More than just a verification package

Andrew Singleton (MET Norway), Alex Deckmyn (RMI), Christoph Zingerle (ZAMG), Bent Hansen Sass (DMI)

So, is this HARPv3?

So, is this HARPv3?

Kinda. Sorta. Not really...

So, is this HARPv3?

Kinda. Sorta. Not really...

It's harp [as in not HARP]

So, is this HARPv3?

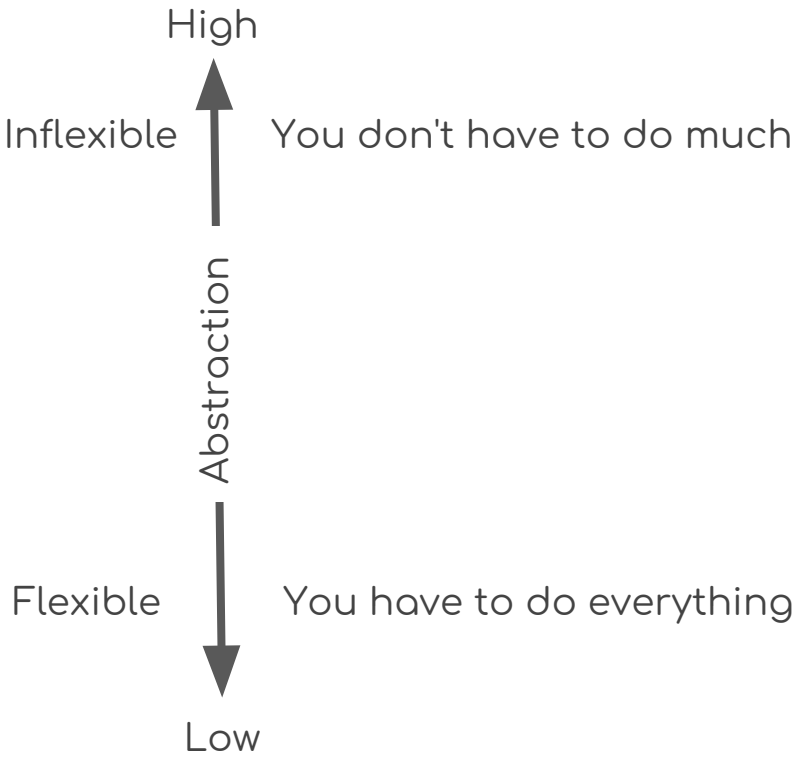
Kinda. Sorta. Not really...

It's harp [as in not HARP]

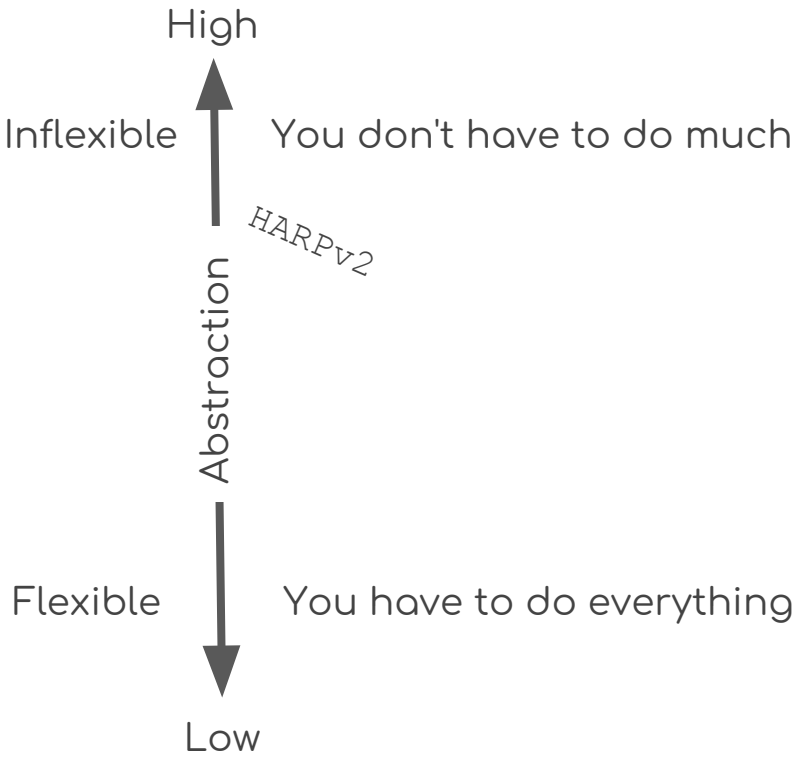
harp refers to a set of R packages that are designed to fulfill tasks related to specific themes. Each harp package has its own version control. harp itself is a collection of packages or an **ecosystem**.

What's different from HARPv2 ?

