

Méthodologie d'analyse statistique des mesures de pièces fabriquées en usine pour identifier et tolérer les composants géométriques critiques responsables d'une mauvaise prestation.



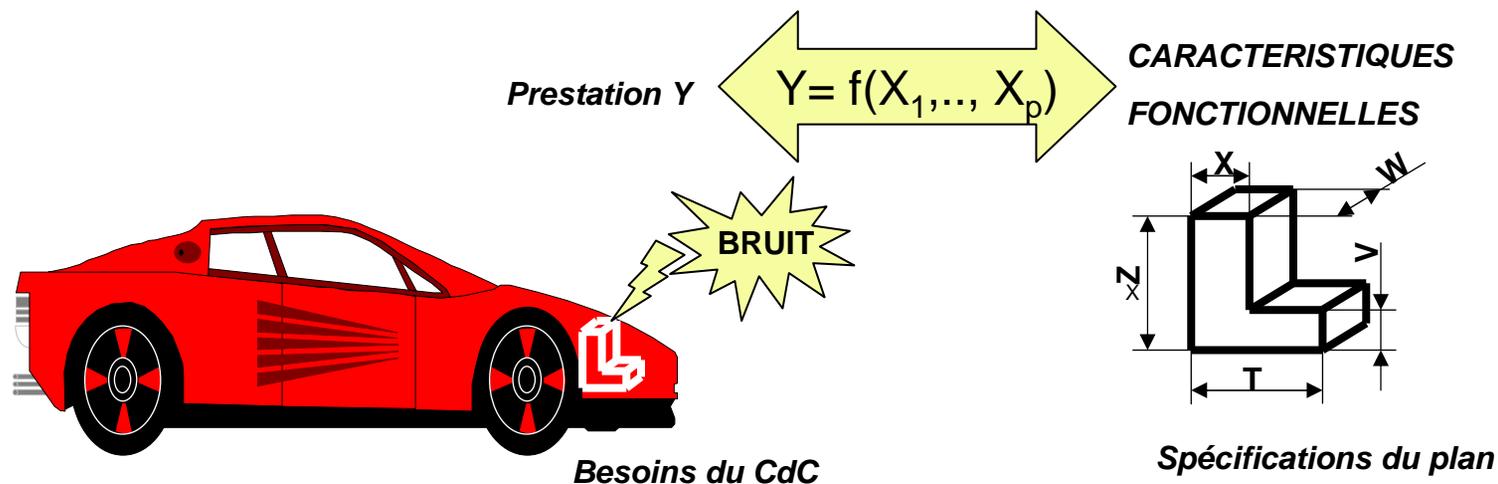
RENAULT

JMS 2009 ASIEM

T. CEMBRZYNSKI J. DEMONSANT – 24 mars 2009



- La qualité d'une prestation Y délivrée par un système physique résulte :
 - des spécifications établies par le BE (nominaux et tolérances des caractéristiques X_i des composants)
 - du respect de ces spécifications par la fabrication.



- Pour améliorer la qualité G. Taguchi préconise le Robust Engineering fondé sur les plans d'expériences pour établir une relation $Y=f(X_1, \dots, X_p)$ puis définir les nominaux et les tolérances optimales des caractéristiques X_i .

