# Plaquette 4 – Interpolation et extrapolation

## Statistiques à deux variables — Première générale

Corrections détaillées

**Boîte à outils**

**Interpoler, extrapoler, et savoir s'arrêter.**

— **Interpolation.** Estimer $y$ pour une valeur de $x$ **située dans** l'intervalle $[x_{\min}\,;x_{\max}]$ des données observées. L'estimation est encadrée par des relevés réels : elle est **fiable**.

— **Extrapolation.** Estimer $y$ pour une valeur de $x$ **hors** de cet intervalle. Rien ne garantit que la tendance observée se poursuive : l'estimation doit être annoncée comme une **prévision**, d'autant plus fragile qu'on s'éloigne.

— **Méthode d'estimation.** Dans les deux cas, on remplace $x$ dans l'équation de la droite :

$$y = ax + b$$

— **Question inverse** (« à partir de quand ? ») : on **résout** l'équation ou l'inéquation. Comme le résultat porte souvent sur un rang entier (une année, un mois), on **teste les deux entiers voisins** avant de conclure.

— **Les trois signes qu'un modèle a cessé d'être valable.**

— la valeur estimée est **impossible** : effectif négatif, consommation négative, note supérieure au maximum, pourcentage hors de $[0\,;100]$ ;

— la valeur est physiquement irréaliste (population multipliée par mille, surface supérieure au terrain) ;

— le nuage montre un **changement de tendance** que la droite ne peut pas reproduire.

— **Bon réflexe de rédaction.** Toute estimation s'accompagne de deux phrases : l'une donne le résultat avec son unité, l'autre précise s'il s'agit d'une interpolation ou d'une extrapolation, et quelle confiance lui accorder.

— **Ajuster sur une partie du nuage.** Quand la tendance change en cours de série, on peut restreindre l'ajustement aux points de la zone étudiée : le modèle devient plus juste là où on l'utilise, au prix d'un domaine de validité plus étroit.

## Correction de l'exercice 1 – Interpolation ou extrapolation ?

Le critère est unique : l'intervalle des données observées est ici $[1\,;6]$.

$$x \in [x_{\min}\,;x_{\max}] \Rightarrow \text{interpolation} \qquad x \notin [x_{\min}\,;x_{\max}] \Rightarrow \text{extrapolation}$$

a) $3{,}5 \in [1\,;6]$ : c'est une **interpolation**. La valeur est même celle de $\overline{x}$, au cœur du nuage : l'estimation $3{,}83 \times 3{,}5 + 7{,}60 \approx 21{,}01$ est très fiable.

b) $8 \notin [1\,;6]$ : c'est une **extrapolation**, de deux rangs au-delà du dernier relevé. L'estimation $3{,}83 \times 8 + 7{,}60 = 38{,}24$ reste plausible, mais n'est plus garantie.

c) $0 \notin [1\,;6]$ : c'est aussi une **extrapolation**, cette fois en dessous des données. Elle donne 7,60 milliers d'euros, la valeur de $b$ — à interpréter avec prudence, rien n'ayant été mesuré avant l'année 1.

d) $5{,}2 \in [1\,;6]$ : c'est une **interpolation**, encadrée par les relevés des rangs 5 et 6.

**Le point de vocabulaire.** Extrapoler n'est pas interdit : c'est le but même d'un modèle de prévision. Ce qui est fautif, c'est d'extrapoler **sans le dire**, comme si l'estimation avait

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

la même valeur qu'une interpolation.

