photobiology Version 0.2.19 User Guide

Pedro J. Aphalo April 28, 2014

1 Introduction

We have developed a set of packages to facilitate the calculation of many different quantities that can be derived from spectral irradiance data. The base package in this suite is called photobiology, and is the package described here. There other specialized packages for quantification of ultraviolet radiation (photobiologyUV), visible radiation (photobiologyVIS), or based on Phytochrome (photobiologyPhy), Cryptochrome (photobiologyCry) (both photoreceptors present in plants), and spectral data for filters (photobiologyFilters). In the future it will be submitted to CRAN (Comprehensive R archive network), it is meanwhile available from https://bitbucket.org/aphalo/photobiology/downloads. There is also a public Git repository at https://bitbucket.org/aphalo/photobiology from where the source code of the current an earlier versions can be cloned.

2 Installation and use

The functions in the package photobiologyPhy are made available by installing the packages photobiology (once) and loading it from the library when needed. To load the package into the workspace we use library(photobiology).

```
library(photobiology)
library(photobiologyUV)
library(photobiologyVIS)
```

3 Spectral data

If our spectral irradiance data is in W m-2 nm-1, and the wavelength in nm, as in the case of the Macam spectroradiometer, the functions will return the effective irradiance in W m-2. In this example we calculated a biologically effective irradiance.

If, for example, the spectral irradiance output by our model of spectroradiometer is in mW m-2 nm-1, and the wavelengths are in Ångstrom then to obtain the effective irradiance in W m-2we will need to convert the units.

```
energy_irradiance(wavelength/10, irrad/1000)
```

In this example, we take advantage of the behaviour of the S language: an operation between a scalar and vector, is equivalent to applying this operation to each member of the vector. Consequently, in the code above, each value from the vector of wavelengths is divided by 10, and each value in the vector of spectral irradiances is divided by 1000.

If the spectral irradiance is in mW m-2 nm-1 then values should be multiplied by 10 to convert them to W m-2 nm-1.

It is very important to make sure that the wavelengths are in nanometres as this is what the functions expect. If the wavelengths are in the wrong units, the action spectra will be wrongly calculated, and the returned value for effective irradiance will be completely wrong.

Here we just use the example data supplied with the package.

4 Defining wavebands

All functions use wavebands as definitions of the range of wave lengths and the spectral weighting function (SWF) to use in the calculations. A few other bits of information may be included to fine-tune calculations. The waveband definitions do NOT describe whether input spectral irradiances are photon or energy based, nor whether the output irradiance will be based on photon or energy units. Waveband objects belong to the S3 class "waveband".

When defining a waveband which uses a SWF, a function can be supplied either based on energy effectiveness, on photon effectiveness, or one function for each one. If only one function is supplied the other one is built automatically, but if performance is a concern it is better to provide two separate functions. Another case when you might want to enter the same function twice, is if you are using an absorptance spectrum as SWF, as the percent of radiation absorbed will be independent of whether photon or energy units are used for the spectral irradiance.

5 Calculating irradiance or exposure

There is one basic function for these calculations <code>irradiance()</code>, it takes an array of wavelengths (sorted in strictly increasing order), and the corresponding values of spectral irradiance. By default the input is assumed to be in energy units, but parameter <code>unit.in</code> cab be used to adjust the calculations to expect photon units. The type of unit used for the calculated irradiance (or exposure) is set by the parameter <code>unit.out</code> with no default. If no <code>w.band</code> parameter is supplied, the whole spectrum spectrum input is used, unweighted, to calculate the total irradiance. If a <code>w.band</code> is supplied, then the range of wavelengths specified and SWF if present are used for calculating the irradiance. If the waveband definition does not include a SWF, then the unweighted irradiance is returned, if the definition includes a SWF, then a weighted irradiance is returned.

The functions photon_irradiance() and energy_irradiance(), just call irradiance() with the unit.out set to "photon" or "energy" respectively.

