

Actualisation de la classification des sites de mesure

Mathieu Joly

24 mai 2019

Table des matières

1	Données utilisées	2
2	Comparaison à la précédente version	3
3	Quantification des anomalies	3
4	Évolution du jeu de stations classifiées	3
5	Cartographie du résultat	3
6	Conclusion	9

1 Données utilisées

- Cette nouvelle version utilise des données AQeR pré-traitées par l'INERIS, et ne comprend pas le NO. Pour la première fois des données « non validées » ont été intégrées (2018) pour prendre en compte le réseau de stations le plus récent.
- 8 années sont considérées : données AirBase V8 de 2011 à 2012, données AQeR « validées » de 2013 à 2017, et données AQeR « temps réel » pour 2018.
- Ne sont pas pris en compte les sites d'altitude supérieure à 1400 m (altitude à partir de laquelle le nombre de stations diminue fortement). En Europe, ces stations sont peu nombreuses, mais ne peuvent pas être confondues avec les sites de plaine pour l'analyse.
- Les stations renseignées comme « industrielles » ne sont pas prises en compte. La variabilité temporelle de ce type de mesure est très difficile à caractériser, et la méthode n'est pas suffisamment robuste pour appréhender le comportement potentiellement erratique des indicateurs calculés.

À partir des métadonnées, on dérive la typologie simplifiée suivante :

Type R : sites qualifiés *background* et *rural*.

Type S : sites qualifiés *background* et *suburban*.

Type U : sites qualifiés *background* et *urban*.

Type T : sites qualifiés *traffic* et *urban*.

Type O : toutes les autres stations, qui ne seront pas prises en compte pour l'Analyse Discriminante, mais qui seront classifiées *a posteriori*.

La figure 1 montre qu'en dépit du renouvellement du quart des données (période d'étude décalée de 2 ans), le réseau de mesure évolue peu. Le CO et l'ozone font toujours figure d'anomalie, avec beaucoup de stations T dans un cas, et beaucoup de stations U et S dans l'autre. Par ailleurs, en dehors des PM dont le nombre de stations augmente significativement depuis plusieurs années, certains polluants (en particulier le CO) connaissent une diminution significative du nombre de sites, alors même qu'on s'attendrait à un étoffement du réseau de mesure au fil des années.

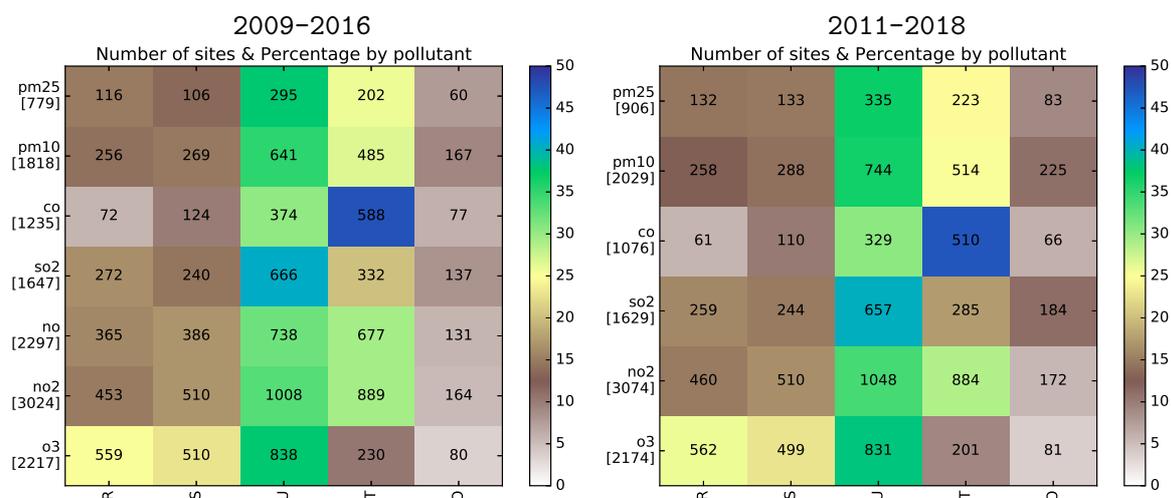


Figure 1 – Nombre de stations sélectionnées (données suffisantes), par type de métadonnée. Les couleurs correspondent au pourcentage par polluant.

