

# Supplemental Appendix in support of PAPER TITLE

Subodh A, Sanford Eigenbrode, Erich Seamon

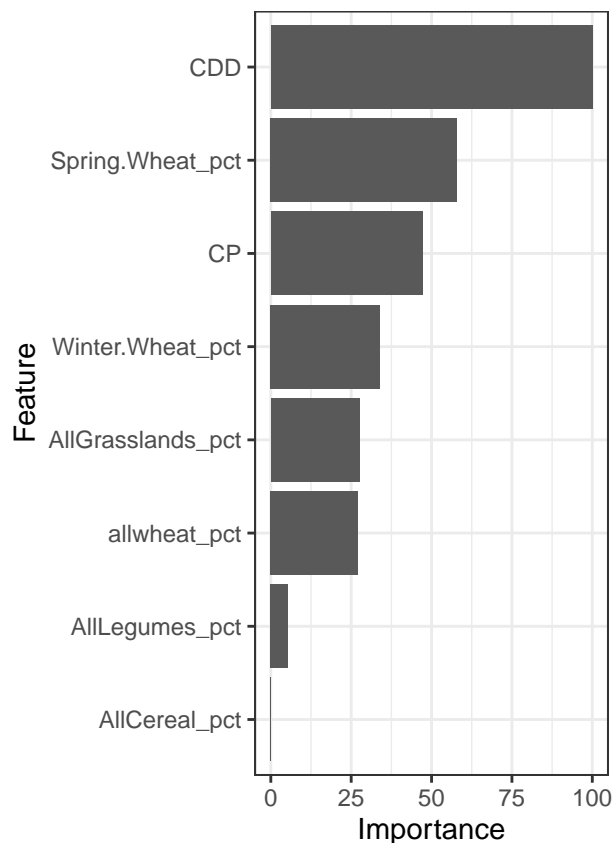
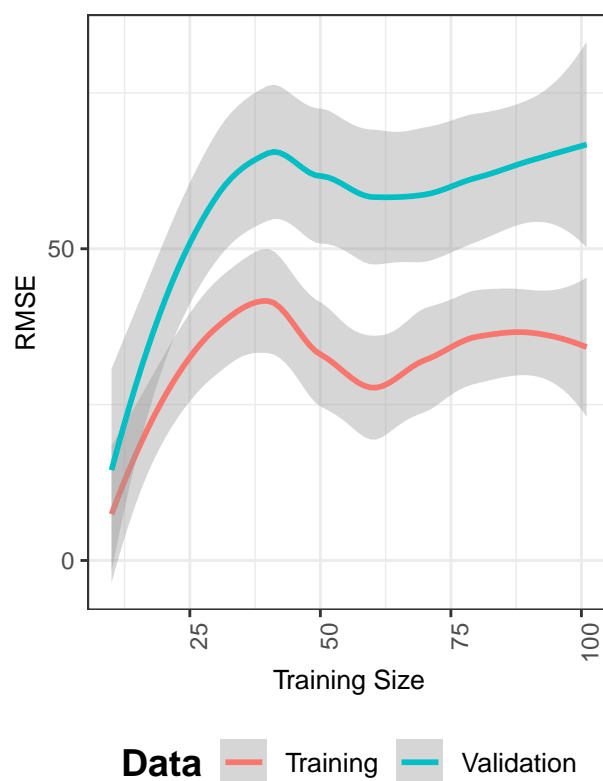
10/22/2021

## Contents

Total Aphids Model . . . . .	2
Total Aphids No Mfc Model . . . . .	3
Relative Abundance Mfc Model . . . . .	4
Individual Aphids: Rp . . . . .	5
Individual Aphids: Sa . . . . .	6
Individual Aphids: Md . . . . .	7
Individual Aphids: Mfc . . . . .	8
Individual Aphids: Rp . . . . .	9

