



HAL
open science

Utilisation d'un Système d'Information Géographique pour représenter la distribution spatiale de composés chimiques ou organiques selon le type de compartiments d'un organisme. Exemple de la répartition de la concentration de glucides non structuraux dans un arbre.

Erwin Thirion

► **To cite this version:**

Erwin Thirion. Utilisation d'un Système d'Information Géographique pour représenter la distribution spatiale de composés chimiques ou organiques selon le type de compartiments d'un organisme. Exemple de la répartition de la concentration de glucides non structuraux dans un arbre.. Les Cahiers des techniques de l'INRA, 2021, pp.1-11. hal-03224041

HAL Id: hal-03224041

<https://hal.science/hal-03224041>

Submitted on 11 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Utilisation d'un Système d'Information Géographique pour représenter la distribution spatiale de composés chimiques ou organiques selon le type de compartiments d'un organisme.

Exemple de la répartition de la concentration de glucides non structuraux dans un arbre.

Erwin Thirion¹



Erwin THIRION – BTS Gestion Forestière (2007), Licence Ecologie et Aménagement (2008), Licence professionnelle Création et Administration de Système d'Information Géographique (2010) et master Agrosceience, Environnement, Territoire, Paysage, Forêt (Validation d'Acquis d'Expérience 2020).

2014 – aujourd'hui Technicien des Milieux Naturels et Géo-Anthropisés au sein de l'UMR Silva de Nancy.

Compétences en pédologie et croissance racinaire, en botanique et en dendrochronologie. Utilisation des Systèmes d'Informations Géographiques et du logiciel de programmation R lors de la préparation des campagnes de terrain, l'analyse de données et l'archivage des données.

Résumé. La gestion de données spatiales et leur représentation géographique s'est énormément développée grâce à l'essor des Systèmes d'Informations Géographiques (SIG). Cet article propose d'utiliser les fonctionnalités de ces outils pour créer une illustration afin de représenter spatialement l'intensité d'un processus ou la distribution d'un composé au sein d'un organisme vivant (comme par exemple les parties aériennes et souterraines, organes d'âges différents ...). Cette tâche n'est pas aisée car aucun outil spécifique n'est disponible. Cet article présente une méthode reposant sur le logiciel SIG ArcGIS et les bibliothèques SIG de R pour réaliser cette représentation d'information à l'aide d'une couche d'information de type vecteur (découpage en entités distinctes qui s'oppose au type raster qui représente l'espace dans un maillage régulier). Techniquement, la démarche proposée consiste 1) à utiliser un SIG (ArcGIS) pour vectoriser une image représentant l'organisme, puis 2) produire une illustration à partir de la couche vecteur créée selon le découpage souhaité. Plus particulièrement, l'image choisie pour représenter l'organisme étudié est vectorisée en objets de type polygone correspondant à des compartiments d'intérêt de l'organisme. Enfin la représentation de l'intensité du phénomène selon les différents compartiments est accomplie à l'aide de deux logiciels, ArcGIS et les bibliothèques SIG du logiciel libre de statistique R. Ce deuxième présente l'avantage, une fois le fichier de forme créé, de pouvoir être implémenté dans un script d'analyse de données R, logiciel utilisé couramment dans la communauté scientifique. Cette seconde approche a permis, dans l'exemple présenté dans cet article, de représenter la répartition des concentrations en azote ou en réserves glucidiques dans les différents compartiments d'un hêtre à différentes dates et soumis à des traitements contrastés.

