

Suivi pendant 56 ans de l'effort de reproduction chez le chêne dans 17 régions polonaises

Pr Jean R. LOBRY

Ces séries temporelles d'une longueur exceptionnelle dans le domaine donnent une illustration très didactique du phénomène dit de mast-ing. Ces données mettent en évidence la synchronisation de l'effort de reproduction chez le chêne (*Quercus spp.*) à des distances de l'ordre de 100 km.

1 Les données

1.1 Import et sélection

LES données sont disponibles dans MASTREE+ [3], le compilateur responsable de l'extraction et de l'incorporation de ces données était Andrew HACKET-PAIN. On pourra en importer une version¹ directement dans le tableau de données `mastree` de  de la façon suivante :

```
chmin <- "http://pbil.univ-lyon1.fr/R/donnees/mastree.Rda"
load(url(chmin))
```

LE sous-ensemble des données qui nous intéresse porte la clef d'identification D00013 dans la colonne `Study_ID` et correspond à 56 années de suivi du chêne *Quercus spp.*, plus précisément à un mélange de *Q. robur* et de *Q. petraea*. Pour une description des étonnantes relations entre ces deux espèces on pourra se délecter de la lecture des travaux d'Antoine KREMER et collaborateurs [5, 6].

```
Kanto <- subset(mastree, Study_ID == "D00013" & Length == 56)
nrow(Kanto) # 17*56 = 952
[1] 952
unique(Kanto$Species)
[1] "Quercus spp."
unique(Kanto$Site)
[1] "Bialystok"    "Gdansk"      "Katowice"    "Krakow"      "Krosno"
[6] "Lublin"      "Lodz"        "Olsztyn"     "Pila"        "Poznan"
[11] "Radom"       "Szczecin"   "Szczecinek" "Torun"       "Warszawa"
[16] "Wroclaw"     "Zielona Gora"
```

¹Pour plus de détails voir <http://pbil.univ-lyon1.fr/R/pdf/MASTREE.pdf>

Fructification	%	Description
Nulle	0	Misère : aucun arbre ne donne des glands valant la peine d'être collectés
Faible	10	Survie : seuls les arbres en lisière donnent des glands
Moyenne	30	Glandée d'oiseaux (calque de l'allemand <i>vogelmast</i> [7]) : beaucoup d'arbres en lisière et quelques arbres au centre du massif donnent des glands
Bonne	100	Glandée : les arbres du massif donnent des glands à profusion

TABLE 1 : Quantification de l'effort de reproduction annuel sur un site d'après la table 1 de [4]. Les valeurs sont ensuite agrégées au niveau de chaque région (RDLP de la figure 1 page 3) en calculant le pourcentage moyen.

1.2 Description des données

LA nature des données a été décrite par Władysław KANTOROWICZ [4]. Ce sont des données populationnelles qui ont été agrégées au niveau des régions forestières polonaises (voir la figure 1 page 3). Chaque région comporte en moyennes 26 sites d'observation, et pour chaque site on note avec un pourcentage (voir la table 1 page 2) l'intensité de l'effort de reproduction. Il s'agit plus d'une variable qualitative ordonnée à 4 modalités, mais son codage sous la forme d'un pourcentage permet de calculer facilement pour chaque année l'intensité moyenne de l'effort de reproduction de chaque région.

