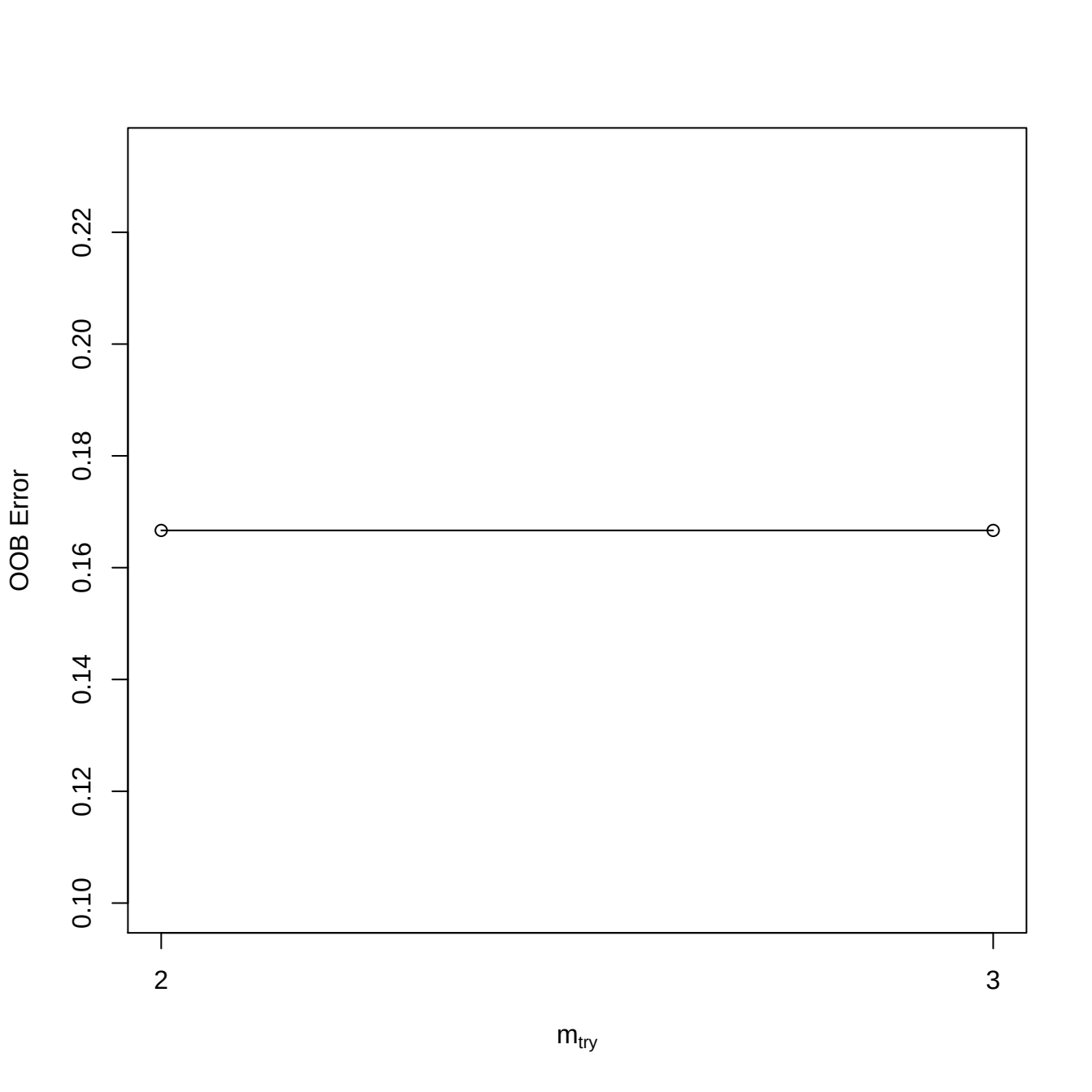
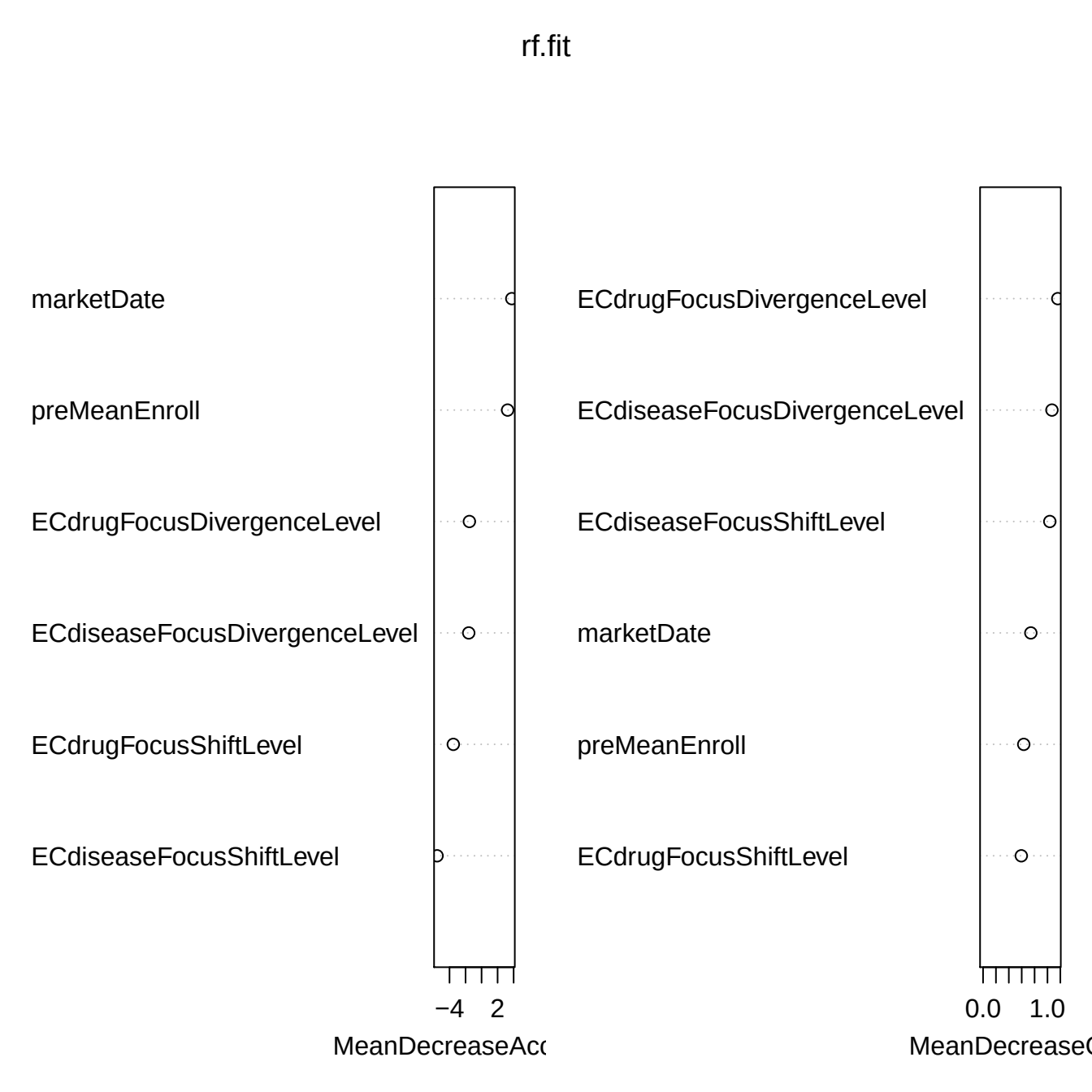
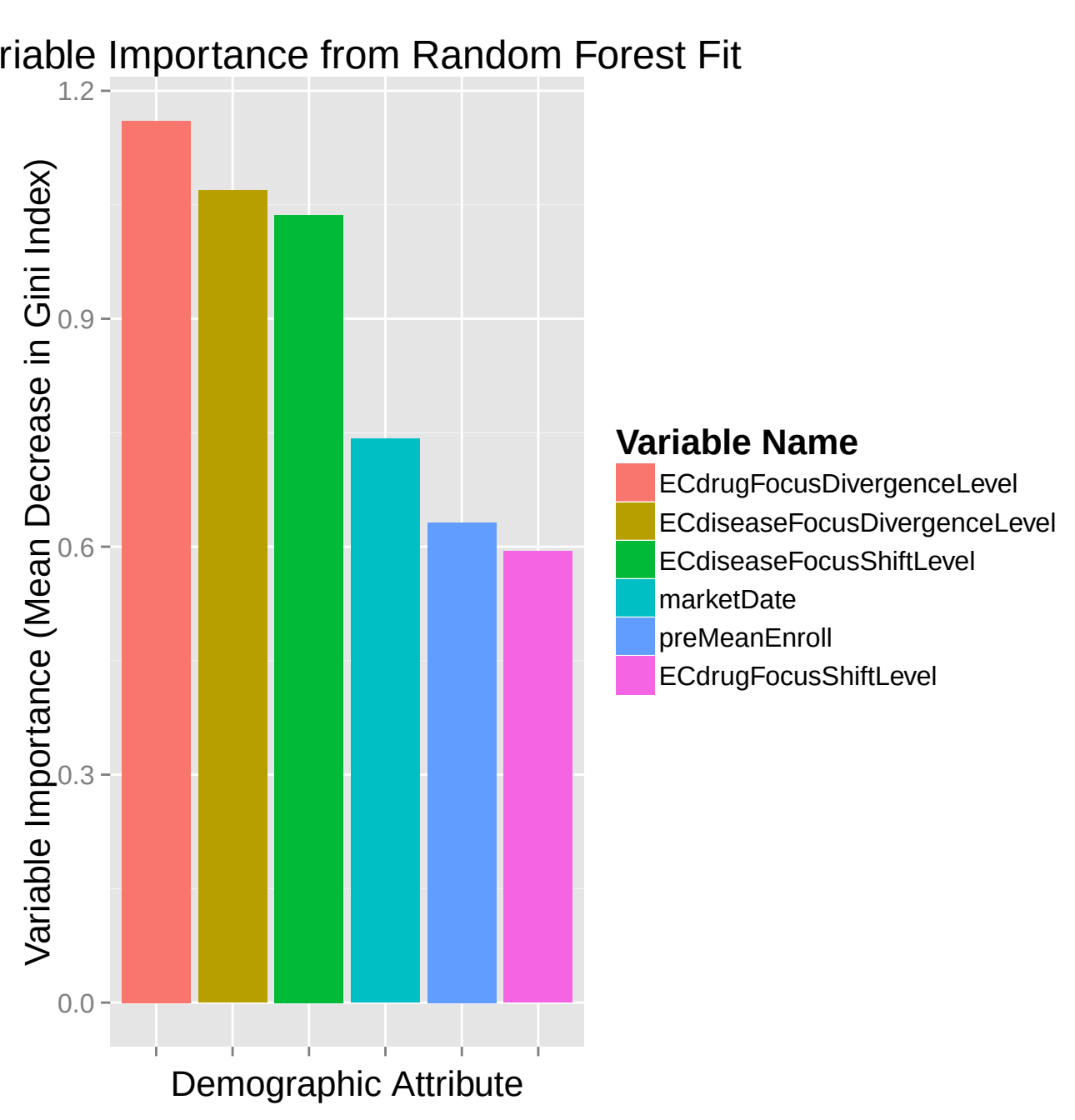
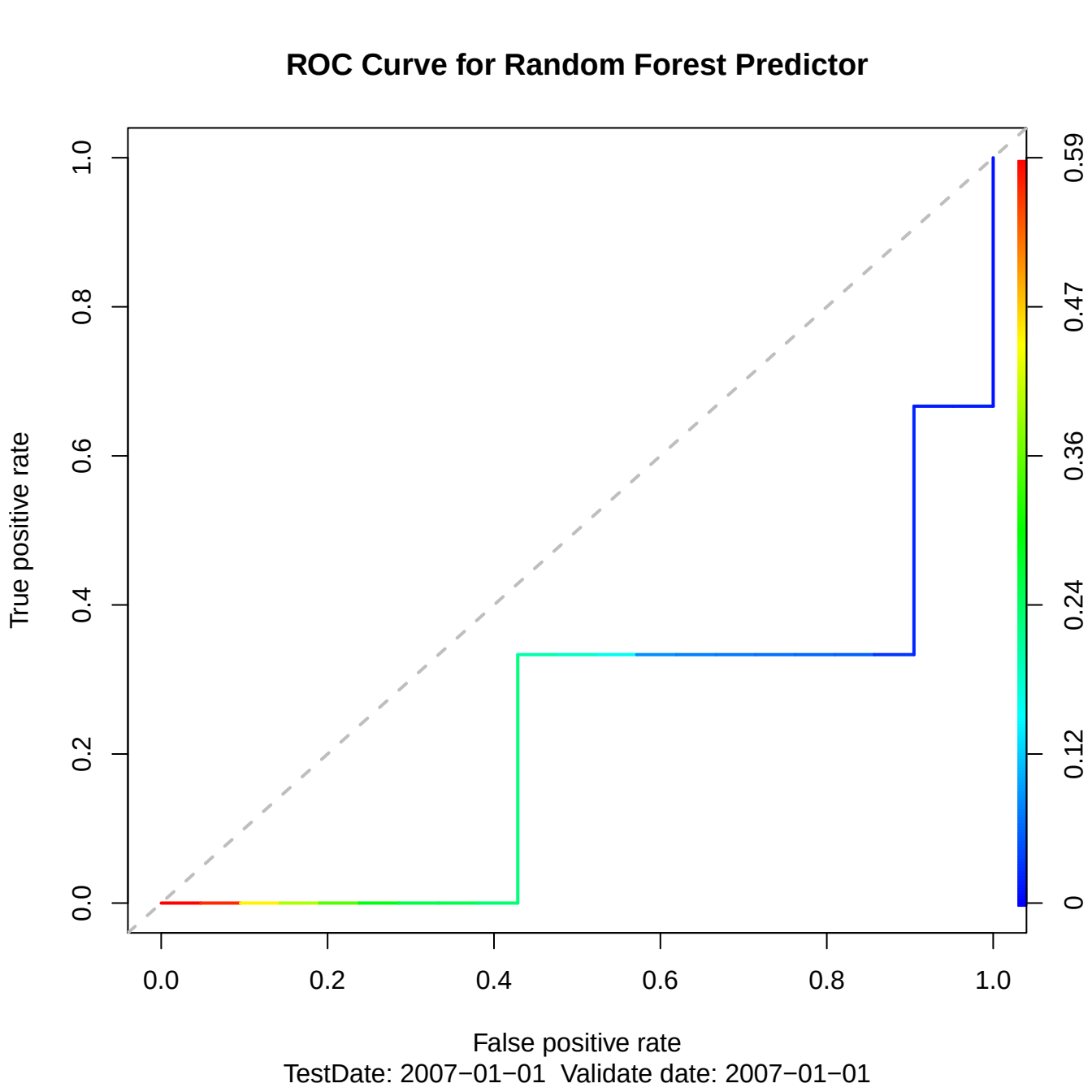

