



## Exemples d'application de l'I.A. pour les sites et sols pollués

Muzammel MUHAMMAD<sup>a,b</sup>, Tanguy WESTELYNCK<sup>a</sup>, Alice OTHMANI<sup>b</sup>, Aurélien TRIGER<sup>a</sup>

Mardi 21 Juin 2022



## Laboratoire d'essais dédié aux SSP

*Essais de faisabilité / traitabilité*



## Transition numérique



# Transition numérique S.S.P

ACTIVATEUR



FRANCE NUM

Transition numérique = Gain d'efficacité en utilisant la puissance du numérique

Diagnostic de site = Beaucoup d'informations non bancarisées :

- ❌ Non évolutives (compilation de rapports)
- ❌ Difficilement accessibles (par des non experts)
- ❌ Non sécurisées (perte de données dans le temps)
- ❌ Non exploitables-valorisables



# Transition numérique S.S.P: Notre démarche

Rapport PDF



BDD exploitable

- ✓ Architecture adaptée
- ✓ Outils d'exploitation
- ✓ Interface utilisateur

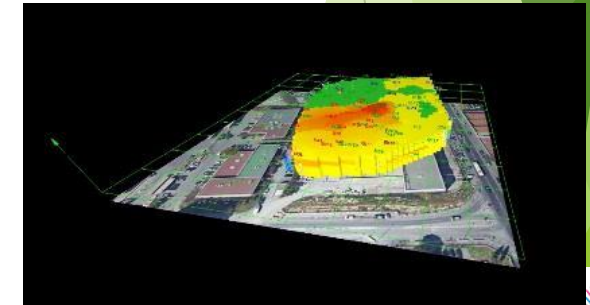
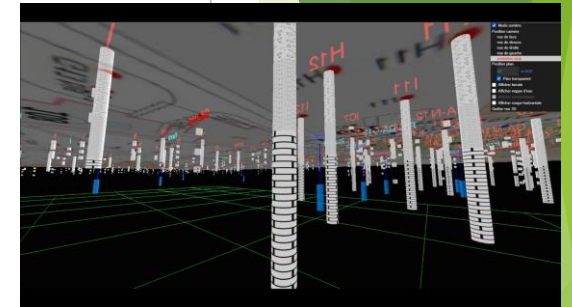
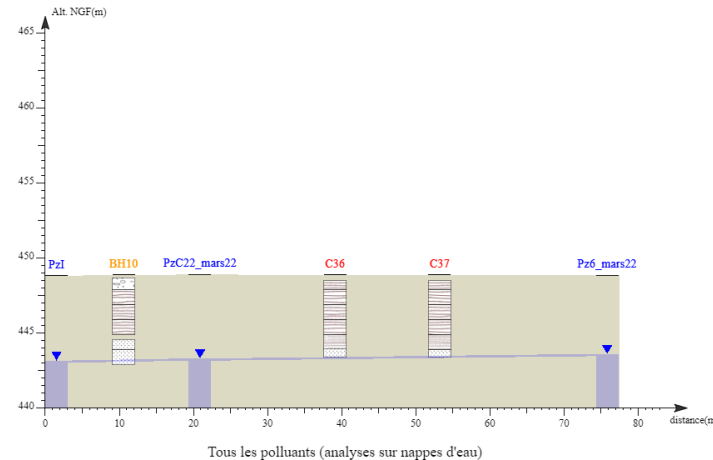


# Transition numérique S.S.P: Notre démarche



## A l'échelle d'un site

- ✓ Visualisation d'informations 3D
- ✓ Interpolation de valeurs:
  - Calculs d'incertitudes
  - Bilan de masses
  - Plan de traitement
  - Plan d'excavation
- ✓ Analyse statistique



# Transition numérique S.S.P: Notre démarche



A l'échelle de plusieurs sites: Base de données

Quelles valorisations possibles ?

