

On n'a pas tous les jours la possibilité de mener une étude sur un concours avec 72 concurrents jugés par 5 juges sur une même carrière. Quelles leçons peut-on tenter de tirer d'une petite analyse statistique des résultats?

A la demande de Jean-Claude Grognet, je me suis livré à une analyse statistique des notes de dressage attribuées lors du mondial 1 cheval de Pratoni.

Le jugement est-il équitable en fonction des juges et des écarts de note ?

Comme bien souvent lorsque l'on regarde ses protocoles de dressage, on est surpris par les écarts entre les notes attribuées par les différents juges. Cela est bien souvent ressenti comme une injustice personnelle.

Ne sachant pas trop bien par quel bout prendre ce problème, j'ai pensé que la représentation graphique des données retravaillées sous différents angles pourrait apporter un éclairage intéressant sur quelques phénomènes qui me paraissaient presque évidents en regardant les résultats à la première lecture, évidences qui à ma grande surprise se sont révélées être des à priori non confirmés par l'analyse.

Pour cette analyse,

j'ai utilisé une approche sur la notation individuelle de chaque juge comparée à la notation moyenne des 5 juges pour chaque concurrent, en valeurs brutes et en écart de points.

j'ai utilisé une approche sur les classements relatifs de chaque concurrent dans le jugement de chaque juge par rapport au classement final obtenu par le concurrent, en valeurs brutes et en écart de classement.

Pour lire les résultats de cette analyse quelques clés :

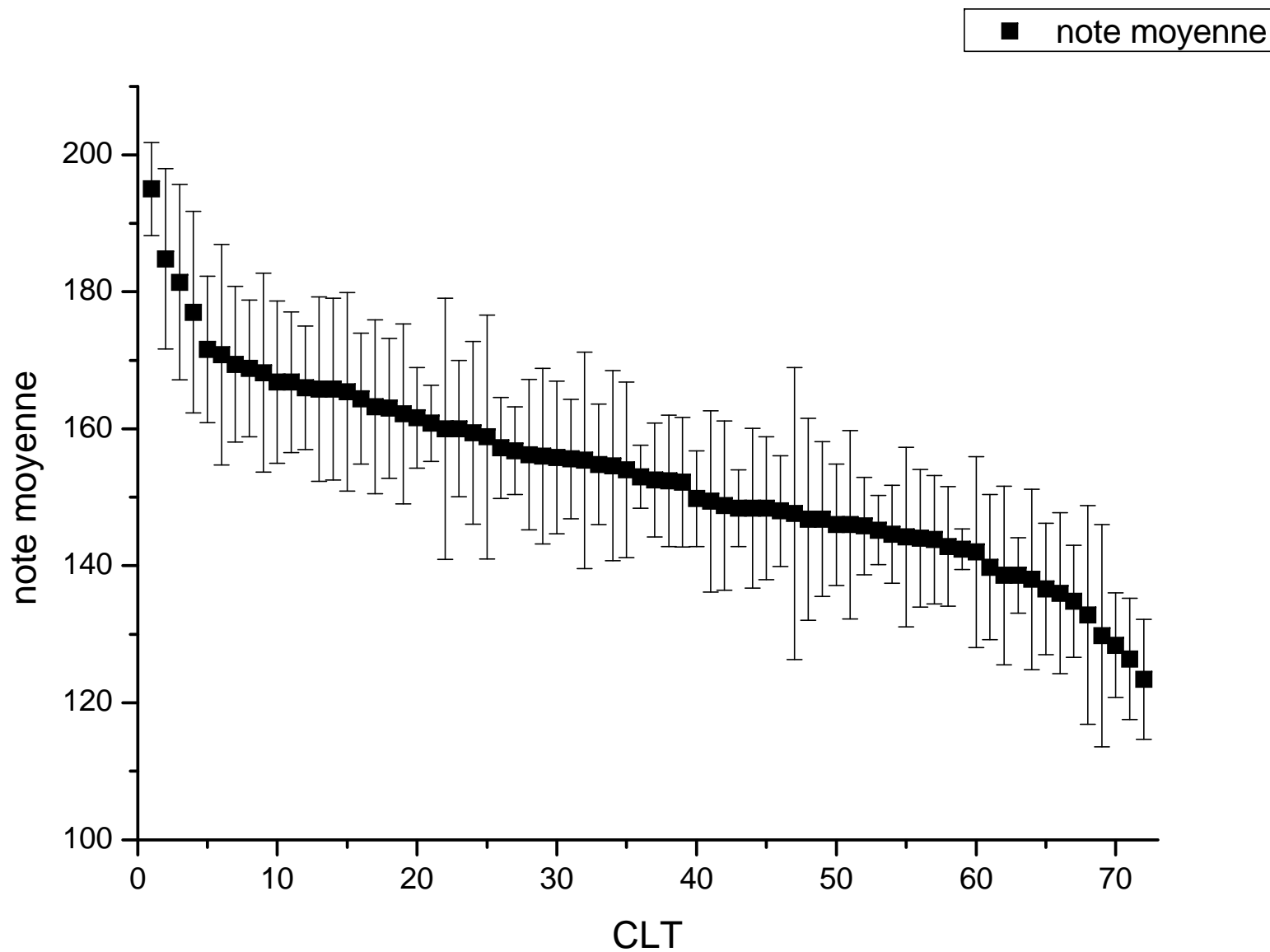
Lorsque l'on représente une corrélation entre deux séries de valeurs, une distribution des points sur la bissectrice est une corrélation parfaite. Tout point s'écartant de la bissectrice représente un défaut de la corrélation

Deux façons d'aborder les statistiques :

