

HMMs Tuco

Número de Capturas

Ao todo foram capturados 22 tucos. Desse número nem todos receberam acelerômetro e lightlogger. As tabelas abaixo mostram quantos animais foram capturados por campanha e quantos animais receberam acelerômetro ou lightloggers.

```
## # A tibble: 5 x 4
##   Month      n Accelerometer Lightlogger
##   <chr>    <dbl>         <dbl>         <dbl>
## 1 February     5             5             5
## 2 July         8             8             6
## 3 March        2             2             0
## 4 October      7             7             2
## 5 TOTAL       22            22            13
```

Hidden Markov Models

O HMM foi ajustado aos dados dos tucos usando o método de *no pooling*. Assumiu-se um número de 3 estados a priori de acordo com o conhecimento que temos dos tucos, da distribuição dos dados e do nosso objetivo de tentar separar os dados em mais categorias além de atividade/repouso. Dessa maneira assumimos que os animais compartilham os mesmo parâmetros para as 3 *state-dependent distribution*, que são as distribuições que fazem parte da mistura. A distribuição escolhida para o s ajustes foi a gamma.

Construi 2 modelos diferentes de acordo com o que foi visto na análise exploratória. Essas modelos são: (i) um modelo nulo/vazio, que é um modelo sem nenhuma covariáveis explicatórias e (ii) um modelo com a covariável 'season'. Nesse segundo modelo os parâmetros dos *state-dependent distributions* é mantido constante e o que é livre são os probabilidades de transição. Isso quer dizer que estamos assumindo que os comportamentos não mudam entre os tucos mas que as probabilidades de transição entre um comportamento e outro podem mudar de acordo com a estação/mês do ano.

Essas dois modelos foram comparados usando o critério de Akaike e o modelo com 'estação como 'covariável' se sai melhor em explicar os dados.

```
m1 = readRDS(paste0(dir, "/03_analysis/hmm/m1.rds")) # modelo vazio
m2 = readRDS(paste0(dir, "/03_analysis/hmm/m2.rds")) # modelo com estação

AIC(m1, m2)
```

```
##   Model      AIC
## 1    m2 -935716.3
## 2    m1 -934905.1
```

Especificações do Modelo:

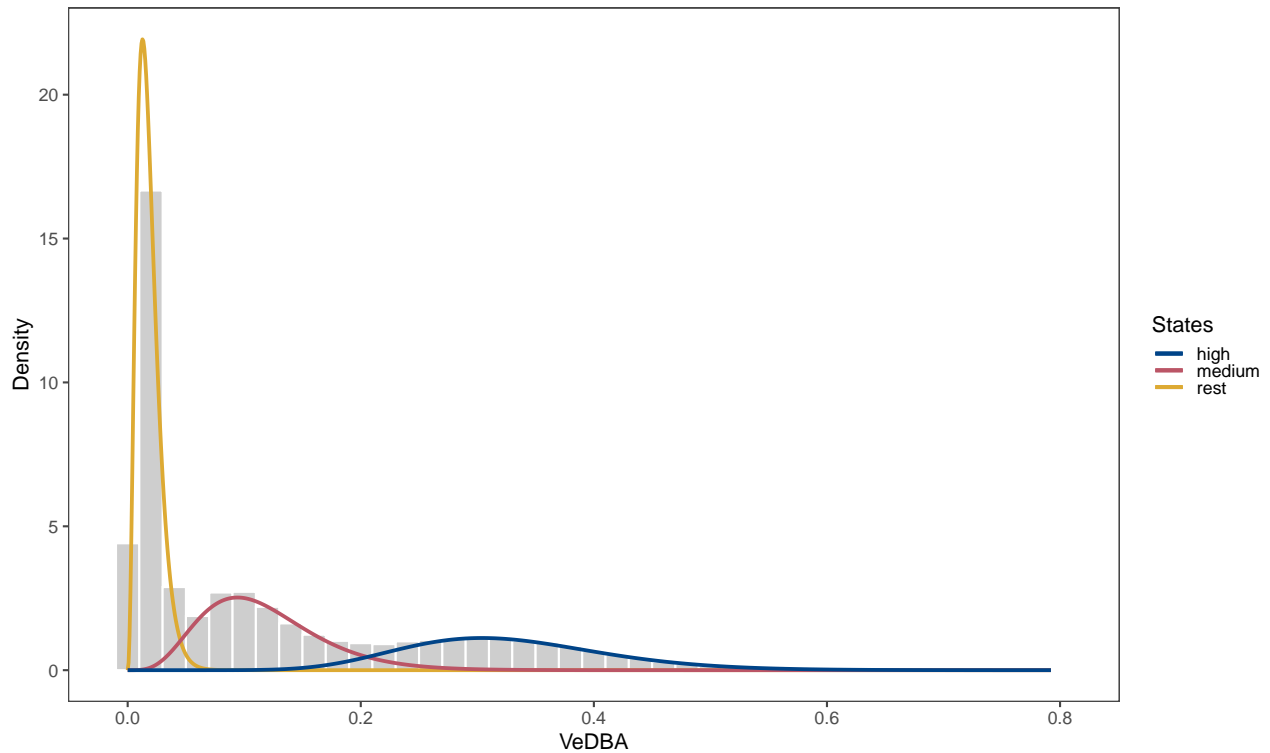
```

## Value of the maximum log-likelihood: 467893.1
##
##
## vedba parameters:
## -----
##           rest      medium      high
## mean      0.0181305210 1.163597e-01 3.268437e-01
## sd         0.0098282133 5.079873e-02 8.651540e-02
## zeromass   0.0008155059 9.880822e-06 3.159508e-13
##
## Regression coeffs for the transition probabilities:
## -----
##           1 -> 2      1 -> 3      2 -> 1      2 -> 3      3 -> 1
## (Intercept) -3.17798818 -5.23124564 -2.7725838 -2.31659494 -4.2745506
## seasonJuly   -0.27634827 -0.84741669 -0.2314373 -0.20701772 -0.2802444
## seasonOctober 0.05792022 0.69249453 0.2929987 0.40836047 0.5907193
## seasonFebruary -0.09521610 -0.03606858 -0.1198916 0.06814938 0.4363291
##           3 -> 2
## (Intercept) -2.00735123
## seasonJuly   0.24929499
## seasonOctober -0.03880555
## seasonFebruary -0.15827581
##
## Transition probability matrix (based on mean covariate values):
## -----
##           rest      medium      high
## rest      0.95509500 0.03979824 0.00510676
## medium    0.05382811 0.86124552 0.08492637
## high      0.01212119 0.11699770 0.87088112
##
## Initial distribution:
## -----
##           rest      medium      high
## 6.026298e-01 3.973259e-01 4.432879e-05

