Code merge - Scatterplots

Emmanuelle Comets

26/01/2021

Summary

Objective

Check scatterplots produced by **npde.plot.scatter** and auxiilary functions.

Current issues

- sorting the covariates in the facet plot
- labels on axes, titles
- box.plot: the width of the boxplots doesn't seem right (much thicker at the beginning but the groups should be equal size)

Requirements

- main function **npde.plot.scatterplot()**
  - input: npdeObject, which.x, which.y (X and Y variables to plot), ref.prof (refrence profile), "..."
  - output: one plot
- auxiliary functions **aux.npdeplot.scatter()**
  - input:
    - a dataframe obs.mat: matrix of Y observations to plot (Y= yobs, npde, npd, pd, tnpde, tnpd) versus X (X=independent variable (eg time), predictions (pred), covariates (cov)), with the following columns
      - x,y: values of X and Y
      - grp: grouping factor - used to sort groups if plot.box=TRUE
      - cens: 1 if censored, 0 otherwise
      - loq: value of LOQ for the observation (used in VPC plot to optionally plot the line y=LOQ)
      - category: value of the category ("all" or "none" if not split by a covariate)
    - pi.mat: matrix of PI and empirical percentiles to plot for each bin, with columns
    - grp: grouping factor (same as obs.mat) - not used
    - xcent: X-value used as center of each bin (one value of xcent per grp)
    - category: covariate category ("all" if over all)
    - 3 prediction intervals: pinf, pmid, psup (mid=middle, inf, sup= extreme PIs)
      - for each PI, 3 quantiles: lower, median, upper (ie usually 0.025, 0.5, 0.975)
      - for each PI, the empirical percentile for the observed data: obs.inf, obs.median, obs.sup
    - plot.opt: a list of graphical options used in the plot
  - output
    - distribution plots: histogram (aux.plot.hist), ecdf or qqplot (aux.plot.dist)
    - note: QQ-plot for uniform variable is similar to ecdf (theoretical quantiles correspond to the cumulative density function)

Defining folders, loading libraries

Loading functions

Run npde

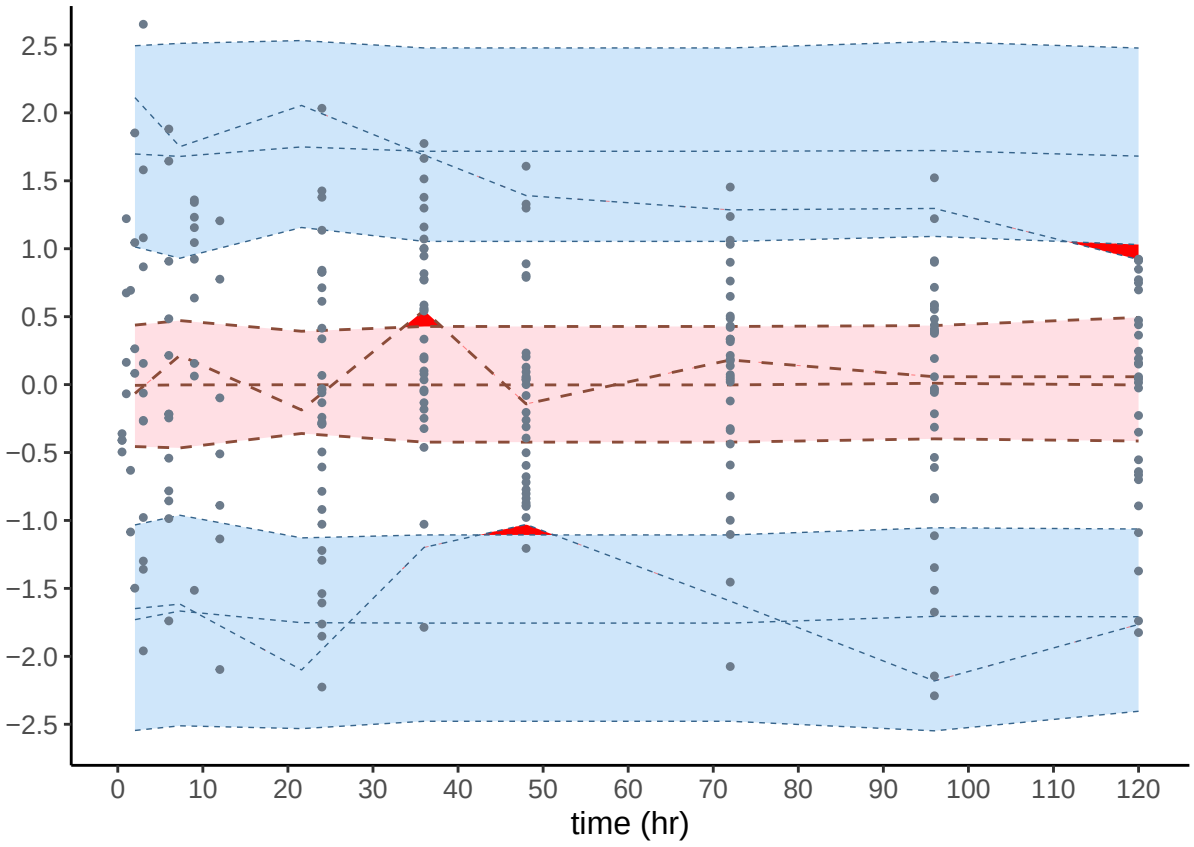
Prepare obsmat for npde/pd versus x, with or without covariates sex and weight

