

Cuvier’s beaked whale foraging dives identified via machine learning using depth and triaxial acceleration: electronic supplement

David A. Sweeney et al.

Table of Contents

Create Models Using Testing and Training Datasets	2
Code S1.....	2
Figure S1.....	4
Figure S2.....	5
Comprehensive Model With Orientation-Dependent Predictors	6
Code S2.....	6
Comprehensive Model Without Orientation-Dependent Predictors	8
Code S3.....	8
Figure S3.....	10
Dives for Which Model and K-Means Classifications Differ	12
Figures S4-S23	12

Create Models Using Testing and Training Datasets

Code S1

```

# make partition
set.seed(987654321)
indexes <- createDataPartition(gtdives$Foraging,
                               times = 500,
                               p = 2/3,
                               list = TRUE)

#set up repeated cross validation
train.control <- trainControl(method = "repeatedcv",
                              number = 10,
                              repeats = 10,
                              search = "grid")

#set parameter grid search
tune.grid.tree <- expand.grid(eta = seq(.5, 1, .1),
                              max_depth = 1:3,
                              min_child_weight = seq(.25, 1, .25),
                              colsample_bytree = 1,
                              gamma = seq(.25, 1, .25),
                              nrounds = 1:4,
                              subsample = 1)

#run models
models <- train.acc <- test.acc <- besttune <- combined.acc <- list()
for (i in 1:length(indexes)) {
  #create train and test datasets
  foraging.train <- gtdives[indexes[[i]],]
  foraging.test <- gtdives[-indexes[[i]],]

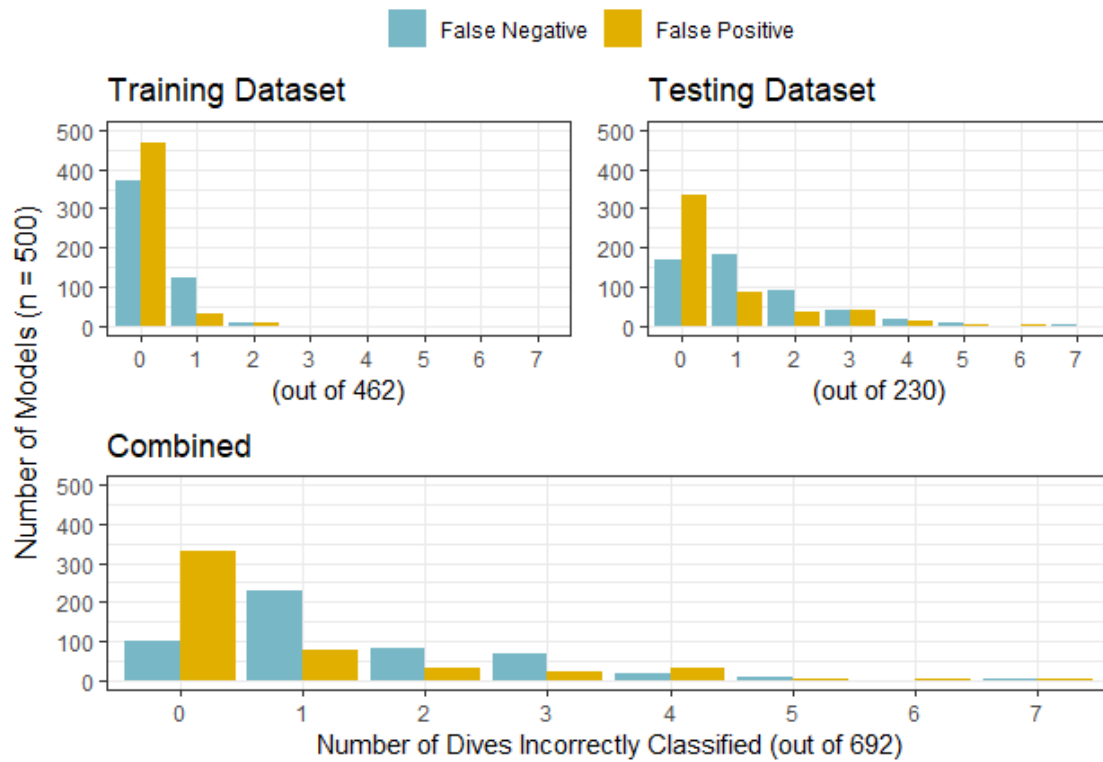
  #train in parallel.
  #register cluster so that caret will know to train in parallel.
  cl <- makeCluster(6, type = "SOCK")
  registerDoSNOW(cl)

  #train the xgboost model with parameters from full model
  set.seed(987654321)
  models[[i]] <- train(Foraging ~ DiveDepth + DiveDuration +
                      BottomPhaseVerticalSpeed +
                      BottomPhaseDuration +
                      BottomVertDirectionChangesProp +
                      DescBottomJerkPeaks +
                      DescRate + AscRate +
                      DescBottomRollVar,
                      data = foraging.train,
                      method = "xgbTree",
                      tuneGrid = tune.grid.tree,
                      trControl = train.control,
                      metric = "Accuracy")
}
stopCluster(cl)

```

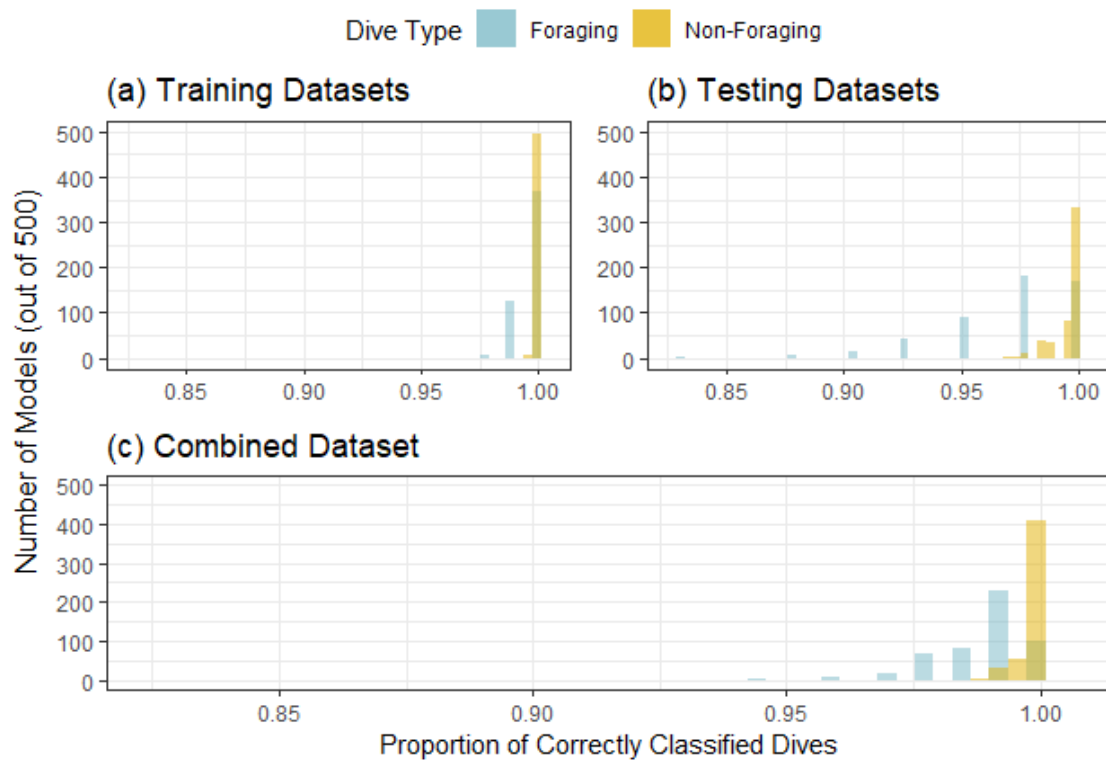
```
#determine accuracy of model against its own data and testing data
train.acc[[i]] <- confusionMatrix(predict(models[[i]], foraging.train),
                                foraging.train$Foraging)
test.acc[[i]] <- confusionMatrix(predict(models[[i]], foraging.test),
                                foraging.test$Foraging)
combined.acc[[i]] <- confusionMatrix(predict(models[[i]], gtdives),
                                    gtdives$Foraging)
besttune[[i]] <- models[[i]]$bestTune
}
```

Figure S1



Supplementary Figure 1: Bar graphs of the number of dives incorrectly classified (broken up by dive type) across the 500 models using randomly partitioned training and testing datasets. The top left panel shows the inaccuracies of the 500 models when fitting dives from the training datasets. The top right panel shows the inaccuracies of the 500 models when predicting dives from the testing datasets. The bottom panel shows the inaccuracies of the 500 models when classifying all dives for which acoustic data was available. Blue bars represent inaccuracies when classifying foraging dives (sensitivity), while the yellow bars represent inaccuracies when classifying non-foraging dives (specificity).

Figure S2



Supplementary Figure 2: Histograms of the proportion of correctly classified dives (broken up by dive type) across the 500 models using randomly partitioned training and testing datasets. The top left panel shows the ability of the 500 models to accurately fit the training datasets. The top right panel shows the ability of the 500 models to accurately predict the testing datasets. The bottom panel shows the ability of the 500 models to correctly classify all dives for which acoustic data was available. Blue bars represent the accuracy levels for classifying foraging dives (sensitivity), while the yellow bars represent the accuracy levels for classifying non-foraging dives (specificity).

Comprehensive Model With Orientation-Dependent Predictors

Code S2

```
#set up repeated cross validation
train.control <- trainControl(method = "repeatedcv",
                              number = 10,
                              repeats = 10,
                              search = "grid")

#set parameter grid search
tune.grid.tree <- expand.grid(eta = seq(.5, 1, .1),
                              max_depth = 1:3,
                              min_child_weight = seq(.25, 1, .25),
                              colsample_bytree = 1,
                              gamma = seq(.25, 1, .25),
                              nrounds = 1:4,
                              subsample = 1)

#train in parallel.
#register cluster so that caret will know to train in parallel.
cl <- makeCluster(6, type = "SOCK")
registerDoSNOW(cl)

#train the xgboost model with parameters from full model
set.seed(987654321)
foragingmod.bottom.tree <- train(Foraging ~ DiveDepth + DiveDuration +
                                BottomPhaseVerticalSpeed +
                                BottomPhaseDuration +
                                BottomVertDirectionChangesProp
+
                                DescBottomJerkPeaks +
                                DescRate + AscRate +
                                DescBottomRollVar,
                                data = gtdives,
                                method = "xgbTree",
                                tuneGrid = tune.grid.tree,
                                trControl = train.control,
                                metric = "Accuracy")

stopCluster(cl)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Yes  No
##           Yes 125  0
##           No   0 567
##
##           Accuracy : 1
##           95% CI : (0.9947, 1)
##           No Information Rate : 0.8194
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 1
##
## Mcnemar's Test P-Value : NA
##
##           Sensitivity : 1.0000
##           Specificity : 1.0000
##           Pos Pred Value : 1.0000
##           Neg Pred Value : 1.0000
##           Prevalence : 0.1806
##           Detection Rate : 0.1806
##           Detection Prevalence : 0.1806
##           Balanced Accuracy : 1.0000
##
##           'Positive' Class : Yes
##
##           nrounds max_depth eta gamma colsample_bytree min_child_weight subsa
mple
## 908           4           3 0.9 0.25           1           0.75
1
```