2 Représentation des séries temporelles

POUR une première représentation directe des données, on s'inspire d'un graphique déjà publié [9] sur une série analogue (voir la figure 2 page 4). On représente les séries temporelles des 17 régions en gris et on met en évidence la moyenne en noir. Comme point de repère on indique en bleu les seuils de la table 1 page 2.

```
Sites <- unique(Kanto$Site) ; ns <- length(Sites)
par(mar = c(5, 5, 1, 1) + 0.1) ; deb <- 1945
plot.new() ; plot.window(xlim = c(deb, max(Kanto$Year)), ylim = c(0, 100))
axis(1, at = seq(deb, 2005, by = 5), cex = 0.75) ; axis(2, las = 1)
rug(1950:2005, col = grey(0.95), ticksize = 1)
title(xlab = "Année", ylab = "Effort reproducteur [%]")
for(i in seq_len(ns)){
  the_site <- Sites[i]
  tmp <- subset(Kanto, Site == the_site)
  points(tmp$Year, tmp$Value, col = rgb(0.5, 0.5, 0.5, 0.5), type = "l", )
}
Years <- sort(unique(Kanto$Year)) ; ny <- length(Years)
moy <- numeric(ny)
for(i in seq_len(ny)){
  moy[i] <- mean(Kanto[Kanto$Year == Years[i], "Value"])
```



FIGURE 1 : Copie d'écran de la figure 1 de [4]. Les 17 régions *Regionalna Dyrekcja Lasów Państwowych*, RDLP, sont l'équivalent des régions forestières de l'ONF. Leurs frontières ont varié au cours du temps mais les données ont été ré-assignées pour respecter la configuration représentée ici. C'est à l'échelle de ces régions que les données ont été agrégées, sachant qu'en moyenne chaque RDLP comporte 26 sites. La surface d'une RDLP est de l'ordre de celle d'un carré de 100 km de côté.

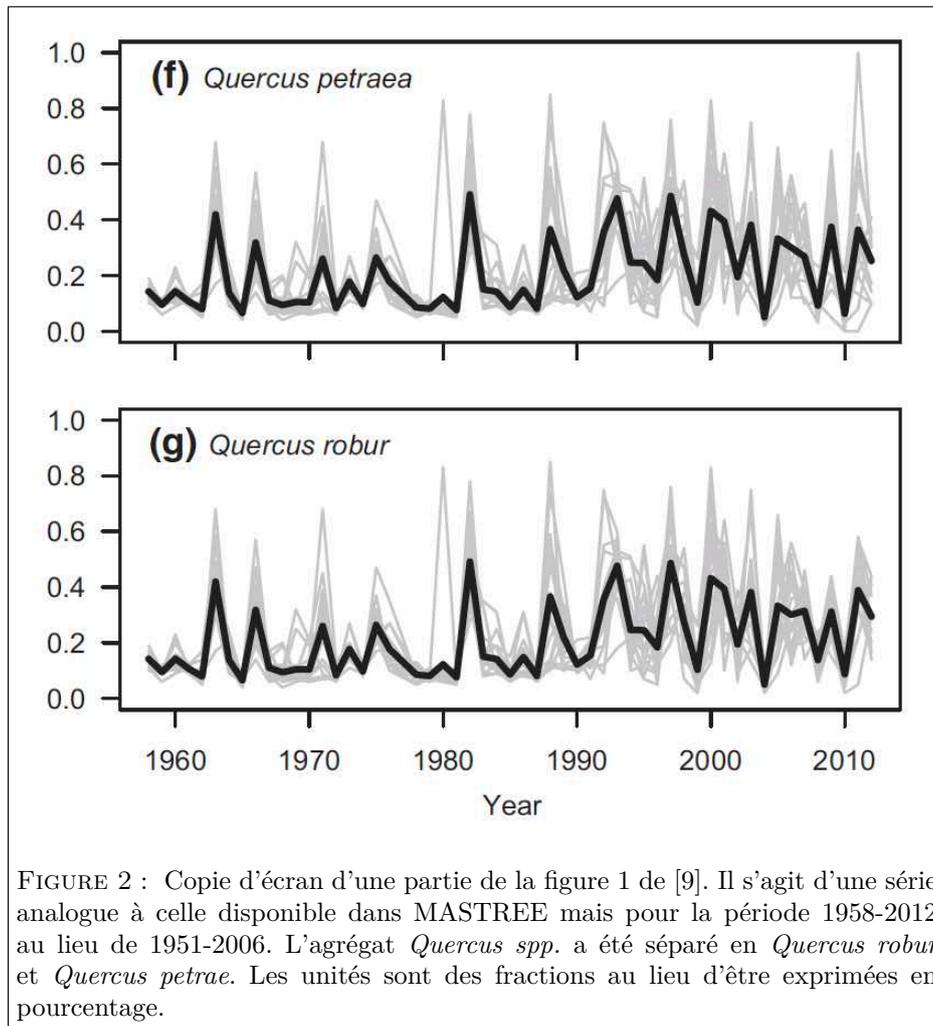
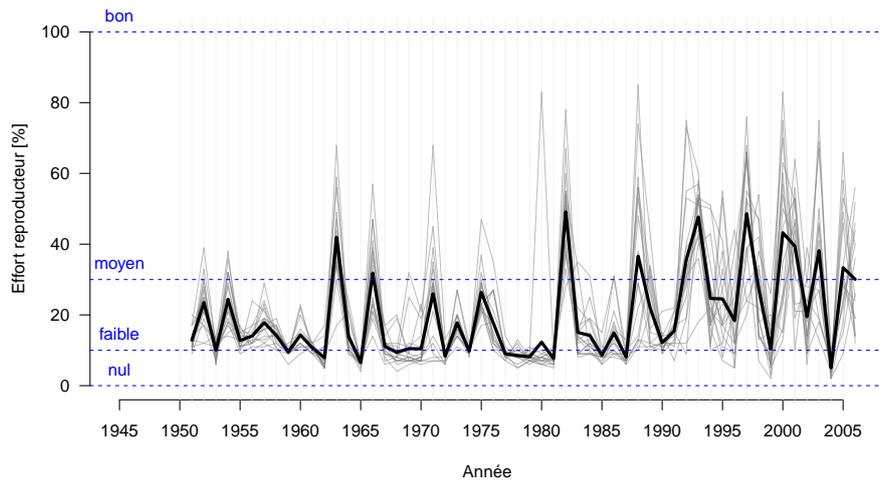