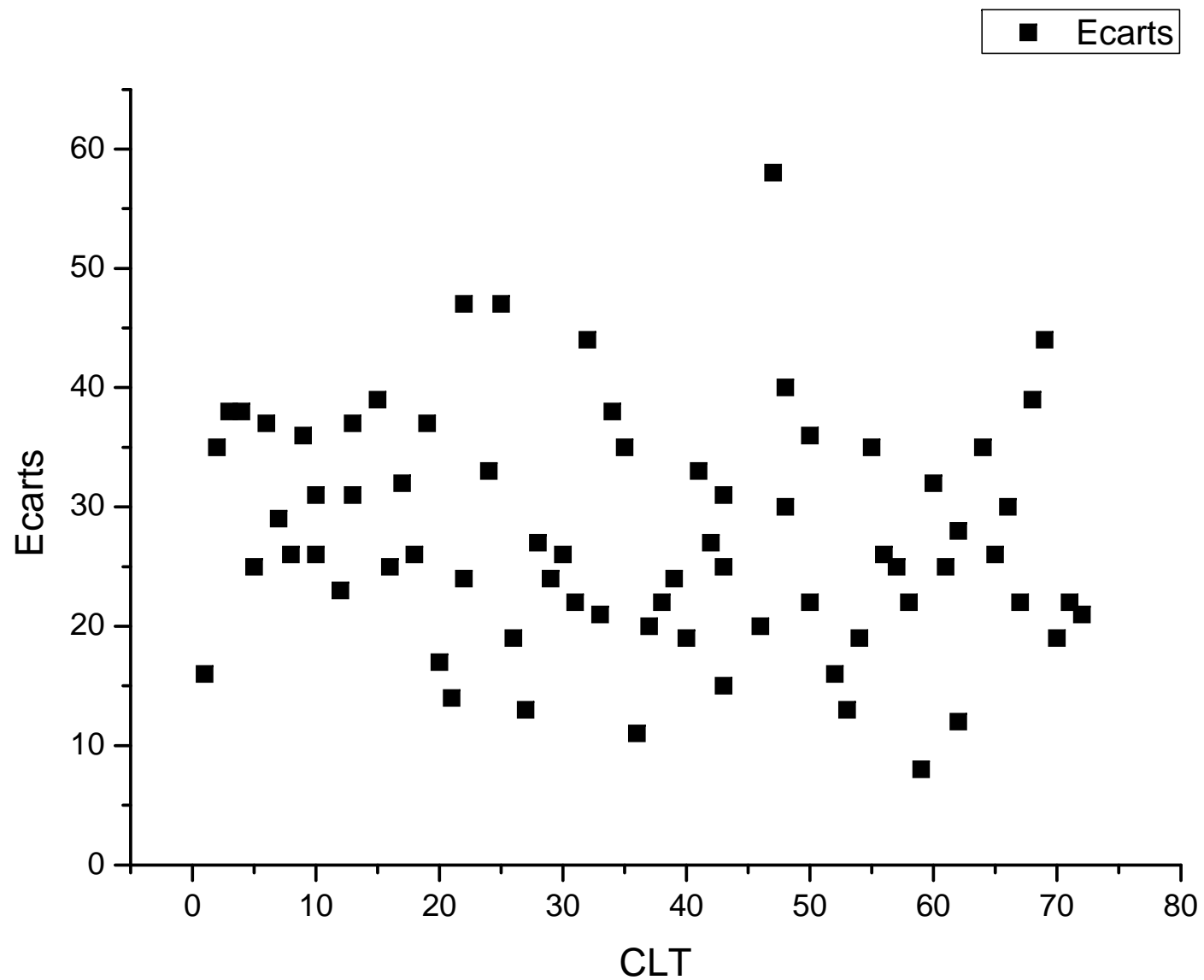
Statistiques paramétriques : moyennes et écart-types

Statistiques non paramétriques : minimum, quartiles médiane et maximum caractérisent la dispersion des valeurs

Moyenne des notes de dressage et écarts-types
dispersion des valeurs en fonction des juges et du classement considéré

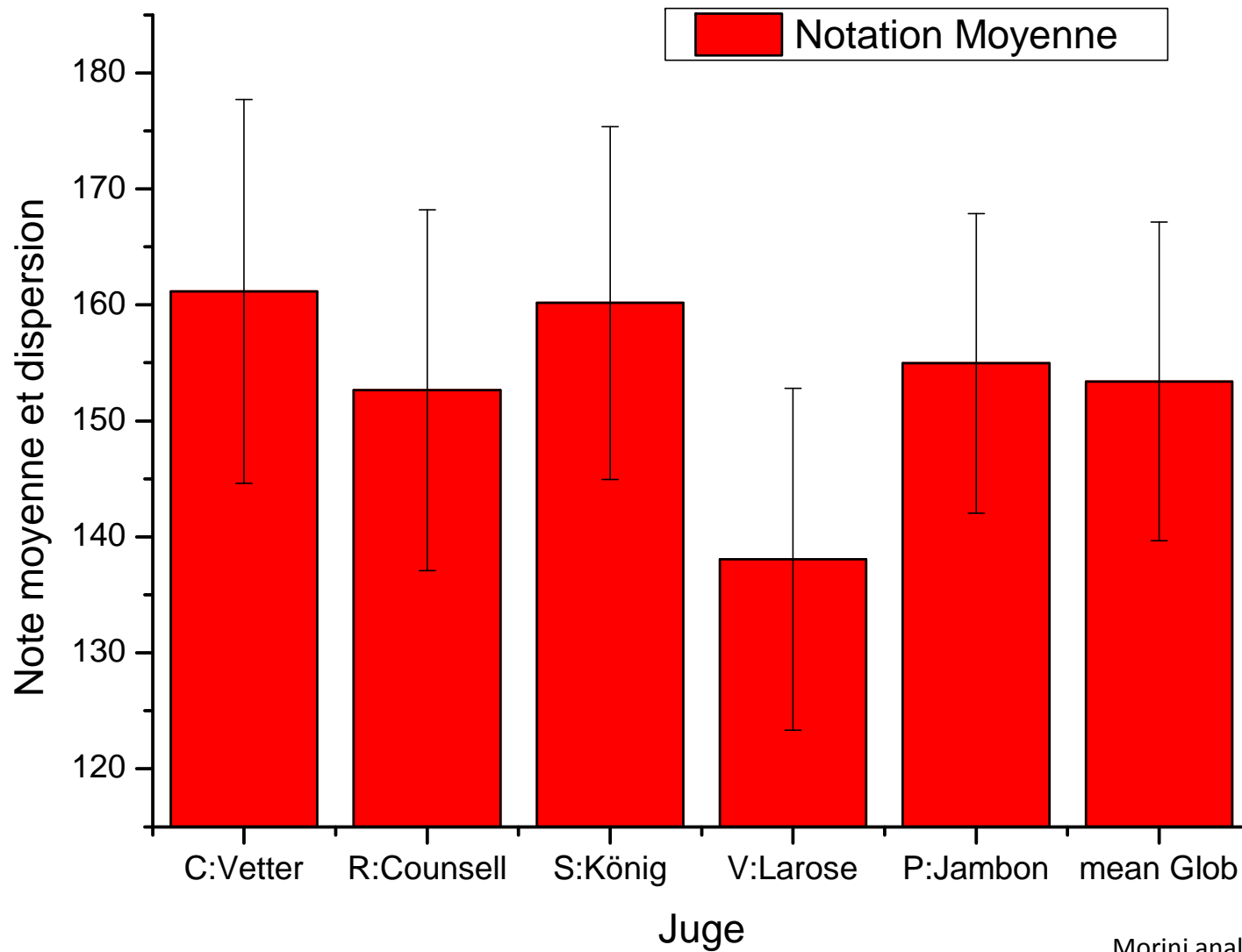


Les écarts entre note maximale et note minimale attribuées ne varient pas de façon significative avec le classement des équipes