```

Distribuições do HMM

O gráfico abaixo mostra o histograma da distribuição dos valores de VeDBA sobreposto com as distribuições dos três estados identificados pelo HMM de acordo com os dados. Esse gráfico mostra que cada estado tem um range de valores distribuídos na forma das curvas especificadas pelos modelos.



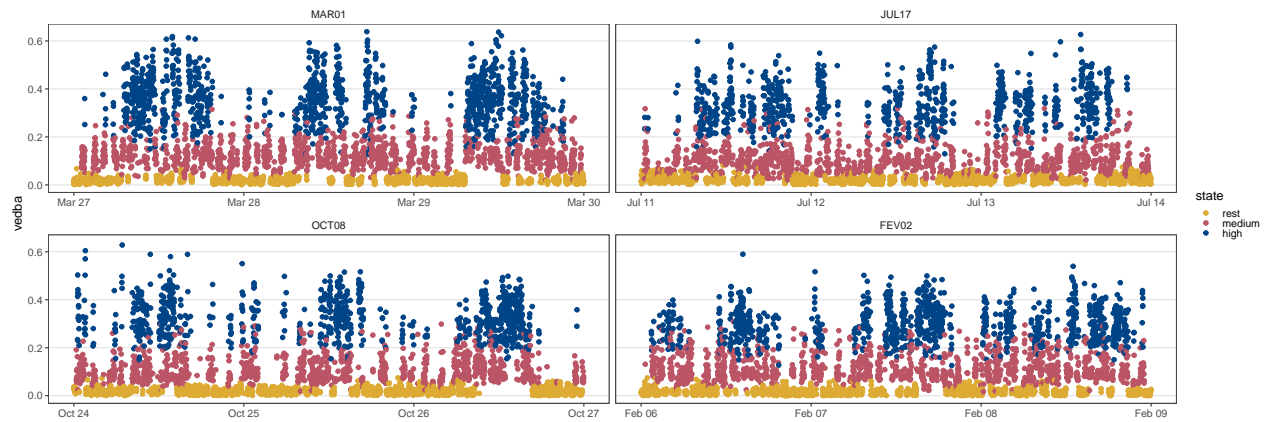
Classificação dos Dados

Com o modelo já feito agora podemos usá-lo para classificar os dados coletados nos três estados definido. Para isso é usado um algoritmo chamado de viterbi. Esse algoritmo classifica cada ponto de acordo tendo como referência o conjunto de dados de cada animal. A figura ilustra como é a classificação de cada ponto dentro da série temporal de quatro animais.

