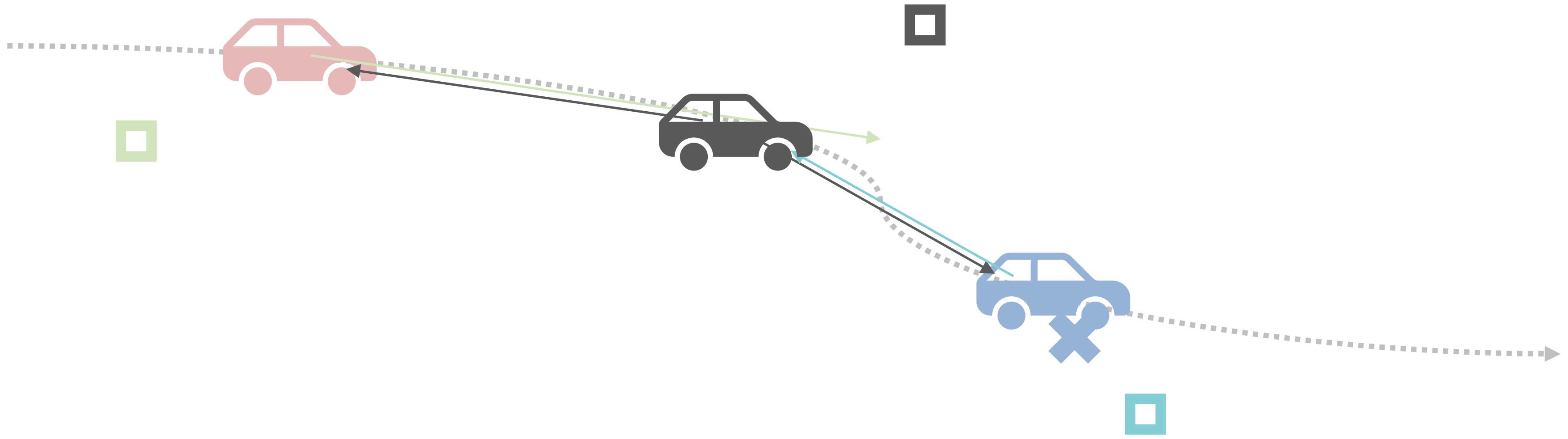
A Data-Fusion Method using Bayesian Approach to Enhance Raw Data Accuracy of Position and Distance Measurements for Connected Vehicles

May 21, 2021

Hyeonsup Lim, Oak Ridge National Laboratory
Bumjoon Bae, The Korea Transport Institute
Lee D. Han, University of Tennessee
Shih-Miao Chin, Oak Ridge National Laboratory
Ho-Ling Hwang, Oak Ridge National Laboratory

Multi-source Data in Connected Environment

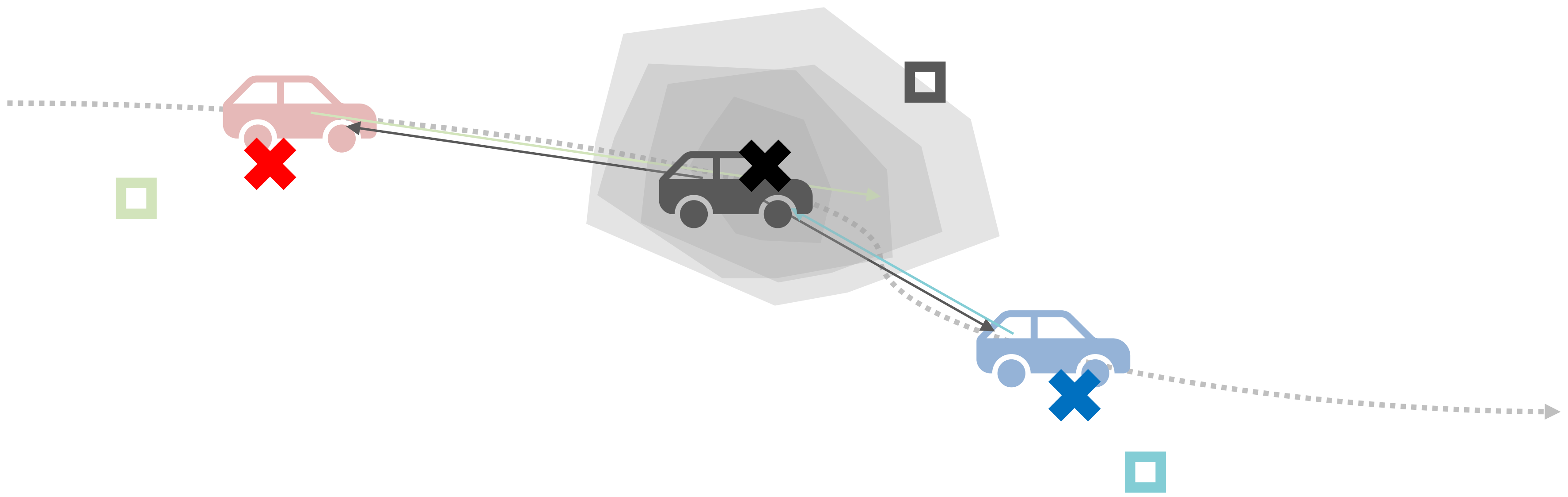
- Each vehicle has three sensor measurements:
  - position of the designated vehicle,
  - distance to the preceding vehicle, and
  - distance to the following vehicle.



Oak Ridge National Laboratory

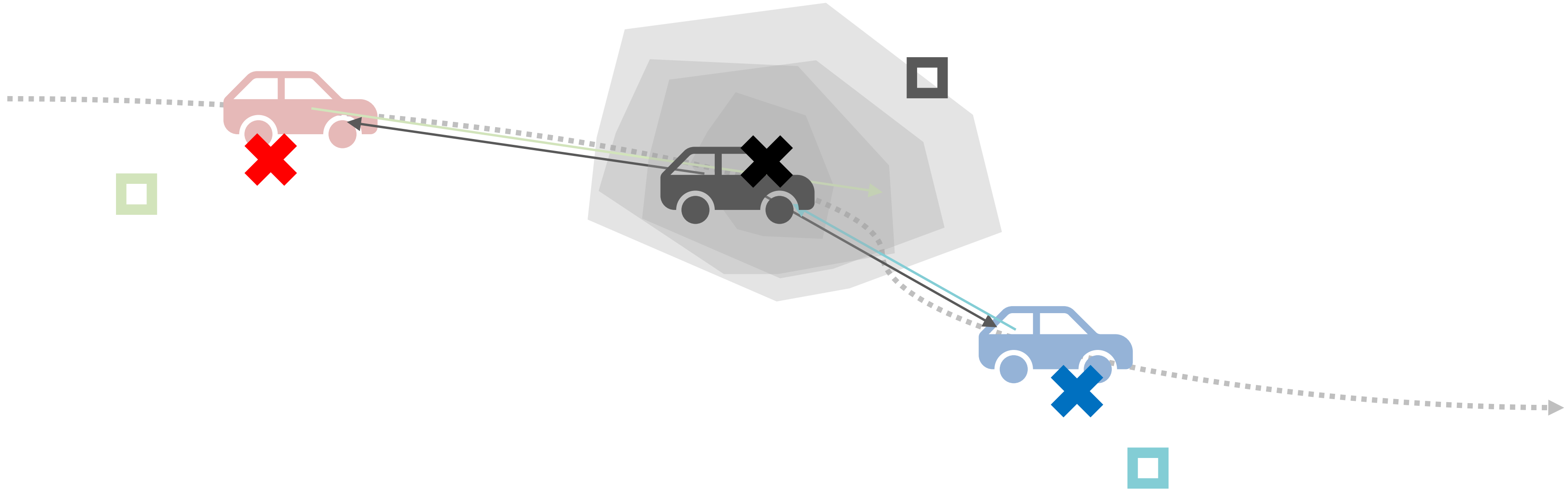
Problem to Solve

- How can we improve the accuracy of the position and distance measurements by aggregating the multi-source data?