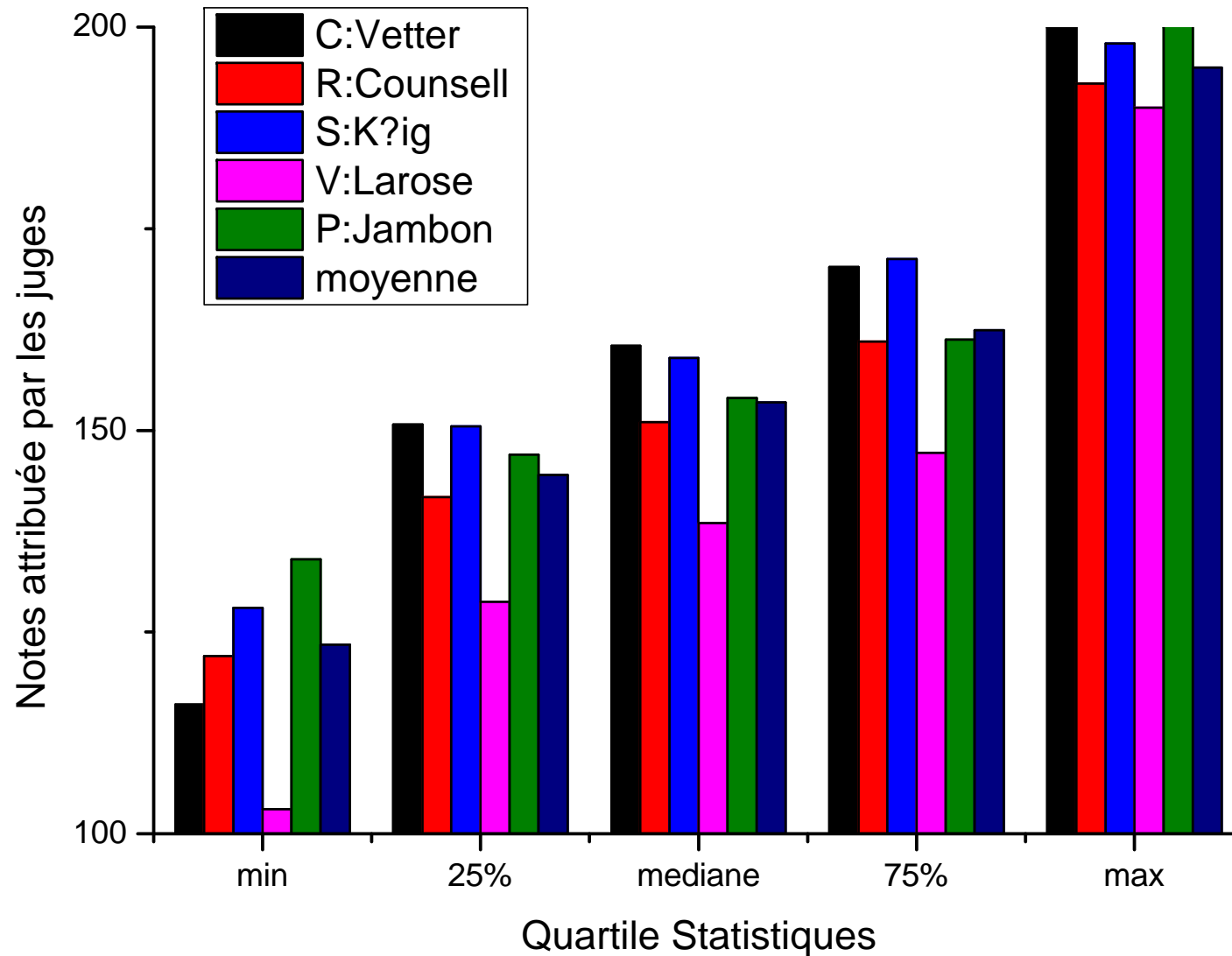


Le comportement individuel des juges

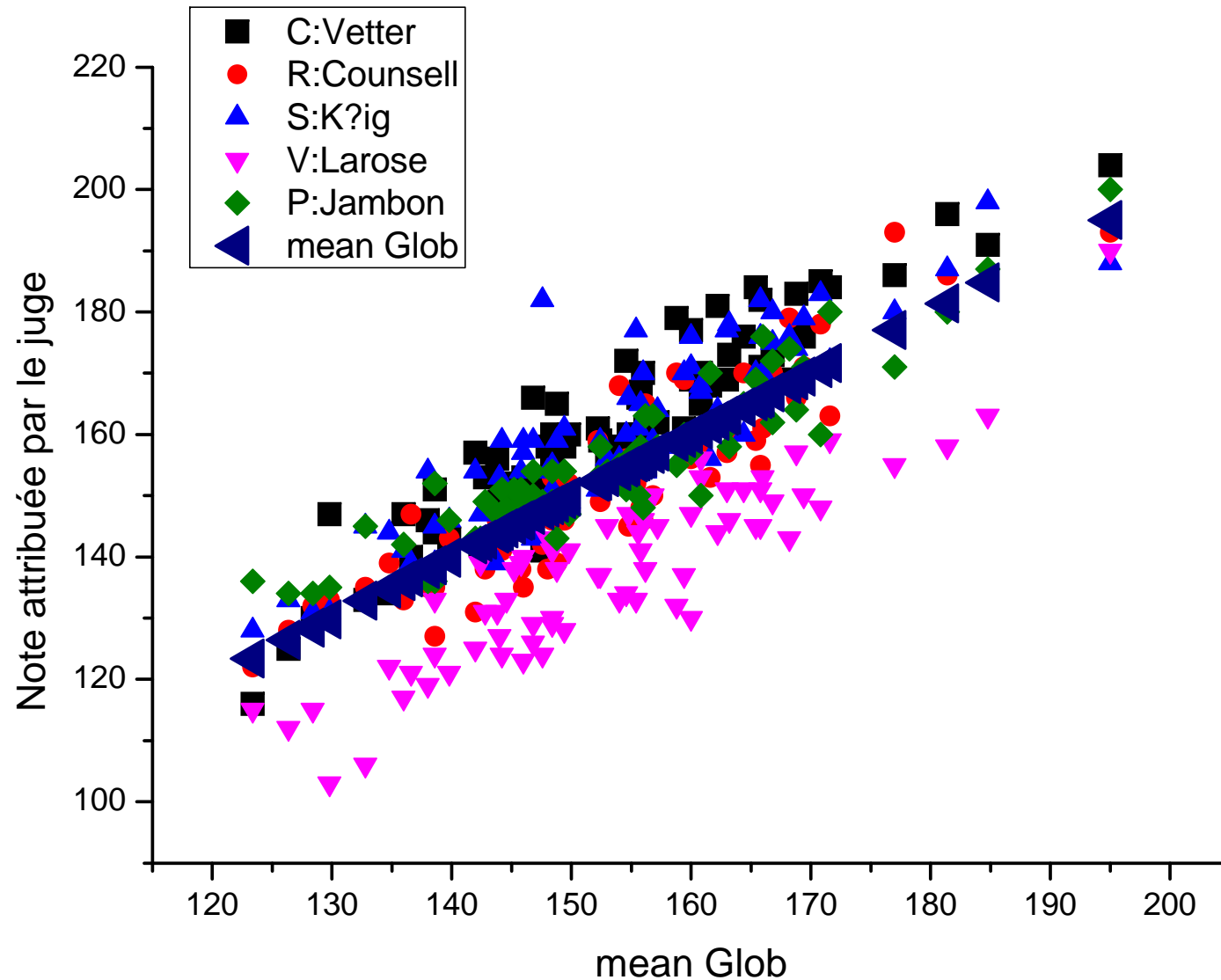
Moyennes et Dispersion (ecart-type) des notes attribuées par chaque juge



