RAM_SR_analysis

Hao Ye

17 November, 2015

Contents

Setup

Load in the packages we need and the analysis functions.

```
library(rEDM)
library(parallel)
library(quantreg)

source("my_functions.R")
```

Pre-processing

Read the data and documentation from the RAM database website.

```
extract_data() # pull data from RAM website
process_data() # process SR data
summarize_data()
```

Check overlap between web files and zip file from Steve.

```
get_doc_info()
load("stock_ids.Rdata")
both <- intersect(andi_stock_ids, web_stock_ids)
a_not_w <- setdiff(andi_stock_ids, web_stock_ids)
w_not_a <- setdiff(web_stock_ids, andi_stock_ids)
```

Analysis

Do univariate analysis. For each stock, do simplex and s-map, find whether AR-0 or AR-1 fits better using AICc, compute simplex and s-map for 1000 AR surrogates.

```
run_univariate_analysis()
append_univariate_analysis()
```

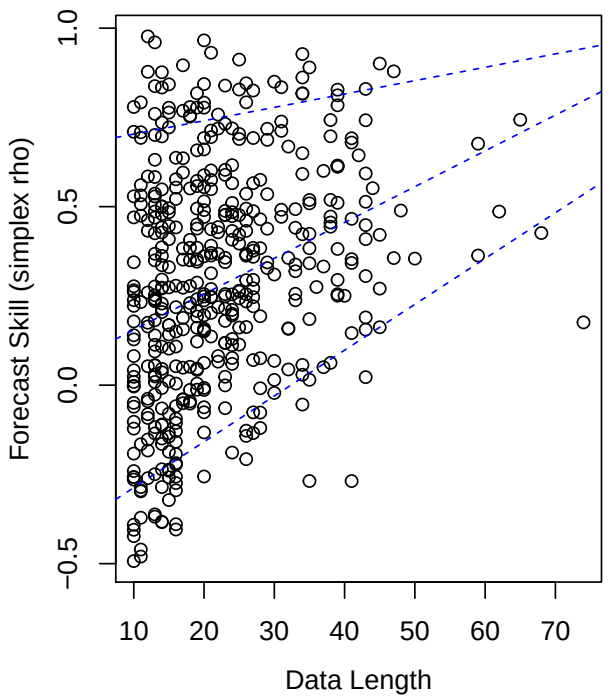
```
load("sr_results.Rdata")

results <- do.call(rbind, lapply(sr_results, function(x) {
    if(class(x) == "try-error")
        return(NA)

    # pull out EDM performance
    simplex_temp <- x$simplex_out[x$simplex_out$tp == 1,]
    smap_temp <- x$smap_out[x$smap_out$tp == 1,]
    n <- sum(is.finite(x$rec))
    best_E <- x$best_E
    simplex_rho <- simplex_temp$rho[simplex_temp$E == best_E]
    simplex_mae <- simplex_temp$mae[simplex_temp$E == best_E]
    smap_rho <- max(smap_temp$rho)
    smap_mae <- min(smap_temp$mae)

    # compute AR performance
    if(x$ar_0$aicc < x$ar_1$aicc)
    {
        ar_pred <- x$rec - x$ar_0$residuals
    } else {
        ar_pred <- x$rec - x$ar_1$residuals
    }
    ar_rho <- cor(x$rec, ar_pred, use = "pairwise")
    ar_mae <- mean(abs(x$rec - ar_pred), na.rm = TRUE)

    # compute EDM p-values
    simplex_rho_p <- (sum(simplex_rho < x$simplex_null$rho)+1) / (NROW(x$simplex_null) + 1)
    smap_rho_p <- (sum(smap_rho < x$smap_null$rho)+1) / (NROW(x$smap_null) + 1)
    simplex_mae_p <- (sum(simplex_mae > x$simplex_null$mae)+1) / (NROW(x$simplex_null) + 1)
    smap_mae_p <- (sum(smap_mae > x$smap_null$mae)+1) / (NROW(x$smap_null) + 1)

    return(data.frame(n = n,
                      simplex_rho = simplex_rho,
                      simplex_mae = simplex_mae,
                      smap_rho = smap_rho,
                      smap_mae = smap_mae,
```

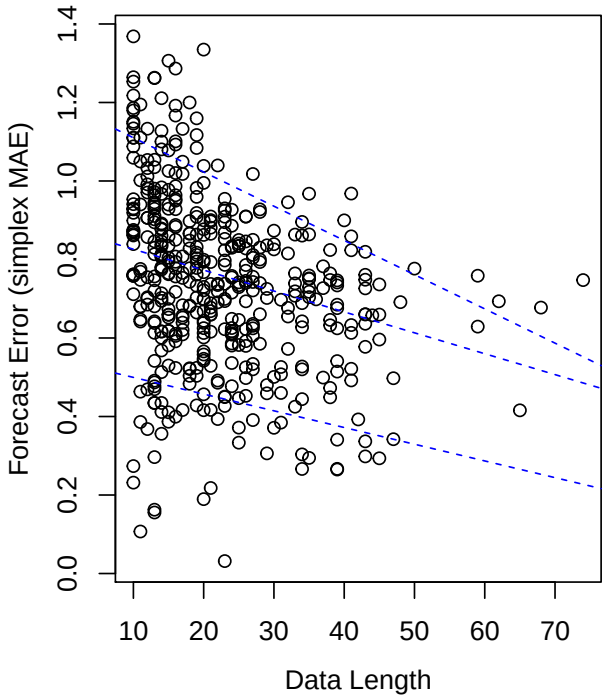
```
                        ar_rho = ar_rho,
                        ar_mae = ar_mae,
                        simplex_rho_p = simplex_rho_p,
                        smap_rho_p = smap_rho_p))
}))
```