We first load some data to play with.

```
data(sun.data)
attach(sun.data)
```

Then we compare the calculations based on the different wavebands defined in the previous section and the predefined functions in the packages photobiologyUS and photobiologyUV. The predefined functions have the advantage of allowing the specification of parameters to modify the w.band created. In the example bellow, we use this to set the normalization wavelength.

This is how CIE() is defined in photobiologyUV.

The generic functions print(), min(), max(), range(), mean() and labels() call the corresponding special functions defined for wavebdand objects: print.waveband(), range.waveband(), etc.

```
PAR()

## PAR

## low (nm) 400

## high (nm) 700

str(PAR())

## List of 9

## $ low : num 400

## $ high : num 700
```

```
## $ weight : NULL
## $ SWF.e.fun: NULL
## $ SWF.q.fun: NULL
## $ SWF.norm : NULL
## $ norm : NULL
## $ hinges : num [1:4] 400 400 700 700
## $ name : chr "PAR"
## - attr(*, "class")= chr "waveband"
range(PAR())
## [1] 400 700
min(PAR())
## [1] 400
max(PAR())
## [1] 700
mean(PAR()) # gives wavelength at center
## [1] 550
labels(PAR())
## $name
## [1] "PAR"
CIE()
## CIE98.298
## low (nm) 250
## high (nm) 400
## weighted SWF
## normalized at 298 nm
str(CIE())
## List of 9
## $ low : num 250
## $ high : num 400
## $ weight : chr "SWF"
## $ SWF.e.fun:function (w.length)
## $ SWF.q.fun:function (w.length)
## $ SWF.norm : num 298
## $ norm : num 298
## $ hinges : num [1:6] 250 250 298 328 400 ...
## $ name : chr "CIE98.298"
## - attr(*, "class")= chr "waveband"
labels(CIE())
## $name
## [1] "CIE98.298"
```

```
labels(CIE(300))
## $name
## [1] "CIE98.300"
labels(UVB())
## $name
## [1] "UVB.ISO"
labels(UVB("medical"))
## $name
## [1] "UVB.medical"
labels(UVB("ISO"))
## $name
## [1] "UVB.ISO"
labels(new_waveband(100, 200))
## $name
## [1] "range.100.200"
labels(new_waveband(100, 200, wb.name = "ANY STRING"))
## [1] "ANY STRING"
```

Now the example calculations.

```
photon_irradiance(w.length, s.e.irrad, my_PAR)
## range.400.700
## 0.0008938
photon_irradiance(w.length, s.e.irrad, my_PARx)
## my_PARx
## 0.0008938
photon_irradiance(w.length, s.e.irrad, PAR())
##
       PAR
## 0.0008938
photon_irradiance(w.length, s.e.irrad, my_CIE_1)
## range.250.400.wtd
       2.037e-07
photon_irradiance(w.length, s.e.irrad, my_CIE_2)
## range.250.400.wtd
## 2.037e-07
```

```
photon_irradiance(w.length, s.e.irrad, my_CIE_3)
## range.250.400.wtd
   2.037e-07
energy_irradiance(w.length, s.e.irrad, my_CIE_1)
## range.250.400.wtd
##
          0.08178
energy_irradiance(w.length, s.e.irrad, my_CIE_2)
## range.250.400.wtd
##
      0.08178
energy_irradiance(w.length, s.e.irrad, my_CIE_3)
## range.250.400.wtd
           0.08178
energy_irradiance(w.length, s.e.irrad, CIE())
## CIE98.298
## 0.08178
energy_irradiance(w.length, s.e.irrad, CIE(298))
## CIE98.298
## 0.08178
energy_irradiance(w.length, s.e.irrad, CIE(300))
## CIE98.300
## 0.1261
```

Lists of wavebands are also accepted as argument.

