
Contrôle de qualité des données STD (IML)

INFORMATIONS GÉNÉRALES

Le contrôle de qualité des données STD est essentiellement basé sur les algorithmes du manuel de l'UNESCO :

UNESCO (1990) GTSP Real-Time Quality Control Manual, Intergovernmental Oceanographic Commission, Manuals and Guides 22, SC/90/WS-74 : 121 p.

Les limites de certains tests GTSP ont été adaptées aux conditions de l'estuaire et du golfe du St-Laurent. Quelques tests supplémentaires ont également été ajoutés de manière à obtenir un meilleur contrôle du profil STD dans son ensemble.

Actuellement, le contrôle de qualité permet de déterminer la validité des données de pression (ou profondeur), température, salinité et sigma-T ainsi que des métadonnées principales de position temps-espace d'un profil STD. Éventuellement, un contrôle de l'oxygène dissous, de la fluorescence et de la transmission de la lumière pourrait y être ajouté.

Le contrôle de qualité STD se divise en 5 étapes :

Étape 1 : Contrôle des métadonnées importantes telles que le temps et la position.

Étape 2 : Contrôle des enregistrements d'un profil les uns par rapport aux autres.

Étape 3 : Comparaison des profils avec une climatologie.

Étape 4 : Comparaison des profils d'une même mission entre eux.

Étape 5 : Visualisation de la trajectoire du navire et visualisation des profils.

Toute métadonnée peut être modifiée, en particulier la coordonnée temps-espace sans quoi un profil est inutilisable.

Aucune donnée n'est modifiée par le contrôle de qualité lui-même. Un sémaphore de qualité est ajouté (voir section suivante) pour qualifier les données de bonnes, douteuses, erronées ou manquantes seulement pour les tests de l'étape 2. Si les données doivent être modifiées pour quelques raisons que ce soit, ces modifications sont apportées en dehors du contrôle de qualité et le sémaphore de qualité doit être ajusté en conséquence.

DESCRIPTION DES SÉMAPHORES DE QUALITÉ INDIVIDUELS

Les tests de l'étape 2 ajoutent des sémaphores de qualité aux données de pression, profondeur, température, salinité et sigma-T. Le sémaphore de qualité est un entier compris entre 0 et 9. La signification des sémaphores de qualité est présentée dans le tableau suivant:

Sémaphore	Signification
0	aucun contrôle de qualité
1	la donnée semble correcte
2	la donnée semble incohérente par rapport aux autres données
3	la donnée semble douteuse
4	la donnée semble erronée
5	la donnée a été modifiée
6 à 8	réservé à un usage futur
9	la donnée est manquante

La priorité des sémaphores est croissante, le sémaphore 0 étant le moins prioritaire et le sémaphore 9 étant le plus prioritaire. Par conséquent, si un test a jugé une donnée douteuse (sémaphore 3) et que le test suivant juge la même donnée erronée (sémaphore 4), c'est le sémaphore 4 qui sera conservé.

Le tableau suivant montre comment sont distribués les sémaphores de qualité en fonction de la dépendance des canaux de données. Les canaux dépendants sont les canaux calculés à partir des données mesurées par le CTD. Le CTD mesure la température (TE90), la conductivité et la pression (PRES). Les données calculées sont : la profondeur (DEPH), la salinité (PSAL) et le sigma-T (SIGT). La salinité est calculée à partir de la conductivité, de la température et de la pression. La densité est calculée à partir de la salinité, de la température et de la pression. La profondeur est calculée à partir de la pression et de la latitude.

PRES	DEPH	TE90	PSAL	SIGT
[4]	4	-	4	4
[3]	3	-	3	3
-	-	[4]	4	4
-	-	[3]	3	3
-	-	-	[4]	4
-	-	-	[3]	3
-	-	-	-	[3]

[]: Résultat d'un test

Par exemple, si la donnée de température d'un enregistrement de données a été jugée erronée (sémaphore 4) par un test de contrôle de qualité, le sémaphore 4 sera ajouté à droite de la donnée de température en erreur. Un sémaphore 4 sera également ajouté à droite des données de salinité et de sigma-T du même enregistrement.