Comprehensive Model Without Orientation-Dependent Predictors

Code S3

```
#set up repeated cross validation
train.control <- trainControl(method = "repeatedcv",
                              number = 10,
                              repeats = 10,
                              search = "grid")

#set parameter grid search
tune.grid.tree <- expand.grid(eta = seq(.5, 1, .1),
                              max_depth = 1:3,
                              min_child_weight = seq(.25, 1, .25),
                              colsample_bytree = 1,
                              gamma = seq(.25, 1, .25),
                              nrounds = 1:4,
                              subsample = 1)

#train in parallel.
#register cluster so that caret will know to train in parallel.
cl <- makeCluster(6, type = "SOCK")
registerDoSNOW(cl)

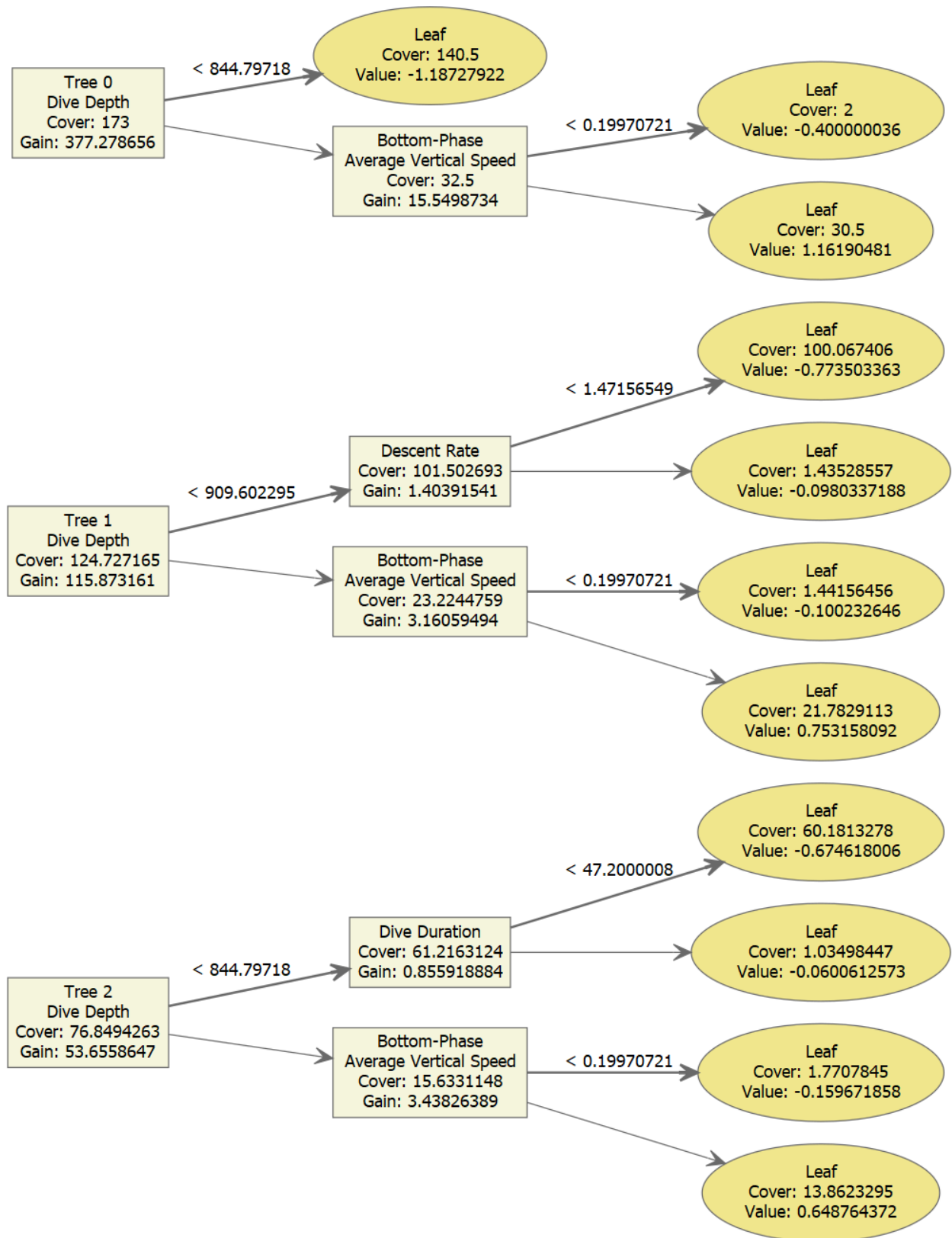
#train the xgboost model with parameters from full model
set.seed(987654321)
foragingmod.bottom.tree.npr <- train(Foraging ~ DiveDepth + DiveDuration +
                                     BottomPhaseVerticalSpeed +
                                     BottomPhaseDuration +
                                     BottomVertDirectionChanges
Prop +
                                     DescBottomJerkPeaks +
                                     DescRate + AscRate,
  data = gtdives,
  method = "xgbTree",
  tuneGrid = tune.grid.tree,
  trControl = train.control,
  metric = "Accuracy")

stopCluster(cl)
```



```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Yes  No
##           Yes 122  0
##           No   3 567
##
##           Accuracy : 0.9957
##           95% CI : (0.9874, 0.9991)
##           No Information Rate : 0.8194
##           P-Value [Acc > NIR] : <2e-16
##
##           Kappa : 0.9852
##
## McNemar's Test P-Value : 0.2482
##
##           Sensitivity : 0.9760
##           Specificity : 1.0000
##           Pos Pred Value : 1.0000
##           Neg Pred Value : 0.9947
##           Prevalence : 0.1806
##           Detection Rate : 0.1763
##           Detection Prevalence : 0.1763
##           Balanced Accuracy : 0.9880
##
##           'Positive' Class : Yes
##
## nrounds max_depth eta gamma colsample_bytree min_child_weight subsa
mple
## 287      3          2 0.6  0.5                1                1
1
```

Figure S3

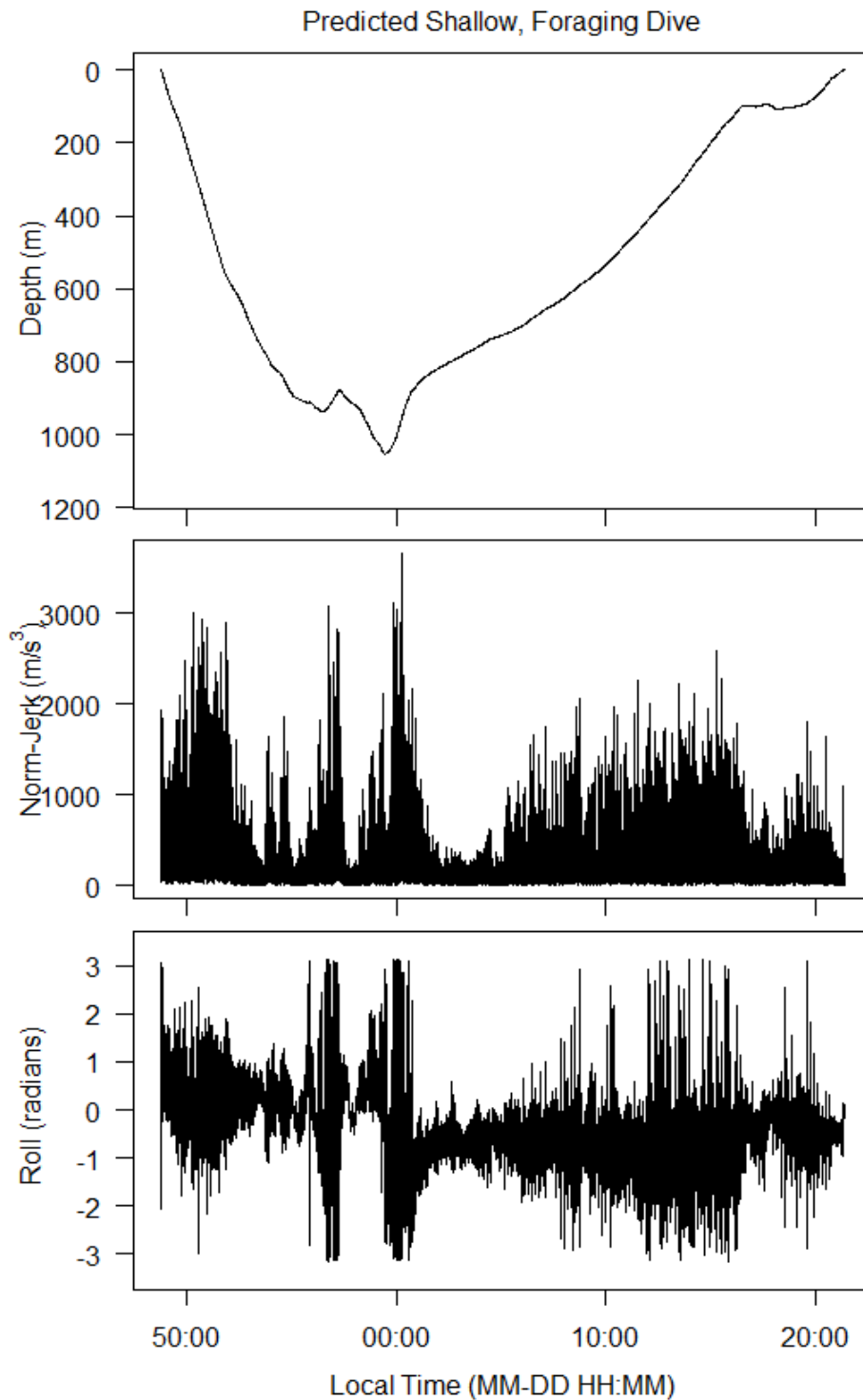


Supplementary Figure 3: Diagram showing optimized decision trees from the comprehensive model without the inclusion of orientation-dependent predictors. Rectangles represent decision nodes from which a set of arrows point to guide dive

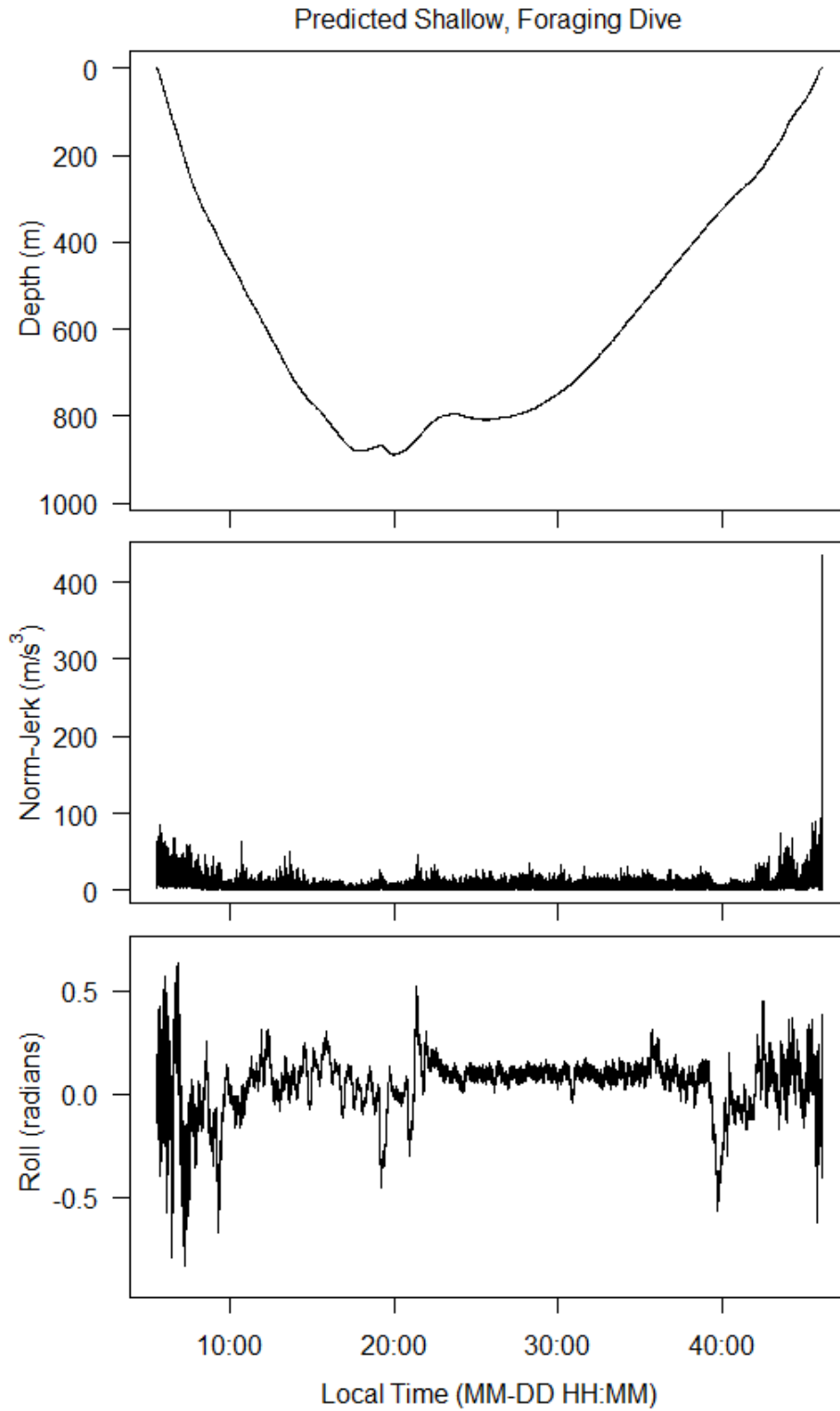
classifications along towards resulting leaves (yellow ellipses) at the end of each decision tree. When classifying each dive via these four decision trees, the top arrow coming from each node is followed if the condition (shown above each top arrow) is met for the variable listed within the node. Otherwise, the bottom arrow is followed. Cover is the sum of the second order gradient of training data classified to the leaf. Gain is a quantification representing the information gained from a split (thus corresponding to the importance of the node in the model). Value represents the marginal value that the leaf may contribute to predictions (positive values contribute to foraging dive classifications and negative to non-foraging classifications). Optimized hyperparameters for the comprehensive model were: nrounds = 3, max_depth = 2, eta = 0.6, gamma = 0.5, colsample_bytree = 1, min_child_weight = 1, and subsample = 1. Variable units for decision thresholds are as follows: dive depth (m), dive duration (min), descent rate (m/s), bottom-phase average vertical speed (m/s).