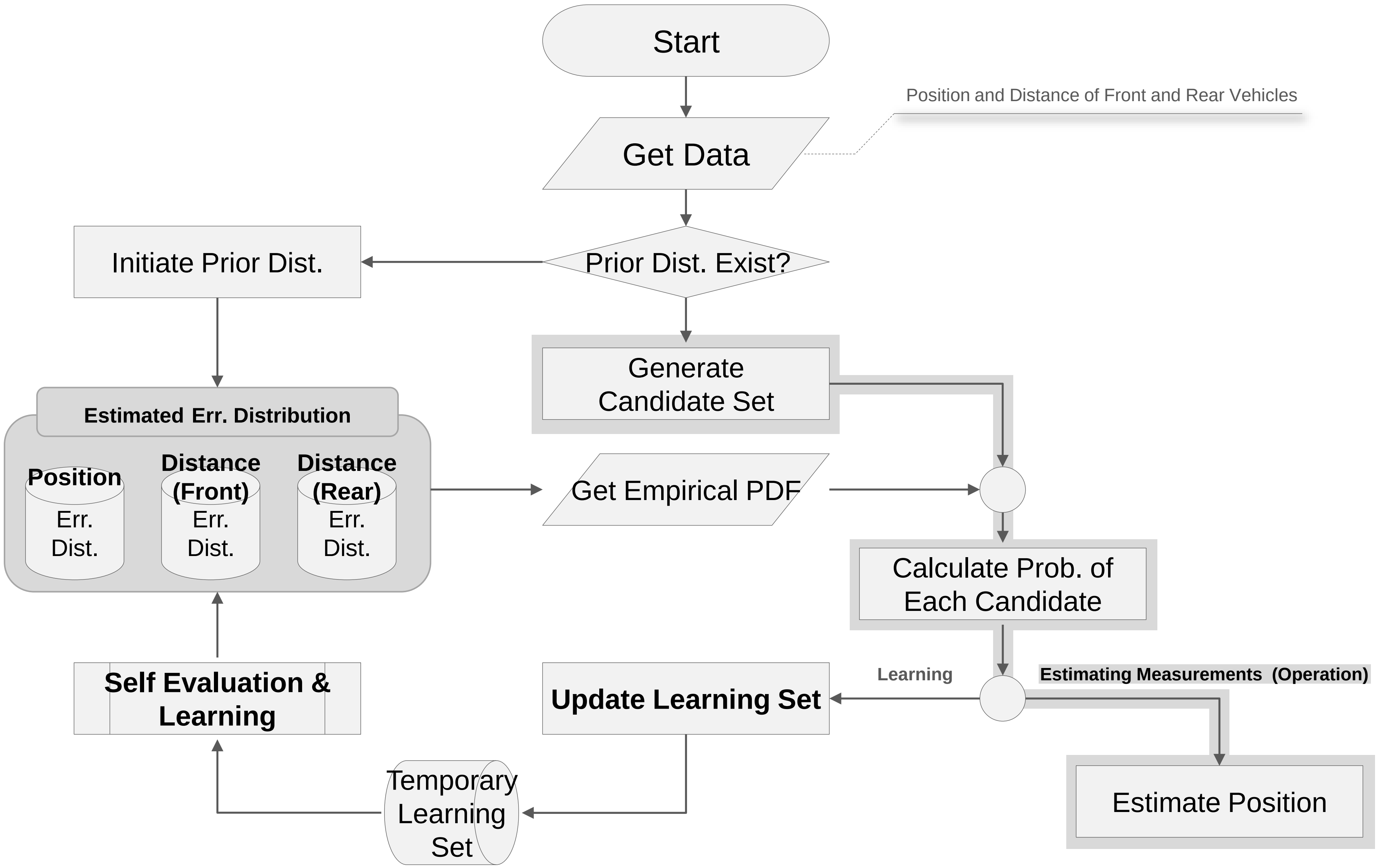


Key Ideas of Proposed Method

- Utilize the given information as much as possible.
- Estimate error distribution of each sensor measurement.
- Apply self-learning scheme that continuously update the estimated error distribution.



Flowchart of Proposed Method



OAK RIDGE National Laboratory

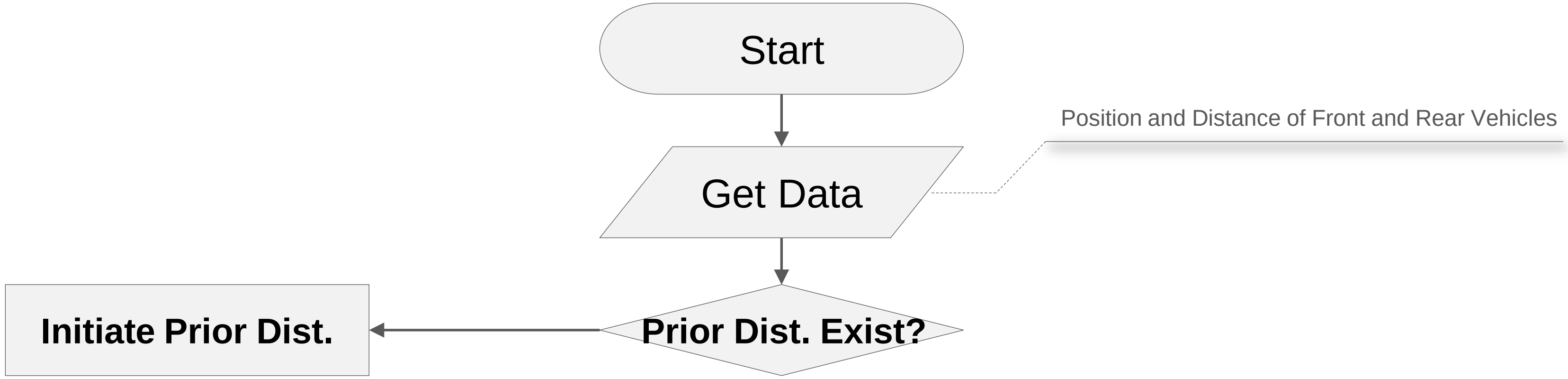
Flowchart of Proposed Method

(1) position of the designated vehicle;
(2) distance measured from the designated vehicle to the preceding vehicle;
(3) distance measured from the designated vehicle to the following vehicle;

