Energy norm 1

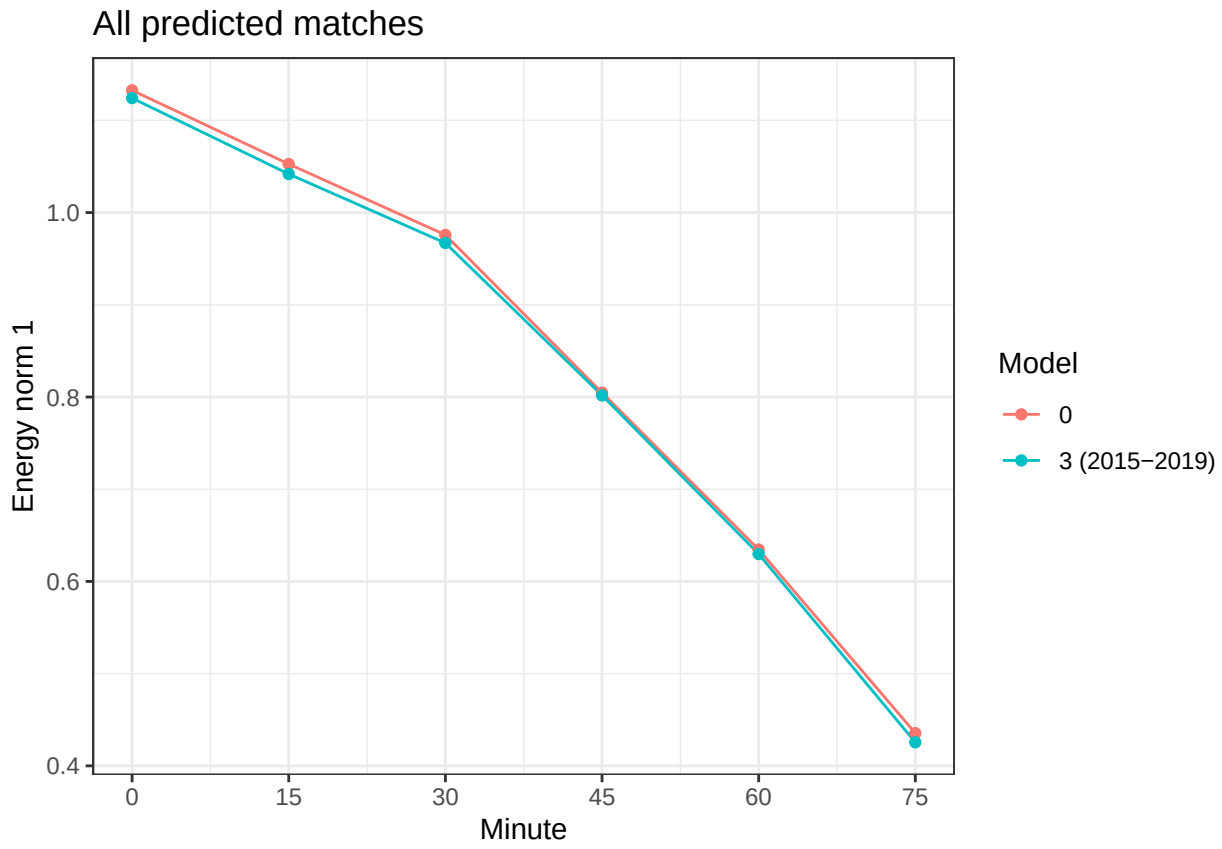
```
library(dplyr)
library(ggplot2)
library(tidyr)
library(knitr)

load("data/HDA2.RData")
load("~/GitHub/soccer-live-predictions/soccer-live-predictions/scrape/data/reds.RData")

nrow(HDA2)
```

```
## [1] 340
```

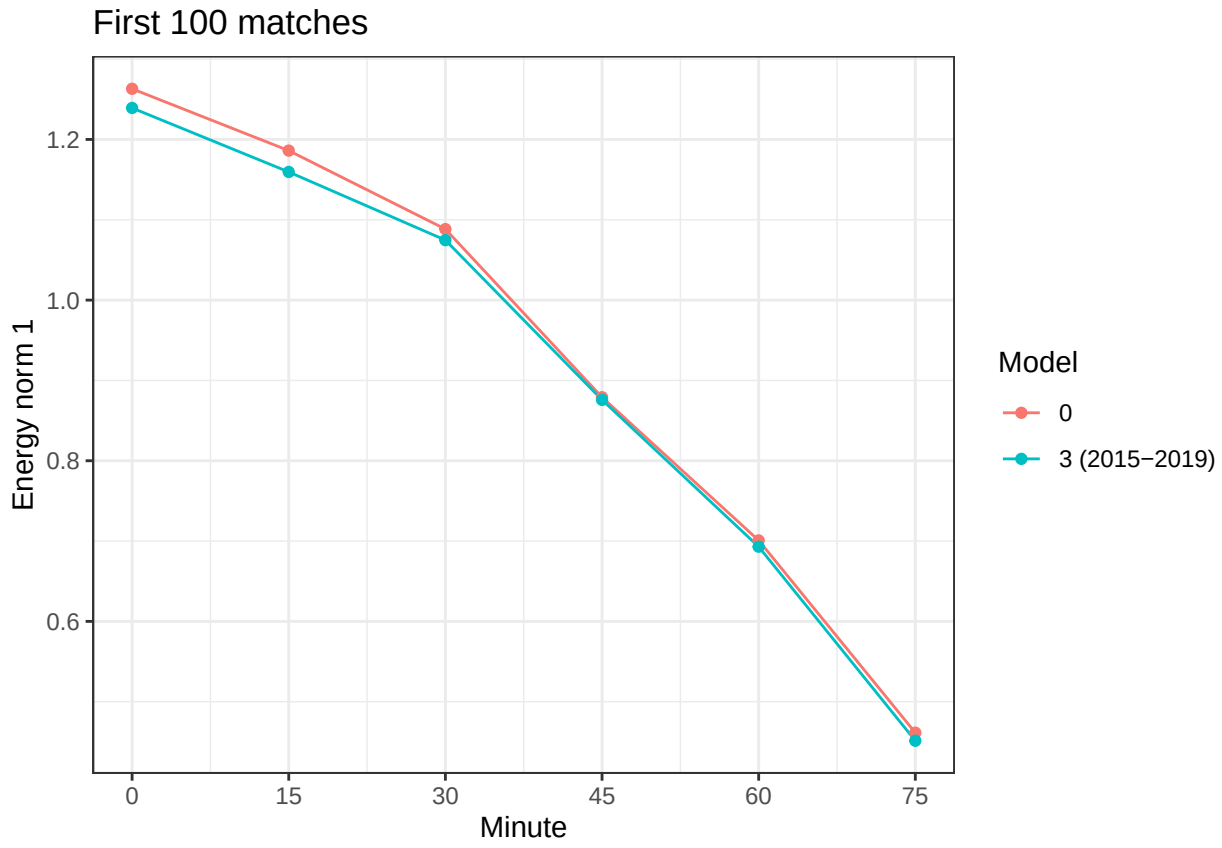
```
all = tibble(ENERG1 = apply(HDA2[,c(81:92)], 2, mean),
             Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
             Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                            levels = c("0", "3 (2015-2019)")))
all %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches") +
  ylab("Energy norm 1")