Dives for Which Model and K-Means Classifications Differ

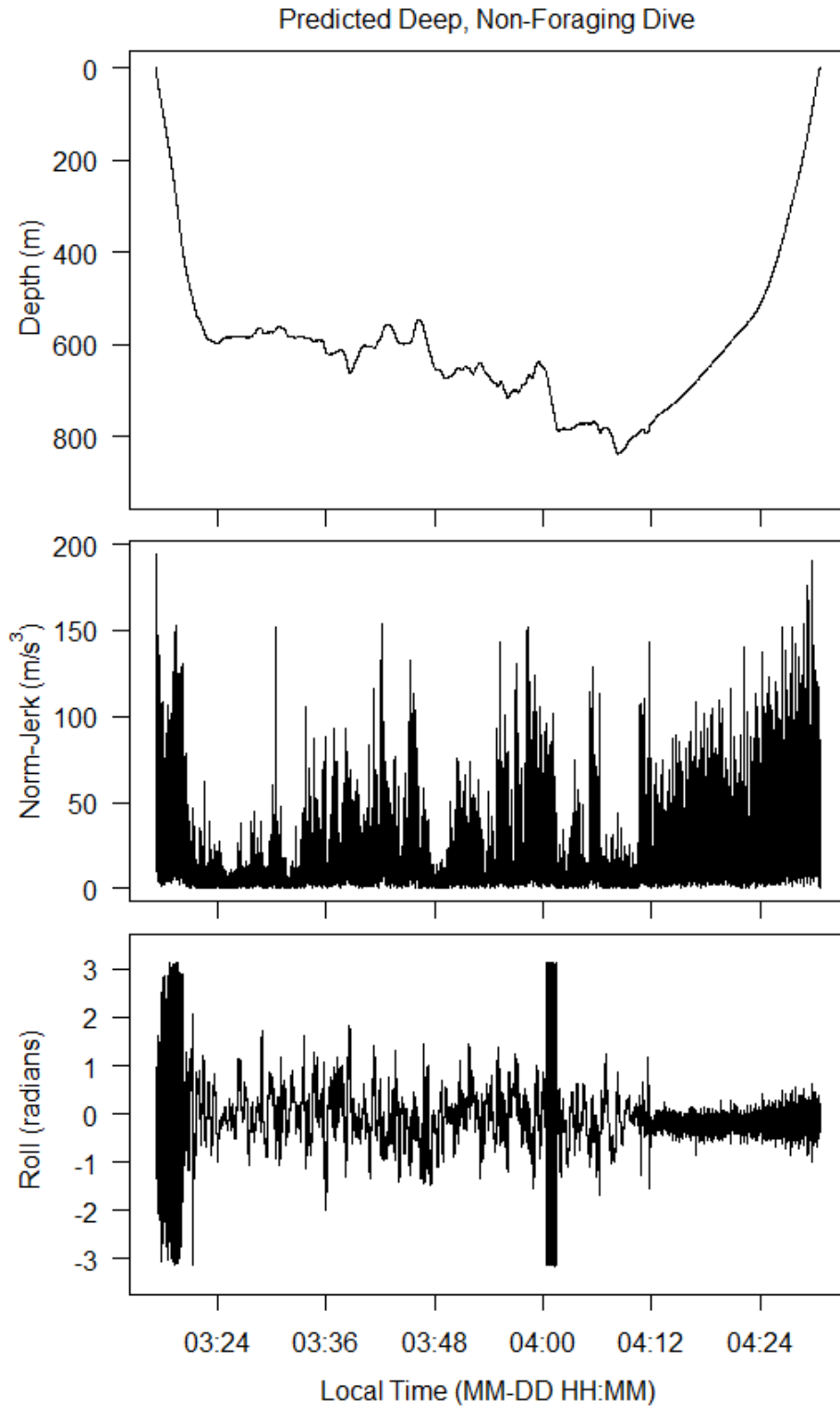
Figures S4-S23 (figure legend below last figure)



