont définis pour chaque série chronologique et chaque *groupe temporel complet* impliqués dans le problème d'équilibrage. Un total temporel donné correspond à la somme des valeurs d'une série chronologique donnée à travers toutes les périodes d'un groupe temporel complet donné.
 - Groupes de traitement** : ensembles de périodes qui doivent être traitées (optimisées) simultanément avec la procédure OPTMODEL. Un groupe de traitement correspond soit à une période unique (traitement période par période) ou bien à l'ensemble des périodes qui composent un groupe temporel complet lorsque les totaux temporels sont préservés.

Trois paramètres de la macro **GSeriesTSBalancing** permettent de spécifier le type de traitement temporel désiré : `timeVarName`, `periodInterval`, et `temporalGrpInterval`. Le tableau suivant résume les scénarios possibles lorsque les données d'entrée contiennent plus d'une période :

Le problème comprend une var. de temps	Intervalle associé aux groupes temporels (valeur de <code>temporalGrpInterval</code>)	Scénario de traitement temporel	Nombre de groupes de traitement
Non	<i>Non spécifié</i>	Période par période	Nombre de périodes

	<code>_ALL_</code>	Toutes les périodes simultanément (préservation de totaux temporels)	1
Oui	<i>Non spécifié</i>	Période par période	Nombre de périodes
	<code>_ALL_</code>	Toutes les périodes simultanément (préservation de totaux temporels)	1
	Fréquence plus basse que <code>periodInterval</code> (ex : <code>periodInterval=MONTH</code> <code>temporalGrpInterval=YEAR</code>)	<ul style="list-style-type: none"> Dans les groupes temp. complets : toutes les périodes simultanément (préservation de totaux temporels) Dans les groupes temp. incomplets : période par période 	Nombre de groupes temporels complets + Nombre de périodes dans les groupes temp. incomplets
	Fréquence plus haute que <code>periodInterval</code> (ex : <code>periodInterval=QTR</code> <code>temporalGrpInterval=MONTH</code>)	Les intervalles spécifiés ne sont pas compatibles avec la notion de préservation de totaux temporels, ce qui se traduit par un traitement période par période .	Nombre de périodes
	Même valeur que <code>periodInterval</code>		
	Incompatible avec <code>periodInterval</code> (ex : <code>periodInterval=WEEK</code> <code>temporalGrpInterval=MONTH</code>)		

La documentation SAS énumère tous les intervalles de date, temps et date-temps qui sont disponibles.

- La variable macro `_GSERIESRC_` contient le code retour de la macro suite à son exécution : « 0 » lors d'un succès et « VA### » lors d'un échec, où « VA### » réfère à un message d'erreur de G-Series.

Comparaison avec PROC TSRaking

- PROC TSRaking** réconcilie des tableaux à 1 et 2 dimensions (avec préservation de totaux temporels si nécessaire) alors que la macro **GSeriesTSBalancing** peut traiter des problèmes d'équilibrage plus généraux (ex : réconciliation de tableaux à plus de 2 dimensions, solutions non négatives, contraintes linéaires générales d'égalité et d'inégalité par opposition à de simples règles d'agrégation, etc.).
- Bien que les deux outils permettent la préservation de totaux temporels, le traitement de la notion de temps n'est pas incorporé dans **PROC TSRaking**. Ceci signifie que l'utilisateur doit lui-même construire les groupes de traitement et effectuer les appels séparés à **TSRaking** pour chacun des groupes de traitement.
- GSeriesTSBalancing** permet la spécification de problèmes creux dans leur forme réduite, ce qui n'est pas le cas de **TSRaking** pour lequel les règles d'agrégation doivent toujours être pleinement spécifiées.
- Les valeurs négatives ne sont pas traitées de façon identique par les deux outils. Alors que les solutions produites par **GSeriesTSBalancing** et **TSRaking** sont identiques lorsque les données d'entrées sont toutes positives, elles seront différentes si certaines de ces données sont négatives.
- La macro **GSeriesTSBalancing** est habituellement plus efficace que **PROC TSRaking** mais est généralement plus sensible aux légères incohérences qui peuvent parfois exister dans des problèmes de réconciliation de tableaux avec préservation de totaux temporels pleinement spécifiés (sur-spécifiés).