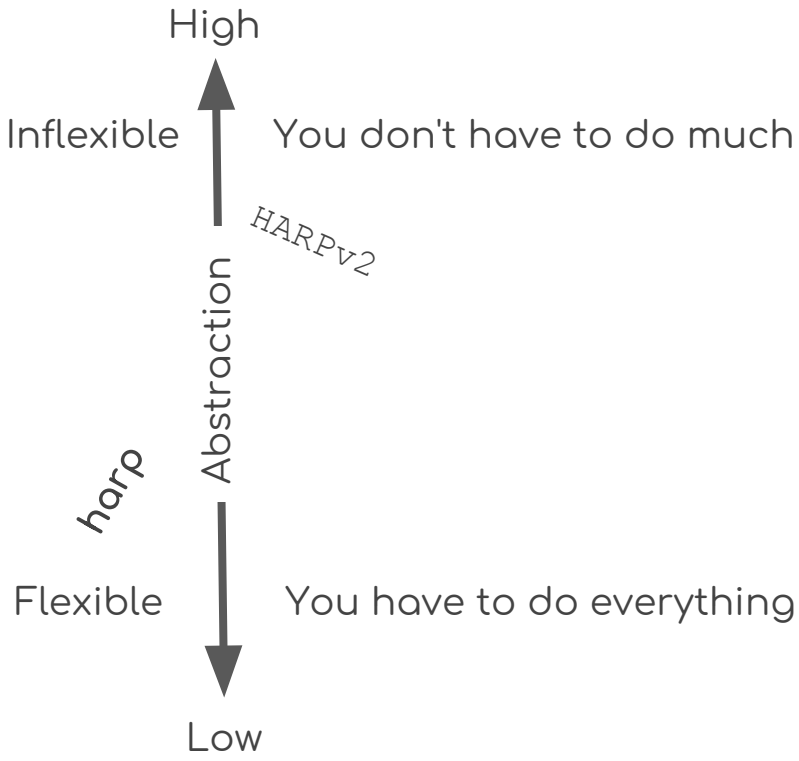
Abstraction Continuum / Spectrum



Abstraction Continuum / Spectrum



Abstraction Continuum / Spectrum




harp

Data IO & interpolation

Data IO & interpolation

Data IO & interpolation

Point
verification and wrangling

Data IO & interpolation

Point
verification and wrangling

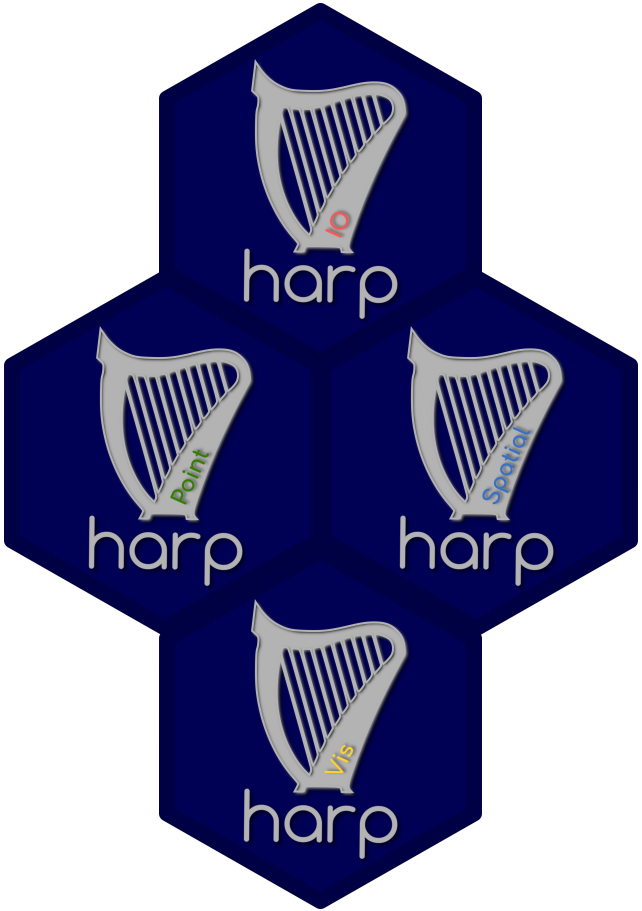
Data IO & interpolation

Point
verification and wrangling

Spatial
verification and wrangling

Data IO & interpolation

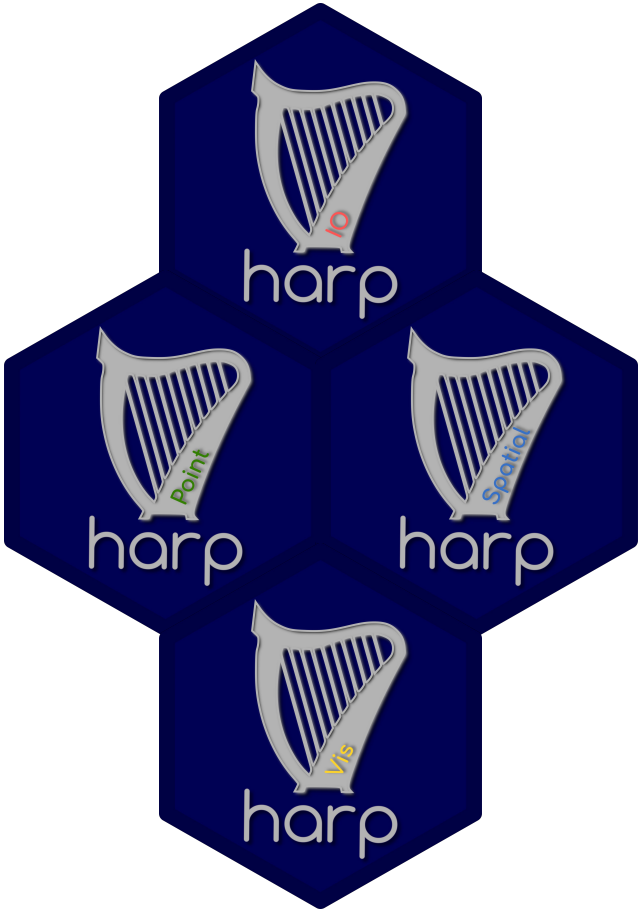
Point
verification and wrangling



Spatial
verification and wrangling

Data IO & interpolation

Point
verification and wrangling



Spatial
verification and wrangling

(Interactive) data visualisation


io
harp

Forecast formats

Grib (1 & 2)

FA

netCDF (currently MET Norway only for interpolating)

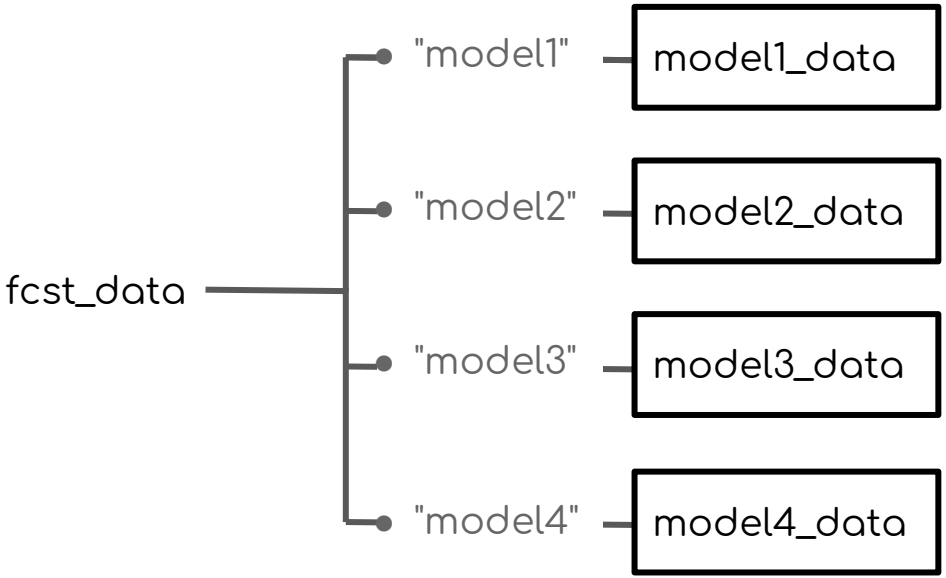
vfld

SQLite

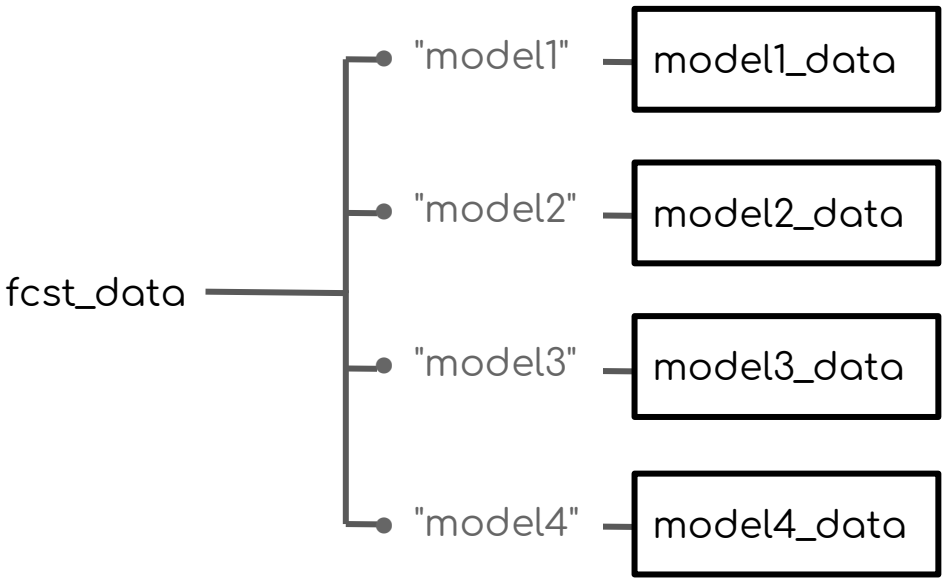
Generic read/write functions


```
read_eps(...) / read_eps_interpolate(...) / read_eps_profile(...)

read_det(...) / read_det_interpolate(...) / read_det_profile(...)

read_point_forecast(...) / read_profile_forecast(...)
```

Data structure



Data structure



Methods

- ❏ filter
- ❏ select
- ❏ mutate
- ❏ transmute
- ❏ arrange
- ❏ join

- ❏ select_members
- ❏ lag_forecast

Observation formats

Point observations	Spatial observations
vobs	HDF5
SQLite	INCA
BUFR (currently MET Norway only)	netCDF

Generic read/write functions


```
read_obs_convert(...) ##For point observations

read_point_obs(...) / read_profile_obs(...)

read_???(...) ##For spatial observations
```


Point
harp

Point verification


```
ens_verify(...)

det_verify(...)
```

Point verification

```
ens_verify(...)

det_verify(...)
```

Get categorical scores by adding thresholds

```
ens_verify(..., thresholds = seq(-10, 20, 5))

det_verify(..., thresholds = seq(-10, 20, 5))
```

