

# Code merge - Scatterplots

Emmanuelle Comets

26/01/2021

## Summary

### Objective

Check scatterplots produced by **npde.plot.scatter** and auxiliary 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 (reference 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: pminf, 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 pimat for npde/pd versus x, with or without covariates sex and weight

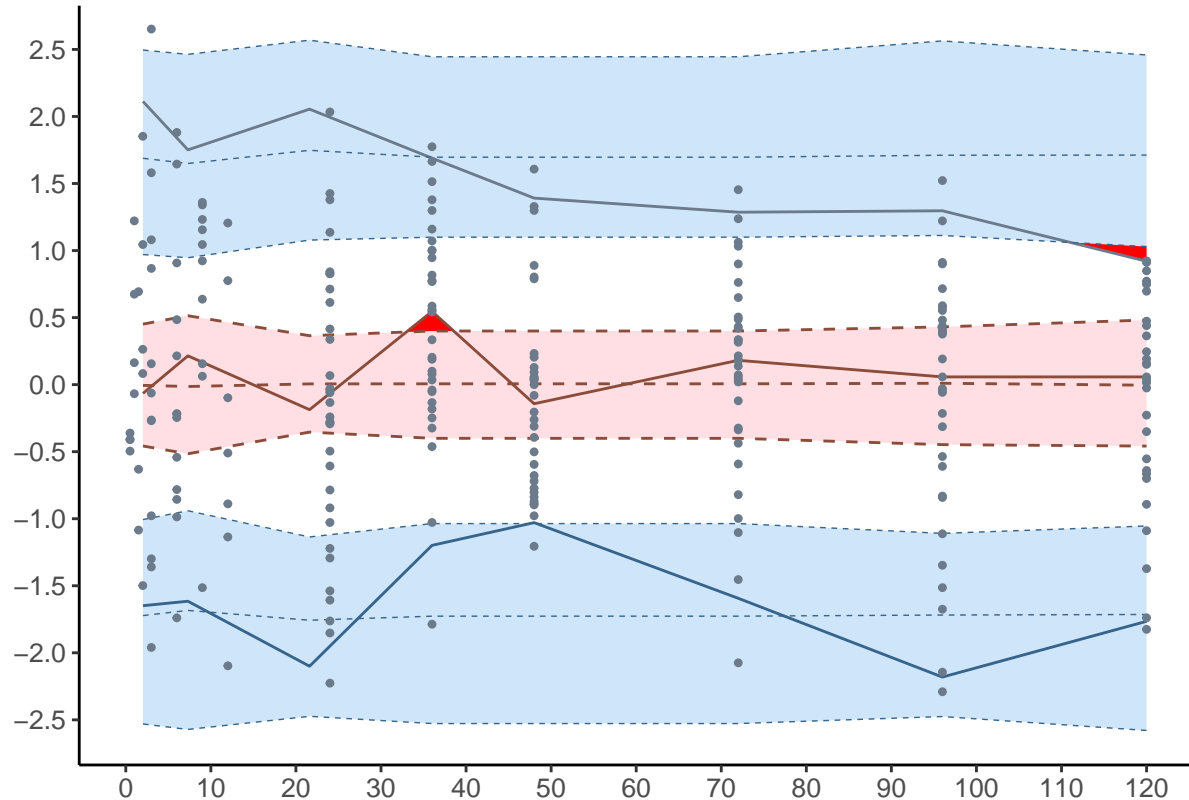
Plots for npde/pd versus cov, for covariates sex and weight

- 2 versions for continuous covariate
  - with PI for weight, computed as if  $x=\text{time}$
  - without PI for weight, weight in 3 categories
- for categorical covariate
  - remove bands and interpolation ?
  - boxplot for categorical covariate

Debug functions

Debug auxiliary plots

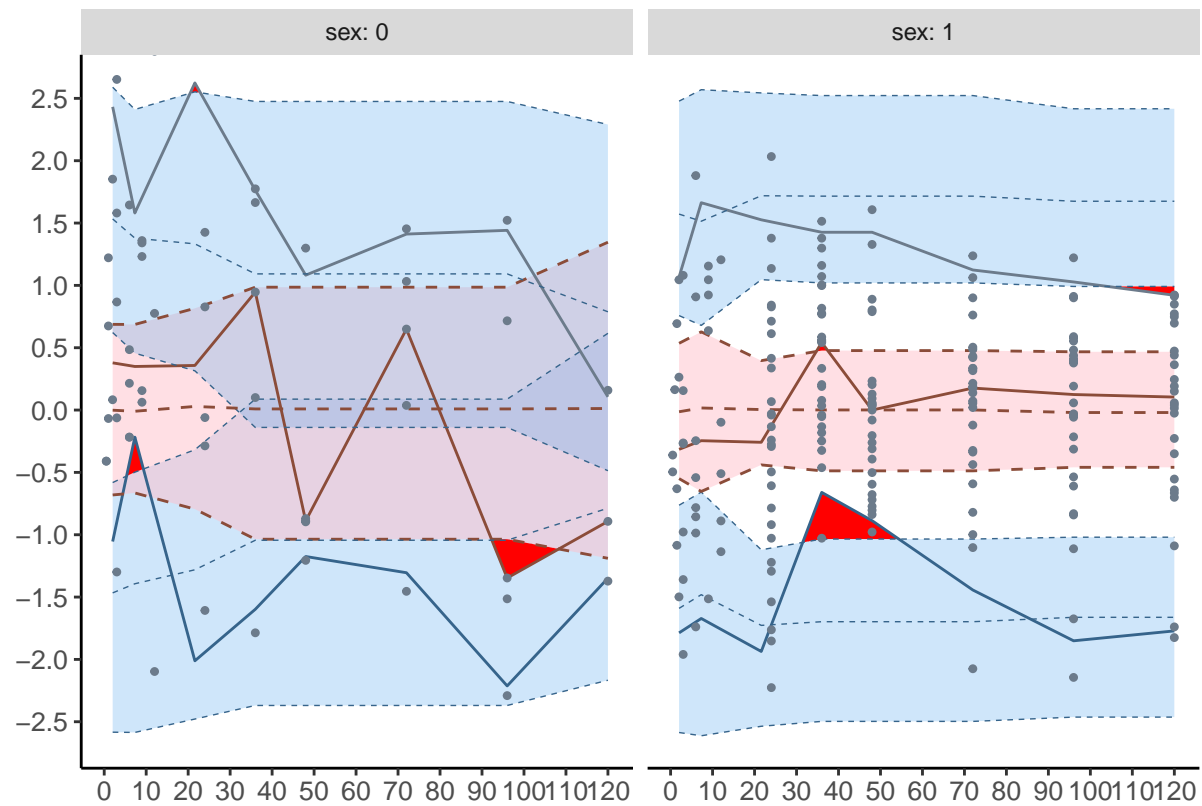
```
plot.opt2<-plot.opt  
aux.npdeplot.scatter(obs.nocov, pimat.nocov, plot.opt2)
```



```

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

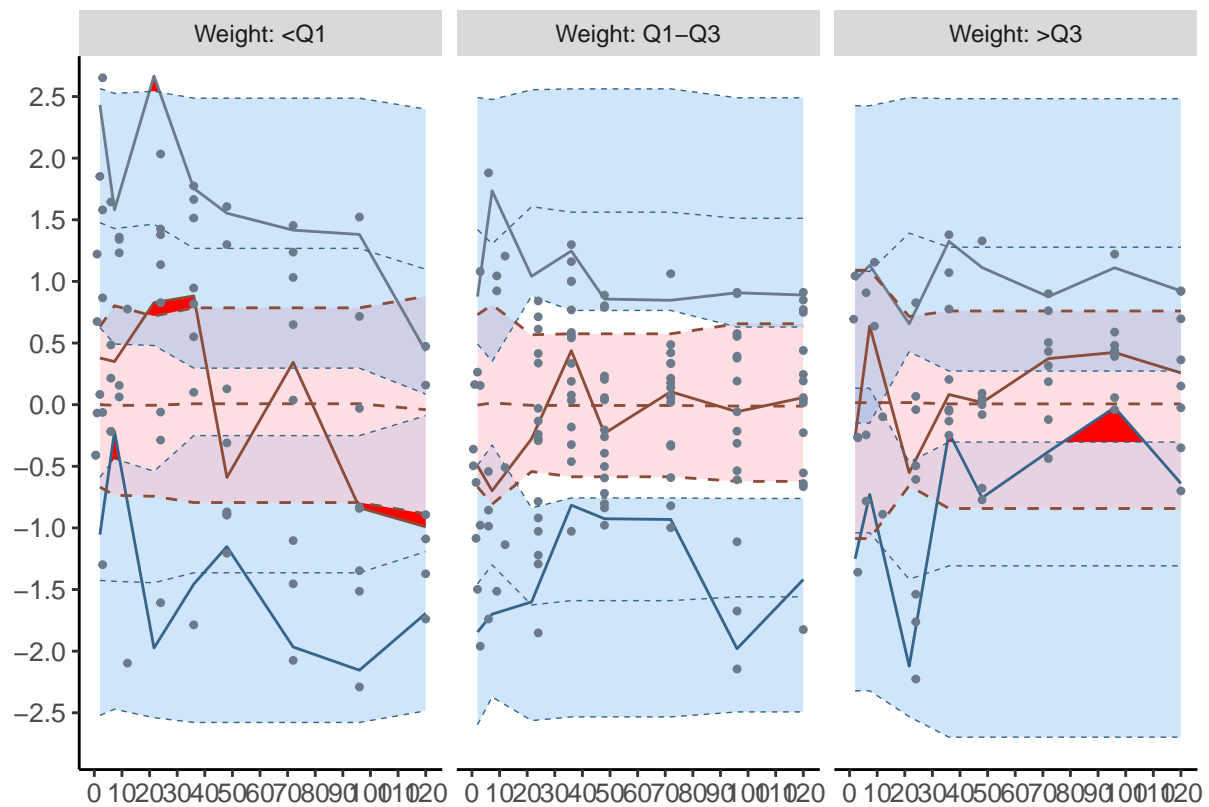
```



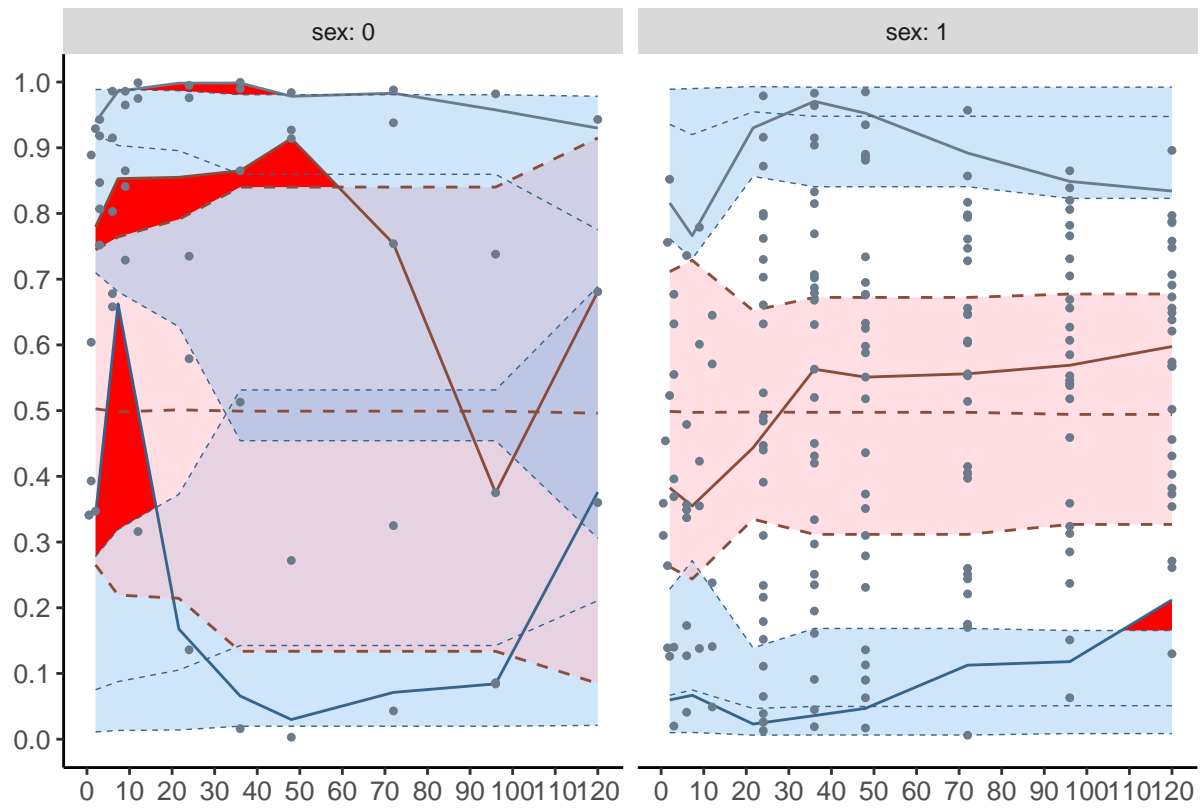
```

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

