

### Chapitre 3, exercice 3

#### Instructions pour employer CH03EX03.XLS et étudier les séries du répertoire CH03EX03

*Le fichier CH03EX03.XLS comporte un exercice de base destiné à tous les apprenants et un exercice avancé réservé aux seuls apprenants de la version avancée*

*Le répertoire CH03EX03 comporte un exercice de base destiné à tous les apprenants*

#### Exercice de base (Pour tous les utilisateurs du cours)

##### Préalable



Le chapitre 3 du cours de base doit avoir été suivi jusqu'à la page 60 pour la partie 1, jusqu'à la page 73 pour la partie 2, jusqu'à la page 73 pour la partie 3, et jusqu'à la page 76 pour la partie 4.

##### Objectif



Le but est d'introduire une méthode simple, la méthode des trois points, déjà rencontrée dans le chapitre 2 pour la régression linéaire simple. Ces méthodes utilisent la médiane revue dans le chapitre 2.

##### Données



Les données sont des données macro-économiques: le produit intérieur brut (en milliards d'euros) pour les années 1972 à 1979. Deux pays sont considérés parce qu'ils correspondent à des situations différentes: la Belgique et la République Fédérale d'Allemagne (R.F.A.). La période a été choisie à cheval sur la première crise pétrolière 1973-1974 et avant la seconde 1980 parce que cette période a été marquée par une récession qui a ralenti la croissance de l'activité économique, de sorte que la considération de courbes de croissance avec plafond a été raisonnable durant cette période.

### *Structure de l'exercice*

L'exercice comporte quatre parties :

- Dans la partie 1, le but de l'exercice est d'employer Microsoft Excel pour réaliser des ajustements par quelques courbes de croissance : la tendance linéaire, la tendance parabolique et la tendance exponentielle.
- Dans la partie 2, le but de l'exercice est de décrire le principe de la méthode jusqu'à la phase de résolution du système d'équations non linéaire, non comprise.
- Dans la partie 3, le but de l'exercice est d'examiner les résultats relatifs à la qualité de l'ajustement de la méthode de la partie 1 et les aspects de résistance de la méthode des trois points, tout au moins quand on emploie la médiane.
- Dans la partie 4, le but de l'exercice est d'appliquer la méthode des trois points au moyen du logiciel Time Series Expert for Windows.

**Partie 1** Dans cette partie, le but de l'exercice est d'introduire les données et d'employer Microsoft Excel pour réaliser des ajustements par quelques courbes de croissance : la tendance linéaire, la tendance parabolique et la tendance exponentielle.

### 1.1 PRISE DE CONNAISSANCE AVEC LES DONNEES

- ⇒ Les données sont présentées dans un tableau (cliquez sur la zone marquée "Données") accompagnées de leurs logarithmes.
- ⇒ Consultez le graphique en échelle métrique (cliquez sur l'onglet "Metric").



Lequel des deux pays semble être le plus prospère?



*1.1.1 Votre réponse*

- ⇒ Regardez maintenant le graphique en échelle logarithmique (c'est-à-dire que les logarithmes du produit intérieur brut est représenté en fonction du temps pour chacun des deux pays). A cette fin cliquez sur l'onglet "Logarithmic".



A quelle courbe de croissance fait penser le fait que, pour les deux pays, la trajectoire en échelle logarithmique soit à peu près une ligne droite?



*1.1.2 Votre réponse*



Un des pays domine-t-il réellement l'autre autant qu'il apparaissait dans la figure en échelle métrique?



Peut-on dire que les deux pays ont connu à peu près le même taux de croissance durant la période considérée?



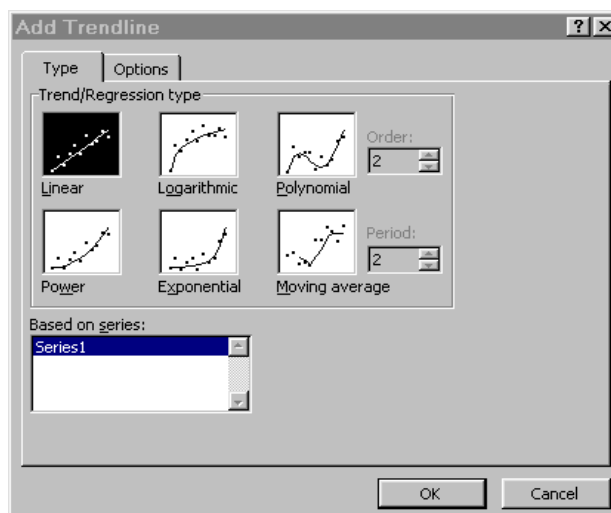
### 1.1.3 Vos réponses

## 1.2 L'AJUSTEMENT DES DONNEES EN ECHELLE METRIQUE

- ⇒ Dans la feuille Main, descendez vers la ligne 31 afin de voir le tableau des données. Vous pouvez aussi cliquer sur la zone “Données” en haut de la feuille.
- ⇒ Cliquez ensuite sur l’onglet Metric afin de visualiser le graphique des données. Cliquez sur une des série, pare exemple la série du bas qui est celle pour la Belgique.

Nous allons procéder comme dans l’exercice 3 du chapitre 2, en commençant par une fonction quadratique, ou parabolique, c’est-à-dire un polynôme de degré 2.

- ⇒ Cliquez sur le menu Chart ⇒ Add Trendline.



- ⇒ Cliquez sur l’onglet Options de la boîte de dialogue et cochez les options Display equation on chart et Display R-squared value on chart.
- ⇒ Cliquez sur l’onglet Type puis sur le type Polynomial.

**?**

Notez les informations qui se sont ajoutées au graphique. Il s'agit de l'équation de la fonction quadratique de régression et de la valeur du coefficient de détermination  $R^2$ .

*1.2.1 Votre réponse*

Essayons maintenant un ajustement par une exponentielle.



Cliquez sur l'onglet Type puis sur le type Exponential.

**?**

Notez les informations qui se sont ajoutées au graphique. Il s'agit de l'équation de la fonction exponentielle de régression et de la valeur du coefficient de détermination  $R^2$ .

*1.2.2 Votre réponse*



### 1.3 L'AJUSTEMENT DES DONNEES EN ECHELLE LOGARITHMIQUE



Cliquez ensuite sur l'onglet Logarithmic afin de visualiser le graphique des données en échelle logarithmique. Cliquez encore une fois sur la série du bas, relative à la Belgique.



