



Quadriges² - Référentiel National de gestion des données de la surveillance littorale

Florian Enez
Emilie Gauthier
Anne Pellouin-Grouhel

Septembre 2018

■ **Procédure de qualification des données Quadriges²**

Données chimie :

- RNOMV
- RNOSED
- **RNOPHY**

SOMMAIRE

1. Qualification : définition et processus pour RNOPHY	3
1.1. Définition de la qualification	3
1.2. Qualification des données RNOPHY	3
1.3. Processus de qualification automatique	4
2. Critères de qualification RNOPHY	4
2.1. Définition de la thématique	4
2.1.1. Liste de programmes	4
2.1.2. Liste de PSFMUs	4
2.2. Liste des entités à qualifier	6
2.3. Identification des intervenants	6
2.4. Anomalies recherchées	6
2.5. Règles de contrôle Quadrige	8
3. Qualification experte.....	8

1. Qualification : définition et processus pour RNOPHY

1.1. Définition de la qualification

La qualification a pour objectif d'attribuer un niveau de qualité aux données : Bon, Douteux ou Faux. Le processus de qualification consiste à rechercher des anomalies plus ou moins complexes dans les données stockées dans la base de données Quadrige², les anomalies étant définies par les responsables de programme Quadrige ou responsables thématiques. Ces recherches d'anomalies sont faites en lien avec la cellule quadrige².

La qualification des données peut se faire :

- de façon ponctuelle : identification d'un problème de processus analytique pour quelques données
- en routine : un processus de qualification dite « automatique » permet de qualifier les enregistrements individuellement (cohérence des informations de chaque résultat individuellement). Puis une qualification dite « experte » permet de vérifier la cohérence d'un résultat par rapport aux autres résultats (séries temporelles, analyse graphique ou statistique).

Le présent document décrit les processus de qualification **en routine** des données du réseau **ROCCH – volet IMPOSEX** (programme RNOPHY de Quadrige).

1.2. Qualification des données RNOPHY

L'imposex est la réponse biologique à une contamination du milieu par le tributylétain (TBT) utilisé dans les peintures antisalissure des carènes de bateaux. Il provoque une masculinisation des femelles de certains gastéropodes marins (Imposex). L'usage du TBT est désormais interdit et le TBT est une des substances prioritaires de la Directive Cadre sur l'Eau. Dans le cadre de la Convention pour la protection du milieu marin de l'Atlantique du Nord-Est (OSPAR), un suivi de l'imposex est réalisé chaque année depuis 2003.

Les données issues de cette surveillance sont bancarisées dans le Système d'Information Quadrige (**programme RNOPHY**) et font l'objet de rapports diffusés sur le site EnvLit (<http://envlit.ifremer.fr/documents/bulletins/rno>).

La surveillance est coordonnée par Ifremer (réseau ROCCH – coordinatrice : Anne Grouhel). Le protocole de prélèvement et la méthodologie d'analyse ont été développés par Martial Huet (LEMAR/IUEM/UBO Brest) et réalisés par cet expert jusqu'en 2009. Depuis 2010 les suivis sont réalisés par un bureau d'étude (TOXEM) formé par M. Huet.

Les données sont fournies à Ifremer une fois par an, puis saisies par la coordinatrice A. Grouhel. Le producteur de données réalise un rapport détaillé (cf. EnvLit) qui inclut l'analyse des données : il est donc considéré que les données fournies ont déjà été examinées par un expert et peuvent être qualifiées bonnes.

Par conséquent, la qualification des données dans Quadrige correspond à la seule recherche d'erreurs de saisie dans la base ou dans les fichiers fournis à Ifremer. **Le processus mis en place est uniquement de la qualification « automatique »**. Cette qualification est lancée une fois par an sur **les données de l'année n-2** de façon à s'assurer que toutes les données ont été saisies et validées.