2 Comparaison à la précédente version

Pour les stations en commun dans les deux classifications, la figure 2 compare les classes obtenues. C'est pour le SO₂ et les PM_{2,5} que les deux versions sont le plus différentes. Pour tous les autres polluants, la corrélation est supérieure à 0.92.

La figure 3 compare les « validations croisées » par rapport aux types dérivés des métadonnées. La cohérence entre les classifications subjective (métadonnées) et objective est stable, sauf pour les PM_{2,5}, dont la performance se dégrade du fait d'un petit nombre de stations urbaines et trafic qui obtient des classes basses (1-3).

3 Quantification des anomalies

Nous allons nous intéresser aux comportements marginaux de la figure 3 :

- le pourcentage des stations R qui se retrouvent dans les classes 6-10.
- le pourcentage des stations S, U et T qui se retrouvent dans les classes 1-3.

	O ₃	NO ₂	SO ₂	CO	PM ₁₀	PM _{2,5}
R 6-10	5 → 6	2 → 2	26 → 27	16 → 9	18 → 16	21 → 28
S+U+T 1-3	8 → 9	3 → 3	10 → 11	3 → 2	7 → 6	9 → 11

Tableau 1 – Pourcentage des anomalies (cf. paragraphe ci-dessus). Évolution entre l'ancienne et la nouvelle classification (en vert pour une amélioration, en rouge pour une détérioration, et surligné de jaune quand plus de 2% des stations sont affectées).

Le tableau 1 montre une amélioration pour les stations rurales du CO et des PM₁₀, mais confirme la dégradation pour les PM_{2,5}, dont les performances rejoignent celles du SO₂.

Les cartes 4 cartographient les anomalies du tableau 1. L'analyse est difficile, car il faudrait regarder localement la configuration de chacun de ces sites « douteux », et les sources de pollution environnantes. On notera néanmoins que pour les PM, les stations problématiques (cf. § ci-dessus) se trouvent majoritairement en France et en Belgique.

4 Évolution du jeu de stations classifiées

La figure 5 permet de suivre l'évolution du jeu de données classifiées. On notera l'apparition de nouvelles stations en Turquie autour du Bosphore. NB : certains ajouts correspondent en réalité à des changements de codes.

5 Cartographie du résultat

Les cartes de la figure 6 illustrent la classification obtenue pour chaque polluant.