SÉMAPHORE GLOBAL QCFF

Le sémaphore QCFF est un sémaphore qui permet de retracer la provenance du sémaphore de qualité. Il ne s'applique qu'aux tests de l'étape 2 du contrôle de qualité. À chaque test de cette étape est associé un nombre 2^x où x est un entier positif. Avant d'exécuter le contrôle de qualité, on attribue une valeur de QCFF de 0 à un enregistrement. Lorsqu'un test échoue, la valeur 2^x qui lui est associée est ajoutée au QCFF. On peut ainsi facilement retracer les tests échoués d'un enregistrement en décomposant le QCFF sous forme binaire. Si un sémaphore de qualité d'un enregistrement est modifié à la main, une valeur de 1 est ajoutée au QCFF.

Voici la liste des tests disponibles actuellement (cette liste anglaise est ajoutée à l'en-tête des fichiers de données lorsque le contrôle de qualité a été effectué ; la description de chacun des tests, dans la prochaine section, est en français) :

TEST	DESCRIPTION (QCFF)
Test 1.1	GTSP Platform Identification
Test 1.2	GTSP Impossible Date/Time
Test 1.3	GTSP Impossible Location
Test 1.4	GTSP Position on Land
Test 1.5	GTSP Impossible Speed
Test 1.6	GTSP Impossible Sounding
Test 2.0	IML Minimum Descent Rate (2)
Test 2.1	GTSP Global Impossible Parameter Values (4)
Test 2.2	GTSP Regional Impossible Parameter Values (8)
Test 2.3	GTSP Increasing Depth (16)

Test 2.4	GTSP Profile Envelope (Temperature and Salinity) (32)
Test 2.6	GTSP Freezing Point (128)',
Test 2.7	GTSP Spike in Temperature and Salinity (one point) (256)
Test 2.8	GTSP Top and Bottom Spike in Temperature and Salinity (512)
Test 2.9	GTSP Gradient in Temperature and Salinity (1024)
Test 2.10	GTSP Density Inversion (point to point) (2048)
Test 2.11	IML Spike in Pressure, Temperature and Salinity (one point or more) (4096)
Test 2.12	IML Density Inversion (overall profile) (8192)
Test 3.5	IML Petrie Monthly Climatology (Temperature, Salinity and Sigma-T)
Test 4.1	GTSP Profile Consistency
Test 4.2	IML Annual Deep Water Profile Consistency
Test 5.1	GTSP Cruise Track Visual Inspection
Test 5.2	GTSP Profile Visual Inspection

Si la salinité d'un enregistrement échoue le test 2.7, alors le QCFF sera de 256. Si en plus, le test 2.0 avait déjà été échoué, alors le QCFF sera de $2+256=258$.

DESCRIPTION DES TESTS DU CONTRÔLE DE QUALITÉ

Test 1.1 : Identification du navire

Ce test vérifie seulement si tous les profils de la mission ont été échantillonnés sur le même navire.

Test 1.2 : Date et heure

Ce test vérifie si la date et l'heure du début et de la fin du profil sont comprises à l'intérieur du temps alloué pour la mission.

Test 1.3 : Latitude et longitude

Ce test vérifie si la position du profil est possible c'est-à-dire si la position est comprise entre -90 et 90 de latitude ou -180 et 180 de longitude.

Test 1.4 : Station à terre

Ce test vérifie si la position du profil se trouve à l'intérieur des polygones de terre de l'estuaire et du golfe du St-Laurent. L'aire couverte par ce test s'étend de -70 à -56 en longitude et de 45 à 52 en latitude.

