

# Global soil water characteristics datasets and parameters

Surya Gupta, Andreas Papritz, Peter Lehmann, Tom Hengl, Sara Bonetti, Dani Or

- Bulk density vs Data sources
- Organic carbon vs Data sources
- Sand content vs Data sources
- Silt content vs Data sources
- Clay content vs Data sources
- pH vs Data sources
- Field measured saturated hydraulic conductivity vs Data sources
- Lab measured saturated hydraulic conductivity vs Data sources
- Porosity vs Data sources
- VG shape parameter vs Data sources
- VG shape parameter vs Data sources
- Saturated water content vs Data sources
- Residual water content vs Data sources

Plots of soil water characteristics data set. Total curves = 15259

```
library(ggplot2)
WRC<- read.csv("C:/Users/guptasu.D/Downloads/GSRP_Figures/WRC_dataset_surya_et_al_2021_final.csv")
dim(WRC)

## [1] 136989 54

colnames(WRC)

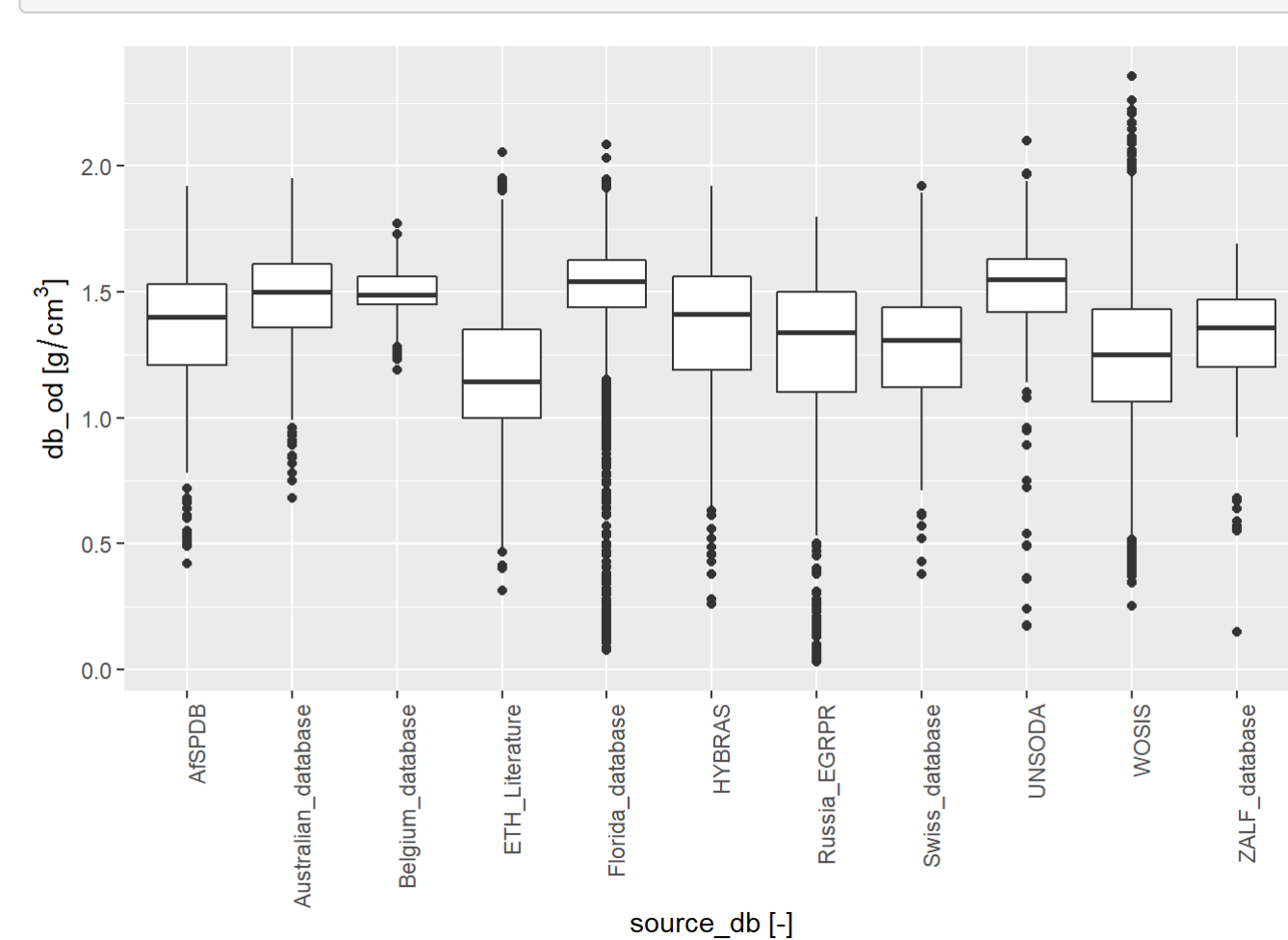
## [1] "layer_id" "disturbed_undisturbed"
## [3] "climate_classes" "profile_id"
## [5] "reference" "DOI_URL"
## [7] "method" "method_keywords"
## [9] "latitude_decimal_degrees" "longitude_decimal_degrees"
## [11] "hm_design" "hzt_tot_psa"
## [13] "hzt_bot" "hzt_top"
## [15] "db_od" "oc"
## [17] "tc_psd" "sand_tot_psa"
## [19] "silt_tot_psa" "clay_tot_psa"
## [21] "ph_h2o" "kcat_field"
## [23] "kcat_lab" "porosity"
## [25] "w6_33kpa" "lab_head_m"
## [27] "lab_wrc" "field_head_m"
## [29] "field_wrc" "reference_total_porosity"
## [31] "SMCC.classes" "source_db"
## [33] "location_accuracy_min" "location_accuracy_max"
## [35] "road_accuracy_classes" "alpha"
## [37] "se_alpha" "n"
## [39] "se_n" "theta_r"
## [41] "theta_s" "q2_5_alpha"
## [43] "q97_5_alpha" "q2_5_n"
## [45] "q97_5_n" "q30_alpha"
## [47] "q98_alpha" "q30_n"
## [49] "q99_n" "q25_alpha"
## [51] "q75_alpha" "q25_n"
## [53] "q75_n" "data_flag"
```

```
unique(WRC$source_db)

## [1] "WOSIS" "Russia_EGRRP" "ETH_Literature"
## [4] "Australian_database" "AFSPDB" "ETH_database"
## [7] "Florida_database" "HYBRAS" "UNSOQA"
## [10] "Belgium_database" "ZALF_database"
```