## Total Aphids Model

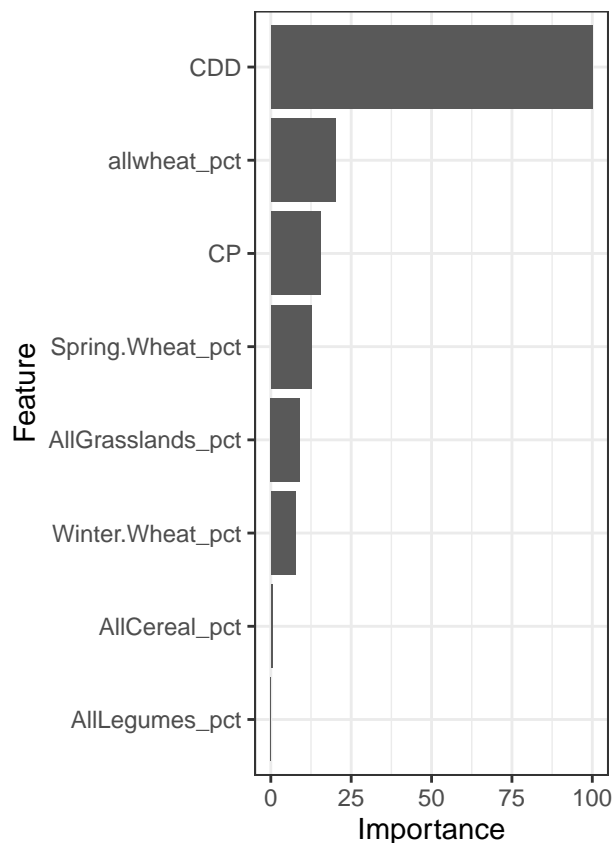
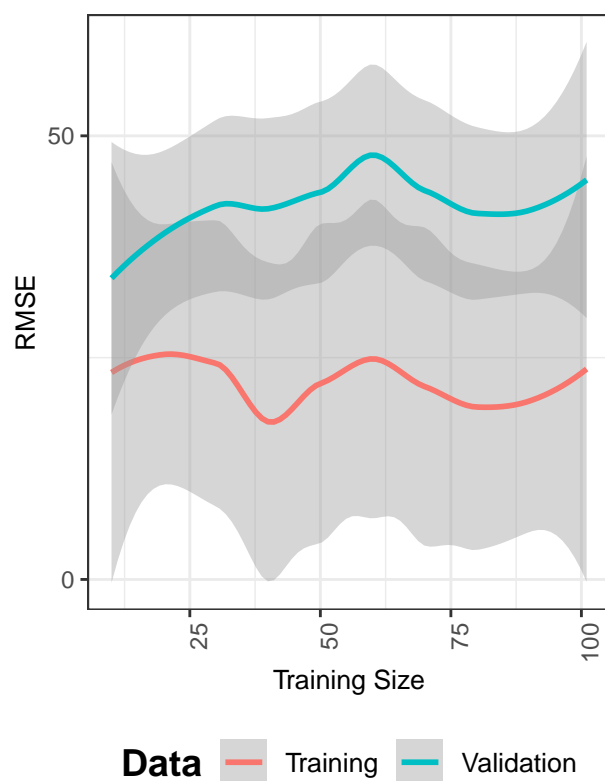
### Learning Curve Analysis



```
## Random Forest
##
## 101 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 90, 92, 91, 91, 90, ...
## Resampling results across tuning parameters:
##
## mtry  