```




```
all %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.132575	1.052664	0.9757626	0.8046796	0.6346003	0.4354614
3 (2015-2019)	1.124165	1.041780	0.9670173	0.8017906	0.6297095	0.4255535

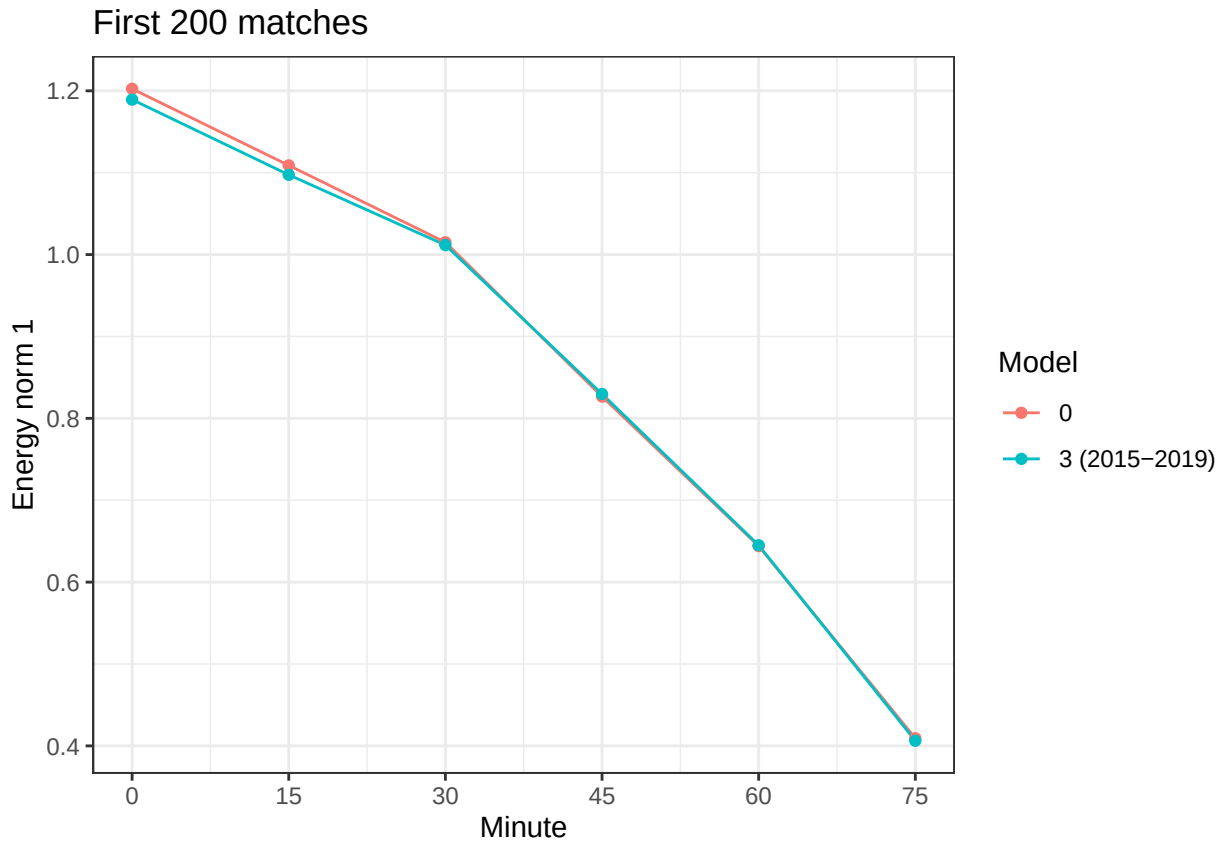
```
first_100 = tibble(ENERG1 = apply(HDA2[c(1:100), c(81:92)], 2, mean),
                   Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
                   Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                                  levels = c("0", "3 (2015-2019)")))
first_100 %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("First 100 matches") +
  ylab("Energy norm 1")
```




```