The are also convenience functions for calculating how 'total' irradiance is split among different contiguous bands of the spectrum. The functions <code>split_photon_irradiance()</code> and <code>split_energy_irradiance()</code>, just call <code>split_irradiance()</code> with the unit.out set to "photon" or "energy" respectively.

```
split_energy_irradiance(w.length, s.e.irrad, c(300, 400, 500,
   600, 700, 800))
## range.300.400 range.400.500 range.500.600 range.600.700
                                                  58.54
##
         28.32
                      69.63
                                    68.53
## range.700.800
##
          43.90
split_energy_irradiance(w.length, s.e.irrad, c(400, 500, 600,
   700), scale = "percent")
## range.400.500 range.500.600 range.600.700
         35.40
                      34.84
split_photon_irradiance(w.length, s.e.irrad, c(400, 500, 600,
   700), scale = "percent")
## range.400.500 range.500.600 range.600.700
   29.41 35.14 35.45
```

6 Calculating ratios

The function waveband_ratio() takes basically the same parameters as irradiance, but two waveband definitions instead of one, and two unit.out definitions instead of one. This is the base function used in all the 'ratio' functions in the photobiology package.

The derived functions are: photon_ratio(), energy_ratio(), and photons_energy_ratio. The packages photobiologyVIS and photobiologyUV use these to define some convenience functions, and here we give and example for a function not yet implemented, but which you may find as a useful example.

If for example we would like to calculate the ratio between UVB and PAR radiation, we would use either of the following function calls, depending on which type of units we desire.

```
photon_ratio(w.length, s.e.irrad, UVB(), PAR())
## [1] 0.001708
energy_ratio(w.length, s.e.irrad, UVB(), PAR())
## [1] 0.00299
```

If we would like to calculate a conversion factor between PPFD (PAR photon irradiance in mol s-1 m-2) and PAR (energy) irradiance (W m-2) for a light source for which we have spectral data we could use the following code.

```
conv.factor <- photons_energy_ratio(w.length, s.e.irrad, PAR())

PPFD.mol.photon <- 0.001

PAR.energy <- PPFD.mol.photon/conv.factor
print(conv.factor)</pre>
```

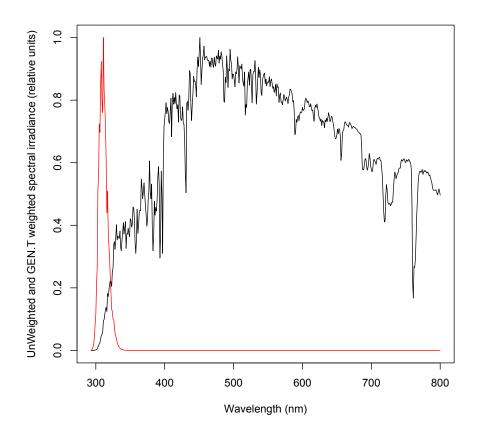
```
## [1] 4.544e-06
print(PPFD.mol.photon * 1e+06)
## [1] 1000
print(PAR.energy)
## [1] 220.1
```

The 'ratio' functions do not accept lists of waveband objects as the 'irradiance' functions do. This is a feature, as otherwise it would be too easy to make mistakes. It is possible to use the 'irradiance' functions to calculate several ratios in one go.

7 Calculating weighted spectral irradiances

This calculation is not very frequently used, but it is very instructive to look at spectral data in this way, as it can make apparent the large effect that small measuring errors can have on the estimated effective irradiances or exposures.

We here plot weighted and unweighted irradiances using simulated solar spectral irradiance data.



8 Auxiliary functions

Sometimes it is needed to add (parallel sum) between two spectra, even if the two spectra have been measured at different wavelengths or wavelength steps. This can happen for example when we would like to calculate the spectrum of a combination of two light sources from the individual spectra. The function sum_spectra() can be used in such a case. It's use is described in User Guide of the photobiologyLamps package.

A function interpolate_spectrum is also included to facilitate interpolation of spectral values. It is used in the function described in the previous paragraph, but also be used by itself when interpolation is needed. Under the hood it uses R's spline function if there are fewer than 25 data points, and uses approx otherwise. It allows easier control of values to be used for extrapolation.