Cliquez sur le menu Chart ⇒ Add Trendline.



Cliquez sur l'onglet Options de la boîte de dialogue et cochez les options Display equation on chart et Display R-squared value on chart.



Cliquez sur l'onglet Type puis sur le type Linear.

**?**

Notez les informations qui se sont ajoutées au graphique. Il s'agit de l'équation de la fonction linéaire de régression sur le logarithmes et de la valeur du coefficient de détermination  $R^2$ .



1.3.1 Votre réponse

Facultatif



Eventuellement après avoir pris connaissance de la remarque ci-dessous, essayez de réconcilier les résultats de la réponse 1.3.1 avec la réponse 1.2.2.



1.3.2 Votre réponse



### Remarques

1. Notons que la transformation logarithmique employé ici était en base 10 et que  $10^{0,0593x + 1,4558} = 10^{1,4558} 10^{0,0593x} = 28,563 \cdot 10^{0,0593x} = 28,563 \cdot e^{0,1365x}$  parce que le logarithme népérien de  $10^{0,0593} = 1,1463$  vaut 0,1365.

2. Le coefficient de détermination affiché dans le cas de la courbe exponentielle en échelle métrique est le même que pour la droite en échelle logarithmique. Ce n'est toutefois pas correct.

A la question 1.1.3, on demandait quel pays de la Belgique et de la R.F.A. avait le plus grand taux de croissance. Un taux de croissance est simplement le paramètre  $\beta$  de l'exponentielle – ici estimée par la méthode des moindres carrés – moins 1 et exprimé en pourcentage.



Quels sont donc les taux de croissance des deux pays ?



1.3.3 Votre réponse

## SYNTHESE

Dans cette partie, nous avons vu l'utilisation de quelques courbes de croissance au moyen des graphiques d'Excel et des possibilités d'ajustement qu'ils permettent. Notons qu'aucune de ces courbes ne traduit l'apparence de plafonnement qui est sensible dans le graphique des données.

**Partie 2** Dans cette partie, on voit le principe de la méthode jusqu'à la phase de résolution du système d'équations non linéaire, non comprise. La suite est traitée dans la partie avancée du cours. Les courbes de croissance considérées sont l'exponentielle (à l'aide de la méthode des moindres carrés sur les logarithmes des données), l'exponentielle modifiée, la courbe en S de Gompertz et la courbe logistique.

### 2.1 LES COURBES DE CROISSANCE A PLAFOND

**?** Si on regarde les déviations par rapport à la droite, lequel des deux pays semble avoir le plus tendance à un plafonnement durant la période?



#### 2.1.1 Vos réponses

### 2.2 LES COURBES DE CROISSANCE EXPONENTIELLE MODIFIEE, DE GOMPERTZ ET LOGISTIQUE

Le but est de déterminer les paramètres alpha ( $\alpha$ ), beta ( $\beta$ ) et gamma ( $\gamma$ ) à partir des données de manière à réaliser un ajustement de la courbe de croissance aux données.

Conformément à la théorie, on se ramène à l'étude de l'exponentielle modifiée :

- pour la courbe de Gompertz, on travaille avec les logarithmes népériens (ou en base e) des données et non plus en logarithmes décimaux (parce que la définition de la courbe comporte une exponentielle en base e) ;
- pour la courbe logistique, on travaille avec les inverses des données.

⇒ Revenez à la feuille de calcul. Les calculs sont présentés dans un tableau. Pour l'atteindre, pressez la touche F5 et sélectionnez TRANSFO ou cliquez au début de la zone marquée "Données transformées".

### 2.3 DEMARRAGE DE LA METHODE DES TROIS POINTS

Nous avons vu que la méthode des trois points appliquée dans le cas de la régression linéaire simple n'emploie en fait que deux points. La raison est

simple: une droite est déterminée par deux paramètres. Ces deux paramètres peuvent être la pente et l'ordonnée à l'origine. Les courbes de croissance considérées ici (sauf l'exponentielle, évidemment) dépendent toutes de trois paramètres: alpha ( $\alpha$ ), beta ( $\beta$ ) et gamma ( $\gamma$ ). On utilise donc trois points déterminés à partir des données pour définir la courbe. L'idée est que la courbe cherchée passe par les trois points. La méthode nous intéresse ici par sa simplicité mais aussi par sa relative résistance aux données aberrantes, conséquence de l'emploi comme paramètre de position de la médiane au lieu de la moyenne. Nous considérons en fait deux versions de la méthode des trois points: une qui utilise les médianes et une autre qui utilise les moyennes.

La méthode des trois points avec emploi de la médiane consiste en les étapes suivantes:

1) Répartir les données dans l'ordre du temps en trois sous-ensembles I, II et III, approximativement de même taille, en veillant à ce que les moyennes des temps  $t_I$ ,  $t_{II}$  et  $t_{III}$  soient équidistantes. Plus précisément :

(a) si  $T$  est multiple de 3,  $T = 3\tau$ , il n'y a pas de difficulté : les trois ensembles seront de taille  $T/3 = \tau$ ;

(b) si  $T$  est multiple de 3 plus 1,  $T = 3\tau + 1$ , mettre l'ensemble de taille  $\tau + 1$  en II et les deux ensembles de taille  $\tau$  en I et en III;

(c) si  $T$  est multiple de 3 plus 2,  $T = 3\tau + 2$ , mettre l'ensemble de taille  $\tau$  en II et les deux ensembles de taille  $\tau + 1$  en I et en III.



**Remarque.**

Notez la légère différence par rapport à la répartition basée sur les quantiles d'ordre 1/3 et 2/3 de  $t$ , utilisée au chapitre 2. Le fait que les moyennes des temps soient équidistantes intervient dans la partie A de l'exercice avancé.