first_100 %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.263071	1.186114	1.088375	0.8790907	0.7007795	0.4614044
3 (2015-2019)	1.239198	1.159543	1.074646	0.8757628	0.6929703	0.4513803

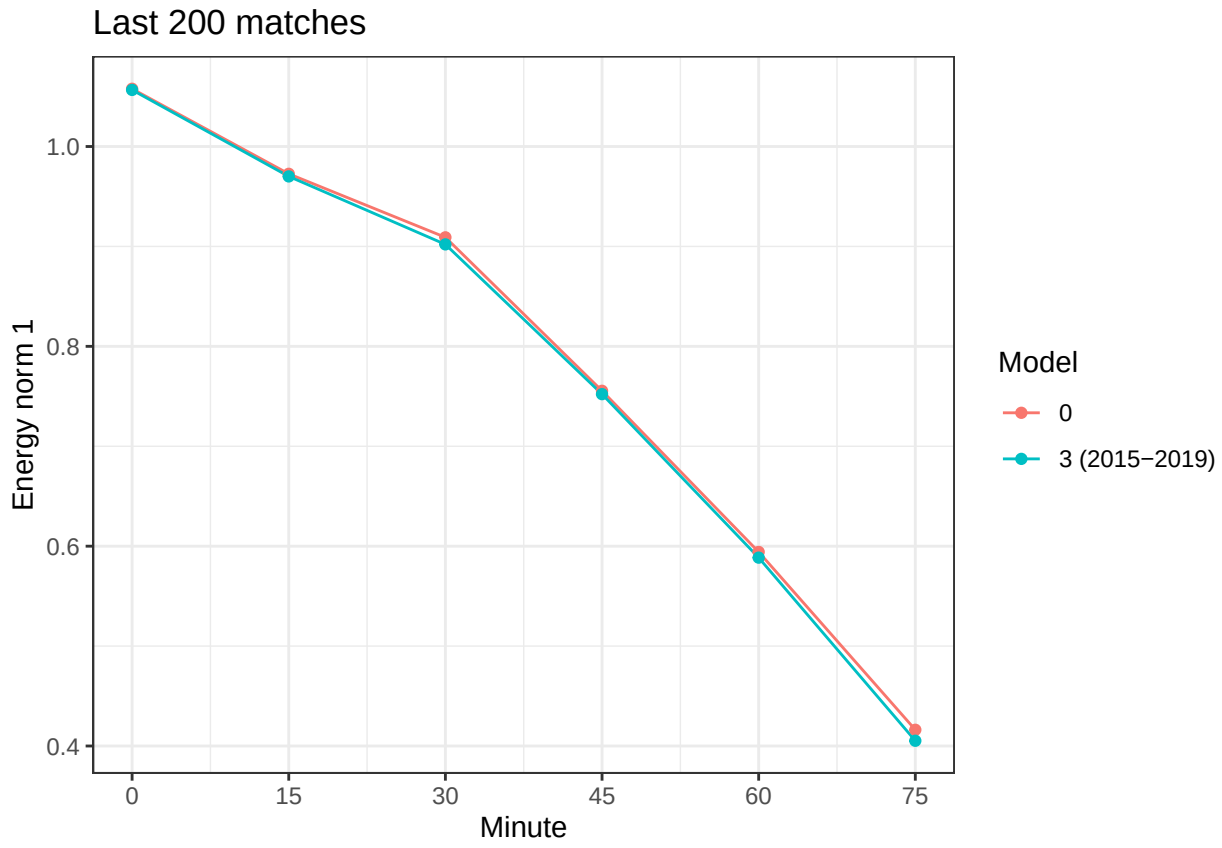
```
first_200 = tibble(ENERG1 = apply(HDA2[c(1:200), c(81:92)], 2, mean),
                   Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
                   Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                                  levels = c("0", "3 (2015-2019)")))
first_200 %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("First 200 matches") +
  ylab("Energy norm 1")
```




```
