Validation d'une méthode analytique

Exercices d'application avec WinNonlin

- Exemple 1 : Construction d'une droite du jour et son rejet en cas d'inadéquation des résultats obtenus par étalonnage inverse (back calculation)
- Exemple 2 : Construction d'une droite du jour et acceptation sur la base des résultats obtenus par étalonnage inverse
- Exemple 3 : test de linéarité d'une droite (test de lack-of-fit) Test de la composante quadratique (test réalisés lors de la validation initiale)
- Exemple 4 Essai de répétabilité et de reproductibilité : calculs à effectuer à la main pour comprendre la démarche générale
- Exemple 5 : Essai de répétabilité et de reproductibilité. Analyse d'un essai avec trois niveaux de concentrations.

P.L. Toutain – A. Bousquet-Mélou – D. Concordet Ecole Nationale Vétérinaire de Toulouse UMR 181 de Physiopathologie et Toxicologie Expérimentales INRA/ENVT

Mai 2008

1. EXEMPLE N°1

1.1. Objectifs:

L'objectif de ce premier chapitre est de montrer

- 1- comment construire une droite ou une courbe d'étalonnage avec WinNonlin (WN version 5.2 ou supérieure).
- 2- Comment utiliser les différents outils statistiques disponibles sur WinNonlin pour évaluer la validité de la courbe de calibration vis-à-vis des objectifs assignés à une technique analytique.

Une connaissance minimale de WinNonlin est nécessaire pour réaliser les différents exemples.

1.2. Les données

Soit la gamme d'un analyte obtenue le 13/09/1993 (tableau 1.1.)

Tableau 1.1: Concentrations en analyte et réponses chromatographiques

	Concentration (µg/mL)	Réponses chromatographiques
1	0.10	0.33
2	0.15	0.72
3	0.25	0.64
4	0.50	1.51
5	1.25	4.45
6	2.50	10.04
7	5.0	21.29
8	10	51.73
9	20	90.29

Le niveau de quantification (LOQ) a été fixé à $0.10~\mu g/mL$. Pour cette concentration, précision et exactitude doivent être de $\pm 20\%$ par rapport à la valeur nominale : pour les autres concentrations, la tolérance est de $\pm 15\%$.

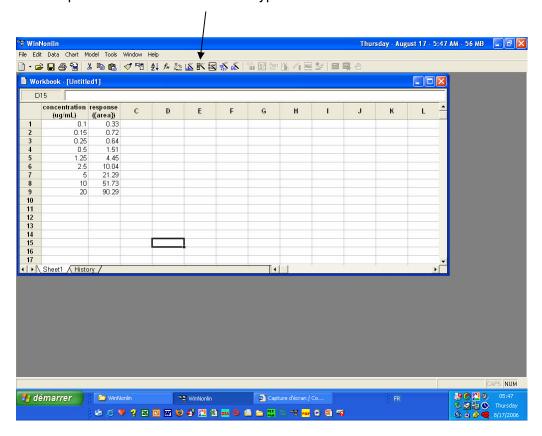
1.3. Ouvrir ou entrer les données dans une feuille de WN

Ouvrir un nouveau dossier dans WinNonlin ou ouvrir le workbook de WinNonlin contenant les données (fig. 1.1). Elles sont placées sur 2 colonnes intitulées : concentrations et réponse.

Nous allons utiliser ces données pour effectuer différents ajustements avec une droite simple avec ou sans pondération.

Figure 1.1 : Concentrations en analyte et réponses entrées dans le workbook de WN et correspondant au tableau 1.1

Bouton pour accéder aux différents types de modèles



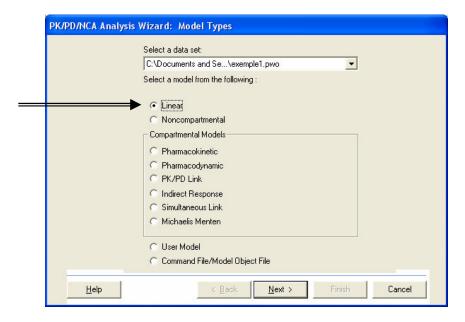
Editer les colonnes en allant dans "Data" puis "column properties".