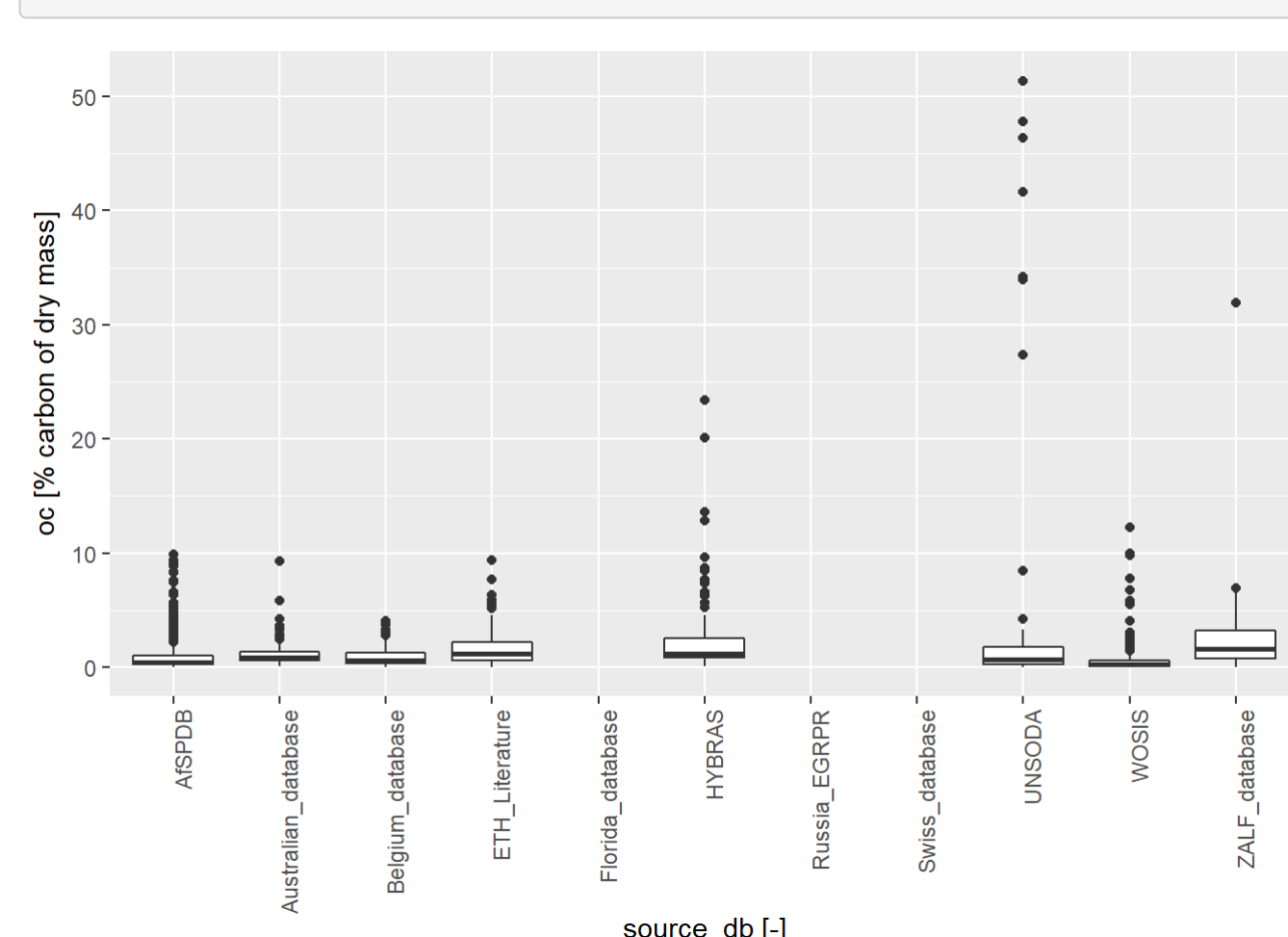
## Bulk density vs Data sources

```
options(warn=-1)
ggplot(WRC, aes(x=source_db, y=db_od)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("db_od [", g/cm^3, "]")), x = expression(paste("source_db [-]" )))
```



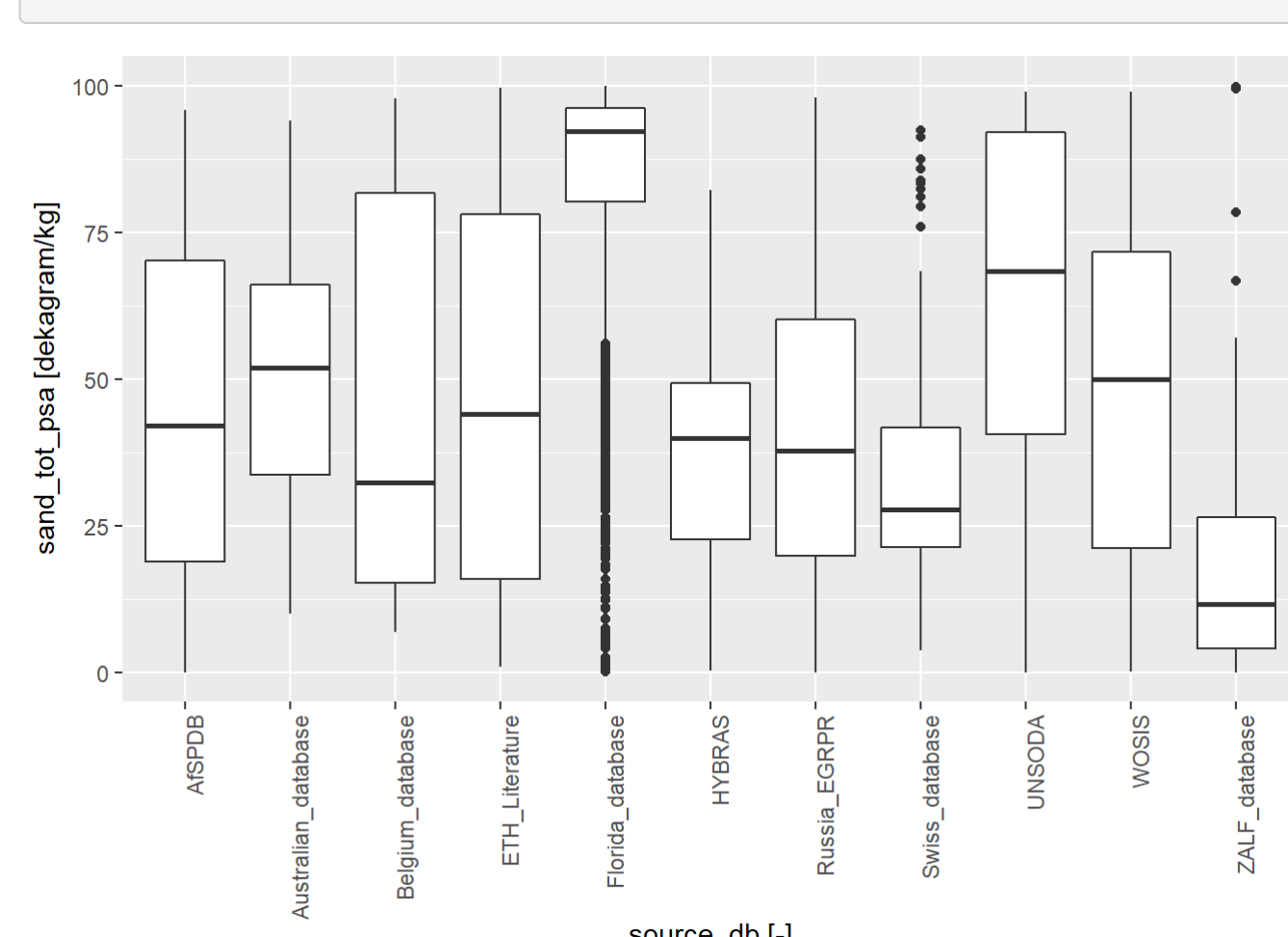
## Organic carbon vs Data sources

```
ggplot(WRC, aes(x=source_db, y=oc)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("oc [% Carbon of dry mass]")), x = expression(paste("source_db [-]" )))
```



