

Direction de la recherche forestière

Titre :	Utilisation des données d'accroissement potentiel en surface terrière ou d'IQS potentiel à des fins de classification du potentiel de productivité des polygones forestiers
Responsable :	Catherine Périé, biol., <i>Ph. D.</i>
Collaboration :	Marie-Claude Lambert, statisticienne, M. Sc.
Date :	Janvier 2017

1- Contexte

En 2012, la Direction de la recherche forestière (DRF) a publié un Mémoire de recherche forestière intitulé *Prédire la croissance potentielle des arbres au Québec à l'aide des caractéristiques cartographiables des polygones et des stations* (Périé *et al.* 2012). Dans cette étude, des modèles linéaires mixtes ont été étalonnés à l'échelle provinciale pour prévoir l'accroissement potentiel en surface terrière (APST) des 10 espèces les plus abondantes de la forêt commerciale du Québec, ainsi que de 4 groupes d'espèces qui n'ont pas été modélisées individuellement. À l'échelle provinciale et selon les espèces (ou les groupes d'espèces), les coefficients de détermination (R^2) des modèles varient de 0,27 à 0,58, tandis que l'erreur (la racine carrée de l'erreur quadratique moyenne associée à la prédiction) varie de 17 à 35 %. Par ailleurs, la validation des modèles a démontré que les biais de prédiction sont négligeables à l'échelle provinciale. Toutefois, cette absence de biais n'a pas été démontrée à une autre échelle spatiale. Cela signifie qu'à l'échelle locale ou régionale, ces modèles pourraient systématiquement surestimer ou sous-estimer l'APST des espèces.

En 2010, la Direction des inventaires (DIF) a fait un exercice similaire pour les IQS potentiels (Müssenberger *et al.* 2010). Ces données d'IQS potentiels et d'APST sont rendues disponibles à l'ensemble des utilisateurs du ministère des Forêts, de la Faune et des Parcs (MFFP) dans le système de diffusion des données écoforestières (DDE).

Cet avis technique explique comment utiliser les prédictions d'APST ou d'IQS potentiels et les valeurs des intervalles de confiance qui leur sont associés pour localiser sur le territoire d'étude les polygones forestiers qui présentent, en moyenne, le meilleur potentiel de productivité pour chaque essence.

On peut citer tout ou partie de ce texte en indiquant la référence
© Gouvernement du Québec

2- Intervalles de confiance : une information indispensable pour déterminer si une valeur est statistiquement différente d'une autre

Les données de productivité potentielle (IQS potentiel ou APST) de chaque polygone de la carte écoforestière sont issues de modèles statistiques. Ainsi, pour une essence donnée, la valeur de productivité associée à chaque polygone est une **estimation**, nécessairement entachée d'erreur, de la valeur réelle. Cette erreur est inversement proportionnelle à la précision du modèle prédictif mesurée par le R^2 des différents modèles. Une valeur de R^2 qui se rapproche de 1 indique que l'erreur est petite et que le modèle est précis. À l'inverse, une valeur de R^2 qui se rapproche de 0 indique que l'erreur est grande et que le modèle manque de précision.

Pour pouvoir comparer deux valeurs estimées de productivité potentielle, il est nécessaire de considérer à la fois les valeurs estimées et les erreurs qui leur sont associées.

L'intervalle de confiance à 95 %* permet d'évaluer la précision de l'estimation d'une moyenne. Plus un intervalle de confiance est étroit autour de la valeur estimée, plus l'estimation est précise. Lorsque les intervalles de confiance de deux prédictions ne se chevauchent pas, les deux prédictions sont statistiquement différentes l'une de l'autre.

3- Différence entre l'intervalle de confiance et l'intervalle de prévision

Afin d'interpréter correctement un intervalle de confiance, il est nécessaire de bien comprendre la différence entre un intervalle de confiance et un intervalle de prévision.

Un **intervalle de confiance** est l'étendue dans laquelle il est probable de retrouver la réponse moyenne pour des valeurs données des variables explicatives du modèle. L'intervalle de confiance présente donc l'étendue possible des valeurs de la moyenne, mais ne donne pas d'information sur la distribution des prévisions individuelles. Prenons l'exemple fictif d'un modèle de prévision de l'APST d'une espèce donnée qui aurait comme seule variable explicative la surface terrière marchande. Si l'APST prédit avec ce modèle est de $12 \text{ cm}^2\cdot\text{an}^{-1}$ pour une surface terrière de $20 \text{ m}^2\cdot\text{ha}^{-1}$, avec un intervalle de confiance de 10 à $14 \text{ cm}^2\cdot\text{an}^{-1}$, nous pouvons être confiants à 95 % que cette étendue inclut la moyenne des APST de tous les polygones ayant une surface terrière marchande de $20 \text{ m}^2\cdot\text{ha}^{-1}$. Toutefois, cette étendue ne donne aucune information sur la valeur possible de l'APST de l'un de ces polygones en particulier.