Test 1.5 : Vitesse du navire

Ce test vérifie la vitesse du navire entre deux profils consécutifs. La vitesse de navire est calculée à partir de la position temps-espace de la mise à l'eau et la position temps-espace de la sortie de l'eau du profil précédent. Dans le cas où la position ou la date et heure de sortie de l'eau du profil précédent sont manquantes, le test utilise les coordonnées de la mise à l'eau du profil précédent pour déterminer la vitesse du navire. Le test compare la vitesse calculée à la vitesse de croisière du navire utilisé.

Test 1.6 : Profondeur sondée

Le test 1.6 reporte la profondeur sondée sur une grille bathymétrique aux 3 km de l'estuaire et du golfe du St-Laurent pour déterminer sa validité. Une profondeur est considérée valide si elle se situe à ± 20 m de la

profondeur correspondante de la grille bathymétrique. L'aire couverte par ce test s'étend de -70 à -56 en longitude et de 45 à 52 en latitude.

Test 2.0 : Vitesse de descente de la sonde

Le test 2.0 est exécuté seulement si la variable DPDT (vitesse de descente) a été sauvegardée ou s'il est possible de la calculer. Lorsque la vitesse de descente est inférieure à la vitesse minimale indiquée dans l'historique (habituellement 0.1 m/s), un sémaphore 3 est attribué à l'enregistrement de salinité.

Test 2.1 : Valeurs impossibles globalement

Le test 2.1 vérifie si les données de pression, température et salinité sont possibles globalement suivant les critères du tableau ci-dessous. Si une donnée est jugée impossible donc erronée, son sémaphore de qualité est alors remplacé par 4.

Code	Variable	Unité	Valeur minimale	Valeur maximale
BATH	Profondeur sondée	M	0	10000
TEMP	Température	°C	-2.5	35
PSAL	Salinité	(psu)	0	50
PRES	Pression	db	0	10000

Test 2.2 : Valeurs impossibles régionalement

Le test 2.2 vérifie si les données de pression, température et salinité sont possibles régionalement suivant les critères du tableau ci-dessous. Si une donnée est jugée impossible donc erronée, son sémaphore de qualité est alors remplacé par 4.

Code	Variable	Unité	Valeur minimale	Valeur maximale
BATH	Profondeur sondée	m	0	600
TEMP	Température	°C	-2.5	35
PSAL	Salinité	(psu)	0	35
PRES	Pression	db	0	600

La région est délimitée par les coordonnées (longitude, latitude) :
 (-56.0, 52.0), (-73.0, 49.5), (-73.0, 46.0), (-64.5, 46.0), (-62.3, 45.2), (-56.0, 48.2), (-56.0, 52.0)

Test 2.3 : Profondeur décroissante

La pression testée est douteuse lorsqu'elle est inférieure à la pression de l'enregistrement précédent et son sémaphore de qualité est remplacé par 3. Les données de pression erronées ou manquantes ne sont pas considérées.

Test 2.4 : Enveloppe du profil

Le test 2.4 vérifie si les données de température et salinité se situent à l'intérieur des limites par intervalle de profondeurs présentées au tableau ci-dessous. La donnée est jugée douteuse si sa valeur ne se situe pas à l'intérieur de l'intervalle permis et son sémaphore de qualité est alors remplacé par 3. Les données erronées ou manquantes de température ou salinité ne sont pas considérées.

Code	Intervalle de profondeurs (m)	Unité	Valeur minimale	Valeur maximale
TEMP	0 – 50	°C	-2.5	35
TEMP	50 – 100	°C	-2.5	30
TEMP	100 – 400	°C	-2.5	28

TEMP	400 – 1100	°C	-2	28
PSAL	0 – 50	(psu)	0	35
PSAL	50 – 100	(psu)	1	35
PSAL	100 – 400	(psu)	3	35
PSAL	400 – 1100	(psu)	10	35

Test 2.6 : Point de congélation

La température de congélation est calculée à partir de la salinité et de la pression. Une valeur de température plus basse que la température de congélation correspondante est jugée erronée et son sémaphore est remplacé par 4. Les données de température préalablement jugées erronées ou manquantes ne sont pas considérées.