2) déterminer trois points  $(t_I, y_I)$ ,  $(t_{II}, y_{II})$  et  $(t_{III}, y_{III})$  comme points *médians* des sous-ensembles I, II et III respectivement, c'est-à-dire en prenant comme abscisse la *médiane* des abscisses de ces points et comme ordonnée la *médiane* de leurs ordonnées;

3) exprimer que la courbe de croissance passe par les trois points ce qui peut s'exprimer en la résolution de trois équations (non linéaires parce que la fonction décrivant la courbe est non linéaire en les paramètres) à trois inconnues, les trois paramètres: alpha ( $\alpha$ ), beta ( $\beta$ ) et gamma ( $\gamma$ );

La méthode des trois points avec emploi de la moyenne consiste en les



mêmes étapes sauf que l'étape 2) est remplacée par la suivante:

2') déterminer trois points  $(t_I, y_I)$ ,  $(t_{II}, y_{II})$  et  $(t_{III}, y_{III})$  comme points *moyens* des sous-ensembles I, II et III respectivement, c'est-à-dire en prenant comme abscisse la *moyenne* des abscisses de ces points et comme ordonnée la *moyenne* de leurs ordonnées;

Il y a 8 années allant de 1972 à 1979 et donc  $t = 1, 2, \dots, 8$ . Les années sont de même longueur, en négligeant les années bissextiles, bien sûr. La variable étant une variable de flux, on place les dates de chaque année au milieu de ces années, donc au 1<sup>er</sup> juillet. Ces dates sont donc équidistantes.



Vérifiez que les quantiles d'ordre  $1/3$  et  $2/3$  valent bien  $t_{1/3} = 3$  et  $t_{2/3} = 6$ . Combien y a-t-il donc d'années dans chaque sous-ensemble I, II et III?

### 2.3.1 Vos réponses



Les calculs sont présentés dans un tableau. Pour atteindre ce tableau avec les médianes et les moyennes, pressez la touche F5 et sélectionnez MEDAVG ou cliquez au début de la zone marquée "Méthode des trois points".

Le tableau comporte quatre parties. Les deux pays, Belgique et R.F.A., sont côte à côte. Les parties du tableau employant les médianes sont au-dessus des parties du tableau employant les moyennes. Choisissez un des pays pour les médianes.



Suivez les trois premières étapes pour les médianes à l'aide des formules dans les différentes colonnes pour un des pays au choix. Arrêtez-vous avant la ligne marquée  $\ln(\beta)$ .



Vérifiez les valeurs de  $t_I$  (la médiane de 1, 2 et 3),  $t_{II}$  (la médiane de 4 et de 5), et  $t_{III}$  (la médiane de 6, 7 et 8).



## 2.3.2 Vos réponses



**?** Vérifiez les valeurs de  $y_I$ ,  $y_{II}$  et  $y_{III}$ , calculées par des médianes pour le pays choisi. Regardez également les formules employées dans Excel.

## 2.3.3 Vos réponses



**?** Vérifiez les valeurs de  $y_I$ ,  $y_{II}$  et  $y_{III}$ , calculées par des moyennes pour le pays choisi. Regardez également les formules employées dans Excel.

## 2.3.4 Vos réponses

La suite est l'objet du cours avancé parce que la résolution fait appel à des connaissances mathématiques. Contentez-vous de vérifier le résultat à savoir que la courbe de croissance dont les paramètres alpha ( $\alpha$ ), beta ( $\beta$ ) et gamma ( $\gamma$ ) sont spécifiés passe effectivement par les trois points.

⇒ A cette fin, ouvrez un nouveau classeur. Placez dans une colonne les valeurs de  $t$  : 2, 4, 5 et 7. Dans ce qui suit, il est important d'employer tous les chiffres donnés pour les paramètres parce qu'un arrondi à trois ou quatre décimales donnera des résultats faux. Dans une deuxième colonne, entrez la formule qui correspond à une des courbes, par exemple, pour l'ajustement par une exponentielle modifiée pour la Belgique : l'équivalent en Excel de  $\alpha\beta^t + \gamma$ .



**?** En remplaçant  $t$  par 2,  $\alpha$  par 177.3132181,  $\beta$  par 1.036758222 et  $\gamma$  par -153.8882353, que trouvez-vous ?



### 2.3.5 Votre réponse

#### Remarques

Dans Excel, le symbole ^ indique l'élévation à un exposant et EXP(...) est la fonction exponentielle.

### SYNTHESE

Dans cette partie, nous avons vu l'utilisation de trois courbes de croissance pour l'ajustement de données macro-économiques : le produit intérieur brut des années 1972-1979 pour deux pays. Pour la courbe de croissance exponentielle modifiée, nous avons estimé les trois paramètres par la méthode des trois points en utilisant la médiane ou la moyenne. L'ajustement par la courbe de Gompertz et la courbe logistique sont réalisés en appliquant la méthode des trois points, respectivement, sur les logarithmes et les inverses des données.

Les aspects de résistance aux données aberrantes seront vus dans la partie suivante en même temps que la qualité de l'ajustement.

**Partie 3** Dans cette partie, on examine les résultats et en particulier la qualité de l'ajustement, graphiquement et au moyen d'un critère, le critère MAPE, calculé sur l'ensemble des années. On examine également les aspects de résistance de la méthode des trois points quand on emploie la médiane au lieu de la moyenne. Les courbes de croissance considérées sont l'exponentielle (à l'aide de la méthode des moindres carrés sur les logarithmes des données), l'exponentielle modifiée, les courbes en S de Gompertz et logistique.

### 3.1 AJUSTEMENT PAR UNE COURBE EXPONENTIELLE

Nous avons déjà réalisé ceci dans la partie 1 de l'exercice. Nous allons le retrouver par une autre manière. Ici, l'ajustement par une exponentielle a été obtenu en employant la méthode des moindres carrés sur les logarithmes des données (voir chapitre 2 pour les détails) mais en employant les fonctions d'Excel.

⇒ Pour atteindre ces résultats, pressez la touche F5 et sélectionnez le champ REGRESS ou cliquez au début de la zone marquée "Régression". Notez que les coefficients pour la Belgique sont en colonne BA et pour la R.F.A. en colonne BF.

Les résultats sont donnés dans la partie inférieure du tableau, marquée Log – la partie supérieure se rapporte à un ajustement linéaire sur les données, sans intérêt ici. Les lignes où sont la constante et le coefficient de régression sont respectivement les lignes 12 et 18.

? Suivez les différentes étapes de calcul à l'aide des formules pour un des pays au choix. Cela semble-t-il correct ?



*3.1.1 Votre impression*

? Est-ce identique à ce que vous avez trouvé à la réponse 1.2.1?



*3.1.2 Votre réponse*

### 3.2 QUALITE DE L'AJUSTEMENT

⇒ Le résultat des ajustements par plusieurs courbes de croissance sont donnés dans un tableau. Pressez la touche F5 et sélectionnez ADJBELMED ou cliquez sur la zone marquée "Ajustement Belgique (méd)".