1.4. Ajustement des données

a) Ajustement à une droite y = ax + b sans pondération

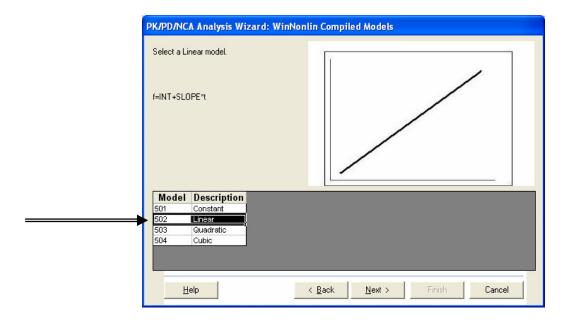
Se placer dans le module de WinNonlin soit en allant dans "*Tools*", soit avec l'icône correspondante (fig.1.1).

Figure 1.2 : écran d'accueil pour sélectionner le module d'analyse de régression linéaire et faire "Next".



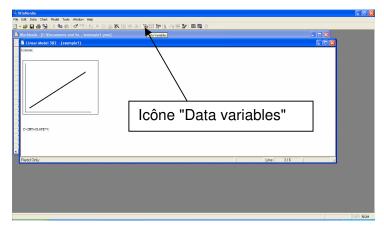
L'écran d'accueil du modèle apparaît, cliquer alors sur "Linear" (fig 1.3)

Figure 1.3.: écran d'accueil du module de régression linéaire après la sélection du modèle "*Linear*" (modèle 502)



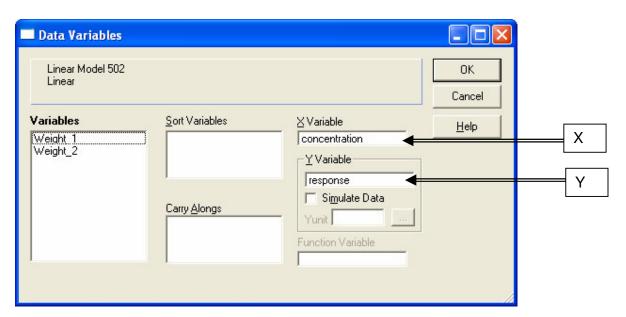
Faire "Next" et "Finish". Apparaît alors l'écran de la figure 1.4.

Figure 1.4 Pour sélectionner ses variables, aller dans "*Model*" puis "*Data Variables*" ou cliquer sur l'icône



Pour réaliser une régression, il convient de déclarer le nom de la variable indépendante et de la variable dépendante. Pour cela il faut aller dans "Data variable" soit en cliquant sur l'icône (fig. 1.5) soit en allant dans le menu Model>Data variable puis par simple glissement, placer le vecteur "concentration" dans la boîte "X variable" et le vecteur "Response" dans la boîte "Y variable" (fig.1.5).

Figure 1.5.: Ecran d'accueil de "*Data variable*". Les noms des variables dépendantes (response) et indépendantes (concentration) sont placés dans les fenêtres *ad hoc* par glisser-coller pour déclarer les variables.



Faire OK

Ensuite, aller vérifier qu'aucun facteur de pondération n'est utilisé en allant sur l'icône "Model options" ou en allant dans le menu Model > Model options (fig 1.6). Dans la fenêtre Model options (fig 1.6) cliquer sur "Weight" et vérifier que le programme est configuré par défaut "uniform weighting". Cliquer sur OK. On peut donner un titre aux futures figures, tableaux etc. en éditant le tableau qui s'ouvre en cliquant sur "Title" (fig. 1.6). De même on peut changer l'algorithme d'optimisation en allant dans "Linear setting" (ici on gardera l'option "Gauss Newton" offerte par défaut).