first_200 %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.202478	1.108932	1.015021	0.8266335	0.6439369	0.4093750
3 (2015-2019)	1.189279	1.097404	1.011607	0.8297361	0.6450285	0.4062678

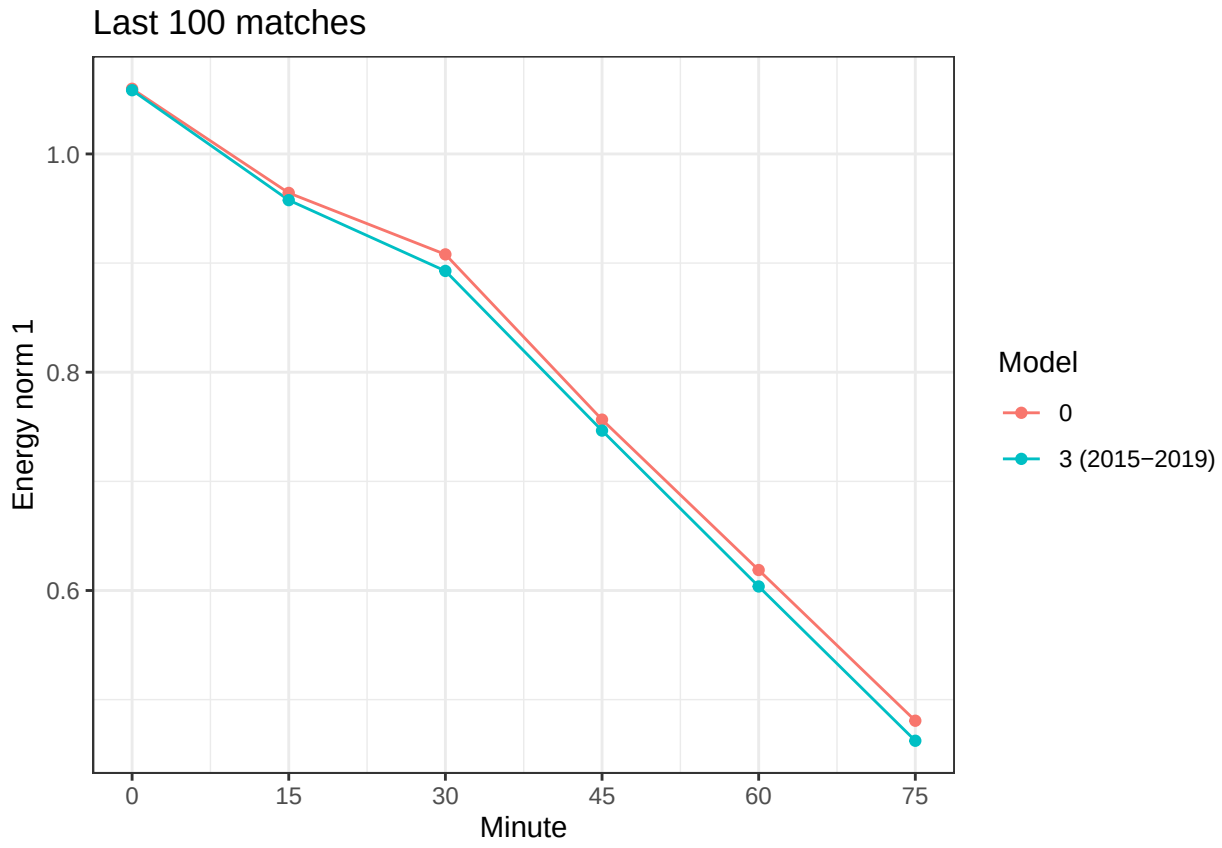
```
last_200 = tibble(ENERG1 = apply(HDA2[c(141:340), c(81:92)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)")))
last_200 %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("Last 200 matches") +
  ylab("Energy norm 1")
```




```
last_200 %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.057760	0.9726480	0.9090899	0.7554823	0.5942355	0.4162339
3 (2015-2019)	1.056618	0.9700675	0.9020345	0.7521446	0.5885169	0.4052530