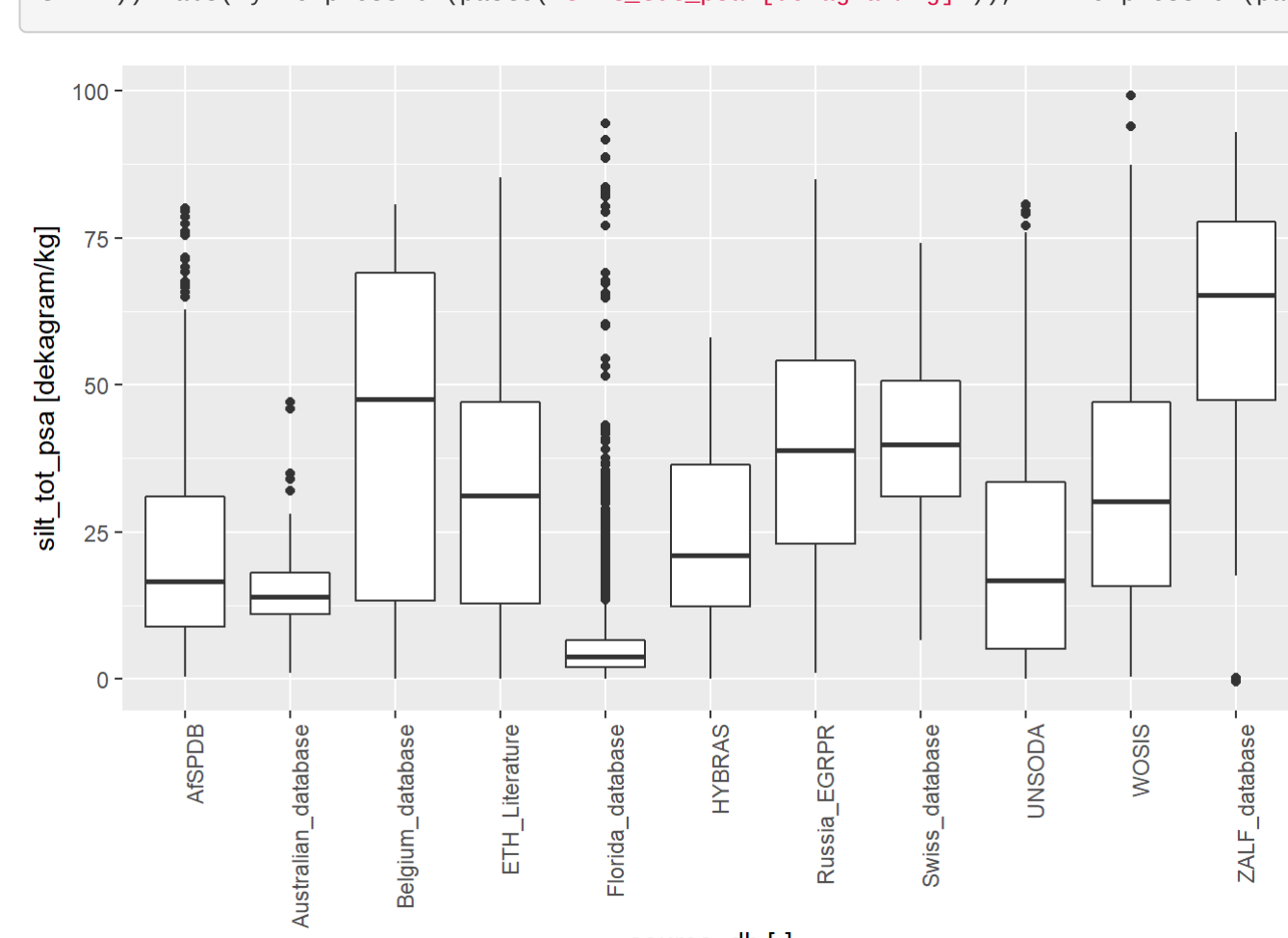
## Sand content vs Data sources

```
ggplot(WRC, aes(x=source_db, y=sand_tot_psa)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("sand_tot_psa [dekagram/kg]")), x = expression(paste("source_db [-]" )))
```