Figure 1.6.: écran d'accueil pour sélectionner des options de modélisation dont l'utilisation d'un facteur de pondération (*Weight*)

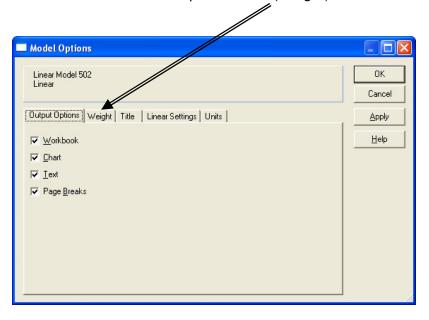
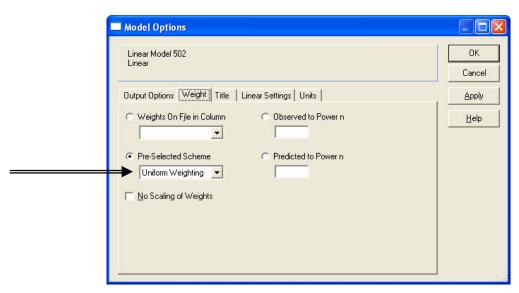
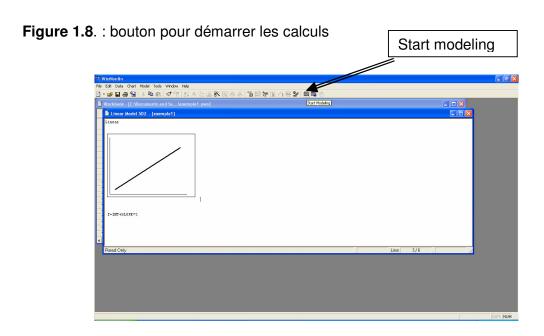


Figure 1.7.: Écran d'accueil pour la gestion des facteurs de pondération ; par défaut, il n'y a pas de pondération et la boite « *Pre-Selected Scheme* » indique "*Uniform Weighting*" ce qui veut dire que le même poids sera donné à toutes les données



Pour exécuter les calculs cliquer sur l'icône "Start modeling" ou en allant dans le menu *Model > Start modeling* (fig. 1.8)



Après l'exécution des calculs, apparaissent 3 feuilles intitulées respectivement "Linear workbook", "Linear text", et "Linear Chart" (fig 1.9). Pour bien visualiser ces 3 feuilles aller dans le menu Window > Cascade. Décaler les feuilles avec le pointeur.

En allant dans l'onglet "*Non-transposed final parameters*" (fig 1.11) on peut visualiser les résultats avec les deux paramètres estimés à savoir l'ordonnée à l'origine = intercept : INT = -0.414471 et la pente : Slope = 4.647302.

Figure 1.9. : cascade de fenêtres montrant les 3 feuilles de résultats pour l'analyse des données par régression linéaire des données du tableau 1.1.

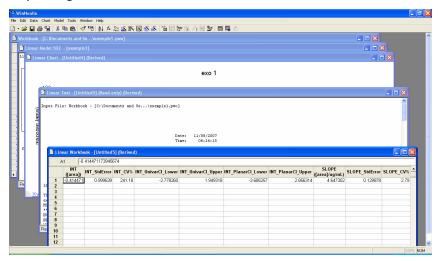
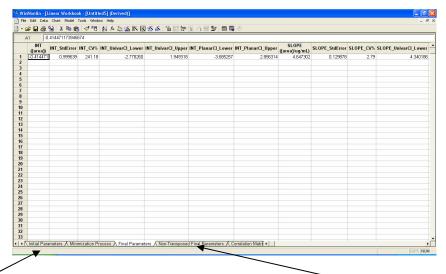


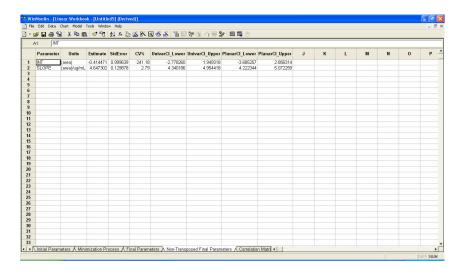
Figure 1.10: Première fenêtre des tableaux de résultats de l'analyse des données du tableau 1.1 avec un modèle de régression linéaire simple et sans pondération. Les onglets situés au bas de la feuille permettent de faire défiler les différents tableaux de résultats.