RMSE      Rsquared  MAE
## 2     63.73568  0.1811534  47.38910
## 5     65.99397  0.1735441  48.38993
## 8     68.40480  0.1574900  49.77404
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

## Total Aphids No Mfc Model

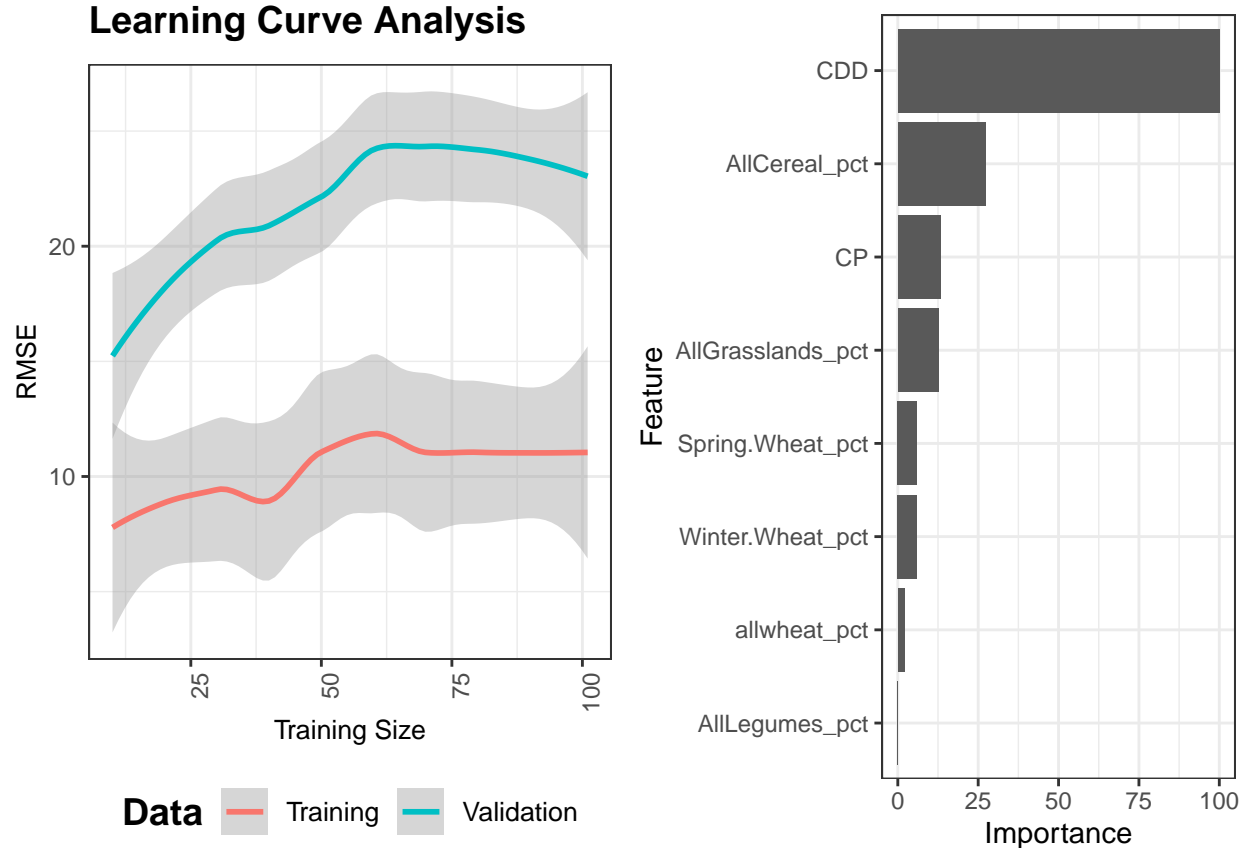
### Learning Curve Analysis