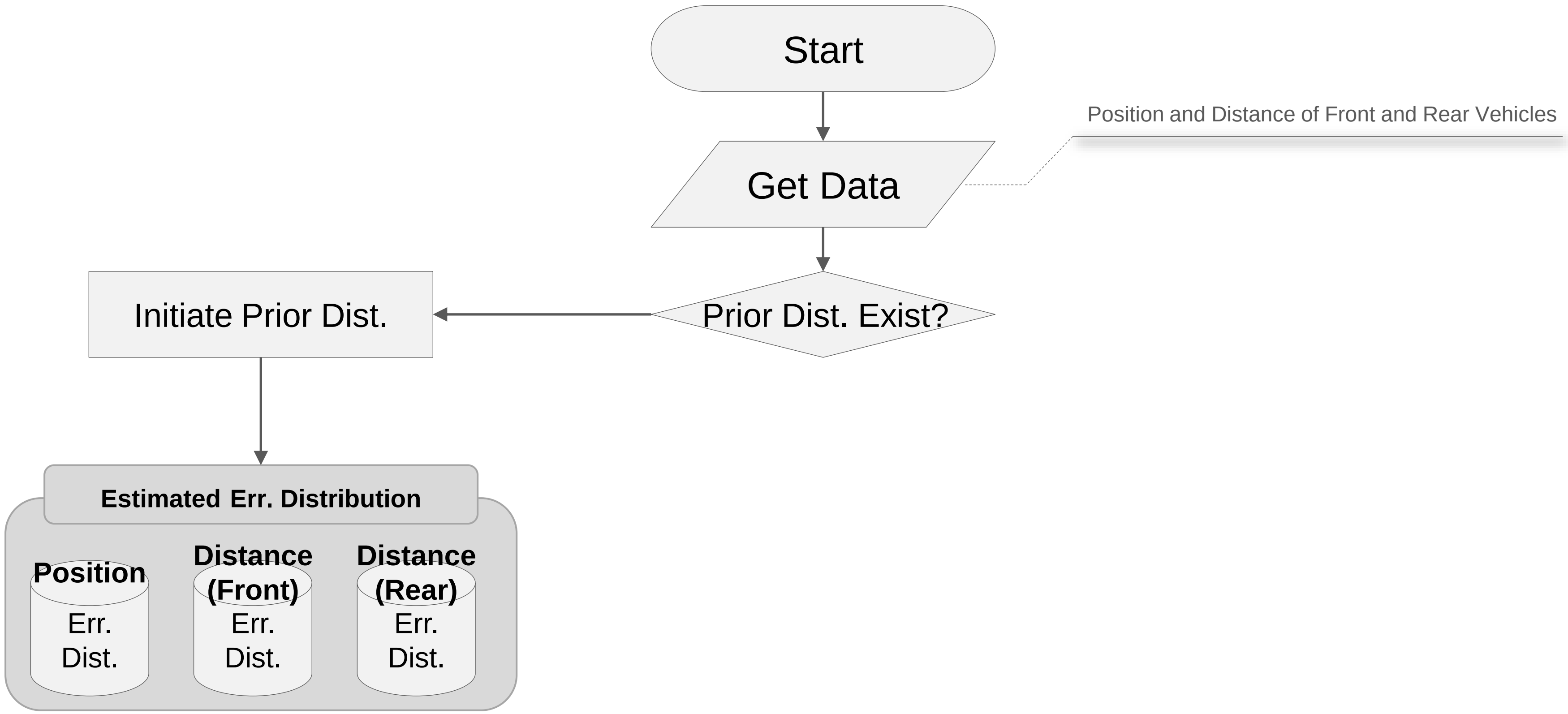
(4) position of the preceding vehicle;
(5) distance measured from the preceding vehicle to the designated vehicle;

(6) position of the following vehicle; and
(7) distance measured from the following vehicle to the designated vehicle.