Onglets de défilement pour les tableaux de résultats

"Non transposed final parameters" donnant l'ordonnée à l'origine de la pente

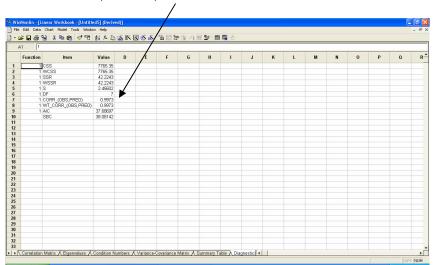
Figure 1.11 : fenêtre intitulée "Non transposed final parameters" donnant la valeur de la pente (Slope= 4.6473) et de l'ordonnée à l'origine (INT=-0.4144)



Sur cette fenêtre sont également donnés des intervalles de confiance (CI) des estimateurs. On peut constater que pour l'intercept, l'intervalle de confiance est de - 2.778 et +1.949 ce qui veut dire que l'intercept n'est pas significativement différent de 0. Cela dit, il est usuel de garder l'ordonnée à l'origine comme variable d'ajustement pour avoir les meilleures prédictions possibles sur l'intervalle d'interpolation.

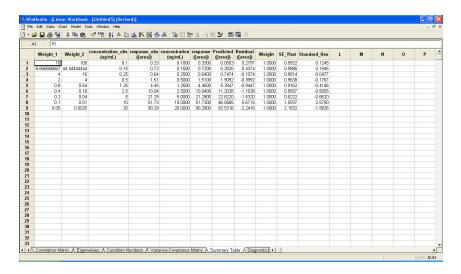
Toujours sur cette feuille, aller à l'onglet "*Diagnostic*" (fig. 1.12) pour avoir la valeur du coefficient de corrélation : CORR-(Obs,pred)=0.997. On notera que le coefficient de corrélation est très élevé mais qu'il ne préjuge en rien de la qualité de la droite de régression.

Figure 1.12 : fenêtre ouverte avec l'onglet « *Diagnostic* » et donnant la valeur du cœfficient de corrélation (r=0.997)



En cliquant sur l'onglet "Summary table" on pourra visualiser le tableau des valeurs observées et des valeurs prédites correspondantes (figure 1.13). On notera aussi que tous les poids (colonne Weight) sont égaux à 1 (pas de pondération).

Figure 1.13 : fenêtre ouverte avec l'onglet « *Summary Table* » donnant les valeurs prédites des réponses chromatographiques par le modèle de régression linéaire.



Avant de travailler avec cette fenêtre, notamment pour faire les "Back calculations", nous allons inspecter la feuille donnant les graphiques. Pour la faire réapparaître faire *Window > Cascade* et sélectionner *3-Linear Chart* (fig.1.14) ou encore réduire la feuille de résultats pour faire apparaître les graphiques.

La première figure qui apparaît est celle des "X vs. observed Y and predicted Y" (fig. 1.15A). L'inspection visuelle de ce premier graphique plaide pour un ajustement satisfaisant. Pourtant on doit considérer que la gamme des concentrations était plutôt large (de 0.1 à 20) et que ce type de graphique est incapable de permettre de faire un bon diagnostic de la qualité de la régression car les grandes valeurs "écrasent" les petites valeurs. La seule façon de juger visuellement de la qualité de l'ajustement est d'inspecter les résidus. Pour cela cliquer sur l'onglet "X vs. weighted residual Y" (fig 1.15B).

Figure 1.14: Ecran montrant comment sélectionner une feuille de résultats à partir de la cascade donnée par WinNonlin

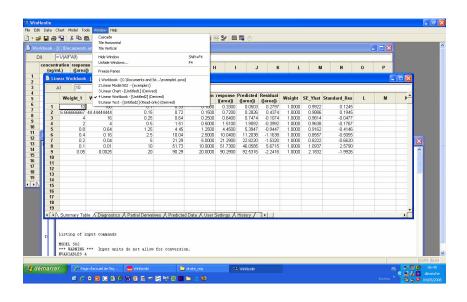


Figure 1.15A : figure montrant l'ajustement des données du tableau 1.1 par régression linéaire ; à première vue cet ajustement paraît satisfaisant.