```
## Random Forest
##
## 101 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 90, 92, 91, 91, 90, ...
## Resampling results across tuning parameters:
##
## mtry  RMSE      Rsquared  MAE
## 2      49.33028  0.2563659  36.33553
## 5      49.26247  0.2559647  35.33321
## 8      49.22627  0.2541812  34.64219
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 8.
```

## Relative Abundance Mfc Model

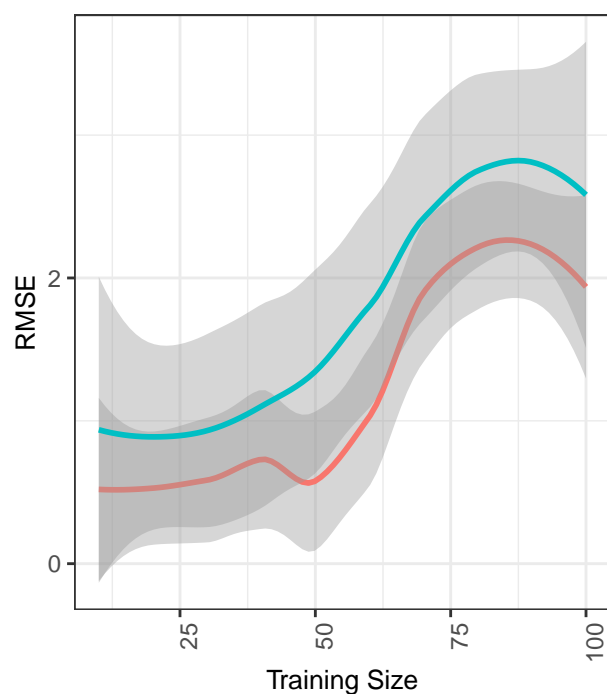
### Learning Curve Analysis



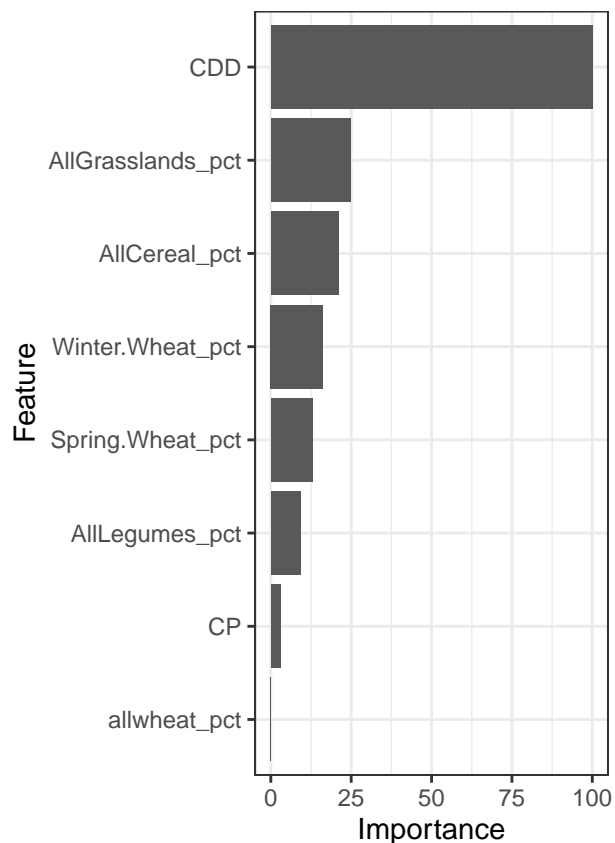
```
## Random Forest
##
## 101 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 90, 92, 91, 91, 90, ...
## Resampling results across tuning parameters:
##
## mtry RMSE      Rsquared  MAE
## 2     23.75841  0.2415788  19.97434
## 5     24.29576  0.2339301  20.53495
## 8     24.53101  0.2530602  20.70043
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

