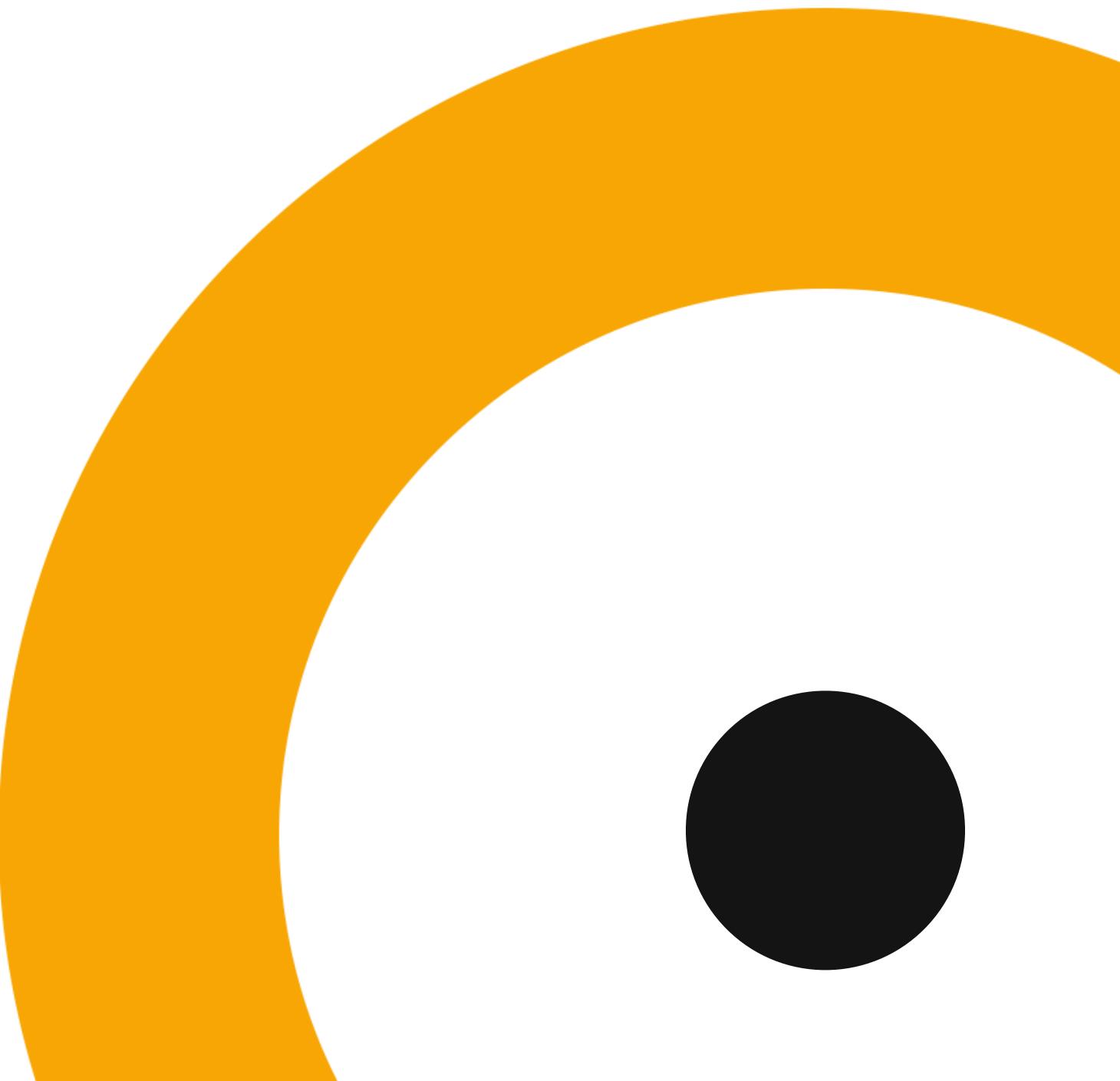


Mestrado em Data Science

Séries Temporais

Deteção de Anomalias em
Séries Temporais



3

Deteção de Anomalias em Séries Temporais



Deteção de Anomalias

O que é uma anomalia?

- Observação, ou conjunto de observações, que não seguem o padrão normal
- Também referidos como outliers
- Não confundir com ruído!
 - Por exemplo, erros de leitura
- **Observações raras, mas tipicamente relevantes**
 - Dados bastante desbalanceados
- Definição concreta varia de domínio para domínio



Deteção de Anomalias

Aplicações

- Mercados financeiros
 - Mudança abrupta no preço de um instrumento financeiro
- Manutenção preditiva
 - Falhas em equipamentos
- Atividade fraudulenta
 - Padrões de uso do cartão de crédito
- Sistema de saúde
 - Monitorização de pacientes hospitalares

