Chapter 4

An RSSI-based localisation method for fine-scale wildlife tracking using an Automated Radio Telemetry System (ARTS)

Abstract

Abstract here

4.1 Introduction

Movement is fundamental to every aspect of life on earth. It is universal among organisms, shapes individual fitness, directs evolutionary development, and influences ecological processes, including responses to anthropogenic change (Nathan et al. 2008; Kays et al. 2015; Nathan et al. 2022). Understanding how, where, when and why animals move is therefore essential to answer questions in ecology, animal behaviour and conservation sciences. Tracking wild animals has revealed detailed patterns of habitat use; discovered long-term migrations across the globe; uncovered mechanisms of navigation, memory and cognition; and characterised animal communication and social behaviour (Strandburg-Peshkin et al. 2015; Zúñiga et al. 2016; Torney et al. 2018; Imlay et al. 2020; Lourie et al. 2021; Heathcote et al. 2023; Hurme et al. 2025). Despite this importance, the study of animal

movement has proved challenging in many taxa, due to the difficulty associated with collecting individual-level data from wild animals to an accurate degree (Kays et al. 2015; Coulson 2020).

Recent advances in wildlife tracking technologies, data processing tools and analytical techniques has allowed researchers to monitor many different terrestrial and aquatic organisms in almost continuous space and time, resulting in the growth of the field of movement ecology (Joo et al. 2020; Joo et al. 2022; Kranstauber et al. 2024). Sensors that are widely used to track wild animals include Global Positioning Systems (GPS) tags, proximity loggers, passive integrated transponder (PIT) tags, very high frequency (VHF) or ultra high frequency (UHF) radio tags, wireless sensor networks and bio-acoustic microphone arrays (Kays et al. 2015; Levin et al. 2015; Verreycken et al. 2021; Meier et al. 2024). Each method comes with major strengths and weaknesses, meaning there is no one-perfect solution. In general, the main trade-off's of tracking involve tag mass, sampling frequency, data retrieval options, cost, accuracy, and battery life (Kenward 2000; Tomkiewicz et al. 2010; Williams et al. 2020; He et al. 2023). For example, PIT tags utilise radio-frequency identification detection (RFID) antennas which detect changes in the electromagnetic field. Although these tags are cheap and lightweight, they require the animal to come into close proximity ($\sim$10cm) of the antenna, which limits the ecological systems that this method can be applied to and inncurs high risk of sampling biases (Kidd et al. 2015). Bluetooth low-energy (BLE) beacons are another low cost alternative, which utilise Apple's Find My network to locate tags (Farine et al. 2014). Although high-levels of accuracy can be obtained in a phone-rich urban environment, this strong sampling bias which affects both the spatial and temporal likelihood of tag localisation makes implementing this method unfeasible for animals in wild and remote spaces.

Tags which utilise GPS remain the primary choice for animal tracking due to their high spatio-temporal accuracy and global application (Kays et al. 2015). However due to the high energy demand needed for accurate localisation, GPS tags require heavy batteries and therefore remain too large to track around 65% of the worlds mammal species and 70% of birds (Hebblewhite et al. 2010; Bridge et al. 2011). The launch of the international cooperation for animal research using space (ICARUS) initiative is a promising move in the continued miniaturization of GPS tags and worldwide tracking (Curry 2018; Krondorf

et al. 2022). Although in the future GPS tags may be reduced to an appropriate mass for small vertebrates, the limitations of GPS signals in complex habitats such as trees, caves or underground burrows make this option infeasible for animals which utilise these spaces (He et al. 2023). Other solutions with competing spatio-temporal resolution have become available, such as the Wildlife Biologging Network (WBN; Ripperger et al. (2020)) and the ATLAS system (Advanced Tracking and Localisation of Animals in real-life Systems; Bijleveld et al. (2021) and Beardsworth et al. (2022)), but cost and installation effort remain high and technical expertise are required to operate the systems.

Radio transmitters have been used to monitor animal movement since the early 1960's (Marshall et al. 1962; Craighead et al. 1965; Cochran et al. 1965). Today they remain a popular, and often the only, option for tracking small animals due to the low-cost and low-weight of the tags. Traditional radio tracking involves locating and concurrently triangulating individually tagged animals in the landscape by trained field ecologists using hand-held receivers (Kenward 2000). This method is both logistically challenging and labour intensive and often the resulting data is irregularly sampled, of low spatial-temporal accuracy, and limited to few individuals (Billington 2004; Tavares et al. 2025). Further to this, there is often no attempt to evaluate the efficacy of traditional radio tracking when utilised in the field and there has been little work to evaluate the implemented techniques of positional error estimation, which is a requirement for applying many spatial models (White et al. 1990; Montgomery et al. 2010; Bartolommei et al. 2013).

A major leap forward in the use of radio telemetry for wildlife tracking has been the development of Automated Radio Telemetry Systems (ARTS), which have enabled often hundreds of animals to be simultaneously and continuously tracked across large areas in a range of habitats (Cochran et al. 1965; Kays et al. 2011; Weiser et al. 2016; Gottwald et al. 2019; Griffin et al. 2020; Wallace et al. 2022). Different ARTS designs have been implemented, but in general the main framework consists of a network of static radio receivers which localise the transmitter via indirect measures of the received signal (e.g., received signal strength indicator (RSSI), time difference of arrival)(Kays et al. 2011; Taylor et al. 2017). ARTS are generally split into omni-directional and directional systems, which vary in their receiver type, detection capabilities and cost (Höchst et al. 2021; Wallace et al. 2022; Paxton et al. 2022). Omni-direction systems typically have

simpler receivers which use one isotropic antenna to detect in an approximately uniform pattern (Krull et al. 2018; Wallace et al. 2022). Directional systems utilise more complex receivers which have multiple Yagi-Uda antennas that are each set in a different orientation (Taylor et al. 2017; Höchst et al. 2021). The different ARTS designs come with their own drawback, the most notable being the complexity of installation, optimization and operation; and additionally the difficulty associated with estimating transmitter locations from RSSI (Kays et al. 2011; Gottwald et al. 2019; Beardsworth et al. 2022).

The relationship between tag signal strength and the distance to the receiver can theoretically be described as an exponential decay function (Bensky 2008; Zhu et al. 2015; Luomala et al. 2019). In practice, the signal attenuation from environmental conditions (e.g., precipitation, habitat type) and multipathing effects, such as shadowing (signal blocked by object) and reflection (signal bounce), adds significant noise to RSSI making the true relationship difficult to predict in the environment (Zekavat et al. 2019). A diverse number of range-based localisation methods have been developed from this RSSI-Distance relationship, which includes lateration techniques (Paxton et al. 2022; Wallace et al. 2022). Lateration is commonly applied to omni-directional system designs and requires simultaneous detections from at least three receivers to estimate a location using the receivers overlap (Mendoza-Silva et al. 2019). Directional ARTS typically apply direction-based techniques for localisation, such as angulation estimation using the angle of arrival (AOA) (Gottwald et al. 2019; Kułakowski et al. 2010). This method estimates the AOA of the transmitter to the receiver by comparing the relative RSSI between antennas of the same receiver. AOA's from at least two receivers are then used in triangulation (Kays et al. 2011; Obeidat et al. 2021). Due to the requirements of multiple receiver detections, location drop out rate is often high for both lateration and angulation techniques. Positional error remains substantially higher than GPS-based systems, but techniques to minimise error have been implemented, including the use of RSSI-filters, grid search algorithms and trajectory smoothing with Kalman Filter (Hartmann et al. 2016; Paxton et al. 2022; Burcher et al. 2025; Ripperger et al. 2020).

Location fingerprinting is a range-free localisation technique that has been widely implemented in indoor positioning systems (Fu et al. 2009; Altaf Khattak et al. 2022). The basic principal of this method is to develop a site-specific radio map of known reference

points to train a machine learning model, which is then used to estimate the locations of unknown signals. The radio map characterises the RSSI at known locations, which encompasses the effect of the surrounding environment on signal noise (Singh et al. 2015; Khodayari et al. 2010). For outdoor application, promising work has been conducted in the recent emergence of Internet of Things (IoT) technologies, which is connecting everyday objects and device in a large low-powered communication network (Janssen et al. 2020b). Studies using an open-source dataset from the Sigfox IoT has shown fingerprinting to be a highly accurate method when receiver coverage is high (Janssen et al. 2018; Anagnostopoulos et al. 2019; Janssen et al. 2020a). The Sigfox IoT has been used in animal tracking however accuracy remains coarse in remote spaces (1 to 10km resolution), tag weights remain too large for small vertebrates (>1g) and data coverage is mostly restricted to countries in Europe (Wild et al. 2023). Both Wallace et al. (2022) and Tyson et al. (2024), have successfully implemented location fingerprinting methods for an omni-directional system, which were appropriate for tracking small vertebrates in a compact study area. Collecting the suitable training datasets for these systems can be extremely labour intensive and requires constant updating when the network setup is altered. As shown by Osta et al. (2024), an alternative approach is to implement a general fingerprinting model to receivers of the same design and in a similar environment. This approach comes at a cost of overall accuracy and has to date only been implemented in a directional system.

In this study, we present an open-source localisation method to estimate radio transmitter positions from an automated radio telemetry system. We have developed two machine learning pipelines which consist of a shared general receiver model that can pool training data across similar receivers, and an extended receiver model which utilises the concept of fingerprinting to improve accuracy. We demonstrate the utility of this methodology using data collected from an ARTS with multiple receiver types in an $\sim 4,500$ha region of Devon, South-West England, which is tracking a bat species of conservation concern in Britain, the greater horseshoe bat (*Rhinolophus ferrumequinum*). Using a series of controlled transect, training and test datasets were created to compare the performance of our general and extended receiver models with two other published fingerprinting-based methods: fingerprinting and receiver offsetting. We additionally evaluated the systematic error in the ARTS and utilised the resulting models to predict the associated error of

an unknown location. Finally, we developed a rule-based filtering approach to improve positional error and assessed the effect on mean and median error and location drop out rate.

4.2 Methods

4.2.1 Training and test data collection

This study was conducted in the region surrounding Chudleigh Caves and Woods SSSI (see section 2.2.1). The ARTS was deployed as described in section 2.2.2, however due to the corruption of the received signal strength indicator values in the SensorStation data during the first season, only data collected during the second study season was used in the development of this methodology.

To collect training and test data from known locations, a series of controlled transects were carried out across the study region. Transects were completed on foot, in a vehicle or by flying a drone, with the aim to collect as much data as possible for producing the fingerprinting radio map. The transects were designed to include a range of habitats, building densities, elevations and variable distances from the receivers in the network. The majority of the transects were completed in an ∼2,000ha section of the full study area. For transects completed on foot, a CTT PowerTag (battery powered, UHF digitally encoded radio transmitters: 10s pulse rate, 434MHz (La Puma 2021b)) was secured to a medium sizes plum (*Prunus domestica*) to mimic the mass of a small bat. This was then mounted to the top of a 3m bamboo cane with the transmitter antenna horizontally orientated, to imitate the position of a bat in flight. Transects were located using a sports watch (Suunto 9 Baro: 1m GPS and barometric accuracy, °5 compass accuracy), and activity states were simulated by travelling at a range of speeds (1 to 4m/s; active state) and periodically stopping at points on route (inactive state). At randomised stopping points, the bamboo cane was spun on the same spot to give readings from different aerial orientations, which may occur while bats are hanging in the roost or during foraging.

When transects were carried out using a vehicle, the bamboo cane was secured to the exterior of the vehicle and was driven at a maximum speed of ∼ 18m/s (above the maximum recorded speed of GHS REFERENCE). Time-stamped GPS locations were

collected using a smartphone device running GPS Logger (BasicAirData on Android), with stopping points being included on the transects. To get recordings at varying heights and to access difficult topologies, some transects were completed using a drone (DJI Mini 2, Shenzhen, China). Both the CTT PowerTag and plum were secured to the top of the drone, which were then flown on the route. The drone was programmed to spin at randomised points on each transect and locations were recorded using the drones internal Global Navigation Satellite System (GPS, GLONASS and GALILEO).

4.2.2 Machine Learning Based Localisation Methods

For the task of localisation, we developed two sets of ML models. The first set of models was a shared general receiver model, which was applied to receivers in the network with limited or no training data (Sulaiman et al. 2023). Due to device failure, re-deployment in the network and limited timing and resources for training data collection, some receivers in the network had no data from known locations. The second set of models extended the general receiver model to utilised the full fingerprinting radio map. This model uses the data per individual receiver, as well as a binary matrix of all other receiver detections in the network to predict a location. Each of these localisation methods consisted of four ML models, which were split by receiver type (SensorStation and Node) and the location estimate (Bearing and Distance). We created a machine learning pipeline to automate the process of deploying our ML models which consisted of multiple linear steps, including: (1) data pre-processing; (2) data splitting for validation and testing; (3) model selection and training; (4) hyperparameter tuning and model evaluation; and (5) model inference and deployment. We additionally compared our methodologies to two other published fingerprinting-based methods, by comparing the mean positional errors. We finally assessed the systematic errors in the radio network and used these models to develop a set of rule-based filters for error handling to improve positional error.

4.2.3 Data processing

To create labelled training and test data, all location data were downloaded from the GPS tracking devices and processed using R statistical software version 4.4.3 (R Core Team 2025) and the packages *tidyverse* (Wickham et al. 2019), *sf* (Pebesma et al. 2023; Pebesma 2018) and *RPostgres* (Wickham et al. 2025). Data from the 'track points' layer of the GPX

files were extracted using the '`st_read`' function and were transformed into a data frame which included the variables 'point_id', 'time', 'latitude' and 'longitude'. Detection data from the ARTS was time-matched to the closest matching location using a rolling join with the package *data.table* (Barrett et al. 2025). The resulting dataset was then filtered to remove any duplicate detections and those that were outwith the transects start and end times, resulting in a total of 10,016 time-matched locations.