FIGURE 2 : Copie d'écran d'une partie de la figure 1 de [9]. Il s'agit d'une série analogue à celle disponible dans MASTREE mais pour la période 1958-2012 au lieu de 1951-2006. L'agrégat *Quercus spp.* a été séparé en *Quercus robur* et *Quercus petraea*. Les unités sont des fractions au lieu d'être exprimées en pourcentage.

```

}
points(Years, moy, lwd = 3, type = "l")
abline(h = c(0, 10, 30, 100), lty = 2, col = "blue")
text(deb, 0, "nul", col = "blue", pos = 3)
text(deb, 10, "faible", col = "blue", pos = 3)
text(deb, 30, "moyen", col = "blue", pos = 3)
text(deb, 100, "bon", col = "blue", pos = 3, xpd = NA)

```



ON retrouve un profil moyen très similaire, et c'est normal, à celui de la figure 2 page 4. Pour les profils par région en gris on retrouve les mêmes singularités, par exemple la région qui fait bande à part en 1980 avec un très gros effort de reproduction alors que les autres sont à leur étiage. On note qu'il n'y a pas de valeurs nulles, ce qui signifie qu'au niveau régional il y a toujours une production résiduelle de glands, même si elle est très faible, la ligne basale étant de l'ordre de 10 %. Du côté des fortes glandées, on note des pointes à 80 %, mais la valeur de 100 % n'est jamais atteinte. On ne peut pas interpréter ceci comme un défaut de synchronisme car on ne connaît pas la proportion de sites n'ayant pas encore atteint la maturité de reproduction. On note à partir de 1990 un changement de rythme avec une augmentation de l'effort de reproduction. Avant de foncer tête baissée pour y voir l'impact du réchauffement global on gardera à l'esprit la réflexion de Władysław KANTOROWICZ [4] : « [t]his could possibly be related to industrial pollution, which kept on increasing between 1960 and 1989, to drop drastically after the change of our political system, first as a consequence of drop in industrial output and later due to the introduction of new more clean technologies. »