OOB Error
0.10
0.12
0.14
0.16
0.18
0.20
0.22
2
3
m_{try}

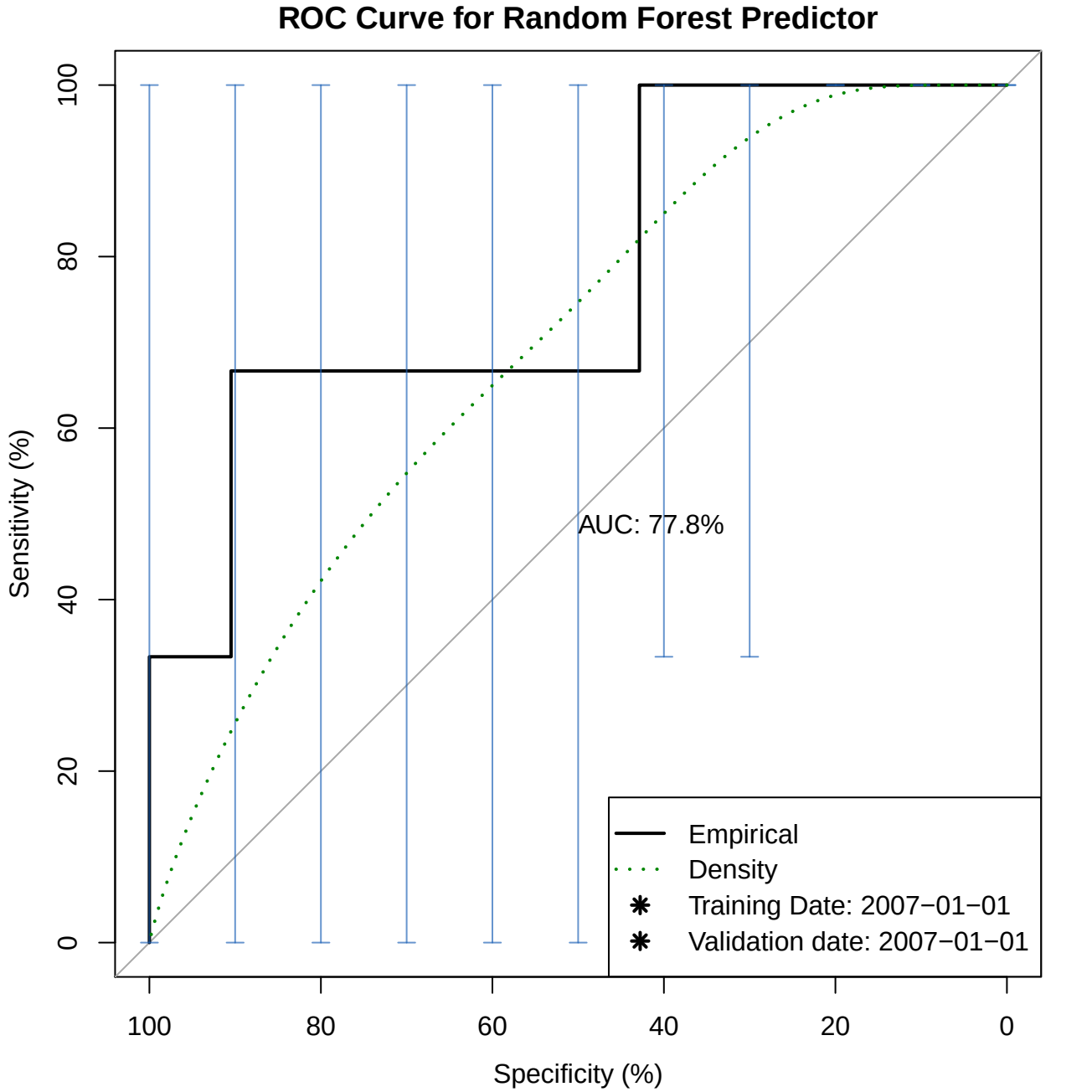

rf.fit
marketDate
preMeanEnroll
ECdrugFocusDivergenceLevel
ECdiseaseFocusDivergenceLevel
ECdrugFocusShiftLevel
ECdiseaseFocusShiftLevel
−4
2
MeanDecreaseAcc
ECdrugFocusDivergenceLevel
ECdiseaseFocusDivergenceLevel
ECdiseaseFocusShiftLevel
marketDate
preMeanEnroll
ECdrugFocusShiftLevel
0.0
1.0
MeanDecreaseG