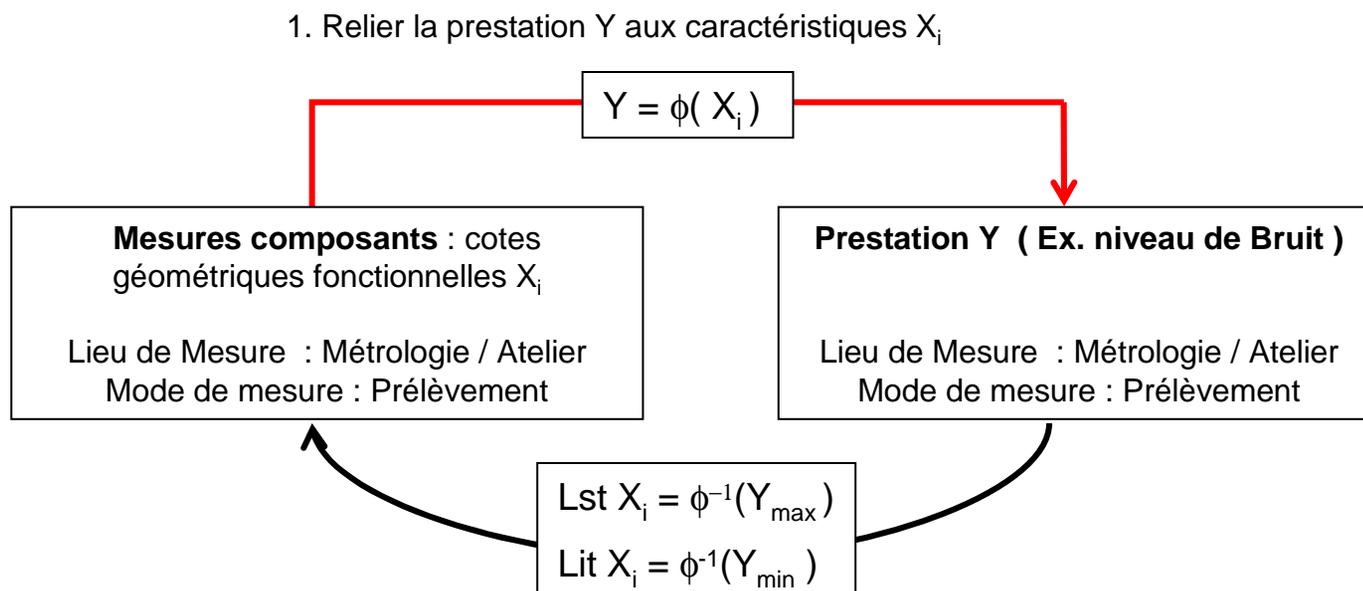
Mais les PE ne sont pas toujours réalisables.

- Il faut alors appliquer notre démarche d'analyse statistique à base de régression PLS pour identifier et optimiser les caractéristiques fonctionnelles X_i vis-à-vis de la prestation Y à partir de pièces usine.



La méthodologie comporte deux phases

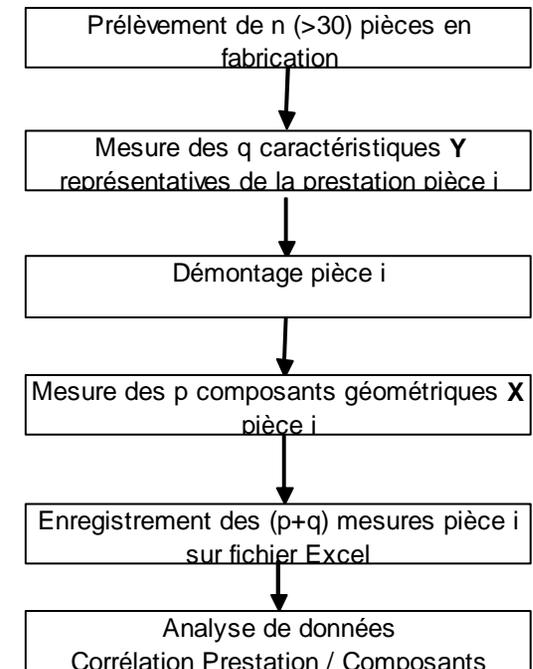
1. Lier la prestation mesurée Y (ex. un niveau de bruit sur véhicule) aux cotes géométriques responsables X_i du système : $Y = f(X_1, \dots, X_p)$
2. Identifier et tolérancer les cotes les plus critiques en regard du cahier des charges maximum Y_{\max} ou minimum Y_{\min} acceptable.