Exemple 1

Réconciliation d'un tableau à une dimension

Dans cet exemple, nous disposons des estimations des ventes de voitures (*cars*) pour les régions *alb*, *sask* et *man* ainsi que des ventes totales (*tot*), pour chacun des trois premiers trimestres de 2011 (pas de préservation des totaux annuels). Les ventes totales sont contraignantes (spécifié avec le paramètre `alterNeg=0`). La contrainte est

- $\text{cars_alb} + \text{cars_sask} + \text{cars_man} = \text{cars_tot}$.

Le fichier de spécification du problème, le fichier de données d'entrée et l'appel macro sont

```
data myProblem;
  length _type_ $2 _col_ $9 _row_ $38;
  input _type_ = _col_ = _row_ = _coef_ =;
  datalines;
_type_=eq      _col_=      _row_=Agrégation - total des voitures (cars)      _coef_=
_type_=      _col_=cars_alb      _row_=Agrégation - total des voitures (cars)      _coef_=1
_type_=      _col_=cars_sask      _row_=Agrégation - total des voitures (cars)      _coef_=1
_type_=      _col_=cars_man      _row_=Agrégation - total des voitures (cars)      _coef_=1
_type_=      _col_=cars_tot      _row_=Agrégation - total des voitures (cars)      _coef_=-1
;

data myData;
  input date cars_alb cars_sask cars_man cars_tot;
  informat date yyq6.;
  format date yyq6.;
  datalines;
2011q1 20 18 12 53
2011q2 16 16 19 44
2011q3 14 15 16 50
;

%GSeriesTSBalancing
(
  inTS          = myData,
  inProblemSpecs = myProblem,
  outTS         = outBalanced,
  timeVarName   = date,
  alterNeg      = 0, /* ventes totales contraignantes (CARS_TOT) */
  language      = FR /* afficher les messages du journal (log) en français */
)
```

Exemple 2

Réconciliation d'un tableau à deux dimensions avec contraintes de non négativité

Dans cet exemple, nous disposons des estimations des ventes de voitures (*cars*) et de fourgonnettes (*vans*) pour les régions *alb*, *sask* et *man* du deuxième trimestre de 2010 jusqu'au premier trimestre de 2011 et nous voulons préserver les totaux annuels pour 2011. Les totaux des marges du tableau sont contraignants de même que les totaux annuels. Les contraintes d'équilibrage sont

- $\text{cars_alb} + \text{cars_sask} + \text{cars_man} = \text{cars_tot}$;

- $\text{vans_alb} + \text{vans_sask} + \text{vans_man} = \text{vans_tot}$;
- $\text{cars_alb} + \text{vans_alb} = \text{alb_tot}$;
- $\text{cars_sask} + \text{vans_sask} = \text{sask_tot}$;
- $\text{cars_man} + \text{vans_man} = \text{man_tot}$.
- cars_alb , cars_sask , cars_man , vans_alb , vans_sask , vans_man , cars_tot , vans_tot , alb_tot , sask_tot et man_tot doivent être ≥ 0 .

Un coefficient d'altérabilité de 0 est utilisé pour *vans_sask* pour le deuxième trimestre de 2010 et la valeur par défaut 1 est utilisée pour toutes les autres périodes. La non négativité des séries chronologiques est spécifiée avec le paramètre `lowerBound=0`. Alors que les totaux annuels sont contraignants par défaut, les totaux des marges contraignants sont spécifiés avec le paramètre `alterNeg=0`. Le niveau d'information à afficher spécifié est 4 afin d'afficher les résultats d'équilibrage dans la fenêtre de sortie (*Output*). Trois fichiers de sortie optionnels sont demandés : les fichiers du *résumé d'OPTMODEL*, des *résultats d'équilibrage détaillés* et des *totaux temporels*.