¹ INRAE, UMR SILVA, F-54260 Champenoux ; erwin.thirion@inrae.fr

Mots clés : illustration, SIG, vectorisation, R, programmation, script

Abstract. The management of spatial data and their geographical representation has developed thanks to Geographical Information System (GIS). This article proposes to use the functionalities of these tools to create an illustration in order to spatially represent the intensity of a process or the distribution of a compound within a living organism (such as aerial and undergrounds parts, organs of different ages ...). This task is not easy because no specific tools is available. This paper presents a method based on ArcGIS GIS software and R GIS libraries to realize this information representation using a vector type information layer (division into distinct entities which is opposed to the raster type which represents space in a regular mesh). Technically, the proposed approach consists of 1) using a GIS (ArcGIS) to vectorize an image representing the organism, then 2) producing an illustration from the vector layer created according to the desired division. More specifically, the image chosen to represent the organism under study is vectorised into polygon-type objects corresponding to compartments of interest of the organism. Finally, the representation of the intensity of the phenomenon according to the different compartments is accomplished using two softwares, ArcGIS and the GIS libraries of the free statistical software R. This second one has the advantage, once the shape file is created, of being able to be implemented in a script for R data analysis, a software widely used in the scientific community. This second approach allowed, in the example presented in this article, to represent the distribution of nitrogen concentrations or carbohydrate reserves in the different compartments of a beech tree at different dates and subjected to contrasting treatments.

Keywords: picture, GIS, vectorisation, R, programming, script

Introduction

Dans le cadre du projet MEPiB-Death (Morphological and Ecophysiological Processes leading to Beech tree Death, financé par le Labex ARBRE), les chercheurs voulaient comprendre, chez le hêtre, les mécanismes physiologiques et les modifications morphologiques induits par de fortes contraintes (défoliation ou sécheresse). Mille hêtres de 10 ans plantés en pépinière ont été placés sous un toit afin d'exclure les pluies. Les chercheurs ont analysé l'impact de ces contraintes sur les métabolismes carbonés et azotés des arbres. Après une phase de quantification des composés azotés et des réserves glucidiques dans les différents compartiments des arbres, les chercheurs avaient besoin de représenter au sein de la structure d'un arbre les composés qu'ils étudiaient à différentes dates.

Cet article propose de représenter les compartiments de la structure aérienne et souterraine d'un arbre à l'aide d'un SIG car dans notre cas celui-ci permet grâce à ses fonctionnalités de gérer, stocker et restituer l'information des données. Le choix d'utiliser une couche vecteur SIG, découpage en objets distincts dans un espace discret, de type polygone a été favorisé à une couche raster qui représente l'espace dans un maillage régulier continu. Une couche SIG est agrémentée d'une table attributaire, celle-ci contient les données non spatiales des objets ainsi créés. Le travail en couche vectorielle est plus simple à mettre en œuvre et il permet grâce à la symbologie (apparence visuelle de la couche vecteur) de délimiter précisément les compartiments étudiés. Cet article propose ensuite deux possibilités pour définir la symbologie 1) de continuer avec un système d'information géographique (SIG), ou 2) d'utiliser les données de type vectoriel avec les bibliothèques SIG du logiciel R pour représenter la distribution au sein de l'arbre des composés carbonés et azotés.

Le Cahier des Techniques de l'Inra 2021 (104)

Méthodologie

Sélectionner une image

La première étape de la méthode consiste à choisir une image de l'organisme à représenter. Une image d'arbre ressemblant à un hêtre et libre d'accès a été retenue. Cette image est contenue dans la base d'images du site Dreamstime (<https://fr.dreamstime.com>) qui dispose d'un panel important. L'image sélectionnée est présentée dans la figure 1.

Cette image d'arbre présente ainsi tous les compartiments étudiés de l'arbre (racines, tronc, branches et feuilles) dont avaient besoin les chercheurs pour la spatialisation des processus étudiés.



Figure 1. Image initiale de l'arbre à vectoriser issue du site Dreamstime

La vectorisation sous SIG

L'étape de vectorisation de ce travail est une phase qui consiste à transformer une information contenue dans une image en vecteur (ici de type polygone : l'information est représentée par un dessin et elle est contenue dans une table attributaire). La vectorisation a été produite à partir du logiciel ArcGIS (<https://my.esri.com>, version 10.3.1) disponible au sein de notre institut de recherche. Pour disposer d'ArcGIS, contactez votre référent SIG de centre. Il est aussi possible d'utiliser QGIS qui est un logiciel libre.