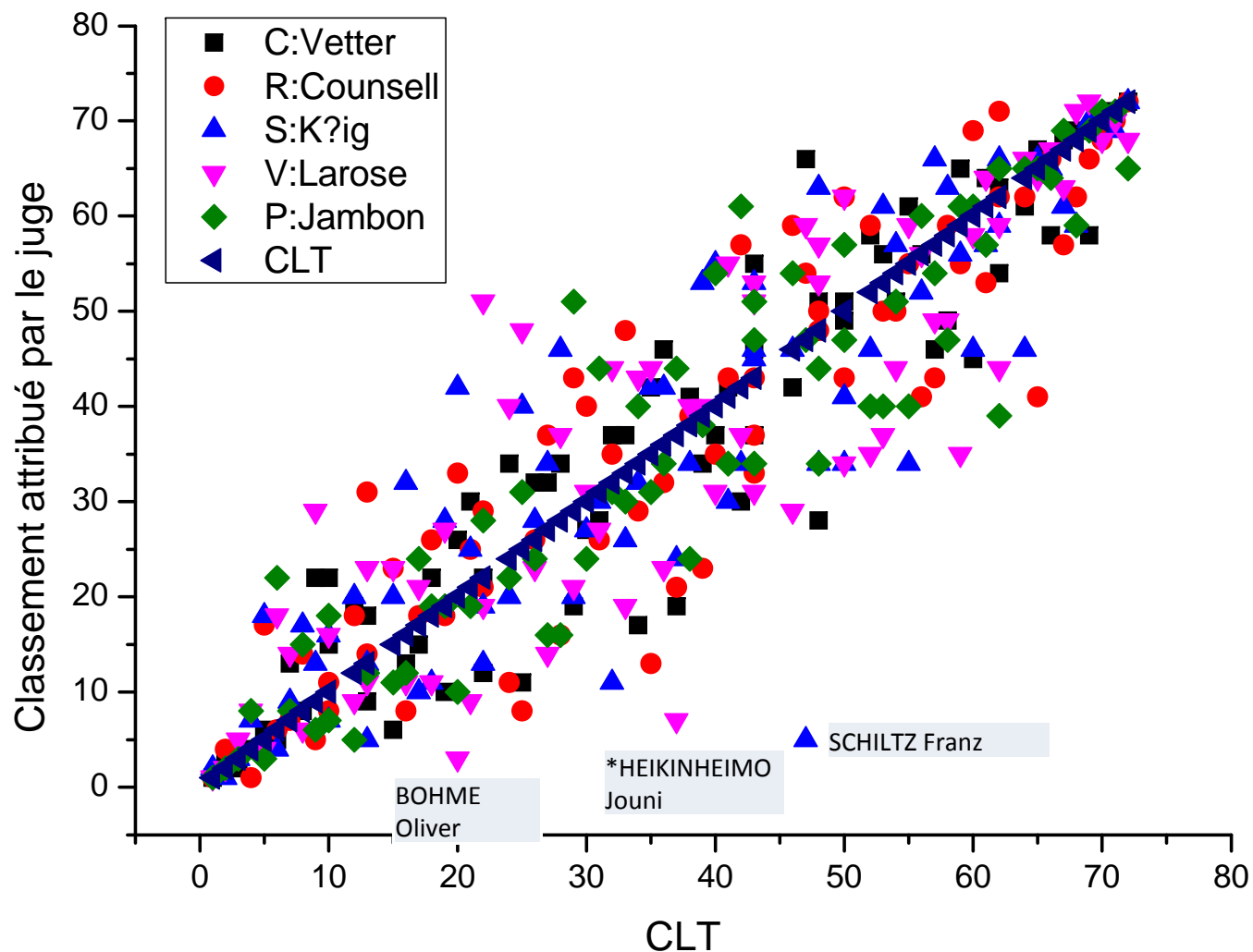
Le comportement individuel des juges : Le juge Larose a régulièrement sous noté, les juges Vetter Jambon et König ont régulièrement surnoté. Les notes du juge Counsell sont plutôt dans la moyenne. La position du juge et sa vision de la reprise influe t'elle sur la notation? Pour cela il faudrait avoir le détail des notes par figure de dressage sur les protocoles