```



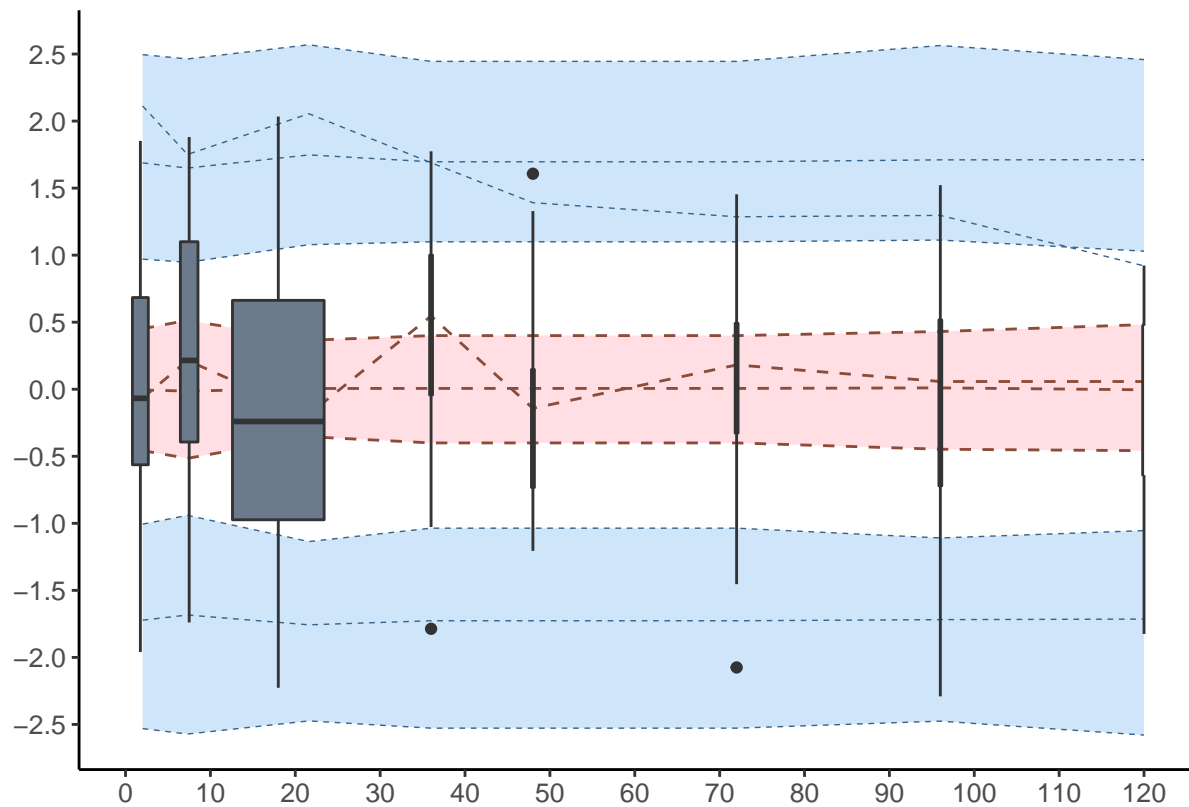
```
plot.opt2<-plot.opt
plot.opt2$which.cov<-c("sex")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(obsdpd.sex, pimat.pdsex, plot.opt2)
```



```
# 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).
```

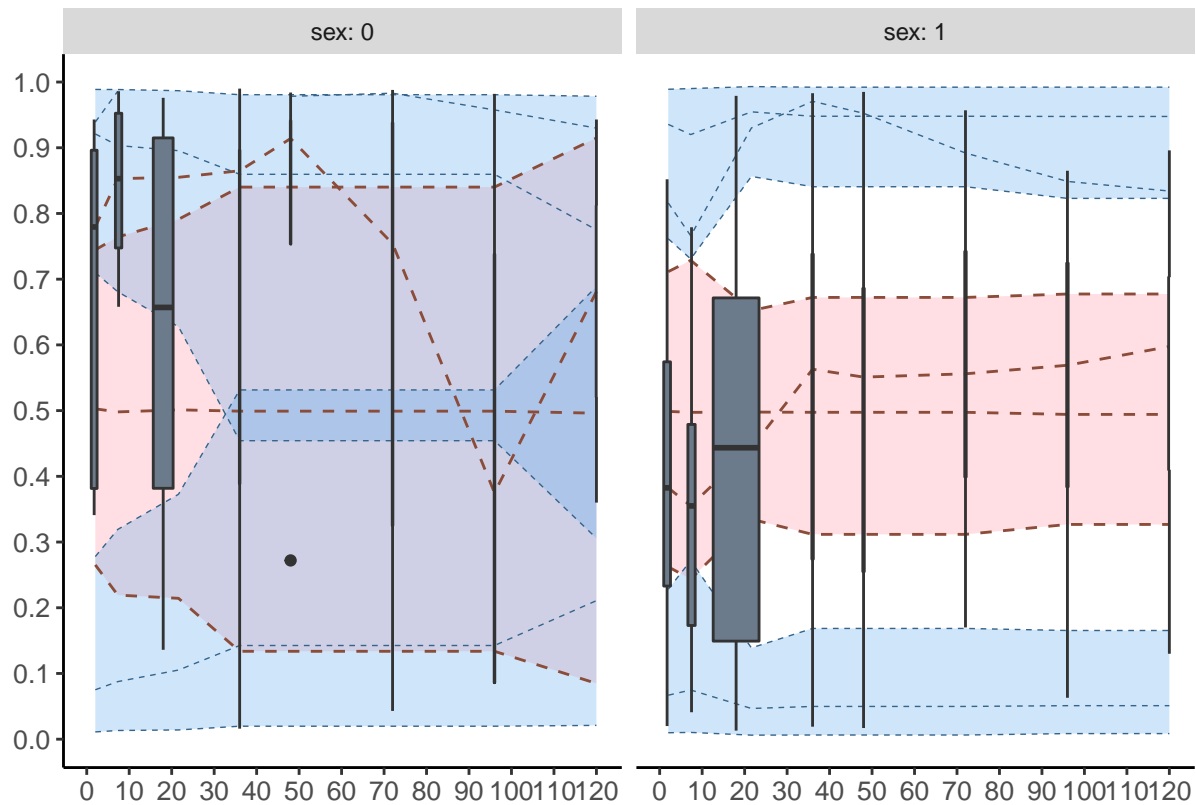


```
plot.opt2$which.cov<-c("sex")
plot.opt2$covsplit<-TRUE
aux.npdeplot.scatter(ospd.sex, pimat.pdsex, plot.opt2)
```

```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```

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

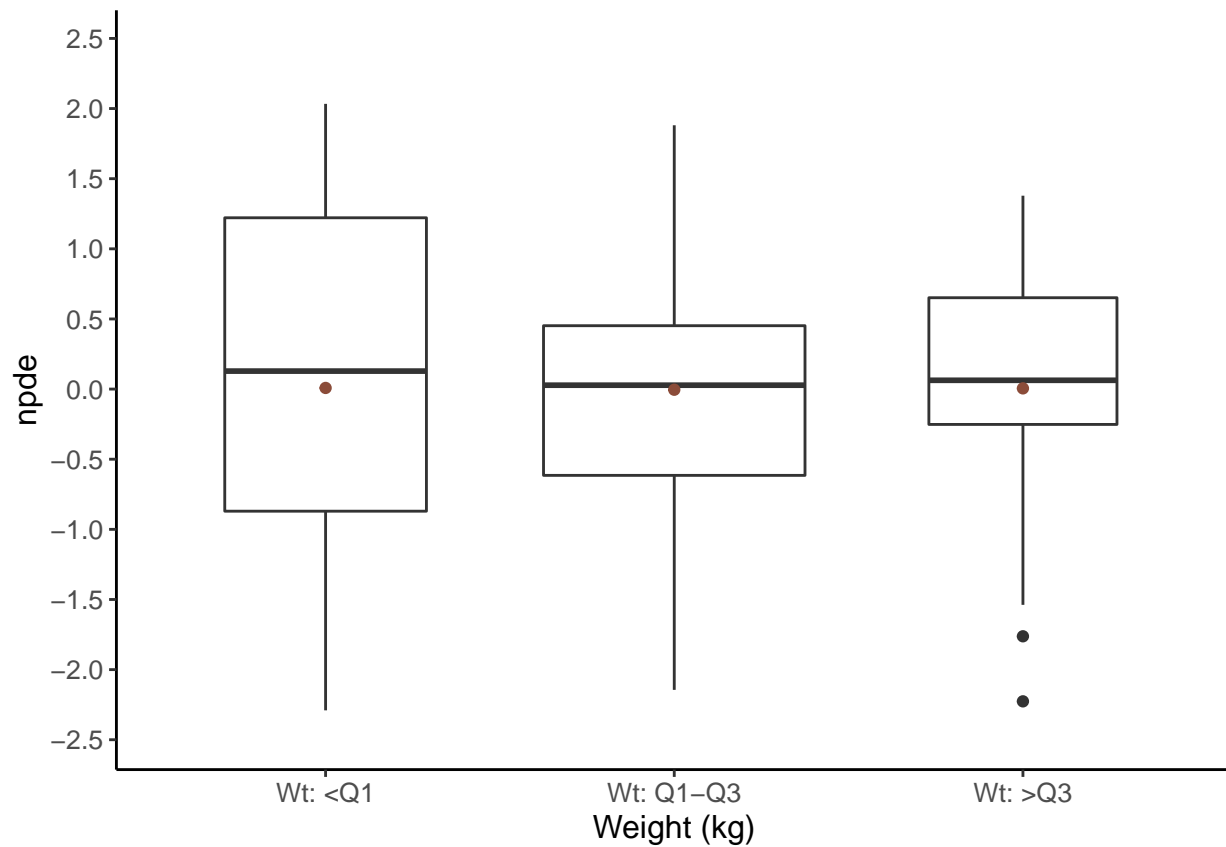
```
## Warning: Removed 1 rows containing missing values (geom_segment).
```



Covariate plots

```
# Boxplot versus weight
lcov<-"Wt"
obsxcov.weight2$grp<-factor(obsxcov.weight2$grp, labels=paste(lcov,c("<Q1","Q1-Q3",">Q3"),sep=": "))
pimatxcov.weight2$grp<-factor(pimatxcov.weight2$grp,labels=paste(lcov,c("<Q1","Q1-Q3",">Q3"),sep=": "))
plot.opt2<-plot.opt
plot.opt2$covsplit<-FALSE
plot.opt2$xlab<-"Weight (kg)"
plot.opt2$ylab<-"npde"
aux.npdeplot.boxcov(obsxcov.weight2,pimatxcov.weight2, plot.opt2)
```

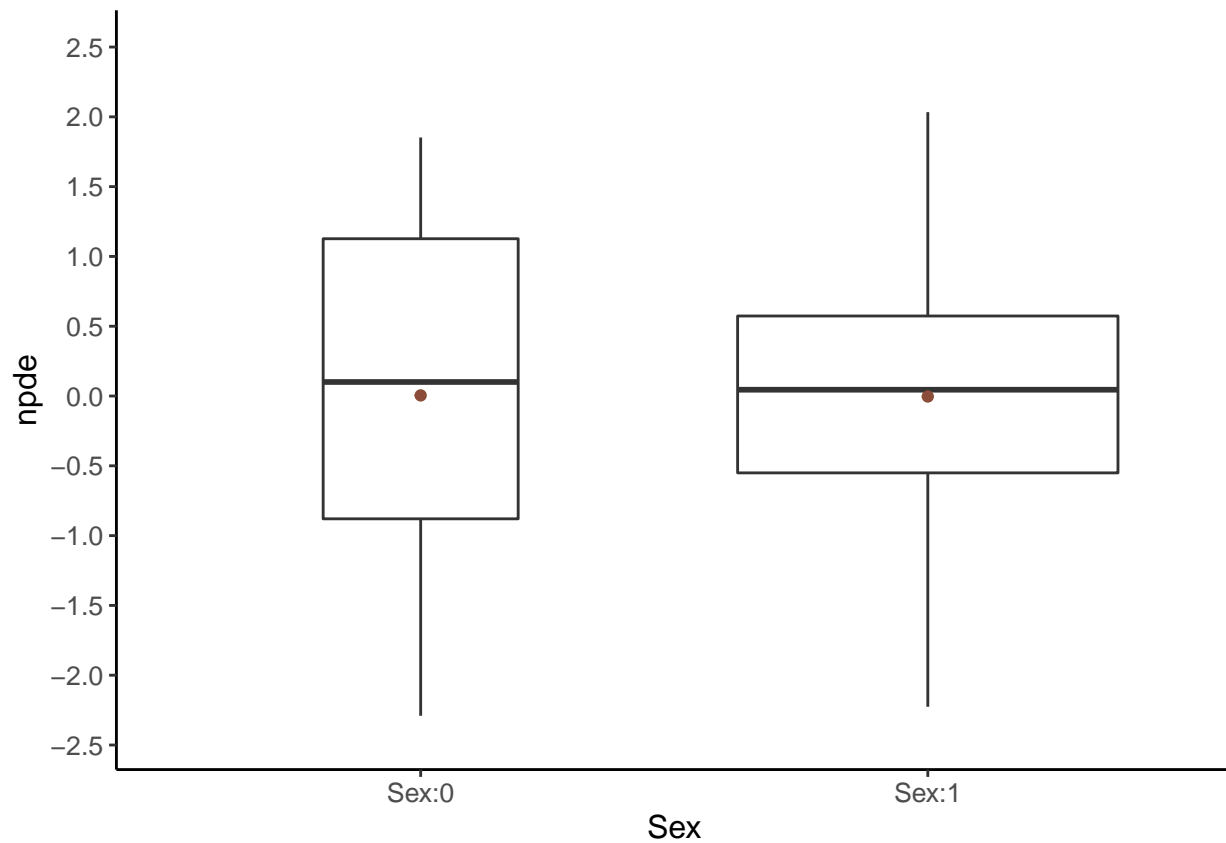
```
## Warning: Removed 2 rows containing non-finite values (stat_boxplot).
```



