

1 Performance Metrics Across All Sheep

The following tables report the performance metrics (in meters) for all individual desert bighorn sheep (*Ovis canadensis mexicana*) using the subset of the GPS data for which each model could produce a prediction. The best performing model for each sheep is highlighted in bold. Desert bighorn sheep (9 F, 7 M) were tracked from November 2022 through May 2024 in the Sierra El Alamo mountains, 200 km from Carboca, Sonora, Mexico.

Table 1: RMSE (Longitude Models)

ID	BART	TGP	LM	GAM	RF
20020	208.7	756.2	245.6	247.2	66.6
20030	244.0	401.4	268.9	274.1	71.3
20038	186.2	423.2	200.2	200.5	50.4
20044	235.0	477.8	267.1	267.4	72.4
22952	235.1	602.4	281.6	281.9	89.2
22953	205.8	373.9	227.4	229.8	60.8
27861	254.8	799.9	278.3	279.4	68.5
39962	233.9	208.1	274.8	279.1	89.3
39964	222.1	243.0	248.8	248.8	65.4
39965	142.1	107.0	181.7	181.7	51.2
39966	214.6	137.3	240.2	240.4	65.6
39967	199.7	120.1	218.4	219.4	55.3
39968	276.8	287.8	352.4	353.7	109.6
39969	181.7	123.6	216.7	216.9	59.6
39970	186.8	210.6	220.5	222.3	62.3
39971	164.4	115.9	194.5	195.6	66.1

Table 2: RMSE (Latitude Models)

ID	BART	TGP	LM	GAM	RF
20020	179.8	947.6	207.5	208.6	63.1
20030	229.6	420.1	254.4	255.4	75.9
20038	147.9	330.3	162.5	163.1	44.6
20044	203.3	370.8	225.9	226.0	60.1
22952	237.8	748.6	276.2	276.4	81.7
22953	194.4	375.9	212.4	212.7	54.3
27861	212.3	602.3	228.3	230.5	62.8
39962	215.4	133.1	262.3	262.7	77.6
39964	201.3	207.1	217.8	218.3	57.9
39965	99.8	211.7	176.9	177.0	65.4
39966	198.6	132.1	220.2	220.5	65.2
39967	167.8	96.4	177.8	178.0	47.2
39968	238.0	195.5	260.0	262.0	76.9
39969	133.8	101.8	190.1	190.1	56.4
39970	138.5	174.9	164.3	165.7	46.1
39971	120.8	82.9	164.4	165.1	53

Table 3: MAE (Longitude Models)

ID	BART	TGP	LM	GAM	RF
20020	129.1	487.3	137.6	138.6	34.6
20030	146.0	254.4	150.1	150.5	39.1
20038	107.2	277.7	109.5	109.6	28.2
20044	145.0	288.3	157.6	157.9	41.9
22952	134.6	290.5	134.1	134.4	40.2
22953	121.6	225.0	125.7	124.7	32.6
27861	152.8	424.0	157.3	157.0	38.9
39962	141.5	128.9	136.7	138.2	42.3
39964	130.7	145.4	138.1	138.2	36.1
39965	90.8	66.9	102.6	102.6	29.5
39966	123.3	80.7	123.9	123.7	36.3
39967	127.3	83.8	134.4	136.2	32.9
39968	168.4	133.4	166.7	167.6	48
39969	99.9	71.3	108.5	108.1	28.6
39970	105.6	118.7	107.1	107.9	30.5
39971	98.7	75.9	116.7	115.8	39.2

Table 4: MAE (Latitude Models)

ID	BART	TGP	LM	GAM	RF
20020	111.4	465.6	106.2	106.3	34.2
20030	130.4	232.9	126.1	126.8	39.1
20038	83.9	208.3	85.2	85.4	24.6
20044	118.8	234.9	124.1	124.1	34.9
22952	131.0	388.2	126.2	126.3	39.2
22953	112.6	215.7	113.5	113.3	30
27861	121.7	318.5	123.8	124.1	33.6
39962	122.7	78.1	123.5	123.7	38.2
39964	123.3	132.4	121.8	122.3	34.3
39965	63.4	61.2	79.1	78.9	25.6
39966	113.5	81.2	109.5	109.2	34.2
39967	101.5	67.6	103.1	103.2	29.6
39968	138.6	122.6	132.9	133.3	40.9
39969	77.5	56.6	84.0	84.0	25
39970	83.1	98.1	86.0	86.6	22.5
39971	72.7	56.6	93.9	93.0	31.8