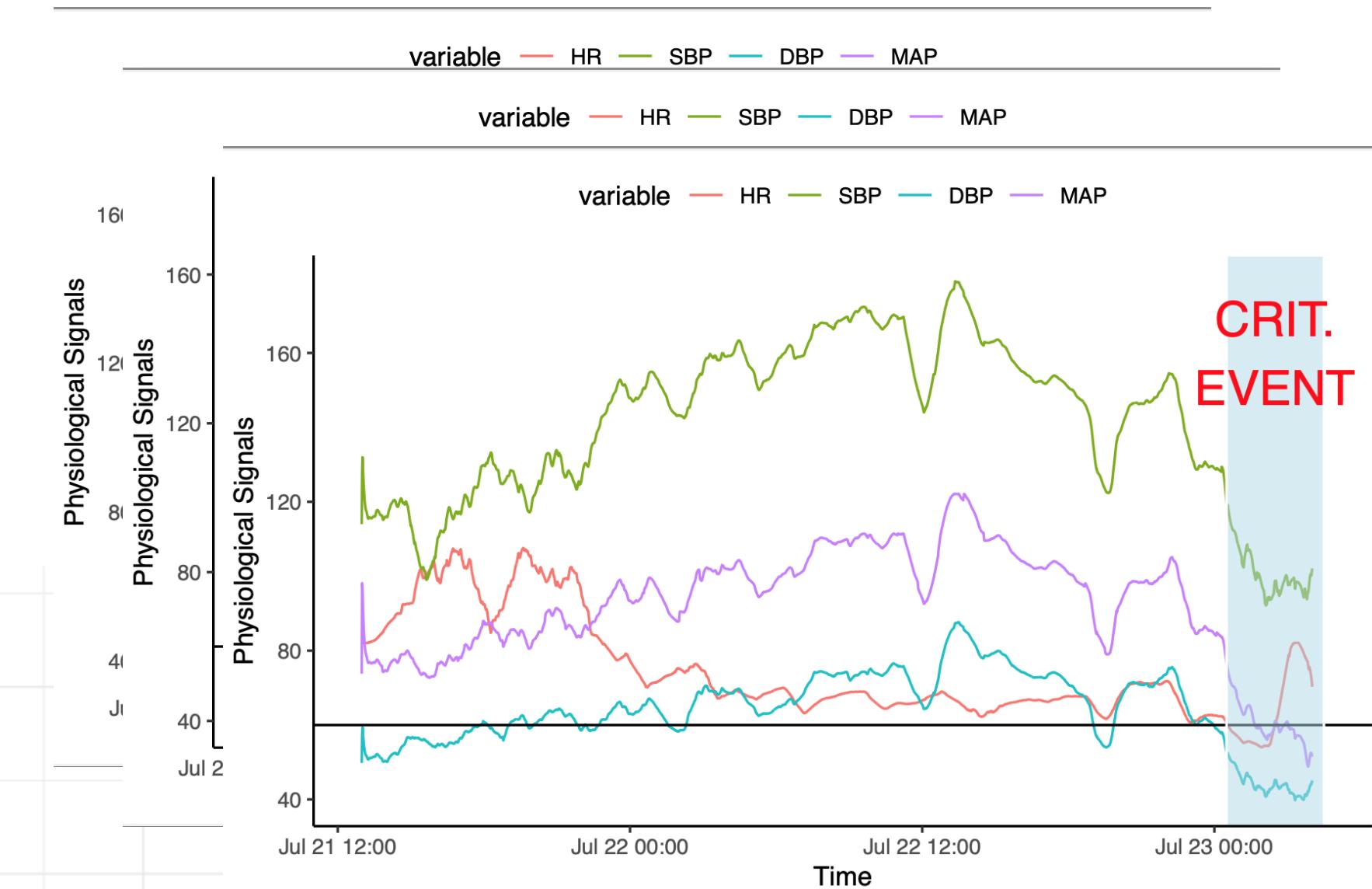
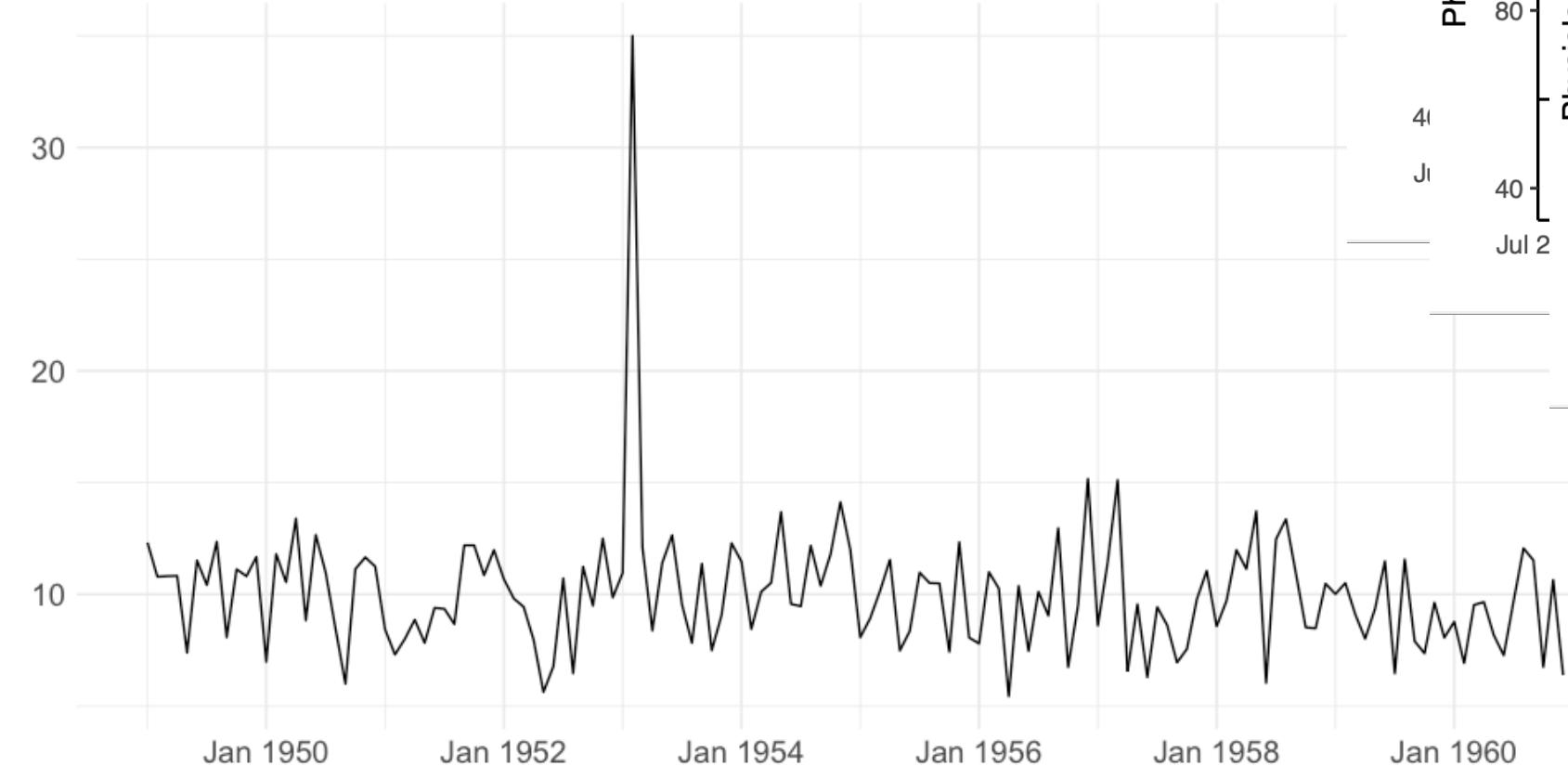
Tipos de Anomalias