* Intervalle de confiance à 95 % : si l'on tire 100 échantillons et que l'on construit 100 intervalles de confiance sur la base de ces échantillons, alors la « vraie » valeur aurait été comprise dans 95 de ces intervalles (en moyenne).

L'**intervalle de prévision** est l'étendue dans laquelle il est probable de retrouver la réponse d'une observation en particulier pour des valeurs données des variables explicatives du modèle. À partir de l'exemple précédent, avec un intervalle de prévision de 4 à 20 cm²·an⁻¹, nous pouvons être confiants à 95 % que cette étendue inclut la valeur de l'APST d'un polygone particulier choisi sur une carte et ayant une surface terrière marchande de 20 m²·ha⁻¹.

L'intervalle de prévision est toujours plus grand que l'intervalle de confiance correspondant, car il y a plus d'incertitude à prédire la valeur d'un polygone en particulier que la moyenne d'un groupe de polygones. L'intervalle de confiance n'inclut que la variabilité associée aux paramètres de la régression, tandis que l'intervalle de prévision inclut aussi la variabilité des observations autour du modèle de régression. Plus les observations varient autour du modèle, plus la valeur de R² est faible et plus l'intervalle de prévision est grand.

Pour illustrer cela, prenons l'exemple de l'unité d'aménagement UA 06152 qui contient 17 159 polygones forestiers ne présentant pas de contraintes majeures à l'exploitation forestière. La figure 1 présente la prévision de l'APST du bouleau jaune (BOJ) de chacun de ces polygones, avec l'intervalle de confiance et l'intervalle de prévision de chacun d'eux.