Point verification

```
ens_verify(...)

det_verify(...)
```

Aggregate scores in any way using groupings

```
ens_verify(..., groupings = c("leadtime", "fcst_cycle", "SID"))

det_verify(..., groupings = c("leadtime", "fcst_cycle", "SID"))
```

Take account of errors

Error bars and statistical significance of differences by bootstrapping

```
bootstrap_score(..., ens_crps)
```

Observation errors

```
jitter_fcst(..., function(x) x + rnorm(1, 0, 0.5))

check_obs_against_fcst(...)
```

Take advantage of parallel processing


```
library(furrr)

parameters <- c("T2m", "RH2m", "Pmsl", "AccPcp3h")

plan(multiprocess)

verif <- future_map(
    parameters,
    ~ens_read_and_verify(
        ...,
        parameter = .x
    )
)
```


Vis
harp

Forecast time series for EPS

Spaghetti plots

Ribbon plots

Box whisker plots

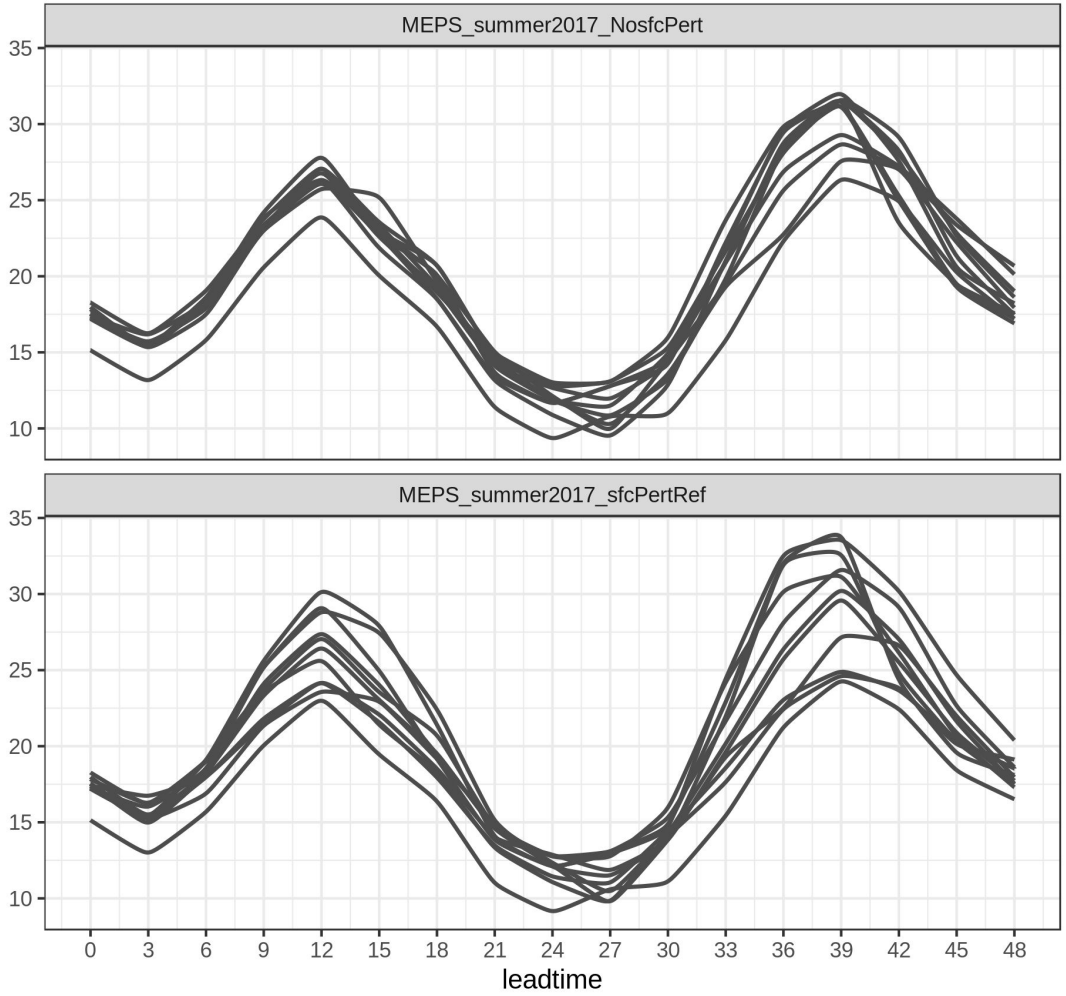
Violin plots

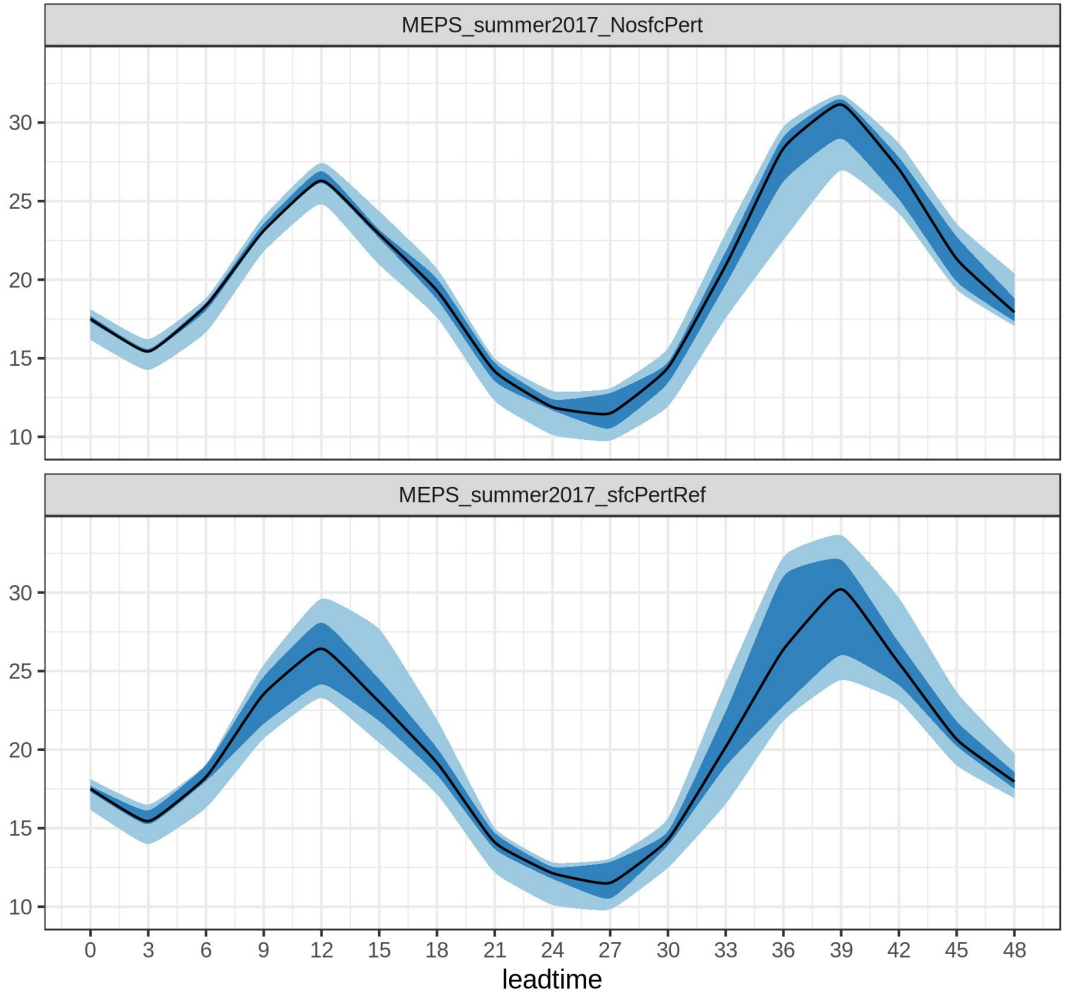
HOPs

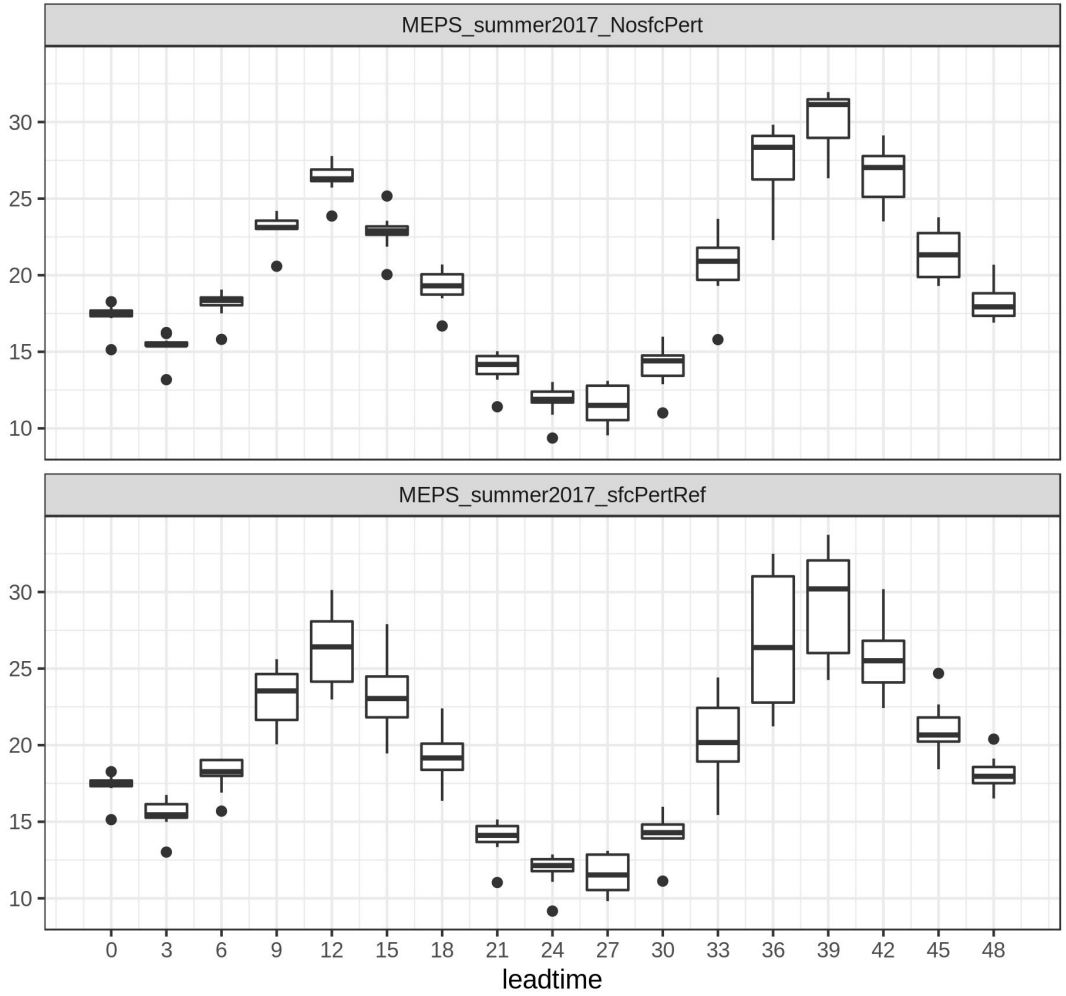
Ridgeline plots

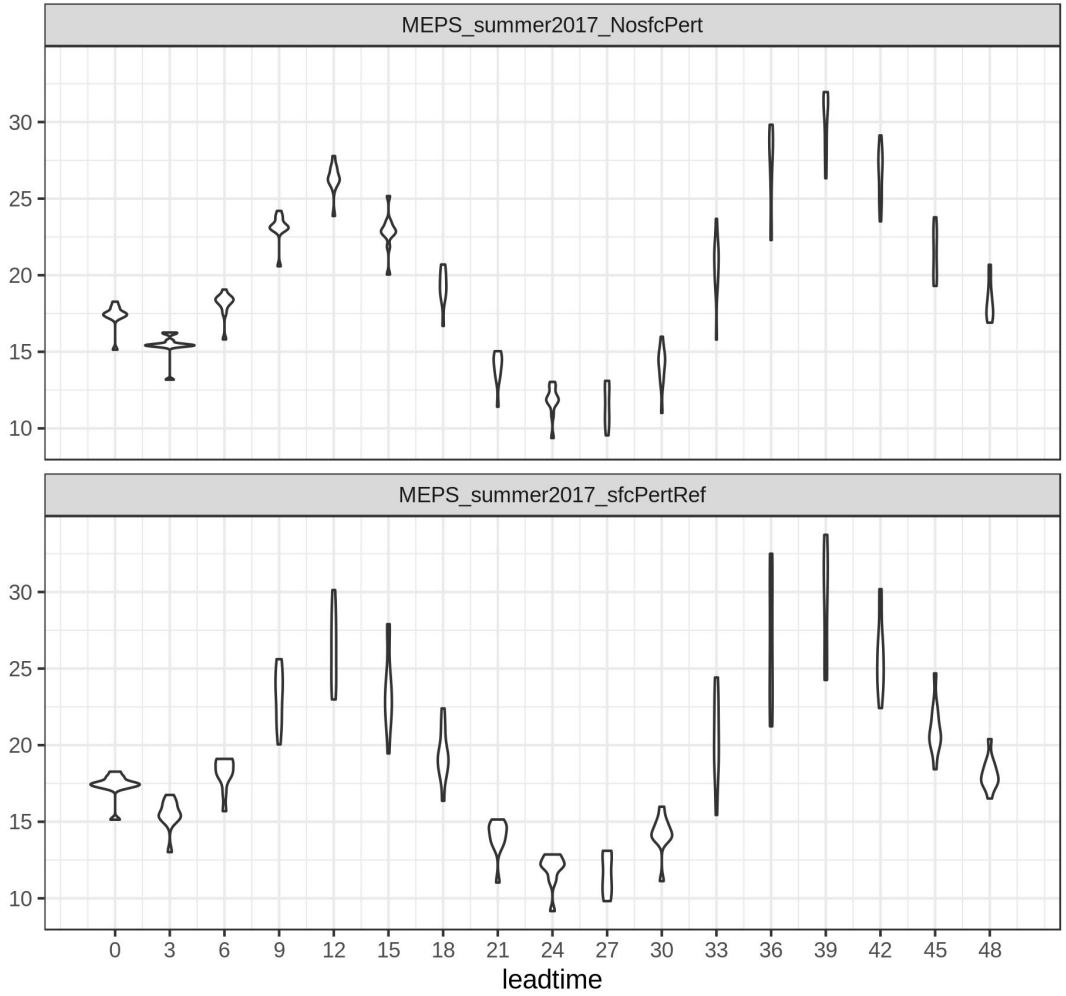
Stacked probability (bar / area)