Le volume de données est faible (38 passages par an), mais le nombre de résultat a évolué depuis 2015 : de 2003 à 2014, seules les moyennes par échantillon étaient saisies (sexe-ratio, longueur moyenne des pénis des femelles, longueur moyenne des pénis des mâles, indices RPSI (Relative Penis

Size Index) et VDSI (Vas Deferens Sequence Index) calculés pour l'échantillon. A partir de 2015 les mesures individuelles (sexe, longueur de pénis et VDSI) de chaque individu sont saisies (40 individus par échantillon).

1.3. Processus de qualification automatique

Les règles de qualification automatique ont été définies par Didier Claisse en 2014 et complétées en 2017 par Anne Grouhel (coordinatrice du ROCCH et qualificatrice).

Les anomalies sont recherchées par la Cellule Quadriges via des requêtes SQL lancées manuellement. Les anomalies trouvées sont envoyées sous forme de fichier type Excel à la qualificatrice qui y indique en retour les corrections apportées et/ou le niveau de qualité à attribuer aux données.

La Cellule Quadriges qualifie les données en conséquence et produit un rapport de qualification.

2. Critères de qualification RNOPHY

2.1. Définition de la thématique

2.1.1. Liste de programmes

Un seul programme Quadriges est concerné par cette qualification : **le programme RNOPHY** : Volet du Réseau national RNO – nouvellement ROCCH : surveillance de paramètres physiologiques (imposex, développement larvaire,...) pour l'évaluation des effets biologiques des contaminants sur les organismes marins.

2.1.2. Liste de PSFMUs¹

La liste des Paramètres – Support – Fraction – Méthode – Unité (PSFMU) utilisés dans le cadre du programme RNOPHY est détaillée dans le Tableau 1. Tous les résultats associés sont qualifiés. Les mesures de TBT par analyse chimique dans l'eau ou le sédiment ont été qualifiées douteuses car le protocole de prélèvement n'a pas été correctement décrit.

¹ Pour les définitions des concepts liés aux données Quadriges, consulter le dictionnaire SANDRE « Processus d'acquisition des données physico-chimiques, microbiologiques, biologiques et déchets - version 1 – Eaux Littorales » (<http://www.sandre.eaufrance.fr/notice-doc/processus-dacquisition-des-donn%C3%A9es-physico-chimiques-microbiologiques-biologiques-et>).

Tableau 1 : Liste des PSFMUs du programme RNOPHY (au 27/10/2017). En italique : PSFMUs « historiques » qui ne sont plus saisis dans les données actuelles.