Une phase préliminaire à la vectorisation consiste en la création sous ArcCatalog d'un fichier de formes (nommé ici arbre) au format vecteur shapefile (.shp) de type Polygone. Ce type d'objet est le seul adapté pour la représentation graphique de notre exemple car il correspond à des surfaces closes. Ainsi l'arbre pourra être compartimenté en polygones de différentes tailles et formes adaptées pour représenter les différents organes de l'arbre (système racinaire, tronc, branches et feuilles).

Il n'est pas nécessaire de définir de système de coordonnées car il ne s'agit pas d'une entité géographique réelle projetée dans un référentiel cartographique défini. L'utilisation des coordonnées X et Y sans projection est suffisante pour l'utilisation du fichier de formes par la suite. Dans notre cas, le système d'informations géographiques est employé pour réaliser un dessin assisté par ordinateur et non pour recueillir, traiter, analyser et gérer des données géographiques.

Pour réaliser une vectorisation interactive dans ArcGIS, il faut réaliser un clic-droit dans la barre d'outils et sélectionner l'onglet « Editeur » (étape 1, figure 2). Dans le menu déroulant « Editeur » (étape 2, figure 2) « Ouvrir une session de mise à jour » ce qui permet d'ouvrir la fenêtre « Créer des entités » et d'utiliser l'outil de construction « Polygone » (étape 3, figure 2). Ensuite le SIG a été utilisé comme un Dessin Assisté par Ordinateur (DAO) et l'image a été vectorisée selon les besoins cités précédemment.

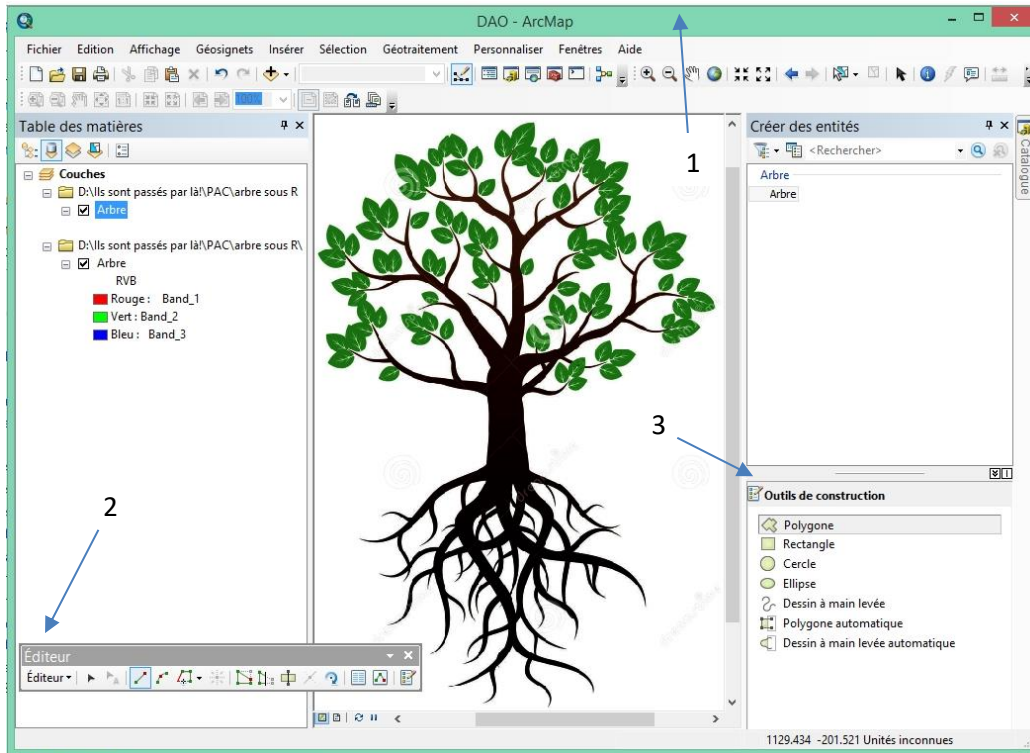


Figure 2. Description des 3 étapes employées sous ArcGIS pour la vectorisation de l'image.

