

Что такое осень? Что такое весна?
Analysis of Phenological Data with the ebfk package

Elie Gurarie - University of Maryland

November 5, 2017

The efbf package

- The efbf package is an in-house package for the analysis of the phenological data in R.
- In this presentation, I show a few possible analyses WITH ALL THE CODE.
- If you install the package, you can easily recreate the figurs and replicate the results
- Пакет efbf - данные и функции для рассматривания, обработки и анализа фенологический данных в (бесплатном!) прогамме R.
- В этом докладе, я демонстрирую пару возможных анализов и ВКЛЮЧАЮ СКРИПТЫ.
- Если вы инсталируете пакет, то сможете запросто поврторить анализ и создать соответственные картины.

loading

Install and load library

```
library(ebfb)
```

Инсталироавать и загружать библиотеку

The package - first and foremost - contains the relevant data

В пакете, в первую очередь, держатся данные:

Data: sites

Table of site information: Name, Coordinates, N. Observations

```
data(sites)
```

Detailed explanation:

```
?sites
```

Таблица мест наблюдений (в основном, ООПТ).
Загружается одной строкой:

Подробное описание:

Data: sites

The first few rows:

```
head(sites)
```

Первые ряды:

В сумме: X локаций

Data: sites

Map of sites:

Kapra:

```
require(maps); require(mapdata)
map("worldHires", xlim = c(2,160), ylim = c(40,75))
points(sites$East, sites$North, col="darkblue", pch = 19))
```

Data: phenodata

This is the main dataset that contains all the of the phenological events. To load the data, type:

```
data(phenodata)
```

Detailed explanation:

```
?phenodata
```

Главные данные всех фено-явлений.
Загружается одной строкой:

Подробное описание:

phenodata

The first few rows:

```
head(phenodata, 5)
```

Each row represents a single event for a single species at a single site in a single year. The main response variable is day (of year)

Первые ряды:

Каждый ряд отражает одно явление, для одного вида, в одном месте, в одном году. Главная переменная - День (года).

phenodata\$Group

```
table(phenodata$Group)  
barplot(sort(table(phenodata$Group)), horiz = TRUE, las = 1)
```

Много растений!
Мало слизняков!

phenodata\$Events

Most common events:

Самые многочисленные явления в данных:

```
sort(table(subset(phenodata, Group == "Plants")$Event), decreasing = TRUE)[1:10]  
sort(table(subset(phenodata, Group == "Birds")$Event), decreasing = TRUE)[1:10]
```

phenodata\$Species

```
table(subset(phenodata, Group == "Plants")$Species)
```

989 plants in total

Most common plants:

Наиболее многочисленные растения:

Question 1 | Вопрос 1

Has flowering time of birches changed in the past century?

Изменился ли срок распускания цветов березы за последнее столетие?



First flowering time of birch / Распускание цветов березы *Betula pendula*

```
birches <- subset(phenodata, Species == "Betula pendula" & Event == "onset of blooming")
ddply(birches, "Site", summarize, n = length(Day),
      firstyear = min(Year), lastyear = max(Year),
      mean.day = mean(Day), sd.day = sd(Day)) %>% arrange(mean.day)
```

Earliest average (April 24) in Браславские Озера, Беларусь.

Latest (May 25) in Столбы, Сибирь.

Change in time | Изменение по времени - пример

```
with(subset(birches, Site == "Centralno-Chernozemnyj"), plot(Year, Day))
```

Hard to see a trends!

With linear model | C трендом

```
require(ggplot2)
ggplot(subset(birches, Site == "Centralno-Chernozemnyj"), aes(Year, Day)) +
  geom_point() + stat_smooth(method = "lm") + ggtitle("Centralno-Chernozemnyj")
```


Everywhere (and all birches) | Всюду (и все березы - *betula* spp.)

```
allbirches <- subset(phenodata, grepl("Betula", Species) & Event == "onset of blooming")
ggplot(allbirches, aes(Year, Day, col = Site)) + geom_point() + stat_smooth(method = "lm") + theme_few()
```

Negative trend | Отрицательная тенденция?

Simple linear model | Простая линейная модель

```
lm(Day ~ Year, data = allbirches)
```

NEGATIVE and SIGNIFICANT slope: -0.1 days / year
in all study domain.

Simple linear model | Простая линейная модель

```
lm(Day ~ Year, data = allbirches)
```

NEGATIVE and SIGNIFICANT slope: -0.1 days / year
in all study domain.

Mixed effects linear model | Смешанная линейная модель (без подробностей - более правильная модель)

```
require(lme4)  
lmer(Day ~ Year + (1 | Site), data = allbirches)
```

even MORE NEGATIVE and MORE SIGNIFICANT
slope: -0.14 days / year.