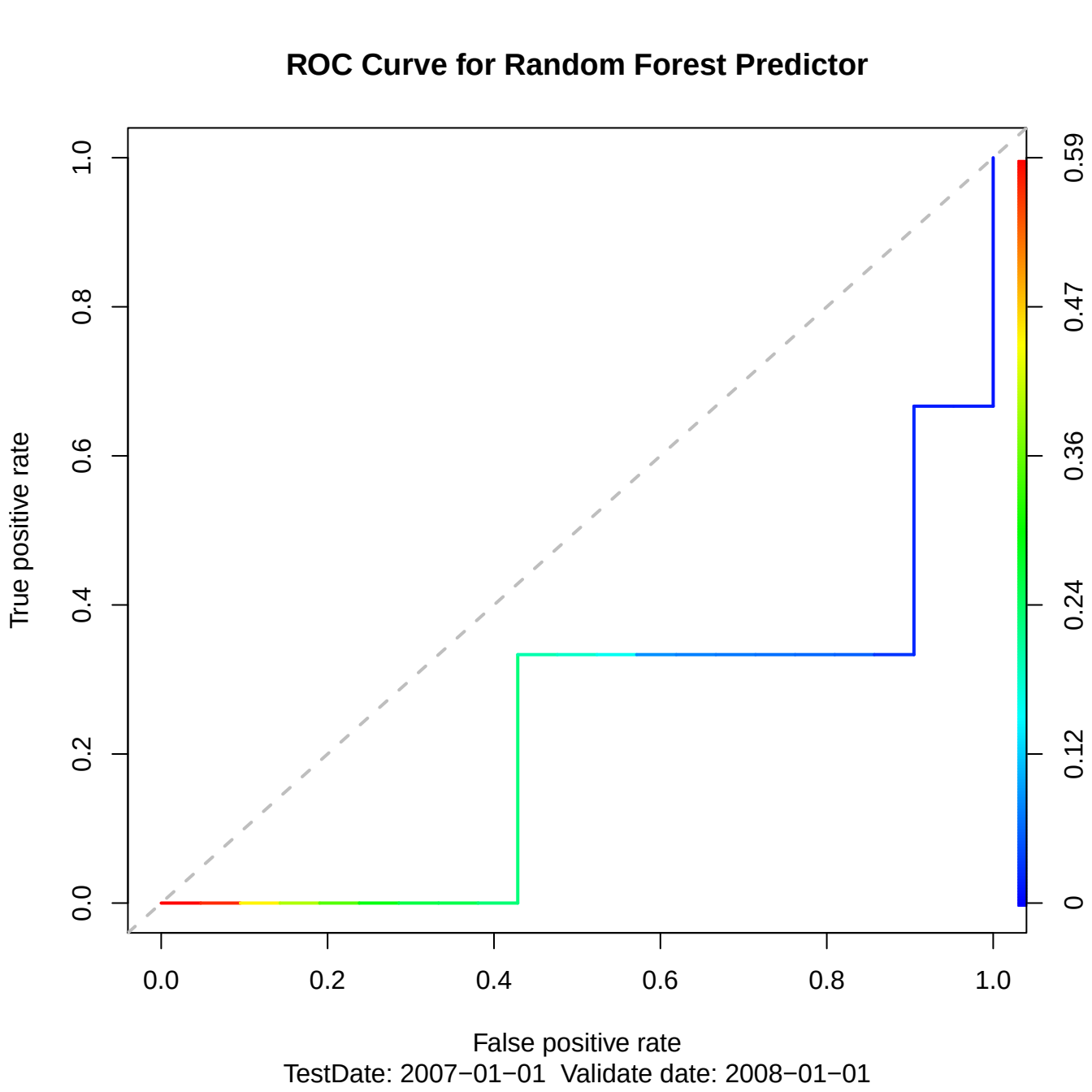
	MeanDecreaseAccuracy	MeanDecreaseGini
marketDate	3.77415236777845	0.741862414318381
preMeanEnroll	3.25046603467431	0.631759214297541
ECdiseaseFocusDivergenceLevel	−1.6023312129689	1.06925405451662
ECdiseaseFocusShiftLevel	−5.55493093144753	1.03644652527182
ECdrugFocusDivergenceLevel	−1.52032961219381	1.16049815065396
ECdrugFocusShiftLevel	−3.52631796869043	0.593929640941685

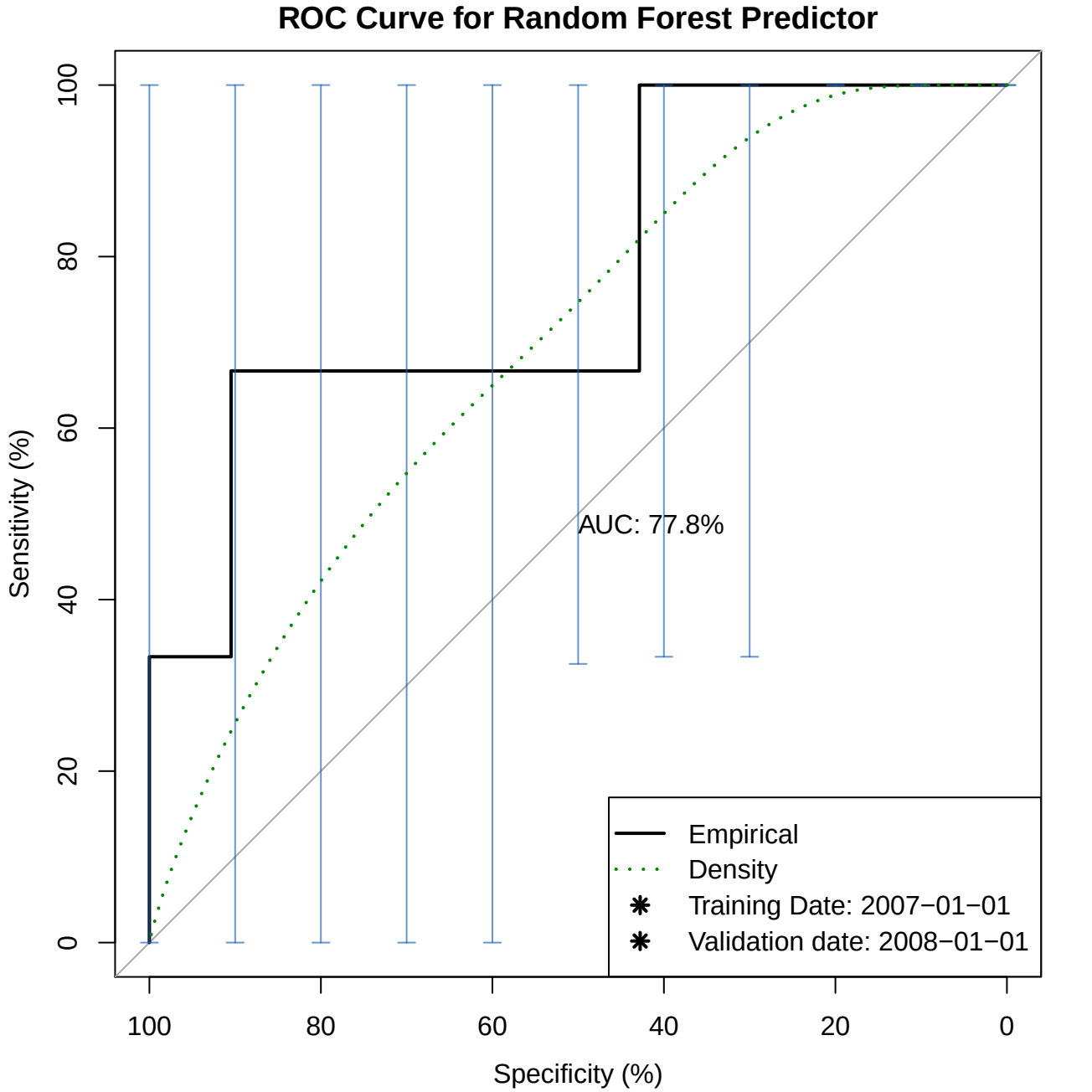

riable Importance from Random Forest Fit
Variable Importance (Mean Decrease in Gini Index)
1.2
0.9
0.6
0.3
0.0
Demographic Attribute
Variable Name
ECdrugFocusDivergenceLevel
ECdiseaseFocusDivergenceLevel
ECdiseaseFocusShiftLevel
marketDate
preMeanEnroll
ECdrugFocusShiftLevel

