

Possibilités offertes par l'intelligence artificielle pour le contrôle qualité des données Argo

Atelier IA ODATIS, Marseille, 27-28/01/2026

Thierry Carval – IRSI/ISI
Pierre-Yves Colson - IRSI/ISI
Vanessa TOSELLO - IRSI/ISI

QC Argo – contexte et enjeux

Processus actuel

- Profils T/S soumis à une série de tests automatiques : date, position, pics, inversion densité, pression croissante, valeurs constantes, et climatologie min-max
- Alertes générées ensuite analysées par inspection visuelle par un expert

Enjeux

- Volume important d'alertes à examiner
- Expertise humaine indispensable pour les cas complexes

Apport potentiel de l'IA

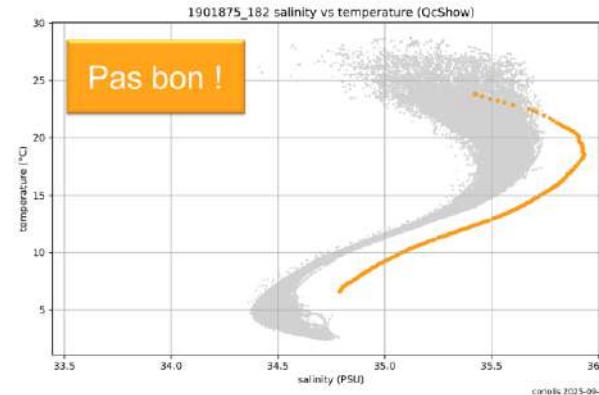
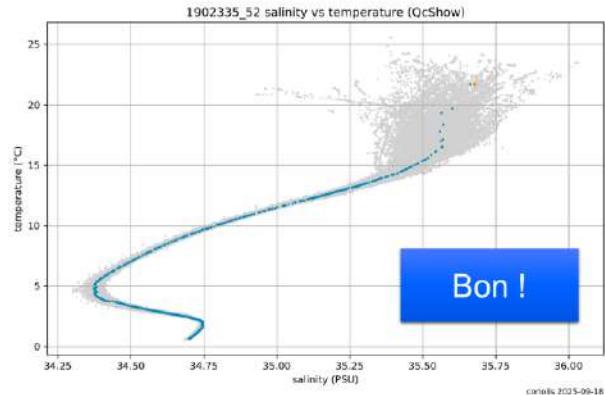
- Aider à l'analyse des alertes QC
- Réduire les fausses alertes tout en conservant un bon niveau de détection
- Mieux cibler l'intervention de l'expertise humaine
- Apprendre à reproduire certains comportements experts récurrents sur des alertes QC

→ Trois approches parmi d'autres sont explorées

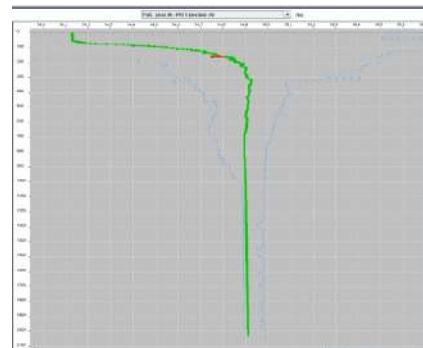
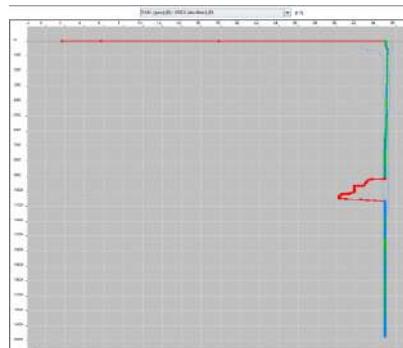
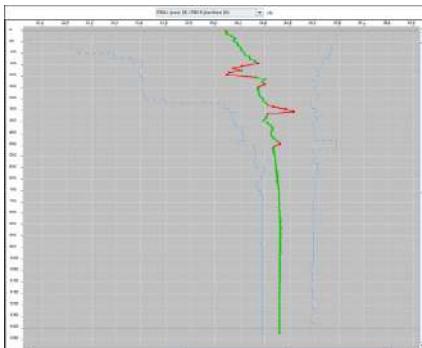


Deux échelles d'analyse QC

Analyse par profil: évaluation de la cohérence globale du profil



Analyse par point : décisions locales



Méthode 1 : Interpolateur T/S

Objectif: comparer les valeurs T/S d'un profil à des valeurs de référence prédites par un interpolateur

Principe

- Prédiction, **point par point**, d'une valeur attendue de température et de salinité à partir d'une position (longitude, latitude, profondeur)
- Calcul de l'écart entre la valeur prédite et la valeur réelle
- Identification des points présentant un écart significatif

Décision QC

- Les écarts sont calculés localement, point par point
- Une alerte peut être levée lorsque plusieurs points d'un même profil présentent des écarts significatifs
- Les critères de décision sont définis explicitement, en dehors du modèle

