Energy norm 1

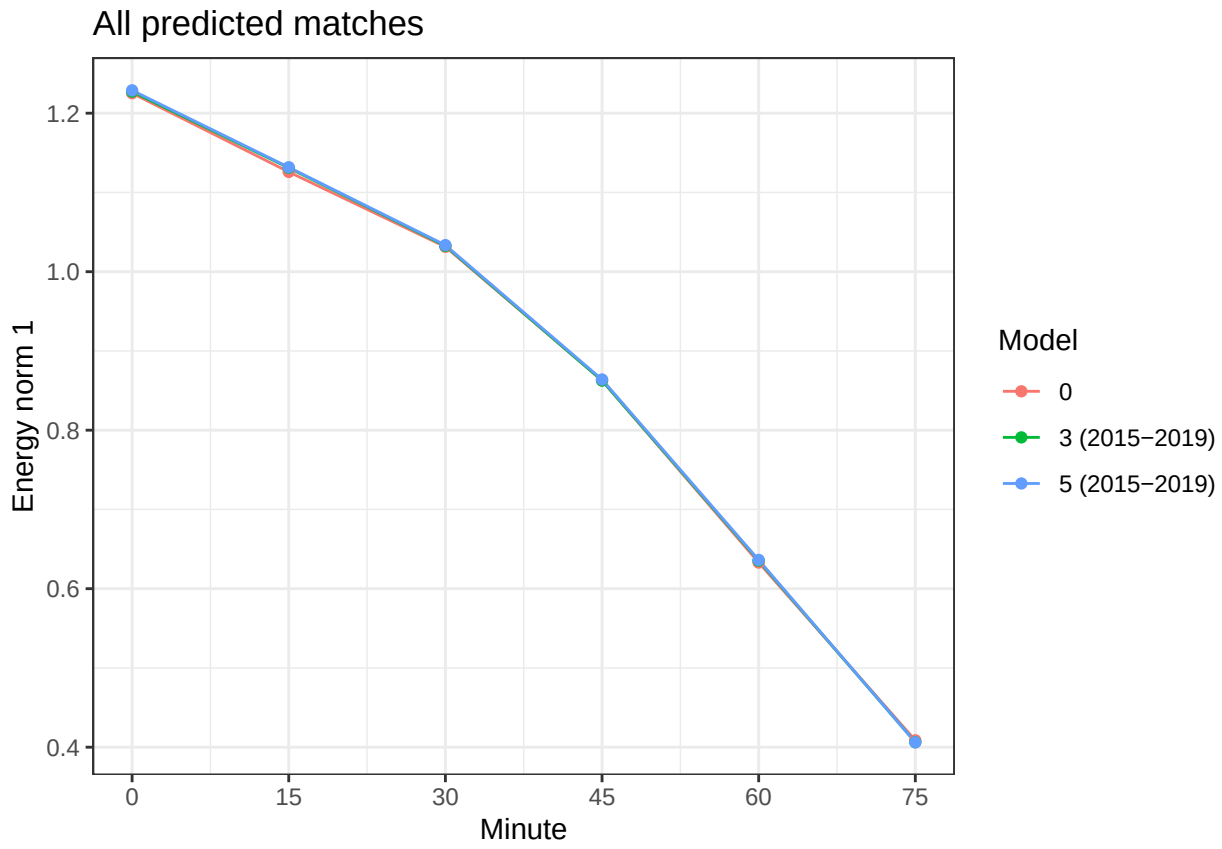
```
library(dplyr)
library(ggplot2)
library(tidyr)
library(knitr)

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

nrow(HDA2)
```

```
## [1] 333
```

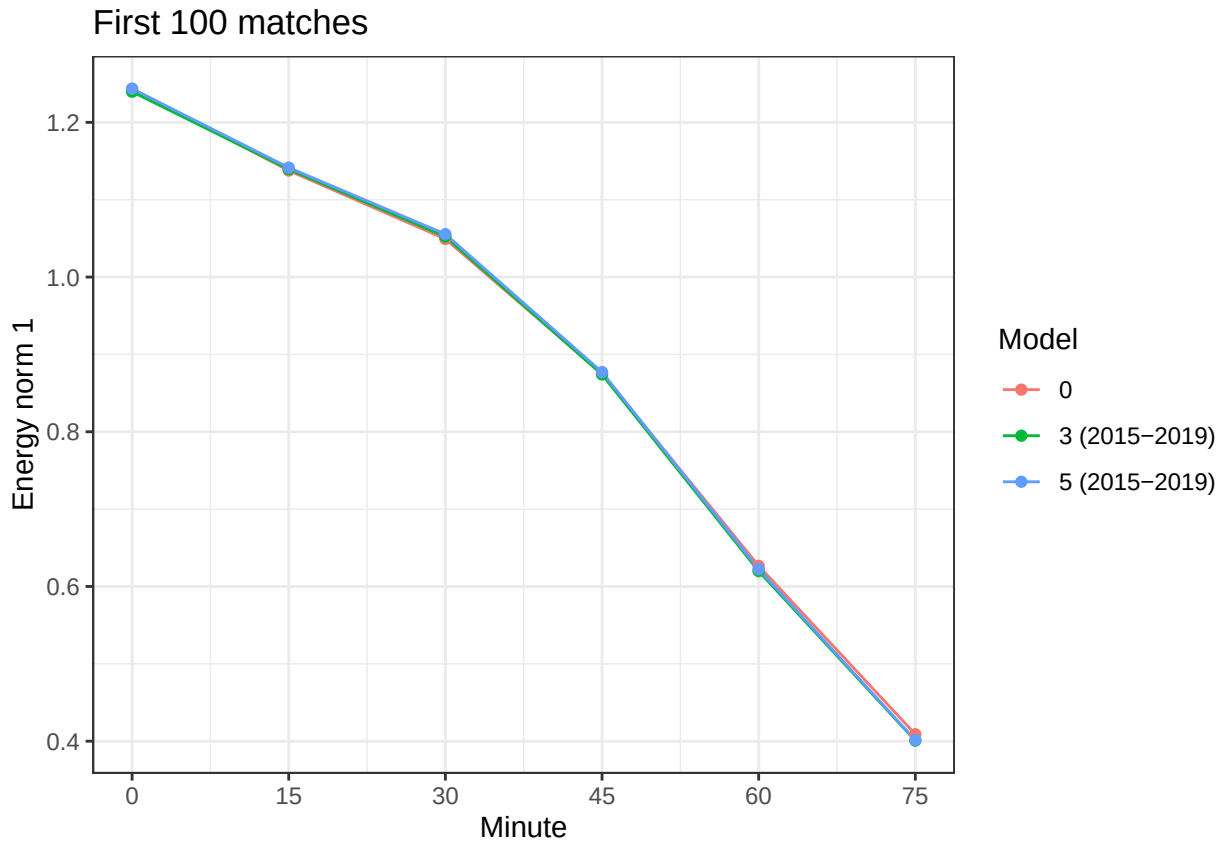
```
all = tibble(ENERG1 = apply(HDA2[,c(189:194, 207:218)], 2, mean),
             Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
             Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                              rep("5 (2015-2019)", 6)),
                            levels = c("0", "3 (2015-2019)", "5 (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	1.225244	1.125903	1.031395	0.8628340	0.6329710	0.4086262