**Retour d'expériences « manuels »:** consultation de la base de données (statistiques)

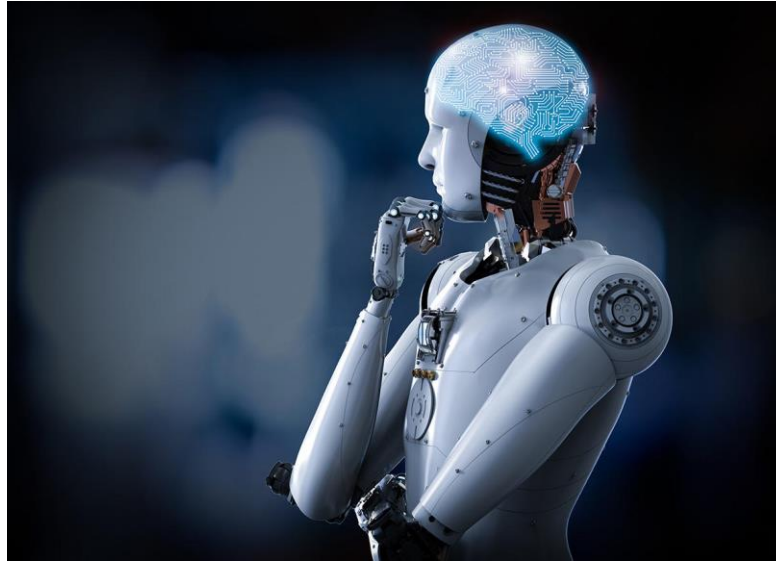
**Retour d'expériences « automatiques »:** Apprentissage Automatique (Machine Learning)



# Valorisation de données environnementales

Valorisation par apprentissage automatique (Machine Learning):

Base de données SSP



Qu'est ce qu'un algorithme d'I.A. peut apprendre à partir d'une BDD SSP ?

# Etapes du projet

- ✓ Développement d'un logiciel de bancarisation des données: KiWi-Maps
- ✓ Création d'une base de données
- ✓ Développement d'algorithmes spécifiques : Brevet GeoPolDeepNet
- **Compréhension et validation de l'apprentissage automatique**
- Application à grande échelle: développement commercial



# Démarche utilisée

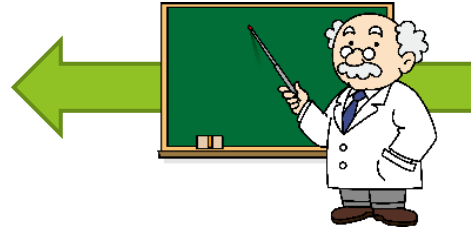
## Apprentissage automatique

### Données d'entrainement

- 54 sites
- > 1 500 sondages
- > 400 000 analyses



BDD SSP



Entrainement de l'algorithme  
GeoPolDeepNet

Algorithme entrainé

« Qu'a-t-il appris ?  
Que peut-on en tirer pour la  
gestion de sites pollués ? »

# Comment évaluer l'apprentissage de l'algorithme?



Tests de prédiction de concentrations en polluants sur plusieurs sites

Utilisation des résultats d'analyses existants pour évaluer l'apprentissage de nos algorithmes



# Evaluation de l'apprentissage de l'algorithme



Exemple:

Site:

80 sondages

4 000 analyses (C10-C40; BTEX; HAP...)

*Recommence 80X*

- 1) Isole un sondage: Par exemple T1
- 2) Demande à l'algo de prédire une concentration: [C10-C40] (0-1 m)
- 3) Compare le résultat avec la réalité (mesure en T1)

# Evaluation de l'apprentissage automatique

Statistique de performance (sur l'ensemble des données disponibles: **Trois indicateurs**)

**Précision ( $>L.D.$ ) : %**

**Précision ( $<L.D. (=0)$ ): %**


**Moyenne quadratique des écarts (RMS) : mg/kg MS**

**Méthodes pour évaluer:**

- L'efficacité de l'apprentissage de nos algorithmes
- Performance d'une méthode déterministe (Polygone de Thiessen)



- 128 sondages
- 10 000 analyses



```
graph LR; A[Isole un sondage  
Prédit sa valeur  
Compare prédiction vs réalité] --> B[128 X];
```