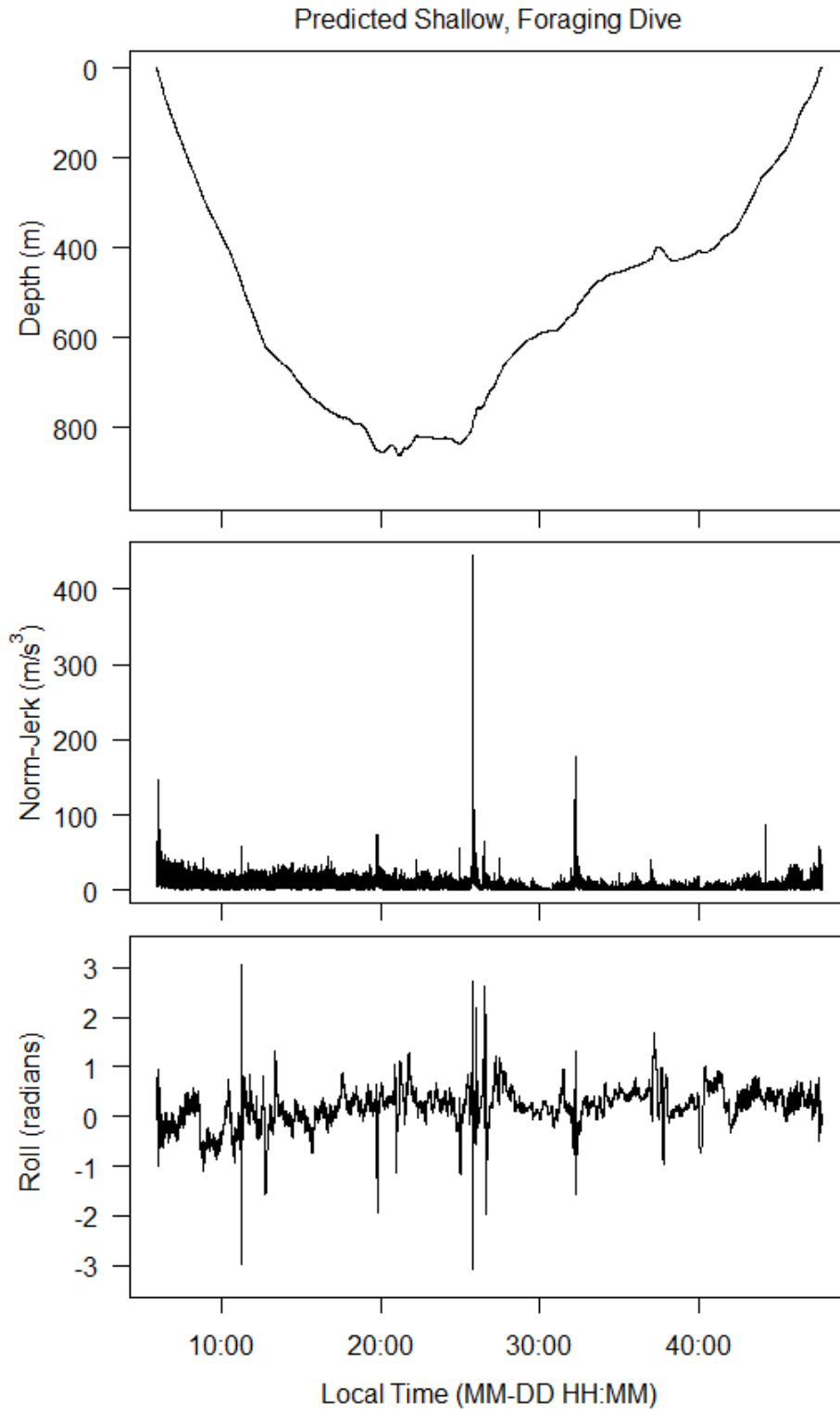
Supplementary Figure 4



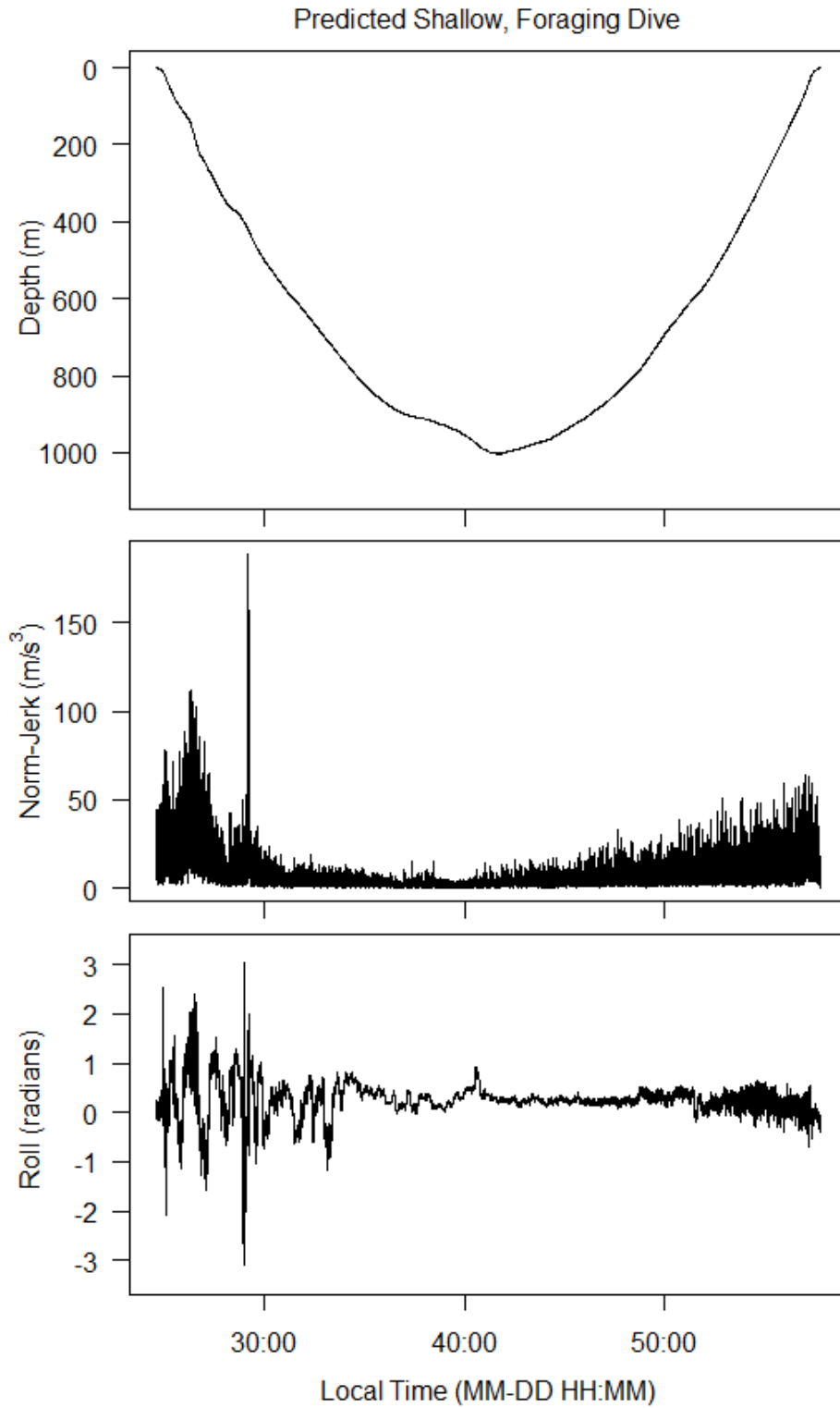
Supplementary Figure 5



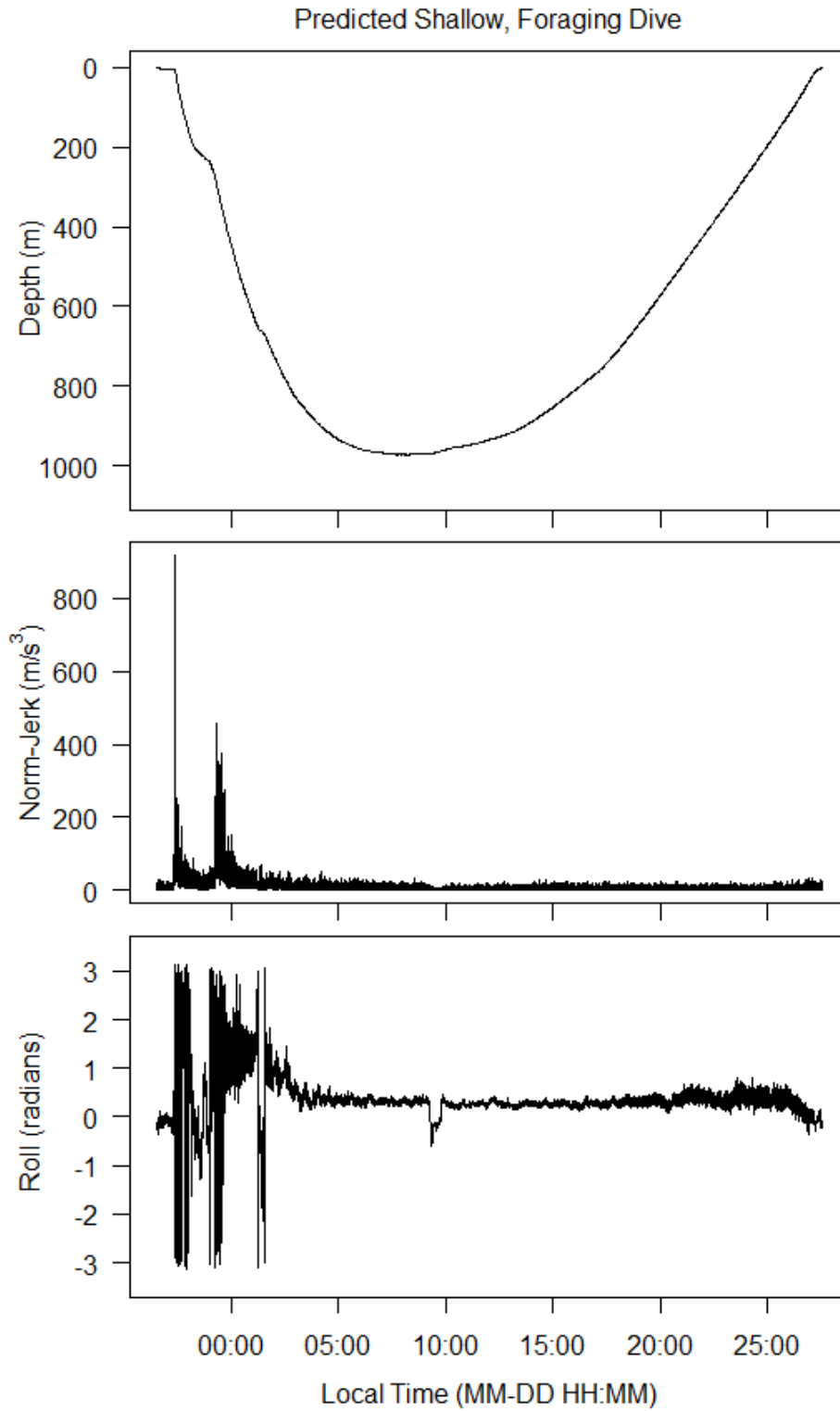
Supplementary Figure 6



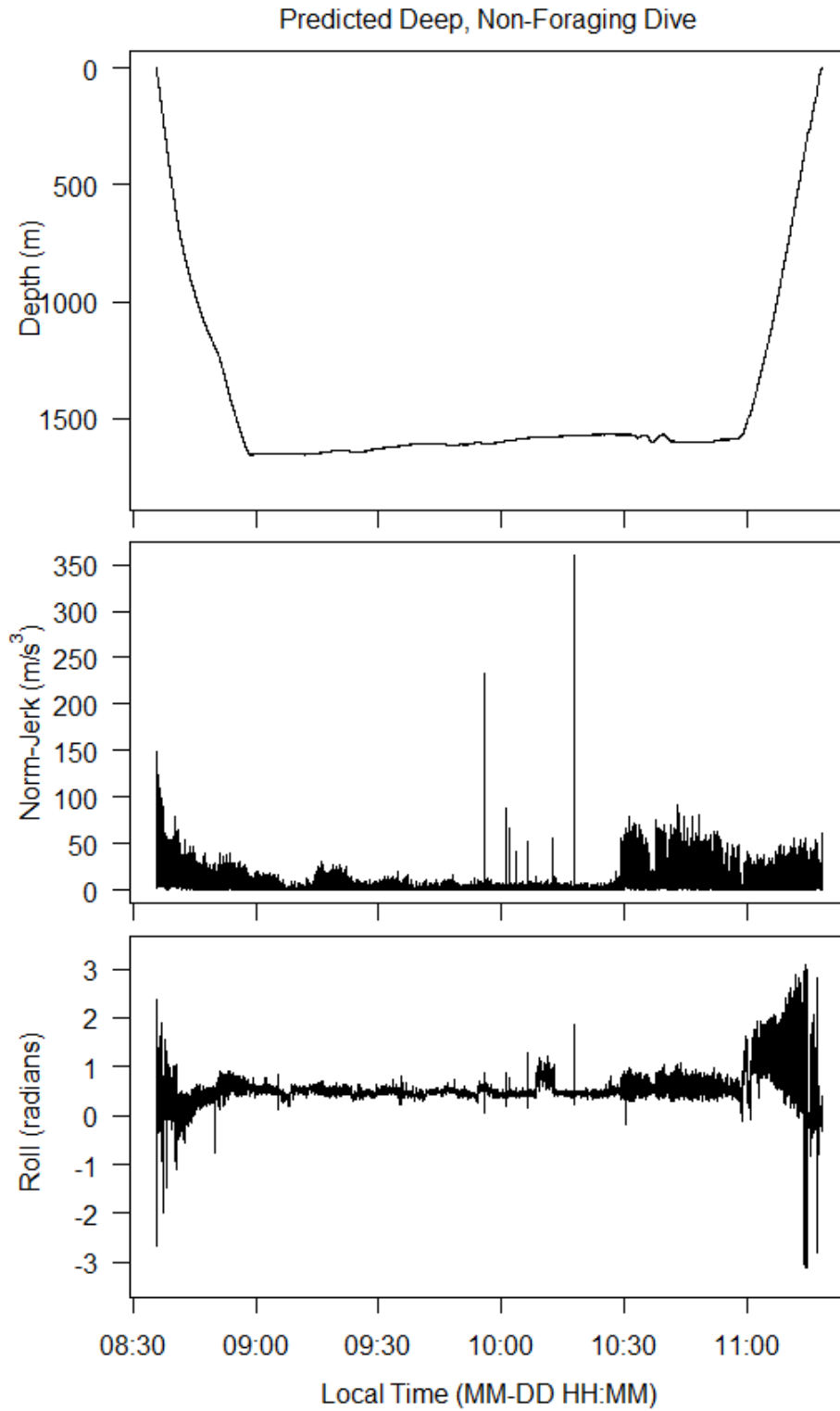
Supplementary Figure 7



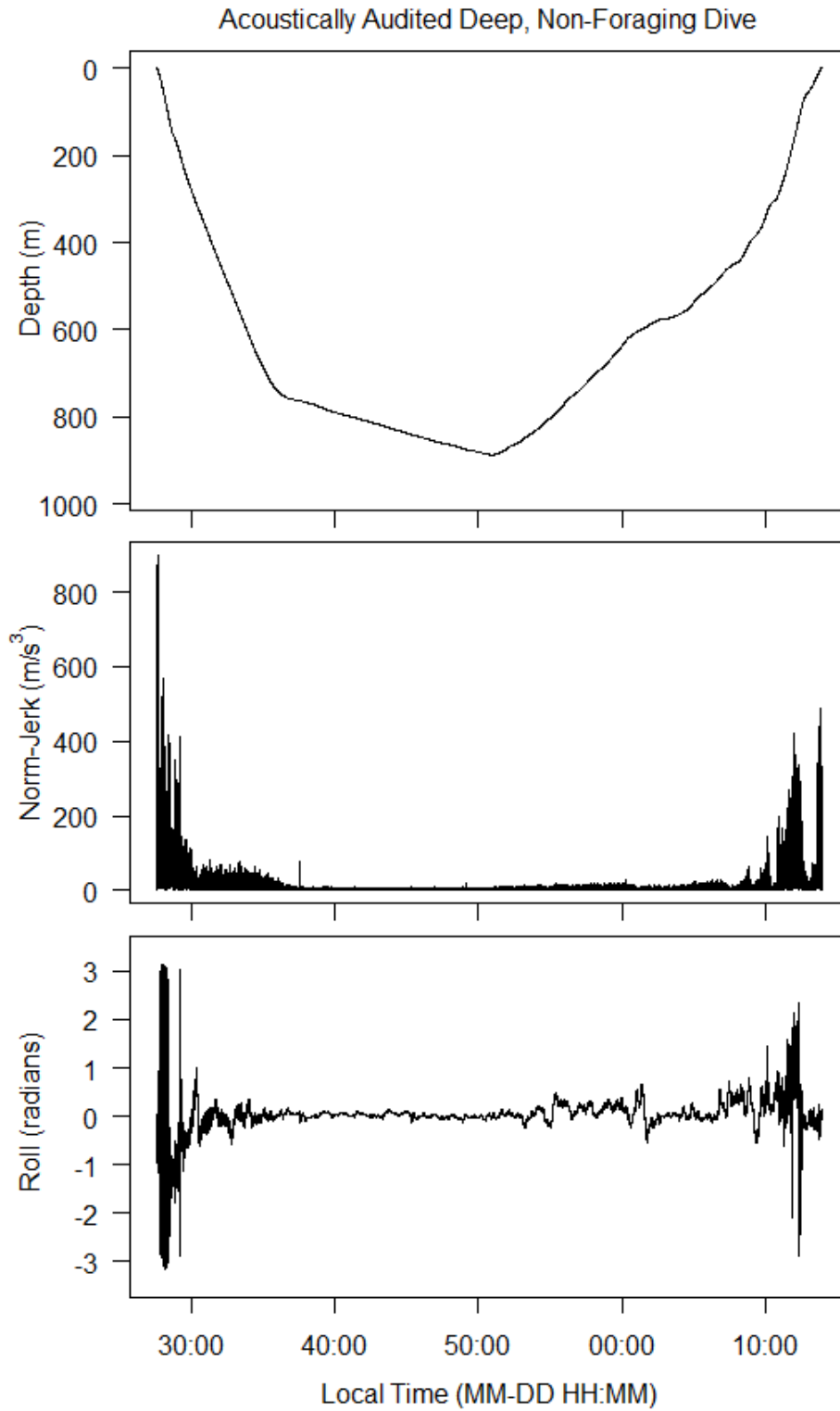
Supplementary Figure 8