Prepare obsmat for npde/pd versus x, with or without covariates sex and weight

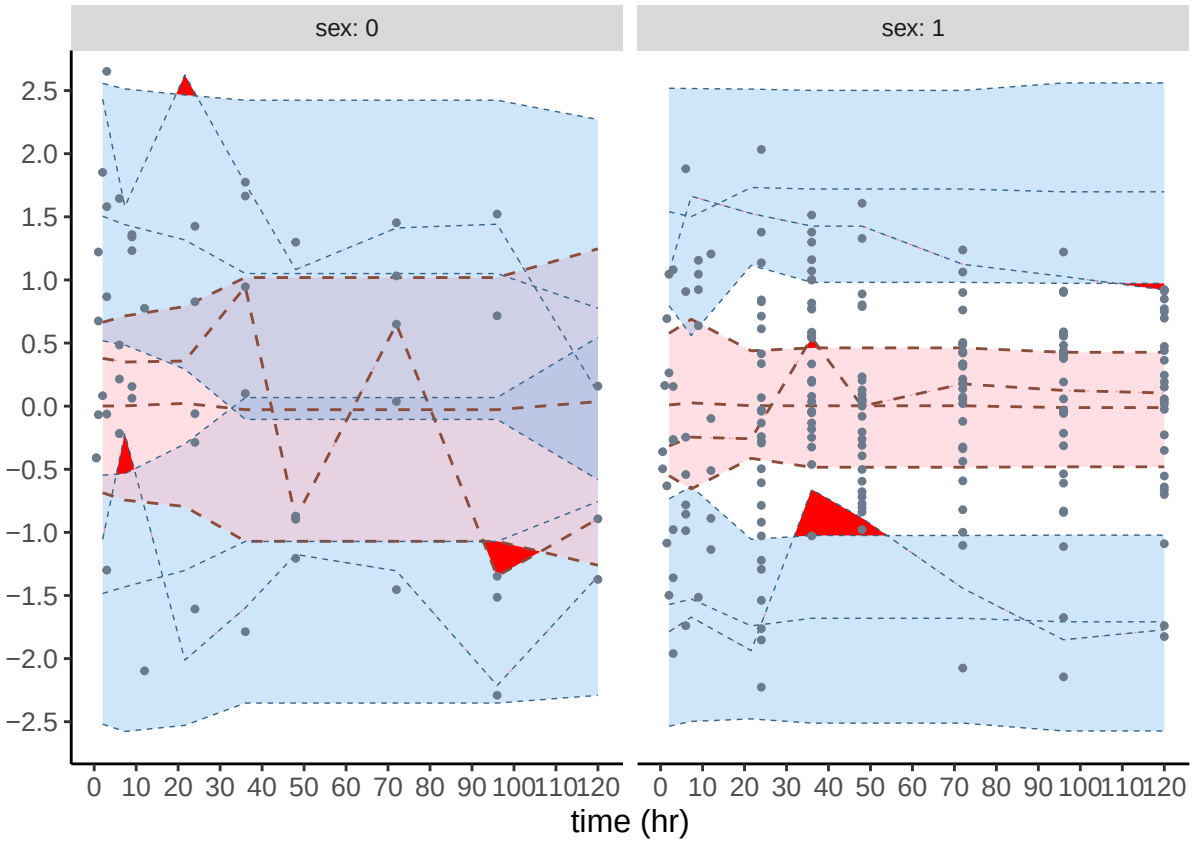
Debug functions

Debug auxiliary plots

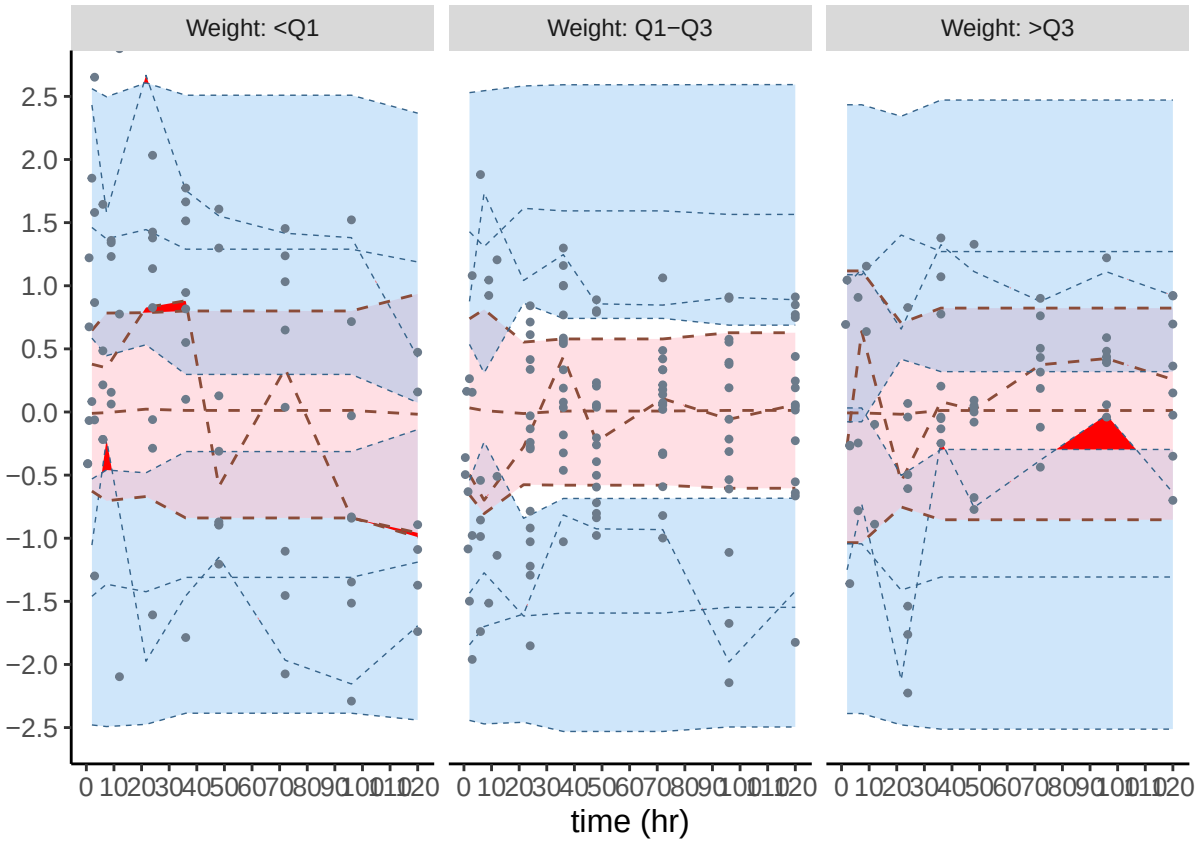
```
plot.opt2<-plot.opt
aux.npdeplot.scatter(obs.nocov, pimat.nocov, plot.opt2)
```



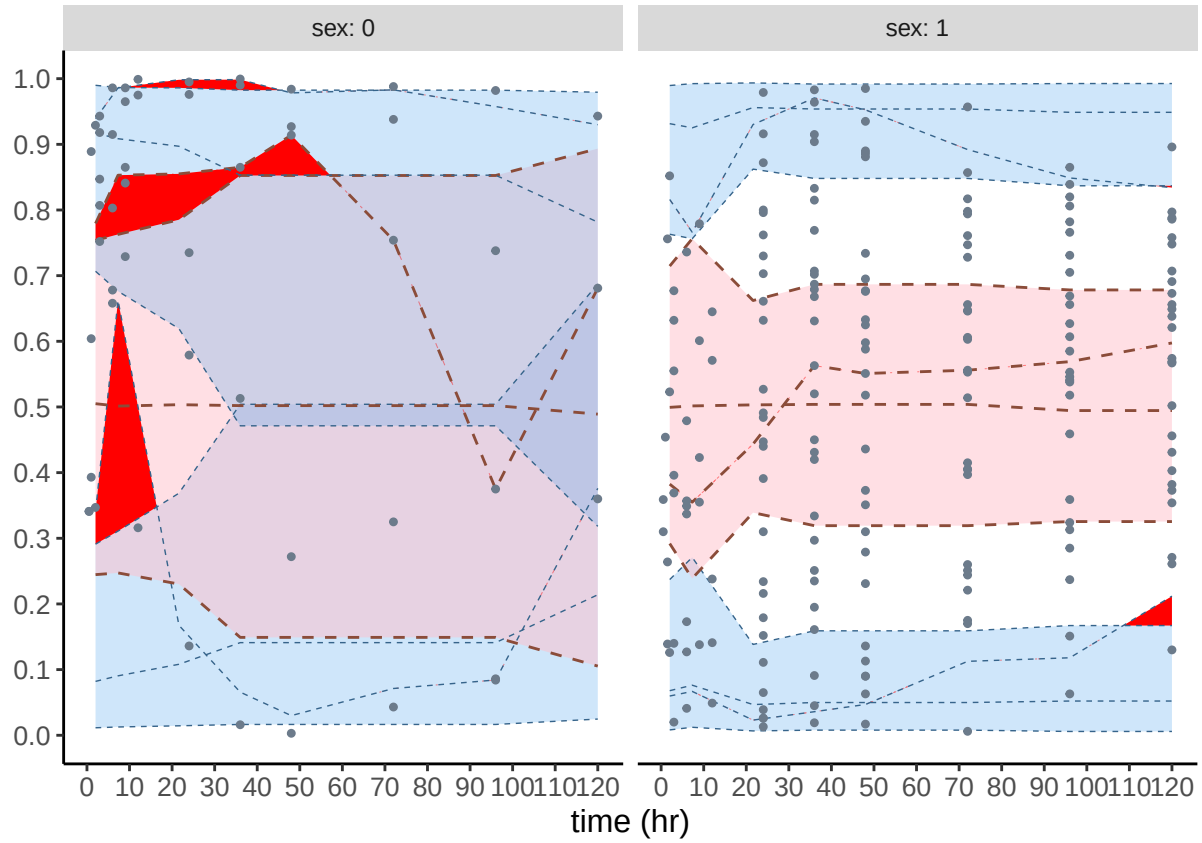

```
plot.opt2<-plot.opt
plot.opt2$which.cov<-c("sex")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(obs.sex, pimat.sex, plot.opt2)
```




```
plot.opt2<-plot.opt
plot.opt2$which.cov<-c("wt")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(obs.weight, pimat.weight, plot.opt2)
```



