

**КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ  
УНИВЕРСИТЕТ**

---

**Институт экологии и природопользования**

**АНАЛИЗ ДРЕВЕСНО-КОЛЬЦЕВЫХ  
ХРОНОЛОГИЙ С ПОМОЩЬЮ  
ПАКЕТА dplR**

**Учебно-методическое пособие**



**КАЗАНЬ  
2022**

*Печатается по решению  
Методической комиссии Института экологии и природопользования  
(протокол № 4 от 20 июня 2022 г.)*

**Авторы-составители:**

кандидат биологических наук, доцент **Н. А. Чижикова**,  
кандидат биологических наук, доцент **Д. В. Тишин**,  
магистрант **Л. Э. Закирова**,  
доктор биологических наук, профессор **А. А. Савельев**,  
кандидат биологических наук, доцент **С. С. Мухарамова**,  
старший преподаватель **А. Г. Пилюгин**

**Научный редактор:**

кандидат физико-математических наук, доцент **Т. В. Никоненкова**

Анализ древесно-кольцевых хронологий с помощью пакета dplR /  
Н.А. Чижикова [и др.] – Казань: Казанский федеральный университет, 2022. –  
69 с.

Пособие предназначено для студентов бакалавров и исследователей, заинтересованных в изучении методов дендрохронологии. Пособие включает обзор функций пакета dplR среды статистического программирования R, реализующих управление данными, вычисление индексов и хронологий радиального прироста, статистический и графический анализ древесно-кольцевых серий. Даны примеры использования функций и упражнения, отражающие этапы построения хронологий.

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	4
ФУНКЦИИ УПРАВЛЕНИЯ ДАННЫМИ.....	9
read.rwl().....	9
read.ids() .....	11
combine.rwl().....	14
i.detrend.series().....	16
detrend.series().....	19
detrend().....	24
chron().....	26
strip.rwl().....	28
ФУНКЦИИ ВЫЧИСЛЕНИЯ ОПИСАТЕЛЬНЫХ СТАТИСТИК.....	33
rwl.stats() и summary().....	33
ccf.series.rwl().....	34
gini.coef().....	38
corr.rwl.seg().....	39
rwi.stats.running() и rwi.stats().....	42
ФУНКЦИИ ПОСТРОЕНИЯ ГРАФИКОВ.....	47
seg.plot() .....	47
crn.plot().....	48
spag.plot() .....	50
УПРАЖНЕНИЕ. АНАЛИЗ РАДИАЛЬНОГО РОСТА СОСНЫ.....	52
Исходные данные.....	52
Схема анализа.....	52
Анализ .....	52
СПИСОК ЛИТЕРАТУРЫ.....	65
ПРИЛОЖЕНИЕ. Описание функции ar() пакета stats.....	67

## ВВЕДЕНИЕ

Древесно-кольцевые хронологии являются косвенными источниками данных высокого разрешения для реконструкции климатических изменений. Число работ в области изучения хронологий радиального роста колец, опубликованных за последнее десятилетие, резко возросло из-за всё более явных климатических изменений и более значительной и масштабной реакции древесных видов на эти изменения. Древесно-кольцевые хронологии полезны как для понимания прошлой климатической изменчивости, так и для оценки реакции растений в будущем в ответ на продолжающиеся климатические сдвиги.

Годичные кольца дают возможность оценить состояние дерева и древостоя в целом, возраст дерева, начало воздействия позитивного или негативного фактора, отразившегося на их размерах и других их свойствах (плотности древесины, количестве сосудов, толщине клеточной стенки, толщине ранней и поздней древесины).

В настоящем пособии раскрыта практическая сторона анализа, а именно инструменты пакета `dplR` (Bunn A.G., 2008; Bunn et al., 2021) среды статистического анализа R (R Core Team, 2022), полезные для работы с данными измерений древесных колец.

Исследователь должен владеть базовыми концепциями и методами, используемыми для построения древесно-кольцевой хронологии, чтобы соотнести радиальный рост деревьев с климатическими изменениями. Данные концепции и методика древесно-кольцевого анализа кратко рассмотрены в учебно-методическом пособии Тишина Д. В. и Чижиковой Н. А. (2018) «Дендрохронология», а в более подробном освещении могут быть найдены в следующих работах:

- Шиятов С.Г., Ваганов Е.А. и др. «Методы дендрохронологии. Часть 1. Основы дендрохронологии. Сбор и получение древесно-кольцевой информации» (2000). Пособие предназначено для студентов, научных сотрудников и аспирантов, специализирующихся в сфере экологии,

дендрохронологии. В пособии поэтапно описан процесс дендрохронологического анализа, начиная с выбора места и вида дерева/кустарника до измерения количественных характеристик годовых колец.

- «Growth Dynamics of Conifer Tree Rings. Images of Past and Future Environments» – это расширенная и обновленная версия книги Е. А. Ваганова и А. Шашкина, которая была опубликована на русском языке. В данном издании читатель знакомится с семейством процесс-ориентированных моделей роста древесных колец.
- Учебное пособие Шишова В.В., Тычкова И.И. и Попковой М.И. «Методы анализа дендроклиматических данных и их применение для территории Сибири» (2015), содержит обзор классических и новых методов анализа данных дендрохронологии, дендрэкологии и дендроклиматологии. В пособии есть примеры для практического применения методов и способы интерпретации полученной информации.
- Bitvinskas T. и др. «Methods of Dendrochronology – Applications in the Environmental Sciences» (1990). В этой книге представлен обзор и описание классических методов анализа годовых колец с особым акцентом на приложениях в науках об окружающей среде.
- Esper J., Gartner H. «Interpretation of tree-ring chronologies» (2001). Большинство хронологий показывают климатические изменения в длинах волн от года до десятилетий. Тенденции более длительных временных масштабов, как и вариации в масштабе столетий, как правило, труднее выделить и часто отсутствуют в хронологиях древесных колец. Как указывают авторы, цель статьи сформировать базовое понимание дендрохронологических методов для успешной интерпретации древесно-кольцевых хронологий.

- Cook E.R. «The Decomposition of Tree-Ring Series for Environmental Studies» (1987). В статье разбираются подходы по выделению различных сигналов в хронологиях.
- Cook E., Peters K. «Calculating unbiased tree-ring indices for the study of climatic and environmental change» (1997). В статье рассматриваются подходы для вычисления индексов радиального прироста.
- Cook E., Peters K. «The Smoothing Spline: A New Approach to Standardizing Forest Interior Tree-Ring Width Series for Dendroclimatic Studies» (1981), здесь можно почерпнуть дополнительные сведения о методах сглаживания и стандартизации древесно-кольцевых хронологий. Сглаживающий сплайн является хорошим методом стандартизации для временных серий колец деревьев, поскольку его процедура расчета превосходит ортогональные многочлены (полиномы), потому что не делает никаких предположений о форме кривой, необходимой для стандартизации.

Среда статистического анализа R является свободно распространяемой и может быть получена на главной странице проекта <https://www.r-project.org/>. В данном пособии не рассматриваются основы среды статистического анализа R, базовые сведения о среде и особенностях работы в ней могут быть получены читателями, к примеру в следующих учебных пособиях:

- А. Б. Шипунов и др. «Наглядная статистика. Используем R!» (2017). Один из источников, который могут использовать исследователи, начинающие знакомство со средой статистического программирования R. Здесь представлена вся необходимая информация о статистическом анализе и среде R в целом. В предисловии авторы кратко описывают актуальность издания и его предназначение: «Эта книга написана для тех, кто хочет научиться обрабатывать данные». В книге можно найти много полезного, начиная с истории развития языка, информация о самых распространенных методах статистики для анализа данных и заканчивая наглядными примерами по всем разделам книги.

- Мостицкий С. Э., Шитиков В. К. «Статистический анализ и визуализация данных с помощью R» (2015). Книга посвящена классическим и современным методам анализа данных в среде R, однако в её начале дается описание среды и языка R, его базовых возможностей.

Пакет `dplR` содержит функции, которыми пользуются для выполнения рутинных процедур при анализе древесно-кольцевых хронологий. В данном пособии не затрагиваются инструменты для перекрестной датировки колец. Эта и другая более полная информация о пакете `dplR` может быть найдена в статьях его создателей (некоторые из этих статей легли в основу данного пособия):

- Andy Bunn и Mikko Korpela «An Introduction to `dplR`» (2018). Эта статья описывает основные функции `dplR`, объясняя начальные шаги, которые аналитик должен выполнить при работе с дендрэкологическим набором данных. Кроме базовых шагов здесь описывается несколько доступных методов для анализа и дается объяснение, как использовать базовые описательные статистики, как построить хронологию, используя ширину кольца и т.д.
- Andy Bunn и Mikko Korpela «`dplR` Basics. Using R for Tree-Ring Analysis» (2018) представляет интернет-ресурс с расширенными и дополненными материалами предыдущей статьи. Здесь рассматриваются основные действия по работе с древесно-кольцевыми данными в том порядке, которому может следовать пользователь. Например, чтение данных, удаление линии тренда, построение хронологии и проведение предварительного исследовательского анализа данных с помощью описательной статистики, выполнение более сложного анализа с помощью функций пакета `dplR`.
- Andy Bunn и Mikko Korpela «Time Series Analysis in `dplR`» «Анализ временных рядов в `dplR`» (2018) раскрывает вопрос о анализе временных рядов в пакете `dplR`.

- Andrew G. Bunn «Statistical and visual crossdating in R using the dplR library» (2010). Статья посвящена демонстрации функций dplR, которые позволяют выполнять статистическое и визуальное перекрестное датирование временных серий годовых колец. Корректно датированные серии ширины колец являются необходимой основой для изучения с климатом. В статье описываются возможности для изучения корреляции между временными рядами, способы вычисления взаимной корреляции, графические инструменты, позволяющие идентифицировать проблемы датировки.
- Andy Bunn и Mikko Korpela «Crossdating in dplR» (2018) – этот интернет-ресурс также посвящен перекрестной датировке с помощью инструментов пакета dplR.

Команды, с помощью которых выполняются действия в среде R выделяются в тексте пособия следующим шрифтом:

```
> install.packages("dplR")
```

Текст, выводимый средой R на экран консоли в результате выполнения команд, выделяется в тексте следующим шрифтом:

```
There does not appear to be a header in the rwl
file
There are 5 series
1      72      1785      2009      0.01
2      74      1646      2009      0.01
```

Текущую версию пакета dplR можно установить из сети архивов (The Comprehensive R Archive Network, CRAN) с помощью команды:

```
> install.packages("dplR")
```

Каждый раз, когда начинается новый сеанс работы в среде R перед использованием функций пакета dplR нужно его инициализировать командой:

```
> library(dplR)
```



## ФУНКЦИИ УПРАВЛЕНИЯ ДАННЫМИ

### `read.rwl()`

Эта функция считывает данные о ширине колец, находящиеся в файле типа «.rwl». По сути это оболочка для вызова других функций, которые фактически реализуют операцию чтения `rwl` файла, имеющего внутреннюю структуру в соответствии с различными стандартами, существующими в дендрохронологии: `read.tucson`, `read.compact`, `read.tridas`, `read.fh`, `csv2rwl`.

Ниже приведен синтаксис функции `read.rwl()`, а описание её аргументов см. табл. 1:

```
read.rwl(fname, format = c("auto", "tucson", "compact", "tridas", "heidelberg",  
                           "csv"), ...)
```

Таблица 1  
Аргументы функции `read.rwl()`

<code>fname</code>	Имя файла в формате <code>rwl</code> .
<code>format</code>	Символьный вектор, указывающий формат структуры файла. Может принимать значения "auto" (автоматическое распознавание), "taxon", "compact", "tridas", "Heidelberg" или "CSV". Автоматическое определение формата файла "auto" используется по умолчанию, но в результате работы этой опции формат может быть распознан некорректно, поэтому в случае неудачи с аргументом, заданным по умолчанию, можно попробовать запустить функцию считывания файла с различными другими вариантами формата.
...	Аргументы, специфичные для функции, реализующей операцию для выбранного формата (в зависимости от выбранного формата функции <code>read.tucson</code> , <code>read.compact</code> , <code>read.tridas</code> , <code>read.fh</code> , <code>csv2rwl</code> ).

### Пример 1

Считаем файл `01.rwl` (для упражнений) со структурой данных в формате `tucson`, поместим полученный объект в переменную под названием **data1**. Для этого действия можно использовать любую из трёх команд, приведённых ниже:

```
> data1 <- read.rwl(fname = "01.rwl",  
                   format = "tucson")  
> data1 <- read.tucson(fname = "01.rwl",  
                      header = NULL)  
> data1 <- read.rwl(fname = "01.rwl",  
                   format = "auto")
```

В результате любой из этих команд будет создана переменная `data1`, являющаяся объектом формата `rw1`. В результате работы первых двух команд на экран будет выведено сообщение:

```
There does not appear to be a header in the rw1 file
There are 5 series
1          72          1785      2009      0.01
2          74          1646      2009      0.01
3          78          1695      2009      0.01
4          79          1950      2009      0.01
5          79A         1580      1810      0.01
```

Это сообщение содержит информацию:

- в файле не найден заголовок (описание данных);
- в созданном объекте **`data1`** находится информация о пяти деревьях (с именами 72, 74, 78, 79, 79A, см. второй столбец сообщения на экране);
- о дате начала (третий столбец сообщения) и конца измерений (четвертый столбец сообщения) в каждой серии, а именно 1785-2009, 1646-2009, 1695-2009, 1950-2009, 1580-1810;
- в пятом столбце таблицы сообщения находится величина, на которую нужно умножить значения ширины колец (0.01), чтобы получить истинные значения. Во многих исследованиях измерения ширины колец указываются в мм, умноженных на 100, поэтому требуется данные в файлах умножать на 0.01.

В результате работы третьей команды будет выведено похожее сообщение, в котором дополнительно сообщается, что формат данных в файле был распознан автоматически и было сделано предположение, что это формат файлов «Tucson». Все остальные строки сообщения идентичны предыдущим:

```
Attempting to automatically detect format.
Assuming a Tucson format file.
There does not appear to be a header in the rw1 file
There are 5 series
1          72          1785      2009      0.01
2          74          1646      2009      0.01
3          78          1695      2009      0.01
4          79          1950      2009      0.01
5          79A         1580      1810      0.01
```

## read.ids()

Данная функция и её синоним *autoread.ids* считывают данные о местоположении, самом дереве и ID керна из файла формата *data.frame* (таблица данных, в колонках которой может храниться информация разных форматов – текстовая, числовая, логическая и т.д.).

Синтаксис функций представлен ниже, описание аргументов см. табл. 2:

```
read.ids(rwl, stc = c(3, 2, 3), ignore.site.case = FALSE, ignore.case = FALSE,  
fix.typos = FALSE, typo.ratio = 5, use.cor = TRUE)
```

```
autoread.ids(rwl, ignore.site.case = TRUE, ignore.case = "auto",  
fix.typos = TRUE, typo.ratio = 5, use.cor = TRUE)
```

Таблица 2  
Аргументы функции *read.ids()*

<i>rwl</i>	Имя таблицы данных <i>data.frame</i> , в которой содержатся временные серии измерений кернов. В каждом столбце такой таблицы содержатся измерения керна, а каждая строка соответствует году формирования измеренного кольца.
<i>stc</i>	Вектор из трех целых чисел или символьной строки "auto", используется для того, чтобы разделить на части имя керна (колонки), т.к. часто имя представляет собой аббревиатуру названия места произрастания дерева и идентификаторов (часто порядковых номеров) дерева, с которого был отобран керн, и идентификатора керна (иногда у одного дерева отбирается несколько кернов). Если аргумент <i>stc</i> задается как вектор из трех целых чисел <i>stc = c(stc1, stc2, stc3)</i> , то числа указывают количество символов для деления кода местоположения <i>stc</i> ), идентификатора дерева <i>stc2</i> и идентификатора керна <i>stc3</i> . По умолчанию аргумент <i>stc = "auto"</i> , что соответствует вектору <i>c(3, 2, 3)</i> , т.е. по умолчанию считается, что в имени колонки первые три символа это аббревиатура места, следующие два символа – это идентификатор дерева и следующие три символа – это идентификатор керна.
<i>use.cor</i>	Логическая константа. Если аргумент <i>use.cor = true</i> и <i>stc = "auto"</i> , то для определения размера дерева и его частей используется кластеризация, основанная на корреляции между кернами. (примечание: возможно эта функция полезна в случае, если данные представляют собой части кернов без указания, какие части к каким кернам или деревьям относятся).

ignore.site.case	Логическая константа. Если ignore.site.case = TRUE, то функция не различает прописные и строчные буквы в названиях местоположений.
ignore.case	Логическая константа или "auto". Если значение ignore.case = TRUE, функция не различает прописные и строчные буквы в идентификаторах дерева / керна. Значение по умолчанию в read.ids = FALSE, то есть прописные и строчные буквы трактуются как разные буквы. Значение по умолчанию в read.ids = "auto", что означает, что функция чувствительна к регистру. В режиме "auto" функция обычно игнорирует различия регистров, если только это не приведет к дополнительным повторяющимся комбинациям идентификаторов дерева и керна.
typo.ratio	числовое значение больше 1, влияющее на готовность функции исправлять опечатки. Значение по умолчанию - 5.

Поскольку дендрохронологи часто берут более одного керна на дерево, иногда бывает полезно вычислить сходство кернов внутри и между деревьями. Международный Банк данных древесных колец (International tree ring data base, ITRDB) использует обозначения кернов, состоящие из восьми символов, но часто названия кернов короче. Обычно создатели gw1-файлов используют метод логической маркировки, который позволяет пользователю по названию колонки (названию, подписи керна) определить дерево и идентификатор керна. Аргумент stc показывает, как название каждой серии разделяется на местообитание, дерево и основные идентификаторы. Например, название керна может быть "ABC011", указывающий на место отбора керна "ABC", дерево № 1, керн № 1. Если нет возможности определить схему названия керна (а соответственно и нет возможности в автоматическом режиме считывать идентификаторы деревьев и кернов), то результирующий набор данных можно построить вручную.

Функция autoread.ids – это оболочка для функции read.ids с параметром stc= "auto", то есть автоматическое определение схемы названия «место отбора / дерево / керн», и различные значения по умолчанию некоторых параметров. В автоматическом режиме имена в одном и том же файле gw1 могут даже следовать различным схемам названия «место отбора / дерево / керн».

Поскольку существует множество возможных схем кодирования для именования рядов измерений, функция не всегда может дать правильный результат.

## Пример 2

Вызовем из памяти набор данных `ca533`, встроенный в пакет `dplR`:

```
> data(ca533)
```

В этом наборе данных представлена ширина годичных колец сосны остистой *Pinus longaeva*, произрастающей на горе Кампито в Калифорнии, США (Graybill и др., 1983). Таблица `ca533` состоит из 1358 строк, которые представляют года формирования древесного кольца, и 34 колонок, в которых находятся временные серии (измерения кернов). Каждая строка – это год измерения. Самое первое измеренное кольцо в этой таблице сформировано в 626 году. Год формирования самого последнего кольца – 1983. Каждая колонка – это серия ширины колец (или другая характеристика древесного кольца). Названия колонок имеют названия «CAM011», «CAM021», «CAM031», «CAM032» и т.д. В этих названиях зашифрованы коды места отбора CAM, идентификаторы дерева 01, 02, 03 и т.д., и идентификаторы кернов 1, 2 и т.д. Так, например, у дерева под номером 03 было отобрано два керна с идентификаторами 1 и 2.