9 Dealing with real 'noisy' spectral data

The first thing to do is to think whether any part of the spectral measurements can be *a priori* known to be equal to zero. For example for the solar spectrum at ground level it is safe to assume that the spectral irradiance is zero for all wavelengths shorter than 293 nm. If the data are noisy, it is best to discard these data before calculating any effective UV doses.

In the following example we use a longer wavelength (297 nm) just to show how the function works, because the example spectral data set starts at 293 nm.

Sub-setting can be easily done as follows if the data are in a data.frame (of course, replacing w.length with the name used in your data frame for the wavelengths array):

```
subsetted.sun.data <- subset(sun.data, w.length >= 297)
head(subsetted.sun.data)
##
      w.length s.e.irrad s.q.irrad
## 5
           297 0.0001533 3.807e-10
## 6
           298 0.0003670 9.141e-10
## 7
           299 0.0007845 1.961e-09
           300 0.0012646 3.171e-09
## 8
## 9
           301 0.0026237 6.602e-09
## 10
           302 0.0039226 9.903e-09
```

And then just use the sub-setted data in your calculations.

If the data are not in a dataframe, then there are different options: 1) create a data frame from your data, 2) use the function trim_tails() from this package, or 3) just use R commands. Here we give examples of the use of trim_tails(), and just use the same data as in earlier examples. First we delete all data for wavelengths shorter than 293 nm.

```
trimmed.sun.data <- trim_tails(w.length, s.e.irrad, low.limit = 297)</pre>
head(trimmed.sun.data)
##
    w.length s.irrad
       297 0.0001533
## 1
## 2
         298 0.0003670
## 3
         299 0.0007845
## 4
         300 0.0012646
## 5
         301 0 0026237
## 6
     302 0.0039226
```

```
tail(trimmed.sun.data)

## w.length s.irrad
## 499    795    0.4147

## 500    796    0.4081

## 501    797    0.4141

## 502    798    0.4236
## 503    799    0.4186
## 504    800    0.4069
```

This function returns a new data.frame and uses always the same variable names for the columns.

```
trimmed.both.sun.data <- trim_tails(w.length, s.e.irrad, low.limit = 297,</pre>
   high.limit = 550)
head(trimmed.both.sun.data)
   w.length s.irrad
       297 0.0001533
## 1
## 2
         298 0.0003670
## 3
         299 0.0007845
## 4
        300 0.0012646
## 5
       301 0.0026237
## 6
        302 0.0039226
tail(trimmed.both.sun.data)
      w.length s.irrad
##
## 249
        545 0.7272
## 250
           546 0.6993
## 251
           547 0.7119
           548 0.6937
## 252
## 253
           549 0.7021
## 254
       550 0.7047
```

If we supply a different value than the default NULL for the parameter fill, the w.length values are kept, and the trimmed spectral irradiance values replaced by the value supplied.

```
trimmed.na.sun.data <- trim_tails(w.length, s.e.irrad, low.limit = 297,</pre>
   fill = NA)
head(trimmed.na.sun.data)
## w.length s.irrad
## 1
         293
                    NA
## 2
         294
                    NA
## 3
         295
                    NA
## 4
         296
                    NΑ
## 5
         297 0.0001533
## 6
      298 0.0003670
```

```
trimmed.both.na.sun.data <- trim_tails(w.length, s.e.irrad, low.limit = 297,
    high.limit = 400, fill = NA)
head(trimmed.both.na.sun.data)</pre>
```

```
##
     w.length
                s.irrad
## 1
          293
                       NΑ
## 2
           294
                       NA
## 3
           295
                       NA
## 4
           296
                       NA
## 5
           297 0.0001533
## 6
           298 0.0003670
tail(trimmed.both.na.sun.data)
##
       w.length s.irrad
## 503
            795
                       NA
## 504
             796
                       NA
## 505
             797
                       NA
## 506
             798
                       NΑ
             799
## 507
## 508
            800
                       NA
```

