

#### Demonstrations of new tools for spectral data analysis & NIRS of Waters (treated with Magnetic Fields)

Bernhard Pollner, Zoltan Kovacs

Aquaphotomics: Understanding Water in the Biological World at the 5th Kobe University Brussels European Center Symposium 14th October 2014, Brussels, Belgium

### **Outline & Resume**

- <u>R-Package</u>,,<u>Aquaphotomics</u>"
- Experiment-design, Data-Import, Data-Analysis
- Prelim. Results:
  - distinguish between different water species
  - show effect of different magnetic treatments of water

#### Online "Aquagram-Maker"

- upload data to make Aquagram
- easy and fast

```
the universal input, checking for the class of the input-object
kPeaks <- function(ObjectToPickPeaks, bandwidth=25, comps=1:4, discrim=FALSE) {</pre>
                                                                                                                                                                                                                                                                                                         ### universal peak picker
if (class(ObjectToPickPeaks) == "mvr") {
               allColNames <- colnames(ObjectToPickPeaks$coefficients[,,1:Object
                                                                                                                                                                                                                                                                                                              ([מי
               lastName <- allColNames[length(allColNames)]</pre>
                                                                                                                                                                                                                                                                    omnonent -- We
               if (!is.character(lastName)) {
                                                                                                                                             ## problem if there is
                                                                                                                                                                                                                                                                                                                                    not get a name bac
                            lastName <- "1 comps"</pre>
              mat <- ObjectToPickPeaks$coefficients[,, ObjectToPickPea</pre>
               dfToPickPeaks <- data.frame(X=mat)</pre>
               colnames(dfToPickPeaks) <- lastName</pre>
 }
if (class(ObjectToPickPeaks) == "PCA") {
               dfToPickPeaks <- as.data.frame(ObjectToPickPeaks$loadings[,comps])</pre>
 }
if (class(ObjectToPickPeaks) == "data.frame") {
               dfToPickPeaks OhiectToPickPeaks
EOF Repair of the solution of 
needs the pick results object created by pickPeaks a
                                                                                                                                                                                                                                                          he vector that was used for pick
                                                                                                                                                                                                                                                                                                                                                                               hg !
if ( and ( in the subscript of the subsc
                                                                                                                                                                                                                                                                                                                                                                                 ####
a <- (colnames(pickResults$rawVector)</pre>
b <- substr(a, stngs$nrCharPrevWL+1, nchar(a)) ## to get rid of the "w" in front of the numbers</pre>
wave
res
colPus _____
colNeg <- stngs$colNegPeaks</pre>
positionTable <-res[1: (nrow(res)/2) ,]</pre>
heigthTable <-res[((nrow(res)/2)+1):nrow(res) , ]</pre>
Yrange <- range(pickResults$rawVector)[2] - range(pickResults$rawVector)[1]</pre>
onSub <- paste(onSub, ", bw=", pickResults$picks$pickWindow, sep="")</pre>
```



About R What is R? Contributors Screenshots What's new?

Download, Packages CRAN

R Project Foundation Members & Donors Mailing Lists Bug Tracking Developer Page Conferences Search

Documentation <u>Manuals</u> <u>FAQs</u> <u>The R Journal</u> <u>Wiki</u> <u>Books</u> <u>Certification</u> <u>Other</u>

Misc

Bioconductor Related Projects User Groups Links

#### The R Project for Statistical Computing



Getting Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To <u>download R</u>, please choose your preferred <u>CRAN</u> <u>mirror</u>.
- If you have questions about R like how to download and install the software, or what the license terms are, please read our <u>answers to frequently asked questions</u> before you send an email.

News:

- R version 3.1.1 (Sock it to Me) has been released on 2014-07-10.
- R version 3.0.3 (Warm Puppy) has been released on 2014-03-06.
- <u>The R Journal Vol.5/2</u> is available.
- useR! 2014, took place at the University of California, Los Angeles, USA June 30 July 3, 2014.
- useR! 2015, will take place at the University of Aalborg, Denmark, June 30 July 3, 2015.



# Why R-Project ?

## high level programming language

## • open source - freeware !

### powerful - extendable

cross-platform

possible: something really new

efore you send an email.