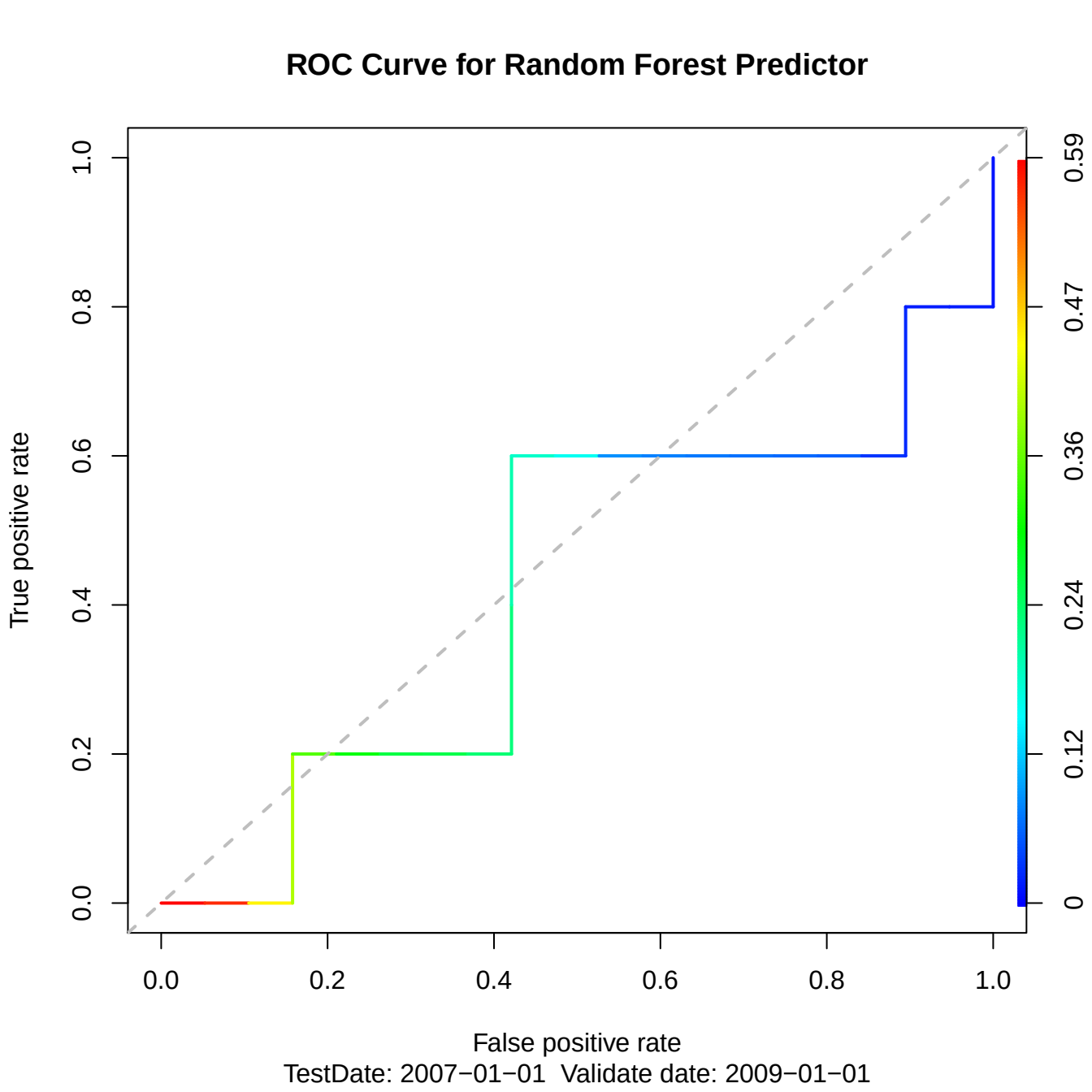
	0	1	MeanDecreaseAccuracy	MeanDecreaseG
marketDate	3.51071616168414	1.23854556090151	3.77415236777845	0.7418624143183
preMeanEnroll	4.05469076302061	−1.66678995450696	3.25046603467431	0.6317592142975
ECdiseaseFocusDivergenceLevel	−2.21176962771227	0.0867113229354418	−1.6023312129689	1.069254054516
ECdiseaseFocusShiftLevel	−4.88320106816811	−4.00445923636101	−5.55493093144753	1.036446525271
ECdrugFocusDivergenceLevel	−1.45961699717998	−1.0696563746014	−1.52032961219381	1.160498150653
ECdrugFocusShiftLevel	−2.89806738879407	−2.60123178137872	−3.52631796869043	0.5939296409416

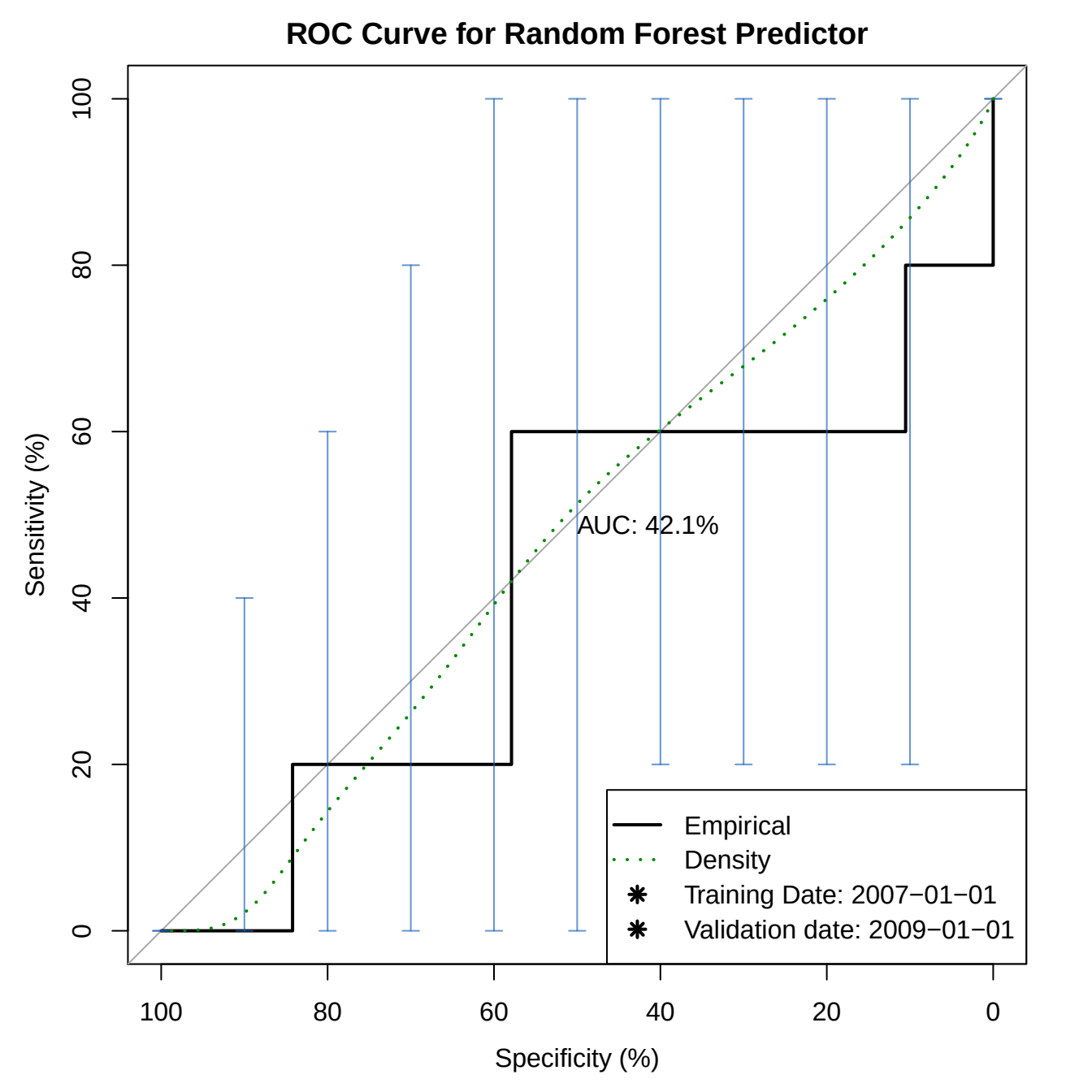

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
TestDate: 2007−01−01 Validate date: 2007−01−01