## Individual Aphids: Rp

### Learning Curve Analysis



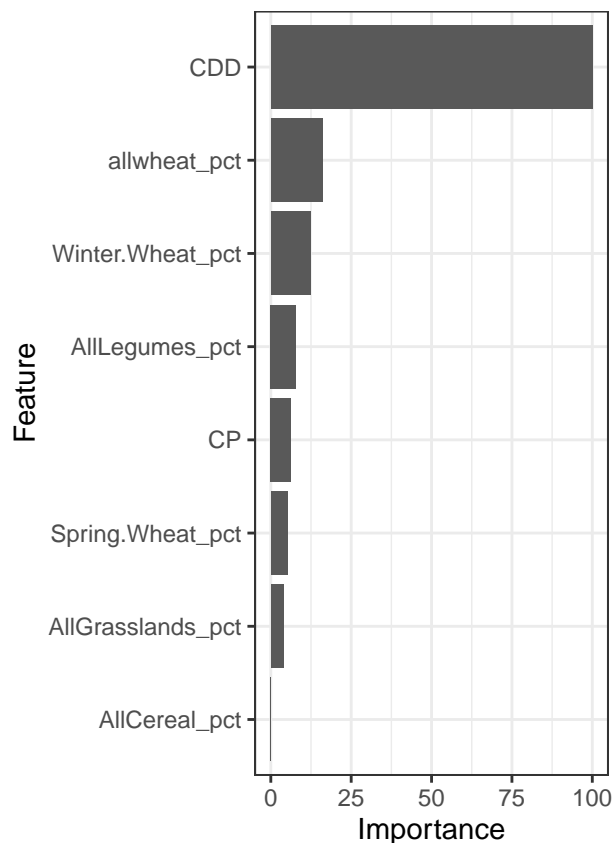
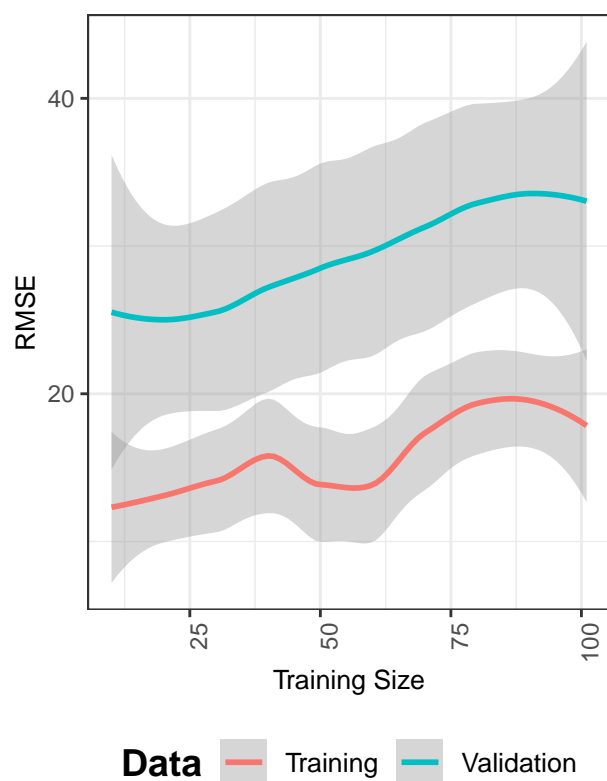
**Data** Training Validation



```
## Random Forest
##
## 100 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 91, 90, 89, 91, 90, ...
## Resampling results across tuning parameters:
##
## mtry RMSE Rsquared MAE
## 2 2.431611 0.09381127 1.541037
## 5 2.575098 0.18266845 1.536897
## 8 2.872344 0.16819091 1.625897
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

## Individual Aphids: Sa

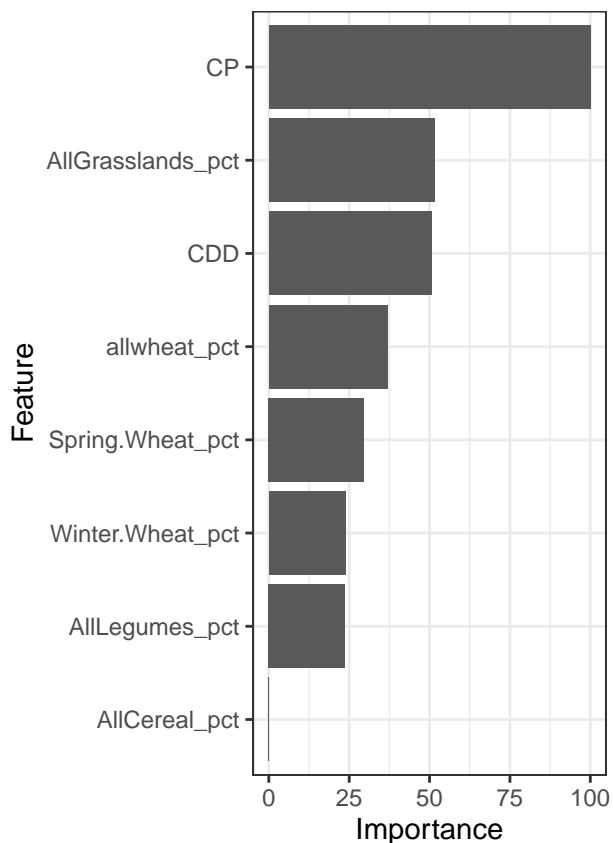
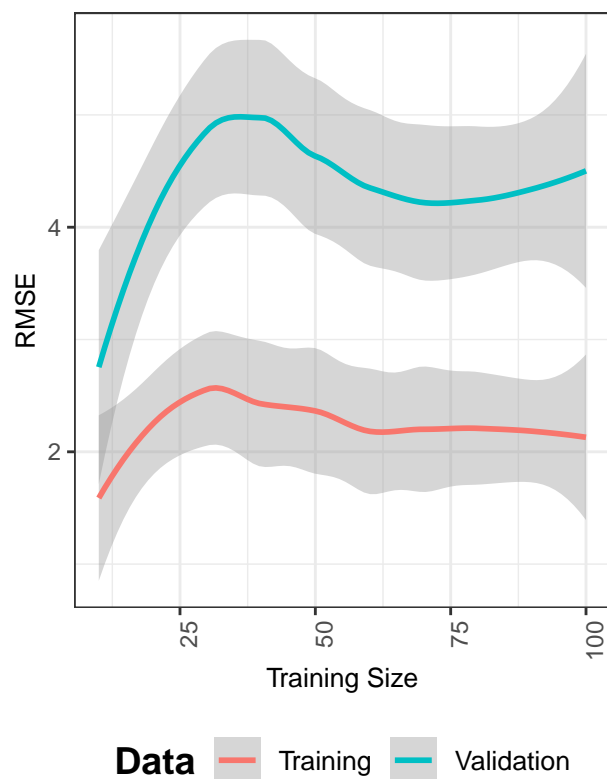
### Learning Curve Analysis



