3. We will compare SVM with an RBF kernel and LSSVM in Matlab. To do so, we will run a Monte Carlo simulation of 100 runs of each LSSVM and SVM. Our test/train split will be 75%, i.e. we will withhold a random 25% of our data for prediction and misclassification rate. With our 100 misclassification rates for each classifier, we can create a box plot displaying their distribution. Furthermore, we will create an ROC curve for each model to show each’s sensitivity. Note: the hyperparameters for SVM will be optimized through 10-fold cross validation.

From the results, we can see that the misclassification rate for LSSVM is slightly lower with an average of about 0.065%, versus about 0.072% for SVM. Additionally, the ROC for LSSVM has an AUC of 0.99535 versus 0.82972 for SVM, indicating that LSSVM is much less sensitive. As such, we can conclude that LSSVM is the better choice for the classifying the ozon data.