The remainder of our machine learning pipeline was produced using Python version 3.12.8 (imported modules: *datetime*, *time*, *utm*, *multiprocessing*, *math*, *itertools*, *collections* and *os*; imported libraries: *NumPy* (Harris et al. 2020), *pandas* (McKinney 2010; The pandas development team 2024), *Dask* (Rocklin 2015; Dask Development Team 2016), *scikit-learn* (Pedregosa et al. 2011), *GeoPy* (GeoPy Development Team 2023), *GeographicLib* (Karney 2013; Karney 2025), and *SQLAlchemy* (Bayer 2012)). To reduce the effect of influential outlier values that are caused by multipathing and shadowing, it is common practice to average the RSSI values of a radio map over a set time period (Alhomayani et al. 2020; Wang et al. 2020; Paxton et al. 2022). To examine the effect of RSSI averaging on location error, we resampled our dataset at varying time periods (10, 30, 60, 120 and 180 seconds) using the '`Grouper`' function. For the longer sampling frequencies (>10 seconds), we averaged the point location latitude and longitude to give a single location for model training. We ran our full ML pipeline on each of the five datasets, which allowed us to compare localisation error from raw detections and those that were averaged over different sampling frequencies. Each group in the datasets were given a unique ID (defined here as t), which corresponded to the radio tag ID and the timestamp of the start time of the group.

For each t, explanatory variables describing the RSSI value at each antenna at each receiver in the network were created, following similar methodologies to those used in Osta et al. (2024). For the 10 second sampling frequency, this meant there was one variable at Nodes and five variables at SensorStations which corresponded to the raw RSSI value at each antenna. Three variables were calculated for each antenna at the longer sampling frequencies, which were the mean RSSI value, the standard deviation (SD) of the RSSI values, and a count of the received number of detections. This meant there were three variables per Node and 15 variables per SensorStation in our general receiver ML models.

Missing values in datasets are incompatible with ML estimators, therefore any NA's were filled for each variable with inconceivable values, which were set as 1 less than the minimum observed RSSI value in the entire dataset for the raw or mean RSSI; 100 for the RSSI SD; and 0 for the received number of detections. To assess if the inclusion of the additional RSSI variables describing the SD and count improved model fit, we ran our full ML pipeline on two versions of the averaged datasets, the most simple only included the mean RSSI values from each antenna as explanatory variables, whilst the other included all three variables.

To create our full fingerprinting radio map, we used the raw or mean RSSI variable to create a binary matrix of all the receivers in the network. For each t, a single variable was created for each receiver (in our network we had calibration data for 22 receivers, which gave 22 variables). If any of the antennas on an individual receiver made any detections at a given t, then a value of 1 was assigned to that receiver, otherwise a value of 0. The resulting radio map was merged with the RSSI explanatory variables for each t per receiver. This resulted in there being a total of 23 or 25 explanatory variables for Nodes and 27 or 37 explanatory variables for SensorStations in our extended fingerprinting radio map ML models (raw or mean RSSI with radio map, or mean RSSI, SD, and count with radio map). After all explanatory variables were calculated, the datasets were filtered to remove any rows from receivers where all antennas had no detections. The datasets were split by their receiver types (Nodes and SensorStations) and were reshaped into a wide format so that each row corresponded to a single tag-receiver combination per t.

Due to having two different receiver designs in our network and based on their varying success in other location methods, we used the bearing and distance of the transmitter in relation to each receiver as the response variables for our sets of ML models (Wang et al. 2013; Kays et al. 2011; Gottwald et al. 2019; Paxton et al. 2022). To calculate these two variables, we used the '`WGS84.Inverse`' function to compute the geodesic between the latitude and longitude of the receiver and the latitude and longitude of the transmitter. By default, this function calculates the distance in meters and the azimuth in degrees from North between the two given points. The azimuth was transformed to a bearing of $\pm 180°$ with $0°$ set to North. This process was carried out for each receiver in the network that detected the transmitter at the given t, which when combined described the estimated transmitter location in relation to each receiver.

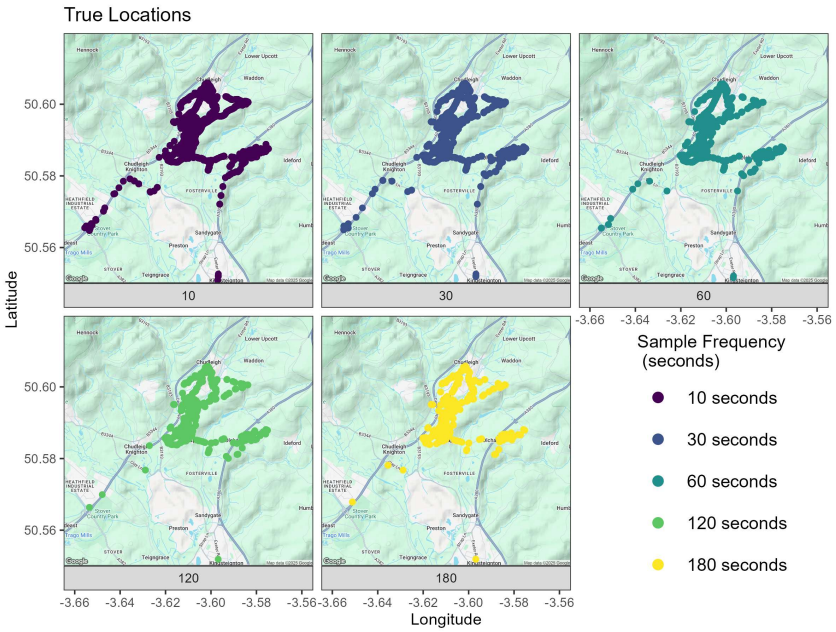

Figure 4.1: True locations that were input into the ML models for each sampling frequencies (seconds). Each panel and colour represents a different sampling frequency.

4.2.4 Machine Learning model

To improve ML model performance, prior to model fitting we standardized (0 mean, scaled to unit variance) the RSSI explanatory variables with the '`StandardScaler`' function. The binary radio map and the response variables remained in the previously described format and were not transformed further. To evaluate our ML models, we performed 10-fold cross validation by randomly splitting our data into 80% training data and 20% test data 10 times. To prevent data leakage, we used the '`GroupShuffleSplit`' and '`splitter.split`' functions with the groups argument set to the unique ID. This ensured that any data from the same tag at the same time interval across different receivers were always either in the training or the test data when randomly split. The number of unique locations in the test data per fold per sampling frequency were as follows: 1,047 for 10 seconds; 383 for 30 seconds; 201 for 60 seconds; 106 for 120 seconds; and 73 for 180 seconds.

We used k Nearest Neighbours (kNN) algorithm for our ML models, due to its simplicity and widespread use in both indoor and outdoor RSSI-based positioning (Janssen et al. 2018; Janssen et al. 2020a; Obeidat et al. 2021; Tyson et al. 2024). Whilst commonly associated with classification tasks, kNN can also be utilised for predicting continuous values or regression. The algorithm estimates the likely target value of unknown data (here, bearing

or distance) by averaging the target value of k nearest neighbours from the training data (Cover et al. 1967). We trained independent models per location estimate (Bearing and Distance) and receiver type (SensorStation and Node), using the '`KNeighborsRegressor`' function, which resulted in four ML models per pipeline (general receiver and radio map).

To select the optimal values for our kNN models we performed hyperparameter tuning for 'k', 'weights' and 'algorithm'. We tested values of 1 to 30 for k, which is the number of neighbours to use in estimation. The weights hyperparameter applies a weighted function in prediction, which can either be 'uniform' (all points in the neighbourhood are equal) or 'distance' (applies an inverse distance weighting which prioritises closer neighbours). 'Brute Force', 'KD Tree' and 'Ball Tree' are the different algorithms that can perform the nearest neighbour search, which primarily alter computation time. For the 'distance' hyperparameter, we set this to 'Euclidean' in all models, based on the work of Janssen et al. (2018) and Anagnostopoulos et al. (2019). All sets of ML models were fit with each set of hyperparameter values to the 10-fold of training data and predictions were made for the test data. We calculated the Root Mean Squared Error (RMSE) and R^2 to assess model performance, which were averaged across the 10-fold of data. Hypermarameters were selected based on the combinations which resulted in the lowest RMSE and the highest R^2. Additionally, bearing and distance error were calculated as the absolute difference between the predicted value and the true value.

4.2.5 Localisation Estimates

Using the bearing and distance predictions from the 10-folds of test data of the best performing models, location estimates were calculated per receiver. This was achieved by inputting the receiver latitude and longitude into the '`distance.distance`' and '`destination`' functions, along with the model predictions. These functions in combination when given a start point, bearing and distance, will calculate the destination point latitude and longitude. This was repeated for each individual receiver that detected the transmitter at any given t, which gave a predicted location in relation to each receiver.

To calculate the final transmitter estimate for t where multiple receivers were involved, the centroid between each receivers predicted location was used. First, the predicted latitudes and longitudes of the transmitter were transformed to eastings and northings using

the '`from_latlon`' function. Latitude and longitude are a geographic coordinate system (GCS) which is based on a spherical model and uses angular units. Eastings and northings are a projected coordinate system (PCS) and use linear units on a flat surface. By reprojecting the location estimate, we then calculated the centroid by averaging the easting and northing of all receiver location estimates per t. The final transmitter locations were converted back to latitude and longitude using the '`to_latlon`' function. The positional error for each location was calculated as the distance in meters of the predicted location from the true location, using the previously described '`WGS84.Inverse`' function.

4.2.6 Comparative Methods: Fingerprinting and Offsetting

We compared the results of our general receiver pipeline and radio map pipeline to two other published fingerprinting-based methods (described below; Tyson et al. (2024) and Osta et al. (2024)). Although one of the published methods utilised the H20 AutoML algorithm for their method, we used only kNN for comparison for reproducibility and due to the differences associated with ensemble modelling (Barai et al. 1999; LeDell et al. 2020). We followed all data pre-processing steps that were outlined in each publication, then applied the steps of our machine learning pipeline to each method to apply appropriate scaling, 10-fold cross validation, hyperparameter tuning and model evaluation. Importantly, when splitting our dataset into training and test data for our comparative methods, we ensured the same random seed values were used to include the same groups of t per method.

Fingerprinting

Fingerprinting is a range-free localisation technique which has widely been applied to indoor positioning systems (Singh et al. 2015). It utilises non distance-based estimations for localisation and instead uses network connectivity, direct signal measurements, or other non distance dependent features (Altaf Khattak et al. 2022). The general principal of fingerprinting is to create a radio map of a unique set of location-dependent signal parameters in an 'offline' training phase. Receiver locations are not required, it is however expected that they remain constant in the network. During the 'online' phase, unknown locations are estimated by matching the observed signal parameters to those fingerprints in the radio map, usually by probabilistic or machine learning approaches (Fu et al. 2009;

Pinto et al. 2021; Wallace et al. 2022).

Tyson et al. (2024) developed an RSSI-based fingerprinting method for outdoor use in a Node network. We adapted the pre-processing steps from this methodology to allow the inclusion of both Nodes and SensorStations in location estimation, as each receiver type required a different number of variables for each antenna (1 vs. 5). Our constructed radio map contained a total of 62 variables, each describing the RSSI at a particular antenna per individual receiver. We followed our previously described pipeline to train two kNN models which used the direct RSSI values in the radio map to predict the northing or easting of the transmitter. This was carried out for each t in all 10-folds of test data and positional error was calculated as previously described.

Receiver Offsetting

Osta et al. (2024) developed a receiver-centric RSSI-based localisation method for outdoor use in a directional receiver network. True fingerprinting is a range-free method, as this methodology uses the distance of the transmitter from each receiver to estimate locations we have here named this method 'Receiver Offsetting' to distinguish from the previously described comparative fingerprinting method.

As per the fingerprinting method, we adapted the pre-processing steps for use with both Nodes and SensorStations. We first calculated the three response variables describing RSSI (raw or mean RSSI, SD and count) per antenna per receiver, as described in section 4.2.3. The two response variables used in this methodology were the X and Y offset of the transmitter in metres to the centre of each individual receiver. This was calculated by first reprojecting the coordinates of the receivers and training data locations from a GCS to a PCS. For each receiver detecting the transmitter at each t, the easting and northing of the receiver was subtracted from the easting and northing of the transmitter location. This therefore normalised the transmitted location as the distance in metres in the X (west to east) and Y (south to north) axes in relation to each receiver (receiver location is grid origin at 0,0).

We followed our previously described pipeline but trained two separate models for each receiver type which predicted the X or Y offset (4 ML models in total). As previously mentioned, we used the kNN algorithm for this methodology for simplicity in comparison,

but are aware that estimates for all methods may be improved by implementing ensemble modelling. We followed the authors steps for location averaging across multiple receivers, whereby the geographic midpoint was calculated by averaging the easting and northing of all receiver location estimates per *t*. This process was repeated for all transmitter locations in all 10-folds of test data and positional error was calculated as previously described.

4.2.7 Model Evaluation and Systematic Error

We evaluated the four localisation methods by comparing the mean positional error of the 10-fold test data. Having an estimate of error on an unknown location is a requirement for fitting certain spatial models (Calabrese et al. 2016). To do this, we additionally created sets of models which could be used to predict the possible positional error of a location by using variables which described the systematic error. We calculated numerous variables which we believed may, or have been previously shown to, describe the systematic error in a network (Tyson et al. 2024; Osta et al. 2024). These variables were: (1) the number of Nodes that detected the signal; (2) the number of SensorStations that detected the signal; the (3) minimum, (4) maximum and (5) mean distance between the receivers that detected the signal; (6) the minimum distance to a receiver that detected the signal; (7) the mean distance to receivers that detected the signal; the (8) distance to the centre of the Node network and the (9) SensorStation network; and (10) average RSSI of the receivers that detected the signal.

We fit three different sets of models to examine the systematic error in the system. The first sets of models investigated the overall systematic error; the second set of models investigated the effect of measurable variables (description below) on locations that were estimated by multiple receivers; the final set of models were of estimates of locations from single receivers. All statistical analysis were performed using the R statistical software version 4.4.3 (R Core Team 2025). All models were Generalized Linear Mixed Models (GLMMs, described below), and were fitted using the R package *glmmTMB* (Brooks et al. 2017; McGillycuddy et al. 2025). We fit separate models for each sampling frequency and all predictor variables were 0-1 normalized (except for average RSSI which was standardized) prior to model fitting.