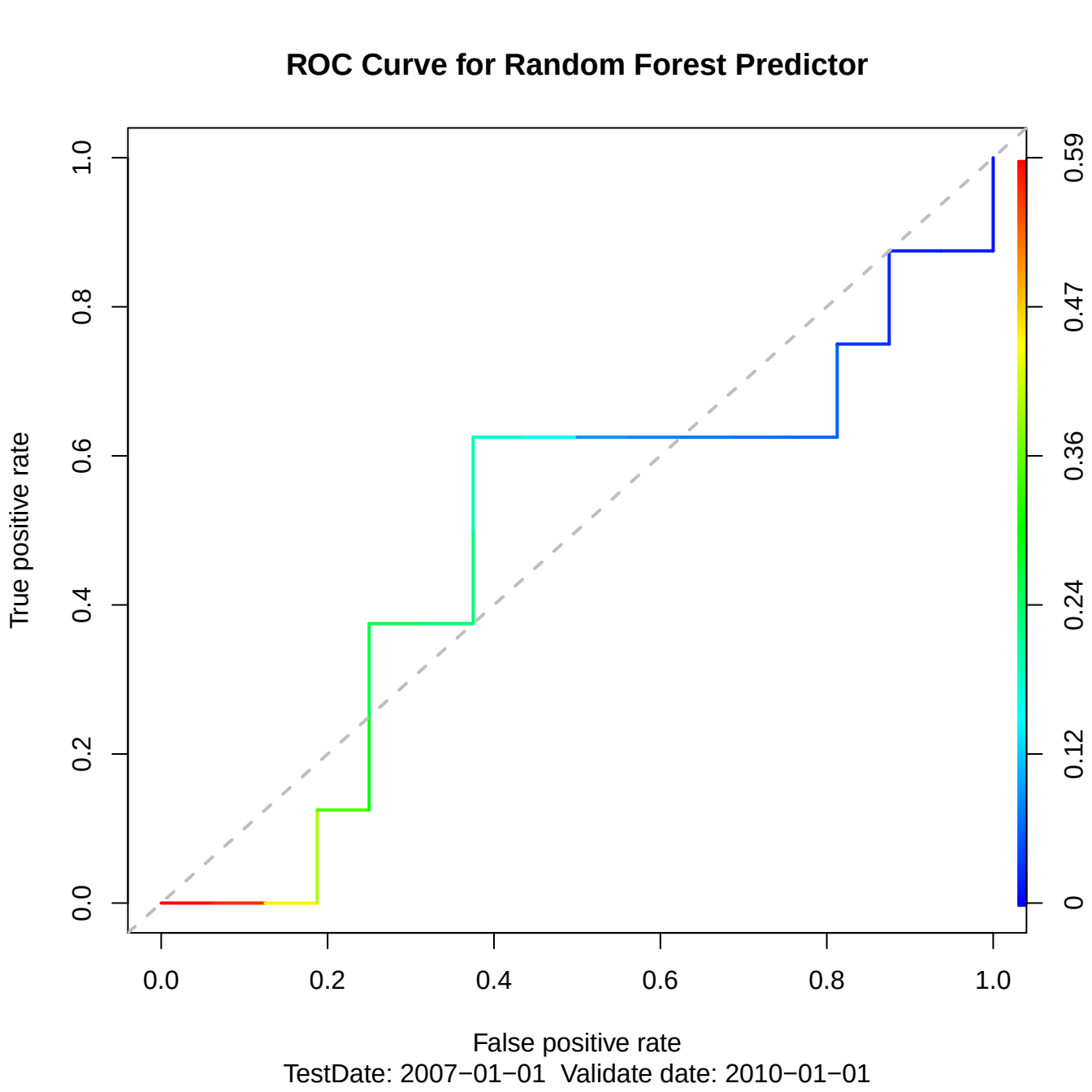

ROC Curve for Random Forest Predictor
Sensitivity (%)
100
80
60
40
20
0
Specificity (%)
100
80
60
40
20
0
AUC: 77.8%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2007−01−01

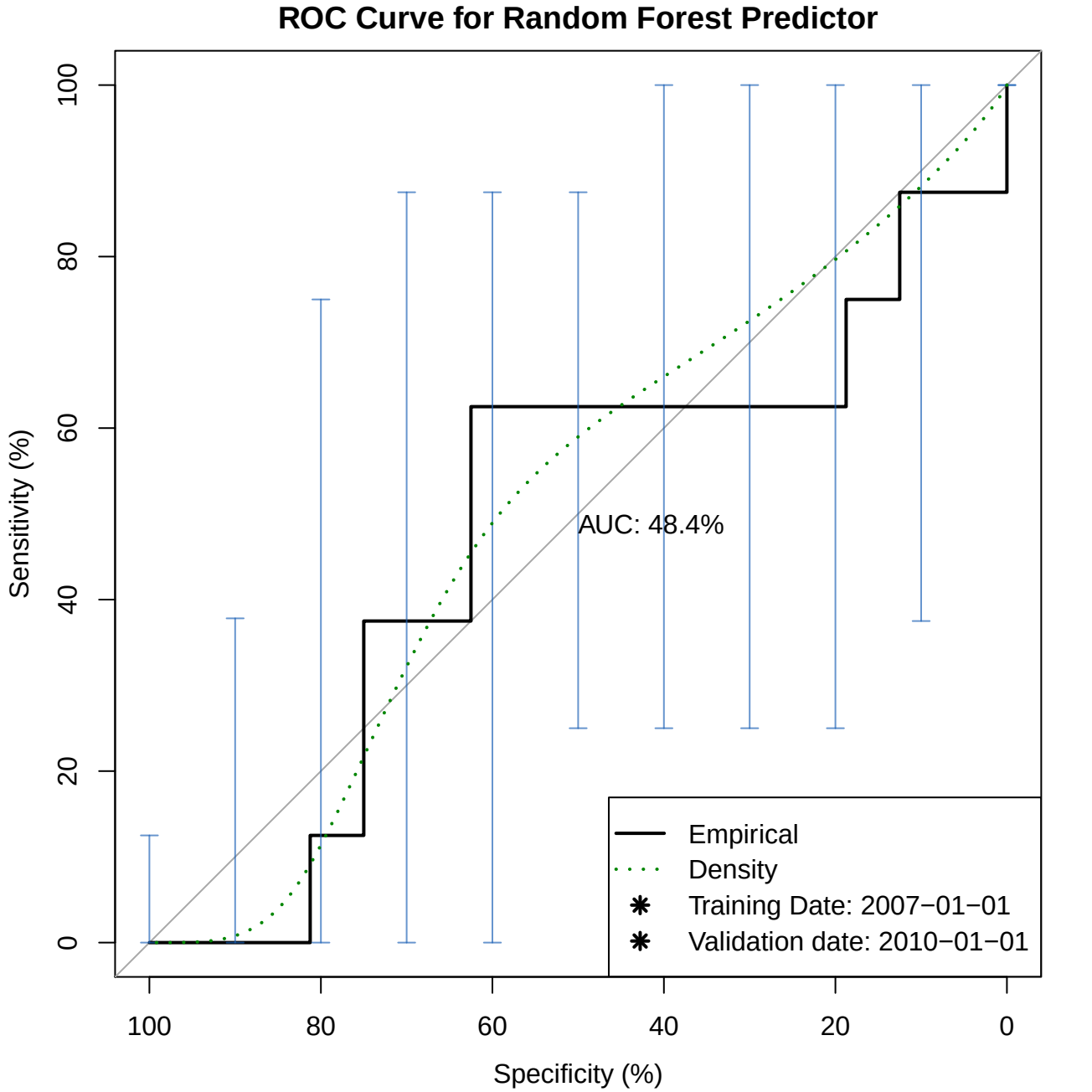

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
False positive rate
TestDate: 2007−01−01 Validate date: 2008−01−01

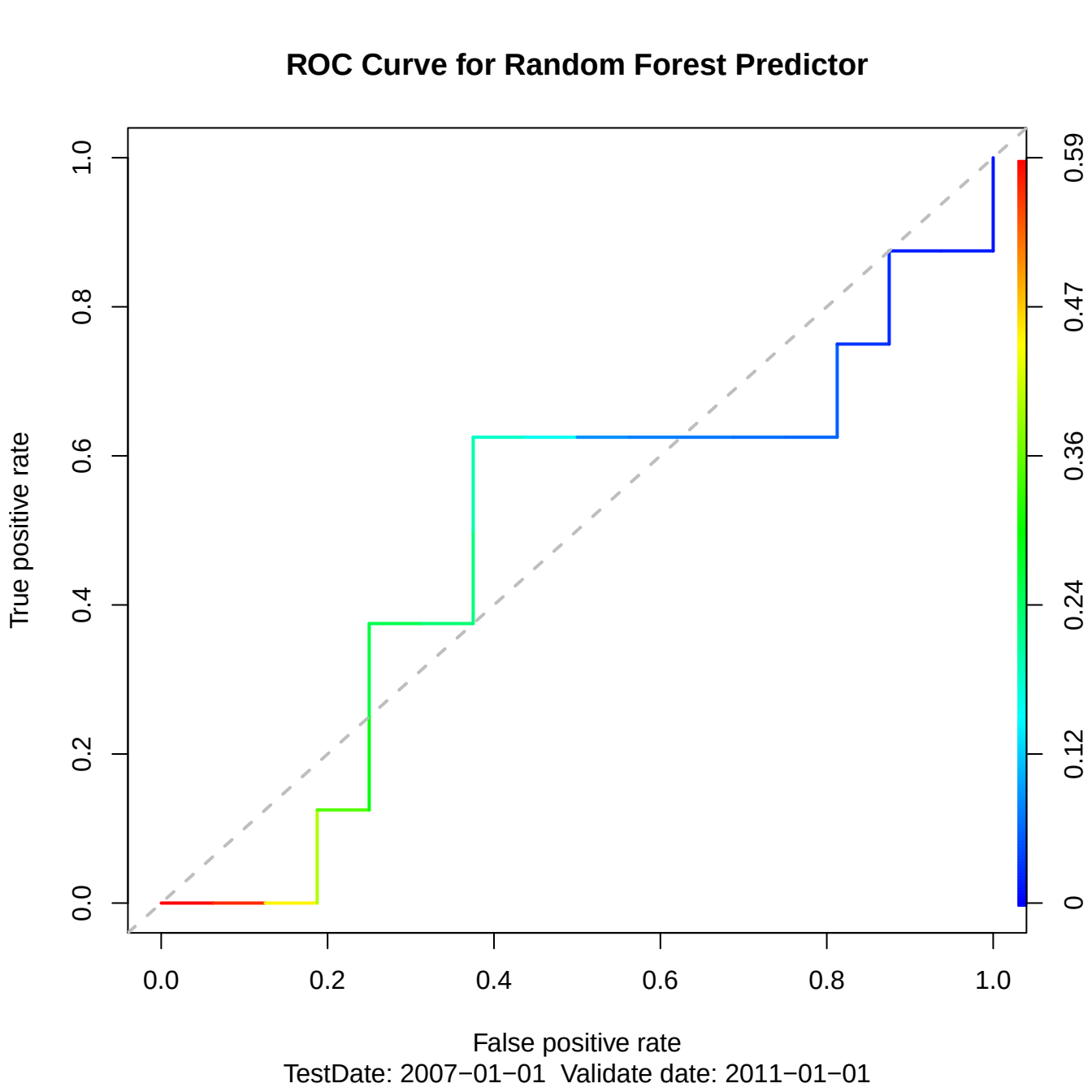

ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 77.8%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2008−01−01