POUR faciliter le repérage des régions, on peut proposer une représentation où les courbes sont translatées les unes par rapport aux autres d'un facteur arbitraire sur l'axe des ordonnées. On voit ainsi que c'est la région de SZCZECINEK qui se singularise en 1980.

```

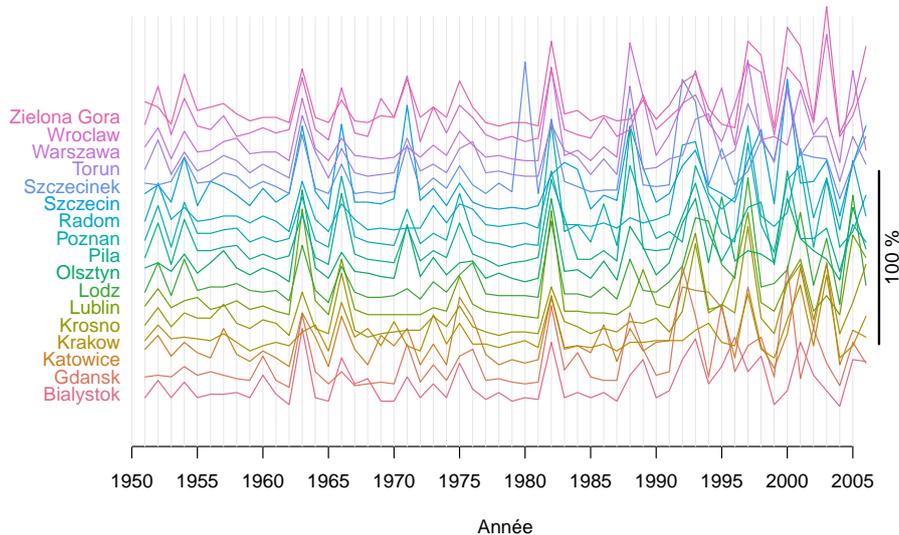
par(mar = c(5, 5, 1, 1) + 0.1)
col <- hcl.colors(ns, "Dark 3")
plot.new() ; plot.window(xlim = range(Kanto$Year), ylim = c(0, 230))
axis(1, at = seq(1950, 2005, by = 5))
rug(1950:2005, col = grey(0.9), ticksize = 1)
title(xlab = "Année")
for(i in seq_len(ns)){
  the_site <- Sites[i]
  tmp <- subset(Kanto, Site == the_site)

```

```

points(tmp$Year, tmp$Value + 10*i, col = col[i], type = "l", xpd = NA)
text(1950, 10*i + 10, the_site, col = col[i], pos = 2, xpd = NA)
}
segments(2007, 50, 2007, 150, xpd = NA, lwd = 2)
text(2008, 100, "100 %", srt = 90, xpd = NA)

```



3 Cartographie

DANS la base de données MASTREE+ on dispose des coordonnées géographiques des régions (longitude et latitude décimales) sur lesquelles portent les observations :

```

régions <- Kanto[!duplicated(Kanto$Site), c("Site", "Longitude", "Latitude")]
rownames(régions) <- NULL
head(régions)

```

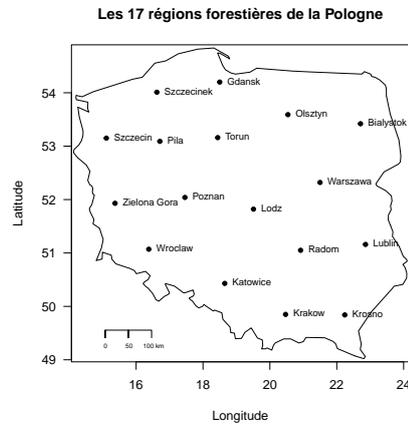
	Site	Longitude	Latitude
1	Bialystok	22.71	53.42
2	Gdansk	18.50	54.20
3	Katowice	18.65	50.43
4	Krakow	20.47	49.85
5	Krosno	22.24	49.84
6	Lublin	22.86	51.16