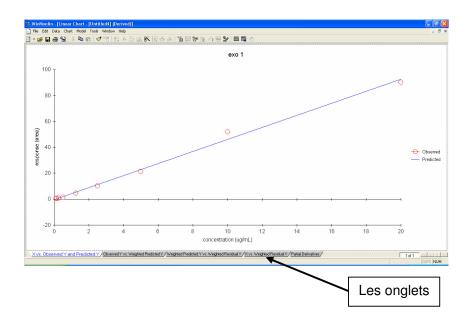
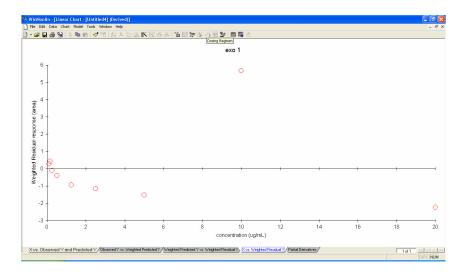


Figure 1.15B : représentation des valeurs des résiduelles en fonction des concentrations testées (X) ; on notera la présence de « structure » c'est-à-dire d'une répartition des résiduelles qui n'est pas aléatoire de part et d'autre de la ligne horizontale d'ordonnée 0.



Les résidus sont la différence entre les valeurs observées et les valeurs prédites par le modèle. Les résiduelles peuvent être négatives ou positives et un bon ajustement implique une alternance régulière de résiduelles positives et négatives. Ici, on remarque que les résiduelles ne se répartissent pas de façon aléatoire autour de la ligne horizontale d'ordonnée 0. Au contraire, on a une tendance claire à la décroissance des résiduelles pour les concentrations allant de 0.1 à 6 et une remontée des résiduelles pour les concentrations égales à 10. On dit qu'il y a de la "structure" dans le graphique des résiduelles. Pour mieux le visualiser on peut éditer les symboles des points en double-cliquant sur l'un des points. Cela fait apparaître une fenêtre (fig. 1.16) qui permet d'éditer la figure en y ajoutant une ligne (onglet Lines) ou en changeant les symboles des points (à gauche aller dans Series cliquer sur Series 1 > data point > markers > filled circle). Le résultat est donné par la figure 1.17.

Figure 1.16 : fenêtre permettant d'éditer la figure en y ajoutant une ligne entre les points et en changeant le symboles des points pour les rendre plus visibles.

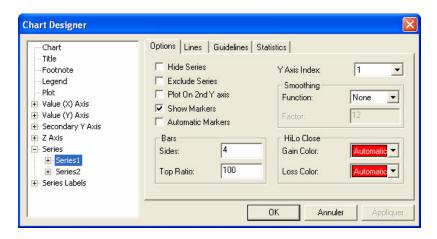
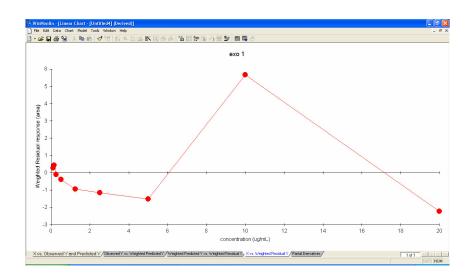


Figure 1.17 : après édition ; la présence d'une ligne pour rejoindre les points des résiduelles facilite l'indentification d'une tendance dans la répartition des résidus



En retournant sur le tableau "Summary table" (fig 1.13) on peut voir la valeur des résidus. On remarquera immédiatement que les résiduelles des 2 premières concentrations (0.27, 0.437) sont positives alors qu'elles sont négatives pour toutes les autres concentrations à l'exception de la concentration "10". Cela suggère déjà un mauvais ajustement.

Dans un ajustement idéal, on aurait une alternance régulière de résiduelles positives et négatives.

L'inspection des résidus suggère que l'ajustement à une droite n'est pas acceptable.

Pour mieux se convaincre du caractère inacceptable de cette droite qui nous était pourtant apparue comme adéquate et qui nous avait donné un cœfficient de corrélation élevé (r=0.997), nous allons "recalculer" les concentrations nominales avec la droite ajustée c'est-à-dire recalculer les valeurs de X avec cette droite. On parle de "back calculation" ou encore d'étalonnage inverse.