Conjunto de séries vs Série individual

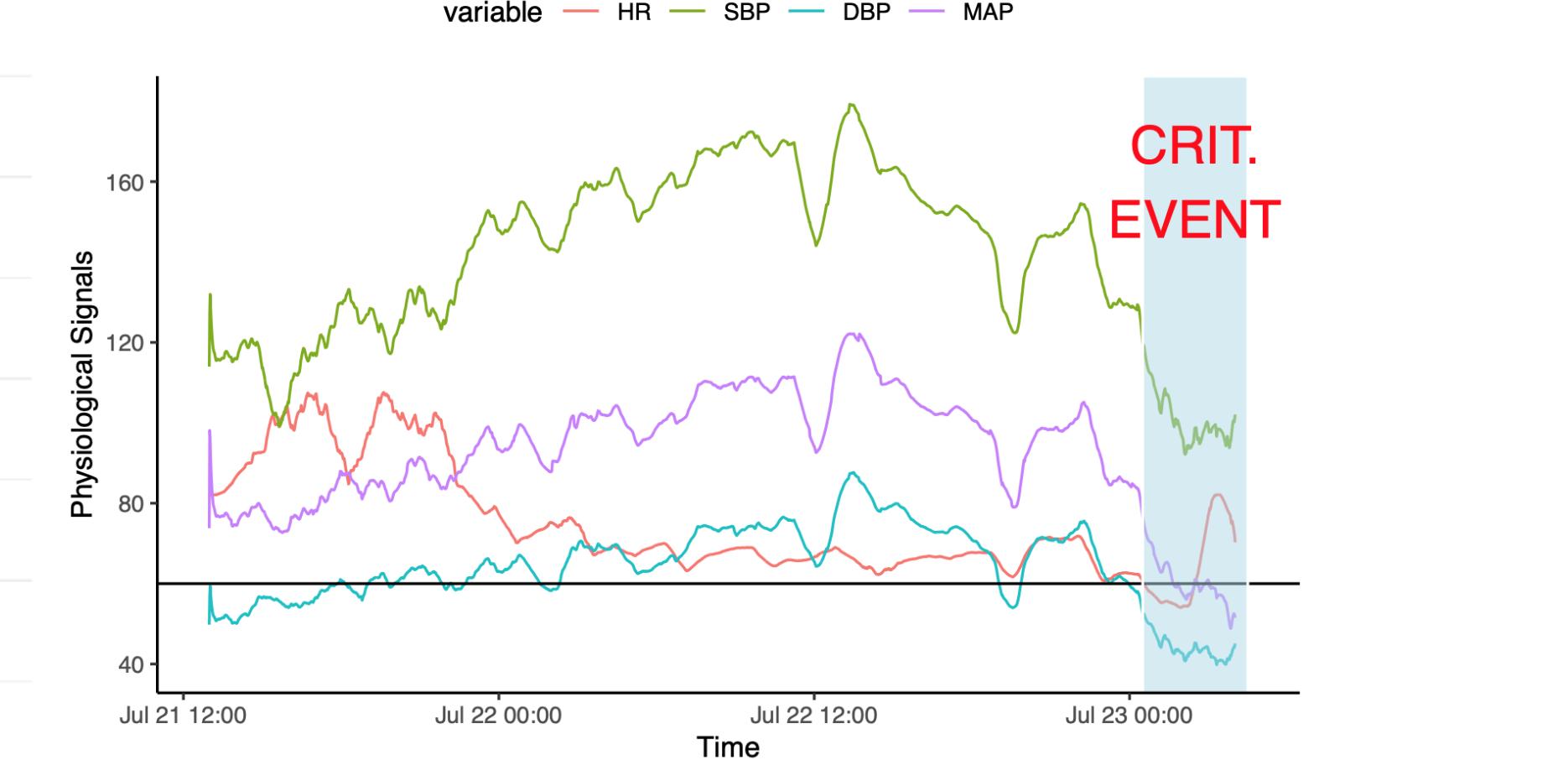
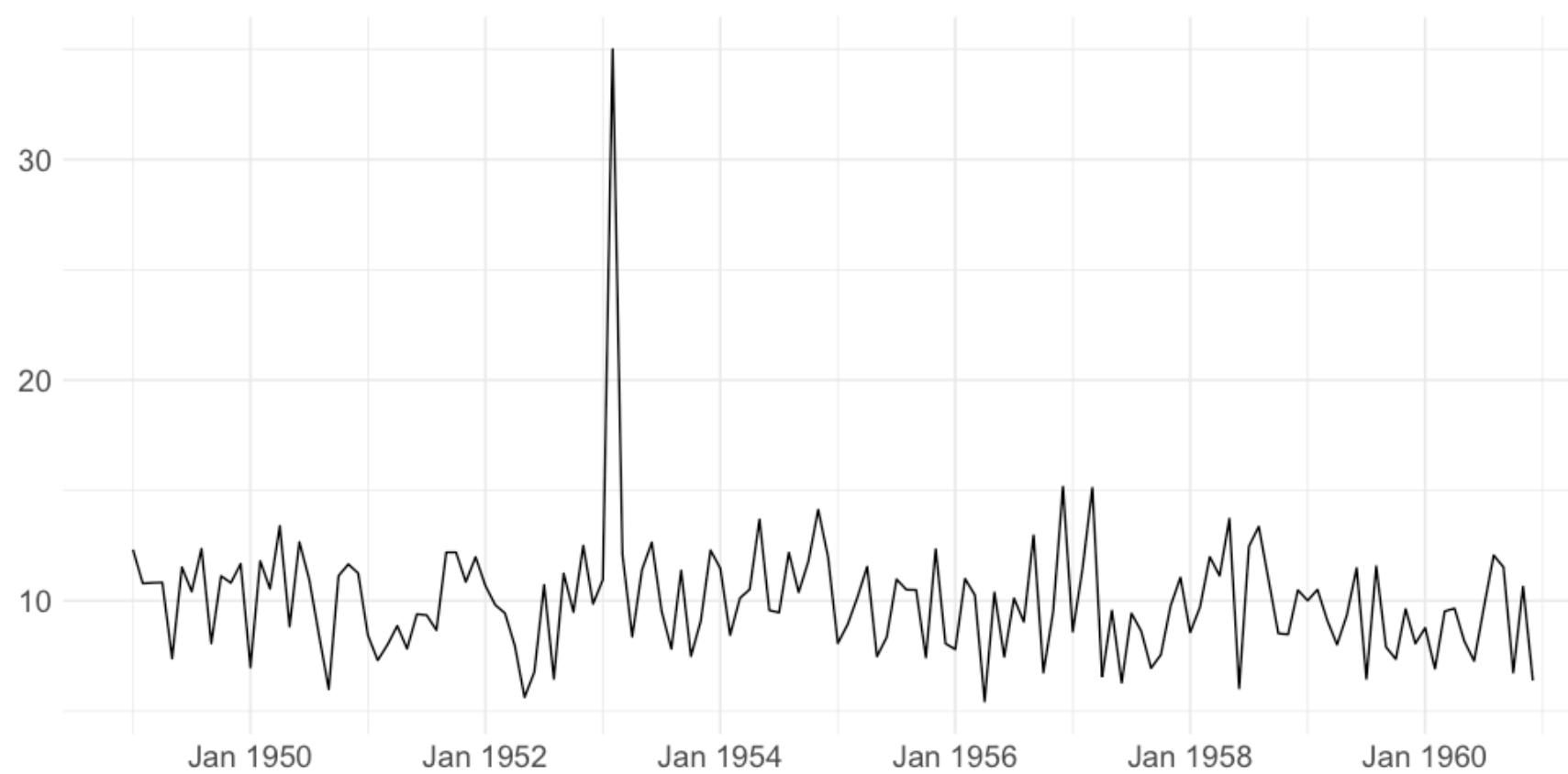
- Base de Dados de Séries Temporais