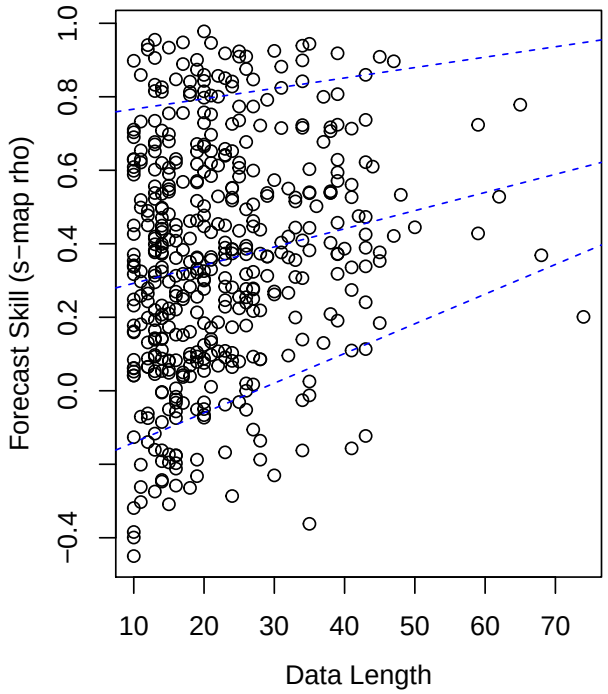
Plots

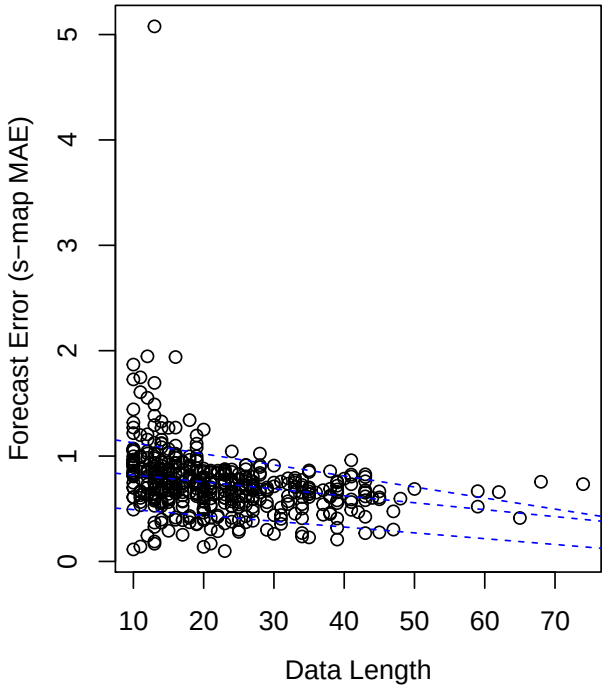
Determine how predictability changes with data length.

```
## Warning in rq.fit.br(x, y, tau = tau, ...): Solution may be nonunique
```