**Réponse.** a) interpolation — b) extrapolation — c) extrapolation — d) interpolation.

## Correction de l'exercice 2 – Interpoler une valeur manquante

a) On remplace $x$ par 4 dans l'équation de la droite :

$$y = 3{,}86 \times 4 + 7{,}67 = 15{,}44 + 7{,}67 = 23{,}11$$

Le chiffre d'affaires estimé de l'année 4 est donc d'environ 23,1 **milliers d'euros**.

b) La valeur 4 est comprise entre 1 et 6, l'intervalle des abscisses observées : c'est une **interpolation**. C'est la situation la plus favorable, puisque l'estimation est encadrée par des données réelles de part et d'autre.

c) Oui. Le relevé de l'année 3 vaut 19 et celui de l'année 5 vaut 27 : l'estimation 23,1 tombe **au milieu** de ces deux valeurs, ce qui est cohérent pour une année intermédiaire dans une série régulière.

$$\frac{19 + 27}{2} = 23$$

**Le contrôle par la moyenne des voisins.** Sur une série presque alignée, l'interpolation doit donner une valeur proche de la moyenne des deux relevés encadrants. Un écart important signalerait une erreur de calcul, ou un nuage moins régulier qu'annoncé.

**Réponse.** a) environ 23,1 milliers d'euros — b) une interpolation — c) oui, c'est presque la moyenne des valeurs voisines 19 et 27.

## Correction de l'exercice 3 – Extrapoler avec prudence

a) On remplace $x$ par chaque valeur dans l'équation de la droite :

$$y = 3{,}83 \times 8 + 7{,}60 = 30{,}64 + 7{,}60 = 38{,}24$$

$$y = 3{,}83 \times 20 + 7{,}60 = 76{,}60 + 7{,}60 = 84{,}20$$

Soit environ 38 milliers d'euros pour l'année 8, et 84 pour l'année 20.

b) La première. L'année 8 n'est qu'à **deux rangs** du dernier relevé (l'année 6), alors que l'année 20 en est à **quatorze rangs**. Or la fiabilité d'une extrapolation décroît avec la distance aux données : plus on s'éloigne, plus l'hypothèse « la tendance se poursuit à l'identique » devient hasardeuse.

c) Pour $x = 50$ :

$$y = 3{,}83 \times 50 + 7{,}60 = 191{,}5 + 7{,}60 = 199{,}1$$

soit environ 199 milliers d'euros, presque **sept fois** le chiffre d'affaires de la dernière année observée. Ce résultat n'a aucune valeur : il suppose une croissance rigoureusement linéaire pendant 44 ans de plus, sans saturation du marché, sans crise et sans changement de stratégie. Le modèle a été ajusté sur six années : l'utiliser sur cinquante n'a pas de sens.

**La règle empirique.** On n'extrapole raisonnablement que sur une **fraction** de l'étendue des données : ici, deux ou trois rangs au-delà de 6, pas quarante.

**Réponse.** a) environ 38 et 84 milliers d'euros — b) celle de l'année 8, plus proche des données — c) environ 199, sans valeur : l'extrapolation est bien trop lointaine.

## Correction de l'exercice 4 – Quand le modèle devient impossible

a) On remplace $x$ par 9 :

$$y = -1{,}25 \times 9 + 22{,}71 = -11{,}25 + 22{,}71 = 11{,}46$$

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

La consommation estimée est d'environ 11,5 m³ par jour. C'est une **interpolation** (9 est entre 2 et 14) : l'estimation est fiable.

b) Pour $x = 25$ :

$$y = -1{,}25 \times 25 + 22{,}71 = -31{,}25 + 22{,}71 = -8{,}54$$

Le modèle prévoit une consommation de $-8{,}54$ m³ par jour, ce qui est **impossible** : une consommation est toujours positive ou nulle. L'extrapolation est ici franchement invalide.

c) On cherche les températures pour lesquelles l'estimation devient négative :

$$-1{,}25x + 22{,}71 < 0 \Leftrightarrow 22{,}71 < 1{,}25x \Leftrightarrow x > \frac{22{,}71}{1{,}25} = 18{,}168$$

Le modèle donne donc des valeurs négatives dès que la température dépasse environ 18,2 **°C**.

**L'explication physique.** Au-delà de 18 °C, on ne chauffe plus : la consommation de gaz tombe à une valeur résiduelle (eau chaude, cuisine) et n'évolue plus avec la température. La relation cesse d'être affine, et le modèle avec elle. Son domaine de validité est celui des relevés, soit $[2\,;14]$ °C.