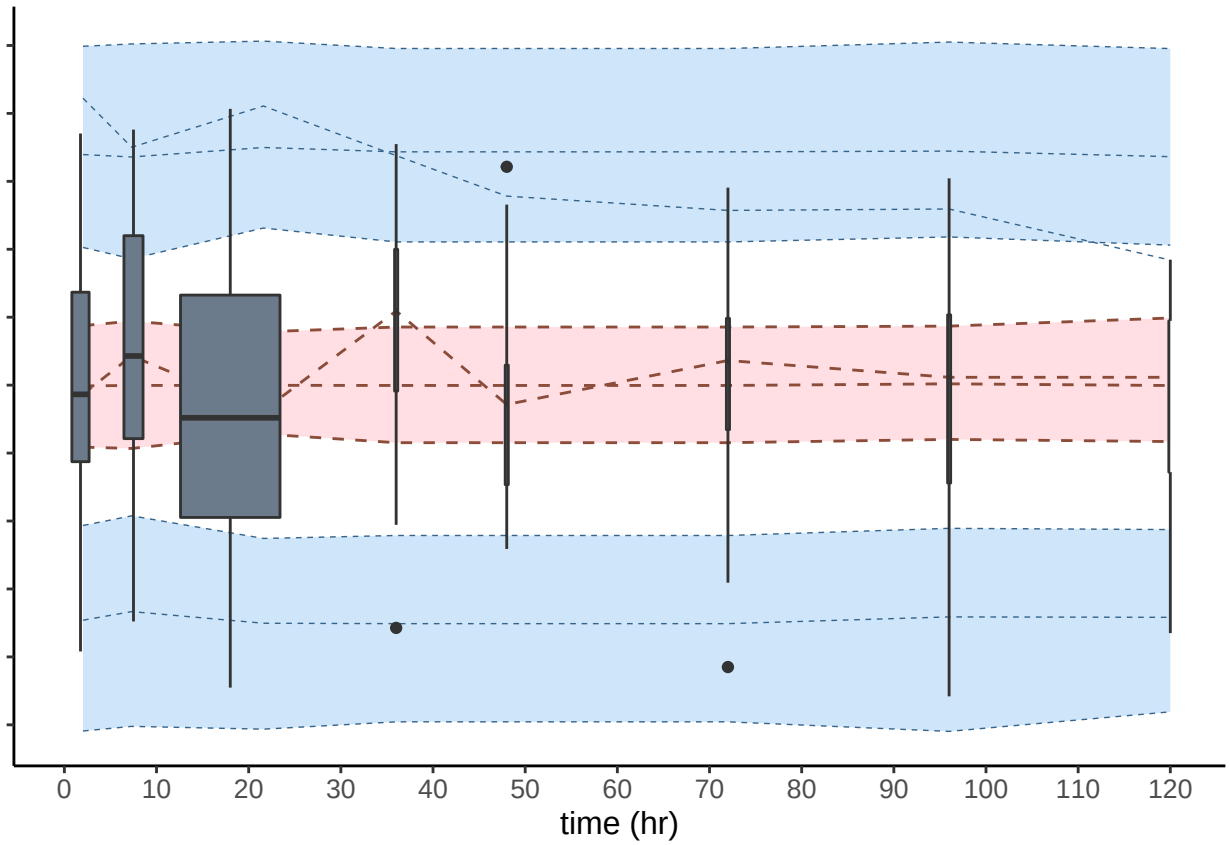

```
plot.opt2<-plot.opt
plot.opt2$which.cov<-c("sex")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(obspd.sex, pimat.pdsex, plot.opt2)
```




```
# box plots
plot.opt2<-plot.opt
plot.opt2$plot.box<-TRUE
aux.npdeplot.scatter(obs.nocov, pimat.nocov, plot.opt2)
```

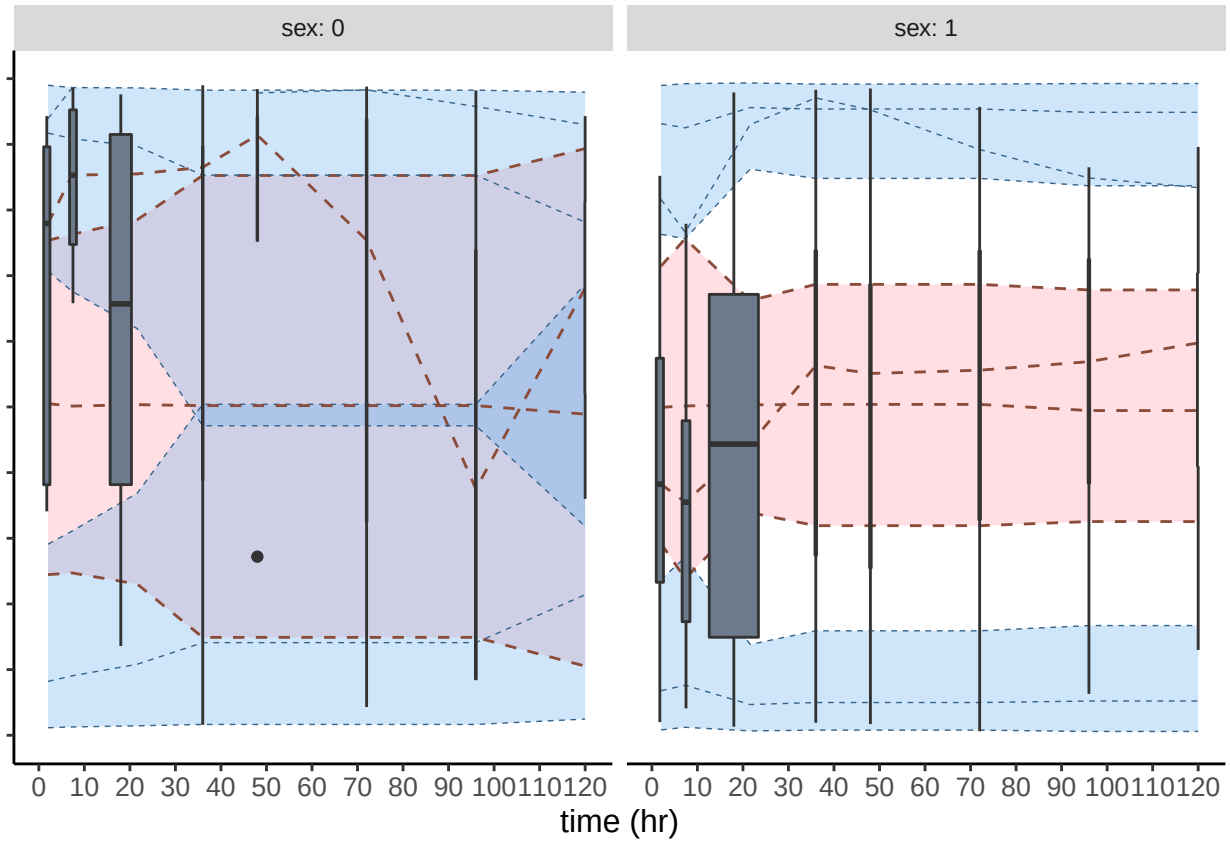
```
## Warning: Removed 2 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 1 rows containing missing values (geom_segment).
```



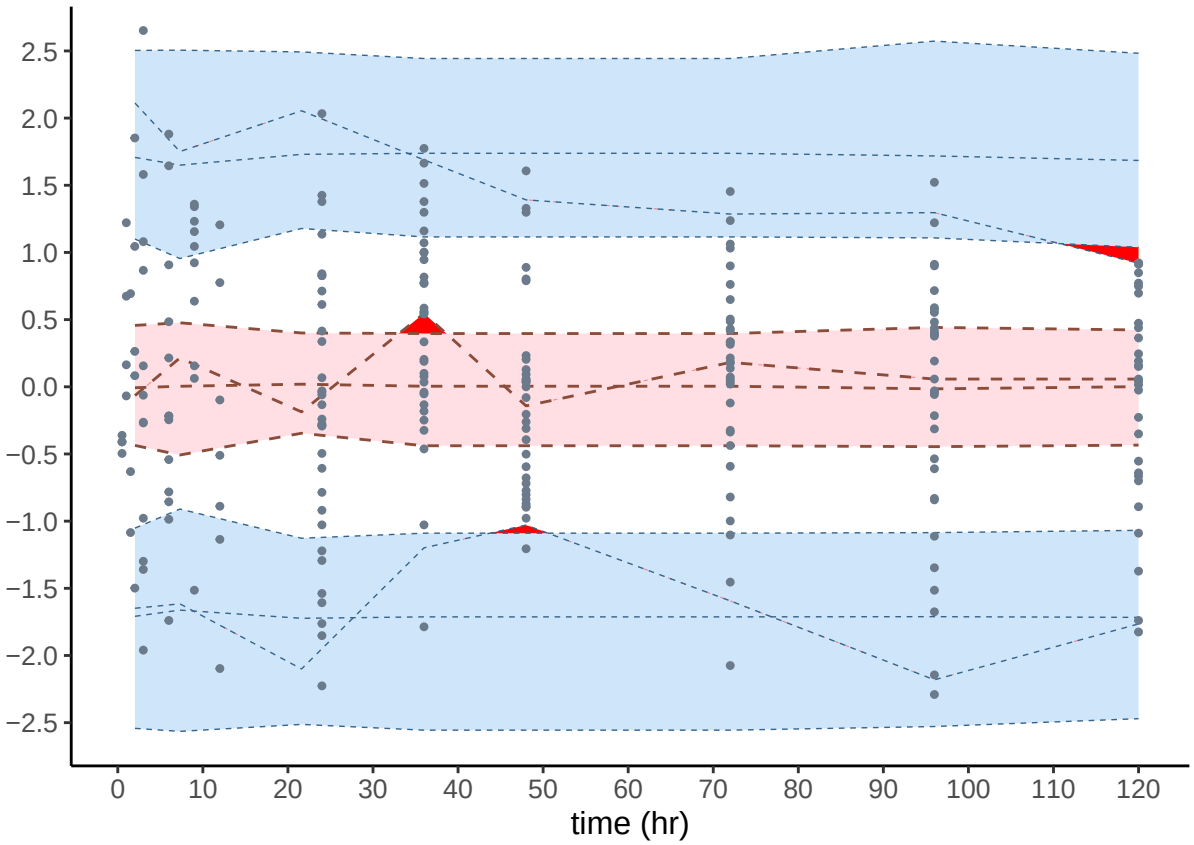

```
plot.opt2$which.cov<-c("sex")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(obspd.sex, pimat.pdsex, plot.opt2)
```

```
## Warning: Removed 4 rows containing non-finite values (stat_boxplot).

## Warning: Removed 1 rows containing missing values (geom_segment).

## Warning: Removed 1 rows containing missing values (geom_segment).
```