3 (2015-2019)	1.227224	1.131052	1.032587	0.8626418	0.6353844	0.4062852
5 (2015-2019)	1.228852	1.131752	1.033521	0.8639664	0.6362743	0.4061601

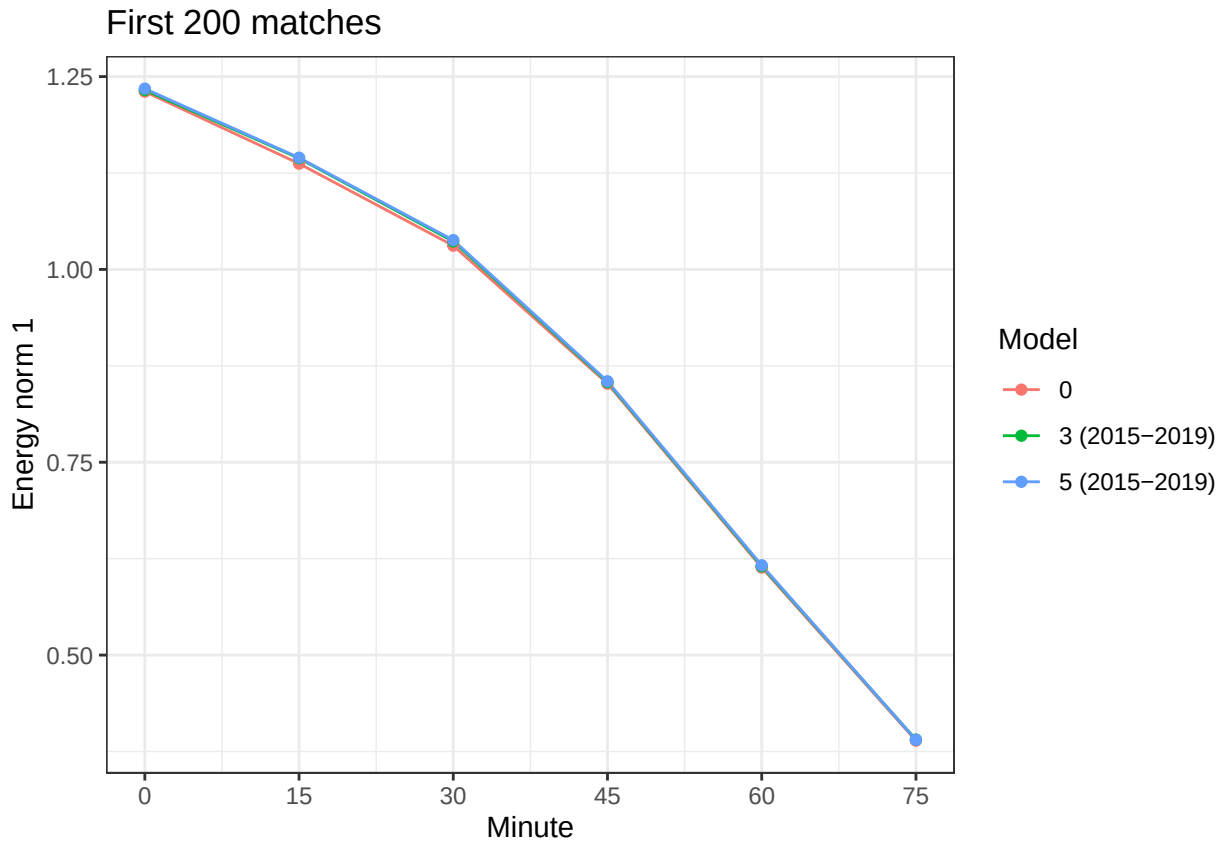
```
first_100 = tibble(ENERG1 = apply(HDA2[c(1:100), c(189:194, 207:218)], 2, mean),
                   Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
                   Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                                    rep("5 (2015-2019)", 6)),
                                  levels = c("0", "3 (2015-2019)", "5 (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	1.243311	1.137666	1.049205	0.8750657	0.6266661	0.4089884