Test 2.7 : Pics ponctuels anormaux en température et salinité (premier et dernier enregistrement exclus)

Un pic est détecté en comparant la valeur d'une donnée (V2) avec la valeur précédente (V1) et suivante (V3). Si $(|(V2 - (V3+V1))/2| - |(V1-V3)/2|) > \text{Valeur limite}$ alors V2 échoue le test. Les valeurs limites de température et salinité se trouvent dans le tableau ci-dessous. Les données qui échouent le test sont douteuses et par conséquent leurs sémaphores de qualité sont remplacés par 3. Les données de température ou salinité préalablement jugées erronées ou manquantes ne sont pas considérées.

Code	Unité	Valeur limite
PRES	m	5
TEMP	°C	2
PSAL	(psu)	0.3

Test 2.8 : Pics anormaux en température et salinité au fond et à la surface

Un pic à la surface est détecté en comparant la valeur de la donnée testée (V1) à celle de l'enregistrement suivant (V2) de telle sorte que si $| (V1 - V2) | < \text{Valeur limite}$ alors aucun pic n'est détecté.

Un pic au fond est détecté en comparant la valeur de la donnée testée (V2) à celle de l'enregistrement précédent (V1) de telle sorte que si $| (V2 - V1) | < \text{Valeur limite}$ alors aucun pic n'est détecté.

Si un pic est décelé, le sémaphore 3 y est attribué. Les données de température ou salinité préalablement jugées erronées ou manquantes ne sont pas considérées.

Code	Unité	Valeur limite
PRES	m	25
TEMP	°C	10
PSAL	(psu)	5

Test 2.9 : Gradients de température et salinité

Les gradients verticaux de température et de salinité sont calculés afin de déterminer s'ils excèdent les valeurs permises du tableau ci-dessous. Les données de température ou salinité préalablement jugées douteuses, erronées ou manquantes ne sont pas considérées.

Si $| V2 - (V1 + V3) / 2 | / (\text{différence de pression}) > \text{Gradient limite}$ alors V2 échoue le test. V1, V2 et V3 sont 3 valeurs successives de température ou de salinité. Un sémaphore 3 est donc associé à V2.

Code	Unité	Gradient limite
TEMP	°C /m	10
PSAL	(psu)/m	5

Test 2.10 : Inversions de densité (1 point)

Ce test compare la densité d'un enregistrement (V2) à celle de l'enregistrement précédent (V1). Si $V2-V1 < -0.05$, le test est échoué et la densité est considérée douteuse (sémaphore de qualité de 3). Les données de densité préalablement jugées douteuses, erronées ou manquantes ne sont pas considérées.

Test 2.11 : Pics anormaux (plusieurs points)

Le test 2.11 permet de détecter les pics anormaux de pression, température et salinité à enregistrements multiples en considérant la différence entre les données du profil et le résultat de l'application d'un filtre médian à ces mêmes données. La fenêtre du filtre médian est déterminée à travers un algorithme qui augmente la fenêtre du filtre jusqu'à l'obtention du profil de densité dont les inversions sont inférieures à 0.01 kg/m^3 . La fenêtre ainsi obtenue est alors utilisée pour filtrer les données de pression, température et salinité. Lorsque la différence entre la mesure de la variable filtrée et non filtrée excède la valeur indiquée dans le tableau ci-dessous, la donnée est jugée douteuse (sémaphore 3). Les données de pression, température ou salinité préalablement jugées douteuses, erronées ou manquantes ne sont pas considérées.