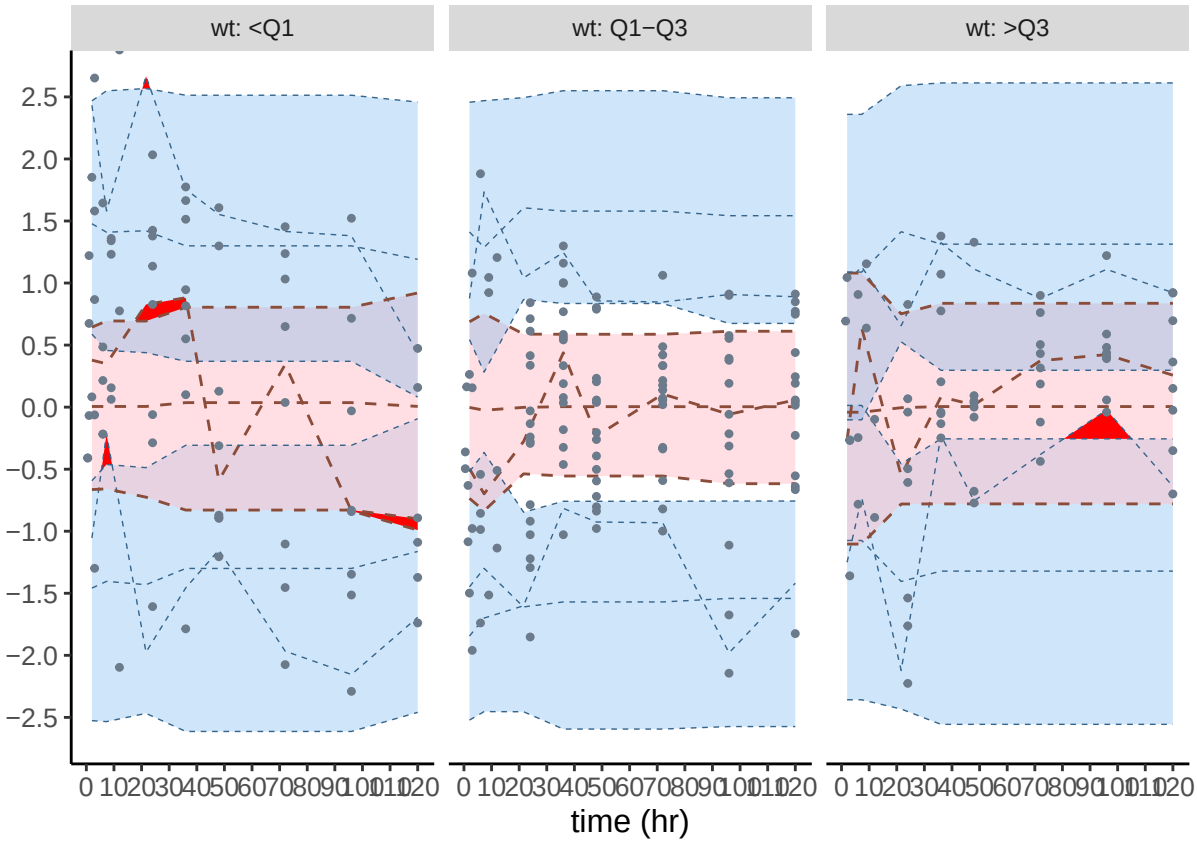


Debug main scatterplot function

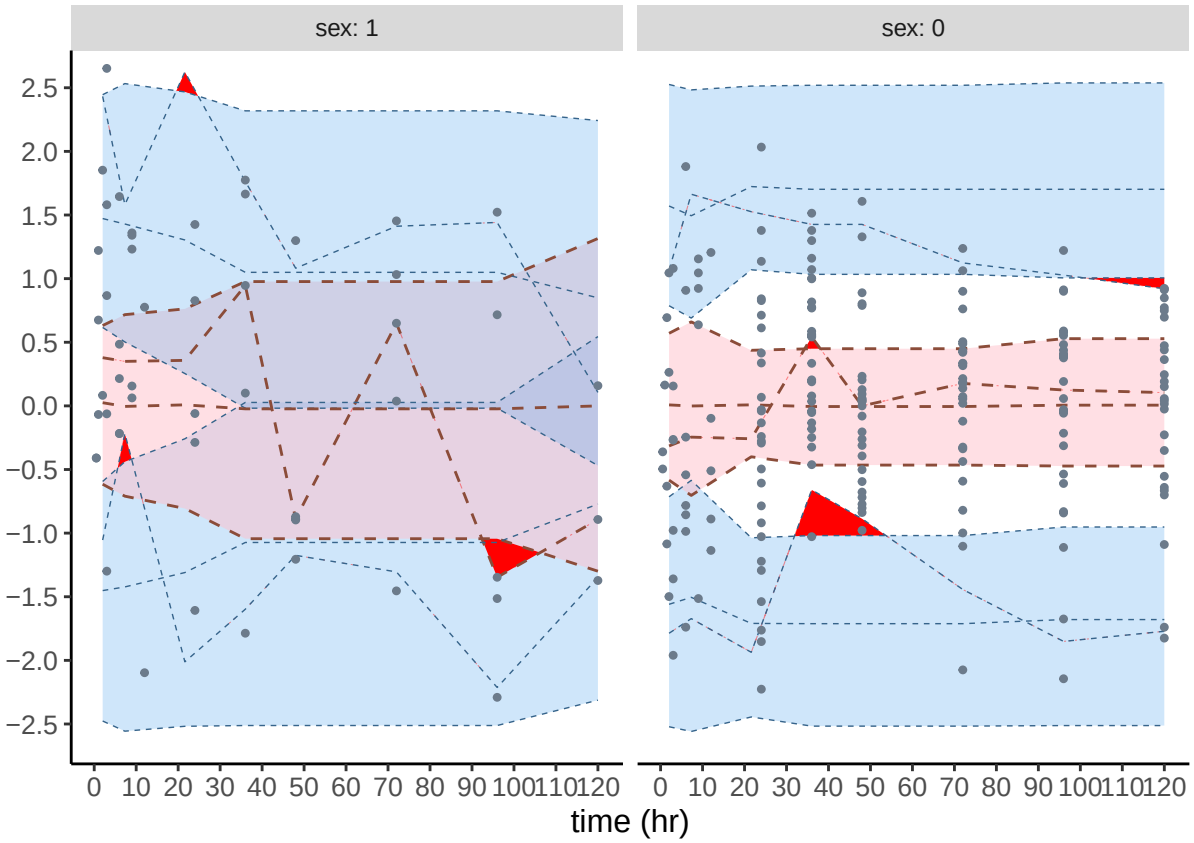
```
source(file.path(workDir,"npde","R","plotNpde-scatterplot.R")) # renamed from plotNpde-unitFunctionsAux
npde.plot.scatterplot(wbase,  which.x="x", which.y="npde")
```