Le fichier de spécification du problème, le fichier de données d'entrée et l'appel macro sont

```
data myProblem;
  length _type_ $5 _col_ $9 _row_ $43;
  input _type_ = _col_ = _row_ = _coef_ = _timeVal_ =;
  informat _timeVal_ yyq6.;
  format _timeVal_ yyq6.;
  datalines;
_type_=eq      _col_=      _row_=Agrégation - total des voitures (cars)      _coef_=      _timeVal_=
_type_=      _col_=cars_alb  _row_=Agrégation - total des voitures (cars)      _coef_=1      _timeVal_=
_type_=      _col_=cars_sask _row_=Agrégation - total des voitures (cars)      _coef_=1      _timeVal_=
_type_=      _col_=cars_man  _row_=Agrégation - total des voitures (cars)      _coef_=1      _timeVal_=
_type_=      _col_=cars_tot  _row_=Agrégation - total des voitures (cars)      _coef==1      _timeVal_=
_type_=eq      _col_=      _row_=Agrégation - total des fourgonnettes (vans) _coef_=      _timeVal_=
_type_=      _col_=vans_alb  _row_=Agrégation - total des fourgonnettes (vans) _coef_=1      _timeVal_=
_type_=      _col_=vans_sask _row_=Agrégation - total des fourgonnettes (vans) _coef_=1      _timeVal_=
_type_=      _col_=vans_man  _row_=Agrégation - total des fourgonnettes (vans) _coef_=1      _timeVal_=
_type_=      _col_=vans_tot  _row_=Agrégation - total des fourgonnettes (vans) _coef==1      _timeVal_=
_type_=eq      _col_=      _row_=Agrégation - total de l'Alberta      _coef_=      _timeVal_=
_type_=      _col_=cars_alb  _row_=Agrégation - total de l'Alberta      _coef_=1      _timeVal_=
_type_=      _col_=vans_alb  _row_=Agrégation - total de l'Alberta      _coef_=1      _timeVal_=
_type_=      _col_=alb_tot   _row_=Agrégation - total de l'Alberta      _coef==1      _timeVal_=
_type_=eq      _col_=      _row_=Agrégation - total de la Saskatchewan _coef_=      _timeVal_=
_type_=      _col_=cars_sask _row_=Agrégation - total de la Saskatchewan _coef_=1      _timeVal_=
_type_=      _col_=vans_sask _row_=Agrégation - total de la Saskatchewan _coef_=1      _timeVal_=
_type_=      _col_=sask_tot  _row_=Agrégation - total de la Saskatchewan _coef==1      _timeVal_=
_type_=eq      _col_=      _row_=Agrégation - total du Manitoba      _coef_=      _timeVal_=
_type_=      _col_=cars_man  _row_=Agrégation - total du Manitoba      _coef_=1      _timeVal_=
_type_=      _col_=vans_man  _row_=Agrégation - total du Manitoba      _coef_=1      _timeVal_=
_type_=      _col_=man_tot   _row_=Agrégation - total du Manitoba      _coef==1      _timeVal_=
_type_=alter    _col_=      _row_=Coefficient d'altérabilité      _coef_=      _timeVal_=
_type_=      _col_=vans_sask _row_=Coefficient d'altérabilité      _coef_=0      _timeVal_=2010q2
;

data myData;
  input date cars_alb cars_sask cars_man vans_alb vans_sask vans_man alb_tot sask_tot
        man_tot cars_tot vans_tot;
  informat date yyq6.;
  format date yyq6.;
  datalines;
2010q2 14 18 14 20 21 27 32 38 43 58 55
2010q3 17 14 16 29 25 26 41 35 39 44 71
2010q4 14 19 18 20 28 27 41 40 45 58 68
2011q1 20 18 12 20 22 26 32 39 43 53 61
2011q2 16 16 19 21 26 21 40 30 33 44 59
2011q3 14 15 16 19 25 19 34 47 40 50 71
2011q4 19 20 14 21 18 27 44 44 38 52 74
2012q1 16 15 19 27 25 28 35 38 32 51 54
;

```



```
%GSeriesTSBalancing
(
    inTS                = myData,
    inProblemSpecs      = myProblem,
    outTS               = outBalanced,
    timeVarName         = date,
    periodInterval      = QTR,
    temporalGrpInterval = YEAR, /* préserver les totaux annuels */
    displayLevel        = 4,    /* afficher les résultats d'équilibrage */
    outOptModelSummary  = outSummary,
    outDetailedResults  = outResults,
    outTemporalTotals   = outAnnTotals,
    alterNeg            = 0,    /* totaux des marges contraignants */
    lowerBound          = 0,    /* contraintes de non négativité */
    language            = FR    /* afficher les messages du journal (log) en français */
)
```