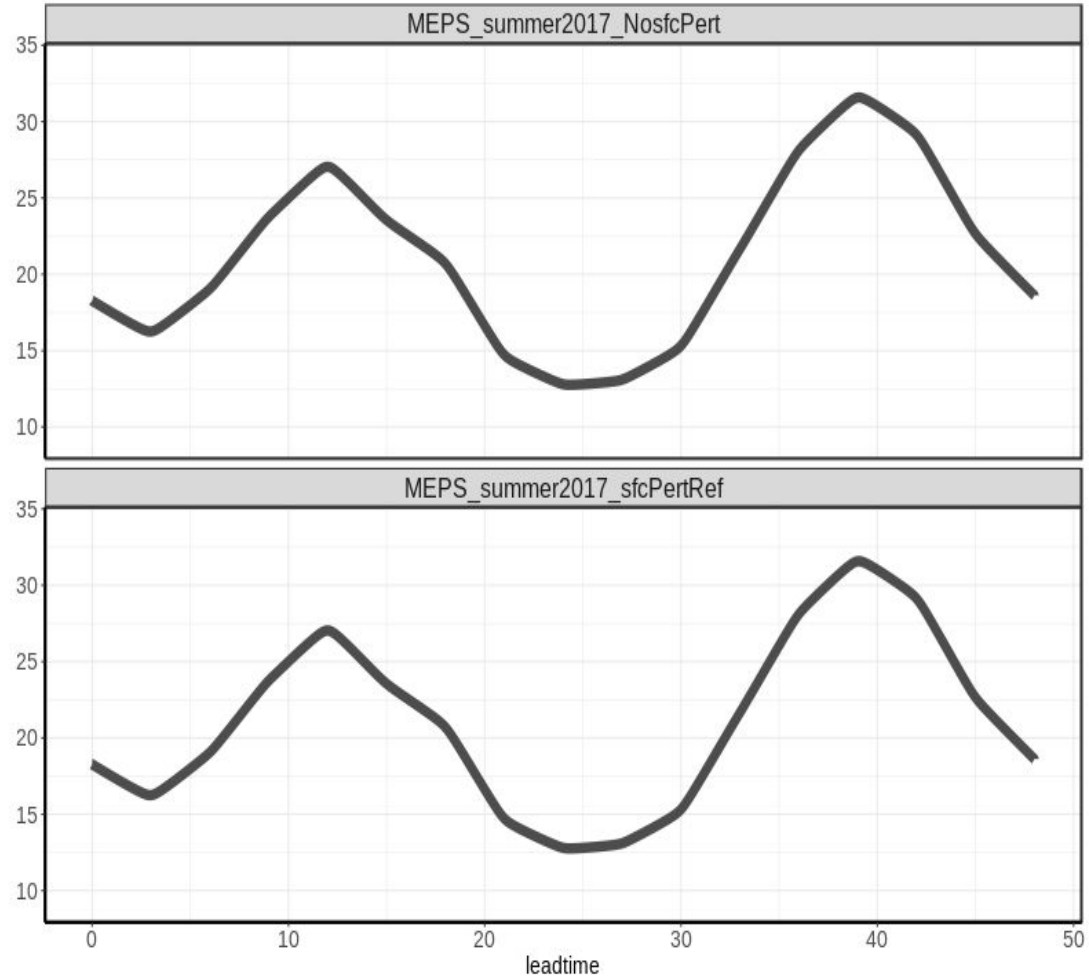
```
plot_station_eps(..., type = "...")
```

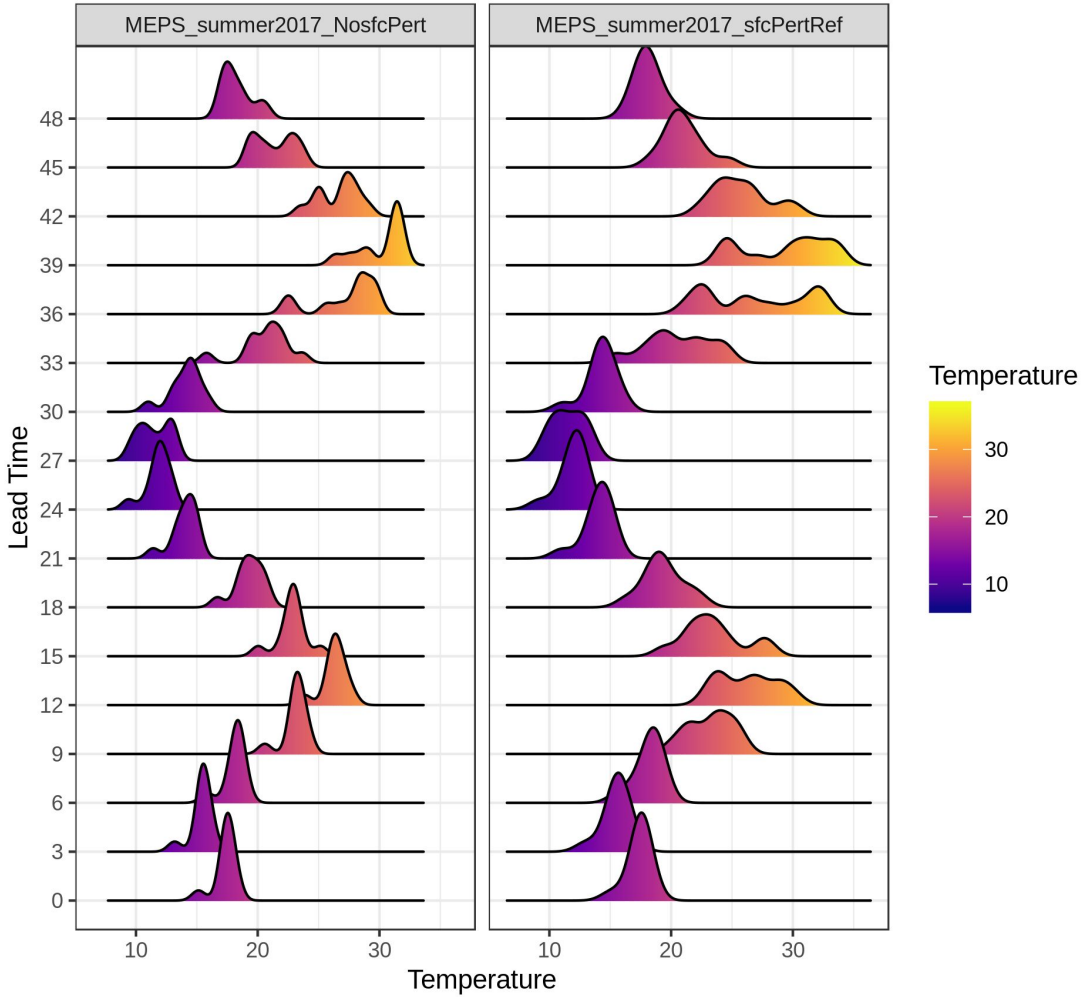

MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
35
30
25
20
15
10
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
leadtime