Pour cela, nous allons utiliser directement WinNonlin qui permet de faire les calculs simplement. Si on veut utiliser la feuille "Summary Table" contenant les "Residuals" il faut pouvoir l'éditer. Pour cela il faut d'abord la déverrouiller avec la fonction "Detach" qui est obtenu par un click droit de la souris.

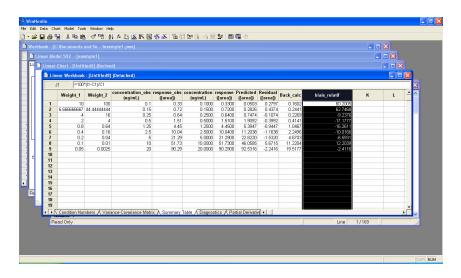
Remarque cette fonction "Detach" n'est pas accessible par le menu "Edit".

Ensuite on peut gérer cette feuille exactement comme une feuille EXCEL. Enlever les colonnes inutiles avec "*Delete*". Ensuite, calculer les concentrations prédites par la droite avec l'équation (fig. 1.19) :

$$X = \frac{Y - b}{a}$$

soit $X = (Y + 0.4145)/4.6473$

Figure 1.19: Calcul des valeurs prédites par la droite de calibration pour les concentrations ayant servi à construire la droite et biais relatif par rapport aux valeurs nominales.



Ensuite, calculer le pourcentage de déviation c'est-à-dire (tableau 1.2) :

Deviation (%) =
$$\frac{[(concentration_observ\acute{e}e) - (concentration_pr\acute{e}dite)]}{(concentration_observ\acute{e}e)} \times 100$$

Tableau 1.2 : calcul des concentrations théoriques de la gamme de l'exemple 1.1 à partir de la droite ajustée par régression linéaire et sans pondération.

Concentrations théoriques (mg/mL)	Réponses chromato- graphiques	Concentrations théoriques recalculées avec les paramètres de la droite issus du tableau 1.1	Déviation (%)
0.10	0.33	0.160	+60
0.15	0.72	0.244	+62.7
0.25	0.64	0.226	-9.6
0.5	1.51	0.414	-17.2
1.25	4.45	1.046	-16.3
2.50	10.04	2.249	-10.04
5.0	21.29	4.670	-6.6
10	51.73	11.22	+12.2
20	90.29	19.52	-2.4

Concentration calculée = (réponse + 0.41447)/4.647

L'inspection du tableau 1.2 montre clairement que cette droite est incapable de prédire les concentrations ayant servi à construire la gamme avec une exactitude appropriée (<15%) notamment pour les plus petites concentrations pour lesquelles le biais est de l'ordre de 60%.

Nous allons tenter d'améliorer cette droite en calculant ses paramètres avec un facteur de pondération c'est-à-dire donner plus de poids aux petites concentrations pour qu'elles "pèsent" davantage sur l'estimation de *a* et *b* c'est-à-dire la pente et l'ordonnée à l'origine.

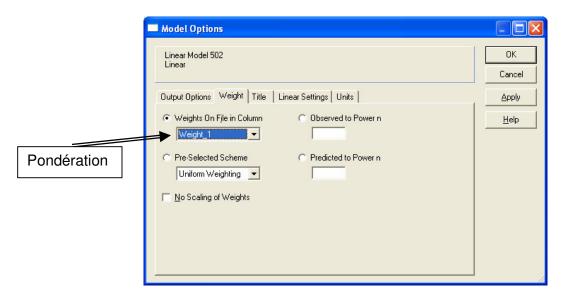
1.5. Ajustement des données du tableau 1.1 à une droite y = aX+b et pondération par 1/X

Pour réaliser une régression pondérée avec WinNonlin, on va reprendre le même modèle dans WinNonlin mais en prenant en compte une pondération. Pour cela repartons dans notre feuille initiale de données et saisissons 2 nouvelles colonnes une nommée 1/X et l'autre 1/X² dans lesquelles on calculera les

valeurs de 1/X et 1/X². Nous allons utiliser les 2 vecteurs de pondération 1/X (weight 1) et 1/X² (weight 2). (Aller dans l'onglet *Window > Cascade > workbook*)

En reprenant le modèle de régression on va refaire tous les calculs précédents mais avec la pondération. Nous allons sélectionner le vecteur "Weight_1" dans la cellule "Weights on file in column".