2. Spécification des cotes critiques X_i par rapport au CdC de Y



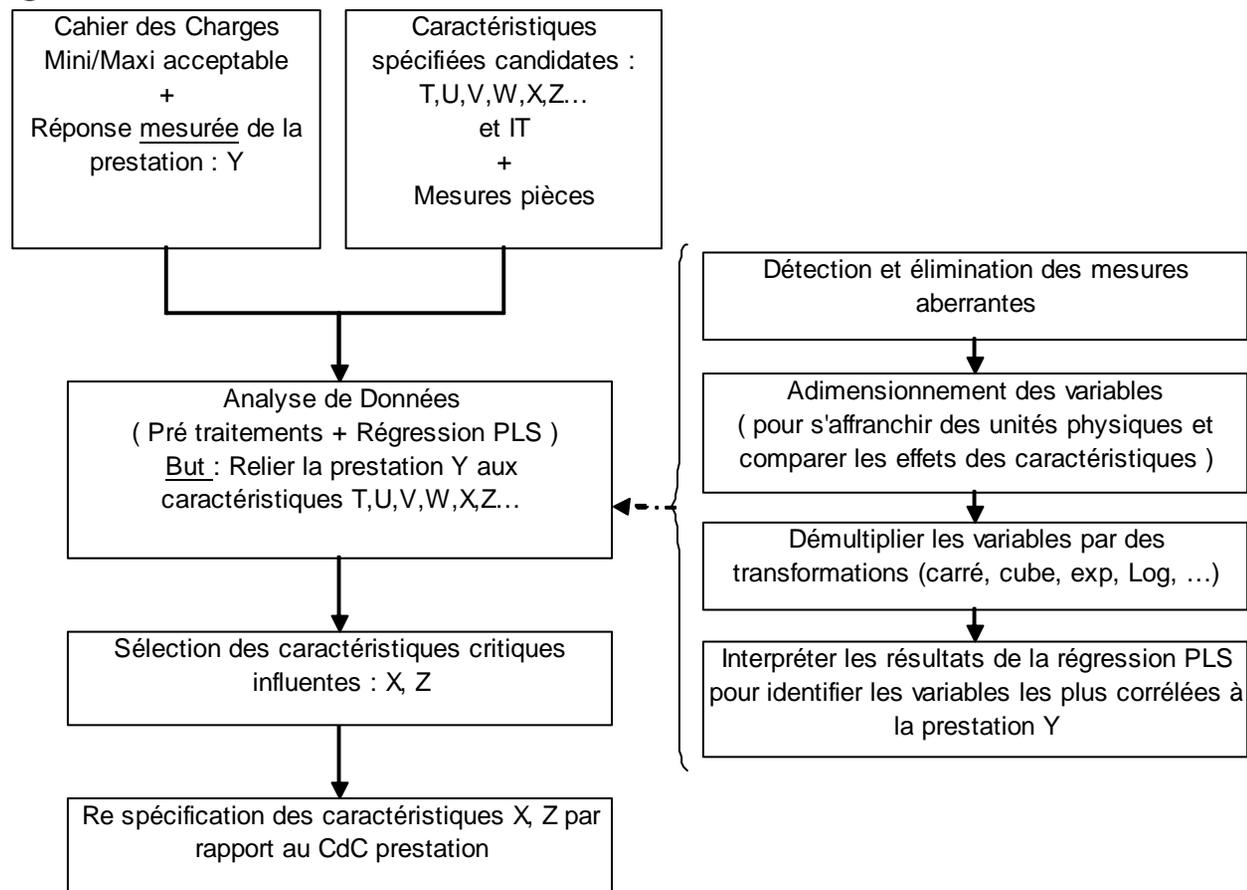
- La démarche requiert un tableau rectangulaire de n individus, $p+q$ variables où ont été mesurées simultanément :
 - q variables Y représentatives de(s) la prestation(s) à objectiver,
 - p cotes géométriques X explicatives candidates.
- Taille d'échantillonnage requise : $n > \min(30, p+q+1)$
- Les limites de spécifications inférieures L_{it} , et supérieures L_{st} des cotes X sont documentées.
- Le CdC de la prestation Y est connu.





La démarche d'analyse

- Fondée sur la modélisation linéaire PLS (Partial Least Squares) de la prestation Y en fonction des cotes candidates X.
 - ➔ *Sélection des paramètres critiques influents,*
 - ➔ *(Re)spécification des paramètres critiques par rapport au Cahier des Charges*





Etape 1 : Détection et élimination des données aberrantes

- Différentes méthodes sont possibles, comme l'ajustement de la loi de distribution selon le symbole de cotation ISO (Loi Normale, Loi de Défaut de Forme, Loi de Rayleigh NF60-181) et filtrage des mesures à faible probabilité (0,1%, 99,9%) ...
- Solution « simple » : hypothèse de normalité de la variable $X_i = N(\mu, \sigma^2)$, détection des mesures aberrantes par Intervalle de Confiance à 99,8%.
 - Estimer s par le moyenne des étendues $w_i / 1,128$ de 2 mesures consécutives)