The following workflow was carried out for each sampling frequency, however the in-

clusion of the RSSI SD and count variables was additionally assessed in the models for the longer sampling frequencies. For the first set of models, we first fitted a GLMM with positional error as the response variable (metres), using a Tweedie error distribution with a log link function. Tweedie error was used as the response variable is continuous data and is primarily positive, but does in fact have a mass at 0 (Dunn et al. 2018). All previously described variables were included in the global model as an interaction with localisation method. There was a single three-way interaction between localisation method, the number of Nodes and the number of SensorStations in the model, to account for the expected differences between the receiver types. Random intercepts for the unique point ID and the random seed value used for splitting training and test data were included to account for non-independence of repeat sampling. Model fit and the distribution of the scaled residuals were assessed by simulating from the fitted model (n = 1,000) using the '`simulateResiduals`' function from the R package *DHARMa* (Hartig 2024). Recommended model checks of inspection of Q-Q plots, model dispersion, frequency of outliers and zero-inflation were performed. Candidate models were generated by dropping a single interaction from the global model and an information theoretic approach was used to select the best performing model by ranking the Akaike Information Criterion (AIC) (Newland 2019). All results were transformed to the response scale using functions from the R package *boot* (Davison et al. 1997; Canty et al. 2024) and outputs were plotted using functions from the packages *ggnewscale* (Campitelli 2025), *MetBrewer* (Mills 2022), *viridis* (Garnier et al. 2024), *rstatix* (Kassambara 2023b), *ggmap* (Kahle et al. 2013) and *ggforce* (Pedersen 2024).

For the second set of models, we assessed the effect of variables that could be measured from an unknown location on positional error. To do this, we subset each of our datasets to include only locations that had been estimated by multiple receivers. We then fitted a GLMM with positional error as the response variables (metres), using a Tweedie error distribution with a log link function. The three-way interaction between localisation method, the number of Nodes and the number of SensorStations was included in the model, along with interactions between localisation method and the minimum, maximum and mean distance between the receivers that detected the signal, and the average RSSI of the receivers that detected the signal. The rest of the model structure, model checks, selection process and visualisations were as previously described.

The final set of models also assessed the effect of measurable variables on positional error but instead for locations that had been estimated by a single receiver. The dataset was subset and a GLMM with positional error as the response variable (metres), was fit with a Tweedie error distribution with a log link function. A single three-way interaction between localisation method, the receiver type and the raw or average RSSI of the receiver that detected the signal (raw for Nodes, average for SensorStations) was included in the global model. The random intercepts for unique point ID and random seed value were included and the previously described model checks and selection process were used.

4.2.8 Error handing

To improve positional error, we created a series of rule-based filters based on the results of the measurable variables models. To prevent high computation load, we tested these rule-based filters only on our two localisation methodologies (general receiver and radio map) and at the 10 second sampling frequency (raw data). We created three separate filtering techniques which were based on the maximum RSSI at each receiver; the mean distance (metres) between the receivers at each unique ID point; and the number of Nodes and SensorStations at each unique ID point. We tested the effect of each filter individually and simultaneously on both the mean positional error and location drop out rate. We additionally assessed the change in variance by testing the equality of the coefficient of variation (CV) between the filtering types by applying a Modified signed-likelihood ratio test using the '`mslr_test`' function from the *cvequality* package (Marwick et al. 2019; Krishnamoorthy et al. 2014).

For the RSSI filter, we assume from our models that Nodes and SensorStations have different levels of error associated with RSSI. We therefore tested the impact of these filters independently on each receiver type and in varying levels of combinations. We filtered receivers to obtain only those that received a maximum RSSI above a set threshold. We tested RSSI threshold values of -115dB (minimum RSSI observed; no filter), -110dB, -105dB, -100dB and -95dB at Nodes and SensorStations, which gave a total of 24 different combinations of RSSI-rule-based filters.

The next two rule-based filters were only applied to location estimates where multiple receivers were used for inference. For the mean distance filter, we assume from our models

that the positional error increases as the mean distance between the receivers used in location estimation increases. We filtered point locations where the mean distance between the receivers was greater than a set threshold. We tested threshold values of 3,500m, 3,000m, 2,500m, 2,000m, 1,500m, and 1,000m. To reduce the location drop out rate, for each distance-based filter when a unique ID group had more than 2 receivers, we calculated a $n \times n$ receiver matrix (where n is the number of receivers that detected the transmitter at the unique ID point). We then calculated the distance between each and every receiver in the matrix and removed the most distant receiver from the group. The mean distance between the remaining receivers was then recalculated. This process was repeated until the mean distance between the receivers at the unique ID point was below the set threshold, or the point was removed if the threshold was never met.

For the Node and SensorStation filter, we assume from our models that when at least 2 Nodes are used for localisation estimates, the additional use of information from SensorStations increases the positional error. To implement this filter, for each unique point ID we counted the number of Nodes that were used in localisation. If the number was equal to or greater than a set threshold, then location inputs from SensorStations were removed from the unique point ID. We tested thresholds of 2, 3, 4, 5, 6 and 7 Nodes.

For our top five performing RSSI filters, we tested how including the mean distance and Node and SensorStation filters altered positional error and drop out rate. We selected the top two performing distance based filters and the top three performing Node and SensorStation filters and simultaneously applied all three to the datasets.

4.3 Results

4.3.1 Positional Error

The mean positional error varied by both sampling frequency and localisation method (Figure 4.2). At the 10 second sampling frequency, mean positional error was highest for the General Receiver method at 677m, followed by the Receiver Offsetting method at 492m, then the Radio Map method at 332m, and lowest for the Fingerprinting method at 207m (Table 4.1). When using only the mean RSSI as an explanatory variable for ML inference, positional error decreased with increasing sampling rate for the General

Receiver and Receiver Offsetting methods (Figure 4.2). In contrast, both the Radio Map and Fingerprinting methods showed lower positional error at the 30 second sampling frequency compared to the 10 seconds, followed by increasing error at longer sampling frequencies (60-180 seconds). All four localisation methodologies were able to produce estimates for all test locations across all the sampling frequencies (Figure 4.4).

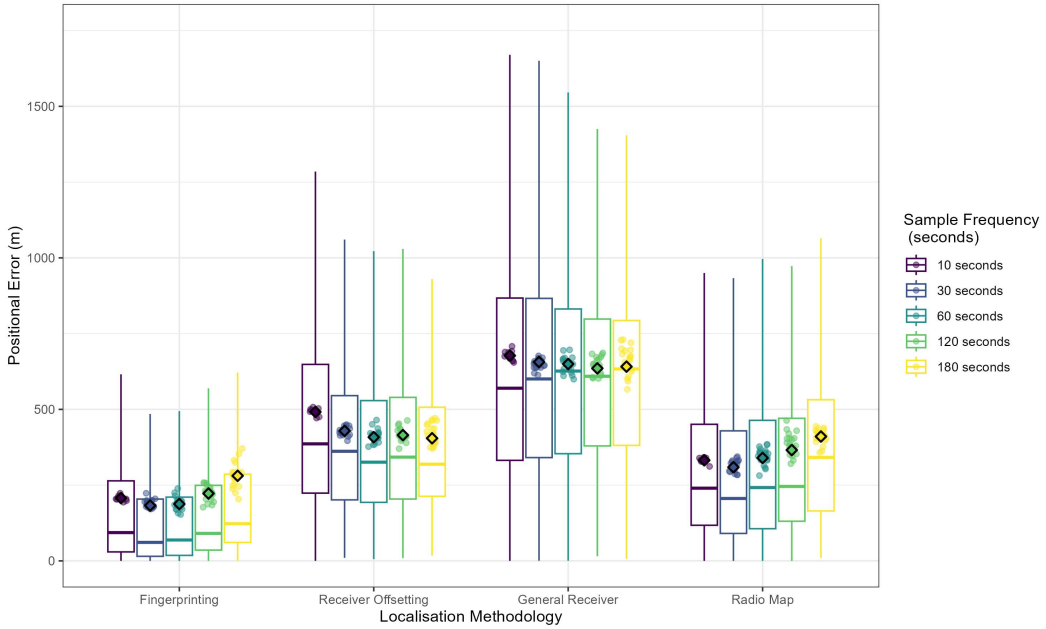

Figure 4.2: Comparison of the positional error (m) of each localisation method across different sampling frequencies (seconds). The lower and upper hinges correspond to the 25th and 75th percentiles, with the median indicated by the solid line. The end of the whiskers indicate ±1.5*IQR (inter-quartile range) and the diamonds mark the mean ±SEM. Colours represent the different sampling frequencies and the results shown in the figure are those of the 'Raw or Mean RSSI' only. Points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

Including additional RSSI-based predictors (RSSI SD and detection count) had varying effects on mean positional error, with improvements observed in some localisation method-sampling frequency combinations and worsening in others (Figure 4.3). At the 30 and 60 second sampling frequencies, there were no differences (0 - 10m) or marginal differences (11 - 20m) in the mean and median positional error across all localisation methods when including the additional RSSI-based predictors. At the 120 second sampling frequency, the mean positional error had no or marginal differences in the General Receiver, Fingerprinting and Receiver Offsetting methods, but worsened by ∼40m for the Radio Map method

when including the additional RSSI-based predictors. The median positional errors for the Radio Map, Receiver Offsetting and Fingerprinting methods worsened by 85m, 30m and 25m respectively. At the 180 second sampling frequency, using only the mean RSSI improved the mean positional error by ∼30m in the General Receiver and Receiver Offsetting methods. No or marginal differences were observed in the median positional error in all localisation methods.

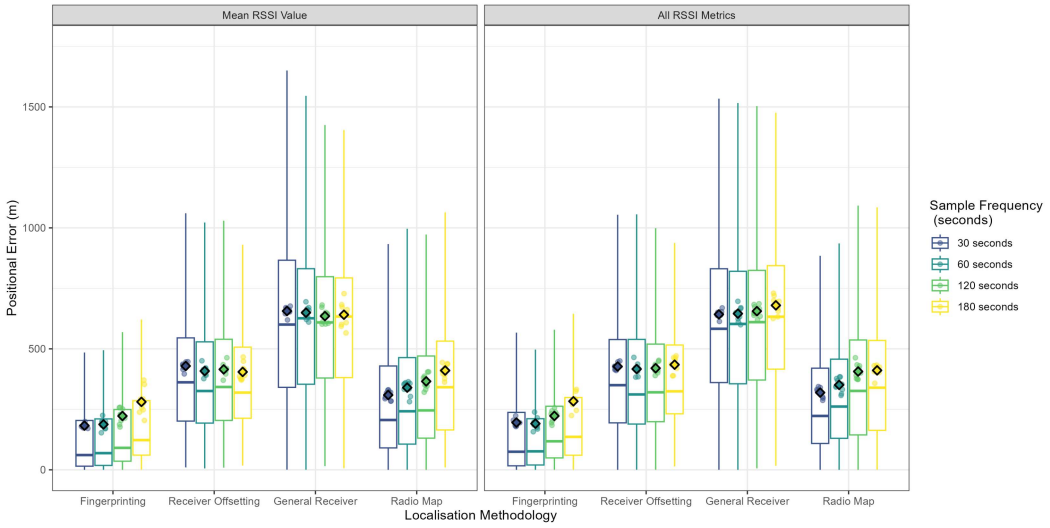

Figure 4.3: Comparison of the positional error (m) of each localisation method with different RSSI variable inputs. The lower and upper hinges correspond to the 25th and 75th percentiles, with the median indicated by the solid line. The end of the whiskers indicate ±1.5*IQR (inter-quartile range) and the diamonds mark the mean ±SEM. Colours represent the different sampling frequencies and panels display the different RSSI variable inpute (Mean RSSI, SD and count). Points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

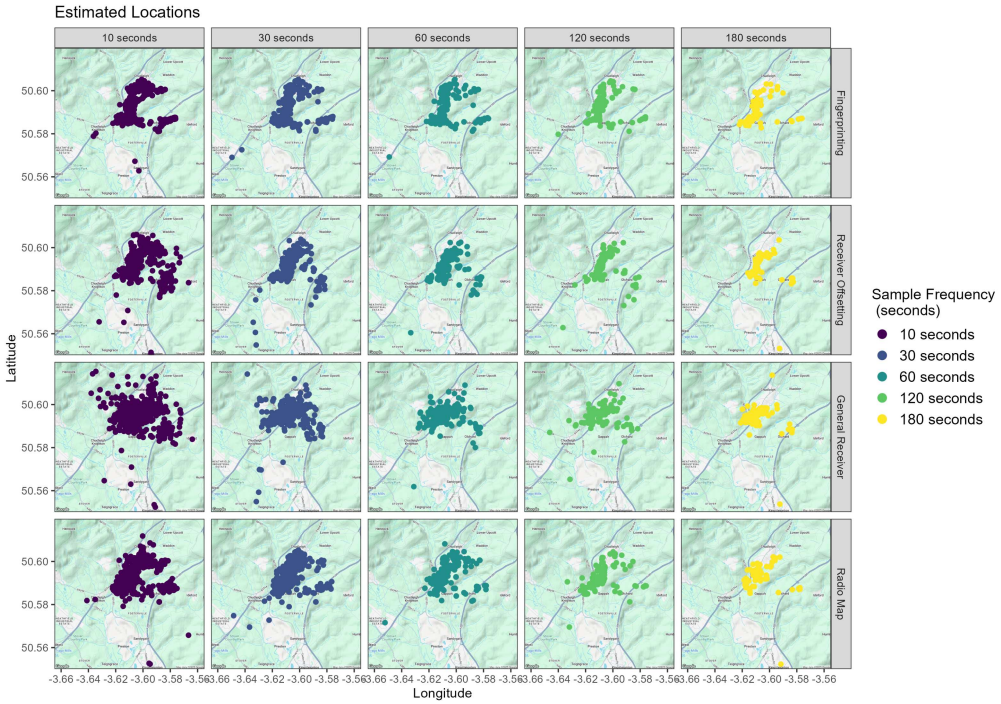

Figure 4.4: Estimated point locations of each localisation method across different sampling frequencies (seconds). Each panel shows the estimated locations for one localisation method, sampling frequency and a single fold of test data (random seed = 949). Colours represent the different sampling frequencies and the results shown in the figure are those of the 'Raw or Mean RSSI' only.