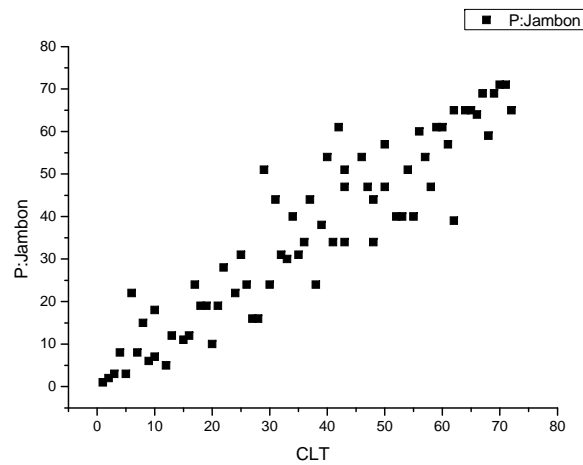
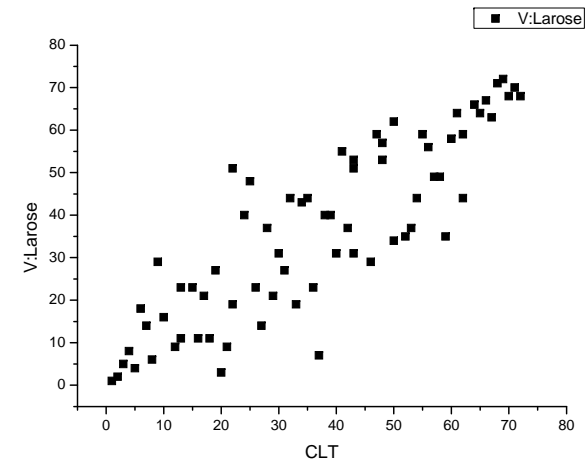
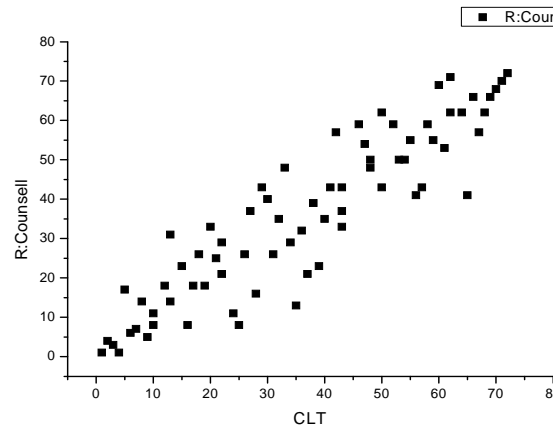
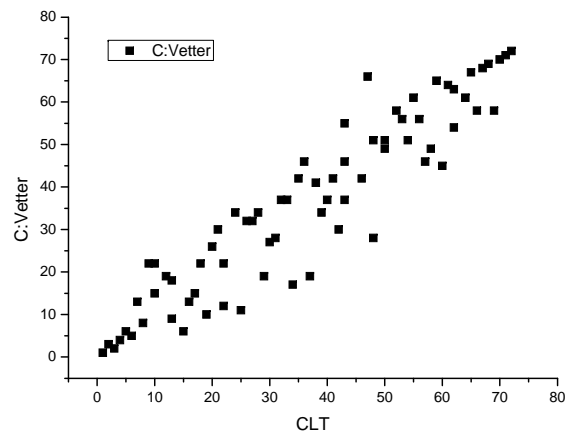
Le juge Larose a régulièrement sous noté, les juges Vetter Jambon et Konig ont régulièrement surnoté. Les notes du juge Counsell sont plus aléatoires. La position du juge et sa vision de la reprise influe t'elle sur la notation? Pour cela il faudrait avoir le détail des notes par figure de dressage sur les protocoles



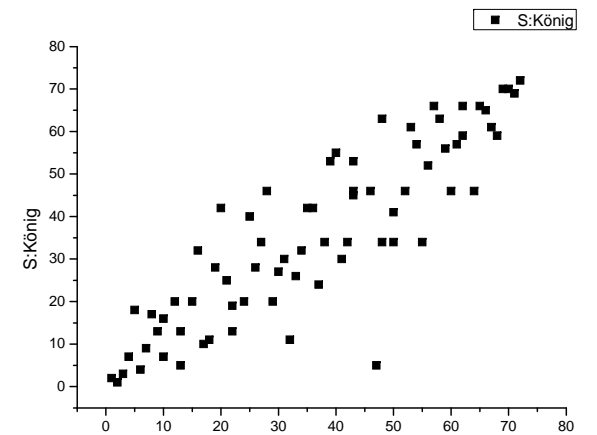
Alors que le positionnement relatif des notes paraît relativement systématique, le classement attribué par les juges est très cohérent pour le 5 premiers et les 5 derniers qui font l'unanimité sur leur sort! La dispersion des classements attribués ne varie pas beaucoup entre le 6^{ème} et le 60^{ème}. Il n'est pas observé de caractère systématique en fonction des juges.



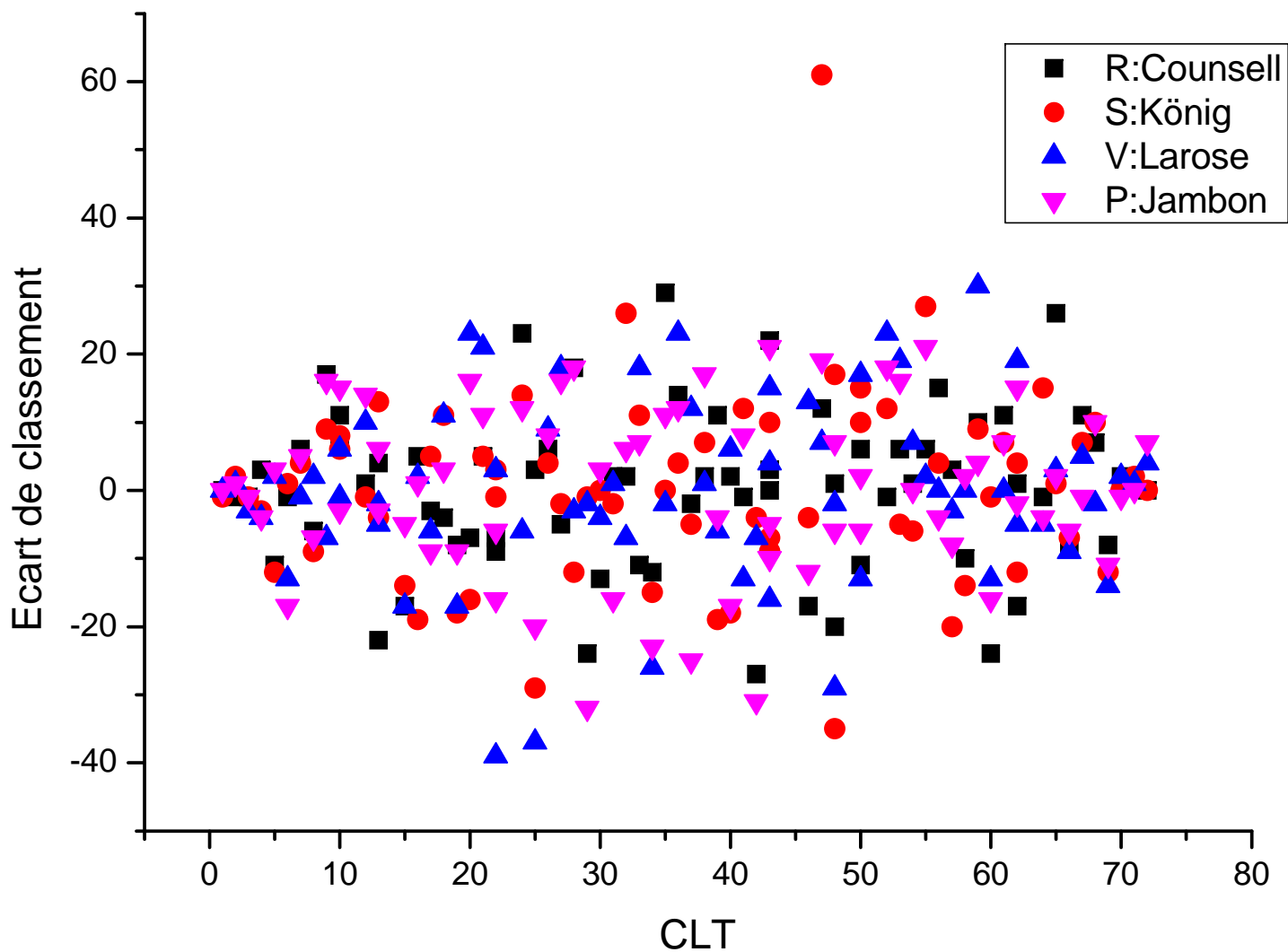
Dispersion des classements attribués par juge par rapport aux classement final du dressage
Difficile de se faire une philosophie...!
Il apparait seulement que la dispersion entre les juges est plus grande dans les classements intermédiaires