- Conjunto de pacientes hospitalares
- Conjunto de equipamentos

- Série Temporal Individual



Séries podem ser univariadas ou multivariadas



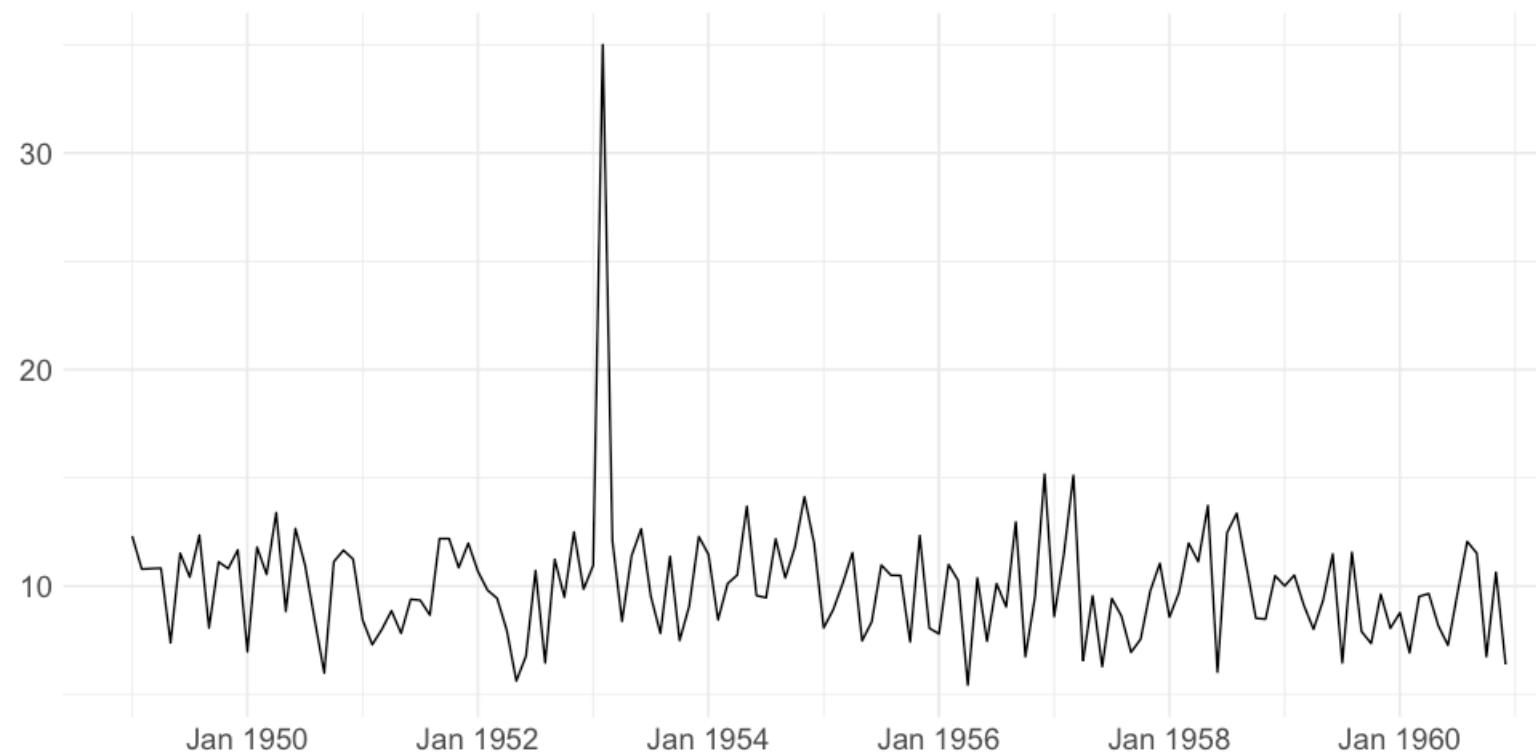
Tipos de Anomalias

Anomalias pontuais ou em subsequência, e contexto