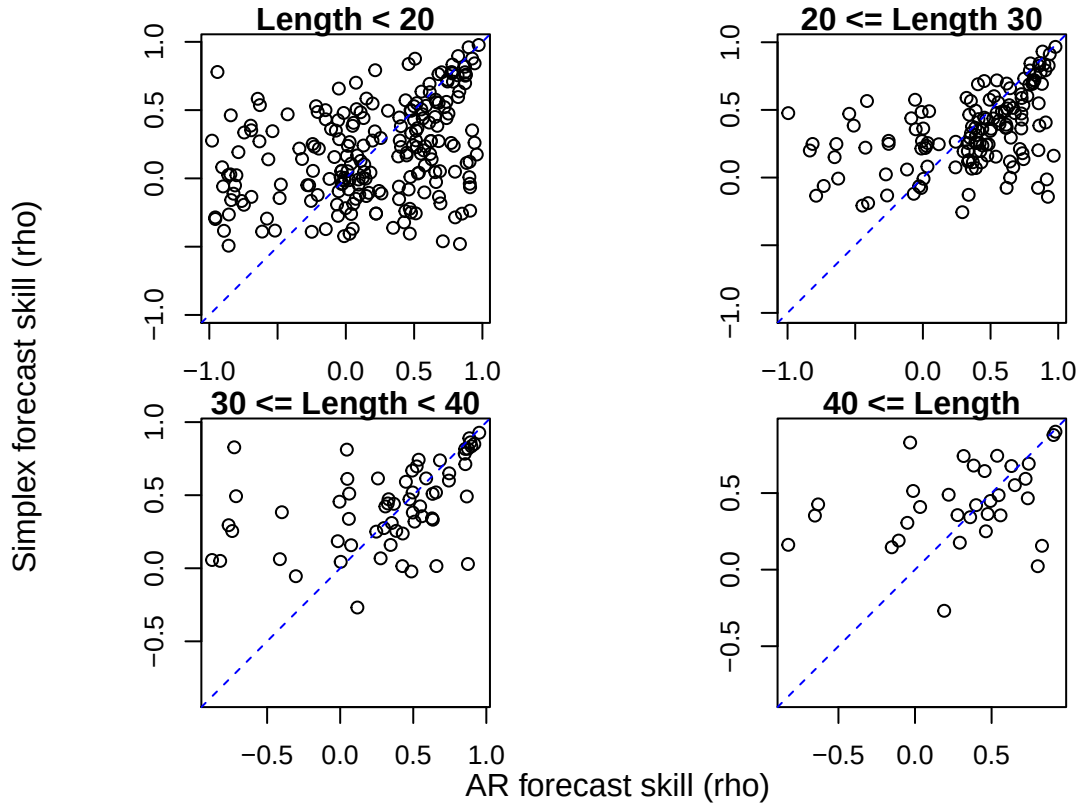


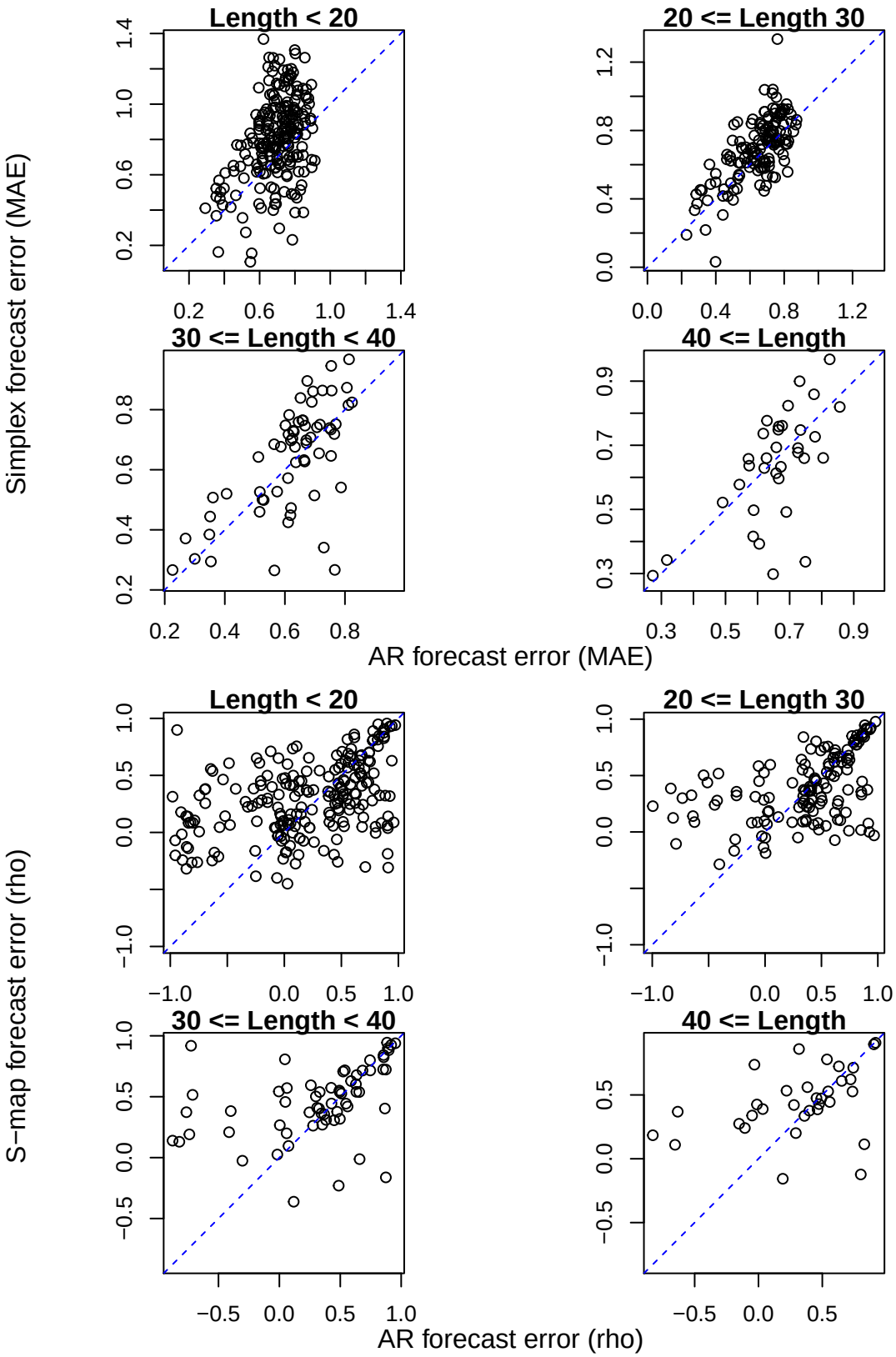


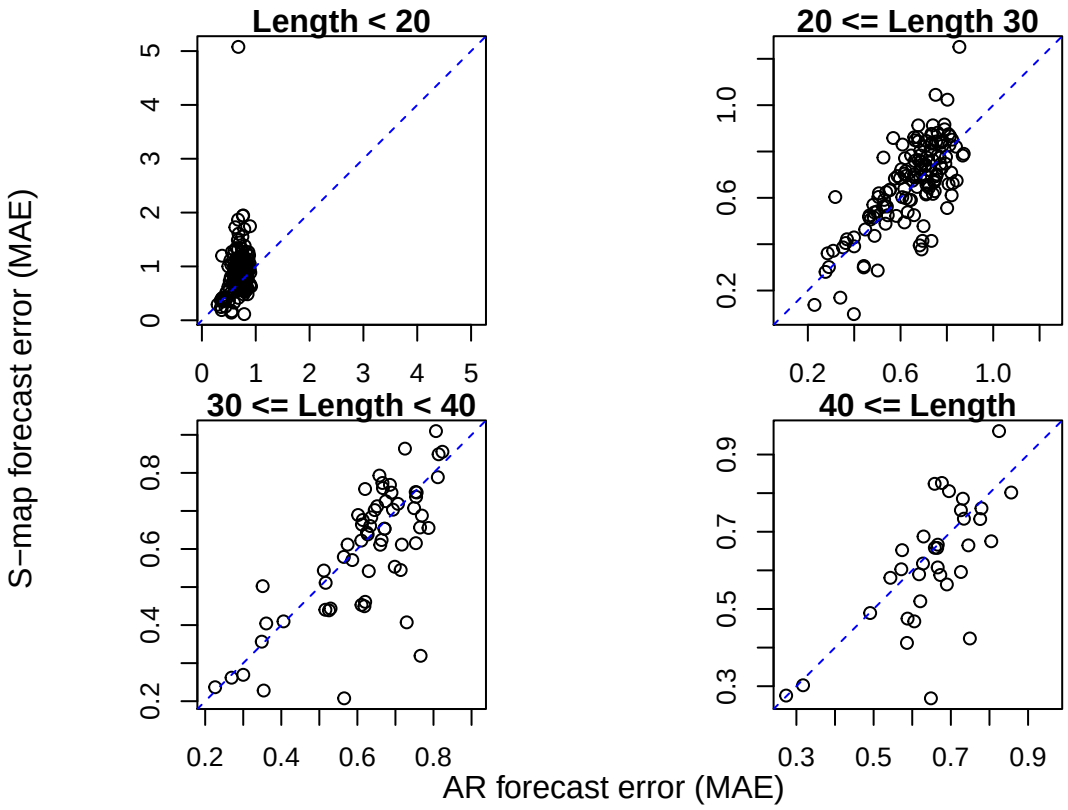




