

# A Noise Robust Automatic Radiolocation Animal Tracking System

Liang Wang (✉ [l.wang@csiro.au](mailto:l.wang@csiro.au))

CSIRO: Commonwealth Scientific and Industrial Research Organisation <https://orcid.org/0000-0003-3529-1680>

Foivos Diakogiannis

UWA: University of Western Australia

Scott Mills

James Cook University

Nigel Bajema

James Cook University

Ian Atkinson

James Cook University

Greg Bishop-Hurley

Commonwealth Scientific and Industrial Research Organisation

Ed Charmley

Commonwealth Scientific and Industrial Research Organisation

---

## Research Article

**Keywords:** radiolocation, machine learning, encoder/decoder, convolutional neural network

**Posted Date:** February 19th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-195951/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Animal Biotelemetry on August 1st, 2021.  
See the published version at <https://doi.org/10.1186/s40317-021-00248-w>.

## RESEARCH

# A noise robust automatic radiolocation animal tracking system

Liang Wang<sup>1,2\*</sup>, Foivos Diakogiannis<sup>3,4</sup>, Scott Mills<sup>5</sup>, Nigel Bajema<sup>5</sup>, Ian Atkinson<sup>5</sup>, Greg J. Bishop-Hurley<sup>2</sup> and Ed Charmley<sup>1</sup>

This document is the result of a research project funded by Advance Queensland Innovation Partnerships (AQIP) - Smart Ear Tag for Livestock, 2016.

## Abstract

Agriculture is becoming increasingly reliant upon accurate data from sensor arrays, with localization an emerging application in the livestock industry. Ground-based Time Difference of Arrival (TDoA) radiolocation methods have the advantage of being lightweight and exhibit higher energy efficiency than methods reliant upon Global Navigation Satellite System (GNSS). Such methods can employ small primary cell batteries, rather than rechargeable cells, and still deliver a multi-year deployment. In this paper, we present a novel deep learning algorithm adapted from a one-dimensional U-Net like a convolutional neural network (CNN) model, originally developed for the task of semantic segmentation. This model both converts TDoA sequences directly to positions and reduces positional errors introduced by sources such as multipathing. We have evaluated the model by using simulated animal movements in the form of TDoA position sequences in combination with known distributions of TDoA error. When errors with a standard deviation of 50 m and 100 m are added to simulated TDoA transmissions the model is able to reduce this error to 22 m and 27 m (RMSE) respectively. Without correction, the standard deviation of these errors is on the order of 90 and 200 m. Accordingly, the model can reduce the error by greater than 80 m (> 80%), demonstrating the effectiveness of this novel 1D CNN U-Net like encoder/decoder for error correction of TDoA position estimates.

**Keywords:** radiolocation; machine learning; encoder/decoder; convolutional neural network

## Introduction

The development and implementation of precision farming practices are enabled by location-aware platforms. Such platforms can track assets across holdings enabling more efficient management strategies.

Animal tracking has increasingly become an active area of both research and applied innovation. Trade-offs exist between the types of geolocation employed including price, precision, accuracy, power consumption, and the frequency of position updates. Geolocation systems include Angle of Arrival (AoA), doppler approaches, Power on Arrival (PoA), Time of Arrival (ToA) and Time Difference of Arrival (TDoA).

The oldest forms of radio-tracking animals used AoA to triangulate the position of individuals fitted with a radio transmitter with the first publications appearing in the 1960s, such as work conducted on the early summer activities of porcupines [15]. Modern versions of this technique estimate the AoA by comparing the amplitude variation of an antenna array at a signal receiver and typically six or more antennas are used for this purpose. These techniques can yield very low power use depending upon the duty cycle and strength of the transmission.

Satellite tracking of wildlife has a long history starting with Craighead Jr et al [6] tracking elk via the Nimbus meteorological satellites and a bulky 11.3 kg collar in April of 1970. However, it was the creation of the ARGOS system in 1978 [5], utilizing the doppler shift of a carrier frequency over successive transmissions, that initiated the first generation of relatively accessible animal satellite tracking devices. Early deployments using ARGOS to track wildlife include Priede [20] tracking Basking sharks in June 1982, Jouventin and Weimerskirch [11] tracking wandering albatrosses in 1989.

The ARGOS system employs a repetition period between two consecutive payloads, of between 45 to 200

\* Correspondence: [l.wang@csiro.au](mailto:l.wang@csiro.au)

<sup>1</sup> Agriculture & Food, CSIRO, 4814 Townsville, Australia

Full list of author information is available at the end of the article

s for as little as 360 milliseconds (PPT-A3 Argos Specification), to estimate the location of a platform transmitter terminal. The accuracy of this system is based upon seven location qualities ranging from 150 m to tens of kilometers, with a low power consumption using as little as 0.8 joules per location estimate (assuming two ARTIC R2 transmissions plus amplification to 0.5 Watt).

Power of Arrival (PoA) localization methods rely upon the received signal strength at a minimum of two receivers. Implementation of PoA in LPWANs (Low Power Wide Area Networks) over long distances is not practical due to the inverse square law and signal attenuation, where the received signal strength quickly becomes too weak for the receiver to meaningfully differentiate small changes in received signal strength. These techniques are most suitable for factory-scale localization.

Time of Arrival (ToA) localization methods use the time of reception of signals received by a roaming device from multiple transmitters of a known location, the most ubiquitous implementation of this technology is Global Navigation Satellite Systems (GNSS). Efficient implementations of GNSS networks, such as those using the Ublox Zoe-M8B, require around 1.8 joules to acquire a location from a cold start, this signal then needs to be re-transmitted adding additional energy overhead. For many animal tracking systems using GNSS, it is the single largest power drain on the system. The advantage of GNSS based systems is the lack of ground-based infrastructure, however they require expensive spaced based infrastructure. Global Navigation Satellite Systems yield excellent spatial fidelity of around 2.5 m, with some systems achieving cm accuracy. Modern approaches to ToA use LPWAN to unload the on-device processing to remote services over communication protocols such as LoRaWAN, for example, Kolmostar's<sup>[1]</sup> JEDI-200 module.

Time Difference of Arrival (TDoA) localization methods require no on-device processing and only short transmission bursts, positions are estimated by examining the time difference of a transmitted signal arriving at multiple fixed time-synchronized receivers. In this paper, we have chosen to examine Taggle's<sup>[2]</sup> proprietary TDoA localization system, each location requires only 0.12 joules of energy making it suitable for tracking solutions using non-rechargeable batteries, allowing for a greater number of location transmissions per energy consumed. Taggle's radio transmitter operates in a frequency range of 912-927 MHz

using direct sequence spread spectrum modulation at a power output of 14 dBm. Its high capacity receivers can accommodate at least 14,000 device transmission per hour. Along with a message header, single transmissions can hold 12 to 15.5 bytes of user data with a transmission duration of around 300 ms. The theoretical TDoA localization accuracy of the Taggle's system is approximate:

$$accuracy = \frac{speed\ of\ light}{4 \times bandwidth} \approx 5m \quad (1)$$

However, the clock synchronisation of the receivers is  $\pm 20$  nanoseconds resulting in a potential error of 12 m. Menzies et al [16] conducted a small-scale trial with Taggle's localization system using twelve Taggle tags on a plot whose size was approximately 5 ha. They found that the positions had a mean precision of  $\pm 22$  m with an SD of 49 m. Ground-based TDoA location systems can experience substantial noise due to multipathing where the time differences are exaggerated due to signal paths that are not line-of-sight, this is the kind of error we seek to address in this study. The examined localization method uses TDoA among four fixed receivers to estimate the origin of the transmission, the two most basic analytical methods are the Taylor series method [17] and Chan method [3] that can localize objects with minimal error in the absence of multipathing. To address noise in TDoA location estimates machine learning-based methods are introduced for localization in an arbitrarily complex system, current approaches to this problem falls into two main methodologies.

The first uses fingerprint references and machine learning models to derive insights about the geometrical structure of the environment which can provide information about TDoA error and improve localization accuracy. de Sousa and Thomä Electronic [7] applied the Random Forest algorithm embedded in a machine learning framework to extract a reference data set of TDoA fingerprints in outdoor scenarios. In the experiment, four TDoA sensors were deployed in an area of 2 km<sup>2</sup> in the City of Ilmenau in Germany, representative of a typical suburban scenario with small buildings and spaced streets. The empirical cumulative density function (CDF) used in the experiment showed 210 m of error for 65% of location estimates, compared to 300 m for the raw TDoA calculations. Similarly, Alonso-González et al [1] implemented a neural network model to estimate the positions for TDoAs using an indoor fingerprint approach to predict a transmitters positions in a 3D environment. They tested their model in a 4 × 4 × 3 m room, their experimental results indicated a substantial improvement in accuracy, with a best average error of 390 μm.

<sup>[1]</sup>Kolmostar 48531 Warm Springs Blvd, Suite 407, Fremont, CA 94539 ([www.kolmostar.com](http://www.kolmostar.com))

<sup>[2]</sup>Taggle Systems Pty Ltd. Sydney, N.S.W. 2000, Australia ([taggle.com](http://taggle.com))

The second approach applies denoising neural networks to reduce the TDoA error to improve localization accuracy. Wu et al [27] proposed a Radial Basis Function (RBF) neural network to improve localization accuracy. They tested the model on simulated TDoAs within a 2D  $500 \times 500$  m space with seven receivers, the Root Mean Square Error (RMSE) of the localization was 17 m while the Chan algorithm leads to an  $\approx 30$  m RMSE. Zhang et al [29] presented a novel localization algorithm based upon a neural network ensemble to estimate the positions of objects in indoor multipathing environments. The ensemble method was tested on the simulated TDoA in a 2D  $60 \times 60$  cm space. The best RMSE of localization was  $< 1$  cm and the ensemble method also showed better generalization and stability than a single neural network.

These approaches demonstrate the utility of machine learning models for localization in reducing the initial TDoA error or for correcting location estimates from noisy TDoAs. In this work, we present a novel denoising 1D convolutional neural network. The denoising encoder/decoder we propose has many similarities with a denoising autoencoder, however, the autoencoder lacks the skip connections we employ. Denoising autoencoders are an extension of simple autoencoders and were originally invented to reduce the risk of overfitting [2, 26]. Denoising autoencoders can be applied to remove the effect of stochastic noise to inputs, for example, to clean the noise from corrupted images. Convolutional layers emphasize local features of structured data, such as those evident in images or sequence data. The information from these local features help the model to reduce noise from TDoA sequences and their associated movement sequences. Diakogiannis et al [8] proposed a novel deep learning framework for semantic segmentation of remotely sensed data, this framework consisted of stacked CNN layers in a U-Net-like backbone [23]. We propose a denoising encoder/decoder algorithm based on this framework with 1D CNN layers. The algorithm is a deep learning approach for TDoA localization error correction, using noisy TDoA tracks to correct for multipathing. The performance of this algorithm was tested on simulated animal track TDoA sequences with added noise derived from real-world data. The results show that this algorithm can recover animal tracks from noisy TDoAs.