```
# Boxplot versus sex
obsxcov.sex$grp<-factor(obsxcov.sex$grp, labels=c("Sex:0", "Sex:1"))
pimatxcov.sex$grp<-factor(pimatxcov.sex$grp, labels=c("Sex:0", "Sex:1"))
plot.opt2<-plot.opt
plot.opt2$xlab<-"Sex"
plot.opt2$ylab<-"npde"
aux.npdeplot.boxcov(obsxcov.sex,pimatxcov.sex, plot.opt2)
```

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

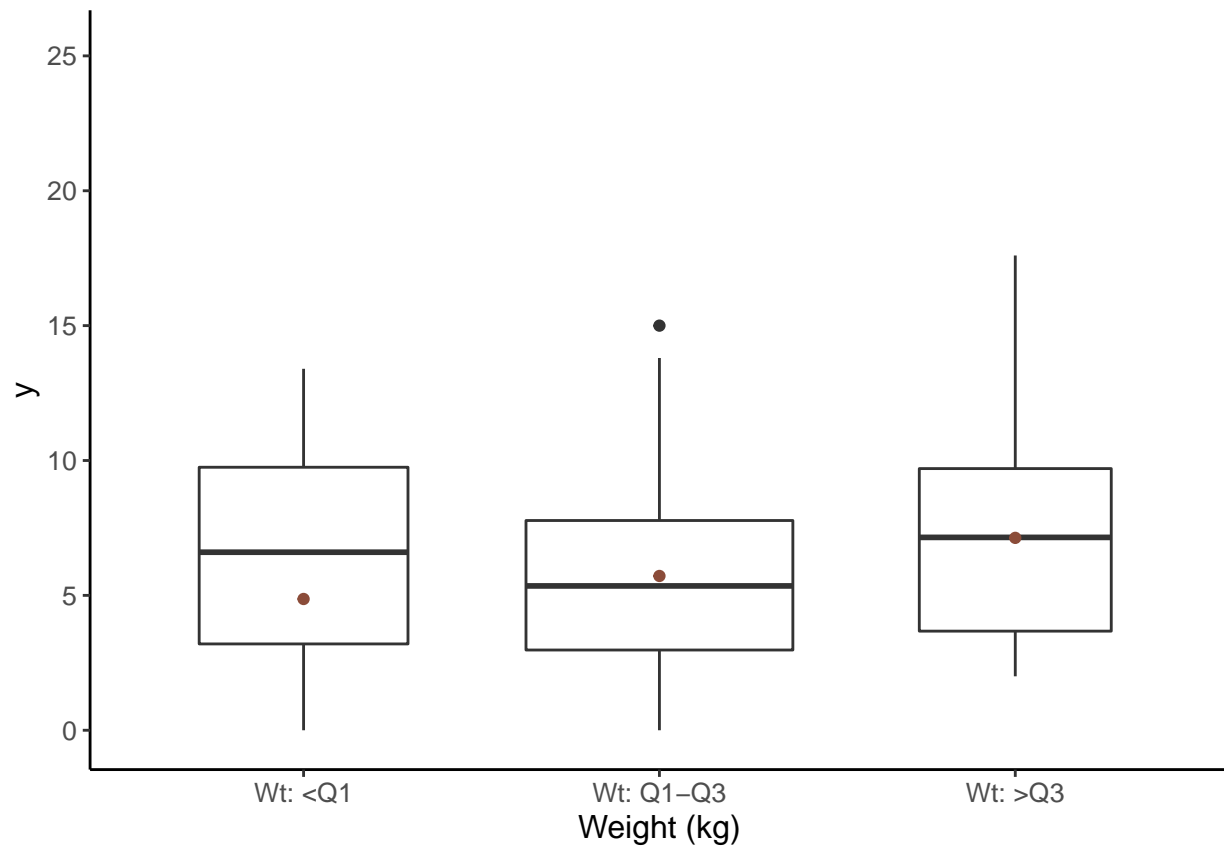




```
if(FALSE) {
  obsmat<-obsxcov.sex
  pimat<-pimatxcov.sex
}

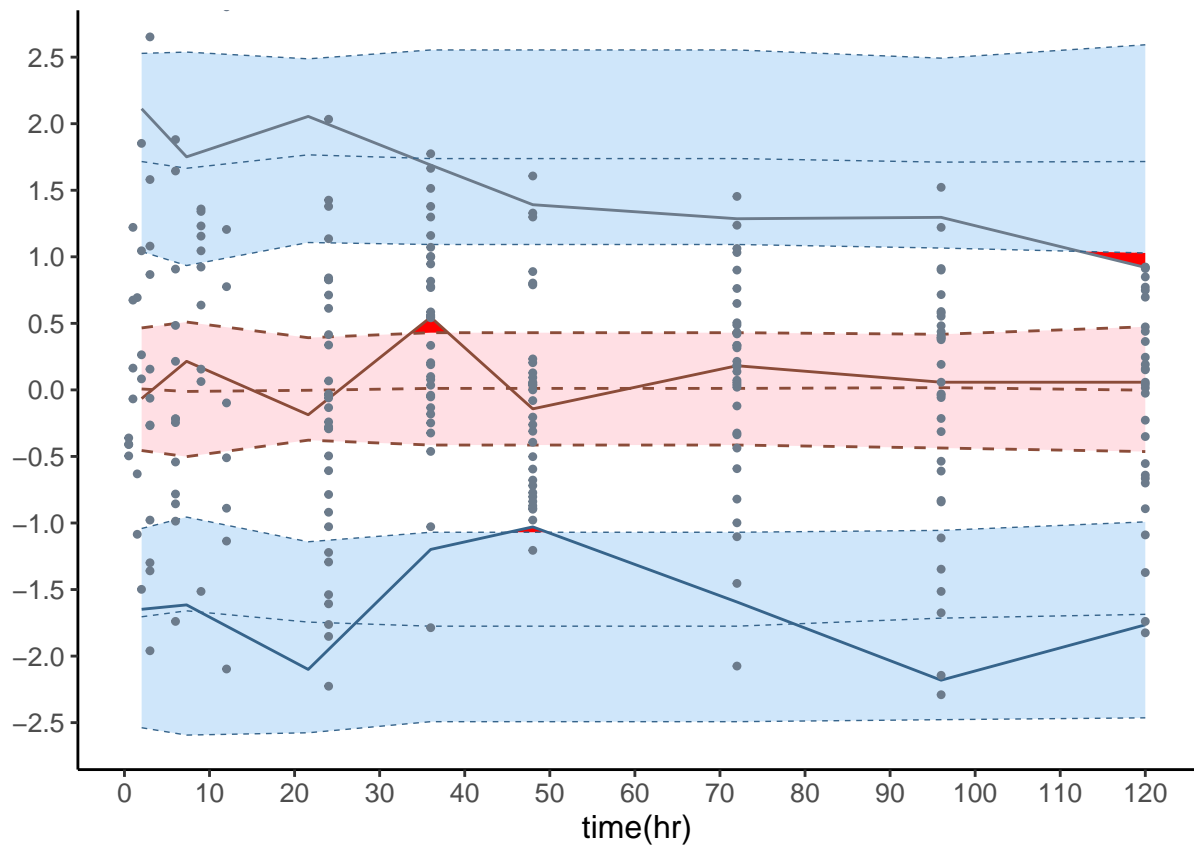
# Same, with observations (not sure that makes sense)
obsycov.weight<-obsxcov.weight2
obsycov.weight$y<-wbase@data@data$dv
pimatycov.weight<-aux.npdeplot.pimat(obsycov.weight, xcent=sort(unique(obsxcov.weight$x)), distrib="norm",
plot.opt2<-plot.opt
plot.opt2$xlabel<-"Weight (kg)"
plot.opt2$ylabel<-"y"