La dernière phase de la vectorisation sous SIG consiste à renseigner les polygones avec les informations sur les compartiments ainsi créés en fonction de l'organe considéré (figure 3). Pour cela un champ doit être créé et ajouté dans la table attributaire du fichier de forme « arbre ». Il est nécessaire de « Quitter la session de mise à jour » dans « Editeur » pour réaliser cette étape (étape 1, figure 3). Ensuite clic-droit sur le fichier de formes et « Ouvrir la table attributaire » (étape 2, figure 3). Puis dans « Options de la table », il suffit d'« Ajouter un champ ... » (étape 3, figure 3) au format texte de la longueur choisie (ici Loc pour localisation). Pour la suite de la démarche, il faut cocher la case « Afficher les enregistrements sélectionnés » (étape 4, figure 3). Après avoir sélectionné un vecteur à nommer, clic-droit sur le nouveau champ (étape 5, figure 3) et choisir « Calculateur de champs ». Ceci permet d'ouvrir une fenêtre (étape 6, figure 3) et de renseigner le nom du vecteur dans le champ (étape 7, figure 3). Il est impératif d'utiliser les guillemets car il s'agit d'une valeur alphanumérique ("Trunk 2010", ici en anglais pour les besoins du projet, "Tronc 2010" en français). Pour les besoins de cette étude, la vectorisation de l'image a été réalisée en fonction du nombre d'organes désirés pour la représentation des processus. Ainsi la division de l'arbre a été faite selon les éléments suivants : i) racines que l'on voulait compartimenter en 3 catégories : racines fines, racines latérales et racines principales, ii) tronc que l'on voulait compartimenter par unité de croissance annuelle, iii) les branches que l'on voulait compartimenter en 3 catégories d'unités de croissance : croissance de l'année n (dernière année de croissance), croissance de l'année n-1 et croissance des années < n-1 , iv) feuilles.

Le Cahier des Techniques de l'Inra 2021 (104)