Com os dados classificados agora podemos dividir as análises em duas partes:

- Atividade Geral:
 - Timing da Atividade:
 - * Que horários cada estado acontece? Existe um padrão diário?
 - * Esse padrão temporal muda ao longo do ano?
 - * Qual o período dessa atividade?
 - * *Nesse momento estamos analisando só o timing de cada estado*
 - Intensidade da Atividade:
 - * Existe diferença na proporção de tempo em que o animal passa em cada estados?
 - * Essa proporção muda ao longo do ano?
 - * *Com esse segundo passo saberemos também se a 'intensidade' é a mesma. P.e. se algum animal apresenta mais atividade intensa ou se essa proporção muda ao longo do ano.*

- Atividade de Superfície
 - Timing da Atividade de superfície:
 - * Como as saídas à superfície estão organizadas no tempo? (O pico é maior no meio do dia no Inverno?)
 - Intensidade da atividade de superfície:
 - * Quais estados são mais frequentes na superfície?



Atividade Geral

- Atividade Geral:
 - Timing da Atividade:
 - * Que horários cada estado acontece? Existe um padrão diário?
 - * Esse padrão temporal muda ao longo do ano?
 - * Qual o período dessa atividade?
 - * *Nesse momento estamos analisando só o timing de cada estado*
 - Intensidade da Atividade:
 - * Existe diferença na proporção de tempo em que o animal passa em cada estados?
 - * Essa proporção muda ao longo do ano?
 - * *Com esse segundo passo saberemos também se a ‘intensidade’ é a mesma. P.e. se algum animal apresenta mais atividade intensa ou se essa proporção muda ao longo do ano.*

Padrão Diário

Gráficos dos padrões temporais diários das atividades de média e alta intensidade.

Alta Intensidade

A alta intensidade parece sim ter um padrão diário. Há uma concentração maior de episódios de alta intensidade durante o dia. Porém, ainda há muitos episódio de alta intensidade durante a noite. As curvas desses gráfico mostram o padrão temporal e são normalizadas entre os animais.

```
## Warning: Removed 42 rows containing missing values (geom_bar).
```