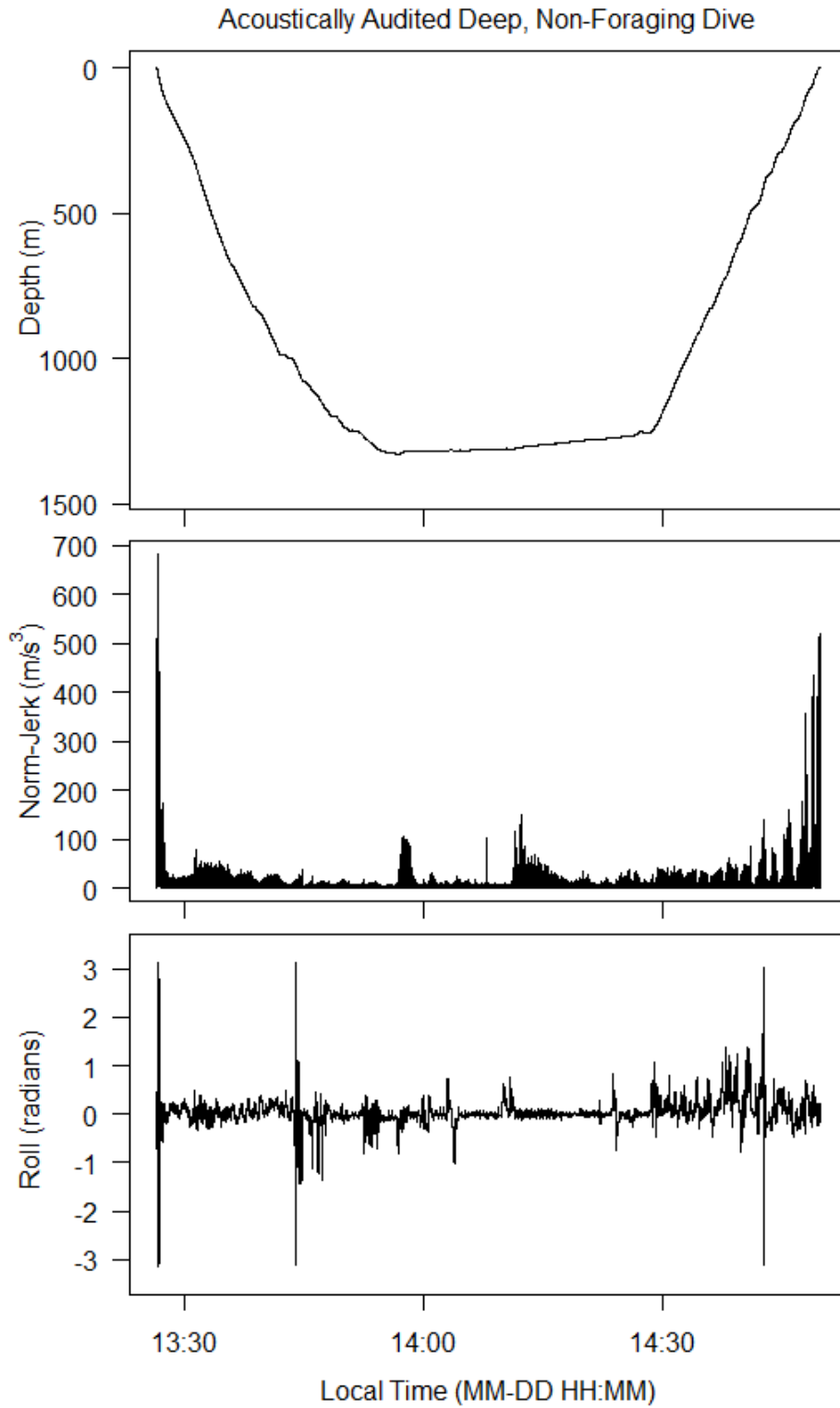
Supplementary Figure 9



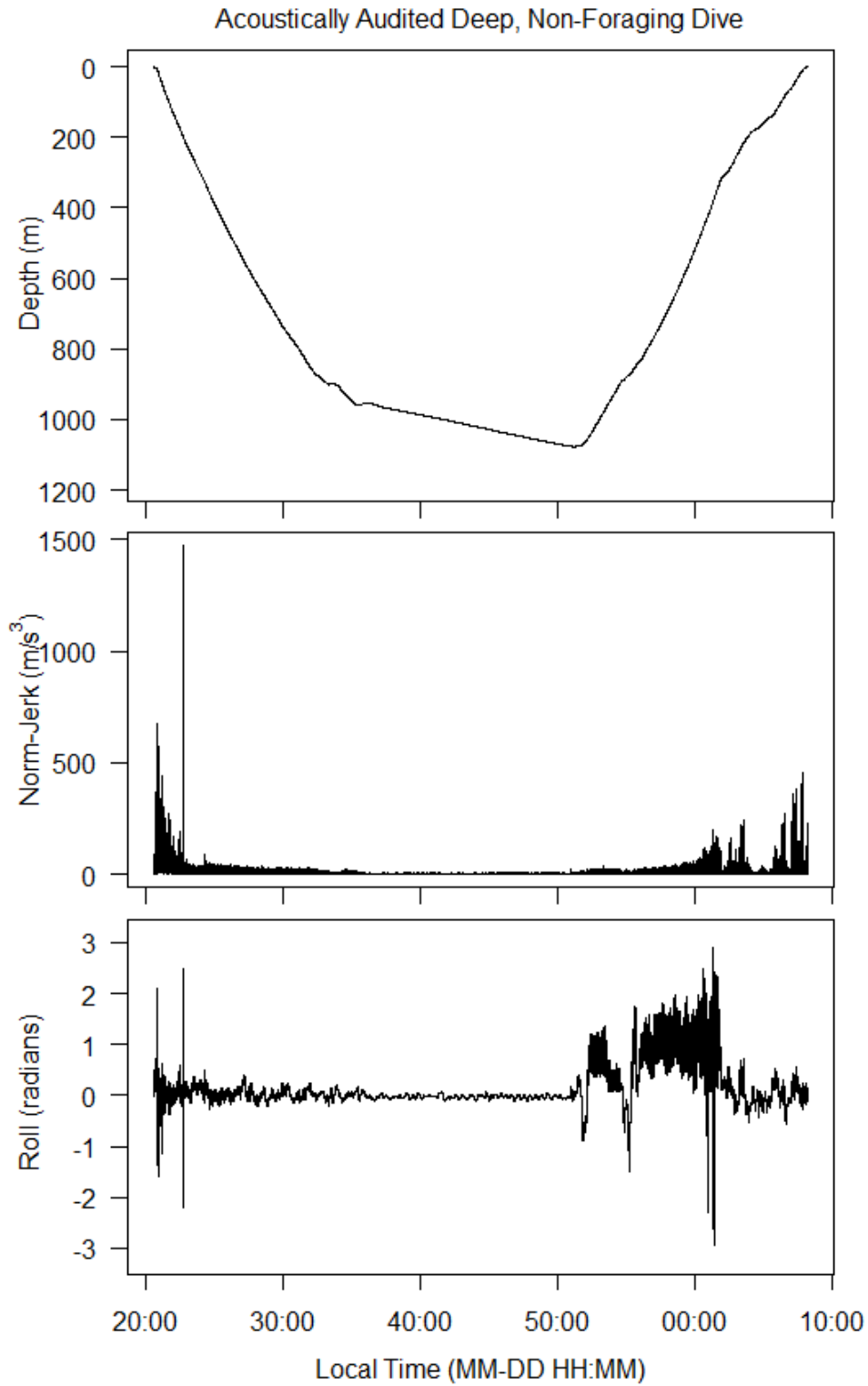
Supplementary Figure 10



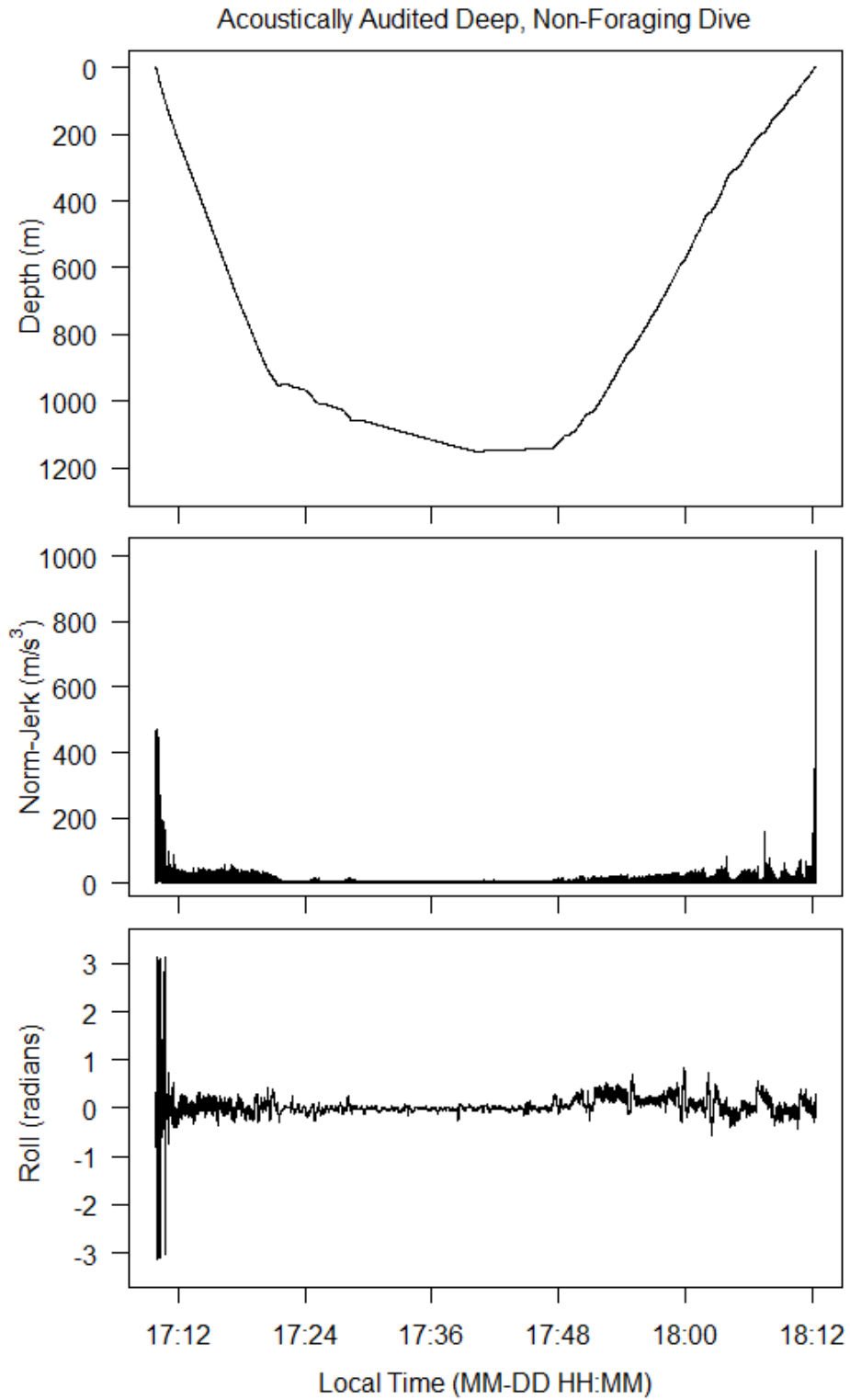
Supplementary Figure 11



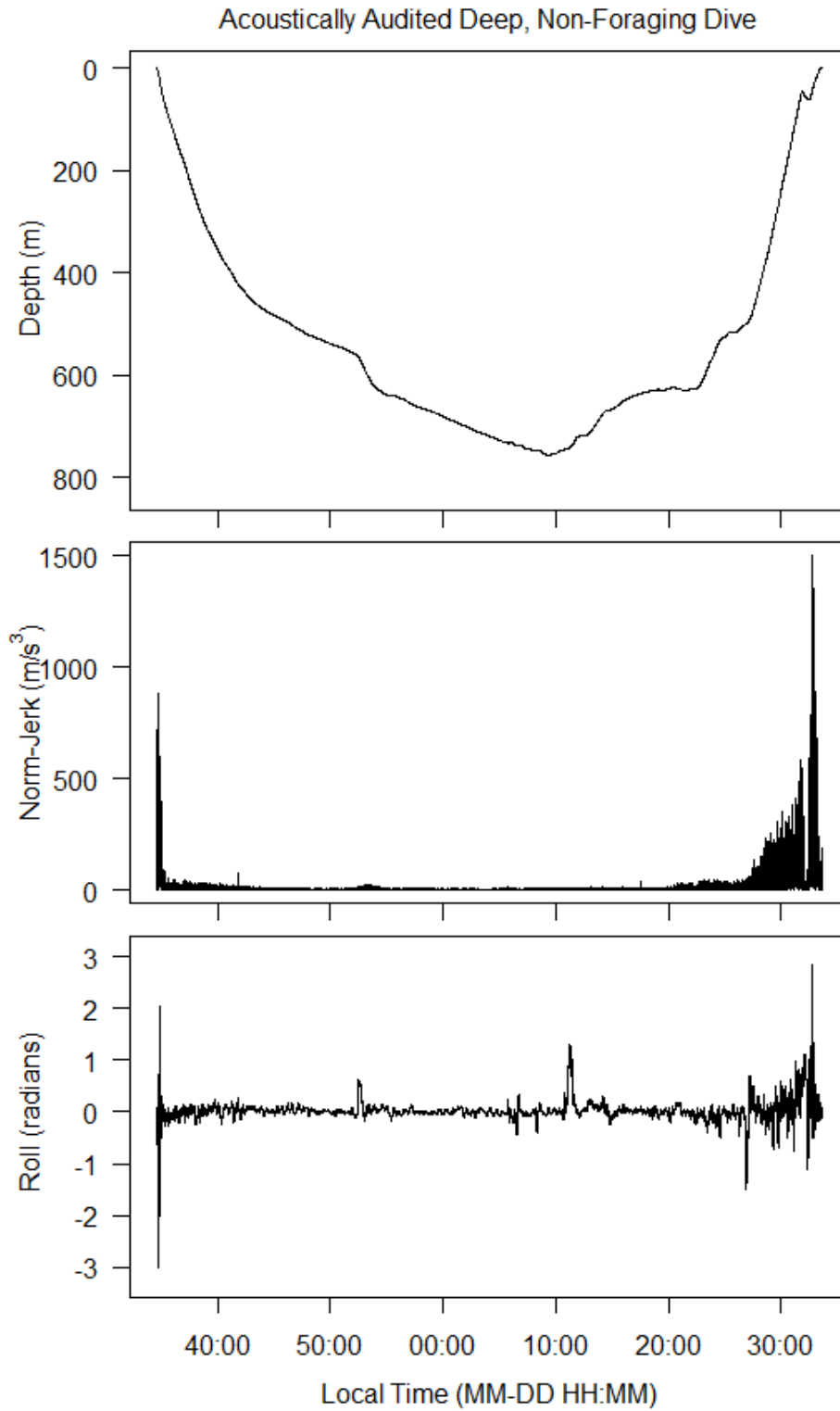
Supplementary Figure 12



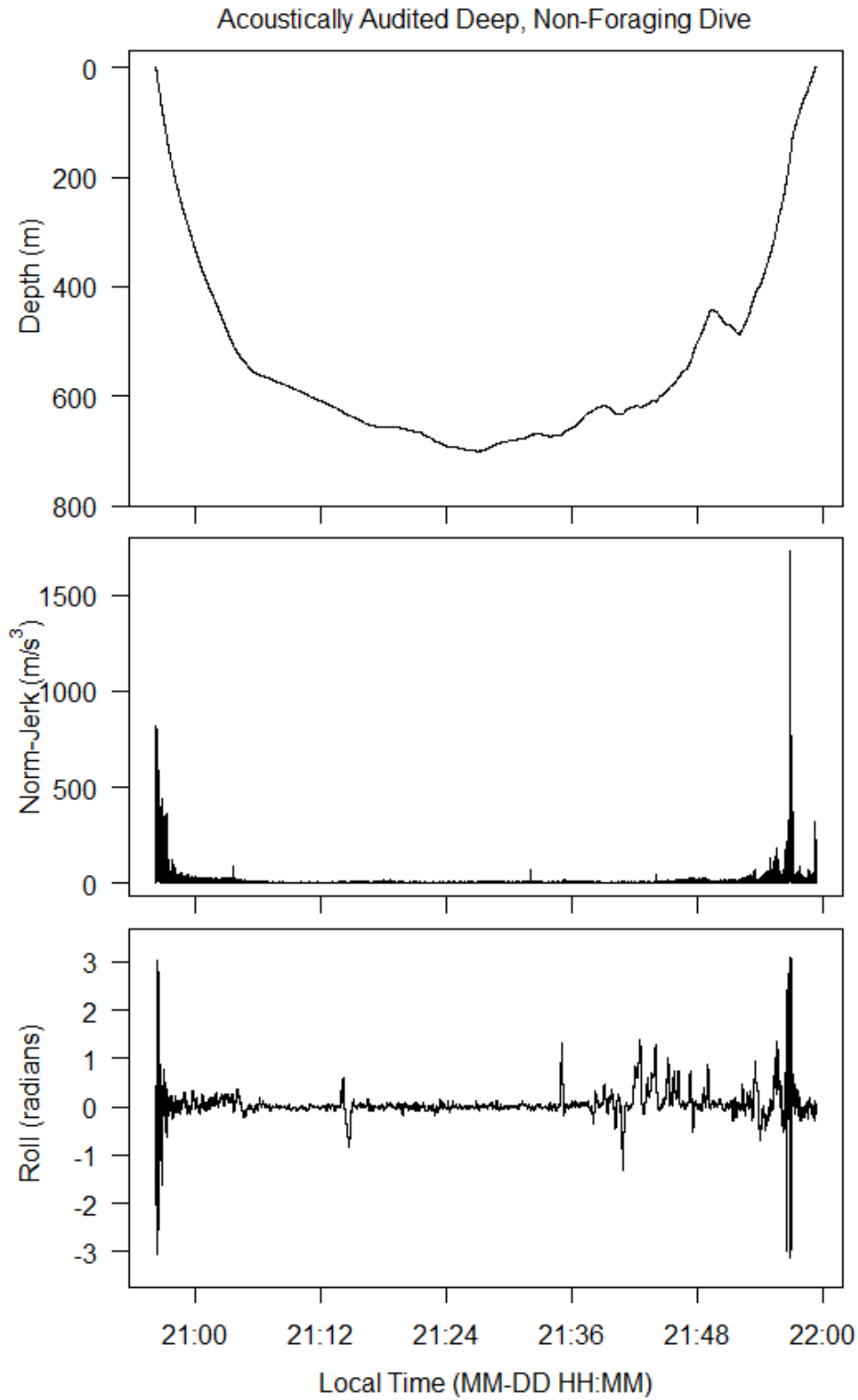
