



# *R Package AQEval*

---

## *AQEval*

*R package for change-point evaluation of air quality time-series data  
And de-seasonalisation, de-weathering, background subtraction  
(Developed in collaboration with Defra and IPSOS Mori)*

## *Methods*

*How AQEval works: Break-points, Break-segments and Signal Isolation*

## *Other Information*

*Code Examples: Package Installation and Standard Usage*

## *Beta Testing*

*We are looking for third-party feedback on the package prior to release  
(The insights that only come from fresh eyes and novel applications)  
We appreciate that we are asking to your time and your input  
But our objectives are to use this process to refine AQEval  
And, at the same time, get you early access...*

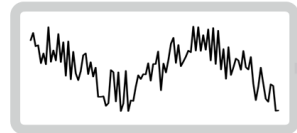


# AQEval: Overview

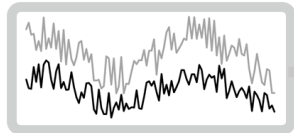
Improved non-seasonal event detection by local contribution isolation, and break-point and change-segment analysis of the local contribution

Input Data:

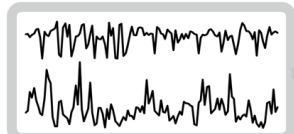
Intervention Site Air Quality



Background Air Quality



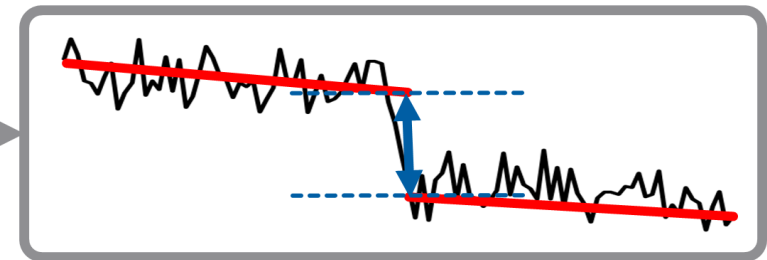
Local Meteorology



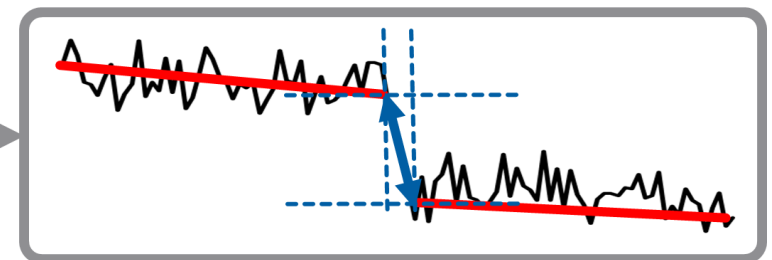
Local Contribution Isolation:



Local Break-Point Analysis:



Local Change-Segment Analysis:



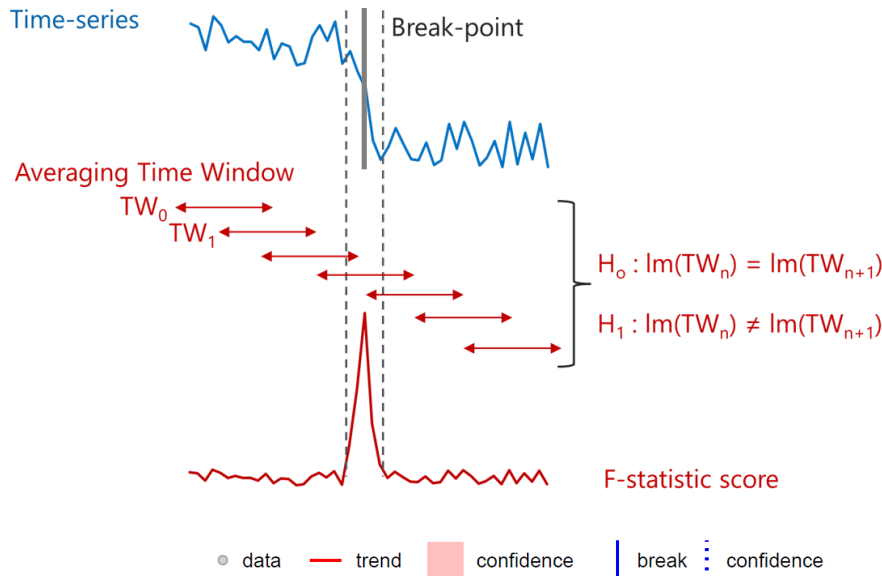
Methods for third party, so

- Accessible (high coverage) data sources
- Robust, evidence-driven and documented methods, and open software (R package)
- BUT also a need to refine methods, so some new elements, e.g. background correction, change-segment analysis...



# Break-point Detection

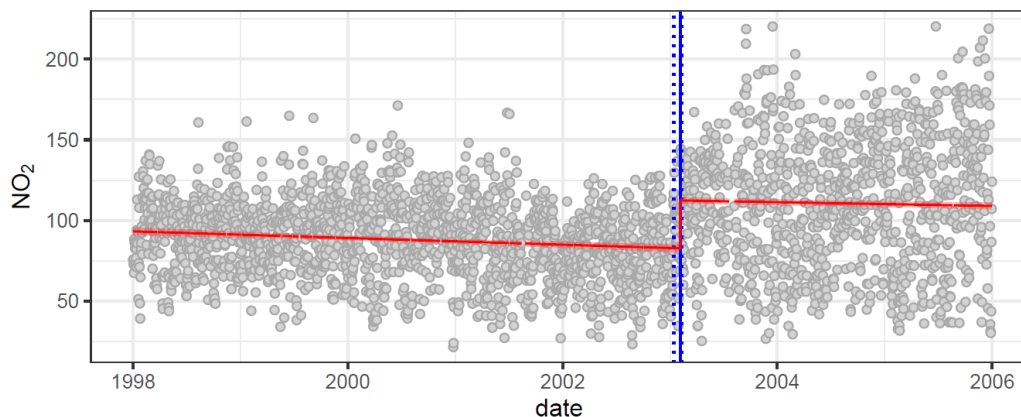
Break-point analysis and other similar change detection methods seek to find statistically significant step-changes in the properties of a time-series



Methods apply a rolling window strategy

- Assume first window is change-free
- Compare that and next window (here by f-statistic by linear models)
- And so on through time-series...
- Assigning point (or points) of change and associated confidence intervals

[REF: Zeileis et al, 2002, J. Stat. Software, 7(2)]



Application of break-point methods to London Marylebone 1998-2005 NO<sub>2</sub> time-series

Late OCT 2003 Break-point  
82.66→112.4; 29.69  $\mu\text{g.m}^{-3}$  (36%)