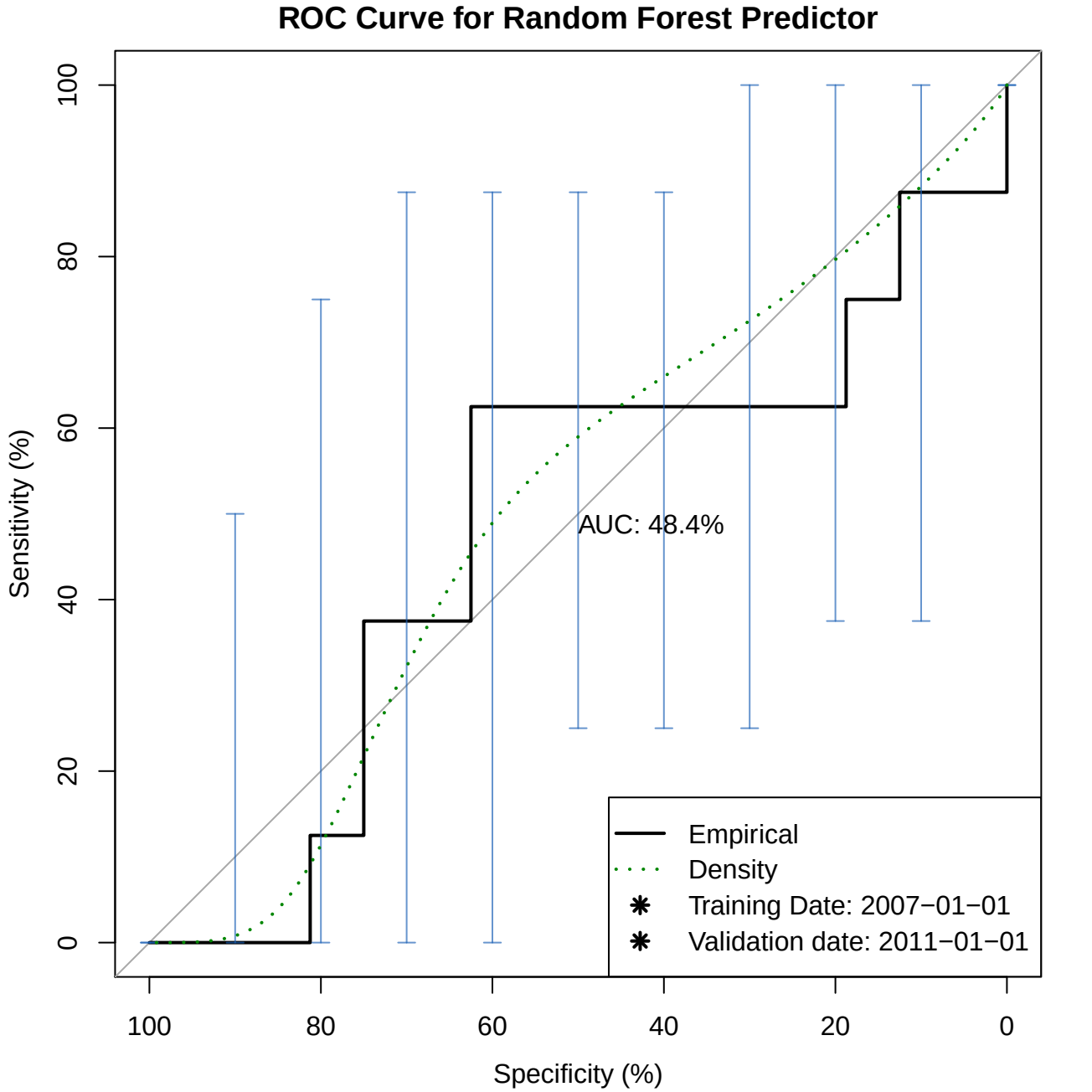

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
False positive rate
TestDate: 2007−01−01 Validate date: 2009−01−01

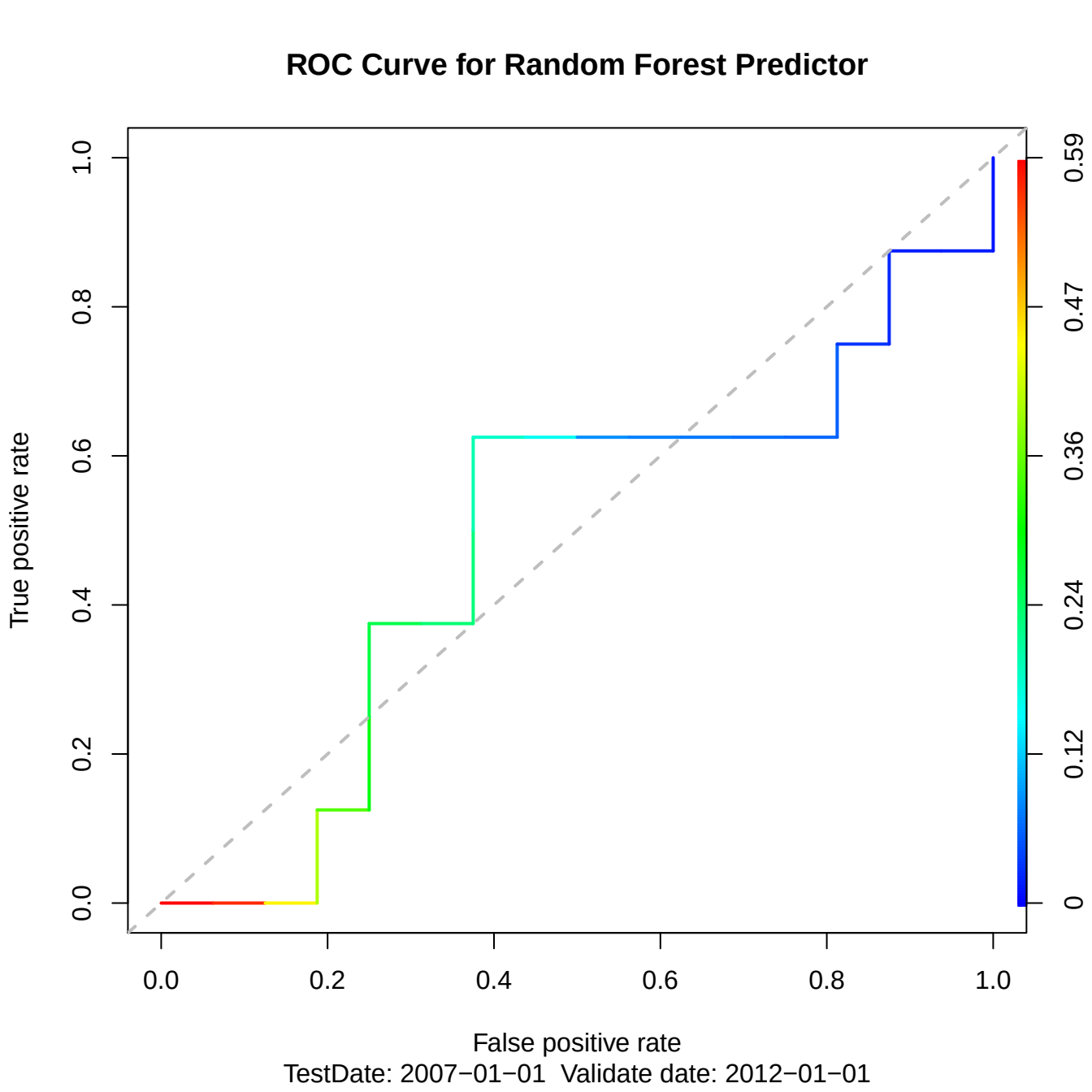

ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 42.1%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2009−01−01