Прочитаем с помощью функции `read.ids` данные из набора `ca533`, и сохраним информацию о кернах в переменную по имени `data2` с помощью любой из приведенных ниже команд:

```
> data2 <- read.ids(ca533, stc = c(3, 2, 3))
> data2 <- autoread.ids(ca533)
```

Параметр `stc` имеет значения `c(3, 2, 3)`, т.е. 3 символа может содержаться в индексе местоположения (CAM), следующие 2 символа могут содержать информацию о порядковом номере дерева и еще 3 знака включают в себя номер керна. Выведем на экран содержимое первых пяти строк таблицы `data2`:

```
> head(data2)
```

На экране появится текст:

```
      tree core
CAM011     1     1
```

```

CAM021    2    1
CAM031    3    1
CAM032    3    2
CAM041    4    1
CAM042    4    2

```

Видно, что таблица **data2** состоит из двух колонок, название каждой строки – это имя керна, в колонке **tree** содержатся идентификаторы деревьев, а в колонке **core** содержатся идентификаторы кернов. Выполните команду, которая подсчитает число кернов, приходящееся на каждое дерево:

```
> table(data2$tree)
```

Глядя на этот список, ответьте на вопрос: у каких деревьев было отобрано более одного керна?

### **combine.rwl()**

Эта функция объединяет любое количество таблиц о ширине древесных колец типа `data.frames` в одну таблицу. Для успешной работы функции требуется, чтобы в каждой таблице строки были расположены так, чтобы года формирования колец шли в порядке увеличения. Если объединяемые таблицы данных имеют другой формат, результат будет неправильным. Также необходимо, чтобы номера лет являлись именами строк (`rownames`) объединяемых таблиц.

Синтаксис функции, описание параметров см. табл. 3:

*combine.rwl(x, y)*

Таблица 3  
Аргументы функции `combine.rwl()`

x	объект (таблица данных) типа <code>data.frame</code> , которую надо связать с таблицей <code>y</code> .
y	объект (таблица данных) типа <code>data.frame</code> , которую надо связать с таблицей <code>x</code> .

### **Пример 3**

Вызовем из памяти наборы данных **ca533** и **co021**, предоставленные пакетом `dplR`:

```

>data(ca533)
>data(co021)

```

Именами строк в эти таблицы являются годы формирования измеренных колец, а в каждой колонке обеих таблиц находятся измерения кернов. В таблице **ca533** находятся ряды длиной 1538 лет (таблица содержит 1538 строк) с 626 года по 1983 год:

```
> nrow(ca533)
[1] 1358
> rownames(ca533)
[1] "626" "627" "628" "629" "630" "631"
[7] "632" "633" "634" "635" "636" "637"
...
[1351] "1976" "1977" "1978" "1979" "1980" "1981"
[1357] "1982" "1983"
```

В таблице **co021** находятся ряды длиной 788 лет (таблица содержит 788 строк) с 1176 года по 1963 год:

```
> nrow(co021)
[1] 788
> rownames(co021 )
[1] "1176" "1177" "1178" "1179" "1180" "1181"
[7] "1182" "1183" "1184" "1185" "1186" "1187"
...
[781] "1956" "1957" "1958" "1959" "1960" "1961"
[787] "1962" "1963"
```

Объединим таблицы **ca533** и **co021** в одну таблицу. Попробуем два способа указания таблиц в качестве аргументов функции `combine.rwl`, результат работы функций сохраним в переменные **combi1** и **combi2**:

```
>combi1 <- combine.rwl(list(ca533, co021))
>combi2 <- combine.rwl(ca533, co021)
```

Оба способа выдадут идентичный результат. Проверим, являются ли результаты работы функций, сохраненные в переменные **combi1** и **combi2**, идентичными:

```
>identical(combi1, combi2)
TRUE
```

На экран выводится сообщение TRUE, означающее, что обе переменные идентичны. В работе можно использовать любой способ записи аргументов для функции `combine.rwl`.

### **i.detrend.series()**

Функция позволяет интерактивно удалить (возрастной) тренд из одной временной серии ширины колец деревьев одним из пяти методов: с помощью сглаживающего сплайна, отрицательной экспоненты, линейной модели или среднего значения. Эта функция позволяет пользователю выбрать метод удаления тренда на основании графика, который в действительности создается функцией `detrend.series()`. Также функция `i.detrend.series()` выполняет стандартизацию временной серии. Стандартизация осуществляется путем деления каждой серии на тренд биологического роста для получения безразмерного индекса ширины кольца (ring-width index, RWI). Результатом работы функции `i.detrend.series()` является числовой вектор, который содержит индекс ширины кольца.

В настоящее время в пакете `dplR` доступно шесть методов удаления тренда, хотя, безусловно, существующих методов гораздо больше. Методы, реализованные в пакете `dplR`: сглаживающий сплайн «Spline», модифицированная отрицательная экспонента «ModNegExp», простая горизонтальная линия «Mean», невязка авторегрессионной модели "Ar", «супер сглаживатель» Фридмана «Friedman» и модифицированная кривая Хугершоффа «ModHugershoff».

Пользователь интерактивно указывает свой выбор метода удаления тренда с помощью клавиатуры. На выходе функцией создается вектор индекса роста, содержащий ряд измерений ( $y$ ), из которых был удален возрастной тренд и который был стандартизован. Имя ряда можно нанести на график, имя ряда можно узнать с помощью функции `colnames(y)`.

Ниже представлен синтаксис функции, описание параметров см. табл. 4:

$$i.detrend.series(y, y.name = NULL, nyrs = NULL, f = 0.5, \\ pos.slope = FALSE)$$

Таблица 4  
Аргументы функции `i.detrend.series()`

$y$	Числовой вектор. Обычно это серия измеренных ширины колец деревьев.
-----	---



y.name	Имя ряда для подписи ряда на графике.
make.plot	Логическая константа. Выводить ли на экран графики.
nyrs	Число от 0 до 1, задающее жесткость сглаживающего сплайна, если метод удаления тренда "Spline". По умолчанию nyrs = 0.67 длины ряда.
f	Число от 0 до 1, задающее частотную характеристику или длину волны отсечения, если метод удаления тренда "Spline". По умолчанию 0.5.
pos.slope	Логическая константа для метода "ModNegExp". Если pos.slope = TRUE, то в оценку возрастного тренда роста будет добавлена линейная компонента с положительным наклоном. Если значение pos.slope = FALSE, то будет добавлена линейная компонента (горизонтальная линия).

#### Пример 4

Возьмем объект **data1** из предыдущего Примера 1. Выберем из таблицы одну серию измерения керна, к примеру из второй колонки этой таблицы, сохраним эти данные в вектор под названием **series**. Преобразуем числовой вектор **series** в именованный вектор, подписав каждый элемент вектора именами строк из таблицы **data1**:

```
> series <- data1[, 2]
> names(series) <- rownames(data1)
```

В этой в векторе **series** находятся ширины годовых колец одного из кернов, измеренные в миллиметрах. Выведем первые и последние пять значений этого вектора (аббревиатура NA обозначает неизвестное, отсутствующее значение):

```
> head(series)
1580 1581 1582 1583 1584 1585
   NA   NA   NA   NA   NA   NA
> tail(series)
2004 2005 2006 2007 2008 2009
0.52 0.46 0.49 0.40 0.50 0.56
```

Удалим возрастной тренд из этой вектора **series**. Имя серии измерений **series** возьмём как имя второй колонки **data1**. Результат вычисления индекса ширины годовых колец сохраним в переменную под названием **rwil**:

```
> rwil <- i.detrend.series(series,
                           y.name = colnames(data1)[2])
```

В появившемся окне (рис. 1) изображены результаты применения шести методов удаления тренда для того, чтобы пользователь мог выбрать подходящий метод. В консоли появится предложение указать метод удаления тренда:

```
Choose a detrending method for this series CAM021.  
Methods are:  
1: Spline  
2: ModNegExp  
3: Mean  
4: Ar  
5: Friedman  
6: ModHugershoff  
Enter a number
```

Какой именно метод выбрать исследователь решает сам, на основании внешнего вида полученных индексов. Из методов удаления тренда рекомендуется выбирать те, что позволяют получить индексы радиального роста, у которых среднее и дисперсия сохраняются постоянными.

Наиболее стабилизированными индексами радиального роста на рис. 1 выглядят результаты применения методов удаления тренда Spline, ModNegExp, Friedman и ModHugershoff.

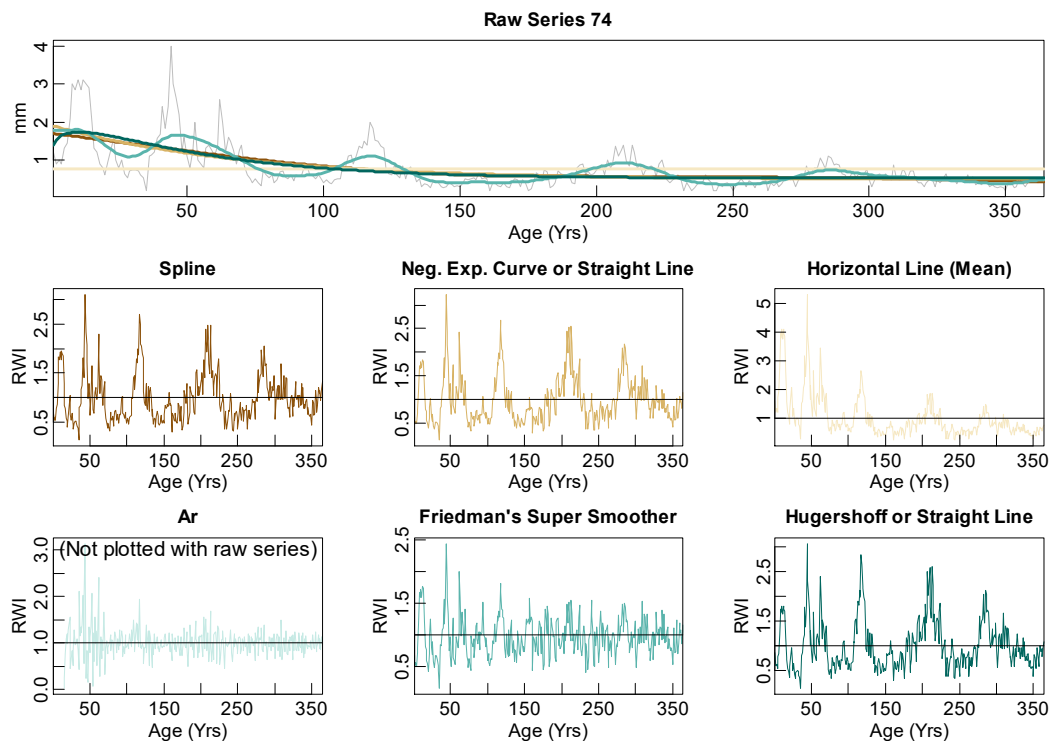


Рис. 1. Визуализация возможных вариантов удаления тренда с помощью функции `i.detrend.series()`

Выберем, к примеру, метод «Friedman», введём с клавиатуры цифру 5:

```
Enter a number 5
```

В результате работы функции `i.detrend.series` будет создан вектор, содержащий безразмерные индексированные значения ширины годовых колец, который мы сохранили в вектор `rwil`. Выведем первые и последние пять значений этого вектора:

```
> head(rwil)
1580 1581 1582 1583 1584 1585
   NA   NA   NA   NA   NA   NA
> tail(rwil)
2004      2005      2006      2007      2008      2009
1.18490 1.02724 1.07281 0.85894 1.05344 1.15803
```

### **detrend.series()**

Эта функция аналогична функции `i.detrend.series()`, но вычисляет индекс ширины годовых колец, удаляя биологический тренд роста и стандартизирует серии измерений в автоматическом режиме. Функция не является интерактивной, при вызове этой функции требуется сразу указать метод удаления тренда. Синтаксис функции:

```
detrend.series(y, y.name = names(rwl), make.plot = FALSE,
method = c("Spline", "ModNegExp", "Mean", "Ar", "Friedman",
"ModHugershoff"), f = 0.5, pos.slope = FALSE)
```

Основные аргументы функции аналогичны `i.detrend.series()` (табл. 5), но требуется указание метода удаления тренда.

Таблица 5  
Аргументы функции `detrend.series()`

<code>y</code>	Числовой вектор. Обычно это серия измеренных ширин колец деревьев.
<code>y.name</code>	Имя ряда для подписи ряда на графике.
<code>make.plot</code>	Логическая константа. Выводить ли на экран график. По умолчанию <code>make.plot = TRUE</code> , что означает «да, выводить».
<code>method</code>	Символьный вектор, с помощью которого указывают методы удаления тренда. Возможные значения вектора – это один, или несколько элементов ряда <code>c("Spline", "ModNegExp", "Mean", "Ar", "Friedman", "ModHugershoff")</code> . Можно указать одно, или несколько значений, т.е. выбрать один, или несколько методов. По умолчанию считается, что пользователь выбирает все шесть методов.

nuys	Число от 0 до 1, задающее жесткость сглаживающего сплайна, если метод удаления тренда "Spline". По умолчанию nuys = 0.67 длины ряда.
f	Число от 0 до 1, задающее частотную характеристику или длину волны отсечения, если метод удаления тренда "Spline". По умолчанию 0.5.
pos.slope	Логическая константа для метода "ModNegExp". Если pos.slope = TRUE, то в оценку возрастного тренда роста будет добавлена линейная компонента с положительным наклоном. Если значение pos.slope = FALSE, то будет добавлена линейная компонента (горизонтальная линия).
return.info	Логическое значение. Если return.info = TRUE, то результатом работы функции помимо индексов будет ещё и информация о моделях трендов (в частности – элемент curves, содержащий кривую тренда, которая вычиталась из исходных измерений).

Если пользователь выбрал один метод удаления тренда, то на выходе будет получен вектор индекса ширины годовых колец. Если были выбраны несколько методов, то результат работы функции – таблица формата data.frame, где в каждой колонке будет вычисленный индекс, соответствующий применению одного из заданных методов.

### Пример 5

Возьмем объект с ширинами годовых колец **series** из предыдущего Примера 4. Вычислим индекс ширины годовых колец с помощью метода удаления тренда «Friedman» и сохраним результат работы функции в переменную rwi2:

```
> rwi2 <- detrend.series(series,
  y.name = colnames(data1)[2],
  method = "Friedman")
```

Т.к. по умолчанию make.plot = TRUE, то появится график (рис. 2);, где будет показана исходная серия данных; тренд, оцененный с помощью выбранного метода; и вид итогового индекса ширины колец.

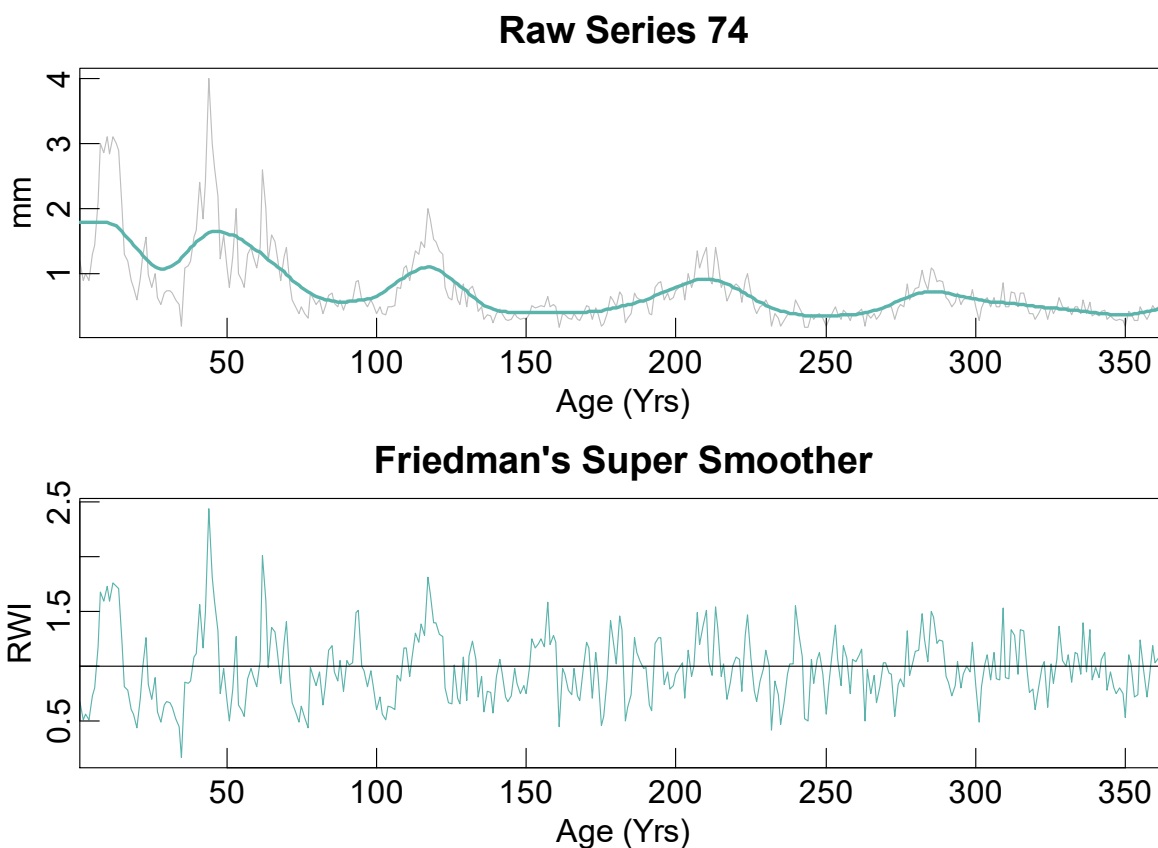


Рис. 2. Визуализация работы функции `detrend.series()`

Созданный объект `rwi2` является именованным вектором, в котором находится безразмерный индекс ширины годовых колец. Т. к. в этом упражнении был использован тот же самый метод удаления тренда, что и в Примере 4, то результат должен быть идентичным. Выведем первые и последние пять элементов вектора `rwi2`, они идентичны результату `rwi1` Примера 4, сравните:

```
> head(rwi2)
1580 1581 1582 1583 1584 1585
   NA   NA   NA   NA   NA   NA
> tail(rwi2)
      2004      2005      2006      2007      2008      2009
1.18490 1.02724 1.07281 0.85894 1.05344 1.15803
```

### Пример 6

Возьмем объект с ширинами годовых колец `series` из Примера 4. Вычислим индекс ширины годовых колец с помощью двух методов удаления тренда: «Spline» и «Friedman», сохраним индексы и информацию о моделях в переменную `rwi3`:

```
> rwi3 <- detrend.series(series,
  y.name = colnames(data1)[2],
  method = c("Spline", "Friedman"),
  return.info = TRUE)
```

Результат работы функции также изобразится графически (рис. 3).

