mean_transition_times

June 21, 2024

[98]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

0.1 Finding transition weeks

0.1.1 What is a transition week?

Transition weeks are the starting and ending week of a phenophase, for example, the first and last weeks of a trees 'budding flowers' period. We can think of transition weeks as the first week that the percent of observations observing an attribute (say budding flowers) exceeds some small threshold, such as 10%.

0.1.2 Our approach

We assign each week w a score, $s(w)$. Then, for each year Y, we normalize the scores to get the probability w is a transition week, $p(w)$.

$$p(w) = \frac{s(w)}{\sum_{w \in Y} s(w)}$$

Let $y(k)$ be the percentage of observations observing an attribute for a species each week. To identify weeks at the beginning of a phenophase, we use the following score function:

$$s(w) = w_1 \left| \max_{w-L \leq k \leq w} y(k) - \min_{w-L \leq k \leq w} y(k) \right|^{-1} + w_2 \prod_{k=w+1}^{w+M} e^{y(k)-y(k-1)}$$

The first term we call 'stagnation'. It is large when there is little change in the M weeks before w. The second term we call 'spike'. It is large if there are large increases in percentage after week w. Weeks at the end of plateaus in the percentage plots which experience spikes directly afterwards should have high scores

0.1.3 Next Steps

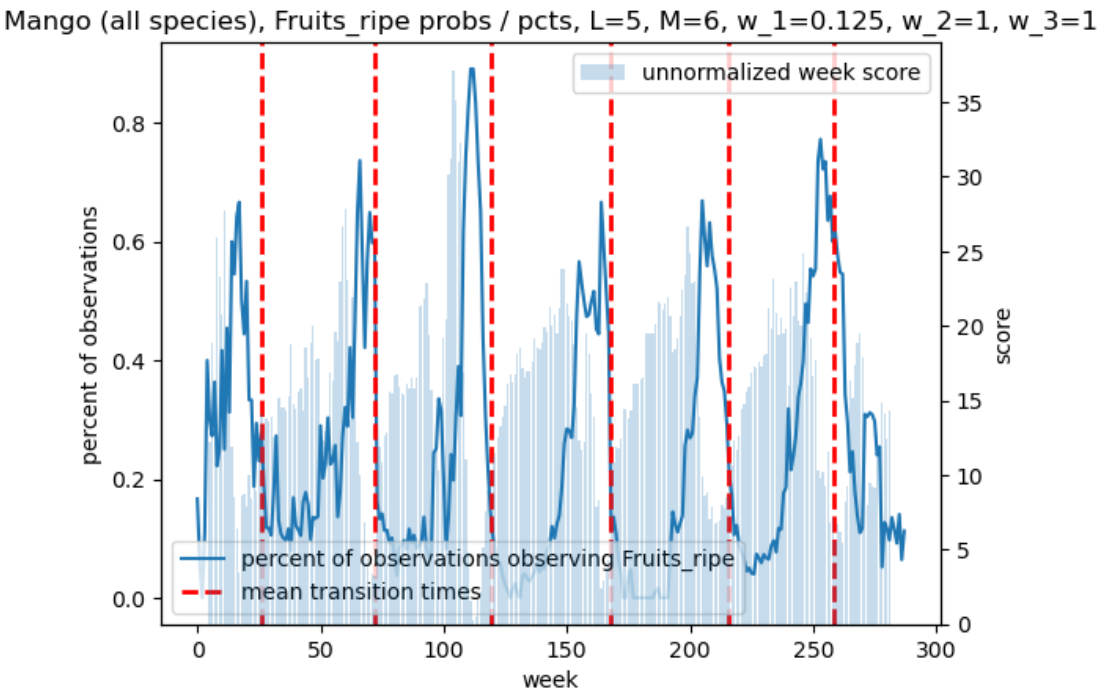
We have tried tuning the parameters of this formula on mango for the attribute Fruits_ripe, and it has proved difficult to tune. We are unsure of whether we should modify this formula or seek a different approach.

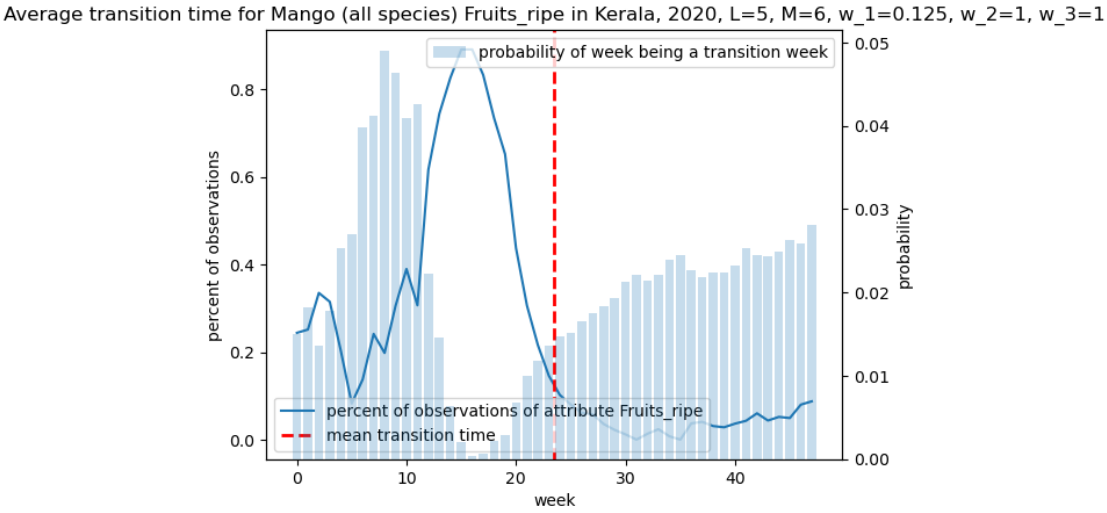
```
    #     twin.bar(range(48), spikes_for_year, label='unnormalized week spike',␣
↪alpha=0.25)
    #     twin.legend(loc='upper right')
    #     twin.set(ylabel='score')
    #     plt.title(f"Average transition time for Mango (all species) {attr} in␣
↪Kerala, {year}, L={L}, M={M}, w_1={w_1}, w_2={w_2}")
```