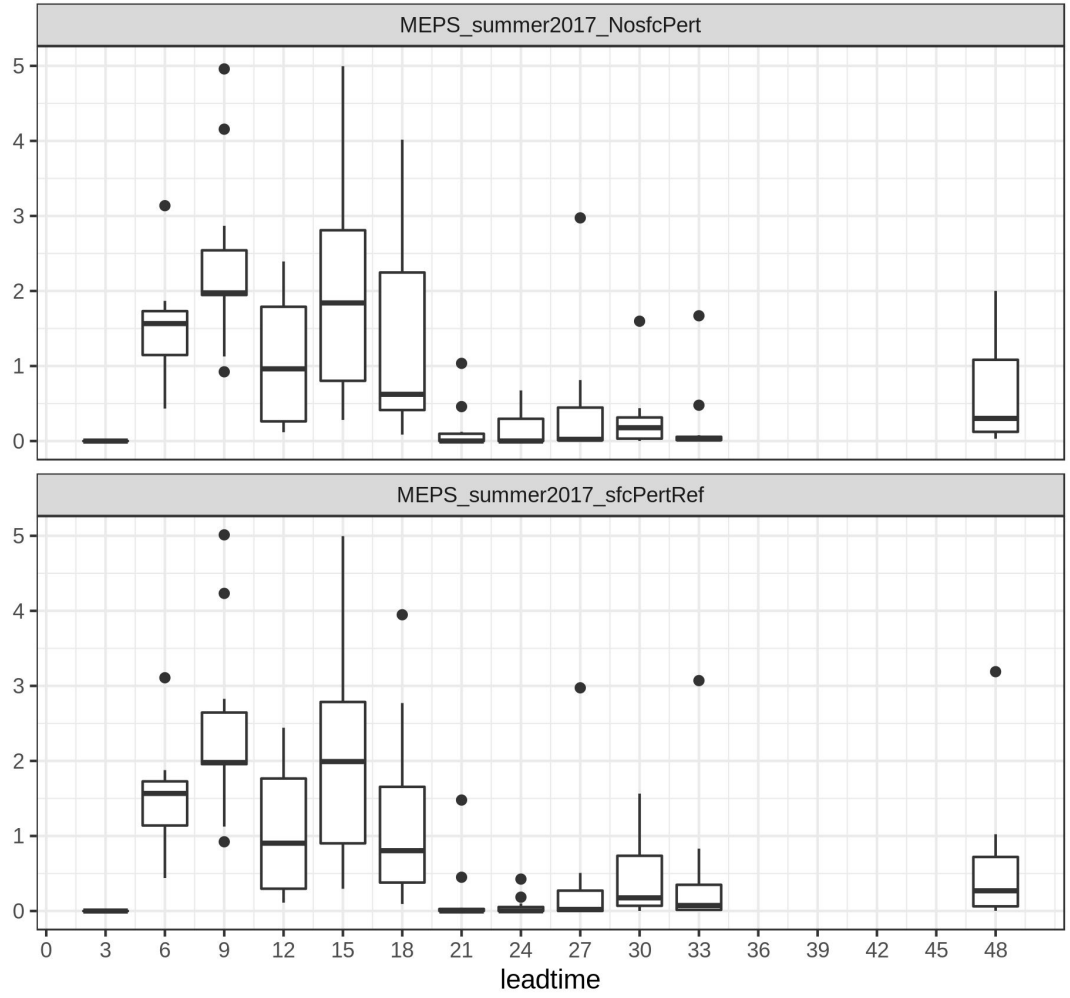

MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
30
25
20
15
10
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
leadtime

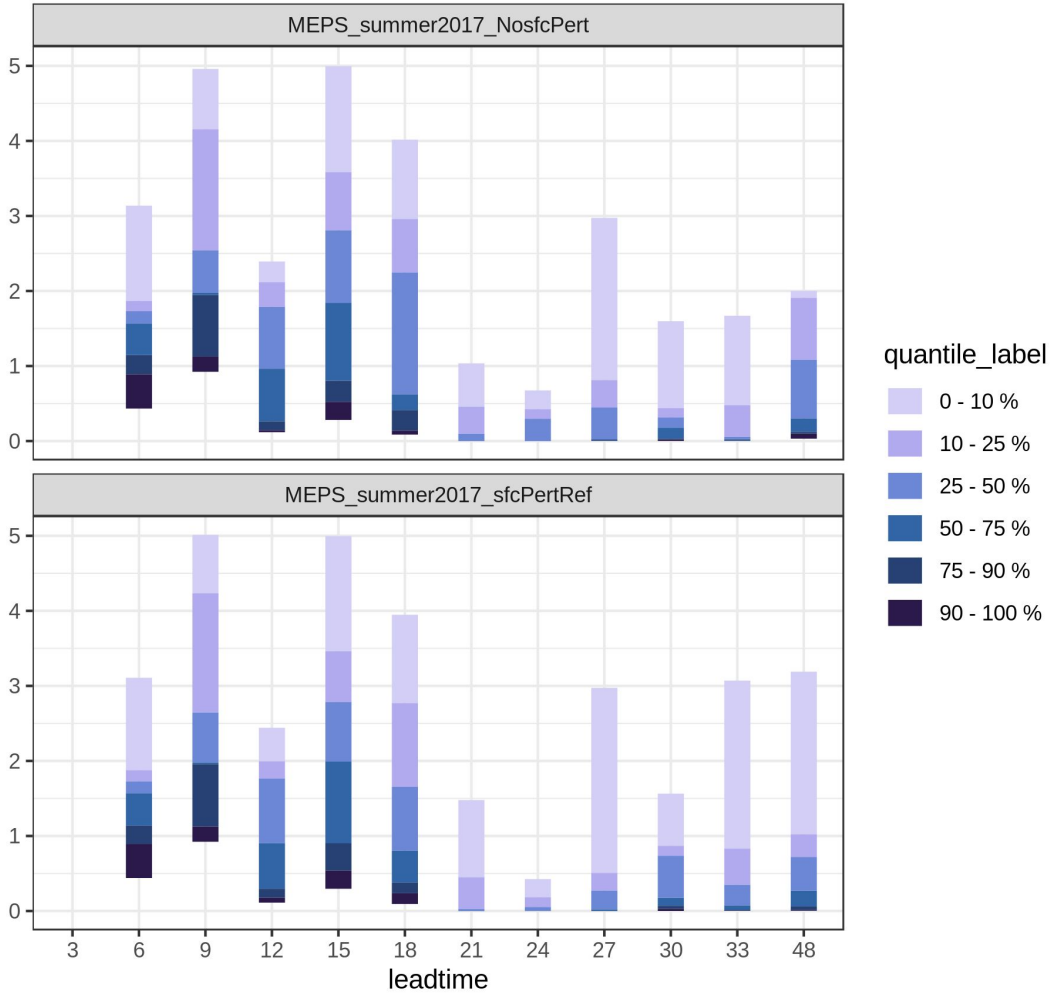

MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
30
25
20
15
10
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
leadtime


MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
30
25
20
15
10
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
leadtime


MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
35
30
25
20
15
10
0
10
20
30
40
50
leadtime


MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
Lead Time
48
45
42
39
36
33
30
27
24
21
18
15
12
9
6
3
0
10
20
30
Temperature
Temperature
30
20
10

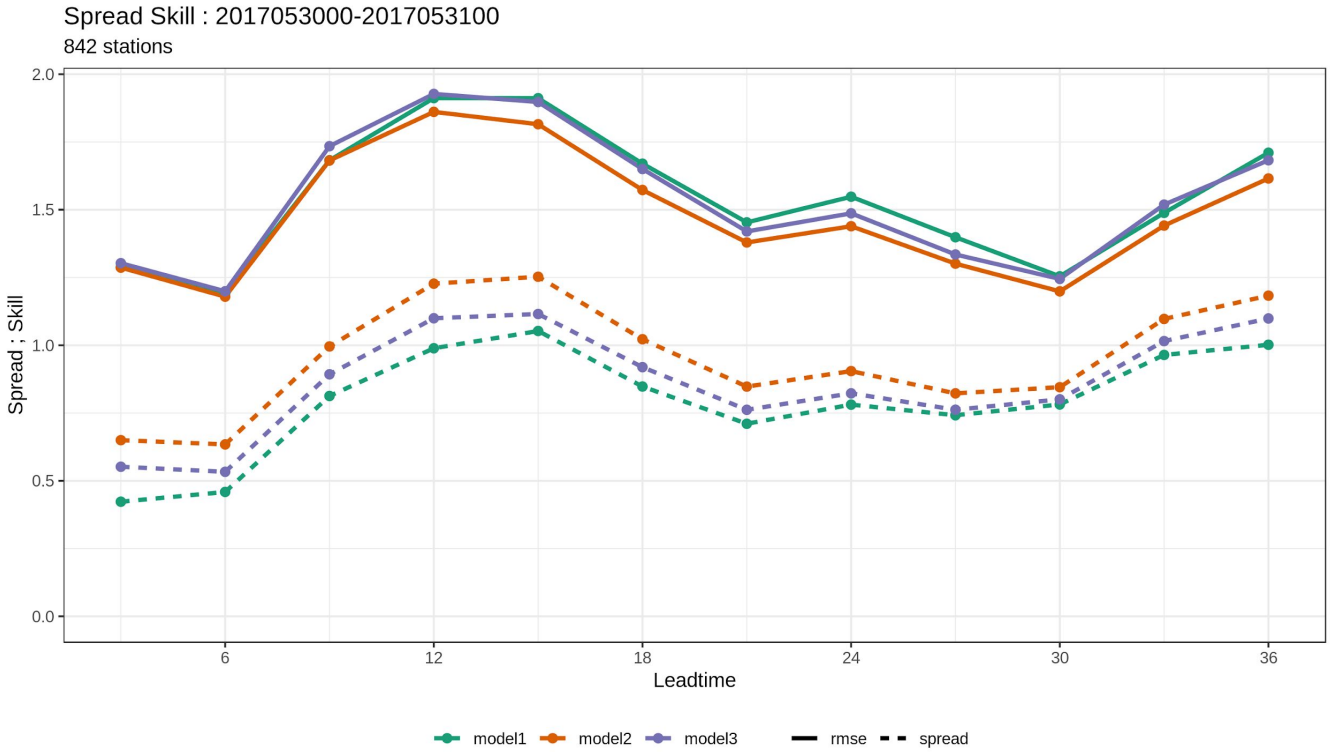

MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
0
1
2
3
4
5
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
leadtime