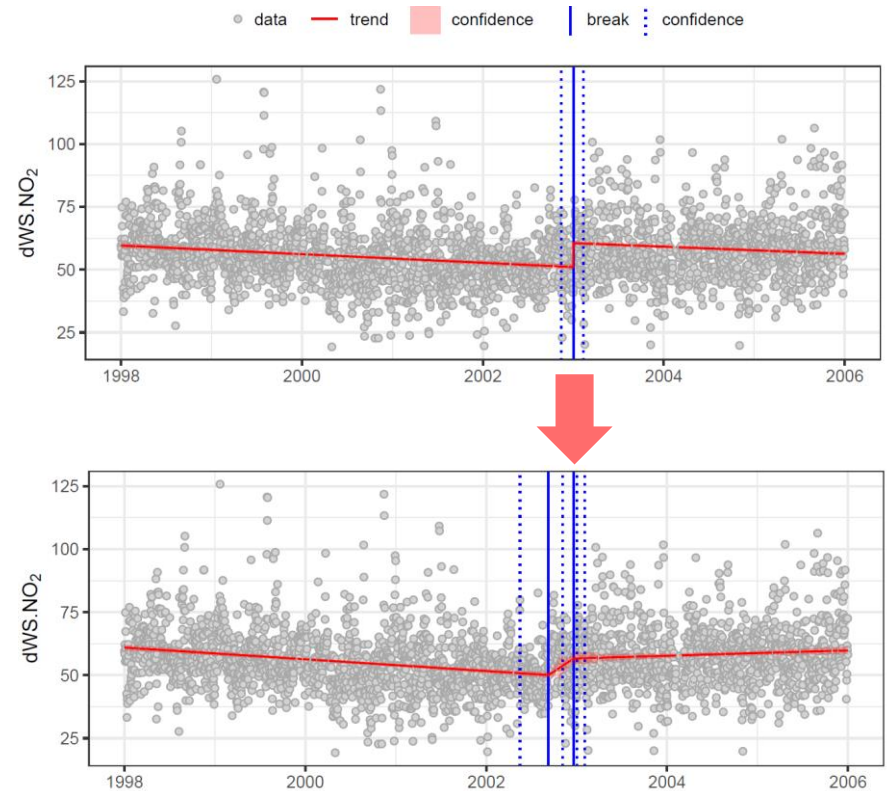
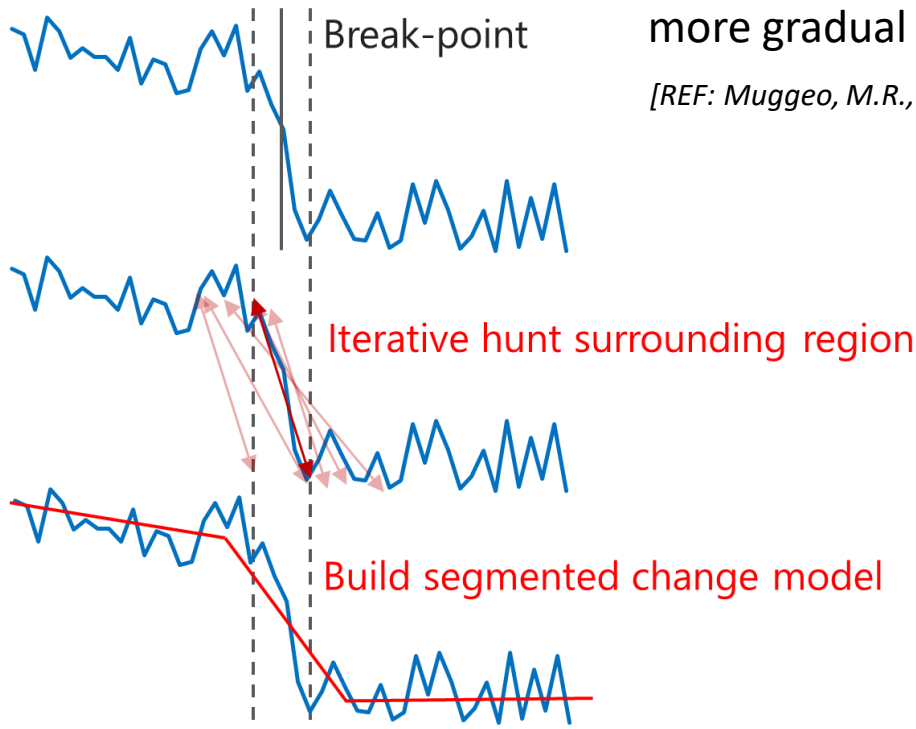


# Break-segment Fitting and Quantification

However, break-points are instantaneous changes and no real-world intervention is ever likely to deliver such an abrupt change...

