

model predict

Shengzhi Luo

05/05/2022

```
beta.res.postmean = beta.res.postmean %>% rename(beta_0 = intercept, beta_1 = Wind_prev, beta_2 = Lat_change, beta_3 = Long_change, beta_4 = Wind_change)

dt_res = merge(dt_test_id, beta.res.postmean, by = "ID")

dt_res = dt_res %>%
  mutate(Wind_kt_pred = beta_0*intercept+beta_1*Wind_prev
    +beta_2*Lat_change+beta_3*Long_change+beta_4*Wind_change) %>%
  group_by(ID) %>%
  mutate(r_square = 1-(sum((Wind_kt_pred-Wind.kt)^2))/(sum((Wind.kt-mean(Wind.kt))^2)),
    rmse = rmse(Wind.kt,Wind_kt_pred))

dt_rmse=
dt_res %>%
dplyr::select(ID, r_square, rmse) %>%
distinct() %>%
mutate(r_square = round(r_square, 3),
  rmse = round(rmse,3)) %>%
filter(r_square > 0 && r_square < 1) %>%
arrange(rmse)
mean(dt_rmse$r_square)

## [1] 0.82245

mean(dt_rmse$rmse)

## [1] 4.51023

df = head(dt_rmse,10)
png("rmse.png", height = 25*nrow(df), width = 100*ncol(df), bg = "white")
grid.table(df)
dev.off()

## pdf
## 2

png("q4_prediction_plot.png", width = 2000, height = 1000)
dt_res %>% filter(ID %in% c( "ERIN.2013", "EDITH.1967", "DEBBY.2006", "GUSTAV.1996")) %>%
  mutate(index = 1:n()) %>%
  ggplot(aes(y = Wind.kt, x = index))+
  geom_line(aes(color = "Actual Wind Speed"))+
  geom_line(aes(y = Wind_kt_pred, color = "Predicted Wind Speed"))+
  facet_wrap(~ID, nrow = 2, scales = "free") +
  theme_classic()+
  labs(x = "Index",
    y = "Wind Speed",
```

```
      title = "Actual Wind Speed vs. Predicted Wind Speed")
dev.off()

## pdf
## 2
```