Supplementary Figure 13



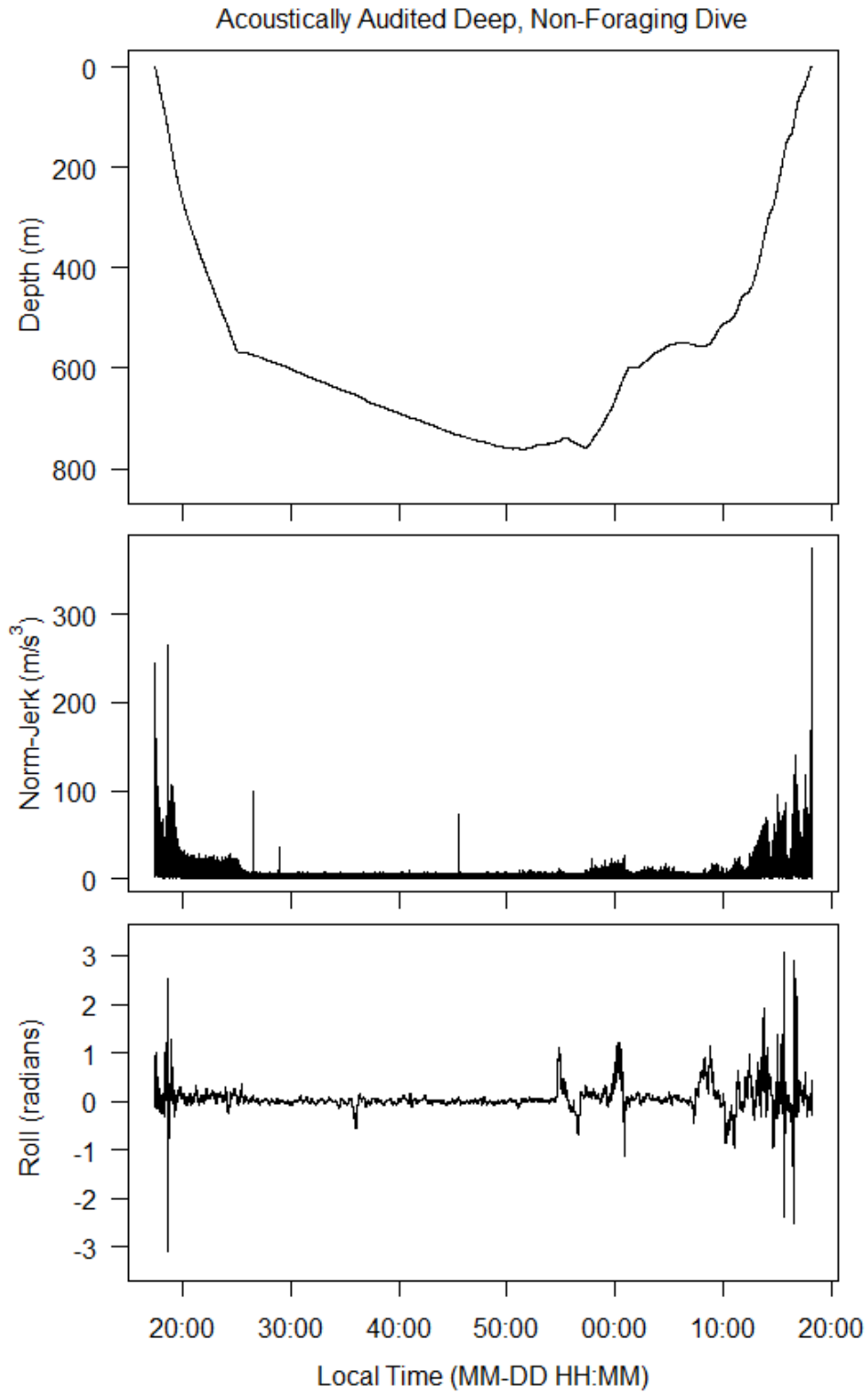
Supplementary Figure 14



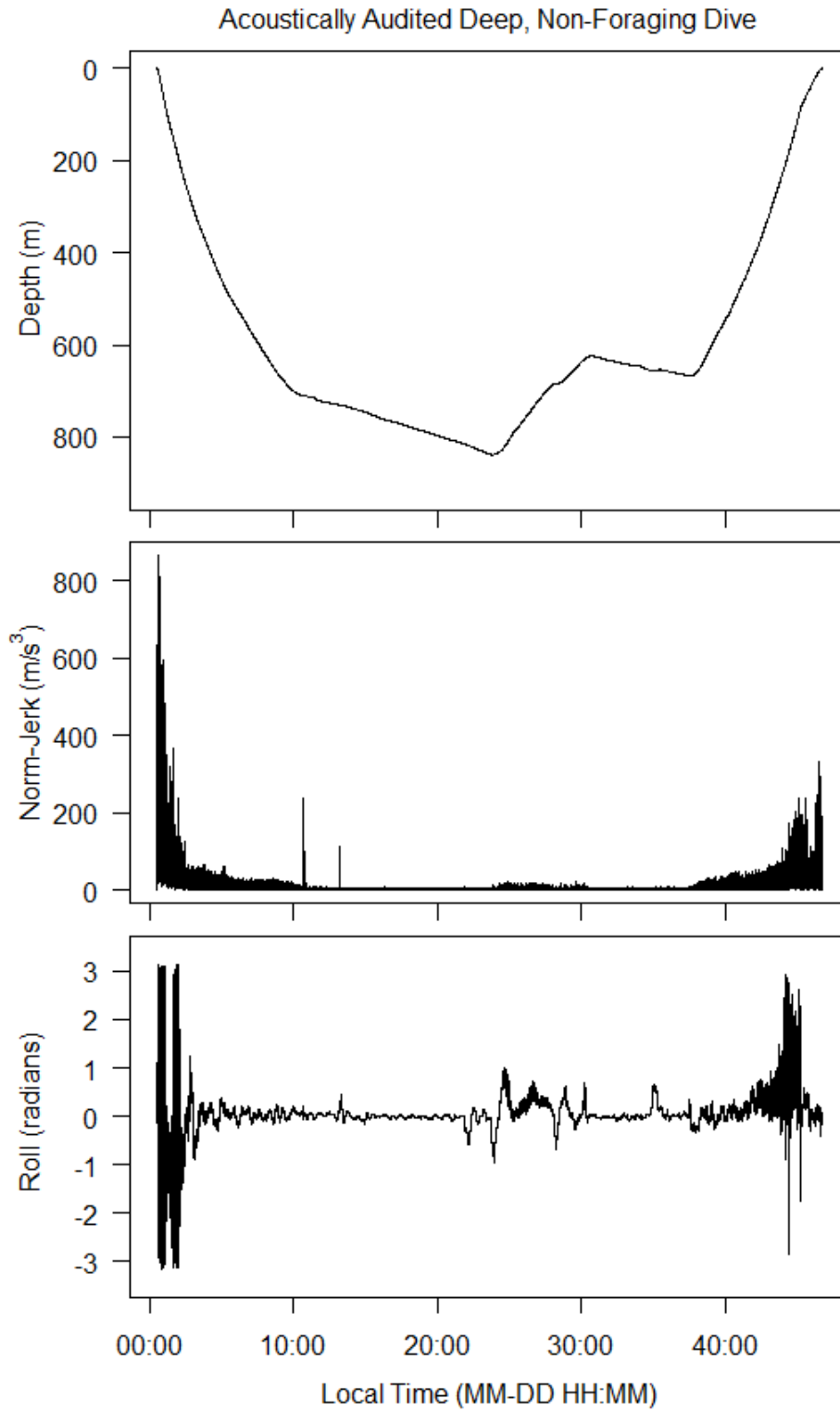
Supplementary Figure 15



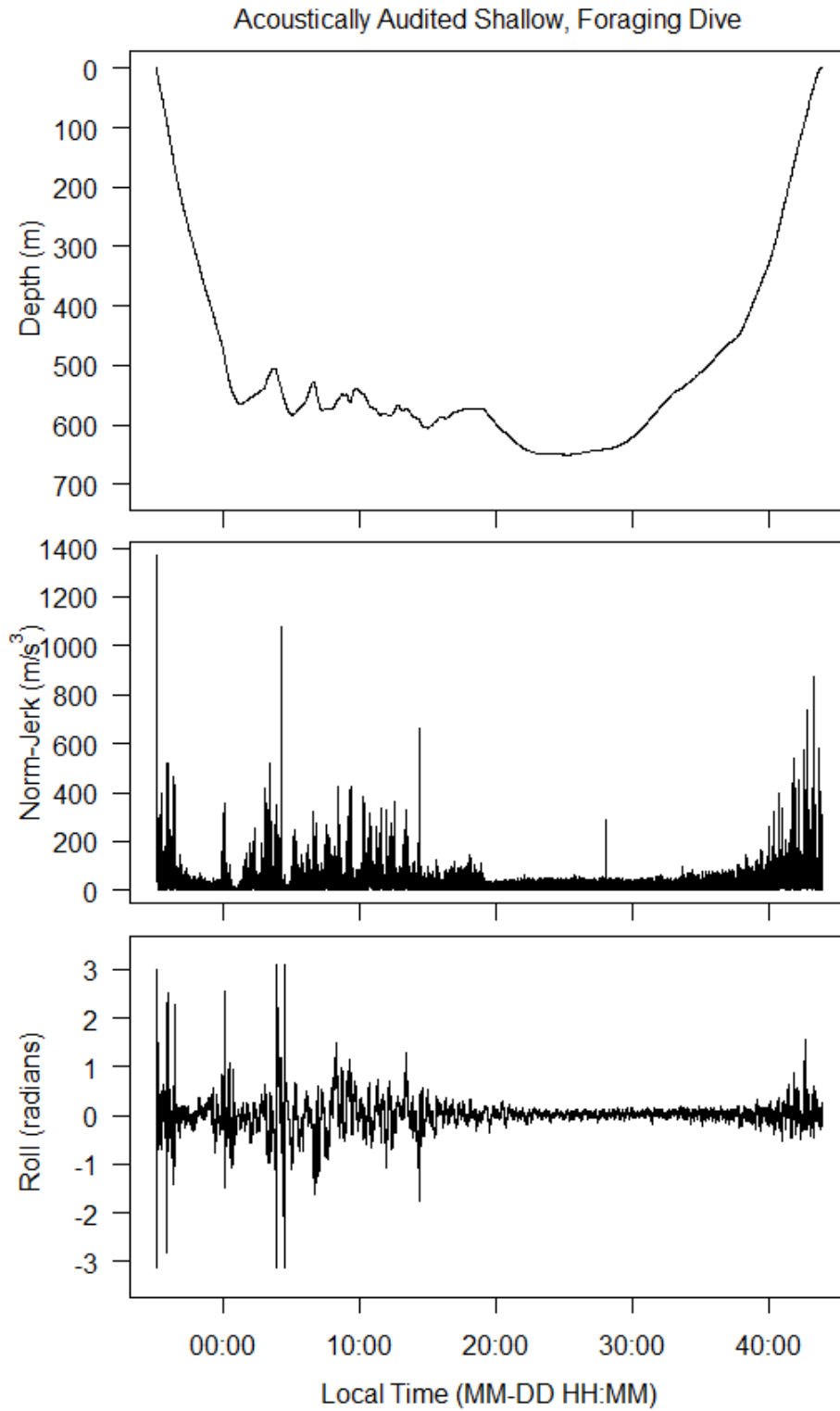
Supplementary Figure 16



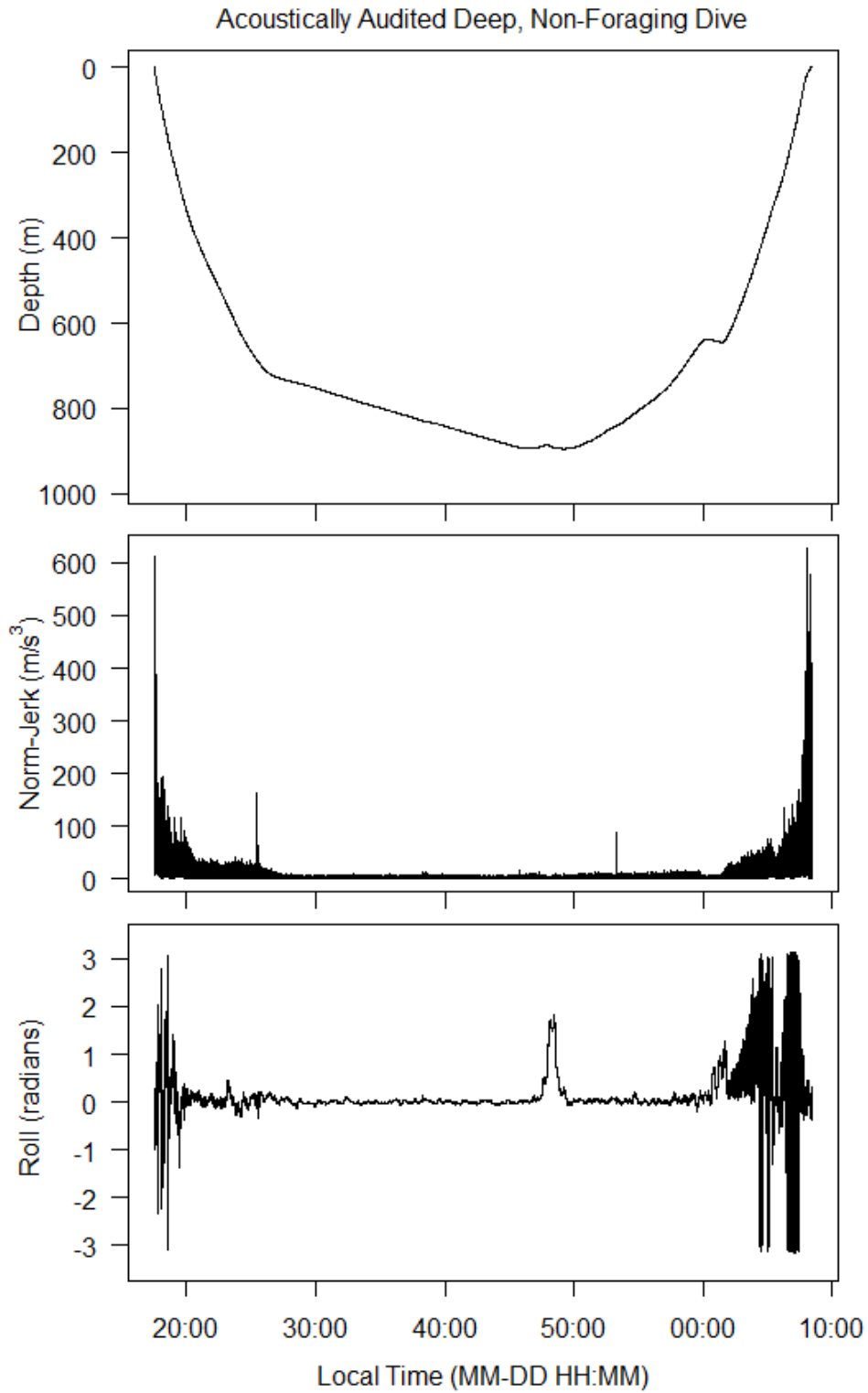
Supplementary Figure 17



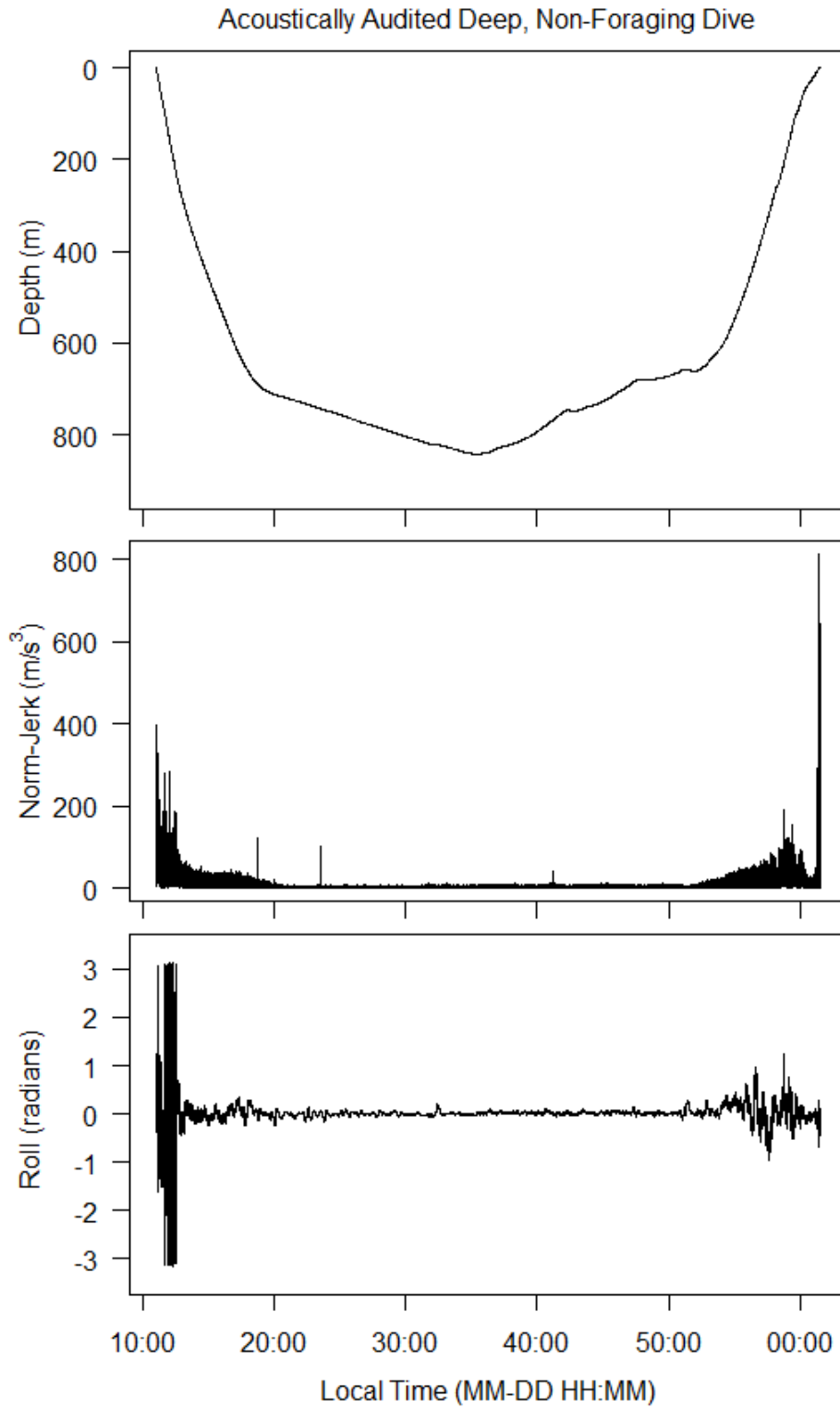