Isole un sondage  
Prédit sa valeur  
Compare prédiction vs réalité

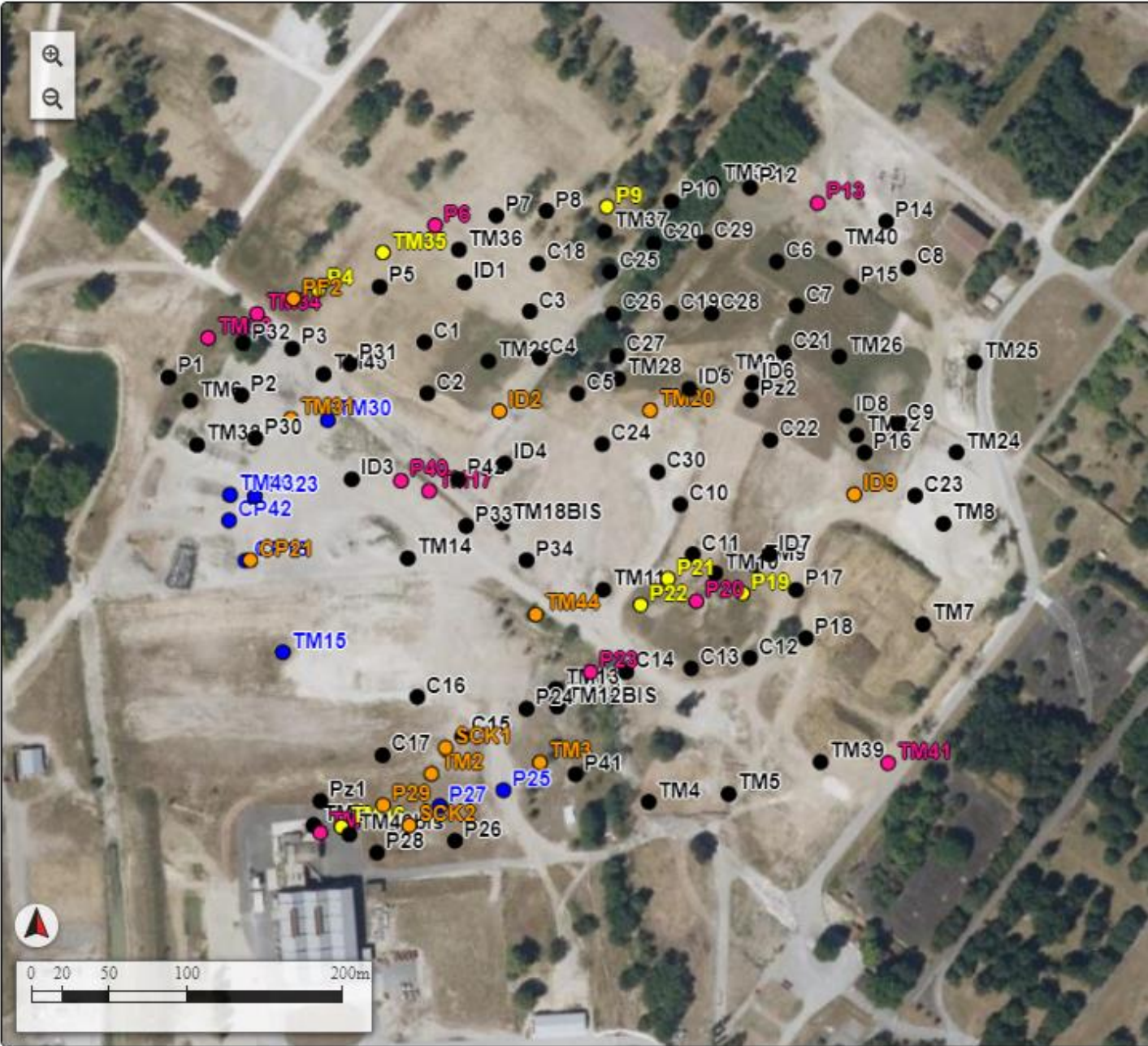
128 X

Précision (> L.D.)	53%
Précision (< L.D. (=0))	73%
RMS	74 mg/kg MS

128 X



# Résultats: Site 1



Site 1 :