Notez que les données, inverses et logarithmes sont dans les colonnes CQ, CR et CS pour la Belgique, respectivement, et dans les colonnes CU, CV et CW pour la R.F.A. Les paramètres estimés sont dans l'ordre: beta ( $\beta$ ), alpha ( $\alpha$ ), et gamma ( $\gamma$ ). Ils sont sur les lignes 175 à 177, quand les médianes sont employées et sur les lignes 192 à 194, quand les moyennes sont employées.

? Examinez les formules des valeurs ajustées pour la Belgique. Cela semble-t-il correct?



#### 3.2.1 Votre avis

Les valeurs ajustées sont fournies dans chaque colonne. On peut d'ailleurs retrouver la valeur 36,7 pour la courbe exponentielle modifiée au temps  $t = 2$ . En bas de la colonne, on présente le critère MAPE calculé sur les 8 années. Il n'était pas possible ici, compte tenu du petit nombre de données, de travailler en prévision ex-post c'est-à-dire de sacrifier des données lors de l'ajustement afin d'examiner les performances en prévision.

? Quelle est la courbe de croissance qui réalise le meilleur ajustement pour la Belgique?



#### 3.2.2 Votre réponse

? Quelle est la courbe de croissance qui réalise le plus mauvais ajustement pour la Belgique?



### 3.2.3 Votre réponse



Comment se comportent les courbes en  $S$  (logistique et Gompertz) par rapport aux courbes de croissance de type exponentiel pour la Belgique?



### 3.2.4 Votre réponse



Cliquez sur l'onglet Belgium, afin de visualiser le graphique.



Confirmez-vous vos impressions à l'aide du graphique?



### 3.2.5 Votre réponse

Procédez de même pour la R.F.A.



Le résultat des ajustements par plusieurs courbes de croissance sont donnés dans un tableau. Pressez la touche F5 et sélectionnez ADJFRGMED ou cliquez sur la zone marquée "Ajustement R.F.A (méd)"



Quelle est la courbe de croissance qui réalise le meilleur ajustement pour la R.F.A.?



### 3.2.6 Votre réponse



Quelle est la courbe de croissance qui réalise le plus mauvais ajustement pour la R.F.A.?



### 3.2.7 Votre réponse



Résumez les valeurs de MAPE obtenues pour les quatre méthodes. Comment se comportent les courbes en  $S$  par rapport aux courbes de croissance de type exponentiel pour la R.F.A.?



### 3.2.8 Votre réponse



Cliquez sur l'onglet Germany, afin de visualiser le graphique.



Confirmez-vous vos impressions à l'aide du graphique?



### 3.2.9 Votre réponse

## 3.3 ÉTUDE DE LA RESISTANCE DES METHODES

La méthode des trois points nous intéresse ici par sa simplicité mais aussi par sa relative résistance aux données aberrantes, conséquence de l'emploi comme paramètre de position de la médiane au lieu de la moyenne. Nous avons décrit deux versions de la méthode des trois points: une qui utilise les médianes et une autre qui utilise les moyennes. Il y a trop peu de données dans l'exemple pour que l'aspect de résistance soit vraiment mis en valeur. Néanmoins, un changement important pour certaines données aura un effet différent dans les deux cas. Il ne faut pas modifier les données du sous-ensemble II ni les valeurs des médianes des sous-ensembles I et III, correspondant aux dates  $t = 2$  et  $t = 7$ . En revanche les données pour  $t = 1, 3, 6$  et  $8$  peuvent être modifiées. Les modifications doivent être effectuées dans les zones (E31:E38) ou H31:H38)



Par exemple, changez très franchement la donnée pour  $t = 8$  et la Belgique en remplaçant 79,0 par 179,0 et regardez l'effet sur les deux variantes de la méthode des trois points, celle employant les

moyennes et celle employant les médianes, en cliquant sur les zones marquées "Ajustement Belgique (moy)" et "Ajustement Belgique (méd)".



Qu'observez-vous ?



3.3.1 Votre réponse



Comme second exemple, changez la première valeur de R.F.A en 500 et la dernière valeur en 700 et regardez l'effet sur les deux variantes de la méthode des trois points, celle employant les moyennes et celle employant les médianes, en cliquant sur les zones marquées "Ajustement R.F.A. (moy)" et "Ajustement R.F.A. (méd)".



Qu'observez-vous ?



3.3.2 Votre réponse



### Remarque

Compte tenu du petit nombre de données, l'aspect de résistance aux donnée aberrantes n'est pas très fort ici, surtout quand les données concernées appartiennent au sous-ensembles II. En effet celui-ci ne comporte que deux points. La médiane est donc en fait une moyenne. En revanche, il doit y avoir plus de résistance de la méthode par rapport à des données aberrantes dans les sous-ensembles I et III.



Pour revenir aux données originelles de produit intérieur brut après un changement de ces données, vous pouvez activer la macro-instruction CH03EX03Initial (ou la combinaison de touches CTRL SHIFT I si le classeur est le seul qui soit ouvert).

## SYNTHESE



Dans cette partie, nous avons examiné les résultats de l'ajustement par plusieurs courbes de croissance et ceci en comparant la critère MAPE.

En plus des courbes pour lesquelles nous avons appliqué la méthode des trois points dans la partie 2, nous avons retrouvé la courbe exponentielle dont les paramètres sont obtenus par la méthode des moindres carrées sur les logarithmes des données.

Nous avons vu aussi que la méthode des trois points utilisant les médianes est plus résistante aux données aberrantes que celle employant les moyennes.

**Partie 4** Dans cette partie, on emploie un logiciel conçu pour le traitement de données temporelles. Nous l'emploierons dans plusieurs chapitres de la suite du cours. Il permet notamment d'appliquer la méthode des trois points pour l'ajustement de courbes de croissance.

#### 4.1 EXAMEN DES DONNEES

Nous allons effectuer cette étude en employant Time Series Expert for Windows, en abrégé TSE.