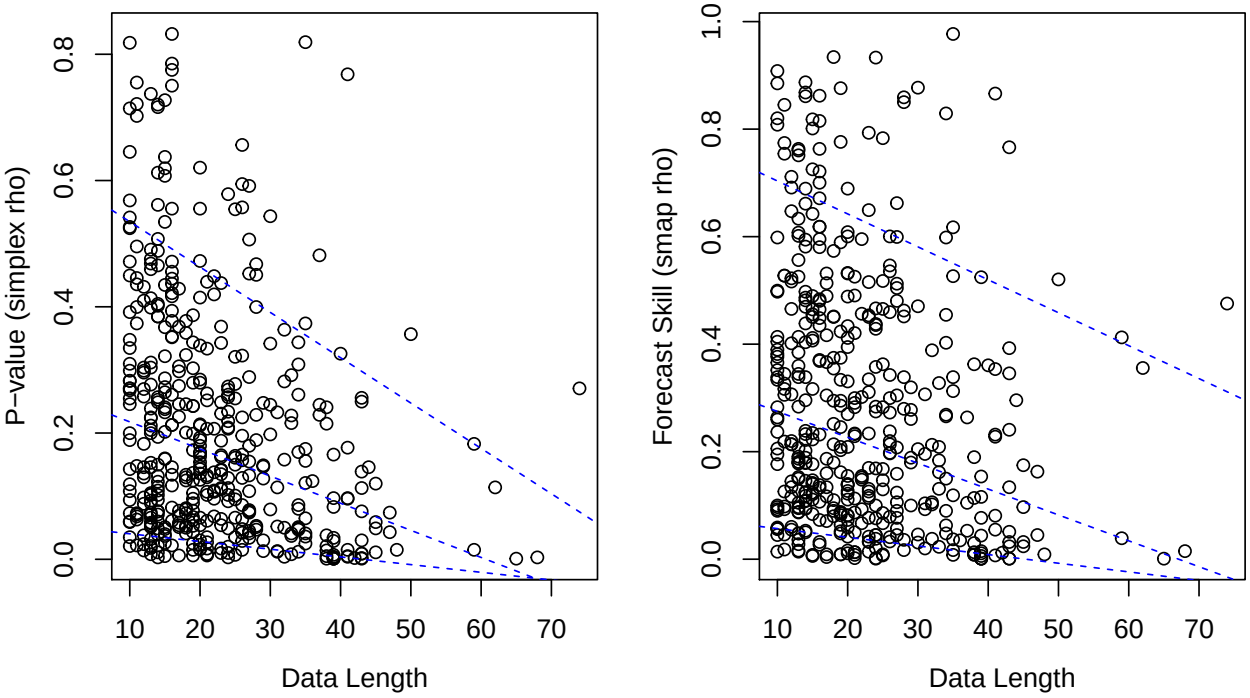
Comparisons of EDM performance vs. AR models




Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
Simplex forecast error (MAE)
AR forecast error (MAE)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast error (rho)
AR forecast error (rho)



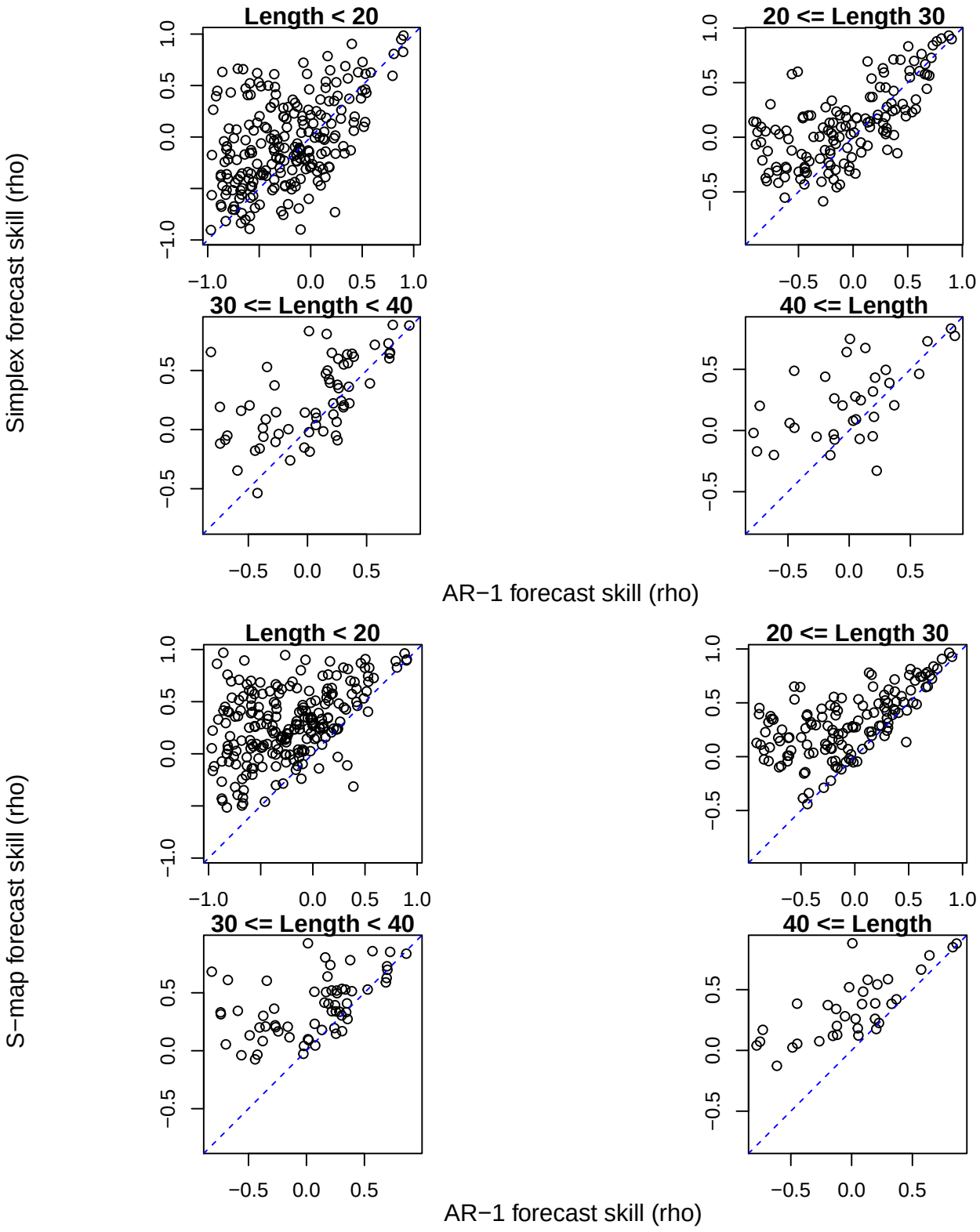
P-values (probability that recruitment can be modeled using AR)