The remainder of the article is organized as follows. In Section , we discuss the problem of TDoA localization in terrestrial systems. Section describes the model architecture and the methodology of the experiment. Section describes the animal movement simulation and the method for generating the TDoA data. The final section (Section ) presents the performance

of the developed algorithm in comparison to the non-corrected TDoA location estimates.

## Problem overview and formulation

Classical TDoA localization methods assume that radio signals travel without obstruction in line-of-sight with localization solutions based on solving the following hyperbolic equations:

$$TDoA_{ij} = ToA_i - ToA_j, \quad (2)$$

Where  $ToA_i$  is the Time of Arrival to the  $i$ th receiver, defined as

$$ToA_i = \frac{1}{c} \sqrt{(x - X_i)^2 + (y - Y_i)^2}. \quad (3)$$

Here  $X_i$ , and  $Y_i$  are the coordinates of the  $i$ th receiver station, while  $x$  and  $y$  are the coordinates of the transmission origin track, and  $c$  is the speed of light in air.

The height of receivers and transmitters in TDoA networks can vary adding systematic error, however, in most practical cases the introduced error is negligible. For instance, the error introduced by a 100 m elevation across 1 km is  $< 5$  m. Of most concern is the occlusion of line-of-sight between transmitters and receivers due to vegetation, topography, or man-made objects. Such obstructions can lead to increased path length, and hence the time of arrival, between a transmitter and receiver. This multipathing phenomenon can add significant error to TDoA localization estimates.

In multipathing scenarios signals always travel along a longer path than line-of-sight, therefore the error of ToAs is always positive, but the error for TDoAs, according to the Equation 2 can be either positive or negative. Since the real path of each signal traveled corresponds to a unique multipathing scenario, it is not possible to predict or model the error of TDoA from a single transmission.

To get an estimate of a real-world TDoA error distribution we placed a single static transmitter, tag ID is 130114, in Warina Park, Townsville Australia (Lat. -19.279, Long. 146.771) and made 2215 localization transmissions at two-minute intervals (Approximately 3 days). The TDoA was examined by looking at two receivers in Townsville's Taggle network, towers taggle-058 and taggle-067. The resulting error distribution in meters is shown in Fig. 1, using a bin size of 8 m. The noise is normally distributed and the problem can be framed as reducing Gaussian noise from TDoA measurements.

## Model framework

This section gives an overview of the architecture of the model **ResUnet-1d** (section ) which reduces the

**Figure 1** Histogram (bin 8 m) of sampled error distribution of TDoAs measured on a static reference tag (tag ID: 130114) deployed in Warina Park, Townsville, Australia. A total of 2215 transmissions were used to estimate the error distribution using two Taggle receivers (tower ID: taggle-067 and taggle-058). The fitted distribution demonstrates the Gaussian nature of the error with an  $\bar{X} \sim 0$  and  $\sigma \sim 100$ .

localization errors of animal tracks. The **ResUnet-1d** model combined two tasks, converting the TDoAs to positions and localization denoising. Section introduces the process of model training and the use of the trained model.

### Architecture

In this study, we implement a modified 1D version of the **ResUnet-a** model [8] that is designed for semantic segmentation of mono temporal very high-resolution aerial images. **ResUnet-a** uses a UNet encoder/decoder backbone and residual building blocks with atrous convolutions for feature extraction. In the middle and at the end of the network, a pyramid scene parsing pooling layer is implemented. The network implements a conditioned multitasking approach, estimating the semantic classes, their boundaries, and their distance transforms.

This model was chosen as it has a few advantages for the problem of TDoA positioning and localization error reduction. Firstly, the U-Net backbone architecture is recognized in the field of computer vision for achieving state-of-art image denoising [13, 14]. Secondly, the residual connections [9] allows for the efficient gradient propagation in deep architectures, thus guaranteeing fast convergence and improved performance. Heinrich et al [10] integrated ResNet into the fully-convolutional neural network (FCN) with U-Net architectures for Low-Dose Computerized Tomography (CT) image denoising, showing that U-Net combined with ResNet yields the most promising result with an enhanced peak signal to noise ratio. Therefore, we have migrated this successful framework from the domain of image denoising and applied it to our TDoA animal tracking problem. The architecture of the framework is shown in Fig. 2.

In the proposed model, **ResUnet-1d**, we introduce several changes to the original **ResUnet-a** that make it suitable for our application. Firstly, the TDoAs and the positions are multi-channel one-dimension time-series data, the 2D convolution layers in the model are replaced by 1D convolutions. Secondly, compared with semantic segmentation tasks, the time series denoising tasks should be simpler in both the input data format and the difficulty of the tasks. Therefore, in **ResUnet-1d**, the encoder part only consists

of three **ResBlock-a** building blocks followed by a 1D **PSPPooling** layer. Each feature extraction unit is a standard residual unit (we did not use atrous convolutions). This shallower model can potentially prevent overfitting issues while reducing the computational burden. Lastly, as **PSPPooling** doesn't perform well on regression problems [8], the last **PSPPooling** layer is replaced by an attention block which is embedded in the **HEAD** block for increased performance, details of this block are illustrated in Fig. 3.

### Methodology

As the received TDoAs of each transmitter denote time-series data, we needed to segment the continuous time-series data into fixed-length sequences. The proposed model expects 256 records in one piece of track data. But it is a free parameter determined by the specific task, the only requirement is that the length of the training time series tracks should match the length of the tracks on which the model will be applied.

The input data is a sequence of noised TDoAs, with the shape  $256 \times N_t$ , where  $N_t$  is 3, which is the number of the TDoAs at each transmission. The simulated ground truth animal tracks, that were used to generate the noisy TDoAs are the ground truth values that the algorithm is trying to recover. The shape of the output track is  $256 \times 2$ , as each position only has two coordinates,  $x$ , and  $y$ . The output of the model can predict denoised tracks by using noised TDoAs as an input, hence reducing TDoA multipathing error.

### Data simulation and preprocessing

Within the field of animal movement behavior, the modeling of movement data is implemented in many ways.

Quaglietta and Porto [21] introduced an algorithm, SimRiv, to simulate individual-based, spatially-explicit movements in river networks and heterogeneous landscapes. In this study, we simulate a cows' movement using this approach on a totally homogeneous landscape.

### Animal track modeling

Animal movements are considered to be Brownian motion and multistate. The main states of the movement include random walking, correlated random walking, and rest. The random walk state is a random movement state, in which the direction of the steps is completely independent. The correlated random walk means the direction taken in one step by an individual animal should be correlated with the direction of the previous step [22, 25]. The correlation, which is the turning angle concentration parameter of the wrapped normal distribution, has a value between  $[0, 1]$ , where 0



**Figure 2** ResUnet-1d architecture. The left (downward) branch is the encoder. The right (upward) branch is the decoder. Conv1D is the standard 1D CNN layer, and Conv1DN is the standard 1D CNN layer with batch normalization. The  $B$  in data size represents the batch size, and the  $N_{\text{time}}$  represents the sequence length of the input data.

**Figure 3** The architecture of the HEAD block. The two inputs of this block are the outputs of the first Conv1DN in the encoder branch named *First* and the output of the final ResUnit in the decoder branch named *Final*. The  $B$  and  $N_{\text{time}}$  have same meaning as in Fig. 2.

means there is no correlation between two steps (yielding a random walk state), and 1 means the direction does not change. In this study, we use 0.98 (as chosen by Quaglietta and Porto [21]) as the value of correlation. The resting state corresponds to a state where the individual animal does not change position.

Following Quaglietta and Porto [21], we assume the cows are Lévy-like walkers who alternate between random walks and correlated random walks. This multi-state movement simulation required specifying the probabilities of transition between the random walk state and correlated random walk state [18]. We used a transition matrix to define the probabilities of transition between states. The transition matrix is a square matrix where all values are probabilities, and the element at row  $i$  column  $j$  defines the probability of the individual changing from state  $i$  to state  $j$ . The transition matrix in this study is as same as the example in Quaglietta and Porto [21], which is

$$\begin{pmatrix} 0.995 & 0.005 \\ 0.01 & 0.99 \end{pmatrix}. \quad (4)$$

The step length, in meters, for both states are set to a random number from the uniform distribution  $U(0, 1)$ .

#### TDoA simulation

In the simulation, the coordinates of the receivers are:  $(-2000, 2000)$  m,  $(-2000, -2000)$  m,  $(2000, -2000)$  m, and  $(2000, 2000)$  m, an area of 16,000 hectares. To mimic real-world animal movements at different time scales the simulations recorded the positions of the track at different step intervals,  $N_s$ . We converted the recorded positions into ToA using equation 3. The TDoAs were obtained by the substitution of two ToAs using equation 2. The input TDoAs in our model are  $TDoA_{12}$ ,  $TDoA_{13}$ , and  $TDoA_{14}$ . For computational simplicity, we multiplied the derived times by the speed of light,  $c$  for all the input TDoAs to get distances. We added Gaussian noise  $N(0, \sigma)$ , based upon the static tag observation, into the simulated TDoAs to generate noised TDoAs, where  $\sigma$  is the standard deviation of the TDoA error distribution.

#### Training data simulation and preprocessing

For the training simulation dataset, the first step was to generate multiple animal movement tracks within a virtual paddock. Each track started from a random position within the paddock and picked a direction for the first step with equal probabilities. The subsequent steps were generated with the method discussed in section . As discussed in section , we assumed the device records the position of the animal in every  $N_s$  steps. The number of steps in each simulated track sequence is  $N_s \times 256$ . We then down-sample the track by extracting the first position in every  $N_s$  steps. The down-sampling mimics receivers only recording the animal's positions at set time intervals, the length of all tracks after down-sampling is 256. We saved these down-sampled tracks as the target outputs of the model. When the tracks were ready, we can use the method detailed in section to calculate the TDoA sequence of each track and then add random Gaussian noise.

The generated noised TDoAs (model input) and the corresponding ground truth positions (model output) pairs are split into a training dataset and a testing dataset in the fraction of 8 : 2. Meeting the requirements of the ResUnet-1d, the values of the model inputs and the model outputs should be in the range of  $[0, 1]$ , we rescale the input and output data in both training and testing datasets through min-max normalization. As all data used in this work is simulated, we don't need to tackle the missing data issue. However, the missing data issue is very common in practise, we will discuss this issue in section .

#### Evaluation

In this section, the predictive accuracy of the ResUnet-1d is evaluated by using simulated animal movement data described in the previous section. The localization results from our model are compared with the results of the analytical method implemented in Menzies et al [16].

#### Design of experiment

In this work, we implemented the ResUnet-1d model to reduce localization error, the main source of this error in real-world TDoA deployments is multipathing. We evaluated our model on the simulated Lévy-like tracks with different recording steps  $N_s$  and different TDoA error standard deviations  $\sigma$ . We aimed to investigate if ResUnet-1d could reduce the localization

**Table 1** Hyperparameters of the ResUnet-1d

Hyperparameters	Value
Depth	3
Number of filters	32
Batch size	256
Learning rate	$10^{-4}$
Optimizer	Adam
Loss Function	L1Loss

error significantly and how the ResUnet-1d model performance varied with step interval  $N_s$  and the original TDoA error's standard deviation  $\sigma$ .