```
last_100 = tibble(ENERG1 = apply(HDA2[c(241:340), c(81:92)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)")))
last_100 %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("Last 100 matches") +
  ylab("Energy norm 1")
```




```
last_100 %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

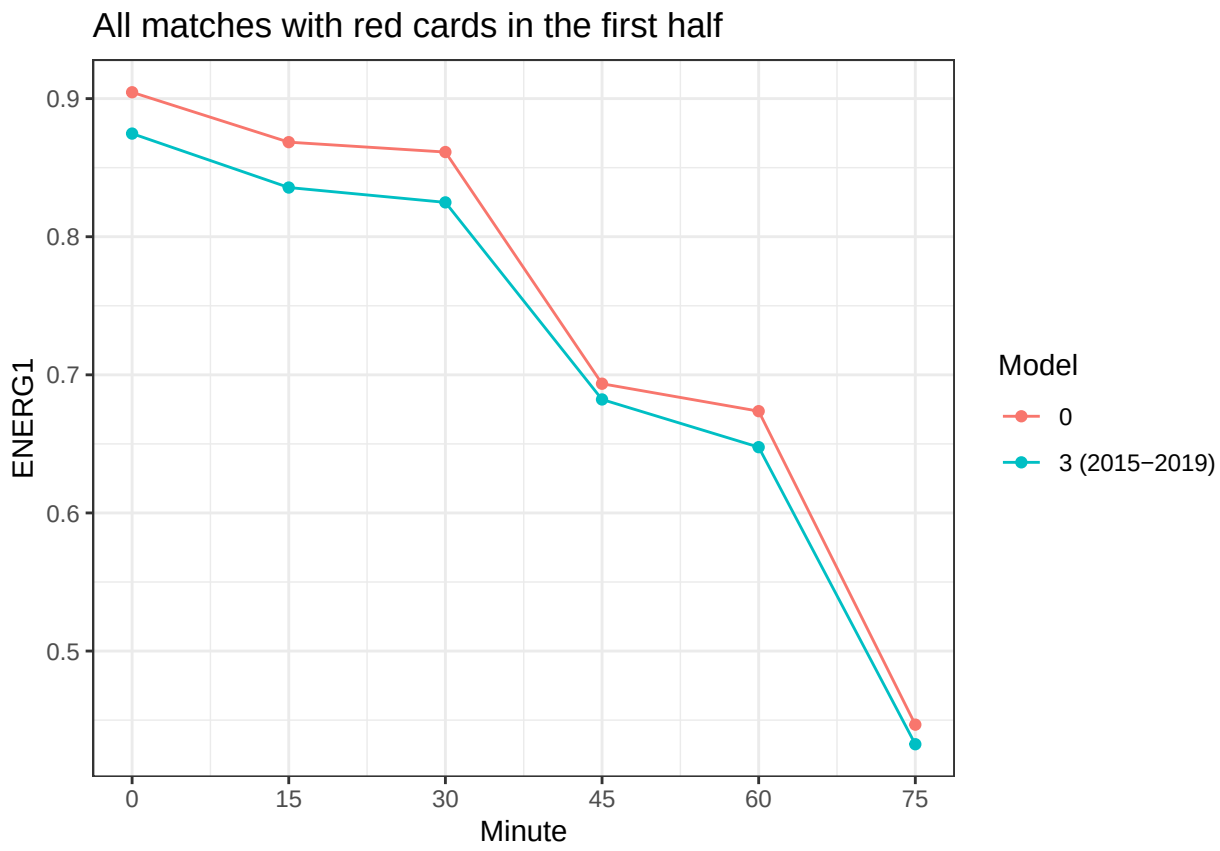
Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.059690	0.9642762	0.9079622	0.7565532	0.6187554	0.4806699
3 (2015-2019)	1.058366	0.9576515	0.8927868	0.7465129	0.6036935	0.4623193

```
matches = reds %>%
  filter(Season == 2020, Half == 1) %>%
  .$Match
length(matches)
```

```
## [1] 23
```

```
HDA2_reds = HDA2 %>%
  filter(Match %in% matches)

all_reds = tibble(ENERG1 = apply(HDA2_reds[,c(81:92)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 2)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)")))
all_reds %>%
  ggplot(aes(x = Minute, y = ENERG1, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All matches with red cards in the first half") +
  ylab("ENERG1")
```



```
all_reds %>%
  pivot_wider(id_cols = "Model", values_from = "ENERG1", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	0.9046019	0.8684704	0.8613110	0.6935566	0.6736865	0.4467058
3 (2015-2019)	0.8746847	0.8356248	0.8248627	0.6821695	0.6476292	0.4325317