- ⇒ Suivez les instructions rappelées en annexe du document introductif du cours afin de lancer le logiciel.
- ⇒ Choisissez le répertoire de données approprié sur votre disque (pas sur le CD-ROM): menu File ⇒ Open. Choisissez DATA puis CHAP03 puis CH03EX03.
- ⇒ Chargez le problème déjà préparé : CH03EX03. Cliquez sur Open. Vous devez alors voir dans le bas de l'écran que la variable dépendante est BGIP, que l'échantillon d'estimation est 1972 – 1979 et que les prévisions seront calculées jusqu'en 1980.
- ⇒ Pour visualiser le tableau des données: menu Data ⇒ Spreadsheet.
- ⇒ Pressez la touche fonction F3 pour charger une série dans la colonne A du tableau. Sélectionnez BGIP. Cliquez sur Open.
- ⇒ Cliquez dans la colonne B.
- ⇒ Pressez la touche fonction F3 pour charger une série dans la colonne B du tableau. Sélectionnez FRGGIP. Cliquez sur Open.



Reconnaissez-vous les données ?



4.1.1 Votre réponse

Nous aurons besoin des deux séries en logarithmes. En restant dans le tableur, faites ce qui suit, en commençant par la R.F.A., colonne B.



- ⇒ Utilisez le menu Transform ⇒ Natural Logarithm.
- ⇒ Pressez la touche fonction F10 pour renommer la série FRGGIP. Tapez L au début du nom pour avoir LFRGGIP comme nouveau nom.
- ⇒ Pressez la touche fonction F2 pour sauver la série LFRGGIP.
- ⇒ Cliquez en colonne A. Refaites les trois étapes précédentes pour la Belgique mais en changeant le nom BGIP en LBGIP et sauvez.



Reconnaissez-vous les données en logarithmes?

4.1.2 Votre réponse



- ⇒ Quittez le tableur par le menu File ⇒ Exit TSE Spreadsheet.
- ⇒ Pour visualiser graphiquement les séries: menu Graphics ⇒ Series.
- ⇒ Cliquez sur Select. Sélectionnez BGIP (cliquez sur ce nom et cliquez) puis FRGGIP (enfoncez la touche Ctrl puis cliquez sur ce nom). Cliquez sur Open puis sur OK.



Reconnaissez-vous les données ?

4.1.3 Votre réponse

- ⇒ Recommencez les étapes précédentes avec les séries LBGIP et LFRGGIP. Avec la souris, approchez-vous des points. Ceci permet de visualiser la valeur numérique des valeurs en logarithmes.



Reconnaissez-vous les données en logarithmes?



#### 4.1.4 Votre réponse

### 4.2 AJUSTEMENT PAR LES COURBES DE CROISSANCE POUR LA BELGIQUE

Nous commençons par traiter l'ajustement pour la Belgique.

La variable BGIP est définie comme variable dépendante, comme on peut le constater dans la fenêtre du bas.

⇒ Fermez le graphique. Pour estimer les paramètres d'une courbe de croissance exponentielle modifiée, menu Methods ⇒ Growth curves ⇒ Modified exponential.

⇒ Sur la ligne Save residuals, vous devez voir "BRES1". Sur la ligne Save forecasts, vous devez voir "BFOR1". Cliquez OK.

Consultez la sortie. Vous devez voir les estimations des paramètres, puis les valeurs ajustées. Pour la commodité, nous reprenons ici les résultats qui ont été obtenus dans Excel pour la méthode des trois points, avec calcul des médianes.

Médianes		B E L G I Q U E		
t		$y_t$	$1/y_t$	$\ln(y_t)$
2	I	36.7	0.02725	3.6028
4.5	II	54.7	0.01843	3.9978
7	III	74.4	0.01344	4.3095
2.5	$\Delta$			
	$\ln(\beta)$	0.036098751	-0.227883167	-0.094898917
	$\beta$	1.036758222	0.796217279	0.909464861
	$\alpha$	177.3132181	0.032028309	-2.261454542
	$\gamma$	-153.8882353	0.006943227	5.47328536



Pour l'exponentielle modifiée, trouvez-vous les mêmes résultats pour les estimations que dans Excel?



#### 4.2.1 Votre réponse

B E L G I Q U E						
t	année	$Y_t$	Expon. M.C.	Expon. modif.	Logis- tique	Gompertz
1	1972	31.3	32.7	29.9	30.8	30.5
2	1973	36.7	37.5	36.7	36.7	36.7
3	1974	44.4	43.0	43.7	43.3	43.5
4	1975	49.8	49.3	51.0	50.5	50.7
5	1976	59.6	56.5	58.5	58.2	58.3
6	1977	68.0	64.8	66.3	66.2	66.3
7	1978	74.4	74.3	74.4	74.4	74.4
8	1979	79.0	85.2	82.8	82.5	82.7
MAPE			3.59%	2.17%	1.86%	1.99%

?

Pour l'exponentielle modifiée, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?



#### 4.2.2 Votre réponse

Il est instructif de regarder les valeurs ajustées et la prévision pour 1980.

- ⇒ Quittez la sortie.
- ⇒ Pour obtenir le graphique des valeurs ajustées et prévisions en parallèle avec les données, procédez comme suit : menu Graphics ⇒ Predictions/Forecasts. Cliquez OK.

Le graphique est complété par un intervalle de prévision obtenu par la méthode vue au chapitre 1. Comme pour tous les graphiques de TSE, il est possible de focaliser sur un point.

Recommencez pour la courbe de Gompertz.

- ⇒ Fermez le graphique précédent.
- ⇒ Pour estimer les paramètres d'une courbe de croissance de Gompertz, menu Methods ⇒ Growth curves ⇒ Gompertz curve.
- ⇒ Sur la ligne Save residuals, vous voyez "BRES1". Pour le changer en "BRES2", cliquez à droite du nom puis pressez la

touche de correction Backspace pour effacer 1 puis tapez 2. Cliquez OK.

⇒ De même sur la ligne Save forecasts, changez pour avoir “BFOR2”. Cliquez OK.

?

Pour la courbe de Gompertz, trouvez-vous les mêmes résultats pour les estimations que dans Excel?



4.2.3 Votre réponse

?

Pour la courbe de Gompertz, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?



4.2.4 Votre réponse

Recommencez pour la courbe de logistique.

⇒ Pour estimer les paramètres d’une courbe de croissance de Gompertz, menu Methods ⇒ Growth curves ⇒ Logistic curve.

⇒ Sur la ligne Save residuals, vous voyez “BRES2”. Pour le changer en “BRES3”, cliquez à droite du nom puis pressez la touche de correction Backspace pour effacer 2 puis tapez 3. Cliquez OK. De même sur la ligne Save forecasts, changez pour avoir “BFOR3”. Cliquez OK.