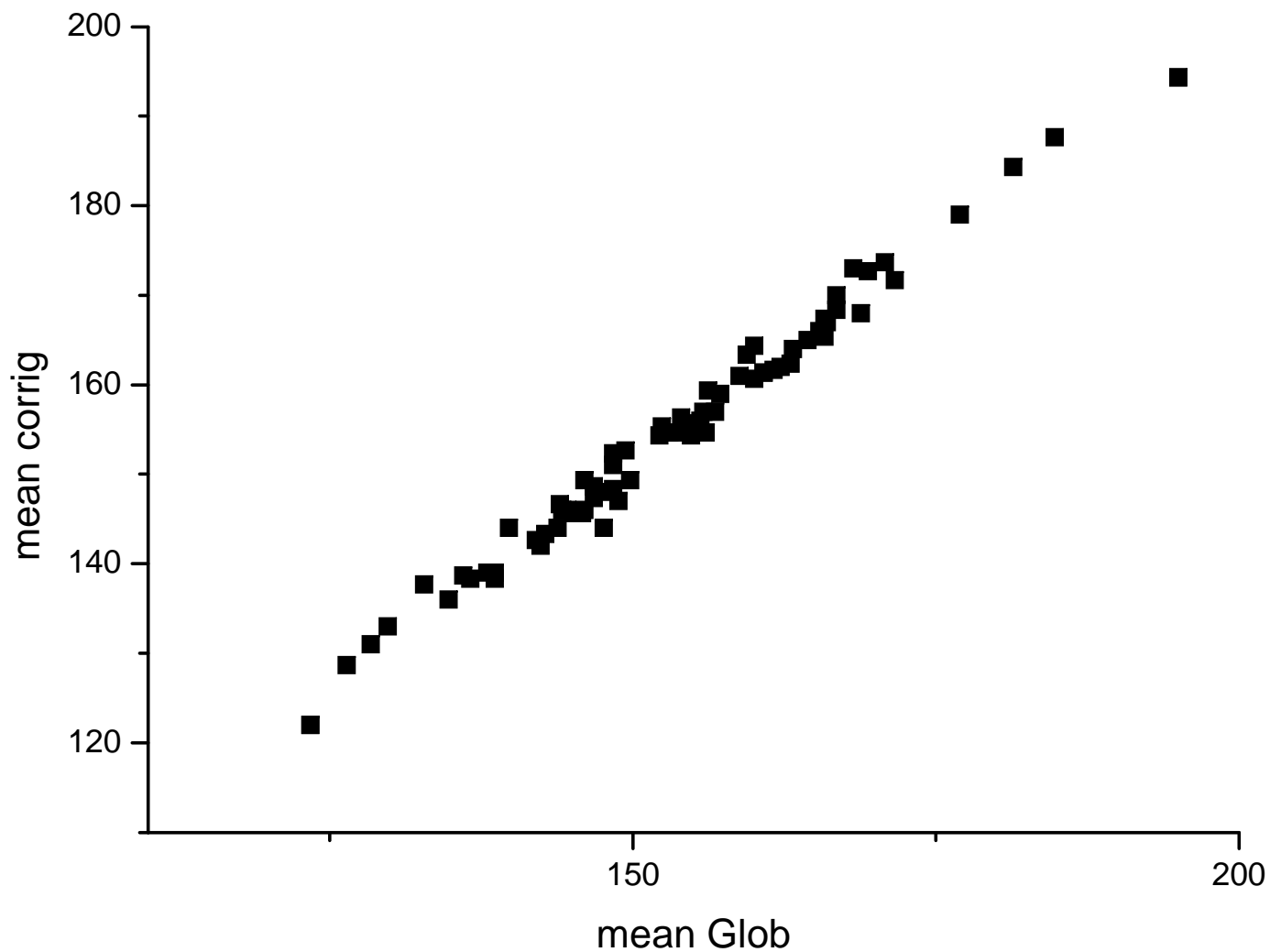
Cette planche reprend par juge les données de la planche précédente