As observed from the static tag in Townsville, the error distribution of TDoAs is Gaussian, and the standard deviation of the error in the range of  $\sim 100$  m. We use this value to simulate a city-like dataset. However, we suspect that the city environment has increased multipathing issues due to large metallic moving objects, such as vehicles, and other effects of the built environment. In more remote agricultural areas these aforementioned obstructions are greatly reduced, thus we can hypothesize that the standard deviation of the TDoA error in these locations is likely to be lower. To mimic this expectation, we choose to halve the standard deviation of the error to 50 m to simulate a remote area-like dataset. The choice of this value for the SD is corroborated by Menzies et al [16].

In our simulation, the step size was selected from a uniform distribution  $U(0, 1)$  m to mimic continuous movement. As discussed in Section , we down sampled the track positions by recording only one position in every  $N_s$  positions, where  $N_s \in [10, 20, 40, 60, 80, 100, 200, 300, 500]$  which is equivalent to a time interval ranging from 0.6 – 30 minutes when considering the average speed of a cow [19] [24], but is likely to be much higher for grazing individuals.

The optimized hyperparameters of ResUnet-1d are summarized in Table . For all models, we used the Adam [12] optimizer, with a learning rate of  $10^{-4}$ . We chose the L1Loss loss function to obtain the best training performance for ResUnet-1d. Our model was built and trained using the MXNet deep learning library [4], under the GLUON API. Each of the models was trained on 8,000 simulated tracks with a batch size of 256 on a single NVIDIA Tesla P100 GPU using the CSIRO's HPC facilities. We used 2,000 simulated tracks to test the performance of our model.

### Performance of ResUnet

Fig. 4 and Fig. 5 illustrate simulated tracks, the orange lines are the ground truth movement track generated from the animal movement simulations and the faded blue points are the measured positions calculated with noised TDoAs. The green lines represent the recovered tracks by the ResUnet-1d model, they reproduce the

shape of the ground truth tracks and recovered most of their features.

To evaluate the performance of the model quantitatively, we measure the root mean square (RMSE) error of noised tracks and recovered tracks to the ground truth tracks in the following way

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \quad (5)$$

where the  $(x, y)$  represents the position of noised tracks or recovered tracks, and  $(\hat{x}, \hat{y})$  represents the position of ground truth tracks. The probability density function (PDF) of the errors is illustrated in Fig. 6 and Fig. 7. In both figures, the distribution of the error of the recovered tracks is significantly narrowed and of a lower value compared with the distribution of the original error of the noised tracks. Therefore, we demonstrate that our ResUnet-1d model can effectively reduce localization error introduced by random Gaussian errors in TDoA position estimates.

Fig. 8 compares the RMSE of noised and recovered tracks from two simulations with different TDoA errors.

The RMSE of the two recovered tracks are significantly lower than the RMSE of the two noised tracks. When the original localization error was higher, orange lines, the efficacy of the recovered tracks dropped slightly. The proposed model reduced the localization error from  $\sim 60$  and  $\sim 150$  m to  $\sim 20$  and  $\sim 25$  m on average. The downsampling interval was not so important, when increased from 20 to 100 the RMSE of the recovered tracks improved becoming asymptotic thereafter.

### Discussion

Conventional analytic methods for calculating locations based upon TDoAs work well in the absence of multipathing; in real-world settings multipathing degrades the performance of these systems. Previous machine learning methods have shown a reduction in localization noise but have only considered a single static position.

The method proposed in this work implements a CNN layer to consider the connection between neighboring positions for animal movement tracks. The combination of the CNN layer and U-Net like encoder/decoder architecture enhances the ability of the model to reduce noise in the data.

By testing this algorithm on a correlated random walk simulation we have made the assumption of a homogeneous paddock environment using SimRiv. It is recognised that SimRiv does not take into account

**Figure 4** Illustrations of simulated ground truth (orange) and recovered (green) tracks. The blue points are the positions from noised TDoA. The black vertical and horizontal lines indicate 100 m distance. The standard deviation of the raw TDoA noise is  $\sigma = 50$  m. Inset black numbers indicate step down-sampling increment.

**Figure 5** Illustrations of simulated ground truth (orange) and recovered (green) tracks. The blue points are the positions from noised TDoA. The black vertical and horizontal lines indicate 100 m distance. The standard deviation of the raw TDoA noise is  $\sigma = 100$  m. Inset black numbers indicate step down-sampling increment.

the impacts of a heterogeneous landscape, where the willingness of an animal to cross a specific environment is not taken into account. We anticipate a heterogeneous environment, such as creek lines, would elicit less random animal movements resulting in a higher correlation between neighboring positions. This kind of correlation should translate to increased performance when compared to animal tracks simulated under the assumption of a homogeneous environment.

The use of this method requires a sequence of the TDoAs without missing data. Missing data is likely to be common in TDoA localization networks. In real deployments, the signal transmission can not only be reflected by objects but also blocked by them. Moreover, network communication outages along with transmitter maintenance or failure will also lead to data gaps.

It would be possible to implement a simple interpolation preprocessing step to the proposed model to recover missing data. Zhang et al [28] proposed a new sequence-to-sequence imputation model for recovering missing data in wireless sensor networks. This method could be implemented in the data preprocessing pipeline to address missing data.

We intend to take this simulated model and apply it to a working cattle station to see if the demonstrated gains hold true for a real-world scenario.

## Conclusion

In this paper, we developed and investigated a 1D CNN-based U-Net like encoder/decoder model for denoising TDoA position estimates for animal tracking using simulated animal movement data. We have demonstrated that our model can successfully recover simulated animal movement tracks from noised TDoA sequences, with a reduction of the mean error of between 10 % ( $\bar{X}$  error 100 m) and 20 % ( $\bar{X}$  error 50 m). As the results for the tracks with different step intervals don't show a clear trend, it is possible the algorithm can be implemented for animal tracks with a lower-frequency transmission, at down sampling intervals greater than 500 m. These findings need to be validated in a working cattle station in conjunction with an assessment of missing data preprocessing.

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Availability of data and material

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

### Competing interests

The authors declare that they have no competing interests.

### Funding

Not applicable

### Author's contributions

L. Wang and F. Diakogiannis conceived of the presented idea. L. Wang, F. Diakogiannis, S. Mills, and N. Bajema planned the experiments. L. Wang carried out data simulation, deep learning model developing, training and data interpretation. F. Diakogiannis supported the development of deep learning model architecture. L. Wang took the lead in writing the manuscript in consultation with F. Diakogiannis, S. Mills and N. Bajema. All authors discussed the results and contributed to the final manuscript.

### Acknowledgements

The authors acknowledge the support of the Scientific Computing team of CSIRO. We would also like to acknowledge Gordon Foyster and Richard Keaney for their useful discussions and comments on this manuscript.

### Author details

<sup>1</sup>Agriculture & Food, CSIRO, 4814 Townsville, Australia. <sup>2</sup>Agriculture & Food, CSIRO, 4067 Brisbane, Australia. <sup>3</sup>ICRAR, The University of Western Australia, 6009 Perth, Australia. <sup>4</sup>Data61, CSIRO, 6014 Perth, Australia. <sup>5</sup>eResearch Centre, James Cook University, 4814 Townsville, Australia.

### References

- Alonso-González I, Sánchez-Rodríguez D, Ley-Bosch C, Quintana-Suárez MA (2018) Discrete indoor three-dimensional localization system based on neural networks using visible light communication. *Sensors* 18(4):1040.
- Bengio Y, Delalleau O (2009) Justifying and generalizing contrastive divergence. *Neural computation* 21(6):1601–1621.
- Chan YT, Ho K (2005) Joint time-scale and tdoa estimation: analysis and fast approximation. *IEEE Transactions on Signal processing* 53(8):2625–2634.
- Chen T, Li M, Li Y, Lin M, Wang N, Wang M, Xiao T, Xu B, Zhang C, Zhang Z (2015) Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. *arXiv preprint arXiv:151201274* URL <https://arxiv.org/abs/1512.01274v1>
- Clark DD (1989) Overview of the argos system. In: *Proceedings OCEANS*, vol 3, pp 934–939.
- Craighead Jr FC, Craighead JJ, Cote CE, Buechner HK (1972) Satellite and ground radiotracking of elk. *NASA, Washington Animal Orientation and Navigation* URL <https://ntrs.nasa.gov/citations/19720017418>
- de Sousa MN, Thomä Electronic RS (2019) Enhanced localization systems with multipath fingerprints and machine learning. In: *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, pp 1–6.