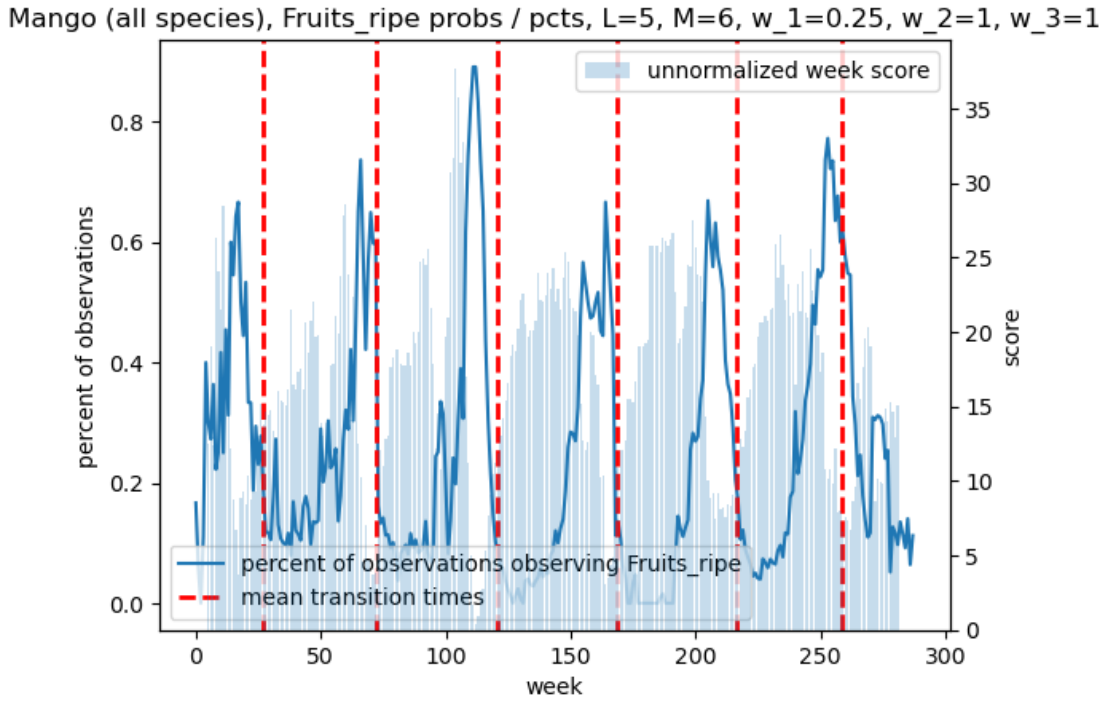
0.5 Initial Experiments

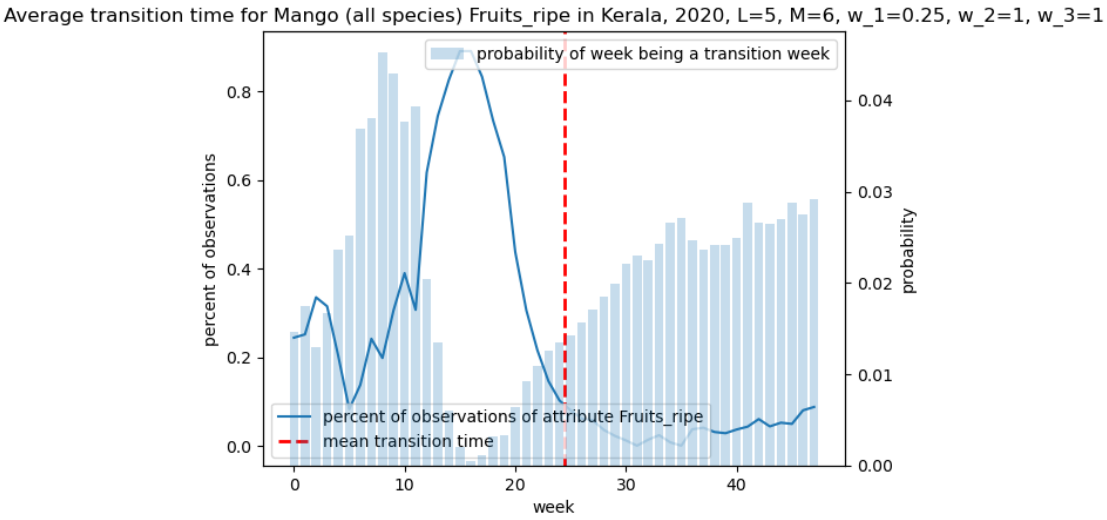
- Experiment with different values of the stagnation weight
- Plots of score function with weeks from all years aggregated into the same plot
- Plots of probability function for the year 2020

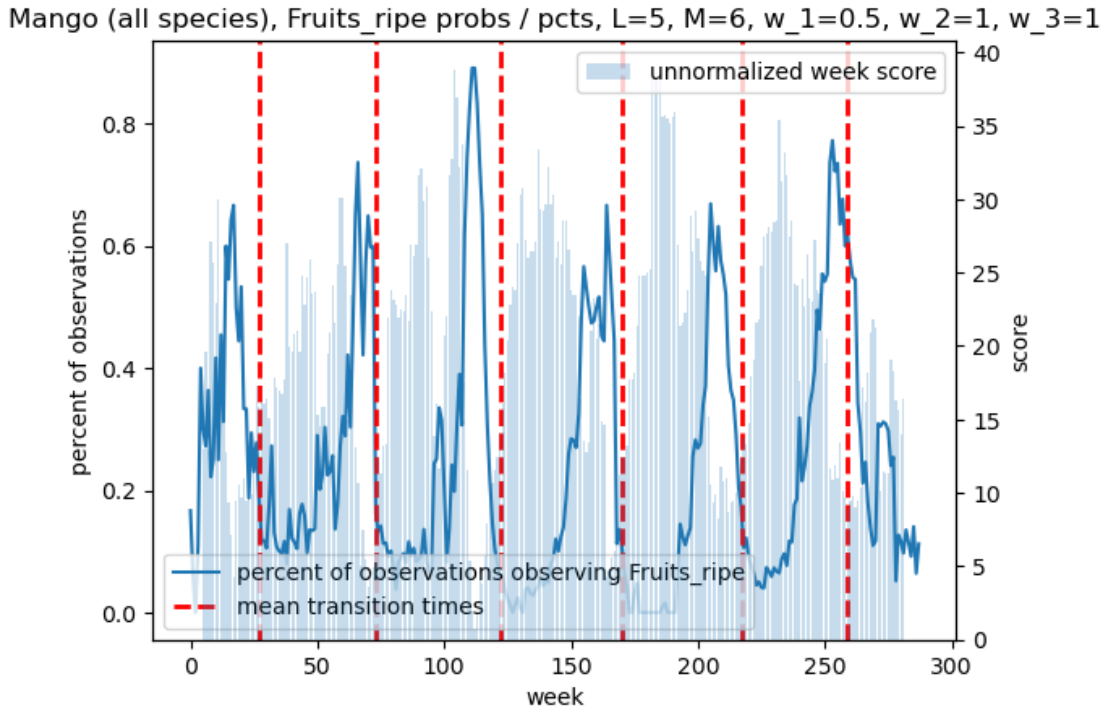
[496]:
```
year = 2022
species_id = 1090
for w_1 in [0.125, 0.25, 0.5, 1, 2, 4, 8]:
    for w_2 in [1]:
        for M in [6]:
            for L in [5]:
                for attr in ['Fruits_ripe']:
                    # plot_stagnations(species_id, attr, L=L, M=M, w_1=w_1,␣
↪w_2=w_2)
                    # plot_spikes(species_id, attr, L=L, M=M, w_1=w_1, w_2=w_2)
                    plot_score_and_pcts(species_id, attr, L=L, M=M, w_1=w_1,␣
↪w_2=w_2, size_term=False)
                #plot_score_for_one_year(year, 1161, attr, L=5, M=M, w_1=w_1,␣
↪w_2=w_2)
```