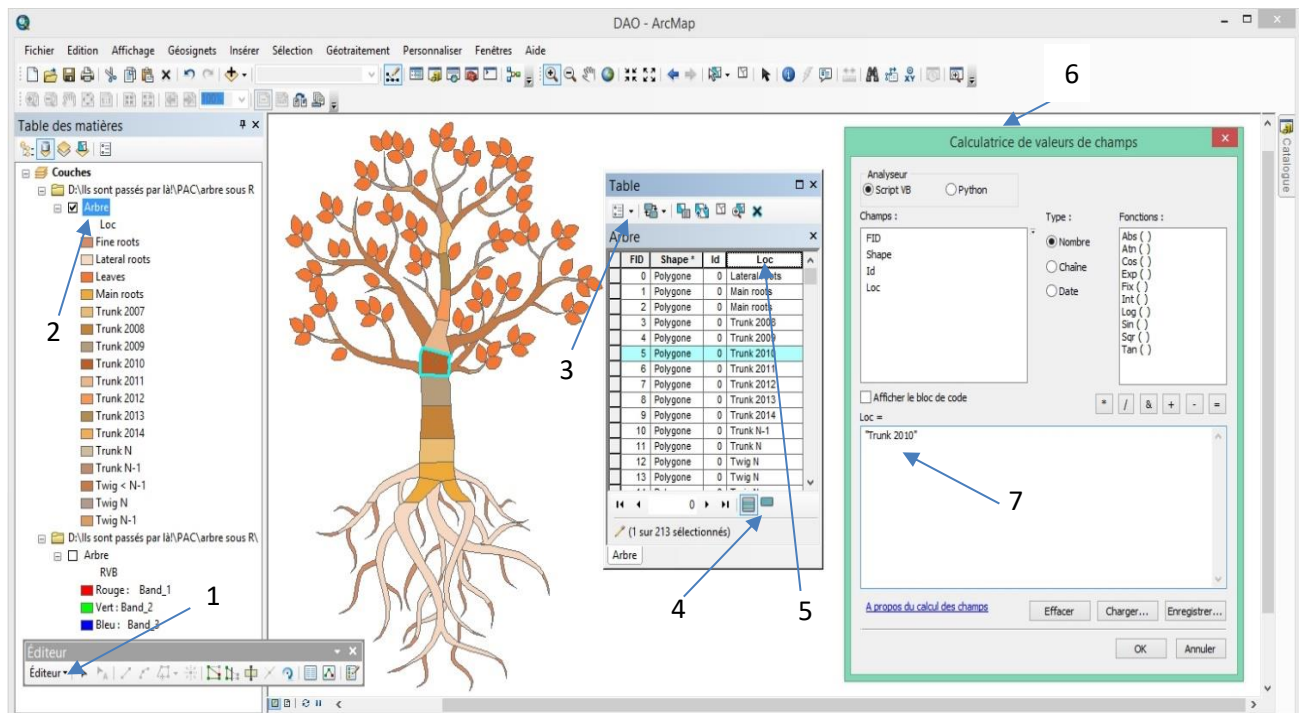


Figure 3. Phase de création de la table attributaire en 7 étapes. Cette phase permet de nommer chaque compartiment correspondant à une topologie dans l'arbre. Les différents compartiments créés pour les besoins de l'étude sont : Fine roots (racines fines), Lateral roots (racines latérales), Main roots (racines principales), Trunk 2007 (tronc 2007), Trunk 2008 (tronc 2008), Trunk 2009 (tronc 2009), Trunk 2010 (tronc 2010), Trunk 2011 (tronc 2011), Trunk 2012 (tronc 2012), Trunk 2013 (tronc 2013), Trunk 2014 (tronc 2014), Trunk n-1 (tronc année N-1), Trunk N (tronc année N), Twigs < N-1 (branches années < N-1), Twigs N-1 (branches année N-1), Twigs N (branches année N), Leaves (feuilles).

Visualisation sous SIG ArcGIS

Pour illustrer cette étape, nous utiliserons un tableur Excel (nommé TableauDeComposes) dans lequel seront renseignées les informations de compartimentations (racines, tronc, feuilles) et de valeurs aléatoires (illustrant les teneurs en azote ou les réserves glucidiques).

Pour permettre une représentation de ce nouveau jeu de données, une jointure tabulaire entre ce tableau (TableauDeComposes) et le fichier de formes (arbre) est réalisé selon la procédure suivante : clic-droit sur le fichier de forme, sélectionnez « Jointures et relations » puis « Joindre ... ». Cette jointure se décompose en trois temps 1) « Choisissez dans la couche le champ sur lequel portera la jointure » (loc pour localisation), 2) « Choisissez la table à joindre à cette couche » (TableauDeComposes) et 3) « Choisissez dans la table le champ sur lequel portera la jointure » (loc pour localisation). La table attributaire de notre fichier de forme contient maintenant les valeurs à illustrer. La mise en forme des données s'exécute en double cliquant sur le fichier de formes (étape 1, figure 4) dans l'onglet « Symbologie ». Sélectionnez « Quantité – Couleurs graduées » (étape 2, figure 4) et le champ sur lequel s'effectue la symbologie (étape 3 figure 4). Enfin modulez le nombre de classe désirée (étape 4 figure 4).