POUR faire une représentation graphique simple des données je recommande dans le cas présent d'utiliser le paquet **R maps** [2] qui permet de disposer facilement du fond de carte d'un pays donné. On pourra vérifier que la représentation ci-après est cohérente avec la figure 1 page 3.

```

library(maps)
map("world", regions = "Poland", mar = c(5, 4, 4, 2) + 0.1)
map.axes(las = 1)
map.scale(ratio = FALSE, cex = 0.5, relwidth = 0.1)
title(xlab = "Longitude", ylab = "Latitude",
      main = "Les 17 régions forestières de la Pologne")
with(régions, {
  points(Longitude, Latitude, pch = 19, cex = 0.75)
  text(Longitude, Latitude, Site, pos = 4, cex = 0.75)
})

```



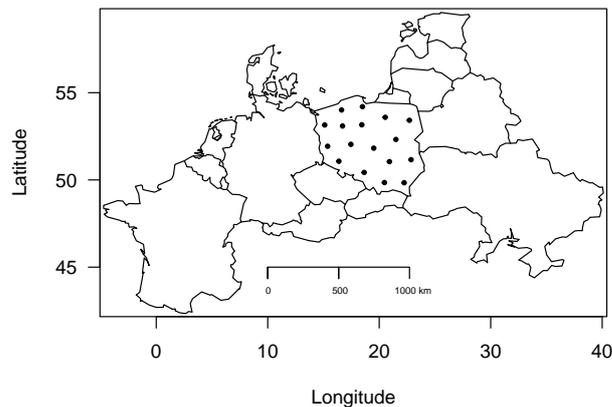
La fonction `map()` utilise par défaut une projection rectangulaire telle que les échelles en abscisse et ordonnée soient identiques au centre de la carte. Elle est compatible avec les paramètres graphiques de base de `R`, comme on peut le voir ici avec `mar = c(5, 4, 4, 2) + 0.1` pour retrouver les marges habituelles². La fonction `map.scale()` permet de préciser l'échelle. On peut facilement représenter un pays dans un contexte plus général³.

```
target <- c("Poland", "Germany", "France(!:Corsica)", "Belgium",
            "Netherland", "Ukraine", "Belarus", "Czech Republic",
            "Slovakia", "Russia:32", "Austria", "Lithuania", "Latvia",
            "Estonia", "Denmark")
map("world", regions = target, mar = c(5, 8, 4, 0))
map.axes(las = 1)
map.scale(ratio = FALSE, cex = 0.5, relwidth = 0.2, x = 10, y = 45)
title(xlab = "Longitude", ylab = "Latitude",
      main = "Les 17 régions forestières de la Pologne")
with(régions, {
  points(Longitude, Latitude, pch = 19, cex = 0.5)
})
```

²Pour une exploration systématique des paramètres graphiques de base dans `R` voir la fiche <http://pbil.univ-lyon1.fr/R/pdf/tdr75.pdf>

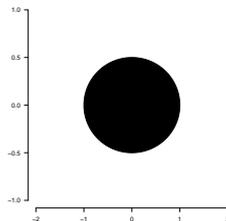
³On peut facilement avoir le nom des pays avec la fonction `map.text()` ou bien en exploitant la valeur de retour de `map()` avec l'argument `namesonly = TRUE`. Les « pays » dans `maps` sont dégraissés de leurs territoires ultra-périphériques ce qui évite bien des surprises lors de la projection des entités surfaciques qui nous intéressent. On peut utiliser des notations telles que `France(!:Corsica)` pour exclure de représenter la Corse dans la France métropolitaine ou bien `Russia:32` pour inclure l'exclave russe de Kaliningrad. Dans la version 3.4.1 de `maps` les données géopolitiques sont relativement récentes puisqu'elles ont été importées en 2013 de <https://www.naturalearthdata.com>. Elles sont instables à l'aune du temps de reproduction moyen d'un chêne mais suffisantes à l'échelle des mortels qui s'y intéressent.