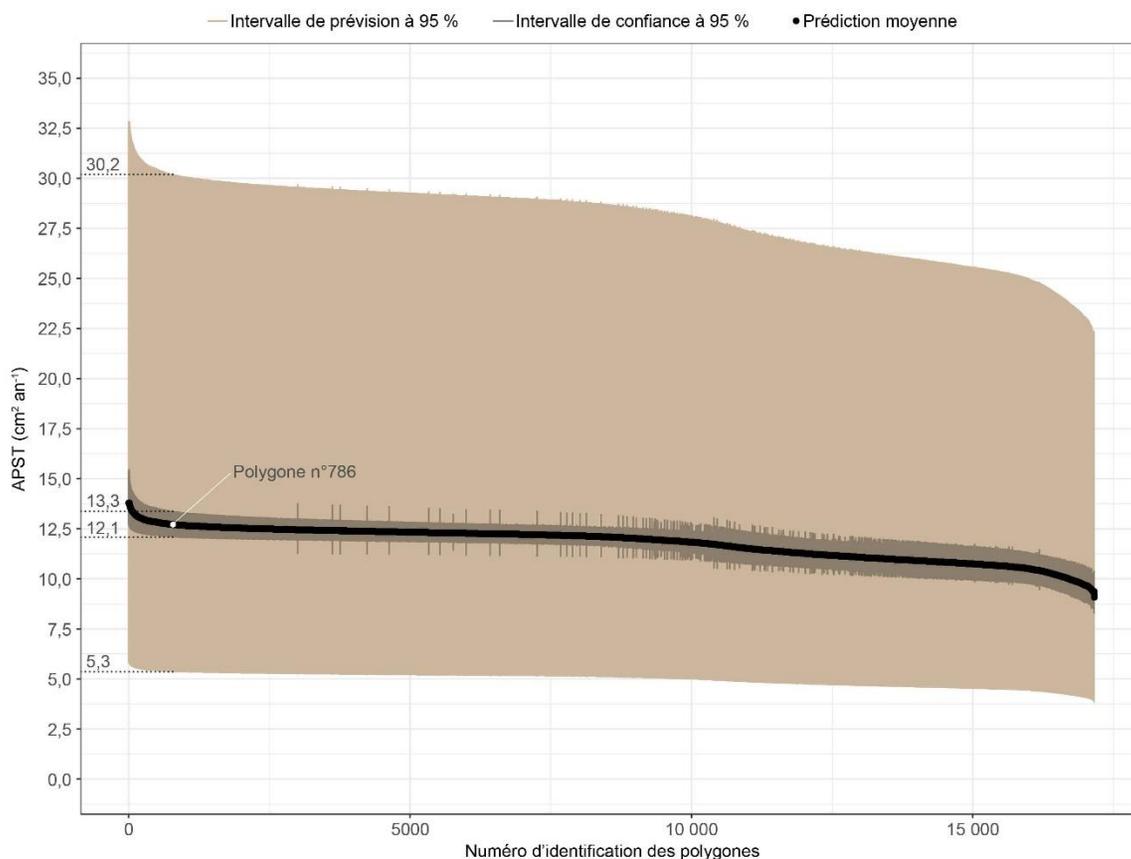


Figure 1 : Intervalles de confiance, intervalles de prévision et prédiction moyenne de l'accroissement potentiel en surface terrière (APST) du BOJ dans l'unité d'aménagement 06152.

Afin d'aider le lecteur à bien comprendre cette figure, nous avons choisi un polygone au hasard parmi les 17 159 polygones de la figure 1 : le polygone 786, dont le géocode est 477536,43+347991,57. Si l'estimation provient seulement de ce polygone unique, l'intervalle de prévision avec 95 % de confiance de la valeur estimée de l'APST pour ce polygone va de 5,3 à 30,2 cm²·an⁻¹. Par contre, si plusieurs polygones ayant les mêmes caractéristiques sont utilisés, la valeur moyenne estimée de l'APST sera incluse dans l'intervalle de confiance allant de 12,1 à 13,3 cm²·an⁻¹.

4- L'outil CLASS pour classer les polygones forestiers en fonction de leur potentiel de productivité

Dans le cadre de l'exercice de classification du potentiel de productivité des polygones forestiers, c'est la notion d'intervalle de confiance, et non celle d'intervalle de prévision, qui a été retenue. Ceci implique que les polygones identifiés comme les plus productifs le sont en moyenne, mais qu'il n'est pas possible de dire si chacun d'eux atteint individuellement un haut potentiel de productivité.

L'application CLASS[†] développée par la DRF, en collaboration avec la DIF, permet de localiser sur le territoire des ensembles (groupes) de polygones qui, en moyenne, ont des potentiels de croissance statistiquement différents. Certains auront donc, en moyenne, des valeurs de productivité potentielle meilleures que d'autres. Toutefois, l'outil ne permet pas d'estimer précisément la productivité potentielle des polygones individuels.

Reprenons le même exemple qu'à la section précédente. Supposons que l'on veuille localiser les polygones qui, en moyenne, ont le meilleur APST pour le BOJ et que la superficie cumulative recherchée de cet ensemble de polygones couvre au moins 20 % de la superficie forestière favorable à l'espèce dans l'unité d'aménagement. Pour ce faire, il suffit de classer les polygones par ordre décroissant d'APST, puis d'additionner les superficies des polygones (en commençant par celui qui a le meilleur APST) jusqu'à ce qu'elles couvrent 20 % de la superficie forestière favorable à l'espèce. Ces polygones (Figure 2, groupe A, en vert foncé) représentent ceux avec, en moyenne, le meilleur APST.

Pour identifier l'ensemble des polygones qui ont, en moyenne, un APST significativement plus faible que ceux du groupe A (Figure 2), il faut calculer la valeur minimale de la borne inférieure des intervalles de confiance pour ce groupe. Dans notre exemple, l'ensemble des meilleurs APST regroupe 3 586 polygones, et le minimum des bornes inférieures des intervalles de confiance d'APST du groupe A est de 11,3 cm²·an⁻¹. Les polygones qui ont un APST significativement plus faible que ceux du groupe A sont ceux pour lesquels la borne supérieure de l'intervalle de confiance est plus petite que 11,3 cm²·an⁻¹. Cet ensemble correspond au groupe D de la figure 2, constitué de 1 154 polygones dont la prévision de l'APST

[†] L'application web CLASS développée pour effectuer cette classification du potentiel de productivité est disponible à cette adresse : <http://214710PT/CLASS/>. La façon de l'utiliser est expliquée dans l'application web elle-même.

varie, en moyenne, de 8,3 à 11,2 $\text{cm}^2 \cdot \text{an}^{-1}$. Les APST moyens des polygones du groupe B ne sont pas considérés comme statistiquement différents de ceux du groupe A ni du groupe C (puisque leurs intervalles de confiance se chevauchent), mais ils le sont de ceux du groupe D. Les APST moyens des polygones du groupe C ne diffèrent pas statistiquement de ceux des 3 autres groupes, puisque leurs intervalles de confiance chevauchent tous les autres (Figure 2).