| Libellé | Variables explicatives X composants géométriques | | | Variables Prestation Y à corrélérer avec X |
|------------|--|------------|------------|--|
| | X1 | X2 | X3 | Y1 |
| Lit | -1 | 9,5 | 8 | |
| Lst | 1 | 10,5 | 10 | |
| HCPP | 2 | 2 | 2 | |
| 1 | 2,1 | 10,08 | 9,29 | 20,34 |
| 2 | 0,23 | 9,99 | 8,95 | 14,03 |
| 3 | -0,35 | 9,85 | 8,8 | 9,38 |
| 4 | -0,19 | 9,83 | 9,5 | 15,45 |
| 5 | -0,16 | 10,05 | 9,52 | 21,26 |
| 6 | -0,94 | 9,95 | 8,98 | 13,15 |
| 7 | 0,09 | 9,72 | 8,6 | 6,38 |
| 8 | 0,07 | 10,27 | 8,68 | 24,42 |
| ... | -0,1 | 9,83 | 9,26 | 18,26 |
| 30 | -0,28 | 9,89 | 8,8 | 10,11 |
| min | -0,94 | 9,72 | 8,6 | 6,38 |
| max | 2,1 | 10,27 | 9,52 | 24,42 |
| moyenne | 0,047 | 9,946 | 9,038 | 14,778 |
| ecart type | 0,47478329 | 0,15847888 | 0,32013396 | 5,70906842 |

$$\bar{x} \rightarrow \mu \quad s \rightarrow \sigma$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad s = \frac{1}{n-1} \sum_{i=2}^n \frac{w_i}{1,128} \quad \text{où } w_i = |x_i - x_{i-1}|$$

Si $\langle x_i \notin [\bar{x} - T_{99,9\%,n-1} \cdot s, \bar{x} + T_{99,9\%,n-1} \cdot s] \rangle \rightarrow x_i$ aberrant

$T_{99,9\%,n-1} > 3,09$ quantile 99,9% de la Loi de Student à $n-1$ ddl

Mesure $X_1 = 2,1$ aberrante

- Faire diagnostiquer les mesures aberrantes par le service mesure.



Etape 2 : Adimensionnement des variables X et Y

- Adimensionnement des variables X et Y de manière à s'affranchir des unités physiques et à pouvoir comparer les effets α_i de chaque variable X'_i sur Y.

$$x'_i = \frac{(x_i - \bar{x})}{S}$$

| Libellé | Adimensionnement des Variables explicatives X composants géométriques | | | Variable Prestation Y à corrélérer avec X |
|-------------------|--|-------------------|-------------------|---|
| | X1' | X2' | X3' | Y1' |
| Lit | -4 | 9,5 | 8 | |
| Lst | 4 | 10,5 | 10 | |
| HCPP | 2 | 2 | 2 | |
| 1 | | 0,8455 | 0,7872 | 0,9742 |
| 2 | 1,3125 | 0,2776 | -0,2749 | -0,131 |
| 3 | -0,5392 | -0,6058 | -0,7434 | -0,9455 |
| 4 | -0,0284 | -0,732 | 1,4431 | 0,1177 |
| 5 | 0,0674 | 0,6562 | 1,5056 | 1,1354 |
| 6 | -2,4227 | 0,0252 | -0,1812 | -0,2852 |
| 7 | 0,8655 | -1,4261 | -1,3682 | -1,471 |
| 8 | 0,8017 | 2,0444 | -1,1183 | 1,6889 |
| ... | 0,2589 | -0,732 | 0,6935 | -0,2659 |
| 30 | -0,3157 | -0,3534 | -0,7434 | -0,8176 |
| min | -2,4227 | -1,4261 | -1,3682 | -1,471 |
| max | 1,3125 | 2,0444 | 1,5056 | 1,6889 |
| moyenne | 0 | -0,00004 | 0 | 0 |
| ecart type | 1,04363918 | 1,00000174 | 1,00730275 | 0,99999551 |

Les mesures sont sensiblement à la même échelle après le centrage et la réduction.



Etape 3 : Déploiement des variables X adimensionnées pour rechercher d'éventuels effets non linéaires.

- Déploiement de transformations des variables X_i en $\exp(X_i)$, X_i^2 , X_i^3 ... pour rechercher d'éventuels effets non linéaires, (quadratique, cubique, exponentiel) de l'impact sur la variable de prestation Y.