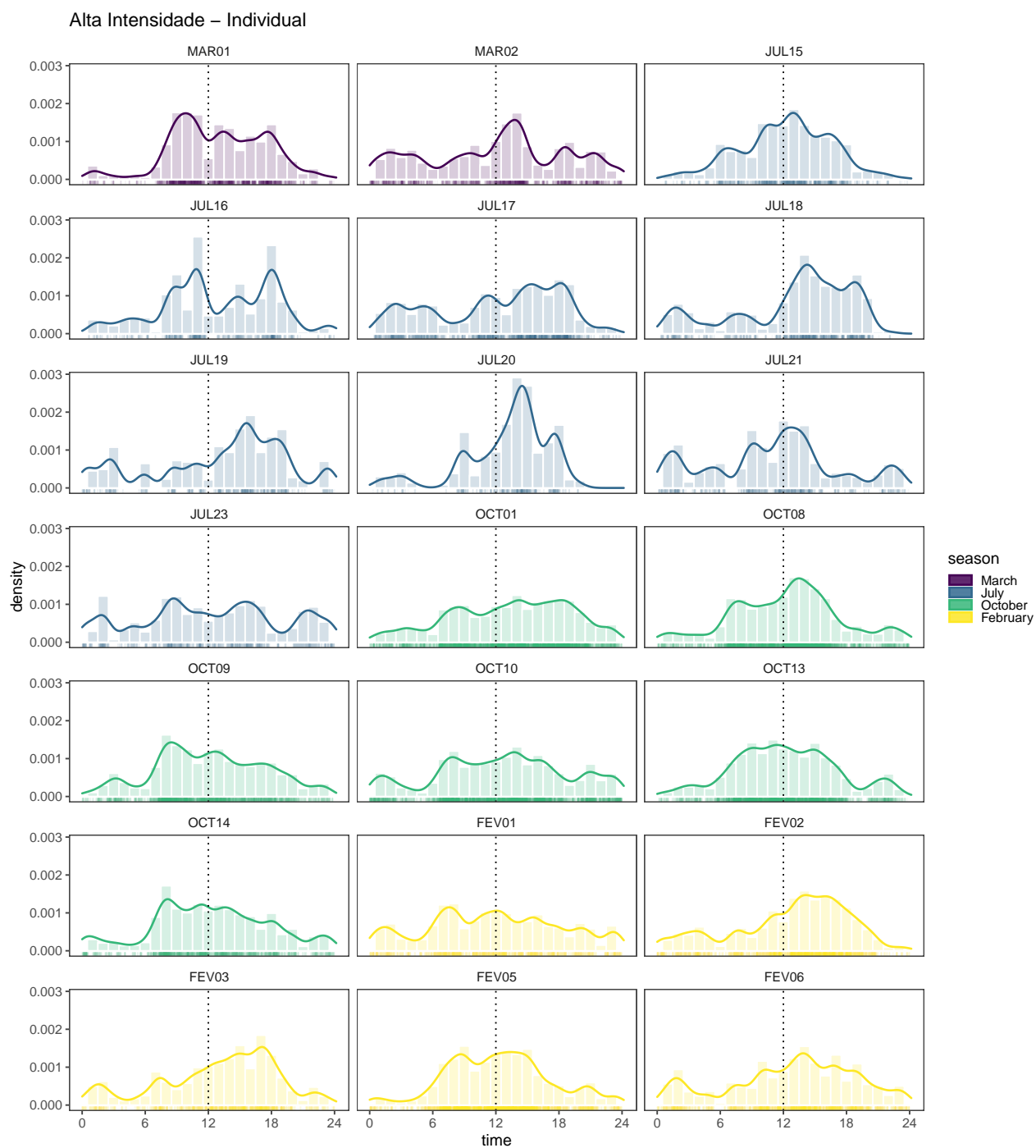


Figure 1: Curvas de densidade que mostram o padrão diário das atividades de alta intensidade por indivíduo.

Warning: Removed 8 rows containing missing values (geom_bar).

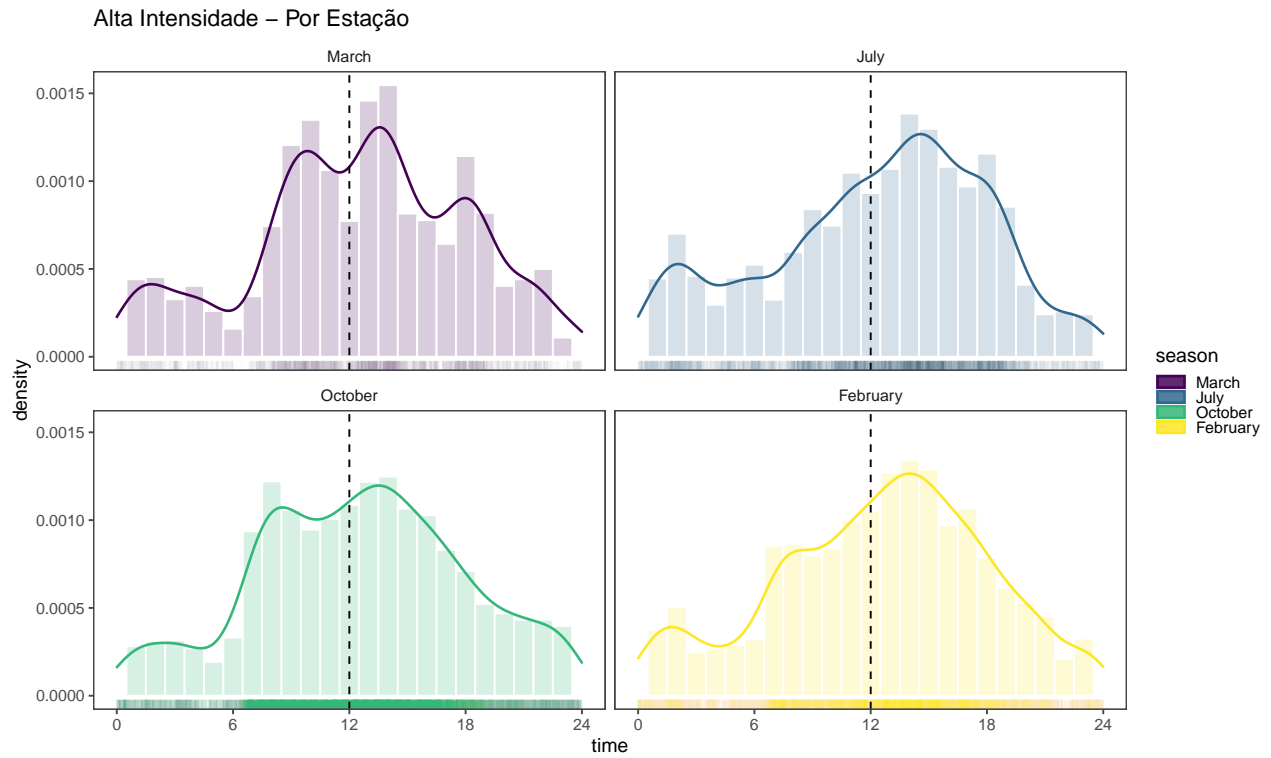


Figure 2: Curvas de densidade que mostram o padrão diário das atividades de alta intensidade por estação.