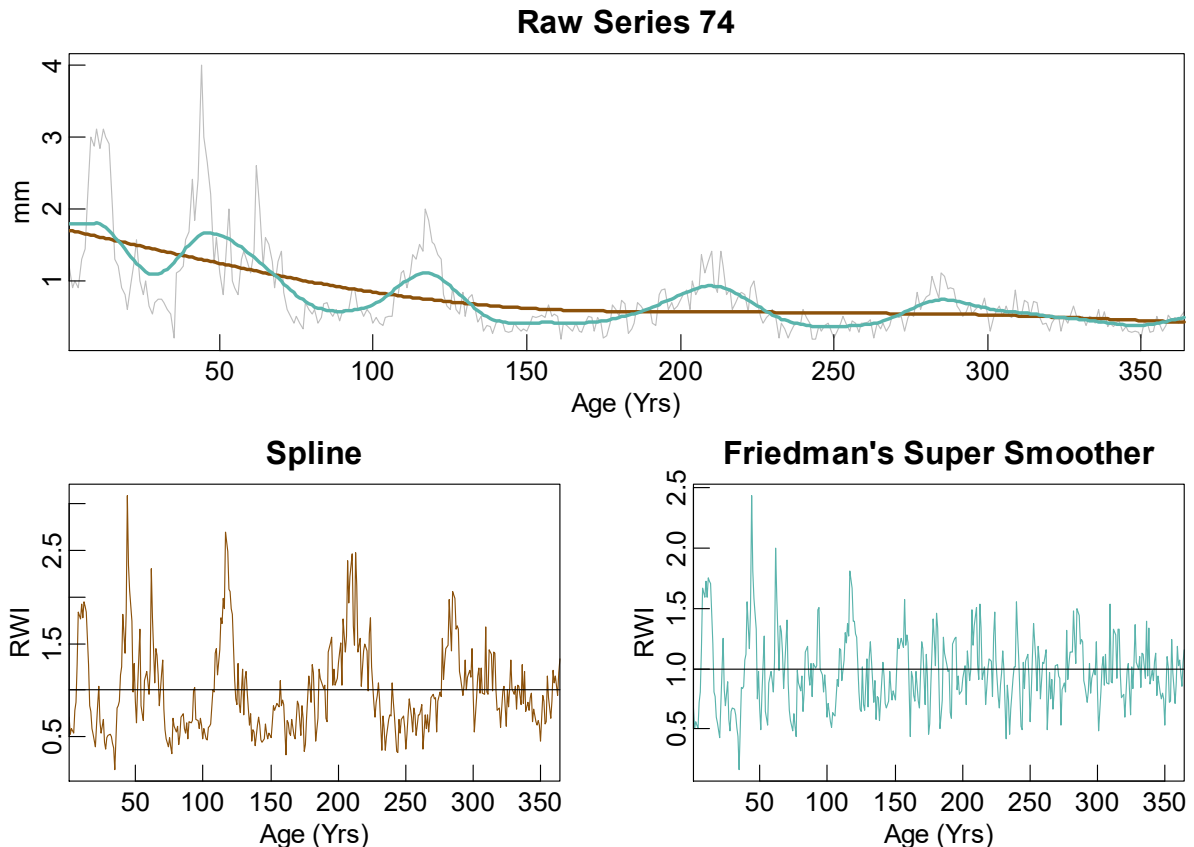


Рис. 3. Визуализация работы функции `detrend.series()` в случае применения двух методов удаления тренда

Созданный объект `rwi3` является уже не вектором, а списком.

Посмотрим его структуру, из каких элементов он состоит:

```
> str(rwi3)
List of 4
 $ series :'data.frame'   430 obs. of  2 variables:
  ..$ Spline  : num [1:430] NA NA NA NA NA NA NA NA
  ..$ Friedman: num [1:430] NA NA NA NA NA NA NA NA
 $ curves  :'data.frame'   430 obs. of  2 variables:
  ..$ Spline  : num [1:430] NA NA NA NA NA NA NA NA
  ..$ Friedman: num [1:430] NA NA NA NA NA NA NA NA
 $ model.info:List of 2
  ..$ Spline  :List of 5
  .. ..$ method   : chr "Spline"
  .. ..$ nyrs     : num 243
  .. ..$ f        : num 0.5
```

```

.. ..$ n.zeros      : int 0
.. ..$ zero.years: chr(0)
..$ Friedman:List of 6
.. ..$ method      : chr "Friedman"
.. ..$ wt          : chr "default"
.. ..$ span        : chr "cv"
.. ..$ bass        : num 0
.. ..$ n.zeros     : int 0
.. ..$ zero.years: chr(0)
$ data.info :List of 2
..$ n.zeros      : int 0
..$ zero.years: chr(0)

```

Список состоит из 4 элементов (List of 4):

- Элемент `$series` – это таблица, в её колонках находятся два вычисленных индекса ширины годовых колец. Первый индекс – с использованием метода `Spline`, второй индекс – с использованием метода `Friedman`.
- Элемент `$curves` – это таблица, в её колонках находятся значения кривых трендов биологического роста, оцененных с помощью методов `Spline` и `Friedman`. Эти тренды вычитались из исходных значений ширины годовых колец и использовались для стандартизации (обезразмеривания) рядов данных.
- Элемент `$model.info` – это список, в нём перечислены характеристики моделей трендов.
- Элемент `$data.info` – список, где находятся характеристики серий. Количество нулевых значений (отсутствующих колец) и их положения в ряде чисел.

Построим линейный график тренда биологического роста, оцененного с помощью метода `Spline` (рис. 4):

```

> plot(x = rownames(rwi3$curves),
       y = rwi3$curves$Spline,
       xlab = "Год ", ylab = "RWI",
       main = "Метод Spline", las = 1, type = "l")

```

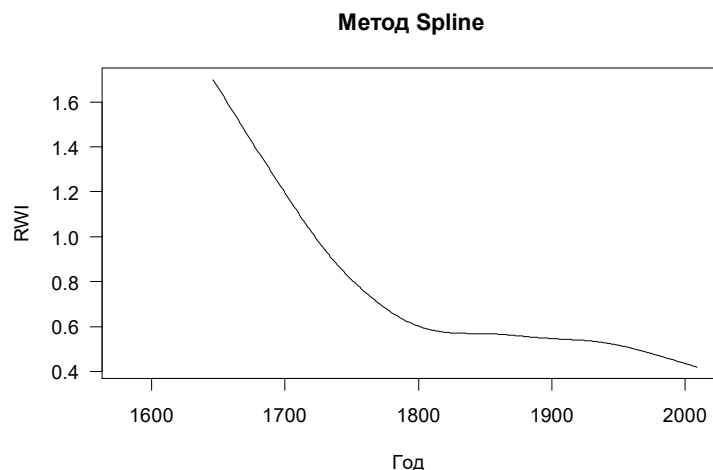


Рис. 4. Тренд биологического роста, оцененных с помощью метода Spline

### **detrend()**

Эта функция аналогичная функции `detrend.series()`. Она в автоматическом режиме вычисляет индексы ширины годовых колец, но не для одной серии как функция `detrend.series()`, а для нескольких серий одновременно, находящихся в таблице данных. Метод оценки биологического тренда роста нужно указать сразу при запуске функции, и этот выбранный метод будет применен к каждой серии исходной таблицы данных. Синтаксис функции приведён ниже, а расшифровка её основных аргументов дана в табл. 6.

```
detrend(rwl, y.name = names(rwl), make.plot = FALSE,
method = c("Spline", "ModNegExp", "Mean", "Ar", "Friedman",
"ModHugershoff"), f = 0.5, pos.slope = FALSE)
```

Основные аргументы функции аналогичны `detrend.series()` (табл. 5), но требуется указание метода удаления тренда.

Таблица 5  
Аргументы функции `detrend.series()`

<code>rwl</code>	Таблица данных <code>data.frame</code> , где в каждой колонке – измерения годовых колец керна, а имена строк - годы.
<code>y.name</code>	Имя ряда для подписи ряда на графике.
<code>make.plot</code>	Логическая константа. Выводить ли на экран график. По умолчанию <code>make.plot = TRUE</code> , что означает «да, выводить».
<code>method</code>	Символьный вектор, с помощью которого указывают методы удаления тренда. Возможные значения вектора – это один, или несколько элементов ряда <code>c("Spline", "ModNegExp", "Mean", "Ar", "Friedman", "ModHugershoff")</code> . Можно указать одно, или



	несколько значений, т.е. выбрать один, или несколько методов. По умолчанию считается, что пользователь выбирает все шесть методов.
parms	Число от 0 до 1, задающее жесткость сглаживающего сплайна, если метод удаления тренда "Spline". По умолчанию parms = 0.67 длины ряда.
f	Число от 0 до 1, задающее частотную характеристику или длину волны отсечения, если метод удаления тренда "Spline". По умолчанию 0.5.
pos.slope	Логическая константа для метода "ModNegExp". Если pos.slope = TRUE, то в оценку возрастного тренда роста будет добавлена линейная компонента с положительным наклоном. Если значение pos.slope = FALSE, то будет добавлена линейная компонента (горизонтальная линия).
return.info	Логическое значение. Если return.info = TRUE, то результатом работы функции помимо индексов будет ещё и информация о моделях трендов (в частности – элемент curves, содержащий кривую тренда, которая вычиталась из исходных измерений).

Если пользователь выбрал один метод удаления тренда, то на выходе будет получена таблица формата data.frame индексов ширины годичных колец, соответствующая применению этого метода к каждой временной серии исходной таблицы данных. Если были выбраны несколько методов, то результат работы функции – список, элементы которого – таблицы формата data.frame, каждая таблица соответствует индексам радиального роста, полученным с помощью одного из указанных методов.

### Пример 7

Вычислим индексы ширины годичных колец для измерений кернов в объекте data1 (см. Пример 1), укажем метод «Spline» оценки тренда биологического роста, сохраним в переменную data2:

```
> data2 <- detrend(data1, method = "Spline")
```

Сейчас в объекте data2 находится таблица из 430 строк (в соответствии с годами наблюдений) и 5 колонок (в соответствии с числом кернов), в ячейках таблицы – безразмерные индексы радиального прироста, полученные путем вычитания модели биологического роста и стандартизации. Выведем в подтверждение слов структуру объекта data2:

```
> str(data2)
```

```
'data.frame': 430 obs. of 5 variables:
 $ 72 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 74 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 78 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 79 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 79A: num 1.75 1.46 1.76 2.56 1.18 ...
```

### chron()

Функция строит мастер-хронологию средних значений, используя, как правило, таблицу типа `data.frame`, в которой находятся индексы ширины колец, вычисленные с помощью функции `detrend()`. При построении мастер-хронологии происходит либо (1) усреднение строк исходной таблицы данных с использованием среднего или робастного среднего (так называемая стандартная хронология, *standard chronology*), либо (2) усредняются остатки индексов, полученные после учёта авторегрессии (остаточная хронология, *residual chronology*).

Синтаксис функции приведён ниже, описание параметров см. табл. 6:

*chron(x, prefix = "xxx", biweight = TRUE, prewhiten = FALSE, ...)*

Таблица 6  
Аргументы функции `chron()`

x	Таблица данных типа <code>data.frame</code> , в которой содержатся индексы радиального прироста. Именами строк таблицы содержат годы, а имена колонок содержат идентификаторы серии индексов (керна, или деревьев). Обычно эти имена считываются в исходную таблицу данных из файла данных с помощью функции <a href="#">read.rwl</a> .
prefix	Символьная строка, которая может содержать менее 4 символов. По умолчанию используется " xxx". Это будущее название (идентификатор) мастер-хронологии, указывается по желанию пользователя.
biweight	логическая константа. Если <code>biweight = TRUE</code> , то с помощью функции <code>tbrm</code> вычисляется робастное среднее (стандартная хронология, <i>standard chronology</i> ). Это значение используется по умолчанию.

prewhiten	логическая константа. Если prewhiten = TRUE, то перед усреднением каждый ряд «отбеливается» (whitening) с помощью подходящей модели автокорреляции для временных рядов, т.е. создается выбеленная хронология, или хронология остатков (residual chronology). По умолчанию используется значение FALSE, если не задано иное.
-----------	---

Функции `chron` можно передать аргументы, которые использует функция `ar()` из пакета `stats` для построения модели авторегрессионной модели в случае когда аргумент `prewhiten = TRUE`. Например, аргументы `aic` и `order.max` (см. Приложение).

### Пример 8

Построим мастер-хронологию по измерениям ширины годичных колец таблицы `ca533`, предоставленной пакетом `dplR`. Инициализируем набор данных:

```
> data(ca533)
```

Эта таблица состоит из строк, которые представляют года, и колонок, в которых находятся временные серии. Каждая строка – это год измерения. Каждая колонка – это серия ширины кольца. Первый год измерений 626. Последний год измерений – 1983.

Выполним удаление тренда с помощью функции `detrend()`, используем метод оценки тренда биологического роста с помощью отрицательной экспоненциальной функции и затем применим саму функцию `chron`:

```
> ca533.rwi <- detrend(rwl = ca533,
                      method = "ModNegExp")
> ca533.crn <- chron(ca533.rwi, prefix = "CAM")
```

Мастер-хронология будет находиться в объекте по имени `ca533.crn`. Имя этой хронологии "CAM". Напрямую мы можем обратиться с таблице данных этого объекта. Например, можно узнать структуру этой таблицы, число строк и столбцов:

```
> str(ca533.crn)
Classes 'crn' and 'data.frame': 1358 obs. of 2 variables:
 $ CAMstd      : num  0.347 0.266 0.287 0.391 0.453.
 $ samp.depth: num  1 1 1 1 1 1 1 1 1 1 ...
```

В таблице имеется две колонки, в первой колонке `CAMstd` находится сама стандартная мастер-хронология (т.к. вычисление опирается на

робастное среднее). Во второй колонке под названием `samp.depth` находится число кернов, или временных серий, которые были использованы для вычисления среднего в этот год.

Можно визуализировать информацию, содержащуюся в мастер-хронологии. Построим график хронологии (рис. 5):

```
> plot.crn(ca533.crn)
```

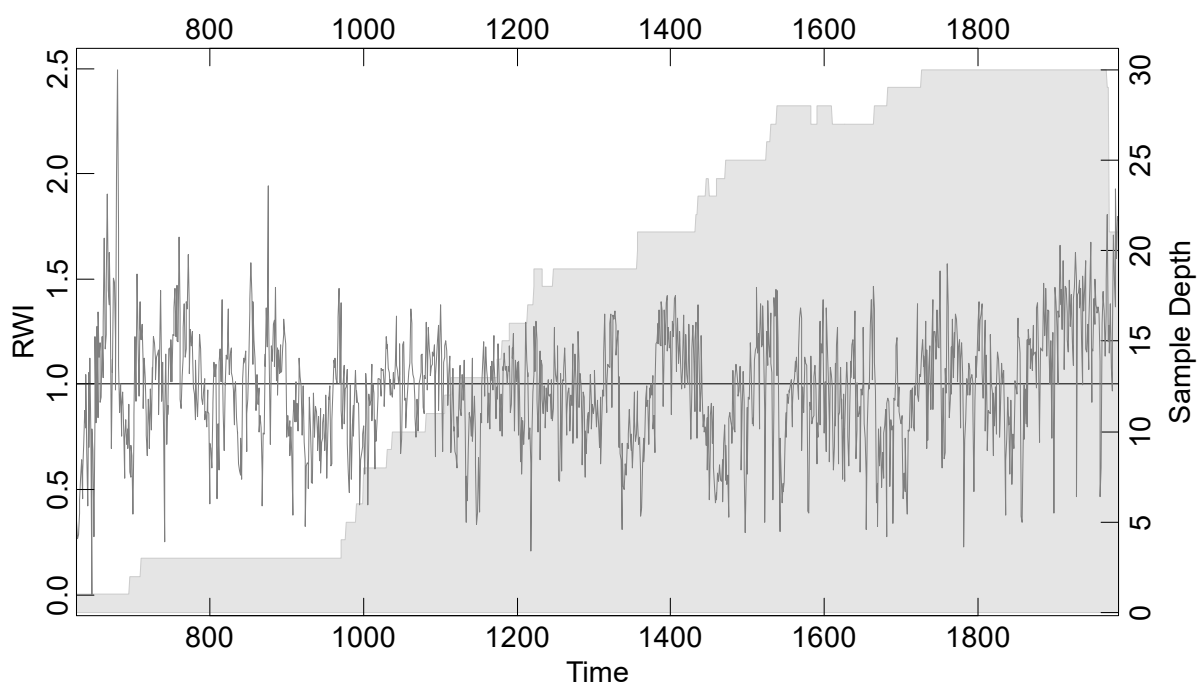


Рис. 5. График стандартной хронологии набора данных **ca533**. На левой оси ординат откладывается значение среднего индекса радиального роста (хронологии), на правой оси откладывается «Sample depth» – число кернов, или серий, которые участвовали в усреднении в текущий год. Хронология показана линией, число кернов показано заливкой

### **strip.rwl()**

Данная функция вычисляет мастер-хронологию так, чтобы улучшить её чувствительность, или индекс EPS. EPS (expressed population signal) или выраженный сигнал популяции – это критерий оценки надежности древесной хронологии. Значение этого сигнала показывает, насколько хорошо хронология подходит для целей реконструкции климата. Функция `strip.rwl()` пытается убрать из набора данных для вычисления мастер-хронологии те индексы прироста, которые понижают EPS мастер-хронологии (понижают её чувствительность). На вход функция `strip.rwl()` принимает исходные ширины радиальных колец (не индексы прироста) и сама убирает тренд биологического роста с помощью двойного удаления тренда, используя

сплайн и частотные отсечки. После этого алгоритм строит хронологию по всем имеющимся данным и вычисляет её EPS. Затем на каждом шаге алгоритм вычисляет значения EPS хронологии, которая получена с использованием всех индексов, кроме одного. Так, по очереди строится набор хронологий, путем удаления одного из индексов. Индекс, удаление которого увеличивает общий EPS больше всего, отбрасывается. Полученный EPS после удаления одного индекса прироста берётся как эталонный и алгоритм переходит к следующему шагу. На нём повторяется процедура поочередного удаления индексов и вычисления новых хронологий, в поиске следующего индекса, удаление которого поможет повысить EPS. Это повторяется до тех пор, пока невозможно будет достигнуть дальнейшего увеличения EPS путем отбрасывания одной серии. Затем процедура повторяется в обратном направлении, т. е. происходит повторное включение каждой ранее удаленной серии в хронологию. В каждом цикле повторно в хронологию включается тот индекс, который позволяет увеличить EPS больше всего. Когда улучшение EPS становится невозможно, алгоритм завершает свою работу. Происходящие изменения могут выводиться алгоритмом на экран, там публикуются достигнутые EPS и исключенные или добавленные серии индексов. Если удаление или повторная вставка одной серии приводит к улучшению EPS, эта серия помечается звездочкой. Результат работы функции – таблица серий ширины годовых колец, из которых исключены те серии, которые приводят к уменьшению EPS.