- 128 sondages
- 10 000 analyses

Performance de l'Algorithme entraîné

Isole un sondage (X,Y,Z)  
Prédit la valeur [C10-C40]  
Compare prédiction vs réalité

128 X

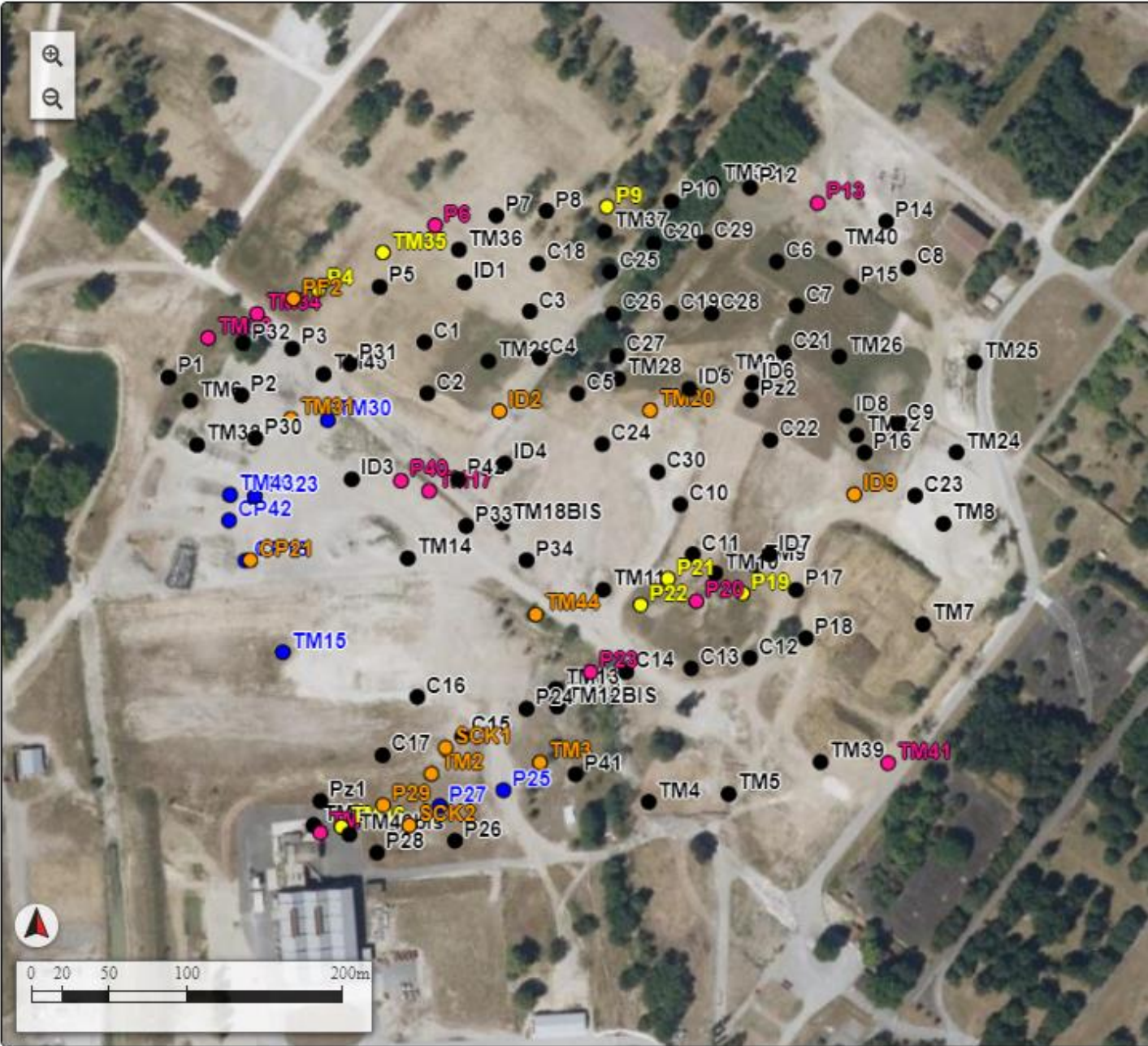
A-t-il appris à faire des interpolations entre valeurs ?