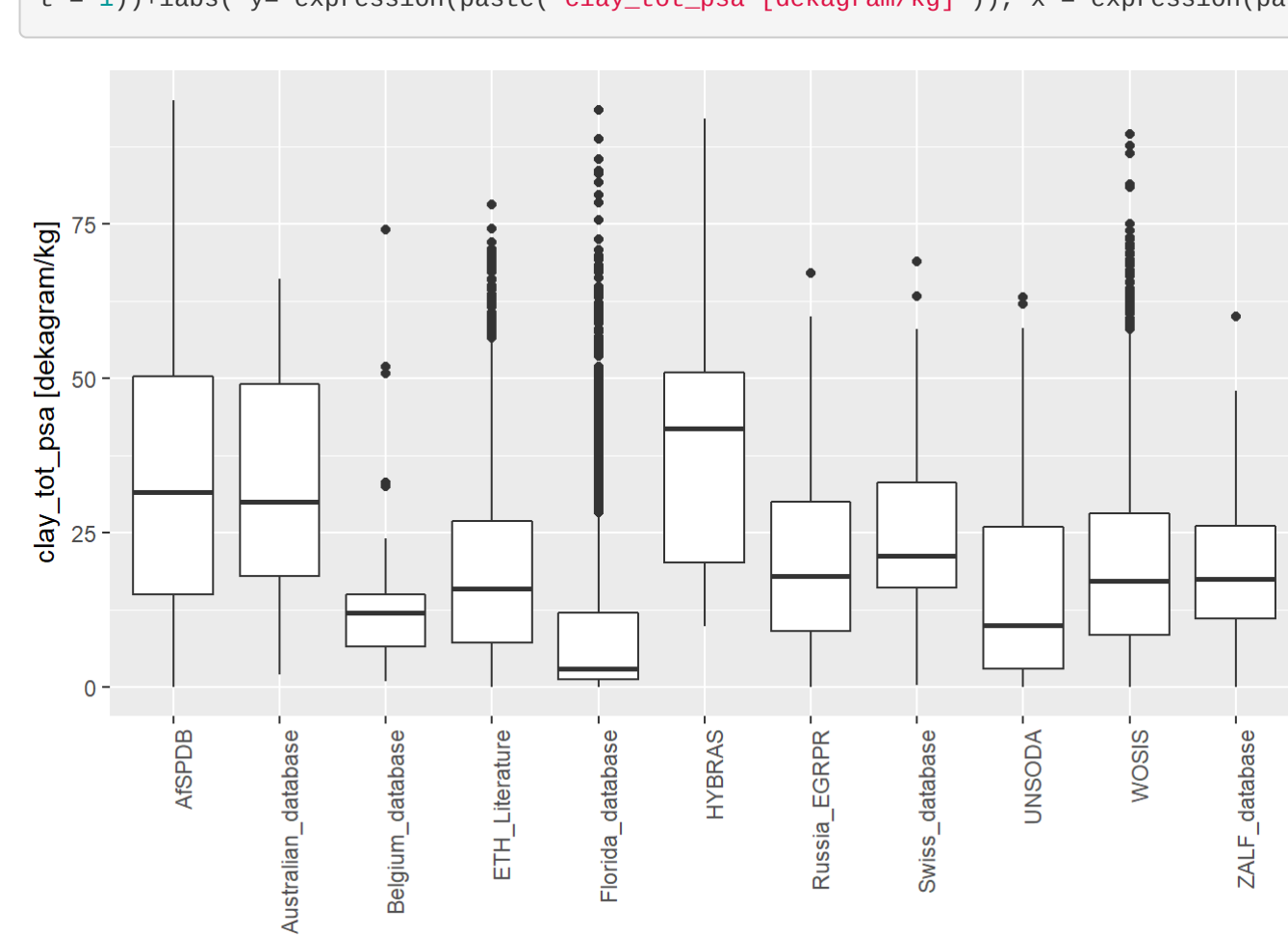
## Silt content vs Data sources

```
ggplot(WRC, aes(x=source_db, y=silt_tot_psa)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("silt_tot_psa [dekagram/kg]")), x = expression(paste("source_db [-]" )))
```



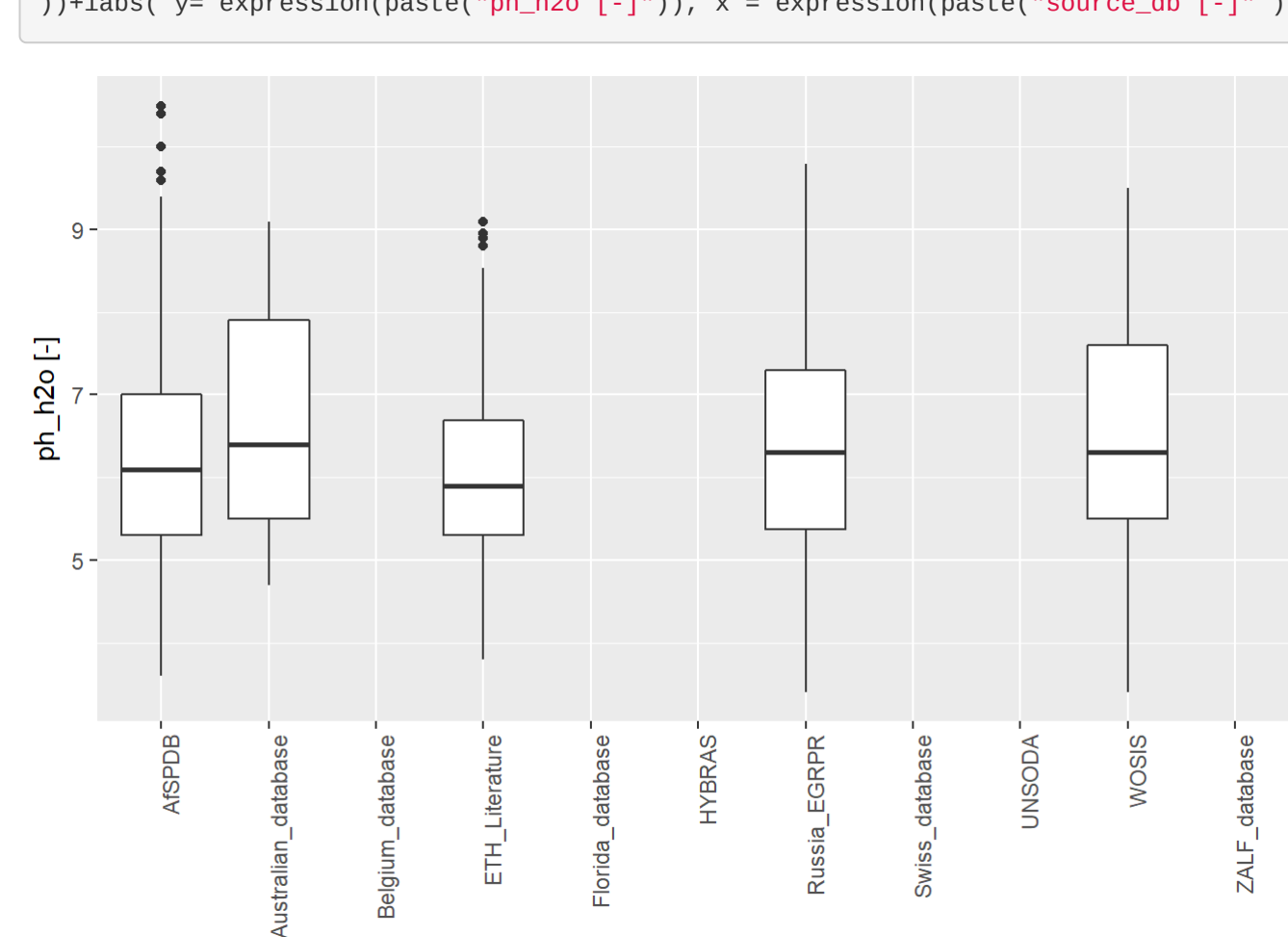
## Clay content vs Data sources

```
ggplot(WRC, aes(x=source_db, y=clay_tot_psa)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("clay_tot_psa [dekagram/kg]")), x = expression(paste("source_db [-]" )))
```