**Réponse.** a) environ 11,5 m³/jour — b) $-8{,}54$ m³/jour, impossible : extrapolation invalide — c) dès que $x > 18{,}2$ °C environ.

## Correction de l'exercice 5 – Une note plafonnée

a) On remplace $x$ par 6,5 :

$$y = 1{,}476 \times 6{,}5 + 1{,}357 = 9{,}594 + 1{,}357 = 10{,}951$$

La note estimée est d'environ 11 **sur** 20. C'est une interpolation (6,5 est entre 1 et 8) : l'estimation est raisonnable, à la dispersion du nuage près.

b) Pour $x = 15$ :

$$y = 1{,}476 \times 15 + 1{,}357 = 22{,}14 + 1{,}357 = 23{,}497$$

Le modèle annonce une note de 23,5 sur 20, ce qui est **impossible** : la note est bornée par 20. L'extrapolation est invalide, et l'énoncé de l'impossibilité fait partie de la réponse attendue.

c) On résout l'inéquation :

$$1{,}476x + 1{,}357 > 20 \Leftrightarrow 1{,}476x > 18{,}643 \Leftrightarrow x > \frac{18{,}643}{1{,}476} \approx 12{,}63$$

Le modèle dépasse donc 20 dès environ 12,7 **heures** de révision. Au-delà, il n'a plus aucun sens.

**Ce que la borne révèle.** Dans la réalité, les gains de note diminuent quand les heures s'accumulent : la relation n'est affine que sur une plage limitée. C'est le cas de tous les phénomènes **bornés** — un pourcentage, une note, une proportion, un taux de remplissage.

**Réponse.** a) environ 11 / 20 — b) 23,5 / 20, impossible : extrapolation invalide — c) dès environ 12,7 heures.

## Correction de l'exercice 6 – Atteindre un seuil

a) On isole $x$ :

$$8{,}571x + 25 = 100 \Leftrightarrow 8{,}571x = 75 \Leftrightarrow x = \frac{75}{8{,}571} \approx 8{,}75$$

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

b) Le rang d'un trimestre est un **entier**. La valeur 8,75 signifie que le seuil est franchi entre le huitième et le neuvième trimestre : c'est donc au **neuvième trimestre** que les ventes dépasseraient 100 vélos.

c) On calcule les deux estimations encadrantes :

$$x = 8 : 8{,}571 \times 8 + 25 = 68{,}57 + 25 = 93{,}57 \approx 94 \text{ vélos}$$

$$x = 9 : 8{,}571 \times 9 + 25 = 77{,}14 + 25 = 102{,}14 \approx 102 \text{ vélos}$$

Le huitième trimestre reste sous 100, le neuvième le dépasse : la réponse est confirmée ✓

**La réserve à formuler.** Le neuvième trimestre est trois rangs au-delà du dernier relevé : il s'agit d'une **extrapolation**, à présenter comme une prévision. Si la tendance s'infléchit, le seuil sera atteint plus tard.

**L'erreur d'arrondi à éviter.** Arrondir 8,75 à 9 « parce que c'est plus proche » donne ici la bonne réponse, mais par hasard : la règle est de tester les entiers voisins, car pour un seuil il faut le premier rang qui **dépasse**, et non le plus proche.

**Réponse.** a) $x \approx 8{,}75$ — b) le neuvième trimestre — c) 94 vélos au huitième, 102 au neuvième ✓

## Correction de l'exercice 7 – Loyers : deux estimations très inégales

a) On remplace $x$ par chaque surface :

$$y = 13{,}343 \times 45 + 217{,}905 \approx 600{,}44 + 217{,}905 \approx 818{,}34$$

$$y = 13{,}343 \times 10 + 217{,}905 = 133{,}43 + 217{,}905 \approx 351{,}34$$

Soit environ 818 **€** pour 45 m² et 351 **€** pour 10 m².

b) L'intervalle des données est $[20\,;70]$ m².