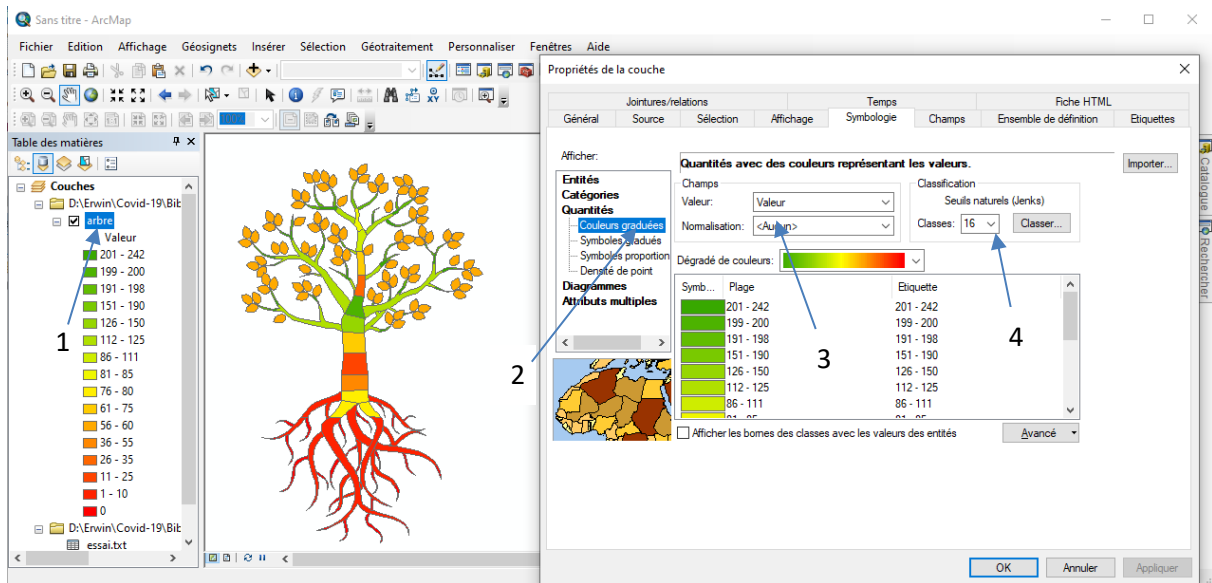


Figure 4. Phase de mise en pages du fichier de formes.

Visualisation sous SIG R

L'ensemble de la communauté scientifique n'utilise pas les logiciels dédiés aux systèmes d'informations géographiques, il est possible pour un scientifique analysant des données d'utiliser le script suivant dans le logiciel R avec les librairies SIG, une fois le fichier de forme récupéré auprès d'un tiers.

L'écriture de ce paragraphe est faite sous le format d'un script R (<http://www.r-project.org>, version 3.5.1, 2018-07-02) permettant d'utiliser le fichier de forme (ici arbre) avec un jeu de couleurs. Il est libre d'adaptation pour toute autre utilisation. En programmation R, tout ce qui est écrit après le # n'est pas pris en compte dans le programme et permet l'insertion de commentaires sur le fonctionnement du code.

```
#####arbre découpé en tronçons#####
#####avec couleurs selon concentration#####
#####
```

```
#derrière chaque # il s'agit d'explications pour la compréhension du script
```

```
#installation du package nécessaire à la lecture de données SIG dans R
install.packages("rgdal")
#ouverture d'un second package pour la jointure de deux dataframes
install.packages("sp")
#ouverture des packages
library(rgdal)
library(sp)
```

```
##définition du dossier de travail
setwd("D:/arbre sous R")
##placer le fichier de forme "arbre" issu du SIG dans ce dossier
```

```
#ouverture du fichier de forme arbre (fichier de type « .shp ») toujours nommé « arbre » dans R
arbre<-readOGR("arbre.shp",layer="arbre")
# Ici arbre est un objet de type 'SpatialPolygonsDataFrame' du package sp et contient l'information issue du fichier de forme
```

```
#visualisation brute
```

Le Cahier des Techniques de l'Inra 2021 (104)

plot(arbre)

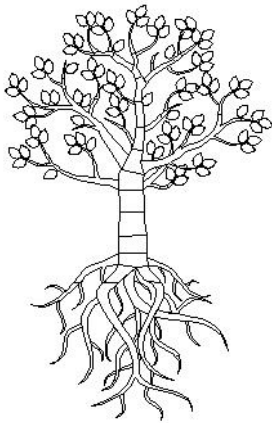


Figure 5. Représentation sous R de l'arbre vectorisé en compartiments sous ArcGIS

```
#lecture du fichier TableauDeComposes
```

```
TableauDeComposes=read.table("TableauDeComposes.txt",header=TRUE,sep="\t")
```

```
#TableauDeComposés = tableau de données avec nos valeurs à représenter (qui est ici créé ou qui peut être importé)
```

```
#visualisation des premières lignes du tableau
```

```
head(TableauDeComposes)
```