Table 4.1: Summary statistics of the positional error (metres) for the different localisation methodologies, split by both sampling frequency and RSSI variable input. Values are rounded to 1 d.p..

Method	Frequency	Inputs	Mean	SEM	Median	SD
General Receiver	10	Raw	677.3	4.6	587.7	472.6
	30	Mean	656.3	6.8	611.8	417.8
		All	642.5	6.4	590.6	398.7
	60	Mean	649.6	8.8	630.1	392.8
		All	645.6	8.9	611.1	400.7
	120	Mean	635.3	10.9	614.3	354.6
		All	655.8	12.1	619.6	394.9
	180	Mean	641.1	13.1	640.2	352.3
		All	679.8	14.5	640.1	390.6
Radio Map	10	Raw	332.0	3.1	241.3	317.9
	30	Mean	309.1	5.2	206.3	323.9
	-	All	319.2	5.2	223.8	321.9
	60	Mean	339.8	7.7	243.0	344.9
		All	350.9	7.3	263.9	325.8
	120	Mean	365.7	11.4	249.4	371.7
		All	406.4	11.1	331.9	361.8
	180	Mean	410.4	12.4	341.8	333.0
		All	411.1	12.1	341.3	327.3
Fingerprinting	10	Raw	207.3	2.8	93.8	289.4
	30	Mean	182.7	4.7	61.4	301.2
		All	195.4	5.0	75.4	307.4
	60	Mean	187.7	7.1	69.2	320.0
		All	191.0	7.1	76.6	320.3
	120	Mean	222.4	10.5	91.9	340.7
		All	222.7	9.5	119.7	309.3
	180	Mean	280.9	17.1	124.3	461.4
		All	283.3	17.0	141.8	459.0
Receiver Offsetting	10	Raw	491.6	3.7	393.4	383.3
	30	Mean	429.1	5.3	363.8	327.2
		All	426.9	5.5	353.6	339.8
	60	Mean	408.3	7.1	331.8	316.2
		All	416.6	7.9	313.5	353.4
	120	Mean	415.1	9.7	347.8	313.9
		All	420.0	10.6	328.2	344.7
	180	Mean	404.6	11.3	319.6	305.6
		All	434.6	13.0	327.5	349.4

4.3.2 Systematic Error

At the 10 second sampling frequency, the global model (which included all variables and their interactions) best described the systematic error across the four localisation methods (n=41,856, random intercept - unique point ID: variance=0.25, SD=0.50, random seed: variance=0.00, SD=0.00). Interestingly, the magnitude of each variables effect was consistently greater for the General Receiver and Receiver Offsetting methods than the Radio Map and Fingerprinting methods (Table 4.2). Of the unmeasurable variables included in the model, the mean distance from the test location to the receivers that detected the signal had a negative effect on the predicted positional error in all four localisation methods (Figure 4.5). At a mean distance of ∼518m, the mean positional errors for the General Receiver, Radio Map, Receiver Offsetting and Fingerprinting methods were 504m, 227m, 320m and 112m respectively. At a mean distance of ∼1,002m the mean positional errors increased to 604m, 279m, 440m and 146m respectively, and to 877m, 423m, 848m and 254m at ∼2000m. At approximately 2,100m the impact on the mean positional error becomes greater for the Receiver Offsetting method when compared to the General Receiver method (∼2,091: GR-907m, RO-900m; ∼2,151m: GR-928m, RO-936m).

The distance from the test location to the centre of the Node network had the greatest effect on the mean predicted positional error of the unmeasurable variables. For example, for a distance of ∼400m from the test location to the centre of the Node network, the mean positional errors for the General Receiver, Radio Map, Receiver Offsetting and Fingerprinting methods were 470m, 252m, 339m and 127m respectively. When distances doubled to ∼800m, the mean positional errors increased to 544m, 261m, 382m and 132m respectively; and increased further to 856m, 291m, 622m and 170m at ∼2,040m. The minimum distance from the test location to the receivers that detected the signal and the distance from the test location to the centre of the SensorStation network had little effect on the mean predicted positional error. Similar effects were observed in all four unmeasurable variables at the 30 and 60 second sampling frequencies, however the magnitude of observed effects varied by localisation methodologies (Figure A4.1). At the 120 second sampling frequency, only the distance from the test location to the centre of the Node network had a notable effect on the mean positional error. In addition to this variable, at the 180 second sampling frequency the minimum distance from the test location to

the receivers that detected the signal increased positional error with increasing distance (Figure A4.2).

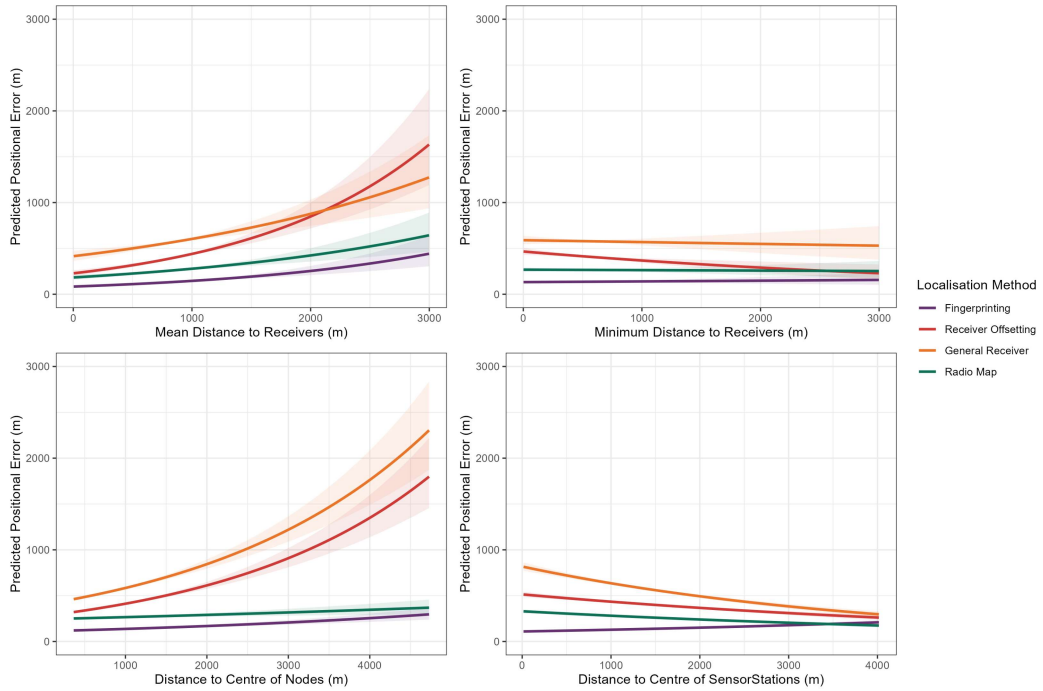

Figure 4.5: Mean positional error in relation to the unmeasurable variables across the four localisation methodologies for the 10 second sampling frequency. Each panel displays one of the four variables (Top left to bottom right: (1) the mean distance to receivers that detected the signal; (2) the minimum distance to a receiver that detected the signal; the (3) distance to the centre of the Node network and the (4) the distance to the centre of the SensorStation network). Colours represent the four localisation methodologies and 95% confidence intervals are indicated by the ribbon.

Table 4.2: Estimated regression parameters of GLMM evaluating the overall systematic error for the 10 second sampling frequency. Estimates of coefficients are on the log scale and the change in model AIC for dropping each parameter is reported. Any repeated or inappropriate terms are left blank, '-', and values are rounded to 2 d.p..

Parameter	Estimate	SE	z-ratio	ΔAIC
Intercept	4.86	0.05	96.41	-
Offsetting (Localisation method)	0.84	0.04	19.39	-
General Receiver (Localisation method)	1.47	0.04	34.39	-
Radio Map (Localisation method)	0.48	0.04	10.80	-
Number of Nodes	-1.54	0.13	-11.64	-
Number of SensorStations	-2.34	0.14	-16.76	-
Minimum distance between receivers	-0.71	0.27	-2.63	-
Mean distance between receivers	2.03	0.60	3.39	-
Maximum distance between receivers	-2.03	0.38	-5.33	-
Mean distance to receivers	1.66	0.27	6.21	-
Minimum distance to receiver	0.17	0.26	0.64	-
Distance to centre of Nodes	0.97	0.14	7.04	-
Distance to centre of SensorStations	0.65	0.09	7.51	-
Average RSSI	-0.32	0.01	-25.11	-
Number of Nodes × Number of SensorStations	3.70	0.35	10.69	-
Offsetting × Number of Nodes	0.58	0.12	5.00	-
General Receiver × Number of Nodes	0.52	0.11	4.54	-
Radio Map × Number of Nodes	2.59	0.12	21.72	-
Offsetting × Number of SensorStations	1.65	0.12	13.58	-
General Receiver × Number of SensorStations	1.19	0.12	9.94	-
Radio Map × Number of SensorStations	2.39	0.12	19.40	-
Offsetting × Minimum distance between receivers	0.68	0.23	2.90	17.41
General Receiver × Minimum distance between receivers	1.07	0.23	4.69	-
Radio Map × Minimum distance between receivers	0.90	0.24	3.70	-
Offsetting × Mean distance between receivers	-2.24	0.53	-4.25	23.96
General Receiver × Mean distance between receivers	-2.69	0.51	-5.22	-
Radio Map × Mean distance between receivers	-1.47	0.54	-2.71	-
Offsetting × Maximum distance between receivers	2.22	0.34	6.57	105.27
General Receiver × Maximum distance between receivers	2.78	0.33	8.45	-
Radio Map × Maximum distance between receivers	0.53	0.34	1.53	-
Offsetting × Mean distance to receivers	0.30	0.24	1.29	18.94
General Receiver × Mean distance to receivers	-0.54	0.23	-2.35	-
Radio Map × Mean distance to receivers	-0.41	0.24	-1.69	-
Offsetting × Minimum distance to receiver	-0.87	0.23	-3.85	15.04
General Receiver × Minimum distance to receiver	-0.27	0.22	-1.22	-
Radio Map × Minimum distance to receiver	-0.22	0.23	-0.96	-
Offsetting × Distance to centre of Nodes	0.90	0.11	7.88	215.24
General Receiver × Distance to centre of Nodes	0.77	0.11	6.89	-
Radio Map × Distance to centre of Nodes	-0.56	0.11	-4.87	-
Offsetting × Distance to centre of SensorStations	-1.33	0.07	-17.93	553.03
General Receiver × Distance to centre of SensorStations	-1.66	0.07	-22.86	-
Radio Map × Distance to centre of SensorStations	-1.29	0.07	-17.30	-
Offsetting × Average RSSI	0.10	0.01	8.30	89.24
General Receiver × Average RSSI	0.04	0.01	3.27	-
Radio Map × Average RSSI	0.01	0.01	0.73	-
Offsetting × Number of Nodes × Number of SensorStations	-2.13	0.31	-6.95	139.51
General Receiver × Number of Nodes × Number of SensorStations	-1.70	0.30	-5.65	-
Radio Map × Number of Nodes × Number of SensorStations	-3.64	0.31	-11.70	-

For the second set of models at the 10 second sampling frequency, the model that best predicted the mean positional error from an unknown location included the single three-way interaction, along with all other interactions except the mean distance between the receivers that detected the signal (n=27,328, random intercept - unique point ID: variance=0.00, SD=0.00, random seed: variance=0.18, SD=0.42; Table 4.3). For the three-way interaction involving localisation method, the number of Nodes and the number of SensorStations, varying effects were observed between localisation method (Figure 4.6). For the General Receiver method, when a single Node was used in localisation, the mean positional error decreased as the number of SensorStations increased (1 SS: 487m, 2 SS: 375m, 3 SS: 289m, 4 SS: 223m). When no SensorStations were used in localisation, as the number of Nodes increased the overall positional error decreased. When both Nodes and SensorStations were used in localisation, as the number of Nodes increased overall mean positional error increased and the magnitude of error reduction associated with additional SensorStations decreased. For example, when six Nodes were used in localisation, the mean positional errors were 528m, 482m, 440m and 402m for 1, 2, 3 and 4 SensorStations respectively. Across one to four SensorStations, the reduction in mean positional error at one Node (∼264m) was approximately twice that observed when using six Nodes (∼125m). For the Radio Map method, the opposite trend was observed when a single Node was used in localisation, with the mean positional error increasing as additional SensorStations were included (1 SS: 177m, 2 SS: 189m, 3 SS: 201m, 4 SS: 215m). Increasing the number of Nodes used in localisation overall increased the mean positional error, however the relationship with SensorStation number reversed at higher Node use. For example, when six Nodes were used in localisation, the mean positional error decreased from 457m with one SensorStation to 310m with four SensorStations.

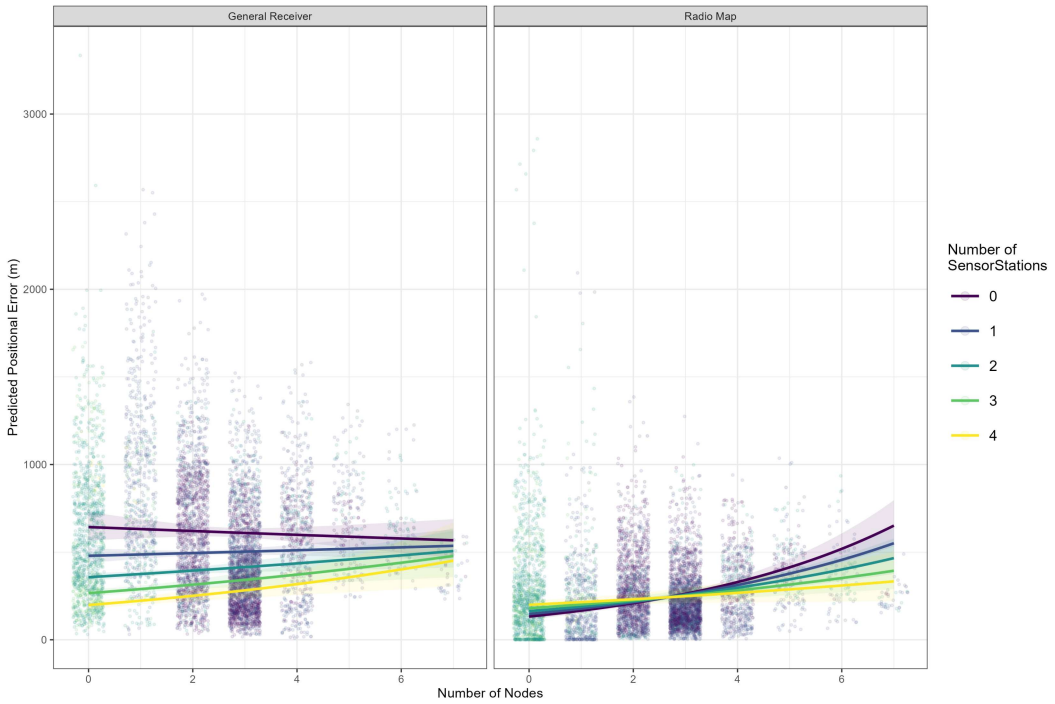

Figure 4.6: Mean positional error in relation to the number of Nodes and SensorStations for the General Receiver and Radio Map methodologies for the 10 second sampling frequency. Colours represent the number of SensorStations that detected the tag at a given location. Raw data points have been jittered for visualisation purposes and the 95% confidence intervals are indicated by the ribbon.