?

Pour la courbe logistique, trouvez-vous les mêmes résultats pour les estimations que dans Excel?

4.2.5 Votre réponse



?

Pour la courbe logistique, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?

#### 4.2.6 Votre réponse



#### 4.3 AJUSTEMENT PAR LES COURBES DE CROISSANCE POUR LA R.F.A.



Chargez le problème déjà préparé: menu File  $\Rightarrow$  Open. Choisissez FRGGIP. Cliquez sur Open. Vous devez alors voir dans le bas de l'écran que la variable dépendante est FRGGIP, que l'échantillon d'estimation est 1972 – 1979 et que les prévisions seront calculées jusqu'en 1980.

Procédez comme au paragraphe 4.2 avec des adaptations évidentes des noms : FRGRES1, FRGRES2 et FRGRES3 pour les résidus, FRGFOR1, FRGFOR2 et FRGFOR3 pour les prévisions. Voici les résultats.

Médianes		R. F. A.		
t		$y_t$	$1/y_t$	$\ln(yt)$
2	I	280.4	0.00357	5.6362
4.5	II	369.0	0.00273	5.9074
7	III	504.4	0.00198	6.2234
2.5	$\Delta$			
	$\ln(\beta)$	0.170014082	-0.046717166	0.061202641
	$\beta$	1.185321543	0.954357284	1.063114323
	$\alpha$	118.9958822	0.008347609	1.45113434
	$\gamma$	113.2123134	-0.00403665	3.996127617

t	année	$Y_t$	R. F. A.			
			Expon. M.C.	Expon. modif.	Logis- tique	Gompertz
1	1972	230.9	241.3	254.3	254.5	254.4
2	1973	280.4	272.8	280.4	280.4	280.4
3	1974	320.1	308.4	311.4	310.6	311.0
4	1975	339.1	348.7	348.1	346.2	347.2
5	1976	398.8	394.3	391.6	388.8	390.2
6	1977	453.3	445.8	443.2	440.4	441.9
7	1978	504.4	504.0	504.4	504.4	504.4
8	1979	557.6	569.8	576.9	585.5	580.5
MAPE			2.34%	2.87%	3.20%	3.02%

?

Pour l'exponentielle modifiée, trouvez-vous les mêmes résultats pour les estimations que dans Excel?



4.3.1 Votre réponse

?

Pour l'exponentielle modifiée, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?



4.3.2 Votre réponse

?

Pour la courbe de Gompertz, trouvez-vous les mêmes résultats pour les estimations que dans Excel?



4.3.3 Votre réponse

?

Pour la courbe de Gompertz, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?







4.3.4 Votre réponse



Pour la courbe logistique, trouvez-vous les mêmes résultats pour les estimations que dans Excel?

4.3.5 Votre réponse



Pour la courbe logistique, obtenez-vous les mêmes valeurs ajustées et la même valeur du critère MAPE que dans Excel?

4.3.6 Votre réponse



Pour quitter TSE, menu File ⇒ Exit. Confirmez en cliquant sur Yes.

## SYNTHESE

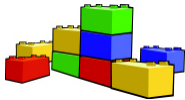
Dans cette partie, nous avons examiné les résultats de l'ajustement par plusieurs courbes de croissance au moyen de la méthode des trois points et nous avons retrouvé les mêmes résultats que dans les parties 2 et 3 de l'exercice.



## Exercice avancé

(Pour les utilisateurs de la version avancée du cours)

### Préalable



Le chapitre 3 du cours de base et du cours avancé doit avoir été suivi jusqu'à la page 74 pour la partie A, et jusqu'à la page 81 pour la partie B.

### Objectif



Le but est d'approfondir la méthode des trois points et d'envisager une alternative : une méthode itérative au moyen du module Solver de Microsoft Excel.

### Données



Les mêmes données que pour l'exercice de base.

### Structure de l'exercice

L'exercice comporte deux parties :

- Dans la partie A, le but de l'exercice est de compléter la partie 1 de l'exercice du cours de base en expliquant la méthode de résolution du système de trois équations non linéaires à trois inconnues et de discuter les cas où il n'y a pas de solution.
- Dans la partie B, le but de l'exercice est d'appliquer la méthode itérative au moyen du module Solver de Microsoft Excel.

**Partie A** Il s'agit de résoudre le système de trois équations (non linéaires) à trois inconnues de manière à achever la résolution entamée dans la partie de base du cours.

### A.a PRINCIPE DE LA METHODE

Nous allons employer ici l'hypothèse que les moyennes des temps des trois sous-ensembles  $t_I$ ,  $t_{II}$  et  $t_{III}$  sont équidistantes. Cette hypothèse n'est pas remplie s'il y a des données manquantes. Nous devons essentiellement discuter l'étape 3) de la méthode. Que l'étape 2) ou l'étape 2') ait été utilisée est sans importance. Voici l'étape 3.

3) Résoudre le système de trois équations (non linéaires) à trois inconnues suivant :

$$y_I = \alpha \beta^{t_I} + \gamma \quad (1)$$

$$y_{II} = \alpha \beta^{t_{II}} + \gamma \quad (2)$$

$$y_{III} = \alpha \beta^{t_{III}} + \gamma. \quad (3)$$

En dépit du fait que les équations sont non linéaires, le système peut être résolu de la façon suivante.

3a) Effectuer les différences membre à membre des équations (3) et (2) et des équations (2) et (1) :

$$y_{III} - y_{II} = \alpha (\beta^{t_{III}} - \beta^{t_{II}}) \quad (4)$$

$$y_{II} - y_I = \alpha (\beta^{t_{II}} - \beta^{t_I}). \quad (5)$$

3b) Diviser membre à membre l'équation (4) par l'équation (5), en notant  $\Delta = t_{III} - t_{II} = t_{II} - t_I$  :

$$\frac{y_{III} - y_{II}}{y_{II} - y_I} = \frac{\beta^{t_{III}} - \beta^{t_{II}}}{\beta^{t_{II}} - \beta^{t_I}} = \frac{\beta^{t_{II}} (\beta^{\Delta} - 1)}{\beta^{t_I} (\beta^{\Delta} - 1)} = \beta^{\Delta},$$

(c'est ici qu'intervient de façon cruciale l'hypothèse que  $t_{III} - t_{II} = t_{II} - t_I$ ), d'où

$$\ln(\beta) = \frac{1}{\Delta} \ln \left( \frac{y_{III} - y_{II}}{y_{II} - y_I} \right).$$

3c) Déterminer  $\beta$  puis, revenant à l'équation (4),

$$\alpha = \frac{y_{III} - y_{II}}{\beta^{t_{III}} - \beta^{t_{II}}}$$

et, grâce à l'équation (1) :

$$\gamma = y_I - \alpha \beta^{t_I}.$$

### A.b MISE EN ŒUVRE DE LA METHODE

⇒ Les calculs sont présentés dans un tableau de la feuille Main. Pour atteindre ce tableau avec les médianes et les moyennes, pressez la touche F5 et sélectionnez MEDAVG ou cliquez au début de la zone marquée "Méthode des trois points".