Média Intensidade

As curvas do padrão temporal para a atividade de média intensidade são espalhadas ao longo do dia e não possuem um padrão diário (de 24 horas).

```
## Warning: Removed 42 rows containing missing values (geom_bar).
```




Figure 3: Curvas de densidade para o padrão temporal das atividades de média intensidade por indivíduo.

Warning: Removed 8 rows containing missing values (geom_bar).

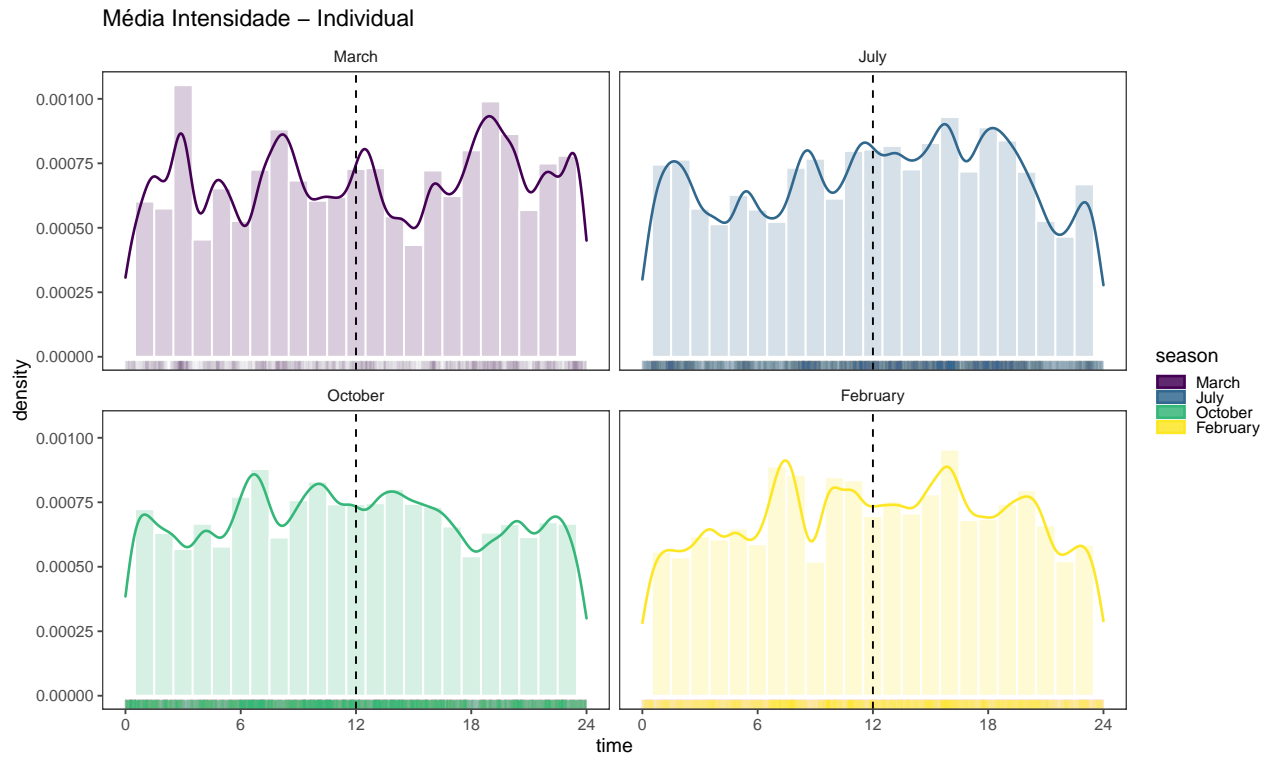


Figure 4: Curvas de densidade para o padrão temporal das atividades de média intensidade agrupadas por estação.

Tempo em Cada Estado

Com os dados classificados agora podemos calcular quanto tempo ou proporção do total de registros que cada animal passou nos diferentes estados. Vimos os episódios de alta intensidade estão mais concentrados durante o dia. Porém, a quantidade de episódios de alta intensidade varia ao longo do ano. Aparentemente os animais se movem mais em Outubro do que no inverno, por exemplo. Ou seja os animais passam mais tempo no estado de alta intensidade em outubro e menos no inverno.