So, we use the break regions as a starting point to test for more gradual change, and build a more realistic model

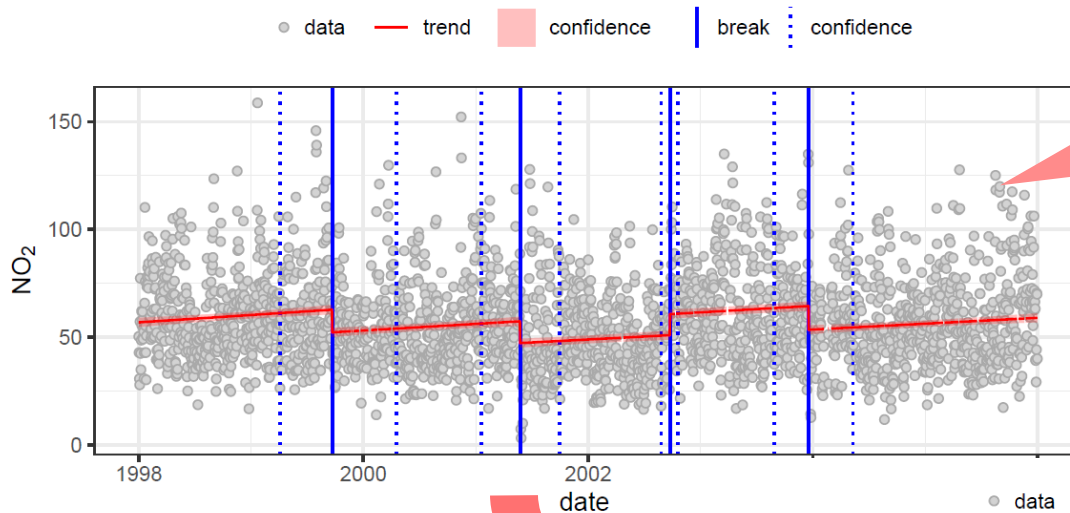
[REF: Muggeo, M.R., 2008. *R News*, 8/1, 20-25]





# *(Data Pre-processing) Signal Isolation*

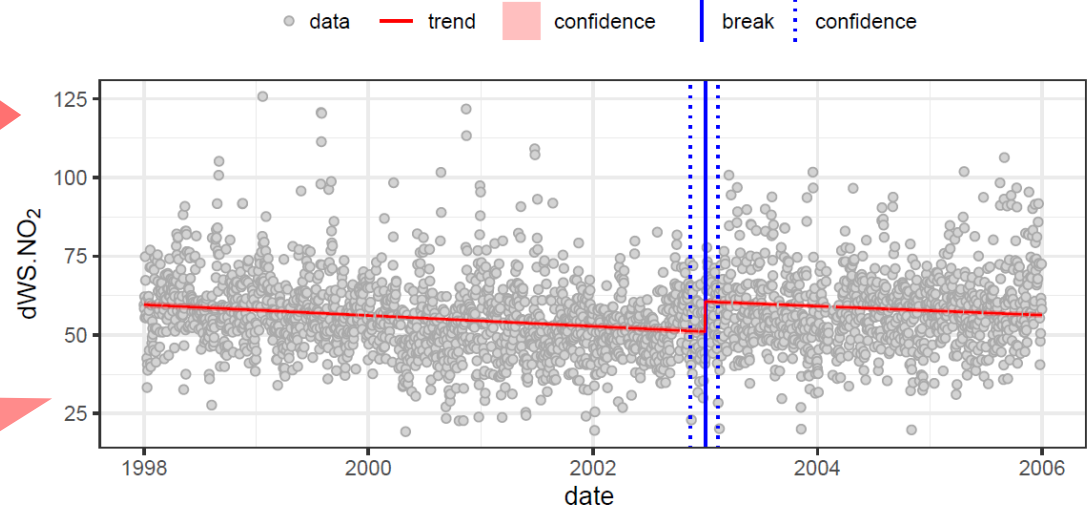
Multiple inputs often contribute to an air quality time-series (e.g. other sources, meteorology, etc) and any of these can produce change events...