Table 5: Mean Euclidean Distance (MED)

ID	BART	TGP	LM	GAM	RF
20020	189.2	750.7	192.3	193.2	54.4
20030	218.2	383.5	218.4	219.2	62.3
20038	150.6	385.8	153.3	153.6	41.8
20044	208.5	413.5	222.6	222.8	61.2
22952	209.7	549.5	205.5	205.8	63.3
22953	185.2	347.2	188.9	188.0	49.9
27861	216.5	585.7	221.0	221.0	57.5
39962	206.5	164.6	203.0	204.4	63.4
39964	200.8	219.7	205.1	205.5	56.1
39965	120.9	103.9	142.9	142.8	44.2
39966	187.4	128.0	185.4	185.0	56.4
39967	180.2	119.0	187.0	188.5	49.6
39968	240.6	201.3	235.6	236.8	70.8
39969	140.9	101.6	153.2	152.8	43.2
39970	150.0	174.2	153.7	154.7	42.9
39971	136.1	104.9	166.9	165.5	56.2

2 Performance Metric Boxplots

The following boxplots display the distribution of the performance of our 5 model classes across the 16 sheep. For a given model, the performance metrics were calculated between the subset of data for which all models could produce a valid output and the predictions made by the given model.

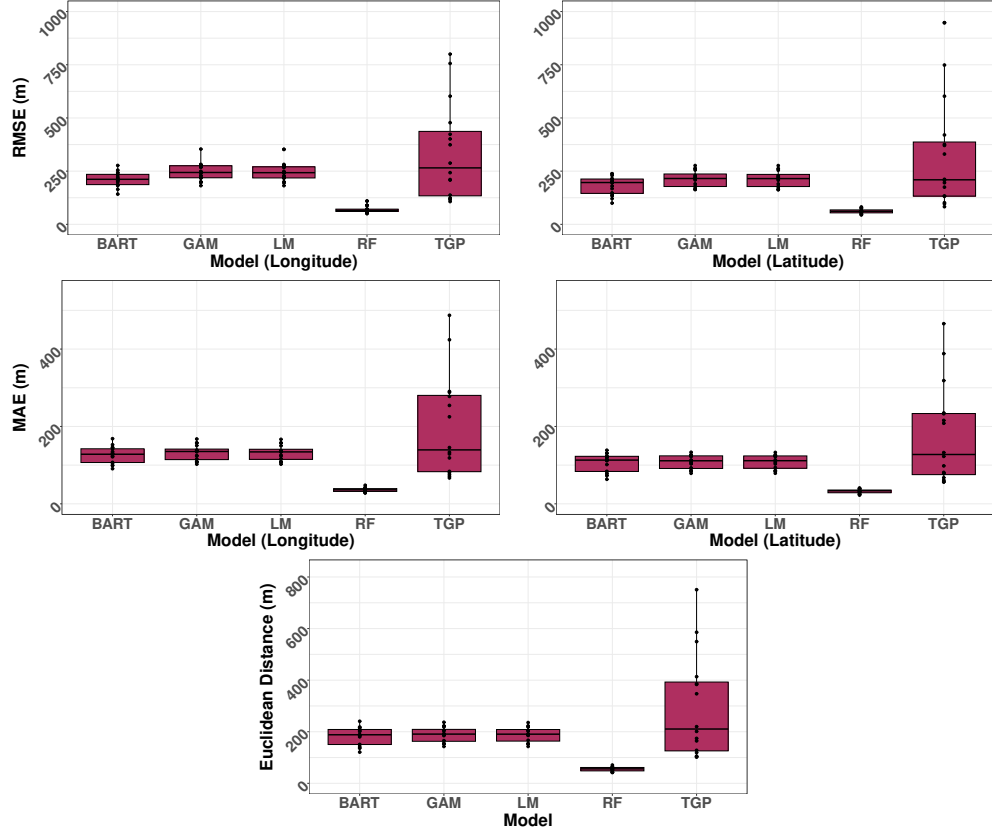


Fig. 1: Across all the performance metrics we used, RF had the best predictions for fixes with a known location and varied little around the mean of the metrics for a given model.