Evaluation

- Recherche de corrélations entre le QC existant et l'erreur de reconstruction du profil de référence



Méthode 1 : Interpolateur T/S

Données et entraînement du modèle

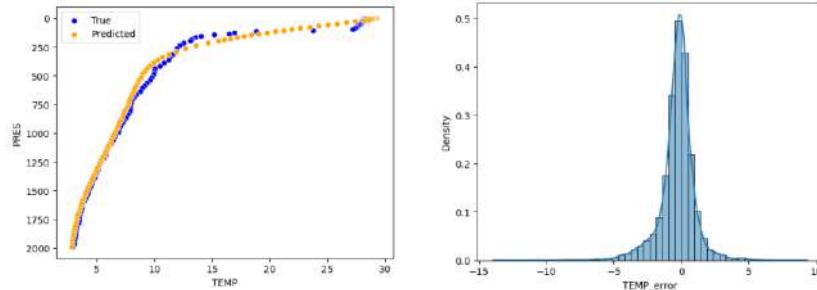
Jeu de données

- Apprentissage et validation: Profils Argo 2014–2023
- Test: Profils Argo 2024 (temporellement indépendant)
- Données considérées comme valides (QC = 1)

Apprentissage

- Architecture MLP
- Entraînement réalisé à ce stade sur ~1500 profils et 1 M de points
- Total train: ~1,5 M de profils, 760 M de points

Exemple de profil reconstruit et distribution des erreurs sur le jeu de test



Méthode 2 : Vision par ordinateur

Approche profil par profil : représentation visuelle des profils

Principe

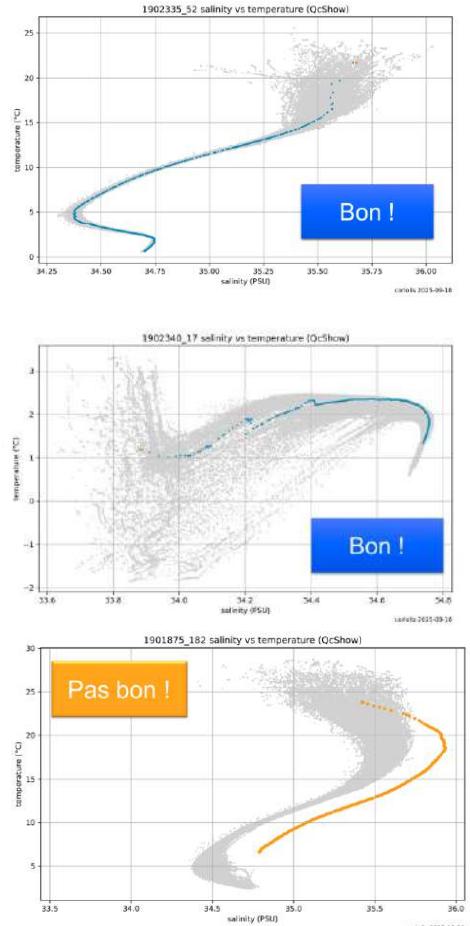
- Conversion des profils T/S en représentations visuelles (images)
- Le profil d'intérêt est tracé avec des profils proches géographiquement afin de fournir un contexte local
- Analyse des images par un réseau de type **YOLO** (*You Only Look Once*)
- Apprentissage supervisé à partir de profils annotés par des experts
- Repérage de profils visuellement atypiques

Évaluation

- Comparaison entre les sorties du modèle et les annotations expertes

Philosophie de l'approche

Le modèle n'intègre aucune connaissance océanographique explicite ; il apprend des régularités visuelles présentes dans des profils déjà expertisés.



Méthode 2 : Vision par ordinateur

Données et entraînement du modèle YOLO

Jeu de données

- Profils T/S Argo – année 2021
- **4 000 profils** représentés sous forme d'images
- Annotations issues du QC expert existant