Annexe – Formulation mathématique du problème d'équilibrage

Notation

Symbole	Description
P	Nombre total de problèmes d'optimisation indépendants (groupes de traitement) à résoudre avec la procédure OPTMODEL afin d'équilibrer le système de séries chronologiques. L'ensemble des périodes incluses dans le groupe de traitement j est donné par T_j (voir ci-dessous).
K	Nombre total de séries chronologiques (variables du problème) dans le système.
T	Nombre total de périodes dans le système de séries chronologiques.
T_j	Représente la partition de la dimension temporelle du système de séries chronologiques. Lorsque les totaux temporels sont préservés, chaque T_j correspond soit à l'ensemble des périodes d'un groupe temporel complet (ex : toutes les périodes d'une année complète) soit à une période spécifique pour les groupes temporels incomplets (ex : les périodes individuelles des années incomplètes). Sinon, quand les totaux temporels ne sont pas préservés, T_j est équivalent à une période spécifique.
x_{kt}	Valeurs initiales (non équilibrées) des séries chronologiques, $k = 1, \dots, K$; $t = 1, \dots, T$.
$\theta_{kt}^{(x)}$	Valeurs équilibrées des séries chronologiques, $k = 1, \dots, K$; $t = 1, \dots, T$.
$c_{kt}^{(x)}$	Coefficient d'altérabilité associé à x_{kt} (coefficients d'altérabilité réguliers).
a_{kj}	Totaux temporels initiaux pour la série chronologique x_k , $k = 1, \dots, K$; j appartenant à l'ensemble des groupes de traitement qui correspondent aux groupes temporels complets (ex : années complètes).
$\theta_{kj}^{(a)}$	Totaux temporels finaux pour la série chronologique x_k , $k = 1, \dots, K$; j appartenant à l'ensemble des groupes de traitement qui correspondent aux groupes temporels complets (ex : années complètes).
$c_{kj}^{(a)}$	Coefficient d'altérabilité associé à a_{kj} (coefficients d'altérabilités des totaux temporels).
$I^{(eq)}$	Nombre total de contraintes linéaires d'égalité.
$b_i^{(eq)}$	Valeur du membre de droite (constante) de la contrainte linéaire d'égalité i , $i = 1, \dots, I^{(eq)}$.

Symbole	Description
$G_{ik}^{(eq)}$	Coefficient de la série chronologique x_{kt} dans la contrainte d'égalité i . La valeur est la même pour chaque période t , $t = 1, \dots, T$.
$I^{(le)}$	Nombre total de contraintes linéaires d'inégalité de type <i>inférieur ou égal à</i> .
$b_i^{(le)}$	Valeur du membre de droite (constante) de la contrainte linéaire d'inégalité i de type <i>inférieur ou égal à</i> , $i = 1, \dots, I^{(le)}$.
$G_{ik}^{(le)}$	Coefficient de la série chronologique x_{kt} dans la contrainte d'inégalité i de type <i>inférieur ou égal à</i> . La valeur est la même pour chaque période t , $t = 1, \dots, T$.
$I^{(ge)}$	Nombre total de contraintes linéaires d'inégalité de type <i>supérieur ou égal à</i> .
$b_i^{(ge)}$	Valeur du membre de droite (constante) de la contrainte linéaire d'inégalité i de type <i>supérieur ou égal à</i> , $i = 1, \dots, I^{(ge)}$.
$G_{ik}^{(ge)}$	Coefficient de la série chronologique x_{kt} dans la contrainte d'inégalité i de type <i>supérieur ou égal à</i> . La valeur est la même pour chaque période t , $t = 1, \dots, T$.

Problème de minimisation

Équilibrer un système de séries chronologiques avec des données couvrant T périodes, $\{1, 2, \dots, T\} = \bigcup_j T_j$, peut être formulé comme une séquence de P problèmes de minimisation quadratique à résoudre de façon indépendante. En utilisant la notation ci-dessus, pour chaque groupe de traitement j , $j = 1, \dots, P$, le problème de minimisation est défini comme suit.