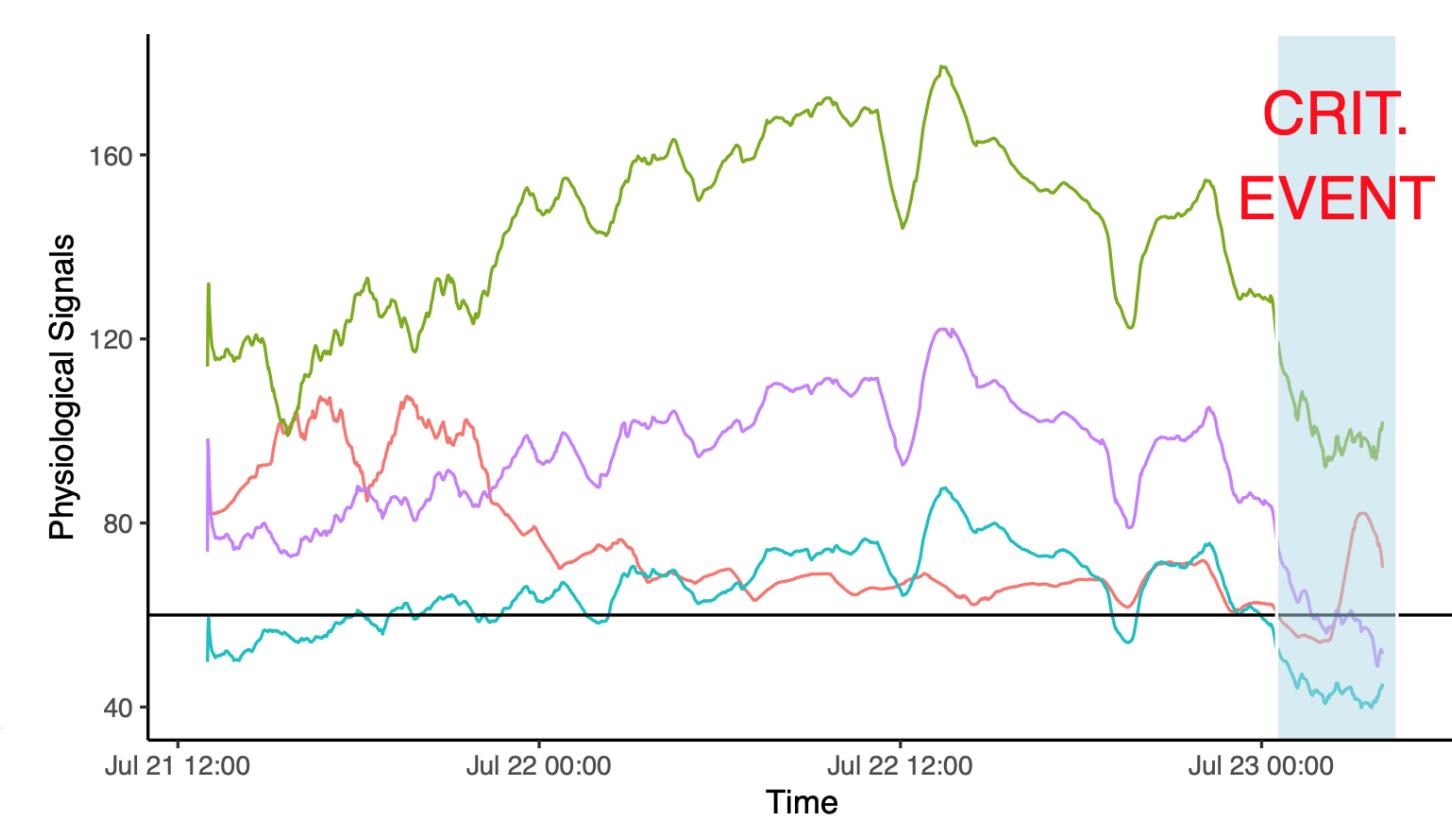
- **Anomalias em Subsequência**

- Abrangem vários pontos

- **Anomalia pontual**



variable — HR — SBP — DBP — MAP

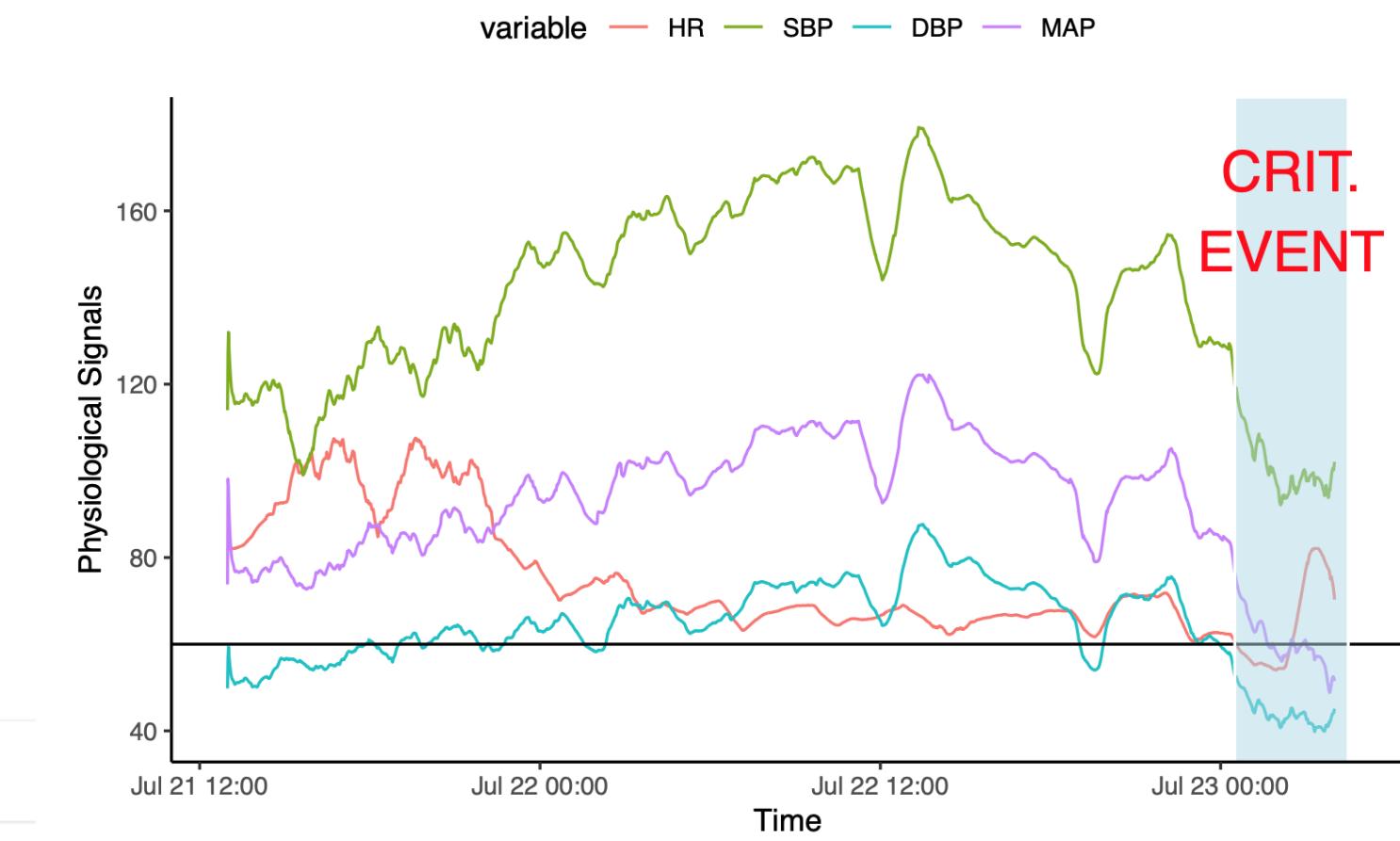
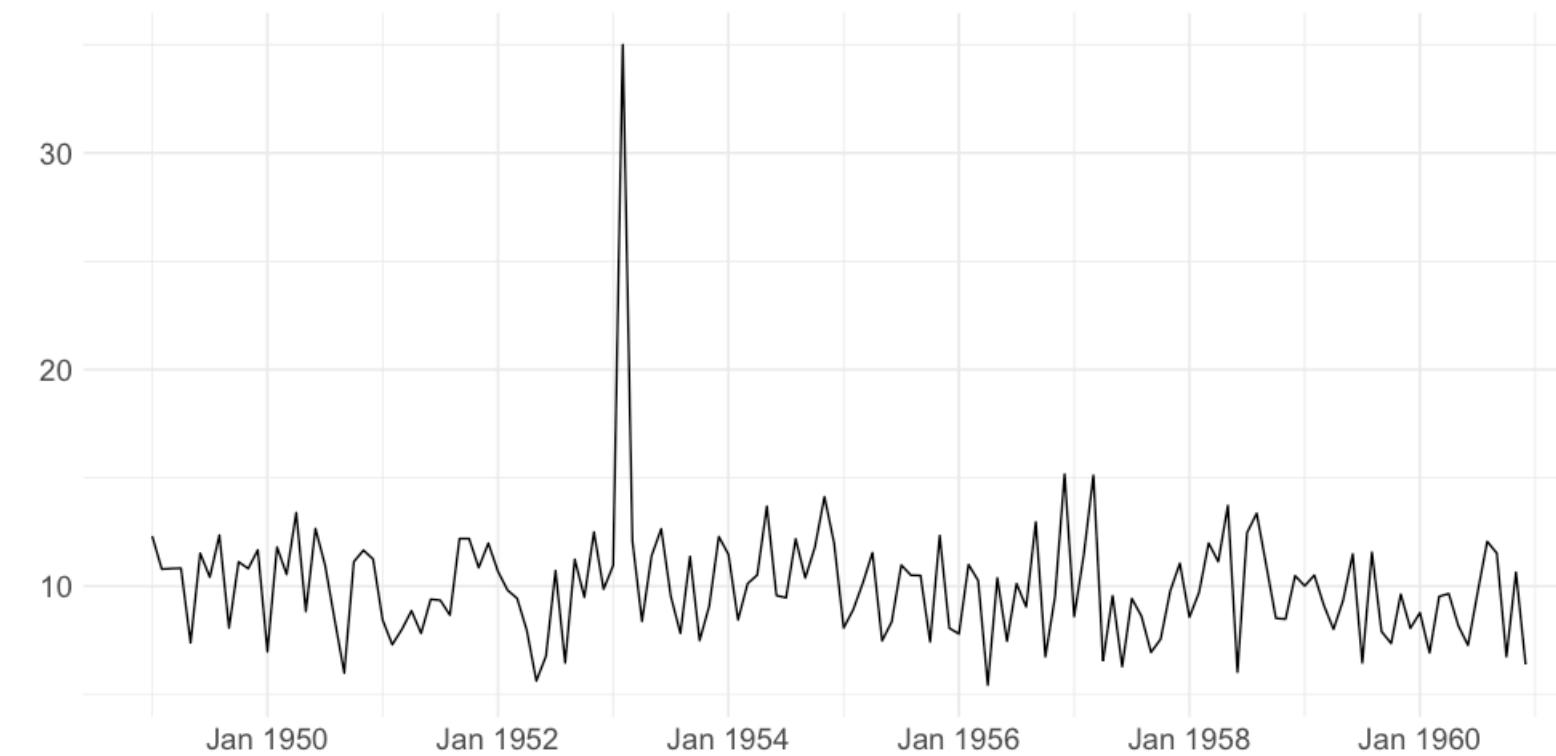