The minimum and maximum distance between the receivers that detected the signal had no effect on the mean positional error for any localisation method at the 10 second sampling frequency (Figure 4.7 and Figure 4.8). As the mean distance between the receivers that detected the signal increased the mean positional error increased and no difference in the magnitude of the effect were observed between localisation method (Mean distance=1.76 $\pm$0.45SE). The mean positional error decreased as the average RSSI of the receivers that the detected the signal increased, with varying magnitudes of the effect across the four methods.

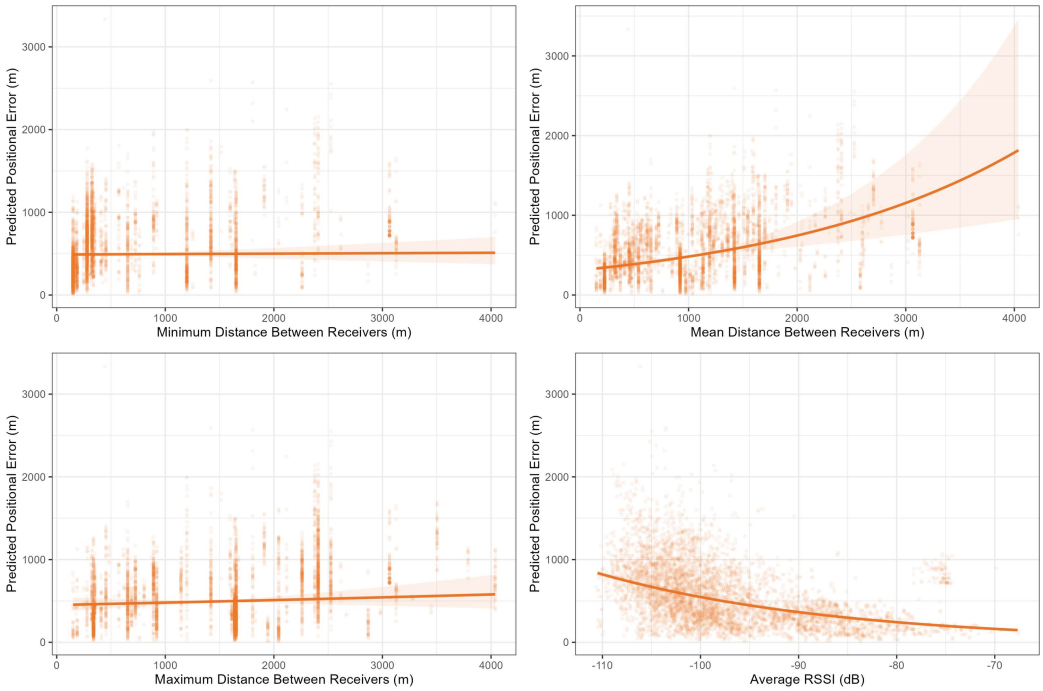

Figure 4.7: Mean positional error in relation to the variables used to predict error of an unknown location for the General Receiver method at the 10 second sampling frequency. Each panel displays one of the four variables (Top left to bottom right: (1) the minimum, (2) the mean, and (3) the maximum distance between the receivers that detected the signal and (4) the average RSSI of the receivers that detected the signal). Raw data points have been jittered for visualisation purposes and the 95% confidence intervals are indicated by the ribbon.

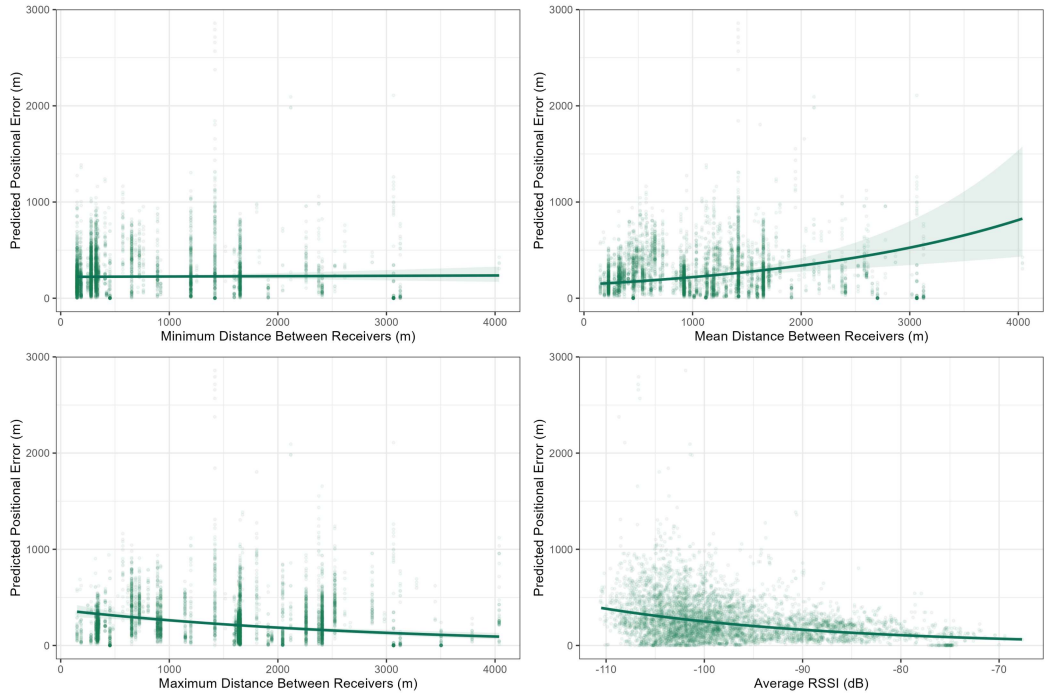

Figure 4.8: Mean positional error in relation to the variables used to predict error of an unknown location for the Radio Map method at the 10 second sampling frequency. Each panel displays one of the four variables (Top left to bottom right: (1) the minimum, (2) the mean, and (3) the maximum distance between the receivers that detected the signal and (4) the average RSSI of the receivers that detected the signal). Raw data points have been jittered for visualisation purposes and the 95% confidence intervals are indicated by the ribbon.

Table 4.3: Estimated regression parameters of GLMM evaluating the systematic error of measurable variables by multiple receivers for the 10 second sampling frequency. Estimates of coefficients are on the log scale and the change in model AIC for dropping each parameter is reported. Any repeated or inappropriate terms are left blank, '-', and values are rounded to 2 d.p..

Parameter	Estimate	SE	z-ratio	ΔAIC
Intercept	5.37	0.07	74.33	-
Mean distance between receivers	1.76	0.45	3.96	13.63
Offsetting (Localisation method)	0.15	0.06	2.44	-
General Receiver (Localisation method)	0.54	0.06	9.08	-
Radio Map (Localisation method)	-0.45	0.06	-7.26	-
Number of Nodes	-1.66	0.18	-9.25	-
Number of SensorStations	-1.01	0.12	-8.12	-
Minimum distance between receivers	-0.38	0.20	-1.88	-
Maximum distance between receivers	-1.84	0.28	-6.53	-
Average RSSI	-0.35	0.01	25.98	-
Number of Nodes × Number of SensorStations	2.81	0.35	7.94	-
Offsetting × Number of Nodes	1.46	0.15	9.86	-
General Receiver × Number of Nodes	1.54	0.14	10.66	-
Radio Map × Number of Nodes	3.25	0.15	21.64	-
Offsetting × Number of SensorStations	0.53	0.11	4.79	-
General Receiver × Number of SensorStations	-0.17	0.11	-1.56	-
Radio Map × Number of SensorStations	1.42	0.11	12.47	-
Offsetting × Minimum distance between receivers	0.39	0.12	3.29	2.05
General Receiver × Minimum distance between receivers	0.43	0.12	3.64	-
Radio Map × Minimum distance between receivers	0.45	0.13	3.54	-
Offsetting × Maximum distance between receivers	1.87	0.10	19.34	64.86
General Receiver × Maximum distance between receivers	2.09	0.09	22.06	-
Radio Map × Maximum distance between receivers	0.45	0.10	4.49	-
Offsetting × Average RSSI	0.08	0.01	6.12	52.45
General Receiver × Average RSSI	0.02	0.01	1.83	-
Radio Map × Average RSSI	0.01	0.01	0.64	-
Offsetting × Number of Nodes × Number of SensorStations	-2.16	0.33	-6.53	137.41
General Receiver × Number of Nodes × Number of SensorStations	-1.86	0.32	-5.75	-
Radio Map × Number of Nodes × Number of SensorStations	-3.89	0.34	11.49	-

For the final set of models, the three-way interaction between the localisation method, the receiver type and the raw RSSI best described the mean positional error for points that had been estimated by a single receiver at the 10 second sampling frequency (n=14,528, random intercept - unique point ID: variance=0.00, SD=0.00, random seed: variance=0.47, SD=0.69; Table 4.4). Overall, the mean positional error decreased with increasing RSSI, however the rate of this decrease varied by both localisation method and receiver type

(Figure 4.9). For the General Receiver and Radio Map methods, at RSSI values less than $\sim -105dB$ the mean positional error was greater for SensorStations than Nodes, for RSSI values greater than this threshold Nodes had greater positional error than SensorStations.

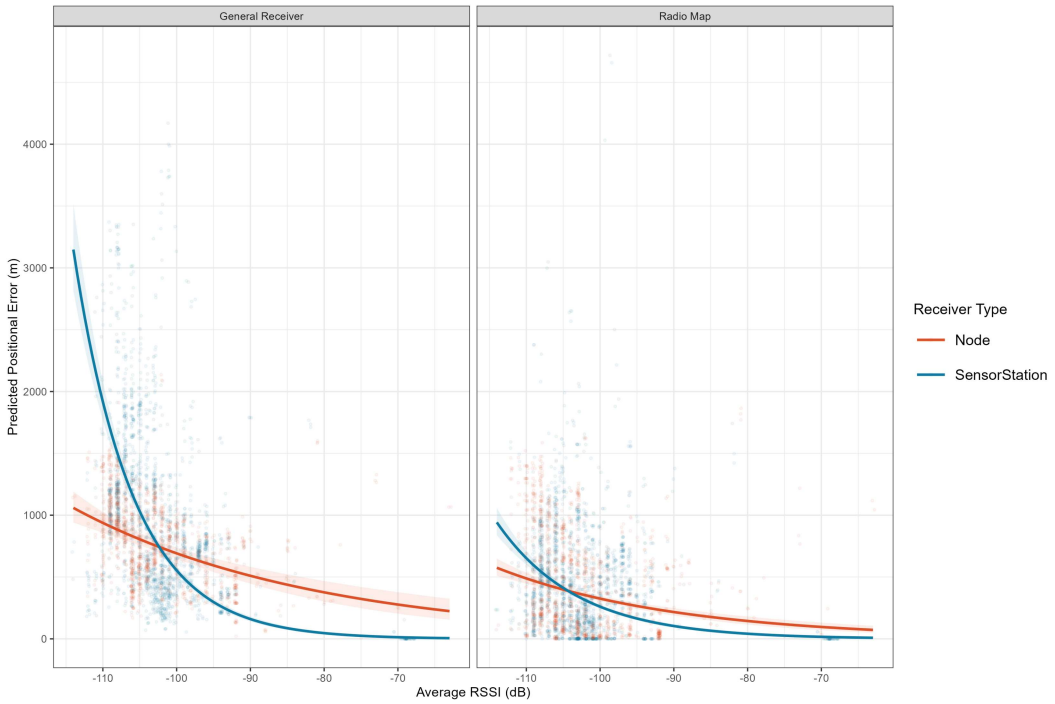

Figure 4.9: Mean positional error in relation to the receiver type and raw RSSI for two of the localisation methods at the 10 second sampling frequency. Each panel displays results for a different localisation method and colours represent the different receiver types. Raw data points have been jittered for visualisation purposes and the 95% confidence intervals are indicated by the ribbon.

Table 4.4: Estimated regression parameters of GLMM evaluating the systematic error of measurable variables by single receivers for the 10 second sampling frequency. Estimates of coefficients are on the log scale and the change in model AIC for dropping each parameter is reported. Any repeated or inappropriate terms are left blank, '-', and values are rounded to 2 d.p..