— $45 \in [20\,;70]$ : **interpolation**, estimation fiable.

— $10 \notin [20\,;70]$ : **extrapolation**. Elle reste plausible (un studio de 10 m² à 351 € est un loyer réaliste), mais aucun appartement de moins de 20 m² n'a été relevé.

c) Pour $x = 150$ :

$$y = 13{,}343 \times 150 + 217{,}905 = 2\,001{,}45 + 217{,}905 \approx 2\,219{,}36$$

soit environ 2 219 €. C'est une extrapolation **lointaine** : la surface dépasse de plus du double la plus grande observée. Or le prix au mètre carré n'est pas constant — il baisse généralement sur les grandes surfaces, et le marché des 150 m² n'obéit pas aux mêmes logiques que celui des studios. L'estimation donne un ordre de grandeur, pas une prévision.

**Le contrôle par le prix au m².** Le modèle donne ici $\frac{2\,219}{150} \approx 14{,}8$ € le m², contre $\frac{818}{45} \approx 18{,}2$ € pour 45 m² et $\frac{351}{10} \approx 35{,}1$ € pour 10 m². La décroissance est cohérente avec le marché, mais son ampleur sur les très grandes surfaces reste à vérifier.

**Réponse.** a) environ 818 € et 351 € — b) interpolation pour 45 m², extrapolation pour 10 m² — c) environ 2 219 €, extrapolation lointaine à ne pas prendre pour une prévision.

## Correction de l'exercice 8 – Extrapoler sur la taille d'un nouveau-né

a) On remplace $x$ par 53 :

$$y = 0{,}170 \times 53 - 5{,}206 = 9{,}01 - 5{,}206 = 3{,}804$$

La masse estimée est d'environ 3,80 **kg**. Il s'agit d'une interpolation (53 est entre 52 et 54) : l'estimation est fiable, d'autant que $r \approx 0{,}998$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

b) Pour $x = 60$ :

$$y = 0{,}170 \times 60 - 5{,}206 = 10{,}2 - 5{,}206 = 4{,}994$$

soit environ 5,0 kg. Le résultat n'est pas absurde en soi — mais aucun nouveau-né mesurant 60 cm n'a été relevé, et une taille de 60 cm correspond plutôt à un nourrisson de deux mois, dont la corpulence ne suit pas la même relation. L'extrapolation est donc douteuse.

Pour $x = 30$ :

$$y = 0{,}170 \times 30 - 5{,}206 = 5{,}1 - 5{,}206 = -0{,}106$$

Le modèle donne une masse **négative**, ce qui est impossible. L'extrapolation est ici franchement invalide.

c) Le modèle n'est raisonnablement utilisable que sur l'intervalle des tailles observées, soit environ $[46\,;54]$ cm, éventuellement élargi de un ou deux centimètres de chaque côté. En dehors, la relation entre taille et masse change de nature.

**Le lien avec la plaquette 3.** C'est le même constat que pour l'ordonnée à l'origine $b = -5{,}206$ : le modèle n'a de sens qu'au voisinage des données. Un excellent coefficient de corrélation ne change rien à cette limite.

**Réponse.** a) environ 3,80 kg — b) 5,0 kg à 60 cm (douteux) et $-0{,}1$ kg à 30 cm (impossible) — c) environ $[46\,;54]$ cm.