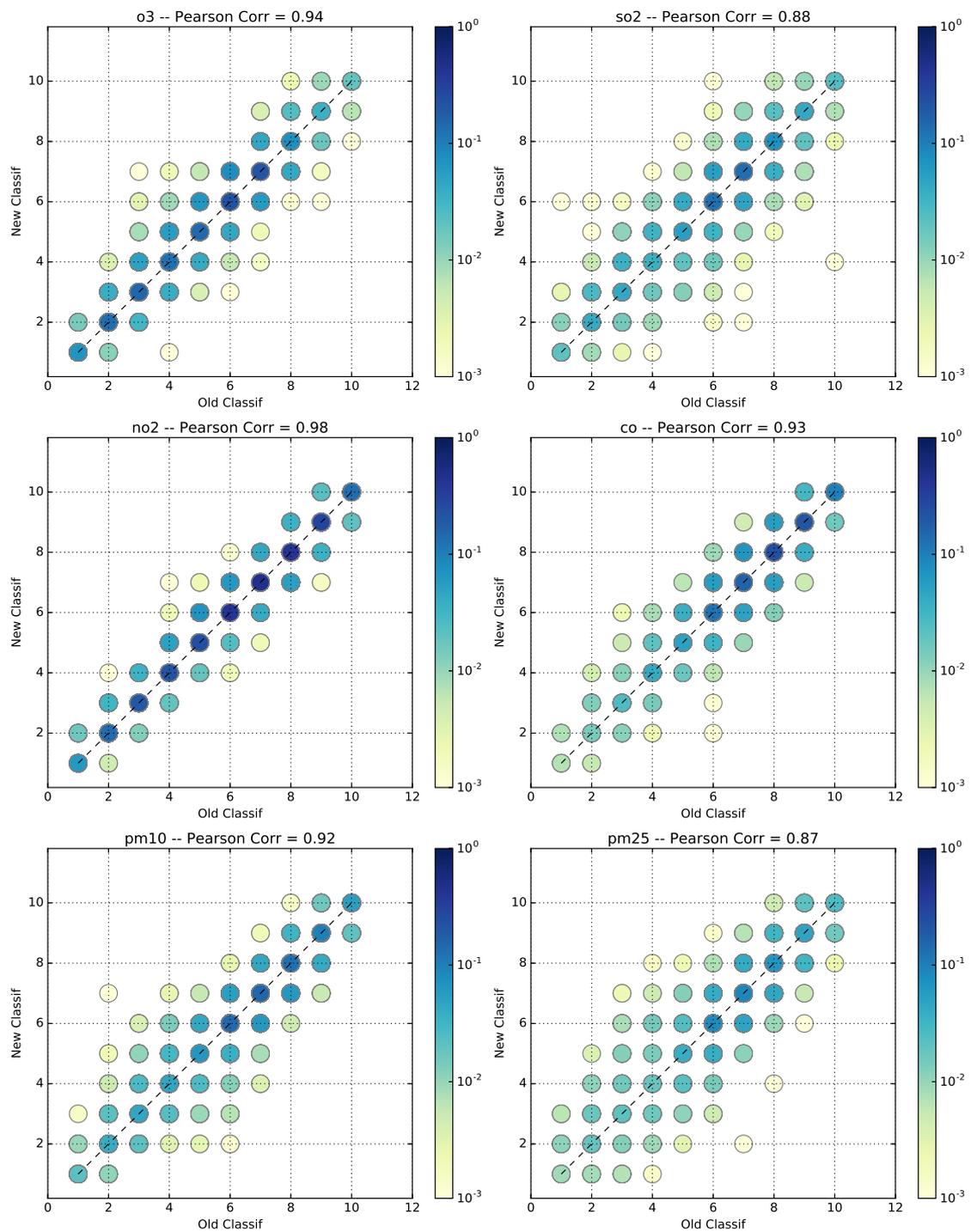


Figure 2 – Scatter Plot des classes obtenues avec l’ancienne et la nouvelle classification. La couleur indique la fréquence d’occurrence.