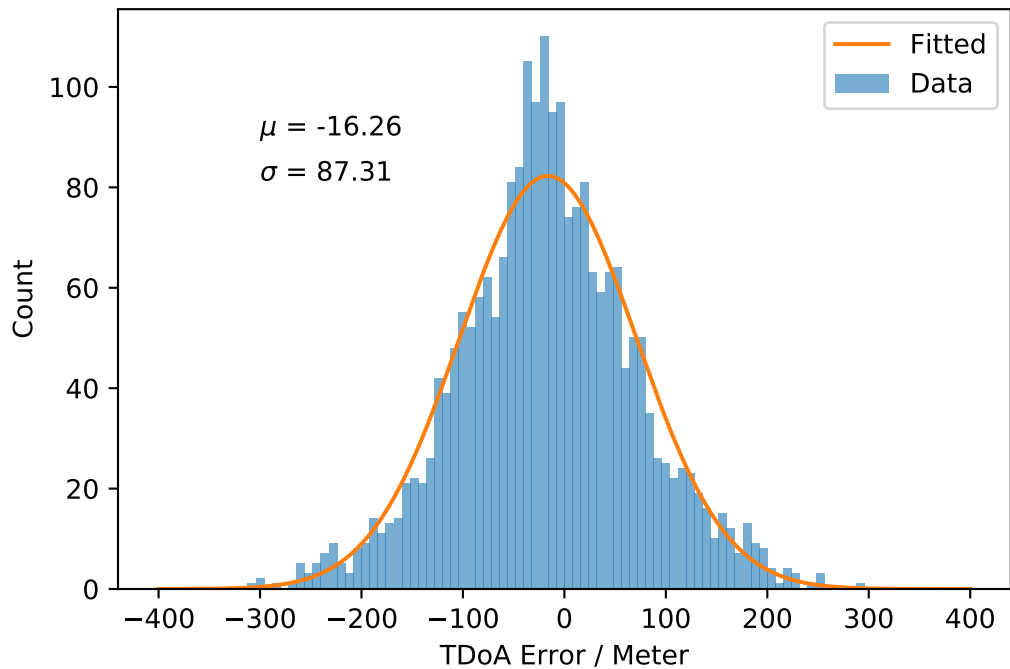


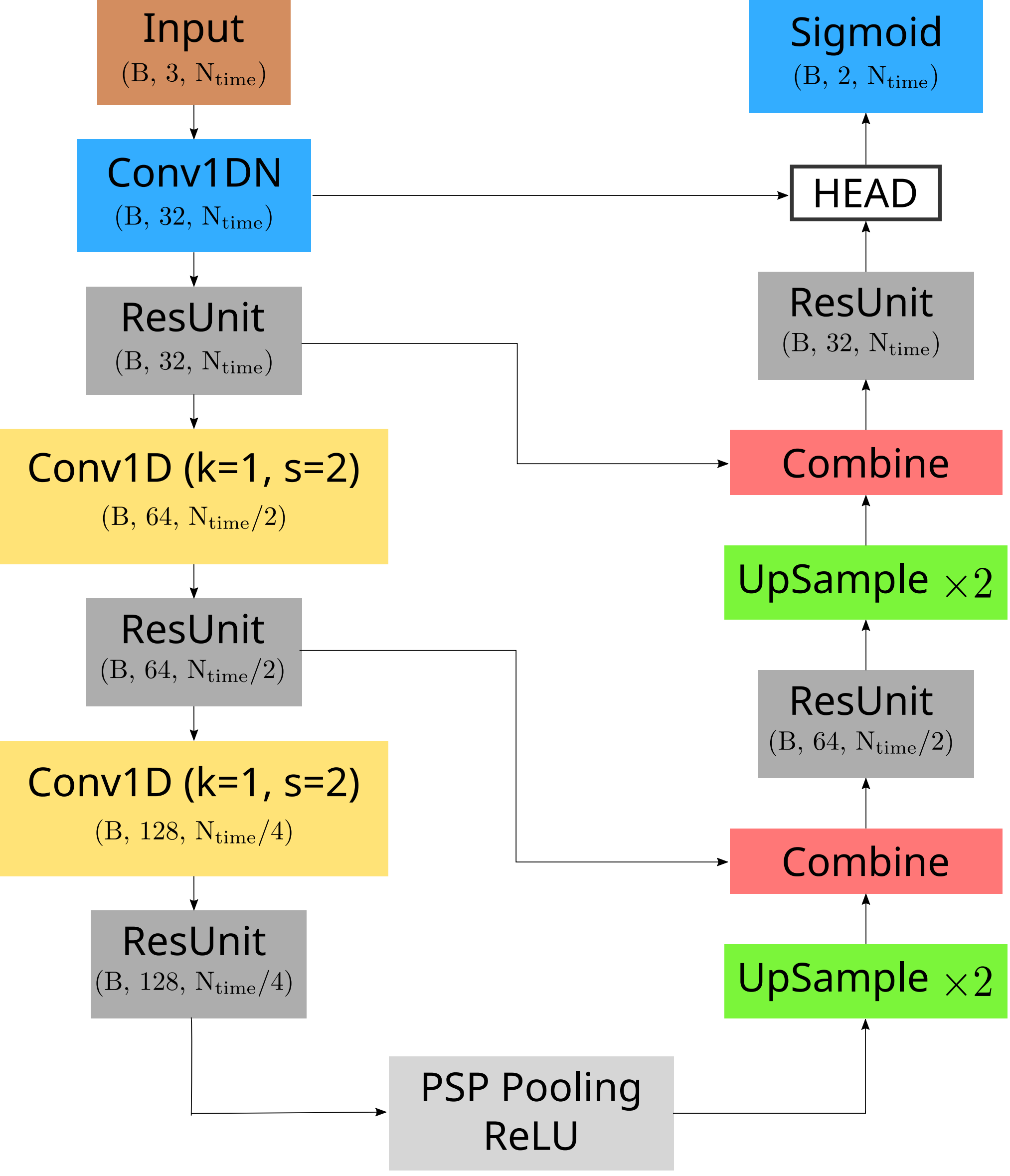
**Figure 6** localization error distribution for tracks calculated with original TDoAs with 50 m error standard deviation and recovered tracks.

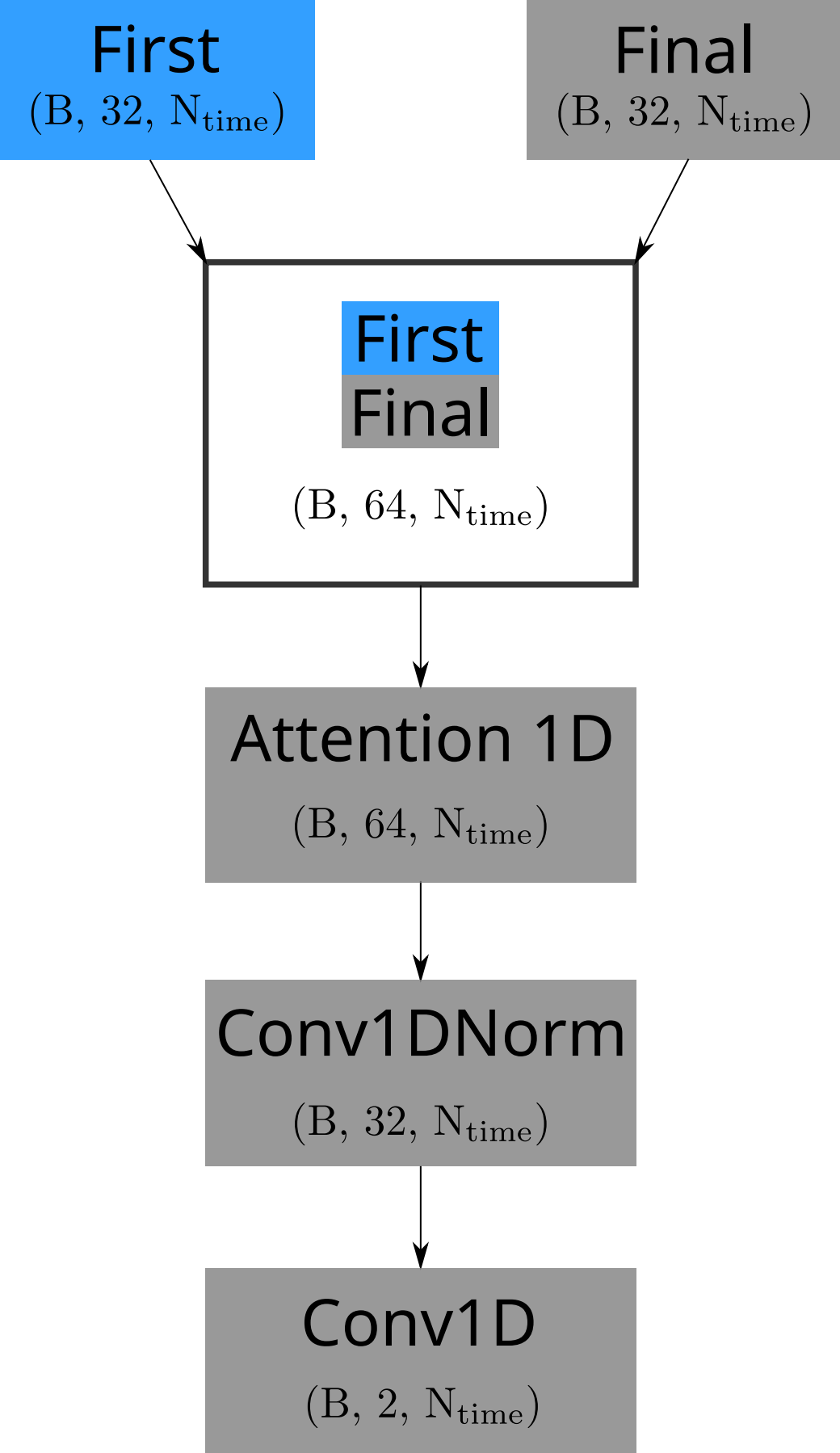
**Figure 7** localization error distribution for tracks calculated with original TDoAs with 100 m error standard deviation and recovered tracks.

**Figure 8** Comparison of localization errors for tracks calculated with original TDoAs and recovered tracks.

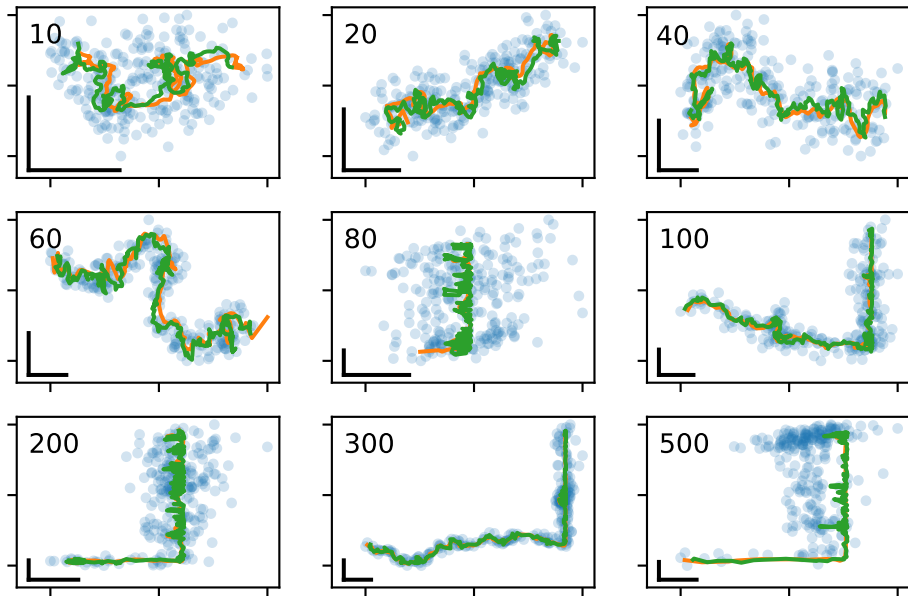
8. Diakogiannis F, Waldner F, Caccetta P, Wu C (2020) Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* 162:94–114,
9. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778,
10. Heinrich MP, Stille M, Buzug TM (2018) Residual u-net convolutional neural network architecture for low-dose ct denoising. *Current Directions in Biomedical Engineering* 4(1):297–300,
11. Jouventin P, Weimerskirch H (1990) Satellite tracking of wandering albatrosses. *Nature* 343(6260):746–748,
12. Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* URL <https://arxiv.org/abs/1412.6980v9>
13. Komatsu R, Gonsalves T (2019) Effectiveness of u-net in denoising rgb images. *Comput Sci Inf Techn* pp 1–10,
14. Liu D, Wen B, Liu X, Wang Z, Huang TS (2017) When image denoising meets high-level vision tasks: A deep learning approach. *arXiv preprint arXiv:1706.04284*
15. Marshall WH, Gullion GW, Schwab RG (1962) Early summer activities of porcupines as determined by radio-positioning techniques. *The Journal of Wildlife Management* 26(1):75–79,
16. Menzies D, Patison KP, Fox DR, Swain DL (2016) A scoping study to assess the precision of an automated radiolocation animal tracking system. *Computers and Electronics in Agriculture* 124:175–183,
17. Mo J, Deng Z, Jia B, Bian X (2017) A pseudorange measurement scheme based on snapshot for base station positioning receivers. *Sensors* 17(12):2783,
18. Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85(9):2436–2445,
19. O'Driscoll K, Schutz MM, Lossie A, Eicher S (2009) The effect of floor surface on dairy cow immune function and locomotion score. *Journal of Dairy Science* 92(9):4249–4261,
20. Priede I (1984) A basking shark (*Cetorhinus maximus*) tracked by satellite together with simultaneous remote sensing. *Fisheries Research - FISH RES* 2:201–216,
21. Quaglietta L, Porto M (2019) Simriv: an r package for mechanistic simulation of individual, spatially-explicit multistate movements in rivers, heterogeneous and homogeneous spaces incorporating landscape bias. *Movement Ecology* 7,
22. Reynolds AM (2014) Mussels realize weierstrassian lévy walks as composite correlated random walks. *Scientific reports* 4:4409,
23. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp 234–241,
24. Swain D, Wark T, Bishop-Hurley G (2008) Using high fix rate gps data to determine the relationships between fix rate, prediction errors and patch selection. *Ecological Modelling* 212(3):273 – 279,
25. Turchin P (1998) Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates,
26. Vincent P, Larochelle H, Bengio Y, Manzagol PA (2008) Extracting and composing robust features with denoising autoencoders. In: *Proceedings of the 25th international conference on Machine learning*, pp 1096–1103,
27. Wu L, Li F, Chen S (2013) A improved wireless location algorithm in nlos environment. *Information Technology Journal* 12(24):8563,
28. Zhang YF, Thorburn PJ, Xiang W, Fitch P (2019) Ssim—a deep learning approach for recovering missing time series sensor data. *IEEE Internet of Things Journal* 6(4):6618–6628,
29. Zhang Z, Jiang F, Li B, Zhang B (2018) A novel time difference of arrival localization algorithm using a neural network ensemble model. *International Journal of distributed sensor networks* 14(11):1550147718815,798,