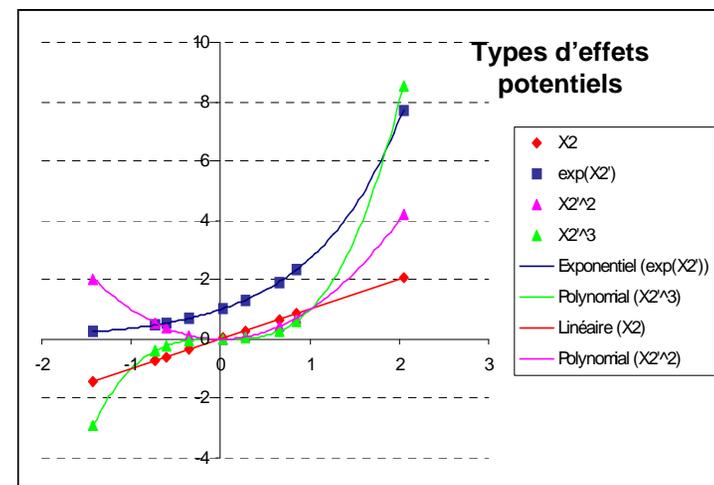
| Libellé | Adimensionnement des Variables explicatives X composants géométriques | | | Déploiement des tranormées de X | | | | | | | | | Variable Prestation Y à corrélér avec X |
|------------|--|------------|------------|---------------------------------|------------|-------------|------------|------------|------------|------------|------------|------------|---|
| | X1' | X2' | X3' | exp(X1') | X1'^2 | X1'^3 | exp(X2') | X2'^2 | X2'^3 | exp(X3') | X3'^2 | X3'^3 | Y1' |
| Lit | -4 | 9,5 | 8 | | | | | | | | | | |
| Lst | 4 | 10,5 | 10 | | | | | | | | | | |
| HCPP | 2 | 2 | 2 | | | | | | | | | | |
| 1 | | 0,8455 | 0,7872 | | | | 2,3291 | 0,7149 | 0,6044 | 2,1972 | 0,6197 | 0,4878 | 0,9742 |
| 2 | 1,3125 | 0,2776 | -0,2749 | 3,7155 | 1,7227 | 2,261 | 1,32 | 0,0771 | 0,0214 | 0,7596 | 0,0756 | -0,0208 | -0,131 |
| 3 | -0,5392 | -0,6058 | -0,7434 | 0,5832 | 0,2907 | -0,1568 | 0,5456 | 0,367 | -0,2223 | 0,4755 | 0,5526 | -0,4108 | -0,9455 |
| 4 | -0,0284 | -0,732 | 1,4431 | 0,972 | 0,0008 | 0 | 0,4809 | 0,5358 | -0,3922 | 4,2338 | 2,0825 | 3,0053 | 0,1177 |
| 5 | 0,0674 | 0,6562 | 1,5056 | 1,0697 | 0,0045 | 0,0003 | 1,9275 | 0,4306 | 0,2826 | 4,5069 | 2,2668 | 3,4129 | 1,1354 |
| 6 | -2,4227 | 0,0252 | -0,1812 | 0,0887 | 5,8695 | -14,22 | 1,0255 | 0,0006 | 0 | 0,8343 | 0,0328 | -0,0059 | -0,2852 |
| 7 | 0,8655 | -1,4261 | -1,3682 | 2,3762 | 0,7491 | 0,6483 | 0,2402 | 2,0338 | -2,9003 | 0,2546 | 1,872 | -2,5612 | -1,471 |
| 8 | 0,8017 | 2,0444 | -1,1183 | 2,2293 | 0,6427 | 0,5153 | 7,7245 | 4,1796 | 8,5447 | 0,3268 | 1,2506 | -1,3985 | 1,6889 |
| ... | 0,2589 | -0,732 | 0,6935 | 1,2955 | 0,067 | 0,0174 | 0,4809 | 0,5358 | -0,3922 | 2,0007 | 0,4809 | 0,3335 | -0,2659 |
| 30 | -0,3157 | -0,3534 | -0,7434 | 0,7293 | 0,0997 | -0,0315 | 0,7023 | 0,1249 | -0,0441 | 0,4755 | 0,5526 | -0,4108 | -0,8176 |
| min | -2,4227 | -1,4261 | -1,3682 | 0,0887 | 0,0008 | -14,22 | 0,2402 | 0,0006 | -2,9003 | 0,2546 | 0,0328 | -2,5612 | -1,471 |
| max | 1,3125 | 2,0444 | 1,5056 | 3,7155 | 5,8695 | 2,261 | 7,7245 | 4,1796 | 8,5447 | 4,5069 | 2,2668 | 3,4129 | 1,6889 |
| moyenne | 0 | -0,00004 | 0 | 1,45104444 | 1,04963333 | -1,21844444 | 1,67765 | 0,90001 | 0,5502 | 1,60649 | 0,97861 | 0,24315 | 0 |
| ecart type | 1,04363918 | 1,00000174 | 1,00730275 | 0,94572252 | 1,48778812 | 3,58414229 | 1,96329787 | 0,97176911 | 2,52012411 | 1,30779157 | 0,81474586 | 1,41207644 | 0,99999551 |