aux.npdeplot.boxcov(obsycov.weight,pimatycov.weight, plot.opt2)
```

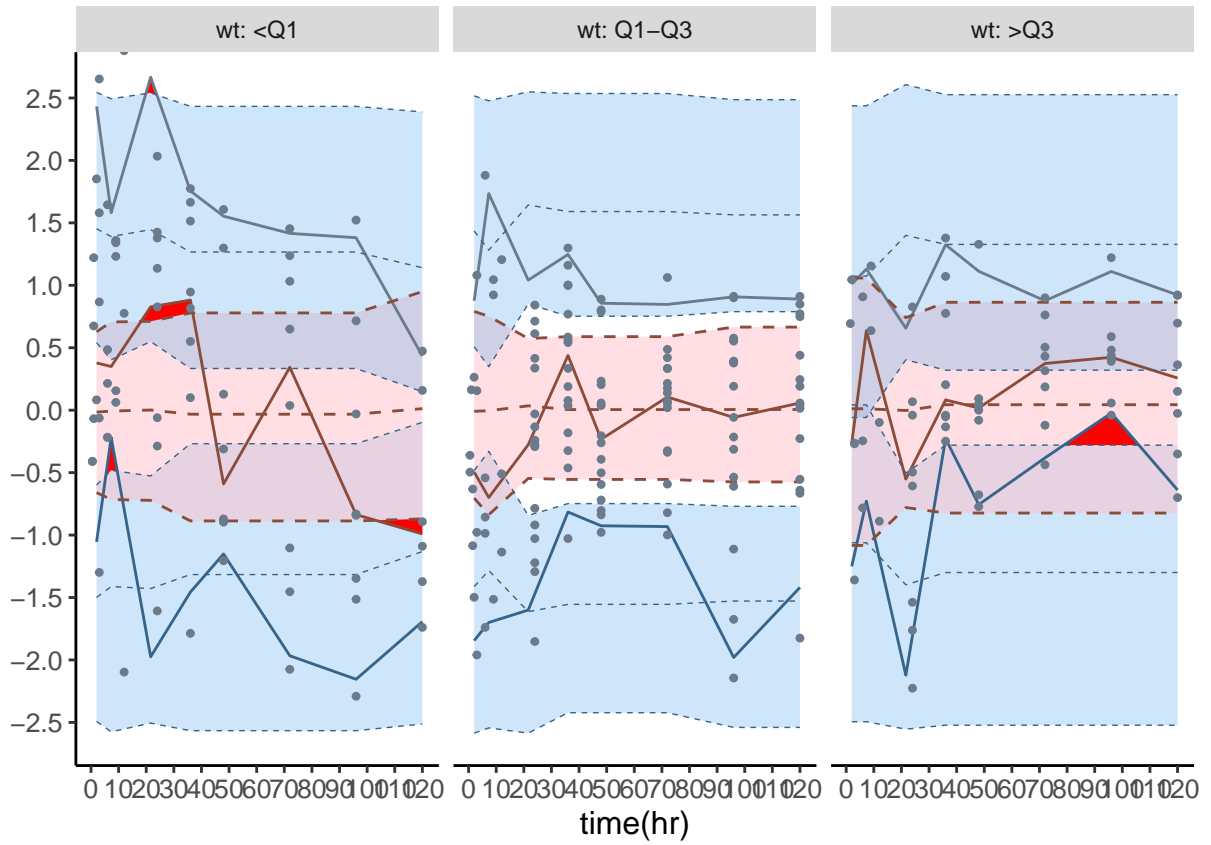


Debug main scatterplot function

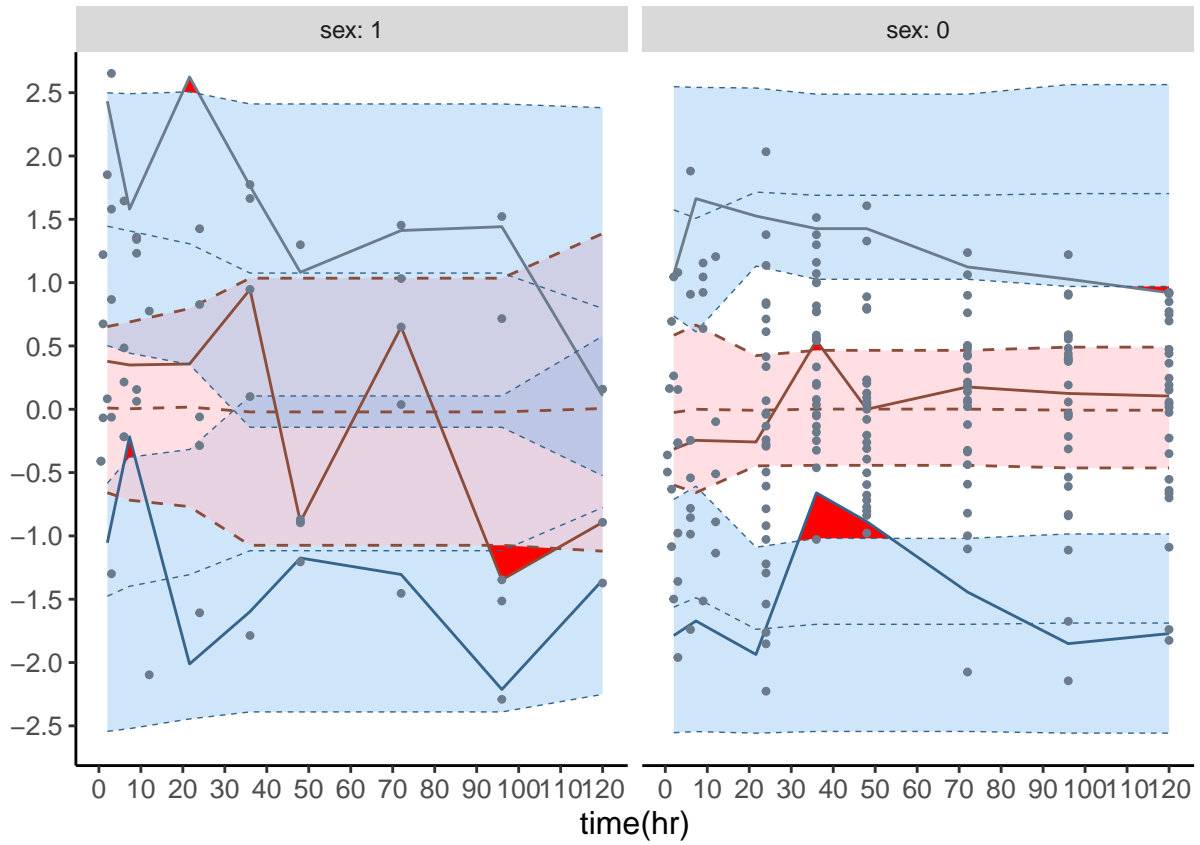
```
source(file.path(workDir,"npde","R","plotNpde-scatterplot.R")) # renamed from plotNpde-unitFunctionsAux
npde.plot.scatterplot(wbase, which.x="x", which.y="npde")
```



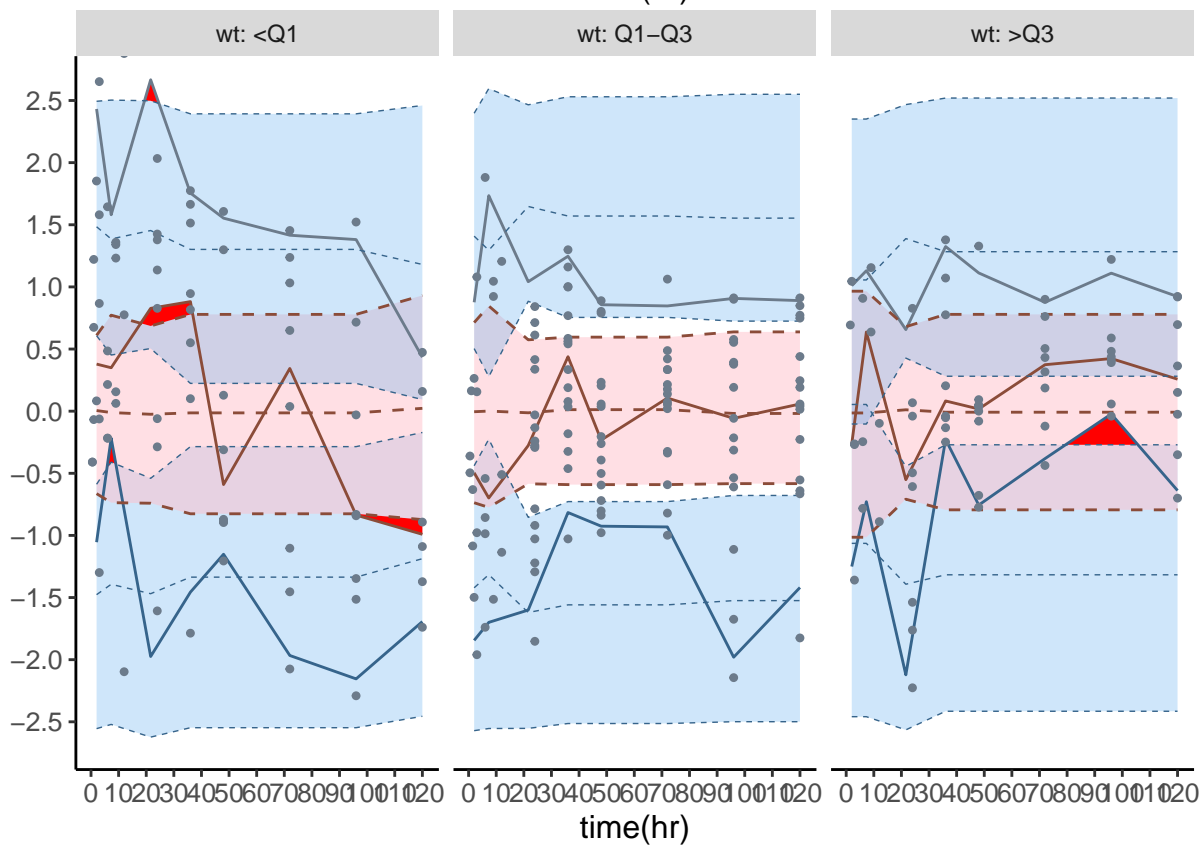
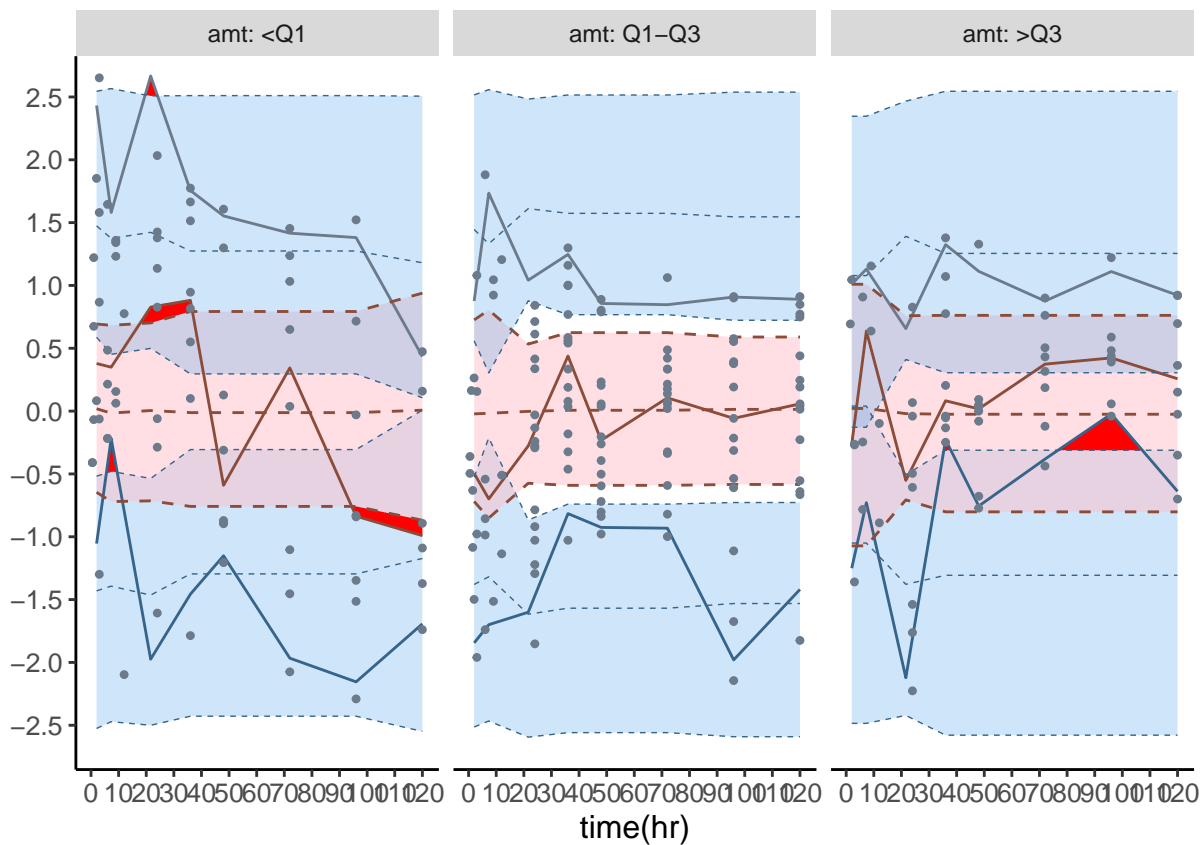
```
# Split by covariate
npde.plot.scatterplot(wbase, which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("wt"))
```

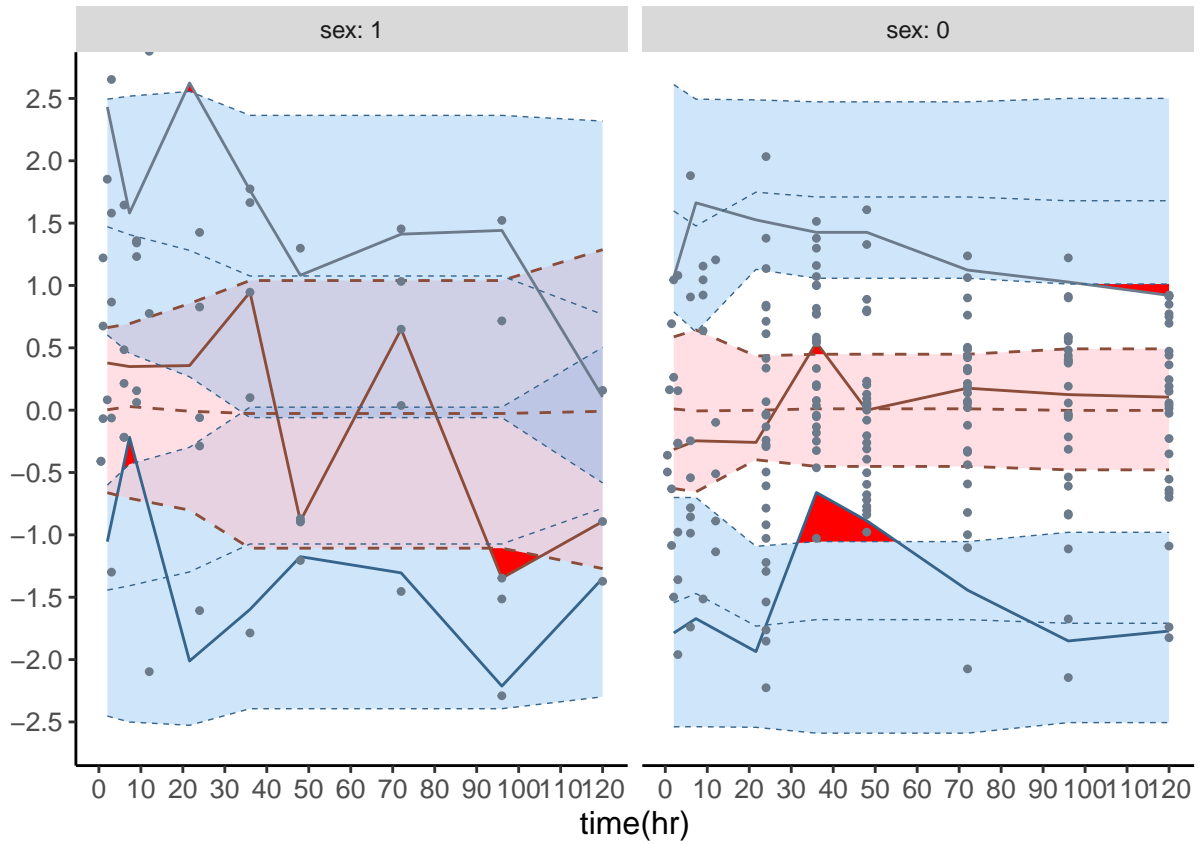