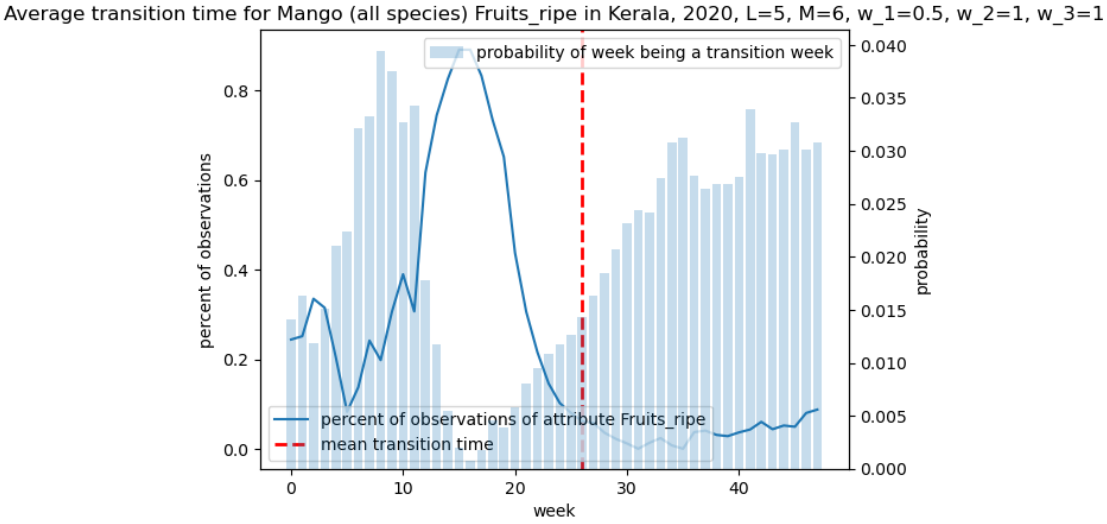


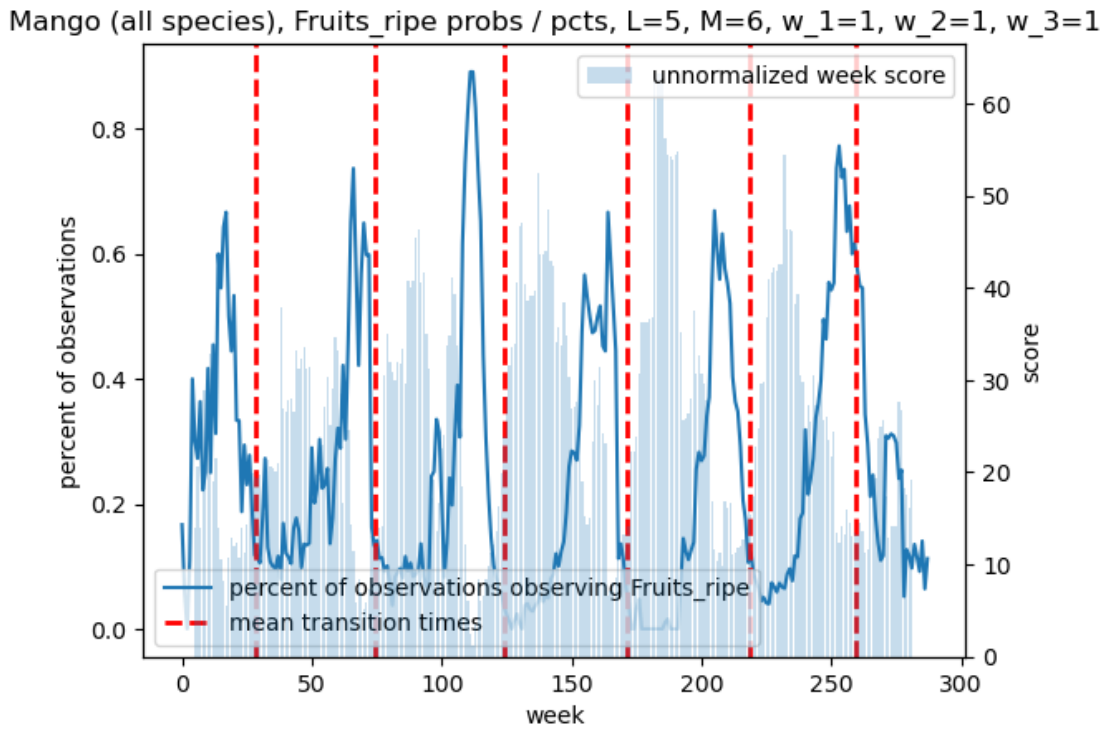

Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=0.125, w_2=1, w_3=1
probability of week being a transition week
percent of observations
probability
week
percent of observations of attribute Fruits_ripe
mean transition time

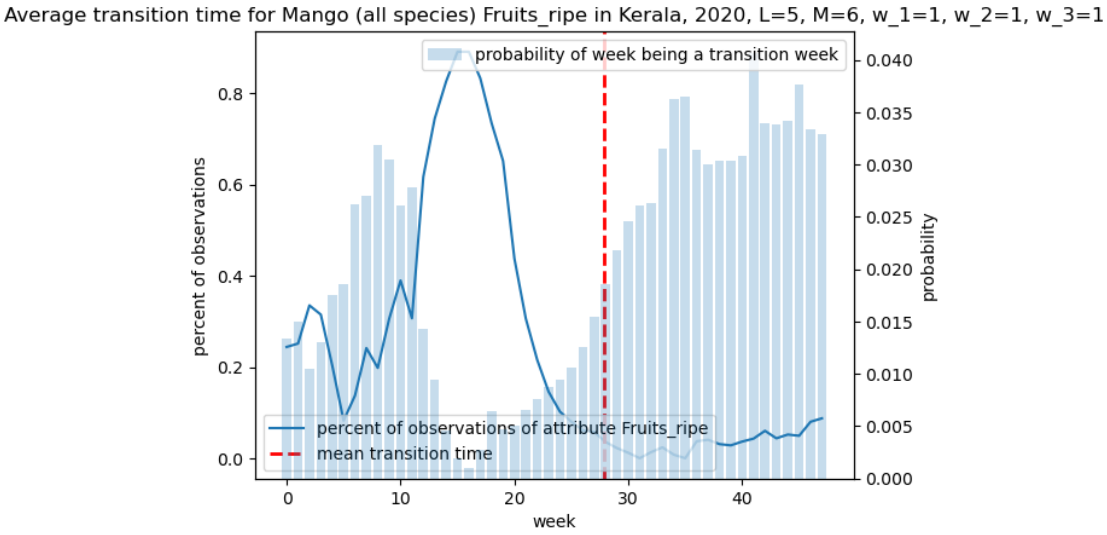

Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=0.25, w_2=1, w_3=1
unnormalized week score
percent of observations
score
week
percent of observations observing Fruits_ripe
mean transition times