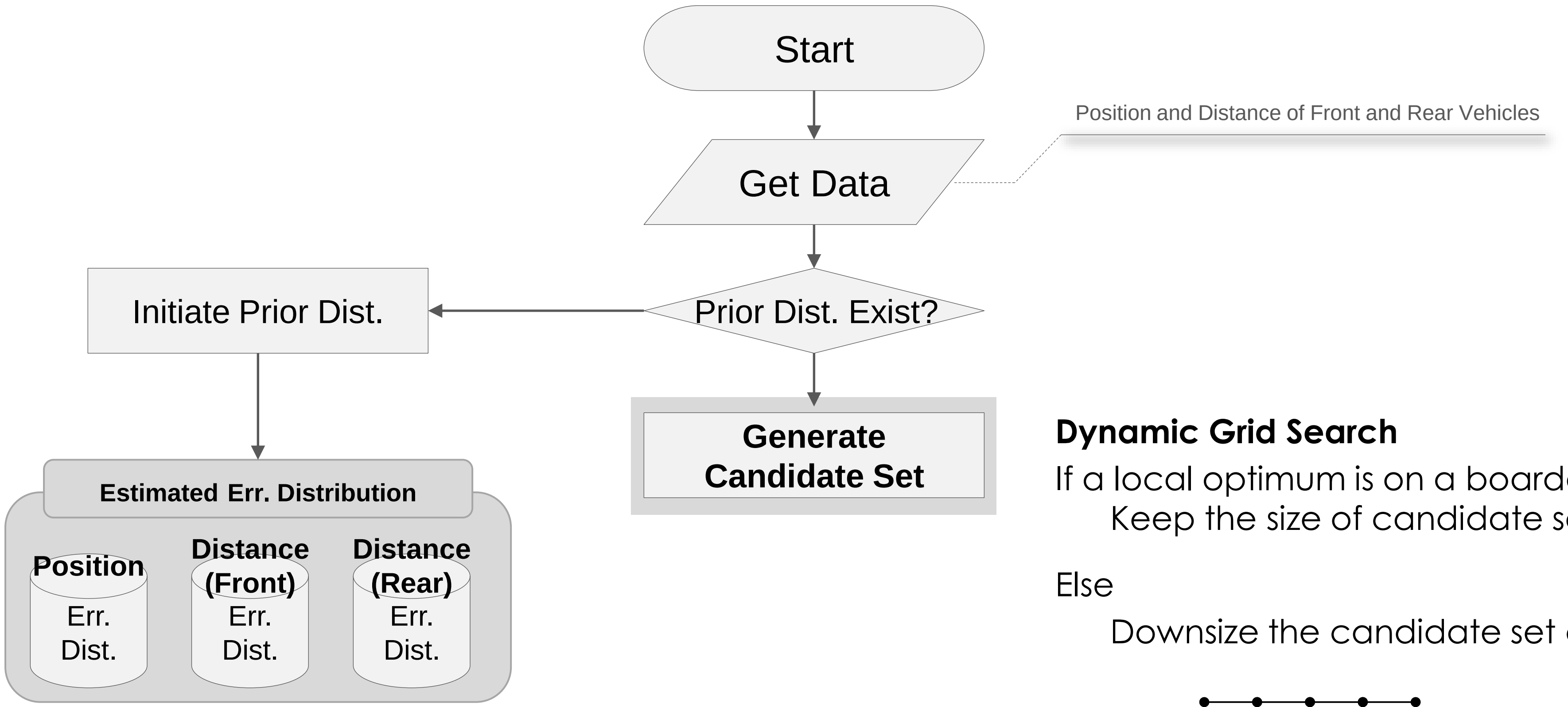
Flowchart of Proposed Method



Flowchart of Proposed Method



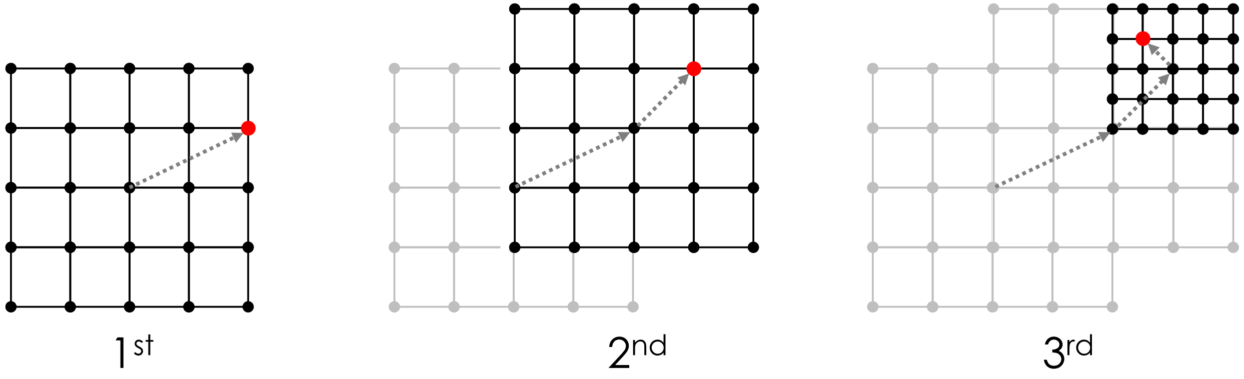
Flowchart of Proposed Method



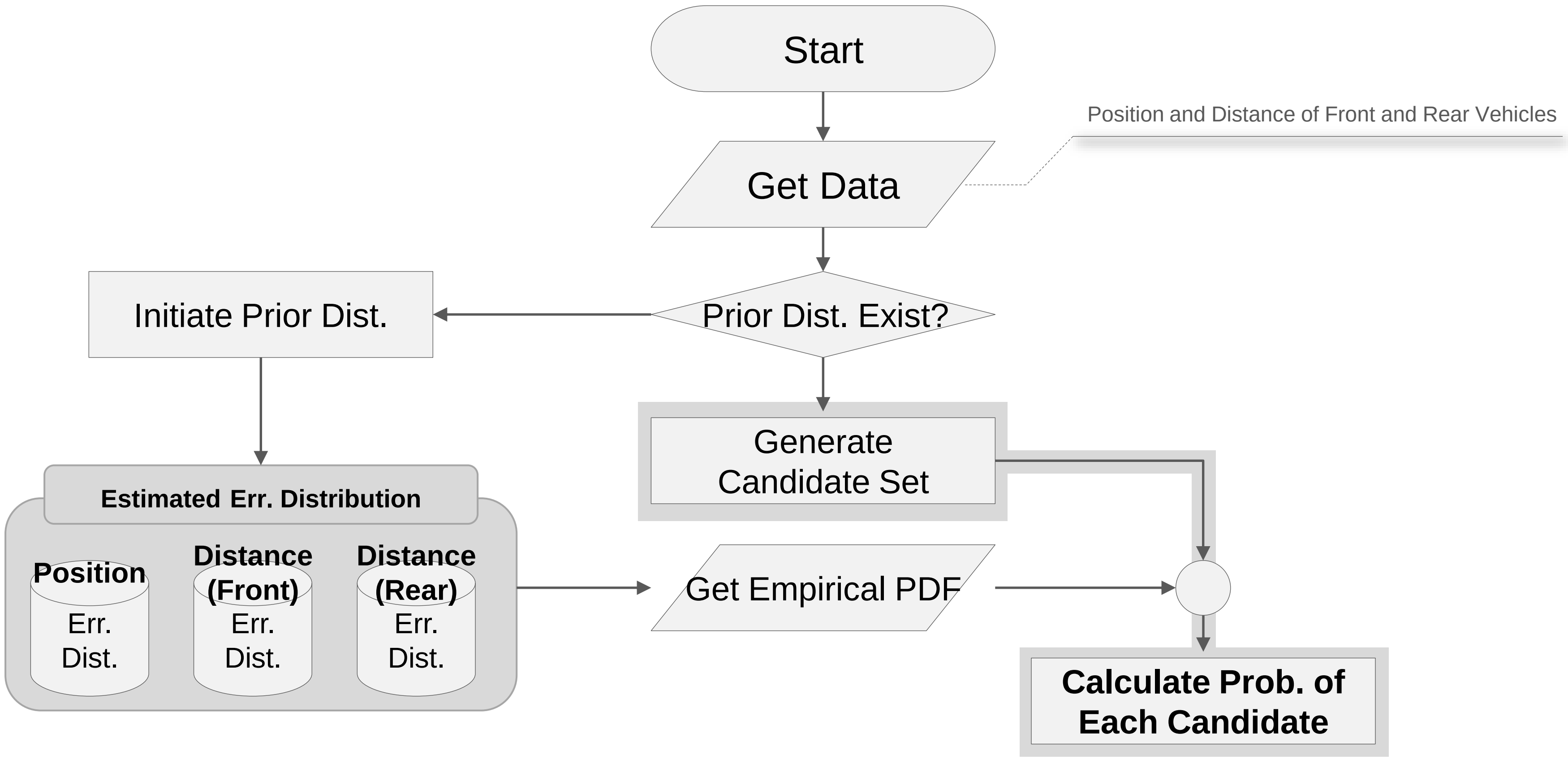
Dynamic Grid Search

If a local optimum is on a boarder,
Keep the size of candidate set area

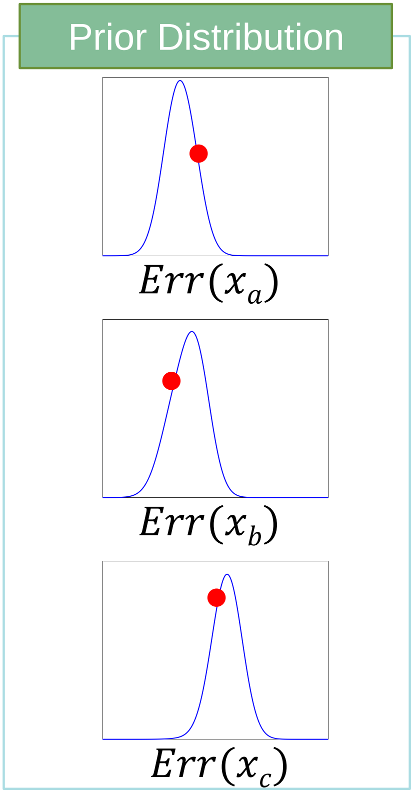
Else
Downsize the candidate set area



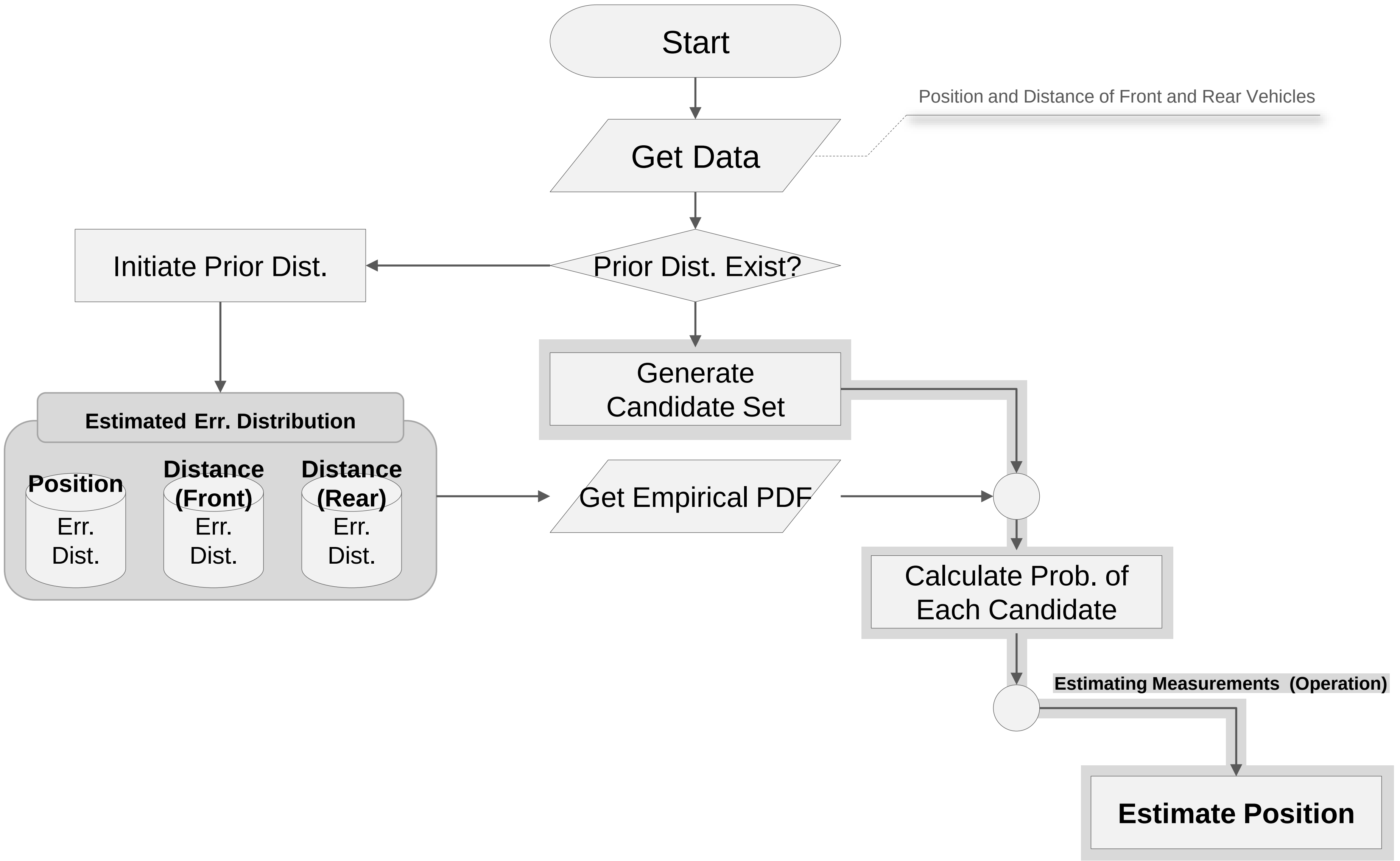