#### • widely used in academia

- sion 3.1.1 (Sock it to Me) has been released on 2014-07-10.
- R version 3.0.3 (Warm Puppy) has been released on 2014-03-06.
- useR! 2014, took place at the University of California, Los Angeles, USA June 30 July 3, 2014.

# R-Package "Aquaphotomics'

#### **Experiment Design**

- randomize samples
- facilitates time resolved experiments
- Data Import

#### • e.g. straight from binary (.da) file

 controlled source and structure for class- and numerical variables (pickResults, onMain="", onSub="", pcaVariances=NULL, customColor=NULL) {

#### Data Analysis

easy splitting and grouping of data

### parallelization of different data pre-treatments

#### tools to compare those parallel results

<pre>cPeaks &lt;- function(ObjectToPickPeaks, bandwidth=25, comps=1</pre>	
<pre>if (class(ObjectToPickPeaks) == "mvr" {</pre>	NDARt
allColNames <- colnames(ObjectToP	<pre>ettTwpickPeaks\$ncomp])</pre>
lastName (- allColNames[length(allColNames)]	
<pre>S C_ExpName : Factor w/ 12 levels "DZ1hSeq", "DZ12h",: 1 1 1 1</pre>	e is only one component we do not get a name ba
<pre>C_ExpClass : Factor w/ 3 levels "ImmCuv1","ImmCuv2",: 3 3</pre>	
Y_SampleNr : num 1111122222	
<pre>\$ C_Time : Factor w/ 17 levels "Seq1", "Seq2",: 1 1 1 1 1</pre>	eaks\$ncomp]
\$ C_ECRM : Factor w/ 2 levels "MQ","RM": 1 1 1 1 1 2 2 2 7 2	
<pre>\$ C_water : Factor w/ 2 levels "MQ", "water": 1 1 1 1 1 2 2 2 \$ C_water : Factor w/ 2 levels "MQ", "water": 1 1 1 1 1 2 2 2</pre>	<ul> <li>standardized data</li> </ul>
\$ C_waterNames: Factor w/ 6 levels "MQ", "StU", "OD",: 1 1 1 1 1 \$ C Repl : Factor w/ 13 levels "MO" "R1" "R2" · 1 1 1 1 1	Standar Gized Gata
<sup>1</sup> f \$ C_Group : Factor w/ 4 levels "Cont", "GDum",: 4 4 4 4 4	structure
\$ C_all : Factor w/ 1 level "all": 1 1 1 1 1 1 1 1 1 .	ngs[,comps])
\$ Y_Temp : num 24.2 24.2 24.2 24.2 24.2 23.4 23.4 23.4	
\$ C_SampleNr : Factor w/ 626 levels "1","10","1001",: 1 1 1	encapsulated and
\$ C_ConSNr : Factor w/ 7 levels "1","2","3","4",: 1 2 3 4	
<pre>\$ C_Temp : Factor w/ 8 levels "22","22.5","23",: 5 5 5 5 5 \$ C PolHum : Factor w/ 9 lovels "54" "57" "60" : 1 1 1 1 1 2</pre>	"transportable color-
\$ Time : POSIXct, format: "2014-07-09 09:14:51" "2014-07-0	discrim)
absTime : num 55695 55695 55696 55696 55697	coding
<pre>\$ chron : int 1 2 3 4 5 6 7 8 9 10 \$ rumPed</pre>	
	• nice to look at:
\$ C_ExpClass : num 3 3 3 3 3 3 3 3 3 3	peavariances=nole, castomeoror=NOLE) {
\$ C_Time : chr "blue4" "blue4" "blue4" "blue4"	view of NIR-data is
\$ C_ection : num 1111122222	
\$ C_waterNames: num 1111122222	minimized (but of course
\$ C_Repl : chr "blue4" "blue4" "blue4" "blue4"	set still possible) in front of the numbers
$$C_all$ : num 111111111	suil possible)
\$ C_SampleNr : chr "blue4" "blue4" "blue4" "blue4"	(regress on absTime)
ConSNr : num 1 2 3 4 5 1 2 3 4 5	(regress on abstime)
CO S RelHum : chr "#0000FF" "#0000FF" "#0000FF" "#0000FF" .	
NIR : num [1:3676, 1:4200] 0.0199 0.0201 0.0203 0.0204	
he attr(*, "dimnames")=List of 2	
\$ : chr "w400" "w400.5" "w401" "w401.5"	<pre>lts\$rawVector)[1]</pre>
onSub ke pasteronany, jow j prekkesoropperkesoprekwinde	w, sep="")