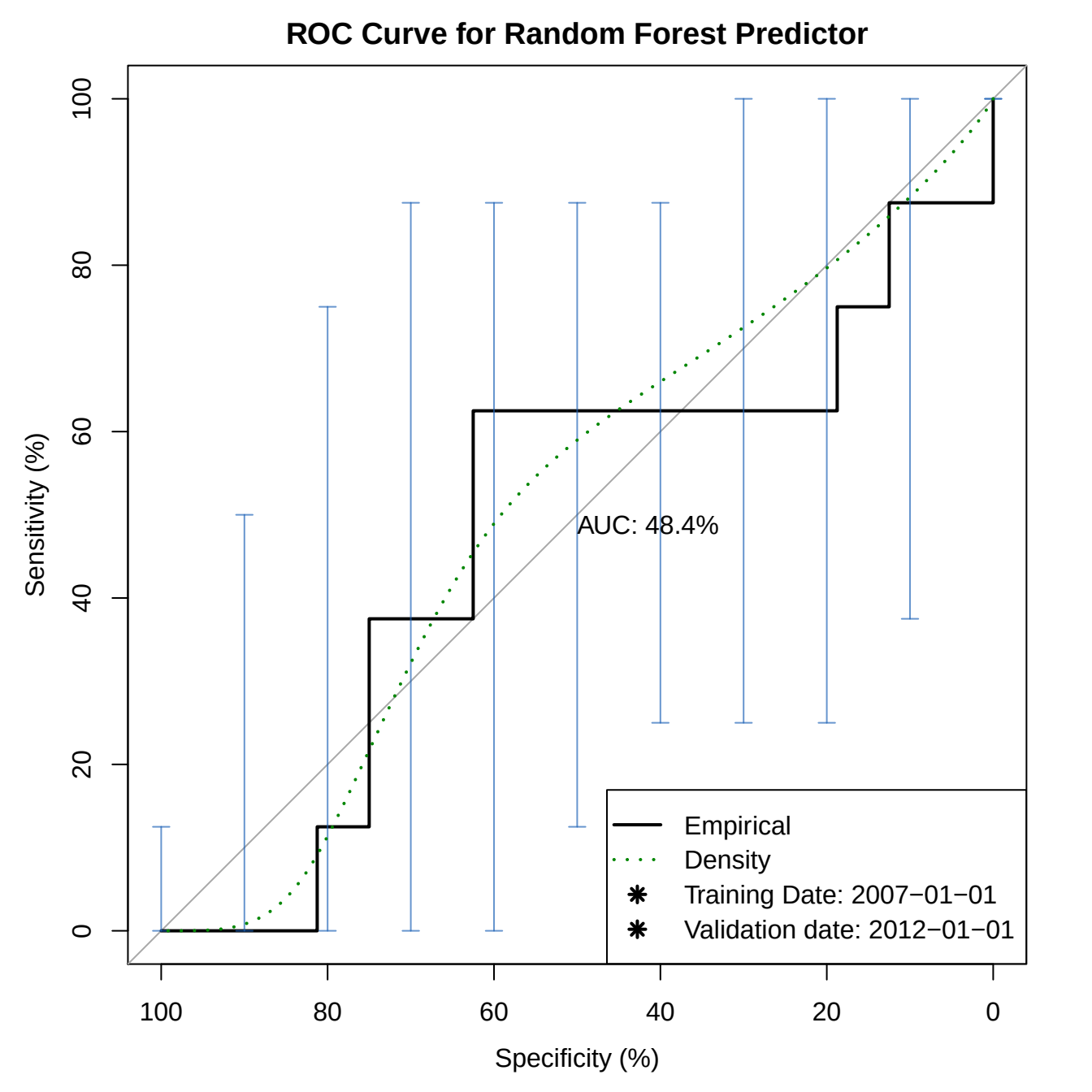

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
False positive rate
TestDate: 2007−01−01 Validate date: 2010−01−01

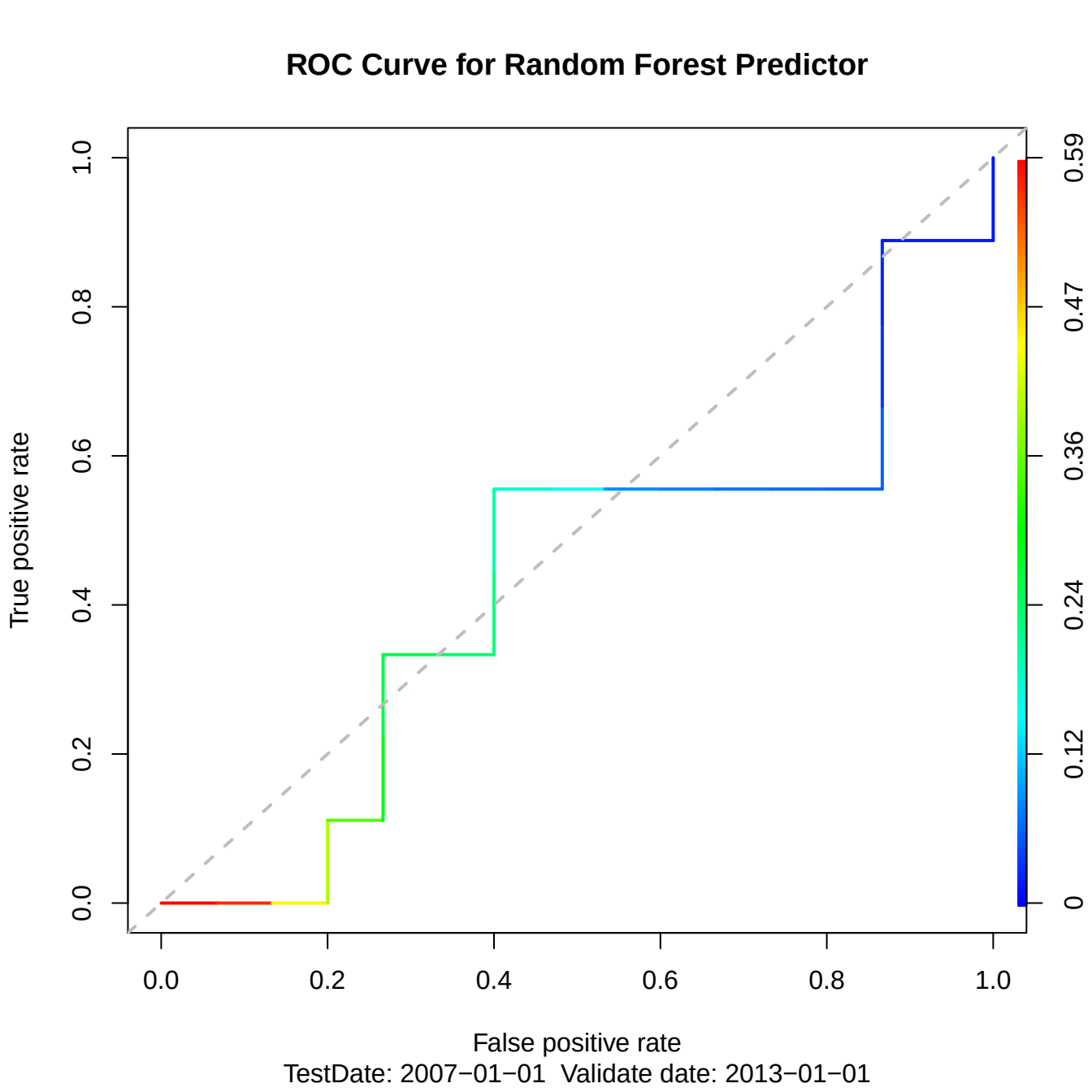

ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 48.4%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2010−01−01

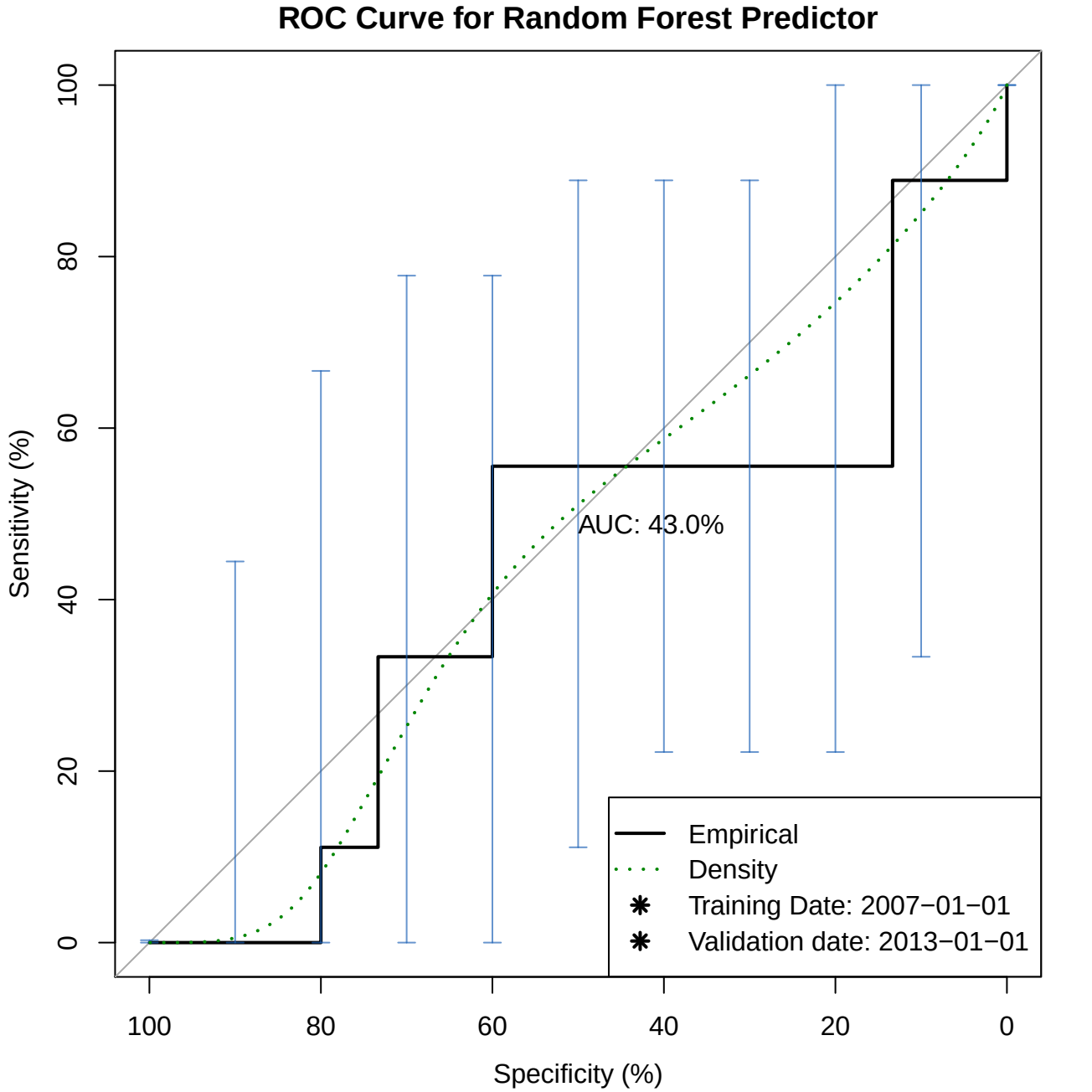

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
TestDate: 2007−01−01 Validate date: 2011−01−01