Direct use with ambient data is often hindered by variance from other site contributions, e.g. weather, temperature, seasonal cycles, etc.

So, we also include functions for various signal isolations:

- deseasonalisation
- deweathering,
- background subtraction





# *AQEval Pre-release R Package Installation*

---

Once released, the intention is to make AQEval freely available under General Public License via on-line archives (CRAN and GitHub).

In the meantime, the pre-release version is shared as an 'tar ball' bundle for testing and evaluation as part of software development, and for informal review and feedback

When you want to use AQEval, load it like any other specialist packages:

## **Installation**

- For A bundled version of AQEval (AQEval\_0.1.3.tar.gz or later)
- Copy file to a known directory, run R (or RStudio) and install using the following:

### **In R (or RStudio):**

(if you do not have devtools package, install that first, then...)

```
devtools::install_local(file.choose())  
#and select the AQEval file
```

```
require(AQEval)
```

(Help documentation like other R packages)



## Basic AQEval Usage (1)

AQEval functions are coded in a similar style to openair (Carslaw & Ropkins, 2012), so the two can be easily used in combination

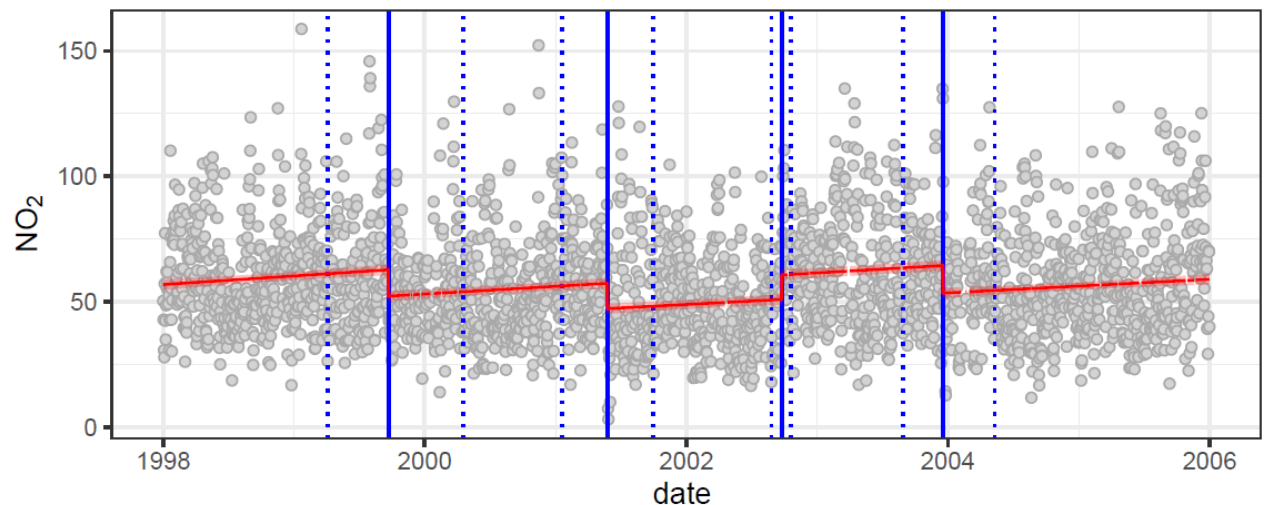
Example using 1998-2005 data from the London Ealing EA2 AURN site. Get data EA2 from the online KCL archive and convert to daily time-series, using openair:

```
require(openair)
ea.1998.2005 <- importKCL("ea2", year=1998:2005, met=TRUE)
ea.1998.2005.day <- timeAverage(ea.1998.2005, avg.time="day")
```

Apply AQEval function quantBreakPoints in form:

```
quantBreakPoints(ea.1998.2005.day, "no2", h=0.1)
```

The inputs are your data set, the time-series to be analysed and the break time-window to apply