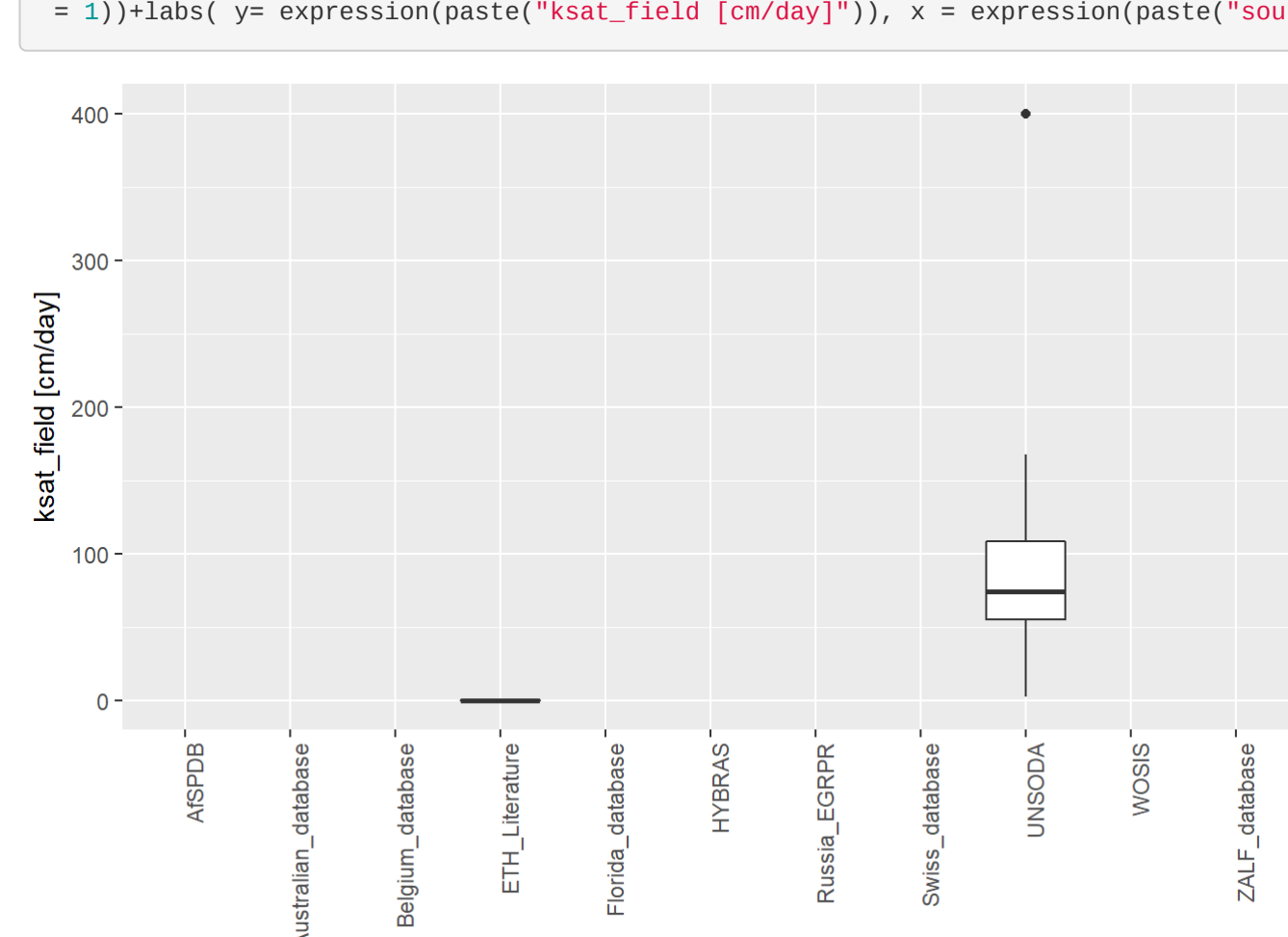
## pH vs Data sources

```
ggplot(WRC, aes(x=source_db, y=ph_h2o)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("ph_h2o [-]")), x = expression(paste("source_db [-]" )))
```