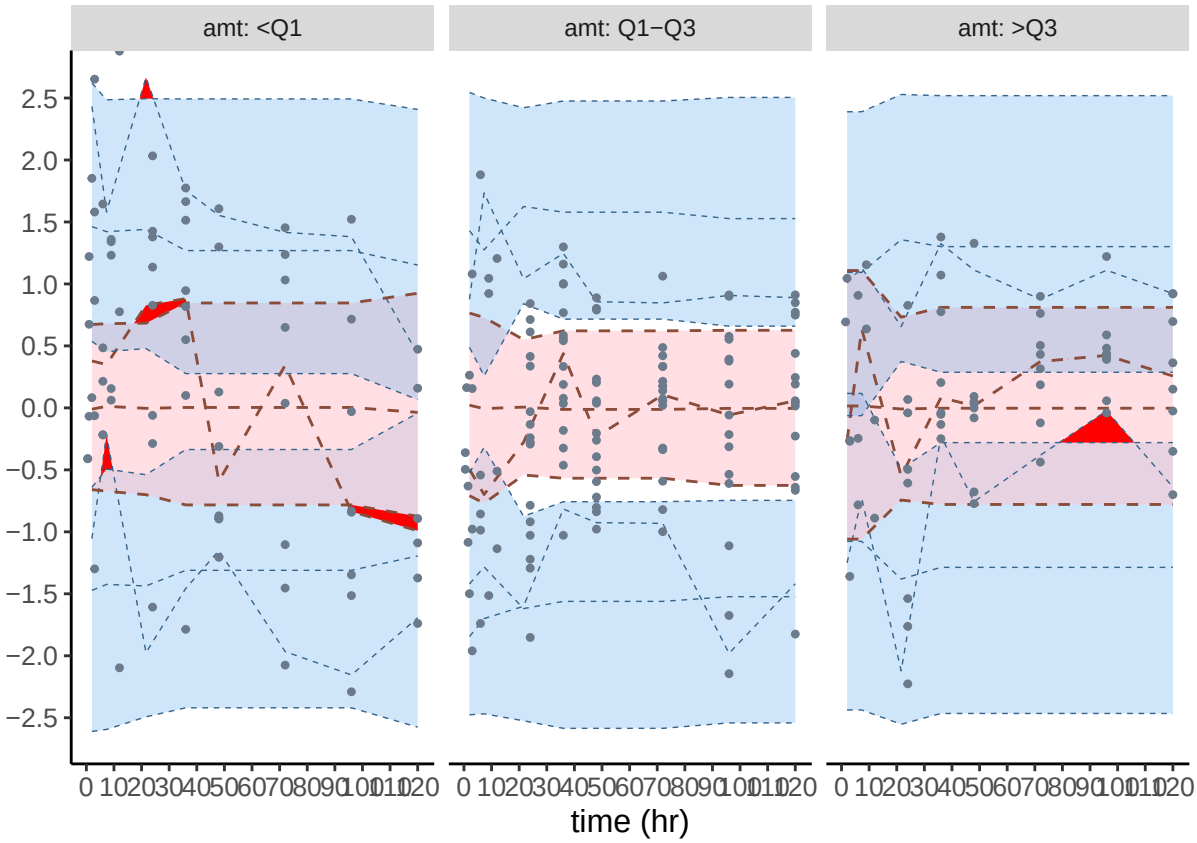

```
npde.plot.scatterplot(wbase,  which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("wt"))
```

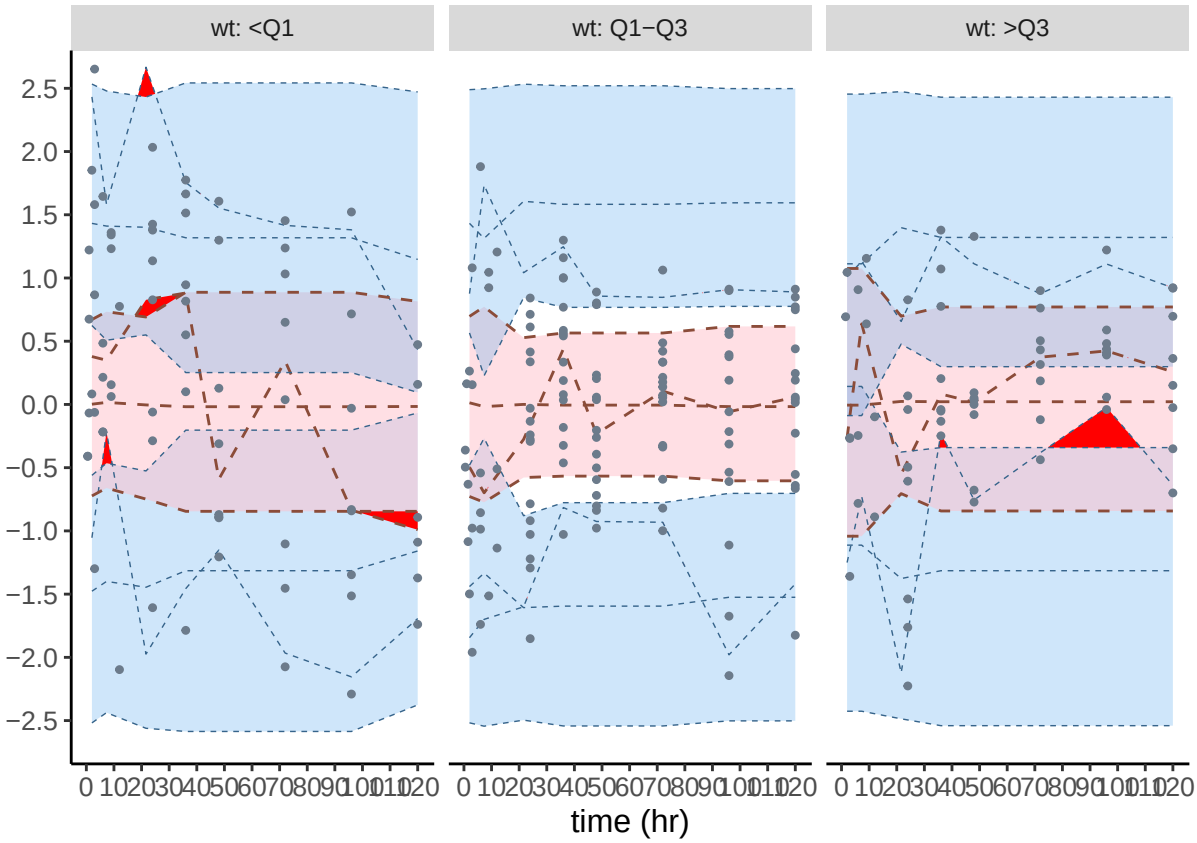


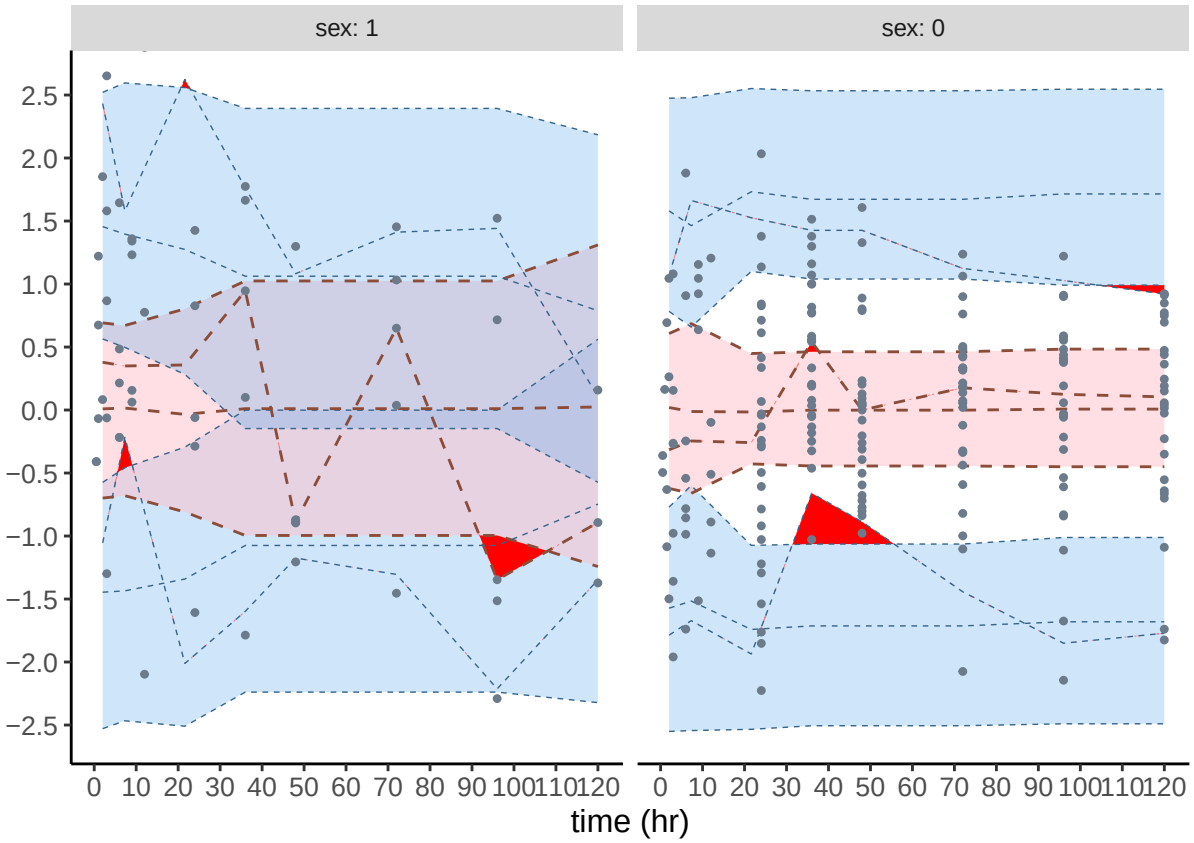

```
npde.plot.scatterplot(wbase,  which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("sex"))
```