In addition to NA we can supply an arbitrary numeric value.

```
trimmed.zero.sun.data <- trim_tails(w.length, s.e.irrad, low.limit = 297,</pre>
   fill = 0
head(trimmed.zero.sun.data)
     w.length s.irrad
##
## 1
          293 0.0000000
## 2
          294 0.0000000
## 3
          295 0.0000000
## 4
          296 0.0000000
## 5
          297 0.0001533
          298 0.0003670
## 6
```

10 Optimizing performance

When developing the current version of photobiology quite a lot of effort was spent in optimizing performance, as in one of our experiments, we need to process several hundreds of thousands of measured spectra. The defaults should provide good performance in most cases, however, some further improvements are achievable, when a series of different calculations are done on the same spectrum, or when a series of spectra measured at exactly the same wavelengths are used for calculating weighted irradiances or exposures.

In the case of doing calculations repeatedly on the same spectrum, a small improvement in performance can be achieved by setting the parameter check.spectrum=FALSE for all but the first call to irradiance(), or photon_irradiance(), or energy_irradiance(), or the equivalent function for ratios. It is also possible to set this parameter to FALSE in all calls, and do the check beforehand by explicitly calling check_spectrum().

In the case of calculating weighted irradiances on many spectra having exactly the same wavelength values, then a significant improvement in the performance can be achieved by setting use.cached.mult=TRUE, as this reuses

the multipliers calculated during successive calls based on the same waveband. However, to achieve this increase in performance, the tests to ensure that the wavelength values have not changed, have to be kept to the minimum. Currently only the length of the wavelength array is checked, and the cached values discarded and recalculated if the length changes. For this reason, this is not the default, and when using caching the user is responsible for making sure that the array of wavelengths has not changed between calls.

You can use the package microbenchmark to time the code and find the parts that slow it down. I have used it, and also I have used profiling to optimize the code for speed. The examples below show how choosing different values from the defualts can speed up calculations when the same calculations are done repeatedly on spectra measured at exactly the same wavelengths, something which is usual when analysing spectra measured with the same instrument. The choice of defaults is based on what is best when processing a moderate number of spectra, say less than a few hundreds, as opposed to many thousands.

```
library(microbenchmark)
```

The convenience functions are slightly slower than the generic **irradiance** function by about 2% if using the cache.

```
res <- microbenchmark(irradiance(w.length, s.e.irrad, PAR(),
    unit.out = "photon", use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res

## Unit: microseconds
## expr min lq median uq max neval
## 105.8 107.4 111 114.9 317.7 100

res <- microbenchmark(photon_irradiance(w.length, s.e.irrad,
    PAR(), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res

## Unit: microseconds
## expr min lq median uq max neval
## 110.2 111.8 112.2 114.3 214.8 100</pre>
```

Using the cache when repeatedly applying the same unweighted waveband halves the execution time.

```
PAR(), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res

## Unit: microseconds
## expr min lq median uq max neval
## 105.4 106.6 107.4 114.3 297 100</pre>
```

Disabling the checking of the spectrum halves once again the execution time.

Saving a waveband object and reusing it, gives a speed up of about 12% when all other optimizations are also used. The reduction in execution time from the original default values is about 70%.

```
myPAR <- PAR()
res <- microbenchmark(photon_irradiance(w.length, s.e.irrad,</pre>
   PAR(), use.cache = TRUE, check.spectrum = FALSE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                            uq max neval
       68.08 68.89 69.5 70.51 149.9 100
res <- microbenchmark(photon_irradiance(w.length, s.e.irrad,</pre>
   myPAR, use.cache = TRUE, check.spectrum = FALSE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
## Unit: microseconds
## expr min lq median
                            uq max neval
    55.92 56.73 57.14 57.54 129.3 100
```