3 (2015-2019)	1.239591	1.139408	1.052311	0.8740950	0.6198390	0.4007079
5 (2015-2019)	1.243538	1.141478	1.055467	0.8771162	0.6223633	0.4014967

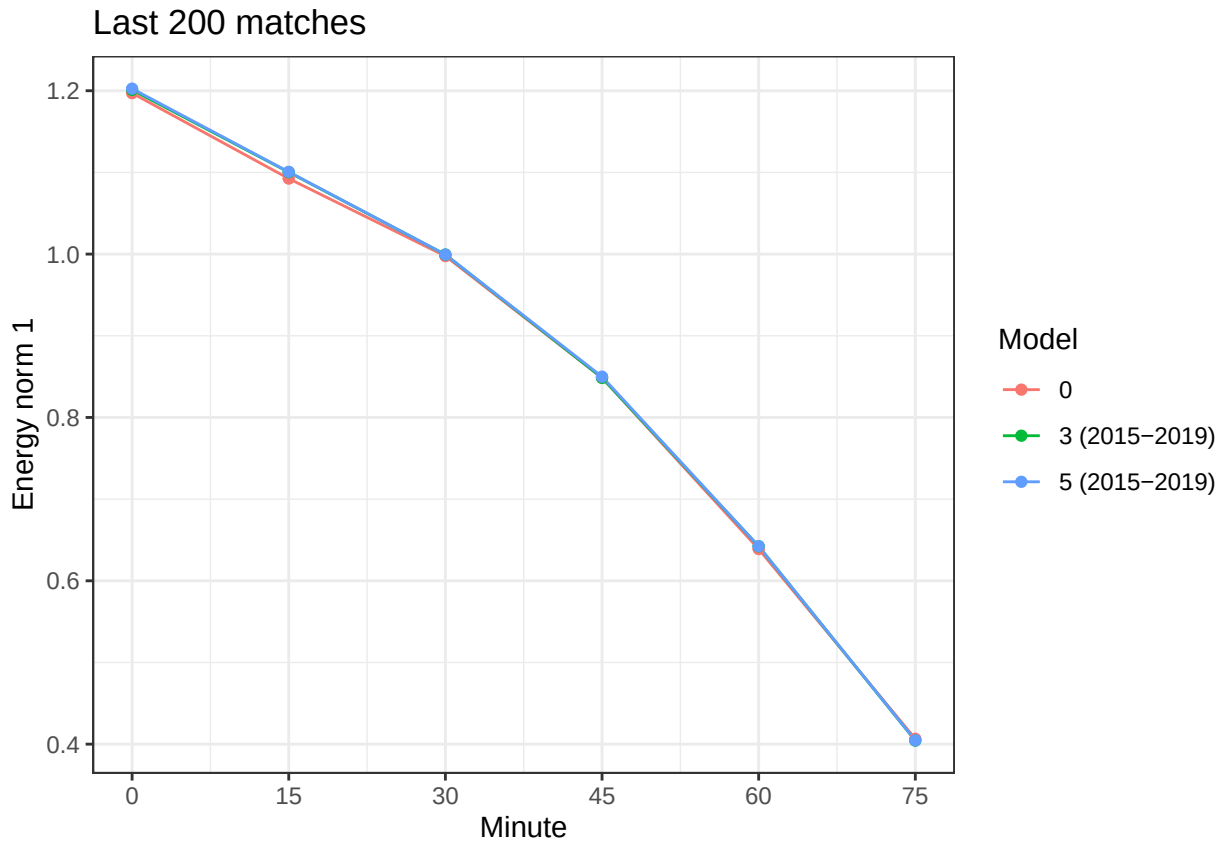
```
first_200 = tibble(ENERG1 = apply(HDA2[c(1:200), c(189:194, 207:218)], 2, mean),
                   Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
                   Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                                    rep("5 (2015-2019)", 6)),
                                  levels = c("0", "3 (2015-2019)", "5 (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	1.230538	1.137180	1.030834	0.8519525	0.6134826	0.3890600