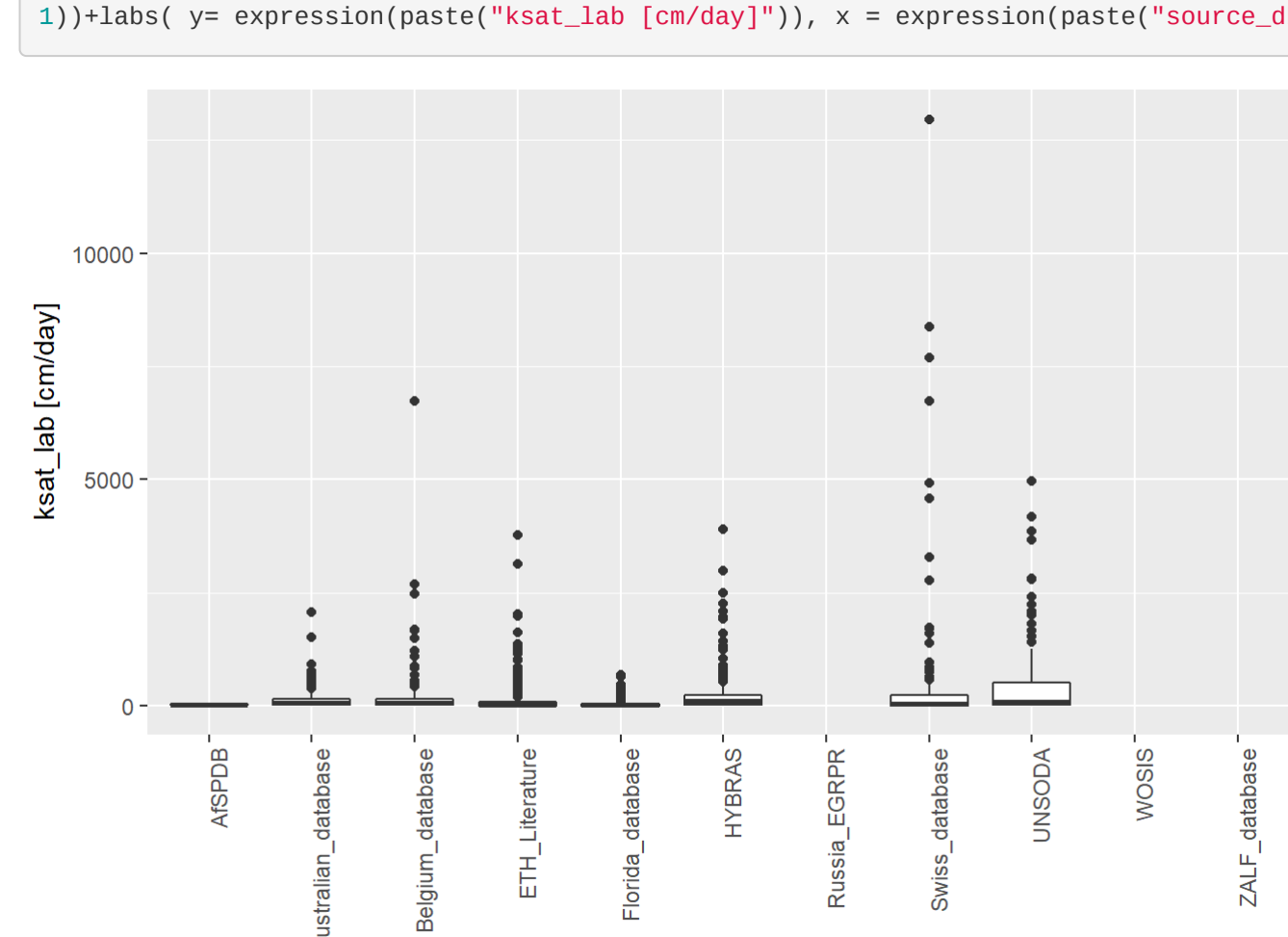
## Field measured saturated hydraulic conductivity vs Data sources

```
ggplot(WRC, aes(x=source_db, y=kcat_field)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("kcat_field [cm/day]")), x = expression(paste("source_db [-]" )))
```



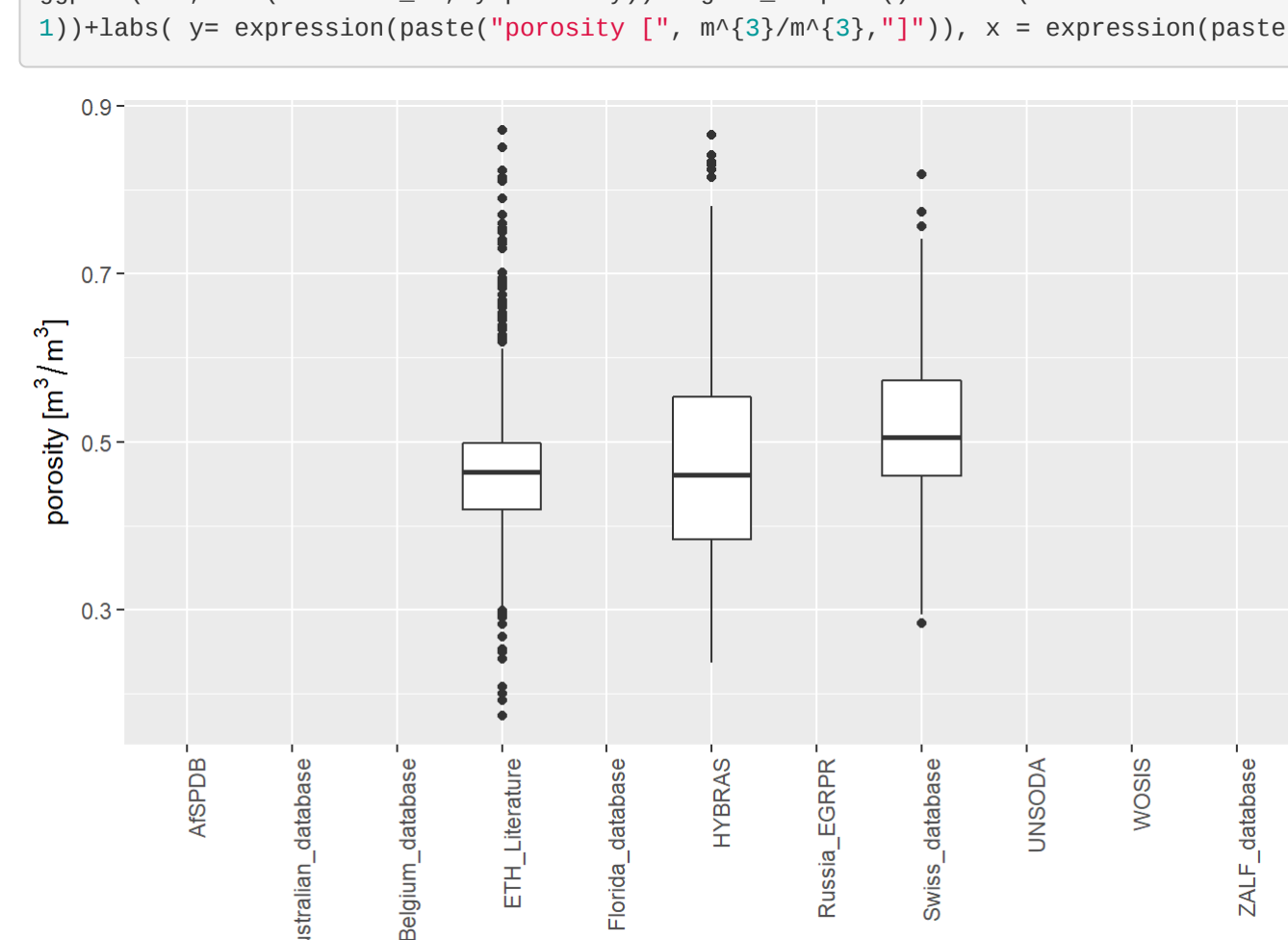
## Lab measured saturated hydraulic conductivity vs Data sources

```
ggplot(WRC, aes(x=source_db, y=kcat_lab)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("kcat_lab [cm/day]")), x = expression(paste("source_db [-]" )))
```