Déploiement de X_1 en $\exp(X_1)$, X_1^2 , X_1^3 ...

⇒ 12 variables explicatives potentielles

⇒ 1 donnée manquante en X1





Etape 4 : Modélisation PLS $Y=\beta X$ pour relier la prestation Y aux composants X

- La Régression Multiple (OLS) est le plus simple des modèles linéaires reliant des variables explicatives (numériques) X ou "prédicteurs", à une ou plusieurs variables « à expliquer » Y ($Y=\beta X+\varepsilon$) ; mais elle souffre de graves défauts :
 - *Incapacité à prendre en compte les données manquantes, ce qui conduit souvent le praticien à rejeter beaucoup d'observations incomplètes qui contiennent pourtant de l'information utile dans les champs renseignés.*
 - *Grande sensibilité à la collinéarité entre les variables explicatives X : la collinéarité exacte rend la régression multiple impossible, et la collinéarité approchée la rend numériquement instable et fait perdre l'interprétation des coefficients β .*
 - *Indétermination lorsque le nombre d'observations n est inférieur au nombre des variables explicatives p , situation pourtant fréquemment rencontrée en pratique.*
- La régression PLS est la technique de modélisation prédictive dans les situations où les p prédicteurs X sont plus nombreux que les n observations, fortement corrélés, voire avec des données manquantes.
- Mais il n'existe pas de méthode de sélection de variables "automatique" "pas à pas", on obtient après déploiement un modèle à $4p$ coefficients !!



La sélection s'appuie :

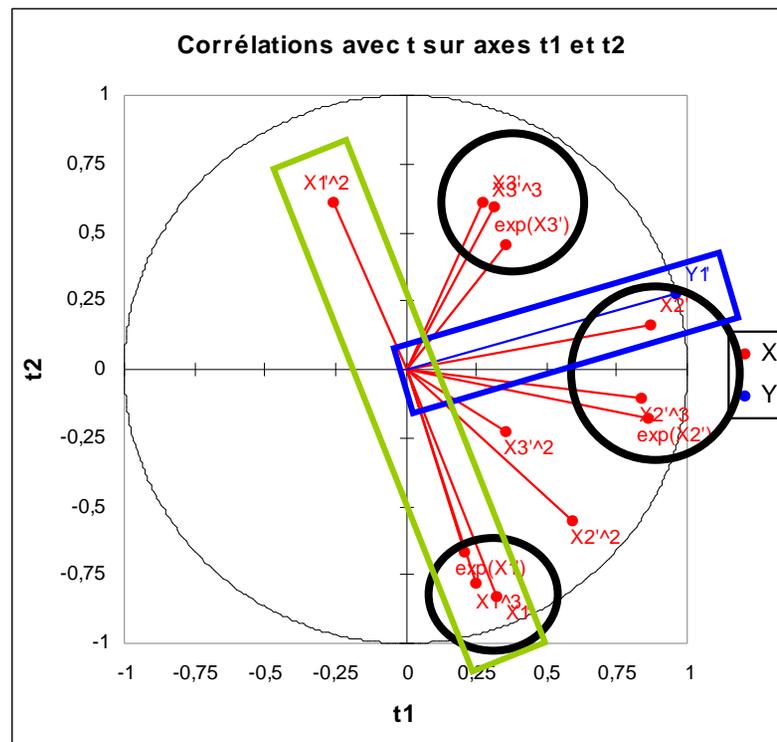
- Sur la significativité des corrélations $\rho(X_i, Y)$ fondée sur le test Fishérien de la régression linéaire $Y = aX_i + b$;

$$(n-2) \times \frac{\rho_{X,Y}^2}{1-\rho_{X,Y}^2} \approx F_{1,n-2}$$

- Sur l'analyse « visuelle » des corrélations des variables X_i, Y dans le plan (T_1, T_2) de la régression PLS ;