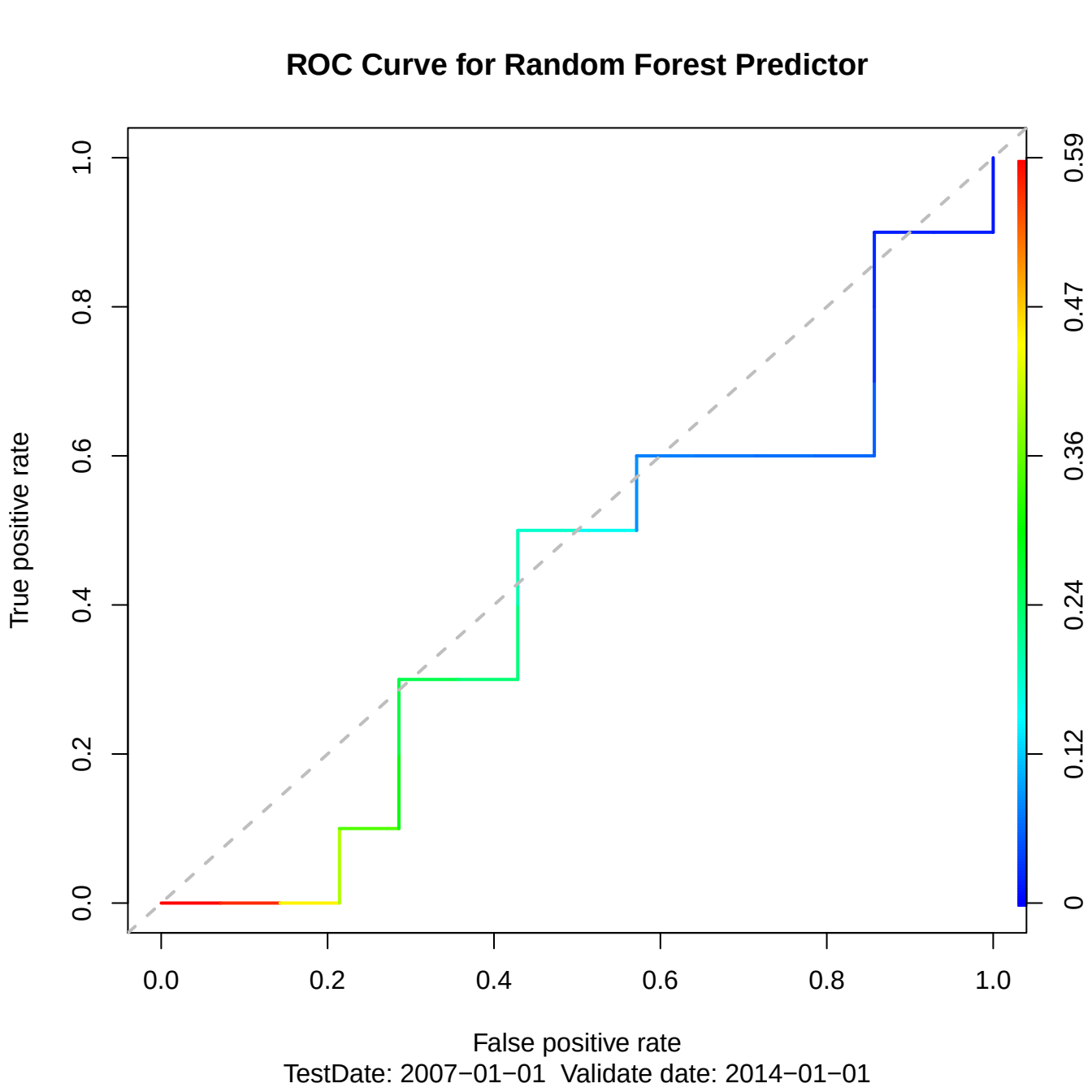

ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 48.4%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2011−01−01

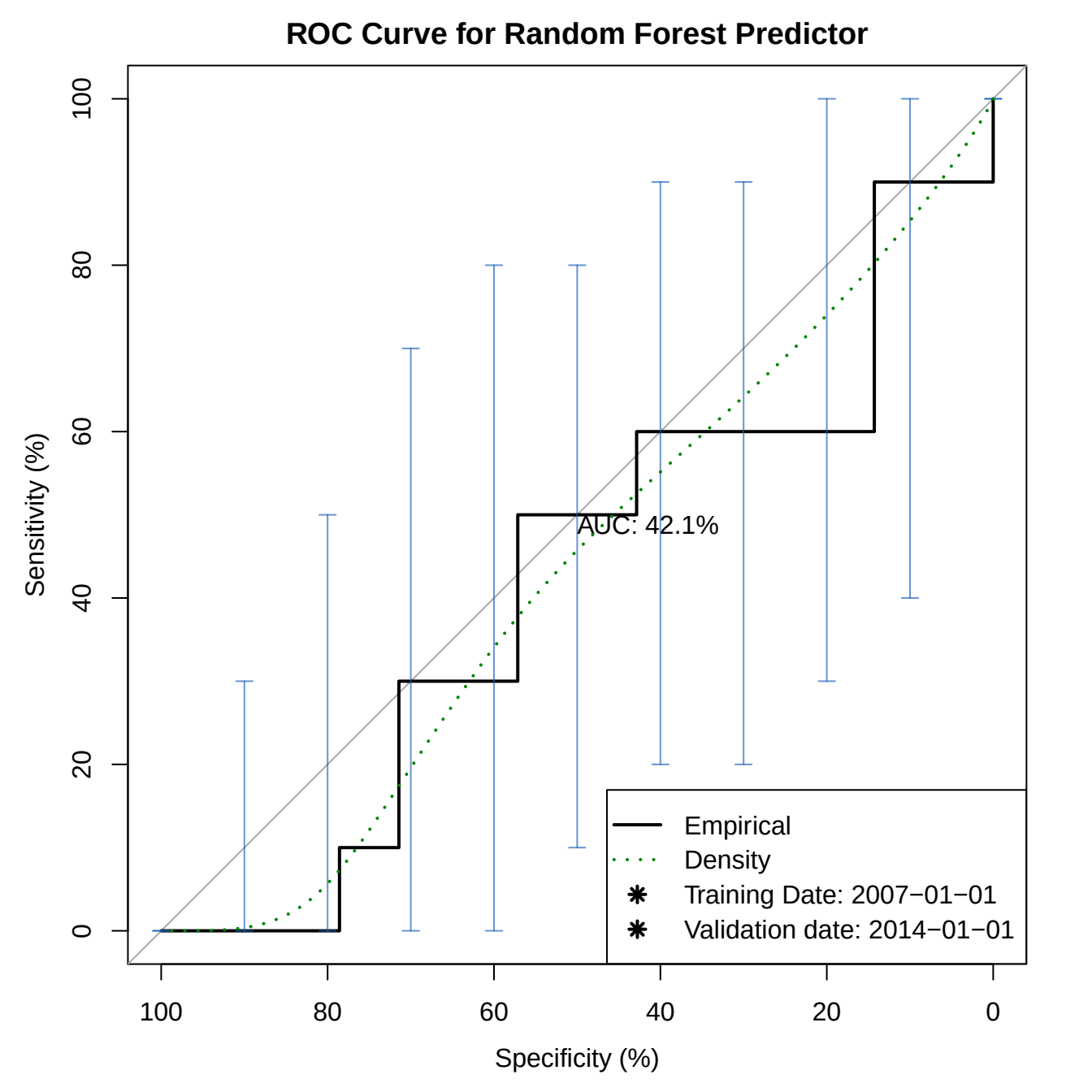

ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
TestDate: 2007−01−01 Validate date: 2012−01−01


ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 48.4%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2012−01−01


ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
False positive rate
TestDate: 2007−01−01 Validate date: 2013−01−01


ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 43.0%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2013−01−01


ROC Curve for Random Forest Predictor
True positive rate
0.0
0.2
0.4
0.6
0.8
1.0
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
0.59
0.47
0.36
0.24
0.12
0
TestDate: 2007−01−01 Validate date: 2014−01−01


ROC Curve for Random Forest Predictor
Sensitivity (%)
Specificity (%)
100
80
60
40
20
0
AUC: 42.1%
Empirical
Density
Training Date: 2007−01−01
Validation date: 2014−01−01