Même si l'ensemble des polygones du groupe A ont, **en moyenne**, un meilleur APST pour le BOJ que ceux du groupe D, on ne peut garantir que l'APST de chacun des polygones de l'ensemble A soit meilleur que celui d'autres polygones ailleurs sur le territoire. **On ne peut pas non plus s'attendre à ce que tous les polygones du groupe A aient un bon potentiel de productivité.**

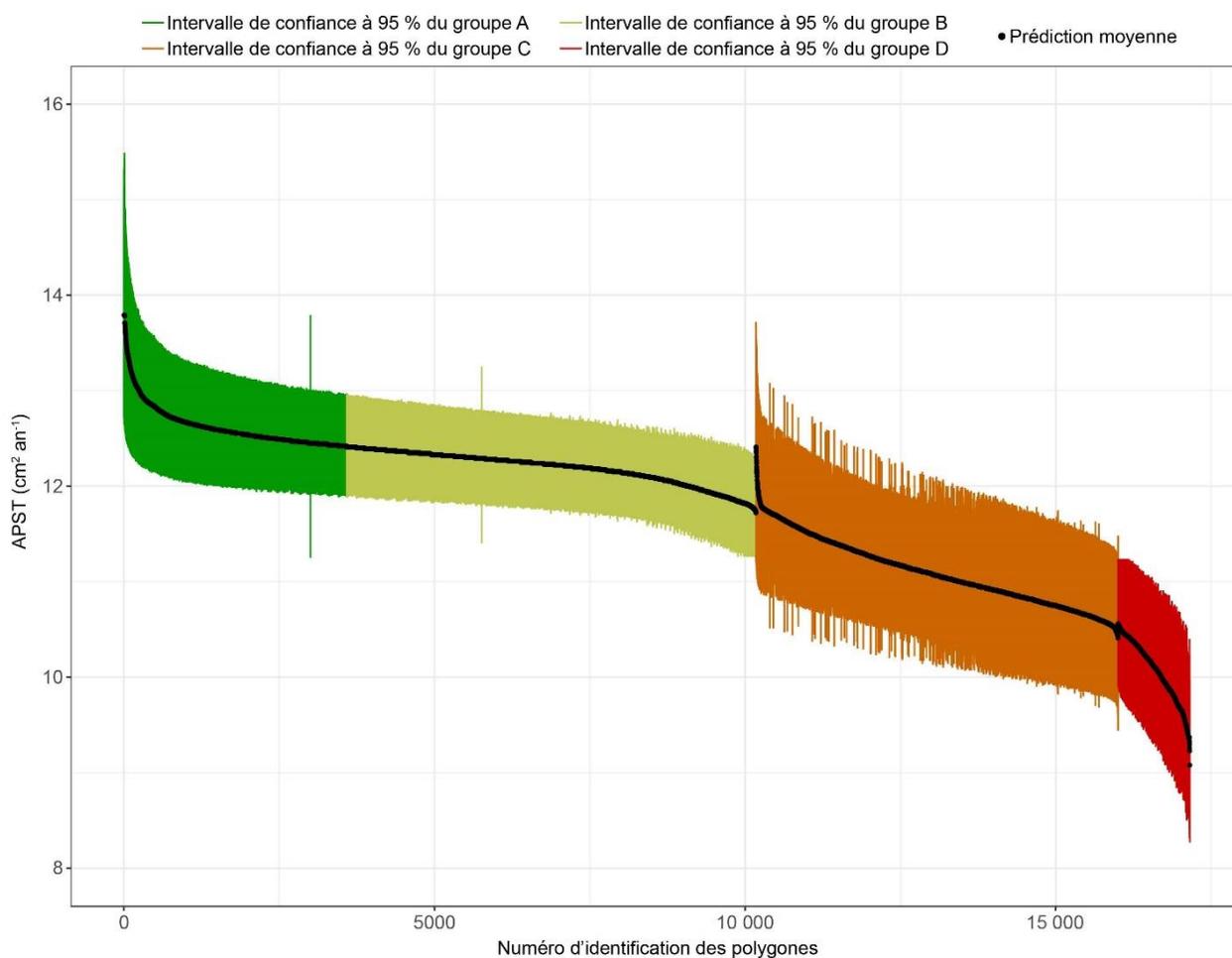


Figure 2 : Prédiction moyenne et intervalle de confiance à 95 % de l'accroissement potentiel en surface terrière (APST) du BOJ de chaque polygone de l'UA 06152. En moyenne, les polygones du groupe A ont un meilleur APST que ceux du groupe D. L'APST moyen des polygones du groupe B diffère aussi significativement de celui du groupe D, mais pas de celui des polygones des groupes A et C.