Paramètre code	Paramètre libellé	Support	Fraction	Méthode	Unité	Statut	Année début	Année fin	Analysé sur	Commentaire
ETAT_IMPOSEX	Etat du pénis du gastéropode - suivi IMPOSEX	Gastéropode	Pénis	Mesure biométrique à la loupe binoculaire imposex - sans unité	sans unité	Actif	2015	2015	Individu	
FEMELP	Longueur moyenne du pénis chez les femelles	Gastéropode	Sans objet	Mesure biométrique à la loupe binoculaire imposex - mm	mm	Actif	2003	2015	Echantillon	
<i>FEMETAIL</i>	<i>Taille moyenne des femelles</i>	<i>Gastéropode</i>	<i>Sans objet</i>	<i>Mesure au pied à coulisse</i>	<i>mm</i>	<i>Gelé</i>	<i>2003</i>	<i>2005</i>	<i>Echantillon</i>	
<i>FRINF250</i>	<i>fraction < 250 µm</i>	<i>Sédiment, substrat meuble</i>	<i>Frac. tot.</i>	<i>Tamissage voie sèche après lyophilisation - %</i>	<i>%</i>	<i>Actif</i>	<i>2003</i>	<i>2003</i>	<i>Echantillon</i>	
INDVSEX	Sexe de l'individu	Gastéropode	Sans objet	Evaluation visuelle - sans unité	sans unité	Actif	2015	2015	Individu	
<i>INDVTAIL</i>	<i>Taille de l'individu</i>	<i>Gastéropode</i>	<i>Sans objet</i>	<i>Mesure au pied à coulisse</i>	<i>mm</i>	<i>Actif</i>	<i>2003</i>	<i>2005</i>	<i>Echantillon</i>	
LONGUEUR	Longueur ou distance	Gastéropode	Pénis	Mesure biométrique à la loupe binoculaire imposex - mm	mm	Actif	2015	2015	Individu	
MALELP	Longueur moyenne du pénis chez les mâles	Gastéropode	Sans objet	Mesure biométrique à la loupe binoculaire imposex - mm	mm	Actif	2003	2015	Echantillon	
<i>MALETAIL</i>	<i>Taille moyenne des mâles</i>	<i>Gastéropode</i>	<i>Sans objet</i>	<i>Mesure au pied à coulisse</i>	<i>mm</i>	<i>Gelé</i>	<i>2003</i>	<i>2005</i>	<i>Echantillon</i>	
RPSI	Relative Penis Size Index	Gastéropode	Sans objet	RPSI (Relative Penis Size Index) - %	%	Actif	2003	2015	Echantillon	
SEXRATIO	Rapport mâles/adultes	Gastéropode	Sans objet	Rapport - %	%	Actif	2003	2015	Echantillon	
<i>TBT</i>	<i>Tributylétain</i>	<i>Eau filtrée</i>	<i>Sans objet</i>	<i>Chromatographie gaz organostanniques - eau - ng[Sn]/l</i>	<i>ng[Sn]/l-1</i>	<i>Actif</i>	<i>2003</i>	<i>2005</i>	<i>Echantillon</i>	<i>Douteux (acquis selon un protocole non décrit)</i>
<i>TBT</i>	<i>Tributylétain</i>	<i>Sédiment, substrat meuble</i>	<i>Frac. tot.</i>	<i>Chromatographie gaz organostanniques (Michel et Averty (1991) - µg[Sn]/kg</i>	<i>µg[Sn].kg⁻¹</i>	<i>Gelé</i>	<i>2003</i>	<i>2003</i>	<i>Echantillon</i>	<i>Douteux (acquis selon un protocole non décrit)</i>
VDSI	Vas deferens Sequence Index	Gastéropode	Sans objet	VDSI (Vas Deferens Sequence Index) - unité	unité	Actif	2003	2015	Echantillon	
VDS_QUAL	Vas Deference Sequence	Gastéropode	Sans objet	VDS (détermination du stade Vas Deference Sequence) - sans unité	sans unité	Actif	2015	2015	Individu	

2.2. Liste des entités à qualifier

Parmi la liste des entités « qualifiables » dans Quadrige², celles concernées par la qualification RNOPHY sont :

- Les passages = SURVEY
- Les prélèvements = SAMPLING_OPERATION
- Les échantillons = SAMPLE
- Les résultats de mesure = MEASUREMENT

Il n'y a ni résultats sur taxons, ni fichier de mesure, ni photos, ni habitats observés, ni populations initiales.

2.3. Identification des intervenants

Les qualificateurs pour cette thématique sont le coordinateur du réseau ROCCH : Didier Claisse de 2003 à 2014, Anne Grouhel depuis 2015.

Les producteurs de données (Martial Huet jusqu'en 2009, puis un bureau d'études depuis 2010) possèdent l'expertise scientifique nécessaire à la désignation de données douteuses ou fausses et en donnent l'information aux coordinateurs lors de la transmission annuelle des données. Dans ce cas, une qualification ponctuelle peut être réalisée.

Dans le cadre de la qualification automatique seuls les coordinateurs sont sollicités. Ils peuvent toutefois contacter les producteurs de données si des anomalies hors erreur de saisie sont détectées.

2.4. Anomalies recherchées

Les anomalies sont définies par rapport aux consignes de saisies et par rapport aux risques d'erreurs de saisie des données suspectées ou identifiées.

Ces anomalies sont liées à chaque entité à qualifier : Passage (PASS), prélèvement (PREL), échantillon (ECHA) et résultat (RESU), et des anomalies Hors Pack (HPAC). Les anomalies Hors Pack sont des anomalies générales, liées au référentiel Quadrige² ou qui nécessitent une modification profonde des données (saisie de données supplémentaires ou suppression de doublons, rattachement des données aux différents programmes, etc...) ou encore qui concernent plusieurs entités simultanément.