## Correction de l'exercice 9 – Une population qui sature

a) L'année 2024 correspond à $x = 9$, et 2035 à $x = 20$ :

$$y = 46 \times 9 + 1\,240 = 414 + 1\,240 = 1\,654 \text{ habitants}$$

$$y = 46 \times 20 + 1\,240 = 920 + 1\,240 = 2\,160 \text{ habitants}$$

La première estimation est une extrapolation de trois ans, la seconde de quatorze ans : la première est nettement plus crédible.

b) On résout l'équation correspondant au plafond :

$$46x + 1\,240 = 2\,500 \Leftrightarrow 46x = 1\,260 \Leftrightarrow x = \frac{1\,260}{46} \approx 27{,}4$$

Le plafond serait donc atteint au cours de la $28^{\text{e}}$ année après 2015, soit en **2043**. Contrôle : pour $x = 27$, $y = 2\,482$ habitants (sous le plafond) ; pour $x = 28$, $y = 2\,528$ (au-dessus) ✓

c) En réalité, la croissance **ralentira** avant d'atteindre le plafond : dès que les terrains constructibles se raréfient, le rythme de 46 habitants par an ne peut plus se maintenir. La population tendra vers une valeur limite au lieu de continuer à croître linéairement.

Le modèle affine n'est donc pertinent que **tant que rien ne freine** la croissance, c'est-à-dire sur les quelques années suivant les relevés. Un modèle qui plafonne serait plus réaliste à long terme.

**Ce que le correcteur attend.** Non pas seulement le calcul, mais la **critique** : nommer le mécanisme de saturation (terrains disponibles, équipements, urbanisme) et dire explicitement que la droite ne peut pas le représenter.

**Réponse.** a) environ 1 654 habitants en 2024 et 2 160 en 2035 — b) en 2043 ($x = 28$) — c) la croissance ralentira avant le plafond : le modèle affine n'est valable qu'à court terme.

## Correction de l'exercice 10 – Arrondir dans le bon sens

a) On isole $x$ :

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

$$4{,}2x + 18 = 100 \Leftrightarrow 4{,}2x = 82 \Leftrightarrow x = \frac{82}{4{,}2} \approx 19{,}52$$

b) Le nombre de mois est un entier, et la contrainte est de **ne pas dépasser** 100 € : il faut donc le plus grand entier inférieur à 19,52, c'est-à-dire 19 **mois**. Contrôle par les deux entiers voisins :

$$x = 19 : 4{,}2 \times 19 + 18 = 79{,}8 + 18 = 97{,}8 \text{ € } \leq 100 \text{ ✓}$$

$$x = 20 : 4{,}2 \times 20 + 18 = 84 + 18 = 102 \text{ € } > 100 \text{ ✗}$$

c) L'élève a arrondi 19,52 **à l'entier le plus proche**, soit 20, comme on le fait pour un résultat numérique ordinaire. Mais ici la question impose une **contrainte** : rester sous 100 €. Or 20 mois coûtent 102 € et dépassent le budget. Dans un problème de seuil, on n'arrondit pas « au plus proche », on arrondit dans le sens qui **respecte la contrainte**.

**La règle générale.** « Ne pas dépasser » : on arrondit **vers le bas**. « Atteindre au moins » : on arrondit **vers le haut**. Et dans les deux cas, on vérifie en testant les deux entiers voisins — c'est le seul moyen de ne jamais se tromper.

**Réponse.** a) $x \approx 19{,}52$ — b) 19 mois (97,80 €) — c) il a arrondi au plus proche, alors que 20 mois coûtent 102 € et dépassent le budget.

## Correction de l'exercice 11 – Deux horizons de prévision

a) On remplace $x$ par chaque valeur :

$$y = 1{,}8 \times 14 + 12 = 25{,}2 + 12 = 37{,}2$$

$$y = 1{,}8 \times 60 + 12 = 108 + 12 = 120$$

Soit environ 37 200 abonnés au bout de 14 mois, et 120 000 au bout de 60 mois.

b) Les deux sont des extrapolations, mais d'ampleur très différente : 14 mois est à **deux mois** du dernier relevé, 60 mois à **quarante-huit mois**. La première prolonge légèrement une tendance observée sur un an ; la seconde suppose que le rythme de 1 800 nouveaux abonnés par mois se maintiendra pendant quatre ans de plus, sans aucune inflexion.