```
npde.plot.scatterplot(wbase, which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("sex"))
```

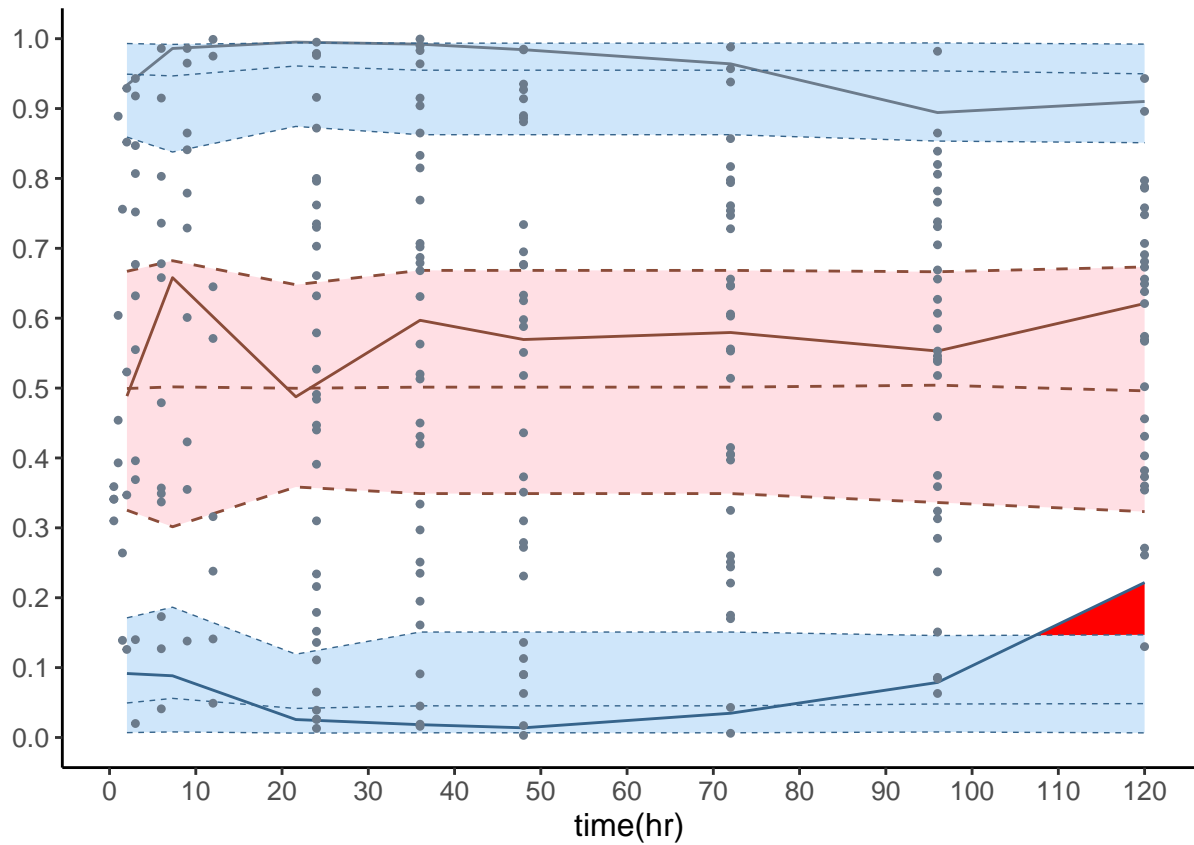


```
npde.plot.scatterplot(wbase, which.x="x", which.y="npde", covsplit=TRUE, which.cov=c("all"))
```





```
# PD vs time
npde.plot.scatterplot(wbase, which.x="x", which.y="pd")
```

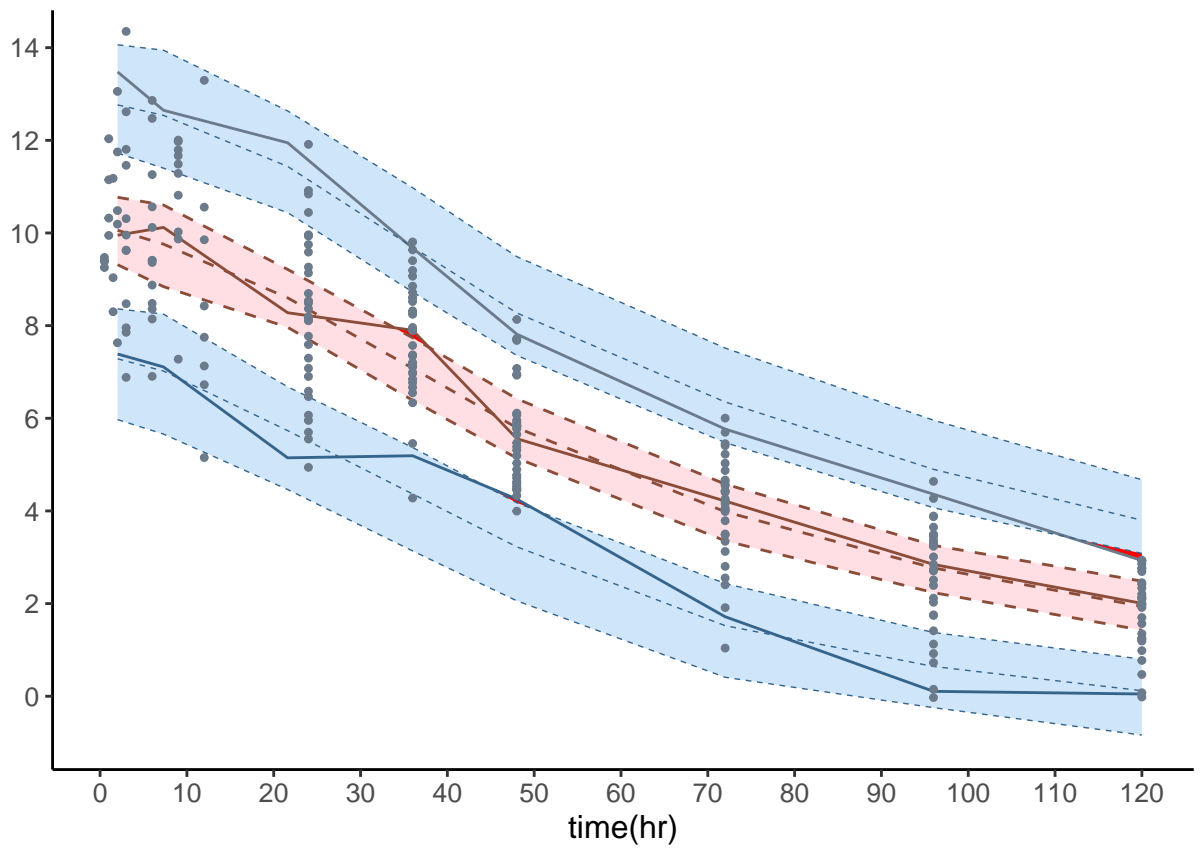


*# Reference profile*

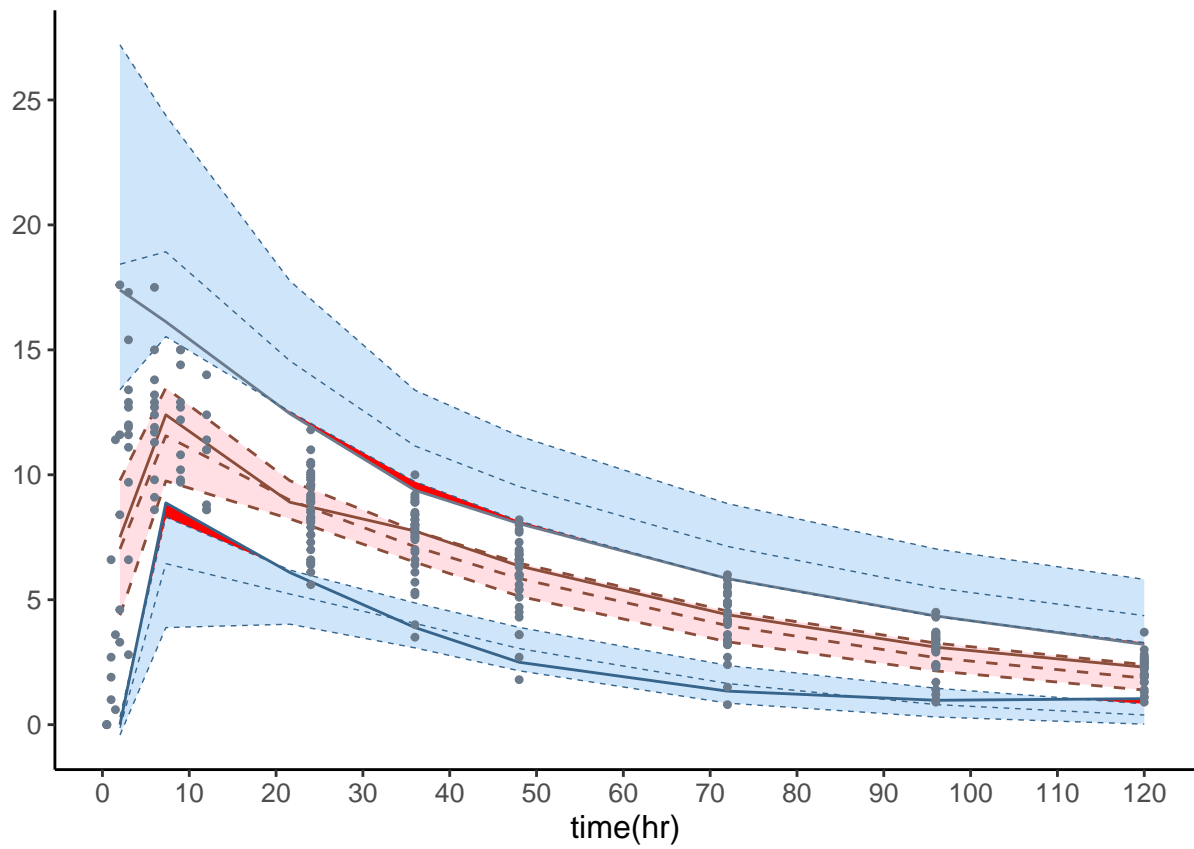
```
npde.plot.scatterplot(wbase, which.x="x", which.y="npde", ref.prof=list(id=1))
```

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

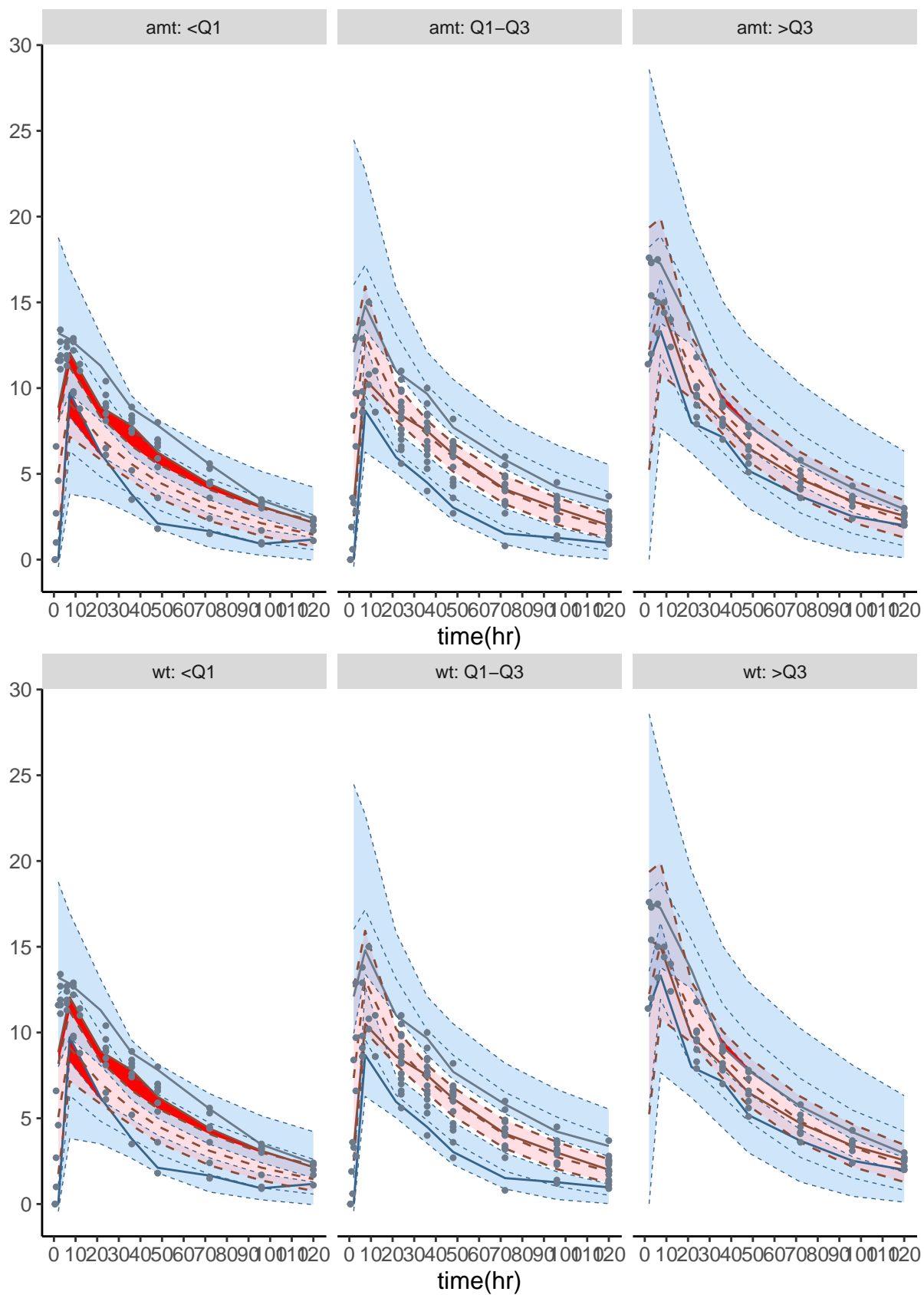


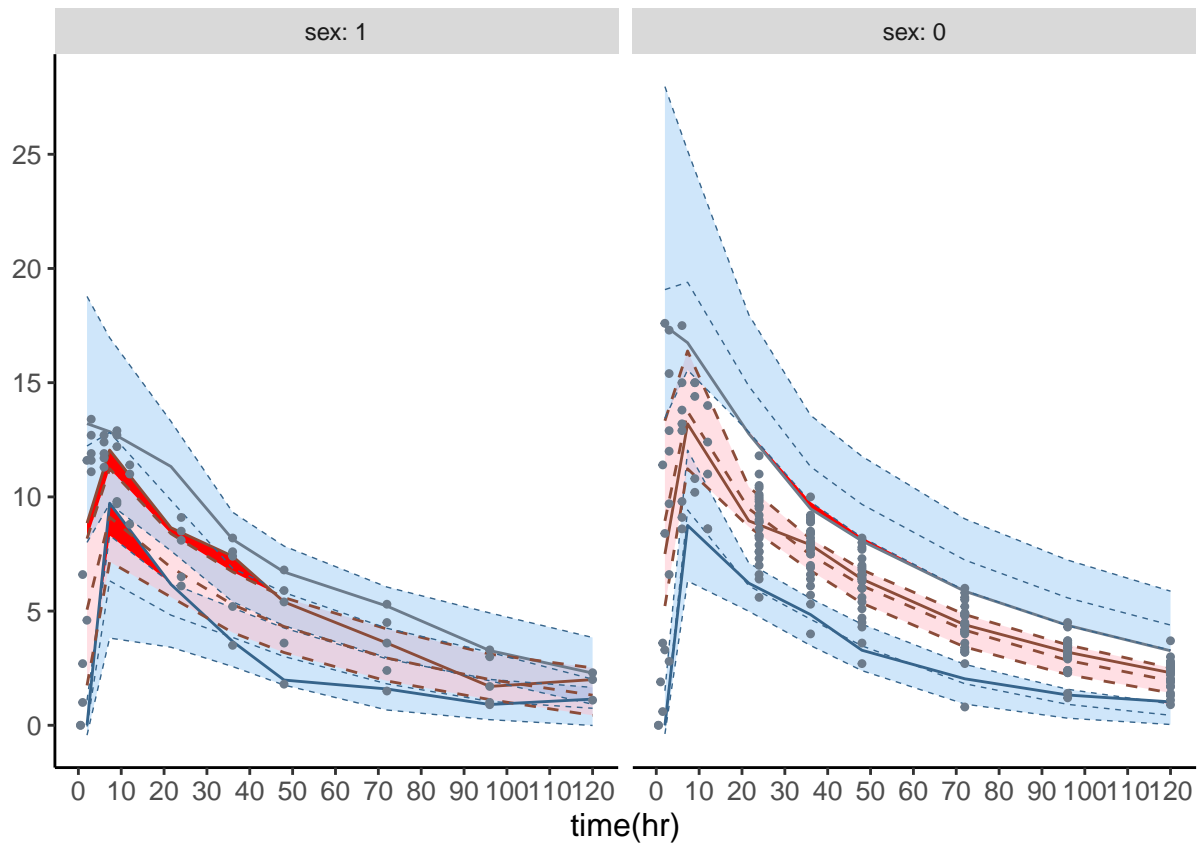