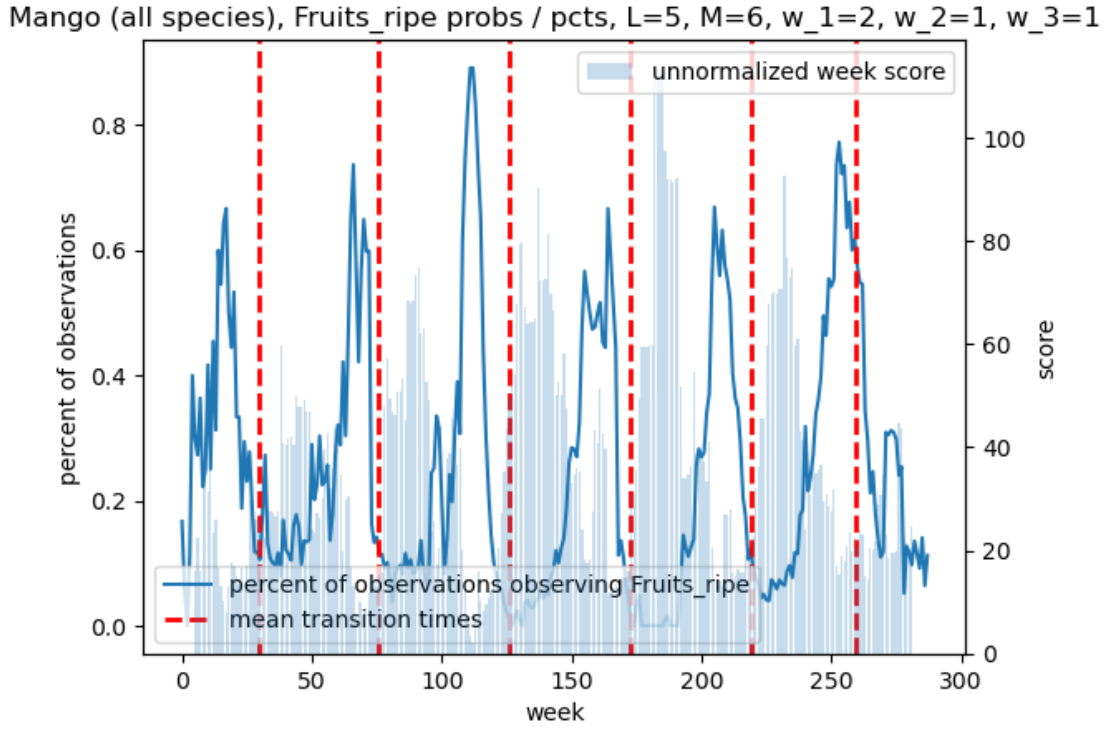

Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=0.25, w_2=1, w_3=1
probability of week being a transition week
percent of observations
probability
week
percent of observations of attribute Fruits_ripe
mean transition time

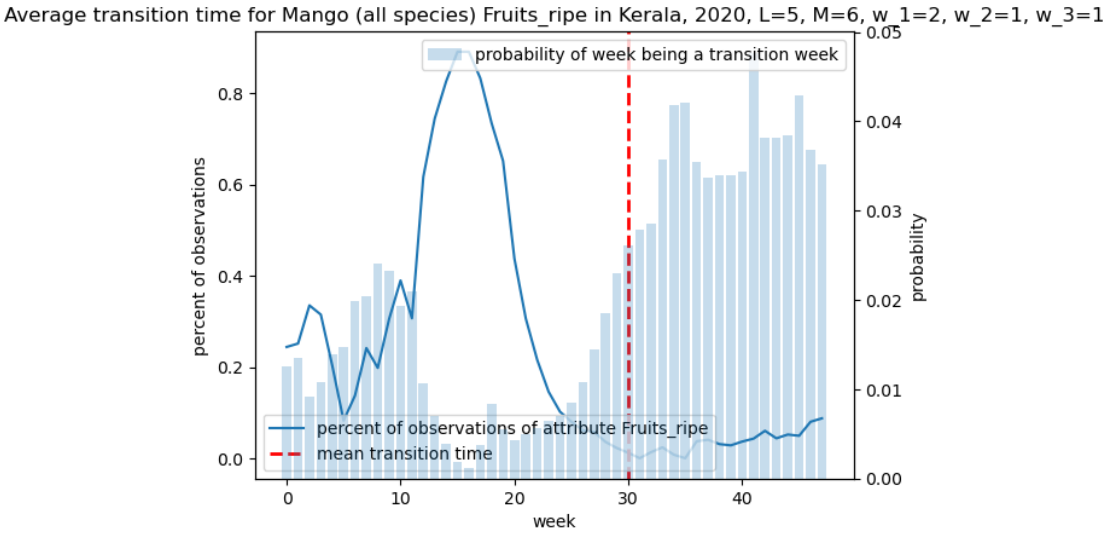

Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=0.5, w_2=1, w_3=1
unnormalized week score
percent of observations
score
week
percent of observations observing Fruits_ripe
mean transition times

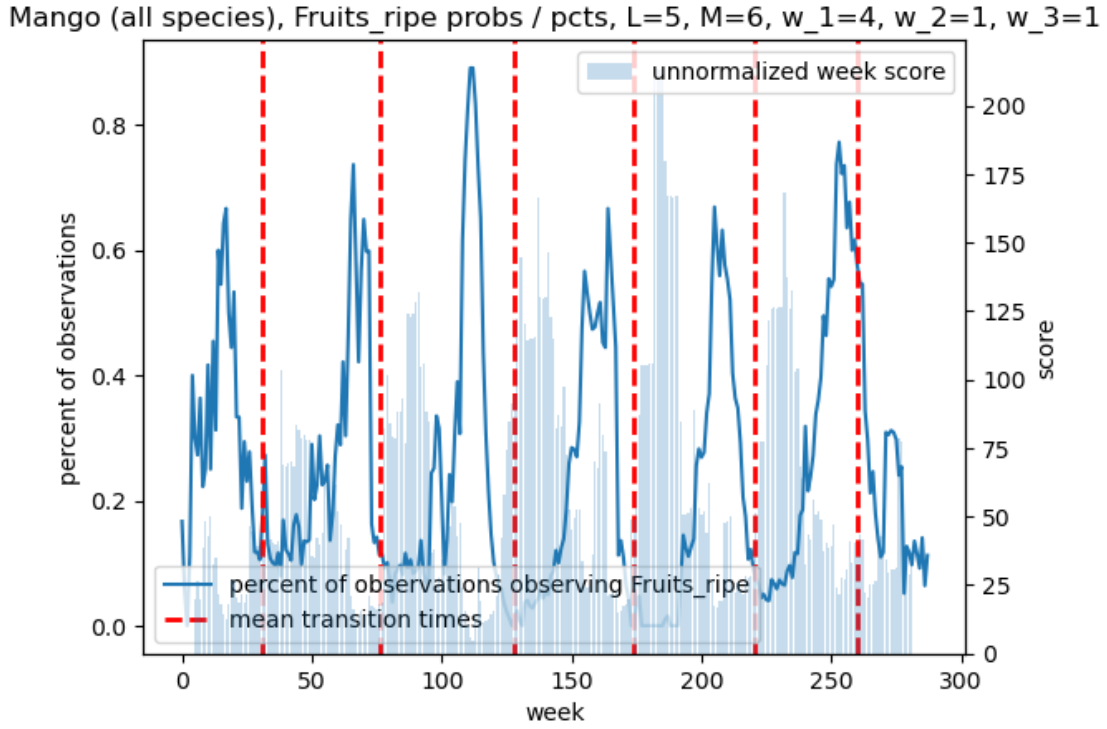

Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=0.5, w_2=1, w_3=1
probability of week being a transition week
percent of observations
probability
week
percent of observations of attribute Fruits_ripe
mean transition time