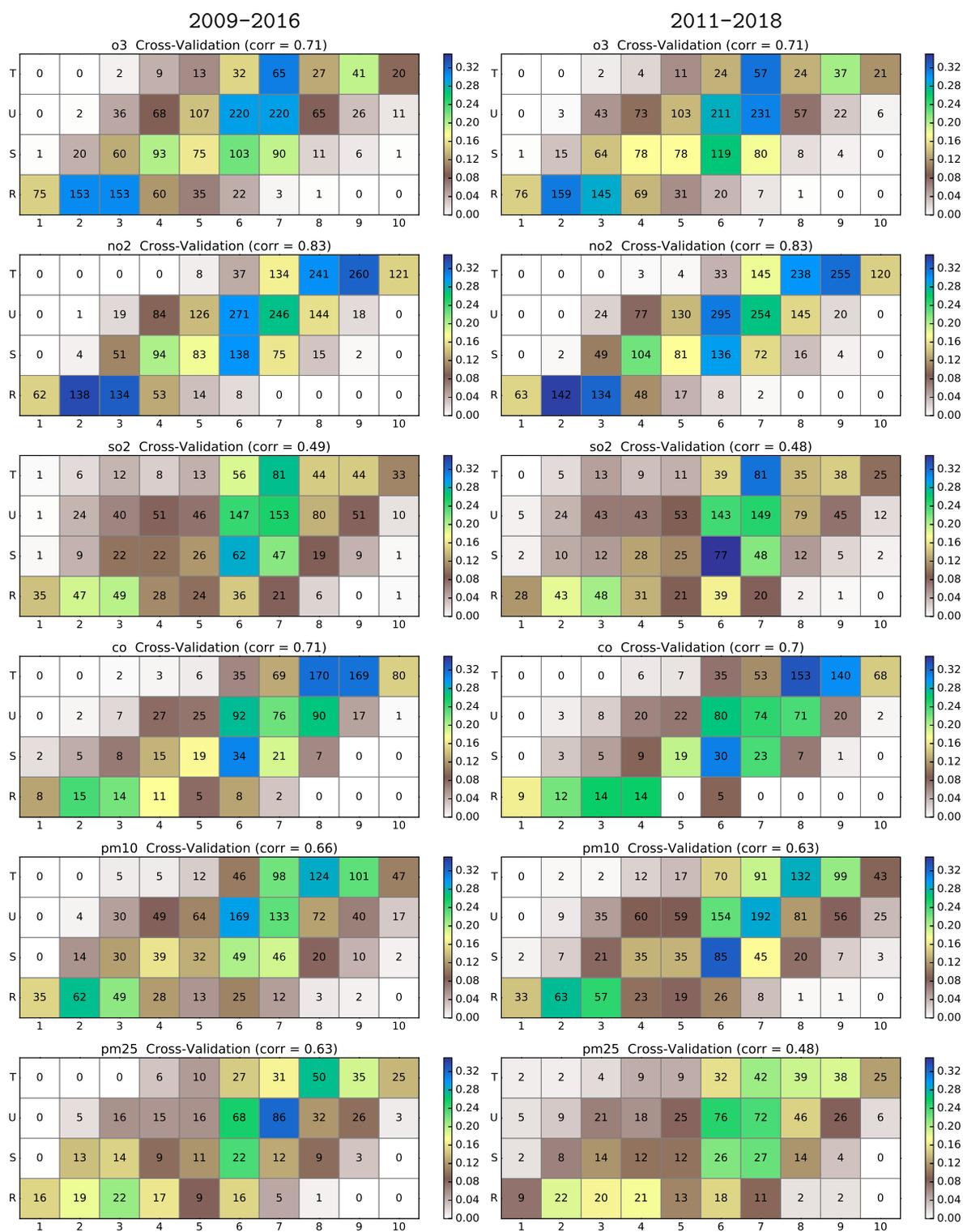


Figure 3 – Validation croisée : nombre et pourcentage (en couleur) dans chaque classe pour chaque type de station. À gauche, pour la précédente classification ; et à droite, pour la nouvelle version.