3 Prediction Uncertainty Metric Boxplots

The following boxplots display the distribution of the maximum, average, and minimum prediction uncertainty of our 5 model classes across the 16 sheep.

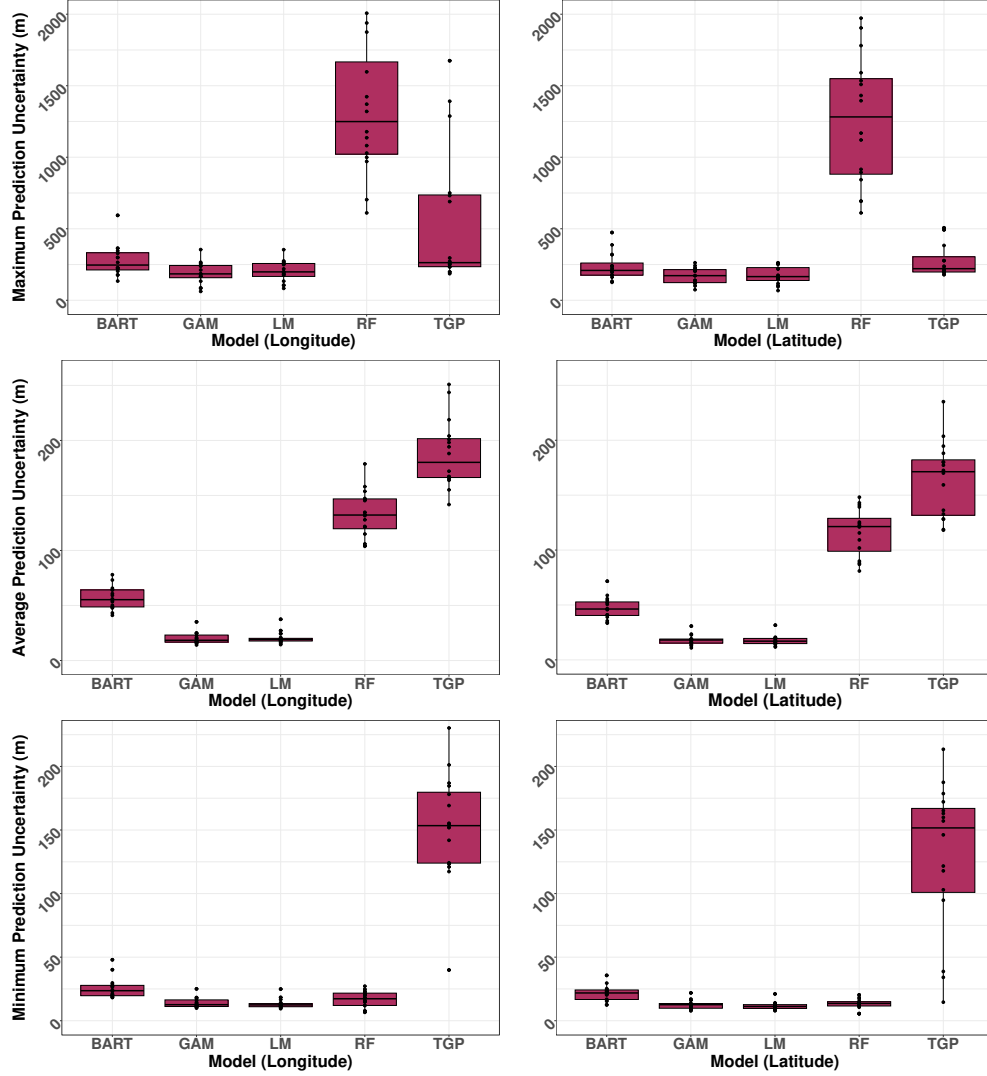


Fig. 2: The above plots show the mean and spread of the maximum, average, and minimum prediction uncertainty across each model, for each sheep. In other words, there are 16 data points for each model's box.

Table 6: # of Fixes Used to Train/Test Each Model

	BART	GAM	LM	RF	TGP
Count	54,381	54,312	54,312	51,119	55,277