**NON**

	Algo entraîné	Polyg. Thiessen
Précision (> L.D.)	39%	53%
Précision (< L.D. )	37%	73%
RMS (mg/kg MS)	82	74



# Résultats: Site 1



Site 1 :

- 128 sondages
- 10 000 analyses

Performance de l'Algorithme entraîné

Isole un sondage [HAP]  
Prédit sa valeur [C10-C40]  
Compare prédiction vs réalité

128 X

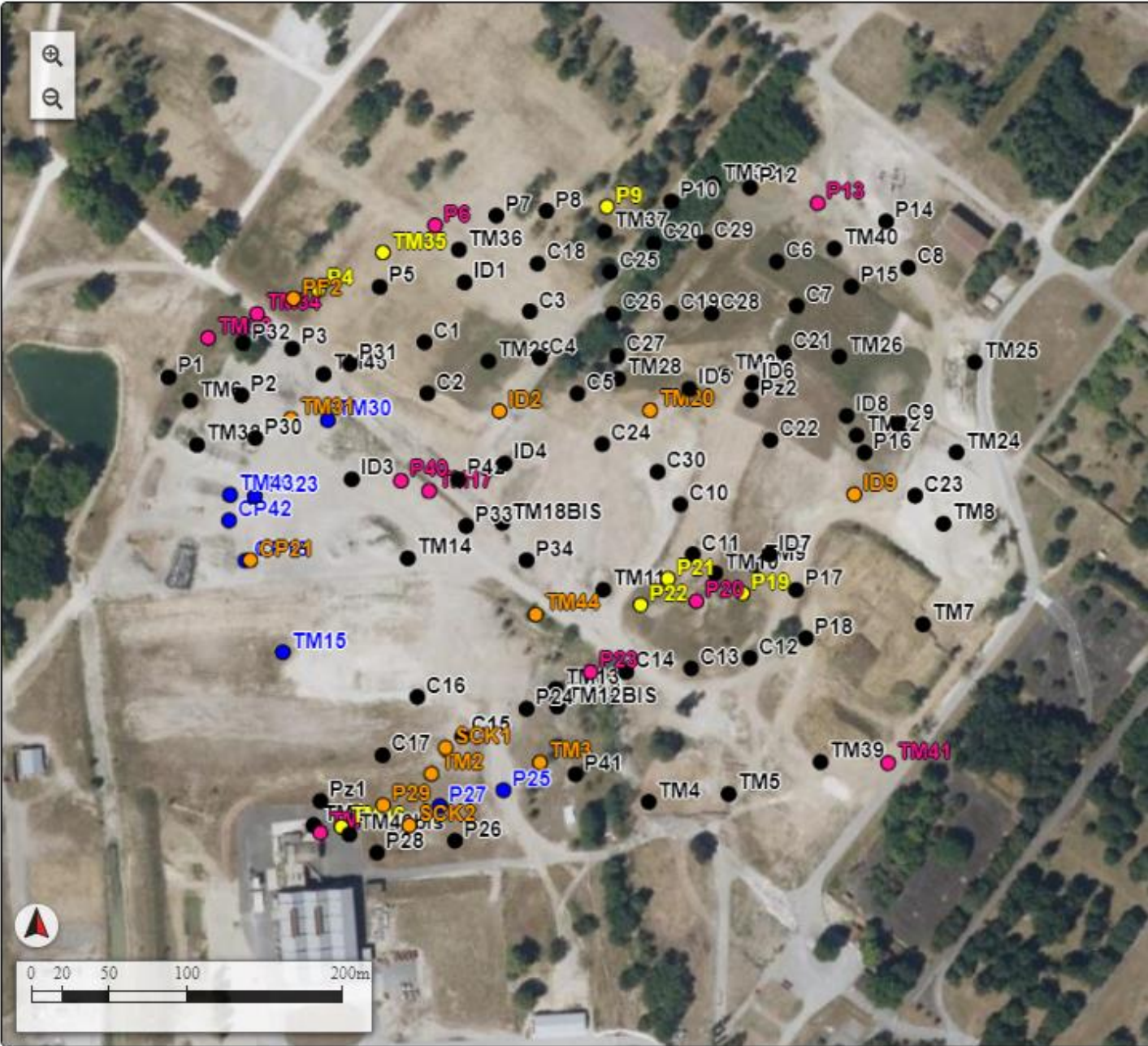
A-t-il appris à faire un lien entre [HAP] et [C10-C40] ?