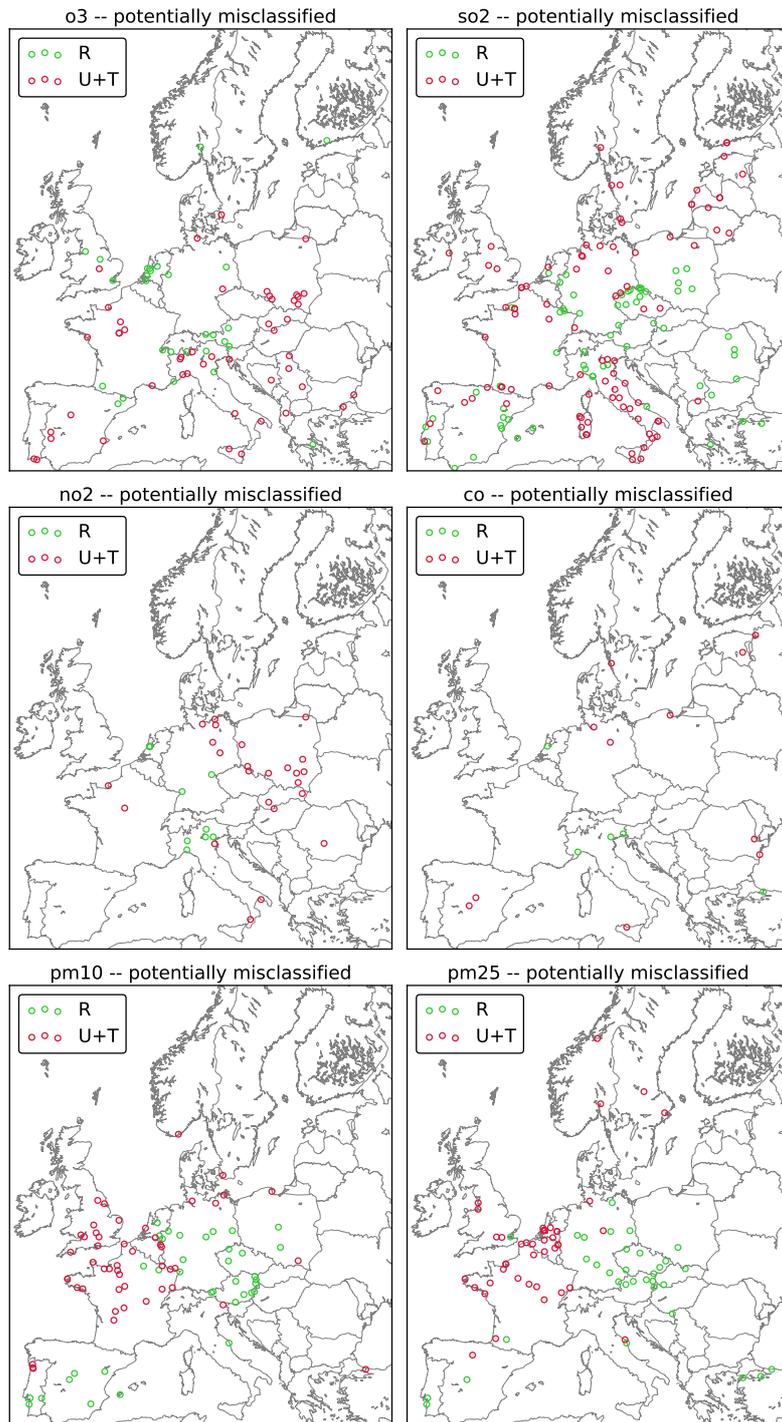


Figure 4 – Stations R qui se retrouvent dans les classes 6-10, et stations U et T qui se retrouvent dans les classes 1-3.