Tipos de Anomalias

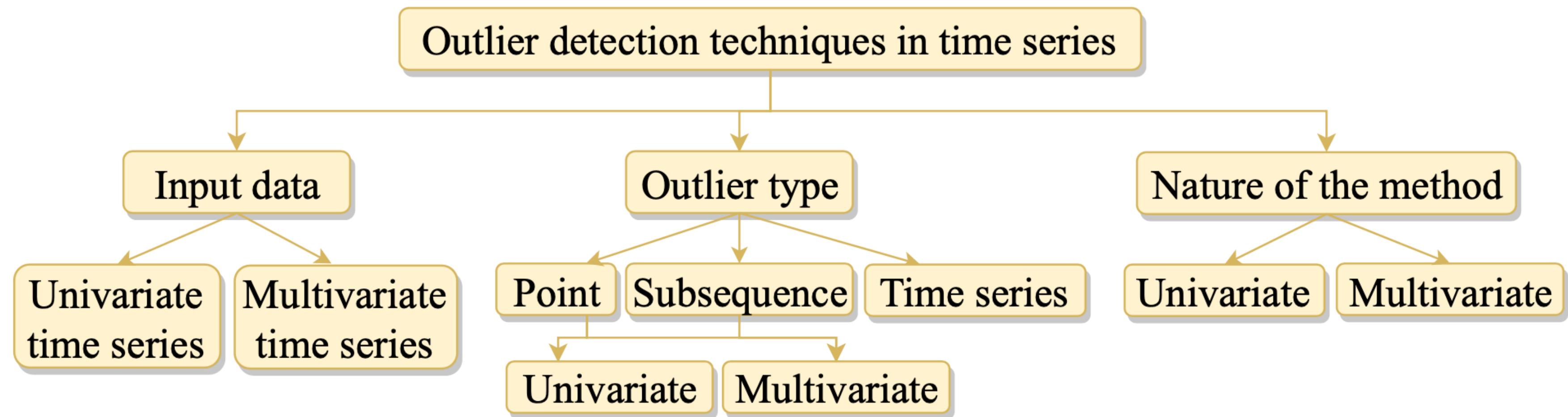
Anomalias pontuais ou em subsequência, e contexto

Contexto

- Uma temperatura de 30º é normal no verão, mas uma anomalia no inverno.



Taxonomia

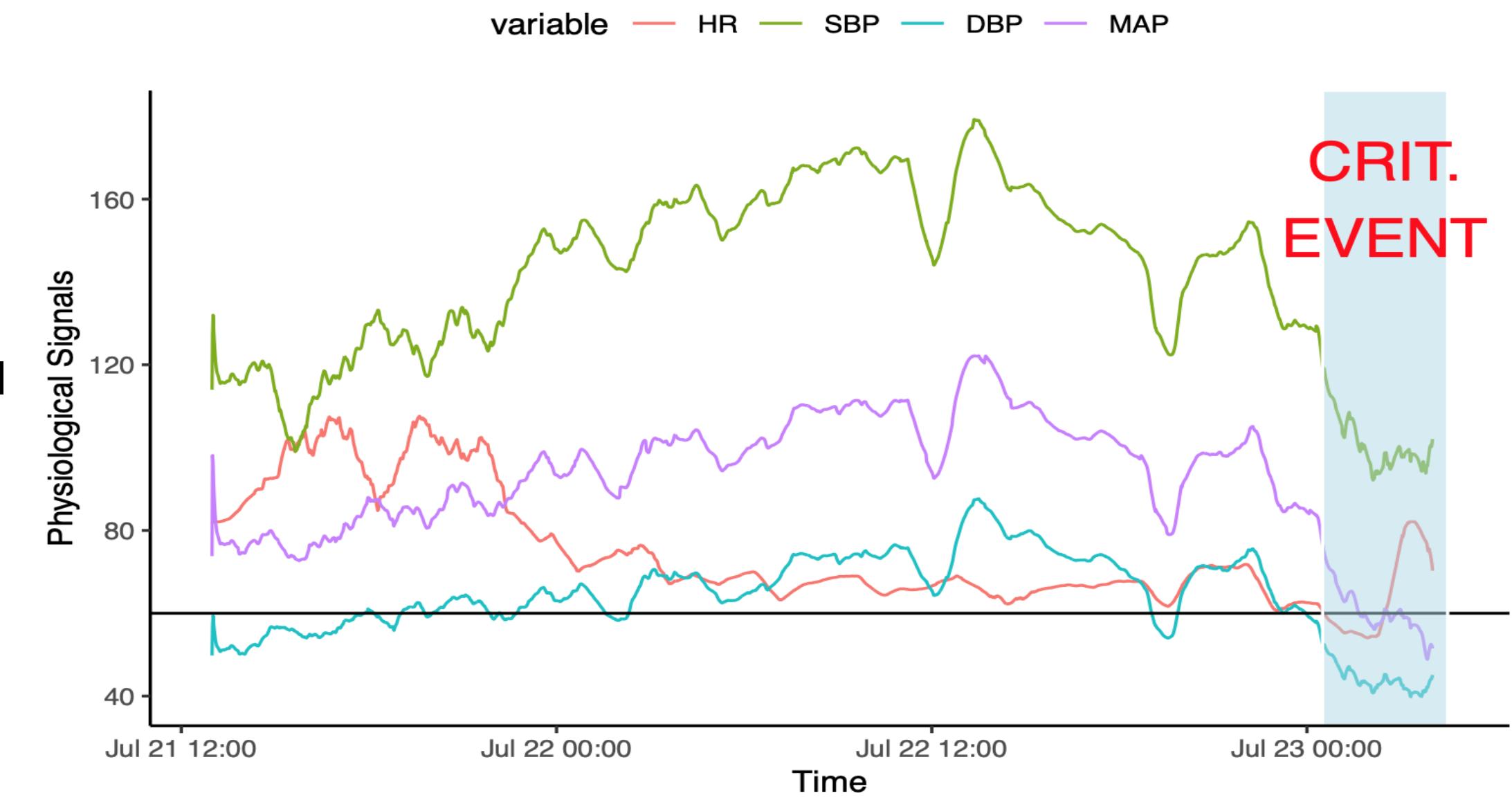


Fonte da imagem: Blázquez-García, Ane, et al. "A review on outlier/anomaly detection in time series data." *ACM Computing Surveys (CSUR)* 54.3 (2021): 1-33.