```
## Random Forest
##
## 101 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 90, 92, 91, 91, 90, ...
## Resampling results across tuning parameters:
##
## mtry RMSE Rsquared MAE
## 2 32.39298 0.4948489 21.99353
## 5 31.82151 0.5220429 21.25240
## 8 33.24165 0.5102561 21.71563
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 5.
```

## Individual Aphids: Md

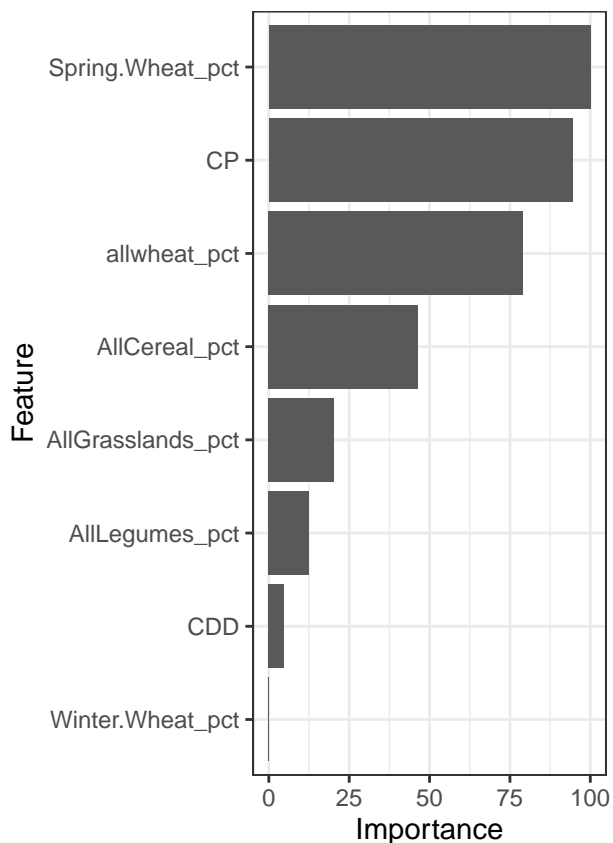
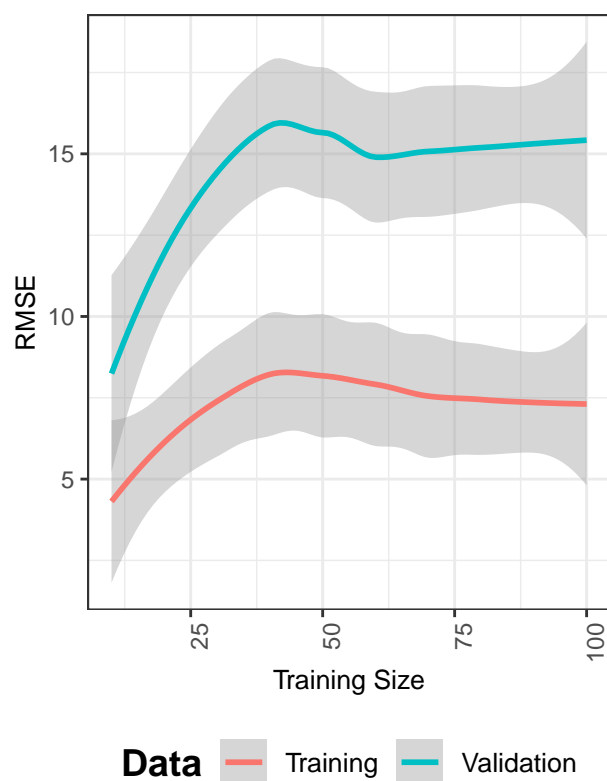
### Learning Curve Analysis