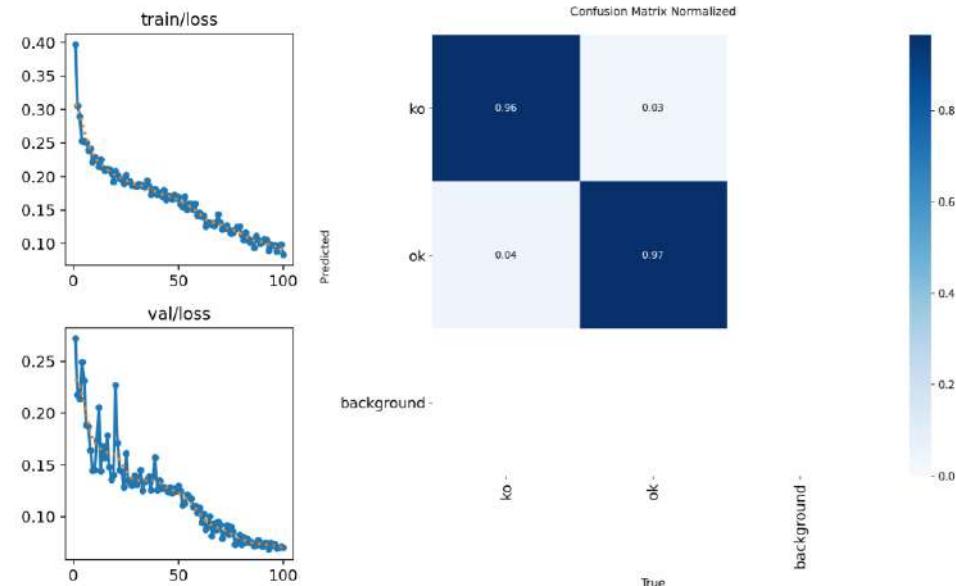
Apprentissage

- Entraînement : **80 %** des profils
- Validation : **20 %** des profils
- Environnement GPU (Datarmor)

Matrice de confusion

Matrice de confusion illustrant la cohérence entre sorties du modèle et annotations expertes, sur le jeu utilisé:

- 97% de détection « OK », 3% de fausses alertes
- 96% de détection « KO », 4% d'erreurs non détectées



Méthode 2 : Vision par ordinateur

Application du modèle sur le jeu de données EasyOneArgoTS

Jeu utilisé

- [EasyOneArgoTS](#) (années 2021–2022)
- Jeu de profils Argo considérés comme valides

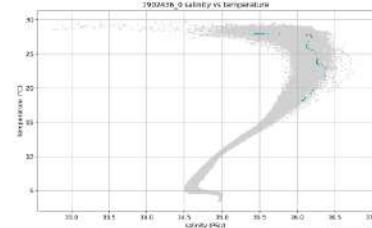
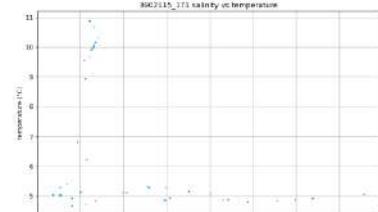
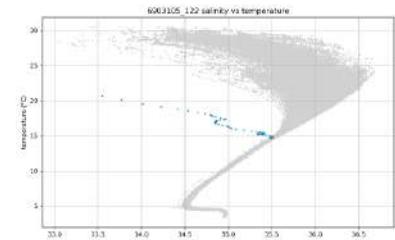
Principe

- Application du modèle à grande échelle (~283 000 profils)
- Attribution d'un score par profil
- Sélection des profils les plus atypiques selon différents seuils

Résultats observés

- Selon le seuil retenu, seuls quelques centaines de profils sont signalés
- Parmi ces profils, certains présentent des caractéristiques visuelles atypiques, bien que le jeu soit considéré comme valide

| Inférences | argyolo 202512 | % |
|------------------------|--------------------------|-------|
| données | années Argo 2021 et 2022 | |
| nb images | 283 498 | |
| nb images en alerte | 10 934 | 4% |
| alertes percentile 100 | 323 | 0,11% |
| alertes percentile 90 | 2671 | 0,94% |
| alertes percentile 80 | 1941 | 0,68% |
| alertes percentile 50 | 5999 | 2,12% |
| alertes percentile <50 | 0 | 0,00% |



Méthode 3 : Post-traitement des alertes QC

Objectif

- Réduire le nombre de fausses alertes générées par les tests de QC automatique
- Concentrer l'expertise humaine sur les alertes pertinentes

Référence existante

- Mieruch S et al. (2025) *SalaciaML-2-Arctic* — a deep learning quality control algorithm for Arctic Ocean temperature and salinity data. doi: 10.3389/fmars.2025.1661208

Principe (approche hybride QC traditionnel + ML)

- Modèle ML appliqué uniquement aux alertes générées par le QC automatique, **point par point**
- Apprentissage supervisé à partir de décisions expertes existantes
- Requalification de certaines alertes jugées non pertinentes

Évaluation

- Comparaison des sorties du modèle ML avec les décisions expertes
- Analyse de la réduction du nombre d'alertes transmises à l'expert
- Vérification du maintien du taux de détection des alertes pertinentes

→ Résultats mitigés lors de tests préliminaires, nécessité d'adapter le principe au cadre spécifique du QC Argo

Conclusion

Apports

- Approches IA complémentaires au QC Argo
 - Vision par ordinateur : exploration visuelle et repérage de profils atypiques à l'échelle du profil
 - Interpolateur : estimation d'une référence locale et comparaison de profils à cette référence
 - Post-traitement QC : pour filtrer certaines alertes issues du QC classique
- Objectif : mieux cibler les cas à investiguer et de rationaliser l'effort d'expertise
- Le QC et l'expertise humaine restent centraux
- Les modèles n'apportent pas de décision QC

Vigilances

- Perte d'explicabilité
- Risque de lever de nouvelles fausses alertes ou de supprimer/corriger de vraies alertes
- Déséquilibre des classes
- Évaluation et comparaison des méthodes (IA vs QC classique)