Code	Unité	Différence permise ¹
PRES	m	0.2
TEMP	°C	0.2
PSAL	(psu)	0.015

¹ Les différences permises sont parfois modifiées en fonction de l'intervalle d'échantillonnage

Test 2.12 : Inversions de densité (plusieurs points)

Le test 2.12 détecte les diminutions de densité de 0.05 kg/m^3 par décibar. Si le test détecte une inversion, la densité en défaut est remplacée par la précédente et le test est refait jusqu'à ce qu'il n'y ait plus d'inversion. Les données de pression et sigma-T préalablement jugées douteuses, erronées ou manquantes ne sont pas considérées. Il est préférable d'effectuer le test 2.11 au préalable puisqu'un pic positif de densité associé à une augmentation douteuse de la salinité, par exemple, ne sera alors pas considéré dans le traitement. Le sémaphore 3 est associé à une inversion de densité.

Test 3.5 : Climatologie mensuelle de Petrie

Cette climatologie a été compilée par Petrie et al. (1996) pour le golfe du St-Laurent. Les températures, salinités et densités moyennes à des profondeurs fixes de 21 régions du golfe y sont calculées pour chaque mois ainsi que leurs écarts-types respectifs. Ce test utilise cette climatologie pour déterminer la validité des observations d'une mission. Si l'écart entre les observations et la climatologie excède 3 écarts-types, un avertissement est émis. La responsabilité revient donc à l'utilisateur de déterminer s'il doit rejeter un lot d'observations ou seulement ajouter des sémaphores à quelques observations en particulier. Le problème avec ce test est qu'une donnée ne peut pas être rejetée simplement parce qu'elle échoue le test. Il se peut que le profil reflète un événement particulier. Il se peut également qu'il y ait un problème de sonde qui rende les observations inutilisables.

Petrie, B., K. Drinkwater, A. Sandström, R. Pettipas, D. Gregory, D. Gilbert and P. Sekhon (1996)
Temperature, salinity and sigma-t atlas for the Gulf of St. Lawrence, Can. Tech. Rep. Hydrogr.
Ocean Sci., 178 : v+256 pp.

Test 4.1 : Comparaison des profils d'une même mission entre eux

Le test 4.1 compare les enregistrements de température et salinité des profils séparés par une certaine distance temporelle et spatiale. Les écarts de température et de salinité de 2 profils situés à l'intérieur de 20 km et dans un intervalle de 2 jours l'un de l'autre devraient être inférieurs aux écarts présentés au tableau ci-

dessous pour des enregistrements à plus de 200 m de profondeur. Seul un avertissement est émis si l'écart entre deux profils est considéré trop élevé. Aucun sémaphore de qualité n'est directement modifié à partir du résultat de ce test.

Code	Unité	Écarts tolérés
TEMP	°C	0.5
PSAL	(psu)	0.3

Test 4.2 : Comparaison de la température et de la salinité des eaux profondes dans un intervalle d'un an

Le test 4.2 est basé sur la conservation à court et moyen terme (maximum 1 an) des propriétés des masses d'eau profondes. Une banque de température et salinité est constituée avec les données de 300 m et plus qui ont été préalablement validées (thermomètre à renversement et bouteille de salinité). Les enregistrements de température et salinité des profils d'une mission sont alors comparés avec les données de la banque s'ils sont situés à l'intérieur d'un rayon de 10 km. Les écarts maximums ne devraient pas dépasser les valeurs du tableau ci-dessous. Un avertissement est émis si l'écart entre 2 profils est considéré trop élevé et la moyenne des écarts est calculée à titre indicatif. Aucun sémaphore de qualité n'est modifié à partir du résultat de ce test.

Code	Unité	Écarts tolérés
TEMP	°C	0.5
PSAL	(psu)	0.3

Test 5.1 : Visualisation de la trajectoire du navire

Ce test permet de visualiser la trajectoire du navire et ainsi d'identifier des positionnements en erreur.

Test 5.2 : Visualisation des profils

Cette étape du contrôle de qualité est primordiale. Elle permet de voir et de comparer le profil STD original et le profil résultant du contrôle de qualité (celui auquel les enregistrements douteux, erronés et manquants ont été soustraits). La validité d'un jeu de données est déterminée à cette étape du contrôle de qualité.