Y (m)

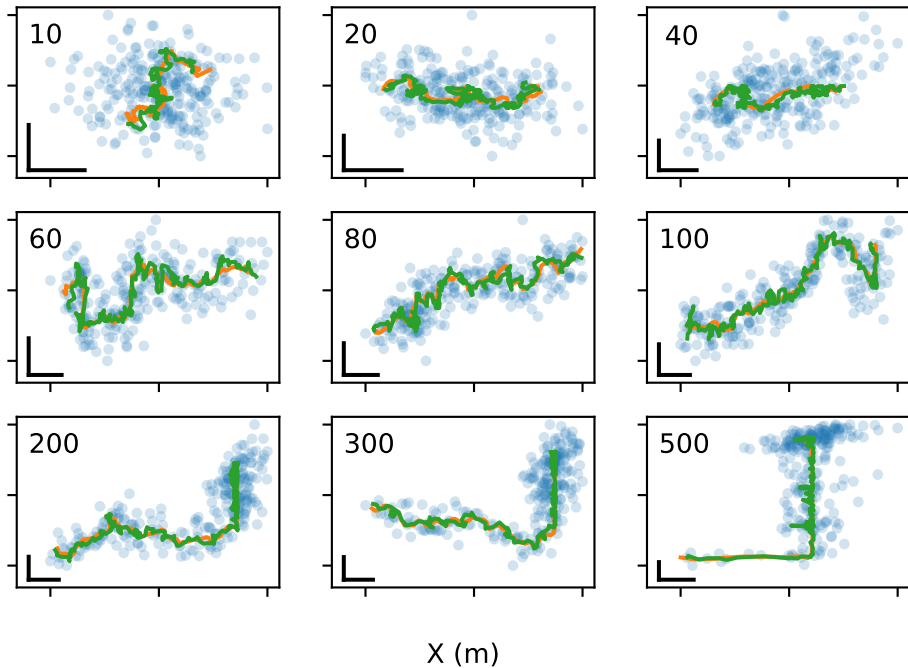


X (m)

• Uncorrected    — Ground Truth    — Corrected

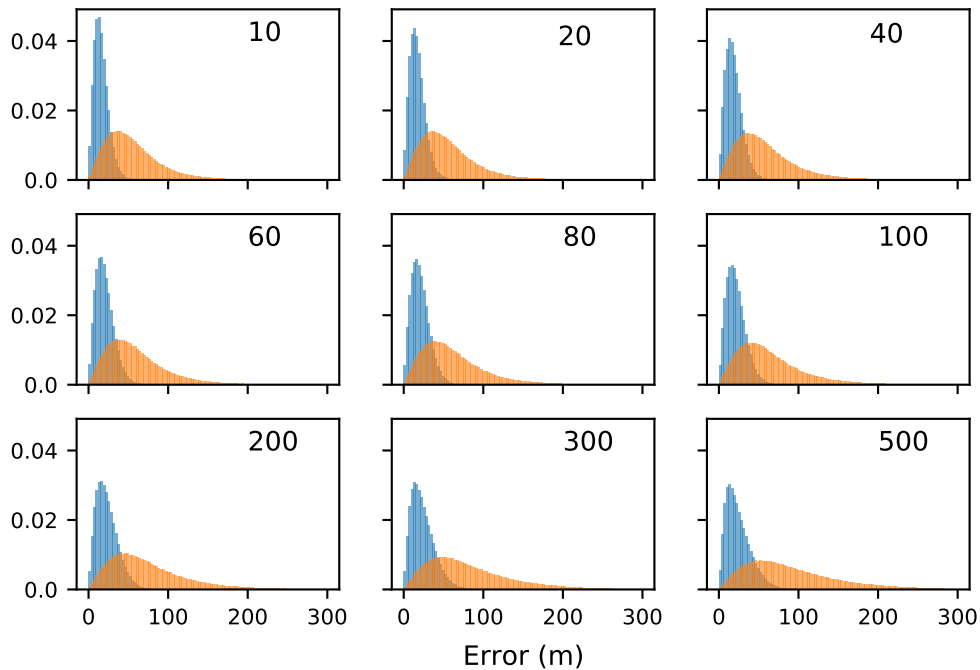


Y (m)



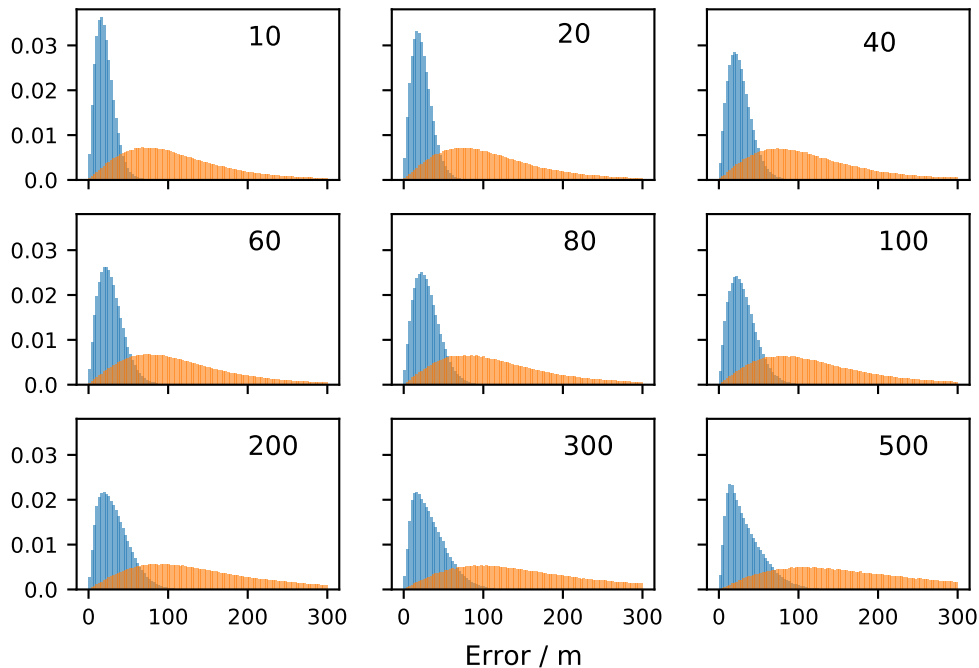
• Uncorrected      — Ground Truth      — Corrected