Flowchart of Proposed Method



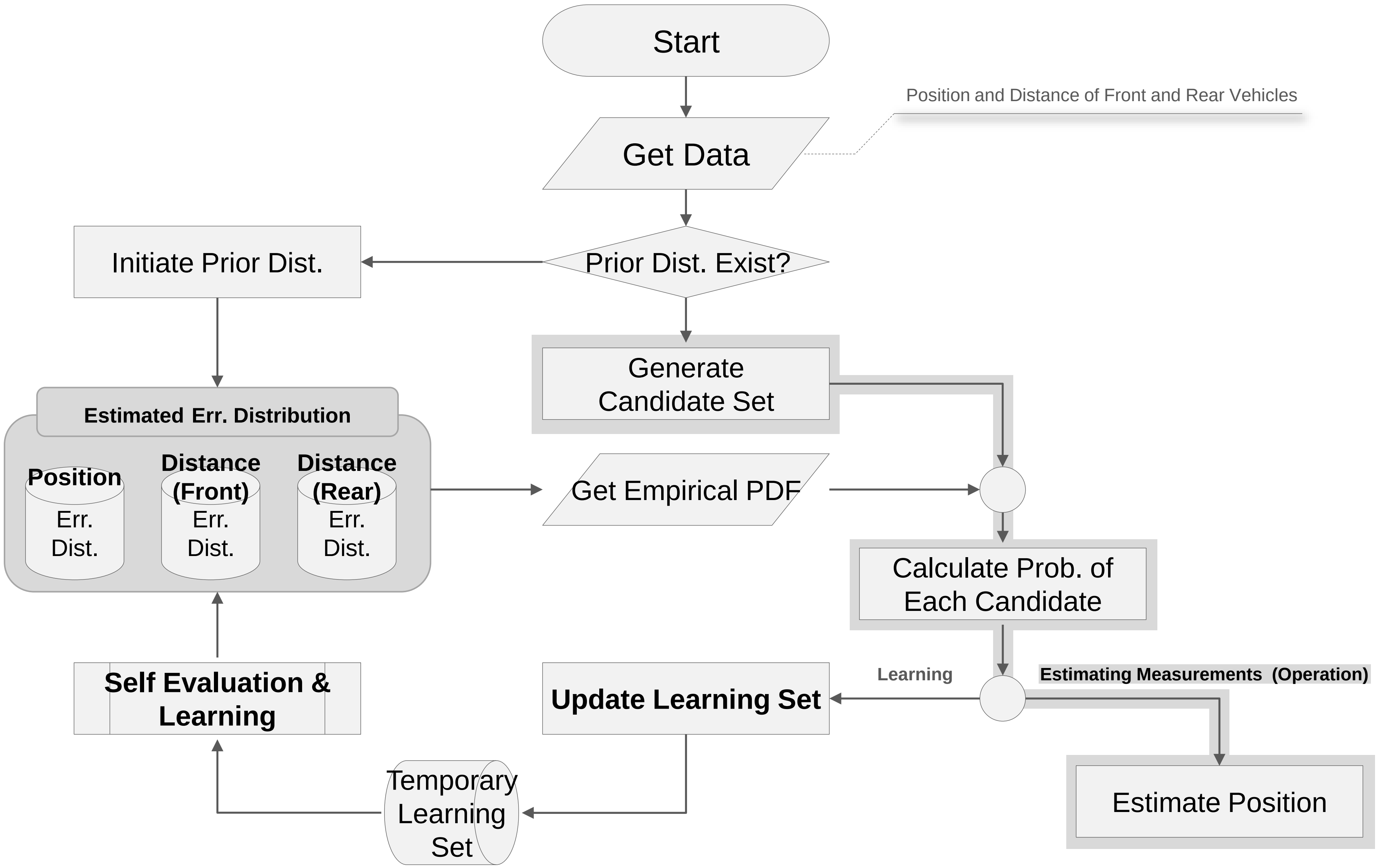
Candidate	$Err(x_a)$	$Err(x_b)$	$Err(x_c)$	Prob.
$\hat{x}_{T1}$	$x_a - \hat{x}_{T1}$	$x_b - \hat{x}_{T1}$	$x_c - \hat{x}_{T1}$	$P(\hat{x}_{T1})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{\hat{x}_{Tn}}$	$\boldsymbol{x_a - \hat{x}_{Tn}}$	$\boldsymbol{x_b - \hat{x}_{Tn}}$	$\boldsymbol{x_c - \hat{x}_{Tn}}$	$\boldsymbol{P(\hat{x}_{Tn})}$