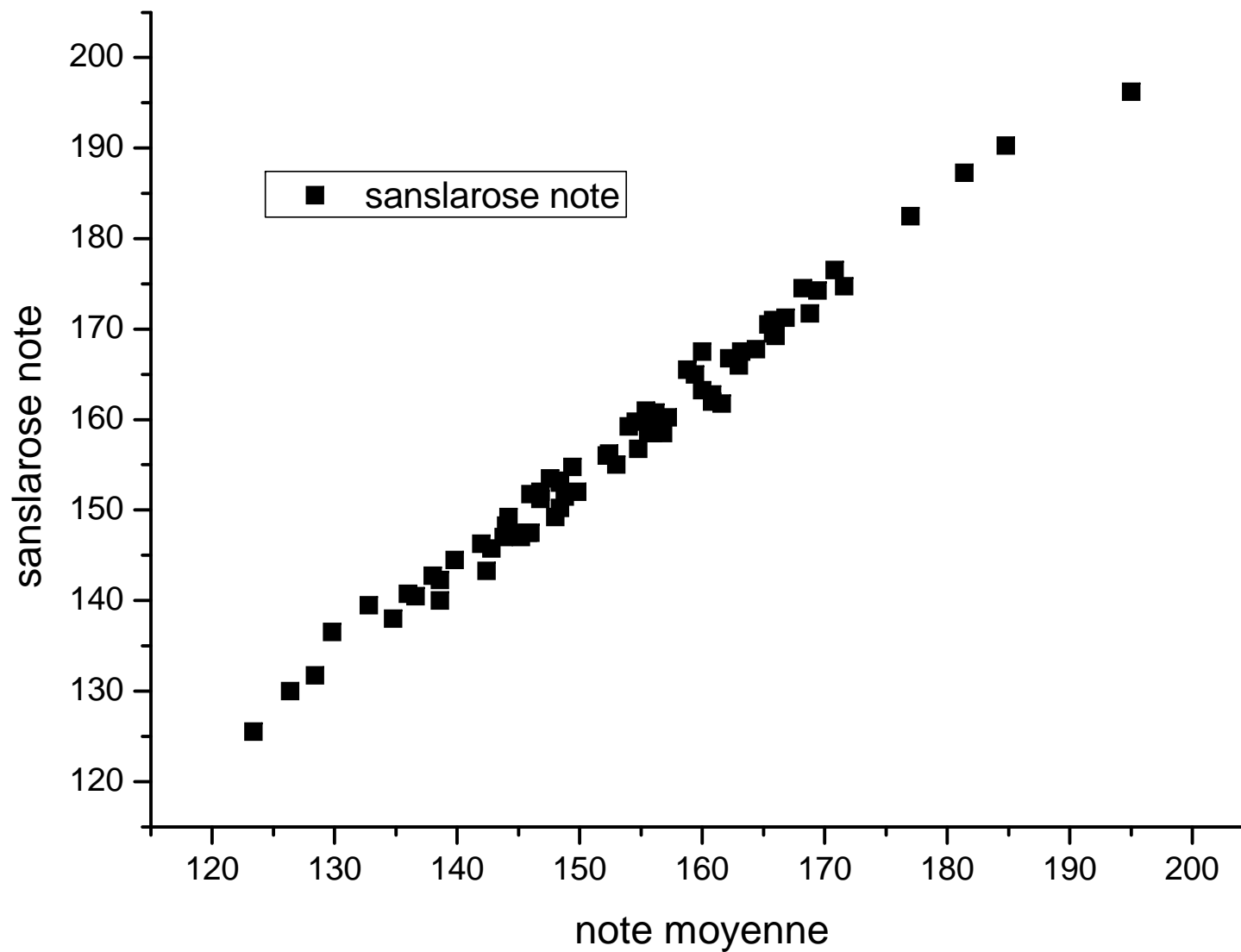
Il semble que Monsieur Vetter soit le juge de référence. Si tel est le cas, voici en fonction du classement final des équipages les écarts de classement donnés par les juges par rapport au classement attribué par Mr Vetter



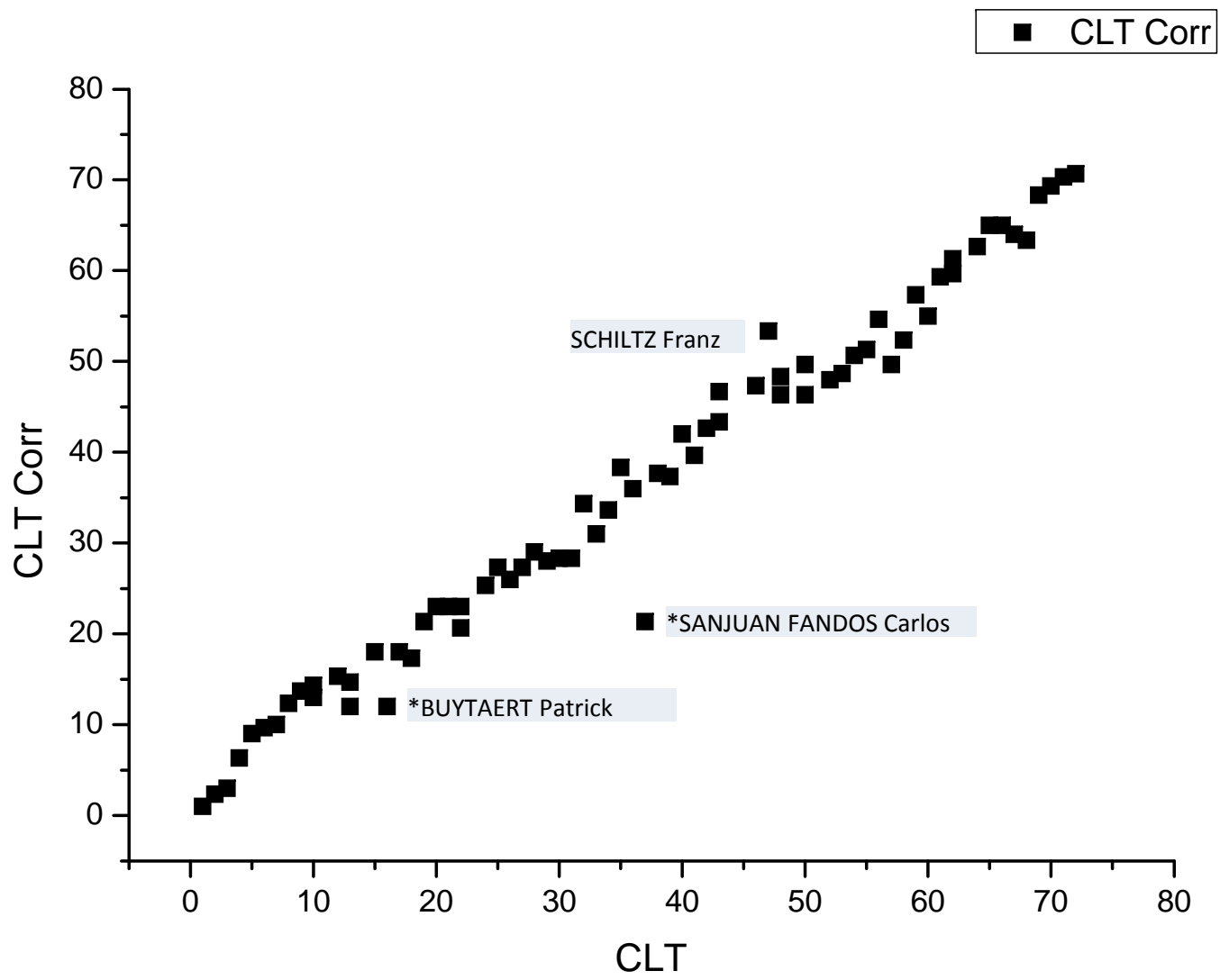
Moyenne des notes de dressage (mean glob) versus moyenne des notes après avoir retiré la meilleure et la moins bonne note. Globalement peu d'écart à la bissectrice donc peu d'impact de cette action sur la notation