On peut mesurer l'écart de crédibilité en rapportant la distance d'extrapolation à l'étendue des données : $\frac{2}{12} \approx 17$ % pour la première, contre $\frac{48}{12} = 400$ % pour la seconde.

c) 120 000 abonnés n'est pas un nombre absurde en soi pour un service national. Mais le modèle ignore deux mécanismes essentiels : la **saturation** du marché (le nombre de clients potentiels est fini) et le **départ** d'abonnés existants. Un modèle réaliste devrait comporter une valeur limite, vers laquelle le nombre d'abonnés tend sans la dépasser.

**Le repère à retenir.** Extrapoler au-delà de 20 à 30 % de l'étendue des données demande toujours une justification explicite. Au-delà de 100 %, l'estimation ne vaut plus que comme ordre de grandeur.

**Réponse.** a) environ 37 200 et 120 000 abonnés — b) la première est bien plus fiable (2 mois contre 48 au-delà des données) — c) le modèle ignore la saturation du marché et les départs d'abonnés.

## Correction de l'exercice 12 – Une extrapolation raisonnable

a) On remplace $x$ par chaque masse :

$$y = 0{,}24 \times 180 + 0{,}5 = 43{,}2 + 0{,}5 = 43{,}7 \text{ mm}$$

$$y = 0{,}24 \times 300 + 0{,}5 = 72 + 0{,}5 = 72{,}5 \text{ mm}$$

La première est une interpolation ($180 \in [50\,;250]$), la seconde une extrapolation

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

modérée de 50 g au-delà du dernier relevé.

b) Pour trois raisons qui se cumulent :

— le coefficient de corrélation, $r \approx 0{,}9997$, indique un alignement quasi parfait : la relation affine n'est pas une approximation commode, elle décrit le phénomène ;
— il existe une **loi physique** qui la justifie — la loi de Hooke : l'allongement d'un ressort est proportionnel à la force appliquée ;
— l'extrapolation est **courte** : 50 g au-delà de 250 g, soit 25 % de l'étendue des données.

Contrairement à une série économique, où la tendance peut s'inverser à tout moment, on dispose ici d'un mécanisme connu qui garantit la poursuite de la relation.

c) Oui. Tout ressort possède une **limite d'élasticité** : au-delà d'une certaine masse, il se déforme définitivement et la relation cesse d'être affine (puis il casse). Le modèle reste valable tant qu'on demeure dans le domaine élastique, que seul le fabricant du ressort peut indiquer.

**Ce que l'exercice enseigne.** Une extrapolation est crédible quand elle est **courte** et adossée à un **mécanisme** connu. C'est ce qui distingue les sciences expérimentales des séries statistiques sans théorie sous-jacente.

**Réponse.** a) 43,7 mm et 72,5 mm — b) $r$ quasi égal à 1, loi physique connue, extrapolation courte — c) oui, la limite d'élasticité du ressort.

## Correction de l'exercice 13 – Ajuster seulement une partie du nuage

a) Sur les six relevés, la consommation **baisse** d'abord de 5,2 à 5,0 entre 50 et 70 km/h, puis augmente continûment ensuite. Le nuage comporte donc deux tendances opposées, et une droite croissante ne peut pas décrire les deux.

En écartant le relevé à 50 km/h, on se restreint à la zone où la consommation **croît** avec la vitesse : le modèle affine y est cohérent. Ce n'est pas une donnée « gênante » qu'on supprime, c'est un **domaine d'étude** qu'on précise, celui des vitesses supérieures à 70 km/h.

b) Le coefficient passe de 0,890 à 0,971 : l'ajustement devient **nettement meilleur**, et franchit le seuil de 0,95. C'était attendu, puisqu'on a retiré le seul point qui contredisait la tendance.

c) Avec le nouveau modèle :

$$y = 0{,}047\,5 \times 120 + 1{,}37 = 5{,}70 + 1{,}37 = 7{,}07 \text{ L/100 km}$$

L'estimation passe donc de 6,9 à 7,07 **L/100 km**, soit 0,17 L de plus. Le second modèle est préférable pour cette question, car 120 km/h se situe au cœur de la zone sur laquelle il a été ajusté.

