

# Cycle Formalisation de la sélection (novembre 2023)

①

Suite de la formalisation de la comparaison d'octobre 2023.

2 types de variables  $X$  sur lesquelles ont fait les comparaisons

- $X$  scalaire, comme la conductivité hydraulique. La comparaison se fait sur la proximité des valeurs. Il faudra néanmoins trouver une façon de comparer les scalaires sur une échelle qui va de 0 à 1 comme la USE.

on pourrait penser par exemple à  $1 - \frac{X^p - x_i}{\max(X^p, x_i)}$

Cette variable doit tendre vers 1 quand  $X^p = x_i$   
et vers 0 quand  $X^p \gg x_i$   
et quand  $x_i \gg X^p$

- $X$  chimique, comme le débit ou la recharge. La comparaison se fait sur l'USE ou l'USE log (à préférer pour les débits) comme formalisé précédemment.

Comparer les critères et sélectionner les seuls BV et arriver à leur proche par un BVcible ( $BV^p, T^p$ )  
et des BVs + Arrive  $\{(BV_i, T_i)\}_{i \in \llbracket 1, N \rrbracket}$

Algorithme

Par formalisation, les critères de comparaison sur actifs  $T_{i \in \llbracket 1, N \rrbracket}$

ou  $N$  est le nombre de variables comparées -

Si on utilise les débits, la recharge et les conductivités hydrauliques,  $N = 3$ .

## ① Méthode des "seuls indépendants".

Par chacun des variables  $x_i$ , on définit un seul  $T_i^+$

On sélectionne les simulations pour lesquelles  $T_i > T_i^+$

$N^p$  simulations sont ainsi sélectionnées.

-  $\Delta$  Il y a peut-être aussi une dépendance au jour cible -

$P$   $\nearrow$  opposant par le BV cible et l'arrivée.  
 $\nwarrow$  indice pour les BVs de la base de données.

On sélectionne ensuite les simulations qui respectent le seuil pour chacune des variables

(2)

$$\{ (BV_i^P, T_i^P) \mid i = [1, N^P] \}$$

où les  $i$  simulations sélectionnées vérifient chacune:

$$T_i^P > T_i^+ \quad \forall i \in [1, N^P]$$

on note  $N^P$  le nombre de BV + Arrivées ~~de~~ autres sélections ( $N^P$  est un indicateur de qualité de la méthode).

plusieurs cas de figure:

-  $N^P > 0$  et  $\frac{N^P}{N} = 99\% \rightarrow$  Méthode ok.

-  $N^P = 0$  ou  $\frac{N^P}{N} \leq \frac{10}{10^4} \rightarrow$  Méthode trop contraignante.

### ② Méthode des "scout-s" de sélection globale indépendante.

Pour chacune des variables  $X_i$ , on définit les  $p_i$ :  $N$  simulations qui sont les plus proches avec  $p_i$  le pourcentage de sélection de la variable  $i$ .

On regarde ensuite les ~~simulations~~ <sup>cas</sup> communes, celles ~~qui sont~~ qui font partie des cas les plus proches pour chacune des variables.

On en déduit les valeurs des fonctions objectives de ces simulations.

$\langle T_i^P \rangle$  pour chacune des variables.

on peut aussi calculer la moyenne de la "sélectionnalité".

$$\sum_{i=1}^{n_n} \frac{\langle T_i^P \rangle}{\langle T_i^{P'} \rangle}$$

où les  $T_i^{P'}$  sont les ~~cas~~ cas sélectionnés sur une sélection unique.

2 cas de figure:  $N^P > 0$  et  $\frac{N^P}{N} = 99\% \rightarrow$  Méthode ok.

$N^P = 0$  ou  $\frac{N^P}{N} \leq \frac{10}{10^4} \rightarrow$  Méthode trop contraignante.

### ③ Méthode des "quêtes complètes"

Pour chaque  $(BV_j, T_j)$  on définit  $T_j = \sum_{i=1}^{n_n} T_{ij}$

Puis on prend les  $N^P$  simulations les cas les plus proches.

On en a donc les mêmes indicateurs que précédemment.

3

Inconvénient: les ~~se~~ en algorithmes peuvent être sous-optimal pour l'un ou l'autre des variables.

Avantage: toujours le même nombre d'algorithmes -  
quelque fois  $\alpha$  s'écrit avec le signe de  $\frac{J_i}{|J_i|}$

Cela permet également d'évaluer directement

- la cohérence - des informations entre elles -
- l'irréductibilité -