Синтаксис функции дан ниже, описание аргументов см. табл. 7:

*strip.rwl(rwl, ids = NULL, verbose = FALSE, comp.plot = FALSE,  
legacy.eps = FALSE)*

Таблица 7  
Аргументы функции `strip.rwl()`

rwl	Таблица данных с исходными данными о ширинах годовых колец в столбцах и годами в подписи строк.
ids	дополнительные данные типа <code>data.frame</code> , где первый столбец это идентификатор дерева, и второй столбец с числовым идентификатором для каждого керна. Такая таблица создается

	с помощью функции <code>read.ids</code> , при условии наличия такой информации. Этот аргумент не является обязательным и может отсутствовать.
<code>verbose</code>	логическая константа, указывающая на то, должны ли выводиться на экран значения EPS и происходящих изменений во время процесса улучшения хронологии. Значение <code>TRUE</code> – будут выводиться на экран, <code>FALSE</code> – не будут выводиться на экран.
<code>comp.plot</code>	логическая константа, указывающая на то, должен ли быть нарисован диагностический график с годичной EPS.
<code>legacy.eps</code>	логическая константа, указывающая на то, будет ли EPS рассчитываться с помощью функции <code>rwi.stats</code> ( <code>FALSE</code> , значение по умолчанию) или <code>rwi.stats.legacy</code> ( <code>TRUE</code> ).

### Пример 9

Получим набор данных, позволяющих вычислить мастер-хронологию, обладающую максимальным EPS из возможных, на примере набора данных `ca533`, встроенного в пакет `dplR`. Инициализируем данные и применим функцию `strip.rwl`:

```
> data(ca533)
> eps_ca533 <- strip.rwl(ca533)
REMOVE -- Iteration 1: leaving series CAM041 out.
EPS improved from 0.936 to 0.937.
REMOVE -- Iteration 2: leaving series CAM152 out.
EPS improved from 0.937 to 0.938.

REMOVE -- Iteration 3: no improvement of EPS.
Aborting...
REINSERT -- Iteration 1: no improvement of EPS.
Aborting...
```

В консоли выведены сообщения, из которых следует, что удаление двух хронологий `CAM041` и `CAM152` позволили несколько увеличить EPS с 0.936 до 0.938.

### Пример 10

Получим набор данных, позволяющих вычислить мастер-хронологию, обладающую максимальным EPS из возможных, на примере набора данных `anos1`, встроенного в пакет `dplR`. Этот набор данных содержит измерения

ширины колец норвежской ели *Picea abies* L. из Баварии, Германия (Zang, 2010): 20 кернов, относящихся к 10 деревьям в столбцах, измерения покрывают 98 лет. Инициализируем данные:

```
> data(anos1)
```

Выведем на экран названия колонок объекта `anos1`, чтобы проверить свои предположения о системе кодификации:

```
> names(anos1)
[1] "ANOS111K" "ANOS101G" "ANOS101K" "ANOS102G"
[5] "ANOS102K" "ANOS105G" "ANOS105O" "ANOS106G"
[9] "ANOS106K" "ANOS107G" "ANOS107K" "ANOS108G"
[13] "ANOS108K" "ANOS111G" "ANOS114K" "ANOS110G"
[17] "ANOS110K" "ANOS112G" "ANOS112K" "ANOS114G"
```

В названии колонки первые четыре символа обозначают код местоположения, следующие четыре символа кодируют идентификатор дерева, и один символ является идентификатором керна. Прочитаем из таблицы используя заголовки колонок информацию о деревьях и относящихся к ним кернам, параметр `stc` зададим как вектор трех чисел `c(4, 3, 1)`:

```
> anos1.ids <- read.ids(anos1, stc = c(4, 3, 1))
```

Построим мастер-хронологию, максимизируя EPS, сохраним результат в переменной **`epschron`**:

```
> epschron <- strip.rwl(anos1, ids = anos1.ids,
                        verbose = TRUE)
REMOVE -- Iteration 1: Initial EPS: 0.931
Leave-one-out EPS:
ANOS111K: 0.929
ANOS101G: 0.93
ANOS101K: 0.925
ANOS102G: 0.931
ANOS102K: 0.926
ANOS105G: 0.923
ANOS105O: 0.922
ANOS106G: 0.922
ANOS106K: 0.926
ANOS107G: 0.926
ANOS107K: 0.928
ANOS108G: 0.927
ANOS108K: 0.926
ANOS111G: 0.923
ANOS114K: 0.928
```

```
ANOS110G: 0.927
ANOS110K: 0.926
ANOS112G: 0.931
ANOS112K: 0.928
ANOS114G: 0.926
```

```
***
```

```
REMOVE -- Iteration 1: no improvement of EPS.
Aborting...
```

Сообщения на экране консоли свидетельствуют о том, что на первом этапе была предпринята попытка удалить одну хронологию, но ни одно удаление не привело к значительному росту EPS, поэтому алгоритм завершил работу. Таблица с данными не изменилась.



# ФУНКЦИИ ВЫЧИСЛЕНИЯ ОПИСАТЕЛЬНЫХ СТАТИСТИК

## `rwl.stats()` и `summary()`

Эти функции вычисляет различные описательные статистики, часто используемые в дендрохронологии. На вход обычно этим функциям подают таблицу данных, содержащую исходные ширины годичных колец, которые обычно порождаются с помощью функции `read.rwl` и др.

Результат работы обеих функций одинаков, это таблица типа `data.frame`, которая содержит описательные статистики по каждой серии измерений (по каждой хронологии). Таблица с описательными статистиками содержит следующие колонки: название серии, год формирования первого `first` и последнего `last` измеренного кольца, количество лет измерений `year`, средняя `mean` ширина кольца, медиана `median` распределения ширины колец, стандартное отклонение `stdev` распределения ширины колец, асимметрия распределения ширины колец `skew`, коэффициент неоднородности Джини (Gini) `gini.coef` и коэффициент автокорреляции первого порядка `ar1`. Синтаксис функций представлен ниже:

```
rwl.stats(rwl)
```

```
summary(rwl, ...)
```

Обе функции на вход принимают таблицу `rwl` с данными. В таблицы строки соответствуют годам измерений, а колонки – ширине колец.

### Пример 11

Расчитаем описательные статистики для набора данных **anos1**. Вызовем из памяти набор данных **anos1** и рассчитаем описательные статистики:

```
> data(ca533)
> rwl.stats(ca533)
```

```
  series first last year mean median stdev skew gini ar1
1 ANOS111K 1955 2007  53 3.890  3.860 1.128 0.265 0.161 0.358
2 ANOS101G 1932 2007  76 2.253  2.010 1.184 1.716 0.267 0.616
3 ANOS101K 1919 2007  89 1.793  1.710 0.876 0.361 0.277 0.748
4 ANOS102G 1924 2007  84 1.942  1.910 0.968 0.320 0.278 0.748
5 ANOS102K 1910 2007  98 2.504  2.445 1.140 0.850 0.249 0.636
```

6	ANOS105G	1943	2007	65	1.566	1.340	0.740	1.118	0.245	0.767
7	ANOS105O	1913	2007	95	1.849	1.770	0.688	0.681	0.206	0.614
8	ANOS106G	1910	2007	98	1.794	1.720	0.904	0.791	0.278	0.678
9	ANOS106K	1927	2007	81	2.668	2.320	1.318	0.428	0.279	0.686
10	ANOS107G	1916	2007	92	2.147	2.100	0.724	0.607	0.186	0.652
11	ANOS107K	1916	2007	92	1.912	1.595	0.986	0.757	0.285	0.764
12	ANOS108G	1913	2007	95	2.275	2.070	1.482	0.581	0.364	0.828
13	ANOS108K	1918	2007	90	2.321	2.385	1.070	-0.040	0.260	0.669
14	ANOS111G	1954	2007	54	2.928	3.090	1.108	-0.161	0.214	0.611
15	ANOS114K	1911	2007	97	2.259	2.030	1.112	1.077	0.263	0.676
16	ANOS110G	1930	2007	78	1.940	1.905	1.166	0.726	0.332	0.721
17	ANOS110K	1932	2007	76	2.589	2.270	1.436	0.498	0.311	0.748
18	ANOS112G	1922	2007	86	2.761	2.430	1.636	1.602	0.309	0.666
19	ANOS112K	1916	2007	92	2.666	2.430	1.144	0.746	0.237	0.504
20	ANOS114G	1911	2007	97	2.366	2.270	1.061	1.794	0.224	0.687

### **ccf.series.rwl()**

Данная функция вычисляет перекрестные корреляции между выбранной серией ширины древесных колец и мастер-хронологией, построенной из таблицы данных типа `rwl`, в заданных пользователем временных лагах и временных отрезках.

Корреляции рассчитываются для первого сегмента лет, затем для второго сегмента и так далее. Корреляции рассчитываются только для сегментов, полностью покрытой значениями мастер-хронологии.

Результат работы функции может использоваться для выявления выпавших колец (`missing rings`).

Синтаксис функции представлен ниже, описание основных аргументов см. табл. 8:

*ccf.series.rwl(rwl, series, series.yrs = as.numeric(names(series)),  
seg.length = 50, bin.floor = 100, n = NULL, prewhiten = TRUE,  
biweight = TRUE, pcrit = 0.05, lag.max = 5, make.plot = TRUE,  
floor.plus1 = FALSE, series.x = FALSE)*

Таблица 8  
Аргументы функции `ccf.series.rwl()`

<code>rwl</code>	Таблица данных <code>data.frame</code> с хронологиями ширины годовых колец в столбцах и строками, соответствующими годам формирования колец.
------------------	--

series	Это может быть число, означающее номер колонки (серии данных), для которой нужно посчитать корреляцию с мастер-хронологией. Либо символьное название этой колонки.
series.yrs	числовой вектор, содержащий год серии. Игнорируется, если серия является индексом столбца <code>rw1</code> .
seg.length	Длина временных отрезков в годах (например, 20, 50, 100 лет). Это должно быть четное число.
bin.floor	неотрицательное интегральное значение, год для отсчёта первого сегмента (например, 1600, 1700, 1800). Чаще: 0, 10, 50, 100.
n	нулевое или интегральное значение, дающее длину фильтра Ханнинга, используемого для удаления низкочастотных колебаний.
prewhiten	Логическая константа. Если ее значение =TRUE, то каждая серия выбеливается с помощью учета автокорреляции первого порядка <code>ar</code> .
biweight	Логическая константа. Если ее значение =TRUE, то мастер-хронология вычисляется робастное среднее с использованием функции <code>tbrm</code> .
pcrit	Число между 0 и 1, дающее критическое значение для теста на значимость отличия корреляции от нуля.
lag.max	Число, означающее максимальный размер лага для которого вычисляется кросс-корреляция между серией и хронологией.
make.plot	логическая константа, указывающая, нужно ли строить график коэффициентов корреляции.
floor.plus1	логическая константа. Если ее значение =TRUE, то к базовому местоположению первого сегмента добавляется один год (1601, 1701, 1801 AD).
series.x	логическая константа, указывающая, следует ли сделать ряд аргументом <code>x</code> для <code>ccf</code> . <i>Примечание: Функция <code>ccf</code> вычисляет кросс-корреляцию или кросс-ковариацию двух одномерных рядов.</i>

## Пример 12

Покажем работу функции `ccf.series.rw1` на примере набора данных **co021** о ширине годовых колец пихты Дугласа *Pseudotsuga menziesii* в Меса-Верде в Колорадо, США (Schulman, 1963), а также то, как результат может быть использован для нахождения выпавших колец. Для этого вначале вызовем из памяти среды набор данных **co021**:

```
> data(co021)
```

Выберем одну временную серию ширины годовых колец, которую будем исследовать, сохраним её в вектор с именем `flagged`. Например, выберем колонку с названием «641143»:

```
> flagged <- co021$"641143"
```

Удалим информацию о выбранной серии данных из исходной таблицы:

```
> co021$"641143" <- NULL
```

Удалим любое измерение из выбранной серии ширины годовых колец, `flagged`, симулировав ситуацию наличия выпавшего кольца (при работе с реальными наблюдениями этого делать не требуется, наоборот, о наличии выпавших колец обычно ничего неизвестно, и обычно стоит задача их найти). Например, удалим измерение, которое находится в строке 325:

```
> flagged <- c(NA, flagged[-325])
```

Сделаем серию `flagged` именованным вектором (имена элементов вектора – годы формирования колец, их можно взять из подписей строк исходной таблицы данных):

```
> names(flagged) <- rownames(co021)
```

Вычислим с помощью функции `ccf.series.rwl()` корреляцию выбранной серии с мастер-хронологией, полученной по исходным данным (не включающим выбранную серию); используем временные отрезки длительностью сто лет:

```
> ccf.series.rwl(rwl = co021,
                 series = flagged,
                 seg.length = 100,
                 make.plot = TRUE)
```

NB: With `series.x = FALSE` (default), negative lags indicate missing rings in `series$ccf`

```
$ccf
      1300.1399 1350.1449 1400.1499 1450.1549 1500.1599 1550.1649 1600.1699
lag.-5      -0.05      0.14      -0.06      0.04      0.15      0.14      0.10
lag.-4       0.01      0.03      -0.11      -0.20      -0.10      0.02      0.02
lag.-3       0.08      0.11      0.06      -0.07      0.05      0.14     -0.03
lag.-2       0.07      0.06     -0.20     -0.22     -0.20     -0.10     -0.05
lag.-1       0.88      0.85      0.91      0.46     -0.11     -0.13     -0.13
lag.0        0.04     -0.05     -0.20      0.33      0.87      0.87      0.86
lag.1       -0.01      0.02      0.11      0.02     -0.02     -0.03     -0.06
lag.2        0.00     -0.04     -0.17     -0.22     -0.07      0.05     -0.04
lag.3       -0.12      0.03     -0.10     -0.12     -0.02      0.04     -0.03
lag.4       -0.10     -0.09     -0.04      0.00     -0.09     -0.08      0.00
lag.5        0.01     -0.04     -0.14     -0.07      0.01      0.02      0.02
$bins
```

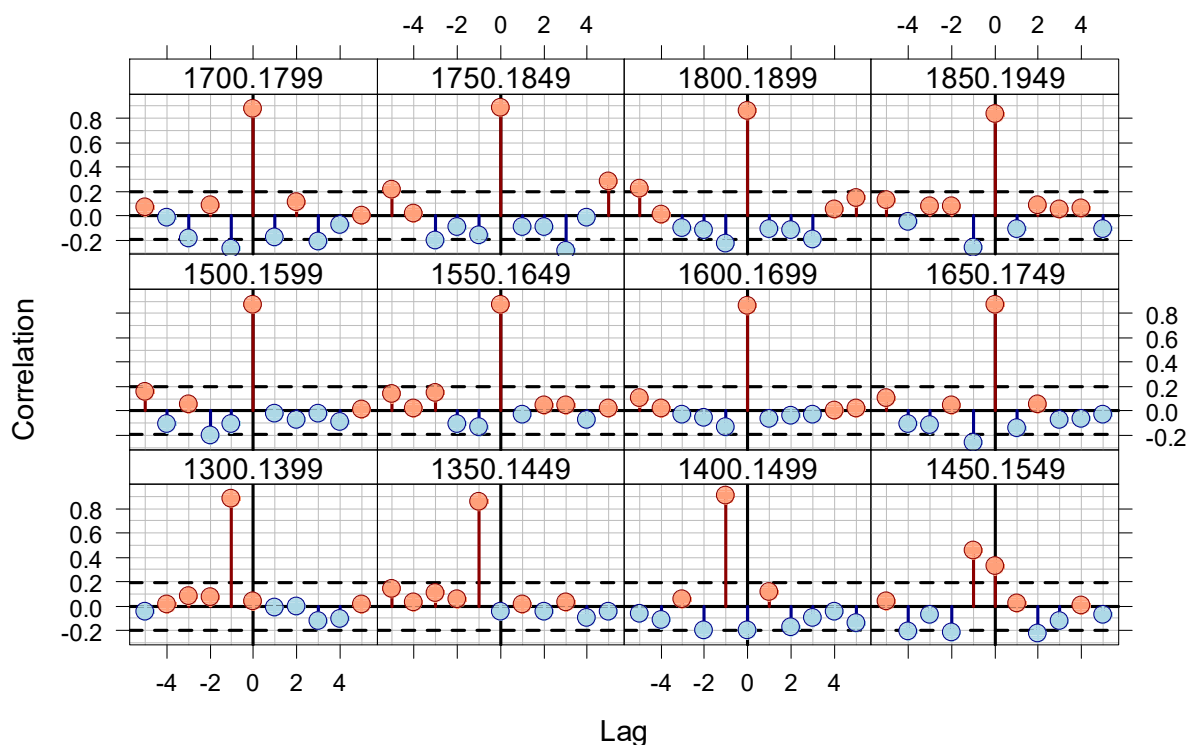
```
[,1] [,2]
[1,] 1300 1399
[2,] 1350 1449
[3,] 1400 1499
[4,] 1450 1549
[5,] 1500 1599
[6,] 1550 1649
[7,] 1600 1699
[8,] 1650 1749
[9,] 1700 1799
[10,] 1750 1849
[11,] 1800 1899
[12,] 1850 1949
```

Результат работы функции не был сохранен в переменную, а был выведен на экран (см. выше). При этом было выведено замечание «Обратите особое внимание» (NB, nota bene), которое обращает внимание на отрицательные корреляции: их наличие свидетельствует выпавших кольцах. Также на экран был выведен список, состоящий из двух элементов:

- таблицы `$ccf`, в которой находятся коэффициенты корреляции между серией и хронологией для всех временных отрезков и всех временных лагов;
- таблицы `$bins`, в которой перечислены все временные отрезки, годы начала и конца периодов, для которых рассчитывались корреляции.

Также функция `ccf.series.rwl()` построила график с иллюстрацией этих коэффициентов корреляции (рис. 10), т.к. аргумент функции `make.plot = TRUE`, при этом на рисунке также присутствует замечание обратить внимание, что отрицательные корреляции могут соответствовать выпавшим кольцам.

На графике и в таблице корреляций есть свидетельства выпавших колец, обратите внимание на низкие корреляции на лаге 0 для временных отрезков 1300-1399, 1350-1449, 1400-1499, 1450-1549. Вероятно выпавшее кольцо находится где-то между 1300 и 1549 годами. Также обратите внимание, что в эти годы наибольшая корреляция сместилась с лага 0 на один лаг, на лаг -1, что может свидетельствовать об одном выпавшем кольце.



NB: With series.x = FALSE (default), negative lags indicate missing rings in series

Рис. 6. График кросс-корреляции выделенной серии данных с мастер хронологией таблицы co021, построенный с помощью функции `ccf.series.rwl()`

### **gini.coef()**

Функция `gini.coef()` вычисляет коэффициент Джини (Gini) для серий ширины годовых колец, индексов радиального роста или мастер-хронологий. Коэффициент Джини используется в качестве меры разнообразия размеров годовых колец, обычно стандартизованных. Коэффициент Джини изменяется в интервале от 0 до 1. Значения коэффициента близкие к 0 указывают на однородное распределение размеров (размеры всех колец одинаковы), тогда как значения коэффициента близкие к 1 свидетельствуют о наличии очень небольшого числа колец, размеры которых значительно превышают размеры всех остальных.

Синтаксис функции:

*gini.coef(x)*

Функция имеет один аргумент `x`, который является числовым вектором, неоднородность которого нужно оценить.