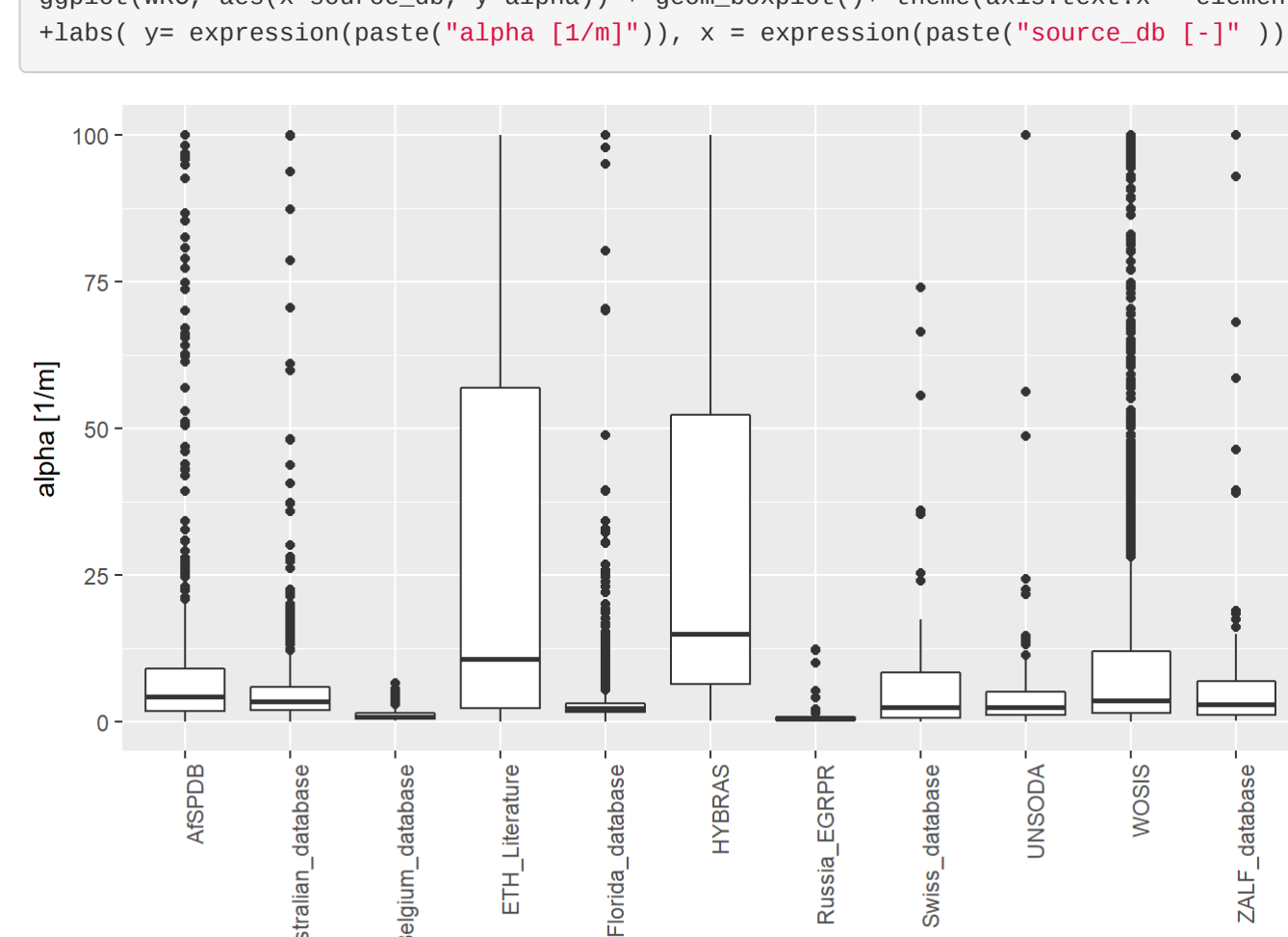
## Porosity vs Data sources

```
ggplot(WRC, aes(x=source_db, y=porosity)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("porosity [-]")), x = expression(paste("source_db [-]" )))
```



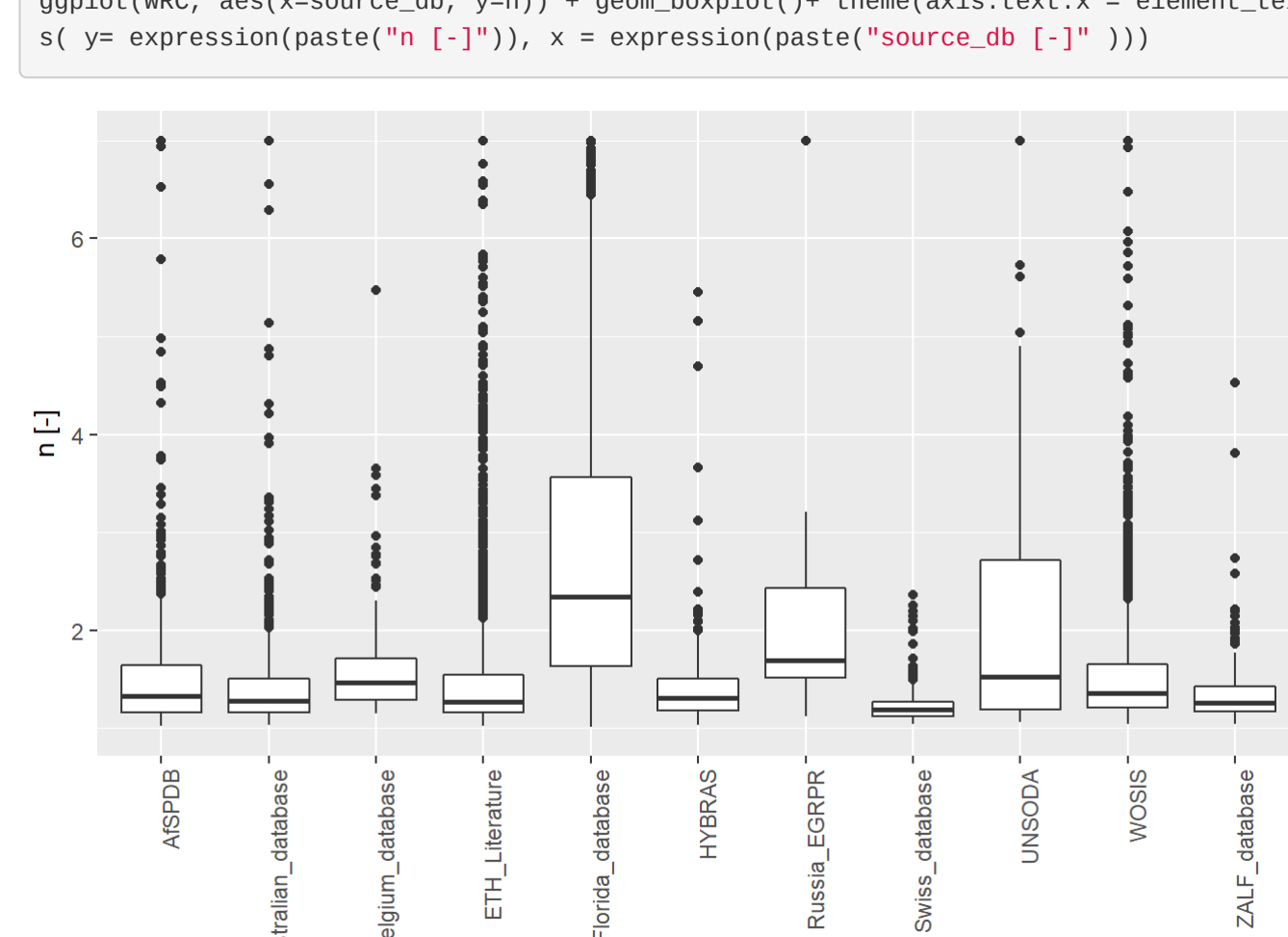
## VG shape parameter vs Data sources

```
ggplot(WRC, aes(x=source_db, y=n)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("n [-]")), x = expression(paste("source_db [-]" )))
```