Inserting 'hinges' to reduce integration errors slows down the computations a lot, increasing the execution time 5.6 times. If the spectral data is measured

with a small wavelength step, the errors are rather small. By default the use of 'hinges' is automatically decided based on the average wavelength step in the spectral data.

```
energy_irradiance(w.length, s.e.irrad, PAR(), use.cache = TRUE)
## PAR
## 196.7
energy_irradiance(w.length, s.e.irrad, PAR(), use.cache = TRUE,
   use.hinges = TRUE)
## PAR
## 196.6
res <- microbenchmark(energy_irradiance(w.length, s.e.irrad,</pre>
   PAR(), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median uq max neval
       111 112.2 112.7 116.7 305.1 100
res <- microbenchmark(energy_irradiance(w.length, s.e.irrad,</pre>
   PAR(), use.cache = TRUE, use.hinges = TRUE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median uq max neval
## 501.7 517.9 553.1 587.8 1024 100
```

When using BSWFs the speed up by use of the cache is more important, with CIE the execution time is reduced by 75%.

Saving a waveband object that uses weighting and reusing it, gives a speed up of about 25% when all other optimizations are also used. when using CIE the reduction in execution time from the original default values is about 85%.

```
myCIE <- CIE()
res <- microbenchmark(photon_irradiance(w.length, s.e.irrad,</pre>
   CIE(), use.cache = TRUE, check.spectrum = FALSE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                            uq max neval
##
        81.85 82.67 83.47 86.31 527.2 100
res <- microbenchmark(photon_irradiance(w.length, s.e.irrad,</pre>
   myCIE, use.cache = TRUE, check.spectrum = FALSE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median uq max neval
   59.16 60.38 60.78 61.8 134.1 100
```

Inserting 'hinges' to reduce integration errors slows down the computations a lot. If the spectral data is measured with a small wavelength step, the errors are rather small. By default the use of 'hinges' is automatically decided based on the average wavelength step in the spectral data.

```
energy_irradiance(w.length, s.e.irrad, CIE(), use.cache = TRUE)
## CTE98.298
## 0.08178
energy_irradiance(w.length, s.e.irrad, CIE(), use.cache = TRUE,
   use.hinges = TRUE)
## CIE98.298
   0.08181
res <- microbenchmark(energy_irradiance(w.length, s.e.irrad,</pre>
   CIE(), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                            uq max neval
       124.8 126.4 127 133.3 555.1 100
res <- microbenchmark(energy_irradiance(w.length, s.e.irrad,</pre>
   CIE(), use.cache = TRUE, use.hinges = TRUE), times = 100L,
   control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median uq max neval
## 514.2 528.2 542.8 562.6 1081 100
```

Using the cache also helps with split_irradiance.

```
res <- microbenchmark(split_photon_irradiance(w.length, s.e.irrad,</pre>
   c(400, 500, 600, 700)), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                           uq max neval
        618 677.7 692.9 724.1 3142
res <- microbenchmark(split_photon_irradiance(w.length, s.e.irrad,</pre>
   c(400, 500, 600, 700), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                            uq
                                max neval
       338 356.6 360.8 369.2 871.2 100
```

Using hinges slows down calculations increasing the execution time by 2.4 times, which is less than if calculating the same three wavebands separately, as all hinges are inserted in one operation.

```
res <- microbenchmark(split_photon_irradiance(w.length, s.e.irrad,</pre>
   c(400, 500, 600, 700), use.cache = TRUE), times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median uq max neval
        355.8 359.2 365.7 372 1054
res <- microbenchmark(split_photon_irradiance(w.length, s.e.irrad,</pre>
   c(400, 500, 600, 700), use.cache = TRUE, use.hinges = TRUE),
    times = 100L, control = list(warmup = 10L))
res$expr <- ""
res
## Unit: microseconds
## expr min lq median
                            uq max neval
   763.8 779 821.8 850.1 2428 100
```

Disabling checking of the spectrum reduces the execution time, but proportionally not as much as for the irradiance functions, as the spectrum is checked only once independently of the number of bands into which it is split.

As all the execution times are in microseconds, all the optimizations discussed above are totally irrelevant unless you are planning to repeat similar calculations on thousands of spectra.

```
detach(sun.data)
```