### **Пример 13**

Оценим неоднородность индексов колец мастер-хронологии для набора данных **ca533**. Для этого инициализируем набор данных:

```
> data(ca533)
```

Вычислим индекс радиального прироста, используя удаление тренда с помощью функции `detrend()` и метода "ModNegExp".

```
> ca533.rwi <- detrend(rwl = ca533,  
                      method = "ModNegExp")
```

Затем построим хронологию средних значений с помощью функции `chron()`:

```
> ca533.crn <- chron(ca533.rwi, prefix = "CAM")
```

Выполним саму функцию `gini.coef()`:

```
> gini.coef(ca533.crn)  
[1] 0.6172432
```

Полученный коэффициент более близок к единице, чем к нулю, что свидетельствует о неоднородном распределении размеров колец в мастер-хронологии. Коэффициент не достигает единицы, а это значит, что выделяется ещё достаточно много колец, размеры которых значительно превышают другие кольца, т.к. чем ближе к единице, тем меньше колец, отличающихся большим размером.

### **corr.rwl.seg()**

Эта функция последовательно вычисляет корреляцию между каждым индексом радиального роста и мастер-хронологией, построенной с использованием всех других индексов таблицы данных (объекте типа `rwl`). При необходимости пользователь может указать в качестве аргумента основную хронологию (вектор). В этом случае для всех рядов в объекте типа `rwl` при расчете корреляции будет использоваться одна и та же мастер-хронология.

Корреляции подсчитываются для каждого сегмента серии, где сегменты перекрываются друг с другом на половину длины сегмента (например, сегменты длиной 100 лет будут идти со смещением по отношению друг к другу с шагом 50 лет). Корреляции рассчитываются для первого сегмента,

затем для второго сегмента и так далее. Корреляции рассчитываются только для сегментов с полным покрытием мастер-хронологии.

Эта функция обычно вызывается для создания графика, в котором каждый сегмент для каждой серии окрашивается в соответствии с его корреляцией с основной хронологией. Зеленые сегменты – это те, которые не покрываются мастер-хронологией полностью. Синие сегменты – это те, которые коррелируют мастер-хронологией выше заданного пользователем критического значения. Красные сегменты – это те, которые коррелируют ниже заданного пользователем критического значения и могут указывать на проблему с датировкой.

Синтаксис функции представлен ниже, описание основных аргументов см. табл. 11:

```
corr.rwl.seg(rwl, seg.length = 50, bin.floor = 100, n = NULL,
prewhiten = TRUE, pcrit = 0.05, biweight = TRUE,
method = c("spearman", "pearson", "kendall"),
make.plot = TRUE, label.cex = 1, floor.plus1 = FALSE, = NULL,
master.yrs = as.numeric(if(is.null(dim(master))))
```

Таблица 9  
Аргументы функции corr.rwl.seg()

rwl	Объект data.frame с сериями ширины годовых колец в виде столбцов и годами в виде строк (необработанные данные, таких как те, которые были созданы read.rwl).
seg.length	Чётное число, задающее длину временных отрезков в годах (например, 20, 50, 100 лет).
bin.floor	Неотрицательное целое число, дающее точку для определения местоположения первого сегмента (1600, 1700, 1800 AD). Чаще: 0, 10, 50, 100.
n	Нулевое или число, задающее длину фильтра Ханнинга, используемого для удаления низкочастотных колебаний.
prewhiten	логическая константа. Если ее значение =TRUE, то каждая серия выбеливается (удаляется автокорреляция).
pcrit	число между 0 и 1, дающее критическое значение для теста по оценке значимости отличия коэффициента корреляции от нуля.
biweight	логическая константа. Если ее значение =TRUE, то мастер-хронология вычисляется как робастное среднее с использованием функции tbrm.



method	Может принимать одно из значений ряда: "pearson", "kendall", или "spearman". Указывает на используемый коэффициент корреляции. По умолчанию используется "spearman".
make.plot	логическая константа, указывающая, нужно ли строить график.
label.cex	Число, размер меток осей на графике.
floor.plus1	Логическая константа. Если ее значение =TRUE, то к базовому местоположению первого сегмента добавляется один год (1601, 1701, 1801 AD).
master	Числовой вектор или числовая матрица, таблица данных типа data.frame. Если значение вектора master = NULL, то ряд мастер-хронологий строится из таблицы, заданной аргументом rwl с использованием принципа "leave-one-out" (убирать по одному). Если это вектор, то функция использует его в качестве основной хронологии. Если это матрица или данные типа data.frame, этот объект используется для построения главной хронологии.
master.yrs	Числовой вектор, содержит годы.

### Пример 14

Вызовем из памяти набор данных (co021), предоставленный пакетом dplR для примера:

```
> data(co021)
```

Выполним функцию corr.rwl.seg(), результат будет представлять собой график на рис. 7:

```
> corseg <- corr.rwl.seg(co021, seg.length = 100,
                        label.cex = 1.25)
```

Результат работы функции был помещен в объект с именем corseg, это список, который включает в себя много элементов: сами коэффициенты корреляции (spearman.rho), наблюдаемые уровни значимости (p.val), вычисленные индексы радиального роста (rwi) и т.д. Выведем имена элементов этого списка:

```
> names(corseg)
[1] "spearman.rho" "p.val"
[3] "overall"      "avg.seg.rho"
[5] "flags"        "bins"
[7] "rwi"          "seg.lag"
[9] "seg.length"  "pcrit"
[11] "label.cex"
```

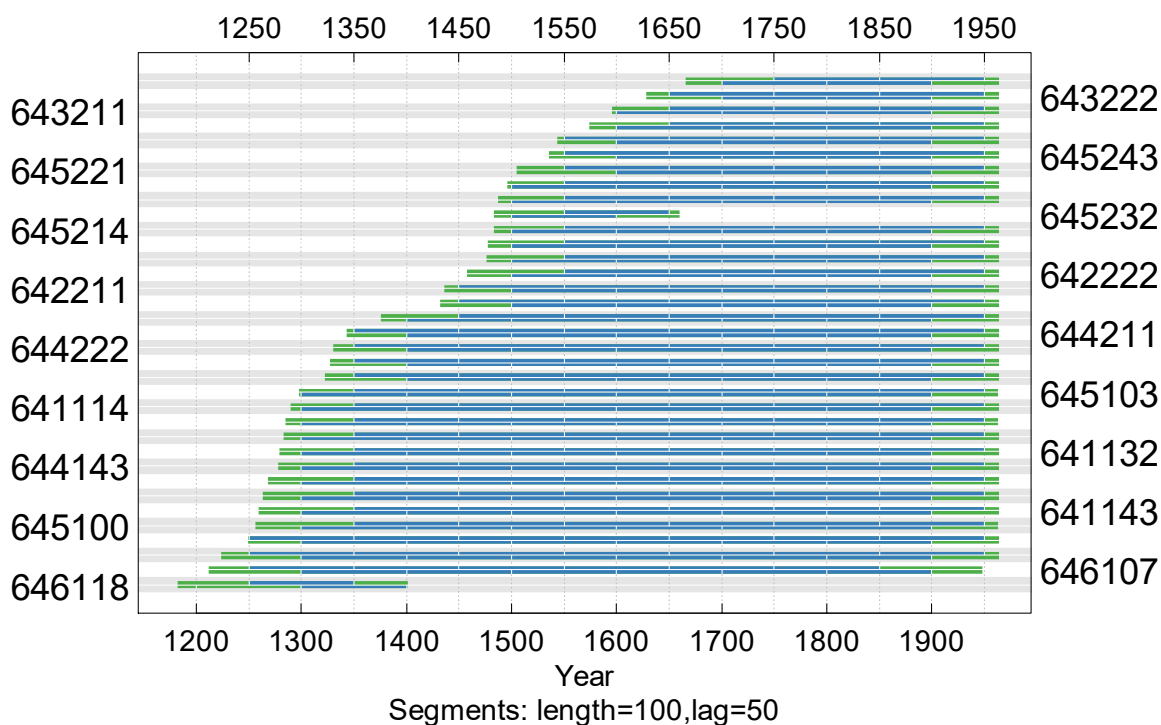


Рис. 7. График корреляции между каждым индексом радиального роста набора данных со021 с мастер-хронологией этого набора, построенный с помощью функции `corr.rwl.seg()`

Визуальный анализ графика (рис. 7) позволяет сделать вывод, что все индексы хорошо коррелируют с мастер-хронологией, что может свидетельствовать об отсутствии проблем с датировкой, а также о хорошем согласии радиального роста деревьев в обследованном участке леса.

### **`rwi.stats.running()` и `rwi.stats()`**

Данные функции вычисляют описательные статистики индексов прироста, принятые в дендрохронологии.

Функция `rwi.stats()` вычисляет статистики для индекса целиком, тогда как функция `rwi.stats.running()` вычисляет набор статистик в скользящем окне. Возможно регулировать размер скользящего окна и шаг смещения для функции `rwi.stats.running()`. Входные данные могут быть отфильтрованы таким образом, чтобы вычисления производились только над высокочастотными компонентами индексов прироста.

Для правильного расчета статистики изменчивости внутри и между временными рядами необходимо указать идентификаторы деревьев, поскольку дендрохронологи часто берут более одного керна на дерево, и соответственно на одно дерево может приходиться несколько индексов радиального роста.

Функции `rwi.stats` – это оболочка, которая вызывает `rwi.stats.running` с аргументом функции `running.window = FALSE` («нет скользящего окна»).

Синтаксис функций, и описание основных аргументов см. табл. 14:

```
rwi.stats.running(rwi, ids = NULL, period = c("max", "common"),
method = c("spearman", "pearson", "kendall"), prewhiten=FALSE,
n=NULL, running.window = TRUE, window.length = min(50, nrow(rwi)),
window.overlap = floor(window.length / 2), first.start = NULL,
min.corr.overlap = min(30, window.length), round.decimals = 3,
zero.is.missing = TRUE)
```

```
rwi.stats(rwi, ids=NULL, period=c("max", "common"),
method = c("spearman", "pearson", "kendall"), ...)
```

Таблица 10  
Аргументы функции `rwi.stats.running()`

<code>rwi</code>	Таблица типа <code>data.frame</code> с индексами радиального роста в столбцах и годами в строках. Такие таблицы индексов обычно производятся функцией <code>detrend()</code> .
<code>ids</code>	Дополнительные данные типа <code>data.frame</code> , где первый столбец это идентификатор дерева, и второй столбец идентификатор для каждого керна. По умолчанию считается, что на одно дерево приходится один керн.
<code>period</code>	Символьная строка, имеющая значения либо <code>"common"</code> , либо <code>"max"</code> , указывающая, следует ли ограничивать корреляции полными наблюдениями за период. Т.е. брать ли при расчете корреляций те строки (годы с кольцами), которые являются общими для всех кернов (т. е. строки, общие для всех выборок, <code>"common"</code> ), или брать в каждом случае строки с максимальным попарным перекрытием ( <code>"max"</code> ). По умолчанию используется значение <code>"max"</code> .

method	Может иметь значения либо "pearson", либо "kendall" или "spearman", что указывает на используемый коэффициент корреляции. По умолчанию используется "spearman".
n	Нулевое или интегральное значение, дающее длину фильтра для фильтра Ханнинга, используемого для удаления низкочастотных колебаний.
prewhiten	Логическая константа. Если ее значение истина (TRUE), то каждая серия выбеливается с помощью функции ar().
running.window	Логическая константа, указывающая, следует ли использовать скользящее окно (TRUE, по умолчанию) или игнорировать другие параметры окна и использовать одно окно, охватывающего все годы (FALSE).
window.length	Число, указывает длину скользящего окна в годах. Значение по умолчанию равно 50 годам или количеству лет (строк), в зависимости от того, что меньше.
window.overlap	Число, указывающее на перекрытие последовательных позиций окна, т. е. шаг движения скользящего окна. По умолчанию используется половина длины окна.
first.start	Необязательное числовое значение, задающее положение первого окна (год).
min.corr.overlap	Число, устанавливающее минимальное число общих лет в любой паре рядов ширины кольца, необходимое для включения их корреляции в расчеты. Считается, что меньшие перекрытия дают ненадежные значения корреляции, которые игнорируются. По умолчанию используется минимум их чисел 30 и длины окна. Один из способов снять ограничение и включить все корреляции - установить min.corr.overlap = 0.
round.decimals	Неотрицательное целое число, задающее желаемое количество десятичных знаков в результатах.
zero.is.missing	Логическая константа, указывающая, следует ли рассматривать нули как отсутствующие значения (TRUE, значение по умолчанию) или включать их в вычисления (FALSE).

Выходные данные являются таблицей типа data.frame, могут содержать следующие столбцы (каждая строка соответствует одной позиции окна), табл. 11:

Таблица 11  
Результаты функции rwi.stats.running()

start.year	Первый год в окне.
mid.year	Средний год в окне, округление в меньшую сторону.

end.year	Последний год в окне.
n.cores	Число кернов.
n.trees	Число деревьев.
n	Среднее число деревьев .
n.tot	Общее число корреляций, рассчитанных как $n.wt + n.bt$ . Равно $n.cores * (n.cores - 1) / 2$ , если существует перекрытие между всеми выборками.
n.wt	Количество вычисленных внутри дерева корреляций.
n.bt	Количество вычисленных корреляций между деревьями.
rbar.tot	Среднее значение всех корреляций между различными кернами.
rbar.wt	Среднее значение корреляций между рядами из одного дерева по всем деревьям.
rbar.bt	Средняя корреляция между всеми рядами из разных деревьев.
c.eff	Эффективное число кернов (учитывает количество корреляций в каждом дереве).
rbar.eff	Эффективный сигнал rbar, рассчитанный как $rbar.bt / (rbar.wt + (1 - rbar.wt) / c.eff)$ .
eps	Выраженный популяционный сигнал EPS, вычисляется с использованием среднего числа деревьев как $n * rbar.eff / ((n - 1) * rbar.eff + 1)$ .
snr	Отношение сигнал / шум SNR вычисляется с использованием среднего числа деревьев в виде $n * rbar.eff / (1 - rbar.eff)$ .

### Пример 15

Используем данные из объекта **data1** (см. Пример 1). Используем их для расчета индексов радиального роста, которые сохраним в объект **rwil**:

```
> rwil <- detrend(data1, method = "Spline")
```

Применим к индексам радиального роста **rwil** функции `rwi.stats.running()`, `rwi.stats()` (результат см. на следующей странице).

По результатам функции `rwi.stats()` видно, что было использовано 5 кернов, относящихся к пяти деревьям. Средняя межсерияльная корреляция индексов составила 0.35, что говорит о довольно низкой согласованности индексов прироста. Коэффициент EPS мастер-хронологии, полученной по индексам радиального роста, равен 0.59, это очень мало и также свидетельствует о плохой согласованности индексов (качественные мастер-хронологии характеризуются  $EPS > 0.85$ ). Коэффициент SNR мастер хронологии равен 1.47.

```
> rwi.stats(rwil)
n.cores n.trees      n n.tot n.wt n.bt rbar.tot rbar.wt rbar.bt c.eff rbar.eff eps snr
1         5         5 2.78      8  0    8      0.35      NA    0.35    1    0.35 0.59 1.47
```

```
> rwi.stats.running(rwil, round.decimals = 2)
start.year mid.year end.year n.cores n.trees      n n.tot n.wt n.bt rbar.tot rbar.wt rbar.bt c.eff rbar.eff eps snr
1         1635         1659         1684      2      2 1.78      1  0    1      0.26      NA    0.26    1    0.26 0.38 0.62
2         1660         1684         1709      3      3 2.30      1  0    1      0.32      NA    0.32    1    0.32 0.52 1.09
3         1685         1709         1734      3      3 2.80      3  0    3      0.52      NA    0.52    1    0.52 0.75 2.99
4         1710         1734         1759      3      3 3.00      3  0    3      0.51      NA    0.51    1    0.51 0.75 3.08
5         1735         1759         1784      3      3 3.00      3  0    3      0.68      NA    0.68    1    0.68 0.86 6.31
6         1760         1784         1809      4      4 3.50      3  0    3      0.78      NA    0.78    1    0.78 0.92 12.16
7         1785         1809         1834      4      4 3.52      3  0    3      0.53      NA    0.53    1    0.53 0.80 4.00
8         1810         1834         1859      4      4 3.02      3  0    3      0.58      NA    0.58    1    0.58 0.80 4.13
9         1835         1859         1884      3      3 3.00      3  0    3      0.44      NA    0.44    1    0.44 0.70 2.35
10        1860         1884         1909      3      3 3.00      3  0    3      0.07      NA    0.07    1    0.07 0.18 0.22
11        1885         1909         1934      3      3 3.00      3  0    3      0.64      NA    0.64    1    0.64 0.84 5.27
12        1910         1934         1959      4      4 3.20      3  0    3      0.31      NA    0.31    1    0.31 0.59 1.44
13        1935         1959         1984      4      4 3.70      6  0    6      0.24      NA    0.24    1    0.24 0.53 1.14
14        1960         1984         2009      4      4 4.00      6  0    6      0.28      NA    0.28    1    0.28 0.60 1.53
```

Статистики, вычисленные в скользящем окне, показывают, что есть отдельные временные промежутки, когда наблюдается более высокое согласие между индексами радиального роста деревьев. Например в период с 1760 по 1809 годы средняя межсерийная корреляция  $\bar{r}$  достигает значения 0.78, а EPS равен 0.92.

## ФУНКЦИИ ПОСТРОЕНИЯ ГРАФИКОВ

### `seg.plot()`

Функция создает график сегментов, иллюстрирующих длину во времени имеющихся измерений кернов. На данном графике каждая временная серия представлена отрезком (сегментом), который показывает год начала и конца серии. Данный график может быть полезен для визуализации пробелов в данных. Синтаксис функции дан ниже:

`seg.plot(rwl, ...)`.

Её единственный аргумент таблица данных типа `data.frame`, где в каждой колонке находится временная серия (обычно ширины годовых колец), а строки соответствуют годам. Данная таблица может быть получена, например, путем чтения исходных данных из файла с помощью функции `read.rwl`.