Si j correspond à un groupe temporel complet ($\text{card}(T_j) > 1$) :

$$\min_{\theta} \sum_{k=1}^K \sum_{t \in T_j} \frac{(x_{kt} - \theta_{kt}^{(x)})^2}{s_{kt}^{(x)}} + \sum_{k=1}^K \frac{(a_{kj} - \theta_{kj}^{(a)})^2}{s_{kj}^{(a)}},$$

$$\text{où } s_{kt}^{(x)} = \begin{cases} 1 & \text{si } |c_{kt}^{(x)} x_{kt}| = 0 \\ |c_{kt}^{(x)} x_{kt}| & \text{sinon} \end{cases}, \quad a_{kj} = \sum_{t \in T_j} x_{kt}, \quad s_{kj}^{(a)} = \begin{cases} 1 & \text{si } |c_{kj}^{(a)} a_{kj}| = 0 \\ |c_{kj}^{(a)} a_{kj}| & \text{sinon} \end{cases}$$

sujet à :

$$\left. \begin{aligned} \sum_{k=1}^K G_{ik}^{(eq)} \theta_{kt}^{(x)} &= b_i^{(eq)} & i = 1, \dots, I^{(eq)}, t \in T_j \\ \sum_{k=1}^K G_{ik}^{(le)} \theta_{kt}^{(x)} &\leq b_i^{(le)} & i = 1, \dots, I^{(le)}, t \in T_j \\ \sum_{k=1}^K G_{ik}^{(ge)} \theta_{kt}^{(x)} &\geq b_i^{(ge)} & i = 1, \dots, I^{(ge)}, t \in T_j \end{aligned} \right\} \text{équivalent à } G\theta \text{ OP } b \text{ où OP est } =, \leq \text{ ou } \geq.$$

$$\begin{aligned} \theta_{kt}^{(x)} &= x_{kt} & \text{si } |c_{kt}^{(x)} x_{kt}| = 0, t \in T_j \\ \sum_{t \in T_j} \theta_{kt}^{(x)} &= \theta_{kj}^{(a)} & k = 1, \dots, K \\ \theta_{kj}^{(a)} &= a_{kj} & \text{si } |c_{kj}^{(a)} a_{kj}| = 0, k = 1, \dots, K \end{aligned}$$

Sinon, si j correspond à une période unique ($\text{card}(T_j) = 1$) :

$$\min_{\theta} \sum_{k=1}^K \sum_{t \in T_j} \frac{(x_{kt} - \theta_{kt}^{(x)})^2}{s_{kt}^{(x)}}, \quad \text{où } s_{kt}^{(x)} = \begin{cases} 1 & \text{si } |c_{kt}^{(x)} x_{kt}| = 0 \\ |c_{kt}^{(x)} x_{kt}| & \text{sinon} \end{cases}$$

sujet à :

$$\left. \begin{aligned} \sum_{k=1}^K G_{ik}^{(eq)} \theta_{kt}^{(x)} &= b_i^{(eq)} & i = 1, \dots, I^{(eq)}, t \in T_j \\ \sum_{k=1}^K G_{ik}^{(le)} \theta_{kt}^{(x)} &\leq b_i^{(le)} & i = 1, \dots, I^{(le)}, t \in T_j \\ \sum_{k=1}^K G_{ik}^{(ge)} \theta_{kt}^{(x)} &\geq b_i^{(ge)} & i = 1, \dots, I^{(ge)}, t \in T_j \end{aligned} \right\} \text{équivalent à } G\theta \text{ OP } b \text{ où OP est } =, \leq \text{ ou } \geq.$$

$$\theta_{kt}^{(x)} = x_{kt} \quad \text{si } |c_{kt}^{(x)} x_{kt}| = 0, t \in T_j$$

Interprétation

L'idée générale consiste à minimiser la distance entre les valeurs initiales des séries chronologiques, notées x_{kt} dans le problème de minimisation, et les valeurs équilibrées (finales) $\theta_{kt}^{(x)}$ le tout sujet aux deux types de contraintes suivants :

- **Contraintes d'équilibrage** : contraintes linéaires multivariées fournies par l'utilisateur et définissant les relations devant être satisfaites (après l'équilibrage) entre les différentes séries chronologiques, pour chaque période individuelle du problème de minimisation (pour chaque $t, t = 1, \dots, T$). Ce sont les contraintes $G\theta \text{ OP } b$, où OP est l'opérateur ($=, \leq$ ou \geq). On les appelle parfois contraintes *transversales* ou *contemporaines*.
- **Contraintes temporelles implicites** : contraintes d'agrégation temporelle univariées automatiquement ajoutées au problème de minimisation quand les totaux temporels sont préservés. Des contraintes temporelles sont définies pour chaque série chronologique du problème de minimisation et imposent que la somme des valeurs équilibrées $\theta_{kt}^{(x)}$ d'une série chronologique k pour les périodes t d'un groupe temporel complet j (pour tout $t \in T_j$) soit égale au total temporel $\theta_{kj}^{(a)}$. Ce sont les contraintes $\sum_{t \in T_j} \theta_{kt}^{(x)} = \theta_{kj}^{(a)}$, où la distance entre le total temporel initial (a_{kj}) et final ($\theta_{kj}^{(a)}$) de la série chronologique k pour le groupe temporel complet j est minimisée dans la fonction objective. Ces contraintes temporelles implicites *n'existent pas pour les groupes temporels incomplets ou quand les totaux temporels ne sont pas préservés*.