3 (2015-2019)	1.232560	1.144086	1.036289	0.8538265	0.6151745	0.3904222
5 (2015-2019)	1.234353	1.144860	1.038013	0.8549890	0.6164406	0.3901203

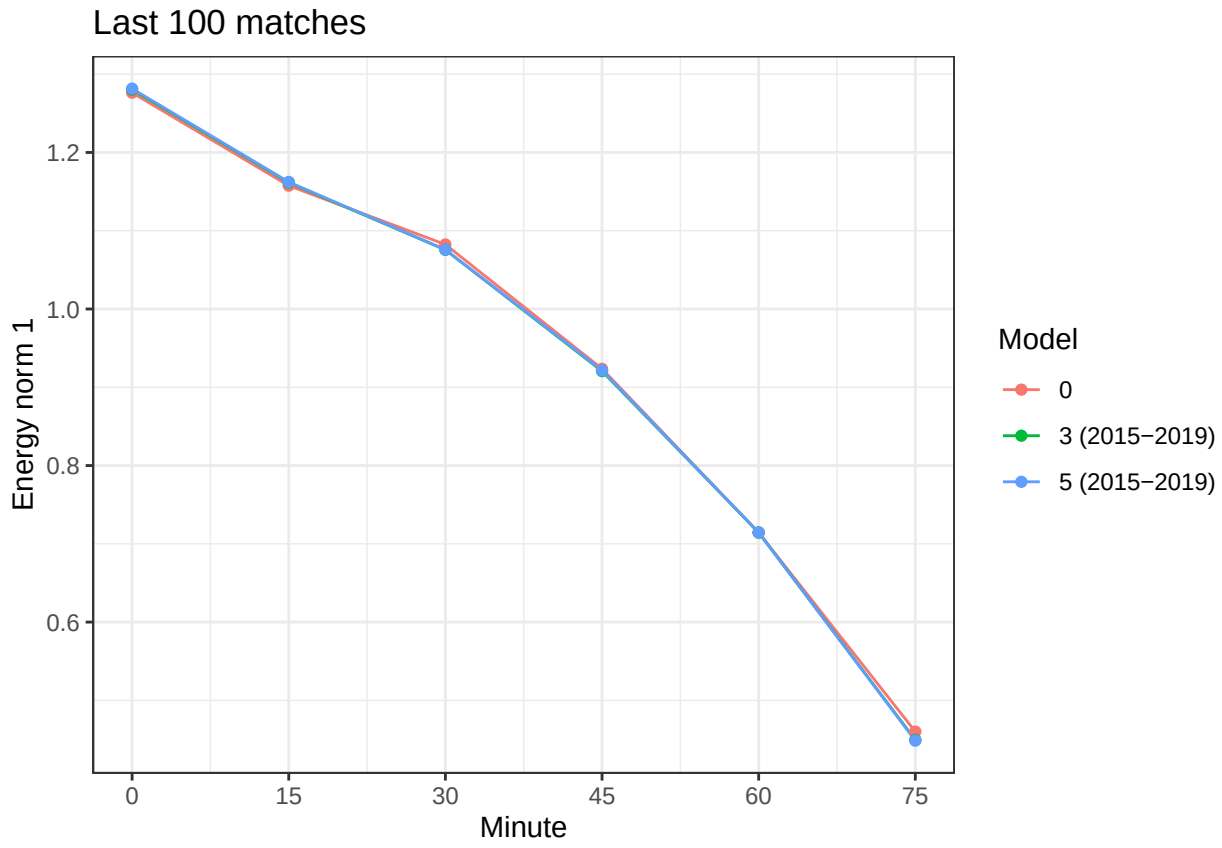
```
last_200 = tibble(ENERG1 = apply(HDA2[c(134:333), c(189:194, 207:218)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                                   rep("5 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)", "5 (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	1.197050	1.092439	0.9975346	0.8485017	0.6388620	0.4065127
3 (2015-2019)	1.201009	1.100260	0.9994170	0.8485142	0.6422552	0.4043070