La figure 3 illustre la répartition spatiale de ces différents ensembles de polygones.

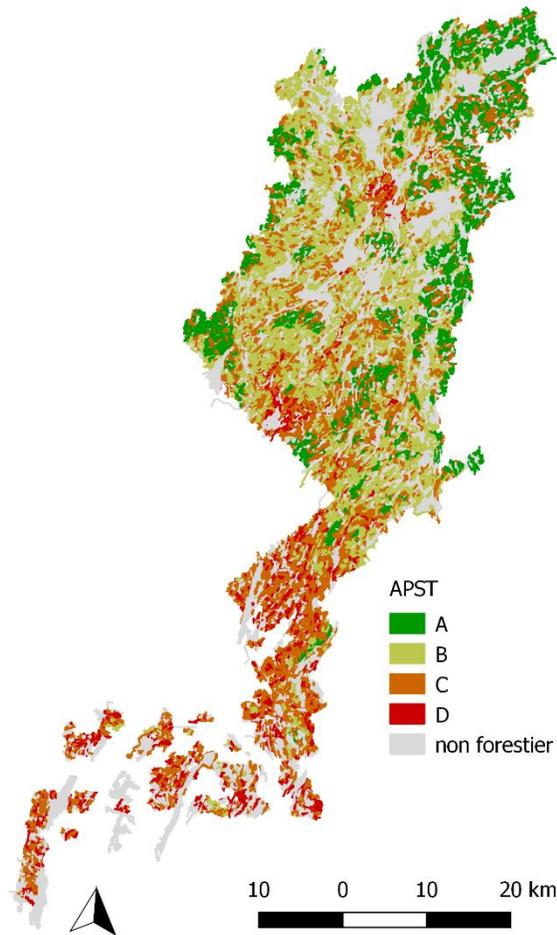


Figure 3. Répartition spatiale des polygones de l'UA 06152 classés en 4 groupes selon l'estimation de l'accroissement potentiel en surface terrière (APST) moyen de chaque polygone pour le BOJ. Les groupes A, B C et D sont décrits à la section 4 et à la figure 2.

5- Conclusion

L'application web CLASS permet de localiser, sur un territoire donné et pour une essence donnée, un ensemble de polygones qui ont, en moyenne, le meilleur potentiel de productivité. Cependant, il n'y a aucune garantie que le potentiel de productivité de chacun des polygones individuels dans cet ensemble soit meilleur que celui d'un autre polygone ailleurs sur le territoire. **On ne peut pas non plus s'attendre à ce que chacun des polygones de cet ensemble ait un bon potentiel de productivité.**

Si l'on voulait s'assurer de localiser les meilleurs polygones de cet ensemble, il faudrait compléter l'analyse avec d'autres sources de données, par exemple des données de productivité **mesurées** (et non estimées à partir de modèles) acquises par des inventaires de terrain.

6- Références

Mussenberger, F., S. Miron, M. Riopel, V. Laflèche, J. Bégin et J.-P. Saucier, 2010. *Le potentiel d'accroissement d'une station écoforestière : prédiction de l'indice de qualité de station et de l'accroissement annuel moyen potentiel par polygone forestier pour la forêt publique commerciale du Québec*. Université Laval. Québec, QC. 251 p.

[\[http://www2.sbf.ulaval.ca/dendro/Recherche/Rapports/Rapport_IQS_FINAL.pdf\]](http://www2.sbf.ulaval.ca/dendro/Recherche/Rapports/Rapport_IQS_FINAL.pdf)

Périé, C., L. Duchesne et M.-C. Lambert, 2012. *Prédire la croissance potentielle des arbres au Québec à l'aide des caractéristiques cartographiables des polygones et des stations*. Gouvernement du Québec, ministère des Ressources naturelles, Direction de la recherche forestière. Mémoire de recherche forestière n° 164. 44 p.

[\[https://www.mffp.gouv.qc.ca/publications/forets/connaissances/recherche/Perie-Catherine/Memoire164.pdf\]](https://www.mffp.gouv.qc.ca/publications/forets/connaissances/recherche/Perie-Catherine/Memoire164.pdf)