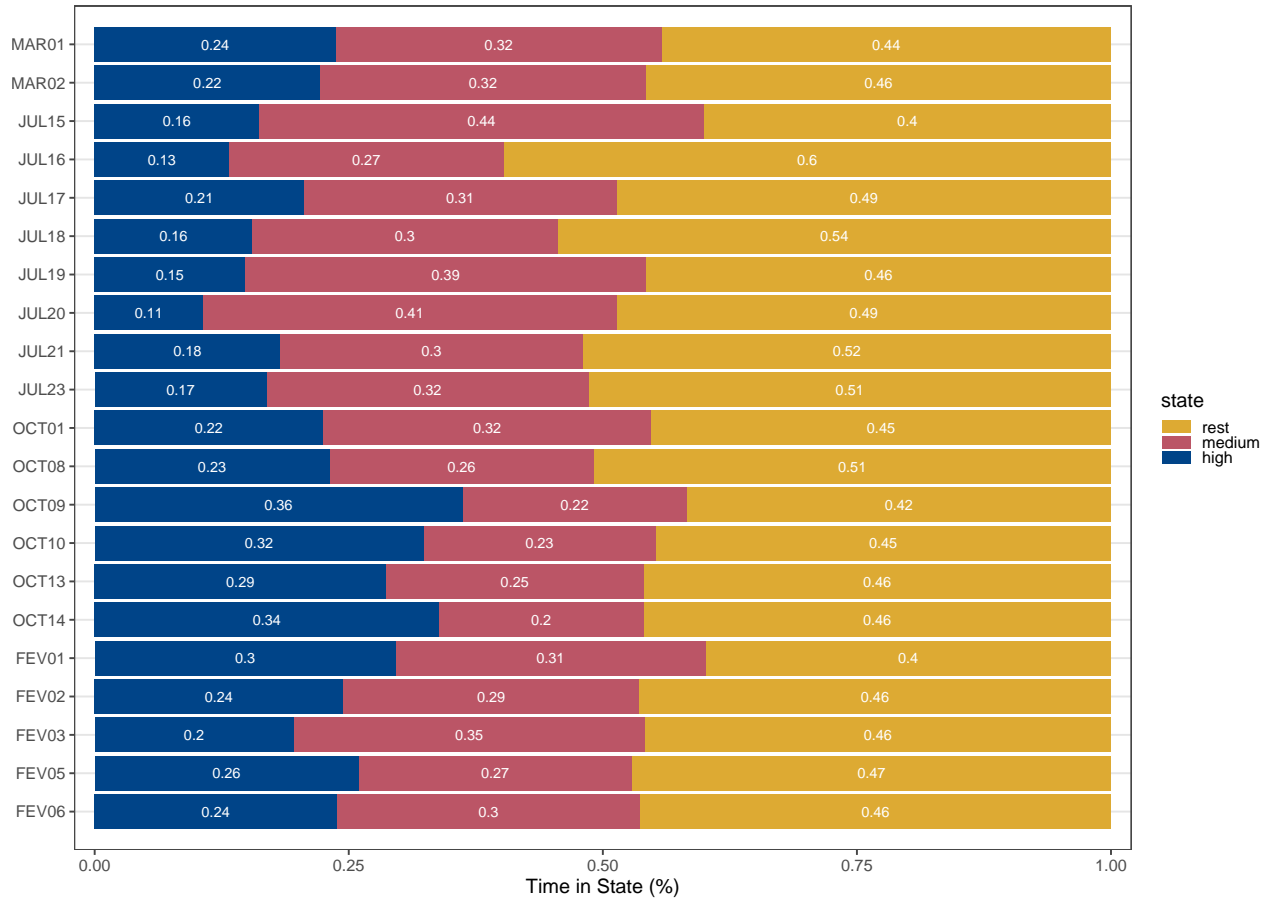
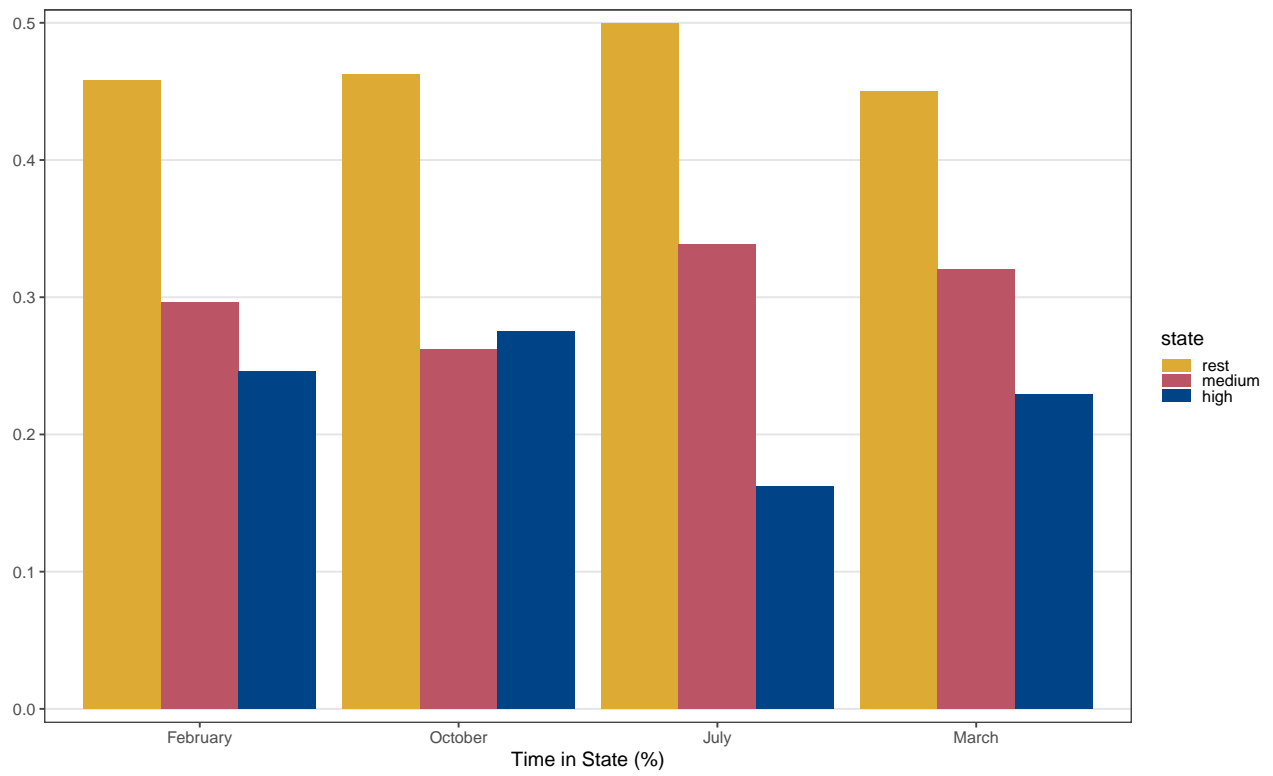