### Пример 16

Вызовем из памяти набор данных `co021`, предоставленный пакетом `dplR`:

```
> data(co021)
```

Эта таблица состоит из строк, которые представляют года, и колонок, в которых находятся временные серии. Построим график сегментов (рис. 8):

```
> seg.plot(co021)
```

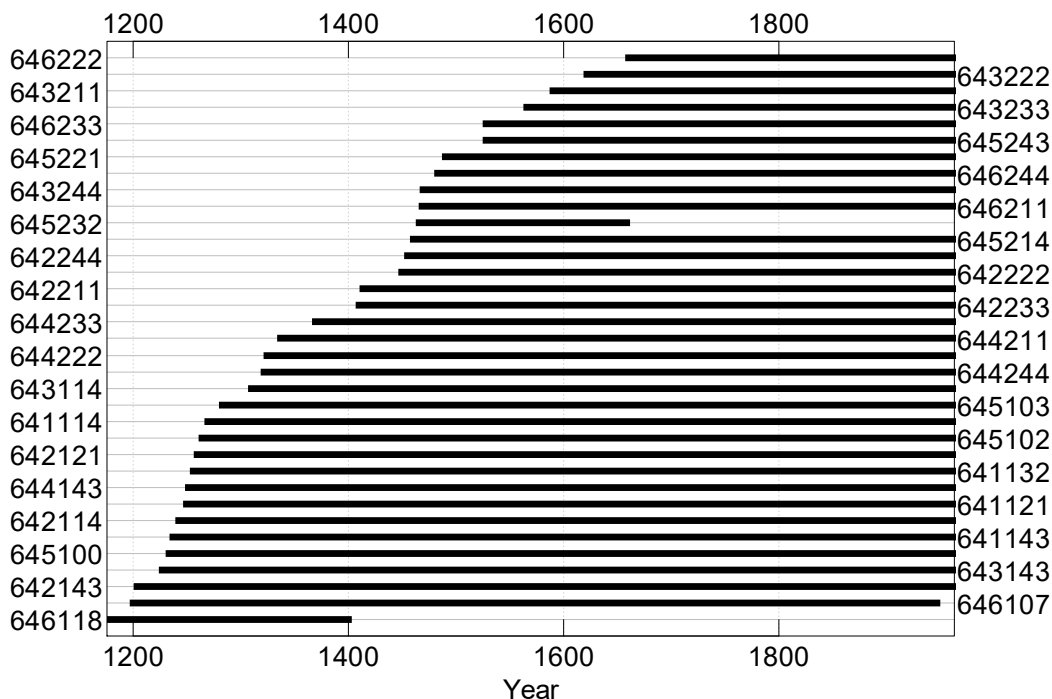


Рис. 8. График сегментов набора данных `co021`, построенный с помощью функции `seg.plot()`

На графике сегментов (рис. 8) видно, что есть два керна, в которых мало измерений (коды 646118 и 645232). Больше всего обеспечен информацией информации временной период после 1600 года.

### **crn.plot()**

Эта функция создает график мастер-хронологии по умолчанию из объекта типа `chron`.

Синтаксис функции представлен ниже, описание её аргументов см. табл.

17:

```
crn.plot(crn, add.spline = FALSE, nyrs = NULL, f = 0.5, crn.line.col = 'grey50', spline.line.col = 'red', samp.depth.col = 'grey90', samp.depth.border.col = 'grey80', crn.lwd = 1, spline.lwd = 1.5, abline.pos = 1, abline.col = 'black', abline.lty = 1, abline.lwd = 1, xlab = "Time", ylab="RWI')
```

Таблица 12  
Аргументы функции `crn.plot()`

<code>x, crn</code>	Мастер-хронология, объект, созданный функцией <code>chron</code> , либо таблица типа <code>data frame</code> . Строки таблицы соответствуют годам, столбец – хронологии. Последний столбец в таблице мастер-хронологии может содержать глубину выборки (число кернов) с именем <code>"samp.depth"</code> . Если данные получены через <code>chron</code> , то объект будет относиться к классу <code>"crn"</code> и может быть использован универсальный метод построения графика.
<code>add.spline</code>	Логическая константа. Добавляет линию со сглаживающим сплайном используя функцию <code>ffcsaps</code> (сглаживающий сплайн с заданной пользователем жесткостью и частотным срезом).
<code>nyrs</code>	Число, задающее жесткость сглаживающего сплайна. Значение по умолчанию равно 0.33 длины первой хронологии, если <code>nyrs = NULL</code> .
<code>f</code>	Число от 0 до 1, задающее частотную характеристику или длину волны отсечения. По умолчанию 0.5.
<code>crn.line.col</code>	Цвет линии, отражающей мастер-хронологию.
<code>spline.line.col</code>	Цвет линии сплайна (если он добавлен).
<code>samp.depth.col</code>	Цвет для полигона глубины выборки (если такие сведения имеются).



samp.depth.border.col	Цвет границы для полигона глубины выборки (если присутствует).
crn.lwd	Толщина линии оси x.
spline.lwd	Толщина линии сплайна (если представлена).
abline.pos	Положение опорной линии на оси Y (референсная линия). Линия не добавляется, если значение abline.pos равно NULL (1 по умолчанию).
abline.col	Цвет референсной линии.
abline.lty	Тип референсной линии.
abline.lwd	Толщина референсной линии.
xlab	Подпись оси X по умолчанию.
ylab	Подпись оси Y по умолчанию.
...	дополнительные аргументы, чтобы редактировать график. Масштаб вторичной оси y (глубина выборки) можно регулировать с помощью аргумента ylim 2.

### Пример 17

Создадим мастер-хронологию для измерений ширины годичных колец, содержащихся в файле 01.rwl. Считаем вначале данные из файла:

```
> data1 <- read.tucson(fname = "01.rwl",
                      header = NULL)
```

Выполним удаление тренда с помощью метода ModNegExp, стандартизацию измерений, получив индексы прироста:

```
> rwil <- detrend(rwl = data1,
                 method = "ModNegExp")
```

Затем выполним функцию chron():

```
> crn1 <- chron(rwil)
```

Построим график хронологии с помощью команды (рис. 9), добавим на него линию сглаживающего сплайна, чтобы проследить периода спада и снижения размеров колец:

```
> crn.plot(crn1, add.spline = TRUE, f = 0.9,
           spline.lwd = 3, spline.line.col = "red",
           crn.line.col = "black", xlab = "Год")
```

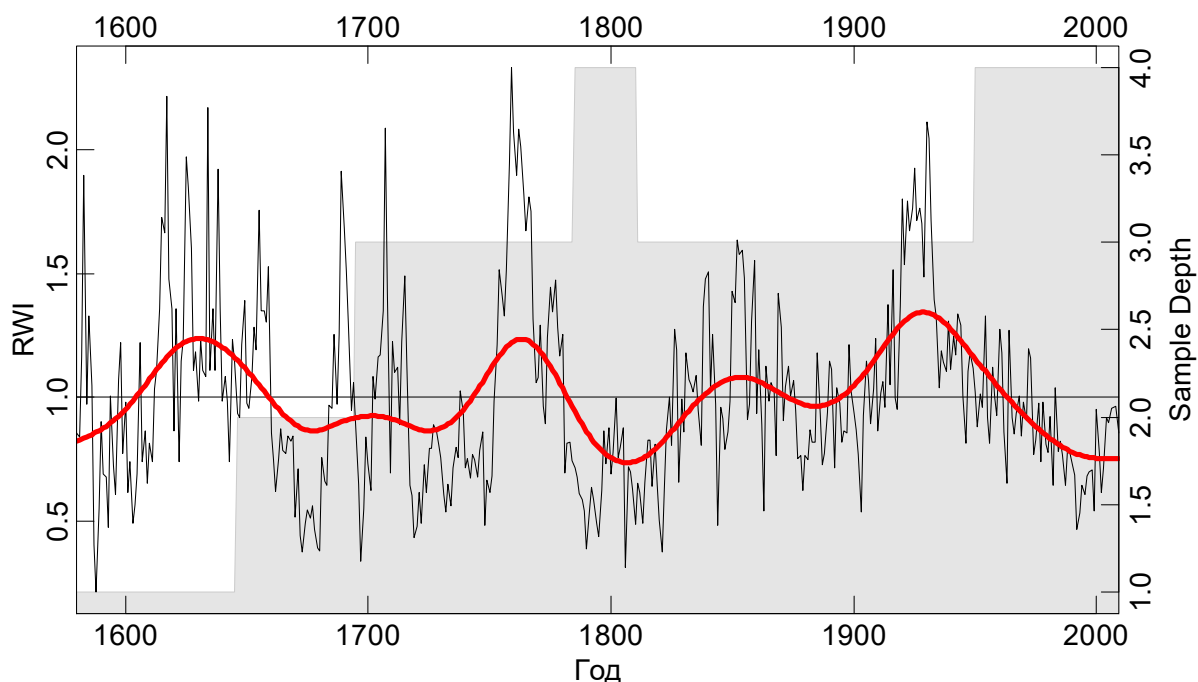


Рис. 9. График мастер-хронологии, построенный с помощью функции `crn.plot`. Черная линия – мастер-хронология, утолщенная красная линия – сглаживающий сплайн, серая заливка – число кернов в основе хронологии

### **`spag.plot()`**

Данная функция строит график, на котором изображены поведения всех временных серий в наборе данных (всех серий ширины колец, все индексы прироста и т.д.). Для нанесения всех кривых на один график каждый временной ряд сначала центрируется путем вычитания среднего значения столбца. Данный график удобно использовать, чтобы визуальнo оценить синхронность и согласованность радиального роста нескольких деревьев сразу. Синтаксис функции дан ниже, описание её аргументов см. табл. 13:

*`spag.plot(rwl, zfac = 1, useRaster = FALSE, res = 150, ...)`*

Таблица 13

Аргументы функции `spag.plot()`

<code>rwl</code>	Таблица данных типа <code>data.frame</code> , где в каждой колонке находится временная серия, а строки соответствуют годам. Данная таблица может быть получена например путем чтения данных с помощью функции <code>read.rwl</code> , или после вычисления индексов прироста с помощью <code>detrend()</code> .
<code>zfac</code>	Множитель для <code>rwl</code> . Позволяет повысить четкость при построении графика (растянуть или сплюснуть по вертикали линии временных серий).

useRaster	Логическая константа. Если useRaster = TRUE, то строится растровое изображение, а не векторное.
res	Числовой вектор длины 1 (из одного элемента, одного числа). Разрешение (пиксели на дюйм) растрового изображения.

### Пример 18

Построим график «спагетти» ширины годичных колец для набора данных co021. Инициализируем данные и построим два варианта графика (рис. 10 А, рис. 10 Б):

```
> data(co021)
> spag.plot(co021, zfac = 1)
> spag.plot(co021, zfac = 2)
```

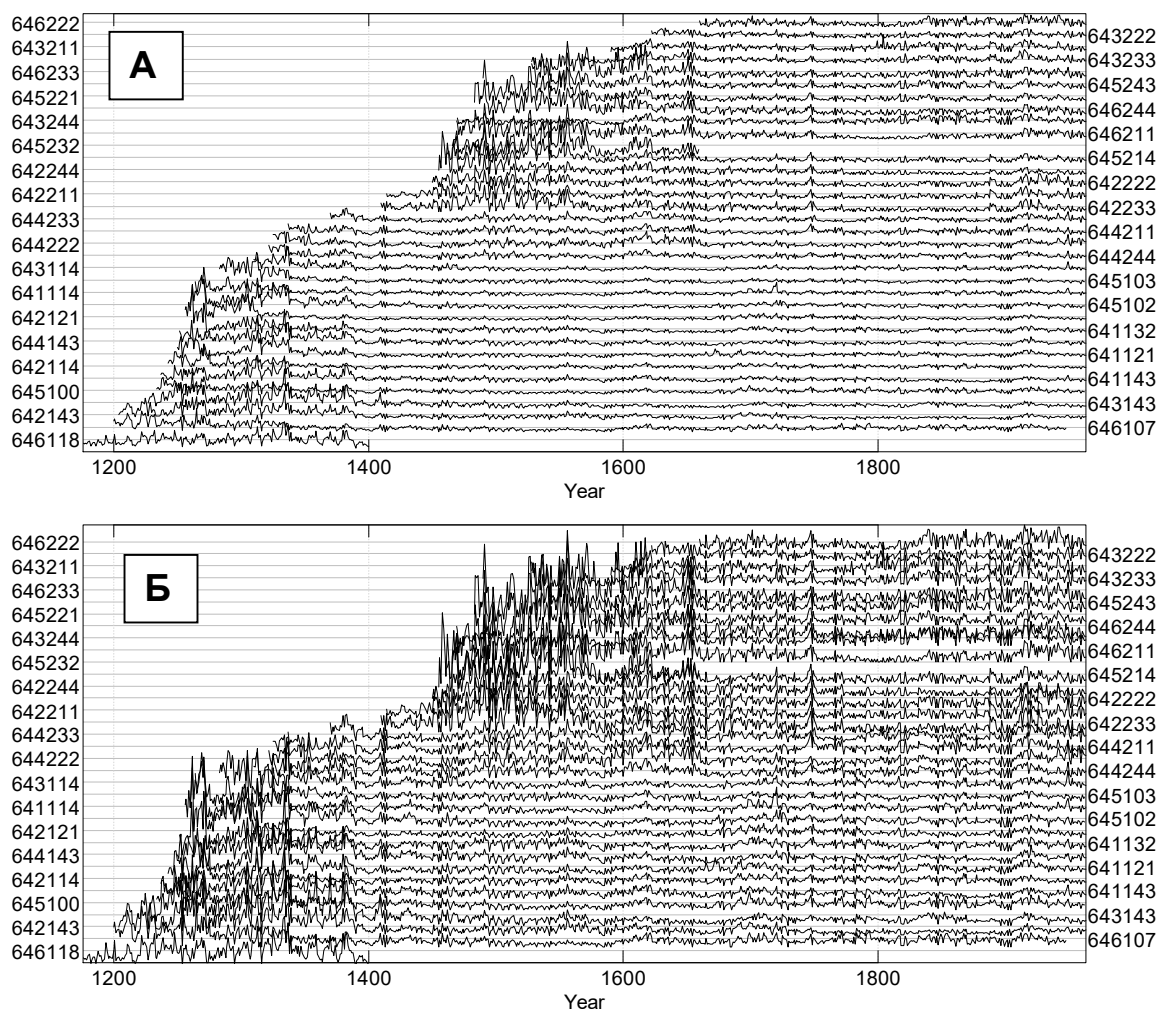


Рис. 10. График, построенный с помощью функции spag.plot.  
А: аргумент zfac = 1. Б: zfac = 2

## УПРАЖНЕНИЕ. АНАЛИЗ РАДИАЛЬНОГО РОСТА СОСНЫ

### Исходные данные

Данные, взятые для упражнения, являются частью базы данных Dendrochron (Тишин, Чижикова, 2013). Идентификатор местообитания в базе: 136, координаты: 33.830635 в.д., 66.345592 с.ш. Керны отобраны в 2009 г. на острове Сидоров, Керетский архипелаг, Карелия. Тип местообитания: торфяное болото, сфагновый сосняк.

Данные помещены в файл с названием 01.rwl.

### Схема анализа

Стандартная схема анализа данных о радиальном приросте деревьев включает следующие действия:

1. Загрузка данных об отобранных кернах.
2. Анализ индивидуальных временных серий (вычисление статистик серий и их визуализация).
3. Удаление возрастного тренда и стандартизация индивидуальных временных серий, получение индивидуальных индексов радиального прироста.
4. Анализ индивидуальных индексов радиального прироста (вычисление статистик серий и их визуализация).
5. Вычисление мастер-хронологии.
6. Визуализация и анализ мастер-хронологии.

### Анализ

*Этап 1. Загрузка данных об отобранных кернах*

В первую очередь, после открытия среды R загрузим библиотеку dplR:

```
> library(dplR)
```

Укажем путь к рабочей папке, где находится файл с данными (у читателей адрес может отличаться):

```
> setwd("C:/Users/Documents/")
```

Далее считаем данные о ширине колец. Используем для этого, к примеру, функцию read.rwl(). Известно, что файл был создан в соответствии

со стандартом tucson. Поместим полученный объект в переменную под названием **data1** формата rwl:

```
> data1 <- read.rwl(fname = "01.rwl",
                    format = "tucson")
There does not appear to be a header in the rwl file
There are 5 series
1  72      1785 2009 0.01
2  74      1646 2009 0.01
3  78      1695 2009 0.01
4  79      1950 2009 0.01
5  79A     1580 1810 0.01
```

Сообщение на экране консоли означает, что из файла было прочитано пять серий ширины годовых колец, или измерения пяти кернов. Названия серий 72, 74, 78, 79, 79A. Известно, что эти названия не включают в себя идентификатора или кода местообитания, два символа отданы для обозначения номера дерева, а один символ хранит идентификатор керна. Используем функцию `read.ids()`, чтобы выделить информацию о деревьях, и относящихся к ним кернам, сохраним информацию в таблицу `ids1` и выведем её на экран:

```
> ids1 <- read.ids(data1, stc = c(0, 2, 1))
> ids
      tree core
72     72    1
74     74    1
78     78    1
79     79    1
79A    79    2
```

Видно, что у дерева под номером 79 было измерено 2 керна.

*Этап 2. Анализ индивидуальных временных серий (вычисление статистик серий и их визуализация)*

Для визуализации пробелов в данных может быть полезен график сегментов, отражающий даты начала и конца каждой временной серии (рис. 11):

```
> seg.plot(data1)
```

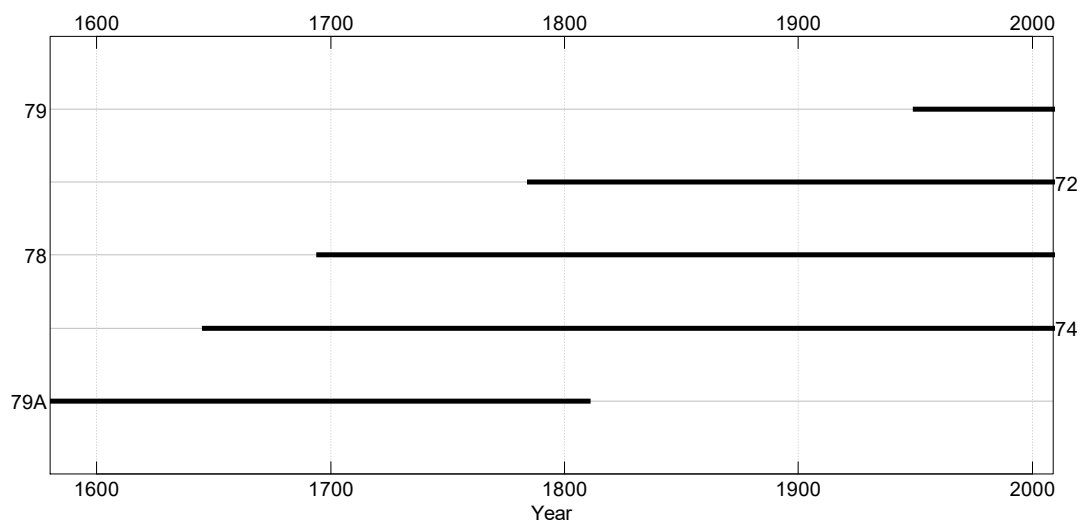


Рис. 11. График сегментов для набора данных 01.rwl (data1), построенный с помощью функции `seg.plot`

На рис. 11 видно, что у керна 79А отсутствует часть измерений после 1800 года, а керн 79 является наиболее коротким и покрывает порядка 50 лет после 1950 года. Керна 72, 78 и 79 имеют сравнимую длину и покрывают близкие временные отрезки.