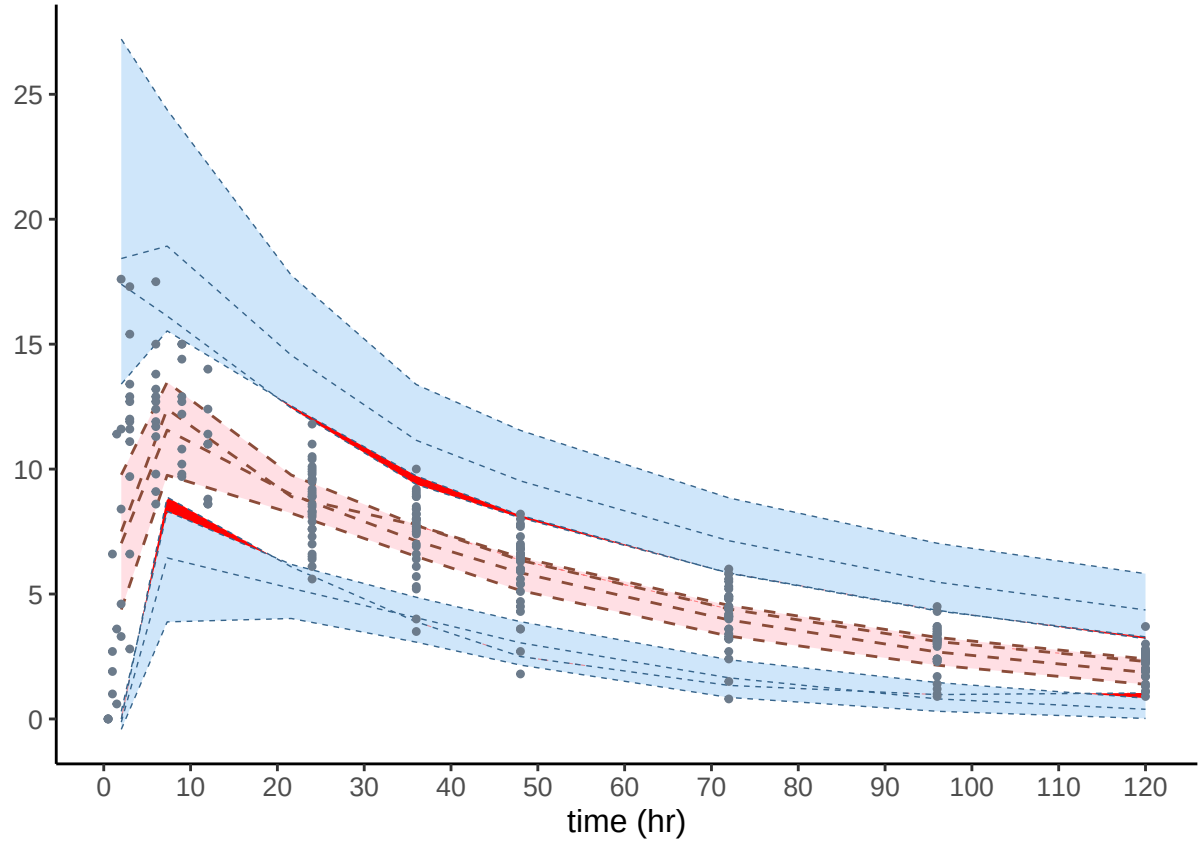

```
npde.plot.scatterplot(wbase,  which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("all"))
```