Les 17 régions forestières de la Pologne



POUR représenter l'intensité de l'effort de reproduction je vais utiliser des cercles dont la surface est proportionnelle à la valeur observée. Je me suis inspiré de la fonction `circle()` du paquet `seqinr` [1] que j'ai simplifiée et complétée avec le paramètre `xyr` pour pouvoir faire un cercle qui ait l'aspect d'un rond même dans un repère non orthonormé.

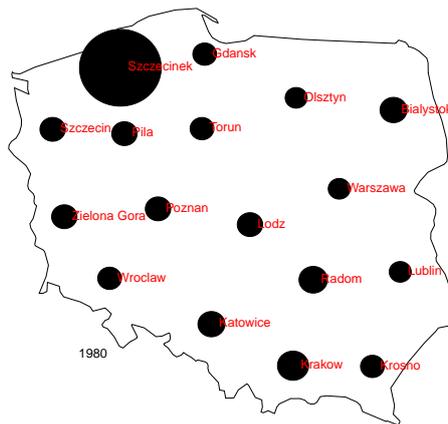
```
circle <- function (x = 0, y = 0, r = 1, xyx = 1, n = 100, ...)
{
  a <- seq(0, 360, length = n)
  xx <- x + r * cos(a * 2 * pi/360)
  yy <- y + xyx*(r * sin(a * 2 * pi/360))
  polygon(xx, yy, ...)
}
par(mar = c(3, 3, 0, 0))
plot.new() ; plot.window(xlim = c(-2, 2), ylim = c(-1, 1))
axis(1) ; axis(2, las = 1)
circle(xyx = 1/2, col = "black")
```



ON peut maintenant définir la fonction utilitaire `plotyear()` pour représenter les données d'une année. Le paramètre d'échelle `scale` doit être ajusté empiriquement pour avoir des cercles suffisamment gros tout en restant non chevauchants. Dans le cas présent une valeur de 0.1 semble un bon compromis.