Figure 5: Proporção do tempo de registro total que cada animal passou nos estados identificados pelo HMM.



Atividade de Superfície

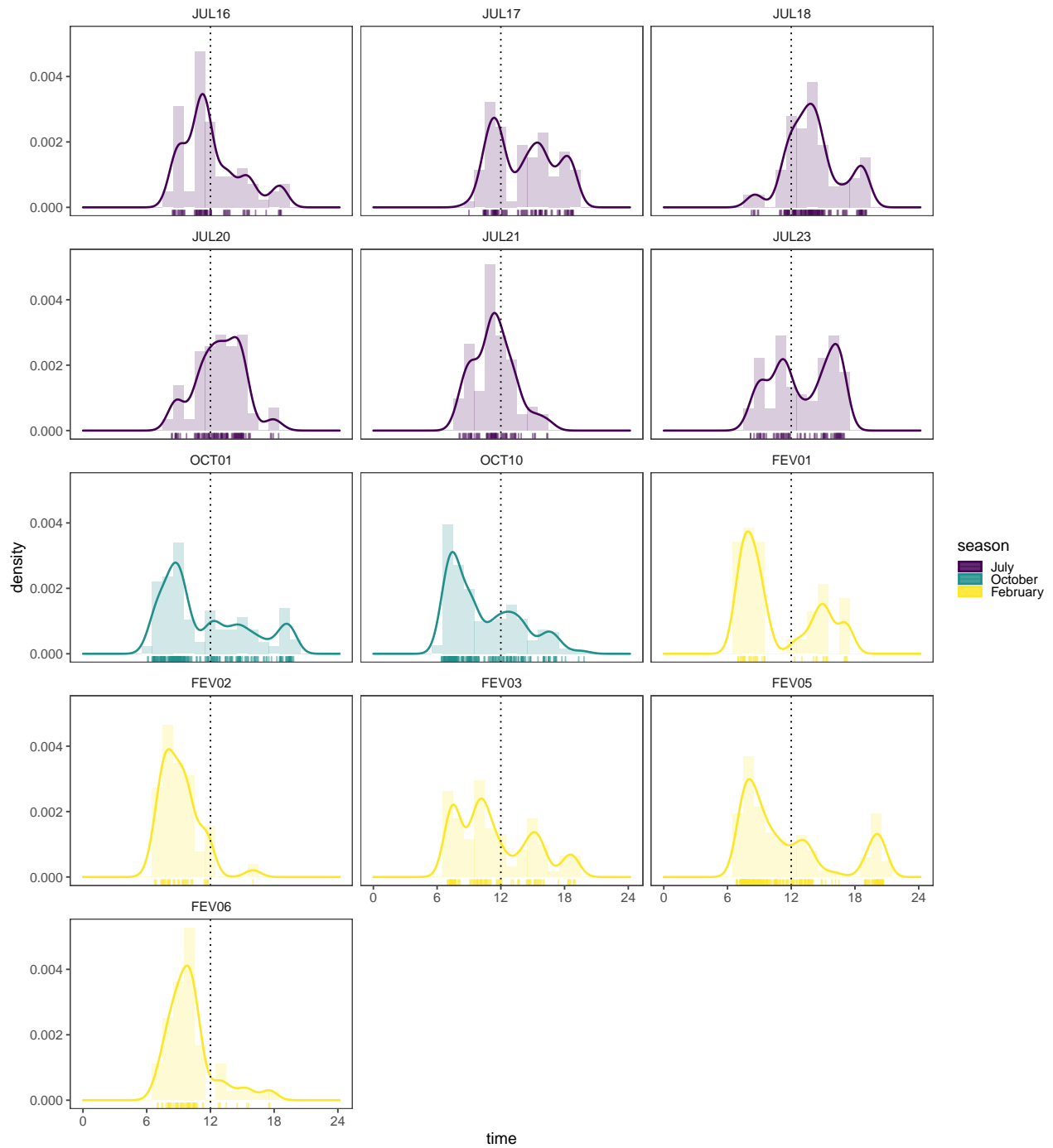
Visto como se comportam os dados da atividade geral agora vamos conferir os dados da atividade de superfície.