## Basic AQEval Usage (2)

Deseasonaling and deweathering data before break-point testing

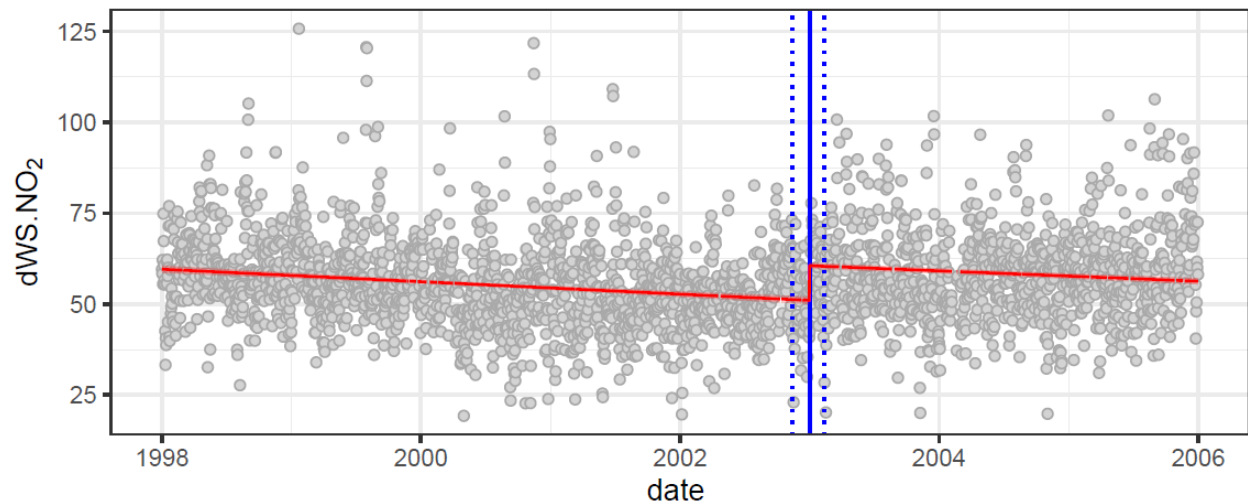
Using same data set and AQEval function isolateContribution:

```
ea.1998.2005$dWS.no2 <- isolateContribution(ea.1998.2005, "no2")
```

(By default, this builds and subtracts a model of wind speed/direction, hour-of-day and day-of-year, so applying it at 1-hour resolution, then handling the deweathered/deseasonalised, dWS, data as NO<sub>2</sub> was before)

```
ea.1998.2005.day <- timeAverage(ea.1998.2005, avg.time="day")  
quantBreakPoints(ea.1998.2005.day, "dWS.no2", h=0.1)
```

Note: This requires  
openair-friendly  
date (time stamp in  
date column, wind  
speed and direction  
as wd and ws)