Probability Density Function

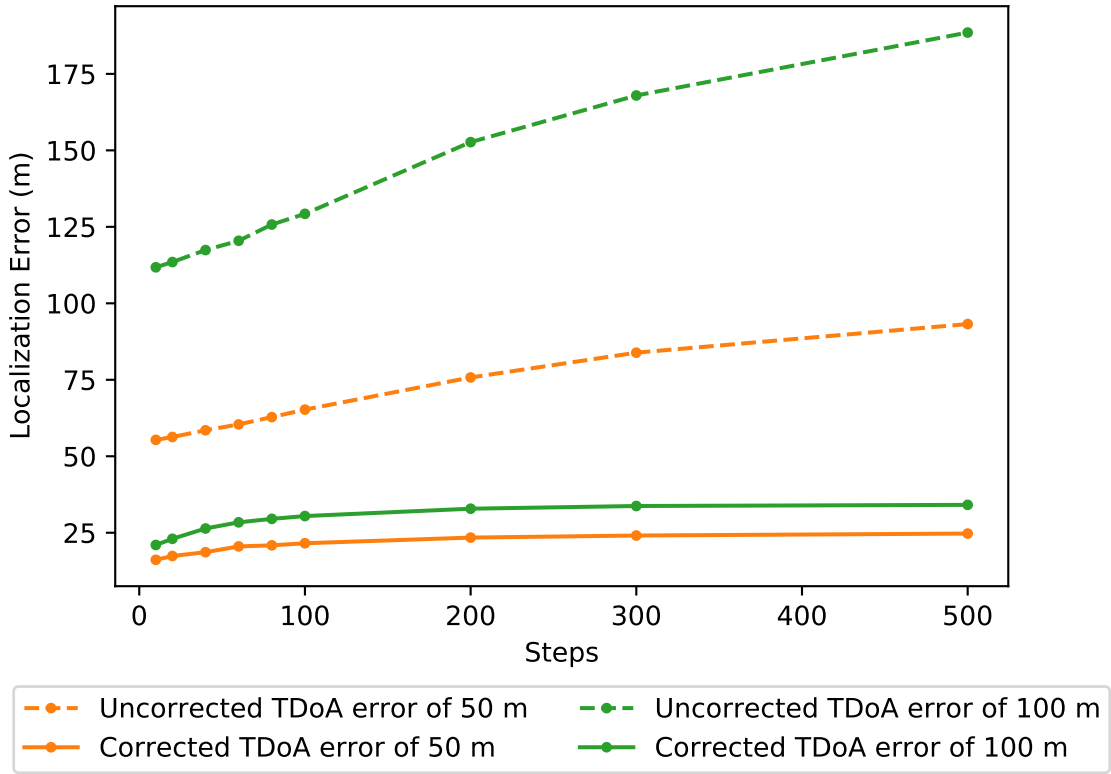


Corrected      Uncorrected

Probability Density Function



Corrected      Uncorrected



## Figures

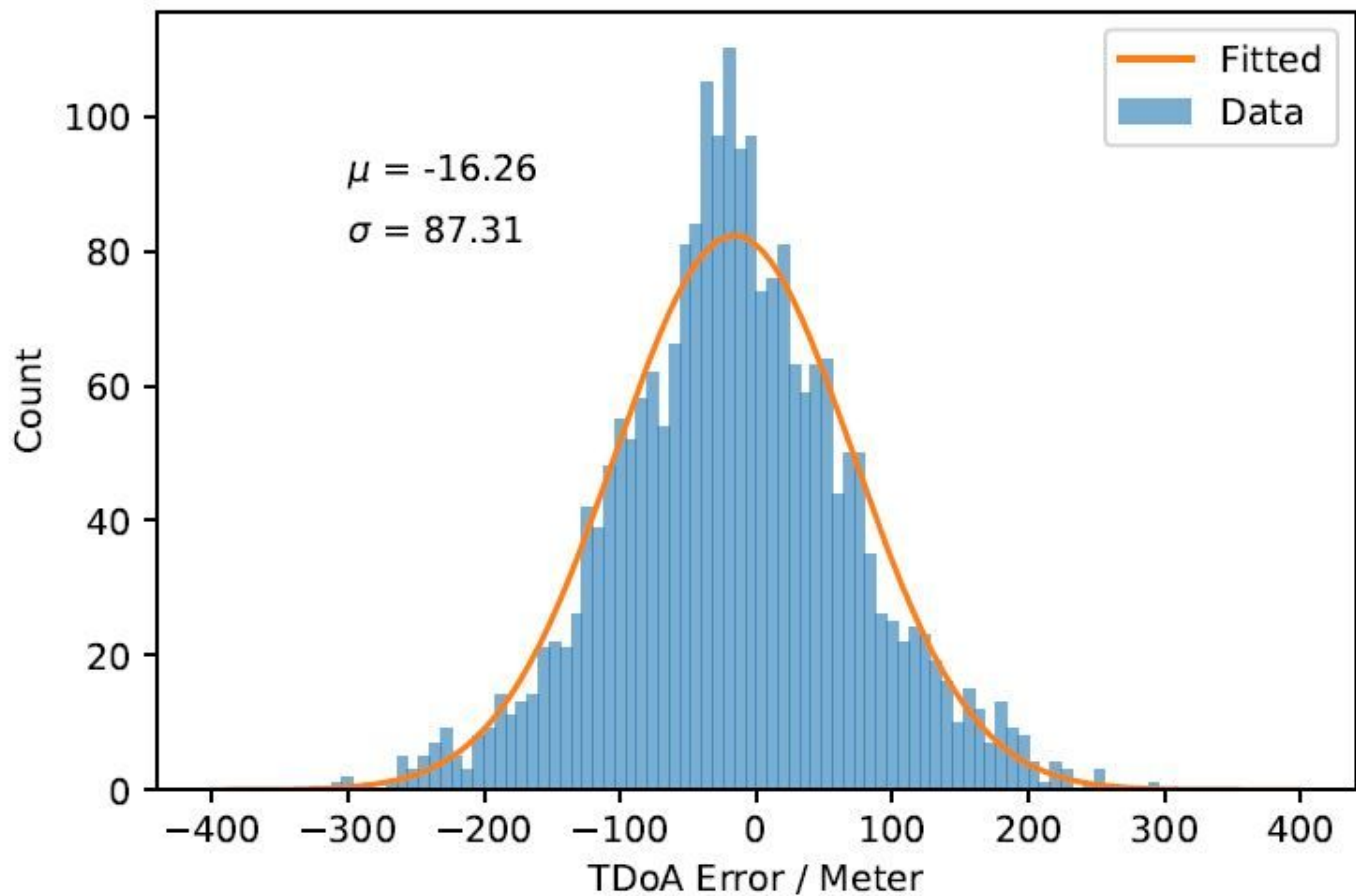
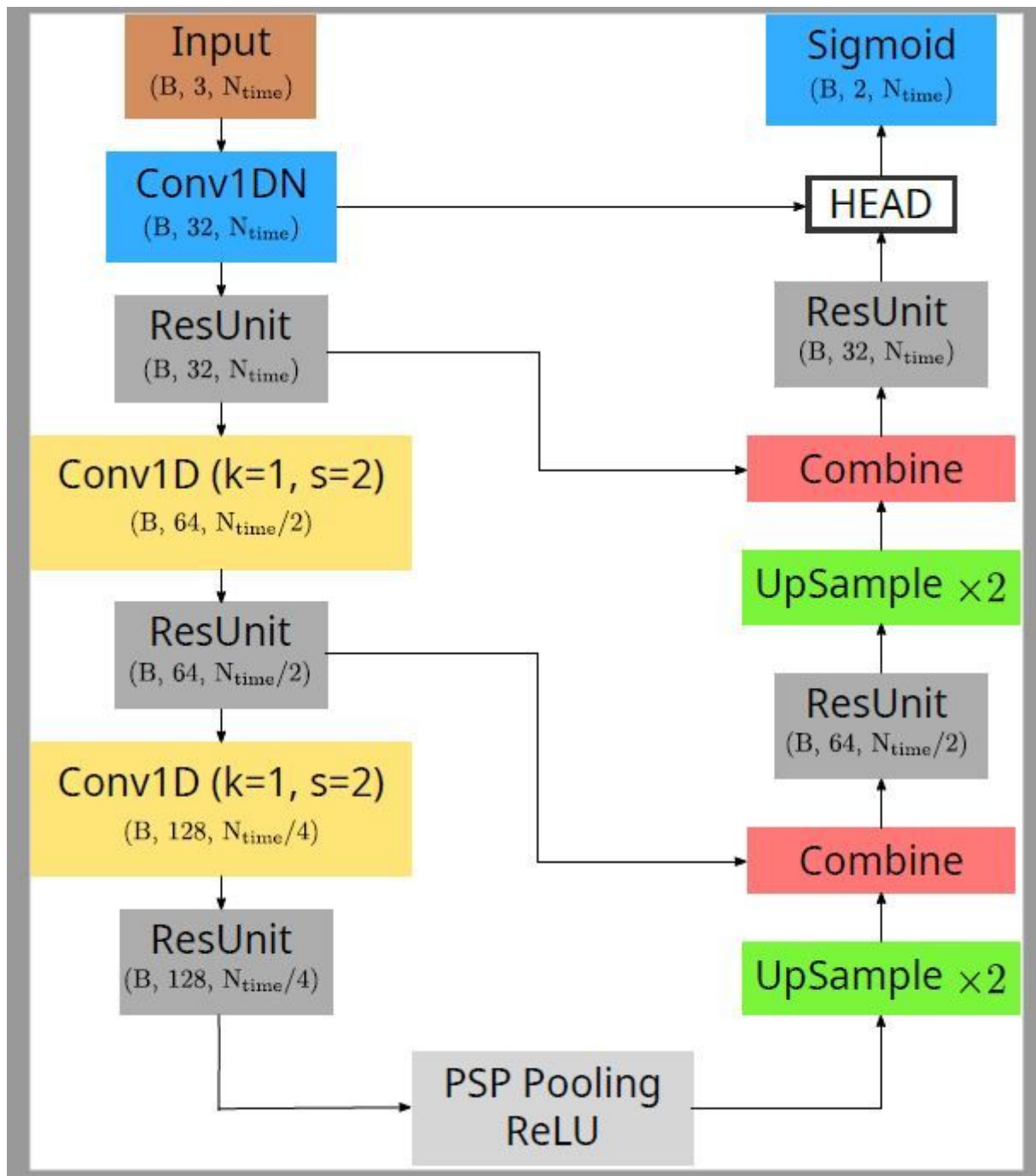


Figure 1

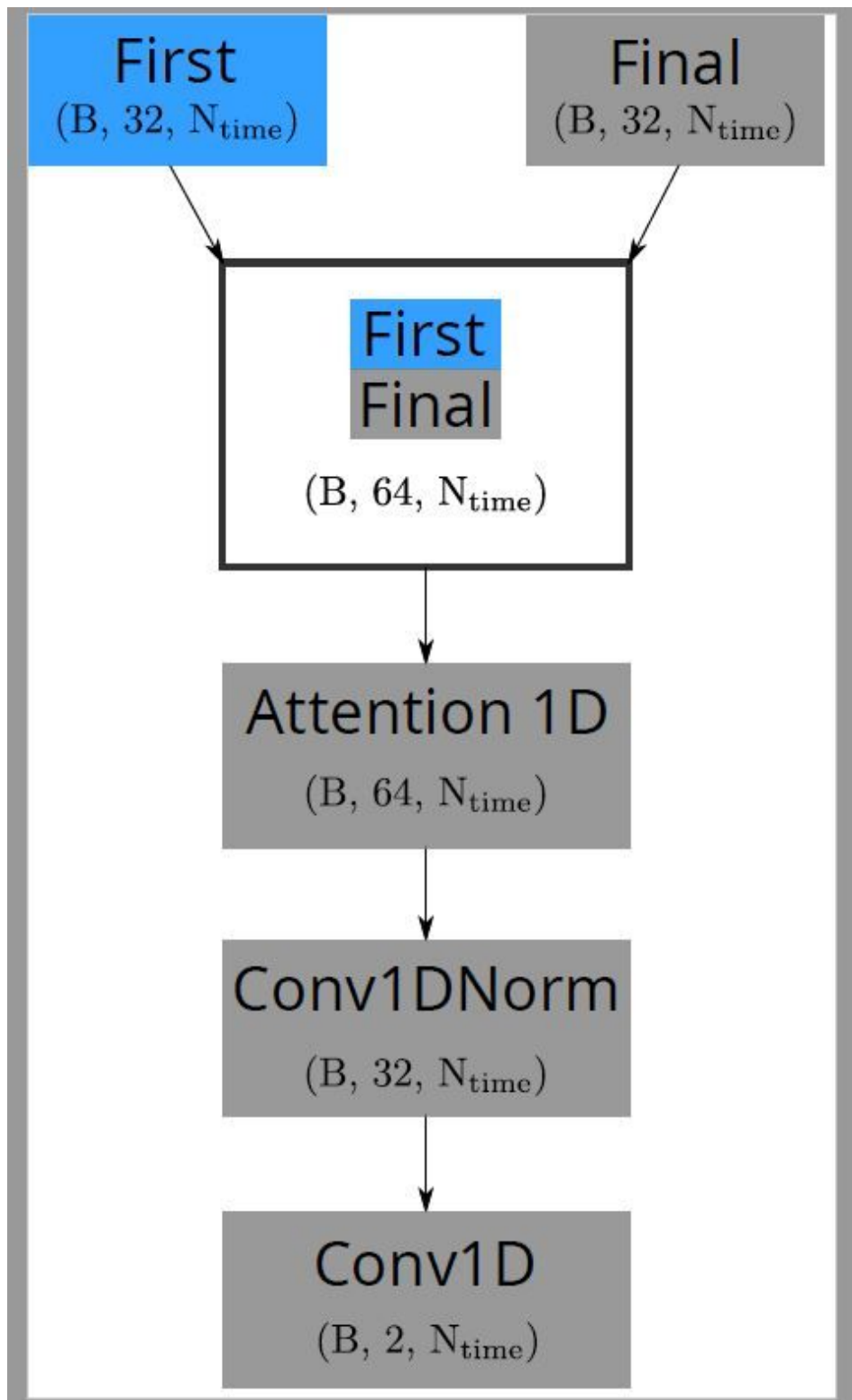
Histogram (bin 8 m) of sampled error distribution of TDoAs measured on a static reference tag (tag ID: 130114) deployed in Warina Park, Townsville, Australia. A total of 2215 transmissions were used to estimate the error distribution using two Taggle receivers (tower ID: taggle-067 and taggle-058). The fitted distribution demonstrates the Gaussian nature of the error with an  $\mu \approx 0$  and  $\sigma \approx 100$ .