Можно визуализировать все индивидуальные хронологии (рис. 12). Эту функцию выполняет команда `spag.plot()`. Для повышения выраженности вариации линий на графике зададим множитель `zfac` равный 1:

```
> spag.plot(data1, zfac = 1)
```

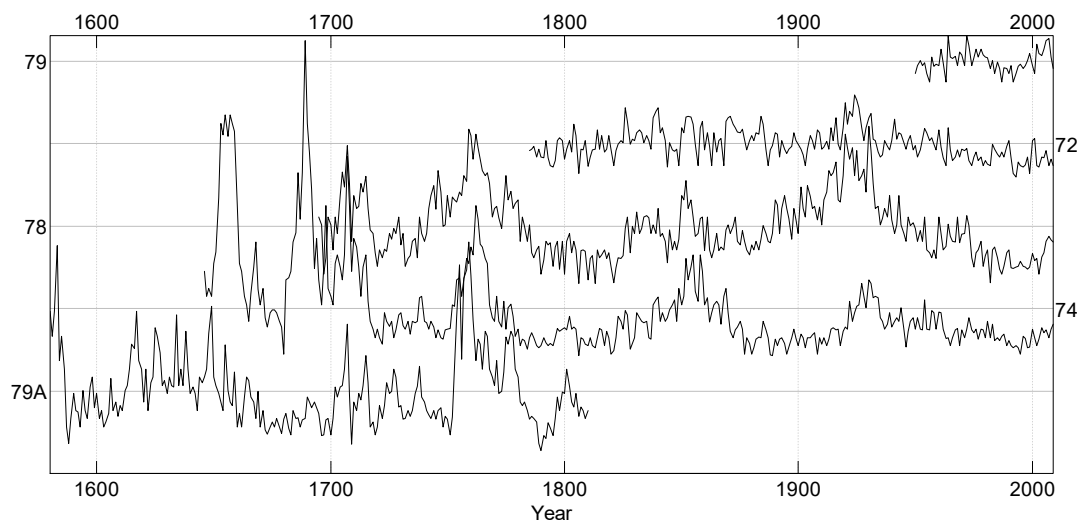


Рис. 12. График «спагетти» индивидуальных хронологий ширины годичных колец сосны обыкновенной, построенный с помощью функции `spag.plot()`

На рис. 12 видны синхронные «всплески» радиального роста, которые прослеживаются у всех хронологий: в районе 1760 года, 1850 года, и в

районе 1940 года. Они могут быть связаны как с вариациями климатом, так и с периодическим выпадением конкурентов внутри сообщества и высвобождением ресурсов для роста.

Вычислим описательные статистики хронологий ширины древесных колец с помощью `rw1.stats()`:

```
> rw1.stats(data1)
  series first last year  mean median stdev  skew  gini  ar1
1      72  1785 2009   225 0.565   0.54 0.188 0.472 0.186 0.581
2      74  1646 2009   364 0.753   0.57 0.565 2.456 0.346 0.891
3      78  1695 2009   315 0.886   0.80 0.385 0.773 0.240 0.810
4      79  1950 2009    60 0.450   0.43 0.132 0.314 0.162 0.291
5     79A  1580 1810   231 0.834   0.70 0.446 1.439 0.279 0.736
```

Выведенная на экран информация содержит годы первого и последнего измеренного кольца в хронологиях, длительности хронологий, так хронология 79 является самой короткой (60 лет), а хронология 74 самой длинной (364 года). Средний и медианный размер колец менее 1 мм, что характерно для роста сосны на обедненных субстратах (сфагновых болот). Коэффициент Джини близок к 0, что свидетельствует о достаточно однородном распределении размеров колец. Для всех хронологий, за исключением 79, свойственна высокая автокорреляция первого порядка (`ar1`), т. е. размер колец текущего года значительно связан с размером колец предыдущего года.

*Этап 3. Удаление возрастного тренда и стандартизация индивидуальных временных серий, получение индивидуальных индексов радиального прироста*

Процесс удаления тренда можно осуществить интерактивно, с помощью функции `i.detrend.series()`, обрабатывая по очереди по одной хронологии и выбирая для каждого случая подходящий метод удаления тренда, или в автоматическом режиме с помощью функции `detrend()` сразу для всех хронологий, выбрав один метод для всех. Выберем функцию `detrend()` и метод `Friedman super smoother`:

```
> rwil <- detrend(rwl = data1, method = "Friedman")
```

*Этап 4. Вычисление статистик индивидуальных индексов радиального прироста*

Визуализируем с помощью графика «спагетти» полученные на предыдущем этапе индивидуальные индексы радиального роста (рис. 13):

```
> spag.plot(rwil)
```

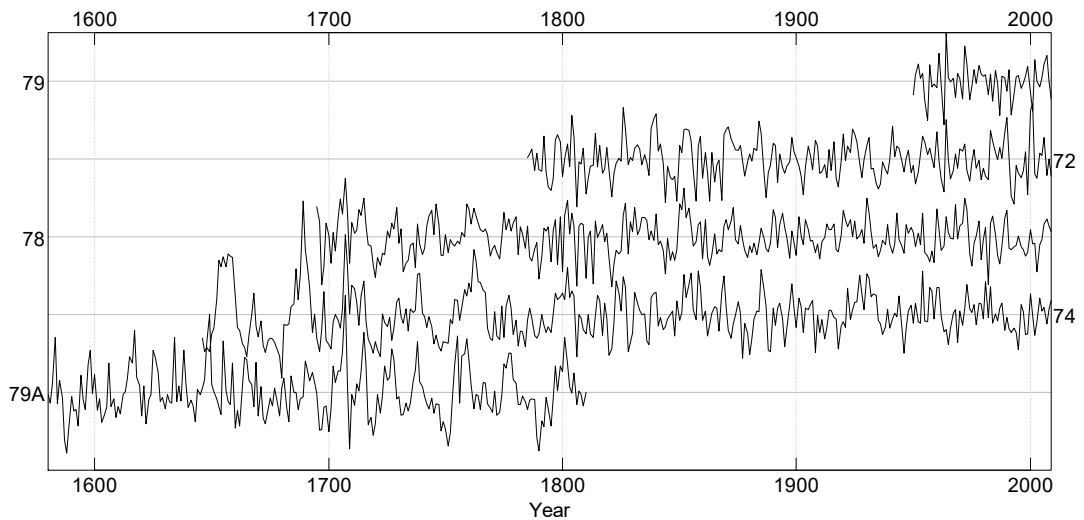


Рис. 13. График «спагетти» индексов радиального роста сосны обыкновенной, построенный с помощью функции `spag.plot()`

После удаления тренда синхронность радиального роста деревьев уже практически незаметна (рис. 13).

Для оценки синхронности роста вычислим корреляцию каждого индекса с мастер-хронологией (рис. 14):

```
> corrwil <- corr.rwl.seg(rwil, seg.length = 50)
```

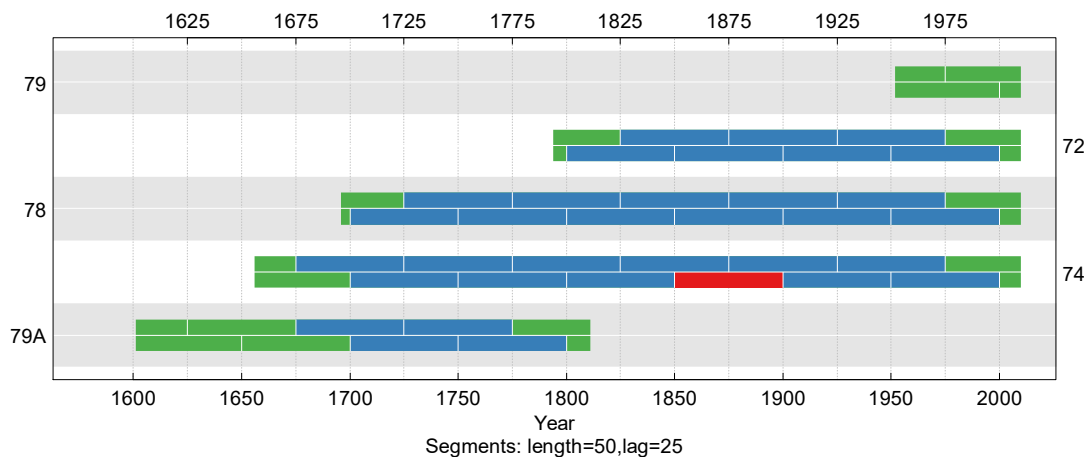


Рис. 14. График корреляции между каждым кольцевым рядом деревьев из набора данных 01.rwl (`data1`), построенный с помощью функции `corr.rwl.seg()`

На графике (рис. 14) преобладают синие сегменты – это те временные периоды, во время которых индексы роста значимо коррелируют с мастер-



хронологией. Зеленые сегменты – это те, которые не перекрываются полностью. У 74 керна имеется период, когда корреляция с мастер-хронологией низкая (красный сегмент на рис. 14). Обычно красные сегменты, иллюстрирующие корреляцию ниже заданного пользователем критического значения, могут указывать на проблему с датировкой, но не в этом случае. Проблема с датировкой в середине хронологии (неустановленные выпавшие или ложные кольца) неизбежно привели бы к проблемам и во всех более ранних сегментах хронологии, однако этого не наблюдается. В данном примере «красный» сегмент один и находится в середине ряда. Вероятно, в этот временной период что-то могло повлиять на рост дерева 74, не затронув деревья 78 и 72.

Вычислим некоторые статистики для индексов прироста:

```
> rwl.stats(rwil)
series first last year mean median stdev skew gini arl
1 72 1785 2009 225 1.003 0.986 0.253 0.122 0.142 0.267
2 74 1646 2009 364 0.976 0.952 0.307 0.687 0.174 0.569
3 78 1695 2009 315 0.989 0.984 0.238 -0.053 0.136 0.286
4 79 1950 2009 60 0.996 1.034 0.223 -0.138 0.121 -0.181
5 79A 1580 1810 231 0.982 0.930 0.327 0.564 0.184 0.456
```

Получим подробный отчет с оценками согласованности индексов прироста и качестве будущей мастер-хронологии в различные временные периоды:

```

> rwi.stats.running(rwil, ids = ids1, round.decimals = 2)
  start.year mid.year end.year n.cores n.trees      n n.tot n.wt n.bt rbar.tot rbar.wt rbar.bt c.eff rbar.eff  eps  snr
1      1635   1659   1684      2      2 1.78      1  0      1   -0.09      NA   -0.09      1   -0.09 -0.17 -0.15
2      1660   1684   1709      3      3 2.30      1  0      1    0.27      NA    0.27      1    0.27 0.47 0.87
3      1685   1709   1734      3      3 2.80      3  0      3    0.46      NA    0.46      1    0.46 0.70 2.36
4      1710   1734   1759      3      3 3.00      3  0      3    0.44      NA    0.44      1    0.44 0.70 2.39
5      1735   1759   1784      3      3 3.00      3  0      3    0.46      NA    0.46      1    0.46 0.72 2.53
6      1760   1784   1809      4      4 3.50      3  0      3    0.53      NA    0.53      1    0.53 0.80 3.98
7      1785   1809   1834      4      4 3.52      3  0      3    0.51      NA    0.51      1    0.51 0.79 3.70
8      1810   1834   1859      4      4 3.02      3  0      3    0.58      NA    0.58      1    0.58 0.81 4.22
9      1835   1859   1884      3      3 3.00      3  0      3    0.46      NA    0.46      1    0.46 0.72 2.59
10     1860   1884   1909      3      3 3.00      3  0      3    0.23      NA    0.23      1    0.23 0.48 0.91
11     1885   1909   1934      3      3 3.00      3  0      3    0.30      NA    0.30      1    0.30 0.56 1.26
12     1910   1934   1959      4      4 3.20      3  0      3    0.30      NA    0.30      1    0.30 0.58 1.40
13     1935   1959   1984      4      4 3.70      6  0      6    0.26      NA    0.26      1    0.26 0.56 1.28
14     1960   1984   2009      4      4 4.00      6  0      6    0.22      NA    0.22      1    0.22 0.54 1.16

```

Межсерийные корреляции (rbar) характеризуются малыми значениями в современные отрезки времени (1860-2009 годы) и в наиболее дальние годы (1635-1684). Коэффициент EPS ни в одном временном периоде не достигает значения, превышающего 0.85, что свидетельствует о том, что во все периоды времени не наблюдается синхронности, достаточной для получения качественной мастер-хронологии.

### *Этап 5. Вычисление мастер-хронологии*

Для вычисления мастер-хронологии можно воспользоваться двумя функциями. Функция `chron()` принимает на входе индексы прироста и использует весь набор данных, который подается, чтобы вычислить мастер-хронологию. Функция `strip.rwl` принимает на вход измерения ширины годовых колец, и пытается убрать те серии данных, которые понижают чувствительность вычисляемой хронологии, максимизируя EPS итоговой мастер-хронологии. Результатом функции `strip.rwl` не является хронология радиального прироста, хотя она вычисляется во время её работы как промежуточный результат. Результат функции `strip.rwl` – набор серий, которые можно использовать для построения мастер-хронологии.

Помимо результата на выходе, отличие сравниваемых функций – это способ удаления биологического тренда, который может быть применен к данным при вычислении индексов роста, которые идут в основу для мастер-хронологии. Т.к. функция `chron` принимает на входе индексы прироста, которые обычно вычисляют с помощью функции `detrend`, которая может использовать различные функции (сплайн, отрицательная экспонента, «сглаживатель» Фридмана и др.) для удаления тренда биологического роста. Функция `strip.rwl` принимает на вход неизменные измерения ширины годовых колец и использует единственный способ удаления биологического тренда при расчете мастер-хронологии: сплайн. Второе отличие функций, это возможность вычислять хронологию с помощью робастного среднего у функции `chron` есть, чего нет у функции `strip.rwl`. Третье отличие – это возможность вычислять остаточную хронологию функцией `chron` с помощью вычитания автокорреляции, что невозможно сделать с помощью функции `strip.rwl`. Тем не менее, учёт автокорреляции имеет смысл в том случае, если она высока (способ оценки – см. описание функции `rwl.stats` и Приложение), и, если есть убеждение, что состояние дерева в предыдущие годы влияет на размер колец в последующие годы, и эта зависимость в данных нежелательна. Учтите, что при вычислении индексов роста уже предпринимается попытка

вычесть из данных биологические тренды роста с помощью различных функций для того, чтобы полученные индексы отражали в большей степени климатический сигнал, а не закономерности биологического роста, поэтому вопрос выбеливать ли хронологию, вычитая автокорреляционную составляющую, оставляется на усмотрение исследователя.

В учебных целях воспользуемся обеими функциями `chron` и `strip.rwl` и сравним результат.

### Построение хронологии с помощью функции `chron`

Построим робастную выбеленную (остаточную) мастер-хронологию с помощью функции `chron`, используем индексы радиального роста `rwil`, полученные в предыдущем этапе. Укажем аргументы для выбеливания хронологии и вычисления робастного среднего:

```
> rwi.crn1 <- chron(rwil, ids = ids1,
                   prewhiten = TRUE,
                   biweight = TRUE,
                   prefix = "CH1")
```

Созданный объект содежит три колонки:

```
> str(rwi.crn1)
'data.frame':  430 obs. of  3 variables:
 $ CH1std      : num  0.958 0.849 1.088 1.687 0.84 ...
 $ CH1res      : num  NA NA NA NA NA NA NA NA NA ...
 $ samp.depth : num  1 1 1 1 1 1 1 1 1 1 ...
```

В первой колонке `$CH1std` находится хронология, вычисленная как робастное среднее, а во второй колонке `$CH1res` – выбеленная (остаточная) хронология. Можно изобразить обе хронологии на одном рисунке с помощью команды `plot.crn(rwi.crn1)`, но для учебных целей нарисуем каждую хронологию с помощью отдельной команды. Изобразим хронологию из первой колонки:

```
> plot.crn(rwi.crn1[, -2], add.spline = TRUE,
           crn.line.col='black',
           spline.line.col='red', spline.lwd=2,
           xlab = "Год", ylab = "RWI",
           main = colnames(rwi.crn1)[1])
```

Изобразим хронологию из второй колонки колонки:

```
> plot.crn(rwi.crn1[, -1], add.spline = TRUE,
           crn.line.col='darkblue',
           spline.line.col='red', spline.lwd=2,
           xlab = "Год", ylab = "RWI",
           main = colnames(rwi.crn1)[2])
```

### Предварительный отбор данных для хронологии с помощью strip.rwl

Используем функцию `strip.rwl`, чтобы получить ещё набор измерений ширины годовых колец, позволяющий максимизировать EPS. Используем данные о ширине годовых колец, сохраним данные в новую переменную таблицу **data2**:

```
> data2 <- strip.rwl(data1, ids = ids1,
                    verbose = TRUE)

REMOVE -- Iteration 1: Initial EPS: 0.527
Leave-one-out EPS:
72: 0.452
74: 0.478
78: 0.334
79: 0.61 *
79A: 0.447
***

REMOVE -- Iteration 1: leaving series 79 out.
EPS improved from 0.527 to 0.61.

REMOVE -- Iteration 2: Initial EPS: 0.61
Leave-one-out EPS:
72: 0.513
74: 0.596
78: 0.482
79A: 0.571
***

REMOVE -- Iteration 2: no improvement of EPS.
Aborting...
REINSERT -- Initial EPS: 0.61
Back-in EPS:
79: 0.527
REINSERT -- Iteration 1: no improvement of EPS.
Aborting...
```

Сообщения на экране содержат следующую информацию. Стартовый EPS хронологии, вычисленной по всем временным сериям, равен 0.527. Далее предпринимается попытка удалить одну серию, и при удалении серии 79 EPS

улучшается до 0.61. На следующем этапе EPS 0.61 берется как точка отсчёта, происходит попытка удалить ещё одну серию, однако ни одно последующее удаление не позволяет улучшить EPS «REMOVE -- Iteration 2: no improvement of EPS. Aborting...». Следующий этап – предпринимается попытка добавить единственно возможную серию, которая была уже удалена, но это не помогает увеличить EPS: «REINSERT -- Iteration 1: no improvement of EPS. Aborting...». На этом алгоритм завершает работу. Созданная таблица содержит меньше колонок, чем исходные данные. Выведем на экран структуру объекта **data2**, он не содержит колонки с серией 79:

```
> str(data2)
Classes 'rwl' and 'data.frame': 430 obs. of 4
variables:
 $ 72 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 74 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 78 : num NA NA NA NA NA NA NA NA NA NA NA ...
 $ 79A: num 1.8 1.5 1.8 2.6 1.2 1.5 1.1 0.4 0.2 ...
```

Удалим тренд из **data2**, используем метод “Spline”:

```
> rwi2 <- detrend(data2, method = "Spline")
```

Используем полученные индексы роста, чтобы вычислить хронологию:

```
> rwi.crn2 <- chron(rwi2, ids = ids1,
                   prewhiten = TRUE,
                   biweight = TRUE,
                   prefix = "CH2")
```

Построим графики хронологий (стандартная и остаточной) (рис. 16):

```
> plot.crn(rwi.crn2[, -2], add.spline = TRUE,
           crn.line.col='black',
           spline.line.col='red', spline.lwd=2,
           xlab = "Год", ylab = "RWI",
           main = colnames(rwi.crn2)[1])
> plot.crn(rwi.crn2[, -1], add.spline = TRUE,
           crn.line.col='darkblue',
           spline.line.col='red', spline.lwd=2,
           xlab = "Год", ylab = "RWI",
           main = colnames(rwi.crn2)[2])
```

Хронологии на рис. 15 и 16 внешне имеют отличия (особенно различаются стандартные хронологии CH1std и CH2std). Различия обусловлены способом удаления биологического тренда (на рис. 15 – сглаживатель Фридмана, на рис. 16 – сплайн), а также набором данных (на рис.