Le tableau comporte quatre parties. Les deux pays, Belgique et R.F.A., sont côte à côte. Les parties du tableau employant les médianes sont au-dessus des parties du tableau employant les moyennes. Choisissez un des pays pour les médianes. Dans la partie 2 de l'exercice de base, nous avons détaillé les calculs jusqu'à la ligne précédant la ligne marquée  $\ln(\beta)$ .

? Que vaut  $\Delta$  ?

A.b.1 Votre réponse



⇒ Suivez les étapes de calcul dans les formules de la feuille (vous pouvez évidemment vous limiter à un pays et une seule courbe), le principe étant identique pour les autres cas.

? Tout est-il correct?

A.b.2 Votre réponse



### A.c CAS DIFFICILES

Nous avons déjà évoqué les cas où les temps  $t_I$ ,  $t_{II}$  et  $t_{III}$  ne seraient pas équidistants. Ce sera généralement le cas si les données ne sont pas équidistantes dans le temps. Ce sera le cas en particulier si toutes les données sont équidistantes dans le temps mais s'il y a une donnée manquante.

Un autre problème peut se produire même si toutes les conditions sont satisfaites par ailleurs. Il peut se faire qu'il n'y ait pas de solution au système d'équations. Par exemple, dans le cas d'une courbe logistique, si le 2ème point est plus haut que les deux autres,  $y_{III} - y_{II}$  est négatif et  $y_{II} - y_I$  est positif de sorte que le rapport de (4) sur (5) est négatif. Comme  $\beta$  est positif, il n'y a pas de solution.



Changez des données dans le but de « casser » l'une ou l'autre des méthodes.



Que se passe-t-il alors ?



*A.c.1 Votre réponse*

Ceci se voit par l'apparition de messages d'erreur dans les cellules contenant les valeurs estimées des paramètres.

## SYNTHESE

Dans cette partie, nous avons complété la partie 2 de l'exercice de base en expliquant comment résoudre le système de trois équations à trois inconnues. Nous avons également expliqué que la méthode des trois points peut ne pas aboutir.

**Partie B** Dans cette partie, on reprend l'ajustement des courbes de croissance mais cette fois au moyen d'une méthode itérative au lieu de la méthode des trois points. Les courbes de croissance considérées sont l'exponentielle modifiée, les courbes en S de Gompertz et logistique. La méthode itérative est appliquée au moyen du module Solver de Microsoft Excel.

### B.a EXAMEN DES PREPARATIFS

Nous traitons ici uniquement la R.F.A. Un tableau a été préparé.

⇒ Pour atteindre ce tableau, pressez la touche F5 et sélectionnez le champ ITERATIVE ou cliquez au début de la zone marquée "Méthode itérative".

Ce tableau reprend les données pour la R.F.A, l'ajustement par une exponentielle mais surtout les résultats d'ajustement pour trois courbes de croissance : l'exponentielle modifiée, la courbe de Gompertz et la courbe logistique.

? Vérifiez le calcul des valeurs ajustées qui doit être similaire à celui de la partie 3, sauf que les paramètres sont dans des cellules situées en dessous de chaque colonne, sur les lignes 67, 68 et 69. Cela semble-t-il correct ?



*B.a.1 Votre impression*

Notez que les valeurs des paramètres ne sont pas celles qui ont été obtenues par la méthode des trois points, en employant les médianes. Les calculs du MAPE sont effectués de manière cachée dans les quatre colonnes qui suivent.

? Est-ce identique à ce que vous avez trouvé à la réponse 3.2.8 ?



*B.a.2 Votre réponse*

### B.b EXAMEN DES RESULTATS DU MODULE SOLVER D'EXCEL

Le module Solver de Microsoft Excel permet de maximiser ou de minimiser une fonction de plusieurs variables évaluée par des formules dans le tableur. La fonction est ici le critère MAPE, à titre d'exemple parce qu'on aurait pu employer un autre critère tel que MSE ou MAE. Il faut le minimiser. Les variables sont ici les paramètres des courbes de croissance. On doit donc minimiser MAPE en jouant sur les trois paramètres.

Les paramètres sont dans l'ordre: beta ( $\beta$ ), alpha ( $\alpha$ ), et gamma ( $\gamma$ ). Ils sont sur les lignes 67 à 69, comme déjà remarqué dans le paragraphe 4.1. Les valeurs sont celles fournies par la méthode des trois points qui peut constituer un bon point de départ pour la méthode itérative.

Nous allons ici exposer l'emploi d'une routine d'optimisation, le module Solver inclus dans Excel, pour réaliser le choix optimal des paramètres d'une courbe de croissance.

Le module Solver produit un rapport. Nous avons fait appel à ce module pour obtenir la solution optimale pour la courbe logistique au sens du critère MAPE. Cliquez sur l'onglet Answer Report 2 pour voir les résultats.



Quelle étaient les valeurs initiales des trois paramètres? Les valeurs finales étaient-elles différentes?

*B.b.1 Vos réponses*



Quelle était la valeur initiale du critère MAPE? Quelle est la valeur finale? La valeur finale est-elle plus grande ou plus petite que la valeur initiale?