OUI !

	Algo entraîné	Polyg. Thiessen
Précision (> L.D.)	68%	53%
Précision (< L.D. )	81%	73%
RMS (mg/kg MS)	49	74



# Résultats: Site 1



Site 1 :

- 128 sondages
- 10 000 analyses

Performance de l'Algorithme entraîné

Isole un sondage **Géologie**  
Prédit sa valeur [C10-C40]  
Compare prédiction vs réalité

128 X

A-t-il appris à faire un lien entre  
Géologie et [C10-C40] ?



	Algo entraîné	Polyg. Thiessen
Précision (> L.D.)	56%	53%
Précision (< L.D. )	77%	73%
RMS (mg/kg MS)	47	74



# Résultats: Site 2



## Site 2

21 sondages  
70 analyses de sulfates sur brut et sur éluats

## Performance de l'Algorithme entraîné

Isole un sondage **[Sulfate]<sub>brut</sub>**  
Prédit sa valeur **[Sulfate]<sub>eluat</sub>**  
Compare prédiction vs réalité

21 X

A-t-il appris à faire le lien entre  
**[Sulfate]<sub>brut</sub>** et **[Sulfate]<sub>eluat</sub>** ?

**OUI !**

	Algo entraîné
Précision (> L.D.)	100%
Précision (< L.D. )	100%
RMS (mg/kg MS)	18

# Conclusion

- ▶ Intérêt technique et scientifique ! → Passionnant
- ▶ Intérêt économique: Diminuer:
  - ▶ Le nombre d'analyses: [Polluant 1]  $\leftrightarrow$  [Polluant 2]
  - ▶ Le nombre de sondages: Géophysique (profils géologiques)
  - ▶ Les délais d'analyses: [Sulfates] brut  $\rightarrow$  [Sulfates] eluats
- ▶ Démocratiser le développement de laboratoires d'analyse mobiles

# Etapes du projet

- ✓ Développement d'un logiciel de bancarisation des données: KiWi-Maps
- ✓ Création d'une base de données
- ✓ Développement d'algorithmes spécifiques : Brevet GeoPolDeepNet
- ▶ Compréhension et validation de l'apprentissage automatique
- ▶ Application à grande échelle: développement commercial

# D.I.A: Diagnostic assisté par l'I.A.

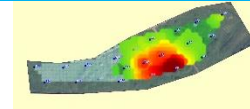
## 1) Numérisation des données historiques existantes

Mémoire numérique des sites



## 2) Traitement des données

Estimation des incertitudes



## 3) Diagnostic de sites

Stratégie d'investigation: Algorithmes brevetés



Utilisation de données « secondaires »

- Mesures PID de terrain
- Relevés Géophysique
- Analyses de surfaces (REMSCAN, Niton..)
- ...



# Recherche de partenariats

- Sites pilotes
- Données de diagnostics
- Mise en œuvre sur site
- ...

**intersol'2022**

Congrès-Exposition International sur les Sols, les Sédiments et l'Eau  
International Conference-Exhibition on Soils, Sediments and Water

21, 22 & 23 juin 2022 - Lyon, France

[www.kiwi-maps.com](http://www.kiwi-maps.com)

[www.evaldepol.com](http://www.evaldepol.com)

06-14-28-63-68

[contact@evaldepol.com](mailto:contact@evaldepol.com)