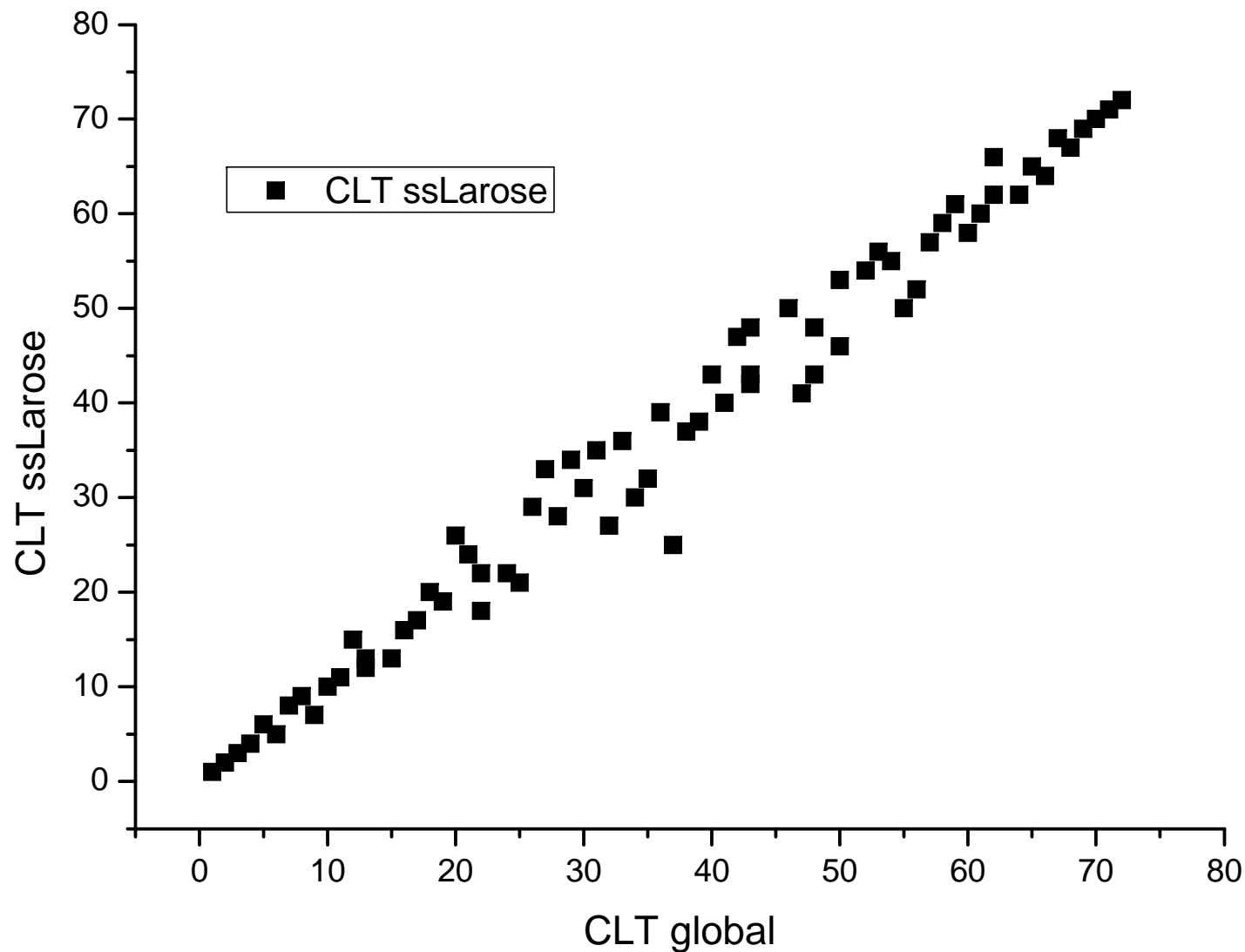
Influence du jugement de Mr Larose sur la notation



Moyenne des places attribuées (CLT) versus moyenne des Classements ayant retiré le meilleur et le moins bon. Globalement peu d'écart à la bissectrice donc peu d'impact de cette action sur la notation. Seuls 3 équipages s'écartent de la bissectrice et n'ont sans doute pas laissé un juge indifférent!



Influence du jugement de Mr Larose sur le Classement moyen
Peu d'impact entre la 1^{ère} et la 20^{ème} place et entre la 55^{ème} et la 72^{ème} place
Impact plus dispersé entre la 20^{ème} et la 55^{ème} place



Mon avis personnel sur la question

Il faut aller chercher l'origine de la variabilité des jugements dans les classements moyens dans la méthode de notation dont les incréments par points sont générateurs d'assez forte variabilité surtout dans une reprise avec 25 notes attribuées.

Si vous vous référez à la planche N°2, vous vous rendrez compte que l'écart type à la moyenne des 5 jugements est en moyenne de 10 points, et que globalement celui-ci varie peu avec le classement des équipages. Cela est assez rassurant sur la qualité et l'impartialité du jugement effectué et souligne l'efficacité du jugement mis en place malgré l'impact de la méthode de notation par points entiers sur la variabilité apparente.

Concernant l'impact individuel des juges :

Avec 5 juges sur une carrière, l'impact individuel est assez lissé. Seuls 3 cas sur 72 concurrents semblent ne pas avoir laissé un juge indifférent (planche 7) ce phénomène apparait donc limité.

Le juge Vetter est souvent qualifié de référence en matière de jugement l'étude réalisée ne permet pas de dégager un impact particulier de son jugement par rapport à celui des autres juges.

Le juge Larose a régulièrement sous noté. Il n'a pas d'impact particulier sur le classement moyen résultant de chaque équipage

Il n'est pas possible avec les notations disponibles et sur 1 seul concours de réaliser l'analyse de l'impact de la position du juge sur la carrière de dressage sur la variabilité apparente. En effet la vision de la reprise et de la qualité des allures et du tracé ne peut être évaluée de la même façon selon la position du juge.

Seul l'accès à l'ensemble des notes par figures pour l'ensemble des concurrents permettrait d'analyser ce point qui est à mon sens en couplage avec la notation par point entier à l'origine des écarts de jugements observés.

Voilà un travail qui pourrait être suggéré et entrepris à plus grande échelle et de façon systématique sur un grand nombre de concours internationaux. Peut-être cela existe-t-il déjà dans quelques cartons fédéraux.

Une telle étude systématique servirait la condition des juges bien souvent critiquée par des sentiments subjectifs que la statistique bien conduite pourrait sans doute dissiper....

Les outils d'analyse existent, par contre la disponibilité des données d'entrée est moins évidente et demanderait un peu de volontarisme pour centraliser les protocoles de dressage et leur saisie!

Ce n'est que mon avis mais la statistique dans ce domaine pourrait faire taire bien des critiques sur la rigueur et l'impartialité des méthodes de jugement sur nos carrières de dressage