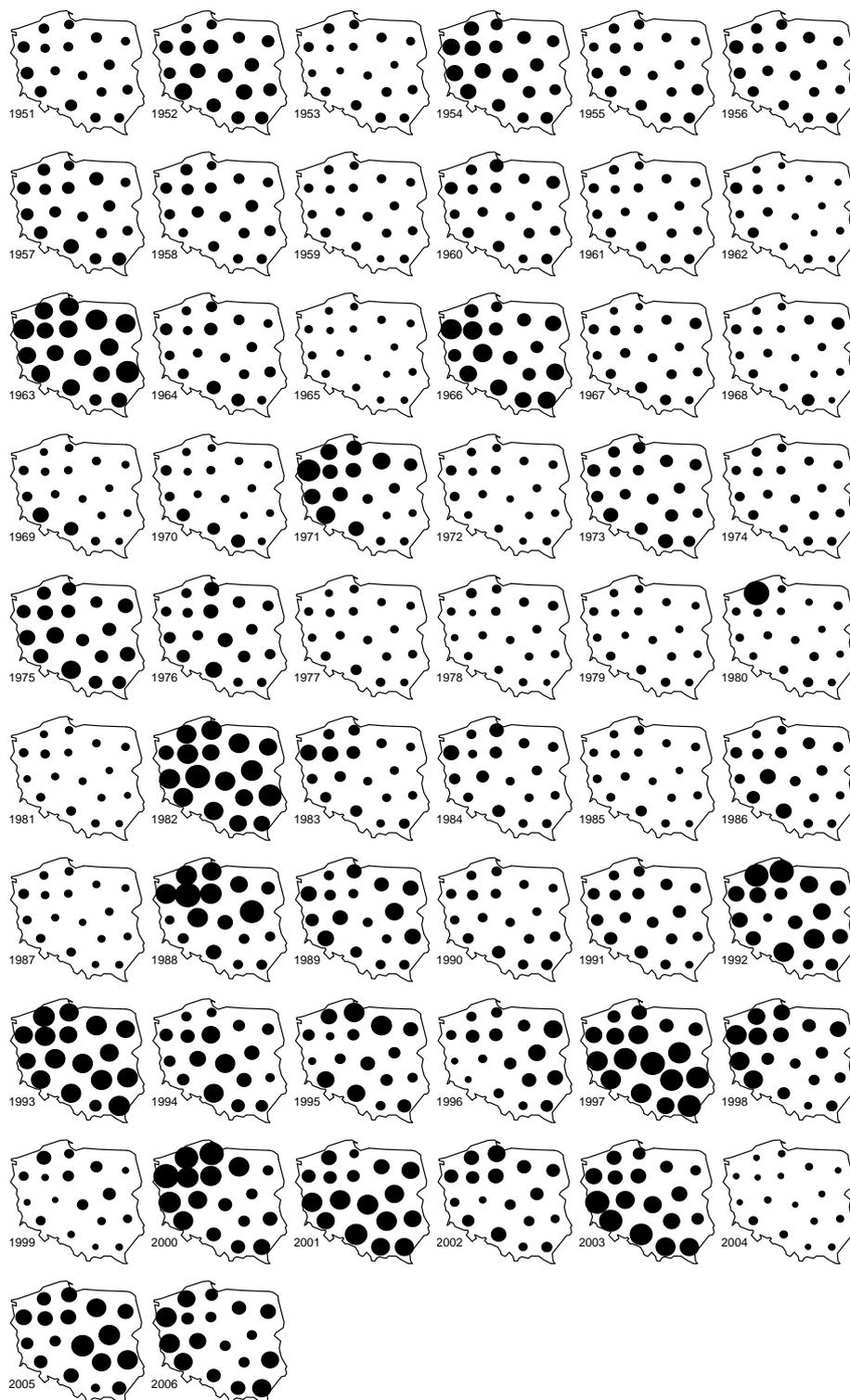
On teste la fonction avec l'année 1980 pour vérifier que c'est bien la région de SZCZECINEK qui se singularise.

```
plotyear <- function(the_year, scale = 1){
  require(maps)
  map('world', regions = "Poland", mar = c(0.5, 0.5, 0, 0))
  target <- c("Site", "Longitude", "Latitude", "Value")
  tmp <- Kanto[Kanto$Year == the_year, target]
  points(tmp$Longitude, tmp$Latitude, pch = 19)
  pu <- par("usr")
  deltax <- pu[2] - pu[1] ; deltay <- pu[4] - pu[3]
  xyr <- deltax/deltay
  for(i in seq_len(nrow(tmp))){
    x <- tmp[i, "Longitude"]
    y <- tmp[i, "Latitude"]
    r <- scale*sqrt(tmp[i, "Value"])
    circle(x, y, r, xyr, col = "black")
  }
  text(16.5, 50, the_year, pos = 2)
}
plotyear(1980, scale = 0.1)
with(régions, text(Longitude, Latitude, Site, pos = 4, col = "red"))
```



IL ne reste plus qu'à faire une boucle sur les 56 années disponibles pour avoir une belle illustration du phénomène dit de *masting* chez le chêne à l'échelle de la Pologne.

```
par(mfrow = c(10, 6))
for(i in 1951:2006) plotyear(i, scale = 0.1)
```



NOTEZ que grâce au paquet `gifski` [8] il est facile de produire une image animée.

```
for(i in 1951:2006){
  png(paste0("image", i, ".png"), width = 150, height = 150)
  plotyear(i, scale = 0.1)
  dev.off()
}
library(gifski)
png_files <- list.files(path = ".", pattern = ".png", full.names = TRUE)
gifski::gifski(png_files, gif_file = "KantorowiczW2000.gif",
               delay = 0.2)
```