```
# VPC
npde.plot.scatterplot(wbase, which.x="x", which.y="yobs")
```



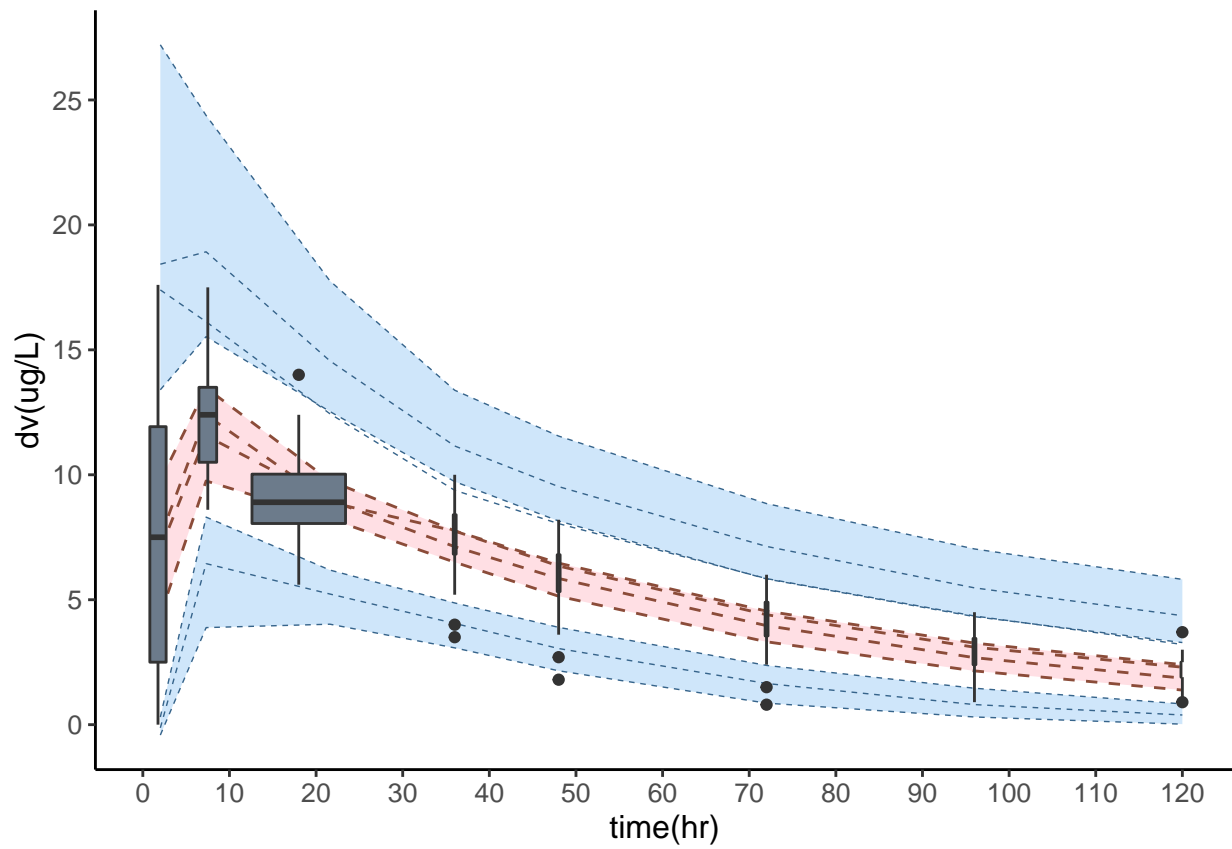
```
npde.plot.scatterplot(wbase, which.x="x", which.y="yobs", covsplit=TRUE)
```



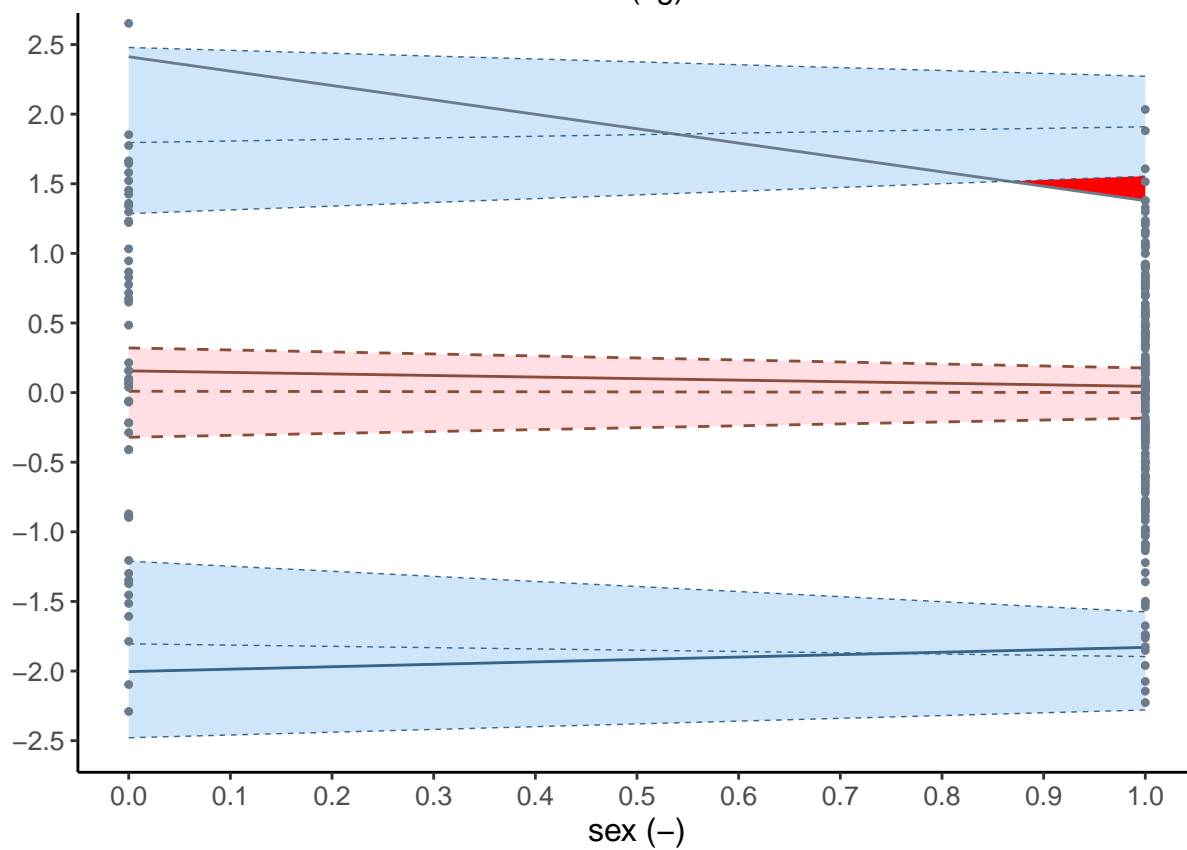
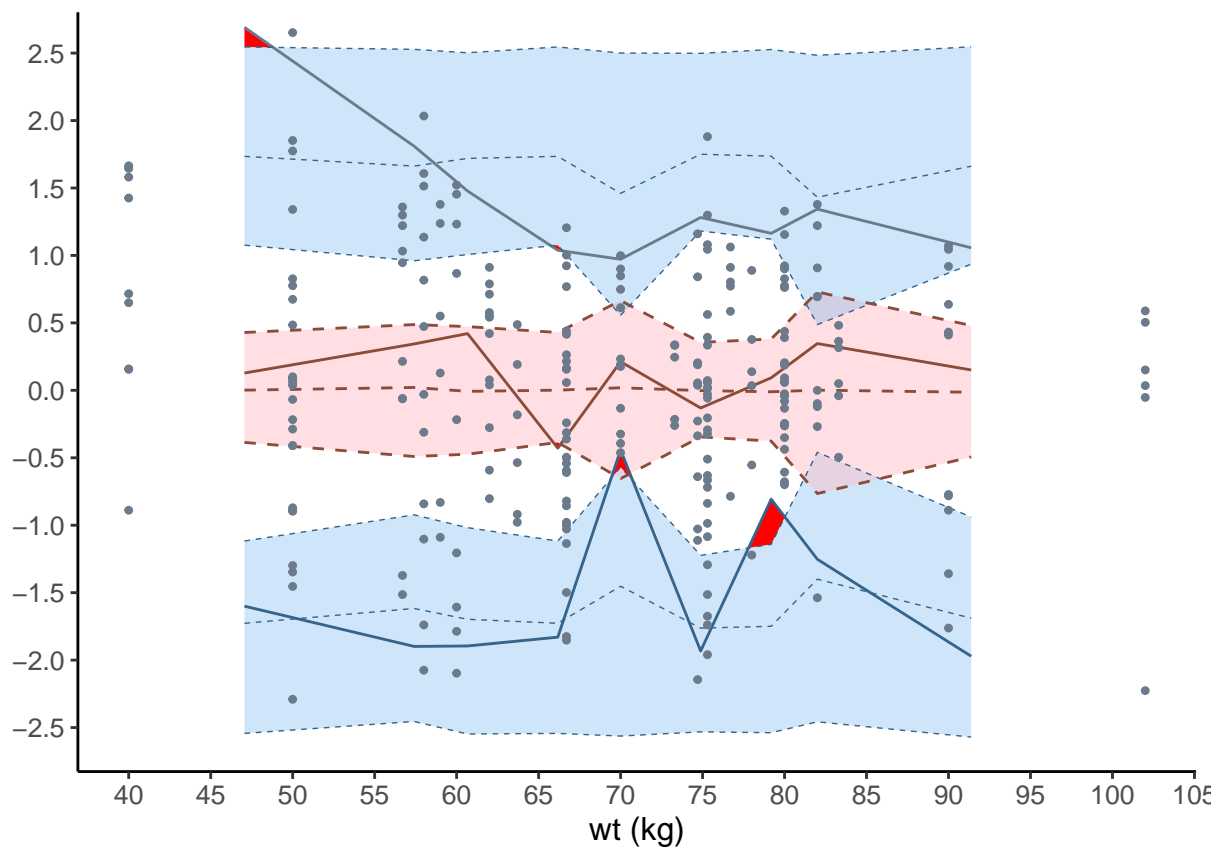