Parameter	Estimate	SE	z-ratio	ΔAIC
Intercept	5.48	0.03	170.92	-
Offsetting (Localisation method)	0.58	0.02	25.39	-
General Receiver (Localisation method)	1.14	0.02	51.08	-
Radio Map (Localisation method)	0.41	0.02	17.78	-
SensorStation (Receiver type)	0.27	0.04	6.48	-
Average RSSI	-0.21	0.03	-6.65	-
SensorStation × Average RSSI	-0.39	0.05	-8.59	-
Offsetting × SensorStation	0.09	0.03	3.08	-
General Receiver × SensorStation	-0.25	0.03	-8.81	-
Radio Map × SensorStation	-0.36	0.03	-12.13	-
Offsetting × Average RSSI	0.06	0.02	2.67	-
General Receiver × Average RSSI	0.01	0.02	0.48	-
Radio Map × Average RSSI	-0.06	0.02	-2.44	-
Offsetting × SensorStation × Average RSSI	-0.16	0.03	-4.57	86.46
General Receiver × SensorStation × Average RSSI	-0.23	0.03	-6.92	-
Radio Map × SensorStation × Average RSSI	0.05	0.03	1.41	-

4.3.3 Error handing

Filtering data using RSSI prior to localisation greatly reduced both the mean and median positional error (Figure 4.10). For the General Receiver method, the mean and median error and the variance were most improved when filtering both the SensorStations and Nodes by the -95dB threshold, but this also resulted in the largest location drop out rate losing over half of the points (5,769 points lost)(Table 4.5 and Figure 4.11). The same RSSI filter resulted in the best improvement to the mean error and the variance in the Radio Map method, however there was a marginally better improvement in the median error when filtering the SensorStations by -100dB and the Nodes by -95dB (Table 4.6 and Figure 4.12). There was a significant difference in the coefficient of variation for both methodologies, with the majority of RSSI filtering thresholds causing a reduction in variance.

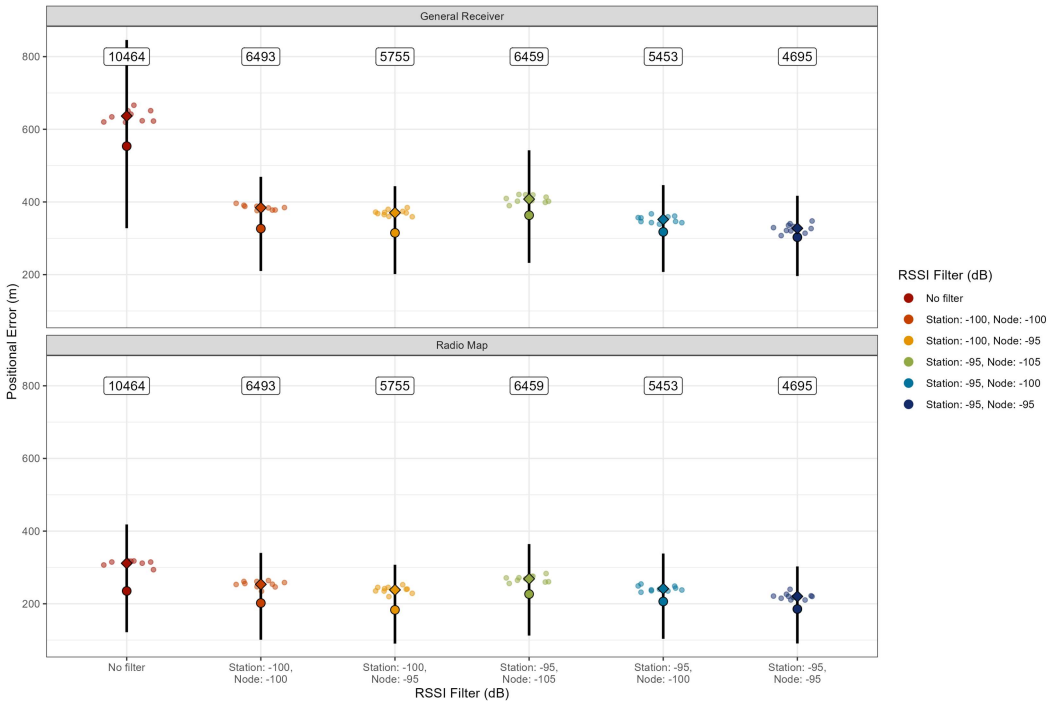

Figure 4.10: Comparison of positional error (m) of the top five performing RSSI filters. The median is indicated by the large point and the diamond marks the mean. The end of the whiskers correspond to the 25th and 75th percentiles and the number of points are labelled. Colours represent the different RSSI filters and points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

Table 4.5: Summary statistics of the positional error (metres) for the top five RSSI filters for the General Receiver methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included. Receiver types have been abbreviated to SS for SensorStation and N for Node.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	553.54	3.54	636.46	435.31	0.76	573.83	$\mathbf{p < 0.0001}$
SS: -100, N: -100	6493	326.60	2.83	384.09	291.52	0.65		
SS: -100, N: -95	5755	314.83	3.11	370.20	301.99	0.68		
SS: -95, N: -105	6459	363.47	2.44	407.77	254.02	0.82		
SS: -95, N: -100	5453	317.51	2.29	351.75	223.50	0.64		
SS: -95, N: -95	4695	302.95	2.35	327.42	214.13	0.62		

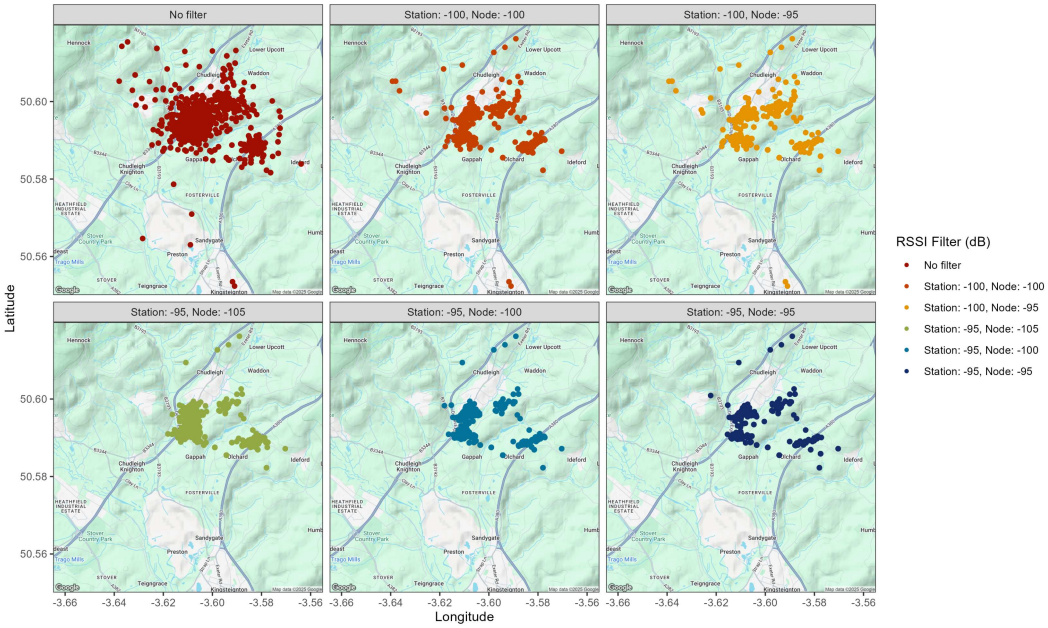

Figure 4.11: Estimated point locations of the top five performing RSSI filters for the General Receiver method. Each panel and colour represents a different RSSI filter.

Table 4.6: Summary statistics of the positional error (metres) for the top five RSSI filters for the Radio Map methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included. Receiver types have been abbreviated to SS for SensorStation and N for Node.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	235.20	2.33	311.42	286.01	1.02	653.82	$\mathbf{p < 0.0001}$
SS: -100, N: -100	6493	202.00	2.50	253.50	257.77	0.83		
SS: -100, N: -95	5755	183.17	2.70	238.22	262.30	0.92		
SS: -95, N: -105	6459	226.70	2.06	268.65	215.23	1.10		
SS: -95, N: -100	5453	206.30	1.98	241.45	193.05	0.80		
SS: -95, N: -95	4695	185.54	2.02	220.32	183.30	0.80		

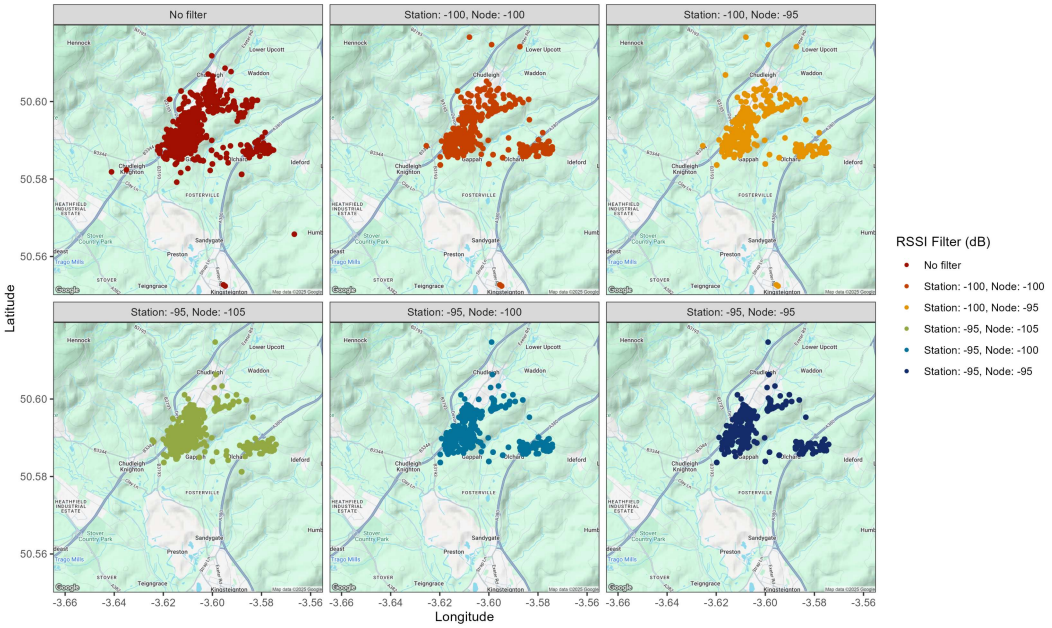

Figure 4.12: Estimated point locations of the top five performing RSSI filters for the Radio Map method. Each panel and colour represents a different RSSI filter.

Filtering data using the mean distance between the receivers prior to localisation had little effect of the mean or median positional error (Figure 4.13). For the General Receiver method, the mean and median positional error were most improved at the lowest distance threshold of 1,000m: showing decreases of 35m and 27m respectively (Table 4.7). There was no significant change in the coefficient of variation for any of the mean distance filters when used on this method. Interestingly the same mean distance filter caused the mean and median positional error to increased for the Radio Map method (Table 4.8). Significant differences were observed in the coefficient of variation, however all filtering methods increased the variance when compared to no filtering. Overall the drop out rate for the mean distance filtering method was less extreme than the RRSI based filter, with a maximum of 1,678 points lost.

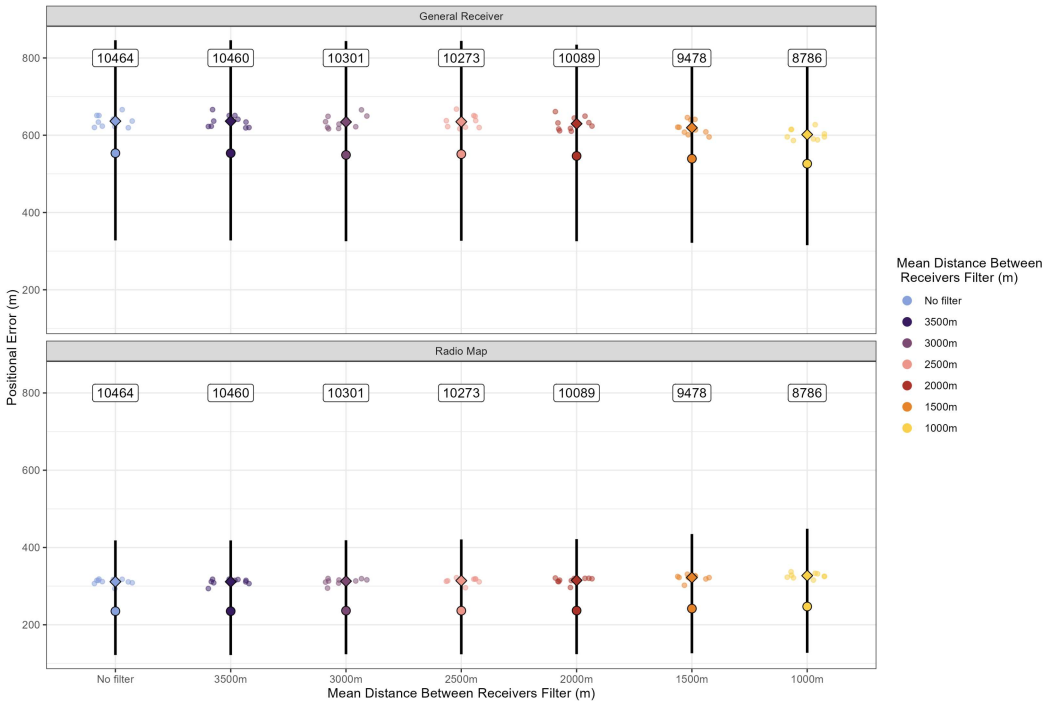

Figure 4.13: Comparison of positional error (m) of the mean distance filters. The median is indicated by the large point and the diamond marks the mean. The end of the whiskers correspond to the 25th and 75th percentiles and the number of points are labelled. Colours represent the different mean distance filters and points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

Table 4.7: Summary statistics of the positional error (metres) for the mean distance filters for the General Receiver methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	553.54	3.54	636.46	435.31	0.68	4.46	$p = 0.61$
3500m	10460	553.51	3.54	636.36	435.33	0.68		
3000m	10301	548.94	3.57	634.39	436.53	0.69		
2500m	10273	551.04	3.57	634.98	435.75	0.69		
2000m	10089	546.31	3.58	629.75	432.79	0.69		
1500m	9478	539.19	3.63	619.39	428.97	0.69		
1000m	8786	526.18	3.65	601.68	419.83	0.70		

Table 4.8: Summary statistics of the positional error (metres) for the mean distance filters for the Radio Map methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	235.20	2.33	311.42	286.01	0.92	12.57	$\mathbf{p = 0.05}$
3500m	10460	235.17	2.33	311.40	286.05	0.92		
3000m	10301	236.45	2.33	312.85	285.01	0.91		
2500m	10273	236.43	2.36	314.17	287.51	0.92		
2000m	10089	236.75	2.39	314.99	289.05	0.92		
1500m	9478	241.86	2.50	322.19	295.12	0.92		
1000m	8786	247.24	2.51	327.14	288.81	0.88		

Filtering data using the Node and SensorStation filter prior to localisation had varying effects on the mean and median positional error across the two methods (Figure 4.14). For the General Receiver method, the mean and median positional error were most improved with the 2 Nodes threshold: giving a decrease from 636m to 597m and 553m to 510m respectively (Table 4.9). Interestingly the coefficient of variation for this filtering level significantly increased, resulting in larger variability in the positional error when compared to the mean. The Node and SensorStation filter increased both mean and median positional error in the Radio Map method (Table 4.10). There was a significant decrease in the coefficient of variation, but this appears to be due to the overall increase in the mean positional error, rather than an improvement to the variability. Overall the Node and SensorStation filtering method caused no points to be lost from the dataset.

