Software Tutorial: Verification and post-processing

Alexander Jordan, Sebastian Lerch

ScienceFore Summer School, October 2017

Some introduction?

absolute forecast evaluation, scoringRules introduction, some simulations?

Data

Set path to directory containing the data

```
data_dir <- "/path/to/data/"
```

and load the data set

```
load(paste0(data_dir, "HDwind.Rdata"))
```

Contents of the data set

The data set contains objects `ensfc`, `obs` and `dates`

The vector `dates` contains dates in 2015 and 2016, at which the forecasts and corresponding observations are valid (at 12 UTC).

```
str(dates)
```

```
##  Date[1:731], format: "2015-01-01" "2015-01-02" "2015-01
```

```
range(dates)
```

```
## [1] "2015-01-01" "2016-12-31"
```

Contents of the data set: Forecasts

The matrix `ensfc` contains ECMWF ensemble forecasts of wind speed for a grid point close to Heidelberg.

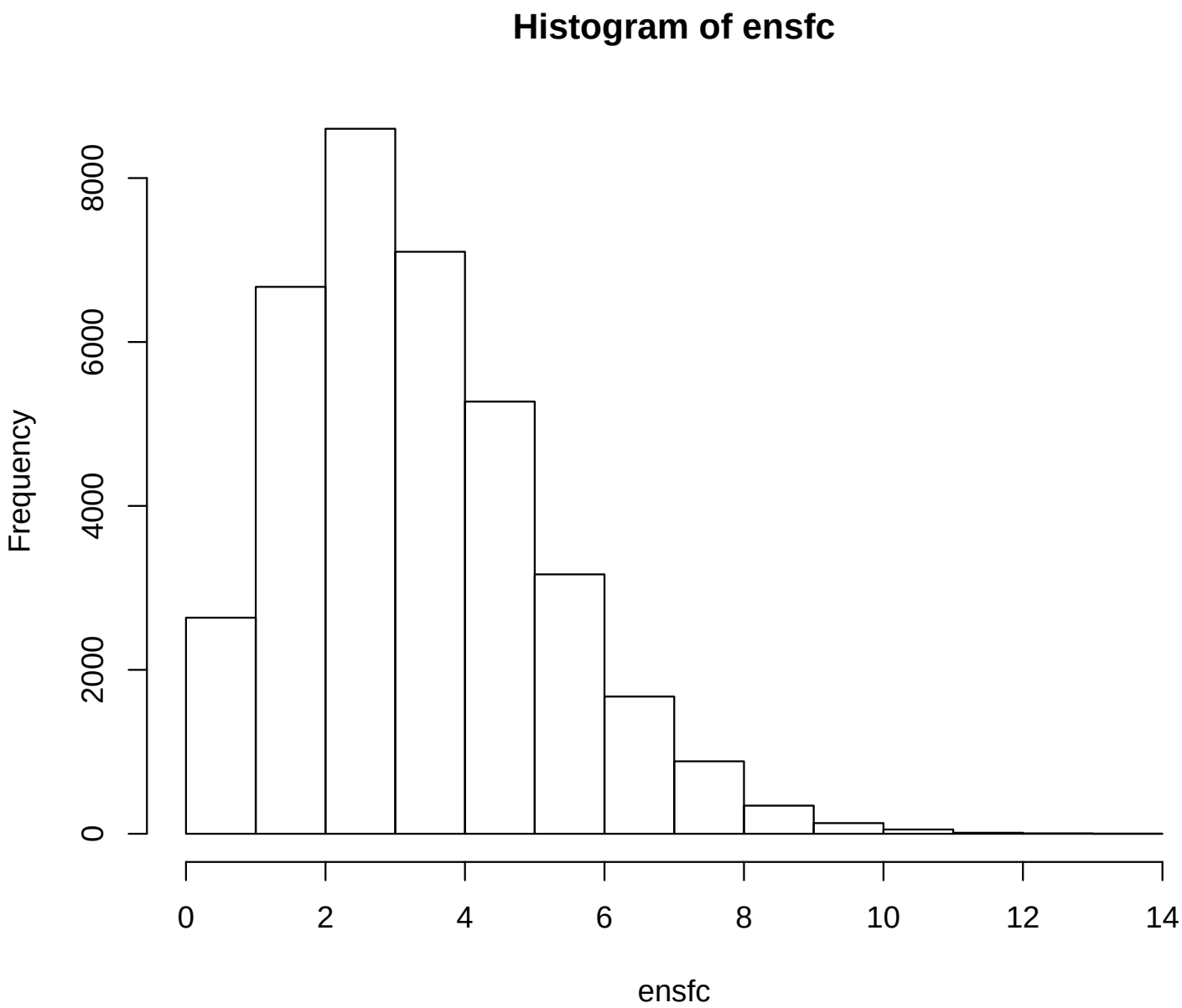
```
str(ensfc)
```

```
##  num [1:731, 1:50] 1.43 4.66 3.34 3.49 1.82 ...
```

For each of the 731 dates, an ensemble with 50 member forecasts is available.