What about geographic factors (North / East)?

Combining data:

```
allbirches <- merge(allbirches, sites, by="SiteID", all = TRUE)
ggplot(allbirches, aes(Year, Day, col = North, size = East)) + geom_point() + theme_few()
```

- More eastern birches flower later ... North, less clear.
- Более восточные березы позже распускают цветы ... Северные, не так ясно ().

```
lmer(Day ~ Year + North + East + (Year | Site.x), data = allbirches)
```

All effects (YEAR - NORTH - EAST) significant.
About 1.5 days later per degree north, 1 day later
for 4 degrees east.

Все переменные (ГОД - СЕВЕР - ВОСТОК)
значимые. Примерно 1.5 дней позже на г.N и
один день на 4 г.E.

(Preliminary) Conclusions | (Предварительные) Заключение

Spring - as measured by birch flowering - is very variable, but moving earlier and earlier in the year (about 2 weeks in 100 years), even accounting for geographical variation.

(But need to look at interactions, model comparison, non-linear effects, etc.)

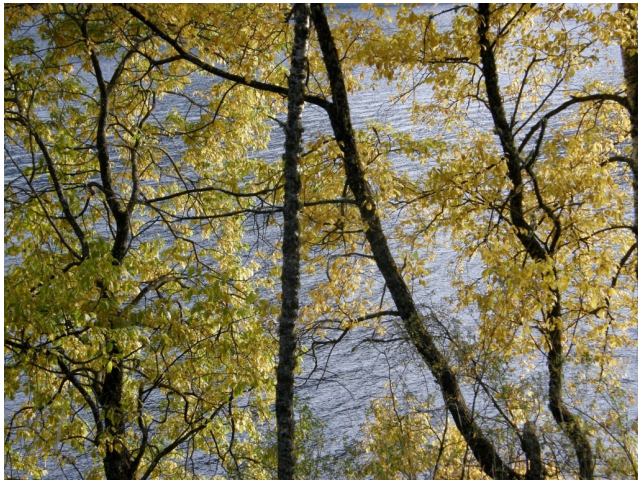
Весна - по крайней мере индекс расцветания березы - не смотря на широкий разброс - крадется все раньше и раньше (пр. 2 недели на 100 лет), даже если включать географические эффекты.

(Но надо по-подробнее моделировать).



Question 2 | Вопрос 2

- How long is fall?
- Сколько осень продолжается?



Birches again | Опять березы

Beginning and end of leaf fall:

```
birch.start <- subset(phenodata,  
                      Event == "onset of leaf fall" & grepl("Betula", Species))  
birch.end <- subset(phenodata,  
                    Event == "leaf fall end" & grepl("Betula", Species))
```

Merging the two by Site:

```
birch.fall <- merge(birch.start, birch.end, by = c('Site', 'Year'))
```

No obvious patterns:

```
with(birch.fall, plot(Day.x, Day.y, xlab = "fall start", y  
abline(0,1, col=2, lwd=2)
```

Note: a few places, fall ended before it began!

Fall duration against fall start

```
birch.fall <- mutate(birch.fall, fall.start = Day.x, fall.end = Day.y, fall.duration = fall.end - fall.start)
birch.fall <- merge(birch.fall, sites, by = "Site")
```

Against start of fall

```
ggplot(birch.fall, aes(fall.start, fall.duration,
                      size=East, col=North)) +
  geom_point(alpha = 0.3) + theme_few()
```

The later the fall starts, the shorter it is. The more East, the more North, the shorter the fall.

Against Year:

```
ggplot(birch.fall, aes(Year, fall.duration,
                      size=East, col=Site)) +
  geom_point(alpha = 0.3) + theme_few()
```

No obvious trends in time

```
lmer(fall.duration ~ fall.start + North + East + Year + (1|Site), data = birch.fall)
```

- Lots of significant effects! It seems like fall is getting a bit LONGER each year, at about the same rate as spring is moving forward.
- Много значимых эффектов, включая удлинения осени со временем - примерно таким же темпом как и весна наступает раньше.

Actual conclusions | Реальные выводы

- The package is available to any collaborator (contact me).
 - It should make it relatively EASY to explore and analyze this insanely rich dataset.
 - It will be grow, with more phenological data, meteorological data, and functions with time. We will try to keep you up to date on its progress (and availability).
 - USE IT!
- Пакет `ebfb` доступен любим сотрудником проекта.
 - Он создан для того что-бы упростить аналитическую работу с этим массивном сбором объединенных данных.
 - Пакет будет расти - добавлять данные и фенологические и метеорологические, и с новыми функциями. Но мы постараемся держать коллектив в курсе.
 - ИСПОЛЬЗУЕТЕ ЕГО!

Thanks | Спасибо