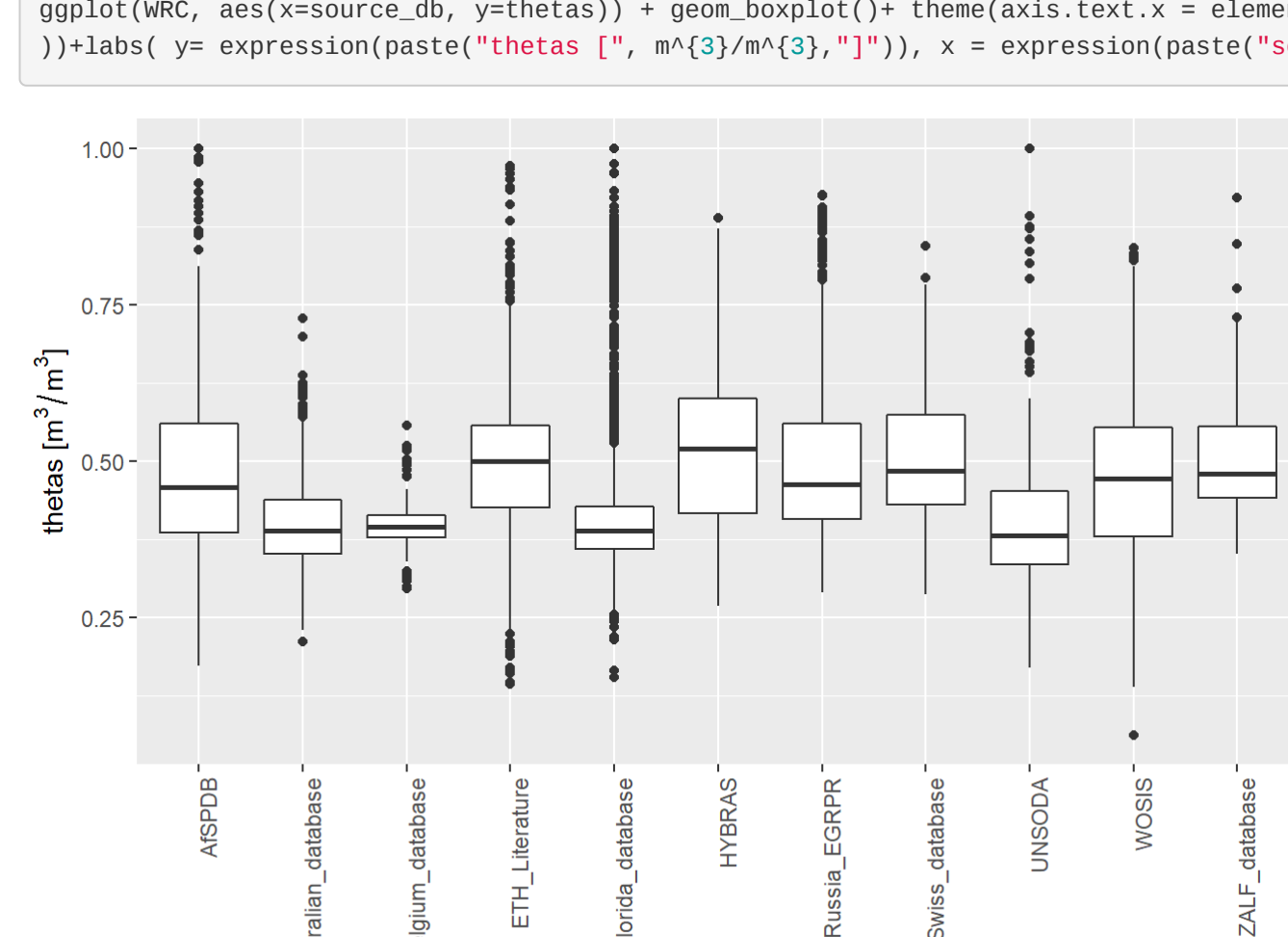
## VG shape parameter vs Data sources

```
ggplot(WRC, aes(x=source_db, y=theta_s)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("theta_s [m^3/m^3]")), x = expression(paste("source_db [-]" )))
```



## Saturated water content vs Data sources

```
ggplot(WRC, aes(x=source_db, y=theta_r)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("theta_r [m^3/m^3]")), x = expression(paste("source_db [-]" )))
```



## Residual water content vs Data sources

```
ggplot(WRC, aes(x=source_db, y=theta_r)) + geom_boxplot() + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
labs(y = expression(paste("theta_r [m^3/m^3]")), x = expression(paste("source_db [-]" )))
```