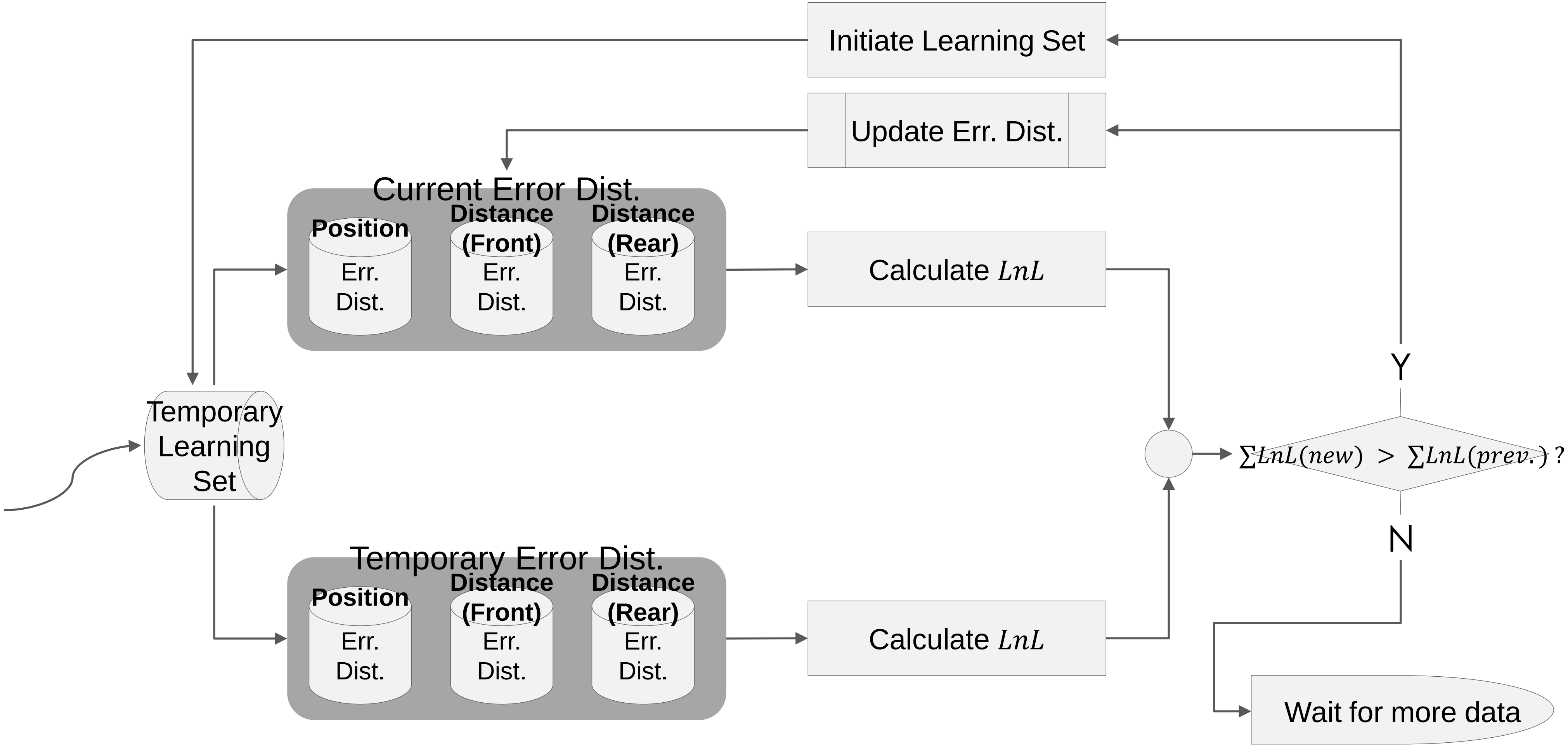
Flowchart of Proposed Method



Flowchart of Proposed Method

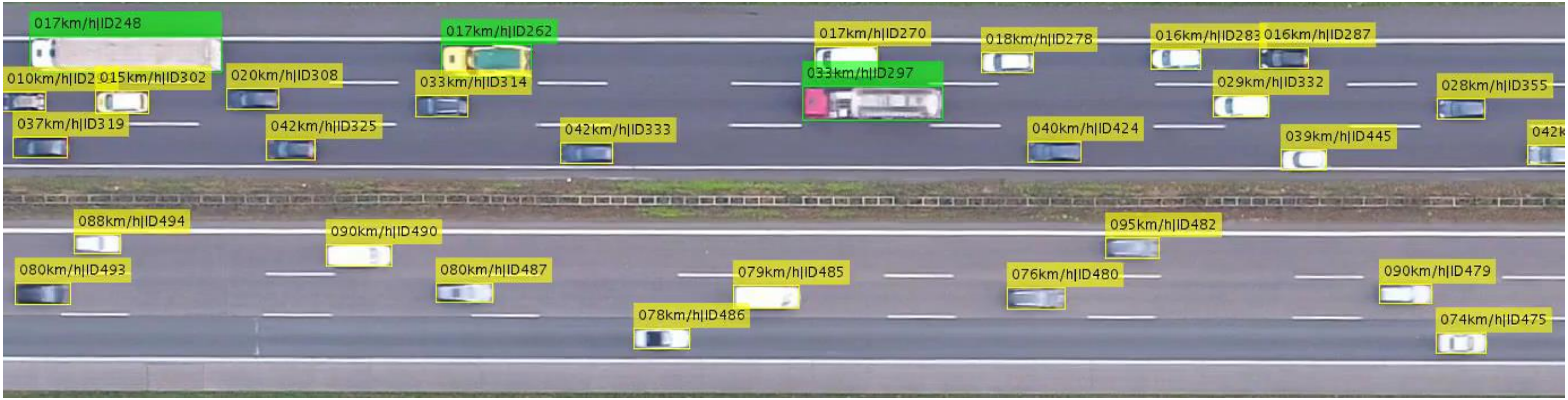


Self-Evaluation and Learning



Study Data – HighD Vehicle Trajectory

- The ‘HighD’ data is a drone-based vehicle trajectory dataset for more than 100,000 vehicles, collected from six different locations in German highways.
- To minimize the impact of specific location and time period of the data collection, the study randomly selected 30,000 frames of data as for each iteration of new training dataset.
- Still, we DO NOT have the ground truth!