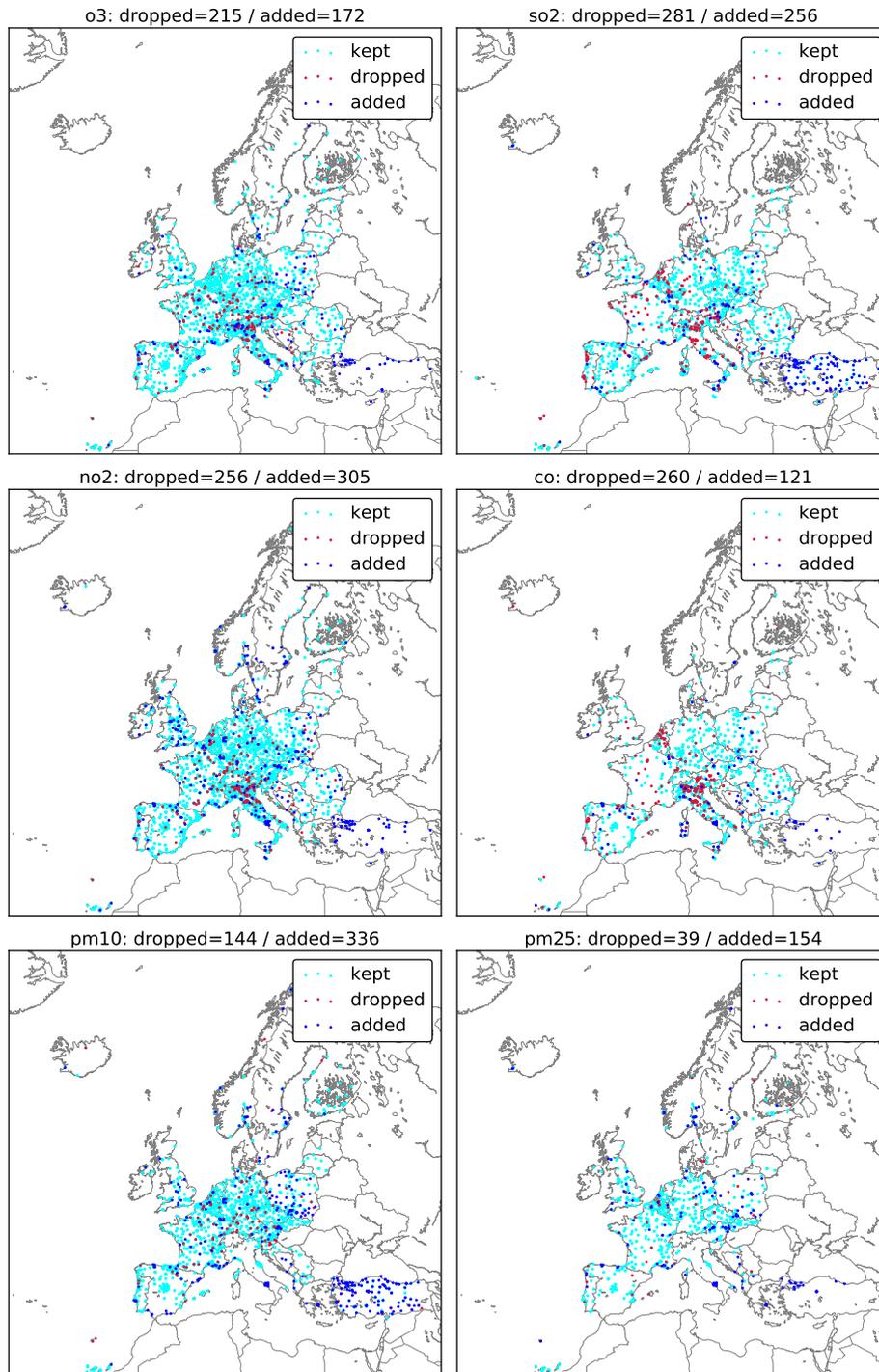


Figure 5 – Stations qui disparaissent (rouge), ou qui apparaissent (bleu) dans la nouvelle version.