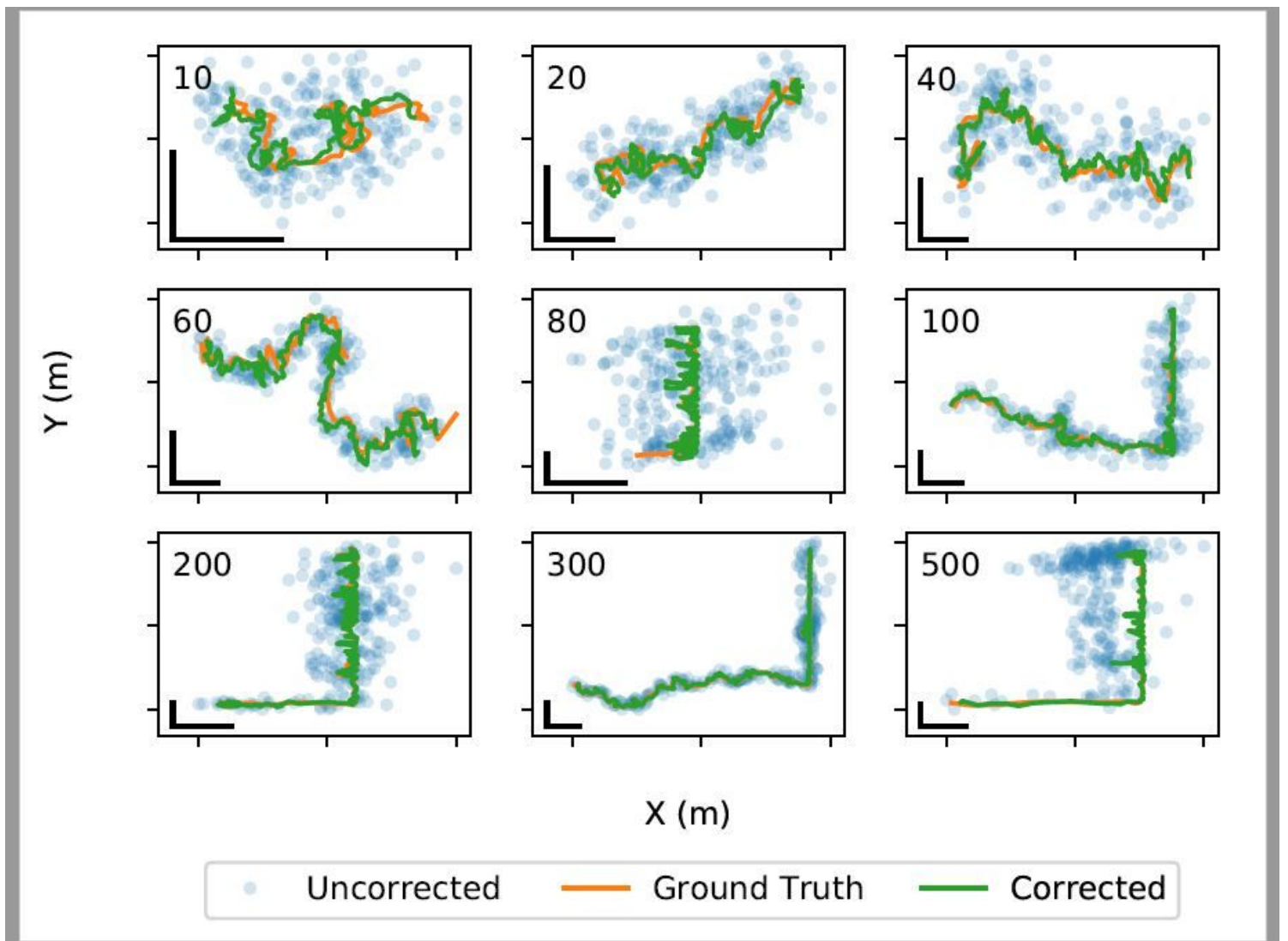
**Figure 2**

ResUnet-1d architecture. The left (downward) branch is the encoder. The right (upward) branch is the decoder. Conv1D is the standard 1D CNN layer, and Conv1DN is the standard 1D CNN layer with batch normalization. The B in data size represents the batch size, and the Ntime represents the sequence length of the input data.



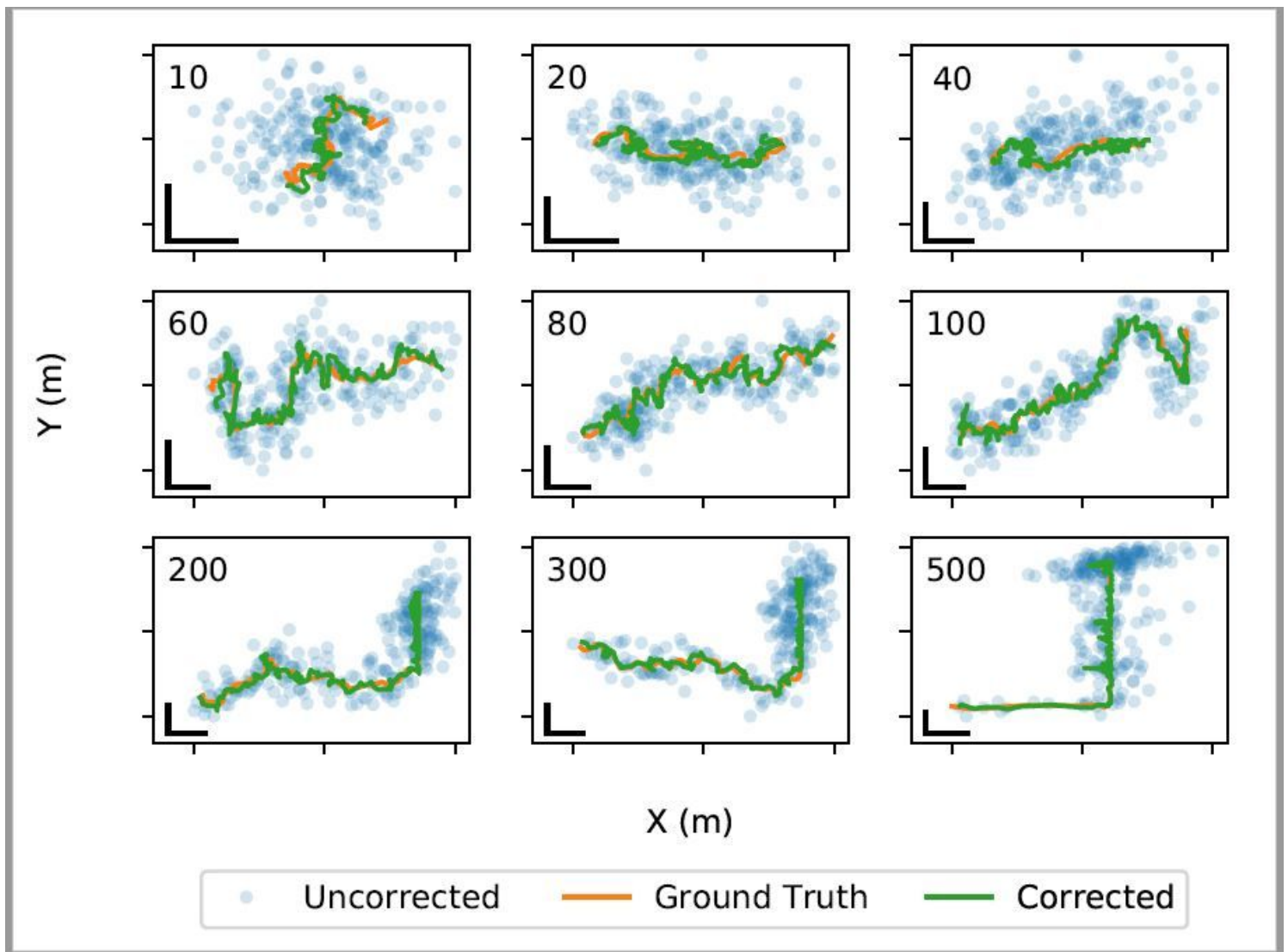
**Figure 3**

The architecture of the HEAD block. The two inputs of this block are the outputs of the last Conv1DN in the encoder branch named First and the output of the final ResUnit in the decoder branch named Final. The B and Ntime have same meaning as in Fig. 2.