- Timing da Atividade de superfície:
 - Como as saídas à superfície estão organizadas no tempo? (O pico é maior no meio do dia no Inverno?)
- Proporção da atividade de superfície:
 - Quais estados são mais frequentes na superfície?

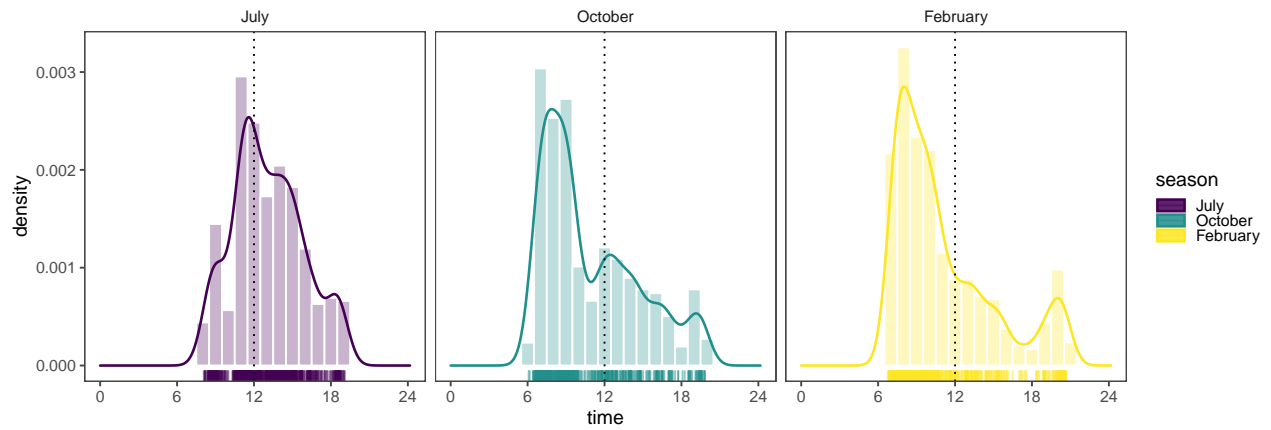
Padrão Diário de Saídas

Aqui estão coisas legais. Os dados de campo confirmar que os picos de saída estão mais concentrados no meio do dia em Julho do que em outras estações. Nas outras estações, porém, o segundo pico não é tão proeminente quanto o primeiro.

```
## Warning: Removed 26 rows containing missing values (geom_bar).
```



Warning: Removed 6 rows containing missing values (geom_bar).



Proporção dos Comportamentos que Acontecem na Superfície

Com os dados classificados também podemos calcular a proporção dos estados que acontecem na superfície. Aqui os resultados parecem mostrar que quando estão na superfície os tucos estão mais ativos.