MEPS_summer2017_NosfcPert
MEPS_summer2017_sfcPertRef
0
1
2
3
4
5
3
6
9
12
15
18
21
24
27
30
33
48
leadtime
quantile_label
0 - 10 %
10 - 25 %
25 - 50 %
50 - 75 %
75 - 90 %
90 - 100 %

Plotting verification scores


```
plot_point_verif(...)

plot_spatial_verif(...)

plot_scatter(...)
```

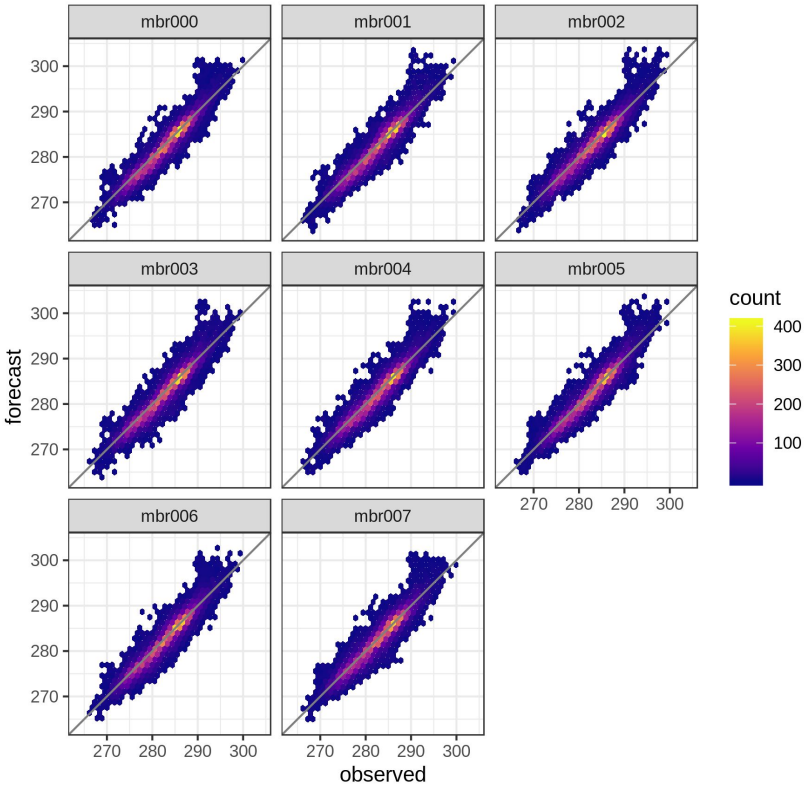
```
plot_point_verif(..., spread_skill)
```



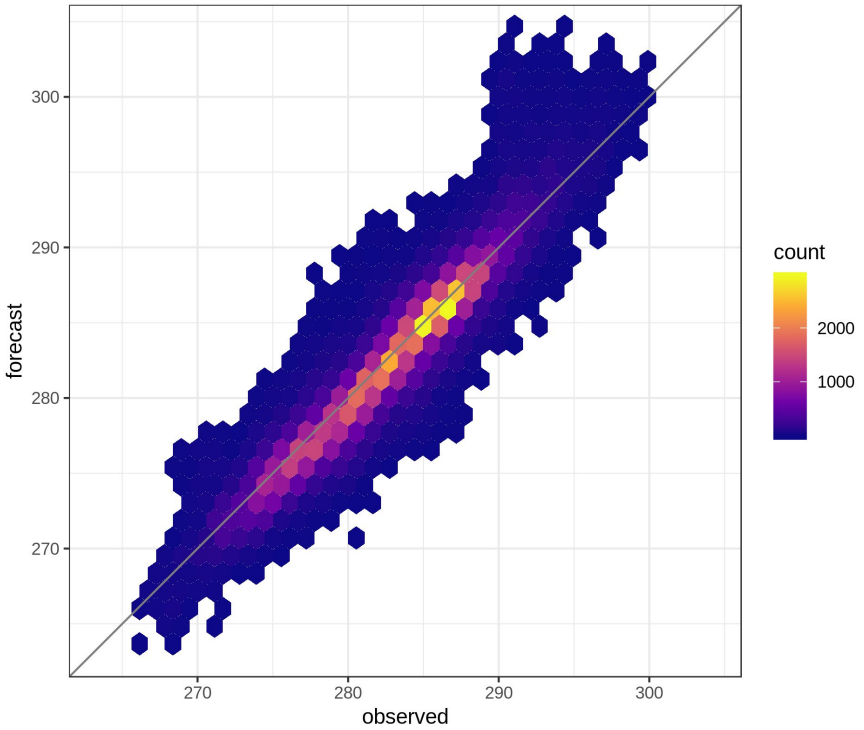
Verification for T2m

```
plot_point_verif(..., rank_histogram, facet_by = vars(leadtime))
```

```
plot_scatter(...)
```