amt: <Q1
amt: Q1−Q3
amt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

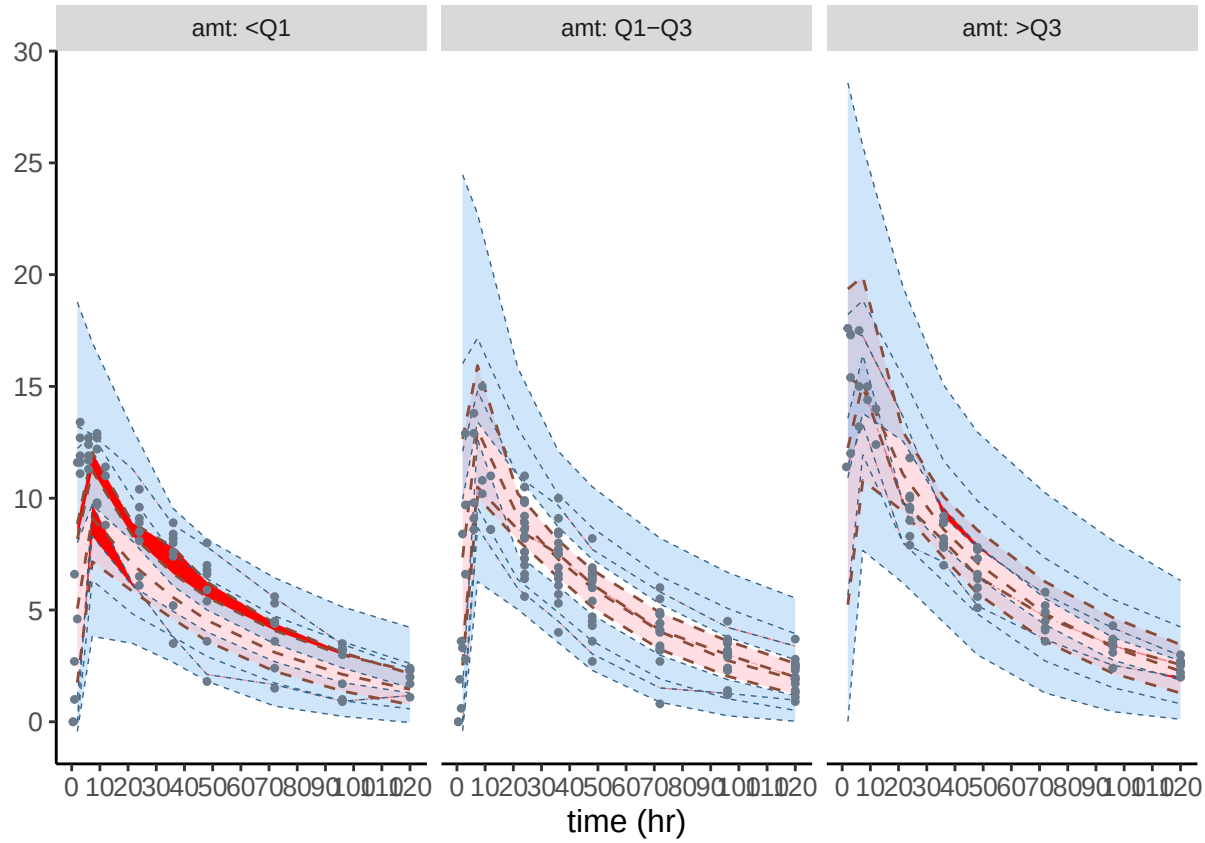

wt: <Q1
wt: Q1−Q3
wt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

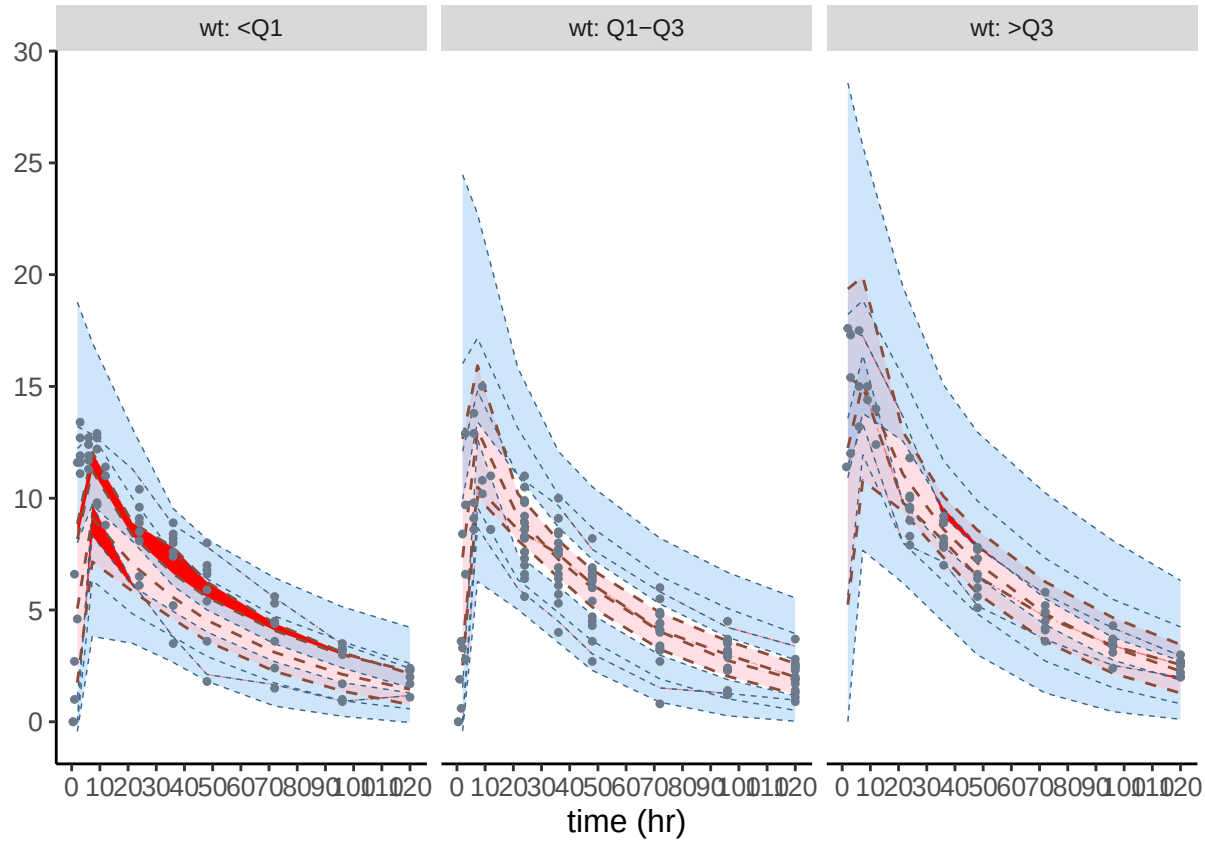


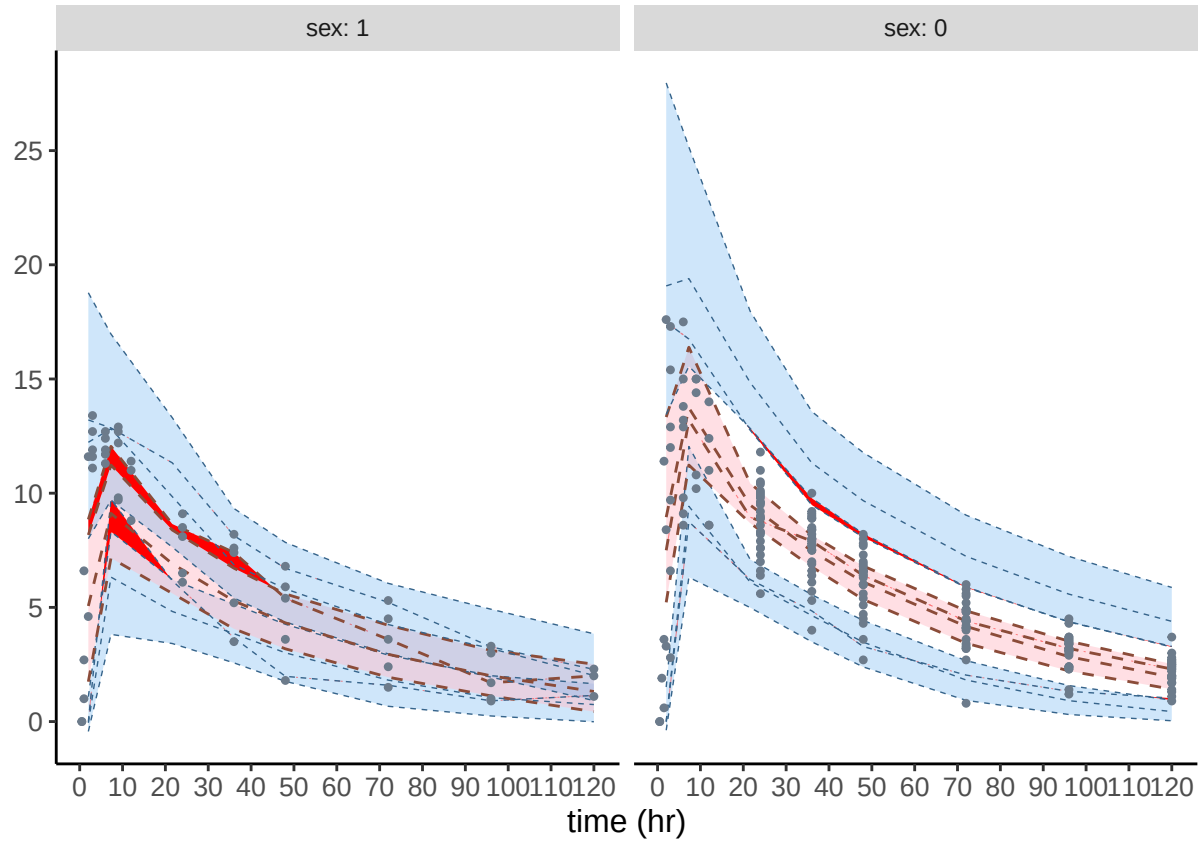

```
npde.plot.scatterplot(wbase,  which.x="x", which.y="yobs")
```




```
npde.plot.scatterplot(wbase,  which.x="x", which.y="yobs", covsplit=TRUE)
```

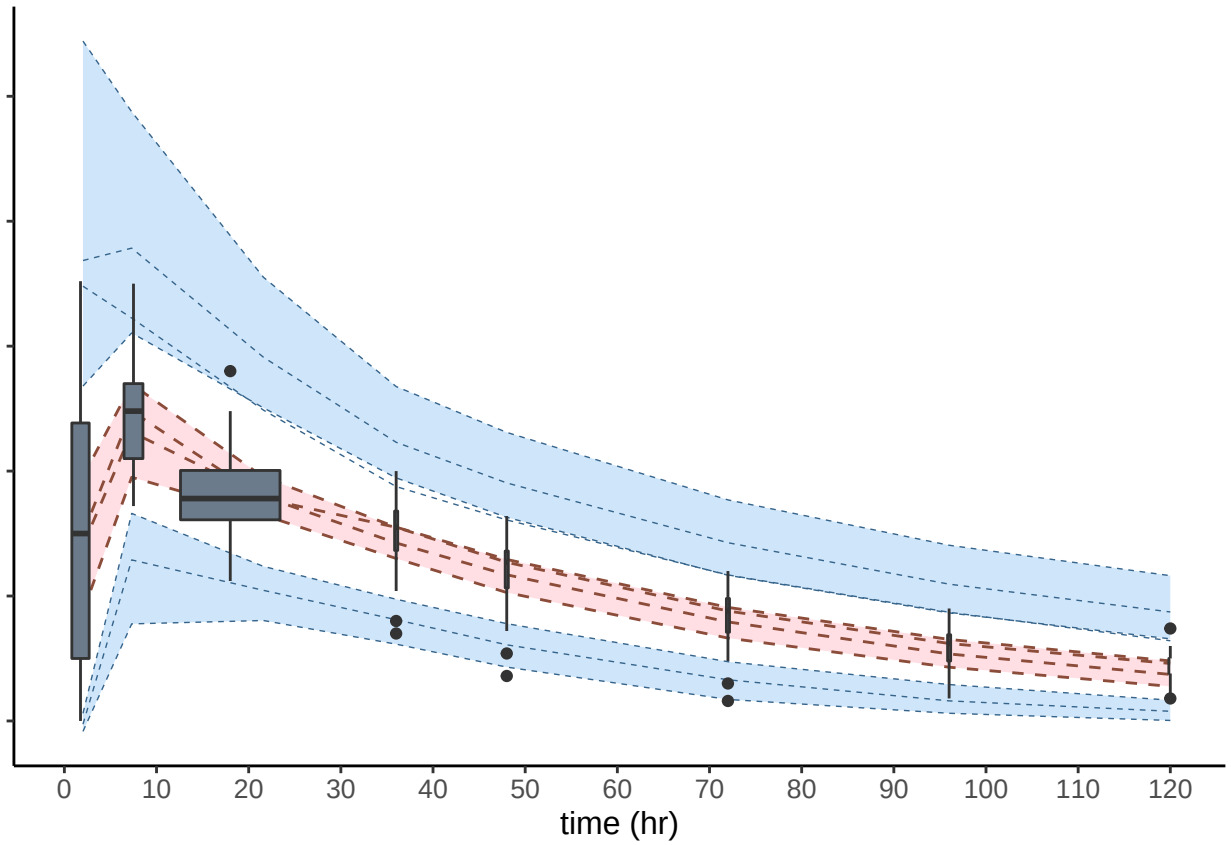

amt: <Q1
amt: Q1−Q3
amt: >Q3
30
25
20
15
10
5
0
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)