5 (2015-2019)	1.202526	1.100657	0.9993805	0.8499282	0.6423479	0.4045639

```r
last_100 = tibble(ENERG1 = apply(HDA2[c(234:333), c(189:194, 207:218)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                                   rep("5 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)", "5 (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")
```




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

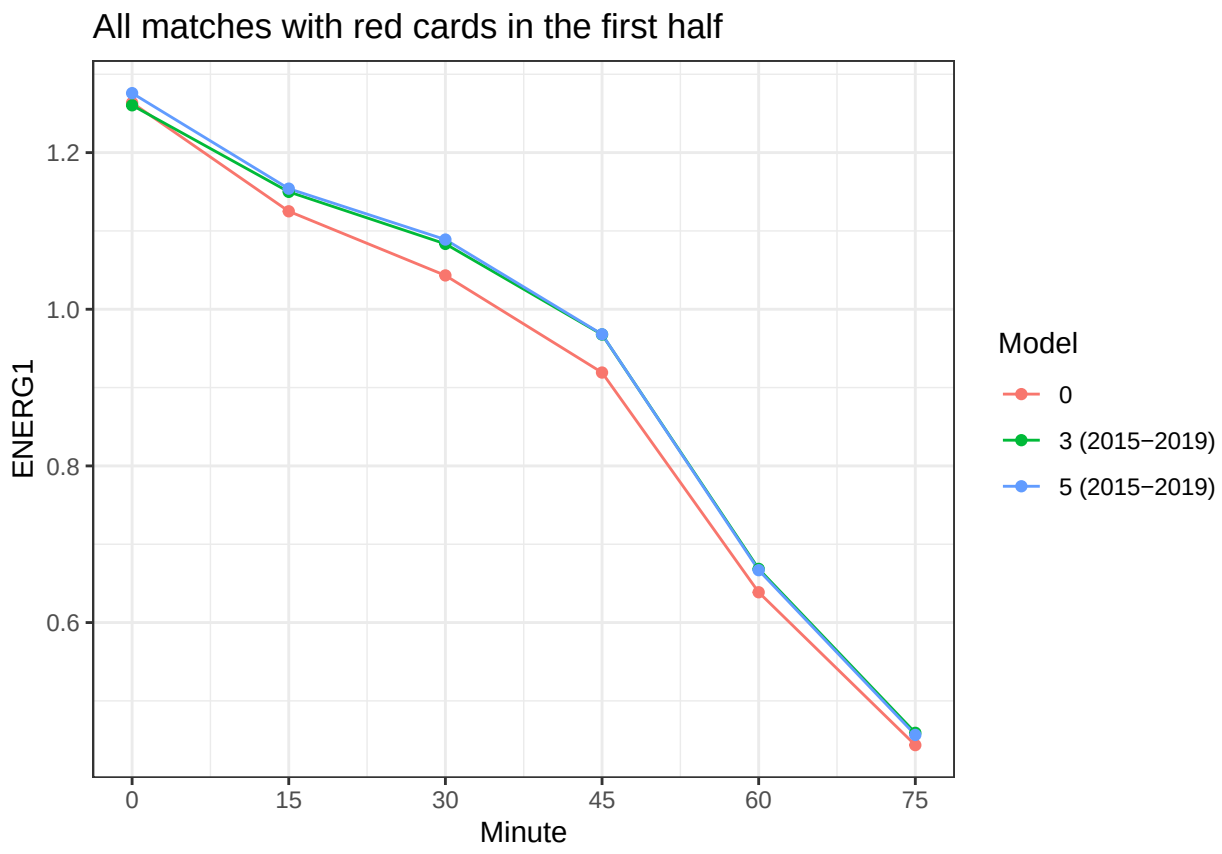
Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
0	1.276390	1.157596	1.082380	0.9235864	0.7143573	0.4600691
3 (2015-2019)	1.280337	1.161512	1.075603	0.9206749	0.7144677	0.4494584
5 (2015-2019)	1.281503	1.162295	1.075820	0.9215679	0.7144267	0.4485382

```
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(189:194, 207:218)], 2, mean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 3)),
                  Model = factor(c(rep("0", 6), rep("3 (2015-2019)", 6),
                                   rep("5 (2015-2019)", 6)),
                                 levels = c("0", "3 (2015-2019)", "5 (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	1.263613	1.125077	1.043175	0.9192029	0.6386901	0.4435420
3 (2015-2019)	1.260414	1.149901	1.083449	0.9677256	0.6685340	0.4592170
5 (2015-2019)	1.275920	1.153975	1.088968	0.9680765	0.6668733	0.4565331