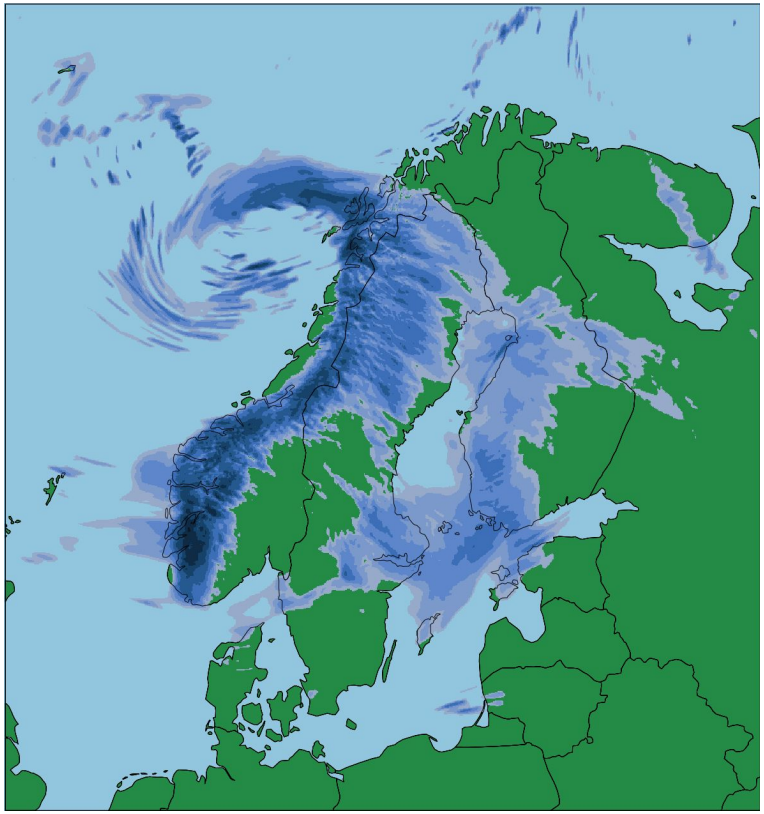


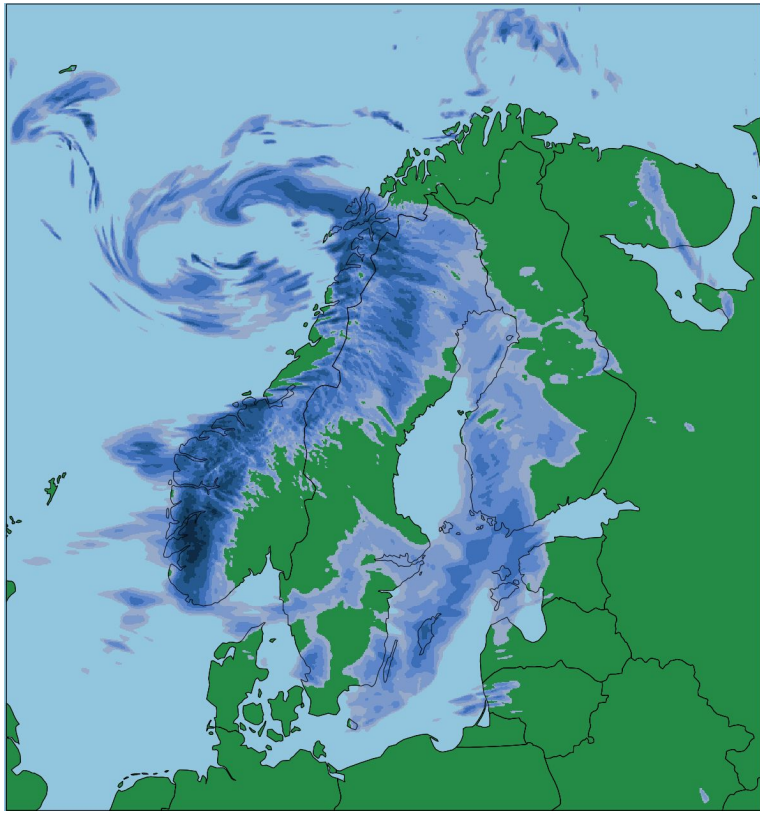
```
plot_scatter(..., facet_members = FALSE)
```



Maps*

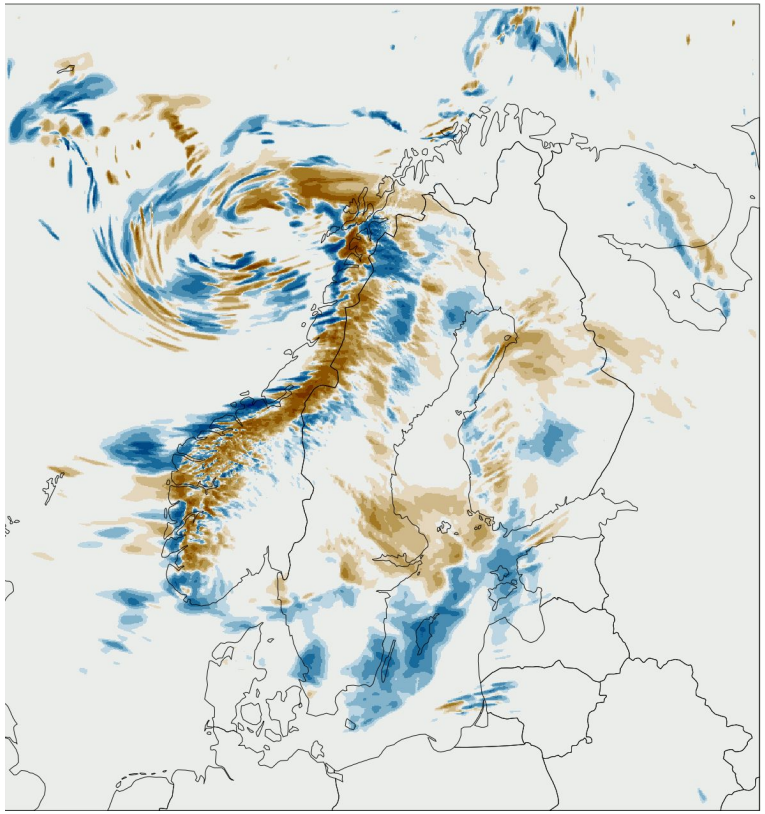
```
plot_field(..., member = c(0, 3))
```





Maps*

```
plot_field(..., member = c(0, 3), diff = TRUE)
```



Cross sections*

```
plot_x_section(..., point1 = ..., point2 = ...)
```

Interactive visualisation

```
shiny_plot_point_verif(options_bar = "all_dates")
```

127.0.0.1:4448

harp : : Point Verification

Select Verification Directory

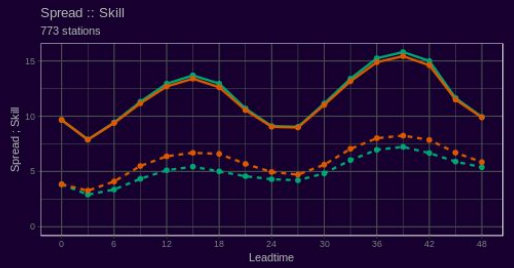
Model combination: MEPS_summer2017_NosfcPert + MEPS_summer2017_sfcPertRef

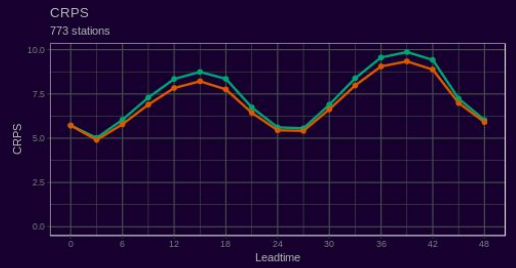
Dates: 2017052600-2017061300

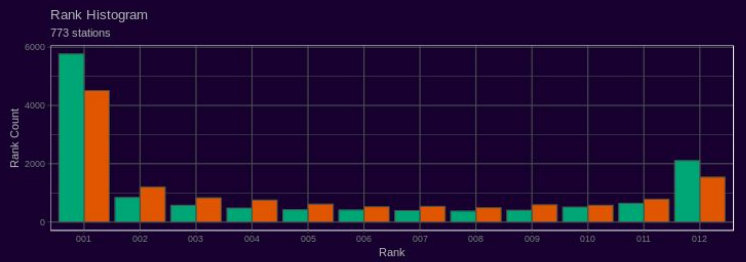
Parameter: RH2m

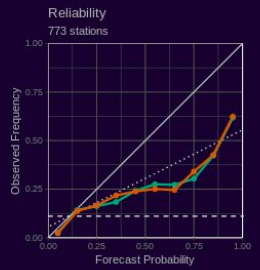
Load

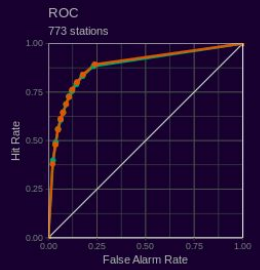
Dashboard | Interactive

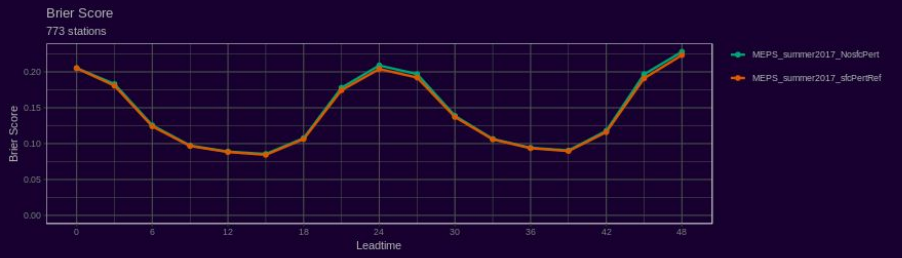












Do I need to know R?

- A tiny bit will help, but just learning how to use the harp functions should be enough. Becoming proficient opens up many more possibilities.
- harp documentation mainly through examples.
- Excellent free online materials for learning R
- harp training course in Autumn(?)