	Loc	valeur
1	Fine roots	0
2	Lateral roots	10
3	Main roots	80
4	Leaves	60
5	Trunk 2007	55
6	Trunk 2008	25

```
#jointure du tableau de données (« TableauDeComposes ») avec la table attributaire du fichier de forme (« arbre@data»), cette opération est implémentée dans la fonction merge du package sp  
arbre = sp::merge(x=arbre, y = TableauDeComposes, by="Loc")
```

```
#les données de composés sont ainsi jointes au fichier de forme
```

```
#la colonne Loc est utilisée comme clé. Elle est identique aux deux fichiers à fusionner
```

```
#visualisation du début du tableau fusionné
```

```
head(arbre@data)
```

	Id	Loc	valeur
0	0	Lateral roots	10
1	0	Main roots	80
2	0	Main roots	80
3	0	Trunk 2008	25
4	0	Trunk 2009	75
5	0	Trunk 2010	150

```
##création d'une palette continue de couleurs afin de coloriser les différents compartiments selon leur valeur
```


#cette commande permet de créer une fonction implémentant un dégradé de couleur en interpolant dans l'espace RGB les couleurs fournies en argument. Cette fonction myPal prend en argument un entier correspondant au nombre de couleurs à retourner

```
myPal <- colorRampPalette( c("red","orange","khaki","green4","darkgreen" ) )
```

```
#Ici, on retient un dégradé transitionnel selon 50 couleurs
#50 est un nombre de couleurs choisi et modulable dans le dégradé de couleur
col=myPal(50)
```

```
#Définition des bornes des classes pour le découpage des valeurs du composé à représenter
# On choisit ici un découpage avec un intervalle régulier. Le nombre de classe souhaité correspond au nombre de
classe de couleurs défini précédemment (ici 50)
my.at <- seq(0, (max(arbre@data$Valeur)+1),
             by = (max(arbre@data$Valeur)+1)/50)
```

```
# #visualisation schématique du résultat
spplot(arbre,"Valeur",
       at=my.at,
       col.regions=col,
       main="Arbre sous R")
```

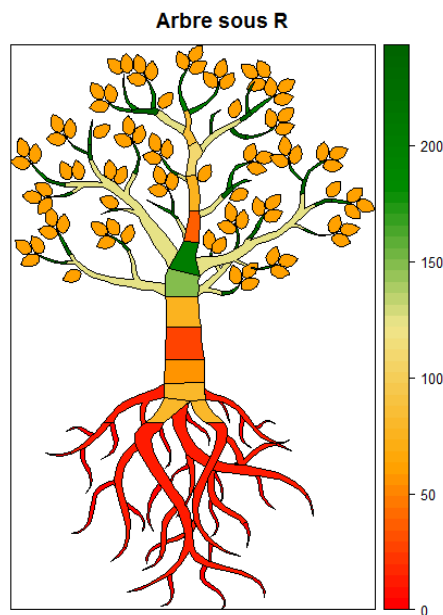


Figure 6. Arbre colorisé en fonction de la concentration d'un composé dans les différents compartiments. La légende à droite représente le gradient de couleurs en fonction des teneurs du composé.

Une démarche générique à partager

Le logiciel de programmation et de statistiques R est très utilisé par la communauté scientifique. Les demandes techniques d'éléments visuels à but illustratif sont souvent très utiles et délicates à mettre en œuvre. L'utilisation d'un système d'information géographique (et du logiciel R si nécessaire) a permis de créer une illustration spatialisée d'un composé (figure 6). Ce travail original a été utilisé dans un poster primé lors d'un colloque international (Chuste et al, 2017), dans la présentation des résultats d'une thèse (Chuste, 2018) et dans la publication d'un article scientifique (Chuste et al, 2019, figure 7, mise en forme sous PowerPoint).