<pre>cs &lt;- function(ObjectToPickPeaks,</pre>	bandwidth=25	, comps=1:4, discrim	=FALSE) {	
<pre>(class(ObjectToPickPeaks) == "mvr</pre>	Jota	Analyci		
allColNames <- colnames(ObjectTo	Jala	<b>MIIAIYS</b>	colickPeaks\$nd	

- splitting / slicing / grouping of data
- recyclable, separate data-analysis procedure file
- strict separation of calculations and plotting
- customizable parallel processing (if applicable; on UNIX)
- various methods implemented

```
    and much more already on the "list"

                                                                                                         sdrc calc plot LDA
                                                                                                           groupAvgSubuaci
                                                                                                            plotAquagram
                                                                                                           2 plotCompScores
                                                                                                        sc e plotCompScores simple
                                                                                                        s c plotGroupAvgSubtract
                                                                                                        scrc plotPCAScores
                                                                                                        s rc plotRaw
                                                                                                        s rc plotRMSEx
                                                                                                         s rc plotSimcaClassDist
                                                                                                          e plotVector
                                                                                                        so c_plotVectorCollection
                                                                                                        scrc printSvmTable
                                                                                                        strc showGroups
                                                                                                         electWLs
                                                                                                          iowAllClassRepresentations
                                                                                                          iowSingleClassRepresentation
                                                                                                         <u>nca</u> mmp
                                                                                                         litSequence
                                                                                                         SSC
```

Calculate and / or plot a plsr Ida. Subtractions in raw-spectra Plot an Aquagram Plot comparison scores; advanced grouping Plot comparison scores; simple grouping Plot averaged subtracted spectra Plot PCA scores Plot Spectra Error / calibration / validation plots of PLSR models Plot SIMCA interclass distances (sdrc) Plot Vector Plot Vector Collection Classify using Support Vector Machines Show SDRC set by grouping Select Wavelengths Show the character representation of a factor Show the character representation of a factor Plot SIMCA interclass distances (dataset) Manually split within sequence according to class-membership he universal input, checking for the class of the input-object

Peaks <- function(ObjectToPickPeaks, bandwidth=25, comps=1:4, discrim=FALSE)</pre>

### Data Analysis - Implemented Methods

if (!is.character(lastName)) { ## problem if there is only one component -- we do not get a name bac

#### Data Pre-Treatment

- mat <- ObjectToPickPeaks\$coefficients[,, ObjectToPickPeaks\$ncomp]</pre>
- smoothing & derivatives, SNV, MSC, EMSC

#### • special Aquaphotomics data pretreatments (parallelized)

Evaluation / Analysis:

#### subtractions in

raw-spectra

• PCA, PLSR

SIMCA
(SVM, bootstrap, ...)
Aquagram !

sdra calc plat I DA rc groupAvgSubtract s rc plotAquagram s lrc plotCompScores f <u>trc plotCompScores simple</u> drc plotGroupAvgSubtract sdrc plotPCAScores drc plotRaw drc plotRMSEx drc plotSimcaClassDist lrc plotVector drc plotVectorCollection sdrc printSvmTable drc showGroups selectWLs showAllClassRepresentations howSingleClassRepresentation imca mmp plitSequence

Calculate and / or plot a plantida Subtractions in raw-spectra Plot an Aquagram Plot comparison scores; advanced grouping Plot comparison scores; simple grouping Plot averaged subtracted spectra Plot PCA scores Plot Spectra Error / calibration / validation plots of PLSR models Plot SIMCA interclass distances (sdrc) Plot Vector Plot Vector Collection Classify using Support Vector Machines Show SDRC set by grouping Select Wavelengths Show the character representation of a factor Show the character representation of a factor Plot SIMCA interclass distances (dataset) Manually split within sequence Select observations (rows)-accord

bw=", pickResurcespicks\$pickWindow, sep=""



nSub <- paste(onSub, ", bw=", pickResults\$picks\$pickWindow, sep="")

# Preliminary Results

### **Different Water Species**

#### Spectrometer:

- FOSS XDS
- transmittance, pathlength = 1mm
- 400nm 2500nm; Δ=0.5nm
- T = 28.6°C
- up to 5 different water species
- MilliQ as standard
- Abbreviations showing origin of water
  - StU: St. Ulrich
  - Ob: Oberndorf
  - MQ: MilliQ

Analysis in 1st overtone of OH/water (1300-1600nm)

#### 3(4) Different Water Species - no Treatment





Wavelength Aquagram Spectra no treatment (not N corr.)