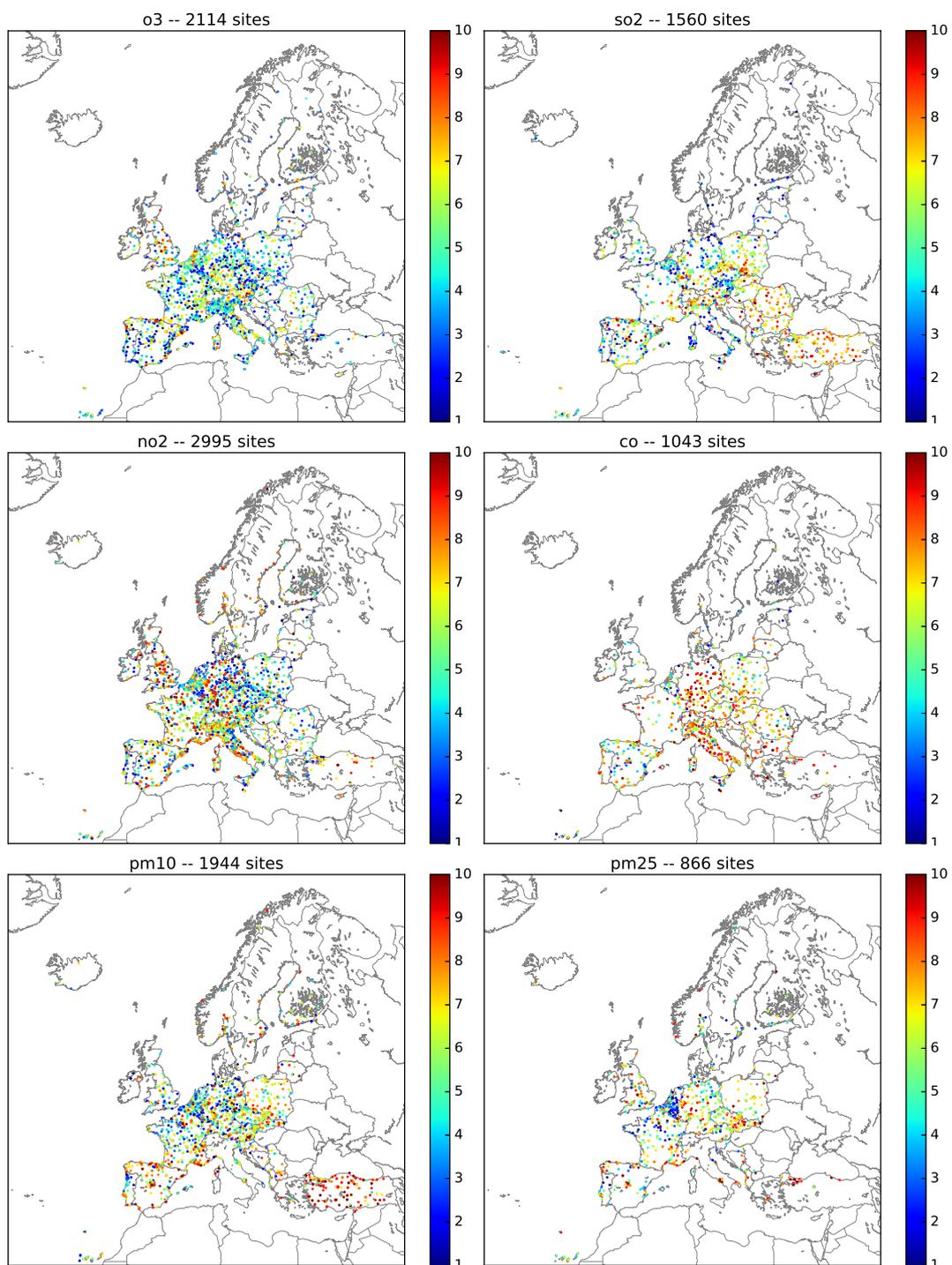


Figure 6 – Cartographie de la classification obtenue.

6 Conclusion

Les 8 années 2011–2018 considérées pour cette mise à jour de 2019 comprennent pour la première fois des données « non validées » récentes. Par ailleurs, en dehors des années 2011–2012 qui proviennent encore d’AirBase V8, les données AQeR ont été pré-traitées par l’INERIS, et ne comprennent pas le NO.

Le décalage de deux ans de la période d’étude fait évoluer substantiellement la classification :

- En dehors des PM, dont le nombre de stations augmente, certains polluants (en particulier le CO et le SO₂) connaissent ces dernières années une diminution significative du nombre de sites répertoriés.
- C’est pour le SO₂ et les PM_{2,5} que la méthode semble le moins robuste (nombreuses différences pour les sites en commun dans l’ancienne et la nouvelle classification).
- La cohérence entre les métadonnées et la classification objective est globalement stable pour la plupart des polluants, s’améliore pour le CO et les PM₁₀, mais se dégrade pour les PM_{2,5}.

Cela confirme les précautions nécessaires lors de l’utilisation de la classification objective pour le SO₂ et les PM_{2,5}.