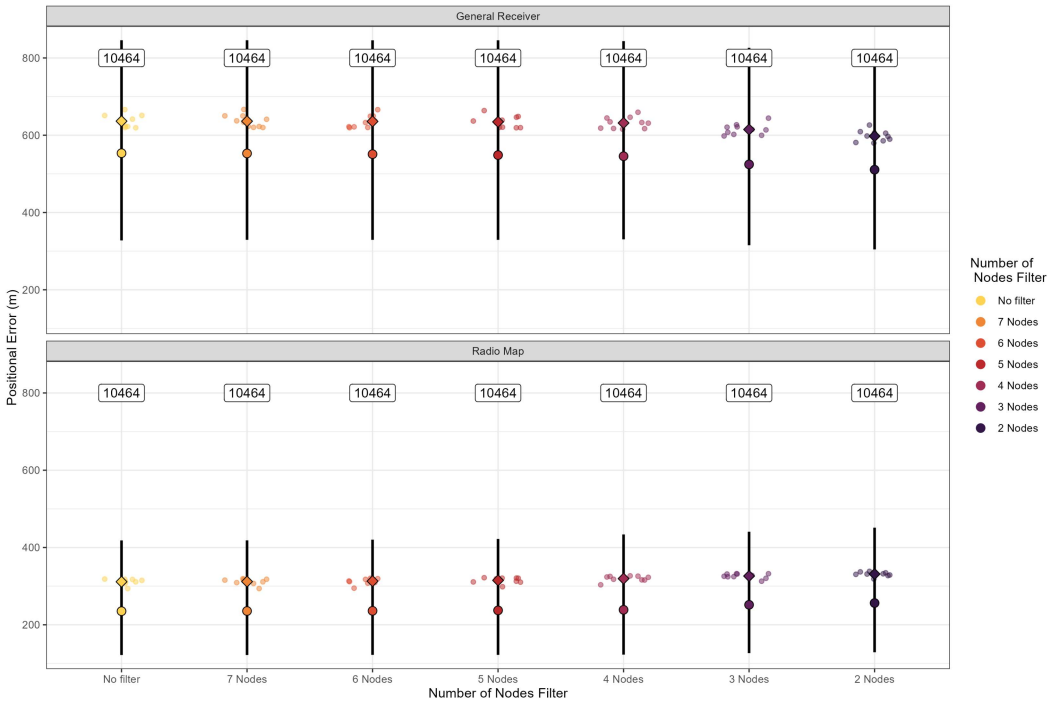

Figure 4.14: Comparison of positional error (m) of the Node and SensorStation filters. The median is indicated by the large point and the diamond marks the mean. The end of the whiskers correspond to the 25th and 75th percentiles and the number of points are labelled. Colours represent the different Node and SensorStation filters and points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

Table 4.9: Summary statistics of the positional error (metres) for the Node and SensorStation filters for the General Receiver methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	553.54	3.54	636.46	435.31	0.68	33.12	$\mathbf{p < 0.0001}$
7 Nodes	10464	552.93	3.54	636.35	435.30	0.68		
6 Nodes	10464	551.02	3.54	635.72	435.23	0.68		
5 Nodes	10464	548.60	3.54	634.60	434.50	0.68		
4 Nodes	10464	545.66	3.51	631.68	431.79	0.68		
3 Nodes	10464	524.56	3.53	614.86	434.10	0.71		
2 Nodes	10464	510.83	3.48	597.65	427.42	0.72		

Table 4.10: Summary statistics of the positional error (metres) for the Node and SensorStation filters for the Radio Map methodology. Values are rounded to 1 d.p. and results for the modified signed-likelihood ratio test are included.

Filter	N	Median	SE	Mean	SD	CV	mSLRT	p-value
No filter	10464	235.20	2.33	311.42	286.01	0.92	18.68	$\mathbf{p < 0.01}$
7 Nodes	10464	235.66	2.33	311.89	286.00	0.92		
6 Nodes	10464	236.10	2.33	312.95	286.90	0.92		
5 Nodes	10464	237.32	2.35	314.98	288.62	0.92		
4 Nodes	10464	238.71	2.37	319.48	290.89	0.91		
3 Nodes	10464	251.70	2.37	326.16	291.07	0.89		
2 Nodes	10464	256.30	2.37	331.36	291.78	0.88		

Applying the three filtering techniques simultaneously had little to no effect on the mean or median positional error across both localisation methodologies (Figures 4.15, 4.16, 4.17 and 4.18). No significant changes to the coefficient of variation were observed for any combination of filtering when all three rule-based methods were applied . Applying all three filters induced a low amount of additional location drop out rate, ranging from an additional 52 to 460 points.

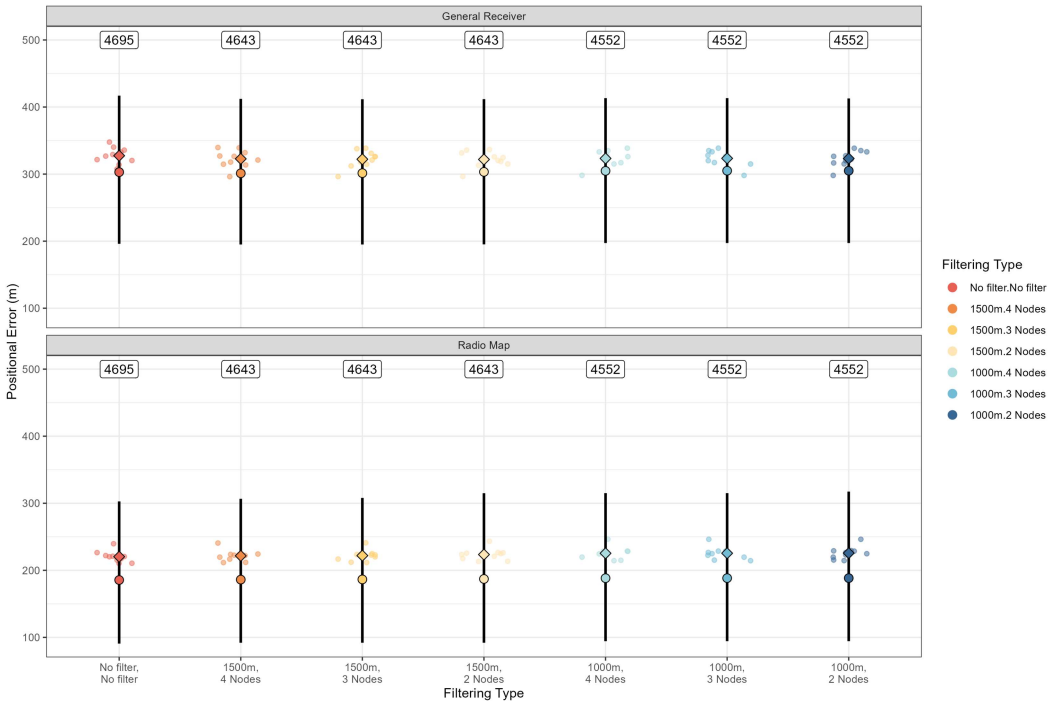

Figure 4.15: Comparison of positional error (m) of the 'SensorStation: -95, Node: -95' RRSI filter, with the top two performing mean distance filters and three Node and SensorStation filters. The end of the whiskers correspond to the 25th and 75th percentiles and the number of points are labelled. Colours represent the different mean distance-Node and SensorStation filters and points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

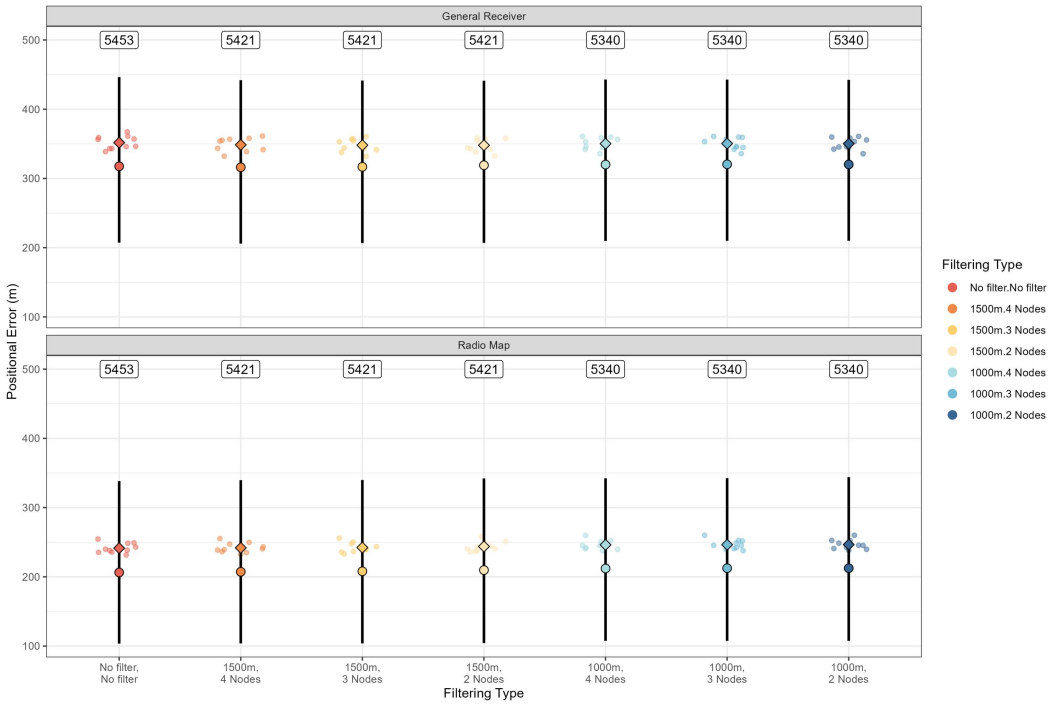

Figure 4.16: Comparison of positional error (m) of the 'SensorStation: -95, Node: -100' RRSI filter, with the top two performing mean distance filters and three Node and SensorStation filters. The end of the whiskers correspond to the 25th and 75th percentiles and the number of points are labelled. Colours represent the different mean distance-Node and SensorStation filters and points are the mean positional error per fold of test data and have been jittered for visualisation purposes.

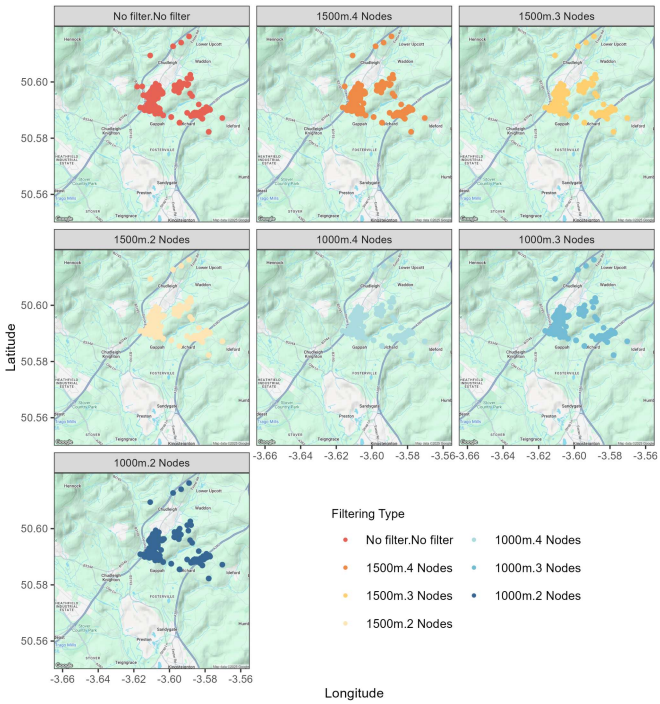

Figure 4.17: Estimated point locations of the 'SensorStation: -95, Node: -100' RRSI filter, with the top two performing mean distance filters and three Node and SensorStation filters, for the General Receiver method. Each panel and colour represents a different mean distance-Node and SensorStation filtering combination.

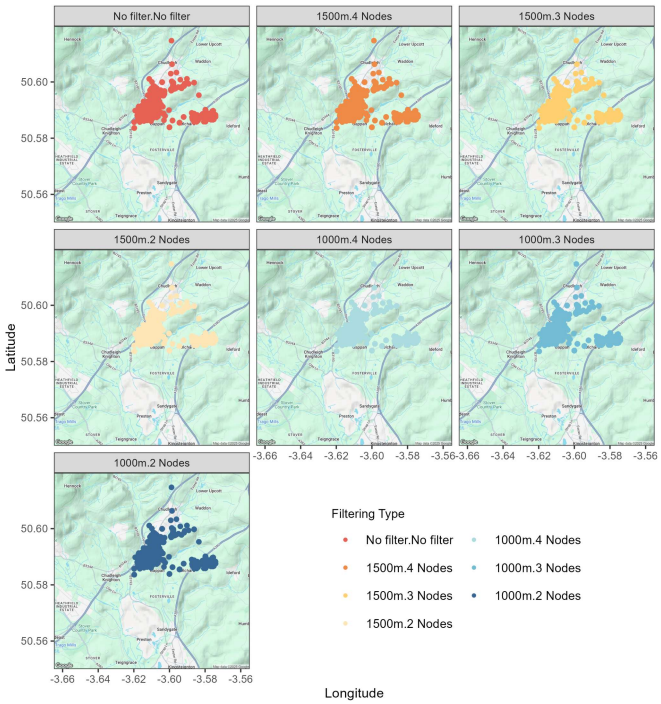

Figure 4.18: Estimated point locations of the 'SensorStation: -95, Node: -100' RRSI filter, with the top two performing mean distance filters and three Node and SensorStation filters, for the Radio Map method. Each panel and colour represents a different mean distance-Node and SensorStation filtering combination.