Il y a autant de problèmes de minimisation indépendants à résoudre qu'il y a de groupes de traitement j à équilibrer. Le nombre total de groupes de traitement P dépend du nombre de périodes fourni dans l'ensemble de données d'entrée et aussi de la nécessité (ou non) de préserver les totaux temporels.

Les **coefficients d'altérabilité** sont des nombres non négatifs changeant le coût relatif de modifier une valeur initiale. En changeant la fonction objective à minimiser, ils permettent la production d'un large éventail de solutions. Comme pour les contraintes, il y a deux types de coefficients d'altérabilité : les *coefficients d'altérabilité réguliers* ($c_{kt}^{(x)}$) et les *coefficients d'altérabilité des totaux temporels* ($c_{kj}^{(a)}$). Comme les coefficients d'altérabilité se trouvent au dénominateur des mesures de distance de la fonction objective, plus le coefficient est grand, moins il est coûteux de modifier l'élément correspondant (c.-à-d. la valeur ou le total temporel d'une série chronologique). Réciproquement, plus le coefficient est petit plus le coût est grand. Il en découle que les éléments avec les plus grands coefficients d'altérabilité changeront proportionnellement davantage que ceux qui ont les plus petits coefficients d'altérabilité. Dès qu'un coefficient d'altérabilité est nul (0), une contrainte supplémentaire sur l'élément concerné est ajoutée au problème (contraintes $\theta_{kt}^{(x)} = x_{kt}$ et $\theta_{kj}^{(a)} = a_{kj}$) et celui-ci est dit *contraignant*, référant au fait qu'il ne peut être modifié. Inversement, l'élément est dit *non contraignant* quand son coefficient d'altérabilité n'est pas nul (différent de 0). Les totaux temporels des séries chronologiques sont habituellement contraignants, de même que les valeurs des séries chronologiques correspondant aux totaux des marges dans les problèmes de réconciliation de tableaux (*raking problems*), c.-à-d. les séries chronologiques avec un coefficient de -1 dans les contraintes d'agrégation d'équilibrage. À l'opposé, les valeurs des séries composantes dans les problèmes de réconciliation de tableaux (c.-à-d. les séries chronologiques avec un coefficient de 1 dans les contraintes d'agrégation d'équilibrage) sont généralement non contraignantes. En pratique, un élément *quasi contraignant* peut être obtenu en spécifiant un très petit coefficient d'altérabilité (presque 0). Cette approche est parfois utile afin d'éviter les problèmes irréalisables pouvant découler de légères incohérences dans les données

(ex : problèmes de réconciliation de tableaux à plusieurs dimensions entièrement spécifiés avec préservation des totaux temporels).

La **préservation des totaux temporels** fait référence au fait que nous essayons toujours de demeurer « aussi près que possible » des totaux temporels initiaux. La *préservation pure* est atteinte pour les totaux temporels contraignants tandis que le changement est minimisé pour les totaux temporels non contraignants.

Notes

Ce document est un guide pour l'utilisation de la macro **GSeriesTSBalancing**. Pour plus de renseignements, veuillez contacter l'équipe de soutien de G-Series à l'adresse courriel [G-Series](mailto:G-Series@statcan.gc.ca) (G-Series@statcan.gc.ca) ou consultez le site web [G-Series](#) (sur l'intranet de Statistique Canada seulement).

Bibliographie

Bikker R., Daalmans J. et Mushkudiani N. (2013). Benchmarking large accounting frameworks: A generalized multivariate model. *Economic Systems Research*, DOI:10.1080/09535314.2013.801010.

Bérubé J. et Fortier S. (2009). PROC TSRAKING: An in-house SAS® procedure for balancing time series. *JSM Proceedings, Business and Economic Section*. Alexandria, VA: American Statistical Association.