15 полный набор данных, на рис. 16 – исключена серия 79). Какие именно хронологии использовать в работе – выбор на усмотрение пользователя.

Сравним EPS наборов данных, использованных для хронология на рис. 15 и 16:

```
> rwi.stats(rwi1, ids1)$eps  
[1] 0.542  
> rwi.stats(rwi2, ids1)$eps  
[1] 0.696
```

Хронология CH2 на рис. 16, полученная с помощью сплайна и работы функции `strip.rwl`, имеет более высокий EPS, по сравнению с первой хронологией CH1, однако заметим, что его значение всё равно мало для того, чтобы считать хронологию качественной.

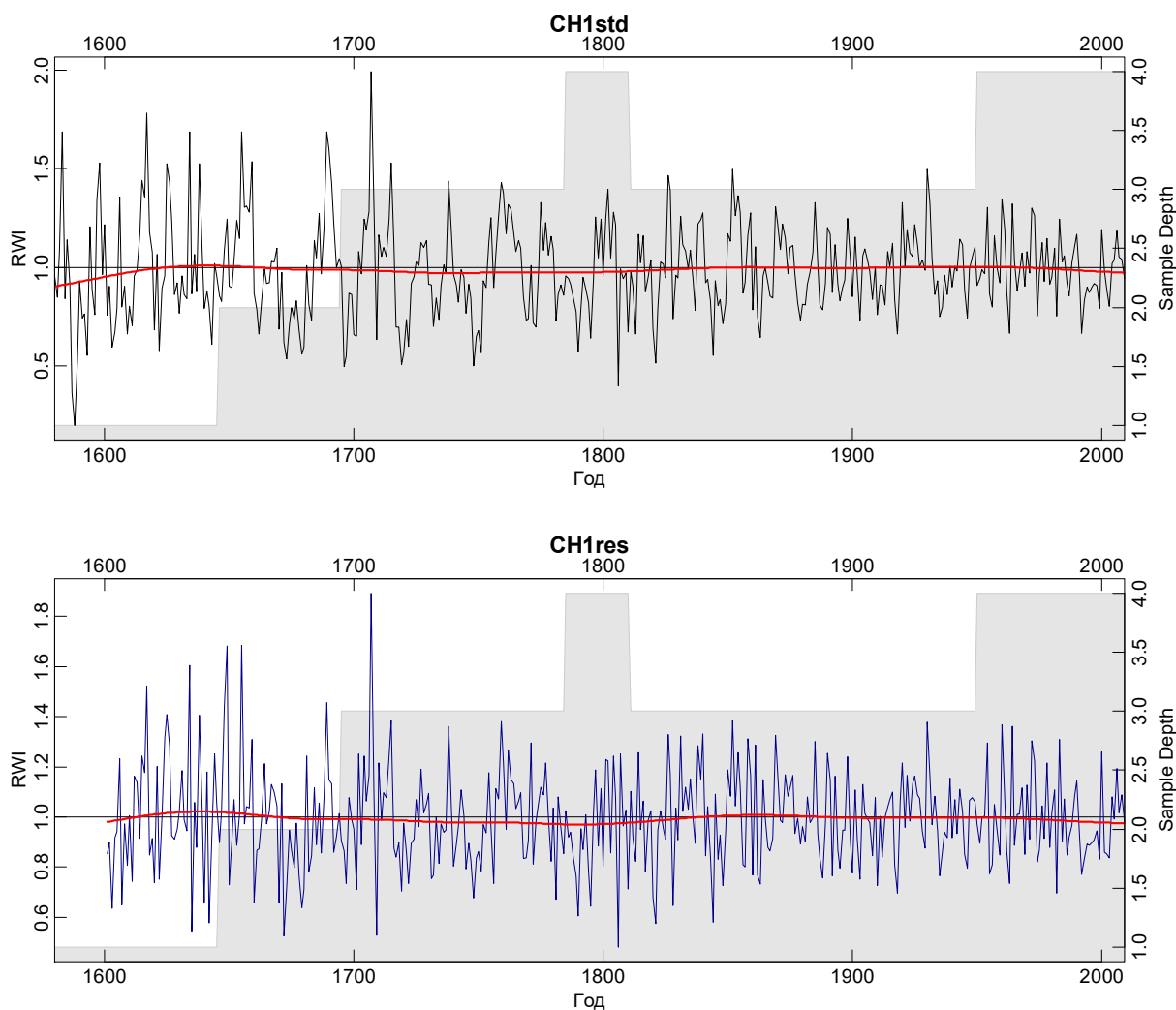


Рис. 15. Графики хронологии средних значений CH1std и остаточной хронологии CH1res набора данных о росте сосны обыкновенной 01.rwl. Красная линия – сглаженное среднее. Серая заливка – Sample depth – число индексов роста в основе хронологии

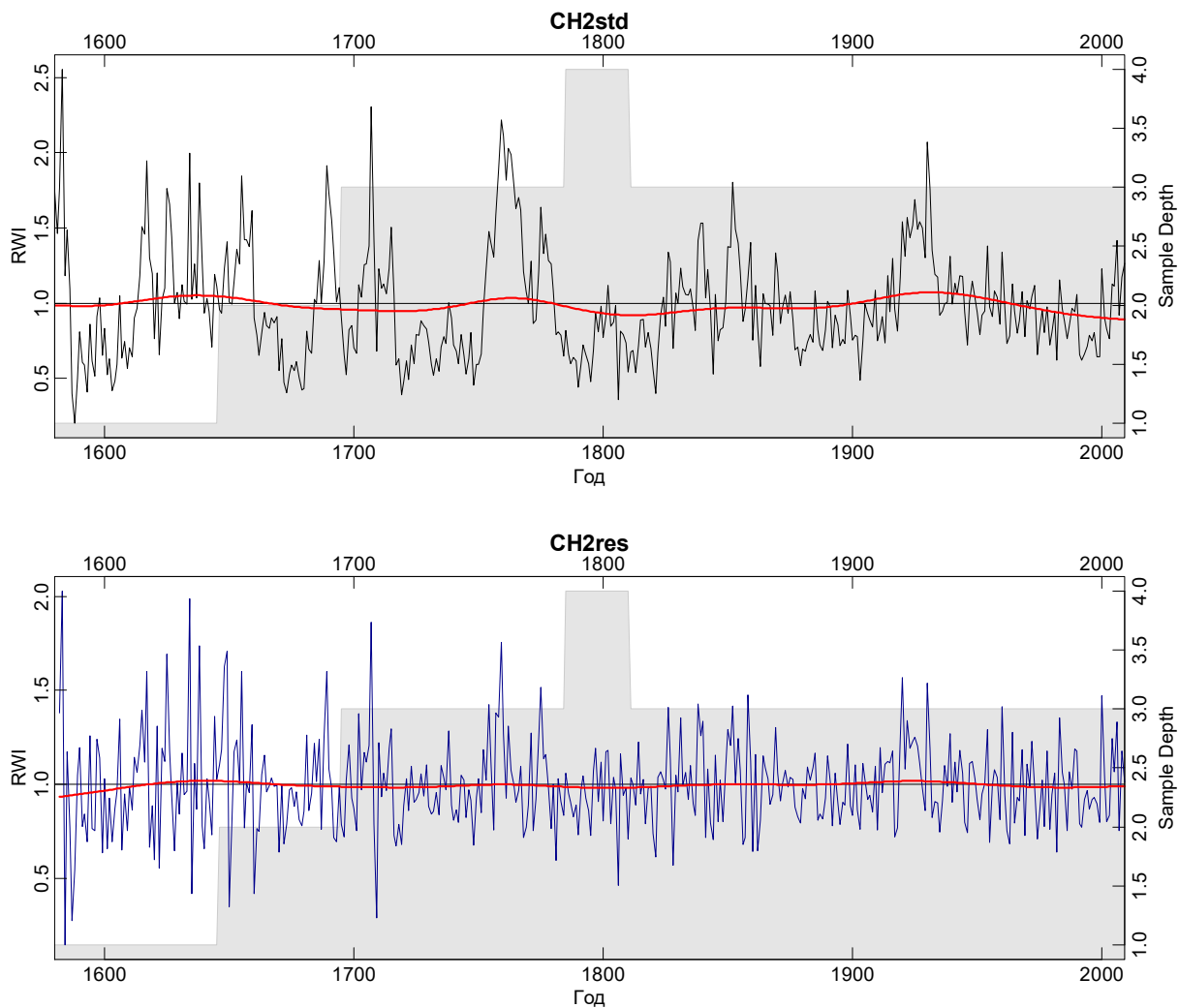


Рис. 16. Графики хронологии средних значений CH2std и остаточной хронологии CH2res набора данных о росте сосны обыкновенной 01.rwl. Красная линия – сглаженное среднее. Серая заливка – Sample depth – число индексов роста в основе хронологии



## СПИСОК ЛИТЕРАТУРЫ

1. Мастицкий С. Э., Шитиков В. К. Статистический анализ и визуализация данных с помощью R. – М.: ДМКПресс, 2015. – 496 с.
2. Тишин Д.В., Чижикова Н.А. Дендрохронология. – Казань: Казанский федеральный унт-т. – 2018. – 34 с.
3. Тишин Д.В., Чижикова Н.А. Свидетельство о государственной регистрации базы данных №2013621269 «DENDROCHRON». 2013.
4. Шипунов А. Б., Балдин Е. М., Волкова П. А., Коробейников А. И., Назарова С. А., Петров С. В., Суфиянов В. Г. Наглядная статистика. Используем R! Москва: ДМК Пресс. – 2017. – 298 с.
5. Шишов В.В., Тычков И.И., Попкова М.И. Методы анализа дендроклиматических данных и их применение для территории Сибири. – Красноярск: Сибирский федеральный ун-т. – 2015. – 210 с.
6. Шиятов С.Г., Ваганов Е.А., Кирдянов А.В., Круглов В.Б., Мазепа В.С., Наурзбаев М.М., Хантемиров Р.М. Методы дендрохронологии. Часть I. Основы дендрохронологии. Сбор и получение древесно–кольцевой информации: Учебно–методическое пособие. – Красноярск: КрасГУ. – 2000. – 80 с.
7. Bitvinskas T., Briffa K., Cook E., Downing D.J., Dubinskaite J., Eckstein D., Fritts H.C., Gordon G.A., Guiot J., Jones J. I. PH., Jagels R., Jones PH., Kairiukstis L., Kincaid W.B., Kocharov G., Kolishchuk V.G., Mazepa V.S., Mclaughlin S.B., Molenaar J., Nash III T.H., Norton D., Ogden J., PILCHER J., Robinson W.J., Schweingruber F., Serre-Bachet F., Shiyatov S.G., Telewski F., Tessier L., Vaganov E.A., Visser H., Worbes M. Methods of dendrochronology: applications in the environmental sciences. – Dordrecht: Kluwer Academic Publishers. – 394 p.
8. Bunn A.G. A Dendrochronology Program Library in R (dplR) // Dendrochronologia, 2008. – 26. – P. 115–124.
9. Bunn A.G. A Statistical and visual crossdating in R using the dplR library // Dendrochronologia, 2010. – 28 (4). – P. 251–258.
10. Bunn A.G., Korpela M. An introduction to dplR [Электронный ресурс]. – 2018. Точка доступа: <https://cran.r-hub.io/web/packages/dplR/vignettes/intro-dplR.pdf> (дата посещения 20.05.2022).
11. Bunn A.G., Korpela M. Crossdating in dplR [Электронный ресурс]. – 2018. Точка доступа: <https://cran.r-hub.io/web/packages/dplR/vignettes/xdate-dplR.pdf> (дата посещения 20.05.2022).
12. Bunn A.G., Korpela M. dplR Basics. Using R for Tree-Ring Analysis. [Электронный ресурс] – 2018. – Точка доступа: <https://rpubs.com/andybunn/dplR-basics> (дата посещения 20.05.2022).

13. Bunn A.G., Korpela M. Time Series Analysis in dplR. [Электронный ресурс]. – 2018. Точка доступа: <https://cran.r-hub.io/web/packages/dplR/vignettes/timeseries-dplR.pdf> (дата посещения 20.05.2022).
14. Bunn A.G., Korpela M. Biondi F., Campelo F., Merian P., Qeadan F. and Zang C. dplR: Dendrochronology Program Library in R. R package version 1.7.2. [Электронный ресурс]. – 2021. Точка доступа: <https://CRAN.R-project.org/package=dplR> (дата посещения 20.05.2022).
15. Cook E. R. A Time Series Analysis Approach to Tree–Ring Standardization: Ph. D. Dissertation. – Tucson, AZ: University of Arizona, 1985. – 171 p.
16. Cook E.R. The decomposition of tree-ring series for environmental studies // Tree-Ring Bulletin. – 1987. – 47. Pp. 37-59.
17. Cook E.R., Peters K. Calculating unbiased tree-ring indices for the study of climatic and environmental change // The Holocene. – 1997. – V. 7 (3) – Pp. 361-370.
18. Cook E.R., Peters K. The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies // Tree-Ring Bulletin. – 1981. – 41. – Pp. 45-53.
19. Fowler A. and Boswijk G. Chronology stripping as a tool for enhancing the statistical quality of tree-ring chronologies // Tree-Ring Research. – 2003. – V. 59 (2). – Pp. 53–62.
20. Graybill D. A. and LaMarche Jr. V. C. Campito Mountain Data Set. IGBP PAGES // World Data Center for Paleoclimatology Data Contribution Series 1983-CA533.RWL. – NOAA/NCDC Paleoclimatology Program, Boulder, Colorado, USA. – 1983.
21. R Core Team. R: A language and environment for statistical computing. [Электронный ресурс] – R Foundation for Statistical Computing, Vienna, Austria. – 2022. Точка доступа: <https://www.R-project.org/> (дата посещения 20.05.2022).
22. Schulman E. Schulman Old Tree No. 1 Data Set. IGBP PAGES //World Data Center for Paleoclimatology Data Contribution Series 1983-CO021.RWL. – NOAA/NCDC Paleoclimatology Program, Boulder, Colorado, USA. – 1963.
23. Zang C. Growth reaction of temperate forest tree species to summer drought – a multispecies tree-ring network approach. – Ph.D. thesis, Technische Universität München. – 2010. – 164 p.

## ПРИЛОЖЕНИЕ. Описание функции `ar()` пакета `stats`

Функция позволяет выбрать подходящую модель авторегрессии для временных рядов. Синтаксис функции, описание аргументов и результатов работы функции в таблице:

$$ar(x, aic=TRUE, order.max=NULL, method = c("yule-walker", "burg", "ols", "mle", "yw"), na.action, series)$$

Таблица  
Аргументы функции `ar()`

<code>x</code>	Одномерный или многомерный временной ряд.
<code>aic</code>	Логическая константа. Если <code>aic = TRUE</code> , тогда используется информационный критерий Акаике, который применяется для выбора порядка построения авторегрессионной модели. Если константа <code>aic = FALSE</code> , тогда используется порядок авторегрессии, указанный аргументом <code>order.max</code> .
<code>order.max</code>	Максимальный порядок (или порядок) модели, которая подгоняется к данным. По умолчанию используется меньшее из чисел $N-1$ и $10\log_{10}(N)$ , где $N$ -число не отсутствующих наблюдений, за исключением метода <code>"mle"</code> , когда <code>order.max</code> вычисляется как минимум из значений $N-1$ , $10\log_{10}(N)$ и числа 12.
<code>method</code>	Символьная строка, указывающая метод, соответствующий модели. Должна быть одной из строк: <code>"yule-walker"</code> , <code>"burg"</code> , <code>"ols"</code> , <code>"mle"</code> , <code>"yw"</code> (достаточно первых нескольких символов). По умолчанию используется вариант <code>"yule-walker"</code> .
<code>na.action</code>	Функция, вызываемая для обработки пропущенных значений. Возможные варианты значений: <code>na.omit</code> (исключить), <code>na.exclude</code> (исключить), <code>na.pass</code> (оставить). В настоящее время метод Yule-Walker может обрабатывать пропущенные значения в случае <code>na.action = na.pass</code> , однако пропущенные значения должны быть согласованы во времени: либо все переменные должны быть пропущены в момент времени, либо в данный момент времени не должно быть ни одного пропуска.
<code>demean</code>	Аргумент указывает на то, следует ли оценивать среднее значение.
<code>series</code>	Названия для серии значений. По умолчанию используется функция <code>deparse1(substitute(x))</code> .
<code>var.method</code>	Метод оценки дисперсии ошибок.
<code>object</code>	Объект, полученный функцией <code>ar()</code> .

newdata	Данные, к которым можно применить прогноз.
n.ahead	То количество шагов, на которое можно составить прогноз.
se.fit	Логическая константа, указывает нужно ли вычислять стандартные ошибки прогноза.

Уравнение авторегрессии имеет следующий вид:

$$X_t - \mu = a_1(x_{t-1} - \mu) + \dots + a_p(x_{t-p} - \mu) + e_t$$

Т.е. размер кольца в момент времени  $t$  зависит от размеров колец в предыдущие годы, умноженные на некие коэффициенты. И оценка этих коэффициентов и выбор числа лет  $p$  (максимального порядка, `order.max`) модели по имеющимся данным временной серии наблюдений, обычно входит в задачи построения авторегрессионной модели. Выбор порядка модели (максимальное число лагов, или шагов, на протяжении которых прослеживается автокорреляция) осуществляется с помощью критерия AIC, если аргумент `aic = TRUE`. При этом только метод `ml` может использовать этот критерий, т.к. выполняет оценку максимального правдоподобия.

Следует помнить, что функция `ar()` по умолчанию включает константу в модель, удаляя общее среднее значение серии  $x$  перед оценкой коэффициентов авторегрессионной модели.

Таблица 2

Выходные значения, создаваемые функцией `ar()`

order	Порядок модели. Выбирается путем минимизации AIC, если <code>AIC = TRUE</code> , в противном случае порядок задается самим пользователем как <code>order.max</code> (табл. 10).
ar	Расчетные коэффициенты авторегрессии для подогнанной модели.
var.pred	Дисперсия прогноза: оценка части дисперсии временного ряда, которая не объясняется авторегрессионной моделью.
x.mean	Оценка средней ряда, используемого при подгонке и для использования в прогнозировании.
x.intercept	Оценивается только в случае, если выбран метод <code>ar.ols</code> . Это свободный член ( <code>intercept</code> ) в модели ( $x$ – среднее).
aic	Различия в AIC между каждой моделью и лучшей моделью. Обратите внимание, что последняя может иметь AIC равный <code>-Inf</code> .

n.used	Количество наблюдений во временных рядах, в том числе отсутствующих.
n.obs	Число не пропущенных наблюдений во временном ряду.
order.max	Значение аргумента order.max.
partialacf	Оценка частичной автокорреляционной функции с точностью до порядка order.max.
resid	Остатки построенной модели, обусловленные наблюдениями первого порядка.
method	Значение аргумента метода.
series	Название временного ряда.
frequency	Частота временного ряда.
asy.var.coef	Матрица дисперсии асимптотической теории оценок коэффициентов. (Одномерный случай, порядок $> 0$ .)