Source: https://www.highd-dataset.com/

OAK RIDGE National Laboratory

Scenarios

- Measurement errors of sensors could vary depending on many factors, such as manufacturer, device, environment, signal interruptions, etc.
- To evaluate the impact of proposed method based on different circumstances, multiple scenarios were prepared based on the following three error ranges (in meters):
  - Bias: [-0.1 ~ +0.1], (Co)-Variance: [+0.1 ~ +0.5]
  - Bias: [-0.5 ~ +0.5], (Co)-Variance: [+0.1 ~ +1.0]
  - Bias: [-1.0 ~ +1.0], (Co)-Variance: [+0.1 ~ +3.0]
- The bias and the variance of true error distribution for each sensor were randomly selected from the uniform distribution within the above ranges.

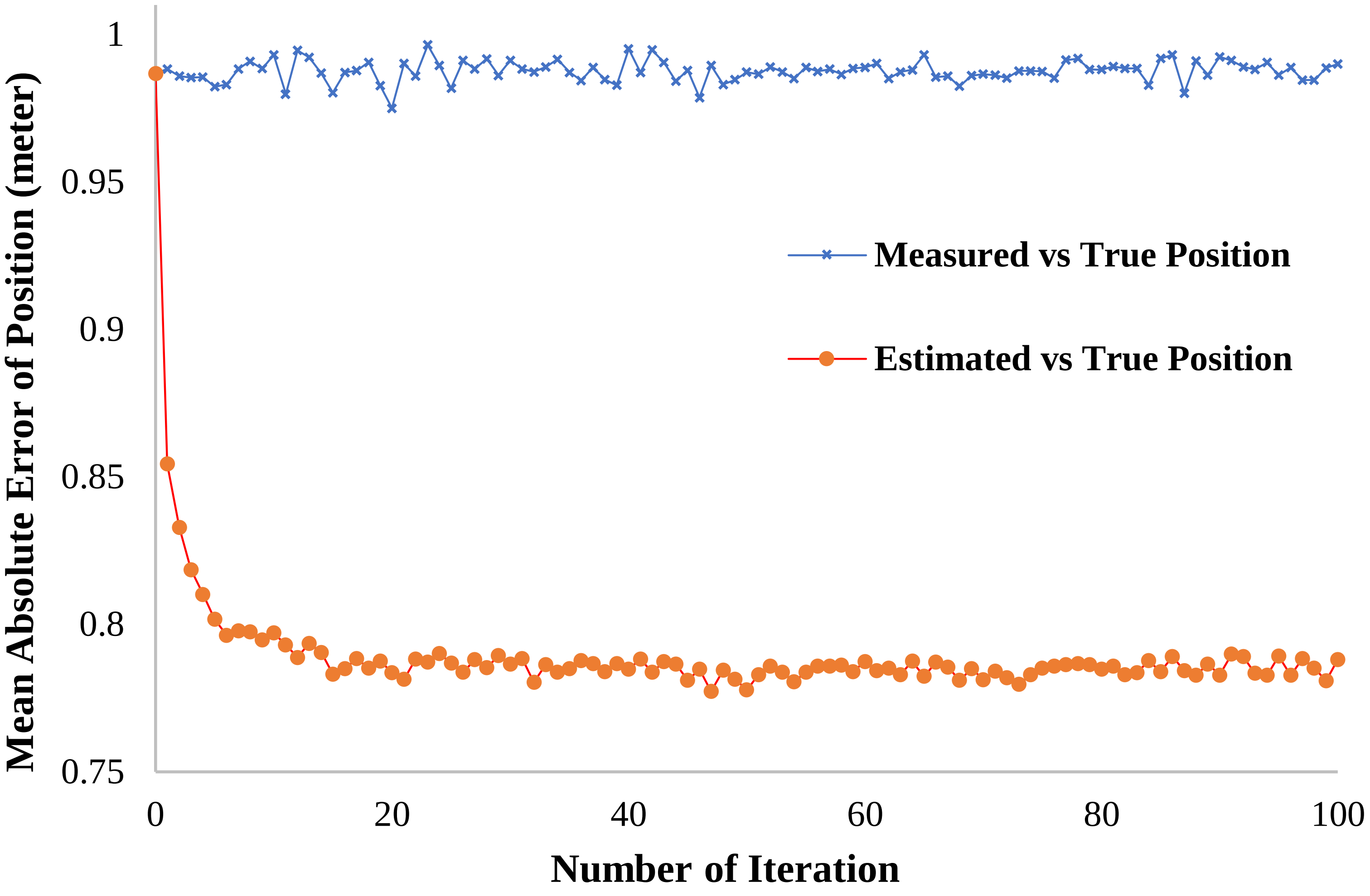
Results - Improvement of Vehicle Position Measurements

- Overall, the accuracy of position has been improved by 5 to 23%.
- Relatively, the accuracy improvement appears to be more significant when the distance measure has less bias/variance and the position measure has higher bias/variance.

Distance	**Position (X/Y coordinate)**		
	ε: [-0.1 ~ +0.1] v: [+0.1 ~ +0.5]	***ε: [-0.5 ~ +0.5] v: [+0.1 ~ +1.0]***	***ε: [-1.0 ~ +1.0] v: [+0.1 ~ +3.0]***
ε: [-0.1 ~ +0.1] v: [+0.1 ~ +0.5]	0.67 → 0.58 (- 13%)	0.98 → 0.75 (- 23%)	1.70 → 1.29 (- 24%)
ε: [-0.5 ~ +0.5] v: [+0.1 ~ +1.0]	0.68 → 0.60 (- 12%)	0.99 → 0.79 (- 20%)	1.71 → 1.34 (- 22%)
ε: [-1.0 ~ +1.0] v: [+0.1 ~ +3.0]	0.68 → 0.64 (- 5%)	0.99 → 0.87 (- 12%)	1.68 → 1.35 (- 20%)