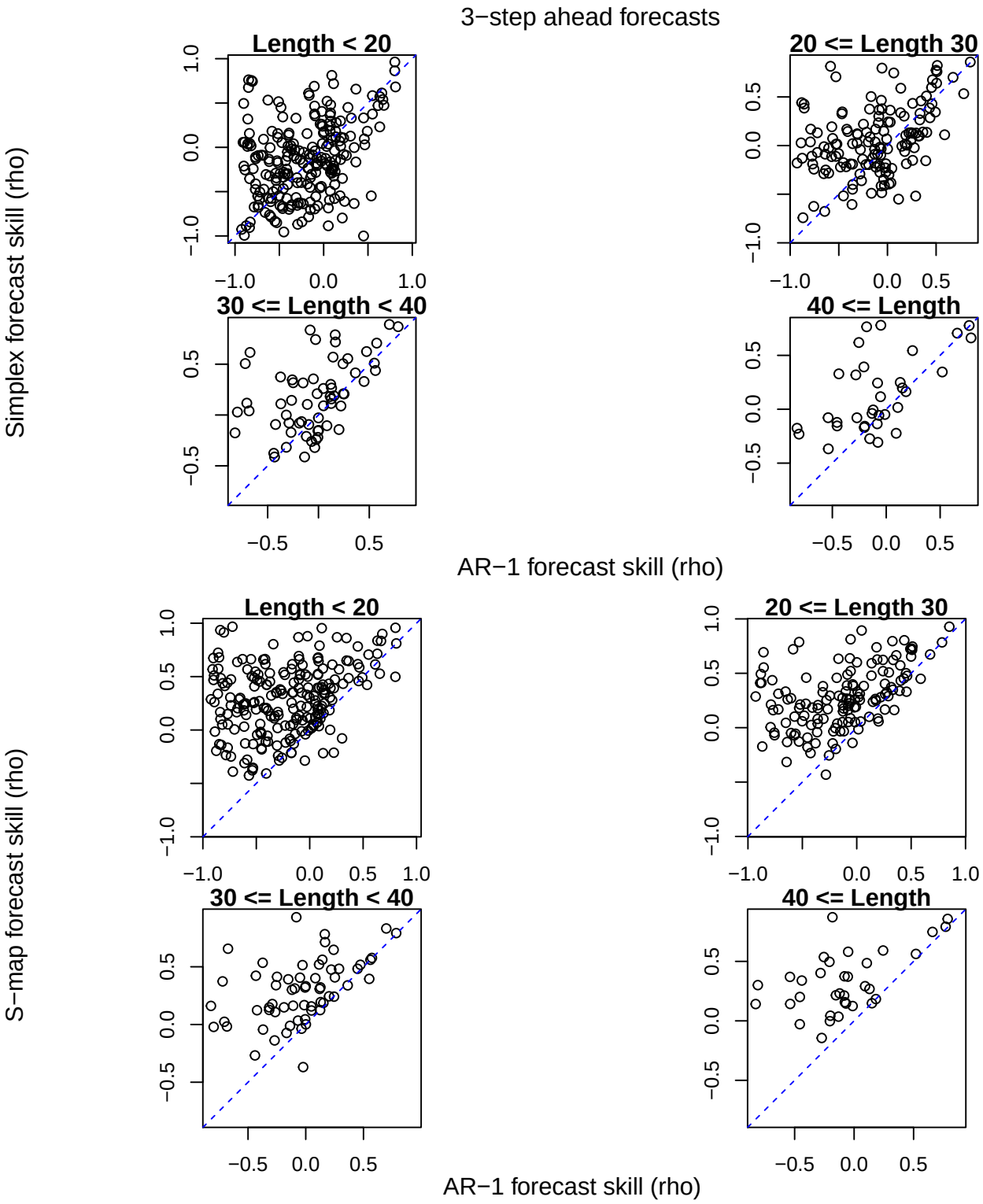


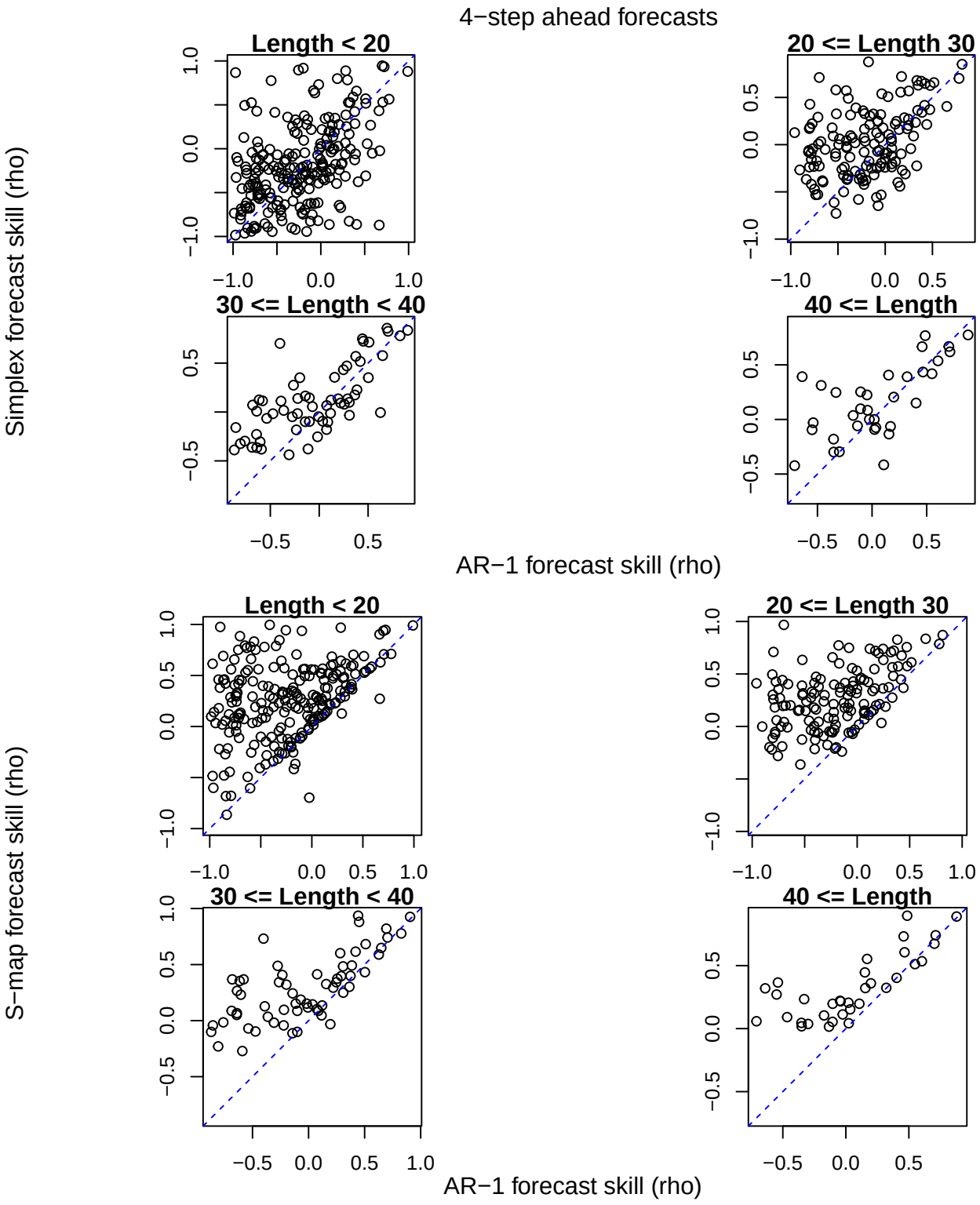
Plots (part 2)