Active	N°anomalie	Niveau priorité	Description
✓	HPAC-001	0	Vérification de la liste des PSFMs (état actif / gelé, cohérence paramètre – support – fraction – méthode)
✓	HPAC-002	0	Conflits thématiques : recherche des autres programmes liés aux données RNO, et des résultats d'autres thématiques rattachés aux données in situ RNO

Active	N°anomalie	Niveau priorité	Description
✓	HPAC-003	0	Champs vides : vérification que les champs qui n'ont pas lieu d'être remplis dans le cadre du RNO sont bien vides. Ont été contrôlés : - Passages : observations terrain, habitat observé, zone de destination dragage, nombre d'individus - Prélèvements : Lot, taille de prélèvement et nombre d'individus - Echantillons : groupe de taxon et taille de l'échantillon
✓	HPAC-004	1	Support-taxon : vérification de la cohérence des supports d'échantillon avec le taxon support associé.
✓	HPAC-005	2	Niveau de saisie : vérification que toutes les PSMFs des stratégies RNO ont le même niveau de saisie (échantillon en l'occurrence), puis vérification que tous les résultats sont saisis au niveau prévu dans la stratégie
✓	HPAC-006	3	Doublons de passages : vérification qu'il existe un seul passage par lieu – date.
✓	HPAC-007	3	Doublons de résultats : vérification qu'il existe un seul résultat de mesure par : lieu – date – taxon – paramètre – méthode – n° d'individu (depuis 2015).
✓	HPAC-008	4	Nombre de résultats saisis : 1) Nb résultats par paramètre > nombre d'individus de l'échantillon (depuis 2015) 2) Nb résultats INDVSEX différent du nombre d'individus de l'échantillon
✓	HPAC-009	5	Données non validées
✓	PASS-001	1	Heure de passage = 00:00:00 (à supprimer)
✓	PASS-002	2	Unité de sonde du passage <> « m »
✓	PREL-001	1	Heure de prélèvement = 00:00:00 (à supprimer)
✓	PREL-002	1	Unité d'immersion « m » (l'unité est « cm » pour RNOSED puisqu'il s'agit d'horizons sédimentaires).
✓	PREL-003	2	Combinaison engin – niveau – immersion – support : vérification de la cohérence des 4 valeurs
	<i>PREL-004</i>	3	<i>Prélèvement portant un résultat TBT (à qualifier Douteux – protocole non précisé)</i>
	<i>ECHA-001</i>	1	<i>Echantillon Gastéropode : à qualifier Bon (2003-2013)</i>
	<i>ECHA-002</i>	1	<i>Echantillon Sédiment sans TBT (2003-2013) : à qualifier Bon</i>
	<i>ECHA-003</i>	1	<i>Echantillon portant un résultat TBT support Eau ou sédiment (2003-2013) : à qualifier Douteux</i>
✓	ECHA-004	1	Echantillon dont le nombre d'individus est soit vide, soit supérieur à 40 individus (depuis 2015)
	<i>RESU-001</i>	1	<i>Résultats sur gastéropodes (2003-2013) : Bon</i>
	<i>RESU-002</i>	1	<i>Résultats FRINF250 (2003-2013) : Bon</i>
	<i>RESU-003</i>	1	<i>Résultats TBT (2003-2013) : Douteux (protocole non précisé)</i>
✓	RESU-001	1	Résultats VDS et sexe des individus : pour les individus femelle, le résultat VDS doit être présent. Pour les mâles il ne doit pas y avoir de résultat VDS.
✓	RESU-002	1	Résultats dont le nombre de décimales est >2
✓	RESU-003	1	Longueur de pénis hors des bornes [0 ; 6,5 cm]

2.5. Règles de contrôle Quadrigé

Aucune règle de contrôle à la saisie n'est appliquée.

3. Qualification experte

Les données RNOPHY ne font pas l'objet d'une qualification experte, étant considéré que le producteur de données est l'expert et qu'il signale toute anomalie dans les jeux de données en amont de la bancarisation.

Si un problème survient au cours de l'acquisition des données, il est tracé dans Quadrigé dès que les données sont saisies et validées via une qualification manuelle des données.