Mean Absolute Error of Position over Number of Iterations

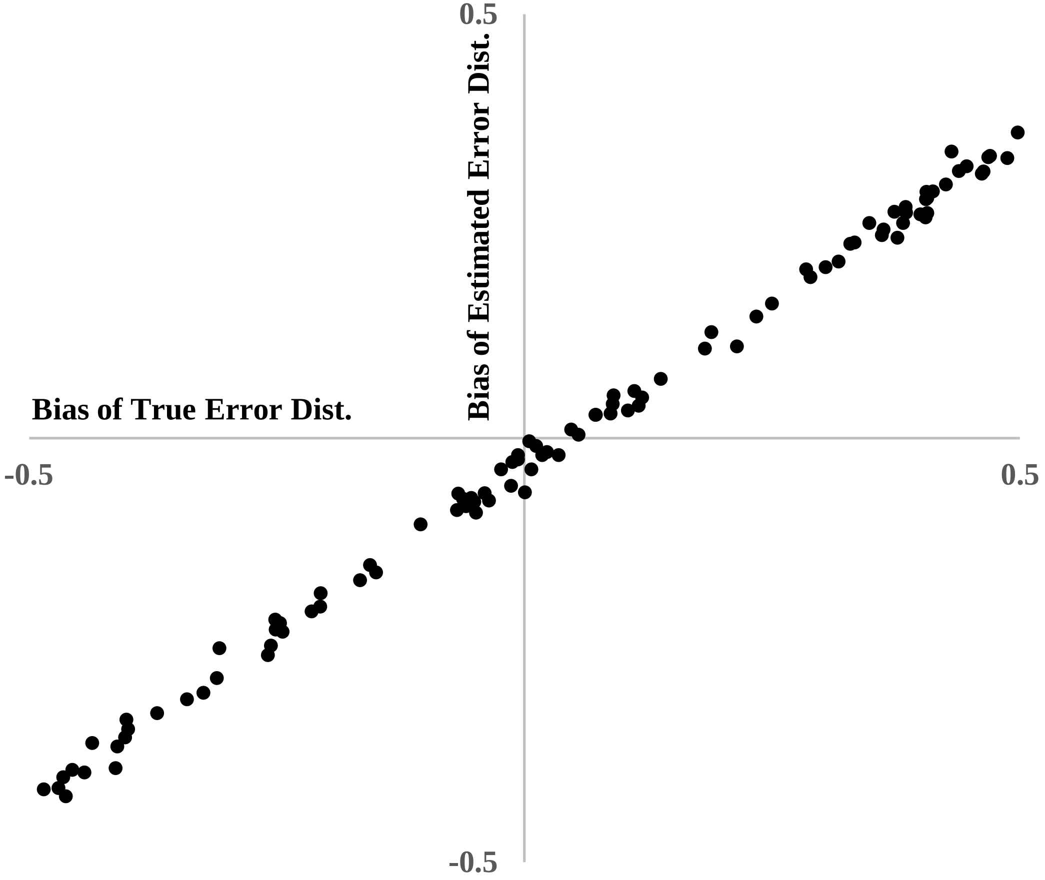
- In terms of the learning speed, the mean absolute error of position has decreased significantly from the first several iterations.



Oak Ridge National Laboratory

Estimated Bias of Position and Distance Errors

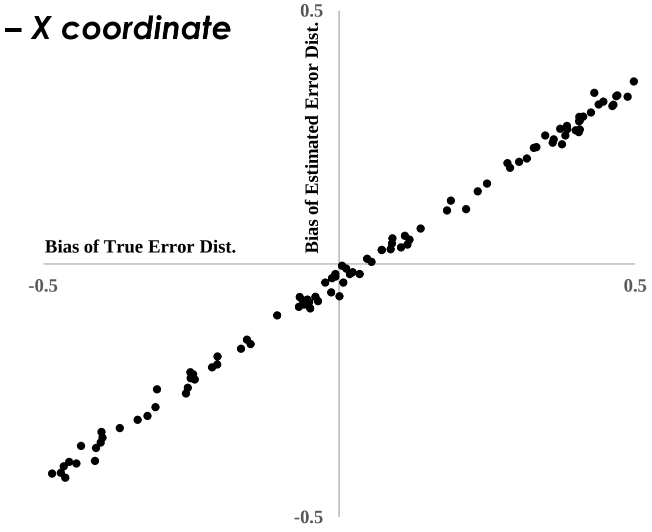
- After enough learning, the biases of x/y coordinates are very closely estimated to the true biases, with the R-squared over 0.98.



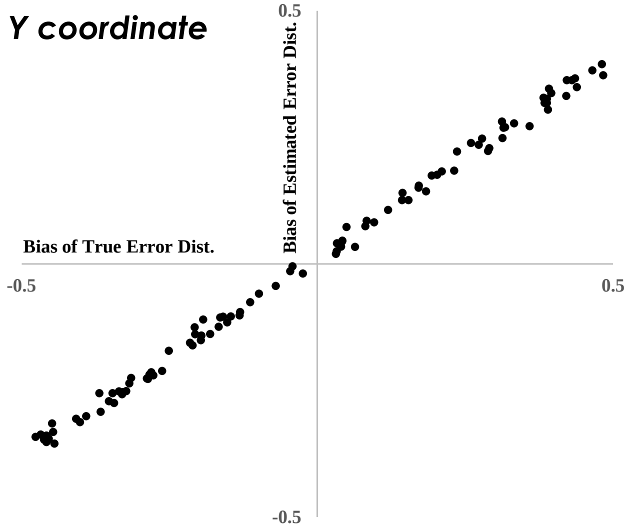
Estimated vs True Bias of Position – X coordinate

Estimated Bias of Position and Distance Errors

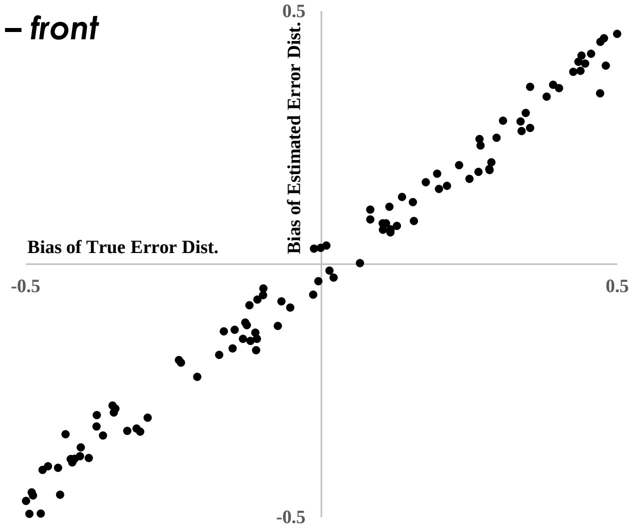
Position – X coordinate



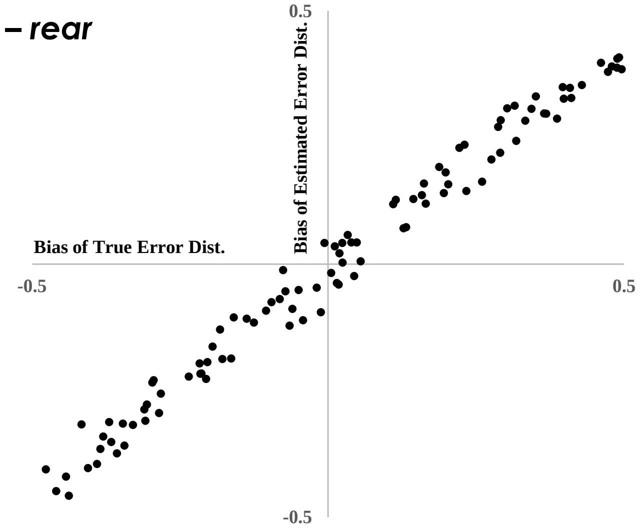
Position - Y coordinate



Distance – front



Distance – rear



OAK RIDGE National Laboratory

Conclusion

- The proposed model noticeably improved the accuracy of position and distance measurements.
- The key components to improve the measurement accuracy are:
  - combining the information from multiple data sources,
  - estimating the error distribution of each data source, and
  - utilizing the power of the multi-source data as more vehicles are connected.
- Potentially, the process can be applied to update the estimated error distribution where the error distribution of sensor measurement may be changed over time.
- Two critical limitations of the proposed method:
  - It is not able to estimate the exact true values of measurements even if the learning time goes infinite.
  - It assumes that overall error distributions of all sensors tend to be unbiased.

Thank you so much!

Questions?

Hyeonsup Lim (limh@ornl.gov)