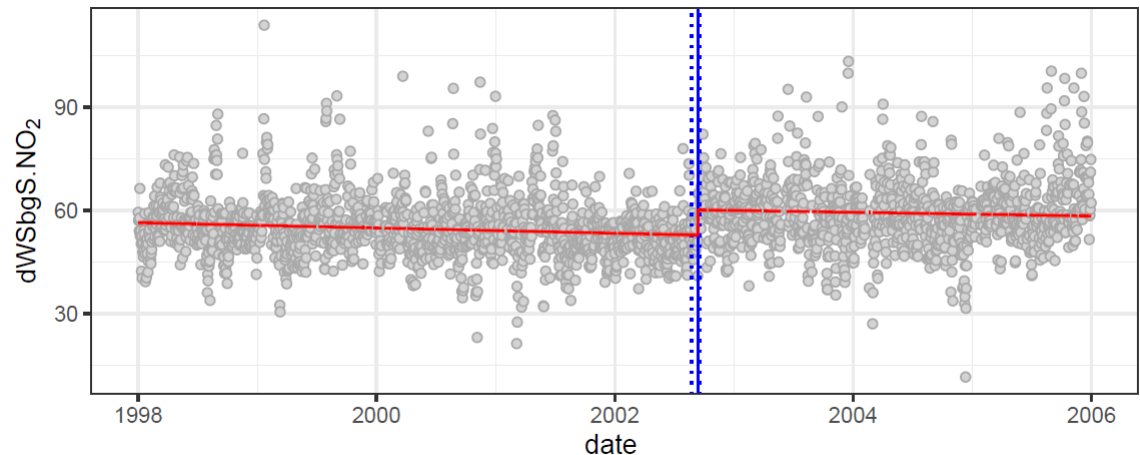


## Basic AQEval Usage (3)

### Expanding the local isolation model

Subtracting a local background, here London Kensington & Chelsea KC1, getting that from AURN (rather than KCL archive) and merging using dplyr before AQEval analysis:

```
require(dplyr)
kc.1998.2005 <- importAURN("kc1", year=1998:2005, meta=TRUE)
temp <- select(kc.1998.2005, date, no2) %>% rename(bg.no2=no2)
ea.1998.2005 <- left_join(ea.1998.2005, temp)
ea.1998.2005$dWSbgS.no2 <- isolateContribution(ea.1998.2005,
                                                "no2", background="bg.no2")
ea.1998.2005.day <- timeAverage(ea.1998.2005, avg.time="day")
quantBreakPoints(ea.1998.2005.day, "dWSbgS.no2", h=0.1)
```



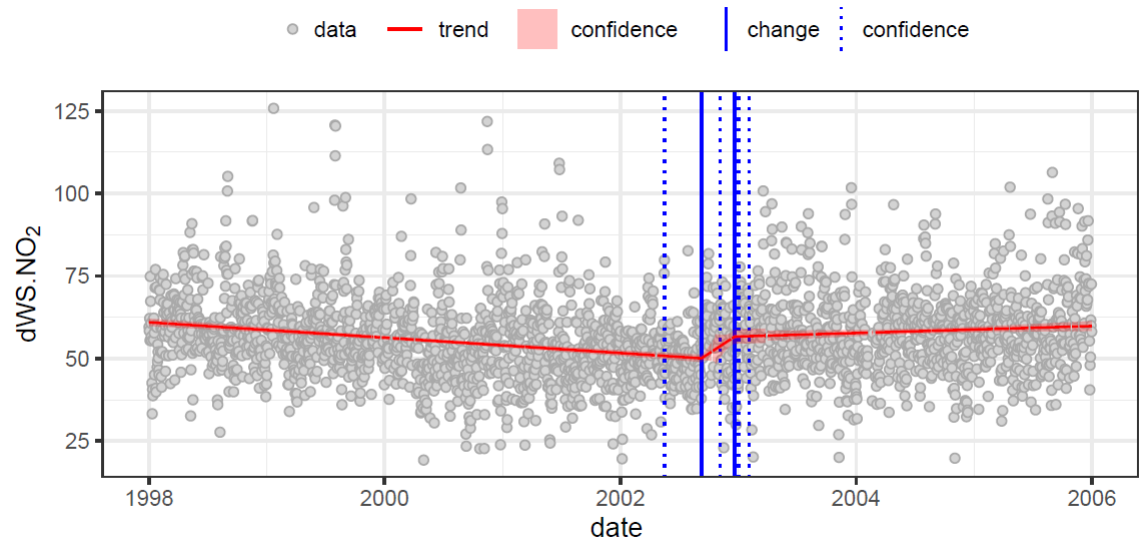


## Basic AQEval Usage (4)

And a break-segment instead of a break-point..?

Use `qauntBreakSegments` instead of `quantBreakPoints`:

```
quantBreakSegments(ea.1998.2005.day, "dWSbgS.no2", h=0.1)
```



(If you have the time, there is a lot more in the package, but hopefully that is enough to get you started...)