**Modelagem**

1. Dados foram separados 70-30
2. Modelos PLSR calibrados com 70% dos dados

model1 - COS ~ SMO (355:2500 nm)

model2 - COS ~ SGD (355:2500 nm)

model3 - COS ~ index

model4 - COS ~ SGD (355:2500 nm) + index

model5 - argila ~ SMO (355:2500 nm)

model6 - argila ~ SGD (355:2500 nm)

model7 - argila ~ index

model8 - argila ~ SGD (355:2500 nm) + index

1. Parametros control

Para encontrar o melhor número de componentes

Foi usado a função trainControl do pacote **caret**, validação cruzada, 10 kfolds.

O melhor número de componentes foi usado no modelo final.

Covariáveis com variância próxima de zero foram excluídas.

1. Validação 30%
2. Métricas

SECONDS – quando tempo o modelo demora para rodar em segundos – converter na internet para HH mm SS

Olhar a definição de cada métrica no pacote

<https://cran.r-project.org/web/packages/Metrics/Metrics.pdf>

Arquivos de saída

**Model\_eval –** tabela com as métricas de cada modelo

**Model\_importance** – tabela com a importância de cada variável –ordenar conforme coprimento de onda antes de plotar.

**Model\_pred** – dados medidos e preditos conjunto de validação

**RESULTADOS:**

Apresentar histogramas

Apresentar estatísticas das propriedades soc e argila em uma tabela

Apresentar correlações – espectro e índices

Apresentar tabela de acurácia dos modelos

Apresentar figura predito (V1) vs observado (V2)