Figure 1.20: Fenêtre pour le calcul d'une droite pondérée avec 1/X. On doit d'abord déclarer le vecteur des poids dans la boîte « *Weights on File in Column* » avec l'ascenseur.



Faire OK et lancer les calculs. Aller inspecter les résultats de la nouvelle feuille dans l'onglet "*Non transposed final parameters*".

Les paramètres estimés avec une droite pondérée par 1/X sont la pente (4.6125) et l'ordonnée à l'origine (-0.2606). La figure 1.21 montre la droite de régression obtenue. Ici encore, la droite paraît satisfaisante mais l'inspection des résidus suggère que cette droite obtenue avec un facteur de pondération n'est toujours pas acceptable (fig 1.22).

Figure 1.21: droite obtenue avec les données du tableau 1.1 ajustée par un modèle de type y = aX + b et avec une pondération de 1/X. L'inspection visuelle de la figure suggère un bon ajustement

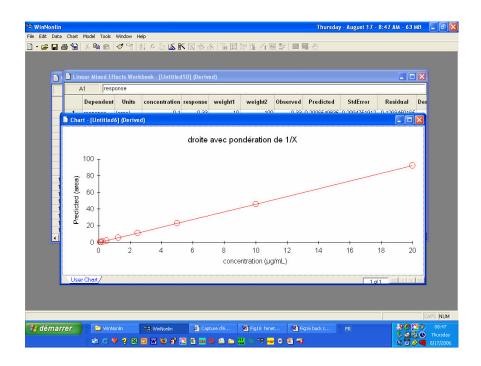
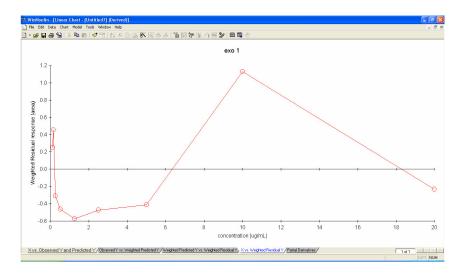


Figure 1.22 Inspection des résidus pondérés (weighted_residuals)

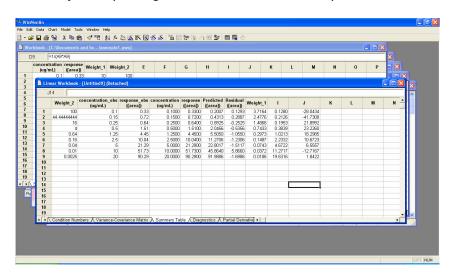


L'inspection de la figure 1.22 montre une mauvaise répartition des résidus pondérés.

Le mauvais ajustement est confirmé par un calcul "back calculation" = étalonnage inverse (fig 1.23).

L'inspection du tableau montre que les concentrations calculées par étalonnage inverse sont meilleures qu'avec celles calculées avec la droite non pondérée mais les déviations restent inacceptables.

Figure 1.23 : Calcul des concentrations théoriques de la gamme de l'exemple 1.1 à partir de la droite ajustée par régression linéaire avec une pondération 1/X



Nous allons reprendre tous les calculs avec une droite pondérée par 1/X² (fig. 1.24). Pour cela il faut re-sélectionner la feuille des données avec *Window > workbook*. Puis dans la fenêtre des "*Model > options*" remplacer "*Weight_1*" par "*Weight_2*" et lancer les calculs.

Les paramètres obtenus sont **b = -0.0738** et **a = 4.1098**. Le tableau des résidus pondérés est donné dans la figure 1.23. La figure 1.24 donne une représentation graphique des résidus pondérés. L'inspection de la figure 1.24 montre à nouveau que la répartition des résidus n'est pas satisfaisante.