#### 3(4) Different Water Species - no Treatment

C\_ConSNr ProxCuv@1300-to-1600



#### <u>PCA:</u>

 biggest effect: consecutive scans (98.7%)

#### 3(4) Different Water Species - no Treatment

C\_waterNames ProxCuv@1300-to-1600



### Loadings

#### ProxCuv@1300-to-1600



wavelengths no treatment PCA , bw=25 ,2 PCs for 99% var.

MQ StU Ob MQs no Treatment ProxCuv@1300-to-1600



no Treatment ProxCuv@1300-to-1600



Ob@5 (classic) no treatment class distance (Ob, MQs) = 2.72

no Treatment ProxCuv@1300-to-1600



StU@5 (classic) no treatment class distance (StU, MQs) = 1.72

no Treatment ProxCuv@1300-to-1600



#### <u>MQ / MQs:</u>

 same water different container

MQ@5 (classic) no treatment class distance (MQ, MQs) = 0.61

## SIMCA - Classification

Apparent error rate 0.0371				
	Predicted			
Actual	MQ	StU	Ob	MQs
MQ	208	7	0	
StU		253	0	0
Ob	0	5	185	0
MQs	12	0	0	28

50% new data - Apparent error rate 0.0657					
	Predicted				
Actual	MQ	StU	Ob	MQs	
MQ	107	3	0	0	
StU		127	0	0	
Ob	0	3	89	0	
MQs	15		0	4	

# PLSR (number of consec. scan)



no treatment color by C\_ExpName, N=254

predicted

## PLSR (number of consec. scan)

StU@1300-to-1600



no treatment mvr regressed on Y\_ConSNr , RMSECV= 0.6442 , R2CV= 0.856 , bw=25

### Aquagram

#### Fusion–AllWaters no Treatment, ProxCuv@1300–to–1600

#### Fusion-AllWaters no Treatment, ProxCuv@1300-to-1600

- MQ

MQs

• • Ob

-

C04

C03

C05

C02

C06

StU

![](_page_23_Figure_3.jpeg)

no treatment grouping by C\_waterNames (not N corr.)

no treatment grouping by C\_waterNames (not N corr.)

aucs.dce-diff

Classic

# Aquagram

Fusion-AllWaters no Treatment, ProxCuv@1300-to-1600

#### New Aquagram Algorythm

- completely different mathematics,
- but showing high similarity with results of ,,classic" code
- independent, stable scale
- unit: °C

#### now possible:

• use fix-scale aquagrams in time-resolved experiments

![](_page_24_Figure_9.jpeg)

no treatment grouping by C\_waterNames (not N corr.)

aucs.dce-diff

## Effect of Magnetic Fields on Water

#### Spectrometer:

- FOSS XDS
- transmittance, pathlength = 1mm
- 400nm 2500nm; ∆=0.5nm
- T = 28.6°C
- 2 different kinds of magnetic treatment
- MilliQ as standard
- Analysis in 1st overtone of OH/water (1300-1600nm)
- Preliminary Results: as this was an industry-financed experiment, I can not give out hardcopies of even the preliminary results before we published them — I humbly ask for your understanding. Thank you.

![](_page_26_Figure_0.jpeg)

## Online Aquagram Maker

water-temperature

®.Tsenkova: 1st overtone free OH stretch (OH-(H2O)4)

Aditional information should be provided by the site editors.

- upload own data
- compare with available data
- easy to use
- in development -- coming soon...
- <u>http://aquaphotomics.com</u>

### Acknowledgements

- Prof. Roumiana Tsenkova & Zoltan Kovacs
- Zoltan Kovacs
- George Bázár & Alexandar Slavchev
- everybody at the Biomeasurement Technology Laboratory

![](_page_28_Picture_5.jpeg)

![](_page_28_Picture_6.jpeg)

![](_page_29_Picture_0.jpeg)

## Thank You !