*B.b.2 Vos réponses*



### B.C EMPLOI DU MODULE SOLVER D'EXCEL

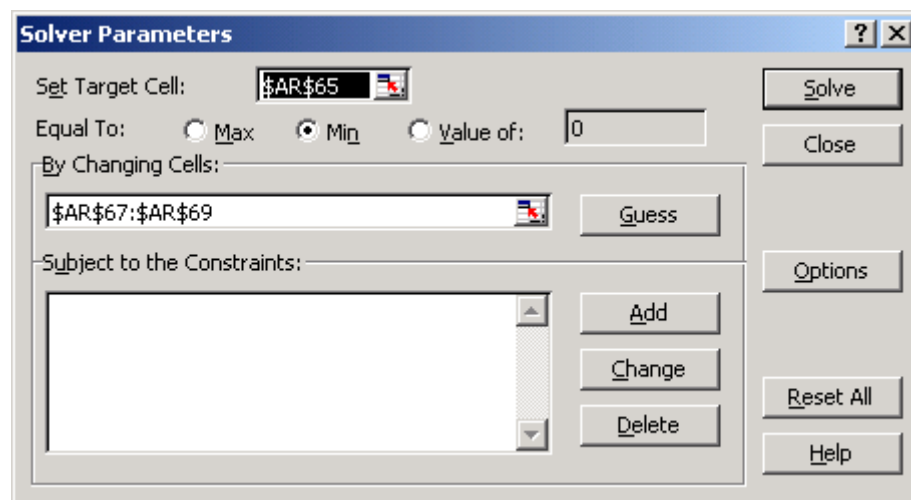
Ce paragraphe facultatif porte sur une utilisation relativement avancée de Microsoft Excel : le module optionnel d'optimisation appelé Solver.

*facultatif*

- ⇒ Placez-vous dans la feuille principale Main.
- ⇒ Déprotégez la feuille: menu Tools, commande Protection et sélectionnez Unprotect sheet.
- ⇒ Choisissez le menu Tools, commande Solver.

#### Remarque

Il est possible que le module Solver ne soit pas installé sur votre système. Cliquez sur le menu Tools, commande Add-Ins et vérifiez dans la liste des programmes auxiliaires que Solver est bien coché. Dans le cas contraire, cochez-le. Si vous avez une installation réduite, il est possible qu'il faille auparavant installer l'ensemble des Add-Ins d'Excel.



- ⇒ Dans la zone "Set target Cell", se trouve le critère MAPE pour la courbe logistique.
- ⇒ Dans la zone, "By Changing Cells", se trouve la plage avec les trois paramètres. La zone "Subject to the Constraints" est vide.
- ⇒ Cliquez sur le bouton Solve. Cliquez Continue pour continuer.
- ⇒ Quand la fenêtre Solver Results apparaît, choisissez "Restore Original Values" et cliquez sur "Answer", puis sur OK. Une feuille



“Answer Report xx” s’ajoute à votre classeur.



Consultez-la afin de déterminer la valeur optimale du MAPE.

*B.c.1 Votre réponse*



Comparez le contenu de la feuille avec le contenu de la feuille Answer Report 2.

*B.c.2 Votre réponse*

### *Remarque*



Il arrive fréquemment que le module Solver ne parvienne pas à quitter les valeurs initiales fournies. Il faut alors les modifier manuellement et relancer le module Solver.

### **B.d RESULTATS DU MODULE SOLVER D'EXCEL**



Nous avons fourni les résultats les meilleurs que nous avons obtenu pour les trois modèles. Ces résultats sont meilleurs que pour la méthode des trois points. C’est assez naturel puisque la méthode des trois points n’essaye pas de minimiser le critère MAPE.



Quelle est la courbe de croissance qui réalise le meilleur ajustement pour la R.F.A. ?

*B.d.1 Votre réponse*



Quelle est la courbe de croissance qui réalise le plus mauvais ajustement pour la R.F.A. ?



*B.d.2 Votre réponse*



### **B.e EXPERIMENTATION AVEC LE MODULE SOLVER D'EXCEL**

Vous pouvez procéder de même pour les deux autres courbes de croissance. D'une fois à l'autre, le Solver propose les mêmes spécifications. Pour changer d'objectif, vous devez changer la cellule mentionnée sous "Set target Cell". Vous devez aussi changer de variables en modifiant "By Changing Cells".

*facultatif*

Afin de faciliter ces manipulations, nous avons préparé trois modèles à partir de la ligne 83. Pour charger le modèle pour la courbe exponentielle modifiée, par exemple, lancez le Solver, cliquez sur le bouton Options, puis sur bouton Load Model. Choisissez la plage de trois cellules allant de AQ83 à AQ85. Cliquez sur OK puis une seconde fois pour accepter de modifier les sélections de cellules. Cliquez encore sur OK puis enfin sur Solve.

## **SYNTHESE**

Dans cette partie, nous avons examiné les résultats de l'ajustement par trois courbes de croissance en employant la méthode itérative avec comme critère la minimisation du critère MAPE.

Nous avons vu qu'il est possible d'améliorer les résultats par rapport à la méthode des trois points mais aussi que la méthode itérative implantée dans le module Solver de Microsoft Excel nécessite des valeurs initiales bien choisies et assez bien de tâtonnement dans ce choix.

[Retour au chapitre 3](#)



*B.d.1 Votre réponse*

?

Quelle est la courbe de croissance qui réalise le plus mauvais ajustement pour la R.F.A. ?



*B.d.2 Votre réponse*

*facultatif*

### **B.e EXPÉRIMENTATION AVEC LE MODULE SOLVER D'EXCEL**

Vous pouvez procéder de même pour les deux autres courbes de croissance. D'une fois à l'autre, le Solver propose les mêmes spécifications. Pour changer d'objectif, vous devez changer la cellule mentionnée sous "Set target Cell". Vous devez aussi changer de variables en modifiant "By Changing Cells".

Afin de faciliter ces manipulations, nous avons préparé trois modèles à partir de la ligne 83. Pour charger le modèle pour la courbe exponentielle modifiée, par exemple, lancez le Solver, cliquez sur le bouton Options, puis sur bouton Load Model. Choisissez la plage de trois cellules allant de AQ83 à AQ85. Cliquez sur OK puis une seconde fois pour accepter de modifier les sélections de cellules. Cliquez encore sur OK puis enfin sur Solve.

## **SYNTHÈSE**

Dans cette partie, nous avons examiné les résultats de l'ajustement par trois courbes de croissance en employant la méthode itérative avec comme critère la minimisation du critère MAPE.

Nous avons vu qu'il est possible d'améliorer les résultats par rapport à la méthode des trois points mais aussi que la méthode itérative implantée dans le module Solver de Microsoft Excel nécessite des valeurs initiales bien choisies et assez bien de tâtonnement dans ce choix.

[Retour au chapitre 3](#)