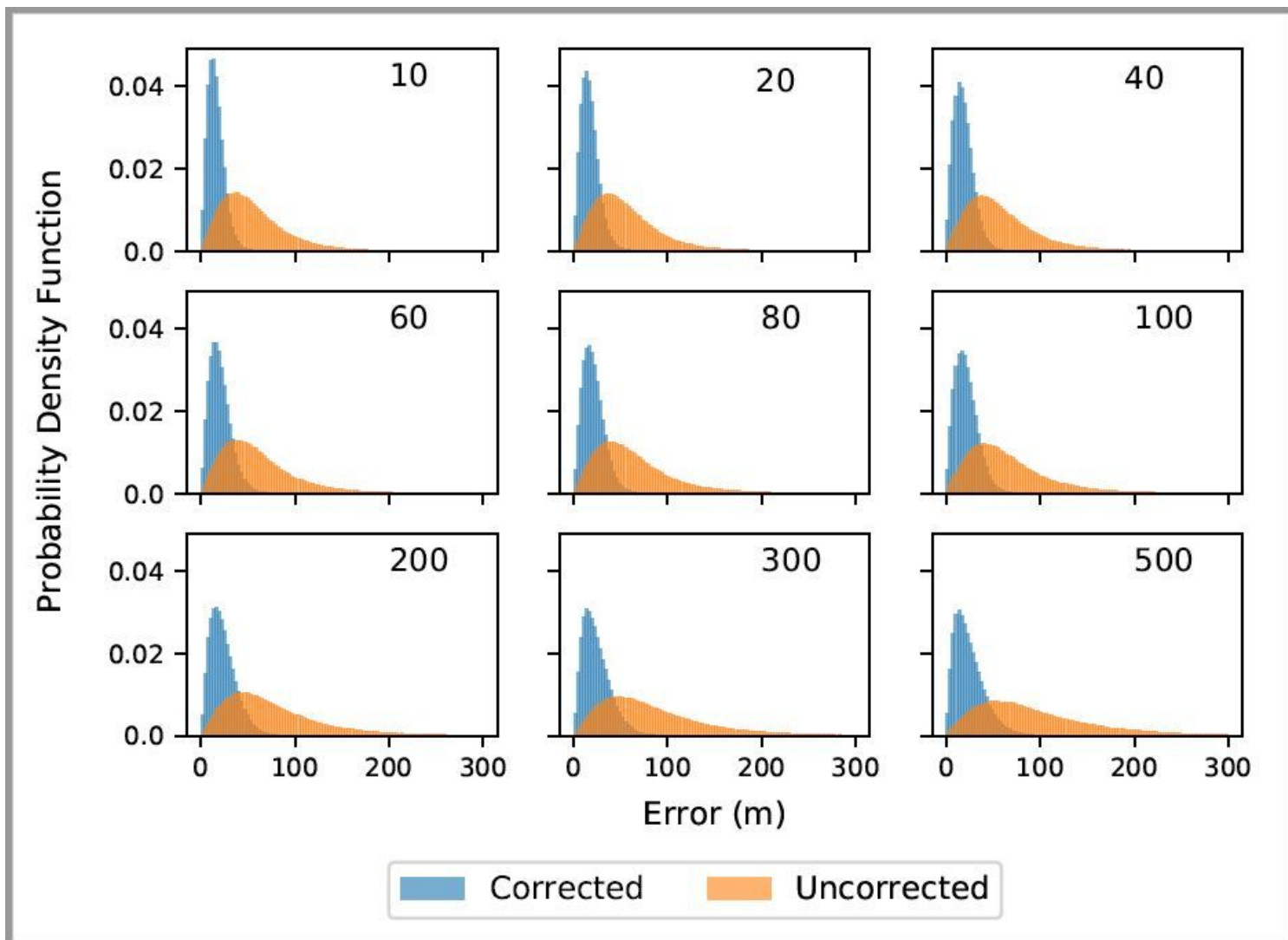
**Figure 4**

Illustrations of simulated ground truth (orange) and recovered (green) tracks. The blue points are the positions from noise TDoA. The black vertical and horizontal lines indicate 100 m distance. The standard deviation of the raw TDoA noise is  $\sigma = 50$  m. Inset black numbers indicate step down-sampling increment.



**Figure 5**

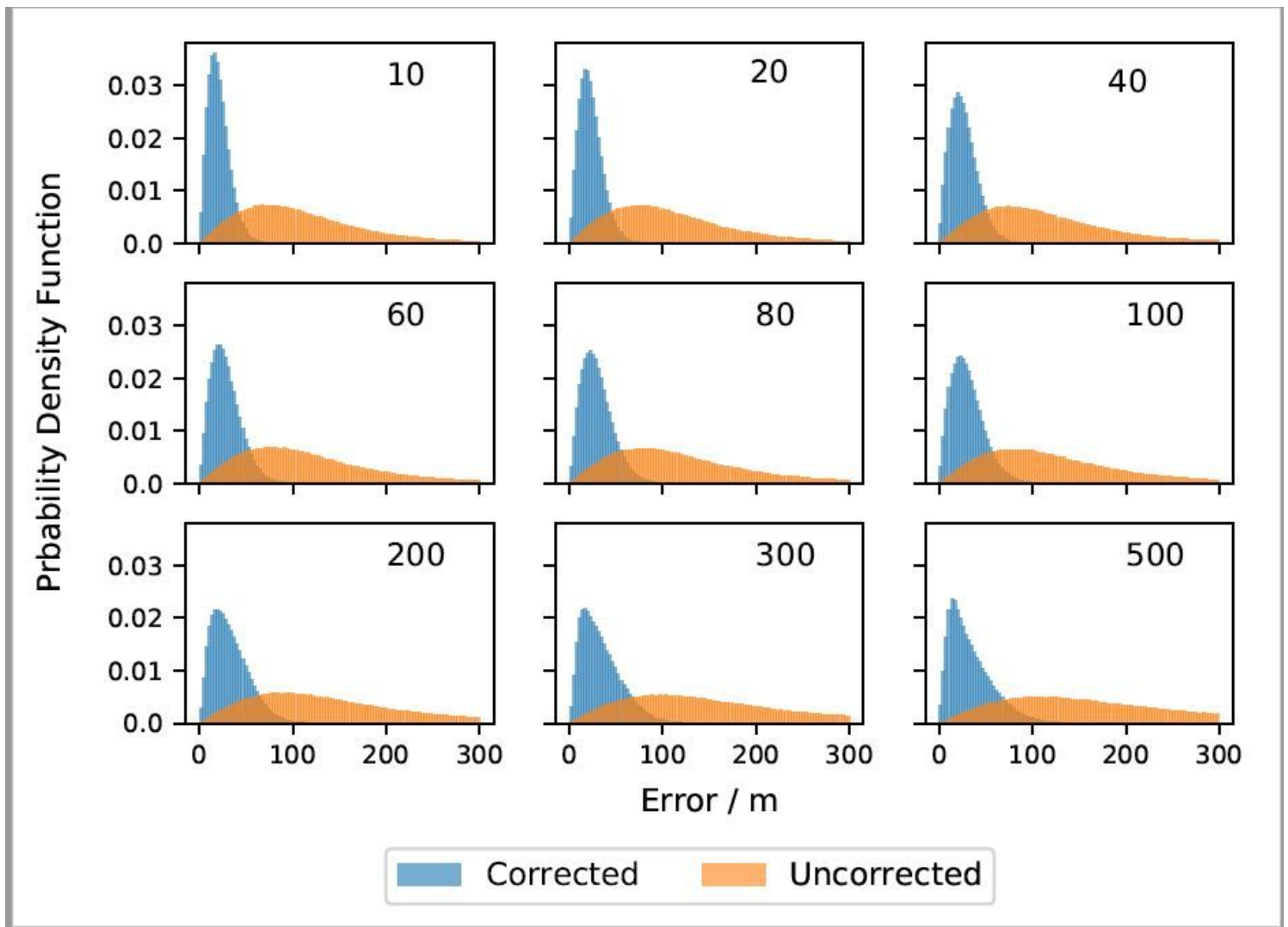
Illustrations of simulated ground truth (orange) and recovered (green) tracks. The blue points are the positions from noise TDoA. The black vertical and horizontal lines indicate 100 m distance. The standard deviation of the raw TDoA noise is  $\sigma = 100$  m. Inset black numbers indicate step down-sampling increment.



**Figure 6**

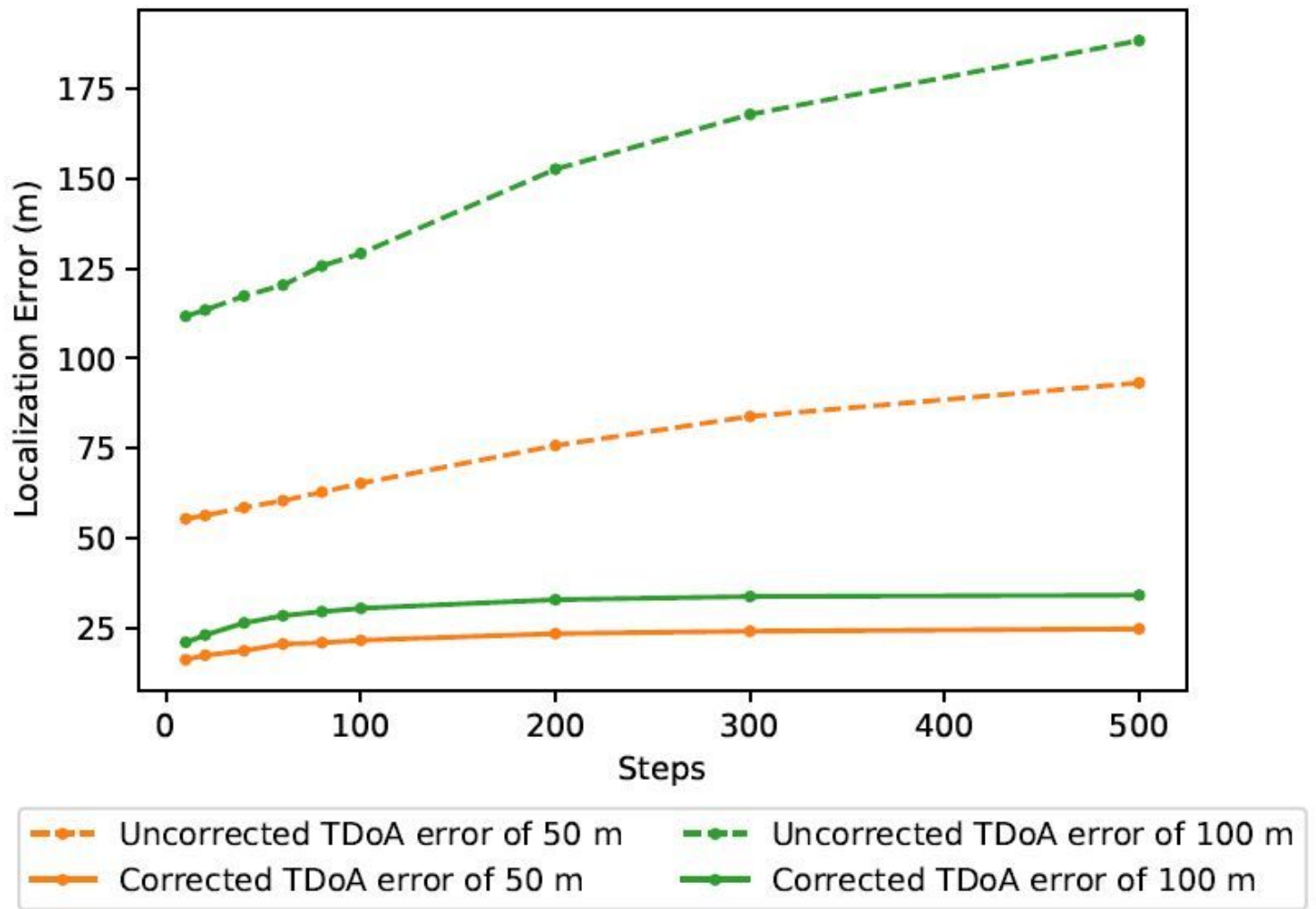
localization error distribution for tracks calculated with original TDoAs with 50 m error standard deviation and recovered tracks.





**Figure 7**

localization error distribution for tracks calculated with original TDoAs with 100 m error standard deviation and recovered tracks.



**Figure 8**

Comparison of localization errors for tracks calculated with original TDoAs and recovered tracks.