Comparisons of EDM performance vs. AR models (2-, 3-, 4-step ahead forecasts)

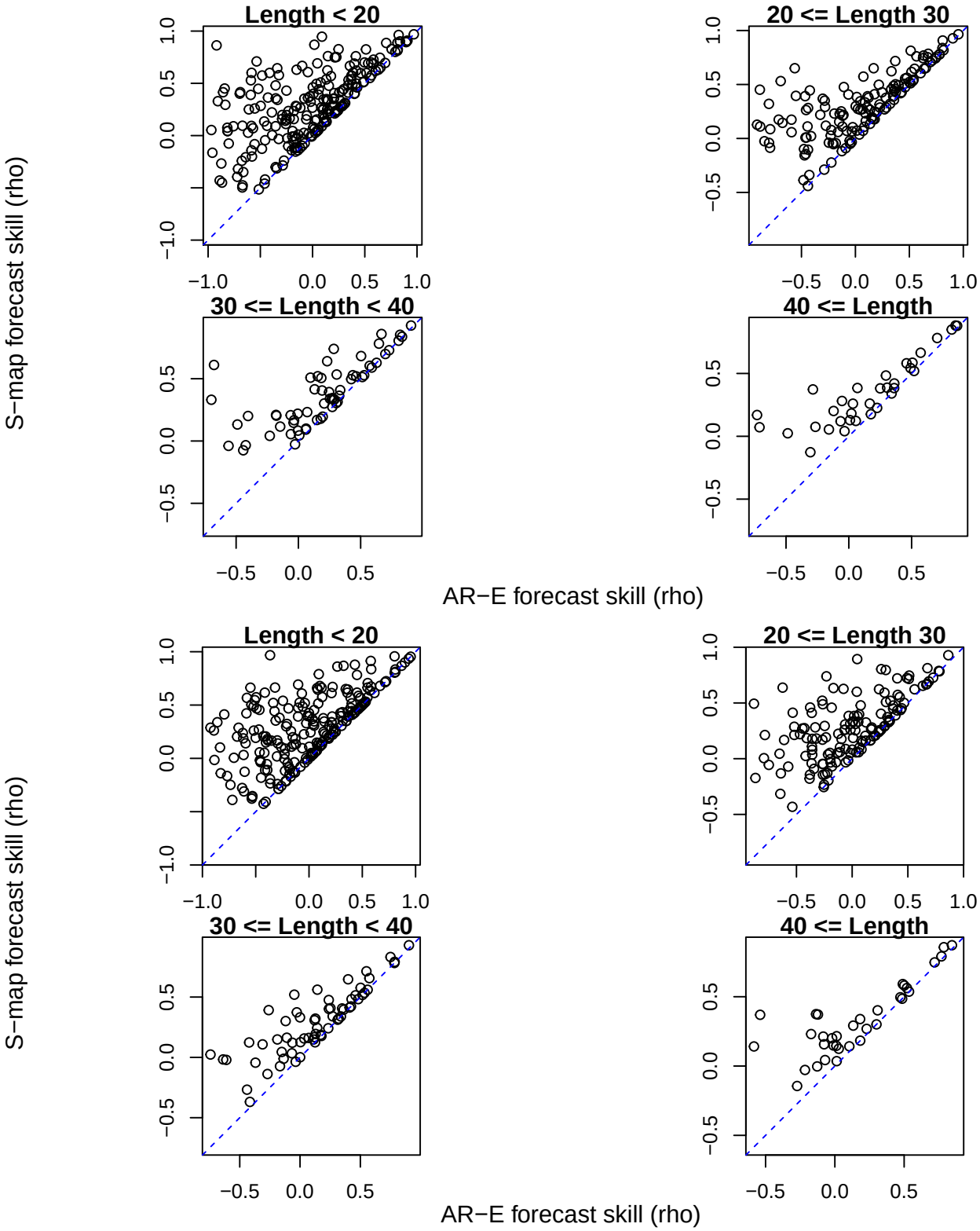
```
get_results <- function(selected_tp = 2)
{
    return(do.call(rbind, lapply(sr_results, function(x) {
        if(class(x) == "try-error")
            return(NA)

        # pull out EDM performance
        simplex_temp <- x$simplex_out[x$simplex_out$tp == selected_tp,]
        smap_temp <- x$smap_out[x$smap_out$tp == selected_tp,]
        n <- sum(is.finite(x$rec))
        best_E <- x$best_E
        simplex_rho <- simplex_temp$rho[simplex_temp$E == best_E]
        simplex_mae <- simplex_temp$mae[simplex_temp$E == best_E]
        smap_rho <- max(smap_temp$rho)
        smap_mae <- min(smap_temp$mae)
        ar1_rho <- x$ar1_out$rho[x$ar1_out$tp == selected_tp]
        ar1_mae <- x$ar1_out$mae[x$ar1_out$tp == selected_tp]

        return(data.frame(n = n,
                          simplex_rho = simplex_rho,
                          simplex_mae = simplex_mae,
                          smap_rho = smap_rho,
                          smap_mae = smap_mae,
                          ar1_rho = ar1_rho,
                          ar1_mae = ar1_mae))
    }))
}
```