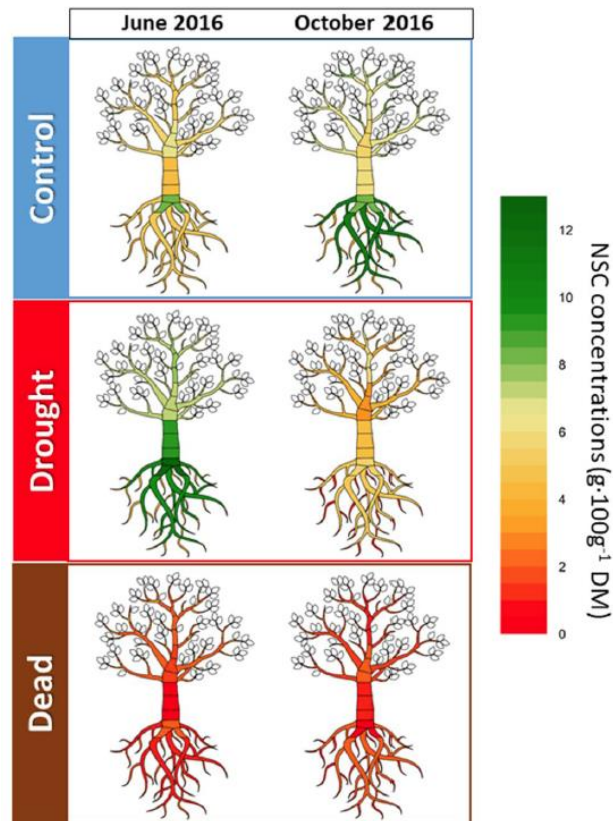


Figure 7. Schéma de la répartition des glucides non structuraux (NSC) en g pour 100 g de matière sèche (Chuste & al, 2019).

Cette méthodologie n'est pas cantonnée à la création d'un arbre. Sa genericité permet de l'utiliser dans de nombreuses autres applications (illustrations en biologie animale et végétale, plan non géométrique, archéologie, ...) dès que le besoin de support visuel apporte une plus-value au message diffusé et que sa création est compliquée. Vous pouvez contacter l'auteur (erwin.thirion@inrae.fr) pour partager le script. Vous pourrez ainsi utiliser et modifier ce protocole selon les besoins de votre étude.

Remerciements

Tous mes remerciements à Catherine MASSONNET et Nathalie BREDA pour leurs lectures et corrections de cet article. Je remercie également Nicolas SABY pour les relectures et les améliorations qu'il m'a proposées. Enfin je remercie Pierre-Antoine CHUSTE de m'avoir fait confiance pour la création de cet arbre R. Ce travail a été financé par l'Agence Nationale de la Recherche (Laboratoire d'Excellence ARBRE) et par la région Lorraine (projet Survival).

Cet article est publié sous la licence Creative Commons (CC BY-SA).



<https://creativecommons.org/licenses/by-sa/4.0/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Le Cahier des Techniques de l'Inra », la date de sa publication et son URL).

Bibliographie

Chuste, P-A, Massonnet, C, Zeller, B, Bréda N, Tillard P, Wortemann R, Thirion E, Maillard, P (2017) Whole tree nitrogen dynamics across seasons in response to defoliation and drought in 10 year-old beech trees. International Plant Nutrition Colloquium 21-24 August 2017, Copenhagen, Denmark.

Chuste, P-A (2018) Etude de la sensibilité du hêtre lorrain à un événement climatique extrême. Quels sont les rôles des métabolismes carboné et azoté dans la mort des arbres? Thèse de Biologie Végétale Forestière, Université de Lorraine, 238 p.

Chuste P-A, Maillard P, Bréda N, Levillain J, Thirion E, Wortemann R, Massonnet C (2019) Sacrificing growth and maintaining a dynamic carbohydrate storage are key processes for promoting beech survival under prolonged drought conditions. *Trees* 1-14. doi:10.1007/s00468-019-01923-5