**Le prix à payer.** Le nouveau modèle a un domaine de validité **plus étroit** : $[70\,;130]$ km/h au lieu de $[50\,;130]$. Il ne dit plus rien de la consommation à basse vitesse — et c'est honnête, puisque la droite précédente n'en disait rien de juste.

**Réponse.** a) le relevé à 50 km/h relève d'une tendance inverse — b) $r$ passe de 0,890 à 0,971 : meilleur ajustement — c) 7,07 L/100 km, contre 6,9 avec les six points.

## Correction de l'exercice 14 – Prévoir un budget

a) On remplace $x$ par 9 :

$$y = 2{,}4 \times 9 + 15 = 21{,}6 + 15 = 36{,}6$$

Les dépenses estimées de l'année 9 sont donc d'environ 36,6 **milliers d'euros**, soit 36 600 €.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

b) Le budget devient insuffisant dès que les dépenses dépassent 40 :

$$2{,}4x + 15 > 40 \Leftrightarrow 2{,}4x > 25 \Leftrightarrow x > \frac{25}{2{,}4} \approx 10{,}42$$

Le budget serait donc insuffisant à partir de l'**année** 11. Contrôle : pour $x = 10$, $y = 39$ milliers d'euros (budget encore suffisant) ; pour $x = 11$, $y = 41{,}4$ (dépassement de 1 400 €) ✓

c) L'estimation de l'année 11 est une **extrapolation** de quatre rangs au-delà des données (relevées jusqu'à l'année 7). Elle indique une tendance, pas une certitude : le dépassement pourrait survenir plus tôt (vieillissement accéléré de la flotte, hausse des tarifs) ou plus tard (renouvellement des véhicules).

La recommandation raisonnable est donc d'**anticiper** : prévoir une revalorisation du budget à l'horizon de l'année 10, et surtout **actualiser le modèle** chaque année avec les nouveaux relevés, plutôt que de s'appuyer sur un ajustement vieux de quatre ans.

**Ce que vaut un modèle en gestion.** Il ne sert pas à prédire, mais à **alerter** : il dit vers quand une contrainte va devenir critique, ce qui laisse le temps de décider.

**Réponse.** a) environ 36 600 € — b) à partir de l'année 11 (41,4 contre 39 l'année 10) — c) anticiper dès l'année 10 et réactualiser le modèle chaque année.

## Correction de l'exercice 15 – Deux modèles, deux prévisions

a) Pour $x = 4$, qui est une **interpolation** :

$$\text{régression: } 5{,}1 \times 4 + 20 = 20{,}4 + 20 = 40{,}4$$

$$\text{Mayer: } 4{,}8 \times 4 + 21{,}5 = 19{,}2 + 21{,}5 = 40{,}7$$

L'écart est de 0,3, soit moins de 1 % : les deux modèles sont **pratiquement équivalents**.

b) Pour $x = 25$, qui est une **extrapolation lointaine** :

$$\text{régression: } 5{,}1 \times 25 + 20 = 127{,}5 + 20 = 147{,}5$$

$$\text{Mayer: } 4{,}8 \times 25 + 21{,}5 = 120 + 21{,}5 = 141{,}5$$

L'écart atteint cette fois 6, soit environ 4 % : il a été multiplié par vingt.

c) À l'intérieur du nuage, le choix de la méthode **importe peu** : toutes les droites d'ajustement passent par ou près du point moyen, et se croisent donc dans la zone des données. En revanche, plus on extrapole loin, plus une petite différence de **pente** se traduit par un gros écart d'estimation.

$$\text{écart} = |a_1 - a_2| \times x + |b_1 - b_2| \quad \text{croît avec } x$$

Pour une prévision lointaine, il faut donc à la fois retenir la droite de **régression** (la seule optimale) et rester conscient que l'incertitude sur la pente se propage — ce qui est une raison supplémentaire de ne pas extrapoler trop loin.