4.4 Discussion

Here we have presented an open-source localisation method to estimate radio transmitter positions from an automatic radio telemetry system with multiple receiver types to a high degree of accuracy. We have designed and implemented two alternative machine learning pipelines, which when used in parallel could overcome the issues associated with cost, labour-intensity and equipment failure in collecting site specific training data. We have shown how this method can be applied to an ARTS in an outdoor, real-world example, which covers a large area of both complex habitats and topography, and have additionally assessed it's accuracy in relation to other published methodologies. While it is common practice to average RSSI values when implementing localisation methods, we have given evidence that important consideration should be given when selecting a sampling frequency for the system being studies and localisation method being applied. Finally, our study gives guidance on the systematic error that is apparent in ARTSs and additionally shows how basic rule-based filtering can substantially improve location accuracy.

Recent technological and analytical advancements in the application of ARTS makes this a promising solution for tracking a range of small mammals across a variety of habitat (Cochran et al. 1965; Kays et al. 2011; Weiser et al. 2016; Taylor et al. 2011; Scardamaglia et al. 2022). In this study, we have deployed an ARTS in a real-world example to monitor the winter activity of the greater horseshoe bat in a complex landscape. We were able to simultaneously and continuously track up to 20 individual bats at once to a resolution of localisations every 10 seconds for a maximum of nearly 60 days per individual (see Chapter 2). While traditional and drone based radio tracking have their advantages in spatial flexibility, obtaining a dataset of this detail would be impossible given the logistical challenges that these methods incur (Kenward 2000; Desrochers et al. 2018; Tremblay et al. 2017). Additionally, utilising an ARTS approach ensures we can validate the efficacy of the applied methodology in our survey design, which when tracking a mammal as small and fast moving as a bat is near impossible with traditional methods. Reverse-GPS based systems produce tags of a low enough weight to track bats at a high degree of accuracy (Beardsworth et al. 2022). The localisation estimation used in these systems, such as ATLAS, are often based on time of flight measurements, which are affected by vegetation and topology to a high degree making these methods better applied in open and flat landscapes (Toledo et al. 2016).

Previous studies have utilised machine learning approaches to localise radio tags to a high degree of accuracy however they substantially differ from the present study in their habitat complexity, spatial coverage, network density, location dimensionality and ARTS design (Harbicht et al. 2017; Wallace et al. 2022; Tyson et al. 2024). In this study, Fingerprinting was the most accurate method for location estimation, performing more than 3-fold better than our General Receiver method. Whilst this indicates that our localisation method is worse performing, we would like to highlight that a site-specific radio map of immense detail is required for Fingerprinting to be an effective localisation method (Singh et al. 2015). Additionally, if there is significant alterations to the environment in which the system is deployed (such as weather and vegetation changes with seasonality), then the radio map may no longer represent the site (He et al. 2016). Although this method performed well on our training data, there are large proportions of our site that were inaccessible due to land access permissions and therefore lack any calibration data (e.g., private shooting estate, clay pit quarry). This lack of data for large pockets of space on our

study site will cause the fingerprinting model to pull estimates towards the training data in these locations, thereby resulting in reduced accuracy in these spaces (Dong et al. 2017; Li et al. 2024). As our General Receiver and Radio Map methods use indirect measures of the transmitter location relative to each receiver, instead of a direct estimate of the coordinate, these methodologies should be less impacted by this phenomenon. Additionally, locations are able to be estimated when only one receiver in the networks detected a tag at any time point with our two methodologies, which with other localisation methods is not possible (Gottwald et al. 2019; Wallace et al. 2022). We also highlight that using these indirect measures also ensures that location estimates can be obtained from receivers using different pipelines, which will reduce the requirements for training data collection when devices have to be redeployed in the system.

Averaging RSSI values prior to localisation is a commonly applied practice when processing data from an ARTS (Paxton et al. 2022; Tyson et al. 2024; Osta et al. 2024). Here we have given evidence that averaging input data is method-dependent and may not always be beneficial to location accuracy. For the two general receiver methods, mean positional error decreased when RSSI values were averaged over longer sampling frequencies, whilst the opposite pattern was observed for the fingerprinting based methods. One explanation could be that for range-based localisation methods which are based on a smooth RSSI-Distance relationship, averaging the data removes outliers and random noisy spikes, resulting in better inference (Zhu et al. 2015; Luomala et al. 2019). Whilst in contrast, free-range methods (such as fingerprinting) utilise these observed changes and features in the RSSI signal to characterise the environment at specific locations, which is used for interpolation. Averaging the data across too long a sampling frequency could be removing these important differences of the received signals into a non-informative smoothed value, resulting in lower model performance. It is additionally worth noting that as we used a moving transmitter to train and test our system, instead of stationary point tests, the error associated with the longer sampling frequencies could be up to an order of magnitude larger depending on how far an animal could move in the set time interval. For example, if a bat were travelling at 10m/s then the distance covered at our longest sampling frequency could be up to 1,800m at a single location. Due to computational limitations we only measured positional error relative to the input data and not to all possible locations within the time interval.

In addition to evaluating the accuracy of our ML pipeline as a localisation method in a real-world outdoor tracking scenario, we have assessed how the systematic error in a network can contribute to positional error. Given the theoretical relationship between signal strength and distance it is unsurprising that in our study the mean distance between the receivers and the mean distance from the tag to the receivers that detected the signal increased positional error (Bensky 2008). Further to this, we observed in our models a reduction in positional error as average RSSI increased, both of which match the findings from Osta et al. (2024). In this work we also found that the distance from the test location to the centre of the Node network had the greatest effect on the mean predicted positional error. These finding corroborates the work from Paxton et al. (2022) in a Node network in Guam, who observed increased localization error as the distance to the edge of the network decreased. Similar patterns have been observed in the Wadden Sea ATLAS project, indicating that this observation is ubiquitous across localisation methods in ARTS (Beardsworth et al. 2022). These observed relationships highlight the important considerations that must be given to receiver configuration and placement when deploying an effective ARTS as the increased receiver density improves localisation accuracy. There is however an inherent trade-off as this comes as either a cost for additional receiver installations in the network or a reduction to overall spatial coverage. ARTS receivers can vary in price depending on complexity and design, with simpler models costing as little as USD$375 to more complex designs costing around USD $6,000 (Ripperger et al. 2020; Cellular Tracking Technologies 2024).

Of the four localisation methodologies that were tested in this study, our General Receiver method was the worst performing, reporting a mean positional error between 635m - 677m. Applying our RSSI-rule-based filter, improved mean localisation accuracy by up to 2-fold in this methodology, making estimates nearly match those of fingerprinting. Similar improvements were not observed in our Radio Map method and data drop out rate at most stringent filter resulted in over a 50% loss of data. We found it surprising that despite the relationships between distance to and between receivers and also in number of Nodes and SensorStations used in localisation, our filtering approaches based on these aspects had little effect on the positional error. Improvements of only 25m and 40m in the mean error were observed at the highest filtering threshold for each method respectively, with little change in the median which indicates that these methods likely target more

extreme errors. We stress that careful consideration and validation should be implemented before applying these filters to data collecting from an ARTS as differences in design, spatial configuration and localisation method may give differing results.

Positional error could be improved in our methodology by implementing alternative approaches to location estimation, such as the use of a weighted averaging, or by implementing additional models or filters post location estimation (e.g., state-space modelling, splines, weighted averaging, grid search, Kalman Filter) (Subedi et al. 2019; Wang et al. 2020; Ripperger et al. 2020; Burcher et al. 2025; Rueda-Uribe et al. 2024). Other machine learning models or ensemble modelling could be applied to improve location estimates in any of the four methods that were tested in this study (Adege et al. 2018; Alhomayani et al. 2020; LeDell et al. 2020; Janssen et al. 2020a). Due to its widespread use in both indoor and outdoor localisation method; ease of implementation and training; and low computation time, we opted to utilise kNN algorithm in the development of our methodology (Janssen et al. 2018). Although this algorithm may give lower positional accuracy in comparison to other more complex methods, given that a lack of technical support is an often reported barrier in conservation, we believe the simpler nature of kNN could be of benefit in developing an open-source localisation tool that could be used by non-specialists (Hahn et al. 2022).

Collecting effective training data which well represents the target species is extremely important for model accuracy when deploying an ARTS. Radio transmission power has been shown to be altered by animal body mass, attachment methodology, antenna orientation, animal mobility (e.g., speed, irregularity) and behaviour (e.g., flight height, perching), making these important aspects to mimic in training data collection (Naef-Daenzer et al. 2005; Levin et al. 2015; Bircher et al. 2020; Rueda-Uribe et al. 2024). Manual radio tracking of tagged animals has been used to collect calibration data for ARTS, but for many species this is unfeasible due to complex behaviours and lack of ability to verify true locations (Osta et al. 2024; Montgomery et al. 2010). Alternatively using an artificial substitute for the animal is common practice, with methods including attaching transmitters to PVC tubing; saline filled water bottles, balloons and rubber gloves; and on occasion a dead animal (Naef-Daenzer et al. 2005; Taylor et al. 2011; Janaswamy et al. 2018; Krull et al. 2018; Gottwald et al. 2023). Point testing is often the go to approach

when calibrating an ARTS, but this can be a time consuming process even for small study sites and we argue that this does not effectively mimic the more inconsistent radio signals that are produced from a moving target (Tyson et al. 2024; Wallace et al. 2022; Paxton et al. 2022; Tran et al. 2024). In this study, we carried out a short preliminary field test prior to training data collection and we observed substantial differences in RSSI (up to 10 dB) when radio transmitters were and were not attached to our animal substitute (here a medium sized plum) before mounting to the top of the bamboo cane (K. Allan). Despite the importance of these aspects in training data collection there has been little empirical testing on the effectiveness of substitutes and many recommendations are based on anecdotes, we therefore recommend future work be carried out in this area.

Reverse GPS systems, such as ATLAS, can track small animals at high spatial-temporal resolution to accuracy's of under 5m (Beardsworth et al. 2022). Whilst our lowest reported mean positional error of 182m (Fingerprinting at 30 second sampling frequency) does not match the accuracy of other available tracking systems, we highlight that the cost effectiveness of this system in a complex environment may be of a better practical application sense. ATLAS base stations cost an estimated 4 to 5-fold more than the Cellular Tracking Technologies SensorStations and some 30-fold more than the Nodes that were used in the current study (Bijleveld et al. 2021; Cellular Tracking Technologies 2024; Cellular Tracking Technologies 2022). Implementing such a cost for monitoring applications in conservation or development is highly unfeasible for many projects and a team of technical experts are required to install, operate and manage both the hardware and data analyses. The Wildlife Biologging Network utilises a RSSI-based Direction-of-Arrival (DOA) localisation method, which has been reported to achieved a mean positional error of 7.3m when tracking mouse-eared bats in deciduous forest in Bavaria, Germany (Duda et al. 2018; Ripperger et al. 2020). To obtain the level of Node density and positional accuracy that was used in the study, would require the deployment of over 10,000 localization nodes to cover less that a 3rd of our study site. In a real-world tracking study, many locations are inaccessible due to terrain, access permission and habitat type, making a deployment of this magnitude unfeasible, even if monetarily possible.

Whilst we have shown that an ARTS can be used in a real-world scenario to track a small vertebrate in an anthropogenically altered landscape, our network was far from a per-

fect array. We experiences both logistical and technical challenges throughout this study and stress that multi-disciplinary collaboration from a range of technical backgrounds is of extreme benefit when utilising this method. Working in both a topologically and environmentally challenging landscape made finding locations for suitable receiver deployment often difficult. In addition to this, much of England's landscape is privately owned with limited right of way in accessing land (Evans 2019). This gave an additional barrier of landowner permission and much more consideration of equipment safety when deploying receivers in the landscape (1 Node was stolen which was deployed in a public park). We found that required maintenance of the hardware in the system was more frequent than we had anticipated, with continuous supervision and trips to recharge batteries; replace solar panels; repair broken aerials and masts; and fix failing equipment being required. Due to local planning regulations our deployed receivers had a particular criteria they were allowed to meet in terms of permanence, which inherently increased our supervision needs. Utilising a more stable structure or framing to deploy our receivers would have reduced these requirements. During our study we experienced a firmware failure which altered the RSSI-values that were reported by the SensorStations, which we stress can be avoided with effective preliminary testing prior to system deployment. The Nodes in this network utilise GPS to set their internal clock and during this study we experienced frequent synchronisation failures with these devices, mostly related to battery supply and lack of sunlight, which resulted in timestamp mismatches and loss of data. We recommend that the use of stationary beacons deployed in an ARTS could both prevent mishaps in device synchronisation but also assist in radio map updating (Cassano et al. 2009; He et al. 2016).

Machine learning-based methods are proving to be a powerful tool in wildlife tracking and in combination could provide greater insight into animal behaviour (Schofield et al. 2018; Nathan et al. 2022; Gottwald et al. 2023; Tyson et al. 2024). In the correct context, fingerprinting is a versatile and highly accurate method and in this study it outperformed all the other tested localisation methodologies. Time and money are the inherent obstacle in wildlife conservation and research, which are both depleted through additional labour and hardware requirements with increased training data collection. The general receiver approach that we utilised in this study along with basic rule-based filtering has shown that effective location estimates in a real-world tracking scenario can be created which

could overcome some of these barriers. Given the growing concerns surrounding the environmental impacts associated with machine learning and AI (Lacoste et al. 2019; Hamdan et al. 2024), we recommend that future researchers considers the spatial accuracy and extent that is required to answer the ecological question at hand and that simpler localisation methods should first be considered if they are appropriate. Overall, we believe that automated radio tracking systems have immense application in both the field of movement ecology and in practical applications for both conservation and development. We recommend that future research should aim to improve localisation methods in real-world ARTS to better develop these methodologies in applied spaces. The production of easy to use, open-source pipelines along with multi-disciplinary collaboration are inherent to the continued development of this powerful tool.