wt: <Q1
wt: Q1−Q3
wt: >Q3
30
25
20
15
10
5
0
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

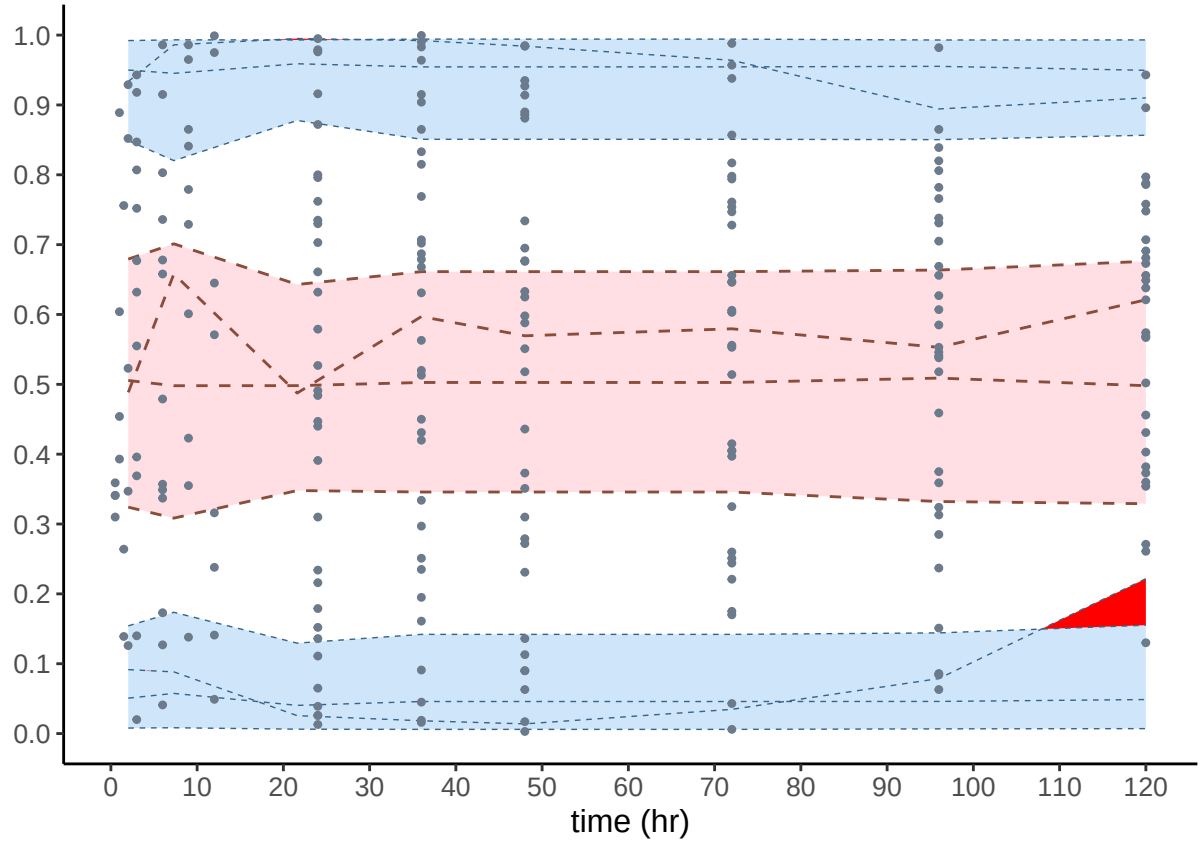



```
npde.plot.scatterplot(wbase,  which.x="x", which.y="yobs", plot.box=TRUE)
```

```
## Warning: Removed 1 rows containing missing values (geom_segment).
```

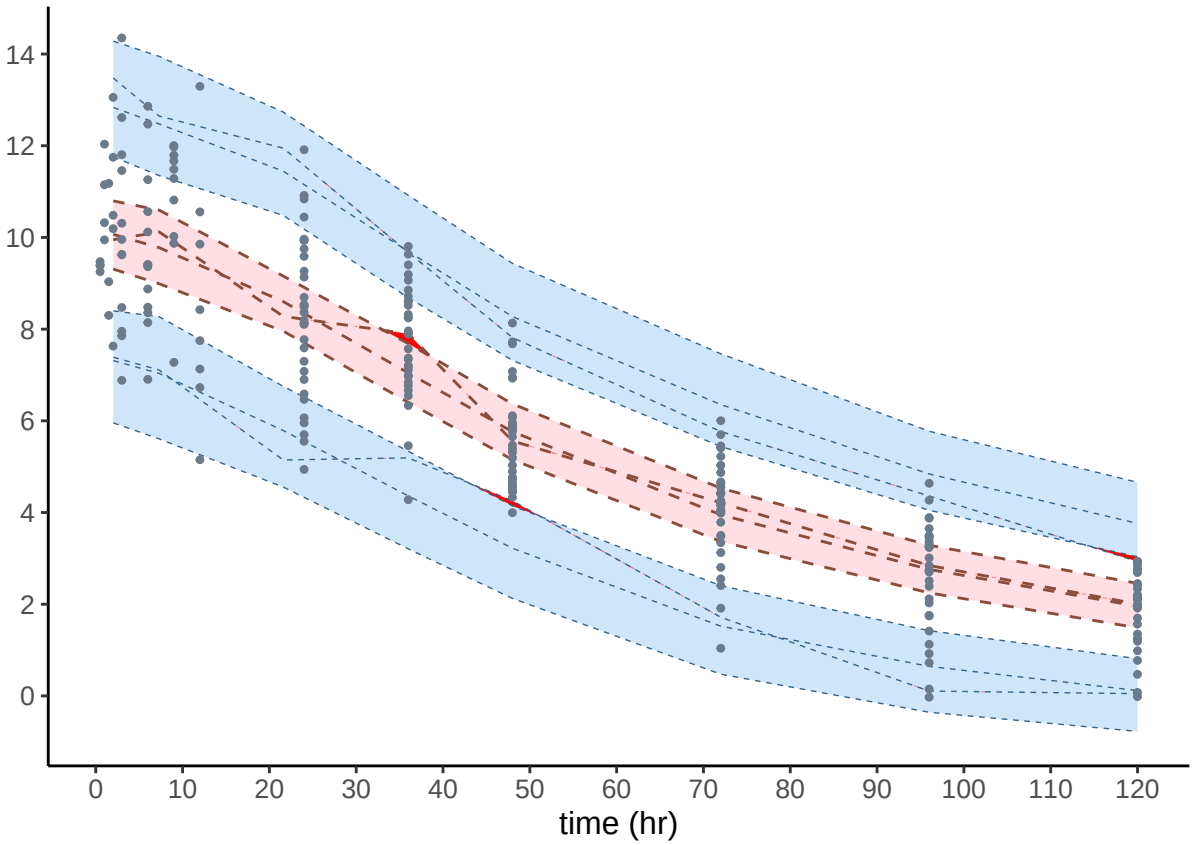



```
npde.plot.scatterplot(wbase,  which.x="x", which.y="pd")
```




```
npde.plot.scatterplot(wbase,  which.x="x", which.y="npde", ref.prof=list(id=1))
```

```
## Not all time points/bins are represented in the subset used for the reference profile: spline interpo
```


14
12
10
8
6
4
2
0
0
10
20
30
40
50
60
70
80
90
100
110
120
time (hr)