Importância do Aspetto Temporal

Alarmes Atempados

- Prever atempadamente é fundamental em domínios onde anomalias têm efeitos devastadores
- Importante detetar anomalias assim que possível



Abordagens

1. Estatísticas em Janela

- a. Observação é considerada anómala se o valor se desviar significativamente da média recente

2. Baseada em previsão

- a. Comparação da previsão do modelo e o valor real
- b. Grande discrepância indica anomalia

3. Densidade

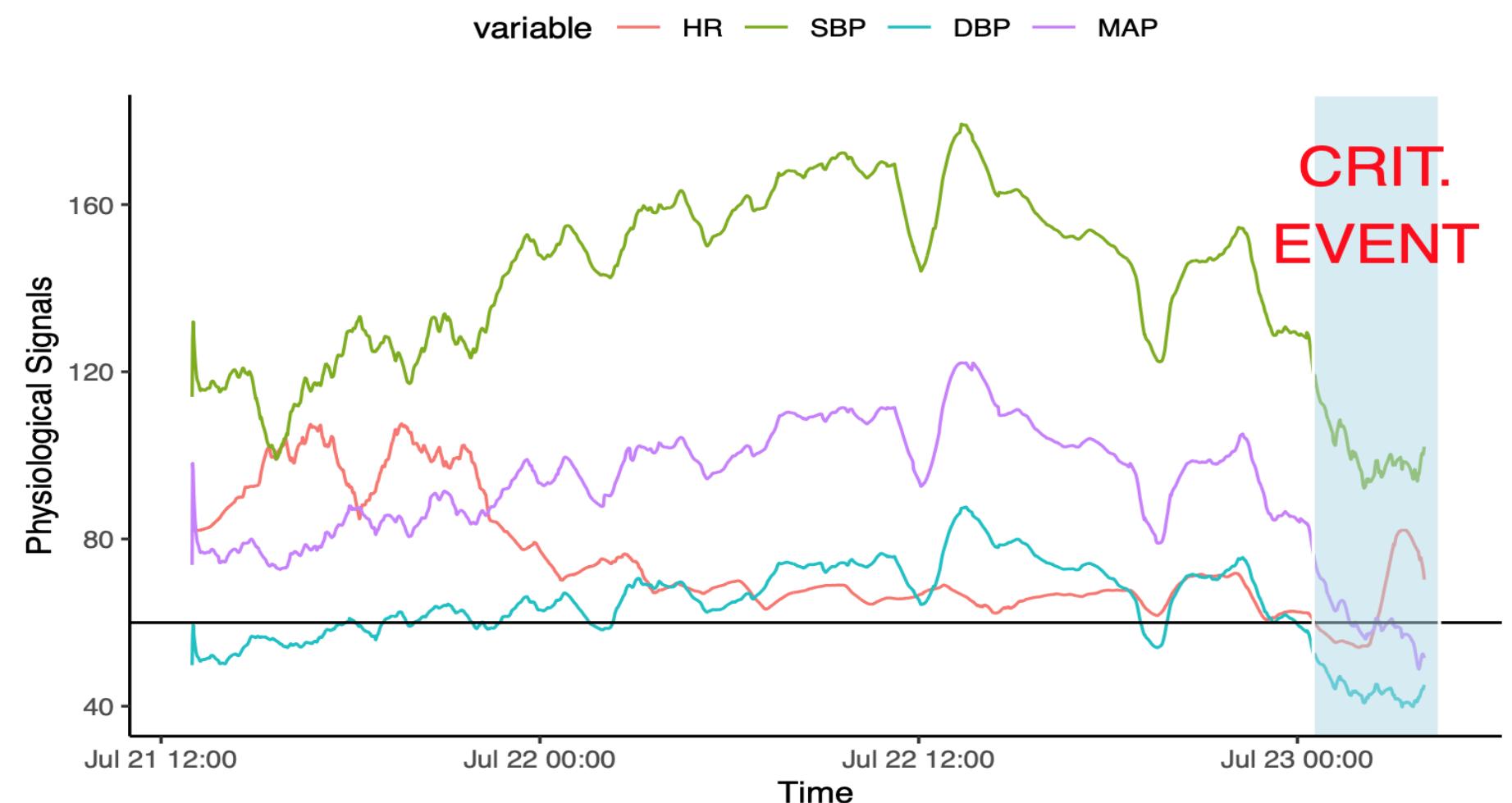
- a. Análise da distância para os vizinhos mais próximos
- b. Observações isoladas são consideradas anómalias

4. Histogramas

- a. Representação da série em histograma, com e sem uma dada observação
- b. Diferenças na representação indicam anomalia