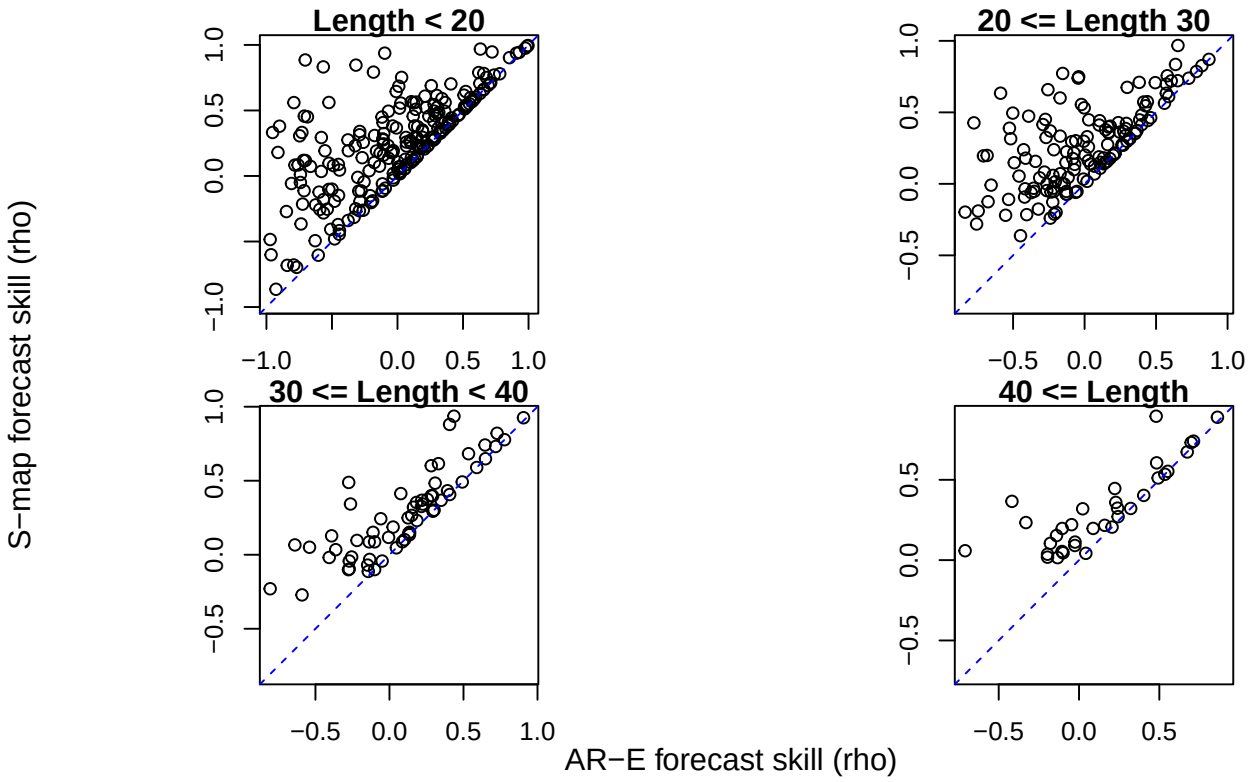

Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
Simplex forecast skill (rho)
AR−1 forecast skill (rho)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast skill (rho)
AR−1 forecast skill (rho)


3−step ahead forecasts
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
Simplex forecast skill (rho)
AR−1 forecast skill (rho)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast skill (rho)
AR−1 forecast skill (rho)


4−step ahead forecasts
Simplex forecast skill (rho)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
AR−1 forecast skill (rho)
S−map forecast skill (rho)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
AR−1 forecast skill (rho)

Plots (part 3)


Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast skill (rho)
AR−E forecast skill (rho)
Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast skill (rho)
AR−E forecast skill (rho)


Length < 20
20 <= Length 30
30 <= Length < 40
40 <= Length
S−map forecast skill (rho)
AR−E forecast skill (rho)
−1.0
−0.5
0.0
0.5
1.0