```
npde.plot.scatterplot(wbase, which.x="x", which.y="yobs", plot.box=TRUE)
```

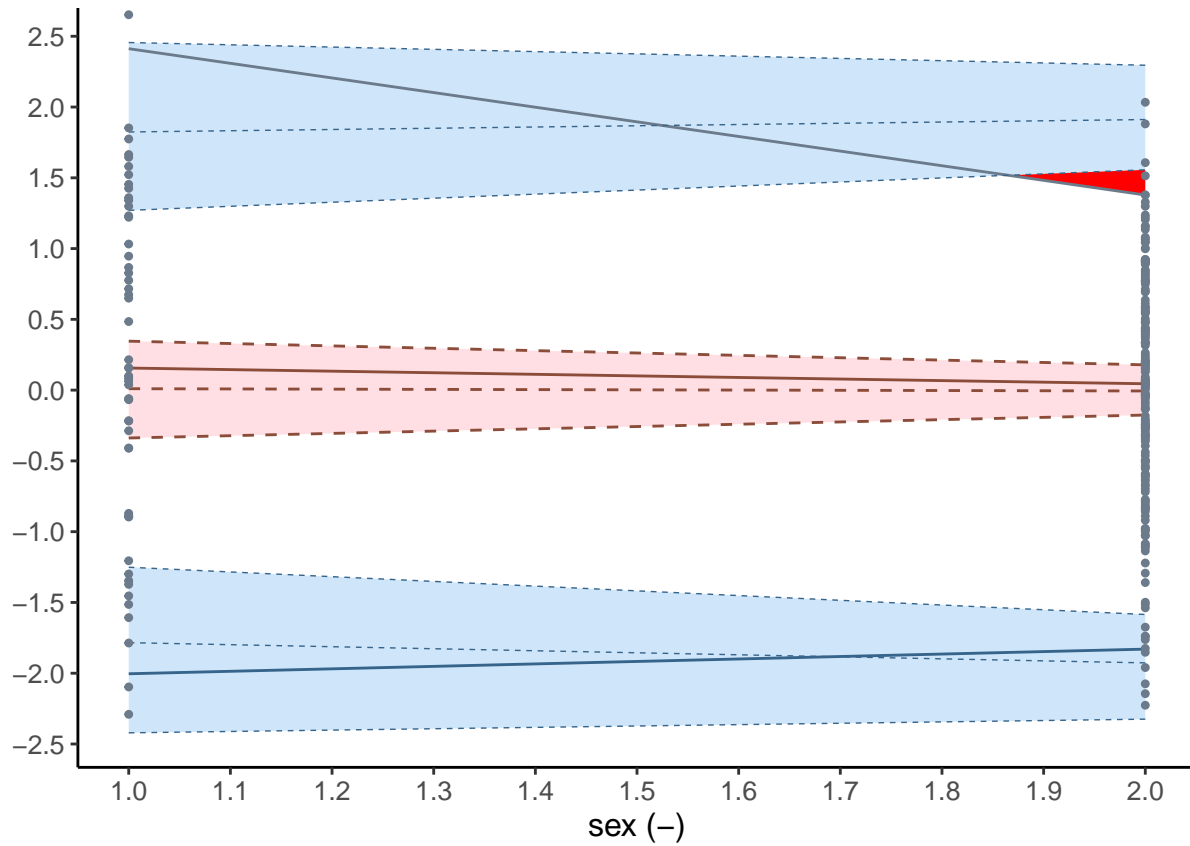
```
## Warning: Removed 1 rows containing missing values (geom_segment).
```



```
# Scatterplot versus covariates
npde.plot.scatterplot(wbase, which.x="cov", which.y="npde", which.cov=c("wt","sex"))
```



```
# Treating sex as factor
wbase@data@data$sex<-factor(wbase@data@data$sex)
npde.plot.scatterplot(wbase, which.x="cov", which.y="npde", which.cov=c("sex"))
```

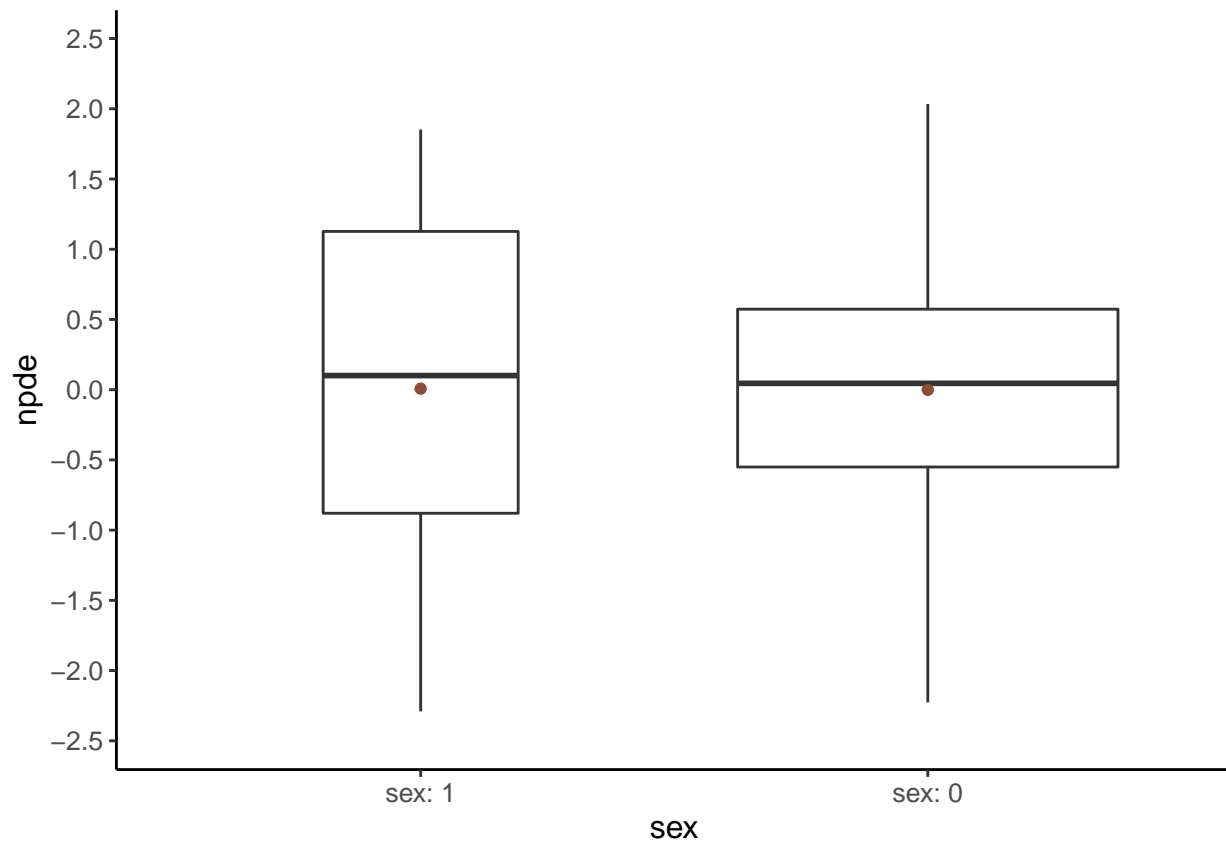


```
if(FALSE) {
  npdeObject<-wbase
  which.x<-"cov"
  which.y<-"npde"
  plot.opt<-npdeObject@prefs
  ref.prof<-NULL
}
```

Debug boxplot of npde versus covariates (Brendel et al. 2010)