- Sur les « gros » coefficients α_i sélectionnés par intervalles de confiance selon la valeur de k
 - $k > 1,65 \rightarrow$ variable très influente
 - $1,04 < k < 1,65 \rightarrow$ variable influente
 - $k < 1,04 \rightarrow$ variable non influente

$$k = \frac{\alpha_i - \bar{\alpha}}{s_{\alpha}}$$



Indépendance entre X et Y

Forte liaison entre X et Y



Bruit de GEP (sirénement) : contexte et enjeux

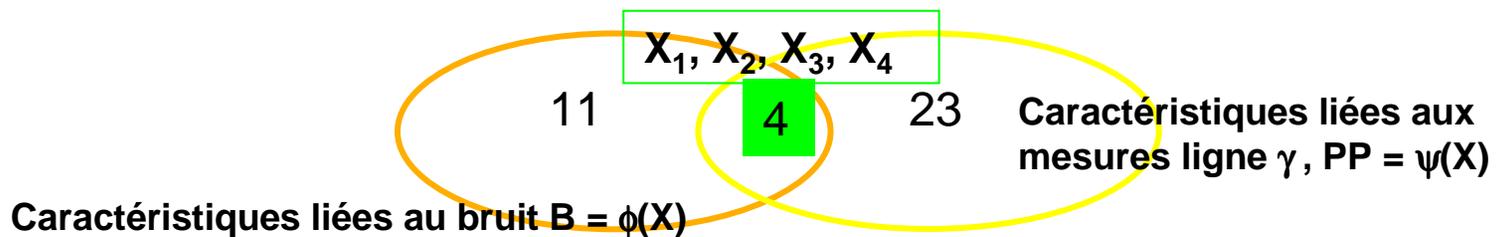
- Le GEP est une pompe hydraulique à engrenages qui assiste la direction ; il génère un bruit de sirénement.
- Pour éviter de livrer des véhicules bruyants ; après une campagne de mesures sur véhicule série, on a défini une limite d'acceptation LSC par carte de contrôle sur le paramètre d'accélérométrie γ (g) mesuré à 100% chez le fournisseur.
 - *Rejet du GEP si $\gamma > LSC$ soit environ 1% de la production.*
- Compte tenu du coût du contrôle Qualité, nous avons engagé avec le fournisseur une démarche de recherche des caractéristiques fonctionnelles influentes sur le bruit.
- Enjeux :
 - *Mieux spécifier (conception des plans)*
 - *Mieux surveiller (optimiser le plan de surveillance).*