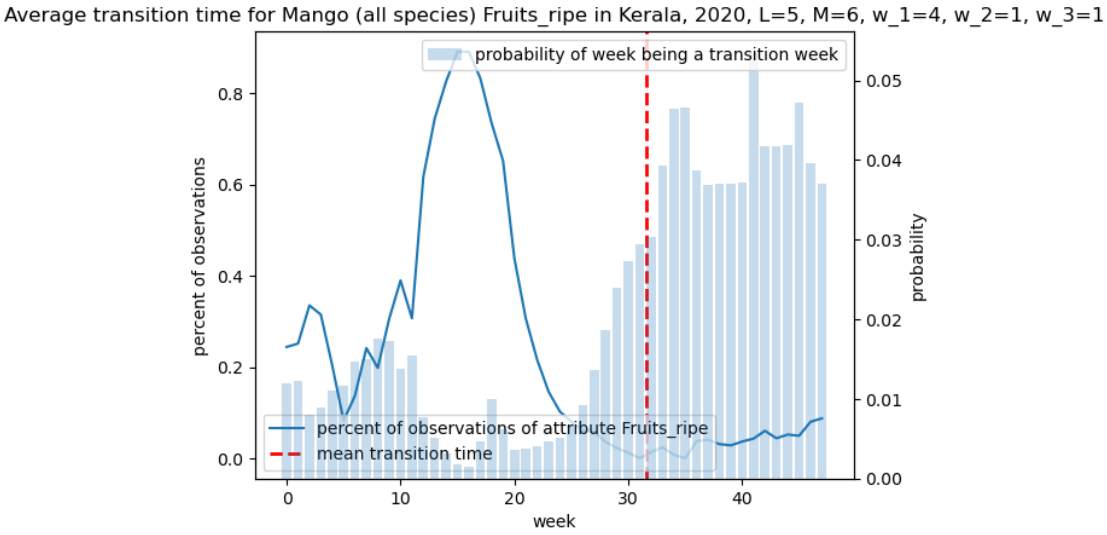

Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=1, w_2=1, w_3=1
unnormalized week score
percent of observations
score
week
percent of observations observing Fruits_ripe
mean transition times

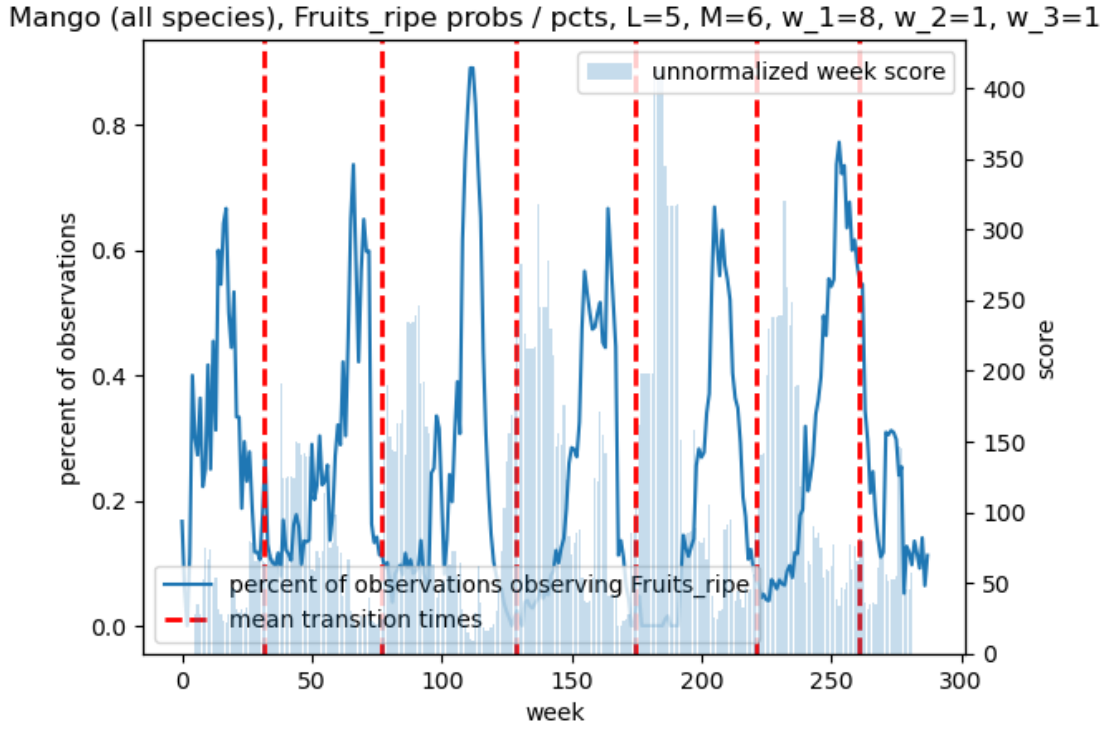

Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=1, w_2=1, w_3=1
probability of week being a transition week
percent of observations
probability
week
percent of observations of attribute Fruits_ripe
mean transition time

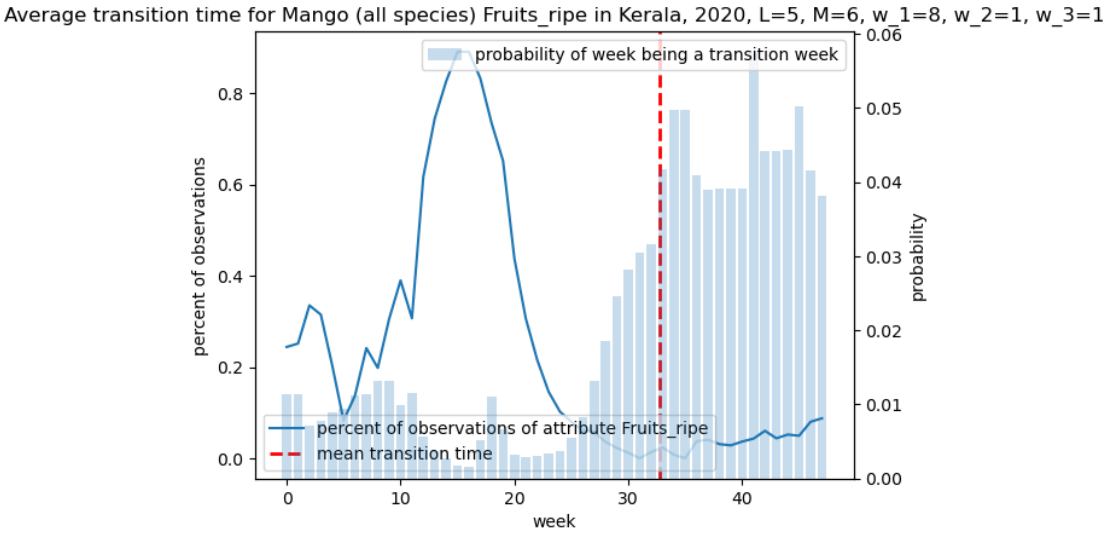

Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=2, w_2=1, w_3=1
unnormalized week score
percent of observations
score
week
percent of observations observing Fruits_ripe
mean transition times


Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=2, w_2=1, w_3=1
probability of week being a transition week
percent of observations of attribute Fruits_ripe
mean transition time
percent of observations
probability
week


Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=4, w_2=1, w_3=1
unnormalized week score
percent of observations observing Fruits_ripe
mean transition times
percent of observations
score
week


Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=4, w_2=1, w_3=1
probability of week being a transition week
percent of observations
probability
week
percent of observations of attribute Fruits_ripe
mean transition time


Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=8, w_2=1, w_3=1
unnormalized week score
percent of observations
score
week
percent of observations observing Fruits_ripe
mean transition times



0.6 Experimenting with size term

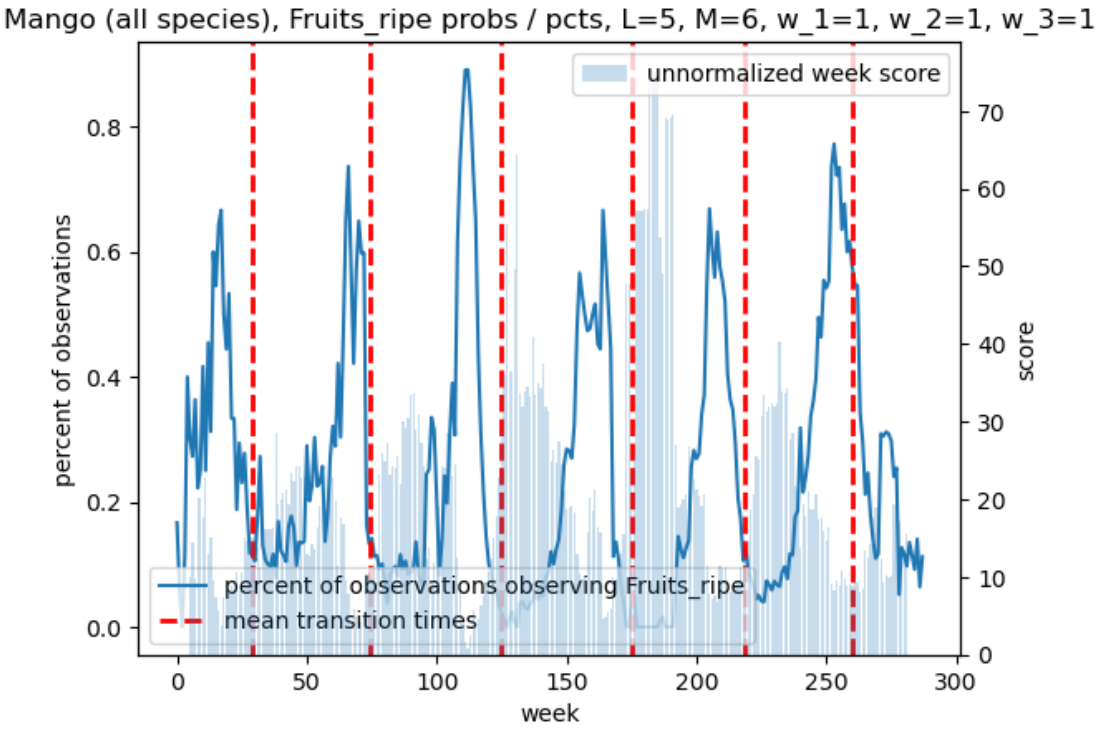
Add a third term to the expression for $s(w)$ which assigns higher scores to weeks with lower percentages.

$$s(w) = w_1 \left| \max_{w-L \le k \le w} y(k) - \min_{w-L \le k \le w} y(k) \right|^{-1} + w_2 \prod_{k=w+1}^{w+M} e^{y(k)-y(k-1)} + w_3 \frac{1}{y(w)}$$

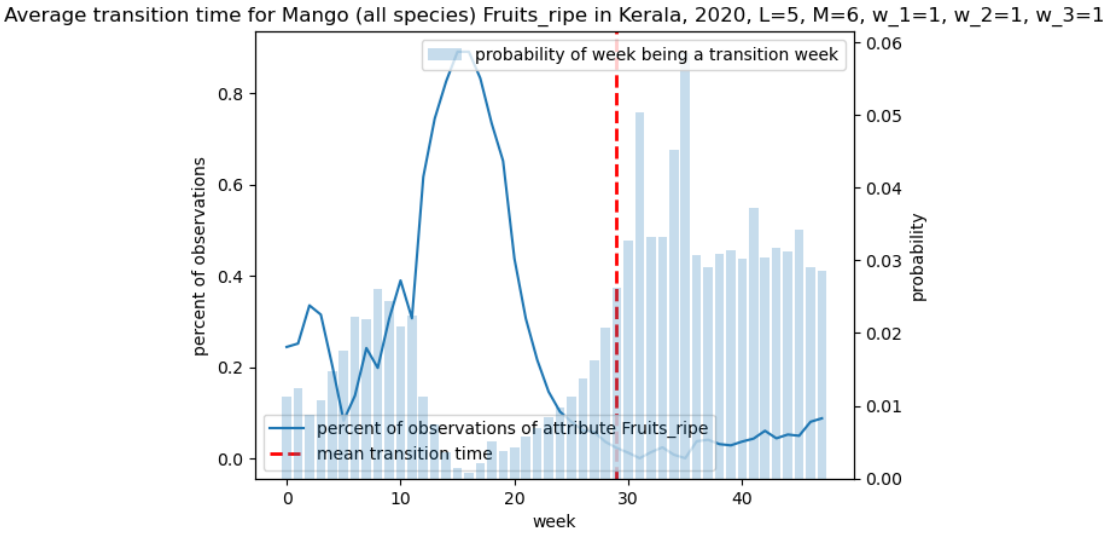
We experiment with different values of the size weight, w_3, and fix $w_1 = w_2 = 1$.