Supplementary Figure 18



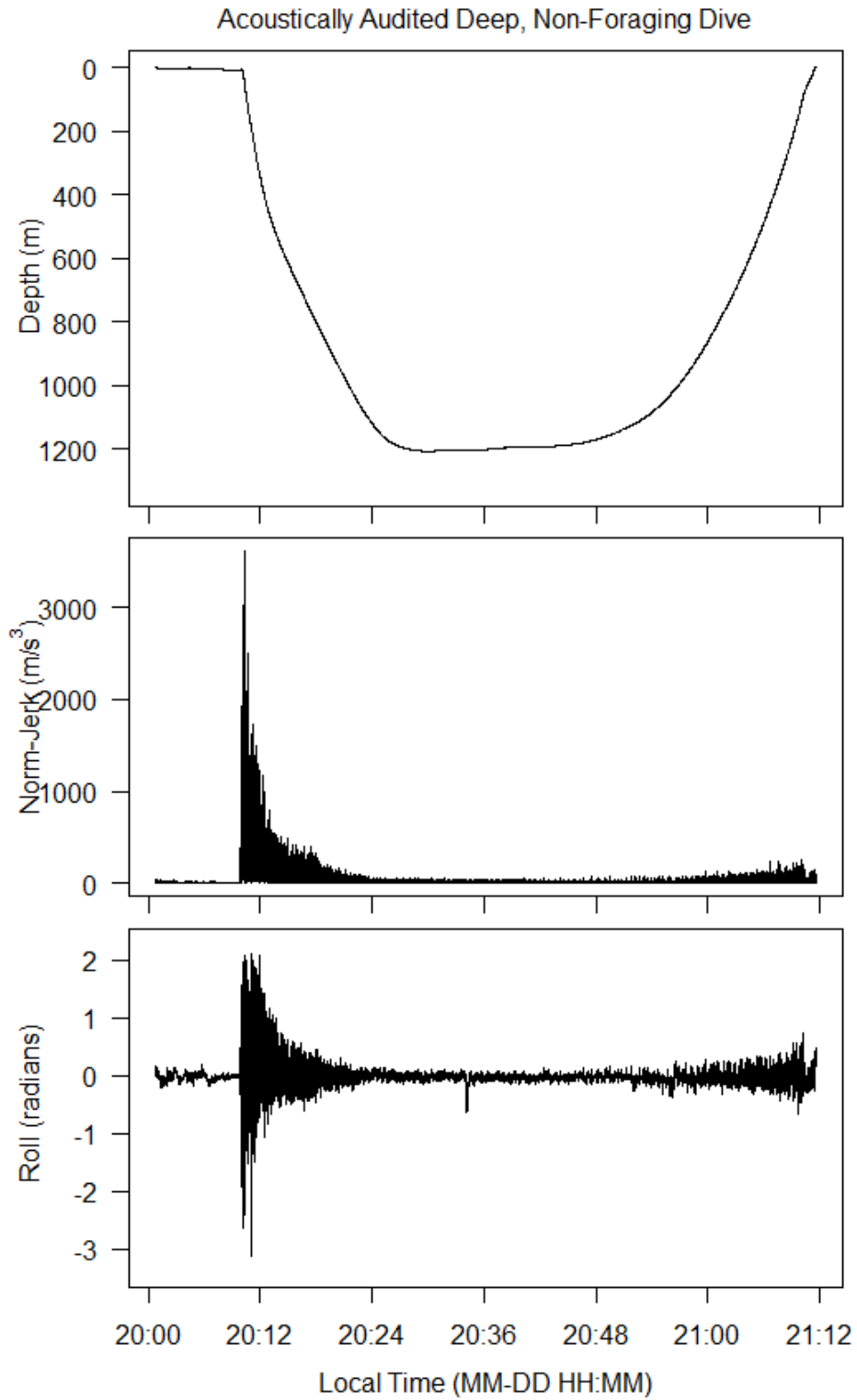
Supplementary Figure 19



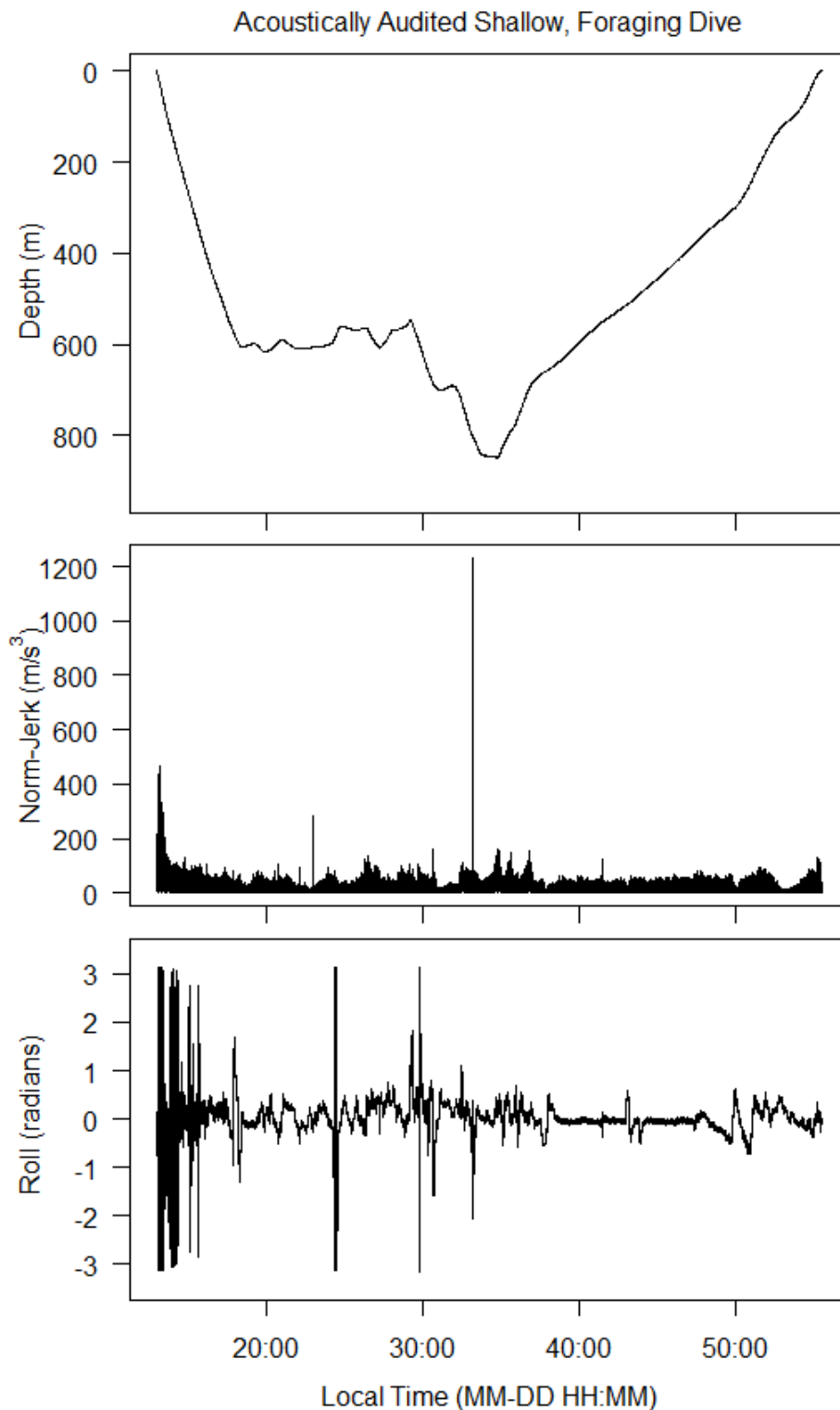
Supplementary Figure 20



Supplementary Figure 21



Supplementary Figure 22



Supplementary Figure 23: Supplementary Figures S4-S23 show dive profiles of the 20 dives for which K-means classification differed from foraging predictions from the comprehensive model. Each respective figure is labeled with its K-means classification, comprehensive model prediction, and whether it was acoustically audited or not. Note that the norm-jerk and roll columns show the raw data prior to being decimated/resampled to 16Hz.