Résolution et résultat n°1 : Optimisation de la surveillance

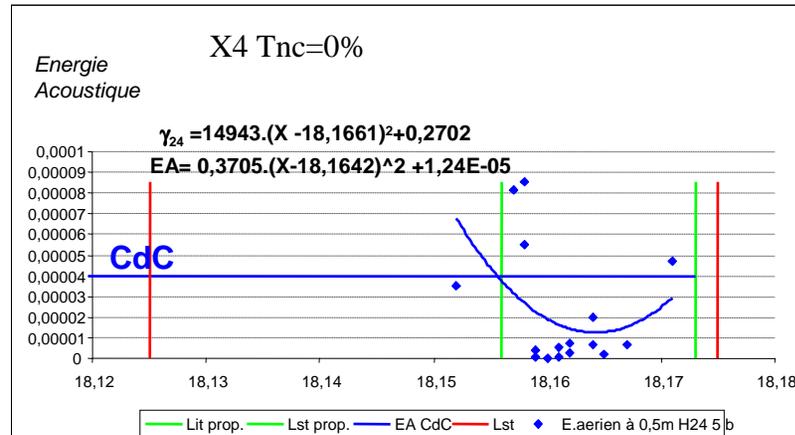
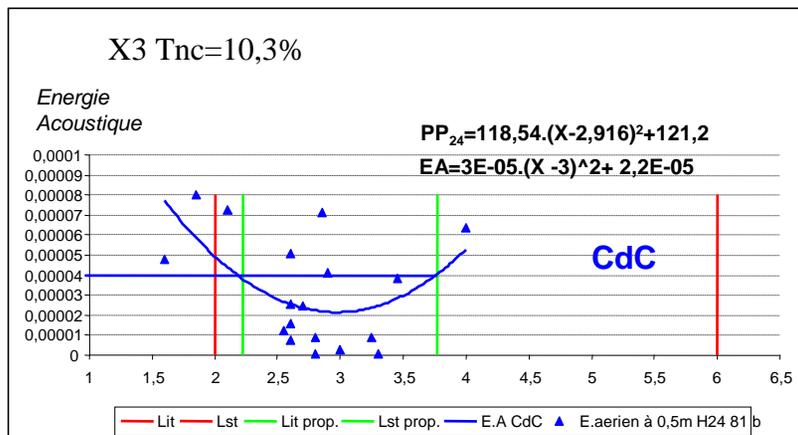
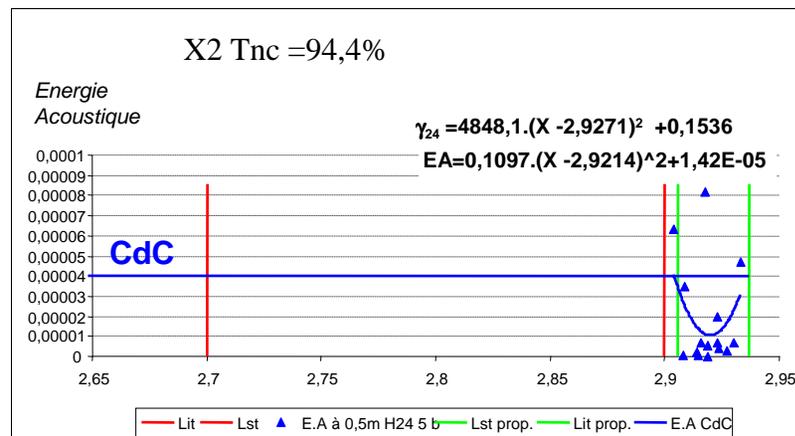
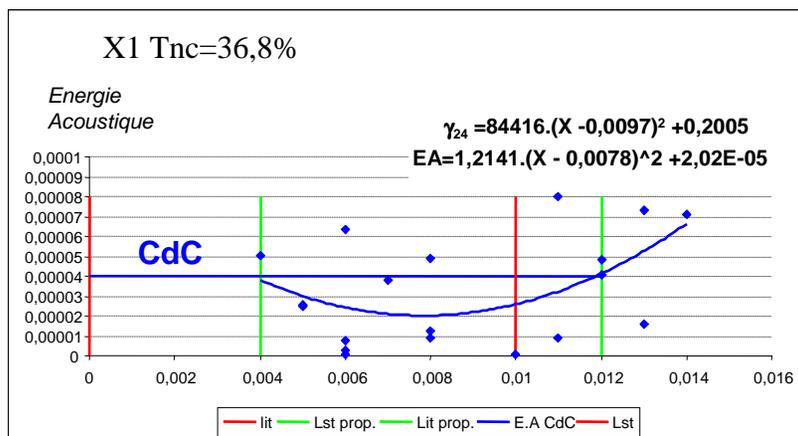
- Les experts métiers ont retenu 51 caractéristiques pièces (dentures, paliers, carters...) sources potentielles de bruit.
 - Le fournisseur a mesuré sur 19 GEP de série :
 - en ligne (100%), les Pressions Pulsées (N/m^2) et les Accélérométries γ (g),
 - les bruits aériens (dB) sur banc convertis en Energies Acoustiques (Pa^2) :
 $EA=4E-10 \times 10^{dB/10}$
 - les 51 caractéristiques fonctionnelles pièces en métrologie.
 - La régression PLS, a permis de relier les réponses (bruit, mesures ligne γ , PP) aux caractéristiques influentes X et/ou à leurs transformées.
1. Sélection PLS des 4 caractéristiques les plus impactantes sur le bruit et sur les mesures lignes γ (g) et de Pression Pulsée (N/m^2).



Pour les 4 caractéristiques, on a optimisé la surveillance et redéfini les limites de contrôle sur les mesures ligne γ et Pressions Pulsées (contrôlées à 100%) à partir du CdC bruit acceptable.



Résultat n°2 : Affiner la conception du GEP (Paramètre Design: recentrage de nominal & Tolerance Design: réduction d'IT)



IT initiaux et Propositions d'IT



Conclusion Questions

FIN

QUESTIONS