```
## Random Forest
##
## 100 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 90, 89, 89, 89, 91, 90, ...
## Resampling results across tuning parameters:
##
##  mtry  RMSE      Rsquared  MAE
##  2     4.277962  0.2618058  3.343674
##  5     4.335352  0.2462028  3.387907
##  8     4.410895  0.2265757  3.438736
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

## Individual Aphids: Mfc

### Learning Curve Analysis

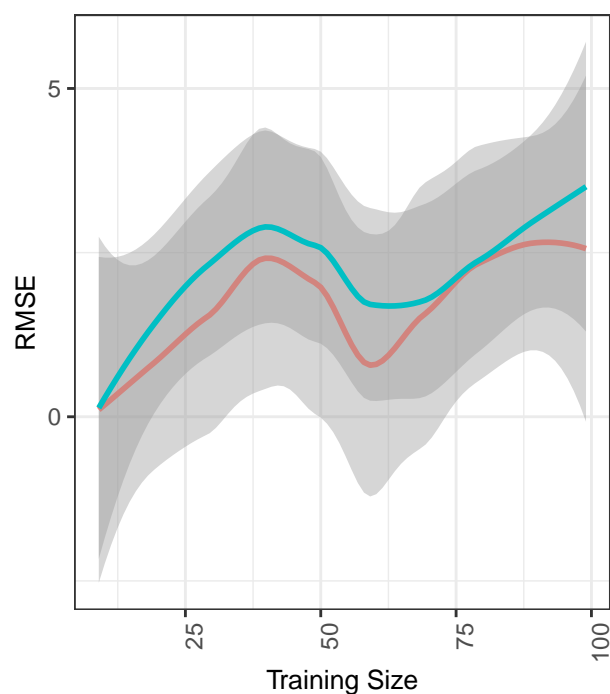


```
## Random Forest
##
## 100 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 90, 89, 89, 90, 90, 91, ...
## Resampling results across tuning parameters:
##
##  mtry  RMSE      Rsquared  MAE
##  2      14.57525  0.1221381  11.71662
##  5      15.00601  