Contents of the data set: Forecasts

```
hist(ensfc)
```



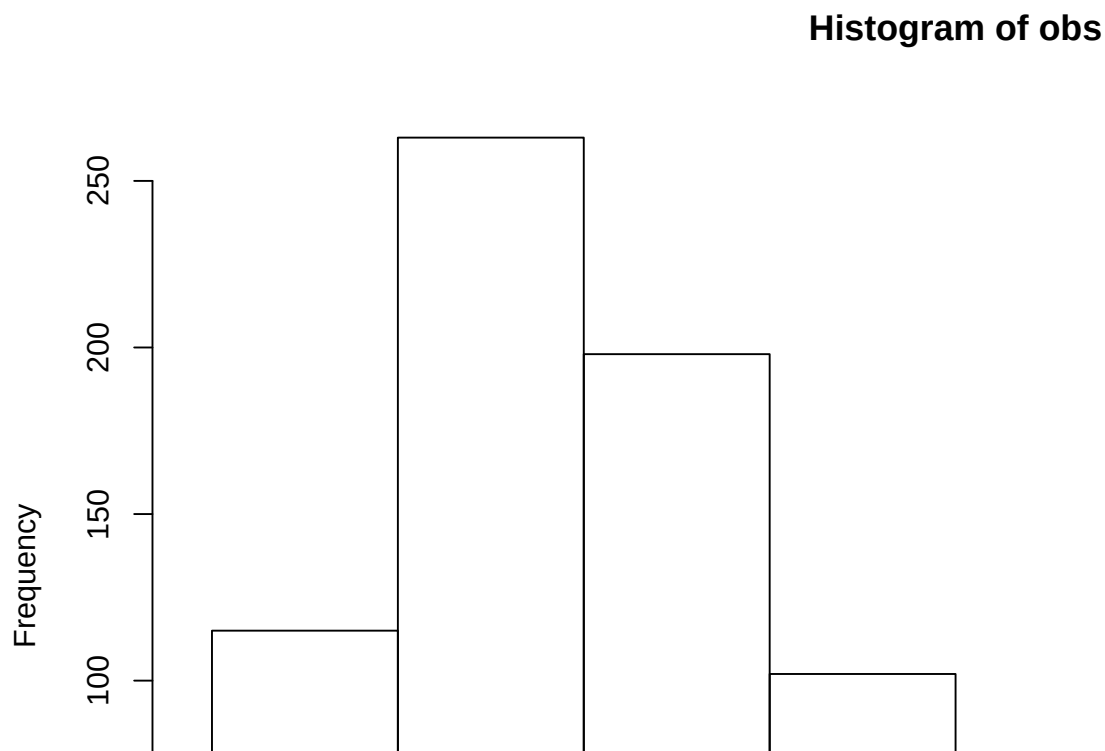
Contents of the data set: Observations

The vector obs contains reanalysis values corresponding to the valid times of the ensemble forecasts.

```
str(obs)
```

```
##  num [1:731] 2.21 7.93 4.92 5.02 3.37 ...
```

```
hist(obs)
```



The scoringRules package

To install the `scoringRules` package

```
install.packages("scoringRules")
```

To load the `scoringRules` package

```
library(scoringRules)
```

Check if version is $\geq$ 0.9.3

```
packageVersion("scoringRules") >= "0.9.3"
```

```
## [1] TRUE
```

If this is not the case, re-install the package from CRAN.

Documentation of the scoringRules package

The documentation of individual functions can be accessed via e.g.

```
?crps_sample
```

To browse the documentation of the functions available in the package use

```
help.start()
```

and navigate to 'packages' - 'scoringRules'

Vignettes with introductions and background information provided with the package can be accessed via

```
browseVignettes("scoringRules")
```

More information is available in our working paper 'Evaluating probabilistic forecasts with the R package scoringRules' available at https://arxiv.org/abs/1709.04743.