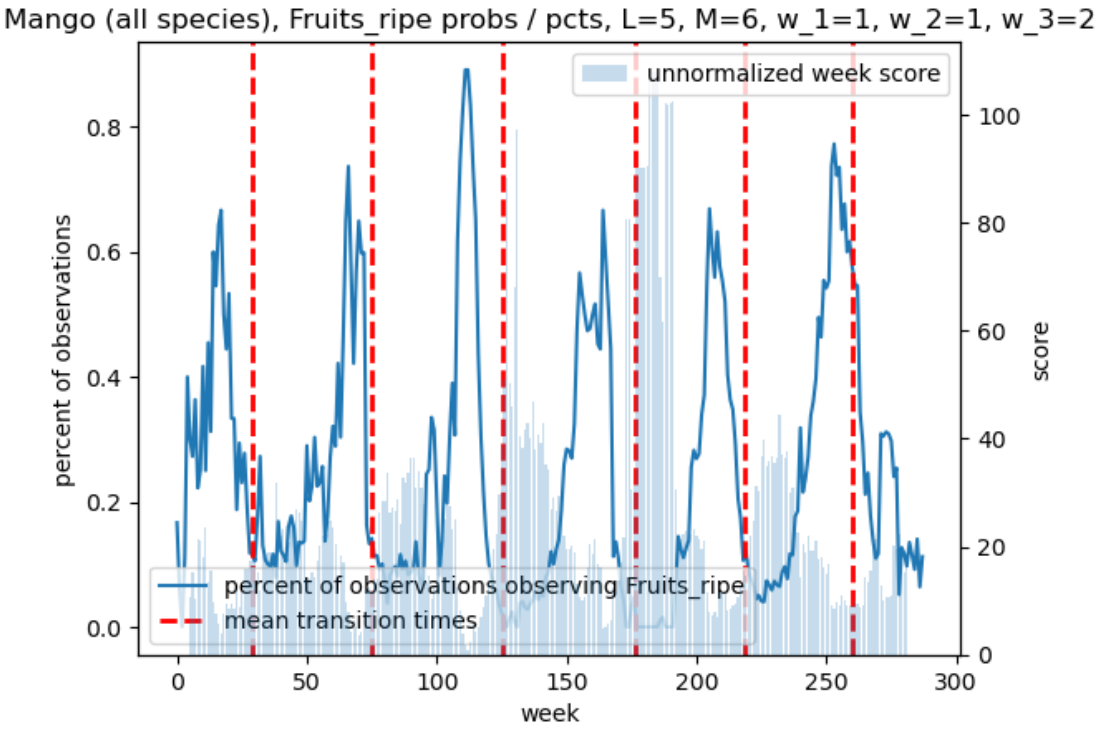
[492]:
```
year = 2022
species_id = 1090
for w_1 in [1]:
    for w_2 in [1]:
        for w_3 in [1, 2, 4, 8]:
            for M in [6]:
                for L in [5]:
                    for attr in ['Fruits_ripe']:
                        # plot_stagnations(species_id, attr, L=L, M=M, w_1=w_1,␣
↪w_2=w_2)
                        # plot_spikes(species_id, attr, L=L, M=M, w_1=w_1,␣
↪w_2=w_2)
                        plot_score_and_pcts(species_id, attr, L=L, M=M,␣
↪w_1=w_1, w_2=w_2, w_3=w_3, size_term=True)
                #plot_score_for_one_year(year, 1161, attr, L=5, M=M, w_1=w_1,␣
↪w_2=w_2)
```

```
len scores 288, len pcts 288
```