**L'image à garder.** Deux droites presque confondues au centre du nuage s'écartent comme les branches d'un compas dès qu'on s'éloigne : c'est la géométrie de l'incertitude en extrapolation.

**Réponse.** a) 40,4 contre 40,7 : écart négligeable — b) 147,5 contre 141,5 : écart de 6 — c) le choix de la méthode compte peu en interpolation, beaucoup en extrapolation lointaine.

## Correction de l'exercice 16 – Synthèse : production d'électricité solaire

a) Avec $n = 7$ :

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

$$\overline{x} = \frac{0+1+2+3+4+5+6}{7} = \frac{21}{7} = 3$$

$$\overline{y} = \frac{120+145+178+205+236+262+295}{7} = \frac{1\,441}{7} \approx 205{,}857$$

Le point moyen est environ $G(3\,;205{,}86)$.

b) On remplace $x$ par 3 dans l'équation $y = 29{,}179x + 118{,}321$ :

$$y = 29{,}179 \times 3 + 118{,}321 = 87{,}537 + 118{,}321 = 205{,}858$$

On retrouve $\overline{y} \approx 205{,}86$ : la droite passe bien par le point moyen ✓ C'est le contrôle systématique de la droite des moindres carrés.

c) L'année 2023 correspond à $x = 5$ :

$$y = 29{,}179 \times 5 + 118{,}321 = 145{,}895 + 118{,}321 \approx 264{,}22$$

Le modèle donne 264 MWh, le relevé 262 MWh : l'écart est de 2 MWh, soit moins de 1 %. C'est un **résidu** faible, qui confirme la qualité de l'ajustement. Il s'agit d'une interpolation, puisque 5 appartient à $[0\,;6]$.

d) L'année 2028 correspond à $x = 10$, et 2050 à $x = 32$ :

$$y = 29{,}179 \times 10 + 118{,}321 = 291{,}79 + 118{,}321 = 410{,}11 \text{ MWh}$$

$$y = 29{,}179 \times 32 + 118{,}321 \approx 933{,}73 + 118{,}321 = 1\,052{,}05 \text{ MWh}$$

La première est une **extrapolation modérée** (quatre ans au-delà des relevés, soit 67 % de l'étendue) : elle est plausible. La seconde est une **extrapolation lointaine** (26 ans, plus de quatre fois l'étendue) : elle suppose que la commune continuera d'installer environ 29 MWh de capacité chaque année pendant un quart de siècle, alors que les toitures disponibles sont en nombre fini. Elle ne vaut que comme ordre de grandeur.

e) On résout :

$$29{,}179x + 118{,}321 > 500 \Leftrightarrow 29{,}179x > 381{,}679 \Leftrightarrow x > \frac{381{,}679}{29{,}179} \approx 13{,}08$$

Le seuil serait donc franchi pour $x = 14$, soit en **2032**. Contrôle : $x = 13$ donne 497,6 MWh (sous le seuil) et $x = 14$ donne 526,8 MWh ✓

La réserve : il s'agit d'une extrapolation de huit ans, qui suppose un rythme d'installation constant. Elle sera d'autant plus fausse que la commune approchera de la **saturation** de ses surfaces exploitables — moment où la croissance cessera d'être linéaire.

**Le bilan de la plaquette.** Une estimation se juge sur trois critères : est-elle **dans** ou **hors** de l'intervalle des données ? à quelle **distance** ? le résultat est-il **possible** ? Répondre aux trois, c'est toute la différence entre un calcul et une réponse.

**Réponse.** a) $G(3\,;205{,}86)$ — b) la droite donne 207,0, l'écart venant des arrondis — c) 264 contre 262 MWh : résidu de moins de 1 % — d) 410 MWh en 2028 (extrapolation modérée) et 1 052 MWh en 2050 (extrapolation lointaine, peu crédible) — e) en 2032 ( $x = 14$), sous réserve de saturation.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).