Figure 1.24. : résidus pondérés (*weighted_residuals*) calculés à partir de l'ajustement réalisé avec un facteur de pondération de 1/X²

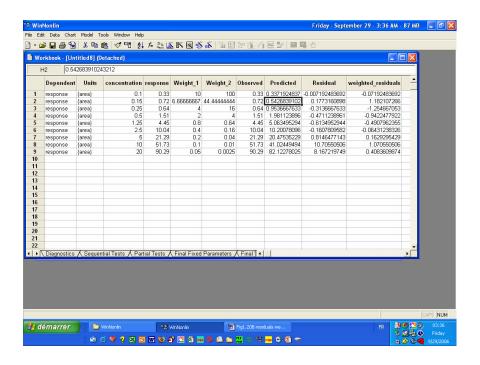
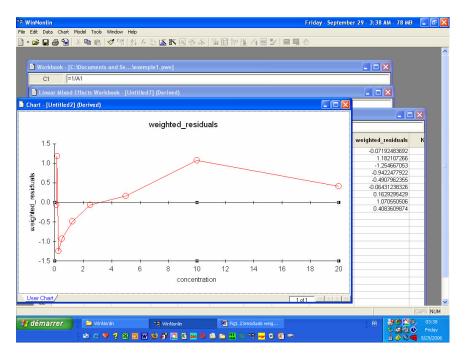


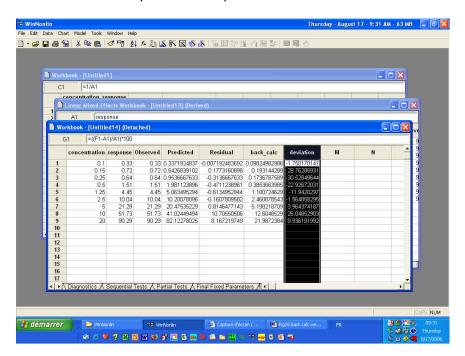
Figure 1.25::Inspection des résidus pondérés (weighted_residuals) calculés à partir de l'ajustement réalisé avec un facteur de pondération de 1/X²



L'inspection du tableau de la figure 1.26 des "Back-calculations" montre que les valeurs prédites sont meilleures que celles obtenues précédemment mais qu'elles ne sont pas encore satisfaisantes car les déviations de plusieurs d'entre elles dépassent encore 20% c'est-à-dire la tolérance acceptable.

Après avoir tenté les différents schémas de pondération, il apparaît impossible d'ajuster correctement les données du tableau 1.1 avec une droite pondérée ou non.

Figure 1.26. : concentrations obtenues par étalonnage inverse avec la droite obtenue avec une pondération par $1/X^2$.



En fait, l'inspection des données brutes du tableau 1.1 montre que les observations 2 et 3 méritent d'être reconsidérées dans la mesure où la réponse chromatographique est plus élevée avec une concentration de 0.15 (0.72) qu'avec la concentration de 0.25 mg/mL (0.64).

Selon les recommandations internationales, deux points de la gamme peuvent être rejetés. Lorsque les observations 2 et/ou 3 sont enlevées, on peut recalculer la droite et en évaluer comme précédemment, les performances en terme d'aptitude à prévoir les concentrations de la gamme. De même, on peut réduire l'étendue de la gamme en excluant les concentrations les plus élevées (10 et 20 µg/mL). Cela a été réalisé et aucune des droites calculées n'a été acceptable. Il est donc réaliste de mettre en

doute les données brutes c'est-à-dire la validité les réponses obtenues pour réaliser des dosages expérimentaux.

1.6. Conclusion

Sur la base des arguments statistiques usuels, les différentes droites obtenues avec ou sans pondération sont à rejeter car l'exactitude des valeurs prédites sera inacceptable vis-à-vis de nos objectifs analytiques.

L'analyste devra donc reconstruire une nouvelle droite.

Cet exemple nous aura permis de montrer comment on peut construire une droite avec WinNonlin et en discuter la valeur vis-à-vis des objectifs analytiques.

Dans l'exemple suivant nous allons construire une droite du jour c'est-à-dire celle qui est réalisée en routine.