```
source(file.path(workDir,"npde","R","plotNpde-covplot.R"))
npde.plot.covariate(wbase, which.y="npde", which.cov="sex")
```

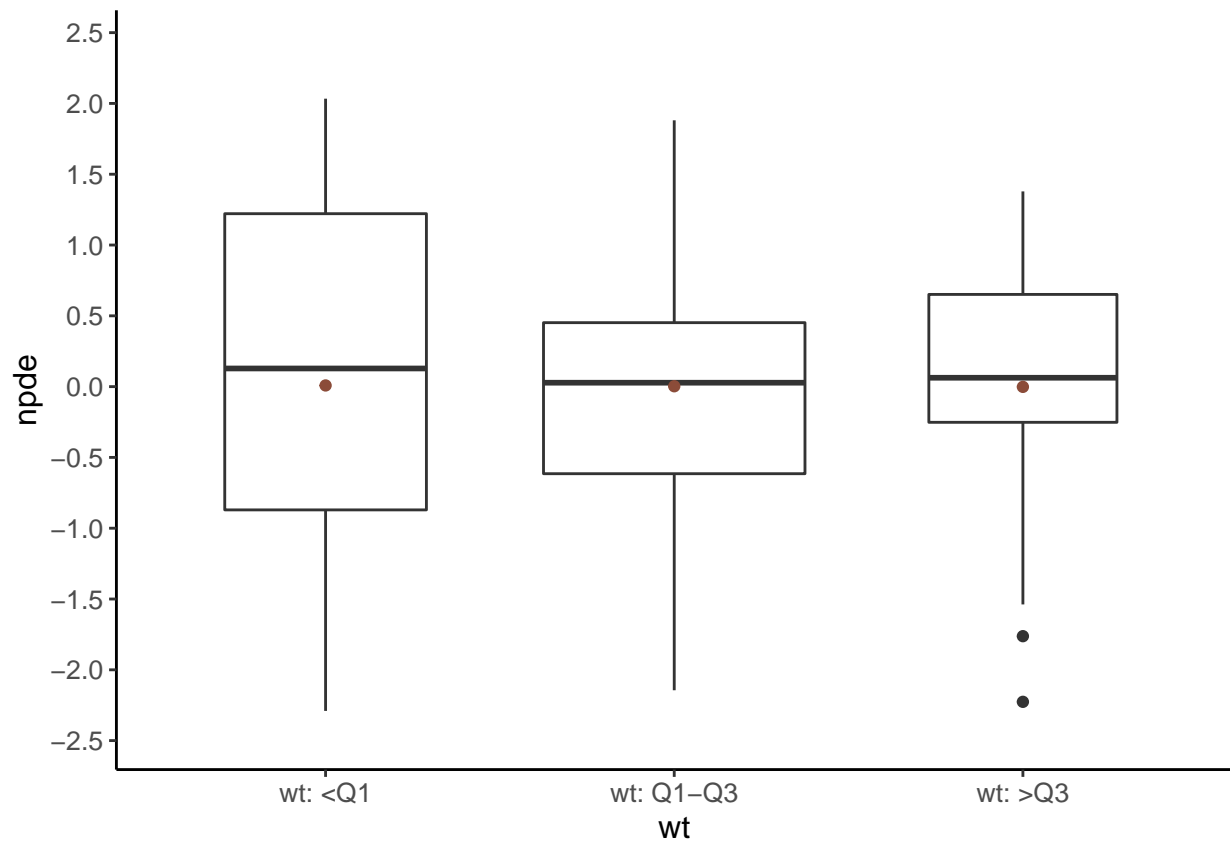
```
## Warning: Removed 2 rows containing non-finite values (stat_boxplot).
```



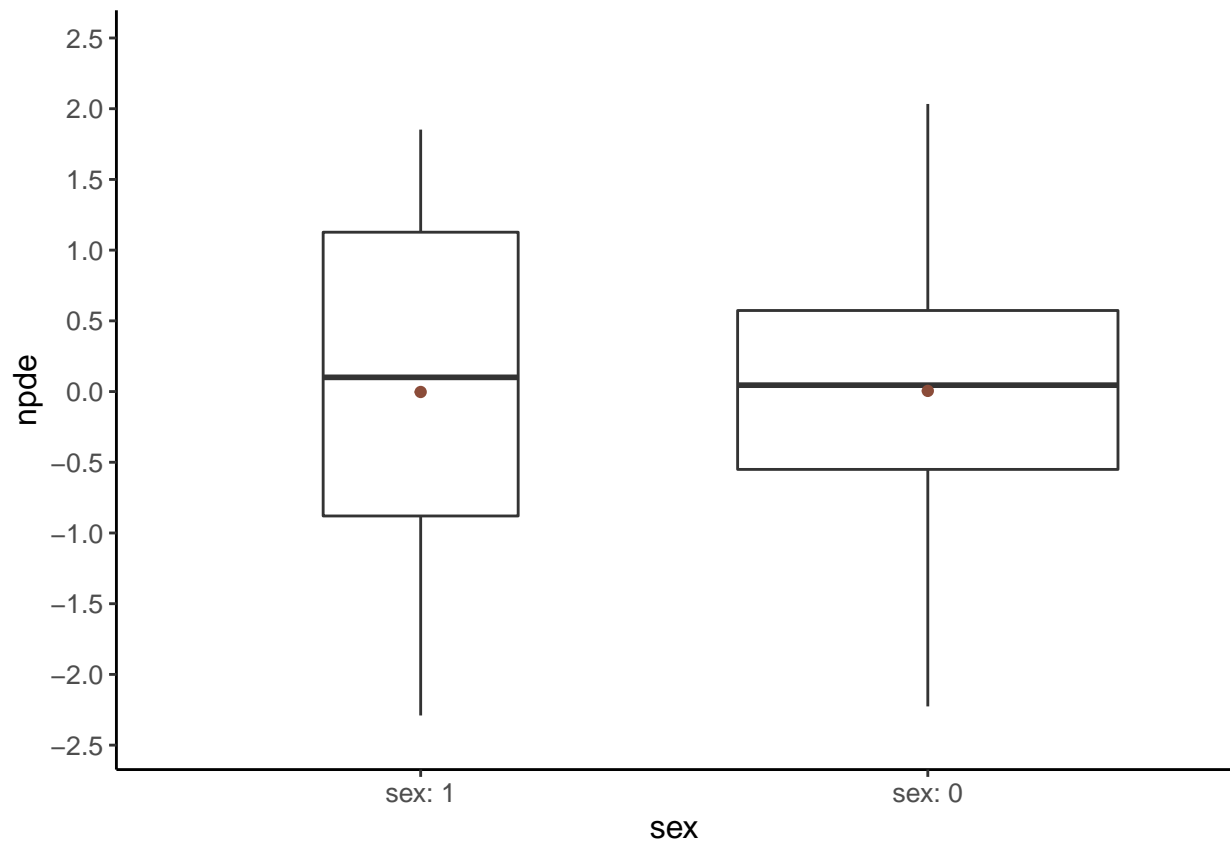
```
npde.plot.covariate(wbase, which.y="npde", which.cov=c("wt", "sex"))
```

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

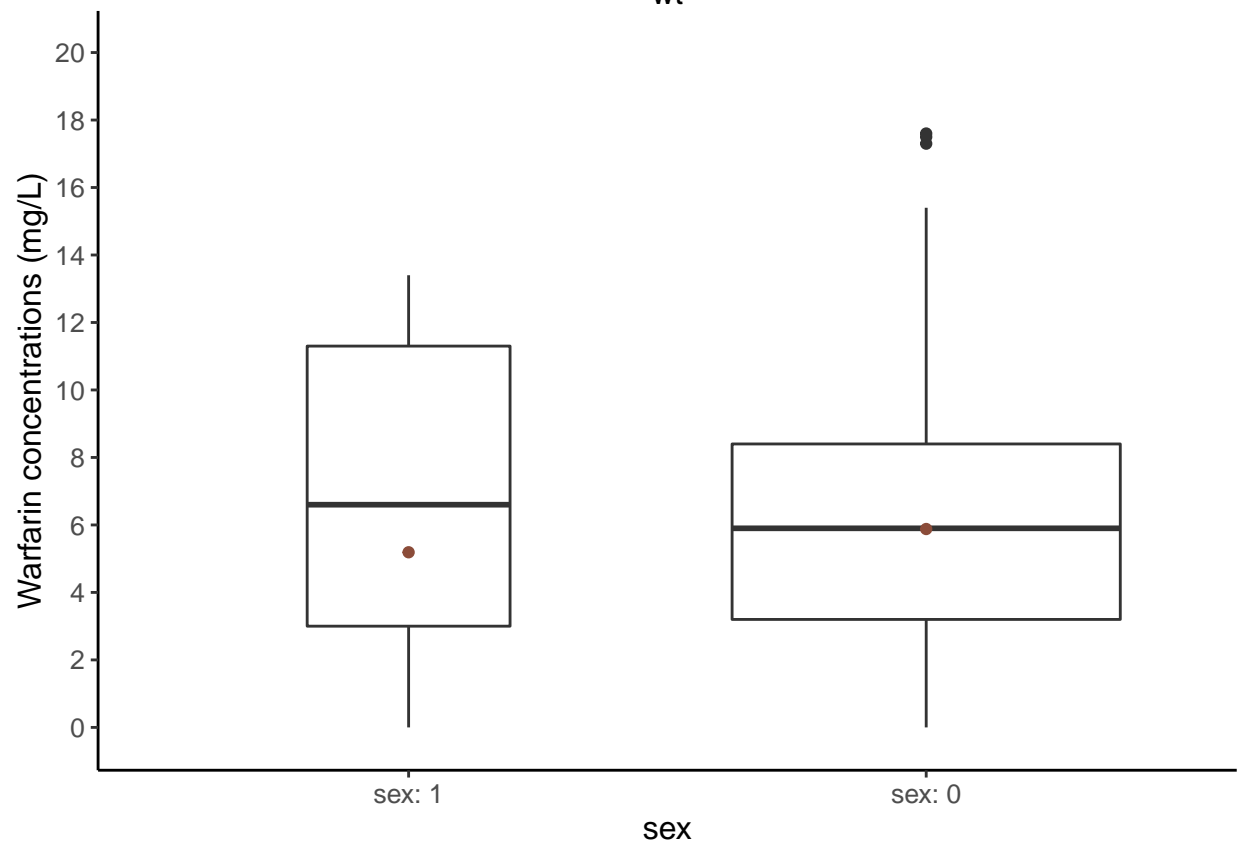
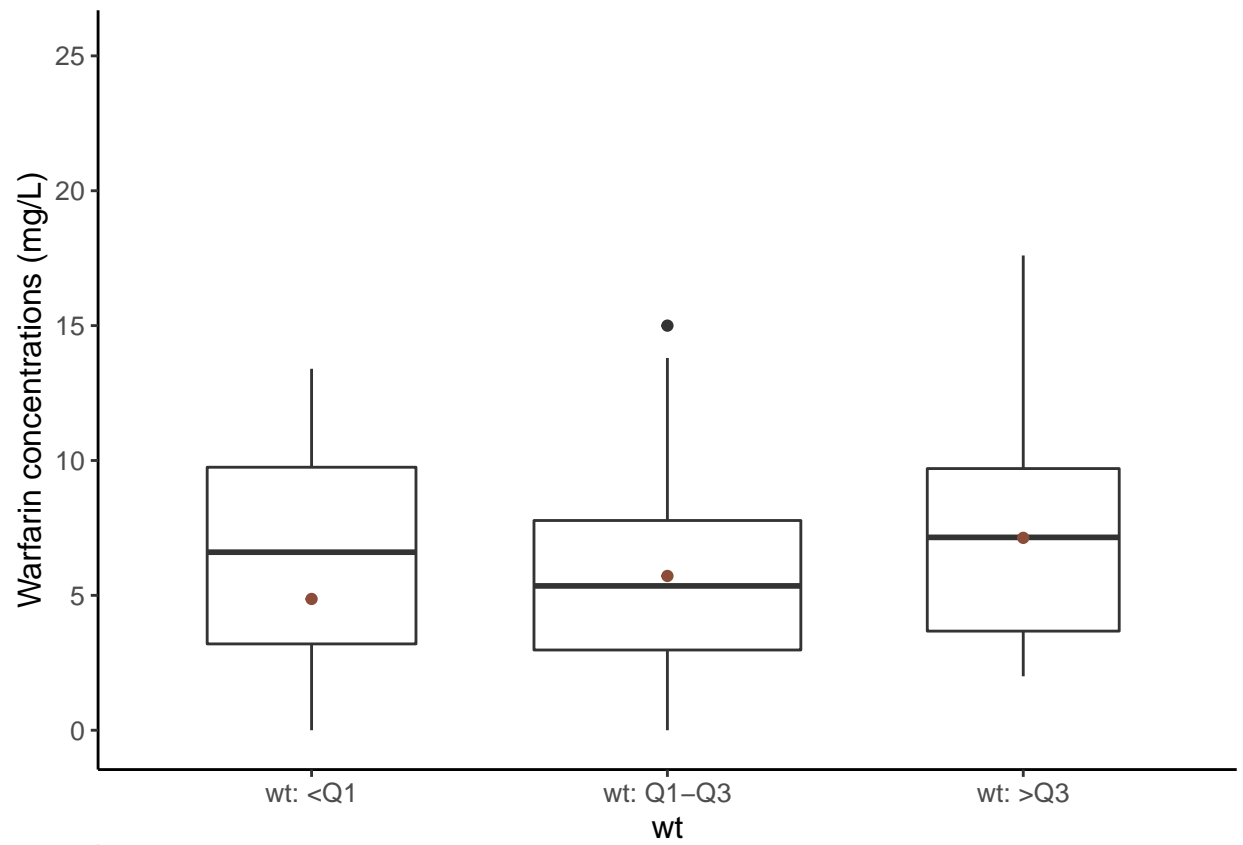




## Warning: Removed 2 rows containing non-finite values (stat\_boxplot).



```
npde.plot.covariate(wbase, which.y="yobs", which.cov=c("wt", "sex"), ylab="Warfarin concentrations (mg/L)
```



```
if(FALSE) {  
  npdeObject<-wbase  
  which.y<-"npde"  
  plot.opt<-npdeObject@prefs  
}
```