Références

- [1] D. Charif and J.R. Lobry. SeqinR 1.0-2 : a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In H.E. Roman U. Bastolla, M. Porto and M. Vendruscolo, editors, *Structural approaches to sequence evolution : Molecules, networks, populations*, Biological and Medical Physics, Biomedical Engineering, pages 207–232. Springer Verlag, New York, USA, 2007. ISBN 978-3-540-35305-8.
- [2] Original S code by Richard A. Becker, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka, and Alex Deckmyn. *maps : Draw Geographical Maps*, 2022. R package version 3.4.1.
- [3] A. Hackett-Pain, J.J. Foest, I.S. Pearse, J.M. LaMontagne, W.D. Koenig, G. Vacchiano, M. Bogdziewicz, T. Caignard, P. Celebias, J. van Dormolen, M. Fernández-Martínez, J.V. Moris, C. Palaghianu, M. Pesendorfer, A. Satake, E. Schermer, A.J. Tanentzap, P.A. Thomas, D. Vecchio, A.P. Wion, T. Wohlgemuth, T. Xue, K. Abernethy, M.-C. Aravena A., M.D. Barrera, J.H. Barton, S. Boutin, E.R. Bush, S.D. Calderón, F.S. Carevic, C.V. de Castilho, J.M. Cellini, C.A. Chapman, H. Chapman, F. Chianucci, P. da Costa, L. Croisé, A. Cutini, B. Dantzer, R.J. DeRose, J.-T. Dikangadissi, E. Dimoto, F.L. da Fonseca, L. Gallo, G. Gratzer, D.F. Greene, M.A. Hadad, A.H. Herrera, K.J. Jeffery, J.F. Johnstone, U. Kalbitzer, W. Kantorowicz, C.A. Klimas, J.G.A. Lageard, J. Lane, K. Lapin, M. Ledwo, A.C. Leeper, M.V. Lencinas, A.C. Lira-Guedes, M.C. Lordon, P. Marchelli, S. Marino, H. Schmidt Van Marle, A.G. McAdam, L.R. . Momont, M. Nicolas, L.H. de Oliveira Wadt, P. Panahi, G. Martínez Pastur, T. Patterson, P. Luis Peri, Ł. Piechnik, M. Pourhashemi, C. Espinoza Quezada, F.A. Roig, K. Peña Rojas, Y. Micaela Rosas, S. Schueler, B. Seget, R. Soler, M.A. Steele, M. Toro-Manríquez, C.E.G. Tutin, T. Ukizintambara, L. White, B. Yadok, J.L. Willis, A. Zolles, M. Żywiec, and D. Ascoli. Mastree+ : time-series of plant reproductive effort from six continents. *Global Change Biology*, 00 :1–17, 2022.
- [4] W. Kantorowicz. Half a century of seed years in major tree species of Poland. *Silvae Genetica*, 49 :245–248, 2000.
- [5] A. Kremer, R. Petit, and A. Ducouso. Biologie évolutive et diversité génétique des chênes sessile et pédonculé. *Revue forestière française*, 54 :111–130, 2002.
- [6] A. Kremer, C. Plomion, Th. Leroy, A. Hermine, L. Truffaut, S. Delzon, A. Ducouso, B. Musch, J.M. Torres-Ruiz, and S. Wagner. Évolution passée et contemporaine des chênes. *Revue forestière française*, 72 :297–318, 2020.
- [7] C. Michalski. Le droit à la glandée. *Revue forestière française*, 57 :377–391, 2005.
- [8] Jeroen Ooms. *gifski : Highest Quality GIF Encoder*, 2023. R package version 1.12.0-1.

- [9] M.B. Pesendorfer, M. Bogdziewicz, J. Szymkowiak, Z. Borowski, W. Kantorowicz, J.M. Espelta, and M. Fernández-Martínez. Investigating the relationship between climate, stand age, and temporal trends in masting behavior of european forest trees. *Global Change Biology*, 26(3) :1654–1667, 2020.