0.1049100  12.11249
##  8      15.07874  0.1190409  12.22297
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

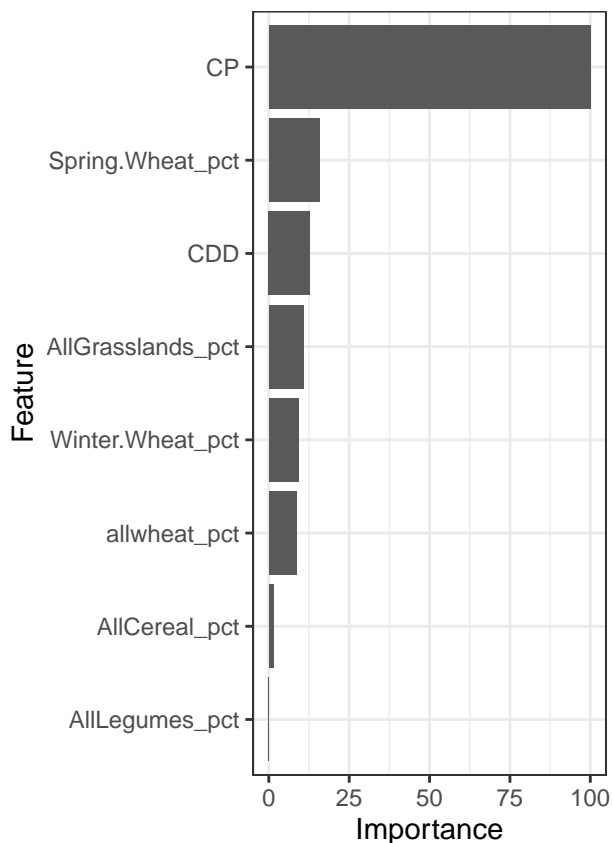


## Individual Aphids: Rp

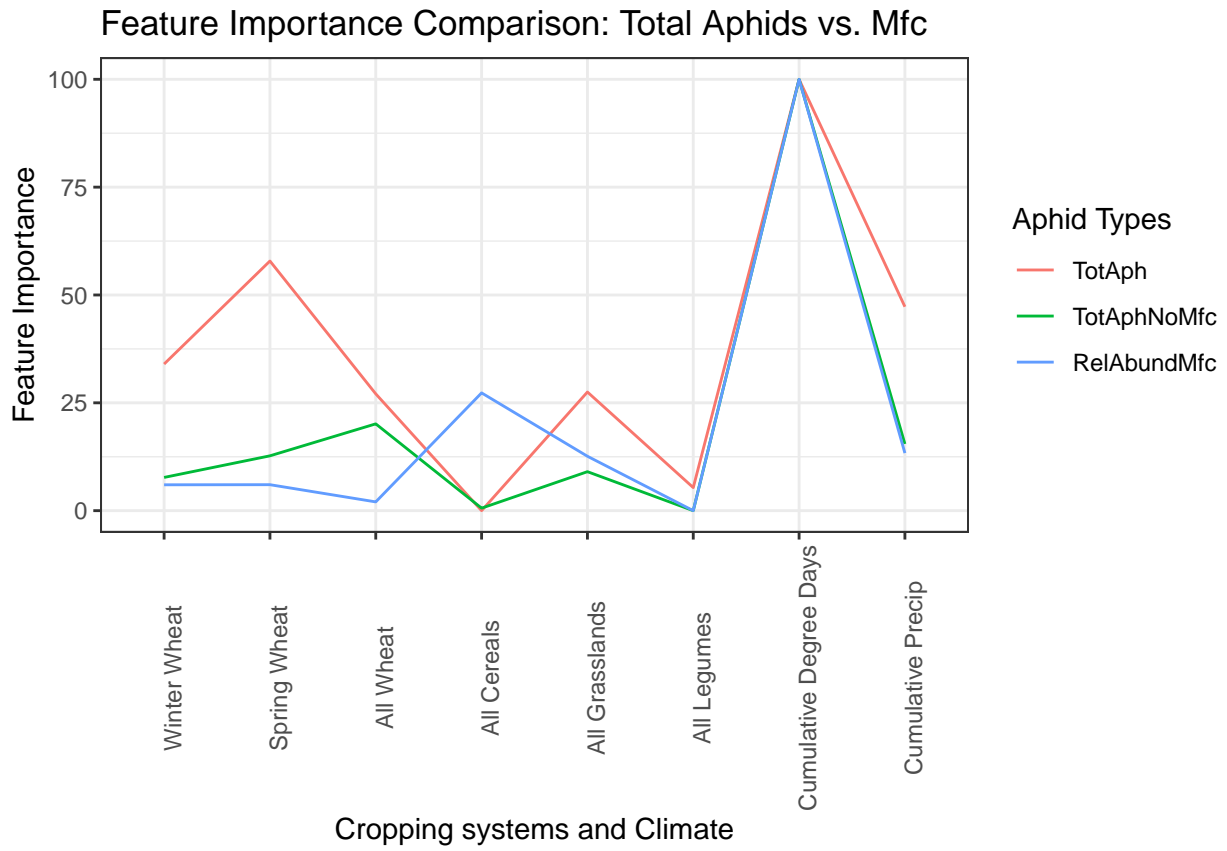
### Learning Curve Analysis



**Data** — Training — Validation



```
## Random Forest
##
## 99 samples
## 8 predictors
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 89, 89, 89, 89, 89, 89, ...
## Resampling results across tuning parameters:
##
##  mtry  RMSE      Rsquared  MAE
##  2     3.267462  0.2959596  1.431210
##  5     3.406572  0.2503583  1.420783
##  8     3.694826  0.2584216  1.461916
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```



Feature Importance Comparison: Other Aphids