5. Frequência (para subsequências)

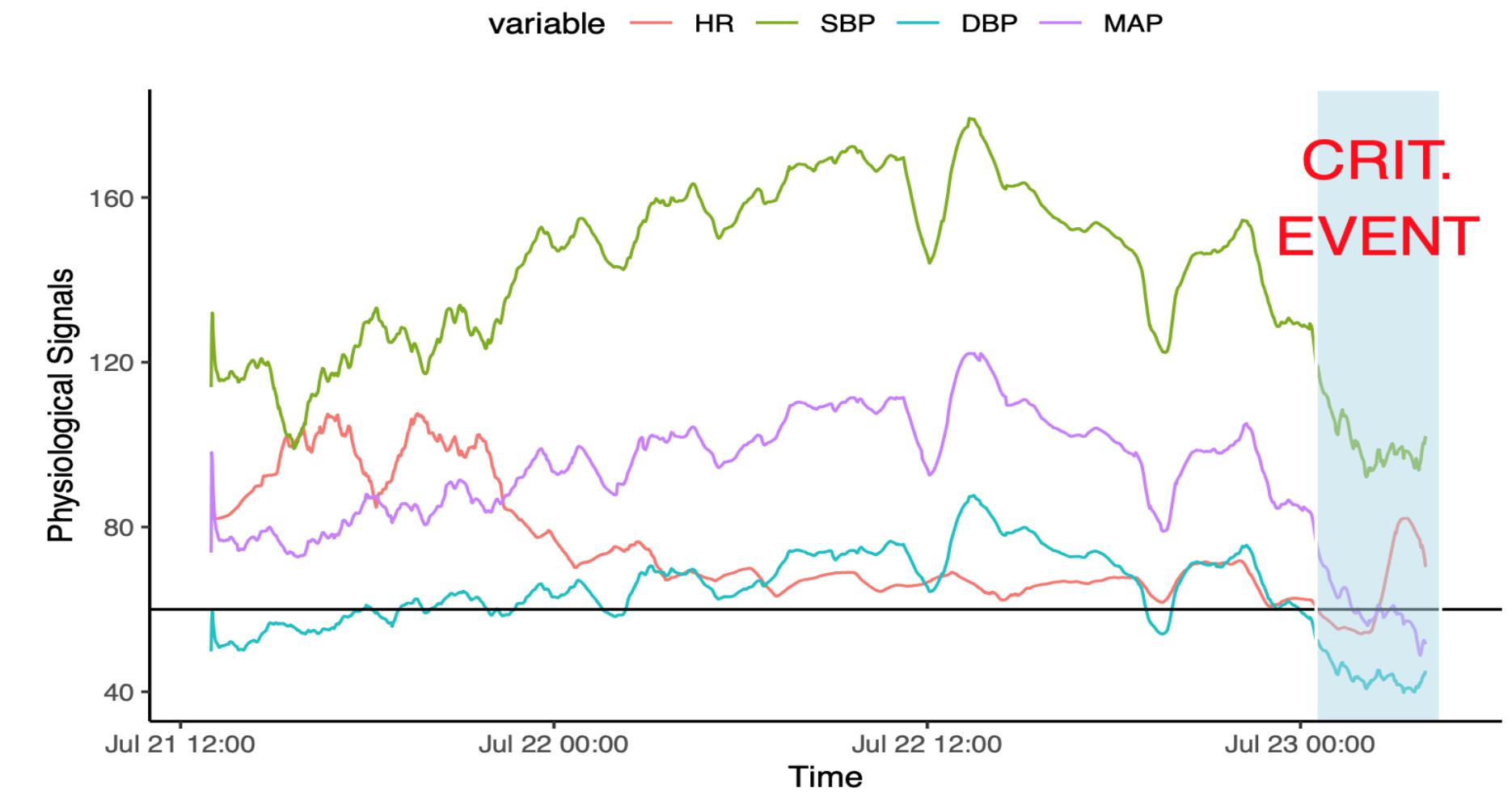
- a. Comparação da subsequência atual com uma subsequência de referência
- b. Grande discrepancia indica anomalia



Abordagens

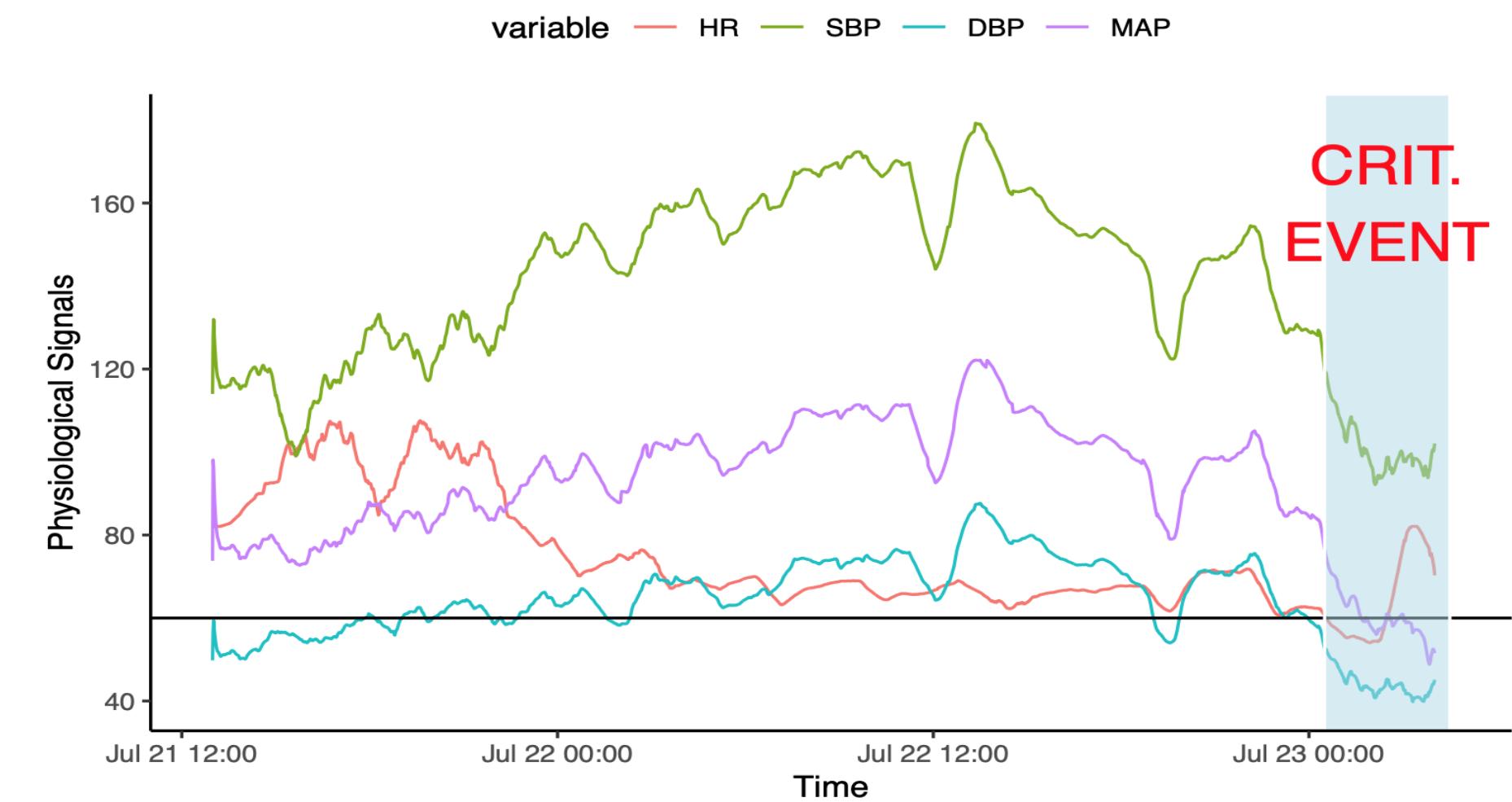
Discriminativa (Modelação preditiva)

- a. Criar um modelo de classificação binária que distingue explicitamente o comportamento normal do anómalo;
- b. Produz probabilidades de uma dada observação ser anómala



Abordagens

- **Global**
 - a. Construir um modelo com base em todas as séries disponíveis
- **Individual**
 - a. Construir um modelo apenas com base no histórico da série objetivo



Construção do Conjunto de Dados

Abordagem discriminativa

- **Janela de observação**
 - a. passado recente
- **Janela de espera**
 - a. período necessário para que as previsões sejam úteis
- **Janela objetivo**
 - a. período onde existe ou não anomalia

Signal — HR — SBP — DBP — MAP

Janelas com tamanhos dependentes do domínio