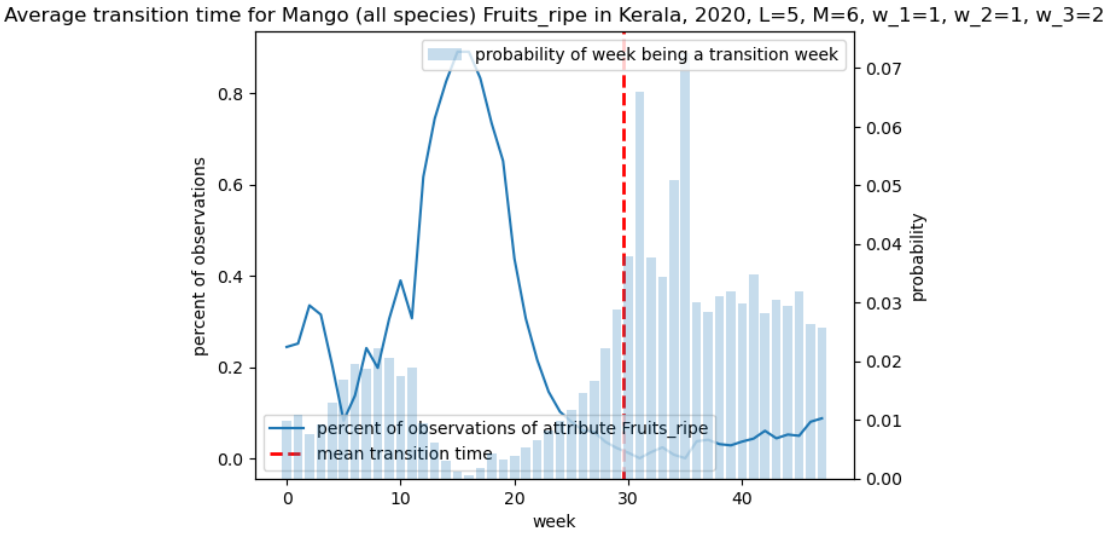


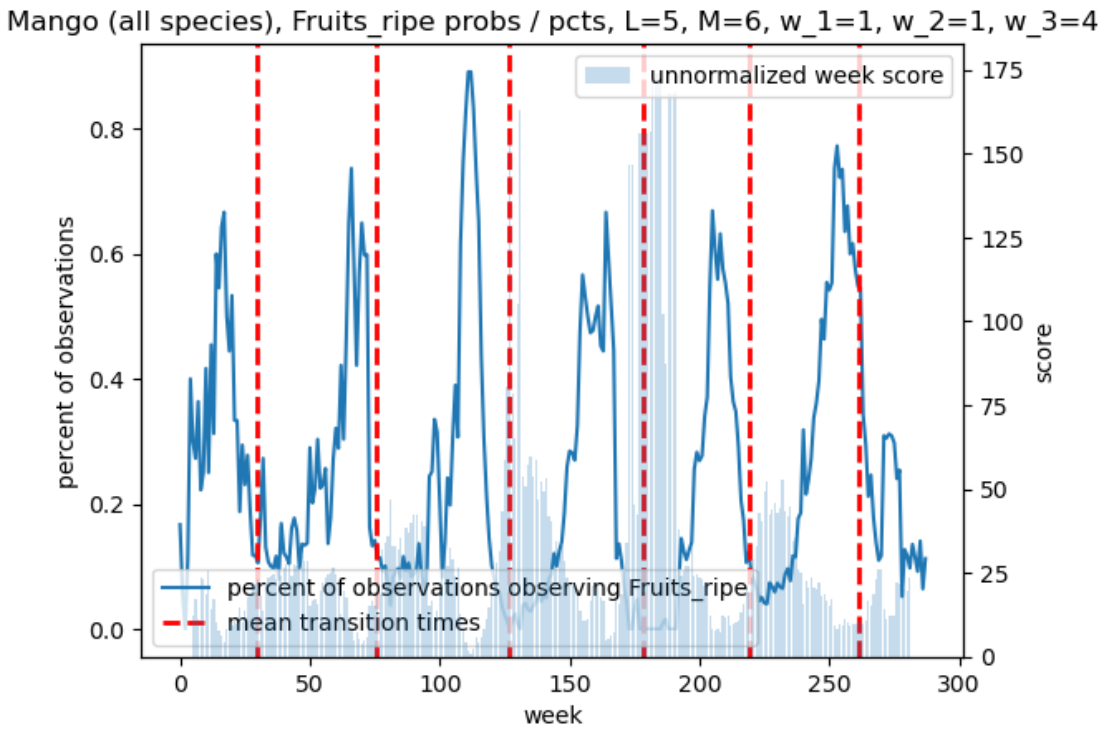

```
len scores 288, len pcts 288
```



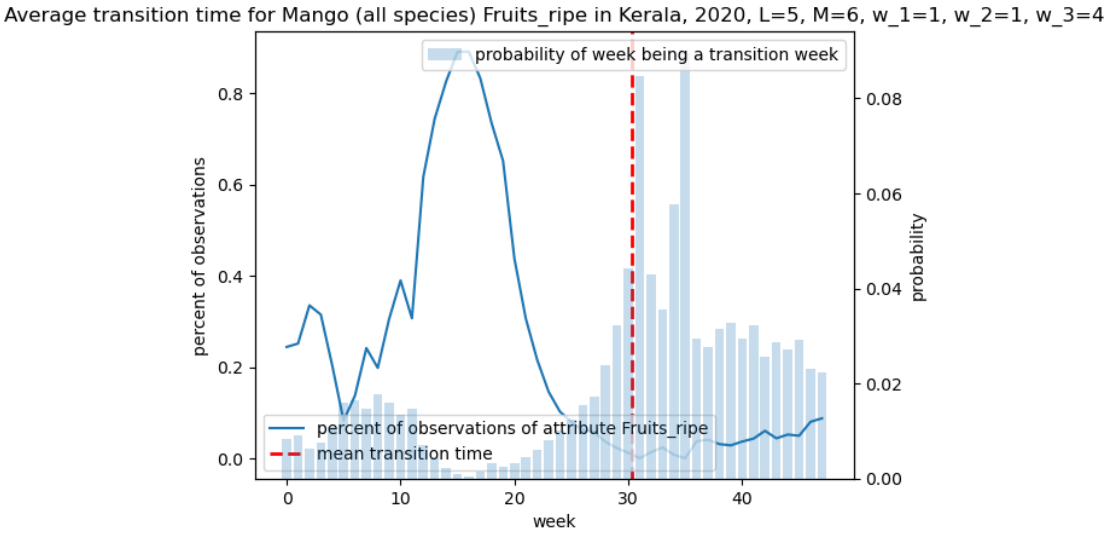


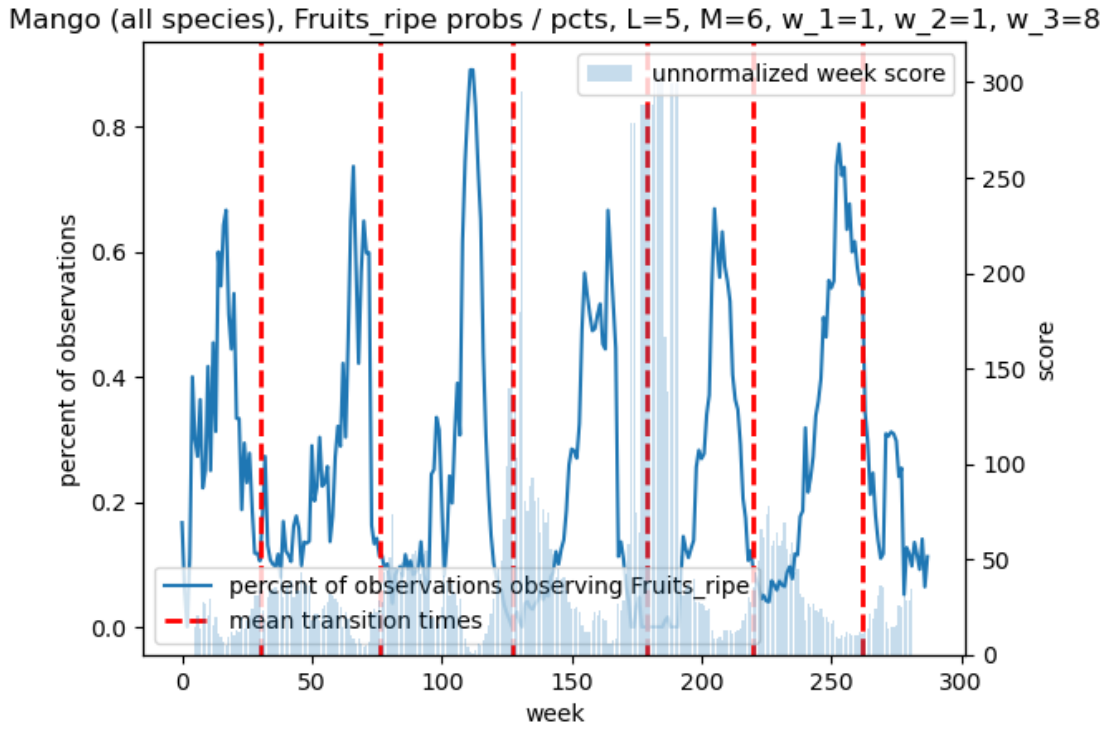

```
len scores 288, len pcts 288
```

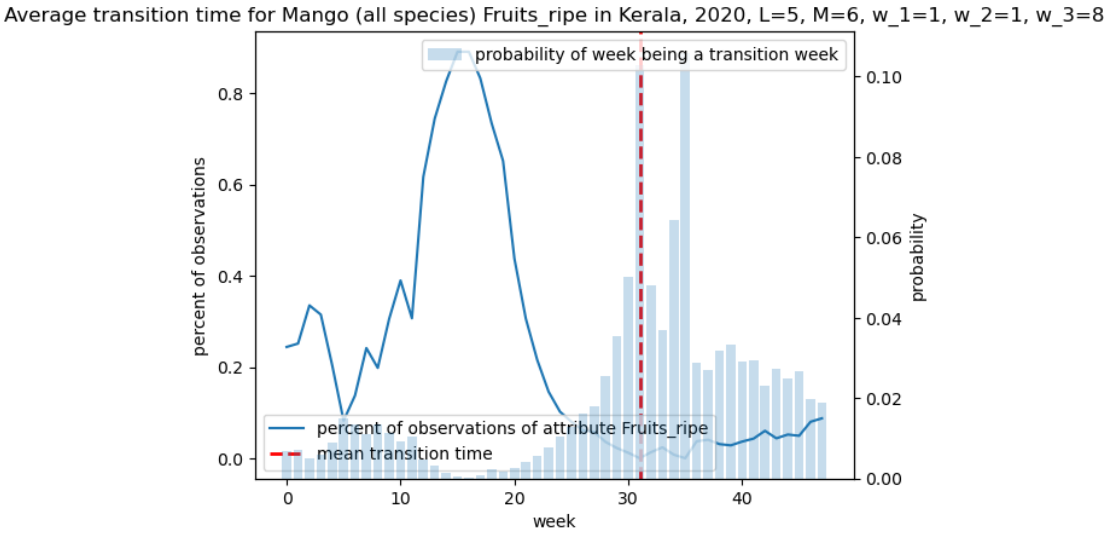





```
len scores 288, len pcts 288
```




Mango (all species), Fruits_ripe probs / pcts, L=5, M=6, w_1=1, w_2=1, w_3=8
unnormalized week score
percent of observations observing Fruits_ripe
mean transition times
percent of observations
score
week
0.0
0.2
0.4
0.6
0.8
0
50
100
150
200
250
300


Average transition time for Mango (all species) Fruits_ripe in Kerala, 2020, L=5, M=6, w_1=1, w_2=1, w_3=8
probability of week being a transition week
percent of observations of attribute Fruits_ripe
mean transition time
percent of observations
probability
week
0.0
0.2
0.4
0.6
0.8
0.00
0.02
0.04
0.06
0.08
0.10
0
10
20
30
40