Compute the CRPS of the ensemble forecast

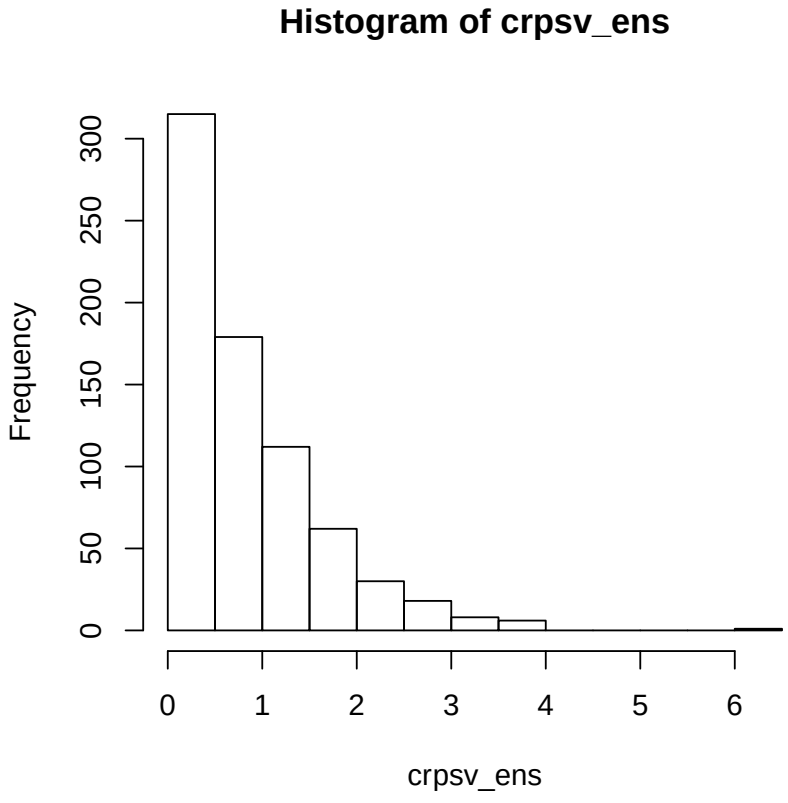
.. or: show simulation examples and use this as exercise?

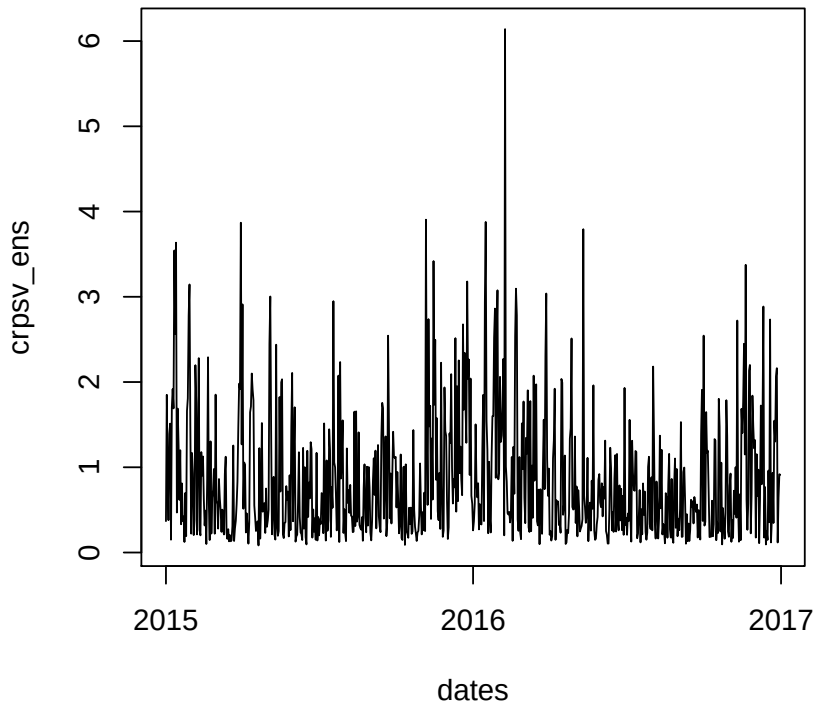
```
crpsv_ens <- crps_sample(y = obs, dat = ensfc)
summary(crpsv_ens)
```

```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.0832  0.2746  0.5988  0.8583  1.1984  6.1404
```

Plot CRPS values

```
par(mfrow = c(1,2))
hist(crpsv_ens)
plot(dates, crpsv_ens, type = "l")
```





Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3